

一种基于模式识别分类提取磁共振成像脑激活区的方法

申请号 : 200710098691.3

申请日 : 2007-04-25

申请(专利权)人 中国科学院自动化研究所

地址 100080北京市海淀区中关村东路95号

发明(设计)人 田捷 甄宗雷

主分类号 A61B5/055(2006.01)I

分类号 A61B5/055(2006.01)I G01R33/54(2006.01)I
G06F17/00(2006.01)I G06T7/60(2006.01)I

公开(公告)号 101292871

公开(公告)日 2008-10-29

专利代理机构 中科专利商标代理有限责任公司

代理人 周国城



(12) 发明专利

(10) 授权公告号 CN 101292871 B

(45) 授权公告日 2010.05.26

(21) 申请号 200710098691.3

(22) 申请日 2007.04.25

(73) 专利权人 中国科学院自动化研究所

地址 100080 北京市海淀区中关村东路 95
号

(72) 发明人 田捷 颖宗雷

(74) 专利代理机构 中科专利商标代理有限责任
公司 11021

代理人 周国城

(51) Int. Cl.

A61B 5/055 (2006.01)

G01R 33/54 (2006.01)

G06F 17/00 (2006.01)

G06T 7/60 (2006.01)

(56) 对比文件

CN 1939211 A, 2007.04.04, 全文 .

EP 1219243 A1, 2002.07.03, 全文 .

CN 1294875 C, 2007.01.17, 全文 .

CN 1626031 A, 2005.06.15, 全文 .

US 6292683 B1, 2001.09.18, 全文 .

JP 特开 2007-44104 A, 2007.02.22, 全文 .

WO 03/017177 A2, 2003.02.27, 全文 .

CN 1620666 A, 2005.05.25, 全文 .

审查员 沈显华

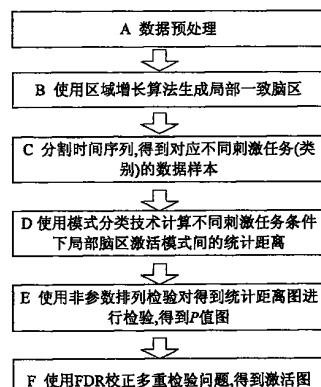
(54) 发明名称

一种基于模式识别分类提取磁共振成像脑激活区的方法

(57) 摘要

本发明公开一种基于模式识别分类提取磁共振成像脑激活区的算法,基于局部脑区包含的精细活动模式间的多元模式距离来提取脑激活区,对 fMRI 图像进行预处理;聚类得到局部一致性脑区;利用局部一致脑区内多个体素的联合活动构造多元模式;利用模式分类方法构造多元距离函数来度量不同刺激条件下局部脑区活动的可分离性质,判断脑区是否激活。本发明直接使用局部脑区内多体素构成的多元模式信息来表示不同刺激条件下的脑活动,多元模式可以全面的反映局部脑活动状态,而多元统计距离可以有效的集成包含局部脑区内的信息来度量不同脑激活状态间的差异,两者一同保证了本发明可以比传统 fMRI 分析技术更完整地检测到精细脑激活模式。

权利要求书 2 页 说明书 10 页 附图 5 页



1. 一种基于模式识别分类提取磁共振成像脑激活区的方法,其特征在于:包含步骤:
预处理步骤 A:对采集到的磁共振图像进行预处理,去除混淆因素,用于获得标准化图像数据;

构造局部一致脑区步骤 B:对图像数据的每个体素 v_0 使用区域增长算法,以 v_0 作为种子点,得到包含 K 个体素的局部一致脑区 $N(v_0)$;

分割时间序列步骤 C:根据实验设计对应的刺激条件类别分割局部一致脑区 $N(v_0)$ 内多个体素的时间序列,则每类刺激条件对应着一组不同的多元数据样本;

构造多元模式统计距离步骤 D:设不同刺激条件下,局部一致脑区 $N(v_0)$ 内的 K 个体素构成的空间模式对应着不同的 K 维随机分布;使用模式识别方法构造多元统计距离函数,用以度量局部一致脑区 $N(v_0)$ 内不同条件对应的数据样本间的统计距离;通过对每个体素迭代以上过程,计算多元统计距离,得到多元统计距离参数图;

多元距离统计参数图假设检验步骤 E:假设在各局部一致脑区 $N(v_0)$ 活动模式在不同刺激条件下的活动相同,使用非参数序列检验对每个体素对应的统计距离参数进行检验,得到显著性非参数序列检验值图;

假设检验多重比较校正步骤 F:利用错误发现率 FDR 方法,用于消除多重比较,得到能显著区分不同试验条件的局部脑区,即对应激活区。

2. 如权利要求 1 所述提取磁共振成像脑激活区的方法,其特征在于,所述获取标准化图像数据的预处理步骤具体包括:

步骤 A. 1 对图像数据不同样本进行标准化,使其均值为 0, 标准差为 1;

步骤 A. 2 对图像数据不同特征进行标准化,使其均值为 0, 标准差为 1。

3. 如权利要求 1 所述提取磁共振成像脑激活区的方法,其特征在于,所述构造局部一致脑区步骤具体包括:

步骤 B. 1 使用体素活动时间序列的皮尔逊 Pearson 相关系数作为准则,来度量不同体素间的活动相似性质;

步骤 B. 2 使用体素 v_0 作为种子点,区域增长算法每次选择同种子点活动最相似的邻域体素作为加入到局部一致脑区中;区域增长迭代进行,达到指定的局部区域大小后停止;从而得到的局部区域内的体素活动均基本一致。

4. 如权利要求 1 所述提取磁共振成像脑激活区的方法,其特征在于,所述构造多元模式统计距离步骤具体包括:

步骤 D. 1:体素 v_0 局部一致脑区 $N(v_0)$ 内的在不同刺激条件 X, Y 下,测得的局部区域 K 个体素的联合活动,活动是从两个不同的多元随机变量 $X = (X_1, X_2, \dots, X_i, \dots, X_K)^T$, $Y = (Y_1, Y_2, \dots, Y_i, \dots, Y_K)^T$, $i = 1, 2, \dots, K$ 采样得到的模式样本;

步骤 D. 2:设刺激条件 X 和 Y 分别对应样本集合 S_x 和 S_y ;根据模式识别方法,根据 Fisher 线性判别分析导出多元距离函数来度量样本集合 S_x 和 S_y 间的距离;

步骤 D. 3:根据 Fisher 线性判别分析,集合 S_x 和 S_y 间的 Fisher 线性判别分析通过下式计算:

$$D^2 = (\mathbf{u}_{S_x} - \mathbf{u}_{S_y})^T \Sigma^{-1} (\mathbf{u}_{S_x} - \mathbf{u}_{S_y})$$

其中, \mathbf{u}_{S_x} , \mathbf{u}_{S_y} 分别是样本 S_x 和 S_y 的平均向量即:

$$\mathbf{u}_{S_x} = \frac{1}{n_x} \sum_{j=1}^{n_x} \mathbf{X}(j), \mathbf{u}_{S_y} = \frac{1}{n_y} \sum_{j=1}^{n_y} \mathbf{X}(j)$$

Σ 是样本混合协方差 :

$$\Sigma = \frac{(\mathbf{n}_x - 1)\Sigma_{S_x} + (\mathbf{n}_y - 1)\Sigma_{S_y}}{(\mathbf{n}_x + \mathbf{n}_y - 2)}$$

