January 10, 2012

Statistical Power Maps for Sparse Representation of Subcortical Brain Structures

Seung-Goo Kim[a], Stacey M. Schaefer[b], Moo K. Chung[b,c,1],
Carien M. van Reekum[e] and Richard J. Davidson[b,d]

[a]Department of Brain and Cognitive Sciences, Seoul National University, Korea

[b]Waisman Laboratory for Brain Imaging and Behavior, [c]Department of Biostatistics and Medical Informatics, [d]Department of Psychology and Psychiatry, University of Wisconsin, Madison, WI, USA.

[e]Centre for Integrative Neuroscience and Neurodynamics, School of Psychology and Clinical Language Sciences, University of Reading, UK.

OBJECTIVES

It is difficult to determine the sample size required to obtain a specific statistical power in brain imaging studies due to the massive multiple comparisons. Instead of trying to compute the power analytically using non-central random fields [1], which is not publicly available yet, we computed the power numerically using permutations. The power analysis was then used in demonstrating a new sparse data representation [2] may require substantially smaller sample size in achieving the same power compared to the traditional least squares estimation (LSE) techniques. The proposed sparse representation and power computation were applied to amygdala and hippocampus shapes.

METHODS

We have T1-weighted MRI of 52 normal adults (age range 37-74; 16 men and 36 women). Amygdalae and hippocampi were manually segmented. We extracted brain volume using BET/FSL [3], and performed a nonlinear registration using ANTS [4]. We used the length of displacement of warping an individual structure to a template as a main anatomical feature of interest. The processing steps are illustrated in Figure 1.

The Laplace-Beltrami (LB) eigenfunctions is then used to parametrically represent the length as a series expansion. In almost all previous studies, LSE was used in estimating the coefficients of the expansion, while high frequency terms were truncated to reduce noise. However, some lower frequency terms may not necessarily contribute significantly in the reconstruction. So we performed the l1-norm minimization to sparsely filter out insignificant terms [2]. Figure 2 shows the first few LB-eigenfunctions and LSE and l1-minimization results.

We tested the gender effect in the general linear model (GLM) at each point x:

$$\text{Length}(x) = a(x) + b(x)*\text{Brain} + c(x)*\text{Age} + d(x)*\text{Gender} + \text{noise}(x),$$

[1] Corresponding address: Moo K. Chung, Waisman Center #281, 1500 Highland Ave. Madison, WI. 53705. Telephone: 608-217-2452, mailto:mkchung@wisc.edu http://www.stat.wisc.edu/~mchung/

where Brain is the total brain volume. As an illustration, we will only consider one sided test here. The null hypothesis H0 for the multiple comparison is that d(x) = 0 for every point x, while the alternative hypothesis H1 is d(x) = const. > 0 for any point x. Let M be the region where H1 is true. As a test statistic, we use a t-random field T(x). Thus the thresholding h for significance 0.05 is given by the probability

$$\Pr(\max T(x) > h) = 0.05,$$

where the maximum is taken over the whole region using random filed theory [1]. For this given threshold h, we can compute the power as the probability

$$\Pr(\max T(x) > h),$$

where the maximum is restricted to the region M. To estimate power, we opted for numerical simulation with 2000 permutations.

RESULTS

We performed the power computation in three ways.

(1) Simulation. We generated 50 Gaussian random fields with zero mean and unit variance on the left hippocampus surface. The effect size and signal region were estimated from [2]. A synthetic signal with the effect size of 0.86 standard deviation over a 5-ring neighbor patch was added for 25 cases. l1-minimization shows higher power than LSE (max 47% more), illustrating the advantage of sparse regression over LSE (Figure 3a).

(2) Real dataset. We used the 52 samples from [2]. We assumed the region of the significant gender difference observed in [2] to be the true signal region M. l1-minimization shows higher power than LSE (max 13% more for the right amygdala and max 8% more for the left hippocampus; Figure 3b).

