LNCS 3617

Fabio Roli
Sergio Vitulano (Eds.)

Image Analysis and Processing – ICIAP 2005

13th International Conference
Cagliari, Italy, September 2005
Proceedings

6 – 8 septiembre

ICIAP 2005

Springer

Volume Editors

Fabio Roli
University of Cagliari
Department of Electrical and Electronic Engineering
Piazza d'Armi, 09123 Cagliari, Italy
E-mail: roli@diee.unica.it

Sergio Vitulano
University of Cagliari
Department of Medical Sciences
Via S. Giorgio 12, 09124 Cagliari, Italy
E-mail: vitulano@pacs.unica.it

Library of Congress Control Number: 2005931520

CR Subject Classification (1998): I.4, I.5, I.3.5, I.2.10, I.2.6, F.2.2

ISSN 0302-9743
ISBN-10 3-540-28869-4 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-28869-5 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media

springeronline.com

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 11553595 06/3142 5 4 3 2 1 0

Study of the Navigation Parameters in Appearance-Based Navigation of a Mobile Robot

Luis Payá, Oscar Reinoso, Arturo Gil, Nicolás García, and Maria Asunción Vicente

Departamento de Ingeniería de Sistemas Industriales,
Miguel Hernández University, Avda. De la Universidad s/n,
03202 Elche (Alicante), Spain
{lpaya, o.reinoso, arturo.gil, nicolas.garcia, suni}@umh.es
http://lorca.umh.es/isa/es

Abstract. Recently, appearance-based approaches have attracted the interests of computer vision researchers. Based on this idea, an appearance-based navigation method using the View-Sequenced Route-Representation model is proposed. A couple of parallel cameras is used to take low-resolution fontal images along the route to follow. Then, zero mean cross correlation is used as image comparison criterion to perform auto-location and control of the robot using only visual information. Besides, a sensibility analysis of the navigation parameters has been carried out to try to optimize the accuracy and the speed in route following. An exhaustive number of experiments using a 4-wheel drive robot with synchronous drive kinematics have been carried out.

1 Introduction

There are three families of techniques to carry out the navigation of a mobile robot in indoor environments and without a previous map: model-based techniques [3], [4], appearance-based techniques [8] and optical flow-based techniques [11]. Optical-flow based techniques take into account the relative movement of the scene elements to calculate the control law for the robot. Model-based techniques use natural or artificial landmarks from the scene as references to guide the robot through the desired route. The recognition of patterns is achieved comparing features of the input image with features that have been previously stored. These techniques suppose high complexity due to the difficulty in the features extraction and the comparison of patterns in realistic and changing environments. At last, appearance-based techniques use the appearance of the whole scene. This approach consist on two phases, the learning one, where the robot stores general visual information from several points of view along the route to follow, and the autonomous navigation, where the robot follows the route comparing the current visual information with the stored one. The main disadvantage of these techniques is that they require huge amounts of memory and they suppose high computational cost to model the route and make the necessary comparisons during the autonomous navigation.

In structured environments, auto-location can be performed distinguishing landmarks inside the scene, so we can use model-based techniques. As an example, [12] presents the application of the visual servoing approach to a mobile robot. The robot

F. Roli and S. Vitulano (Eds.): ICIAP 2005, LNCS 3617, pp. 1109 – 1116, 2005.

is controlled according to the position of some features in an image and obstacle detection is outperformed throughout a combination of vision and ultrasonic sensors. Nevertheless, precision in the navigation will depend on the ability to recognize this set of features, what may become very difficult in most situations. Besides, additional problems may arise in non-structured environments, where artificial landmarks cannot be allocated or there are no natural landmarks that could be segmented with precision. In this case, using the properties of global appearance of the images could result more adequate. Also, due to the constant improvement of computer technology, appearance-based approaches have become a feasible technique despite its computational cost. Anyway, the key points to carry out a successful navigation with these methods reside on the quantity of information to store and in how to make the comparison between the current view and the stored information to reduce the computing time.

Researchers have proposed several methods to outperform auto-location and navigation based on global appearance. Matsumoto et al. [7], [8] addressed the VSRR method, consisting on the direct comparison of low-resolution images. Jones et al. [2] proposed a method based on the same concept but using a couple of cameras and odometer information to carry out navigation. Other approach makes use of the color histogram to perform auto-location. The problem is that the histogram does not contain any spatial information, so it is necessary to extract other features, as texture, density of edges, etc., as in [13], [15]. Regini et al. [10] proposed a method that calculates spatial relationships between the regions of color. Also, the complexity of the problem can be reduced working in the PCA space as in [1], [6], or trying to extract the regions of the image that contain the most significant information as in [14].