$\Sigma_{S_x}, \Sigma_{S_y}$ 表示 S_x 与 S_y 的估计协方差阵, n_x, n_y 指样本 S_x 及 S_y 的大小; Fisher 线性判别分析体现了在最佳判别轴上, 两类刺激条件引起的局部脑活动模式的可分离程度。

5. 如权利要求 1 所述提取磁共振成像脑激活区的方法, 其特征在于, 多元距离统计参数图假设检验步骤具体包括 :

在各局部脑区 $N(v_0)$ 活动模式在不同刺激条件下的脑活动相同即 : 不可显著分离的零假设下, 使用非参数序列检验对每个体素对应的统计距离参数进行检验, 得到显著性非参数序列检验值图 :

通过如下的磁共振图像数据重采样来进行非参数序列检验 : 数据空间模式不变, 磁共振图像数据对应的时间序列进行随机重排; 随机重排破坏了磁共振图像信号和试验条件的相关性, 保留了完整的空间结构, 从而计算不同局部脑活动模式间的多元距离统计量对应的非参数序列检验值。

6. 如权利要求 1 所述提取磁共振成像脑激活区的方法, 其特征在于, 假设检验多重比较校正步骤具体包括 :

步骤 F. 1, 选择一个错误发现率 FDR 的边界值 $q (0 < q < 1)$, 是所期望的最大的错误发现率 FDR;

步骤 F. 2, 由非参数排列检验得到非参数序列检验值图是非参数序列检验 p 值图, 对非参数序列检验 p 值图进行从小到大对值排序 : $p(1) \leq p(2) \leq \dots \leq p(m)$, 设 $p(i)$ 对应着体素统计量 $v(i)$, m 是所检验的 fMRI 数据共有的体素个数;

步骤 F. 3, 假设 r 是满足不等式 $p(i) \leq \frac{i}{V} \frac{q}{c(m)}$ 的最大 i 值, 其中 $c(m)$ 是预先设定的常量, 它的选择和体素的分布情况有关, 不同条件下有 2 种选择 : $c(m) = 1, c(m) = \sum_{i=1}^V \frac{1}{i}$;

步骤 F. 4, 得到结果 : 拒绝空假设条件即真正激活的体素为 $v(1), v(2), \dots, v(r)$, 即是统计量大于 $v(r)$ 的体素是激活的体素, 称 $v(r)$ 为多检验确定的阈值;

步骤 F. 5, 根据 FDR 校正得到的多检验阈值 $v(r)$, 对统计参数图进行阈值分割, 得到那些能显著区分不同试验条件的脑激活区。

一种基于模式识别分类提取磁共振成像脑激活区的方法

技术领域

[0001] 本发明属于神经影像数据分析技术领域,具体涉及磁共振成像 fMRI 激活区提取算法,尤其涉及使用多元模式识别方法进行 fMRI 脑激活区。

背景技术

[0002] 随功能磁共振成像 (functional Magnetic Resonance Imaging, fMRI) 以其高时空分辨率,非侵入式等特点在人脑功能研究中得到了广泛应用。fMRI 一般是指基于血氧水平依赖 (blood oxygen level-dependent, BOLD) 的 fMRI 成像,它通过测量由神经活动引起的脑血流和脑血氧等成分变化而造成的磁共振信号变化来反应脑活动。随着近年来试验数据的急剧增长,合理有效的 fMRI 数据分析技术显的越发重要。功能数据分析的核心问题是根据测得的 fMRI 数据寻找其活动能显著区分不同试验条件 (如,刺激条件和基线条件) 的脑区。合理的 fMRI 分析方法需同时考虑到人脑功能一般组织原则以及 fMRI 数据本身一些特性。

[0003] 一般认为,脑功能遵循两个基本组织原则 :功能集成化和功能特异化。在大空间刻度 (规模) 上,一个复杂的脑功能可能会由许多功能特异的脑区通过相互作用 (集成) 来完成 ;反过来,一个特异性脑区也会对许多不同的认知任务进行表示或加工,通过精细空间刻度上不同的分布式脑活动来对外部不同刺激进行表示。另一方面,随着 MRI 技术的进步, fMRI 图像空间分辨率一直在提高。传统 fMRI 数据在体素各维宽度约为 4 毫米时可获得较高的信噪比。当前,各维宽度为 2 毫米的体素已经可以在标准 3T 临床磁共振 (MRI) 机器上可靠地获得。随着超高场 ($\geq 4T$) 磁共振技术的使用, fMRI 图像的空间分辨率正在向亚毫米的刻度上逼近。fMRI 图像空间分辨率的提高,为我们更精细的空间刻度上研究脑功能提供了可能 :然而,传统的 fMRI 数据分析大多基于广义线性模型 (General Linear Model, GLM), GLM 本质上是一种一元统计技术,它通过孤立地分析每个体素的时间序列来决定该体素是否被激活,完全忽略了不同体素间的相互关系 (尤其是隐含在局部脑区内的模式信息)。因而传统一元分析技术无法检测到试验刺激引起的精细刻度上的脑活动模式,及构成这些模式的对刺激反应微弱的体素。另外,为了提高信噪比和统计效力,一元分析技术通常会采用高斯核对 fMRI 数据进行空间平滑。数据平滑本质上相当于低通滤波,它会模糊那些含有神经科学相关信息的精细的脑活动信号。从而,大量的精细信息将会被滤除掉, fMRI 提供的高空间分辨率信息仍远没有被利用到。

发明内容

[0004] 为了能在更精细的刻度上区分由不同刺激引起的脑活动模式,充分利用 fMRI 数据中包含的信息,本发明描述了一类全新的,基于模式识方法检测脑激活区的算法 (local multivariate distance mapping brainactivation, LMDM)。本发明直接使用局部脑区内多体素构成的多元模式信息来表示不同刺激条件下的脑活动,进而以不同模式间的多元距离作为统计量度量局部脑区活动在不同刺激条件下的可分离性。多元模式可以全面地反

映局部脑活动状态,而多元统计距离作为统计量则能更有效地集成局部脑区内的多体素信息用以区分不同脑激活状态。从而,相对于一元分析技术,本发明可以更精确地提取脑激活区,判别不同认知条件下脑活动状态。

[0005] GLM 方法假定各个体素是相互独立的,从而在进行脑激活区定位时,完全忽略了包含在局部脑区内的模式信息。从认知神经科学的角度说,它忽略了人脑特定功能和 fMRI 数据中单个体素间不存在一一对应关系的可能性。为了解决现有 fMRI 数据分析技术种中存在的缺点,能在更精细的刻度上区分由不同刺激引起的脑活动模式,充分利用 fMRI 数据中包含的信息,本发明的目的是提供一类能更精确地提取脑激活区,判别不同认知条件下脑活动状态的基于模式识方法的脑激活区检测算法。

[0006] 为了实现本发明的目的,本发明提供一类基于模式识方法的脑激活区检测算法,包括步骤:

[0007] 预处理步骤 A :对采集到的磁共振图像进行预处理,去除混淆因素,用于将获得标准化图像数据;

[0008] 构造局部一致脑区步骤 B :对图像数据的每个体素 v_0 使用区域增长算法以 v_0 作为种子点,得到包含 K 个体素的局部一致脑区 $N(v_0)$;

[0009] 分割时间序列步骤 C :根据实验设计对应的刺激条件类别分割局部一致脑区 $N(v_0)$ 内多个体素的时间序列,则每类刺激条件对应着一组不同的多元数据样本;

