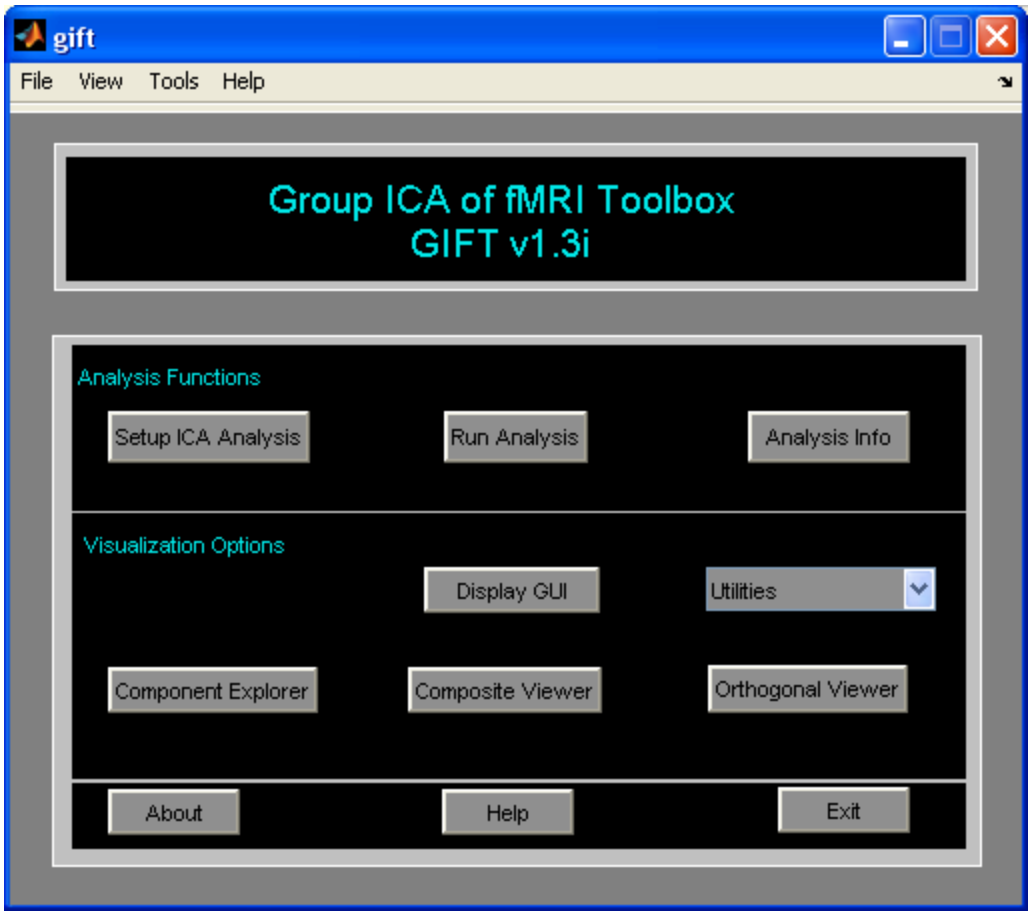
Group ICA of fMRI Toolbox (GIFT) Manual

Srinivas Rachakonda[1], Eric Egolf[2], Nicolle Correa[3] and Vince Calhoun[14]
July 08, 2011



[1] *The MIND Research Network, 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

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. In section 2.14, Source Based Morphometry is discussed.

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 ([28]) 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) ([20]) 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 [30], a model was proposed to extend ICA to group studies. Group ICA can be run using a batch script (Section 2.13.1) 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

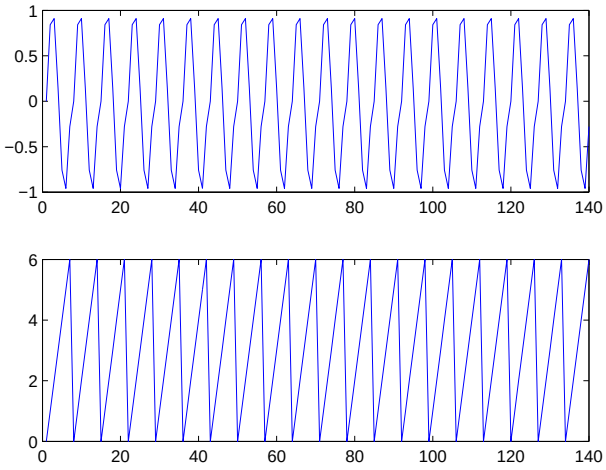


Figure 1.1: Unknown source signals.

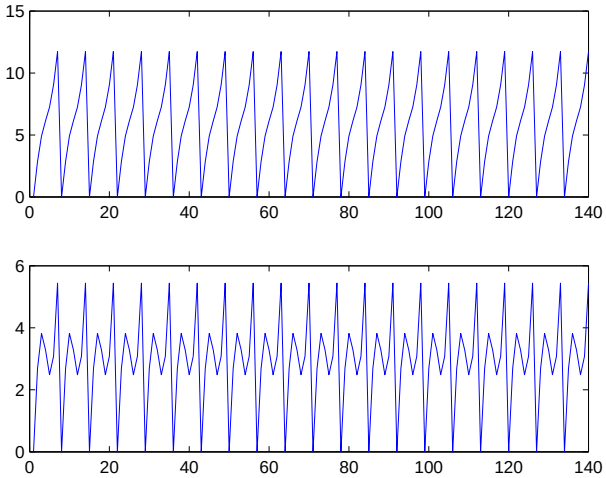


Figure 1.2: Mixed signals.

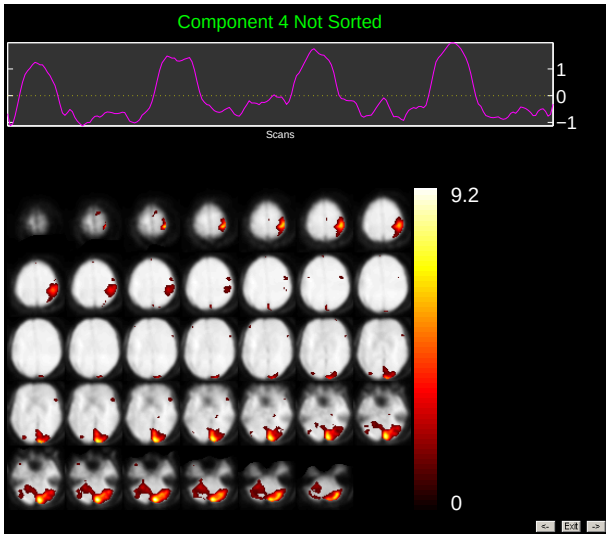


Figure 1.3: A component consists of a time course and a spatial map.

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 ([30]). 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`

Please send comments and bug reports to:

`vcalhoun@unm.edu or srachakonda@mrn.org`

2.1 Installing GIFT

Unzip the file `GroupICATv2.0e.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 by default will be set at the bottom of the MATLAB path. You can create `gift_startup.m` (Appendix 4.6) 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 2.13.1). Batch script is very useful for running large data-sets.

Note:

- GIFT toolbox can also be invoked by using the statement `groupica` and clicking on the *fMRI* button or by typing `groupica fmri` at the MATLAB command window.
- 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 [30] 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 Changes from the release GroupICATv2.0d

- Mancovan toolbox is integrated in GIFT. Mancovan toolbox does multivariate tests on ICA timecourse spectral power, spatial map intensity and functional network connectivity to determine the significant covariates which will be used later in the univariate tests. Please see [12] for more information.
- SBM toolbox is added to do source based morphometry. Source based morphometry is a useful tool to study the gray matter differences between patients and controls ([1] and [18]).
- Options are provided in the Group ICA Toolbox to write only the necessary output components information which will be used later in the display.
- SPM8 volume functions are used to read and write image data.
- While using scaling timecourses option in GIFT, average of top 1% voxels is used instead of the maximum spatial intensity.
- Default mask used in the dimensionality estimation step is generated using all subjects in the analysis.

2.5 GIFT Updates

We post the updates to the software that contain new features or any bug fixes in the updates section of the project web page. Please see `Updates_Readme.txt` for more details.

2.6 Things to do before configuring the analysis

2.6.1 Compiling MEX files

- SPM MEX binaries - We use SPM8 volume functions to read and write image data.
- Group ICA MEX binaries - C-MEX files are provided for computing the eigen values of a symmetric matrix. The compiled MEX binaries are used only when you select packed storage scheme for computing covariance matrix during the second or third data reduction stage.

To compile the MEX source code, type `icatb_compile_mex_files` at the MATLAB command prompt. Please see below to copy the SPM MEX binaries manually if the source code fails to compile:

- Copy SPM8 MEX binaries like `spm_bwlabel*`, `spm_existfile*`, `spm_sample_vol*`, `spm_slice_vol*` to the directory `icatb\icatb_spm8_files` and prefix them with `icatb_`.
- Copy SPM8 MEX binaries like `spm8\@file_array\private\*` to `icatb_spm8_files\@icatb_file_array\private` and prefix them with `icatb_`.

  Note: * refers to the MEX file extension on your Operating System. Type `mexext` on the MATLAB command window to get the MEX file extension on the Operating System you are working on.

2.6.2 Memory Requirements

Since PCA and ICA are Multi-variate approaches unlike General Linear Model, there are some memory requirements to do the group ICA analysis. We added a script `icatb_mem_ica.m` which will give a close estimate of RAM required to do the group ICA analysis. Please enter the parameters like number of voxels, time points, subjects, sessions, components and reduction steps in the input parameters section of the script and run the script to get an approximate amount of RAM required.

2.6.3 Organizing Data

Organizing data reduces the amount of selection. GIFT has two ways to enter the data for GUI and four ways to enter the data using the 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 2.13.1)is recommended.
- Third method uses regular expressions to get the subject and session directories. This option is useful in matching directories that have nested paths. However, this option still requires a common file pattern for all the subjects. Please see Section 2.13.1 for more information.
- Option is provided to paste the file names in the fourth method (Section 2.13.1).

2.6.4 Defaults

Configure the specified defaults before using setup ICA as needed.

- `FUNCTIONAL_DATA_FILTER` - Set the variable to *.nii if you would like to write components as 4D Nifti files.
- `ZIP_IMAGE_FILES` - By default, components will be compressed in zip format. If you want to turn off the compression, set variable value to 'no'. Uncompressed data is faster to load in memory.
- `SPM_STATS_WRITE_TAL` - If you plan to do one sample t-tests on components over subjects using SPM5 or SPM8, set variable value to 1.
- `CENTER_IMAGES` - By default, subject component spatial maps after the scaling components step is centered based on the skewness of the distribution of the mean component maps. If you want to turn off this option, set the variable to 0.
- `MAX_AVAILABLE_RAM` - This variable is used during the run analysis step and best PCA settings for each option (maximize performance or less memory usage) are used. Set this variable value to the maximum available RAM.
- `WRITE_ANALYSIS_STEPS_IN_DIRS` - Organize analysis results in directories. When this variable is set to 1, analysis steps are saved in separate directories. The directories naming are as follows:
  - Data reduction files are stored in `*_data_reduction_files`
  - ICA files are stored in `*_ica_files`
  - Back-reconstruction files are stored in `*_back_reconstruction_files`
  - Scaled components are stored in `*_scaling_components_files`
  - Group stats are stored in `*_group_stats_files`
- `CONSERVE_DISK_SPACE` - Set the variable value as needed. The following are the options:
  - 0 - Analysis runs faster with this option. All the intermediate analysis files are written.
  - 1 - Only the required analysis files are written which will be used during the post-processing (display, components sorting, remove components, etc).
  - 2 - All the intermediate analysis files (Data reduction, back-reconstruction, scaled components MAT files) are cleaned up after the end of the group stats step. Only the basic post-processing steps like display and components sorting will work with this option.
- `DEFAULT_MASK_OPTION` - By default, first file of each subject is used to generate the default mask. If you want to use all files, set the variable value to `all_files`.
- `REMOVE_CONSTANT_VOXELS` - Constant voxels are removed in the fMRI data when this variable is set to 1.
- `DEFAULT_MASK_SBM_MULTIPLIER` - Default mask multiplier in SBM. Defaults is 1% of mean i.e., voxels greater than or equal to 1% of mean will be used.

2.7 Spatial Templates

The following example spatial templates are provided in `icatb\icatb_templates`:

- `ref_right_visuomotor.nii` - Mask containing the right visual regions.
- `ref_left_visuomotor.nii` - Mask containing the left visual regions.
- `ref_default_mode.nii` - Mask containing the default mode network regions. Please see [5] for more information.
- `*DMN_ICA_REST*.nii` - This mask was created from a data-set of 42 subjects. During a fMRI scan, subjects were asked to relax and passively stare at a fixation cross. A pooled group ICA was performed and the default mode component network was selected to create this mask. We provide templates like `rDMN_ICA_REST_3x3x3.nii` and `rDMN_ICA_REST_3x3x4.nii` in the GIFT. Please see [4] for more information.
- `*DMN_MASK.WFU*.nii` - This mask was constructed using the Wake Forest Pick atlas toolbox. A binary mask was created by selecting the anatomical regions that have been most commonly reported to comprise the default mode network. The labels from the Wake Forest Atlas that constituted this mask included posterior cingulate (BAs 23/31), inferior and superior parietal lobes (BAs 7/39/40), superior frontal gyrus (BAs 8/9/10), and anterior cingulate cortex (BAs 11/32). In addition, a larger weight was given to the anterior and posterior cingulate cortex, which are believed to be the central nodes of the default mode network. A higher correlation was observed when giving a higher weight to these two central nodes. We provide templates like `rDMN_MASK.WFU.3x3x3.nii` and `rDMN_MASK.WFU.3x3x4.nii` in the GIFT. Please see [4] for more information.

2.8 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
    - Batch - Option is now provided to select the input files for batch analysis. More information on input file is provided in Section 2.13.1.
    - Remove Component(s) - Removes a component or components from the data after you have selected the parameter file. Please see section 2.12.1 for more information.
    - Mancovan - Mancovan toolbox works on the ICA output. Multi variate stats is used to determine the significant covariates which will be used later in the univariate tests. Please see 2.12.3 for more information.
    - Component viewer - Orthogonal views of the selected component and its power spectra are displayed.
    - ICASSO - ICASSO plugin is used to assess the stability of components. Please see Section 2.12.2 for more information.

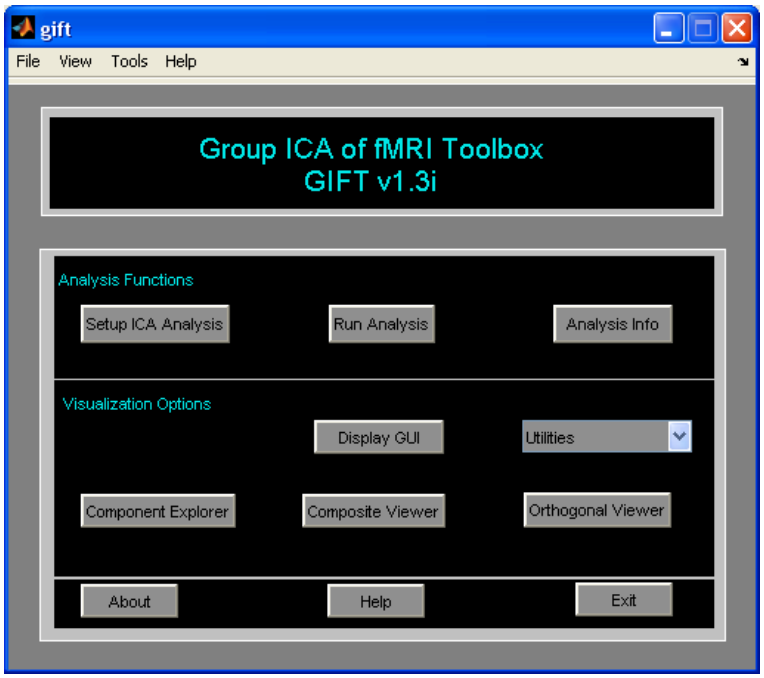


Figure 2.1: GUI for running group ICA.

* `Ascii_to_spm.mat` - Creates `SPM.mat` from ascii file (containing regressor time courses) that is needed during temporal sorting. Please see Section 2.12.4 for more information.
* Event Average - Event average is calculated for the ICA time courses. Please see Section 2.12.5 for more information.
* Calculate Stats - Mean, standard deviation, t-maps are calculated for components over sessions, subjects or subjects and sessions.
* Spectral Group Compare - This utility is used to compare the power spectra between groups. Please see Section 2.12.7 for more information.
* Stats On Beta Weights - This utility is provided for doing statistics on the time courses (beta weights). Please see Section 2.12.8 for running this utility.
* SPM Stats - This utility is used to do one sample t-test or two sample t-test on the component images. Please see Section 2.13.1 for more information.
* Spatial-temporal regression - Given a set of GLM or ICA spatial maps and the original data of the subjects, you could use this utility to back reconstruct subject components (Section 2.12.11).
* Write Talairach Table - Talairach daemon client is used to generate the talairach tables for the selected image. Please see Section 2.12.10 for more information.
* Single Trial Amplitudes - We provide the option for calculating single trial amplitudes (Section 2.12.12) in GIFT.
* Z-shift - Please see Section 2.12.13 for more information.
* Percent Variance - This utility can be used after running the group ICA analysis. Please see Section 2.12.14 for more information.

2.9 Analysis Functions

2.9.1 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

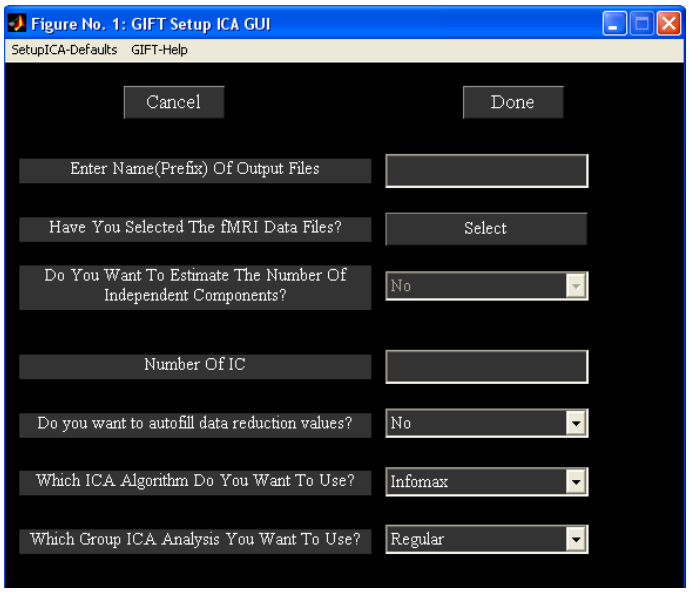


Figure 2.2: Initial parameter selection window.

