An independent component analysis based tool for exploring functional connections in the brain

S. M. Rolfe[a], L. Finney[b], R. F. Tungaraza[b], J. Guan[b], L.G. Shapiro[b], J. F. Brinkely[b], A. Poliakov[c], N. Kleinhans[d], E. Alyward[d]

[a] Electrical Engineering, Box 352350, U. of Washington, Seattle, WA, USA 98195;
[b]Computer Science and Engineering, Box 352350, U. of Washington, Seattle, WA, USA 98195;
[c]Neurological Surgery, Children's Hospital, 4800 Sand Point Way, NE, Seattle WA, USA 98105
[d]Radiology, Box 375115, U. of Washington, Seattle, WA, USA 98195;

ABSTRACT

This work introduces a MATLAB-based tool developed for investigating functional connectivity in the brain. Independent component analysis (ICA) is used as a measure of voxel similarity which allows the user to find and view statistically independent maps of correlated voxels. These maps of correlated voxel activity may indicate functionally connected regions. Specialized clustering and feature extraction techniques have been designed to find and characterize clusters of activated voxels, which allows comparison of the spatial maps of correlation across subjects. This method is also used to compare the ICA generated images to fMRI images showing statistically significant activations generated by Statistical Parametric Mapping (SPM). The capability of querying specific coordinates in the brain supports integration and comparison with other data modalities such as Cortical Stimulation Mapping and Single Unit Recordings.

Keywords: Functional imaging

1. DESCRIPTION OF PURPOSE

The Structural Informatics Group at the University of Washington has developed software tools for processing, integrating, and visualizing multimodality data for surgical planning and the study of language mapping in the brain [6]. Data for this project was collected from patients who were about to undergo surgery for intractable epilepsy in order to remove the foci of the seizures. Three main data types were collected from each patient. Before surgery, fMRI volumes showing brain activations for a language task and structural MRI volumes showing the detailed anatomy of the brain were collected. During surgery, data was collected using Cortical Stimulation Mapping (CSM) to pinpoint language areas to be avoided during surgery.

One goal of the project is to integrate these data types to allow for analysis of the relationships between the data sources. Identification of regions in an fMRI image volume associated with the language areas identified by CSM could provide a non-invasive alternative to CSM for epileptic patients and would advance the study of language mapping in the normal brain. Since it is expected that language regions behave as a functionally connected network, it is useful to be able to search for functional connectivity between voxels of interest in fMRI image volumes and to compare networks of connectivity across subjects [3]. The tool developed for this purpose has also been used to analyze fMRI scans from patients with autism. In this application, the tool is useful for identifying independent activation patterns of interest in one subject and investigating the rate of occurrence of this activation pattern among subjects with the same condition and a control group.

Six queries have been designed to allow the user to explore regions of similar activation within and across subjects and to compare these to fMRI images showing statistically significant activations generated by Statistical Parametric Mapping (SPM):

1. Starting with SUR, CSM, or fMRI data, select an (x, y, z) coordinate of interest in the brain. Get the raw data fMRI time series of the voxel at that location. Use signal similarity measures to find correlated voxels within the subject's brain.
2. Starting with one subject's voxel correlations, find SPM-generated statistical activation images from that patient with similar spatial patterns of activation.

3. Starting with one subject's voxel correlations, search for other subjects who have a similar correlation pattern for a voxel in the same region.
4. For a given subject's statistical activation image and given location, find other subjects who have greater than or equal activation values at that location, by searching the SPM images showing statistically significant activations.
5. For a given subject's statistical activation image, find other subjects who have similar spatial patterns of activation, by searching the SPM images showing statistically significant activations.
6. For a given subject's statistical activation image, find signal-similarity-generated correlation patterns from that subject with similar spatial patterns of activation.

2. METHODS

There are two main tasks associated with the queries supported by this interface. The first task is the measurement of voxel similarity. This is required by the first query, which allows the user to select a region of interest (ROI) in a subject's brain and returns a map of other regions in the same subject's brain with similar activation patterns. The second task is the measurement of similarity between spatial activation maps. This measurement is needed for the remainder of the queries, which allow the user to select an ICA or SPM activation map.   It returns similar maps of the selected data type from the database, using a clustering algorithm and feature extraction method developed for this application.

