

# 视像整体特征在人类初级视皮层上的稀疏表象: 脑功能成像的证据

赵松年, 姚力\*, 金真, 熊小云, 邬霞, 邹琪, 姚国正, 蔡晓红, 刘一军

中国科学院大气物理研究所大气边界层物理和大气化学国家重点实验室, 北京 100029;  
北京师范大学信息科学与技术学院, 认知神经科学与学习国家重点实验室, 北京 100875;  
中国人民解放军北京 306 医院脑功能中心磁共振成像研究室, 北京 100101;  
国家自然科学基金委员会电子和信息系统学部, 北京 100085;  
北京交通大学计算机科学技术学院, 北京 100044;  
北京大学信息科学中心, 北京 100871;  
University of Florida McKnight Brain Institute, FL 32610, USA

\* 联系人, E-mail: [yaoli@bnu.edu.cn](mailto:yaoli@bnu.edu.cn)

2008-03-03 收稿, 2008-04-25 接受

国家自然科学基金(批准号: 60371045, 60628101)和国家高技术研究发展计划(批准号: 2006AA01Z132)资助项目

**摘要** 在视皮层表象的 fMRI 研究中设计了两类视觉刺激图像: 一类是高分辨率人脸图像和对应的线条图; 另一类是完整的几何视错觉图和对应的非完整的几何视错觉图. 这两类视觉刺激图像的特点是, 各自具有相同的大尺度特征和不同的细节, 能够保证实验结果的可比性. 在时段 fMRI 实验中, 采用第一类视觉刺激; 在事件相关的 fMRI 实验中, 采用第二类视觉刺激. 实验结果表明, 视皮层对具有相同的轮廓与形状, 而细节不同的视觉刺激表现出近似不变性. 本质上, 这是视皮层对于图像中的整体特征的同步化响应与高级皮层区的形状感知对 V1 皮层反馈作用的协同效果, 是整体特征整合的体现, 同时也是稀疏化表象策略和分布式群编码的体现.

关键词

视像  
整体特征  
fMRI 实验  
初级视皮层  
稀疏表象  
整合  
反馈  
可比性

外部世界或视觉图像在视皮层上是如何表象的? 在Marr<sup>[1]</sup>之后的 20 多年中, 这仍然是当前在系统水平上研究视觉信息处理的重要的并受到广泛关注的问题. 无论是Hubel和Wiesel<sup>[2]</sup>发现的V1 皮层简单细胞的局域、朝向和带通的选择特性, 还是V1 皮层的稀疏表象策略<sup>[3-6]</sup>或者是同步化响应的整合作用<sup>[7,8]</sup>, 本质上都与视觉如何处理图像中的大尺度特征有关, 也就是说, 视觉表象是如何以多种尺度把图像的性质表达清楚的<sup>[1]</sup>, 特别是V1 皮层是怎样处理视像中的边缘、轮廓甚至形状特征的, 是一个值得研究的重要课题.

对视网膜 外膝体 视皮层的信息传输通路, 电生理实验已证实, 视觉图像的轮廓、边缘等特征主要是由M细胞传递的; 而细节信息(如纹理特征等)则

主要是由P细胞传送的<sup>[9]</sup>. 在视皮层V1 上先出现边缘、轮廓特征, 以便首先检测外界客体的基本特征, 识别和区分它所属的类别, 随后(平均约 100 ms)出现的细节信息则用于进一步地做出类内不同个体的识别<sup>[10]</sup>, 这个特点早已在 10 多年前由Young和Yamane<sup>[11]</sup>在颞下皮层的微电极实验中所证实.

目前一些相关的研究认为, 从视网膜到视皮层, 神经元群逐渐增加了组合特征的选择性<sup>[9,12]</sup>, 初级视皮层对来自视网膜的图像采用了在空间上分布式的稀疏编码策略<sup>[13]</sup>, 为后续的信息加工提供了更加有效的表象<sup>[14]</sup>. 视觉计算模型所得的结果表明, 稀疏编码(也就是用数目最少的活性神经元来表示一幅视像)对记忆而言, 可以增加存储容量, 便于大脑实现按内容检索<sup>[15,16]</sup>. 由于高级皮层区形状感知的反馈

作用, 视皮层对外部刺激的活性降低了 [17-19], 人类视皮层细胞的感受野对朝向、带宽和局域特征有很强的选择性, 它更易于对视像中的整体特征做出响应, 因而是稀疏表象的基础 [20]。现在, 基于稀疏表象的视觉不变性理论 [21] 开始得到一些间接实验证据的支持。不过, 由计算模型得出的V1 皮层稀疏编码的结果是高度假设性的 [22], 这类研究尚未指明视觉系统是如何理解或感知外部世界自然场景的统计结构的, 稀疏编码的策略在生物视觉中又是如何实现的, 许多研究工作仍然是阐述实现稀疏编码的神经网络学习模型与算法 [23,24]。

在探讨V1 皮层对视像是否采用了稀疏表象策略时, 为了避免计算模型设定的过多假设 [3-6,22], 我们进行了初步的fMRI实验研究, 其目的是为V1 皮层的稀疏表象提供一个直接的证据。在实验中, 将V1 皮层的稀疏表象策略与同步化响应的整合作用的思路结合起来, 设计了两种视觉刺激图像, 一种是人脸灰度图像和对应的人脸线条图; 另一种是视错觉图。作为视觉刺激的人脸灰度图像, 既包含了客体的细节特征信息(纹理和明暗等细节), 又包含了客体的整体特征信息(轮廓与边沿); 而与其对应的黑白线条图只包含该客体的整体特征信息。在线性模型的假设之下, 可以用叠加原理 [25,26] 对fMRI实验结果进行分析, 也就是说, 可以将视皮层上对人脸灰度图像和线条图像形成的大量神经元群的活性模式进行对比, 灰度图像与相应的线条图像二者在视皮层上分别引起的大量神经元群的发放模式之间的差别(也就是由fMRI信号构建的、与人脸灰度图像对应的大脑活性模式与对应于人脸线条图的大脑活性模式之间的差值或者说血氧水平依赖BOLD信号间的差值)反映了这两种刺激图像之间的差别, 也就是细节信息。同样, 也可以分析视错觉图形引起的大脑V1 皮层的活性模式。在实验中使用了两种视错觉图形, 一种是包含了一个完整的视错觉图形(以下简称为“闭合的视错觉图形”), 而另一种则是包含了不完整的视错觉图形(以下简称为“非闭合的视错觉图形”)。我们想知道, 闭合的视错觉图形与非闭合的视错觉图形之间的差异能不能在由它们引起的视皮层V1 的活性模式中反映出来, 由此, 便能够根据fMRI实验结果来直接检验视皮层对视觉图像的表象。

