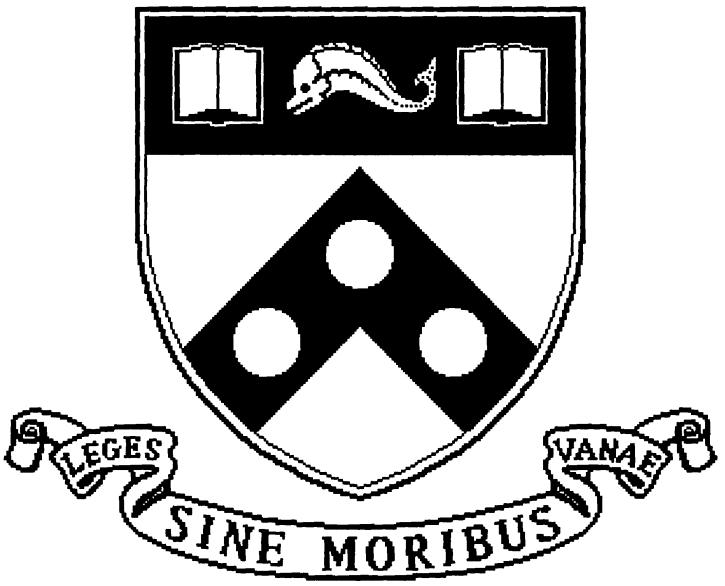
Feature-Based Localization using Fixed Ultrasonic Transducers*

MS-CIS-95-02

Robert Mandelbaum and Max Mintz [1]



University of Pennsylvania
School of Engineering and Applied Science
Computer and Information Science Department
Philadelphia, PA 19104-6389

December 14, 1994

[1]Portions of this research were supported by the following grants and contracts: ARPA Contracts N00014-92-J-1647, and DAAH04-93-G-0419; ARO Contracts DAAL03-89-C-0031PRI, and DAAL03-92-G0153; NSF Grants CISE/CDA-88-22719, IRI92-10030, IRI92-09880, IRI93-03980, and IRI93-07126.

Feature-Based Localization using Fixed Ultrasonic Transducers *

Robert Mandelbaum and Max Mintz

General Robotics and Active Sensory Perception (GRASP) Laboratory
Department of Computer and Information Science
University of Pennsylvania
Philadelphia, PA 19104

December 14, 1994

Abstract

We describe an approach for mobile robot localization based on geometric features extracted from ultrasonic data. As is well known, a single sonar measurement using a standard POLAROID™ sensor, though yielding relatively accurate information regarding the *range* of a reflective surface patch, provides scant information about the location in *azimuth* or *elevation* of that patch. This lack of sufficiently precise localization of the reflective patch hampers any attempt at data association, clustering of multiple measurements or subsequent classification and inference.

In previous work [15, 16] we proposed a multi-stage approach to clustering which aggregates sonic data accumulated from *arbitrary* transducer locations in an *sequential* fashion. It is computationally *tractable* and *efficient* despite the inherent exponential nature of clustering, and is *robust* in the face of noise in the measurements. It therefore lends itself to applications where the transducers are *fixed* relative to the mobile platform, where remaining stationary during a scan is both impractical and infeasible, and where deadreckoning errors can be substantial.

*Portions of this research were supported by the following grants and contracts: ARPA Contracts N00014-92-J-1647, and DAAH04-93-G-0419; ARO Contracts DAAL03-89-C-0031PRI, and DAAL03-92-G0153; NSF Grants CISE/CDA-88-22719, IRI92-10030, IRI92-09880, IRI93-03980, and IRI93-07126.

In the current work we apply this feature extraction algorithm to the problem of localization in a partially known environment. Feature-based localization boasts advantages in robustness and speed over several other approaches. We limit the set of extracted features to planar surfaces. We describe an approach for establishing correspondences between extracted and map features. Once such correspondences have been established, a least squares approach to mobile robot pose estimation is delineated. It is shown that once correspondence has been found, the pose estimation may be performed in time *linear* in the number of extracted features. The decoupling of the correspondence matching and estimation stages is shown to offer advantages in speed and precision.

Since the clustering algorithm aggregates sonic data accumulated from *arbitrary* transducer locations, there are no constraints on the trajectory to be followed for localization except that sufficiently large portions of features be ensonified to allow clustering. Preliminary experiments indicate the usefulness of the approach, especially for accurate estimation of orientation.

1 Introduction

1.1 Overview

In [11], a distinction is drawn between *continuous localization* and *relocation*. The difference "rests on the use of *a priori* knowledge of the vehicle position estimate in achieving correspondence, and the weighted inclusion of this *a priori* position in the updated estimate" [11]. *Continuous* localization is seen as "the normal mode of operation, with relocation used for initialization and error recovery".

In this paper both forms of localization are addressed. However, unlike [11], our approach does not employ extended Kalman filtering for the estimation phase; for this reason, no weighted inclusion of a priori position information in the updated estimate is necessary in either type of localization.

The problem of *geometric* model-based mobile robot localization (whether relocation or continuous) using the ultrasound modality can be subdivided into the following subproblems:

1. Extraction of geometric features,
2. Localization of extracted features within the local coordinate frame (calculation or updating of location and orientation parameters of extracted features within the local coordinate frame),
3. Establishment of correspondences between extracted features and model features, and
4. Estimation of robot location and orientation within the global coordinate frame.

Many approaches in the literature omit stages 1 and 2 above completely. Instead, they rely on establishing correspondences directly between actual sensor measurements and map features. The two primary difficulties with such an approach are

- For certain sensor modalities, a single measurement is insufficient to allow correspondence matching. An example is the ultrasound modality: the wide-beam nature of wave propagation results in large uncertainty in azimuth and orientation of the reflecting surface associated with a single measurement. This hampers direct correspondence matching with map features.
- Even for sensor modalities where direct correspondence matching between raw data and map features is possible, this correspondence has to be established *on every cycle*. While the process may be expedited with the aid of the previous pose estimate, the essential difficulty of correspondence matching cannot be averted.

By matching extracted *features* with map *features*, we avert these difficulties in this work. While stages 1 and 2 are necessary for the *relocation* problem, once correspondences have been established, the same correspondences may be used for many cycles, with new data readings being used to update the parameters of the extracted features. It may be argued that establishing correspondences between new measurements and extracted features is no less difficult a task. While this is true in essence, at least the correspondence problem has been decoupled from the localization problem. This decoupling allows for pose estimation techniques which are both fast and precise. Further, since localization is now performed using extracted *features*, it is much less sensitive to individual noisy or spurious measurements; each extracted feature contains the combined information from *many* measurements over *relatively long* periods. In theory, false matches become less frequent, and localization benefits from the resultant improved robustness.

We employ the algorithm described in [15, 16] for the first two subproblems above. The algorithm accepts a stream of transducer locations, orientations and corresponding measurements, and outputs a list of planar and corner features extracted from the input data. The location and orientation parameters of the extracted features are given in the local coordinate frame. For a more detailed exposition of the operation of the feature extraction algorithm, the reader is referred to [15, 16].

The current paper focuses on an approach to subproblems 3 and 4 for both continuous localization and relocation. In section 2 we describe our approach to correspondence matching, while the pose estimation problem is addressed in section 3. We show that once correspondence has been achieved, the pose estimation phase may be performed in time *linear* in the number of extracted features.

1.2 Notation

Define the *pose* $p \in \Re^2 \times [0, 2\pi]$ of a mobile platform to be a combination of its current position and orientation within a global coordinate frame (GCF). Denote the set of possible poses by $\mathcal{Q}$. If the extent of the platform's motion is confined to the region $[0, X] \times [0, Y]$, then $p \in \mathcal{Q} = [0, X] \times [0, Y] \times [0, 2\pi]$.

Let $\Re^+$ represent the non-negative real numbers. $\mathcal{P}(A)$ denotes the power set of A. $\lfloor x \rfloor$ represents the largest integer less than or equal to x. $\mathbb{Z}$ is the set of integers. $c[A] \subseteq \mathcal{B}$ denotes the image of set $A \subseteq \mathcal{A}$ under the mapping $c : \mathcal{A} \rightarrow \mathcal{B}$. Similarly, $c^{-1}[B]$ denotes the inverse image of set $B \subseteq \mathcal{B}$ under the mapping c.

$\mathbf{A}^\sharp$ denotes the pseudo-inverse of a non-square matrix $\mathbf{A}$.

1.3 Related work

The issue of localization (geometric as well as other) has been addressed extensively in the literature. Approaches may be divided into the following four broad categories according to the basic primitives used for correspondence matching between the local and global coordinate frames:

1. **Rastor-rastor** localization: Many approaches which use an *occupancy grid* representation of the environment [8, 9] attempt localization by finding the optimal match of the local bitmap of occupied space with a global occupancy grid. The goodness of a match is a function of three parameters x, y, and θ: It is a measure of the (discrete) correlation between the 2D function approximated by the occupancy grid and the function approximated by the local bitmap translated by (x, y) and rotated by θ. Drawbacks of this approach are the computational complexity of correlation as well as the trade-off between efficiency and precision, embodied in the choice of grid resolution. In [18] it is concluded that even for systems using an occupancy grid representation, more reliable position estimates are obtained by extracting segments from the grids and performing segment-segment matching. Of course, extracting features such as segments from an occupancy grid can be computationally expensive. It is also often conceptually inefficient: in constructing the occupancy grid, one discards relevant data such as surface orientation. One then attempts to re-extract this information with some operator on the grid.

2. **Rastor-feature** localization: This category comprises approaches which represent the environment parametrically in features; Localization is performed by establishing correspondences between individual unclustered sensor measurements and the map features. Once correspondence has been established, some form of optimization is performed to minimize some function of the vector of spatial discrepancies between measurements and features. Approaches differ in how the correspondences are established, the definition of spatial discrepancy and the function to be minimized.