2.1 Voxel Similarity Measure

Independent component analysis (ICA) is a technique used to separate statistically independent source signals from a set of observations [2]. When applying ICA to fMRI data, temporal or spatial independence can be used to model the independent components [5]. Spatial ICA models the entire observed data set of N volume images as a mixture of N independent voxel maps. Each voxel map is associated with one time course. The observed data X can be modeled as a matrix of independent spatial component maps, $C,$ multiplied by a random mixing matrix $M,$

$$X_{ji} = \sum_{k=1}^{V} M_{jk} C_{ki} .$$

Algorithms which perform ICA estimate both the mixing matrix and the independent components from the observed data. The independent components are contained in the TxV matrix, C, where each of the T rows contains a spatial volume of V voxels. Each volume represents a statistically independent pattern of activity. The voxels activated in the IC pattern share the activation pattern contained in the associated time course over the duration of the time course. Regions where significant clusters of activation have been identified in one independent component are determined to be similar, as they are likely to share one source of activation.

2.2 Spatial Map Similarity Measure

 For this application, similarity between two spatial maps is defined as similarity between the locations and characteristics of their activated voxel clusters. The first step in measuring spatial map similarity is to identify the voxel clusters in a spatial map and to extract their relevant cluster features. Before clustering, the image is preprocessed to improve identification of cluster centers. A threshold based on the number of standard deviations from the mean voxel value is applied to the image to locate activated voxels. Voxels determined to be activated are assigned a value of one and the remainder zero. This binary labeling eliminates the influence of activation level when identifying the clusters. The most interesting activations often make up only five to ten percent of the raw data, so weighting by activation value is not useful for finding cluster centers. After the threshold has been applied, the image is filtered to emphasize the significance of voxels located in groups. A brain activation is expected to involve a cluster of voxels, so an activated voxel is more likely to be significant when its neighbors are also activated. To utilize this information, each activated voxel x_i is weighted using the non-linear function

$$x_i = e^n ,$$

where n is the number of neighboring voxels activated. After the preprocessing, modified k-means clustering is used to find the activation centers. Unlike standard k-means clustering, the contribution of each point to the mean used to find the new bin location is dependent on the weight x_i at that location. This makes use of a priori information about activation structure to locate the centers. Each activation location i in a bin is weighted by

$$w_i = \frac{x_i}{\sum_{j=1}^{n} x_j},$$

where n is the number of activation locations in the bin, and x_j is the value at voxel location j in the weighted, thresholded image. The weights in each bin sum to one, so the sum of the weighted locations is the weighted mean. This process is repeated until the bin centers converge. Once the activation clusters have been located, five features are extracted from each cluster and are stored in a feature vector.

1. Location of bin center: This feature gives the *(x, y, z)* coordinates of the cluster center found by the clustering algorithm.
2. Number of activation locations in bin
3. Average distance from bin center
4. Average bin weight
5. Weighted variance of distances: The variance is weighted so that locations with larger weights contribute more to the variance calculation and is given by

$$\sum_i \frac{dist_i * weight_i - E[dist_i * weight_i]}{\sum_j dist_j * weight_j},$$

where $dist_i$ is the distance between the bin center and activation location i and $weight_i$ is the value at location i.

2.3 Extracting Feature Weights

Once the feature vectors have been extracted, each feature is assigned a weight [4]. The elements in the feature vector have different scales and distributions. Before a feature vector difference can be calculated, the quantities should be normalized so that one feature does not dominate the difference calculation. Each feature is therefore weighted according to its mean and standard deviation

$$w_k = \frac{1}{F_k} std(\frac{F_k}{F_k}),$$

where w_k is the calculated weight for feature k, and F_k is the set of all values for feature k. In the interest of having a dynamic system, the weights are calculated on a per-subject basis. F_k encompasses all values for all IC activation maps for that subject.

2.4 Feature distance measure

Once the elements of the feature vector have been normalized for comparison, the distance between IC activation maps can be calculated. To compare an activation map i to a second activation map j, the minimum distance to each cluster k in activation map i will be found using

$$d_{ik} = \min_n(abs(f_{ik} - f_{jn}))$$

where f_{ik} is the feature vector for cluster k in activation map i and n is the total number of clusters in activation map j. If the distance d_{ik} is less than a threshold value of 50, cluster k in activation map i is determined to have a match in activation map j. The total quality of the match between activation maps i and j is given by

$$m_{ij} = \frac{1}{abs(n_i - n_j)} \sum_k d_{ik} * binsize_k * averagebinweight_k,$$

where n_i is the number of clusters in activation map i and $binsize_k$ and $averagebinweight_k$ are two elements of the feature vector f_{ik} which describe the size and weight of the cluster. The cluster matches are weighted by the size and weight of the cluster in i to determine their relative importance to the overall image match. The final distance measurement between the activation maps i and j is calculated by repeating this process to find m_{ji}, the distance between maps j and i, and averaging the value of the matches from each direction,

$$Dist_{ij} = \frac{m_{ij} + m_{ji}}{2}.$$

3. RESULTS

The system described in this paper was evaluated using fMRI datasets from twenty-three subjects. Only normal subjects were used in the evaluation to control for differences due to brain abnormalities.

