2011 IEEE/RSJ International Conference on
Intelligent Robots and Systems
September 25-30, 2011. San Francisco, CA, USA

# Adaptive Predictive Gaze Control of a Redundant Humanoid Robot Head

Giulio Milighetti  Luca Vallone  Alessandro De Luca

*Abstract*— **A general concept for the gaze control of a redundant humanoid robot head is presented. It is based on an adaptive Kalman filter that predicts the next state of the moving target, processing the position information provided by a head-mounted stereo camera. The trajectory tracking control at the task level combines a proportional feedback and a feedforward term. The gains of both control actions are adapted in order to provide optimal dynamic response for unknown arbitrary target trajectories. Inverse differential kinematics is evaluated so that human-like joint motions are achieved. To exploit kinematic redundancy, a weighted pseudoinverse is realized that takes into account different optimization criteria. Additional self-motions of the head are also considered. Experimental results on the head of the humanoid robot ARMAR-III are presented.**

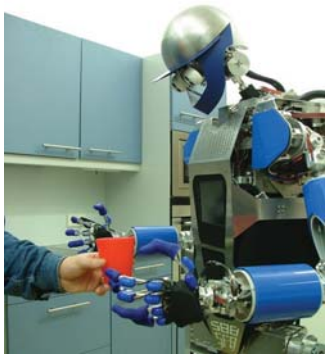


Fig. 1: The ARMAR-III humanoid robot grasping a cup

## I. Introduction

Research in humanoid robotics is engaged with the development of robots with anthropomorphic shape, accompanying the human in his everyday activities and executing a wide spectrum of human-like actions. Advances in many interdisciplinary fields, including mechanics, control, sensing and artificial intelligence, are essential for the success of autonomous humanoid robots.

Like the human, such robots use visual perception as a fundamental component of their sensing capabilities for physical and cognitive interaction with dynamic environments. Vision helps in successfully achieving complex skills dealing with varying and uncertain boundary conditions. A basic behavior used in many everyday tasks (e. g., the interactive grasping in Fig. 1) is the gaze direction control of robot eyes for tracking moving targets. Pointing to an object in the 3D space is a two-dimensional robot task, and requires thus at least two degrees of freedom (dof) devoted to it in the robotic structure. When additional dof are present, the robot becomes kinematically redundant for this task [1].

Many of the developed humanoid robots have heads with only 2 dof and fixed eyes [2], [3], [4]. Heads with human-like motion capabilities have been built, but mainly for studying emotional human-robot interaction [5], [6] or cognitive processes [7], [8], [9]. Robotic heads have been used also to investigate and validate neurophysiological models of human motions, see, e. g., [10], [11]. However, only a reduced number of robot axes are involved and general tracking functionalities are not completely implemented. The approaches in [12] and [13] are based on the consideration of biological experiments. In both cases, a 4-dof head is used and kinematic redundancy is exploited by activating different motion behaviors that reproduce human-like reactions. Because of the specialized decision units, the proposed methods are not easily generalizable to robot heads with a larger number of dof. In [14], the lack of additional dof in the 2-dof head is compensated by using other joints in the robot body, and resorting to optimization criteria aimed at minimizing the motion of parts with larger inertia or maximizing the distance to the joint limits. However, the use of distal dof located in the body rather than local to the head results in less natural motion or inefficient solutions. Finally, in [15] a novel approach has been developed for the ARMAR-III full 7-dof head. Two virtual prismatic joints are introduced, connecting the eyes to the target point to be tracked. However, the two additional dof introduced in the kinematic structure of the head have to comply with an artificial vector constraint imposed in the task space.

In this paper, we present a general concept for the gaze control of the redundant robotic head of the ARMAR-III humanoid robot described in Sect. II. The target objects to be tracked have unknown and arbitrary trajectories with high dynamics. The control concept, detailed in Sect. III, combines a feedback on the task error with varying proportional gain and a feedforward action based on an adaptive Kalman filter that predicts the next state of the moving target. Redundancy is used for shaping the inverse kinematic solution by weighted pseudoinversion of the task Jacobian, taking into account local optimization criteria associated to human-like motion. Tracking experiments are presented in Sect. IV. Section V discusses also on-going research.

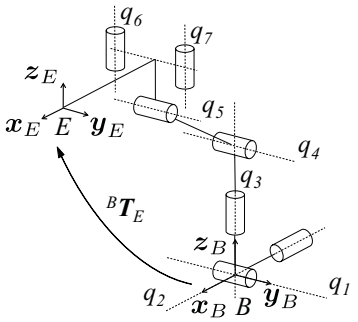
G. Milighetti is with the Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB, Karlsruhe, Germany (giulio.milighetti@iosb.fraunhofer.de)

L. Vallone was with the Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB, Karlsruhe, Germany (luca.vallone85@gmail.com)

A. De Luca is with the Dipartimento di Informatica e Sistemistica, Università di Roma "La Sapienza", Rome, Italy (deluca@dis.uniroma1.it)

978-1-61284-455-8/11/$26.00 ©2011 IEEE

(a) Stand-alone robot head




(b) Full kinematics of the head

Fig. 2: The head of the ARMAR-III humanoid robot

## II. Experimental Setup

The humanoid robot head used for the validation of the proposed gaze control concept has been developed within the SFB 588 "Humanoid Robots". It has been designed to be used both as part of the humanoid robot ARMAR-III and as a stand-alone robot head for studying various perception tasks in the context of object recognition and human-robot interaction [16] (see Fig. 2a). The head has seven dof and its full kinematic scheme is shown in Fig. 2b. The neck motion is realized by four joints, with lower pitch $q_1$, roll $q_2$, yaw $q_3$, and upper pitch $q_4$ angles. The motion of the two eyes is realized by three more joints, with a shared common tilt axis ($q_5$) and two independent rotations around a vertical axis ($q_6$ and $q_7$). All seven joints are driven by DC motors.

