

Bearing-only SLAM in Indoor Environments Using a Modified Particle Filter

N. M. Kwok and G. Dissanayake

ARC Center of Excellence in Autonomous Systems (CAS)

Faculty of Engineering

University of Technology Sydney

Broadway, NSW, 2007, Australia

ngai.kwok,gdissa@eng.uts.edu.au

Abstract

The implementation of a particle filter (PF) for vision-based simultaneous localisation and mapping (SLAM) for a mobile robot in an unstructured indoor environment is presented in this paper. Variations to standard PF are proposed to remedy the sample impoverishment problem in bearing-only SLAM. A CCD camera mounted on the robot is used as the measuring device and image quality is incorporated into data association, PF update and map management. A passive path control strategy to maintain the accuracy of the SLAM process is also illustrated. Experimental results from an implementation using real-life data acquired from a Pioneer robot are included to demonstrate the effectiveness of our approach.

1 Introduction

In mobile robot applications, it is a fundamental requirement that the robot should be able to know its position within its operating environment. This can be achieved by using odometry measurements and referring to a map of the environment. However, odometry feedback tends to accumulate errors in time, therefore other measuring devices are needed. A detail and precise map is also not always available and we have to build the map as the robot operates. Therefore, we have to solve the simultaneous localisation and mapping (SLAM) problem. Solutions to SLAM [Dissanayake et. al., 2001] in indoor environment use a laser range-and-bearing scanner and put the problem in the extended Kalman filter (EKF) framework. The laser scanner is accurate but is heavy and expensive. The EKF operates satisfactorily when the system is not severely non-linear and non-Gaussian. If this does not hold true, one has to employ particle filters (PF) [Gordon et. al., 1993] for a better representation of the probability density function.

Apart from using a laser scanner as the measuring device, cameras are popular alternatives because of their

reduced cost, weight and power consumption. Related work can be found in [Davison, 2002] where a pan-tilt camera is used for SLAM and also uses the EKF. Work on a combination of laser and vision is reported in [Perez et. al., 1999] and [Arras et. al., 2000]. Use of vision and odometry is reported in [Crowley, 1992], where landmarks are known and listed in a database. Similarly, in [Mufioz and Gonzalez, 1998], pre-stored landmarks are also used in localising a mobile robot using a single image. Another approach by pre-recording image sequences is found in [Jeon and Kim, 1999] but this restricts the operating area of the robot. Although the use of vision has its many advantages, detection and selection of landmarks from image sequences is challenging [Livatino and Madsen, 1999]. On the other hand, some research work for outdoor SLAM can be found in [Fitzgibbons and Nobet, 2001].

The PF is based on the Monte Carlo simulation technique and uses samples to represent the system probability distribution function (PDF). This filter out-performs the EKF for non-linear systems. However, the PF is limited by its need for a large number of samples to represent the PDFs making it computational expensive. The PF also suffers from the so-called sample impoverishment problem that samples tend to converge to a confined region in the solution space, making state estimations trapped in local optimums. Methods to remedy the problems can be found in [Gordon et. al., 1993].

This paper presents research on the use of a fixed camera mounted on a mobile robot for simultaneous localisation and mapping in an unstructured indoor environment. Landmarks are extracted from edges in the scene and stored as subsequent matching templates while the robot is moving.

To ensure that the PF operates over an extended period of time, we tackle the sample impoverishment problem by introducing techniques from genetic algorithms (GA). Similarities between PF and GA have been pointed out in [Higuchi, 1997]. Further reference is also found in [Bienvenue et. al., 2001] where multi-

objective estimation is treated. Variations to the PF using GA techniques are proposed in [Deutscher et. al., 2001] where enhancements to the PF is achieved with a crossover operator. Another way of modifying PFs is found in [Meier and Ade, 1999] where the PDF is truncated resulting in an extended tail. The PDF is truncated that allows for dynamic inclusion of newly observed landmarks. Here, we use the BLX- α operator which fills the undefined supports of the PDF and expands it [Herrera et. al., 1998].

