Biomimetic Eye-Neck Coordination

Manuel Lopes, Alexandre Bernardino, José Santos-Victor, Kerstin Rosander and Claes von Hofsten

***Abstract*—We describe a method for coordinating eye and neck motions in the control of a humanoid robotic head. Based on the characteristics of human oculomotor behavior, we formulate the target tracking problem in a state-space control framework and show that suitable controller gains can be either manually tuned with optimal control techniques or learned from bio-mechanical data recorded from newborns subjects. The basic controller relies on eye-neck proprioceptive feedback. In biological systems, vestibular signals and target prediction compensate for external motions and allow target tracking with low lag. We provide ways to integrate inertial and prediction signals in the basic control architecture, whenever these are available. We demonstrate the ability of the method in replicating the behavior of subjects with different ages and show results obtained through a real-time implementation in a humanoid platform.**

I. Introduction

There is an increasing interest in advanced human-robot interfaces due to a growing need for "service robots", designed to perform a variety of tasks in human inhabited environments. Head and eye movements are particularly important for human-humanoid interaction, because they constitute a highly attended and communicative part of the human body. A key issue for the acceptance of such systems by generic users lies on the ability to generate human-like motions and postures. In this paper we address the generation of human-like head-eye movements in the control of a humanoid robot head.

Two types of ocular motions are predominant in the visuomotor control system. *Smooth-pursuit* motions control the eyes' velocity to keep a moving target stable on the retina, but their maximum velocities are limited to about 30 deg/sec [1]. *Saccadic* motions perform fast changes of the gaze direction whenever the target moves quicker or when the focus of attention has to be changed between interest points. In both types of motions, the neck and the eyes move coordinately to perform the task. Due to different bio-mechanical dynamic characteristics, the eyes are quicker in achieving the goals set by the neurological commands. In the control of saccades, eyes rapidly achieve the target point. Then, the head moves slower toward the target, while the eyes counter rotate to compensate head motion [2]. Studies in the control of smooth pursuit in infants show that head movements lag with respect to the eyes, even with predictable target trajectories [3]. We propose a model for head-eye coordination capable of generating and explaining these two types of behaviors. The same type of controller, but with different gains, can be used for smooth-pursuit and saccadic eye-head coordination.

In humans several stimuli and (neuro-)physiological subsystems contribute to coordinate eyes and neck motions: visual information, acoustic cues, afferent and efferent copy signals, the vestibular system, motion prediction, etc. The overall system is of great complexity and we do not aim, in this paper, to provide a neuro-biological compatible and complete model. Instead we intend to model the behavior of biological gaze control systems as observed by an external agent. For the design of human-humanoid interfaces, it is the external robot behavior that influences human perception, rather than the driving internal models.

In newborns tracking capabilities develop very fast. Neonates can follow small targets only with saccadic motions but already exhibit smooth pursuit motions in the presence of large moving objects. In the beginning these behaviors are limited to very low angular speeds and no prediction seems to be present [3]. In the first weeks of age, newborns increase significantly the smooth pursuit gain and become able to coordinate eyes and neck. The neck still moves much less than the eyes but its gain increases with age. Interesting to note is that around 3 months, the vestibulo-ocular reflex (VOR) system is already fully functional, including inhibition when both the target and the head are moving coordinately [4]. Without such inhibition the tracking quality would be reduced due to the VOR operation (actuation of the eye in opposition to head movements).

There are a few attempts to model and implement biomimetic eye-head coordination systems in real robotic systems. A mathematical model based on psychophysical studies of head-free gaze saccades in human subjects is presented in [5]. It models the coordination between eye and head movements during rapid orienting gaze shifts for targets presented in a two-dimensional frontal plane, under both aligned and unaligned initial fixation conditions. Recent work [6] has implemented that model in a 7 degree of freedom anthropomorphic robot head that mimics the physical dimensions, performance and functional abilities of humans. The implemented model shows coherent behavior with respect to the reported patterns of eye-head coordination in humans. However the model has not been applied to the execution of smooth-pursuit movements. In [7], a model of oculomotor control includes the vestibular ocular reflex, the optokinetic nystagmus, predictive control and feedback-error learning. However, in that work, head-eye coordination was performed with simple trajectory planning methods and no concern on modeling human behavior.

With respect to the state-of-the-art, this paper contributes

A. Bernardino, M. Lopes and J. Santos-Victor are with the Institute for Systems and Robotics, Instituto Superior Técnico, Lisbon, Portugal. Email: {macl,alex,jasv}@isr.ist.utl.pt

K. Rosander and C. von Hofsten are with the Dept. of Psychology, Uppsala University, Sweden. Email: {kerstin.rosander,claes.von_hofsten}@psyk.uu.se

978-1-4244-4118-1/09/$25.00 2009 IEEE

Authorized licensed use limited to: UNIVERSIDADE TECNICA DE LISBOA. Downloaded on December 29, 2009 at 13:00 from IEEE Xplore. Restrictions apply.

in the following points: (i) proposal of a basic eye-head coordination scheme formulated with state-space methods suitable for both saccadic and smooth pursuit movements (ii) a method for learning of controller gains from human data recordings, (iii) incorporation of inertial and prediction measurements as additions to the basic controller (iv) operation under horizontal and vertical displacements, and (v) easy implementation on current humanoid robot technology.

