Simple ITK multi atlas registration and segmentation: Methods and tutorials

Brian B. Avants

The University of Pennsylvania, Philadelphia, PA 19104.

Expected cost (include direct and indirect/overhead costs) for the proposed work.

Total cost = $99,995, Direct cost = $62,497, Indirect cost = $37,498

Short description or abstract of the proposed work

We propose to augment Simple ITK with an integrated framework for registration and segmentation of biomedical images. Three components will interact: (1) Scalable registration methods that employ point set representations of *both* fixed and moving images will, for suitable forms of data, allow accurate registration with reduced memory footprint; (2) Registration methods will feed data into classical prior-based image segmentation based on expectation-maximization algorithms; (3) Registration methods will also initialize multi-atlas segmentation, specifically the class of Joint Fusion multi-atlas segmentation methods. This contribution would extend the applicability of the current ITK version 4 analysis framework to include full image quantification pipelines that are appropriate for diverse application areas such as template-based brain mapping or quantitative high-resolution microscopy. We will provide Simple ITK tutorials, unit tests as well as the necessary JSON and other metadata needed for Simple ITK bindings.

Address of the Offering Institution (including e-mail, fax, phone, and surface-mail address).

Office of Research Services, 3451 Walnut Street, Suite P-221

Philadelphia PA 19104-6205

Email: PennAORS@lists.upenn.edu; Phone Number: 215-746-0234; Fax Number: 215-898-9708

List of all named personnel/investigators in the proposal

- Brian B. Avants, Ph.D., Department of Radiology, University of Pennsylvania.

- James C. Gee, Ph.D., Department of Radiology, University of Pennsylvania.

- Paul A. Yushkevich, Ph.D., Department of Radiology, University of Pennsylvania.

- Nicholas J. Tustison, D.Sc., Department of Radiology and Medical Imaging, University of Virginia (consultant).

Technical Proposal

Purpose of the proposed work

Annotations and anatomical labelings are critical to the interpretation and statistical analysis of biomedical images. However such expert labeling is both expensive and time consuming to obtain and often out of reach for many institutions due to lack of local expertise and/or funding to hire such expertise. However, libraries of expert annotations are becoming available in ever increasing numbers. Expert annotation on biomedical imaging may exist at different scales from labeling cell types in histology to labeling microscopic structures in optical images all the way up to millimeter or centimeter scale neuroanatomy visible in magnetic resonance images. These digital annotations attached to biomedical images can take several hours to several weeks or more in order to create. Therefore, it is of critical importance that the potential benefit of these resources is maximized. This requires that such datasets are distributed along with powerful open-source software that brings their value to as many researchers as possible.

Multi-atlas labeling (MAL) is the current state of the art for propagating expert labelings from a reference library onto new instances of unlabeled data. The typical procedure for multiple atlas fusion involves common steps regardless of the underlying method. These steps include: (1) registration, (2) label transformation and (3) local statistical prediction models. The end product of this procedure is, for a given new subject, the algorithmic "best guess" at the full anatomical label set across the image. The quality of this guess differs across algorithms and substantial effort has been invested in identifying the state-of-the-art in MAL.

Joint Fusion (JF) [1, 2] is perhaps *the state-of-the-art algorithm* for MAL that is capable of predicting anatomical labels with accuracy that rivals expert anatomists [Yushkevich:2010aa]. JF has proven its effectiveness in lung [3], cardiac data [**???**], the human brain [**???**] and in multiple modality canine MRI [**???**]. This technology has previously been made available only through a research-grade C++ implementation available within Advanced Normalization Tools (ANTs [**???**]). Its inclusion within ANTs is sensible in that the success of JF and related methods depends heavily upon high quality deformable maps that transform the anatomical library image set into the target subject's physical space.

This proposal will bring JF to the ITK and SimpleITK user community. We will also document and promote the models with tutorial material that will illustrate the application of JF to template based segmentation of T1-weighted MRI of the brain. Furthermore, we will support the successful and efficient application of JF to both standard size and large datasets via novel improvements to the existing ITK segmentation and registration frameworks. For segmentation, we will add a complete implementation of (spatial) prior-based statistical segmentation via guassian mixture or non-parametric models. This framework will build upon existing ITK resource and will be extensible with other likelihood models. For registration, we will extend the current framework with more memory efficient image data representations. This approach will build upon existing sub-sampling frameworks within ITK version 4 registration to allow registration to be driven by a sparse image representation which will allow adaptive registration strategies that can be customized for large datasets and will extend the reach of these methods.

- Joint label fusion is a multi-atlas segmentation method.

- It performed well in several recent competitions (SATA 2012, SATA 2013)

- We use it regularly in our studies to build template priors and to label cortical or deep structures in the brain.

