Robotic manipulation with stereo-visual servoing

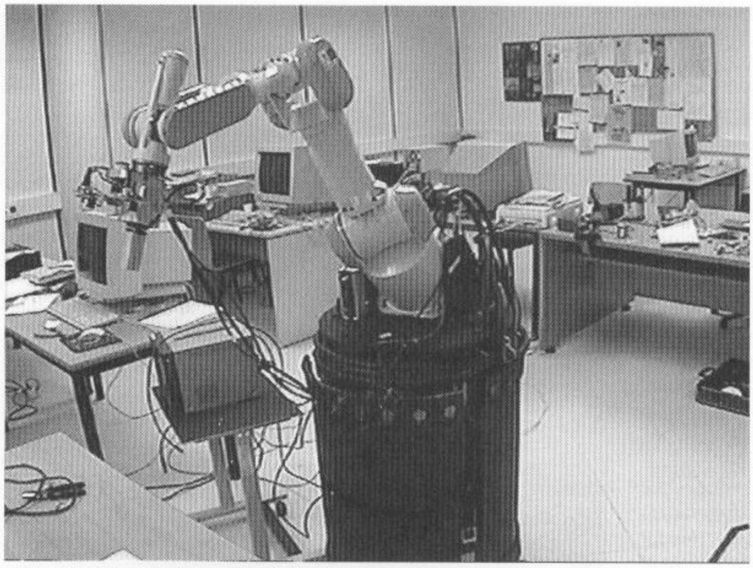

Figure 1. Shown is the set-up of the visual servoing system: a stereo rig mounted on the end-effector of a mobile manipulator.

Visual servoing is an approach to robot control based on visual perception. It involves the use of cameras in the control loop of the robot relative to its environment as required by the task. Its essence is the computation of the matrix of derivatives (the Jacobian) of the visual feature vector with respect to the motion of the robot. Whether to use raw pixel data or estimated 3D-point coordinates is a matter of choice: both approaches require the estimation of camera parameters. However, the resulting dynamic properties of the task may differ. Stereo visual servoing offers some advantages over the classical monocular 2D- and 3D-visual-servoing approaches. Depth information can be recovered without the need for any geometrical model of the observed object. It should be noted that, even in 2D visual servoing, this information would be needed for the computation of the image Jacobian.

In our work we have achieved some theoretical results concerning the modeling of a positioning task.[1] We have tested many approaches to define this kind of task when using a stereo-vision system. The vision system segments the observed object from the scene and computes its center of gravity. The image coordinates of this point, in both cameras, are the output of the vision system to the robot controller. Since 3D coordinates of the observed point can be computed from the image data (and an estimation of the extrinsic and intrinsic parameters of the cameras), they can also be used in the control law, thus getting a linear Jacobian matrix. As a result, some theoretical properties of the trajectory of the end-effector can be obtained.[2] To deal with real objects, orientation has to be taken ii»o account. Classically, axes of inertia are computed from image segmentations. These axes provide information for robot orientation with regard to the object. Though visual servoing formalisms with lines and orientation are theoretically sound, real applications are uncommon due to the difficulty of robustly extracting such primitives.

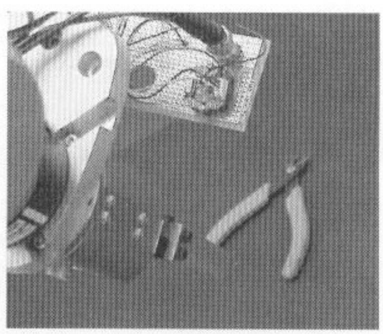

Figure 2. Close view of the end-effector, targeted to a real object (pliers).

The mobile manipulator of the Robotic Intelligence Lab consists of a Nomad XR4000 platform and a Mitsubishi PA-10 arm. Attached to the end-effector of the arm is a stereo rig with two miniature CMOS NTSC color cameras, in an eye-in-hand configuration, linked to two video boards that deliver visual features at video rate. Figure 1 shows the configuration of the system. At this moment only the arm is controlled by vision: in future we hope to control the mobile base this way too. The cameras are coarsely positioned, mounted with approximately the same orientation and at equal distances from the end-effector's origin. No calibration procedure is used. The vision system is made of two inexpensive, off-the-shelf boards (*Cognachrome* by Newton Research Labs). These are video-rate color segmentation systems that extract colored regions from an image and deliver the coordinates of its centroid, its aspect ratio, and the orientation of its major axis of inertia. Each camera is connected to it's own processor.

The work-place is depicted in Figure 2 where pliers lie on a black surface and the robot is observing the object. Pliers are orange-colored, thus the blobs corresponding to the arms are segmented based on color information. Blobs are not symmetric along their inertia axes, but the system is expected to be robust against minor deviations.

Theoretical developments have shown us how to extract 3D control features from stereo images: the Jacobian matrix is computed for raw pixels, 3D coordinates are estimated, and a new feature vector that uses stereo disparity is obtained. Real experiments in adverse conditions (large rotation, noisy images, coarse calibration) show that the trajectory of the end-effector relies strongly on the features chosen for the control loop. Future work should allow us to state more precisely the robustness of the different approaches with respect to camera parameters and signal loss. We are also interested in considering other visual features and integrating them with other types of sensors.[3]

Enric Cervera*, Philippe Martinet†, and F. Berry†
*Robotic Intelligence Lab
Jaume-I University
12071 Castelló, Spain
E-mail: ecervera@icc.uji.es

†LASMEA - GRAVIR
Blaise Pascal University of Clermont-Ferrand
63177 Aubiere-CEDEX, France
E-mail: {martinet, berry}
@lasmea.univ-bpclermont.fr

References

1. E. Cervera, F. Berry, and P. Martinet, *Image-based Stereo Visual Servoing: 2D vs 3D features,* **Proc. IEEE Int'l Conf. on Robotics and Automation**, Washington DC, 2002, *(in press)*.
2. P. Martinet and E. Cervera, *Stacking Jacobians Properly in Stereo Visual Servoing,* **Proc. IEEE Int'l Conf. on Robotics and Automation**, pp. 717-722, Seoul, Korea, 2001.
3. E. Cervera and A. P. del Pobil, *Sensor-Based Learning for Practical Planning of Fine Motions in Robotics,* **Int'l J. of Information Sciences,** Special Issue on Intelligent Learning and Control of Robotics and Intelligent Machines in Unstructured Environments, 2002, *(in press)*.