Biomimetic Gaze Stabilization based on Feedback-Error-Learning with Nonparametric Regression Networks

T. Shibata[1]

[1] Kawato Dynamic Brain Project
Japan Science and Tech. Corp.
tom@erato.atr.co.jp
http://www.erato.atr.co.jp/~ tom/

S. Schaal[1,2]

[2] Computer Science and Neuroscience
Univ. of Southern California
sschaal@usc.edu
http://www-slab.usc.edu/sschaal/

Abstract

Oculomotor control in a humanoid robot faces similar problems as biological oculomotor systems, i.e., the stabilization of gaze in face of unknown perturbations of the body, selective attention, stereo vision, and dealing with large information processing delays. Given the nonlinearities of the geometry of binocular vision as well as the possible nonlinearities of the oculomotor plant, it is desirable to accomplish accurate control of these behaviors through learning approaches. This paper develops a learning control system for the phylogenetically oldest behaviors of oculomotor control, the stabilization reflexes of gaze. In a step-wise procedure, we demonstrate how control theoretic reasonable choices of control components result in an oculomotor control system that resembles the known functional anatomy of the primate oculomotor system. The core of the learning system is derived from the biologically inspired principle of feedback-error learning combined with a state-of-the-art nonparametric statistical learning network. With this circuitry, we demonstrate that our humanoid robot is able to acquire high performance visual stabilization reflexes after about 40 seconds of learning despite significant nonlinearities and processing delays in the system.

1 Introduction

The goal of our research is to investigate the interplay between oculomotor control, visual processing, and limb control in humans and primates by exploring the computational issues of these processes with a humanoid robotic system and by comparing predictions of our theories with data from neurobiology. In this paper, we will present the first step towards these goals, focussing on oculomotor control. In many biological organisms, oculomotor is the interface to one of the most important sensory systems, the visual system, and through this system it affects biological information processing in various ways. Firstly, oculomotor control is needed to select which visual information should be processed by the limited sensory resources of the eye and the following brain circuits. This role is particularly important in primates that employ foveal vision. Furthermore, as eye movements must be executed in a sequential manner, it is crucial to focus visual attention at the right time on the right targets such that subsequent information processes, in particular motor planning and execution, receive relevant information sufficiently fast to update on-going computations. Secondly, in order to obtain full understanding of limb motor control or whole-body control, it may be necessary to study oculomotor control because oculomotor control may be a crucial constraint in how movement of other body parts are planned. Recent physiological research provides several insights that are consistent with the above statements. For instance, Gauthier et al. demonstrated a tight coupling of oculomotor and limb-motor systems through oculo-manual tracking experiments of visual targets[Gauthier *et al.*(1988),], and Miyashita et al. showed that anticipatory saccades in sequential procedural learning in monkeys are tightly coupled to the limb-motor system [Miyashita *et al.*(1996),].

Many artificial vision systems [Ballard & Brown(1993), Aloimonos *et al.*(1988),] have been developed to include oculomotor control techniques. To contrast these "moving" vision systems from research conducted using purely "static" vision systems, the term

"active" vision was introduced [Ballard & Brown(1993),]. Although most of these approaches are inspired by biology's active vision [Ferrell(1996), Takanishi *et al.*(1995), Capurro *et al.*(1996), Berthouze *et al.*(1996a), Murray *et al.*(1992),], only few implementations of oculomotor systems can be found that try to emphasize biological plausibility.

In this paper, we will focus on an active vision system that employs as much as possible computational mechanisms that have been discovered in neurophysiological research on primates. We will restrict our scope to the most basic oculomotor behaviors, the stabilization of visual information by means of oculomotor reflexes. Successful visual perception requires that retinal images remain constant, at least for a certain amount of time. Since an oculomotor system usually resides in a moving body, one of the most basic and phylogenetically oldest functions of oculomotor control is visual stabilization, an area that has been studied intensively in neurobiology and that is known to be implemented as a combination of reflexes.

Our work is particularly inspired by research in the vestibulocerebellum that suggested the idea of feedback-error-learning (FEL) as a biologically plausible and control theoretically sound adaptive control concept [Kawato(1990),]. While FEL has been pursued in several robotic studies before (e.g.,[Bruske *et al.*(1997), Berthouze *et al.*(1996b),]), there has never been any emphasis on developing synthetic FEL learning approaches that consider fast nonlinear learning and coping with unknown delays in the sensory feedback pathway.

In the context of learning oculomotor reflexes, we will demonstrate in this paper how nonparametric regression networks in conjunction with FEL can be employed to learn a biomimetic oculomotor controller that leads to accurate control performance and very fast learning convergence for a nonlinear oculomotor plant with temporal delays in the feedback loop. For this purpose, first, we will review some of the elementary principles of biological gaze stabilization reflexes. Second, we will focus on exploring how, from a computational point of view, efficient learning of these reflexes could be achieved based on parallels that have been developed in research on the vestibulocerebellum. At last, we describe the experimental setup we have developed to explore the feasibility of biomimetic oculomotor control and results obtained with this system.

2 Biological VOR-OKR Models

2.1 The VOR and the OKR

When we wave our hand in front of our eyes at a moderate frequency, it seems blurred, while there is no blurring when we look at the non-moving hand during movement of our head at a similar frequency. The former phenomenon involves the Opto-Kinetic-Response (OKR), one of the basic reflexes to stabilize images on the retina. As sensory input, the OKR receives retinal slip information. Retinal slip is usually defined as the overall velocity with which an image drifts on the retina. As the goal of the OKR is to keep the image still on the retina, retinal slip can be used as an error velocity signal for the OKR system. We will also use the positional error signal in our experiments below, and thus, we will explicitly distinguish between retinal slip error and retinal slip error velocity when necessary. Otherwise, retinal slip will denote both position and velocity errors.

The second phenomenon above which resulted in non-blurred visual perception is due to the Vestibulo-Ocular-Reflex (VOR). The VOR uses as sensory input the head velocity signal acquired by the vestibular organ in the semicircular canal. This signal is available with short latency (15-30ms) and can thus stabilize the image on the retina more efficiently than the OKR, whose retinal slip-based negative feedback system operates with about 80-100ms latency. The VOR inverts the sign of the measured head velocity and, with the help of a well-tuned feedforward controller, rapidly generates the appropriate motor commands for the eyes. This fast control loop allows for non-blurred perception at much higher frequencies than the OKR.

For biological systems it is known that VOR and OKR cooperate to achieve visual stabilization. The output of the vestibular system has been shown to be proportional to the head velocity in the frequency range between 0.02 and 2 Hz [Melvill Jones & Milsum(1971),]. However, its output is attenuated during rotational head motion with constant velocity, which leads to a vanishing of the VOR response. In this situation, the slip of the image on the retina will elicit a response from the OKR such that the OKR instead of the VOR will stabilize the retinal image. The OKR is well suited to fill in the missing VOR performance as its preferred frequency range is much lower than that of the VOR.

2.2 An overview of previous learning models of the VOR and OKR

VOR performance is not genetically preprogrammed but rather adaptive. During experimental manipulations of retinal slip using magnifying spectacles, inversion prisms, or rotating visual screens, the VOR gain (the ratio of eye to head velocity) changes after some adjustment time such that the previous performance of keeping images still on the retina is recovered. This learning is assumed to happen in the the vestibulocerebellum. Especially for the horizontal VOR, the neural control circuitry has been studied intensively since it involves only a small number of brain areas, i.e., a three neuron reflex arc and the H-zone [Ito(1984),] of the cerebellar flocculus, which is also known to participate in the OKR.

