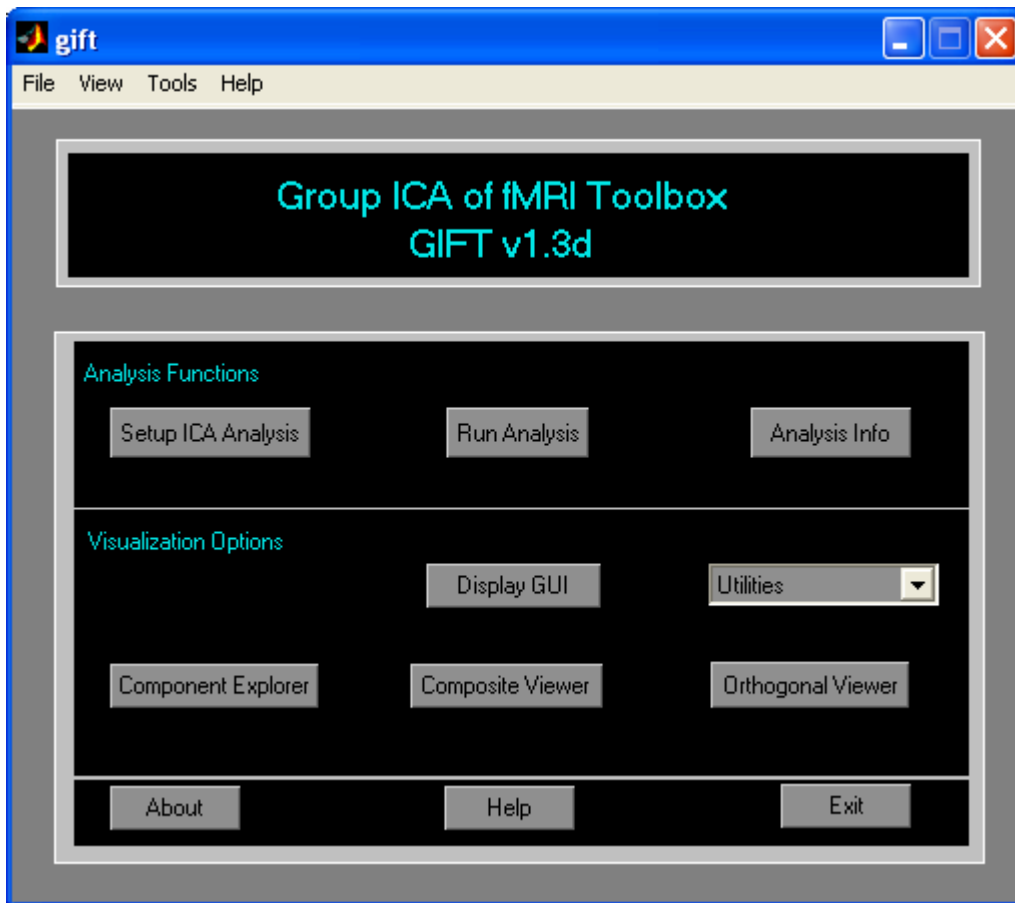


Group ICA of fMRI Toolbox (GIFT) Manual

Srinivas Rachakonda¹, Eric Egolf², Nicolle Correa³ and Vince Calhoun¹⁴
December 18, 2007



¹ The MIND Institute, Albuquerque, NM

² Olin Neuropsychiatry Research Center, Hartford, CT

³ University of Maryland Baltimore

⁴ Dept. of Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM

Contents

1	Introduction	1
1.1	What is GIFT	1
1.2	Why ICA on fMRI data	1
1.3	Why Group ICA	1
2	Getting Started with GIFT	5
2.1	Installing GIFT	5
2.2	Installing Example Subjects	5
2.3	HTML Help	5
2.4	Organizing Data	5
2.5	Menus	6
2.6	Important Information	6
2.7	Analysis Functions	8
2.7.1	Setup ICA Analysis	8
2.7.2	Run Analysis	10
2.7.3	Analysis Info	12
2.8	Visualization Methods	13
2.8.1	Component Explorer	14
2.8.2	Subject Component Explorer	15
2.8.3	Orthogonal Explorer	15
2.8.4	Composite Viewer	18
2.9	Sorting Components	18
2.9.1	Temporal Sorting	18
2.9.2	Spatial Sorting	24
2.10	Utilities	26
2.10.1	Remove Components	26
2.10.2	ICASSO	29
2.10.3	Ascii to SPM.mat	30
2.10.4	Event Related Average	30
2.10.5	Calculate Stats	31
2.10.6	Stats On Beta Weights	31
2.11	More Information	35
2.11.1	Batch Script	35
2.11.2	Output Files Naming	38
3	Process involved in Group ICA	39
3.1	Data Reduction	39
3.2	Independent Component Analysis	39
3.2.1	Infomax	39
3.2.2	Fast ICA	40
3.2.3	JADE OPAC	40
3.2.4	SIMBEC	40
3.2.5	AMUSE	40
3.2.6	ERICA	41
3.2.7	EVD	41

3.2.8	Constrained ICA (Spatial)	41
3.3	Back Reconstruction	41
3.4	Calibrating Components	41
3.5	Group Stats	41
4	Appendix	45
4.1	Experimental Paradigms	45
4.2	Defaults	45
4.3	Options Window	47
4.4	Interactive Figure Window	47
4.5	GIFT Startup File	49

Chapter 1

Introduction

This manual is divided mainly into three chapters. Motivation for using the group ICA of fMRI Toolbox (GIFT) is discussed in this Chapter. In Chapter 2, quick start to the toolbox is discussed. Chapter 3 focusses on the process involved in group ICA.

1.1 What is GIFT

GIFT is an application developed in MATLAB 6.5 that enables group inferences from fMRI data using Independent Component Analysis (ICA). A detailed explanation of ICA is explained in the next section. GIFT is used to run both single subject and single session analysis as well as group analysis.

1.2 Why ICA on fMRI data

Functional Magnetic Resonance Imaging (fMRI) is a modality for studying the brain function. fMRI techniques use Blood Oxygenation Level Dependent (BOLD) signal as a measure for detecting the neural activity. Since fMRI uses an indirect measure of neural activity, mathematical models are needed to analyze the data. Many fMRI experiments use a block design in which the subject is instructed to perform experimental and control tasks in an alternating sequence of 20-40 second blocks. The resulting activity is recorded for each volume element (voxel) of the brain. Based on events of the experimental task and knowing shape of the haemodynamic response, a reference function is constructed.

Model based techniques such as SPM ([18]) use this reference function to separate the signals of interest and the signals not of interest. A general method is necessary which does not depend on the prior information of the experiment or task. In this respect, a statistical technique called Independent Component Analysis (ICA) ([11]) is proposed that allows the extraction of signals of interest and not of interest without any prior information about the task. Thus ICA analysis could reveal characteristics of the brain function that cannot be modeled due to lack of prior information.

Independent Component Analysis (ICA) is a method of blind source signal separation i.e., ICA allows one to extract or "unmix" unknown source signals (Figure 1.1) which are linearly mixed together (Figure 1.2). For fMRI data, temporal and spatial ICA are possible, but spatial ICA is by far the most common approach. The GIFT implements spatial ICA of fMRI data. In spatial ICA, spatially independent brain sources or components are calculated from fMRI data. Figure 1.3 shows a component extracted from the fMRI data. For a complete description of this experiment please see Appendix 4.1.

1.3 Why Group ICA

ICA has been successfully applied to single subject and single session analyses. Group analysis of fMRI is important to study specific conditions within or between groups of subjects. It is not clear how ICA can be applied on a group of subjects as different individuals in the group will have different time courses. In [20], a model was proposed to extend ICA to group studies. Group ICA can be run using a batch script (Section 2.11.1) or using GIFT. If you run batch script, please check all the sub-folders of `icatb` are added to the MATLAB path. The GIFT contains an implementation of ICA for analyzing the fMRI data. Specifically, the GIFT implements both analysis and display

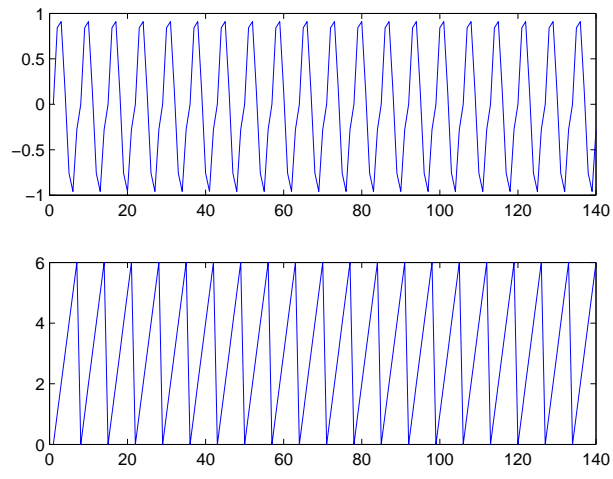


Figure 1.1: Unknown source signals.

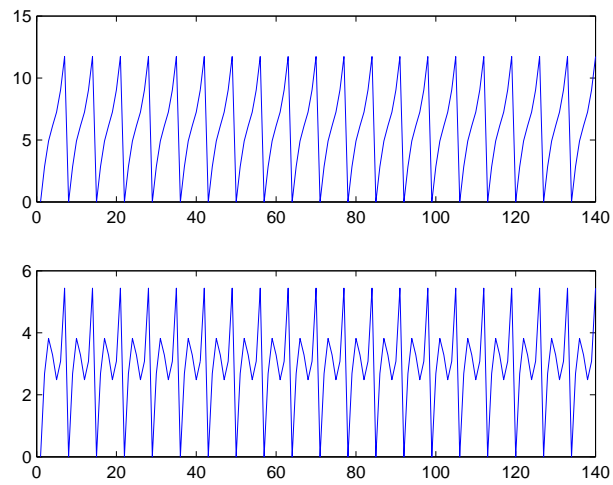


Figure 1.2: Mixed signals.

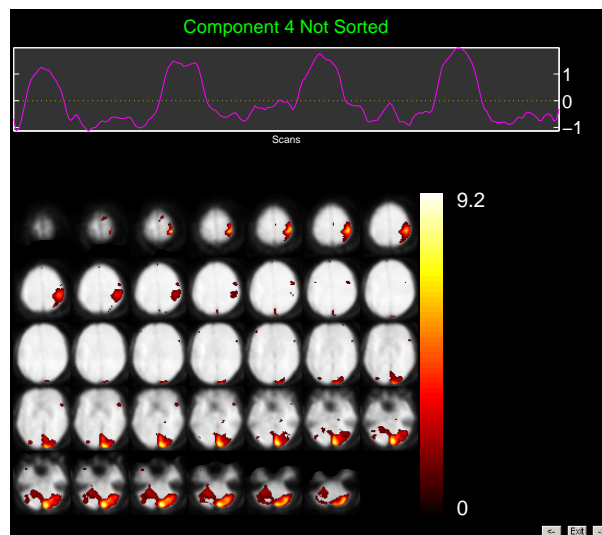


Figure 1.3: A component consists of a time course and a spatial map.

tools, each using standard input and output file types (Analyze or Nifti format). There are three main stages to Group ICA; Data Compression, ICA, and Back Reconstruction ([20]). The outputs from these stages are multiple time courses. Each time course has an image map associated with it. A detailed explanation of the process involved in group ICA is discussed in the Chapter 3.

Chapter 2

Getting Started with GIFT

The source code and the example subjects fMRI data need to be installed. These are available on project web page:

<http://icatb.sourceforge.net>

Also available on this web page are posters on GIFT at the conferences like Society of Biological Psychiatry and Human Brain Mapping. The mailing list is:

icatb-discuss@lists.sourceforge.net

For more information, please send comments or suggestions to:

vcalhoun@unm.edu or srachakonda@mrn.org

2.1 Installing GIFT

Unzip the file `giftv1.3d.zip` and copy the folder `icatb` onto your local machine. Add GIFT directories to the MATLAB search path or run the `gift.m` file to automatically add GIFT directories. The GIFT path will be set at the bottom to fix the menu problem for figures. You can create `gift_startup.m` (Appendix 4.5) file for setting the path according to your needs. After the GIFT path is set, GIFT toolbox (Figure 2.1) opens in a new figure window. There is also an option to run group ICA using a batch script (Section 2.11.1). Batch script is very useful for running large data-sets.

Note: If you have downloaded an older version of GIFT, make sure that there is only one version on path at a time.

2.2 Installing Example Subjects

Download the `example_subjects.zip` file and unzip into an appropriate directory. Included in this file are three subjects pre-processed fMRI data from a visuomotor task (See [20] for a complete description of the task). Whole brain and single slice data (for rapid testing) are provided. More information on the task the subject performed while in the scanner is given in the Appendix 4.1. Each example subject also contains a functional data-set, which contains only one brain slice. If you are having problems with this toolbox you may test this with a smaller data-set.

2.3 HTML Help

When you click *Help* button (Figure 2.1), HTML help manual is opened in the default web browser. "GIFT-help" menu is plotted on some figures which will directly open a particular topic.

2.4 Organizing Data

Organizing data reduces the amount of selection. GIFT has two ways to enter the data for both GUI and batch script.

- First method requires the data to be in one root folder with a common file pattern (like `sw*.img`) for all subjects and sessions. Each subject folder can have session folders and the number of session folders with the matching file pattern should be the same over subjects.
- Second method does not require the data to be in one root folder or have a common file pattern but the selection process through GUI can be tedious for large data-sets and therefore batch script (Section 2.11.1) is recommended.

2.5 Menus

Menus are provided as a shortcut to the user-interface controls in the GIFT toolbox (Figure 2.1). The function of each menu is given below:

- File
 - New - Setup ICA GUI will open after you have selected the output directory.
 - Open - Setup ICA GUI will open showing the values for the parameters after you have selected the subject file that has suffix `Subject.mat`.
 - Close - Closes the GIFT Toolbox and the figures generated by the GIFT.
- View
 - Analysis Info - Analysis information will be shown after you have selected the parameter file that has suffix `ica_parameter_info.mat`.
- Tools
 - Run Analysis - Analysis will be run after you have selected the parameter file.
 - Display GUI - Display GUI will be opened after you have selected the parameter file.
 - Utilities
 - * Remove Component(s) - Removes a component or components from the data after you have selected the parameter file. Please see section 2.10.1 for more information.
 - * `Ascii_to_spm.mat` - Creates `SPM.mat` from ascii file (containing regressor time courses) that is needed during temporal sorting. Please see section 2.10.3 for more information.
 - * Event Average - Event average is calculated for the ICA time courses. Please see section 2.10.4 for more information.
 - * Calculate Stats - Mean, standard deviation, t-maps are calculated for components over sessions, subjects or subjects and sessions.
 - * Stats On Beta Weights - This utility is provided for doing statistics on time courses (beta weights). Please see section 2.10.6 for running this utility.

2.6 Important Information

- Right-left text plotted during display is changed in this version to make it consistent with SPM convention (Neurological convention).
- Flip parameter for analyze images is set to 0 in `icatb_defaults.m` as in previous versions. This variable is stored in ICA parameter file and will give a warning message whenever the flip parameter is changed during display.

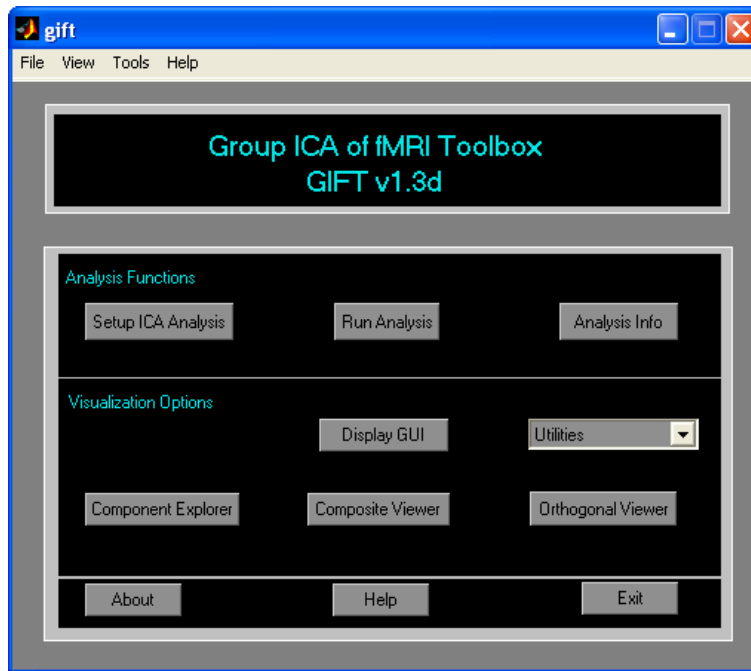


Figure 2.1: GUI for running group ICA.

