Published in final edited form as:

Brain Imaging Behav. 2014 June ; 8(2): 300–310. doi:10.1007/s11682-013-9273-9.

Graphical Neuroimaging Informatics: Application to Alzheimer's Disease

John Darrell Van Horn¹, Ian Bowman¹, Shantanu H. Joshi², and Vaughan Greer¹

John Darrell Van Horn: jvanhorn@usc.edu

¹The Institute for Neuroimaging and Informatics, Keck School of Medicine, University of Southern California, 2001 North Soto Street – SSB1-102, Los Angeles, CA 90032, Phone: (323) 442-7246

²Department of Neurology, UCLA School of Medicine, 635 Charles E. Young Drive, NRB #225, Los Angeles, 90095

Abstract

The Informatics Visualization for Neuroimaging (INVIZIAN) framework allows one to graphically display image and meta-data information from sizeable collections of neuroimaging data as a whole using a dynamic and compelling user interface. Users can fluidly interact with an entire collection of cortical surfaces using only their mouse. In addition, users can cluster and group brains according in multiple ways for subsequent comparison using graphical data mining tools. In this article, we illustrate the utility of INVIZIAN for simultaneous exploration and mining a large collection of extracted cortical surface data arising in clinical neuroimaging studies of patients with Alzheimer's Disease, mild cognitive impairment, as well as healthy control subjects. Alzheimer's Disease is particularly interesting due to the wide-spread effects on cortical architecture and alterations of volume in specific brain areas associated with memory. We demonstrate INVIZIAN's ability to render multiple brain surfaces from multiple diagnostic groups of subjects, showcase the interactivity of the system, and showcase how INVIZIAN can be employed to generate hypotheses about the collection of data which would be suitable for direct access to the underlying raw data and subsequent formal statistical analysis. Specifically, we use INVIZIAN show how cortical thickness and hippocampal volume differences between group are evident even in the absence of more formal hypothesis testing. In the context of neurological diseases linked to brain aging such as AD, INVIZIAN provides a unique means for considering the entirety of whole brain datasets, look for interesting relationships among them, and thereby derive new ideas for further research and study.

Introduction

Recent advancements in imaging protocols combined with a reduction in storage costs have led to an upsurge of neuroimaging data in both clinical as well as research settings.

Increasingly, neuroimaging databases are capable of containing image volumes in excess of hundreds or thousands of subjects (Van Horn, Wolfe et al. 2005, Van Horn and Toga 2009, Biswal, Mennes et al. 2010). Moreover, brain databases are often limited to text-based

metadata searches of their contents thus limiting the user interaction considerably. Once the complete set of image data files has been downloaded, a user is obliged to conduct a formal analysis of the data simply to discover if the data contain any particular effects of interest. Finally, many commonly available visualization tools ideally designed for neuroimaging focus on single subject data and are not conducive to plotting such multi-subject relationships. In the context of data mining and exploratory inspection of database content, this process is inefficient and time consuming.

Many neuroscience database uses follow a typical processing and analytical approach which begins with a text-based search for relevant data and then downloading the raw imaging files for analysis. The commonly employed neuroimaging processing framework involves the fitting of the imaging data from each subject to a common spatial frame of reference in the form of a standardized brain atlas (Collins, Neelin et al. 1994, Nowinski and Belov 2003). The individual brain volumes are thereby subjected to warping with respect to specific template brain volume in order to remove the anatomical differences in brain structure (Lancaster, Fox et al. 1999, Nowinski and Thirunavuukarasuu 2001), permitting population-level averaging (Van Horn and Toga 2009), and comparisons of anatomy between groups of subjects irrespective of overall brain size (Mega, Dinov et al. 2005). Final steps involve the specification of an experimental design and inferential statistical modeling, used to assess morphological differences according to phenotypic observations of interest, such as cortical thickness in relation to patient diagnosis (Thompson, Mega et al. 2001).

Such analysis-driven techniques can provide useful average summaries of a collection of scans from a database (Van Horn and Toga 2009). However, these mappings often do not allow a user to view the structure and variation of multiple data sets across different subject-based categories. For instance, there are hippocampal volume deficits (Carmichael, Aizenstein et al. 2005) as well as white matter alterations (Carmichael, Schwarz et al. 2010) in the case of Alzheimer's disease which may correlate with disease duration, severity, genetic susceptibility, and other factors (Stein, Medland et al. 2012).

While correlational analyses representing this result could be obtained by merging data from multiple subjects into an atlas space, extracting the hippocampus, and conducting a formal statistical analysis, this would be time consuming, computationally expensive, and might prohibit independent visualization of individual subjects with respect to one another, per se. Consequently, such an approach tends to prohibit or severely limit a more informal or exploratory examination of database contents. In other words, data mining or exploratory analyses are not easily performed and a user must commit to a more formal image processing and statistical analysis. As a means for browsing database contents, this approach is costly and counter to the notion of exploring databases from a content-driven point of view.

