A Monocular Vision Based Localizer

Zhengzhi Zhang and K. R. S. Kodagoda
ARC Centre of Excellence for Autonomous Systems,
University of Technology Sydney
Faculty of Engineering PO Box 123 Broadway NSW 2007,
{ z.zhang, s.kodagoda }@cas.edu.au

Abstract

Utilization of low cost sensing is feasible in office like environments due to the regular man made structures. In this paper, we presented an algorithm, which exploits floor boundaries as a feature for robot localization. A colour based floor detection technique along with laser-like scan generation algorithm is proposed for feature extraction. Then an AMCL based localizer is used for robot localization. Experimental results in an office like environment shows the feasibility of the monocular vision based localization.

1 Introduction

In the past decade, there is a significant continual increase in robotics applications for indoor activities, such as floor cleaning, surveillance, hazard inspection, search and rescue and fire fighting. Typical indoor environment consists of man made structures, which can be geometrically represented. Complexity of sensing can be compromised by utilizing the monotonous properties of the environment. Further, it allows the utilization of low cost sensing.

The floor boundaries or any other object boundaries on the floor are important features for mobile robot navigation. Those boundaries can be used for path planning, obstacle avoidance, navigation and localization. In this paper, we made an attempt to extract floor boundaries using a monocular camera and localize the robot in a given map.

Given the map of the environment, Monte Carlo Localization (MCL) [Menegatti, 2004], [Frank et. al, 1999], [Juan and Franciso, 2000] is one of the robust localization techniques used in the literature. It uses a particle filter based methodology for robot localization. MCL with different sensors such as, laser range finders, sonars and cameras can be found in the literature.

Laser based MCL as in [Sebastian et al., 2001] provides satisfactory results, however the utilization is restricted by cost of the sensor. Low cost sensors, such as sonar can be used for MCL [Sebastian et al., 2001], however, it compromises the accuracy. Stereo camera based MCL [Juan and Franciso, 2000] can be a better alternative for reasonable localization of the robot. However, the cost of sensing can be further reduced by utilizing a monocular camera. A Single camera based MCL utilizing omnidirectional camera is proposed in the [Menegatti, 2004]. A monocular vision, ceiling based localizer is proposed by [Frank et. al, 1999]. [Nguyen, et., al, 2004] used a monocular camera to detect landmarks for localization. The research directions in these applications motivated us using a single camera for robot localization.

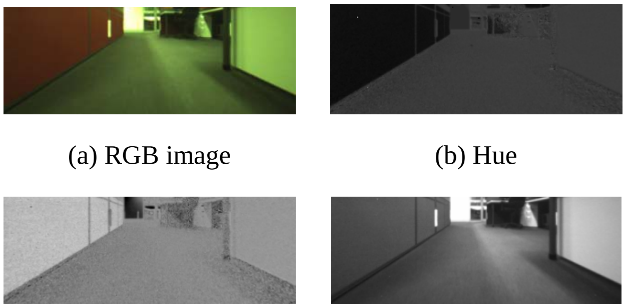
In this paper, a horizontally mounted monocular camera is used for robot localization. The colour images acquired by the camera are processed for floor boundary detection. Those boundaries in image plane are transformed to a laser-like, range bearing information in the robot coordinate system. The range, bearing data and the map of the environment are used for robot localization.

This paper is organized as follows. In section 2, floor detection using monocular vision and laser-like scan generation are presented. Section 3 describes the Monte Carlo Localization. Experiments carried on a Pioneer ® robot in an office like environment are presented in section 4. Section 5 concludes the paper.

2 Floor Detection Using Monocular Vision

Regularities inherently presented in man made indoor environments can be effectively and efficiently extracted using vision sensing. One such very useful feature is the floor. It can be used for localization, navigation and obstacle avoidance. Here an attempt is made to extract the floor boundaries. Fig. 1 (a) shows an image taken from a camera mounted on the Pioneer® robot. One way of detecting floor boundaries is to do edge detection on the grey level image. Unfortunately, the shadows caused by various lightings in indoor environments can introduce unwanted edges as in Fig. 2 causing floor detection difficult.

This can be overcome by transforming RGB image into another colour scheme, such as HSV, which decouples colour information from intensity and brightness. Fig. 1 (d) clearly shows the intensity component, which is decoupled from colour. Now, either hue image or saturation image can be edge detected and it has minimal edges due to intensity variations. However, it may contain some noise due to camera quality as well as various colour patterns in the environment. Further, hue and saturation can be combined to enhance the edges by subtracting saturation image from hue image as shown in Fig. 3.



