

电子科技大学

UNIVERSITY OF ELECTRONIC SCIENCE AND TECHNOLOGY OF CHINA

专业学位硕士学位论文

MASTER THESIS FOR PROFESSIONAL DEGREE



论文题目 轮廓检测和整合的脑认知机理及信息传输机制

专业学位类别 工 程 硕 士

学 号 201322090508

作者姓名 陈祥惠

指导教师 颜红梅 教授

分类号 _____ 密级 _____

UDC ^{注1} _____

学 位 论 文

轮廓检测和整合的脑认知机理及信息传输机制

(题名和副题名)

陈祥惠

(作者姓名)

指导教师

颜红梅

教 授

电子科技大学

成 都

(姓名、职称、单位名称)

申请学位级别 硕士 专业学位类别 工 程 硕 士

工程领域名称 生物医学工程

提交论文日期 2016.3.15 论文答辩日期 2016.5.24

学位授予单位和日期 电子科技大学 2016 年 6 月

答辩委员会主席 _____

评阅人 _____

注 1：注明《国际十进分类法 UDC》的类号

Brain cognition and information transmission mechanism of contour detection and integration

A MasterThesis Submitted to
University of Electronic Science and Technology of China

Major: **Master of Engineering**

Author: **Xianghui Chen**

Supervisor: **Hongmei Yan**

School: **School of Life Science and Technology**

独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。据我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得电子科技大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

作者签名：_____ 日期： 年 月 日

论文使用授权

本学位论文作者完全了解电子科技大学有关保留、使用学位论文的规定，有权保留并向国家有关部门或机构送交论文的复印件和磁盘，允许论文被查阅和借阅。本人授权电子科技大学可以将学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

（保密的学位论文在解密后应遵守此规定）

作者签名：_____ 导师签名：_____

日期： 年 月 日

摘 要

轮廓检测是高级动物视觉系统中一项最为基础且重要的功能，轮廓整合则被认为是物体或者情景感知所必需的，尽管在生理视觉皮层的研究上人们取得了显著的进展，但是其具体的脑认知机理和信息传输机制还不是很明朗。本文围绕轮廓检测和整合问题，通过眼动行为实验与 fMRI 实验，从自下而上的刺激驱动和自上而下的意识驱动两方面，对轮廓整合的脑认知机理和信息传输机制进行了研究，研究内容包括以下两个方面：

第一，本研究设计了无轮廓背景 (NC)、有显著轮廓 (OB) 和不显著轮廓 (UO) 三种刺激形式，研究了注视条件下人们在不同任务难度的轮廓检测与整合感知的正确率和反应时差异，同时用眼动仪监测了被试的眼动轨迹情况。实验结果发现，反应时由快到慢依次是：显著轮廓，不显著轮廓，无轮廓。反应正确率由高到低依次是：显著轮廓，无轮廓，不显著轮廓。此外，眼动轨迹状态还表明，三种情况下眼动轨迹存在明显差异，表明了人们可能对自下而上的刺激驱动和自上而下的意识驱动这两种类型的轮廓检测与整合过程的认知机制是不同的。

第二，为进一步对轮廓整合进行研究，我们设计了磁共振实验，并在检测轮廓错误的情况下加入 feed-back 对受试者提示，然后分析了五种情况下的功能性核磁共振成像 (Functional magnetic resonance imaging, fMRI) 信号，通过样本 T 检验。结果发现，当轮廓清晰，任务简单时，V4 区有明显激活；当任务难度增大时，额叶眼动区 (frontal eye fields, FEF) 区和 V4、V2 等区域则有较强的激活，提示 FEF 区和 V4、V2 可能参与了前馈和反馈的相互作用机制。当任务难度较大，被试无法完成正确的轮廓检测整合，并接受特定的 feed-back 后，其前额叶 (prefrontal cortex, PFC) 区和颞下回区 (inferotemporal cortex, IT) 活动增强，这表明轮廓检测相关的反馈可能是起源于 PFC 并且回流作用于 IT 区或 V4 区。

我们的证据表示，轮廓的检测和整合可能是通过自上而下和自下而上相结合的机制共同完成的。轮廓检测相关的高级脑区有 PFC 和 FEF，FEF 的作用更多的是通过眼动控制进行前期的检测，而 PFC 则通过接受外界的提示或记忆。此外，V1，V2，V4 间的相互作用可以完成简单轮廓的整合，任务难度较大时，则需要 IT 区和 LOC 区作为桥梁连通高级意识区和枕叶的视觉皮层区相互作用来完成轮廓的检测整合。

关键词：轮廓整合检测，功能性磁共振，眼动，top-down 影响，脑认知信息传输

ABSTRACT

Contour detection is one of the most basic and important functions in the advanced animal visual system. Contour integration is thought to be necessary for the perception of objects or situations. In this thesis, we combine the contour detection and the integration problem creatively through the research of eye movement behavior and fMRI together, in view of the bottom-up stimulus driven and top-down driven by two aspects of consciousness, the mechanism of brain cognition and the transmission of information is studied, and the research contents include the following two aspects:

Firstly, the target in the visual scene is simple or complex. In this paper, a new stimulus is divided into three kinds of situations: no contour background (NC), obvious contour (OB) and unobvious contour (UO), the three stimuli were randomly presented in fixation, of people in different stimulation difficulty of integration correct rate and reaction time differences also observed the correct detection and error of eye movement trajectory, help us to understand the whole process of contour perception. The experimental results show that: the reaction time is from fast to slow in order: significant contour, no significant contour, no contour. The correct rate of reaction from high to low is: significant contour, no contour, no significant contour. In addition, the eye movement track found that there were significant differences in the contour detection of the eye movement in three cases. It shows that the integration process is different.

Secondly, in order to further study on the contour integration, We designed the magnetic resonance experiment, and in the case of detection of contour error to join feed - back to subjects, and then analyzes the five areas of the brain fMRI signal through the sample T-test. The results show that when the contour is clear and the task is simple, the human V4 area is enough to complete the contour detection integration; when the task difficulty increases, the human FEF area and V4, V2 and other regions have a strong activation. When the task is difficult, the participants cannot complete the correct integration of contour detection, and receive specific feed - back, the prefrontal cortex (prefrontal cortex, the PFC) and temporal gyrus (inferotemporal cortex, IT) area increased activity, suggesting that contour detection related feedback may be originated in the PFC and reflux in the IT area or V4 area.

Our evidence suggests that detection and contour integration need top-down and bottom-up mechanisms to work together. Contour detection is related to the higher brain area is located in PFC and FEF, and the effect of FEF is more early detection through eye movement control, while PFC by accepting external cues or memory. In addition, V1, V2, V4 interactions can complete the simple contour integration. When the task is more difficult, it needs the IT and LOC areas as the bridge connecting the high consciousness area and the occipital lobe of the visual cortical area to complete the contour detection integration.

Keywords: Contour integration detection, fMRI, eye movement, top-down effect, Brain cognition information transmission

目 录

第一章 绪论	1
1.1 研究工作的背景与意义	1
1.2 轮廓检测和整合脑机制的国内外研究历史与现状	1
1.2.1 Bottom-up 及 top-down 知觉组织机制的研究历史与现状	1
1.2.2 轮廓检测和整合相关的生理神经机制研究现状	6
1.3 开展本研究的主要理由	9
1.4 本文的主要贡献与创新	10
1.5 本论文的结构安排	10
第二章 轮廓检测和整合的脑认知理论基础	12
2.1 格式塔原则	12
2.2 功能特化理论	13
2.3 TOP-DOWN 机制在轮廓检测和整合处理中的影响	15
第三章 眼动实验	17
3.1 实验目的	17
3.2 实验设计	17
3.2.1 实验被试	17
3.2.2 实验仪器	17
3.2.3 实验刺激	18
3.2.4 实验流程	20
3.2.5 结果数据分析和展示	21
3.3 本章小结	26
第四章 磁共振实验	27
4.1 实验目的	27
4.2 实验设计	29
4.2.1 被试	29
4.2.2 刺激	29
4.3 实验结果	31
4.3.1 结果预处理	31
4.3.1.1 任务态实验数据的预处理	31
4.3.2 实验结果	31

4.3.2.1 OB（显著轮廓）与 NC（无轮廓）的差异	31
4.3.2.2 UOW（不显著轮廓被试判断错误）与 NC（无轮廓）的差异	33
4.3.2.3 UOR（不显著轮廓观察者判断正确）与 OB（显著轮廓）的差异 ..	33
4.3.2.4 UOWF（不显著判断错误接受反馈后）与 UOW（不显著判断错误） 的差异	34
4.3.2.5 UOWF（不显著判断错误接受反馈后）与 OB（显著轮廓）的差异	37
4.4 本章小结	40
第五章 全文总结与展望	42
5.1 全文总结	42
5.2 后续工作展望	42
致谢	44
参考文献	45
攻读硕士学位期间取得的成果	48

第一章 绪论

1.1 研究工作的背景与意义

轮廓检测和整合作为动物和人类视觉感知的基础，这一共识已经被人们普遍接受，其在生理视觉皮层上的研究已经有显著的进展，但是其具体的神经机制还不是很明朗，更为严格的信息传输导向更是人们研究的热点。V1 区是感受野与视觉皮层连接的首要门槛，其在视觉感知，特别是轮廓整合中的作用不言而喻，但是 V1 到底能否独立完成复杂图形的感知，这仍然是人们所争论的问题。Bottom-up 和 top-down 的感知机制是如何进行分配的，其与轮廓相关的脑区有哪些？在本文中，我们的研究就是从这两个方面更为深刻的剖析轮廓检测和整合的动态机制，进一步从基础理论研究方面去为图形识别，轮廓提取等图像处理的计算模型奠定一定的理论基础。

