Table 6 (abstract A16). User feedback

User Level	EOU	Time	Notes
Novice 1	8	00:35:25	Required walk through support
Novice 2	8	00:52:55	Required support for basic terminal commands only; then was able to complete independently
Moderate 1	3	00:23:45	Required no support
Moderate 2	4	00:22:10	Required no support
Expert 1	3	00:11:34	Required no support
Expert 2 - DE	3	02:00:00	Getting scripts to run took several minutes but reorganizing data and troubleshooting with freesurfer took significant time
Expert 3 – DE	2	01:15:00	Required walk through support

EOU: Ease of Use score (1–10) 1 = easiest, 10 = hardest. Time: the time it took for the user to setup and learn to use the scripts. DE: User's expertise is with a different computational environment than the one required by the scripts

A17

A cortical surface-based geodesic distance package for Python

Daniel S. Margulies, Marcel Falkiewicz, Julia M. Huntenburg
Max Planck Research Group for Neuroanatomy & Connectivity, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
Correspondence: Daniel S. Margulies (margulies@cbs.mpg.de) – Max Planck Research Group for Neuroanatomy & Connectivity, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
GigaScience 2016, **5(Suppl 1):**A17

Introduction

The human cerebral cortex, whether tracing it through phylogeny or ontogeny, emerges through expansion and progressive differentiation into larger and more diverse areas. While current methodologies address this analytically by characterizing local cortical expansion in the form of surface area [1] several lines of research have proposed that the cortex in fact expands along trajectories from primordial anchor areas [2,3] and furthermore, that the distance along the cortical surface is informative regarding cortical differentiation [4]. We sought to investigate the geometric relationships that arise in the cortex based on expansion from such origin points. Towards this aim, we developed a Python package for measuring the geodesic distance along the cortical surface that restricts shortest paths from passing through nodes of non-cortical areas such as the non-cortical portions of the surface mesh described as the "medial wall'.

Approach

The calculation of geodesic distance along a mesh surface is based in the cumulative distance of the shortest path between two points. The first challenge that arises is the sensitivity of the calculation to the resolution of the mesh: the coarser mesh, the longer the shortest path may be, as the distance becomes progressively less direct. This problem has been previously addressed and subsequently implemented in the Python package gdist [https://pypi.python.org/pypi/gdist/], which calculates the exact geodesic distance along a mesh by subdividing the shortest path until a straight line along the cortex is approximated [5]

The second challenge, for which there was no prefabricated solution, was ensuring that the shortest path only traverses territory within the cortex proper, avoiding shortcuts through non-cortical areas included in the surface mesh — most prominently, the non-cortical portions along the medial wall. Were the shortest paths between two nodes to traverse non-cortical regions, the distance between nodes would be artificially decreased, which would have artifactual impact on the interpretation of results. This concern would be especially relevent to the 'zones analysis' described below, where the boundaries between regions would be altered. It was therefore necessary to remove mesh nodes prior to calculating the exact geodesic, which requires reconstructing the mesh and assigning the respective new node indices for any seed regions-of-interest.

Finally, to facilitate applications to neuroscience research questions, we enabled the loading and visualization of data from commonly used formats such as FreeSurfer and the Human Connectome Project (HCP). A Nipype pipeline for group-level batch processing has also been made available [6]. The pipeline is wrapped in a command-line interface and allows for straightforward distance calculations of entire FreeSurfer-preprocessed datasets. Group-level data are stored as CSV files for each requested mesh resolution, source label and hemisphere, facilitating further statistical analyses.

Results

The resultant package, SurfDist, achieves the aforementioned goals of faciliating the calculation of exact geodesic distance on the cortical surface. We present here the distance measures from the central and calcarine sulci labels on the FreeSurfer native surfaces (Fig. 14b). The distance measure provides a means to parcellate the cortex using the surface geometry. Towards that aim, we also implement a 'zones analysis', which constructs a Voronoi diagram, establishing partitions based on the greater proximity to a set of label nodes (Fig. 14c).

Surface rendering of the results draws from plotting functions as implemented in Nilearn [7] and exclusively relies on the common library matplotlib to minimize dependencies. The visualization applies sensible defaults but can flexibly be adapted to different views, colormaps and thresholds as well as shadowing using a sulcal depth map.

Conclusions

The SurfDist package is designed to enable investigation of intrinisic geometric properties of the cerebral cortex based on geodesic distance measures. Towards the aim of enabling applications specific to neuroimaging-based research question, we have designed the package to facilitate analysis and visualization of geodesic distance metrics using standard cortical surface meshes.

