A FSL-based pipeline for fMRI data preprocessing

Huang, L., Parks, E.L., Baliki, M., Hashmi, J. A., Petre, B., Brain, A. T., and Apkarian A. V.
Physiology Department, Feinberg School of Medicine, Northwestern University, 303 E. Chicago Ave, Chicago, IL 60611

In fMRI data analysis, preprocessing aims to remove various artifacts, which negatively impact the functional signal-to-noise ratio. FMRIB's Software Library (FSL), with an easy-to-use and powerful graphical user interface (GUI), identifies and removes artifacts by removing user-visually-identified Probabilistic Independent Component Analysis (PICA) components. However, this GUI preprocessing procedure is only suitable for small numbers of data sets. When data sets become large, especially for our study (100 subjects for Chronic Back Pain study and each subject is scanned 6 times during one year period), the GUI processing is time consuming and prone to operator-dependent mistakes. Furthermore, manually labeling artifact components affects the consistence of further data analysis. Here, using tools from FSL, we develop an automated pipeline in Bash Shell environment incorporating a procedure of automated selection of PICA artifact components.

Figure 1 is a flow diagram of the pipeline. Two input text files store the paths of to-be-processed fMRI functional and corresponding T1 data sets. The output is an email that informs users of the preprocessing status. For the procedure of automated selection artifact components, we extract 7 features from temporal and spatial components. The maximum correlation between a given temporal component and six motion parameters is a key feature. The other 6 features include correlation between a given temporal component and an average white matter and ventricle signal, and the percentage of activated voxels in edge, white matter, ventricle, and gray matter.

In our preliminary assessment, we compare 2 fMRI data sets processed by the pipeline or by lab members. Four experienced raters labeled artifact components. The algorithm successfully identified all “signal” components and picks up over half of the artifact components. The variance from the algorithm is less than that among raters. We conclude that the pipeline is more efficient than a manual approach, and improves the quality of fMRI data processing.

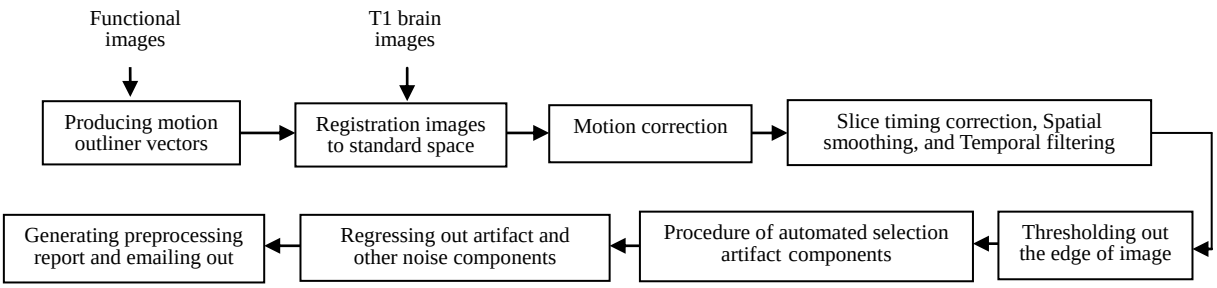

Fig. 1 Flow diagram of the proposed pipeline