JLF theory: "Multi-Atlas Segmentation with Joint Label Fusion"

A matrix M_x is defined by the number of atlas segmentations one has.

$M_x(i, j)$ measures joint atlas errors wrt a target segmentation at a voxel.

Entries in M_x relate to the likelihood two atlases make the same error.

> The key difference between joint label fusion and other label fusion methods is that it explicitly considers correlations among atlases, i.e., the dependence matrix, into voting weight assignment to reduce bias in the atlas set.

JLF theory: "Multi-Atlas Segmentation with Joint Label Fusion"

The expected label difference between the consensus solution obtained from weighted voting and the target segmentation is: $w_x^T M_x w_x$.

Find atlas weights, w_x, for each of A^i atlases, st

$$w_x^T (M_x + \alpha \text{Id}) w_x$$

is minimized subject to $\sum_{i=1}^{n} w_x(i) = 1$.

JLF theory: "Multi-Atlas Segmentation with Joint Label Fusion"

Define

$$K_m = \langle \mid A_N^{i,m} - T_N^m \mid, \mid A_N^{j,m} - T_N^m \mid \rangle$$

then

$$M_x(i, j) = (\sum_m K_m)^{\beta}$$

with N representing a neighborhood calculation, $A^{i,m}$ representing the ith atlas and the mth modality. Lagrange multipliers yield:

$$w_x = \frac{M_x^{-1} 1_n}{1_n^t M_x^{-1} 1_n}$$

Finally, local patch search is used to improve the neighborhoods that correspond.

JLF Example

Suppose that a pair of atlases A_1 and A_2 produce statistically independent label errors for a given target image. If A_1 produces a wrong label 50% of the time and A_2 produces a wrong label 20% of the time, we have

$$M_x = \begin{bmatrix} 0.5 & 0.1 \\ 0.1 & 0.2 \end{bmatrix}$$

The optimal voting weights are then $w_x = [0.2, 0.8]^t$.

Simple ITK Integration: This proposal may appear ambitious given the modest budget. However, our proposal builds upon existing work within ANTs to make these goals achievable. Primarily, we will port the near-ITK quality existing code into the ITK ecosystem for review by other developers via the Gerrit code review system. We are strongly familiar with this system as the team members, particularly lead developer Nicholas Tustison, regularly contribute to ITK. Thus, the effort, here, will follow a natural progression:

- Augment existing ANTs classes that implement prior-based and JF segmentation with ITK-style tests; * Pass these for review to Gerrit; *Implement code refactoring/documentation requested by ITK core;* Implement appropriate wrapping, documentation and metadata for Simple ITK.

Beyond these first steps, we will also (most importatly, perhaps) implement step-by-step tutorial material based on the Simple ITK platform. We feel this is best illustrated with a detailed example of this tutorial material which follows here.

Simple ITK tutorial material for JF-augmented brain mapping

We will employ freely available *real neuroimaging data* on which to base the tutorial material and to promote reproducibility and transparency. We will employ the Pediatric Template of Brain Perfusion (PTBP) at figshare which includes free multiple modality MRI data with demographics and psychometrics. The data is accompanied by an organized csv file with full data available at figshare. We will use a lightly processed version of the data to make examples quick enough for tutorial material.

We will use the sample PTBP subjects to:

- Build a template based on deformable registration done in Simple ITK

- Construct template priors with Joint Fusion and based on freely available labeled data

- Normalize and segment the population data based on spatial prior-based gaussian mixture modeling.

This will all be put together to create a reproducible analysis for a subset of the PTBP. This material will be invaluable for better introducing the neuroscience and larger biomedical image analysis community to this broadly applicable and powerful quantification technology.

Relationship and benefits of the project to the SimpleITK/ITK effort

Registration and segmentation are complementary tools that have often existed along independent software development paths. Registration teams rarely intersected with segmentation teams. Joint Fusion and related methods, however, serve as a meta-algorithm which uses components of both classic registration and classic segmentation to improve upon the results obtained by either independently. Making these bleeding edge algorithms available to the ITK and Simple-ITK community will break down barriers between "code aware" and "code naive" user bases further allowing the computational scientist to communicate effectively with biological and medical scientists.

Advantages that the proposed work derives from SimpleITK/ITK

We are excited about the opportunity to port our current JF and Atropos segmentation framework into the core of ITK. Our prior experience has shown that the quality and lifespan of such algorithms is greatly improved when they are merged within the ITK ecosystem. The deep testing, consistent use of valgrind to find memory defects and cross-platform CMake-based ctests are all integral to the further improvement of the C++ backbone of these methods.

