

FMRI ICA Training: ICA-based BOLD fMRI Artifact Remover Training under Matlab - version 1.2 (beta). Documentation

Jussi Tohka

October 27, 2008

1 INTRODUCTION

Blood oxygenation level dependent (BOLD) signals in functional magnetic resonance imaging (fMRI) are often small compared to the level of noise in the data. The sources of noise are numerous including different kinds of motion artifacts and physiological noise with complex patterns. It has been suggested that the contributions of some of these artifacts to fMRI timeseries can be detected using independent component analysis (ICA). If an independent component reflecting an obvious artifact can be automatically detected, then its removal from the subsequent statistical analysis of the data becomes straightforward. This package of the Matlab functions implements an automatic classifier trained to distinguish between independent components (ICs) reflecting an obvious artifact and ICs possibly related to true BOLD signals. The method has been described in [1].

Please read carefully this document and the article [1] before using this software. For licensing conditions, please see the m-files in this package.

This is still considered to be a beta release.

2 REQUIREMENTS

- FMRI ICA training software runs under Matlab (Mathworks, Natick, MA). Hence, Matlab is needed. The Matlab versions 6.x, and 7.x should be OK although the versions 5.x have not been tested.
- FMRI ICA Classify software is needed. It is available free of charge from the same site as this classifier training tool.
- FMRI ICA training and classify softwares are designed to work with the Melodic ICA-software from the FSL. The FSL software is freely available for research purposes. Note that the software currently relies on the naming conventions etc. of the Melodic 2.0 while the current version is 3.0. The version of this software which will be compatible with Melodic 3.0 will be made available later on. The component identification method is independent from the ICA algorithm used, but the training of the classifiers was performed based on the output of the Melodic 2.0. This may have an effect on the classifiers.
- FMRI ICA classify software also requires Nifti tools by Jimmy Shen. These can be downloaded at <http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=8797&objectType=FILE>
- FMRI ICA classify software has been tested under Linux and Unix.

training_data_fn	The name of the file specifying the training data. The training data consists of the ICA decompositions of FMRI timeseries and classifications of each independent components (either noise or signal) for these components. The structure of the file is described in more detail later on.
np_thr	The desired Neyman Pearson thresholds.
np_testlevels	The NP thresholds for which the element classifiers are trained. Defaults to [0.0100 0.0150 0.0200 0.0300 0.0400 0.0500 0.0700 0.0900]. If you are happy with the default, you can give an empty matrix.
blocked	0 if event related design (target frequency assumed to be [0.01 Hz - 0.1 Hz]) 1 if blocked design with the target frequency of 10 cycles n if blocked design with the target frequency of n cycles
tr(optional)	The repetition time (TR) in ms. Defaults to 2000. If you're using a considerably different TR than 2000, the classifier may not work well for the event related case.
flimits(optional)	A 3 component vector which sets 1) low frequency limit, 2) the lower limit of the target frequency and 3) the upper limit of the target frequency. All should be given in Hertz. 2) and 3) are ignored for the blocked designs. Give an empty matrix if happy with the defaults.
input_rejected_clusters (optional)	This a new option in the version 1.2. You can define the noise classes of the rejected components. This is a cell vector with the number of cells equaling the number of training 4-D time-series. Each cell is a matrix n by 2 matrix whose each row gives 1) the number of the rejected component and 2) its noise class. To use this option, I would suggest that you 1) run the algorithm once to get the noise classes of the rejected components cell vector as an output, 2) change the matrix in those parts that you need to, and 3) re-train the classifier with the new rejected components matrix.

Table 1: Input arguments

3 INSTALLATION

Just copy the zip-package into the same directory where the NifTI and FMRI ICA classify tools have been placed and unzip.

4 INTENDED USAGE

FMRI ICA Training is to be used to train classifiers for the FMRI ICA classify. For the intended usage of FMRI ICA classify tools, consult the documentation of FMRI ICA classify.

5 USING THE SOFTWARE

5.1 Basic usage

The main Matlab function is `fmri_ica_training`. Use it as

```
[global_decision_tree_classifier, rejected_clusters] =
fmri_ica_training(training_data_fn,np_thr,
                  np_testlevels,blocked,[tr],[flimits], [input_rejected_components]);
```

The details about the input arguments are given in Table 1. The outputs are described in the preamble of the file `fmri_ica_training`.

Training data files: The training data files are ASCII-text files listing 1) directories of the Melodic ICA decompositions, 2) rejected (noise) components for each decomposition. Each line of the file should start with a name of directory containing ICA decomposition (see the section 'FILE NAMING CONVENTIONS') followed by the indexes of the noise components in that decomposition. See the file 'training_data.txt' in this package for an example. **Important!.** It is imperative that the file contains no empty lines because an empty line is judged as the end of the input by the program.

File types: The 4-D timeseries files are expected to be either in 1) Analyze 7.5 or 2) NiFTI format. The program deduces automatically which type of files you use. The program supports the use of gzip compressed files under Unix based operating systems.

6 FILE NAMING CONVENTIONS

The FMRI ICA training software assumes the following file naming conventions, which follow the default settings of the Melodic 2.0. You may need to edit the source code, if the conventions are not equal of the desired settings. The following convention is in effect :

- The `ica_directory` is expected to contain following files: 1) `melodic_IC.nii` (or `melodic_IC.img` and `melodic_IC.hdr`) 2) `mask.nii` (or `mask.img` and `mask.hdr`), 3) `melodic_mix`, 4) `melodic_FTmix`; These file names can be easily adapted by editing the function `fmri_readicani` in the file `fmri_ica_training.m`.

7 NEW IN VERSION 1.2

- The warnings created by the possibility that there are classes with no training data were erased. The classifier is fully capable dealing with such a possibility..
- The possibility to define the noise classes in training data was added. This could be important if there are very few training samples and/or if only certain noise classes are to be removed. The classifier should work well for removing only ceratin noise classes.
- Minor bug corrections took place. Note that these may yield slightly altered classifiers.
- Added a warning displayed if the training algorithm has not found any classifiers meeting the required NP threshold(s).

References

- [1] J. Tohka, K. Foerde, A.R. Aron, S.M. Tom, A.W. Toga, and R.A. Poldrack. Automatic independent component labeling for artifact removal in fmri. *NeuroImage*, 39:1227 –1245, 2008.