Position Estimation Using Principal Components of Range Data

James L. Crowley, Frank Wallner and Bernt Schiele

Project PRIMA-IMAG, INRIA Rhône-Alpes
655, avenue de l'Europe
38330 Montbonnot Saint Martin, France

Abstract[1]

This paper describes a new approach to mobile robot position estimation based on principal component analysis of laser range data. An eigenspace is constructed from the principal components of a large number of range data sets. The structure of an environment, as seen by a range sensor, is represented as a family of surfaces in this space. Subsequent range data sets from the environment project as a point in this space. Associating this point to the family of surfaces gives a set of candidate positions and orientations (poses) for the sensor. These candidate poses correspond to positions and orientations in the environment which have similar range profiles. A Kalman filter can used to select the most likely candidate pose based on coherence with small movements.

The first part of this paper describes how a relatively small number of depth profiles of an environment can be used to generate a complete eigenspace. This space is used to build a representation of the range scan profiles obtained from a regular grid of positions and orientations (poses). This representation has the form of a family of surfaces (a manifold). This representation converts the problem of associating a range profile to possible positions and orientations into a table lookup. As a side benefit, the method provides a simple means to detect obstacles in a range profile. The final section of the paper reviews the use of estimation theory to determine the correct pose hypothesis by tracking.

1. Introduction

Navigation is generally defined as controlling motion to arrive at a given position. In the case of mobile robots, navigation can be decomposed into the subproblems of position estimation, path planning and local motion control. Hypotheses of the vehicle position and orientation are commonly derived from the interpretation of sensor signals provided by range sensors, beacon detectors or computer vision. It is common to maintain an estimate of position and orientation using odometry. Errors in this estimate are corrected using beacons or range sensors.

A common method for position estimation from range

---

sensors is to construct a structural description from sensor data and to match this description to a previously acquired model [Crowley 85]. An alternative is to project individual range measurements onto a previously acquired model [Leonard and Durrant-Whyte 91]. It is also possible to fuse range measurements directly using occupancy grids [Elfes 86], [Schiele 94]. Recently it has been shown that raw range data from nearby scans can be registered using a technique known as scan correlation [Weiss et al 95]. The correction vector from this technique provides a correction to position estimation. A thorough review in the state of the art in position estimation is provided in [Borenstein et al 96].

All of the above techniques attempt to register the observed sensor signal to an a priori model within a 3D Cartesian coordinate frame defined by the external environment. However, alternative coordinate spaces are possible. For example, for the case of a profile of range measurements, each measurement can be considered as a dimension in a space. In such an approach, a range scanner with a 0.5 degree resolution and a 180 degree field of view generates range observations which are points in a 360 degree space. Analysis of the date from such a scanner will show that adjacent range measurements are highly correlated. This is to be expected as most environments, particularly man made, have regular structure. Thus it is possible to use principal components analysis to determine a linear subspace with a minimum number of dimensions for representing an environment with a range sensor. This approach is well founded mathematically, which means that performance can be analyzed and predicted. Unlike previous approaches to position estimation, no ad hoc "tricks" are necessary. In this paper we show that such an approach can lead to a system which is reliable, robust to occlusions, and computationally simple.

The approach described here is a form of appearance based method. Similar methods have been adapted for navigation using computer vision [Jones et al. 97], for object recognition [Murase and Nayar 95] [Schiele 97] and for interpretation of deformable objects such as faces [Turk and Pentland 91]. In the work described here, we show how this approach can be adapted to the problem of position estimation using range sensors.

