Spatial Alignment of Functional Regions in fMRI

by

Gabriel Andres Tobón

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Engineering in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2011

Author .....
Department of Electrical Engineering and Computer Science
May 20, 2011

Certified by .....
Polina Golland
Associate Professor
Thesis Supervisor

Accepted by .....
Christopher J. Terman
Chairman, Masters of Engineering Thesis Committee

Spatial Alignment of Functional Regions in fMRI

by

Gabriel Andres Tobón

Submitted to the Department of Electrical Engineering and Computer Science
on May 20, 2011, in partial fulfillment of the
requirements for the degree of
Master of Engineering in Electrical Engineering and Computer Science

Abstract

An essential step for discovering a common structure in brain activation regions from multi-subject fMRI data is the ability to find spatial correspondences across subjects. This has proven to be a challenging problem due to the lack of a ground truth and variability in anatomical brain structure, functional activation, and spatial locations of functional regions. Standard methods rely on the correspondences given by anatomical registration to a common space, but fail to account for spatial variability of the functional regions relative to anatomy. We develop a clustering method that relies on the alignment of both the anatomical structure and the functional landmarks. The method is shown to improve over standard group analysis techniques that rely on anatomical alignment only. The validation of our method confirms that peaks of activation exhibit consistent spatial structure. Furthermore, our work creates a framework for future testing of different metrics for similarity of brain activation regions across subjects.

Thesis Supervisor: Polina Golland
Title: Associate Professor

Acknowledgments

To my advisor, Polina Golland, thank you for your invaluable support throughout my project. I appreciate the many times you challenged me as a student and researcher, as well as all the opportunities and advice you gave me to succeed. To my collaborators, Nancy Kanwisher and Evelina Fedorenko, thank you for your always timely help and feedback, even when you were busy with your own deadlines and commitments. Also, thank you for your patience as I learned the relevant Neuroscience concepts. To my fellow members of Medical Vision lab, especially George Chen and Danial Lashkari, who frequently helped me with my project, thank you for making me feel welcome and helping me whenever I needed it.

To Berthold Horn, who had me as a teaching assistant and student for his Machine Vision course, thank you for your advice and support both as a teaching assistant and in my research during my fall semester. To my supervisors and collaborators on other research projects and internships throughout my time at MIT, especially Kelly Davis Orcutt, Rafiou Oketokoun, and Andrew Rothbart, thank you for helping me develop into the researcher that I am now.

To Mom, Dad, Ana, and the rest of my incredible family, thank you for all of your unconditional love and support that has helped me through the ups and downs of MIT. I cannot thank you all enough. To my apartment mates this past year, Jess Kim, Dustin Kendrick, and David Wen, thank you for putting up with my living habits and, of course, for being my friends. To all of my friends, especially Angela Yen, Mingwei Gu, Kent Willis, Lindsay Willis, Dember Giraldez, Aubrey Tatarowicz, Neel Hajare, and Josh Wilemon, thank you for always being there for me and giving me so many great memories to carry with me.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The understanding of functional specificity in the human brain has been one of the main goals of Neuroscience since its inception. Functional Magnetic Resonance Imaging (fMRI) provides indirect measurements of brain activity which can be used for discovering functional specificity. The high dimensionality and high level of noise in fMRI data present a number of interesting challenges that require a combination of neuroscientific experimental design and sophisticated computational techniques for analysis.

1.1 Discovering Common Functional Specificity

Inference of spatial locations in the brain that are highly correlated with experimental conditions from fMRI data is tremendously useful for answering questions about the human brain. For example, by presenting conditions such as pictures of faces and pictures of objects to subjects while they are being scanned, the inference procedure identifies regions of the brain responsible for visual face recognition. It can also be used for computer aided diagnoses in a medical setting. The focus of this work is to improve such inference. Methods for determining spatial correlations with the experimental conditions in one subject's brain have been widely demonstrated in the literature. Assuming this subject level analysis has been done, the next step is to compare subjects and make inferences about function that hold for the population in

general.

1.2 Problem of Spatial Correspondences

Before a set of subjects' fMRI data can be analyzed together, a basis for comparison must be established. Anatomical characteristics of the brain, spatial location of functional activation, and the presence of functional activation all vary across individuals; this variability must be accounted for in a group level analysis. The standard approach in the literature assumes that spatial variability of functional regions is small enough to be discounted. Higher resolution anatomical scans of the subjects are used to align the subjects' brains to a common space. The transformations used for this alignment are then applied to the functional data to establish spatial correspondences across subjects. Although this assumption has been effective for a variety of different experiments, anatomy is not always a good predictor of function. For some functionally specific regions of interest, the locations of active regions can vary as much as the size of those regions. This is the case for language processing in the brain, motivating our work presented in this thesis. Under such circumstances, the standard methods may completely fail to detect regions of robust activation in the brain simply because they do not overlap anatomically.

1.3 A Landmark Based Approach For Functional Alignment

We develop an improvement over the standard method for group analysis by using functional information for alignment in addition to the anatomical registration. Our landmark-based approach reduces the inherently high dimensional fMRI data to a manageable level. Peaks of functional activation have been suggested as landmarks that are robust across subjects [13]. The proposed method successfully focuses on peaks as landmarks for alignment. Since peaks of activation also represent main points of interest for an fMRI study, we choose to ignore other regions of subjects' brains for

the purposes of matching. Instead, spatial correspondences of the peaks are found, and these correspondences are validated by demonstrating that the spatial locations of the peaks are better predictors of function than anatomy alone in the case of language areas in the brain.

1.4 Thesis Outline

Chapter 2 provides background useful for understanding the problem and content of this thesis. It describes fMRI, Random Effects Analysis, and current methods of fMRI group analysis. Chapter 3 describes the method of finding spatial correspondences of functional regions developed in this thesis. Chapter 4 provides details of the relevant fMRI experiments and the pre-processing steps used to prepare the raw fMRI data for analysis. It also presents the method used for validation of the spatial correspondences found by the method in Chapter 3. Chapter 5 summarizes the main results from the methods described in Chapters 3 and 4. Chapter 6 provides a discussion of these results and presents possibilities for future work.

Chapter 2

Background

2.1 Functional Magnetic Resonance Imaging

fMRI is used to indirectly access changes in neural activity in the brain. It achieves this by measuring a blood oxygen level dependent (BOLD) signal which is known to be correlated with neural activity. Specifically, the BOLD signal is the magnetic resonance contrast for deoxygenated hemoglobin, which is found in greater concentrations at sites of increased neural activity. It is a non-invasive imaging technique, widely used in studies of brain activity. The raw data for our work include magnetic resonance scans of BOLD signal over time, where each scan is a 3-dimensional image.

This technology affords experiments that study variation in BOLD signal with respect to different stimuli. Controlled experiments can be designed to shed light on the neural basis of different cognitive functions. In the experimental validation, we focus on language comprehension.

2.2 Functional Localization

fMRI localization studies seek function-specific regions of the brain that appear robustly across subjects. Such regions can be targeted through the use of contrasts. The underlying assumption of this method is that the response to a particular stimulus is expected to be a superset of the functional response of interest. In order to isolate

regions of interest, the non-specific estimated response is removed using other stimuli. In the case of language, for example, the response to a collection of non-words is subtracted from the response to sentences [5]. In the case of face selectivity, the response to objects is often subtracted from the response to faces [8].

The detailed mapping of the language system is still a topic of active research. Broca's area, for example, was discovered by Pierre Paul Broca in two subjects with lesions in the left inferior frontal gyrus of their brains and the corresponding speech impairment [3]. Since this discovery, decades of studies on patients with similar lesions have supported the specificity of this region for language. Broca's area has also been implicated in music and action representation [4] using fMRI. Futhermore, consistent results are lacking for the localization of semantic, syntactic, and phonological processing [5].

One possible explanation for difficulty in language localization is the deficiency in the methods used for group analysis in the presence of spatial variability. This spatial variability can be difficult to evaluate, but it is suspected to be on the order of centimeters [12], such as in the case of verbal working memory [16]. The remainder of this chapter reviews the standard approach of Random Effects Analysis as well as more recent methods that motivated our work on improving spatial alignment of function.

2.3 Group Analysis of fMRI Data

Group analysis of fMRI aims to extract a common activation pattern from the fMRI data in multiple subjects who participated in the study. A number of challenges arise during inference of localized brain function in a population. We focus on the challenge of identifying spatial correspondences. In particular, it is important to know whether a functional region in one subject corresponds to a functional region in another subject. This problem is related to the well studied task of registration in anatomical MRI, where two images are aligned with each other. In contrast, fMRI images are acquired over time and depend on the experiment.

This section describes current approaches to group analysis of fMRI data and reviews how they approach the challenge of spatial correspondences. We present the standard method of Random Effects Analysis and more recently developed methods. From these, we derive a set of desirable method characteristics.

