Assessment of non-negative matrix factorization for the preprocessing of long-term ECG

Pauline Guyot, El-Hadi Djermoune, Thierry Bastogne

ASSESSMENT OF NON-NEGATIVE MATRIX FACTORIZATION FOR THE PREPROCESSING OF LONG-TERM ECG

P. GUYOT[1,2,3], E.-H. DJERMOUNE[1], T. BASTOGNE[1,2,3]
[1]CRAN UL-CNRS UMR 7039, [2]INRIA BIGS, [3]CYBERnano

Background. With the advent of health information technology and wearable acquisition systems, **long-term ECG** are more and more used in long-term cardiac tolerability studies of new compounds. Nevertheless, the accurate analysis of such long signals requires reliable and fast signal processing algorithms.

Objectives

The objective of this study is to assess the practical relevance of innovative matrix factorization **methods for the preprocessing of long-term ECG**. Those signals are generally noisy with complex baseline wander and require preprocessing, such as filtering, to perform a correct analysis. Our goal is to present two innovative algorithms of matrix factorization to detect R-peaks in long-term ECG.

Methods

The two tested methods are :

- **Independent Component Analysis (ICA)**
- **Non-Negative Matrix Factorization (NMF)**

They are both source separation methods whose goal is to isolate each component of an ECG: R-peak, P and T waves, noise, baseline wander etc. On the one hand, ICA assumes that all subcomponents of the signal are statistically independent from each other and on the other hand, Non-Negative Matrix Factorization is a method that uses the non-negativity of the spectrogram of the ECG to separate the different time-frequency patterns. The two signal processing methods were implemented in the Matlab computing environment.

Results

The proposed approaches are tested on the MIT-BIH Arrythmia and Noise Stress Test databases: ICA shows a strong drawback by returning the different sources in a random order making compulsory a reconnaissance step or the action of a specialist. Whereas NMF achieves high results in terms of sensitivity and specificity in general even in case of complex baseline wander and highly noisy signals.

Conclusion

This study emphasizes promising results of a new long-term ECG preprocessing technique based on a matrix factorization method. This approach simultaneously undertakes three tasks: denoising, baseline wander removal and peak R detection.

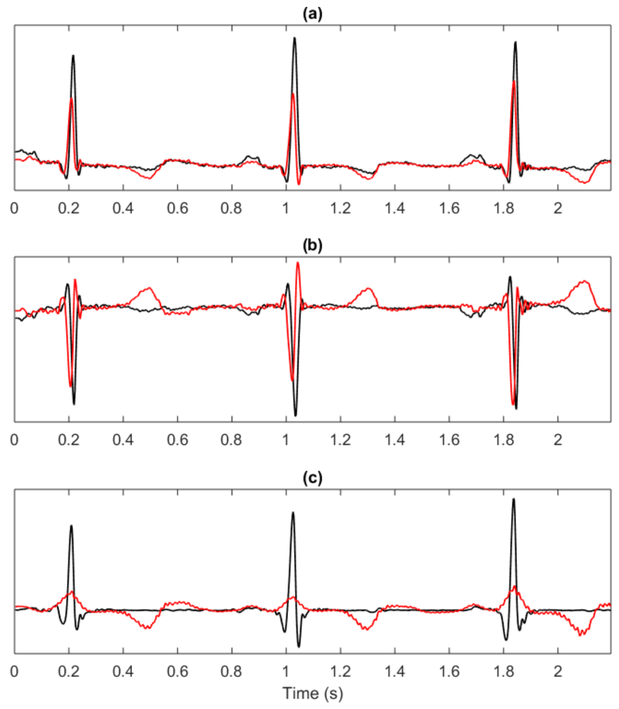

Fig.1: Comparison of separation source methods applied on the first ECG signal of the MIT-BIH Arrhythmia Database : (a) Original two-leads ECG signal (b) ICA components; R-peaks are shared in both sources when P waves can be mostly found in the first source and T-waves in the second source (c) NMF components; the first source only contains R-peaks while the second source contains P and T waves and some low-frequency parts of QRS complexes.