The approach addressed in this paper is based on the VSRR model with a couple of cameras, but with the objective of following pre-recorded routes using only visual information, with no odometer information. New control laws for the linear and steering velocities are proposed and it is also carried out a study of the parameters that affect the navigation to optimize the accuracy and speed in the route following.

The paper is organized as follows. Section 2 describes the key points of the developed application. Section 3 studies several parameters that have influence over the robot navigation and their optimization. Finally, section 4 exposes the conclusions and the future work that can be developed to improve navigation.

2 Visual Navigation Using Appearance Features

To develop the application, the B21r mobile robot has been used. This robot has 4-wheel drive with synchronous drive kinematics. It means that the four wheels can spin around the vertical axis which passes through their centers, remaining parallel all time. The separation between the driving and the steering systems makes possible to control independently the linear velocity v (robot advance) and the steering velocity ω (robot direction). These two variables will be used to control the robot movement.

On the top of the robot, there is a pan-tilt unit with a couple of Sony XC999 cameras with their optical axis aligned. The maximum resolution of these cameras is 640x480 pixels but an image server that provides images of less resolution has been added. The simultaneous use of two cameras will make our method more robust.

Fig. 1. Material used to develop the application. B21r Robot and Sony XC999 cameras.

The proposed approach is based on the VSRR model [4]. This model solves the problem of the huge memories using low resolution images to model the route.

With the purpose of following pre-recorded routes two phases need to be accomplished: a learning phase in which some visual information about the route is stored and an autonomous navigation phase, in which the current position of the robot must be estimated to drive it through the learnt route. To avoid the vulnerability to environment changes such as unknown obstacles, walking persons and cast shadows, a new phase is being implemented, to detect obstacles using just the two cameras to sense the environment.

2.1 Learning Phase

Firstly, the route is decomposed in straight segments, and the robot is guided in a teleoperated way through these segments, taking images simultaneously with both cameras in several points of the route. In these points, it is also stored, in a qualitative way, the next action the robot must take to follow the route. This action is stored as 'f' when the robot has to drive forward, 'r' when the robot has to stop and turn right, 'l' when it has to turn left or 's' when the end of the route has been reached and the robot must stop. All this information (images and actions) is stored in text files. In the turning points, just two pairs of views are stored, one at the beginning and one at the end of the turn.

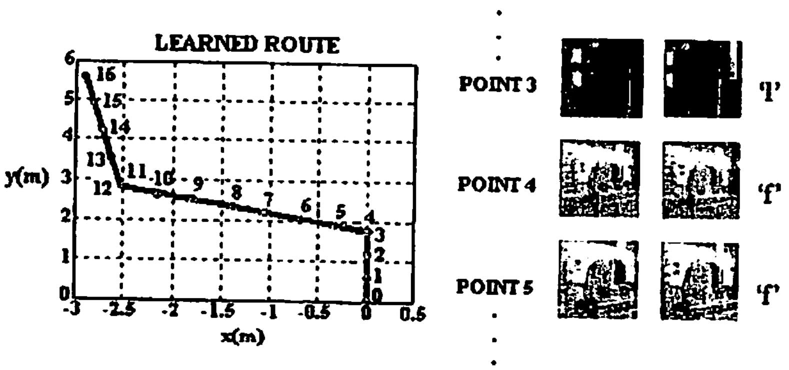

Fig. 2. Form of the database created during the learning phase

2.2 Autonomous Navigation

During the autonomous navigation, the robot is located in a point near the learnt route. Then, it has to recognize which of the stored positions is the nearest to the current and drive to tend to the route, following it till the end. To do this, two processes, that are executed successively, have been implemented: auto-location and autonomous navigation. The robot will drive forward in the straight segments with a correcting turning speed to tend to the route, and it will have a pure rotation movement in turning points.

