Article:

Anderson, S.R., Dean, P., Kadirkamanathan, V. et al. (2 more authors) (2007) System identification from multiple short-time-duration signals. IEEE Transactions on Biomedical Engineering, 54 (12). pp. 2205-2213. ISSN 0018-9294

https://doi.org/10.1109/TBME.2007.896593

System Identification From Multiple Short-Time-Duration Signals

Sean R. Anderson*, Paul Dean, Visakan Kadirkamanathan, *Member, IEEE*, Chris R. S. Kaneko, and John Porrill

***Abstract*—System identification problems often arise where the only modeling records available consist of multiple short-time-duration signals. This motivates the development of a modeling approach that is tailored for this situation. An identification algorithm is presented here for parameter estimation based on minimizing the simulated prediction error, across multiple signals. The additional complexity of estimating the initial states corresponding to each signal is removed from the estimation algorithm. A numerical simulation demonstrates that the proposed algorithm performs well in comparison to the often-used least squares method (which leads to biased estimates when identifying systems from measurement noise corrupted signals). The approach is applied to the identification of the passive oculomotor plant; parameters are estimated that describe the dynamics of the plant, which represent the time constants of the visco-elastic elements that characterize the plant connective tissue.**

***Index Terms*—Initial conditions, oculomotor plant, output error, parameter estimation, state-space.**

I. Introduction

A COMMON approach to system identification is to assume that the model structure corresponds to an autoregressive with exogenous inputs (ARX) description, for which parameters are estimated using least squares (LS) [1]. However, the LS method can lead to significant bias in the parameter estimates if the signals are corrupted by measurement noise [2]. There are limited conceivable scenarios where measurement noise-free signals are recorded in a physiological context; hence, use of the ARX model is often inappropriate (except where the noise model has the same poles as the system). Therefore, an output error model is utilized in the modeling problem discussed here as an alternative to the ARX model, which results in correct treatment of the measurement noise entering the system description.

There is no closed-form solution for the output error model parameter estimation problem. Therefore, the parameters must be estimated via a nonlinear search. An estimate of the parameters can be obtained by minimizing the simulated prediction error (SPE) [1]. However, the SPE is significantly dependent on the initial system conditions for short time periods after excitation [3]. Therefore, if the initial conditions are unknown, which is generally true if they are unobserved, then these should be estimated along with the model parameters if the signal is of short time duration.

A further complication arises when modeling a system from multiple signals: the number of initial conditions to estimate increases in proportion to the number of experimental signals: an nth-order system will have n initial conditions corresponding to a single experimental signal; if there are M experimental signals available for modeling, then there will be Mn additional parameters to estimate. There are many cases, arising especially in the biological sciences, where multiple short-time-duration signals are collected for the purposes of system identification, for instance, when modeling muscle dynamics [4], [5] and the oculomotor plant [6], [7].

This paper takes the approach of solving the output error model identification problem via a separable least squares (SLS) estimation algorithm [8], [9]. The optimal estimate of the initial states is rewritten as a function of the model parameters. This removes the additional computational burden of estimating the initial state vector corresponding to each signal. Specifically, the state-space output error (SSOE) model representation is used, which leads to natural inclusion of the unknown states.

The proposed SLS method is applied to the practical problem of identifying the passive oculomotor plant using multiple signals. The modeling of the oculomotor plant dynamics is important for a number of reasons, including relating eye movement to oculomotor firing patterns [10], [11] and also for understanding the underlying algorithms governing eye-movement control [12], [13].

The identification of the oculomotor plant that is conducted here involves a reanalysis of signals that were previously modelled in [6], where a specialized estimation algorithm was developed. The aim of applying the proposed SLS algorithm to the problem of modeling the oculomotor plant is to demonstrate that this generic method works successfully on a real-world problem, which is validated by reference to the previous study described in [6].

There are alternative approaches to solving the joint state-parameter estimation problem, such as subspace methods [14], [15]. An advantage of the SLS approach proposed here is that it makes the parameterization of the model particularly straightforward so that states and model parameters can be easily related to physical quantities of interest. For instance in the application demonstrated here (modeling of eye movement dynamics), the model states are the extensions of each visco-elastic element representing the connective tissue and the model parameters are

Manuscript received September 1, 2006; revised March 3, 2007. This work was supported by EPSRC (UK) grant GR/T10602/01 under their Novel Computation Initiative. *Asterisk indicates corresponding author.*

*S. R. Anderson is with the Neural Algorithms Research Group, Department of Psychology, University of Sheffield, Western Bank, Sheffield S10 2TP, U.K. (e-mail: s.anderson@sheffield.ac.uk).

P. Dean and J. Porrill are with the Neural Algorithms Research Group, Department of Psychology, University of Sheffield, Sheffield S10 2TP, UK.

V. Kadirkamanathan is with the Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield S1 3JD, U.K.

C. R. S. Kaneko is with the Department of Physiology and Biophysics, School of Medicine, University of Washington, Seattle, WA 98195-7290 USA.

Digital Object Identifier 10.1109/TBME.2007.896593

the poles relating to the time-constants of the aforementioned visco-elastic elements.

The paper is structured as follows. Section II reviews the LS method in the context of identifying systems from signals corrupted by measurement noise; motivation is then demonstrated for the approach of minimizing the SPE as an alternative. Section III provides background on the use of SLS in system identification. The SLS approach to modeling multiple signals is derived in Section IV. The proposed method is analyzed and compared with LS in Section V. In Section VI, the method is applied to the estimation of the parameters that characterize the passive oculomotor plant. The paper is concluded in Section VII.

II. Problem Motivation

This section demonstrates the bias inherent in using LS to estimate the parameters of a model from an output signal corrupted by measurement noise. The minimization of SPE is then shown to be an appealing alternative, which provides motivation for the parameter estimation approach subsequently developed in the later sections.

A. Parameter Estimation via LS

It is often assumed that the structure that describes a linear time-invariant system (in discrete-time) is that of an ARX model. In fact, if no process noise is present and the observed system output is corrupted by measurement noise, then the correct system description is the output error (OE) model [1]. The purpose of this background section is to demonstrate how the LS parameter estimation of an OE model is biased by measurement noise. The single-input single-output OE model is described as

$$z_t + a_1 z_{t-1} + \cdots + a_n z_{t-n_a} = b_1 u_{t-1} + \cdots + b_m u_{t-n_b} \tag{1}$$

$$y_t = z_t + v_t \tag{2}$$

where z_t is the system output at time t, y_t is the observed system output corrupted by measurement noise, u_t is the system input, and $v_t \sim \mathcal{N}(0, \sigma^2)$ is a zero-mean normally distributed white process noise signal.