The paper is organized as follows. Next section, Sec. II, presents our basic controller model. Then, Section III explains how its gains can be tuned either from optimal control methods or learned from human subjects' recordings. Section IV extends the basic model with inertial measurements and predictive tracking methods. In Section V we present the iCub platform [8] used in our experiments and results obtained in a real-time implementation. Finally, in Section VI we present the conclusions and ideas for future work.

II. Eye-Neck Coordination Controller

In this section we present our controller. In accordance with [6] we consider eye and head plant dynamics as pure integrators. Both for the control of saccades and smooth-pursuit, only one of the eyes's measurements is considered. Since saccades and smooth-pursuit are conjugate eye movements, the information from one eye only is enough to drive the full system.

A. Head-Eye State-Space Model

We adopt a parameterization for orientations in azimuth and elevation angles, and so consider them as decoupled. Later, we will relate these coordinates with the specific eye and neck degrees of freedom. We explicitly represent the orientation of gaze $\vec{g}$ and neck/head $\vec{n}$ in a fixed (body centered) reference frame. The relative orientation of the eye with respect to the head $\vec{e}$ is implicitly defined by the subtraction of the neck orientation from the gaze orientation:

$$\vec{e} = \vec{g} - \vec{n}$$

In short we define our state vector as:

$$X = \begin{bmatrix} \vec{g} \\ \vec{n} \end{bmatrix}$$

For the control design, we use as input variables the eye and neck azimuth and elevation velocities. We denote the gaze, neck and eye angular velocities as ω_g, ω_n and ω_e. For the implementation on the robot we have to transform these velocities to the actual joint level velocities.

Considering that eye and head plants respond as pure integrators (angular position is the time integration of the angular velocity commands) we use a zero-hold discretization method with sampling period T and obtain the following state-space model describing the dynamics of our system:

$$\begin{aligned} X(n+1) &= X(n) + B_d U(n) \\ &= X(n) + \begin{bmatrix} T & T \\ 0 & T \end{bmatrix} \begin{bmatrix} \omega_e(n) \\ \omega_n(n) \end{bmatrix} \end{aligned} \quad (1)$$

where n is the discrete time variable ($t = nT$).

B. Controller Structure

We adopt the full-state feedback controller structure with reference input presented in [9] and illustrated in Fig. 1. The

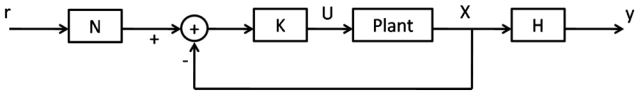

Fig. 1. Gaze feedback controller with reference input: general structure.

command input U is computed by multiplying a gain vector K by the difference between the current state X and a desired state depending on the reference command r:

$$U(n) = -K\left(X(n) - Nr(n)\right) \quad (2)$$

The reference command is set at each time step to the measured orientation of the target, $\vec{t}$ in body centered coordinated. In our problem, we are willing to control both the gaze and the neck direction towards the target, so $N = \begin{bmatrix} 1, & 1 \end{bmatrix}^T$. This provides a frontal gaze behavior where the relative eye angle tends to zero in steady state. However, by appropriately tuning vector N other configurations can be achieved. For instance, $N = \begin{bmatrix} 1, & 0.5 \end{bmatrix}^T$ will provide a solution where half of the displacement is performed by the neck and the remaining performed by the eyes.

Vector K sets the feedback interactions between gaze and neck position to eye and neck velocities:

$$K = \begin{bmatrix} k_{ge} & k_{ne} \\ k_{gn} & k_{nn} \end{bmatrix}$$

Inputs to the plant are then computed by:

$$\begin{aligned} \omega_e &= k_{ge}e + k_{ne}n \\ \omega_n &= k_{ne}e + k_{nn}n \end{aligned}$$

Substituting (2) in (1) we obtain the controlled system equations:

$$X(n+1) = X(n) - B_d K X(n) + B_d K N r(n) \quad (3)$$

III. Controller Design

The dynamical properties of the full control system can be determined by its gain vector K. There are many control design paradigms that allow the computation of suitable gains for linear systems (see [9] for a review). In our case, we want the full control system to match the characteristics of human behavior. To do so we can either try to define some appropriate constraints in the control system to follow human characteristics, or use human recordings to learn the best matching K. In the following part we will pursue the first approach. The subsequent part will address learning from human recordings.

Authorized licensed use limited to: UNIVERSIDADE TECNICA DE LISBOA. Downloaded on December 29, 2009 at 13:00 from IEEE Xplore. Restrictions apply.

A. Optimal Control

We would like our controller to fulfill the following constraints:

- minimum gaze error
- eye dynamics faster than neck dynamics
- maintain target image in the center of the retina
- limited energy expenditure (control values).

