

Locally-Optimized Inter-Subject Alignment of Functional Cortical Regions

Marius Cătălin Iordan

Ph.D. Candidate, Computer Science, Enrolled in CS 231A
Stanford University, Stanford, CA

mci@stanford.edu

Armand Joulin

Postdoctoral Fellow, Computer Science, Project Mentor
Stanford University, Stanford, CA

ajoulin@stanford.edu

Abstract

Inter-subject registration of cortical areas is necessary in functional imaging (fMRI) studies for making inferences about equivalent brain function across a population. However, many high-level visual brain areas are defined as peaks of functional contrasts and it is usually difficult to identify clear anatomical landmarks and boundaries for these areas, due to large variability in their cortical position. As a consequence, previous methods usually fail to accurately map such functional regions of interest (ROIs) across participants. To address this problem, we propose a locally optimized registration method that directly predicts the location of a seed ROI on a separate target cortical sheet by maximizing the functional correlation between regions and simultaneously constraining the global structure of the mapping, while allowing for non-local deformations in its topology. Our registration method outperformed two canonical baselines (anatomical landmark-based AFNI alignment and cortical curvature-based FreeSurfer alignment) in the percentage of overlap between predicted region and ground truth LOC. Furthermore, the maps obtained using our method are more consistent across subjects than both baseline measures. Consequently, our method has the ability to directly and immediately improve the quality of group maps for high-level visually in countless fMRI studies. This would dramatically increase the statistical power of such studies, as a more accurate mapping to a common space implies less smoothing and larger effect sizes.

1. Introduction

Inter-subject registration of cortical areas is necessary in functional imaging (fMRI) studies for making inferences

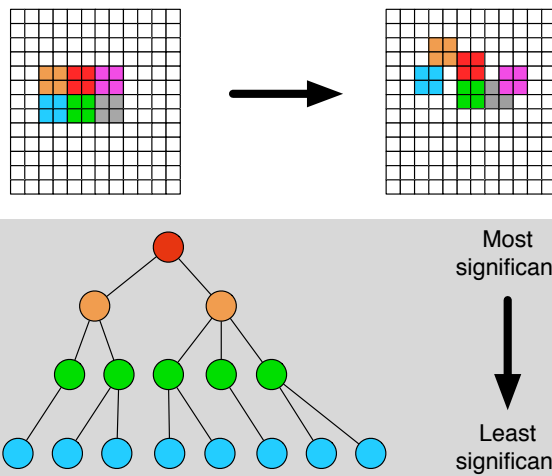


Figure 1. Descriptions of our two methods: (Top) Shallow tree co-localization. (Bottom) Deep tree hierarchical representation.

about equivalent brain function across a population. Most state-of-the-art alignment methods define transformations between entire cortical volumes that attempt to preserve anatomical landmarks, cortical curvature, or functional connectivity, and subsequently check whether specific regions of interest (ROIs) are accurately matched between subjects [11, 8, 1]. However, many high-level visual brain areas are defined as peaks of functional contrasts (i.e. higher activation for scenes versus objects) and it is usually difficult to identify clear anatomical landmarks and boundaries for these areas, due to large variability in their cortical position. As a consequence, although they provide a reasonable global matching, previous methods usually fail to accurately map such functional ROIs across participants.

The goal of our project is to increase the reliability of

inter-subject mapping for these cortical functional peaks and for the visual areas they define using fMRI data.

An improved solution to this problem would directly and immediately improve the quality of group maps for high-level visual areas in countless fMRI studies. Consequently, this would dramatically increase the statistical power of such studies, as a more accurate mapping to a common space implies less smoothing and larger effect sizes.

Moreover, obtaining a reliable functional region mapping across subjects would represent a significant step towards obviating the need for running time- and resource-consuming localizer scans for every study participant (the method currently used to identify these cortical areas). Such a mapping is also useful in settings where one needs to compare analyses and hypotheses between datasets where functional localizers are missing and gathering extra sessions of data is either expensive (large number of participants) or impossible (unavailability of former subjects).

