Statistical Inference Models for Image Datasets with Systematic Variations

Won Hwa Kim[1,3], Barbara B. Bendlin[4,3], Moo K. Chung[2] Sterling C. Johnson[4,3] Vikas Singh[2,1,3]
[1]Dept. of Computer Sciences, University of Wisconsin - Madison. [2]Dept. of Biostatistics & Med. Informatics, University of Wisconsin - Madison.
[3]Wisconsin Alzheimer's Disease Research Center, University of Wisconsin - Madison. [4]GRECC, William S. Middleton VA Hospital.

Statistical analysis of longitudinal or cross sectional brain imaging data to identify effects of neurodegenerative diseases is a fundamental task in various studies in neuroscience. However, when there are systematic variations in the images due to parameter changes such as changes in the scanner protocol, hardware changes, or when combining data from multi-site studies, the statistical analysis becomes problematic. In this scenario, the goal of this paper is to develop a unified statistical solution to the problem of systematic variations in statistical image analysis. Based in part on recent literature in harmonic analysis on diffusion maps, we propose an algorithm which compares operators that are resilient to the systematic variations. These operators are derived from the empirical measurements of the image data and provide an efficient surrogate to capturing the actual changes across images. To evaluate the proposed ideas, we present various experimental results on detecting changes in simulations as well as show how the method offers improved statistical power in the analysis of real longitudinal PIB-PET imaging data acquired from participants at risk for Alzheimer's disease (AD).

High level description of the framework. Consider two unknown latent functions f and f', and let α and β denote two parameters such that they modify the form of the function. Then, we are provided with the measurements, f_α and f'_β, i.e., both the parameter and the function changes. It is clearly not possible to verify the true difference between the latent functions, unless we also know the relationship between the transformations induced by α and β. Assume that an oracle provides us an operator $\mathcal{T}$ with the interesting property that it is *invariant* to the parameter space $\mathcal{P}$ from which α and β are drawn. That is, if we construct a pair of operators from the empirical measurements of f_α and f'_β, the operators will reflect the true difference between f and f'. Since the operator $\mathcal{T}$ only offers invariances to the parameter space $\mathcal{P}$, in this case, the operators $\mathcal{T}_{f_\alpha}$ and $\mathcal{T}_{f'_\beta}$ *cannot* be compared. Nonetheless, these operators provide a mapping to two *different* spaces, S_{f_α} and $S_{f'_\beta}$, since f and f' are distinct. Because of the invariance to $\mathcal{P}$, if we plug a *known* function (such as an impulse function) at all locations in the original space into the two operators, we will obtain its transformed representations in S_{f_α} and $S_{f'_\beta}$. Once these transformed forms of the impulse functions are mapped from $S_{f'_\beta}$ to S_{f_α}, we can calculate the distance [1].

Key Idea. We assume that the two images f_α and f'_β are spatially registered so that the *only* variations in the measurements comes from the parameters, $\alpha, \beta \in \mathcal{P}$. We also assume that the relationship of measurements at *each* grid point to measurements at *other* grid points in the same image is preserved and is independent of changes in the parameter space, $\mathcal{P}$. When we place a unit energy at location p, the propagation of the energy will now show *different* patterns for the two image-derived operators: capturing the difference between those patterns is an excellent surrogate for detecting the difference between the two original functions f and f'.

Interestingly, the process of constructing an operator and applying it on a δ function is the same as the implementation of a mother wavelet function. In a graph setting (with N number of nodes), this mother wavelet is constructed using bases from spectral graph theory. Let λ, and χ denote the eigenvalues and eigenvectors from a graph Laplacian, then the actual mother wavelet $\psi_{s,p}$ at scale s centered at p is realized by a delta function δ_p as

$$\psi_{s,p}(q) = \mathcal{T}_g^s \delta_p = \sum_{l=0}^{N-1} g(s\lambda_l)\chi_l(p)\chi_l(q) \tag{1}$$

where $g()$ is a band-pass filter from a filter bank [2]. Here, a wavelet function $\psi_{s,p}(q)$ can be viewed as a kernel function as $\psi_s(p,q)$, defining a relationship between vertex p and q. For our application, it is used to define Wavelet Kernel Distance (WKD) $d_s(p,q)$ at scale s, a measure between two points p and q defined as ℓ^2-norm of the wavelet density difference as

$$d_s(p,q)^2 = \sum_{l=0}^{N-1} g(s\lambda_l)^2(\chi_l(p) - \chi_l(q))^2 \tag{2}$$

It can be interpreted as if we were comparing the effect of the same wavelet function dissipating from locations p and q to their neighbors by the wavelet kernel function $g()$, thereby measuring the effect of the propagation.