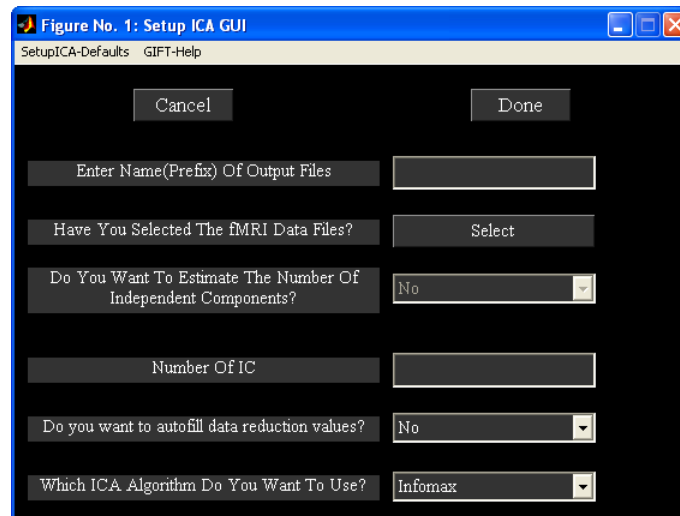


Figure 2.2: Initial parameter selection window.

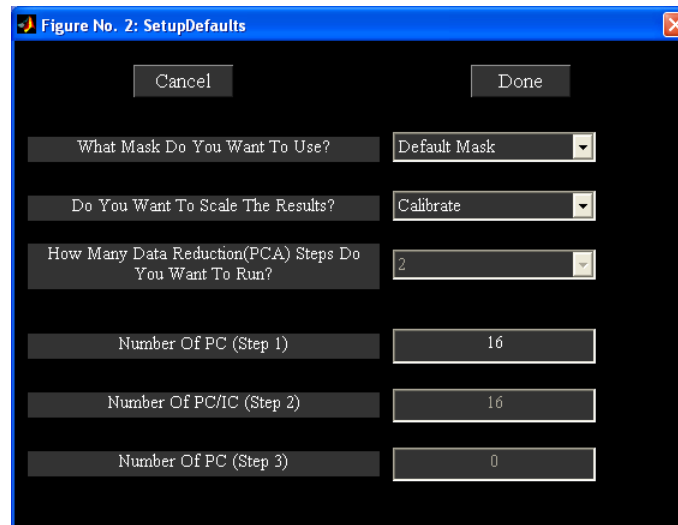


Figure 2.3: Hidden user interface controls.

2.7 Analysis Functions

2.7.1 Setup ICA Analysis

When you click *Setup ICA* button (Figure 2.1), *Setup ICA GUI* (Figure 2.2) will open after you have selected the output directory for the analysis. *Setup ICA* is the GUI used for entering parameters required for group ICA. Figure 2.2 shows the main user interface controls. Some of the parameters are plotted in "Setup ICA-Defaults" menu (Figure 2.3). It is recommended that after entering the parameters in the main figure window parameters plotted in menu be changed. The parameters are explained below:

Main User Interface Controls

- 'Enter Name (Prefix) Of Output Files' is the prefix string to all the output files created by GIFT. This should be a valid character name as the files will be saved using this prefix.
Note: Avoid characters like \, /, :, *, ?, ", < and > in the prefix.
- 'Have You Selected the fMRI Data Files?' Click on the push button *Select* to select the data. There are two options for selecting the data as explained in Section 2.4. After the data is selected, the push button *Select* will be changed to popup with 'Yes' and 'No' as the options.
 - 'Yes' - Data reduction steps will be enabled if you have selected the parameters previously with the same output prefix.
 - 'No' - the data can be selected again.

Note: After the data-sets are selected, a file will be saved with suffix `Subject.mat`. This MAT file contains information about number of subjects, sessions and files.

- 'Do you want to estimate the number of independent components?' Components are estimated ([22]) from the fMRI data using the MDL criteria. All the data-sets or a particular data-set can be used to estimate the components. When all the data-sets are used estimated components are calculated by using the mean of the estimated components of all data-sets.
- 'Number of IC' refers to the number of independent components that will be extracted from the data.

Note: If you have selected Constrained ICA (Spatial) algorithm, the number of independent components is set to the number of spatial reference files selected. Example reference files like `ref_right_visuomotor.nii`, `ref_left_visuomotor.nii` and `ref_default_mode.nii` are provided in folder `icatb/icatb_templates`.

Table 2.1: Number of reduction steps possible for the number of data-sets selected.

Number of data-sets(N)	Number of reduction steps
1	1
$N < 4$	2
$4 \leq N \leq 10$	User specified number (2 or 3)
$N > 10$	3

- 'Do you want to auto fill data reduction values?' By default this option is set to 'Yes' when the data is selected and the 'Number of IC' is set to 20.
- 'Which ICA Algorithm Do You Want To Use?' There are 11 ICA algorithms available like Infomax, FastICA, ERICA, SIMBEC, EVD, JADE OPAC, AMUSE, SDD ICA, Semi-blind Infomax, Constrained ICA (Spatial) and Radical ICA.

Note: If the auto fill data reduction steps drop down box is set to 'No' after entering the prefix, check the numbers for principal components by clicking the "Setup-ICA Defaults" menu. (Figure 2.3).

Hidden User Interface Controls

- 'What Mask Do You Want To Use?' There are two options like 'Default Mask' and 'Select Mask'.
 - 'Default Mask' - Mask is calculated using all the files for subjects and sessions or only the first file for each subject and session depending upon the variable `DEFAULT_MASK_OPTION` value in defaults. Boolean AND operation is done to include the voxels that surpass the mean of each subject's session.

Note: By default first file for each subject session is selected because using all the files is time consuming. You can use all the files for each subject and session by setting variable `DEFAULT_MASK_OPTION` value to `all_files`.
 - 'Select Mask' - You can specify a mask containing the selected regions for the analysis. This mask must be in Analyze or Nifti format.
- 'Do You Want To Scale The Results?' The options available are 'No', 'Calibrate' and 'Z-Scores'.
 - 'Calibrate' - Each subject component image and time course will be scaled to represent percent signal change.
 - 'Z-Scores' - Each subject component image and time course will be converted to z -scores. Standard deviation of image is calculated only for the voxels that are in the mask.
- 'How Many Reduction (PCA) Steps Do You Want To Run?' A maximum of three reduction steps is provided. The number of reduction stages depends on the number of data-sets (Table 2.1). For the example data-set two reduction steps are automatically selected.
- 'Number Of PC (Step 1)' - Number of principal components extracted from each subject's session. For one subject one session this control will be disabled as the number of principal components extracted from the data is the same as the number of independent components.
- 'Number Of PC (Step 2)' - Number of principal components extracted during the second reduction step. This control will be disabled for two data reduction steps as the number of principal components is the same as the number of independent components.
- 'Number Of PC (Step 3)' - Number of principal components extracted during the third reduction step. This control will be disabled for three data reduction steps as the number of principal components is the same as the number of independent components.

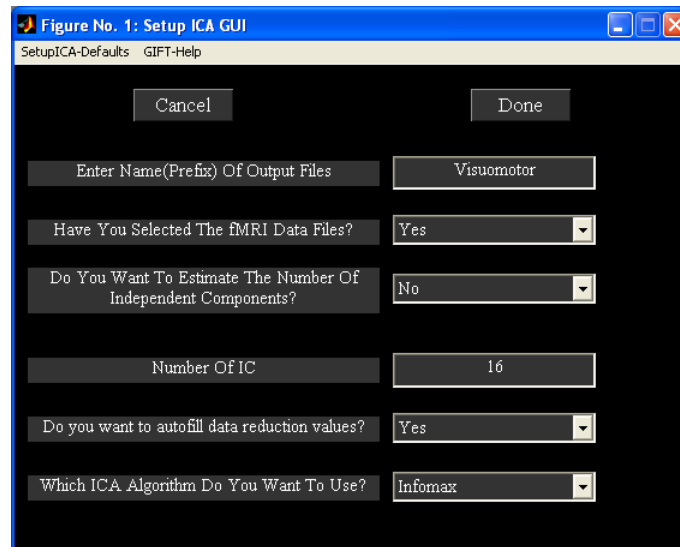


Figure 2.4: Completed parameter selection window.

Completing Entering Parameters

Figure 2.4 shows the completed parameters window. Press *Done* button after selecting all the answers for the parameters. This will open a figure window (Figure 2.5) to select the ICA options. You can select the defaults, which are already selected in the dialog box or you can change the parameters within the acceptable limits that are shown in the prompt string. ICA options window can be turned off by changing defaults (`icatb_defaults.m`). Currently, the dialog box is only available for the Infomax, FastICA, SDD ICA, Semi-blind ICA and Constrained ICA (Spatial) algorithms. After selecting the options, parameter file for the analysis is created in the working directory with the suffix `ica_parameter_info.mat`.

Note:

- Different analyses for the same functional data can be run by copying subject and parameter files to a different directory.
- All the parameters can also be entered by using an input file. This input file should be specified while running batch script. This is the best option for running large data-sets and having different design matrix (required during temporal sorting) for each subject.

2.7.2 Run Analysis

Run Analysis is the step used to perform the group ICA on fMRI data. The following are the steps involved in run analysis:

- The parameter file required for the analysis should be selected. This file is the same file where you entered all the analysis information. It is named as `ica_parameter_info.mat`. Once the parameter file is selected, a figure (Figure 2.6) will pop up showing the options for run analysis. The options are as follows:
 - 'All***'
 - * All the group ICA steps are run at once. The analysis can also be run using one step at a time and should be in order.
 - 'Parameter Initialization'
 - * All the variables that are needed later on or during the analysis are declared and initialized.
 - * A parameter error check is also preformed to try and catch errors before the group analysis begins.
 - 'Data Reduction'

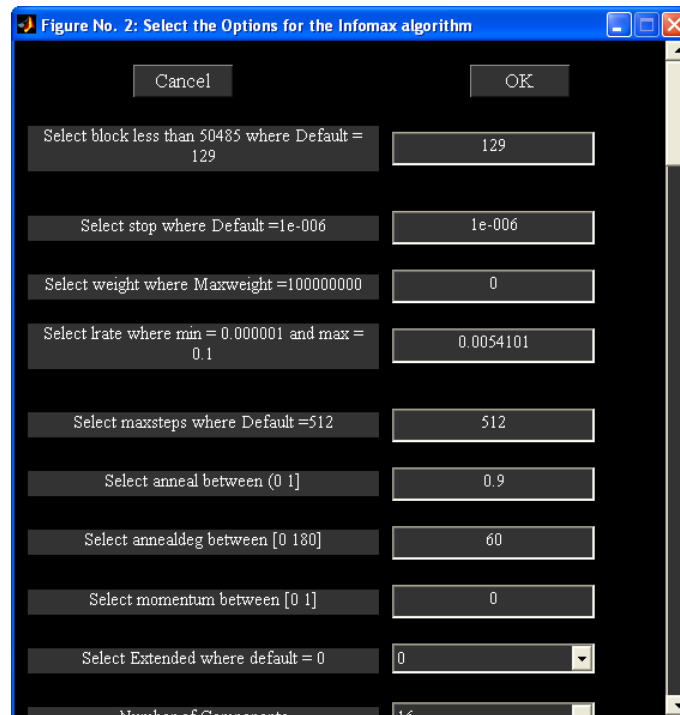


Figure 2.5: ICA Options Window

- * Each data-set is reduced using Principle Components Analysis (PCA). These reduced data-sets are then concatenated into a group or groups depending on the number of data reductions steps selected, this process is repeated.
- * Each reduced data is saved in a MAT file and will be used in the back reconstruction step.
- ‘Calculate ICA’
 - * The concatenated data from the data reduction step is used and the aggregate ICA components are saved in both MAT and Analyze (or Nifti) format.
- ‘Back Reconstruction’
 - * The aggregate components and the results from data reduction are used to compute the individual subject components.
 - * The individual subject components are saved in Analyze (or Nifti) format.
- ‘Calibrating Components’
 - * The functional data is used to scale the components from arbitrary units to percent signal change or z -scores.
- ‘Group Stats’
 - * The individual back reconstructed components are used to compute a mean spatial map and time course, a standard deviation spatial map and time course and a t -statistic spatial map. The time course used for the t -statistic component is the mean time course. These group stats components are calculated for each session and are saved in Analyze (or Nifti) format.
- Results during each of the steps are printed to the MATLAB command window. After the analysis is completed, *Display GUI* (Figure 2.8) will open automatically for visualizing components.

Note:

- All the analysis information is stored in the `_results.log` file. This file gets appended each time the analysis is run with the same prefix for the output files.

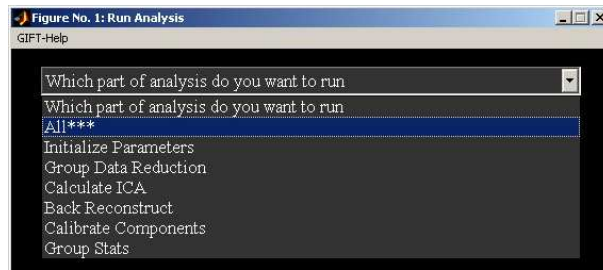


Figure 2.6: Figure window used to run the analysis.

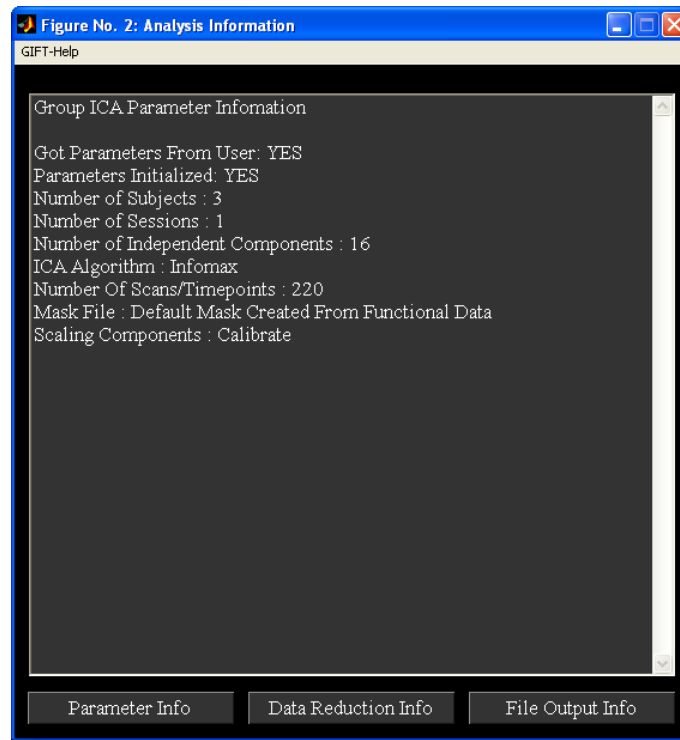


Figure 2.7: Analysis Info shows the information of the completed analysis.

- You need not run the whole analysis if you have changed the scaling method or ICA algorithm as data reduction is the time consuming step for large analysis. For running the analysis with changed scaling method, you need to select 'Parameter Initialization', 'Calibrate Components' and 'Group Stats' one at a time. For running the analysis with changed ICA algorithm, you need to run all the steps except the data reduction.
- By default, the component spatial maps and time courses will be compressed to a zip file according to their viewing set name like subject 1 session 1, mean for session 1, etc. The variable used for compressing files is ZIP_IMAGE_FILES (Appendix 4.2).

2.7.3 Analysis Info

Analysis Info contains the information about the parameters, data reduction and the output files. Once the analysis is done, click on the *Analysis Info* button on the GIFT main window and select the parameter file that you want to look at. Then a figure (Figure 2.7) will pop up showing the information contained within this window.

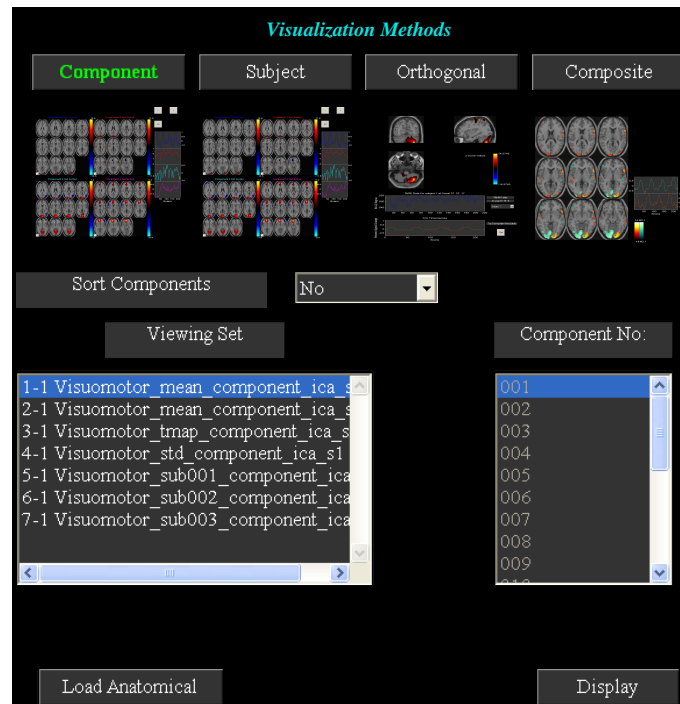


Figure 2.8: *Display GUI* contains four visualization methods like Component Explorer, Subject Explorer, Orthogonal Viewer and Composite Viewer. Component Explorer is plotted by default after you had selected the parameter file.