The vision system consists of two cameras per eye, one with wide-angle lens for peripheral vision and one with narrow-angle lens for foveal vision. The used Dragonfly®2 cameras capture color images at a frame rate of up to 60 Hz. They have a resolution of 640×480 pixel and are controlled through a FireWire interface [17]. The head is also equipped with an audio system, with a six channel microphone array for 3D localization of acoustic events.

The head control program has been implemented within the MCA2 framework [18] and is executed on a PC with 2 GHz Pentium 4 processor, 2 GB RAM, running under Debian Linux (kernel 2.6.8). The input data for the head control are produced by an image processing algorithm developed at Fraunhofer IOSB. The target object is localized using stereo triangulation between the two narrow-angle cameras, on the basis of its color features in the HSV space. The collected data are sent via UDP connection to the head control PC. The image processing algorithm provides new data every 50 ms, while the control algorithm has a cycle time of $T = 30$ ms and provides commanded joint velocities to the low-level servos, which work at 1 kHz.

## III. Control Concept

The desired task is the visual tracking of a moving target with unknown and arbitrary trajectory by the robot eyes (*robot gazing*), using all available dof of the robot head so as to obtain a natural human-like behavior. The proposed gaze control concept is presented in Fig. 3. It is composed by a feedback and a feedforward (predictive) control action: the feedback term guarantees zero error of the pointing angle vector $\boldsymbol{\theta}$ in static conditions, while the feedforward term anticipates the target motion thus avoiding the head to lag behind. The visual input data are processed by an adaptive Kalman filter that estimates the actual absolute Cartesian position $\boldsymbol{p}_t^f(k)$ of the target at sampling time $t = kT$ and predicts its next position $\boldsymbol{p}_t^f(k+1)$ at time $t = (k+1)T$, both expressed with reference to the head base frame. The kinematic redundancy of the robot head w.r.t. the pointing task is exploited using a weighted pseudoinverse of the task Jacobian, and introducing robot self-motions in the null space of the task Jacobian. The main components of this control concept are illustrated next.

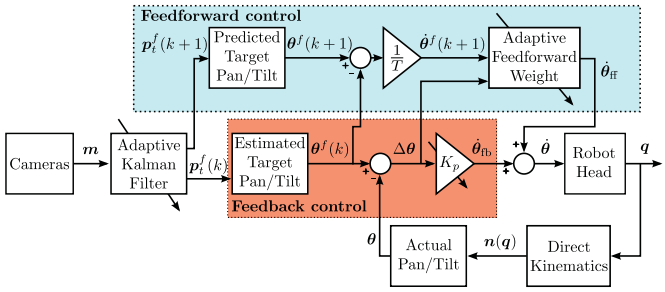



Fig. 3: Proposed gaze control scheme

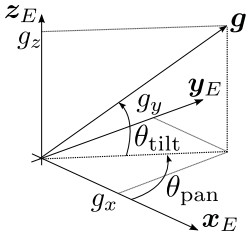



Fig. 4: Definition of pan and tilt angles

### A. Head Kinematics

The *gaze direction* of the robot head is a unit vector $\boldsymbol{g}$ from the origin of the end-effector (camera) frame and pointing to the target. With reference to Fig. 4, this direction is characterized by a two-dimensional vector of pointing angles $\boldsymbol{\theta} = (\theta_{\text{pan}}\ \theta_{\text{tilt}})^T \in (-\pi/2, \pi/2) \times (-\pi/2, \pi/2)$ as

$$\boldsymbol{g} = \begin{pmatrix} g_x \\ g_y \\ g_z \end{pmatrix} = \begin{pmatrix} \cos\theta_{\text{pan}}\ \cos\theta_{\text{tilt}} \\ \sin\theta_{\text{pan}}\ \cos\theta_{\text{tilt}} \\ \sin\theta_{\text{tilt}} \end{pmatrix}. \tag{1}$$

Excluding the boundary values $\pm\pi/2$ of the tilt angle, the adopted description does not suffer any singularity. The pointing angles are evaluated as:

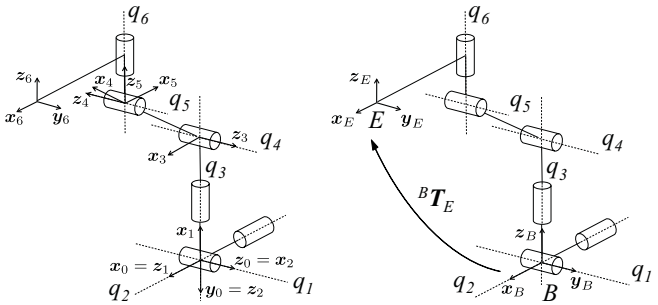
$$\theta_{\text{pan}} = \text{atan}\left(\frac{g_y}{g_x}\right), \quad \theta_{\text{tilt}} = \text{atan}\left(\frac{g_z}{\sqrt{g_x^2 + g_y^2}}\right). \tag{2}$$

The *pointing direction* $\boldsymbol{n}$ of the robot head is chosen as the $\boldsymbol{x}$ axis of the end-effector (camera) frame, and is determined

once the kinematic model of the head is established. In order to avoid the implementation of an on-line calibration of the stereo system, motions of the two eyes are kept parallel. Thus, a simplified kinematic scheme of the head with only 6 dof has been considered (see Fig. 5), without loss of the generality for the proposed control scheme. The direct kinematics of the head is defined by the homogeneous matrix

$$
\begin{aligned}
{}^B\boldsymbol{T}_E(\boldsymbol{q}) &= {}^B\boldsymbol{T}_0\,{}^0\boldsymbol{T}_6(\boldsymbol{q})\,{}^6\boldsymbol{T}_E \\
&= \begin{pmatrix} \boldsymbol{n}(\boldsymbol{q}) & \boldsymbol{s}(\boldsymbol{q}) & \boldsymbol{a}(\boldsymbol{q}) & \boldsymbol{p}(\boldsymbol{q}) \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad (3)
\end{aligned}
$$

where ${}^0\boldsymbol{T}_6(\boldsymbol{q})$ is calculated using the Denavit-Hartenberg parameters in Tab. I, ${}^B\boldsymbol{T}_0$ is a (constant) rotation by $-\pi/2$ around the $\boldsymbol{x}_B$ axis, and ${}^6\boldsymbol{T}_E = \boldsymbol{I}$.