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 2.6.3. 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 ([32]) 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.

- '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. If there are more than one data reduction step, initial PC numbers are set to 1.5 times the number of final components.

- 'Which ICA Algorithm Do You Want To Use?' There are 14 ICA algorithms available like Infomax, FastICA, ERICA, SIMBEC, EVD, JADE OPAC, AMUSE, SDD ICA, Semi-blind Infomax, Constrained ICA (Spatial), Radical ICA, Combi, ICA-EBM and FBSS.

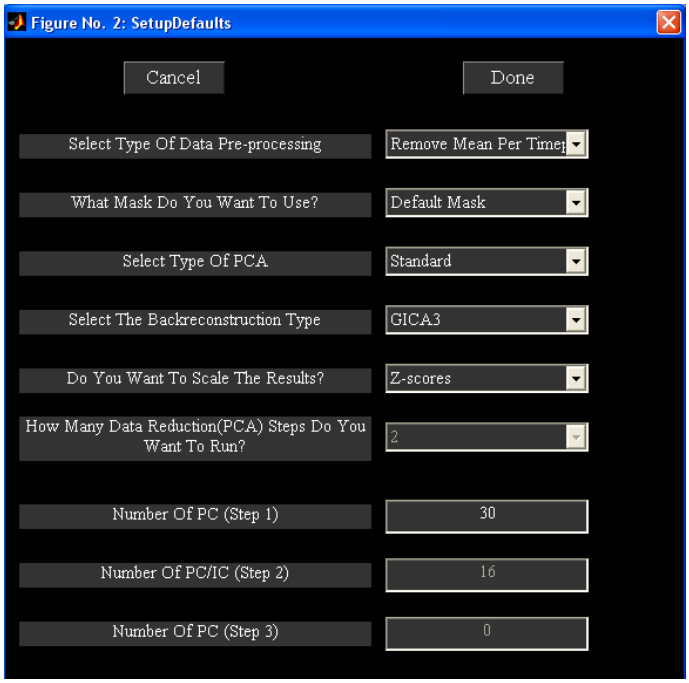


Figure 2.3: Hidden user interface controls.

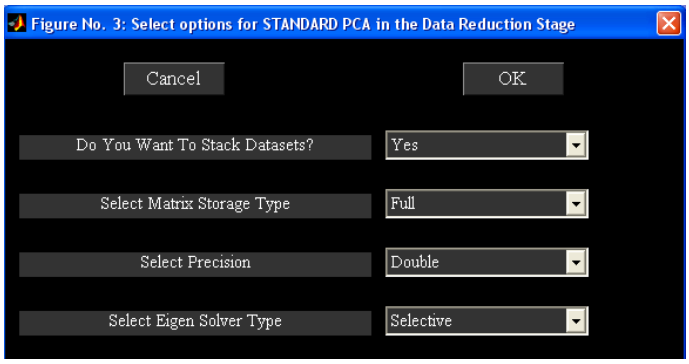


Figure 2.4: PCA Options.

- 'Which Group ICA Analysis You Want To Use?' Options are 'Regular' and 'ICASSO'. When you select 'ICASSO', ICA is run several times and the best estimate for each component is used (See Section 2.12.2).

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

- 'Select Type Of Data Pre-processing' - Data is pre-processed prior to the first data reduction. Options are discussed below.
  - 'Remove Mean Per Timepoint' - At each time point, image mean is removed.
  - 'Remove Mean Per Voxel' - Time-series mean is removed at each voxel.
  - 'Intensity Normalization' - At each voxel, time-series is scaled to have a mean of 100. Since the data is already scaled to percent signal change, there is no need to scale the components.
  - 'Variance Normalization' - At each voxel, time-series is linearly detrended and converted to z-scores.
- '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.
- 'Select Type Of PCA' - There are three options like 'Standard', 'Expectation Maximization' and 'SVD'. PCA options window (Figure 2.4) will change depending on the type of PCA selected.
  - Standard
    - 'Do You Want To Stack Datasets?' - Options are 'Yes' and 'No'.
      - 'Yes' - Data sets are stacked to compute covariance matrix. This option assumes that there is enough RAM available to stack the data sets and for computing covariance matrix. Please note that full storage of covariance matrix is required when you select this option.
      - 'No' - A pair of data sets are loaded at a time to compute covariance matrix. This option uses less memory but it requires N_{C_2} loops to compute the covariance matrix where N is the number of data sets.
    - 'Select Matrix Storage Type' - Options are 'Full' and 'Packed'. You have the option to store only lower triangular portion of the symmetric matrix with the packed storage scheme.
    - 'Select Precision' - Options are 'Double' and 'Single'. Single precision uses 50% less memory required when compared to double precision. Single precision is accurate up to 7 digits after decimal point.
    - 'Select Eigen Solver Type' - Options are 'Selective' and 'All'. These options will be used only for the packed storage scheme.
      - 'Selective' - Only a few desired eigen values are computed. This option will compute eigen values faster when compared to 'All' option. However, if there are convergence issues use option 'All' to compute eigen values.
      - 'All' - All eigen values are computed. We recommend to use this option for computing eigen values only when the selective eigen solver doesn't converge.
  - Expectation Maximization (EM PCA) has fewer memory constraints and is advantageous over standard PCA when only few eigen values need to be computed from a large data-set ([27]). PCA options of this approach are discussed below:
    - 'Do You Want To Stack Datasets?' - Options are 'Yes' and 'No'.

        - 'Yes' - This option assumes that there is enough RAM available to stack the data sets.
        - 'No' - A data-set is loaded at a time to compute transformation matrix at each iteration. This option may take days to solve the problem if there are very large data-sets.
    - 'Select Precision' - Options are 'Double' and 'Single'.
    - 'Select Stopping Tolerance' - Norm of residual error is used. Residual error is computed by subtracting the transformation matrix at the current iteration from the previous iteration.
    - 'Enter Max No. Of Iterations' - Maximum number of iterations to use.
  - SVD - Singular value decomposition (SVD) is preferable when the data is ill-conditioned. Memory requirements of SVD are similar to covariance based PCA.
    - 'Select Precision' - Options are 'Double' and 'Single'.
    - 'Select Solver' - Options are 'Selective' and 'All'.

  Note:

    - Before setting up analysis, please see `icatb_mem_ica.m` script to get a close estimate of the RAM required for all the analysis types. In general for better performance, stack data-sets using single precision. However, if memory is an issue don't stack data-sets and use slower ways to compute PCA (EM PCA or packed storage scheme of standard PCA).
    - By default, GIFT will save MAT files in the uncompressed format (`-v6`). Always use uncompressed format if you want a better performance during the analysis phase.

- 'Select The Backreconstruction Type' - Options are 'Regular' (GICA2), 'Spatial-temporal regression', 'GICA3' and 'GICA'. GICA2 and GICA3 are not not shown in the GUI but can be called in the batch script.
  - 'Regular' - Regular or GICA2 has one desirable property that the sum of the reconstructed subject spatial maps equals the aggregate spatial map. However, product of time courses and spatial maps doesn't estimate the PCA reduced data.
  - 'Spatial-temporal Regression' - Back reconstruction is done using a two step multiple regression (See [22]). In the first step, aggregate component spatial maps are used as basis functions and projected on to the subject's data resulting in subject component time courses. In the second step, subject component time courses are used as basis functions and projected on to the subject's data resulting in component spatial maps for that subject.
  - 'GICA3' - GICA3 has two desirable properties that the sum of the subject spatial maps is the aggregate spatial map and the product of the time courses and spatial maps estimate the data to the accuracy of the PCA's. Please see [14] for more information.
  - 'GICA' - GICA ([30]) is a more robust tool to back reconstruct components when compared to GICA2 and GICA3 for low model order.

  Note:

  - GICA, GICA2 and GICA3 back reconstruction methods use the PCA whitening and dewhitening matrices to reconstruct subject spatial maps and timecourses.
  - GICA and GICA2 timecourses are similar to the timecourses obtained using Spatial-temporal Regression.
  - Spatial maps obtained using GICA2 are exactly equal to the GICA3 method.
  - All the back reconstruction methods give the same spatial maps and timecourses for one single subject single session analysis.
  - GICA, GICA2 and Spatial-temporal Regression component timecourses are equivalent when 100% variance is retained in the first step PCA.

- 'Do You Want To Scale The Results?' The options available are 'No Scaling', 'Scale To Original Data(%)', 'Z-Scores', 'Scaling in Timecourses' and 'Scaling in Maps and Timecourses'.
  - 'Scale To Original Data(%)' - 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.

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
$N \geq 4$	User specified number (2 or 3)

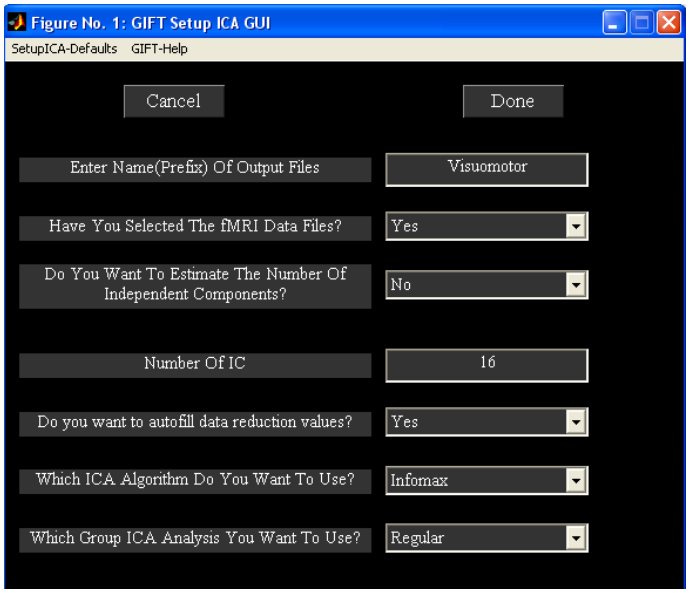


Figure 2.5: Completed parameter selection window.

  - 'Scaling in Timecourses' - Spatial maps are normalized using the average of top 1% voxels and the resulting value is multiplied to the timecourses.
  - 'Scaling in Maps and Timecourses' - Spatial maps are scaled using the standard deviation of timecourses and timecourses are scaled using the maximum spatial intensity value.

Note: By default, subject component images are centered based on the peak of the distribution. Please see variable `CENTER_IMAGES` in `icatb_defaults.m`.

- '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.

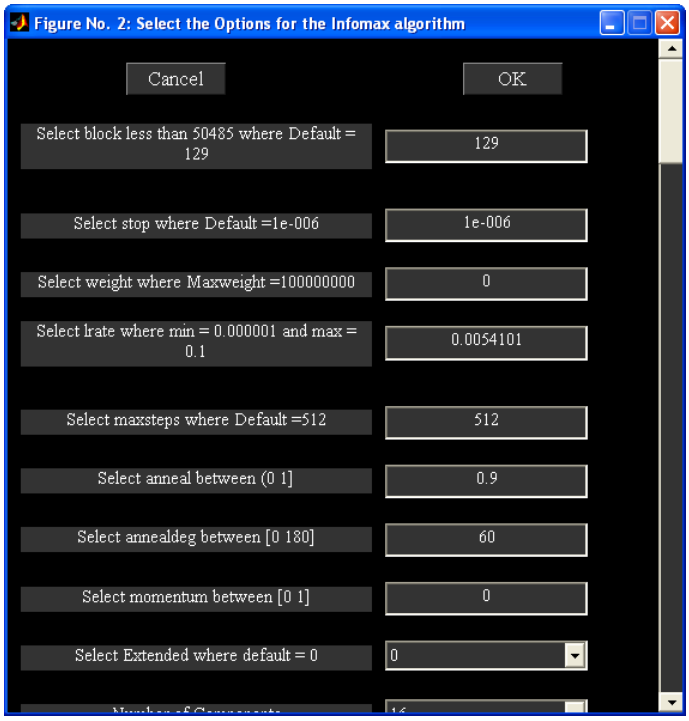


Figure 2.6: ICA Options Window

Completing Entering Parameters

Figure 2.5 shows the completed parameters window. Press *Done* button after selecting all the answers for the parameters. This will open a figure window (Figure 2.6) 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, Constrained ICA (Spatial) algorithms and FBSS. 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.

2.9.2 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 2.7) 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 by selecting steps from 'Parameter Initialization' to 'Group Stats' in order.
  - 'Resume'

    - Resume option used to handle interrupted analysis. Resume option also detects changes in the user input and runs the appropriate group ICA steps. For example, if you changed back-reconstruction approach resume option runs the steps from back-reconstruction to group stats.
  - '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'
    - 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'
    - · For 'GICA', 'GICA2' and 'GICA3', the aggregate components and the results from data reduction are used to compute the individual subject components.
      · 'Spatial-temporal Regression' - The aggregate components and the original data are used to compute the individual subject components.
    - The individual subject components are saved in Analyze (or Nifti) format.
  - 'Calibrating Components'
    - By default, components are in arbitrary units. Components are scaled to percent signal change, z-scores, scaling in timecourses or scaling in maps and timecourses.
  - '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 2.9) will open automatically for visualizing components.

- Group PCA Performance Settings - There are three options like 'Maximize Performance', 'Less Memory Usage' and 'User Specified Settings'. Best match for each option is selected based on the variable `MAX_AVAILABLE_RAM`. PCA types selected will be between covariance based and expectation maximization approaches.
  - 'Maximize Performance' - Reduced data-sets from the first data reduction step are stacked by default.
  - 'Less Memory Usage' - Slower ways of computing PCA are used.
  - 'User Specified Settings' - User specified PCA options are selected.

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.
- Run analysis steps can also be accessed from the command line.

```
load(param_file); % Load parameter file (*ica*param*mat)
sesInfo = icatb_runAnalysis(sesInfo, 1); % Run All Steps
sesInfo = icatb_runAnalysis(sesInfo, 2); % Parameter Initialization
sesInfo = icatb_runAnalysis(sesInfo, 3); % Data Reduction
sesInfo = icatb_runAnalysis(sesInfo, 4); % ICA
sesInfo = icatb_runAnalysis(sesInfo, 5); % Back reconstruction
sesInfo = icatb_runAnalysis(sesInfo, 6); % Scaling components
sesInfo = icatb_runAnalysis(sesInfo, 7); % Group Stats
sesInfo = icatb_runAnalysis(sesInfo, 8); % Resume interrupted analysis
```

- Option is provided in the GIFT to run a particular data reduction step. This is useful when a particular data reduction step was already done and you would like to go to the next step without re-running the earlier step.

```
load(param_file); % Load parameter file (*ica*param*mat)
sesInfo.reductionStepsToRun = 2; %Run 2nd reduction only
sesInfo = icatb_runAnalysis(sesInfo, 3); % Call Data Reduction
```

