Published in final edited form as:

Proc IEEE Int Symp Biomed Imaging. 2018 April ; 2018: 6–9. doi:10.1109/ISBI.2018.8363511.

JOINT EXPLORATION AND MINING OF MEMORY-RELEVANT BRAIN ANATOMIC AND CONNECTOMIC PATTERNS VIA A THREE-WAY ASSOCIATION MODEL

Jingwen Yan[1,★], **Kefei Liu**[2], **Huang Li**[3], **Enrico Amico**[4], **Shannon L. Risacher**[2], **Yu-chien Wu**[2], **Shiaofen Fang**[3], **Olaf Sporns**[5], **Andrew J. Saykin**[2], **Joaquín Goñi**[4], and **Li Shen**[6,★]

[1]BioHealth Informatics, Indiana University Indianapolis, IN, USA

[2]Radiology and Imaging Sciences, Indiana University School of Medicine, IN, USA

[3]Computer and Information Science, Purdue University Indianapolis, IN, USA

[4]Industrial Engineering, Purdue University West-Lafayette, IN, USA

[5]Psychological and Brain Sciences, Indiana University Bloomington, IN, USA

[6]Biostatistics, Epidemiology and Informatics, University of Pennsylvania, PA, USA

Abstract

Early change in memory performance is a key symptom of many brain diseases, but its underlying mechanism remains largely unknown. While structural MRI has been playing an essential role in revealing potentially relevant brain regions, increasing availability of diffusion MRI data (e.g., Human Connectome Project (HCP)) provides excellent opportunities for exploration of their complex coordination. Given the complementary information held in these two imaging modalities, we hypothesize that studying them as a whole, rather than individually, and exploring their association will provide us valuable insights of the memory mechanism. However, many existing association methods, such as sparse canonical correlation analysis (SCCA), only manage to handle two-way association and thus cannot guarantee the selection of biomarkers and associations to be memory relevant. To overcome this limitation, we propose a new outcome-relevant SCCA model (OSCCA) together with a new algorithm to enable the three-way associations among brain connectivity, anatomic structure and episodic memory performance. In comparison with traditional SCCA, we demonstrate the effectiveness of our model with both synthetic and real data from the HCP cohort.

Index Terms

three way sparse association; brain connectome; memory performance

★Correspondence to Li Shen (Li.Shen@pennmedicine.upenn.edu) or Jingwen Yan (jingyan@iupui.edu).

1. INTRODUCTION

Memory loss is among the earliest symptoms for many brain diseases and its clinical evaluation is essential for early diagnosis. Although the main cause is known due to the physical brain changes, the underlying mechanism remains largely unknown. Brain imaging has been playing an essential role in tackling this challenge. For example, using structural magnetic resonance imaging (MRI), anatomical changes in hippocampus and entorhinal cortex have been found to be highly associated with memory performance [1]. Recent availability of diffusion MRI data (e.g., Human Connectome Project (HCP) (www.humanconnectome.org)) provides excellent opportunities for exploration of the complex coordination among brain regions underlying memory performances.

We propose to jointly exploit anatomic and connectomic data and explore their association for more valuable insights of the memory mechanism. Compared to sparse regression models that search for multiple-to-one relationship, bi-multivariate sparse association models provide a more powerful option to explore the multi-to-multiple relationship between anatomic and connectomic features[2]. However, existing association methods, e.g. sparse canonical correlation analysis (SCCA) [3], can only handle two-way associations. The biomarkers and their associations identified are not necessarily related to memory performance unless the input features include only candidate biomarkers [4].

To overcome this limitation, we propose an outcome-relevant SCCA (OSCCA) to explore the three-way association among connectivity, anatomic structure and memory performance. We perform an empirical comparison between OSCCA algorithm and a widely used SCCA implementation [3]. Based on both synthetic data and real data from the HCP cohort, the empirical results show that the proposed OSCCA outperformed traditional SCCA on both association performance and variable selection.

