MULTI-RESOLUTION STATISTICAL ANALYSIS ON GRAPH STRUCTURED DATA IN NEUROIMAGING

Won Hwa Kim Vikas Singh Moo K. Chung Nagesh Adluru Barbara B. Bendlin Sterling C. Johnson

University of Wisconsin - Madison

http://pages.cs.wisc.edu/~wonhwa

ABSTRACT

Statistical data analysis plays a major role in discovering structural and functional imaging phenotypes for mental disorders such as Alzheimer's disease (AD). The goal here is to identify, ideally early on, which regions in the brain show abnormal variations with a disorder. To make the method more sensitive, we rely on a multi-resolutional perspective of the given data. Since the underlying imaging data (such as cortical surfaces and connectomes) are naturally represented in the form of weighted graphs which lie in a non-Euclidean space, we introduce recent work from the harmonics literature to derive an effective multi-scale descriptor using wavelets on graphs that characterize the local context at each data point. Using this descriptor, we demonstrate experiments where we identify significant differences between AD and control populations using cortical surface data and tractography derived graphs/networks.

Index Terms— wavelets, wavelets on graphs, cortical thickness, brain network, Alzheimer's disease

1. INTRODUCTION

The representation of an image at multiple resolutions is fundamental to a broad spectrum of approaches in computer vision and image processing [1]. Ideas based on the Laplacian of Gaussians and Wavelets drive various applications like interest point detection, matching/registration, texture analysis, compression and denoising. Interestingly, while such multi-resolution concepts are central to image registration in Neuroimaging, their use in (the downstream) statistical analysis of brain imaging data (except perhaps spherical harmonics [2]) has been limited. The defacto analysis is performed at a single (given) resolution and based on the univariate intensity measurement at each voxel. For example, we may perform statistical hypothesis testing at a specific voxel across a cohort of images of disparate groups (diseased/healthy) and check if it is statistically different across groups; by repeating the test at every voxel we obtain disease affected 'regions'. Such techniques broadly fall under "Voxel based analysis" (VBA). Frequently, VBA is adapted with little modification to statistical analysis tasks which involve brain meshes as well as brain connectivity graphs/networks. This work investigates new models which may offer improvements in this graph structured data regime.

Our paper is inspired from the observation that statistical inference on signals/functions may be more meaningful using multivariate descriptors that characterize the *local context* around each measurement location rather than a single univariate measurement, as in VBA. This intuition is very similar to Scale Space theory [3] — what we investigate and present here are analogous ideas when the underlying data are *not images*, instead graphs structured data like meshes and connectivity. Specifically, if the domain of the image can be represented as a regular grid lattice, for multi-resolution analysis, we can easily write the wavelet expansion and proceed with the analysis of the data. Instead, many recent brain imaging datasets contain data that live in a non-Euclidean space, for example cortical thickness on brain surfaces from Freesurfer [4] or tractography derived connectivity measures from diffusion tensor imaging (DTI) [5]. In such settings, it is problematic to perform Wavelet analysis since the domain has arbitrary structure. The main contribution of this work is to show using three interesting applications, how one can adapt recent results from the harmonic analysis literature [6] to derive representations and improve statistical power for common brain image analysis pipelines involving graph-structured data.

2. METHOD

We will derive a multi-variate descriptor at each measurement location using a suitable Wavelet parameterization of the observed function. We first describe the Wavelet formulation and then the scheme for deriving a descriptor on graph nodes.

2.1. Preliminaries: Continuous Wavelets Transform

The Wavelet transform is conceptually similar to the Fourier transform [7]. While the Fourier bases is localized in frequency only, wavelets can be localized in both time and frequency. The traditional construction of wavelet transform requires a mother wavelet function $\psi_{s,a}(x) = \frac{1}{s}\psi(\frac{x-a}{s})$ defined by two parameters, the scale s and translation a. The mother wavelet $\psi_{s,a}(x)$ is a localized oscillating function at a with finite duration acting as a local support controlled by s. The terms $\psi_{s,a}(x)$ at multiple scales are used to approxi-

mate a signal using a wavelet expansion, moreover, they form band-pass filters in the frequency domain.