(3) Local power map. For the real data, we also computed local power maps were estimated following the framework [1]. As seen in the overall power curves, l1-minimization gives higher power with a larger extent for a given sample size and it needs smaller sample size to achieve a given power level than LSE, again demonstrating the advantage of the proposed method (Figure 4).

CONCLUSIONS

We showed the proposed sparse representation method shows higher overall statistical power for a given sample size than LSE. Additionally, the power maps and sample size maps demonstrated that the sparse representation has advantage not only in the overall power, but also in the discoverable extent of signals.

ACKNOWLEDGEMENT

This research was supported by WCU program through the KOSEF funded by the MEST, Korea (R31-10089) to MKC and by grants from the National Institute of Aging (P01-AG20166) to RJD. We also like to thank Seongho Seo of the Department of Brain and Cognitive Sciences, Seoul National University and Shubing Wang of Merck for the discussion on the sparse shape representation.

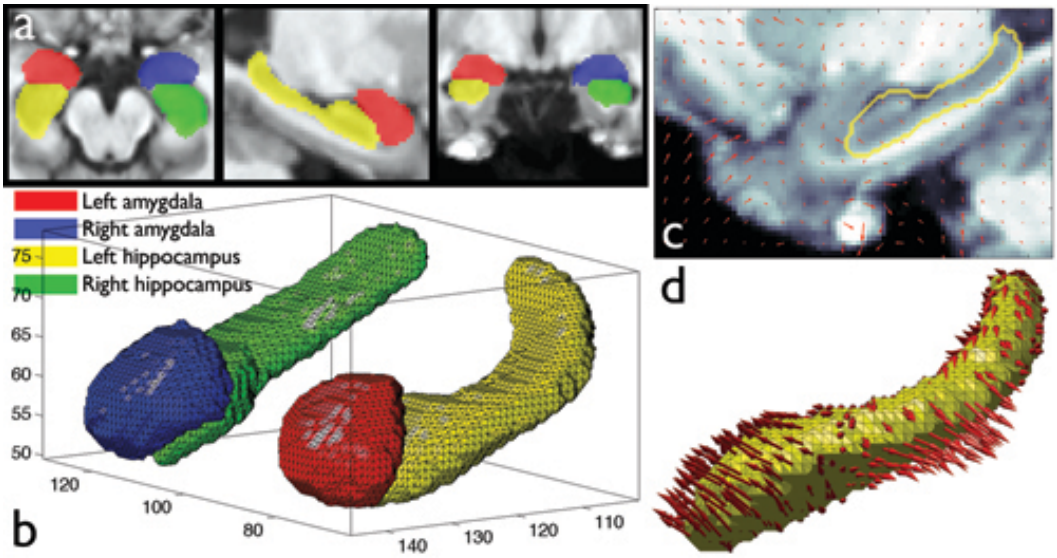

Figure 1: Processing of subcortical structures. Manual segmentations of amygdala and hippocampus (a) were used to construct template surfaces (b). Displacement vector field obtained during non-linear normalization (c) is interpolated onto the template surface (d). The length of displacement is used as an anatomical measure.

