Associazione Italiana per
l'Intelligenza Artificiale

International Association for Pattern
Recognition – Italian Chapter

Atti della Giornata di Lavoro dell'Associazione Italiana
per l'Intelligenza Artificiale (AI*IA) e dell'International Association
for Pattern Recognition - Italian Chapter (IAPR-IC)

*Proceedings of Joint Workshop of the Italian Association for
Artificial Intelligence (AI*IA) and the International Association
for Pattern Recognition - Italian Chapter (IAPR-IC)*

IAPRVA98

TECNICHE DI INTELLIGENZA ARTIFICIALE E PATTERN RECOGNITION PER LA VISIONE ARTIFICIALE

*Artificial Intelligence and Pattern Recognition Techniques
for Computer Vision*

6–7 Aprile 1998

Università degli Studi di Ferrara - Facoltà di Ingegneria

A Biologically-Inspired Visual Homing Method for Robots

R. Cassinis, A. Rizzi, G. Bianco, N. Adami, P. Mosna

Dept. of Electronics for Automation - University of Brescia
Via Branze 38, I-25123 Brescia - Italy
Phone +39-30-3715.457
E-mail: { cassinis, rizzi, bianco }@bsing.ing.unibs.it

Abstract

This paper presents an algorithm that uses only visual information to drive an autonomous robot towards a previously visited place. The proposed mechanism, called "homing", is inspired by the behaviour of social insects, like bees and it is used for a precise positioning in the latter phase of robot navigation.

The algorithm uses an affine model and a best matching criteria, on color images, to estimate the robot distance from the goal position. It has been implemented and tested in an indoor environment and results are presented.

1 Introduction

A lot of research has been done about insect homing strategies and several interesting models have been developed [4, 5, 1, 8, 6], but few of them have been applied to real robotic navigation with color images. This paper presents a homing method inspired by bee's behaviour.

A robot navigation task is normally composed of two subtask: a coarse approach to the target and a precise positioning to it. In this paper the latter phase, called "homing", will be investigated.

The described homing method can be divided in two phases. In the first matching phase the perceived image is compared with goal image in order to estimate the parameters of the affine model and in the second navigation phase from these parameters is computed a navigation vector to drive the robot. The matching and the navigation phases have been implemented and tested.

2 Matching phase

In order to drive the robot to the goal[1], comparisons are made between what the robot "sees" from its actual position and what it "saw" when it was standing at the goal position. In fact the robot displacement from the goal position results in a shift and in a variation of the object size in the actual image that can be described using a simple affine model [2]:

$$\begin{cases} S_X(X,Y) = a_{0X} + a_{1X} \cdot X + a_{2X} \cdot Y \\ S_Y(X,Y) = a_{0Y} + a_{1Y} \cdot X + a_{2Y} \cdot Y \end{cases}$$

[1] that must have been already reached at least once

where S_X and S_Y are the displacement components in x and y directions respectively, (X, Y) are the pixel co-ordinates, a_{0X}, a_{0Y} are the translation parameters and a_{1X}, a_{2X}, a_{1Y}, a_{2Y} are the parameters for rotations, compression and deformation respectively.

The camera heading is kept constant in order to eliminate fixed objects rotation and the fixed height of the robot eliminates pure vertical shift.

Thus all the possible changes due to robot movement can be described with this simplified version of the previous model:

$$\begin{cases} S_X(X,Y) = a_{0X} + a_{1X} \cdot X \\ S_Y(X,Y) = a_{2Y} \cdot Y \end{cases}$$

An apparent vertical shift in the image plane can be introduced by the change of object distance and is described by the compression/dilation parameters a_1X and a_2Y.

The parameters that maximize the intercorrelation value between the goal image and the actual one, modified by the model, are used to compute the navigation vector. The maximum intercorrelation value is found minimizing the following Mean Square Error (MSE):

$$MSE = \frac{1}{M} \sum_{(x,y)\in S} E_r(X,Y) + E_g(X,Y) + E_b(X,Y)$$

with

$$E_{r,g,b}(X,Y) = [I_{1(r,g,b)}(X,Y) - I_{2(r,g,b)}(X+S_X, Y+S_Y)]^2$$

where M is the number of couples $(X+S_X, Y+S_Y)$ still in the image plane $\mathcal{S}$, I_1 and I_2 are the images, r, g and b are the chromatic components, S_X and S_Y are the estimate displacement vectors for each pixel.

In order to speed up the parameters estimation and at the same time to allow the estimation of large displacement vectors, a multi resolution pyramidal technique has been implemented as in Wittman [8]. Before each subsampling operation a low pass seven coefficients gaussian filtering is applied in order to exclude possible spatial aliasing [3]. The estimation ranges and the multi resolution levels are chosen according to the maximum displacement considered.

3 Navigation phase

Using a projection model it is possible to estimate a navigation vector. In fact a point in the real world can be linked to its projection in the image plane in the following way [7]. Let p be a point of the environment and P its projection on the camera image plane. The relationship between the two points is:

$$\begin{cases} X = F \cdot \dfrac{x}{z} \\ \\ Y = F \cdot \dfrac{y}{z} \end{cases}$$

where X and Y are the co-ordinates of the point P in the image plane, x, y, z are the co-ordinates of the point p in the 3D space and F is the focal length.

