Mobile Robot Localisation on Reconstructed 3D Models

João Gomes-Mota, Maria Isabel Ribeiro

Instituto Superior Técnico/Instituto de Sistemas e Robótica

Av. Rovisco Pais, 1049-001 LISBOA, PORTUGAL, Fax: +351.1.8418291

jsgm@isr.ist.utl.pt, mir@isr.ist.utl.pt

Keywords: mobile robot localisation, laser scanner, 3D scene reconstruction, frame localisation

Abstract

This paper presents a two-layer mobile robot localisation solution based on a Laser Range Scanner applied on reconstructed 3D models. This solution encompasses a feature matching layer performing without an initial posture estimate; a second layer based on coordinate transforms to refine an initial posture estimate, and a third module to assess the likelihood of the results computed by the two layers.

1 Introduction

This paper presents a novel strategy to perform mobile robot localisation using a Laser Range Scanner and reconstructed 3D models of real indoor environments. It is based on a two-layer and three module solution, each module processing the data differently. This approach overcomes most problems of individual algorithms, increases robustness while keeping the development and analysis fairly simple, due to its modular nature. In Fig. 1, the first layer, Frame Localisation (FrameLoc), uses the 3D reconstructed map and a 2D laser scan only, to compute a coarse estimate, which is verified by a Likelihood Test (LT). The valid results are fed into the second layer, the Approximate Localisation (ApproxLoc), along with the same map and range scan, to produce a *refined estimate*, which is also verified by the Likelihood Test. In case an initial posture is available from other sources (*e.g.*, odometry), the first layer is bypassed and the second layer is initialised with the external posture estimate.

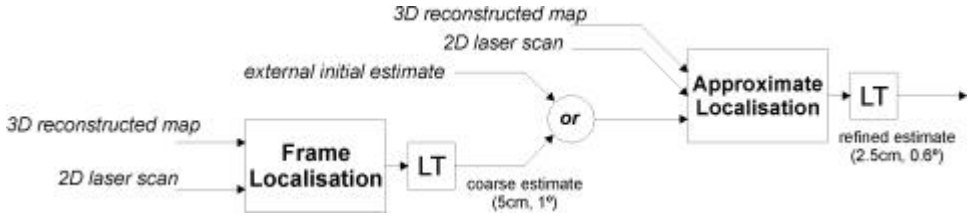


Fig. 1 - Two layers-three modules strategy

The paper is organised as follows: Section 2 describes the RESOLV project, the framework in which this research was performed. Section 3 summarises some localisation algorithms using laser range scanners developed in recent years and states the novelty of the proposed approach. Section 4 presents Frame Localisation, to be used in case no a priori posture estimate is available. Section 5 presents the Likelihood Test, which is an error estimate rule. Section 6 presents Approximate

Localisation, to be used if an initial posture estimate is available. Sections 4 to 6 are illustrated with data of the main stages of a real experiment. Section 7 analyses the interactions between the three modules. Section 8 concludes the paper and comments on known issues and directions for further research.

2 The RESOLV project

The **RESOLV-RE**construction using **S**canned **L**aser and **V**ideo project's goal is to create a 3D reconstruction of a real indoor environment using laser range images and video images to be shown and operated through the Internet, [7]. The 3D models, such as the one shown in Fig. 2a, are represented in VRML language and may be explored with a regular Web browser. The main RESOLV sensor is shown in Fig. 2b: it consists of a video camera and a laser camera with a rotating mirror on the front, both mounted on top of a Pan & Tilt motorised unit. This compound sensor is connected to a computer where the 3D models are generated. The computer also runs the Web server where the models are hosted. The system is fully autonomous since all computations are done on-board.

To reconstruct a complex scene, multiple acquisition points are required to resolve occlusions and maintain a good level of detail. Therefore, the 3D reconstruction system must be moved between consecutive acquisition points. A mobile robot (Fig. 2c) is the perfect candidate for the task; this robot, termed **AEST - A**utonomous **E**nvironment **S**ensor for **T**elepresence, is based on a mobile platform featuring a ring of ultra-sound sensors for obstacle detection and odometers for in-motion localisation. For localisation purposes, the Laser Scanner is the sensor to use whereas odometry is supplemental since the RESOLV system may also be mounted on a manual trolley. The input data required for localisation is a 2D horizontal laser scan and the 3D reconstructed maps, containing the 3D surface description. The nature of the task increases the difficulties for localisation: the map is unknown beforehand; it is created as the system moves between acquisition points; therefore the environment around the system may have large unmapped or occluded areas.

(a) (b) (c)

Fig. 2 - (a) sample reconstructed 3D environment, [3] (b) sensing device, (c) RESOLV autonomous system

The localisation solution should provide a 2D position plus orientation posture estimate, (**x**,**y**,**q**), with less than 5 cm position error and 0.02rad (≈1º) heading error, within the RESOLV system constraints mentioned and described in [7] and [9] in detail. The proposed solution meets all the requirements.

3 Related research

In recent years, several systems based on laser scanners have been developed for indoor localisation, under different world models and applications. Some solutions are briefly described below and compared with the RESOLV solution.

Some methods use the laser range data associated with other sensors, usually odometry, to compute incremental updates on the robot posture. Such is the case presented in [6], where significant features, such as straight lines, are extracted in different moments and the displacement between them is used to compute the robot's movement. This solution, although very robust, is insufficient for RESOLV since it requires an initial posture estimate.