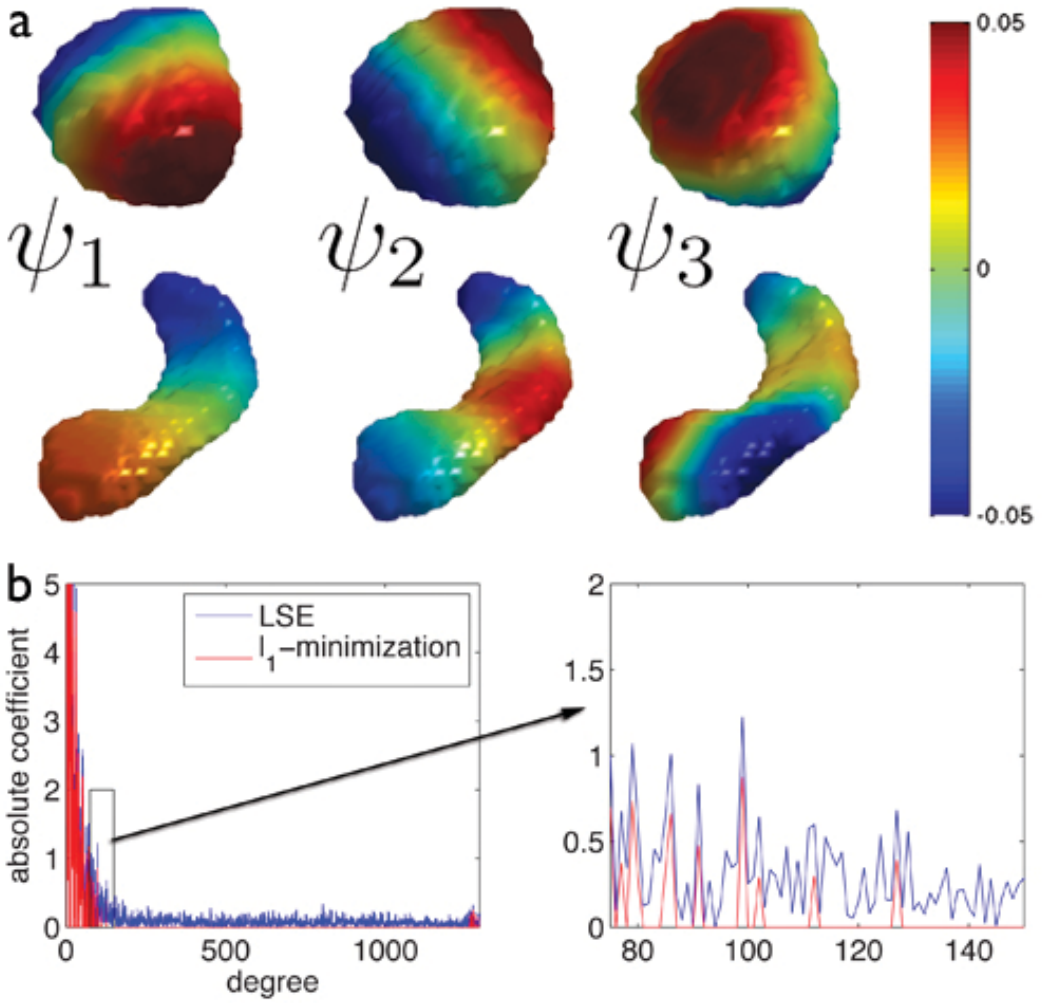

Figure 2: Examples of LB-eigenfunctions (a) and absolute values of Fourier coefficients estimated by LSE (blue) and l_1-minimization (red).

