Scan Matching in the Hough Domain

Andrea Censi*, Luca Iocchi* and Giorgio Grisetti*‡

* *Dipartimento di Informatica e Sistemistica "A. Ruberti",*
University of Rome "La Sapienza", Via Salaria 113, I-00198 Rome, Italy
‡ *University of Freiburg, Department of Computer Science,*
Georges-Köhler-Allee 79, D-79110 Freiburg, Germany
E-mail: {acensi,iocchi,grisetti}@dis.uniroma1.it

***Abstract*— Scan matching is used as a building block in many robotic applications, for localization and simultaneous localization and mapping (SLAM). Although many techniques have been proposed for scan matching in the past years, more efficient and effective scan matching procedures allow for improvements of such associated problems. In this paper we present a new scan matching method that, exploiting the properties of the Hough domain, allows for combining advantages of dense scan matching algorithms with feature-based ones.**

***Index Terms*— Scan matching, Hough Transform**

I. Introduction

The 2D scan matching problem can be expressed as: *given two sets of 2D data (i.e. a reference scan and a current scan), determine a 2D rigid motion (a translation T and a rotation R_ϕ) that makes the scan data overlapping the reference data.*

The typical input for a scan matching algorithm is composed by: i) a 2D range sensor (e.g a laser range finder) producing a set of 2D points representing the contour of the visible environment around the sensor; ii) a reference scan representing the environment in which the robot moves and that can vary depending on the application (e.g. a pre-built map, a previous scan, etc.).

Many techniques have been proposed for scan matching in the past years; however, a definitive solution does not exist because scan matching is used in a vast range of operative conditions.

Scan matching methods differ largely based on the availability of a guess of the solution with relatively tight bounds (e.g. pose tracking) or the complete lack of a guess (e.g. global localization).

Another distinction can be made with respect to the assumptions about the sensor data: presence of sensor noise and feature-richness of the environment. Feature-based scan matching is computationally more efficient, but suffers from two limitations: 1) features must exist in the environment, 2) feature extraction produces information loss. On the other hand, dense scan matching is more robust to noise but can be very costly.

In addition to this, not all methods provide a multi-modal solution. This is important in a SLAM setting because it provides a hook to perform explicit handling of ambiguities or uncertainty in the mapping process.

In this paper we present a new scan matching method based on the Hough Transform (HT), called Hough Scan Matching (HSM), having the following properties.

- Despite the use of the HT, it does not extract line features from the scan (as in previous approaches [1]), but it matches dense data that can be interpreted as "feature distributions". It does not rely on specific features, therefore it is robust to sensor noise. The Hough domain enhances the properties of linear segments, resulting in better performance when they are present.
- HSM is able to perform a multi-modal, global search in the reference space, thus it is suitable for processing requiring extensive search (e.g. global localization). However, it can be restricted to perform local searches reducing computational time, when global search is not necessary (e.g. position tracking).
- The method is sound, in the sense that if the scan matching problem has only one exact solution (i.e. there exists only one rigid transformation that allows for perfectly overlapping the two scans), then HSM, without any initial guess, will include that one in the solution set.

II. Scan Matching

We first list some scan matching approaches that need a guess of the solution to operate.

If the environment is rich in roto-translations-invariant features (corners, straight walls [2]) and these are preserved notwithstanding the sensor noise, then it is possible to do simple filtering on the scan, extract them and find a solution, sometimes in closed form and linear time with respect to the number of points in the scan [3]. This approach has been proved to be efficient and robust when sufficient features are present in the environment.

In unstructured environments with relatively low noise, it is possible to employ algorithms of the ICP family [4] that do not assume the existence of features. These employs a two-step process: first, a series of heuristic correspondences between points in the two scan is established. Then a roto-traslation that approximately satisfies the series of constraints is found. The solution is obtained by iteratively executing the two steps until the error drops below a given threshold. In order to achieve convergence, such an approach requires the two scans to be taken at close enough positions. Modifications of the original algorithm allow for efficient implementations.

If data noise is higher, there exists dense methods that search in the solution space and do not require the establishment of any feature-to-feature or point-to-point correspondence. In [5] the solution is searched by performing gradient descent on a score function. Such a function is built by convolving the reference scan with a Gaussian kernel and then correlations it with the sensor scan. This approach has been subsequently refined in [6] by defining a closed form for a potential function that can be minimized using the Newton algorithm. These approaches require a number of iterations that depends on the input configuration and the entity of the error.