   In [5], the entire environment space of the robot is searched for locations which would yield sensor readings consistent with the measured range readings. It is recognized that due to noise, there may be no location at which *all* readings are consistent; the location which yields the greatest number of consistent readings is selected. While mention is made of the need for a metric over pose space $\mathcal{Q}$, no such metric is described. Despite numerous tools borrowed from computational geometry, the algorithm is computationally prohibitively expensive: it is stated to be $O(m^2n^2\log(mn))$ where m is the number of range readings taken and n is the complexity of the map (compare with $O(n)$ for pose estimation using the approach described in section 3. The rastorized version runs in time $O(mr^2e^2l)$ where the rastorized environment is r by r, e is the radius of an "error ball", and l is the size of the greatest range reading. Hence, this algorithm is not practical for *continuous*

localization. Moreover, the approach does not address orientation uncertainty at all. Inclusion of this dimension would make it even more computationally infeasible, so that the utility of the approach is questionable even for the less constrained problem of *relocation*. Furthermore, the algorithm is based more on intuitive arguments than mathematical rigor, nor does it deal with uncertainty and sensor error in a satisfactory manner.

In [13], the pose estimation problem is formulated as an iterative optimization in terms of the extent to which the map explains the observed measurements. The approach assumes a coarse initial position estimate is available, and estimates the correct position *assuming the orientation is known*. Each data point, representing the position of a reflecting point in the environment, is classified to a target line segment. A *correction vector* is then associated with each data point, and *weighted voting* of the correction vectors yields an overall translation vector to correct position. Various estimators are introduced to indicate whether a calculated pose is probable. For pose calculations deemed reliable, orientation is then corrected by maximizing one of the estimators with respect to rotation. The new orientation estimate is used as an initial orientation estimate for the next iteration.

Though the general approach presented in [13] seems valid, it suffers from a few drawbacks. A pencil-beam model for ultrasound propagation is assumed and forms an integral part of the entire approach; this is an inadequate model, glossing over the issues introduced by the wide-beam nature of the propagation. Further, many of the estimators used are heuristic and ad hoc. Though "correct" behavior is shown for a single example, no general behavior is proved, either theoretically or empirically. The orientation estimation procedure is particularly deficient in this regard: Global maximization is performed on a poorly characterized function.

Methods employing Kalman filters to model both robot pose and map features abound [1, 4, 3, 6, 11, 12]. These approaches may be categorized as rastor-feature localization methods, since each unclustered measurement is used individually to update the various Kalman filters. Though features are extracted from the data, these features are not used for correspondence.

A consistent shortcoming of systems employing Kalman filters is that often the system in question is not shown to meet the underlying assumptions of the Kalman filter. One of these assumptions is the Gaussian nature of both process and measurement noise: many researchers *assume* this noise model with no empirical or theoretical justification at all.

3. **Feature-feature** localization: This category consists of approaches in which measurements are clustered into features. Correspondence is established between extracted features and map features. A new estimate for location is obtained by minimizing a function on the vector of feature-feature discrepancies. These discrepancies may be based not only on *spatial* distances between extracted and map features, but also on *any* of the features' other parameters such as size, shape or orientation.

New data is incorporated into already existing clusters, and is used to update the clusters' parameters, prior to another phase of feature-feature localization. In this way, the correspondence and localization problems are decoupled.

The main advantages of feature-feature localization are twofold: Firstly, it is robust in the face of noise since each feature's parameters are derived from multiple measurements. This reduces the sensitivity of the localization to noisy or spurious measurements. Secondly, the same correspondence match may be used over many cycles, reducing overhead and speeding up the localization procedure. The major disadvantage of feature-feature localization is the longer start-up period during which features are extracted. During this period, some other form of localization such as odometry must be employed.

The approach described in this paper is a member of the feature-feature category.

In [10], regions of common depth (RCD's as defined in [11]) are extracted from ultrasonic data from a rotating scanning transducer. Each RCD is classified into one of four types according to neighboring RCD's and the width of the RCD. Pairs of RCD's are matched with pairs in the map; each match of a pair has an associated transformation under which the local coordinate frame is brought into correspondence with the global frame. The set of transformations is analyzed for clusters, and the centroid of the largest cluster selected as the "optimal" transformation. Though at first glance this approach appears to belong to the feature-feature category of localization algorithms, in essence it belongs more to the rastor-feature category: though features (RCD's) are extracted from the data before correspondence is attempted, new features are extracted during each cycle. New data is not analyzed for correspondence with previously found clusters. One of the main advantages of feature-feature localization is not exploited in that previous correspondence matches are not re-used. In effect, rastor-feature localization is being performed, though the "rastor" data is slightly more refined than individual sonar measurements.

The are several other difficulties with the approach described in [10]. First of all, the search for clusters is performed on a 4-dimensional space. Each coordinate is assumed to be independent of the remaining three coordinates; in fact, two of the coordinates are non-linearly dependent, and related to each other by a trigonometric relationship. Points within the space are assumed to be uniformly distributed, though no justification for the assumption is presented, either theoretic, intuitive or empirical. Similarly, no justification is given for the assumptions that the largest cluster in the space of transformations is the "correct" one, or that the centroid of this cluster is a reasonable choice for the localization transformation. Finally, the approach is suitable only for systems using scanning rotating sonars.

4. **Landmark-landmark** localization. Many systems base the localization procedure on *landmarks* detected in the environment. The distinction drawn here between a feature and a landmark lies in the amount of knowledge

about identity: A *feature* is a summary description of a cluster of data points; its identity within a large class of similar features is not known. A landmark, on the other hand, is a feature with a unique identity based on some distinguishing characteristic such as spatial location or some sensed property.

Landmark-based localization, then, differs from feature- or rastor-based localization in that *no correspondence matching is necessary.* This phase is rendered superfluous by a more comprehensive *recognition* phase.

Various approaches to landmark localization are explored in [2, 19, 20, 21] among others.

Localization based on the detection of beacons — easily recognizeable features placed in the environment — falls into the landmark-landmark or feature-feature category depending on whether detection of a beacon uniquely determines its identity or whether correspondence must still be established between extracted and map beacons.

2 Establishing correspondences

2.1 Problem statement

In this section we describe an approach to establishing correspondences between extracted features and model features. In the present work, extracted features comprise planar surfaces only.

In 2D, a planar feature is a line-segment. We use a redundant *normal parametrization* representation ([7], page 336) for line-segments: the line of which the segment forms a part is represented by the pair (ρ, ϕ) where ρ is the perpendicular distance of the line from some fixed $(0,0)$ point; ϕ is the inclination of the normal to the line relative to the x-axis in a counter-clockwise direction. This pair is supplemented with parameters to specify the locations of the end-points. Only two extra parameters are necessary, but for simplicity we use three: two for the location of the center-point of the line-segment (c_x, c_y) and one for the length of the segment l. Hence, each line-segment is represented by the 5-tuple $(\rho, \phi, c_x, c_y, l) \in \Gamma \times \Phi \times C_x \times C_y \times L$.

Let $\mathcal{M} = \{m_i = (\rho_i, \phi_i, c_{x_i}, c_{y_i}, l_i) \mid 1 \leq i \leq |\mathcal{M}|\}$ be the set of planar features in the given map. Let $\mathcal{F} \subseteq \mathcal{M}$ be the set of planar features within the mobile robot's current ultrasonic "field of view". Let $\mathcal{E}$ be the set of features extracted by the clustering algorithm described in [16]. Let $\mathcal{C}$ be the class of functions $\mathcal{C} = \{c : \mathcal{E} \rightarrow \mathcal{F} \cup \bar{0}\}$. Each element $c \in \mathcal{C}$ represents an assignment of correspondences: for a specific $e \in \mathcal{E}$, $c(e)$ is the map feature to which e corresponds under this assignment c. Note that we do not insist that elements of $\mathcal{C}$ are either injective or surjective: multiple elements of $\mathcal{E}$ may correspond to the same feature in $\mathcal{F}$ since the clustering algorithm may extract multiple

sections of the same underlying planar surface; further, not all elements of $\mathcal{F}$ will necessarily be detected. Note also the augmentation of $\mathcal{F}$ with the zero element $\bar{0}$ to form the range of elements of $\mathcal{C}$. The zero element $\bar{0}$ may be made the image of spurious extracted features. In this way, phantom extracted features need not to be mapped to elements of $\mathcal{F}$, though elements of $\mathcal{C}$ are still well-defined.

In essence, these properties of elements of $\mathcal{C}$ highlight some of the difficulties with which we are faced in the correspondence problem: Some features may be detected multiple times, while others are not detected at all, and the set of extractions may include phantom features which should not be associated with any real feature. The set $\mathcal{E}$ must be mapped to $\mathcal{F}$ despite these structural differences in the sets.

Let $q : \mathcal{C} \rightarrow \Re^+$ be a quality measure of assignments. The correspondence problem may be then be formulated as a search over $\mathcal{C}$ for an element $\hat{c} \in \mathcal{C}$ which maximizes q. A combinatorial analysis shows that $|\mathcal{C}| = (|\mathcal{F}|+1)^{|\mathcal{E}|}$. Hence the correspondence problem is inherently exponential. We describe here an approach to find $\hat{c}$ in polynomial time.

2.2 The 1-dimensional case

Consider the special case where no two line-segments in $\mathcal{F}$ are parallel. We deal with more general cases in subsections 2.3 and 2.4. As a working example, let $\mathcal{F}$ be the set of line-segments making up an arbitrary triangle, as shown in figure 1. Let the normals to the three segments have orientations α_1, α_2 and α_3 relative to the global coordinate frame (GCF). Then the histogram of Φ-values of $\mathcal{F}$ has value 1 at α_1, α_2 and α_3, and is 0 elsewhere on the interval $[0, 2\pi)$. Denote this histogram $\mathcal{H}_{\mathcal{F}}^{\Phi}$. Similarly, let $\mathcal{H}_{\mathcal{E}}^{\Phi}$ denote the (discrete) histogram of Φ-values of $\mathcal{E}$. However, instead of each element of $\mathcal{E}$ contributing to the histogram equally, let the contribution of $e \in \mathcal{E}$ be the *number of ultrasound readings* which were clustered to form e (see [16] for an exposition of the clustering algorithm). The width of the "bins" in $\mathcal{H}_{\mathcal{E}}^{\Phi}$, w, represents the "slop" in orientation of extracted features which is to be tolerated. Examples of $\mathcal{E}$ and $\mathcal{H}_{\mathcal{E}}^{\Phi}$ are shown in figure 1. Note that the intervals between local modes of $\mathcal{H}_{\mathcal{E}}^{\Phi}$ are invariant under translations and rotations of the local coordinate frame (LCF) relative to the GCF. It is this invariance we exploit in order to find a good correspondence assignment $c \in \mathcal{C}$.

