Estimation of Time-Varying Delay Time in Nonstationary Linear Systems: An Approach to Monitor Human Operator Adaptation in Manual Tracking Tasks

Erwin R. Boer, *Member, IEEE*, and Robert V. Kenyon, *Member, IEEE*

***Abstract*— Adaptability is one of man's advantages over machines. Perhaps one of the reasons for our limited understanding about human adaptation during manual tracking tasks is that we have only limited tools to identify the model coefficients (especially delay time) of an adapting human operator. In this paper, we introduce a discrete time recursive delay identifier (RDI) capable of simultaneously estimating a human operator's nonstationary delay time and linear model coefficients. At its core lies the extended Kalman filter (EKF). Our goal to obtain fractional delay time estimates was realized by using the bicubic interpolation scheme as part of the EKF to provide subsample magnitude and derivative estimates of the observed input/output time series. While this theoretically limits the RDI's applicability to bandlimited or differentiable signals, this is seldom a concern in practice. Based on data from simulated and experimental time varying tracking tasks, we show the RDI's potential to substantially increase our understanding about human adaptations thus perhaps offering new avenues for machine adaptation.**

I. Introduction

MODELING and identifying the characteristics of a human controller made great strides in the late 1960's and early 1970's [1]. Human operator models were refined as new identification techniques became available. These efforts produced extremely useful models such as the linear crossover model of McRuer [2] and the optimal control model for compensatory tracking of Baron *et al*. [3], [4], which was later extended by Tomizuka to cover preview tracking [5], [6]. What made this work such a challenging endeavor is that the human operator (HO) is a nonlinear, often time-varying controller whose operating coefficients and delay characteristic can vary rapidly with factors such as task demands, motivation, workload and fatigue [7], [8]. These factors limited the range of applications for which the various models would hold. These same characteristics make identification of the HO equally challenging. Even though many different identifiers have been applied to this problem since the introduction of the Kalman filter (KF) in 1960 [9], a satisfactory solution to recursively estimate time varying HO delay-time remained a problem in that the assumption of a fixed delay could not be circumvented within the existing paradigms. While many solutions or fixes have been reported (Section II), they do not provide a straight forward or otherwise satisfactory solution to tackle the problem of identifying an adapting or, in more general terms, a time varying HO. For example, as a HO's attention and/or motivation fluctuates, the rate and manner in which sensory information is processed and acted on may also fluctuate thus introducing more or less unintended changes in the HO. These changes may not necessarily qualify as adaptation. The difficulties encountered in estimating time varying HO delay time have limited our ability to understand how the forces driving human adaptation affect specific aspects of human control.

Delay time in a human operator is primarily a result of transport delays and central nervous system latencies. Depending on the input bandwidth, it may also include neuromuscular lag, high-frequency lead equalization, as well as a time varying component that depends on attention level, task difficulty, etc. Since differentiation between these sources is generally very difficult or impossible, delay time (sometimes referred to as effective delay time) is defined as that portion of the phase lag that linearly increases with frequency over the measurement bandwidth. Therefore, delay time is considered a model parameter and not necessarily one particular physiological property.

In this paper, a new recursive delay-time identifier (RDI) is introduced that simultaneously estimates the varying delay-time and linear model coefficients of an adapting system like a HO responding to changing conditions. This new identifier is based on the extended Kalman filter (EKF) [10], [11] which is commonly used to identify parameters of a nonlinear model.

Since the delay time is modeled as a shift operator in the discrete time domain, which is a nonlinear operation, it was natural to resort to the EKF. To further bypass the fact that the delay time is expressed in integer multiples of the sampling interval, we adopted a scheme to obtain fractional delay time estimates by embedding the bilinear interpolation scheme in the EKF to estimate subsample magnitudes and derivatives of the observations.

To elucidate these issues, consider a highly simplified case in which the delay time d_n of a noise free pure gain system $y_n = b_0 r_{n+d_n}$ needs to be estimated. Nearly all identification methods are based on minimizing a quadratic cost function with as one of its terms the estimation error $e_n = y_n - \hat{y}_n$,

Manuscript received April 14, 1996; revised November 25, 1996. This work was supported by the NSF under Grant IRI-9213822.

E. R. Boer is with Nissan Cambridge Basic Research, Cambridge, MA 02142 USA (e-mail: erwin@pathfinder.cbr.com).

R. V. Kenyon is with the Department of Electrical Engineering and Computer Science, University of Illinois at Chicago, Chicago, IL 60607 USA.

Publisher Item Identifier S 1083-4427(98)00135-0.

which for this simple example is $e_n = y_n - b_0 r_{n+\hat{d}_n}$. Coefficient estimation relies on the partial derivative of e_n with respect to each of the unknown coefficients. In this example, with $\hat{d}_n$ as the only unknown, the partial derivative becomes $b_0 \dot{r}_{n+d_n}$, where the dot above the r denotes the time derivative estimate. This partial derivative is the sensitivity of the output estimation error to changes in delay time and is used to guide gradient searches to the true coefficient values (delay time) very much like the approach taken in the EKF (see below). Based on the specification that d_n be a real instead of an integer, calculation of this quantity as well as e_n itself require that the continuous time versions of time series $\{y_n\}$ and $\{r_n\}$ are computable. This simple example shows that the only true issue is differentiability of the input signal. In the area of manual control, and many others, observations are often bandlimited and noise is seldom truly white in the sense that it spans all frequencies up to the Nyquist frequency. Keeping this in mind, continuous time magnitude and derivative estimates of the observations at subsample times can be computed (see Appendix).

Following the background section, the RDI is introduced for the general case of this simple example (i.e., identification of higher order systems with time varying coefficients and delay time based on bandlimited noisy observations) and applied to simulated and experimental data. We believe that this new approach will greatly enhance our ability to quantify the behavior of human adaptation mechanisms.

II. Background

Many different approaches have been proposed to circumvent the problem that the delay time is modeled as a nonlinear operator in a discrete time system representation and has to be fixed before applying identification schemes such as least squares. A frequently used solution is to approximate the delay time by an all-pass linear system thus reducing delay time identification to linear coefficient estimation [12]. Jex used the first order Pade approximation to characterize human operator delay time in his Critical Tracking Task experiments [13]. Even if the delay time is known to its nearest whole number of sample intervals, the remaining unaccounted for fractional delay time can have a significant effect on the moving average (MA) portion of a linear model [14] (see Section III-A for a brief introduction to ARMAX models). In case of low bandwidth systems that are excited with only limited bandwidth signals, as is generally the case in manual control, it is often difficult to distinguish between a pure delay and a phase lag resulting from: i) a pole that is associated with a cutoff frequency outside the system bandwidth or ii) a zero in the right half s-plane (i.e., nonminimum phase system). In these cases, the MA or autoregressive (AR) coefficients can easily model part of the pure delay [15], [16]. The problem with modeling the effect of a pure delay time with a linear coefficient model is that the total number of model coefficients increases which results in difficulties similar to those associated with trying to identify over parameterized systems [17].

Direct delay time identification based on minimizing a cost function requires differentiation with respect to the delay time (see Section IV-C). Banyasz and Keviczky [18] provide an insightful derivation based on minimizing a quadratic error function based on a pure gain pure delay system. Unfortunately, they fail to recognize many of the problems associated with such an approach such as the fact that: i) the cost function is not unimodal with respect to delay time [19], ii) derivation with respect to delay time involves differentiation of the observations which can not be blindly performed in case of stochastic signals, and iii) problems associated with identification of linear ARMAX models in case of small time steps [17] on which their derivations were based.

The most promising optimization based approach so far is one in which the estimation error is filtered such that the resulting generalized criterion function with respect to the delay time becomes unimodal are given in [19], [20]. However, their multistep iterative methods do not apply to time varying systems. Tuch *et al.* [21] apply straightforward minimization of a quadratic error function to continuous time linear time-invariant systems. They assume noise-free observations and initially known linear model coefficients. For this highly specialized case, nonbiased delay time estimates are obtained. The fact that their method only applies to time invariant systems limits its applicability to nonadaptive systems. While these methods appear to be mathematically elegant ways to simultaneously identify delay time and linear coefficients in stationary systems, they are not applicable to adapting human operators.