Gomi and Kawato proposed a model of VOR-OKR learning, simulated simultaneous adaptation of VOR and OKR [Gomi & Kawato(1992),], and compared it with the biological data. Based on Ito's research, learning in their model takes place in the cerebellar flocculus according to the feedback-error learning theory [Kawato(1990),]. Lisberger and his colleagues have proposed VOR learning models where learning occurs both in the cerebellar flocculus and the brainstem [Coenen & Sejnowski(1996), Raymond & Lisberger(1998),] using findings from a previous line of research [Stone & Lisberger(1990a), Stone & Lisberger(1990b),]. Their scheme is more complex than the Gomi and Kawato model as it differentiates frequency dependent sites of learning in the brain in order to explain a larger set of experimental VOR stimuli. Quinn et al. provided a simulation of adaptive mechanisms in the vestibulo-ocular reflex employing a learning algorithm similar to feedback-error learning, however, without investigating the cooperation and coordination between the OKR with the VOR [Quinn *et al.*(1992),]. Another application of FEL to learning ocolomotor control was reported by Berthouze et al. [Berthouze *et al.*(1996a),]. These authors trained a 3-layer perceptron in a smooth pursuit task to demonstrate principles of learning by demonstration.

3 A Computational Model of the VOR-OKR

3.1 Research objectives

From a control theoretic point of view, the VOR-OKR system corresponds to a negative feedback controller based on retinal slip information, augmented with a feedforward controller based on vestibular input. Control with such a system is straightforward if the dynamics of the eye system is known and the feedback pathways have no delays. However, the opposite is true for biological and even artificial oculomotor systems: retinal slip information is significantly delayed due to the overhead of complex computations in the visual pathway, and the oculomotor plant is nonlinear. The nonlinearities in artificial systems result from distortions of the lens of camera-eyes and nonlinear spring terms added by the relatively heavy cables attached to the cameras. In both biology and artificial oculomotor system, the offset between the rotational axes of the eye-balls and the head causes an additional nonlinearity [Coenen & Sejnowski(1996), Panerai & Sandini(1998),].

In order to cope with these problems, two routes can be taken. In classical control engineering, manual system identification would be employed to model the nonlinearities and delays in the control system as well as possible. Afterwards, a nonlinear controller could be designed to achieve appropriate VOR-OKR performance [Slotine & Li(1991),]. Despite this approach has been successful, it cannot cope with the inevitable changes of dynamical systems over time, and it requires a fair amount of manual work. Therefore, we will follow an adaptive control strategy in this paper, similar to biology organisms. Biology apparently equips motor systems with some crude initial performance, and fine-tunes the performance during development by means of adaptive neural networks. Feedback error learning mimics such a process.

In the following sections, we will develop an oculomotor control system that resembles the functionality of the brainstem and cerebellar pathways in biological oculomotor control. Our goal is to obtain a computationally efficient, control theoretically sound, and as much as possible biologically plausible VOR-OKR circuit. Equipped with a state of the art nonparametric neural network, our approach will be demonstrated to have fast learning convergence, excellent performance, and the ability to deal with unknown nonlinearities and time delays in the control system. During our development, for the sake of clarity it is initially assumed that the oculomotor system is linear. This assumption is also made in most biological research, since reflex loops and antagonistic activation can strongly contribute to a linearization of the nonlinear muscle properties, and the oculomotor system is unaffected by interaction forces from limb movement due to its low inertia. In later sections, we will relax the linearity

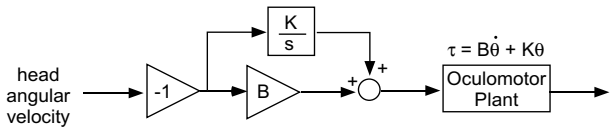

Figure 1: Inverse model control as a most basic VOR system. For the moment, the oculomotor plant is assumed to be a first order linear system in terms of its angular position θ, where τ is the torque input and K and B correspond to the mass and viscous components of the plant respectively. K/s is an integrator $(1/s)$ with the position gain K.

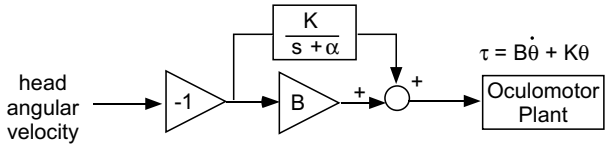

Figure 2: Inverse model control with a leaky integrator which implements a forgetting term with time constant α.

assumption.

3.2 A simple VOR system

A most basic model of the VOR system can be synthesized as a feedforward open-loop controller using an inverse model control (Fig. 1). Here, the oculomotor plant is a first order linear system as in most biological studies. The inverse dynamics model requires two gains, a position gain in the integrator pathway, and a velocity gain in the velocity pathway. If these gains closely approximate the stiffness and damping terms of the oculomotor dynamics, perfect feedforward control is achieved.

Obviously, the control system in Fig. 1 is only marginally stable due to the floating integrator. One possibility to add robustness is to employ a leaky integrator instead of a perfect integrator, a strategy that is also employed in biology (Fig. 2), largely because of the imperfect realization of integration in biological neurons. Typical value for α would be 0.001, i.e., a time constant of $1,000$ seconds. The appropriate value of the leakage term depends on the dynamics of the input signal and the plant dynamics.

3.3 Merging the VOR with an OKR-like pathway

Fig. 3 suggests an alternative solution to stabilize the control system of Fig. 1 by employing negative

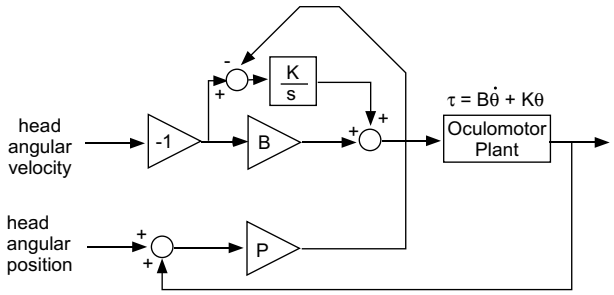

Figure 3: Using the retinal slip information to stabilize the integrator. The summation of the head angular position and the oculomotor angular position form the retinal slip. The retinal slip is multipled by a feedback gain P to stabilize the integrator.

feedback from a simple feedback controller, e.g., a proportional controller in our first order dynamics example in Fig. 3. The output of the feedback controller is interpreted as a velocity signal and added to the input of the integrator. Even if the feedback pathway is delayed and has low gain, excellent and stable VOR performance can be achieved with this new control system.

With one more modification, the final basic design of an VOR-OKR controller can be accomplished. The feedback error signal can be replaced by retinal slip error and retinal slip error velocity which are computed by the visual systems, as has been made explicit in Fig. 4. After multiplying these signals with appropriate gains and summing them up, the output of the feedback controller is added to the velocity signal from the vestibular system, as shown in Fig. 4. The resulting control circuit shows a surprising similarity with the Final Common Path (FCP) hypothesis as proposed by Robinson [Robinson(1975),]. Robinson suggested that the outputs of all oculomotor behaviors converge as velocity commands in a direct pathway and an integrated indirect pathway. The final motor command is generated in these two final pathways and subsequently relayed to the motoneurons. A large amount of evidence supports that oculomotor behaviors employ velocity control [Carpenter(1977),], thus indicating that velocity signals may be the major component of the final common path. Such velocity commands could also carry an inverse dynamics model of the eye-muscle system.

