Thalamocortical Connection Topography Mapping in Human by Co-clustering

Hewei Cheng1,2, Jie Liu3,* and Yiying Tao1,* b

1Department of Biomedical Engineering, School of Bioinformatics, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
2Center for Biomedical Image Computing and Analytics, Department of Radiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia 19104, USA
3Research Institute of Education Development, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

*E-mail: ai.jieliu218@163.com; b15552092163@163.com

Abstract. Connection topography mapping is crucial for understanding how information is processed in the brain, which is an essential precursor for revealing principles of brain organization. However, existing connectomic mapping methods are dependent on prior knowledge, or not completely driven by data. Accordingly, the constructed connection topographies by these methods are biased towards hypotheses, or deviate from data. For these challenges, we propose a novel co-clustering based method for connection topography mapping in a fully data-driven manner. The proposed method aims to construct the connection topography between two ROIs of a certain neural circuit in consideration by leveraging the power of co-clustering. More precisely, the proposed method parcellates one ROI into subregions and identified their respective connected subregions from the other ROI simultaneously. The effectiveness of our method was validated on the mapping of the human thalamocortical system for 57 subjects based on their resting state fMRI data. The validation experiment results have demonstrated that our method can construct neurobiologically meaningful thalamocortical connection topography. Compared with existing methods, our method yields more meaningful and interpretable connection topography.

1. Introduction

Brain connectivity analysis, derived from resting state fMRI and DTI, has shown promising role in non-invasively revealing the connectomic organization of the human brain [1], such as organization of the subcortical-cortical system. The thalamocortical system is a subsystem of the subcortical-cortical system. The thalamus within the thalamocortical system is responsible for varieties of functions, such as motor control, sensory relay, attention and awareness, which are supported by its globally distributed cortical connections [2]. For this reason, connectivity mapping of the thalamocortical system is vitally important for unveiling neural mechanisms of this system. So far, however, the majority of previous work is not fully automatic, which can’t guarantee that the mapped connectivity of the thalamocortical system is neurobiologically meaningful. In addition, the mapped thalamocortical connectivity system is optimal for hypothesis-driven or partially data-driven approaches, rather than for a completely data-driven method. Such approaches have two main shortcomings. Firstly, hypothesis-driven methods require the usage of prior cortical targets to map the
thalamocortical connectivity system, causing biases towards hypotheses. Furthermore, a large number of existing data-driven methods don’t have a tightly unified framework, which can’t effectively utilize the full information from data. Consequently, the thalamocortical connectivity system, mapped through these methods, might deviate from data. To address these issues, we resort to a fully data-driven strategy for mapping neurobiologically interpretable thalamocortical connectivity system.

In this work, we propose a novel method to map the intrinsic functional connectivity of the human thalamocortical system using co-clustering in a fully data-driven manner. This approach simultaneously parcellates thalamic subregions and their connected cortical networks from resting state fMRI, i.e., fully automatically mapping the intrinsic functional connectivity of the thalamocortical system. Particularly, we adopt spectral co-clustering to map the connectivity of the thalamocortical system, which is achieved by clustering rows and columns of the thalamocortical functional similarity matrix calculated based on functional connectivity between functional signals. The used spectral co-clustering is a bipartite graph partitioning problem, which can be solved by classic spectral clustering [3].

The remaining paper is organized as follows. Firstly, we review previous connectopic mapping approaches. Secondly, we describe our proposed co-clustering based connectopic mapping method. Thirdly, we present experimental results. Lastly, we give closing remarks in the section of Discussion and conclusions.

2. Previous work
There are a large number of connectopic mapping methods, which have been proposed to unveil the connectopic organization of two brain regions in human, and connectivity mapping of the thalamocortical/subcortical-cortical systems has gained increasing interest in recent decades [4-6]. These previous methods are mainly derived from techniques, including prior connectivity mapping, and independent component analysis (ICA). However, such methods, i.e., prior connectivity mapping based methods, and ICA based methods, have one of the weaknesses stated in section 1.