A one-step-ahead prediction model (based on the ARX structure) may be formed to estimate the model parameters from the observed signals, which is

$$y_t = \boldsymbol{\psi}_t \boldsymbol{\theta} + \varepsilon_t, \quad \text{for } t = 1, \ldots, N \tag{3}$$

where ε_t is the one-step-ahead residual modeling error, N is the number of data samples, and

$$\boldsymbol{\psi}_t = [-y_{t-1} \quad \ldots \quad -y_{t-n_a} \quad u_{t-1} \quad \ldots \quad u_{t-n_b}] \tag{4}$$

$$\boldsymbol{\theta} = [a_1 \quad \ldots \quad a_{n_a} \quad b_1 \quad \ldots \quad b_{n_b}]^T. \tag{5}$$

In order to track the influence of the measurement noise signal on the LS estimate, the prediction model (3) can be separated into

$$y_t = \boldsymbol{\phi}_t \boldsymbol{\theta} + \mathbf{v}_t \boldsymbol{\theta} + \varepsilon_t \tag{6}$$

where

$$\boldsymbol{\phi}_t = [-z_{t-1} \quad \ldots \quad -z_{t-n_a} \quad u_{t-1} \quad \ldots \quad u_{t-n_b}] \tag{7}$$

$$\mathbf{v}_t = [-v_{t-1} \quad \ldots \quad -v_{t-n_a} \quad 0 \quad \ldots \quad 0]. \tag{8}$$

This form of the prediction model immediately shows that the problem formulation is incorrect, for this system, because the prediction of the system output is partially dependent on measurement errors at previous sample times.

The parameter estimation cost function corresponding to the prediction model (6) is

$$V(\boldsymbol{\theta}) = \frac{1}{N}\sum_{t=1}^{N}(y_t - \boldsymbol{\phi}_t\boldsymbol{\theta} - \mathbf{v}_t\boldsymbol{\theta})^2. \tag{9}$$

Minimizing $V(\boldsymbol{\theta})$ leads to the LS estimate

$$\boldsymbol{\theta}_{\mathrm{LS}} = \left(\frac{1}{N}\sum_{t=1}^{N}\boldsymbol{\phi}_t^T\boldsymbol{\phi}_t + \frac{1}{N}\sum_{t=1}^{N}R_t\right)^{-1}\frac{1}{N}\sum_{t=1}^{N}\boldsymbol{\phi}_t^T y_t \tag{10}$$

where

$$\mathbb{E}[R_t] = \mathbb{E}\left[\mathbf{v}_t^T\mathbf{v}_t\right] = \begin{bmatrix} R_{1,1} & 0 \\ 0 & 0 \end{bmatrix} \tag{11}$$

where $\mathbb{E}[\cdot]$ denotes expectation and each diagonal element of $R_{1,1} \in \mathbb{R}^{n_a \times n_a}$ is the variance σ^2 of the measurement noise, that is

$$R_{1,1} = \begin{bmatrix} \sigma^2 & & 0 \\ & \ddots & \\ 0 & & \sigma^2 \end{bmatrix}. \tag{12}$$

This analysis demonstrates that the LS estimate is biased, when estimating parameters from measurement noise corrupted signals, because the solution includes the additional term R_t, which is related to the variance, or power, in the noise signal.

Intuitively, it may be supposed that for high SNR, the LS estimate will not incorporate significant bias. However, it is the case that if any of the singular values of R_t are of similar magnitude to any of the singular values of $\boldsymbol{\phi}_t^T\boldsymbol{\phi}_t$, the parameter estimates will be significantly biased. This point is demonstrated on an example problem in Section V-B.

B. Parameter Estimation via Minimization of SPE

The bias in the LS estimate of the model parameters originates from the specification of the prediction structure, as seen in (6). A way of overcoming the measurement noise problem is to generate the system model predictions by some alternative method; for instance via model simulation. This leads to an optimization problem where the SPE is minimized (by a nonlinear search). This is the method typically used to identify the output error model [1].

The minimization of SPE is accomplished by filtering the input sequence through a transfer function model of the system $H(\boldsymbol{\theta}, \mathbf{x}_1)$, which for short-duration signals is dependent on the parameter vector $\boldsymbol{\theta}$ and the initial state $\mathbf{x}_1$

$$\hat{y}_t = H(\boldsymbol{\theta}, \mathbf{x}_1)u_t. \tag{13}$$

The parameter vector $\boldsymbol{\theta}$ and initial state $\mathbf{x}_1$ can be updated by a nonlinear least squares routine, minimizing the simulated prediction error η_t, which is defined as

$$\eta_t = y_t - \hat{y}_t \quad \forall t. \tag{14}$$

It is apparent that this approach does not suffer from the bias of the LS method because the recorded output variables are not used in *predicting* the model output, in contrast to (6). Instead, the recorded (noisy) output signal is only used to obtain the residual error η_t; the model predictions are obtained independently of the recorded output signal using the input sequence and the system model $H(\boldsymbol{\theta}, \mathbf{x}_1)$.

III. Separating the Optimization Problem: Model Parameters and Initial States

The problem of estimating the model parameters and initial states is simplified by recognizing that for any given model parameters the corresponding initial states can be obtained from a closed form solution. Therefore, an SLS problem can be constructed, where the model parameters are estimated via nonlinear least squares and the initial states are obtained via the use of a state-space prediction model [9].

A linear discrete-time invariant system can be represented by the SSOE model

$$\mathbf{x}_t = A(\boldsymbol{\theta})\mathbf{x}_{t-1} + B(\boldsymbol{\theta})\mathbf{u}_{t-1} \tag{15}$$
$$\mathbf{y}_t = C(\boldsymbol{\theta})\mathbf{x}_t + \mathbf{v}_t \tag{16}$$

where $A(\boldsymbol{\theta}) \in \mathbb{R}^{n\times n}$ is the state transition matrix, $B(\boldsymbol{\theta}) \in \mathbb{R}^{n\times n_u}$ is the input matrix, $C(\boldsymbol{\theta}) \in \mathbb{R}^{n_y\times n}$ is the measurement matrix, $\mathbf{x}_t \in \mathbb{R}^n$ is the system state, $\mathbf{y}_t \in \mathbb{R}^{n_y}$ is the system output, $\mathbf{v}_t \in \mathbb{R}^{n_y}$ is the measurement error, $\mathbf{u}_t \in \mathbb{R}^{n_u}$ is the system input, and $\boldsymbol{\theta}$ is a vector of unknown model parameters.

