

一种基于先验信息的脑功能网络提取方法*

杜宇慧^{1,3†}, 桂志国¹, 隋婧^{2,3}

(1. 中北大学信息与通信工程学院, 太原 030051; 2. 中国科学院自动化研究所模式识别国家重点实验室, 北京 100190; 3. The Mind Research Network 研究所医学图像分析实验室, 阿尔伯克基 87131)

摘要: 提出一种基于先验信息的脑功能网络提取方法。该方法基于先验信息得到初始的目标和背景种子点, 然后基于图论将整个脑图像构建图, 最后利用半监督聚类技术提取脑功能网络。基于不同信噪比的模拟数据, 对提出方法、基于种子点的方法, 独立成分分析方法以及两种聚类方法(归一化最小化割和 K-均值方法)进行比较。基于真实脑静息态功能核磁共振数据, 使用提出方法对默认模式网络进行提取。基于模拟数据的实验结果表明, 提出的算法相对于传统的方法可以得到更为准确且鲁棒的脑功能网络。基于静息态功能核磁共振数据得到的默认模式网络在一些重要脑区具有高的稳定性, 且不同地点采集数据得到的结果具有较强的一致性。提出的方法是一种有效的脑功能网络提取方法。

关键词: 磁共振成像; 脑功能网络; 图论; 半监督聚类; 先验信息

中图分类号: TP391.41 **文献标志码:** A **文章编号:** 1001-3695(2016)04-1258-04

doi:10.3969/j.issn.1001-3695.2016.04.066

Method for brain functional network extraction based on prior information

Du Yuhui^{1,3†}, Gui Zhiguo¹, Sui Jing^{2,3}

(1. School of Information & Communication Engineering, North University of China, Taiyuan 030051, China; 2. National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China; 3. Medical Image Analysis Laboratory, Institute of the Mind Research Network, Albuquerque 87131, USA)

Abstract: This paper proposed a prior information based method to identify brain functional networks. The method firstly obtained objective and background seeds based on prior information, and then constructed a graph for the brain image using graph theory, finally extracted the brain functional network using a semi-supervised clustering technique. Based on the simulated data with different signal to noise ratios, it compared the proposed method, the seed based method, independent component analysis, and two clustering methods (normalized cuts and K-means). Based on the real resting-state functional magnetic resonance imaging (fMRI) data, it extracted default mode network using the proposed method. Simulation based experimental results show that compared to the traditional methods, the proposed method could obtain more accurate and robust brain functional network. Real fMRI based experimental results illustrate that the default mode network had high stability in important regions as well as great consistency among data from different sites. The proposed method is an effective method for brain functional network extraction.

Key words: magnetic resonance imaging; brain functional network; graph theory; semi-supervised clustering; prior information

0 引言

功能核磁共振成像(fMRI)可揭示脑区之间功能信号的一致波动^[1-3], 这种一致波动是脑功能网络存在的体现, 也正是脑功能网络提取方法所能利用的信息。脑功能网络分析已被广泛应用于探索神经精神疾病的医学影像标志^[4,5], 因此稳定可靠的脑功能网络提取方法具有重要的应用价值。

常用的脑功能网络提取方法包括假设驱动和数据驱动两类^[6-8], 它们用不同的技术手段来探索脑区之间的一致波动以提取脑功能网络。在假设驱动方法中, 基于种子点的方法^[9]最为常用, 该方法以种子点或感兴趣区为中心计算种子点的功能信号与脑内其它体素的相关系数, 并采用阈值方法得到最终激活的脑功能网络。基于种子点的方法受所选择的种子点影响大, 而且对采用的阈值敏感。在数据驱动方法中, 独立成分