Availability of supporting data

More information about this project can be found at: http://github.com/margulies/surfdist

Competing interests

None.

Author's contributions

DSM, MF, and JMH wrote the software and report.

Acknowledgements

Report from 2015 Brainhack Americas (MX). The authors would like to thank the organizers and attendees of Brainhack MX. The visualization functions were originally developed during the Nilearn coding sprint 2015 in Paris, for which we would also like to thank the organizers and participants of this event.

References

1. Winkler Anderson M, Sabuncu Mert R, Yeo BT Thomas, Fischl Bruce, Greve Douglas N, Kochunov Peter, Nichols Thomas E, Blangero John, Glahn David C. Measuring and comparing brain cortical surface area and other areal quantities. Neuroimage. 2012; 61: 1428–1443.
2. Sanides Friedrich. Comparative architectonics of the neocortex of mammals and their evolutionary interpretation. Ann N Y Acad Sci. 1969; 167: 404–423.
3. Buckner R,y L, Krienen Fenna M. The evolution of distributed association networks in the human brain. Trends Cogn Sci. 2013; 17: 648–665.
4. Wagstyl Konrad, Ronan Lisa, Goodyer Ian M, Fletcher Paul C. Cortical thickness gradients in structural hierarchies. Neuroimage. 2015; 111: 241–250.
5. Mitchell Joseph S B, Mount David M, Papadimitriou Christos H. The Discrete Geodesic Problem. SIAM J Comput. 1987; 16: 647–668.
6. Gorgolewski Krzysztof, Burns Christopher D, Madison Cindee, Clark Dav, Halchenko Yaroslav O. Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in Python. Frontiers in Neuroinformatics. 2011; 5.
7. Abraham Alex,re, Pedregosa Fabian, Eickenberg Michael, Gervais Philippe, Mueller Andreas, Kossaifi Jean, Gramfort Alex,re, Thirion Bertr, Varoquaux Gaël. Machine learning for neuroimaging with scikit-learn. Frontiers in Neuroinformatics. 2014; 8: 1–10.

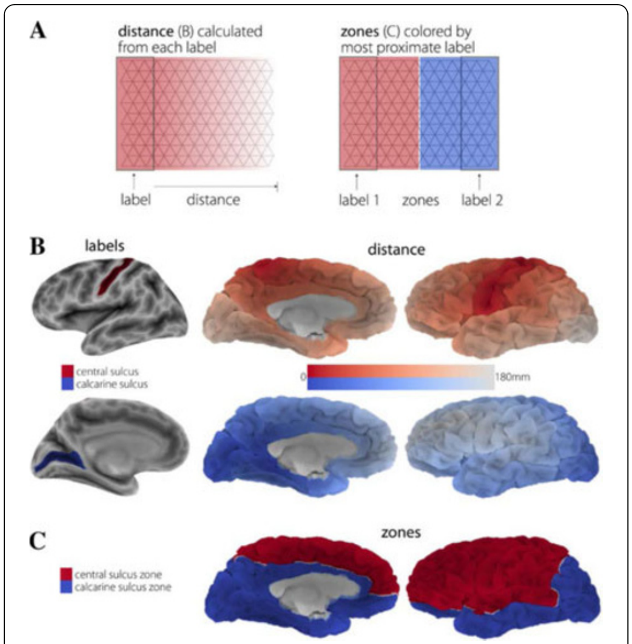

Fig. 14 (abstract A17). a Schematic illustrating the distance (**b**) and zone (**c**) analyses. **b** FreeSurfer labels from the central and calcarine sulci depicted on the individual inflated surface (left), and the exact geodesic distance from the two labels presented on an individual pial surface (right). **c** Zones delineated based on proximity to the central (red) or calcarine (blue) sulci

A18

Sharing data in the cloud

David O'Connor[1,2], Daniel J. Clark[2], Michael P. Milham[1,2], R. Cameron Craddock[1,2]
[1]Center for Biomedical Imaging and Neuromodulation, Nathan Kline Institute for Psychiatric Research, Orangeburg, NY, USA; [2]Center for the Developing Brain, Child Mind Institute, New York, NY, USA
Correspondence: David O'Connor (david.oconnor@childmind.org) – Center for Biomedical Imaging and Neuromodulation, Nathan Kline Institute for Psychiatric Research, Orangeburg, NY, USA
GigaScience 2016, **5(Suppl 1):**A18