3.1 Query 1

Starting with SUR, CSM, or fMRI data, select an (x, y, z) coordinate of interest in the brain. Get the raw data fMRI time series of the voxel at that location. Use signal similarity measures to find correlated voxels within the subject's brain.

This query is used to locate regions in the brain that may be functionally connected to a region of interest. Functional connectivity, the correlations between spatially remote neural events, is observed when activity in two brain regions covaries. This covariance is measured by comparing the voxel time courses, the raw data value at a voxel position at each sample point in one scan, aligned in time. To execute this query the user chooses a subject from the database and uses the tool interface to scroll through sagital, coronal, and axial views of the structural MRI. Once a region of interest is identified, the user draws a two dimensional box around this area on any of the three views. This box outlines the length and width of a three-dimensional region of interest with a predefined depth of five samples, centered on the slice where the user has drawn the box. When the query is executed, ICA-generated activation images from the selected subject with a cluster of voxels activated above a predefined threshold in the region of interest are returned to the user in the re-wrapped, three-dimensional format. These patterns describe independent activation maps that are significantly contributing to observed activations in that region.
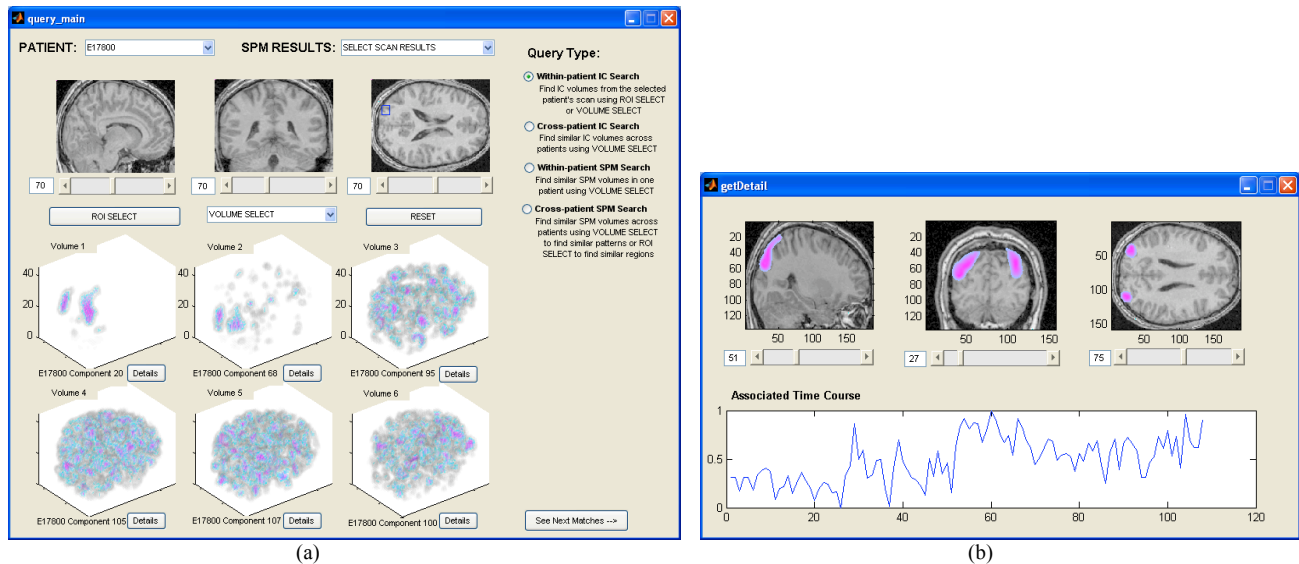

Figure 1.(a) The main user interface used to execute queries showing query type 1 and (b) the detail browser window showing the first match returned in the sample query.

