Change-Point Detection in Kinetic Signals

Gerhard Staude and Werner Wolf

Institut fuer Mathematik und Datenverarbeitung
Universitaet der Bundeswehr Muenchen, D-85577 Neubiberg, Germany
gerhard.staude@unibw-muenchen.de

Abstract. A method to precisely determine the onset of voluntary discrete movements in kinetic signals (e.g. joint angle) is presented. The movement onset is identified as an abrupt change in the (time varying) parameters of a statistical process model. An adaptive Kalman whitening filter transforms the digitized kinetic signal into a sequence of innovations which is examined for possible change-points by a generalized log-likelihood-ratio test. The accuracy of the algorithm is assessed by statistical simulations and compared to the accuracy of a standard threshold criterion. Results show that the method provides accurate change time estimates even for weak and highly variable response profiles.

1 Introduction

An important aspect of studying motor control mechanisms in humans is the analysis of reaction time (RT). RT (i.e., the time interval between stimulus presentation and the instant of movement initiation) has usually been measured with mechanical switches or photoelectric devices that indicate the onset of a voluntary motor response. Alternatively, the movement onset may be derived from the digitized kinetic signal by some automatic procedure (e.g., a threshold operation). However, when using simple amplitude threshold methods, particularly weak and abnormal response profiles which are typical for a variety of central motor disorders may introduce high RT variability as well as systematic errors [1]. This paper describes a new method for computerized onset detection which is based on statistical signal processing and includes a priori knowledge on the generator process of kinetic signals.

2 The Process Model

The method is based upon the process model in Fig. 1 which was derived from the nonlinear single-joint model established in [2]. The digitized kinetic signal is approximated by a discrete random process $(Y_k)_k$ which is generated by a Gaussian white noise process $(X_k)_k$ driving a linear filter $H(z)$. The former reflects the discharge timing and recruitment of independent signal sources involved (twitch forces contributed by single motor units), the latter describes the shape of the twitch forces as well as the mechanical properties of the joint. A distortion component $(N_k)_k$ considers measurement errors and "noise" contributed by other biological signal generators. $H(z)$ is modeled by an all-pole representation (autoregressive (AR) model) of order p

$$H(z) = \frac{1}{1 + a_1 z^{-1} + a_2 z^{-2} + \ldots + a_p z^{-p}} . \tag{1}$$

R.W. Brause and E. Hanisch (Eds.): ISMDA 2000, LNCS 1933, pp. 43−48, 2000.
© Springer-Verlag Berlin Heidelberg 2000

$a_1, a_2, \ldots a_p$ are the AR parameters and z is the complex frequency of the z-transform. The time-domain representation of the signal generator is

$$X_k = \mu(k) + W_k, \qquad \tilde{Y}_k = -\sum_{i=1}^{p} a_i \tilde{Y}_{k-i} + X_k, \qquad Y_k = \tilde{Y}_k + N_k \tag{2}$$

where $\tilde{Y}_k$ denotes the undisturbed output signal of the filter. Sequences $(W_k)_k$ and $(N_k)_k$ are uncorrelated zero mean Gaussian white noise signals with constant variances q and r, respectively. As an essential property of the model, the mean $\mu(k)$ of the white noise excitation $(X_k)_k$ is time variant. This yields the motor system the control over the biological signal generator in order to produce a particular movement. Within this framework, the response onset t_0 is defined as an abrupt change between two time varying profiles $\mu_0(k)$ and $\mu_1(k, t_0)$, as illustrated in Fig. 1.

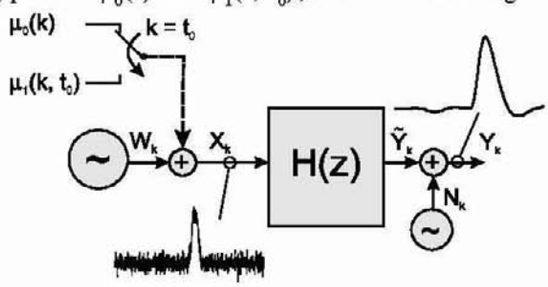

Fig.1 Process model for the kinetic signal of a responding subject.