[0010] 构造多元模式统计距离步骤 D :设不同刺激条件下,局部一致脑区 $N(v_0)$ 内的 K 个体素构成的空间模式对应着不同的 K 维随机分布;使用模式识别方法构造多元统计距离函数,用以度量局部一致脑区 $N(v_0)$ 内不同条件对应的数据样本间的统计距离;通过对每个体素迭代以上过程,计算多元统计距离,得到多元统计距离参数图;

[0011] 多元距离统计参数图假设检验步骤 E :假设在各局部一致脑区 $N(v_0)$ 活动模式在不同刺激条件下的活动相同,使用非参数序列检验对每个体素对应的统计距离参数进行检验,得到显著性非参数序列检验值图;

[0012] 假设检验多重比较校正步骤 F :利用错误发现率 FDR 方法,用于消除多重比较,得到能显著区分不同试验条件的局部脑区,即对应激活区。

[0013] 根据本发明的实施例,预处理步骤所述标准化图像数据,还包括:

[0014] 步骤 A. 1 :对图像数据不同样本进行标准化,使其均值为 0,标准差为 1;

[0015] 步骤 A. 2 :对图像数据不同特征进行标准化,使其均值为 0,标准差为 1。

[0016] 根据本发明的实施例,所述构造局部一致脑区步骤,还包括:

[0017] 步骤 B. 1 :使用体素活动时间序列的皮尔逊 Pearson 相关系数作为准则,来度量不同体素间的活动相似性质;

[0018] 步骤 B. 2 :使用体素 v_0 作为种子点,区域增长算法每次选择同种子点活动最相似的领域体素作为加入到局部一致脑区中;区域增长迭代进行,达到指定的局部区域大小后停止;从而得到的局部区域内的体素活动均基本一致。

[0019] 根据本发明的实施例,所述构造多元模式统计距离步骤,还包括:

[0020] 步骤 D. 1 :体素 v_0 局部一致脑区 $N(v_0)$ 内的在不同刺激条件 X, Y 下,测得的局部区域 K 个体素的联合活动,活动是从两个不同的多元随机变量 $X = (X_1, X_2, \dots, X_i, \dots, X_K)^T$, $Y = (Y_1, Y_2, \dots, Y_i, \dots, Y_K)^T$, $i = 1, 2, \dots, K$ 采样得到的模式样本;

[0021] 步骤 D. 2 :设刺激条件 X 和 Y 分别对应样本集合 S_x 和 S_y ;根据模式识别方法,构造不同的多元距离函数来度量样本集合 S_x 和 S_y 间的距离,这个多元距离是根据 Fisher 线性判别分析导出的 Fisher 距离函数,根据支持向量机导出的最大边界距离函数,及由其它模式分类算法导出的多元距离函数;

[0022] 步骤 D. 3 :这里以 FLDF 为例进行说明,根据 Fisher 线性判别分析,集合 S_x 和 S_y 间的 FLDF 通过下式计算:

$$[0023] D^2 = (\mathbf{u}_{S_x} - \mathbf{u}_{S_y})^T \Sigma^{-1} (\mathbf{u}_{S_x} - \mathbf{u}_{S_y})$$

[0024] 其中, \mathbf{u}_{S_x} , \mathbf{u}_{S_y} 分别是样本 S_x 和 S_y 的平均向量即:

$$[0025] \mathbf{u}_{S_x} = \frac{1}{n_x} \sum_{j=1}^{n_x} \mathbf{X}(j), \quad \mathbf{u}_{S_y} = \frac{1}{n_y} \sum_{j=1}^{n_y} \mathbf{X}(j)$$

[0026] Σ 是样本混合协方差:

$$[0027] \Sigma = \frac{(\mathbf{n}_x - 1) \Sigma_{S_x} + (\mathbf{n}_y - 1) \Sigma_{S_y}}{(\mathbf{n}_x + \mathbf{n}_y - 2)}$$

[0028] Σ_{S_x} , Σ_{S_y} 表示 S_x 与 S_y 的估计协方差阵, n_x, n_y 指样本 S_x 及 S_y 的大小。FLDF 体现了在最佳判别轴上,两类刺激条件引起的局部脑活动模式的可分离程度;使用模式分类中的多元距离函数来度量不同刺激条件激活的局部脑活动模式间的差异,选择其它的多元距离函数来替换 FLDF 进行相关分析,其原理和步骤和使用 FLDF 相同。

[0029] 根据本发明的实施例,多元距离统计参数图假设检验步骤:

[0030] 在各局部脑区 $N(v_0)$ 活动模式在不同刺激条件下的活动相同即:不可显著分离的零假设下,使用非参数序列检验对每个体素对应的统计距离参数进行检验,得到显著性非参数序列检验值图;

[0031] 通过如下的磁共振图像数据重采样来进行非参数序列检验:数据空间模式不变,磁共振图像数据对应的时间序列进行随机重排;随机重排破坏了磁共振图像信号和试验条件的相关性,保留了完整的空间结构,从而计算不同局部脑活动模式间的多元距离统计量对应的非参数序列检验值。

[0032] 根据本发明的实施例,假设检验多重比较校正步骤,还包括:

[0033] 步骤 F. 1 :选择一个错误发现率 FDR 的边界值 $q (0 < q < 1)$, 是所期望的最大的错误发现率 FDR;

[0034] 步骤 F. 2 :对由非参数排列检验得到的非参数序列检验 p 值图进行从小到大对值排序: $p(1) \leq p(2) \leq \dots \leq p(m)$, 设 $p(i)$ 对应着体素统计量 $v(i)$, m 是所检验的 fMRI 数据共有的体素个数;

[0035] 步骤 F. 3 :假设 r 是满足不等式 $p(i) \leq \frac{i}{V c(m)} q$ 的最大 i 值,其中 $c(m)$ 是预先设定的

常量,它的选择和体素的分布情况有关,不同条件下有 2 种选择: $c(m) = 1$, $c(m) = \sum_{i=1}^r \frac{1}{i}$;

[0036] 步骤 F. 4 :得到结果:拒绝空假设条件即真正激活的体素为 $v(1), v(2), \dots, v(r)$,即是统计量大于 $v(r)$ 的体素是激活的体素,称 $v(r)$ 为多检验确定的阈值;

[0037] 步骤 F. 5 :根据 FDR 校正得到的多检验阈值 $v(r)$,对统计参数图进行阈值分割,得

到那些能显著区分不同试验条件的脑激活区。

[0038] 本发明的有益效果是：

[0039] 无论在大空间刻度层次还是在精细刻度层次上, fMRI 数据都包含着丰富的脑活动信息。然而, 直到最近, 关于如何直接提取, 利用隐藏在 fMRI 数据中的精细信息的研究还比较少。传统一元分析技术仅依靠单个体素的时间序列信息来提取脑激活区, 全完忽略了包含在局部脑区内的空间模式信息。另外, 一元分析技术通常采用的空间平滑预处理也会不合理地滤除掉许多隐含在数据中的精细刻度上的高频信息。本文提出了一种新的激活区检测的算法, 该方法首先使用多个体素构成的空间模式向量间来表示局部脑区内脑活动状态, 进而使用多元距离统计量来集成局部脑区信息, 判别不同刺激条件下脑活动差异。多元模式可以全面的反映局部脑活动状态, 而多元统计距离可以有效的集成包含局部脑区内的信息用以度量不同脑激活状态间的差异, 两者一同保证了 LMDM 方法可以比传统一元分析技术更完整地检测到精细脑激活模式。仿真试验和真实数据试验结果均证明该算法比传统一元分析方法能更完整地检测到脑激活区, 为脑功能数据分析提供了一种新途径。

