Proceedings of the First

Euromicro Workshop on Advanced Mobile Robots (EUROBOT '96)

Kaiserslautern, Germany
October 9–11, 1996

Sponsored by Deutsche Forschungsgemeinschaft (DFG)

IEEE Computer Society
50 Years of Service • 1946-1996

The Institute of Electrical and Electronics Engineers, Inc.

Using Colour Information in an Omnidirectional Perception System for Autonomous Robot Localization

R. Cassinis, D. Grana, A. Rizzi

University of Brescia
Department of Electronics for Automation
Via Branze, 38 I-25123 Brescia, Italy
Phone:+39 - 30 - 3715.453; Fax:+39 - 30 - 380014
e-mail: cassinis@bsing.ing.unibs.it

Abstract

A localization system for mobile robots is presented. The system enables an autonomous mobile robot to locate itself along previously learned routes in a dynamic environment. The localization is based on data provided by an omnidirectional visual perception system employing optical pre-processing and neural network learning. New results have been achieved using colour information.

1. Introduction

This paper focuses on enhancements and tests that are being carried on the POLLICINO system [1], developed at the Department of Electronics for Automation, University of Brescia (Italy). The system is derived from the COPIS system presented by Yagi et al. [2][3].

In previous experimentation POLLICINO was provided with a B/W CCD sensor [4]. Thus, the only kind of data used for image processing and neural network training was the brightness, scaled on the available grey-level range. With a colour CCD we can obtain more information about the environment, that can be used in different ways according to the neural network structure.

Work is still in progress, but some significant results have been already achieved. This paper mainly focuses on the problem of evaluating colour images obtained by the optical system, with the aim of improving self-localization skills.

2. Aims and features of POLLICINO

The aim of the system is the localization of a mobile robot moving autonomously in a working area in which it has been previously trained. The system operation involves two phases: the supervised learning of the working areas and the autonomous navigation of the robot in the learned areas. The localization system is devised to take advantage of natural, pre-existent fixed reference objects in the environment (walls, openings…), without requiring their explicit identification.

Another feature of the localization system is its ability to work even if the learned environment changes. This holds true, provided that changes affect only limited portions of the whole omnidirectional vision field (e.g. people walking around, objects theat were not present in the training phase…).

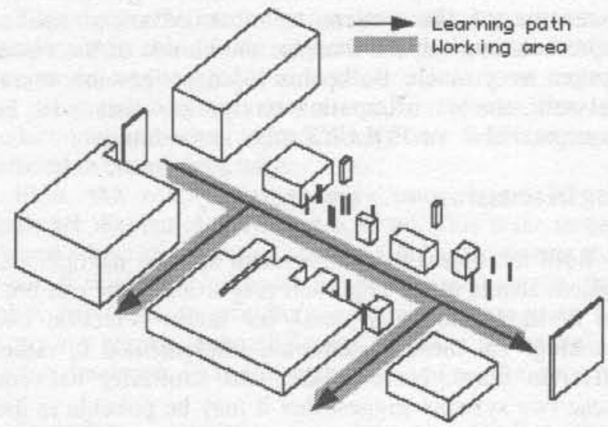

Fig. 1. Working area and learning paths.

3. System structure

The localization system is composed of the following sub-systems:

- an omnidirectional visual perception system (CCD camera + conical mirror);
- an image pre-processing system;
- a learning system based on a neural network.

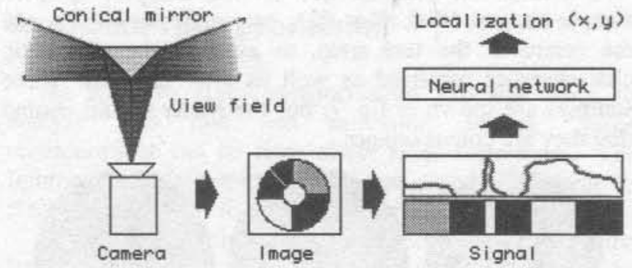

Fig. 2. Structure of the system.

The perception system generates an image of the omnidirectional view of a section of the environment around the robot. It is composed of a CCD camera facing upwards, which records the image of a cone-shaped mirror, placed at a known distance, coaxially to the camera lens.

The conical mirror does not have to be a perfect mirror, because the perception system is not designed to obtain a perfect image of the environment. The large amount of image detail given by a perfect mirror would increase the image elaboration complexity. As high spatial frequencies are less influent on the performance of the learning system, it is convenient to enhance the image blur. Blurring acts as a low-pass filter and is responsible for the loss of information related to those objects in the environment that are small or far away from the robot.

