Feature Extraction and Scene Interpretation for Map-Based Navigation and Map Building

Author(s):
Arras, Kai O.; Siegwart, Roland

Publication date:
1998

Permanent link:
https://doi.org/10.3929/ethz-a-010111542

Originally published in:
Proceedings of SPIE 3210, https://doi.org/10.1117/12.299565

Feature Extraction and Scene Interpretation for Map-Based Navigation and Map Building

Kai Oliver Arras, Roland Y. Siegwart

Swiss Federal Institute of Technology Lausanne
Institute of Microengineering
CH–1015 Lausanne, Switzerland
{arras, siegwart}@imt.dmt.epfl.ch

ABSTRACT

A scheme for extracting environment features from 1D range data and their interpretation is presented. Segmentation is done by deciding on a measure of model fidelity which is applied to adjacent groups of measurements. The extraction process is considered to include a subsequent matching step where segments which belong to the same landmark are to be merged while keeping track of those which originate from distinct features. This is done by an agglomerative hierarchical clustering algorithm with a Mahalanobis distance matrix. The method is discussed with straight line segments which are found in a generalized least squares sense using polar coordinates including their first-order covariance estimates. As a consequence, extraction is no longer a real time problem on the level of single range readings, but must be treated on the level of whole scans. Experimental results with three commercially available laser scanners are presented. The implementation on a mobile robot which performs a map-based localization demonstrate the accuracy and applicability of the method under real time conditions.

The collection of line segments and associated covariance matrices obtained from the extraction process contains more information about the scene than is required for map-based localization. In a subsequent reasoning step this information is made explicit. By successive abstraction and consequent propagation of uncertainties, a compact scene model is finally obtained in the form of a weighted symbolic description preserving topology information and reflecting the main characteristics of a local observation.

Keywords: feature extraction, range image segmentation, hierarchical clustering, localization, map building, scene modelling, symbolic description, relocation

1. INTRODUCTION

Reliable localization belongs to the basic skills of a mobile system whose task demands autonomous navigation in an unmodified environment. It can be approached qualitatively by a topology-based navigation scheme guided by static and locally unique features, or quantitatively by representing the position as a state vector in a global reference frame which is then to be estimated. In order to provide *precise* localization it is indispensible to represent uncertainties and to propagate them from single measurements to all system stages involved in the estimation process. Map-based localization with a Kalman filter where a measurement prediction based on odometry is matched with its corresponding observation in order to compute the best linear estimate of the position has become a widely used technique. The same method can also be used for feature extraction where the model parameters constitute the state vectors in question[1,2]. Due to its Markov property to represent all knowledge of the past in the present state, a Kalman filter can be recursively implemented such that the best estimate is instantly available each time a new measurement arrives. This can be perceived as a real time problem in the level of single range readings. The notion ‘real time’ is understood as the property of a system to process the incoming information in the speed with which it arrives. However, for the initialization of the filter which has to be done when starting and each time a terminating segment edge was detected, a first estimate of the feature position has to be found without the Kalman Filter. This is a regression problem which is nonlinear in the parameters if geometric meaningful errors are to be minimized. In the extraction process presented hereafter only such a method is utilized.We consider the extraction process as not yet finished with the segmented range image. Due to outliers, obstacles which partially occlude a landmark or discontinuities in the landmark’s structure or surface (e.g. a glass door in a wall) several segments are extracted belonging to the same environmental structure. Hence, it is required to associate these segments in order to involve as much information as available for a precise estimate of the final model parameters before using them for position update. This leads to the problem of feature discriminance and segment correspondence which is addressed hereafter.

In: Proceedings of SPIE, Mobile Robotics XII, Vol. 3210, 1997.

2. ESTIMATING A LINE AND ITS COVARIANCE MATRIX USING POLAR COORDINATES

A frequent case in regression is finding the best line to some data in Cartesian coordinates which have errors in only one direction, where 'best' is with respect to unweighted squared algebraic errors. This is a problem with known solution which is linear in the parameters. One dimensional range data are acquired in polar coordinates and if they have been obtained by a laser range finder it is a good approximation to neglect the angular uncertainty. In this case a reparametrization of the line model can be made such that the fit problem is lead back to the standard case in regression[2]. In general, range data uncertainties are in both coordinates and the question is which errors are then to be minimized by the fitting algorithm. The perpendicular distances from the points to the line is a choice which makes sense from the geometrical point of view because we don't want a solution in *some* algebraic sense but a solution which keeps track of the spatial or geometrical character of the problem. This yields a non-linear regression problem which is to be solved for polar coordinates. Under the assumption of known uncertainties we can further determine a weight for each measurement point, and fit the line in the generalized least squares sense. It can be shown that the solution is

$$\tan 2\alpha = \frac{\frac{2}{\sum w_i}\sum_{i<j}\sum w_i w_j \rho_i \rho_j \sin(\theta_i + \theta_j) + \frac{1}{\sum w_i}\sum (w_i - \sum w_j) w_i \rho_i^2 \sin 2\theta_i}{\frac{2}{\sum w_i}\sum_{i<j}\sum w_i w_j \rho_i \rho_j \cos(\theta_i + \theta_j) + \frac{1}{\sum w_i}\sum (w_i - \sum w_j) w_i \rho_i^2 \cos 2\theta_i} \tag{1}$$

$$r = \frac{\sum w_i \rho_i \cos(\theta_i - \alpha)}{\sum w_i} \tag{2}$$

where (ρ_i, θ_i, w_i) is measurement i and α and r are the parameters of the line model

$$\rho \cos(\theta - \alpha) - r = 0 . \tag{3}$$

For the implementation we suggest to use the identical Cartesian form of (1) and (2) which is of lower computational complexity

$$\tan 2\alpha = \frac{-2\sum w_i (\bar{y}_w - y_i)(\bar{x}_w - x_i)}{\sum w_i [(\bar{y}_w - y_i)^2 - (\bar{x}_w - x_i)^2]} \tag{4}$$

$$r = \bar{x}_w \cos\alpha + \bar{y}_w \sin\alpha \tag{5}$$

where $\bar{x}_w = 1/\sum w_i \sum w_i \rho_i \cos\theta_i$ and $\bar{y}_w = 1/\sum w_i \sum w_i \rho_i \sin\theta_i$ are the weighted means. For the final result of (1) or (4), the four-quadrant arc tangent has been taken. This sometimes generates (α, r)–pairs with negative r values. They should be detected in order to change the sign of r and to add π to α. All α–values lie then in the interval $-\pi/2 < \alpha \le 3\pi/2$. Starting from Cartesian coordinates, a third approach by solving a constrained least squares problem with a QR-decomposition and a singular value decomposition is to be considered[3]. Comparison of all three solutions showed that (4) and (5) is clearly the most efficient implementation regarding the number of floating point operations.