附图说明

- [0040] 图 1 是本发明算法计算流程示意图。
- [0041] 图 2a 和图 2b 是本发明中提取局部一致脑区示意图。
- [0042] 图 3 是本发明中用以度量不同激活模式间统计距离的模式分类概念示意图。
- [0043] 图 4 是实例中本发明算法和传统激活区检测算法的对比。
- [0044] 图 5 是实例中本发明算法和传统激活区检测算法的对比。
- [0045] 图 6 是实例中本发明算法和传统激活区检测算法的对比。
- [0046] 图 7 是实例中本发明算法和传统激活区检测算法的对比。
- [0047] 图 8 是实例中本发明算法和传统激活区检测算法的对比。

具体实施方式

[0048] 下面结合附图 1 和实施例对本发明作进一步说明, 应指出的是, 所描述的实施例仅旨在便于对本发明的理解, 而对其不起任何限定作用。

[0049] 下面对本发明的方法涉及的步骤进行逐一详细说明 :

[0050] 首先介绍, 具体实施例的形式如下所述 :

[0051] 基于局部脑区包含的精细活动模式间的多元模式距离来提取脑激活区, 对 fMRI 图像进行预处理; 聚类得到局部一致性脑区; 利用局部一致脑区内多个体素的联合活动构造多元模式; 利用模式分类方法构造多元距离函数来度量不同刺激条件下局部脑区活动的可分离性质, 判断脑区是否激活。本发明直接使用局部脑区内多体素构成的多元模式信息来表示不同刺激条件下的脑活动, 多元模式可以全面的反映局部脑活动状态, 而多元统计距离可以有效的集成包含局部脑区内的信息来度量不同脑激活状态间的差异, 两者一同保证了本发明可以比传统 fMRI 分析技术更完整地检测到精细脑激活模式。

[0052] 一、预处理步骤所述标准化图像数据包括 :

- [0053] 步骤 A. 1 对图像数据不同样本进行标准化, 使其均值为 0, 标准差为 1;
- [0054] 步骤 A. 2 对图像数据不同特征进行标准化, 使其均值为 0, 标准差为 1。

[0055] 二、构造局部一致脑区步骤,还包括:

[0056] 步骤 B. 1 使用体素活动时间序列的皮尔逊 Pearson 相关系数作为准则,来度量不同体素间的活动相似性质;

[0057] 步骤 B. 2 使用体素 v_0 作为种子点,区域增长算法每次选择同种子点活动最相似的领域体素作为加入到局部一致脑区中;区域增长迭代进行,达到指定的局部区域大小后停止;从而得到的局部区域内的体素活动均基本一致。

[0058] 三、构造多元模式统计距离步骤,还包括:

[0059] 步骤 D. 1 :体素 v_0 局部一致脑区 $N(v_0)$ 内的在不同刺激条件 X, Y 下,测得的局部区域 K 个体素的联合活动,活动是从两个不同的多元随机变量 $X = (X_1, X_2, \dots, X_i, \dots, X_K)^T$, $Y = (Y_1, Y_2, \dots, Y_i, \dots, Y_K)^T$, $i = 1, 2, \dots, K$ 采样得到的模式样本;

[0060] 步骤 D. 2 :设刺激条件 X 和 Y 分别对应样本集合 S_x 和 S_y ;根据模式识别方法,构造不同的多元距离函数来度量样本集合 S_x 和 S_y 间的距离,这个多元距离是根据 Fisher 线性判别分析导出的 Fisher 距离函数,根据支持向量机导出的最大边界距离函数,及由其它模式分类算法导出的多元距离函数;

[0061] 步骤 D. 3 :这里以 FLDF 为例进行说明,根据 Fisher 线性判别分析,集合 S_x 和 S_y 间的 FLDF 通过下式计算:

$$[0062] D^2 = (\mathbf{u}_{S_x} - \mathbf{u}_{S_y})^T \Sigma^{-1} (\mathbf{u}_{S_x} - \mathbf{u}_{S_y})$$

[0063] 其中, \mathbf{u}_{S_x} , \mathbf{u}_{S_y} 分别是样本 S_x 和 S_y 的平均向量即:

$$[0064] \mathbf{u}_{S_x} = \frac{1}{n_x} \sum_{j=1}^{n_x} \mathbf{X}(j), \quad \mathbf{u}_{S_y} = \frac{1}{n_y} \sum_{j=1}^{n_y} \mathbf{X}(j)$$

[0065] Σ 是样本混合协方差:

$$[0066] \Sigma = \frac{(n_x - 1)\Sigma_{S_x} + (n_y - 1)\Sigma_{S_y}}{(n_x + n_y - 2)}$$

[0067] Σ_{S_x} , Σ_{S_y} 表示 S_x 与 S_y 的估计协方差阵, n_x, n_y 指样本 S_x 及 S_y 的大小。FLDF 体现了在最佳判别轴上,两类刺激条件引起的局部脑活动模式的可分离程度;使用模式分类中的多元距离函数来度量不同刺激条件激活的局部脑活动模式间的差异,选择其它的多元距离函数来替换 FLDF 进行相关分析,其原理和步骤和使用 FLDF 相同。

[0068] 四、多元距离统计参数图假设检验步骤:

[0069] 在各局部脑区 $N(v_0)$ 活动模式在不同刺激条件下的活动相同即:不可显著分离的零假设下,使用非参数序列检验对每个体素对应的统计距离参数进行检验,得到显著性非参数序列检验值图;

[0070] 通过如下的磁共振图像数据重采样来进行非参数序列检验:数据空间模式不变,磁共振图像数据对应的时间序列进行随机重排;随机重排破坏了磁共振图像信号和试验条件的相关性,保留了完整的空间结构,从而计算不同局部脑活动模式间的多元距离统计量对应的非参数序列检验值。

[0071] 五、假设检验多重比较校正步骤,还包括:

[0072] 步骤 F. 1,选择一个错误发现率 FDR 的边界值 q ($0 < q < 1$),是所期望的最大的错误发现率 FDR;

[0073] 步骤 F. 2, 对由非参数排列检验得到的非参数序列检验 p 值图进行从小到大对值排序 : $p(1) \leq p(2) \leq \dots \leq p(m)$, 设 $p(i)$ 对应着体素统计量 $v(i)$, m 是所检验的 fMRI 数据共有的体素个数 ;

[0074] 步骤 F. 3, 假设 r 是满足不等式 $p(i) \leq \frac{i}{V c(m)} q$ 的最大 i 值, 其中 $c(m)$ 是预先设定的常量, 它的选择和体素的分布情况有关, 不同条件下有 2 种选择 : $c(m) = 1, c(m) = \sum_{i=1}^r \frac{1}{i}$;

[0075] 步骤 F. 4, 得到结果 : 拒绝空假设条件即真正激活的体素为 $v(1), v(2), \dots, v(r)$, 即是统计量大于 $v(r)$ 的体素是激活的体素, 称 $v(r)$ 为多检验确定的阈值 ;