2. OUTCOME-RELEVANT SCCA (OSCCA)

In this section, we denote vectors as boldface lowercase letters and matrices as boldface uppercase ones. For a matrix $\mathbf{M} = (m_{ij})$, we denote its i-th row and j-th column to $\mathbf{m}^i$ and $\mathbf{m}_j$ respectively. Let $\mathbf{X} = [\mathbf{x}_1; \ldots; \mathbf{x}_n] \subseteq \Re^{n \times p}$ be the anatomic structure data and $\mathbf{Y} = [\mathbf{y}_1; \ldots; \mathbf{y}_n] \subseteq \Re^{n \times q}$ be the connectivity data, where n is the number of participants, p and q are the number of regions of interest (ROIs) and fibers.

Sparse canonical correlation analysis (SCCA) is a bi-multivariate method that explores the linear transformations of variables **X** and **Y** to achieve the maximal correlation between **Xu** and **Yv**, which can be formulated as:

$$\max_{\mathbf{u},\mathbf{v}} \mathbf{u}^T\mathbf{X}^T\mathbf{Y}\mathbf{v} \qquad (1)$$
$$s.t.\ \mathbf{u}^T\mathbf{X}^T\mathbf{X}\mathbf{u} = 1, \mathbf{v}^T\mathbf{Y}^T\mathbf{Y}\mathbf{v} = 1, ||\mathbf{u}||_1 \leq c_1, ||\mathbf{v}||_1 \leq c_2$$

where **u** and **v** are canonical loadings or weights, reflecting the contribution of each feature in the identified associations. The L_1 penalty over **u** and **v** encourages the global sparsity for

easy interpretation where only a small set of features in both modalities can be selected and thus can avoid the overfitting problem when the feature size outnumbers the sample size.

However, SCCA can only manage to explore the two-way associations. It can identify a list of highly associated ROIs and fibers, but they may not all be related to memory. To address this problem, we propose a novel outcome-relevant SCCA (denoted as OSCCA) model to enable three-way association analysis. An extra penalty term $P(\mathbf{u})$ will be introduced to incorporate the outcome, or the 3rd modality, to help yield outcome-relevant biomarkers and associations.

$$P(\mathbf{u}) = ||\mathbf{u}||_O = \mathbf{u}^T\mathbf{X}^T\mathbf{L}\mathbf{X}\mathbf{u} = \frac{1}{2}\sum_{i,j=1}^{p}\mathbf{S}_{ij}\left|\left|(\mathbf{X}^i\mathbf{u} - \mathbf{X}^j\mathbf{u})\right|\right|_2^2 \quad (2)$$

Here, $\mathbf{S} \subseteq \Re^{n\times n}$ is the similarity matrix. Let $\mathbf{M} \subseteq \Re^{n\times k}$ be the outcome data, where k is the total number of outcome features. $\mathbf{S}_{ij}$ indicates the similarity between subjects i and j based on their outcome profiles $\mathbf{M}^i$ and $\mathbf{M}^j$. $\mathbf{L}$ is the Laplacian matrix of $\mathbf{S}$. This new penalty term $||\cdot||_O$ incorporates the similarity matrix to encourage the closeness of subjects with similar outcome profiles in $\mathbf{Xu}$ space. L_1 norm is kept to ascertain the selection of only a few brain ROIs and fibers. This model is expected to jointly perform bi-multivariate association and outcome-relevant biomarker selection. The final objective function of OSCCA can be formulated as follows:

$$\max_{\mathbf{u},\mathbf{v}} \mathbf{u}^T\mathbf{X}^T\mathbf{Y}\mathbf{v} - \frac{\beta}{2}P(\mathbf{u}) \quad (3)$$
$$s.t.\ \mathbf{u}^T\mathbf{X}^T\mathbf{X}\mathbf{u} = 1, \mathbf{v}^T\mathbf{Y}^T\mathbf{Y}\mathbf{v} = 1, ||\mathbf{u}||_1 \le c_1, ||\mathbf{v}||_1 \le c_2$$