The prior connectivity mapping based methods construct the subcortical-cortical connectivity system by labeling each subcortical voxel to one of the cortical target regions with strongest connection, and the employed cortical target regions need to be determined in advance [5-16]. Obviously, these methods require prior defined targets, which might result in biases due to incompletely reliable prior knowledge. In addition to the prior connectivity mapping based methods, the existing ICA based methods use subcortical-cortical connections to map connectivity of the subcortical-cortical system [2, 17, 18]. These existing ICA based methods are not fully data-driven approaches, and require postprocessing, such as thresholding. For these challenges, we propose a fully automatic connectopic mapping method not depending on prior hypotheses, which is completely a data-driven approach by leveraging the power of co-clustering. In this method, we adopt the state-of-the-art spectral co-clustering algorithm to map connection topography of the human brain.

3. Proposed method
The proposed approach consists of three fundamental steps, illustrated in figure 1. In the first step, two regions of interest (ROIs) are extracted for mapping their connectopic topography (figure 1(a)). Most commonly, one ROI is a subcortical structure, and the other is its cortical circuit. In the second step, functional similarity matrix is calculated between the two ROIs, and corresponding bipartite graph model is also built for co-clustering (figure 1(b)). In the third step, the modeled bipartite graph is co-partitioned for connectopic mapping (figure 1(c)).
3.1. Regions of interest extraction
In order to obtain highly precise ROIs, the ROIs, i.e., brain structure and its main circuit under consideration, are segmented from each subject’s structural magnetic resonance imaging (sMRI) data. Specifically, for each subject $s_i, i = 1, 2, ..., N$ in a dataset, FreeSurfer is adopted to extract subcortical structures and cerebral cortex from his or her sMRI data [19]. For example, the extracted substructure thalamus and its cortical circuit of one randomly selected subject are shown in figure 1(a).

3.2. Bipartite graph modeling for co-clustering
In the proposed method, the spectral co-clustering algorithm is adopted to map connection topography [3]. For this novel connectopic mapping method, it is firstly required to build the spectral co-clustering’s bipartite graph model $G_d = (V_r, V_c, E_d)$, and at the same time calculate functional similarity matrix $M$ between two ROIs of a certain neural circuit in consideration. Here, the bipartite graph is constructed by modeling all the voxels within each of the two ROIs as graph node sets $V_r = \{u_i | i = 1, 2, ..., n_r, n_r = |V_r|\}$ and $V_c = \{v_i | i = 1, 2, ..., n_c, n_c = |V_c|\}$, respectively, $|\cdot|$ is the cardinality of a set; and connecting each pair of voxels from the two ROIs comprising graph edge set $E_d = \{(u, v) | u \in V_r, v \in V_c\}$. The graph edge weights $\{m(u, v) | u \in V_r, v \in V_c\}$ form the functional similarity matrix $M \in \mathbb{R}^{n_r \times n_c}$.

In this paper, an edge weight is defined by measuring functional similarity between functional signals. The functional similarity is usually defined based on the functional connectivity measure, i.e., Pearson correlation coefficient between functional signals [20]

$$r(u, v) = \frac{(1/T) \sum_{t=1}^{T} (I(u, t) - \bar{I}(u))(I(v, t) - \bar{I}(v))}{\sqrt{\sum_{t=1}^{T} (I(u, t) - \bar{I}(u))^2} \sqrt{\sum_{t=1}^{T} (I(v, t) - \bar{I}(v))^2}}.$$ (1)

Where $u \in V_r, v \in V_c, I(u, t)$ and $I(v, t)$ are functional signals of voxels $u$ and $v$ at time point $t$, $T$ denotes the number of time points, $\bar{I}(\cdot) = (1/T) \sum_{t=1}^{T} I(\cdot, t)$, and $S_I(\cdot) = \sqrt{1/(T-1) \sum_{t=1}^{T} (I(\cdot, t) - \bar{I}(\cdot))^2}$. The bipartite graph model requires edge weights to be non-negative, so that the graph edge weight, i.e., functional similarity here, based on functional connectivity measures can be defined as
\[ m(u, v) = \begin{cases} r(u, v) & \text{if } r(u, v) > 0 \\ 0 & \text{otherwise} \end{cases} \] (2)

Where \( u \in V_r, v \in V_c \), \( r(u, v) \) is the functional connectivity measure between voxels \( u \) and \( v \) as defined in equation (1). The defined graph edge weights \( \{m(u, v)\} u \in V_r, v \in V_c \) form the functional similarity matrix. Figure 1(b) shows the heat map of the functional similarity matrix \( M \).