3 The Algorithm

The properties of the sequence $(Y_k)_k$ are fully described by the (time varying) vector

$$M = \left[a_1 \;\; a_2 \;\; \cdots \;\; a_p \;\; q \;\; r \;\; \mu(k)\right]^T \tag{3}$$

comprising the parameters of the process model in Fig. 1. Thus, statistically optimal onset detection can be achieved with the aid of a binary hypothesis test between the two statistical models M_0 and M_1 which describe the signal properties before and after the change, respectively. The method consists of two stages, (i) an adaptive whitening filter $H_w(z)$, and (ii) a statistical decision element, as illustrated in Fig. 2. The whitening filter $H_w(z)$ adapted to the M_0 model transforms the measured signal $(Y_k)_k$ into an uncorrelated („white") sequence $(\varepsilon_k)_k$ of innovations. The new series $(\varepsilon_k)_k$ sensitively reflects deviations of $(Y_k)_k$ from the M_0 model. It is examined by a decision rule based upon the log-likelihood ratio test, which signals a possible change in model parameters (alarm time t_a) and, in addition, computes an estimate $\hat{t}_0$ of the unknown change time t_0.

3.1 The Adaptive Whitening Filter

The properties of $(Y_k)_k$ are determined to a large extend by the transfer function $H(z)$. According to the model, however, the change to be detected predominantly affects the excitation $(X_k)_k$, whereas $H(z)$ remains comparably constant.

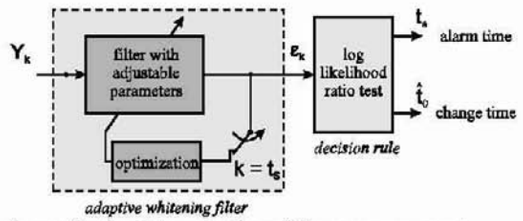

Fig. 2 Detection scheme for the determination of the response onset

This implies that all the information for event detection available in $(Y_k)_k$ is comprised by the mean profile $\mu(k)$ of the excitation, but the information contained in the AR parameters a_i is not relevant for correct detection of changes. In order to remove this irrelevant component from the signal before it enters the decision stage, an adaptive whitening filter is used. The filter with transfer function

$$H_w(z) = \frac{1 + c_1 z^{-1} + c_2 z^{-2} + \ldots + c_p z^{-p}}{1 + b_1 z^{-1} + b_2 z^{-2} + \ldots + b_p z^{-p}} \tag{4}$$

transforms the measured sequence $(Y_k)_k$ into a new series $(\varepsilon_k)_k$ called „the innovations". The filter coefficients are related to the M_0 model according to

$$c_i = a_i \,, \qquad b_i = \begin{cases} c_i(1 - l_1) + l_{i-1} & when\ 1 \le i < p \\ c_i(1 - l_1) & when\ i = p \end{cases} \tag{5}$$

where $L = [l_1\, l_2 \cdots l_p]^T$ is the steady-state Kalman gain vector which depends upon the variances q and r of the process model [3].
Under noise-free conditions ($r = 0$), (4) reduces to a pure moving average (MA) filter

$$H_w(z) = 1 + c_1 z^{-1} + c_2 z^{-2} + \ldots + c_p z^{-p} \tag{6}$$

Thus, for $c_i = a_i$ and $r = 0$, the whitening filter represents an ideal inverse filter $H_w(z) = H^{-1}(z)$ with respect to the transfer function $H(z)$. The excitation $(X_k)_k$ can completely be reconstructed and the detection task reduces to the detection of a change in the dynamic mean $\mu(k)$ of a series of statistically independent Gaussian random variables with variance q. For $r > 0$, the filter still produces an independent Gaussian sequence, but with different mean profile $\mu^*(k)$. Since $\mu^*(k)$ can be obtained by filtering $\mu(k)$ with a filter

$$H_\mu(z) = \frac{1}{1 + b_1 z^{-1} + b_2 z^{-2} + \ldots + b_p z^{-p}} \,, \tag{7}$$

detection for $r > 0$ can be performed as in the ideal noise-free situation by using the modified (filtered) profile $\mu^*(k)$.

Usually, the parameters of the M_0 model are unknown. In this case, the coefficients b_i and c_i of the whitening filter are individually determined from the pre-stimulus interval $k < t_s$ which, by definition, only contains data produced by the M_0 model. As illustrated in Fig. 2, before stimulus presentation ($k < t_s$), model parameters are subjected to a standard least squares optimization technique [3]. At $k = t_s$, optimization

is stopped and a filter with constant parameters is used to compute the innovations $(\varepsilon_k)_k$ for the remaining part of the record.

