Reaching a Moving Target Using Optical Flow *

G. Metta(1), A. Oddera(1), and M. Salganicoff(2) and G. Sandini(1)

(1) DIST - University of Genova
Laboratory for Integrated Advanced Robotics (LIRA - Lab)
Via Opera Pia 11A - I16145 Genoa, Italy
e-mail:giulio@vision.dist.unige.it

(2) GRASP-Lab.
University of Pennsylvania
Philadelphia, - U.S.A.
e-mail: sal@grip.cis.upenn.edu

1 Introduction

Most manipulation tasks involving the use of an end effector coupled to a robot arm, require, as a basic skill, the ability to reach a point in space. For example most of the tasks proposed by the organizers of this workshop (such as inserting, catching, digging, unlocking a door, feeding someone with a spoon and others), require, as a part of the overall task, to bring the end effector close to some point in space. This observation becomes even more important considering that, during the approaching phase, vision is the only sensory information usable to adapt the system to changes in environment conditions. The simplest possible example is that of "reaching" a target which moves unpredictably in the environment.

Many solutions have been presented in the literature to solve this problem even in complex, dynamic situations [1, 2]. Almost all systems, however, are bounded by the use of quantitative geometric measures relying, one way or the other, on the knowledge of some calibration parameters. Even if in many practical situations this is not really a constraint, the flexibility of the resulting system is reduced because the procedure required to set-up the system has to undergo a calibration phase involving the extraction of the camera's intrinsic and extrinsic parameters as well as the relationship between the camera and the arm coordinate systems. It is worth noting that, following this approach, even in a "simple" reaching task, part of the calibration procedure must be repeated whenever the end effector is changed or a different tool has to be manipulated (e.g. a spoon or a fork, an hammer or a wrench). This solution is not feasible

*The research described in this paper has been supported by the ESPRIT project SECOND and by a grant from HP-Italia

whenever the manipulated tool is not known a-priori or if the camera and the arm are not mechanically connected (e.g. in case the camera and/or the arm are mounted on moving platforms).

2 The approach proposed

The approach presented in this paper is based on a control structure using visual measures not requiring camera calibration or the knowledge of the camera-arm relationship [3, 4, 5]. Similar approaches have been presented in the past within the framework of active perception [6, 7] and visually guided manipulation [8, 9].

With particular reference to the current presentation relative to a reaching action, the goal of the visually-guided controller is to keep the end effector on an ideal linear trajectory connecting, at each instant of time, the point to be reached with the end effector. The projection of this 3D trajectory on the image plane represents the 2D trajectory that must be followed by the *image* of the moving end-effector in order to reach the goal point (see Figure 1). This is the first constraint: keep the image of the end-effector moving along this ideal 2D trajectory. Reasoning in terms of optical flow vectors, this constraint can be satisfied by minimizing the component of the flow field perpendicular to this ideal trajectory. The component of the optical flow along the ideal trajectory, on the other hand, can be used to control the approaching speed of the end effector by computing the time-to-impact with respect to the goal point. Both measures do not require a calibration of the eye-hand system and of the camera itself. Using these two measures the control is limited to 2 degrees of freedom because the relative distance of the end effector and the point to be reached cannot be computed from a single camera. To overcome this constraints a binocular system is required. However, in order to comply with the initial assumption of an un-calibrated system, depth cannot be computed directly and, for the task described, it is not strictly necessary. In fact, if we assume to be able to retrieve the *disparity* of the image points, the task will be successfully executed if the control action minimizes the difference in disparity between the end effector and the goal point.

Summarizing the overall approach, then, the visual measures necessary to drive the controller are the 2D image velocity field (used to drive the end effector toward the goal point and to control the approaching speed) and the difference in disparity between the end effector and the point to be reached (used to bring the end effector and the goal point on the same disparity plane).

To test this approach experiments have been realized controlling in real-time a robot arm performing a reaching task in unpredictable environments.

The experimental set-up is shown in Fig. 2 and, at this stage, is based on a single camera and the use of optical flow. Two constraints have been used: the arm and the goal points are on the same 2D plane (this plane, however, does not need to be parallel to the image plane); the position of the target in the

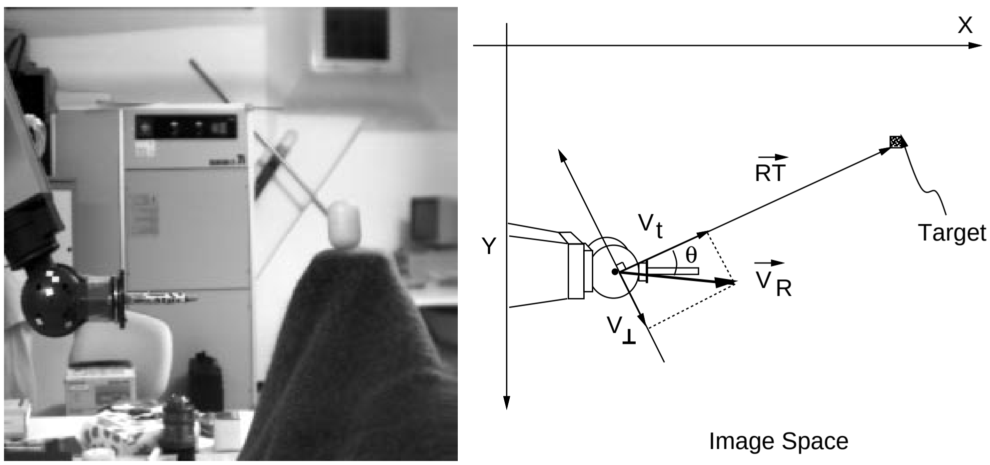

Figure 1: Reference frame. Left: sample image of the puma arm during a reaching task. The end effector and the target lies on the same vertical plane. Right: schematics of the measures used. $\vec{RT}$ is the desired trajectory of the end effector on the image plane. $\vec{V}_R$ is the istantaneous velocity vector obtained by averaging the optical flow of the end effector (This is used to control its approaching velocity). θ is the angle between the desired trajectory and the actual velocity (this controler zeros this angle).

image plane is determined on the basis of its gray level.