A class of solutions assumes the map is known beforehand; for instance, in [5] a subset of points in the map are associated to a numerical pattern ("Principal Components of Range Data") extracted from the range profile at each of those points. A space is created upon these data and the posture estimate from a particular range scan is mapped into the space. A Kalman-Bucy filter is used to extract the most likely solution. This solution is inadequate for RESOLV because it requires a closed and fully described environment and also a set-up phase.

In [10] a method closer to RESOLV's is proposed. The laser data is transformed into a set of lines or circle arcs, which are combined to resemble natural landmarks, such as corners. If the robot displacement is approximately known, a feature extracted from the range data should be close to the corresponding feature on the map. Once the two features are matched the robot posture is known. In RESOLV's solution the range data is also transformed into a set of lines, albeit in a different manner, but since the map is previously unknown, the system must find its own "landmarks" and decide how to associate extracted landmarks to landmarks on the map. A method inspired by the previous one is described in [1]. Straight lines, extracted from the range data, are combined to form natural landmarks. Laser intensity is also measured to enhance the statistical representation of data and to complement or replace natural landmarks with artificial landmarks such as reflective adhesive bands. In RESOLV's solution the laser intensity is also used to enhance range data analysis but artificial landmarks are unavailable.

Briefly, RESOLV's solution differences are: it requires no initial posture estimate, thus odometry is a supplement, whereas it is necessary for other methods; it includes two complementary algorithms, enhancing its overall performance and minimising each algorithm's pitfalls, and, last but not the least, performs correctly even though only a small fraction of the surrounding scene is known.

In [11] a solution for 3D mapping is presented. This work is fairly related to RESOLV's goals, although the paradigm is different: it uses a 3 camera video system as an acquisition tool. Surface boundaries are its distinctive features and range data is translated into a similar description. To perform localisation the lines extracted from the range scan are compared to the lines in the world model and a Mahalanobis distance criterion is used to find a match. This solution complies with the constraints of the RESOLV system, except for the acquisition sensors. However, it should be interesting to try adapting it to RESOLV.

4 Frame Localisation Algorithm (FrameLoc)

The FrameLoc algorithm requires two inputs: a 2D horizontal laser scan to perceive the surroundings and a map, which results from the RESOLV 3D reconstructed maps [9]. The localisation refers the system location to the 3D reconstructed model since there is no external map.

The algorithm flow is represented in Fig. 3: first, the data from the 3D reconstructed map and the 2D laser scan acquired for localisation are expressed in a common form, denoted as the *frame*. The frames from each source are then compared. For each match, a candidate posture estimate is computed. Finally, the set of solutions is analysed and clustered to yield a reduced set of plausible solutions.

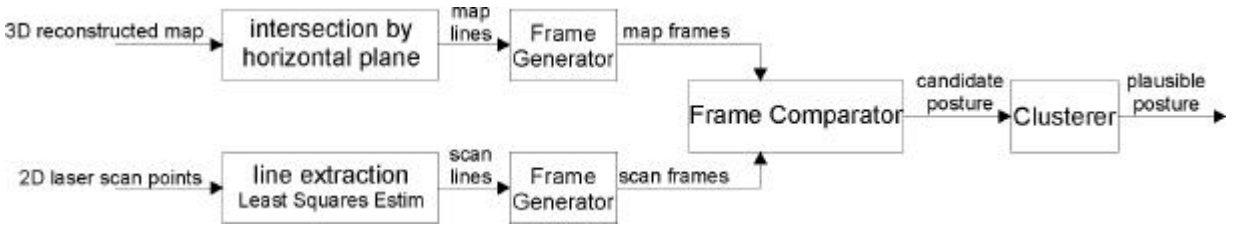


Fig. 3 - FrameLoc operation flow

The first step is to transform the input data into line segments: the 3D reconstructed map (Fig. 4a) is intersected by a horizontal plane at the height of the laser scan, resulting in a 2D map expressed as a set of segments (Fig. 4b); likewise, the 2D range scan points (Fig. 5a) are passed through a line extraction filter to obtain a set of line segments (Fig. 5b). Although the current implementation uses line segments only, the method is readily extendible to biquadratic curves.

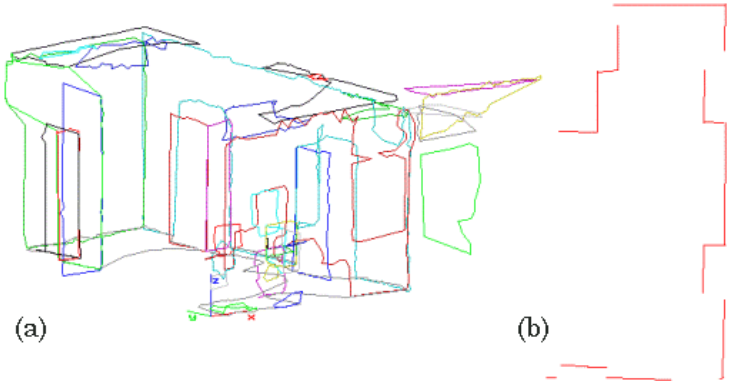


Fig. 4 - (a) 3D reconstructed map,
(b) extracted lines from the 2D map

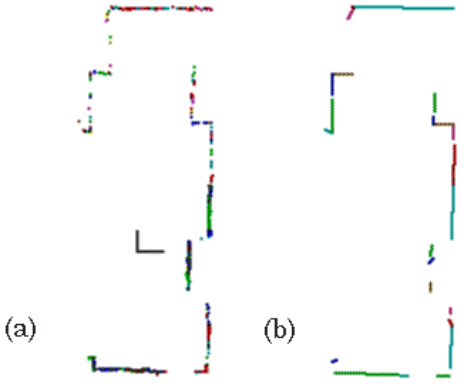
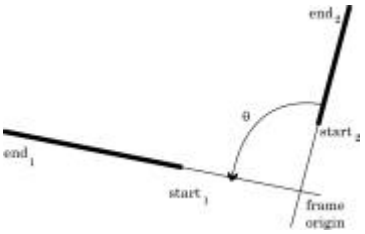


