

基于典型相关的脑功能网络构建方法^①

汪成林, 曾卫明, 时莹超

(上海海事大学 信息工程学院, 上海 201306)

摘要: 人脑是自然界中最复杂的系统之一, 在该系统中, 各个脑区之间相互作用构成脑网络. 为方便科学研究, 学者们通过不同的方法构建脑功能网络. 本文基于自动结构标签模板和典型相关来构建脑功能网络, 并研究了所构建网络的某些关键的拓扑性质, 如小世界性, 全局效率, 局部效率等. 与此同时, 比较了正常被试与心理异常海员脑功能网络的这些性质, 结果表明在统计意义上正常被试和心理异常海员某些性质存在显著性差异, 说明此种脑功能网络构建方法是可行的.

关键词: 典型相关; 脑功能网络; 小世界性; 心理异常海员

Approach for Estimating Functional Brain Networks Based on Canonical Correlation

WANG Cheng-Lin, ZENG Wei-Ming, SHI Ying-Chao

(College of Information Engineering, Shanghai Maritime University, Beijing 201306, China)

Abstract: Human brain is one of the most complex systems in the world, and the interaction between different regions has constituted brain network. To aid scientific research, scholars build brain functional networks with variety of methods. In this paper, brain functional networks were estimated by using AAL (Anatomical Automatic Labeling) template and Canonical Correlation, and some topology properties such as small-world property, global efficiency, local efficiency, etc. were studied. The result indicates some characteristics present significantly statistical difference between normal population and abnormal psychological sailors. The prominent conclusion demonstrates the proposed way of building functional networks is feasible.

Key words: canonical correlation; brain functional networks; small-world property; abnormal psychological sailors

现实世界中很多复杂系统都可以用网络来进行描述. 人脑是自然界中最复杂的系统之一, 在此系统中, 存在着很多个区域, 这些区域之间的相互作用是人脑进行信息处理和认知表达的结构基础^[1,2]. 因此在脑科学研究中, 可以将人脑构造成一个复杂网络^[3].

本文是基于静息态下的 fMRI(functional magnetic resonance imaging, 功能磁共振成像)数据构建脑功能网络. 首先, 因为静息态数据能很好地描述神经元的自发活动. 其次, 基于血氧水平依赖(Blood Oxygenation Level Dependent, BOLD)的 fMRI 主要是依赖血管内脱氧血红蛋白与含氧血红蛋白比例的变化来获知周围神经元活动的方法, 同时具有较高的时间分辨率和空

间分辨率, 所以在研究人类脑功能时成为了重要的检查手段^[4].

一个网络一般由节点和边组成. 在静息态脑功能网络中, 脑区域代表节点, 区域之间的关系表示连接. 如何定义节点和边是网络分析中核心的问题. 在定义网络节点方面, 本文使用 AAL(Anatomical Automatic Labeling, 自动结构标签)^[5]脑模板将全脑分割成 90 个脑功能区域, 将每个区域定义成一个网络节点; 在定义节点之间连接时, 使用典型相关计算出每个脑区时间过程之间的相关系数来度量脑网络节点间的连接; 接着设定网络的稀疏度, 生成一系列描述脑功能网络的阈值关联矩阵; 最后统计分析网络的各个特征.

^① 基金项目: 国家自然科学基金项目(31170952)

收稿时间: 2014-02-19; 收到修改稿时间: 2014-03-10

1 方法

图论^[6-8](Graph Theory)是 18 世纪发展起来的数学领域的分支之一,在图论中一个复杂网络可以表述为一个图 $G(V, E)$,它由节点(vertex 或 node)集合 V 和边(edge 或 link)集合 E 构成,节点集合 V 的大小 N 表示网络的规模,边集合 E 的大小 M 表示网络边的总数.该理论主要用于描述及分析网络的局部及整体特征,是目前研究复杂网络不可或缺的数学工具,由于它能够定量的分析节点之间的关系,因此在研究大脑功能区域之间的连接性中起到至关重要的作用.