2.3.1 Random Effects Analysis

Random Effects (RFX) models are widely used in fMRI group analysis to make inferences about a population from a set of subjects. The Random Effects model is a two-level hierarchical model that captures inter-subject variability at the group-level and intra-subject variability for each individual subject. The model assumes that the pre-whitened data $\mathbf{y}$ can be thought of as a linear combination of subject-level effects from the design matrix X and noise η:

$$\mathbf{y}_i = X\beta_i + \eta_s \qquad \eta_\mathbf{s} \sim N(\mathbf{0}, \sigma_s^2 I) \tag{2.1}$$

where i is a subject index. $N(x; \mu, \sigma^2)$ represents a Gaussian probability distribution for a random variable x with mean μ and variance σ^2. Columns of matrix X are sometimes called explanatory variables, conditions, and predictors in the literature because they function as a low dimensional characterization of the data. In fMRI, the effects of interest are the experiment stimuli. At the group-level, the coefficients β of the subject effects are assumed to be a noisy realization of population ground truth coefficients μ:

$$\beta_i = \mu + \eta_g \qquad \eta_\mathbf{g} \sim N(\mathbf{0}, \sigma_g^2 I) \tag{2.2}$$

This model also assumes that the data has been temporally filtered to ensure that the subject-level noise is indeed white. The model is fit to every voxel independently. This step requires voxel-wise correspondences across subjects, a important assumption discussed in detail in the following section. The vector $\mathbf{y}_i$ contains all elements of the time course for the voxel of interest. The columns of matrix X correspond to different effects and each row indicates which effects, or in our case stimuli, were present at a

particular time point. By substituting (2.2) into (2.1), we obtain the likelihood of the data:

$$\begin{aligned} \mathbf{y}_i &= X\mu + X\eta_g + \eta_s \qquad \mathbf{y}_i \sim N(X\mu, \Lambda), \\ \Lambda &= \sigma_g^2 \mathrm{diag}(XX^T) + \sigma_s^2 I \end{aligned}$$

The maximum likelihood (ML) estimator of the group mean μ and its variance are needed to compute the t-statistics for the effects. We can obtain the ML estimates analytically. Formally,

$$\begin{aligned} p(\mathbf{y}; \mu, \sigma_s^2, \sigma_g^2) &= \prod_{i=1}^{N_s} p(\mathbf{y}_i; \mu, \sigma_s^2, \sigma_g^2) \\ &= \prod_{i=1}^{N_s} \frac{1}{(2\pi)^{\frac{N_s N_t}{2}} |\Lambda|^{\frac{1}{2}}} \exp(-\frac{1}{2}(\mathbf{y_i} - X\mu)^T \Lambda^{-1} (\mathbf{y_i} - X\mu)), \\ \frac{\partial \ln p(\mathbf{y}; \mu, \sigma_s^2, \sigma_g^2)}{\partial \mu} &= \sum_{i=1}^{N_s} (\mathbf{y}_i - X\mu)^T \Lambda^{-1} X = \mathbf{0}, \\ \hat{\mu} &= \frac{1}{N_s} (X^T \Lambda^{-1} X)^{-1} X^T \Lambda^{-1} \sum_{i=1}^{N_s} \mathbf{y}_i. \end{aligned}$$

We observe that this estimator is equal to the average of subject-level estimates of β_i:

$$\hat{\beta}_i = (X^T \Lambda^{-1} X)^{-1} X^T \Lambda^{-1} \mathbf{y}_i$$

Full estimation of the parameters μ, σ_s^2, and σ_g^2 requires an iterative approach using parametric empirical Bayes [10] or EM [2]. For more information about such estimation, see [17]. In practice, an approximation is used that is consistent with iterative inference. In this approach, the subject-level model and group-level model are solved separately. Namely, a set of summary statistics from the subject-level estimates is used to derive group-level estimates. Before we setup this model, we formalize the estimation at the subject-level model.

To generalize the subject-level model to temporally correlated fMRI data, an autoregressive model is used. The subject level noise is no longer white, but is assumed

to be zero mean and have a covariance structure V that is shared across voxels. A covariance estimate $\hat{V}$ is derived and used to construct a filtering matrix $F = \hat{V}^{-\frac{1}{2}}$. Thus, the subject level model becomes $F\mathbf{y}_i = FX\beta_i + F\eta_i$ where $F\eta_i \sim N(\mathbf{0}, \sigma_i^2, I)$. With this pre-whitened model, we can once again derive the ML estimates for β_i and σ_i^2.

$$\hat{\beta}_i = (X^T F^T F X)^{-1} X^T F^T F \mathbf{y}_i \tag{2.3}$$

$$\hat{\sigma}_i^2 = \frac{\epsilon_i^T \epsilon_i}{\mathrm{tr}(R)} \tag{2.4}$$

where $\epsilon_i = (I - FX(X^T F^T F X)^{-1} X^T F^T F)\mathbf{y}_i = R\mathbf{y}_i$ is the residual error.

To isolate the effect of language in this model, a contrast is defined on the experimental conditions, such as sentences minus nonwords. Such a contrast can be defined using a multiplier $\mathbf{c}$ (e.g, 1 for sentences, -1 for nonwords, and zero for everything else) on the coefficients β_i. From these estimates, $\mathbf{c}^T\hat{\beta}_i$ and $\mathrm{Var}(\mathbf{c}^T\hat{\beta}_i)$ are treated as summary statistics to solve the group-level model. The new two level model follows:

$$\begin{aligned} \mathbf{c}^T\hat{\beta}_i &= \mathbf{c}^T\beta_i + \eta_i \qquad \eta_i \sim N(0, \sigma_i^2) \\ \mathbf{c}^T\beta_i &= \mu + \gamma \qquad\quad \gamma \sim N(0, \delta^2) \end{aligned}$$

In the RFX case, the individual subject variance is ignored. If $i \in \{1, ..., S\}$, then the following RFX t-statistic is constructed for each voxel:

$$t_{S-1} = \sqrt{S}\frac{\hat{\mu}}{\sqrt{\hat{\delta}^2}},$$

where $\hat{\mu} = \frac{1}{S}\sum_{i=1}^{S} \mathbf{c}^T\hat{\beta}_i$ and $\hat{\delta}^2 = \frac{1}{S}\sum_{i=1}^{S}(\mathbf{c}^T\hat{\beta}_i - \hat{\mu})^2$. μ is the true group effect for $\mathbf{c}^T\beta_i$, which is in turn the true subject effect for contrast $\mathbf{c}$ in subject i. η_i and γ are subject and group-level noise, respectively.

We also investigate the use of an approximation to Mixed Effects Analysis (MFX) The corresponding statistic is referred to as a pseudo-MFX statistic by [14] and was

found to have the best reproducibility of recent methods of group analysis. It will be referred to as Weighted Random Effects Analysis (WRFX) in this thesis. Its results, however, are not as easily interpretable as those of RFX or MFX. The two-level model for this weighted least squares setup and its t-statistic are:

$$\begin{aligned} \mathbf{c}^T\hat{\beta}_{sv} &= w_{sv}\mathbf{c}^T\beta_{sv} + \eta_{sv} \quad \eta_{sv} \sim N(0, \sigma_{sv}^2), \\ \Psi &= \sqrt{S}\frac{\hat{\mu}_v}{\sqrt{Var(\hat{\mu}_v)}}, \end{aligned}$$

where $\mathbf{c}^T\beta_i = \mu$, $w_i = \frac{Var(\mathbf{c}^T\hat{\beta}_i)}{\sum_{i=1}^{S} Var(\mathbf{c}^T\hat{\beta}_i)}$, $\hat{\mu} = \sum_{i=1}^{S} \frac{1}{w_i}\mathbf{c}^T\hat{\beta}_i$, and $Var(\hat{\mu}) = \frac{1}{S}\sum_{i=1}^{S}(\hat{\mu} - \mathbf{c}^T\hat{\beta}_i)^2$. Ψ is assumed to have a Student's t distribution with $S-1$ degrees of freedom under the null hypothesis.

2.3.2 Recent Alternatives to Voxel-wise Correspondences

As suggested in the previous section, a significant assumption of RFX is voxel-wise correspondences of data across subjects. These correspondences are typically derived by registering an anatomical image of each subject to a common space and using the same transformation to map all fMRI data to that common space. Finding spatial correspondences of functional regions faces a number of challenges, some of which are described in [12]. To summarize,

1. Anatomical registration of a subject to a template is likely not perfect, and it is unclear if a perfect registration even exists.

2. Even if the anatomical registration were perfect, there is no guarantee that the same transformation exists for fMRI.

To increase the robustness of Random Effects Analysis, the spatial resolution of fMRI data is often sacrificed by smoothing the images with large Gaussian smoothing kernels (10mm FWHM or more), meant to increase functional overlap. Ideally, however, the true alignment of functional regions could be found across subjects. One approach is to align the subject-level estimated activations.

In [12], an improvement of classical Random Effects Analysis is demonstrated, based on a combination of functional and anatomical information. The subjects' β coefficients are pooled and clustered under the spatial constraint that members of the same cluster must be close to each other in the common anatomical space. An embedding is then defined for each subject's voxel β coefficients containing a mixture of functional and anatomical distances. Each voxel is assigned a cluster ID based on distances in the embedded space. This enforces a matching property where at least one voxel from each subject is represented in each cluster. The results suggest that using functional information can improve sensitivity of functional localization. However, it is not clear what features other than functional homogeneity can improve the sensitivity of the analysis. Furthermore, it might be better to relax the constraint that every subject has a particular activation region.

