Abstract

Electromyogram (EMG) contamination has been shown to affect electroencephalogram (EEG) signals. Therefore, methods of isolating and removing EMG contamination are a focus of research. One of the most common ways to eliminate this contamination is through independent component analysis (ICA). Also, surface Laplacian (SL) has been proven to isolate the distant sources of EEG signals. The objective of this thesis is to demonstrate the effects of EMG contamination on EEG signals using the Neurophysiological Biomarker Toolbox (NBT) and the impact of applying ICA, and ICA + SL on raw data. In this thesis, the method for preparing the data is ICA with an auto-pruned method and SL using a flexible spherical spline. The thesis has two main sections designed to demonstrate the objective. The first describes the use of random sampling of subjects who were assigned three tasks during EEG recording (eyes closed, eyes open, and solving a maze) and comparing them, under three types of data preprocessing, using Student's paired t-test and normalised amplitude of delta (1–4 Hz), alpha (8–13 Hz), and gamma (30–45 Hz). Second, machine learning was used to classify three neuropsychiatric diseases (anxiety, depression, and epilepsy) against control subjects under the three types of data pre-processing and raw data + SL. The data has been split into one second segments and classified according to features extracted from the NBT, which are the amplitude and the normalised amplitude for all frequency bands. Principal component analysis (PCA) was used for reducing the features, and 10x10-fold cross-validation and artificial neural networking were the methods used for the classification.

The results in the first section show that EMG contamination affected the EEG signal in the gamma bands, that ICA eliminated the EMG contamination, and that ICA + SL improved the reading of brain signals; and the delta and alpha bands were not affected by ICA or ICA + SL. The results in the second section show a high percentage of accuracy in ICA + SL in all frequency bands. However, ICA in general has a percentage quite similar to the raw data, while SL, as well as ICA with a small percentage improved more than ICA and raw data. Overall, the gamma band for both amplitude and normalised amplitude in ICA + SL showed the best results, with accuracy over 87%, when comparing it with all disease classifications. Both results indicate that ICA + SL eliminate and isolate EMG contamination. However, the classification of ICA shows no significant change in the percentage of accuracy.