To improve the robustness of SLAM, we incorporate the image quality [Wang and Bovik, 2002] for data association in addition to the usual nearest-neighbour validation test. Image quality is used in the update step of the PF as motivated by the data association method in [Dezert and Bar-Shalom, 1993]. Map management as in [Dissanayake et. al., 2002] is also used where we remove landmarks by an index combined from image quality and detected edge height. We also consider the effect of robot path that affects the SLAM performance. Related work can be found in [Bianco and Zelinsky, 2000] where navigation guidance is treated.

The rest of this paper is arranged as follows. In section 2, we briefly review the SLAM problem. Then, the PF is introduced in section 3. In section 4, modifications to the particle filter will be presented. Data association using images will be discussed in section 5 followed by the map management in section 6. In section 7, we will propose a passive path control strategy. Experiment results will be presented in section 8 and the discussion and conclusion in section 9.

2 SLAM Problem

Given a mobile robot deployed in its operating environment, we denote its position at time zero as the origin of the world co-ordinate system. That is

$$X_v = [x_v, y_v, \phi_v]^T = [0, 0, 0]^T \quad (1)$$

where x_v, y_v is the location in Cartesian co-ordinates and ϕ_v is the heading with reference to the x-direction

Fig.1 depicts the system definitions. When a control u , consisting of speed and turn rate, is issued; the robot moves according to a transition equation

$$\begin{bmatrix} x_v(k+1) \\ y_v(k+1) \\ \phi_v(k+1) \end{bmatrix} = \begin{bmatrix} x_v(k) + v t \cos(\phi_v(k) + \gamma(k)) + v_x \\ y_v(k) + v t \sin(\phi_v(k) + \gamma(k)) + v_y \\ \phi_v(k) + \gamma(k)t + v_\phi \end{bmatrix} \quad (2)$$

where v is the velocity command, γ is the turn rate, t is the time step and $v_{xy\phi}$ are noise terms lumping in the effects of control response, wheel slip, etc., the noises are characterised by $N(0, \sigma_{xy\phi})$

While moving, the robot makes measurements to landmarks in the operating environment. The observation is

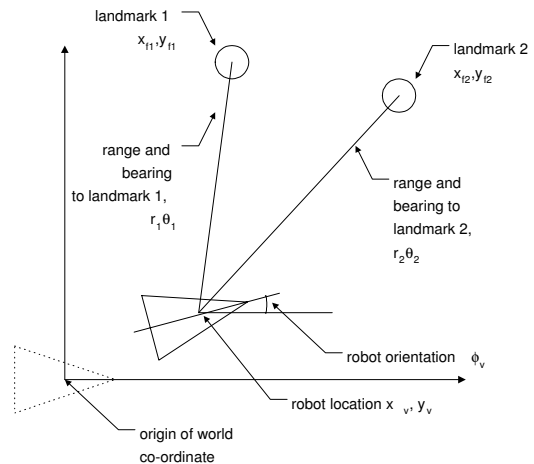


Figure 1: System definition

given by

$$\theta = \arctan\left(\frac{x_{fi} - x_v}{y_{fi} - y_v}\right) - \phi_v + v_\theta \quad (3)$$

where x_{fi}, y_{fi} is the location of a stationary landmark, v_θ is the measurement noise $N(0, \sigma_\theta)$

The SLAM problem is to estimate the aggregated robot and landmark locations, in the form of a state vector, provided with the measurement. The state vector is given as

$$X = [X_v, X_{fi}]^T = [x_v, y_v, \phi_v, x_{fi}, y_{fi}]^T \quad (4)$$

3 Particle Filter

The particle filter (PF) is based on the Bayesian estimation framework. Information about the system is embedded in a probability distribution function (PDF). Using a large number of samples (particles) as in the Monte Carlo methods, the PDF is represented and evolves according to the system transition equation. Upon reception of a measurement, the PDF is updated and the estimation process repeats.