Finally, the relationship between peaks of functional contrasts and the computation performed by cortex surrounding them is not well understood. Therefore, a method that improves the quality of functional ROI mapping between subjects would also become especially useful for investigating the key complex relationship between anatomy, functional contrast peaks (ground-truth ROIs), and cortical computation (measured fMRI BOLD response).

To address the problem of computing correspondences between equivalent functional regions, we use the key observation that peak functional contrast points (where ROIs are centered) must share similar function between subjects, even though functional gradients of selectivity surrounding the peaks may not be spatially organized identically across subjects [6]. Thus, a key requirement for our methods must be to allow for at least a small degree of non-smoothness in the local deformations afforded by the mapping between the two cortical surfaces. Furthermore, we will focus our efforts on directly capturing the correspondence between regions of interest by finding a locally strong functional match between regions, rather than seeking a perfect one-to-one correspondence between entire cortical volumes (a hallmark of many previous approaches).

The first method we propose is a weak tree representation, where the region of interest to be matched is split into several sub-regions. These smaller regions are allowed to match independently to regions in the target map, with the overall constraint that they must remain within close proximity of each other (a constraint given by the root of the two-level tree). This method is illustrated in Fig. 1 (Top). The second method we propose leverages a similar structure, but allows for a deeper hierarchical representation to be built on top of each area. Here, we define a 3+ level tree where each level corresponds to decreasing significance thresholds for the contrast map originally used to define the

functional area. This extension of the previous method is shown in Fig. 1 (Bottom).

2. Related Work

Many cortical alignment methods have been previously proposed, virtually all of which define transformations between entire cortical volumes and subsequently check whether specific regions of interest (ROIs) are accurately matched between subjects.

Anatomical alignment relies on large scale correspondences between all human brains, such as the reliable presence and relatively consistent positions of primary features such as major sulci and gyri on the cortical surface. Of this type, the most widely used registration method is Talairach alignment, which warps the seed cortical volume to a pre-set fixed atlas based on manually or automatically selected anatomical landmarks [10]. Similarly, the AFNI tool [2] also attempts to preserve anatomical landmarks, except it allows direct warping between two subjects' brains, without the need for common alignment to a known atlas. A more complex method, which achieves a more precise anatomical alignment of secondary features was proposed in [11], however, it suffers the same shortcomings in matching functional areas previous anatomical alignment methods.

Given that the main obstacle in aligning the cortical surface between subjects is its folding variability, methods have been proposed that warp gray matter meshes by taking into account local curvature properties of the cortex. A widely used tool is Freesurfer [4], which is capable of very high quality matching of major, as well as secondary cortical features. FreeSurfer relies on extracting the boundary between gray and white matter, and then aligns these boundaries between brains by minimizing the differences between corresponding normal vectors to the boundary surface. In our work, we test our method against both AFNI and FreeSurfer standard alignment techniques and show that they perform poorly for high-level visual functionally-defined areas of interest.

Another recent method that goes beyond anatomical features to perform alignment incorporates functional connectivity constraints in the mapping [1]. This method leverages the fact that connectivity between brain regions during resting states (i.e. no active task being performed: subject keeps eyes closed and allows his or her mind to *wander*) are thought to be fairly consistent across the population. These methods show improved ability to align functional areas that are part of known large intertwined networks in the brain (i.e. the *default mode network*). However, many functional areas, especially high level occipito-temporal areas involved in vision which we are interested in aligning, are not usually a strong part of these networks and thus receive little benefit from such methods.

Finally, alignment has also been performed by incorpo-

rating functional correlation constraints. Of note is a recent method that starts with FreeSurfer alignment as an initialization step and then maximizes local functional correlation across the cortical surface to *nudge* the vertices of the map into a new alignment that takes into account functional responses [8]. This method performs well for early visual areas (which are also more easily aligned using anatomical landmarks), but shows limited ability to match functional regions as distance from the occipital pole increases. This approach is similar to our own, however, we take a localized approach and we also allow non-smooth local deformations of the maps, whereas previous work always uses continuous maps.