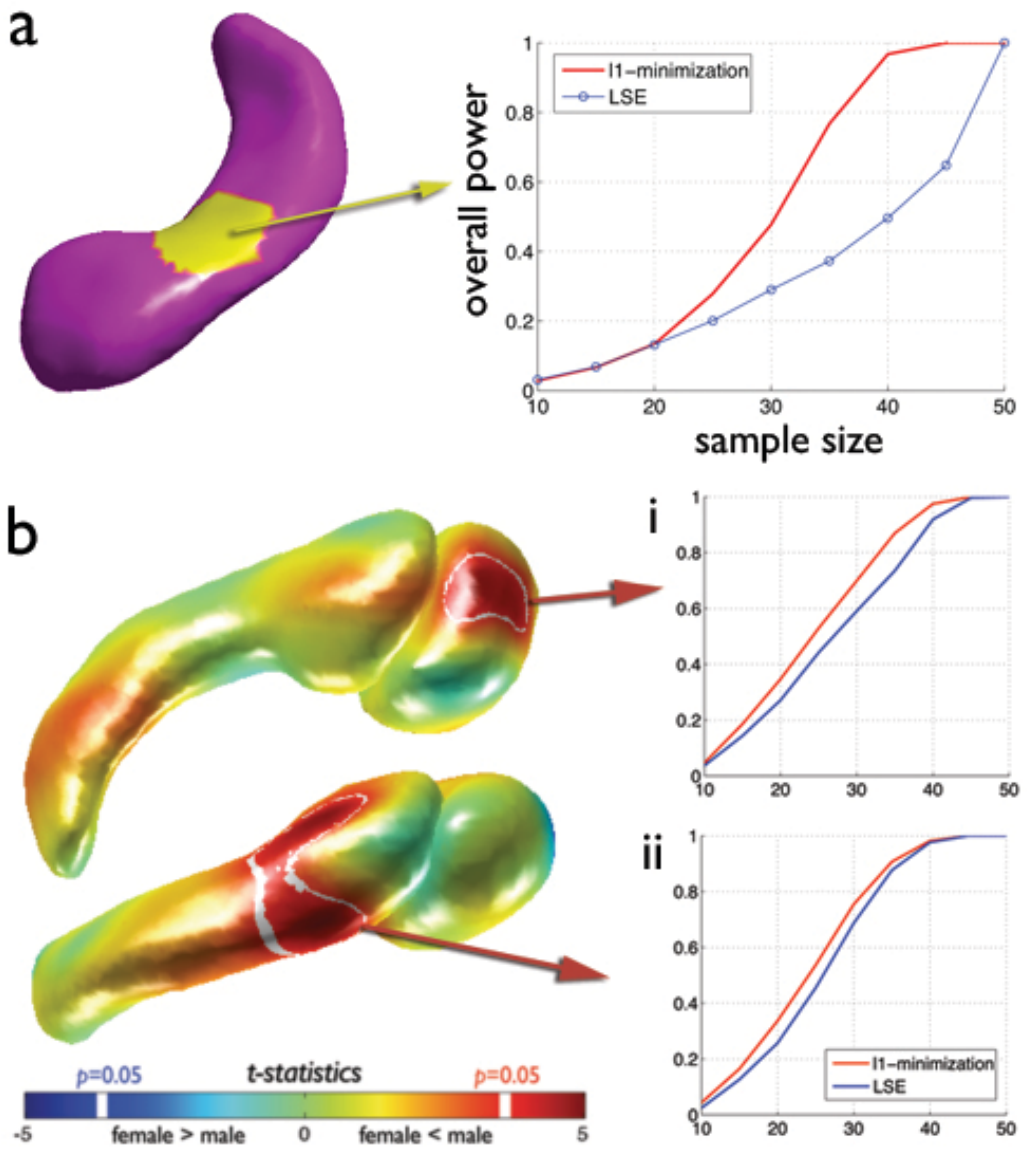

Figure 3: Overall power curves are given over sample sizes for simulation data (a) and real data (b). For simulations, 5-ring neghbiors of a certain vertex on the left hippocampus surface were taken as true signal region (yellow area). For real data, the regions showed significant gender effect with 52 total subjects on the right amygdala (i) and the left hippocampus (ii) were taken as signal regions (highlighted by white boundaries).

