pp. 367-370

OF THE
PROCEEDINGS OF

FIRST INTERNATIONAL SYMPOSIUM ON COMMUNCIATION SYSTEMS & DIGITAL SIGNAL PROCESSING

6-8 APRIL 1998

SHEFFIELD HALLAM UNIVERSITY, UK

ISBN 0-86339-7719

PUBLISHED IN UNITED KINGDOM BY
SHEFFILED HALLAM UNIVERSITY PRESS,
LEARNING CENTRE

CONSIDERATIONS ON ADAPTIVE AUTOREGRESSIVE MODELLING IN EEG ANALYSIS

Alois Schloegl and Gert Pfurtscheller
[1]Ludwig Boltzmann Institute for Medical Informatics and Neuroinformatics,
[2]Department of Medical Informatics, Institute of Biomedical Engineering,
Graz University of Technology, Austria.
Brockmanngasse 41, A-8010 Graz, Austria
a.schloegl@ieee.org

INTRODUCTION

For a Brain Computer Interface it is important that the electroencephalogram (EEG) can be analysed on-line and on a single trial basis. The EEG characteristic changes with respect to different mental activity are exploited. In other words, the event-related desynchronization (ERD) [1], which is an attenuation of frequency components, has to be evaluated.

For the evaluation of the spectra an autoregressive model can be used. This offers the advantage that the spectrum can be estimated based on a Maximum entropy (MESE) [2]. The time-variation is taken into account using an adaptive model.

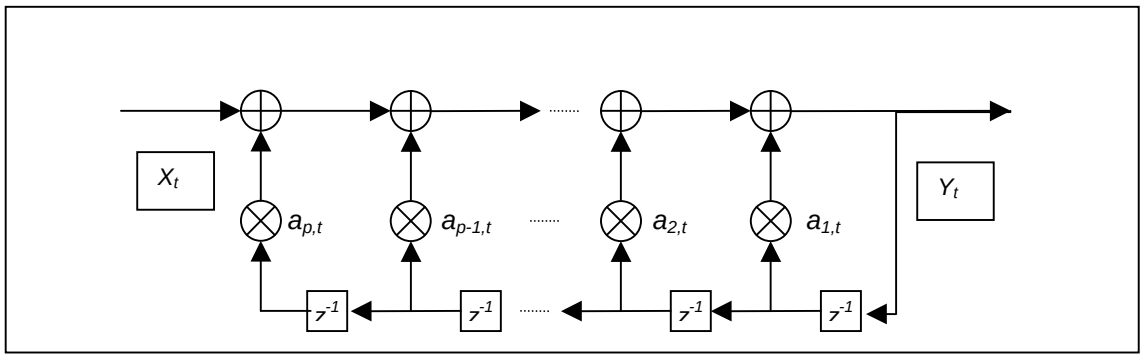

Figure 1: An Adaptive Autoregressive model is shown. One can imagine that the model output Y(t) is a white noise process which is filtered. The characteristics of the signal are described by the AR parameters. For the evaluation of time-varying behaviour an adaptive AR model is used; the model parameters are varying with time.

$$Y_t = a_{1,t}Y_{t-1} + a_{2,t}Y_{t-2} + ... + a_{p,t}Y_{t-p} + X_t = \Sigma\, a_{i,t} * Y_{t-i} + X_t \qquad (1)$$

An adaptive autoregressive model is described by equation (1). Note, that the model coefficients are also time-varying. In the ideal case, X_t is a purely random (white noise) process with zero mean and variance σ_x^2. X_t is uncorrelated with the signal, and the cross-covariance function $E\{Y_t * X_{t+k}\}$ is zero for every k. The difference to an AR model is that the parameters $a_{1,t}$... $a_{p,t}$ can vary with time, however it is assumed that the parameters change only ”slowly”.

METHOD

For the model identification, the AAR parameters have to be identified. Several methods, e.g. the least-mean-squares (LMS) and the recursive-least-squares (RLS) method, can be used for the estimation of the AAR parameters [2,3]. In the following sections, the AR parameters and the past p samples of the time series are defined as vectors:

$$\hat{a}_t = [\hat{a}_{1,t} \ldots \hat{a}_{p,t}]^T \qquad (2)$$
$$\mathbf{Y}_t = [Y_t \ldots Y_{t-p+1}]^T \qquad (3)$$

where p is the order of the autoregressive model, ^ denotes the estimates of parameter and ^{T} denotes the transpose of the vector; bold letters indicate vectors.

