

RESEARCH ARTICLE

Hierarchical ordering with partial pairwise hierarchical relationships on the macaque brain data sets

Woosang Lim¹, Jungsoo Lee², Yongsub Lim³, Doo-Hwan Bae¹, Haesun Park⁴, Dae-Shik Kim², Kyomin Jung^{5*}

1 School of Computing, KAIST, Daejeon, Korea, **2** School of Electrical Engineering, KAIST, Daejeon, Korea, **3** Department of Computer Science and Engineering, Seoul National University, Seoul, Korea, **4** School of Computational Science and Engineering, Georgia Tech, Atlanta, United States of America, **5** Department of Electrical and Computer Engineering, Seoul National University, Seoul, Korea

* kjung@snu.ac.kr



OPEN ACCESS

Citation: Lim W, Lee J, Lim Y, Bae D-H, Park H, Kim D-S, et al. (2017) Hierarchical ordering with partial pairwise hierarchical relationships on the macaque brain data sets. PLoS ONE 12(5): e0177373. <https://doi.org/10.1371/journal.pone.0177373>

Editor: Kewei Chen, Banner Alzheimer's Institute, UNITED STATES

Received: July 8, 2016

Accepted: April 26, 2017

Published: May 18, 2017

Copyright: © 2017 Lim et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Two data sets which are macaque vision and somatosensory-motor data sets are provided in the paper of Felleman and Van Essen (1991): Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral cortex*, 1(1), 1-47.

Funding: W. Lim acknowledges support from the 2015 KAIST Graduate Research Fellowship via Kim-Bo-Jung Fund and the 2016 Google Ph.D. Fellowship in machine learning. K. Jung was

Abstract

Hierarchical organizations of information processing in the brain networks have been known to exist and widely studied. To find proper hierarchical structures in the macaque brain, the traditional methods need the entire pairwise hierarchical relationships between cortical areas. In this paper, we present a new method that discovers hierarchical structures of macaque brain networks by using partial information of pairwise hierarchical relationships. Our method uses a graph-based manifold learning to exploit inherent relationship, and computes pseudo distances of hierarchical levels for every pair of cortical areas. Then, we compute hierarchy levels of all cortical areas by minimizing the sum of squared hierarchical distance errors with the hierarchical information of few cortical areas. We evaluate our method on the macaque brain data sets whose true hierarchical levels are known as the FV91 model. The experimental results show that hierarchy levels computed by our method are similar to the FV91 model, and its errors are much smaller than the errors of hierarchical clustering approaches.

Introduction

Hierarchical organization in the brain networks has been known to enable the efficient processing of information to support complex brain functions, and it has been studied in various ways to understand structural and functional brain networks [1, 2]. Among the recent studies, most of the work can be roughly categorized into the following two types of approaches: finding hierarchical modularity in the brain [3–7], and hierarchical ordering of cortical areas in the brain [8–14]. The former computes a hierarchy of modules by partitioning the organization into submodules without using pairwise hierarchical relationships, but the result may not reflect the information flow in the brain. Whereas the latter computes hierarchy levels of cortical areas which successfully reflect the information flow, but it needs pairwise hierarchical relationships.

supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (2016R1A2B2009759), and supported by the Brain Korea 21 Plus Project in 2016. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Hierarchical ordering of cortical areas in the brain was proposed to provide the understanding and insight about cortical structure and function by Felleman and Van Essen (1991) [8]. To compute hierarchy orders of cortical areas, they first obtained the connectivities of cortical areas which imply the existence of information flow between two cortical areas from tract tracing experiments. Then, they derived 305 pairwise hierarchical relationships for the 32 cortical areas by observing differential laminar source and termination patterns, and the pairwise hierarchical relationships consists of three types on the pair of cortical areas which have a connection between them: feedforward, feedback, and lateral. Specifically, the lateral pairwise hierarchical relationship corresponds to the flow of information between two cortical areas whose hierarchy levels are the same. The feedforward pairwise hierarchical relationship corresponds to the flow of information from a cortical area of lower hierarchy level to a cortical area of higher hierarchy level, and the reverse direction flow corresponds the feedback. Especially, feedforward has two different types of weights which are ascending and strongly ascending, and feedback also has two different types of weights which are descending and strongly descending [8, 12]. Based on these extracted pairwise hierarchical relationships, Felleman and Van Essen proposed the FV91 model which describes 10 discrete hierarchy levels of information processing and a global view of areal relations (Table 1) [8]. Similarly, other related studies also compute hierarchy levels based on the entire pairwise hierarchical relationships [9, 12, 14–17].

Although traditional methods including the FV91 model successfully compute hierarchical orders of cortical areas in the brain, they need almost the entire pairwise hierarchical relationships to compute hierarchy levels, and the technique of obtaining pairwise hierarchical relationships is complicated. For example, the technique using invasive method and anatomical criteria is not applicable to the in vivo human brain, thus there is a very limited amount of pairwise hierarchical information for human brain. Consequently, we have a question; How can we compute hierarchy orders of cortical areas by using a restricted amount of pairwise hierarchical relationships with small errors?

In this paper, we propose a novel method which computes hierarchy orders with the connectivities of cortical areas and a smaller amount of pairwise hierarchical relationships. More specifically, since the connectivities of cortical areas do not imply the pairwise hierarchical relationships between cortical areas, we use the pseudo hierarchical distance to overcome the lack of pairwise hierarchical relationships. Although we use a restricted amount of pairwise hierarchical relationships, our method computes the hierarchy levels of cortical areas with small errors. Whereas the traditional approaches need the entire pairwise hierarchical

Table 1. The FV91 model: Hierarchy levels of macaque vision and somatosensory-motor data sets [8].

Hierarchy Level	The FV91 Model (vision)	The FV91 Model (somatosensory-motor)
9	46 TF TH	35 36
8	STPa AITd AITv	SMA 6 Id
7	7a FEF STPp CITd CITv	4 Ig
6	VIP LIP MSTd MSTl FST PITd PITv	7a 7b
5	DP VOT	SII
4	MDP MIP PO MT V4t V4	RI
3	PIP V3A	5
2	V3 VP	2
1	V2	1
0	V1	3a 3b

<https://doi.org/10.1371/journal.pone.0177373.t001>

relationships of cortical areas to compute accurate hierarchy levels. In the later sections, we will provide the motivation and the detailed description of our method. We will also evaluate our method on the macaque brain data sets, and compare the experimental results with the variants of FV91 models using the entire pairwise hierarchical relationships and the results of hierarchical clustering approaches.

Methods

In this section, we propose a novel method that can solve the hierarchical ordering problem by using the connectivities of cortical areas and the partial information of their pairwise hierarchical relationships. We note that the connectivities of cortical areas corresponds to the existences of information flow between two cortical areas, and do not imply the pairwise hierarchical relationships between cortical areas. Thus, we define *pseudo hierarchical distance* to overcome the lack of pairwise hierarchical relationships, and compute hierarchy levels of cortical areas by minimizing the hierarchical distance errors. Our method is displayed in Alg 1, and it consists of three parts: manifold learning, modeling the hierarchical distance, minimizing the hierarchical distance errors. We will discuss them in details in the following sections.