本研究结果表明, 具有相同的轮廓与形状而细节不同的视觉刺激所引起的初级视皮层的活性模式

是相似的, 表现出近似不变性。初步分析认为, 这是视皮层对图像中整体特征的同步化响应与高级皮层区的形状感知对V1 皮层反馈作用的协同效果, 是整体特征整合的体现; 同时也是稀疏化表象策略的体现, 即用最少的活性神经元(也就是视皮层区消耗最少的能量)来表达一幅视像。

## 1 方法

采用两类不同的视觉刺激图像分别进行了 fMRI 实验。实验 1 采用了时段设计方法(blocked design), 实验 2 采用了事件相关设计方法(event-related design)。通过预实验检验了这两种方法的实际效果, 效果基本相同。利用人脸图像与线条图像作为实验刺激图像对受试者进行视觉刺激(由安放在足部的投影显示器显示), 呈现模式为随机的和固定的两种。实验参数是按检测效率、检测能力和人眼日常观看的习惯三者折中考虑并通过预实验之后确定的。

( ) 实验 1. 视觉刺激图形如图 1 所示。实验采用 Block 设计, 如图 1(c), “+”表示屏幕上呈现一个“+”字。第一个“+”的 block 长 36 s(处理数据时去除前

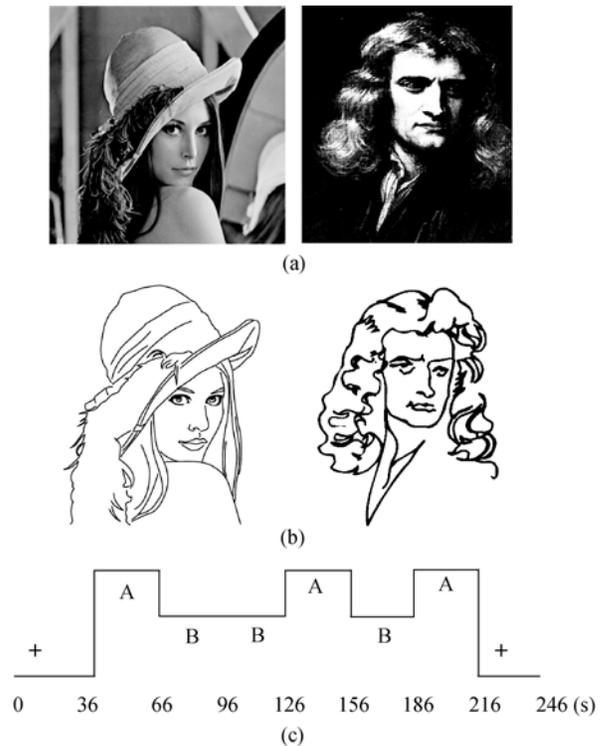


图 1 视觉刺激图像与刺激呈现模式

(a) 高分辨率人脸图。(b) 与高分辨率人脸图相对应的人脸线条图, 每种情况共 30 种不同的人脸图。每幅图像的大小及在屏幕上呈现的位置均一致。(c) Block 设计

6 s 的数据), 其他的 block 时长 30 s. 每 850 ms 出现一幅图像, 空白 150 ms. 每个 block 中出现 10 幅图像, 每幅图像重复出现 3 次. 在相邻的 A 与 B 的 block 中, 图像呈现的顺序相同.

( ) 实验 2. 视觉刺激图形如图 2 所示, 实验采用事件相关设计方法, 每幅图形随机出现, 呈现时间为 3 s, 间隔分别为随机的 3, 6, 9 s, 间隔中呈现“+”. 要求被试根据呈现的刺激图中是否出现闭合的图形通过按键做出回应.

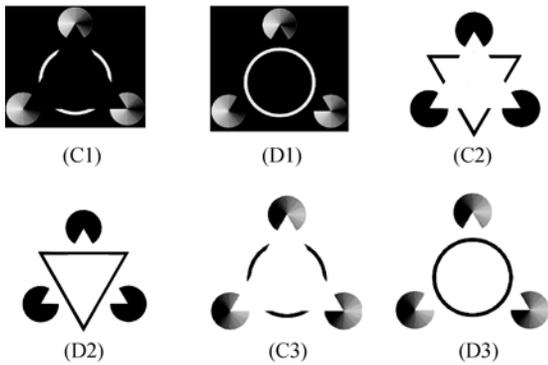


图 2 非闭合几何图形与相对应的闭合几何图形

C1~C3 是非闭合几何图形; D1~D3 是相对应的闭合几何图形, 每种情况共 8 种不同的图形

( ) 被试. 实验 1 与实验 2 中使用相同的被试, 均为视力或矫正视力正常的大学生, 平均年龄 22 岁, 男女各 5 名, 右利手, 母语为汉语. 刺激图像用液晶显示投影仪投影到受试者足部后方的屏幕上, 被试通过安装在头部线圈上能改变倾斜角度的镜面注视屏幕, 要求被试用双手拇指按键做出判断. 所有参加者均阅读了磁共振实验的有关条款和注意事项, 完成了全部同样的实验.