2.8 Visualization Methods

GIFT contains three main ways of visualizing the components after analysis like the Component Explorer, the Composite Viewer and the Orthogonal Viewer. These three visualization options can be used independently or collectively using the *Display GUI*. This visualization tool provides the user an easy way to explore all the three visualization options along with subject component explorer.

The selected visualization method will be highlighted in colored text. By default, Component Explorer visualization method will be selected after selecting the parameter file. Figure 2.8 shows the main user interface controls. Some of the user-interface controls are shielded from the user and plotted in the "Display Defaults" menu (Figure 2.9). All the display parameters are explained below followed by explanation of visualization methods:

Main user interface controls

- 'Sort Components' - Components will be sorted spatially or temporally. This control will be enabled only for Component Explorer visualization method.
- 'Viewing Set' - This is the component viewing set to look at like subject 1 session 1, mean for session 1, etc and will be disabled for Subject Explorer visualization method.
- 'Component number' - Component number/numbers to look at. This will be disabled for Component Explorer visualization method.
- *Load Anatomical* - *Load Anatomical* button is used to select anatomical image. Component images will be overlaid on this anatomical image. By default, first image of functional data will be used as an anatomical image.
- *Display* - *Display* button is used to display the components of different visualization methods.
- Display Defaults menu - Hidden display parameters will be shown in a figure (Figure 2.9) when you click on Display Defaults menu. This figure contains parameters like 'Image Values', 'Anatomical Plane', 'Threshold', 'Slice Range' and 'Images Per Figure'. "Display GUI Options" menu can be used to change design matrix and

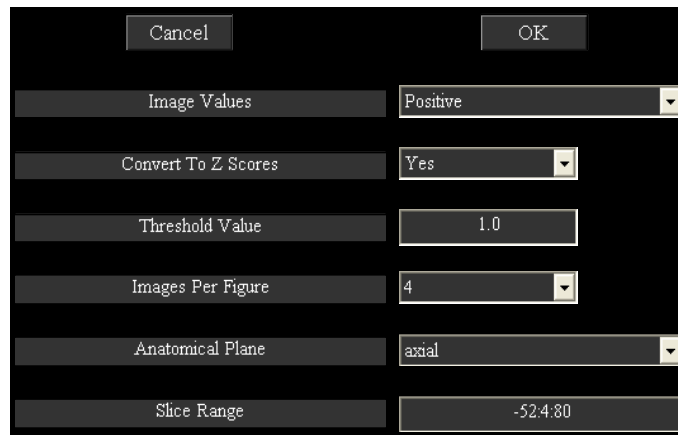


Figure 2.9: User interface controls plotted in "Display Defaults" menu

selecting the text file (See Appendix 4.2) that contains regressor information for temporal sorting. Select 'Design Matrix' for selecting design matrix for temporal sorting. There are three options for selecting design matrix like 'Same regressors for all subjects and sessions', 'Different regressors over sessions', 'Different regressors for subjects and sessions'. The options are explained below:

- 'Same regressors for all subjects and sessions' - The regressors used will be the same over data-sets. This will open a figure window for selecting SPM design matrix.
- 'Different regressors over sessions' - The regressors used will be the same over subjects but different over sessions. This will open a figure window for selecting SPM design matrix.
- 'Different regressors over subjects and sessions' - Different regressors can be used for each subject's session. This will open a figure window for selecting a design matrix for each subject.

User Controls in Display Defaults menu

- 'Image Values' - There are four options like 'Positive', 'Positive and Negative', 'Absolute', 'Negative'. 'Positive' and 'Negative' refer to activations and de-activations on spatial map. You should also look at the time course (flipped or un-flipped) to make the conclusion.
- 'Convert To Z-scores' - Converts spatial maps to z-scores.
- 'Threshold Value' - This is the z threshold value.
- 'Images Per Figure' - Number of images per figure for Component Explorer and Subject Explorer visualization methods.
- 'Anatomical Plane' - This is the anatomical plane to look at for Component Explorer, Subject Explorer and Composite Viewer.
- 'Slices Range' - Slices plotted in mm. Slices in mm are calculated based on the anatomical data. You can change this setting to not use the slices based on the anatomical data by setting `USE_DEFAULT_SLICE_RANGE` variable value to 1 and specify the slices you want to plot in variable `SLICE_RANGE`.

2.8.1 Component Explorer

- Component Explorer is used to display all components of a particular viewing set. Therefore, 'Component Number' control (Figure 2.8) will be disabled for this visualization method.
- Figure 2.8 shows the selected parameters for the Component Explorer. Click on *Display* button and wait for the figures containing spatial maps to pop up. Figure 2.10 shows all the components of mean for all subjects and sessions in groupings of four. By default all the slices in axial plane are plotted. You can change these parameters by clicking on menu "Display Defaults" (Figure 2.8).

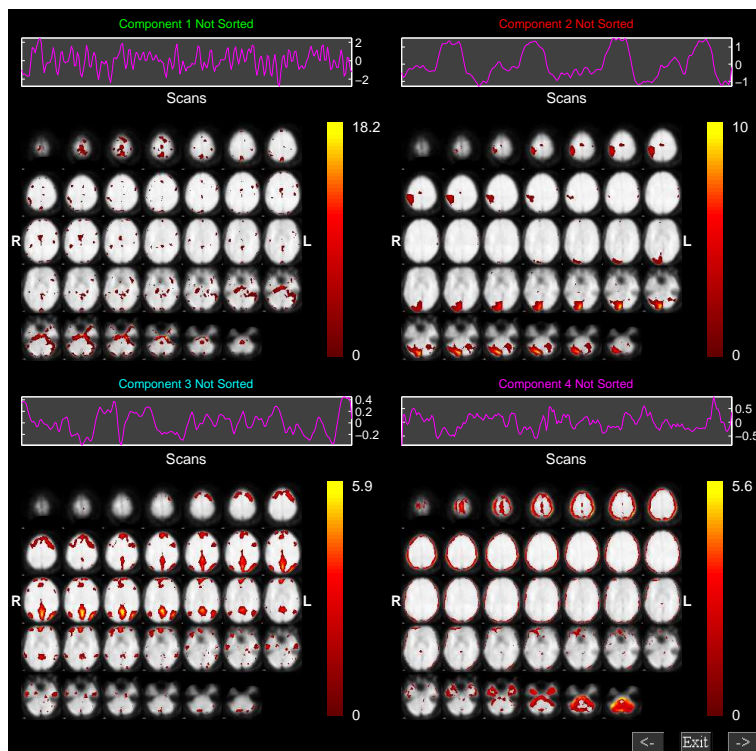


Figure 2.10: Figure shows the components not sorted in groupings of four.

- The time course for each component is displayed on the top of the figure (Figure 2.10). The color bar for each component is displayed next to it. Click on the time course for an enlarged view. Look through the components by clicking on the arrow keys at the bottom of each figure. Find the components of interest and take a note. With this data-set you should find two task related components and one transiently related component. The task related components show activation in the visual cortex. The act of classifying components becomes more difficult with more complex tasks and is the motivation for adding the sorting option.

2.8.2 Subject Component Explorer

- Displays a specific component for all subjects, sessions, mean etc.
- When you click *Subject* button, the 'Component Number' user interface control will be enabled. Figure 2.11 shows the selected parameters.
- Click *Display* button. Figure 2.12 shows the component '001' of all the entries in the 'viewing set'.

2.8.3 Orthogonal Explorer

- Orthogonal viewer is used to look at a component and compare it to the functional data.
- Figure 2.13 shows the selected parameters. Click on Display button. Figure 2.14 shows one of the task related components.
- Upper plot is the BOLD time course for the selected data-set in the popup window at the current voxel. You can interactively select voxel by clicking on any of the slices. Lower plot shows the ICA time course for the maximum voxel (red), minimum voxel (dotted red) and the selected voxel (green).
- When you click on *Plot* button top five components (of the selected viewing set in *Display GUI*) for the selected voxel will be displayed. The maximum voxel and the location will be printed to the command prompt. Option (Click on "Options" menu) is provided in the figure 2.14 to enter the voxel (real world coordinates) instead of navigating around the brain.

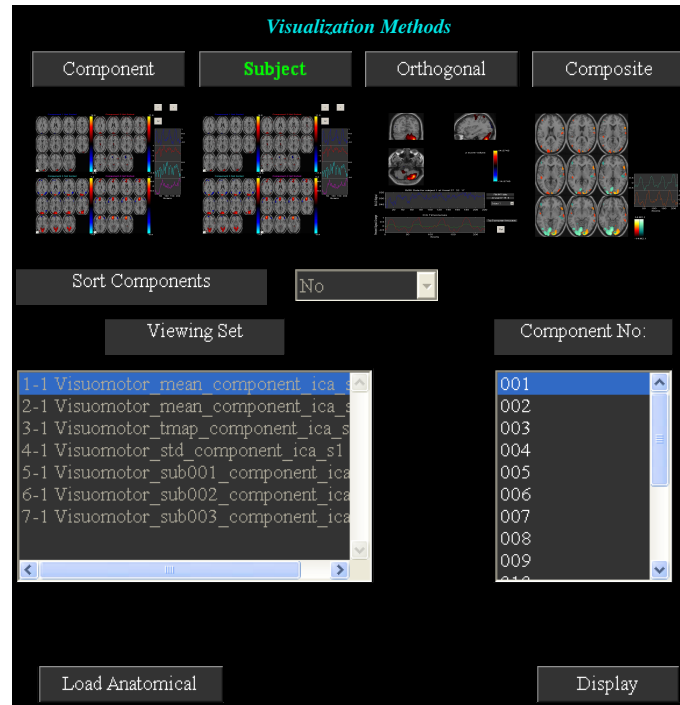


Figure 2.11: Selected options of the Subject Explorer method.

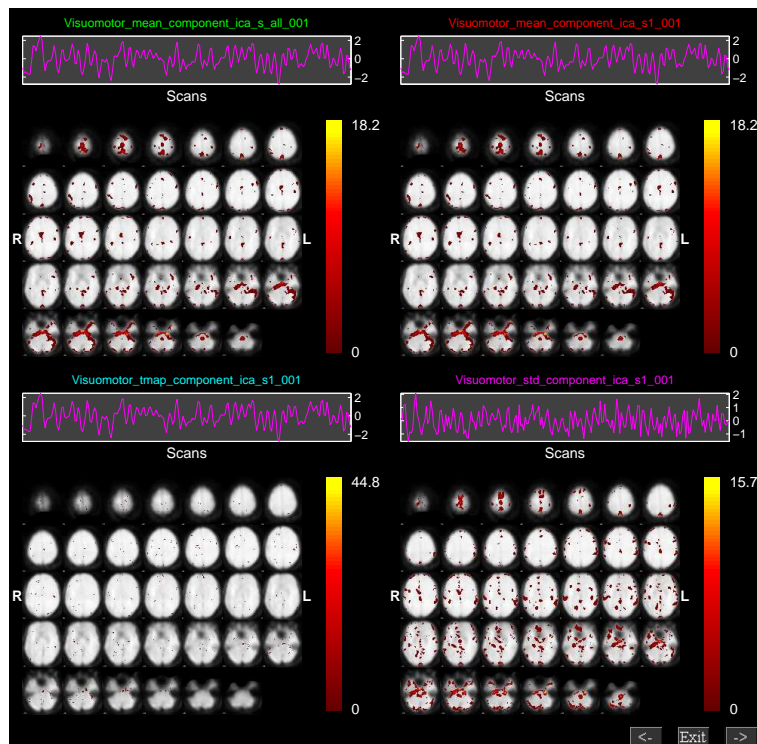


Figure 2.12: Component '001' of all the viewing sets. Here, only the first four components are displayed.

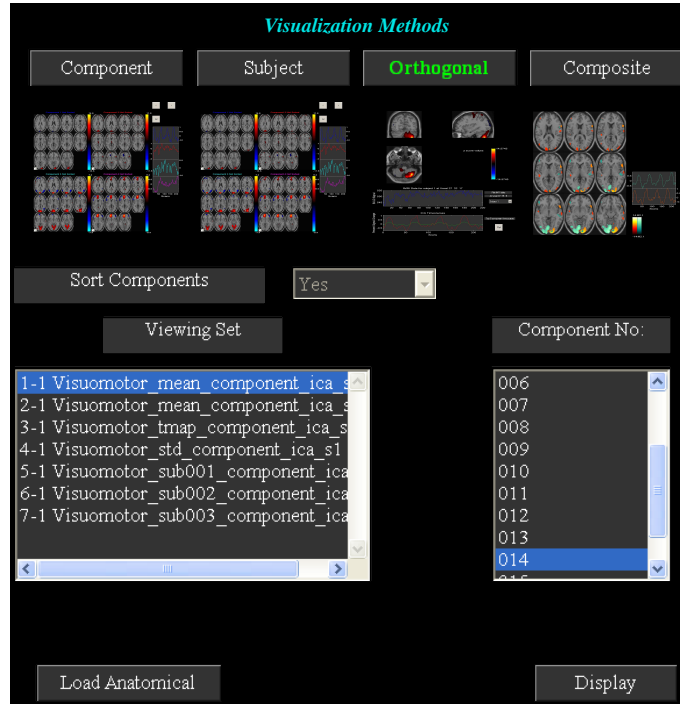


Figure 2.13: Figure shows the options of the Orthogonal Viewer method.

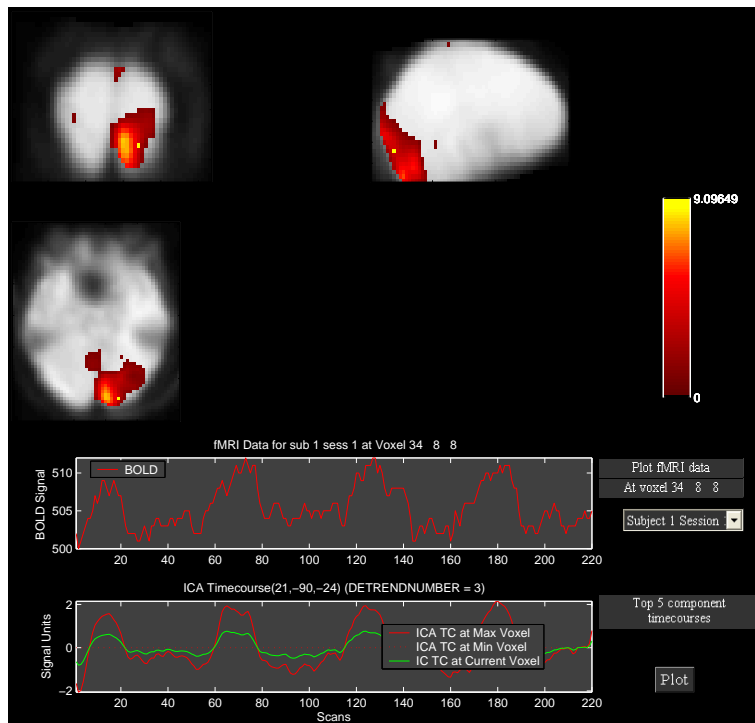


Figure 2.14: One of the task-related components plotted using Orthogonal Viewer.

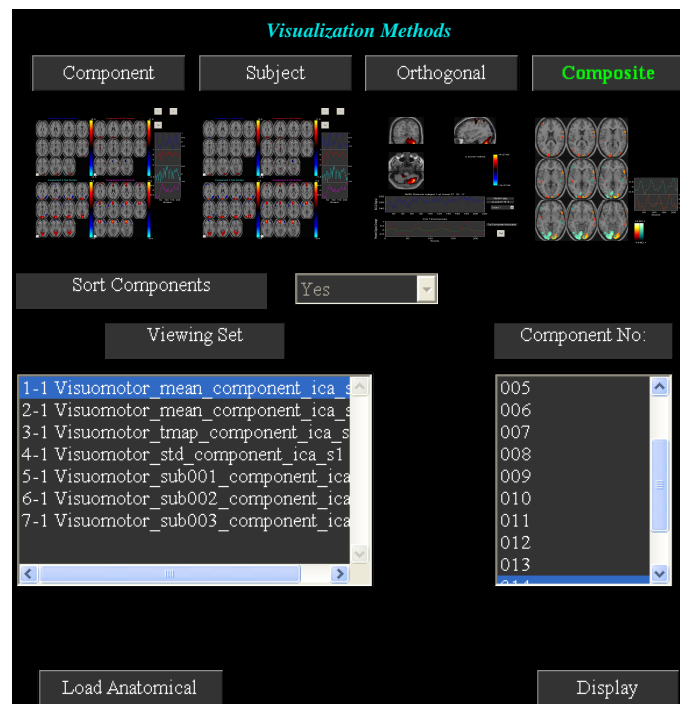


Figure 2.15: Selected options of the Composite Viewer method.

2.8.4 Composite Viewer

Composite viewer is used to look at multiple components of interest. Use the component explorer to find the task related components. In the 'Component Number' user interface control, select the two components that are task related. At most five different components can be overlaid on one another. Figure 2.15 shows the selected parameters. We used anatomical image `nsingle_subj_T1_2_2_5.nii` from folder `icatb/icatb_templates`. When you click *Display* button, figure 2.16 will open in a new window.