A sample query of type 1 is shown in Figure 1. The selected subject's structural MRI is shown in the top row of the main user interface window in Figure 1 (a). The region of interest was chosen on the axial view and can be seen outlined in blue. The database was then queried for ICA-generated activation patterns from the same subject with significant activation in this area. The results are displayed below the structural data in Figure 1 (a). The user can choose to map any of the results back onto the subject's structural MRI in a detail browser window, shown in Figure 1 (b). In this example, Volume 1 from the main interface is shown in the detail browser in Figure 1 (b). The associated time course for the selected activation pattern is shown below the structural images. Comparing the plots from the detail browser to the query image, it can be seen that the returned result has significant activation in the user-selected region of interest.

3.2 Query 2

Starting with one subject's voxel correlations, find SPM-generated statistical activation images from that subject with similar spatial patterns of activation.
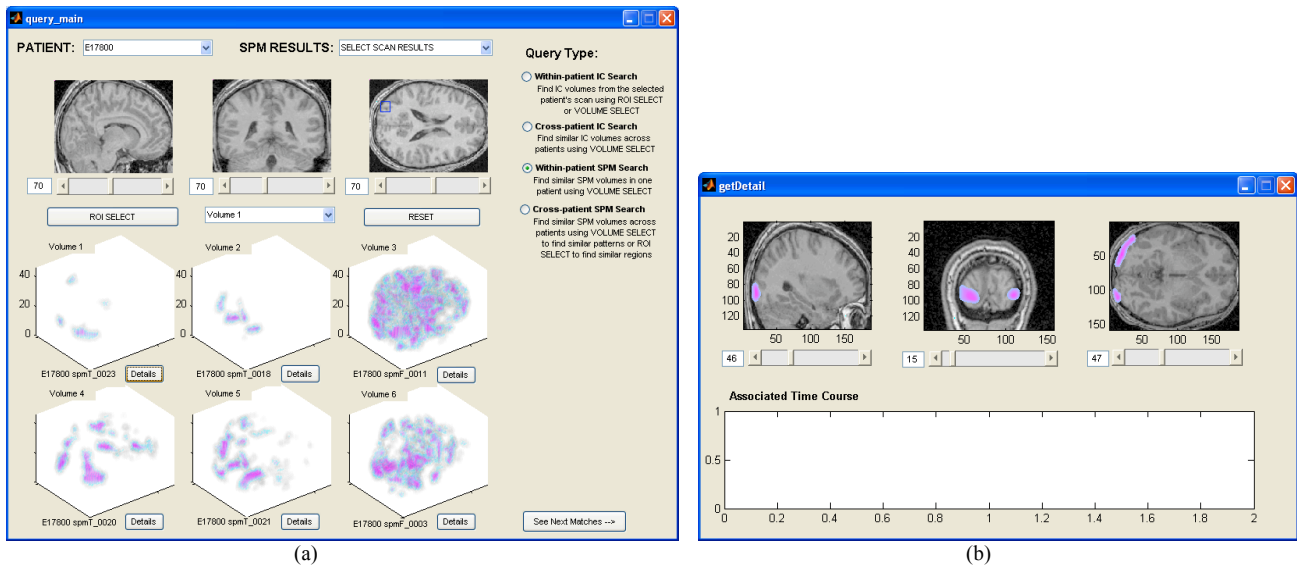
Figure 2.(a) The main user interface used to execute queries showing query type 2 and (b) the detail browser window showing the first match returned in the sample query.

This query allows the user to compare an ICA-generated activation image to SPM-generated activation images for that same subject. A query such as type 1 must first be done to find an interesting ICA-generated map, which will be used as the new query image. The user can then use the interface to query the database for SPM activation images for that subject that are similar to the ICA-generated result selected.

A sample query of type 2 is shown in Figure 2. This query was done using the results from the query of type 1 shown in Figure 1. The ICA-generated activation image Volume 1, shown in detail in Figure 1 (b) was selected and the database was searched for SPM-generated activation images similar to this pattern. The results are displayed below the structural data in Figure 2 (a). The best match from this query, Volume 1, was mapped back onto the subject's structural image, shown in Figure 1 (b). There is no associated time course for this activation pattern since these only exist for those generated by ICA. Comparing the detail browser plots of the SPM-generated pattern in Figure 2 (b) to the detail browser plots of the ICA-generated pattern shown in Figure 1 (b), similarities in the location and shape of the activations can be seen.