Thus, to rapidly understand and explore the image-centric data contained within these repositories can often become both computationally demanding as well as operationally overwhelming. What is needed is a means by which users may explore relationships in these archives through interaction with the data themselves in contrast to only text-based searches and then inspection of data one subject at a time. Methods for data mining and exploratory

analysis have been proposed (Megalooikonomou, Ford et al. 2000) and would seem ideally suited for examining database contents. For instance, in the context of neuroanatomical imaging data, it would greatly assist database users if they could directly view representations of the subject's data in an interactive way, discover the basic relationships between neuroanatomical features and subject metadata, as a prelude to downloading the data for a formal statistical examination. Dimension reduction, graphical display, and data-mining approaches form a useful combination of tools which, when used collectively, can be employed to interactively examine potentially large amounts of neuroimaging data at once. This has clear advantages for database interactivity, the discovery of interesting effects present therein, and hypothesis generation (Mennes, Biswal et al. 2012, Voytek and Voytek 2012).

We recognized that the ability to display multiple brain data sets simultaneously represents an essential step for data exploration and mining, allowing users to examine patterns of similar brain anatomy in advance of subsequent processing, atlas fitting, averaging, and computational analysis (Joshi, Van Horn et al. 2009). Through the process of exploring the similarity space of brain anatomy from across a large collection of preprocessed MR data, a user might be ideally positioned to identify interesting patterns in the data which would not be appreciated by averaging or by inspection of the metadata alone.

With this idea in mind, we developed a novel framework for the interactive discovery of brain morphological feature and meta-data attribute relationships drawn from a large-scale archive of structural MRI data. The goals of our work in the area are as follows: 1) To demonstrate a visually compelling interface for displaying multiple cortical surfaces computed from a repository of raw brain imaging data in which users may dynamically interact with the complete collection of surfaces and rapidly examine feature relatedness; 2) Illustrate INVIZIAN's text query-based framework which provides for the immediate interactive investigation of how subject metadata varies with regard to extracted morphological metrics; and 3) showcase graphical meta-analytics for visualizing the basic relationships between morphological metric values and subject metadata which may help to set the stage for subsequent formal statistical modeling.

In what follows, we illustrate the use of the Informatics Visualization for Neuroimaging (INVIZIAN) framework (Bowman, Joshi et al. 2012) in the use-case context of neuroimaging data relevant to human brain aging and Alzheimer's Disease. The INVIZIAN software can be accessed via our project website, <http://invizian.loni.usc.edu>, with easy-to-use installation packages available to deploy the software on PC and Linux platforms, and links to the source code repository (<http://bitbucket.org/uscini/invizian>). INVIZIAN is also represented on the Neuroimaging Tools and Resources Clearinghouse (NITRC) website (<http://www.nitrc.org/projects/invizian>).

Background and Related Work

Graphical visualization, dimension reduction, and query based database searches are widely available individually but have not been used together under a single framework previously.

In this section we review several of these elemental approaches for data exploration noting their strengths and weaknesses:

Visualization in Neuroscience

The prevailing graphical brain imaging tools such as FreeSurfer (Dale, Fischl et al. 1999), BrainVoyager (Goebel, Esposito et al. 2006), and BrainSuite (Shattuck and Leahy 2002) are specifically designed to display brain surface models from a single subject at a time. This is because their focus is on analysis and visualization of subject-specific cortical data along with the feature values such as thickness, volume, curvature etc. In these approaches, while important for processing and visualization of brain surface-based results at the individual subject level, the examination of multiple subjects at a time is not possible.

Dimensionality Reduction

Dimension reduction is widely used as a primary step to reduce the overall complexity of large volumes of data while minimizing information-loss. There are a number of well-known approaches. Principal Component Analysis (PCA) (Rencher 2002) transforms the data to a reduced dimensional representation using a set of orthogonal vectors which seek to explain variance components. Multidimensional scaling (MDS) (Kruskal and Wish 1978) is a strategy aiming to replicate the similarity between data in native representation using a lower dimensional representation. Kohonen's Self Organizing Maps (SOM) (Kohonen 1998) are a form of artificial neural network that produce a topology-preserving two-dimensional representation of the input data. While dimension reduction successfully reduces structural complexity of the input, in neuroimaging the relationship between the restructured output and original data is sometimes unclear.

Prior work has implemented an interactive component on top of the dimension reduction to improve the focus and intuition of the calculations. An MDS system is proposed in (Williams and Munzner 2004) that allows a user to iteratively reduce the dimensionality of data, while targeting reduction computation toward regions of interest. Visual Hierarchical Dimension Reduction (VHDR) (Yang, Ward et al. 2003) constructs and arranges dimensions into a hierarchy, presented as a radial interface facilitating user exploration of dimensional configuration. Dimension Ordering Spacing and Filtering (DOSFA) (Yang, Peng et al. 2003) extends VHDR by adding automated components that attempt to identify lower dimensional subspace structure. Johansson and Johansson (2009) propose an interactive dimension reduction system driven by user-defined quality metrics. Any dimension reduction approach can lead to information loss, but in this interactive system the user is in effect able to decide what dimensions are important by appropriately weighting correlation, clustering and outlier significance. Once the user assigns weights and a dimensionality threshold, the system provides a graph illustrating information loss per variable reduced. In another compelling recent paper Endert et al. (2011) consider a means for users to graphically manipulate various dimension reduction model parameters through clicking and dragging of a two dimensional data plot. The users of the observation-level system operate entirely with the spatial interface and do not need to modify numerical parameters directly to alter the visualization. In addition to dimension reduction, visualization frameworks such as

Ggobi (Cook and Swayne 2007) provides the ability to link associated graphical displays together using a coordinated multiple view approach.