We now attempt to find the best correlation of $\mathcal{H}_{\mathcal{F}}^{\Phi}$ with a shifted version of $\mathcal{H}_{\mathcal{E}}^{\Phi}$. Let $\mathcal{H}_{\mathcal{E}}^{\Phi}(\beta)$ denote $\mathcal{H}_{\mathcal{E}}^{\Phi}$ shifted in the positive Φ-direction by β, with wrap-around at 2π. Assuming that the pairwise separations of α_1, α_2 and α_3 are all greater than the bin width w, the product $\mathcal{H}_{\mathcal{F}}^{\Phi} \cdot \mathcal{H}_{\mathcal{E}}^{\Phi}(\beta)$ will consist of at most three non-zero bins for each value of β. Let $\bar{p}(\beta)$ represent the three-vector of bin values of $\mathcal{H}_{\mathcal{E}}^{\Phi}(\beta)$ "picked out" by $\mathcal{H}_{\mathcal{F}}^{\Phi}$ for each $\beta = iw$, $0 \leq i \leq \lfloor \frac{2\pi}{w} \rfloor$. Examples of $\bar{p}(\beta)$ are shown in figure 1.

The final step consists of evaluating all the vectors $\bar{p}(\beta)$, $\beta = iw$, $0 \leq i \leq \lfloor \frac{2\pi}{w} \rfloor$, and selecting the "optimal" shift β. In other words, if $h : \mathbb{Z}^{|\mathcal{F}|} \rightarrow \Re$ represents an evaluation function of vectors of length $|\mathcal{F}|$ of integers, we search for a value of β, call it $\hat{\beta}$, which maximizes $h(\bar{p}(\beta))$. The choice of h depends on which properties of a correspondence assignment we choose to emphasize. In our case, we chose to stress two properties:

1. We wish to account for the greatest possible number of ultrasound readings. In other words, we wish to reward a high *mean* value of components of $\bar{p}(\beta)$.
2. We prefer the correspondence matches to be spread *evenly* over the elements of $\mathcal{F}$. In other words, we wish to reward low *standard deviation* among components of $\bar{p}(\beta)$.

Thus, the function h we selected for our experiments has the form $h(\bar{p}) = \lambda_1 \cdot \boldsymbol{m}(\bar{p}) - \lambda_2 \cdot \boldsymbol{s}(\bar{p})$ where $\boldsymbol{m} : \mathbb{Z}^{|\mathcal{F}|} \rightarrow \Re$ and $\boldsymbol{s} : \mathbb{Z}^{|\mathcal{F}|} \rightarrow \Re$ are, respectively, functions giving the mean and standard deviation of the components of a vector of length $|\mathcal{F}|$, and λ_1 and λ_2 are non-negative weighting factors.

Once $\hat{\beta}$ has been found, the correspondence matching is complete: use that $c \in \mathcal{C}$ which makes the following assignments:

- Tor each element $f \in \mathcal{F}$ with orientation ϕ_f in the GCF, map all elements in $\mathcal{E}$ which contributed to the bin in $\mathcal{H}_{\mathcal{E}}^{\Phi}(\hat{\beta})$ picked out by f. In other words, let the inverse image of f under c be

$$c^{-1}[\{f\}] = \{e \in \mathcal{E} \mid iw \leq \phi_e \leq (i+1)w,\ i = \lfloor \frac{\phi_f + \hat{\beta}}{w} \rfloor\}$$

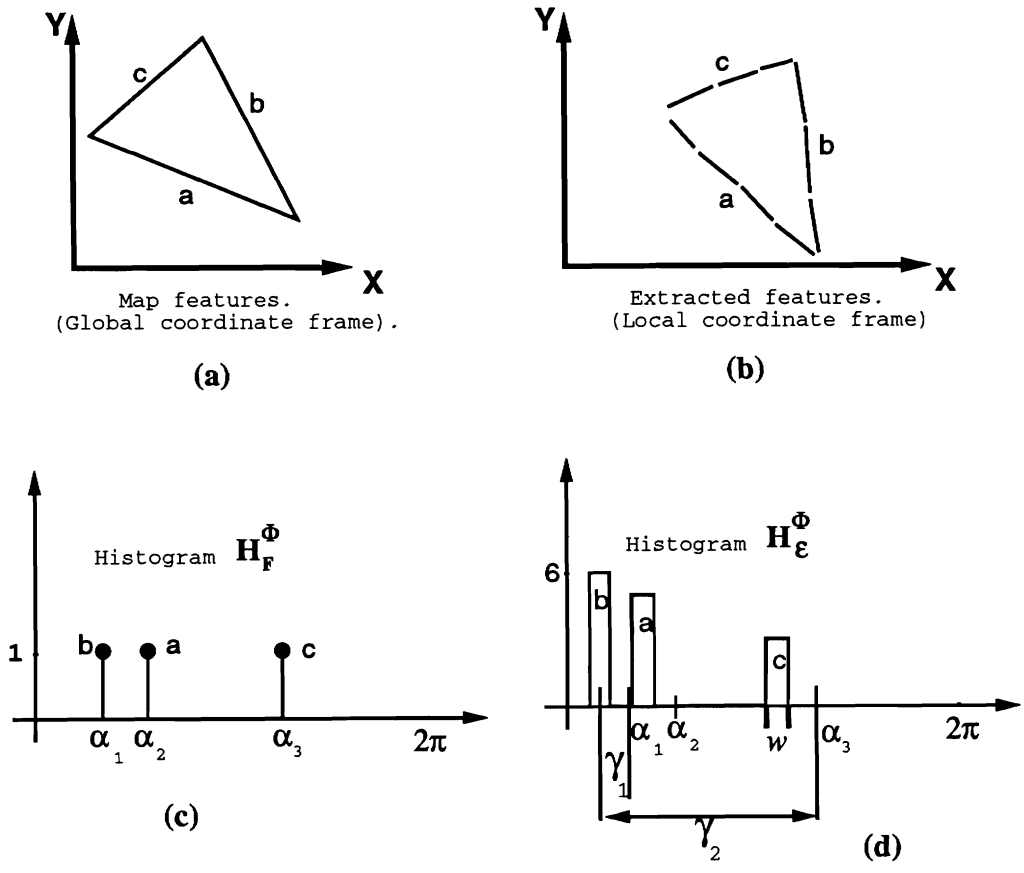
  where ϕ_e is the orientation of element e in the LCF.

- Map all remaining elements of $\mathcal{E}$ to the zero vector in $\mathcal{F} \cup \bar{0}$. i.e.

$$c\left[\mathcal{E} - c^{-1}[\mathcal{F}]\right] = \bar{0}$$

2.3 The 2-dimensional case

A difficulty arises when either $\mathcal{F}$ or $\mathcal{E}$ contains parallel line-segments (or at least line-segments whose angular separation is smaller than w so that they fall into the same bin in $\mathcal{H}_{\mathcal{F}}^{\Phi}$ or $\mathcal{H}_{\mathcal{E}}^{\Phi}$) but whose perpendicular spatial separation is substantial. In this case, projecting $\mathcal{F}$ and $\mathcal{E}$ into the Φ axis in the construction of $\mathcal{H}_{\mathcal{F}}^{\Phi}$ and $\mathcal{H}_{\mathcal{E}}^{\Phi}$, erases the distinction between these line-segments. This, in turn, leads to ambiguous correspondence matching if $\mathcal{F}$ contains the parallel segments, or at least one mismatched "phantom" segment if it is $\mathcal{E}$ that contains the parallel segments.



(e)
$$\bar{p}(0) = (0,0,0)$$
$$\bar{p}(\gamma_1) = (6,5,4)$$
$$\bar{p}(\gamma_2) = (0,0,6)$$

Figure 1: The 1-dimensional histogram case: (a) Map features in the GCF, (b) Extracted features in the LCF, (c) The histogram $\mathcal{H}_{\mathcal{F}}^{\Phi}$ with three non-zero values, one for each feature in $\mathcal{F}$. The labels represent correspondences between map features and histogram points. (d) The histogram $\mathcal{H}_{\varepsilon}^{\Phi}$. The labels represent correspondences between map feature clusters and histogram points. (e) Sample values of $\bar{p}(\beta)$.