Using ψ, the wavelet transform of a signal $f(x)$ is defined as a projection of f onto the wavelet bases as

$$W_f(s,a) = \langle f, \psi \rangle = \frac{1}{s} \int f(x) \psi^*(\frac{x-a}{s}) \mathrm{d}x \tag{1}$$

yielding wavelet coefficient $W_f(s,a)$ at scale s and at location a, where ψ^* is the complex conjugate of ψ. Such a transform is invertible, that is

$$f(x) = \frac{1}{C_\psi} \iint W_f(s,a) \psi_{s,a}(x) \mathrm{d}a \, \mathrm{d}s \tag{2}$$

where C_ψ is the so-called admissibility condition constant. Unlike the single set of $\sin()$ basis in the Fourier transform, wavelet transforms have many possible basis functions depending on their shapes and use. The wavelet transform above is not directly applicable when the domain is arbitrarily structured, such as graphs. Next, we review how an analogue of the wavelet transform can be defined on graphs [6].

2.2. Wavelets in Arbitrary Structured Domain

The basic idea in defining a Wavelet transform on graph is to obtain a wavelet basis using spectral graph theory concepts. We describe this construction below. A graph $G = \{V, E, \omega\}$ is defined by vertex set V, edge set E and edge weights ω, which may be given as a $N \times N$ adjacency matrix $A = \{a_{ij}\}$ when $|V| = N$. Here, a_{ij} gives the connection weight between the ith and jth vertices, i.e., the edge weight ω_{ij}. A (diagonal) degree matrix D has at the ith diagonal, the sum of all the edges connected to the ith vertex. From these two matrices, a graph Laplacian L is defined as $L = D - A$. The matrix L is positive semi-definite, therefore, has ordered eigenvalues $0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{N-1}$ and corresponding eigenvector χ. Using the $\{\lambda_l, \chi_l\}$ pairs, the forward and inverse graph Fourier transformation is defined as

$$\hat{f}(l) = \langle \chi_l, f \rangle = \sum_{n=1}^{N} \chi_l^*(n) f(n), \text{ and } f(n) = \sum_{l=0}^{N-1} \hat{f}(l) \chi_l(n) \tag{3}$$

Then, the spectrum of the Laplacian corresponds to the frequency domain, where scales are defined over band-pass filters $g()$ which are dual representation of wavelets. From the scaling property of the Fourier transform, scales can be defined with $g()$ in the frequency domain. Using (3), we construct spectral graph wavelets by applying band-pass filters at multiple scales and localizing it with an impulse function as,

$$\psi_{s,n}(m) = \sum_{l=0}^{N-1} g(s\lambda_l) \chi_l^*(n) \chi_l(m) \tag{4}$$

where m and n are vertex indices on the graph. Now, the wavelet coefficients of a function $f(n)$ can be easily obtained by the inner product of the wavelets and the given function,

$$W_f(s,n) = \langle \psi_{s,n}, f \rangle = \sum_{l=0}^{N-1} g(s\lambda_l) \hat{f}(l) \chi_l(n), \tag{5}$$

which finally defines spectral graph wavelet transform (SGWT) [6]. The coefficients obtained from the transformation yield the Wavelet Multiscale Descriptor (WMD) as a set of wavelet coefficients at each vertex n for each scale s as

$$\mathrm{WMD}_f(n) = \{W_f(s,n) | s \in S\} \tag{6}$$

which is a *multi resolution* descriptor defining local context at each vertex n on the graph [8, 9]. Our statistical analysis will operate on this descriptor in the next section.

3. RESULT

In this section, we demonstrate various applications of multi-resolutional representation of functions given in non-Euclidean space. First, we show results of cortical surface and signal smoothing, then exhibit results from group analysis on cortical thickness and DTI tractography which show significant improvements in statistical power.

3.1. Smoothing via Wavelets

Existing methods for smoothing cortical surfaces and the function defined on it, such as spherical harmonics, suffer from loss of information due to the ballooning process which maps the convoluted brain surface to a sphere. If we can define bases directly on the brain surface, obtaining its smoothed approximation will be much easier. Using SGWT, we can define wavelet transform on the arbitrarily structured surfaces, and inverse wavelet transformation provides the smoothed estimate of the cortical surface at various scales. Let us rewrite (2) in terms of the graph Fourier basis,

$$\frac{1}{C_g} \sum_l \left(\int_0^\infty \frac{g^2(s\lambda_l)}{s} \mathrm{d}s \right) \hat{f}(l) \chi_l(m) \tag{7}$$

where the set of scales controls the spatial smoothness of the surface. Coarser spectral scales overlap less and smooth higher frequencies, and at finer scales, the complete spectrum is used and recovers the original surface to high definition. Smoothing at multiple resolutions is shown in Fig. 1 where we can see the reconstruction of a brain mesh surface (and the cortical thickness signal) from coarse to finer scales.

