Classification of Cerebral Blood Flow Oscillation

Péter Somogyi

Department of Control Engineering and Information Technology, Budapest University of Technology and Economics, Magyar tudósok krt. 2, H-1117 Budapest, Hungary, psomogyi@bio.iit.bme.hu

Balázs Benyó

Department of Informatics, Széchenyi István University Egyetem tér 1, H-9026 Győr, Hungary, benyo@sze.hu

Béla Paláncz

Department of Photogrammetry and Geoinformatics, Budapest University of Technology and Economics, Magyar tudósok krt. 2, H-1117 Budapest Hungary palancz@epito.bme.hu

Abstract: Oscillation of the cerebral blood flow (CBF) is a feature in several physiological or pathophysiological states of the brain. In order to distinguish between different physiological states, two different classification methods have been developed; a Radial Basis Function based Neural Network and a Support Vector Classifier with Gaussian kernel. In order to describe the temporal blood flow patterns, two feature extraction procedures were applied; spectral matrix and wavelet subband analysis.

Keywords: Biomedical systems, Classification, Neural-network models, Radial base function networks, Support vector machine

1 Introduction

Low frequency spontaneous oscillations in cerebral hemodynamics have been observed and linked to certain physiological and patophysiological states, such as epilepsy. Therefore it is worthwhile to investigate the possibilities of classification of the temporal patterns of this vasomotion. Three classes of CBF signals have been distinguished experimentally, and in relation to consecutive administration of two different drugs:

(a) Normal blood flow signals before applying any drugs, that does not exhibit low frequency oscillations (LFO-s), referenced as class A;

(b) Slight oscillation after the administration of L-NAME, a NO synthase inhibitor reportedly evoking CBF oscillations, referenced as class B;

(c) More pronounced oscillation observed after the administration of U-46619 for stimulating thromboxane receptors, having the effect of also inducing LFO, referenced as class C.

To identify the different states of CBF oscillation described above, different classification methods have been employed, using neural network and support vector machine classifiers (SVMC). However, these approaches were only partly successful because the two-dimensional feature vector could not characterize all the features of the time series. Even the most promising technique, the SVMC suffered from overlearning.

The separation of the first class from the two latter has been carried out successfully using two feature vector containing elements derived from the measured signal. However, the second and third classes cannot be effectively distinguished due to the highly overlapping regions (stars and squares), as seen on the feature map Fig. 1. Hence the discrimination of the two LFO classes, or cerebral blood flow states, is the subject of this paper.

Two different feature extraction methods have been applied to characterize the given time signals, based on spectral and wavelet subband analysis.

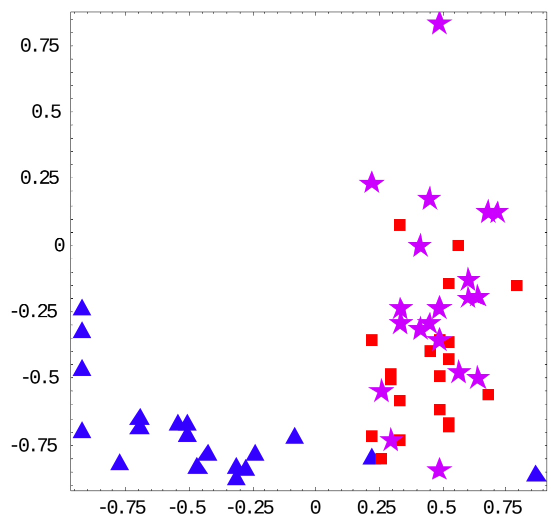

Figure 1

Normalized dimensionless feature map of cerebral blood flow: normal blood flow, class A (triangle), before administration of U-46619, class B (square) and after administration of U-46619, class C (star) from [1]

2 Feature Extraction

2.1 Using Spectral Analysis

According to [2], the two greatest components of a wavelet decomposition do not represent adequately the signals derived from drug induced oscillations. A different approach is an eigenvalue based characterization. In order to obtain the singular values being characteristic of the different states, a matrix has to be derived from the time signal. This is obtained by creating a spectral matrix. Given the time series of data d_i, where i = [1…70,000] are the sample points, we form window vectors of size n and of range m. By choosing $n << m$, the following window vectors can be constructed:

$$\begin{aligned} \underline{u}_1 &= (d_1, d_2, d_3, d_4, d_5, d_6, \ldots, d_n) \\ \underline{u}_2 &= (d_2, d_3, d_4, d_5, d_6, d_7, \ldots, d_{n+1}) \\ &\vdots \\ \underline{u}_{m-n+1} &= (d_{m-n+1}, d_{m-n+2}, d_{m-n+3}, \ldots, d_m) \end{aligned}, \qquad (1)$$

The matrix is built from these window vectors as columns:

$$\underline{\underline{A}} = \left[\underline{u}_1^T \, \underline{u}_2^T \cdots \underline{u}_{m-n+1}^T\right], \qquad (2)$$

and our spectral matrix is $\underline{\underline{A}}^T \underline{\underline{A}}$.