At time zero, we have the system PDF given by $p(x_0)$. The system then evolves according to the transition equation, giving $p(\tilde{x}_{k+1}|x_k)$. When a measurement is available, a likelihood of obtaining that measurement can be found as $p(z_{k+1}|\tilde{x}_{k+1})$. Finally, we combine the PDF according to the Bayes' rule and get a recursive equation

$$p(x_{k+1}|z_{k+1}) \propto p(z_{k+1}|\tilde{x}_{k+1})p(\tilde{x}_{k+1}|x_k)p(x_k|z_k) \quad (5)$$

Details of theoretical development can be found in [Gordon et. al., 1993]. Implementation of the particle filter can be described as follows.

1. randomly initialise particles, x^i

2. in every time step, propagate particles through the transition equation, $\tilde{x}^i \leftarrow x^i$
3. calculate the likelihood (weight) from the measurement, $p(z^i|\tilde{x}^i)$
4. re-sample (select) particles according to their weights, $x^j \leftarrow x^i$

The implementation of a PF includes propagating the particles through the transition equation. However, in the SLAM problem, the landmarks are stationary and the landmark particles¹ will soon collapse onto a single point, which is called sample impoverishment.

4 Modified Particle Filter

There are several research works on the similarity between particle filtering and genetic algorithms (see introduction). We will make use of GA techniques to modify the PF implementation.

4.1 Crossover

We will adopt the BLX- α operator [Herrera et. al., 1998] to *move* the landmark particles in each iteration. The procedure is given below.

1. select 2 particles x^i, x^j , with equally likely probability
2. calculate the distance d^{ij} between the 2 particles in Cartesian co-ordinate
3. move the particles according to the product of the α parameter and the distance, giving $\tilde{x}^{ij} = x^{ij} \pm (\alpha \times d^{ij})$
4. further move the particles by a uniform random parameter $r \in [0,1]$, giving $x_{ij} = r \times \tilde{x}^{ij} + (1-r) \times x^j$

Here, we choose parameter $\alpha = 0.5$. This choice results in spreading of particles and helps in reduce the sample impoverishment problem by introducing new potential solutions to the estimation. Referring to [Gordon, 1994], this corresponds to jittering and prior editing.

To minimise the effect of the sample redistribution on the shape of the PDF, we limit the crossover using a small crossover probability ($p_c = 0.1$) and this appears to work well in practice. We are currently in the process of developing a more rigorous algorithm to find an appropriate value for p_c .

Landmark particles far away from the robot have smaller bearing error than those particles near the robot provided they have the same perpendicular displacement from the bearing measurement. The result is that the re-sampling process favours particles far away and makes the estimation biased. Using the same method as above, we again move the particles as follows.

¹Note that a particle represents the system vector, we violate the terminology here for the ease of description

1. separate landmark particles into 2 groups with equally likely probability
2. for 1 group, rotate the particles by 180° centred at the expectation value of the 2 groups, that is

$$\begin{bmatrix} \tilde{x}_i \\ \tilde{y}_i \end{bmatrix} = \begin{bmatrix} \cos(\pi) & -\sin(\pi) \\ \sin(\pi) & \cos(\pi) \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix} \quad (6)$$

When we rotate the particles at the expectation value, this value is un-altered. The rotation by 180° does not involve any scaling or translation and this retains the covariance between the particles. The result is that we have redistributed the particles leaving their statistics un-altered. This reduces far away particles to the re-sampling process and consequently reduces the bias. Fig.2 through Fig.4 shows typical landmark particle histograms (PDF) before and after rotate and move.

4.2 Batch Update

Each update in the particle filter involves re-sampling, thus increases the possibility of sample impoverishment. We try to minimise this possibility by performing batch update by accumulating a series of measurements before each update.

4.3 Measurement Quality

The likelihood function represents the probability of obtaining the real-life measurement given the expected robot and landmark locations. When using a camera as a measuring device, we can incorporate the image quality into the calculation of the likelihood function. We use the following equation to calculate the likelihood.

$$w_i \propto \frac{1}{\sqrt{2\pi\sigma_\theta^2}} \exp\left(-\frac{(\theta - \hat{\theta}_i)^2}{2\sigma_\theta^2}\right) \quad (7)$$

where $\hat{\theta}_i$ is the expected bearing measurement for a particle and $Q \in [0,1]$ is the quality of the landmark image (see section5). When the image quality is poor, most particles will be retained through the re-sampling process. That is, we take a rather conservative approach when measurement is not sure then reduce the effect of updating.