These criteria can be formulated as the minimization of a cost function J weighing the commanded values with weight matrix R and weighing tracking error by matrix Q:

$$J = \sum_{n=0}^{\infty} \left(X^T Q X + U^T R U \right) \tag{4}$$

The entries of R will describe the extra energy expenditure in the neck due to the larger inertia that it has to move. By minimizing cost function J we will obtain a control gain K providing the best trade-off between the constraints. Such solution can be obtained by standard optimal control techniques (LQR - see [9]) with proven stability guarantees

B. Controller Learning from Recordings

The controller design approach explained above requires some manual tuning of the gain matrices Q and R in (4) to achieve a reasonable human-like behavior. Given the availability of human biomechanical signals it is, in principle, possible to estimate controller gains that best approximate the acquired data. With a time sequence of state variables we can estimate the gain matrix K that best fits the data. This can be seen as an inverse optimal control problem [10]. After some manipulation Eq. 3 reduces to:

$$B_d^+ \left(X^{n+1} - X^n \right) = K \left(Nr - X^n \right)$$

A least-squares approximation can be used to estimate K. B_d^+ represents the pseudo-inverse of B_d. As it is not square it is computed using the Moore-Penrose pseudo inverse.

It remains to know if the structure of the proposed controller is flexible enough to provide good approximations to the human motions. In Section V we will show that good approximations to human behavior can be achieved using this approach.

IV. Adding VOR and Prediction

Two important characteristics of the human oculomotor control system are the effective rejection of external disturbances and the ability to control gaze without lag, whenever target trajectories are regular. The first characteristic is provided by the vestibulo-ocular reflex (VOR) while the latter exploits target motion prediction. In this section we describe how the basic feedback control model can be extended to incorporate these two characteristics.

A. Inertial compensation (VOR)

The vestibulo-ocular reflex compensates involuntary head motions by moving the eyes with a velocity opposite to the measured disturbance. This can not be a simple reactive behavior that couples head accelerations (measured by the vestibular systems in the inner ear in humans) because in that case a voluntary motion of the neck would be automatically compensated by the eye. The system must be able to distinguish what accelerations are due to self controlled motions from the ones that are externally driven.

In robots the vestibular information is gathered by an inertial measurement unit (IMU) usually installed in a head fixed reference frame, thus measuring the head velocity ω_h. This velocity can be decomposed in the neck self-generated motion ω_n and the independently generated body velocity ω_b, taken here as an external disturbance:

$$\omega_h = \omega_n + \omega_b$$

Therefore, a body motion velocity estimate can be obtained by discounting the afferent signal (the motion command of the previous time step) from the inertial sensor measurement:

$$\hat{\omega_b}(n) = \omega_i(n) - \omega_n(n-1)$$

A command velocity with opposite signal is then added to the eyes control value, and the original control rule of (2) becomes:

$$U(n) = -K \left(X(n) - Nr(n) \right) - D\hat{\omega_b}(n) \tag{5}$$

Matrix D, in the present case, is the vector $[1,\, 0]'$, because we are considering that all external disturbances are compensated by the eyes (vestibulo-ocular reflex). Notwithstanding, this formulation allows the implementation of head stabilization by neck compensation of external disturbances (the vestibulo-collic reflex [11]), if the values of vector D are appropriately set. A diagram of the full control system with inertial compensation is shown in Fig. 2.

B. Prediction

When target trajectories are regular, humans have an unmatched ability to predict the short-term future locations of the target and show an anticipatory behavior. In many cases smooth pursuit movements show no lag between target and gaze directions, which is an essential mechanism to keep the image of the target stable on the retina. Here we describe how our basic model can be extended with prediction mechanisms and explore two control theoretical mechanism that provide low steady state tracking error: integral control and predictive control.

1) Integral Control: Linear control theory shows that the number of integrators in the feedback loop of a system determines the classes of signals that can be tracked with zero steady-state error [9]. For instance a system with one integrator in the loop is able to asymptotically converge to a step-input, and a system with two integrators is able to converge to a ramp input. In our setting, the eye and neck plants have one integrator each. They are therefore able to match the target orientation when it is still. If it moves with a constant velocity then we have to add an additional integrator to the feedback loop of each plant. At first sight, having to change the control structure just to be able to track a single type of movement may seem a bit limiting. However many classes of motions

Authorized licensed use limited to: UNIVERSIDADE TECNICA DE LISBOA. Downloaded on December 29, 2009 at 13:00 from IEEE Xplore. Restrictions apply.