In [13], peaks of activation are used to determine spatial correspondences of functional regions. The peaks are matched based on relative distance to other peaks in the same subject. Promising results suggest that peaks of activation are good landmarks for activation regions. Validation was done on synthetic data and on real fMRI scans for reproducibility across subjects. Validation on real fMRI data is desirable, but reproducibility alone does not necessarily guarantee improved alignment. We employ larger data sets in individual subjects that allow for splitting of the fMRI time courses such that alignment can be derived on one half of the data and applied to the other half.

In [19], a generative model for functional profile is designed to fully explain various layers of variance in a population. It develops a 4-level Bayesian hierarchical model. At the lowest level, individual functional activation maps are allowed to vary at each voxel. Second, individual component means are allowed to vary in each subject with Gaussian mixtures serving as priors for these means. This second level is most similar to the peaks of activation used in our work. Third, the centers of the Gaussian mixture are allowed to vary around activation region centers. The final level models the distribution of individual centers around population centers. If this model could be estimated perfectly and truly represented the underlying structure, it would be

crucial for understanding the hierarchy of signal variability in fMRI. Unfortunately, the complexity of the model makes it difficult to estimate so many parameters together and still achieve a global minimum. The method developed in this work takes a simpler approach that isolates one aspect of spatial variability of functional regions.

Further work has been done to use more sophisticated registration algorithms to enhance anatomical alignment using fMRI time courses [11]. This approach has been shown to improve over standard RFX analysis. Sophisticated registration algorithms are likely to outperform our method due to the large space of nonlinear warps it can discover. Using such registration algorithms, however, makes it difficult to understand the underlying spatial variability of functional regions.

Chapter 3

Spatial Correspondences of Functional Regions

We focus on developing a method for identifying spatial correspondences of functional regions. As discussed in the previous chapter, a good method of alignment will have certain properties. First, resulting activation regions given by alignment should be more **sensitive**, where sensitivity refers to method's level of detection of robust regions of activation. Second, the method should also introduce a minimal number of arbitrary constraints. For example, every subject should not need to have a particular functional region. Third, it is reasonable to expect that aligned peaks will be close to each other spatially in a common space. Finally, the method should reveal some inherent variability and structure of the data.

In this work, we use the locations of activation peaks within each region as landmarks to align activation regions.

3.1 Region Matching Algorithm

To improve on the standard method of alignment that is limited to anatomical information, we develop a method that takes advantage of both anatomical and functional information. Our algorithm works by clustering peaks of activation spatially across subjects. The clusters themselves represent an answer to the challenging question

of finding spatial correspondences across subjects.

In subsequent sections we let p_{sr} be the peak r in subject s. p_{sr} has coordinates $C(p_{sr})$ in the MNI152 space. It is a valid peak if its significance value is greater than the significance value at all voxels within a surrounding sphere with a 2mm radius. The significance values are assumed to be derived from individual level analysis of the contrast of interest. The matching can be completely described by a function $M(\cdot,\cdot)$ where $M(p_{ab}, p_{cd}) = 1$ iff p_{ab} and p_{cd} are matched and is zero otherwise.

3.1.1 Identifying Pairwise Correspondences

Before we begin to identify spatial correspondences between two subjects, we motivate the algorithm by presenting our assumptions about the spatial patterns of activation. The first assumption is that the activation peaks represent robust functional constructs. In other words, given a cluster of peaks across subjects, if a new subject were introduced and it had a peak of activation spatially close to the cluster in the common space, it identifies the "same" peak of activation as those in the cluster. The second, and perhaps the strongest assumption made by out method, is a one-to-one mapping of peaks across subjects. In other words, in any given matching, a peak in subject A is either present once or not present at all in subject B. The method does not allow the mapping of two peaks in subject B to one peak in subject A.

Once all peaks have been found in a pair of subjects, a cost function $D(\cdot,\cdot)$ quantifies the quality of a match between a peak in subject A and a peak in subject B.

$$D(p_{ab}, p_{cd}) = \begin{cases} \text{dist}^2(p_{ab}, p_{cd}) & : \text{dist}^2(p_{ab}, p_{cd}) \leq \gamma \\ k & : \text{dist}^2(p_{ab}, p_{cd}) > \gamma \end{cases} \tag{3.1}$$

where $\text{dist}^2(p_{ab}, p_{cd}) = ||C(p_{ab}) - C(p_{cd})||^2$. Parameter γ defines the furthest distance between two peaks allowable if they are matched. We set γ to 10mm, which is approximately equal to the diameter of language activation regions seen in the thresholded activation map. Parameter k is a large "infinity-like" value which is many times larger than any distance squared possible in the MNI152 space. Since the number of peaks

in subject A and subject B can differ, we add fictional peaks to the subject with the fewer number of peaks at k distance from all other peaks. Thus, when a matching with k or greater distance occurs, it is ignored.

The desired pairwise matching minimizes the total cost of peak matches between two subjects s_a and s_c:

$$T(s_a, s_c) = \sum_{i,j \text{ s.t. } M(p_{ai}, p_{cj})=1} D(p_{ai}, p_{cj}). \tag{3.2}$$

This is achieved using the Hungarian algorithm [9], which we review in Appendix A. Using recent implementations, the minimum of the total cost can be found in $O(n^3)$ where n is the maximum number of peaks in a subject. This is very reasonable considering the number of peaks that typically exist for each subject is about 100.

3.2 Deriving Group Correspondences

One method of deriving a group-wise clustering using an engine that computes pairwise matches is to select one subject to serve as a template. If subject s_a is used as a template, matchings $M_{b \to a}$ can be found for all other subjects to the template. Thus, for any peak, p_{ar}, in subject s_a, a cluster label can be assigned to all peaks from other subjects matched to p_{ar}. Such a cluster is fully defined by $\mathbf{m}_{ar}$, a vector indexed the same way subjects are indexed via $s \in \{1, ..., S\}$. Element i of vector $\mathbf{m}_{ar}$ contains the index of the peak in subject i matched to peak p_{ar}. If no such match exists, element i is assigned to a fictitious index, such as -1.

The above method for deriving a group-wise clustering is strongly biased towards the peak configuration of the template. Furthermore, it is possible that the subject chosen as a template lacks some peaks that might be robustly present in other subjects. To avoid such a bias, we construct clusters using every subject as a template, creating a vector $\mathbf{m}_{sr}$ for every subject s and every peak r in that subject. Since the most robust peaks across subjects are desired, we expect that the best clusters are robust to the choice of a template. We define function $H(\cdot, \cdot)$ that provides a similarity score

between two cluster assignments $\mathbf{m}_{ar}$ and $\mathbf{m}_{cq}$. We define $H(\mathbf{m}_{ar}, \mathbf{m}_{cq})$ as the number of identical elements between $\mathbf{m}_{ar}$ and $\mathbf{m}_{cq}$.

To assign each cluster a score that measures how robustly it appears using different subjects as the template, we iterate over all subjects and define the overall score, $S(a, r)$.

$$S(a,r) = \sum_{s=1}^{S} \max_{j} H(\mathbf{m}_{ar}, \mathbf{m}_{sj}) \tag{3.3}$$

This procedure computes the total number of elements equal to $\mathbf{m}_{ar}$ for the best matching cluster from each subject. This score is higher when the cluster is found in more templates and when the cluster itself has peaks from more subjects.

We rank the clusters and select the best clusters with at least one quarter of the subjects represented. Since similar clusters with good scores can be derived from multiple subjects as templates, only the best of the redundant set of clusters is chosen. Two clusters are considered redundant if they share peaks. The resulting clusters were the robust, unique clusters. These clusters come from a variety of subject templates.

The algorithm produces clusters using both functional information from the peaks of activation in the individual-level analysis and anatomical information from spatial proximity in the MNI space after registration.

Chapter 4

Pre-processing and Validation

Before deriving the activation maps and peaks, we process data from each subject using standard fMRI techniques for denoising and intra-subject image alignment. Furthermore, we developed a method of validating our inter-subject peak matchings that does not rely on availability of ground truth.

4.1 fMRI Pre-processing Pipeline

The raw fMRI time courses were pre-processed using a combination of Freesurfer [7] and FSL [6] software tools. A flow chart of the pipeline is provided in Figure 4-1. Motion correction of the fMRI time courses for each subject was performed to ensure that functional scans in each subject are properly aligned with each other. Volumetric smoothing with a 5mm FWHM Gaussian kernel was used to denoise the data. The motion correction, smoothing, and intensity normalization across time points were performed using Freesurfer's FsFast package [7]. Registration of functional data to anatomical data is done using Freesurfer's boundary-based registration. Affine registration of the subject's anatomical to MNI152 space is done using FSL's FLIRT algorithm, and a nonlinear warp to improve that registration is found using FSL's FNIRT algorithm. The transformations derived during registration steps are used to map all summary statistics from the individual-level GLM into a common space for all subjects.