Query-Based Search

A query-based search system can be thought of as one facilitating guided discovery of significant data values through user query-result interaction. This can be performed through text-based phrases or via a user interface for selecting various criteria upon which to search. Results from such searches can be displayed as a list or table, but more interestingly as a graphical rendering. One approach for search visualization was pioneered by Keim and Kriegel (1994) in their method entitled VisDB for building a multidimensional view of a relational database repository. The authors defined type-specific distance metrics which were employed against all data to rank relevance to each user query. Feedback was then presented as a pixel map, with a carefully chosen color scheme depicting distance of query result from significant near matches. In a related approach, the program called Scout (McCormick, Inman et al. 2004), used GPU hardware acceleration to enable the guidance of visualization processing via mathematical evaluation. Users interacted with the system through a customizable API which manipulated pixel parameters during visualization to identify statistically interesting combined parameter values.

INVIZIAN System Architecture

In designing the INVIZIAN program, we sought to develop a graphically compelling system in which users could easily visualize a set of cortical surfaces as an entire collection and interact with them in a fluid manner using only their computer mouse (Bowman, Joshi et al. 2012). By combining tools for viewing relationships between clusters of similarly featured cortical surfaces, enabling searches over those data, permitting grouping and meta-analytic assessment, the INVIZIAN interface is specifically developed for users to explore and discover relationships that might not be appreciated in a strictly text-based search as a prelude to more formal statistical treatment. Using high-density surface meshes in an OpenGL-based (<http://www.opengl.org>) interface, users may zoom in, rotate, and examine in detail the regional cortical patterns of within the clusters of similar brains. A summary of the INVIZIAN visual analytics system is illustrated in Figure 1 and briefly consists of the following basic elements:

Feature Analytics, Extraction, and Representation

Image pre-processing for T1-weighted anatomical volumes and subsequent feature extraction takes advantage of FreeSurfer's ReconAll cortical surface reconstruction and cortical parcellation algorithm (Dale, Fischl et al. 1999, Fischl, Sereno et al. 1999). These steps are implemented in the LONI Pipeline workflow design and execution environment (Dinov, Van Horn et al. 2010) to take advantage of efficient parallel computing of surfaces and other results. Feature metrics, such as thickness, volume, etc are extracted for each brain parcellation. Such features have previously served as useful biomarkers when evaluating the effect of disease (Narr, Bilder et al. 2005, Lerch, Pruessner et al. 2008), but are also recently being investigated as potential genetic phenotypes (Winkler, Kochunov et al. 2010) that influence heritability of the brain. As described previously (Joshi, Bowman et al. 2011),

using these regional feature vectors, subjects are systematically compared to all other subjects in the collection using a probabilistic “distance” metric, and then this set of pairwise comparisons is decomposed using multi-dimensional scaling (MDS) to reduce its dimensionality and create a new set of coordinates in which brains which are most similar are positioned close to one another while those most dissimilar are positioned furthest from each other. These new coordinates form the basis for the simultaneous rendering of the entire set of extracted brain surfaces in the INVIZIAN user interface (Bowman, Joshi et al. 2012).

Multiresolution Brain Surface Attribute Mapping

Fluid interaction with the displayed cortical surfaces is an important design characteristic of INVIZIAN which employs a multi-resolution technique to facilitate interactivity when viewing data. This has several advantages for when displaying multiple brain surfaces at once. For instance, a fully zoomed out view of the entire dataset collection loads surface meshes decimated by 95 percent. The triangle count of such data is approximately 30k, and the disk size of such a dataset is slightly less than 1 MB. However, when INVIZIAN determines that the user is viewing less than six surfaces in the field of view it seamlessly replaces them with meshes that have been decimated by 80 percent. By and large, these decimated cortical surface models tend to be visually indistinguishable from the non-decimated versions, but afford considerable reduction in load-times by having triangle counts of around 100k at ~3.5 MB in size while providing a finer resolution surface for the user to inspect. When a single brain surface is selected, it is seamlessly replaced by its full resolution version, thereby enabling detailed inspection of its surface characteristics. Additional advantage is taken of GPU-based processing to interpolate and color the brain surfaces to present an aesthetically compelling experience for the user.

Query-Based Surface Identification, Clustering, and Coordinated Viewing

Within INVIZIAN, a simple metadata dialogue is available if the user clicks on a surface with their mouse (Figure 3). The dialogue box lists patient data at the time the subject scan was conducted and typical information includes age, sex, weight, disease diagnosis information, as well as MRI scanning parameters. In addition to simply viewing the cortical neuroanatomy and meta-data for individual subjects, a user might be interested in how cortical feature values relate to clinical subject metadata. The user enters expressions such as “Subject Age > 70” into the query text box and INVIZIAN renders a bubble around matching subjects (Figure 4, top). Clustered matches indicate potential correlations between subject metadata and the anatomical feature similarity. The user can then have INVIZIAN place a circular glyph around the surfaces to indicate group membership (Figure 4, bottom) and store these groupings for later recall.