To obtain a blurring effect on the image we make use of the conical mirror surface roughness, instead of applying diffusion algorithms on the pixel matrix of the image, because one of the effects originated by the surface roughness is an increased diffuse reflection component [5].

Images obtained by the perception system are further pre-processed to simplify and enhance useful information that is used, during the learning phase, to train the neural network.

In the execution phase, the mobile robot attempts to follow the previously learned path. The localization system takes advantage of the information organised and stored in the neural network to obtain an approximation of the actual robot position in a relatively small subsection of the whole path.

Only the perception system (CCD camera + conical mirror) has so far been mounted on POLLICINO, while the image pre-processing and the neural network training were performed off-line. The images were grabbed and pre-processed on Silicon Graphics workstations. The neural network has been simulated with the SNNS simulator (vers. 4.1).

The tests on the system, limited to only a subsection of a longer path, were performed in the following way. As a first step, images were taken corresponding to known position given by a sampling grid of the work area. Then, the input–output patterns for the neural network were obtained by pre-processing the images and associating them with the corresponding coordinates.

Only a subset of these data was used to train the neural network. Data not used for training were then used to check the system ability in self-localization by measuring the mean localization error between the output of the neural network and the actual position.

4. Image characteristics

The image of the environment, as it is reflected by the conical mirror, has a typical circular structure with many different bright and/or coloured sectors. They collect, in a single omnidirectional image, information about the angular position of objects in the surrounding environment.

Moreover, the blurring effect and the shape of the mirror make the vertical dimension collapse along each ray of the image.

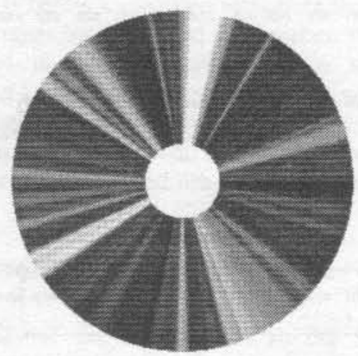

Fig. 3. A typical environment image from the conical mirror.

A reference system transformation from rectangular to polar coordinates is applied to this image, in order to simplify further processing. Pixels around the centre of the image are discarded from the pre-processing, avoiding problems due to the presence of a great number of sectors in a relatively small area.

For this purpose, a selected ring of the circular image is mapped into a rectangle. Each vertical stripe of the rectangle contains all the pixels representing the chromatic characteristics perceived in the related direction.

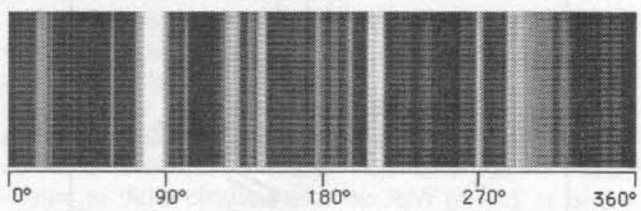

Fig. 4. Image obtained transforming the ring of Fig. 3.

Selected chromatic characteristics of all the pixels in each vertical stripe are extracted with this transformation, greatly reducing image size while preserving useful omnidirectional information about the environment [6].

A test environment was set up to test the results obtainable with different configurations of the system varying colour space mapping, image pre-processing and network structures.

5. Description of the scene and of perceived images

A test scene containing some coloured patches was used to perform tests. Thirty images were taken in a rectangular test area, 3 cm apart from each other, along three asymmetric lines, as shown in Fig. 5. The sampling grid dimensions were chosen considering the proximity of the surrounding patches in order to have a scale corrispondance between the test scene and a possible indoor environment.

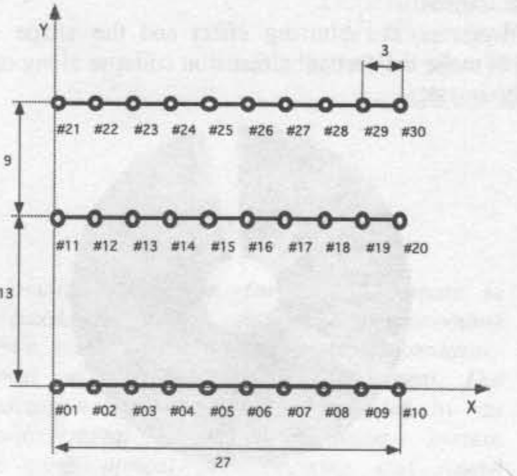

Fig. 5. Sampling grid. Lengths are in cm.