1. **Manifold Learning.** In this part, we compute the k -dimensional embedding of cortical areas in the brain by using connectivity of cortical areas without the pairwise hierarchical relationships, so that we can capture the unseen relationships between cortical areas by using k -dimensional vectors. Especially, we use Laplacian eigenmap to compute k -dimensional representative vectors given the adjacency matrix of brain connectivity graph.
2. **Modeling Hierarchical Distance.** We assume that if the information is processed in step-wise fashion, then the direct interaction between two cortical areas which have a large hierarchical distance would be relatively rare. Based on this assumption, in this part we compute similarities $s(u_i, u_j)$ for every pair of nodes by using k -dimensional embedding, and define pseudo hierarchical distance $\mu(u_i, u_j)$ by using similarities $s(u_i, u_j)$ to overcome the lack of pairwise hierarchical relationships.
3. **Minimizing the Hierarchical Distances Errors.** In this part, we set the hierarchy level $h(u_i)$ for a few nodes, e.g., $h(V1) = 0$, and compute hierarchy levels of cortical areas by minimizing the sum of the square of the hierarchical distance errors for all the pairs.

Algorithm 1 Hierarchical Ordering with Partial Pairwise Hierarchical Relationships

Input: The $n \times n$ weighted adjacency matrix \mathbf{W}

Output: Hierarchy levels $\mathbf{h} = (h(u_1), \dots, h(u_n))$ for n cortical areas

1: Manifold Learning.

Eigen-decompose a directed graph Laplacian \mathbf{L}_d defined in Eq (1) with rank- k and construct \mathbf{Y} in Eq (2)

Compute \mathbf{Z} in Eq (3) by normalizing row vectors of \mathbf{Y}

2: Modeling Hierarchical Distance.

Compute a regularized Tanimoto similarity $s(u_i, u_j)$ for every pair by using \mathbf{Z} (Eq (4))

Compute a pseudo hierarchical distance $\mu(u_i, u_j)$ by using $s(u_i, u_j)$ (Eq (5))

3: Minimizing the Error Function.

Fix the hierarchy level for some cortical areas, e.g., $h(u_i) = 0$

Compute $\mathbf{h} = (h(u_1), \dots, h(u_n))$ by minimizing $\sum_{u_i, u_j} (|h(u_i) - h(u_j)| - \mu(u_i, u_j))^2$ in the interval $[0, K]$

Map $h(u_i)$ into discrete domain $\{0, 1, \dots, K\}$

Manifold learning based on directed normalized Laplacian

Before computing similarities between cortical areas, we first use manifold learning to compute k -dimensional vectors which represent cortical areas in the brain. To compute low-dimensional representative vectors, we use graph Laplacian eigenmap which is one of the manifold learning techniques, and we will introduce it in this section. We note that we do not use pairwise hierarchical relationships to compute such low-dimensional embedding, and we only need connectivity matrix of cortical areas whose weight elements are 0 or 1, where the positive weight 1 implies the existences of information flow between two cortical areas.

Suppose that two cortical areas are connected or they have relatively many common connected cortical areas. Then, we argue that their k -dimensional representative vectors should be close to each other. Manifold learning is suitable to compute such representative vectors, since it computes a low-dimensional representation of the data set which preserves locality properties [18, 19]. That is, if two points \mathbf{x}_1 and \mathbf{x}_2 are similar or close to each other, then $\mathbf{f}(\mathbf{x}_1)$ and $\mathbf{f}(\mathbf{x}_2)$ are close to each other, where $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_k(\mathbf{x}))$ is a k -dimensional representative vector computed by the manifold learning given the data \mathbf{x} .

Laplacian eigenmap is one of the effective manifold learning techniques given the graph $G = (V, E)$, and it is widely used in data mining and machine learning, where $n = |V|$ [18, 20–22]. Among its variants, combinatorial Laplacian $\mathbf{L} = (\mathbf{D} - \mathbf{W})$ is the most basic Laplacian operator given the symmetric graph, where \mathbf{W} is the $n \times n$ adjacency matrix such that its element W_{ij} is 1 when there exist an edge from node u_i to u_j and zero when there is no edge, and \mathbf{D} is the diagonal matrix consisting of $d(u_i) = \sum_j W_{ij}$. We can compute k -dimensional representative vectors by computing the first k eigenfunctions of Laplacian operator \mathbf{L} . There are also normalized Laplacian matrices of symmetric graph which are $\mathbf{L}_{rw} = \mathbf{I} - \mathbf{D}^{-1} \mathbf{W}$ and $\mathbf{L}_{sym} = \mathbf{I} - \mathbf{D}^{-\frac{1}{2}} \mathbf{W} \mathbf{D}^{-\frac{1}{2}}$. Since the brain networks are directed, we use directed normalized Laplacian which reflects information flows to compute k -dimensional representative vectors.

Normalized directed Laplacian \mathbf{L}_d is defined as

$$\mathbf{L}_d = \mathbf{I} - \frac{\boldsymbol{\Psi}^{1/2} \mathbf{P} \boldsymbol{\Psi}^{-1/2} + \boldsymbol{\Psi}^{-1/2} \mathbf{P}^\top \boldsymbol{\Psi}^{1/2}}{2}, \tag{1}$$

where the Perron vector $\boldsymbol{\psi}$ is the unique left eigenvector corresponding the largest eigenvalue ρ of the transition matrix $\mathbf{P} = \mathbf{D}^{-1} \mathbf{W}$, and $\boldsymbol{\Psi}$ is a $n \times n$ diagonal matrix with $\Psi_{ii} = \psi_i$. Let \mathbf{Y} be the matrix which consists of the first k eigenvectors of \mathbf{L}_d , then row vectors of \mathbf{Y} are k -dimensional vectors

$$\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_k] \text{ subject to } \mathbf{L}_d \mathbf{y}_i = \lambda_i \mathbf{y}_i, \tag{2}$$

where eigenvalues of \mathbf{L}_d satisfy $0 = \lambda_1 < \lambda_2 \leq \dots \leq \lambda_{n-1}$, and \mathbf{y}_i is the first i -th eigenvector of \mathbf{L}_d . Especially, we normalize row vectors of \mathbf{Y} so that each row vector lies on the surface of k -dimensional sphere as

$$\mathbf{Z} = (\text{diag}(\mathbf{Y}^\top \mathbf{Y}))^{-1} \mathbf{Y}. \tag{3}$$

That is, the row vectors of \mathbf{Z} in Eq (3) are k -dimensional representative vectors of cortical areas in the brain.

Modeling the hierarchical distance

In this section, we compute similarities between cortical areas by using k -dimensional representative vectors, and model the hierarchical distance to overcome the lack of pairwise hierarchical relationships.

We use Tanimoto similarity and k -dimensional representative vectors to compute similarities between cortical areas. Tanimoto Similarity is a generalized version of Jaccard similarity, which measures the similarities between vectors [23]. For two vectors $\mathbf{f}(u_i)$ and $\mathbf{f}(u_j)$, it is defined as $\frac{\mathbf{f}(u_i)^\top \mathbf{f}(u_j)}{\|\mathbf{f}(u_i) - \mathbf{f}(u_j)\|_2^2 + \mathbf{f}(u_i)^\top \mathbf{f}(u_j)}$. To limit the maximum and minimum values, we define a similarity $s(u_i, u_j)$ of two cortical areas u_i and u_j by using regularized Tanimoto similarity

$$s(u_i, u_j) = \frac{\mathbf{f}(u_i)^\top \mathbf{f}(u_j)}{\|\mathbf{f}(u_i) - \mathbf{f}(u_j)\|_2^2 + \mathbf{f}(u_i)^\top \mathbf{f}(u_j) + \epsilon}, \tag{4}$$

where $\epsilon = 1/\delta$ and $s(u_i, u_j) \in [-\delta, \delta]$.

