Copyright © 2002 IFAC
15th Triennial World Congress, Barcelona, Spain

EXEMPLIFYING THE USE OF VISUAL REFERENCES IN THE NAVIGATION OF MOBILE ROBOTS

Mário Sarcinelli-Filho
Roger Alex de Castro Freitas
Renato de Sousa Dâmaso
Teodiano Freire Bastos-Filho

Departamento de Engenharia Elétrica, Universidade Federal do Espírito Santo
Av. Fernando Ferrari, s/n 29060-900 Vitória/ES BRAZIL

Abstract: This paper presents an experiment on mobile robot navigation based on the recognition of references in its working environment. It is discussed an example in which the mobile robot should recognise an opened door and go through it to get out the lab. The sensorial system used in the experiment is described, including the subsystem in charge of recognising the references from images acquired by a CCD camera onboard the robot. The distributed control system implemented to accomplish the experiment is also addressed. *Copyright © 2002 IFAC*

Keywords: Object recognition; Sensor systems; Image segmentation; Autonomous mobile robots; Computer vision; Computer controlled systems.

1. INTRODUCTION

In many applications, a mobile robot should either to navigate towards a specific point in its working environment or to park in an appropriate site. An example is a transportation vehicle in an automated manufacturing environment, which is expected to dock close to a machine to receive a piece the machine has just produced. Another example is the mobile robot docking for battery charging.

Autonomous robots, for having the capability of navigating autonomously, need some suitable sensors to get precise information about their surrounding environment to allow the control of its navigation and correct docking. Thus, they sometimes need to recognise a landmark as a previous step in making a decision or before starting a new behaviour.

The most sophisticated sensor available onboard a mobile robot is the vision system, which allows getting very rich information of the environment surrounding the robot. When using landmarks, for example, the use of a vision system to perfectly identify a certain landmark is fundamental (Betke and Guwits, 1997; Asensio, *et al.*, 1998; Gaspar, *et al.*, 2000). As an example, in Asensio, *et al.* (1998) a stereo trinocular vision system is used so that the robot can locate a door and go through it. The construction of a 3D model of the environment is implemented from the three camera views, so that the robot can plan its trajectory towards the door.

However, stereo vision systems are too much expensive, besides being computationally intensive, which makes them not suitable for many applications. This way, most vision systems onboard mobile robots are monocular vision systems, once they are cheaper for using only a single CCD camera.

In some special situations, like when the landmarks are bigger (e. g., a corridor) one can use an omnidirectional vision system, in which a single CCD camera in connection with a special mirror gets ground views (Gaspar, *et al.*, 2000). In this case, however, the cost of the vision system is still high, once the mirror used should have a particular geometry to get a bird's eye view of the ground without distortion (Hicks and Bajcsy, 1999). A simplification corresponding to use a different mirror (e. g., a spherical one) could be used, but it would demand an image transformation to eliminate the distortion included in the omnidirectional view (Gaspar, *et al.*, 2000). However, the image-acquiring rate should be decreased in order to account for the additional computation of this transformation.

Thus, the cheapest configuration of the monocular onboard vision system is a single CCD camera directly facing the environment in front of the robot. Sometimes it can have a pan-tilt-zoom movement, but the most common is a fixed mounting (the image plane is a fixed vertical one). This is the case of the mobile robot used in this work, the Pioneer 2 DX mobile platform, which is shown in Figure 1. It only can take 2D views of a narrow area (delimited by a certain angle) in front of it, which is its visual field. This arrangement imposes some limitations on the visual-based control of the robot navigation, because of the loss of depth information and the limitation of the visual field. When necessary, however, the depth information is recovered by using ultrasonic, infrared or laser sensors in connection to the CCD camera.

For the robot used in this paper, as one can see from Figure 1, ultrasonic sensors are available, which will be used in connection to the CCD camera in a sensorial integration scheme known as guidance (Hong, 1999). What has been implemented is a system in which the camera is activated only when the ultrasonic sensors detect an obstacle close to the robot. This connection between camera and ultrasonic sensors is enough for applications in which it is necessary to recognise a landmark, as it is here reported, in spite of the narrow visual field of the robot.

Other techniques based only in the visual information, like the optical flow technique, can also be adopted to control the navigation of the mobile robot, like in Dev, *et al.* (1997) and Soria, *et al.* (2001). However, in this case it is not common to recognise landmarks.

Fig. 1. The mobile robot used in this work.

In this work it is adopted the so-called 2 ½ D vision system (Bastos-Filho, *et al.*, 1999), which is characterised by the use of ultrasonic sensors to guide the action of the camera, as detailed ahead.