- You could also switch between PCA types using command line. For example, the first data reduction could be done using Standard PCA and the memory intensive second data reduction could be done using Expectation Maximization.

```
load(param_file); % Load parameter file (*ica*param*mat)
%% Run 1st data reduction using Standard PCA
sesInfo.pcaType = 'standard'; % Standard PCA
sesInfo.reductionStepsToRun = 1; % First reduction
sesInfo = icatb_runAnalysis(sesInfo, 3); % Call data reduction

%% Run 2nd data reduction using EM PCA
sesInfo.pcaType = 'expectation maximization'; % EM PCA
sesInfo.pca_opts.precision = 'single'; % Use single precision
sesInfo.reductionStepsToRun = 2; % Second reduction
sesInfo = icatb_runAnalysis(sesInfo, 3); % Call 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).

2.9.3 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 2.8) will pop up showing the information contained within this window.

2.10 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 2.9 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 2.10). 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.

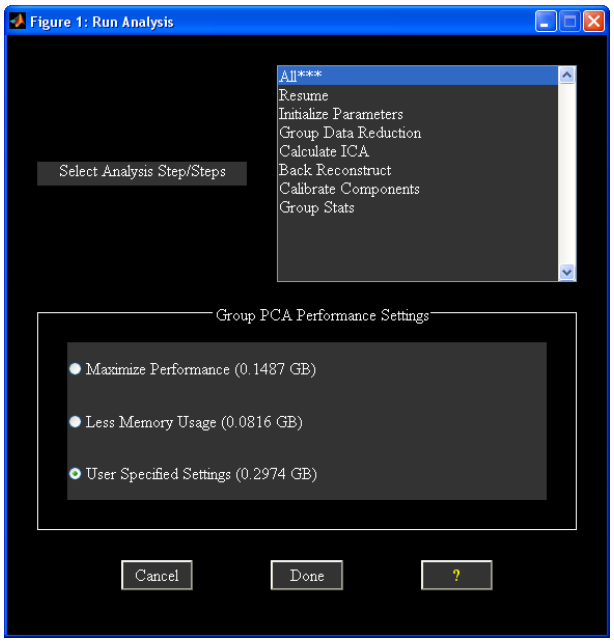


Figure 2.7: Figure window used to run the analysis.

Figure 2.8: Analysis Info shows the information of the completed analysis.

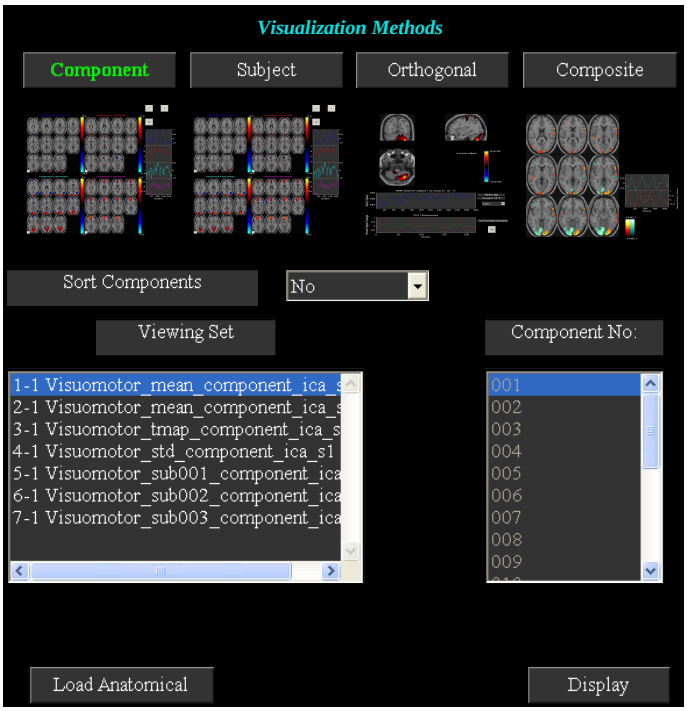


Figure 2.9: *Display GUI* contains four visualization methods like Component Explorer, Subject Explorer, Orthogonal Viewer and Composite Viewer. Component Explorer is plotted by default after you had selected the parameter file.

Figure 2.10: User interface controls plotted in "Display Defaults" menu

- *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 2.10) 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 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`.

2.10.1 Component Explorer

- Component Explorer is used to display all components of a particular viewing set. Therefore, 'Component Number' control (Figure 2.9) will be disabled for this visualization method.
- Figure 2.9 shows the selected parameters for the Component Explorer. Click on *Display* button and wait for the figures containing spatial maps to pop up. Figure 2.11 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 2.9).
- The time course for each component is displayed on the top of the figure (Figure 2.11). 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.

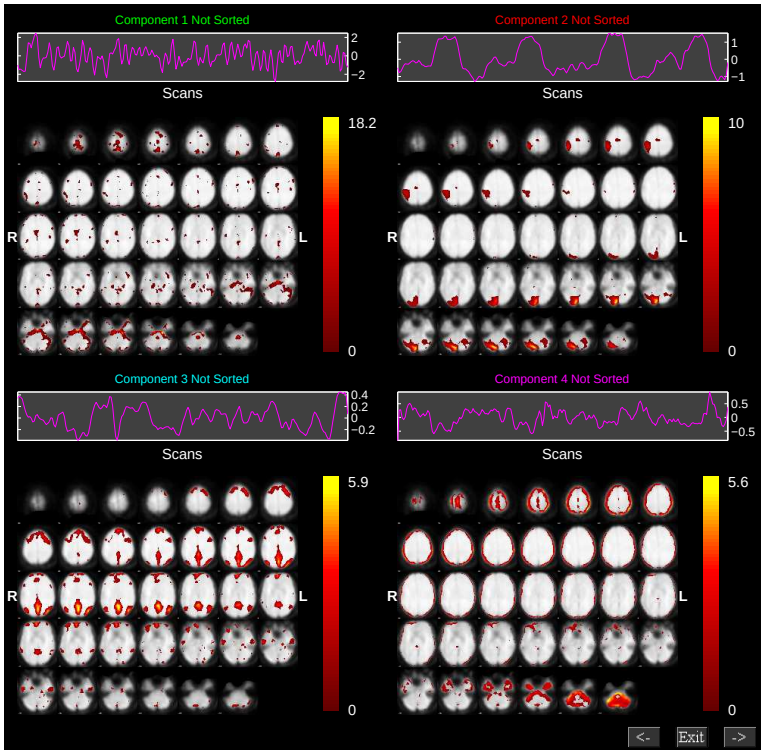


Figure 2.11: Figure shows the components not sorted in groupings of four.

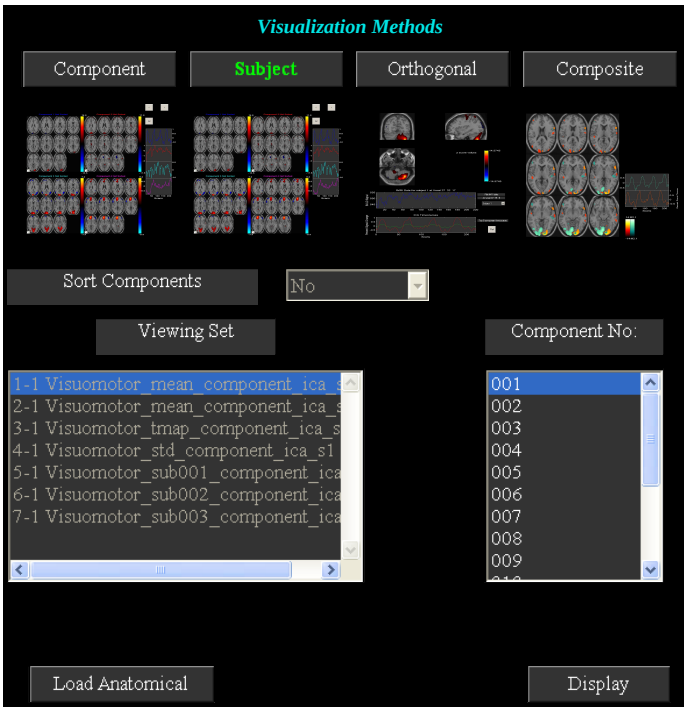


Figure 2.12: Selected options of the Subject Explorer method.

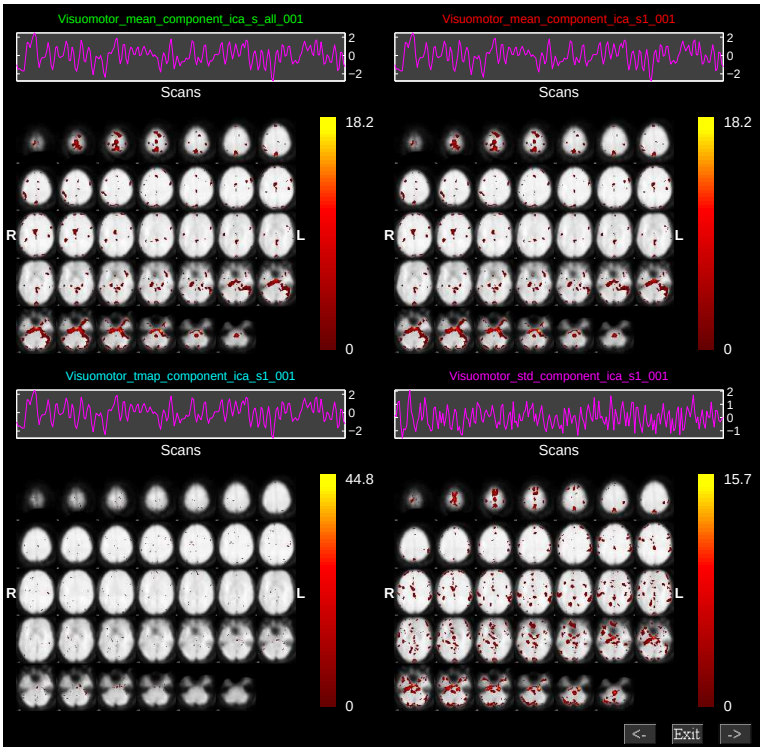


Figure 2.13: Component '001' of all the viewing sets. Here, only the first four components are displayed.

2.10.2 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 2.12 shows the selected parameters.
- Click *Display* button. Figure 2.13 shows the component '001' of all the entries in the 'viewing set'.

2.10.3 Orthogonal Explorer

- Orthogonal viewer is used to look at a component and compare it to the functional data.
- Figure 2.14 shows the selected parameters. Click on Display button. Figure 2.15 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 2.15 to enter the voxel (real world coordinates) instead of navigating around the brain.

2.10.4 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 2.16 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 2.17 will open in a new window.

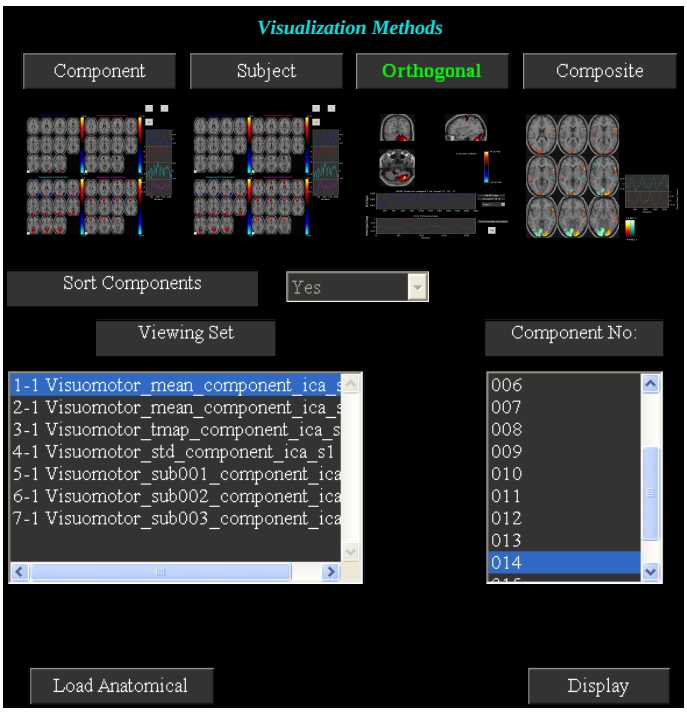


Figure 2.14: Figure shows the options of the Orthogonal Viewer method.

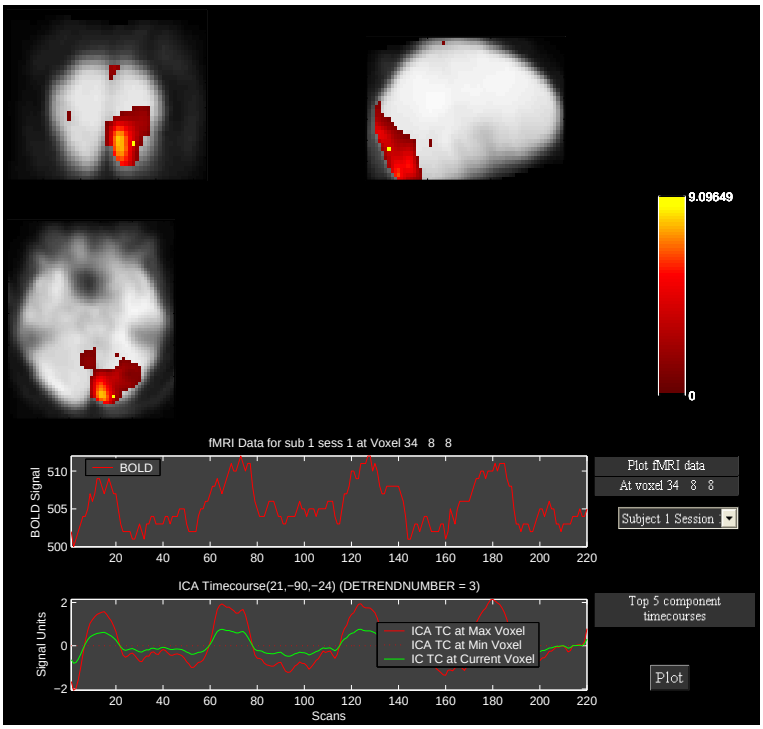


Figure 2.15: One of the task-related components plotted using Orthogonal Viewer.

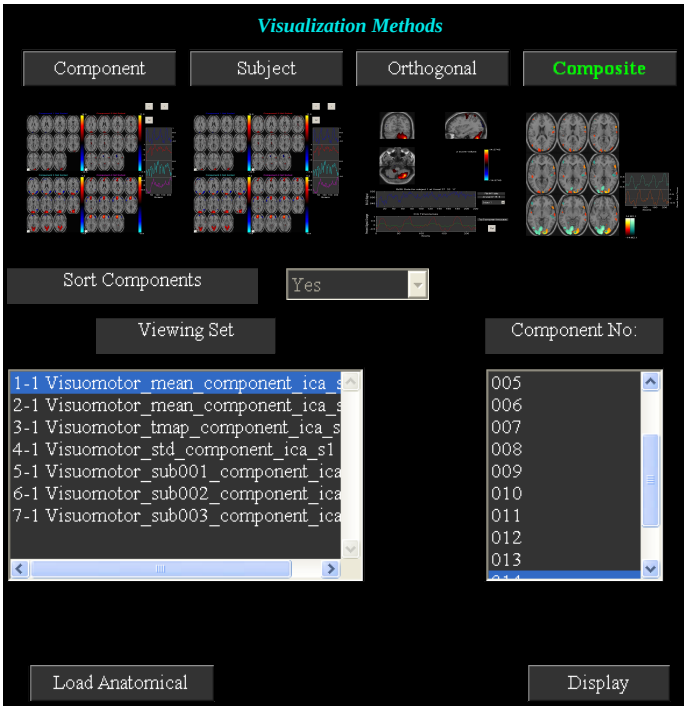


Figure 2.16: Selected options of the Composite Viewer method.

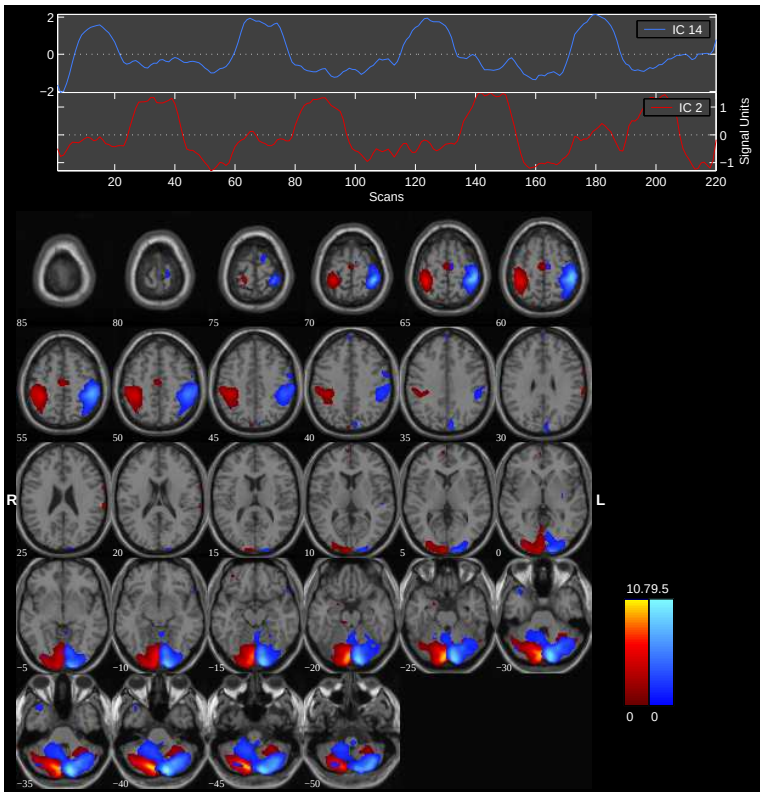


Figure 2.17: Task-related components overlaid on one another. At most five different components with different color bars can be overlaid.

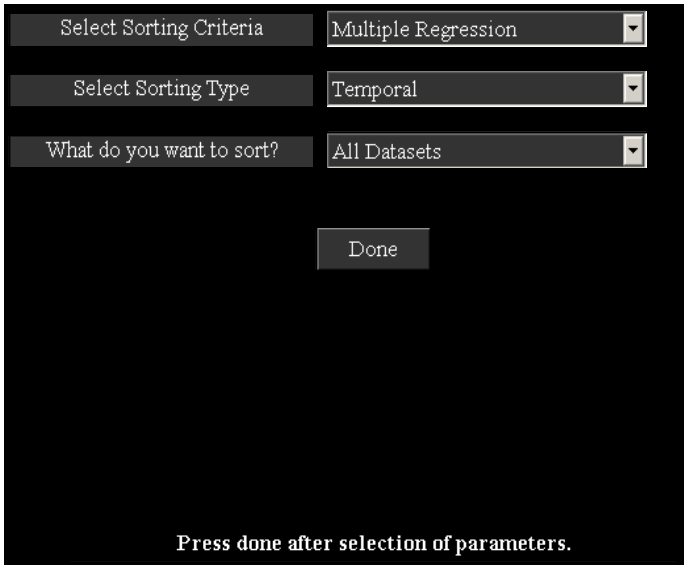


Figure 2.18: Selected parameters for sorting the components temporally.

2.11 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 2.18) 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:

2.11.1 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 2.19 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 2.20) 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:

- 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 2.21) will be shown. Click on sub menu "Event Average" and select "right*bf(1)" reference function to plot the event averages (Figure 2.22) of the ICA time courses. Explanation of the event average is given in Section 2.12.5.
- 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

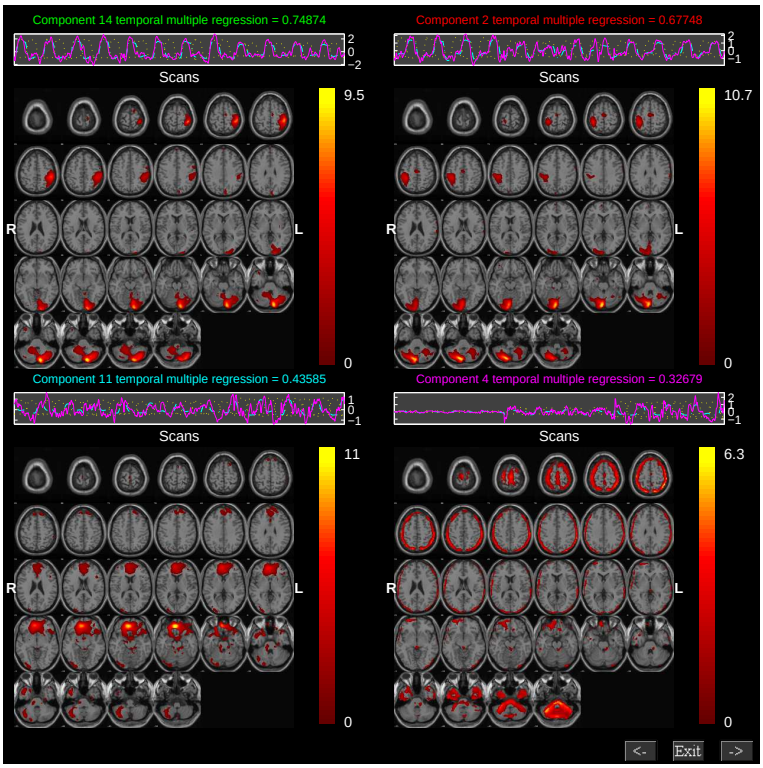


Figure 2.19: Components are sorted based on Multiple Regression criteria in groupings of four.

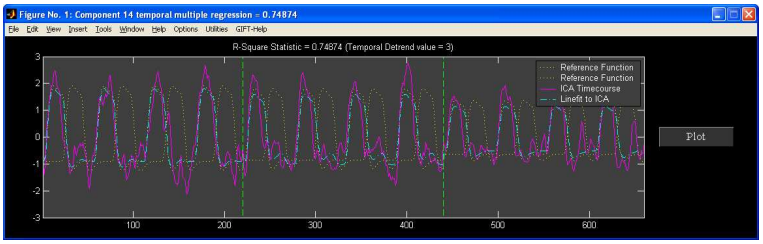


Figure 2.20: Expanded view of the "right*bf(1)" time course.

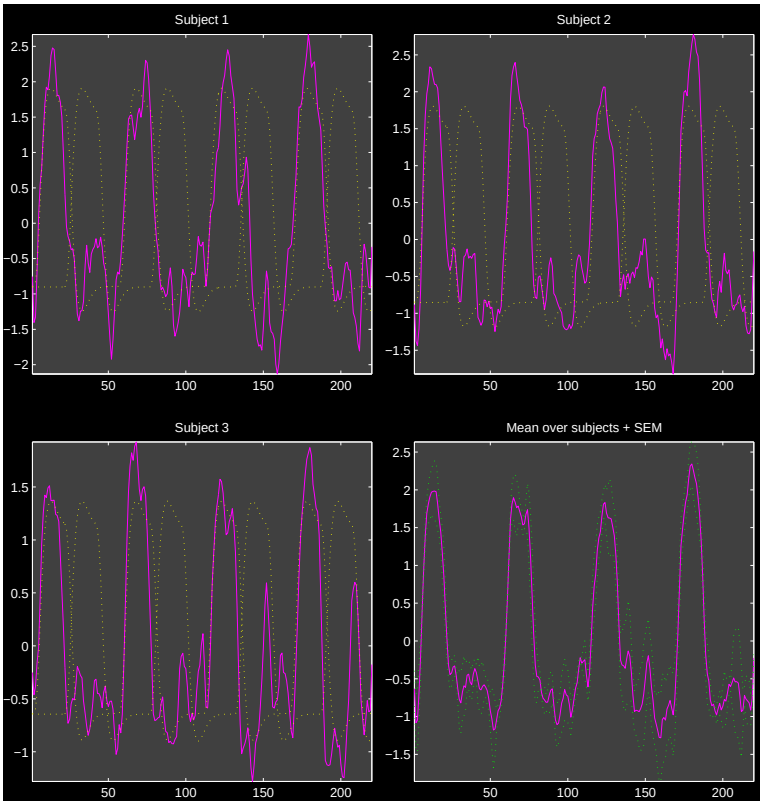


Figure 2.21: Figure shows the split of the concatenated time courses of all the data-sets. Mean is calculated over sessions and subjects as well as over all data-sets.

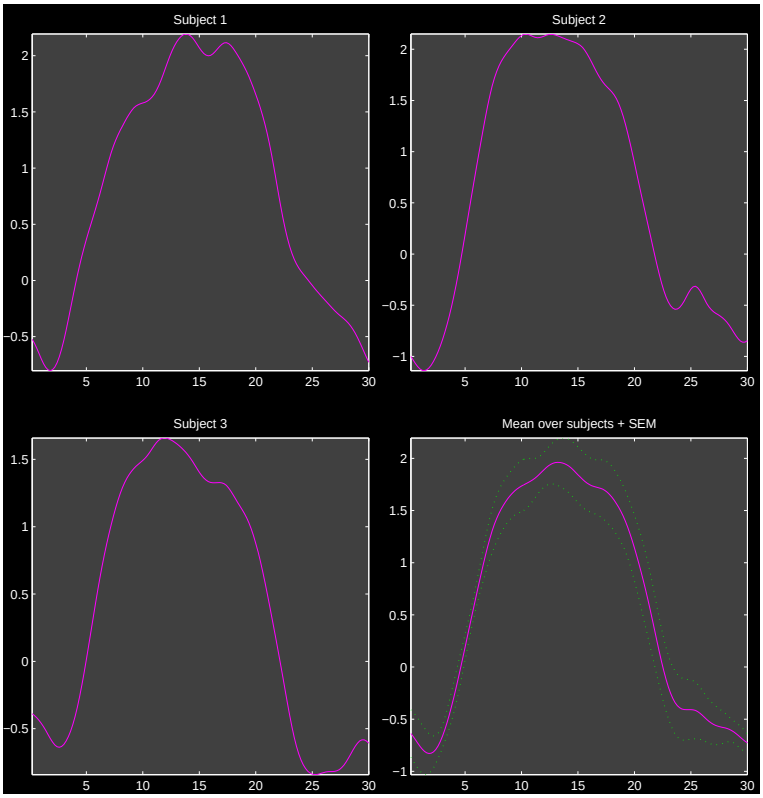


Figure 2.22: Figure shows the event related averages of the selected component for different subjects.

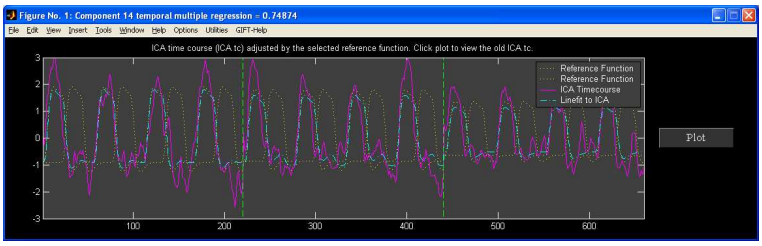


Figure 2.23: Expanded view of ICA time course after removing the variance of other than selected regressor ("left*bf(1)").

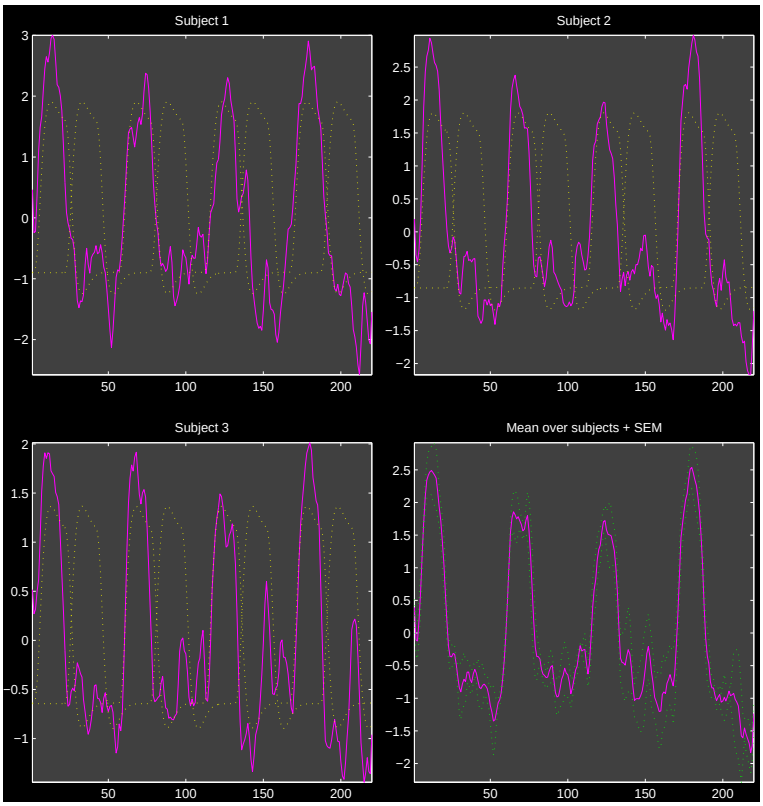


Figure 2.24: Figure shows the split of the concatenated time courses of all the data-sets after adjusting.

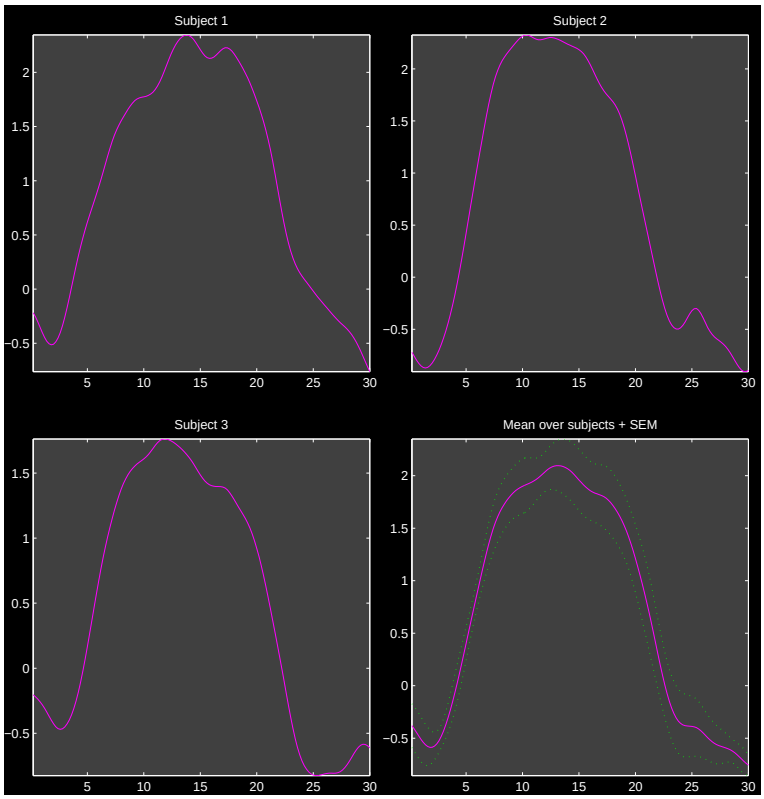


Figure 2.25: Figure shows the event averages of the adjusted ICA time courses.

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 2.23). When you click on the sub menu "Split-time courses" in "Utilities" menu, a new figure window (Figure 2.24) 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 2.25) of the new ICA time courses.

Note: Event average can also be done without sorting components (Section 2.12.5). 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 ([10], pp 12-14).

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 provide an option in GIFT to do one sample t-test and two sample t-test on individual subject component maps using SPM5 (Section 2.13.1).

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

- We provide utility to do statistics on the regression parameters. Please see section 2.12.8 for more information.

2.11.2 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 2.26 shows the components of subject 1 session 1 sorted based on the MLR sorting criteria in groupings of four. The templates used are `RighTemplate.nii` and `LeftTemplate.nii`. Here, you can see that the first two components are task related.
- Figure 2.27 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`.

2.12 Utilities

2.12.1 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) ([11]), Principal Component Analysis or Maximum Signal Fraction ([17]) 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).

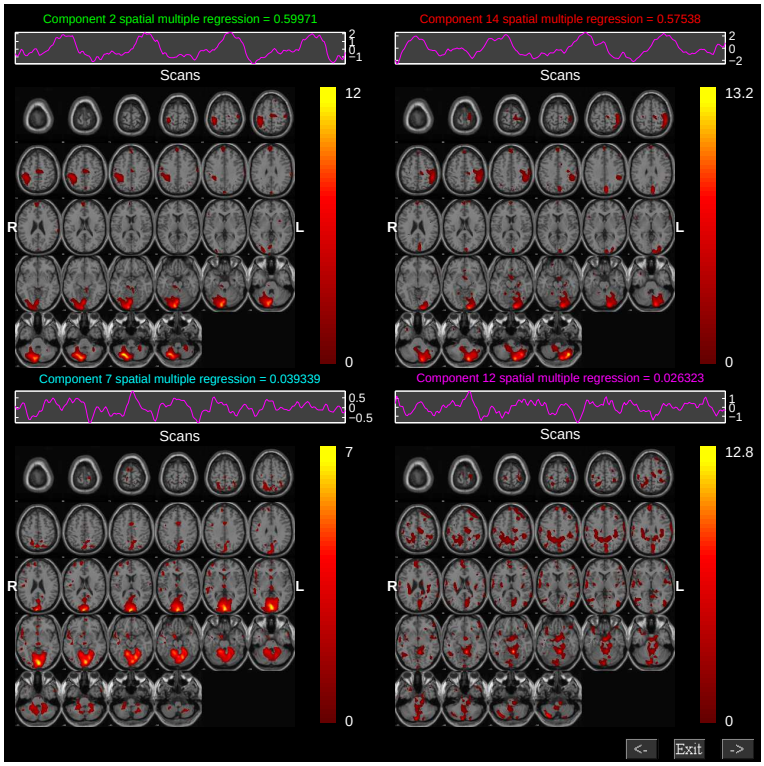


Figure 2.26: Figure shows the components spatially sorted using Multiple Regression sorting criteria.

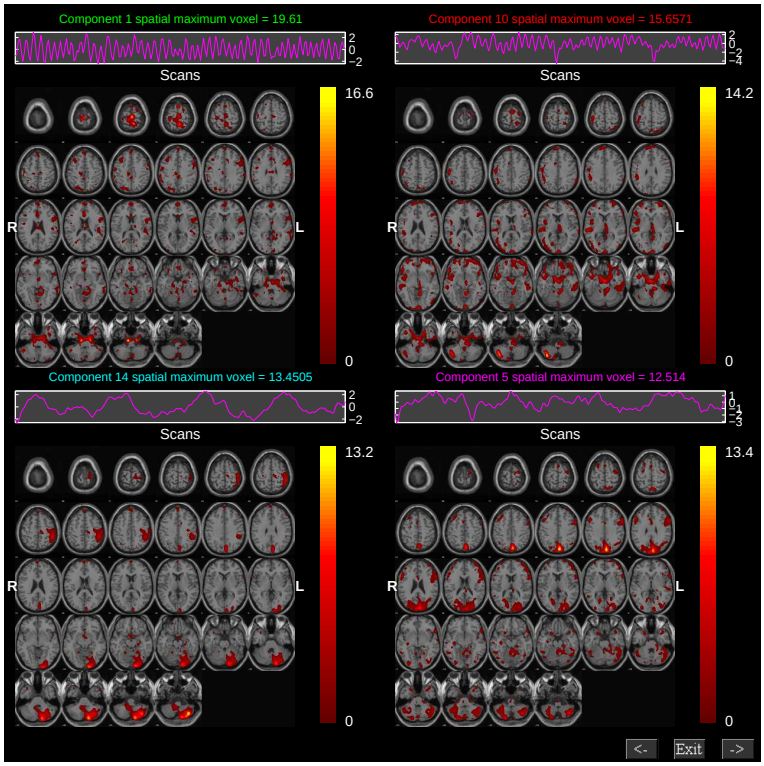


Figure 2.27: Figure shows the components spatially sorted using Maximum Voxel sorting criteria.

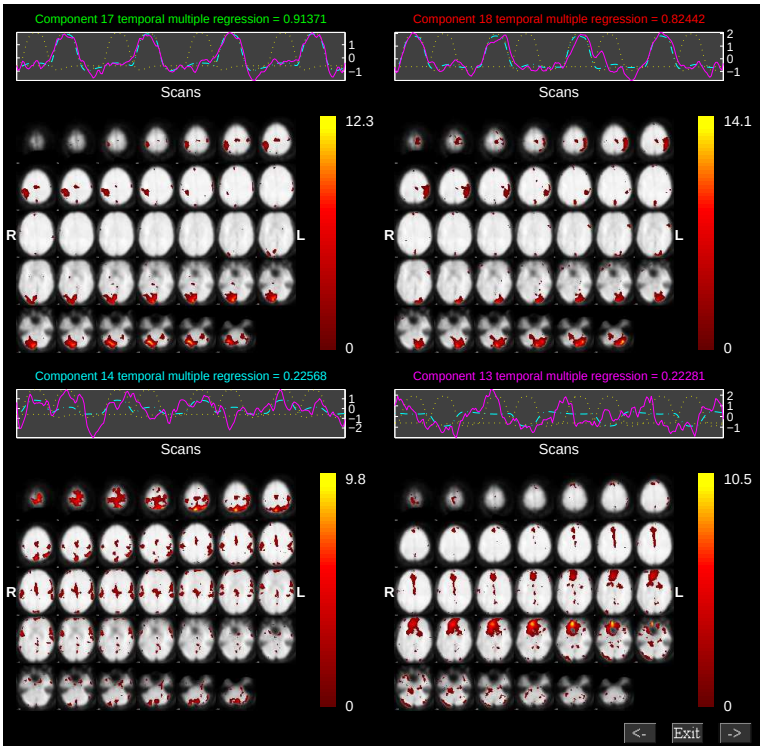


Figure 2.28: First four components that were sorted temporally using Multiple Regression and the regressors used are "right*bf(1)" and "left*bf(1)".

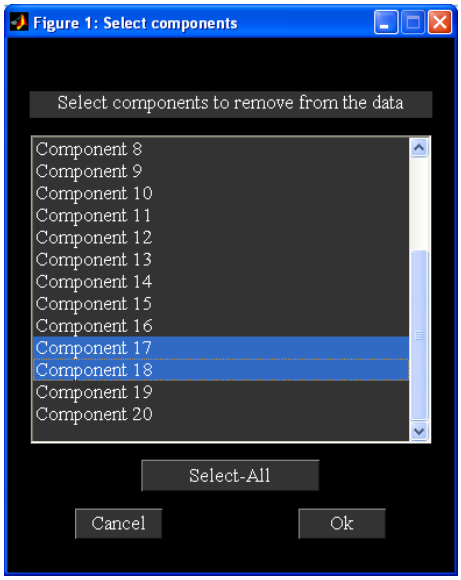


Figure 2.29: Dialog box showing component/components to be removed from the data.

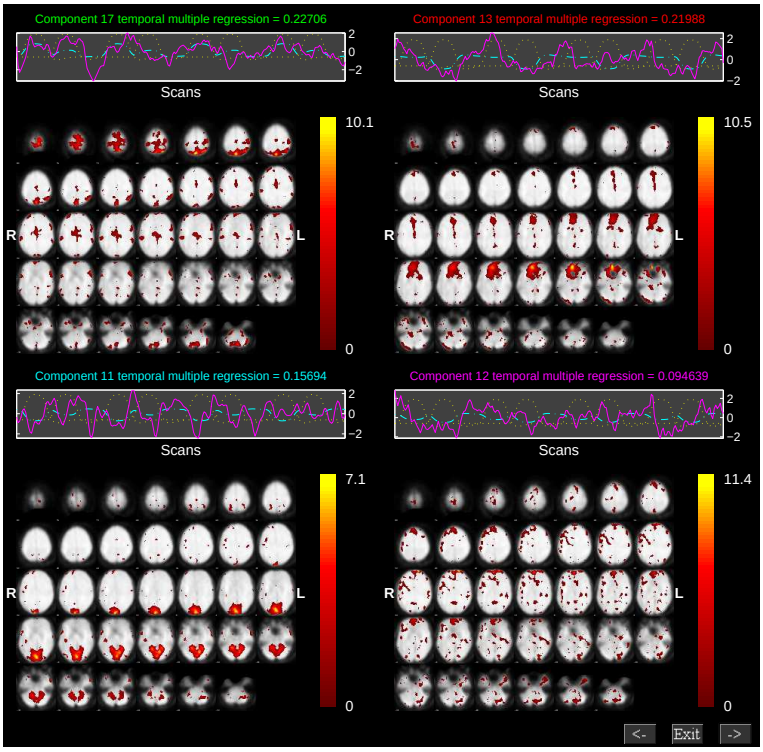


Figure 2.30: Results obtained after removing components from the data.

- 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 2.29) will open to select the components to be removed from the fMRI Data. We selected the first two components in the figure 2.28 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 2.30 shows that both left and right visual components are removed 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 2.31) and after (Figure 2.32) removing the IC from the fMRI data. Both the results are obtained by applying a t-threshold of 3.3.

2.12.2 ICASSO

ICASSO toolbox ([15]) 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 2.33 will open when you click on "Utilities" drop down box (Figure 2.1). The parameters in the figure 2.33 are as follows:

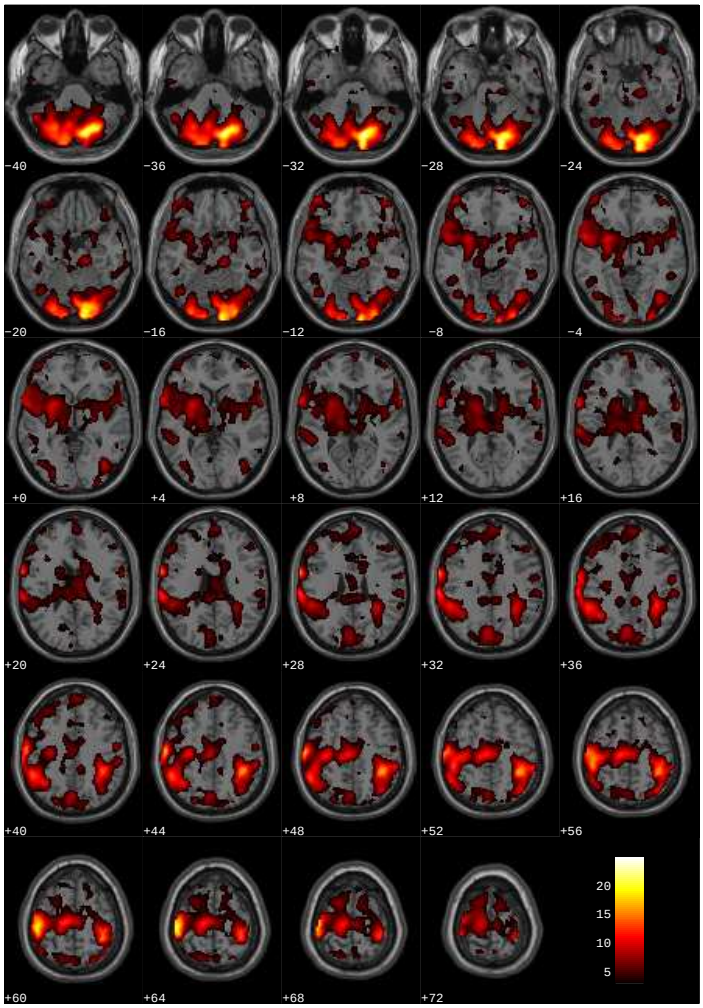


Figure 2.31: Left-right visual before removing the IC from the data.

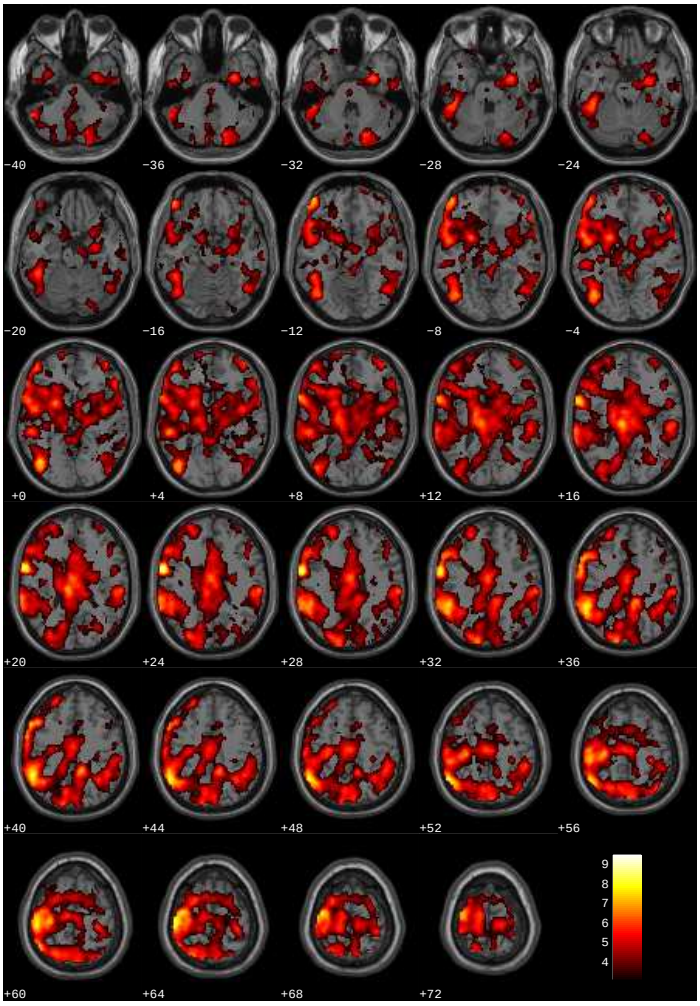


Figure 2.32: Left-right visual after removing the IC from the data.