Fig. 5 - (a) 2D laser scan range data,
(b) extracted lines from 2D laser scan

The second step of the FrameLoc is to associate the segments on each set in a new geometric entity, termed *frame*. A frame is defined by two non-parallel segments, called *axes*, and the point where the lines defined by the two segments intersect, called *origin* (Fig. 6a). A frame has two types of parameters: five local parameters (Fig. 6a), which are independent of the coordinates their axes are expressed in, and three global parameters relating the frame origin to the axes' coordinate system (Fig. 6b).

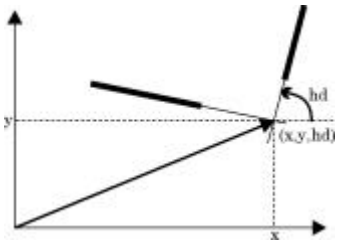


(a)

start$_1$, start$_2$: from the origin to the beginning of axis 1 and 2,

end$_1$, end$_2$: from the origin to the end of axis 1 and 2,

θ: frame inner angle



(b)

x, y, hd: frame origin x,y,heading-coordinates expressed in its reference: the world for the map frames and the AEST for the scan frames

Fig. 6 - (a) Frame Local Parameters, (b) Frame Global Parameters

The local parameters are: the distance from beginning of axis and end of axis to the origin, $\mathbf{start_i}$ and $\mathbf{end_i}$ for both axes (i=1,2) and the axes' inner angle $\mathbf{q}$. The global parameters enclose the information relating the frame to its coordinate system. Given the position and heading of the frame origin (**x**, **y**, **hd**) and the local parameters, it is possible to restore the axes to their original posture.

To generate the frame sets, each set of line segments is sorted by decreasing size, because longer segments are more relevant to the environment description and their statistics are more accurate. Also, the set of candidate segments is reduced to the cases where the lines intersect with $|\sin(\mathbf{q})|>0.5$, ignoring the segment pairs which are close to parallel. The frame generation process is repeated for a sufficient number of map and scan segments, yielding a map frame set and a scan frame set.

In normal operation, if a surface is visible for localisation it has been visible before to the 3D reconstruction process. Therefore, to verify if a scan frame corresponds to a given map frame, it must fit within the map frame (Fig. 7). The frame match test, (1), requires the local parameters only,

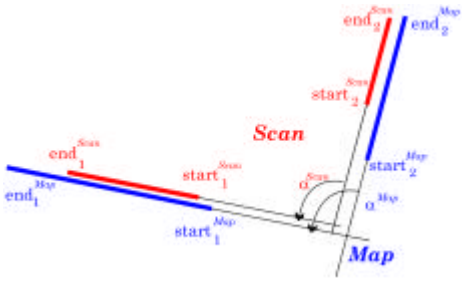


Fig. 7 - Frame match

$$\begin{cases} start_1^{Map} \le start_1^{Scan} \le end_1^{Scan} \le end_1^{Map} \\ start_2^{Map} \le start_2^{Scan} \le end_2^{Scan} \le end_2^{Map} \\ \left|\boldsymbol{a}^{Scan} - \boldsymbol{a}^{Map}\right| < \boldsymbol{d} \end{cases} \quad (1)$$

where δ is a heuristic limit to the angle difference. The angle δ depends on surface characteristics which are difficult to model and on the laser, to a lesser extent. Nevertheless, a conservative value ($\delta<0.05$rad ≈ 3º) accommodates almost every case, and has been used on the experiments.

If the frame match, (1), is valid, the canonical coordinate transform equation (2) applied to the frames' global parameters, yields a candidate solution (x_R, y_R, q_R), expressing the system posture in the 3D reconstructed model reference.

$$\boldsymbol{q}_R = hd_1^{Scan} - hd_1^{Map}$$

$$\begin{bmatrix} x^{Map} \\ y^{Map} \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\boldsymbol{q}_R & -\sin\boldsymbol{q}_R & x_R \\ \sin\boldsymbol{q}_R & \cos\boldsymbol{q}_R & y_R \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x^{Scan} \\ y^{Scan} \\ 1 \end{bmatrix} \quad (2)$$

Fig. 8 - Some possible postures

The FrameLoc is intrinsically exponential. Therefore, it is vital to keep the number of instances and the computation effort as low as possible. To attain these goals, all sets are sorted and the algorithm is developed in an iterative manner. The frame generator uses only the segments longer than a given threshold and the frame sets are also sorted by decreasing size of the longer axis. The frame matching is applied iteratively to the elements in the frame lists, until a coherent pattern of postures becomes apparent in the **(x, y, q)** solution space. In the process, *false hits* due to environment symmetry (Fig. 8) spread sparsely, while the *true hits* concentrate in close neighbourhoods (A and B in Fig. 8).

