International Society for Advanced Research

ISSN 1828-6984

INTERNATIONAL JOURNAL OF FACTORY AUTOMATION, ROBOTICS AND SOFT COMPUTING

Issue 3, July 2007

Editor-in-Chief: Salvatore Pennacchio

www.internationalsar.org

Robust Vision Landmark Acquisition and Data Association for an Autonomous Mobile Robot

R. VÁZQUEZ-MARTÍN, P. NÚÑEZ, R. MARFIL, A. BANDERA and F. SANDOVAL
Grupo de Ingeniería de Sistemas Integrados, Dpto. Tecnología Electrónica
E.T.S.I. Telecomunicación, University of Málaga
Campus de Teatinos s/n, 29071-Málaga - SPAIN
rvmartin@dte.uma.es http://www.grupoisis.uma.es

Abstract: **This paper presents a visual attention mechanism for the robust acquisition and association of landmarks. The proposed attention mechanism consists of two consecutive selection stages and a final attentive stage. The first stage employs preattentive saliency computations to select a reduced set of regions of interest from the whole input image. The second stage selects from the output of the first selection stage the region that can be considered as a landmark. This is the input of the attentive stage, that characterizes the detected landmark and compares it with the ones stored in a landmark map by means of a robust M-estimator criterion. In our tests, this map must be generated a priori and it stores the location and color template associated to a set of reference landmarks. The attentive stage also returns the confidence value associated to the matching. If this matching confidence is low, the robot can solve data association ambiguity or reject spurious observations. Experimental results show that the proposed method is able to detect and correctly mapping landmarks in indoor and outdoor scenes.**

Key-Words: - **Mobile robotics, SLAM, visual landmark, data association, template matching**

1. Introduction

Reliable localization and navigation are two essential components of an autonomous robot. In order to perform these tasks correctly, the robot needs: i) to detect obstacles and free space in its environment reliably and fast; and ii) to represent this information into a navigation map. This navigation map can be built with features present in the environment that are automatically detected by external sensors. In these cases, recognizable features are essential since they will be used to bound the navigation errors. In the past, a variety of approaches for mobile robot localization and navigation has been developed. These approaches mainly differ in the methods employed to represent the belief of the mobile robot about its current pose or to find and track a safe path to a goal, and according to the type of sensor information that they use. In this paper, we consider the problem of vision-based landmark acquisition and association. Landmark association can be defined as the problem of relating observations with landmarks and, traditionally, it is considered as one of the main problems of the EKF-based SLAM approaches. Compared with active ranging devices, visual image data provides a huge amount of information (color, texture or shape) that can permit to disambiguate features for data association purposes. Besides, cameras are passive sensors, so that vision-based systems do not suffer from the interference often observed when using sonar or laser sensors. On the other hand, to perform the data association, visual-based systems need to compare images obtained by the robot's cameras with images stored in a database. This matching process has a high computational complexity, consuming a big amount of computational resources. So, it must be carefully designed in order to avoid a slow response of the system.

There are a variety of different approaches from computer vision to compare perceived images with previously stored images. Iconic methods directly compare the raw data with the stored one [1]. In order to avoid the direct matching of raw image pixels, feature-based methods consider mainly the prominent features [2]. Thus, image-retrieval approaches try to find images in a database that look similar to the given query image [2]. The key idea is to compute image features that are invariant with respect to image rotations, translations and limited scale. Other feature-based methods try to find potential objects of interest in the environment. This selection permits to reduce the overall computational cost of the process, but it also implies a preliminar step to obtain these potential features of the environment. One popular technique to implement that selection process is to model a visual attention mechanism. If visual systems have been usually discarded because of its high demand of computational resources, attention is a mechanism that show the quadrangular vertically-oriented objects detected by the robot when it was in the poses marked as #1 to #8 in Fig. 3.3a. Like in Fig. 3.2, all observations are marked by black boxes in the input frames, while the observation associated to a landmark is marked by a black and white bounding box.

4. Conclusion

This chapter describes a visual attention mechanism dedicated to the landmark acquisition for a mobile robot in structured or unstructured environments. Experimental results show that the proposed visual attention mechanism can be used to detect good candidates for landmarks, avoiding an exhaustive search in the whole image. The main contribution of this work is the proposed model of visual attention, that integrates bottom-up and top-down processing. This model avoids the problems related to the detection as landmarks of large landscape elements (e.g., lakes, rivers, clouds) that although may be visually salient, cannot be used as punctual landmarks. Besides, a template-based approach for data association has been integrated in the same framework with the proposed attention mechanism. A robust M-estimator criterion is employed to perform the template matching. This method will permit to avoid data association failures in a SLAM process.