1.2 轮廓检测和整合脑机制的国内外研究历史与现状

1.2.1 Bottom-up 及 top-down 知觉组织机制的研究历史与现状

自二次世界大战期间移民美国的一批德国心理学家，包括 Lurt Koffka 和 Wolfgang Kohler 等为代表的格式塔心理学家，是最早的系统的探索知觉分析和知觉组织的研究者，针对知觉组织分离，即指人类能够分辨哪些信息是属于一个整体进而把这些信息组织成一个独立物体的能力，提出了一系列的组织原则，拉开了人们研究轮廓检测和整合，进而是计算机物体识别的序幕。

格式塔心理学家最初假定，把各个知觉元素组合在一起这一过程是发生在视觉加工早期，并且很多证据也支持这一观点，但是在 1990 年洛克报告了不一致的实验结果，他们做了一个这样的实验：在黑暗中平行的呈现一排排的亮珠，这些亮珠在水平方向的距离小于垂直方向距离。当画面从正面向后倾斜时，在二维视网膜像上，这些亮的珠竖直距离渐渐的比横向间的还要小，但是在三维图像上还是水平之间的要小。受试者看到的是它们以水平的方向组织在一起的。洛克等人因此得出结论：整合（Grouping）是基于三维空间所感知到的接近性而不是在二维视网膜上的实际接近性，可见由接近性而形成的组合一定发生在深度知觉之后。因此，整合可以发生在比格式塔心理学家最初所认为的更晚些的加工阶段^[1]。

与轮廓检测与整合的后续相关理论还有统一连通性。Geisler、Perry、Super 和 Gallogly（2001）对人类是如何感知连贯性（即准确判断物体的轮廓）很感兴趣^[2]。

他们认为最好以视觉环境中的真实物体为切入点开展研究。因此，他们仔细地反复检查了花朵、树、山脉，河流等的轮廓。他们从这些材料中发现两个关键的原则：

(1)轮廓的相邻线段一般会有非常相似的朝向；(2)轮廓中相隔较远的线段一般具有相对不同的朝向。

Geisler 等 (2001) 向观察者同时呈现两个复杂的图形，并要求他们判断哪个图形包含一个环绕轮廓。结果表明，上述两个原则可以很好的预测观察者的成绩。这些发现提示，“轮廓整合机制的发展……由自然界中的各种图像的发生概率驱动” (Sekuler 和 Blake, 2002)。因此，当对轮廓进行判断时，我们会广泛利用关于真实物体的知识，这些知识就是我们的先验知识。以上观点是 bottom-up 和 top-down 共同作用的最初的原型。

Palmer 和 Rock (1994) 提出的新的知觉组织原则，统一连通性。它指出任何具有统一视觉特征（如颜色、纹理、亮度）的连通区域将倾向于被组织为一个知觉单元。他们认为，统一连通性会比格式塔原则的接近性和相似性更为有效^[3]。

不仅如此，在轮廓检测和整合的脑机制研究中，最早的格式塔心理学家还认为，各种知觉组织原则都是以自下而上（刺激驱动）的方式来运作的。北师大李武教授等人在 2014 年发表在《Neuron》上的文章 “Incremental integration of global contours through interplay between visual cortical areas” 中讲到，要想产生一个视觉目标的连贯表征，全局轮廓必须能够被正确组装，而这个轮廓整合的过程是遵循良好的格式塔原则的，经典的 bottom-up 观点指出：V1 具有较小的感受野且作用在局部的朝向分析上，而线段整合成复杂的形状是要发生在高级视皮层，随着大脑皮层层次的提升，神经特性也要有相应的复杂程度的提高。

当然针对上述观点，越来越多的研究者持有反对的意见，近来，就有很多的研究者发现，仅仅凭借刺激驱动方式是不能够完成一些复杂场景下的轮廓提取和物体识别的。有些学者发现：轮廓整合的中级处理过程在 V1 就已经被发现，V1 形态学上的长距离水平连接也可以用于调制具有相同朝向选择性的神经元特性^[4]。

那么，不可否认，在现实生活中，轮廓整合在复杂情况下会变得更具有挑战性，因为要将属于同一物体的段（元素）连接起来，从背景部分整合划分出来，这一过程在现实中将存在很多歧义，所以仅仅凭借前馈处理是不能够完成的。举个例子，如图 1-1 所示，图中的小鸟由于其羽毛色彩与背景几乎一致，背上的羽毛颜色同落叶树枝的颜色便非常容易混淆，所以，没有看见过鸟的观察者是不可能从该背景中将其识别出来的。有很多人，只有在外界提示情况下才能够感知到图片中小鸟的存在。甚至部分观察者，即便事先知道小鸟的存在也需要别人进行详细描

述才能确定鸟的位置和轮廓。



图 1-1 复杂背景中的目标物体

为了更好的了解和分析轮廓整合在皮层上的表现，Li 等人借助微电极阵列的方法记录并研究了猴子在观看轮廓时的 V1 和 V4 区活动（找到了孤立的峰值，两个电极：一个位于 V4 颞上沟和月状沟之间，另一个在 V1），检测了猴子对于相同轮廓的 V1 和 V4 的时空特性。结果发现，对于视觉刺激的响应延时，V1 延时在 44ms，V4 延时在 58ms，就是说，视觉刺激引起的信息从 V1 到 V4 的正馈需要 14ms。之后，他们又测量了与轮廓相关的延时和强度（V1 延时在 90ms，而 V4 的延时在 59ms），结果发现，V4 出现的轮廓相关的响应较 V1 早，且激活强度更强。

上述结果表明：轮廓分组最初是在 V4 有一个快速的前馈，接着是一个逆向流，使 V1 神经元将轮廓信号放大，同时抑制背景信号，当 V1 的轮廓响应启动，在 V1 和 V4 的轮廓信号建立并达到一个最大值。这意味着这个增量涉及到视觉皮层的双向互动。为增强轮廓信息，同时抑制背景噪声的两个逆向处理流，导致了皮层间循环的全局轮廓信息增量。前馈为简单刺激赋予神经元选择性，而反馈连接能根据刺激环境和行为目标（任务要求）动态调制其响应特性。真正的形成意识，构成感知应该是反馈来完成的，无论场景复杂程度怎样，任务难度怎样^[5]。

上述的结果我们可以看出, Li 等人认为, 对于轮廓的检测和整合是皮层间正向流和逆向流共同作用的结果, 且无论任务的难易程度, 刺激的复杂程度如何, 都是存在正馈和反馈的, 与轮廓相关的响应最早是从高级视觉皮层开始的。同时, 这种皮层间的相互作用最终会引起一个皮层增量。

在轮廓处理中皮质神经元的相互作用是复杂的: 1) V1 和 V4 神经元与轮廓相关的特性不同, V1 在任意朝向易化全局轮廓, 而 V4 只是在最优朝向轮廓有易化, V1 和 V4 不同响应特性表明在轮廓整合检测中各自的贡献, V4 具有较大的感受野, V4 的轮廓信号指示一个粗糙的信号(大致的位置和形状), 但是足够给 V1 一个轮廓加强, V1 可能提供了轮廓更高的分辨率和具体的位置信息, 对齐, 轮廓边界, 轮廓形状的细节。

以上处理机制与通用的理论架构是一致的, 基于 bottom-up 和 top-down 重复性处理分离和区分复杂视觉场景, 先前的研究发现类皮层分层网络能可靠快速的识别物体, 有些模型也研究了 top-down 机制是如何由 V1 局部回路操作的, 解释了轮廓加强和背景抑制。

众多的研究就衍生出了关于 V1 的上下文信号的来源问题, 存在以下观点:

1) 来源于许多朝向依赖性的影响, 包括格式塔连续性组织, 认为是 V1 水平连接调节引起的, 因为那些连接的神经元有非重叠感受野但是有相似的朝向选择性。

2) 归功于远程上下文影响。如环绕抑制, 面分割, 边界所有权, 高级皮层反馈等。

有的研究者还认为 V1 区长距离上下文调制是来源于反馈的, 是基于需要一个相对于 V1 视觉响应的额外延时获得一些上下文的影响, 然而这种观点被质疑, 存在一个大的重叠皮层区域的延时, 在相互联系的皮质回路有复杂的复发性神经相互作用在许多节点。另外一个起源于反馈的解释, V1 响应存在一个固定不变的延时, 与周边环境的距离无关(独立于上下文的距离影响), 就是一个单纯 top-down 机制反馈^[6]。

李武等人的研究同样发现了轮廓易化(或背景抑制)的固定延时, 无论 V1 感受野是落在轮廓上还是背景上, 这个固定的延时表明有轮廓引起的推挽式调节效应不需要轮廓信号在 V1 的一个区域到另一个区域。他们发现的 V1 中轮廓引发的调节效应具有朝向独立性, 这可以被 V1 水平连接所解释, 水平连接是将具有相似朝向的皮质连接起来, 横向连接则对局部朝向无特异性, 对所有朝向都有差不多的水平。

分层皮质网络中视觉场景成分由粗到细的增量整合, 包含两个处理的逆向流,

还有一个逆向的全局到局部的处理过程，被认为是视觉场景组织和分割的通用模型。携带着全局信息的相似的延时响应成分在清醒猴子的 V1 区被发现，用于轮廓整合，表面分割和感知填补主观轮廓。