It follows that the navigation vector estimation can be computed as:

$$\Delta_x = K \cdot a_{0X}$$
$$\Delta_z = H \cdot a_{1X}$$

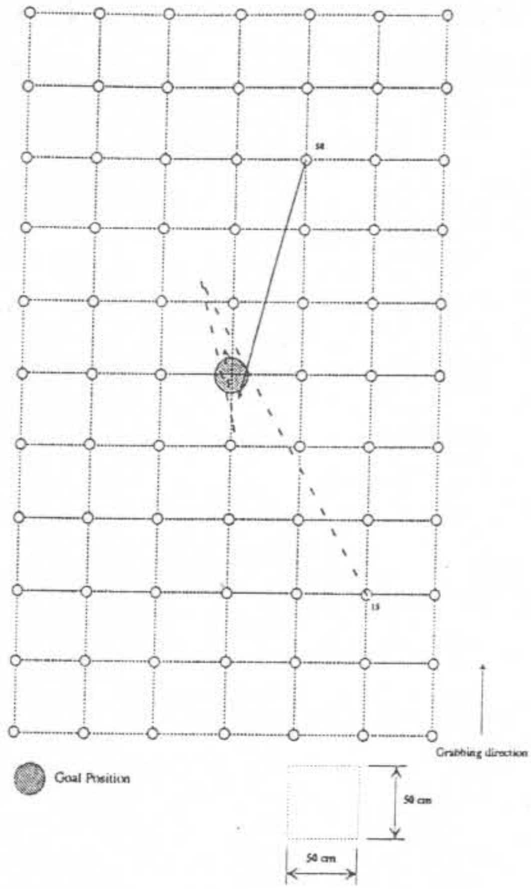

Figure 1: Navigation tests

The values of K and H are determined for a more efficient navigation through an initial calibration phase, consisting of a series of known displacements around the goal position.

Due to the null value of the vertical shift, vertically decimated images of a factor $2k$ have been used.

The noise in the acquired images is substantially due to the following factors: people moving in the room, changes of displays on computer monitors and lighting being switched on and off.

The main error contribution in the matching of two images acquired in the same point derives from the moving objects (people walking, etc.) and occlusions that cause a "punctual" error higher than the average. For this reason in the final matching phase the MSE over the whole image is substituted by a new MSE computed only over the pixels with error lower than the whole MSE previously computed.

4 Tests of the system

The system, tested in an indoor environment, has shown its ability to estimate the robot position with respect to the goal point. The robot driven with the estimated distance vector can move itself toward the goal position in a very limited number of steps and with a very low error, as visible in Fig. 1.

The algorithm has also been tested computing the first step from a wider set of starting points. The results (Fig. 2) show that the algorithm can drive the robot towards the goal point in the major part of the tested area, expecially when the goal point is beyond or around the robot camera. In Fig. 2 only the direction of the navigation vector is plotted.

In Fig. 3 are plotted the estimated target point of the first navigation step. The accuracy in the estimation can suggest the use of this algorithm also for robot self-localisation.

The algorithm has shown different results according to different navigation areas with respect to