3.3 Query 3
Starting with one subject's voxel correlations, search for other subjects who have a similar correlation pattern for a voxel in the same region.
This query allows the user to compare an ICA-generated activation map to all other ICA-generated activation maps in the database. Before executing, a query such as type 1 must be done to find an ICA-generated activation map to use as the new query pattern. The user may then select an interesting map and search for similar ICA-generated maps across subjects. This can reveal whether the pattern is specific to the subject, or if there is evidence that other people may have had a similar activation during the scan. This information allows the user to explore the frequency with which this pattern occurs in the pool of subjects in the database, and look for other commonalities between the returned subjects.

A sample query of type 3 is shown in Figure 3. This query was done using the ICA-generated activation map shown in the detail browser in Figure 3 (a). The database was searched for ICA-generated activation images similar to this pattern across subjects. The results of this query are displayed below the structural data in Figure 3 (b). The best match from this query, Volume 1, was mapped back onto that subject's structural image, shown in Figure 3 (c). Comparing the detail browser plots in Figure 3 (a) to the detail browser plots in Figure 3 (c), it can be seen that the location and shape of the activations are highly similar.
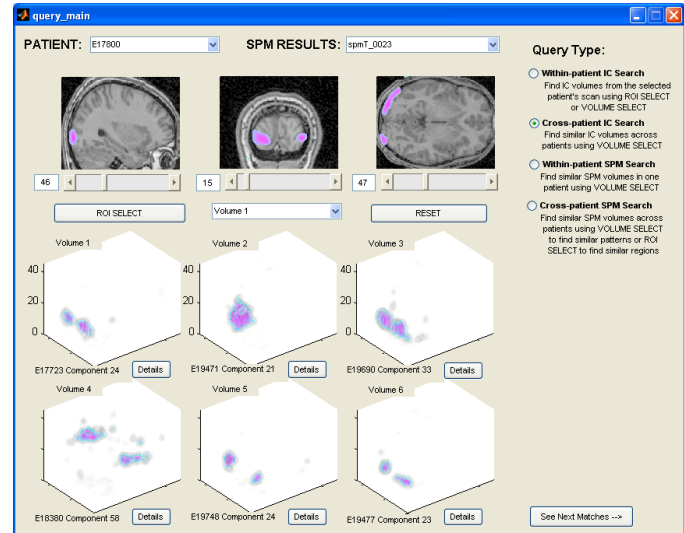
3.4 Query 4
For a given subject's statistical activation image and given location, find other subjects who have greater than or equal activation values at that location, by searching the SPM images showing statistically significant activations.

(a)

(b)

(c)
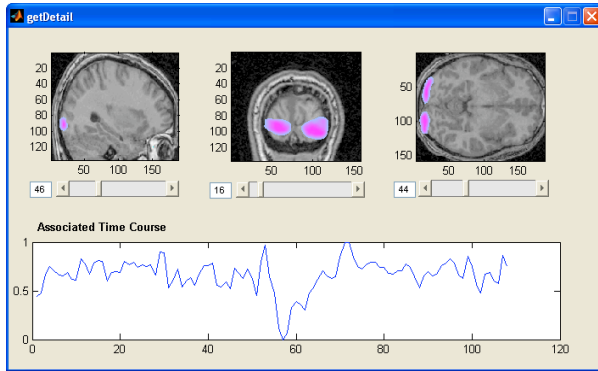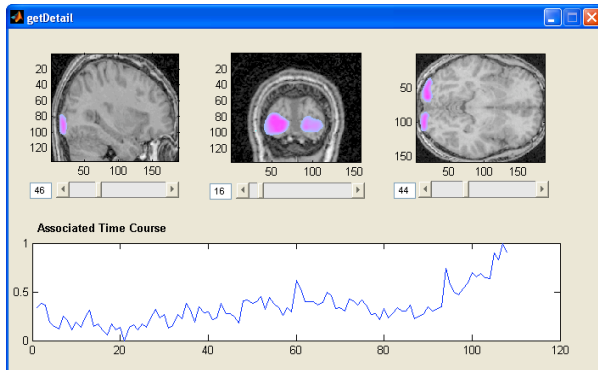
Figure 3. (a) The detail browser window showing the IC-generated activation volume used as a query image, (b) the main user interface used to execute queries showing query type 3 and (c) the detail browser window showing the first match returned in the sample query.

This query allows the user to look for SPM activation images from other subjects who have brain activations above a predefined threshold in the user-defined region of interest. The user selects this region by drawing a box on one of the three structural MRI viewers on the tool interface. This box outlines the length and width of a three-dimensional region of interest with a predefined depth of five samples, centered on the slice where the user has drawn the box. When the query is executed, SPM activation images from other subjects with a cluster of activated voxels in the region of interest are returned to the user.