In theory any "local" matcher can be (inefficiently) used for a global search. As for the iterative approaches, one can use as initial guesses random points that represent a fine sampling of the solutions space - finer than the realignment margins of the local scan matcher. Analogously, with features based matcher, a combinatorial approach to feature correspondences can be employed. The following methods are designed for global search.

In [4] a technique for dealing with arbitrary orientation errors has been proposed. This is done by computing an histogram of the local best fitting tangent lines in the two scans to align. Assuming a small translational error, the rotational invariance of such a normal histogram allows to determine the heading component of the two scans. The translational pose component is then computed by minimizing the distance among correspondent scan points. In [7] a correlation based approach that is well suited in polygonal environments is presented. The orientation of each scan point is found and the circular histogram of these is build. Then the rotation is found by correlating the normal histograms. This can be done since the relative angles of a set of lines does not change when changing the reference system. Once the heading is known, the translational component of the displacement is recovered by determining pair of non parallel lines in the reference scan, and computing the translational error among the normal directions of those lines. Both these methods need that point orientation be extracted from data therefore are not reliable to sensor noise.

In [8] global localization is performed by using a two-dimensional correlation operator. This method evaluates every point in the search space and is therefore very robust to noise. Nevertheless, it can be implemented efficiently by the use of processor-specific SIMD (single instructions, multiple data) extensions.

HSM is a global, multi-modal, non-iterative matcher that can work in unstructured environments.

III. Overview of HSM

We will define a "spectrum" function, computed from the HT, which has these properties: 1) it is invariant to input translations, 2) it is (circularly) shifted on input rotation. We are therefore able to get several estimates of the rotation by comparing the spectra of sensor and reference data - this is done via a cross-correlation operator. Compared to the angle-histogram method this first part of the algorithm does not need to know the surface direction therefore it is very robust to sensor noise.

At this point we have a list of hypotheses for orientation. Other existing algorithms could be used from here, to be run in a reduced search space to find T. We propose one that exploits the already computed HT and that can produce a list of solutions or a continuous distribution.

HSM will provide a ranking of the solutions that in most cases seems quite right for localization; however in the general case another likelihood should be used (e.g. log likelihood).

IV. The Hough Transform

The Hough Transform (HT) has been used in the computer vision community since the sixties (proposed by Hough and generalized in works such as [9]) as a method for detecting geometric curves (lines, circles) in digital pictures. Although originally used for bitmap images (2D), the HT can be generalized to n-dimensional continuous spaces. In this section we will introduce a generalization of the HT and briefly describe some of its properties.

The HT maps an input $i(\mathbf{s}),\ \mathbf{s} \in \mathcal{S}$ (the *input space*) to a function $\mathrm{HT}\{i\}(\mathbf{p}),\ \mathbf{p} \in \mathcal{P}$ (the *parameter space*). Let $\mathcal{F}_{\mathcal{P}}$ be a family of sets $\mathcal{F}_{\mathbf{p}} \subset \mathcal{S}$ indexed by $\mathbf{p}$. We define the Hough Transform of input $i(\mathbf{s}), \mathbf{s} \in \mathcal{S}$ as:

$$\mathrm{HT}[\mathcal{F}, i](\mathbf{p}) = \int_{\mathcal{F}_{\mathbf{p}}} i(\mathbf{s}) d\mathbf{s} \tag{1}$$

The parameter space that we choose to use here is the one representing lines in $\mathbb{R}^2$. As we will show later, this choice does not impose a requirement about the presence of lines in the environment; however, with this parametrization we can both exploit some important properties (described later in this Section) and take advantage of the presence of lines for increasing performance.

Different parameterizations can be found for representing a line in $\mathbb{R}^2$. The common choice is to use the polar representation $x \cos\theta + y \sin\theta = \rho$. With this representation ρ is the distance of the line from the origin, and θ the direction of the normal vector (see Fig. 1). The family of sets for applying the Hough Transform is therefore the family of lines indexed by ρ and θ: $\mathcal{F}_{(\theta,\rho)} = \{(x, y) \mid x \cos\theta + y \sin\theta = \rho\}$

In this paper we assume the input space to be a finite set of points $P = \{p_j\}$ (i.e. the output of a range sensor). Thus, we define $\mathbf{s} := (x, y) \in \mathcal{S} := \mathbb{R}^2$ and $i(\mathbf{s}) = \sum_j \delta(\mathbf{s} - \mathbf{p}_j)$ where δ is the Dirac impulse distribution.

The properties of the Hough Transform for lines are:

1) $\mathrm{HT}(\theta, \rho)$ is 2π-periodic.

