Automatic selection of frequency and time intervals for classification of EEG signals

M. Dalponte, F. Bovolo and L. Bruzzone

A novel technique is presented for the automatic selection of time and frequency intervals to be used in feature extraction on multidimensional signals acquired by an electroencephalogram (EEG). This technique is completely automatic, adaptive (task independent), and does not require any specific prior domain knowledge. Experimental results obtained by integrating the proposed technique in a system for brain computer interface (BCI) confirm its effectiveness.

Introduction: Automatic analysis of EEG signals is very important both for supporting diagnosis of brain diseases and for contributing to a better understanding of cognitive processes. In recent years, increasing attention has been devoted to the analysis of EEG signals in motor imagery problems related to BCI applications [1]. In this context, a major issue is related to the development of completely automatic systems for EEG signal classification to be implemented on computer architectures. However, the design of an effective classification system is complex, given the high variability of the EEG signals in the presence of different subjects and target events (classes).

One of the most complex problems to address in the system design consists of the selection of proper time and frequency intervals in which to filter the EEG signals before feature extraction. Generally, these intervals are identified by a visual analysis on the data carried out by experts. This affects the system-design phase, requiring a time-consuming expert-dependent process for each set of events to be classified and each subject. Only a few papers have partially addressed the semi-automatic selection of either frequency bands or time intervals. However, application independent, fully automatic and adaptive techniques for the selection of both frequency and time intervals are not available in the literature. In this Letter, we propose a system for EEG signal classification that exploits a novel fully automatic technique for the selection of time and frequency intervals for effective feature extraction in EEG signals for BCI applications.

Proposed system: The architecture of the proposed system is shown in Fig. 1. It is made up of a pre-processing module (based on the standard common average reference (CAR) algorithm [1]), a feature-extraction module, and a classification module based on a supervised support vector machine (SVM) classifier [2]. The feature extraction module includes: (i) the frequency and time filters with the proposed technique for the selection of their parameters; and (ii) the common spatial subspace decomposition (CSSD) algorithm [3] for feature computation. In the following we focus our attention on the proposed technique for the selection of the filter parameters, which is the main novel contribution of this Letter.

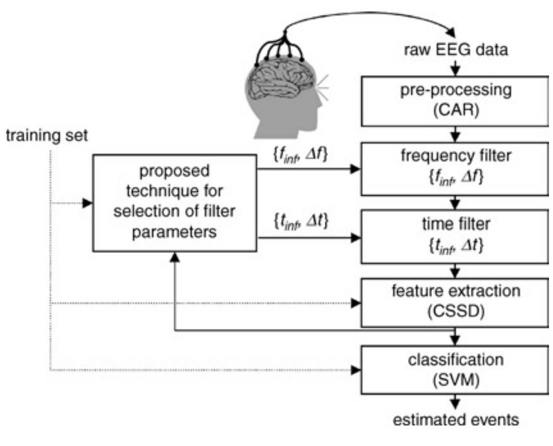

Fig. 1 *Architecture of system used for assessing effectiveness of proposed technique for automatic selection of frequency and time intervals for classification of EEG signals*

Let us assume that the analysed signals show significant frequencies in $[f_{min}, f_{max}]$ and have been observed for an interval $[t_{min}, t_{max}]$. The two intervals represent the search domains for the filter parameters. Let Δf denote the bandwidth of the bandpass filter and $f_{inf} \in [f_{min}, f_{max} - \Delta f]$ be its lower cutoff frequency. Let Δt be the width of the analysis window of the time filter and $t_{inf} \in [t_{min}, t_{max} - \Delta t]$ be its lower bound. Finally, let $\Omega = \{\omega_1, \ldots, \omega_M\}$ be the set of the M ($M \geq 2$) classes (events) to be recognised. The goal of the proposed technique is to automatically identify the set of parameter values $\boldsymbol{P}_{opt} = \{f_{inf\ opt}, \Delta f_{opt}, t_{inf\ opt}, \Delta t_{opt}\}$ of the filters that define the most effective intervals for the discrimination of the events (classes) under investigation. This is done according to the use of: (i) a *search strategy*; and (ii) a *statistical distance measure*. The search strategy properly generates candidate combinations of parameter values $\boldsymbol{P} = \{f_{inf}, \Delta f, t_{inf}, \Delta t\}$ to be tested. Then, after filtering with these values, features are extracted with the CSSD algorithm and a statistical measure of distance among class distributions is computed in the feature space. The set of values $\boldsymbol{P}_{opt}$ associated with the features that result in the highest distance (d_{opt}) among classes is automatically selected. The proposed search strategy and statistical distance measure are described in the following.

