Published in final edited form as:

Proc IEEE Symp Vis Anal Sci Technol. 2010 December 10; 2010(25-26 Oct. 2010): 237. doi:10.1109/VAST.2010.5652532.

Large-scale Neuroanatomical Visualization Using a Manifold Embedding Approach

Shantanu H. Joshi, Ian Bowman, and John Darrell Van Horn

Laboratory of Neuro Imaging, Department of Neurology, UCLA School of Medicine, University of California Los Angeles, CA 90034

Shantanu H. Joshi: sjoshi@loni.ucla.edu; Ian Bowman: ian.bowman@loni.ucla.edu; John Darrell Van Horn: jvanhorn@loni.ucla.edu

Abstract

We present a unified framework for data processing, mining and interactive visualization of large-scale neuroanatomical databases. The input data is assumed to lie in a specific atlas space, or simply exist as a separate collection. Users can specify their own atlas for comparative analyses. The original data exist as MRI images in standard formats. It is uploaded to a remote server and processed offline by a parallelized pipeline workflow. This workflow transforms the data to represent it as both volumetric and triangular mesh cortical surfaces. We use multiresolution representations to scale complexity to data storage availability as well as graphical processing performance. Our workflow implements predefined metrics for clustering and classification, and data projection schemes to aid in visualization. Additionally the system provides a visual query interface for performing selection requests based on user-defined search criteria.

Index Terms

I.3 [Computer Graphics]: Three-Dimensional Graphics and Realism; I.3.3 [Viewing Algorithms]; I.3.8 [Applications]

1 Introduction

Neuroimaging repositories continue to grow rapidly. A large number of such databases reside in established research sites both nationally and internationally. The interface to such systems is usually in the form of a textual or web-based query, providing sequential access to collections of data. Such serialized lookup does not facilitate a comprehensive view of the data, and thus may present incomplete information to the user. As typical for such systems, the user has no choice but to cycle through each and every object in the population in order to get a better understanding of the data before using it for analysis. With this background in mind, we claim that there is a pressing need for a graphics-based, informatics approach for interacting with neuroimaging data. Our goal is to enable large-scale informatics, mining, and visualization of the existing neuroimaging data repositories by developing streamlined data processing work-flows and visualizing the results in an interactive application.

Recently, Herskovits et al. [2] have developed an open source database system with data-mining capabilities for managing, querying, analyzing and visualizing brain-MR images (BRAID). However, their focus is on manipulation, visualization, and analysis of individual images. Our approach is different in that we construct compact representations of brain features by their cortical surfaces, and present the user with interfaces for graphically navigating, browsing, and querying *aggregates* of such repositories. To our knowledge,

there is currently no effort in the neuroinformatics or even the visualization community in this direction.

2 Methods

We adopt a three-stage approach for constructing the proposed visual informatics framework: i) the input-processing stage consisting of feature extraction and representation, ii) the data-analytics stage consisting of modeling, regression, and clustering, and finally iii) the visualization stage that gives unrestricted, multi-faceted, 3D navigable and selectable views of the neuroimaging data. The input data comprises of structural neuroimaging acquisitions that are generated from 3D MRI scanners. The primary source of visualization is a surface geometry representation of the cortex, obtained after pre-processing, skull removal, segmentation, cortical extraction, and topology correction [4]. After preprocessing, we calculate several features on cortical surfaces, e.g. shape foldedness index, principal curvatures, and curvedness. Additionally we incorporate several volumetric features such as cortical thickness and gray matter volumes, thereby defining a feature n -dimensional vector at each point on each cortex. Finally, for visualization purposes we project distances between feature vectors in the Euclidean space after approximation, using multi-dimensional scaling. The visualization environment is a dynamic view of the database, enabling the user to navigate, discover, and verify the brain surface geometry simultaneously in relation to its neighbors. Each brain surface is accompanied by an XML description of its meta-data that the user can optionally display.

2.1 Extracting Local and Global Features

In order to conduct a large-scale discriminative analysis on a population of brains, we extract pertinent local and global features which serve as compact representations of underlying dataset patterns. The local features are given by local curvatures (Gaussian or Mean), shape index, and curvedness, whereas the global features are represented by areas and volumes. Furthermore, in order to improve the resolution, we parcellate an individual cortex into 56 different labeled regions. [6] For each region we compute local spatial values and their averages. As an example, Figure 1 shows a cortical surface overlaid with the local shape-index, curvedness, Gaussian curvature, and mean curvature calculated at each vertex.