Note that $||\cdot||_O$ is only applied to one modality because maximizing the correlation between $\mathbf{Xu}$ and $\mathbf{Yv}$ will finally encourage $\mathbf{Yv}$ to yield a similar pattern as $\mathbf{Xu}$. This can significantly reduce the computation cost as there will be less parameters to tune and no need to calculate $\mathbf{Y}^T\mathbf{L}\mathbf{Y}$ in the following algorithm. Eq. (3) is known as a bi-convex problem, which can be easily solved using an alternating algorithm [3]. By fixing $\mathbf{u}$ and $\mathbf{v}$ respectively, we will have the following two minimization problems shown in Eq. (4) and (5). To simplify the problem, we follow [3] and set $\mathbf{X}^\mathbf{T}\mathbf{X} = \mathbf{I}$ and $\mathbf{Y}^\mathbf{T}\mathbf{Y} = \mathbf{I}$.

$$\min_{\mathbf{u}} -\mathbf{u}^T\mathbf{X}^T\mathbf{Y}\mathbf{v} + \frac{\beta}{2}P(\mathbf{u}), s.t.\ \mathbf{u}^T\mathbf{u} = 1, ||\mathbf{u}||_1 \le c_1 \quad (4)$$

$$\min_{\mathbf{v}} -\mathbf{u}^T\mathbf{X}^T\mathbf{Y}\mathbf{v}, s.t.\ \mathbf{v}^T\mathbf{v} = 1, ||\mathbf{v}||_1 \le c_2 \quad (5)$$

Eq. (4) can be solved using the Nesterov's accelerated algorithm [5] and Eq. 5 can be solved using Algorithm 3 in [3]. Let $g(\mathbf{u}) = -\mathbf{u}^T\mathbf{X}^T\mathbf{Y}\mathbf{v} + \frac{\beta}{2}P(\mathbf{u})$ in Eq. 4. The Lipschitz constant of $g'(\mathbf{u})$ is the spectral norm of $g''(\mathbf{u}) = \beta\mathbf{X}^T\mathbf{L}\mathbf{X}$ [6], which can be used as the stepsize to further accelerate the Nesterov's algorithm by removing the iterative searching (Steps 4–10 of Algorithm 3 in [5]). The detailed proof is not included due to the space limitation. Algorithm 1 summarizes the optimization procedure. π_G is the Euclidean projection function defined in [5]. The convergence is based on the value changes of canonical loadings, where $|\mathbf{u}_t - \mathbf{u}_{t-1}| \leq 10^{-4}$ and $|\mathbf{v}_t - \mathbf{v}_{t-1}| \leq 10^{-4}$ were used as stop criteria.

Algorithm 1

Outcome-relevant SCCA (OSCCA)

Require:

$\mathbf{X} = \{x_1, \ldots, x_n\}$, $\mathbf{Y} = \{y_1, \ldots, y_n\}$, $\mathbf{L} \subseteq \Re^{n \times n}$

Ensure:

Canonical vectors $\mathbf{u}$ and $\mathbf{v}$.

1: $t = 1$, Initialize $\mathbf{u}_0 \in \Re^{p \times 1}$, $\mathbf{v}_0 \in \Re^{q \times 1}$;

2: **while** not converge **do**

3: $i = 1$, Initialize $\mathbf{w}_1 = \mathbf{w}_0 = \mathbf{u}_{t-1}$, $k_{-1} = 0$, $k_0 = 1$, $\ell = \|\beta\mathbf{X}^T\mathbf{L}\mathbf{X}\|_2$

4: **while** not converge **do**

5: Set $\alpha_i = \frac{k_{i-2} - 1}{k_{i-1}}$, $\mathbf{s}_i = \mathbf{w}_i + \alpha_i(\mathbf{w}_i - \mathbf{w}_{i-1})$

6: $\mathbf{w}_{i+1} = \pi_G(\mathbf{s}_i - \frac{1}{\ell}(-\mathbf{X}^T\mathbf{Y}\mathbf{v}_{t-1} + 2\beta\mathbf{X}^T\mathbf{L}\mathbf{X}\mathbf{s}_i))$

7: Set $k_i = \frac{(1 + \sqrt{1 + 4k_{i-1}^2})}{2}$

8: $i = i + 1$.