Now we model the hierarchical distance of every pair based on the similarity $s(u_i, u_j)$ defined in Eq (4). Let $h(u_i) \in \mathbb{R}$ be a hierarchical level of cortical area u_i , and $z(u_i, u_j) \in \mathbb{R}$ be an absolute value of difference between hierarchical levels of cortical areas u_i and u_j , i.e., $z(u_i, u_j) = |h(u_i) - h(u_j)|$. If we set $\max_{u_i} h(u_i) = K$ and $\min_{u_i} h(u_i) = 0$, then $h(u_i) \in [0, K]$ and $z(u_i, u_j) \in [0, K]$. Since we do not know the exact $z(u_i, u_j)$, we introduce a pseudo hierarchical distance $\mu(u_i, u_j)$ which is an estimated mean of $z(u_i, u_j)$. We define $\mu(u_i, u_j)$ as

$$\mu(u_i, u_j) = \delta - s(u_i, u_j), \tag{5}$$

where $\delta = K/2$ and $\epsilon = 1/\delta$. We note that the pseudo hierarchical distance $\mu(u_i, u_j)$ is in the interval $[0, K]$. The pseudo hierarchical distance $\mu(u_i, u_j)$ between cortical areas u_i and u_j is small when their similarity $s(u_i, u_j)$ is large, and $\mu(u_i, u_j)$ is large when their similarity $s(u_i, u_j)$ is small.

Minimizing hierarchical distance errors

In this section, we define the objective function based on the pseudo hierarchical distance, and compute hierarchy levels of all cortical areas by minimizing the objective function.

We note that the hierarchical distances in the graph are symmetric and nonnegative, even if the graph is directed. Thus, we consider a symmetric objective function, whereas we used the directed normalized Laplacian in the previous sections. Basically, we want to minimize the sum of the squared error of $z(u_i, u_j)$ and $\mu(u_i, u_j)$ for all pairs. However, we do not know the exact hierarchy levels of all cortical areas. Approximate hierarchy levels $\tilde{\mathbf{h}} = (\tilde{h}(u_1), \dots, \tilde{h}(u_n))$ for n cortical areas. That is, we set the objective function as

$$\begin{aligned} & \underset{\mathbf{h}}{\text{minimize}} \sum_{u_i, u_j} (z(u_i, u_j) - \mu(u_i, u_j))^2 \\ & = \underset{\mathbf{h}}{\text{minimize}} \sum_{u_i, u_j} (|h(u_i) - h(u_j)| - \mu(u_i, u_j))^2, \end{aligned} \tag{6}$$

where \mathbf{h} consists of hierarchy levels of cortical areas such that $\mathbf{h} = (h(u_1), \dots, h(u_n))$. The reason of considering all pairs is that the recent studies suggest the importance of role of long-distance connections in the brain [1, 24–27], and we interpret long-distance connections as paths between all cortical areas.

We fix the hierarchy levels for specific cortical areas before solving Eq (6), e.g., $h(V1) = 0$. Then, we solve Eq (6) to compute hierarchy levels in the continuous domain, i.e., $h(u_i) \in [0, K]$. We can use an average of local minimum, since the FV91 model is also just one of the 150,000 equally plausible solution [15]. Next, we map the computed hierarchy levels to the discrete domain $\{0, 1, 2, \dots, K\}$, e.g., rounding to integer.

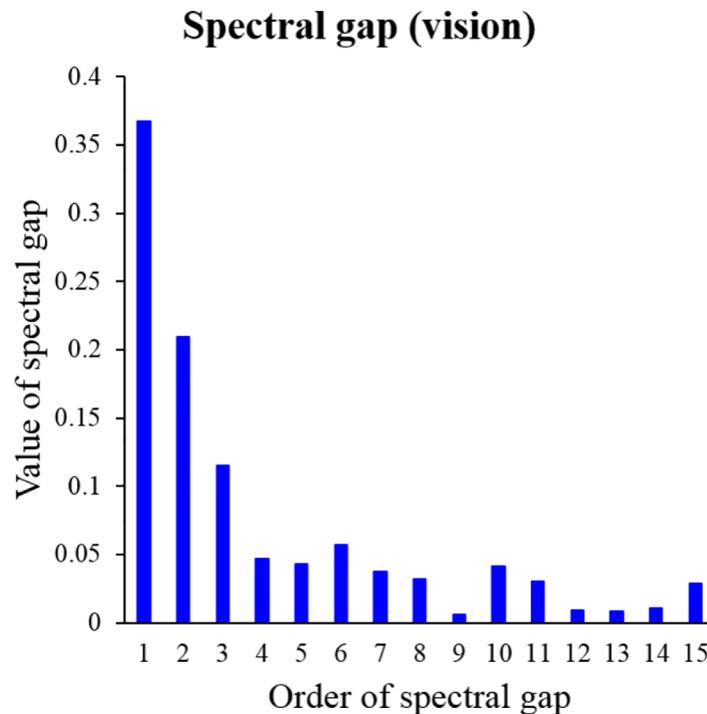


Fig 1. The spectral gaps of directed Laplacian of macaque vision data set. We use the spectral gaps to select parameters k and K . We compute the spectral gaps by using the eigen-decomposition of graph Laplacian, and do not use any pairwise hierarchical relationships of cortical areas.

<https://doi.org/10.1371/journal.pone.0177373.g001>

In the experiment section, we will show that we can similarly compute hierarchy levels compared to the FV91 model with partial pairwise hierarchical relationships. We will assume that we can relatively easily infer the cortical areas which have the smallest or the largest hierarchy level. That is, we will use the partial pairwise hierarchical relationships about only some cortical areas of the minimum hierarchy level or the highest hierarchy level, and we will fix the hierarchy levels of a few cortical areas as 0 or K to minimize Eq (6).

Selecting parameters

In Alg 1, we need to select two parameters k and K which are the dimension of Laplacian eigenmap and the maximum hierarchical level, respectively. We can select them by analyzing the spectral gap γ_i of Laplacian which are the differences between eigenvalues s.t. $\gamma_i = \lambda_i - \lambda_{i+1}$, and we can compute them by using the eigen-decomposition of graph Laplacian [22]. We note that we do not use pairwise hierarchical relationships of cortical areas to compute two parameters k and K , since we need only the connectivity matrix of cortical areas whose weight are 0 or 1 to construct graph Laplacian.

For example, the spectral gaps of directed Laplacian L_d of macaque vision network are displayed in Fig 1. The 1st, 2nd, 3rd spectral gaps in Fig 1 are relatively large than others, thus we can notice that macaque vision cortical areas can be categorized as high-, medium- or low-level. However, since we want to obtain a more separated hierarchical structure, we select $k = 4$. We can also determine the maximum hierarchy level K by using the spectral gaps. In Fig 1, the five consecutive spectral gaps from the 4th spectral gap are relatively similar, and the

9th spectral gap is relatively very small. Thus, we can select 8 as the total number of hierarchy levels, then the maximum hierarchical level K is 7 when the minimum level is 0, *i.e.*, $h(u_i) \in \{0, 1, 2, \dots, 7\}$. Meanwhile, $K = 9$ or $K = 10$ are also possible candidates for K according to the distribution of spectral gap. We note that the maximum level of the FV91 model is 9, and the maximum level of the modified FV91 model is 10. To compare the FV91 model, we set $K = 9$ in the experiments.