To compute the most likely solutions a weighted clustering procedure is used, where the weights are the product of the scan frames axes' lengths, since long, balanced axes are supported on well defined surfaces in the environment. The clustering process aggregates all solutions in a single neighbourhood into a single weighted posture (with weight equal to the sum of the weights of all cluster elements), reducing the frame match set to a few plausible candidates with most of the weight and many clusters with residual weight, which are clearly false hits due to environment symmetry.

Otherwise, the FrameLoc algorithm is extended, gathering more line segments from the original sets, building up the associated frames and matching the new instances with both the previous and the new ones. The new match candidates are added to the solution space without losing the existing data.

The FrameLoc final result is a short list of postures, sorted by decreasing weight. Most often, the first element is the correct posture and takes more than 50% of the total weight. The clustering procedure is very stringent, therefore clusters A and B in Fig. 8, although only 8 cm apart, are not merged. This option favours homogeneous clusters with small variance, at the expense of a longer list of postures and diluted weights. Even when FrameLoc best solution is less clear, the Likelihood Test, described next, will assess very accurately the quality of each candidate posture.

Except for very odd or symmetric environments or insufficient data, the algorithm converges swiftly to a reduced set of candidate clusters. For instance, the current example was based on 107 scan segments and 38 map segments and 16 frames were sufficient to compute the correct solution (A in Fig. 8).

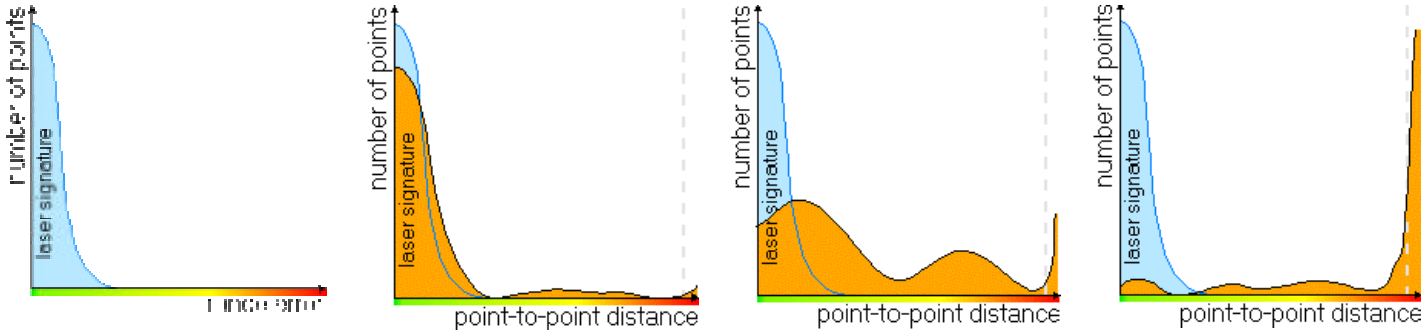
5 Likelihood Test

The solutions computed by FrameLoc and ApproxLoc (described in Section 6) are based on subsets of the laser range scan and the 2D map data. To assess the quality of each candidate posture, a likelihood criterion is defined: measure the difference between the whole laser range scan (Fig. 5a) and the 2D map (Fig. 4b) extracted from the reconstructed 3D model.

The first step is to express both profiles in a common form. For each candidate posture (x_R, y_R, q_R), a *simulated laser scan* profile is evaluated using the 2D map assuming the RESOLV system is located at (x_R, y_R, q_R). This simulated scan technique will be used also for ApproxLoc.

The simulated laser scan replicates the main characteristics of the real laser scan, i.e., it is centred on the RESOLV system and taken at the sensor head distance from the ground, it has the same angular width and number of points. An histogram on the distance error (the modulus of the range

error) is restricted to non-negative values and should resemble the normal curve with twice the amplitude. This histogram may be regarded as the laser scanner "signature" (Fig. 9a).



(a) (b) (c) (d)

Fig. 9 - (a) The laser signature and likelihood histograms: (b) correct solution, (c) to be refined, (d) wrong solution

Once the real laser data and the 2D map data are expressed as "scans" their point-to-point difference is computed. This measure establishes a likelihood criterion relating the laser scan to the map. If the posture estimate (x_R, y_R, q_R) were exact, the difference between the real scan and the simulated scan would be due only to the laser errors, the errors in the simulated scan model and the errors on the 3D reconstructed map. Since the first error source is much stronger than the others, drawing a histogram on the point-to-point distance should reveal a pattern close to the laser signature.

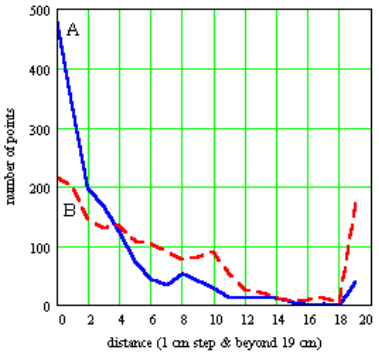
For the sake of simplicity, the simulated laser scan doesn't consider the detailed sensor characteristics, namely beam divergence, sizeable footprint, angle of incidence, surface reflectance and transitions since these approximations have a minor influence in the range error pattern. Although these inaccuracies degrade the absolute likelihood measure, the relative values remain unchanged, since they affect all difference measures likewise. Different laser cameras are used in the RESOLV project, and their errors follow a normal distribution with zero mean and 1 to 3 cm standard deviation [8].

