MARCO: A mobile robot with learning capabilities to perceive and interact with its environment*

A. Sanfeliu†, J. Andrade-Cetto†, R. Alquézar‡, J. Aranda§,
J. Climent§, A. Grau§, F. Serratosa¶ and J. Vergés-Llahí†.
†Instituto de Robótica e Informática Industrial, UPC-CSIC.
‡Depto. de Sistemas y Lenguajes Informáticos, UPC.
§Depto. de Ingeniería de Sistemas, Automática e Informática Industrial, UPC.
¶Depto. de Ingeniería Informática y Matemáticas, URV.

Abstract

Marco is the name of a research mobile robot that is being developed at the Instituto de Robótica e Infomática Industrial of UPC-CSIC. It is designed with learning abilities to acquire information about indoor environments using various perception sensors. Marco uses video cameras and ultrasonic sensors to perceive the world, and pattern recognition and computer vision techniques to extract knowledge about it. In this paper we explain the objectives of this project and the techniques developed so far.

Keywords: automatic learning techniques, computer vision, pattern recognition, mobile robots.

1 Introduction

There exist numerous contributions to mobile robot perception and automatic robot guidance, however few robots use automatic learning techniques to acquire knowledge about its environment. The work described in this paper will illustrate the objectives of the Marco Mobile Robot project, and the techniques developed to date.

The fundamental work that is being developed started in 1992, when we came up with new techniques for automatic learning of objects based on grammatical inference. These learning techniques were extended to learn 3D objects based on the graph synthesis of appearance views of a 3D object [1]. Lately, we have developed new techniques to learn not only 3D objects, but also the robot environment using visual scene features.

* Work partially funded by CICYT project TAP98-0473.

J.S. Sánchez and F. Pla editors, Proc. 9th Spanish Sym. Pattern Recog. Image Anal.
Benicassim, May 2001, Vol. 2, pp. 219-224.

2 Objectives of the MARCO Mobile Robot Project

The fundamental objective of the Marco Mobile Robot project is to develop new automatic learning techniques to acquire and interact with the information perceived by external sensors of any kind for a mobile robot platform. These learning capabilities in conjunction with reliable sensory data processing techniques can be applied to different mobile robotics platforms, including autonomous vehicles.

The specific short and medium-term goals of the Marco Mobile Robot project are: (1) to develop the basic tools and techniques in order for Marco to navigate in an indoor environment – real-time stereo vision, path planning and obstacle avoidance, among others–; (2) to learn and identify the 3D objects that limit the robot pathway; (3) to autonomously learn and identify visual (and probably audible) landmarks in the pathway for self-location purposes; (4) to track moving objects in order to be able to follow and identify them; and (5) to track and learn human faces in its pathway in order to identify them.

Marco is a mobile robot (Pioneer 2 DX) with 2 sonar arrays totalling 16 ultrasonic sensors firing at 25MHz, two bumpers with 5 contact switches each, and two colour video cameras mounted on a computer controlled Pan and Tilt head (see Figure 3). Inside Marco, there is a Pentium II based PC (266MHz CPU and 128 MB RAM) connected to a wireless Ethernet network at 2.4GHz. Video acquisition takes place on external computers with Matrox Meteor and a Genesis boards, receiving the video signals from the robot using a pair of microwave wireless video senders. Learning, navigation, path planning, and image analysis routines take place on external computers on the Ethernet network. Marco's internal computer is used solely for sensor monitoring and control, i.e., to read and process data from the sonar rings, to control the Pan and Tilt unit, to read the dead-reckoning data, and to control the internal parameters of the cameras. Navigation, path planning, image analysis and other high level tasks take place in the external computers.

3 3D object learning and recognition

This is probably one of the most ambitious tasks in Marco's project. Its difficulty stands from the fact that to learn and recognise an object, it needs to be initially segmented out from the scene. In other words, it is necessary to acquire, segment and extract the fundamental features that provide robust object learning and recognition. The segmentation of a colour scene is a typical problem in computer vision, and it still has not been solved for unstructured environments. Moreover, the robustness of the techniques developed so far, is still very poor. However, some results can be achieved if some restrictions are imposed on the objects to be learned and on the environment illumination characteristics.

In order to overcome some of the difficulties during segmentation, we have been working in colour constancy algorithms, colour image segmentation and the fusion of depth information with colour segmentation results. For colour constancy we still cannot keep invariant the colour of the scene objects when the robot is moving.