Search strategy: We adopt an exhaustive iterative search strategy based on a hierarchical analysis of quantised frequency and time intervals. Let Δf_{max}, Δf_{min}, and Δt_{max}, Δt_{min} be the maximum and the minimum width of the frequency and time windows, respectively, considered in the search algorithm. Let S_f and S_t be the shifts applied to f_{inf} and t_{inf}, and $S_{\Delta f}$ and $S_{\Delta t}$ the variation steps of Δf and Δt at each iteration. Let $d(\boldsymbol{P})$ denote the distance measure between class (event) distributions after feature extraction with the set of filter parameters $\boldsymbol{P}$. The pseudo-code procedure of the adopted search strategy is as follows:

$d_{opt} = 0$
FOR $\Delta f = \Delta f_{max}$ **TO** Δf_{min}
  FOR $\Delta t = \Delta t_{max}$ **TO** Δt_{min}
    FOR $f_{inf} = f_{min}$ **TO** $f_{max} - \Delta f$
      FOR $t_{inf} = t_{min}$ **TO** $t_{max} - \Delta t$
        $\boldsymbol{P} = \{t_{inf}, \Delta t; f_{inf}, \Delta f\}$
        Filter signal in frequency in the range $[f_{inf}, f_{inf} + \Delta f]$
        Filter signal in time in the range $[t_{inf}, t_{inf} + \Delta t]$
        Extract features with the CSSD algorithm
        Compute $d(\boldsymbol{P})$
        IF $d(\boldsymbol{P}) > d_{opt}$
        THEN $d_{opt} = d(\boldsymbol{P})$
          $\boldsymbol{P}_{opt} = \boldsymbol{P}$
        END IF
        $f_{inf} = f_{inf} + S_f$
      END FOR
      $t_{inf} = t_{inf} + S_t$
    END FOR
    $\Delta f = \Delta f - S_{\Delta f}$
  END FOR
  $\Delta t = \Delta t - S_{\Delta t}$;
END FOR

The selected filter parameters are those in $\boldsymbol{P}_{opt}$ at the convergence of the algorithm.

Statistical distance measure: The separability measure $d(\boldsymbol{P})$ adopted in the proposed technique is the Jeffries-Matusita (JM) distance [4]. The JM distance between the classes ω_i and ω_j is defined as follows:

$$JM_{ij}(\boldsymbol{P}) = \left\{ \int_x \left[\sqrt{p(\boldsymbol{x}/\omega_i)} - \sqrt{p(\boldsymbol{x}/\omega_j)} \right]^2 d\boldsymbol{x} \right\}^{1/2} \tag{1}$$

where $p(\boldsymbol{x}/\omega_i)$ and $p(\boldsymbol{x}/\omega_j)$ are the conditional probability density functions for the CSSD feature vector $\boldsymbol{x}$ given the classes ω_i and ω_j. These density functions depend on the selected filter values in $\boldsymbol{P}$ (which affect the CSSD features). Equation (1) can be rewritten as:

$$JM_{ij}(\boldsymbol{P}) = \sqrt{2\left\{1 - \exp\left[-B_{ij}(\boldsymbol{P})\right]\right\}} \tag{2}$$

where B_{ij} is the Bhattacharyya distance. Under the assumption of Gaussian distributed features (approximation reasonable for many features extracted from EEG signals, including the CSSD measures), the

Bhattacharyya distance can be easily computed as follows:

$$B_{ij}(\boldsymbol{P}) = \frac{1}{8}(\boldsymbol{m}_i - \boldsymbol{m}_j)^T\left((\Sigma_i + \Sigma_j)/2\right)^{-1}(\boldsymbol{m}_i - \boldsymbol{m}_j) + \frac{1}{2}\ln\left(\frac{|(\Sigma_i + \Sigma_j)/2|}{|\Sigma_i|^{1/2}|\Sigma_j|^{1/2}}\right) \quad (3)$$

where $\boldsymbol{m}_i$, $\boldsymbol{m}_j$ and Σ_i, Σ_j are the mean and the covariance matrices, respectively, for classes ω_i and ω_j. When $M > 2$ classes are considered, the average JM distance can be computed as follows [4]:

$$JM_{ave}(\boldsymbol{P}) = \sum_{i=1}^{M}\sum_{j>i}^{M} P(\omega_i)P(\omega_j)JM_{ij}(\boldsymbol{P}) \quad (4)$$

where $P(\omega_i)$ and $P(\omega_j)$ are the prior probabilities of classes ω_i and ω_j. The use of the JM distance is particularly suitable for the solution of the considered problem when $M > 2$, as pairwise JM distances have a saturating behaviour when classes are well separated. This avoids bias effects of well separated classes on overlapped ones in the selection of filter parameters.

