33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

EEG Signal Classification Using Time-Varying Autoregressive Models and Common Spatial Patterns

D. Gutiérrez, *Member, IEEE*, and R. Salazar-Varas

Abstract— **The performance of EEG signal classification methods based on Common Spatial Patterns (CSP) depends on the operational frequency bands of the events to be discriminated. This problem has been recently addressed by using a sub-band decomposition of the EEG signals through filter banks. Even though this approach has proven effective, the performance still depends on the number of filters that are stacked and the criteria used to determine their cutoff frequencies. Therefore, we propose an alternative approach based on an eigenstructure decomposition of the signals' time-varying autoregressive (TVAR) models. The eigen-based decomposition of the TVAR representation allows for subject-specific estimation of the principal time-varying frequencies, then such principal eigencomponents can be used in the traditional CSP-based classification. A series of simulations show that the proposed classification scheme can achieve high classification rates under realistic conditions, such as low signal-to-noise ratio (SNR), a reduced number of training experiments, and a reduced number of sensors used in the measurements.**

I. Introduction

Brain-computer interfaces (BCI) based on electroencephalographic (EEG) data rely on accurate classification methods in order to use brain's electrical activity to control computerized devices in real-time [1]. In the last few years, many classifiers have been proposed in the literature, some of which report excellent performance in classifying different motor and cognitive tasks for BCI applications [2]. Frequency-domain characteristics (such as changes in the mu, beta, or gamma rhythms) and time-varying characteristics (such as movements-related potentials) are the most common extracted features in BCI.

Even though the common spatial patterns (CSP) method is not, strictly speaking, a classification but a signal enhancement method, it has been used in multiple BCI applications by itself and in combination with other processing tools. In the CSP method, optimal spatial filters are constructed such that each filter enhances the variance of one feature of interest while blocking the other features [3]. CSP has been used in many BCI applications for the discrimination of motor imaginary tasks (see e.g., [4], [5]). However, its performance is dependent on its operational frequency band, then BCI systems relying on CSP-feature classification generally yield poor accuracies when the EEG measurements are either unfiltered or have been filtered with an inappropriately selected frequency range [6]. Hence, setting a broad frequency range or manually selecting a subject-specific frequency range is a common practice with the CSP algorithm [7].

The problem of manually selecting the operational subject-specific frequency band of the CSP has been addressed in several ways. Two of the most recent approaches rely on the construction of filter banks that decomposes the EEG measurements into multiple sub-bands, then the CSP algorithm is used on each of the sub-bands. In [6], the construction of the sub-bands is performed by a Gabor filter bank, while in [8] a zero-phase Chebyshev Type II infinite impulse response filter bank is used. In both cases, the performance of the method relies on the number of filters that are stacked and the criteria used to determine their cutoff frequencies and the overlapping between them (if any).

In order to avoid the use of fixed frequency bands, we propose to substitute the frequency bank-based sub-band decomposition by an eigenstructure decomposition. Therefore, we present a method where the EEG signals are decomposed by means of non-stationary time series models. Specifically, time-varying autoregressive (TVAR) models are used to obtain a representation of the time-frequency structure of the EEG signals. TVAR models were first introduced in [9] and have thereafter been partially reformulated and applied to various fields such as seismology, geology, and economics. In the case of EEG studies, TVAR models have been applied to the analysis of non-stationary human seizure EEG data [10]. In this paper, the assessment of changes over time in the TVAR models obtained from the EEG data allows for the identification of time-varying principal frequencies (most likely related to the physiological events of interest) from which the most significant ones, in an eigenstructure sense, are then used in the traditional CSP-based classification.

In Section II, we present the proposed classification scheme, from which the basis of TVAR modeling is reviewed, as well as the dynamic eigenstructure decomposition of the corresponding TVAR evolution matrix. Section II includes a brief description of the CSP algorithm as well. In Section III, we show the applicability of our methods through numerical examples using simulated EEG data. Finally, in Section IV, we discuss the results, limitations, and future work.

II. Methods

In this section, we describe each of the steps involved in our proposed method for EEG signal feature extraction, which is illustrated in Figure (1).

This work was supported by the National Council of Science and Technology (CONACyT-Mexico) under Grant 101374.