One semi-successful computationally expensive approach is one in which a large set of models each associated with a different delay time are recursively identified and a criterion function is used to pick the one that represents the observations best at a given time instance [10], [22], [23]. The recursive nature of these methods makes them applicable to time varying systems. These methods do not produce fractional delay time estimate even though fixes can be devised to obtain such estimates (e.g., subsequent smoothing of the delay time series, followed by resampling of the observations to obtain the final linear coefficient estimates based on a zero delay time) [24].

If the input and output signals of a system differ only by a time shift but not in spectral composition, as is often the case in communication applications, a wide variety of delay time identification schemes are available many of which are based on correlation analysis [25]–[27]. These approaches are applicable in a system identification setting when the system at hand is known except for its delay time. An example is the correlation based approach combined with the least squares method in Zheng and Feng [28]. Inspired by Box and Jenkins [29], they first identify the AR coefficients, then delay time followed by the MA coefficients. Again, these methods are not applicable to systems quickly changing over time.

III. Recursive Delay Identifier

Most recursive identification methods, including the EKF, are based on minimizing a cost function, which means that the derivative of this function with respect to the model coefficient needs to exist in order to be able to search for a minimum. One of the fundamental problems with delay time identification in the discrete time domain is the fact that the delay time is

modeled as a shift operator with an integer argument (whole number of time steps) against which one can not differentiate. In the RDI, this issue is bypassed by incorporating a bicubic interpolation filter to produce subsample estimates of the discrete time signal's magnitudes [30] and derivatives (see the Appendix). The validity of this extension is based on the differentiability of the system's input and output signals thus limiting the RDI's applicability to bandlimited signals sampled at a frequency significantly higher than their bandwidth. The delay time estimation component of the RDI operates on these subsample magnitude and derivative estimates hence extending the otherwise unit shift operator to a fractional shift operator against which can be numerically differentiated. For linear coefficient estimation, the observations are decimated and locally resampled based on the current best fractional delay time estimate and then used to update the linear coefficient estimates.

Taking the derivative with respect to the delay shift involves differentiation of the input signal with respect to time. For this to be possible, a continuous time representation of the reference signal needs to be constructed. For deterministic input signals or equivalently, stochastic signals whose bandwidth is much lower than the Nyquist frequency, such a representation can be obtained through interpolation. Ito calculus provides a theory for differentiation of stochastic processes; this approach was not pursued because most practical input signals are deterministic, and even if this were not the case, once a bandlimited stochastic process is sampled at a frequency well beyond its bandwidth, it can without additional loss of information be treated as a deterministic signal [31].

By appending the Rauch–Tung–Striebel nonlinear fixed interval optimal smoother to the EKF (i.e., RDI) [11], [32], [33], more accurate model coefficient estimates are obtained because more information is used [34]. Theoretical properties and efficient implementations of the optimal smoother were established in [35]–[37]. Surprisingly few applications have been reported on their use in system identification perhaps due to their non-realtime character. Norton used them to study daily rainfall in-flow dynamics of the King River in SW Tasmania [38]. To the best of our knowledge, they have not been applied to HO data. To enhance ones understanding about a system, using optimal smoothing in the model identification is certainly advantageous. It not only improves model coefficient estimates, it also eliminates identification lags introduced by the EKF's own time constants (determined by the *a priori* noise covariance matrices) thus improving the interpretation of observed changes in model coefficients [39]. In studying adaptation, it is important to keep in mind that the identifier used is also a filter with its own time constants. Ideally, the identifier's time constants should be smaller than those of the adaptation mechanisms involved so that the identifier can converge fast enough to track changes in the model caused by adaptation.

A. Kalman Filter

State space descriptions of linear models form the basis for much of the modeling, control and system identification literature. The linear ARMAX model has become a standard tool for both system description and control design. ARMAX models are ARMA or Auto Regressive Moving Average models eXtended with a Moving Average external disturbance term. An example of an ARMAX representation is the following:

$$y_n = a_1 y_{n-1} + a_2 y_{n-2} + b_0 r_n + b_1 r_{n-1} + v_n + c_1 v_{n-1} + c_2 v_{n-2}$$

where a_1 and a_2 are the AR coefficients, b_0 and b_1 the MA coefficients and c_1 and c_2 the X coefficients. The input time series is denoted by $\{r_n\}$, the output by $\{y_n\}$, and the output observation noise by $\{v_n\}$. Detailed information pertaining to ARMAX state space system representations as well as Kalman filter principles can for instance be found in [11] and [17]. The linear state space model equations are

$$\theta_n = \theta_{n-1} + w_n \tag{1a}$$

$$y_n = \theta_n^T \phi_n + v_n \tag{1b}$$

where θ_n is the model coefficient vector (assuming a single output system). The white Gaussian noise process $\{w_n\}$, with zero mean and covariance matrix Q_n, represents the model coefficient noise. The diagonal entries in Q_n can also be interpreted to indicate the rate at which these model coefficients are expected to change over time (see below). The vector ϕ_n holds the input and output samples on which corresponding entries in θ_n operate. The Gaussian noise process $\{v_n\}$, with covariance matrix (scalar in case of a single-output system) R_n, represents the observation noise in the output time series $\{y_n\}$. In terms of the above ARMAX model example, θ_n and ϕ_n become

$$\theta_n = [a_1, a_2, b_0, b_1, c_1, c_2]$$
$$\Phi_n = [y_{n-1}, y_{n-2}, r_n, r_{n-1}, v_{n-1}, v_{n-2}].$$

To clarify the relationship between expected model coefficient adaptation and the selection of $\{w_n\}$'s covariance matrix Q_n, a few words about one approach to derive the Kalman filter algorithm, namely through minimization of a cost function that assigns a cost to variations in the estimated model coefficients $\hat{\theta}_n$ and one to the error between the observed output y_n and estimated output $\hat{y}_n = \hat{\theta}_n^T \varphi_n$. These costs Q_n^{-1} and R_n^{-1} are the inverse of the covariance matrices of the noise processes $\{w_n\}$ and $\{v_n\}$, respectively [10], [40]. The cost function is therefore

$$J = \sum_{n=0}^{N-1} [\tfrac{1}{2}(\hat{\theta}_n - \hat{\theta}_{n-1})^T Q_n^{-1} (\hat{\theta}_n - \hat{\theta}_{n-1}) + \tfrac{1}{2}(y_n - \hat{\theta}_n^T \phi_n)^T R_n^{-1} (y_n - \hat{\theta}_n^T \phi_n)].$$

From this equation, it is clear that an increase in Q_n, decreases the cost on changing the model coefficients, thus allowing for faster adaptation. As was mentioned earlier, the dynamics of the Kalman filter itself are constrained by these two covariance matrices. These issues are especially important if we want to identify step changes in system behavior because the estimated model coefficient time series are partially affected by the Kalman filter's own bandwidth (slew rate).

B. Extended Kalman Filter

The EKF was developed for nonlinear systems of the following form:

$$\theta_{n+1} = f_n(\theta_n) + w_n \tag{2a}$$

$$y_n = h_n(\theta_n) + v_n \tag{2b}$$

which is the discrete time representation given in [10]. Their nomenclature is adjusted to conform with the rest of this paper, plus the equations are reconfigured for system identification purposes. The concept of the EKF is to represent these equations by their first order Taylor expansions about the conditional means $\hat{\theta}_{n/n}$ and $\hat{\theta}_{n/n-1}$

$$f_n(\theta_n) = f_n(\hat{\theta}_{n/n}) + F_n(\theta_n - \hat{\theta}_{n/n})$$
$$h_n(\theta_n) = h_n(\hat{\theta}_{n/n-1}) + H_n(\theta_n - \hat{\theta}_{n/n-1})$$

where $\hat{\theta}_{n/n-1}$ denotes the estimate of θ at time step n whereby all observations up to and including $n-1$ are utilized. Also

$$F_n = \left.\frac{\partial f_n(\theta)}{\partial \theta}\right|_{\theta=\hat{\theta}_{n/n}}$$

and

$$H_n = \left.\frac{\partial h_n(\theta)}{\partial \theta}\right|_{\theta=\hat{\theta}_{n/n-1}}.$$

In this paper, the system model coefficients are assumed to vary according to (1a) so that (2a) reduces to (1a). The EKF equations become

$$\hat{\theta}_n = \hat{\theta}_{n-1} + K_n[y_n - h_n(\hat{\theta}_{n-1})]$$
$$K_n = P_{n-1}H_n[H_n^T P_{n-1} H_n + R_n]^{-1}$$
$$P_n = [I - K_n H_n^T]P_{n-1} + Q$$

which is the basic form used in the RDI. The delay time estimate is assumed to be the first entry in θ_n.