5 Data Association

Data association maintains the correspondence between a measurement and a landmark. It is crucial for the operation of any estimation. In this work, we will use a hybrid association that the nearest neighbour and image quality are combined. We also use hysteresis thresholds to reduce ambiguities.

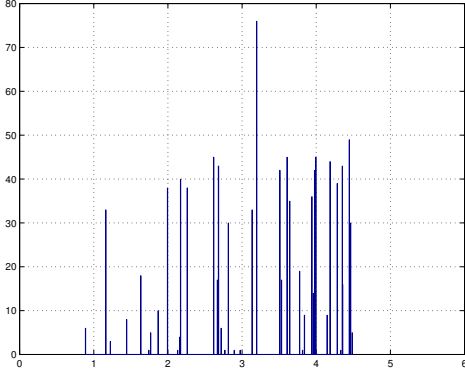


Figure 2: Landmark PDF before rotate

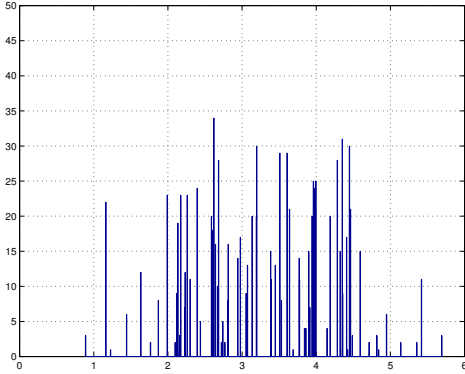


Figure 3: Landmark PDF after rotate and before move

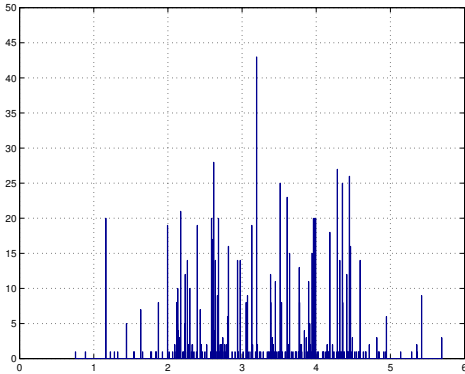


Figure 4: Landmark PDF after move

5.1 Nearest Neighbour

This is a classical approach that determines the difference between the real-life and the expected measurements. The result is thresholded against some confidence level in the χ^2 sense. That is, if

$$\eta_1 = \nu^2 / \sigma < \gamma_1 \quad (8)$$

then declare the measurement from a particular landmark. Where $\nu = \theta - \hat{\theta}$ is the innovation, σ is the error covariance between the robot and the landmark, $\gamma_1 \approx 3.8$ for 95% confidence with 1 degree of freedom (bearing-only measurement).

Otherwise, if

$$\eta_2 = \nu^2 / \sigma > \gamma_2 \quad (9)$$

then declare the measurement from a new landmark. Where $\gamma_2 \approx 10.8$ for 99.9% confidence.

5.2 Image quality

A camera is used in our work as the measuring device providing the bearing-only measurement for SLAM. Image quality then plays a critical role in matching image patches for data association. We follow the criteria proposed in [Wang and Bovik, 2002] using correlation, contrast and intensity characteristics as a metric between successful image frames I_i and I_j . Image I_i is the patch of image in the current frame and I_j is the previous image frame stored during PF iterations. The image quality $Q \in [0,1]$ is defined as

$$Q = \frac{1}{2} \left(1 + \frac{4\sigma_{ij}\bar{i}\bar{j}}{(\sigma_i^2 + \sigma_j^2)(\bar{i}^2 + \bar{j}^2)} \right) \quad (10)$$

where σ_{ij} is the correlation between the images I_i and I_j , σ_i , σ_j are the variances of the 2 images, \bar{i} and \bar{j} are the mean values respectively

