Data quality assessment of capacitively-coupled ECG signals

Aims: This study aims to build a classification method to assess the quality level of capacitively-coupled ECG (ccECG) signals. This is important since ccECG is particularly sensitive to motion artefacts.

Methods: 10.000 15-second ccECG segments from different setups were included. From these, half had an ECG reference. 5 annotators labelled the segments in three quality levels. After outlier rejection, a Fleiss' Kappa of 0.80 was achieved. Two binary datasets were created: one with a 'low threshold' (level1 vs level2-3) -datasetL- and another with a 'high threshold' (level1-2 vs level3) -datasetH-. 48 features were extracted from each segment, and the dataset was divided in 70% training and 30% test. Feature selection using Neighborhood Component Analysis (NCA) and a method based on Random Forest (RF) classification was applied to the training sets. Different supervised classifiers were tested using the selected features.

Results: For 9910 segments at least three annotators agreed. Similarity metrics between the ccECG and reference ECG presented a histogram in accordance to the annotations. The NCA-based and RF-based methods partially agreed on the selected features. The evaluation of the classification models using the selected features resulted in balanced accuracies of 91.53% (NCA selection) and 91.58% (RF selection) for the 'datasetL', and of 92.40% (NCA selection) and 94.35% (RF selection) for the 'datasetH'. In all the evaluations except the last RF reported value, the maximum performance was obtained using fine K-nearest neighbor classification. In the latter case, it was obtained with a coarse SVM using a Gaussian kernel.

Conclusion: The selected features allow a quality-based ccECG classification with high accuracy. This allows to process long-term ccECG recordings. Furthermore, multi-electrode ccECG systems can benefit from this as part of real-time electrode selection schemes. This has the potential to increase the time coverage of ccECG signals in realistic scenarios.