[0076] 步骤 F. 5, 根据 FDR 校正得到的多检验阈值 $v(r)$, 对统计参数图进行阈值分割, 得到那些能显著区分不同试验条件的脑激活区。

[0077] 其二, 介绍仿真数据实验 A

[0078] A. 1 仿真数据设计

[0079] 由于真实磁共振成像 fMRI 数据的激活区位置和形状一般无法精确获知, 故我们首先设计仿真数据 (预先设定激活区位置, 形状及体素的活动情况), 并使用其对不同激活区检测方法进行性能比较。仿真数据包含两类刺激条件, 采用慢速呈现事件相关设计。每个刺激呈现 500 毫秒, 不同事件起始时间间隔在 16 ~ 20 秒间变动。每类刺激条件包含 30 个事件, 呈现顺序随机。通过刺激事件序列卷积标准血液动力学函数 (canonical Hemodynamic Response Function, cHRF) 仿真各类刺激引起的脑激活对应的 fMRI 时间序列。数据具体参数如下 : 层数 (number of slice) 为 5, 体素大小 (voxel sizes) 为 $3 \times 3 \times 3$ 立方毫米, 重复时间 (TR) 为 2000 毫秒, 矩阵 (matrix) 为 64×64 。

[0080] 五个形状和大小均不相同的激活区被仿真 (图 6a), 包含体素个数分别为 10, 30, 90, 180, 270。在激活区内, 每类刺激引起的空间效应服从高斯白噪声分布。在激活区外, 刺激不引起效应。以高斯白噪声仿真各类刺激引起的激活模式可使效应信号的能量均匀散布于不同的空间频段, 从而使仿真激活脑区的局部平均信号 (低频成分) 和精细空间结构 (高频成分) 中均含有刺激条件相关信息, 更加接近真实 fMRI 数据。然后, 刺激引起的脑活动信号被加到时空噪声背景上。为了模拟真实 fMRI 数据中噪声的局部空间相关特性, 噪声背景按如下方式产生 : 首先仿真得到时空高斯白噪声, 然后使用半高宽 (Full Width at Half Maximum, FWHM) 为 3.5 毫米的高斯核对产生的高斯白噪声进行平滑, 得到具有局部空间相关特性的噪声背景。为检验激活区提取算法在不同对比噪声比 (Contrast to Noise Ratio, CNR) 下的性能, 五组具有不同对比噪声比的数据被产生, 对比噪声比分别取值为 0.2, 0.4, 0.6, 0.8, 1.0。对比噪声比定义为各个脑激活区内空间平均信号幅度的最大绝对值同背景噪声标准差之比。

[0081] A. 2 仿真数据分析

[0082] 采用如下三种方式对仿真数据进行分析 :

[0083] (1) 原始数据 -GLM 分析 ;

[0084] (2) 高斯核平滑数据 -GLM 分析 ;

[0085] (3) LMDM 分析。

[0086] 高斯核平滑数据 -GLM 分析中, 半高宽为 6 毫米和 9 毫米的高斯核滤波器被分别用

来平滑数据。之所以采用这两个高斯核滤波器是因为在 fMRI 数据分析和文献报道中它们被普遍采用。LMDM 分析采用的局部区域大小 K 分别为 10 个和 30 个体素，这保证了 LMDM 分析使用到的局部体素个数分别近似等于高斯核平滑数据 -GLM 分析中两个高斯核所能涵盖的局部体素个数，使我们可以客观地比较 LMDM 分析和高斯核平滑数据 -GLM 分析对局部空间信号集成的效率。

[0087] 操作曲线 ROC(Receiver Operating Characteristic, ROC) 曲线被用来定量对比不同分析方法的性能。对特定分析方法计算得到的统计参数图，设定不同的阈值进行分割，便得到不同的激活图。ROC 曲线刻画了当阈值从参数图中最小值到最大值变动时，分析方法检测到的激活图对应的真激活率 (true activation rate) 和假激活率 (false activation rate) 随阈值变动的情况。真激活率定义为正确识别为激活体素的体素数比上真实激活体素数；而误识别为激活体素的体素数比上真实非激活体素数定义为假激活率。根据信号检测理论，ROC 曲线体现了信号检测方法敏感性 (sensitivity) 和特异性 (specificity) 间的制约关系，可用于不同信号检测方法的定量对比。这里，真激活率对应着激活区提取算法的敏感性，而假激活率则对应着激活区提取算法的特异性。ROC 曲线包围面积 (ROC area) 可作为一个综合指标来衡量信号检测方法在多大程度上能同时获得高敏感性和特异性。就是说，这个指标可以衡量由特定方法计算得到的统计参数可以在多大程度上正确地把仿真激活体素同背景噪声区分开来。

[0088] 图 2，本发明中提取局部一致脑区示意图，图 2a 标出了在进行一次区域增长时使用的种子体素，图 2b 显示了该次区域增长得到的局部一致区域。

[0089] 图 3，本发明中用以度量不同激活模式间统计距离的模式分类概念示意图，(此处以线性判别分析为例)。图 3 中，空心圆圈表示在刺激条件 X 下测得的样本点，实心圆圈表示在刺激条件 Y 下测得的样本点。为便于在二维图上显示，这里假设局部空间模式由两个体素 (X 轴和 Y 轴)。线性判别分析通过把两类样本点投影到一个最优判别轴上，使两类样本在该轴上具有最好的可分离性，并可通过 Fisher 线性判别函数来度量两类样本在该轴上的分离性。

[0090] 图 4 实例中本发明算法和传统激活区检测算法的对比，曲线均由 30 次仿真实验平均得到，给出了三种分析方法对应的 ROC 曲线包围面积随着对比噪声比变动的情况。可以清楚地观察到，在任一对比噪声比下，数据平滑均降低了 GLM 分析的检测性能。数据平滑越严重 (平滑窗越大)，性能降低越多，这是因为高斯核平滑数据时，很多精细刻度上的高频信息被滤除了。LMDM 分析的性能在各种情况下都要优于 GLM 分析，因为通过直接使用多个体素构成的空间模式，LMDM 完整地描述和利用了包含在局部区域内的脑活动信息。

[0091] 图 5，实例中本发明算法和传统激活区检测算法的对比。显示了三种分析方法在一例对比噪声比为 0.6 的仿真数据下的 ROC 曲线。可以看到 LMDM 分析对应的 ROC 曲线在所有情形下均处在最左上方，高于所有 GLM 分析对应的 ROC 曲线。原始数据 -GLM 分析得到的 ROC 曲线位于 LMDM 分析和高斯核平滑数据 -GLM 分析得到的 ROC 曲线之间，表明原始数据 -GLM 分析的性能要比 LMDM 分析差，但优于高斯核平滑数据 -GLM 分析。仿真数据在所有低于 Nyquist 频率 (由体素大小决定) 的频段具有等同的能量级别，高斯核平滑数据 -GLM 分析滤除了高频信息，消减了信号能量，从而无法利用这些高频信息，降低了检测性能，故在三者中性能最差；原始数据 -GLM 分析没有进行滤波，保留了这部分高频信息，但由于只

利用了单个体素的信息,没有利用局部脑区内的信息,所以性能要低于 LMDM 分析;而 LMDM 分析正是有效利用了分布在局部空间模式内的所有频段的信息来判别不同的脑活动状态,故其性能最优。