For better understanding, the paper is hereinafter structured in four sections. Section 2 describes the sensing apparatus used, as well as how it is integrated in order to implement the 2-½ D vision system. Section 3 describes the image-processing task implemented in order to recognise some references in the environment. Section 4 describes the whole problem, which consists in guiding the robot to cross an open door. The agent-based distributed control structure (Xavier, *et al.*, 1998) implemented to accomplish the task is also presented. Finally, in Section 5, some conclusions are presented.

2. THE SENSING SYSTEM ADOPTED

The platform used in the experiment here implemented posses an external sensing system composed by eight ultrasonic sensors and a single CCD camera. As one can see in Figure 2, six ultrasonic sensors are in the frontal part of the robot (in different angular positions related to the axis of movement), while the two other are in the sides of the platform.

This sensor array is able to detect the presence of obstacles close to the robot, their distance to the robot and their orientation relative to the axis of movement of the robot (Bastos-Filho, *et al.*, 1999). In the present case, the minimal distance in order to consider the presence of an obstacle is 60 cm, relative to any of the frontal ultrasonic sensors. For the two side sensors (angles of ± 90°), this distance is 45 cm. For the robot does not have any rear sensing apparatus, rear manoeuvres should be avoided. However, just small rear translations are considered in this work, to allow positioning the robot when an obstacle is detected.

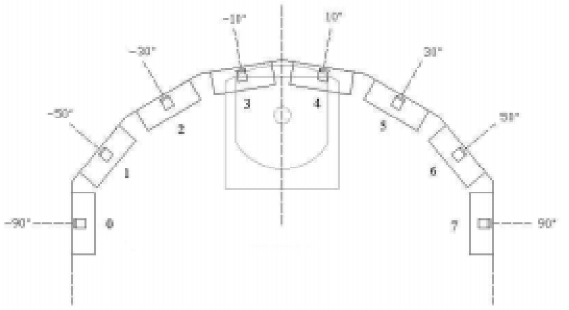


Fig. 2. A sketch (top view) showing ultrasonic sensors and CCD camera positioning.

The CCD camera is in the middle of the robot head, as shown in Figure 2. It is a Sony D31 colour camera, and acquires images as 640x480-pixel bitmaps with 24 bits per pixel. The first step of the image-processing task, therefore, is to transform these images to the 8 bits grey scale. The camera is mounted on a moving support (PTZ) that allows an angular displacement of $\pm 100^{o}$ in the horizontal direction and $\pm 25^{o}$ in the vertical direction. However, for the objectives of this work it will remain fixed in the angles of zero degrees in both directions, with a zoom adjustment of 1X.

The scheme integrating the ultrasonic sensors and the CCD camera in order to recognise the references is quite similar to that proposed in Bastos-Filho, *et al.* (1999). A difference is that in the present case the robot used posses only eight ultrasonic sensors, which correspond to a semicircle in its frontal part, instead of the ring there used (see Figure 2). Another difference is that instead of rotating the camera towards a detected obstacle (Bastos-Filho, *et al.*, 1999), the robot rotates in order to align the camera and the detected object. It should be mentioned that this alignment is necessary in order to get frontal images of the detected obstacles. The integrated sensing system operates in the following way: whenever an object is detected by an ultrasonic sensor the robot rotates around its central axis and then positions itself at a distance of 50 cm from the object. This final positioning is performed by a sequence of frontward and backward small linear displacements. Then, a frontal image of the object is acquired and used to identify it. The system is programmed to recognise table legs, chair legs, walls, corners, edges and doorframes either corresponding to open or closed doors. Any other obstacle is an unknown one, and their presence activates obstacle-avoidance behaviour. However, whenever one of the above mentioned obstacles is detected, it is possible to activate two distinct behaviours: to cross the doorframe when an open door is detected and to avoid the obstacle otherwise. More details on the control system implemented to accomplish the objective proposed are given in Section 4.

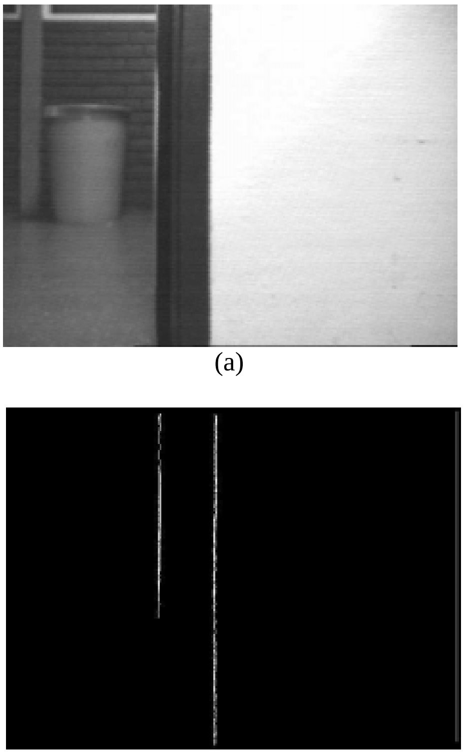
3. RECOGNISING THE REFERENCES

The program generated to perform the object recognition was implemented as a C language project compiled as a dynamical loading library (DLL). This function is called by another program written in Colbert, the dedicated language used to program the robot (ActivMedia Robotics, 1998 and 2000).