2.9 Sorting Components

Sorting is a way to classify the components. The components can be sorted either spatially or temporally. For every independent component spatial maps and time courses are generated. Temporal sorting is a way to compare the model's time course with the ICA time course whereas spatial sorting classifies the components by comparing the component's image with the template. When you click *Component* button, 'Sort Components' popup box will be enabled. Select 'Yes' for 'Sort Components'. Click *Display* button then a figure (Figure 2.17) will open in a new window. We have implemented three different types of sorting criteria like Correlation, Kurtosis and Multiple Linear Regression (MLR). MLR can be a very useful method in separating the two task related components. First, temporal sorting is explained followed by spatial sorting. The following are the steps involved in sorting components:

2.9.1 Temporal Sorting

Multiple regression sorting criteria is used to explain the temporal sorting. We select all data-sets (concatenated ICA time courses) and correlate with model time course. The regressors selected are "right*bf(1)" and "left*bf(1)" time courses. After the calculation is done, components are sorted based on the R-square statistic. The R-square statistic values and the slopes of the regressors are printed to a text file with the suffix `regression.txt`. Partial correlations and the slopes of the regressors are printed to a text file with the suffix `partial_corr.txt`. Figure 2.18 shows the components sorted based on the MLR sorting criteria in groupings of four. Here you can see that the first two components are task related. For a larger view of the time course plot (Figure 2.19) click on the time course plot in the main window. A list of menus is plotted on the time course plot. The explanation of each menu will be explained below:

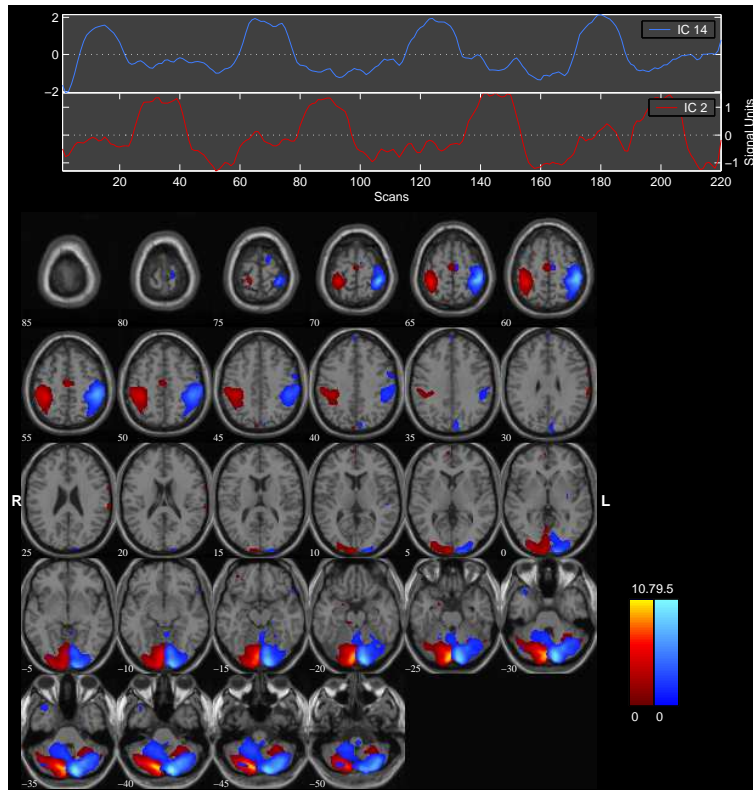


Figure 2.16: Task-related components overlaid on one another. At most five different components with different color bars can be overlaid.

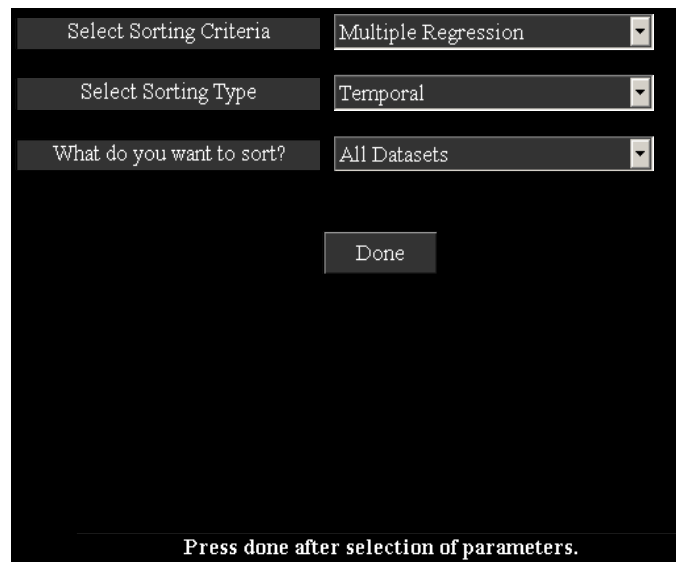


Figure 2.17: Selected parameters for sorting the components temporally.

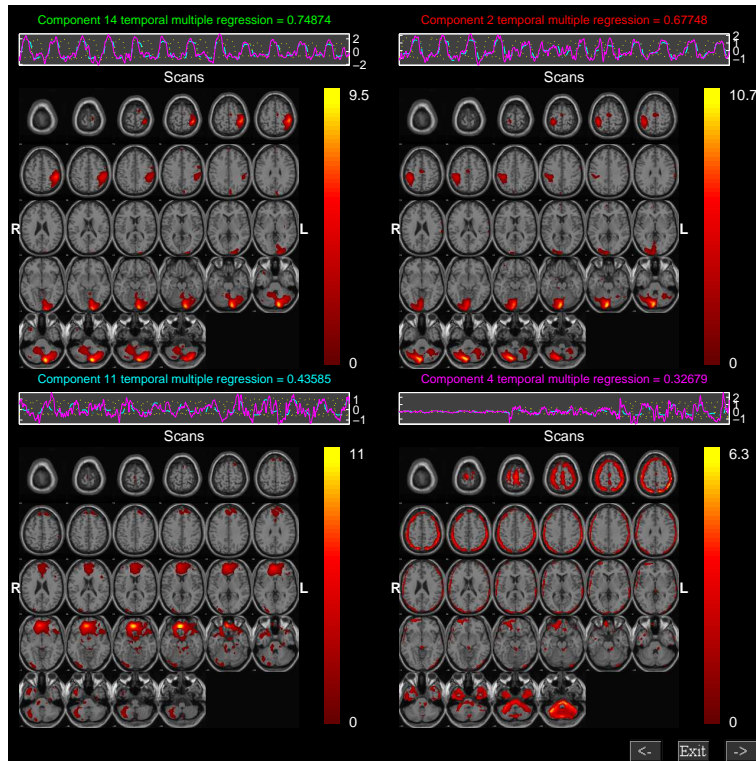


Figure 2.18: Components are sorted based on Multiple Regression criteria in groupings of four.

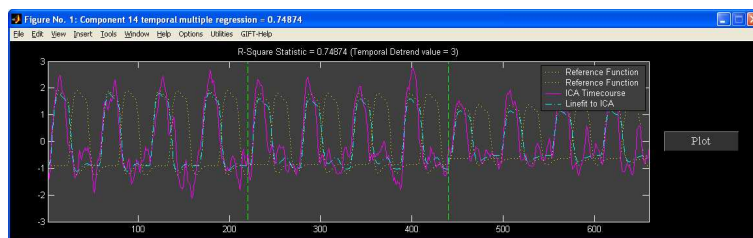


Figure 2.19: Expanded view of the "right*bf(1)" time course.

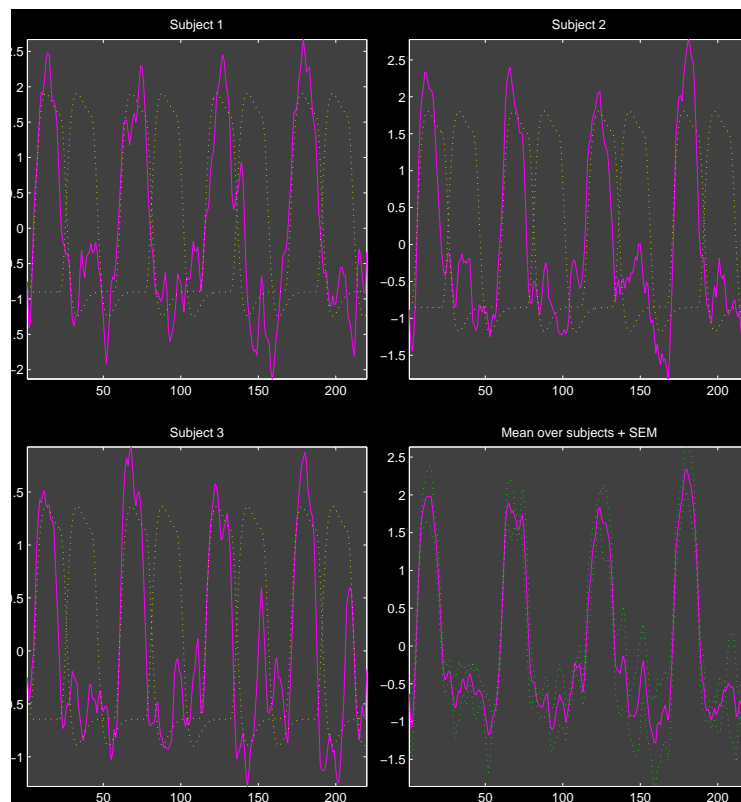


Figure 2.20: Figure shows the split of the concatenated time courses of all the data-sets. Mean is calculated over sessions and subjects as well as over all data-sets.

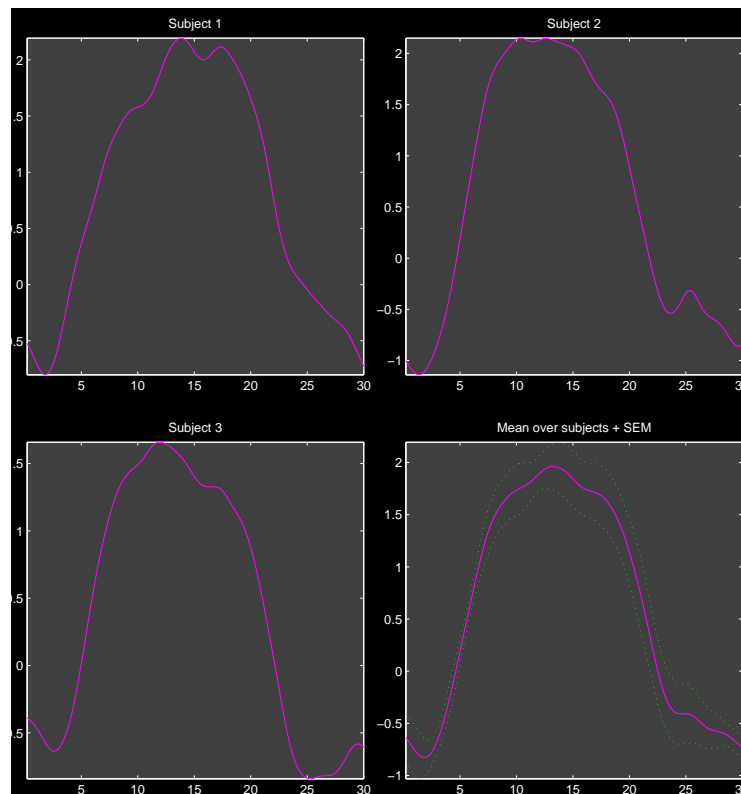


Figure 2.21: Figure shows the event related averages of the selected component for different subjects.

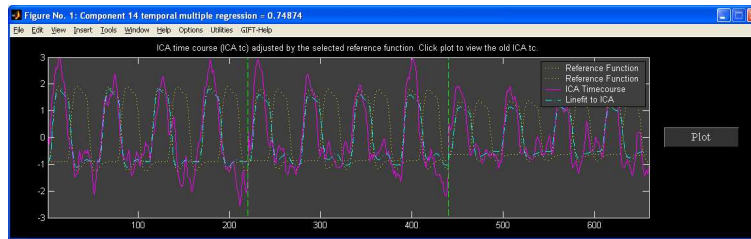


Figure 2.22: Expanded view of ICA time course after removing the variance of other than selected regressor ("left*bf(1)").

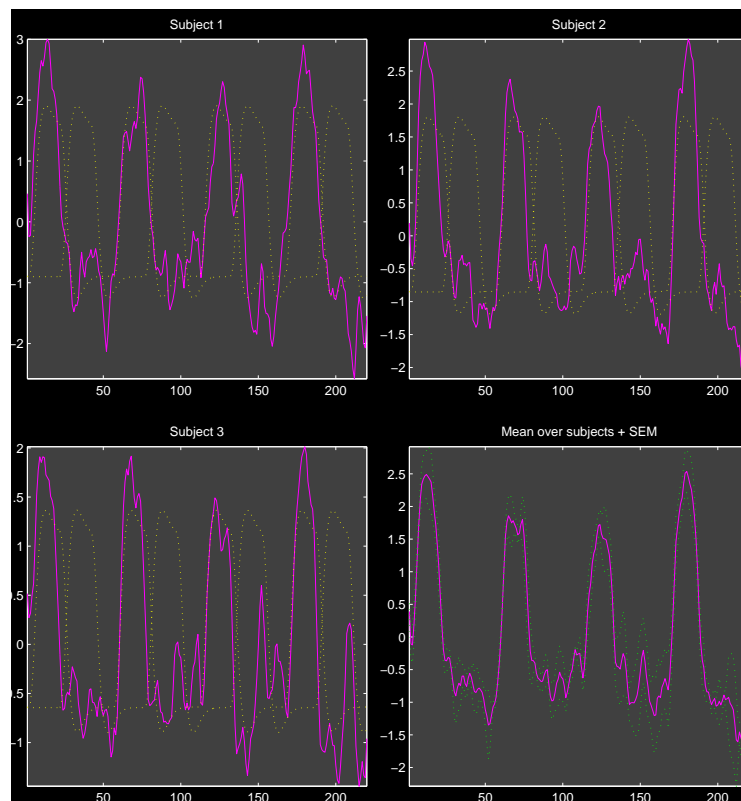


Figure 2.23: Figure shows the split of the concatenated time courses of all the data-sets after adjusting.

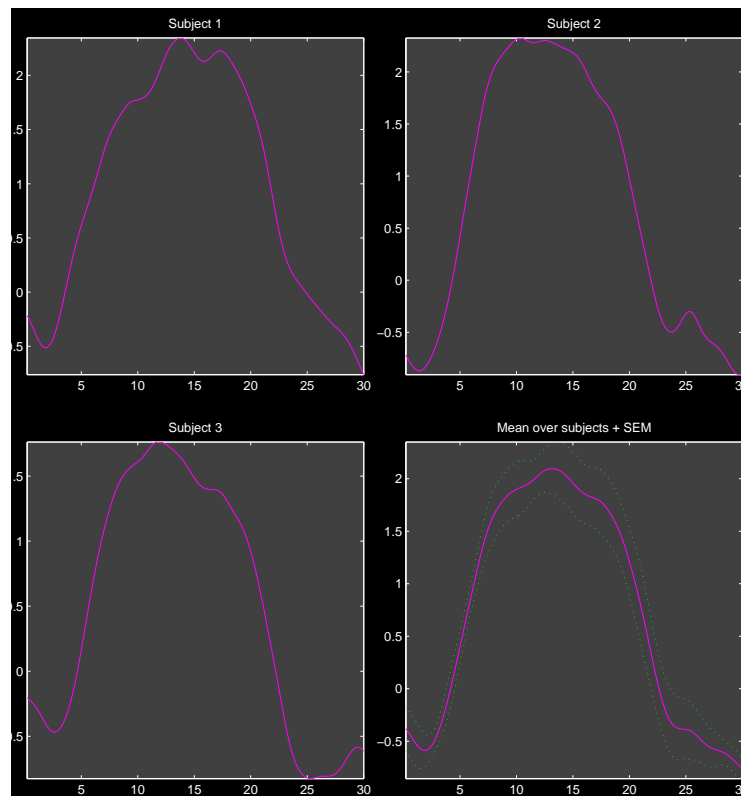


Figure 2.24: Figure shows the event averages of the adjusted ICA time courses.

- Utilities: Utilities contain sub menus like "Power Spectrum", "Split-time courses" and "Event Average". When you click "Split-time courses" sub menu, split of the time courses (Figure 2.20) will be shown. Click on sub menu "Event Average" and select "right*bf(1)" reference function to plot the event averages (Figure 2.21) of the ICA time courses. Explanation of the event average is given in Section 2.10.4.
- Options: "Options" menu has sub menus like "Timecourse Options", and "Adjust ICA".
 - Timecourse Options: When you click on sub menu "Timecourse Options", a new figure window will open that has options for detrending the ICA time course, model time course and options for event average. Explanation of this figure window is given in the Appendix 4.3. Leave the defaults as shown in the figure.
 - Adjust ICA: Option is provided in this sub-menu to remove the variance of other than selected regressor. When you click on sub menu "Adjust ICA", a list dialog box will open to select the reference function. For now select the "right*bf(1)" time course. The ICA time course is adjusted by removing the line fit of the model with the ICA time course where model contains nuisance parameters and other than the selected reference function. After the ICA time course is adjusted the plot is shown in the expanded view time course plot (Figure 2.22). When you click on the sub menu "Split-time courses" in "Utilities" menu, a new figure window (Figure 2.23) showing the split of the adjusted ICA time courses will be shown. Similarly click on sub menu "Event-Average" in "Utilities" menu and select the "right*bf(1)" reference function to view the event averages (Figure 2.24) of the new ICA time courses.