构建脑功能网络主要包括两个关键步骤:一是如何确定网络节点;二是如何定义网络连接边.在此之前需获取 fMRI 数据,然后对 fMRI 数据进行预处理,整个网络的构建流程包括获取数据、预处理、时间过程提取、脑功能网络构建及脑功能网络拓扑性质分析等五个模块,如图 1 所示.

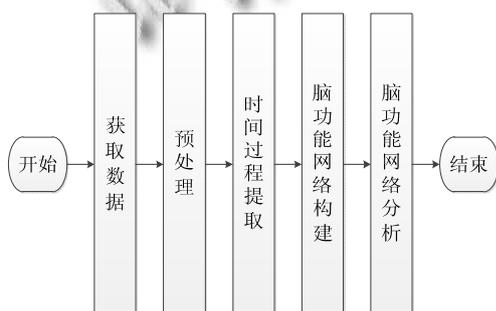


图 1 脑功能网络构建流程图

1.1 数据获取

本文所采用的数据集包括正常被试数据集和筛选出的异常海员数据集,其中海员被试数据集采集于华东师范大学的上海市核磁共振重点实验室,所用仪器为西门子磁共振仪.数据获取过程中被试要求保持大脑清醒,平躺于磁共振仪器内,不做任何定性思考,对应的参数如下:采用单次激发敏感梯度回波平面成像,切片数为 37,覆盖整个脑区,TR 为 2.0s,扫描分辨率为 64×64 ,片内分辨率为 $4 \text{ mm} \times 4 \text{ mm}$,片厚度为 4 mm ,片间间隔为 1 mm .

本文所使用的正常被试静息态数据集从神经影像公共数据库免费下载获取的(http://www.nitrc.org/projects/fcon_1000/).数据集是由 James J.Pekar 博士和 Stewart H. Mostofsky 博士发布,该数据集采用单次激发敏感梯度回波平面成像,切片数为 47 片,覆盖整

个脑区,敏感加速因子为 2.0,TR=2.5s,扫描分辨率为 96×96 ,片内分辨率为 $2.67 \text{ mm} \times 2.67 \text{ mm}$,片厚度为 3 mm .

1.2 数据预处理

fMRI 技术需要在 TR 时间内获得多个脑断层面的一系列反映脑功能变化的图像,同一层面上不同时刻的脑图像体现在脑功能的变化上.但是由于数据获取过程中由于受被试运动的影响,往往会造成图像上相应点的位置与相对应的解剖位置并不完全吻合.其次,由于 fMRI 对微小的运动很敏感,还可能受到各种随机噪声和生理噪声的影响,导致真实的神经活动信号可能会被噪声所掩盖.

综上所述,为了保证脑激活体素定位的准确性,在进行 fMRI 数据分析前需要对数据进行相应的预处理来解决图像信噪比低、图像不稳定,不同样本之间的差异性问题.本实验所有数据集的预处理采用 SPM(<http://www.fil.ion.ucl.ac.uk/spm>)软件包完成,主要包括时间层矫正(slice timing),头动矫正(realign)、图像标准化(normalization)以及平滑处理(smooth)等操作.

1.3 脑功能网络节点的定义

本文采用目前使用广泛的标准脑模板 AAL 定义节点,此模板将全脑分割成 90 个区域,即每个区域代表脑功能网络中的一个节点,区域内所有体素的时间过程代表该区域神经元活动的情况.

1.4 脑功能网络连接的定义

网络中连接的构建过程如下:

首先,相关分析.脑功能网络的连接是指不同区域神经活动信号之间的动态协调性,一般通过皮尔森相关,偏相关^[9]等统计方法来度量.考虑到每个节点区域包含多个体素,用每个区域的所有体素的时间过程的一个线性组合来代表区域的整体时间过程,因此在建立连接时本文采用方便计算多变量之间相关性的典型相关来构建网络连接,具体原理如下文所述^[10].

典型相关分析的基本原理:对于两组随机变量 $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_p)^T$, $\mathbf{Y} = (\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_q)^T$ ($p \leq q$) 像主成分分析^[11]那样,考虑 $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_p)^T$ 的一个线性组合 U 及 $\mathbf{Y} = (\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_q)^T$ 的一个线性组合 V ,希望找到的 U 和 V 之间有最大可能的相关系数,以充分反映两组变量间的关系.

