TIME-FREQUENCY-SPACE KERNEL FOR SINGLE EEG-TRIAL CLASSIFICATION

Gary Garcia Molina and Touradj Ebrahimi

Swiss Federal Institute of Technology EPFL, CH 1015 Lausanne Switzerland
gary.garcia@epfl.ch, touradj.ebrahimi@epfl.ch

ABSTRACT

In the framework of the research on Brain-Computer Interface systems, the classification of single EEG trials occupies a central place. In this paper we propose a technique of classification consisting on the analysis of EEG from a joint time-frequency and space point of view.

1. INTRODUCTION

Research on Human-Computer Interfaces (HCI) for disabled people has lead to the so called Brain-Computer Interface (BCI) systems that use brain activity for communication purposes [1]. When the brain activity is monitored through electroencephalogram measurements (EEG) one has an EEG-based BCI, henceforth called simply BCI.

In this research, the EEG signals are measured at the scalp by affixing an array of electrodes according to the 10-20 international system and with reference to digitally linked ears [9][2]. An EEG signal is therefore multivariate.

Among the commonly used EEG signals in BCI systems, we focused our attention in the spontaneous signals [3]. These signals are spontaneous because they do not constitute the response to a particular stimulus.

A BCI system involves two entities: a human subject and a machine. The subject performs a mental activity (MA) to control a machine action. The MAs are uniquely characterized by the presence of patterns in the EEG signals. A time segment of EEG (EEG-trial) is therefore classified in order to determine the MA that originated it.

The classification method for EEG-trials constitutes a very important part of a BCI system. In this paper, we present a classification algorithm relying on the analysis of the joint time, frequency and space correlations.

The application that motivated our research was the design of an immersive 3D environment where people could interact, between themselves and the environment, by merely thinking.

2 CLASSIFICATION METHOD

The single EEG-trial classification problem can be stated as follows.

Given a training set Υ of labeled EEG-trials $\gamma = \left\{ S_{w_1}^q(t), S_{w_2}^q(t), ..., S_{w_W}^q(t) \; ; \; q = 1, ..., Q \right\}$; where W is the number of MAs (classes) Q the number of labeled trials per class, and $S_{w_k}^q(t)$ the q^{th} trial belonging to class w_k. We wish to characterize each class by a model so that we can compare an unknown trial $S(t)$ to each class-model and assign $S(t)$ to its most likely class.

In order to build such a model we chose to analyze the EEG trials with respect to their correlative time-frequency representation (CTFR).

The CTFRs can provide a measure of the interaction strength between groups of neurons as a function of the time and frequency. Precedent studies emphasized the importance of these parameters for characterizing the brain activity [4]. The CFTRS are commonly characterized by the ambiguity function [5].

The ambiguity function of a multivariate signal $S(t) = \left[s_1(t) \quad s_2(t) \quad ... \quad s_N(t) \right]^t$ is defined in Equation 1. The signals $s_1(t), \quad s_2(t), \quad ... \quad s_N(t)$ are the univariate components (also called spatial components) of $S(t)$.

$$A_S(\theta, \tau) = \int S\left(t + \frac{\tau}{2}\right) \cdot S^H\left(t - \frac{\tau}{2}\right) e^{j\theta t} dt \quad (1)$$

In Equation 1, ^{H} stands for the conjugate transpose and. θ, τ are the frequency and time lags respectively. The Equation 1 can be written in a matrix form (Eq. 2).

$$\begin{aligned} A_S(\theta, \tau) &= \left[A_{kl}(\theta, \tau) \right] \qquad 1 \le k, l \le N \\ A_{kl}(\theta, \tau) &= \int s_k\left(t + \frac{\tau}{2}\right) s_l{*}\left(t - \frac{\tau}{2}\right) e^{j\theta t} dt \end{aligned} \quad (2)$$

The diagonal terms of $A_S(\theta, \tau)$ are called the auto ambiguity functions and the off diagonal terms are the cross-ambiguity functions.

In [6] L. Cohen defined the characteristic function in the univariate case as the product between the ambiguity function and a two-dimensional function called the kernel. In an analog way we can define the multivariate characteristic function in a matrix form as follows.

$$M_S(\theta, \tau) = \left[\phi_{kl}(\theta, \tau) A_{kl}(\theta, \tau) \right] \qquad 1 \le k, l \le N \quad (3)$$

The two-dimensional functions $\phi_{kl}(\theta,\tau)$ are the kernels and there are N^2 of these kernels.