2) the two points in the $\mathcal{P}$ space: (θ, ρ) and $(\theta + \pi, -\rho)$ represent the same line in $\mathbb{R}^2$:

$$\mathrm{HT}(\theta, \rho) = \mathrm{HT}(\theta + \pi, -\rho) \tag{2}$$

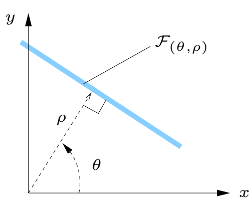


Fig. 1. HT.

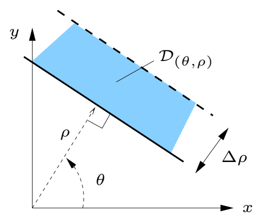


Fig. 2. DHT.

It follows that for representing the set of lines in $\mathbb{R}^2$ only a subset of $\mathcal{P}$ is needed: two possible choices are $\mathcal{P} = \Theta \times \mathcal{R} = [0, \pi) \times \mathbb{R}$ and $\Theta \times \mathcal{R} = [0, 2\pi) \times [0, \infty)$; both can be used depending on the application.

It is also interesting to analyze the behavior of the HT with respect to rigid transformations of the input space. Let $i(\mathbf{s})$ and $i'(\mathbf{s})$ be two inputs such that $i'(\mathbf{s}) = i(R_\phi \cdot \mathbf{s} + T)$, with R_ϕ and T denoting a rigid 2D transformation, and let $\mathrm{HT}(\theta, \rho)$ and $\mathrm{HT}'(\theta, \rho)$ be their transforms, then, if we remain within the boundaries of $\mathcal{P}$, we have:

$$\mathrm{HT}'(\theta, \rho) = \mathrm{HT}(\theta + \phi, \rho + (\cos\theta \quad \sin\theta)\, T) \qquad (3)$$

Note that:

- If $T = \mathbf{0}$, then $\mathrm{HT}'(\theta, \rho) = \mathrm{HT}(\theta + \phi, \rho)$, that is, the $\mathcal{P}$ space translates in the θ direction (Fig. 3).
- If $\phi = 0$, then the $\mathcal{P}$ space bends only in the ρ direction, because $(\cos\theta \quad \sin\theta)\, T$, the displacement for the θ column, does not depend on ρ (Fig. 4).

A. The Discrete Hough Transform

The algorithmic implementation of HT uses the *Discrete Hough Transform* (DHT), which is a discrete approximation of the HT and is quite easy and efficient to compute.

Given the family of sets for the HT (Fig. 2):

$$\mathcal{D}_{(\theta,\rho)} = \{(x, y) \mid \rho \leq x\cos\theta + y\sin\theta < \rho + \Delta\rho\} \qquad (4)$$

and a sampling $\Delta\rho$, $\Delta\theta$ of $\mathcal{P}$, the parameter space is discrete and finite once appropriate bounds are chosen.

A fast way to compute the DHT [9] is to set up a bi-dimensional array of accumulators, whose size depends on the discretization for ρ and θ, that represents the whole $\mathcal{P}$ space. The accumulators are initialized at 0, then for each p_j in the input space, the curve $\rho = (\cos\theta \quad \sin\theta)\, p_j$ is drawn on the buffer, increasing the accumulators by one. Fig. 5 shows that the loss of information when performing the DHT is small.

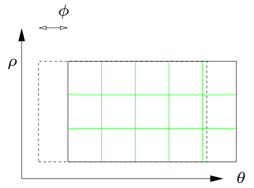


Fig. 3. Deformation if $T = 0$.

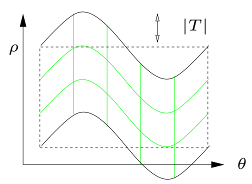
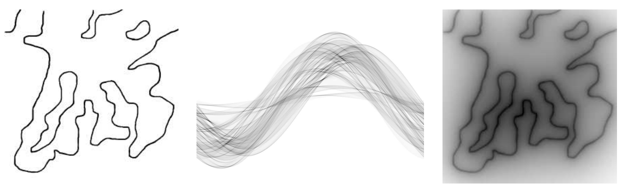


Fig. 4. Deformation if $\phi = 0$.



(a) i (b) $\mathrm{DHT}[i]$ (c) i'

Fig. 5. DHT introduces small loss of information. Here $i' = \mathrm{DHT}^{-1}[\mathrm{DHT}[i]]$.