( ) MRI 扫描. 整个实验是在中国人民解放军北京 306 医院核磁共振室的 2.0T fMRI (Elscent/Prestige whole-body system) 上完成的. 功能像(GE-EPI): TR = 3000 ms; TE = 60 ms, Flip = 90°, FOV (field of view) = 375 mm × 210 mm; 像素矩阵 = 128 × 72; 轴状切面, 共 20 层, 切片厚度 4 mm, 无间隙. 扫描时间: 与刺激呈现时间相对应. 三维解剖像(GE T1WI): TR = 25 ms, TE = 6 ms, Flip = 28°, FOV = 220 mm × 220 mm; 像素矩阵 = 220 × 220; 矢状切面, 共 80 层, 切片厚度 2mm, 无间隙.

( ) 数据处理. 用 SPM2 软件做数据处理 (statistical parametric mapping software package). 首

先进行预处理, 重排功能像并得到一个均值图像, 把解剖像对应到功能均值像上; 然后, 把数据正态化到一个标准空间里, 并重新切片为 3 mm × 3 mm × 4 mm 大小的体元; 再使用一个 8 × 8 × 8 的 FWHM (full-width at half maximum) 的高斯核对数据进行空间平滑来提高信号的信噪比, 然后按照一般的 SPM2 步骤进行建模、估计和得到结果, 然后把结果图叠加到标准结构像上呈现 fMRI 图像结果,  $P = 0.001$ , 不校正.

## 2 实验及其结果

### 2.1 实验 1: 高分辨率人脸图与人脸线条图的比较

实验中让 10 名被试分别观看两种不同的视觉刺激图(图 3(a), (b)), (a)表示被试观看高分辨率的人脸图, 每一组刺激图由 30 种包含许多细节的高分辨率的不同人脸图片组成, 实验中使用了 5 组这样的图片序列, 它们之间的区别仅在于图片排列的顺序不同, 是按随机方式排列的(任务 A); (b)表示让同样的 10 名被试人观看 5 组人脸线条图, 每一组刺激图由 30 种不同的人脸线条图片组成, 5 组线条图以及其中各图片的呈现顺序与任务 A 是完全对应的和完全相同的, 从而保证实验结果在统计上的可比性(任务 B). 图 3(c)表示大脑的两种激活模式的差值(A-B). 整个实验是时段设计, 每张图片呈现时间为 850 ms, 图片之间的间隔为 150 ms, 各组之间间隔为 6 s, 实验开始时有 6 s 的间隔时间, 其扫描数据不参加数据处理.

### 2.2 实验 2: 无视错觉图与有视错觉图的比较

实验中让 10 名被试分别观看两种不同的视觉刺激图, 图 4(a)表示被试人观看具有完整的几何图形(如圆和三角形)的视错觉图形, 每一组刺激图由 30 种不同几何图形组成, 实验中使用了 5 组这样的图形序列, 它们之间的区别仅在于图形的排列顺序不同, 是按随机方式排列的(任务 A); 图 4(b)表示让同样的 10 名被试人观看 5 组没有完整几何图形的视错觉图, 每一组刺激图由 30 种不同视错觉几何图形组成, 它们是将(a)中所对应的各个闭合的几何图形(如三角形和圆)去掉一部分而变成非闭合的几何图形, 5 组视错觉图形以及其中各图形的呈现顺序与(a)是完全对应的和完全相同的, 从而保证实验结果在统计上的可比性(任务 B); 图 4(c)表示大脑的两种激活模式的差值(A-B). 整个实验采用事件相关方法设计, 任务 A 和 B 随机出现, 每次呈现 3 s, 间隔为随机的 3, 6, 9 s,

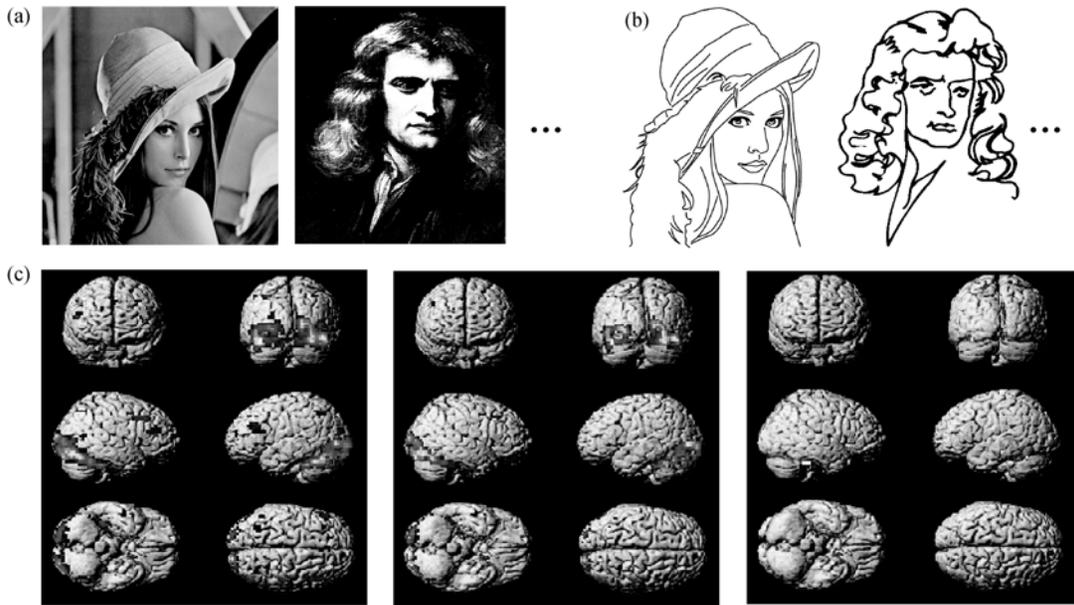


图3 视觉刺激在大脑中的不同表象

(a) 高分辨率人脸像; (b) 人脸线条图; (c) 两种不同表象模式之间的差别(与任务相关的脑的活性由 SPM2 确定,  $P < 0.001$ )