9: **end while**

10: $\mathbf{u}_t = \mathbf{w}_i$

11: Solve Eq. (5) using step 2(a) in Algorithm 3 in [3] and obtain $\mathbf{v}_t$;

12: $t = t + 1$.

13: **end while**

3. RESULTS

In this section, we evaluate OSCCA model and compare it with SCCA [3]. We use 5-fold nested cross-validation to tune the parameters and apply the same fold partition in both methods for fair comparison.

3.1. Results on Synthetic Data

We generated the anatomic data $\mathbf{X}$ and connectivity data $\mathbf{Y}$ ($n = 90$, $p = 100$, $q = 120$), assuming the subjects belonging to three groups. The similarity matrix is built that the subjects in the same group have value 1 and −1 otherwise. We set 10 anatomic and 12

connectivity features to be related. 5 anatomic features are outcome-relevant and generated as $\mathbf{x}_j = u_j\mathbf{z} + \mathbf{n}_j$, $j = 1, \ldots, 5$, where $\mathbf{n}_j \sim \mathcal{N}(0, \sigma^2\mathbf{I}_{n\times n})$ is a noise vector and $\mathbf{z}$ is a random vector generated using a 3-component Gaussian mixture model. Subjects form three clusters (30 each) with centers $\mu = -5, 0, -5$ respectively. The other 5 anatomic features are outcome-irrelevant and generated as $\mathbf{x}_j = u_j\mathbf{w} + \mathbf{n}_j$, $j = 6, \ldots, 10$, where $\mathbf{w}$ are i.i.d. drawn from a zero-mean unit-variance Gaussian distribution. 6 connectivity features are outcome-relevant and generated as $\mathbf{y}_\ell = v_\ell\mathbf{z}' + \mathbf{n}_\ell$ $\ell = 1, \ldots, 6$, where $\mathbf{z}' = c\mathbf{z} + \sqrt{1-c^2}\varepsilon$ and $c = 0.8$ is the true correlation. ε is a random vector independent of $\mathbf{z}$ and have the same Euclidean norm such that corr $(\mathbf{z}', \mathbf{z}) \approx c$. The other 6 connectivity features are outcome-irrelevant and generated as $\mathbf{y}_\ell = v_\ell\mathbf{w}' + \mathbf{n}_\ell$ $\ell = 7, \ldots, 12$. $\mathbf{w}'$ is generated similarly as $\mathbf{z}'$. All other features are generated according to $\mathcal{N}(0, \sigma^2\mathbf{I}_{n\times n})$. In both training and test data, we observed comparable association performance of OSCCA and SCCA, both close to the ground truth. However, most features identified in OSCCA are outcome-relevant, whereas those in SCCA are mostly outcome-irrelevant. We further evaluated the classification performance of the canonical components, **Xu** and **Yv**, for each pair of groups. For OSCCA, **u** and **v** obtained from training data were used to generate **Xu** and **Yv** for testing data. With these two features, support vector machine (SVM) was then applied (LIBSVM package) (OSCCA+SVM). We evaluated the performance of SCCA using the same strategy (SCCA+SVM). In addition, the classification power of original data was examined using elastic net (Elnet). Table. 1 summarizes 5-fold classification performance in test data. The canonical components generated by OSCCA are more discriminative than those by SCCA. It also outperforms elastic net and shows a potential in both feature extraction and multi-class classification.