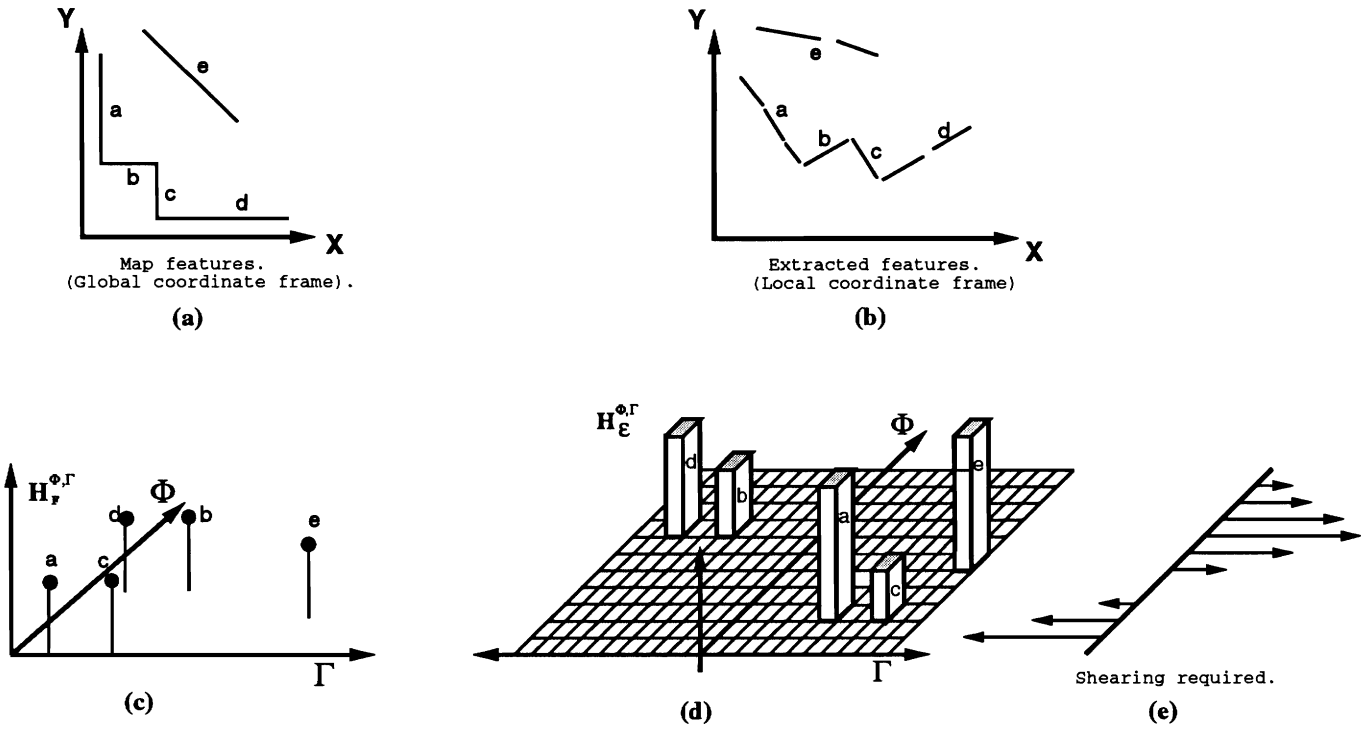


Figure 2: The 2-dimensional histogram case: (a) Map features in the GCF, (b) Extracted features in the LCF, (c) The histogram $\mathcal{H}_{\mathcal{F}}^{\Phi,\Gamma}$ with five non-zero values, one for each feature in $\mathcal{F}$. The labels represent correspondences between map features and histogram points. (d) The histogram $\mathcal{H}_{\mathcal{E}}^{\Phi,\Gamma}$. The labels represent correspondences between map feature clusters and histogram points. (e) Example shearing necessary to bring $\mathcal{H}_{\mathcal{E}}^{\Phi,\Gamma}$ into good correlation with $\mathcal{H}_{\mathcal{F}}^{\Phi,\Gamma}$ after a suitable (rigid) shift in the Φ direction.

In order to distinguish between parallel line-segments, we extend the histogram matching approach to the 2-dimensional case: Let $\mathcal{H}_{\mathcal{F}}^{\Phi,\Gamma}$ be the 2-dimensional histogram of $\mathcal{F}$ with respect to Φ and Γ. A point f in $\mathcal{F}$ contributes to the bin in $\mathcal{H}_{\mathcal{F}}^{\Phi,\Gamma}$ at location $\left(\lfloor\frac{\phi_f}{w_\Phi}\rfloor, \lfloor\frac{\rho_f}{w_\Gamma}\rfloor\right)$, where ϕ_f and ρ_f are the Φ- and Γ-values of f, and w_Φ and w_Γ are the resolutions of $\mathcal{H}_{\mathcal{F}}^{\Phi,\Gamma}$ in the Φ and Γ directions respectively. Similarly, let $\mathcal{H}_{\mathcal{E}}^{\Phi,\Gamma}$ denote the 2-dimensional histogram of $\mathcal{E}$. See figure 2 for a graphical depiction of the current running example.

As in the 1-dimensional case, the task is now to find shifts in $\mathcal{H}_{\mathcal{E}}^{\Phi,\Gamma}$ which will bring it into greatest correlation with $\mathcal{H}_{\mathcal{F}}^{\Phi,\Gamma}$. Unfortunately, however, in the 2-dimensional case, the relative locations of points in $\mathcal{H}_{\mathcal{E}}^{\Phi,\Gamma}$ are *not* invariant under translations and rotations of the LCF relative to the GCF. In fact, a straightforward trigonometric argument (see figure 3) shows that a rotation of the LCF by β followed by a spatial translation of (x, y) transforms a line with

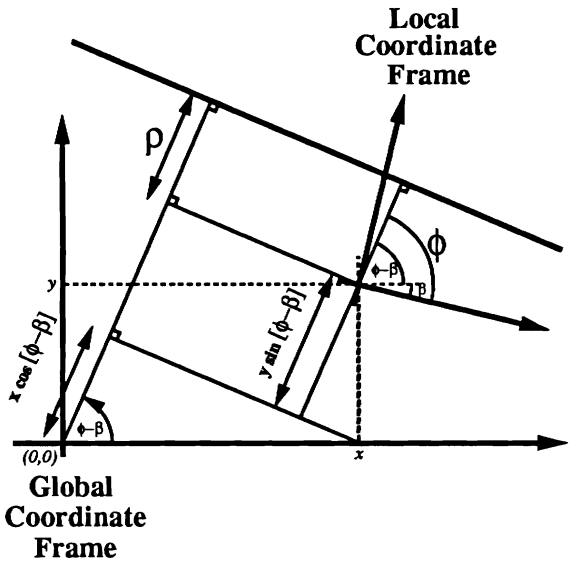


Figure 3: Rotation of the Local Coordinate Frame by β followed by spatial translation of (x, y) transforms a line with normalized parameters (ρ, ϕ) into the line with parameters $(\rho + x\cos(\phi - \beta) + y\sin(\phi - \beta), \phi - \beta)$.

normalized parameters (ρ, ϕ) into the line (ρ', ϕ'), where

$$\begin{pmatrix} \rho' \\ \phi' \end{pmatrix} = \begin{pmatrix} \rho + x\cos(\phi - \beta) + y\sin(\phi - \beta) \\ \phi - \beta \end{pmatrix} \tag{1}$$

Hence, relative differences in Φ are preserved. Similarly, for any fixed value of $\phi \in \Phi$, relative differences in Γ are preserved, since all lines with orientation ϕ are shifted in Γ by a constant factor $x\cos(\phi - \beta) + y\sin(\phi - \beta)$. However, lines with different Φ values are shifted by different amounts in the Γ direction. In other words, a **rotation** of the LCF corresponds to a rigid shift of $\mathcal{H}_{\mathcal{E}}^{\Phi,\Gamma}$ in the Φ direction; a *translation* of the LCF with respect to the GCF, however, corresponds to a *non-linear shearing* of $\mathcal{H}_{\mathcal{E}}^{\Phi,\Gamma}$. See figure 2 for a graphical depiction.

Let $\mathcal{H}_{\mathcal{E}}^{\Phi,\Gamma}(\beta, x, y)$ denote the 2-dimensional histogram of $\mathcal{E}$ after a rotation of β and a translation of (x, y). We describe here an approach to find the triple $(\hat{\beta}, \hat{x}, \hat{y})$ for which $\mathcal{H}_{\mathcal{E}}^{\Phi,\Gamma}(\hat{\beta}, \hat{x}, \hat{y})$ is in best correlation which $\mathcal{H}_{\mathcal{F}}^{\Phi,\Gamma}$. Once this has been done, we deduce the correspondence assignment $c \in \mathcal{C}$ in much the same way as in the 1-dimensional case.

Our approach to finding the triple $(\hat{\beta}, \hat{x}, \hat{y})$ is to decouple the searches for the rotation $\hat{\beta}$, and the translation $(\hat{x}, \hat{y})$.

The search for rotation $\hat{\beta}$: After a pure rotation, $\mathcal{H}_{\mathcal{E}}^{\Phi,\Gamma}(\beta, 0, 0)$ is related to $\mathcal{H}_{\mathcal{E}}^{\Phi,\Gamma}(0, 0, 0)$ by a rigid shift of β in the Φ direction. A heuristic approach to finding $\hat{\beta}$ consists of projecting $\mathcal{H}_{\mathcal{F}}^{\Phi,\Gamma}$ onto $\mathcal{H}_{\mathcal{F}}^{\Phi}$ and $\mathcal{H}_{\mathcal{E}}^{\Phi,\Gamma}$ onto $\mathcal{H}_{\mathcal{E}}^{\Phi}$ and then finding the best correlation of $\mathcal{H}_{\mathcal{F}}^{\Phi}$ with a shifted version of $\mathcal{H}_{\mathcal{E}}^{\Phi}$ as in the 1-dimensional case. We note that it is possible for the highest correlation estimate of $\mathcal{H}_{\mathcal{F}}^{\Phi}$ with $\mathcal{H}_{\mathcal{E}}^{\Phi}$ to result in a *non*-optimal correlation of $\mathcal{H}_{\mathcal{F}}^{\Phi,\Gamma}$ with $\mathcal{H}_{\mathcal{E}}^{\Phi,\Gamma}$. Cases for which this occurs usually involve multiple elements in $\mathcal{F}$ being regularly spaced in the Φ and Γ directions; we address this issue in section 2.4. For many cases, however, the computational efficiency afforded by the heuristic outweighs the sacrifice of a guarantee of optimality.

The search for translation $(\hat{x}, \hat{y})$: Once a suitable rotation of the LCF relative to the GCF has been found, it is possible to use a spatial representation of elements of $\mathcal{F}$ and $\mathcal{E}$ to find the translation necessary for a good correspondence match. However, the rotation of all elements of $\mathcal{E}$ may be computationally prohibitive. We describe here an alternative approach which finds the parameters of shearing of $\mathcal{H}_{\mathcal{E}}^{\Phi,\Gamma}(\hat{\beta}, 0, 0)$ necessary to bring it into best correspondence with $\mathcal{H}_{\mathcal{F}}^{\Phi,\Gamma}$. These parameters correspond to the translation $(\hat{x}, \hat{y})$.