To reduce ambiguity, hysteresis threshold is used. If $Q > \gamma_3$ then declare as matched landmark, otherwise, if $Q < \gamma_4$ then declare as new landmarks. Where

$$\gamma_3 = \text{mean}(Q) + 0.5 \times \text{std}(Q) \quad (11a)$$

$$\gamma_4 = \text{mean}(Q) \quad (11b)$$

5.3 Edge Detection

Landmarks from the image frame are detected by finding their edges using the *sobel* mask. Threshold is also used that we accept edges when their edge heights are above the mean values from an image frame. Point landmarks may be detected, however, they are liable to occlusions. We also use spatial separation to resolve ambiguities that we only extract edges from a horizontal strip of the middle of the image. This also speeds up the image processing speed. Fig.5 and Fig.6 show typical scenarios in edge detection, initialising new landmarks and tracking landmarks.

5.4 Overall Data Association

We summarise the overall data association below. First, we pick measurements from nearest neighbour. Second,



Figure 5: Previous frame with detected edges

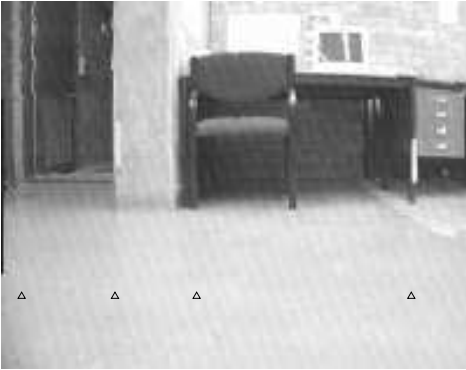


Figure 6: Current frame, initialised (left) and detected edges

compute the image quality and pick the matched edges. Among the measurements, pick the one with

$$\min(\eta_1/(Q > \gamma_3)) \quad (12)$$

as matched measurement and tracked landmark. Otherwise, initialize new landmarks by picking the measurements corresponding to

$$\max(\eta_2/(Q < \gamma_4)) \quad (13)$$

6 Map Management

It has been shown in [Dissanayake et. al., 2002] that it is possible to trim the system state by removing some of landmarks that is not measured and is of low certainty. In this work, we make use of edge height and image quality as the criteria to remove landmarks. We set a maximum number of landmarks, $m = 15$, in our implementation. When the number of landmarks is more than m , then remove those tracked landmarks with a smallest rank. That is, remove the landmark using

$$Q_L = e_g Q \quad (14)$$

where e_g is the stored edge height during tracking, Q is the store image quality

When a new landmark is found, it is initialised by spatial separation that a certain angle span, θ_s , among the tracked landmarks has to be met. It is a reasonable assumption that landmarks are distributed uniformly within an indoor environment where the robot operates. Employing our strategy, landmarks with high edges and high image qualities will remain being tracked.

7 Path Control

With bearing-only SLAM, appropriate path following strategy should be derived for the success of SLAM. In this case, we are inferring the 2-dimensional robot and landmark position from the 1-dimensional bearing-only measurement. The convergence to the true estimation depends on the bearing sustained between the robot and landmarks. A $\pm 90^\circ$ measurement maximises the rate of convergence. However, our camera has a field-of-view of 90° and maximum convergence rate cannot be obtained.