INVIZIAN also provides a “top-down” or coordinated viewing window for the entire set of data which is linked to user mouse movements. When the user pans, zooms, or rotates, the projection renders a rectangle on top of the dot plot, indicating the region of display currently occupying the view. The coordinated projection frees the user of either memorizing subject placements, for having to continually zoom out to reorient the view, or to still appreciate subject-to-subject relationships. Examples of coordinated projection sub-

plots can be seen in Figures 6-8. An example data set for use in exploring INVIZIAN is available by contacting the corresponding author.

Use Case: Graphical Informatics Applied to Examining Alzheimer's Disease

In order to illustrate how the INVIZIAN framework can be utilized to conduct exploratory analyses on neuroimaging data, we describe an example of the interaction with neuroimaging data from patients diagnosed with Alzheimer's Disease (AD), those having Mild Cognitive Impairment (MCI), and an otherwise set of older healthy subjects. We specifically chose this demonstration due to the well-known observation that AD has significant alterations in brain cortical morphometry, in particular, the atrophic effects of AD on hippocampal anatomy (Jack, Bernstein et al. 2008).

Subjects

The neuroimaging datasets used consisted of a set of 120 MRI volumetric images (MPRAGE) from the Alzheimer's Disease Neuroimaging Initiative (ADNI; <http://www.adni-info.org/>) cohort maintained at the Laboratory of Neuro Imaging (LONI) at the University of California Los Angeles. Data belonging to 40 Alzheimer's disease (AD) patients, 40 mild cognitively impaired (MCI) subjects, and 40 healthy, otherwise normal control (NC) subjects were identified. Each group had a mean age of 76 years.

Data Pre-Processing

All the MRI images underwent skull stripping and further processing for cortical surface extraction using FreeSurfer (<http://surfer.nmr.mgh.harvard.edu>) via a LONI Pipeline (Dinov, Van Horn et al. 2009) data processing workflow from which the cortex was parcellated into 34 regions, and volumetric feature values such as gray matter thickness, volume and surface area were determined. Using these metrics as regional feature vectors, we computed pair-wise probabilistic difference measurements - the Jensen-Shannon divergence - between each pair of concordant regions, integrated over all region-pairs, to compute subject-by-subject "distances" (Joshi, Bowman et al. 2011). The Jensen-Shannon divergence is an information theoretical metric of (dis)similarity which has been used in medical image processing (Iglesias, Konukoglu et al. 2011, Lu, Zheng et al. 2012), biomarker discovery for cancer (Berretta and Moscato 2010), genetics (Chen, Guo et al. 2013), as well as text mining (Sasahara, Hirata et al. 2013). It has been used successfully in morphological analysis of neuroimaging data (Guo, Rangarajan et al. 2005). However, other measures of relative distance between subjects would be possible to reflect different aspects of data (dis)similarity suitable for various pattern recognition techniques (see for instance, Kuriakose, Ghosh et al. 2004, Xu, Thompson et al. 2006).

Once this matrix of NxN subject-by-subject distances was obtained, we utilized classical MDS (Kruskal and Wish 1978, Chen, Härdle et al. 2008) to project the data in a lower dimensional space and determination of new subject-specific coordinates as discussed above. Each subject's brain surface was then positioned in the INVIZIAN interface at these newly derived coordinates. Incidentally, to enable other spatial arrangements of brain surfaces, any set of 2D or 3D sets of coordinates can be utilized for surface placement within

INVIZIAN permitting planar, scatter cloud, or in any other configuration (e.g. helical, spherical, etc) depending on the interests of the researcher.

Visualization of Diagnostic Trends using Cortical Features

In the following subsections we illustrate several examples for visualizing differences between AD patients and NC subjects and making predictive inferences that might merit further examination of the data under more formal analytic frameworks.

Cortical Gray Matter Thickness

Gray matter thickness was measured as the mean distance from the gray matter/cerebrospinal fluid interface to the gray/white matter surface, and vice versa. Figure 6 (left) shows the color map subdivided into quintiles used to map colors to the (0 to 5mm) thickness valued features on the cortical surface. Figure 6 (middle) shows all 120 subjects colored according to thickness. Simple visual examination indicates that there is a general trend of colors from more blue on the left, to more red and yellow on the right. This describes the subjects having large thickness deficits on the left that gradually decrease toward the right. Using the query input dialogue box, it can be quickly seen that that the subjects towards the left are the AD patients, while the subjects on the right are NC subjects. Utilizing the grouping functionality of INVIZIAN, Figure 6 (right) shows the subjects grouped according to diagnosis meta-data value, with red glyphs representing patients with an AD value, and blue glyphs representing subjects with an HC value.

Cortical Gray Matter Volume

Gray matter volume was defined as the product of the thickness and the area of the surface layer midway between the gray/CSF and white/gray matter boundaries. Following identical steps discussed above for the rendering and display of cortical thickness we find a similar distribution of the subjects based on cortical volume. Figure 7 (left) illustrates the color mapping of volumetric feature values on the brain surfaces. Based upon the feature distribution and the query and grouping tool, Figure 7 (middle) shows the subjects grouped according to the categories AD (red glyphs) and NC (blue glyphs). However, from observing the group clustering, it can be noted that many subjects with differing diagnosis are in overlapping clusters in contrast to the clearer separation seen when considering the distributions of cortical thickness. This clustering hints that there is a slightly decreased discriminatory capability of gray matter volume as a feature as compared to cortical thickness.