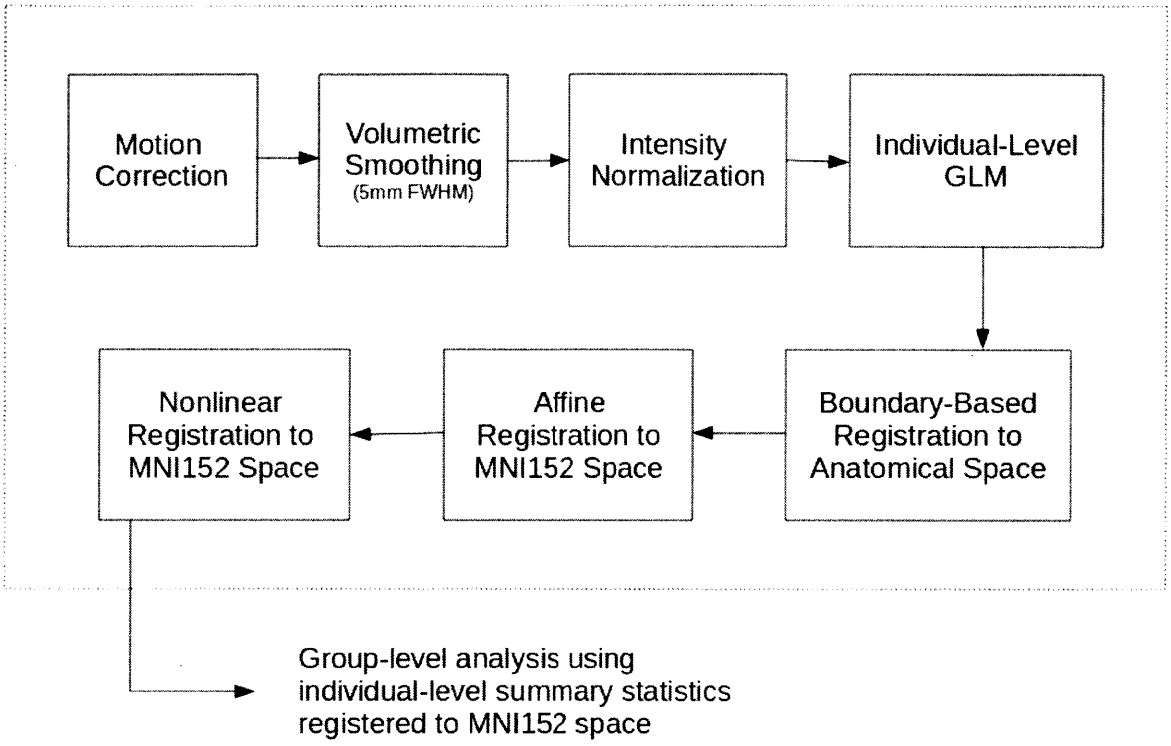


Figure 4-1: Pipeline for pre-processing fMRI data. The dotted line encloses the pre-processing steps done using Freesurfer [7] and FSL [6] tools on the fMRI data.

Based on the estimated regression coefficients β_i, we construct a t-statistic as follows.

$$t_\nu = \frac{\mathbf{c}^T \hat{\beta}_i}{\sqrt{Var(\mathbf{c}^T \hat{\beta}_i)}},$$

Where X is the design matrix, F is the whitening matrix, and $Var(\mathbf{c}^T \hat{\beta}_i) = \hat{\sigma}_i^2 \mathbf{c}^T (X^T F^T F X)^{-1} \mathbf{c}$. We estimate the effective degrees of freedom ν using the Satterthwaite approximation [18]:

$$\nu \approx \frac{(trace(R))^2}{trace(R^2)}.$$

Under the null hypothesis, the contrast effect t-statistic is distributed according to a Student's t distribution with ν degrees of freedom. From the t-statistic at every

voxel, p-values and significance values ($-\log_{10}p$) were derived and used as "data" for the matching algorithm. Figure 4-2 shows an example of a thresholded significance map generated from this statistical test.

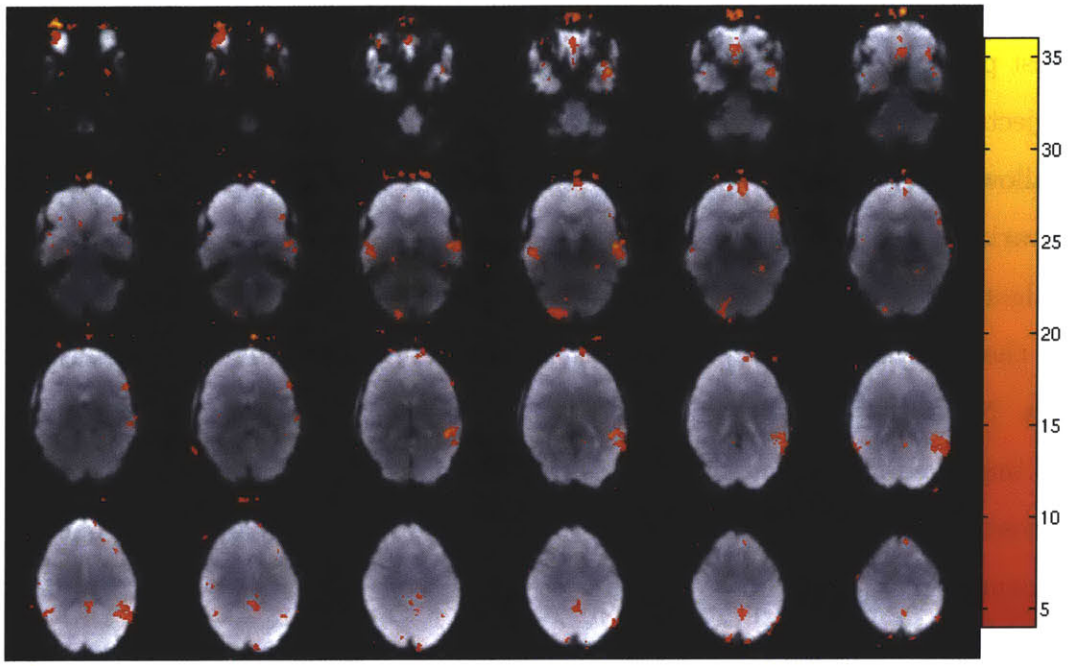


Figure 4-2: Example statistical significance map for the sentences minus nonwords contrast in one subject presented as axial slices and thresholded at $p = 10^{-4}$.

The contrast maps $\mathbf{c}^T\hat{\beta}_i$ and the residuals $Var(\mathbf{c}^T\hat{\beta}_i)$ were stored for each subject as summary statistics to be used for group analysis.

4.2 Validation of Clusters

To validate that peak clusters derived using our matching algorithm indeed provide reasonable spatial correspondences, we developed a method for validation that compares our method to the standard RFX analysis. If a cluster of peaks does capture spatial correspondences, we hypothesize that local alignment of these peaks should perform better than alignment solely based on anatomical information. To test this hypothesis, we realign the functional summary statistics using the estimated matching.

First, a sphere of radius 5mm was constructed around each peak in each cluster.

Then, for any given cluster, the spheres were aligned together around the mean peak location for that cluster. In other words, the translation which would move a peak to the mean peak location of its cluster is used to translate every voxel around that peak within the surrounding sphere. Since not all subjects were represented in the cluster, the nearest peak to the mean from the unthresholded significance maps were used to align subjects not present in the cluster. This enforces the "onto property" described by [12], allowing for valid statistics to be computed at each voxel and not relying on uniform variance of the summary statistics across voxels in the sphere.

The algorithm estimates clusters from the complete data set using the significance maps for the contrast of interest. The clusters are then used to align voxels around the peaks. Note that this does not necessarily validate the proposed algorithm as a general improved alignment, but it does indicate whether alignment based on the localizer contrast can improve the alignment of the corresponding areas that are marked as significant. To compare our results to the results of using standard anatomical alignment, the significance values from voxels of the average spherical neighborhood for each cluster are estimated via Random Effects Analysis.

Chapter 5

Results

The first section of this Chapter provides details about the fMRI experiments used for this work. It is followed by the results of the preliminary analysis that motivate the need for better cross-subject alignment techniques for our data. The next section presents the results clustering from our matching algorithm described in Chapter 3. Without a reliable ground truth for these correspondences, the derived clusters were evaluated with the indirect validation procedure desribed in Chapter 4. The last section presents the results of this procedure applied to the derived clusters.

5.1 Language Experiments

5.1.1 Experimental Details

The language fMRI data set was collected with a Tesla Siemens Trio scanner at the Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research. The BOLD data was collected with a spatial resolution of $3.1 \times 3.1 \times 4$mm voxels and a temporal resolution of TR = $2,000$ms and TE = 30ms. The first 4s of each run were excluded to ensure steady state magnetization. T1-weighted anatomical images were also collected in 128 axial slices with 1.33mm isotropic voxels and TR = $2,000$ms and TE = 30ms. 34 right-handed subjects between the ages of 18 and 30 were scanned. 23 of them were female. 672 time points of fMRI scans were available from each subject.

In the first experiment, all 672 time points were used to produce clusters. In the second experiment, only the first half, or 336, of the time points were used to derive clusters.