在上述的延时期，人们对人类的 V1 区进行了经颅磁刺激，尽管高级视觉皮层被激活，但是被试分类自然场景图像的能力被大大减弱，说明从高级皮层到低级皮层的反馈在全局感知形成中十分重要。

Gilbert 等人在神经科学杂志上发表的《Top-down influence in visual processing》综述中提到皮层间的再进入或者反馈通道是携带着丰富多样的行为上下文信息，包括注意^[7]，期望，感知任务，工作记忆等。根据被执行的任务，神经元作为处理器接受输入承担不同的功能状态。最近的数据表明，对特定输入的选择代表整合野不同的成分，使神经元执行不同的功能，综述中讨论了在视觉皮层通道上不同的 top-down 影响，强调了感受野动态特性，允许神经元携带与当时感知任务有关的信息。皮层神经元的特性不是固定的，而是可以通过行为环境改变他们的功能，皮层神经元受注意，感知任务、期望的 top-down 影响，top-down 是指认知的影响^[7,8]，高级信息侵袭到信息处理的低级，基于皮层层次结构的反馈，用来处理比较复杂的场景，然而，在前馈通路叠加，有一个携带着高级皮层信息的反馈回到初级视皮层。它能够易化我们对视觉场景的翻译，同时可以在眨眼或者是持续快速的眼动时有对物体稳定的呈现。同时，它会易化我们对背景的分割或者多个物体的分割。除此之外，top-down 机制还扮演着对学习信息的编码和回忆^[9]。这与传统的信息处理机制不同，传统的信息处理机制认为，信息是正馈的，且处理的层次是越来越高的，最开始是简单的信息处理，当我们看到视觉场景时，我们会建立逆向流，在这里，我们考虑到感受野受 top-down 影响的特性，再次处理的信息的特性，神经元所携带的信息是如何依赖于行为。在很长一段时间，感受野能改变以适应视觉感受的改变，这些证据都指向感受野特性的这样一个事实，包括上下文影响，强调其动态特性，神经元有不同的特性以应对不同的经历和期望。Top-down 影响是通过丘脑核覆盖了一系列的下行通路贯穿整个大脑皮层的，如图 1-2，正向流是由 LGN 出发，始于 V1 的层次结构，有两条通路：一条是腹侧流，参与物体的识别，一条是背侧流，负责视觉引导运动和注意控制，对每一个正向连接（蓝线）都会有一个携带着上下文的反馈连接（红线）存在^[10]。

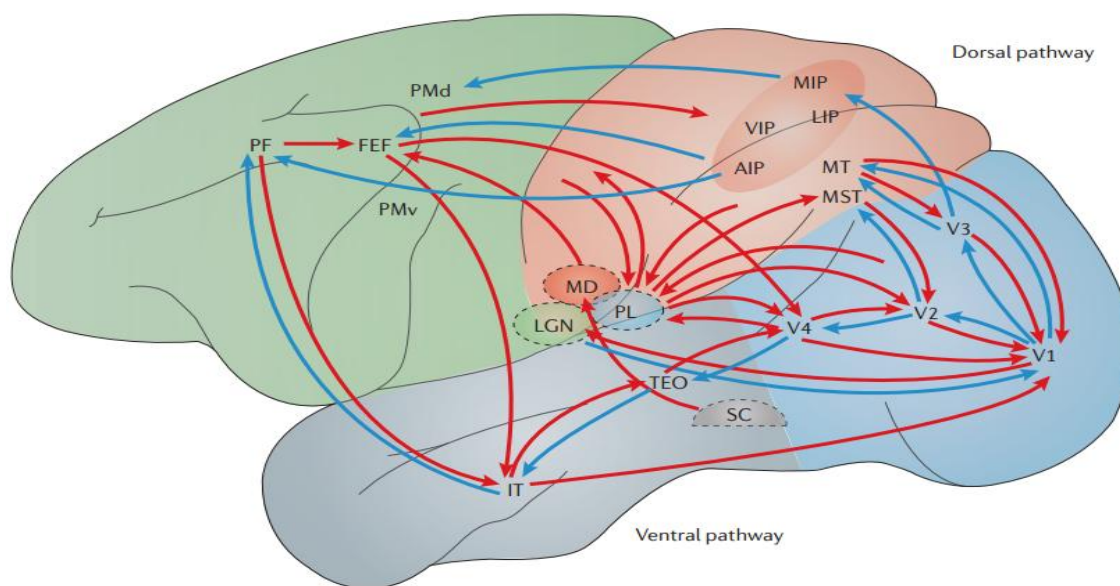


图 1-2 大脑皮层信息流正馈和反馈图（蓝色箭头为正馈、红色箭头为反馈）^[10]

1.2.2 轮廓检测和整合相关的生理神经机制研究现状

为了更好的探究轮廓检测和整合的脑认知机制和信息传输机理，生理学家在动物身上从来没有停止生理学的研究，目前也取得了很多实质性的进展。

了解感知的影响是如何作用在神经元功能特性就需要了解感受野的特性，感受野就是视网膜的一部分，刺激能够引起感受野上神经元一连串的反应。感受野的特性是依赖于刺激的特性，比如，一个简单的朝向刺激将激活视野一小部分的神经元，但是在该区域外的相似的刺激不能引起神经元的响应，却可以影响神经元响应，这些调节可以是易化，也可以是抑制，其大小和正负影响是取决于其侧面刺激与感受野中心的关系，因此，神经元的响应是依赖于全局轮廓的特性。

1968 年，Hubel 和 Wiesel 在 *J. Physiol* 上第一次发表了哺乳动物初级视皮层细胞的经典感受野结构^[11]，近五十年来，神经科学家又发现在经典的感受野外，还存在着一个范围很大的非经典感受野。李朝义院士的实验室在非经典感受野的研究方面起步较早，他们首先观察到，在猫的视网膜和外膝体神经细胞的经典的感受野外，都有一个面积比感受野大十倍以上的“去抑制区”（非经典感受野）（Li et al, 1987,1996）^[12]；在视皮层神经元的经典感受野外，也有一个大范围的非经典感受野（他们称为整合野）^[13]（Li et al, 1994; Yao et al, 2002），用视觉图形单独刺激 V1 区神经元感受野外区域不能引起细胞产生动作电位，但却能对传统感受野内刺激的反应进行调制，这种调制可能是易化或抑制的，调制区域的范围一般为 CRF 大小的 3~5 倍。他们发现整合野对提取图像的形状信息

(Chen et al, 2005) 和检测物体与背景间的相对运动起着至关重要的作用^[14]。这些机制被认为是早期视觉系统对轮廓检测与整合功能的神经基础。此外, 初级视皮层神经元还能够对某些主观轮廓错觉图形(一种纯主观的与高级认知过程有关的边缘轮廓知觉)作出反应^[15]。轮廓检测与识别会受到周围环境的影响。清醒猴实验表明, V1 区神经元对共线轮廓的反应会被周围背景抑制, 如果背景和轮廓越相似, 抑制越强, 轮廓检测和识别越困难, 如果把轮廓周围的背景去除后, V1 区神经元的反应就会加强。这种同质信息对目标知觉的影响, 就是由于初级视皮层神经元对相同特征响应时的抑制作用。Li Wu, Gilbert 等人通过行为实验, 详细研究了轮廓识别与整合的影响因素和轮廓与背景的相互作用。实验任务是让被试在杂乱短线段组成的背景中寻找共线短线段组成的轮廓, 实验结果发现, 当共线短线段之间的距离增加时轮廓识别的难度增大。虽然短线段的间距变小能够帮助轮廓识别, 但是当小到一定程度时对轮廓识别的促进作用几乎不会增加。当共线的短线段个数增加, 轮廓识别的正确率又增高, 但是当共线短线的长度达到 30° 视角, 再增加共线元素, 轮廓识别的正确率也不会再提高。背景杂乱时短线与共线短线的相对距离也会影响到共线短线轮廓形成。他们推断这些特性可能是由于初级视皮层(V1 区)存在水平连接的细胞, 会把局部的特征整合在一起, 使得较大空间范围的轮廓最终可以从复杂背景中被识别^[16]。随后他们在清醒猴电生理实验中证实了他们的推断, 即在清醒猴 V1 区的电生理实验中发现初级视皮层就能进行一定程度的轮廓整合, 表明基于物体特征的自下而上的信息传输机制在轮廓提取和边缘检测中起到了一定作用。我们最近在双眼竞争的行为实验中也发现, 有意识的同质轮廓元素形成的外周干扰, 会降低中心轮廓识别的正确率, 并且发现, 被抑制的外周轮廓元素在无意识下也可以被加工, 会形成外周的无意识轮廓线索, 并且促进中心共线轮廓的整合, 表明轮廓整合可能是起源于与意识相关的高级视觉处理过程以前。

电生理实验进一步研究证实了, 初级视皮层轮廓整合功能会受到高级皮层的调控。当轮廓混在杂乱的光栅组成的背景中时, 对于学习过轮廓的清醒猴子, V1 区皮层会产生更为强烈的神经反应, 表明初级视皮层中轮廓识别等过程相关的神经通路会受到学习及高级皮层信息反馈的调控^[17]。另一方面, 在复杂背景中进行轮廓检测和整合时, 需要将相关特征联系起来, 并与其它感知输入区别开来, 同时赋予其某种几何意义, 这需要空间和时间的信息以及认知机制才能整合成形状, 这些任务一般被认为是较高级皮层的处理过程。Field 等人通过行为实验结果最早猜想可能存在一个联合区域(Association field)网络, 能够将识别类似特征的相邻神经元信息整合起来, 这种机制能够把朝向或其他信息相似的视觉信息整合成一个