3.3. Bipartite graph co-partitioning for connectopic mapping

Subsequently, the established bipartite graph model \( G_d = (V_r, V_c, E_d) \) is solved by using spectral co-clustering algorithm for connectopic mapping. Particularly, the bipartite graph model is firstly transformed into an undirected weighted graph model \( G_o = (V, A, E_o) \). The transformed graph \( G_o = (V, A, E_o) \) is constructed by modeling all the voxels from two ROIs of a certain neural circuit in consideration as graph node set \( V = V_r + V_c \) and connecting each pair of graph nodes consisting of graph edge set \( E_o = \{(u, v)|u, v \in V\} \), and the number of elements in the graph node set \( V \) is \( n = n_r + n_c \). The graph edge weights \( \{a(u, v)|u, v \in V\} \) form an affinity matrix, defined as

\[ A = \begin{bmatrix} 0 & M \\ M^T & 0 \end{bmatrix} \] (3)

Where \( M \) is the functional similarity matrix, \( M^T \) denotes the transpose of a matrix, and \( A \in \mathbb{R}^{n \times n} \) is a symmetric block matrix.

The bipartite graph co-partitioning problem is then transformed into a classical spectral clustering problem. Suppose the graph \( G_o \) is partitioned into \( k \) clusters \( V = \{V_1, V_2, ..., V_k\} \). In the spectral clustering problem, the normalized cut on the graph \( G_o = (V, A, E_o) \) is defined as

\[ \text{Ncut}(G_o) = \min_{V_1, V_2, ..., V_k} \sum_{i=1}^{k} \frac{\text{cut}(V_i, V - V_i)}{\text{assoc}(V_i, V)} \] (4)

Where \( \text{cut}(V_i, V - V_i) = \sum_{u \in V_i, v \in V - V_i} a(u, v) \), \( \text{assoc}(V_i, V) = \sum_{u \in V_i, v \in V} a(u, v) \). \( a(u, v) \) is the value of an element of the affinity matrix \( A \) at location \( (u, v) \). According to the \( k \)-cluster partition \( V = \{V_1, V_2, ..., V_k\} \), two ROIs, from a certain neural circuit under consideration, is correspondingly parcellated into \( k \)-pair connected subregions. More precisely, for each pair of the connected subregions, one subregion is parcellated from one ROI, and the other is parcellated from the other ROI.

The \( k \)-pair connected subregions is the final mapped connection topography between the two ROIs. Intuitively, co-clustering based connectopic mapping is achieved by clustering the rows and columns of the functional similarity matrix \( M \) between the two ROIs simultaneously. Figure 1(c-③) is the reordered functional similarity matrix using co-clustering, which shows that the reordered functional similarity matrix contains clearly identifiable clusters in the row direction and in the column direction, respectively. For instance, figure 1(c-②) shows the mapped thalamocortical connection topography using the proposed method, i.e., identified thalamic subregions and their respective connected cortical networks.

4. Experimental results

4.1. Data and preprocessing

To validate the proposed method, 57 healthy subjects’ resting state fMRI data of the first scan session from Beijing Normal University 1 (BNU 1) dataset were employed for mapping thalamocortical connection topography, which are publicly available from Consortium for Reliability and Reproducibility (CoRR) [21]. Although the human thalamus is comprised of multiple functionally heterogeneous subregions, which was simply mapped into 4-pair connected subregions according to non-human primate studies [15], considering that the fMRI dataset have relatively low spatial resolution of 3.125×3.125×4.2 mm³, each consisting of 200 volumes at repetition time (TR) = 2 s.
The fMRI data was preprocessed as follows: (1) removal of first 5 volumes to allow for signal equilibration, (2) head motion correction, (3) global 4D mean intensity normalization, (4) temporally filtering (0.01-0.08Hz) after intensity scaling, (5) regression out of a series of nuisance signals including signal averaged over the white matter, signal averaged over the cerebrospinal fluid, global signal averaged over the whole brain, and six motion parameters, (6) spatially smoothing with a 6 mm full width at half-maximum (FWHM) Gaussian kernel, and nonlinear registration to MNI152 template with spatial resolution 3×3×3 mm³. The preprocessing procedure was carried out by using tools from the FMRIB Software Library (FSL; https://fsl.fmrib.ox.ac.uk/fsl).