B. The Hough Spectrum

Definition of Hough Spectrum. Let g be a translation invariant functional (that is, $f'(\tau) = f(\tau + a) \Rightarrow g[f] = g[f']$). We define the Hough Spectrum as

$$\mathrm{HS}_g[i](\theta) := g\Big[\mathrm{HT}[i](\theta, \cdot)\Big] \qquad (5)$$

Spectrum invariance. *Let $\mathcal{R} = \mathbb{R}$ and*

$$i'(\mathbf{s}) = i(R_\phi \cdot \mathbf{s} + T) \qquad (6)$$

Then

$$\mathrm{HS}_g[i](\theta) = \mathrm{HS}_g[i'](\theta + \phi) \qquad (7)$$

Proof: *It follows from (3).*

A broad class of functionals can be chosen as g; we will use the energy of the sequence in the discrete domain:

$$g[f] = \sum_i {f_i}^2 \qquad (8)$$

Unfortunately, property (7) doesn't hold exactly when considering the DHT, because for a single column $\mathrm{DHT}(\theta, \cdot)$ the translational invariance holds exactly only if the quantity $(\cos\theta \quad \sin\theta)\, T$ is a multiple of $\Delta\rho$; however, the subsequent algorithm steps are robust to this small variation of $g[\mathrm{DHT}'(\theta, \cdot)]$.

V. The Hough Scan Matcher

A. Overview of the algorithm

1) The DHT and DHS are computed for reference and sensor data.
2) Local maxima of the spectra cross-correlation are used to produce hypotheses on ϕ.
3) For each hypothesis ϕ, linear constraints for T are produced by correlating columns of the DHT.
4) Two alternatives:
    a) The linear constraints are combined to produce multiple hypotheses for T and solutions are ordered with a likelihood function.
    b) The results of the correlations are accumulated in a buffer producing a dense output.

B. Spectra correlation and ϕ estimation

Because of (7), we can estimate ϕ by correlation of $\mathrm{DHS}^S(\theta)$ and $\mathrm{DHS}^R(\theta)$. Local maxima of the correlation are found and ordered, then for each maximum, beginning with the greatest and up to an arbitrary number, a new hypothesis for ϕ is generated.

Let $\mathrm{DHS}^S(\theta)$ and $\mathrm{DHS}^R(\theta)$ be the discrete spectra of sensor and reference data. Let Φ be the search domain for ϕ; if one has an estimate ϕ_U of the upper bound of ϕ (for example from odometry data), then $\Phi = [-\phi_U, \phi_U]$ (local search), whereas $\Phi = [-\pi, \pi]$ for the uninformed case (global search). The correlation of the two spectra is defined by:

$$\mathrm{corr}_{\mathrm{DHS}}(\phi) = \sum_{\theta \in \Theta} \mathrm{DHS}^S(\theta) \cdot \mathrm{DHS}^R(\theta - \phi) \qquad (9)$$

where $\phi \in \Phi$. The hypotheses for ϕ are given by $\{\phi_1, \dots, \phi_k\} = \text{localMaxOf}\{\text{corr}_{\text{DHS}}(\phi)\}$[1]. An example of spectra correlation is shown in fig. 6.

Note that if $\mathcal{R} = \mathbb{R}$, because of (2), (9) is π-periodic therefore can be computed for half of the domain; on the other hand we get information only about the direction and not the heading of the robot. If $|T|$ is small, it is useful to use $\mathcal{R} = \mathbb{R}^+$ for which (9) is 2π-periodic, therefore giving directly heading information.

ϕ hypotheses can be created directly regardless of T, but for large environments this would produce too many hypotheses; a large map should be divided in smaller patches and the matching process be carried out separately for each patch.

C. T estimation

At this point we have solved half of the original problem; several 2D estimation algorithms could be plugged here after an hypothesis for ϕ exists, however we will propose one that exploits the already computed DHT. For the sake of notation clarity and without loss of generality, we will assume that $\phi = 0$ (this is equivalent to shift the columns of either DHT^S or DHT^R by the previous computed ϕ).

Choose an arbitrary $\theta = \hat{\theta}$ and consider the columns $\text{DHT}(\hat{\theta}, \rho)$ for both the reference and sensor reading. Because of (3) and $\phi = 0$:

$$\text{DHT}^R(\hat{\theta}, \rho) = \text{DHT}^S(\hat{\theta}, \rho + (\cos\hat{\theta} \quad \sin\hat{\theta})T) \tag{10}$$