完整的轮廓(Field et al, 1993)。Shpaner 等人利用 ERP 实验方法, 结果表明, 轮廓整合更加依赖高级皮质层的神经元反应, 例如: 侧枕叶皮层 (LOC)^[18]。fMRI 实验也表明, 当受试者检测到有共线 Gabor 光栅组成的轮廓时, 很多高级脑区被激活, 涉及的脑区包括侧颞枕叶皮层 (LOC)、初级视皮层 (V1)、前纹状皮层 (V2)、腹侧第三视觉复合体 (VP) 等^[19]。实验推测, 视觉特征性信息通过视觉通路传入初级视皮层, 初级视皮层进行简单的组合后信息流入到 V2、VP 区等皮层, 并且高级皮层会给予初级皮层强烈的反馈, 初级视皮层对轮廓的处理主要是信息特征的处理, 包括方向、频率等等, 而高级皮层会整合知觉的形状等整体的大范围信息, 然后反馈给初级皮层以调节信息的处理过程和范围, 最终完成轮廓的识别。Van Humber 等人通过行为学实验表明, 眼动在轮廓检测任务中也起着重要作用^[20]。轮廓检测和整合的联合区域猜想最近在清醒猴行为实验和电生理相结合的实验中得到了证实。实验使用延迟匹配任务, 让猴子对之前学过的线索刺激任务作反应, 发现猴子的初级视皮层会对较复杂的形状做出反应, 而且不同的最优特征下的形状的识别会激活联合区域网络的不同区域, 他们推测, 初级皮层的横向连接提供了一个潜在的广泛连接, 高级皮层和学习等自上而下的信息能调控初级视皮层动态的特征识别过程^[21]。但是, Bushnell 等人对猴子进行实验发现, 在灵长类动物中, 在负责场景中基于轮廓的封闭物体信息是在 V4 上进行编码的^[22]。

Gilbert (2013) 强调了期望在场景分割中的重要性, 由于场景的复杂性, 目标感知最困难的不是一个目标的识别, 而是与目标相关的轮廓元素或者面的整合, 并且将他们从背景中分离出来。尽管在理论上, 认知可以单独在前馈机制完成, 但是 top-down 处理或者逆向流机制需要确定在多个目标组成的复杂场景中的场景分割, 反复性处理过程的模型可以解决视觉机制非常具有挑战性的任务: 元素的组处理和分割。

另外一个任务依赖性的例子就是曲线跟踪任务, 当感受野与轮廓重合时与不重叠时比在 V1 有一个响应的加强, 在复杂背景中的轮廓检测能够引起与轮廓相关的神经元响应的易化, 随着轮廓线的增加, 感知能力也提高。在动物中, 进行轮廓检测任务时共线元素的易化影响比进行不相关的任务时大的多。任务依赖的自上而下的影响, 是对视野中刺激特性的调优改变。

不仅如此, top-down 的影响还与感知任务有关, 研究发现猴子被训练判断图片上是狗还是猫, 是跑车还是轿车, 图像从狗演变成猫, 从跑车演变成轿车^[23]。其前额叶的神经元在不同任务时的 firing rate 是不同的。支持了神经元的多任务处理特性^[10, 24-26]。

最近的电生理研究也表明前额叶眼动区 (frontal eye fields, FEF) 是注意错误

信息选择性的皮质区，对人类此位置的经颅磁刺激表明和解剖学都发现，额皮质到 V4 或者是 MT，可能携带着对不同刺激特点相关的不同自上而下的信号^[27]。

对前额叶（PFC）的记录显示，神经元可以表现出不同的调优以进行分类，所以，当任务改变时，同一神经元可以表现出不同的分类表现，建立了这样的行为普遍性，神经元的选择性是依赖于不同的行为上下文（比如，神经元的形状选择性是依赖于动物不同的形状搜索任务的）^[28]。

同时，认知活动与目标期望也是存在关系的，当动物被暗示寻找特定的形状，V1 神经元的形状选择性便变成接近暗示形状或者综合部分形状的状态。特定几何形状选择性的证据来源于一个实验：动物被训练成从复杂背景中找到嵌入的轮廓，轮廓是直线、圆或者波形，测量 V1 神经元的形状选择性是在每个感受野中有错误和正确的目标呈现，而动物会做一个扫视去看正确的目标。这个实验重要的发现：1) V1 神经元对复杂的几何图形具有选择性（不光是单一的直线）2) 改变对形状的判断期望，选择性可以从单个神经元或者一个表层神经元的切换。这过程表明对目标的期望创建了一组滤波器对目标的成分进行选择，同时这还需要 top-down 机制的参与。该思想在共享成分的目标感知学习转移当中被进一步的证实^[10]。

大脑皮层对刺激的响应是依赖于先验知识和当前任务的，通常颞中回（middle temporal, MT）的神经元对特定方向的移动刺激有响应，但是接受过训练的动物其 MT 对静止的刺激刺激也有响应，这表明这活动不仅反映了刺激状态，也包括认知状态和视觉想象。研究发现，MT 只是对移动物体有响应，但是训练后，对静止的但有方向标示的物体变得也有响应。

1.3 开展本研究的主要理由

边缘检测与轮廓提取是高级动物视觉系统的一项基本任务，深入了解视觉系统在轮廓检测方面的神经机制和认知机理，对进一步建立基于生物视觉神经机制的轮廓检测模型和算法，具有重大基础研究和应用研究意义。我们提出本研究的理由如下：

(1) 建立基于生物视觉神经机制的轮廓检测模型和算法对于机器视觉有着重要的意义。

(2) 生物视觉的轮廓检测与特征整合具有非常复杂的脑机制，涉及多个脑区的共同作用，信息传递既有前馈也有反馈调控。要使计算机能够像人眼一样快速准确地处理图像边缘或轮廓等复杂信息，需要在多个研究层次上深入了解轮廓检测的神经机制和认知机理。本研究中提出围绕轮廓检测这一中心任务，是轮廓检测

从生理、行为过渡到模型三维一体的研究方案的中间环节，至关重要且可能产生创新性成果。

1.4 本文的主要贡献与创新

轮廓检测整合在动物身上的生理学研究为人类的研究提供了很好的参考价值，不同的皮层之间的相互作用是任何感知活动的前提，但是感知活动是受很多条件和环境影响的，V1 区是感受野与视觉皮层连接的首要门槛，其在视觉感知，特别是轮廓整合中的作用不言而喻，但是 V1 到底能否独立完成复杂图形的感知，这仍然是人们所争论的问题。Bottom-up 和 top-down 的感知机制是如何进行分配的，其与轮廓相关的脑区有哪些？Top-down 影响的起源在哪里，接受外界的提示和自发的调用先验知识，这两者是不是一样的作用机制和信息传输导向呢？为了更好的去研究轮廓整合这一特异性的感知过程。本论文结合轮廓最基本的形成特性，共线性和接近性，设计出直观的轮廓刺激，以行为眼动技术、任务态磁共振技术针对轮廓检测和整合问题的脑认知机制和信息传输为重点研究内容，主要创新点与贡献如下：

采用全新的轮廓刺激，利用格式塔原理的接近律和相似律基本原则，巧妙的实验流程范式，加入自发的注意控制和来自外界的自上而下的反馈信息作用比较，分别在宏观认知行为层面和生理皮层角度全面模拟人类在视觉场景中的轮廓检测。具有“生命科学与信息科学相结合、宏观认知与微观电生理相结合、数学模型与生理实验相结合”的研究特色，研究成果将会为机器视觉和图像处理技术提供重要的理论基础和技术支撑。

1.5 本论文的结构安排

本文的章节结构安排如下：

第一章是绪论，主要介绍了本文的研究背景，即在轮廓检测和整合领域目前已有的研究成果、亟待解决的焦点问题等。

第二章是与认知相关的经典脑认知理论基础。包括了格式塔原则和功能特化理论。

第三章是眼动行为实验，目的是了解我们的刺激是否达到了预期的任务难度区分要求，同时获得人们在观看不同刺激时的行为表现，比如正确率、反应时和眼动图，为功能磁共振实验的设计提供基础，做好准备。

第四部分是利用 fMRI 成像技术，深入分析观察者在面对轮廓刺激时，能否形成意识，不同形式的完成任务时，其脑区激活程度差异变化，最终得到与轮廓整

合相关的 bottom-up 和 top-down 脑区及其相互作用的信息流推断结果。

第五部分是对本文的总结和展望，主要讨论了该论文中结果分析的局限性以及未解决的问题。同时还有针对性的提出了该课题在未来的探究中可以采纳的方法和方向。

第二章 轮廓检测和整合的脑认知理论基础

近年，研究者在轮廓检测和整合领域已经取得了相当大的进展。这方面的进展很多来自认知神经科学家的努力，他们的工作使我们能够更多的了解轮廓整合所涉及的大脑系统。我们将首先介绍轮廓加工所涉及的经典脑认知理论。为后面我们详细讨论知觉过程做基础准备。本章，我们重点介绍参与轮廓检测和整合的主要大脑系统。为了完全了解大脑中的脑知觉加工过程，我们有必要考虑眼睛和大脑皮层所发生的信息加工过程。