The paper is structured as follows: in the next section

the technique of principal component analysis of laser range data is described. The manner in which the resulting eigen-scans represent a sensor position is discussed. The third section compares different techniques for searching position hypotheses in eigenspace. The search is based on the detection of surrounding training scans of which the position is known. The question of detection and recovery from noisy data and scene changes is discussed in section 4. The final section shows how sequences of position hypothesis may be integrated using a probability model obtained by tracking hypotheses.
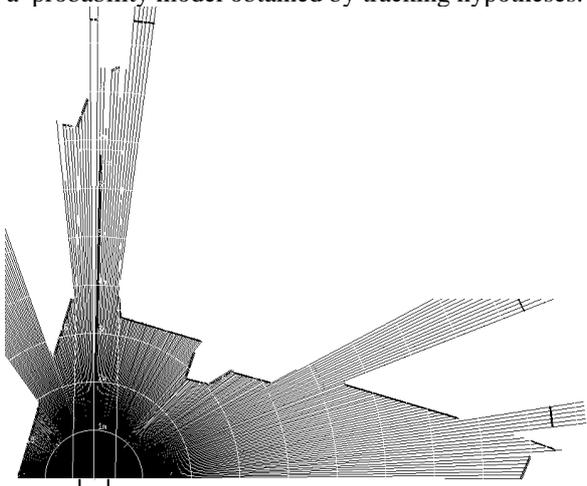

Figure 1a. An example of a scan from a laser range sensor. The scan is composed of 361 range readings over an angular field of 180°.

Figure 1b. An image of the center part of the scene observed in figure 1a.

The approach is currently implemented and tested using a commercial time-of-flight laser range scanner which gives data sets (called scans) which consist of 361 range values covering an angular field of 180°. Figure 1 illustrates the measuring properties of the laser scanner which detects obstacles between 4 cm and 50m.

2. Principal Component Analysis of Laser Range Data

With a scanning range sensor, each individual range measurement can be considered as an independent dimension. Such a view is not commonly applied because of the obvious correlation between adjacent measurements. This redundancy means that the dimensions in such a space are not orthogonal and that the space is not minimal. Principal component analysis (also known as Karhunen-Loëve expansion [Kirby and Sirovich 90]) provides a method to automatically determine a linear subspace with a minimal number of dimensions. Such a representation is optimal in the least-square error sense.

The principal component vectors from a covariance matrix of range data form an orthogonal basis set, which are the eigen-modes (or eigen-scans in the case of sets of range data) of the range sensor in that environment. By using only the most important eigen-vectors as defined by the corresponding eigen-values, the dimension of the eigenspace can be significantly reduced with a minimal loss in precision and reliability. This reduction can greatly reduce the memory and computing requirements.

For a N scans S_n, n=0,..., N-1 the mean scan is a vector is defined using the expectation operator E{}:

$$\mu = E\{S_n\} = \frac{1}{N} \sum_{n-0}^{N-1} S_n$$

the covariance matrix **A** is defined as the expectation of the outer product of the scan vector after subtraction of the mean.

$$\mathbf{A} = E\{(S_n - \mu)(S_n - \mu)^T\}$$

Eigen vector analysis determines a rotation matrix φ which projects **A** into an diagonal matrix Λ. The rotation matrix, φ and the diagonal matrix Λ are determined by solving the equation :

$$\varphi^T \mathbf{A} \varphi = \Lambda$$

Such a solution is easily obtained using the Jacobi Transformation or the more efficient Householder transformation, as described in [Press et al 92], and found in standard mathematical libraries. The matrix φ is composed of N rotation vectors (eigenvectors), $_n$. Each vector has unit length and is orthogonal to the other vectors.

$$\varphi = [\quad _0 | \quad _1 | \cdots | \quad _{N-1}]$$

The diagonal matrix Λ is composed of N diagonal terms $_n$, which are the variances (or eigenvalues) of the data in each dimension or row of φ. Principal

components analysis algorithms commonly provide the principal values in decreasing order of magnitude. The rows of the matrix φ constitute an orthogonal transformation for a data scan, S. Let S_O represent a range scan. A normalized scan, S, is obtained by subtracting the mean vector, μ,

$$S = S_O - \mu \qquad (1)$$

Multiplication of S by the matrix φ gives a vector

$$= [\ _0, \dots \ _{N-1}\ ] \text{ of N coefficients}$$

$$= \varphi^T S \qquad (2)$$

If all dimensions are preserved, this transformation is reversible, and the scan can be recovered as a sum of the eigen vectors.

$$\hat{S} = \sum_{n=0}^{N-1}\ _n\ _n \qquad (3)$$

When scans are correlated, the true dimension of the data is much less than N. In this case, The first M principal components can be used to form an M row by N column transformation matrix. The number of dimensions is commonly determined by eliminating dimensions for which the eigenvalues are small. The quality of the reconstruction can be evaluated by the energy in the error between the reconstruction and the observed scan.