Cortical Gray Matter Surface Area

Lastly, an average surface area measurement was calculated at each point on the cortical surface mesh through a processing of averaging over the areas of those adjacent triangles including that point. Displaying these color mapped brain surfaces illustrates a further reduced discriminating ability as illustrated in Figure 8. This is also evident observing the color-coded distribution of the area on the individual cortical surfaces. Little variation of surface area is present locally on the individual cortical surfaces as well as globally throughout the population.

Visualization of Local Structural Features

The diagnostic trend findings illustrated above focus on global feature patterns in the population. To gain a better understanding of the local structural changes due to the disease, we use the coordinated projection to zoom in on a neighborhood of brains that share similar feature characteristics. Figure 9 (left) shows a screenshot of the zoomed-in view of the brain surfaces from AD patients (red dots in the coordinated projection plot), while Figure 9 (right) is a zoomed-in view of the NC subjects (blue dots in the coordinated projection plot). From the two side-by-side views one can compare the detailed cortical structure and immediately appreciate the localized thickness deficits along the cortex. In the case of AD patients, there is a noticeable decrease in cortical thickness overall, and particularly in the frontal, temporal, and some portions of the parietal cortex.

Implications of Results

From visualizing this collection of cortical surface data in INVIZIAN, a user can rapidly examine the properties of the data itself to derive potentially valuable neuroscientific inferences. In regard to the use-case example illustrated here, we list them succinctly as follows:

- Cortical thickness is likely to be the best predictor of Alzheimer's Disease in the observed population sample.
- Although thickness and volume are proportionally related in the cortex, the spatial variation of thickness appears to follow a different distribution compared to regional volume. This may be due to the fact that regional boundaries vary widely from subject to subject.
- Consequently, volume has a lower predictive effect for disease compared to thickness.
- Surface area values do not provide a trend one way or the other with regard to AD.
- While it is understood that the human cortex is thinner in the granular post-central gyrus region, which is apparent graphically for both AD and NC subjects, the AD subjects also show region-specific atrophy particularly in the frontal, temporal and to some extent parietal regions.

Although the data processing was computationally the most expensive task (in excess of a few hours using an 800 CPU cluster grid), followed by feature display and analysis, the entire visualization process consisting of loading the processed data to reaching the above conclusions only required a few minutes. The case studies demonstrate that INVIZIAN can be successfully used by researchers for visualizing database con exploratory analyses using visual tools.

Discussion

In this paper we sought to illustrate how the INVIZIAN system can be used for exploring distributions of neuroanatomical feature and subject metadata values across a collection of brain surfaces drawn from databases containing patients with age-related disease. This

graphical informatics system employs a dynamic and interactive approach for mining large-scale neuroimaging data, allowing the user to discover suggestive relationships between cortical surface feature and meta-data attribute values. This approach builds on pairwise distance calculations and decomposition of dataset features and uses the extracted brain surfaces to create a visualization environment intuitive to users of existing neuroimaging applications. Utilizing a multi-resolution mesh display, the system provides an exploratory visual overview of the data at interactive rates. The display environment anchors query-based grouping of clusters and feature-wise coloring for visual exploration of neuroanatomically similar groups. A collection of coordinated projections and graphical exploratory data analysis tools help the user keep track of global relationships with the data collection while they inspect and examine detailed local cortical structure. All in all, INVIZIAN provides a unique way to visually inspect, explore, and assess collections of data from large-scale databases.

We wish to emphasize that INVIZIAN is not meant to compete with nor replace a more formal statistical analytic treatment of the data. Though this may appear to be a limitation of this approach when, rather, the opposite is the case. INVIZIAN can be used as a useful precursor to explore, mine, and perform basic meta-analyses of neuroimaging archives as a way to generate new hypotheses about the data that might not have been apparent through narrative description of their collection, the samples in question, or previously described effects. Image-centric databases for neuroscience (Bug and Nissanov 2003) will require such tools to move beyond the limits of text-only searches and more readily take advantage of image content. The present situation with neuroimaging archives is not too dissimilar that of satellite imagery and space science telescope databases which found particular utility when their content was re-factored from a catalogue-based system to one driven specifically by what the images contain. Noteworthy examples of this include Microsoft's World Wide Telescope (<http://www.worldwidetelescope.org>; see also Szalay and Gray 2001) and, of course, Google Earth (<http://www.google.com/earth>).

As we develop the INVIZIAN further we will examine methods for visualizing multiple brain surface features per scene. For example, it would be of benefit to color the mesh data based on combinations of feature values. We will make the dimension reduction and offset generation system available interactively during runtime. We will also generalize the methods used to calculate the cortical surface coordinates by incorporating user-selected metadata attributes. While entirely feasible, there is a challenge in maintaining a suitable interface which is intuitive and not overly abstract to the user. Further evaluations will include what distance metrics lend themselves to effective user-based visual comparison of MRI cortical surface data.