Experimental results

In this section, we report the experimental results on the macaque brain data sets whose true hierarchical levels are known as FV91 model. We use the measures which are *Pearson Correlation Coefficient* (PCC), *Mean-Absolute Error* (MAE) and *Root Mean Square Error* (RMSE) for comparison.

$$\begin{aligned} \text{PCC}(\tilde{\mathbf{x}}) &= \frac{\sum_i^n (x(u_i) - m)(\tilde{x}(u_i) - \tilde{m})}{\sqrt{\sum_i^n (x(u_i) - m)^2} \sqrt{\sum_i^n (\tilde{x}(u_i) - \tilde{m})^2}}, \\ \text{MAE}(\tilde{\mathbf{x}}) &= \frac{1}{n} \sum_{i=1}^n |x(u_i) - \tilde{x}(u_i)|, \\ \text{RMSE}(\tilde{\mathbf{x}}) &= \sqrt{\frac{1}{n} \sum_{i=1}^n (x(u_i) - \tilde{x}(u_i))^2}, \end{aligned}$$

where $\tilde{\mathbf{x}} = (\tilde{x}(u_1), \dots, \tilde{x}(u_n))$, $m = \frac{1}{n} \sum_i^n x(u_i)$, and $\tilde{m} = \frac{1}{n} \sum_i^n \tilde{x}(u_i)$. The range of PCC is $[-1, 1]$, and the optimum value of PCC is 1. The minimum of both MAE and RMSE are 0. Suppose that $\mathbf{x} = (x(u_1), \dots, x(u_n))$ be the hierarchy levels of FV91 model, and $\tilde{\mathbf{x}} = (\tilde{x}(u_1), \dots, \tilde{x}(u_n))$ be approximate hierarchy levels computed by other methods. Then, if the approximate solution is similar with the FV91 model, then PCC will be close to 1, and MAE and RMSE will be small and close to 0.

We empirically compare our method with the FV91 model [8], hierarchical ordering with social rank [28], traditional hierarchical clustering [29], and EM-based hierarchical clustering [30]. We use directed normalized Laplacian for embedding directed graph data set. For traditional hierarchical clustering (HC), we use `pdist.m`, `linkage.m`, and `cluster.m` function in MATLAB with 49 different settings. We use 7 different distance metrics in `pdist.m` for computing pairwise distance between object pairs: Euclidean (Euclid), Standardized Euclidean (SEuclid), cosine (Cos), correlation (Corr), Chebychev (Cheb), Mahalanobis (Mah), Hamming (Ham). We also use 7 different methods in `linkage.m` for computing distance between clusters: average, centroid, complete, median, single, ward, weighted.

Data sets

Let us remind that we use the connectivities of cortical areas and the partial information of pairwise hierarchical relationships for our method, and connectivities of cortical areas do not imply the pairwise hierarchical relationships. The connectivity matrix (adjacency matrix) \mathbf{W} consists of 0 or 1 weights, where $\mathbf{W}_{i,j} = 1$ means the existences of information flow from i -th cortical area to j -th cortical area, and $\mathbf{W}_{i,j} = 0$ means the nonexistence of information flow from i -th cortical area to j -th cortical area. Since we want to analyze the hierarchical structure of macaque brain, we use two data sets which are macaque vision and somatosensory-motor data sets are provided in [8]. We note that if we want to apply our method to other data sets, we can use the connectivity obtained from fMRI, EEG, MEG, and DTI by the graph theoretical analysis of structural and functional systems [31].

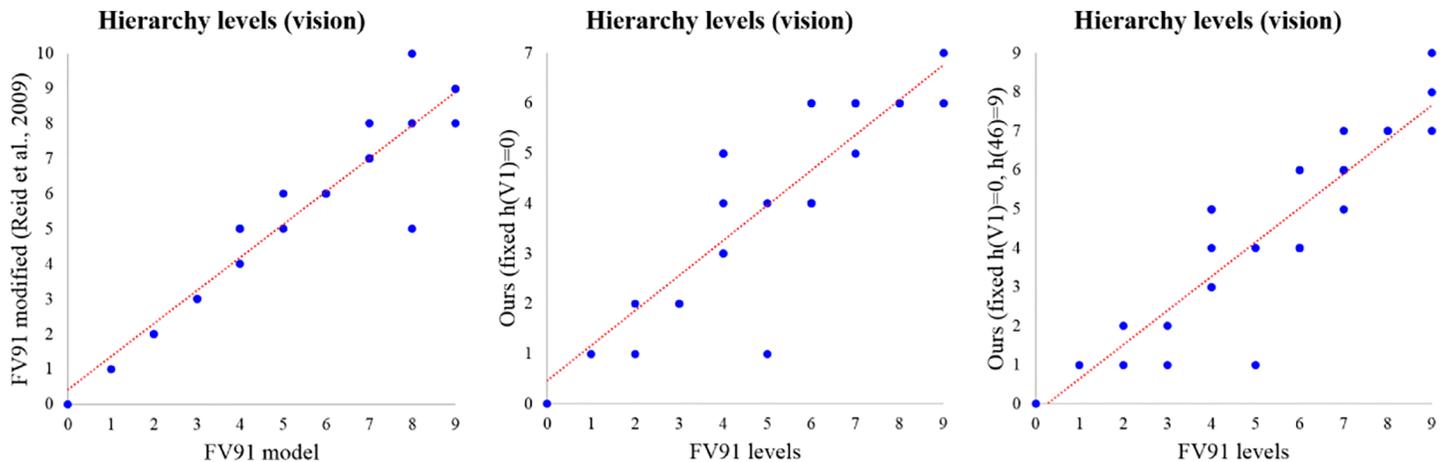


Fig 2. The scatter plot between the FV91 model [8] and FV91 modified model [12] for the data set of macaque brain vision (left). Comparison between the results of our method and the FV91 model on the macaque vision data set (middle and right). We used the adjacency matrix and at most 12.38% of pairwise hierarchical relationships which FV91 and FV91 modified models used. Although we used a small partial pairwise hierarchical relationships, we get similar result with the FV91 and FV91 modified models. The red line is a linear regression line.

<https://doi.org/10.1371/journal.pone.0177373.g002>

For macaque vision data, we use the ‘fve32.mat’ file which includes the 32×32 adjacency matrix corresponding to 32 visual cortical areas and their 315 connections. The weights of connections are 0 or 1. For macaque somatosensory-motor data, we use the ‘macaque47.mat’ file which includes the 47×47 adjacency matrix corresponding to both vision and somatosensory-motor systems. We extract the 16×16 principal submatrix from the original 47×47 adjacency matrix to obtain the adjacency matrix of induced subgraph of 16 somatosensory-motor cortical areas. The weights of connections are also 0 or 1.

Experiments on macaque vision data

In this section, we report the experimental results on the macaque vision data set. We use the adjacency matrix \mathbf{W} which consists of only 0 or 1 values depending on the connection in the brain. The adjacency matrix \mathbf{W} can not describe the pairwise hierarchical relationships such as feedforward, feedback, and lateral. Thus, we use partial pairwise hierarchical relationships in addition.

For macaque visual cortex data set, we use the pairwise hierarchical relationships about only the V1 cortical area whose hierarchy level is the smallest, and fix its hierarchy level as 0 *i.e.*, $h(V1) = 0$. It corresponds to $5.08\% = 16/315$ of pairwise hierarchical relationships which FV91 model used. We run Alg 1 with $h(V1) = 0$ and Eq (6), and display the averaged results as the scatter plot in the middle of Fig 2. The scatter plot shows that hierarchy levels computed by our method are similar to the FV91 model, even though we use a very small amount of pairwise hierarchical information compared to the FV91 model and FV91 modified model.