The RLS algorithm and Kalman filtering are still topics of discussion; seen in various publications on that subject; only one example is [4]. Here a version discussed in [5, 6] was used.

$$E_t = Y_t - \hat{\mathbf{a}}_{t-1}^T \mathbf{Y}_{t-1} \quad (4)$$
$$\mathbf{r}_t = (\mathbf{A}_{t-1} + \mathbf{W}_{t-1})\, \mathbf{Y}_{t-1} \quad (5)$$
$$\mathbf{k}_t = \mathbf{r}_t / (\mathbf{Y}_{t-1}^T \mathbf{r}_t + 1) \quad (6)$$
$$\hat{\mathbf{a}}_t = \hat{\mathbf{a}}_{t-1} + \mathbf{k}_t E_t \quad (7)$$
$$\mathbf{A}_t = (\mathbf{A}_{t-1} + \mathbf{W}_{t-1}) - \mathbf{k}_t \mathbf{r}_t^T \quad (8)$$
$$\mathbf{W}_t = UC * \mathbf{A}_t \quad (9)$$

E_t is the (one-step) prediction error and $\mathbf{k}_t$ the Kalman Gain Vector; $\mathbf{r}_t$ was introduced to make the algorithm computationally more efficient; $\mathbf{A}_t$ is the covariance matrix and $\mathbf{W}_t$ considers that $\mathbf{A}_t$ does not become too small and adaptation is possible. The initial values $\hat{\mathbf{a}}_0$ were some average AR parameters from the EEG and $\mathbf{A}_0 = 0.1 * \mathbf{I}_{pxp}$.

Furthermore, we define a ratio between the mean square error and the total power of the signal

$$REV = MSE / MSY \quad (12)$$

whereby MSE is the mean-square error (or variance of the error process)

$$MSE = N^{-1} \sum_{t=1}^{N} E_t^2 \quad (13)$$

and MSY is the variance or the total power of the EEG signal

$$MSY = N^{-1} \sum_{t=1}^{N} Y_t^2 \quad (14)$$

Assuming that the error process E_t is the white noise or innovation process X_t, and considering that X_t is not correlated with the signal, one can say that the relative error variance REV is the relative power of the innovation process. Furthermore, the power of the one-step prediction process $\Sigma (a_{i,t} * Y_{t-i})$ is MSY-MSE.

REV=1 means that the model parameters are zero and the signal is white noise; REV = 0 means that the signal can be explained completely by the model (only a theoretical consideration, not possible in practice). If 0 < REV < 1, REV tells us how much of the signal is contributed by the white noise process and how much from the one-step prediction process. Consequently, REV expresses how much of the signal is not explained by the model (parameters).

RESULTS

The data of one session of a BCI experiment were used. Trials with 78 left and 78 right imagined hand movements were investigated. The EEG signal was recorded from the left hemisphere of the sensorimotor area. The sampling rate was 128Hz; each trial was of 8s duration, after a cue from second 3-4.25 the subject performed a imaginary hand movement. According to the brain activity the spectral density function changes. For more details see [7].

The frequency components below 2Hz were removed with a symmetric Blackman-128 window. Model orders of 6 and 10 were used; the update coefficient UC was varied from 10^(-3 + k/8) for k = 0..12; REV was calculated from the average variance MSE and MSY over all trials.

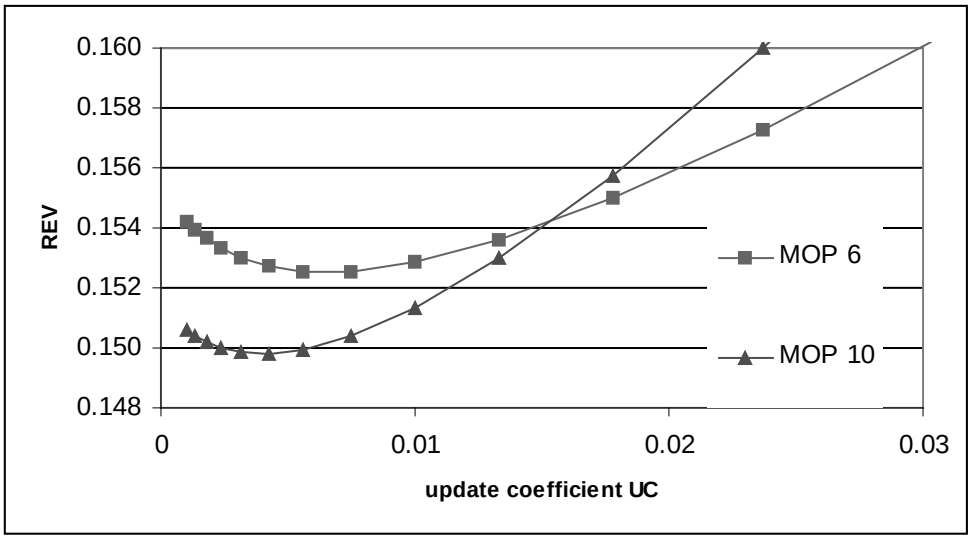

Figure 2: The recording from BCI4C F5-6, bipolar channel C3 was analysed. The REV is smaller with the model order 10. However, the minimum REV is at 0.0056 and 0.0042 for orders 6 and 10 respectively.