The point to point distance measurement must account for invalid points that may exist in both "scans". The points in the simulated scan are invalid when it includes void areas in the reconstructed map. The real scan profile has some invalid points too, which are mainly due to specular reflections, device errors, time-outs and poor reflectance. Once the errors are detected the histogram is computed: for each pair of valid points the Cartesian distance is computed and stored in 1 cm slots, up to a limit. Above that, all points are grouped in one common slot. Fig. 9b to Fig. 9d show typical distance histograms, superimposed on the laser signature. If the histogram is close to the laser signature, the likelihood measure is high: the solution is correct (Fig. 9b). If the likelihood measure falls below a threshold the test fails (Fig. 9d) (*e.g.*, less than 50% pairs below 10 cm for FrameLoc and less than 90% pairs below 5 cm for ApproxLoc). Intermediate cases, such as (Fig. 9c), are subject to refinement by ApproxLoc.

A real example in Fig. 10 illustrates the characteristics of the Likelihood pattern. This experiment was made in an 12x6 meter laboratory room, with a 280º range scan sampled at 1800 points. The two solutions presented in Fig. 8 correspond to the clusters A=(2.296m, 2.381m, 0.022rad (1.24º)) and B=(2.217m, 2.366m, 0.019rad (1.1º)) computed by FrameLoc. This case highlights the sensitivity of the

Likelihood Test, which is able to discriminate two postures only (8 cm, 0.002rad (0.14º)) apart and reacts to a small correction such as A to A' (1.5 cm, 1.7×10^{-4} rad (0.01º)).

The graphics in Fig. 10 show the distance distribution classified in 1 cm slots from 0 to 19 cm and a final class above 19 cm. In Fig. 10a the two lines represent the Likelihood of the solution computed by FrameLoc. Solution A is better, although B is already acceptable: 76.5% pairs are closer than 10 cm. The Approximate Localisation introduced in Section 6 (Fig. 10b), updates A and B to A'=(2.285m, 2.371m, 1.25º) and B'=(2.285m, 2.371m, 1.35º). Comparing the two images it is clear how a slight posture update "pushes" most of the instances in the histogram to the left, increasing the Likelihood measure.

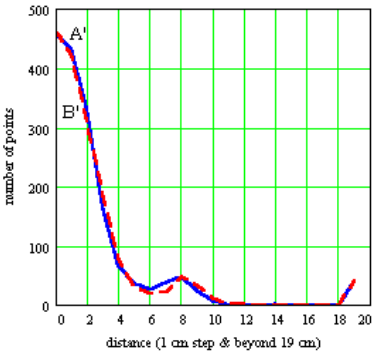


FrameLoc results

distance:	< 5 cm	< 10 cm
A:	84.5%	92.5%
B:	47.5%	76.5%

ApproxLoc results

distance:	< 5 cm	< 10 cm
A':	86.5%	96.9%
B':	87.3%	96.5%



(a) (b)

Fig. 10 - (a) Likelihood after FrameLoc, (b) Likelihood after ApproxLoc

Up to around 4 cm, i.e. twice the range error standard deviation, a good histogram, such as A, is close to the laser signature. A poorer solution, such as B, presents a reduced Likelihood measure which can be enhanced by ApproxLoc up to the signature curve (from B to B'). From 4 cm to 19 cm, the histogram on Fig. 10a accounts for the instances that may be fitted closer by the ApproxLoc. The last slots include the pairs farther from the sensor, and therefore very sensitive to angular errors. It also includes the residual pairs that do not match at all, because of sensor or map errors. The graph on Fig. 10b shows the results after ApproxLoc. It is apparent that A' and B' are very close to the Laser Scanner signature, except for the residual pairs above 19 cm and a minor peak at 8 cm, due to inaccuracies in the 3D reconstruction.

6 Approximate Localisation Algorithm (ApproxLoc)

The second layer of the proposed localisation methodology requires an initial posture estimate, supplied either by FrameLoc, the odometry data in the AEST platform as described in [4], or any other external source. The algorithm uses the same simulated scan technique as described Section 5: to compare the map and the range scan, a simulated scan is taken, assuming the platform is at the given initial posture estimate, (x_I, y_I, q_I). ApproxLoc follows an approximation approach: assuming the initial posture is close to the real posture, (x_w, y_w, q_w), the laser scan should be very close to the simulated laser scan "perceived" in the reconstructed map; thus, a sample point in the simulated scan $(x,y)_i^{SimSc}$, should have its counterpart in the range scan, $(x,y)_j^{AEST}$, in the vicinity of position **i**. In particular, if one natural landmark is identified in the simulated scan, its range scan counterpart

should lie in the vicinity of the same radial direction. In this implementation the corners are chosen as landmarks.

In Fig. 11 the references involved are shown. The simulated scan points refer to the known local reference with origin (x_I, y_I, θ_I); the laser scan points refer to the reference attached to the AEST platform located at the unknown posture (x_w, y_w, θ_w); the vector (x_p, y_p, θ_p) transforms the AEST reference to the SimSc reference.

Assume that the sample points, $(x,y)_i^{SimSc}$ and $(x,y)_j^{AEST}$, are the points closest to a surface transition. These natural landmarks are readily detected by sliding a window through the range profile: when a consistent change of direction occurs, the corner is detected. When $(x,y)_i^{SimSc}$ is detected in the simulated scan, where there is no noise, a similar feature is searched for in the real scan around position i, $i-k \le j \le i+k$, with a fixed value of k, *e.g.* k=5. In Fig. 12 the landmark selection procedure is illustrated. The laser scan and the simulated scan are superimposed, referred to the platform. The lines indicate the directions where the landmarks were detected.