Meanwhile, we get the TH and 46 cortical areas as the top 2 hierarchy levels in the middle of Fig 2. Hence, it would be reasonable to set $h(TH) = K$ or $h(46) = K$ in addition, where $K = 9$. Setting $h(V1) = 0$ and $h(46) = K$ corresponds to 12.38% of pairwise hierarchical relationships which FV91 model used, and the experimental result is displayed as the scatter plot in the right of Fig 2. In the right scatter plot, the cortical areas of high hierarchy levels in the FV91 model have more accurate hierarchy levels compared to the scatter plot in the middle of Fig 2. Despite of using partial pairwise hierarchical relationships, the results of our method displayed in the

right scatter plot have high PCC and small errors: PCC = 0.890, MAE = 1.094, and RMSE = 1.403. The result of FV91 modified model displayed in the left of Fig 2 is slightly better than ours: PCC = 0.940, MAE = 0.377, and RMSE = 0.796. However, the FV91 modified model used the entire pairwise hierarchical relationships, and it can not compute hierarchical levels of MDP and MIP regions [12]. If we consider the potential errors of MDP and MIP cortical areas in FV91 modified model, then error values of FV91 modified model may be worse.

We also compare the results of our method and various versions of traditional hierarchical clustering methods [29], and the experimental results are displayed in Fig 3. Although we select the 7 best results among 49 different settings of hierarchical clustering by considering low RMSE and high PCC, we can see that the traditional hierarchical clustering algorithms can not effectively find the hierarchical structures. Since hierarchical clustering just merges (or splits) clusters in a greedy manner by using cluster dissimilarity, it can not accurately find hierarchical structures which reflect the information flow in the brain. That is, the assigned cluster numbers from 0 to K for n nodes can not properly reflect the information flow in the brain. We can see that the cortical areas of the lowest level in the results of hierarchical clustering have hierarchical levels from 4 to 6 in the FV91 model. In addition, in the results of

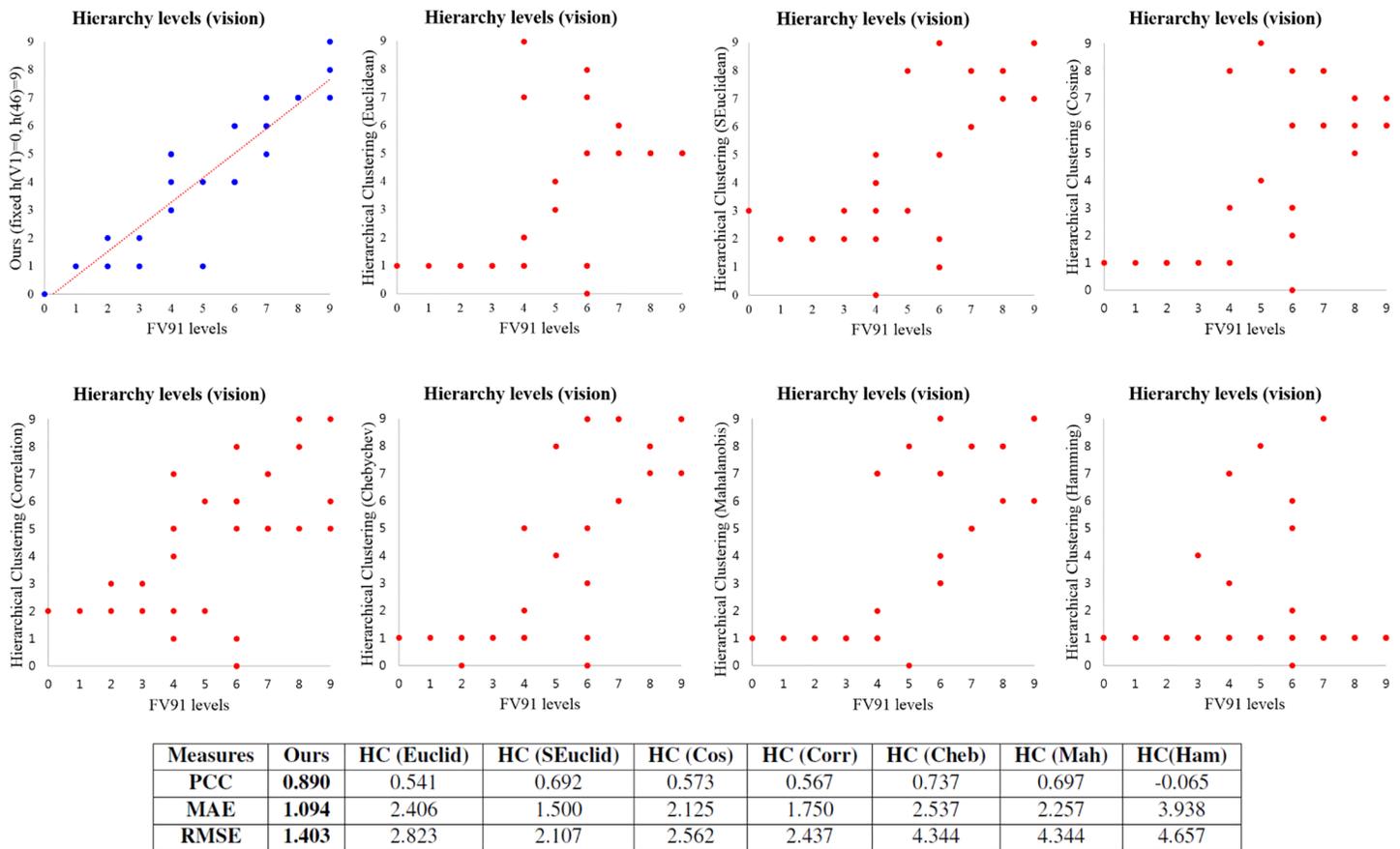


Fig 3. Comparison among the results of the FV91 model [8], our method and traditional hierarchical clustering methods [29] on the macaque vision data set. The performance is measured by using PCC, MAE, and RMSE. Although our method uses 12.38% of pairwise hierarchical relationships compared to the FV91 model, we get similar results to the FV91 model. In addition, the errors of our method are much smaller than the errors of traditional hierarchical clustering methods. The red line is a linear regression line.