The sequence of system outputs can be written as a function of the unknown parameter vector $\boldsymbol{\theta}$ and the initial state vector $\mathbf{x}_1$

$$Y = F(\boldsymbol{\theta})\mathbf{x}_1 + G(\boldsymbol{\theta})U + V \tag{17}$$

where

$$Y = \begin{bmatrix}\mathbf{y}_1^T & \dots & \mathbf{y}_N^T\end{bmatrix}^T \tag{18}$$
$$U = \begin{bmatrix}\mathbf{u}_1^T & \dots & \mathbf{u}_{N-1}^T\end{bmatrix}^T \tag{19}$$
$$V = \begin{bmatrix}\mathbf{v}_1^T & \dots & \mathbf{v}_N^T\end{bmatrix}^T, \tag{20}$$

$$F(\boldsymbol{\theta}) = \begin{bmatrix}C(\boldsymbol{\theta})A^0(\boldsymbol{\theta}) & \dots & C(\boldsymbol{\theta})A^{N-1}(\boldsymbol{\theta})\end{bmatrix}^T \tag{21}$$

$$G(\boldsymbol{\theta}) = \begin{bmatrix} 0 & \dots & 0 \\ g_{1,1} & \dots & g_{1,N-1} \\ \vdots & & \vdots \\ g_{N-1,1} & \dots & g_{N-1,N-1} \end{bmatrix} \tag{22}$$

where $F(\boldsymbol{\theta}) \in \mathbb{R}^{n_y N\times n}, G \in \mathbb{R}^{n_y N\times n_u(N-1)}$ and

$$g_{i,j} = \begin{cases} C(\boldsymbol{\theta})A^{i-j}(\boldsymbol{\theta})B(\boldsymbol{\theta}), & 1 \le j \le i \\ 0, & \text{otherwise.} \end{cases} \tag{23}$$

The sequence of measurement errors V is assumed to be zero mean and normally distributed. Therefore, it is apparent from (17) that for any given parameter vector $\boldsymbol{\theta}$ a closed-form solution can be obtained for the initial state vector $\mathbf{x}_1$.

When the B matrix is unknown, the state-space prediction model defined in (17) may be separated further so that the B matrix is linearly related to the output along with the initial state vector [9]. In practice, a control canonical form may always be used to represent an input-output system where the B matrix is known. Furthermore, for multiple signals, the number of initial conditions will usually be much greater than the number of parameters in the B matrix. Thus, the main computational benefits (from a state-parameter estimation perspective) are derived from separating out the initial states, which is discussed further in the next section.

The parameter vector can be estimated using an iterative nonlinear search routine where the direction of update is dependent at each iteration on the estimate of $\mathbf{x}_1$. The particular nonlinear search used in this investigation was a quasi-Newton method, where the gradient is estimated using a numerical update; see, for instance, [16] and [17].

IV. Parameter and State Estimation From Multiple Signals

A. Problem Definition

The identification task is to estimate the single set of model parameters that describes the dynamic behavior of the system that generated all the available signals and the initial states corresponding to each signal; that is, to minimize the cost function

$$V(\boldsymbol{\theta}, X_1) = \frac{1}{2}||\boldsymbol{\eta}(\boldsymbol{\theta}, X_1)||_2^2 \tag{24}$$

where $\boldsymbol{\eta}(\boldsymbol{\theta}, X_1)$ is the SPE corresponding to all signals and X_1 is the stacked initial state vectors corresponding to each signal

$$X_1 = \begin{bmatrix}\mathbf{x}_1^{(1)T} & \dots & \mathbf{x}_1^{(M)T}\end{bmatrix}^T \tag{25}$$

where M is the number of signals.

As mentioned above, the task of solving the optimization problem is separated into two parts: a closed-form solution of the initial state vector as a function of the model parameters and ii) parameter estimation via non-linear search. These are dealt with in the subsequent section.

B. Parameter and State Estimation

The model prediction error $\boldsymbol{\eta}(\boldsymbol{\theta}, X_1)$ can be expressed using a prediction model that is similar to (17), which is augmented to include all signals,

$$\boldsymbol{\eta}(\boldsymbol{\theta}, X_1) = \Upsilon(\boldsymbol{\theta}) - \Lambda(\boldsymbol{\theta})X_1 - \Gamma(\boldsymbol{\theta}), \tag{26}$$

where the model that predicts the output corresponding to all signals is $\hat{\Upsilon}(\boldsymbol{\theta}) = \Lambda(\boldsymbol{\theta})X_1 + \Gamma(\boldsymbol{\theta})$ and

$$\Upsilon(\boldsymbol{\theta}) = \begin{bmatrix} Y_1^T & \dots & Y_M^T \end{bmatrix}^T \tag{27}$$

$$\Lambda(\boldsymbol{\theta}) = \begin{bmatrix} F_1(\boldsymbol{\theta}) & & 0 \\ & \ddots & \\ 0 & & F_M(\boldsymbol{\theta}) \end{bmatrix} \tag{28}$$

$$\Gamma(\boldsymbol{\theta}) = \begin{bmatrix} U_1^T G_1^T(\boldsymbol{\theta}) & \dots & U_M^T G_M^T(\boldsymbol{\theta}) \end{bmatrix}^T. \tag{29}$$

Substituting (26) in (24) leads directly to the definition of the cost function

$$V(\boldsymbol{\theta}, X_1) = \frac{1}{2}\|\Upsilon(\boldsymbol{\theta}) - \Lambda(\boldsymbol{\theta})X_1 - \Gamma(\boldsymbol{\theta})\|_2^2. \tag{30}$$

Formulating the above optimization task as a separable least squares problem requires finding a closed-form solution for the optimal initial state vector X_1^* as a function of $\boldsymbol{\theta}$. Taking the partial derivative of $V(\boldsymbol{\theta}, X_1)$ with respect to X_1 leads to

$$\frac{\partial}{\partial X_1}V(\boldsymbol{\theta}, X_1) = X_1^T\Lambda^T(\boldsymbol{\theta})\Lambda(\boldsymbol{\theta}) - \Upsilon(\boldsymbol{\theta})^T\Lambda(\boldsymbol{\theta}) + \Gamma^T(\boldsymbol{\theta})\Lambda(\boldsymbol{\theta}). \tag{31}$$

Setting $(\partial/\partial X_1)V(\boldsymbol{\theta}, X_1) = 0$ and solving for the optimal estimate X_1^* leads to

$$X_1^* = \Lambda^\dagger(\boldsymbol{\theta})(\Upsilon(\boldsymbol{\theta}) - \Gamma(\boldsymbol{\theta})) \tag{32}$$

where $\Lambda^\dagger(\boldsymbol{\theta})$ denotes the pseudoinverse of $\Lambda(\boldsymbol{\theta})$.

