

A square root ensemble Kalman filter application to a motor-imagery brain-computer interface

M. Kamrunnahar, and S. J. Schiff

Abstract – We here investigated a non-linear ensemble Kalman filter (SPKF) application to a motor imagery brain computer interface (BCI). A square root central difference Kalman filter (SR-CDKF) was used as an approach for brain state estimation in motor imagery task performance, using scalp electroencephalography (EEG) signals. Healthy human subjects imagined left vs. right hand movements and tongue vs. bilateral toe movements while scalp EEG signals were recorded. Offline data analysis was conducted for training the model as well as for decoding the imagery movements. Preliminary results indicate the feasibility of this approach with a decoding accuracy of 78%-90% for the hand movements and 70%-90% for the tongue-toes movements. Ongoing research includes online BCI applications of this approach as well as combined state and parameter estimation using this algorithm with different system dynamic models.

I. INTRODUCTION

In a brain computer interface (BCI) development, neuronal signals are translated into commands to build a direct interface between the brain and a device. Although invasive techniques have shown promise in the application of BCI, non-invasive scalp EEG based methods can be more easily applied. Feature extraction and pattern discrimination are commonly applied in the design of motor imagery based BCI paradigms, where subjects perform imagery tasks in response to audio-visual cues on the computer screen. Recently, there have been some model-based BCI applications where a generative model that correlates brain activity with intended tasks is developed. Martens and Leiva [1] developed such a model for decoding the visual event-related potential-based brain-computer speller. Geronimo *et al* [2] presented a simplified generative model for motor imagery BCI application and showed significantly higher BCI accuracy performance.

The Kalman filter (KF) has been applied in an encoding-decoding framework in a number of neural interface applications using intracranial signals. Wu *et al* [3-5] used the KF for neural decoding of motor cortical activity and cursor motion. In motor imagery BCI applications, the KF has been applied mainly for optimizing model parameter estimation.

Manuscript received April 15, 2011. This work was supported by the NINDS (NIH) grant K25NS061001(MK).

M. Kamrunnahar is with the Center for Neural Engineering, Dept. of Engineering Science and Mechanics, The Pennsylvania State University, University Park, PA 16803 USA (Phone: 814-865-0663, email: muk11@psu.edu).

S.J. Schiff is with the Center for Neural Engineering, Depts. of Engineering Science and Mechanics, Neurosurgery, and Physics, The Pennsylvania State University, University Park, PA 16803 USA (email: sschiff@psu.edu).

We here applied a KF as a model-based approach for brain state estimation in motor imagery task performance, using non-invasive scalp electroencephalography (EEG) signals. This approach enables encoding imagery hand, tongue, and bilateral toe movements in motor cortex and decoding that movement from recorded EEG signals. A model-based approach in BCI has the benefit of understanding the brain dynamics for further applications in BCI. The application of a Kalman Filter is expected to handle noisy brain signals effectively.

II. METHODS

A. Experimental Paradigm

Five healthy human subjects, 25- 32 years old, four males and one female, none of them under any kind of medication, participated in the motor imagery tasks. The experiments were conducted under Institutional Review Board approval at Penn State University. Each subject conducted one session of tasks that consisted of four runs, each with 40 trials. Each trial was designed as follows: the subject would be quiet and relaxed, a cross would appear on the computer screen, a left, right, up, or down arrow, depending on the task to be performed, would appear during which time the subject would imagine the task, and then both the cross and arrow would disappear to end the trial. Of the four total runs, the first two were designed for imagery of left or right hand movements and the last two runs were for imagery of tongue or bilateral toe movements. Of the 40 trials in each left-right hand movement run, 20 randomly permuted trials showed “left” arrows indicative of imagined left hand movements and the other 20 showed “right” arrows indicative of imagined right hand movements. Similarly, “up” and “down” arrows were used for tongue-toes tasks.

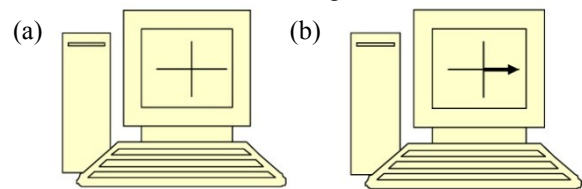


Fig. 1. (a) Fixation cross and (b) Arrow cue for motor imagery tasks.

B. Data Acquisition and Processing

Nineteen monopolar electrode positions (FP1, FP2, F7, F3, Fz, F4, F8, T7, C3, Cz, C4, T8, P7, P3, Pz, P4, P8, O1, and O2 as per the International 10-20 standard electrode locations) referenced to linked earlobe electrodes were selected for acquiring EEG under open loop conditions while the participants performed the imagery tasks. Data were passed through a fourth order band-pass Butterworth

filter of 0.5-60 Hz and sampled at 256 Hz. Data were recorded with g.tec amplifier systems [6]. The same experimental paradigm with the same montage was used in our previous work [7].