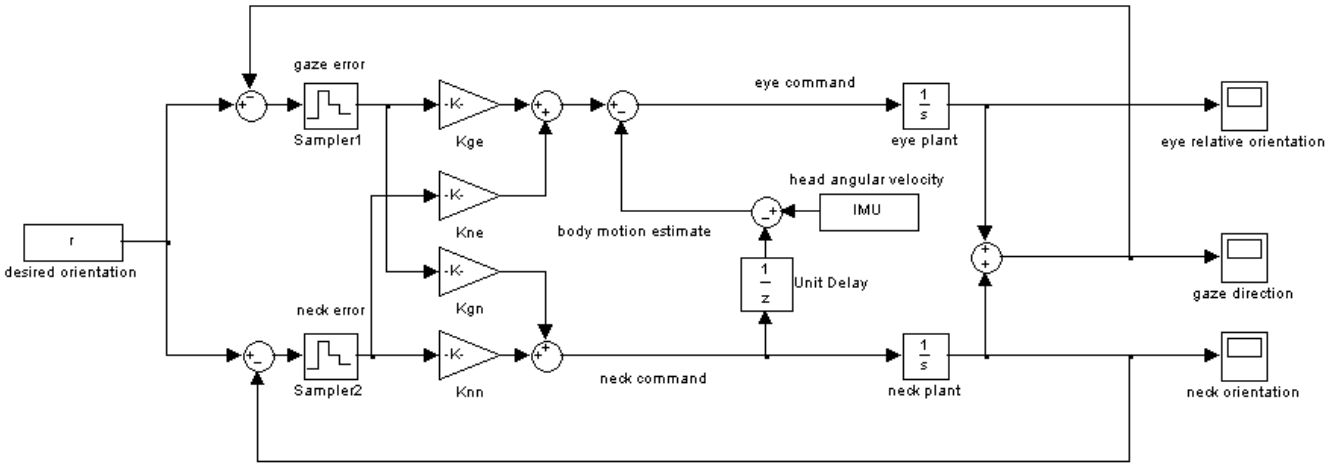

Fig. 2. Block diagram of the eye-head coordination controller with VOR.

can be locally approximated by almost constant velocity (low acceleration) trajectories, thus integral control will improve tracking in such cases.

Integral control requires the accumulation of the gaze error along time:

$$i(n) = i(n-1) + (g(n) - r(n))$$

The controller output is then computed as:

$$\begin{aligned} \omega_e &= k_{ge}e + k_{ne}n + k_{ie}i \\ \omega_n &= k_{ne}e + k_{nn}n + k_{in}i \end{aligned}$$

Gains k_{ie} and k_{in} are manually tuned to avoid overshoot in the response.

2) Predictive Tracking: While the previous approach requires changing the structure of the controller, in predictive tracking we just change the value of the reference orientation to some point in the future. The prediction horizon (how much in the future we predict) depends on the time the system takes to respond to commands, taking into account its latencies, delays and dynamical time constants. No prediction would be required, if the system converges instantaneously to the desired set-points, but, in practice, every causal plant has some kind of delay. Predictive tracking addresses this problem by looking ahead in time and controlling the plant to converge to the future position of the target.

There is a vast literature about trajectory estimation and prediction methods (see [12] for a survey). Most of them have parameterized models of typical trajectories and, given the input data, estimate the parameters. Common models are constant velocity, constant acceleration, circular, curvilinear [12], ballistic [13] and periodic [14]. There are methods that run many models in parallel and choose the best fit, at each time windows, or fuse the outputs of several of them according to their likelihoods [15]. To illustrate the application of this methodology in our architecture, we employ a simple constant acceleration model for target motion:

$$\begin{bmatrix} p_t(n+1) \\ v_t(n+1) \\ a_t(n+1) \end{bmatrix} = \begin{bmatrix} 1 & T & T^2/2 \\ 0 & 1 & T \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_t(n) \\ v_t(n) \\ a_t(n) \end{bmatrix} + \eta(n) \quad (6)$$

where p_t, v_t and a_t are the coordinates (azimuth or elevation) of, respectively, the estimated target angular positions, velocities and accelerations. Noise vector η is assumed Gaussian. The observation measurement is also assumed to be corrupted with white noise:

$$\vec{t}_n = p_t(n) + \xi(n) \quad (7)$$

For this model, standard linear Kalman filtering techniques [15] are employed to estimate the target state and predict its future values in time. The set-point orientation introduced as reference in the controller is no longer the current target prediction $\vec{t}$ but a prediction $p_t^h(n)$, where h is the prediction horizon.

V. RESULTS

In this section we present results illustrating the performance of the proposed methods. We have performed experiments on a real humanoid robot head with controllers either tuned by optimal control methods or learned from human data recordings.

We have used the iCub humanoid robotic head [8], with six degrees of freedom. The head can be positioned with arbitrary orientation through the activation of tilt, pan and swing joints. Eyes have independent pan and a common tilt. The head's mechanical characteristics were designed in order to replicate the typical motions exhibited by young children.

Tests were performed at the BABYLAB of the Uppsala University Department of Psychology, where several devices for generating and measuring precise target and body motions are available. Figure 3 shows the robot prepared for experiments with the BABYLAB setups. One of the setups consists in a board where a hidden motorized arm with a magnet is able to produce arbitrary motions of a visible target. We call this setup

Authorized licensed use limited to: UNIVERSIDADE TECNICA DE LISBOA. Downloaded on December 29, 2009 at 13:00 from IEEE Xplore. Restrictions apply.

the "magnetic board". The second setup is based on a chair inside a cylinder. It has the capability of rotating the chair, the cylinder or both simultaneously, with several periodic motions. This setup generates precise angular speeds and has been extensively used for evaluating infants oculomotor behavior. Chair motions allow the stimulation of the vestibular systems and the verification of VOR performance. We call this setup the "moving drum".