C. Difference Between Linear Coefficient and Delay Time Identification

The momentary convergence rates for delay time and linear coefficient identification differ depending on the current input output signal characteristics. The delay time converges fastest when the derivative of the input signal reaches its maximum, whereas the linear coefficients converge fastest at the highest magnitudes in the observations [highest signal-to-noise ratio (SNR)]. To elucidate this difference, consider the following pure gain pure delay system $y_n = b_{0_n} r_{n+d_n} + v_n$, with unknown gain b_{0_n} and delay d_n. In a KF, the convergence rate of a given model coefficient depends on the sensitivity of the estimation error with respect to this coefficient (Kalman gain). The estimation error e_n is defined as $e_n = y_n - \hat{y}_n$ where $\hat{y}_n = \hat{b}_{0_{n-1}} r_{n+\hat{d}_{n-1}}$ for our example. The sensitivity with respect to the linear coefficient $\hat{b}_{0_{n-1}}$ is

$$\frac{\partial e_n}{\partial \hat{b}_{0_{n-1}}} = -r_{n+\hat{d}_{n-1}}$$

and we see that the effect is largest when the input signal $r_{n+\hat{d}_{n-1}}$ reaches high magnitudes. The effect of a change in delay time on the estimation error is

$$\frac{\partial e_n}{\partial \hat{d}_n} = -\hat{b}_{0_{n-1}} \frac{\partial r_{n+\hat{d}_{n-1}}}{\partial \hat{d}_n}$$

which tells us that the effect is highest when the slope in the input signal is highest.

From an experimental point of view, to assure that the RDI focuses identification equally between the two classes of model coefficients, the input signal should exhibit an even distribution of high peaks and steep slopes. This means that binary signals should not be used on systems with time varying delay time. Note that these signals are also not bandlimited as is required for the RDI to operate properly. Input signal selection in any experiment designed to gain an understanding about a system is very important and should be linked closely to the class of models one believes describes the system best. Ljung provides a good discussion on the practical aspect of input signal selection for linear stationary systems [17].

D. Calculating H_n

This section describes how to obtain H_n, the partial derivative of $h_n(\hat{\theta}_n)$ with respect $\hat{\theta}_n$. The partial derivative of $h_n(\hat{\theta}_n)$ with respect to the AR coefficients and the colored noise coefficients is simply the vector containing the corresponding past output samples. On the other hand, the partial derivatives with respect to the MA coefficients depend on the delay time and subsample estimates of the input signal magnitude while partial differentiation with respect to the delay time depends on the delay time and subsample estimates of the input signal derivative. These estimates are calculated at subsample times $n - m + \hat{d}_n$ where $m \in \{0, 1, \cdots nb\}$ and nb is the order of the MA portion of the model. This is best clarified with an example. Assume the following ARMAX (1, 1, 0) model:

$$\hat{y}_n = h_n(\hat{\theta}_n) = \hat{a}_{1_n} y_{n-1} + \hat{b}_{0_n} r_{n+\hat{d}_n} + \hat{b}_{1_n} r_{n-1+\hat{d}_n}$$

for which $\hat{\theta}_n$ is defined as

$$\hat{\theta}_n = [\hat{d}_n, \hat{a}_{1_n}, \hat{b}_{0_n}, \hat{b}_{1_n}]^T.$$

The partial derivative H_n of $h_n(\hat{\theta}_n)$ with respect to at time step n becomes

$$H_n = [\hat{b}_{0_n} \dot{r}_{n+\hat{d}_n} + \hat{b}_{1_n} \dot{r}_{n-1+\hat{d}_n}, y_{n-1}, r_{n+\hat{d}_n}, r_{n-1+\hat{d}_n}]$$

where $\dot{r}_{n-m+\hat{d}_n}$ stands for the time derivative of the continuous time representation of $r_{n-m+\hat{d}_n}$ at time $(n-m+\hat{d}_n)T_s$. In these parameterizations, $\hat{d}_n$ is a floating point value (fractional number of time steps) which basically means that $r_{n-m+\hat{d}_n}$ needs to be reconstructed as the continuous time representation from the time series $\{r_n\}$. In the Appendix, it is shown that the bicubic interpolation method most accurately reproduces the continuous time signal in magnitude and first order derivative based on a discretely sampled bandlimited signal consisting of a sum of sinusoids (often used in manual control).

The Kalman filter is optimal if the assumption of a Gaussian noise signal holds. It is easily demonstrated that the observed output noise is Gaussian for a linear model if $\{v_n\}$ and $\{w_n\}$ are two independent Gaussian processes [41]. For a nonlinear system, this no longer holds and optimality is no

longer guaranteed [11]. The exact consequences in our case where $h_n(\hat{\theta}_n)$ is linear except in the delay time still needs to be determined; the issue being the delay time coefficient in $\hat{\theta}_n$ is an operator and that subsample magnitude and derivative estimates of the input are used. In principle, the RDI is a hybrid between the discrete and continuous time domains. In effect, it shifts the input sample times around before applying the linear coefficient estimation. To quote Gelb [11] on the use of the EKF: "There is no guarantee that the actual estimate obtained will be close to the truly optimal estimate. Fortunately, the extended Kalman filter has been found to yield accurate estimates in a number of important practical applications. Because of this experience and its similarity to the conventional Kalman filter, it is usually one of the first methods to be tried for any nonlinear filtering problem."

To minimize the problems associated with identification based on over-sampled time series [17], linear coefficient estimation is based on observations decimated by a factor of λ. On the other hand, subsample magnitude and derivative estimation are still calculated based on the original observation sampled at T_S second intervals to minimize interpolation noise. To express the delay time in terms of time steps associated with the decimated observations (i.e., $T_d = \lambda T_s$) while $\partial h_n(\hat{\theta}_{n-1})/\partial \hat{d}_{n-1}$ is calculated based on the non-decimated time series, the expression for $H_n[0]$ becomes $\lambda \partial h_n(\hat{\theta}_{n-1})/\partial \hat{d}_{n-1}$.

E. Optimal Smoother

To reduce the estimation noise and eliminate some of the dynamical effects of the identifier (see above), a non-linear smoother was applied on the output of the extended Kalman filter. The applied fixed interval smoother is the Rauch–Tung–Striebel algorithm and is repeated here in a form suitable for identification [10], [11]:

$$\begin{aligned}\hat{\theta}_{n/N-1} &= \hat{\theta}_n + K_n(\hat{\theta}_{n+1/N-1} - \hat{\theta}_n)\\ K_n &= P_n(P_n + Q_n)^{-1}\\ P_{n/N-1} &= P_n + K_n(P_{n+1/N-1} - P_n - Q_n)K_n^T\end{aligned}$$

where $\hat{\theta}_n$ and P_n are the results obtained from the forward-in-time EKF and where $\hat{\theta}_{n|N-1}$ contains the backward filtered model coefficients and $P_{n|N-1}$ is an estimate of the model coefficient covariance matrix. As Gelb pointed out, the true interpretation of $P_{n|N-1}$ or equivalently P_n is dubious for nonlinear time varying systems [11]. He suggests to use Monte-Carlo simulations to probe the model coefficient space for possibly better solutions. Most importantly, one should use common sense in interpreting the data.