By inspecting Fig. 4 more carefully, and by conceiving of the inverse dynamics controller as a final common path, this control circuit can be interpreted as realizing an VOR-pathway and an OKR-like pathway. In the biological oculomotor system, the OKR

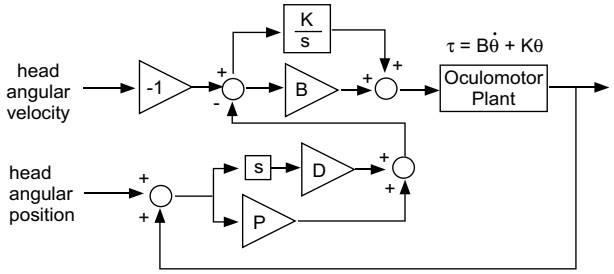

Figure 4: Adding the PD command to the vestibular signal. In addition to Figure 3, a velocity feedback pathway consisting of a differentiator s and a gain D is added.

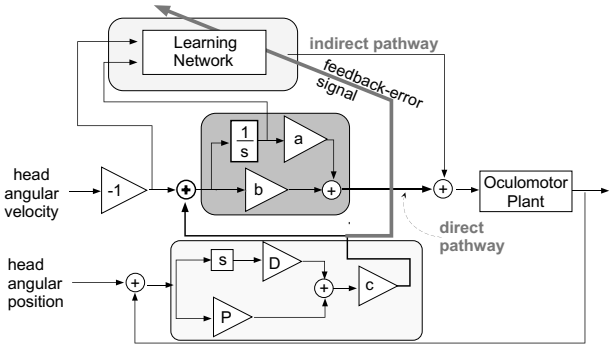

Figure 5: Adding a learning controller as an indirect pathway. This learning controller takes the desired velocity and the estimated desired position of the oculomotor plan, and outputs necessary feedforward torque. Since we use feedback-error-learning, the feedback signal is also fed to the learning controller as a training signal.

is conventionally defined as a compensatory negative feedback controller for the VOR. In this sense, the pathway involving the PD controller based on retinal slip can be called the OKR pathway in our final diagram.

3.4 Adding a learning network

In a last step, a learning component is added to the circuitry in Fig. 4, illustrated in Fig. 5.

The entire control systems is quite similar to what has been discovered in the primate oculomotor system. The a priori existing feedforward controller provides some crude functionality of the VOR. The a priori existing feedback controller provides acceptable OKR performance for slowly changing visual targets and acts as a compensatory negative feedback controller for the VOR. These systems form what is called the "direct pathway" of oculomotor control in biology. By adding a learning controller in the indirect pathway of Fig. 5, trained with the FEL strategy, excellent VOR performance can be accomplished, even if the feedback pathway has large delays, and also the OKR performance will be improved to some extent. This learning network, known to be located in the primate cerebellum, acquires during the course of learning an inverse dynamics model of the oculomotor plant that compensates for the missing performance of the crude feedforward controller in the shaded box of Fig. 5. The coordination of a direct and indirect pathway is analogous to how the cerebellar pathway acts in parallel to the brainstem pathways [Gomi & Kawato(1992),]. As developed in [Shibata & Schaal(2000),], this control system is a control theoretical reasonable solution for both biological and robotic oculomotor control.

From the viewpoint of adaptive control, FEL is a model-reference adaptive controller. The controller is assumed to be equipped a priori with a stabilizing linear feedback controller whose performance, however, is not satisfactory due to nonlinearities in the plant and delays in the feedback signals. Therefore, the feedback motor command of this controller is employed as an error signal to train a neural network controller. Given that the neural network receives the correct inputs, i.e., usually current and desired state of the plant, it can acquire a nonlinear control policy that includes both an inverse dynamics model of the plant and a nonlinear feedback controller. [Kawato(1990),] proved the convergence of this adaptive control scheme and advocated its architecture as an abstract model of learning in the cerebellum. Fig. 5 shows how FEL can be embedded in our oculomotor system.

3.5 Fast and stable learning

Feedback-error learning (FEL) does neither prescribe the type of neural network employed in the control system nor the exact layout of the control circuitry. Rather FEL is a principle of learning motor control. It employs an approximate way of mapping sensory errors into motor errors that, subsequently, can be used to train a neural network by supervised learning. As a computationally efficient learning mechanism, we suggest to use Recursive Least Squares (RLS) for FEL, a Newton-like method with very fast convergence, high robustness, and without the need for elaborate parameter adjustments [Ljung & Soderstrom(1986),]. To apply RLS for FEL, a small modification in the RLS algorithm is required. The normal RLS is formulated as in Equations 1 and 2, where $\boldsymbol{w}$ is the regression vector to be

estimated, $\mathbf{P}$ is the inverted covariance matrix of the input data, $\boldsymbol{x}$ is the input vector, y is the output, and $\hat{y}$ is the predicted output. λ is a forgetting factor, discussed below.

$$\mathbf{P}(t) = \frac{1}{\lambda}[\mathbf{P}(t-1) - \frac{\mathbf{P}(t-1)\boldsymbol{x}(t)\boldsymbol{x}(t)^T\mathbf{P}(t-1)}{\lambda + \boldsymbol{x}(t)^T\mathbf{P}(t-1)\boldsymbol{x}(t)}] \quad (1)$$

$$\boldsymbol{w}(t) = \boldsymbol{w}(t-1) + \mathbf{P}(t)\boldsymbol{x}(t)(y(t) - \hat{y(t)}) \quad (2)$$

$$\hat{y}(t) = \boldsymbol{w}(t)^T\boldsymbol{x}(t) \quad (3)$$

As can be seen in Equation 2, normal RLS requires the presence of a target output y in the update rules. In motor learning, target values for motor commands rarely exist since errors are usually generated in sensory space, not in motor command space. The strategy of FEL can be interpreted as generating a pseudo target for the motor command $y(t-1) = \hat{y}(t-1) + \tau_{fb}(t)$, where τ_{fb} denotes the feedback error signal. Thus, for FEL, Equation 2 needs to be modified to become:

$$\boldsymbol{w}(t) = \boldsymbol{w}(t-1) + \mathbf{P}(t)\boldsymbol{x}(t)\tau_{fb}(t+1) \quad (4)$$

It should be noted that FEL requires the appropriate time alignment of error and state, as shown by the time indices in Equation 4. Moreover, since the feedback error signal is only an approximate value, it is necessary to add a forgetting factor λ in RLS, as shown above. λ lies in the $[0, 1]$ interval. For $\lambda = 1$, no forgetting takes place, while for smaller values, older values in the matrix $\mathbf{P}$ will be exponentially forgotten. This forgetting strategy allows to neglect training data from the early stages of learning, where the feedback error was large and most likely the most inaccurate.

Fig. 6 shows a Matlab/Simulink model for the suggested VOR learning system. The oculomotor plant is described as $M\ddot{\theta} + B\dot{\theta} + K\theta = \tau$ where $(M, B, K) = (3.6889 \times 10^{-4}, 0.0341, 0.4875)$. These values were derived from our real experiments. Since the mass term M is very small, the system can be considered first order. 0.5 Hz sinusoidal head motion was generated, and 33 ms sensor delay was assumed.