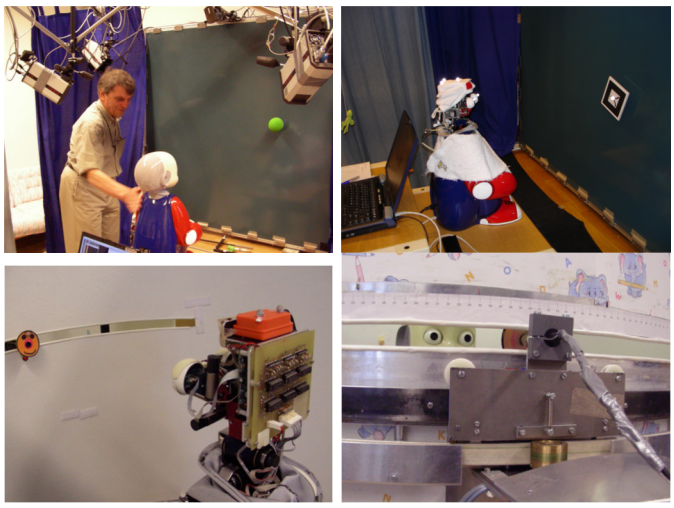

Fig. 3. The iCub Robot Head in BABYLAB (Uppsala 2007-2008). In the top figures, we can see the "magnetic board" setup, where arbitrary motions in a 2D plane can be applied to a target attached to a ferromagnetic material. In the bottom figures we show the "drum setup" that can perform coordinated target and body motions.

The first set of experiments demonstrate the tracking capabilities of the empirically tuned controller gains using optimal (LQR) design. The parameters used were:

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 10 \end{bmatrix} \quad R = \begin{bmatrix} \frac{1}{100} & 0 \\ 0 & 10 \end{bmatrix} \tag{8}$$

resulting in the following gain matrix:

$$K = \begin{bmatrix} 7.8 & -0.91 \\ 0.0069 & 0.97 \end{bmatrix}$$

R is chosen to minimize energy. As the eyes have less inertia, they cost less energy and so the weight for the neck has to be larger. Q is tuned to improve tracking reliability. As the eyes are faster it is better to weight more the eye tracking error, forcing them to move faster, with less expenditure of energy. It is interesting to note that the feedback gain from gaze error to neck velocity is very low, in accordance to the psychophysics results in [5].

Figure 4 show the system's performance when tracking a circular motion in the "magnetic board" (2D plane), eliciting both azimuth and elevation movements. The initial period of the experiment corresponds to a saccade at 5 sec due to system startup. It is visible that gaze has faster dynamics than the neck, exhibiting the typical human behavior. When the system reaches steady state, the tracking error decreases to zero (during the period 10-15 sec target and gaze directions are identical). The saccade has a slow response because the set of controller gains were tuned for smooth-pursuit behavior. In practice, different sets of gains should be used for different types of motions. After that period, the experiment corresponds to a circular motion with constant angular velocity, eliciting sinusoidal smooth-pursuit motions. Notice that neck lags the gaze direction and has smaller gain. This is also typical in human behavior. However, steady state error is not driven to zero because no prediction mechanisms were used at this stage.

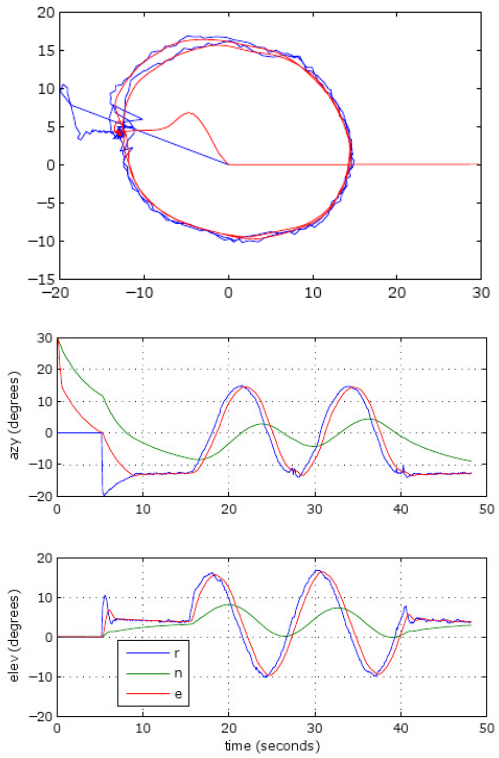

Fig. 4. Tracking a circle. A target, initially at position $(0, 0)$ (azimuth, elevation) moves suddenly (at time 5 sec) to a point in the path of a circular trajectory with diameter 25. At time 15 the target starts moving along the circular trajectory with an angular velocity of about $0.1 cycles/sec$. The plot in the top of the figure shows the reconstruction of the target (blue) and gaze (red) headings (plot units in degrees). The middle and bottom plots show, respectively, the time evolution of azimuth and elevation angles of the target (blue) neck (red) and gaze (red).