Data were epoched from 2 s before to 4 s after the presentation of each arrow cue. Recordings were visually inspected for artifacts, and by using an amplitude threshold (55 μV) criterion, trials that contained artifacts were excluded from further analysis. For each subject, the number of per class trials remaining after artifact exclusion was 70-80 out of signals recorded in 80 trials.

The Laplacian derivations [8-11] were developed for nine inner loop channels (F3, Fz, F4, C3, Cz, C4, P3, Pz, P4) using four channels surrounding the active channel for deriving the weighted average. The Laplacian is a discrete second derivative, calculated as the difference between an electrode potential and a weighted average of the surrounding electrode potentials and is commonly applied to increase the spatial resolution of EEG signals.

C. Dynamic Modeling and Kalman Filter Application

The KF is “essentially a set of mathematical equations that implement a predictor-corrector type estimator that is *optimal* in the sense that it minimizes the estimated *error covariance*—when some presumed conditions are met”[12]. Although the linear KF, in its original format introduced by Kalman [13], works for many applications, a number of nonlinear filters have been developed for non-linear processes. A KF can be applied in order to estimate states or model parameters as well as to estimate both.

We begin with a dynamic state space model (DSSM) with unobserved states x_k that evolve over time (discrete time steps designated by k) and observed outputs y_k that are conditionally independent given the states, as:

$$\begin{aligned} x_k &= f(x_{k-1}, u_k, v_k; w) \\ y_k &= g(x_k, n_k; w) \end{aligned} \quad (1)$$

where v_k and n_k are process and measurement noises respectively, and u_k are external inputs. The state transition function f and the observation function g are parameterized by w , in an augmented format.

We here used a multi layer perceptron (MLP) artificial neural network (ANN) model corresponding to each of the 9 Laplacian electrodes for motor imagery task performance, as our system dynamic model. An ANN is a computational model based on biological neural networks and consists of an interconnected group of artificial neurons [7-10]. It can be treated as non-linear statistical data modeling tools that can be used to model complex relationships between inputs and outputs or to find patterns in data. Within a layer of an ANN, a single neuron makes an operation of a weighted sum of the incoming signals and a bias term, fed through an activation function and resulting in the output value of that neuron (Fig. 2). A popular activation term in dynamic modeling is a non-linear hyperbolic tangent (tanh) function. There are mainly two learning processes of an ANN, namely supervised learning and unsupervised learning. In supervised learning, for a given set of input-output pairs, the aim is to find a function that maps the inputs to the outputs by

minimizing a cost function that is related to the mismatch between the mapped outputs and the target outputs. A commonly used objective is to minimize the average squared error between the network's outputs and the target values. When one tries to minimize this mean-squared error cost function, one obtains the well-known backpropagation algorithm for training neural networks. Typically, the objective is to make outputs o_i and targets t_i identical for $i = 1, \dots, p$ for p variables, by using a learning algorithm. More precisely, the objective is to minimize the error function of the network, defined as

$$E = \frac{1}{2} \sum_{i=1}^p \|o_i - t_i\|^2 \quad (7)$$

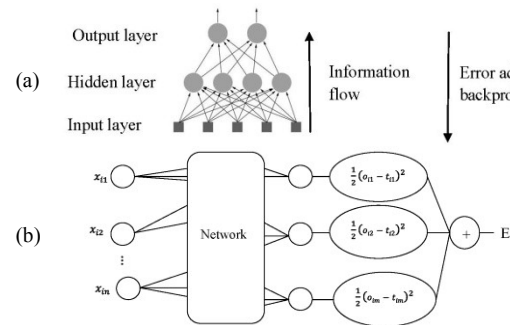


Fig. 2. (a) A simple example of an artificial neural network (ANN), (b) an example of an extended network to estimate the cost functions in ANN backpropagation.

Fig. 2 shows an example of an ANN model and the backpropagation framework. We used a 6-4-1 structure for the model, a hyperbolic tangent (tanh) activation function in the hidden layer, and a linear activation function for the output layer. The preliminary model structure and function were chosen here in an ad-hoc basis with a goal to introduce different model structures and compare them in the ongoing study. The ANN parameters can be estimated by model fitting using backpropagation to the EEG data corresponding to each electrode for the motor imagery task performance. The single channels were used as an initial modeling approach with the goal to apply a multi-variable KF design in the ongoing study.