D. Gutiérrez and R. Salazar-Varas are with the Center of Research and Advanced Studies (Cinvestav) at Monterrey, 66600, Apodaca, Mexico dgtz@ieee.org, rsalazar@cinvestav.mx

978-1-4244-4122-8/11/$26.00 ©2011 IEEE

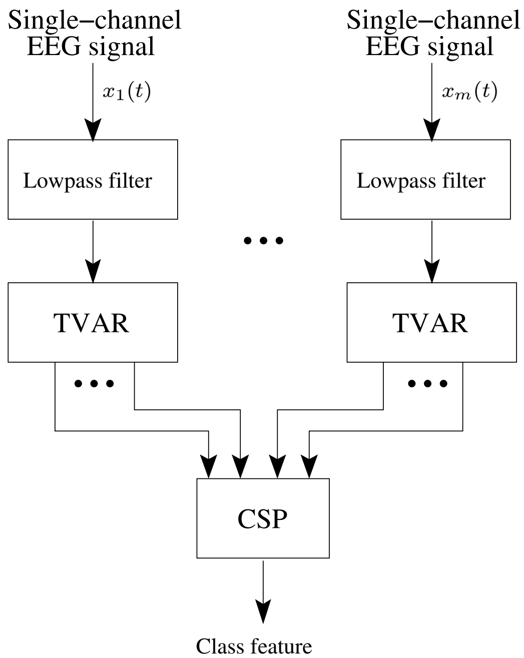

Fig. 1. Proposed processing scheme.

A. TVAR decomposition

In the standard autoregressive framework, a discretely sampled EEG signal is modeled by representing the voltage level at time t as a linear combination of voltage levels at times $t-1, t-2, \dots, t-p$, (where $p > 0$ is the maximum time lag) plus a random component (driving noise or "innovation"). The relationship is assumed to be fixed over time, then the coefficients defining the linear combination are constant for the entire period of recording. In TVAR models, those coefficients are allowed to vary over time then they can adapt to changes evidenced in the series. In particular, such models can respond to and adequately capture the forms of frequency changes seen on EEG oscillations. Furthermore, if a multichannel scheme is considered, then a TVAR approach can be defined where each signal sample is defined versus both its previous samples and the previous samples of other channels [11].

Under those conditions, let us define $x_m(t)$ as the time series resulting of the EEG measurement at sensor $m = 1, 2, \dots, M$ and at time $t = 1, 2, \dots, N$. The corresponding TVAR model of order p is given by

$$x_m(t) = \sum_{i=1}^{p} \varphi_{m,i}(t) x_m(t-i) + n_m(t), \qquad (1)$$

where $\varphi_{m,i}(t)$ are the time-varying coefficients of the model, which are often calculated using the Levinson-Wiggins-Robinson (LWR) algorithm [12], and $n_m(t)$ represents the noise input to channel m. Then, the TVAR coefficients are arranged into a matrix $G_m(t)$ as [13]

$$G_m(t) = \begin{bmatrix} \varphi_{m,1}(t) & \varphi_{m,2}(t) & \cdots & \varphi_{m,p-1}(t) & \varphi_{m,p}(t) \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}. \qquad (2)$$

Once the EEG series are modeled via TVAR models, the focus is on exploring the time-frequency structure of the latent processes underlying the signals using a dynamic model decomposition based on the eigenstructure of $G_m(t)$. Such decomposition is given by

$$x_m(t) = \sum_{i=1}^{p_z} z_{m,i}(t) + \sum_{i=1}^{p_y} y_{m,i}(t), \qquad (3)$$

where p_z is the number of pairs of complex characteristic roots of the polynomial defined by the eigendecomposition of $G_m(t)$, and p_y is the number of real characteristic roots, such that $2p_z + p_y = p$. Each $z_{m,i}(t)$ leads to a time-domain analysis of the time-frequency structure of $x_m(t)$ through an exploration of their frequencies $\omega_{m,i}(t)$ and moduli $r_{m,i}(t)$. Assessment of changes over time in $\omega_{m,i}(t)$ and $r_{m,i}(t)$ are an approach to explore time-variation in spectral density functions, and the corresponding decomposition analysis represents a form of spectral decomposition in the time domain [10].

In order to perform the decomposition in (3), we used a software tool (freely available at http://www.stat.duke.edu/research/software/west/tvar.html), which implements the sequential updating and retrospective smoothing algorithms for calculating the TVAR models. In addition, this software supports computation of the latent processes in the decomposition of the series and the corresponding characteristic wavelengths, moduli and amplitudes related to such processes at each time [14]. In our case, this software implementation decomposes $x_m(t)$ into p_z time-series from which those with the most significant frequency content (based on the eigenvalue assessment of $\omega_{m,i}(t)$) can be used as input in the CSP algorithm. The selected components are then arranged into a spatio-temporal matrix X of size $p_o M \times N$, where $p_o \leq p_z$ is the number of components selected from the decomposition.