4.2. Thalamocortical connection topography and its validation

To examine and validate the mapped connection topography, figure 2 presents the results of a randomly selected subject. As shown in figure 1, the bilateral human thalamocortical subsystems can be successfully co-partitioned into 4-pair connected subregions based on resting state fMRI data using our proposed method, respectively. Particularly, three out of four cortical subregions are well-studied brain networks, and they are visual network, task-positive network, and task-negative network or called default mode network, respectively [22, 23]. The fourth cortical subregion might be an undiscovered brain network, named unknown network here, whose network nodes are distributed throughout the cerebral cortex. In addition, the four co-partitioned thalamic subregions are named visual area, task-positive area, task-negative area, and unknown area, respectively according to their respective connected cortical networks. Since each of the identified 4-pair connected subregions has closely functional association, the co-partitioned thalamic subregions’s brain functions can be inferred from their respective connected cortical networks. These results demonstrate that the mapped thalamocortical connection topography is neurobiologically meaningful.

![Figure 2](imageurl). Mapped connection topography of bilateral thalamocortical subsystems based on resting state fMRI data using our proposed method.
4.3. Symmetry in connection topography
For further evaluating the validity of our proposed method, an index is used to quantify the symmetry of the mapped thalamocortical connection topography between two hemispheres. This index measures the percentage overlap between pairwisely connected subregions on both hemispheres. For each of the pairwisely connected subregions, the symmetry index is defined as [24]

\[ SI_i = \frac{1}{n_i} \sum_{u \in V_i} \begin{cases} 1 & \text{if } l_u = l_u' \\ 0 & \text{otherwise} \end{cases} \] (5)

Where \( i = 1,2,\ldots,k, k = 4 \) is number of pairwisely connected subregions, \( n_i \) is the number of voxels of the \( i \)th pairwisely connected subregions \( V_i \) from one hemisphere, \( u' \) is the mirrored voxel of \( u \) from the other hemisphere, and \( l \) denotes the subregion label of a voxel. Because this measure requires that voxels are mirror symmetrical, only voxels that are symmetrically present on both hemispheres are considered in equation (5).

Figure 3 presents the results of symmetry measures for each pairwisely connected subregions, demonstrating that the mapped thalamocortical connection topography has medium to high symmetry. Specifically, the symmetry measures are 0.47, 0.50, 0.65, and 0.28 for visual area with visual network, task-positive area with task-positive network, task-negative area with task-negative network or called default mode network, and unknown area with unknown network, respectively. The fourth pairwisely connected subregions, i.e., unknown area with unknown network, has relatively lower symmetry, which might be more vulnerable to the noise of the resting state fMRI data. Overall, these results show that the co-clustering based connection topography is fairly similar between two hemispheres. This finding provides sufficient face validity for our proposed co-clustering based connectopic mapping method because the proposed method is completely driven by data.

![Figure 3](image)

**Figure 3.** Symmetry of the mapped thalamocortical connection topography. Pair 1: Visual area with visual network. Pair 2: Task-positive area with task-positive network. Pair 3: Task-negative area with task-negative network. Pair 4: Unknown area with unknown network.