Another type of method that leverages functional correlation to perform cortical matching is hyperalignment [5]. In this case, cortical responses from two brains are modeled as vectors in a high dimensional space (of dimension equal to the number of voxels considered in each map) and then a rigid rotation and reflection transformation is derived such that it optimally aligns the two vectors together. While showing excellent cross-decoding performance across maps, this method essentially represents a point in the target map as a linear combination of (possibly) all voxels in the other map, and thus is not directly amenable to transferring the location of one contained area across maps. Usually, this method is tested by transferring the contrast map used to defined a localizer in the seed map to the target map and then extracting the region of interest in the target map. In our work, we implement a version of hyperalignment for the purpose of comparing our method against it.

3. Cortical Region Alignment

Our goal is to align functionally-defined high-level visual areas between participants. To compute such a correspondence between regions, we reasoned that although two cortical surfaces (corresponding to two separate subjects) must express the same necessary computational units that give rise to observed function, these units might not be perfectly equivalent or identically distributed spatially across the two ROIs. Therefore, we designed an alignment method which optimizes functional correlation between regions while allowing for non-smooth local deformations in the mapping.

Recently, we proposed a locally optimized registration method that directly predicts the location of a seed region of interest (ROI) on a separate target cortical sheet by leveraging the principle detailed above [7]. We refer to this method as *Shallow Tree Alignment* and describe it in depth later in this section. This method represents a reasonable first attempt to solve the functional region registration problem described in the previous sections. For this project, we sought to improve this method in three ways: (1) improve its per-

formance by experimenting with a different way of representing the fMRI data on the cortical surface; (2) generate a more in depth comparison between this method and current state of the art methods; and (3) extend our method by leveraging insights from known neural properties of the areas we are attempting to match.

Throughout this section, we will highlight the innovations implemented as part of this project versus the initial method which was recently accepted for presentation at a human vision conference [7].

3.1. Interpolation

Cortical computation occurs in a thin sheet at the surface of the brain (gray matter) and this sheet is folded in 3D space, which gives it a large surface area compared to the volume it occupies inside the skull. This entails that points that might be close together in 3D could, in fact, be separated by a relatively large distance on the 2D surface, the latter of which is a more accurate measure of neural population distance.

Using functional neuroimaging (fMRI), we can detect changes in blood oxygenation associated with neural activity from cubic voxels (3D pixels) that tile the entire cortical volume at a spatial resolution of 1.75 x 1.75 x 4 mm. However, this means that in order to recover the true distribution of activity on the 2D gray matter surface, we must find a way to accurately estimate the location of the 2D plane, project the 3D volume onto it, and then interpolate our measurements between the cubic data points and the points on the surface.

The Shallow Tree Alignment algorithm used an existing software package to do the projection and interpolation step (AFNI-SUMA [2]), before attempting to find corresponding regions between two cortical surfaces. While SUMA provides an exceptional solution to the problem of estimating the 2D surface, it is unclear how accurate the default linear interpolation step is at creating a reasonable projection of the 2D flat cortical map onto a regular grid suitable for algorithmic manipulation. Consequently, we sought to characterize the sensitivity of our algorithm to the type of interpolation used by computing a Gaussian projection between the 3D volume and the 2D surface. We implemented this new projection by computing the 5 nearest neighbors of each grid corner on the flat map and then using a Gaussian kernel to interpolate between these neighbors to obtain our grid point features.

3.2. Baselines

To compare our method to existing alignment techniques, we initially implemented two canonical widely used baselines: anatomical landmark-based AFNI [2] alignment and cortical curvature-based FreeSurfer [4] alignment. AFNI uses information about brain shape and automatically

defined anatomical points of interest to warp cortical volumes from one subject to another. FreeSurfer also uses brain shape, but instead uses information about cortical curvature (sulci and gyri locations, as well as the distribution of normals to the cortical gray matter surface) to iteratively distort one cortical surface into another.

Although these two methods are the most commonly used cortical registration methods, several other state-of-the-art methods exist that claim to obtain better functional region alignment. As part of this project, we further implemented one such method called *hyperalignment* [5]. Hyperalignment embeds two cortical surfaces into a high dimensional space whose features are determined by functional responses to the same set of stimuli. Then, a series of rotations and translations are estimated which best align the shapes of the two spaces to one another. This alignment can subsequently be used to project a region from one map to the other. Given a seed map X and a target map Y , hyperalignment can be written as the following optimization problem:

$$\begin{aligned} \min_{t,s,R} \quad & \|Y - t - sXR\|_F \\ \text{s.t.} \quad & R^T R = I, \end{aligned}$$

where $\|\cdot\|_F$ represents the Frobenius norm of a matrix and the estimated parameters t, s, R refer to translation, scale, and rotation components, respectively. This optimization problem can be solved exactly using the single value decomposition of the matrix product $X^T Y$.