The approach consists of finding the amount of shearing of $\mathcal{H}_{\mathcal{E}}^{\Phi,\Gamma}(\hat{\beta}, 0, 0)$ for each value of γ, $\gamma = iw_{\Phi}$, $0 \leq i \leq \lfloor \frac{2\pi}{w_{\Phi}} \rfloor$. Let $\mathcal{H}_{\mathcal{F}}^{\Phi,\Gamma}|_{\phi=\gamma}$ denote the 1-dimensional histogram "slice" of $\mathcal{H}_{\mathcal{F}}^{\Phi,\Gamma}$ at $\phi = \gamma$. Similarly, let $\mathcal{H}_{\mathcal{E}}^{\Phi,\Gamma}(\hat{\beta}, 0, 0)|_{\phi=\gamma}$ denote the corresponding 1-dimensional histogram in $\mathcal{H}_{\mathcal{E}}^{\Phi,\Gamma}(\hat{\beta}, 0, 0)$. For each $\gamma = iw_{\Phi}$, $0 \leq i \leq \lfloor \frac{2\pi}{w_{\Phi}} \rfloor$ for which $\mathcal{H}_{\mathcal{F}}^{\Phi,\Gamma}|_{\phi=\gamma}$ is non-zero, we find the best correlation between the 1-dimensional histograms $\mathcal{H}_{\mathcal{F}}^{\Phi,\Gamma}|_{\phi=\gamma}$ and $\mathcal{H}_{\mathcal{E}}^{\Phi,\Gamma}(\hat{\beta}, 0, 0)|_{\phi=\gamma}$ using the method described in section 2.2. Denote by d_i the Γ-shift in $\mathcal{H}_{\mathcal{E}}^{\Phi,\Gamma}(\hat{\beta}, 0, 0)|_{\phi=\gamma}$ necessary to achieve greatest correlation with $\mathcal{H}_{\mathcal{F}}^{\Phi,\Gamma}|_{\phi=\gamma}$.

We note from equation 1 that

$$\begin{aligned} d_i &= (\rho_{\text{GCF}} - \rho_{\text{LCF}}) \text{ at } \phi = iw_{\Phi} \\ &= \hat{x}\cos(iw_{\Phi} - \hat{\beta}) + \hat{y}\sin(iw_{\Phi} - \hat{\beta}) \\ &= a_i\hat{x} + b_i\hat{y} \end{aligned}$$

where $a_i = \cos(iw_{\Phi} - \hat{\beta})$ and $b_i = \sin(iw_{\Phi} - \hat{\beta})$. We may therefore estimate $\hat{x}$ and $\hat{y}$ by a least squares fit:

$$\begin{aligned} \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} &= \mathbf{A}^{\mathbf{I}}\bar{d} \\ &= (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\bar{d} \end{aligned}$$

where $\mathbf{A} = [\bar{a}, \bar{b}]$, $\bar{a} = [a_{i_1}, a_{i_2}, \ldots]^T$, $\bar{b} = [b_{i_1}, b_{i_2}, \ldots]^T$, $\bar{d} = [d_{i_1}, d_{i_2}, \ldots]^T$, and $i_1, i_2, \ldots$ are the values of i for which $\mathcal{H}_{\mathcal{F}}^{\Phi,\Gamma}|_{\phi=iw_{\Phi}}$ is non-zero. As a caveat, we note that a least squares approach is sensitive to outliers. For this

reason, least squares estimation may be replaced here with truncated least squares, least median of squares, or some other more robust estimator.

Once the triple $(\hat{\beta}, \hat{x}, \hat{y})$ has been found, the correspondence matching is complete: use that $c \in \mathcal{C}$ which makes the following assignments:

- For each element $f \in \mathcal{F}$ with Φ and Γ values ϕ_f and ρ_f in the GCF, map all elements in $\mathcal{E}$ which contribute to the bin in $\mathcal{H}_{\mathcal{E}}^{\Phi,\Gamma}(\hat{\beta}, \hat{x}, \hat{y})$ picked out by f. In other words, if f picks out bin (i, j) in $\mathcal{H}_{\mathcal{E}}^{\Phi,\Gamma}(\hat{\beta}, \hat{x}, \hat{y})$ (i.e. $i = \lfloor \frac{\phi_f + \hat{\beta}}{w_\Phi} \rfloor$ and $j = \lfloor \frac{\rho_f - \hat{x}\cos(\phi_f + \hat{\beta}) - \hat{y}\sin(\phi_f + \hat{\beta})}{w_\Gamma} \rfloor$), let the inverse image of f under c be

$$c^{-1}[\{f\}] = \{e \in \mathcal{E} \mid iw_\Phi \leq \phi_e \leq (i+1)w_\Phi \wedge jw_\Gamma \leq \rho_e \leq (j+1)w_\Gamma\}$$

  where ϕ_e and ρ_e are the Φ and Γ values of an element $e \in \mathcal{E}$ in the LCF.

- Map all remaining elements of $\mathcal{E}$ to the zero vector in $\mathcal{F} \cup \bar{0}$. i.e.

$$c\left[\mathcal{E} - c^{-1}[\mathcal{F}]\right] = \bar{0}$$

2.4 The general case

Even in the 2-dimensional case, ambiguity may arise in the search for the triple $(\hat{\beta}, \hat{x}, \hat{y})$. Instances where ambiguity may arise include the following cases:

1. For certain values of $\gamma = iw_\Phi$, $\mathcal{H}_{\mathcal{F}}^{\Phi,\Gamma}|_{\phi=\gamma}$ has regularly spaced non-zero bins, but not all the features corresponding to these bins are detected. In this case, the goodness of a match of $\mathcal{H}_{\mathcal{F}}^{\Phi,\Gamma}|_{\phi=\gamma}$ with $\mathcal{H}_{\mathcal{E}}^{\Phi,\Gamma}(\hat{\beta}, 0, 0)|_{\phi=\gamma}$ may be approximately equal for multiple values of Γ-shift d_i. An example of such an $\mathcal{F}$ is a set of parallel line segments, not all of which are detected. In this case a "mismatch" would be rejected as an outlier by a robust estimator of translation $(\hat{x}, \hat{y})$.

2. $\mathcal{F}$ contains features regularly spaced in Φ. In this case, if not all features in $\mathcal{F}$ are detected, the goodness of match of $\mathcal{H}_{\mathcal{E}}^{\Phi,\Gamma}(\beta)$ with $\mathcal{H}_{\mathcal{F}}^{\Phi,\Gamma}$ may be approximately equal for multiple values of β. An example of such an $\mathcal{F}$ is the set of line-segments representing a square room. In this case, the correlation between $\mathcal{H}_{\mathcal{F}}^{\Phi,\Gamma}$ and $\mathcal{H}_{\mathcal{E}}^{\Phi,\Gamma}(\beta)$ is equally good for four distinct values of β.

In the first case, the ambiguity may be resolved, or at least reduced, by using a different method for finding the best translation $(\hat{x}, \hat{y})$. As suggested in section 2.3, once $\hat{\beta}$ has been found, it is possible to use a spatial representation of

elements of $\mathcal{F}$ and $\mathcal{E}$ to find the translation necessary for a good correspondence match. Translations by (x, y) are evaluated according to the amount by which each element $e \in \mathcal{E}$, rotated by $\hat{\beta}$ and translated by (x, y) *overlaps* its counterpart in $\mathcal{F}$. For the second case, each candidate rotation can be evaluated by finding the best possible translation by the afore-mentioned spatial approach. The drawback of the spatial correlation approach is its computational inefficiency. This is exacerbated in the second case, where rotation and translation are no longer decoupled.

3 Pose estimation

Once correspondences have been established between extracted features and map features, a transformation must be found between the local and global (map) coordinate frames which maximizes some matching evaluation function. This decoupling of the correspondence matching phase from the pose estimation phase improves the precision of localization: correspondence matching is rastor-based and therefore computationally efficient but accurate only to the resolution of the $\mathcal{H}_x^{\Phi,\Gamma}$ and $\mathcal{H}_\varepsilon^{\Phi,\Gamma}$ histograms; we can now afford to use a more precise approach to pose estimation with the knowledge that only "correct" correspondences are contributing to the final pose estimate. As far as orientation estimation is concerned, we describe an approach which is *linear* in the number of extracted features, combining precision with efficiency.

We divide the approach to aligning the LCF with the GCF into three phases: feature merging, translation estimation and rotation estimation.

3.1 Feature merging

The correspondence matching approach described in section 2 produces as output a many-to-one function $c : \mathcal{E} \rightarrow \mathcal{F} \cup \bar{0}$ where $\mathcal{F}$ is the set of planar features within the mobile robot's current ultrasonic "field of view," and $\mathcal{E}$ is the set of features extracted by the clustering algorithm described in [16].

In order to reduce the computational intensity of the translation and rotation estimation phases described in sections 3.2 and 3.3, we merge into a single feature all extracted features in $\mathcal{E}$ which have been deemed to correspond to the *same* underlying feature in $\mathcal{F}$. In the notation of section 2, we construct a new set of features $\mathcal{E}'$ and a new correspondence function $c' : \mathcal{E}' \rightarrow \mathcal{F} \cup \bar{0}$ as follows: Let $Merge : \mathcal{P}(\mathcal{F}) \rightarrow \Gamma \times \Phi \times C_x \times C_y \times L$ be a function which takes a set of extracted line-segments as an argument and produces the line-segment resulting from a merger of scatter matrices, as described in [16]. Begin with $\mathcal{E}' = c^{-1}[\{\bar{0}\}]$ and $\forall e' \in \mathcal{E}'$, $c'(e') = c(e')$. Then, for each feature

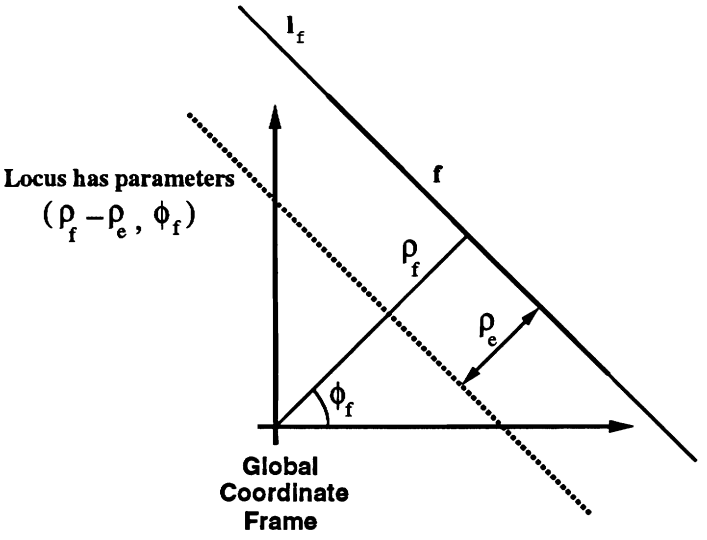


Figure 4: The locus (dotted) of positions (x_p, y_p) of the mobile robot within the GCF for which $l_{e'} = (\rho_{e'}, \phi_{e'})$ (not shown) lies on $l_f = (\rho_f, \phi_f)$.