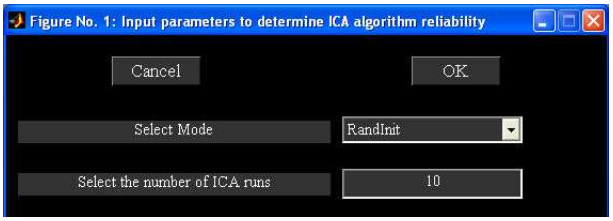


Figure 2.33: ICASSO GUI.

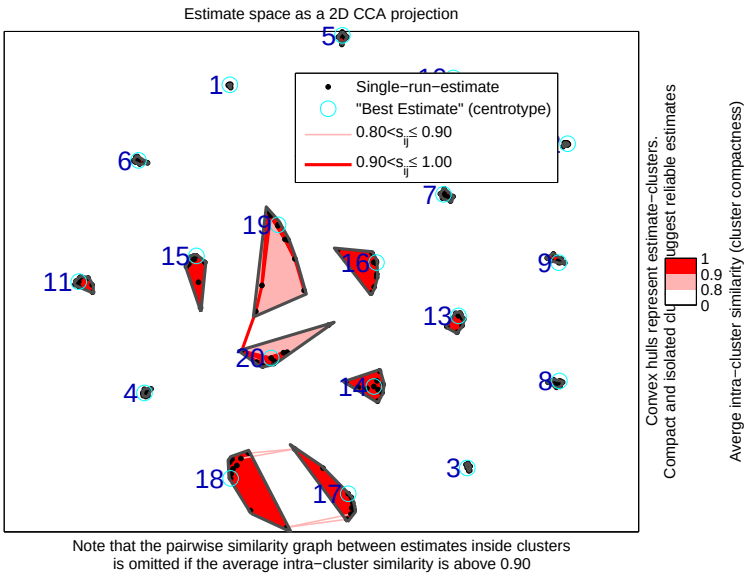


Figure 2.34: ICASSO Results.

- "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.
- "Select number of ICA runs" - Number of times ICA algorithm will be run.

We ran ICA 10 times on FastICA algorithm and extracted 20 components from the data. Figure 2.34 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.

We use centrotype of the cluster for each component instead of the average of the individual ICA runs as it is the best estimate. After the ICASSO step is completed, subsequent group ICA analysis steps like Back Reconstruction, Scaling Components and Group Stats are run.

2.12.3 Mancovan

Mancovan toolbox is based on the paper [12]. This toolbox works on MATLAB versions greater than R2006a and is dependent on MATLAB toolboxes like stats, image processing, signal processing and optimization. Features used are subject component spatial maps, timecourses spectra and FNC correlations. Multivariate tests are done on the features to determine the significant covariates which are later used in the univariate tests on each feature. You could also invoke toolbox using `mancovan_toolbox` at the command prompt. Mancovan toolbox (Figure 2.35) is divided into four parts like create design matrix, setup features, run mancova and display. Each step is explained below:

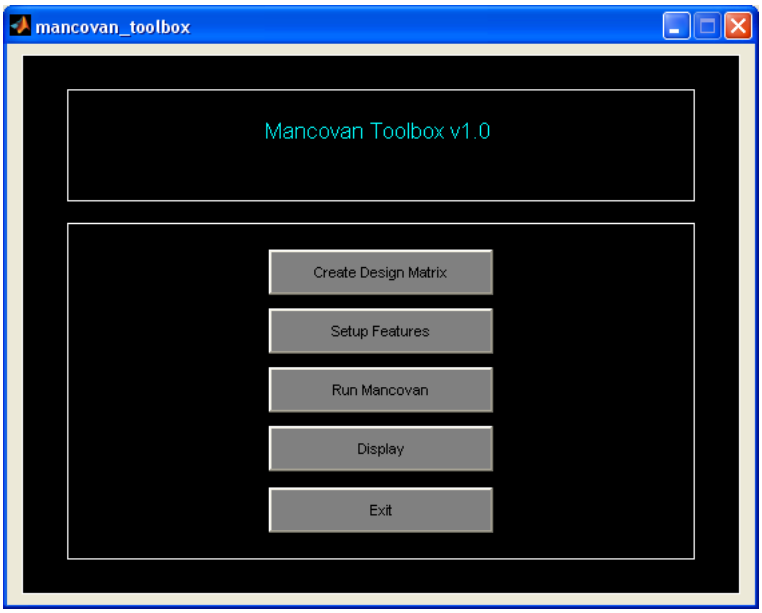


Figure 2.35: Mancovan Toolbox

Create Design Matrix

When you click on *Create Design Matrix* button, a figure window (Figure 2.36) will open to select the ICA parameter file. All the output files will have prefix `prefix_mancovan` where prefix is from ICA parameter file. Add covariates of interest using + button next to the listbox (Figure 2.37). You could type covariate values by hand or use right click on the edit box (Figure 2.38) and this will open a figure window to load the ascii file for continuous covariates. For categorical covariates, specify labels to distinguish the levels. Each covariate vector must have length equal to no. of subjects in the original ICA analysis. Option is also provided to apply transformation function to the continuous covariates. Use *create* button to create the design matrix. Design information is stored in mancovan parameter file `*mancovan.mat`.

Setup Features

When you click on *Setup Features* button (Figure 2.35), a figure window will open to select the mancovan parameter file. The following are the steps involved:

- Features - Select feature or features of interest (Figure 2.39).
- Add components - Enter component networks using + button. The network name will be used in the advanced display in future release.
- p-value significance - Enter p-value significance threshold which will be used in multivariate and univariate tests.
- Enter TR in seconds - Enter TR of the original data in seconds. This information will be used in computing spectra and filtering of timecourses.
- Enter no. of components - Enter no. of components for each feature in a vector. The entered components should not exceed the minimum feature dimensions. Feature data will be reduced in the voxel or spectra dimension which will be later used in the multivariate tests.
- Mancovan defaults - Mancovan defaults (Figure 2.40) will open when you use "Mancovan defaults" menu. The defaults for each feature are as follows:
  - Spatial Maps - You could use the user specified mask or default mask. Default mask includes the voxels based on the distribution of voxelwise t-statistics. Only the voxels with strong and consistent activation across subjects are included.

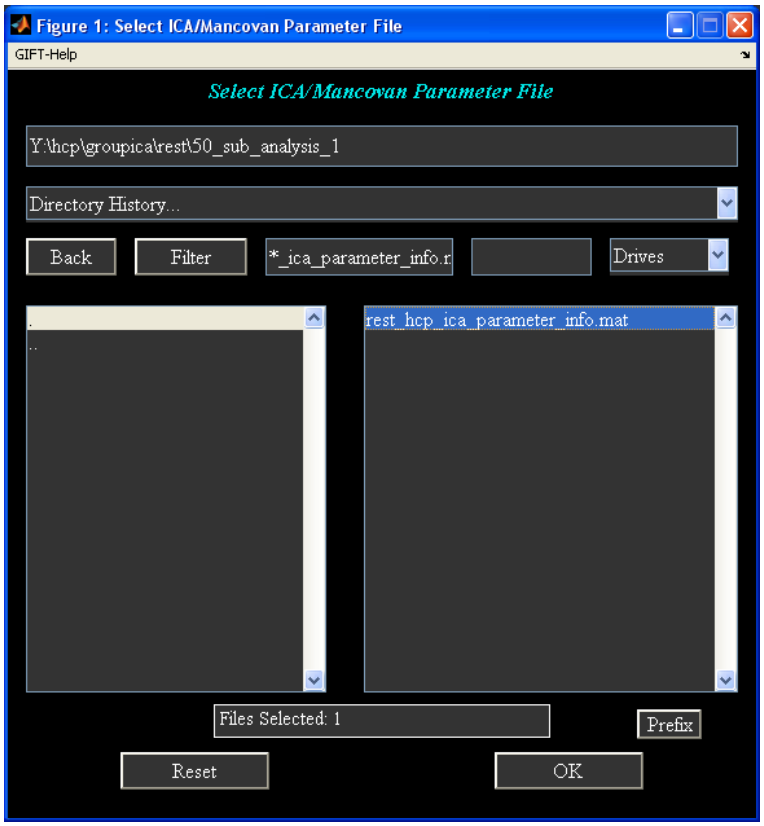


Figure 2.36: Figure window to select the ICA parameter file

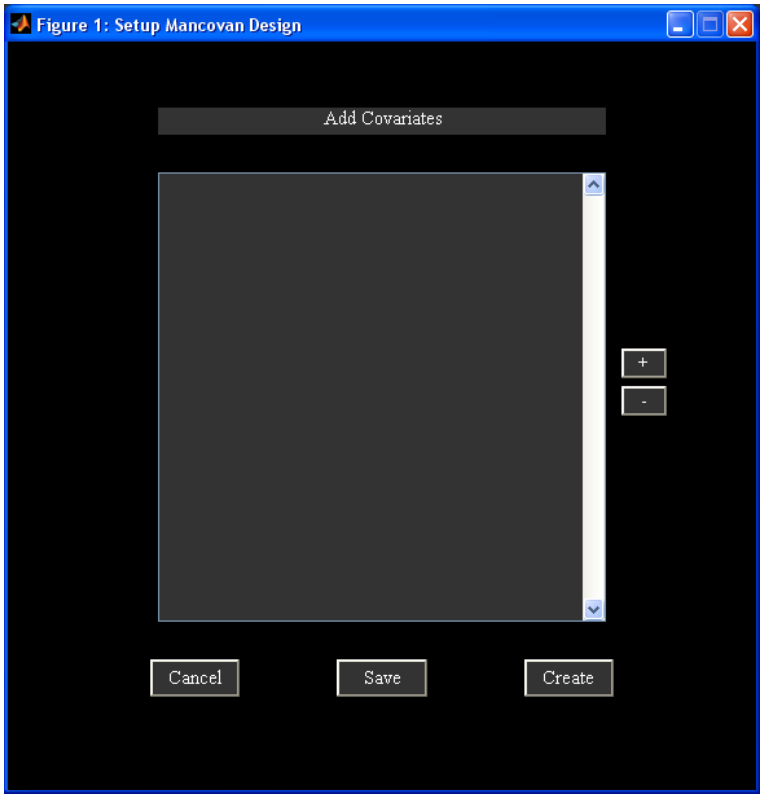


Figure 2.37: Setup Mancovan Design

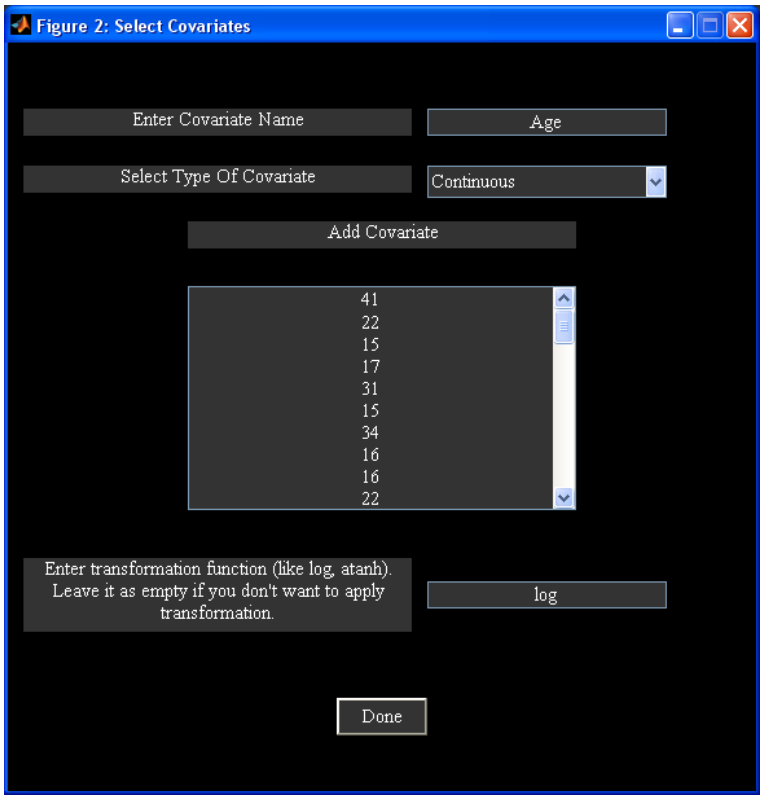


Figure 2.38: Add covariates

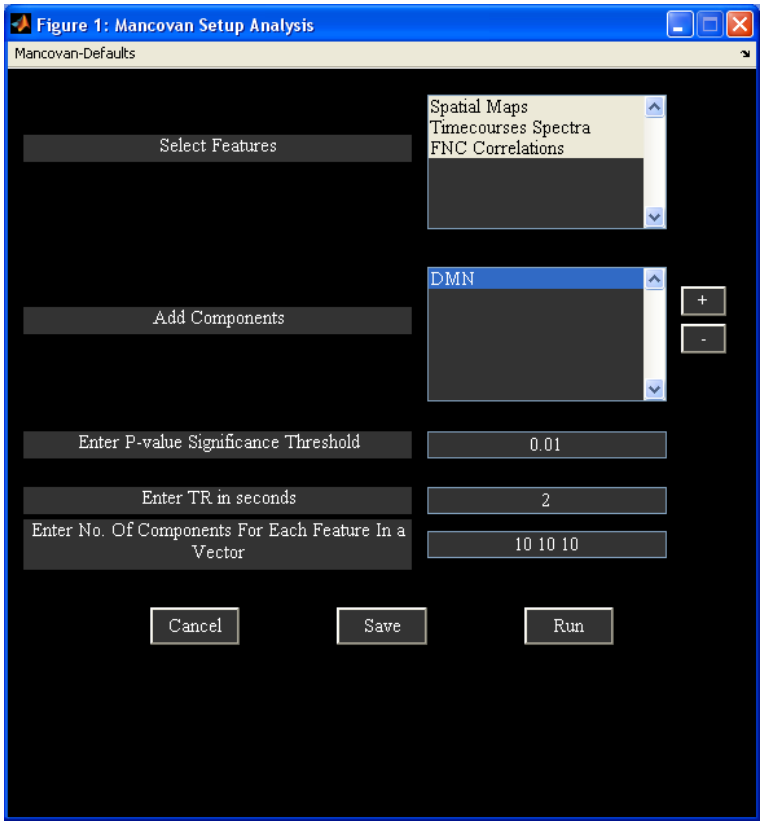


Figure 2.39: Setup features for mancovan analysis

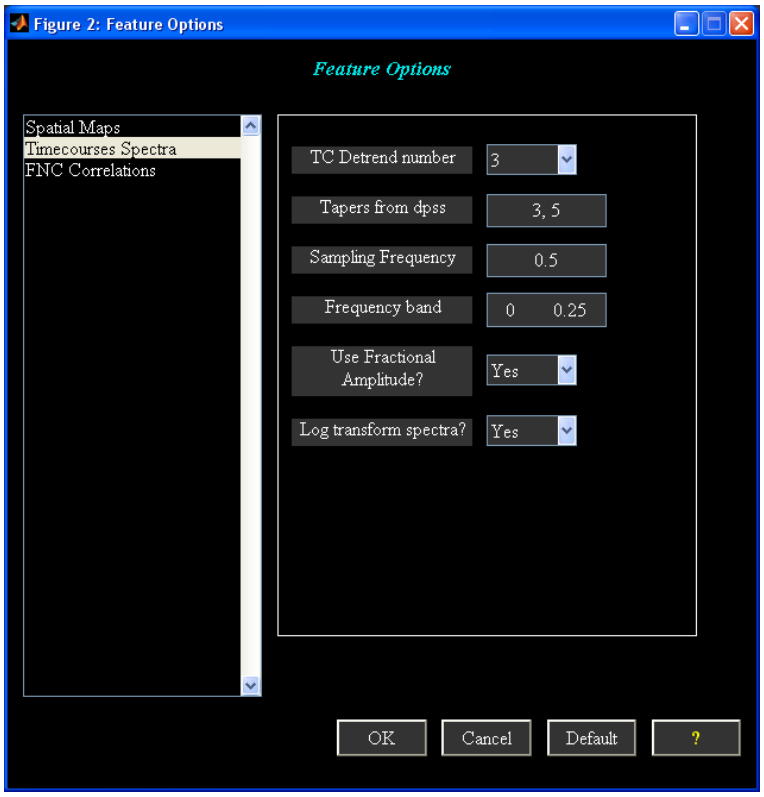


Figure 2.40: Mancovan Defaults

- Timecourses Spectra - The following are the options:
  - TC detrend number - Options are 0, 1, 2, and 3. Timecourses are detrended based on the detrend level.
  - Tapers from dpss - Multi-taper approach is used as implemented in Chronux ([9]), with the time-bandwidth product set to 3 and the number of tapers set to 5.
  - Sampling frequency - Default sampling frequency is set to 1/TR.
