7

The Use of Kalman Filter in Biomedical Signal Processing

Vangelis P. Oikonomou, Alexandros T. Tzallas, Spiros Konitsiotis, Dimitrios G. Tsalikakis and Dimitrios I. Fotiadis
Department of Computer Science, University of Ioannina
GR 45110 Ioannina, Greece

1.Introduction

The Kalman Filter (KF) is a powerful tool in the analysis of the evolution of a dynamical model in time. The filter provides with a flexible manner to obtain recursive estimation of the parameters, which are optimal in the mean square error sense. The properties of KF along with the simplicity of the derived equations make it valuable in the analysis of signals. In this chapter an overview of the Kalman Filter, its properties and its applications is presented. More specifically, we focus on the application of Kalman Filter in the Electroencephalogram (EEG) processing, addressing extensions of Kalman Filter such as the Kalman Smoother (KS) in the time varying autoregressive (TVAR) model. The model can be written in a state – space form and the employment of KF provides with an estimation of the AR parameters which can be used for the estimation of the non – stationary signal. It is also demonstrated how these parameters can be used as input features of the signal in a clustering approach.

The Kalman Filter is an estimator with interesting properties like optimality in the Minimum Mean Square Error (MMSE). After its discovery in 1960 (Kalman, 1960), this estimator has been used in many fields of engineering such as control theory, communication systems, speech processing, biomedical signal processing, etc. An analogous estimator has been proposed for the smoothing problem (Rauch et al., 1963), which includes three different types of smoothers, namely fixed-lag, fixed-point and fixed interval (Anderson & Moore, 1979; Brown, 1983). In this chapter we address the fixed interval smoother. The difference between the two estimators, the Kalman Filter and the Kalman Smoother, it is related on how they use the observations to perform estimation. The Kalman Filter uses only the past and the present observations to perform estimation, while the Kalman Smoother uses also the future observations for the estimation. This means that the Kalman Filter is used for on - line processing while the Kalman Smoother for batch processing. The derivations of these two estimators is presented in (Kay, 1993; Grewal & Andrews, 2001; Haykin, 2001). Both estimators are recursive in nature. This means that the estimate of the present state is updated using the previous state only and not the entire past states. The Kalman Filter is not only an estimator but also a learning method (Grewal & Andrews, 2001; Bishop, 2006). The observations are used to learn the states of the model. The Kalman Filter is also a computational tool and some problems may exist due to the finite precision arithmetic of the computers.

Source: Kalman Filter: Recent Advances and Applications, Book edited by: Victor M. Moreno and Alberto Pigazo,
ISBN 978-953-307-000-1, pp. 584, April 2009, I-Tech, Vienna, Austria

The Kalman Filter and the Kalman Smoother have been extensively used in biomedical signal processing. The general idea is to propose a model for the observations, in most cases linear, where some parameters must be estimated. To be able to apply the Kalman Filter or the Kalman Smoother the model for the observations must be written in a state – space form. A state – space model is represented by two equations. One equation, which describes the evolution of the parameters, and a second equation, which describes the relation of the parameters with the observations:

$$\mathbf{x}_t = A\mathbf{x}_{t-1} + \mathbf{w}_t, \tag{1}$$

$$\mathbf{y}_t = C\mathbf{x}_t + \mathbf{v}_t. \tag{2}$$

These two equations represent a state – space model. In the above model $\mathbf{x}_t$ are the parameters in time t, $\mathbf{y}_t$ are the observations, $\mathbf{w}_t$ is the state noise with zero mean and covariance matrix C_w, $\mathbf{v}_t$ is the observation noise with zero mean and covariance matrix C_v, A is the state transition matrix and C is the observation matrix. In the above model the matrices A and C are assumed to be known, as well as the covariance matrices C_v and C_w. However, in reality we are not able to know exactly the above matrices. In that case some assumptions are considered for the model. For example we can assume that the evolution of the parameters is a random walk procedure (Tarvainen et al., 2004), i.e. $A=I$, where I is the identity matrix, or we restrict the matrix A to be a diagonal one (Oikonomou et al., 2007). Also, these matrices can be estimated through an estimation procedure like the EM algorithm (Shumway & Stoffer, 1982; Khan & Dutt, 2007).
In (Sayadi & Shamsollahi, 2008) the authors proposed a non linear model for the electrocardiogram (ECG) signal. They use the model for ECG denoising and compression. To estimate the model parameters they use a modified version of the Kalman Filter, the Extended Kalman Filter (EKF) (Haykin, 2001). In (Kazemi et al., 2008) the authors use the Kalman Filter to detect and extract periodic noise from the ECG. In (Georgiadis et al., 2005; Georgiadis et al., 2007) they assumed that the Evoked Potentials in the Electroencephalogram can be represented as a linear combination of basis functions. The coefficients of the basis functions are assumed to change with time. This assumption lead to the use of the Kalman Filter to estimate the coefficients of the basis functions. In (Oikonomou & Fotiadis, 2005) the authors use the Kalman Smoother to estimate the sources of a dynamic version of a Bayesian PCA. They applied the method to the extraction of fetal EEG.
Besides these applications of the Kalman Filter and the Kalman Smoother for Biomedical Signal Processing, there is a particular application which has been attracted special interest, especially because at the end a time varying spectrum is obtained. This application concerns the use of parametric models such as the AR and ARMA models. In this chapter we will focus on the time varying AR model. The time varying autoregressive (TVAR) model is an AR model where the AR coefficients evolve in time. The parametric spectrum analysis is used to overcome the limited frequency resolution of FFT based methods. The spectral density can be calculated at each frequency point using the model parameters. The TVAR model has been used for EEG spike detection (Oikonomou et al., 2007), for time varying - spectrum estimation of Event Related Synchronization (ERS) and Desynchronization (ERD) (Khan & Dutt, 2007), for the calculation of coherence in the analysis of biomedical signals such EEG and ECG (Arnold et al., 1998) and for time varying spectrum estimation of

intracranial pressure signals from patients with traumatic brain injury (Aboy et al., 2005). In the above studies the TVAR coefficients have been estimated using the Kalman Filter or the Kalman Smoother, while in (Khan & Dutt, 2007) the EM algorithm is used to estimate the parameters of the model.

2. The time varying AR model and Kalman filter

We assume that our signal can be described by an AR model given by:

$$y(t) = \sum_{i=1}^{p} a(i)y(t-i) + e(t), \tag{3}$$

where p is the order of the AR model, $y(t)$ is the observation in time t, $a(i)$ are the AR coefficients and $e(t)$ is the driving noise, which is Gaussian with zero mean and variance σ_e^2. If we let the AR coefficients to vary in time then we have the following equation for the AR model:

$$y(t) = \sum_{i=1}^{p} a_t(i)y(t-i) + e(t), \tag{4}$$

which can be written in vector notation as:

$$y(t) = \mathbf{h}_t\mathbf{a}_t + e(t), \tag{5}$$

where the vector $\mathbf{h}_t$ contains the past observations $[y(t\text{-}1),\dots, y(t\text{-}p)]$ and the vector $\mathbf{a}_t$ the AR coefficients in time t. This model is called the time-varying AR model (TVAR). If we restrict the time evolution of TVAR coefficients to be linear, then we have:

$$\mathbf{a}_t = A\mathbf{a}_{t-1} + \mathbf{w}_t, \tag{6}$$

where A is the state transition matrix and $\mathbf{w}_t$ is the state noise which follows the Gaussian distribution with zero mean and covariance matrix Q. The equations (5) and (6) correspond to a state - space model. Now, given the observations $\mathbf{Y}_{1:k} = [y(1),\dots,y(k)]$ we want to obtain estimates of the TVAR coefficients. This can be done using the Kalman Filter. The Kalman Filter gives the optimal estimates in the mean square error sense. In the estimation procedure we are interested in two quantities, the expectation of conditional states $\mathbf{a}_{t|k} = E\{\mathbf{a}_t \mid \mathbf{Y}_{1:k}\}$ and the error covariance of them $P_{t|k} = E\{(\mathbf{a}_t - \mathbf{a}_{t|k})(\mathbf{a}_t - \mathbf{a}_{t|k})^T \mid \mathbf{Y}_{1:k}\}$. These quantities are obtained using the following equations:

$$\mathbf{a}_{t|t-1} = A\mathbf{a}_{t-1|t-1}, \tag{7}$$

$$P_{t|t-1} = AP_{t-1|t-1}A^T + Q, \tag{8}$$

$$K_t = P_{t|t-1}\mathbf{h}_t^T(\sigma_e^2 + \mathbf{h}_t P_{t|t-1}\mathbf{h}_t')^{-1}, \tag{9}$$

$$\mathbf{a}_{t|t} = \mathbf{a}_{t-1|t-1} + K_t(y(t) - \mathbf{h}_t\mathbf{a}_{t|t-1}), \tag{10}$$

$$P_{t|t} = (I - K_t\mathbf{h}_t^T)P_{t|t-1}, \tag{11}$$

with initial condition $\mathbf{a}_{1|0} = \boldsymbol{\mu}$ and $P_{t|t-1} = \Sigma$, where $\boldsymbol{\mu}$ and Σ are the initial conditions for the states. For more information on how these equations have been derived the interested reader can look in (Kay, 1993; Haykin, 2001). From these equations we can observe how the Kalman Filter is working. To estimate the current state $\mathbf{a}_{t|t}$ a prediction step to obtain the predictive state $\mathbf{a}_{t|t-1}$ based only on the previous state $\mathbf{a}_{t-1|t-1}$ is performed. After that a correction step takes place using the present observation $y(t)$ and the predictive state. Also, we can observe that the update equation for the covariance matrix is $P_{t|t}$ calculated as the difference of two matrices (see Eq. (11)). This can lead to numerical problems and destroy the symmetry of the matrix. To avoid these problems the update equation of covariance $P_{t|t}$ is replaced with the so called Joseph form (Brown, 1983):

$$P_{t|t} = (I - K_t \mathbf{h}_t^T) P_{t|t-1} (I - K_t \mathbf{h}_t^T)^T + \sigma_e^2 K_t K_t^T. \tag{12}$$

Now, the covariance matrix is calculated by adding two matrices. The state at time t using only the observations until time t, $\mathbf{Y}_{1:t}$, is estimated. When all the observations are available, $\mathbf{Y}_{1:N}$, to estimate the state at time t the Kalman Smoother is used. The update smoothing equations are:

$$J_{t-1} = P_{t-1|t-1} A^T P_{t|t-1}^{-1}, \tag{13}$$

$$\mathbf{a}_{t-1|N} = \mathbf{a}_{t-1|t-1} + J_{t-1}(\mathbf{a}_{t|N} - A\mathbf{a}_{t-1|t-1}), \tag{14}$$

$$P_{t-1|N} = P_{t-1|t-1} + + J_{t-1}(P_{t|T} - P_{t|t-1}) J_{t-1}^T. \tag{15}$$

The derivation of those equations is explained in (Haykin, 2001). The equations of Kalman Filter, together with the above smoothing equations, consist the Kalman Smoother. In general to apply the Kalman Filter or the Kalman Smoother to a model, we must write the model in a state – space form. After that the above equations can be applied easily. However, there are several parameters which are assumed known before the application of the update equations. These parameters are the state transition matrix A, the covariance of the state noise Q, the variance of the driving noise σ_e^2, and the initial conditions, $\theta = \{A, Q, \sigma_e^2, \boldsymbol{\mu}, \Sigma\}$. The parameters θ can be tuned based on some empirical knowledge (Oikonomou et al., 2007) or define a function of parameters and perform optimization to obtain the optimal values of the parameters like the EM algorithm (Khan & Dutt, 2007). In the next section the EM algorithm is described to tune the parameters θ. The EM algorithm was introduced in (Dempster et al., 1977) and has been used for the optimization of linear state – space model for the first time in (Shumway & Stoffer, 1982).

3. The EM algorithm and the Kalman smoother

The EM algorithm is a procedure to perform maximum likelihood estimation. The objective of the algorithm is to maximize the likelihood of the observations, $\mathbf{Y}_{1:N}$, in the presence of hidden variables, $\mathbf{a}_{t|N}$, $t = 1, \ldots, N$. The maximization is performed with respect to the parameters θ, $\theta_{ML} = \max_\theta \log p(\mathbf{Y}_{1:N} | \theta)$, where $p(\mathbf{Y}_{1:N} | \theta)$ is the probability density function of the observations, which is called likelihood, when it is seen as a function of the parameters θ. Direct maximization cannot be performed because the hidden variables are not available.

The maximization must be done with respect to the hidden variables, in our case the time varying AR coefficients, and the model parameters. The EM algorithm is an iterative scheme consisting of two steps, the E-step and the M-step. In the E-step the expected values of the hidden variables are evaluated and in the M-step the maximization is performed with respect to the model parameters. To perform the E-step the expected complete log-likelihood must be calculated as follows:

$$F = E\left\{\log p(\mathbf{Y}_{1:N}, \mathbf{a}_{1:N|N}) \middle| \mathbf{Y}_{1:N}\right\}. \tag{16}$$

The expected likelihood depends on three quantities:

$$\mathbf{a}_{t|N} = E\left\{\mathbf{a}_t \middle| \mathbf{Y}_{1:N}\right\}, \tag{17}$$

$$S_{t|N} = E\left\{\mathbf{a}_t \mathbf{a}_t^T \middle| \mathbf{Y}_{1:N}\right\} = P_{t|N} + \mathbf{a}_{t|N}\mathbf{a}_{t|N}^T, \tag{18}$$

$$S_{t,t-1|N} = E\left\{\mathbf{a}_t \mathbf{a}_{t-1}^T \middle| \mathbf{Y}_{1:N}\right\} = P_{t,t-1|N} + \mathbf{a}_{t|N}\mathbf{a}_{t-1|N}^T \tag{19}$$

The first two quantities can be calculated using the Kalman Smoother equations, while for the calculation of the last quantity we use the following equation (Bishop, 2006):

$$P_{t,t-1|N} = J_{t-1}P_{t|N}. \tag{20}$$

The M - step involves direct differentiation of F with respect to the parameters θ. The estimates for the model parameters θ are given as:

$$A_{new} = [\sum_{t=2}^{N} S_{t,t-1|N}][\sum_{t=2}^{N} S_{t-1|N}]^{-1}, \tag{21}$$

$$Q_{new} = [1/(N-1)][\sum_{t=2}^{N} S_{t|N} - A_{new}\sum_{t=2}^{N} S_{t-1,t|N}], \tag{22}$$

$$\sigma_e^2 = (1/N)\sum_{t=1}^{N}(y_t^2 - 2\mathbf{h}_t^T\mathbf{a}_{t|N}y_t + \mathbf{a}_{t|N}^T S_{t|N}\mathbf{a}_{t|N}), \tag{23}$$

$$\mu_{new} = \mathbf{a}_{1|N}, \tag{24}$$

$$\Sigma_{new} = P_{1|N}. \tag{25}$$

In the presented model the EM algorithm consists of two iterative steps. First, the application of the Kalman Smoother, using the parameters from the previous step, to obtain the expected statistics, and second maximization of the expected log – likelihood with respect to the parameters. These two steps are applied iteratively until convergence of the likelihood. As we can see the use of EM algorithm needs statistics which can be obtained only with the Kalman Smoother, and not using the Kalman Filter. The use of the Kalman Filter for the calculations of these statistics leads to suboptimal procedures. At the end of the EM algorithm we obtain the TVAR coefficients, which can be used to obtain a time varying spectrum, given by:

$$P(t,f)=\frac{\sigma_e^2}{\left|1-\sum_{j=1}^{p}a_{t|N}(j)e^{-i2\pi j\frac{f}{f_s}}\right|^2}, \tag{26}$$

where f_s is the sampling frequency.