Note: Event average can also be done without sorting components (Section 2.10.4). Please see Appendix 4.2 for entering regressors through a text file for large data-sets.

Statistical Testing Of Images and Time Courses

Two Sample t-test: Two sample t -test can be used to compare a component between subjects or sessions. For example for a particular component session 1 images can be treated as group 1 and session 2 images as group 2. You can use these images in SPM to do two sample t -test ([5], pp 12-14).

Table 2.2: Table shows two sorted components parameters. Regression values for each component are shown followed by the beta weight of each condition with the ICA time course.

Component Numbers	7	3
Regression Values	0.74876543	0.67747138
Subject 1 Sn(1) right*bf(1)	2.3585817	0.55115175
Subject 1 Sn(1) left*bf(1)	0.097659759	2.158409
Subject 2 Sn(1) right*bf(1)	2.2468961	0.69208646
Subject 2 Sn(1) left*bf(1)	0.40444896	1.3045916
Subject 3 Sn(1) right*bf(1)	1.4413982	0.46694756
Subject 3 Sn(1) left*bf(1)	-0.045741345	1.6667349

Regression Parameters: After performing temporal sorting, the fit regression parameters are saved in a file with suffix `regression.txt`. This file can be quite useful in performing statistical tests to evaluate the task-relatedness of the components or to perform tests between components or groups. The text file contains a number for each component, regressor, subject and session (Table 2.2). The numbers represent the 'fit parameter' from the regression. In order to use these numbers it is important to scale the ICA data. For example, if 10 subjects were analyzed with ICA and two regressors were used to sort, then one can test the degree to which the average amplitude of group 1 was greater than the average amplitude of group 2 by computing a two-sample t-test between the 10 parameters for group 1 and the 10 parameters for group 2. A separate test can be done for each component, and the component which show a significant difference can be determined.

Note: We have provided an utility to do statistics on regression parameters. Please see section 2.10.6 for more information.

2.9.2 Spatial Sorting

Components can be spatially sorted by defining the regions of interest or a spatial template. Presently, there are four ways of sorting the components spatially like Multiple Regression, Correlation, Kurtosis and Maximum Voxel.

- 'Select Sorting Criteria'
 - The options available are 'Multiple Regression', 'Correlation', 'Kurtosis' and 'Maximum Voxel'. Kurtosis criteria does not need a template for sorting the components. Multiple Regression criteria can be used to select one or more templates.
- 'Select Sorting Type'
 - Options are 'Temporal' and 'Spatial'. Select 'Spatial' option.
- 'Select Template'
 - Template is used to define the regions of interest. For Maximum Voxel and Correlation criteria only one template should be used whereas for Multiple Regression more than one template can be selected. All the templates are located in `icatb_templates` folder.
- 'Select component set to sort'
 - Component set consists of individual subject's sessions, mean over sessions and mean of all subjects and sessions.
- Figure 2.25 shows the components of subject 1 session 1 sorted based on the MLR sorting criteria in groupings of four. The templates used are `RightTemplate.nii` and `LeftTemplate.nii`. Here, you can see that the first two components are task related.
- Figure 2.26 shows the components of subject 1 session 1 sorted based on the Maximum Voxel sorting criteria in groupings of four. The template used is `VisuomotorMask.img` in the analysis directory. The results are stored in a file with the suffix `max_voxel.txt`.

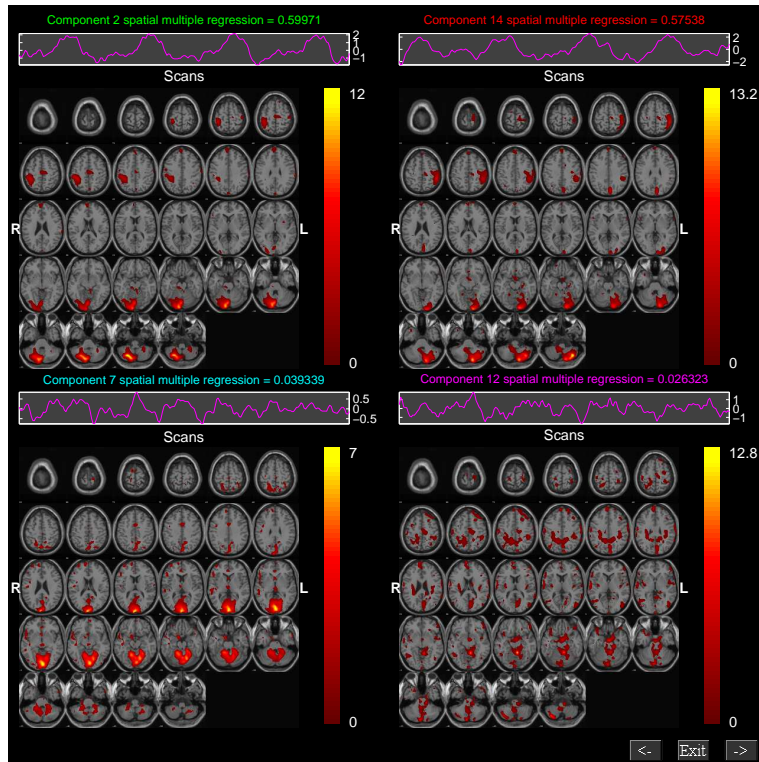


Figure 2.25: Figure shows the components spatially sorted using Multiple Regression sorting criteria.

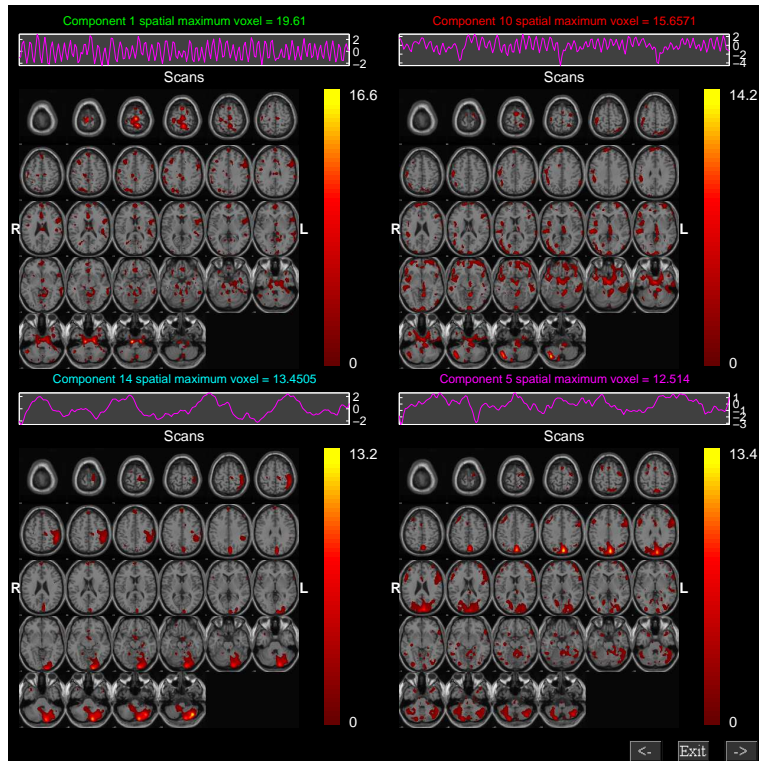


Figure 2.26: Figure shows the components spatially sorted using Maximum Voxel sorting criteria.

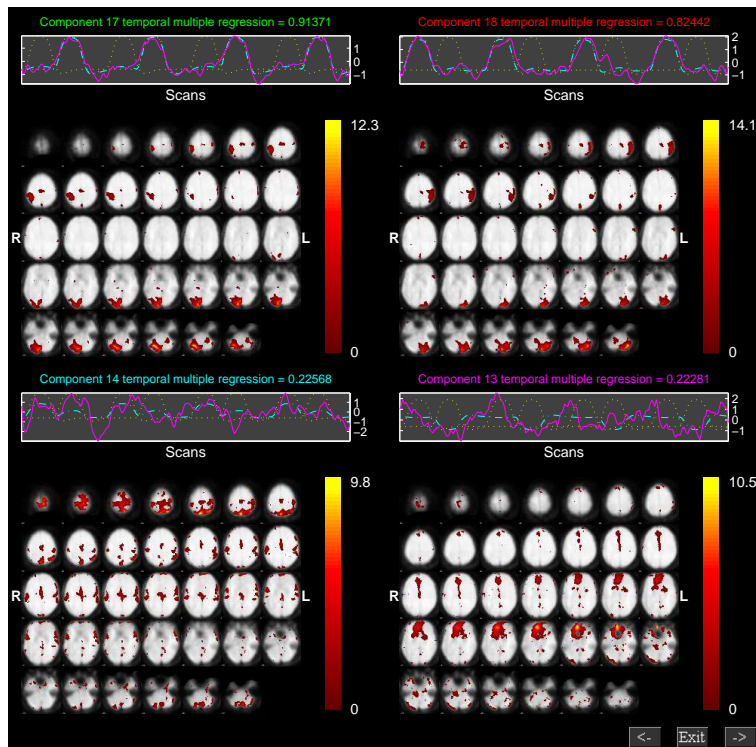


Figure 2.27: First four components that were sorted temporally using Multiple Regression and the regressors used are "right*bf(1)" and "left*bf(1)".

2.10 Utilities

2.10.1 Remove Components

Artifact signals like eye blinks, eye movements, muscle activity, etc make the detection of brain activity difficult. Therefore, signal processing techniques should be used to remove the artifacts from the data. Signal processing techniques like Independent Component Analysis (ICA) ([6]), Principal Component Analysis or Maximum Signal Fraction ([9]) are some of the techniques used to remove the artifacts from the data. Here, we discuss how ICA can be used to remove the components from the fMRI data using the Group ICA of fMRI Toolbox (GIFT).

- ICA was run on single subject single session and 20 components are extracted from the data. After the ICA analysis, the components extracted form a derived measure of the data. In order to remove artifacts, identify the components using any of the visualization methods used in the toolbox (Figure 2.1). Here, we remove the task-related components and these are identified by sorting components temporally.
- After identifying the components, use the "Utilities" drop down box and select "Remove Component(s)" entry. A figure window will open to select the parameter file used for the analysis. This is the same parameter file that you have used for running ICA on the fMRI data.
- After the parameter file is selected, a list dialog box (Figure 2.28) will open to select the components to be removed from the fMRI Data. We selected the first two components in the figure 2.27 to be removed from the data. The components will be removed from the data by zeroing out the corresponding columns of the mixing matrix and the rows of the spatial maps. The modified data is written to the selected output directory. The new set of images will have prefix R_.
- The modified data can now be analyzed using any toolbox that analyzes fMRI data. ICA is used to analyze the modified data and the components are sorted temporally using Multiple Regression and the regressors selected are "right*bf(1)" and the "left*bf(1)". Figure 2.29 shows that both left and right visual components are removed from the data.

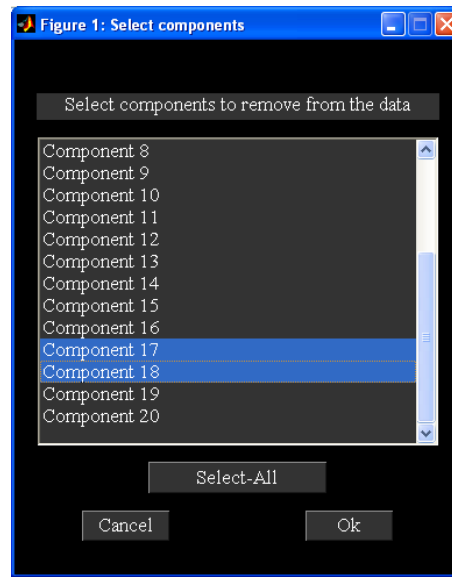


Figure 2.28: Dialog box showing component/component(s) to be removed from the data.

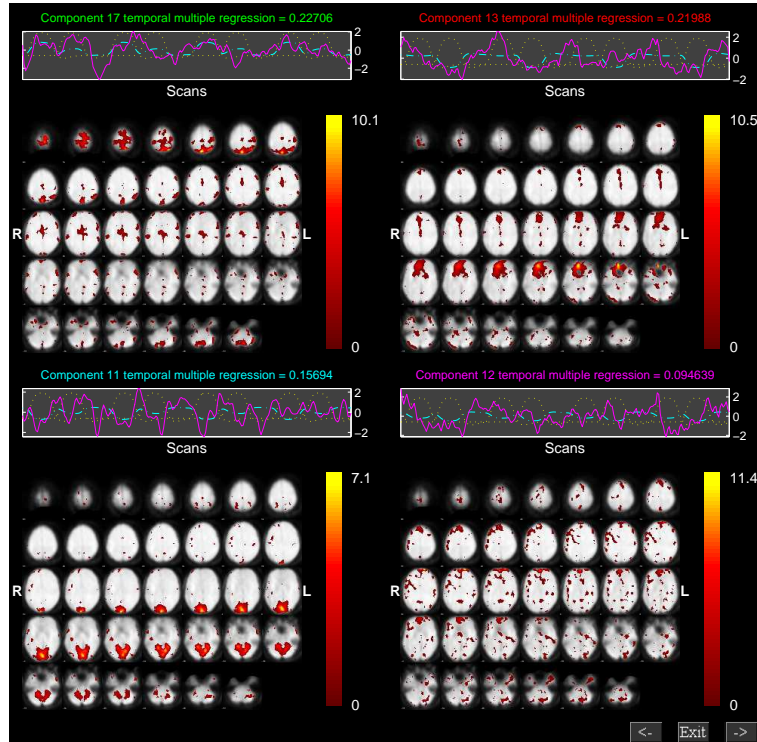


Figure 2.29: Results obtained after removing components from the data.

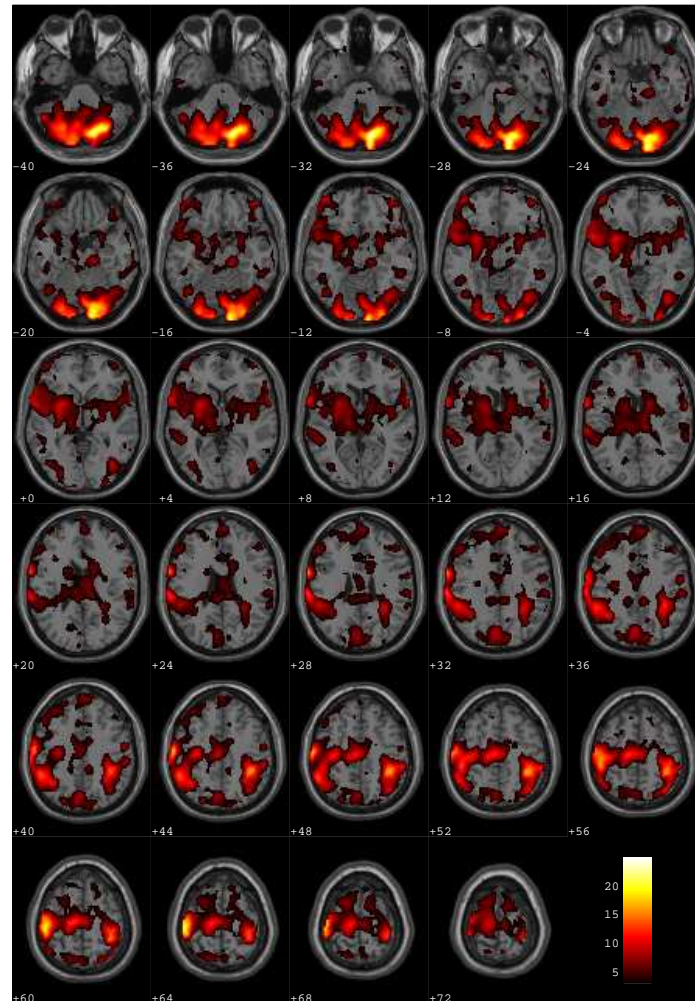


Figure 2.30: Left-right visual before removing the IC from the data.