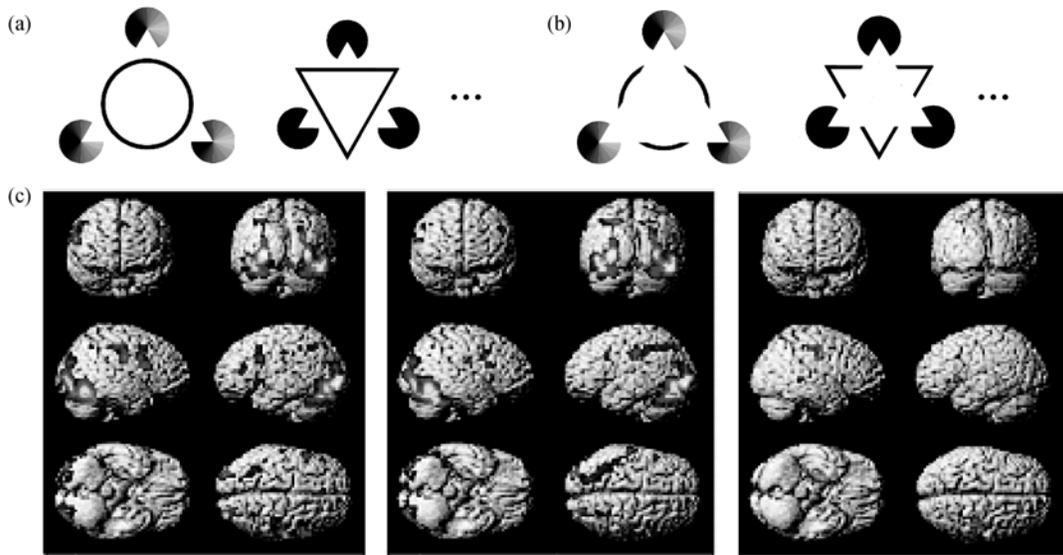


图4 视觉刺激在大脑中的不同表象

(a) 具有闭合的几何图形的视错觉图形; (b) 具有非闭合的几何图形的视错觉图形; (c) 两种不同表象模式之间的差别(与任务相关的脑的活性由 SPM2 确定,  $P < 0.001$ )

整个任务持续时间为 375 s。要求被试在观看两种图形的同时判断是否出现完整的三角形或圆, 并通过按键回应, 前 15 s 的扫描数据不参加数据处理。

在实验 1 中, 大脑活性区主要集中在视区 V1, V2 和 V3, 只是在观看高分辨率人脸图时, 颞下皮层 (inferotemporal cortex, area 37), 后枕区 (posterior parietal cortex, areas 5, 7) 以及前额叶 (prefrontal cortex,

area 9) 有少量的活性, 而在观看人脸线条图, 这些皮层区的活性几乎消失, 任务 A 和 B 在视皮层 V1 上的表象与活性模式几乎没有达到统计上有意义的差别, 这就说明高分辨率人脸图像的细节部分并没有引起 V1 皮层神经元群的强烈发放。由于在实验 2 中, 任务 A 和 B 的视觉刺激图形之间的差别比实验 1 要小, 以图 2 中的 C1 与 D1 为例, 在 C1 图中, 有一个虚假的

三角形部分地遮挡了白色的圆环;而在 D1 图中,一个完整的白色圆环在虚假的三角形的前面,图 D1 和 C1 之间的差别只是由完整的白色圆环变成 3 个白色的弧形线段.

通常认为,作为视觉刺激,这些差别会在 V1 皮层上引起相应的活性模式的显著差别,可是实验结果并非如此.在实验 2 的任务 A 中,尽管颞下皮层(area 37),后枕区(areas 5, 7)以及前额叶(area 9)也有少量的活性,不过在任务 B 中,这些皮层区的活性则非常少.其次,大脑高级皮层区的认知功能将会把非闭合的几何图形(即 3 个圆弧线段)所缺失的部分(即被遮挡的 3 段弧线)联系起来,如同有一个闭合的几何图形,这时,由高级皮层区反馈到初级视皮层的信息将会降低它的细胞群体的活性.因此,就目前 fMRI 方法所能达到的时间分辨率与空间分辨率而言,任务 A 和 B 在 V1 皮层上呈现的活性模式之间的差异尚未达到统计上有意义的程度,如果根据 fMRI 比较其结果,我们可以将任务 A 和 B 引起的视皮层活性模式看作是相似的,是对这两种视觉刺激的近似不变的表象.

除了全脑活性模式的对比外(图 3 和图 4),还对 A 组中表现出活性的 12 个水平切片与 B 组中对应的切片进行比较,其差值几乎为零(图 5),这个结果表

明,无论是实验 1 还是实验 2,与 A-B 对应的视皮层区域的活性模式并没有出现在统计上有意义的差别,这是一个令人惊讶的结果.

从以上实验结果可知,尽管被试人之间存在差异,枕叶激活的区域并不完全相同,但是被试人在观看高分辨率人脸图时,枕叶激活的范围比观看人脸线条图要略为大一些,激活的强度也要稍高一些(图 5(A1), (B1), (A1-B1)),被试人在观看具有闭合的几何图形的视错觉图时,枕叶激活的范围要比观看具有非闭合几何图形的视错觉图时稍许大一些,激活的强度也要略高一些(图 5(A2), (B2), (A2-B2)).不过,二者之间的差异(A2-B2)尚未达到统计上有意义的程度.