B. CSP

The normalized spatial covariance of X can be obtained from

$$C = \frac{XX^T}{\text{tr}\{XX^T\}}, \qquad (4)$$

where $\text{tr}\{\cdot\}$ denotes the trace. Considering the case of two-class discrimination (e.g., contra-lateral brain activity from either left or right motor cortex), a *composite* spatial covariance can be computed as $C_c = \overline{C}_1 + \overline{C}_2$, where $\overline{C}_l$ is the average spatial covariance obtained from independent trials of class $l = 1, 2$. Furthermore, C_c can be factored as $C_c = U_c D_c U_c^T$, where U_c is the matrix of eigenvectors and D_c is the diagonal matrix of eigenvalues arranged in descending order. Hence, the whitening transformation

$$P = D_c^{-1/2} U_c^T \quad (5)$$

will equalize the variances in the space spanned by U_c, i.e., all eigenvalues of PU_cP^T will be equal to one. Therefore, the transformation $S_l = P\overline{C}_l P^T$ for $l = 1, 2$, will produce matrices that share common eigenvectors:

$$\text{if } S_1 = BD_1B^T \text{then } S_2 = BD_2B^T, \quad (6)$$

where $D_1 + D_2 = I$. Therefore, if we define the projection matrix $W = (P^T B_o)^T$ such that B_o contains the first and last eigenvectors in B (which correspond to the largest eigenvalues for each class), then the filter W will produce feature vectors that are optimal for discriminating between the two classes in the least square sense [4]:

$$Z = WX. \quad (7)$$

Under these conditions, the columns of W^{-1} become the CSPs. Hence, Z is computed by (7) and from its rows $\boldsymbol{z}_k^T$, for $k = 1, 2, \ldots, p_o M$, we construct the feature vector $\boldsymbol{v} = [v_1, \ldots, v_{p_oM}]^T$ to be used in the classification:

$$v_k = \log \left[\frac{\operatorname{var}(\boldsymbol{z}_k^T)}{\sum_{j=1}^{p_oM} \operatorname{var}(\boldsymbol{z}_j^T)} \right], \quad (8)$$

where var$(\cdot)$ denotes the variance of the vector's elements.

III. Numerical Examples

We performed a series of numerical experiments for simulated EEG data corresponding to contra-lateral activation of either the left or right motor cortex. The data was generated based on the method of phase-resetting [16] whose computer implementation is freely available at http://www.cs.bris.ac.uk/ rafal/phasereset/. Further details on the use of this computer tool for EEG data generation can be found in [17].

We simulated measurements for each class (left or right brain hemisphere) using an array of $M = 31$ sensors, while sampling at a frequency of 250 Hz. The data was chosen to simulate motor-related activation of the mu rhythm, which has been extensively examined in BCI applications [18]. Furthermore, the data was added with different levels of uncorrelated noise in order to produce mean signal-to-noise ratios (SNRs) of -10 and -4.5 dB (i.e., with high and moderate noise conditions, respectively). Note that the SNR is defined as the ratio (in decibels) of the Frobenious norm of the signal data matrix to that of the noise matrix.

The process of adding noise to the simulated data was repeated with independent noise realizations to obtain 150 trials corresponding to left-side activity and 150 trials of right-side activity, where each trial had $N = 200$ time samples. From each side, 100 trials were destined to be classified (testing data) and 50 trials were used as training data (i.e., to compute $\overline{C}_1$ and $\overline{C}_2$).

Under those conditions, we evaluated the performance of the processing method described in Section II (starting with a lowpass filter with cutoff frequency of 20 Hz) for the case when the feature vectors in (8) were discriminated using a Mahalanobis distance-based classifier [19] and the evaluation was made in terms of its receiver operating characteristics (ROC) curve. This procedure has been previously used as a generalized evaluation framework in BCI applications (see [20]) . Furthermore, the proposed method was compared against the filter bank common spatial pattern (FBCSP) method described in [8]. In order to provide an evaluation closer to real-life conditions, we evaluated the performance for the cases when, from the original $M = 31$ sensors, only a subset of them was used. In the first case, only six sensors were used which, according to the international 10-20 sensor arrangement, were: C3, C4, CP3, CP4, FC3, and FC4. The second case corresponded to ten sensors: those previously selected plus TP7, TP8, O1, and O2.