4.4. Comparison with other alternative methods
We compared the proposed connectopic mapping method with other alternative methods, including prior connectivity mapping based methods, and ICA based methods. Compared with these alternative methods, the proposed co-clustering based connectopic mapping method has two main advantages.
Firstly, the proposed method doesn’t depend on prior knowledge, so that the mapped connection topography isn’t biased towards prior targets. In detail, anatomically defined cortical regions, such as prefrontal/temporal zone, primary motor/premotor zone, somatosensory zone, and posterior parietal/occipital zone, are usually adopted as targets in prior connectivity mapping based methods [9-12, 14-16], while cortical regions automatically identified by our method are well-studied brain networks, such as visual network, task-positive network, task-negative network. Therefore, the co-partitioned thalamic subregions with our method are much more meaningful than prior connectivity mapping based methods. Furthermore, the proposed method is fully driven by data, so the mapped connection topography doesn’t deviate from data. In contrast, existing ICA based methods are not completely data-driven approaches, usually requiring postprocessing, such as thresholding [2, 17, 18]. These comparisons demonstrate that our proposed connectopic mapping method is superior to these alternative methods, and the connection topography mapped by our method is more meaningful and interpretable.

5. Discussion and conclusions
We have proposed a novel method for unveiling the connectopic organization of the human brain. By leveraging the power of co-clustering, our method is able to map the connection topography between two ROIs of a certain neural circuit in consideration in a fully data-driven way. More importantly, the proposed connectopic mapping method parcels one ROI into subregions and identified their respective connected subregions from the other ROI simultaneously. The effectiveness of the proposed method is validated on the connection topography mapping of thalamocortical system for 57 subjects based on their resting state fMRI data. The validation experiment results indicate that our method can construct neurobiologically meaningful thalamocortical connection topography. The validity of the proposed method is further strengthened by the medium to high symmetry of the constructed connection topography between two hemispheres. Compared with existing connectopic mapping methods, our method is completely driven by data and not dependent on prior knowledge, yielding more meaningful and interpretable connection topography.

In conclusion, a novel, completely data-driven, co-clustering based connectopic mapping method is proposed, and outperforms existing methods. The proposed method can construct neurobiologically meaningful connection topography, which provides an important tool for connection topography mapping. The proposed method also offers a new perspective to revealing the connectopic organization of the human brain, which might facilitate better understanding the brain organization under normal and disease states.

Acknowledgments
This work was supported by the State Scholarship Fund of China Scholarship Council (201807845018), and the Natural Science Foundation of Chongqing (cstc2018jcyjAX0163).

References
[1] Arslan, S., Ktena, S.I., Makropoulos, A., Robinson, E.C., Rueckert, D., and Parisot, S., 2017, Human brain mapping: A systematic comparison of parcellation methods for the human cerebral cortex. *NeuroImage*, 170, p. 5-30
[2] Yuan, R., Di, X., Taylor, P.A., Gohel, S., Tsai, Y.-H., and Biswal, B.B., 2016, Functional topography of the thalamocortical system in human. *Brain Structure and Function*, 221 (4), p. 1971-1984
[3] Dhillon, I.S., 2001, Co-clustering documents and words using bipartite spectral graph partitioning. *Proc. of the 7th Int. Conf. on Knowledge and Data Mining*, p. 269-274
[4] Wu, L., Calhoun, V.D., Jung, R.E., and Caprihan, A., 2015, Connectivity based whole brain dual parcellation by group ICA reveals tract structures and decreased connectivity in schizophrenia. *Human Brain Mapping*, 36 (11), p. 4681-4701
[5] Cunningham, S.I., Tomasi, D., and Volkow, N.D., 2017, Structural and functional connectivity of the precuneus and thalamus to the default mode network. *Human Brain Mapping*, 38 (2), p. 938-956
[6] Ji, J.L., Spronk, M., Kulkarni, K., Repovš, G., Anticevic, A., and Cole, M.W., 2019, Mapping the human brain's cortical-subcortical functional network organization. *NeuroImage*, 185, p. 35-57

[7] Alcauter, S., Lin, W., Smith, J.K., Short, S.J., Goldman, B.D., Reznick, J.S., Gilmore, J.H., and Gao, W., 2014, Development of thalamocortical connectivity during infancy and its cognitive correlations. *Journal of Neuroscience*, 34 (27), p. 9067-9075

[8] Nanda, P., Banks, G.P., Pathak, Y.J., and Sheth, S.A., 2017, Connectivity-based parcellation of the anterior limb of the internal capsule. *Human Brain Mapping*, 38 (12), p. 6107-6117