$$E = S - \hat{S} \qquad (4)$$

An estimate of the quality of the reconstruction and thus the reliability of the projection is give by the energy in the error vector.

$$e = \|\ E\ \|$$

3. Acquiring a Sample Grid of Scans

Principal components will only be useful for position estimation if they reliably represent all possible sensor scans. When principal components are computed from scans taken at random positions (e.g. positions the robot has visited), scene characteristics which have often been observed will dominate. In order to obtain principal components which include all possible environmental appearances, the positions where data are collected should form a relatively dense sample grid.

The collection of a dense data grid at precise positions and orientations presents practical difficulties. The effort of training data collection can be greatly reduced by computing the training set synthetically. This can be done by assembling a composite description of the environment from an overlapping set of range scans taken at arbitrary positions and orientations. The technique of scan correlation described in [Weiss et al 95] can be used to determine the relative positions and orientations of these scans. The scans are then assembled into a composite scan in a common reference frame. The composite scan can be used to predict new scans for the dense grid of positions and orientations (poses). Note that this process is robust because only raw sensor data are used.

Given a set of overlapping data scans taken from unknown positions and orientations, scan correlation will determine the relative rotation and position at which the scans best overlap. Applying these transformations to the data projects the set of range readings to a common coordinate system. An example of a composition of four range scans is shown in figure 2.

Having constructed a composite scan, it is possible to generate a synthetic scan from any position and orientation. Each measurement in the synthetic scan corresponds to the nearest intersection of a ray with a data point from the composite scan, as shown in figure 3.

We obtain a dense sample of possible range data measurements by placing the synthetic scan generator at a grid of positions ever 50 cm. That is a total of 4 synthetic data sets per square meter. From each synthetic data set we can extract 720 different synthetic scans of 360 measurements each. Each synthetic scan covers 180 degrees, with half degree shifts in direction, thus simulating our range sensor. This gives a total of 2880 synthetic sub scans with half degree angular precision per square meter. We compute the mean and covariance of these synthetic sub-scans. The covariance matrix has 360 x 360 dimensions, but this one-time calculation is well within the bounds of Householders method for principal components analysis.
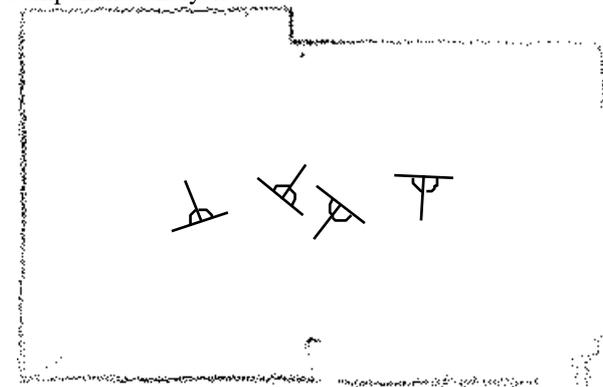

Figure 2. The composition of 4 range scans from the interior of the office of second author. The range scans were made at the 4 points marked with an icon. Each range scan covers 180° with 361 range measurements.

In our experiments, the eigen-value terms beyond the 3rd or 4th term are negligibly small. This is natural, as the underlying environment is a rectilinear structure in 2

dimensions. The range data is dominated by flat walls, concave and convex corners, and hallways. In a non-rectilinear environment, a larger number of dimensions would be necessary to adequately represent the range data. In our environment, 2 or 3 dimensions are generally sufficient for a more than adequate reconstruction.
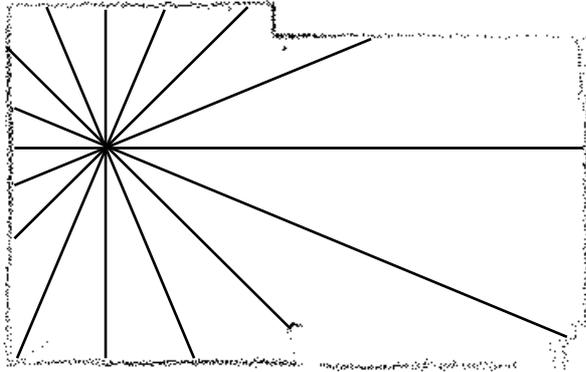

Figure 3. Synthetic range measurements every $\pi/8$ made from a composite scan. The actual scan synthesis program gives samples 720 measurements over an angle of 2π.

