Notes on WP1 and WP2 cross-actions

Giorgio Cannata: UG-DIST

Introduction

These notes are a summary of the discussion held during the III Eyeshots technical meeting (sept. 16th, 2009) and related to activities coupling parts of WP1 and WP2.

Eliminato: hold

Coordinated Control of Ocular Movements

The problem proposed by KUL is that of training a neural-network (NN), based on stereo-vision data, and to use the output of the trained network to drive (using a continuous closed loop control strategy) the gaze direction onto a point of interest in the scene.

The proposed control problem will be first formulated in a kinematic framework and simulated within the integrated simulation environment (currently under development by UG).

The rationale is the following. The NN is modeled as a mapping from stereo input to camera joint coordinates (assuming a generic tilt-pan camera), as sketched in Figure 1:

Commento [PDM1]: Actually, we have only the vergence angle as output since we do not control the gaze (it is assumed constant).

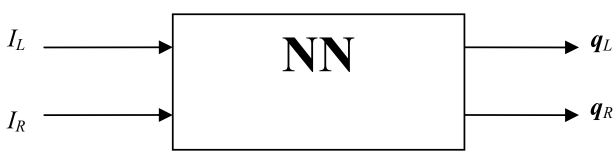

Figure 1

The current approach adopted by KUL to compute the *exact values* of the joint vectors $\boldsymbol{q_L}$ and $\boldsymbol{q_R}$ is a recursive method which could be approximated by the closed loop simulation scheme shown in Figure 2,

Commento [PDM2]: You mean the vergence angle (see my previous comment)

Commento [PDM3]: What does 1/s stand for: integrative control?

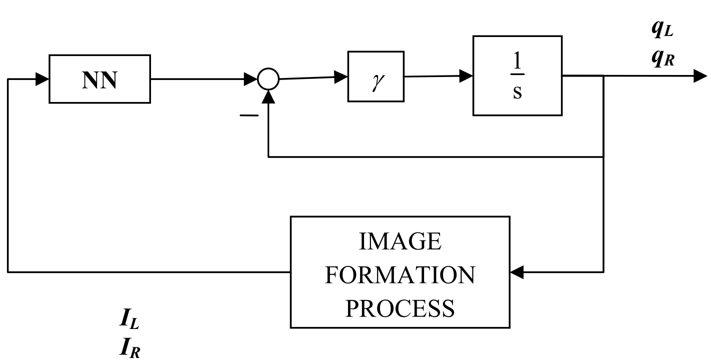

Figure 2

where the feedback gain $\gamma > 0$ will define the closed loop rate of motion of the cameras.

The feedback scheme sketched above mimics the currently adopted method used by KUL, but in a realistic simulation environment could lead to *jerky* camera motions (depending upon the magnitude of the feedback gain γ.

A slightly different, although more satisfactory[1] scheme has been also discussed during the meeting. The idea is that of changing the approach to the design of the NN assuming to associate to it output a *measure of mismatch* of the current images from the *desired* (target) view. Following this idea the output of the net will be a sort of error function based on image (and possibly also camera position) feedback.

According to this assumption, by following a *task function* control loop design (based on the task error defined by the NN output) it would be possible to design the feedback control loop sketched in Figure 3:

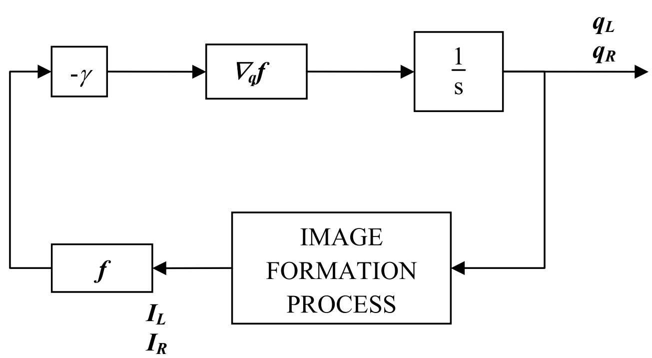

Figure 3

In this scheme according to the (known or estimated) properties of the NN mapping $\boldsymbol{f}$ (and of its gradient) it is possible to ensure a *smooth* convergence of the camera orientations to the target (a priori unknown[2] posture). The drawback of the scheme is related to the requirement of changing the role of the NN mapping, and the need of providing (an estimate of) the Jacobian of the mapping $\boldsymbol{f}$ with respect to the joint angles of the two cameras.

Commento [PDM4]: Well I understood that we need to give a rotational speed (of the eyemotors) as NN output, since now our NN output is in angular terms…

Actions

Implementation of (at least) the first control scheme is expected to start (with support from UG) is expected to start as long as the first release of the integrated virtual simulator will be available[3].

Estimation of relative eye orientation

The coordinated control of the robot eyes (as part of WP1) could require the measurement or the estimate of the orientation of the eyes (gaze) with respect to some (head) fixed reference frame. As a matter of fact the present design does not plan the use of any orientation sensor, furthermore, assuming to use similar control strategies on different robotic eyes (e.g. tilt-pan cameras), uncertainties related to camera placement as well as on the kinematics of the structure may require the adoption of image based estimation procedures to compute at any image time[4] *sampling time* the homogeneous transformation (or at least the rotation matrices) of the left and right eyes with respect to the head, i.e. the matrices:

Eliminato: ag

[1] In terms of predictability of closed loop performances and *quality* of camera motions.
[2] At least in terms of the camera joint angles.
[3] Tests on the kinematic and dynamic ocular models are in progress. Integration with the virtual visual environment should be completed by mid November 2009.
[4] As a matter of fact the control of the robot movements as well as the image acquisition are discrete time processes.

Actually, we had already written minutes in that direction (prepared by Karl), including space-variant mapping:

KUL's approach for disparity estimation under limited accuracy of the gaze direction will be enhanced with the inclusion of a space-variant mapping. This will enable a reduction in the number of scales considered in the coarse-to-fine disparity refinement procedure. The gaze correction procedure will provide gaze information (corrected fundamental matrix) to Giorgio as required for the interactive exploration of the fragment. This will be the subject of Deliverable 2.2a.

$$ {}^H_L H = \begin{bmatrix} {}^H_L R & {}^H p_{L/H} \\ 0^T & 1 \end{bmatrix} $$

$$ {}^H_R H = \begin{bmatrix} {}^H_R R & {}^H p_{R/H} \\ 0^T & 1 \end{bmatrix} $$

Formattato: Abbassato 36 pt

Said T_C the (principal) control loop sampling period, and T_I the sampling time of the image frames, and considering the computational burden related to image processing we shall assume: $T_I = m\ T_C$ where m is a positive integer. Under these assumptions the controller will estimate (on a model basis) the change of posture (rotation and translation during each inter-frame interval, being reset at any image sampling time (i.e. every T_I sec), when a new estimate of the transformation matrices (based on visual feedback is available.

Eliminato: o

Actions

Implementation of the visual based posture estimation algorithms within the simulator and tests with the candidate control algorithms proposed in WP1.