The results in the univariate case [7] show that it is possible to design a kernel that is optimized for the classification in the $\theta-\tau$ plane. The proposed kernel in [7] has a radially gaussian shape and can be written as shown in Eq. 4.

$$\phi(\theta,\tau)=e^{-\frac{\rho^2}{2\sigma(\varphi)^2}}$$
$$\rho^2=\theta^2+\tau^2;\ \varphi=\tan^{-1}\left(\frac{\theta}{\tau}\right) \tag{4}$$
$$\sigma(\varphi)=a_0+\sum_{p=1}^{p_{\max}}\left(a_p\cos(2p\varphi)+b_p\sin(2p\varphi)\right)$$

We generalized the results in [7] to the multivariate case by considering N^2 radially gaussian kernels. The parameters of these kernels should be optimized for the classification. In order to reduce the complexity of this optimization we propose to decompose it in a space and a time-frequency optimizations.

2.1. Optimization in the spatial domain

The ambiguity function of a multivariate signal provides the information about the joint time, frequency and space correlations. In fact, the off diagonal terms of the matrix $M_S(\theta,\tau)$ (Eq. 2) represent the spatial correlations. If we transform the multivariate signal so as to obtain spatially decorrelated components we can concentrate on the optimization of the diagonal kernels of the transformed components.

The decorrelating transformation is designed so as to obtain transformed components that are maximally separated with respect to the classes being classified as explained in [8]. However the technique in [8]works only for the two class classification problem. In the case of W classes we should perform pair-wises classifications.

The decorrelating transform in the case of two classes can be obtained as follows.

Let R_{w_1} and R_{w_2} (Eq. 5) represent the correlation matrices of the training trials corresponding to classes w_1 and w_2.

$$R_{w_k}=\frac{1}{Q}\sum_{q=1}^{Q}S_{w_k}^q\cdot\left(S_{w_k}^q\right)^t\ \ ;\ k=1,2 \tag{5}$$

We note P_{12} the linear transformation such that:

$$P_{12}^t\cdot R_{w_1}\cdot P_{12}+P_{12}^t\cdot R_{w_2}\cdot P_{12}=I \tag{6}$$

The matrix P_{12} can therefore be expressed as

$$P_{12}=\begin{bmatrix}\vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_N\end{bmatrix}\cdot diag\begin{pmatrix}\frac{1}{\sqrt{\lambda_1}} & \frac{1}{\sqrt{\lambda_2}} & \dots & \frac{1}{\sqrt{\lambda_N}}\end{pmatrix} \tag{7}$$

where $\vec{v}_i$ are the eigenvectors and λ_i are the eigenvalues of $R_{w_1}+R_{w_2}$. Generally P_{12} does not diagonalize either R_{w_1} or R_{w_2}. It does however transform the multivariate signal $S(t)$ into $Y(t)=\begin{bmatrix}y_1(t) & y_2(t) & \dots & y_N(t)\end{bmatrix}^t$ (Eq. 8).

$$Y(t)=P_{12}^t\cdot S(t) \tag{8}$$

The transformation in Eq. 8 is applied to all the trials in the training set. The correlation matrices $R_{w_1}^{'}$ and $R_{w_2}^{'}$ of the transformed trials are then given by

$$R_{w_k}^{'}=P_{12}^t\cdot R_{w_k}\cdot P_{12}\ \ ;\ k=1,2 \tag{9}$$

By virtue of Eq. 6 we have

$$R_{w_1}^{'}+R_{w_2}^{'}=I \tag{10}$$

The eigenvectors of $R_{w_1}^{'}$ and $R_{w_2}^{'}$ are identical. If $\vec{e}$ is an eigenvector of $R_{w_1}^{'}$ corresponding to an eigenvalue λ then $\vec{e}$ is also an eigenvector of $R_{w_2}^{'}$ corresponding to an eigenvalue $(1-\lambda)$. As $R_{w_1}^{'}$ and $R_{w_2}^{'}$ are positive semidefinite the eigenvalues must lie in the interval $[0\ ;\ 1]$. As a result, the eigenvectors that are best for the representation of class w_1 are worst for the representation of class w_2 and vice versa.