4. Representing the Pose-Space in Eigenspace

Consider the first few principle components from the covariance matrix of the synthetic scans from a collection of rooms. These components are orthonormal vectors which define a normed orthogonal space. For purposes of discussion, we will describe an eigenspace composed of three dimensions. A range scan at a particular position and orientation in the set of rooms is a point in this space. The set of possible range scans from a region of an environment at a particular orientation defines a surface in this space. Rotating the sensor deforms this surface. Thus the set of all range scans from an environment projects into a family of surfaces, or a manifold. This manifold is generally continuous but can contain discontinuities. To use this manifold for position estimation we must sample it at a discrete grid of positions and orientations.

The gird of positions and orientations constitutes a discrete sampling of a three dimensional Euclidean "pose-space" whose dimensions are (x, y, θ). For each pose-space sample, the range data set projects to a point in eigenspace. This collection of eigenspace points makes up a curved manifold in eigenspace. In most cases this manifold is continuous, although it is possible to have discontinuities near doorways and around convex corners. We can use a local linear approximations to this manifold as a model of the environment as observed by a range sensor.

Movements along a straight line in pose-space will not generally give a straight line in eigenspace. The

projection of the grid in eigenspace is non-Euclidean. The curvature of the grid in eigenspace determines the precision of a linear approximation, and thus the required sample density in pose-space. We can use the synthetic scan generator to sample the manifold to any required precision.

For purposes of discussion, let us consider a three dimensional eigenspace whose axes are defined by the first three principal components: (ϕ_0, ϕ_1, ϕ_2). The environment is described by a grid of synthetic scans generated at a discrete set of positions and orientations (x, y, θ) with a sample density of Δx, Δy, and $\Delta\theta$. In the experiments described below we used a fix sample grid of $\Delta x = \Delta y = 50cm$ and $\Delta\theta = 0.5$ degrees. Apart from the limited number of discontinuities, the mapping from pose-space to eigenspace is conformal: that is, continuous but curved. Scans from a triangle of points in pose-space will project to a non-Euclidean triangle of points in eigenspace. Assuming that curvature is small with respect to the triangle, the borders can be approximated by straight lines, as shown in figure 4.
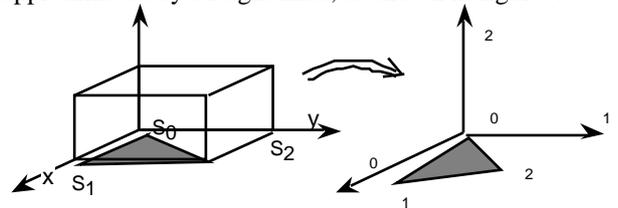

Figure 4. A triangle in pose-space maps to a triangle in eigenspace.
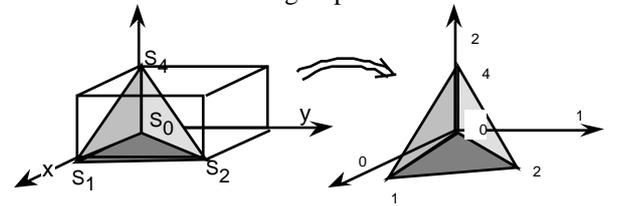

Figure 5. A tetrahedron in pose-space maps to a tetrahedron in eigenspace.

Four non-planar points in 3-D pose space define a tetrahedron, composed of four triangles. Assuming that curvature is small over the sides of the triangles, the corresponding volume in eigenspace can be approximated by a tetrahedron, as shown in figure 5. It is possible to refine the position estimation by interpolating in eigenspace, but such an interpolation is a linear approximation to a non-linear function.

5. Using the Range Manifold for Position Estimation

A range scan from the real sensor projects to a point in eigenspace. Listing all sample points within a given distance of the observed point gives a list of candidate

poses (positions and orientations) for the sensor. Multiple hypothesis points occur because different parts of the environment look similar to local view of the range sensor. Thus the projection of the grid representing the pose-space folds on itself. Some points in eigenspace may be traversed by several parts of the manifold. Thus indexing into the pose-space with a range scan can generate multiple hypotheses of pose.