The recursive form of the optimal Kalman update of the conditional mean of the state random variable, $\hat{x}_k = E[x_k | y_{1:k}]$ and its covariance, P_{x_k} is written as:

$$\begin{aligned} \hat{x}_k &= (\text{prediction of } x_k) + K_k (y_k - (\text{prediction of } y_k)) \\ &= \hat{x}_k^- + K_k (y_k - \hat{y}_k^-) \end{aligned} \quad (2)$$

$$P_{x_k} = P_{x_k}^- + K_k P_{y_k} K_k^T \quad (3)$$

The optimal terms here are:

$$\hat{x}_k^- = E[f(x_{k-1}, u_k, v_{k-1})] \quad (4)$$

$$\hat{y}_k^- = E[g(x_k^-, n_k)] \quad (5)$$

$$\begin{aligned} K_k &= E[(x_k - \hat{x}_k^-)(y_k - \hat{y}_k^-)^T] E[(y_k - \hat{y}_k^-)(y_k - \hat{y}_k^-)^T]^{-1} \\ &= P_{x_k \tilde{y}_k} P_{\tilde{y}_k}^{-1} \end{aligned} \quad (6)$$

The *Sigma-point Kalman filters* (SPKF) are a group of Kalman filter algorithms that fall into a general deterministic sampling framework known as the *sigma-point approach* for the calculation of the posterior mean and covariance of the

pertinent Gaussian approximate densities in the Kalman framework recursion [14, 15]. The SPKF was introduced as a better alternative to the extended Kalman filter (EKF) for Gaussian approximate probabilistic inference in general nonlinear DSSMs. The underlying unifying sigma-point approach that is common to all SPKFs, was introduced as a method to calculate the statistics of a random variable that undergoes a nonlinear transformation. These calculations form the core of the optimal Kalman time and measurement update equations, which are simply the original (optimal) recursive Bayesian estimation integral equations recast under a Gaussian assumption.

The unscented KF (UKF) [16-18] and the central difference KF (CDKF) [14, 15], although derived from different starting assumptions, both employ the sigma-point approach as their core algorithmic components for calculating the posterior Gaussian statistics necessary for Gaussian approximate inference in a Kalman framework.

In practice, there are many situations where linearizing the underlying model equations, or linearizing the control law, as is done in an extended Kalman Filter (EKF) application, is not suitable. The central difference filter (CDF) of Ito and Xiong [19] and the divided difference filter (DDF) of Nørgaard [20] are based on polynomial approximations of the nonlinear transformations obtained with a multidimensional extension of *Stirling's polynomial interpolation formula* [21, 22]. The CDKF (either CDF or DDF) is a derivativeless Kalman filter for nonlinear estimation, based on polynomial interpolation and approximation of the derivatives using difference equations (central or divided, as are termed). A particularly useful idea of the CDKF is to directly update the *Cholesky factors* of the covariance matrices, which are used in the *a-priori* and *a-posteriori* updates of the outputs and states. Nørgaard *et al.* [20] showed how the CDKF has a slightly smaller absolute error (compared to the UKF) in the fourth order term and also guarantees positive semi-definiteness of the posterior covariance. The CDKF and UKF perform equally well with negligible difference in estimation accuracy and have been found to be superior to the EKF.

The square-root form of CDKF increases the numerical robustness of the filter, as well as reduces the computational cost for certain DSSMs. This form propagates and updates the square-root of the state covariance directly in *Cholesky* factored form, using the *sigma-point* approach and linear algebra techniques such as *QR decomposition*, *Cholesky factor updating* and *pivot-based least squares* [14]. The choice of a SR-UKF or SR-CDKF to be used is a matter of implementational choice. However, there is one advantage CDKF has over the UKF: CDKF uses only a single scalar scaling parameter, the central difference interval size h , as opposed to three parameters in UKF [14]. This h parameter determines the *spread* of the sigma-points around the prior mean. For Gaussian random variables, the optimal value of h is $\sqrt{3}$.

We here applied the square root CDKF (SR-CDKF) algorithm which lies in the family of ensemble KF or sigma point KF developed by Van der Merwe [14, 23], for the state estimation of brain dynamics corresponding to motor

imagery. The essential steps and the detailed derivations of SR-CDKF algorithm can be found in the literature.

We used a state dimension of order six, Gaussian noise with 5dB of signal to noise ratio (SNR), zero initial state value, and initial noise variance to be unity. We used the ReBEL toolkit [24] for the KF application.