The procedure of obstacle identification implemented uses quite simple descriptors for the objects it is expected to identify, which are obtained from an image-segmentation process applied to the acquired image. The first descriptor is the number N of vertical lines present in the binary image obtained by using the same filters used in Bastos-Filho, *et al.* (1999). For instance, a wall generates no vertical lines, while table legs or doorframes generate two vertical lines and an edge or a corner generates just a single vertical line. Thus, the first step is to segment the acquired image in order to define the meaningful vertical lines, where meaningful applies for vertical lines spanning over more than 60% of the image height. The reason for considering only the meaningful vertical lines is to eliminate objects in the image background. Besides the height of the vertical lines itself, close vertical lines are considered as a single vertical line (the minimum distance accepted is 20 pixels).

In addition, to distinguish objects like doorframes, chair legs or table legs, it is necessary to consider the distance D, in pixels, between two adjacent meaningful vertical lines. This distance is a valid descriptor just because the images are always acquired considering a standard camera-object distance, which characterises the 2½-D vision system.

Figure 3 shows the case in which the robot is close to an open door. The original acquired image and the final binary image showing the meaningful vertical lines detected are presented (the obstacle is correctly identified, in this case). The number of vertical lines and the distance between them allows finding that the obstacle is a doorframe. In order to distinguish the open door from the closed door a third descriptor is used: the average grey level in both sides of the outer vertical lines in the original image. These values are calculated considering a horizontal line about the


(a)

(b)

Fig. 3. Recognising an open door. The original image and the result of the image-processing task implemented are shown.

bottom third of the height of the image. The grey level of each pixel on this horizontal line is considered in the two regions, thus resulting in the average values T_1 and T_2. For the example, one can notice from Figure 3 that the average grey level in the right and the left sides are quite close. If the door were closed, however, they would be quite different, once the door is much darker than the background of the left side of Figure 3-a.

After getting the values of N, D, T_1, T_2 and DT (the absolute value of the difference between T_1 and T_2) for the image under analysis, the robot should now determine which object is in the image. For doing that, a kind of rule base is mounted (Bastos-Filho, *et al.*, 1999) regarding these variables, whose output is the most likely objet described by a given set of the input variables. An important observation is that it is necessary to recalibrate such rule base whenever the robot moves to a different working environment.

4. AN ILLUSTRATIVE EXAMPLE

In this section, the operation of the 2-½ D vision system and the control of the robot in order to accomplish a certain task is described. This whole system is implemented as an agent-based control system intended to guide the robot to cross an open door. The robot is programmed to avoid any obstacle while going ahead, but the open door. When recognising the open door it starts a specific behaviour to go through the door opening.

4.1 The Control Structure Implemented

In the above-mentioned illustrative case it is used the behaviour-based control approach implemented as a subsumption architecture (Brooks, 1986). The basic behaviour, which runs until being overcome by a higher level behaviour, is to go ahead with a given linear speed and an angle of zero degrees relatively to the axis of movement. Other higher level behaviours are to avoid obstacles other than the open door, to position the robot whenever an obstacle is detected and to cross the doorframe whenever an open door is recognised.

The sensorial information is the responsible for starting one of these higher level behaviours, which will assume the control of the robot actuators thus suppressing the basic lower level behaviour (Brooks, 1986). In order to process the sensorial information as described in Sections 2 and 3, it is implemented an agent-based structure like described in Xavier, *et al.* (1998). Such structure is shown in Figure 4, where the scheme of agent activation is also presented. In this case, eight primitive sensor agents (Xavier, *et al.*, 1998) $S_0, S_1, \ldots, S_7$, are used, which are associated to the physical ultrasonic sensors available onboard the robot. These eight agents send the information on the distance robot-obstacle, as measured by each ultrasonic sensor, to a virtual sensor agent (Xavier, *et al.*, 1998) called *Ultrasound*. It will arbitrate which sensor has detected the closer obstacle and which is the distance from this sensor to the obstacle.