The following experiments test the performance of the two proposed techniques for prediction: integral control and predictive tracking. In the first test, a target with a triangular wave motion profile is put in front of the robot. Tracking performance is shown in Figure 5. It is noticeable that in the constant velocity periods, the error decreases. A residual error remains for two reasons. First the duration of the ramp is not long enough to allow the error further decrease. Secondly, there is a computational delay that makes the robot perceive an error smaller that in reality. In fact, the error in the retinal plane was very close to zero but due to the computational delay, in reality the target is already ahead of the position where the image was acquired. Since the plots were obtained with an external motion capture measurement system, this delay is noticeable.

In Figure 6 we show results using the predictive tracking

Authorized licensed use limited to: UNIVERSIDADE TECNICA DE LISBOA. Downloaded on December 29, 2009 at 13:00 from IEEE Xplore. Restrictions apply.

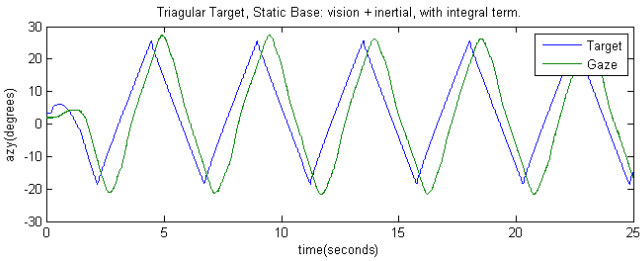

Fig. 5. Integral control. A target with triangular trajectory (piecewise constant velocity is tracked by the robot). During the constant velocity segments the tracking error is reduced via the integral effect.

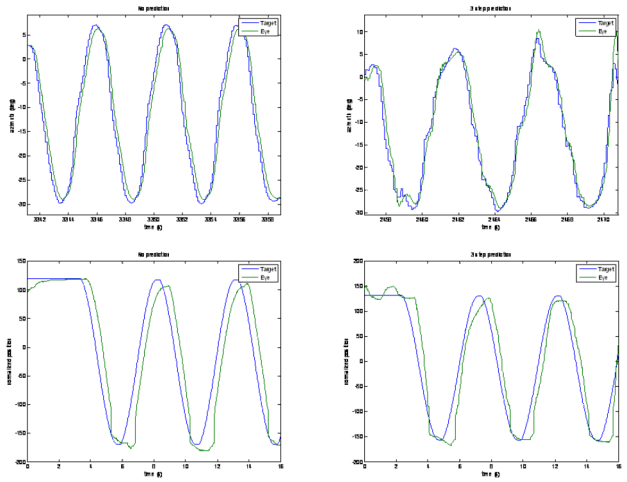

Fig. 6. Predictive tracking. Plots in the top row show values computed by the robot perception. In the bottom row orientations were measured with an external motion capture device. Left: no prediction. Right: 3 steps ahead prediction. Note: Time scales are not aligned.

method with a constant acceleration model and a sinusoidal target motion with a period of 5 sec. We compare the situations without prediction and when the prediction horizon is 1, 2, and 3 time steps. Figure 6 shows the results obtained both from an external measuring device (motion capture) and from a robot's perception point of view, i.e. the target orientation in the fixed reference frame is computed by composing the target orientation in the eye frame with the current eye orientation. Such computations introduce noise in the measurement due to some lack of synchronism in image and head angles' measurements which is reflected in the control for large prediction horizons (prediction involves differentiation which amplifies noise). By computing the phase of the signals' Fourier transforms we were able to estimate the average delay in tracking. The values are shown in Table I.

TABLE I
AVERAGE DELAYS IN TRACKING A SINUSOID FOR DIFFERENT PREDICTION HORIZONS.

Prediction horizon	Average delay
0	112 ms
1 step (67 ms)	96 ms
2 steps (133 ms)	70 ms
3 steps (200ms)	52 ms

The delay decreases steadily when the prediction horizon increases but for prediction horizons bigger that 3, noise degrades significantly the control signal. To address this issue, a better synchronization in the control architecture must be achieved. Again, due to the computational delay, the measurements with the external motion capture device show an additional lag (around 100 ms).

A. VOR

To evaluate the implemented VOR in the compensation of external motions we have performed experiments where the body was moved with a sinusoidal motion profile and the target was static. Figure 7 shows results of the tracking system with and without VOR compensation. The values shown in the plot are written in a body (moving) centered reference frame. The tracking error reduces significantly (about 8 dB) when VOR is active.

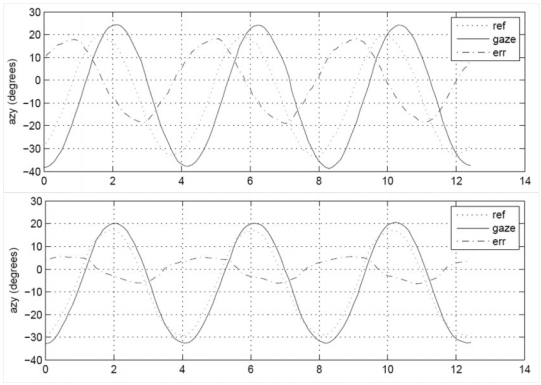

Fig. 7. VOR compensation. Tracking a static target from a rotating platform.