The fMRI contrast used in this work to localize regions of interest was sentences minus nonwords in auditory experiments. It is commonly used to localize language comprehension, which is known to be highly variable spatially across subjects [5]. The language experiments included a sentences condition, a words condition, and a nonwords condition in a blocked design. The experimental design is described in [5]. The sentences condition is designed to engage lexical and structural processing. An example of the sentences condition would be "THE DOG CHASED THE CAT ALL DAY LONG." The words condition is a scrambled set of words with no sentence structure, but still requiring lexical processing. The nonwords condition is not meant to engage lexical nor structural processing. An example of the nonwords condition would be "CRON DACTOR DID MAMP FAMBED BLALK THE MALVITE."

From these experiments, the language localizer contrast was constructed by subtracting the nonwords response from the sentences response. This contrast was used by our matching algorithm to identify spatial correspondences across subjects.

5.2 Preliminary Analysis of Spatial Variability of Functional Regions

While no ground truth exists for the alignment of functional regions, some preliminary analysis can still be done to evaluate the spatial variability of functional regions across subjects. In theory, anatomical variation can fully account for the spatial variability of function, and one would expect that improving the registration algorithm would significantly improve the sensitivity of the Random Effects Analysis.

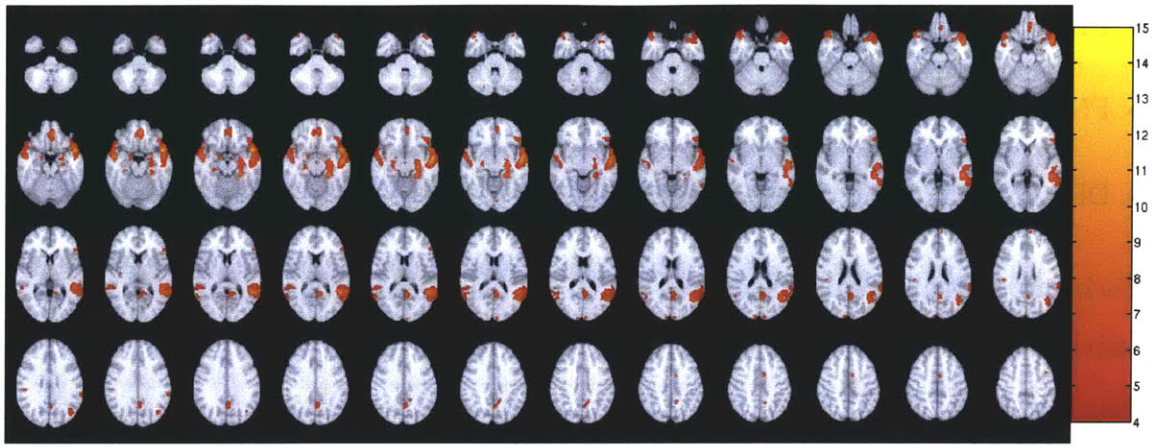
To explore this further, we performed the Random Effects Analysis for progressively more complex warps from subject space to the MNI152 template. We compared three different registration algorithms:

- FLIRT - affine registration provided by FSL [6]
- FNIRT - nonlinear registration provided by FSL [6]
- DD - diffeomorphic demons algorithm [15]

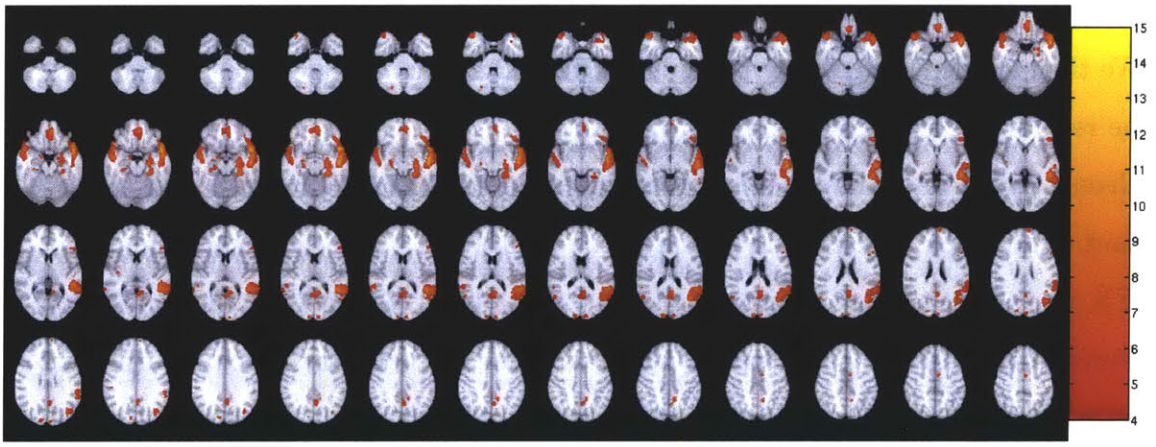
Each subject's anatomical scan was registered using each of the algorithms above to derive a transformation, which we used as a non-linear registration block in the pipeline in Figure 4-1. The resulting total transformation was applied to the subject's summary statistics from the individual-level GLM analysis to put them all in MNI152 space.

We then applied the standard Random Effects Analysis as described in Chapter 2, to the results of each registration algorithm.

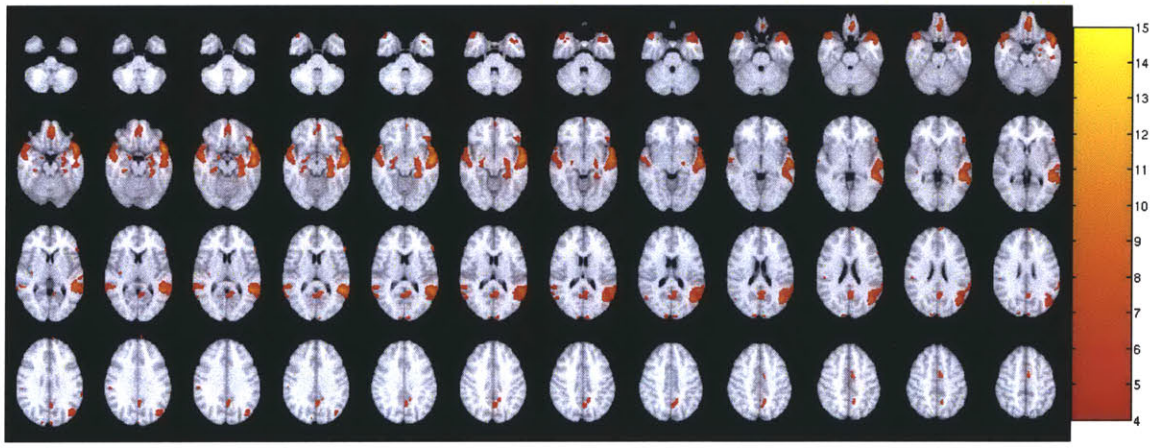
Group significance maps on both statistics were computed for each registration method. The maps for RFX are displayed in Figure 5-1. The significance maps for WRFX were very similar in magnitude and structure. Notice that the sensitivity, or overall magnitude of significances at each voxel, do not vary much for different registration methods. The structure of the thresholded regions are very similar and no heatmap is significantly brighter than the rest.



(a) FLIRT RFX

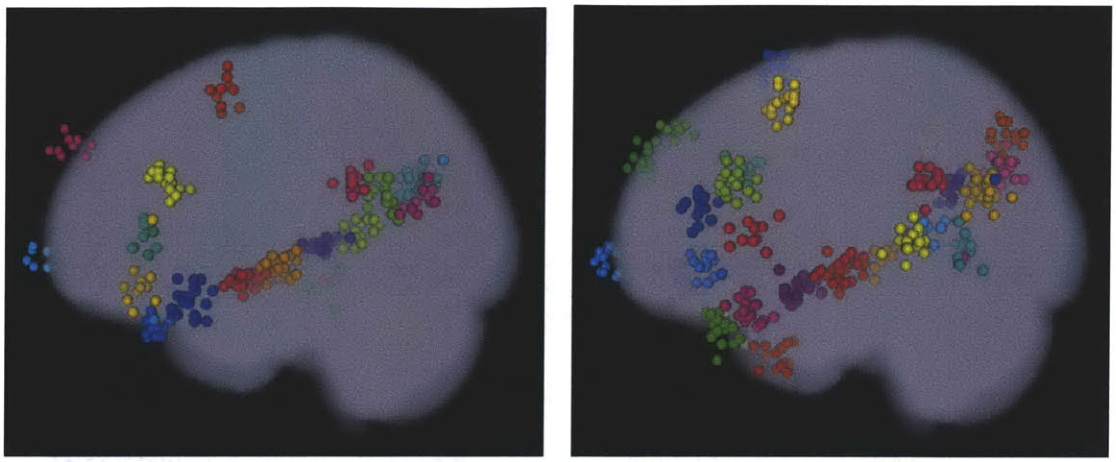


(b) FNIRT RFX



(c) DD RFX

Figure 5-1: Group significance maps for the sentences minus nonwords contrast computed using all 32 subjects. No method yields much higher significances than the others.