Substituting (32) in (30) has the desired effect of removing the unknown initial state vector from the cost function $V(\boldsymbol{\theta}, X_1)$, leading to a new cost function that is only a function of the parameter vector, that is

$$V(\boldsymbol{\theta}) = \frac{1}{2}\|[I - \Psi(\boldsymbol{\theta})]\Upsilon(\boldsymbol{\theta}) - [\Psi(\boldsymbol{\theta}) - I]\Gamma(\boldsymbol{\theta})\|_2^2 \tag{33}$$

where

$$\Psi(\boldsymbol{\theta}) = \Lambda(\boldsymbol{\theta})\Lambda^\dagger(\boldsymbol{\theta}). \tag{34}$$

The optimal model parameters $\hat{\boldsymbol{\theta}}$ are obtained by minimization of (33)

$$\hat{\boldsymbol{\theta}} = \mathrm{argmin}_{\boldsymbol{\theta}} V(\boldsymbol{\theta}). \tag{35}$$

Note that this minimization problem is not a function of the initial state vector. Hence, the number of parameters to estimate via a non-linear search is reduced from $Mn + n_\theta$ to just n_θ, where M is the number of signals, n is the system order and n_θ is the number of model parameters.

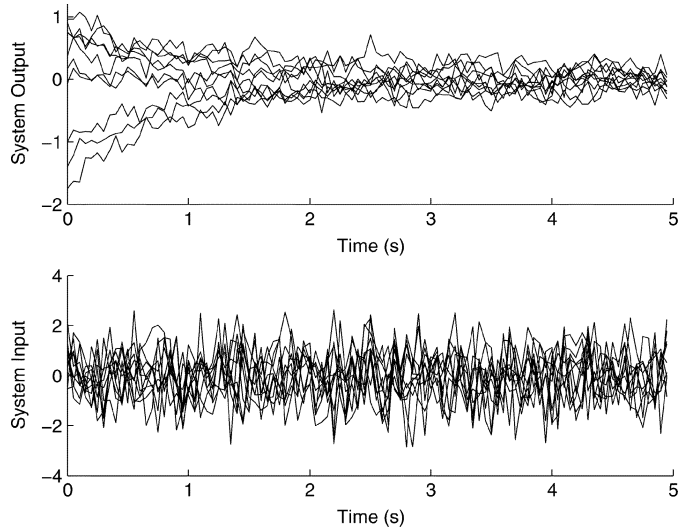

Fig. 1. One set of input/output training signals used in the test identification problem.

V. Analysis of SLS Parameter Estimation Algorithm Performance

This section describes the application of the SLS identification algorithm developed above to a test problem, focusing on the parameter estimation problem; a comparison is made with LS.

A. Problem Definition

The discrete-time test system was described in input–output form as

$$G(q) = \frac{0.04818q - 0.04583}{q^2 - 1.88q + 0.8825}. \tag{36}$$

A total of $M = 10$ signals were generated using a set of input records (normally distributed zero-mean white noise) and varying initial conditions in each case. The transfer function $G(q)$ was mapped into state-space control canonical form (of system order $n = 2$) for the straightforward inclusion of initial conditions; each initial state was defined as a random number, drawn from a uniform distribution in the range (0, 1).

To simulate measurement noise each signal was corrupted by normally distributed zero-mean white noise (adjusted to an SNR of 20 dB for each separate signal). The duration of the simulation was 5 s, and the sample rate was defined to be 20 Hz. The training signals are shown in Fig. 1.

To demonstrate the consistency of each estimation algorithm, the initial ten signals were duplicated 200 times with different measurement noise corruption. Each estimation algorithm (LS and SLS) was then applied to these different data sets. The LS estimate was obtained by concatenation of the regression matrix pertaining to each signal as described in [1].

The SLS estimation procedure was initialized by first low-pass-filtering the signals (cut-off at 5 Hz) and then obtaining the LS estimate from this filtered set of signals. This method has been demonstrated to lead to improved results in the initialisation of parameter estimates for output-error-type modeling approaches, over applying LS to the raw data [2].

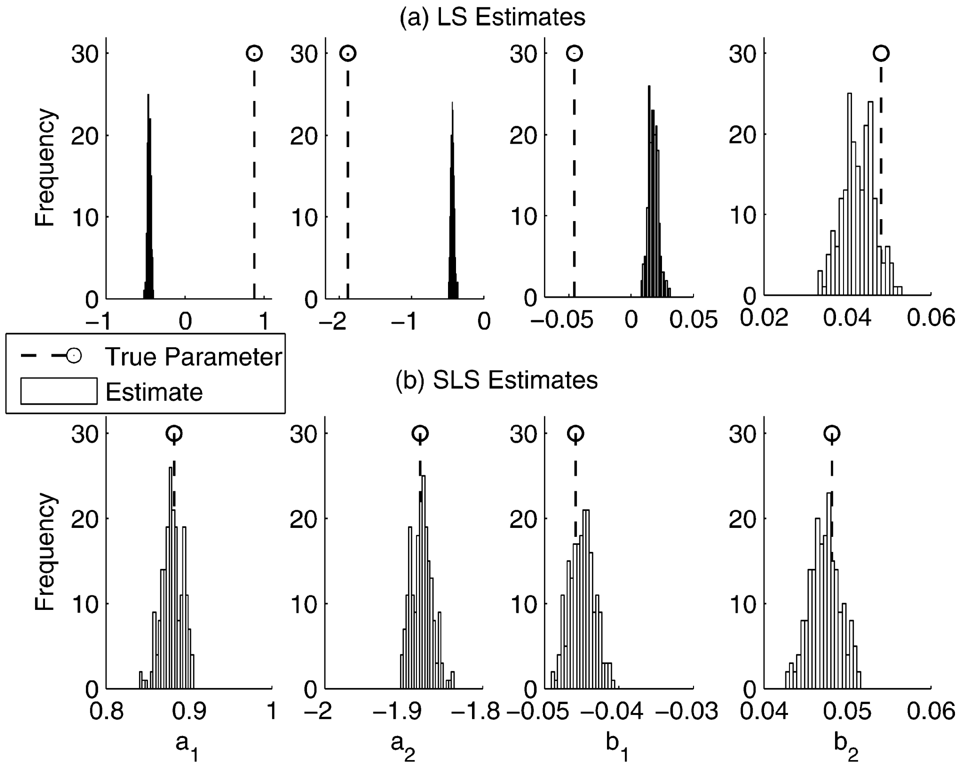

Fig. 2. Histograms that compare the true and estimated parameter values using (a) LS and (b) the SLS algorithm.