B. Human-data matching

The BABYLAB of the Uppsala University investigates how smooth pursuit and visual-vestibular interactions develop in young infants [4]. For that research several data sets were acquired, with the trajectories of targets, eyes and head recorded along time. In some cases the targets are moving; in other cases the infant's body is rotated and the target is still; on other cases both the target and the chair are moving. We have used some of these recordings, with static body, to learn the gains of our biomimetic controller for smooth-pursuit motions, using the methodology explained in Section III-B. In Figure 8 we can observe temporal plots of babies of different ages (2, 4, and 7 weeks) tracking a sinusoidal target, and the values predicted by the learned controller.

We also show the overall motion decomposition into eye and neck motions, as performed by the babies and by our model. It is visible that the controller reasonably matches babies performance, despite the noise and irregularities in babies signals. This shows that our controller has enough structure to represent the behavior of human tracking. The approximation quality increases with babies age, mainly because baby tracking becomes more regular (less noisy). Table II quantitatively presents the error between the model and the newborn data.

Authorized licensed use limited to: UNIVERSIDADE TECNICA DE LISBOA. Downloaded on December 29, 2009 at 13:00 from IEEE Xplore. Restrictions apply.

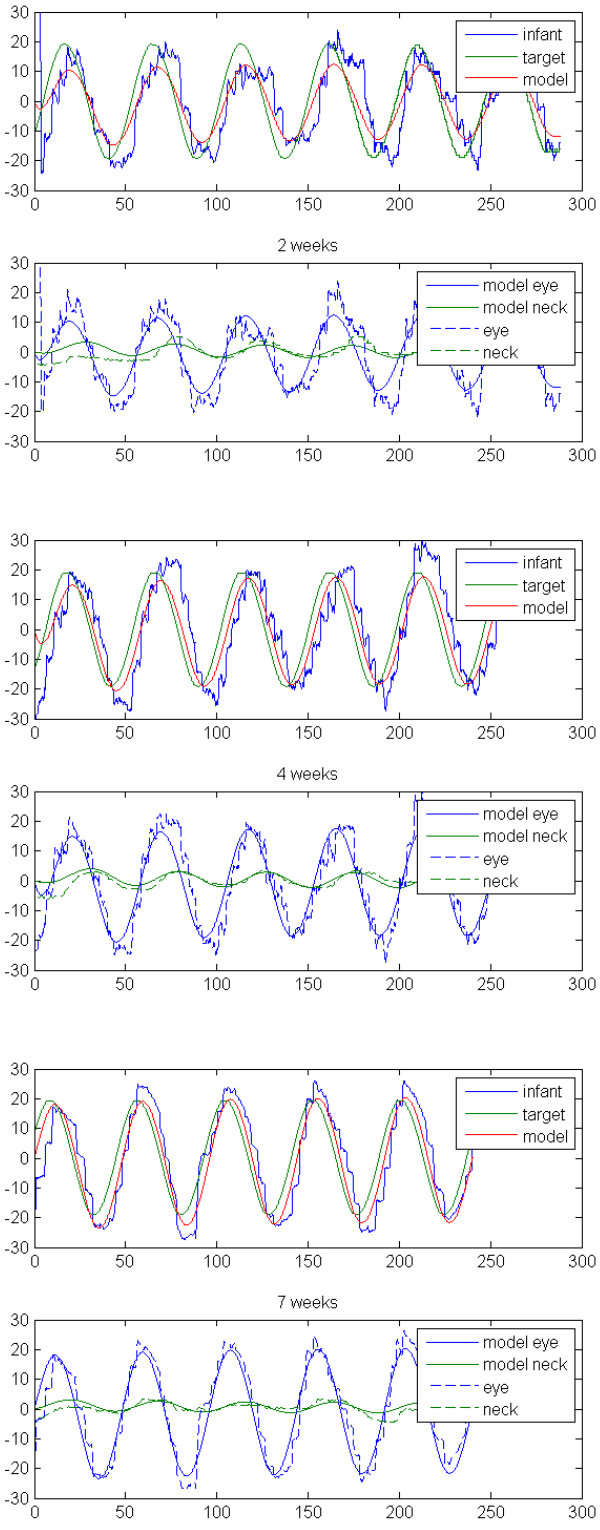

Fig. 8. Matching between model and infant data at 2, 4 and 7 weeks. For each age we show the tracking behavior of infant and model, and the distribution between eyes and neck.

TABLE II
ROOT-MEAN-SQUARE TRACKING ERROR WITH AGE. UNITS ARE DEGREES.