<https://doi.org/10.1371/journal.pone.0177373.g003>

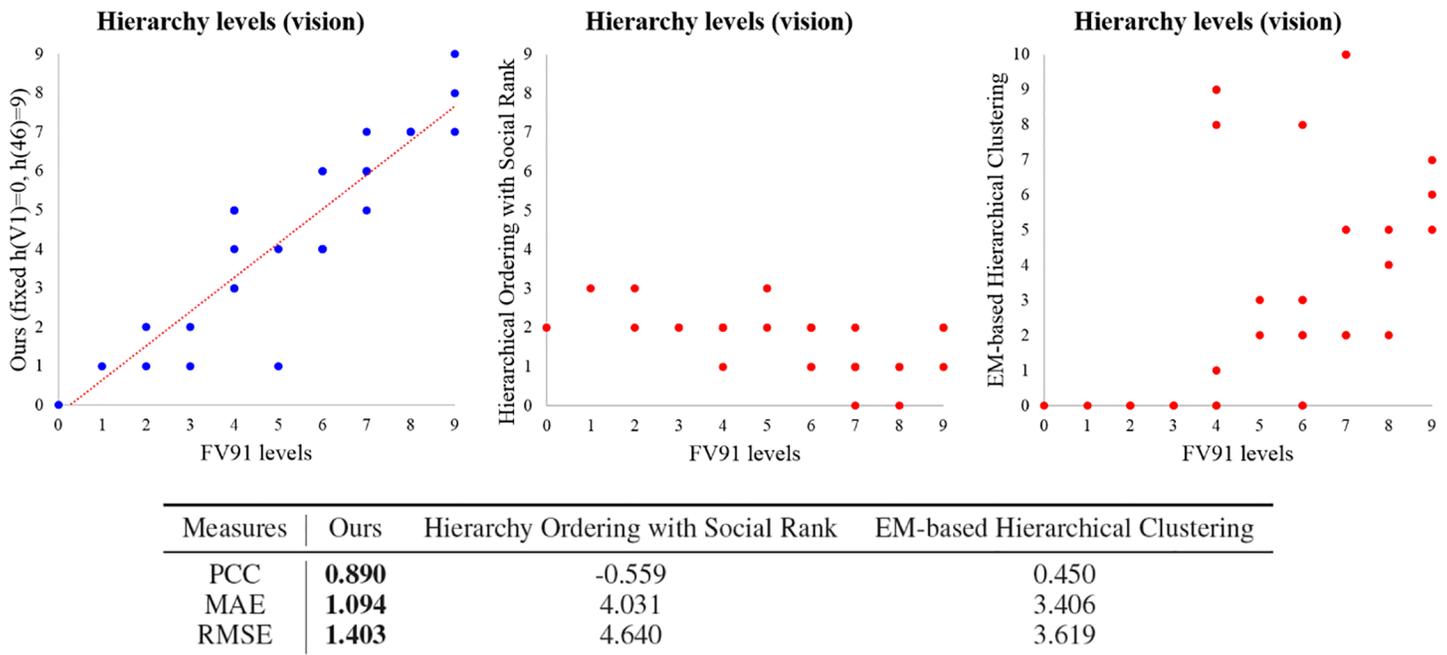
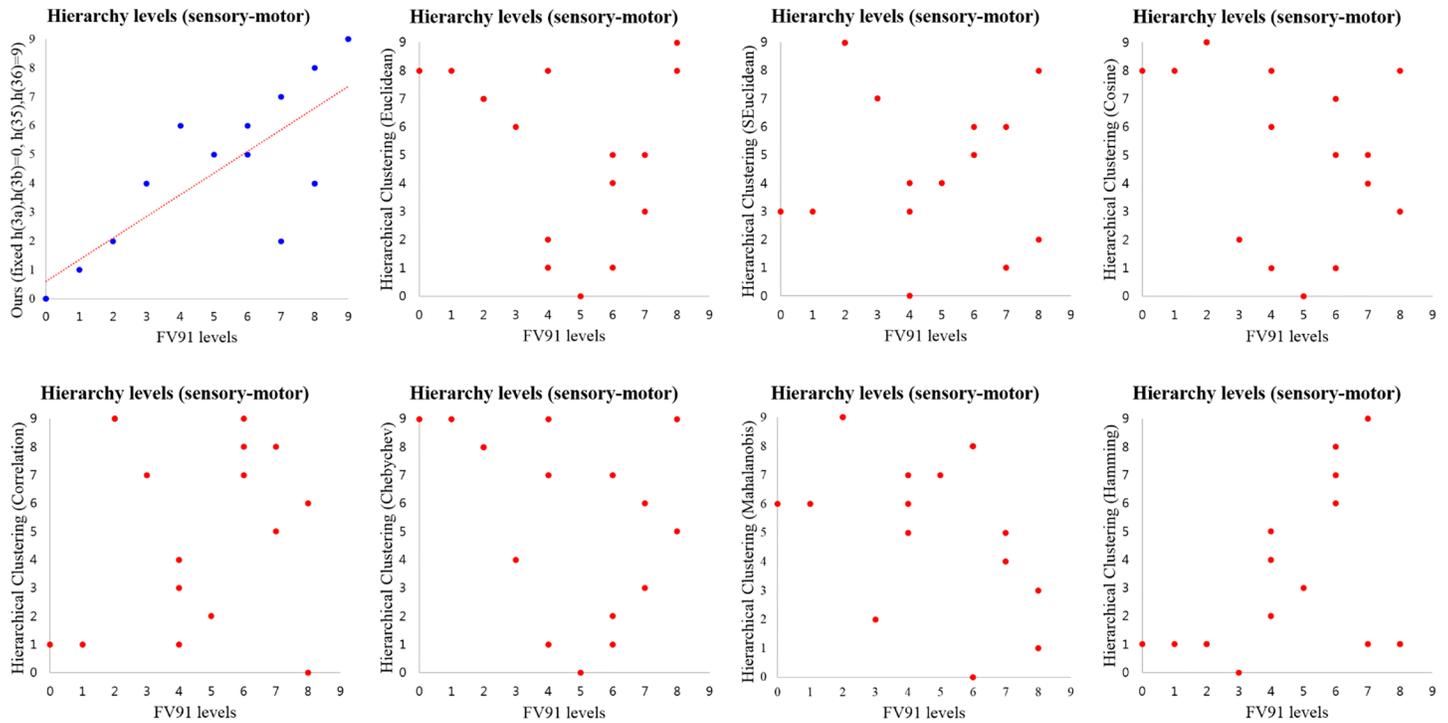


Fig 4. Comparison among the results of the FV91 model [8], our method, hierarchical ordering with social rank [28], and EM-based hierarchical clustering [30] on the macaque vision data set. The performance is measured by using PCC, MAE, and RMSE. The result of our method is similar to the FV91 model. In addition, the errors of our method are much smaller than the errors of hierarchical ordering with social rank, and EM-based hierarchical clustering.

<https://doi.org/10.1371/journal.pone.0177373.g004>

hierarchical clustering, we can hardly find the differences of hierarchy levels among the cortical areas which have hierarchical levels from 0 to 4 in the FV91 model. Whereas the result of our method shows an apparent hierarchical structure which is much closer to the FV91 model than the results of hierarchical clustering. There are also big differences between performances of our method and hierarchical clustering in terms of PCC, MAE, and RMSE. PCC of our method is close to the optimum 1, and is much higher than PCC of various hierarchical clustering results. Both MAE and RMSE of our method are much smaller than the results of hierarchical clustering.

Finally, we compare the results of our method, hierarchical ordering with social rank [28], and EM-based hierarchical clustering [30], and the experimental results are displayed in Fig 4. Although we select the best result of hierarchical ordering with social rank and EM-based hierarchical clustering by considering low RMSE and high PCC respectively, both methods can not effectively find the hierarchical structures. We can see that the maximum level in the result of hierarchical ordering with social rank is just 3, and we can hardly find any similar tendency between the FV91 model and the results of hierarchical ordering with social rank. We guess the reason is that the hierarchical property in the social network is quite different with the hierarchical property in the brain network. We also can not find any advantages of EM-based hierarchical clustering for hierarchical ordering of cortical areas in Fig 4. The PCC, MAE, and RMSE of EM-based hierarchical clustering are worse than the results of traditional hierarchical clustering with SEuclid distance. Thus, there are large differences among the performances of our method, hierarchical ordering with social rank, and EM-based hierarchical clustering in terms of PCC, MAE, and RMSE. PCC of our method is much higher than PCC of the two



Measures	Ours	HC (Euclid)	HC (SEuclid)	HC (Cos)	HC (Corr)	HC (Cheb)	HC (Mah)	HC (Ham)
PCC	0.803	0.276	0.251	0.565	0.503	0.402	0.038	0.184
MAE	1.063	2.813	2.750	1.813	2.188	2.438	3.063	3.500
RMSE	1.984	3.562	3.518	2.817	3.112	3.400	3.976	4.257

Fig 5. Comparison among the results of the FV91 model [8], our method and various versions of hierarchical clustering [29] on the macaque somatosensory-motor data set. The performance is measured by using PCC, MAE, and RMSE. Although our method uses 36.45% pairwise hierarchical relationships compared to the FV91 model, we get similar results to the FV91 model. In addition, the errors of our method are much smaller than the errors of hierarchical clustering. The red line is a linear regression line.