(a) DH-frames (b) Global frames

Fig. 5: Simplified kinematics of the head

TABLE I: Denavit-Hartenberg parameters of the head

| joint | $\alpha_i$ | $a_i$ | $d_i$ | $\theta_i$ |
|---|---|---|---|---|
| 1 | $-\pi/2$ | 0 | 0 | $q_1$ |
| 2 | $\pi/2$ | 0 | 0 | $q_2$ |
| 3 | $\pi/2$ | 0 | $d_3\,(<0)$ | $q_3$ |
| 4 | $\pi$ | $a_4$ | 0 | $q_4$ |
| 5 | $\pi/2$ | 0 | 0 | $q_5$ |
| 6 | 0 | $a_6$ | 0 | $q_6$ |

The head gazes perfectly at the target when $\boldsymbol{n}(\boldsymbol{q}) = \boldsymbol{g}$. Evaluating $\boldsymbol{n}(\boldsymbol{q})$ from eq. (3) and setting $\boldsymbol{g} = \boldsymbol{n}(\boldsymbol{q})$ in eq. (2), in the spirit of the task function approach (see, e. g., [19], [20] for visual servoing applications) the task kinematics is defined as

$$
\boldsymbol{\theta} = \boldsymbol{f}(\boldsymbol{q}) = \begin{pmatrix} \theta_{\mathrm{pan}}(\boldsymbol{q}) \\ \theta_{\mathrm{tilt}}(\boldsymbol{q}) \end{pmatrix}. \qquad (4)
$$

The associated differential kinematics is

$$
\dot{\boldsymbol{\theta}} = \frac{\partial \boldsymbol{f}(\boldsymbol{q})}{\partial \boldsymbol{q}}\,\dot{\boldsymbol{q}} = \boldsymbol{J}(\boldsymbol{q})\dot{\boldsymbol{q}}, \qquad (5)
$$

where the elements $J_{ij}$ of the $2 \times 6$ task Jacobian $\boldsymbol{J}$ are computed as

$$
J_{ij} = \frac{\partial \theta_i}{\partial q_j}, \quad i = \{\text{pan, tilt}\}, \quad j = 1, \ldots, 6. \qquad (6)
$$

As a result, the degree of redundancy of the robot head in the gazing task is equal to four.

### B. Adaptive Kalman Filter

Kalman filter has been extensively used in the past in order to reduce the noise of input data from stereo cameras, cope with their asynchronous entering, and estimate target motions (see, e. g., [21], [22]). Revisiting this idea, a linear discrete-time Kalman filter (cf. [23, p. 362]) has been implemented. Since we wish to track a target that moves along a priori unknown trajectories with high dynamics, a critical issue is the choice of an appropriate model of target motion to be used within the filter. A target motion model based on constant accelerations [24] has been found suitable for our purposes, in association with a proper adaptation of some relevant filter parameters. Defining the state vector

$$
\boldsymbol{\xi} = \begin{pmatrix} x & \dot{x} & \ddot{x} & y & \dot{y} & \ddot{y} & z & \dot{z} & \ddot{z} \end{pmatrix}^T, \qquad (7)
$$

and denoting with $\boldsymbol{m}$ the output measures of the camera, the discrete-time linear model of the observed target motion is

$$
\begin{aligned}
\boldsymbol{\xi}(k+1) &= \boldsymbol{A}\,\boldsymbol{\xi}(k) + \boldsymbol{v}(k) \qquad (8) \\
\boldsymbol{m}(k) &= \boldsymbol{H}\,\boldsymbol{\xi}(k) + \boldsymbol{w}(k), \qquad (9)
\end{aligned}
$$

where the state transition matrix $\boldsymbol{A} = \mathrm{diag}\{\boldsymbol{A}_i\}$ has

$$
\boldsymbol{A}_i = \begin{pmatrix} 1 & T & T^2/2 \\ 0 & 1 & T \\ 0 & 0 & 1 \end{pmatrix}, \qquad i = x, y, z,
$$

the measurement output matrix $\boldsymbol{H} = \mathrm{diag}\{\boldsymbol{H}_i\}$ has

$$
\boldsymbol{H}_i = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}, \qquad i = x, y, z,
$$

and where the state and measurement noises $\boldsymbol{v}$ and $\boldsymbol{w}$ are white, zero-mean Gaussian stochastic processes driving the system. Equations (8–9) are used within the Kalman filter, having state $\hat{\boldsymbol{\xi}}$, both for the estimation of the current target position at $t = kT$, $\boldsymbol{p}_t^f(k) = (\hat{x}(k)\ \hat{y}(k)\ \hat{z}(k))^T$, and for the prediction of the next target position at $t = (k+1)T$, $\boldsymbol{p}_t^f(k+1) = (\hat{x}(k+1)\ \hat{y}(k+1)\ \hat{z}(k+1))^T$.

The noises $\boldsymbol{v}$ and $\boldsymbol{w}$ are characterized by block diagonal covariance matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$, respectively. In particular, for the state covariance matrix $\boldsymbol{Q} = \mathrm{diag}\{\boldsymbol{Q}_i\}$ it is

$$
\boldsymbol{Q}_i = \sigma_{\boldsymbol{Q}_i}^2 \begin{pmatrix} T^5/20 & T^4/8 & T^3/6 \\ T^4/8 & T^3/3 & T^2/2 \\ T^3/6 & T^2/2 & T \end{pmatrix}, \quad i = x, y, z,
$$

where $\sigma_{\boldsymbol{Q}_i}$ is the standard deviation (see [25]).