2.1 格式塔原则

Max Wertheimer 认为具体的轮廓是因为其几何关系而被感知为一个整体的，他们在轮廓检测和整合方面做了如下图所示的研究。

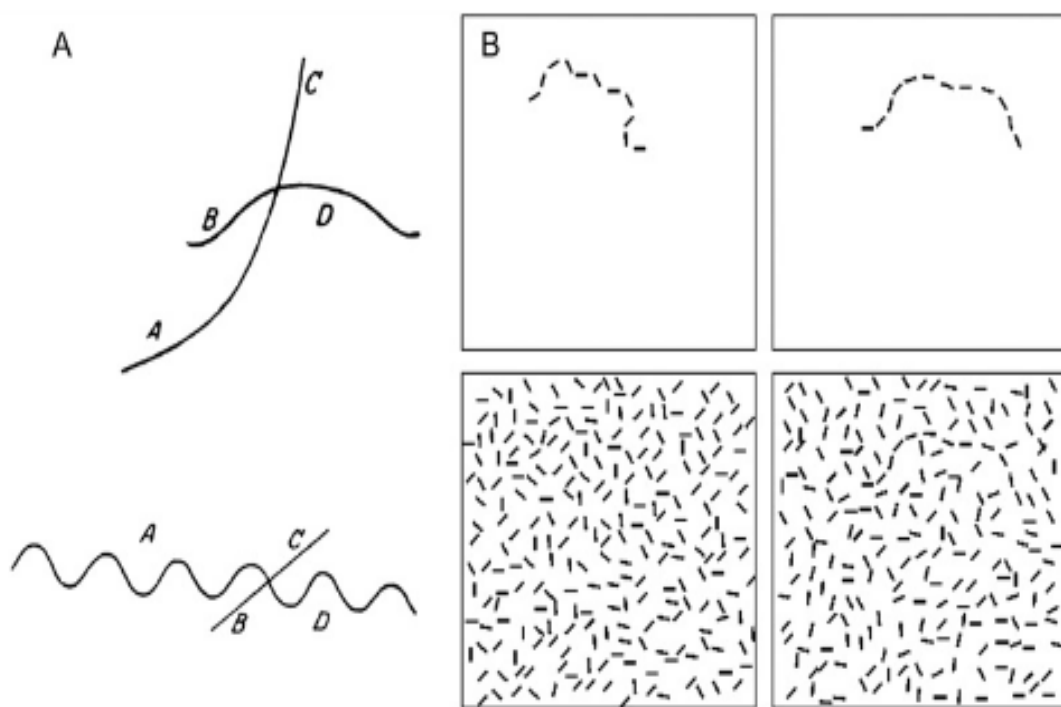


图 2-1 感知组织的格式塔原则^[32]，A 图为格式塔原则的“最优延续”，B 图为相似性原则在轮廓显著性上的适用

Max Wertheimer (1938) 的研究如上图 2-1 (A) 所示，人们自然的会认为，A 部分和 C 部分同属于一个轮廓而分离段 A 和 D 为单独的轮廓，因为 A 和 C 保持

同一方向和曲率。这是被称为原则的“最优延续”；图 A) 下则表示，轮廓线可以遵循复杂的组织形式，AD 部分是正弦线。

图 2-1(B)是相似性原则在轮廓显著性上的适用，表明相邻直线段的相对方向及其接近程度可以影响其是否能从背景中凸显出来形成轮廓。右边的轮廓线，由类似面向的连续线元素组成，当嵌入在随机取向和定位线的背景中时，能够被弹出。而左边的轮廓线段元素有更多的变化，在相邻的行元素的方向，不能被弹出，需要一个扩展的搜索过程才能从下图中找到。

生活中不光全是这种具象轮廓，还有很多抽象轮廓存在，如考尼饶三角形，是在三个缺角圆盘之间有个白色的正三角形，并可以清楚的看到三角形的三条边，如果注视更久一些，会感觉这个三角形的白色和背景的白色多少有些不同。但是如果遮挡掉部分黑色图形，就再也看不到白色的背景之上的白色三角形了，更不要说白色的边啦，为什么这种错觉会栩栩如生呢？这是因为，在观看目标物体时，大脑会把视觉刺激里面的各种细节组织成一些整体，有时为了组织成整体，大脑甚至会补充上一些缺失的成分。这样就造成了错觉。至于更加细致的解释，目前还在争论中。但是可以肯定的一点是我们的大脑是根据已有的经验做出了这种判断。

电生理实验表明，在初级视觉皮层旁边的高级视觉皮层中枢有许多细胞，不仅对真实的边缘轮廓有反应，而且对这种错觉边界也有反应，然而，初级视觉皮层中的细胞则只对前者有反应，而对后者没有反应。

2.2 功能特化理论

ZeKi (1992,1993) 提出了一个功能特化理论 (functional specification theory)，该理论认为，视觉皮层的各个部分分别负责不同的视觉功能。打个比方，视觉系统就像一支工作队伍，每个队员都负责解决某一个复杂问题的一部分。这些人工作的结果被整合起来就产生了一个最终的解决方法，即统一的视知觉^[29]。

ZeKi 列出了恒河猴的一些主要的视觉皮质。视网膜主要与初级视觉皮层 V1 相连接。V1 是十分重要的，这一点，我们可以从对视网膜至 V1 区这一通路上的任一点进行损毁都可导致所影响的 V1 区会发生全盲的证据中看出来。而 V2 到 V5 等在视觉加工中的重要性也不容忽视。ZeKi 对这些区域的功能进行了描述：

V1 和 V2：这两个区域参与视觉信息的早期处理。V1 和 V2 区由对颜色和形状做出反应的不同细胞群组成，而这些区域还包含一些负责在信息传送到各特定视觉皮层前对各类信息进行分类的分装站。

V3 和 V3A 区：这两个区域主要对形状，特别是运动物体的形状做出反应。

V4 区：该区域中的绝大部分细胞对颜色做出反应。此外，还有很多细胞对线条的朝向做出反应。猴的这一区域有些与众不同，它与颞叶和顶叶皮质具有广泛的连接。

颞下回区（inferotemporal cortex, IT）：Sugase、Yamane 和 Kawano（1999）向猴大量呈现人脸，猴脸和简单的几何图形（如方块、圆等）。结果表明，在刺激呈现大约 50ms 后，神经活动会受到所呈现刺激类型的影响。在刺激呈现数百毫秒后，神经活动会受到刺激的更为具体的特征（如面部表情）的影响^[30, 31]。

Kobatake（1994）在猴身上发现，检测负责物体的神经元在 IT 区前部最常见。相反，对中等复杂性做出反应的神经元多位于 IT 区后部和 V4 区^[32]。Young 和 Yamane（1992）报道，在 IT 区一组相对小数量的面孔识别神经元能够对一定的面孔进行编码并产生一个激活模式^[33]。

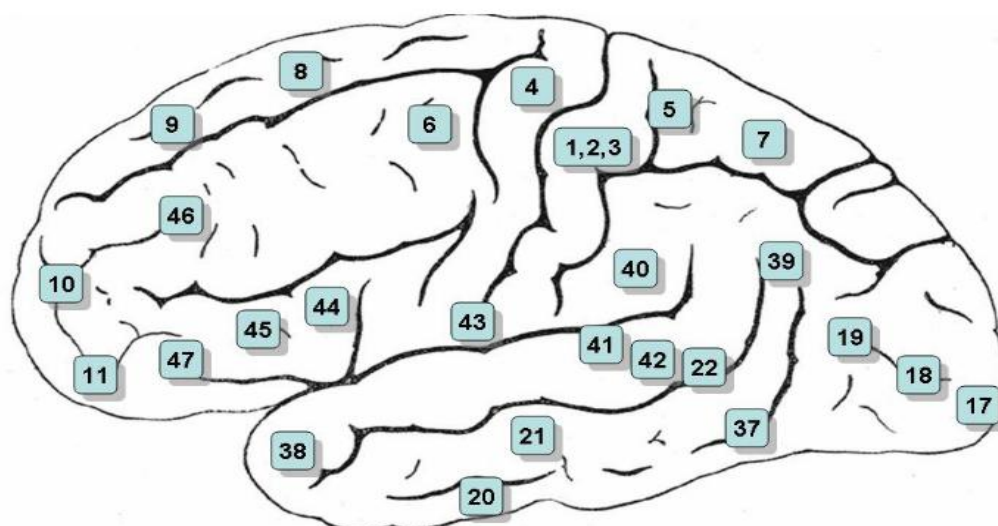


图 2-2 大脑皮层 Bradman's Areas 划分（图片来自网络）

Bradman 对脑区的划分与功能特化的脑区划分之间存在对应关系，如图 2-2 所示，其中 V1: BA17，V2: BA18，LOC: BA19，IT: BA20。

前额眼区(Frontal Eye Fields, FEF)：这个叫做前额叶眼动区的皮层在视觉注意和眼动方面扮演着重要的角色，在 FEF 区进行电刺激可以引发眼跳和微眼动。FEF 有地形结构和代表在视网膜的坐标扫视目标。额叶眼领域大致位于地区 BA4，BA6、BA8 之间。