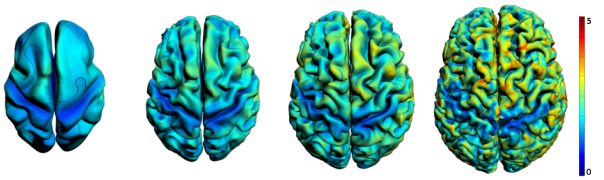

Fig. 1. Smoothing of cortical surface and the cortical thickness values on the cortical surface, demonstrated from coarse to fine scales.

3.2. Cortical Thickness Analysis on Brain Surfaces

Cortical thickness is a distinctive biomarker implicated in brain disorders such as Alzheimer's disease (AD). Given that brain function manifests strongly as changes in the cortical thickness, the statistical analysis of such data to find clinically meaningful group level differences is critical in structural brain imaging studies. Typical analysis in clinical studies

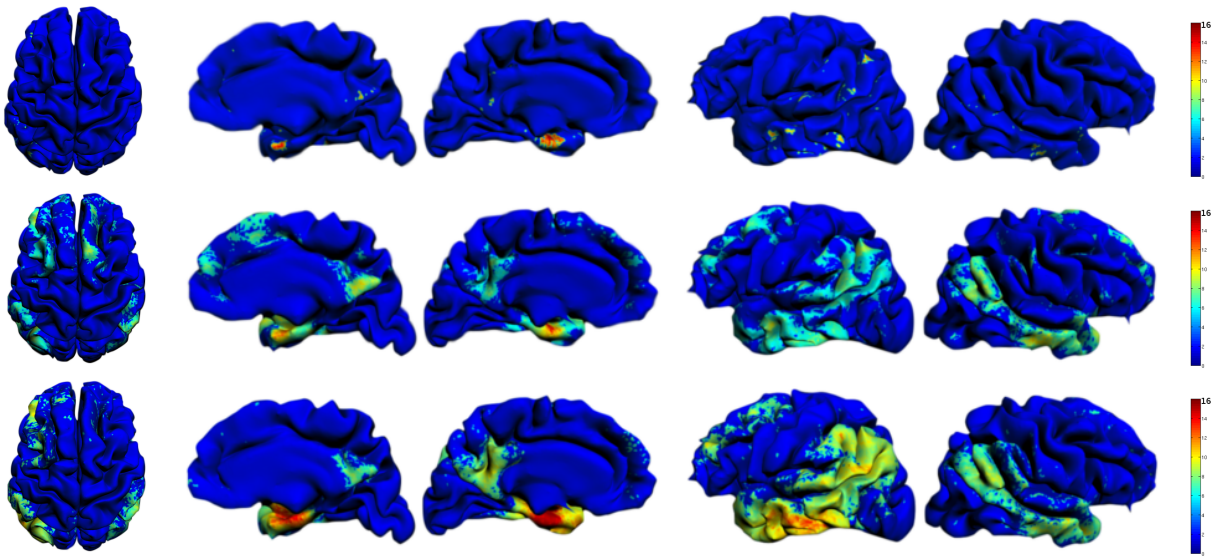

Fig. 2. Cortical thickness group difference analysis (Top row: cortical thickness, Middle row: SPHARM, Bottom row: WMD). Resultant p values in $-\log_{10}$ scale after FDR correction at $\alpha = 0.001$ are mapped on a template brain, identifying brain regions that are affected by AD.

use cortical thickness values obtained from the segmentation directly for a vertex-wise $t-$test, or smooth the signal first using spherical harmonic (SPHARM) or spherical wavelet approach. Employing multiple comparisons correction on the p-values obtained from the tests and mapping them to the original brain surface reveals disease specific regions.

Our multi-scale descriptor from the Wavelet coefficients characterizes the shape (and the signal) on the native graph domain itself. Since WMD is a multivariate descriptor, we will use Hotelling's T^2 test for hypothesis testing, which is a generalization of the student's t test. Using the raw cortical thickness and SPHARM representation as the baseline, we next describe our experimental design and results.

