VISUAL NAVIGATION OUTDOORS: THE ARGOS PROJECT

ENRIC CELAYA and CARME TORRAS
Institut de Robòtica i Informàtica Industrial (CSIC-UPC)
Llorens i Artigas 4-6, 08028-Barcelona
celaya@iri.upc.es, torras@iri.upc.es

Abstract. Visual navigation in unstructured, previously unknown, environments is investigated in the project ARGOS. This paper describes the goals of the project, the guidelines we follow to face each aspect of it, and the motivations that led us to work with a walking robot. The current state of the project in the areas of legged locomotion, landmark-based navigation and vision is succintly described.

1 Introduction

Visual navigation, that is, reaching a target using visual information, is a capability that most animals exhibit in different degrees, and should be considered as a basic one for any autonomous mobile robot. The goal of the ARGOS project[1] is to develop a system that, after placing the robot in an unstructured, previously unknown environment, allows the user to select a target in the scene as captured by the cameras of the robot, and start a navigation process to reach the target with no further human intervention.

Posed in this way, the problem faced is so general that it cannot be expected to be solved in all cases. Our purpose is to push the set of solvable situations from the simplest (e.g., target always visible and no obstacles in the way), towards the more complex ones (difficult terrain forcing long detours where the target gets occluded).

In this paper we describe the approach we follow to advance in our understanding of visual navigation, and the current state of our work in the aspects of legged locomotion, landmark-based navigation, and vision. While our work in each one of these aspects has been reported in specific papers, our purpose here is to give a global view of the project as a whole.

2 Robot autonomy

It is agreed that autonomy is a desirable property for a robot, and specially for a mobile one. It is clear that autonomy implies the ability to work without close control or supervision by a human, but in our opinion, there are other aspects that must also be considered. A first aspect is related to the diversity of environmental situations the robot can face. Only if the situations to which the robot is able to respond in a sensible way are varied enough, or even unpredictable, we say that the robot is autonomous.

A second aspect of robot autonomy has to do with the complexity of the robot itself. On the one hand, the robot sensory system has to be complex enough to distinguish

[1] Carried out jointly with the Institut d'Investigació en Intel·ligència Artificial (CSIC), and financed by the Spanish research agency (CICYT) under project DPI2000-1352-C02-01.

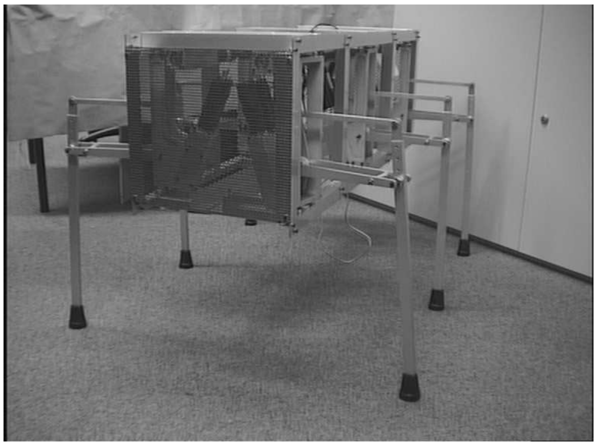

Figure 1: The six-legged robot ARGOS.

between the different environmental situations that require different actuations. On the other hand, the robot's repertoire of actions must be varied enough so that its response can be tailored to each situation.

We are interested in autonomous robot navigation in unstructured outdoor environments. The choice of this kind of environments grants the diversity of situations needed to look for autonomy. Concerning the robot, we prefer to use legged locomotion since it provides a much richer set of motor actions than wheeled locomotion, and thus, the opportunities to deal with autonomy are increased.

To deal with the problem of controlling a complex robot in a complex environment, we have decomposed the task of visual navigation in 3 main subtasks: Piloting, landmark-based navigation, and vision. Each of these subtasks is further decomposed in more specialized processes, as described in the following sections.

3 Legged robot locomotion: The pilot

The task assigned to the pilot is to move the robot in the direction specified by the navigation level while negotiating obstacles and terrain difficulties found in the way. In the pilot, no map information is used, but only reactive behaviors responding to the information provided by local sensors.