分析(independent component analysis, ICA)^[10-12]可以得到互相独立的脑功能网络, 该方法应用非常广泛。ICA的缺点在于成分个数的不可预知性和输出成分的无序性, 使用不同的成分个数计算会得到不同的脑功能网络模式, 输出成分的无序性使得不同被试成分之间的对应性难以建立。聚类方法也是一种常用的数据驱动脑功能网络提取方法, 它将三维脑图像分割成几个互不相交叠的类别, 使得每个类别所包含体素的功能信号具有一致的波动, 被分到一个类别的脑区域被认为属于同一个脑功能网络。已有方法包括 C-均值^[13]、层次聚类^[14,15]、自组织映射^[16]、谱聚类^[17]等。基于聚类方法得到的脑功能网络所包含的区域具有一致的功能, 然而类别个数的选择变得困难。

1 基于先验信息的脑功能网络提取方法

为克服已有方法的缺陷, 本文提出一种基于先验信息利用

收稿日期: 2015-01-26; 修回日期: 2015-03-30 基金项目: 国家自然科学基金面上项目(81471367)

作者简介: 杜宇慧(1980-), 女(通信作者), 山西高平人, 讲师, 博士(后), 主要研究方向为医学图像分析、脑功能图像处理(duyuhui@nuc.edu.cn); 桂志国(1972-), 男, 教授, 博导, 主要研究方向为图像处理; 隋婧(1980-), 女, 山东潍坊人, 研究员, 博士, 主要研究方向为脑图像处理。

图论提取脑功能网络的方法。该方法利用先验信息得到初始种子点,然后将整个脑图像建立一个图,图的节点是体素,节点间的边反映体素功能信号之间波动的异同,再利用半监督聚类技术得到共激活的脑功能网络。

a) 利用先验知识选取可靠的种子点以保证得到的脑功能网络是可靠的。本文利用已知的结构信息来决定功能网络的目标种子点,选取距离要提取的功能网络较远且与目标种子点的功能信号相关性小的体素作为背景种子点。

b) 利用功能信号对三维的脑图像构建图 $G(V, E)$ 。将每个体素看做是图的一个节点 V , 用两两体素间的功能信号的相似性度量图的一条边 E 。本文中体素 i 和 j 的相似性表示为 $\text{corr}(T_i, T_j)$, 其中 T_i 表示体素 i 的功能时间信号, corr 表示计算 Pearson 相关系数。在得到所有体素的相关系数矩阵以后, 为保证算法的有效性和抗噪性, 本文采用 k 近邻图对图进行稀疏化, 使得稀疏化后的图的每个节点只与 k 个节点相连。

c) 基于种子点和构建的图, 利用半监督聚类算法来提取脑功能网络。大多数半监督聚类方法都定义了一个连续的分类函数 F , 通过最小化这个分类函数得到聚类结果。在半监督聚类算法中, 这个分类函数既体现聚类结果的光滑性(即相似的节点更易被分到同一类), 又体现聚类结果与初始种子点类别标号之间的一致性^[18]。已有方法包括高斯随机场和调和函数法^[19]、保持局部整体一致性方法(LGC)^[20]、交替最小化图信息传播方法^[21]等。本文采用的 LGC 方法可以通过优化下面目标函数解决:

$$Q(F) = \sum_{i,j} w_{ij} \left\| \frac{F_i}{\sqrt{D_{ii}}} - \frac{F_j}{\sqrt{D_{jj}}} \right\|^2 + \mu \sum_{i=1}^n \|F_i - Y_i\|^2 \quad (1)$$

其中: w_{ij} 表示节点之间边的权重; F_i 是节点 i 的类别向量; $D_{ii} = \sum_{j=1}^n w_{ij}$ 则表示节点 i 的度(即与其他节点连接边的权重和); Y_i 是节点 i 的初始类别向量; n 是节点数目。式(1)中的第一项是邻近节点应具有相似的类别信息的局部一致性约束条件, 第二项是最终聚类结果和初始类别信息应具有一致性的全局一致性约束条件, 两者通过参数 μ 权衡以得到满足局部和全局一致性的聚类结果。式(1)用矩阵形式可表示为

$$Q(F) = \text{tr}(F^T(I - S)F + \mu(F - Y_{\text{ini}})^T(F - Y_{\text{ini}})) \quad (2)$$