A sample query of type 4 is shown in Figure 4. The selected subject's structural MRI is shown in the top row of the main user interface window in Figure 4 (a). The region of interest was chosen on the coronal view and can be seen outlined in blue. The database was then queried for SPM-generated activation patterns from the same subject with significant activations in this area. The results are displayed below the structural data in Figure 4 (a). Volume 1 in the main interface was mapped back to the subject's structural MRI, which is shown in Figure 4 (b). Comparing the plots from the detail browser to the query image, it can be seen that the returned result has significant activation in the user-selected region of interest.

3.5 Query 5
For a given subject's statistical activation image, find other subjects who have similar spatial patterns of activation, by searching the SPM images showing statistically significant activations.
This query takes a user selected SPM activation image and searches across SPM activation images from all subjects in the database to find similar activation patterns. To execute this query, an SPM activation of interest must be selected for the comparison. This can be done by loading an activation pattern for that subject using the SPM Results menu on the interface, or by first executing a query such as type 4 which returns relevant SPM-generated images.
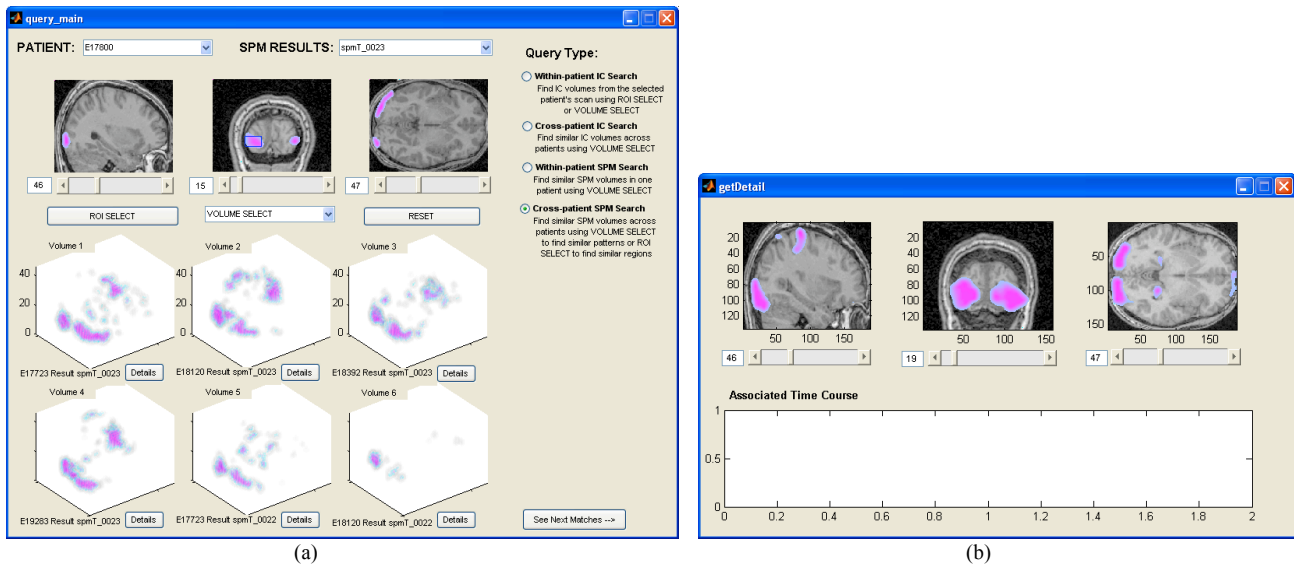
Figure 4.(a) The main user interface used to execute queries showing query type 4 and (b) the detail browser window showing the first match returned in the sample query.