(a) RGB image (b) Hue

(c) Saturation (d) Value

Fig.1 Colour components of an image

Fig.2 Edge detection on grey image

Fig.3 Subtracted (Saturation from Hue) image

Then edge detection is carried out for the subtracted image (Fig. 4). It can be noted that the important floor edges are now visible with some noisy pixels. Then a morphological operation is carried out to open the connected pixel groups, which have less than a predefined number of pixels (10 pixels in this case). Image after such a morphological operation is shown in Fig. 5.

Fig. 4 Edge detection on subtracted image

Now starting from bottom-middle of the image (Fig. 5), rays are emitted in equal angular resolution until the rays hit a white pixel, which is considered to be a floor boundary pixel. It can be perceived as a laser scan on the image plane, which sends rays and calculate the distance to an object. It is to be noted that the boundaries are not continuous, which causes missed boundary detections. Hence morphological dilation is carried out before detecting boundary pixels. Fig. 5 after morphological dilation with a disk structuring element of size 2 is shown in Fig. 6.

Fig. 5 Image after morphological opening

Fig. 6 Dilated image

Knowing the camera calibration parameters, and assuming a flat floor, it is now possible to transform image coordinates to camera coordinates for the floor boundary pixels as (Fig. 7),

$$\frac{r}{f} = \frac{H_c}{y_c} \Rightarrow y_c = \frac{H_c \times f}{r} \tag{1}$$

$$\frac{x_c}{y_c} = \frac{c}{f} \Rightarrow x_c = \frac{y_c \times c}{f} \tag{2}$$

where, H_c is the height of the camera with respect to the robot coordinates. f is the focal length of the camera. (r, c) are the coordinates of an pixel in the image. (x_c, y_c) is a point in camera coordinates. From, equation (1), (2):

$$Y_R = y_c + a = \frac{H_c \times f}{r} + a \tag{3}$$

$$X_R = x_c + b = \frac{y_c \times c}{f} + b \tag{4}$$

$$Z_R = 0 \tag{5}$$

where, a and b are translations of camera coordinate system in the robot coordinates as shown in Fig. 7.

Fig. 8 shows the transformed image floor boundaries in the robot coordinates as *crosses.* For evaluation purposes, the corresponding laser data in the same coordinates are also shown by symbol *dot*. Fig. 9 shows the stereo vision based floor boundaries (*crosses*) and the laser data (*dots*). It is to be noted that the estimated floor boundaries from monocular vision is superior to that are from the stereo pair. It can be true as far as the flat floor assumption is valid.