Auto-location. To carry out auto-location, the current entire images are compared with all those previously stored in the database (the current left image with all stored left images, and the same with the right image). The adopted comparison criterion is the zero mean cross correlation:

$$\gamma_i = \frac{\sum_{x,y}\left[I(x,y)-\overline{I}\right]\cdot\left[A(x,y)-\overline{A}\right]}{E_I\cdot E_A}.$$
$$E_I=\sqrt{\sum_{x,y}\left[I(x,y)-\overline{I}\right]^2}\quad E_A=\sqrt{\sum_{x,y}\left[A(x,y)-\overline{A}\right]^2}. \tag{1}$$

I and A are the two images to compare, whose average values are $\overline{I},\overline{A}$ and whose energies are E_I, E_A. Although this is a computationally demanding way of doing image matching, it has been chosen because of its insensibility to the scene illumination and input noise [5], comparing to other criteria such as the direct difference of the images. The average and the energy of the stored images can be calculated in an off-line process, what can save computational cost during navigation. As in each point we have two images, to obtain a general data of comparison we use the arithmetic average of the correlations: $\gamma_{av}=\left(\gamma_{left}+\gamma_{right}\right)/2$.

At the beginning of the navigation, the current images need to be compared with all the stored ones, recognizing the nearest position by means of the average correlation. In the next iterations, it is necessary to compare the current image only with the previously matched and the following one, because navigation is continuous and robot has to pass through all pre-recorded points successively. Then, once the robot has started navigation, the time of processing is independent of the database size, and so, of the length of the route to be followed. Besides, we can take points frequently during learning step to get more accuracy without increasing computational cost.

Control. In this task, we have to correct the robot steering to make it tend to the route and follow it to the end. The behavior of the robot depends on the action that the currently matched images have got associated in the database.

If the associated action is 'f' (go forward), we have to correct the small fluctuations that the robot position may suffer respect the route to follow. To do this, the current images are compared continuously with the previously matched and the next ones (Fig. 3: Auto-location). Once we have a new matching, we take a sub-window in the left matched image and track it on the right matched image, identifying the sub-window of the right image that better correlates with the left one (Fig. 3: Control, step 1. This can be done off-line). As the optical axes are parallel, the vertical offset of the right sub-window respect the left one must be zero. Then, the same process is carried out between the left matched image sub-window and the left current image (Fig. 3: Control, step 2), and finally, between the right images (Fig 3: Control, step 3). The horizontal offsets x_1 (between the left sub-windows) and x_2 (between the right sub-windows) allow calculating the necessary steering velocity to tend to the route. The linear velocity will be proportional to the average correlation, what means that when we are far from the route, the linear velocity will be low to allow the robot correct its trajectory, but when we are following the route quite good, the robot goes quicker. Then, the proposed control law is:

$$\omega = k_1 \cdot x_1 + k_2 \cdot x_2 . \quad v = k_3 \cdot \gamma_{av} . \tag{2}$$

Being k_i three constants whose value will be studied in the next section.

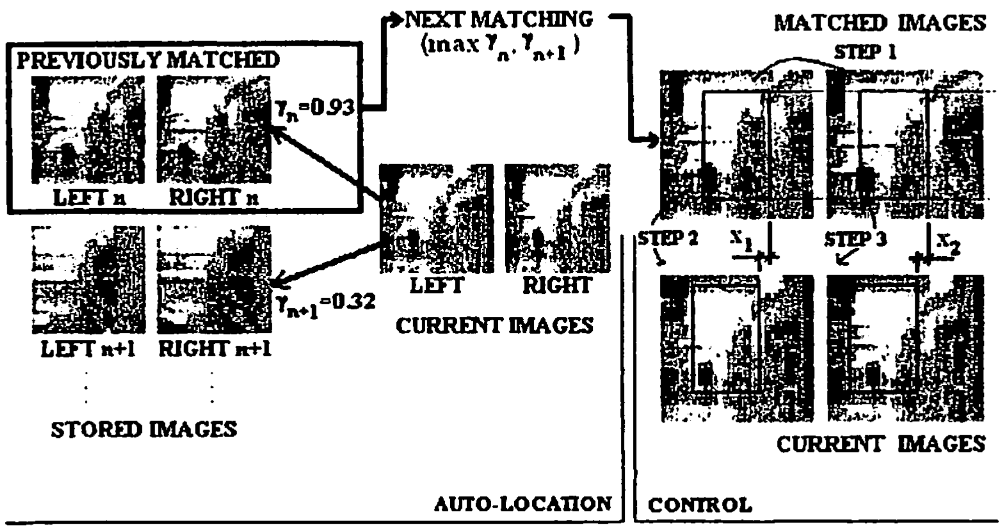

Fig. 3. Tasks performed during autonomous navigation when the action is 'f'. First one, the robot makes auto-location, comparing current images with the previously matched and the next ones. Once we have a match we calculate the linear and steering speeds based on the global correlation and the horizontal displacement of a template.