In conclusion, the Informatics Visualization for Neuroimaging (INVIZIAN) framework graphically displays image and meta-data information from sizeable collections of neuroimaging data as a whole using a dynamic and compelling user interface. While not a replacement for more formal statistical analysis, INVIZIAN provides a means for quickly assessing brain and meta-data relationships as a precursor to further data access and modeling. In our paper, we illustrate the utility of INVIZIAN for simultaneous exploration and mining a large collection of extracted cortical surface data arising in clinical

neuroimaging studies of patients with Alzheimer's Disease, mild cognitive impairment, as well as healthy control subjects. Specifically, we show how cortical thickness and hippocampal volume differences between groups are evident using probabilistic distance and multivariate decomposition even in the absence of more formal hypothesis testing. For neurological diseases linked to brain aging such as AD, as well as across the entire lifespan, INVIZIAN provides a unique means for considering the entirety of whole brain datasets, exploring interesting relationships among them, and, in so doing, deriving new ideas for further research and study.

Acknowledgments

The authors wish to thank the faculty and staff of the Institute of Neuroimaging and Informatics of the University of Southern California and the Laboratory of Neuro Imaging (LONI) at the University of California Los Angeles. This work was supported by RC1 MH088194 and R01 MH100028 (sub-award) grants to JVH.

References

- Berretta R, Moscato P. Cancer biomarker discovery: the entropic hallmark. *PLoS ONE*. 2010; 5(8):e12262. [PubMed: 20805891]
- Biswal BB, Mennes M, Zuo XN, et al. Toward discovery science of human brain function. *Proc Natl Acad Sci U S A*. 2010; 107(10):4734–4739. [PubMed: 20176931]
- Bowman I, Joshi SH, Van Horn J. Visual Systems for Interactive Exploration and Mining of Large-Scale Neuroimaging Data Archives. *Frontiers in Neuroinformatics*. 2012;6. [PubMed: 22470335]
- Bug W, Nissano J. A guide to building image-centric databases. *Neuroinformatics*. 2003; 1(4):359–377. [PubMed: 15043221]
- Carmichael O, Schwarz C, Drucker D, et al. Longitudinal changes in white matter disease and cognition in the first year of the Alzheimer disease neuroimaging initiative. *Arch Neurol*. 2010; 67(11):1370–1378. [PubMed: 21060014]
- Carmichael OT, Aizenstein HA, Davis SW, et al. Atlas-based hippocampus segmentation in Alzheimer's disease and mild cognitive impairment. *Neuroimage*. 2005; 27(4):979–990. [PubMed: 15990339]
- Chen, C-h; Härdle, W.; Unwin, A., et al. *Handbook of Data Visualization*. Springer; Berlin Heidelberg: 2008. Multidimensional Scaling; p. 315–347.
- Chen RL, Guo W, Shi Y, et al. Computational identification of specific splicing regulatory elements from RNA-seq in lung cancer. *Eur Rev Med Pharmacol Sci*. 2013; 17(13):1716–1721. [PubMed: 23852893]
- Collins DL, Neelin P, Peters TM, et al. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *J Comput Assist Tomogr*. 1994; 18(2):192–205. [PubMed: 8126267]
- Cook, D.; Swayne, DF. *Interactive and Dynamic Graphics for Data Analysis: With R and GGobi*. New York: Springer; 2007.
- Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage*. 1999; 9(2):179–194. [PubMed: 9931268]
- Dinov I, Van Horn J, Lozev K, et al. Efficient, distributed and interactive neuroimaging data analysis using the LONI pipeline. *Frontiers in Neuroinformatics*. 2010; 3(22):1–10.
- Dinov ID, Van Horn JD, Lozev KM, et al. Efficient, distributed and interactive neuroimaging data analysis using the LONI pipeline. *Frontiers in Neuroinformatics*. 2009;3. [PubMed: 19212459]
- Endert, A.; Han, C.; Maiti, D., et al. Observation-level interaction with statistical models for visual analytics. In: Miksch, SaWM., editor. *IEEE Symposium on Visual Analytics Science and Technology - VAST*. Providence, RI: IEEE Computer Society; 2011.
- Fischl B, Sereno MI, Dale AM. Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage*. 1999; 9(2):195–207. [PubMed: 9931269]