5.1 Associating a Range Data Scan to the Eigenspace Manifold

The projection of a range scan gives the address of a cell in Eignespace and the cell provides a list of possible poses. If sufficient memory is available, the eigenspace can be represented directly as three dimensional array of cells. Each cell points to a list of poses for which a range scan projects to that cell. Alternatively, it is possible to use a hierarchical tree as an data structure.

The simplest association algorithm is to list the K nearest pose-grid points within a specified volume in eigenspace. These points can be sorted by distance in eigenspace to the scan projection. While the nearest point is very often close to the true pose, this is not always the case, and it is necessary to test the candidates.

The obvious means to reduce the number of pose hypotheses is to use the connectivity of the manifold structure. For example, when a surface passes through the volume around the projection of the scan, only the nearest point on the surface need be considered. We can extend this idea to the manifold using planar patches defined by triangles, or by using volumes defined by tetrahedron, as shown in figure 7.
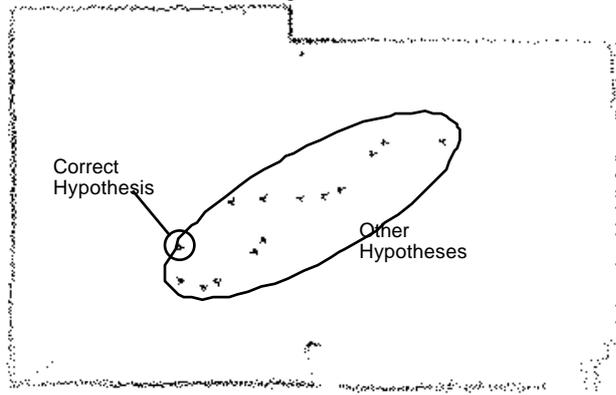

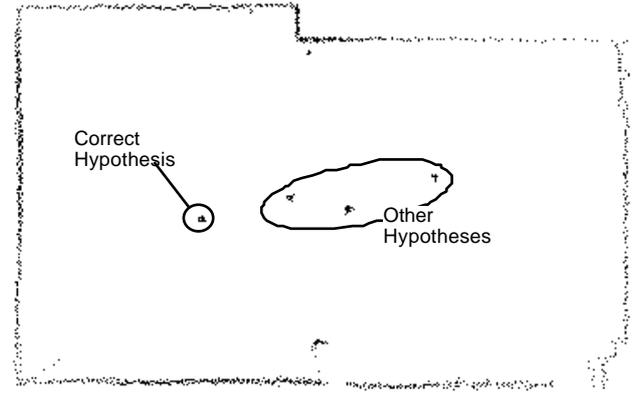**Figure 6.** Hypotheses from nearest neighbor search.

Figure 7. Hypotheses from volume search

5.2 Improving Precision by Interpolation in Eigenspace

The pose-space is sampled relatively densely in orientation, but somewhat sparser in position. We can improve the precision in position by interpolating using the three nearest poses. A simple interpolation algorithm works as follows. Give an associated pose grid cell, p_k, with scan projection, $_k$, determine the three adjacent grid cells in pose-space whose projections enclose . Determine the grid point, p_0, with the smallest values of x, and y. Determine the direction in eigenspace of the grid axes corresponding to a change in x, and a change in y. Project the scan onto these axes and measure the ratio of the distance from the point p_0 to the distance in eigenspace to the next grid sample. These ratios correspond to the correction factors, d_x, and d_y that can be applied to the pose corresponding to p_0.

6. Obstacle Detection

Once a number of potential scan-positions have been estimated, the range data can be analyzed by comparing them with a scan which was predicted for that position. Using the notation introduced above, let S_0 be an observed scan and S be the observed scan after subtraction of the mean vector. Within the precision of the error introduced by the reduction of eigenspace to M dimensions, an error vector can be computed from the difference between an observed scan and a reconstructed scan, as given in equation 4 above.

$$E = S - \hat{S} = S - \sum_{n=0}^{M-1} \quad n \quad n$$

Regions of the scene which are occluded appear as regions of large values in the error vector. The confidence in the projection can be estimated from the energy of the error vector.