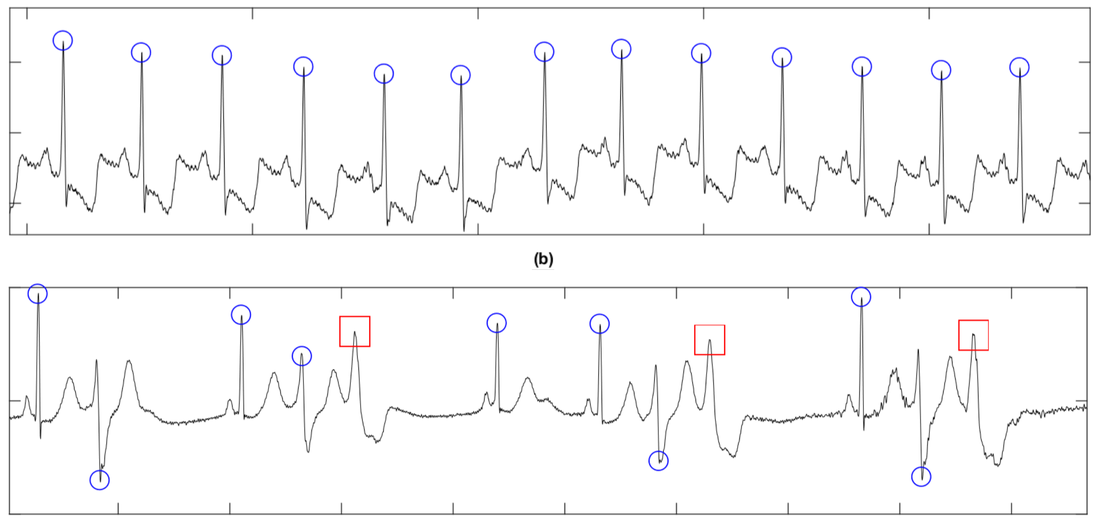

Fig.2: Results of R-peak detection using NMF separation on the MIT-BIH Arrhythmia Database, blue circles represent correctly detected peaks and red squares missed peaks: (a) 10 seconds of ECG signal with all peaks correctly detected (b) 10 seconds of ECG signal with three missed peaks because of arrhythmia.

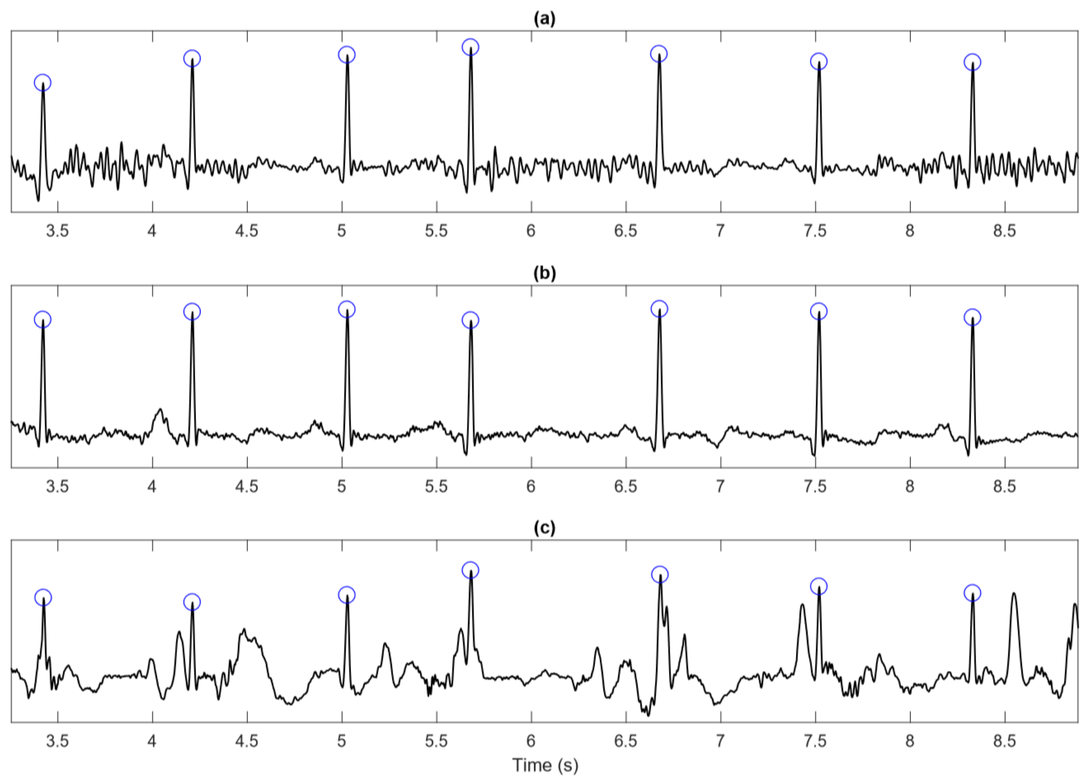

Fig.3: Results of R-peak detection using NMF on the MIT-BIH Arrhythmia Database polluted by differend kind of noises from the Noise Stress Test Database, blue circles represent correctly detected peaks; (a) ECG signal polluted by muscle artifacts (b) ECG signal polluted by baseline wander (c) ECG signal polluted by electrode movements.

Annual Meeting of Safety Pharmacology Society, Washington DC, Sep 30 – Oct 3 2018