Fig. 7 illustrates the time course of the retinal slip from the simulation employing several variations in the control circuitry. (a) shows the time course of the retinal slip acquired by the complete control system as in Fig. 6. In (b), the RLS learning pathway was eliminated. (c) is a case where additionally the vestibular signal was not passed to the integrator. And in (d),

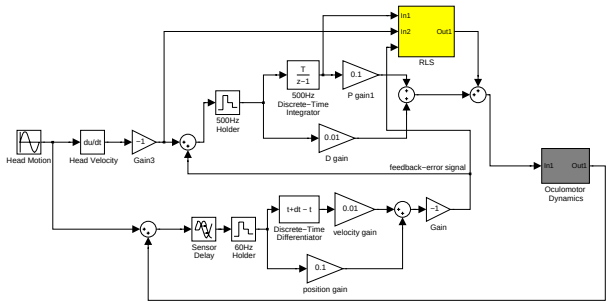

Figure 6: Our VOR-OKR diagram used for the simulation on SIMULINK. This diagram is a hybrid system of the discrete and the continuous oculomotor plant in order to simulate the real robot system as realistically as possible.

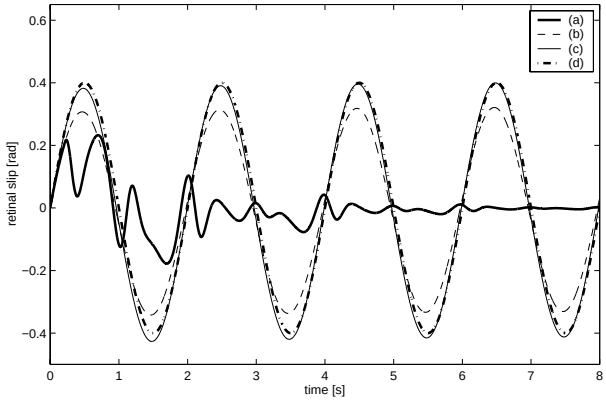

Figure 7: Time course of the retinal slip for several control variants cases described in Figure 6

even the feedback signal was not integrated anymore. Only the complete system can achieve perfect performance, i.e., zero retinal slip.

4 Learning in nonlinear oculomotor systems

4.1 Sources of nonlinearity in oculomotor control

There are three sources of nonlinearities both in biology and artificial oculomotor systems: i) muscle nonlinearities or nonlinearities added by the actuators and the usually heavy cable attached to the cameras, ii) perceptual distortion due to foveal vision, and iii) off-axis effects. Off-axis effects result from the non-coinciding axes of rotation of eye-balls and the head and require a nonlinear adjustment of the feedforward

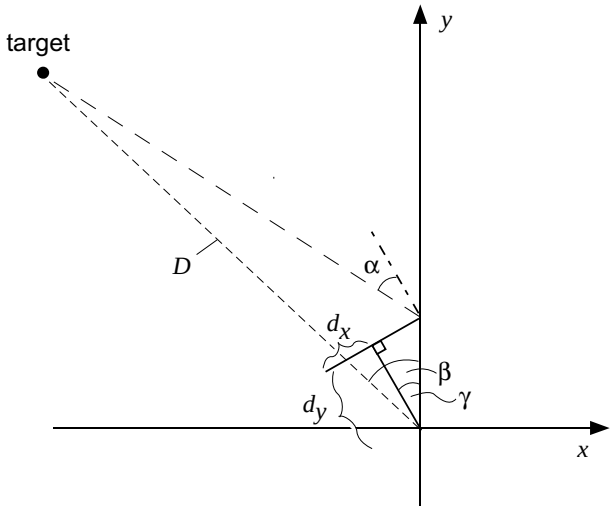

Figure 8: Off-axis configuration

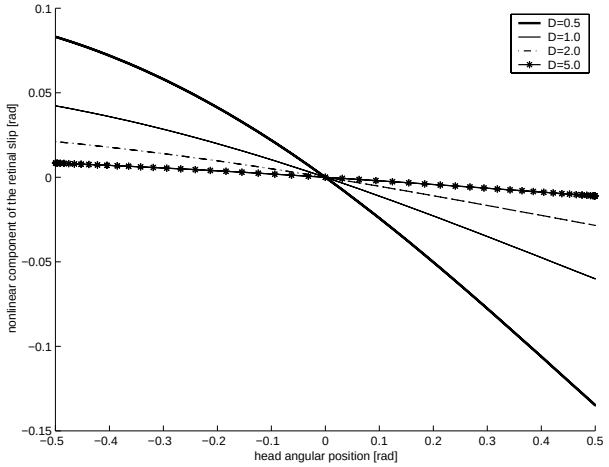

Figure 9: Nonlinear component of the retinal slip in off-axis cases for $d_x = 0.05\ m, d_y = 0.1\ m$

controller as a function of focal length, eye, and head position. Note that this off-axis effect is the most significant nonlinearity in our oculomotor system.

Equation 5 demonstrates the mathematical formula of the off-axis nonlinearity, derived from Fig. 8. $\alpha(t)$ is the appropriate eye angular position for a target at distance D and angular position $\beta(t)$, given an offset d_x, d_y of the eye-axis to the head-axis and the head angular position $\gamma(t)$.

$$\alpha(t) = \tan^{-1}\left(\frac{D\sin(\beta(t)) - d_y\sin(\gamma(t)) + d_x\cos(\gamma(t))}{D\cos(\beta(t)) - d_y\cos(\gamma(t)) - d_x\sin(\gamma(t))}\right) - \gamma(t) \tag{5}$$

Fig. 9 exemplifies the magnitude of the nonlinearity of the off-axis effect. Each curve shows the nonlinear component of the retinal slip, i.e., $\delta\alpha + \delta\gamma = (\alpha(t) - \alpha(0)) + (\gamma(t) - \gamma(0))$. For a linear system, this quantity would be zero, i.e., a change in *gamma* would require an equal change in *alpha*. The axis offset effects that this equality does not hold anymore. The curves in Fig. 9 were calculated for different values of D given that $\beta = 0$, $\gamma(0) = 0$, and $\alpha(0) = \sin^{-1}(d_x/(D - d_y))$. As can be seen, the nonlinearity of the retinal slip become quite significant for small D.

4.2 Learning in nonlinear systems

For nonlinear plants, a nonlinear learning system needs to replace the RLS solution from above. The appealing performance of recursive least squares can be carried over to the nonlinear domain by employing a spatially localized version of RLS, as suggested in the Receptive Field Weighted Regression (RFWR) algorithm [Schaal & Atkeson(1998),]. RFWR approximates nonlinear functions by piecewise linear functions localized in input space. The region of validity of each local model is determined by a receptive field which assigns a weight w_k, i.e., the activation strength of the receptive field, to an input data vector $\mathbf{x}$ according to multidimensional Gaussian kernel function:

$$w_k = exp(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{c}_k)^T\mathbf{D}_k(\boldsymbol{x} - \boldsymbol{c}_k)) \tag{6}$$

The receptive field is thus parameterized by its location in input space, $\boldsymbol{c}_k \in \mathcal{R}^n$, and a positive definite distance metric $\mathbf{D}_k$, deteremining the size and shape of the receptive field.