F. Constraints on the Recursive Delay Time Identifier

The RDI algorithm as it is presented above does not take into account that the delay time estimates may converge to a wrong value because of the multimodality of the cost function with respect to delay time [19]. Furthermore, if the delay time estimate is incorrect, the linear coefficients will not converge to the correct values (and vise versa). Here we present monitors to determine when divergence into a local instead of global minimum is likely to occur. These monitors can also be used to stop identification when the chance of entering a local minimum regime is high, thus constraining the RDI to operate only around the global minimum. Note that this assumes the initial model coefficient estimate to lie within "reach" of the global minimum.

Before we introduce the constraints used for monitoring identifier divergence, a simple example is presented to demonstrate the duality between the mutual effects of errors in delay time and linear coefficient estimation. Assume a noise free pure gain pure delay system with a sinusoidal input. If the true gain is one and the delay time is off by half a period of the sinusoid, the gain will converge to negative one. This is driven by a sign difference between the observed and estimated output signals. Similarly, if the gain is incorrectly estimated to be negative one, the delay time will converge to a value off by half a period of the sinusoid (positive or negative). This is driven by a sign difference between the derivatives of the observed and estimated output signals. One characteristic associated with exact model coefficient estimates (in the noise free case) is that the signs of the magnitudes (derivative) between the observed and estimated output are equal at all time steps. It is easily seen that small estimation errors are characterized by inequalities in these signs whenever the magnitude (derivative) of output signal (observed or estimated) crosses zero. Note that these are exactly the times when the rate of convergence in linear coefficients (delay time) is lowest. By turning the identifier off whenever the magnitude (derivative) signs are unequal, it is guaranteed that the estimates remain close to the global minimum.

The observations made regarding the simple example still hold if the input signal becomes a more complex bandlimited signal and the model becomes a full blown ARMAX model. In summary, when derivative constraint

$$\operatorname{sign}(\dot{y}_n) = \operatorname{sign}(\dot{y}_n(\hat{\theta}_n))$$

or when the magnitude constraint

$$\operatorname{sign}(y_n) = \operatorname{sign}(y_n(\hat{\theta}_n))$$

do not hold, identification is temporarily halted until the constraints hold again. We do not claim mathematical rigor but simply provide an intuitively straightforward method to prevent runaway coefficient estimates. Note that the magnitude constraint can also be used in case of systems with a fixed but unknown delay time. If the delay time is estimated incorrectly one will, depending on the *a priori* noise covariance estimates, either observe variations in linear coefficient estimates that are correlated with the input signal or that the magnitude constraint does not hold for extended periods of time. In fact it is worthwhile to explore how the average duration that the constraint does not hold relates to the necessary delay time correction (this was not pursued here).

Even with the above constraints the RDI has been observed to produce misleading results when: i) the true delay time changes too fast for the time constant of the identifier; ii) the true delay time makes a jump greater than half the period of the bandwidth of the observations hence causing the delay

time to jump out of the global convergence zone; and iii) the initial delay time estimate is off by more than half the period of the bandwidth of the observations which places it outside the global convergence zone. The first case can be solved by increasing the delay time noise variance in the model coefficient covariance matrix Q (i.e., diagonal entry). While the third case simply requires re-estimation of the initial delay time, the second case poses a real problem in that it requires an artificially induced jump in the delay time estimate such that it falls in the global convergence zone again. Such large changes are generally caused by a malfunction in which case the whole system dynamics can change and one may as well reinitialize the identifier and restart the identification at that time step with a new set of model coefficient estimates as well as a coefficient estimation noise covariance matrix P with large diagonal entries to reflect the uncertainty in these new model coefficient settings. As a general rule, if the constraints do not begin to hold for a significant portion of the time, a new set of initial estimates should be tried. Again common sense should be used in selecting these model coefficient estimates.

G. Application

Besides determining the correct model order, selecting the *a priori* model coefficient covariance matrix Q and the observation noise variance R is often difficult [17]. A few guidelines specifically geared to the identification of time varying systems are presented here. First, the covariance matrix Q in the RDI equations is split up into Q_d for the delay time and Q_c for all the linear coefficients. Since the delay time is expressed in terms of fractional time steps, Q_d also has to be expressed in terms of the sampling interval after decimation (i.e., $T_d = \lambda T_s$).

For the RDI (or any recursive identifier) to return meaningful model coefficient estimates of an adapting system, its bandwidth should exceed the bandwidth of adaptation mechanisms. In other words, the RDI should be able to track changes in model coefficients with minimal lag. This can always be accomplished by increasing the diagonal elements (variances) of the *a priori* model coefficient covariance matrix Q. Estimating the *a priori* noise variance $\sigma^2_{\theta(i)}$ in model coefficient $\theta(i)$ requires knowledge about the maximum rate per time step (after decimation) at which $\theta(i)$ is expected to vary (call this $\Delta\theta(i)$). The variance of $\theta(i)$ after n time steps equals $n\sigma^2_{w(i)}$ whereby $\sigma^2_{w(i)}$ is the variance of the Gaussian random variable $\{w(i)_n\}$ in the update equation $\theta(i)_n = \theta(i)_{n-1} + w(i)_n$ [41]. By setting $\sigma_{w(i)}$ to $\Delta\theta(i), Q_d$ becomes $\Delta\theta(i)^2$. If the resulting coefficient time series exhibits large segments with monotone increases or decreases, the respective variance may have been set too low. Similarly, if the coefficient fluctuates wildly around some visible mean trend, the variance was most likely over estimated. Finally, since only the ratio Q/R has an effect on convergence rates, the resulting variance of the estimation error can be used to normalize R to this value and adjust the diagonal entries of Q accordingly [40].

Finally, it is important to realize that the RDI, just like the EKF itself, can be applied to systems with multiple inputs and/or outputs as well as systems with nonlinearities beyond delay time itself. However, as systems and their models become more complex, identifiability becomes more of a problem especially if different components of the model are able to realize the same behavioral characteristics. The optimal control model for human operators [6] is a prime example; for instance, the neuromuscular dynamics can to a large degree also be modeled by the cost assigned to changes in applied control. As shown in the following section, delay time can be modeled in different ways thereby complicating its estimation.

IV. Simulation

To gain insight into the RDI's abilities and limitations, a simulation was designed to look at the interaction between delay time and model coefficient estimation in a time varying linear system.

A. Specification

The simulation was performed on a first order ARMA model

$$y_n = \frac{(b_0 + b_1 q^{-1}) q^{-d_n}}{1 - a_1 q^{-1}} r_n + v_n$$

with fixed linear coefficients and a time varying delay time $d_n = d_{n-1} + w_n$. Note that q indicates a unit shift operator ($q^{-1} r_n = r_{n-1}$). The bandlimited time series $\{r_n\}, \{v_n\}$ and $\{w_n\}$ were obtained by applying an ideal lowpass filter (using the Fast Fourier Transform and its inverse) on normally distributed signals. The cutoff frequency and standard deviation for $\{r_n\}$ were respectively set to 1.0 Hz and 2.0. For $\{v_n\}$, they were set to 3.0 Hz and 0.01, and for $\{w_n\}$ to 0.3 Hz and 0.1. In the noise free simulations, discussed below, σ_v was set to zero. The zero's and pole's associated with the ARMA model's cutoff frequencies were respectively set to 0.75 and 1.25 Hz and a unit dc-gain was used. The simulation was run at a sampling frequency of 600 Hz and subsampled at 60 Hz to produce fractional delay times. While decimation of the input and output signals as well as the AR coefficients affect the effective output noise variance [39], it suffices here to know that the resulting σ_v was about 0.3; this includes the fact that identification of the linear coefficients was performed at a 15 Hz sampling rate (decimation factor λ of 4) while input signal magnitude and derivative estimation (i.e., interpolation) was based on the 60 Hz rate. The true values for the linear coefficients were $a_1 = 0.59, b_0 = 1.51$ and $b_1 = -1.10$.

Negative delay time, as indicated in the figures, represents anticipation. One can think of these delay time fluctuations as caused by a mechanism that scans the input signal without knowing whether it feeds the past, present or future to the controller, thus causing the system to effectively respond to a time advanced or delayed representation of the true input signal. The RDI is used to estimate these time-shift fluctuations.