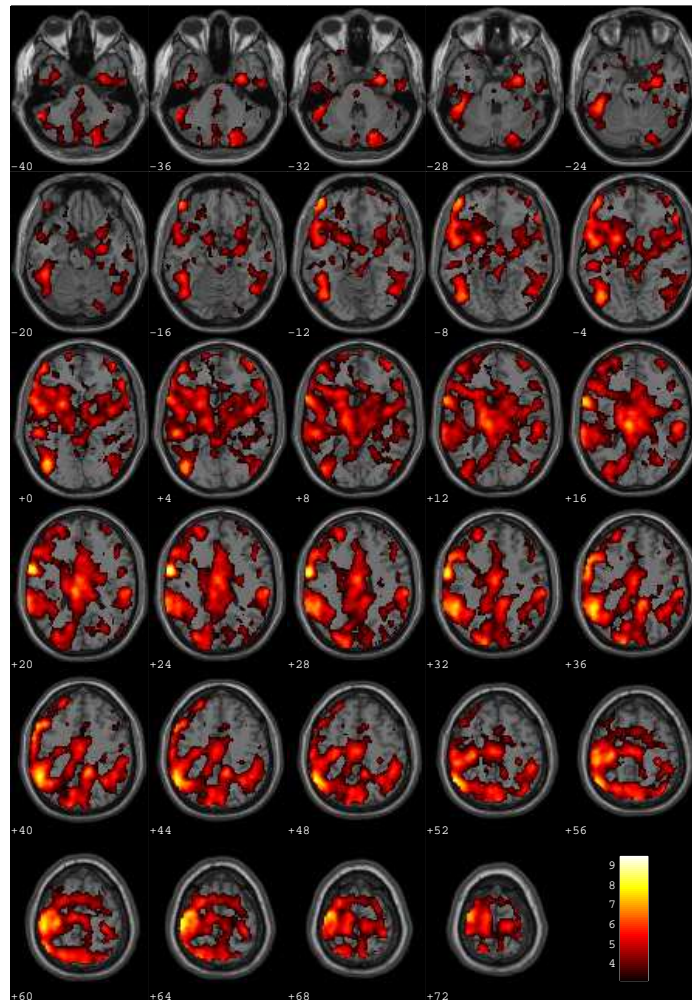


Figure 2.31: Left-right visual after removing the IC from the data.

- We did fMRI data analysis using SPM and the design matrix includes both "right*bf(1)" and "left*bf(1)" regressors. We show the t -maps of the left-right visual fields before (Figure 2.30) and after (Figure 2.31) removing the IC from the fMRI data. Both the results are obtained by applying a t -threshold of 3.3.

2.10.2 ICASSO

ICASSO toolbox ([7]) is used in GIFT to determine the reliability of ICA algorithm. ICA algorithm is run several times to determine the algorithmic reliability or stability. Reliable estimates correspond to tight clusters and unreliable ones do not point to any cluster. Figure 2.32 will open when you click on "Utilities" drop down box (Figure 2.1). The parameters in the figure 2.32 are as follows:

- "Select Mode" - Options available are 'RandInit', 'Bootstrap' and 'both'. The explanation of each option is given below:
 - 'RandInit' - Algorithm starts with different initial values.
 - 'Bootstrap' - Bootstrap technique is used.
 - 'both' - Uses both 'RandInit' and 'Bootstrap' options.

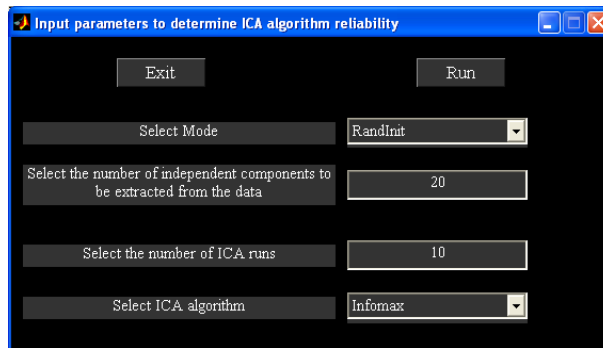


Figure 2.32: ICASSO GUI.

- "Select the number of independent components to be extracted from the data" - Enter number of components that will be extracted from the data.
- "Select number of ICA runs" - Number of times ICA algorithm will be run.
- "Select ICA algorithm" - Select the ICA algorithm. Some of the ICA algorithms are deterministic like ERICA, SIMBEC, JADE OPAC and AMUSE and therefore, will give errors when 'RandInit' option is used from "Select Mode".

We ran ICA 10 times on FastICA algorithm and extracted 20 components from the data. Figure 2.33 shows the results of FastICA algorithm. ICASSO results are written to a MAT file with suffix `_icasso_results.mat`. This MAT file contains the following variables:

- `iq` - Stability index.
- `sR` - Variable containing information about similarity measure, clustering and projection.
- `A` - Mixing matrix.
- `W` - Un-mixing matrix.
- `S` - Source signal.

2.10.3 Ascii to SPM.mat

Temporal sorting is done using a SPM (SPM2 or SPM5) design matrix. You can create SPM design matrix from a ascii file containing regressor time courses using "Utilities" (Figure 2.1) drop down box. Under "Utilities" drop down box select `Ascii_to_spm.mat`. The time course matrix is of dimensions m by n where m is the sum of the number of scans over sessions and n is the number of total regressors. Type `help icatb_formDesignMat` at the MATLAB command prompt for an example. See `icatb_templates` folder for example regressor data files.

2.10.4 Event Related Average

Event related average means the average of the events in the ICA time course. Events are calculated based on the onsets of the selected reference function or model time course. GIFT also provides the user to calculate the event average without sorting the components temporally.

Under "Utilities" drop down box (Figure 2.1) select "Event Average". This will open a figure window to select the parameter file. After selecting the parameter file you can specify the subjects, sessions, component and reference function required for event average.

Note:

- For large data-sets regressors can be entered through a sorting text file (Appendix 4.2). The first regressor specified in the text file will be used for event average.
- ICA time courses are not adjusted when calculating event average using "Utilities" drop down box in Figure 2.1. Please see Section 2.9.1 for calculating event average with the variance removal ("Adjust ICA") tool.

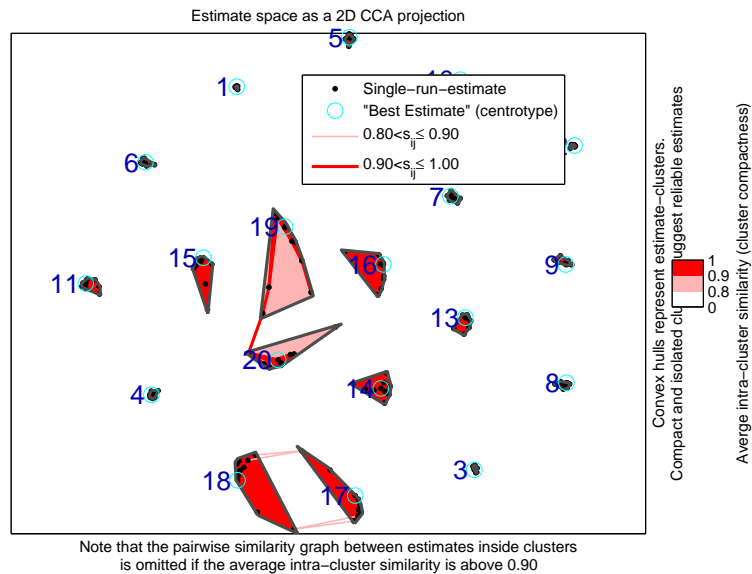


Figure 2.33: ICASSO Results.

2.10.5 Calculate Stats

Option is provided under "Utilities" drop down box (Figure 2.1) to calculate the statistics for the required data-sets over sessions or subjects or subjects and sessions. The resulting set of images can be used in SPM to do multi-group comparisons (between subjects or sessions) like two sample t -test.

Note: The images created using "Calculate Stats" under "Utilities" drop down box cannot be viewed using *Display GUI* and therefore, buttons like *Component Explorer*, *Composite Viewer* and *Orthogonal Viewer* in figure 2.1 must be used.

2.10.6 Stats On Beta Weights

This utility is provided to test the significance of component time courses by doing statistics on beta weights, which are obtained after doing temporal sorting on all data-sets. You need MATLAB R14 or higher version and statistics toolbox in order to run this utility. Design criteria like one sample t -test, two sample t -test, one way anova (groups), one way anova (regressors) and two way anova (groups, regressors) are provided. We use MATLAB statistics toolbox to do one sample t -test and two sample t -test on a condition. Linstats toolbox ([12]) is used to do one way anova, two way anova and anova contrasts. The following are the steps involved:

- When you select "Stats On Beta Weights" option in "Utilities" drop down box (Figure 2.1), a figure window (Figure 2.34) will open to select the regression parameters text file, which is obtained during temporal sorting of all data-sets.
- After selecting regression parameters text file, an option is provided to average beta weights across runs or sessions. GUI for doing stats on beta weights will open after selecting the answer in figure 2.35.
- The parameters in the figure are explained below:
 - "Select design criteria" - Options provided are 'one sample t -test', 'two sample t -test', 'one way anova (groups)', 'one way anova (regressors)' and 'two way anova (groups, regressors)'.
 - Groups Listbox - You can add groups using + button adjacent to groups listbox. Figure 2.36 will open to select subjects and sessions for that group. You can also use edit box below the listbox to select the subjects, if the subject numbers are stored in ascii file by entering `load('c:\sub1.asc')` or using any other valid MATLAB expression. Each group added will be an entry in groups listbox. You need to double click on the entry or press enter key to view or make any changes to the selected group. To remove the selected groups use - button adjacent to groups listbox.

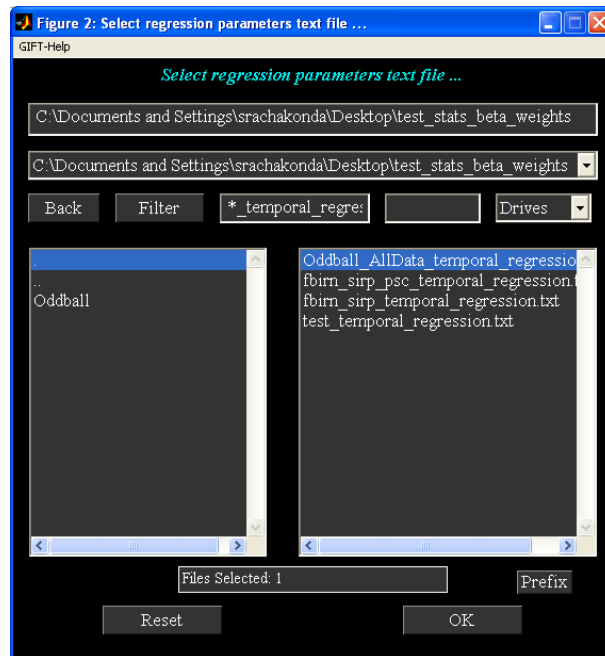


Figure 2.34: Figure window used to select the regression parameters text file.

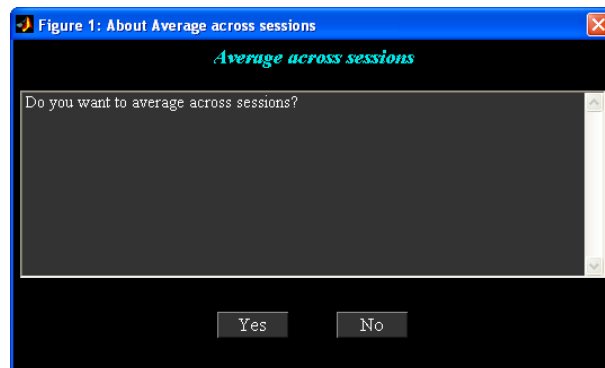


Figure 2.35: Option for averaging beta weights across runs or sessions.

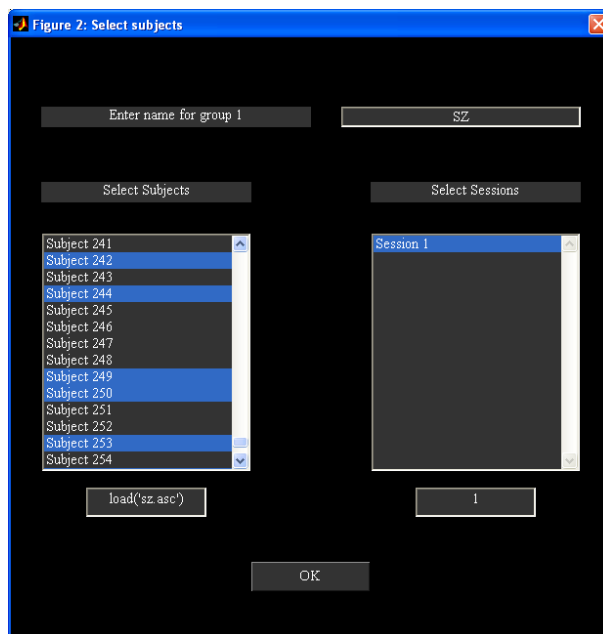


Figure 2.36: Figure window used to enter group name, subject and session numbers for a group.

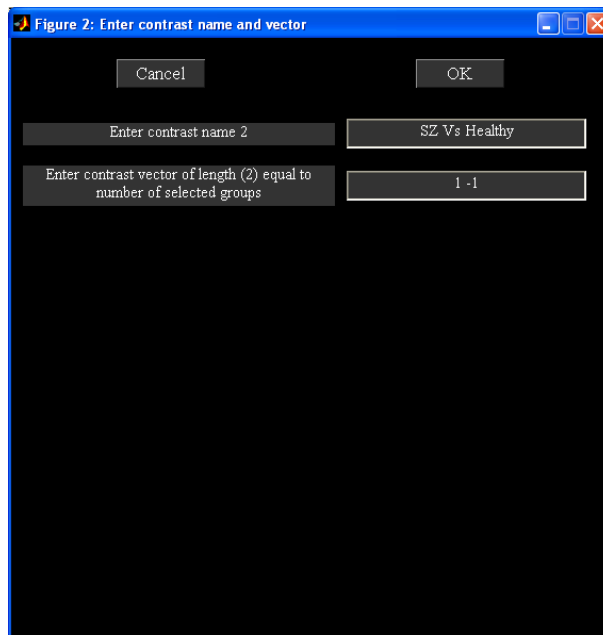


Figure 2.37: Figure window used to enter contrast name and vector.

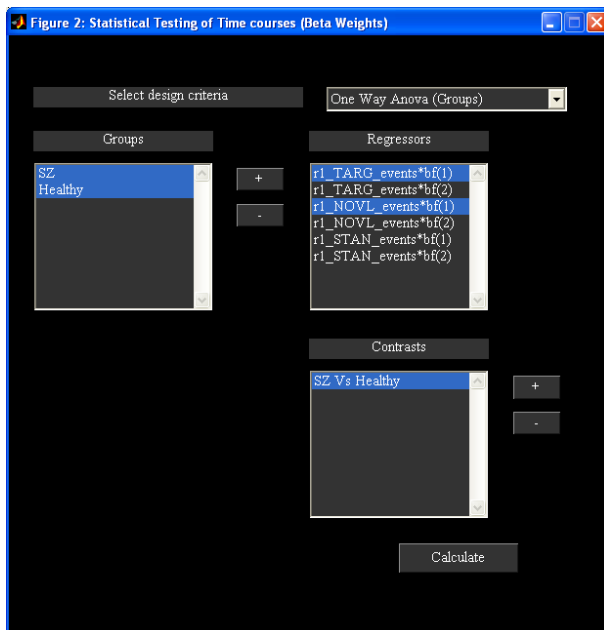


Figure 2.38: GUI shows the completed parameters.

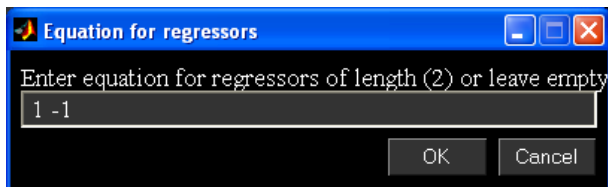


Figure 2.39: Option for doing operation on selected regressors.

- Regressors Listbox - You can select only one regressor for one sample t-test and two sample t-test whereas for anova you can select multiple regressors.
- Contrasts Listbox - This option is provided to do anova contrasts. To add a contrast use $+$ button (Figure 2.37) adjacent to contrasts listbox. Sum of contrast vector must equal to zero. The length of contrast vector to enter will change depending on the design criteria. The description is as follows:
 - * One way anova (Groups) - Length of contrast vector is equal to the number of selected groups. For example if there are 2 groups contrast vector must be of length 2.
 - * One way anova (Regressors) - Length of contrast vector is equal to the number of selected regressors. For example if there are 4 regressors contrast vector must be of length 4.
 - * Two way anova (Groups, Regressors) - Length of contrast vector is equal to the number of selected groups and regressors. For example if there are 2 groups and 4 regressors, contrast vector must be of length 6. First 2 entries correspond to groups and the next 4 entries correspond to regressors.

Note: You need to double click or press enter key on the selected contrast, if you decide to change the name of the contrast or value of the contrast vector.

- Calculate - Statistics are done after selecting the design criteria, groups, regressors and contrasts (Figure 2.38). All the results are printed to a file with suffix `summary.txt`. This file also contains mean and standard deviation for each condition of a group. An additional option is provided for one way anova (groups) to do operation on regressors (Figure 2.39) and treat it as a single regressor. For example, if you have selected two regressors like targets and novels to test the significance of components. You can directly subtract targets and novels and do a one way anova (groups) by specifying `[1 -1]` for the equation of regressors.

Note: Option is provided to do statistics on beta weights using a batch file. Please see section 2.11.1 for more information.

2.11 More Information

2.11.1 Batch Script

Analysis

Batch script is an alternative way to run the group ICA. `Input_data_subjects_1.m` and `Input_data_subjects_2.m` are provided as examples. The function for running the batch file is `icatb_batch_file_run`. The syntax for the function is as given below:

Type `icatb_batch_file_run(inputFile)` at the MATLAB command prompt. Where `inputFile` is the full file path for the input data of the subjects.