Again, by correlation of columns of DHT^R and DHT^S we can estimate $(\cos\hat{\theta} \quad \sin\hat{\theta})T = d(\hat{\theta})$, that is the projection of T in the $\hat{\theta}$ direction. If one has an estimate $|T|_U$ of the upper bound for $|T|$, then the search domain for $d(\hat{\theta})$ is $[-|T|_U, |T|_U]$. By considering two different values for θ and finding the maximum of the correlation, we can build a linear system to determine T. Different pairs (θ_1, θ_2) will give different results for T. One could use the resulting T of each pair as a different hypothesis or combine more than two directions to get an over-constrained system to be resolved by least squares estimation:

$$i = 1 \dots n : \quad (\ \cos\theta_i \quad \sin\theta_i\ )T \quad = \quad d(\theta_i) \tag{11}$$

1) Choice of alignment directions: It is still an open question how to choose the n directions for linear constraints. Of course not all directions can be chosen for performance reason (at least in an on line setting). Even in the $n = 2$ case, there are different needs: to keep as independent as possible the observations $d(\theta_1)$ and $d(\theta_2)$ (perpendicular) and choose directions along which are well defined the local maxima of the correlation. One solution is to choose the local maxima of the sensor DHS: $\{\theta_1, \dots, \theta_n\} = \text{localMaxOf}\{\text{DHS}^S(\phi)\}$. Because of our previous choice of the energy functional for g, we are choosing columns whose correlation will have an high expected energy therefore with clearly definite peaks.

[1]Here we denote by localMaxOf the extraction of the local maxima - an operation that in practice needs carefully chosen thresholds or smoothing.

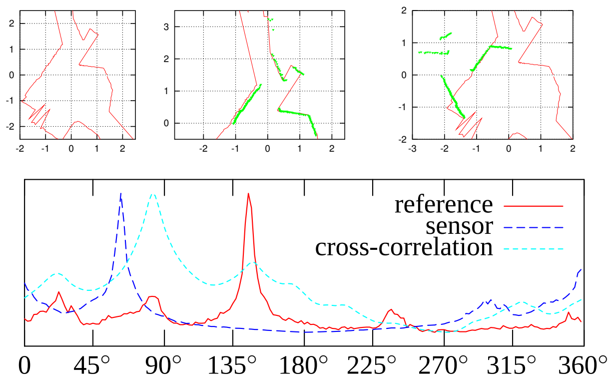


Fig. 6. In (a) it is shown the reference data, resulting from ray-tracing a map of a Nist Rescue arena. In (b) and (c) it is shown the sensor data (laser range finder, shown as points) matched according to the first two solutions given by HSM. (b) is the "exact" match, apart from map inconsistencies, however also (c) is a plausible solution.

2) Dense approach: Alternatively one can accumulate the results of each correlation and produce a dense output: see Fig. 7(b)-7(c). One could use up to every column, and approximate complex distributions (see fig.7(c)). In polygonal environments 2 to 4 directions are usually enough.

3) The under-constrained case: There are situations when the matching problem is ill-posed, for example a robot travelling in a long corridor with limited horizon has essentially only one direction for aligning, and any other choice of alignment direction will cause erroneous results that will deteriorate the estimate of T. One solution for detecting such under-constrained cases is to linearly normalize the DHS in the $[0, 1]$ range and accept as alignment directions only those values over a certain thresholds, say over 0.5 the global maximum. If only one direction is used then an under determined system is produced and the conservative choice for pose tracking is to assume for T the minimum norm solution: $T = (\cos\theta_1 \quad \sin\theta_1)^t\, d(\theta_1)$.

D. Complexity

Let

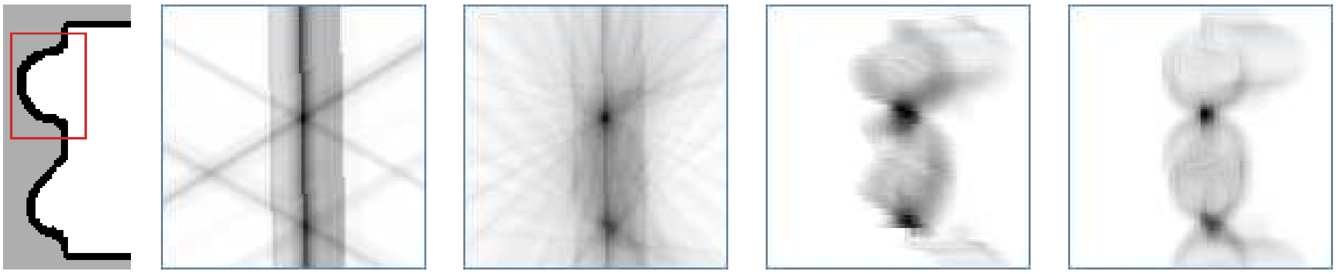
- T_{max} and ϕ_{max} be bounds for $|T|$ and $|\phi|$
- σ_ϕ and σ_T be the required angular/linear discretization (i.e. resolution of the solution)
- S,M the number of points in sensor/reference
- r_S, r_M the minimum radii containing all points for sensor/reference (i.e. range of the scans)
- n_ϕ the number of ϕ hypotheses tracked
- n_c the number of linear correlations.