[0092] 为了直观地理解图 4,图 5 中给出的定量结果,图 6(a)、(b)、(c)、(d) 实例中本发明算法和传统激活区检测算法的对比。给出了三种方法针对图 5 中在一例对比噪声为 0.6 的所用仿真数据上得到的未经阈值分割的参数图(均取自仿真数据中间层)。图 6(a) 仿真激活区轮廓图。图 6(b) 原始数据, GLM 分析结果。图 6(c) 数据经半高宽为 9 毫米的高斯核滤波后, GLM 分析结果。图 6(d) 局部区域大小设为 30 个体素, LMDM 分析结果。图中,真实脑激活区边界已由轮廓线标出。可以看到原始数据 -GLM 分析得到的绝对 t 值图没有很好地检测到仿真激活区内一些活动较弱的体素,同时参数图的整体信噪比也较低。当采用半高宽为 9 毫米的高斯核平滑数据后, GLM 分析提取到的参数图信噪比有所提高,但在真实仿真激活区外也出现了一些明显的随机伪激活体素。相对于两种 GLM 分析, LMDM 分析得到的多元距离统计参数图的信噪比更高,且精确地检测到了真实脑激活区,同时在真实激活区外也没出现伪激活体素。这些未分割的参数图定性地说明了 LMDM 分析通过使用多元统计距离来集成局部脑区内的完整信息,可以更精确地检测到能区分不同脑活动状态的局部脑区。

[0093] 其三,是介绍真实数据实验 B

[0094] B. 1 真实 fMRI 数据获取

[0095] 6 个被试参与了实验,男女各半。实验采用具有两类刺激条件的组块化(block)设计,一类刺激条件是人脸图片,另一类刺激条件为房子图片。在刺激条件组块内呈现中心标有“十”号的刺激(人脸或房子)图片;在基线组块(baseline)内,呈现内容仅是“十”号的图片。要求被试在观看图片时把注视点置于“十”号上。每个刺激组块持续 30 秒,期间共呈现 20 个同类图片,每个图片呈现 500 毫秒,在刺激图片间隔呈现“十”号图片。基线组块长度约为 10 秒,允许基线组块长度可变,以更全面地采集人脑对刺激的血液动力学反应,从而有利于脑激活区的检测。采用 T2* 加权梯度回波平面成像(Echo-Planar Imaging,EPI)序列获取 BOLDfMRI 数据,详细参数如下:层厚(slice thickness)为 4 毫米,视窗(FOV)为 240×240 平方毫米,重复时间(TR)为 2000 毫秒,矩阵(matrix)为 64×64。

[0096] 采用 SPM2(<http://www.fil.ion.ucl.ac.uk/spm/>) 对数据进行预处理,包括:时间对齐,空间重排,去基线漂移,去除非脑部体素。另外 LMDM 分析时, fMRI 时间序列被向前移动 4 秒以抵消血液动力学反应延迟;组块转换处对应的数据被丢弃,仅保留血液动力学稳定时测得的数据。

[0097] B. 2 真实 fMRI 数据分析

[0098] 采用与分析仿真数据相同的方式对真实数据进行分析:(1) 原始数据 -GLM 分析,(2) 高斯核平滑数据 -GLM 分析,(3) LMDM 分析。通过仿真实验可看到,局部区域大小设为 30 个体素时, LMDM 分析检测性能最好,其用到的局部体素个数近似等于半高宽为 9 毫米的高斯核涵盖的局部体素个数。故下面仅使用这两个参数对数据进行分析。结果如图 7 所示,三种分析方法均可检测到对应着人脸图片刺激和房子图片刺激差异的主要脑区(FFA 与 PPA)。图中,所有显著性 p 值图均是通过 1000 次独立排列重采样获得。不同的是高斯核平滑数据 -GLM 分析和 LMDM 分析的显著性 p 值图均已使用 FDR 校正,设定平均 FDR 不超过 q

= 0.05；而原始数据 -GLM 分析给出的显著性 p 值图没有进行 FDR 校正，若使用和其它两者同样的校正，将没有任何体素能达到显著激活。GLM 分析得到的激活图表示该位置人脸图片刺激引起的体素活动显著强于房子图片刺激引起的体素活动，或者是房子图片刺激引起的该位置体素活动显著强于人脸图片刺激引起的活动。LMDM 分析得到的激活图说明对应体素与其周围局部区域内体素活动联合构成的空间模式在两类图片刺激下具有很好的可分性，就是说两类刺激条件引起的局部脑区活动呈现的模式是显著不同的。如图 7a、图 7b、图 7c 实例中本发明算法和传统激活区检测算法的对比所示，原始数据 -GLM 分析可以大致得到脑激活区位置，但是存在着严重的椒盐噪声。在 fMRI 数据中单个体素信号的对比噪声比较低，图 7a 原始数据 -GLM 分析结果仅利用单个体素活动信息来决定体素的状态，而没利用局部脑区内含有的信息，因而得到的参数图信噪比较低。采用半宽高为 9 毫米的高斯核对数据进行滤波后，使用 GLM 分析结果得到的激活图较原始数据 -GLM 分析得到的激活图干净和集中，但只有那些对两类刺激反应差异较强的核心激活体素被检测到（图 7b）。然而，使用局部区域大小为 30 个体素的 LMDM 分析结果比高斯核平滑数据 -GLM 分析检测到更多的激活体素，这些体素分布在高斯核平滑数据 -GLM 分析得到的中心激活区四周。这表明除了高斯核平滑数据 -GLM 分析检测到的核心激活区外，在其四周依然存在着很多包含精细信息，对实验刺激反应较微弱的体素（图 7c），当对这些体素进行联合分析时，同样能很好地区分不同实验刺激条件。GLM 分析无法检测到这些边缘的，含有微弱信息的体素，主要源于两个原因：一是 GLM 分析本质上是一元分析，它在进行激活区域检测时，分离地对每个体素进行处理，无法综合利用包含在局部区域内的空间模式信息；二是高斯核滤波尽管一方面会加强 GLM 分析的统计效力，提高核心激活区内体素信号的信噪比，但同时不合理地滤除了包含在局部区域内的大部分精细脑活动信息。通过分析所有被试数据，我们进一步发现，LMDM 分析检测到的激活区域相对于高斯核平滑数据 -GLM 分析检测到的激活区域更为延展，呈现分布式形态，这和前人的研究结果一致。

[0099] 图 7 中 L 表示该侧对应人脑轴向位的左侧，R 表示该侧对应人脑轴向位的右侧。Face 表示人脸图片刺激条件，house 表示房屋图片刺激条件。图中对人脸图片和房屋图片具有激活差异的体素均经过 permutation 检验。b 和 c 使用 FDR 校正，保证所有激活体素的平均 FDR 小于 0.05。

[0100] 为定量对比不同分析方法给出的参数图在被试间的重复性（相似性或一致性），被试的参数图首先被配准到标准 MNI 模板上，然后计算每对被试参数图的皮尔逊相关系数 (r) 用以度量它们之间的相似性。图 8 实例中本发明算法和传统激活区检测算法的对比，图中给出了三种分析方法对应的参数图在不同被试间的一致性，横轴上 GLM, sm-GLM, LMDM 分别代表原始数据 -GLM 分析，高斯核平滑数据 -GLM 分析，LMDM 分析；纵轴刻度代表皮尔逊相关系数。以均值 \pm 标准平均误差 ($\bar{r} \pm SEM$) 的形式对每种分析方法下 15 对 (C_6^2) 参数图的皮尔逊相关系数进行了统计。三种分析结果中，原始数据 -GLM 分析得到的统计参数图在被试间的相似性 ($r = 0.25 \pm 0.04$) 最差，这是由参数图信噪比太低造成的。LMDM 分析的参数图在被试间的相似性 ($r = 0.49 \pm 0.03$) 比高斯核平滑数据 -GLM 分析的参数图 ($r = 0.57 \pm 0.02$) 要差一些。这一方面是因为在精细空间刻度上，不同被试间脑激活模式本身差异较大；另外还有可能是 LMDM 分析没有采用空间平滑，对空间标准化过程中引入的误差较敏感，导致被试间参数图的一致性下降。因此本发明特别适用于单个被试的 fMRI 脑激活