Batch script and input data files are located in folder `icatb_batch_files`. Parameters in the input file are as follows:

- `dataSelectionMethod` - There are two ways to select the data and the SPM design matrix (required for temporal sorting). Options are 1 and 2. Each option is explained below:
 - 1 - Data will be selected automatically if you specify the root folder for subjects and sessions, file pattern, a flag and file numbers to include. Options for flag are `data_in_subject_folder` and `data_in_subject_subfolder`.
 - * `data_in_subject_subfolder` - Data is selected from the subject sub-folders. Number of sessions is equal to the number of sub-folders containing the specified file pattern.
 - * `data_in_subject_folder` - Data is selected from the subject folders. The number of sessions is 1 and the number of subjects is equal to the number of subject folders containing the specified file pattern.

`spmDesignFilter` - Specify design matrix filter pattern here if you have selected `diff_sub_diff_sess` option for variable `keyword_designMatrix`. It looks for the design matrix in the respective subject folder or session folders.

- 2 - This option can be used when all the data is not in one directory. You need to specify the data directory for each subject and session followed by file pattern. The required variables are `selectedSubjects` and `numOfSess`. `selectedSubjects` contains the names (arbitrary) of subjects (s1 refers to subject1, s2 refers to subject 2, etc) and `numOfSess` contains the number of sessions. Subject 1 session 1 data information must be entered in variable `s1_s1` and subject 2 session 2 information must be entered in variable `s1_s2`. Design matrix information must be entered in `_designMat` variable.
- `keyword_designMatrix` - Design matrix is used for sorting the components temporally during display and will not be used during the analysis stage except for semi-blind ICA. There are four options like `no`, `same_sub_same_sess`, `same_sub_diff_sess` and `diff_sub_diff_sess`.
 - `no` - SPM design matrix is not specified for the analysis.
 - `same_sub_same_sess` - All the subjects and sessions will share the same regressors. Specify location of the SPM design matrix in variable `OnedesignMat`.
 - `same_sub_diff_sess` - Different regressors can be specified over sessions but same over subjects. Specify location of the SPM design matrix in variable `OnedesignMat`.
 - `diff_sub_diff_sess` - Different regressors can be specified over subjects and sessions.
- `outputDir` - Output directory of the analysis.
- `prefix` - All the output files will be pre-pended with this prefix.
- `maskFile` - Specify the location of the mask file or leave it as empty (Default mask).
- `numReductionSteps` - The number of reduction steps used and is dependent on the number of data-sets used. A maximum of three reduction steps is used.
- `doEstimation` - 1 means dimensionality estimation is done and PC step numbers are set to this value.
- `numOfPC1` - Number of PC for reduction step 1.
- `numOfPC2` - Number of PC for reduction step 2.
- `numOfPC3` - Number of PC for reduction step 3.
- `scaleType` - Options are 0, 1 and 2. 0 means don't scale, 1 means scale components to percent signal change and 2 means scale components to z-scores.
- `algoType` - Currently there are 11 ICA algorithms available in the GIFT toolbox. The algorithms are as follows:
 - 1 - Infomax
 - 2 - FastICA
 - 3 - ERICA
 - 4 - SIMBEC
 - 5 - EVD
 - 6 - JADE OPAC
 - 7 - AMUSE
 - 8 - SDD ICA
 - 9 - Semi-blind ICA
 - 10 - Constrained ICA (Spatial)
 - 11 - Radical ICA
- `refFunNames` - Reference function or regressor names to constrain ICA time courses. This is needed when using semi-blind ICA for the analysis.
- `refFiles` - Spatial reference or template files required to constrain the ICA source maps. This variable is required when Constrained ICA (Spatial) algorithm is used for doing ICA. Enter the files in a cell array.

Display

Display methods like Component Explorer, Composite Viewer and Orthogonal Viewer can be accessed through a batch script. The function for running the batch file is `icatb_batch_display`. The syntax for the function is as given below:

Type `icatb_batch_display(inputFile)` at the MATLAB command prompt
Where `inputFile` is the file containing the necessary display parameters.

The display parameters in the input file are as follows:

- `sourceDir` - Directory where fMRI images are located.
- `sourceFilePattern` - File pattern for fMRI images.
- `outputDir` - Directory where the component images are located.
- `compNumbers` - Component numbers to plot.
- `structFile` - All the component images will be overlaid on this anatomical file.
- `returnValue` - Variable used for plotting activations or deactivations or activations and deactivations on spatial map. There are four options like 1, 2, 3 and 4.
 - 1 - Positive and Negative
 - 2 - Positive
 - 3 - Absolute Value
 - 4 - Negative
- `convertToZ` - Convert image values to z -scores. Options are 1 and 0. 1 means convert to z -scores.
- `thresholdValue` - Z -threshold for spatial maps.
- `imagesPerFigure` - Number of images plotted per figure. Options are 1, 4, 9, 16 and 25.
- `anatomicalPlane` - Slice plane used for Component Explorer or Composite Viewer display methods. Options used are 'axial', 'sagittal' and 'coronal'.
- `displayType` - Display method to use. Options are 'component explorer', 'composite viewer' and 'orthogonal viewer'.

Stats On Beta Weights

Statistics on beta weights can also be done using a batch script. Example batch file (`Input_data_stats_beta_weights.m`) is in folder `icatb/icatb_batch_files`. The parameters in the batch file are as follows:

- `averageRuns` - A value of 1 means beta weights will be averaged across runs or sessions. The number of sessions will be set to 1.
- `desCriteria` - Design criteria used to test the significance of components. There are five options like:
 - 1 - One sample t-test
 - 2 - Two sample t-test
 - 3 - One way anova between the groups
 - 4 - One way anova between the regressors
 - 5 - Two way anova where groups and regressors are used as independent variables
- `groupInfo` - Variable containing information about the groups. This variable must be a cell array of size number of groups by 3. Each row must contain a name for group, subject numbers and session numbers. You can also use commands like `load('c:\healthy.asc')` in place of subject or session numbers.

- `selGroups` - Groups used for doing statistics on beta weights.
- `selConditions` - Regressors used for doing statistics on beta weights.
- `eq_regressors` - This option is provided when one way anova (groups) design criteria is selected. For example, you could do a one way anova between groups by subtracting one condition from another.
- `contrastNames` - A cell array containing contrast names.
- `contrastMatrix` - Each row of contrast matrix must equal to zero. The number of contrasts must equal the number of rows of contrast matrix. Contrast matrix depends on the design criteria selected. The description is as follows:
 - One way anova (groups) - The length of contrast vector must equal the number of selected groups like [g1, g2, etc].
 - One way anova (regressors) - The length of contrast vector must equal the number of selected regressors or conditions like [c1, c2, c3, c4, etc].
 - Two way anova (groups, regressors) - The length of contrast vector must equal the number of selected groups and regressors like [g1, g2, c1, c2, c3, c4, etc].

After entering the values for the variables in a batch file, type the following at the MATLAB command prompt:

```
icatb_statistical_testing_TC(regressParamFile, userInputFile);
```

Where `regressParamFile` is the regression parameters text file and `userInputFile` is the batch file.

2.11.2 Output Files Naming

- Parameter file - File used for storing parameters before and after the analysis. `_ica_parameter_info.mat` is the suffix used for parameter file.
- Reduction step file - After PCA, the information is stored in a MAT file. `_pca_r` is the suffix used for reduction files.
- ICA step file - After ICA, the information is stored in a MAT file with the suffix `_ica`. The aggregate images are written in Analyze or Nifti format with the suffix `_agg__component_ica_`.
- Back reconstruction step file - After back reconstruction step, the information is stored in a MAT file with the suffix `_ica_br`.
- Calibrate step file - After scaling, the information is stored in a MAT file with the suffix `_ica_c`.
- Component map file - Component maps are stored with the suffix `_component_ica_` in Analyze or Nifti format.
- Component time course file - Component time course for a particular data-set is stored with the suffix `_timecourses_ica_` in Analyze or Nifti format.

Chapter 3

Process involved in Group ICA

In this chapter, the process involved in group ICA is described. This chapter is divided into four sections. Each section is a step involved in group ICA. In Section 3.1, data reduction step is discussed. In Section 3.2, ICA process and the different ICA algorithms used in GIFT are explained. In Section 3.4, calibrating the components images and time-courses is discussed. Chapter 3.5 is used to describe the statistics involved in the toolbox.

3.1 Data Reduction

Data reduction is a step to reduce the size of the subject's functional data. Principal Components Analysis (PCA) is used as a technique to reduce the dimensions. A single subject might be reduced from $53 \times 63 \times 34 \times 220$ to $53 \times 63 \times 34 \times 50$. After each subject's functional data is reduced, the subjects are then concatenated into groups and put through another data reduction step. The number of subjects to put into each group is called partitions. The number of partitions is equal to one-fourth of the number of data-sets selected. Figure 3.1 shows three data-sets passed through two data reduction steps. The next stage is applying ICA on this reduced data-set.

3.2 Independent Component Analysis

Independent Component Analysis (ICA) is used to find independent components (ICs, Figure 3.2). GIFT includes a number of ICA algorithms that are available online and on ICALAB ([4]). Here we briefly introduce some of the ICA algorithms included in GIFT and cite relevant references. We also discuss a few parameters that affect their separation performance and some of the observations we obtained in our runs with simulated and real fMRI data ([13]) using Infomax, FastICA, Jade, SIMBEC and AMUSE.

3.2.1 Infomax

Infomax ([1]) maximizes the information transfer from the input to the output of a network using a non-linear function. A majority of applications of ICA to fMRI use Infomax since the sources of interest in this case are super gaussian in nature and the algorithm favors separation of super-gaussian sources. However, the artifacts present in fMRI data typically have sub-gaussian distributions. To improve separation of the mixture containing both super-gaussian and sub-gaussian sources, Extended Infomax ([19]) can be used.

To use Extended Infomax, choose '1' instead of the default of '0' while selecting ICA options for the Infomax algorithm. Based on our experiments, we observed that Infomax is a reliable choice for performing ICA on fMRI

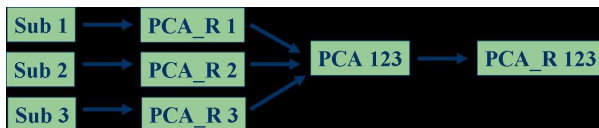


Figure 3.1: First each data-set is reduced using PCA where R term indicates reduced data. These three reduced data-sets are concatenated into one group (or partitions). The grouped data is passed through another data-reduction step.

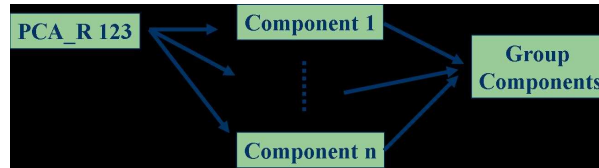


Figure 3.2: Reduced data-set from the last step in the data-reduction stage is used and ICA algorithm is applied. The components represent group and not of individual subjects. The next stage is Back-Reconstruction.

data. Z -scores for Infomax were higher than the other algorithms for the task-related source, indicating that Infomax achieves a higher contrast to noise ratio. Repeated runs showed that the changing initial random condition does not change results significantly. Infomax is much slower than the other algorithms listed in the toolbox.

3.2.2 Fast ICA

FastICA ([2]) maximizes the higher order statistics or negentropy of the output to maximize the non-gaussianity of the estimated sources using fixed point iterations. Two approaches exist: The symmetric approach estimates all the ICs in parallel and the deflationary approach estimates the ICs one at a time. The *tanh*, *pow3*, *gauss* and *skew* non linearities can be used ([3]).

The desired non-linearity can be chosen while selecting ICA options for FastICA. In our experiments, we used FastICA in the symmetric mode and compared results obtained using the non linearities *tanh*, *pow3* and *gauss*. On simulated data, this algorithm does better in terms of spatial correlation of the estimated sources with the original sources, for super-gaussian sources when compared to the other sources. For smaller number of components (number of components=5), FastICA with the *gauss* non-linearity provides better performance compared to the other two non-linearities for the gaussian and sub-gaussian sources. However, for a slightly larger simulated set (number of components=8), all three non-linearities result in very similar performance. Overall, the general performance of FastICA using *tanh* is better than the results obtained using the other two non-linearities. For actual fMRI data, we observed that for the transiently task-related component the spatial extent is slightly higher in FastICA as compared to Infomax. FastICA is slow when it has convergence problems and in this case it would be advisable to use the stabilized version of the algorithm. This can be done by changing the stabilization option to 'ON', while selecting the ICA options.

3.2.3 JADE OPAC

JADE (joint approximate diagonalization of eigenmatrices) ([8]), uses the Jacobi technique, to perform joint approximate diagonalization on fourth order cumulant matrices to achieve spatial independence among the sources. The version of JADE included in GIFT is MATLAB optimized with a reduced number of eigen matrices ([4]).

JADE is fast and the results are comparable to those obtained using Infomax. In the case of transiently task related components this algorithm shows spatial extent of activations to be higher than the obtained using Infomax.

3.2.4 SIMBEC

SIMBEC (simultaneous blind extraction using cumulants) ([17]), uses natural gradient ascent in a Stiefel manifold to simultaneously extract sources using a contrast function based on higher order cumulants with a learning rate that provides fast convergence.

In our simulations, we observed that for smaller number of components, SIMBEC performs well for the sources of interest. However, SIMBEC may prove useful to identify the sub-gaussian sources, i.e. artifacts in fMRI data as its performance for these sources was consistently observed to be very good. SIMBEC is also observed to be one of the faster algorithms.

3.2.5 AMUSE

AMUSE (algorithm for multiple unknown signal extraction) ([10]), is a second order Blind Source Separation algorithm that utilizes the structure within the data to obtain uncorrelated components. It performs singular value decomposition on the shifted cross-variance matrix and the shift should be chosen such that the autocorrelations of the sources at that shift are non-zero and as different from each other as possible. The default shift is set to '1'.

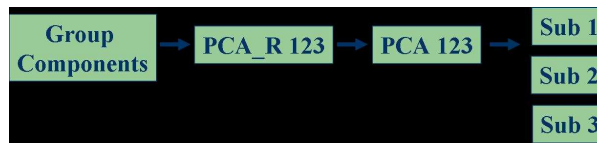


Figure 3.3: The individual subject components are calculated from the group components.

AMUSE is highly dependent on the differentiability of the spectra i.e. the autocorrelation of the sources should be different, for a given delay and its performance suffers greatly when this condition is not met. In our simulations, we observed that this condition is limiting for fMRI data especially when the number of sources is increased.

3.2.6 ERICA

ERICA (equivariant robust ICA) ([16]) novel blind source separation algorithms using cumulants) uses a cumulant based entropy cost function instead of a nonlinearity. The algorithm uses quasi-Newton iterations and converges at a saddle point of the entropy cost function. The algorithm achieves isotropic convergence, is fast and is independent of the source distributions regardless of Gaussian noise.

3.2.7 EVD

EVD (eigen value decomposition) algorithm ([14]) is based on second order statistics and very similar to AMUSE, the main difference being that EVD uses higher order correlations instead of second order correlations and is based on non-smooth optimization theory. To achieve separation, the source signals are required to have linearly independent higher self-correlation functions of even order. The algorithm does not assume non-gaussianity, non-stationarity and independence.

3.2.8 Constrained ICA (Spatial)

Constrained ICA ([15]) is a semi-blind ICA algorithm that utilizes prior information about desired sources as reference signals to extract only the desired sources. This algorithm uses fixed point iteration scheme (fICA-R) for optimizing the constrained ICA contrast function. Compared with the Newton-like ICA with reference, the fICA-R algorithm has the following advantages

- The fICA-R algorithm has no learning rate and is insensitive to the initialization.
- The fICA-R algorithm is simplified since no second derivatives are needed.

The resulting independent components have higher SNR ratio than traditional ICA algorithms.

3.3 Back Reconstruction

The components resulting from ICA represent group components and therefore, is the motivation for adding the Back Reconstruction (BR) step. BR uses the aggregate components of ICA and the results from data reduction step to compute the individual subject components. Figure 3.3 shows the process involved in the BR.

3.4 Calibrating Components

The image maps and time courses of components have arbitrary units after the back reconstruction step. Using the original functional data as a reference, the image maps and time courses can be scaled according to percent signal change. Calibrated image maps (Figure 3.4) are computed using the individual image maps and time courses.

3.5 Group Stats

Statistics is performed on the group of subjects. Mean, standard deviation and t-maps are calculated on the group of subjects. To give an illustration, let us say M data-sets are used and N components are extracted for each data-set.

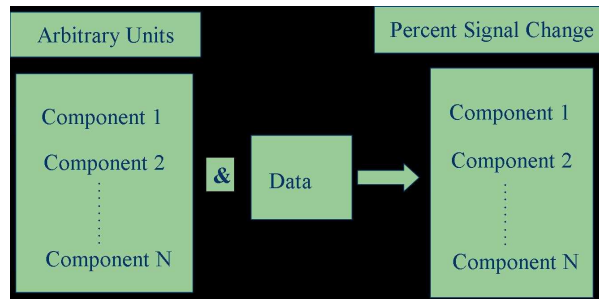


Figure 3.4: The individual subject components are calibrated to percent signal change using the original fMRI data as reference.