However, if we want to determine the covariance matrix we need to know (1) and (2). By writing the model parameters as random variables represented by their first-order Taylor series expansion about the mean point and assuming independence of P and Q, the first-order covariance estimate is

$$C_l = m_{\rho\theta} C_{\rho\theta} m_{\rho\theta}^T \tag{6}$$

$$= m_\rho C_\rho m_\rho^T + m_\theta C_\theta m_\theta^T \tag{7}$$

where

$$m_{\rho\theta} = \begin{bmatrix} m_\rho & m_\theta \end{bmatrix} = \begin{bmatrix} \frac{\partial\alpha}{\partial P} & \frac{\partial\alpha}{\partial Q} \\ \frac{\partial r}{\partial P} & \frac{\partial r}{\partial Q} \end{bmatrix} \tag{8}$$

is the $p \times 2n$ Jacobian matrix containing all partial derivatives of the model parameters with respect to P and Q about the mean

point and $C_{\rho\theta} = diag(C_\rho, C_\theta)$ the $2n \times 2n$ partitioned diagonal data covariance matrix with submatrices $C_\rho = diag(\sigma^2_{\rho_i})$ and $C_\theta = diag(\sigma^2_{\theta_i})$ where $i = 1, \ldots, n$.

According to the noise model for laser range finders whose scans are subsequently considered, the term in (7) which keeps track of angular uncertainties has been omitted yielding a parameter covariance matrix C_l with elements

$$\sigma_{\alpha\alpha} = \frac{1}{(D^2 + N^2)^2} \sum w_i^2 [N(\bar{x}_w \cos\theta_i - \bar{y}_w \sin\theta_i - \rho_i \cos 2\theta_i) - D(\bar{x}_w \sin\theta_i + \bar{y}_w \cos\theta_i - \rho_i \sin 2\theta_i)]^2 \sigma^2_{\rho_i} \tag{9}$$

$$\sigma_{rr} = \sum \left[\frac{w_i}{\sum w_j} \cos(\theta_i - \alpha) + \frac{\partial \alpha}{\partial P_i} (\bar{y}_w \cos\alpha - \bar{x}_w \sin\alpha) \right]^2 \sigma^2_{\rho_i} \tag{10}$$

$$\sigma_{\alpha r} = \sum \frac{\partial \alpha}{\partial P_i} \frac{\partial r}{\partial P_i} \sigma^2_{\rho_i} \tag{11}$$

where N and D are numerator and denominator of the right hand side of (1) or (4) respectively. Note that all elements are of complexity $O(n)$ and allow reuse of already computed expressions.

The weights can be chosen as the inverses of the variances giving most weight to data points with low variability, which seems desirable. By measuring the signal strength of the reflected beam knowing its relation to the variance of radial noise one can determine a weight for each range reading. However not all commercially available laser scanners provide this information. In this case the relationship of range variance to distance can be experimentally identified. It is obvious that all other sources which influence range variance (surface properties or beam splitting) are no longer taken into account.

3. SEGMENTATION

Segmentation of the range image is done by accumulating evidences in the model space similar to the Hough transform[4,5]. This leads to a clustering problem with n points which can be solved efficiently only by the particular character of the problem. n denotes the number of measurements.

The model is fitted into n_f neighboring points and the covariance matrix is computed. This is done for all points of the scan obtaining the same number of points in the model space[†]. When adjacent groups of range readings lie on the same landmark, their associated points constitute a cluster in the model space corresponding to that landmark. Feature extraction is now the task of finding these clusters. In a general case a clustering problem of this size where no a priori knowledge is available would lead to impracticably high computation times under real time conditions. Here the fact can be exploited that points on the same landmark are *almost always* consecutive points. Due to this underlying regularity in the acquisition process, a distance measure in the model space is defined which is applied to n_m adjacent points.

$$d_i = \sum_j (x_j - x_w)^T (C_j + C_w)^{-1} (x_j - x_w) \tag{12}$$

where $j = i - (n_m - 1)/2, \ldots, i + (n_m - 1)/2$ and x_w is the weighted mean

$$x_w = C_w \sum C_i^{-1} x_i, \tag{13}$$

$$C_w^{-1} = \sum C_i^{-1}. \tag{14}$$

Low distance indicates that the points involved have high model fidelity. If d_i is plotted against the measurement index, regions of low value can be expected at the corresponding index places of the sought clusters. A threshold d_m is applied cutting off the regions of low distance. A contributing segment is now defined to be the set of measurement points those representations in the model space satisfy

[†] By talking of 'points in the model space' the mean vectors are denoted although they are normal distributions, spatially diffused over the model space.

$$d_i \leq d_m . \tag{15}$$

If (15) holds only for one point of the model space, the minimal size of a segment is therefore n_f. The best estimate of feature position and covariance matrix at this point is then obtained by a new fit with all contributing points according to equations (4), (5) and (6).

This segmentation method allows also a pipeline implementation where each arriving measurement P_i triggers fitting of point $P_{i-(n_f-1)/2}$ and model fidelity check of point $P_{i-(n_f-1)/2-(n_m-1)/2}$. The delay would therefore be $(n_f + n_m)/2 - 1$ points where n_f and n_m are assumed to be odd numbers. Each time when condition (15) changes from true to false a terminating segment edge is detected and a new fit has to be done with all points constituting the segment. Note that (13) and (14) can also be derived from a special case of the Kalman filter equations or vice versa. Thus, the more efficient recursive implementation which minimizes the number of matrix inversions can be taken from this method.

4. FEATURE DISCRIMINANCE AND SEGMENT MERGING

As stated in section 1, feature extraction is considered to be not yet completed at that point. Valuable information would be lost, e.g., in the matching procedure of a Kalman filter-based localization task where only the observed segment which is closest to the prediction is taken whereas the others belonging to the same landmark would be ignored. On the other hand, two landmarks with differ only slightly in one or more of their model parameters, i.e. in their position, should be identified as being distinct within the limits which are given by sensor noise.

The matching problem, i.e. finding correspondence of objects represented by randomly distributed vectors in some model space, also called *data* or *measurement association*, arises, e.g., in the target tracking domain[7]. The goal is to produce an assignment from a measurement to a predicted target in the presence of clutter causing false measurements. A validation region around the prediction in the model space is used to discard or confirm a measurement. When dealing with multiple targets or clutter its likely that a measurement falls in two validation regions or several measurements lie in a single validation region requiring more complex probabilistic association concepts. In mobile robotics this problem arises not only if the Kalman filter is employed for localization[8] but generally when dealing with uncertain spatial information[9,10], e.g., when matching features in visual images[1,11]. However, opposed to this problem, it is always a pairwise correspondence which is sought, and not the assignment of multiple equally rated estimates to an unknown reference.