In order to increase the robustness of the estimation filter with respect to target trajectories having variable dynamics, an adaptation law for the matrix $\boldsymbol{Q}$ has been implemented. Adapting the values $\sigma_{\boldsymbol{Q}_i}$, and thus the magnitude of $\boldsymbol{Q}$, influences the reliability of the model: the lower are these values, the more reliable is the assumed model (8) of target motion. The adaption law has been chosen as an affine function of the difference $\Delta\boldsymbol{p}_t^f = \boldsymbol{p}_t^f(k) - \boldsymbol{p}_t^f(k-1)$ between the current estimate and the previously predicted target position:

$$
\sigma_{\boldsymbol{Q}_i} = a_{\sigma_{\boldsymbol{Q}}} \left| \Delta p_{t,i}^f \right| + b_{\sigma_{\boldsymbol{Q}}}, \quad \text{for } i = x, y, z. \qquad (10)
$$

The resulting profile is shown in Fig. 6.

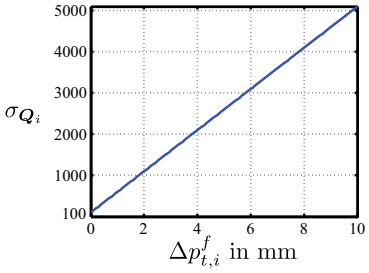



Fig. 6: Adaptation law for standard deviations in (10)

### C. Control Law

The gaze of the robot is controlled by specifying the desired task velocity $\dot{\boldsymbol{\theta}}$ in (5) through a feedback and a feedforward action:

$$\dot{\boldsymbol{\theta}} = \dot{\boldsymbol{\theta}}_{\text{fb}} + \dot{\boldsymbol{\theta}}_{\text{ff}}. \tag{11}$$

*1) Proportional feedback term:* The goal of the proportional feedback term is to correct the current pointing error $\Delta\boldsymbol{\theta}$, given by the difference between the task variable $\boldsymbol{\theta}^f$ computed from the estimated target position $\boldsymbol{p}_t^f$ and its value $\boldsymbol{\theta}$ associated to the current head configuration $\boldsymbol{q}$. More specifically, the value $\boldsymbol{\theta}^f$ is obtained from $\boldsymbol{g}^f = \boldsymbol{p}_t^f - \boldsymbol{p}(\boldsymbol{q})$ by using eq. (3), and then evaluating eq. (2) for $\boldsymbol{g} = \boldsymbol{g}^f / \|\boldsymbol{g}^f\|$. The error $\Delta\boldsymbol{\theta}$ may be due to a (transient) positioning error w.r.t. a static target, as well as to the linearization error involved in the on-line inversion of the differential kinematics (see Sect. III-D) when tracking a moving target.

In a discrete-time implementation, it is:

$$\dot{\boldsymbol{\theta}}_{\text{fb}}(k) = \boldsymbol{K}_p \, \Delta\boldsymbol{\theta} = \boldsymbol{K}_p \, (\boldsymbol{\theta}^f(k) - \boldsymbol{\theta}(k)), \tag{12}$$

where $\boldsymbol{K}_p > 0$ is a $2 \times 2$ diagonal gain matrix.

An adaptive gain has been introduced, in order to achieve a good response both when far from the actual trajectory of the target (large pointing error) and when already following it (small pointing error). Accordingly, the adaptation law has been defined as a quadratic function of the amplitude of the scalar angular errors $\Delta\theta_{\text{pan}}$ and $\Delta\theta_{\text{tilt}}$:

$$K_{p,i} = a_{\boldsymbol{K}_p} \, \Delta\theta_i^2 + b_{\boldsymbol{K}_p} \, \Delta\theta_i + c_{\boldsymbol{K}_p}, \quad \text{for } i = \{\text{pan}, \text{tilt}\}. \tag{13}$$

The resulting profile is shown in Fig. 7.

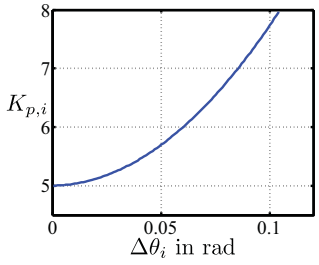



Fig. 7: Adaptation law for proportional gains in (13)

*2) Weighted feedforward term:* The feedforward term uses the task value $\boldsymbol{\theta}^f(k+1)$ associated to the predicted target position $\boldsymbol{p}_t^f(k+1)$ for anticipating the target motion so that the robot head does not lag behind. It is:

$$\dot{\boldsymbol{\theta}}_{\text{ff}}(k) = \boldsymbol{K}_{\text{ff}} \, \frac{\boldsymbol{\theta}^f(k+1) - \boldsymbol{\theta}^f(k)}{T}. \tag{14}$$

In (14), the $2 \times 2$ diagonal gain matrix $\boldsymbol{K}_{\text{ff}}$ has been introduced for coping with possible instabilities due to incorrectly predicted target positions. The chosen adaptive weighting disables the feedforward term (separately for the pan and tilt components, so for each $K_{\text{ff},i} \in [0,1]$) in the following two situations, see Fig. 8:

- when the target task velocity $\dot{\theta}_i^f(k+1)$ is predicted to be *low* and the pointing error $\Delta\theta_i$ is *small*, because the object is moving slowly and the feedback action should be already sufficient to keep following it;
- when the target task velocity $\dot{\theta}_i^f(k+1)$ is predicted to be rather *high* and for *any* level of the pointing error $\Delta\theta_i$, because velocities that are too high to be admissible are typically associated to an error in the estimates and/or measurements.