D. Motor Imagery Task Performance

For the evaluation of the decoding performance of the KF approach, we applied each left, right, tongue, or toes state model to evaluate motor imagery task performance using offline EEG signals. We calculated the distance of the signals from the models corresponding to each of the tasks. We assigned the signals to the model having the shortest distance as the intended task. First order norm of the residual (error between the model and the EEG signal) was considered as the measure of the distance.

III. RESULTS AND DISCUSSION

Fig. 3 presents the proposed SR-CDKF results for decoding motor imagery task performance for Subjects 1-5. The results, with the limited number of subjects under this study, show a decoding accuracy of 78%-90% for the hand movements and 70%-90% for the tongue-toes movements, which are very promising to be applied for an online BCI application that is under study.

Classification accuracies for the four imagery tasks (left vs. right hand and tongue vs. toes movements) by the state estimation model for five subjects were compared with the accuracies using the commonly applied discriminative approach using spectral powers in μ (8-12 Hz) and β (14-20 Hz) frequency bands as the features along with linear discriminant analysis (LDA) as a classification method [25-29]. The best (i.e. highest) classification accuracies (among the four time segments) by each method for all five subjects are shown in Fig. 4.

An analysis of variance (ANOVA) check shows that the classification accuracy using the Kalman model was significantly better (p-value 0.006) than the accuracy using the μ - β (MB) power spectra features.

The advantage of this approach lies in that this model-based approach with dynamic state estimation will help understand the brain dynamics better as well as effectively handle noisy brain signals through the application of a Kalman Filter.

IV. SUMMARY

A square root central difference Kalman filter (SR-CDKF), a member of non-linear sigma point Kalman filters (SPKF) was applied to a brain computer interface (BCI) for brain state estimation in motor imagery task performance, using scalp electroencephalography (EEG) signals. For left vs. right hand movements and tongue vs. bilateral toe movements, preliminary results of offline data analysis indicate the feasibility of this approach with a decoding accuracy of 78%-90% for the hand movements and 70%-

90% for the tongue-toes movements, for the limited number of subjects in this study. Comparison of the imagery classification accuracy using the proposed approach with the accuracy using commonly applied mu-beta (MB) power spectra features shows that the proposed algorithm performance is significantly better than the discrimination approach using MB features. Ongoing research includes online BCI applications of this approach as well as combined state and parameter estimation using this algorithm with different system dynamic models.

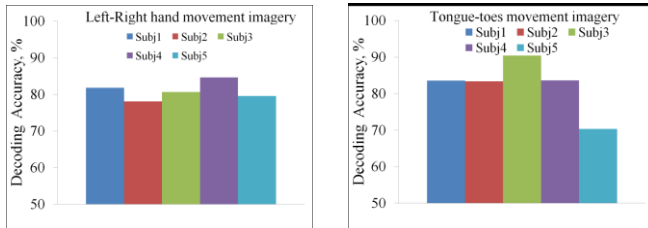


Fig. 3. Decoding accuracy of motor imagery tasks using the square root central difference Kalman Filter estimation algorithm for five subjects. Two pairs of tasks, left vs. right hand movements and tongue vs. bilateral toe movements are shown.

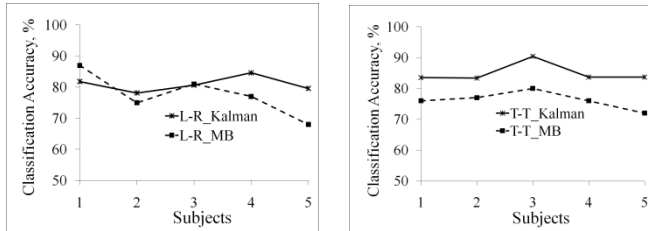


Fig. 4. Classification accuracy of motor imagery tasks using the square root central difference Kalman Filter estimation algorithm in comparison to the accuracy using mu-beta power spectra features. Two pairs of tasks, left vs. right hand movements and tongue vs. bilateral toe movements are shown.