4. Biomedical signal processing: an application to EEG

4.1 The EEG signal

The electroencephalogram can be roughly defined as the signal which corresponds to the mean electrical activity of the brain in different locations of the head. More specifically, it is the sum of the extracellular current flows in a large group of neurons. It can be acquired using either intracranial electrodes inside the brain or scalp electrodes on the surface of the head (Niedermeyer & Lopes da Silva, 1993). The EEG has been found to be a valuable tool in the diagnosis of numerous brain disorders. Nowadays, the EEG recording is a routine clinical procedure and is widely regarded as the physiological "gold standard" to monitor and quantify levels of drowsiness and wakefulness but also for detection of epileptic spikes and seizures and generally for the diagnosis of epilepsy (Tzallas et al., 2006). The electric activity of the brain is usually divided into three categories: 1) bioelectric events produced by single neurons, 2) spontaneous activity, and 3) evoked potentials. EEG spontaneous activity is measured on the scalp or on the brain. Clinically meaningfull frequencies lie between 0.1 Hz and 100 Hz. In more restricted sense, the frequency range is classified into several frequency components, or delta rhythm (δ: 0.5-4 Hz), theta rhythm (θ: 4-8 Hz), alpha rhythm (α: 8-13 Hz), beta rhythm (β: 13-30 Hz), and gamma rhythm (γ: 30-60 Hz) (Niedermeyer & Lopes da Silva, 1993).

The properties of the EEG signal can be described as complex (Niedermeyer & Lopes da Silva, 1993; Thakor & Tong, 2004). The EEG complexity originates from the intricate neural system. Traditionally, the spontaneous EEG is characterized as a linear stochastic process with similarities to noise. From the signal processing view, EEG has the following properties (Thakor & Tong, 2004): (a) Noisy and pseudo-stochastic: The EEG is often between 10-300μV, which is easily affected by various physiological and electrical noises. Meanwhile, artefacts from electrocardiogram (ECG), electrooculogram (EOG), electromyogram (EMG), and recording systems can also contaminate the signals. Even the EEG shows a high degree of randomness and nonstationarity. (b) Time-varying and nonstationary: EEG is not a stationary process; it varies with the physiological states. The waveforms may include a complex of regular sinusoidal waves, irregular spikes/polyspikes, or spindles/polyspindles. In most pathological conditions, such as epileptic seizures, the EEG may show evident singularity or nonstationarity. In practice, EEG is considered as a stationary process over a relatively short period (approximately 3.5sec for routine spontaneous EEG) (Goel et al., 1996). (c) High nonlinearity: Although the traditional linear models of EEG still play significant roles in EEG analysis and diagnosis, EEG is a nonlinear process (Palu, 1996). This kind of nonlinearity is also time-, state-, and site-dependent (Pijn et al., 1991).

One of the most important challenges of EEG analysis is the quantification of the manifestations of epilepsy (Niedermeyer & Lopes da Silva, 1993; Thakor & Tong, 2004). The main goal is to establish a correlation between the EEG and clinical or pharmacological conditions. One of the possible approaches is based on the properties of the inter-ictal EEG (electrical activity measured between seizures), which typically consists of linear stochastic background fluctuations interspersed with transient nonlinear spikes, sharp waves or spikes-and-wave complexes (Tzallas et al., 2006). These transient potentials originate as a result of a simultaneous pathological discharge of neurons within a volume of at least several mm^3. The traditional definition of a spike is based on its amplitude, duration, sharpness, and emergence from its background (Wilson & Emerson, 2002). However, automatic epileptic spike detection systems based on this direct approach suffer from false detections in the presence of numerous types of artefacts and non-epileptic transients (Wilson & Emerson, 2002; Tzallas et al., 2006). This shortcoming is particularly acute for long-term EEG monitoring of epileptic patients, which became common in 1980s (Waterhouse, 2003; Mormann et al., 2007).

There has also been a challenge to find functional cerebral activation indices for cognitive processes involved in a given task. The EEG is a continuous measure over time and can be used to study ongoing activity in the brain while subjects perform long-lasting and/or variable tasks. The alpha rhythm of the EEG is predominantly observed over the posterior cortex (Lähteenmäki et al., 1999). This rhythm correlates with relaxation, and for this reason it has been interpreted as a sign of inhibition of activity in the areas over which it has been recorded. Activation of the cortex causes a desynchronization of the alpha band, i.e. its amplitude decreases, while alpha synchronization denotes the increase of alpha activity (Pfurtscheller & Aranibar, 1977; Pfurtscheller, 1989). When alpha desynchronization or synchronization is related to an internally or externally paced event, it is called as event-related desynchronization (ERD) (Pfurtscheller, 1977) or event-related synchronization (ERS), respectively. The quantification of ERD/ERS requires the comparison of two different experimental conditions. ERD and ERS are defined as the relative difference in the EEG alpha power between the reference recorded before each event and the actual event.

ERD/ERS is, thus, a 'within-subject' measure of cortical activation and is expressed as a percentage. ERD and ERS can be either externally (by stimuli) or internally (by voluntary behavior) paced and they have a specific topographical distribution depending upon the state of the brain, stimulus paradigm and modality (Pfurtscheller, 1989). ERD has been observed e.g. during complex auditory stimulation (Krause et al., 1994), during cognitive and attentional tasks, and during voluntary movement tasks (Pfurtscheller, 1977). The ERD/ERS of the lower alpha frequencies (8–10 Hz) has been claimed to reflect non-specific cognitive functions, such as sustained attention, while that of the upper alpha frequencies (10–12 Hz) appears to reflect stimulus-related, i.e. task-specific cognitive processes.

We apply the TVAR model for EEG spike identification and ERS/ERD frequency tracking. For both problems the algorithms have been initialized as follows. First when using the Kalman Smoother alone the state transition matrix has been set to be the identity matrix. The covariance of the state noise was assumed diagonal with the same element in the diagonal, σ_w^2. This parameter has been adjusted based on visual analysis of the data. When using the Kalman Smoother with the EM, the algorithm has been initialized as follows: To obtain an initial estimate of the covariance of state noise, the signal has been divided into overlapped segments. In each segment we find the AR coefficients using the Matlab's `aryule` function. We set the matrix $A=I$ and from the local estimates of AR coefficients we find an initial estimate for the covariance of state noise. For EEG spike identification experiments the model order was set to 2, while in the ERD/ERS experiments the model order was set to 5.

These values have been found to work well, based on visual analysis of the available datasets.

4.2 Application of the TVAR model to EEG spike identification

The EEG data used is part of the dataset described in (Oikonomou et al., 2007). The EM algorithm for the state – space model in the problem of EEG spike identification is applied. After the calculation of TVAR coefficients the time varying spectrum is calculated. The time varying spectrum is used to identify the EEG spike. In Fig. (1) an EEG segment with four spikes is depicted. Three of them can be identified easily, while one of them is not so easily to be identified. In Fig (1) we show the EEG segment, the time varying spectrum and a zoom of time varying spectrum in the region of interest. In this case the region of interest is the frequency range where a spike can be observed. According to (Pinault et al., 2001) this range is from 5-20 Hz. We can observe that the time varying spectrum give us an indication for the spike position. We observe that the first three spikes are identified in the frequency around 10 Hz, while the last spike is observed around 6-7 Hz. This difference between the spikes can be justified by taking into account the waveform morphology of the spikes. The first three spikes have similar morphology and they differ from the last spike.

In Fig. (2) we show the EEG segment, the time varying spectrum and a zoom of time varying spectrum using the Kalman Smoother. In that case we assume that the TVAR coefficients evolve in time according to a random walk. The covariance of the state noise in that case is tuned based on visual analysis and is set equal to $10^{-3}I$, where I is the identity matrix. We can observe that using only the Kalman Smoother we place some restrictions on the model. The state transition matrix was set equal to the identity matrix and we assume that the covariance of the state noise is diagonal with the same element in the diagonal (isotropic model). These restrictions are not based on some optimization procedure but on assumptions that we make for the model to make it simpler, i.e. it is more easily to find, through visual analysis, the isotropic covariance to have a meaningful solution of the Kalman Smoother than the full covariance. In Fig. (3) the Instantaneous Frequency (IF) using the Kalman Smoother (Fig. 3a) and the EM with Kalman Smoother (Fig. 3b) are shown. In the results we define the IF as the frequency where the maximum value of a time varying spectrum is located. It is clear that the use of EM gives better estimation of IF. It is interesting here to note how the IF is changing during the spike. The slow rise of IF few msec before the appearance of a spike could represent the accumulation of excitatory postsynaptic potentials from dendritic spines of pyramidal neurons from the epileptic brain. In the case of isolated spikes the activity is self-restricted and does not spread to involve adjacent brain areas. It is interesting, in the case of a series of spikes, that the rise in the frequency persists for the duration of the event. This finding might have important implications in the prediction of epileptic crisis or "epileptic events".