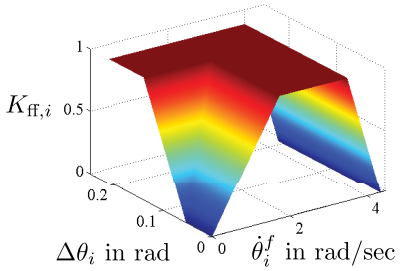



Fig. 8: Weighting law for feedforward gains in (14)

### D. Inverse Differential Kinematics

Once the pan/tilt task velocity control terms in (11) have been evaluated, the joint velocities $\dot{\boldsymbol{q}}$ of the robot head are computed by inverting the task using the differential kinematics (5). Since the task Jacobian in (6) is not square, the simplest way is to resort to the minimum joint velocity norm solution given by its pseudoinverse,

$$\dot{\boldsymbol{q}} = \boldsymbol{J}^{\dagger}(\boldsymbol{q}) \left( \dot{\boldsymbol{\theta}}_{\text{fb}} + \dot{\boldsymbol{\theta}}_{\text{ff}} \right). \tag{15}$$

However, this may not comply with the existence of physical bounds in the joint space and with the sought human-like behavior of the robot head. To these purposes, and wishing to limit additional complexity, redundancy of the head has been exploited mainly through the use of a suitably weighted pseudoinverse, i. e.,

$$\dot{\boldsymbol{q}} = \dot{\boldsymbol{q}}_W = \boldsymbol{J}_W^{\dagger}(\boldsymbol{q}) \left( \dot{\boldsymbol{\theta}}_{\text{fb}} + \dot{\boldsymbol{\theta}}_{\text{ff}} \right), \tag{16}$$

where

$$\boldsymbol{J}_W^{\dagger}(\boldsymbol{q}) = \boldsymbol{W}^{-1}(\boldsymbol{q}) \boldsymbol{J}^T(\boldsymbol{q}) \left( \boldsymbol{J}(\boldsymbol{q}) \boldsymbol{W}^{-1}(\boldsymbol{q}) \boldsymbol{J}^T(\boldsymbol{q}) \right)^{\dagger} \tag{17}$$

with a $6 \times 6$ diagonal, configuration-dependent weighting matrix $\boldsymbol{W} > 0$. Note that pseudo-inversion of the rightmost $2 \times 2$ matrix in (17) is considered in place of its standard inversion, so as to handle possible loss of rank of the task Jacobian $\boldsymbol{J}$. Moreover, the computation of the $\boldsymbol{J}_W^\dagger$ matrix can be performed by a simple modification of the (damped) SVD algorithm for computing $\boldsymbol{J}^\dagger$. The diagonal elements $W_i \in [1, \infty)$ of $\boldsymbol{W}$ are defined using the following criteria:

- *Avoidance of joint limits.* The joints should work as near as possible to the center of their (possibly non-symmetric) working ranges $[q_{\mathrm{min,i}}, q_{\mathrm{max,i}}]$, for $i = 1, \ldots, 6$. According to [26], the separable cost function

$$H(\boldsymbol{q}) = \sum_{i=1}^{6} \frac{(q_{\mathrm{max,i}} - q_{\mathrm{min,i}})^2}{4\,(q_{\mathrm{max,i}} - q_i)\,(q_i - q_{\mathrm{min,i}})} = \sum_{i=1}^{6} H_i(q_i), \tag{18}$$

  can be associated with the closeness to the joint limits. Therefore, its gradient $\nabla_{\boldsymbol{q}} H = (\partial H / \partial \boldsymbol{q})^T$ can be used to weigh the pseudoinverse. Distinguishing whether the joints were moving (at $t = (k-1)T$) towards or away from their limits, the weights $W_i^{\mathrm{pos}}$ are defined as

$$W_i^{\mathrm{pos}} = \begin{cases} 1 + \left| \frac{\partial H_i}{\partial q_i} \right| & \text{if } \dot{q}_i(k-1)\,(q_i(k) - \bar{q}_i) \geq 0 \\ 1 & \text{else,} \end{cases} \tag{19}$$

  being $\bar{q}_i = q_{\mathrm{max,i}} - q_{\mathrm{min,i}}$ the range center of joint $i$.

- *Avoidance of joint velocity limits.* This criterion helps to keep the joint velocities away from their (always symmetric) limits $[-\dot{q}_{\mathrm{max,i}}, \dot{q}_{\mathrm{max,i}}]$. Similarly to (18), we define

$$U(\dot{\boldsymbol{q}}) = \sum_{i=1}^{6} \frac{\dot{q}_{\mathrm{max,i}}^2}{\dot{q}_{\mathrm{max,i}}^2 - \dot{q}_i^2} = \sum_{i=1}^{6} U_i(\dot{q}_i), \tag{20}$$

  and the associated weights $W_i^{\mathrm{vel}}$ are specified, taking into account the previous joint velocity samples, as

$$W_i^{\mathrm{vel}} = \begin{cases} \left| \frac{\partial U_i}{\partial \dot{q}_i} \right| & \text{if } \dot{q}_i(k-1)\dot{q}_i(k-2) \geq 0 \\ 0 & \text{else.} \end{cases} \tag{21}$$

- *Joint prioritization.* In order to realize human-like behavior and take advantage of the faster speed of the eyes (as in saccadic motion), a joint motion priority can be established by tuning a factor $f_i^W \geq 1$ for each joint. Using (19) and (21), the diagonal elements of $\boldsymbol{W}$ will be

$$W_i = f_i^W \left( W_i^{\mathrm{pos}} + W_i^{\mathrm{vel}} \right), \quad i = 1, \ldots, 6. \tag{22}$$

The largest is $f_i^W$, the lowest will be the priority associated to the motion of joint $i$.

