BCI-competition III – the Graz data - dataset IIIb – Algorithm Description

Model Description

Damien Coyle, Girijesh Prasad and Martin McGinnity

Intelligent Systems Engineering Laboratory, School of Computing and Intelligent Systems,
Faculty of Engineering, Magee Campus, University of Ulster, Northland Road, Derry,
Northern Ireland, BT48 7JL, UK. Phone: +44 (0)28 7137 5170. Email: dh.coyle@ulster.ac.uk

Preprocessing

For subjects O3VR and S4 the raw EEG data are fed through prediction neural networks (pNNs). Two neural networks (NNs) are trained to perform one-step-ahead predictions for the EEG time-series data, where one NN is trained on right motor imagery signals and the other on left motor imagery signals (both are trained on trained on data from channels C3 and C4 therefore each pNN has two output channels). Each NN is referred to as LpNN or RpNN corresponding to the type of data it is specialized on (either left or right motor imagery).

To obtain the preprocessed signals the raw EEG are input to both pNNs and each pNN provides predictions for both the C3 and C4 data therefore four predicted signals are produced for each trial. This process is referred to as neural-time-series-prediction-preprocessing (NTSPP). Due to each pNN specialization on a particular signal type (left or right motor imagery) the morphology of the predicted signals (Ys) should be different to the morphology of the original signals (Os) and the Ys signals should be more separable than the Os signals.

For subject X11 no preprocessing was performed because of difficulties in training pNNs which enhanced the data separability.

Feature Extraction

For subject O3VR and X11 features are extracted using time-frequency methods. A main sliding window containing a number of shorter overlapped windows is employed. The frequency spectra of the signal contained within the shorter windows is calculated using the Fourier transform. This is referred to as the short-time-Fourier-transform (STFT) and enables the frequency components to be localized in time within the main sliding window. The spectra obtained within each sliding window are smoothed by an interpolation process. The norm (Euclidean length) of the power of each interpolated spectra within subject-specific frequency bands is used as a feature for the signal.

The number of features, *k,* for each signal at a particular time point *t* is

k=fix((stp-ovlap)/(length(wind)-ovlap));

where *stp* is the length of the main window, *ovlap* is the amount of overlap between consecutive STFT windows and length(*wind*) is the length of the STFT window.

For subject O3VR there are 4*k features for each trial a particular time point because features are extracted from the NTSPP signals.

For subject X11 there are 2*k features for each trial a particular time point because features are extracted from the original signals.

For subject S4 features are derived from the mean of the signals within a main sliding window. The main sliding window length is subject specific and one feature is derived from each signals at a particular time point *t*. Features are extracted from the NTSPP signals therefore there are four features for each trial a particular time point

The data in the main window for all subjects always includes signals preceding and including time point *t* but never exceeding *t*.

Classification

Linear discriminant analysis (LDA) is used for classification. Features are extracted and classified at every time point *t* (at the rate of the sampling interval).