前额叶区（prefrontal cortex, PFC）：它大致位于地区 BA9，BA10、BA11 之间。

Zeki 的功能特化理论是对非常复杂的大脑系统作出的一个相对简化的解

释，但是这种简化也存在很多局限性，它不太好解释信息整合问题，以及对视野中物体的各种信息进行整合的捆绑问题等。

2.3 Top-down 机制在轮廓检测和整合处理中的影响

所有皮层和丘脑的感觉处理水平受强大的自上而下的影响，对图形低级水平的处理更是需要复杂的信息。自上而下的相互作用的多样性的新发现表明，皮质区功能是作为自适应处理器，受到关注，期望和感知任务的影响。大脑的状态是由多个皮质区之间的相互作用和调制固有回路的反馈连接来确定的。在知觉学习当中，学习信息的编码和召回涉及的选择适当的输入，传达的刺激信息被抑制。这种相互作用的中断便可能导致行为障碍，包括精神分裂症^[8]。

一般理论框架认为：虽然越来越多的神经系统科学家开始建立研究怎样的皮层区域的激活和皮层神经元的反应与行为，持久存在的问题是大脑状态的基本特性是什么，在皮质电路产生的认知水平上的基本算法是什么。为了得到这个算法，我们必须分析大脑回路的行为方面。传统的信息学观点认为大脑信息处理是基于一个分层的组织。在视觉系统中，路径从分析非常简单，局部属性的分析，到表示视觉的信息变得越来越复杂，整体是作为一个移动的层次结构。然而，从计算的角度来看，这是不可能的，前馈机制可以实现灵活的、不变的复杂和迅速变化的环境模式识别。最近的研究结果已经从根本上改变了自上向下的相互作用在皮质上的角色，范围和功能多样性。我们已经了解到，大脑皮层的任何区域，包括初级视觉皮层的功能，是受关注、期望和感知任务的自上而下的影响。世界的固有的内在表征，是通过经验获取，来影响我们的大脑分析视觉场景的策略。视觉是一个主动的过程，和任何皮质区的功能是不固定的，每个区域根据上下文和目前的感知需求运行不同的程序。因此，视觉处理是涉及逆向信息流的，初级视觉皮层的操作涉及自底向上的信息和来自高阶皮质区的反馈连接之间的相互作用。

自上而下影响的一般思想是，在更高阶段的复杂信息处理会影响在先前阶段发生的简单处理过程。虽然一些早期的工作空间注意研究表明，在最受关注的自上而下调制的实例中，视觉注意的影响只发生在视觉通路中的高级水平，但是在皮质感觉处理的最早期阶段，神经元的功能特性是受到注意，以及其他形式的自上而下调制的，这一点已经变得越来越清楚。注意在感知中的角色也已经不仅仅局限于是一个聚光灯，探照灯的身份。注意这个概念，可能本身就已经是一个全方位的自上而下的影响。自上而下的影响已出现在一个广泛的种类范畴中，包括功能特征，表面，目标对象，对象类别，时间背景，几乎任何形式的知觉组织。此外，自上而下处理效果的影响可以是许多不同的种类，不仅包括增加或乘法的

响应，也有锐化调整曲线，控制上下文的影响，或者可以作为一个具有可塑性的调制器等。因此，必须扩大关注一系列的自上而下的影响描述的定义。

第三章 眼动实验

眼动记录方法作为分析受试者观看刺激的最主要的方法之一，常用于行文分析和追踪眼动轨迹问题。本文利用眼动跟踪来探究被试观看轮廓刺激时的注视点轨迹规律。

3.1 实验目的

轮廓检测是物体识别的基础，是格式塔原理最直接的表现，那么人类在进行最基础的轮廓整合时的眼睛注视点是怎么移动和变化的呢，随着共线线段距离变大，共线线段识别任务难度加大，被试在反应时和感知正确度上的表现有什么变化？众所周知的包括 V1 区在内的低级视觉皮层调节情况与轮廓感知是否形成以及任务难度存在着关系，眼动实验也能很好的测试我们的实验刺激是否能够代表各种轮廓检测和整合的情况。

我们试图通过眼动仪，结合模拟的虚拟轮廓刺激，通过调整轮廓线段间距，来调整轮廓的复杂程度和识别难度，分析不同难度的刺激形式引起被试的不同眼动以及判断情况，探究当轮廓存在时，被试形成知觉和未形成知觉的行为表现，还有轮廓检测和整合过程中的眼动差异。

3.2 实验设计

利用眼动记录仪，首先需要选取适当的视觉刺激、视觉刺激呈现范式，同时确定实验过程中需要记录的相关数据以备后期数据分析。

3.2.1 实验被试

10 名被试参与了该眼动实验 (6 男 4 女，年龄 24.1 ± 2.23)，六男四女，均为右利手，被试视力正常或矫正后正常，均事先不知道本实验的目的，且自愿参加，在正式实验前，每个被试进行两个 rounds 的训练，训练数据不计在内，每个被试要在三种不同的轮廓下，重复观看 40 个试次。

3.2.2 实验仪器

实验是在眼动仪实验室环境下进行。实验刺激是首先由 matlab 软件编写出如图 3-1 所示的轮廓图片。实验刺激则是在 matlab 和 PTB (Psychtoolbox) 共同编写完成，被试观看的电脑显示器是一台 21 英寸的戴尔 CRT 显示器，其空间分辨率是

1024*768 像素，刷新频率 100HZ。显示器距离被试 58cm，被试在做实验过程中，头部有头托支撑，可有效防止头部移动所引起的数据不准确。

本实验采用眼动仪 eyelink2000 (SR research, Canada) 测量和记录被试右眼数据，采样频率为 1000hz。该眼动仪采用眼角膜反射的原理记录眼动轨迹，简单准确，无需佩戴任何设备，接近自然环境下的注视观看，每个实验开始前都会进行眼动仪的方位九点校正，通过方能进行正式实验。

3.2.3 实验刺激

实验刺激采用 MATLAB 生成的。背景部分，我们采用了随机朝向的线段元素，刺激的轮廓部分，则是共线线段元素组成。刺激背景为黑色，线段元素为白色。

如图 3-1 所示，刺激中的线段元素全部为随机朝向，即无轮廓类型(no contour)。随机生成的图片中，尽管各个线段元素的朝向都是随机的，但是难免会出现几个相邻线段朝向是相似或者相同，为了避免造成混淆，这种图片要手动剔除掉。

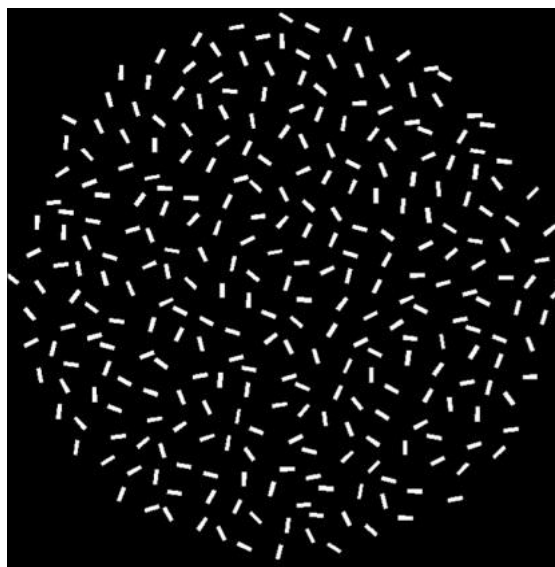


图 3-1 无轮廓刺激示例图 (NC)

如图 3-2 (a)，是任务难度较低，有显著轮廓的情况；图 3-2 (b) 是用红色线段元素表示有标注轮廓位置的图片，每张有轮廓的图片都生成了一张相应的 b) 图，以备被试在观看刺激，做出错误判断后给予反馈使用。

特别的，整个实验为了避免垂直和水平效应，不使用垂直和水平的轮廓刺激。

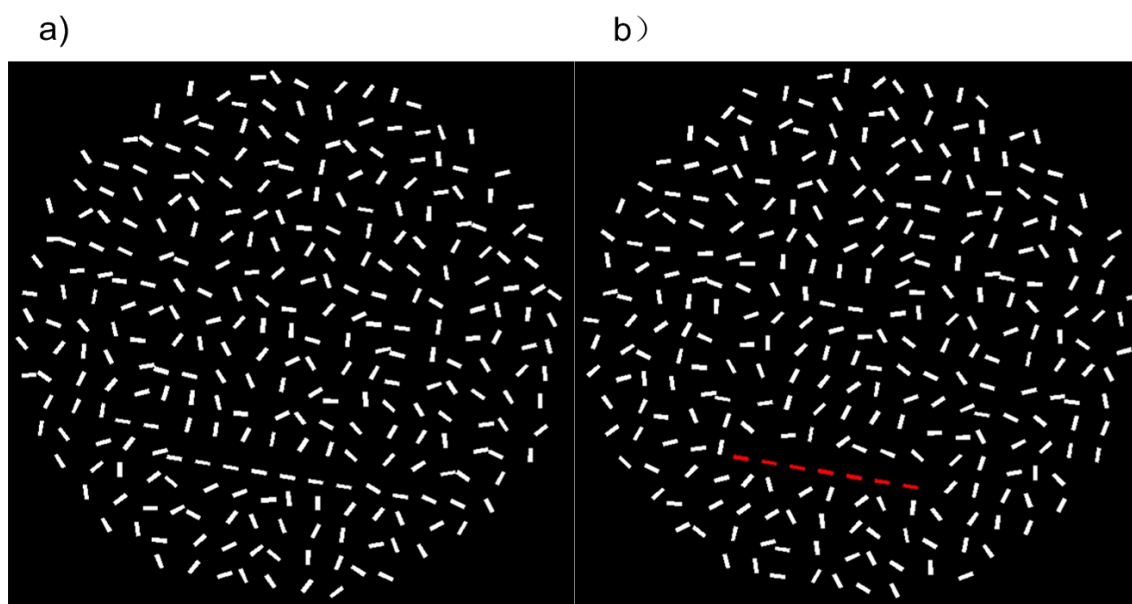


图 3-2 有显著轮廓示例图（SC），（a）图为被试判断刺激图片示例图，（b）为判断错误后进行反馈所使用的示例图

如图 3-3，有不显著轮廓（即轮廓连接间距变长，切断了全局形状只能依靠更高的层信息出现，被试通过跨越更大的元素距离进行耦合）。图 3-3（b）是用红线段元素表示有标注轮廓位置的图片。