<https://doi.org/10.1371/journal.pone.0177373.g005>

methods, and both MAE and RMSE of our method are also much smaller than the results of two methods.

Experiments on macaque sensory-motor data

In this section, we report the experimental results on the macaque somatosensory-motor data set. We compare the FV91 model, hierarchical clustering, and our method. We again use the adjacency matrix W and partial pairwise hierarchical relationships.

We display the experimental results of our method and the traditional hierarchical clustering methods [29] on the macaque somatosensory-motor network as the scatter plots in Fig 5. We run Alg 1 with $h(3a), h(3b) = 0$ and $h(35), h(36) = 9$ which correspond to 36.45% of pairwise hierarchical relationships which FV91 model used. The top left scatter plot in Fig 5 shows that hierarchy levels computed by our method are similar to the FV91 model, even though we use partial pairwise hierarchical relationships. We also run hierarchical clustering with 49 different settings, and select the best result for each distance metric by considering low RMSE and high PCC. We display these 7 results as the scatter plots in Fig 5. We can see that our method outperforms various hierarchical clustering approaches. The figures show that the

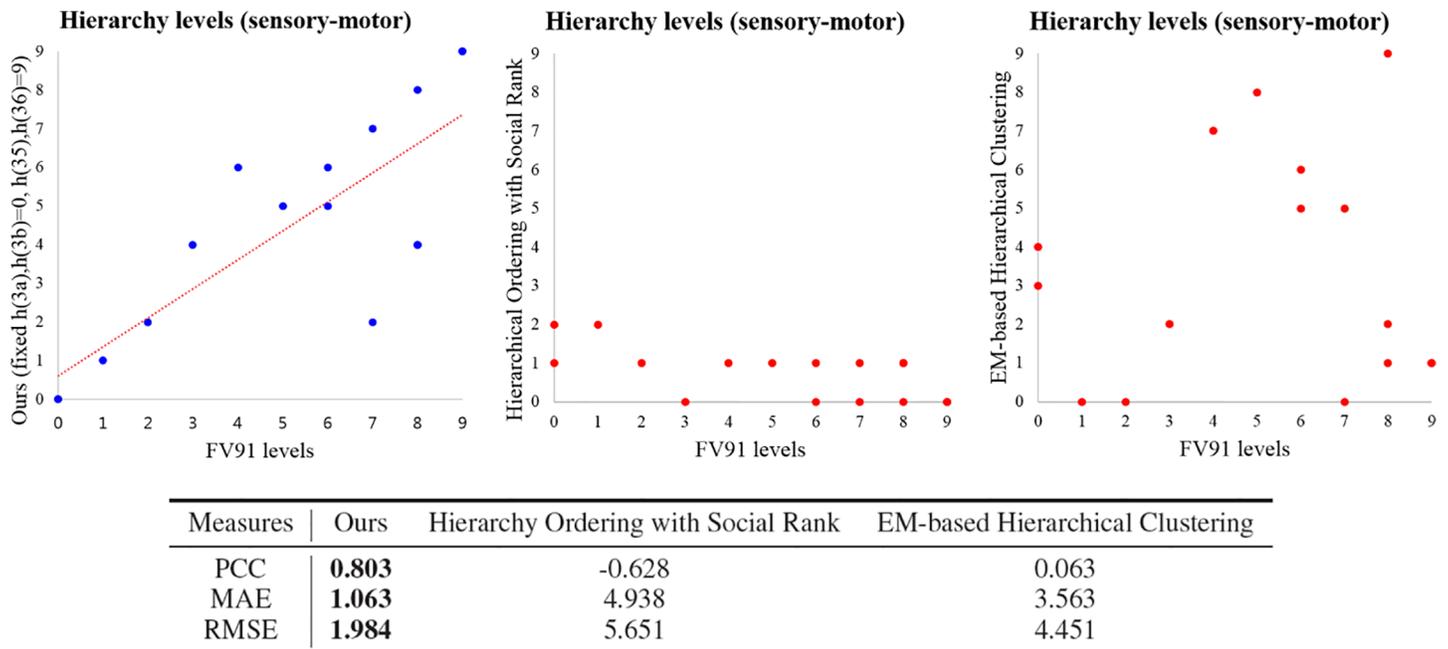


Fig 6. Comparison among the results of the FV91 model [8], our method, hierarchical ordering with social rank [28], and EM-based hierarchical clustering [30] on the macaque somatosensory-motor data set. The performance is measured by using PCC, MAE, and RMSE. The result of our method is similar to the FV91 model. The results of our method is superior than the results of hierarchical ordering with social rank and EM-based hierarchical clustering in terms of PCC, MAE, and RMSE.

<https://doi.org/10.1371/journal.pone.0177373.g006>

traditional hierarchical clustering approach can not effectively find the hierarchical structures, whereas our method finds a similar hierarchical structure to the FV91 model. There are also big differences between performances of our method and hierarchical clustering approach in terms of PCC, MAE, and RMSE. PCC of our method is close to the optimum 1, and is much higher than PCC of various hierarchical clustering results. Both MAE and RMSE of our method are also much smaller than the results of hierarchical clustering.

Finally, we compare the experimental results of our method, hierarchical ordering with social rank [28], and EM-based hierarchical clustering [30] on the macaque somatosensory-motor network in Fig 6. Although we select the best result of hierarchical ordering with social rank and EM-based hierarchical clustering by considering low RMSE and high PCC respectively, both methods can not find the proper hierarchical structures. In Fig 6, we can see that the maximum level in the result of hierarchical ordering with social rank is just 2, and we can hardly find any similar tendency between the FV91 model and the results of hierarchical ordering with social rank and EM-based hierarchical clustering. PCC of our method is much higher than PCC of the two methods, and both MAE and RMSE of our method are also much smaller than the results of two methods. Especially, PCC of hierarchical ordering with social rank and EM-based hierarchical clustering are -0.628 and 0.063 respectively, and they are poor results.

Conclusion

In this paper, we suggested a new framework that compute the hierarchy orders of cortical areas in the macaque brain by using partial pairwise hierarchical relationships. To overcome the lack of pairwise hierarchical relationships, we used a directed Laplacian eigenmap to

exploit the inherent topology of the brain networks as a low-dimensional embedding, and we defined pseudo hierarchical distances for every pair of cortical areas by using the low-dimensional embedding. We computed hierarchy levels of cortical areas by minimizing the sum of squared hierarchical distance errors with the hierarchical information of few cortical areas. The experimental results showed that hierarchy levels computed by our method are similar to the FV91 model, even though we used partial pairwise hierarchical relationships. Furthermore, we showed that our method outperforms hierarchical clustering methods in terms of several error measures. Thus, we conclude that our method is quite a good compromise between the variant of FV91 models and hierarchical clustering method for computing hierarchy orders of cortical areas in the brain.

Acknowledgments

W. Lim acknowledges support from the 2015 KAIST Graduate Research Fellowship via Kim-Bo-Jung Fund and the 2016 Google Ph.D. Fellowship in machine learning. K. Jung was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIP) (2016R1A2B2009759), and supported by the Brain Korea 21 Plus Project in 2016. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author Contributions

Conceptualization: WL JL YL DK KJ.

Data curation: WL JL YL.

Formal analysis: WL HP KJ.

Funding acquisition: DB KJ.

Investigation: WL JL YL.

Methodology: WL YL.

Project administration: WL.

Resources: WL JL YL DB DK KJ.

Software: WL YL.

Supervision: DB KJ.

Validation: WL.

Visualization: WL KJ.