Age (months)	Infant-to-Target	Model-to-Target	Model-to-Infant
2	12.15	6.01	8.61
4	12.19	5.73	7.74
7	9.95	5.60	5.74

VI. CONCLUSIONS

In this work we presented a control architecture for coordinating the eye and neck motions of an anthropomorphic robot head during generic tracking tasks. This architecture consists in a basic controller capable of generating human-like eye-head coordinated motions either with optimal control theoretic methods or by learning the behavior of humans from biometric recordings. We show experiments where the the parameters of our controller are learned with data from children with 2, 4 and 7 weeks of age. The behavior of our controller reliably matches the behavior of 7 week old infants. For younger ages the approximation is less precise, mostly because infants tracking is more irregular. We note that 10 weeks is when children's tracking ability becomes very reliable.

We were able to track targets with generic trajectories in the $3D$ space. The control is performed in decoupled elevation and azimuth coordinates, that are then mapped to the joint level motions using the robot kinematics Jacobian. The basic controller performs eye-head coordination with pure proprioceptive feedback. To improve tracking accuracy, additional layers add vestibular and predictive capabilities to the system. This incremental and modular approach allows the implementation of limited forms of eye-head coordination even in robots lacking inertial sensors or predictive abilities.

In the future we plan to improve our system to take into account additional aspects of human oculomotor control that were not addressed here at a neurologically plausible level. For instance we use retinal error for controlling both saccadic and smooth-pursuit motions whereas it is know that smooth pursuit is controlled by retinal slip. Also we will research prediction methodologies that adapt to the input signal characteristics. We have shown methods based on motion models for constant velocity and constant acceleration targets, but humans are able to track predictively a much wider range of motions. In particular we are targeting methods for detecting and modeling general periodic motions.

ACKNOWLEDGMENTS

This work was partially supported by the European Commission, Project IST-004370 RobotCub, and by the Portuguese Government - Fundação para a Ciência e Tecnologia (ISR/IST plurianual funding) through the POS_Conhecimento Program that includes FEDER funds, and through project BIO-LOOK, PTDC / EEA-ACR / 71032 / 2006.

Authorized licensed use limited to: UNIVERSIDADE TECNICA DE LISBOA. Downloaded on December 29, 2009 at 13:00 from IEEE Xplore. Restrictions apply.

REFERENCES

[1] R. Lencer and P. Trillenberg, "Neurophysiology and neuroanatomy of smooth pursuit in humans." *Brain and Cognition*, vol. 68, pp. 219–228, 2008.

[2] W. H. Zangemeister and L. Stark, "Gaze latency: Variable interactions of head and eye latency," *Experimental Neurology*, vol. 75, pp. 389–406, 1982.

[3] C. von Hofsten and K. Rosander, "Development of smooth pursuit tracking in young infants," *Vision Research*, vol. 37, pp. 1799–180, 1997.

[4] K. Rosander and C. von Hofsten, "Visual-vestibular interaction in early infancy," *Experimental Brain Research*, vol. 133, pp. 321–333, 2000.

[5] H. Goosens and A. V. Opstal, "Human eye-head coordination in two dimension under different sensorimotor conditions," *Experimental Brain Research*, vol. 114, pp. 542–560, 1997.

[6] E. S. Maini, L. Manfredi, C. Laschi, and P. Dario, "Bioinspired velocity control of fast gaze shifts on a robotic anthropomorphic head," *Autonomous Robots*, vol. 25, no. 1-2, pp. 37–58, August 2008.

[7] T. Shibata, S. Vijayakumar, J. Conradt, and S. Schaal, "Biomimetic oculomotor control," *Adaptive Behavior*, vol. 9, no. 3-4, pp. 187–207, March 2001.

[8] R. Beira, M. Lopes, M. Praça, J. Santos-Victor, A. Bernardino, G. Metta, F. Becchi, and R. Saltarén, "Design of the robot-cub (icub) head," in *IEEE - International Conference on Robotics and Automation (ICRA'06)*, 2006.

[9] M. W. G. F. Franklin, J. D. Powell, *Digital Control of Dynamic Systems*. Addison Wesley Longman, 1998.

[10] A. Locatelli, *Optimal Control: An Introduction*. Birkhuser, 2001.

[11] V. J. Wilson, R. Boyle, K. Fukushima, P. K. Rose, Y. Shinoda, Y. Sugiuchi, and Y. Uchino, "The vestibulocollic reflex." *Journal of Vestibular Research*, vol. 5, no. 3, pp. 147–70, 1995.

[12] X. R. Li and V. P. Jilkov, "Survey on maneuvering target tracking - part i: dynamic models," *IEEE Transactions on Aerospace and Electronics Systems*, vol. 39, no. 4, pp. 1333–1364, 2003.

[13] ——, "Survey on maneuvering target tracking - part ii: Ballistic target models," *Proc. 2001 SPIE Conf. Signal and Data Processing of Small Targets*, vol. 4473, pp. 559–581, 2001.

[14] R. Cutler and L. S. Davies, "Robust real-time periodic motion detection, analysis and applications," *PAMI*, vol. 22, no. 8, pp. 781–796, 2000.

[15] P. S. Maybeck, *Stochastic Models, Estimation and Control, Vol 1*. Academic Press, 1979.

Authorized licensed use limited to: UNIVERSIDADE TECNICA DE LISBOA. Downloaded on December 29, 2009 at 13:00 from IEEE Xplore. Restrictions apply.