3.2. Results on Real Data

From the HCP (www.humanconnectome.org), we downloaded the structural MRI (sMRI), diffusion MRI (dMRI) and memory performance data of 93 unrelated subjects (19 males/74 females, age: 29.91 $\pm$ 2.91). The dMRI data was first denoised and corrected for motion and distortion [7]. Tractography was performed in Camino [8] based on white matter fiber orientation distribution function (ODF). Streamlines were modeled with a multi-tensor modeling approach, where voxels will fit up to two fiber orientations. Second, the sMRI image was registered to the b0 volume of dMRI data using the FNIRT toolbox [9] and 278 brain regions of interest (ROIs) were extracted following [10]. The final networks were constructed using streamlines going through white matter and connecting ROIs. Here, we use the fiber density for the following association analysis, which is the fraction between number of streamlines and the average surface of grey-matter regions i and j. 20346 connectivity measures with non-zero variation were included. For anatomic data, we downloaded the Freesurfer results from HCP and extracted cortical thickness and volume measures from 90 ROIs. For memory performance, we focused on episodic memory and downloaded three scores: one from Picture Sequence Memory Test and two from PennWord Memory Test. We calculated the Euclidean distance between subjects based on these scores and then generated the similarity measures by taking the element-wise inverse. All the measures were adjusted for the age and gender, with intracranial volume as extra covariate for brain connectivity and structural measures.

Fig. 1 shows the 5-fold canonical loadings, indicating a set of associated anatomic and connectomic markers relevant to episodic memory. Lateral ventricle, accumbens, and middle posterior corpus callosum (1(a)) were found to be memory-relevant and be attributed to the wiring within subcortical regions. Around 700 connectivity measures are selected and shown in Fig. 1(b) are the canonical loadings in one example fold. It shows that most of the significant connectivities are within subcortical regions. Mapping 278 ROIs to Automated Anatomical Labeling (AAL) atlas, memory-related ROIs, e.g. bilateral hippocampi, are found to be linked by these significant fibers. But the hub regions are postcentral, temporal, and frontal lobe instead. This suggests the essential role of some key brain regions in both anatomic variation and memory performance.

4. CONCLUSION

We proposed a new association model to explore the three-way association among anatomy, connectivity and memory. OSCCA showed superior performance than SCCA in both association and outcome-relevant feature selection. In real data, OSCCA identified a small set of associated regions and fibers. This suggests the potential role of brain wiring mechanism in both anatomic variation and memory.

Acknowledgments

This research was supported by NIH grants R01 AG053993, R01 EB022574, R01 LM011360, R01MH108467, R01 AG019771, P30 AG010133 and K01 AG049050. This project was also funded, in part, with support from the Indiana Clinical and Translational Sciences Institute funded, in part by Grant Number UL1TR001108 from the National Institutes of Health, National Center for Advancing Translational Sciences, Clinical and Translational Sciences Award. Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

References

1. O'Hanlon ErikHowley SarahPrasad SarahMc-Grath JaneLeemans AlexanderMcDonald ColmGaravan HughMurphy Kieran C. Multimodal mri reveals structural connectivity differences in 22q11 deletion syndrome related to impaired spatial working memory. Human Brain Mapping. 2016; 37(12):4689–4705. [PubMed: 27511297]
2. Lin D, Calhoun VD, Wang YP. Correspondence between fMRI and SNP data by group sparse canonical correlation analysis. Med Image Anal. 2013
3. Witten DM, Tibshirani R, Hastie T. A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis. Biostatistics. 2009; 10(3):515–34. [PubMed: 19377034]
4. Du L, Yan JW, Kim S, Risacher SL, Huang H, Inlow M, Moore JH, Saykin AJ, Shen L. Alzheimers Dis Neuroimaging Initia. A novel structure-aware sparse learning algorithm for brain imaging genetics. Medical Image Computing and Computer-Assisted Intervention - Miccai 2014, Pt Iii. 2014; 8675:329–336.
5. Liu JunJi ShuiwangYe Jieping. Multi-task feature learning via efficient l2,1-norm minimization. Proceedings of the twenty-fifth conference on uncertainty in artificial intelligence; 2009; AUAI Press; 339–348.
6. Beck AmirTeboulle Marc. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM journal on imaging sciences. 2009; 2(1):183–202.