首先求出两组随机变量之间的协方差矩阵:

$$A = \begin{bmatrix} \sum_{11} & \sum_{12} \\ \sum_{21} & \sum_{22} \end{bmatrix} \quad (1)$$

其中 $\sum_{11} = \text{cov}(X)$, $\sum_{22} = \text{cov}(Y)$, $\sum_{12} = \sum_{21} = \text{cov}(X, Y)$. 其次寻找变量之间的线性组合

$$U_1 = a_1^T X = a_{11}X_1 + a_{12}X_2 + \dots + a_{1p}X_p \quad (2)$$

$$V_1 = b_1^T Y = b_{11}Y_1 + b_{12}Y_2 + \dots + b_{1q}Y_q \quad (3)$$

使 U_1 和 V_1 的相关系数 $P(U_1, V_1)$ 达到最大, 这里

$$a_1^T = (a_{11}, a_{12}, \dots, a_{1p}) \quad (4)$$

$$b_1^T = (b_{11}, b_{12}, \dots, b_{1q}) \quad (5)$$

由 (1) 式得 $\text{var}(U_1) = a_1^T \sum_{11} a_1$,

$\text{var}(V_1) = b_1^T \sum_{22} b_1$, $\text{cov}(U_1, V_1) = a_1^T \sum_{12} b_1$, 所以 U_1, V_1 的相关系数为

$$\rho(U_1, V_1) = \frac{a_1^T \sum_{12} b_1}{\sqrt{a_1^T \sum_{11} a_1} \sqrt{b_1^T \sum_{22} b_1}} \quad (6)$$

两组随机变量之间的线性组合一般不止一种, 因此相关系数也不止一种, 本研究中取第一个相关系数, 即最大的相关系数.

综上, 用典型相关计算相关性时, 首先是求出 90 个区域各个区域中所有体素所对应的时间过程的线性组合, 用来表示该 90 个节点区域各自的整体时间过程和脑功能活动属性, 这些时间过程间的同步性来表示节点区域间的功能连通关系.

其次, 阈值设定. 在研究过程中, 一般认为脑功能网络特性复杂且稀疏, 以大量的短程边为主, 少量的长程边为辅. 因此, 设定阈值是脑功能网络构建过程中比较灵活的一步. 阈值的设定主要是确定节点与节点之间连接, 它是将相关值小的认为是噪声边, 加以去除. 同时又要保证网络无孤立节点区域. 本文选取阈值的策略是: 脑功能网络平均度值要满足大于 $2 \ln N$, 且网络密度要小于 50%^[12].

最后, 建立连接. 使用典型相关分析计算 90 个区域之间的相关性, 得到的相关系数用来判断脑功能网络节点与节点之间的连接情况. 如果相关系数的值大

于阈值, 认为节点之间存在边, 即建立连接, 将对应位置的置为 1; 反之则不建立连接, 将对应位置的置为 0. 考虑到自相关的问题, 需要将相关系数矩阵对角线元素都置为 0, 即防止出现自身与自身相连的问题.

经过以上步骤, 便建立了邻接矩阵, 即脑功能网络. 在建立网络的过程中, 阈值的选择一定要满足网络密度的要求, 本文网络密度设置在 10%~50%.

1.5 网络特征

本文主要分析网络的聚类系数, 平均最短路径长度, 小世界性, 全局系数, 局部系数^[13], 这五种拓扑性质, 并将正常被试群体和心理异常海员群体的这五个性质相比较, 从而说明本文所构造网络的可行性.