We have built a six-legged robot (Fig. 1) within this project. In order to have full mobility, each leg has three degrees of freedom, which gives a total amount of 18 for the robot. Following a behavior-based approach, we have decomposed the pilot task in 4 main subtasks [1]:

3.1 Posture control

The purpose of posture control is to keep the body in the most appropriate position with respect to the current feet locations. Keeping an optimal posture improves robot stability and mobility at the same time. The action of posture control consists in the execution of simultaneous movements of all legs in such a way that feet positions remain fixed, so that the net effect is a displacement of the body with respect to them.

3.2 Terrain adaptation

The task of terrain adaptation consists, first, in keeping contact with ground at any time with as many feet as possible and, second, in adapting the body position to avoid collisions with obstacles and unstable situations. Whenever a foot lacks ground contact, it is made to move down vertically until the contact is recovered. Note that even though this layer does only provide a mechanism to move legs down, the average leg elevation is maintained constant by virtue of the posture control described above.

When the robot detects an obstacle, its body is displaced in the opposite direction to avoid the contact. Similarly, if the robot stays in an inclined position, the body is moved up-hill in order to increase its stability.

3.3 Gait generation

This is the layer specifically devoted to walking. Two simple rules are used to determine which legs can perform a step: the *stability rule*, that forbids to rise a leg whenever this would result in losing ground contact with two neighboring legs at the same time, and the *efficiency rule*, that gives priority to legs nearer to reach their workspace limit.

3.4 Direction control

The task of this layer is to drive the robot in the direction provided by the navigation level. It is assumed that the driving commands specify the local trajectory as an arc of circumference with given center and radius. The action of this layer consists in defining the target position (usually called the AEP, or anterior extreme position) for each stepping leg, according to the local trajectory.

When an obstacle forbidding the advance of the robot in the desired direction is found, the trajectory is modified to avoid the obstacle.

4 Landmark-based navigation

The whole navigation system is implemented as a multiagent system whose activity is regulated through a bidding mechanism, in which each agent bids to gain access to the limited resources of the robot, according to the priority attributed to its task in any given situation [5].

The navigation task is decomposed in the following three ones: target reaching, map management, and path planning.

4.1 Target reaching

This layer drives the robot directly towards the target when it is perceived by the vision system and there are no big obstacles in the path to it.

4.2 Map management

We have adopted the use of a topological map, instead of a metric or a grid one. The approach we follow is that proposed in [4], that is based on the relative positions of landmarks in order to estimate the location of the target using a method called the *beta-coefficient system*.

4.3 Path planning

When the target is not visible by the robot, its position is estimated from that of the currently visible landmarks, and the robot is directed towards it. If a large obstacle is found ahead, the path is modified by directing the robot towards a provisional new target selected using the topological map.

5 Vision

The vision system we have designed to comply with the specific demands of visual navigation for legged robots is arranged around three main modules: landmark extraction and recognition, egomotion computation, and terrain characterization.

5.1 Landmark extraction and recognition

Detecting *salient* and *stable* landmarks is required to build the topological maps used by the navigation module.

Saliency outdoors should be addressed in a radically different way than in indoor environments, where geometric features (corners, straight lines, etc.) are the usual option. In natural settings, color and texture seem more appropriate characteristics. Inspired by a model of visual attention, we have devised a pyramidal system that computes multiscale saliencies based on color opponencies and texture orientations.

Attaining color *stability* outdoors is a big challenge, since the acquired visual information is strongly dependent on lighting geometry (direction and intensity of light source) and illuminant color (spectral power distribution), which change with time within the day, seasons, and atmospheric conditions. Noting that what was needed was not constancy of colors themselves, but constancy of color opponencies, we have devised a novel procedure called *multiscale color ratios.* For instance, the red-green color ratio is $R_c G_s / R_s G_c$, where R and G are red and green center or surround regions, depending on the subscript. The yellow-blue and white-black opponencies are dealt with likewise. The interesting point about these ratios is that they are intrinsically invariant to intensity and color normalizations and, as a consequence, they are immune to the inclusion/exclusion of objects in the scene [7].