In order to find the optimal window size and range, a series of decompositions have been completed, and the reconstructed signals have been compared to the original recordings. There is a few percent difference; therefore, for the feature extraction, a window size of 50 and a window range of 5000 samples has been selected.

Employing these window parameters, the eigenvalues of the spectral matrix can be computed. As it can be seen, in the case of a class C signal on Fig. 2, the first six values are good candidates to be the elements of the feature vector describing a given signal.

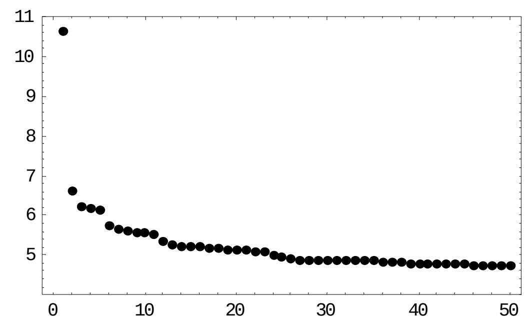

Figure 2
Eigenvalues of the spectral matrix of a class C signal, on a logarithmic scale

2.2 Feature Extraction via Wavelet Transformation

Before the discrete wavelet transformation of the time signal can be computed, we drop the beginning and the end of this raw signal, getting a signal of length of 2^{16} samples.

This transformation decomposes the data into a set of coefficients in the wavelet basis. There are 16 sublists containing the wavelet coefficients in the orthogonal basis of the orthogonal subspaces.

The contributions of the coefficients to the signal at different scales are represented by the phase space plot, see Fig. 3. Each rectangle is shaded according to the value of the corresponding coefficient: the bigger the absolute value of the coefficient, the darker the area. The time unit is 5 msec.

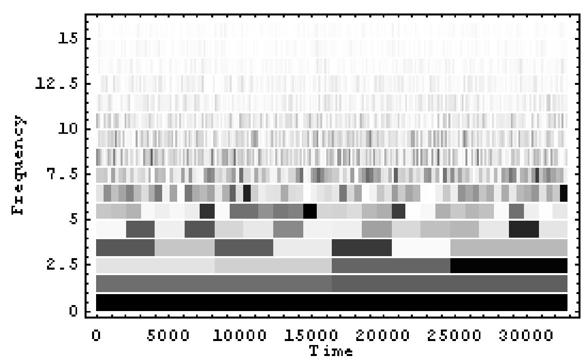

Figure 3
The phase space plot of the DWT of the time signal

Normally, from the wavelet coefficients of each of the 16 resolution levels (subbands) and from sample values of the original time signal, one computes the average energy content of the coefficients at each resolution. There are a total of 17 subbands (16 wavelet subbands and one approximation subband represented by the original signal), from which features are extracted. The ith element of the feature vector is given by,

$$v_i = \frac{1}{n_i} \sum_{j=1}^{n_i} w_{i,j}{}^2, i = 1,2,...,17, \quad (3)$$

where $n_1 = 2$, $n_2 = 2$, $n_3 = 2^2, \ldots, n_{16} = 2^{15}$ and $n_{17} = 2^{16}$, where $w_{i,j}$ is the jth coefficient of the ith subband. In this way, from a time signal having 2^k samples or dimensions, one can extract a feature vector of $k+1$ dimensions.

This technique has been extended for two dimensional signals, for digital images.

In order to study the effect of the dimension of the input space on the quality of the classification as well as to save the morphology of DWT, here we employ a different approach. We consider the wavelet coefficients belonging to a given subband as a feature vector based on this given resolution. It can be a reasonable approach, because the approximated signal representation in the orthogonal subspace corresponding to this subband is given by these coefficients.

In our case, there are two sets of time signals, representing two classes of CBF states and only 40 patterns (2 × 20) are at our disposal. Intuitively, it is possible to shatter two points by any linear manner in the one-dimensional space and three points in two-dimensional space. By analogy, it is possible to shatter N +1 points in the N-dimensional space with probability 1. If the patterns to be classified are independent and identical distributed, then in the 2 N patterns are linearly separable in the N-dimensional space.

The coefficients of the subbands from $n_2 = 2$ up to $n_6 = 2^5 = 32$ as different feature vector components will be employed. Fig. 4 shows the maximums of the magnitude of the wavelet coefficients of different resolutions, except of those belonging to the first (lowest) one. The omitted first wavelet coefficient has a magnitude of about 74276, being significantly larger than the other wavelet coefficients.