区,相对于传统 fMRI 激活区提取方法,该方法可提取单个被试更精细的脑激活信息。

[0101] 上面描述是用于实现本发明的实施例,本领域的技术人员应该理解,在不脱离本发明的范围的任何修改或局部替换,均属于本发明权利要求来限定的范围。

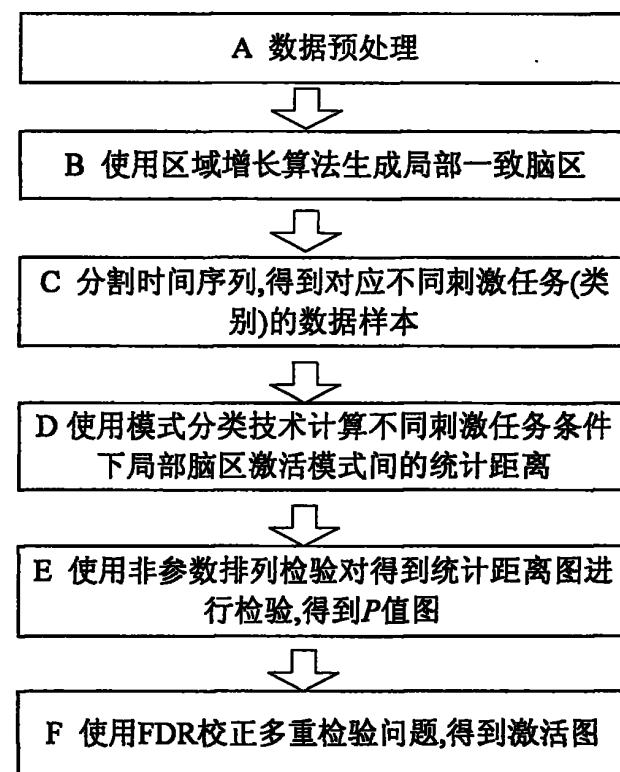


图 1

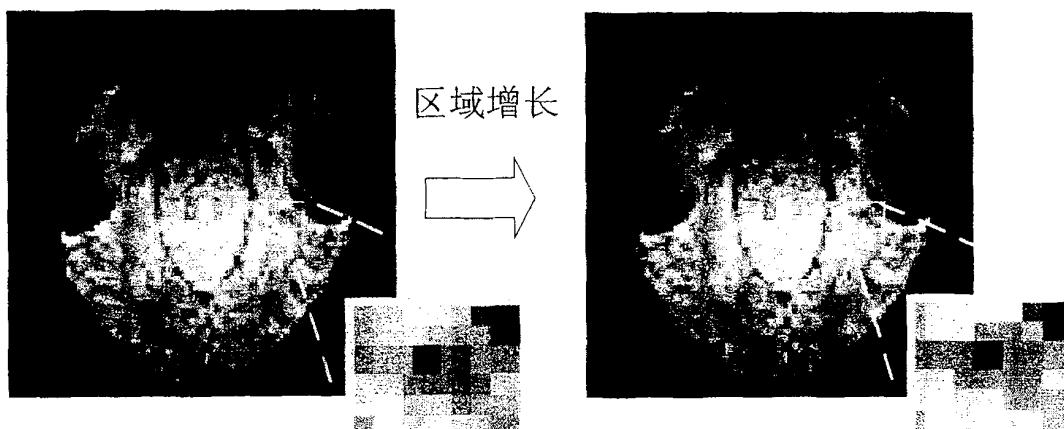


图 2a

图 2b

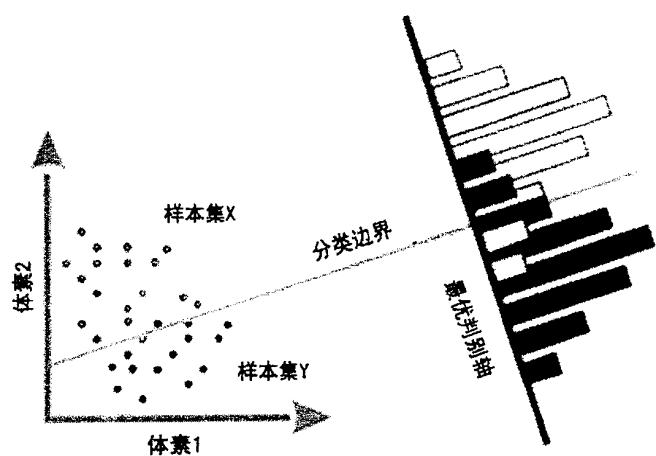


图 3

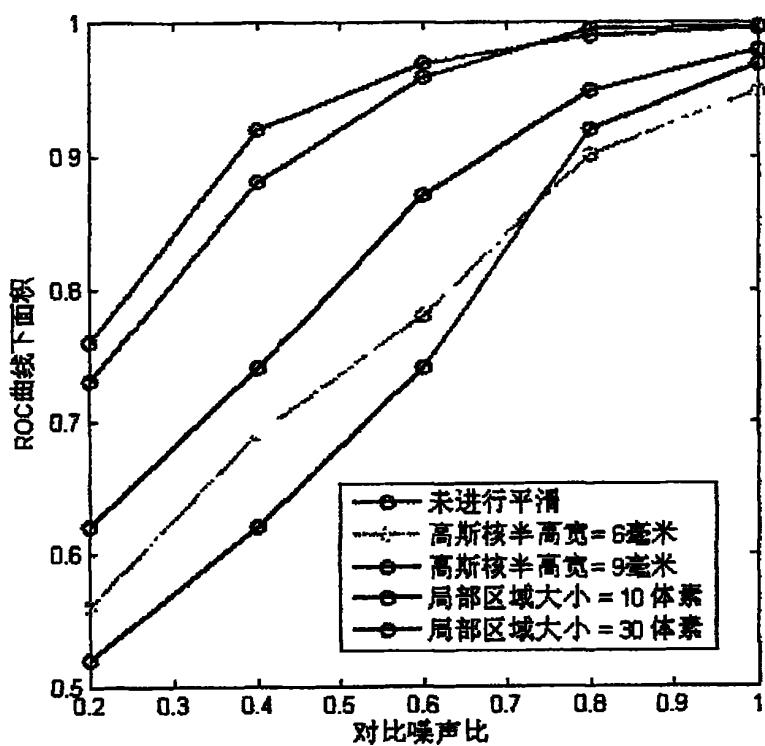


图 4