(a) Complete Data (b) First Half

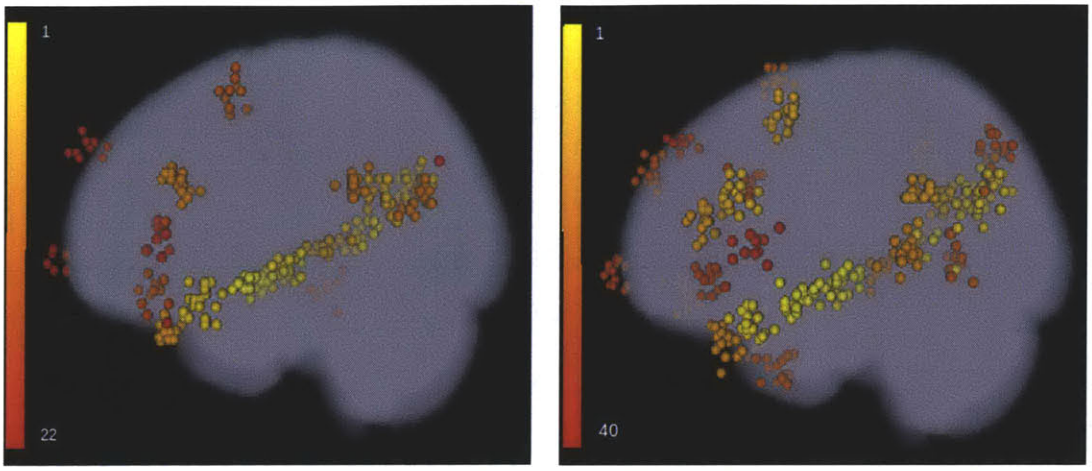
Figure 5-2: Clustering of peaks in the left hemispheres of subjects visualized in the MNI152 space. The complete data used GLM significance maps from all subject time points and the first half used maps from half of each subject's time points. Images were produced using matVTK [1].

5.3 Derived Clusters

The region matching algorithm was applied to the language localizer contrast of sentences minus nonwords to data from 33 subjects. The input to the algorithm were the thresholded significance maps for each subject derived from the individual-level GLM and the output was a set of peak clusters with their respective scores. The significance maps were thresholded at $p > 10^{-4}$. The resulting clusters in the left hemisphere are shown in Figure 5-2.

The locations of clusters agree well with the locations of language regions in the literature [5], particularly along the temporal lobe and parts of the frontal gyrus.

Each cluster was assigned a score based on produced clusters of peaks in the MNI152 space. The score from (3.3) quantifies how consistent the cluster is across different templates as well as how many subjects have a participant peak. A heatmap of this score is shown in Figure 5-3 for both clustering results.



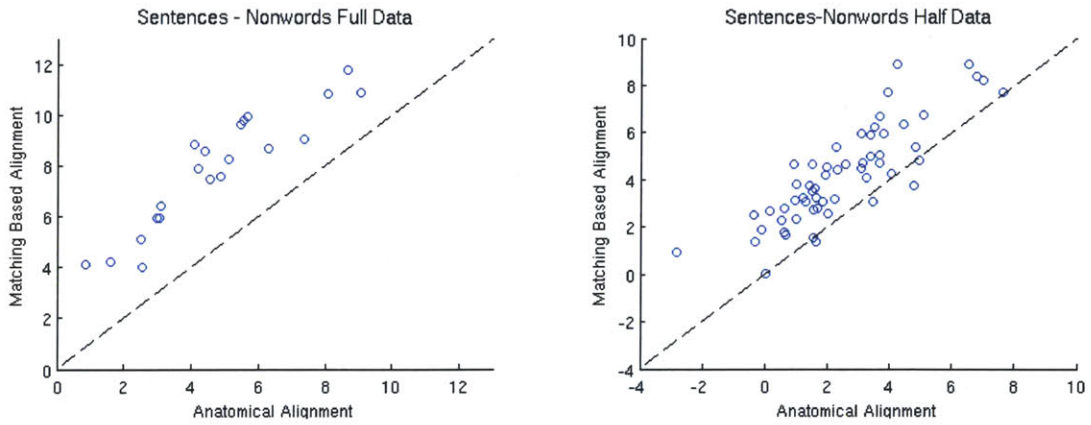
(a) Complete Data (b) First Half

Figure 5-3: Cluster scores by rank. Only clusters with at least 25% of subjects included are shown. Images were produced using matVTK [1].

5.4 Validation of Our Method

The complete results of RFX analysis based on the peak correspondences are shown in a scatter plot in Figure 5-4. The horizontal axis measures average significance of the clusters obtained from the Random Effect Analysis using standard anatomical alignment. The vertical axis measures average significance of the cluster using the matching algorithm's alignment followed by the Random Effects Analysis. As expected, both contrasts improve with the peak matching alignment derived from the data itself. Figure 5-5 illustrates examples of average cluster significances for the results of peak matching and one for the results of standard anatomical alignment.

To evaluate the method itself, the Hungarian-based matching algorithm was run on the sentences minus nonwords localizer contrast from the first half of the data, or 336 time points per subject. The alignment based on those clusters was applied on the second half of the data, or last 336 time points per subject. The alignment was applied to both the sentences minus nonwords contrast and the words minus nonwords contrast. The average significances for each cluster were compared to the results of standard anatomical alignment of the second half of the data similar to the case above. These results are also shown Figure 5-4. Examples of average cluster significances are displayed in Figure 5-5. An improvement is seen again for both contrasts, but it



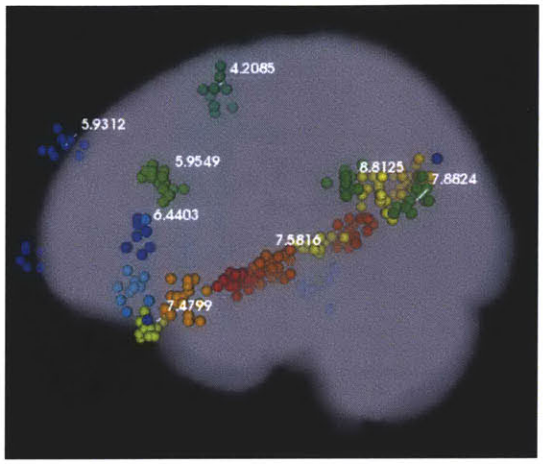
(a) Full Data (b) Half Data

Figure 5-4: Every point corresponds to a single cluster found using the matching algorithm. The horizontal axis represents the average significance obtained via RFX in anatomically aligned subjects. The vertical axis is the average significance value achieved by RFX on locally aligned peaks based on our matching method. This validation was done both on the complete data set and an independent half of the data.

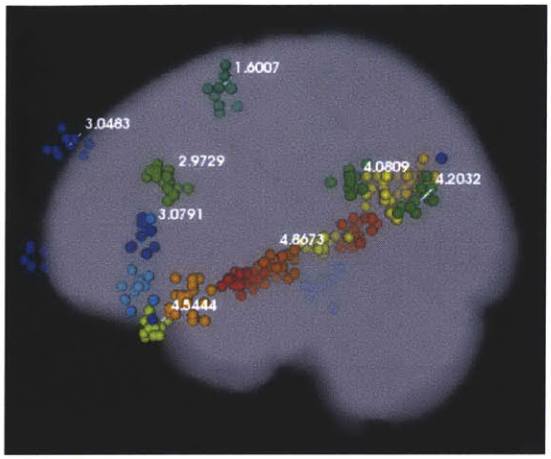
is notably smaller. This reflects the fact that the clusters were not derived from the same data set this time, but it still suggests that improvement can be achieved.

The results of using sentences minus nonwords from the complete and half data sets to align activation peaks from the words minus nonwords contrast are presented in Figures 5-7 and 5-6. These results are very similar to that of alignment on the sentences minus nonwords contrast, but overall, the average significances tend to be lower.

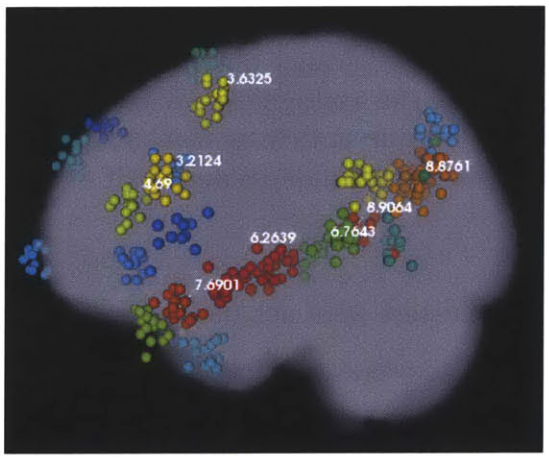
All the Figures use RFX, but a similar analysis was done using WRFX in each setup and the results were very similar. This reflects an additional positive feature of this validation. By simply re-aligning voxels locally around peaks, various existing algorithms of group analysis can be applied without much added difficulty.