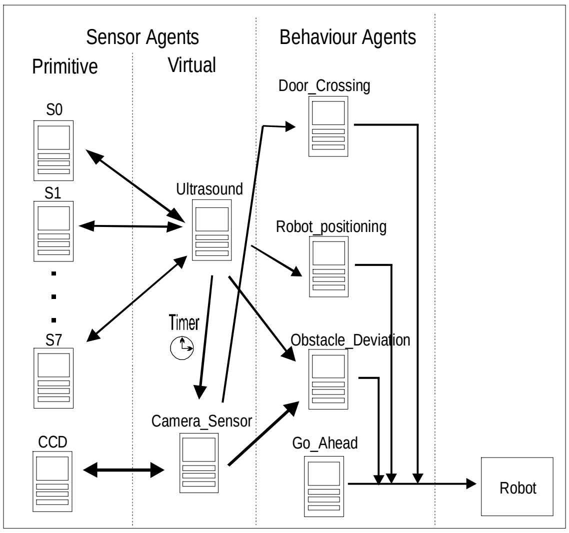


Fig. 4: The agent-based distributed control structure implemented.

Whenever an obstacle is detected, the robot is rotated such that the camera moves towards the obstacle, as well as it moves itself ahead and backwards in order to get stopped about 50 cm far from the obstacle. This behaviour is implemented by the behaviour agent (Xavier, *et al.*, 1998) *Robot_Positioning*. Once the robot is in the right position, the virtual sensor agent *Camera_Sensor*, associated to the CCD camera, is activated. Then, an image of the detected obstacle is got and processed according to Section 3, thus defining which obstacle was detected. A time delay is used when activating it, relative to the activation of the behaviour agent *Robot_Positioning*, in order to guarantee the right positioning of the robot before acquiring the image.

If the obstacle detected is an open door, the behaviour agent *Door_Crossing* is activated, which will guide the robot through the doorframe. Otherwise it is activated the behaviour agent *Obstacle_Deviation*, such that the robot avoid the obstacle and continues navigating.

If no obstacle is detected closer than 60 cm to the robot the active behaviour continues being to go ahead, the same behaviour activated when the agents *Obstacle_Deviation*, *Robot_Positioning* and *Door_Crossing* are deactivated. It corresponds to the behaviour agent *Go_Ahead* shown in Figure 4.

4.2 The Strategy for Door Crossing

After identifying the open door, the robot starts the task of going through it. The behaviour corresponding to this crossing was implemented through direct commands for motion control available in the programming environment dedicated to the Pioneer 2 DX platform (ActivMedia Robotics, 1998). The action of the robot is based on the orientation, alignment and monitoring of the minimum values of the distance from the robot to the doorframe or other obstacle during the crossing. As the minimum limit of distance is a safety parameter, it has a higher priority.

In order to guide the rotation of the robot around its central axis and its translation, the eight ultrasonic sensors shown in Figure 2 are the unique sensing apparatus used. The sensors numbered as 3 and 4 are used for the robot orientation, when necessary, in the direction of the door opening. As the minimum distance continues being monitored, these sensors are able to avoid frontal collisions. After the correct alignment of the robot in the direction of the door opening, the sensors numbered as 2 and 5 are used for small corrections in the robot trajectory, with the objective of avoiding side collisions. Once the robot approaches the door opening the sensors 2 and/or 5 will also cross it, thus changing their distance measure from about a few tens of centimetres to many tens of centimetres. Then, their action will be repeated now using the sensors positioned in the angles of ±50^{o}, and after the sensors positioned in ±90^{o}, until the robot concludes the door crossing. The minimum distance measured by the sensors 0 to 2 and 5 to 7 continue being monitored, thus avoiding side collisions. The sequence of actions thus described is illustrated in Figure 5.

Finally, Figure 6 shows a sequence of image frames taken from a video clip, which illustrates the experiment step in which the robot is close to the open door and some table leg. As it is shown, the system here implemented was effectively able to guide the robot to accomplish the task proposed. However, it is worthwhile to mention that the final trajectory followed by the robot since its initial point till to identify the open door and to cross it is not optimised, which is due to the behaviour-based control approach. One should remember that there is no path planning associated to this control approach: its navigation is totally reactive, where the decision-makings are guided by the sensorial information then available.