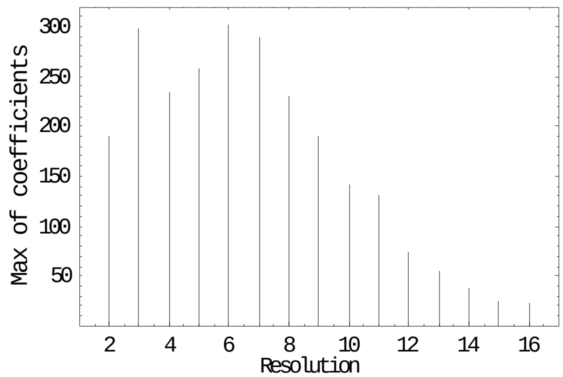

Figure 4
The maximal magnitudes of the wavelet coefficients of different resolutions

3 Classification

3.1 Using Radial Basis Function with Artificial Neural Networks

Considering N patterns of measured CBF signals representing the two overlapping classes, we have $\underline{x}_i \in R^M$ feature vectors derived from time series samples, where $i = 1...N$ are the samples, and M is the dimension of the feature vectors, consisting of several dominant eigenvalues. In our case the number of the measurements were $N = 40$. In order to obtain the minimum size of the feature vector which is required to produce reliable results, up to six eigenvalues were used. The goal of the classification problem is to assign new, previously unseen patterns to their respective classes based on previously known examples: in our case to assign input signals to class B or class C. Therefore the output of our unsupervised learning algorithm is a set of discrete class labels corresponding to the different CBF states. The labelled patterns corresponding to classes B and C, were to be classified. This means, that we are looking for a decision function; the output of this estimating function is interpreted as being proportional to the probability that the input belongs to the corresponding class.

To carry out the systematic classification of CBF signals, an Radial Basis Function was used.

3.2 Support Vector Machine (SVM) Classifier

This kernel based classifier can be trained on any size of training set, while neural networks should have so many input nodes as the dimension of the input space and need definitely more training patterns than the number of these input nodes. Employing kernels, a classification problem can be transferred in a higher dimensional space, where the linear separability is more likely. In addition, the quality of the classification in any dimension can be measured by the geometric margin of the SVM classifier.

Here we used the feature vectors produced by the wavelet subband analysis. Twenty of these vectors represent one CBF state, the other twenty represent the other state. As an example let us load the coefficients of the fifth subband, $n_5 = 2^4 = 16$, for all of the 40 patterns, giving us 40 feature vectors of dimension of 16.

First, these data should be standardized; to be transformed so that their mean is zero and their unbiased estimate of variance is unity.

Let us employ Gaussian kernel, with parameter β = 5,

K[u_,v_]:=Exp[-□ (u-v).(u-v)] (5)

Let the value for the control parameter of regularization be $c = 100$.

To carry out the training of the support vector classifier, we shall employ the algorithm embedded into the function, SupportVectorClassifier developed for Mathematica.

A sample pattern can be considered as support vector, if its contribution (its weighting coefficient α_i) to the decision function is greater than 1% of the maximal contribution.

The geometric margin, γ can indicate the quality of the classification, greater the γ, more reliable the classification is,

$$\gamma = \left(\sum_{i=1}^{N} \alpha_i - \frac{1}{c} \langle \alpha \cdot \alpha \rangle \right)^{-1/2} \tag{6}$$

These computations were carried out for different feature vectors based on the coefficients of the different subbands.

4 Results

Columns 1 and 2 of Table 1 show the result of the ANN classification results using eigenvalue-based feature extraction. It can be clearly seen from the numbers, that it makes no sense to use more than 6 eigenvalues. Comparing different feature extraction methods and classification algorithms by taking different numbers of eigenvalues as feature vectors, the results are very close to that obtained when using wavelet decomposition, see columns 3 and 4 of Table 3. In any case, it is clear, that a merely two element feature vector is insufficient for reliable results; at least a five element feature vector was needed to differentiate class B from class C in the case of ANN classification with eigenvalue-based feature extraction, while in the case of the SVM classification with wavelet-based feature extraction, at least 8-dimensional feature vector should be used.

Table I

Misclassification rate

Number of eigen-values	Eigenvalue feature extraction & ANN classification, misclassifi-cation number	Wavelet coefficient (subband level)	Wavelet feature extraction & SVM classifi-cation, misclassifi-cation number
6	0	16 (5)	0
5	0	8 (4)	0
4	3	4 (3)	0
3	3	-	-
2	6	2 (2)	4

Conclusions

Two feature extraction and classification methods are presented. First an Artificial Neural Network using a radial based function, combined with a spectral matrix based feature extraction was shown. Secondly, a Support Vector Machine Classifier with wavelet subband analysis as feature extraction method was employed. The two mehods can successfully differentiate cerebral blood flow classes B and C, and although the approaches described in this paper are very different, they still produced comparable results for this classification problem.

References

[1] B. Benyó, G. Lenzsér, and B. Paláncz, "Characterization of the Temporal Pattern of Cerebral Blood Flow Oscillations" in *Proceedings of 2004 International Joint Conference on Neural Networks*, Budapest, Hungary, July 25-29, 2004, pp. 2467-2470

[2] B. Benyó, Z. Benyó jr., Z. Benyó, P. Somogyi, B. Paláncz, "Classification of Cerebral Flood Flow Oscillation using SVM classifiers with different kernels" in *Intelligent Systems at the Service of Mankind*, Vol. 2, 2005