Writing – original draft: WL.

Writing – review & editing: WL HP KJ.

References

1. Park HJ, Friston K. Structural and functional brain networks: from connections to cognition. *Science*. 2013; 342(6158):1238411. <https://doi.org/10.1126/science.1238411> PMID: 24179229
2. Bullmore E, Sporns O. The economy of brain network organization. *Nature Reviews Neuroscience*. 2012; 13(5):336–349. PMID: 22498897

3. Zhou C, Zemanová L, Zamora G, Hilgetag CC, Kurths J. Hierarchical organization unveiled by functional connectivity in complex brain networks. *Physical Review Letters*. 2006; 97(23):238103. <https://doi.org/10.1103/PhysRevLett.97.238103> PMID: 17280251
4. Ferrarini L, Veer IM, Baerends E, van Tol MJ, Renken RJ, van der Wee NJ, et al. Hierarchical functional modularity in the resting-state human brain. *Human Brain Mapping*. 2009; 30(7):2220–2231. <https://doi.org/10.1002/hbm.20663> PMID: 18830955
5. Meunier D, Lambiotte R, Fornito A, Ersche KD, Bullmore ET. Hierarchical modularity in human brain functional networks. *Frontiers in Neuroinformatics*. 2009; 3. <https://doi.org/10.3389/neuro.11.037.2009>
6. Boly M, Perlberg V, Marrelec G, Schabus M, Laureys S, Doyon J, et al. Hierarchical clustering of brain activity during human nonrapid eye movement sleep. *Proceedings of the National Academy of Sciences*. 2012; 109(15):5856–5861. <https://doi.org/10.1073/pnas.1111133109>
7. Wang Y, Li TQ. Analysis of whole-brain resting-state fMRI data using hierarchical clustering approach. *PLoS One*. 2013; 8(10):e76315. <https://doi.org/10.1371/journal.pone.0076315> PMID: 24204612
8. Felleman DJ, Van Essen DC. Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*. 1991; 1(1):1–47. PMID: 1822724
9. Batardière A, Barone P, Knoblauch K, Giroud P, Berland M, Dumas AM, et al. Early specification of the hierarchical organization of visual cortical areas in the macaque monkey. *Cerebral Cortex*. 2002; 12(5):453–465. <https://doi.org/10.1093/cercor/12.5.453> PMID: 11950763
10. Vezoli J, Falchier A, Jouve B, Knoblauch K, Young M, Kennedy H. Quantitative analysis of connectivity in the visual cortex: extracting function from structure. *The Neuroscientist*. 2004; 10(5):476–482. <https://doi.org/10.1177/1073858404268478> PMID: 15359013
11. Grant S, Hilgetag CC. Graded classes of cortical connections: quantitative analyses of laminar projections to motion areas of cat extrastriate cortex. *European Journal of Neuroscience*. 2005; 22(3):681–696. <https://doi.org/10.1111/j.1460-9568.2005.04232.x> PMID: 16101750
12. Reid AT, Krumnack A, Wanke E, Kötter R. Optimization of cortical hierarchies with continuous scales and ranges. *Neuroimage*. 2009; 47(2):611–617. <https://doi.org/10.1016/j.neuroimage.2009.04.061> PMID: 19398021
13. Krumnack A, Reid AT, Wanke E, Bezgin G, Kötter R. Criteria for optimizing cortical hierarchies with continuous ranges. *Frontiers in neuroinformatics*. 2010; 4. <https://doi.org/10.3389/fninf.2010.00007> PMID: 20407634
14. Markov NT, Vezoli J, Chameau P, Falchier A, Quilodran R, Huissoud C, et al. Anatomy of hierarchy: feedforward and feedback pathways in macaque visual cortex. *Journal of Comparative Neurology*. 2014; 522(1):225–259. <https://doi.org/10.1002/cne.23458> PMID: 23983048
15. Hilgetag CC, O'Neill MA, Young MP. Indeterminate organization of the visual system. *SCIENCE*. 1996; p. 776–776. <https://doi.org/10.1126/science.271.5250.776> PMID: 8628990
16. Hilgetag CC, O'Neill MA, Young MP. Hierarchical organization of macaque and cat cortical sensory systems explored with a novel network processor. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2000; 355(1393):71–89. <https://doi.org/10.1098/rstb.2000.0550>
17. Barone P, Batardiere A, Knoblauch K, Kennedy H. Laminar distribution of neurons in extrastriate areas projecting to visual areas V1 and V4 correlates with the hierarchical rank and indicates the operation of a distance rule. *The Journal of Neuroscience*. 2000; 20(9):3263–3281. PMID: 10777791
18. Belkin M, Niyogi P. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation*. 2003; 15(6):1373–1396. <https://doi.org/10.1162/089976603321780317>
19. Fogel F, d'Aspremont A, Vojnovic M. Serialrank: Spectral ranking using seriation. In: *Advances in Neural Information Processing Systems*; 2014. p. 900–908.
20. Belkin M, Niyogi P. Laplacian eigenmaps and spectral techniques for embedding and clustering. In: *Advances in Neural Information Processing Systems*. vol. 14; 2001. p. 585–591.
21. Mahadevan S, Maggioni M. Value function approximation with diffusion wavelets and Laplacian eigenfunctions. In: *Advances in Neural Information Processing Systems*; 2005. p. 843–850.
22. Luxburg UV. A tutorial on spectral clustering. *Statistics and Computing*. 2007; 17:395–416. <https://doi.org/10.1007/s11222-007-9033-z>
23. Lind P, Maltseva T. Support vector machines for the estimation of aqueous solubility. *Journal of Chemical Information and Computer Sciences*. 2003; 43(6):1855–1859. <https://doi.org/10.1021/ci034107s> PMID: 14632433
24. Markov N, Misery P, Falchier A, Lamy C, Vezoli J, Quilodran R, et al. Weight consistency specifies regularities of macaque cortical networks. *Cerebral Cortex*. 2011; 21(6):1254–1272. <https://doi.org/10.1093/cercor/bhq201> PMID: 21045004

25. Markov NT, Ercsey-Ravasz M, Lamy C, Gomes ARR, Magrou L, Misery P, et al. The role of long-range connections on the specificity of the macaque interareal cortical network. *Proceedings of the National Academy of Sciences*. 2013; 110(13):5187–5192. <https://doi.org/10.1073/pnas.1218972110>
26. Gallos LK, Makse HA, Sigman M. A small world of weak ties provides optimal global integration of self-similar modules in functional brain networks. *Proceedings of the National Academy of Sciences*. 2012; 109(8):2825–2830. <https://doi.org/10.1073/pnas.1106612109>
27. Hermundstad AM, Bassett DS, Brown KS, Aminoff EM, Clewett D, Freeman S, et al. Structural foundations of resting-state and task-based functional connectivity in the human brain. *Proceedings of the National Academy of Sciences*. 2013; 110(15):6169–6174. <https://doi.org/10.1073/pnas.1219562110>
28. Gupte M, Shankar P, Li J, Muthukrishnan S, Iftode L. Finding hierarchy in directed online social networks. In: *Proceedings of International Conference on World Wide Web*. ACM; 2011. p. 557–566.
29. Murtagh F, Contreras P. Algorithms for hierarchical clustering: an overview. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*. 2012; 2(1):86–97.
30. Li J, Ray S, Lindsay BG. A nonparametric statistical approach to clustering via mode identification. *Journal of Machine Learning Research*. 2007; 8(Aug):1687–1723.
31. Bullmore E, Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*. 2009; 10(3):186–198. <https://doi.org/10.1038/nrn2575> PMID: 19190637