其中: I 为单位矩阵; $S = D^{-1/2}WD^{1/2}$ 为规整化的相似性矩阵, $W = \{w_{ij}\}$; $F \in R^{n \times c}$ 为分类函数矩阵; $Y_{\text{ini}} \in R^{n \times c}$ 为初始类别信息矩阵, c 是类别数目(本文中 $c = 2$)。可以证明, 最小化函数 $Q(F)$ 可以通过式(3)迭代算法得到。

$$F^{m+1} = (1 - \alpha)SF^m + \alpha Y_{\text{ini}} \quad (3)$$

其中: F^m 是在第 m 次迭代后的分类函数矩阵的值; $F^0 = Y_{\text{ini}}$; α 为 $0 < \alpha < 1$, 是与 μ 相关的一个参数。分类函数矩阵迭代的过程就是类别信息传播的过程, 每个节点都会从与它相似的节点类别中获得信息, 同时还要与它的初始类别保持一致。当算法收敛时, 就得到了每个节点的最终类别。

综上所述, 本文提出的脑功能网络的提取方法包括:

a) 基于先验信息, 决定初始的属于脑功能网络的目标种子点和不属于脑功能网络的背景种子点, 由此决定类别标号矩阵 Y_{ini} 。

b) 计算两两体素的功能信号之前的相似性度量, 然后构建 k 近邻图, 得到 S 。

c) 根据经验决定参数 α , 根据式(3)执行上面描述的半监督聚类算法直到收敛。

d) 根据每个体素最后的类别标号来提取脑功能网络。

2 实验验证

2.1 基于模拟数据的实验

为评估提出方法及传统方法的准确性和鲁棒性, 本文产生不同信噪比(signal to noise ratio, SNR)的模拟数据进行实验。首先, 将某正常被试的静息态 fMRI 数据使用 SPM8 进行层间校正、头动校正、配准到 $3 \times 3 \times 3$ 的 EPI 模板, 以及使用 6mm 的高斯核进行空间平滑。然后, 将预处理后的数据作为基础数据, 为三维脑图像中设定的三维长方体区域内的体素添加模拟信号来模拟脑功能网络。区域内每个体素添加的信号是一个相位变化的正弦信号和一个高斯白噪声信号的混合。SNR 设为 5、4、3、2 和 1, 这里 SNR 为信号方差除以噪声方差。图 1 显示了模拟的脑功能网络区域在某层的图像以及 SNR 为 1 时的观测数据在该层的模拟图像。

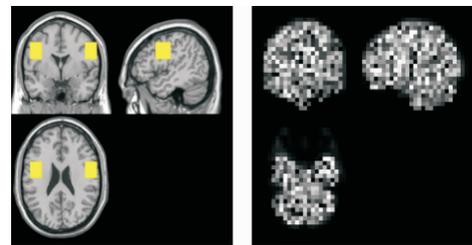


图1 模拟数据

基于模拟数据, 本实验对提出方法、基于种子点的方法、ICA、谱聚类方法中的归一化最小割(normalized cuts, Ncut)^[22]以及 K 均值(K-means)进行了比较。在使用提出方法时, 分别采用模拟区域及其他区域内四个相邻的体素作为目标和背景种子点, 使用不同的 k 值来构建 k 近邻图以评测参数的影响。在使用基于种子点的方法时, 种子点与提出方法使用的目标种子点相同, 并使用不同的阈值(0.3 ~ 0.7)来评价该方法的鲁棒性。ICA 中采用不同的成分个数(5 ~ 30)以及不同的 z 阈值(2 和 3)来评测脑功能网络受这些参数的影响。在使用 Ncut 和 K-means 方法时, 使用不同的类别个数(2 ~ 6)进行评估。另外, 由于在使用 ICA、Ncut 和 K-means 方法时会得到多个成分或类别, 本文通过度量得到结果与模拟网络之间的空间一致性, 得到最匹配的成分和类别作为提取的脑功能网络。最后, 对不同方法提取的脑功能网络的空间特性用空间吻合度进行评估, 计算公式为 $\frac{|A \cap B|}{|A \cup B|}$, 这里 A 和 B 分别表示提取的和模拟的脑功能网络区域。吻合度越接近于 1, 表示得到的结果越接近于模拟的脑功能网络。