Finally, in order to obtain a more comfortable static (or quasi-static) posture, self-motions $\dot{\boldsymbol{q}}_0$ that do not affect the gaze direction have been activated when the head motion is slow or next to stop. Following the human, a comfortable posture has the joint positions closer to the centers of their range position. The cost function to be locally minimized is

$$H_0(\boldsymbol{q}) = \frac{1}{12} \sum_{i=1}^{6} \left( \frac{q_i - \bar{q}_i}{q_{\mathrm{max,i}} - q_{\mathrm{min,i}}} \right)^2, \tag{23}$$

and the joint velocity $\dot{\boldsymbol{q}} = \dot{\boldsymbol{q}}_W - k_0\,\dot{\boldsymbol{q}}_0$ is obtained from (16) and the projected gradient method (see, e. g., [1]) as

$$\dot{\boldsymbol{q}} = \boldsymbol{J}_W^\dagger(\boldsymbol{q}) \left( \dot{\boldsymbol{\theta}}_{\mathrm{fb}} + \dot{\boldsymbol{\theta}}_{\mathrm{ff}} \right) - k_0 \left( \boldsymbol{I} - \boldsymbol{J}_W^\dagger(\boldsymbol{q}) \boldsymbol{J}(\boldsymbol{q}) \right) \nabla_{\boldsymbol{q}} H_0(\boldsymbol{q}). \tag{24}$$

The scalar gain $k_0 \geq 0$ activating the self-motion has been chosen as a cubic function of the non-homogeneous joint velocities $\dot{\boldsymbol{q}}_W$

$$k_0 = a_{k_0} \|\dot{\boldsymbol{q}}_W\|^3 + b_{k_0} \|\dot{\boldsymbol{q}}_W\|^2 + c_{k_0} \|\dot{\boldsymbol{q}}_W\| + d_{k_0}. \tag{25}$$

A typical profile is shown in Fig. 9.

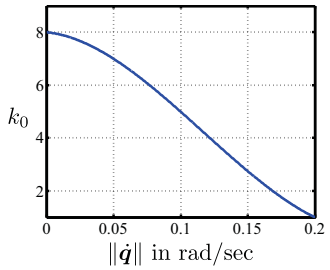



Fig. 9: Gain for self-motion activation in (25)

## IV. Experimental Results

For illustrating the performance of the proposed gaze control concept, we present here representative results from an extensive experimental campaign that has been carried out with different target trajectories. For more details on data and results, please refer to [27].

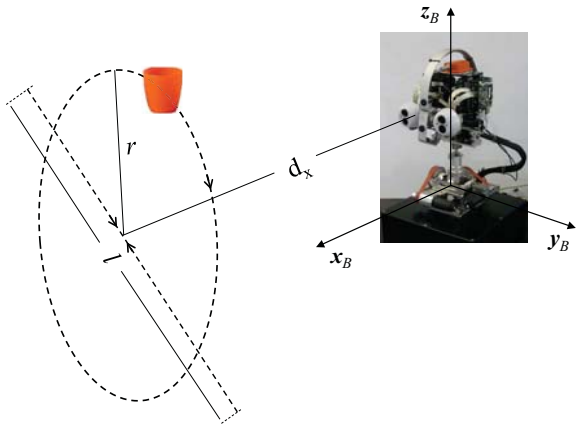



Fig. 10: Target trajectory for the experimental validation

The considered trajectory consists of a series of rotations along an ellipse having major half-axis $r \approx 0.4$ m, followed by some back and forth linear motion along a segment of length $l \approx 1.15$ m, see Fig. 10. The whole path belongs roughly to a plane parallel to $(\boldsymbol{y}_B, \boldsymbol{z}_B)$, placed at a distance $d_x \approx 1.10$ m from the robot head. The motion is executed with varying velocity by a human hand holding a cup (the target). The maximum target velocity is about 2 m/sec.

The gazing task was first executed using only the feedback control action, with equal and constant gains $K_{p,i} = 5\ \mathrm{s}^{-1}$

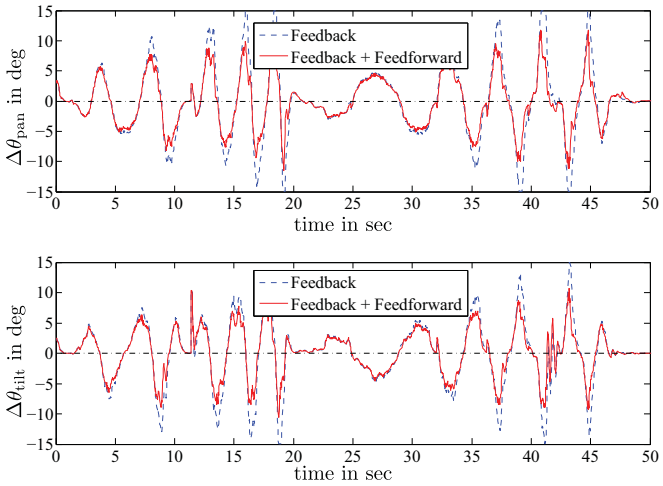



Fig. 11: Errors on task angles, with and without the feedforward control action

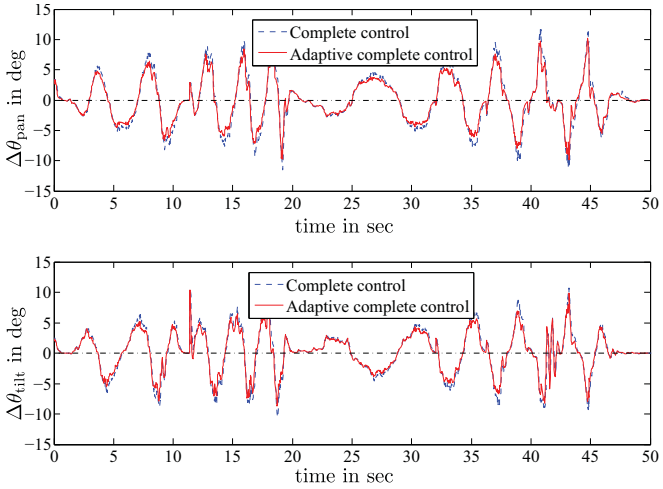



Fig. 12: Errors on task angles, with and without adaptive gains

for both pan and tilt error components. After some preliminary experiments, the standard deviations of the measurement covariance matrices along $x$, $y$ and $z$ were set in the Kalman filter as follows: $\sigma_{\boldsymbol{R}} = (15.6, 8.8, 5.0)$ mm. Moreover, the simple pseudoinversion (15) was used, i. e., with $\boldsymbol{W} = \boldsymbol{I}$. A damping factor $\lambda = 0.1$ has been used in the SVD algorithm. The feedforward control action was added next, with unitary factors $K_{\text{ff},i}$. The comparison of pointing errors in Fig. 11 shows that the feedforward reduces the lag of the head, as expected. The improvement is more noticeable at motion direction changes, where the reduction of both pointing errors is about 5°.