In this case the DHT buffer dimensions are $|\Theta| \propto \sigma_\phi$, $|\mathcal{R}| = r_M \sigma_T$ therefore space occupation is $O(\sigma_\phi r_M \sigma_T)$.

Building the HT and HS takes $O(S\sigma_\phi)$ and $O(M\sigma_\phi)$. Finding ϕ costs $O(\phi_{\text{max}}\sigma_\phi \cdot \sigma_\phi)$ The column correlation is done $n_\phi n_c$ times and each costs $(T_{\text{max}}\sigma_T)(\sigma_T r_S)$. Total cost is:

$$O(S\sigma_\phi + M\sigma_\phi + \phi_{\text{max}}\sigma_\phi^2 + n_\phi n_c T_{\text{max}} r_S \sigma_T^2) \tag{12}$$

When doing position tracking we are interested in increasing the resolution of the sensor and of the solution. In this case the cost grows as $S\sigma_\phi + \sigma_\phi^2 + \sigma_T^2$, that is, linear in the number of data points and quadratic on each



(a) Input (b) HSM ($n_c = 3$) (c) HSM ($n_c = 15$) (d) Log likelihood (e) CBML

Fig. 7. Fig. 7(a) shows the input data: the red mark encloses the part that it is used as sensor data. These pictures show the second part of the algorithm, T estimation with dense buffer, after that HSM has (correctly) found as a solution $\phi = 0$. Fig. 7(b) shows the distribution resulting from correlating the 3 columns to the three peaks of the HS, while 7(c) comes from uniformly sampling 15 columns from the HT. Fig. 7(d) shows the output of the log likelihood for $\phi = 0$, likewise fig. 7(e) shows the output of CBML.

of the discretizations. Note that in this case the search space grows as $\sigma_\phi \sigma_T^2$.

For global localization we are interested in increasing the environment size. In this case the cost of our algorithm grows as $n_\phi n_c T_{\max}$, while the search space grows as $T_{\max}^2$. However, note that in this case also n_ϕ grows and as mentioned before one should divide the map in patches before applying HSM.

VI. Experiments

To gather significant statistical data the experiments have been conducted with a simulator using scan data available through the Internet. Given a scan data set, a reference position and a scan position are randomly sampled from the environment with a fixed displacement. The raw ray-tracing result from the first position is assumed as the reference scan. The sensor scan is simulated using four different sensor models (see Fig. 10):

1) *Ideal-180*, a range finder with $180°$ range and $1°$ resolution and low noise (similar to *Sick LMS*);
2) *Disc.Noise-180*, a range finder with high quantization noise (similar to *Sick PLS*);
3) *Gaus.Noise-180*, a range finder affected by high Gaussian noise with variance growing with distance (the *Hokuyo PB9-01* as described in [10]);
4) *Syst.Noise-360*, a range finder subject to systematic error but with large FOV (e.g., the edge-extracting pipeline of an omnidirectional camera).

We report the results obtained with two exemplar environments: an office-like environment (a large hospital) and an unstructured environment (hand-drawn map of a cave)[2].

In these experiments we are mainly interested in evaluating the ability of the system to deal with large uncertainty in the prior pose. Therefore, we evaluate different situations in which the sensor scan is displaced by the reference position by 0, 0.5, and 1 m, respectively. Note that the *average beam reading distance* in the data sets considered is about 1.6 m, therefore the value 1 m is quite high. Furthermore, we assumed no prior information about scan orientation (in fact, performing a $360°$ matching).

Because this method performs a global search over orientation ϕ, the error distribution $|e(\phi)|$ is multi-modal (due to the presence of orthogonal lines in the hospital environment), and the error distribution $|e(T)|$ is composed of a single definite peak with super-imposed noise. Two typical error distributions are shown in Fig. 9. In other words, error distributions are formed by a Normal distribution and additional noise and in order to make a more precise error analysis we have taken into consideration the following parameters for each error distribution: 1) the amount of samples in the principal mode of the distribution, that is a measure of the reliability of the method; 2) the mean of the principal mode, that shows its precision.