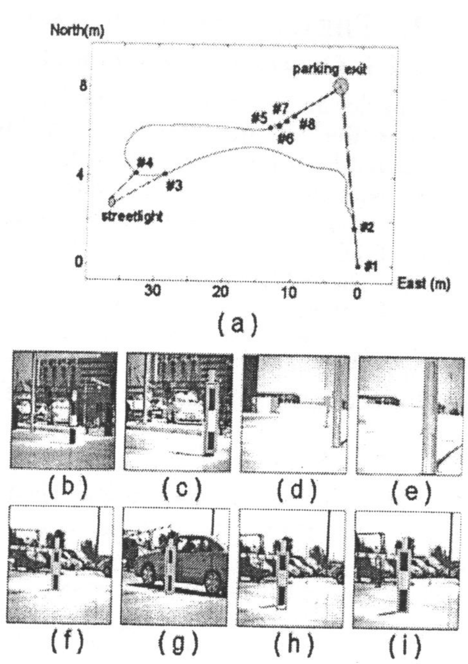

Fig. 3.2. Exploration of a scene: a) Robot trajectory and detected landmarks; b-i) left captured image annotated with the bounding boxes of observations.

Acknowledgement

This work has been partially granted by the Spanish Ministerio de Educación y Ciencia (MEC) and FEDER funds Project no. TIN2005-1359.

References

[1] D. Yuen and B. MacDonald, "Vision-based localization algorithm based on landmark matching, triangulation, reconstruction, and comparison", *IEEE Trans. on Robotics*, vol. 21, no. 2, 2005, pp. 217-226.

[2] J. Wolf, W. Burgard and H. Burkhardt, "Robust vision-based localization by combining an image-retrieval system with Monte Carlo localization", *IEEE Trans. on Robotics*, vol. 21, no. 2, 2005, pp. 208-216.

[3] C. Koch and S. Ullman, "Shifts in selective visual attention: towards the underlying neural circuitry", *Human Neurobiology*, vol. 4, 1985, pp. 219-227.

[4] J. Chen, C. Chen and Y. Chen, "Fast algorithm for robust template matching with M-estimators", *IEEE Trans. on Signal Processing*, vol. 51, no. 1, 2003, pp. 230-243.

[5] Y. Sun and R.B. Fisher, "Object-based visual attention for computer vision", *Artificial Intelligence*, vol. 146, no. 1, 2003, pp. 77-123.

[6] M. Mirza and K. Boyer, "Performance evaluation of a class of M-estimators for surface parameter estimation in noisy range data", *IEEE Trans. on Robotics and Automation*, vol. 9, 1993, pp. 75-85.

[7] R. Vázquez-Martín, J.C. del Toro, A. Bandera and F. Sandoval, "Data- and model-driven attention mechanism for autonomous visual landmark acquisition", *Proc. IEEE Int. Conf. on Robotics and Automation*, 2005, pp. 3383-3388.

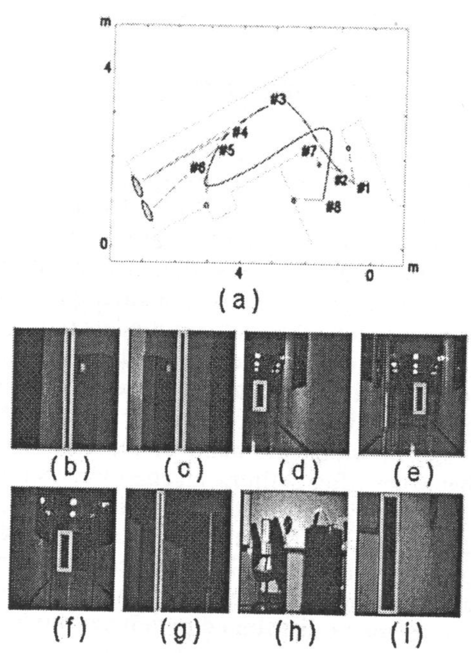

Fig. 3.3. Exploration of a scene: a) Robot trajectory and detected landmarks; b-i) left captured image annotated with the bounding boxes of observations.