The TVAR coefficients, besides the estimation of the time varying spectrum, can be used as features of the signal and can be used for classification or clustering purposes. In the case of EEG signal clustering, the number of clusters is assumed to be two, one cluster for the spike and one cluster for the background activity. For the clustering the `netlab` toolbox is used (Nabney & Bishop, 2004) and the clustering method was based on the gaussian mixture models. In Fig. (4) we see the results of clustering in an EEG signal. We can see that the clustering have correctly identified the EEG samples which belong to the spikes.

4.3 Application to event related synchronization / desynchronization

The event related phenomena represent frequency changes of the ongoing EEG activity. A decrease of power in a given frequency band is called event related desynchronization (ERD) and an increase event related synchronization (ERS). ERS/ERD phenomena are generated by changes in one or more parameters which control oscillations in neuronal networks (Pfurtscheller & Lopes da Silva, 1999). When we refer to ERS/ERD is necessary to specify the frequency band. In our experiments this frequency band is 10 Hz, which corresponds to the alpha rhythm band. The ERS/ERD phenomena are related to frequency changes, which can be detected by frequency analysis. The TVAR model is used here to produce the time varying spectrum, which is helpful in the analysis of the ERS/ERD.

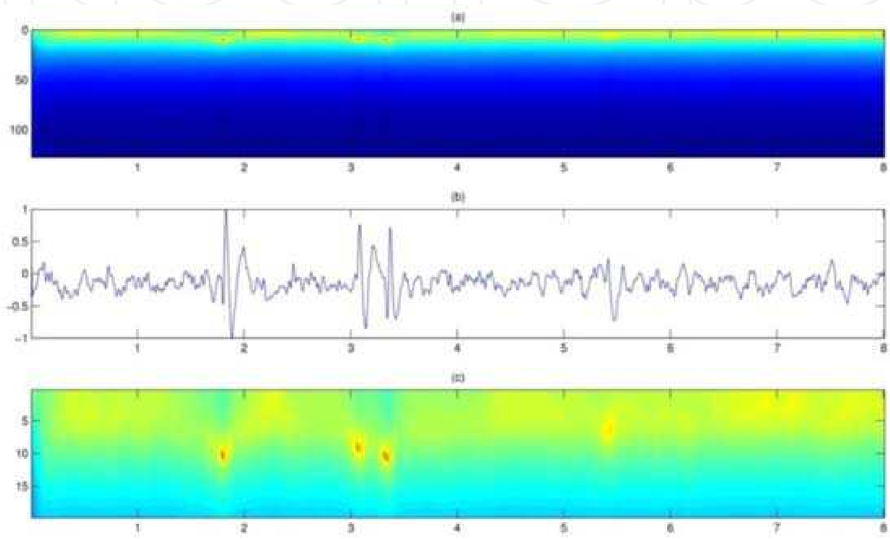

Fig. 1. Example of spike identification using the EM and Kalman Smoother, x – axis corresponds to time and y- axis corresponds to frequency. (a) The time varying spectrum. (b) The EEG signal. (c) Zoom in the frequency range 5-20 Hz.

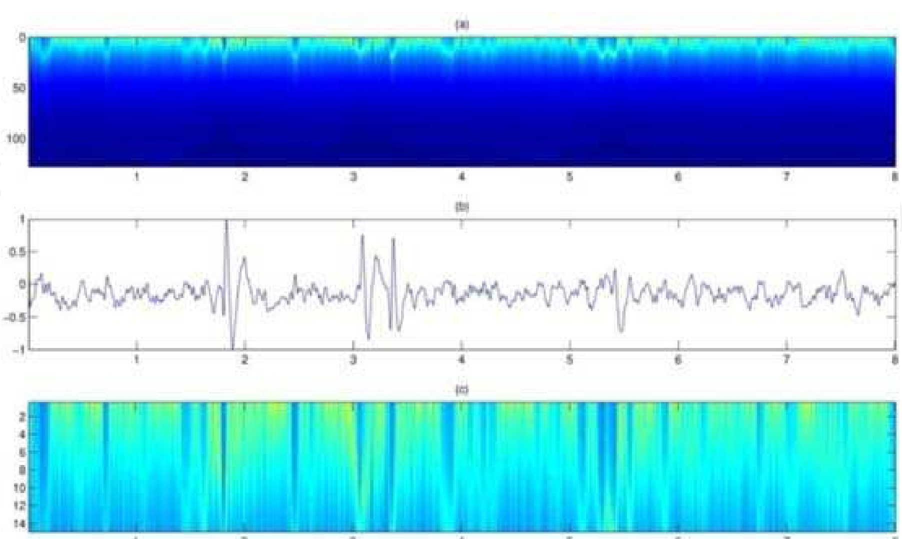

Fig. 2. Example of spike identification using the Kalman Smoother. (a) The time varying spectrum, x – axis corresponds to time and y- axis corresponds to frequency. (b) The EEG signal. (c) Zoom in the frequency range 5-20 Hz.

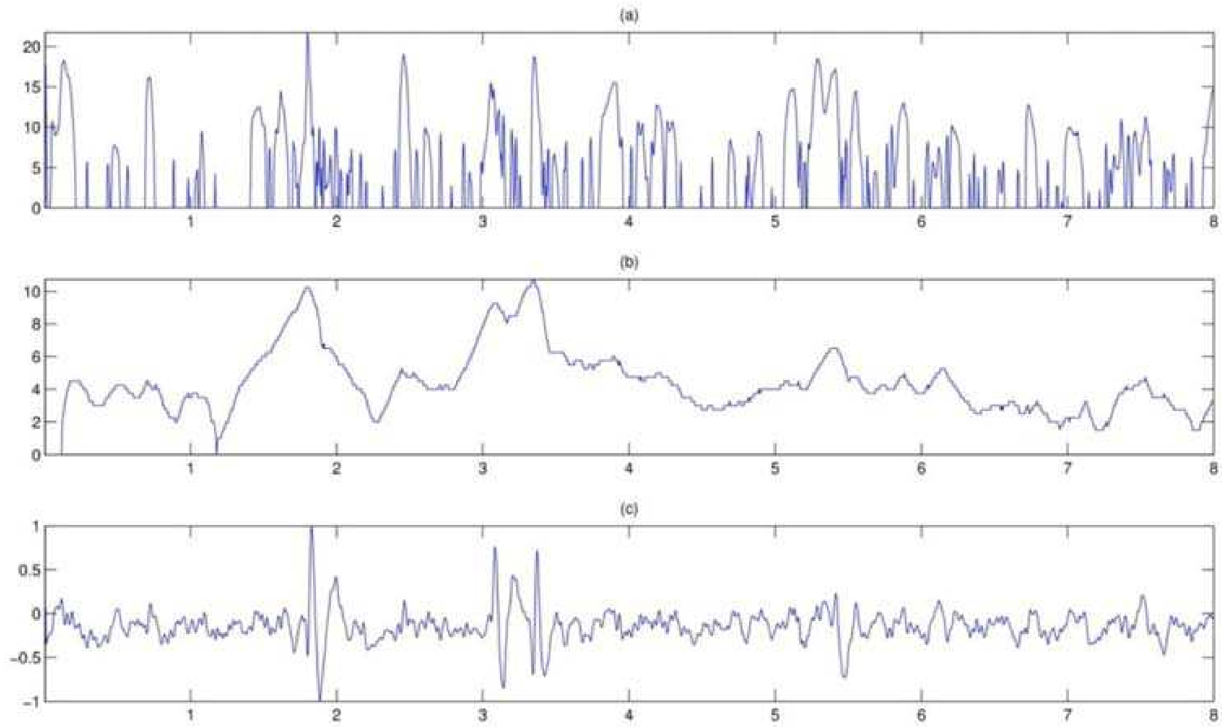

Fig. 3. Instantaneous Frequency. (a) Kalman Smoother, x - axis corresponds to time and y – axis to the IF value, (b) Kalman Smoother with EM, x - axis corresponds to time and y – axis to the IF value, (c) EEG signal, x - axis corresponds to time and y – axis to the signal amplitude.