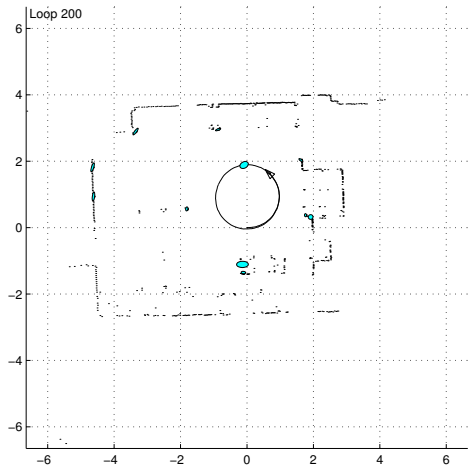
We propose here, to steer the robot in a circular path with *periodic* panning back of the robot. This strategy simulates an extended field-of-view, and the estimation result (see experiment section) outperforms that of pure circular path. By panning the back robot, we may limit the robot uncertainty while it is moving. If this uncertainty is bounded, newly found landmarks will have a lower uncertainty and be kept tracked. Bounding the robot uncertainty also enhances the success rate in data association.

8 Experiment Results

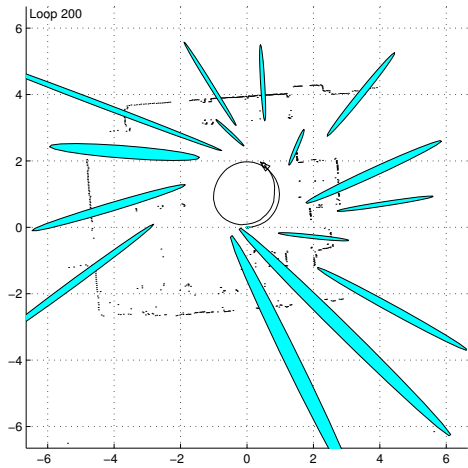
Experiment results from range-and-bearing SLAM using a laser scanner are used as the reference to verify the vision-based bearing-only experiments. The laser scanner is relatively accurate and with laser reflectors as landmarks ensures the robustness of data association and estimation, and is reasonably justified as the reference. However, because of difficulties in detecting laser reflectors as landmark in image frames, it is not possible to make a direct comparison for the landmark locations. As a compromise, we show the laser scans in the figures that illustrate vision-based experiments (note that the laser scans are not used in the estimation) to impose an indication of the operating environment. We have kept 15 landmarks in the estimation and that amounts to 33 states including the robot. In the experiment, we use 10000 particles and obtain extended PF operation through 400 time steps.

8.1 Circular Path

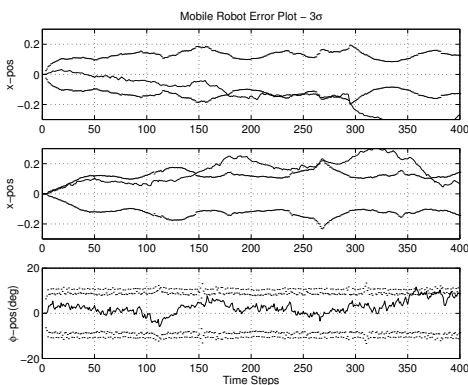
In this experiment, we move the robot in a continuous circular path for about 2 rounds. Fig.7(a) shows a snapshot during the estimation using the laser range-and-



(a) laser range-and-bearing



(b) vision bearing-only



(c) robot position error

Figure 7: Experiment results for Circle Path

bearing measurements. Fig.7(b) shows the corresponding snap-shot for the vision-based bearing-only SLAM. It can be observed that landmark estimations are not as good as the range-and-bearing, however, this is expected as more information is available in the former case. Fig.7(c) shows a plot of the robot position estimation error with the corresponding covariance 3σ bound. Because of extended close-loop time, errors tend to accumulate and it is clear that the estimate becomes inconsistent as time is increased.