2.2 基于真实数据的实验

本实验采用 Fcon1000 的公共数据对提出的算法进行验证, 对 5 个地点采集到的 106 个正常被试的静息态 fMRI 数据进行默认模式网络(default mode network, DMN)的提取, 被试来自 Bangor(19 个)、Berlin(26 个)、Leiden(19 个)、New York(20 个)以及 Oxford(22 个)。DMN 是脑功能网络中研究最为广泛且与神经精神疾病最密切相关的一个脑功能网络^[23], 它包括一些在静息态情况下较任务态更为活跃的脑区, 支持人脑自我意识和自我反省等基本功能。脑科学领域中, 对 DMN 的研究具有重要意义。

本实验根据先验知识将几个位于任务消极区和任务积极区的体素分别当做目标和背景种子点。由于脑图像具有几万个体素数,如果将每个体素作为图的一个节点,则不利于计算和存储构建的图,所以本实验用 27 邻域体素内的功能信号的平均值来表示一个节点信号,并在得到聚类结果后采用 27 邻域投票规则填充离散点。参数 k 设为 100。为了评估实验结果,在得到每组被试的 DMN 后,本实验根据该组所有被试的 DMN,计算每个体素在所有被试中出现的概率以构建 DMN 的概率图,该概率图可以反映 DMN 在被试间的重复性。此外,基于每个被试得到的 DMN,关于脑网络区域内每个体素计算它与网络内所有其他体素的 Pearson 相关系数,并用所有相关系数的平均值反映该体素属于这个网络的程度,因此每个被试将会得到一个三维的相关系数图来反映相应 DMN 的激活程度。进一步,为测量得到 DMN 的稳定性,对于每个地点的所有被试的 DMN 相关系数图进行基于体素的单样本 t 检验,即可得到每个地点数据对应的 DMN 组统计结果。

3 实验结果

3.1 基于模拟数据的实验

图 2 显示了使用提出方法、基于种子点方法、ICA、Ncut 以及 K-means 这些方法在 SNR 为 1 时得到的脑功能网络。从图中可以看出相对于其他方法,提出的方法可以准确地提取模拟的脑功能网络。图 3 表示在使用不同信噪比的模拟数据时,各种方法在不同参数下得到的脑功能网络与模拟网络间的空间吻合度结果。结果表明,提出方法具有更高的精确性,在使用不同 SNR 的数据时表现更为鲁棒,且受参数的影响小。