### 3 讨论

如何来看待我们的实验结果呢?从视觉通路的神经解剖学可知,视网膜的光敏细胞(视杆细胞和视锥细胞)的数目大约是 1.25 亿,它们汇集于约 100 万个神经节细胞,然后经由外膝体投射到视皮层 V1,这种细胞数目逐渐稀疏化的拓扑映射是从视网膜就开始的一种选择性逐渐增强的信息处理策略<sup>[9,20]</sup>,它首先对外部世界客体的整体特征做出响应,而后才是对细节的响应,其中还包括了高级皮层的认知过程对视皮层神经元群活性的反馈作用,从而降低了

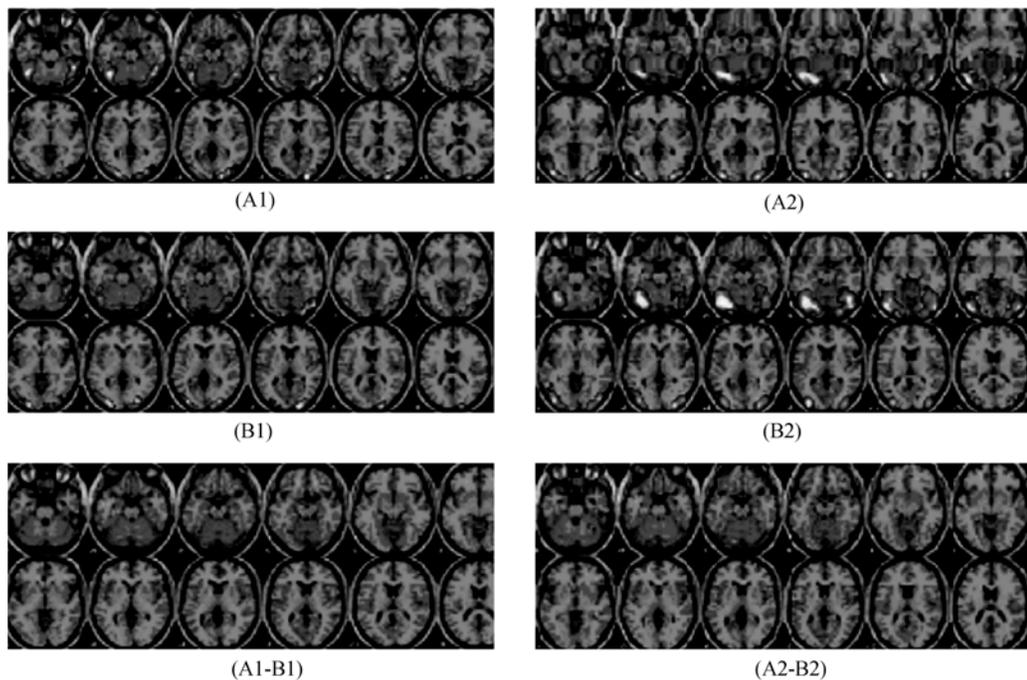


图 5 每个被试在任务 A, B 下大脑皮层各 12 个切片之间的比较

视皮层神经元群对与客体的整体特征无关的、分离的细节特征的响应强度, 这样一来, 具有相同的整体特征而细节不同的视像在视皮层上就会有几乎相同的活性模式.

但是, 在视网膜上的视像必须一一对应地投射到视皮层 V1 上, 后继的所有皮层处理都要从 V1 皮层获取信息. 显然, 在 V1 皮层上是不能损失任何信息的, 视网膜上的视像的每一个像素与 V1 皮层上的图像的每一个像素应当是一一对应的, 只有这样, 才能保证大脑皮层高保真地感知视网膜上的视像. 神经生理学已表明, 视网膜图像拓扑投射到视皮层上, 每一个对应的神经元都会被激励而处于活性状态, V1 皮层整体的活性模式便高保真地复制了视网膜的视像. BOLD-fMRI 信号可以间接地反映在视觉刺激下, 大脑皮层大量神经元群的活性模式, 不同的视觉刺激图像一定会产生不同的皮层神经元群的活性模式, 这和他的知觉一致, 而与本研究的 fMRI 实验结果不一致, 在高分辨率人脸图像与线条图两种情形下, 为什么会得出几乎相同的结果呢?

通常认为, 在 fMRI 实验中出现零结果(null effect)是由于实验中不可控制的因素造成的, 如果用 A 表示视觉刺激图像中大尺度特征 a 产生的 fMRI 信号, 用 B 表示视觉刺激图像细节 b 产生的 fMRI 信号, 那

么(a+b)与 a 本应产生不同的实验结果(A+B)与 A, 但由于不可控的因素 c 的影响(以乘性作用的方式出现; 而加性作用可以滤除, 不予考虑), 使得真实的实验结果变成由(a+b)×c 与 a×c 分别产生, 这时不能排除这两种结果(A+B)C 与 AC 的差别[(A+B)C-AC]很小, 甚至接近于 0, 出现这种情况意味着 BC→0; 但是, 在本文的实验中, 视皮层上与(a+b)×c 和 a×c 对应的不同皮层区的 fMRI 信号是稳定而清楚的, 也就是说 AC 不为 0. 为降低不可控因素 c 的影响, 要求具有正常视力的被试在观看两种图形的同时做出判断并通过按键回应, 只有判断正确的实验结果才进入数据处理. 当然, 一个不敏感的设计也可能出现零结果, 目前尚无简单可行的方法来区分这两种原因中的哪一种原因导致了实验的零结果. 此外, 如果利用小波方法将细节信息从整个图像中分离出去之后, 用所得的轮廓图与细节图分别刺激视觉所得的 fMRI 结果的合成能否等同于包含细节的原图的 fMRI 效果, 也并不清楚, 因此, 在 fMRI 实验中尽量避免出现零结果. 既然本研究的意图是进行对比, 对于特定的零结果进行分析就是必要的, 从中也许能得出有意义的启示; 实际上, 比较不同受试者之间对应的各皮层区的 BOLD-fMRI 信号, 差异明显可见(图 6 与图 7). 而且, 在本研究中, 除了视皮层之外, 高分辨率人脸图