A prediction $\hat{y}$ for a query point $\mathbf{x}$ is calculated from the normalized weighted sum of the individual predictions $\hat{y_k}$ of all receptive fields:

$$\hat{y} = \frac{\sum_{k=1}^{K} w_k\hat{y}_k}{\sum_{k=1}^{K} w_k} \tag{7}$$

For the individual prediction within each receptive field, a linear function models the relationship between input and output data in analogy to RLS:

$$\hat{y}_k = (\boldsymbol{x} - \boldsymbol{c}_k)^T\boldsymbol{b}_k + b_{0,k} = \tilde{\boldsymbol{x}}^T\boldsymbol{\beta}_k, \quad \tilde{\boldsymbol{x}} = ((\boldsymbol{x} - \boldsymbol{c}_k)^T, 1)^T \tag{8}$$

where $\boldsymbol{\beta}_k$ denotes the parameters of the locally linear model and $\tilde{\boldsymbol{x}}$ a compact form of the center-substracted, augmented input vector to simplify the notation.

To clarify the elements and parameters of RFWR, Fig. 10 gives a network-like illustration for a single output system. The inputs are routed to all

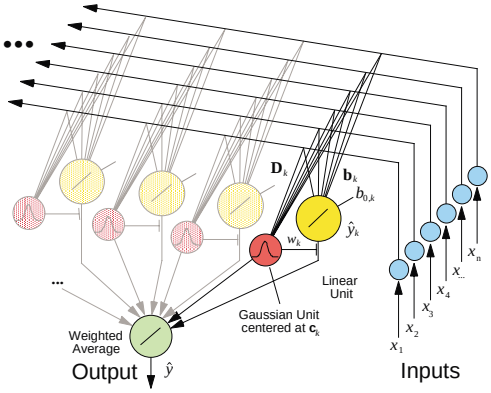

Figure 10: A network illustration of Receptive Field Weighted Regression

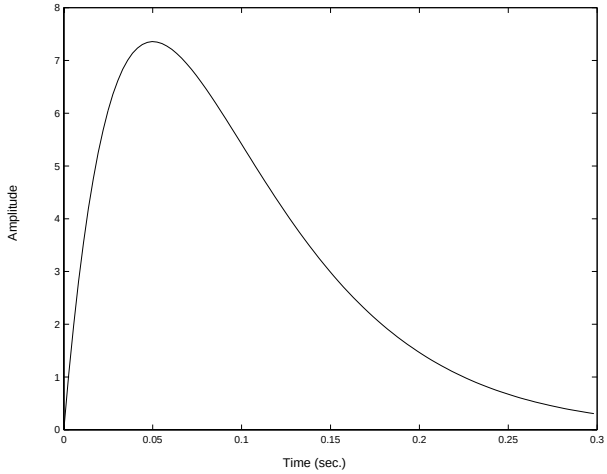

Figure 11: Impulse response of a second order filter

receptive fields, each of which consists of a linear and a Gaussian unit. The learning algorithm of RFWR determines the parameters $\boldsymbol{c}_k, \mathbf{D}_k$, and $\boldsymbol{\beta}_k$ for each receptive field by nonparametric regression techniques [Schaal & Atkeson(1998),].

For updating $\boldsymbol{\beta}_k$, RFWR adopts the local version of the RLS formulae:

$$\boldsymbol{\beta}_k(t) = \boldsymbol{\beta}_k(t-1) + w_k \mathbf{P}_k(t)\tilde{\boldsymbol{x}}(t)(y - \boldsymbol{\beta}_k(t-1)^T\tilde{\boldsymbol{x}}(t)) \tag{9}$$

$$\mathbf{P}_k(t) = \frac{1}{\lambda}[\mathbf{P}_k(t-1) - \frac{\mathbf{P}_k(t-1)\tilde{\boldsymbol{x}}(t)\tilde{\boldsymbol{x}}(t)^T\mathbf{P}_k(t-1)}{\frac{\lambda}{w_k} + \tilde{\boldsymbol{x}}(t)^T\mathbf{P}_k(t-1)\tilde{\boldsymbol{x}}(t)}] \tag{10}$$

In analogy with RLS, in order to use RFWR with FEL, the update equation for $\boldsymbol{\beta}_k$ needs to be adjusted to:

$$\boldsymbol{\beta}_k(t) = \boldsymbol{\beta}_k(t-1) + w_k \mathbf{P}_k(t)\tilde{\boldsymbol{x}}(t)\tau_{fb}(t+1) \tag{11}$$

4.3 Learning with the Delayed-Error Signal

For successful feedback-error-learning, the time-alignment between input signals and the feedback-error signal is theoretically crucial, and, thus, additional techniques are required in the case of delayed sensory feedback. For instance, if a perturbation of the head or body has frequency components that are much faster than the delay in the feedback pathway during VOR learning, the phase delay in the feedback pathway gets large such that learning speed decreases, or learning can even become unstable in the worst case.

To solve this "temporal credit assignment problem", the concept of eligibility traces has been suggested in both biological modeling and machine learning [Barto *et al.*(1983),]. For neurons in the brain, it is assumed that a second messenger would tag a synapse as eligible for modification. This "tag" would decay with an appropriate time constant, thus forming a temporal eligibility window. Schweighofer et al. proposed a biologically plausible learning model for saccadic eye movement, and modeled the second messenger as a second order linear filter (see Fig. 11) of the input signals to the learning system[Schweighofer *et al.*(1996),]. Using a second order filter is important since, in contrast to a first order filter, the impulse response has a unimodal peak at a delay time determined by the time constant of the filter. For successful learning, the delay time only has to roughly coincide with the actual delay of the sensory feedback. Note that first order filters are less appropriate as their impulse response peaks at the stimulus onset, i.e., without any delay. In related work, Fagg proposed a cerebellar learning model where eligibility traces modeled by such a second order filter are employed[Fagg *et al.*(1997),]. Applying this technique to FEL, we result in our final learning control system as illustrated in Fig. 12. This control system will serve for experimental evaluations on our anthropomorphic robot in the following section.

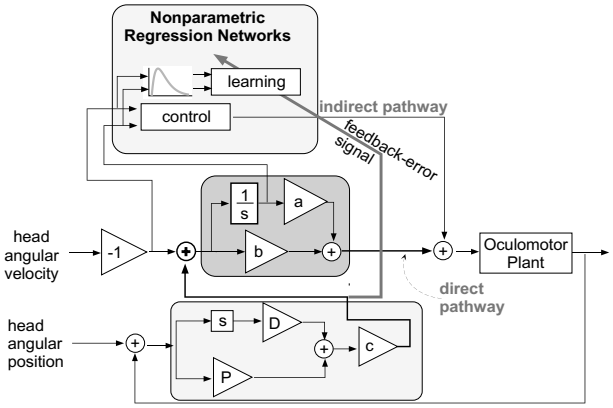

Figure 12: VOR-OKR model with eligibility trace. Compared to Figure 5, control and learning are separated in the learning compartment. The impuluse response of a second order filter is depicted before the learning box to indicate eligibility traces.