From Eq. 10 and the above considerations it is clear that $R_{w_1}^{'}$ and $R_{w_2}^{'}$ can be simultaneously diagonalized by a matrix that we note T_{12}. Such a matrix transforms $Y(t)$ into $Z(t)=\begin{bmatrix}z_1(t) & z_2(t) & \dots & z_N(t)\end{bmatrix}^t$ (Eq. 11) where the $z_i(t)$'s are decorrelated.

$$Z(t)=T_{12}^t\cdot Y(t) \tag{11}$$

Applying the Eq. 8 to Eq. 11 we have.

$$Z(t)=T_{12}^t\cdot P_{12}^t\cdot S(t)=K_{12}\cdot S(t) \tag{12}$$

Therefore, the result of the spatial optimization is a set of transformation matrices K_{ij} for each pair of classes.

2.2. Optimization in the time-frequency domain

From the optimization in the spatial domain (Sect. 2.1) we know that we dispose of a set of transformed training trials whose components are decorrelated.

Without loss of generality we consider the classification between the classes 1 and 2.

Let $\gamma_{12}^{'}=\left\{Z_{w_1}^q(t),Z_{w_2}^q(t)\ ;\ q=1,...,Q\right\}$, the set of labeled transformed trials where $Z_w^q(t)=K_{12}\cdot S_w^q(t)$.

$$Z_w^q(t)=\begin{bmatrix}{}^1z_w^q(t) & {}^2z_w^q(t) & \dots & {}^Nz_w^q(t)\end{bmatrix}^t$$

The radially gaussian kernels (Eq. 4) are separately optimized for each component. The optimization procedure for the kernel associated with the ℓ^{th} component is summarized in Figure 1.

We note ${}^\ell\vec{\sigma}=\begin{bmatrix}{}^\ell a_0 & {}^\ell a_1 & .. & {}^\ell a_{p_{\max}} & {}^\ell b_1 & .. & {}^\ell b_{p_{\max}}\end{bmatrix}^t$ the vector of the parameters associated with the ℓ^{th} gaussian kernel (Eq. 4).

The kernel optimization follows an iterative procedure whose goal is to maximally separate the classes that we want to classify. The separation is measured using the distance between the time-frequency representations (TFR) as explained in Eqs. 13-15.

$$ {}^{\ell}d^{q}_{k1,k2} = \iint \left| {}^{\ell}C^{q}_{w_{k1}}(t,\omega) - \left\langle {}^{\ell}C_{w_{k2}}(t,\omega) \right\rangle \right|^2 d\theta d\tau \quad (13) $$

where ${}^{\ell}d^{q}_{k1,k2}$ is the distance between the TFR of the q^{th} trial of class w_{k_1} (Eq. 14) and the mean TFR of class w_{k_2} (Eq. 15) with respect to the ℓ^{th} component.

$$ {}^{\ell}C^{q}_{w_{k1}}(t,\omega) = \frac{1}{4\pi^2} \iint {}^{\ell}M^{q}_{w_{k1}}(\theta,\tau) \cdot e^{-j\theta t - j\tau\omega} d\theta d\tau \quad (14) $$

$$ \left\langle {}^{\ell}C_{w_{k2}}(t,\omega) \right\rangle = \frac{1}{4\pi^2} \iint \left\langle {}^{\ell}M_{w_{k2}}(t,\omega) \right\rangle \cdot e^{-j\theta t - j\tau\omega} d\theta d\tau \quad (15) $$

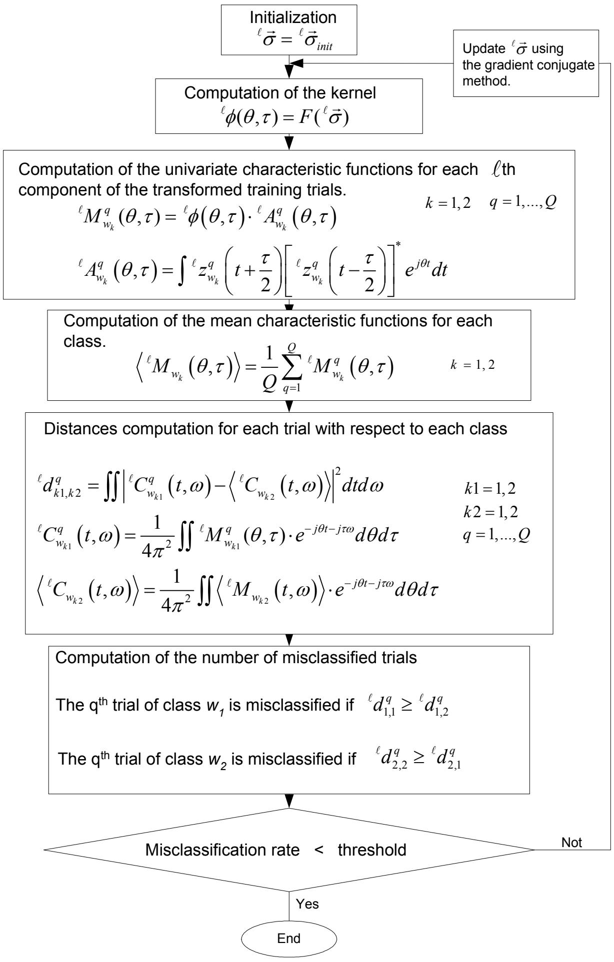

Figure 1. Optimization procedure for the kernel associated with the ℓ^{th} component.