B. Analysis

Since the linear coefficients were fixed, Qc was set to the small value 1.0e-6 I(I is the identity matrix). In case a linear coefficient was assumed to be known, its initial value was set to the correct value and the corresponding diagonal entry in

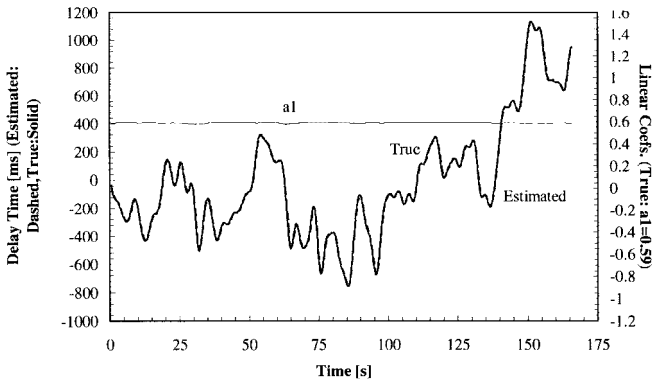

Fig. 1. RDI model coefficient estimations from a noise-free simulation based on the assumption that the MA-coefficients are known.

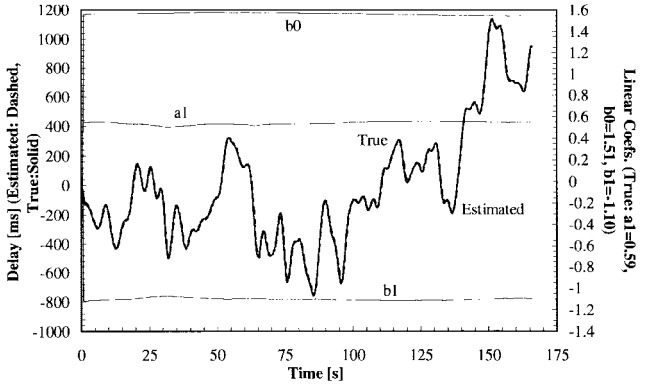

Fig. 2. RDI model coefficient estimations from a noise-free simulation.

Qc was set to zero. Because the delay time standard deviation was set to 0.1 time steps based on a 600 Hz sampling rate plus the fact that the final identification update rate was 15 Hz, Q_d became $40 * 0.1^2 = 0.4$. Finally, R was set to 0.1 in the noisy cases and to $1.0e-4$ in the noise-free cases. Identification was applied under four conditions:

- noise free with known MA-coefficients (Fig. 1);
- noise free with unknown MA-coefficients (Fig. 2);
- noisy with known MA-coefficients (Fig. 3);
- noisy with unknown MA-coefficients (Fig. 4).

To identify a partially known system, diagonal entries in Q_c that correspond to the known coefficients are simply set to zero. This approach does not work for the optimal smoother since $P + Q$, which is inverted, no longer has full rank. This problem can be circumvented by subtracting the known portion of the output from the observed output and removing the known coefficients from the model thus reducing the model order [42]. Since this solution was not applied here, the smoother was not activated in any of these identification runs.

C. Results and Discussion

In general, both noise free conditions resulted in highly accurate delay time as well as linear coefficient estimates (Figs. 1 and 2). Also, identification with noisy observations resulted in meaningful estimates with known MA coefficients (Fig. 3) but the case with unknown MA coefficients produced highly biased linear coefficient estimate (Fig. 4). The delay

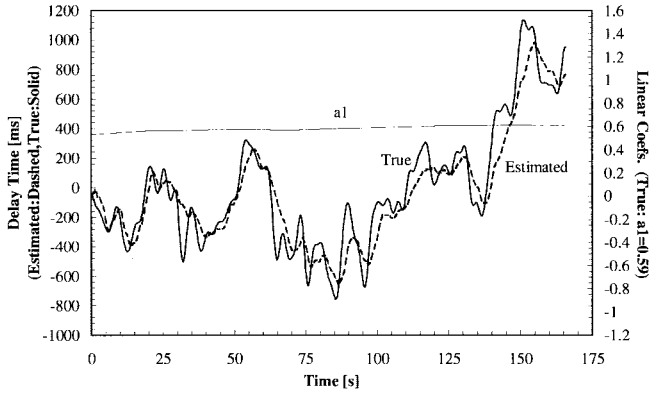

Fig. 3. RDI model coefficient estimations from a simulation with induced observation noise based on the assumption that the MA-coefficients are known.

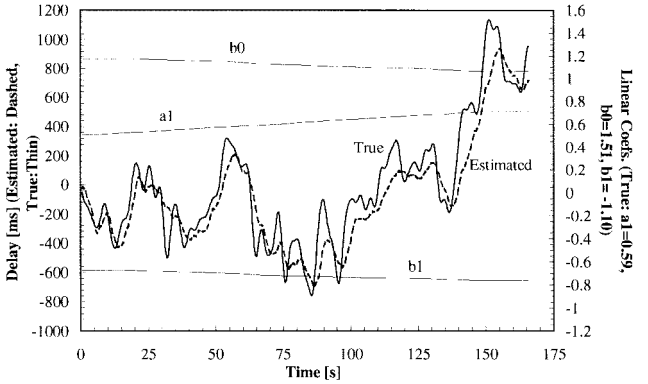

Fig. 4. RDI model coefficient estimations from a simulation with induced observation noise.

time estimates do not appear to be severely biased under any of the conditions. We believe that the reason for these biased MA coefficient estimates lies in the fact that they also approximate fractional delay times. Since it is only an approximation, we would not expect such a bias under noise free conditions because the delay time estimate provides a more accurate solution. This is supported by the fact that Figs. 1 and 2 show no difference in delay time estimates (exact in both cases). Furthermore, a close look at the delay time estimates in Figs. 3 and 4, shows that the dashed line (estimated delay time) in Fig. 4 is always slightly lower than the one in Fig. 3 confirming the assertion that the MA coefficients also model a portion of the delay time. If we would increase the *a priori* noise in the MA coefficients (i.e., Q_c), oscillations in these coefficients would be observed with periods corresponding to the time it takes the true delay time to increase or decrease by one sampling interval. Similar concerns hold for low bandwidth systems in which the AR coefficients can accurately model pure delay (pole slightly beyond the bandwidth). For high-bandwidth systems, Box and Jenkin's approach to first identify the AR coefficients followed by the MA and delay time may offer some advantage [29]. It is important to realize that these issues are general concerns that simply become more important in the RDI because two competing mechanisms are working simultaneously to approximate the pure delay time. At

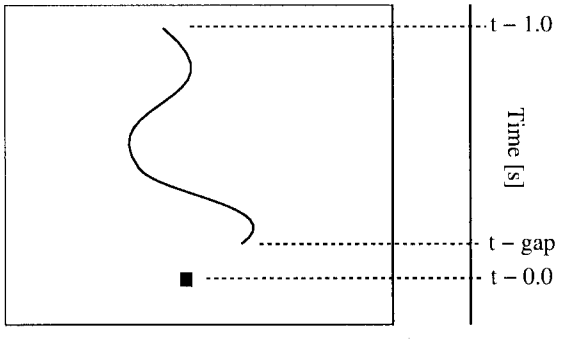

Fig. 5. Illustration of the gap-postview display. The box at the bottom of the screen is controlled by lateral mouse movements. The track is shown from time $t - 1.0$ at the top of the screen to $t - gap$ at the point closest to the box. The box is displayed at current time t. The time axis was not displayed during the experiments.

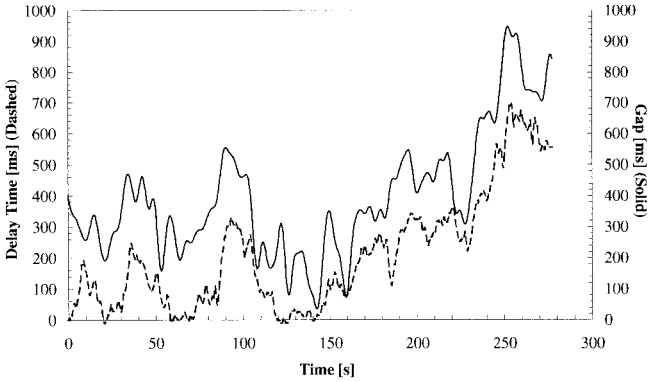

Fig. 6. RDI model coefficient estimations from a typical result of the manual postview tracking task with time varying gap between the control element and the visible part of the track.

higher noise levels, it is more difficult to differentiate between a truly pure delay and a lag introduced by a pole or zero well outside the system's bandwidth. Since the true pole of the simulated system lies outside the input signal bandwidth, this may have caused the slight bias in Figs. 3 and 4.