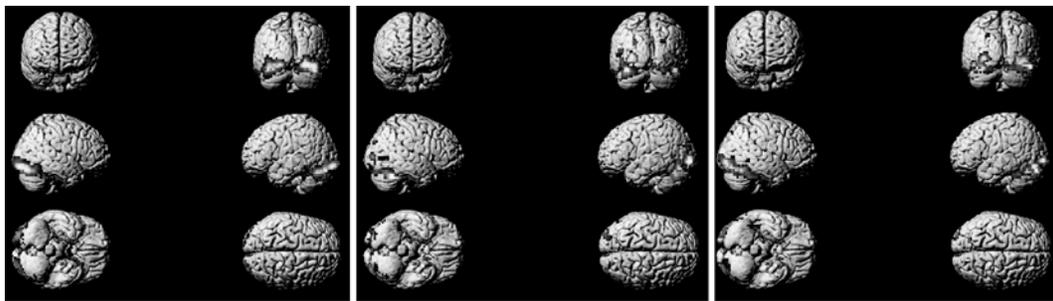


图 6 高分辨率人脸像在 3 位被试大脑中的 BOLD-fMRI 信号

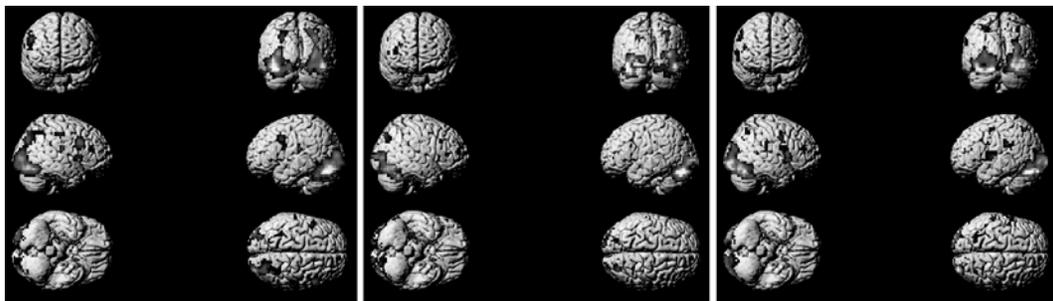


图 7 闭合的几何视错觉图形在 3 位被试大脑中的 BOLD-fMRI 信号

像在所有受试者大脑中的 BOLD-fMRI 信号强过线条图; 同样, 非闭合的几何视错觉图形在所有受试者大脑中的 BOLD-fMRI 信号也明显地强过闭合的几何视错觉图形, 这反映了视觉对复杂图像的信息处理更需要其他相关皮层的参与, 这至少说明, 人脸图像与几何视错觉图形对于受试的大学生而言, 是熟悉的图形, 加之对图形辨认实验还需通过手持按键加以确认, 这些单个的实验是合理的; 至于两类实验以相减的方式进行对比, 主要是简单直观, 二者的差值没有达到在统计上有明显区别的程度, 即在 fMRI 中所谓的“零结果”, 其实, 在当前 fMRI 的词汇判断实验中也经常遇到“零结果”, 如果实验是可重复的, 排除了实验方法的缺陷, 分析“零结果”便是有意义的. 我们认为, 得出本实验结果的可能原因有以下 3 方面:

( ) 从神经生物学的观点和 fMRI 的特性来看, 视觉刺激图像中的细节引起的 V1 皮层细胞的发放远比轮廓, 边缘引起的神经元发放弱得多, 而 BOLD-fMRI 的信噪比和它的空间分辨率都比较低, 还不足以使微弱的活性模式在 fMRI 信号中显示出来, 在 fMRI 的算法中包含了卷积运算, 本质上是滤波作用, 平滑了细节, 突出了轮廓与边缘特征<sup>[25,26]</sup>. 因此, 由 BOLD-fMRI 信号得出的 V1 皮层上的大脑活性模式体现的只是视觉刺激图像的轮廓与边缘等较大尺度的特征, 就目前 fMRI 方法所能达到的时间分辨率与空间分辨率而言, 细节信号引起的 BOLD-fMRI 信号太弱而不能在大脑活性模式中呈现出来, 而高分辨率的人脸图像与其对应的线条图具有相同的轮廓与边缘特征, 自然, 它们分别诱导的 BOLD-fMRI 信号和大脑活性模式基本上是相同的, 这是对上述两种视觉刺激的不变表象的一种解释.

( ) 从稀疏表象的观点和视觉的计算理论来看, 外部世界的客体或者自然界的图像均具有几乎相同的统计特征, 也就是说, 有几乎相同的三阶(偏斜度, skewness)与四阶(陡峭度, kurtosis)统计矩<sup>[3,4]</sup>, 因而可以用一组数目最少的基函数表达(与图像压缩类似)千变万化的外部世界, 其优点是既能节省能量又能适应环境的改变, 神经元的活性模式等价于图像的直方图, 由于细节刺激引发的神经元是低活性的, 占有的能量仅仅是总能量中很小的份额, 细节图只需要少量的低活性神经元即可以表达, 这就使得它们具有相似的直方图, 它所对应的 BOLD-fMRI 信号自然也就是相似的.

( ) 从同步化整合与高级皮层反馈的观点来看, 视觉刺激图像的轮廓与边缘等较大尺度的特征容易被同步振荡所整合而加强(Hebb 学习律), 分离的细节信号则不然, 此外, 高级皮层对视皮层神经元群活性的反馈作用, 降低了视皮层神经元群对细节特征的响应强度, 这样一来, 具有相同的整体特征而细节不同的视像在视皮层上就会有几乎相同的活性模式.