As part of our project, we implemented this state-of-the-art alignment method and we will subsequently use it as an additional baseline for our experiments, one that potentially outperforms canonical ones, and whose performance would be a better benchmark for both our methods (the Shallow Tree and Deep Hierarchical Tree Alignment algorithms).

3.3. Shallow Tree Alignment

We proposed and implemented a locally optimized registration method that directly predicts the location of a seed region of interest (ROI) on a separate target cortical sheet by leveraging the principle detailed above [7].

The algorithm optimized functional correlation between pairs of regions on two-dimensional manifolds of inflated cortical surfaces. The method is inspired by an object colocalization technique [3] and represents regions in each cortical map as grid graphs of $n \times n$ nodes, where each node is associated with a small brain area (i.e. 5×5 pixels cortical surface patch). Correspondences are then established between separate cortical maps by maximizing the sum of feature correlations across all the nodes, while enforcing that the distances between connected nodes in the

graph change by less than a specified amount in each direction (i.e. 5 pixels) after projecting onto the other cortical surface. The initial region parcellation, as well as a potential matching are shown in Fig. 1 (Top). The objective function can be written as follows:

$$\begin{aligned} \min_M \quad & \sum_j d_F(F_i, F_{m_i}) \\ \text{s.t.} \quad & d_s(p_{m_i}, p_{m_j}) \leq \rho, \end{aligned}$$

where:

- $M = \{(k, m_k)\}$ is the collection of correspondences between nodes in the first cortical surface (k) and nodes in the second cortical surface (m_k);
- d_F is the feature distance between the nodes in each correspondence;
- d_s is the distance difference between the original and mapped configuration of each pair of points (point i mapped to point j) in the two maps; and
- ρ is the maximum allowable distance change between neighboring patches across maps.

This optimization is solved using a deterministic grid search through the space of all possible node jitter permutations.

3.4. Deep Tree Hierarchical Alignment

To extend our original method, we propose leverages a similar structure, but allows for a deeper hierarchical representation to be built on top of each area.

In neuroscience, functionally-defined ROI are identified as peaks of a general linear model contrast seeking consistent increased activation for one type of stimulus over another (i.e. which area of the brain responds more to faces than to houses). To obtain the boundaries of such a region, an usually arbitrary statistical threshold is chosen (e.g. $p < 0.001$) for what constitutes significant activation difference. Here, we propose that we build a hierarchical representation of a region of interest by defining a 3+ level tree where each level corresponds to decreasing significance thresholds for the contrast map originally used to define the functional area. Thus, a node closer to the root has a high chance of being representative of the activation in that functional area (high statistical significance), whereas leaves are likely part of the area, but may hold noisier representations and thus more difficult to match across participants. This extension of the previous method is depicted in Fig. 1 (Bottom).

Similarly to [3], we can treat this tree as a pseudo-MRF, and solve for the best matching of each node between cortical maps by using the Max-Product algorithm.