Dataset. We used Magnetic Resonance (MR) images data acquired from the Alzheimer's Disease Neuroimaging Initiative (ADNI). Our data consists of 356 participants: 160 AD (86 male and 74 female; age: 75.53 ± 7.41) and 196 healthy controls (101 male and 95 female; age: 76.09 ± 5.13). The dataset was pre-processed using a standard pipeline. The Freesurfer algorithm [10] was used to segment the cortical surfaces, calculate the cortical thickness values, and provide vertex to vertex correspondences across brain surfaces.

Analysis. Our analysis is summarized in Fig. 2. The first row corresponds to group analysis using the original cortical thickness values (CT), which reveals small regions. The second row shows results using SPHARM which offers an improvement over the baseline. Finally, the bottom row shows that performing the statistical tests using our multi-scale descriptor identifies much larger regions with significantly smaller p-values. The number of vertices showing significant group differences is substantially larger using WMD compared to CT and SPHARM. At FDR 10^{-4} level, there are a total of 6943 (CT), 28789 (SPHARM) and 40548 (WMD) vertices out of 131076 vertices.

We checked whether the regions identified by the methods are expected to be atrophic in AD. All three methods identified the anterior entorhinal cortex in the mesial temporal lobe, but at the prespecified threshold, the WMD method was more sensitive to changes in this location as well as in the posterior cingulate, precuneus, lateral parietal lobe, and dorsolateral frontal lobe. These are regions that are commonly implicated in AD, and strongly tie to known results from neuroscience.

3.3. Brain Connectome Analysis

Recent developments in deriving tractography results from diffusion weighted imaging (DWI) have enabled the analysis of connectivity data for a better understanding of how disease and aging impacts brain connectivity. The tractography solution represents connectivity as a weighted graph, where each vertex corresponds to the anatomically segmented region of interests (ROI) in the brain and the edges denote the relation between the ROIs by clinically meaningful measures such as the strength of the tracks or the mean Fractional Anisotropy (FA) along the fiber bundle. Again, as such studies may involve small sample sizes, the goal is to maximize statistical power of detecting disease specific group differences in the brain connectivity. It is problematic to apply the WMD framework directly to this connectivity setting, since the signal that we want to detect is defined on the connection, not on the vertices. Therefore, we adopt the line graph transform [11, 12] which changes the role of vertices and edges in a graph to obtain a dual representation of the brain network: the measurement at each edge is considered as a function at each vertex. We then proceed with the analysis as described in section 3.2, and the resultant connectivity of significant group differences in fractional anisotropy (FA) between AD and healthy control groups controlled for age and sex are presented.

Dataset. The Wisconsin Alzheimer's Disease Research

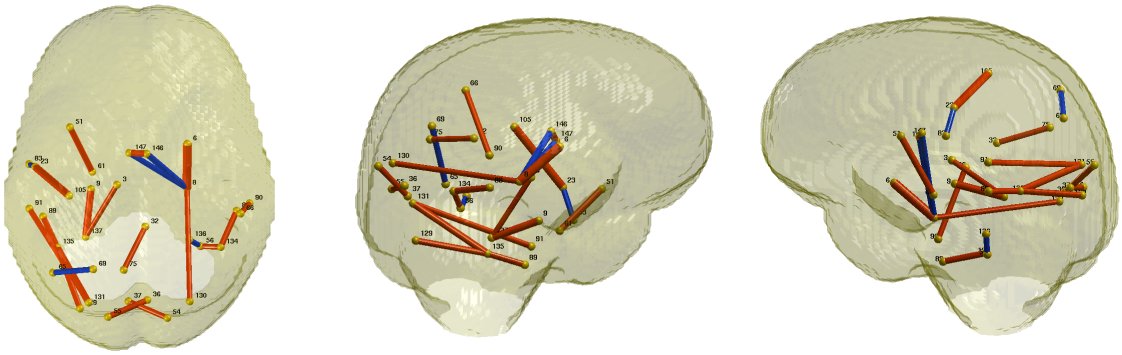

Fig. 3. Significant group differences (controlled for age and sex) from AD vs. control connection analysis using WMD on FA. Those connections with p-values that survive Bonferroni thresholding at 0.001 are exhibited in top, left and right view. The thickness of each connection represents the p-values in $-\log_{10}$ scale (thicker connection corresponds to lower p-value), and the color of each connection represents the direction of the difference (red: stronger in controls group, blue: stronger in AD group). See IIT3 atlas document on NITRC for the region labels.