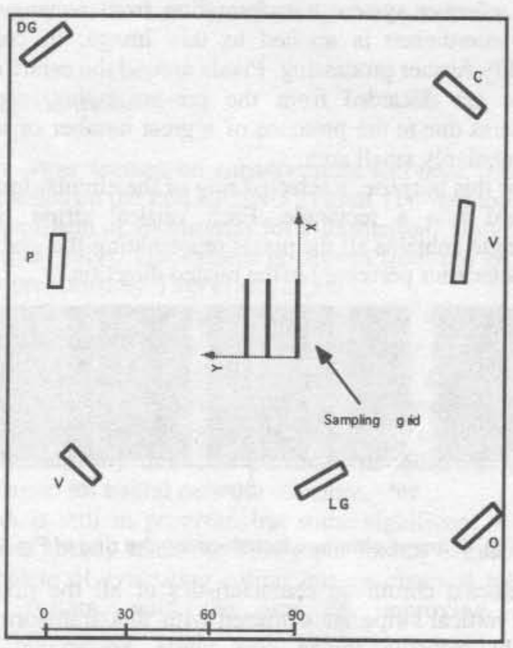

Fig. 6. Disposition of objects in the experimental setup. Letters indicate colours:
V = violet; DG = deep green; LG = light green;
P = polychromatic; O = ochre ; C = cyan.
The background wall colour is beige. Lengths are in cm.

Our attention then focused on the analysis of three sample images, (#10, #16, #21: two opposite corners and the centre of the test area), to evaluate the chromatic characteristics perceived as well as their changes. These samples are shown in fig. 7, but the reader should remind that they are colour images.

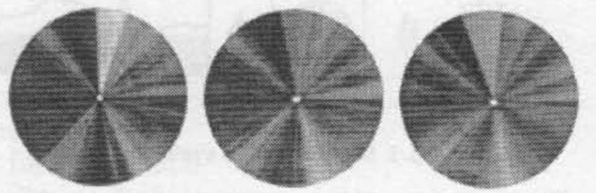

Fig. 7a, 7b, 7c. Sample colour images #10, #16, #21.

R, G, B histograms of the sample images are shown in fig. 8. Pixel values range around the middle of the scale. This is due to the greyness of the mirror surface and to the diffused lighting of the scene.

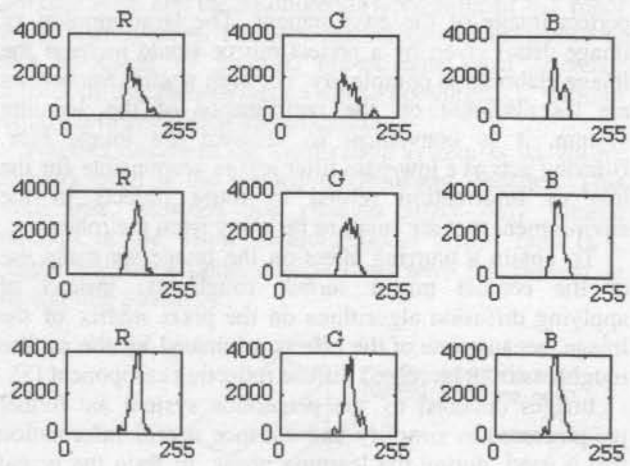

Fig. 8a, 8b, 8c. Sample image #10, #16, #21 RGB histograms.

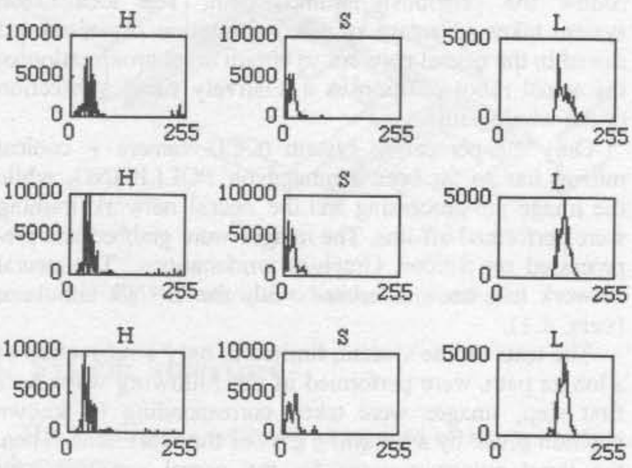

Fig. 9a, 9b, 9c. Sample image #10, #16, #21 HSL histograms.