$f \in \mathcal{F}$, add to $\mathcal{E}'$ the line-segment $e' = Merge\left(c^{-1}[\{f\}]\right)$ and define $c'(e') = f$.

Computational complexity: The weighted summation of 2 × 2 matrices is *linear* in the number of matrices. Computation of the eigenvectors of the resultant 2 × 2 scatter matrix requires a constant number of operations. Computation of endpoints of a line-segment representing a cluster of features may be performed in time *linear* in the size of the cluster. Hence, feature merging is *linear* in the number of extracted features.

3.2 Translation estimation

Let e'_i be the ith element of $\left(\mathcal{E}' - c'^{-1}[\{\bar{0}\}]\right)$. Denote the infinite line of which e'_i is a segment by $l_{e'_i}$. The normal parameters of $l_{e'_i}$ *in the LCF* are $\rho_{e'_i}$ and $\phi_{e'_i}$. Denote the underlying feature corresponding to e'_i by $f_i = c'(e'_i) \in \mathcal{F}$, and the infinite line of which it is a segment by l_{f_i}. The normal parameters of l_{f_i} *in the GCF* are ρ_{f_i} and ϕ_{f_i}.

If the LCF is rotated counter-clockwise by $\phi_{f_i} - \phi_{e'_i}$, then $l_{e'_i}$ can be made to lie on l_{f_i} for a suitable translation of the LCF. The locus of positions (x_p, y_p) of the mobile robot within the GCF for which $l_{e'_i}$ lies on l_{f_i} is the line with normal parameters $\rho_{f_i} - \rho_{e'_i}$ and ϕ_{f_i}. See figure 4. Hence, x_p and y_p are such that

$$x_p \cos\phi_{f_i} + y_p \sin\phi_{f_i} = \rho_{f_i} - \rho_{e'_i}$$

Hence, the set of constraints governing the position (x_p, y_p) of the mobile robot within the GCF may be expressed as

$$\mathbf{A}\begin{pmatrix} x_p \\ y_p \end{pmatrix} = \bar{d}$$

$$\text{where } \mathbf{A} = \begin{bmatrix} \cos\phi_{f_1} & \sin\phi_{f_1} \\ \cos\phi_{f_2} & \sin\phi_{f_2} \\ \vdots & \vdots \\ \cos\phi_{f_n} & \sin\phi_{f_n} \end{bmatrix} \text{ and } \bar{d} = \begin{pmatrix} \rho_{f_1} - \rho_{e'_1} \\ \rho_{f_2} - \rho_{e'_2} \\ \vdots \\ \rho_{f_n} - \rho_{e'_n} \end{pmatrix}$$

where $n = |\left(\mathcal{E}' - c'^{-1}[\{\bar{0}\}]\right)|$. A least squares estimate of the necessary translation of the LCF is, therefore,

$$\begin{aligned} \begin{pmatrix} x_p \\ y_p \end{pmatrix} &= \mathbf{A}^{\dagger}\bar{d} \\ &= \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\bar{d} \end{aligned} \tag{2}$$

Note the similarity in structure with the least squares approximation of translation described in section 2.3. The difference is that in section 2.3, *rastorized* histograms of $\mathcal{F}$ and $\mathcal{E}$ were used to obtain a *coarse* approximation to translation, and to eliminate mismatched features; in the present section, we assume knowledge of correspondence between features, and use *non-discretized* data to obtain a refined estimate of translation.

Once again, least squares estimation may be replaced here with truncated least squares, least median of squares, or some other robust estimator.

Computational complexity: The number of operations required for the evaluation of the matrix $\mathbf{A}$ is *linear* in n. The various matrix multiplications are also *linear* in n. The matrix inversion is performed on a 2×2 matrix, requiring a constant number of operations. Hence, the translation estimation phase is, overall, *linear* in n.

3.3 Rotation estimation

Once a suitable translation of the LCF has been found, all that remains is to find the rotation θ_p of the translated LCF which minimizes some disparity criterion. The criterion we have selected is the sum of squared perpendicular

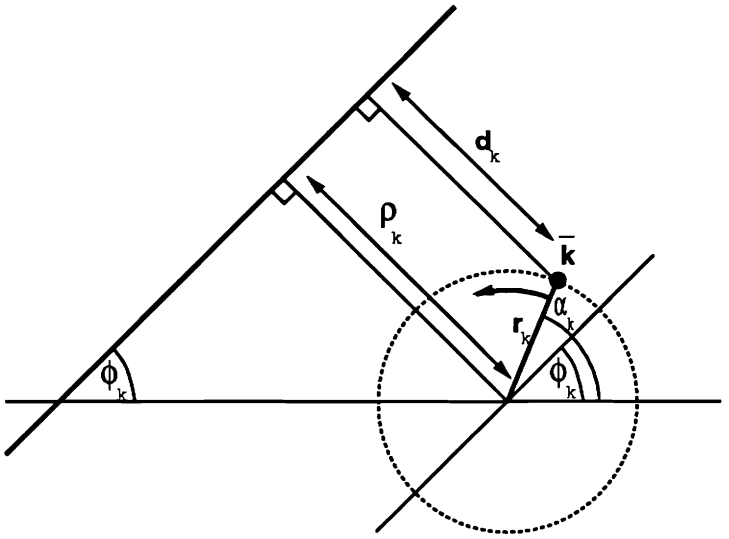


Figure 5: Perpendicular distance d_k of an arbitrary point $\bar{k} = (r_k, \alpha_k)$ (polar coordinates) from a line with with parameters ρ_k, ϕ_k) as a function of rotation θ of $\bar{k}$: $d_k(\theta) = \rho_k - r_k \sin(\alpha_k - \phi_k + \theta)$

distances between the endpoints of extracted features and the infinite lines to which these features have been deemed to correspond.

Let $\bar{k} = (r_k, \alpha_k)$ be an arbitrary point in the LCF expressed in polar coordinates. Let d_k denote the perpendicular distance of $\bar{k}$ from some line in the LCF with parameters ρ_k and ϕ_k. As can be seen from figure 5, after a counter-clockwise rotation of $\bar{k}$ by angle θ,

$$d_k = \rho_k - r_k \sin(\alpha_k - \phi_k + \theta) \tag{3}$$

Now, once again let e'_i be the ith element of $\left(\mathcal{E}' - c'^{-1}[\{\bar{0}\}]\right)$. Let the endpoints of e'_i be $\bar{k}_i$ and $\bar{j}_i$. Let the polar coordinates of $\bar{k}_i$ and $\bar{j}_i$ be (r_{k_i}, α_{k_i}) and (r_{j_i}, α_{j_i}) respectively. Denote the underlying feature corresponding to e'_i by $f_i = c'(e'_i) \in \mathcal{F}$, and the infinite line of which it is a segment by l_{f_i}. The normal parameters of l_{f_i} *in the GCF* are $\rho_{f_i}^{GCF}$ and $\phi_{f_i}^{GCF}$. By equation 1, the parameters of l_{f_i} *in the unrotated LCF* are

$$\begin{pmatrix} \rho_{f_i}^{LCF} \\ \\ \phi_{f_i}^{LCF} \end{pmatrix} = \begin{pmatrix} \rho + x_p \cos\phi_{f_i}^{GCF} + y_p \sin\phi_{f_i}^{GCF} \\ \\ \phi_{f_i}^{GCF} \end{pmatrix}$$

For ease of notation, in what follows denote $\rho_{f_i}^{LCF}$ simply by ρ_i and $\phi_{f_i}^{LCF}$ by ϕ_i. Hence, for $n = |\left(\mathcal{E}' - c'^{-1}[\{\bar{0}\}]\right)|$, the sum of squared perpendicular distances after a counter-clockwise rotation of endpoints by θ is $\sum_{i=1}^{n}\left(d_{k_i}^2 + d_{j_i}^2\right)$,

where, by equation 3,

$$\begin{aligned}
\sum_{i=1}^{n} d_{k_i}^2 &= \sum_{i=1}^{n} \left(\rho_i - r_{k_i} \sin\left(\alpha_{k_i} - \phi_i + \theta\right)\right)^2 \\
&= \sum_{i=1}^{n} \left(\rho_i^2 - 2\rho_i r_{k_i} \sin\left(\alpha_{k_i} - \phi_i + \theta\right) + r_{k_i}^2 \sin^2\left(\alpha_{k_i} - \phi_i + \theta\right)\right) \\
&= \sum_{i=1}^{n} \rho_i^2 - 2\sum_{i=1}^{n} \rho_i r_{k_i} \sin\left(\alpha_{k_i} - \phi_i + \theta\right) + \sum_{i=1}^{n} r_{k_i}^2 \sin^2\left(\alpha_{k_i} - \phi_i + \theta\right) \\
\text{and since } \sin^2\gamma &= \frac{1 - \cos 2\gamma}{2}, \\
\sum_{i=1}^{n} d_{k_i}^2 &= \sum_{i=1}^{n} \rho_i^2 - 2\sum_{i=1}^{n} \rho_i r_{k_i} \sin\left(\alpha_{k_i} - \phi_i + \theta\right) + \sum_{i=1}^{n} r_{k_i}^2 \left(\frac{1 - \cos\left(2\alpha_{k_i} - 2\phi_i + 2\theta\right)}{2}\right) \qquad (4)
\end{aligned}$$

$\sum_{i=1}^{n} d_{j_i}^2$ has similar form. As a function of θ, the second term in equation 4 is the sum of sinusoids of equal frequency 1. It is, hence, also a sinusoid of frequency 1. Similarly, the third term in equation 4 is the sum of sinusoids of frequency 2. It is, therefore, also a sinusoid of frequency 2. Hence, we may write

$$g(\theta) = \sum_{i=1}^{n} \left(d_{k_i}^2 + d_{j_i}^2\right) = B_1 + B_2 \sin\left(\theta + \mu_1\right) + B_3 \sin\left(2\theta + \mu_2\right) \qquad (5)$$

for some constants B_1, B_2 and B_3, μ_1 and μ_2. We seek the value of θ which minimizes $g(\theta)$. This function has at most two troughs in the interval $\theta \in [0, 2\pi)$. It can be shown that if we sample g at four points equally spaced in θ, call them θ_1, θ_2, θ_3 and θ_4, $\theta_j - \theta_{j-1} = \frac{\pi}{2}$, $2 \leq j \leq 4$, then $\min_{1 \leq j \leq 4} g(\theta_j)$ lies in a trough containing a *global* minimum of $g(\theta)$. Hence, an efficient approach to finding θ_p for which $g(\theta_p)$ is a global minimum is

1. Evaluate $g(0)$, $g(\frac{\pi}{2})$, $g(\pi)$ and $g(\frac{3\pi}{2})$ using equation 4.
2. Find the minimum of these four values.
3. Use the value of θ corresponding to this minimum as an initial estimate in a Newton-Raphson (NR) iterative approximation of θ_p.