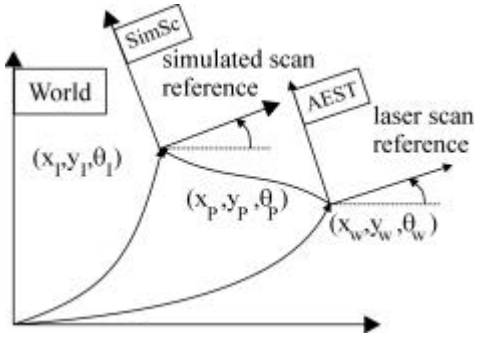


Fig. 11 - References and coordinate transforms

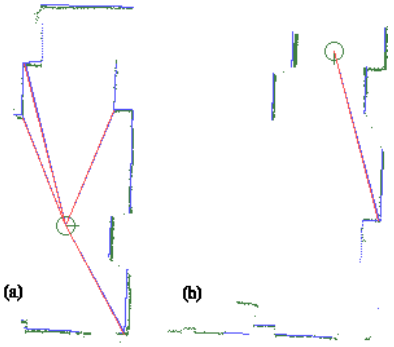


Fig. 12 - Natural landmark detection (a) success: four pairs are found (b) failure: one pair is found while two are required

To determine the three dimensional vector (x_p, y_p, θ_p) a minimum of two landmark pairs are required, each pair supplying two equations. The process is applied to a maximum of 5 to 6 landmarks, as experiments shown to be the balance point between computational burden and equation stability.

If the initial estimate is accurate enough and the 2D map extracted from the reconstructed 3D model includes enough landmarks within the scan width, a sufficient number of landmarks are detected (Fig. 12a). Otherwise, the search fails, such as in (Fig. 12b); in this particular case there are too few landmarks because the reconstructed model perceived from the top of the scene is very incomplete (Fig. 4b). In such cases, detecting more landmarks would require a relaxed confidence criterion. The update transform from the simulated range scan reference (SimSc) to the actual AEST posture can be computed using (2), rewritten as (3). Considering **n** pairs simultaneously, generates a non-linear over determined system, (4).

$$\begin{bmatrix} x \\ y \end{bmatrix}^{AEST} = \begin{bmatrix} \cos(\theta_P) & \sin(-\theta_P) & x_P \\ \sin(\theta_P) & \cos(\theta_P) & y_P \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix}^{SimSc} \qquad (3)$$

$$\begin{bmatrix} x_1 \\ y_1 \\ \vdots \\ x_n \\ y_n \end{bmatrix}^{AEST}_{(2n\times 1)} = \begin{bmatrix} \cos(\theta_P) & \sin(-\theta_P) & x_P \\ \sin(\theta_P) & \cos(\theta_P) & y_P \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ y_1 \\ \vdots \\ x_n \\ y_n \end{bmatrix}^{Sim\ Sc}_{(2n\times 1)} \qquad (4)$$

Assuming the angular correction q_P is small, it is possible to linearise the system around $q_P=0$ (5).

$$\begin{cases}\cos(q_P)\approx 1\\ \sin(q_P)\approx q\end{cases} \quad iff \quad q_P\approx 0 \qquad (5)$$

Replacing the terms in q_P as in (5) and rearranging (4) in order of the unknowns, $(x_P, y_P, \boldsymbol{q}_P)$ yields the over determined system (6a), where **M** is the coefficient matrix (6b) and **D** is the difference matrix (6c).

$$M\cdot\begin{bmatrix}x_P\\ y_P\\ q_P\end{bmatrix}=\Delta \qquad (6a)$$

$$M=\begin{bmatrix}1 & 0 & -y_1^{SimSc}\\ 0 & 1 & -x_1^{SimSc}\\ \vdots & \vdots & \vdots\\ 1 & 0 & -x_n^{SimSc}\\ 0 & 1 & -y_n^{SimSc}\end{bmatrix}_{(2n\times 3)} \qquad (6b)$$

$$\Delta=\begin{bmatrix}x_1^{AEST}-x_1^{SimSc}\\ y_1^{AEST}-y_1^{SimSc}\\ \vdots\\ x_n^{AEST}-x_n^{SimSc}\\ y_n^{AEST}-y_n^{SimSc}\end{bmatrix}_{(2n\times 1)} \qquad (6c)$$

This over determined system may be solved in the form of (7), where the coefficients of $\mathbf{M^TD}$ and $\mathbf{M^TM}$ are defined in (8a) and (8b), respectively.

$$M^T\Delta=M^TM\cdot\begin{bmatrix}x_P\\ y_P\\ q_P\end{bmatrix} \qquad (7)$$

$$M^T\Delta=\begin{bmatrix}\sum(x_{AEST}-x_{SimSc})\\ \sum(y_{AEST}-y_{SimSc})\\ \sum(x_{SimSc}(y_{AEST}-y_{SimSc})-y_{SimSc}(x_{AEST}-x_{SimSc}))\end{bmatrix}_{(3\times 1)} \qquad (8a)$$

$$M^TM=\begin{bmatrix}n & 0 & -\sum y_{SimSc}\\ 0 & n & \sum y_{SimSc}\\ -\sum y_{SimSc} & \sum x_{SimSc} & \sum(x_{SimSc}^2+y_{SimSc}^2)\end{bmatrix}_{(3\times 3)} \qquad (8b)$$

In (8a) and (8b) all sums are from 1 to the number of points, **n**. It is now clear that (7) represents a (3x3) linear system which can be solved by standard Gauss elimination. The result, $(x_P, y_P, \boldsymbol{q}_P)$, is the transform parameters relating the actual AEST posture where the range scan was taken, to the initial posture estimate, $(x_I, y_I, \boldsymbol{q}_I)$, where the simulated laser scan was taken.