8.2 Pan-back Path

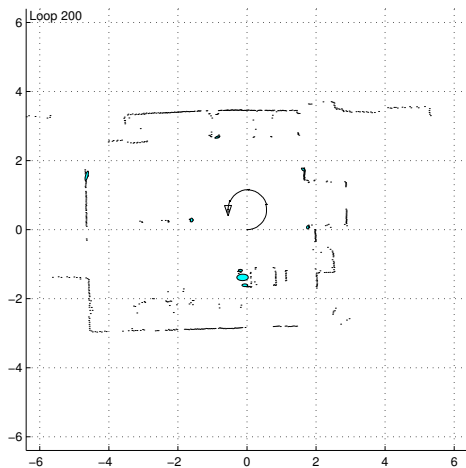
We basically move the robot in a circular path, but we pan back the robot for about 90° when the robot moves about a quarter of a circle. Fig.8(a) shows the laser range-and-bearing experiment result. The vision-based bearing-only experiment snap-shot is shown in Fig.8(b). The result using the laser scanner is quite similar to the previous case. This is due to the fact that the laser scanner is relatively accurate and has a large field-of-view (180°), thus making the estimation rather independent of the path. The vision-based case shows a better performance than the circular path case. This is due to the fact that we periodically pan the robot back and limits the estimation uncertainties. The landmark estimation is inferior to the laser range-and-bearing case, which is again expected. Fig.8(c) shows the robot position estimation error. A better performance is observed than the circle path case. However, the robot heading estimation shows periodic fluctuations, we suspect this is due to errors in the process model that become apparent when the robot turns rapidly during pan back.

9 Discussion and Conclusion

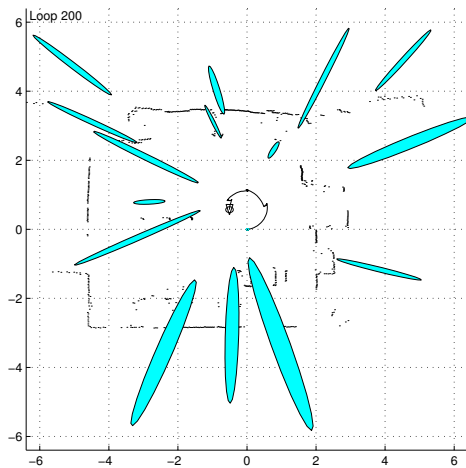
A vision-based bearing-only SLAM is implemented using a particle filter. Preliminary experiment results show that extended operation of the particle filter is possible by manipulating the particle samples by adopting the crossover technique from genetic algorithms. Image qualities are incorporated in filter operation, data association and map management. The results also show that effective path planning is crucial to obtaining acceptable estimation quality, particularly as the field of view of the camera used is limited to 90° .

Acknowledgement

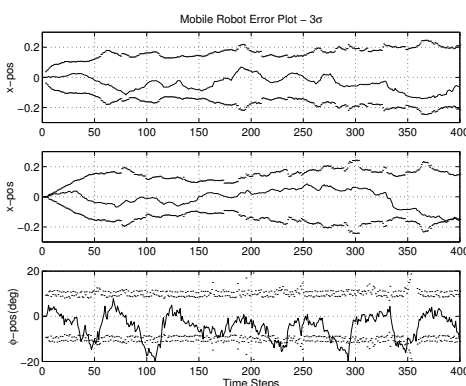
This work is supported by the ARC Centre of Excellence programme, funded by the Australian Research Council (ARC) and the New South Wales State Government.