Since the NR algorithm converges rapidly, only a few evaluations of the function $\sum_{i=1}^{n} \left(d_{k_i}^2 + d_{j_i}^2\right)$ by way of the right hand side of equation 4 will be necessary. Nevertheless, for large values of n, evaluations of $\sum_{i=1}^{n} \left(d_{k_i}^2 + d_{j_i}^2\right)$ may become prohibitively computationally intensive. For this reason, we use the following approach:

We evaluate $\sum_{i=1}^{n} \left(d_{k_i}^2 + d_{j_i}^2\right)$ at five points in order to solve for the five unknown parameters of $g(\theta)$ in equation 5. Let the five values of θ at which we evaluate $\sum_{i=1}^{n} \left(d_{k_i}^2 + d_{j_i}^2\right)$ be 0, $\frac{\pi}{4}$, $\frac{\pi}{2}$, π and $\frac{3\pi}{2}$. Then, substitution into equation 5

yields

$$g(0) = B_1 + B_2 \sin\mu_1 + B_3 \sin\mu_2 \quad (6)$$

$$g\left(\frac{\pi}{2}\right) = B_1 + B_2 \cos\mu_1 - B_3 \sin\mu_2 \quad (7)$$

$$g(\pi) = B_1 - B_2 \sin\mu_1 + B_3 \sin\mu_2 \quad (8)$$

$$g\left(\frac{3\pi}{2}\right) = B_1 - B_2 \cos\mu_1 - B_3 \sin\mu_2 \quad (9)$$

$$g\left(\frac{\pi}{4}\right) = B_1 - B_2 \cos(\frac{\pi}{4} + \mu_1) - B_3 \cos\mu_2 \quad (10)$$

(11)

$$(6)+(7)+(8)+(9) \Rightarrow \quad B_1 = \frac{g(0) + g(\frac{\pi}{2}) + g(\pi) + g(\frac{3\pi}{2})}{4} \quad (12)$$

$$(6)-(8) \Rightarrow \quad B_2 \sin\mu_1 = \frac{g(0) + g\left(\frac{\pi}{2}\right)}{2} \quad (13)$$

$$(7)-(9) \Rightarrow \quad B_2 \cos\mu_1 = \frac{g\left(\frac{\pi}{2}\right) + g\left(\frac{3\pi}{2}\right)}{2} \quad (14)$$

$$(13)\,/\,(14) \Rightarrow \quad \mu_1 = \tan^{-1}\left(\frac{g(0) + g(\pi)}{g\left(\frac{\pi}{2}\right) + g\left(\frac{3\pi}{2}\right)}\right) \quad (15)$$

$$(13) \Rightarrow \quad B_2 = \frac{g(0) + g(\pi)}{2\sin\mu_1} \quad (16)$$

$$(6)+(8) \Rightarrow \quad B_3 \sin\mu_2 = \frac{g(0) + g(\pi) - 2B_1}{2} \quad (17)$$

$$(10) \Rightarrow \quad B_3 \cos\mu_2 = B_2 \sin\left(\frac{\pi}{4} + \mu_1\right) - g\left(\frac{\pi}{4}\right) \quad (18)$$

$$(17)\,/\,(18) \Rightarrow \quad \mu_2 = \tan^{-1}\left(\frac{g(0) + g(\pi) - 2B_1}{2\left(B_2 \sin\left(\frac{\pi}{4} + \mu_1\right) - g\left(\frac{\pi}{4}\right)\right)}\right) \quad (19)$$

$$(17) \Rightarrow \quad B_3 = \frac{g(0) + g(\pi) - 2B_1}{2\sin\mu_2} \quad (20)$$

From equations 12, 15, 16, 19 and 20, we obtain an explicit formulation by which $g(\theta)$ may be evaluated iteratively at low computational cost.

Computational complexity The five evaluations of $\sum_{i=1}^{n}\left(d_{k_i}^2 + d_{j_i}^2\right)$ may be performed in time *linear* in the number of matched extracted features n. The time required for the iterative application of the NR algorithm is negligible compared to this. Thus, rotation estimation may be performed in time *linear* in n.

In this way, a precise estimate of the pose (x_p, y_p, θ_p) is obtained. The overall computational complexity of the pose estimation phase is *linear* in the number of extracted features.

4 Experiments

We are currently involved in the empirical testing and evaluation of the algorithm, both in simulation and on our physical testbed agents. Simulation results are very encouraging. A thorough analysis of the localization algorithm's properties on a real testbed system requires a measure of ground truth against which pose estimates may be compared. A system for establishing this ground truth is currently under development. See [17] for an analysis of the algorithm's empirical performance in terms of speed, precision and region of attraction[1]. See [14] for a comparison of localization characteristics using this modality with localization using various combinations of other modalities.

Two preliminary testbed experiments illustrate the usefulness of our approach:

Region of attraction: In this experiment we show the large region of attraction of the localization algorithm, both for *relocation* and for *continuous localization*, at least for the simple case ($|\mathcal{F}| = 2$) shown in figure 6.

Ultrasound localization versus odometry In this experiment we compare odometry-based against ultrasound feature-based localization. In the absence of a method of establishing ground truth, we proceed as follows:

1. Begin from some marked point in space. Identify the GCF with the LCF.
2. Steer the robot in a loop, and return to the original point in space.
3. Compare the estimates of final location and orientation according to odometry and the ultrasound feature-based localization algorithm.

We note that in our case, *no* continuous localization was performed; *all* data was subject to odometric error, so that the feature-based localization algorithm was at a distinct disadvantage. If continuous localization or some form of ground truth positioning system were to be used instead of odometry to keep track of the robot's pose for the duration of the experiment, we would expect an improvement in the quality of the parameters of the extracted features. This would lead to a corresponding improvement in relocation at the conclusion of the experiment.

Nevertheless, as is shown in figure 7, the relocation algorithm succeeds in reducing odometric error by at least 50%.

[1] The region of attraction for a known pose $q \in \mathcal{Q}$ is defined as the region in pose space $\mathcal{Q}' \subseteq \mathcal{Q}$ such that, for any $q' \in \mathcal{Q}'$, a robot with initial LCF origin at q' is be able to localize itself to within some threshold distance of q. The region of attraction depends on the nature of the environment, the point q, and the features that have been extracted thus far.

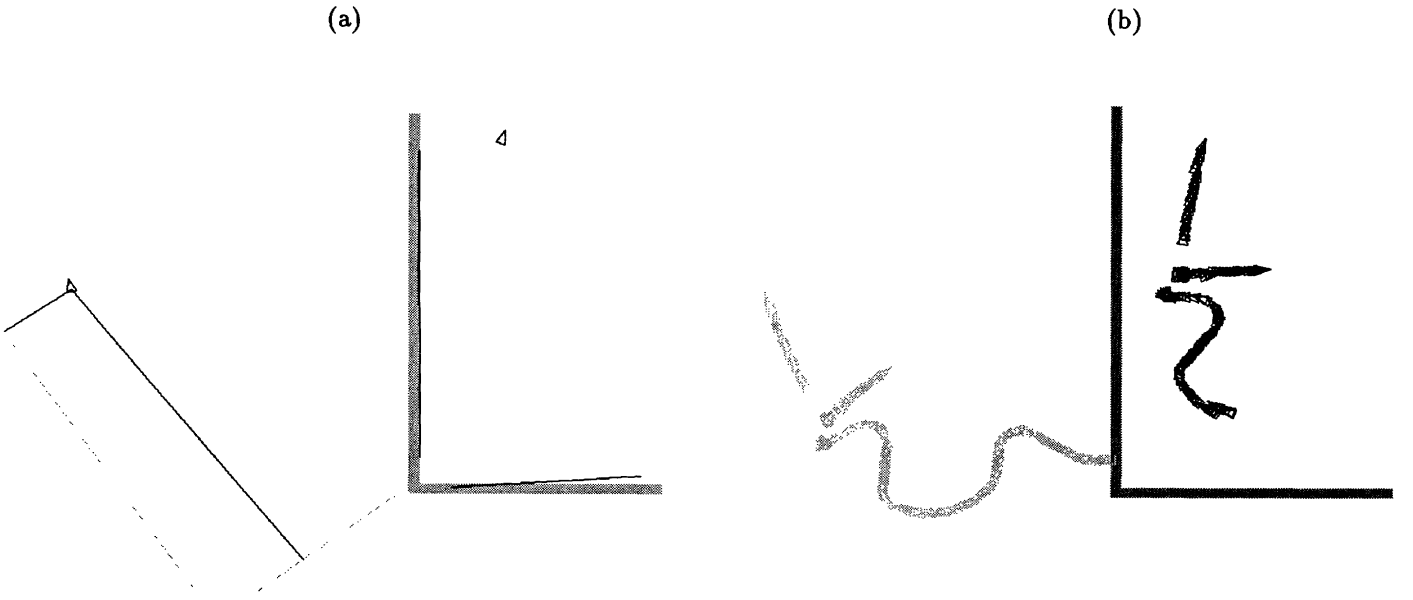


Figure 6: Experiment for which region of attraction is large. (a) Relocation. Light grey represents clusters of extracted features. Darker grey denotes features in the map. Dark lines on left show ρ values of extracted features in the LCF. Robot location in the LCF is shown on left. Dark features on right hand side represent the extracted features cast into the GCF after localization. The relocated robot is shown in its new pose in the upper right corner. (b) Continuous localization. Light grey represents the trajectory followed by the robot in the LCF. Darker grey represents environment structures (features in $\mathcal{F}$). Black represents the trajectory of the robot in the GCF as a result of continuous localization.