A mask, M, can be derived by applying a threshold to

the error vector. Whenever the value of an estimated range reading exceeds the observed range reading be more than the threshold, the corresponding term of M is set to 0. When the difference is less than the threshold, the term is set to 1. This mask can be used to detect unexpected obstacles. We have also experimented with improving the precision of position estimation, however, eliminating occluded range values tends to introduce a bias in position estimation.

7. Confirming Pose Hypotheses by Tracking

In most cases pose hypotheses have very different positions and orientations. If even an approximate estimate of position and orientation is available, this can generally be used to reject incorrect hypotheses. If the sensor is mounted on a vehicle equipped with an odometric position estimation system, this can be sufficient to reject incorrect hypotheses. If, for example, the odometric position estimation system is based on a Kalman filter [Crowley and Reignier 93], then the rejection is accomplished by the validation gate. However, if no prior information about position or orientation is available, then the hypotheses must be sorted by tracking multiple hypotheses while the robot makes local movements.

Integrating successive pose estimations is a text-book case of data association and tracking, using estimation theory. Association of poses to tracks requires an estimate of the precision of each estimate. Such an estimate may be obtained from the spacing of samples in the pose grid x, y, and . As a first approximation, we take the grid sample sizes as the standard deviation of a pose estimate. In fact the true error is somewhat smaller because of interpolation.

$$\begin{matrix} x & & x \\ y & = & y \\ y & & \end{matrix}$$

The covariance matrix, $\mathbf{C_p}$, is an estimation of the error

$$\mathbf{C} = \begin{matrix} x^2 & xy & x \\ yx & y^2 & y \\ x & y & 2 \end{matrix}$$

When a new pose estimate is created, it is assigned a covariance of $\mathbf{C_p}$. If there is no knowledge about the robot's motion, we can estimate a loss of position information from the maximum velocity of the robot in V_x, V_y, and V . This maximum velocity is multiplied by the time step, T, to estimate a growth in the covariance matrix.

$$\mathbf{W} = T^2 \begin{matrix} v_x^2 & v_{xy} & v_x \\ v_{yx} & v_y^2 & v_y \\ v_x & v_y & v^2 \end{matrix}$$

Thus for a hypothesized trajectory, the initial estimated precision when the hypothesis is first created at time T is.

$$\hat{C}_T = \mathbf{C}$$

the predicted precision after a time step t is given by

$$C^*_{t+\ t} = \hat{C}_t + \mathbf{W}$$

Given a set of observed pose hypotheses, Y_k, and a set of tracks X_i, the probability that an observed pose belongs to a track can be computed by a normal probability density function.

$$p(Y_k \mid X_i) = \frac{1}{\sqrt{2}\ \det(\mathbf{C})^{\frac{1}{2}}}\ e^{-\frac{1}{2}(Y_k - X_i)^T C_i^{*-1}(Y_k - X_i)}$$

The observed pose is associated to the track for which this probability is maximum. If the maximum probability is less than a threshold then a new track is created. If more than one pose is associated to a track then, the pose is forked into multiple hypotheses.

The estimated position and orientation for a track is computed using recursive estimation.

$$\hat{C}_t := (C^{*-1}_t + C^{-1})^{-1}$$

The new pose estimate for time t can then be computed as a weighted average:

$$\hat{X}_t := \hat{C}_t (C^{*-1}_t X^*_t + C^{-1} Y_k)$$

Of course, if no pose hypothesis is associated with a track, the predicted pose and covariance for the track are taken as the new estimate.

Scans are integrated in an existing sequence if their difference in position does not exceed a maximum defined by the observation cycle and the maximum robot velocity. To a sequence of scans taken at the positions. Under the assumption that the robot is located inside the grid area, the probabilities of all sequences are normalized to '1'. Tracks are eliminated when their probability drops below a threshold. Figure 9 shows an example of hypothesis generation and trajectory reconstruction. Shown in each figure is the position of all hypothesis which form a trajectory and the actual probability of this trajectory. The robot moves from the low-left corner to the upper right corner.