The AEST posture, $(x_w, y_w, \boldsymbol{q}_w)$, is updated according to (9),

$$\begin{aligned}&\boldsymbol{q}_W=\boldsymbol{q}_I+\boldsymbol{q}_P\\ &\begin{bmatrix}x\\ y\\ 1\end{bmatrix}_W=\begin{bmatrix}\cos\boldsymbol{q}_P & -\sin\boldsymbol{q}_P & x_P\\ \sin\boldsymbol{q}_P & \cos\boldsymbol{q}_P & y_P\\ 0 & 0 & 1\end{bmatrix}\cdot\begin{bmatrix}x_I\\ y_I\\ 1\end{bmatrix}\end{aligned} \qquad (9)$$

If the correspondence between landmarks in the simulated and range scans was perfect, one iteration of ApproxLoc would suffice to find the optimal estimate. However, this is seldom the case since it is very unlikely to have a scan sample exactly on the corner. Therefore, the algorithm operates in closed loop, using the estimated posture of one iteration to start the next one. The Likelihood Test tracks the convergence of the algorithm and the final solution usually occurs within three to five iterations. During the loop, the laser range error induces a uniform error pattern, which can not be

resolved by further iterations. Thus, the boundary for precision is close to the laser range error, which lies between 1 and 3 cm, for the RESOLV equipment. Hence, the typical output accuracy is 2.5 cm, 0.01rad (≈0.6º).

7 Operation Flow

The Localisation operation flow depends on the presence of odometry or other external initial data. If an initial posture estimate is available, then ApproxLoc is called. Otherwise, FrameLoc must be called to compute a first estimate for the AEST posture which is fed into ApproxLoc.

The simple operation sequence shown in Fig. 1 solves the vast majority of cases. However, some difficulties remain especially while the map is very sparse or if the scan resolution or accuracy is diminished. Moreover, calling FrameLoc is seldom necessary when an external initial estimate, usually odometry, is available. Thus, a loop structure (Fig. 13) was created for accuracy and robustness sake.

The start point of the Localisation cycle depends on the presence of odometry. If there is no odometry data the process starts at Frame Localisation, otherwise the posture estimate read from the odometers is used to compute the simulated scan, which is entered in the ApproxLoc, together with the real range scan. The resulting posture estimate is sent to the Likelihood Test to check if the two scans match. In case of success the algorithm ends, otherwise the initial posture was inaccurate and FrameLoc is called.

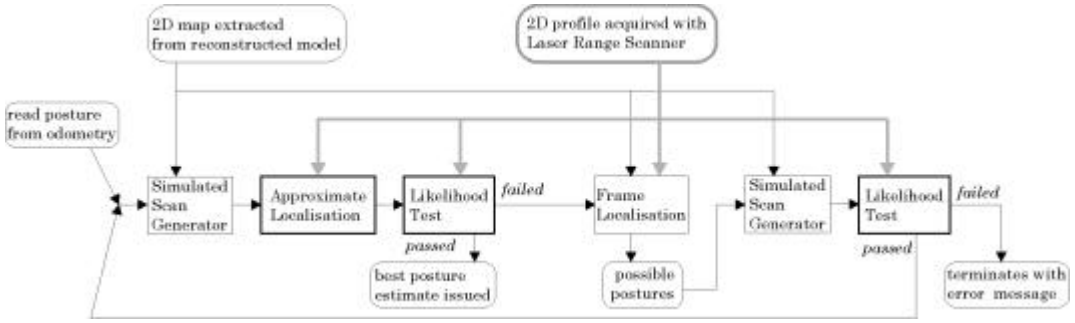


Fig. 13 - Compound algorithm operation flow

FrameLoc operation is based on the environment 3D model and the 2D laser range scan only, thus, it is irrelevant whether ApproxLoc has been called before. The results from FrameLoc are a list of possible postures which are passed, one by one, to the Likelihood Test. The approved postures are sent to ApproxLoc, while the others are discarded.

To avoid an endless loop after ApproxLoc, a flag indicates if FrameLoc has already been called; in that case the algorithm ends with failure. In future extensions a more sophisticated control loop would be inserted at this stage.

It is interesting to notice that a single range scan is used for all the operation loop. However, if the algorithm fails, it is possible to repeat it taking a scan at a different height. This is a likely extension to the compound algorithm and will be discussed in the next section.

8 Conclusions, known issues and extensions

This paper presents a novel approach for indoor localisation using a Laser Range Scanner. It requires only a partial map and no initial localisation estimate. Because one single algorithm would fail to operate properly in some conditions, two complementary algorithms were developed. The presented solution proved adequate for 3D model reconstruction, because it ignores large unknown areas, keeping its "anchor" on the available reconstructed 3D model. Accuracy requirements were exceeded and (2.5 cm, 0.01rad (0.6º)) maximum errors are expectable in standard conditions.

The algorithm uses computer resources with modesty. There are no large arrays, only comprehensive object lists. Since all repetitive tasks include tests and processing of linear equations and systems of reduced dimension, the algorithms are swift. The only computation intensive task is the line extraction algorithm (not described) even though it was developed in an efficient iterative form.