Experimental results: The signals used to test the proposed system are associated with the Data Set IVa of the BCI Competition 2005 [5]. This data set is focused on motor imagery tasks. The data are recorded from five healthy subjects ('aa', 'al', 'av', 'aw' and 'ay'). In each trial a visual cue indicated to subjects which motor task to imagine: move right hand and right foot. The duration of each trial is 3.5 s. There are 140 events for each class analysed, for a total of 280 events for each subject. EEG data were acquired by 118 electrodes, and the signals are sampled at 100 Hz. Table 1 gives the number of training and test samples for each subject (see [5] for greater detail).

Table 1: Number of training and test samples used in experiments for five subjects considered; overall accuracies obtained with proposed automatic technique and reference values also shown

Subject	Training samples	Test samples	Overall accuracy (%)	
			Proposed technique	Reference intervals
aa	168	112	86.6	89.3
al	224	56	100	100
av	84	196	74.0	78.6
aw	56	224	98.7	99.1
ay	28	252	95.6	96.4

We set the limit values for the frequency analysis in the range [f_{min} = 8 Hz; f_{max} = 28 Hz] as these are the most relevant frequencies for motor task problems. For the time analysis, we considered the range [t_{min} = 0 s, t_{max} = 4 s] (i.e. the whole duration of the cue, plus a 0.5 s margin). The search strategy was implemented with values of Δf and Δt ranging from Δf_{max} = 20 Hz to Δf_{min} = 2 Hz (with step $S_{\Delta f}$ = 1 Hz and shift S_f = 0.5 Hz) and from Δt_{max} = 4 s to Δt_{min} = 2 s (with step $S_{\Delta t}$ = 0.1 s and shift S_t = 0.1 s), respectively.

Table 2 gives frequency and time intervals selected in a completely automatic way by the proposed method (without using any specific domain knowledge) and reference values obtained by an optimal manual analysis. It is possible to observe that for four subjects there is a very good agreement between our results and reference intervals. Table 1 shows accuracies obtained after the SVM-based classification with the intervals selected by the proposed technique and with the reference ones. The average accuracy obtained by the presented method is 90.9%, and for three subjects accuracies are higher than 95%. These accuracies are comparable to those obtained with the reference values (see Table 1), and confirm the effectiveness of the proposed technique.

Table 2: Frequency and time intervals obtained with proposed technique for considered five subjects; reference values derived manually also shown

Subject	Proposed technique		Reference values	
	Frequency (Hz)	Time (s)	Frequency (Hz)	Time (s)
aa	11.5–13.5	0.5–3.5	11–16	0.7–3.5
al	10.5–14.5	0.5–3.5	12–14	0.7–4.3
av	11–14	0.8–2.0	9.5–11.5	1.0–3.5
aw	10.5–14.5	0.9–4.0	10–15	0.8–4.0
ay	8.5–21.5	0.8–2.0	9–24	0.7–2.0

Conclusions: We have proposed an EEG-based BCI system, which exploits a novel automatic technique for the selection of frequency and time intervals in feature extraction on EEG signals. The main properties of this technique are: (i) it does not require any prior knowledge on the EEG signal behaviour; (ii) it is adaptive (task independent); (iii) it can be used with any kind of feature extraction and classification method; and (iv) it is easy to implement and to use. Results confirm both the effectiveness of the proposed technique in selecting the most informative frequency and time components, and the capability of the considered system to recognise motor imagery tasks with a high accuracy.

21 August 2007
Electronics Letters online no: 20072428
doi: 10.1049/el:20072428

M. Dalponte, F. Bovolo and L. Bruzzone (*Department of Information and Communication Technologies, University of Trento, Via Sommarive, Povo, Trento 14 I-38050, Italy*)

E-mail: michele.dalponte@dit.unitn.it

M. Dalponte: Also with Centro di Ecologia Alpina, Viote del Monte Bondone, I-38040 Trento, Italy

References

1 McFarland, D.J., Mccane, L.M., David, S.V., and Wolpaw, J.R.: 'Spatial filter selection for EEG-based communication', *Electroencephalogr. Clin. Neurophysiol.*, 1997, **103**, pp. 386–394
2 Vapnik, V.N.: 'The nature of statistical learning theory' (Springer-Verlag, Berlin, Germany, 1999)
3 Wang, Y., Berg, P., and Scherg, M.: 'Common spatial subspace decomposition applied to analysis of brain responses under multiple task conditions: a simulation study', *Clin. Neurophysiol.*, 1999, **110**, (4), pp. 604–614
4 Bruzzone, L., Roli, F., and Serpico, S.B.: 'An extension of the Jeffreys-Matusita distance to multiclass cases for feature selection', *IEEE Trans. Geosci. Remote Sens.*, 1995, **33**, (6), pp. 1318–1321
5 http://ida.first.fraunhofer.de/projects/bci/competition