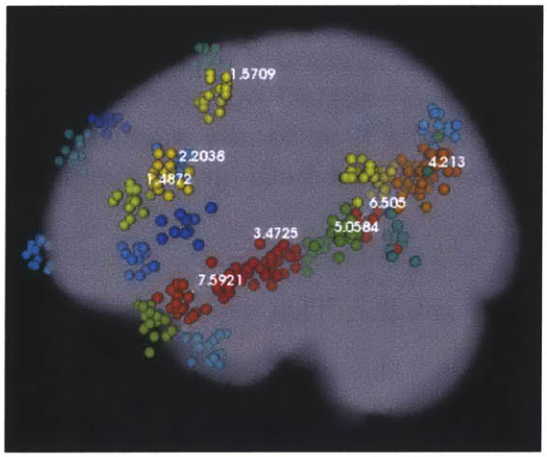
(a) Full Data Functional Alignment



(b) Full Data Anatomical Alignment

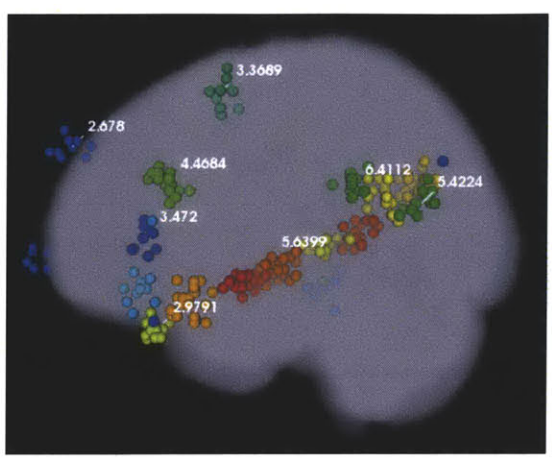


(c) Half Data Functional Alignment

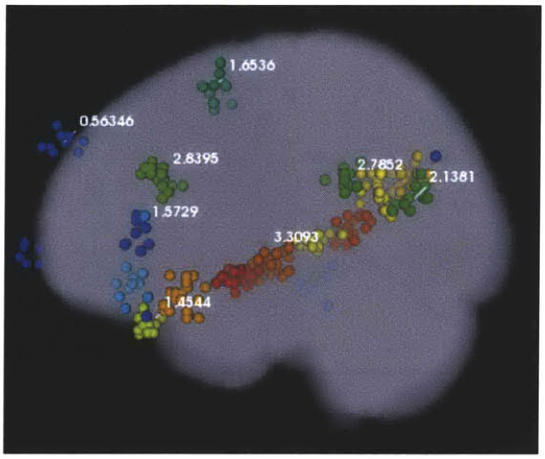


(d) Half Data Anatomical Alignment

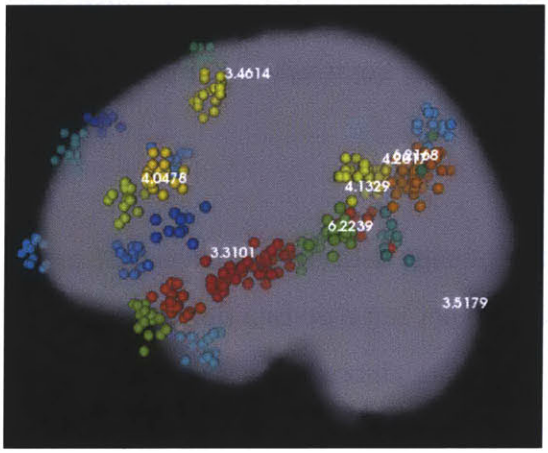
Figure 5-5: Average significance using peak alignment and anatomical alignment annotated on the clusters over the MNI152 brain. The peaks' alignment was derived and tested on both the complete set of subject data and an independent half of the data.



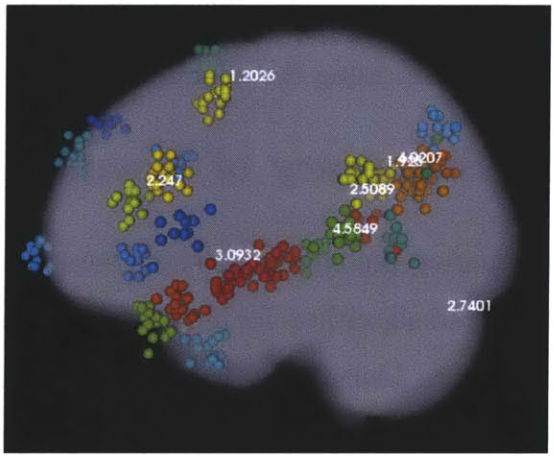
(a) Full Data Functional Alignment



(b) Full Data Anatomical Alignment



(c) Half Data Functional Alignment



(d) Half Data Anatomical Alignment

Figure 5-6: The alignment was derived from complete subject data and an independent half of the data for the sentences minus nonwords contrast but the significances were computed by applying the alignment to the words minus nonwords contrast.

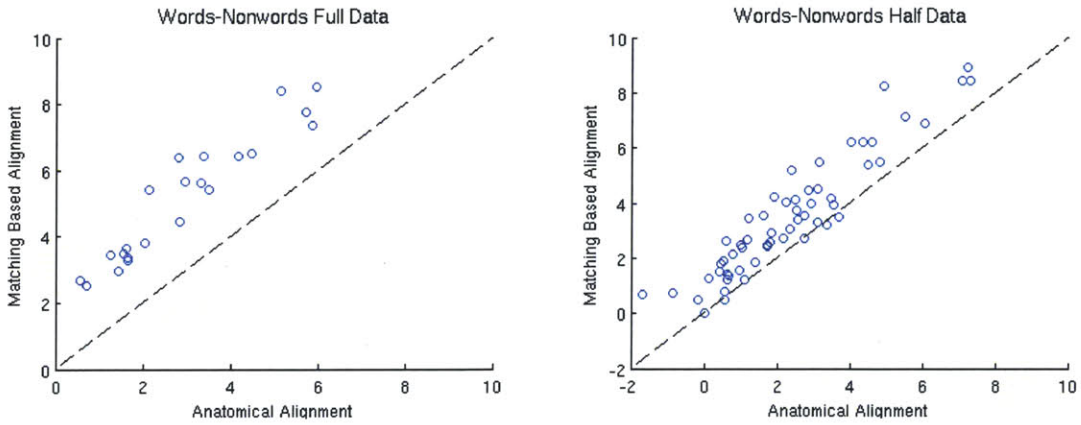


Figure 5-7: The alignment was derived from complete subject data and an independent half of the data on the sentences minus nonwords contrast but the significances were computed from using the alignment on the words minus nonwords contrast.

5.5 Summary of Results

To summarize, preliminary analysis done on our language data demonstrate that more sophisticated algorithms for anatomical alignment do not improve functional prediction for the group. Using clusters derived from our matching algorithm, greater sensitivity to language regions was achieved when the alignment was applied to independent fMRI data. As an example, activation near Broca's area was not detected using standard anatomical alignment after thresholding for significant voxels. However, such activation was recovered from significant voxels generating using our method's alignment.

Chapter 6

Discussion

In this thesis, we developed a method for identifying spatial correspondences of activation regions across subjects. The method relies on activation peaks as landmarks for matching of these regions. We have shown that using our method for alignment outperforms standard anatomical alignment when searching for voxels of significant activation. The method makes a few important assumptions that should be studied in future work, such as the requirement of one-to-one matchings of activation peaks across subjects. This chapter discusses the contributions of our method as well as its potential limitations.

6.1 Analysis of Hungarian Based Region Matching

Our matching algorithm provides a framework for finding spatial correspondences based on a user-specified metric of similarity. In our work, both functional and anatomical features were used in alignment. The subject activation maps are first transformed to a common space using anatomical registration. The anatomical information is represented in the measures of distance in the common space. Functional information is encoded as peaks of activation. The proposed method offers an improvement over using anatomical information only by improving group-level significances as shown in Chapter 4. Compared to the standard RFX on anatomically aligned functional data, the only difference was in the use of peaks of activation from individual level analysis

to further align the data. This suggests that peaks of activation are important spatial indicators of common brain activity across subjects. Indeed, this finding has been suggested in the literature [13]. The intractable combinatorial problem of pairwise matching of all voxels in two brains was reduced to the matching of the activation peaks.

Examining the clustering in Figure 5-2, one noticeable feature is the tight clustering of the peaks in the temporal lobe. The language activation in the temporal lobe was variable across subjects, and many subjects showed bright, contiguous regions of activation with multiple peaks in this region. Figure 5-2 shows that this scenario is a limitation of the algorithm. The high concentration of peaks here leads to a somewhat arbitrary clustering of the peaks due to the thresholded cost function used for matching peaks. As this threshold varies, the results can vary significantly. Moreover, it is not clear that an optimal threshold even exists. The desired properties of a good threshold for the cost function is one that is high enough to capture the spatial variability of functional regions of interest, but not too high so as to group peaks in clearly distinct locations. The response to language stimuli in the frontal lobe is better understood, so the threshold was targeted at distinguishing those regions. In this case, the variability in spatial location of these regions dictated what the threshold should be, and such a qualitative selection would have to be done for any contrast of interest.