6. Colour pre-processing

An RGB to HSL transformation [7] was made to analyse also hue and saturation values. Some considerations can be done about HSL histograms. The lightness (L) histograms (fig. 9) are scaled accordingly to the RGB ones (fig. 8) to make comparisons easier.

Saturation (S) values are very low and similar in all the images. This is due to the reflectance characteristics of the cone surface. On the contrary, hue (H) appears to be subject to strong variations of the heights of the sharp peaks and of their relative distance.

The images were transformed from rectangular to polar coordinates with an angular resolution of 1 degree, as shown in fig. 4. For each chromatic component and angular sector considered, we obtained the following data: number of pixels involved, minimum, maximum, mean, most frequent and second most frequent values.

Selected data for each sector and chromatic component were used to form an input array of 360 values, each one corresponding to a 1° angular sector. This array was fed to the network input layer, which was also composed of 360 input units.

For R, G, B, S and L components the sector mean value was chosen to represent each sector. Concerning the H component, more than one peak was expected in each sector, due to the proximity of different coloured objects. On the contrary, the analysis of collected data revealed a distribution of the H values in each sector having only one peak. Thus the H most frequent value was chosen.

7. The learning system

The learning system is a feed-forward neural network, trained with a back-propagation algorithm [8]. The network used, shown in fig. 9, has two hidden layers and an output layer consisting of two output units that encode the position of the robot in the chosen reference system (i.e. as X, Y coordinates).

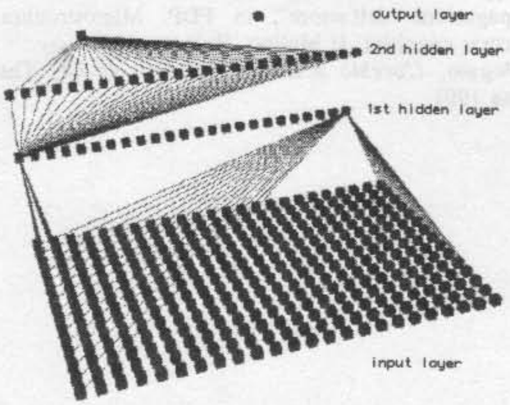

Fig. 9. The structure of the neural network.

Each hidden layer is completely connected to the following layer and is composed of units with a sigmoidal activation function. The total number of units in the network is 418, while the total number of links is 1480.

Each unit in the first hidden layer gets its inputs only from a group of units of the input layer, thus forming a cluster. Moreover, input clusters are partially superposed. This improves the robustness of the system. In this way, small pattern rotations or moderate changes in the perceived environment can be tolerated [9].

A set of images grabbed at each grid position is collected. Data obtained by the pre-processing of only a subset of these images are used for neural network training. The training stops when the error between network output and expected output (X and Y image coordinates) is below a specified minimum threshold.

All the images are then used to simulate the wandering of the robot in the test area. To do so, the data corresponding to each image are fed to the network input layer and the difference between network output and expected output is collected. The mean and maximum value of the results obtained are considered as performance indicators.

8. Test results

The density of the training set in the sampled grid influences the performance of the system. This paper does not focus on this problem, dealing mainly with the evaluation of the system performance with different kinds of chromatic input, but tests were made to determine a maximum sampling grid spacing that allows an acceptable localization error. Three sampling grid spacings were considered, obtaining low, medium and high density learning sets containing respectively 9, 12 or 15 images out of the 30 images in the training set. The pilot simulation runs, made to determine correct values for network parameters, have shown that the network learns the patterns corresponding to the image in the learning set in a small number of iterations.

The network has been tested with the following chromatic inputs: R, G, B, H, S, L. For each component and density, the following data on localization error were collected over all the images in the test set: minimum (Min), maximum (Max), average (Avg) and standard deviation (Dev). R, G and L tests have shown results similar to those obtained with the B/W images of earlier works. The average localization error, related to the length of the path, was around 10%.

Better results were achieved with S or B data and high density sampling, as can be seen in tab. 1 and fig. 10, obtaining an average localization error limited to 5% of the length of test area, while the maximum localization error was around 20%. Similar results were also obtained with S and B medium density sampling.