Clustering techniques provide this treatment of data, and unlike the previous section the problem size has changed. Having gained a number of segments from the method in section 3 or any other scheme providing propagation of the first two moments, the number of points in the model space has been significantly reduced. Starting from n, where n can be large, just the number of segments, n_s, remain which, e.g., for straight lines in structured environments is typically around fifteen, rarely over thirty. Hence, from the viewpoint of clustering techniques, segmentation can be perceived as being a preprocessing step in order to reduce input information leading to small clusters at the same locations where the original clusters have been observed before (figure 5.2). An agglomerative hierarchical clustering algorithm[12] which permits an efficient implementation is utilized and due to its simplicity a short outline is given:

Having computed the $n_s \times n_s$ symmetric distance matrix D, the procedure starts with each point as a separate cluster.

(1) Find the minimal distance d_{ij} of clusters Q_i and Q_j in D and while $d_{ij} \leq d_{\alpha_\varepsilon}$ proceed to

(2) Merge Q_i and Q_j obtaining Q_{ij}. The number of clusters is decreased by one.

(3) Update D by calculating the new distances from Q_{ij} to all other clusters. Only one column and row is changed. Then go back to (1)

The problem of determining the distance between two clusters is not straightforward and there are many propositions for its solution. Often it remains a matter of experience to select the most appropriate method. However, with the covariance estimates there is additional information on the points available which can be exploited for eliminating the usual ambiguities. There is no need of dealing with clusters consisting in more than one point since in the merging step, points are instantly fused by making a new model fit with the combined set of measurements. Hence, the distance between two clusters is always the distance of two normally distributed vectors in the model space and we can refer back to the matching problem where the Mahalanobis distance is widely employed for that purpose

$$d_{ij}^2 = (x_i - x_j)^T (C_i + C_j)^{-1} (x_i - x_j) \tag{16}$$

If x_i and x_j belong to the same landmark, d_{ij}^2 has chi-square distribution and an appropriate threshold $\chi^2_{\alpha_\varepsilon, 2}$ must be chosen. It is to be mentioned that if the points in the model space were truly normally distributed and therefore completely described by the first two moments, merging of step (2) could be efficiently done with equations (13) and (14) without leaving the model space. Experiments showed, however, that iterated fusion according to (13) and (14) leads to accumulating deviations from the (true) fit alternative, which for large structures where many segments have been merged, became non-negligible. The final estimates of the parameters is directly available after exiting the clustering algorithm. The final segments are obtained by combining their measurement points provided that the segments are adjacent and belong to the same line. They are used in section 6 and subsequently denoted when talking of 'segments' or 'joint segments'.

5. EXPERIMENTS, IMPLEMENTATION AND DISCUSSION

Scans of three different commercially available laser scanners are considered with varying specifications. The scan from the Acuity AccuRange3000LV has a medium angular resolution of nearly 1˚ on 360˚ and is of relative low quality (figure 5.1). The Sick PLS100 has a radial quantization of 5 cm and an angular resolution of 0.5˚ on 180˚ (figure 5.4)[††]. The scan from the Leuze RS3 exhibit low noise and has a high angular resolution of 0.25˚ on 180˚ (figure 5.6). Each measurement receives a standard deviation based on a linear relationship to the distance, indicated by 4σ–intervals. It was not considered to be crucial for the subsequent experiments to have elaborated sensor models for the range variances. See e.g. ref. 6 for a careful analysis of a ranging technique and its identification. n_f is chosen according to the ratio of angular resolution versus radial accuracy. The choice of $n_f = 7$ for the PLS100 and $n_f = 15$ for the RS3 yields the similar products with angular resolution of 3.5˚ and 3.75˚, reflecting their similar accuracy. The AccuRange3000LV exhibits more noise but has a lower angular resolution which leads also to the choice of $n_f = 7$. The other parameters have been kept constant in all experiments although tuning wouldn't have yield the suboptimal results for the PLS100-scans (figure 5.5). The intention is to demonstrate insensitiveness to the differing sensor specifications. They are $n_m = 3$, i.e. only the two neighboring points in the model space are taken, $d_m = 3$ and $\alpha_\varepsilon = 0.99998$.

A closer look on the extraction process is taken with the scan from the AccuRange3000LV (figure 5.1a). Figure 5.1b shows

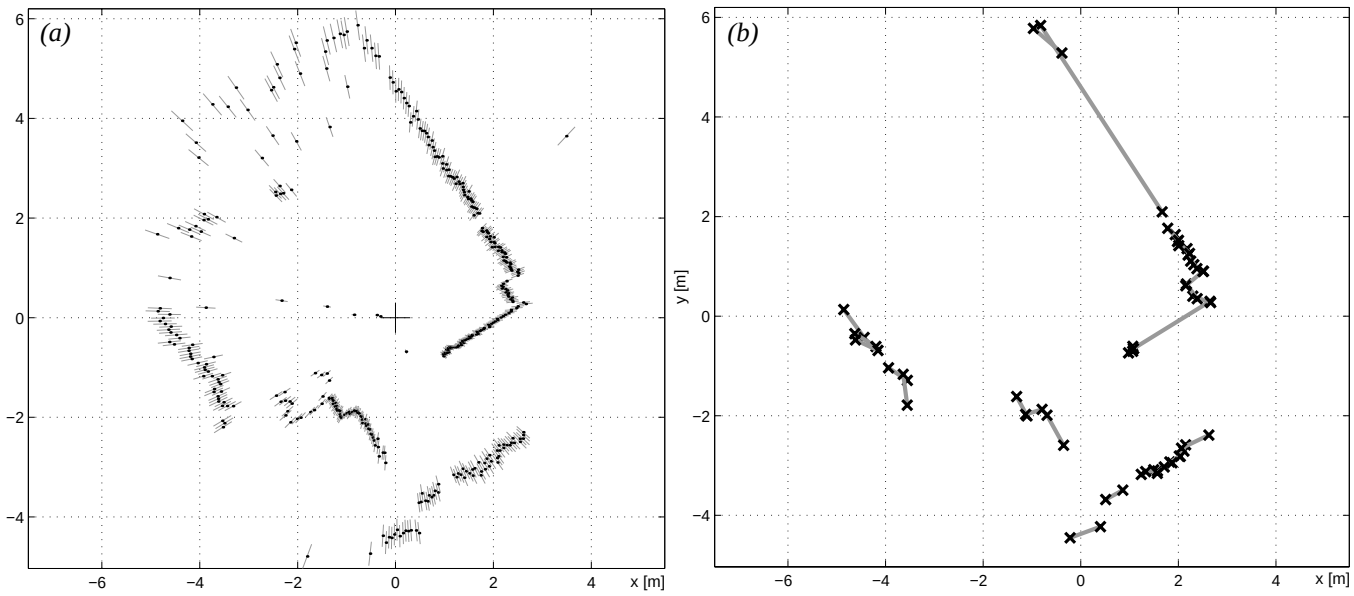

Figure 5.1 ***(a)*** *360 ˚-scan of the Acuity AccuRange3000LV. A scan with medium angular resolution (here 402 measurement points) while radial noise is relatively large. The simulated standard deviations are indicated by* 4σ *-intervals.* ***(b)*** *Range image after segmentation. The definition of a contributing segment allows overlapping segments.*

[††] It is to be mentioned that Acuity has a newer model, the AccuRange4000, and Sick is about to introduce the successor model of the PLS100, the LMS200. Both have better specifications.