Hypothesis which can not be associated with an existing scan sequence, create a new sequence to which a default minimum probability is assigned. A number of refinements which will only briefly be mentioned are the split of sequences in case two hypothesis are available, the unification of hypothesis and a time factor which

decreases the probability of sequences for which no longer hypotheses are generated and finally deletes them. Figure 9 gives an example for the current state of implementation of the approach. The square structure of the room in which the robot moves generates four potential hypothesis. The in symmetries of the room confirm the correct path within few observation cycles.

8. Conclusion and Perspective

This paper presents a new approach to position estimation for a mobile robot based on principal component analysis of range data. From each set of range data, a number of hypothesis which include a position estimation and the probability of this estimation is computed. The hypothesis are combined to sequences of estimation which describe a robot trajectory. The main advantages of this approach are computational efficiency and robustness to occlusion errors.

Because no a priori position knowledge is required, the approach can be used as a stand alone process to re-localize the robot or to track its current position. In this case, it is difficult to guarantee a precision for the final position estimation, especially if the robot moves in a changing scene. None-the-less, the probability model of scan sequences will reliably detect such cases and indicate the confidence of the position. The technique can also be used in combination with other estimation techniques. When combined with other approaches with other estimation techniques such as odometry, it is possible to survey and correct this techniques.

Because of space limitations, this article is restricted to explaining the foundation of the method and presenting examples. A much more complete set of evaluation experiments has been performed and is presented in [Wallner 97]. Currently the technique is applied to larger scenes. In that case, a number of overlapping subspaces are defined for regions such as rooms or hallways and each scan is transformed separately for each subspace. Resulting hypothesis are processed in the same manner as described above, even the ambiguities of the situation are much higher.
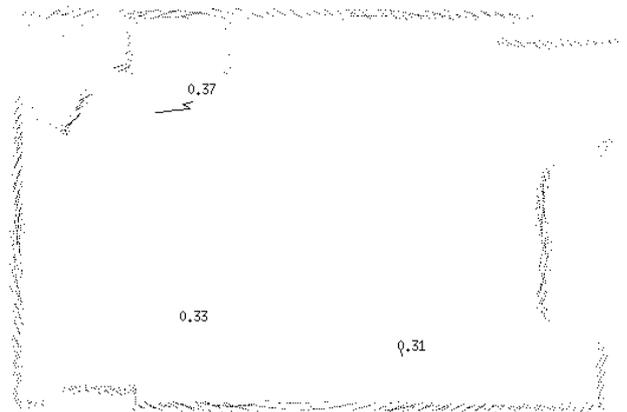

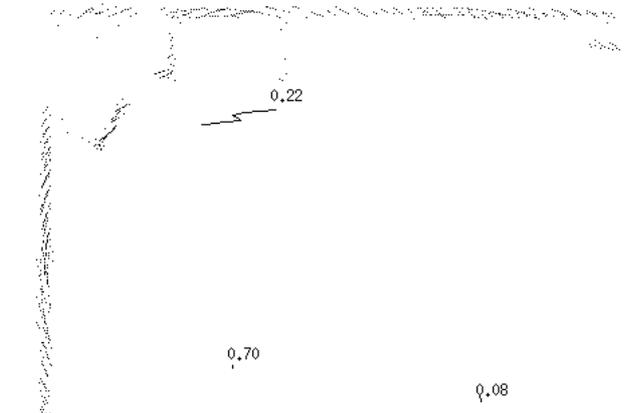**Figure 9a.** Three initial hypothesis trajectories with their probabilities tracked over 3 observations.