2.3. Classification of an unknown trial

2.3.1. Classification into two classes

When classifying an EEG trial $S(t)$ we first transform it using Eq. 12 and obtain a transformed multivariate signal $Z(t) = [z_1(t) \quad z_2(t) \quad ... \quad z_N(t)]^t$. We then compute the distances, with respect to each component to each of the classes. We therefore obtain $2N$ distances (Sections 2.1 and 2.2). In order to provide a global classification result we combined these distances into a single measure D which is defined below.

$$ D = \sum_{\ell=1}^{N} \lambda_\ell \cdot \left({}^{\ell}d_{w_1} - {}^{\ell}d_{w_2} \right) \quad (16) $$

where ${}^{\ell}d_{w_k} = \iint \left| {}^{\ell}C(t,\omega) - \left\langle {}^{\ell}C_{w_k}(t,\omega) \right\rangle \right|^2 dt d\omega$ and λ_ℓ is the eigenvalue associated with the ℓ^{th} component (Sect. 2.1).

We chose to weight the distances associated with each component by their respective eigenvalues because they provide an indication about the importance of each component in the classification (Sect 2.1).

Finally, $S(t)$ is classified into the class w_1 if $D < 0$ and into the class w_2 if $D > 0$.

2.3.2. Classification into several classes

An EEG-trial $S(t)$ is classified by performing pair-wise comparisons between all possible pairs of classes.

Among all the classification results provided by each pair wise classifier, we chose the most represented class as the result of the classification. In the case that two or more classes are equally represented the classification is not defined.

3 RESULTS AND DISCUSSIONS

The EEG signals were gathered from electrodes Fp1, Fp2, C3, C4, P3, P4, O1 and O2 of the 10-20 international system [9]. Among these signals we used the last six for the classification, i.e. $N = 6$.

The signals coming from Fp1 and Fp2 were used for the detection and removal of perturbations in the EEG signals [3].

We tested the classification method with respect to three types of mental activities: mental counting (MA1) imagined left and right index movement (MA2 and MA3 respectively) recorded under different conditions. Three subjects (S1, S2 and S3) participated in the experiences. We recorded 200 half-second EEG trials corresponding to each mental activity. We divided them into 120 trials (training set) for the design of the transformation matrices and kernels and 80 trials for testing (testing set).

We determined the average rate of misclassification (error rate) per task over 40 different choices of the training and testing sets. The results of the classification, in terms of error rate, are reported in Figure 2. The kernels were optimized after applying the spatial transformation as indicated in Sect. 2.1

As it can be seen the radially gaussian kernel with $p_{\max} = 3$ performs better than any other kernel.

We can also note that the unitary kernel, corresponding to the case in which the characteristic function is equal to the ambiguity function, provides the worst classification result for subjects S1 and S3, and a non-optimal result for subject S2. On other hand an increase of the parameter $p_{\max}$ implies an increase on the error rate, this is due to a more sensitivity to noise when $p_{\max}$ takes big values.

In order to evaluate the utility of the spatial optimization we computed the classification error obtained when the kernels are optimized on the original signals $S(t)$. This is equivalent to the case when $K_{ij} = I$ in Eq. 12. The results are reported in Figure 3. As it can be easily noted, the results of the optimization on the transformed signals are much better then those obtained on the original signals.

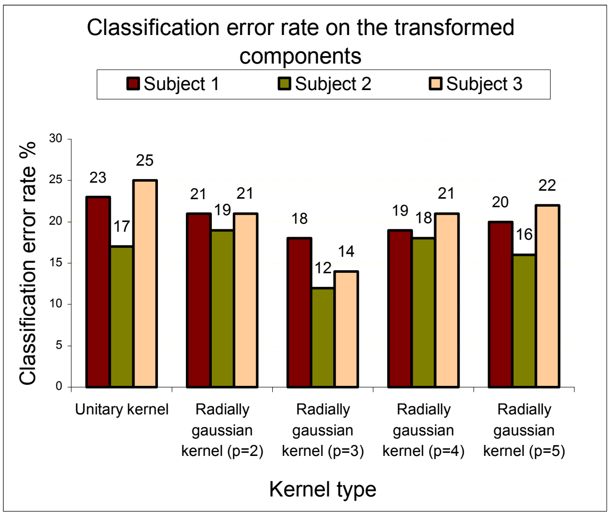

Figure 2. Classification error rate on the transformed components for different choices of the kernel and for each of the three tested subjects.