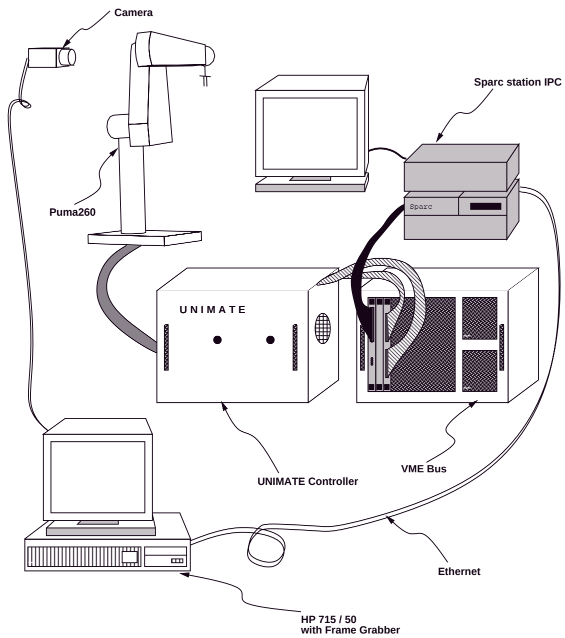

Figure 2: Experimental set-up used for the experiment. The sparc station controls the arm using RCCL and is connected through ethernet to the workstation performing the real-time computation of the optical flow.

The integration of optical flow and disparity has not been completed, however, in addition to what has been presented in a recent paper [4], the computation and use of time-to-impact has been fully tested and included in the control loop to show the importance of this direct measure in the execution of reaching tasks.

The control strategy is implemented as follows: the error

$$e(t) = \theta(t) = atan(\frac{||\vec{V_r} \times \vec{RT}||}{\vec{V_r} \cdot \vec{RT}} sgn(\vec{V_r} \times \vec{RT})) \quad (1)$$

where θ is the angle between $\vec{RT}$ and $\vec{V_r}$.

This is used to control the arm velocity that is useful to be represented in the following way (all these variables are expressed in the robot coordinate system):

$$\begin{cases} u = V cos(\beta) \\ v = V sin(\beta) \end{cases} \quad (2)$$

where V is the amplitude of the velocity, β the angle and (u, v) the cartesian velocity of the end effector. Then the following control law is applied:

$$\beta(t) = k_p e(t) + k_d(e(t) - e(t-1)) \tag{3}$$

The amplitude V is controlled by measuring:

$$e(t) = V - \frac{\|\vec{RT}\|}{t_{cdes}} \tag{4}$$

and acting on the arm acceleration $A = \frac{dV}{dt}$ as follows: (A is expressed in the robot coordinate system):

$$A = k_p e(t) + k_d(e(t) - e(t-1)) \tag{5}$$

this is equivalent to control the time to impact t_c but has a better stability respect to the direct control of the t_c variable:

$$t_c = \frac{\|\vec{RT}\|}{\vec{V}_t} = \frac{\|\vec{RT}\|}{V cos(\theta)} \tag{6}$$

References

[1] A. Conkie and P. Chongstitvatana. An Uncalibrated Stereo Visual Servo System. In *BMVC() Proceedings of the British Machine Vision Conference*, pages 277–280. University of Oxford, September 1990.

[2] Russell L. Andersson. *A Robot Ping-Pong Player.* MIT-Press, 1988.

[3] G. Sandini, F. Gandolfo, G. Metta, and A. Oddera. The role of vision in two-arms manipulation. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, New York, USA, 1993.

[4] M. Salganicoff, G. Metta, A. Oddera, and G. Sandini. A direct approach to vision guided manipulation. In *to appear on: Proc. ICAR-93*, Tokyo - Japan, November 1-2 1993.

[5] M. Salganicoff, G. Metta, A. Oddera, and G. Sandini. A vision-based learning method for pushing manipulation. In *to appear on: AAAI Fall Symposium on Machine Learing in Computer Vision*, Raleigh, NC U.S.A., October 22-24 1993.

[6] R. Bajcsy and C. Tsikos. Perception via manipulation. In *Proc. of the Int. Symp. & Exposition on Robots*, pages 237–244, Sydney, Australia, November 6-10 1988.

[7] P. Allen. *Object Recognition Using Vision and Touch.* Ph.D. Dissertation, University of Pennsylvania, 1985.

[8] N. P. Panikolopoulos and P. K. Khosla. Adaptive robotic visual tracking: Theory and experiments. *IEEE Trans. on Automatic Control*, 38, N. 3:429–445, 1993.

[9] N. P. Panikolopoulos, P. K. Khosla, and T. Kanade. Visual tracking of a moving target by a camera mounted on a robot: A combination of control and vision. *IEEE Trans. on Robotics and Automation*, 9, N. 1:14–34, 1993.