Consider two individual graphs I and J, constructed using functions (or images) f_α and f'_β, where the number of vertices in each is N. Let λ^I, λ^J and χ^I and χ^J denote the eigenvalues and eigenvectors from graph Laplacians of I and J respectively. On these graphs, WKD between a vertex p^I from I and a vertex q^J on J is defined as

$$\begin{aligned} d_s(p^I,q^J)^2 = &\sum_{l_1=0}^{N-1} g(s\lambda_{l_1}^I)^2\chi_{l_1}^I(p)^2 + \sum_{l_2=0}^{N-1} g(s\lambda_{l_2}^J)^2\chi_{l_2}^J(q)^2 \\ &- 2\sum_{l_1,l_2=0}^{N-1} g(s\lambda_{l_1}^I)\chi_{l_1}^I(p)g(s\lambda_{l_2}^J)\chi_{l_2}^J(q)\langle\chi_{l_1}^I,\chi_{l_2}^J\rangle \end{aligned} \tag{3}$$

using wavelet kernel functions ψ_s^I and ψ_s^J.

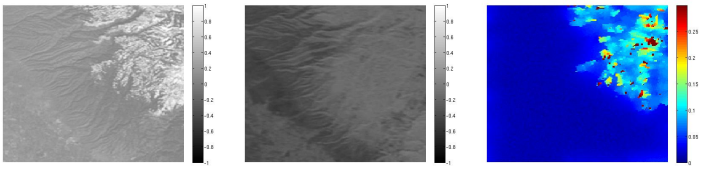

Figure 1: Results from NASA Earth Observatory images. Left: image from 2013, Middle: image from 2014 (intensity transformed), Right: image changes using WKD.

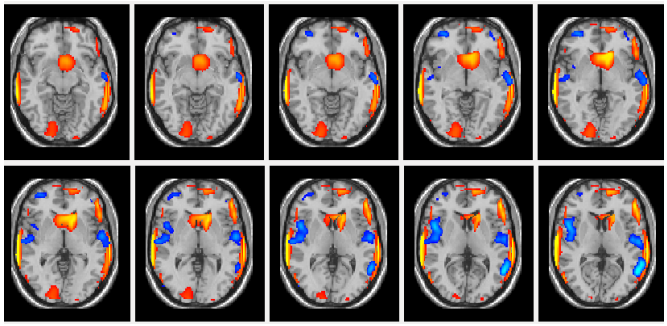

Figure 2: The correlation between the PIB changes and the ratio of total τ-protein and Aβ(1-42). The red-yellow and blue-light blue intensities indicate correlation using WKD and SUVR images respectively in the range of [0.3 0.5].

Results. We applied our algorithm on synthetic images and for statistical analysis of longitudinal Pittsburgh compound B positron emission tomography (PIB-PET) scans. Representative results are shown in Fig. 1 and 2. In Fig. 1, we show a result comparing two different satellite images that cannot be compared directly due to image intensity transformation and in Fig. 2, we show correlation between a well-known AD risk factor (i.e., ratio of total τ-protein and Aβ(1-42)) and subject age when a proper normalization is unavailable. More experimental results are provided in the main paper.

[1] Ronald R Coifman and Matthew J Hirn. Diffusion maps for changing data. *Applied and Computational Harmonic Analysis*, 36(1):79–107, 2014.

[2] D.K. Hammond, P. Vandergheynst, and R. Gribonval. Wavelets on graphs via spectral graph theory. *Applied and Computational Harmonic Analysis*, 30(2):129 – 150, 2011. ISSN 1063-5203. doi: 10.1016/j.acha.2010.04.005.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.