  - Frequency band - Default frequency band is set to [0, 1/(2*TR)].
  - Use Fractional amplitude - If the value is set to 'yes', each subject's spectra is normalized in the spectra dimension.
  - Log transform spectra - Option is provided to do log transformation on the spectra.
- FNC Correlations - Subject specific timecourses will be detrended and despiked using 3dDespike ([2]), then filtered using a fifth-order Butterworth low-pass filter with a high frequency cutoff of 0.15 Hz. You could turn off the default options, if you don't want to do pre-processing on the timecourses.

When you click *Run* button, computation is done on the features and the results are saved in each feature stats directory. The file names stored are as follows:

- Spatial Maps - T-maps are saved as `sm_stats\mancovan*tmap*img` and the spatial map parameters like mask and offset information is stored in `sm_stats\mancovan*results*sm*mat`.
- Timecourses spectra - Spectra information is stored in `spectra_stats\mancovan*results*spectra*mat`.
- FNC correlations - FNC correlations are stored in `fnc_stats\*mancovan_results_fnc.mat`.

Run Mancovan

Select the mancovan parameter file which was created after the end of setup features step. This will run mancova on each feature. Multivariate and univariate test results (`MULT` and `UNI`) are saved in the `*stats\mancovan*results*mat` files.

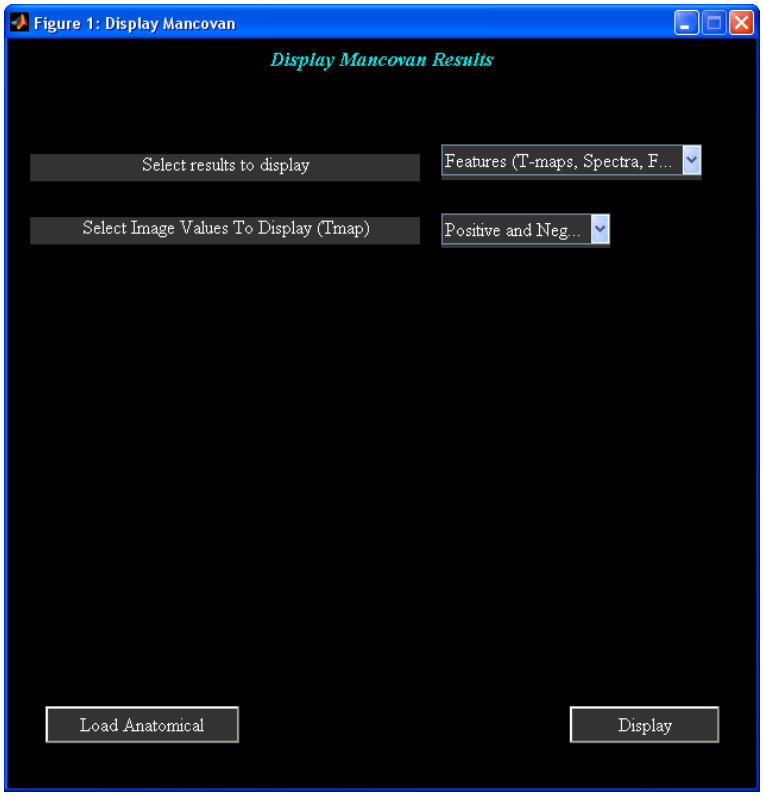


Figure 2.41: Display Mancovan Results

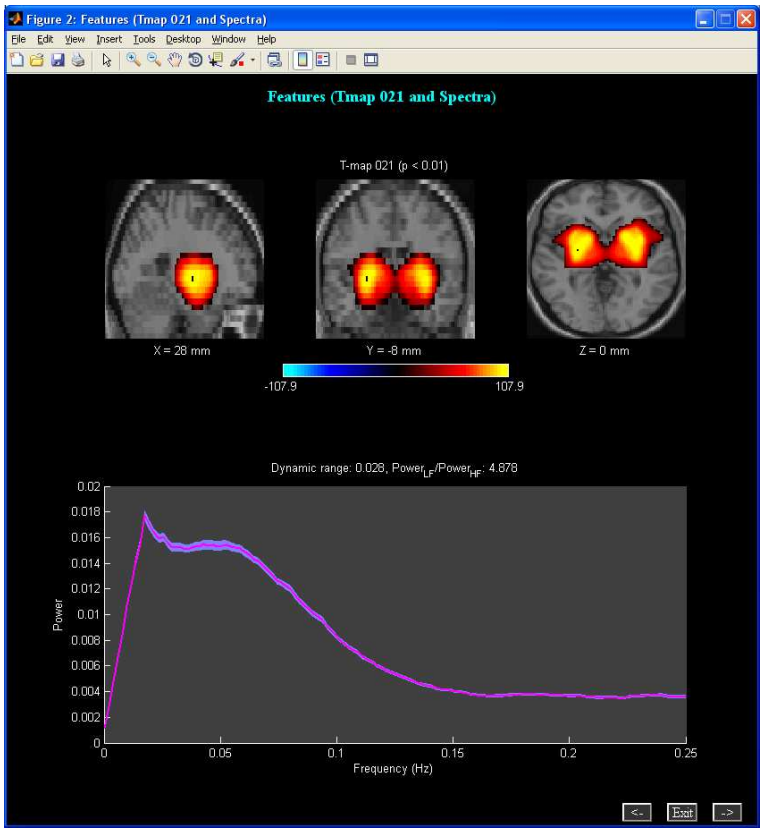


Figure 2.42: T-map and Spectra

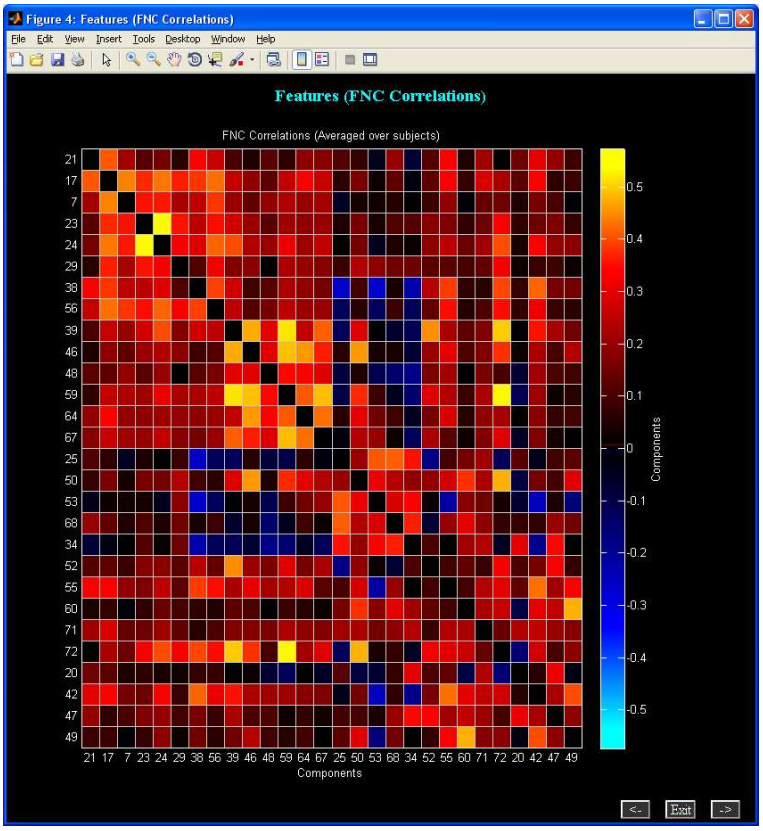


Figure 2.43: FNC Correlations

Figure 2.44: Multivariate results

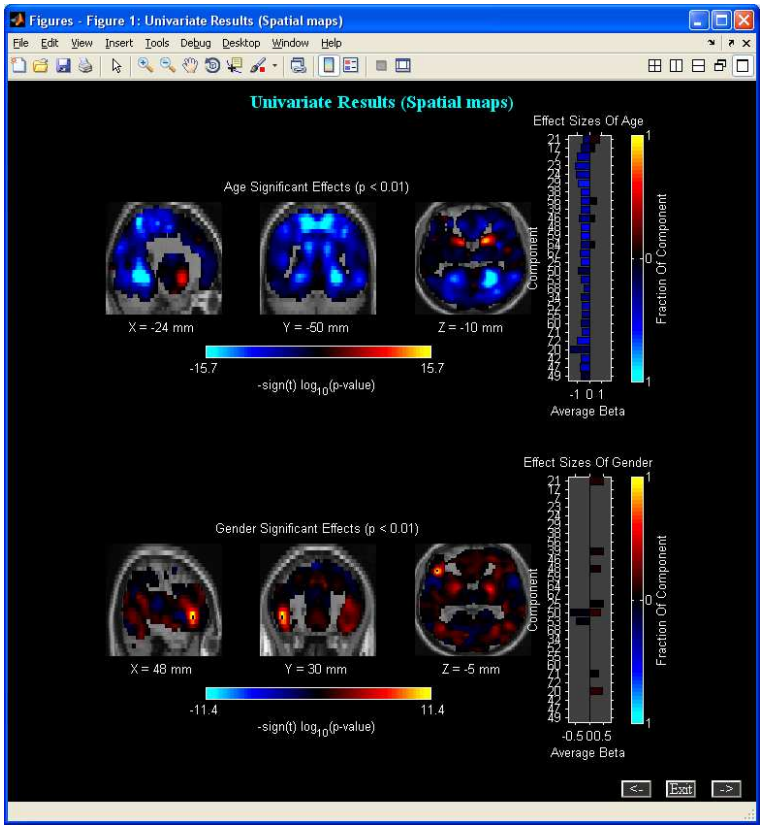


Figure 2.45: Univariate results of spatial maps.

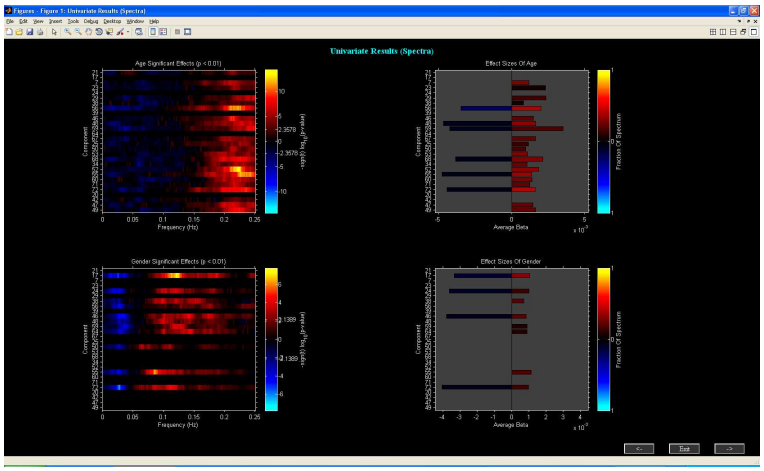


Figure 2.46: Univariate results of spectra.

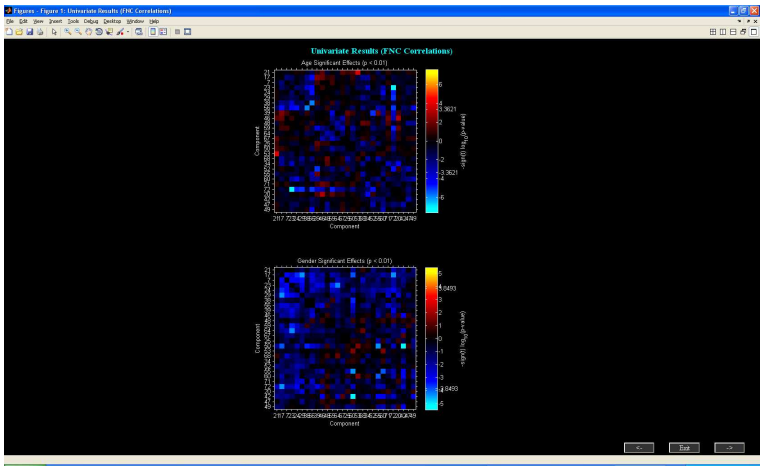


Figure 2.47: Univariate results of FNC.

Display

When you click *Display* button (Figure 2.35), Figure 2.41 will open.

- Features - Features will be displayed when you select features and click on *display* button. Figures 2.42 and 2.43 show ortho views of T-maps, power spectra, and FNC correlations. T-maps are thresholded using the same threshold as selected in the *Setup Features* option and you can interactively browse using mouse within each slice.

- Multivariate results - Significant results of covariates will be plotted in a matrix of dimensions covariates by components (Figure 2.44). The darker the color the more significant is the covariate in that component.

- Univariate results - Univariate results of age and gender covariates are shown for each feature in figures 2.45, 2.46 and 2.47.

  - Spatial maps - On the left side of the plot (Figure 2.45), composite maps of significant effects over all components are displayed as $-sign(t)log_{10}(p)$. Effects are considered significant if test statistics exceeded the FDR threshold ($\alpha = 0.01$) with a cluster extent of at least 27 contiguous voxels. Plots on the right side show average β values. β-values are averaged over significant clusters with effects of the same directionality and the color of the bar is proportional to the fraction of component voxels contributing to each effect.

  - Timecourses spectra - On the left side of the plot (Figure 2.46), covariates significance is shown as a function of frequency for each component displayed as $-sign(t)log_{10}(p)$. Dashed horizontal lines on the colorbar designate the FDR corrected threshold ($\alpha = 0.01$). On the right side of the plot, β-values are averaged over frequency bands with effects of the same directionality where test statistics exceeded the FDR threshold. The color of the bar is proportional to the fraction of contributing frequency bins. The absence of a bar indicates that univariate tests were not performed or test statistics were not significant.

  - FNC Correlations - Significance and direction of age and gender terms for each pairwise correlation, displayed as the $-sign(t)log_{10}(p)$. Dashed horizontal lines on the colorbar designate the FDR corrected threshold ($\alpha = 0.01$).

2.12.4 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.

2.12.5 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 2.11.1 for calculating event average with the variance removal ("Adjust ICA") tool.

2.12.6 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.