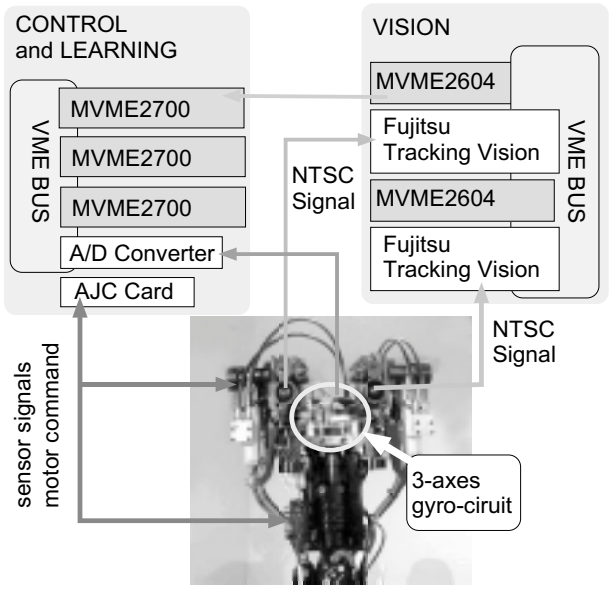

Figure 13: Our experimental setup

5 Experimental Results

5.1 Experimental setup

We implemented an on-line learning system of the VOR-OKR controller for our humanoid robot. Each DOF of the robot is actuated hydraulically out of a torque control loop. Each eye of the robot's oculomotor system consists of two cameras, a wide angle (100 degrees view-angle horizontally) color camera for peripheral vision, and second camera for foveal vision, providing a narrow-viewed (24 degrees view-angle horizontally) color image. This setup mimics the foveated retinal structure of primates, and it is also essential for an artificial vision system in order to obtain high resolution vision of objects of interest while still being able to perceive events in the peripheral environment. Each eye has two independent degrees of freedom, a pan and a tilt motion.

Fig. 13 depicts our experimental system. Two subsystems, a learning control subsystem and a vision subsystem, are setup in each VME rack and carry out all necessary computations out of the real-time operating system VxWorks.

Three CPU boards (Motorola MVME2700) are used for the learning control subsystem, and two CPU boards (Motorola MVME2604) are equipped for the vision subsystem. In the learning control subsystem, CPU boards are used, respectively, for: i) low level motor control of the eyes and other joints of our robot (compute torque mode), ii) visuomotor learning, and iii) data receiving from the vision subsystem. All communication between the CPU boards is carried out through the VME shared memory communication which, since it is implemented in hardware, is very fast. In the vision subsystem, each CPU board controls one Fujitsu tracking vision board in order to calculate retinal slip and retinal slip velocity information of each eye. NTSC video signals from the binocular cameras are synchronized to insure simultenous processing of both eyes' vision data. Vision data are sent via a serial port (115200 bps) to the learning control subsystem. For the experimental demonstrations of this paper, only one peripheral camera is used for VOR-OKR in its horizontal (pan) degree-of-freedom. Multiple degrees of freedom per camera, and multiple eyes just require a duplication of our control/learning circuits. If the image on a peripheral camera is stabilized, the image on the mechanically-coupled foveal vision is also stabilized. In order to mimic the semicircular canal of biological systems, we attached a three-axis gyro-sensor circuit to the head (Murata Manufacturing). From the sensors of this circuit, the head angular velocity signal is acquired through a 12 bit A/D board. The oculomotor and head control loop runs at 480 Hz, while the vision control loop runs at 30 Hz.

We use both visual-tracking and optical flow calculation in order to acquire the retinal slip and the retinal slip velocity, respectively. Both processes are based on the block-matching method[Inoue *et al.*(1993),] which is performed by the Fujitsu Tracking Vision board in real-

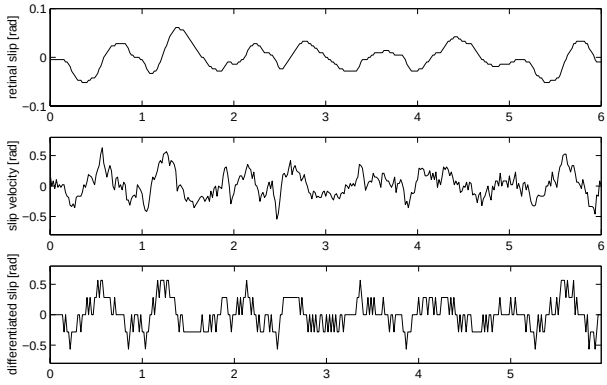

Figure 14: Retinal slip acquired by the real vision system

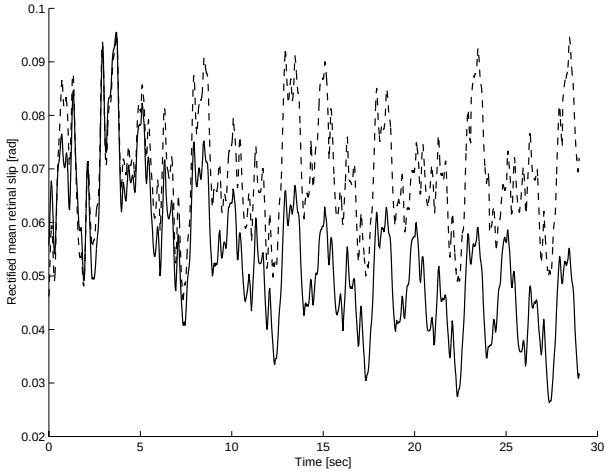

Figure 15: Time course of the rectified mean retinal slip: While the dashed line represents data obtained without eligibility trace, the solid line shows data acquired with eligibility trace.

time. In the beginning of each learning experiment, a template image is sampled from the the center of the image, and stored in memory. During the experiment, visual-tracking of the template image is performed in a pre-specified search area, and its resulting motion-vector is used as a retinal slip. The top row of Fig. 14 shows time course of the acquired retinal slip by our vision system. For this plot, the eye was fixed in the head, and the head was rotated sinusoidally. The bottom row of Fig. 14 shows the time course of the differentiated retinal slip. This retinal slip velocity is too noisy for learning. As on-line temporal filtering would produce too much time lag in the signal, we chose spatial averaging of multiple optical flow detectors to reduce the noise. The mid row of Fig. 14 illustrates samples of the retinal slip velocity acquired by this method. Comparing this data with the data in the bottom row demonstrates the improvements of spatial averaging of flow vectors in order to calculate velocity signal. For the following experiments, one template-tracker and 81 optic flow detectors were run for one peripheral camera. To maintain a 30 Hz vision processing loop rate, pixels were sampled only every three dots. Due to this sampling, the effective angular resolution around the center of the image was about 0.03 rad.

5.2 Experimental results(1): efficacy of eligibility trace

Since the retinal slip signals are generated from visual data, they are delayed by more than 30 ms. This delay affects FEL negatively. In this first set of experiments, we test how the eligibility traces can improve the efficacy of VOR learning.

The following experimental result was obtained from a head movement generated by three superimposed sinusoidal signals with frequencies of 0.6, 2.0, and 3.0 Hz and amplitude of 0.1 rad, respectively. Fig. 15 shows the time course of the rectified retinal slip obtained from a moving average using a one second time window. While the dashed line represents data obtained from learning without eligibility traces, the solid line shows data acquired with eligibility traces. This figure shows that eligibility traces are necessary for successful learning as the the retinal slip does not decrease without using the traces.

5.3 Experimental results(2): off-axis case

The next set of experiments investigated the improvements of using RFWR for learning in a nonlinear oculomotor plant. For this purpose, a large board with texture appropriate for vision processing was placed in front of the robot (see Fig. 16). The distance between a camera and the board was around 50 cm, i.e., a distance that emphasized the off-axis nonlinearities. In this experiment, the head was moved horizontally according to a sinusoidal signal with frequency 0.8 Hz and amplitude 0.25 rad. We compared linear learning based on RLS with nonlinear learning based on RFWR.