A sample query of type 5 is shown in Figure 5. The query SPM activation pattern is shown mapped onto the subject's structural MRI, shown in the top row of the main user interface window in Figure 4 (a). The database was searched for similar SPM activation patterns from the other subjects. The results are displayed below the structural data in Figure 5 (a). Volume 1 in the main interface was mapped back to the subject's structural MRI, which is shown in Figure 5 (b). Comparing the plots from the detail browser to the query image, the similarity between the location and shape of the activation patterns can be seen.
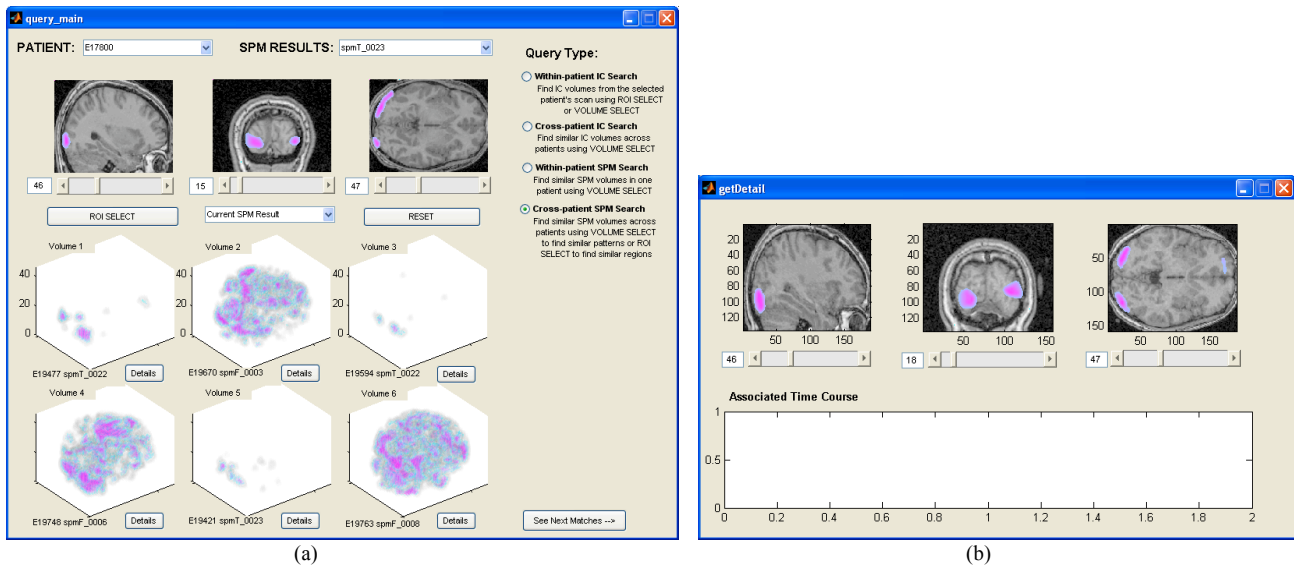

Figure 5.(a) The main user interface used to execute queries showing query type 5 and (b) the detail browser window showing the first match returned in the sample query.

3.6 Query 6
For a given subject's statistical activation image, find signal-similarity-generated correlation patterns from that subject with similar spatial patterns of activation.
This query allows the user to compare the SPM-generated activation images to ICA-generated activation maps for that same subject. To execute, an SPM activation of interest must be selected for the comparison. The user can then use the
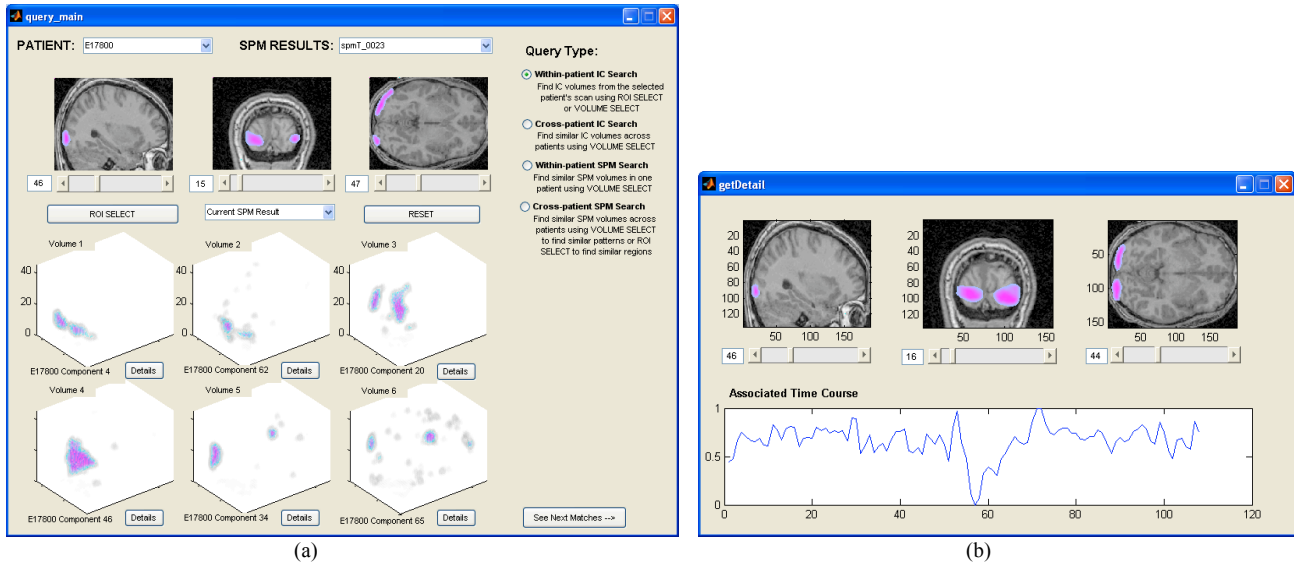
Figure 6.(a) The main user interface used to execute queries showing query type 6 and (b) the detail browser window showing the first match returned in the sample query.