2.12.7 Spectral Group Compare

Group comparison of time courses is done by comparing the power spectra between the groups at different frequency bins. The results of the power spectra comparison is saved in a MAT file having suffix `*comparison_frequency_bins.mat`. The variables stored in the file are as follows:

- `mean_power_group1` - Mean power for group 1.
- `mean_power_group2` - Mean power for group 2.
- `tValues` - T-values. This variable is used for plotting the bar graphs.
- `pValues` - p-values.

Figure 2.48 shows the T-values of the components in groupings of four at different frequency bins. Bins are labeled using the variable `DEFAULT_TR_SPECTRAL_GROUP_COMPARE`.

2.12.8 Stats On Beta Weights

This utility is provided to test the significance of the component time courses by doing statistics on the beta weights, which are obtained after doing temporal sorting on all data-sets. Design criteria like one sample t-test, two sample t-test, one way anova (groups), one way anova (regressors), two way anova (groups, regressors) and Multiple Regression are provided. 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 2.49) 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 2.50.
- 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)', 'two way anova (groups, regressors)' and 'Multiple Regression'.

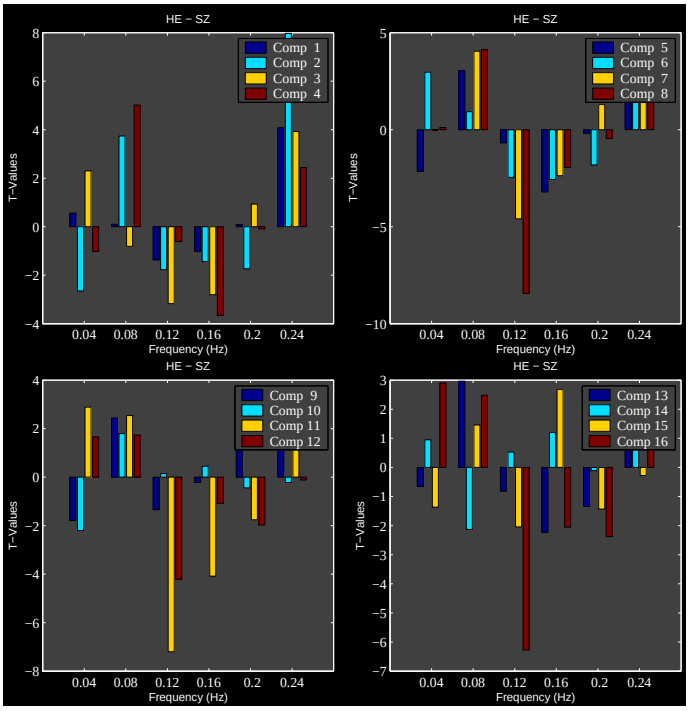


Figure 2.48: Power spectra difference between groups.

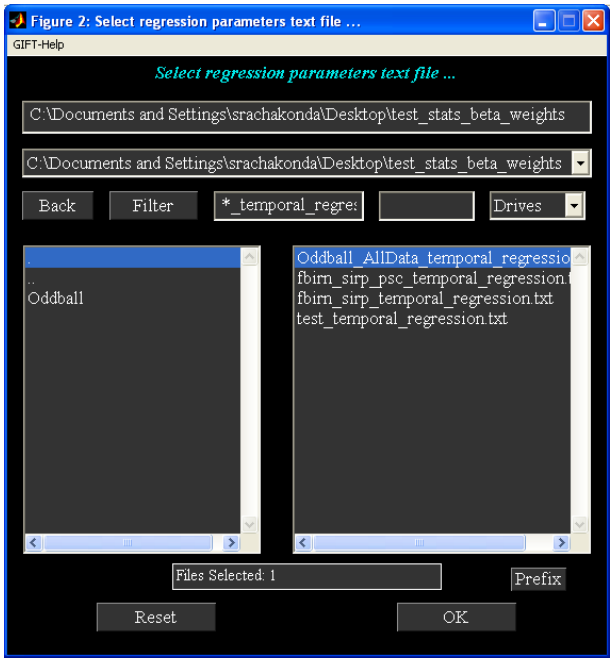


Figure 2.49: Figure window used to select the regression parameters text file.

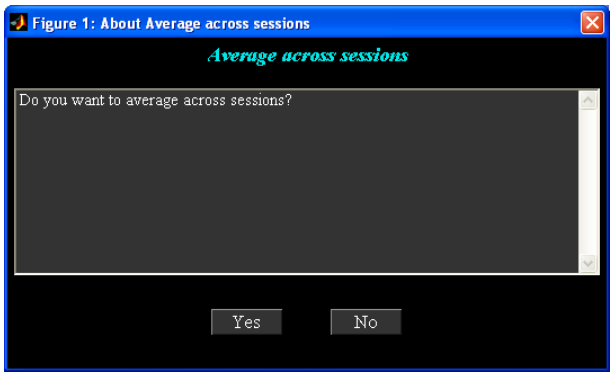


Figure 2.50: Option for averaging beta weights across runs or sessions.

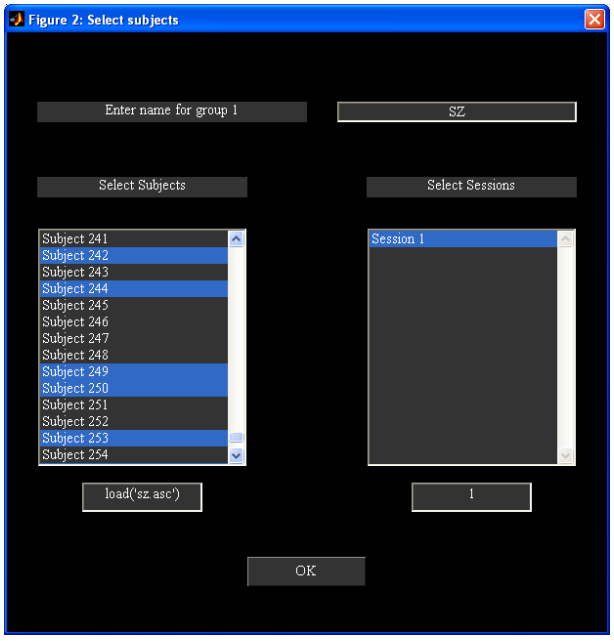


Figure 2.51: Figure window used to enter group name, subject and session numbers for a group.

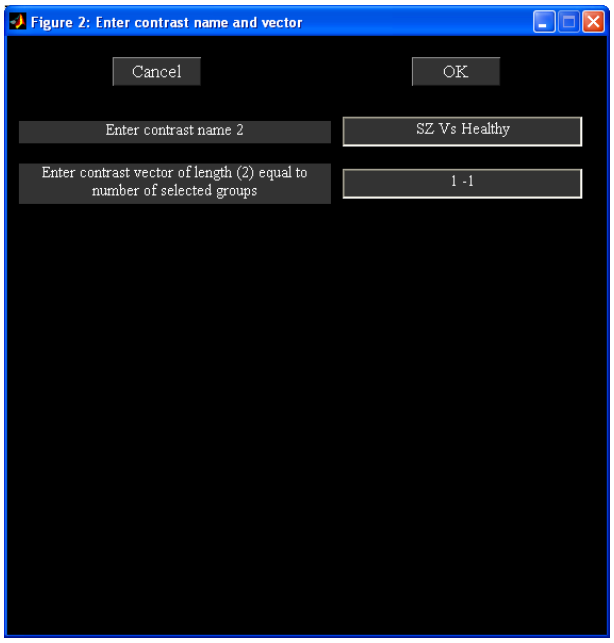


Figure 2.52: Figure window used to enter contrast name and vector.

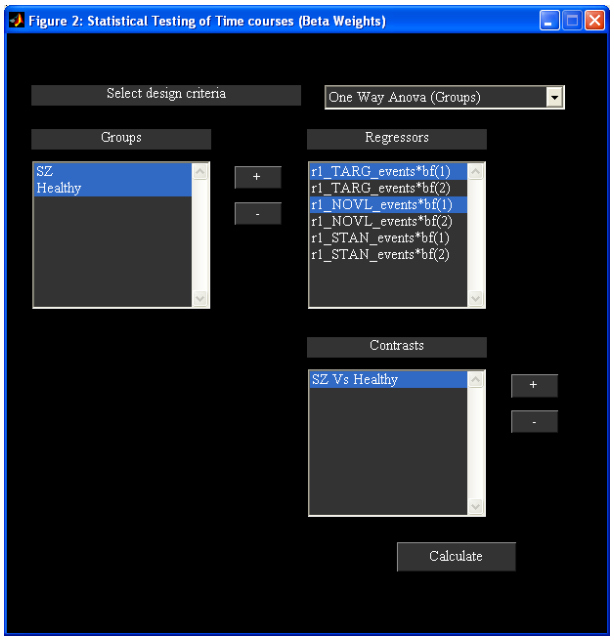


Figure 2.53: GUI shows the completed parameters.

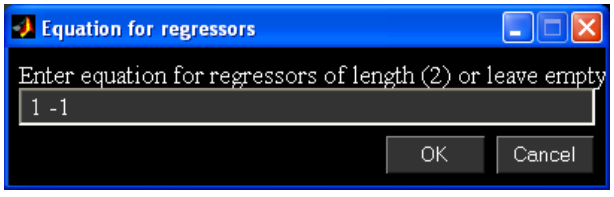


Figure 2.54: Option for doing operation on selected regressors.

- Groups Listbox - You can add groups using + button adjacent to groups listbox. Figure 2.51 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.
- Regressors Listbox - You can select only one regressor for one sample t-test and two sample t-test whereas for Anova and Multiple Regression you can select multiple regressors.
- Contrasts Listbox - This option is provided to do Anova contrasts. To add a contrast use + button (Figure 2.52) 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 use 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 2.53). 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 2.54) 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:

- When you select Multiple regression tool, a figure window will open to select the regressor file or files like age, test scores, etc. After you selected the regressor files, R-square statistic is calculated for each component. We also report the slopes and partial correlations of the regressors.
- Option is provided to do statistics on the beta weights using a batch file. Please see section 2.13.1 for more information.
- If you wish to use your favorite statistics package to do the statistics on the beta weights, `icatb_parse_regression_mat` function is provided that will write regression parameters in a excel file or load it in the MATLAB command window. Type help `icatb_parse_regression_mat.m` for more information.

2.12.9 SPM Stats

Please see section 2.13.1 for more information.

2.12.10 Write Talairach Table

We use Talairach Daemon software for generating talairach tables. When you select "Write Talairach Table" option in "Utilities" (Figure 2.1) drop down box, a figure window will open to select the image. We apply a threshold from variable `TALAIRACH_THRESHOLD` and write talairach coordinates for positive and negative regions in separate excel files.

2.12.11 Spatial-temporal Regression

GLM or ICA spatial maps are used as design matrix and the original data of subjects as observations to reconstruct subject components using Multiple Regression. This utility can also be invoked from the MATLAB command line.

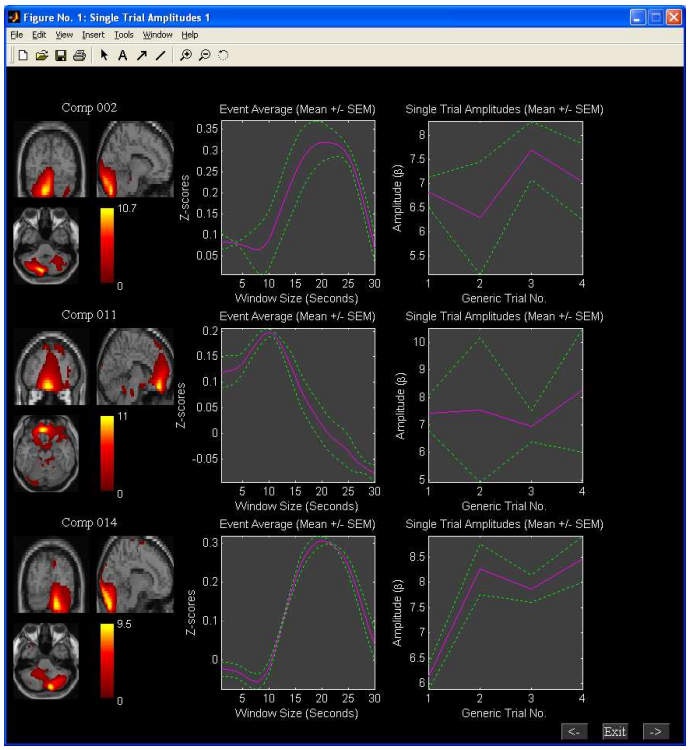


Figure 2.55: Plot contains orthogonal slices at the maximum voxel, event average and single trial amplitudes.