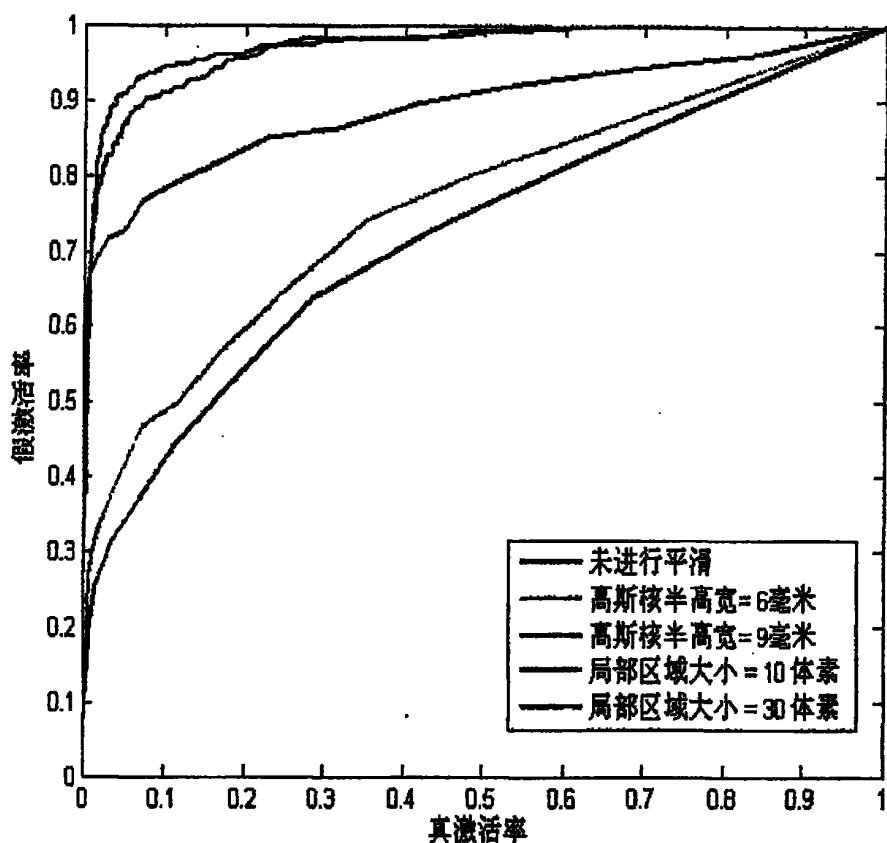


图 5

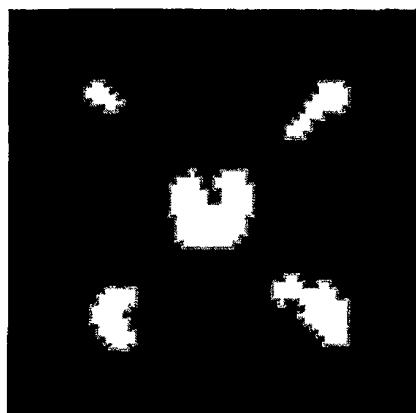


图 6 (a)

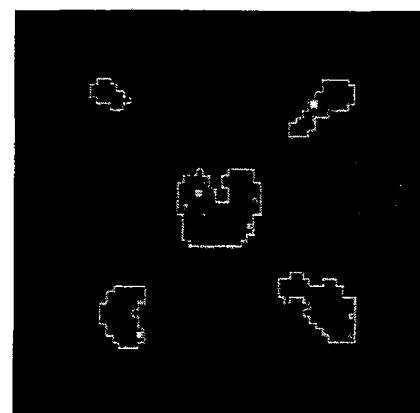


图 6 (b)

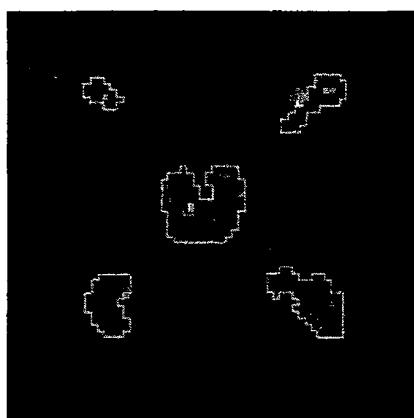


图 6(c)

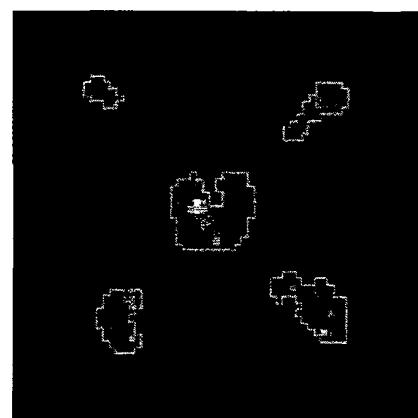


图 6(d)

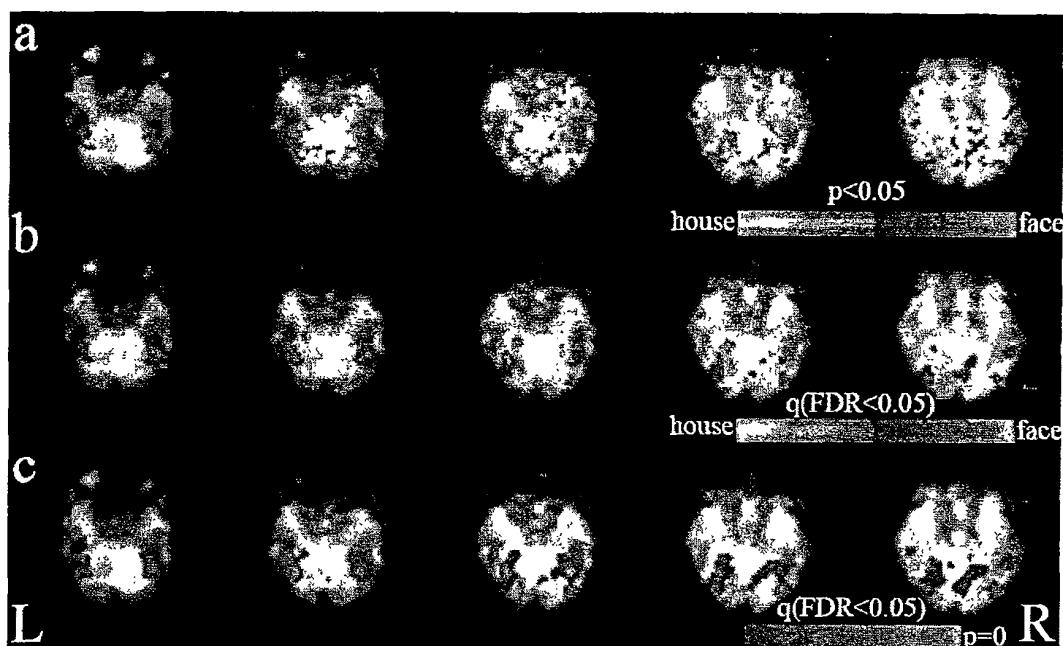


图 7

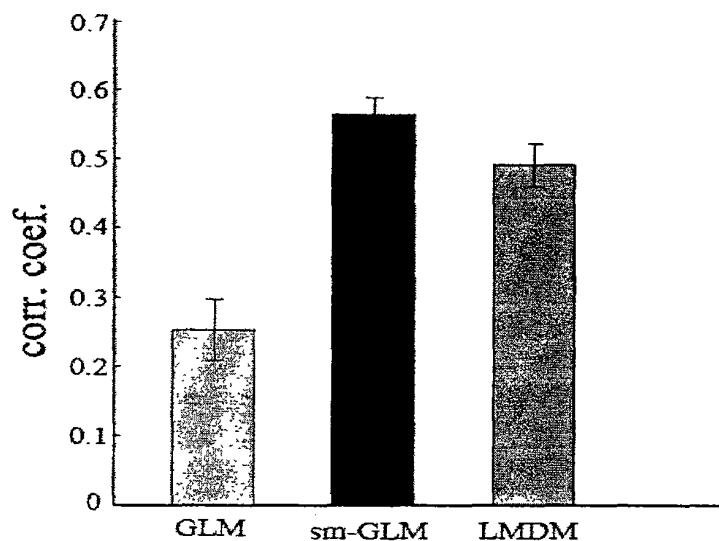


图 8