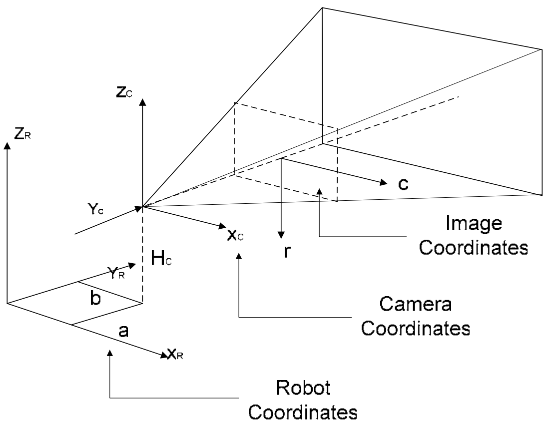


Fig. 7. Coordinate systems

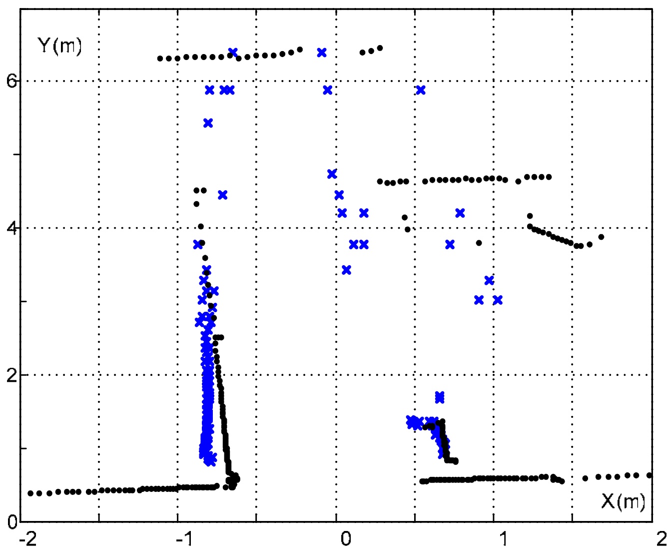


Fig. 8. Floor boundaries: *crosses* – image based, *dots* – laser based

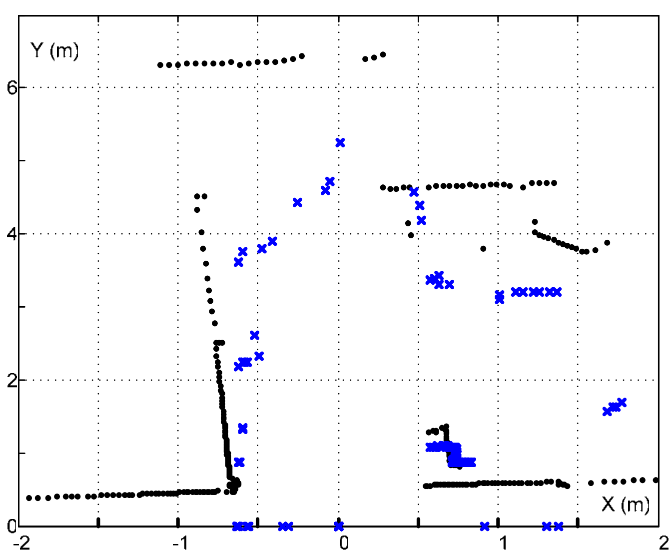


Fig. 9. Floor boundaries: *crosses* – stereo image based, *dots* – laser based

3 Monte Carlo Localization (MCL)

3.1 Introduction

The conventional approach of localizing a robot described the state space by a probability density function. However, the particle filter based localization is achieved by maintaining a set of samples randomly derived from the state space. These density representations are updated by Monte Carlo techniques. One of the advantages of such MCL is that, it can represent multi modal distributions, which can not be accurately represented with conventional Kalman filter based techniques. The robot is given a map with an initial uniformly distributed sample set with same weights. Once an observation is made, the likely hood, $p(z \mid x)$ of the observation is incorporated into the sample set by adjusting the weights. After receiving input, u_{t-1}, a sample set is randomly generated and next location of the robot is predicted $p(x_t \mid x_{t-1}, u_{t-1})$. And this process repeats.

A high number of particles are required for accurately represents the beliefs in the initial stages. Once it attains a reasonably accurate localization, the number of particles required can be low in number. Therefore, adaptive MCL, in which the number of particles is being adaptively determined during the estimation process, is proposed in [Fox, 2003]. Adaptation increases the efficiency of the algorithm and hence it is decided to adopt it.

3.2 Sensor Model

Let σ_r, σ_c be the row and column uncertainties in the image floor boundaries. It should capture the uncertainties due to segmentation and quantization. The uncertainties in the calibration are not introduced for simplicity. The range, R and bearing θ can now be determined using (3) and (4) as,

$$R = \sqrt{X_R^2 + Y_R^2}$$
$$\theta = \tan^{-1}\left(Y_R / X_R \right) \qquad (6)$$

Now the uncertainties of the road boundaries in the robot coordinates can be obtained by,

$$\Lambda_{R,\theta} = J_{r,c} \Lambda_{r,c} J_{r,c}^T \qquad (7)$$

where, $J_{r,c}$ is the Jarcobian of (6) with respect to r and c. $\Lambda_{r,c} = \begin{vmatrix} \sigma_r^2 & 0 \\ 0 & \sigma_c^2 \end{vmatrix}$. σ_c and σ_r were assumed to be 1 pixel each.

3.3 Methodology

The vision based floor boundary detection methodology given in section 2 outputs laser scan like range bearing information to floor boundaries. This information, as well as the map of the environment, and the odometry resulting from the robot, is then passed to the AMCL module for estimating the pose of the robot (Fig. 10).

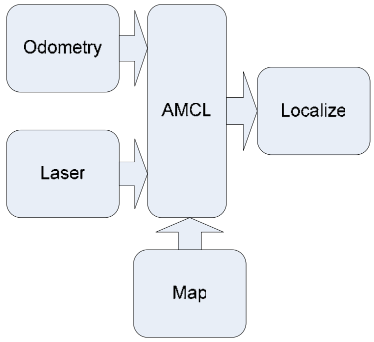


Fig. 10 input and output of AMCL

4 Experiment Results

The platform used in this experiment is a Pioneer ® robot with a SICK laser and a camera on board (see Fig. 11). The algorithm was implemented on Player/Stage environment [Player Stage web-site]. Player is a robot server program, which allows the user's client program to control the behaviour of a robot (here the Pioneer® robot). Stage is a robot simulation platform. It provides an environment, which is identical to the actual environment. User's Player client program can be tested on Stage, and then put into a real environment test with nearly no modification of the client program. Fig.12 shows the basic architecture of Player and user's client program [Player Stage web-site].