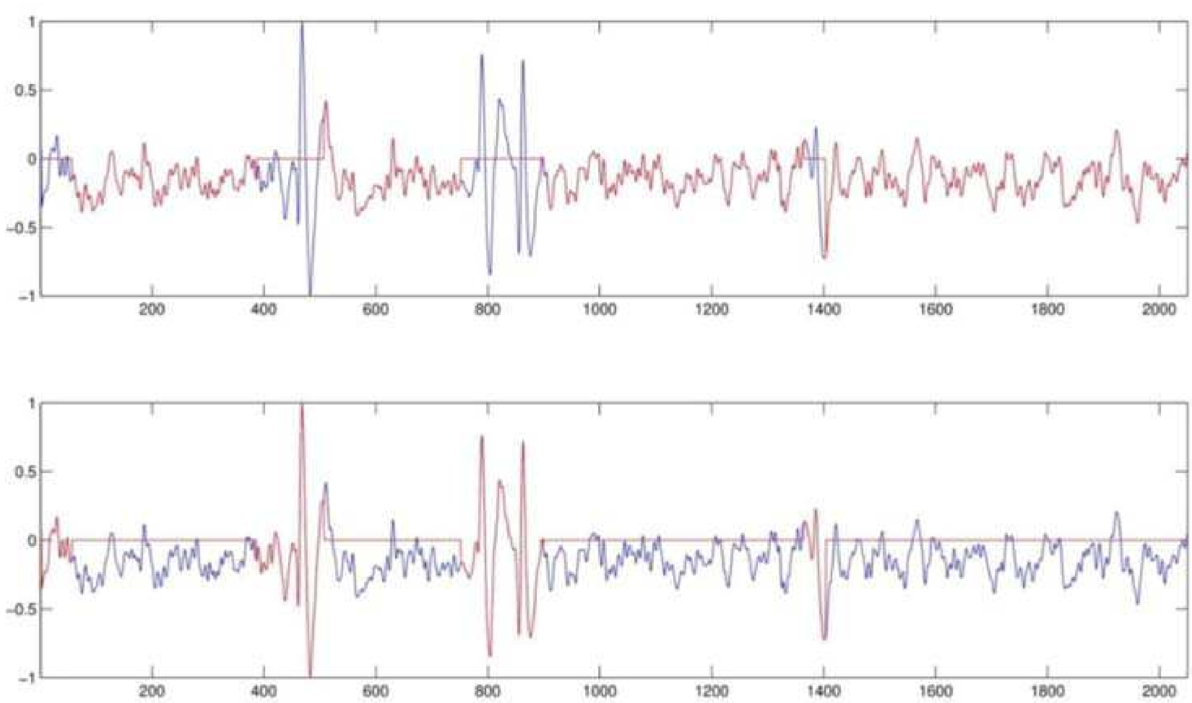

Fig. 4. Clustering using the TVAR coefficients. (a) The EEG signal (blue line) and the cluster (red line) which corresponds to background activity and (b) The EEG signal (blue line) and the cluster (red line) which corresponds to spikes.

The dataset used in this section is the motor imagery dataset provided by the Gratz University of Technology (Blankertz et al., 2004). In this experiment the subject was asked to control a feedback bar by means of imagery left or right hand movements. Two bipolar EEG channels, C3 and C4, have been used to obtain the EEG signals. The dataset consists of 280 trials of 9sec. The first 2sec was quite, at t=2sec an acoustic stimulus indicates the beginning of the trial, and at t=3sec an arrow was displayed as cue. At the same time the subject was asked to move the bar on the direction of the cue. The EEG signal was sampled at 128 Hz and it was filtered between 0.5 and 30 Hz.

The TVAR model has been applied to these trials using the Kalman Smoother only and the Kalman Smoother together with the EM algorithm. In the case of using only the Kalman Smoother some simplifications of the model are introduced. This restricts the number of parameters that must be known before the Kalman Smoother. The state noise is assumed to have diagonal covariance and the same element in the diagonal, i.e. $Q = \sigma_w^2 I$. Based on visual analysis of the data $\sigma_w^2 = 0.001$. Also the state transition matrix has been set to the identity matrix, i.e $A{=}I$. The model order for both methods has been set equal to 5. The initial state and covariance was set to zero and identity matrix, respectively.

A trial from the dataset with the extracted IF using the Kalman Smoother and Kalman Smoother with EM is shown in Fig. (5). It is clear that both methods track well the frequency evolution. However, the Kalman Smoother with EM produces smoother estimates than using the Kalman Smoother alone. Also, at Fig. (6) the time varying spectrums of both methods are presented. It is clear that the Kalman Smoother alone produces a noisy spectrum. All these observations can be justified by the fact that the Kalman Smoother alone imposes constraints on the state evolution model, in contrast to the Kalman Smoother with EM where such constraints do not exist.

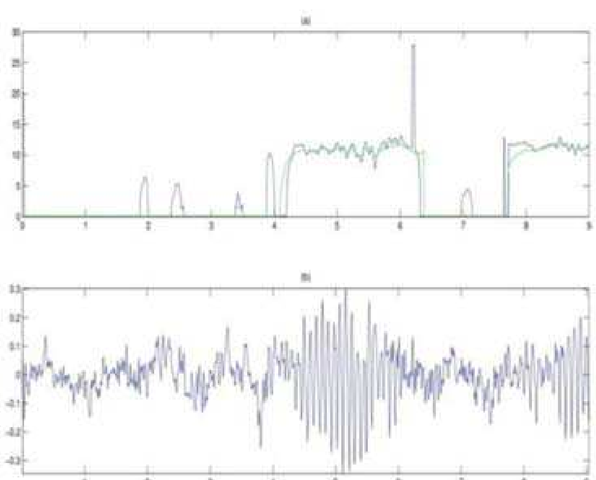

Fig. 5. (a) Instantaneous Frequency using the Kalman Smoother and EM (green line) and the Kalman Smoother only (blue line) (b) Trial data.

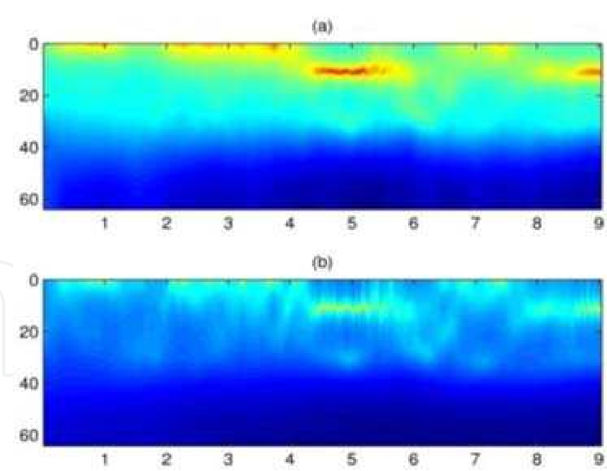

Fig. 6. Time varying spectrum using (a) the Kalman Smoother with EM and (b) the Kalman Smoother only