Figure 9b. The same hypothesis trajectories with their probabilities tracked over 5 observations
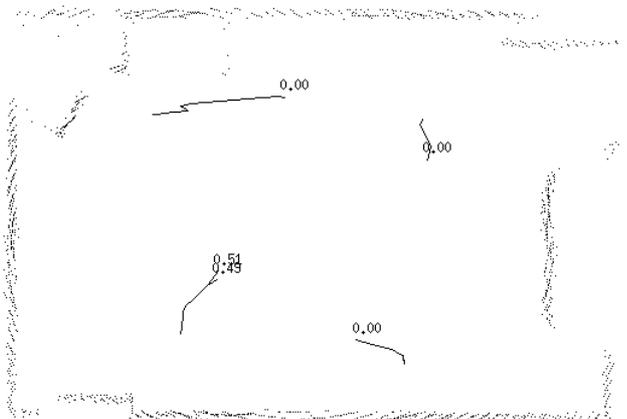

Figure 9c The hypothesis trajectories with their probabilities tracked after 15 observations.
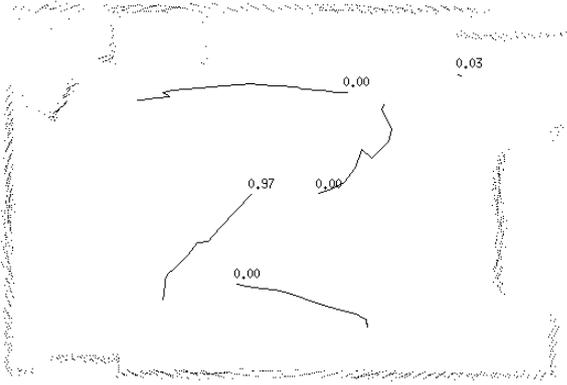
Figure 9d The hypothesis trajectories with their probabilities tracked after 26 observations.

Bibliography

[Borenstein et al 96]   J. Borenstein, H. R. Everett, L. Feng, "Where am I? Sensors and Methods for Autonomous Mobile Robot Positioning", Univ of Michigan Tech Report, 1996.

[Crowley 85]   J. L. Crowley, "Navigation for an Intelligent Mobile Robot", <u>IEEE Journal on Robotics and Automation</u>, JRA 1 (1), March 1985.

[Crowley and Reignier 93] J. L. Crowley and P. Reignier, "Asynchronous Control of Rotation and Translation for a Robot Vehicle", <u>Robotics and Autonmous Systems</u>, Vol 10, No. 1, January 1993.

[Elfes 89]   A. Elfes. Using occupancy grids for mobile robot perception and navigation. <u>IEEE Computer</u>, pages 46-57, June 1989.

[Jones et al 97]   S. D. Jones, C. Andersen and J. L. Crowley, "Appearance Based Processes for Visual Navigation", IROS 97, IEEE/RSJ Conference on Intelligent Robotics Systems, Grenoble, 1997.

[Kirby and Sirovich 90]   M. Kirby and L. Sirovich, "Application of the Karhunen Loeve Procedure for the Characterization of Human Faces", <u>IEEE Transactions on Pattern Analysis and Machine Intelligence</u>, PAMI Vol 12, No. 1, pp 103-108, 1990.

[Leonard and Durrant-Whyte 91]   J. Leonard and H. Durrant-Whyte, "Mobile Robot Localisation by Tracking Geometric Beacons", <u>IEEE Trans on Robotics and Automation</u>, Vol 7 No. 3, pp 376-382, 1991.

[Murase and Nayar 95]   H. Murase and S. K. Nayar, "Visual learning an recognition fo 3-D objects from appearance", <u>International Journal of Computer Vision</u>, Vol 14 pp 5-24, 1995.

[Press et al 92]   W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, <u>Numerical Recipies in C: The Art fo Scientific Computing</u>, Cambridge Univeristy Press, 1992.

[Schiele 94] B. Scheile and J. L. Crowley, "Certainty Grids: Perception and Localisation for a Mobile Robot", <u>Robotics and Autonomous Systems</u>, Vol 12, Nos 3-4, p 163-172, April 1994.

[Turk and Pentland 91]   M. A. Turk, and A. P. Pentland, "Eigenfaces for Recognition", <u>Journal of Cognitive Neuroscience</u>, Vol 3, No. 1, pp 71-86, 1991.

[Wallner 97] F. Wallner, "Position Estimation for a Mobile Robot from Principal Components of Laser Range Data", Doctoral Thesis, INPG, October 1997.

[Weiss et al 95]   G. Weiss, C. Wetzler, adn E. Van Puttkamer, "A Map Based on Laserscans without geometric interpretation", 1995 International Conference on Intelligent Autonomous Systems, IAS-4, Karlsruhe, pp 403-407, 1995.