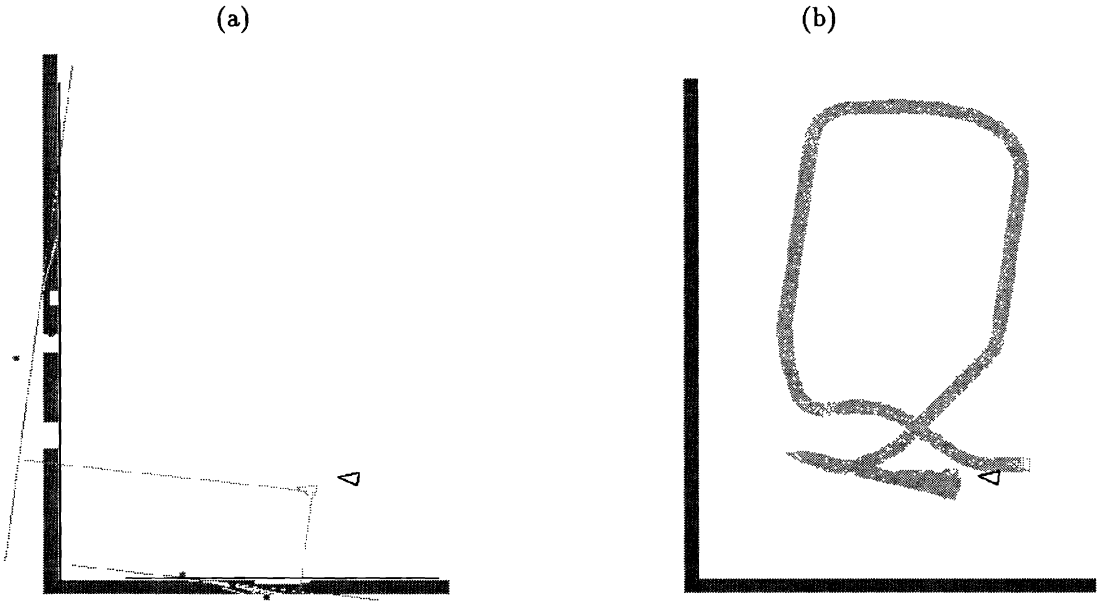


Figure 7: Experiment to compare odometry-based against ultrasound feature-based relocation. (a) Comparison of final orientation estimate. Light grey represents clusters of extracted features in the LCF, ρ values for these features, and the final pose of the robot according to odometry. Darker grey denotes features in the map. Dark features represent the extracted features cast into the GCF after relocation. Note the goodness of fit with map features. The relocated robot is shown in its new pose as a dark triangle. Note that localization algorithm reduces the odometric error in orientation by about 50%: in reality, final orientation was due west. This is in spite of the corruption of each individual ultrasound measurement by odometric error. (b) Comparison of final translation estimate. Light grey represents the trajectory followed by the robot in the LCF (i.e. according to odometry readings). Starting point is at the right. Endpoint cannot be clearly seen as the robot was maneuvered back and forth to ensure that, in reality, it ended at its starting point. Odometric slippage is clearly visible. Darker grey represents environment structures (features in $\mathcal{F}$). Black represents the final estimated pose of the robot in the GCF as a result of relocation. The localization algorithm is seen to reduce translation error by about 50%, placing the robot closer to its known starting point. Once again, this is in spite of the corruption of each individual ultrasound measurement by odometric error.

5 Conclusion

We have described a *feature-based* localization algorithm for mobile robots equipped with fixed ultrasonic transducers. We do not assume the presence of beacons, nor do we require the modification of the environment in any way.

We employ the method delineated in [15, 16] for the extraction of planar features from ultrasound data in the local coordinate system of the mobile robot. The advantages of using extracted features rather than unclustered ultrasonic measurements for localization include

1. Greater robustness and noise immunity: each feature represents the combined information from *many* measurements over *relatively long* periods.
2. Greater speed since the same feature matches may be used over long periods.

We address the issues of correspondence matching and pose estimation. The algorithm described here decouples these two phases. The matching stage is histogram-based, yielding a coarse estimate of pose and a function mapping extracted features to features in the map. Using this function, the pose estimation stage makes use of least squares estimation to yield a refined estimate of translation and rotation. The pose estimation stage is shown to be *linear* in the number of extracted features.

The decoupling of the matching and pose estimation phases allows the speed of rastorized techniques to be combined with the precision and finer resolution of non-discretized estimation. Furthermore, the complete correspondence matching phase need only be invoked for the purposes of *relocation*. For the case of *continuous localization*, the same correspondence matches may be used over long periods; only the pose estimation phase need be invoked during these periods.

The approach we delineate minimizes a sum of squared errors expression. Rapidly convergent gradient-descent techniques are employed for this minimization in the case of orientation estimation; we avoid local minima by selecting a initial estimate which *guarantees* convergence to the *global* minimum. The overall result is a localization algorithm which is both computationally efficient and accurate. Such an algorithm is a key component for the tasks of navigation, exploration of partially known environments, and cooperative material handling by multiple agents.

We are currently involved in the empirical testing and evaluation of the algorithm, both in simulation and on our physical testbed agents. As part of this effort, a system for establishing ground truth is under development. The accuracy and region of attraction of the localization algorithm are under analysis. Further, the algorithm is

being extended to incorporate point-type ultrasound features (corners), as well as features extracted by other sensor modalities. See [14] for details.

References

[1] R. Bauer. Active manoeuvres for supporting the localisation process of an autonomous mobile robot. In *Proceedings of the International Workshop on Intelligent Robotic Systems '94*, July 1994. Grenoble, France.

[2] M. Betke and K. Gurvits. Mobile robot localization using landmarks. In *Proceedings of the IEEE International Conference on Robotics and Automation*, volume 2, pages 135–142, May 1994.

[3] S. Borthwick and H. F. Durrant-Whyte. Simultaneous localisation and map building for autonomous guided vehicles. In *Proceedings of the IEEE International Conference on Intelligent Robots and Systems*, pages 761–768, September 1994. Munich, Germany.

[4] S. Borthwick, M. Stevens, and H. F. Durrant-Whyte. Position estimation and tracking using optical range data. In *Proceedings of the IEEE International Conference on Intelligent Robots and Systems*, pages 2172–2177, July 1993. Yokohama, Japan.

[5] R. G. Brown, L. P. Chew, and B. R. Donald. Mobile robots, map-making, shape metrics, and localization. *Submitted to IEEE ICRA*, 1993.

[6] J. L. Crowley. World modeling and position estimation for a mobile robot using ultrasonic ranging. *1989 IEEE International Conference on Robotics and Automation*, May 1989.

[7] R. O. Duda and P. E. Hart. *Pattern Classification and Scene Analysis.* Wiley-Interscience, 1973.

[8] A. Elfes. Sonar-based real-world mapping and navigation. *IEEE Journal of Robotics and Automation*, RA-3(3):249–265, June 1987.

[9] A. Elfes. Multi-source spatial fusion using bayesian reasoning. In M. A. Abidi and R. C. Gonzalez, editors, *Data Fusion in Robotics and Machine Intelligence*, pages 137–163. Academic Press, 1992.

[10] A. A. Holenstein, M. A. Müller, and E. Badreddin. Mobile robot localization in a structured environment cluttered with obstacles. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 2576–2581, May 1992. Nice, France.

[11] J. J. Leonard. *Directed Sonar Sensing for Mobile Robot Navigation.* PhD thesis, University of Oxford, 1990.

[12] J. J. Leonard and H. F. Durrant-Whyte. Mobile robot localization by tracking geometric beacons. *IEEE Transactions on Robotics and Automation*, 7(3), June 1991.

[13] P. MacKenzie and G. Dudek. Precise positioning using model-based maps. In *Proceedings of the IEEE International Conference on Robotics and Automation*, volume 2, pages 1867–1874, May 1994.

[14] R. Mandelbaum. *Sensor Fusion for Mobile Robot Localization, Exploration and Navigation.* PhD thesis, University of Pennsylvania. In preparation.

[15] R. Mandelbaum and M. Mintz. Sonar signal processing using tangent clusters. Technical report, University of Pennsylvania, 1994. In preparation.

[16] R. Mandelbaum and M. Mintz. Sonar signal processing using tangent clusters. In *Proceedings of the OCEANS '94: special session on Automated Unmanned Vehicles*, September 1994. Brest, France.

[17] R. Mandelbaum and M. Mintz. Testing and evaluation of a feature-based localization algorithm for fixed ultraonic transducers. Technical report, University of Pennsylvania, 1994. In preparation.

[18] B. Schiele and J. L. Crowley. A comparison of position estimation techniques using occupancy grids. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 1628–1634, May 1994.

[19] K. T. Sutherland. Landmark selection for accurate navigation. In *Proceedings of the 1993 Image Understanding Workshop*, pages 485–490, April 1993.

[20] K. T. Sutherland and W. B. Thompson. Inexact navigation. In *Proceedings of the IEEE International Conference on Robotics and Automation*, May 1993.

[21] W. B. Thompson, T. C. Henderson, T. L. Colvin, L. B. Dick, and C. M. Valiquette. Vision-based localization. In *Proceedings of the 1993 Image Understanding Workshop*, pages 491–498, April 1993.