How the proposed work differs from or relates to existing work in ITK and its related software

There is a substantial level-set framework within ITK. There also exists a relatively under utilized statistical framework that allows methods such as k-means clustering followed by Markov random field regularization to be implemented. However, both of these frameworks lack the ability to implement truly Bayesian statistical models that incorporate spatial probability maps and maximize the posterior probability of a segmentation map explicitly. Consider the equation implemented by Atropos. **FIXME.**

Personnel and resources to be committed to the proposed work

The personnel that will implement this plan has extensive background working on C++. Furthermore, the team has many decades of experience contributing specifically to ITK. Dr. Avants made his first commit to ITK on Tue Apr 9 19:09:13 2002. His most recent commit included a bug fix on Nov 29 11:22:12 2012. Despite his personal contributions to code, Dr. Avants has contributed perhaps the most through his leadership of the ITKv4 development team. Specifically, his team implemented a full refactoring of the ITK registration framework. A few contributions of this framework include thread safety, the ability to implement composite transformations, multi-channel registration, extensibility and addition of cutting edge image similarity metrics for both intensity and point set features. In addition, an early version of GPU-based registration was implemented with OpenCL. Dr. Gee was part of the original ITK development core and was also involved in the ITKv4 registration refactoring, as PI. He also obtained several A2D2 grants in collaboration with Drs. Avants and Tustison. Dr. Yushkevich began developing with ITK in 2003 as part of the first round of translationally focused awards which began the eminently popular ITK-SNAP interactive segmentation software. Dr. Yushkevich is a key contributor to the application of ITK core tools to neuroscience goals through his maintenance and continued development of the elegant and easily accessible ITK-SNAP user interface, which uses ITK underneath.

Dr. Avants and Dr. Tustison also promote ITK through their software Advanced Normalization Tools (ANTs, originating at sourceforge.net on 2008-06-26 and now residing at http://stnava.github.io/ANTs/). ANTs is a systematic framework for quantitative biological image analysis based on the Insight ToolKit. ANTs was first created to rapidly disseminate our latest research to the community of scientists who depend on imaging analytics and to allow them to study different organ systems, species or modalities with the same sound foundation. While originally focused on diffeomorphic image registration, ANTs now incorporates novel and cutting-edge methods for segmentation, feature extraction and, more recently, complete statistical pipelines via ANTsR http://stnava.github.io/ANTsR/. In 2014, there were nearly 2,000 citations to ANTs and the software is cloned, downloaded or otherwise accessed over 100-200 times per week, on average at

github. The sourceforge site hosts a similar number of visits and downloads. ANTsR is accessed on average 50 times per week—a substantial number for a new software. There are also over 500 discussion topics on the ANTs sourceforge community site, nearly 100 topics on the github site and over 50 help-focused emails to the personal addresses of developers. Generally, response time to requests for help is within a few hours with rare occasions taking up to a day or two. All of this effort magnifies the impact of ITK.

For institutional resources committed to this work, please see the business portion of the proposal.

Budget for the work.

We request a total of $100,000 which includes both direct and indirect costs.

Dr. Avants (5%) will manage the conduct of the project and oversee the project's successful completion of milestones and deliverables including testing, software development and construction of tutorial material.

Dr. Yushkevich (2%) is a co-inventor of the original Joint Label Fusion algorithm [**???**]. Dr. Yushkevich will contribute to testing and validating the ITK and SimpleITK implementations of JF.

Dr. Gee (2%) has been involved in the Insight ToolKit since its inception and will contribute his substantial registration expertise and project management skills to the team.

Dr. Tustison (consultant) will be the lead software developer for this project. Dr. Tustison is among the top contributors to the Insight ToolKit as judged by the number of lines of code from each author that have survived and are still intact in the current revision of the Insight ToolKit. He is the 11th overall contributor according to this metric as of September 1, 2015.

1. Wang, H., Suh, J. W., Das, S. R., Pluta, J., Craige, C., and Yushkevich, P. A. "**Multi-Atlas Segmentation with Joint Label Fusion**" *IEEE Trans Pattern Anal Mach Intell* (2012): doi:10.1109/TPAMI.2012.143

2. Wang, H. and Yushkevich, P. A. "**Multi-Atlas Segmentation with Joint Label Fusion and Corrective Learning-an Open Source Implementation**" *Front Neuroinform* 7, (2013): 27. doi:10.3389/fninf.2013.00027

3. Tustison, N. J., Qing, K., Wang, C., Altes, T. A., and Mugler, J. P., 3rd. "**Atlas-Based Estimation of Lung and Lobar Anatomy in Proton MRI**" *Magn Reson Med* (Accepted):