Our results are shown in Figure (2), for the case of six channels, and Figure (3) for the case when 10 channels were used. It is clear that, under all conditions, the proposed TVAR-CSP classification procedure has better performance than the filter bank-based approach. The improvement in the performance is more noticeable under high noise conditions (SNR=-10 dB) where the area under the ROC curve for the TVAR-CSP is larger than the one for the filter bank. We can also note that increasing the number of sensors from six to ten while maintaining the same number of training trials (K = 50) affected the performance of both classifiers. This is due to the fact that more data is necessary to warranty that the matrices $\overline{C}_1$ and $\overline{C}_2$ are consistent estimates of the spatial covariance. Nevertheless, the TVAR-CSP approach seems to be more robust to this condition as its performance did not decrease significantly. Finally, we have to mention that a test for SNR=0 dB was also performed and the result was a perfect classification for both the TVAR-CSP and the filter bank-based approach. However, we have to remember that the TVAR-CSP method has the advantage of not requiring any *a-priori* selection of frequency bands.

IV. Concluding Remarks

We presented a classification method based on modeling the EEG signals through TVAR models as preliminary step of the CSP method. Instead of directly decomposing the EEG signals, we studied the dynamics of their corresponding TVAR model coefficients using an eigenstructure time-frequency analysis. By selecting the most significant components and using them as input to the CSP method, we were able to eliminate the need of selecting fixed frequency bands as in the case of filter bank-based approaches.

We evaluated the proposed method under realistic conditions, such as low SNR, a reduced numbers of trials used for training data, and a few sensors used from the measurements. Our results showed that the TVAR-CSP method can achieve high rates of correct classification under those conditions, and outperformed the filter bank-based approach when tested with realistically simulated EEG data. Future work will include a more intensive experimentation using real EEG data corresponding to tasks commonly used in BCI applications. Also, in terms of the implementation, future work will evaluate the classification performance as a function of the position of the sensors, then optimal array of sensors may be found for a given BCI application.

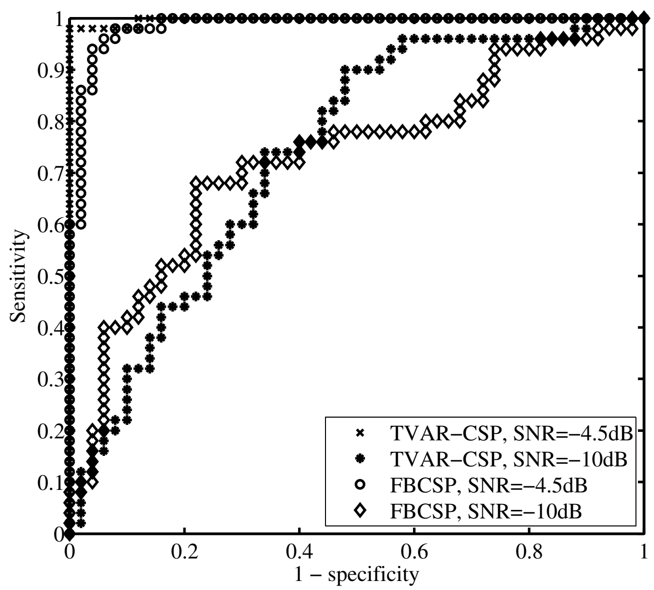

Fig. 2. ROC curves for the EEG classification methods using six channels under different noise conditions.

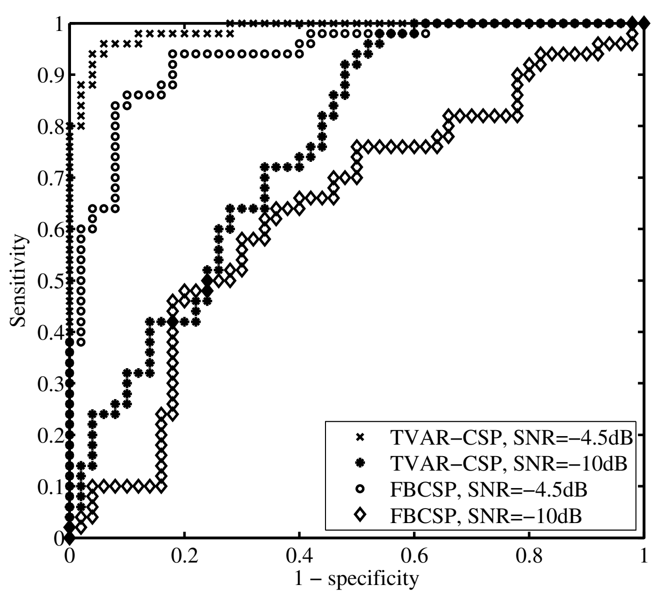

Fig. 3. ROC curves for the EEG classification methods using ten channels under different noise conditions.