2.2 Manifold Embedding of Brain Association Graphs

We derive previously defined local and global features on the surface of the cortex, and then represent these quantities in the space of real numbers and adopt the standard Euclidean metric. As a specific example, for a database with N subjects and L delineated sub-volumes, given by $\{V_k^i\}, i=1, \dots, N, k=1, \dots, L$ we first normalize all the volumes to have a unit scale. Then given a set of features $\{S_i\}, i=1, \dots, N$, we calculate a N -by- N dissimilarity matrix given by,

$$D(S_i, S_j) = \frac{\sum_{k=1}^L (V_k^i - V_k^j)^2}{L} \quad (i, j=1 \dots N, i \neq j) \quad (1)$$

In order to explore the dissimilarities between brain volumes, we need to project the dissimilarity matrix into an appropriate 2D or 3D space. There are numerous techniques to project high dimensional data into lower dimensional spaces for analysis or visualization. As discussed above, one could calculate principal coefficients, or principal factors explaining the maximum observed population variability in terms of a few determining factors. For visualization purposes, only the first 3 eigen projections are needed to display objects in a 3D space. Since the construction of our embedding relies on dissimilarities among neuro-

structures, we will use the multidimensional scaling [3] approach for projecting the dissimilarity matrix into a 3D space. Multidimensional scaling is an optimization technique that projects a high dimensional dissimilarity matrix into a low dimensional space that most accurately represents the pairwise distances between the objects. Furthermore we also attempt to learn the inherent low-dimensional structure that was not exposed using MDS. For this purpose we use the Isomap algorithm [5] that achieves non-linear dimension reduction by performing classical MDS on an adjacency matrix which approximates the geodesic distance between data points. We then render this reduced-dimensional manifold representation of the whole data and use it as the basis for user interaction.

2.3 Interactive Visualization of Neuroimaging Repositories

The main challenge for real-time visualization of such large data is the long-term storage and memory requirements which multiply progressively according to data volume. Real time visualization of large datasets presents numerous difficulties with regards to limited processing power and computer memory. For this reason, it is necessary to represent the data in multiple resolutions. For our visualization interface, we have implemented the quadric error mesh simplification strategy [1] that keeps on contracting edges defined by vertex-pairs until the desired number of faces are achieved. Multi-resolution representation and rendering enables faster response times, and facilitates better user interaction. Currently the surface geometry is stored as triangular meshes with faces, vertices, and colors. We also allow surfaces to be colored according to vertex attributes. These attributes measure features such as cortical thickness, and functional activity. Figure 2 shows a snapshot of this interactive visualization environment that was developed in C++ and OpenGL.

3 Discussion and Conclusion

The proposed architecture is decentralized in nature, with storage, computing and visualization shared across heterogeneous platforms and physical locations. The novelty in our approach lies in the user ability to simultaneously view and interact with one or more atlases, and utilize off the shelf discriminatory analysis and visualization capabilities. The problem of manifold learning by itself is complicated, and unlike most synthetic examples, we do not have ground truth for evaluating our results. Additionally, in the case of human brains there is are vast local and global features that describe underlying patterns. As the next step in neuroimaging informatics, we anticipate the development of graphical visualization tools which enhance scientific interaction with large-scale databases. Though some basic image viewing tools exist, we argue the need for a next generation visual interaction framework. We also demonstrated a content-based solution that can be applied to any such archive in order for researchers to more easily examine dissimilarity between brains, and dynamically visualize their degree of proximity. This may further be indicative of the demographic and clinical attributes of the data themselves. In fact, throughout our approach we have made as few assumptions about the data as possible, allowing the data to segregate itself based upon regional shape and geometry.

References

1. Garland, M.; Heckbert, P. SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co; 1997. Surface simplification using quadric error metrics; p. 209-216.
2. Herskovits EH, Chen R. Integrating data-mining support into a brain-image database using open-source components. *Adv Med Sci* 2008;53(2):172–81. [PubMed: 18467275]
3. Kruskal, JB.; Wish, M. Multidimensional scaling. Beverly Hills: Sage Publications; 1978.
4. Shattuck DW, Leahy RM. Brainsuite: an automated cortical surface identification tool. *Med Image Anal* 2002;6(2):129–142. [PubMed: 12045000]

5. Silva VD, Tenenbaum JB. Global versus local methods in nonlinear dimensionality reduction. *Neural Information Processing Systems 2002*;16:721–728.
6. Tu Z, Narr KL, Dollar P, Dinov I, Thompson PM, Toga AW. Brain anatomical structure segmentation by hybrid discriminative/generative models. *IEEE Trans Med Imaging* 2008;27(4):495–508. [PubMed: 18390346]

Figure 1.
From left, a cortical surface colored with shape-index, curvedness, Gaussian curvatures, and mean curvatures respectively.

Figure 2.

A screenshot of the 3D Visualization environment for the neuroimaging repositories that displays a large collection (N=400) of brain surfaces along with their meta-data.