The scores given to the clusters in the temporal lobe in Figure 5-3 are likely to be greater than desired. A perfect scoring metric would perfectly rank the distinct robust clusters with many member subjects highest. In regions with high concentrations of peaks, however, the scores will always be higher, regardless of how robust the distinct functional regions are. Nevertheless, the score is useful for filtering out the least robust clusters found.

The method for validating a set of correspondences examines whether a derived alignment for a set of subjects can improve upon group analysis of another data set for the same set of subjects. Thus, the validation verifies that the method improves over anatomical alignment, but it does not make generalizations to new subjects. Similar validation techniques have been employed for fMRI group analysis in the

literature [13].

6.2 Related Methods for Group Analysis

The visualization of clusters of peaks derived by this method has similar advantages to the group level overlap maps in [5]. Both can be used to match and align new subjects and create a useful visualization for qualitative analysis. Although our method does not provide the shape and the outlines of regions as in [5], it implies an approach for local alignment of regions that can be exploited for validation. Further work is needed to compare the locations of clustered peaks to group level regions identified by [5].

One advantage of the hierarchical Bayesian model presented in [19] is that it makes a very clear set of assumptions to model subjects' functional activations. Such assumptions include modeling the activation regions as Gaussian mixtures in a common anatomical space. It also avoids making some of the strong assumptions found in the matching method of this work. The use of the Hungarian algorithm enforces one-to-one correspondences between subjects. In other words, no peak from one subject can be matched to multiple peaks in another subject. Nothing in the literature suggests that this is the case and is a limitation of the algorithm. Language Regions in the temporal lobe, for example, could contain two peaks that represent the same entity. Without ground truth, it is challenging to evaluate this assumption. Furthermore, the method of scoring regions is clearly an approximation to the true cost function. An advantage of our method is in its simplicity and ease of implementation. The method presented in [19] suffers from high dimensionality and the slow iterative estimation of model parameters. With our matching method, however, it is clear that peaks of activation are the driving force for the method and thus their use as landmarks can be easily validated.

In [13], a different peak-based statistical method for identifying spatial correspondences across subjects was demonstrated. The next important step for our method is to compare it to the approach of [13], as it has also been shown to improve over standard RFX on anatomically aligned subjects. A comparison of our algorithm to

this and a many other unsupervised methods of group analysis is needed.

6.3 Future Work

As mentioned earlier, the matching method presented in this work creates a framework for enhancing anatomical alignment of subjects' fMRI data. By using a similar validation scheme, this method could be applied using a number of different metrics for region or peak similarity. The method would then become a tool for evaluating the importance of a metric.

Furthermore, our matching method should be compared to other sophisticated methods for group analysis of fMRI data. [13], [19], and [5], for example, have contributed to the evolution of such group analysis. Each algorithm takes a different approach to the problem and makes different assumptions. Future work should use a comparison to discover which assumptions represent the underlying functional organization of the brain better and enable discovery of further structure in the functional activation data.

Appendix A

Hungarian Algorithm

We provide a brief overview of the Hungarian algorithm, also known as the Kuhn-Munkres algorithm. This algorithm solves the linear assignment problem, which is typically presented as the problem of matching workers x_i to jobs y_j. Every worker-job pairing has an associated cost $C(x_i, y_j)$. The Hungarian algorithm finds the bipartite matching that minimizes the total cost $T(\cdot)$ of all matches. We start with definitions.

A feasible labeling $l(\cdot)$ maps all nodes in the bipartite graph to labels. It also has the following property:

$$l(x_i) + l(y_j) \geq C(x_i, y_j)$$

An equality subgraph is defined as a spanning subgraph over the bipartite graph with all of the same vertices, but only the edges which satisfy the following property:

$$C(x_i, y_j) = l(x_i) + l(y_j)$$

Finally, a perfect matching is a matching where every node is matched once. That is, exactly one edge is incident to every node in the graph.

Using these definitions, it is straightforward to prove that, if an equality subgraph G has a perfect matching $P*$ then $P*$ is an optimal matching. Given $P*$ and G, we define $T(P*) = \sum_{e \in P*} C(e)$. We also note that $\sum_{e \in edges} C(e) = \sum_{v \in vertices} l(v)$. Given a perfect matching P in G, $T(P) = \sum_{e \in P} C(e)$, and $\sum_{e \in P} C(e) \leq \sum_{v \in vertices} l(v)$,

$T(P) \leq T(P*)$. The last line establishes the maximum. Based on this result, the Hungarian Algorithm proceeds as follows:

1. Initialize a feasible set of labels and make the corresponding equality subgraph E.
2. Find the maximum matching in E.
3. If the matching is perfect, then it is optimal the algorithm terminates. Otherwise, continue.
4. Add alternating edges from x vertices not in E until an augmenting path is made.
5. Revise the labels $l(\cdot)$ and augment the matching with the augmenting path that is found. Go to step 4.

The Hungarian algorithm was one of the first solutions proposed for the linear assignment problem. It shares many noticeable similarities with network flow algorithms.

Bibliography

[1] Erich Birngruber, René Donner, and Georg Langs. matvtk - 3d visualization for matlab. In *Proceedings of the MICCAI 2009 Workshop on systems and architectures for CAI*, 2009.

[2] A. Dempster, N. Laird, and D. Rubin. Maximum likelihood from incomplete data via the em algorithm. *J. Royal Statistical Society, Series B*, 39(1):1–38, 1977.

[3] N F Dronkers, O Plaisant, M T Iba-Zizen, and E A Cabanis. Paul brocas historic cases: high resolution mr imaging of the brains of leborgne and lelong. *Brain*, 130(Pt 5):1432–41, 2007.

[4] Luciano Fadiga, Laila Craighero, and Alessandro DAusilio. Brocas area in language, action, and music. *Annals Of The New York Academy Of Sciences*, 1169(1):448–458, 2009.

[5] Evelina Fedorenko, Po-Jang Hsieh, Alfonso Nieto-Castanon, Susan Whitfield-Gabrieli, and Nancy Kanwisher. New method for fmri investigations of language: defining rois functionally in individual subjects. *J Neurophysiol*, 104(2):1177–94, 2010.

[6] FMRIB. Fsl. `http://www.fmrib.ox.ac.uk/fsl/`.

[7] Martinos Center for Biomedical Imaging. Freesurfer. `http://surfer.nmr.mgh.harvard.edu/`.

[8] Nancy Kanwisher and Galit Yovel. The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society of London B*, 361:2109–2128, 2006.

[9] H.W. Kuhn. The Hungarian method for the assignment problem. *Naval research logistics quarterly*, 2(1-2):83–97, 1955.

[10] Carl N Morris. Parametric empirical bayes inference: Theory and applications: Comment. *Journal of the American Statistical Association*, 78(381), 2008.

[11] Mert R. Sabuncu, Benjamin D. Singer, Bryan Conroy, Ronald E. Bryan, Peter J. Ramadge, and James V. Haxby. Function-based intersubject alignment of human cortical anatomy. *Cereb. Cortex*, 2010.

[12] B. Thirion, G. Flandin, P. Pinel, A. Roche, P. Ciuciu, and J.-B. Poline. Dealing with the shortcomings of spatial normalization: Multi-subject parcellation of fMRI datasets. *Hum. Brain Mapp.*, 27(8):678–693, Aug. 2006.

[13] B. Thirion, P. Pinel, A. Tucholka, A. Roche, P. Ciuciu, J.-F. Mangin, and J.-B. Poline. Structural analysis of fMRI data revisited: Improving the sensitivity and reliability of fMRI group studies. *IEEE Trans. Med. Imag.*, 26(9):1256–1269, Sep. 2007.

[14] Bertrand Thirion, Philippe Pinel, Sbastien Mriaux, Alexis Roche, Stanislas Dehaene, and Jean-Baptiste Poline. Analysis of a large fMRI cohort: Statistical and methodological issues for group analyses. *Neuroimage*, 35(1):105–120, Mar 2007.

[15] Tom Vercauteren, Xavier Pennec, Aymeric Perchant, and Nicholas Ayache. Diffeomorphic demons: efficient non-parametric image registration. *NeuroImage*, 45(1 Suppl):S61–S72, 2009.

[16] Xingchang Wei, Seung-Schik Yoo, Chandlee C Dickey, Kelly H Zou, Charles R G Guttmann, and Lawrence P Panych. Functional mri of auditory verbal working memory: long-term reproducibility analysis. *NeuroImage*, 21(3):1000–1008, 2004.

[17] M W Woolrich, T E Behrens, C F Beckmann, M Jenkinson, and S M Smith. Multilevel linear modelling for fmri group analysis using bayesian inference. *Neuroimage*, 21(4):1732–47, 2004. Journal ArticleResearch Support, Non-U.S. Gov'tUnited States.

[18] K.J. Worsley and K.J. Friston. Analysis of fMRI time-series revisited - again. *NeuroImage*, 2:173–181, 1995.

[19] Lei Xu, Timothy D. Johnson, Thomas E. Nichols, and Derek E. Nee. Modeling inter-subject variability in fMRI activation location: A Bayesian hierarchical spatial model. *Biometrics*, 65(4):1041–1051, 2009.