Aside from these competing mechanisms in delay time representation, the RDI's own time constants, as dictated by the *a priori* covariance matrix settings, play an important role. The delay time estimate (dashed line) always lags the true delay time (solid line) in Figs. 3 and 4 thus in some instances causing substantial errors. This is not observed in Figs. 1 and 2 because of the significantly smaller value of R which makes the identifier much more responsive (see Section III-G. on selecting Q and R). Overall, we see excellent delay time tracking.

V. Time Varying Gap-Postview Tracking Task

By applying the RDI to human tracking data, we take it into a more practical realm with less controlled system adaptations and noise structures. This section and the next are not meant to provide an in-depth analysis of human adaptation; they are meant to demonstrate the RDI's potential in helping to characterize and understand human adaptation.

A. Experimental Design

The goal here was to design an experiment in which the human delay time would change in a predictable fashion thus enabling us to verify the RDI's estimates on a data set with human induced noise which may not necessarily be Gaussian. The experiment consisted of a subject sitting in front of a 19-in. monitor on which a winding track was displayed between the top of the screen (1.0 s postview) and some distance above the bottom of the screen where a user controlled marker (box) was displayed (Fig. 5). This resulted in a blank space between the marker (0.0 s postview) and the point where the track disappears into this gap. The gap was always greater than 0.0 s and never exceeded 1.0 s. The marker moved laterally in unison with the mouse movements (i.e., pure-gain system). The task was to keep the marker on the spot where the track would intersect the 0.0 s postview or current time line thus forcing the subject to extrapolate across the gap. The sampling and screen update rate were 36 Hz. The gap signal was obtained by adding 500 ms to the delay time signal from the previous simulation experiment and multiplying the result by 0.8 to avoid a gap greater than 1.0 s. The resulting gap signal bandwidth is 0.18 Hz. The bandwidth of the input signal came to (replay of the 1.0 Hz input signal from previous simulation, originally designed for 60 Hz, at 36 Hz) 0.6 Hz.

B. Identifier Initialization

The closed loop system was assumed to be well described by a pure-gain pure-delay model. Such an assumption may not hold for compensatory, error regulating control because of stability related to the feed back loop. However, the postview display used here enables accurate perception of the reference track thus making highly accurate open-loop tracking highly plausible especially for the relatively low bandwidth input signal (0.6 Hz) used here.

Here, Q_d was set to 0.5, Q_c to 1.0e-6, and R to 0.5. This reflects a highly variable delay time, a rather constant gain and a substantial amount of observation noise. While the optimal smoother was turned off for the simulations, it was turned on here to improve model coefficient estimates (minimize dynamical lag). Furthermore, to determine the RDI's response to a highly inaccurate initial delay time estimate, it was set to 0 ms even though the initial gap was about 400 ms.

C. Results and Discussion

The great similarity between the estimated delay time and the gap length (Fig. 6) shows that the RDI does not break down when exposed to human induced noise which may not necessarily be Gaussian. If the naive subject (MG) had simply placed the marker at the lateral position of the bottom of the visible portion of the track, the delay time and gap would have been identical. However, if the subject would have been able to perfectly predict the track's intersection point with the current time line, his delay time would have been around zero. Neither case seems to apply. For the 0.6 Hz bandwidth input track, it is reasonable to assume that humans, using mostly linear extrapolation, are capable of predicting its future fairly accurately up to about one-eighth of the period of the highest frequency component in the signal which corresponds to 0.2 s

in this experiment. The data in Fig. 6 indicates a 0.2 s difference between gap and delay suggesting that our subject may indeed have applied simple linear extrapolation across the gap.

Whether the small fluctuations of the estimated delay time are true or a result of the identifier can be determined by varying Q_d and fixing it at the value below which this "high" frequency noise is eliminated. Going too far down will cause the RDI to loose track (often signified by a sudden jump) and produce meaningless estimates. As mentioned earlier, this approach to tune the R and Q matrices was also suggested by Bryson and Ho in [40].

During the first 10 s, the RDI correctly recovers from the highly inaccurate initial delay time estimate. As was discussed earlier, recovery from an initial delay time error greater than half the period associated with the input signal bandwidth may not result in correct convergence.

The gap in this experiment finds many counter parts in the world of tele-operation [43]. It is expected that embedding the RDI in adaptive controllers for systems with time varying delay time, such as distributed virtual reality with variable transmission delays, will result in enhanced man-machine interaction since it provides estimates about the amount of prediction required in presenting the visuals.

VI. Critical Tracking Task

A. Experimental Design

The critical tracking task (CTT) was originally designed by Jex *et al*. in the 1960's and used to estimate a human operator's effective delay time [13]. The goal in this compensatory tracking task was to keep a gradually becoming more unstable first order plant under control. At the point of control loss (i.e., when the tracking error exceeded a set value), they hypothesized that the human operator's effective delay time and the instability level (λ) were each other's reciprocal. True reciprocity would only be reached in the event of zero gain and phase margin control, which is humanly impossible. To maintain control, subjects have to lower their gain and delay time as the instability increases to preserve maximal gain and phase margins.

Here we repeated this experiment. The subject's task was to minimize lateral movements of a cross displayed in the middle of a computer screen via mouse control. The trial was terminated when the cross deviated from the center by more than a screen width. A 36 Hz update rate was maintained. As disturbance input to the system, the 1.0 Hz bandwidth input signal from the gap-postview experiment was used but its magnitude was first reduced by a factor of 100. The instability level increased linearly from zero at a rate of 0.11/s.

The goal here is to illustrate that the RDI can be used to enhance our understanding about human adaptation during these kinds of tracking tasks, especially since to the best of our knowledge no other methods have successfully been applied to human operator data in which the observed system exhibits substantial adaptations over relatively short periods of time. In depth analysis on CTT and other time varying manual tracking tasks will be reported elsewhere.

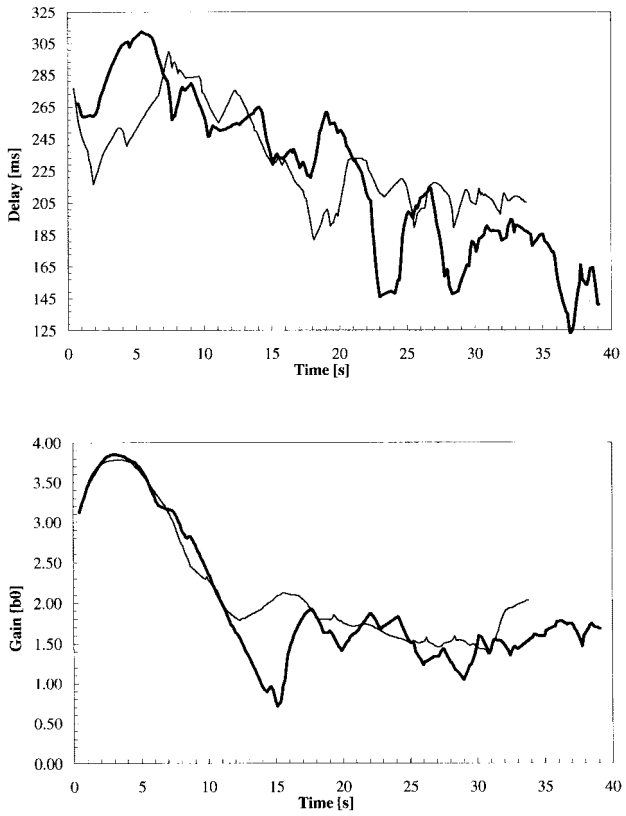

Fig. 7. RDI model coefficient estimations from a time varying manual tracking task modeled after Jex's critical tracking task.