However, some pitfalls remain. One major cause of concern is the close nature of the localisation sensor and the map acquisition sensor. If the Laser Scanner fails to detect a mirror and maps a second scene *behind* the mirror, it is quite likely that the localisation will repeat the error. Strictly speaking, this is not a localisation error since the model includes the *virtual* scene. Nevertheless, it is an undesirable feature. The line extraction algorithm includes some statistical analysis using laser reflectance data to overcome this problem, in particular for mirrored windows. Also, if some surfaces in the scene are too dark, smooth or very detailed, the 3D reconstruction program will fail to extract these surfaces, leading to void areas in the 3D model. However, if there is a sufficient number of well defined surfaces the algorithm will "anchor" to them for localisation.

A subtler problem arises on non-horizontal floors. The 2D map extracted from the 3D model is independent of the AEST location if and only if the plane intersection is parallel to the ground. So, unless the AEST is equipped with an inclinometer of any type, it is restricted to piece-wise horizontal floors.

Finally, problems arise in highly symmetric environments, such as square or round rooms or long corridors with evenly spaced doors, which occur more often. If the corridor length is greater than the Laser Scan operating range, the algorithm will find more than one equally likely solution separated by the distance which correspond to the corridor's repeating feature period.

The algorithm presented may be extended in several ways. One natural extension is the generalisation of the frame elements to biquadratic curves. On another direction, it is also desirable to include a richer control structure in the loop, such as automatically choosing the scan height, repeating localisation at different heights and matching the results.

The major extension to the algorithm is to compute the localisation while the AEST is moving. The authors are convinced that this task will require a new approach, using a narrow laser scan, anchoring the localisation to a few carefully selected landmarks, in the lines suggested by [2].

Acknowledgements

This research work was supported by the project RESOLV of the ACTS Programme, EU.

References

[1] Arras, Kai and Sjur Vestli, *Hybrid, High-Precision Localisation for the Mail Distributing Mobile Robot System MOPS*, Proc. IEEE Int. Conference Robotics and Automation, pp. 3129-3134, Leuven, Belgium, May 1998.

[2] Arsénio, Artur and M. Isabel Ribeiro, *Active Range Sensing for Mobile Robot Localisation*, Proc. IEEE/RSJ Int. Conference Intelligent Robotic Systems, IROS'98, Victoria, Canada, October 1998.

[3] Butterfield, Stuart, Kia Ng, David Hogg, *The RESOLV Texture-Mapping Module*, School of Computer Science, University of Leeds. Report 97.48, December 1997.

[4] Castro, José, Vítor Santos, M. Isabel Ribeiro, *A Multi-Loop Robust Navigation Architecture for Mobile Robots*, Proc. IEEE Int. Conference Robotics and Automation, pp. 970-975, Leuven, Belgium, May 1998.

[5] Crowley, James L, F. Wallner, B. Schiele, *Position Estimation Using Principal Components of Range Data*, Proc. IEEE Int. Conference Robotics and Automation, pp. 3121-3128, Leuven, Belgium, May 1998.

[6] Dubrawski, Artur and Barbara Siemiatkowska, *A Method for Tracking Pose of a Mobile Robot Equipped with a Scanning Laser Range Finder*, Proc. IEEE Int. Conference Robotics and Automation, pp. 2519-2523, Leuven, Belgium, May 1998.

[7] Leevers, David *et al*, *An Autonomous Sensor for 3D Reconstruction*, 3rd European Conference on Multimedia Applications, Services and Techniques, ECMAST 98, Berlin, Germany, May 98.

[8] Vítor Sequeira, *Active Range Sensing for Three Dimensional Environment Reconstruction, Ph.D. Thesis*, Instituto Superior Técnico, Technical University of Lisbon, Portugal, December 1996.

[9] Sequeira, Vitor, Kia Ng, Stuart Butterfield, Joao. G. M. Gonçalves, David. C. Hogg, *Three-Dimensional Textured Models of Indoor Scenes from Composite Range and Video Images*, In R.N. Ellson and J.H. Nurre, editors, Proc. SPIE, Three-Dimensional Image Capture and Applications, Vol. 3313, 1998.

[10] Vestli, Sjur, *Fast, accurate and robust estimation of mobile robot position and orientation, Ph.D Thesis*, Eidgenössische Technische Hochschule Zürich, Switzerland, December 1995.

[11] Weckesser, P., R. Dillman, *Modeling Unknown Environments with a Mobile Robot*, Proc. Symposium on Intelligent Robotic Systems SIRS'97, pp. 69-76, Stockholm, July 1997.

João Gomes-Mota, born 1970, graduated at Instituto Superior Técnico, Technical University of Lisbon, Portugal, where he's now concluding a M.Sc. degree.

His research interests focus on sensor statistics, robot navigation, 3D environment reconstruction and Internet applications for Robotics.

M. Isabel Ribeiro received the Licenciatura, M.Sc and Ph.D. degrees in Electrical Engineering from Instituto Superior Técnico, Technical University of Lisbon, Portugal in 1978, 1983 and 1988, respectively. From 1978 she has had teaching and R&D responsibilities at Instituto Superior Técnico where she is an Associate Professor since 1991. She is a member of the Institute for Systems and Robotics, being responsible for the Mobile Robotics Laboratory. Her research interests include navigation of mobile robots, cooperative robotics, flexible automated guided vehicles and 2D and 3D-environment reconstruction. Dr. Ribeiro is a senior member of IEEE and of Ordem dos Engenheiros, the Portuguese Association of Engineers.