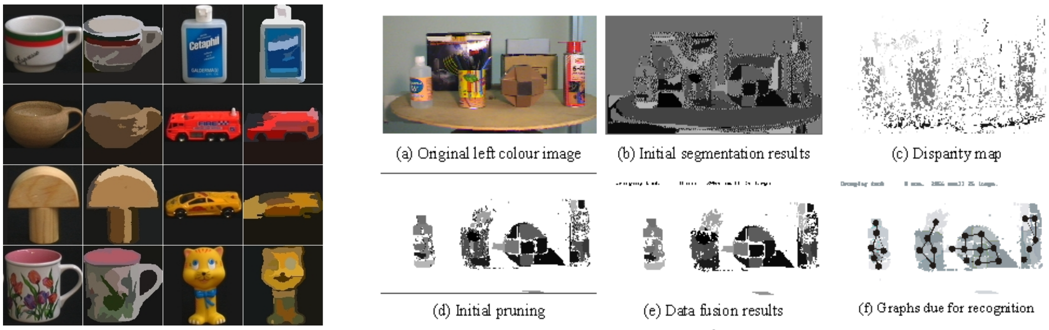
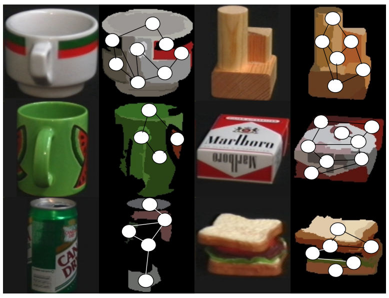


Figure 1. Segmentation results and data fusion steps.

Better results have been obtained on colour segmentation, where we are using a new method based on a greedy approach for graph partitioning into a set of spanning trees [2]. An input image is used to build up a graph where its connected components are object segments. Another approach is to combine depth information with colour segmentation [3, 4]. This approach gives improved segmentation results, and also separates the objects of interest from the rest of objects and from the background using depth cues (Figure 1).

Once an object has been segmented, its constituent parts are used to build an FDG (Function Described Graph [5]) in the following way: from each segmented region on the object we extract some quasi-invariant features (colour, circularity, Euler number etc.) which are then used as node attributes on the FDG. We can also use the distance between region colours as arc attributes on the FDG.

Object learning proceeds as follows: (1) take a sequence of object images in a clockwise or counter-clockwise direction maintaining a close distance from it; (2) for each image, do segmentation, feature extraction, and obtain its associated FDG; (3) apply the synthesis algorithm [5, 6] using all the FDGs. Once an object is learned, we apply the same methods for segmentation and feature extraction to a new scene. Recognition is computed using the distance measure between graphs described in [7].



Average number of vertices and second-order relations of the FDGs.						
Number of FDGs per object	**9**	**6**	**4**	**3**	**2**	**1**
Number of AGs per FDG	**4**	**6**	**9**	**12**	**18**	**36**
Average Num. of FDG vertices	11	13	19	25	31	44
Average Number Antagonisms	22	47	88	98	118	174
Average Number Occurrences	112	121	132	139	141	199
Average Number Existences	0	0	0	0	0	0

Figure 2. Recognition results using FDGs on some color-segmented objects.

Experiments of both processes, learning and recognition, have been done using different types of objects in a controlled environment, where instead of moving the robot around the object, the object is rotated on a round table (Figure 2). Even when the method works well for isolated objects with controlled background and lightning conditions, a lot of work has to be done in order to do the same in a partially controlled environment. For example, in a typical indoor setting, or at a manufacture plant.

4 Learning of visual landmarks for robot self-location

One of the basic tasks for a mobile robot a robot is to be able to identify its position based on self-acquired visual or ultrasonic landmarks from the environment. Towards this end, we are developing a new technique based on the construction of a dynamic 3D map model of the robot environment named *Burrow 3D Map* for robot self-location. The term "burrow" comes intuitively from the analogous exploration techniques used by rodents. The data structure used in this technique resembles a tunnel-like map of the environment that is constructed as the mobile robot navigates.

The approach generates a probabilistic map of the visual 3D path that the robot follows when it moves. The probabilistic 3D map is built using robust features extracted from grey-level images of a stereovision pair. For the time being, we are using as vertices and vertical and horizontal edges as visual cues. For these features to be incorporated in the map, they must be matched in the stereo pair. Once they are matched, their 3D position is computed and included in the Burrow 3D Map as explained in [8]. By using this Borrow 3D Map, the robot can locate its position by computing its distance to the identified 3D landmarks.

5 Object tracking and following

Tracking visual landmarks is an important behaviour that is being incorporated in Marco. We want Marco to perform the following tasks: (1) to track a known object, such as a human face for identification purposes or an object within a fixed window during a learning step; and (2) to detect and track unknown moving objects close to the robot. In our current approach the observer is initially stationary, and we use a model-based technique based on colour histograms to characterize it. The system has two different modules: the first one termed *model learning*, extracts the features that describe the object using two different methods: supervised and unsupervised. The second part carries out *object tracking* using the model learned in the first part.

Object models are characterized by its apparent image size and the information given by the colour histogram of pixels in the learned object. The histogram is used to segment a new image in the sequence into object and background [9]. Then, new features are extracted from the segmented object, the histogram model is updated, and the commands needed to move the Pan and Tilt unit are computed so that it follows the tracked object. Initially, it is necessary to learn the object model. This learning can be *supervised*, i.e., when we know a priori the object that will be tracked; or *unsupervised*, tracking an arbitrary single object in motion. Unsupervised learning needs an initial step of *independent motion detection* to focus the mobile object search and to learn the object model, in our case by frame differencing. Our tracking algorithm has three main steps: object segmentation, feature extraction and, feature adaptation by means of a Kalman filter. In the near future, we will incorporate the object following ability by moving not only Marco's head, but its base as well. The main difficulty for this latter task is the time response of Marco's motion commands. In order for Marco to obey some safety guidelines, the object tracking behaviour cannot supersede the sonar-based obstacle avoidance behaviour. We are currently working on a robust method to subsume both behaviours simultaneously.

