Thesis Overview:

"A High-level Perception Architecture: Real-time Visual Navigation for Autonomous Robots in Structured Environments"

Pedro E. López de Teruel Alcolea
Universidad de Murcia (Spain)
Departamento de Ingeniería y Tecnología de Computadores
Advisors: A. Ruiz García y J.M. García Carrasco
July 25, 2003
pedroe@ditec.um.es

This thesis describes the design and implementation of a mobile robot which is able to perform a structural interpretation of indoor environments, using only visual and proprioceptive sensory information. The desired behaviour is real-time navigation based on this interpretation, instead of a reactive approach. The design is guided by a *predictive criterion*: the system must anticipate the consequences of its actions, showing a certain predictive understanding of the scene in which it moves.

We propose solutions for all perception levels. First, we have developed a segment extraction technique augmented with color information, of remarkable expressive power and computational efficiency. The method is appropriate for real-time visual geometry applications. If the expressive power of raw segments is augmented with robust color information from their two sides, the most relevant geometric and photometric structure in the image can be concisely captured. In this thesis we describe an efficient algorithm to compute this kind of representation, which can be successfully exploited in several projective geometry problems, such as 3D reconstruction, motion estimation, calibration, and in interpretation related tasks, all of them critical for robot navigation. We also show how these enhanced primitives are powerful enough to recover a very acceptable approximation of the original image, especially for partially structured scenes, like interior of buildings, man-made objects, and so on.

To infer euclidean properties of space we present a collection of autocalibration methods based on the projective geometry formalism. These techniques can be applied on-line during robot operation, using features extracted from the natural environment (thus avoiding the need of calibration patterns, typical in classical approximations). The distinctive feature of the proposed methods is that they take advantage of the odometric information of the autonomous agent when solving for the intrinsic and extrinsic parameters of the camera.

High-level perception is solved through generation, tracking and confirmation of hypothesis about the environment, which are maintained in an stable internal representation tuned with the agent movements. This mechanism is characterized by a constant interaction between the bottom-up perceptive processes, guided by sensory stimuli, and the top-down ones, guided by the models. Using the estimated camera parameters, and following reasonable clues obtained from the geometric and color information of the segments, the robot can detect relevant planes (floor, walls, doors) in indoor environments. A calibrated sensor allows for the metric rectification of these planes. The system is then able to construct an abstract interpretation of the scene, which is used to categorize different high-level typical situations (rooms, corridors, corners, and so on), that allow the robot to exhibit a non-reactive behaviour, depending on abstract properties of the environment. Finally, the robot is able to accumulate the reconstruction of the visited environment (in the form of a 3D map) as it navigates, using odometry and tracking.

All these elements are integrated in an efficient, modular and flexible hardware-software architecture, in which action and perception cooperate to achieve the robustness and continuity needed by real world navigation. Our efficient implementation and integration of all the components of the system results in a very fast perception-action cycle (25 fps in a modern 2Ghz PC, including calibration, reconstruction, interpretation, and motion control). This is enough for robust navigation of the agent at approximate speeds of 20 cm/s in real scenarios.