TABLE I
COMPARISON OF TRUE AND MEAN ESTIMATED PARAMETER VALUES

Parameter	True	LS	SLS
a_1	0.8825	-0.4500	0.8800
a_2	-1.88	-0.4316	-1.8775
b_1	-0.04583	0.0179	-0.0449
b_2	0.04818	0.0429	0.0473

B. Parameter Estimation Results

The application of each estimation algorithm led to parameter estimates that were consistent. However, the LS method lead to parameter estimates that were consistently and significantly biased (excepting the parameter b_1, which may have been due to the fact that it was a gain term). In contrast, the application of the proposed SLS algorithm reduced the bias in the estimate considerably. The results of parameter estimation for each method are shown in Fig. 2 in the form of histograms. The mean values of the parameter estimates are given in Table I.

To emphasize the poor performance of LS at even very high SNR levels (as hypothesized in Section II), each estimation algorithm (LS and SLS) was applied to a similar problem as defined above, but varying the SNR (note that only one set of signals was used at each level of SNR). The results confirmed that the LS estimate can be significantly biased at high SNR, whereas the SLS algorithm can reduce the bias to very small magnitudes at both low and high SNR. The results are shown in Fig. 3, in terms of RMSE in the parameter estimates.

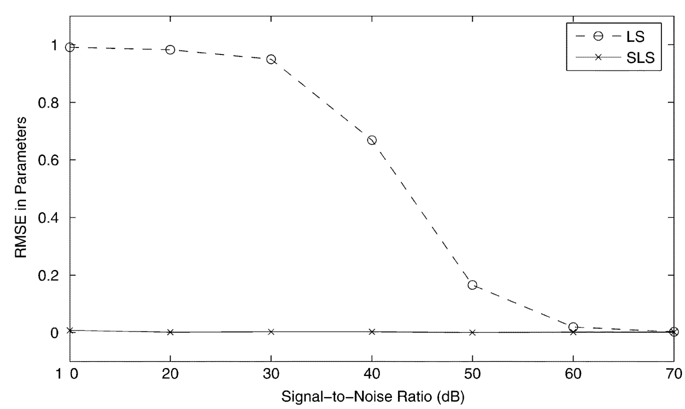

Fig. 3. Accuracy of parameter estimates when varying SNR.

VI. MODELING OF THE OCULOMOTOR PLANT

The modeling results presented here are compared to those previously presented in [6]. That method utilized the continuous-time system relationship $\mathbf{x}_t = e^{At}\mathbf{x}_1$, and hence was only useful in the case of modeling systems with zero input. The validation results presented in [6] show that the method was effective in describing the system dynamics; therefore, the comparison of new results presented here provides a cross-validation of the generic SLS systems modeling algorithm developed above.

A. Data Collection

Measurements were obtained from a juvenile rhesus monkey (Macaca mulatta), which was referred to as animal M in [6]. The animal had been implanted with a scleral search coil, a recording chamber and stabilizing lugs. The location of the nucleus prepositus hypoglossi had been mapped with standard

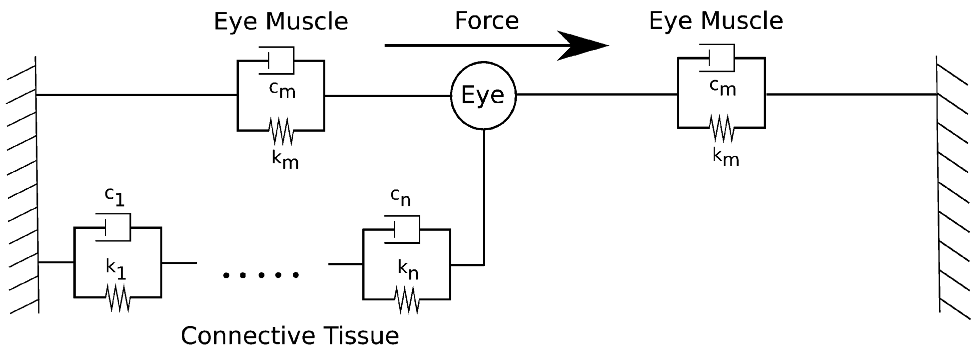

Fig. 4. Mechanical connectivity of the oculomotor plant.

extracellular electrophysiological recording techniques, but the nucleus itself had not been injected with ibotenic acid when these measurements were taken.

Following calibration of eye movements by requiring the animal to fixate targets at known eccentricities, the animal was lightly anaesthetized with ketamine ($\approx$10 mg/kg). Ketamine was chosen so that the animal would tolerate mechanical manipulation of the globe and because it is a dissociative anaesthetic and should thus minimize the affects on normal activity level in brainstem structures. Topical anaesthetic could not be used because completely alert animals do not tolerate manipulation of the globe even when it is anaesthetised. The exact dosage was titrated to the minimum level necessary to allow the animal to tolerate manipulation of the globe. It was low enough to avoid precipitating the vertical nystagmus that often accompanies higher doses (e.g., 25 mg/kg). If the threshold for vertical nystagmus was exceeded measurements were postponed until a future session. Horizontal nystagmus was never observed.

After a sufficient anaesthesia level was attained, the coiled eye was deviated manually with small forceps to between 15° and 45° either medial or lateral in the horizontal plane and abruptly released. Care was taken to avoid vertical deviation by monitoring eye position via the coil output. Trials were excluded if the return to resting position could be seen not to follow a smooth velocity trajectory due to the occurrence of a blink, saccade or slow eye movement. Eye position was sampled at 1 kHz.

B. Modeling

1) Model Structure: Knowledge of the physical system can lead to a useful representation of the system model in an identification context. This section describes the formation of an appropriate model structure based on physical insight.

The connective tissue of the oculomotor plant is a viscoelastic structure [18], which can be represented by a small number of Voigt elements in series [6], [7], as described in Fig. 4. Fig. 4 shows the mechanical structure of the eye connectivity, where c_j is a damping constant and k_j is an elasticity constant (the time constant associated with the jth Voigt element is $\tau_j = (c_j/k_j)$).