The other type of features, namely texture orientations, are computed using Gabor filters. Each feature (color opponency or texture orientation) has an associated Gaussian pyramid that permits computing saliencies at different resolutions. By comparing centers and surrounds at high-resolution levels, e.g. between levels 2 and 5, visual saliencies of relatively small targets are found, while at lower resolution levels, e.g. levels 4 and 7, relatively large salient regions are found. Thus, with this technique, it is possible to detect salient objects within a wide size range, for example, from small stones to big trees.

The salient regions in the image are subsequently analysed to obtain visual signatures (based essentially on textures and contours), capable of identifying them as an existing or a new landmark. The description of the entire pyramidal system, as well as some results can be found in [6].

5.2 Egomotion computation

Since the odometry of legged robots outdoors is quite unreliable, a requirement for the vision module is to provide position estimation in the short run. This is complementary to the long-term localization based on landmarks.

Estimating camera motion from optic flow is computationally costly. Thus, we have explored the cheaper option of deriving the 3D egomotion from the deformation of an active contour fitted to a landmark (e.g., the target), by relating the affine deformation in 2D to the translation and rotation in 3D. Note that the translation along the optical axis is recovered up to a scale factor. Results of this module are reported in [2] and the particular case of a zooming camera appears in these proceedings [3].

5.3 Terrain characterization

Classifying terrains according to their degree of traversability is a very challenging task. In particular, vegetation poses many problems. Although a depth map may be useful, it cannot distinguish between different types of obstacles, some of which may be traversable. So far, we have just worked on the obtention of a depth map, and future efforts will be devoted to complementing it with additional information.

The results of our egomotion computation module permit computing the epipolar lines, which can be used to guide the search for point matches. This leads to the obtention of qualitative depth maps, such as those presented in [2].

6 Conclusions

We are addressing the problem of autonomous visual navigation in outdoor, unstructured environments. Our approach to deal with the challenging complexity of this task is to decompose the task itself in order to solve separately different aspects of it, instead of simplifying the environment on which this task is performed. Legged locomotion provides the required complexity to respond to the diversity of environmental situations that the robot must face.

The main contribution of the project in legged locomotion is an adaptable walking algorithm that allows the robot to follow arbitrary trajectories in very rough terrain. Concerning landmark-based navigation, a multiagent system is used to manage the map and locate the target. In the vision area, a novel technique to detect potential outdoor landmarks has been developed, whose stability in front of lighting variations and slight viewpoint changes has been demonstrated. Moreover, an egomotion estimation procedure based on active contours has proven to be reliable enough, with much less computation than the usual methods based on optic flow.

References

[1] Celaya, E. and Porta, J.M., "A Control Structure for the Locomotion of a Legged Robot on Difficult Terrain", *IEEE Robotics and Automation Magazine*, Vol. 5, No. 2, June 1998, pp. 43-51.

[2] Martínez, E. and Torras, C., "Qualitative vision for the guidance of legged robots in unstructured environments", *Pattern Recognition*, 34(8), 2001, pp. 1585-1599.

[3] Martínez, E. and Torras, C., "3D motion recovery while zooming using active contours", Proc. *7th Intl. Conf. on Intelligent Autonomous Systems (IAS-7)*, Marina del Rey, USA, March 2002.

[4] Prescott, T.J., "Spatial representation for navigation in animats", *Int. Adaptive Behavior*, 4(2), 1996, pp. 85-125.

[5] Sierra, C., López de Mántaras, R., Busquets, D. "Multiagent Bidding Mechanisms for Robot Qualitative Navigation", *Lecture Notes in Artificial Intelligence*. Springer, March 2001.

[6] Todt, E. and Torras, C., "Detection of natural landmarks through multiscale opponent features", *15th International Conference on Pattern Recognition*, Barcelona, Spain, 2000, vol. 3, pp. 988-991.

[7] Todt, E. and Torras, C., "Color constancy for landmark detection in outdoor environments", Proc. of the *4th European workshop on advanced mobile robots* (EUROBOT'01), Lund, Sweden, Sept. 2001.