interface to query the database for ICA generated images from that subject with activations similar to the selected SPM-generated image.

A sample query of type 6 is shown in Figure 6. This query was done using the SPM activation pattern shown mapped onto the subject's structural MRI in the top row of the user interface in Figure 6 (a). The database was searched for ICA-generated activation images similar to this pattern. The results are displayed below the structural data in Figure 6 (a). The best match from this query, Volume 1, was mapped back onto the subject's structural image, shown in Figure 6 (b). Comparing the plots from the detail browser in Figure 6 (b) to the query image shown in Figure 6 (a), it can be seen that the location and shape of the activations are highly similar.

4. DISCUSSION

SPM is a standard way to identify correlated voxels in an fMRI scan. However, unlike the IC-based method, SPM does not look at the relationships between voxels. Instead, it models the hypothesized response to the exam stimuli of the fMRI scan session, and each voxel is independently evaluated for a statistical relationship. This produces a map of the voxels that have a direct relation to the modeled stimuli response. The IC-based method of finding voxel correlations is expected to be much more powerful than the SPM method, since it is not limited by the use of a hypothesized response and can look for functionally connected regions that are indirectly dependent, or independent of the exam stimuli [1]. However, while the IC-based method should return more information, it should still be able to replicate the results found by SPM, since the SPM map indicates voxel regions that vary according to the exam stimuli, a statistically independent source. It is expected that for each SPM result map, there would be a corresponding IC map with a similar activation. It can be seen from the sample queries in Figures 2 and 6 that pairs of similar IC and SPM activation maps could be found for a single subject. In Figures 3, 4, and 5 it can be seen that IC and SPM activation maps similar to these pairs were found across subjects in the database.

5. CONCLUSIONS

This paper introduces a tool that applies independent component analysis as a measure of voxel similarity to allow the user to explore statistically independent maps of correlated voxel activity. A new algorithm for detecting and extracting characteristics from significant clusters of activated voxels and a three-dimensional spatial similarity measure allow comparison of the maps of correlated activity across subjects. Similar search capabilities are provided for spatial maps produced by the popular fMRI analysis tool SPM, so that the relationship between the IC maps and the results produced by SPM may be investigated.

6. REFERENCES

[1]   Beckmann, C.F., DeLuca, M., Devlin, J.T. and Smith, S.M., "Investigations into resting-state connectivity using independent component analysis", Philosophical Transactions of the Royal Society B: Biological Sciences 360(1457), 1001-1013(2005).

[2]   Bell, A.J. and Sejnowski, T.J., "An information-maximization approach to blind separation and blind deconvolution", Neural Computation 7(6), 1129-1159(1995).

[3]   Horwitz, B., Warner, B., Fitzer, J., Tagamets, M.A., Husain, F.T. and Long, T.W., "Investigating the neural basis for functional and effective connectivity. Application to fMRI", Philosophical Transactions of the Royal Society B: Biological Sciences 360(1457), 1093-1108(2005).

[4]   Lai, C., Tax, D., Duin, R., Pekalska, E., Paclik, P., "On combining one-class classifiers for image database retrieval", Lecture Notes on Computer Science 2364, 212-221(2002).

[5]   McKeown, M., Makeig, S., Brown, G., Jung, T., Kindermann, S., Bell, A., Sejnowski, T., "Analysis of fMRI data by blind separation into independent spatial components", Human Brain Mapping 6, 160-188(1998).

[6]   Poliakov, A.V., Hinshaw, K.P., Rosse, C. and Brinkley, J.F., "Integration and visualization of multimodal brain data for language mapping", Proceedings of American Medical Informatics Association Fall Symposium 349-353(1999).