Fig. 11. Pioneer robot

The experiment was carried out in an office like environment, which is shown in Fig. 13. It consists of cubicles, chairs, walls, glass doors, etc. Although, people and chairs were present while doing the experiment, they were not included in the map. The stereo camera is used for data collection however, only the images from left camera were used. The laser scanner data is used for comparison purposes. All the sensors including encoders, laser scanner and the camera were synchronized and operated at 3Hz.

4.1 Robot Localization Performance

True path of the robot is important in analysing the errors of localization. It is generated by an Iterative Closest Point (ICP) based algorithm [Tim, 2002] performed on laser data. A comparison of such a generated true path with the odometry along localization is shown in Fig. 14(a). Without surprise, it can be seen that the odometry along path is erroneous. The problem with odometry is that any error introduced is integrated with time degrading localization accuracy. Fig. 14 (b) shows the monocular vision based AMCL localizer results with the true path on the same plot. It can be seen that the localization is relatively closed to true path, when comparing with the path generated by odometry as in Fig 14 (b).

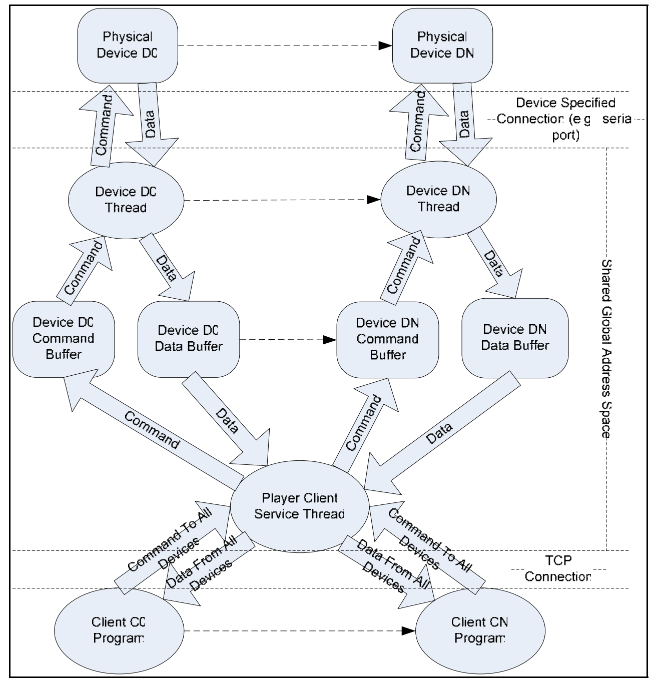


Fig. 12 Architecture of Player and the client program

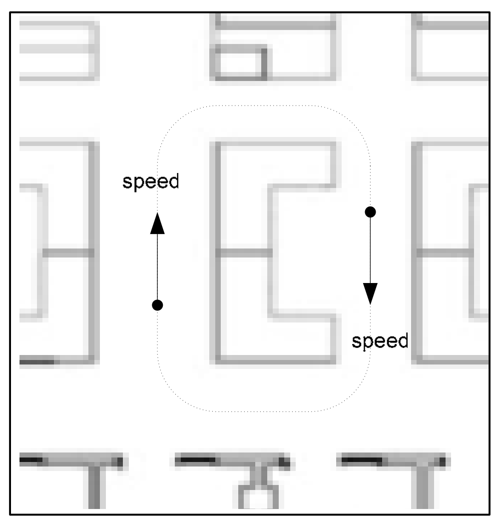
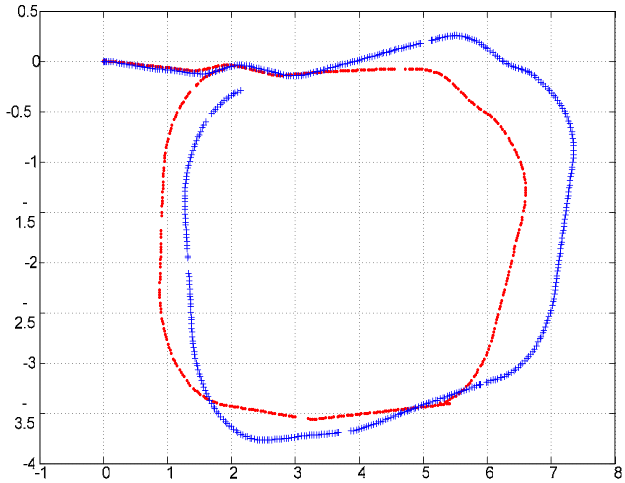
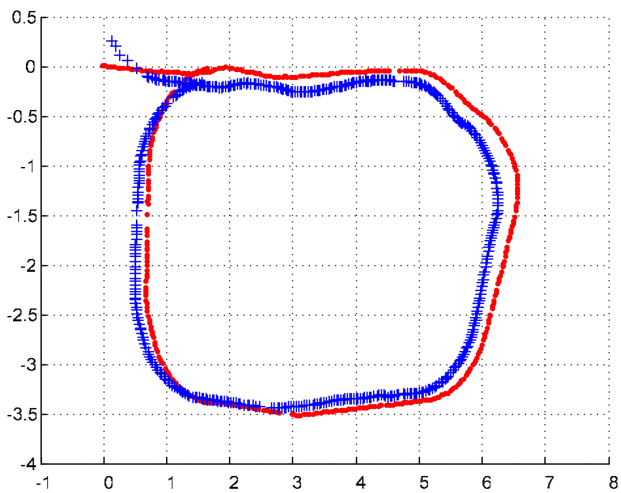


Fig. 13. Map of the office like environment



(a) Odometry (*crosses*) Vs True path (*dots*)



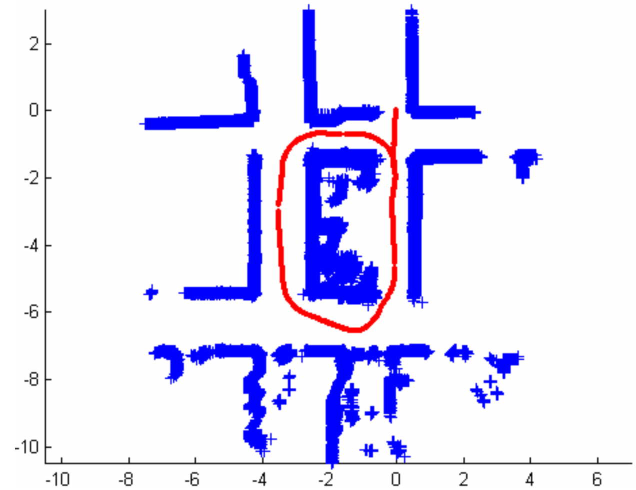
(b) Vision AMCL (*crosses*) Vs True path (*dots*)

Fig.14. Robot localization results

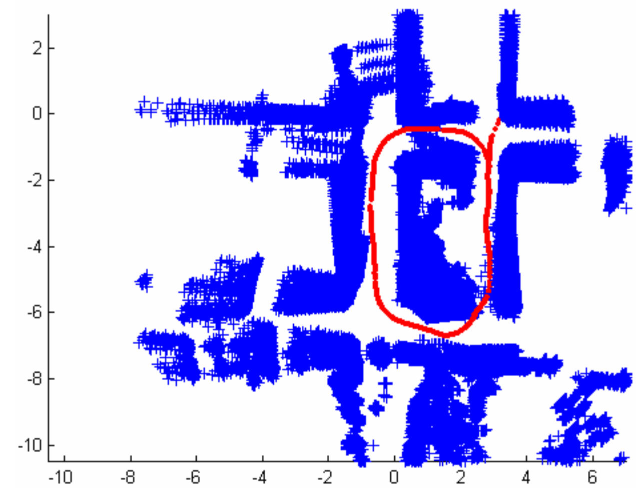
4.2 Map generation

With the availability of laser data it is possible to generate a map looking scans as shown in Fig. 15. Assuming the laser data to be accurate, it is possible to qualitatively assess the localizer performance by generating such maps. Ideally, map generated with true path should have perfectly aligned scans. However, due to the uncertainty of the laser data, a little deviation can be seen. Fig. 15 (a) shows the true path and plotted laser scans to generate a map. It could be noted the 90 deg corners and parallel lines, which are qualitative assessing criteria of the correctness of the map. Fig. 15 (b) shows the scan map generated using laser data and monocular vision AMCL. It shows the acceptability of monocular vision based localizer.