3.2 The Decision Element

From a statistical point of view, event detection represents a binary testing problem between the null hypothesis $\boldsymbol{H_0}$ saying "there is no change in the statistical properties of the sequence $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_k$ actually observed and the mean profile is $\mu_0(i)$ for all times $1 \le i \le k$", and the alternate hypothesis $\boldsymbol{H_1}$ "there is a change in statistical properties at some unknown change time $1 \le j \le k$, and the mean profiles are $\mu_0(i)$ before and $\mu_1(i, j)$ after this change". The adequate tool for binary hypothesis testing is the log-likelihood ratio test which compares the logarithm of the ratio between the two joint probability density functions for either hypothesis with a threshold [4]. The implementation of the test depends on the knowledge about $\mu_0(i)$ and $\mu_1(i,j)$.
If the model parameters (except t_0) are exactly known, the test can be efficiently implemented as a CUSUM type decision rule by comparing the cumulative sum

$$S_j^k = \sum_{i=j}^{k} s_i(j) \,, \qquad s_i(j) = ln \frac{p_{\mu_1}(\varepsilon_i, j)}{p_{\mu_0}(\varepsilon_i)} \tag{8}$$

of the log-likelihood ratios $s_i(j)$ of the single samples with a threshold h [4]. $p_{\mu_0}(\varepsilon_i)$ and $p_{\mu_1}(\varepsilon_i, j)$ are the probability density functions of the i-th random variable before and after a possible change at j, respectively. Assuming a Gaussian density distribution, the individual log-likelihood ratio is explicitly given by

$$s_i(j) = \frac{\mu_1(i,j) - \mu_0(i)}{\sigma^2} \left[\varepsilon_i - \frac{1}{2} \left(\mu_1(i,j) + \mu_0(i) \right) \right] \tag{9}$$

where σ^2 is the (constant) variance of the process, and $\mu_0(i)$ and $\mu_1(i,j)$ denote the mean profiles before and after a change at j, respectively. Since the exact change time in (8) is unknown, it is replaced by its maximum likelihood (ML) estimate, i.e., a maximum operator selects the largest value of the test function with respect to all possible change times $1 \le j \le k$. The resulting CUSUM detection rule is

$$\begin{aligned} t_a &= min \left\{ k \ge 1 : \max_{1 \le j \le k} S_j^k \ge h \right\} , \quad \hat{t}_0 = arg \max_{1 \le j \le t_a} S_j^{t_a} \\ S_j^k &= \sum_{i=j}^{k} \left[\frac{\mu_1(i,j) - \mu_0(i)}{\sigma^2} \left[\varepsilon_i - \frac{1}{2} \left(\mu_1(i,j) + \mu_0(i) \right) \right] \right] \end{aligned} \tag{10}$$

The data sequence is analysed by a growing observation window which comprises all data points available at time k. A test window with the upper bound k fixed at the current observation and with its lower bound j proceeding reversely in time is applied to the data. For each possible change time j, S_j^k is computed from the observations $\varepsilon_j, \varepsilon_{j+1}, \cdots, \varepsilon_k$ within the test window. If the maximum of S_j^k with respect to all hypothetical change times $1 \le j \le k$ exceeds the threshold h, an event alarm is given (alarm time t_a). The time j at which the maximum value is obtained serves as the maximum likelihood estimate $\hat{t}_0$ of the unknown change time t_0.

In case the exact profiles $\mu_0(i)$ and $\mu_1(i,j)$ are unknown, they must be replaced by appropriate estimates. Generally, the mean profiles will depend upon several parameters $\boldsymbol{\theta_0}$ and $\boldsymbol{\theta_1}$ associated with the actual movements before and after the change, respectively. Estimates $\hat{\boldsymbol{\theta}}_0$ of the unknown parameters $\boldsymbol{\theta_0}$ before change can be obtained off-line from the pre-stimulus interval which, by definition, contains signal samples corresponding to the null hypothesis only. During detection, these parameters then are considered to be known. Estimates $\hat{\boldsymbol{\theta}}_1$ of the unknown parameters $\boldsymbol{\theta_1}$ after change, however, must be computed individually for each possible change time $1 \le j \le k$, at each time k a new data point is available. In order to avoid the multiple scheme of growing test windows of such a generalized likelihood ratio (GLR) procedure, an approximated GLR decision rule (AGLR) was used, which splits the detection procedure into separate phases of detection and change time estimation. During detection, a sliding window of fixed size L is continuously shifted along the data sequence. For each location of the window, the ML estimate $\hat{\boldsymbol{\theta}}_1$ of the unknown parameter vector $\boldsymbol{\theta_1}$ is determined from the L data points covered by the window, and the corresponding $\hat{S}_{k-L+1}^{k}$ is computed. After a change has been indicated, the exact change time is estimated off-line by a ML procedure from all possible candidates $1 \le j \le t_a$. The AGLR decision rule can be summarized as

$$t_a = \min\left\{k \ge d:\ \hat{S}_{k-L+1}^{k} \ge h\right\}, \qquad \hat{t}_0 = \arg\max_{1 \le j \le t_a} \hat{S}_j^{t_a+d^*}$$

$$\hat{S}_j^k = \sup_{\boldsymbol{\theta_1}} \sum_{i=j}^{k} \frac{\mu_1(i,j,\boldsymbol{\theta_1}) - \mu_0(i)}{\sigma^2}\left[\varepsilon_i - \frac{1}{2}\left(\mu_1(i,j,\boldsymbol{\theta_1}) + \mu_0(i)\right)\right] \tag{11}$$