聚类系数(clustering coefficient). 聚类系数是衡量网络集团化的程度, 表示某一节点邻居节点 i 之间相互连接程度. 节点 i 的聚类系数 C_i 等于此节点邻居间

实际存在的边数(e_i)与最大可能的存在的边数($N_i(N_i - 1)/2$)的比值(图 2), 如公式(7)所示,

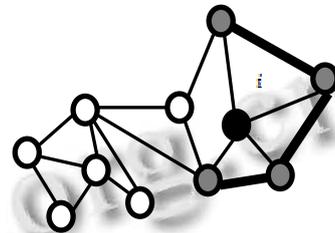


图 2 网络 $G(11,18)$ 中黑色实心节点 i 有 4 个邻居节点(灰色实心节点), 且这 4 个节点之间有 3 条边(黑色粗实线)相连, 故节点 i 的聚类系数 $C_i = 3/C_4^2 = 0.5$

$$C_i = \frac{2e_i}{N_i(N_i - 1)} \quad (7)$$

其中 N_i 为 i 节点的邻居节点的个数. 整个网络的聚类系数为网络中所有节点聚类系数的平均, 如公式(8)所示,

$$C = \frac{1}{N} \sum_{i \in V} C_i \quad (8)$$

不难知一个网络的聚类系数 C 应该满足 $0 < C < 1$.

虽然聚类系数能描述网络的集团化程度, 但是由于

其只考虑了邻居节点之间的连接,于是在此基础上 Latora V, Marchiori M 提出了局部效率(local efficiency)的概念 Eloc 的概念^[14]. 节点 i 的局部效率如公式(9)所示,

$$E(i) = \frac{1}{N_{G_i}(N_{G_i} - 1)} \sum_{j \neq k \in G_i} \frac{1}{l_{jk}} \quad (9)$$

其中, G_i 指节点 i 的邻域所构成的子图, l_{jk} 表示节点 j, k 之间的最短路径长度(即边数最少的一条通路). 网络的局部效率为所有节点局部效率的平均, 如公式(10)所示,

$$E_{loc} = \frac{1}{N} \sum_{i \in V} E(i) \quad (10)$$

聚类系数和局部效率反映了网络传递信息的能力, 在一定程度上还反应了脑功能网络抵御疾病的能力.

最短路径长度(shortest path length). 最短路径长度在信息传递中起着重要作用, 路径越短传递效率越高. 它表示一个节点到另一个节点的最佳路径, 通过最短路径能更快的传输信息, 从而节省资源. 两个节点之间边数最少的一条通路为这两个节点之间的最短路径(图 3). 一个网络的最短路径表示为所有节点到其他节点最短路径的均值, 如公式(11)所示,

$$L = \frac{1}{N(N-1)} \sum_{i, j \in V, i \neq j} l_{ij} \quad (11)$$

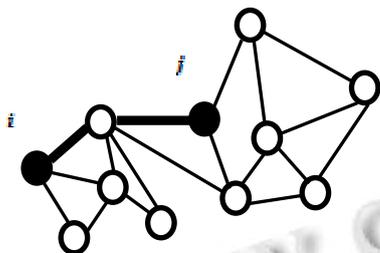


图 3 网络 $G(11, 18)$ 中 2 个实心节点 i, j 之间的最短路径长度(黑色粗实线) $l_{ij} = 2$

当网络中存在不连接的节点时, 其最短路径的便会达到无穷, 因此 Latora V, Marchiori M 等人提出了全局效率(global efficiency) E_{glob} ^[14], 如公式(12)所示,

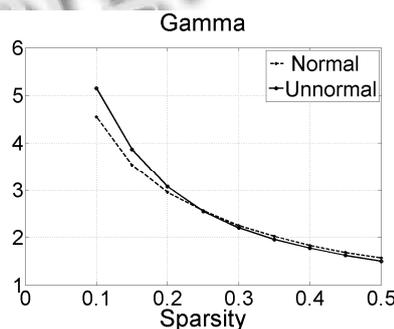
$$E_{glob} = \frac{1}{N(N-1)} \sum_{i, j \in V, i \neq j} \frac{1}{l_{ij}} \quad (12)$$

效率将聚类系数和最短路径融合成了一个整体, 作为新的网络度量指标为复杂网络的研究开辟了新空间.