If the associated action is 'r' or 'l' (turning right or left), the robot keeps going forward according to equation (2), until the average correlation has arrived to a maximum and decreases a 10% respect to this maximum. This is to avoid the fact that the robot begins turning too soon. In this point, the robot stops and begins turning, comparing the current images with the following stored. The steering velocity is inverse proportional to global correlation. When it begins turning, correlation is low, so the robot turns quickly, and as we go closer the final position, correlation increases, so the speed decreases. This supposes a compromise between speed and accuracy.

$$w = k_4 / \gamma_{av} . \quad v = 0. \tag{3}$$

3 Analysis of the Navigation Parameters

In this section we study the influence of the parameters previously presented in the behavior of the robot. To carry out this analysis the effects of the linear speed, shape and size of the comparison template and steering speed have been studied separately.

In all the shown experiments, the learnt route, is composed by 16 couples of simultaneous images along a total length of about 8 meters, with two turning points, one to the left and the other one to the right. A resolution of 32x32 pixels has been chosen to capture the images. This allows a top speed of about 1 m/s. Lower resolutions do not provide information enough so the robot gets lost often. Higher resolutions suppose an important increment of calculations at each iteration so the robot speed has to decrease in the same proportion. As an example, a 64x64 resolution allows a top speed of only about 0.3 m/s.

3.1 Influence of the Advance Speed

During the navigation, the advance speed of the robot is a linear function of the global correlation, according to (2). So the advance speed can be modified simply changing the value of the constant k_3. Fig. 4 shows the value of the average correlation during all navigation (average for all the current images during the navigation) for different values of k_3, This value is taken as a measure of the following accuracy, so the accuracy in the following of the learnt route decreases as the speed is higher. Then, we will have to arrive to an agreement between speed and accuracy, depending on the application. For values higher than 0.8, the navigation is not possible. In this case, the robot runs a too long distance with the same control action so, when it is updated, the robot is already lost.

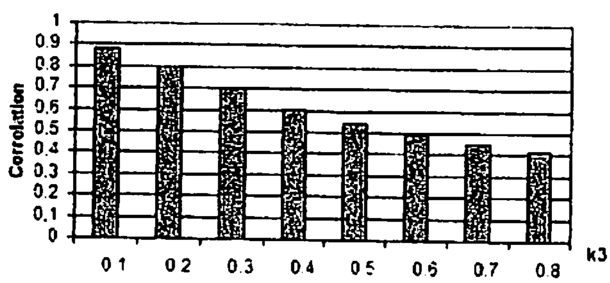

Fig. 4. Results of the navigation accuracy for different linear speeds

3.2 Shape and Size of the Comparison Template

The control during navigation is based on the continuous comparison of a template that is taken on the matched stored images over the current images. So, shape and size of the template result to be parameters of great importance, in part, due to the fact that the computational cost depends directly on the size of this template (for example, using a 24x10 template instead of 16x16 saves a 33% of operations). Exhaustive experimentation makes us arrive to the following conclusions.

The optimal size is the one that has about the half of rows and columns of the complete images. For bigger sizes, the error in the route following increases quickly. In this case, if the deviation respect the learned route is high, part of the template will not appear in the current image. As shown in fig. 5, navigation may become impossible in spite of the fact that the time per iteration is quite low. At last, templates with higher number of rows than columns present worse results.

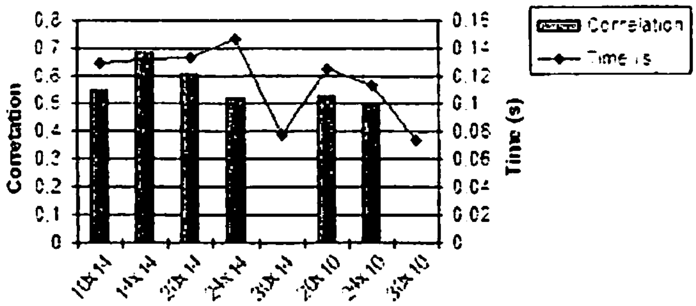

Fig. 5. Results of the correlation and time per iteration for different sizes and shapes of the comparison template. A zero correlation means that navigation has not been possible.