For each trial the time varying spectrum has been calculated using the TVAR coefficients. Also, for each trial the Instantaneous Frequency (IF) has been extracted from the time varying spectrum. The mean spectrum of trials is presented in Figs. (8,9) for the left hand movement. In Fig. (7) the mean Instantaneous Frequency is presented. We can observe that at channel C3 there is a decrease in IF from 8 Hz to 6 Hz at the beginning of the trial, t=2sec. Some milliseconds after that the IF increases to 8 Hz (t=3sec), where it stays for the rest of the trial. For the C4 channel we can observe that the IF decreases at the presentation of the acoustic stimulus, which indicates the beginning of the trial. Then from t=2sec to t=3sec the IF increases, to start to decrease from t=3sec to t=4sec. At t=4sec, the IF increases up to 6 Hz. Comparing the Kalman Smoother and the Kalman Smoother with EM we can observe that the Kalman Smoother gives noisy estimates of the time varying spectrum and IF. This change in frequency can be justified as reorganization of the brain neurons activity due to

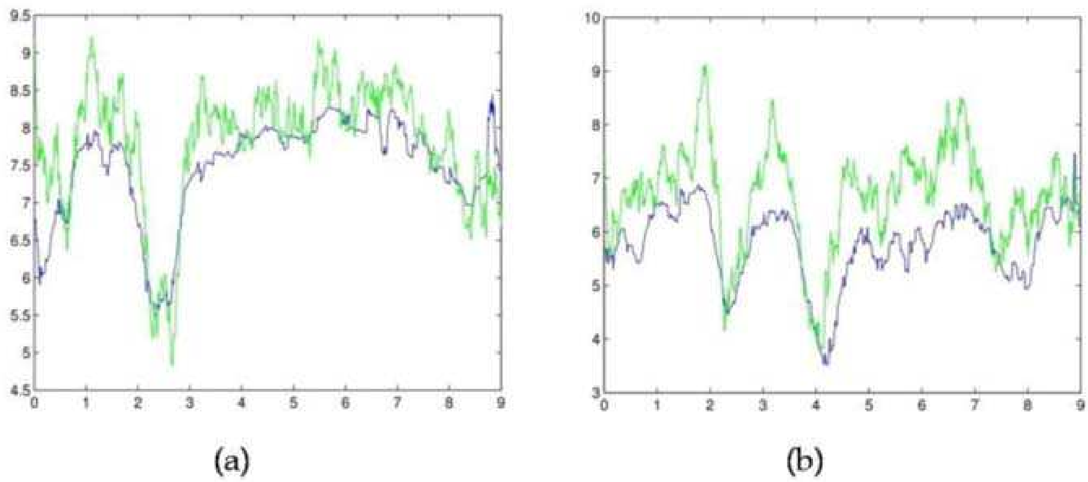

Fig. 7. Mean Instantaneous Frequency for left hand movement. (a) Channel C3. (b) Channel C4. In both figures the blue line corresponds to the Kalman Smoother with EM and green line to the Kalman Smoother only.

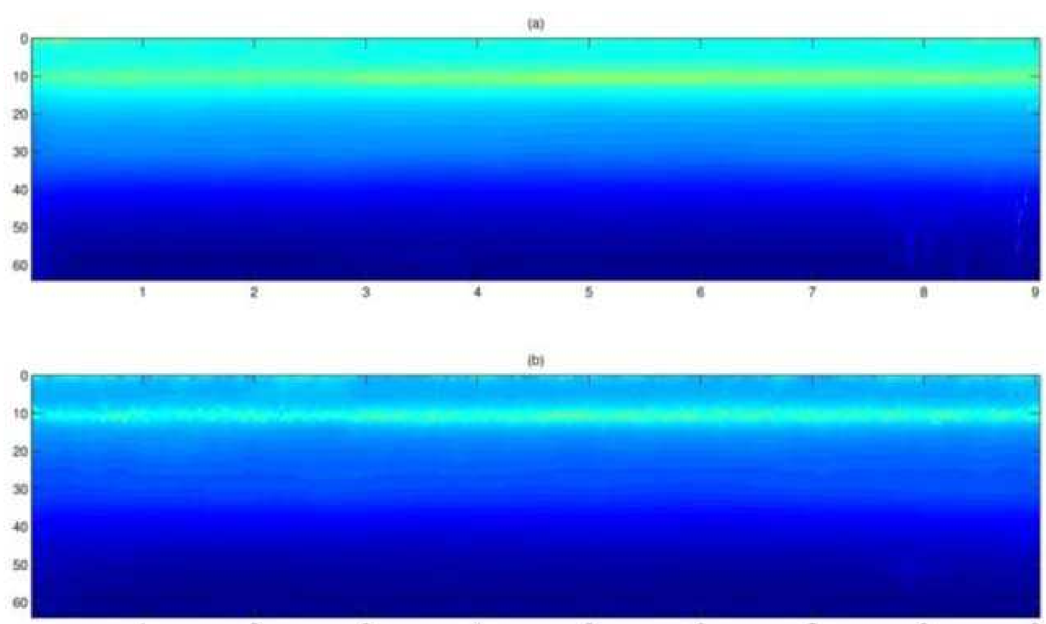

Fig. 8. Mean time varying PSD (Power Spectrum Density) for channel C3 in left hand movement. (a) Kalman Smoother with EM, (b) Kalman Smoother only.

the stimulus. It is clear that a change in IF is observed at t=2sec for both channels C3 and C4. However, at t=3sec there is a difference, a change in IF is observed only for channel C4. At the presentation of cue both brain positions are activated. However, after the cue is presented activation is observed only at C4 (t=3sec). The term activation is used here to describe a change in IF. In Figs. (10,11 and 12) the IF and the mean time varying spectrums for the right hand movements are shown. Similar conclusions can be extracted if we take into the account the hemispherical asymmetry of the brain due to the motor imagery experiment, i.e. we can see similarities in Fig. (7b) and in Fig. (10a). In both cases there is an increase in IF from t=2sec to t=3sec, and after that there is a decrease from t=3sec to t=4sec.

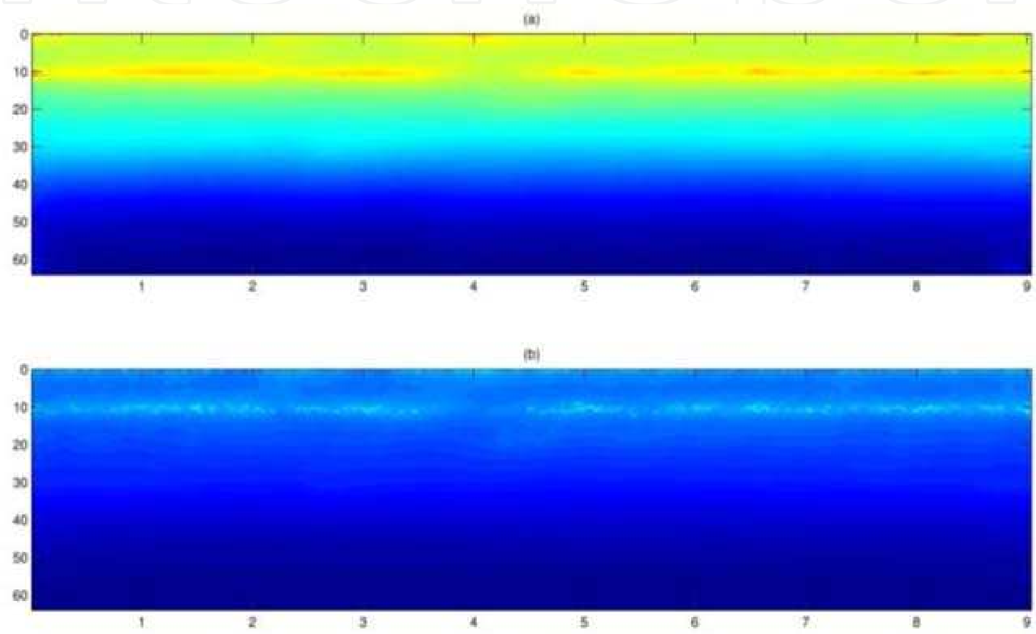

Fig. 9. Mean PSD (Power Spectrum Density) for channel C4 in left hand movement. (a) Kalman Smoother with EM, (b) Kalman Smoother only.