The quantities d and d^* are appropriate dead zones confining parameter estimation to a minimum number of observations.

In this study, the shape $u(i,j)$ of the change in the innovations was assumed to be known but with its exact magnitude unknown, i.e., $\mu_1(i,j,\boldsymbol{\theta_1}) = \theta_1 u(i,j)$. In this case, maximization with respect to θ_1 is explicitly possible, and the log-likelihood ratio in (11) can be written as

$$\hat{S}_j^k = \frac{1}{\sigma^2}\left[\frac{1}{2}\left(\sum_{i=j}^{k} u(i,j)\varepsilon_i\right)^2\left(\sum_{i=j}^{k} u^2(i,j)\right)^{-1} - \sum_{i=j}^{k}\mu_0(i)\left(\varepsilon_i - \frac{1}{2}\mu_0(i)\right)\right] \tag{12}$$

which facilitates efficient implementation of the test.

4 Simulation Results

The method was tested on 4000 simulated kinetic signals each comprising 1000 data points with known change time $t_0 = 500$. Signals were produced by using pseudo-noise sequences driving an AR system $H(z)$ of order $p = 4$ with narrow band transfer characteristic. The movements were initiated by ramp-like impulses in the mean pattern $\mu(k)$ with randomly varying magnitude and slope. The AGLR method was implemented using a ramp template with unit slope ($u(i,j) = i - j$) and a test window size of $L = 25$ samples. Fig. 3 depicts histograms of the estimation error $\Delta = \hat{t}_0 - t_0$ for

several onset detectors. The AGLR method was compared to a simple (adaptive) threshold (ST) criterion. In addition, results of the CUSUM method exactly tuned with the true profiles $\mu_0(i)$ and $\mu_1(i,j)$ are shown as a reference for optimal performance. The ST method employed an adaptive threshold: event alarm occurred when the current sample Y_k exceeded three times the standard deviation of $(Y_k)_k$ estimated from the pre-stimulus period $k < t_s$. The broad shape of the error distribution of the ST method indicates highly variable onset estimates for this standard detection criterion. By contrast, the small error variance of only few samples indicates precise onset estimates of the AGLR method being nearly as accurate as those obtained with the (optimal) CUSUM detector.

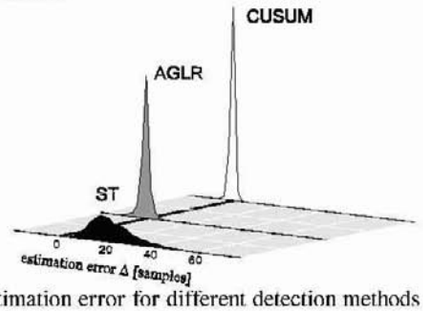

Fig. 3 Histograms of estimation error for different detection methods

5 Discussion

Inclusion of a priori knowledge on the biomechanical generator process of kinetic signals significantly enhances the accuracy of detected movement onsets. When the parameters of the process model are exactly known, the CUSUM detector can serve as a reference criterion, e.g., in order to assess the performance of more „realistic" methods by statistical simulations. When the exact model parameters are unknown, the AGLR method still provides high detection power by computing appropriate estimates of the unknown parameters.

References

1. Staude, G., Wolf, W., Appel, U., Dengler, R.; Methods for onset detection of voluntary motor responses in tremor patients. IEEE Trans. Biomed. Eng., Vol. 43. (1996) 177-88
2. Staude, G., Dengler, R., Wolf, W.: The discontinuous nature of motor execution I. A model concept for single-muscle multiple-task coordination. Biol. Cybern., Vol. 82. (2000) 23-33
3. Goodwin, G.C., Sin, K.S.: Adaptive filtering, prediction and control. Prentice-Hall, Englewood Cliffs, New Jersey. (1984)
4. Basseville, M., Nikiforov, I.V.: Detection of abrupt changes: Theory and application. Prentice-Hall, Englewood Cliffs, New Jersey. (1993) 29-43