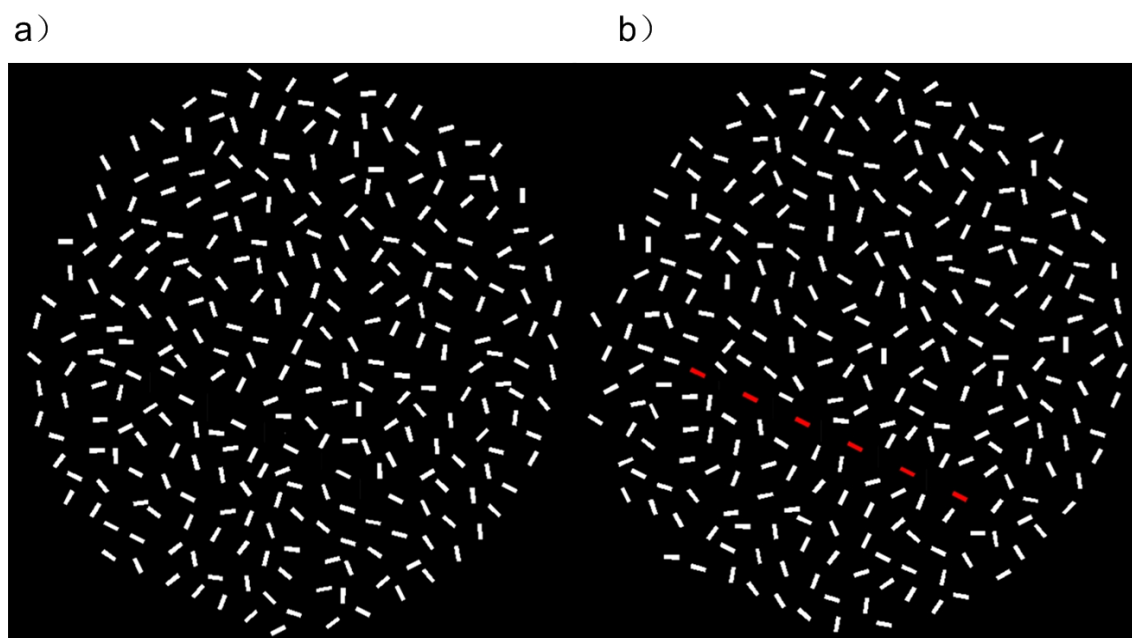


图 3-3 不显著轮廓示例图（NSC）

图 3-3 共由 294 条线段组成，无轮廓的情况每个线段的朝向随机，有轮廓的情

况则是线段元素共线，其他朝向随机线段作为背景。最特殊的情况，便是最后一种，它是通过将线段的中心间距变为原来的 1.5 倍，生成的刺激。我们共生成每类图片 40 张，一共是 120 副图片。同时有轮廓的情况，我们还相应的生成了用红色线段元素表示轮廓的图片，为受试者在实验中提供反馈使用。

3.2.4 实验流程

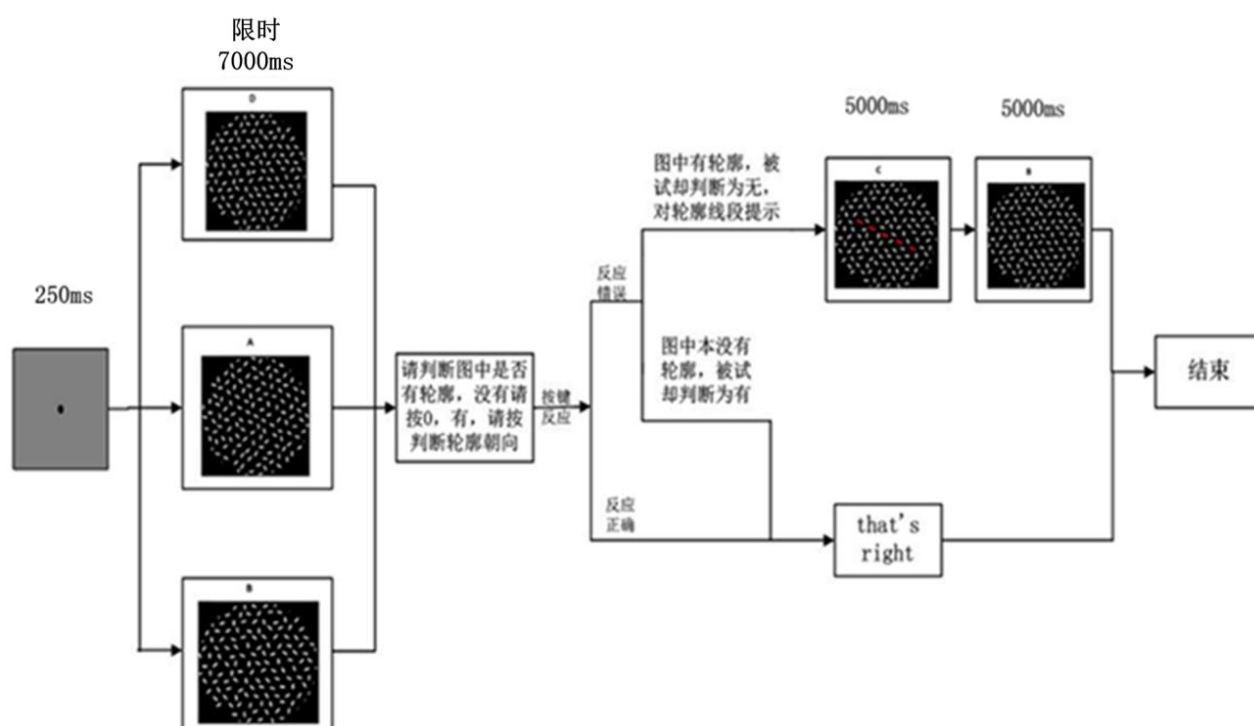


图 3-4 眼动实验的实验流程图

眼动实验具体的刺激呈现流程图如图 3-4 所示。实验准备阶段，要求受试者在电脑前保持坐姿，同时确保能够观看清楚整个屏幕内容。通过 ENTER 键进入实验，每个 RUN 开始前，均需进行眼动校正。每个试次首先是在屏幕中央呈现一个黑色的注视点，持续 250ms，然后在屏幕上随机呈现三种轮廓中的一种，呈现时间持续到被试做出按键反应，（0 键为无轮廓，1 键为轮廓朝向向左，2 键为轮廓朝向向右），时间上限设定为 7000ms。之后我们会对受试者有按键反馈，如果判断正确，则进入下一个试次；若判断错误，屏幕将呈现如图 3-2（b）和图 3-3（b）的反馈，让被试再做一次判断。每个 RUN 由 40 个试次组成，每个受试者共做三个 RUN，且每次呈现的刺激没有重复。在正式实验前，每个被试进行一个 RUN 的预实验。

3.2.5 结果数据分析和展示

在眼动实验中，我们主要记录了每个试次的刺激类型，正确按键，受试者的按键，判断反应时间，眼动坐标等变量。

所有实验数据都采用 **MATLAB** 进行处理，下图是具有代表性的一个被试的反应正确率和反应时间图，分别如图 3-5 所示

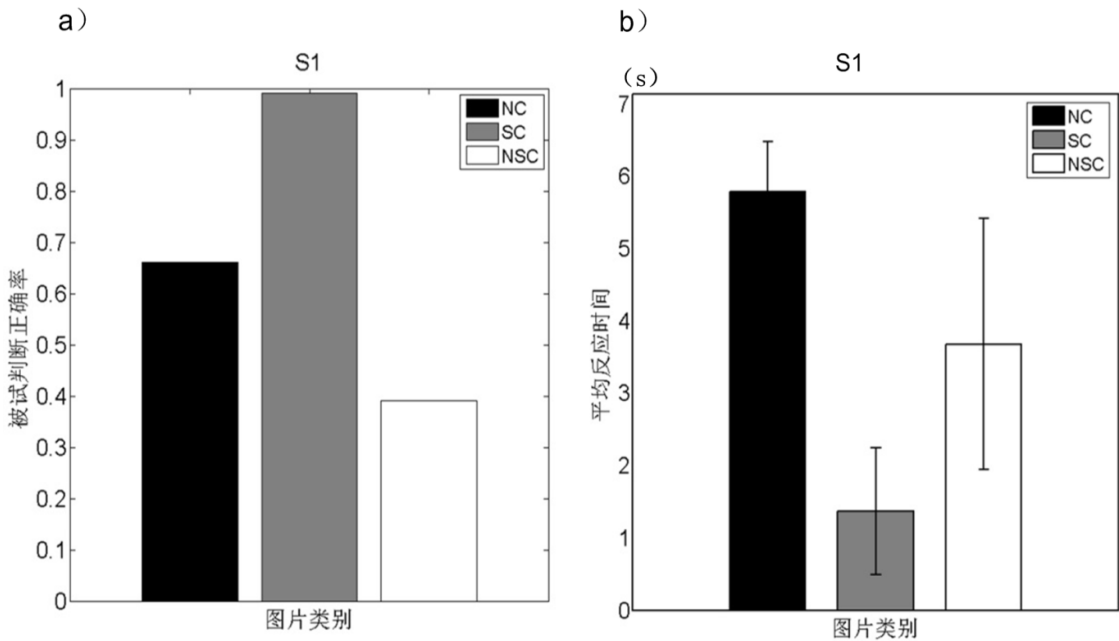


图 3-5 单个被试的反应正确率和反应时。黑灰白三色柱状图分别表示无轮廓，有显著轮廓和有不显著轮廓三种不同的图片刺激类别。

被试 1 在观看无轮廓的刺激时，其正确率在 **chance level** 以上，达到了 67%，而观看显著轮廓刺激时，其正确率 99% 以上，而不显著轮廓时，其正确率只有 40% 不到；（b）图中可以看出，被试 1 在观看显著轮廓时，被试判断反应时间平均只有 1.34s，而特别的，无轮廓的情况下其反应时间最长。