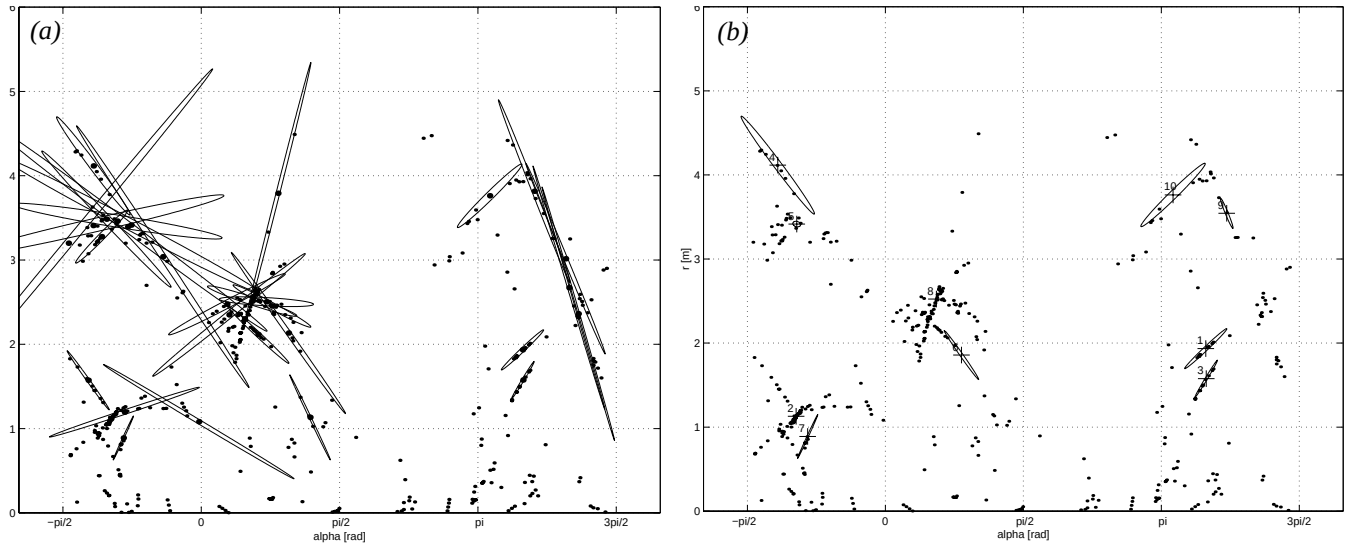

Figure 5.2 ***(a)*** *Segments seen in the model space. They are depicted by 95% error ellipses. Some of them belong to the same environmental structure and are therefore to be merged, others originate from different landmarks and are consequently to be distinguished.* ***(b)*** *Model space after the hierarchical clustering algorithm, only the final estimates remain.*

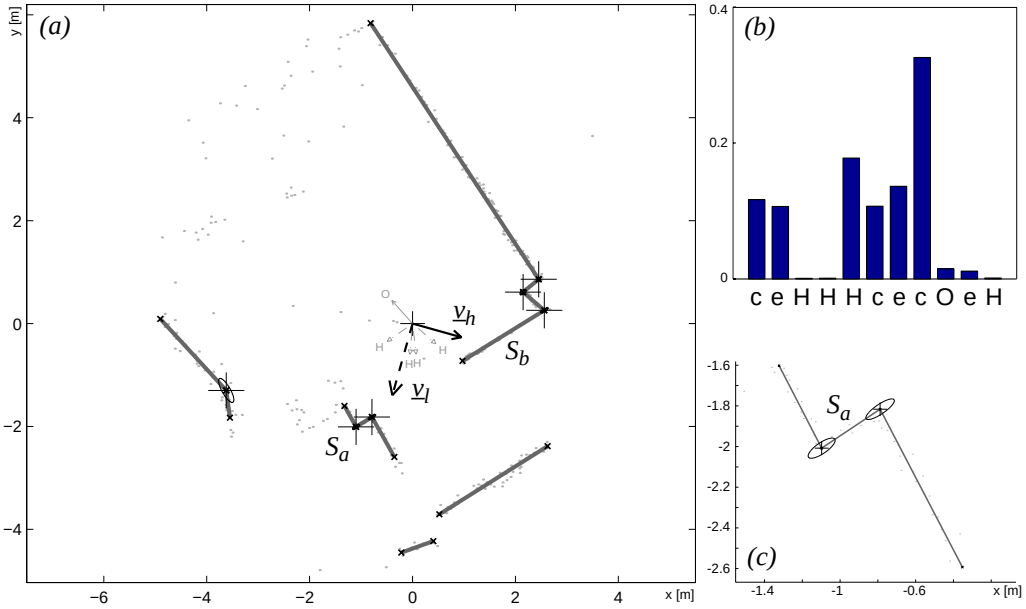

Figure 5.3 ***(a)*** *Extracted lines after the agglomerative hierarchical clustering. Corners which are found according to the scheme in section 6 are depicted by crosses and their 99% error ellipse.* ***(b)*** *Symbolic description with weights. They are a representation of the particularity of the scene and enable the calculation of v_h and v_l (see section 6).* ***(c)*** *Detail from (a). Corner probabilities are diffused in direction of the walls which is to be expected with perpendicular corners.*

the result of segmentation. There are thirty one extracted segments from this scan. Note that the definition of a contributing segment in section 3 permits overlapping segments. When merging two segments with common points, one has to keep track that they are taken only once for the joint fit. Each segment is displayed in the parameter space with its 95% error ellipse (figure 5.2). They constitute groups of ellipses with small Mahalanobis distances. Consider segment S_a in figure 5.3a/c. It has found to lie on the same line as segment S_b which has a great number of contributing points and therefore low uncertainty. The shape of the corner error ellipses in figure 5.3c reflect this by a low variability in the perpendicular direction.

