Dictionary learning for M/EEG multidimensional data
Christos Papageorgakis, Sebastian Hitziger, Théodore Papadopoulo

To cite this version:
Christos Papageorgakis, Sebastian Hitziger, Théodore Papadopoulo. Dictionary learning for M/EEG multidimensional data. International conference on basic and clinical multimodal imaging (BACI), Sep 2015, Utrecht, Netherlands. <hal-01243284>
Dictionary learning for M/EEG multidimensional data

Christos Papageorgakis1,2, Sebastian Hitziger1, Théodore Papadopoulou1
1Inria Sophia Antipolis - Méditerranée, France
2Université Côte d’Azur, France

Sept. 1–5, 2015

1. Jitter-adaptive dictionary learning model (JADL)

JADL is a dictionary learning framework to learn a dictionary for one or multiple channels and latency and phase of atoms $\phi_j$.

- Atoms learned by JADL are defined on the entire signal domain.
- The set of signals of interest can suffer from unknown time delays (jitter).
- The dictionary learning framework is extended to handle the jitter-adaptive dictionary learning method, that is able to handle signals with unknown time delays.

The algorithm solving the JADL problem, is built on [2] common dictionary learning, which iteratively alternates between:

(i) Sparse coding: finding the coefficients $a_{ij}$ and the jitters $\Delta_i$.

Let an "uncropped" version of the dictionary $D$ be a dictionary $D'$ containing all allowed shifts ($S - \{\Delta\}$) of all its atoms.

The sparse coding problem is solved using a modification of least angle regression (LARS) by restricting the problem as follows:

Once an atom $\Delta_i$ is chosen all its shifts are forbidden.

(ii) Dictionary update: finding the shapes $\phi_j$.

Block coordinate descent is used to iteratively solve the constrained minimization problem for each atom.

2. Our modified JADL model

We propose an extension to the jitter-adaptive dictionary learning method, that:

- is able to handle multidimensional measurements such as M/EEG.
- learns a dictionary over M/EEG recordings that have the same waveform and jitter over all the channels in a single trial.
- is still able to account for different jitters across trials.

Significant modifications are applied to the original JADL framework, especially in:

- The dictionary learning framework.
- Sparse coding: finding the coefficients $a_{ij}$ and the jitters $\Delta_i$.

3. Synthetic data generation

- Create a dictionary $K = 3$ synthetic atoms.
- Generate an extended dictionary of $9$ signals.
- Introducing random jitters (from the set $\Delta = \{\Delta_i\}$ of all its atoms).
- Select 3 source groups, each of them containing 3 neighboring sources.
- Each source group is associated to shifted versions of the same atom.

- Combine the generated signals with a lead field model $G$ computed from real EEG measurements [3].
- Introducing random jitters to the dictionary of $K = 3$ synthetic atoms.
- Generated clean M/EEG measurements of $C = 6$ channels, $M = 200$ trials and $N = 515$ time samples.

4. Results on lead field synthetic data

A comparison between the original and our multi-dimensional JADL model shows:

- Similar results when the best channel is used by the single-channel algorithm.
- Worse results for the single-channel algorithm when a medium or the worst channel is used.
- Good fit to the data, where the single-channel algorithm is unable to recover correctly all the atoms of the dictionary used to generate the signals.

5. Results on real data

The multi-dimensional approach is tested using real MEG and EEG data:

- $C = 200$ channels.
- $M = 63$ trials.
- $N = 541$ time samples.
- Contaminated by ambient noise.

6. Conclusions

- The method shows superior performance and less noisy estimated waveforms compared to the original single-channel JADL framework, both on synthetic and real data.
- It is more robust to various levels of noise.
- Using the JADL framework allows one to deal with signal variations such as jitters which is difficult to do with standard methods such as PCA or ICA.
- Not having to select a "best" channel (as with the JADL method) is both a user simplification and allows the exploitation of all the available information for M/EEG trial by trial signal decomposition. Thus it provides better estimations of waveforms in the dictionary.

References