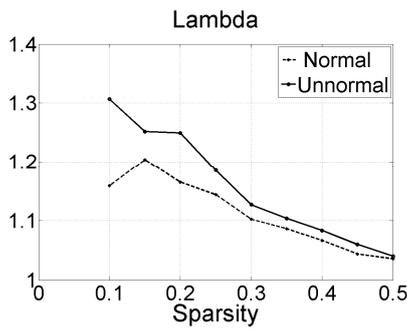
在研究网络时, Watts 和 Strogatz 发现存在这样的网络, 其具有复杂网络的较短的最短路径, 又具有规则网络的较高的聚类系数, 这样的网络被称为“小世界”网络^[15], 为了量化“小世界”网络, Watts 和 Strogatz 将随机网络作为参照, 如果所研究的网络相对随机网络具有较大的聚类系数, 且最短路径长度相近, 则认为该网络具有小世界性, 即 $\gamma = C_{real} / C_{random} > 1$, $\lambda = L_{real} / CL_{random} \sim 1$ (random 表示随机网络, real 表示真实网络). 之后, 为了方便度量小世界特性, Humphries 等人^[16]将两个度量指标统一成一个度量指标来衡量“小世界”特性, 即 $\sigma = \gamma / \lambda$, 如果 $\sigma > 1$ 则网络具有“小世界”性, 且值越大说明网络的“小世界”性越强.

2 实验统计结果

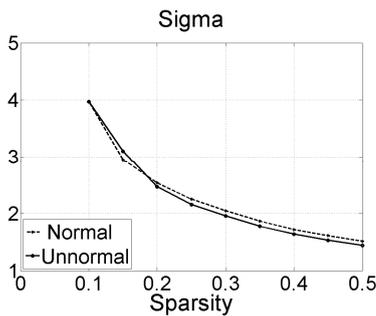
本文采用显著性水平 $\alpha=0.05$ 对区域之间的功能连接显著性进行判定. 稀疏度阈值 s (定义见表 1)从 0.1 开始, 每次增加 0.05, 一直增加到 0.5, 每个阈值对应一个脑功能网络邻接矩阵. 经过统计发现, 当阈值 s 大于 0.1 时平均值大于 $2 \times \ln(90) = 8.9$, 且平均值随着阈值的增大而增大; 当阈值 s 为 0.15 时, 最短路径长度、聚类系数、全局效率、局部效率、小世界性均服从正态分布, 且在置信水平为 0.05 的情况下, 正常人和心理异常海员两种对照组的小世界性没有显著性差异, 但最短路径长度、聚类系数、全局效率、局部效率存在显著性差异, 实验结果如下图所示:



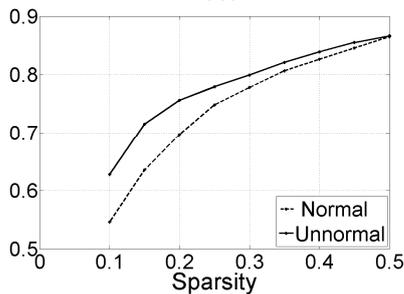
(a) 平均聚类系数的比值



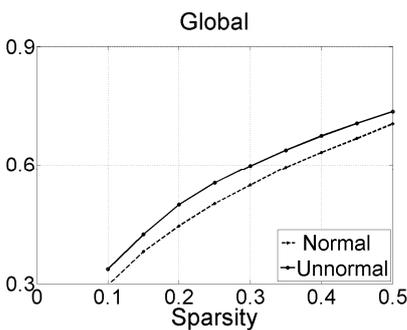
(b) 平均最短路径的比值



(c) 小世界性
Elocal



(d) 局部系数



(e) 全局系数

图 4 网络特征比较

图中各参数所表示的含义见表 1.