(a) laser range-and-bearing



(b) vision bearing-only



(c) robot position error

Figure 8: Experiment results for Pan-back Path

References

- [Arras et. al., 2000] Arras K. O., Tomatis N. and Siegwart R., *Multisensor On-the-Fly Localisation Using Laser and Vision*. Proc. 2000 IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems, 462-467, 2000.
- [Bianco and Zelinsky, 2000] Bianco G. and Zelinsky A., *Real-time Analysis of the Robustness of the Navigation Strategy of a Visually Guided Mobile Robot*. Proc. Intl. Conf. on Intelligent Autonomous Systems, IAS2000, Venice, July 2000.
- [Bienvenue et. al., 2001] Bienvenue A., Joannides M., Bernard J., Fontenas E. and Francois O., *Niching in Monte Carlo Filtering Algorithms*. Proc. 5th Intl. Conf. on Artificial Evolution, 19-30, Creusot, October 2001.
- [Crowley, 1992] Crowley J. L., *Position Estimation for a Mobile Robot Using Vision and Odometry*. Proc. 1992 IEEE Intl. Conf. on Robotics and Automation, 2588-2593, Nice, May 1992.
- [Davison, 2002] Davison A. J., *SLAM With a Single Camera*. SLAM/CML Workshop at ICRA 2002, 11-15, May 2002.
- [Deutscher et. al., 2001] Deutscher J., Davison A. and Reid I., *Automatic Partitioning of High Dimensional Search Spaces Associated with Articulated Body Motion Capture*. Proc. 2001 IEEE Compt. Society Conf. on Comput. Vision and Pattern Recognition, 669-676, December 2001.
- [Dezert and Bar-Shalom, 1993] Dezert J. and Bar-Shalom Y., *Joint Probabilistic Data Association for Autonomous Navigation*. IEEE Trans. on Aerospace and Electronic Systems, 29(4):1275-1285, October 1993.
- [Dissanayake et. al., 2001] Dissanayake G., Newman P., Clark S., Durrant-Whyte H. F. and Csorba M., *A Solution to the Simultaneous Localisation and Map Building (SLAM) problem*. IEEE Trans. on Robotics and Automation, 17(3):229-241, June 2001.
- [Dissanayake et. al., 2002] Dissanayake G., Williams S. B., Durrant-Whyte H. F. and Bailey T., *Map Management for Efficient Simultaneous Localisation and Mapping (SLAM)*. Autonomous Robots, 12:267-286, 2002.
- [Fitzgibbons and Nobet, 2001] Fitzgibbons T. and Nobet E., *Application of Vision in Simultaneous Localization and Mapping*. Proc. 2001 Australian Conf. on Robotics and Automation, 121-127, Sydney, November 2001.
- [Gordon et. al., 1993] Gordon N. J., Salmond D. J. and Smith A. F. M., *Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation*. IEE Proceedings F, Radar and Signal Processing, 140(2):107-113, April 1993.
- [Gordon, 1994] Gordon N. J., *Non-linear / Non-Gaussian Filtering and the Bootstrap Filter*. IEE Colloquium on Non-linear Filters, 1-6, May 1994.
- [Herrera et. al., 1998] Herrera F., Lozano M. and Verdegay J. L., *Tackling Real-coded Genetic Algorithms: Operators and Tools for Behavioural Analysis*. Artificial Intelligence Review, 12(4):265-319, 1998.
- [Higuchi, 1997] Higuchi T., *Monte Carlo Filter Using the Genetic Algorithm Operators*. Journal of Statis. Comput. Simul., (59):1-23, 1997.
- [Jeon and Kim, 1999] Jeon S. H. and Kim B. K., *Monocular-based Position Determination for Indoor Navigation of Mobile Robots*. Proc. IASTED Intl. Conf. on Control and Applications, 408-413, Banff, July 1999.

- [Livatino and Madsen, 1999] Livatino S. and Madsen C. B., *Optimization of Robot Self-Localisation Accuracy by Automatic Visual-Landmark Selection*. Proc. SCIA99 11th Scandinavian Conf. on Image Analysis, 501-506, Kangerlussnag, June 1999.
- [Meier and Ade, 1999] Meier E. B. and Ade F., *Tracking Cars in Range Images Using the Condensation Algorithm*. Proc. 1999 IEEE/IEEJ/JSAI Intl. Conf. on Intelligent Transportation Systems, 129-134, October 1999.
- [Mufioz and Gonzalez, 1998] Mufioz A. J. and Gonzalez J., *Two-Dimensional Landmark-based Position Estimation from a Single Image*. Proc. 1998 IEEE Intl. Conf. on Robotics and Automation, 3709-3714, Leuven, May 1998.
- [Perez et. al., 1999] Perez J. A., Castellanos J. A., Montiel J. M. M., Neira J. and Tardos J. D., *Continuous Mobile Robot Localisation: Vision vs. Laser*. Proc. 1999 IEEE Intl. Conf. on Robotics and Automation, 2917-2923, Detroit, May 1999.
- [Wang and Bovik, 2002] Wang Z. and Bovik A. C., *A Universal Image Quality Index*. IEEE Singal Processing Letters, 9(3):81-84, March 2002.