The model representation of the system with zero input is defined in the state-space form as

$$\mathbf{x}_t = A\mathbf{x}_{t-1} \tag{37}$$
$$\mathbf{y}_t = C\mathbf{x}_t + \mathbf{v}_t. \tag{38}$$

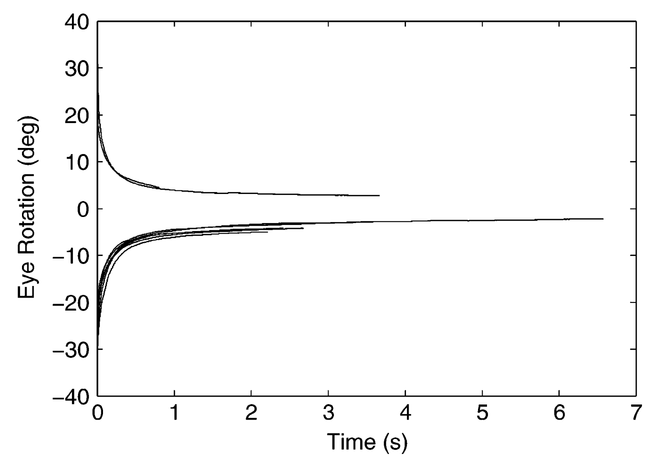

Fig. 5. Eye rotation data used for training and validation.

The states can be defined with physical significance; in this case the extension of a Voigt element. This requires the measured output (the extension of the muscle) to be the sum of the extension of each individual Voigt element (that is, the states). This leads to a natural definition of the state-space matrices as

$$A = \begin{bmatrix} p_1 & & 0 \\ & \ddots & \\ 0 & & p_n \end{bmatrix} \tag{39}$$
$$C = [1 \quad \dots \quad 1] \tag{40}$$

where the poles of the system, p_i, for $i = 1, \dots, n$, in the discrete-time model are related to the time constants of the corresponding Voigt elements (in the continuous-time system model) by the relationship

$$\tau_i = -\frac{T}{\log(p_i)}, \quad \text{for } i = 1, \dots, n \tag{41}$$

where T is the sample time.

To ensure that the model remained stable and nonoscillatory (a known property of the system), the parameter estimates were transformed within the search routine using the expression

$$\tilde{\theta} = \frac{1}{e^{-\theta} + 1} \tag{42}$$

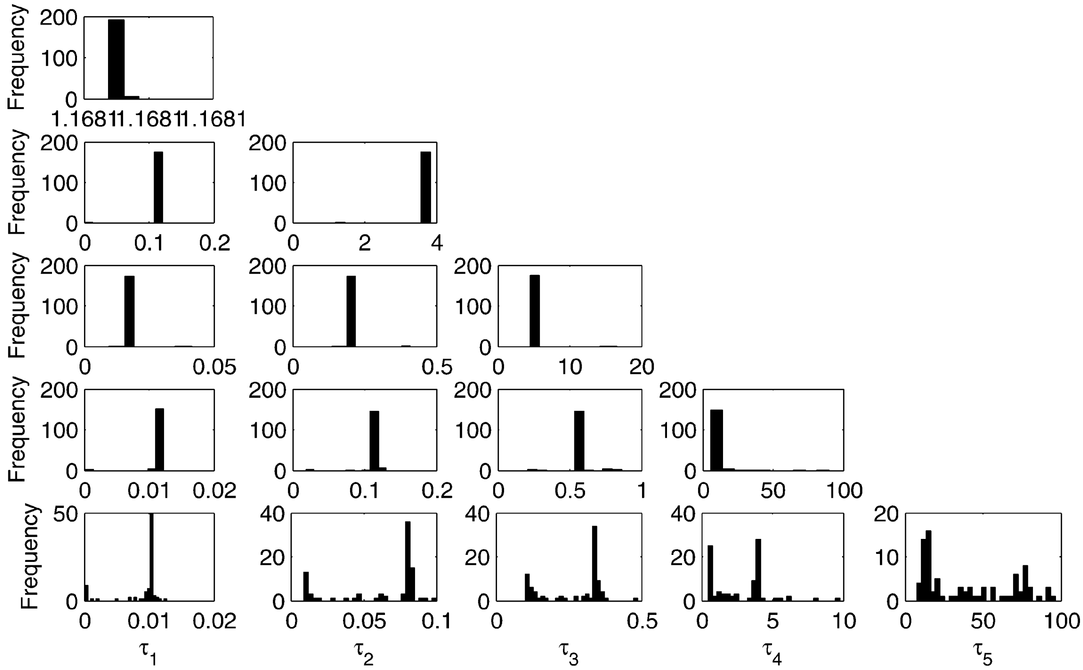

Fig. 6. Converged parameter estimates from 200 random initial values, corresponding to model orders from 1 to 5, where each row of histograms corresponds to a single model structure.

where $\tilde{\theta} \in [0, 1]$ was the transformed parameter vector; note that $\theta \in [-\infty, \infty]$. In practice, this formulation of the problem led to improved numerical properties: the associated problem was due to the presence of a long system time constant, which corresponded to a pole situated close to the edge of the unit circle. This pole would occasionally go unstable during the estimation procedure, probably because such a relatively long time constant acted like a constant offset in the short term and therefore growth or decay (that is, instability or stability) of the pole was insignificant over the recording scale of the data.

2) Data Preprocessing: The resting position of the eye (in each case lateral to the primary position) was estimated from inspection of all the eye-position traces. Traces that were interrupted within 400 ms of release by discontinuities of slope (possibly corresponding to active components such as small saccades that are often associated with ketamine anaesthesia) were excluded from further analysis. Each remaining trace was fitted from the time of its maximum velocity, rather than from the time of release. In theory, these two times should coincide for a pure viscoelastic system released instantaneously, that is, the acceleration time should be zero. However, in the actual traces, the time from release to peak velocity ranged from 8 to 20 ms, reflecting an unknown combination of (small) globe inertia and the release time of the forceps opening. Fitting from the time of peak velocity was an attempt to avoid these complexities: The full set of signals are shown in Fig. 5. Training and validation data sets were formed by partitioning adjacent trials, resulting in a training set of five trials and a validation set of five trials.

3) Modelling Results: The model structure detection problem for a real system is complicated by the fact that the optimal parameters for a given structure are unknown. The approach taken in this investigation was to assess the consistency of the parameter estimates by starting the estimation algorithm at different values. In practice this was accomplished by selecting the initial poles randomly from a uniform distribution in the range (0, 1). This incorporated the known prior information about the system poles, that is, stable and real-valued. The estimation algorithm was run from 200 different starting estimates for each model order.