REFERENCES

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, pp. 767-791, 2002.

[2] A. Bashashati, M. Fatourechi, R. K. Ward, and G. E. Birch, "A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals," *Journal of Neural Engineering*, vol. 4, pp. R32-R57, 2007.

[3] Z. J. Koles, "The quantitative extraction and topographic mapping of the abnormal components in the clinical EEG," *Electroencephalography and Clinical Neurophysiology*, vol. 79, pp. 440-447, 1991.

[4] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imaging hand movement," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, pp. 441-446, 2000.

[5] G. Pfurtscheller and C. Neuper, "Motor imagery and direct brain computer communication," *Proceedings of the IEEE*, vol. 89, no. 7, pp. 539-550, 2001.

[6] Q. Novi, C. Guan, T. H. Dat, and P. Xue, "Sub-band common spatial pattern (SBCSP) for brain-computer interface," in *3rd International IEEE/EMBS Conference on Neural Engineering*, 2007, pp. 204-207.

[7] G. Dornhege, B. Blankertz, M. Krauledat, F. Losch, G. Curio, and K. R. Muller, "Combined optimization of spatial and temporal filters for improving brain-computer interfacing," *IEEE Transactions on Biomedical Engineering*, vol. 53 no. 11, pp. 2274-2281, 2006.

[8] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, "Filter bank common spatial pattern (FBCSP) in brain-computer interface," in *IEEE International Joint Conference on Neural Networks*, 2008, pp. 2390-2397.

[9] T. S. Rao, "The fitting of non-stationary signals," *Journal of the Royal Statistical Society*, vol. B32, pp. 312-322, 1970.

[10] A. D. Krystal, R. Prado, and M. West, "New methods of time series analysis of non-stationary EEG data: eigenstructure decompositions of time varying autoregressions," *Clinical Neurophysiology*, vol. 110, pp. 2197-2206, 1999.

[11] S. Sanei and J. A. Chambers, *EEG Signal Processing*, John Wiley & Sons, Ltd., New Jersey, 2007.

[12] M. Morf, A. Vieria, D. Lee, and T. Kailath, "Recursive multichannel maximum entropy spectral estimation," *IEEE Transactions on Geoscience Electronics*, vol. 16, pp. 85-94, 1978.

[13] R. Prado and M. West, "Exploratory modelling of multiple non-stationary time series: latent process structure and decompositions," in *Modelling longitudinal and spatially correlated data*, Springer Ed., New York, 1997.

[14] R. Prado, "Latent structure in non-stationary time series," PhD Thesis, Duke University, Durham, NC, 1998.

[15] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K. -R. Müller, "Optimizing spatial filters for robust EEG single-trial analysis," *IEEE Signal Processing Magazine*, vol. 25, no. 1, pp. 41-56, 2008.

[16] V. Mäkinen, H. Tiitinen, and P. May, "Auditory event-related responses are generated independently of ongoing brain activity," *NeuroImage*, vol. 24, no. 4, pp. 961 - 968, 2005.

[17] N. Yeung, R. Bogacz, C. B. Holroyd, and J. D. Cohen, "Detection of synchronized oscillations in the electroencephalogram: an evaluation of methods", *Psychophysiology*, vol. 41, pp. 822-832, 2004.

[18] J. A. Pineda, B. Z. Allison, and A. Vankov, "The effects of self-movement, observation, and imagination on μ rhythms and readiness potentials (RPs): toward a braincomputer interface (BCI)," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 219-222, 2000.

[19] F. Babiloni, L. Bianchi, F. Semeraro, J. R. Millán, J. Mourino, A. Cattini, S. Salinari, M. G. Marciani, and F. Cincotti, "Mahalanobis distance-based classifiers are able to recognize EEG patterns by using a few EEG electrodes," in *Proceedings of the 23rd Annual Conference of the IEEE Engineering in Medicine and Biology Society*, Istanbul, Turkey, 2001, pp. 651-654.

[20] D. Gutiérrez and D. I. Escalona-Vargas, "EEG data classification through signal spatial redistribution and optimized linear discriminants," *Computer Methods and Programs in Biomedicine*, vol. 97, no. 1, pp. 39-47, 2010.