5.1 Implementation

The experimental setup for the implementation consists of a Robosoft Robuter and the RS3 from Leuze, configured to have an angular resolution of 2˚ and a radial resolution of 4 mm. The CPU is a VME card with a Motorola 68020@20MHz. The implementation is reduced such that only the first statistical moments are propagated. For all distance measures their Euclidian equivalents have been taken, the thresholds were informally determined. The overall performance especially with respect to

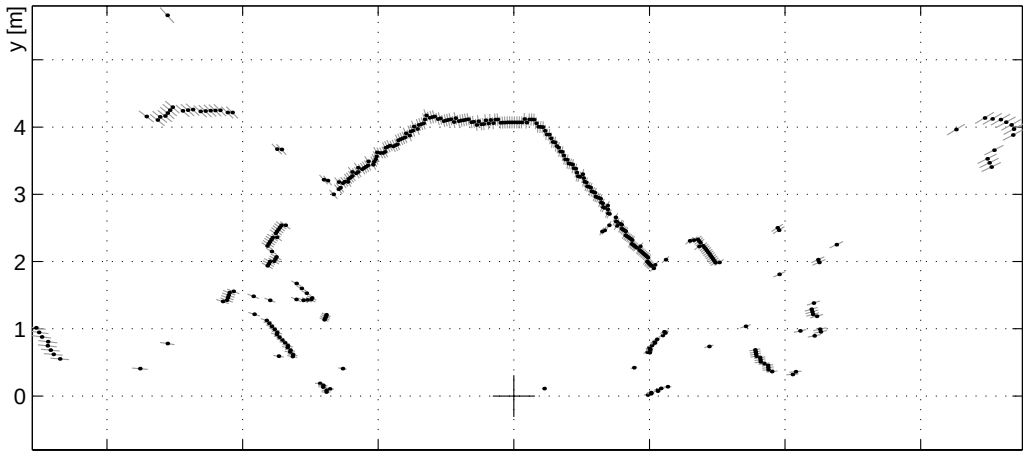

Figure 5.4 *Scan from Sick PLS100 which has a radial resolution of 5 cm. Small regions of constant depth can be identified where the sensor hits a target perpendicularly. Such on the upper horizontal wall at 90˚ and at about 45˚, where the big wall ends. This can lead to suboptimal results if the segment merging condition does not take this sensor characteristics into account.*

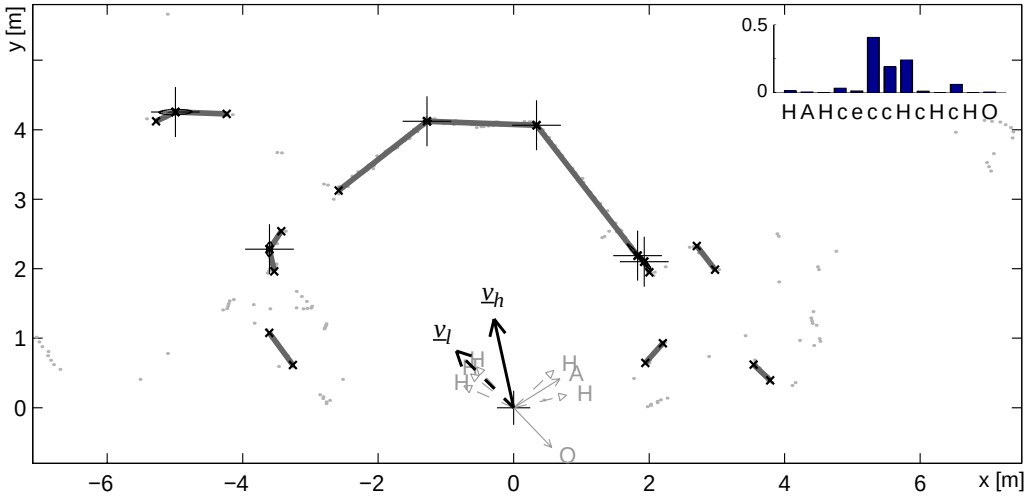

Figure 5.5 *Extracted lines, corners (indicated by crosses and their 99% error ellipse) and symbolic description with their weights (see section 6). The scene is mainly characterized by three big lines constituting two concave corners and adjacent hidden corners. The weights reflect this property causing a similar interpretation of the scene also when the small segments are not reliably found each time.*

robustness against outliers, is clearly inferior compared to the presented probabilistic version. Due to this, a demanding preprocessing which eliminates spurious points and thins out the raw range data became indispensable.

Firstly the extraction accuracy was identified. The robot was put into a corner about two meters away from a metal cupboard and a large cardboard box, where 300 extractions were made. The hit rate was 100%, the standard deviation in r was found to be 5 mm, standard deviation in α was 0.2˚. Computation time including preprocessing was not greater than 240 ms. Then the extraction was used for a map-based localization task with an extended Kalman filter. Position update while moving is accomplished where the robot can be manually pushed away from its given trajectory to which it responds with smoothly returning to its path within a few position update cycles. The localization accuracy which was immediately obtained without further optimization of any vehicle parameters like odometry was below one centimeter when the robot positioned itself at a desired location. As a further estimate for its precision, the noise of the position estimates has been identified when the robot was stationary. Standard deviation in x and y was found to be 6 mm, standard deviation in θ was 0.05˚. The average localization cycle time during a journey through lab and corridors was 445 ms.

5.2 Discussion

For the experiments on a number of scans from these sensors *no preprocessing* was performed. The method exhibits satisfactory robustness against outliers and cluttered regions. However, when the rotating beams hits no target many sensors return some maximal range value, which is the reception radius, r_s, a system parameter which will be used in section 6. In some cases lines were found in these ‘measurement points’ which were obviously wrong. It is justifiable to reject these point since they contain no information on the environment and they slow down the extraction process. The necessity of *postprocessing* depends on the treatment of uncertain features in the subsequent system stages like a matching procedure for localization or its incorporation in a geometrical map as new spatial information. The output of extraction can not be assumed to be perfect. See figure 5.7 where in the lower left-hand part of the image four small segments have been extracted where only three were expected.

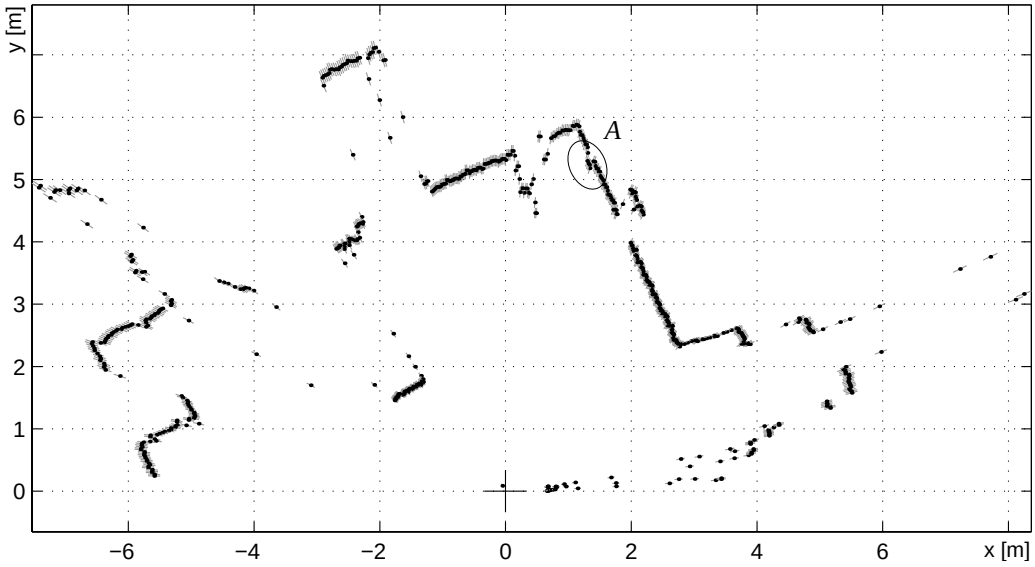

Figure 5.6 *Scan from the Leuze RS3 with an angular resolution of 0.25˚. Due to low noise also on distant targets, details are well reproduced. However, with simulated uncertainties only based on a distance relationship, other phenomen which influence range variance cannot be accounted for. The displaced points at position A originate from black velvet. Their true standard deviations are expected to be greater than the calculated ones.*