[9] Zhang, D., Snyder, A.Z., Shimony, J.S., Fox, M.D., and Raichle, M.E., 2010, Noninvasive functional and structural connectivity mapping of the human thalamocortical system. *Cerebral Cortex*, 20 (5), p. 1187-1194

[10] Hale, J.R., Mayhew, S.D., Mullinger, K.J., Wilson, R.S., Arvanitis, T.N., Francis, S.T., and Bagshaw, A.P., 2015, Comparison of functional thalamic segmentation from seed-based analysis and ICA. *NeuroImage*, 114, p. 448-465

[11] Johansen-Berg, H., Behrens, T.E., Sillery, E., Ciccarelli, O., Thompson, A.J., Smith, S.M., and Matthews, P.M., 2005, Functional-anatomical validation and individual variation of diffusion tractography-based segmentation of the human thalamus. *Cerebral Cortex*, 15 (1), p. 31-39

[12] Zhang, D., Snyder, A.Z., Fox, M.D., Sansbury, M.W., Shimony, J.S., and Raichle, M.E., 2008, Intrinsic functional relations between human cerebral cortex and thalamus. *Journal of Neurophysiology*, 100 (4), p. 1740-1748

[13] Hwang, K., Bertolero, M., Liu, W., and D'Esposito, M., 2017, The human thalamus is an integrative hub for functional brain networks. *Journal of Neuroscience*, 37 (23), p. 5594-5607

[14] Traynor, C., Heckemann, R.A., Hammers, A., O'muircheartaigh, J., Crum, W.R., Barker, G.J., and Richardson, M.P., 2010, Reproducibility of thalamic segmentation based on probabilistic tractography. *NeuroImage*, 52 (1), p. 69-85

[15] Behrens, T.E., Johansen-Berg, H., Woolrich, M., Smith, S., Wheeler-Kingshott, C., Boulby, P., Barker, G., Sillery, E., Sheehan, K., and Ciccarelli, O., 2003, Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *Nature Neuroscience*, 6 (7), p. 750-757

[16] Semedo, C., Cardoso, M.J., Vos, S.B., Sudre, C.H., Bocchetta, M., Ribbens, A., Smeets, D., Rohrer, J.D., and Ourselin, S., 2018, Thalamic nuclei segmentation using tractography, population-specific priors and local fibre orientation. *Proc. of the 21st Int. Conf. on Medical Image Computing and Computer-Assisted Intervention*, p. 383-391

[17] O'Muircheartaigh, J., Vollmar, C., Traynor, C., Barker, G.J., Kumari, V., Symms, M.R., Thompson, P., Duncan, J.S., Koepp, M.J., and Richardson, M.P., 2011, Clustering probabilistic tractograms using independent component analysis applied to the thalamus. *NeuroImage*, 54 (3), p. 2020-2032

[18] O'muircheartaigh, J., and Jbabdi, S., 2017, Concurrent white matter bundles and grey matter networks using independent component analysis. *NeuroImage*, 170, p. 296-306

[19] Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., Van Der Kouwe, A., Killiany, R., Kennedy, D., and Klaveness, S., 2002, Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33 (3), p. 41-55

[20] Biswal, B., Zerrin Yetkin, F., Haughton, V.M., and Hyde, J.S., 1995, Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic Resonance in Medicine*, 34 (4), p. 537-541

[21] Lin, Q., Dai, Z., Xia, M., Han, Z., Huang, R., Gong, G., Liu, C., Bi, Y., and He, Y., 2015, A connectivity-based test-retest dataset of multi-modal magnetic resonance imaging in young healthy adults. *Scientific Data*, 2, p. 150056

[22] Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Van Essen, D.C., and Raichle, M.E., 2005, The brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences*, 102 (27), p. 9673-9678

[23] Raichle, M.E., 2015, The brain's default mode network. *Annual Review of Neuroscience*, 38, p. 433-447

[24] Kahnt, T., Chang, L.J., Park, S.Q., Heinzle, J., and Haynes, J.-D., 2012, Connectivity-based parcellation of the human orbitofrontal cortex. *Journal of Neuroscience*, 32 (18), p. 6240-6250