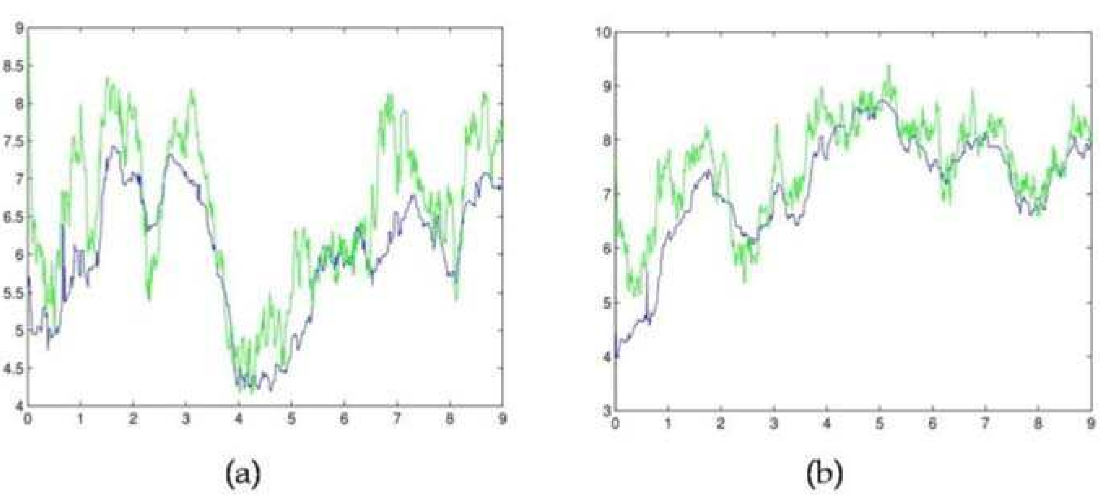

Fig. 10. Mean Instantaneous Frequency for right hand movement. (a) Channel C3. (b) Channel C4. In both figures the blue line corresponds to the Kalman Smoother with EM and green line to the Kalman Smoother only.

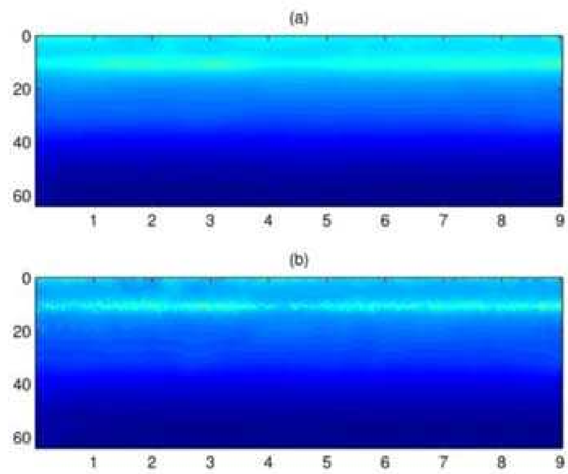

Fig. 11. Mean time varying PSD for channel C3 in right hand movement. (a) Kalman Smoother with EM, (b) Kalman Smoother only.

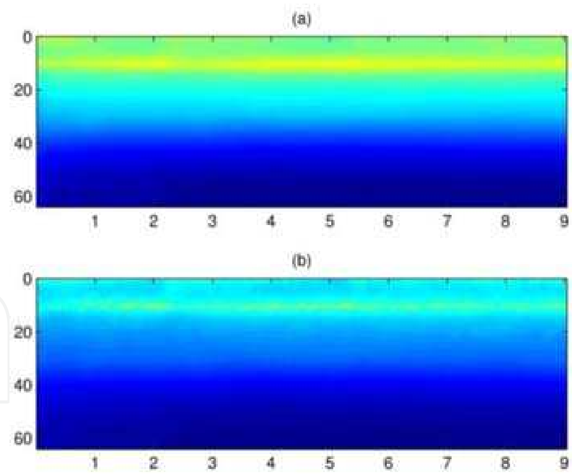

Fig. 12. Mean time varying PSD for channel C4 in right hand movement. (a) Kalman Smoother with EM, (b) Kalman Smoother only.

5. Discussion

In this chapter the Kalman Smoother, with or without the EM algorithm, has been used for the processing of the EEG signal in two cases, epileptic form spike identification and ERD/ERD analysis. Use of the Kalman Smoother forces to some simplifications of the model. This is performed in order to decrease the number of parameters which must be tuned. Based on the assumptions that the state transition matrix is the identity and the covariance is diagonal with the same element on the diagonal, there is one parameter to be tuned, the variance of the state noise. The value of this parameter defines how smooth or rough will be the evolution of states, in our case the TVAR coefficients. Large values of variance indicate rough estimates for the TVAR coefficients. This has as a result a noisy time varying spectrum. Small values indicated smooth estimates for the TVAR coefficients and hence a smooth time varying spectrum. The value of this parameter depends on the problem. In the case where we expect that the time varying spectrum is smooth, a small value for the variance of the state noise is preferable.

However, the parameters can be estimated based on some optimization procedure like the EM algorithm. The EM algorithm provides with estimates of the parameters. So the tuning of the parameters is done automatically based on the dataset, without manual settings. This fact permits the use of full covariance for the state noise and a general transition state matrix. As a consequence the model is more flexible because of the different types of state noise.

We observe that the Kalman Smoother with EM provides with smoother estimates than using the Kalman Smoother alone. This happens because the first approach can capture the patterns of the signal more accurately. In the estimation of the IF in the spike problem it is observed that the IF starts to increase before the appearance of the spike. Also, in the ERD/ERS analysis we observe that the IF is modulated when some events take place on the experiment, like the sound at t=2sec which denotes the beginning of the trial. In both problems we observe that the IF is a good measure to track changes in EEG activity.

The Kalman Smoother or the Kalman Filter can be used for the estimation of time varying AR coefficients. The use of TVAR model is twofold. First, the TVAR coefficients can be used for the estimation of the time varying spectrum. Usually, the Power Spectrum Density (PSD) estimation of biomedical signals is performed using classical techniques based on Fourier Transform which presents two problems: the frequency resolution and the assumption of stationarity. To overcome the problem of frequency resolution a parametric approach based on AR model can be used. However, the assumption of stationarity restricts the use of a parametric model to stationary signals. A solution to these problems is provided by the TVAR model. Second, the TVAR coefficients can be used as input for classification or clustering purposes. The use of AR model is not restrictive and other parametric models such as ARMA can be used for time varying spectrum estimation (Tarvainen et al., 2004).
Besides the particular application of the Kalman Filter or the Kalman Smoother for time varying spectrum estimation there is another application. In this a parameterization of the model is assumed, i.e. the signal is assumed to be a linear combination of basis functions. This signal is observed in additive noise. Assuming that the parameters are evolving in time the Kalman Filter or Smoother can be used to produce the evolution of the parameters in time. At the end an estimate of the signal can produced. This approach have been applied for the estimation of evoked potentials (Georgiadis et al., 2005; Georgiadis et al., 2007).

6. Conclusions

In this chapter we presented how the Kalman Filter and the Kalman Smoother is used for time varying spectrum estimation. The estimation of spectrum is based on a parametric model, a TVAR model. Given the evolution between the states, the Kalman Filter can be used to estimate the AR coefficients which evolve in time. At the end of this procedure a time varying spectrum is calculated. In the case where all the observations are available before the application of the estimation procedure, the Kalman Smoother can be used instead of the Kalman Filter. However, in that case, on line processing of the data cannot be done. When we are interested for batch processing the Kalman Smoother with the EM algorithm is a valuable tool. The use of EM algorithm leads to a process which does not depend on manual settings, which usually are not optimal. The main idea behind the use of Kalman Filter is to propose a model for the problem under consideration, which depends on some parameters. If we let the parameters of the model to change in time then the estimation of the parameters can be performed using the Kalman Filter. If the linearity assumption does not hold then some modifications of the Kalman Filter can be used such as the Extended Kalman Filter.