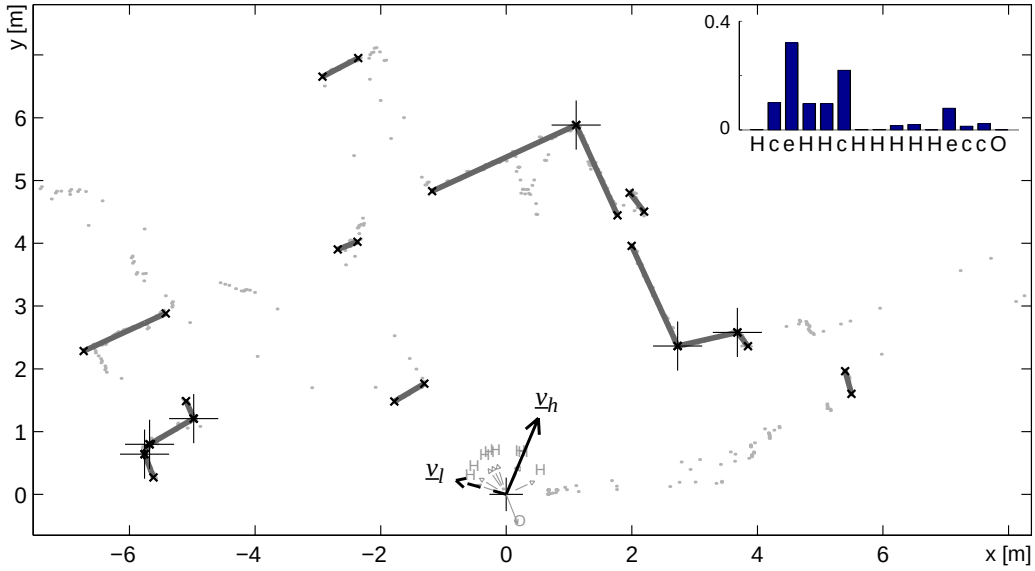

Figure 5.7 *Resulting lines, corners and weighted symbolic description. $\underline{v}_h$ denotes the vector of highest confidence, $\underline{v}_l$ the vector of least confidence. They concisely indicate where in the local environment a robot can expect to find secure terrain e.g. in order to reduce its own position uncertainty or where it can effectively gather new information on the environment.*

Although segments of this kind have a low weight, it is always possible to tighten the model fidelity condition (12) obtaining a conservative extraction where only landmarks with very high evidence will be found.

In general, model spaces are nonlinear manifolds whose points are sets of values with given physical dimensions. It is not easy to define distance measures for n_m points like (12) on such manifolds. Although (12) exhibited considerably better discrimination behavior than the Euclidian distance, the problem with equation (12) is that it has no standard distribution as one could believe. The advantage of the presented procedure is that it provides an understanding of the thresholds compared to an informal choice. However, in all experiments $n_m = 3$ showed good performance for all types of sensors. The problem would be significantly simplified if $n_m = 3$ is the general case. Then the appropriate distance measure is found with the smallest ellipse which goes through the three points. Their Mahalanobis distance to the center of the ellipse is constant and has a chi-square distribution.

6. FROM LINE SEGMENTS TO WEIGHTED SYMBOLS

The localization layer needs no more knowledge on the situation than is already available. It requires the estimated first two moments of the parameter vectors which correspond to the locally visible landmarks. For map building this is not sufficient. There is no higher layer which, based on some a priori knowledge, tells the robot what to do. As long as it cannot refer to already acquired information, the current sensory input has to furnish all the information which is needed for planning the next action.

Its evident that the understanding of scenes from a topological point of view is a central issue since a topology mistake would be a severe deficiency of the map. The detection of openings, doors or interesting sites for a inquiring robot when exploring an unknown environment incorporates often *ad hoc rules* which might work well in many cases. Subsequently it is tried to minimize this kind of knowledge when reasoning about the current scene. The result is finally summarized in a weighted symbolic description.

First a simple frame is derived which is free from heuristics and which serves as the base for following decisions. The procedure is to consider adjacent pairs of line segments and classify their spatial configuration. The classification feature makes use of the known scan direction of the range sensor. It is subsequently called the *directionality of an end point* λ_{E_i} and defined by the scan direction of E_i with respect to the intersection point $C_{i,\,i+1}$.

$$\lambda_{E_i} = \frac{\underline{s}_i^T \underline{d}_i}{|\underline{d}_i|} \tag{17}$$

Vector $\underline{s}_i$ is the rotated normalized perpendicular to the line l_i. Rotation angle is $+\pi/2$ for a counterclockwise scanning laser sensor, $-\pi/2$ otherwise. Vector $\underline{d}_i$ is the difference of E_i, the end point in question, and $C_{i,\,i+1}$. λ takes on two values, -1 and 1 corresponding to the cases 'scan direction of E_i is pointing *towards* $C_{i,\,i+1}$' and 'scan direction of E_i is pointing *away* from $C_{i,\,i+1}$'. It is then possible to formulate sufficient conditions for openings, hidden corners and concave or convex corners. The latter two are identifiable by $\Delta\alpha = \alpha_{i+1} - \alpha_i$. See figure 6.1 for illustration and conditions (18) to (21) for the so gained basic classification scheme.

$$\text{hidden corners, } H \;\Leftrightarrow\; \lambda_i = \lambda_{i+1} \tag{18}$$

$$\text{opening, } O \;\Leftrightarrow\; (\lambda_i = 1) \wedge \neg H \tag{19}$$

$$\text{concave corner, } c \;\Leftrightarrow\; ((0 < \Delta\alpha < \pi) \vee (\Delta\alpha < -\pi)) \wedge \neg O \tag{20}$$

$$\text{convex corner, } e \;\Leftrightarrow\; \neg c \tag{21}$$

Classification according to (18) to (21) is elegant but suffers from a lack of realism. It assumes infinite visibility radius of the sensory system and does not take the vehicle's dimensions into account, which is expected when looking for opening candidates. Concave corners which lie beyond the robot's sight, apertures between two segments or gaps in a wall originating from doors are not recognized as openings. Hence the scheme is to be extended by two system parameters, r_s, the reception radius of the sensor, and w_r, some constant reflecting the dimensions of the platform.