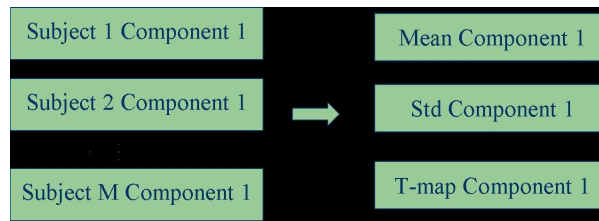


Figure 3.5: Mean, standard deviation and t -maps are calculated for each component over the data-sets used.

The mean, standard deviation and t -map are calculated for each component (Figure 3.5) over the number of data-sets used. This will produce N components for mean, standard deviation and t -map.

Bibliography

- [1] A. J. Bell and T. J. Sejnowski, "An information maximisation approach to blind separation and blind deconvolution," *Neural Comput.*, vol. 7, pp. 1129–1159, 1995.
- [2] A. Hyvarinen and E. Oja, "A fast fixed-point algorithm for independent component analysis," *Neural Comput.*, vol. 9, no. 7, pp. 1483–1492, 1997.
- [3] A. Hyvarinen, "Fast and robust fixed-point algorithms for independent component analysis," *Neural Networks*, vol. 10, pp. 626–634, 1999.
- [4] A. Cichocki, S. Amari, K. Siwek, T. Tanaka et al., "ICALAB Toolboxes," <http://www.bsp.brain.riken.jp/ICALAB>.
- [5] D. Glahn, *SPM fMRI Analysis*, 2000.
- [6] EEGLAB Toolbox, <http://scn.ucsd.edu/eeglab/>
- [7] ICASSO Toolbox, <http://www.cis.hut.fi/projects/ica/icasso/>
- [8] J. F. Cardoso and A. Souloumiac, "Blind beamforming for non gaussian signals," *IEE-Proc-F*, vol. 140, no. 6, pp. 362–370,
- [9] J. Knight, "Signal Fraction Analysis and Artifact Removal in EEG", Colorado State University, Fort Collins, Colorado, MS Thesis 2003.
- [10] L. Tong, V. C. Soon, Y. F. Huang and R. Liu, "Indeterminacy and identifiability of blind identification," *IEEE Trans. Circuits Sys.*, vol. 38, pp. 499–509, 1991.
- [11] M. J. McKneown, Scott Makeig, Greg G. Brown, Tzyy-Ping Jung, Sandra S. Kindermann, Anthony J. Bell and Terrence J. Sejnowski, "Analysis of fMRI Data by Blind Separation Into Independent Spatial Components", vol. 6, pp. 160-188, 1998.
- [12] M. Boedigheimer, Linstats Toolbox, <http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=13493>
- [13] N. Correa, T. Adali, Yi-Ou Li and V. Calhoun, "Comparison of Blind Source Separation algorithms for fMRI using a new MATLAB toolbox: GIFT," to be published in ICASSP 2005.
- [14] P. Georgiev and A. Cichocki, "Blind source separation via symmetric eigenvalue decomposition," *Proc. Sixth International Symposium on Signal Processing and its Applications*, Shangri-La Hotel, Kuala Lumpur, Malaysia, pp. 17–20, Aug. 2001.
- [15] Q. Lin, J. Liu, Y. Zheng, H. Liang and V. D. Calhoun, "Incorporating Spatial Hypotheses within ICA: An Efficient Fixed-point ICA-R Algorithm Applied to fMRI Data", Submitted to *NeuroImage* 2006.
- [16] S. Cruces, L. Castedo, A. Cichocki, "Novel blind source separation algorithms using cumulants," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, pp. 3152–3155, Istanbul, Turkey, June 2000.
- [17] S. Cruces, A. Cichocki and S. Amari, "Criteria for the simultaneous blind extraction of arbitrary groups of sources," *Int. Conf. Independent Component Analysis and Blind Sig. Separation*, San Diego, California, USA, 2001.
- [18] Statistical Parametric Mapping, <http://www.fil.ion.ucl.ac.uk/spm/>.

- [19] T. W. Lee, M. Girolami, and T. J. Sejnowski, "Independent component analysis using an extended Infomax algorithm for mixed sub-Gaussian and super-Gaussian sources," *Neural Comput*, vol. 11, no. 2, pp. 417–441, 1999.
- [20] V.D. Calhoun, T. Adali, G.D. Pearlson, and J.J. Pekar, "A Method for Making Group Inferences From Functional MRI Data Using Independent Component Analysis", *HBM*, vol. 14, pp. 140-151, 2001.
- [21] V.D. Calhoun, T. Adali, J.J. Pekar, and G.D. Pearlson, "Latency (in) Sensitive ICA: Group Independent Component Analysis of FMRI Data in the Temporal Frequency Domain", *NeuroImage*, vol. 20, 2003. 1993.
- [22] Y. Li, T. Adali, and V. D. Calhoun, "Sample Dependence Correction For Order Selection In FMRI Analysis," in *Proc. ISBI*, Washington, D.C., 2006.

Chapter 4

Appendix

4.1 Experimental Paradigms

The GIFT contains an example data-set which employs visuomotor paradigm (Figure 4.1). The paradigm contains two identical but spatially offset, periodic, visual stimuli, shifted by 20 seconds from one another. The visual stimuli were projected via an LCD projector onto a rear-projection screen subtending approximately 25 degrees of visual field, visible via a mirror attached to the MRI head coil. The stimuli consisted of an 8 Hz reversing checkerboard pattern presented for 15 seconds in the right visual hemi-field, followed by 5 seconds of an asterisk fixation, followed by 15 seconds of checkerboard presented to the left visual hemi-field, followed by 20 seconds of an asterisk fixation. The 55 second set of events was repeated four times for a total of 220 seconds. The motor stimuli consisted of participants touching their right thumb to each of their four fingers sequentially, back and forth, at a self-paced rate using the hand on the same side on which the visual stimulus is presented. fMRI data from this paradigm, when analyzed with standard ICA, separates into two different task-related components (one in left visual and motor cortex, one in right visual and motor cortex).

4.2 Defaults

Defaults used in GIFT toolbox are in `icatb_defaults.m` file. The variable names are in capital letters. Explanation of some of the variables is given below:

- `FUNCTIONAL_DATA_FILTER` - Variable used for listing functional data with the specified file pattern. You can also write Nifti images by changing the file pattern to `*.nii`.
- Colors: Colors are RGB values i.e., `[0 0 0]` means black. Background color can be changed by changing variable values. The variable names are listed below for figures and user interface controls:
 - `BG_COLOR`: Figure background color.
 - `BG2_COLOR`: User interface controls background color except push button.
 - `FONT_COLOR`: User interface controls font color except push button.
 - `BUTTON_COLOR`: Push button background color.

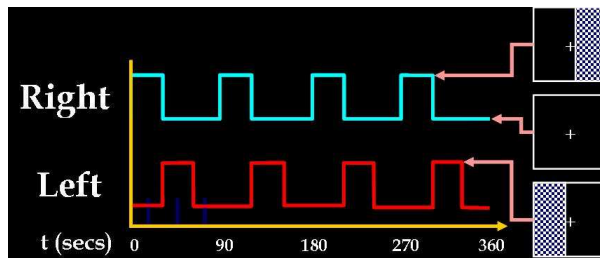


Figure 4.1: Visuomotor paradigm.

- `BUTTON_FONT_COLOR`: Push button font color.
- Fonts for user interface controls:
 - * `UI_FONTNAME`: Font name.
 - * `UI_FONTUNITS`: Font units.
 - * `UI_FS`: Font size.
- Display Defaults: Defaults in display GUI are plotted first. These can be changed by accessing variables like:
 - `SORT_COMPONENTS`: Options available are 'No' to 'Yes'.
 - `IMAGE_VALUES`: Options available are 'Positive', 'Positive and Negative', 'Absolute Value' and 'Negative'.
 - `CONVERT_Z`: Options are 'Yes' and 'No'.
 - `THRESHOLD_VALUE`: Threshold value to use voxels above or equal to the threshold
 - `IMAGES_PER_FIGURE`: Options are '1', '4', '9', '16', '25'.
 - `ANATOMICAL_PLANE`: Options are 'axial', 'sagittal' and 'coronal'.
- `DETRENDNUMBER`: Detrend defaults for ICA time courses can be changed by changing the variable value. Options are 0, 1, 2 and 3.
- Smoothing parameters: When `SMOOTH PARA` variable is changed to 'Yes' the time courses are smoothed by the value indicated in the variable `SMOOTHINGVALUE`.
- `FLAG_ACKNOWLEDGE_CREATORS`: Variable used to display acknowledgement for creators. You can turn off the dialog box by changing value to 'off'.
- `ICAOPTIONS_WINDOW_DISPLAY`: Variable used to display ICA options window. You can turn off the figure window by changing value to 'off'.
- `STORE_DIRECTORY_INFORMATION`: When this variable is changed to 'Yes', directories specified in cell array `DIRS_TO_BE_STORED` are stored in the file selection window.
- `ZIP_IMAGE_FILES`: When this variable is set to 'Yes', the component images will be compressed to a zip file based on their viewing set.
- `METHOD_ENTERING_REGRESSORS`: Variable has three options like 'AUTOMATIC', 'GUI' and 'BATCH' for temporal sorting. Each option is explained below:
 - 'AUTOMATIC' - Regressors in `SPM.mat` file will be used directly. For 'Different regressors over sessions' or 'Different regressors for subjects and sessions' session related regressors will be used.
 - 'GUI' - Figure window will open to select regressors.
 - 'BATCH' - Regressors can be entered using a text file by specifying the file name. This is the best way to enter regressors if you have many data-sets and selected 'Different regressors over subjects and sessions'.

Note:

 - * `TXTFILE_REGRESSORS` is the variable used for specifying the sorting text file.
 - * Sorting text file specified in defaults will not be used if you had changed using *Display GUI*.
- `FLIP_ANALYZE_IMAGES`: Flip parameter for the analyze images. Default value is 0.
- `NUM_RUNS_GICA`: Number of times ICA will be run. Default is 1.
- `DEFAULT_MASK_OPTION`: Mask is calculated by doing a Boolean AND of voxels that surpass or equal the mean. By default first file for each subject is used. You can use all files by changing variable value to 'all_files'.
- `OPEN_DISPLAY_GUI`: Option is provided to open the display GUI automatically after the analysis. You can turn off the option by changing variable value to 0.

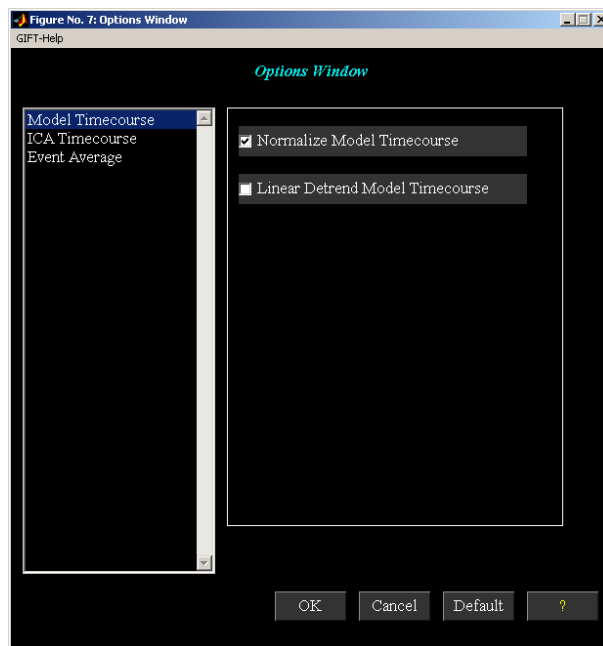


Figure 4.2: Time course options.

4.3 Options Window

When you click "Timecourse Options" sub menu in Options Menu (Figure 2.19), options window (Figure 4.2) appears with the defaults used from the `icatb_defaults.m`. There are three categories in the left list box like Model Time course, ICA Time course" and Event Average. The explanation of each option in the list box is explained below:

- Model Time course:
 - Normalize Model Timecourse: Model time course is normalized and the ICA time course is used as a reference.
 - Linear detrend Model Time course: Linear trend is removed in model time course.
- ICA Time course:
 - Detrend Number: ICA time course is detrended with the number specified in popup control.
 - Flip ICA Time course: ICA time course is flipped.
 - Smoothing Parameter: ICA time course is smoothed if the option is set to 'Yes'.
 - Smoothing Value: ICA time course is smoothed using the value specified in edit control.
- Event Average:
 - Window size in seconds: Window size to display the event average of ICA time course.
 - Interpolation factor: Factor to interpolate the ICA time course.

When you click *OK* button, the values in the controls are used and the time courses are plotted in the expanded view of time course (Figure 2.19). *Cancel* button closes the figure window. *Default* button uses the defaults from `icatb_defaults.m` file.

4.4 Interactive Figure Window

Interactive figure window is used to select a directory, file or files for performing the group ICA. Figure 4.3 shows the file selection window with the ability to select more than one file. The explanation for each user interface control in the figure is explained below:

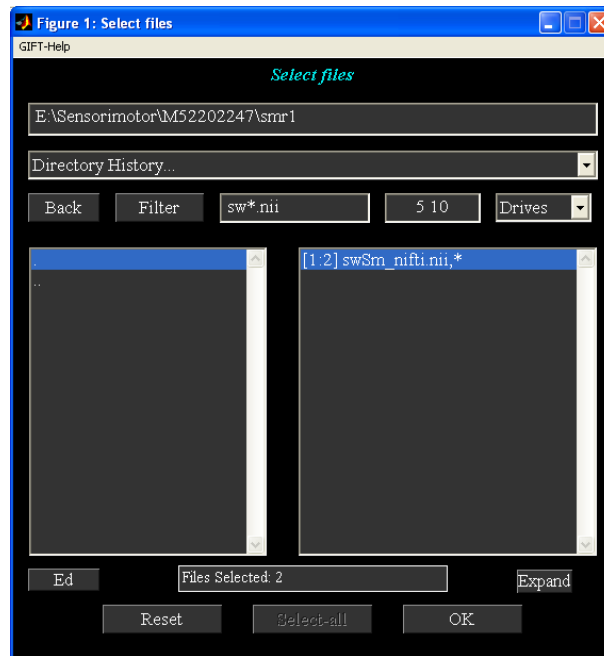


Figure 4.3: Figure window used to select a directory, file or files.

- Directory text box shows the current directory or you can edit the text box to specify the directory. When enter key is pressed the sub-folders on the left list box are displayed and also the files or folders on the right list box are displayed.
- Directory History shows the history of the directories recently visited in the MATLAB along with the path for the GIFT.
- *Back* button is used to return to the previous folder in the same drive.
- *Filter* button resets the filter text box to ***.
- Filter text box filters the files based on the text entered. If you want to list files with different patterns, separate them with semi-colon as delimiter.
- File numbers edit box adjacent to filter text box is used to enter the file numbers for a 4D Nifti file.
- Drives show the drives available for the Windows operating system except the floppy drive. For other operating systems the root directory is displayed.
- Sub-folders list box shows sub-folders in the left list box.
- Files or sub-folders list box plotted on the right shows files in the current directory. In case of directory selection, sub-folders in the current directory will be displayed.
- *Ed* button is used to edit the files.
- *Reset* refreshes the drives in the Windows operating system and de-selects all the files or directories previously selected. It is also used to enable the *Select-all* button.
- *Expand* lists files in expanded mode.
- *Select-All* selects all the files in the current working directory depending on the filter specified.
- *OK* closes the figure window when the entries on the right list box are selected.

4.5 GIFT Startup File

You can add GIFT paths in a `gift_startup.m` file in the sequence you want and add this M file on MATLAB path. `gift_startup.m` will be executed when you run `gift.m` file. You can add the following statements in `gift_startup.m` file:

```
addpath 'C:\giftv1.3c\icatb'  
addpath 'C:\giftv1.3c\icatb\icatb_analysis_functions'  
addpath 'C:\giftv1.3c\icatb\icatb_analysis_functions\icatb_algorithms'  
addpath 'C:\giftv1.3c\icatb\icatb_analysis_functions\icatb_algorithms\bin'  
addpath 'C:\giftv1.3c\icatb\icatb_analysis_functions\icatb_algorithms\icatb_semiblindInfomax'  
addpath 'C:\giftv1.3c\icatb\icatb_batch_files'  
addpath 'C:\giftv1.3c\icatb\icatb_display_functions'  
addpath 'C:\giftv1.3c\icatb\icatb_helpManual'  
addpath 'C:\giftv1.3c\icatb\icatb_helper_functions'  
addpath 'C:\giftv1.3c\icatb\icatb_io_data_functions'  
addpath 'C:\giftv1.3c\icatb\icatb_spm5_files'  
addpath 'C:\giftv1.3c\icatb\icatb_scripts'  
addpath 'C:\giftv1.3c\icatb\icatb_spm2_files'  
addpath 'C:\giftv1.3c\icatb\icatb_templates'
```