4 CONCLUSIONS AND FUTURE WORK

From the results obtained, it appears that the classification strategy consisting in dividing the optimization in the spatial and time-frequency domain gives better results than a simple time-frequency on the original (untransformed) signals.

The radially gaussian kernels with $p_{\max} = 3$ provided the best results in terms of classification error. We also noted that an increase of $p_{\max}$ leaded to an increase in the classification error because of the sensitivity to noise that inevitably contaminates EEG signals.

As the final goal of our research is to apply this classification strategy to a BCI system in which a feedback is usually provided to the user when training the system. It is necessary to design a method for the online updating of the kernels.

On other hand, as the number of mental activities in a BCI system increases the number of pair-wises comparisons increases as well. We thus need to adapt our method and find a way to directly compare all the classes at the same time. This could be achieved by simultaneously diagonalizing the correlation matrices in Eq. 5.

An extension of our work that we are currently exploring is the application of the Time-frequency and space analysis of EEG to the segmentation of sleep periods.

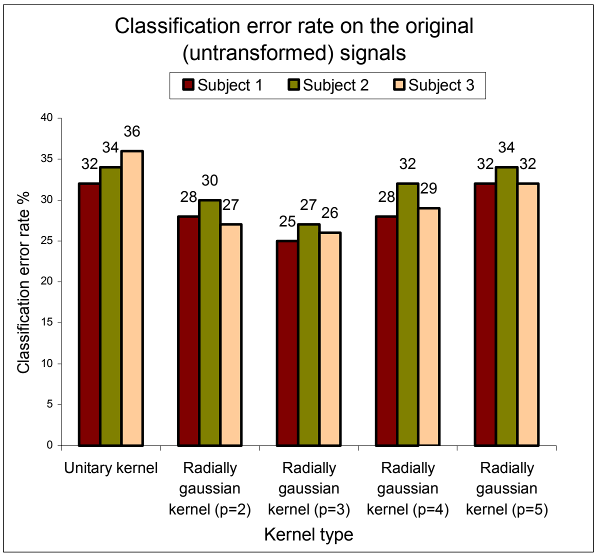

Figure 3. Classification error rate on the original signals for different choices of the kernel and for each of the three tested subjects.

5 REFERENCES

[1] J. R. Wolpaw, et al, “Brain-Computer Interface Technology: A review of the first international meeting”, *IEEE Trans. Rehab. Eng.*, Vol. 8, No. 2, pp. 164-173, June 2000.

[2] D.J. McFarland, et al, “Spatial filter selection for EEG-based communication”, *Electroenceph. Clin. Neurophysiol.,* 103, pp. 386-394, 1997.

[3] G. Garcia, T. Ebrahimi, J.-M. Vesin, “Classification of EEG Signals in the Ambiguity Domain for Brain-Computer Interface Applications”, *IEEE DSP2002, Vol. 1*, pp. 301-306,July 1-3 2002.

[4] P. L. Nunez, et al, “EEG coherency I: statistics, reference electrode, volume conduction, Laplacians, cortical imaging, and interpretation at multiple scales,” *Electroenceph. Clin. Neurophysiol.*, 103, pp. 499-953, 1997.

[5] M. Coates, “Time-frequency modelling”, PhD Thesis university of Cambridge, August 1998.

[6] L. Cohen, Time-Frequency analysis, Prentice Hall Signal Processing Series, New Jersey, 1995.

[7] M. Davy, "Noyaux optimises pour la classification dans le plan temps-frequence - Proposition d'un algorithme constructif et d'une reference bayesienne basee sur les methodes MCMC - Application au diagnostic d'enceintes acoustiques", PhD Thesis, *Université de Nantes,* September 2000.

[8] C. W. Therrien, Decision estimation and classification, John Wiley & Sons, Inc., USA, 1989.

[9] H.H. Jasper, “The Ten-Twenty electrode system of the international federation”, *Electroenceph. Clin. Neurophysiol.,* 10, pp. 371-375, 1958.