REFERENCES

- [1] S. M. M. Martens, and J. M. Leiva, "A generative model approach for decoding in the visual event-related potential-based brain-computer interface speller," *Journal of Neural Engineering*, vol. 7, no. 2, Apr, 2010.
- [2] A. Geronimo, M. Kamrunnihar, and S. Schiff, "A generative model approach to motor-imagery brain-computer interfacing."
- [3] W. Wu, M. J. Black, Y. Gao et al., "Neural decoding of cursor motion using a Kalman filter," *Advances in Neural Information Processing Systems*, vol. 15, pp. 133-140, 2003.
- [4] W. Wu, Y. Gao, E. Bienenstock et al., "Bayesian Population Decoding of Motor Cortical Activity Using a Kalman Filter," *Neural Computation*, vol. 18, pp. 80-118, 2006.
- [5] W. Wu, A. Shaikhouni, J. R. Donoghue et al., "Closed-loop neural control of cursor motion using a Kalman filter." pp. 4126-4129.
- [6] "g.tec medical engineering", <http://gtec.at/Products/Hardware-and-Accessories/g.USBamp-Specs-Features>, 1999-2006.
- [7] M. Kamrunnihar, N. Dias, and S. Schiff, "Toward a Model-Based Predictive Controller Design in Brain-Computer Interfaces," *Annals of Biomedical Engineering*, pp. 1-11, 2011.
- [8] P. L. Nunez, and R. Srinivasan, *Electric Fields of the Brain: The Neurophysics of EEG*, 2nd ed., New York: Oxford University Press, 2006.
- [9] P. L. Nunez, *Electric fields of the brain*: Oxford University Press, 1981.
- [10] R. Srinivasan, "Methods to Improve the Spatial Resolution of EEG," *International Journal of Bioelectromagnetism*, vol. 1, no. 1, pp. 102-111, 1999.
- [11] R. Srinivasan, W. R. Winter, and P. L. Nunez, "Source analysis of EEG oscillations using high-resolution EEG and MEG," *Event-Related Dynamics of Brain Oscillations; Progress in Brain Research*, C. Neuper and W. Klimesch, eds., pp. 29-42: Elsevier, 2006.
- [12] G. Welch, and G. Bishop. "An Introduction to the Kalman Filter," http://www.cs.unc.edu/~tracker/media/pdf/SIGGRAPH2001_CoursePack_08.pdf.
- [13] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Transactions of the ASME - Journal of Basic Engineering*, vol. 82, pp. 35-45, 1960.
- [14] R. Van der Merwe, "Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models," PhD, OGI School of Science & Engineering, Oregon Health & Science University, Portland, OR, 2004.
- [15] R. Van der Merwe, and E. Wan, "Efficient Derivative-Free Kalman Filters for Online Learning." pp. 205-210.
- [16] S. Julier, and J. Uhlmann, "A new extension of the Kalman filter to nonlinear systems."
- [17] S. J. Julier, and J. K. Uhlmann, *A General Method for Approximating Nonlinear Transformations of Probability Distributions*, Dept. of Engineering Science, University of Oxford, Nov 1996.
- [18] S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte, "A new approach for filtering nonlinear systems." pp. 1628-1632 vol.3.
- [19] K. Ito, and K. Q. Xiong, "Gaussian filters for nonlinear filtering problems," *Ieee Transactions on Automatic Control*, vol. 45, no. 5, pp. 910-927, May, 2000.
- [20] M. Nørgaard, N. K. Poulsen, and O. Ravn, "New developments in state estimation for nonlinear systems," *Automatica*, vol. 36, no. 11, pp. 1627-1638, 2000.
- [21] C. E. Fröberg, *Introduction to numerical analysis*, Reading, MA: Addison-Wesley, 1970.
- [22] J. F. Steffensen, *Interpolation*, Baltimore, MD: Williams & Wilkins, 1927.
- [23] R. Van der Merwe, and E. Wan, "The Square-Root Unscented Kalman Filter for State- and Parameter-Estimation." pp. 3461-3464.
- [24] R. Van der Merwe, and E. A. Wan, "ReBEL Toolkit in MATLAB," School of Science & Engineering at Oregon Health & Science University (OHSU); <http://choosh.csee.ogi.edu/rebel/>.
- [25] F. Lotte, M. Congedo, A. Lecuyer et al., "A review of classification algorithms for EEG-based brain-computer interfaces," *Journal of Neural Engineering*, vol. 4, no. 2, pp. R1-R13, Jun, 2007.
- [26] D. J. McFarland, D. J. Krusienski, and J. R. Wolpaw, "Brain-computer interface signal processing at the Wadsworth Center: mu and sensorimotor beta rhythms," *Prog Brain Res*, vol. 159, pp. 411-9, 2006.
- [27] J. Millan, M. Franze, J. Mourino et al., "Relevant EEG features for the classification of spontaneous motor-related tasks," *Biol Cybern*, vol. 86, no. 2, pp. 89-95, Feb, 2002.
- [28] C. Neuper, and G. Pfurtscheller, "ERD and ERS during motor behavior in man," *Journal of Psychophysiology*, vol. 11, no. 4, pp. 375-375, 1997.
- [29] M. Pregenzer, and G. Pfurtscheller, "Frequency component selection for an EEG-based brain to computer interface," *IEEE Trans Rehabil Eng*, vol. 7, no. 4, pp. 413-9, Dec, 1999.