7. Manjón José V, Coupé PierrickConcha LuisBuades AntonioLouis Collins D, Robles Montserrat. Diffusion weighted image denoising using overcomplete local pca. PloS one. 2013; 8(9):e73021. [PubMed: 24019889]
8. Cook PA, Bai Y, Nedjati-Gilani SKKS, Seunarine KK, Hall MG, Parker GJ, Alexander DC. Camino: open-source diffusion-mri reconstruction and processing. 14th scientific meeting of the international society for magnetic resonance in medicine; Seattle WA, USA. 2006; 2759
9. Jenkinson MarkBeckmann Christian F, Behrens Timothy EJ, Woolrich Mark W, Smith Stephen M. Fsl. Neuroimage. 2012; 62(2):782–790. [PubMed: 21979382]
10. Shen XilinTokoglu F, Papademetris XeniosTodd Constable R. Groupwise whole-brain parcellation from resting-state fmri data for network node identification. Neuroimage. 2013; 82:403–415. [PubMed: 23747961]

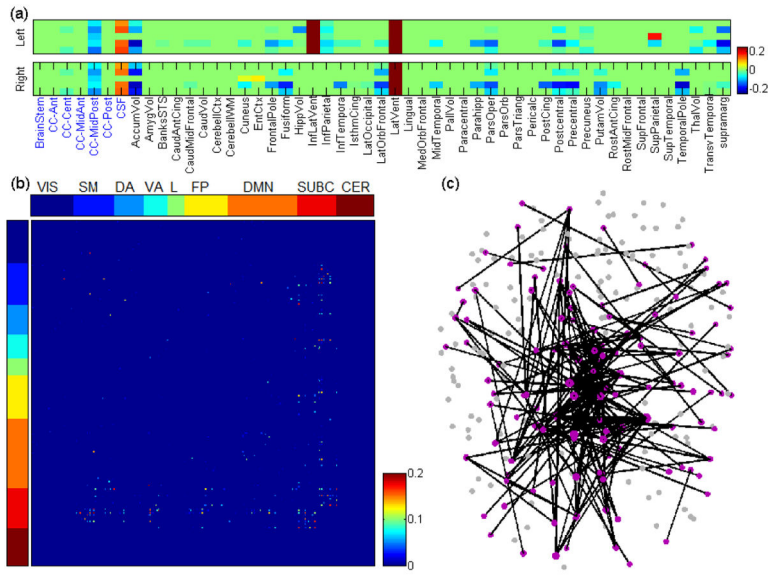

Fig. 1.

Brain anatomic and connectomic markers associated with episodic memory. (a) 5-fold canonical loadings of brain anatomic features. Regions labeled blue are unilateral. (b) Canonical loadings of brain connectomic features in one example fold. Top panel indicates different regions: Visual (VIS), Somato-Motor (SM), Dorsal Attention (DA), Ventral Attention (VA), Limbic system (L), Fronto-Parietal (FP), Default Mode Network (DMN), subcortical regions (SUBC) and cerebellum (CER). (c) brain connectomic features with non-zero weights in (b) mapped to brain.

Table 1

Cross-validation performance on synthetic data: measured by the accuracy in pair-wise classification tasks

	OSCCA+SVM			SCCA+SVM			Elnet		
	1/2	**1/3**	**2/3**	**1/2**	**1/3**	**2/3**	**1/2**	**1/3**	**2/3**
f1	1.00	1.00	1.00	0.92	0.50	0.75	1.00	1.00	0.67
f2	1.00	1.00	1.00	0.92	0.50	0.67	0.83	1.00	0.83
f3	1.00	1.00	1.00	0.83	0.50	0.50	0.92	1.00	0.75
f4	1.00	1.00	1.00	0.92	0.50	0.92	0.92	1.00	1.00
f5	1.00	1.00	1.00	0.67	0.50	0.75	1.00	1.00	0.75
Mean	1.00	1.00	1.00	0.85	0.50	0.72	0.93	1.00	0.80