$$\text{gap, } G \;\Leftrightarrow\; (l_{S_i} = l_{S_{i+1}}) \wedge (d_{E_iE_{i+1}} > w_r) \tag{22}$$

$$\text{hidden corners, } H \;\Leftrightarrow\; (\lambda_i = \lambda_{i+1}) \wedge \neg G \tag{23}$$

$$\text{opening, } O \;\Leftrightarrow\; ((\lambda_i = 1) \vee (r_{C_{i,\,i+1}} > r_s)) \wedge \neg H \tag{24}$$

$$\text{aperture, } A \;\Leftrightarrow\; (d_{E_iE_{i+1}} > w_r) \wedge \neg O \tag{25}$$

$$\text{concave corner, } c \;\Leftrightarrow\; ((0 < \Delta\alpha < \pi) \vee (\Delta\alpha < -\pi)) \wedge \neg A \tag{26}$$

$$\text{convex corner, } e \;\Leftrightarrow\; \neg c \tag{27}$$

$d_{E_iE_{i+1}}$ denotes the distance of two adjacent end points, condition $l_{S_i} = l_{S_{i+1}}$ ensures that both segments under consideration belong to the same line and $r_{C_{i,\,i+1}}$ is the distance of intersection point $C_{i,\,i+1}$ from the origin of the sensor coordinate system.

When the type of spatial configuration is known, a normalized vector $\underline{v}_i$ indicating its direction in the scene is determined. For opening candidates the midpoint between two adjacent endpoints is taken, for corners their arc tangent. A further abstraction is made by concatenation of the associated symbols from the set $\{A, c, e, G, H, O\}$. This yields a rotation variant string with variable length where rotating the robot causes a cyclic shift. Although it is a scene model on a high abstraction level, the string is not free from noise. Lines consisting in only small segments with a low hit rate in successive scans effect the length and structure of the string to change even when the vehicle is not moving. This diminishes the applicability of the string as a instrument for meaningful inference. See e.g. figure 5.5 for a situation which is mainly characterized by three walls constituting two concave corners. If the string sould serve as a reliable tool, e.g., for similarity testing of scenes in a relocation scenario, it has to reflect the particularity of the local observation in a robust manner.

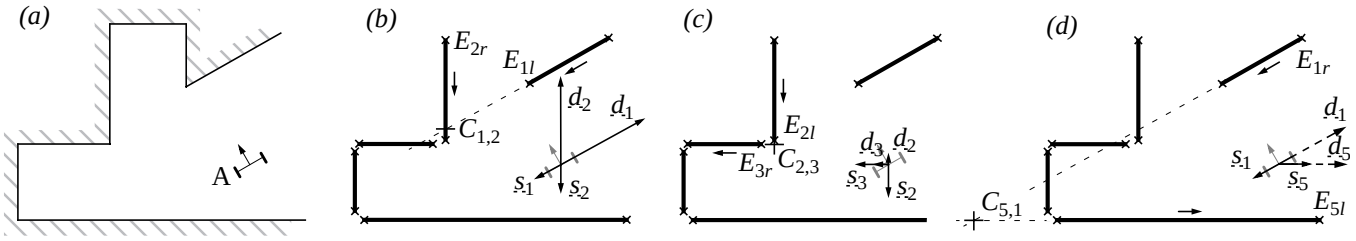

Figure 6.1 *A simple two-dead end situation for illustration of the concept of directionality. The robot is in position* A*, scanning counterclockwise and extracts the displayed segments whose scan directions are indicated with arrows. The directionality* λ *of an end point is defined as its scan direction with respect to the intersection point. For segment* S_1 *and* S_2 *we find the arrows pointing towards* $C_{1,2}$ *thus* $\lambda_{E_{1l}} = \lambda_{E_{2r}} = -1$*, for* $C_{2,3}$ *we find* $\lambda_{E_{2l}} = -1$ *and* $\lambda_{E_{3r}} = 1$ *and for* S_5 *and* S_1*,* $\lambda_{E_{5l}} = 1$ *and* $\lambda_{E_{1r}} = -1$*. According to (23), the spatial configuration of* S_1 *and* S_2 *is classified as 'hidden corners' which seems a reasonable interpretation because two concave corners are not visible from position A. The resulting string for this situation is* $'HeccO'$*.*

Robustness is archieved by exploiting the uncertainty which is associated to all extracted segments. By determining a weight $w_i = f(C_{l_i}, C_{l_{i+1}})$, where C_{l_i}, $C_{l_{i+1}}$ are the nondiagonal covariance matrices of lines l_i and l_{i+1}, a confidence measure for each symbol is gained which gives most weight to symbols with high evidence. For opening candidates (gaps, hidden corners, openings and apertures) the weight is

$$f(C_{l_i}, C_{l_{i+1}}) = 1/tr(C_{l_i} + C_{l_{i+1}}) \quad . \tag{28}$$

For corners (concave and convex corners) the intersection point covariance matrix is used

$$f(C_{l_i}, C_{l_{i+1}}) = 1/tr(f_{\alpha r} \begin{bmatrix} C_{l_i} & 0 \\ 0 & C_{l_{i+1}} \end{bmatrix} f_{\alpha r}^T) \tag{29}$$

where $f_{\alpha r} = [x_{c_i}\ y_{c_i}]^T[\partial/\partial\alpha_i \ \ \partial/\partial r_i \ \ \partial/\partial\alpha_{i+1} \ \ \partial/\partial r_{i+1}]$ is the Jacobian and

$$\begin{bmatrix} x_{c_i} \\ y_{c_i} \end{bmatrix} = \frac{1}{\sin(\alpha_{i+1} - \alpha_i)} \begin{bmatrix} \sin\alpha_{i+1} & -\sin\alpha_i \\ -\cos\alpha_{i+1} & \cos\alpha_i \end{bmatrix} \begin{bmatrix} r_i \\ r_{i+1} \end{bmatrix} . \tag{30}$$

A revealing piece of information on the situation is the weighted vector sum and the inversely weighted vector sum. A normalized vector $\underline{v}_i$ is determined for each symbol, indicating its direction in the scene. Then

$$\underline{v}_h = \frac{1}{\sum w_i} \sum w_i \cdot \underline{v}_i , \tag{31}$$

$$\underline{v}_l = \frac{1}{\sum 1/w_i} \sum \frac{\underline{v}_i}{w_i} \tag{32}$$

are referred to as the *vector of highest confidence* and the *vector of least confidence*. $\underline{v}_h$ heads into the region of highest evidence which enables the robot to approach secure terrain when it is desirable. If position uncertainties are low and the robot can afford to explore uncertain terrain it chooses the opening candidate which is closest to $\underline{v}_l$. See figures of section 5 for example situations.

The string itself does not contain more information than we already have. But it uncovers a different perspective on the same problem which might result in a fruitful change of the way it is approached. In the artificial intelligence domain, a number of concepts have been developed for dealing with symbolic representations. In information theory and genetics various methods exist for determining a similarity measure of two given strings of different length[13]. When an appropriate metric which keeps track of the weighting has been found, an explicit and concise similarity measure for matching and comparing local environments is available.