Further improvements have been obtained through the adaptation of both the proportional gain $\boldsymbol{K}_p$ and the feedforward weight $\boldsymbol{K}_{\text{ff}}$, according to Figs. 7–8. The obtained pointing errors are given in Fig. 12. As before, differences can be noticed especially at motion direction changes, in particular with higher speeds. In this case, peak errors decrease by approximately 3°.

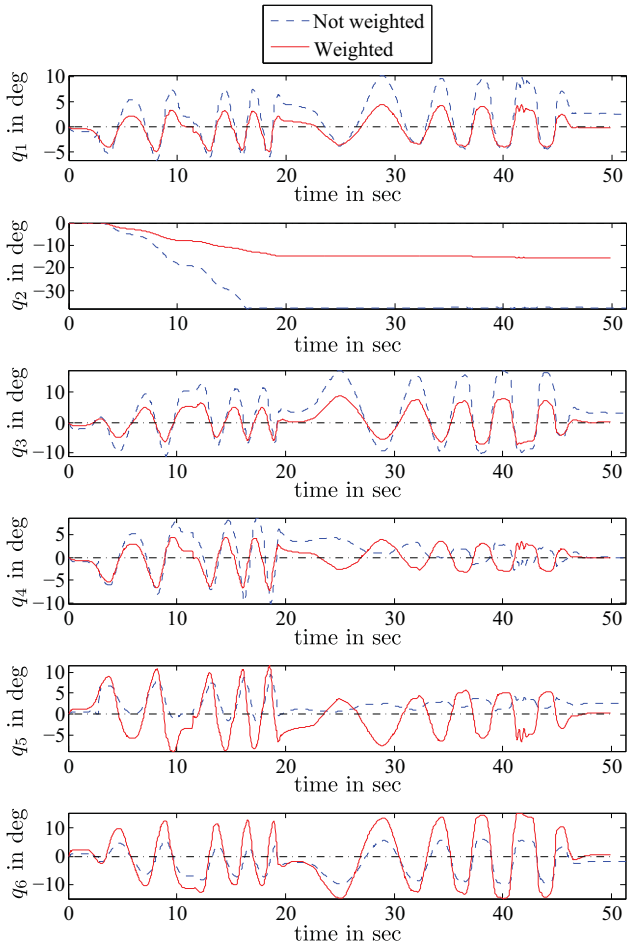



Fig. 13: Joint evolutions, with and without a weighting matrix $\boldsymbol{W}$ in the kinematic pseudoinversion

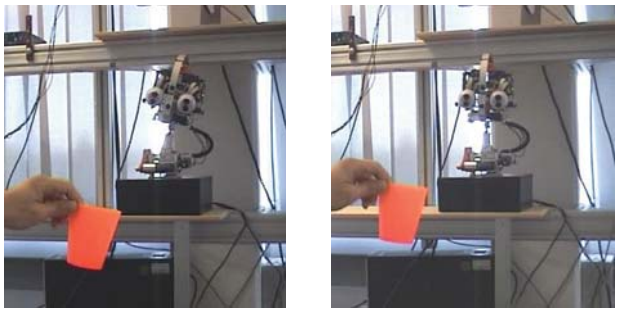
In order to better exploit the redundancy of the robot head, a weighted pseudoinversion and then the use self-motions have been considered, as introduced in Sect. III-D. Figure 13 shows the different behaviors of the joints obtained when weighted pseudoinversion was used. Due to joint prioritization (i. e., using $f_1^W = f_2^W = f_3^W = 4$, $f_3^W = 3$, and $f_5^W = f_6^W = 1$ in eq. (22)), the motion range of the last two joints ($q_5$ and $q_6$) associated to the eyes increases considerably. On the other hand, the second joint avoids now the saturation to its lower position limit.

Finally, the benefit of adding a self-motion is clearly appreciable when the target stops. As an example, Figure 14 shows two different head configurations, one reached when the target ends its motion (Fig. 14a), the other obtained after the execution of a self-motion (Fig. 14b). In the latter case, the head assumes a more natural posture.

The performance reached with the presented control concept can be also appreciated in the accompanying video.

## V. Conclusions

Our new concept for the gaze control of a redundant humanoid robot head uses a Kalman filter to process data from a stereo camera so as to localize the position and predict



(a) Without self-motion (b) With self-motion

Fig. 14: Influence of self-motion on the head configuration

the motion of a target. An adaptation of the state noise covariance matrix increases the robustness of estimations with respect to target trajectories having large dynamic variations. The kinematic control law at the task velocity level uses a feedback term proportional to the pointing error and a feedforward term based on the predicted motion of the target. Adapting both control gains provides better dynamic responses in general conditions. Inverse kinematics is solved at the differential level by means of a suitably weighted pseudoinverse of the task Jacobian, considering joint position and velocity limits, as well as joint prioritization. Experimental results show that the joint motion of the robot head mimics that of a human.

We are currently considering the head as a part of a walking humanoid robot. In this case, the oscillatory movements of the hip additionally affect the accuracy of the gazing task. The egomotion of the head can be estimated by means of image information and compensated using the neck redundancy. In this way, an active image stabilization is performed and human-like behavior of the full body is expected. At present, the functionality of the human hip is provided by a Stewart platform on which the head is placed [28]. First promising results have been obtained using a priority-based multi-task approach.

## VI. Acknowledgements

This work is part of the cooperation between the Fraunhofer Institute of Optronics, System Technologies and Image Exploitation (IOSB) and the Department of Computer and System Sciences (DIS) of the University of Rome "La Sapienza", in the framework of the Collaborative Research Center (SFB) 588 "Humanoid robots" supported by the Deutsche Forschungsgemeinschaft (DFG) [29].