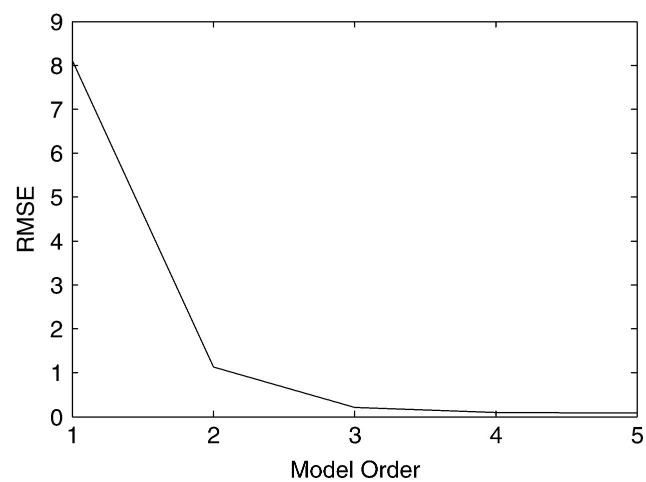

Fig. 7. RMSE corresponding to increasing model order.

Fig. 6 displays the parameter estimates in the form of histograms for model orders $n = 1$ to $n = 5$ (where each row of histograms corresponds to a single model structure). It is apparent that there is a consistency in the estimates at each model order $n = 1$ to $n = 4$; at model order $n = 5$, the consistency begins to degrade, which may be due to the relative high order causing overfitting of the data, and hence redundancy in the system description.

Structure detection focused on obtaining a parsimonious model description consisting of a few real poles (each of which corresponded to the time constant associated with a single Voigt element). Four poles were found to be sufficient to describe the

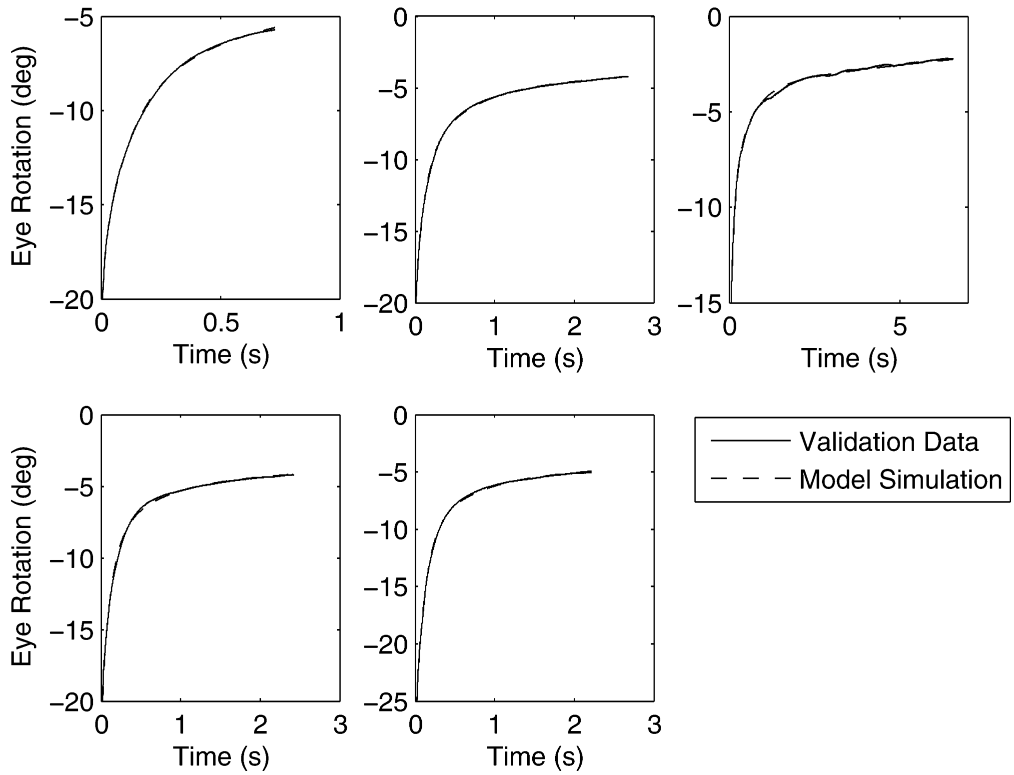

Fig. 8. Model prediction over validation data for the model of oculomotor plant dynamics.

system, evidence for which can be seen in Fig. 7, which shows the root mean square error (RMSE) of prediction for increasing model orders.

The model was validated in the time domain by verifying its prediction accuracy on the reserved independent data set; the model predicted output is shown in Fig. 8. The accuracy of the prediction demonstrates that this model was a good descriptor of the system.

The discrete-time state-space model was identified as

$$A = \begin{bmatrix} 0.9139 & 0 & 0 & 0 \\ 0 & 0.9910 & 0 & 0 \\ 0 & 0 & 0.9982 & 0 \\ 0 & 0 & 0 & 0.9999 \end{bmatrix}. \quad (43)$$

This model description corresponds to continuous-time domain time constants of $\tau_1 = 0.0111$ s, $\tau_2 = 0.110$ s, $\tau_3 = 0.543$ s, and $\tau_4 = 12.5$ s.

These modeling results correspond closely to those presented in [6] in terms of both structure (4 visco-elastic units) and parameter estimates, where the time constants were reported to be approximately 0.01 s, 0.1 s, 1 s, and 10 s. These results provide some validation as to the success of the generic modeling algorithm for multiple short-time-duration signals proposed here.

VII. Conclusion

An SLS parameter estimation method was derived that is useful for identifying a system from multiple short-time-duration signals. The proposed approach utilized an output error model; such a description naturally leads to a correct treatment of measurement noise, unlike the ARX model, which was demonstrated to potentially lead to significant bias at even high SNR levels. Most conceivable physiological signals would contain measurement noise, hence the output error approach would appear to be widely applicable. The passive characteristics of the oculomotor plant were identified using the SLS algorithm; the results were validated by reference to a previous study. This demonstrated the utility of the proposed method because the oculomotor plant is a demanding system to identify, incorporating a wide range of time constants.

References

[1] L. Ljung, *System Identification—Theory for the User*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 1999.

[2] L. Ljung, "Initialisation aspects for subspace and output error identification methods," presented at the 5th Eur. Control Conf., Cambridge, U.K., 2003.

[3] G. F. Franklin, J. D. Powell, and A. Emami-Naeini, *Feedback Control of Dynamic Systems*, 4th ed. Upper Saddle River, NJ: Prentice-Hall, 2001.

[4] K. Hunt, M. Munih, and N. N. Donaldson, "Investigation of the Hammerstein hypothesis in the modeling of electrically stimulated muscle," *IEEE Trans. Biomed. Eng.*, vol. 45, no. 8, pp. 998–1009, Aug. 1998.