Fig. 17 and Fig. 18 present the results of this experiment. Fig. 17 shows the time course of the rectified retinal slip obtained from a moving average over a one second time window. The dashed line corresponds to RLS learning, while the solid line presents the learning

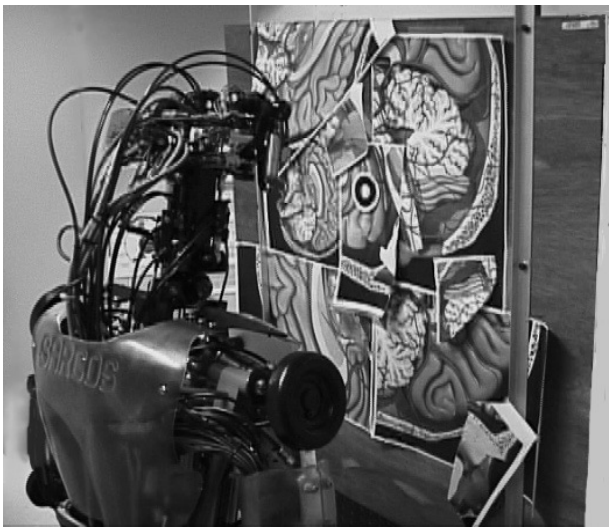

Figure 16: Experimental environment for the off-axis case

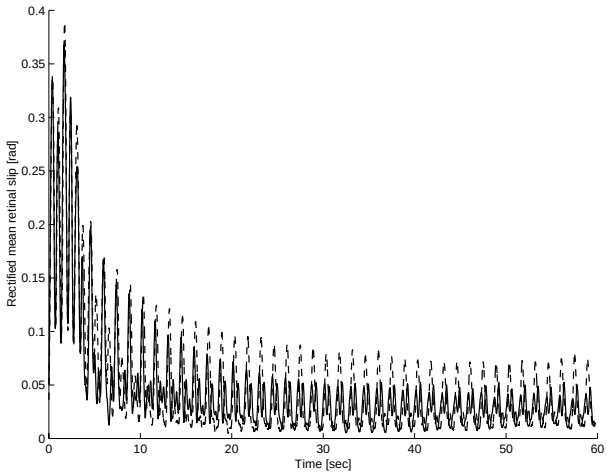

Figure 17: Time course of the rectified mean retinal slip: the dashed line corresponds to a linear learning result, and the solid line corresponds to a nonlinear learning result.

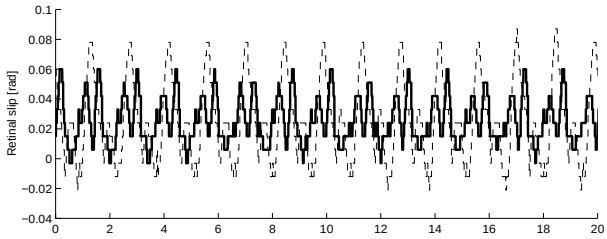

Figure 18: Retinal slip in the last part of learning: the dashed line corresponds to a linear learning result, and the solid line corresponds to a nonlinear learning result.

performance of RFWR, the nonlinear learning system. The need for a nonlinear learning is clearly demonstrated in this plot: each learning curve shows improvement over time, but the final retinal slip out of nonlinear learning is almost half of the remaining slip from linear learning. Fig. 18 shows the time course of the raw retinal slip signals in the end of learning. Since, as mentioned in Section 5.1, the effective angular resolution around the center of the image was 0.03 rad, the learning results shown in Fig. 17 and Fig. 18 are satisfactory as their amplitude is also about 0.03 rad, i.e., the best result achievable with this visual sensing resolution.

As shown in Fig. 9 , the nonlinear component generated by the off-axis effect is around 0.05 rad when the head is rotated 0.25 rad and the visual stimulus is at 0.5 m distance. This difference is consistent with the average difference between the results obtained RLS and RFWR, suggesting that RFWR was able to learn the nonlinear component generated by the off-axis effect.

6 Discussion

Our research objective is to study the computational processes of oculomotor control, visual processing, limb control, and the interdependencies of these three modalities by using a humanoid robot. This paper took a first step towards this goal by exploring adaptive gaze stabilization in a biomimetic artificial oculomotor system. We presented how the strategy of feedback-error-learning together with a state-of-the-art statistical learning network can be used to construct a control theoretically principled learning system for oculomotor control that is surprisingly similar to the adaptive control strategies employed in the primate cerebellum. We also showed how the idea of eligibility traces, a concept from biology and reinforcement learning, works nicely for overcoming unknown delays in the sensory feedback pathway. In experiments with our humanoid oculomotor system, it was shown that this system can acquire good VOR performance after about 10 seconds of learning and converges to excellent performance after about 30 to 40 seconds. This performance remained the same even in the case of nonlinearities of the oculomotor control system due to off-axis effects. Our control and learning strategies should be applicable without modifications

Figure 19: VOR-OKR learning with three template trackers: the monitor output of four cameras(two eyes) are shown. The bottom half images show the peripheral vision output where there are three large white rectangles showing search areas and three small rectangles showing the current targer positions.

to any other oculomotor systems.

In this paper, for the simplicity of presentation, we demonstrated experiments using only one axis camera movements. In our ongoing work, we also applied our techniques in experiments that required binocular, 4-axes simultaneous learning, and learning with multiple visual template trackers. Fig. 19 illustrated the camera images with superimposed tracking windows for this latter experiment. As can be seen in this figure, three visual trackers were located in the peripheral area of the image, which adds robustness in visual processing in so far that attending to a moving object in the center of the image would not interfere with the learning process that is guided by peripheral vision.

Despite the success of our oculomotor control system, several open issues need to be addressed in future work.

6.1 Velocity following VOR

In this paper, we have used a template tracker to acquire the retinal slip error. However, simple template-tracking is not very robust unless tuned with significant manual effort concerning luminance change and image distortion. Another problem is registering the matching template during robot motion, which is particularly problematic when using a normal NTSC camera running in interlace mode. Finding a tradeoff between camera shutter speed and image brightness is a possible but not very satisfactory way out. In biology, it is assumed that the brain only uses retinal slip error velocity for the VOR-OKR system, and that the saccadic system compensates for the drift in the integrator that occurs if a positional stabilization is missing. This strategy only requires optical flow calculation and no template matching, and it is thus very robust. Of course, another oculomotor behavior, i.e., saccades, need to be added to our control scheme, a topic that we will address in future research.

6.2 Improving OKR performance

This paper has described how to build a high-performance VOR system, but not a high performance OKR. The OKR is induced by large range visual stimuli on the retina, and, in biology, it is known that the OKR can be adaptive. The prime objective of the OKR in biology is conceived as compensating for missing VOR performance at low frequency movements, but not to achieve perfect visual tracking. For this purpose, primates possess a specialized visual feedback oculomotor behavior called 'smooth pursuit'. Smooth pursuit allows tracking a moving target with zero-error on the retina. It supresses the VOR such that the two motor behaviors do not interfer.

Kettner et al. proposed a cerebellar learning model for learning smooth pursuit [Kettner *et al.*(1997),]. From the view point of time series prediction, the scheme proposed in Kettner et al. corresponds to learning with tapped-delay-lines: inputs and efferent copies of the output signals are fed to the learning module with increasingly larger delay times. Applying this scheme to create a zero-error smooth pursuit circuit, and combining it with VOR, OKR, and saccades will be our future work.