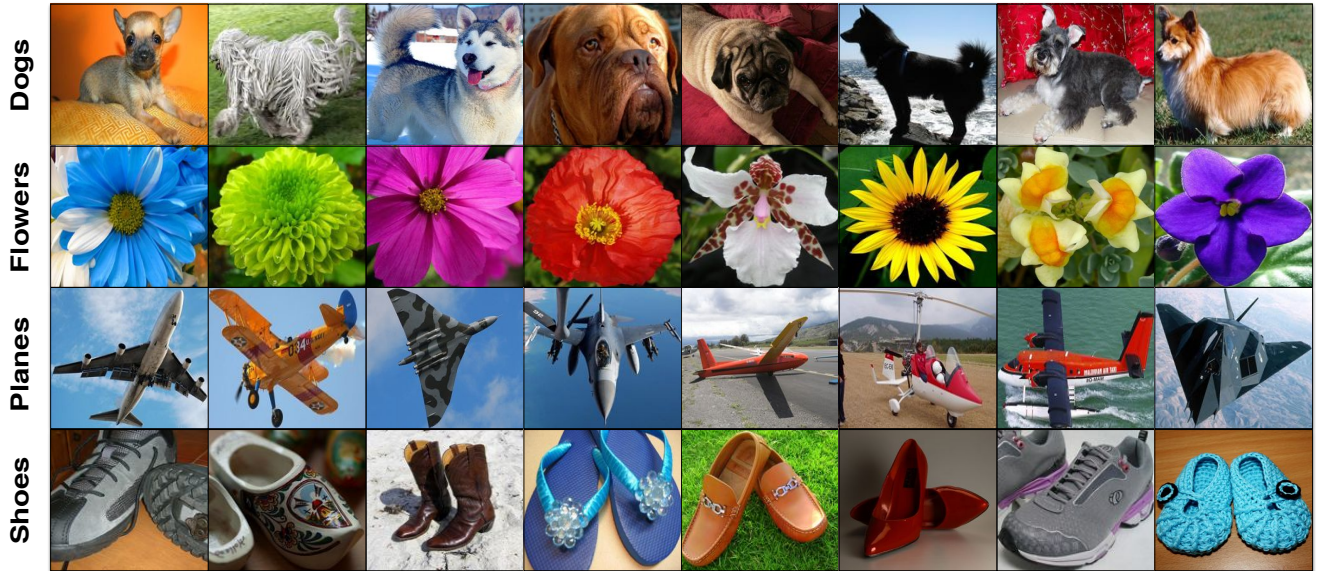


Figure 2. Stimulus set for fMRI experiment used to perform and evaluate cortical alignment. During the experiment, participants were shown images from 32 object categories: 8 breeds of dogs, 8 types of flowers, 8 types of planes, 8 types of shoes (32 images per category; 1,024 images total).

4. Experiments

4.1. Dataset

We tested our method by aligning a difficult to match, functionally defined, object-selective ROI (lateral occipital complex, LOC) between subjects using data from a passive-viewing fMRI experiment where participants were shown 1,024 images of objects from 32 categories: 8 breeds of dogs, 8 types of flowers, 8 types of planes, 8 types of shoes. The dataset contains 512 TRs (17m 4s) of fMRI data per subject. The stimulus set used in the experiment is shown in Fig. 2. We used brain activations elicited by these stimuli as features corresponding to each point on the cortical map (i.e. each cortical point has a 512 feature representation).

4.2. Results

To test how well our method aligns functional ROIs between cortical surfaces, we used two metrics:

- **Accuracy:** percentage of overlap between ground truth region and predicted region after mapping from a different subject’s brain;
- **Consistency:** amount of overlap between predicted regions from multiple subjects aligned to the same target map.

Overlap is computed as intersection over union for predicted and ground truth maps. Ground truth was established in a canonical fashion through a separate localizer scan in

each subject [9]. Below, we show results for two of our baselines (AFNI and Freesurfer), as well as two versions of the Shallow Tree Alignment, the first using linear interpolation and the second using Gaussian interpolation. As of the conclusion of the project, we had completed the implementation of the hyperalignment algorithm, but the result suite had not yet finished running (due to the size of the cortical maps, the analysis requires about 1 week to complete on our server), and the Deep Tree Hierarchical Alignment algorithm was still in the development and debugging stage.

A summary of our most recent results is given in Fig. 3. Our registration method vastly outperformed the two canonical baselines (anatomical landmark-based AFNI alignment and cortical curvature-based FreeSurfer alignment) in the percentage of overlap between predicted region and ground truth LOC: baselines 10-11%, ours 24-25%. Furthermore, the maps obtained using our method are more consistent across subjects than both baseline measures (overlap of region commonly mapped from 3+ subjects: baselines 9-11%, ours 26%). Moreover, our method is not sensitive to the type of interpolation used to generate the 2D cortical grid from the 3D volume representation: we see that when using either linear and gaussian interpolation we obtain very similar results, and in both cases outperform the two shown baselines.