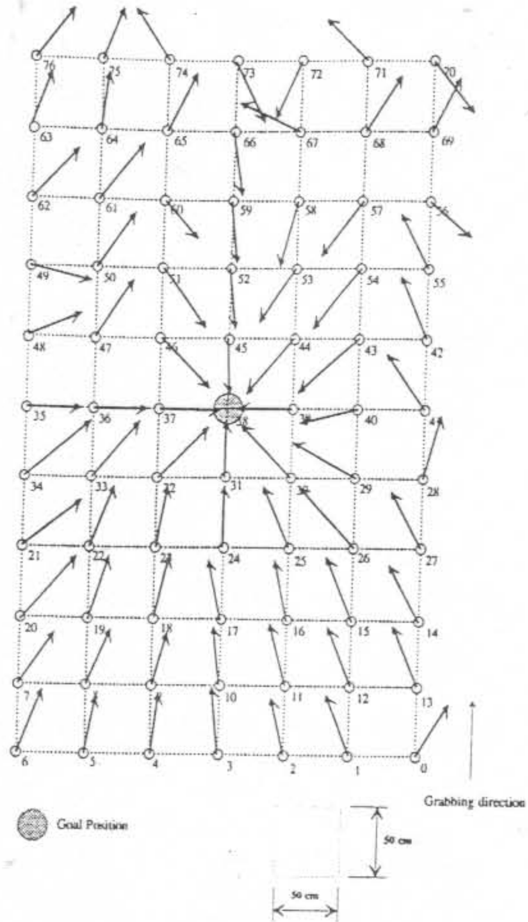

Figure 2: First step tests

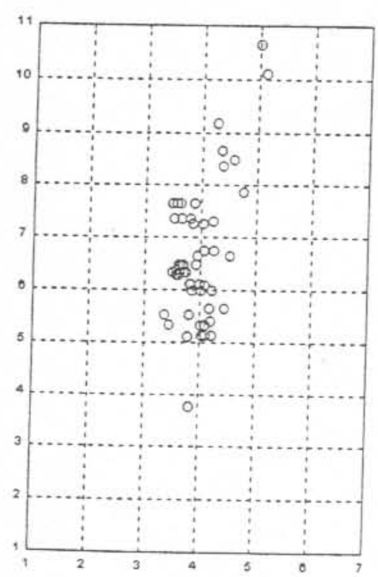

Figure 3: Estimated first step points

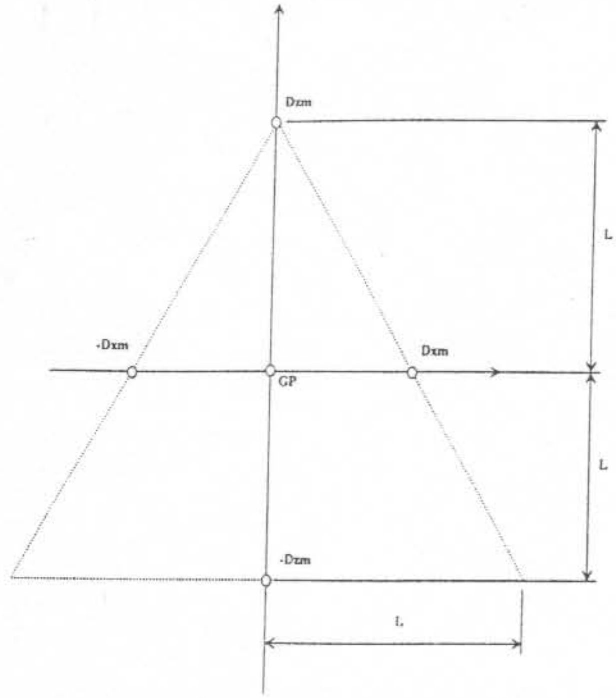

Figure 4: Region of applicability

the same goal position. A reason of this fact could be found investigating the matching algorithm in order to find a "region of applicability" (ROA). This region depends mainly on the distance between the the goal position and the objects in front of the robot camera. This distance influences the parallax angle with the robot visual reference system that the algorithm uses to calculate the goal point position.

In Fig. 4 is shown the generic shape of a ROA where (Dzm,-Dzm) is the maximum distance range with no translation and (Dxm,-Dxm) is the maximum translation range with no distance changing. In this region, except in the case of occlusions, the algorithm has all the visual information to perform correct results. This does not mean that the algorithm fails outside the ROA.

From the following formula, that checks if a point is in the ROA, it can be noticed that also the parameters H and K influence the extension of this region:

$$ROA = \{\langle DX, DZ\rangle \mid -Sxm < \frac{DX}{K} + \frac{DZ}{H} \cdot \frac{C}{2} < Sxm \wedge -Sym < \frac{DZ}{H} \cdot \frac{R}{2} < Sym\}$$

where $\langle DX, DZ\rangle \in R \times R$ are the checked co-ordinates in the environment, Sxm, Sym are the maximum translation and compression, respectively, in the image plane, R, C are the number of row and columns of the image.

The shape of this generic ROA can be easily individuated checking the first step directions in Fig. 2.

5 Conclusions and perspectives

This homing method, intended as final step for robot positioning, originates precise navigation in a small number of steps. It can also be used as module for visual self-localisation.

The influence of noise, robot undetected rotation and color shifting in the illuminant could decrease the algorithm performance and will be investigated.

References

[1] U. Nehmzow. Animal and robot navigation. *Robotics and Autonomous Systems 15*, pages 71–81, 1995.

[2] A.N. Netravali and B.G. Haskell. *Digital Pictures-Representation and Compression.* Plenum Press, New York, 1988.

[3] W.K. Pratt. *Digital Image Processing.* A Wiley interscience publication, New York, 1991.

[4] J. Santos-Victor, G. Sandini, F. Curotto, and S. Garibaldi. Divergent stereo for robot navigation: learning from bees. In *IEEE Computer Society Conference on Computer Vision e Pattern Recognition*, New York, 1993.

[5] J. Santos-Victor, G. Sandini, F. Curotto, and S. Garibaldi. Divergent stereo in autonomous navigation: from bees to robots. *Int. Jour. of Computer Vision 14*, pages 159–177, 1995.

[6] M. V. Srinivasan, S. W. Zhang, M. Lehrer, and T.S. Collett. Honeybee navigation en route to the goal: visual flight control and odometry. *Journal of Experimental Biology 199*, pages 155–162, 1996.

[7] Y.T. Tse and R.L. Baker. Camera zoom/pan estimation and compensation for video compression. *SPIE Image Processing Algorithms and Techniques II, vol. 1452*, pages 468–479, 1991.

[8] T. Wittman. Insect navigation: Models and simulations. Technical report, Zentrum fur Kognitionswissenschaften, University of Bremen, 1995.