7. References

Aboy, M., Marquez, O.W., McNames, J., Hornero, R., Trong, T. & Goldstein, B. (2005). Adaptive modeling and spectral estimation of nonstationary biomedical signals based on Kalman filtering. *IEEE Trans Biomed Eng* 52(8): 1485-9.

Anderson, B.D.O. & Moore, J.B. (1979). Optimal Filtering, Englewood Cliffs, NJ: Prentice-Hall.

Arnold, M., Miltner, W.H., Witte, H., Bauer, R. & Braun, C. (1998). Adaptive AR modeling of nonstationary time series by means of Kalman filtering. *IEEE Trans Biomed Eng* 45(5): 553-62.

Bishop, C.M. (2006). Pattern recognition and machine learning, Springer.

Blankertz, B., Muller, K.R., Curio, G., Vaughan, T.M., Schalk, G., Wolpaw, J.R., Schlogl, A., Neuper, C., Pfurtscheller, G., Hinterberger, T., Schroder, M. & Birbaumer, N. (2004). The BCI Competition 2003: progress and perspectives in detection and discrimination of EEG single trials. *IEEE Trans Biomed Eng* 51(6): 1044-51.

Brown, R.G. (1983). Introduction to random signal analysis and Kalman filtering, Wiley New York.

Dempster, A.P., Laird, N.M. & Rubin, D.B. (1977). Maximum Likelihood from Incomplete Data Via Em Algorithm. *Journal of the Royal Statistical Society Series B-Methodological* 39(1): 1-38.

Georgiadis, S.D., Ranta-Aho, P.O., Tarvainen, M.P. & Karjalainen, P.A. (2007). A subspace method for dynamical estimation of evoked potentials. *Comput Intell Neurosci*: 61916.

Georgiadis, S.D., Ranta-aho, P.O., Tarvainen, M.R. & Karjalainen, P.A. (2005). Single-trial dynamical estimation of event-related potentials: A Kalman filter-based approach. *IEEE Trans Biomed Eng* 52(8): 1397-1406.

Goel, V., Brambrink, A.M., Baykal, A., Koehler, R.C., Hanley, D.F. & Thakor, N.V. (1996). Dominant frequency analysis of EEG reveals brain's response during injury and recovery. *IEEE Trans Biomed Eng* 43(11): 1083-92.

Grewal, M.S. & Andrews, A.P. (2001). Kalman Filtering: Theory and Practice Using MATLAB, Wiley.

Haykin, S.S. (2001). Kalman Filtering and Neural Networks, Wiley Chichester.

Kalman, R.E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Transaction of the ASME* 82: 35-45.

Kay, S.M. (1993). Fundamentals of statistical signal processing. Englewood Cliffs, N.J., PTR Prentice-Hall.

Kazemi, R., Farsi, A., Ghaed, M.H. & Karimi-Ghartemani, M. (2008). Detection and extraction of periodic noises in audio and biomedical signals using Kalman filter. *Signal Processing* 88(8): 2114-2121.

Khan, M.E. & Dutt, D.N. (2007). An expectation-maximization algorithm based Kalman smoother approach for event-related desynchronization (ERD) estimation from EEG. *IEEE Trans Biomed Eng* 54(7): 1191-1198.

Krause, C.M., Lang, H.A., Laine, M., Helle, S.I., Kuusisto, M.J. & Pörn, B. (1994). Event-related desynchronization evoked by auditory stimuli. *Brain Topography* 7(2): 107-112.

Lähteenmäki, P.M., Krause, C.M., Sillanmäki, L., Salmi, T.T. & Lang, A.H. (1999). Event-related alpha synchronization/desynchronization in a memory-search task in adolescent survivors of childhood cancer. *Clinical Neurophysiology* 110(12): 2064-2073.

Mormann, F., Andrzejak, R.G., Elger, C.E. & Lehnertz, K. (2007). Seizure prediction: the long and winding road. *Brain* 130: 314-333.

Nabney, I. & Bishop, C. (2004). "Netlab toolbox." from http://www.ncrg.aston.ac.uk/netlab/.
Niedermeyer, E. & Lopes da Silva, F.H. (1993). Electroencephalography: basic principles, clinical applications, and related fields. Baltimore, William &Wikins
Oikonomou, V.P. & Fotiadis, D.I. (2005). A Bayesian PCA Approach for Fetal ECG extraction for Biomedical Signal Denoising *Conf Proc Eur Med Biol Eng (EMBEC'05)*, Prague, Czech Republic.
Oikonomou, V.P., Tzallas, A.T. & Fotiadis, D.I. (2007). A Kalman filter based methodology for EEG spike enhancement. *Computer Methods and Programs in Biomedicine* 85(2): 101-108.
Palu, M. (1996). Nonlinearity in normal human EEG: cycles, temporal asymmetry, nonstationarity and randomness, not chaos, Springer. 75: 389-396.
Pfurtscheller, G. (1977). Graphical display and statistical evaluation of event-related desynchronization (ERD). *Electroencephalogr Clin Neurophysiol* 43(5): 757-60.
Pfurtscheller, G. (1989). Spatiotemporal analysis of alpha frequency components with the ERD technique. *Brain Topogr* 2(1-2): 3-8.
Pfurtscheller, G. & Aranibar, A. (1977). Event-related cortical desynchronization detected by power measurements of scalp EEG. *Electroencephalogr Clin Neurophysiol* 42(6): 817-26.
Pfurtscheller, G. & Lopes da Silva, F.H. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin Neurophysiol* 110(11): 1842-57.
Pijn, J.P., Van Neerven, J., Noest, A. & Lopes da Silva, F.H. (1991). Chaos or noise in EEG signals; dependence on state and brain site. *Electroencephalogr Clin Neurophysiol* 79(5): 371-81.
Pinault, D., Vergnes, M. & Marescaux, C. (2001). Medium-voltage 5-9-Hz oscillations give rise to spike-and-wave discharges in a genetic model of absence epilepsy: in vivo dual extracellular recording of thalamic relay and reticular neurons. *Neuroscience* 105(1): 181-201.
Rauch, H., Missiles, L., Co, S. & Alto, P. (1963). Solutions to the linear smoothing problem. *IEEE Trans Automatic Control* 8(4): 371-372.
Sayadi, O. & Shamsollahi, M.B. (2008). ECG denoising and compression using a modified extended Kalman filter structure. *IEEE Trans Biomed Eng* 55(9): 2240-8.
Shumway, R.H. & Stoffer, D.S. (1982). An approach to time series smoothing and forecasting using the EM algorithm. *Journal of the Times Series Analysis* 3(4): 253-264.
Tarvainen, M.P., Hiltunen, J.K., Ranta-aho, P.O. & Karjalainen, P.A. (2004). Estimation of nonstationary EEG with Kalman smoother approach: an application to event-related synchronization (ERS). *IEEE Trans Biomed Eng* 51(3): 516-24.
Thakor, N.V. & Tong, S. (2004). Advances in quantitative electroencephalogram analysis methods. *Annu Rev Biomed Eng* 6: 453-95.
Tzallas, A.T., Karvelis, P.S., Katsis, C.D., Fotiadis, D.I., Giannopoulos, S. & Konitsiotis, S. (2006). A method for classification of transient events in EEG recordings: application to epilepsy diagnosis. *Methods Inf Med* 45(6): 610-21.
Waterhouse, E. (2003). New horizons in ambulatory electroencephalography. *IEEE Eng Med Biol Mag* 22(3): 74-80.

Wilson, S.B. & Emerson, R. (2002). Spike detection: a review and comparison of algorithms. *Clin Neurophysiol* 113(12): 1873-81.

Kalman Filter Recent Advances and Applications

Edited by Victor M. Moreno and Alberto Pigazo

ISBN 978-953-307-000-1
Hard cover, 584 pages
Publisher InTech
Published online 01, April, 2009
Published in print edition April, 2009

The aim of this book is to provide an overview of recent developments in Kalman filter theory and their applications in engineering and scientific fields. The book is divided into 24 chapters and organized in five blocks corresponding to recent advances in Kalman filtering theory, applications in medical and biological sciences, tracking and positioning systems, electrical engineering and, finally, industrial processes and communication networks.