Firstly, one can see that for some UC the REV is a minimum. Which means for some UC the AAR model describes the signal best. In case UC is small the AAR estimates do not follow the variations of the signal; In case the signal changes, the parameters describe the actual behaviour of the signal not well.

In case UC is large, the parameters can follow the changes of the signal, but now the estimation accuracy of the parameters is worse because only a few samples are considered for the actual estimation. This means the AAR parameters have a high variance of the estimation error. As a result, the one-step-prediction is worse and the prediction error E_t is increased.

For some UC an optimal REV exists, which is the best tradeoff between adaptation speed and estimation accuracy. We can see that for p=6 the minimum REV occurs at UC=0.0056; for p=10 the optimum UC=0.0042. Furthermore, one can see that for a low UC (<0.015) a model with order p=10 gives a lower REV; for a larger update coefficient the model order p=6 gives a smaller REV.

DISCUSSION

REV can be seen as the part of the signal which is not explained by the model. In an stochastic model like an AR model REV can never vanish completely. But one can say that if REV is smaller the model explains the signal behaviour more accurately. Consequently, one can derive from Figure 3 that for some fixed model order an optimum update coefficient exists; this UCopt can be found by minimising REV. Furthermore, it can be seen that the AAR(10) model describes the signal more accurately, but it is also found that the optimal update coefficient is smaller. This means the changes in the signals cannot be recognised that fast.

This leads to considerations which are related to the principle of uncertainty as discussed by [8]. The uncertainty principle says: "*the more accurately we try to determine the spectrum as function of time, the less accurately we determine it as a function of frequency, and vice a versa.*" In other words: "in determining evolutionary spectra, one can not obtain simultaneously a high degree of resolution in both the time and frequency domain."
We can reformulate the principle of uncertainty to time-varying autoregressive models, if we assume that MOP correlates to the frequency resolution (how many spectral peaks can be distinguished) and UC determines the time resolution (adaptation speed). Conclusively, it must be formulated that, ”one cannot obtain a high number of accurately estimated AAR parameters and a large adaptation rate simultaneously”.

To make it more simple, if for any reason a large update coefficient is needed, the model order should be chosen small and vice versa.

We assumed that the AAR parameters vary only slowly, meaning highly non-stationary (transient) effects are not considered by the model. It can be expected that in case of a highly non-stationary environment with many transient events, the method may fail. In other words, even if only a very low model order is selected, the update coefficient cannot be increased ad infinitum.

In case of a signal with fast and slow changes it can be expected that more than one combination of {UC, MOP} gives an – at least locally – optimum solution. In this case, other criteria must be used. In EEG analysis with slow variations (e.g. alpha wave) and highly transient events (e.g. muscle artefacts, spindles in the sleep EEG, etc.) these considerations might be important.

CONCLUSION

A method is presented which allows to determine the adaptation rate of AAR estimation algorithms. This method is based on the principle of minimising the (relative) error variance. It estimates how much of the EEG signal Y_t is not described by the AAR model coefficients; accordingly it is used as a measure for the goodness-of-fit;
It can be assumed that also other free parameters e.g. the order of the model or modifications of the RLS and LMS algorithm, can be analysed and optimised. It was shown that the principle of minimising the variance of the error process is not only successful in stationary systems, but can also be used in an adaptive and non-stationary environment.

ACKNOWLEDGEMENT

This work was supported by the Austrian *Fonds zur Förderung der wissenschaftlichen Forschung,* Project 11208 and by part of the Biomed-2 Project BMH4-CT97-2040 (SIESTA).
Thanks to Herbert Ramoser and Britta Ortmayr for their assistance and for proof-reading.

REFERENCES

1 Florian, G. and Pfurtscheller, G.: Dynamic spectral analysis of event-related EEG data. *Electroenceph. clin. Neurophysiol.* 1995, **95**, pp.393-396.

2 Haykin, S.; *Adaptive Filter Theory*. (Prentice Hall, 1986).

3 Widrow B. and Stearns S.D. *Adaptive Signal Processing,* (Prentice Hall, 1985).

4 Haykin, S., Sayed, A.H., Zeidler, J. R., Yee, P. and Wie, P.C. Adaptive Tracking of Linear Time-variant Systems by Extended RLS Algorithms, *IEEE Trans. on Signal Processing;* 1997, **45**, pp.1118-1128

5 Schloegl, A.; Time-varying Autoregressive Models with Application to EEG, *MSc-Thesis*, (University of Technology Graz, 1995)

6 Schloegl, A., Flotzinger, D. and Pfurtscheller G.; Adaptive Autoregressive Modeling used for Single-Trial EEG Classification *Biomed. Techn.* 1997, **42,** pp.162-167.

7 Pfurtscheller, G., Neuper, Ch., Flotzinger, D., Pregenzer, M.; EEG-based discrimination between imagination of right and left hand movement. *Electroenceph. clin. Neurophysiol.* in press 1997.

8 Priestley, M.B.; *Spectral Analysis and Time Series*. (Academic Press, 1981)