References

[Aloimonos *et al.*(1988)] Aloimonos, J. Y., Weiss, I., & Bandyopadhyay, A. (1988). Active vision. *International Journal of Computer Vision*, **1**(4), 333–356.

[Ballard & Brown(1993)] Ballard, D. & Brown, C. (1993). Principles of animate vision. In Y. Aloimonos (Eds.), *Active Perception* (pp.245–282). Hillsdale, N.J.: Lawrence Erlbaum Associates.

[Barto *et al.*(1983)] Barto, A., Sutton, R., & Anderson, C. (1983). Neuronlike adaptive elements that can solve difficult learning control problems. *IEEE Transaction on Systems, Man, and Cybernetics*, **SMC-13**, 834–846.

[Berthouze *et al.*(1996a)] Berthouze, L., Bakker, P., & Kuniyoshi, Y. (1996a). Learning of oculo-motor control: a prelude to robotic imitation. In *International Conference on Intelligent Robots and Systems*.

[Berthouze *et al.*(1996b)] Berthouze, L., Rougeaux, S., Chavand, F., & Kuniyoshi, Y. (1996b). A learning stereo-head control system. In *World Automation Congress/International Symposium on Robotics and Manufacturing*.

[Bruske *et al.*(1997)] Bruske, J., Hanse, M., Riehn, L., & G, S. (1997). Biologically inspired calibration-free adaptive saccade control of a binocular camera-head. *Biological Cybernetics*

[Capurro *et al.*(1996)] Capurro, C., Panerai, F., & Sandini, G. (1996). Vergence and tracking fusing log-polar images. In *Proceedings of ICPR'96*, pages 740–744.

[Carpenter(1977)] Carpenter, R. (1977). *Movements of the Eyes*. Pion.

[Coenen & Sejnowski(1996)] Coenen, Olivier, J. & Sejnowski, T. (1996). A dynamical model of context dependencies for the vestibulo-ocular Reflex. In G. Tesauro, D.S. Touretzky, & T.K. Leen (Eds.), *Advances in Neural information Processing Systems 8*, Cambridge, MA: MIT Press.

[Fagg *et al.*(1997)] Fagg, A., Sitkoff, N., Barto, A., & Houk, J. (1997). Cerebellar learning for control of a two-link arm in muscle space. In *Proceedings of IEEE Conference on Robotics and Automation*, pages 2638–2644.

[Ferrell(1996)] Ferrell, C. (1996). Orientation behavior using registered topographic maps. In *Proceedings of the 4th International conference on Simulation of Adaptive Behavior*, pages 124–131.

[Gauthier *et al.*(1988)] Gauthier, G., Fercher, J.-L., Mussa Ivaldi, F., & Marchetti, E. (1988). Oculo-manual tracking of visual targets: control learning, coordination control and coordination model. *Experimental Brain Research*, **73**, 127–137.

[Gomi & Kawato(1992)] Gomi, H. & Kawato, M. (1992). Adaptive feedback control models of the vestibulo-cerebellum and spinocerebellum. *Biological Cybernetics*, **68**, 105–114.

[Inoue *et al.*(1993)] Inoue, H., Inaba, M., Mori, T., & Tachikawa, T. (1993). Real-Time Robot Vision System based on Correlation Technology. In *Proceedings of International Symposium on Industrial Robots (ISIR)*, pages 675–680.

[Ito(1984)] Ito, M. (1984). *The cerebellum and neural control*. New York: Raven Press.

[Kawato(1990)] Kawato, M. (1990). Feedback-error-learning neural network for supervised motor learning. In R. Eckmiller, editor, *Advanced Neural Computers*, pages 365–372. North-Holland: Elsevier.

[Kettner *et al.*(1997)] Kettner, R., Mahamud, S., Leung, H.-C., Sitkoff, N., Houk, J., Peterson, B., & Barto, A. (1997). Prediction of complex two-dimensional trajectories by a cerebellar model of smooth pursuit eye movement. *Journal of Neurophysiology*, **77**(4), 2115–2130.

[Ljung & Soderstrom(1986)] Ljung, L. & Soderstrom, T. (1986). *Theory and practice of recursive identification*. Cambridge, MA: MIT Press.

[Melvill Jones & Milsum(1971)] Melvill Jones, G. & Milsum, J. (1971). Frequency response analysis of central vestibular unit activity resulting from rotational stimulation of the semicircular canals. *Journal of Physiology*, **219**, 191–215.

[Miyashita *et al.*(1996)] Miyashita, K., Kato, R., Miyauchi, S., & Hikosaka, O. (1996). Anticipatory saccades in sequential procedural learning monkeys. *Journal of Neurophysiology*, **76**(2), 1361–1365.

[Murray *et al.*(1992)] Murray, D., Du, F., McLauchlan, P., Reid, I., Sharkey, P., & Brady, M. (1992). Design of stereo heads. In A. Blake & A. Yuille, editors, *Active Vision*. Cambridge, MA: MIT Press

[Panerai & Sandini(1998)] Panerai, F. & Sandini, G. (1998). Oculo-motor stabilization reflexes: integration of inertial & visual information. *Neural Networks*, **11**, 1191–1204.

[Quinn *et al.*(1992)] Quinn, K., Schmajuk, N., Baker, J., & Peterson, B. (1992). Simulation of adaptive mechanisms in the vestibulo-ocular reflex. *Biological Cybernetics*, **67**, 103–112.

[Raymond & Lisberger(1998)] Raymond, J. & Lisberger, S. (1998). Neural learning rules for the vestibulo-ocular reflex. *The Journal of Neuroscience*, **18**(21), 9112–9129.

[Robinson(1975)] Robinson, D. (1975). Oculomotor control signals. In *Basic mechanisms of ocular motility and their clinical implications*(pp. 337–374). Pergamon Press.

[Schaal & Atkeson(1998)] Schaal, S. & Atkeson, C. (1998). Constructive incremental learning from only local information. *Neural Computation*, **10**(8), 2047–2084.

[Schweighofer *et al.*(1996)] Schweighofer, N., Arbib, M., & Dominey, P. (1996). A model of the cerebellum in adaptive control of saccadic gain. *Biological Cybernetics*, **75**, 19–28.

[Shibata & Schaal(2000)] Shibata, T. & Schaal, S. (2000). Biomimetic Gaze Stabilization In J. Demiris & A. Birk (Eds.), *Robot Learning: An Interdisciplinary Approach* (pp. 31–52). World Scientific.

[Slotine & Li(1991)] Slotine, J. E. & Li, W. (1991). *Applied Nonlinear Control*. Prentice Hall.

[Stone & Lisberger(1990a)] Stone, L. S. & Lisberger, S. G. (1990a). Visual responses of purkinje cells in the cerebellar flocculus during smooth-pursuit eye movements in monkeys. i. simple spikes. *Journal of Neurophysiology*, **63**(5), 1241–1261.

[Stone & Lisberger(1990b)] Stone, L. S. & Lisberger, S. G. (1990b). Visual responses of purkinje cells in the cerebellar flocculus during smooth-pursuit eye movements in monkeys. ii. complex spikes. *Journal of Neurophysiology*, **63**(5), 1262–1275.

[Takanishi *et al.*(1995)] Takanishi, A., Ishimoto, S., & Matsuno, T. (1995). Development of an anthropomorphic head-eye system for robot and human communication. In *Proceedings of IEEE International Workshop on Robot and Human Communication*, pages 77–82.