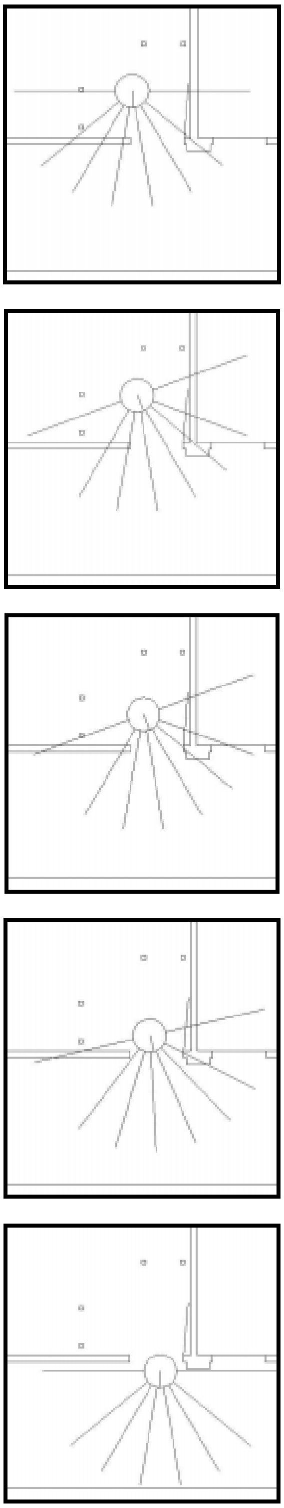


Fig. 5. The sequence of actions illustrating the strategy to guide the robot through the door opening (the radial lines are the angles of the ultrasonic sensors in the robot).

CONCLUSION

This work reported an application in which a distributed agent-based control structure is used to guide a mobile robot starting from a certain position in the lab until going out of it through an open door. The robot goes ahead inside the lab until detecting the presence of an obstacle. An object recognition system is implemented as part of the whole control system, which gives the robot the capability of recognising some obstacles, including the open door. Upon detecting the open door, the robot goes through it, using a specific strategy for nearness navigation here proposed.

Fig. 6. The open door crossing.

The whole control system implemented effectively guided the robot to accomplish the programmed task, although the final trajectory followed is not an optimal one, which was expected as a consequence of the use of the behaviour-based control approach.

ACKNOWLEDGEMENT

The authors thank CNPq – National Council of Scientific and Technological Development, a Brazilian governmental institution promoting the scientific and technological development for granting this work.

REFERENCES

Betke, M. and L. Gurvits (1997). Mobile robot localization using landmarks. *IEEE Transactions on Robotics and Automation*, **13**, 251-263.

Asensio, J.R, J.M.M. Montiel and L. Montano (1998). Navigation among obstacles by the cooperation of a trinocular stereo vision system and laser range finder. In: *Proceedings of the 3rd IFAC Symposium on Intelligent Autonomous Vehicles*, 2, 456-461. Madrid, Spain.

Gaspar, J., N. Winters and J. Santos-Victor (2000). Vision-based navigation and environmental representations with an omnidirectional camera. *IEEE Transactions on Robotics and Automation*, **16**, 890-898.

Hicks, A. and R. Bajcsy (1999). Reflective surfaces as computational sensors. In: Proceedings of the *IEEE Workshop on Perception for Mobile Agents, IEEE International Conference on Computer Vision and Pattern Recognition*, 82-86.

Hong, L. (1999). Sense your world better: multi-sensor/information fusion. *IEEE Journal of Circuits and Systems*, **10**, 7-8, 12-15, 28.

Dev, A., B.J.A. Krose and F.C.A. Groen (1997). Heading direction for a mobile robot from optical flow. In: *Proceedings of the IEEE International Conference on Robotics and Automation*, 558-563. Leuven, Belgium.

Soria, C.M, M. Sarcinelli-Filho, T.F. Bastos-Filho, R. Carelli and O. Nasisi (2001). Obstacle detection and avoidance using optic flow. *Proceedings of the 5th World Multi-Conference on Systemics, Cybernetics and Informatics*, vol. IX, 66-71. Orlando, FL, USA.

Bastos Filho, T. F., M. Sarcinelli-Filho and R. A. C. Freitas (1999). A multi-sensorial integration scheme to help mobile robot navigation through obstacle recognition. *Proceedings of the 7th IEEE International Conference on Emerging Technologies and Factory Automation*, 1, 549-558. Barcelona, Spain.

Xavier, J.E.M and H.A. Schneebeli (1998). A framework based on concurrent object-oriented programming for building behaviour-based control systems for mobile robots. *Journal of the Brazilian Computer Society*, **4**, 27-34.

Brooks, R.A (1986). A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, **RA-2**, 14-23.

ActivMedia Robotics (2000). *Pioneer 2 Mobile Robots Operations Manual*. Peterborough, NH, USA.

ActivMedia Robotics (1998). *Saphira Software Manual Version 6.1e*. Peterborough, NH, USA.