Equivalent to the method of symbolic constraint-propagation which has been used for understanding line drawings of plane-faced objects[14], a real physical entity, the environment, is under consideration. This physical object manifests constraints and regularities which can be exploited when attempting to derive its understanding. Hence the strings, describing sections of the robot's environment, adhere an underlying set of restrictions which limits the amount of possible symbol sequences. E.g. *'OOO'* is impossible since there is no configuration of three walls which result in that string. A number of conflicting combination can be likewise determined in order to enable consistency testing of the actual environment model. As already stated, two types of spatial configurations are to be distinguished. Opening candidates stand for a *lack of knowledge* and corners are considered as *perfect knowledge* requiring no further investigation by the robot. According to this conception the alphabet of terminal symbols $\{c, e\}$ and the auxiliary alphabet of nonterminal symbols $\{A, G, H, O\}$ can be distinguished. This is a view of exploration as a syntactic approach which relies on the definition of a string grammar. Exploration of unknown line environments is then the task of applying a set of production rules from a so called recursively enumerable grammar in order to generate strings which consist only of terminal symbols $\{c, e\}$. If all nonterminal symbols have been substituted, i.e. all opening candidates have been explored and rewritten, the environment is perfectly known. See figure 6.1a where complete knowledge about this part of the environment is gained by maneuvering the robot towards the hidden corners which corresponds to the rule $H \rightarrow cc$.

7. SUMMARY AND OUTLOOK

A scheme for extracting environment features from 1D range data and their interpretation is presented. Segmentation of the range image is done by deciding on a measure of model fidelity which is applied to adjacent groups of measurements. Due to outliers, obstacles which partially occlude a landmark or discontinuities in the landmark's structure or surface, the extraction process is considered to include a subsequent matching step where segments which belong to the same landmark are to be merged while keeping track of those which originate from distinct features. This is done by an agglomerative hierarchical clustering algorithm with a Mahalanobis distance matrix. The method is discussed with straight line segments which are found in a generalized least squares sense using polar coordinates including their first-order covariance estimates. As a consequence, extraction is no longer a real time problem on the level of single range readings, but must be treated on the level of whole scans. Experimental results with three commercially available laser scanners with varying specifications are presented and discussed. The implementation of a reduced version on a mobile robot which performs a map-based localization demonstrate the accuracy and applicability of the method under real time conditions and moderate on-board computing power.

The collection of line segments and associated covariance matrices obtained from the extraction process contains more information about the scene than is required for map-based localization. In the subsequent reasoning step this information was made explicit. The directionality of an endpoint was introduced as a classification feature which provides the basis for the classification of spatial configurations of segments. By successive abstraction and consequent propagation of uncertainties, a compact scene model is finally obtained in the form of a weighted symbolic description. It preserves topology information and reflects the specific characteristics of a local observation which is expected to be robust against noise on the feature level. Furthermore, the vectors of highest and least confidence could be determined which concisely indicate regions of high and low feature evidence in the local observation.

We are confident that the presented method has a fast real time implementation on our new mobile robot *Pygmalion* equipped with a PowerPC@100MHz. The preprocessing step necessary for the reduced version can be discarded due to the increased robustness of the method. Based on measurements with the real time operating system XOberon which is constantly used in our experiments, an acceleration factor of at least hundred, opposed to the CPU which was used in the experimental setup, is expected for pure floating point operations. This remains, however, to be done.

It is further to be investigated how the directions of highest and least confidence can improve navigation in known and support exploration in unknown environment and whether a similarity measure for strings can be used as a reliable tool to determine the uniqueness of a local observation. This would also provide the necessary basis for a topology-based navigation scheme[15] where no or only rough metric information is incorporated.

REFERENCES

1. Ayache N., Faugeras O.D., "Maintaining Representation of the Environment of a Mobile Robot", in *Autonomous Robot Vehicles*, Cox I.J., Wilfong G.T. (eds.), p. 205, Springer, 1990.
2. Taylor R.M., Probert P.J., "Range Finding and Feature Extraction by Segmentation of Images for Mobile Robot Navigation", *Proceedings of the IEEE International Conference on Robotics and Automation*, Minneapolis, 1996.

3. Gander W., Hrebícek J., *Solving Problems in Scientific Computing Using Maple and MATLAB*, p. 71, Springer-Verlag 1993.
4. Hough P.V.C., *Method and Means for recognising complex patterns*, US Patent 3069654, 1962.
5. Vestli S.J., *Fast, accurate and robust estimation of mobile robot position and orientation*, Doctoral Thesis Nr. 11360, ETH Zürich, 1995.
6. Adams M.D., *Optical Range Data Analysis for Stable Target Pursuit in Mobile Robotics*, doctoral thesis, University of Oxford, Department of Engineering Science, UK, 1992.
7. Bar-Shalom Y., Fortmann T.E., *Tracking and Data Association*, Mathematics in Science and Engineering, Vol. 179, Academic Press Inc., 1988.
8. Leonard J.J., Durrant-Whyte H.F., *Directed Sonar Sensing for Mobile Robot Navigation*, Kluwer Academic Publishers, 1992.
9. Smith R., Self M., Cheeseman P., "Estimating Uncertain Spatial Relationships in Robotics", in *Autonomous Robot Vehicles*, Cox I.J., Wilfong G.T. (eds.), p. 167, Springer, 1990.
10. Hebert P., Betge-Brezetz S., Chatila R., "Probabilistic Map Learning: Necessity and Difficulties, in *Reasoning with Uncertainty in Robotics*, Dorst L., van Lambalgen M., Voorbraak F. (eds.), Lecture Notes in Artificial Intelligence, Vol. 1093, Springer, 1995.
11. Hummerl R., "Uncertainty Reasoning in Object Recognition by Image Processing", in *Reasoning with Uncertainty in Robotics*, Dorst L., van Lambalgen M., Voorbraak F. (eds.), Lecture Notes in Artificial Intelligence, Vol. 1093, Springer, 1995.
12. Kaufman L., Rousseeuw P.J., *Finding Groups in Data: An Introduction to Cluster Analysis*, Wiley Series in Probability and Mathematical Statistics, 1990.
13. Aoe J.-I. (ed.), *Computer Algorithms: String Pattern Matching Strategies*, IEEE Computer Society Press, Los Alamitos, CA, 1994.
14. Waltz D., "Understanding Line Drawings of Scenes with Shadows", in *Psychology of Computer Vision*, Winston P. H. (ed.), MIT Press, Cambridge, MA, 1972.
15. Kuipers B.J., Byun Y.T., *"A Robust, Qualitative Approach to a Spatial Learning Mobile Robot"*, Proceedings of the SPIE, Sensor Fusion: Spatial Reasoning and Scene Interpretation, Vol. 1003, 1988.