## References

[1] Y. Nakamura, *Advanced Robotics: Redundancy and Optimization*. Reading, MA, USA: Addison-Wesley, 1991.

[2] S. Sakagami *et al.*, "The intelligent ASIMO: System overview and integration," in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Lausanne, CH, Oct. 2002, pp. 2478–2483.

[3] K. Akachi *et al.*, "Development of humanoid robot HRP-3P," in *Proc. 5th IEEE-RAS Int. Conf. on Humanoid Robots*, Tsukuba, JPN, Dec. 2005, pp. 50–55.

[4] Fujitsu Laboratories. (2008) Humanoid robot HOAP-3. [Online]. Available: http://jp.fujitsu.com/group/labs/downloads/en/business/activities/activities-4/fujitsu-labs-robotics-005-en.pdf

[5] H. Miwa *et al.*, "Development of a new human-like head robot WE-4," in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Lausanne, CH, Oct. 2002, pp. 2443–2448.

[6] L. Aryananda and J. Weber, "MERTZ: A quest for a robust and scalable active vision humanoid head robot," in *Proc. 4th IEEE-RAS Int. Conf. on Humanoid Robots*, Los Angeles, CA, USA, Nov. 2004.

[7] H. Kozima and H. Yano, "A robot that learns to communicate with human caregivers," in *Int. Work. on Epigenetic Robotics*, Lund, S, 2001.

[8] R. Beira *et al.*, "Design of the robot-cub (iCub) head," in *Proc. IEEE Int. Conf. on Robotics and Automation*, Orlando, FL, USA, May 2006, pp. 94–100.

[9] G. Cheng *et al.*, "CB: A humanoid research platform for exploring neuroscience," *Advanced Robotics*, vol. 21, no. 10, pp. 1097–1114, 2007.

[10] T. Shibata and S. Schaal, "Biomimetic gaze stabilization based on feedback-error-learning with nonparametric regression networks," *Neural Networks*, vol. 14, no. 2, pp. 201–216, Mar. 2001.

[11] L. Manfredi *et al.*, "Implementation of a neurophysiological model of saccadic eye movements on an anthropomorphic robotic head," in *Proc. 6th IEEE-RAS Int. Conf. on Humanoid Robots*, Genova, I, Dec. 2006, pp. 438–443.

[12] F. Du *et al.*, "Gaze control for a two-eyed robot head," in *Proc. British Machine Vision Conf.*, Glasgow, UK, Sept. 1991.

[13] L. Gu and J. Su, "Gaze control on humanoid robot head," in *Proc. 6th World Congr. on Intelligent Control and Automation*, Dalian, PRC, June 2006.

[14] F. Faber, M. Bennewitz, and S. Behnke, "Controlling the gaze direction of a humanoid robot with redundant joints," in *Proc. 17th IEEE Int. Symp. on Robot and Human Interactive Communication*, Munich, D, Aug. 2008, pp. 413–418.

[15] D. Omrcen and A. Ude, "Redundant control of a humanoid robot head with foveated vision for object tracking," in *Proc. IEEE Int. Conf. on Robotics and Automation*, Anchorage, AK, USA, May 2010, pp. 4151–4156.

[16] T. Asfour, K. Welke, P. Azad, A. Ude, and R. Dillmann, "The Karlsruhe humanoid head," in *Proc. 8th IEEE-RAS Int. Conf. on Humanoid Robots*, Daejeon, KR, Dec. 2008, pp. 447–453.

[17] Point Grey Reasearch, Inc. (2010) DragonFly2. [Online]. Available: http://www.ptgrey.com/products/dragonfly2/

[18] MCA2, Modul Controller Architecture Version 2. [Online]. Available: http://www.mca2.org/

[19] B. Espiau, F. Chaumette, and P. Rives, "A new approach to visual servoing in robotics," *IEEE Trans. on Robotics and Automation*, vol. 8, no. 3, pp. 313–326, June 1992.

[20] S. Hutchinson, G. Hager, and P. Corke, "A tutorial on visual servo control," *IEEE Trans. on Robotics and Automation*, vol. 12, no. 5, pp. 651–670, Oct. 1996.

[21] Y. Chan, A. Hu, and J. Plant, "A Kalman filter based tracking scheme with input estimation," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 15, no. 2, pp. 237–244, Mar. 1979.

[22] P.-O. Gutman and M. Velger, "Tracking targets using adaptive Kalman filtering," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 26, no. 5, pp. 691–699, Sept. 1990.

[23] P. Antsaklis and A. Michel, *Linear Systems*. New York: McGraw-Hill, 1997.

[24] A. Kiruluta, M. Eizenman, and S. Pasupathy, "Predictive head movement tracking using a Kalman filter," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 27, no. 2, pp. 326–331, 1997.

[25] K. Mehrotra and P. Mahapatra, "A jerk model for tracking highly maneuvering targets," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 33, no. 4, pp. 1094–1105, Oct. 1997.

[26] T. Chan and R. Dubey, "A weighted least-norm solution based scheme for avoiding joint limits for redundant joint manipulators," *IEEE Trans. on Robotics and Automation*, vol. 11, no. 2, pp. 286–292, Apr. 1995.

[27] L. Vallone, "Adaptive predictive gaze control of a redundant humanoid robot head," Master's thesis, Università di Roma "La Sapienza", Rome, I, Sept. 2010.

[28] G. Milighetti *et al.*, "Mobile experimental platform for the development of environmentally interactive control algorithms towards the implementation on a walking humanoid robot," in *Proc. 6th German Conf. on Robotics (Robotik 2010)*, Munich, D, June 2010.

[29] Sonderforschungsbereich 588. (2001) "Humanoid Robots" - Learning and cooperating multimodal robots. Karlsruhe, Germany. [Online]. Available: http://sfb588.ira.uka.de/