- Goebel R, Esposito F, Formisano E. Analysis of functional image analysis contest (FIAC) data with brainvoyager QX: From single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. *Hum Brain Mapp.* 2006; 27(5):392–401. [PubMed: 16596654]
- Guo H, Rangarajan A, Joshi SC. 3-D diffeomorphic shape registration on hippocampal data sets. *Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv.* 2005; 8(Pt 2):984–991.
- Iglesias JE, Konukoglu E, Montillo A, et al. Combining generative and discriminative models for semantic segmentation of CT scans via active learning. *Inf Process Med Imaging.* 2011; 22:25–36. [PubMed: 21761643]
- Jack CR Jr, Bernstein MA, Fox NC, et al. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *J Magn Reson Imaging.* 2008; 27(4):685–691. [PubMed: 18302232]
- Johansson S, Johansson J. Interactive dimensionality reduction through user-defined combinations of quality metrics. *IEEE Transactions on Visualization and Computer Graphics.* 2009; 15:993–1000. [PubMed: 19834164]
- Joshi SH, Bowman I, Toga AW, et al. Brain Pattern Analysis of Cortical Valued Distributions. *Proc IEEE Int Symp Biomed Imaging.* 2011:1117–1120. [PubMed: 21743844]
- Joshi SH, Van Horn JD, Toga AW. Interactive exploration of neuroanatomical meta-spaces. *Front Neuroinformatics.* 2009; 3:38.
- Keim DA, Kriegel HP. VisDB: database exploration using multi-dimensional visualization. *IEEE Transactions on Computer Graphics and Applications.* 1994; 14(5):40–49.
- Kohonen T. The self-organizing map. *Neurocomputing.* 1998; 21(1–3):1–6.
- Kruskal, JB.; Wish, M. *Multidimensional Scaling.* New York: Sage Publications; 1978.
- Kuriakose J, Ghosh A, Ravi Kumar V, et al. Isometric graphing and multidimensional scaling for reaction-diffusion modeling on regular and fractal surfaces with spatiotemporal pattern recognition. *J Chem Phys.* 2004; 120(11):5432–5443. [PubMed: 15267417]
- Lancaster JL, Fox PT, Downs H, et al. Global spatial normalization of human brain using convex hulls. *J Nucl Med.* 1999; 40(6):942–955. [PubMed: 10452309]
- Lerch JP, Pruessner J, Zijdenbos AP, et al. Automated cortical thickness measurements from MRI can accurately separate Alzheimer's patients from normal elderly controls. *Neurobiol Aging.* 2008; 29(1):23–30. [PubMed: 17097767]
- Lu C, Zheng Y, Birkbeck N, et al. Precise segmentation of multiple organs in CT volumes using learning-based approach and information theory. *Med Image Comput Comput Assist Interv.* 2012; 15(Pt 2):462–469. [PubMed: 23286081]
- McCormick, PS.; Inman, JM.; Ahrens, JP., et al. *Visualization '04 (VIS '04).* Washington, DC, USA: IEEE Computer Society; 2004. Scout: A Hardware-Accelerated System for Quantitatively Driven Visualization and Analysis.
- Mega M, Dinov I, Mazziotta J, et al. Automated brain tissue assessment in the elderly and demented population: construction and validation of a sub-volume probabilistic brain atlas. *NeuroImage.* 2005; 26(4):1009–1018. [PubMed: 15908234]
- Megalooikonomou V, Ford J, Shen L, et al. Data mining in brain imaging. *Stat Methods Med Res.* 2000; 9(4):359–394. [PubMed: 11084713]
- Mennes M, Biswal BB, Castellanos FX, et al. Making data sharing work: The FCP/INDI experience. *Neuroimage.* 2012
- Narr KL, Bilder RM, Toga AW, et al. Mapping cortical thickness and gray matter concentration in first episode schizophrenia. *Cereb Cortex.* 2005; 15(6):708–719. [PubMed: 15371291]
- Nowinski WL, Belov D. The Cerefy Neuroradiology Atlas: a Talairach-Tournoux atlas-based tool for analysis of neuroimages available over the internet. *Neuroimage.* 2003; 20(1):50–57. [PubMed: 14527569]
- Nowinski WL, Thirunavuukarasuu A. Atlas-assisted localization analysis of functional images. *Med Image Anal.* 2001; 5(3):207–220. [PubMed: 11524227]
- Rencher, AC. *Methods of Multivariate Analysis.* New York, NY: John Wiley & Sons, Inc; 2002.

- Sasahara K, Hirata Y, Toyoda M, et al. Quantifying collective attention from tweet stream. *PLoS ONE*. 2013; 8(4):e61823. [PubMed: 23637913]
- Shattuck DW, Leahy RM. BrainSuite: an automated cortical surface identification tool. *Med Image Anal*. 2002; 6(2):129–142. [PubMed: 12045000]
- Stein JL, Medland SE, Vasquez AA, et al. Identification of common variants associated with human hippocampal and intracranial volumes. *Nat Genet*. 2012; 44(5):552–561. [PubMed: 22504417]
- Szalay A, Gray J. The World-Wide Telescope. *Science*. 2001; 293(5537):2037–2040. [PubMed: 11557879]
- Thompson PM, Mega MS, Woods RP, et al. Cortical change in Alzheimer's disease detected with a disease-specific population-based brain atlas. *Cereb Cortex*. 2001; 11(1):1–16. [PubMed: 11113031]
- Van Horn, JD.; Toga, AW. Brain Atlases: Their Development and Role in Functional Inference. In: Filippi, M., editor. *Functional MRI Techniques and Protocols*. New York: Humana Press; 2009.
- Van Horn JD, Toga AW. Multisite neuroimaging trials. *Curr Opin Neurol*. 2009; 22(4):370–378. [PubMed: 19506479]
- Van Horn JD, Wolfe J, Agnoli A, et al. Neuroimaging databases as a resource for scientific discovery. *Int Rev Neurobiol*. 2005; 66:55–87. [PubMed: 16387200]
- Voytek JB, Voytek B. Automated cognome construction and semi-automated hypothesis generation. *J Neurosci Methods*. 2012; 208(1):92–100. [PubMed: 22584238]
- Williams, M.; Munzner, T. Steerable, progressive multidimensional scaling. *IEEE Symposium on Information Visualization*; Washington, D.C. IEEE; 2004.
- Winkler AM, Kochunov P, Blangero J, et al. Cortical thickness or grey matter volume? The importance of selecting the phenotype for imaging genetics studies. *Neuroimage*. 2010; 53(3):1135–1146. [PubMed: 20006715]
- Xu M, Thompson PM, Toga AW. Adaptive reproducing kernel particle method for extraction of the cortical surface. *IEEE Trans Med Imaging*. 2006; 25(6):755–767. [PubMed: 16768240]
- Yang, J.; Peng, W.; Ward, MO., et al. Interactive hierarchical dimension ordering, spacing, and filtering for exploration of high dimensional datasets. *Ninth Annual IEEE Conference on Information Visualization (INFOVIZ'03)*; Washington, D.C. IEEE Computer Society; 2003. p. 105–112.
- Yang, J.; Ward, MO.; Rundensteiner, EA., et al. Visual hierarchical dimension reduction for exploration of high dimensional datasets. *Symposium on Data Visualization*; Aire-la-Ville, Switzerland: Euro-Graphics Association; 2003.