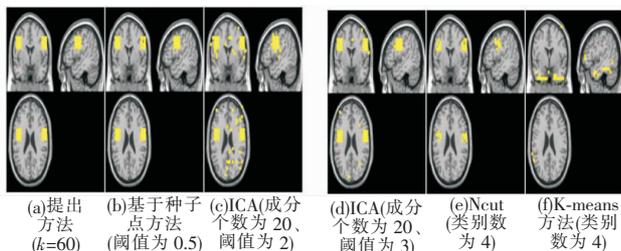


图 2 各种方法在 SNR 为 1 的情况下得到的功能网络区域

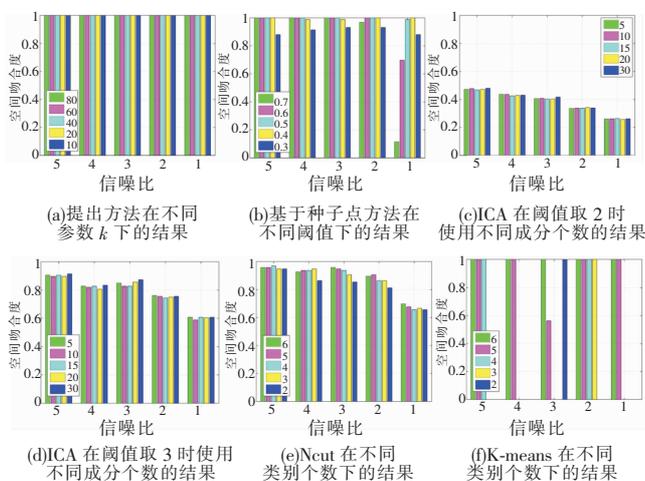


图 3 各种方法在使用不同信噪比数据及不同参数情况下得到的功能网络与模拟网络的吻合度结果

3.2 基于真实数据的实验

图 4(a) ~ (e) 为得到的五个地点被试的 DMN 的概率图。从图中可以看到 DMN 在被试间具有很好的重复性,尤其是在

扣带回前部和扣带回后部等一些重要脑区域,这一结论与已有工作相符。此外,基于同一地点不同被试的 DMN 相关系数图进行单样本 t 检验的结果显示于图 4(f) ~ (j) 中,可以看到不同地点得到的统计结果一致性较高,说明 DMN 是一个稳定性较高的脑功能网络。

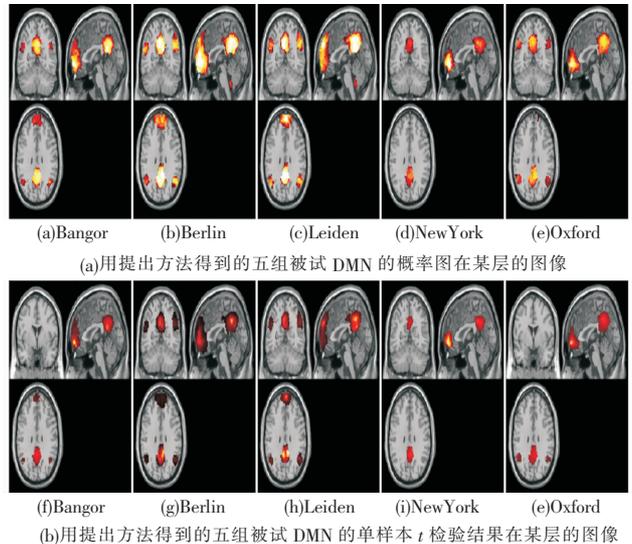


图 4 真实 fMRI 数据得到的 DMN 结果

4 结束语

稳定可靠的脑功能网络提取方法十分重要。基于种子点方法得到的脑功能网络受所选择的种子点影响大,且对使用的阈值敏感。ICA 方法中成分个数的不可预知使得提取可靠的个体脑功能网络变得困难,输出成分的无序性则使得挑选感兴趣的功能网络和建立不同被试网络之间的对应性成为一个难题。基于聚类的方法直接对体素的功能信号进行聚类提取功能网络,但其缺陷是类别个数的选择是一个难点。为克服传统脑功能网络提取方法的缺陷,本文提出一种基于先验信息利用图论知识使用半监督聚类技术进行脑功能网络提取的方法。基于模拟数据的实验表明,相对于传统的基于种子点方法、ICA 以及聚类方法(K-means 和 Ncut),提出方法可得到准确可靠的脑功能网络,且提出方法在使用不同噪声数据时表现更为稳定,受算法参数影响小。基于真实静息态 fMRI 的实验表明,提出的方法得到的 DMN 在扣带回前部和扣带回后部这些重要脑区具有高的可重复性和稳定性,且不同地点得到的 DMN 结果具有高度的一致性,这个结论与已有文献相符。此外,提出的方法的另外一个优点是通过利用先验信息,该方法能够直接建立起感兴趣脑功能网络在不同被试间的对应性。总之,本文提出的基于先验信息的方法是一种有效的脑功能网络提取方法。

参考文献:

- [1] Smith S M, Vidaurre D, Beckmann C F, et al. Functional connectomics from resting-state fMRI[J]. Trends in Cognitive Sciences, 2013, 17(12): 666-682.
- [2] Lee M H, Smyser C D, Shimony J S. Resting-state fMRI: a review of methods and clinical applications[J]. AJNR American Journal of Neuroradiology, 2013, 34(10): 1866-1872.
- [3] Sporns O. Contributions and challenges for network models in cognitive neuroscience[J]. Nature Neuroscience, 2014, 17(5): 652-660.
- [4] Castro E, Gomez-Verdejo V, Martinez-Ramon, et al. A multiple kernel