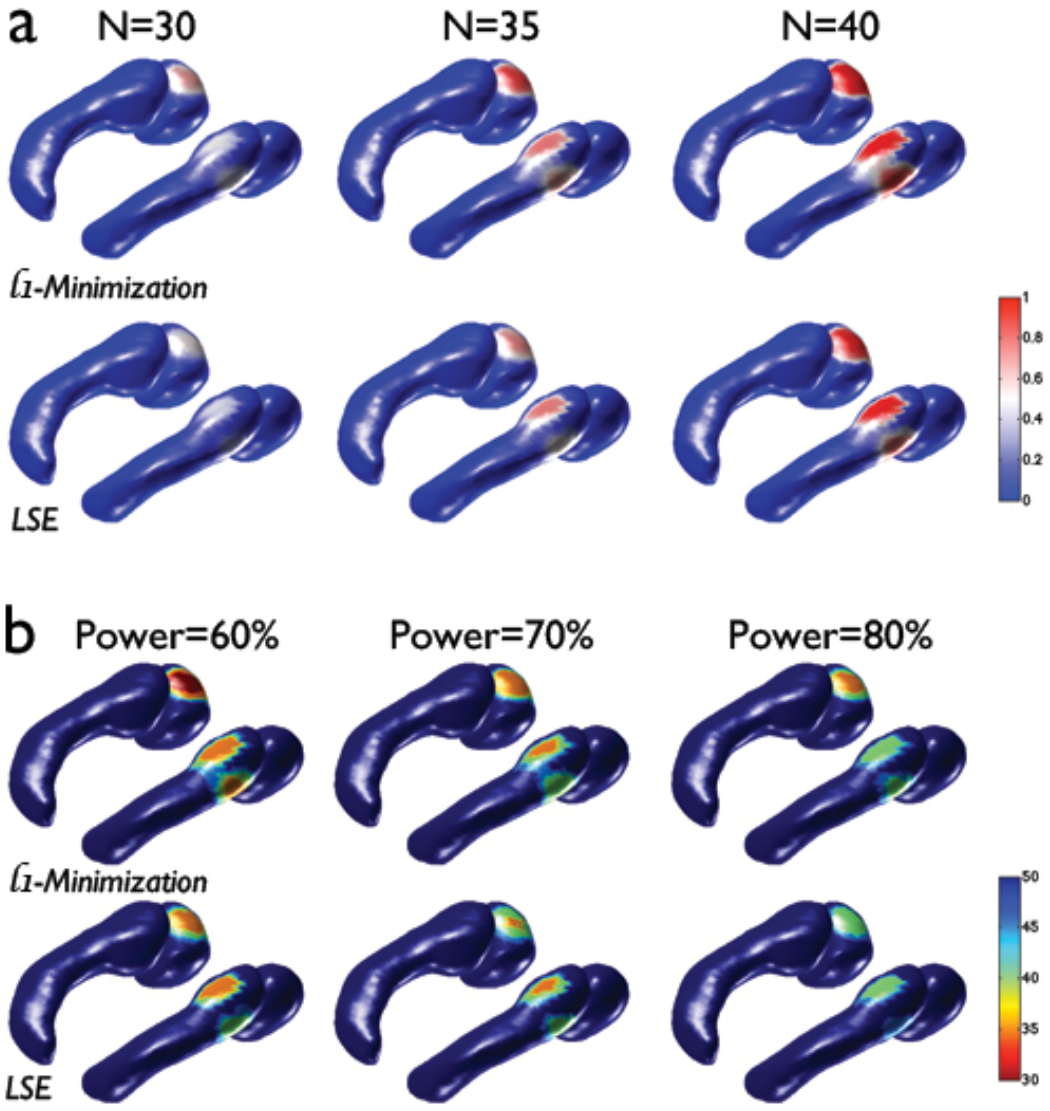

Figure 4: Local power maps (a) and sample size maps (b) are shown for given samplze sizes and power. l1-minimization and LSE are compared in each measures.

Bibliography

1. Hayasaka S., Peiffer A. M., Hugenschmidt C. E., Laurienti P. J. (2007) Power and sample size calculation for neuroimaging studies by non-central random field theory. Neuroimage 37: 721-730.
2. Kim S.-G., Chung M.K., Schaefer S.M., van Reekum C., Davidson R.J. (2012) Sparse Shape Representation using the Laplace-Beltrami Eigenfunctions and Its Application to Modeling Subcortical Structures. IEEE workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA). Colorado.
3. Smith S. M. (2002) Fast robust automated brain extraction. Hum Brain Mapp 17: 143-155.

4. Avants B. B., Epstein C. L., Grossman M., Gee J. C. (2008) Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med Image Anal 12: 26-41.