Figure 1.

A schematic overview of INVIZIAN system architecture operating on MRI cortical surface data. The filled top boxes represent fixed program functionality, while the outlined bottom boxes represent variable inputs defined by user-driven interactivity. Any of these user-defined values may be saved to an INVIZIAN scene file for use across visual analysis sessions.

Figure 2.

The initialized INVIZIAN similarity-based display of cortical feature information. The cortical surfaces are positioned according to feature-descriptive offsets in three dimensions, so that degree of proximity indicates similarity of neuroanatomical feature. Other arrangements are also possible. Feature values on the cortical surfaces are color mapped which, when viewed across subjects, highlights neuroanatomical trends. For instance, the surfaces pictured display a color gradient from blue to yellow along a central axis. This gradient indicates that generally surfaces on the left exhibit proportionally lower cortical feature values than those on the right.

Subject metadata list (right) for surface shown (left). Clicking with the mouse on any cortical surface within INVIZIAN displays this metadata dialogue containing subject-specific information including the subject's age, sex, weight and (by scrolling down in the window shown) diagnostic MRI information. Having the maximum resolution cortical surface rendering available allows the user to dynamically interact with and precisely examine the cortical detail for any selected subject.

Figure 4. Highlighting the subjects matching a search expression

As the subject-specific cortical surfaces are arranged in a feature-similarity dependent manner, the immediate graphical indication of results from a user-provided search illustrates a putative association between neuroanatomy and the subject metadata. (Top) Based on query results the user may define named groups, also assigning a group color, and optional description. (Bottom) INVIZIAN indicates group membership with a circular glyph which allows the user to view surface characteristics and group membership simultaneously.

Figure 5.

(Left) The initialized view of one coordinated projection is a two-dimensional dot plot of the feature-respective offsets, colored by group. (Middle) The coordinated projection renders a small rectangle that updates with zooming and panning to represent what region of the overall environment currently occupies the view. (Right) The user may turn off the group indicator glyphs within the display, but determine group membership by observing group colored plot. All in all, coordinated projections such as this allow the user to investigate cortical regional surface detail and global feature relationships, simultaneously.

Figure 6.

Visualization of cortical thickness: (Left) Quintile-based color map legend representing 0 to 5mm of thickness. Quintile or other sub-divisioning helps the user assess the position of features in the overall distribution of measured values. (Middle) The observed feature trend is an increase in thickness from left to right. (Right) Each brain is annotated by a glyph according to diagnosis, red representing AD (Alzheimer's Disease) and blue representing NC (Normal Control). The clustered groupings of AD and NC subjects on the left and right respectively indicate that cortical thickness is a strong predictor of Alzheimer's disease.

Figure 7.

Cortical surface area: (Left) Quintile-based color map legend representing 0 to 7mm² surface area values. The small quintile distance in the color map indicates small area variation across the surfaces. (Middle) The area remains nearly same across the population. (Right) Each brain is annotated by a glyph according to diagnosis, red representing AD (Alzheimer's Disease) and blue representing NC (Normal Control). There is little separation between groupings, indicating that cortical area is not a predictor for Alzheimer's disease.

Figure 8.

Gray matter volume: (Left) Here, the quintile-based color map legend represents 0 to 10mm³ gray matter volume values. (Middle) The observed trend in features shows an increase in volume from left to right. (Right) Each brain is annotated by a glyph according to diagnosis, red representing AD (Alzheimer's Disease) and blue representing NC (Normal Control). The slightly overlapped groupings indicate that gray matter volume has a reduced predictive power relative to cortical thickness for Alzheimer's disease.

Figure 9.

A comparison of subjects' cortical attributes. Left, the quintile based color mapping of cortical thickness is coded in example AD patients, while, on the right, healthy control subjects are displayed using the same color scaling.