6 Learning people faces that Marco finds in its pathway

We have been developing two approaches to learn human faces. Our first method is presented in the work of [10] where basic features from faces are extracted. These features correspond to the nodes on a graph. We use the Euclidean distance between these image features as the attributes on the arcs. Using this information, we learn the graph attributes for various faces by a synthesis procedure [5]. To identify a face, we compare the extracted information on the scene face to all the faces that have been learned previously. Our second approach is based on the popular PCA methods. We use eigenfaces for the identification of a face in a scene and for person recognition among faces in a database. These methods will be incorporated to Marco in the near future for face tracking and recognition to identify people on its path. Another area of research under development with the aid of Marco is that of 3D reconstruction from stereo (see Figure 3).

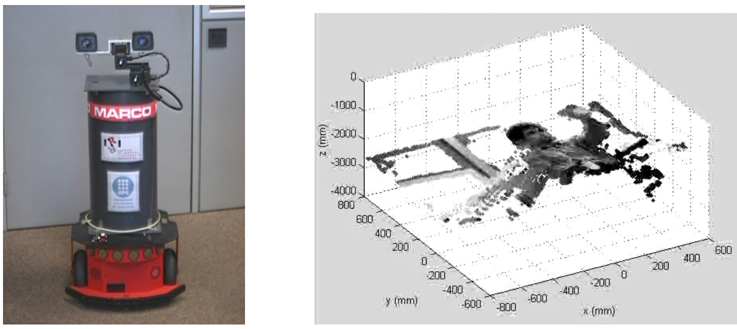


Figure 3. Marco: a mobile robot with learning abilities, and a sample image of 3D reconstruction using its stereo head.

7 Conclusions

Marco is a mobile robot with learning capabilities that use various learning methods to acquire information about its surroundings. The Marco Mobile Robot Project is an ambitious research project that is under development, and that has to face with very complex problems. We foresee that a general solution to some of these problems is unlikely to be found in the near-term. However, we think that if some constraints are imposed, many of the goals can be achieved. This research project is intended for the autonomous learning of indoor environments for a mobile robot in industrial settings.

References

[1] F. Serratosa and A. Sanfeliu. Function-described graphs applied to 3D object recognition. *Proc. 9th Int. Conf. Image Anal. Process.*, Firenze, Vol. 1, pp. 701-708, 1997.

[2] J. Vergés-Llahí and A. Sanfeliu. Colour image segmentation solving hard-constraints on graph partitioning greedy algorithms. *Proc. 15th IAPR Int. Conf. Pattern Recog.*, Barcelona, Vol. 3, pp. 629-632, 2000.

[3] S. Pankanti and A. K. Jain. Integrating vision modules: stereo, shading, grouping, and line labelling. *IEEE T. Pattern Anal.*, 17(8): 831-842, 1995.

[4] J. Andrade-Cetto and A. Sanfeliu. Integration of perceptual grouping and depth. *Proc. 15th IAPR Int. Conf. on Pattern Recog.*, Barcelona, Vol. 1, pp. 295-298, 2000.

[5] R. Alquézar, A. Sanfeliu and F. Serratosa. Synthesis of function-described graphs. *Advances in Pattern Recog., Proc. Joint IAPR Int. Workshops SSPR'98 and SPR'98,* Sydney, Springer LNCS-1451, pp. 112-121, 1998.

[6] A. Sanfeliu, F. Serratosa and R. Alquézar. Clustering of attributed graphs and unsupervised synthesis of function-described graphs. *Proc. 15th IAPR Int. Conf. on Pattern Recog.*, Barcelona, Vol. 2, pp. 1026-1029, 2000.

[7] F. Serratosa, R. Alquézar and A. Sanfeliu. Efficient algorithms for matching attributed graphs and function-described graphs, *Proc. 15th IAPR Int. Conf. on Pattern Recog.*, Barcelona, Vol. 2, pp. 871-876, 2000.

[8] J. Andrade-Cetto and A. Sanfeliu. Topological map learning for a mobile robot in indoor environments. Accepted for presentation *9th Spanish Sym. Pattern Recog. Image Anal.*, Castellón, May 2001.

[9] J. Vergés-Llahí, J. Aranda and A. Sanfeliu. Reliable object tracking system using colour histograms. Accepted for presentation *9th Spanish Sym. Pattern Recog. Image Anal.*, Castellón, May 2001.

[10] J. Vergés-Llahí, A. Sanfeliu, F. Serratosa and R. Alquézar. Face recognition: graph matching versus neural techniques. *Proc. 8th Spanish Sym. Pattern Recog. Image Anal.*, Bilbao, Vol. 1, pp. 259-266, 1999.