A New Approach to Visual Servoing in Robotics

Bernard Espiau, François Chaumette, Patrick Rives

1 ISIA-ENSMP, rue Claude Dauness, Sophia Antipolis, 06565 Valbonne, France.
2 IRISA / INRIA Rennes, campus de Beaulieu, 35042 Rennes-cedex, France.
3 INRIA Sophia Antipolis, 2004 Route des Lucioles, 06565 Valbonne, France.

Abstract. This paper describes a new approach of vision-based control in robotics. The basic idea consists in considering a vision system as a specific sensor dedicated to a task, and included in a control servo-loop. Once the necessary modeling stage is performed, the framework becomes the one of automatic control, and naturally stability and robustness questions arise.

The paper is organized as follows: in the introduction, a short state-of-the-art in the area of visual servoing is reviewed. Then, the basic concepts allowing to model the concerned interactions are given. The Interaction Screw is thus defined in a general way, and the application to images follows. Starting from the concept of task function, the general framework of the control is then described, and stability results are recalled. The concept of hybrid task is also presented and then applied to visual sensors.

The paper ends with the presentation of several simulation and experimental results, and some guidelines for future work are drawn in the conclusion.

1 Introduction

How to use vision has always been a major research area in robotics. Early studies in this field, in the 70's, were mainly motivated by problems of pattern recognition. The obtained results together with the huge improvement of the available computing power lead us to consider that the recognition, the localization or the inspection of a motionless part are now in today's state-of-the-art. In parallel, and partly owing to the attention paid to mobile robots, researches have investigated more complex questions: stereo vision, outdoor scene analysis,... In complement, major works were done in the domain of the analysis of image sequences (dynamic vision) [4]. The original motivations came from telecommunication (motion-compensated coding) or military (target tracking, moving objects recognition) applications. Associated algorithms were issued from the signal processing area, and rapidly allowed to consider the possibility of a 'real-time' implementation, i.e. of working at the video rate.

If we now return to the domain of robotics, a recent trend is to use exteroceptive non-contact sensors inside the control servo-loops themselves, and not only as sources of data used in higher decision levels. Such an approach may prove to be very useful if it is necessary, for example, to compensate for small
positioning errors, to grasp objects moving on a conveyor belt, to track a seam in arc welding or, more generally, to be adaptable to uncertainties of the environment. Applications of this kind were realized in manipulation robotics as well as in teleoperation or mobile robotics. However, the sensors were mainly optical proximeters [3], [8], [10], [12], [28], or acoustic range finders [2], [16].

In fact, it is also possible to consider a mobile vision sensor as a device able to provide useful information for realizing closed loop control schemes with respect to the environment. Taking into account this particular goal requires:

- the ability to extract from an image the information which is sufficient and pertinent for the completion of the task (in general a positioning task). Note that this may require the design of dedicated targets;
- the implementation of control algorithms simple enough to work at a rate compatible with the desired bandwidth of the closed-loop system, but, further, with an acceptable robustness with regard to unavoidable uncertainties existing about the sensor and the environment.

This approach, known as visual servoing, is the central issue of this paper.

It should be emphasized that this approach differs from the ones referred to as Dynamic Vision approaches, which exploit without controlling it the motion of the sensor or of the objects in the scene. In the Visual servoing approach, a first step consists in defining in the image a particular set of characteristic features which constitutes the goal to be reached. A second step will be to design a control which will ensure the convergence towards the configuration corresponding to the goal image, by starting from a different initial condition.

Early studies in this domain, in the 70's, were mainly based on heuristics [1], [17]. More formalized approaches arose around 1982, setting out two kinds of problems:

- the choice and the extraction of the visual feature elements to be used in the control;
- the analysis and the synthesis of the control schemes with the point of view of automatic control theory.

For the first point, it should be noticed that the characteristic parameters allowing to define the goal image were often selected in relation to the existence of algorithms allowing their extraction within a reasonably short time interval. This is why it is not surprising that most of the performed works use low level primitives like contours: points, segments [19], [11], [15] or regions: surface, barycenter, inertia axis [17], [29]. The original approach of [25] should also be cited: here, the image of a polyhedron is described by a graph, the nodes of which represent the surfaces of the faces and the vertices the lengths of the edges.

For the second point, the most relevant results in the literature come from Carnegie Mellon University [26], [29], [30]. Two kinds of studies have been conducted: the analysis of the mapping between the screw space of the camera motion and the space of velocity fields in the image (which will be called later 'interaction screw'), and the design of control schemes with the study of their