Qualitatively, the cortical maps further showcase the strength of our results compared to the AFNI and FreeSurfer baselines. In the first two images (Fig. 3, bottom left) we see that functional regions in other subjects are mapped with

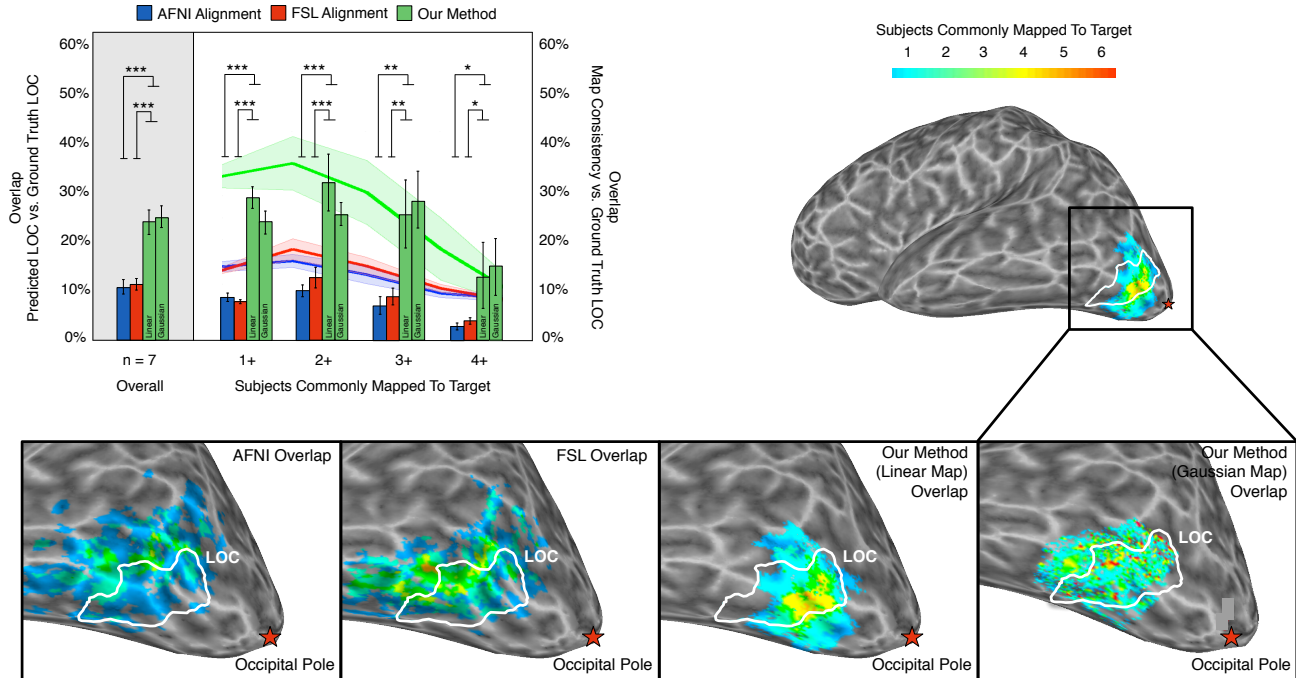


Figure 3. **Alignment Results: Accuracy and Consistency** ($n = 7$ subjects). For every target subject, we align LOC from all other 6 subjects to the target cortical surface using functional data from the above experiment. **(Graph Top Left)** Overlap between predicted LOC and ground truth LOC (i.e. defined using separate standard localizer procedure), measured as intersection over union of surfaces. **(Graph Top Right)** We select the voxels predicted consistently in the target map for $n+$ subjects and compute the overlap between this restricted region and ground truth LOC for $n \in \{1, 2, 3, 4\}$. **(Bottom Cortical Maps)** Consistency of predicted LOC obtained from aligning using AFNI, FSL, and Shallow Tree Alignment with linear and Gaussian interpolation. Heatmap indicates how many subjects’ LOC were mapped to that voxel on the target surface.

a high degree of variance onto the target subject cortical sheet. Often, there is little overlap with the ground truth ROI and, most importantly, the mapping may place the region several centimeters away from its desired location, often on a different gyrus. By contrast, our method (Fig. 3, bottom right) shows much less variance in the predicted area, with the peak of the prediction fully contained within the ground truth region.