我们知道, 小波变换是当前最适合对视像进行多尺度描述的数学方法<sup>[27]</sup>, 以 Lenna 图为例, 可以分解成大尺度基本特征 LL, 水平方向细节 LH, 垂直方向细节 HL 和对角线方向细节 HH, 从能量的分配来看, LL 约占总能量的 98%, 而所有细节(LH + HL + HH)的能量总和仅占总能量的约 2%; LL 中像素的最大值约为 3640, 平均值为 1980; 细节中像素的最大值出现在水平方向细节 LH 中, 约为 94, 平均值为 0.05, 细节的小波系数几乎有 90% 为 0, 由此可知, 人脸线条图与其高分辨率图比较, 对应的 BOLD-fMRI 信号的差异也不会大于 2%, 图像的细节能够激活的神经元数目要比大尺度基本特征激活的神经元数目低约 4 个量级, 显然, 细节信息用少量低活性的神经元足可以表达, 这也是主张 V1 皮层对视像采取稀疏表象策略的理由之一, 作为稀疏表象理论的倡导者之一的 Olshausen<sup>[28]</sup>甚至认为 V1 皮层约有 60% 的神经元处于静息状态.

本研究所关注的问题是 V1 皮层的神经表象, 所得结果初步显示, 视皮层的表象是稀疏化的、分布式的群编码, 因而对于视像的整体特征的表象具有近似的不变性.

当前已有一些研究指出, 由于被试的文化背景<sup>[29]</sup>、受教育程度、种族差别<sup>[30]</sup>等因素的影响, 同一个视觉刺激图像在不同被试的大脑皮层上的 fMRI 也并不完全一致<sup>[31]</sup>, 其相似程度大约在 30% 左右<sup>[32]</sup>. 因此, 对待 fMRI 实验结果要持慎重的态度, BOLD-fMRI 信号只是大脑皮层大量神经元群的活性模式的一种间接反映, 是一种平均的集合的效果.

此外, Scott 等人<sup>[17~19]</sup>提出由于高级皮层区对形状感知的反馈作用, 视皮层对外部刺激的活性降低了, 我们在实验 2 中, 采用了视错觉图(图 2), 现在, 以其中的 C2 与 D2 的三角形视错觉图为例进行分析. 大脑高级皮层区分别感知了一个白色的虚假的三角形平面和一个真实的黑色三角形边框, 这两者对初级视皮层神经元群活性的反馈作用是一样的吗? 就

实验结果而言,它们是几乎是一样的,既然对比较实验中出现“零效果”存有疑问,那么,通过何种可行的实验才能真正判断这个问题,则是一个值得认真仔细探讨的问题。

V1 视皮层在视觉信息处理中具有极为重要的作用,但是,从Hubel和Wiesel<sup>[2]</sup>在 20 世纪 60 年代开创性的研究工作到 90 年代Field和Olshausen提出V1 皮层稀疏化编码猜测<sup>[3]</sup>的 30 多年中,我们在系统水平

上对视皮层的信息处理机制的了解还是相当少的<sup>[28,33]</sup>。即使在数学上,满足一定精度要求时,一幅图像最少需用多少个特征基函数才能够精确地重建,特征基函数在视皮层上与何种细胞相对应,仍然是一个没有解决的问题。尽管如此,当前沿着这条途径已取得的许多成果说明,神经网络计算模型与fMRI实验方法相结合是探索视觉系统信息处理的有效方法<sup>[34-36]</sup>,只是它需要不断地改进与完善。

致谢 感谢 Karl Friston 提出的修改建议。感谢中国科学院图书馆贾苹、姚丹丹和范超英在文献方面给予的帮助。

## 参考文献

- Marr D. Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. San Francisco: W. H. Freeman, 1982
- Hubel D H, Wiesel T N. Receptive fields and functional architecture of monkeys striate cortex. *J Physical*, 1968, 195: 215—243
- Olshausen B A, Field D J. Emergence of simple cell receptive field properties by learning a sparse code for natural images. *Nature*, 1996, 381: 607—609[[doi](#)]
- Simoncelli E P, Olshausen B O. Natural image statistics and neural representation. *Annu Rev Neurosci*, 2001, 24: 193—216
- Olshausen B A. Principles of image representation in visual cortex. In: Chalupa L M, Werner J S, eds. *Visual Neurosciences*. Cambridge, Massachusetts: MIT Press, 2004. 1603—1615
- Foldiak P. Sparse coding in the primate cortex. In: Arbib M A, ed. *The Handbook of Brain Theory and Neural Networks*. Cambridge, Massachusetts: MIT Press, 2002. 895—989
- Engel A K, Konig P, Kreiter A K, et al. Temporal coding in the visual cortex: New vistas on integration in the nervous systems. *Trends Neurosci*, 1992, 15(6): 218—226[[doi](#)]
- Gray C M, Konig P, Engel A K, et al. Oscillatory responses in cat visual cortex inter-columnar synchronization which reflect global stimulus properties. *Nature*, 1989, 338: 334—337[[doi](#)]
- Nicholls J G, Martin A R, Wallace B G, et al. *From Neuron to Brain*. 4th ed. Massachusetts: Sinauer Associates, Inc, 2001
- Sugase Y, Yamane S, Ueno S, et al. Global and fine information coded by signal neurons in the temporal visual cortex. *Nature*, 1999, 400: 869—873[[doi](#)]
- Young M P, Yamane S. Sparse population coding of faces in the inferotemporal cortex. *Science*, 1992, 256: 1327—1331[[doi](#)]
- Brincat S L, Connor C E. Underling principles of visual shape selectivity in posterior inferotemporal cortex. *Nat Neurosci*, 2004, 7: 880—886[[doi](#)]
- Olshausen B A, Field D J. Sparse coding of sensory inputs. *Curr Opin Neurobiol*, 2004, 14: 481—487[[doi](#)]
- Reddy L, Kanwisher N. Coding of visual objects in the ventral stream. *Curr Opin Neurobiol*, 2006, 16: 1—7
- Riesenhuber M, Poggio T. Are cortical model really bound by the “Binding Problem”? *Neuron*, 1999, 24: 87—93[[doi](#)]
- Riesenhuber M, Poggio T. How the visual cortex recognizes objects: Tale of the standard model. In: Chalupa L M, Werner J S, eds. *Visual Neurosciences*. Cambridge, Massachusetts: MIT Press, 2004. 1640—1653
- Vinje W E, Gallant J L. Sparse coding and decorrelation in primary visual cortex during natural vision. *Science*, 2000, 287: 1273—1276[[doi](#)]
- Vinje W E, Gallant J L. Natural stimulation of the non-classical receptive field increases information transmission efficiency in V1. *J Neurosci*, 2002, 22: 2904—2915
- Scott M, Daniel K, Olshausen B A, et al. Shape perception reduces activity in human primary visual cortex. *Proc Natl Acad Sci USA*, 2002, 99(23): 15164—15169[[doi](#)]
- Conner C. Friends and grandmothers. *Nature*, 2005, 433: 1036—1037[[doi](#)]
- Quiroga R Q, Reddy L, Kreiman G, et al. Invariant visual representation by single neurons in the human brain. *Nature*, 2005, 435:

1102—1107[doi]

22 Hyvarinen A, Hoyer P O. A two-layer sparse coding model learn simple and complex cell receptive fields and topography from natural images. *Vision Res*, 2002, 41(18): 2413—2423[doi]

23 Palmer S E. *Vision Science*. Cambridge, Massachusetts: MIT Press, 1999. 186—193

24 van Hateren J H, Ruderman D L. Independent component analysis of natural image sequences yields spatio-temporal filters similar to simple cells in primary visual cortex. *Proc R Soc Lond B*, 1998, 2: 2315—2320[doi]

25 Kiebel S, Holmes A. The general linear model. In: Frackowiak R S J, Friston K J, Frith C D, et al, eds. *Human Brain Function*. 2nd ed. Amsterdam: Elsevier Pte Ltd, 2004. 719—754

26 Friston K J. Experimental design and statistical parametric mapping. In: Frackowiak R S J, Friston K J, Frith C D, et al, eds. *Human Brain Function*. 2nd ed. Amsterdam: Elsevier Pte Ltd, 2004. 597—630

27 Mallat S. Wavelets for vision. *Proceedings of the IEEE*, 1996, 84: 604—614 [doi]

28 Olshausen B A, Field D J. What is the other 85% of V1 doing. *Neur Comp*, 2005, 17: 1665—1699[doi]

29 Chua H F, Boland J E, Nisbett R E. Cultural variation in eye movements during scene perception. *Proc Natl Acad Sci USA*, 2005, 102(35): 12629—12633[doi]

30 Tang Y Y, Zhang W T, Chen K W, et al. Arithmetic processing in the brain shaped by cultures. *Proc Natl Acad Sci USA*, 2006, 103: 10775—10780[doi]

31 Kanwisher N. The ventral visual object pathway in humans: Evidence from fMRI. In: Chalupa L M, Werner J S, eds. *Visual Neurosciences*. Cambridge, Massachusetts: MIT Press, 2004. 1179—1189

32 Heeger D, Ress D. What does fMRI tell us about neuronal activity? *Nat Neurosci*, 2002, 3: 142—151

33 Caradini M, Demb J B, Mante V, et al. Do we know what the early visual system does? *J Neurosci*, 2003, 25(46): 10577—10579[doi]

34 Herz A V M, Gollisch T, Machens C K, et al. Modeling single-neuron dynamics and computations: A balance of detail and abstraction. *Science*, 2006, 314: 80—85 [doi]

35 O'Reilly R. Biologically based computational models of high-level cognition. *Science*, 2006, 314: 91—94[doi]

36 Perrinet L U. On efficient sparse spike coding schemes for learning natural scenes in the primary visual cortex. *BMC Neurosci*, 2007, 8(Suppl 2): P206[doi]

## 《中国科学 D 辑: 地球科学》 第 38 卷 第 6 期 目次

东昆仑造山带中带的锆石 U-Pb 定年与构造演化启示	陈能松 孙 敏 王勤燕 张克信 万渝生 陈海红
中国东部陆相断陷盆地原油中烷基酚形成与分布控制因素分析	周树青 黄海平
青藏高原及邻区的 Rayleigh 面波的方位各向异性	苏 伟 王椿镛 黄忠贤
2007 年云南宁洱 $M_s 6.4$ 地震震源过程	张 勇 许力生 陈运泰 冯万鹏 杜海林
利用 C, N 稳定同位素分析法鉴别家猪与野猪的初步尝试	胡耀武 栾丰实 王守功 王昌燧 Michael P. Richards
中国表土花粉与建群植物地理分布的气候指示性对比	郑 卓 黄康有
许清海 吕厚远 Rachid Cheddadi 罗运利 Celia Beaudouin 罗传秀 郑艳伟 李春海 魏金辉 杜春彬	
植被变化驱动机制的样带研究	张玉进 周广胜
德令哈大气水汽中 $\delta^{18}O$ 的时间变化特征——以 2005 年 7 月~2006 年 2 月为例	
	尹常亮 姚檀栋 田立德 刘东年 余武生 曲东梅
青藏高原植被变化与地表热源及中国降水关系的初步分析	
	华 维 范广洲 周定文 倪长健 李学敏 王永立 刘雅勤 黄先伦
考虑次网格变异性 and 土壤冻融过程的土壤湿度同化方案	田向军 谢正辉
阿拉伯半岛-北太平洋型遥相关及其与亚洲夏季风的关系	孙建奇 袁 薇 高玉中
印度夏季风的爆发与中国长江流域梅雨的遥相关分析	刘芸芸 丁一汇
一次罕见的平流辐射雾过程的特征	濮梅娟 张国正 严文莲 李子华