Results for the two data sets are reported in Table 11. We compute the above mentioned measures for both orientation ϕ and translation T for three scan-reference displacements (0, 0.5, 1 m) and for each of the sensor model considered. Error analysis shows that in a polygonal environment high precision and reliability can be obtained for ϕ and T, and that the method is able to tolerate both various sensor noises and large input differences. In an unstructured environment, we still have good results for limited scan displacements, while performance degrade with increasing distance between current and reference scans.

It is also interesting to notice that, with increasing of input differences, a sensor with large field of view and higher noise has better performance than one with smaller field of view and accurate measures.

VII. Conclusions and future work

In this paper we have proposed a scan matching algorithm (HSM) that works in the Hough domain. The key concept of HSM is to transfer the matching problem to another domain through the use of an information-preserving transformation. The invariance properties of the Hough domain allow to decouple the problem in 1) finding the orientation error, and 2) finding the translation error once the orientation error is known. HSM does not rely on features extraction, but it matches dense data (that can be interpreted as ”features distributions”) in a different

[2] Data sets obtained from the Robotics Data Set Repository (Radish) (radish.sourceforge.net - thanks to Richard Vaughan) and from Map of the Victoria fossil cave in Naracoorte (Australia), courtesy of the Cave Exploration Group of South Australia (www.cegsa.org.au).

Fig. 8. Part of the cave map used in the experiments.

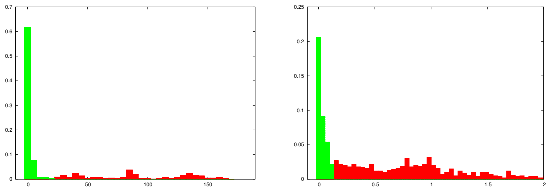


Fig. 9. Typical shape of error distributions (Gaus.Noise-160, cave, $d = 1m$) with principal mode highlighted. $|e(\phi)|$ distribution is multimodal (left), $|e(T)|$ distribution has a peak and uniform noise (right).

Sensor	FOV	res.	quant.	dist.	σ
Ideal-180	180°	1°	0.01m	1	$0.01 \cdot d$
Disc.Noise-180	180°	1°	0.07m	1	$0.03m$
Gaus.Noise-160	160°	1.78°	0.005m	1	$0.01 \cdot d^2 - 0.0017 \cdot d + 0.0075m$
Syst.Noise-360	300°	4°	0.01m	1.15	$0.01 \cdot d$

Fig. 10. Parameters for sensor models. If d is the "real" distance to the obstacle then the perceived distance $\hat{d}$ is given by $\hat{d} = \mathcal{N}(\text{dist} \cdot d, \sigma^2)$ with σ computed as in the table. Parameters for Hokuyo were taken from [10].

	$\|\phi\|$ peak mass	$\|\phi\|$ peak average	$\|T\|$ peak mass	$\|T\|$ peak average	$\|\phi\|$ peak mass	$\|\phi\|$ peak average	$\|T\|$ peak mass	$\|T\|$ peak average
Ideal-180	98%	$< 1^\circ$	97%	$< 1cm$	90%	$< 1^\circ$	89%	$< 1cm$
Disc.Noise-180	97%	$< 1^\circ$	93%	$4cm$	72%	4°	71%	$4cm$
Gaus.Noise-160	94%	$< 1^\circ$	82%	$2cm$	80%	2°	77%	$2cm$
Syst.Noise-360	99%	$< 1^\circ$	98%	$5cm$	97%	1°	97%	$3cm$

(a) Hospital, $d = 0$ (b) Cave, $d = 0$

	$\|\phi\|$ peak mass	$\|\phi\|$ peak average	$\|T\|$ peak mass	$\|T\|$ peak average	$\|\phi\|$ peak mass	$\|\phi\|$ peak average	$\|T\|$ peak mass	$\|T\|$ peak average
Ideal-180	96%	$< 1^\circ$	86%	$1cm$	82%	$< 1^\circ$	70%	$18cm$
Disc.Noise-180	96%	$< 1^\circ$	88%	$5cm$	67%	4°	57%	$8cm$
Gaus.Noise-160	95%	$< 1^\circ$	86%	$3cm$	77%	2°	74%	$10cm$
Syst.Noise-360	98%	$< 1^\circ$	96%	$8cm$	87%	2°	84%	$7cm$