These results suggest that our registration technique increases the reliability of transferring the location of functionally defined ROIs between subjects, which is an important step towards obviating the need for prohibitively expensive or impossible to obtain localizer scans.

5. Conclusion

In this project, we proposed a locally optimized registration method that directly predicts the location of a seed region of interest (ROI) on a separate target cortical sheet by maximizing the functional correlation between regions and simultaneously constraining the global structure of the mapping, while allowing for non-local deformations in its

topology.

Our method vastly outperforms two canonical baselines (anatomical-landmark-based AFNI alignment and cortical-curvature-based FreeSurfer alignment) in overlap percentage between predicted region and ground truth LOC and our predicted maps are more consistent across subjects than both baselines. Therefore, our technique improves the quality and reliability of matching and transferring the location of functional ROIs across subjects, an important step towards obviating the need for additional or impossible to obtain localizer scans.

Furthermore, our work can directly and immediately improve the quality of group maps for high-level visual areas in countless fMRI studies. Consequently, this would dramatically increase the statistical power of such studies, as a more accurate mapping to a common space implies less smoothing and larger effect sizes. Moreover, our work represents a significant step towards obviating the need for running time- and resource-consuming localizer scans for every study participant. Our method is also applicable to settings where one needs to compare analyses and hypotheses between datasets where functional localizers are miss-

ing and gathering extra sessions of data is either expensive (large number of participants) or impossible (unavailability of former subjects).

Finally, the relationship between peaks of functional contrasts and computation performed by cortex surrounding them is not well understood. Since our method improves the quality of functional ROI mapping between subjects, it becomes especially useful for investigating the key complex relationship between anatomy, functional contrast peaks, and cortical computation.

References

- [1] B. R. Conroy, B. D. Singer, J. S. Guntupalli, P. J. Ramadge, and J. V. Haxby. Inter-subject alignment of human cortical anatomy using functional connectivity. *NeuroImage*, 81:400–411, 2013.
- [2] R. W. Cox. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.*, (29):162–173, 1996.
- [3] O. Duchenne, A. Joulin, and J. Ponce. A Graph-matching Kernel for Object Categorization. *International Conference on Computer Vision*.
- [4] B. Fischl. FreeSurfer. *NeuroImage*, 62(2):774–781, 2012.
- [5] J. V. Haxby, J. S. Guntupalli, A. C. Connolly, Y. O. Halchenko, B. R. Conroy, M. I. Gobbini, M. Hanke, and P. J. Ramadge. A common, high-dimensional model of the representational space in human ventral temporal cortex. *Neuron*, 72(2):404–16, 2011.
- [6] A. G. Huth, S. Nishimoto, A. T. Vu, and J. L. Gallant. A continuous semantic space describes the representation of thousands of object and action categories across the human brain. *Neuron*, 76(6):1210–24, 2012.
- [7] M. C. Iordan, A. Joulin, D. M. Beck, and L. Fei-Fei. Locally-Optimized Inter-Subject Alignment of Functional Cortical Regions. *Vision Sciences Society (VSS)*, 2014.
- [8] M. R. Sabuncu, B. D. Singer, B. Conroy, R. E. Bryan, P. J. Ramadge, and J. V. Haxby. Function-based inter-subject alignment of human cortical anatomy. *Cerebral Cortex*, 20(1):130–40, 2010.
- [9] R. Sayres and K. Grill-Spector. Relating retinotopic and object-selective responses in human lateral occipital complex. *Journal of Neurophysiology*, 100:249267, 2008.
- [10] J. Talairach and P. Tournoux. Co-planar stereotaxic atlas of the human brain. *Thieme Publishing Group*, 1988.
- [11] T. Yeo, F. M. Krienen, J. Sepulcre, M. R. Sabuncu, D. Lashkari, M. Hollinshead, J. L. Roffman, J. W. Smoller, L. Zollei, J. R. Polimeni, B. Fischl, H. Liu, and R. L. Buckner. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106(3):11251165, 2011.