[5] J. Bobet, E. R. Gossen, and R. B. Stein, "A comparison of models of force production during stimulated isometric ankle dorsiflexion in humans," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 13, no. 4, pp. 444–451, 2005.

[6] S. Sklavos, J. Porrill, C. R. S. Kaneko, and P. Dean, "Evidence for wide range of time scales in oculomotor plant dynamics: Implications for models of eye movement control," *Vision Res.*, vol. 45, pp. 1525–1542, 2005.

[7] S. Sklavos, D. M. Dimitrova, S. J. Goldberg, J. Porrill, and P. Dean, "Long time-constant behavior of the coulomotor plant in barbiturate anesthetized primate," *J. Neurophysiol.*, vol. 95, no. 2, pp. 774–782, 2006.

[8] G. H. Golub and V. Pereyra, "The differentiation of pseudo-inverses and non-linear least squares problems whose variables separate," *SIAM J. Numer. Anal.*, vol. 10, no. 2, pp. 413–432, 1973.

[9] J. Bruls, C. T. Chou, B. R. J. Haverkamp, and M. Verhaegen, "Linear and non-linear system identification using separable least squares," *Eur. J. Control*, vol. 5, pp. 116–128, 1999.
[10] A. F. Fuchs, C. A. Scudder, and C. R. S. Kaneko, "Discharge patterns and recruitment order of identified motoneurons and internuclear neurons in the monkey abducens nucleus," *J. Neurophysiol.*, vol. 60, no. 6, pp. 1874–1895, 1988.
[11] K. E. Cullen, H. L. Galiana, and P. A. Sylvestre, "Comparing extraocular motoneuron discharges during head-restrained saccades and head-unrestrained gaze shifts," *J. Neurophysiol.*, vol. 83, pp. 630–637, 2000.
[12] P. Dean, J. Porrill, and J. V. Stone, "Decorrelation control by the cerebellum achieves oculomotor plant compensation in simulated vestibulo-ocular reflex," *Proc. R. Soc. Lond. B*, vol. 269, pp. 1895–1904, 2002.
[13] P. M. Blazquez, Y. Hirata, S. A. Heiney, A. M. Green, and S. M. Highstein, "Cerebellar signatures of vestibulo-ocular reflex motor learning," *J. Neurosci.*, vol. 23, pp. 9742–9751, 2003.
[14] M. Verhaegen, "Subspace model identification part 3. Analysis of the ordinary output-error state-space model identification algorithm," *Int. J. Control*, vol. 58, no. 3, pp. 555–586, 1993.
[15] B. R. J. Haverkamp, "State space identification: Theory and practice," Ph.D. dissertation, Technical Univ. Delft, Delft, The Netherlands, 2001.
[16] D. F. Shanno, "Conditioning of quasi-Newton methods for function minimization," *Math. Computat.*, vol. 24, pp. 647–656, 1970.
[17] G. A. F. Seber and C. J. Wild*, Nonlinear Regression*. New York: Wiley, 2003.
[18] D. A. Robinson, "The mechanics of human saccadic eye movement," *J. Physiol.*, vol. 174, pp. 245–264, 1964.

Sean R. Anderson was born in Reading, U.K., in 1979. He received the M.Eng. degree in control systems engineering from the Department of Automatic Control and Systems Engineering, University of Sheffield, U.K., in 2001 and the Ph.D. degree from the Department of Chemical and Process Engineering in 2005.

He is currently a Research Associate in the Neural Algorithms Research Group, Department of Psychology, University of Sheffield, where he studies the central nervous system, using systems engineering techniques to interrogate methods of biological adaptive control. His research interests include identification of continuous- and discrete-time dynamic systems, nonlinear systems modeling, and the study of adaptive control in biological systems.

Paul Dean received the M.A. degree in physiology with psychology from the University of Cambridge, U.K., and the D.Phil. degree from the University of Oxford, U.K.

He is currently a Professor in the Department of Psychology at the University of Sheffield, U.K., and is a member of the recently established Centre for Signal Processing in Neuroimaging and Systems Neuroscience. His major research interest is in producing computational models of neural systems that are based on both biological data and current developments in control engineering, signal processing, and robotics. These models are intended to serve as a vehicle for two-way communication between biological and physical sciences, allowing roboticists to use new discoveries in biology, and biologists to interpret their findings in the light of current developments in signal processing.

Visakan Kadirkamanathan (S'89–M'90) was born in Jaffna, Sri Lanka He received the B.A. and Ph.D. degrees from the Department of Engineering, University of Cambridge, U.K.

Previously, he has held Research Associate positions at the University of Surrey, U.K., and the University of Cambridge, U.K. In 1993, he joined the University of Sheffield as a Lecturer, where he is currently a Professor in the Signal Processing and Complex Systems Research Group of the Department of Automatic Control and Systems Engineering. He has coauthored a book on intelligent control titled *Functional Adaptive Control: An Intelligent Systems Approach* (London, U.K.: Springer, 2001) and has published more than 120 papers in refereed international journals and proceedings of international conferences. His research interests include nonlinear signal processing and control, optimization and decision support, signal and fault detection and neural networks with applications in biomedical systems modeling, aircraft engine health monitoring, and systems biology.

Prof. Kadirkamanathan is the Co-Editor of the *International Journal of Systems Science* and has served as an Associate Editor for the IEEE TRANSACTIONS ON NEURAL NETWORKS and the IEEE TRANSACTIONS ON SYSTEMS, MAN AND CYBERNETICS, PART B.

Chris R. S. Kaneko received B.A. degrees in psychology and zoology from the University of California at Los Angeles in 1968 and the M.A. degree in psychology and the Ph.D. degree in zoology from the University of Iowa, Iowa City, in 1970 and 1973, respectively.

He was a Postdoctoral Fellow in the Division of Cellular Neurobiology of the Department of Neuroscience at Einstein College of Medicine and Neurophysiology at the University of Washington, Seattle. He is currently a Research Scientist in the Department of Physiology and Biophysics of the University of Washington. His research interests are oculomotor system neurophysiology and saccadic eye movement control, in particular.

John Porrill trained in mathematics and theoretical physics at the University of Cambridge, U.K., working with J. Stewart on topics in classical general relativity. He moved to the University of Sheffield, U.K., to work on human and computer vision with J. Mayhew and J. Frisby and has a continuing interest in the psychophysics of human stereo vision. Recent research projects include video tracking of eye movements and building the EyeLab open source model of the extraocular muscle system. He is currently working on computational models of the role of the cerebellum in the control of eye movements.