3.3 Influence of the Steering Velocity

The corrections on the robot trajectory so that it follows the learnt route are calculated as a steering velocity that depends on the horizontal offsets between the current images and the sub-windows on the stored ones. Taking into account this fact, the effect of varying the turning speed can be appreciated changing the values of the constants k_1 and k_2. The experiments show that the navigation is very little sensitive to the variation of these parameters. When they take values in [0.001, 0.040], the changes of these parameters do not affect the global error in a noticeable way.

4 Conclusions and Future Work

In this work, new control laws for appearance based navigation with two cameras have been proposed and tested, carrying out a sensibility analysis of the navigation parameters to try to optimize the accuracy and the speed in the task of following pre-recorded routes. With these laws, the robot is able to find itself and follow the route in a band of about 2 meters around it. It can be done although the scene suffers changes (illumination, position of some objects...). There are some future works that could improve our navigation system, such as the automation of the learning phase and the implementation of continuous navigation. An automated learning phase would imply the robot should just be driven along the desired route, but the image acquisition would be carried out automatically, when the correlation of the current image respect to the last one stored would go down a threshold. On the other hand, the continuous navigation would allow navigation in any route, without decomposing it in straight segments, so, no information about the action that the robot takes during learning must be stored. That would suppose an important change in the control law.

Acknowledgements

This work has been supported by Ministerio de Educación y Ciencia through project DPI2004-07433-C02-01 'Herramientas de Teleoperación Colaborativa. Aplicación al Control Cooperativo de Robots'.

References

1. Artac, M., Jogan, M., Leonardis, A.: Mobile robot localization using an incremental eigenspace model. Proceedings of the IEEE International Conference on Robotics & Automation (2002) 1025-1030
2. Jones, S.D., Andersen, C., Crowley, J.L.: Appearance based processes for visual navigation. Proceedings of the IEEE International Conference on Intelligent Robots and Systems (1997) 551-557
3. Kosaka, A., Kak, A.C., Fast vision-guided mobile robot navigation using model-based reasoning and prediction of uncertainties. Computer Vision, Graphics and Image Processing: Image Understanding. Vol 56, no3 (1992) 271-329
4. Lebegue, X., Aggarwal, J.K.: Significant line segments for an indoor mobile robot. IEEE Transactions on Robotics and Automation. Vol. 9, no 6 (1993) 801-815
5. Lewis, J.P.: Fast normalized cross-correlation. Expanded version of paper from Vision Interface (1995) 120-123
6. Maeda, S., Kuno, Y., Shirai, Y.: Active navigation vision based on eigenspace analysis. Proceedings IEEE International Conference on Intelligent Robots and Systems. Vol 2 (1997) 1018-1023
7. Matsumoto, Y., Ikeda, K., Inaba, M., Inoue H.: Visual navigation using omnidirectional view sequence. Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (1999) 317-322
8. Matsumoto, Y., Inaba, M., Inoue, H.: Visual navigation using view-sequenced route representation. Proceedings of IEEE International conference on Robotics and Automation. Vol 1 (1996) 83-88
9. Ohno, T., Ohya, A., Yuta, S.: Autonomous navigation for mobile robots referring prerecorded image sequence. Proceedings IEEE International Conference on Intelligent Robots and Systems. Vol 2 (1996) 672-679
10. Regini, L., Tascini, G., Zingaretti, P.: Appearance-based robot navigation. Proceedings of the Workshop su agenti robotici, Associazione Italiana per l'Intelligenze Artificiale, VIII Convegno. (2002)
11. Santos-Victor, J., Sandini, G., Curotto, F., Garibaldi, S.: Divergent stereo for robot navigation: Learning from bees. Proceedings IEEE International conference con Computer Vision and Pattern Recognition (1993) 434-439
12. Swain-Oropeza, R., Devy, M., Cadenat, V.: Controlling the execution of a visual servoing task. Journal of Intelligent and Robotic Systems. Vol 25, No 4 (1999) 357-369
13. Ulrich, I., Nourbakhsh, I.: Appearance-based place recognition for topological localization. Proceedings IEEE International Conference on Robotics and Automation (2000) 1023-1029
14. Winters, N., Santos-Victor, J.: Information Sampling for optimal image data selection. Proceed of the 9th International Symposium on Intelligent Robotics Systems (2001) 249-257
15. Zhou, C., Wei, T., Tan, T.: Mobile robot self-localization based on global visual appearance features. Proceedings of the 2003 IEEE International Conference on Robotics & Automation (2003) 1271-1276