B. Identifier Initialization

No particular rigor has been adopted to optimize the selection of the *a priori* covariance matrices Q and R. The main goal was to see whether the RDI would confirm the hypothesis that subjects reduce their delay time and gain as the instability level increases and therefore not always operate at the maximum of their ability. In accordance with Jex's prediction, the human operator was modeled as a pure-gain pure-delay system. Here, Q_d was set to 0.5, Q_c to 0.05, and R to 0.5. Again, the optimal smoother was left on.

C. Results and Discussion

Data from our naive subject (MG) is shown in Fig. 7. The top panel shows human operator delay time estimates for two different trials while the bottom panel shows the corresponding gain estimates. These data do indeed show the trends predicted by Jex *et al*. The delay time drops "gradually" to about 150 ms as the gain drops to about 1.5. Jex in [13] assumed a 110 ms lower limit for a human's effective delay time and a gain of about 1.5 at the instability levels reached by our subject (3.9 and 4.5). The two depicted trials show very similar adaptation profiles. It is interesting to note that the thin gain-line increases toward the end of this trial which may have caused premature loss of control. Other trials show similar terminal values, only the path at which the gain and delay time reached these values differed suggesting variability in the time course of adaptation. Even the two trials shown here show marked differences. For example, the thin delay-line reaches

a plateau after about 15 s upon which the variability seems to decrease indicating that control was forced into a corner as suggested by Jex *et al*. On the other hand, the thick delay-line shows much more fluctuation in an overall significantly lower delay. Such differences were expected because control is easy at low instability levels thus allowing the human operator to control "well enough" under a wide range of gain and phase margins (i.e., gain and delay time settings). In fact, from the thick delay-line data in Fig. 7, it appears that the tolerance in control was tightened after about 20 s based on the relatively sudden decrease in delay time.

The suggested usefulness of the RDI in tracking a human operator's time varying model coefficients is extremely encouraging in that it may open doors to answer many new questions and hypothesis about human adaptation. Attention no longer has to be focused on adaptation to sudden large changes because of analysis constraints [7], [44]–[46], but can be directed to the effect of more gradual changes.

One important component in many human operator models is remnant. Human controller remnant is defined as that portion of the human output that is not related to the system input by the input/output describing function [47]. One source of remnant is attributed to a human operator's involuntary time-varying behavior. To determine the magnitude of this effect, the RDI can be used to quantify that portion of remnant caused by these unintended adaptations.

VII. Conclusions

A new approach to recursively identify a linear system's time varying delay time and linear coefficients was introduced. This RDI is based on the EKF and uses a bicubic interpolation scheme to provide subsample magnitude and derivative estimates of the observations. The encouraging results on simulated as well as manual tracking data suggest that the imposed differentiability constraint on the observations was of little practical consequence.

Initial results from an experiment based on Jex's critical tracking task confirmed his theory about human gain and delay time adaptation. Furthermore, the fact that the RDI produced highly meaningful delay time estimates in a gap-postview tracking tasks with a highly variable display gap suggests great potential to use the RDI to further enhance our understanding about human adaptation mechanisms.

Appendix
Bilinear Interpolation

The theoretically ideal interpolation scheme is the Whittaker interpolation which convolves a sinc function with the sampled time series to obtain subsample magnitude estimates [48]. One practical approach is to approximate this sinc function over a limited interval with polynomials. An example is the bicubic interpolation scheme in which the sinc function is approximated over the interval $[-2, 2]$ with four third-order polynomials. Because of symmetry, the derivation of the polynomial coefficients only need to be determined for the intervals [0, 1] and [1, 2]. Parker *et al*. [30] showed that the lowest approximation error for signals whose bandwidth is much lower than the Nyquist frequency is obtained with the following two approximating polynomials

$$f(x) = 1.5x^3 - 2.5x^2 + 1, \qquad x \in [0,1]$$
$$g(x) = -0.5x^3 + 2.5x^2 - 4x + 2, \qquad x \in [1,2].$$

Estimation of the signal value at time $n+d$, where $d \in [0,1)$, is based on sample values at time steps $n-1, n, n+1$, and $n+2$. Note that the time $n+d$ is central to these three time intervals. It is easily verified that the respective weights are $g(1+d), f(d), f(1-d)$, and $g(2-d)$. The arguments are the distances between $n+d$ and the respective sample times. The fact that the integral over -2 to 2 is exactly one, means that no signal power is dissipated in the interpolation (i.e., filtering) process. The value of $\hat{h}_{n+d}$ is now found with

$$\hat{h}_{n+d} = g(1+d)x_{n-1} + f(d)x_n + f(1-d)x_{n+1} + g(2-d)x_{n+2}.$$

The derivative of the sampled signal at $n+d$ is calculated with

$$\hat{\dot{h}}_{n+d} = g'(1+d)x_{n-1} + f'(d)x_n + f'(1-d)x_{n+1} + g'(2-d)x_{n+2}$$

where the prime stands for derivative.

Boer compared seven interpolation schemes on their accuracy in magnitude and derivative estimation, including the linear, quadratic and cubic polynomial interpolation methods as well as two others that approximate the sinc function [39]. They were all applied to an analytically known signal (sum of sinusoids) with uniform power and a 1.0 Hz bandwidth. Even though the third order polynomial interpolation showed the best results in the magnitude estimates, the bicubic interpolation was superior in estimating the derivative. A practical review of many approximation and interpolation methods can be found in [48].

Acknowledgment

The authors would like to thank the subjects for their time and diligence in performing these experiments. The authors also thank Prof. A. Nehorai for his helpful comments on the manuscript.

References

[1] T. B. Sheridan and W. R. Ferrell, *Man-Machine Systems: Information, Control, and Decision Models of Human Performance*. Cambridge, MA: MIT Press, 1981.
[2] D. McRuer and H. R. Jex, "A review of quasi-linear pilot models," *IEEE Trans. Human Factors Electron.*, vol. HFE-8, pp. 231–249, Sept. 1967.
[3] D. L. Kleinman, S. Baron, and W. H. Levison, "An optimal control model of human response, part 1: Theory and validation," *Automatica*, vol. 6, pp. 357–369, 1970.
[4] S. Baron, D. L. Kleinman, and W. H. Levison, "An optimal control model of human response, part II: Prediction of human performance in a complex task," *Automatica*, vol. 6, pp. 371–383, 1970.
[5] M. Tomizuka and D. E. Whitney, "Optimal discrete finite preview problems (why and how is future information important?)," *J. Dyn. Syst., Meas., Contr.*, vol. 97, pp. 319–325, Dec. 1975.
[6] M. Tomizuka and D. E. Whitney, "The human operator in manual preview tracking (an experiment and its modeling via optimal control)," *J. Dyn. Syst., Meas., Contr.*, vol. 98, pp. 407–413, Dec. 1976.
[7] L. R. Young, "On adaptive manual control," *IEEE Trans. Man–Mach. Syst.*, vol. MMS-10, pp. 292–331, Dec. 1969.
[8] C. R. Kelley, *Manual and Automatic Control, A Theory of Manual Control and its Application to Manual and Automatic Systems*. New York: Wiley, 1968.