```
compFiles={'E:\Multiple_sub_Multiple_sess\sens\Sensorimotor_agg__component_ica_018.img',...
'E:\Multiple_sub_Multiple_sess\sens\Sensorimotor_agg__component_ica_019.img'};

inputFiles={'F:\smr1\swSm_nifti.nii', 'F:\smr2\swSm_nifti.nii'};

icatb_spatial_temp_regress(compFiles, inputFiles, 'outputDir', '.',...
'format', '.nii', 'outputPrefix', 'STR');
```

Where `compFiles` and `inputFiles` must be in a cell array. Each cell in `inputFiles` can contain a character array of 3D analyze images. The resulting subject components are saved in the specified format and the selected output directory. These components can be displayed in GIFT by using display buttons like *Component Explore*, *Orthogonal Viewer* and *Composite Viewer*.

2.12.12 Single Trial Amplitudes

Single trial amplitudes are the beta weights which are obtained by doing a Multiple Regression of the design matrix and the ICA time course where the design matrix is obtained by doing a convolution of the onsets and the event averages. This utility must be used after doing temporal sorting of all data-sets in the GIFT and using only the regressors of interest. The best regressor onsets for each data set are used to compute single trial amplitudes and event averages. We determine the best regressor for each data-set based on the maximum beta weight value (not the absolute value). Figure 2.55 shows the orthogonal slices at the maximum voxel, event average and the single trial amplitudes for the selected components. We also store the single trial amplitude results in a file having suffix `*single_trial_results.mat`.

2.12.13 *Z*-shift

Z-shift is used to center the image distribution to zero based on the peak of the distribution. Please select the file or files of interest. The function used for computing *Z*-shift is `icatb_convert_to_z_shift`.

2.12.14 Percent Variance

Percent variance utility can be used after running a group ICA analysis. Percent variance explained by the components in the data is calculated by doing a Multiple Regression of the BOLD signal and the component time courses at each voxel where the component time courses are treated as a model. After the calculation is done, the final result is printed to the MATLAB command window.

2.13 More Information

2.13.1 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:

- `perfType` - Group PCA performance settings. The best match for each option is dependent on the variable `MAX_AVAILABLE_RAM` in defaults file. Options are 1, 2 and 3.
  - 1 - Maximize performance.
  - 2 - Use less memory.
  - 3 - User specified settings will be selected based on the PCA options.
- `which_analysis` - Options are 1 and 2. If you select 2, group ICA is run using ICASSSO. The following are the options available:
  - `icasso_opts.sel_mode` - Selection mode. Options are "randinit", "bootstrap" and "both".
  - `icasso_opts.num_ica_runs` - Number of times you want ICA to be run.
- `dataSelectionMethod` - There are four ways to select the data and the SPM design matrix (required for temporal sorting). Options are 1, 2, 3 and 4. 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.
  - 3 - This option uses regular expressions to get the data set directories. The required variables are as follows:
    - `input_directory_name` - Full path of the directory where the subjects and sessions are located.

* `subject_dir_regexp` - Regular expression used for matching subject directories. This variable accepts nested paths. For example, to match single slice data in the example subjects directory, you can use `Sub\w+; single\w+` regular expression where semi-colon is used as a delimiter. If there are no subject directories inside the input directory, leave it as empty.
* `session_dir_regexp` - Regular expression used for getting the session directories inside the subject directories. Unlike `subject_dir_regexp` variable this cannot contain nested paths. If there are no session directories, leave this as empty.

  Note: More information on regular expressions is given in the Appendix 4.4. For a detailed information about regular expressions, please refer to the MATLAB help.

* `data_file_pattern` - File pattern used for getting data. Use wild card not regular expressions.
* `file_numbers_to_include` - This option will let you work on a subset of files. Enter the file numbers to include. Leave it as empty if you want to include all the files.
* `spm_stats_dir` - SPM stats directory name relative to subject or session directories. GIFT will automatically search in the appropriate directories to get `SPM.mat` file for the subject. Please note that this variable will be read only when you use `diff_sub_diff_sess` value for the variable `keyword_designMatrix`.

– 4 - This option is useful when you want to specify the file names directly. Each variable is explained below:

* `input_data_file_patterns` - File patterns in a cell of dimensions equal to number of subjects and sessions. Each new subject must be entered in a new row.
* `input_design_matrices` - This variable will be read only when you are using a different design between subjects and sessions.
* `dummy_scans` - Enter number of dummy scans to exclude from the analysis.

- `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).
- `preproc_type` - Specify type of data pre-processing.
  - 1 - Remove mean per time point
  - 2 - Remove mean per voxel
  - 3 - Intensity normalization
  - 4 - Variance normalization
- `pcaType` - Specify PCA type.
  - 1 - Standard
  - 2 - Expectation Maximization
  - 3 - SVD
- `pca_opts` - Optional parameters for PCA.

  - Standard PCA:
    * `pca_opts.stack_data` - Options are as follows:
      · 'Yes' - Data sets are stacked to compute covariance matrix. This option assumes that there is enough RAM to stack the data sets and for computing the covariance matrix. Full storage of covariance matrix is used.
      · 'No' - A pair of data sets are loaded at a time to compute covariance matrix. This option uses less memory usage but slower than the other option. You also have the option to store only the lower triangular portion of the covariance matrix (packed storage) when using this option.
    * `pca_opts.storage` - Options are 'Full' and 'Packed'. Packed storage scheme uses only lower triangular portion of symmetric matrix to compute eigen values.
    * `pca_opts.precision` - Options are 'Double' and 'Single'. Single precision is accurate up to 7 digits after decimal place and uses 50% less memory usage when compared to double precision.
    * `pca_opts.eig_solver` - Options are 'Selective' and 'All'. These options will be used only when you use packed storage scheme. Use option 'All' only when there are convergence issues with the option 'Selective'.
  - Expectation Maximization:
    * `pca_opts.stack_data` - Options are as follows:
      · 'Yes' - This option assumes that there is enough RAM to stack the data sets.
      · 'No' - A data-set is loaded at a time to compute the transformation matrix.
    * `pca_opts.precision` - Options are 'Double' and 'Single'.
    * `pca_opts.tolerance` - Enter stopping tolerance. Default is 1e-4.
    * `pca_opts.max_iter` - Enter maximum number of iterations to use.
  - SVD:
    * `pca_opts.precision` - Options are 'Double' and 'Single'.
    * `pca_opts.solver` - Options are 'Selective' and 'All'.

- `backReconType` - Options are 1, 2, 3 and 4.
  - 1 - Regular
  - 2 - Spatial-temporal regression
  - 3 - GICA3
  - 4 - GICA

- `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. You could select max, mean or median for a particular PC step in variable `estimation_opts`.

- `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, 2, 3 and 4.
  - 0 - No scaling
  - 1 - Scale components to percent signal change
  - 2 - z-scores
  - 3 - Scaling in timecourses
  - 4 - Scaling in maps and timecourses

- `algoType` - Currently there are 14 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
  - 12 - Combi
  - 13 - ICA-EBM
  - 14 - FBSS

- `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 six 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
  - 6 - Multiple Regression
- `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.
- `multi_regress_files` - You can specify regressor files like age, test scores, etc in a cell array. The format for the text files is ASCII or MAT.
- `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.

Mancova

Batch template for mancova is located at `icatb\icatb_batch_files\input_mancovan.m`. The following are the variables:

- `outputDir` - Output directory to place the analysis results.
- `ica_param_file` - Full path to ICA parameter file.
- `features` - Options are spatial maps, timecourses spectra and FNC correlations.

- `covariates` - Enter covariates information in a cell array of dimensions no. of covariates by 4. The four columns are covariate name, type, vector and transformation function. You could use file name instead of entering covariate vector. This file must be ascii for continuous covariates and a text file with new line, comma or tab as delimiter for categtorical covariates. Transformation function will be used only for continuous covariates.
- `interactions` - Specify pairwise model interactions. If you don't want to use interactions, you could leave it as empty.
- `comp_network_names` - Component networks. Use name by values where values are component numbers. The names information will be used in future release for advanced plotting.
- `numOfPCs` - Number of principal components used to reduce the data in voxel or spectra dimension which will be used in multi-variate tests. The vector length must match the no. of features.
- `p_threshold` - Significance threshold used in multivariate and univariate tests.
- `TR` - TR of the experiment in seconds.

After entering the parameters, use `icatb_mancovan_batch(inputFile);` at the command prompt.

Stats On Individual Subject Component Maps

You can test the significance of components by doing a one sample t-test or two sample t-test on images. We now provide an option in GIFT to do one sample t-test or two sample t-test on images using SPM5.

- One sample t-test - One sample t-test for each component will be automatically calculated if you set variable `SPM_STATS_WRITE_TAL` in `icatb_defaults.m` to 1 or 2. If you set variable `SPM_STATS_WRITE_TAL` value to 1 only one sample t-test will be calculated whereas a value of 2 will also write talairach tables for the t-map.

  Note: An option is also provided in "Utilities" (Figure 2.1) drop down box to calculate t-maps for the selected data-sets.

- Two sample t-test - When you click "Utilities" (Figure 2.1) drop down box and have selected "SPM Stats" as the option, a figure window will open to select the design criteria. When you select design as "Two sample t-test", two sample t-test for a component will be calculated between the selected groups by using a explicit mask. The explicit mask is calculated by applying a threshold (`SPM_STATS_TTEST_THRESHOLD`) on t-map obtained by doing a one sample t-test on data-sets.

  Note: When you set `SPM_STATS_TTEST2_EXPLICIT_MASK` variable in `icatb_defaults.m` file to 0, explicit mask is not used.

2.13.2 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.

  Note: To load the image files, use `icatb_loadData(file_name)`.

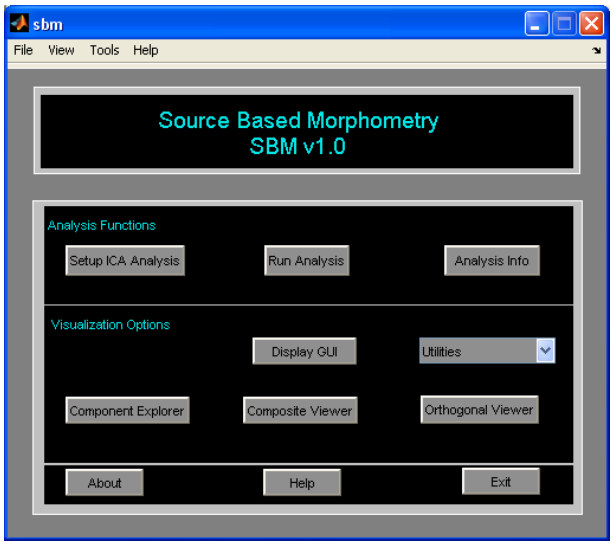


Figure 2.56: Source Based Morphometry

2.14 Source Based Morphometry

Source Based Morphometry (SBM) is a multivariate tool to study the gray matter differences between the patients and controls ([1] and [18]). ICA is used on the subject images to determine the maximally independent sources. Basically ICA decomposes data into subject loading coefficients and component maps. It is similar to doing single subject single session analysis in the GIFT except the timepoints are subject images. To invoke SBM (Figure 2.56), type `sbm` or `groupica smri` at the MATLAB command prompt. The following are the differences between the GIFT and SBM:

- Setup ICA - Default mask used in the SBM includes voxels greater than or equal to 1% of mean of the data.
- Display - Subject explorer is excluded in the display GUI.
- Output files naming - Component maps are stored with the suffix `*group*component*ica*` in Analyze or Nifti format. Subject component loading coefficients are stored with the suffix `*group*loading*coeff*` in Analyze or Nifti format. To load the image files, use `icatb_loadData(file_name)`.
- Batch template is provided in `icatb\icatb_batch_files\Input_sbm.m`. Specify `modalityType` as 'smri' and enter the parameters similar to one subject one session analysis as in the GIFT. After entering the parameters, use `icatb_batch_file_run(inputFile)` at the MATLAB command prompt.
- Only remove components, write talairach tables and ICASSO utilities are included.

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*63*34*220 to 53*63*34*50. Two or three data reduction steps are used for multiple subjects (See Table 2.1).

- Two data reduction steps - After each subject's functional data is reduced, the subjects are then concatenated into one group and put through another data reduction step. Figure 3.1 shows three data-sets passed through two data reduction steps.
- Three data reduction steps - 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. Reduced data from the groups are stacked into one group and put through the final data reduction.

  Note: Two data reduction steps is the recommended approach for group PCA.

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 ([8]). 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 ([21]) using Infomax, FastICA, Jade, SIMBEC and AMUSE.

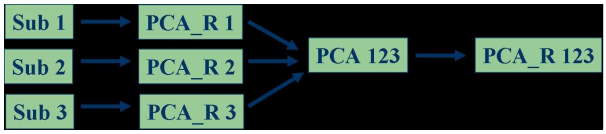


Figure 3.1: First each data-set is reduced using PCA where R term indicates reduced data. These three reduced data-sets are concatenated into one group (or partitions). The grouped data is passed through another data-reduction step.

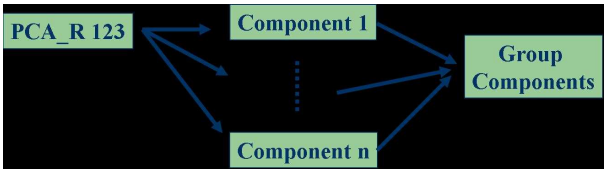


Figure 3.2: Reduced data-set from the last step in the data-reduction stage is used and ICA algorithm is applied. The components represent group and not of individual subjects. The next stage is Back-Reconstruction.

3.2.1 Infomax

Infomax ([3]) 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 ([29]) 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 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 ([6]) 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*, *pow*3, *gauss* and *skew* non linearities can be used ([7]).
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*, *pow*3 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) ([16]), 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 ([8]).
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) ([26]), 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

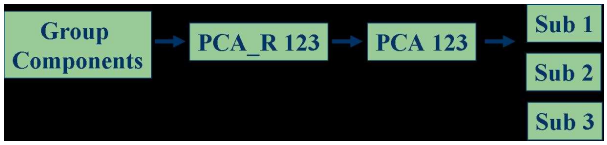


Figure 3.3: The individual subject components are calculated from the group components.

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) ([19]), 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'. 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) ([25]) 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 ([23]) 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 ([24]) 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.

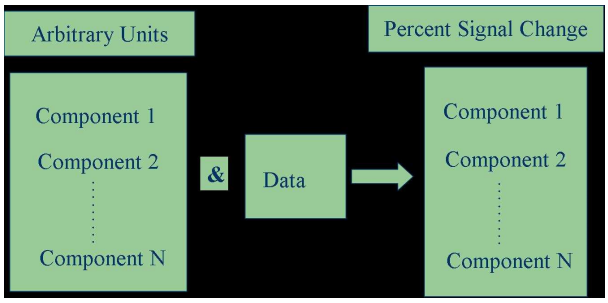


Figure 3.4: The individual subject components are calibrated to percent signal change using the original fMRI data as reference.

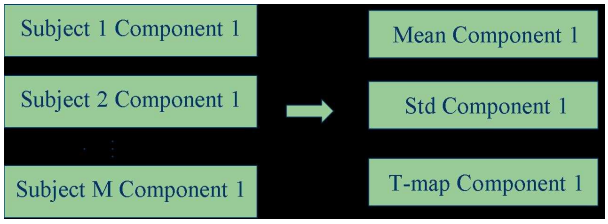


Figure 3.5: Mean, standard deviation and t-maps are calculated for each component over the data-sets used.

3.4 Scaling Components

The spatial maps and time courses of components have arbitrary units after the back reconstruction step. Component spatial maps and timecourses can be scaled using percent signal change, z-scores, scaling in timecourses or scaling in timecourses and maps. Figure 3.4 shows the spatial maps and timecourses scaled to percent signal change.

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. 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. Caprihan, C. Abbott, J. Yamamoto, G. D. Pearlson, N. Bizzozero, J. Sui, and V. D. Calhoun, "Source-based morphometry analysis of group differences in fractional anisotropy in schizophrenia," Brain Connectivity, In Press

[2] http://afni.nimh.nih.gov/afni

[3] 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.

[4] A. R. Franco, A. Pritchard, V. D. Calhoun, A. R Mayer, "Inter-rater and Inter-method Reliability for Selecting the Default Mode Network during Data-Driven Analyses", HBM, vol. 30, pp. 2293-2303, 2009.

[5] A. Garrity, G. D. Pearlson, K. McKiernan, D. Lloyd, K. A. Kiehl, and V. D. Calhoun, "Aberrant default mode functional connectivity in schizophrenia" Am.J.Psychiatry, vol. 164, pp. 450-457, 2007.

[6] A. Hyvarinen and E. Oja, "A fast fixed-point algorithm for independent component analysis," Neural Comput., vol. 9, no. 7, pp. 1483-1492, 1997.

[7] A. Hyvarinen, "Fast and robust fixed-point algorithms for independent component analysis," Neural Networks, vol. 10, pp. 626-634, 1999.

[8] A. Cichocki, S. Amari, K. Siwek, T. Tanaka et al., "ICALAB Toolboxes," http://www.bsp.brain.riken.jp/ICALAB.

[9] http://chronux.org

[10] D. Glahn, SPM fMRI Analysis, 2000.

[11] EEGLAB Toolbox, http://sccn.ucsd.edu/eeglab/

[12] E. Allen, E. Erhardt, E. Damaraju, W. Gruner, J. Segall, R. Silva, M. Havlicek, S. Rachakonda, J. Fries, R. Kalyanam, A. Michael, J. Turner, T. Eichele, S. Adelsheim, A. Bryan, J. R. Bustillo, V. P. Clark, S. Feldstein, F. M. Filbey, C. Ford, K. Hutchison, R. Jung, K. A. Kiehl, P. Kodituwakku, Y. Komesu, A. R. Mayer, G. D. Pearlson, J. Phillips, J. Sadek, M. Stevens, U. Teuscher, R. J. Thoma, and V. D. Calhoun, "A baseline for the multivariate comparison of resting state networks," Frontiers in Human Neuroscience, vol. 1, p. 12, 2011

[13] E. Allen, E. Erhardt, T. Eichele, A. R. Mayer, and V. D. Calhoun, "Comparison of pre-normalization methods on the accuracy of group ICA results," in Proc. HBM, Barcelona, Spain, 2010.

[14] E. Erhardt, S. Rachakonda, E. Bedrick, T. Adali, and V. D. Calhoun, "Comparison of multi-subject ICA methods for analysis of fMRI data," in Proc. HBM, Barcelona, Spain, 2010.

[15] ICASSO Toolbox, http://www.cis.hut.fi/projects/ica/icasso/

[16] J. F. Cardoso and A. Souloumiac, "Blind beamforming for non gaussian signals," IEE-Proc-F, vol. 140, no. 6, pp. 362-370,

[17] J. Knight, "Signal Fraction Analysis and Artifact Removal in EEG", Colorado State University, Fort Collins, Colorado, MS Thesis 2003.

[18] L. Xu, K. Groth, G. Pearlson, D. Schretlen, and V. Calhoun, "Source Based Morphometry: The Use of Independent Component Analysis to Identify Gray Matter Differences with Application to Schizophrenia," Hum Brain Mapp, vol. 30, pp. 711-724, 2009.

[19] 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.

[20] 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.

[21] 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.

[22] N. Filippini, B. J. MacIntosh, M. G. Hough, G. M. Goodwin, G. B. Frisoni, S. M. Smith, P. M. Matthews, C. F. Beckmann, and C. E. Mackay, "Distinct patterns of brain activity in young carriers of the APOE-epsilon4 allele," Proc Natl Acad Sci USA, vol. 106, pp. 7209-7214, Apr. 2009.

[23] 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.

[24] 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.

[25] S. Cruces, L. Castedo, A. Cichochki, "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.

[26] 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.

[27] S.Roweis, "EM algorithms for PCA and sensible PCA", Advances in Neural Information Processing Systems. 1998.

[28] Statistical Parameteric Mapping, http://www.fil.ion.ucl.ac.uk/spm/.

[29] 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.

[30] 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.

[31] 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.

[32] 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.

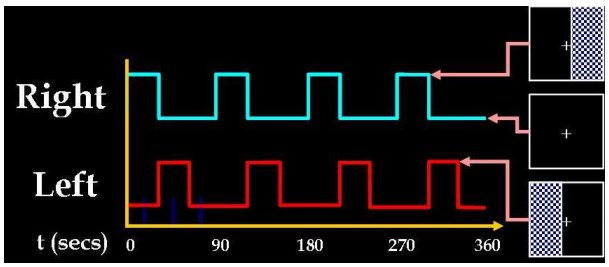


Figure 4.1: Visuomotor paradigm.

  - 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 SMOOTHPARA 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.

- CENTER_IMAGES: A value of 1 means subject spatial maps will be centered based on the peak of the distribution.

- PREPROC_DEFAULT: Data pre-processing default.

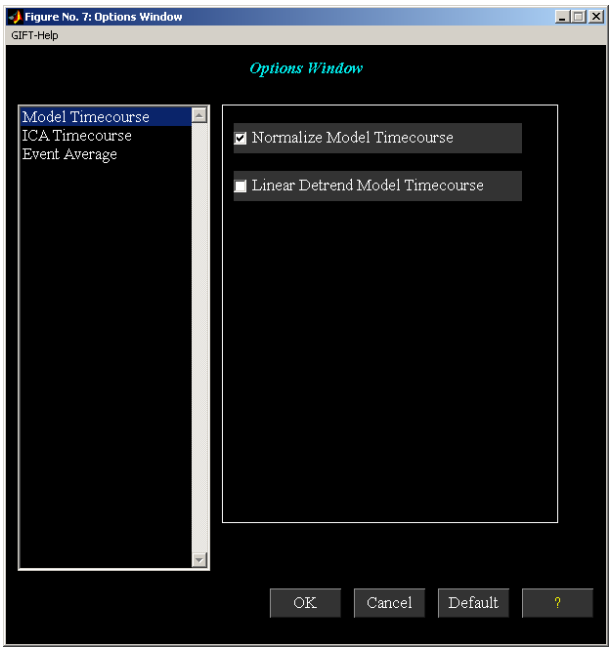


Figure 4.2: Time course options.

  - 1 - Remove mean per time point
  - 2 - Remove mean per voxel
  - 3 - Intensity normalization
  - 4 - Variance normalization

- `PCA_DEFAULT` - PCA default.
  - 1 - Standard
  - 2 - Expectation Maximization

- `BACKRECON_DEFAULT`: Backreconstruction default.
  - 1 - Regular
  - 2 - Spatial-temporal Regression
  - 3 - GICA3
  - 4 - GICA

- `SCALE_DEFAULT` - Scaling components default.
  - 0 - No scaling
  - 1 - Scale To Original Data(%)
  - 2 - Z-scores
  - 3 - Scaling in Timecourses
  - 4 - Scaling in Maps and Timecourses

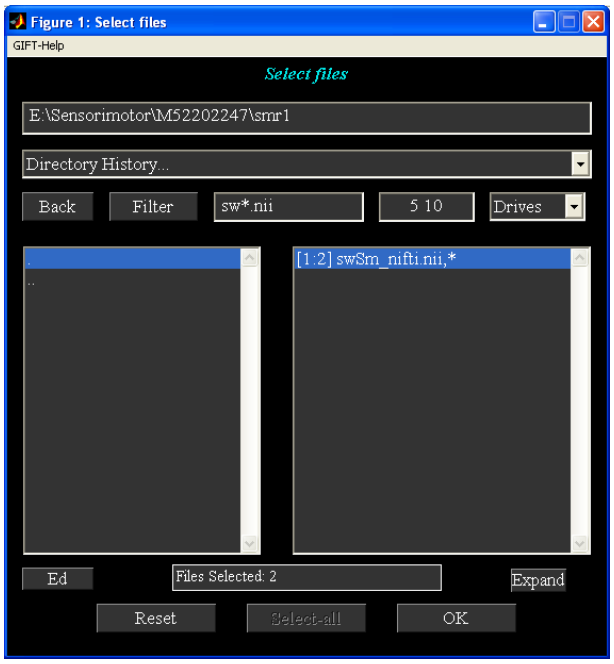


Figure 4.3: Figure window used to select a directory, file or files.

4.3 Options Window

When you click "Timecourse Options" sub menu in Options Menu (Figure 2.20), 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 2.20). *Cancel* button closes the figure window. *Default* button uses the defaults from `icatb_defaults.m` file.

4.4 Regular Expressions

Regular expressions are used for pattern matching. Some of the regular expressions are given below:

- `\w+` - Match strings containing alphabets, numerals or underscore characters like `M8710345`, `M8710g45_aod`, etc.
- `^Study\w+` - Match strings that start with `Study` followed by alpha numerals or underscore characters.
- `\<\d\>` - Match strings containing exactly one numeral like 1, 2, etc.
- `aod.*` - Match strings that contain `aod` followed by zero or more characters.
- `aod.+` - Match strings that contain `aod` followed by one or more characters.
- `^S1.*V1$` - Match strings that start with `S1` and end with `V1`.

4.5 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:

- 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.

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