(c) Hospital, $d = 0.5$ (d) Cave, $d = 0.5$

	$\|\phi\|$ peak mass	$\|\phi\|$ peak average	$\|T\|$ peak mass	$\|T\|$ peak average	$\|\phi\|$ peak mass	$\|\phi\|$ peak average	$\|T\|$ peak mass	$\|T\|$ peak average
Ideal-180	91%	$< 1^\circ$	72%	$2cm$	74%	$< 1^\circ$	28%	$8cm$
Disc.Noise-180	91%	$< 1^\circ$	71%	$6cm$	58%	4°	40%	$11cm$
Gaus.Noise-160	89%	$< 1^\circ$	68%	$3cm$	70%	2°	47%	$9cm$
Syst.Noise-360	95%	$< 1^\circ$	77%	$10cm$	72%	2°	54%	$10cm$

(e) Hospital, $d = 1$ (f) Cave, $d = 1$

Fig. 11. Tables show statistical information about the distribution of errors on heading ($|e(\phi)|$) and position ($|e(T)|$) for the two environments (office-like and cave-like), four sensor models (see fig. 10) and three values for displacement of sensor and reference scan ($d = 0, 0.5, 1$ m), with unknown orientation (1000 iterations performed for each combination). The faster discrete-constraint HSM has been used with these parameters: $\Delta\theta = 0.5, \Delta\rho = 4cm$ for Gaus.Noise-160, $\Delta\theta = 0.5$, $\Delta\rho = 2cm$ for the others. Shown in the columns: i) ϕ error samples in the first mode; ii) average of these samples iii) T error samples in the first mode; iv) average of these samples

convenient parameter space; this allows to match non-linear surfaces and makes the process robust to noise.

Derivation of the algorithm was done in a continuous input space $i(s)$. This could allow in the future to relate in an analytical way the matching error to the sensor model, that is to express the uncertainty of the matching to the noise parameters of the sensor's model.

The employment of HSM in large scale localization must be investigated further. CBML [8] has complexity $O(S \cdot \sigma_T^2 T_{\max}^2 \cdot \sigma_\phi \phi_{\max})$ therefore neither HSM nor CBML dominates the other. Moreover both methods are based on a simple operation (bi-dimensional correlation for CBML, unidimensional correlation for HSM) and the use of processor-specific extensions can considerably change the running time. Moreover, a proper comparison should employ some kind of metrics on the resulting distribution, and would probably result in a trade-off of speed (HSM) vs. thoroughness of the search (CBML).

In this paper we have dealt with the 2D scan matching, where the input space is bi-dimensional and the solution space is three-dimensional (ϕ, T_x, T_y), effectively subdividing the original problem in simple uni-dimensional sub-problems (finding maxima of a cross-correlation). 2D scan matching is a low-dimensional problem when compared with the vast range of complex sensors and robot mobilities available nowadays: for example scan matching for a snake-like robot in an unstructured environment with a stereo camera would be a problem with 3D input space and 6 degrees of freedom solution. In the near future we intend to investigate an extension of HSM to higher dimensions, attempting to exploit and refine the ability of subdividing the search problem into low-dimensional problems.

REFERENCES

[1] G. Grisetti, L. Iocchi, and D. Nardi, "Global hough localization for mobile robots in polygonal environments," in *Proc. of IEEE International Conference on Robotics and Automation (ICRA'2002)*, 2002.

[2] J. S. Gutmann and C. Schlegel, "AMOS: Comparison of scan matching approaches for self-localization in indoor environments," in *1st Euromicro Workshop on Advanced Mobile Robots (EUROBOT)*, 1996.

[3] K. Lingemann, H. Surmann, A. Nüchter, and J. Hertzberg, "Indoor and outdoor localization for fast mobile robots," in *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'04)*, 2004, pp. 2185–2190.

[4] F. Lu and E. Milios, "Robot pose estimation in unknown environments by matching 2d range scans," in *CVPR94*, 1994, pp. 935–938.

[5] D. Hähnel, D. Schulz, and W. Burgard, "Map building with mobile robots in populated environments," in *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2002.

[6] P. Biber and W. Strasser, "The normal distributions transform: a new approach to laser scan matching," in *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2003.

[7] G. Weiß, C. Wetzler, and E. von Puttkamer, "Keeping track of position and orientation of moving indoor systems by correlation of range-finder scans," in *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 1994.

[8] K. Konolige and K. Chou, "Markov localization using correlation," in *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI'99)*, 1999.

[9] R. Duda and P. Hart, "Use of the hough transformation to detect lines and curves in the pictures," *Communications of the ACM*, vol. 15, no. 1, pp. 11–15, 1972.

[10] H.-H. Kim, Y.-S. Ha, and G.-G. Jin, "A study on the envornmental map building for a mobile robot using infrared range-finder sensors." in *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'03)*, 2003, pp. 711–716.