[9] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng., Trans. ASME*, vol. 82D, pp. 35–45, Mar. 1960.
[10] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Prentice-Hall Information and Systems Sciences Series, T. Kailath, Ed. Englewood Cliffs, NJ: Prentice-Hall, 1979.
[11] A. Gelb, *Applied Optimal Estimation*. Cambridge, MA: MIT Press, 1989.
[12] M. E. Salgado, C. E. de Souza, and G. C. Goodwin, "Issues in time delay modeling," in *IFAC Identification System Parameter Estimation*, Beijing, China, 1988, pp. 549–554.
[13] H. R. Jex, J. D. McDonnell, and A. V. Phatak, "A "critical" tracking task for manual control research," *IEEE Trans. Human Factors Electron.*, vol. HFE-7, Dec. 1966.
[14] R. Iserman, "Practical aspects of process identification," *Automatica*, vol. 16, pp. 575–587, 1980.
[15] A. Abdel-Malek and V. Z. Marmarelis, "Modeling of task-dependent characteristics of human operator dynamics pursuit manual tracking," *IEEE Trans. Syst., Man, Cybern.*, vol. 18, pp. 163–172, Jan./Feb. 1988.
[16] F. Osafo-Charles, G. C. Agarwal, W. D. O'Neill, and G. L. Gottlieb, "Application of time-series modeling to human operator dynamics," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-10, pp. 849–860, Dec. 1980.
[17] L. Ljung, *System Identification, Theory for the User*, T. Kailath Ed. Englewood Cliffs, NJ: Prentice-Hall, 1987.
[18] C. Banyasz and L. Keviczky, "A new recursive time delay estimation method for armax models," in *IFAC Identification System Parameter Estimation*, Beijing, China, 1988, pp. 1055–1060.
[19] R. Pupeikis, "Recursive estimation of the parameters of linear systems with delay time," in *Proc. 7th IFAC Symp. Identification System Parameter Estimation*, Oxford, U.K.: Pergamon, 1985, pp. 787–792.
[20] W. X. Zheng and C. B. Feng, "Optimizing search-based identification of stochastic time-delay systems," *Int. J. Syst. Sci.*, vol. 22, no. 5, pp. 783–792, 1991.
[21] J. Tuch, A. Feuer, and Z. J. Palmor, "Time delay estimation in continuous linear time-invariant systems," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 823–827, Apr. 1994.
[22] R. M. De Keyser, "Adaptive dead-time estimation," in *IFAC Adaptive Systems Control Signal Processing*, Lund, Sweden, 1986, pp. 385–389.
[23] X. Hongyao, "Robust protection-time delay tracing adaptive control and its application," in *IFAC Identification System Parameter Estimation*, Beijing, China, 1988, pp. 769–774.
[24] E. R. Boer, "Investigative research after subjectively and objectively measured parameters during manual control tracking tasks with disorienting visual stimuli," M.S. thesis, Univ. Twente, Enschede, The Netherlands, 1990.
[25] H. Meyr and G. Spies, "The structure and performance of estimators for real-time estimation of randomly varying time delay," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, pp. 81–94, Feb. 1984.
[26] C. L. Nikias and R. Pan, "Time delay estimation in unknown Gaussian spatially correlated noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 1706–1714, Nov. 1988.
[27] P. C. Ching and Y. T. Chan, "Adaptive time delay estimation with constraints," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 599–602, Apr. 1988.
[28] W. Zheng and C. B. Feng, "Unbiased identification of linear time-delay systems corrupted by correlated noise," in *IFAC Identification System Parameter Estimation*, Beijing, China, 1988, pp. 477–482.
[29] G. E. P. Box and G. M. Jenkins, *Time Series Analysis, Forecasting and Control*. San Francisco, CA: Holden-Day, 1969.
[30] A. Parker, R. V. Kenyon, and D. E. Troxel, "Comparison of interpolating methods for image resampling," *IEEE Trans. Med. Imag.*, vol. MI-2, pp. 31–39, Mar. 1983.
[31] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C, The Art of Scientific Computing*. Cambridge, U.K.: Cambridge Univ. Press, 1990.
[32] H. E. Rauch, "Solutions to the linear smoothing problem," *IEEE Trans. Automat. Contr.*, vol. AC-8, pp. 371–372, Oct. 1963.
[33] H. E. Rauch, F. Tung, and C. T. Striebel, "Maximum likelihood estimates of linear dynamic systems, *AIAA J.*, vol. 3, pp. 1445–1450, Aug. 1965.
[34] J. P. Norton, "Recursive identification and prediction of time-varying systems using extra information," in *Proc. 4th IFAC Symp.*, Tbilisi, Russia, Sept. 1976, and *Identification and System Parameter Estimation*, N. S. Rajbman, Ed. Amsterdam, The Netherlands: North-Holland, 1978, pp. 1881–1890.
[35] D. Q. Mayne, "A solution of the smoothing problem for linear dynamic systems," *Automatica*, vol. 4, pp. 73–92, 1966.
[36] B. D. O. Anderson, "Properties of optimal linear smoothing," *IEEE Trans. Automat. Contr.*, vol. AC-14, pp. 114–115, Feb. 1969.
[37] J. B. Moore, "Discrete-time fixed-lag smoothing algorithms," *Automatica*, vol. 9, pp. 163–173, 1973.
[38] J. P. Norton, "Optimal smoothing in the identification of linear time-varying systems, *Proc. Inst. Elect. Eng.*, vol. 122, June 1975.
[39] E. R. Boer, "Identification of time varying systems with applications to adaptive human control," Ph.D. dissertation, Univ. Illinois, Chicago, 1995.
[40] A. E. Bryson and Yu-Chi Ho, *Applied Optimal Control, Optimization, Estimation, and Control*. New York: Wiley, 1975.
[41] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill, 1965.
[42] C. A. Canundas de Wit, *Adaptive Control for Partially Known Systems, Theory and Applications, Studies in Automation and Control*. Amsterdam, The Netherlands: Elsevier, 1988, vol. 7.
[43] T. B. Sheridan, *Telerobotics, Automation, and Human Supervisory Control*. Cambridge, MA: MIT Press, 1992.
[44] J. D. McDonnel, "A preliminary study of human operator behavior following a step change in the controlled element," *IEEE Trans. Human Factors Electron.*, vol. HFE-7, pp. 125–128, Sept. 1966.
[45] R. A. Miller and J. I. Elkind, "The adaptive response of the human controller to sudden changes in controlled process dynamics," *IEEE Trans. Human Factors Electron.*, vol. HFE-8, pp. 218–223, Sept. 1967.
[46] A. V. Phatak and G. A. Bekey, "Model of the adaptive behavior of the human operator in response to a sudden change in the control situation," *IEEE Trans. Man–Mach. Syst.*, vol. MMS-10, pp. 72–80, Sept. 1969.
[47] W. H. Levison, S. Baron, and D. L. Kleinman, "A model for human controller remnant," *IEEE Trans. Man-Mach. Syst.*, vol. MMS-10, no. 4, pp. 101–108, 1969.
[48] A. C. R. Newbery, "Numerical analysis," in *Handbook of Applied Mathematics, Selected Results and Methods*, 2nd ed., C. E. Pearson, Ed. New York: Van Nostrand Reinhold, 1983, ch. 18.

Erwin R. Boer (S'92–M'92) was born in Enschede, The Netherlands, in 1965. He received the M.S. degree in electrical engineering from Twente University, The Netherlands, in 1990, and the Ph.D. degree in electrical engineering from the University of Illinois, Chicago, in 1995.

Since 1990, he has been affiliated with the Center for Clouds Chemistry and Climate, Scripps Institution of Oceanography, conducting research in the areas of image processing and data visualization. Currently, he is a Research Scientist at Nissan Cambridge Basic Research, Cambridge, MA. His main research interests are in the areas of system identification, adaptive visuomotor control, and image processing.

Robert V. Kenyon (S'71–M'78) received the B.S. degree in electrical engineering from the University of Rhode Island, Kingston, in 1970, the M.S. degree in bioengineering from the University of Illinois, Chicago, in 1972, and the Ph.D. degree in physiological optics from the University of California, Berkeley, in 1978.

From 1979 to 1986, he was a Faculty Member with the Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge. He is currently an Associate Professor of Electrical Engineering and Computer Science at the University of Illinois, Chicago. His research interests include sensory-motor adaptation, visuomotor control, flight simulation, virtual environments, computer graphics, and effects of microgravity on vestibular development. He also participated in several shuttle experiments that studied the effects of microgravity on human/animal orientation: Spacelab-1, German Space-lab (D-1), and STS-29 (Chix in Space). A majority of his research has examined the contribution of visual information in the control of motor coordination. Some work examined the influence of visual information on eye movement control and how disease affects information processing. Other work has examined how human manual control characteristics change as the operator is deprived of various visual cues such as peripheral vision. Recent work has focused on motor control and training in virtual environments.