Introduction

Cloud computing resources, such as Amazon Web Services (AWS) [http://aws.amazon.com], provide pay-as-you-go access to high-performance computer resources and dependable data storage solutions for performing large scale analyses of neuroimaging data [1]. These are particularly attractive for researchers at small universities and in developing countries who lack the wherewithal to maintain their own high performance computing systems. The objective of this project is to upload data from the 1000 Functional Connectomes Project (FCP) [2] and International Neuroimaging Datasharing Initiatives (INDI) [3] grass-roots data sharing initiatives into a Public S3 Bucket that has been generously provided by AWS. This will make the data more quickly accessible for AWS-based analysis of these data, but will also improve the speed and availability of access to this data for analyses performed outside of the cloud. To begin with, we focused on the following collections:

- The autism brain imaging data exchange *(ABIDE)* consists of structural MRI and resting state functional MRI from 1113 individuals (164 F, 948 M, 6–64 years old, 539 with autism spectrum disorders, 573 typical controls) aggregated from 20 different studies [4]
- The *ADHD-200* contains structural MRI and resting state functional MRI from 973 individuals (352 F, 594 M, 7–21 years old, 362 with attention deficit hyperactivity disorder (ADHD), 585 typically developing controls) collected from 8 sites [5]
- The *Consortium for Reliability and Reproducibility (CoRR)* consists of 3,357 structural MRI, 5,093 resting state fMRI, 1,302 diffusion MRI, and 300 cerebral blood flow scans from 1629 subjects (673 F, 956 M, 6–84 years old, all typical controls) acquired in a variety of test-retest designs at 35 sites [6]
- The *Enhanced Nathan Kline Institute - Rockland Sample (ENKI-RS)* consists of structural MRI, resting state functional MRI, diffusion MRI, cerebral blood flow, and a variety of task functional MRI scans and deep phenotyping on over 700 participants from across the lifespan and a variety of phenotypes acquired at a single site [7] The acquisition of this collection is ongoing.
- The *Addiction Connectome Preprocessed Initiative (ACPI)* [http://fcon_1000.projects.nitrc.org/indi/ACPI/html/index.html] consists of 216 structural MRI and 252 functional MRI from 192 subjects (44 F, 148 M, 18–50 years old) from three datasets generated by NIDA investigators.

Approach

Data for the ADHD-200, ABIDE, CoRR, and Rockland Sample data collections are currently downloadable from NITRC [http://fcon_1000.projects.nitrc.org/] as a series of large (>2GB) tar files. The process of uploading the data involved downloading and extracting the data from these tar files, organizing the individual images to the standardized INDI format [http://fcon_1000.projects.nitrc.org/indi/indi_data_contribution_guide.pdf] and then uploading the data to S3. We developed a S3 upload script in python using the Boto AWS software development kit [https://aws.amazon.com/sdk-for-python/] to facilitate this process. We also developed a download script in python that provides basic query functionality for selecting the data to download from a spreadsheet describing the data.

Results

The entirety of the CoRR, ABIDE, ACPI, and ADHD-200 data collections and ENKIRS data for 427 individuals were uploaded during the OHBM Hackathon event. The data are available as individual files to make it easily indexable by database infrastructures such as COINS [8] LORIS [9] and others. Additionally, this makes it easy for the users to download just the data that they want. The data in the bucket can be browsed and downloaded using a GUI based S3 file transfer software such as Cyberduck [http://cyberduck.io] (see Fig. 14), or using the Boto Python library [https://github.com/FCP-INDI/INDI-Tools]. One can connect to the bucket using the configuration shown in Fig. 15. The data is structured as follows: bucketname/data/Projects/ProjectName/DataType. For example you can access raw data from the ENKI-RS, as shown in Fig. 15, by specifying the following path in CyberDuck: https://s3.amazon.com/fcp-indi/data/Projects/RocklandSample/RawData

Conclusions

Uploading data shared through the FCP and INDI initiatives improves its accessibility for cloud-based and local computation. Future efforts for this project will include uploading the remainder of the FCP and INDI data and organizing the data in the new brain imaging data structure (BIDS) format [10].

Availability of supporting data

More information about this project can be found at: https://github.com/DaveOC90/INDI-Organization-Scripts

Competing interests

None.

Author's contributions

DO performed quality control, and uploaded the data. DJC wrote code to interact with AWS, preprocessed and uploaded data. MPM and RCC lead the data collection and sharing projects. All of the authors contributed to writing the project report.

Acknowledgements

Report from 2015 OHBM Hackathon (HI). The authors would like to thank the organizers and attendees of the OHBM Brainhack in Hawaii. This project was made possible by the S3 public bucket generously provided by Amazon Web Services.