	Avg			Dev		
	Low	Mid	High	Low	Mid	High
R	43.7	41.26	32.84	30.23	23.23	23.3
G	42.67	33.13	34.41	35.26	17.54	28.79
B	33.04	22.1	24.16	15.5	13.46	14.59
H	84.01	71.35	77.63	50.12	31.36	57.59
S	38.77	34.09	27.27	35.19	22.45	14.84
L	41.52	37.82	31.75	25.69	18.63	24.78

Tab. 1. Test results. Localization errors in mm.

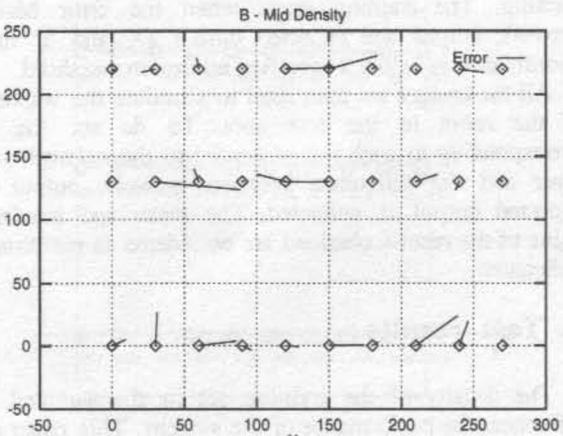

Fig. 10. Blue mid density test results. X and Y are in mm.

8. Sensitivity to robot rotation

The system does not measure the robot rotation in the fixed reference system defined in the learning phase. Tests were made to investigate the sensitivity of the system to rotations.

As it can be seen in tab. 2, small rotations up to about four degrees are well tolerated. Higher rotations could be managed if the robot keeps track of its heading in the reference system. If the robot heading is known, then a simple circular shift of the pre-processing output data could take into account the rotation effects.

Rot.	0°	2°	4°	6°	8°	10°
Avg.	14.78	19.81	30.04	40.47	52.14	64.91
Dev.	14.09	12.36	16.42	22.14	29.63	38.40

Tab. 2. Rotation test results. Localization errors in mm.

9. Perspectives

As it can be seen in the results, the use of different chromatic components from the same image set leads to different system performances. But there are no reasons to think that those performances should be strictly related with any particular chromatic channel. On the contrary, we expect better results from the joined use of more than one chromatic component.

In planned experimentations other topics will be focused. The system performance on different scales will be tested on images obtained by varying the sampling grid spacing. Also the network structure will be changed to take advantage of more than one chromatic component at the same time. The system behaviour is now being tested in noisy conditions simulating moving objects, occlusions and changes in lighting condition. Other test environments will be used in order to test the system performance under different real-world conditions.

10. Bibliography

[1] R. Cassinis, D. Grana and A. Rizzi, "A perception system for mobile robot localization", in WAAP95 Workshop congiunto su Apprendimento Automatico e Percezione, Ancona (Italy) 22/23 June 1995

[2] Y. Yagi and M. Yachida, "Real-time generation of environment map and obstacle avoidance using omnidirectional image sensor with conic mirror", in *Proceedings of the IEEE Conference on Computer Vision & Pattern*, pp 160-165, 1991

[3] Y. Yagi, Y. Nishizawa and M. Yashida, "Map based navigation of a mobile robot using omnidirectional vision system COPIS", in *Proceedings of the IEEE Int. Conference on Robotics & Automation,* Nice, France, may 1992

[4] E. Bideaux, P. Baptiste, C. Day and D. Harwood, "Mapping with a backpropagation neural network", in *Proceedings of the IMACS-IEEE/SMC Symposium,* Lille, France, April 1994

[5] S.K. Nayar, K. Ikeuchi and T. Kanade, "Surface reflection: physical and geometrical perspectives", in *IEEE Trans. Pattern Anal. Machine Intell.* , vol. 13, 7, Luglio 1991

[6] A.K. Jain, *Fundamentals of digital image processing*, pp. 418-420, Prentice Hall, Englewood Cliffs 1989

[7] G. Wyszecki and W.S. Stiles, *Color Science*, Wiley, New York 1982

[8] D.E. Rumelhart and J.L. McClelland, "Un quadro generale del processamento distribuito in parallelo", "L'apprendimento di rappresentazioni interne attraverso la propagazione dell'errore", in PDP. Microstruttura dei processi cognitivi, Il Mulino, Bologna 1991

[9] T. Poggio, *L'occhio e il cervello*, pp. 44-45, Theoria, Roma 1991