- learning approach to perform classification of groups from complex-valued fMRI data analysis: application to schizophrenia[J]. *Neuroimage*, 2014, **87**:1-17.
- [5] Khadka S, Meda S A, Stevens M C, *et al.* Is aberrant functional connectivity a psychosis endophenotype? A resting state functional magnetic resonance imaging study[J]. *Biological Psychiatry*, 2013, **74**(6):458-466.
- [6] Li Kaiming, Guo Lei, Nie Jingxing, *et al.* Review of methods for functional brain connectivity detection using fMRI[J]. *Computerized Medical Imaging and Graphics*, 2009, **33**(2):131-139.
- [7] Smith S M, Miller K L, Salimi-Khorshidi G, *et al.* Network modelling methods for FMRI[J]. *Neuroimage*, 2011, **54**(2):875-891.
- [8] 杜宇慧, 桂志国, 刘迎军, 等. 静息态脑功能网络分析的假设驱动和数据驱动方法综述[J]. *北京生物医学工程*, 2013, **32**(3):307-311.
- [9] Biswal B, Yetkin F Z, Haughton V M, *et al.* Functional connectivity in the motor cortex of resting human brain using echo-planar MRI[J]. *Magnetic Resonance in Medicine*, 1995, **34**(4):537-541.
- [10] Du Yuhui, Fan Yong. Group information guided ICA for fMRI data analysis[J]. *Neuroimage*, 2013, **69**:157-197.
- [11] 杜宇慧, 桂志国, 刘迎军, 等. 基于独立成分分析的脑功能网络分析方法综述[J]. *生物物理学报*, 2013, **29**(4):266-275.
- [12] Calhoun V D, Adali T. Multisubject independent component analysis of fMRI: a decade of intrinsic networks, default mode, and neurodiagnostic discovery[J]. *IEEE Reviews in Biomedical Engineering*, 2012, **5**:60-73.
- [13] Goutte C, Toft P, Rostrup E, *et al.* On clustering fMRI time series[J]. *Neuroimage*, 1999, **9**(3):298-310.
- [14] Bassett D S, Bullmore E, Verchinski B A, *et al.* Hierarchical organization of human cortical networks in health and schizophrenia[J]. *Journal of Neuroscience*, 2008, **28**(37):9239-9248.
- [15] Ferrarini L, Veer I M, Baerends E, *et al.* Hierarchical functional modularity in the resting-state human brain[J]. *Human Brain Mapping*, 2009, **30**(7):2220-2231.
- [16] Peltier S J, Polk T A, Noll D C. Detecting low-frequency functional connectivity in fMRI using a self-organizing map (SOM) algorithm[J]. *Human Brain Mapping*, 2003, **20**(4):220-226.
- [17] Van Den Heuvel M, Mandl R, Pol H H. Normalized cut group clustering of resting-state FMRI data[J]. *PLoS One*, 2008, **3**(4):e2001.
- [18] Zhu Xiaojin. Semi-supervised learning literature survey[J]. *Computer Sciences*, 2008, **37**(1):63-77.
- [19] Zhu Xiaojin, Ghahramani Z, Lafferty J. Semi-supervised learning using gaussian fields and harmonic functions[C]//Proc of the 20th International Conference on Machine Learning. 2003.
- [20] Zhou Dengyong, Bousquet O, Lal T N, *et al.* Learning with local and global consistency[J]. *Advances in Neural Information Processing Systems*, 2004, **16**:321-328.
- [21] Wang Jun, Jebara T, Chang S F. Graph transduction via alternating minimization[C]//Proc of the 25th International Conference on Machine Learning. 2008.
- [22] Shi Jianbo, Malik J. Normalized cuts and image segmentation[J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2000, **22**(8):888-905.
- [23] Buckner R L. The brain's default network: origins and implications for the study of psychosis[J]. *Dialogues Clinical Neuroscience*, 2013, **15**(3):351-358.