表 1 特征参数

参数	拓扑性质	含义	范围
Gamma	γ	构建的网络与随机网络平均聚类系数的比值	$(0, \infty)$
Lambda	λ	构建的网络与随机网络平均最短路径的比值	$(0, \infty)$
Sigma	σ	构建的网络的小世界性	$[1, \infty)$
Elocal	E_{local}	构建的网络的平均局部效率	$[0, 1]$
Global	E_{global}	构建的网络的平均全局效率	$[0, 1]$
Sparsity	S	网络稀疏度	$[0.2, 0.4]$

3 应用展望

本文构建了心理异常海员脑功能网络, 并通过与正常被试对照组对比分析, 发现了脑功能网络最短路径长度、聚类系数、全局效率等参数存在显著性差异, 从而有望发展一种海员心理异常检测的定量性生物标记. 除此以外, 本文脑功能网络构建方法还有望应用于不同类型的脑疾病研究中, 如精神分裂症、儿童多动症、抑郁症等. 通过分析由疾病引起的脑功能网络拓扑性质的异常变化, 从而在系统水平上为揭示脑疾病的病理及生理机制提供新的启示, 并在此基础上建立描述疾病的脑功能网络影像学标记, 为病人的早期诊断和疗效评价等提供重要的辅助工具.

4 结语

本文探讨了一种基于 AAL 模板和典型相关构建脑功能网络的方法. 实验结果表明, 构建出来的脑功能网络具有“小世界”特性, 可以很好地描述脑内信息加工处理的高效率. 此外该方法已经在心理异常海员数据上得到了验证, 其结果是可靠的. 然而本文所研究的脑功能网络为无向网络, 虽然研究起来比较容易, 但是由于无向网络忽略了脑功能网络中非常重要的方向信息和时间先后信息, 因此构建有向功能网络来刻画不同脑区连接的方向性及神经活动之间的因果关系是我们今后的研究方向.

参考文献

1 Sporns O, Chialvo DR, Kaiser M, et al. Organization, development and function of complex brain networks. Trends Cogn Sci, 2004, 8: 418-425.

- 2 Wang F, Kalmar JH, He Y, et al. Functional and structural connectivity between the perigenual anterior cingulate and amygdala in bipolar disorder. *Biological Psychiatry*, 2009, 66(5): 516-521.
- 3 Bullmore E, Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 2009, 10(3): 186-198.
- 4 Ogawa S, Lee TM, Kay AR, et al. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc. of the National Academy of Sciences*, 1990, 87(24): 9868-9872.
- 5 Tzourio-Mazoyer N, Landeau B, Papathanassiou D, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, 2002, 15(1): 273-289.
- 6 Reijneveld JC, Ponten SC, Berendse HW, et al. The application of graph theoretical analysis to complex networks in the brain. *Clinical Neurophysiology*, 2007, 118(11): 2317-2331.
- 7 Huber W, Carey VJ, Long L, et al. Graphs in molecular biology. *BMC Bioinformatics*, 2007, 8(S6): S8.
- 8 Bullmore E, Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 2009, 10(3): 186-198.
- 9 Liang X, Wang J, Yan C, et al. Effects of different correlation metrics and preprocessing factors on small-world brain functional networks: a resting-state functional MRI study. *PLoS One*, 2012, 7(3): e32766.
- 10 Friman O, Cedefamn J, Lundberg P, et al. Detection of neural activity in functional MRI using canonical correlation analysis. *Magnetic Resonance in Medicine*, 2001, 45(2): 323-330.
- 11 Lai SH, Fang M. A novel local PCA-based method for detecting activation signals in fMRI. *Magnetic Resonance Imaging*, 1999, 17(6): 827-836.
- 12 薛绍伟,唐一源,李健,等.一种基于 fMRI 数据的脑功能网络构建方法. *计算机应用研究*, 2010, 27(11).
- 13 梁夏,王金辉,贺永.人脑连接组研究:脑结构网络和脑功能网络. *科学通报*, 2010(16): 1565-1583.
- 14 Latora V, Marchiori M. Efficient behavior of small-world networks. *Physical Review Letters*, 2001, 87(19): 198701.
- 15 Watts DJ, Strogatz SH. Collective dynamics of "small-world" networks. *Nature*, 1998, 393(6684): 440-442.
- 16 Humphries MD, Gurney K, Prescott TJ. The brainstem reticular formation is a small-world, not scale-free, network. *Proc. of the Royal Society B: Biological Sciences*, 2006, 273(1585): 503-511.