Center (WADRC) dataset included 102 subjects, which comprised 44 AD (31 male and 13 female; age: 77.05 ± 9.35) and 58 CN (33 male and 25 female; age: 74.05 ± 6.82). Participants were diagnostically characterized in the WADRC's multidisciplinary consensus conferences using standard procedures for the diagnosis of AD [13, 14]. For tractography, a total of 164 regions were defined [15], and two regions labeled `unknown` were excluded in the analysis giving us 162×162 symmetric matrices per subject with mean FA values as elements. A full index of the ROIs can be found in IIT3 atlas documentation on NITRC.

Analysis. We first note that we could not identify any significant connections using the *raw* FA values when controlled for age and sex. However, using our framework, we discover 22 significant connections, after controlling for age and sex, even after Bonferroni correction at 0.001. As demonstrated in in Fig. 3, these connections involve representative brain regions discovered in many AD studies such as left and right Hippocampus, left parietal superior, right precuneus, left and right temporal regions and others, matching our previous result in section 3.2. Among the 22 connections, 17 connections showed stronger FA in the CN group, and 5 connections show stronger FA in the AD group.

4. CONCLUSION

In this paper, we demonstrated how a multi-resolutional framework based on Wavelets can help facilitate statistical analysis of (brain image-derived) graph structured representations. We show via various experiments on surface and signal smoothing on brain surfaces, statistical analysis on cortical thickness on surfaces and brain connectivity, that statistical power can be improved with only minor changes in the analysis design (i.e., involves substituting univariate tests with their multivariate versions). Our results strongly suggests that higher sensitivity can be obtained in many of these applications with only a small additional computational load. Code/other details available on the first author's homepage.

5. REFERENCES

[1] D. Lowe, "Object recognition from local scale-invariant features," in *ICCV*, 1999, vol. 2, pp. 1150–1157.

[2] M.K. Chung, K.M. Dalton, and R.J. Davidson, "Tensor-based cortical surface morphometry via weighted spherical harmonic representation," *TMI*, vol. 27, no. 8, pp. 1143 –1151, aug. 2008.

[3] T. Lindeberg, "Scale-space theory: A basic tool for analyzing structures at different scales," *Journal of Applied Statistics*, vol. 21, no. 1-2, pp. 225–270, 1994.

[4] B. Fischl, "FreeSurfer," *NeuroImage*, vol. 62, no. 2, pp. 774–781, 2012.

[5] P. Hagmann, L. Cammoun, X. Gigandet, et al., "Mapping the structural core of human cerebral cortex," *PLoS Biology*, vol. 6, no. 7, pp. e159, 2008.

[6] D. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 129 – 150, 2011.

[7] S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. on PAMI*, vol. 11, no. 7, pp. 674 –693, 1989.

[8] W. H. Kim, V. Singh, M. K. Chung, et al., "Multi-resolutional shape features via non-euclidean wavelets: Applications to statistical analysis of cortical thickness," *NeuroImage*, vol. 93, pp. 107–123, 2014.

[9] W.H. Kim, D. Pachauri, C. Hatt, et al., "Wavelet based multi-scale shape features on arbitrary surfaces for cortical thickness discrimination," in *NIPS*, 2012, pp. 1250–1258.

[10] M. Reuter, H. Rosas, and B. Fischl, "Highly accurate inverse consistent registration: A robust approach," *NeuroImage*, vol. 53, no. 4, pp. 1181–1196, 2010.

[11] F. Harary, *Graph Theory*, Addison-Wesley, 1969.

[12] W.H. Kim, N. Adluru, M. Chung, et al., "Multi-resolutional brain network filtering and analysis via wavelets on non-euclidean space," in *MICCAI*, pp. 643–651. Springer, 2013.

[13] G. McKhann, D. Drachman, M. Folstein, et al., "Clinical diagnosis of AD report of the NINCDS-ADRDA work group under the auspices of Dept. of Health and Human Services task force on AD," *Neurology*, vol. 34, no. 7, pp. 939–939, 1984.

[14] G. McKhann, D. Knopman, H. Chertkow, et al., "The diagnosis of dementia due to AD: Recommendations from the NIA-Alzheimers Association workgroups on diagnostic guidelines for AD," *Alzheimer's & Dementia*, vol. 7, no. 3, pp. 263–269, 2011.

[15] A. Varentsova, S. Zhang, and K. Arfanakis, "Development of a high angular resolution diffusion imaging human brain template," *NeuroImage*, vol. 91, pp. 177–186, 2014.