The localization errors of vision AMCL in x, y and orientation are shown in Fig. 16.

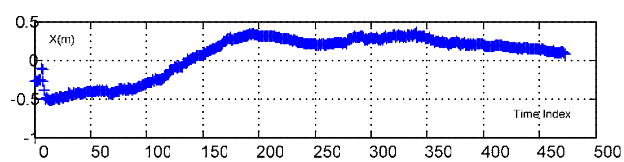


(a) Map generated by true path and laser data

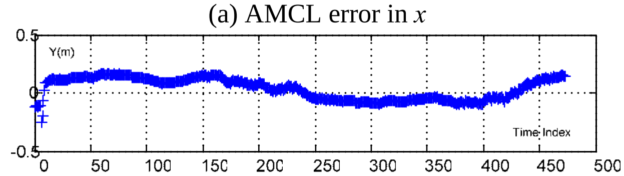


(b) Map generated by vision AMCL and laser data

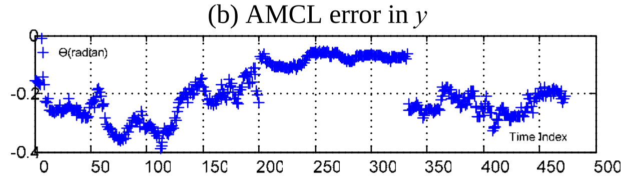
Fig. 15 Mapping Results: Laser based map(*crosses*), Robot path (*dots*)



(a) AMCL error in x



(b) AMCL error in y



(c) AMCL error in orientation

Fig. 16 Error plots

5 Conclusion

With the increase in demand for indoor type robots, low cost sensing and computation are essential for feasible cost effective robot solutions. In this paper, we made an attempt to utilize a single camera for robot localization. We utilized colour based segmentation of the floor as a feature for localization. Once the floor boundary is segmented, a laser-like range/bearing data is generated. It is noted that the data generated is superior to stereo vision based laser-like scan. The laser-like scan is fed to an AMCL for robot localization. Experiments were carried out in an office like environment. The vision AMCL localization results were compared with a true path generated by laser based ICP algorithm. Further, it is qualitatively analysed by generating maps using vision AMCL robot pose with laser scans.

One of the shortcomings of using colour for floor segmentation is the less robustness with the introduction of multi colour floors, for example tiled floors. However, we are currently working on incorporating other cues such as texture, edges and geometry to overcome such deficiencies. For example, first a small region of interest closer to the robot can be analysed for colour or texture. Depending on the sparsity of data the best cue for that particular floor can be determined. We are also in the process of implementing the algorithm on an Amigobot® equipped with a webcam.

Acknowledgements

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

References

[Frank et al., 1999] Frank Dellaert, Wolfram Burgard, Dieter Fox, Sebastian Thrun. Using the CONDENSATION algorithm for robust, vision-based mobile robot localization. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Volume 2, 23-25 June 1999.

[Fox, 2003] Dieter Fox. Adapting the sample size in particle filters through KLD-Sampling. *The International Journal of Robotics Research*. Vol. 22, No. 12, 985-1003, 2003.

[Juan and Francisco, 2000] Juan Manuel Sáez and Francisco Escolano. Monte Carlo Localization in 3D maps using stereo vision. *Proceedings of the 8th Ibero-American Conference on AI: Advances in Artificial Intelligence*, pages 913-922, 2000.

[Menegatti et al., 2004] E. Menegatti, A. Pretto, and E. Pagello. Testing omnidirectional vision-based Monte-Carlo Localizaion under occlusion. *Proceedings of 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Japan, 2004.

[Nguyen et al., 2004] Nguyen Xuan Dao, Bum-Jae You, Sang-Rok Oh, and Young Jin Choi. Simple Visual Self-Localization for Indoor Mobile Robots using Single Video Camera. *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems*. pages, 3767-3772, 2004

[Player Stage web-site.] Avaiable as a hypertest document at: : http://playerstage.sourceforge.net/doc, last check Sept, 2005

[Sebastian et al., 2001] Sebastian Thrun, Dieter Fox, Wolfram Burgard, Frank Dellaert. Robust Monte Carlo localization for mobile robots. *Artificial Intelligence*. pages 99-141, 2001.

[Tim, 2002] Tim Bailey, Mobile Robot Localisation and Mapping in Extensive Outdoor Environments, *PHD thesis*, university of Sydney, Sydney, Australia, 2002