我们总共采集了 10 个被试的数据，其不同图片类型下的反应时和正确率如表 3-1 所示

表 3-1 10 个被试在三种刺激情况下的正确率和反应时的表现 (AC 正确率, RT 反应时)

		S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
NC	AC	0.81	0.66	0.63	0.63	0.64	0.68	0.75	0.64	0.68	0.72
	RT	4.32	5.79	5.81	4.86	4.67	4.77	6.08	5.29	5.11	5.24
SC	AC	0.97	0.99	0.94	0.94	0.97	0.97	0.94	0.95	0.97	0.98
	RT	1.26	1.37	1.73	1.58	1.47	1.23	1.33	1.56	1.62	1.69
NSC	AC	0.49	0.39	0.16	0.41	0.33	0.28	0.35	0.36	0.35	0.28
	RT	2.7	3.68	2.75	3.39	3.02	2.44	2.97	3.18	3.63	3.55

在反应时上的表现趋势完全相同, 因此最后我们做了一个平均正确率和平均反应时。三种轮廓的正确率和反应时分别是 (0.684 ± 0.088 , 0.962 ± 0.024 , 0.34 ± 0.14 ; 5.194 ± 0.733 , 1.484 ± 0.212 , 3.131 ± 0.481)。如图 3-6 所示

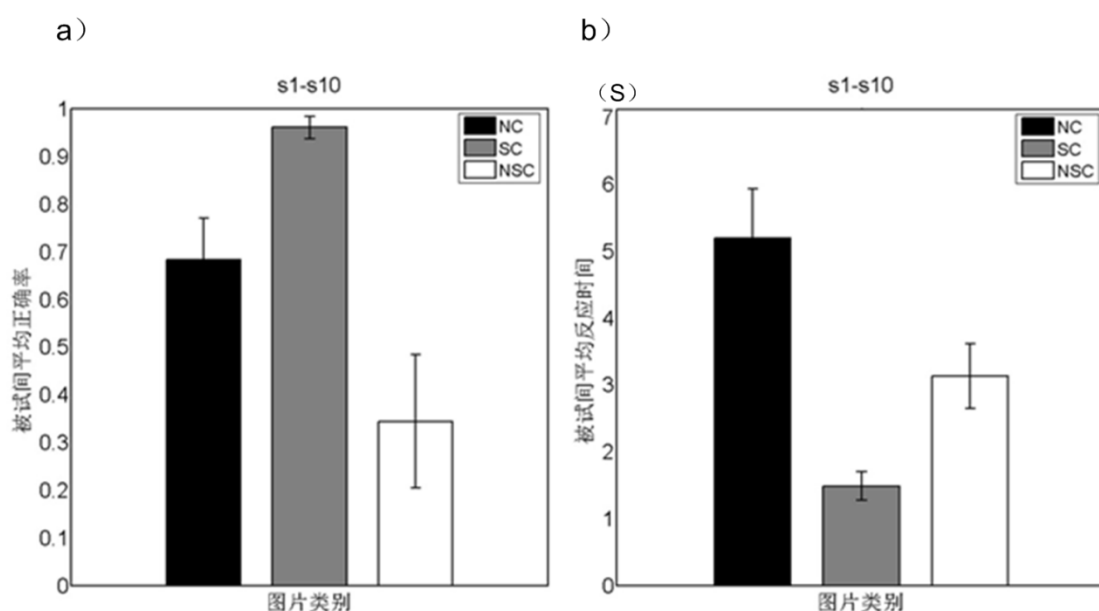


图 3-6 被试总体的平均正确率和反应时间

我们可以看到显著轮廓 OB 条件下, 整合感知所用时间最短, 同时正确率稳定的达到 90% 以上; 是不是感知时间短证实了刺激简单, 同时正确率会越高呢? 比较无轮廓 NC 与不显著轮廓 NSC, 我们发现, 无轮廓 NC 情况下被试所需的时间较不显著轮廓 NSC 长, 反应出任务难度是较后者大的, 但是其正确率却较后者高,

说明人们倾向于感知背景，除非真正检测整合到确切轮廓（如同 OB 情形）；同时，被试在呈现不显著轮廓 NSC 刺激条件下，正确判断的反应时约为 3.131 ± 0.481 ，也就是刺激呈现的时间的一半以下。换句话说，被试如果不能从复杂背景快速识别出轮廓的话，他将失去信心和耐心，放弃进行组织感知，这与现实生活中，我们在玩图片细节抓取类任务或者游戏时的表现是一致的。

表 3-2 正确率在三种不同刺激条件下的单因素方差分析结果

Multiple Comparisons							
Dependent Variable: 正确率							
	(I) 图 片类型	(J) 图 片类型	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
LSD	1	2	-.277000 [*]	.068235	.003	-.43136	-.12264
		3	.340500 [*]	.068235	.001	.18614	.49486
	2	1	.277000 [*]	.068235	.003	.12264	.43136
		3	.617500 [*]	.068235	.000	.46314	.77186
	3	1	-.340500 [*]	.068235	.001	-.49486	-.18614
		2	-.617500 [*]	.068235	.000	-.77186	-.46314
*. The mean difference is significant at the 0.05 level.							

从该结果表中可以看出，方差检验的相伴概率最大为 0.03，其相伴概率小于显著性水平，表示拒绝零假设，也就是说，控制变量不同的图片刺激类型造成了被试在轮廓检测正确率判断上的显著性差异，同样，在反应时间上也存在显著性的差异，如表 3-3 所示。

表 3-3 反应时在三种不同刺激条件下的单因素方差分析结果

Multiple Comparisons							
Dependent Variable:反应时间							
	(I) 图 片类型	(J) 图 片类型	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
LSD	1	2	3.709500*	.368417	.000	2.87608	4.54292
		3	2.064250*	.368417	.000	1.23083	2.89767
	2	1	-3.709500*	.368417	.000	-4.54292	-2.87608
		3	-1.645250*	.368417	.002	-2.47867	-.81183
	3	1	-2.064250*	.368417	.000	-2.89767	-1.23083
		2	1.645250*	.368417	.002	.81183	2.47867

最后，我们还分别就被试在观看并正确判断三类图片时的眼动做了分析。
无明显轮廓以及有轮廓但被试判断错误时候其眼动情况如图 3-7 所示

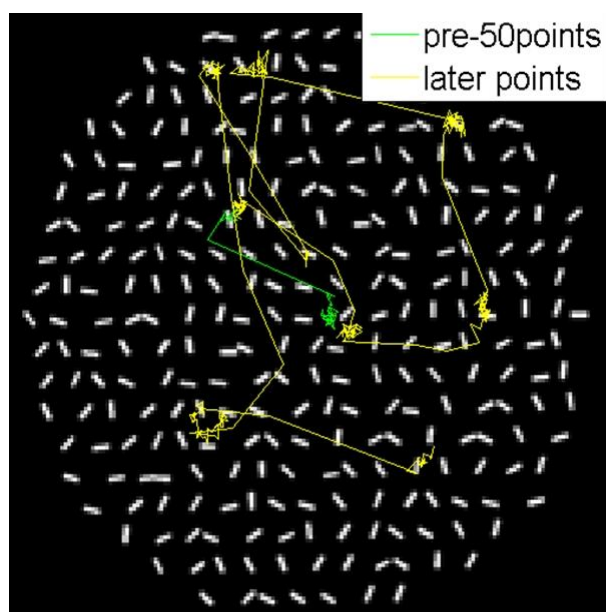


图 3-7 无轮廓及有轮廓被试判断错误时的典型眼动轨迹图（绿色部分代表前 50 个时间点的眼动轨迹，黄色部分是剩余时间点的眼动轨迹图，为了区分开轨迹的先后，我们以 50 个时间点作为界线）

上图我们看出，被试没有形成轮廓整合时，其眼动轨迹是杂乱的且在整個图片范围内进行检索的，分布散乱，无规律。

当有显著轮廓或者不显著轮廓但被试能够很快感知其轮廓情况及位置时，其眼动情况如图 3-8 所示

受试者在这种情况下，其眼睛的注视点会从中心注视点，快速定位到轮廓所处的位置，然后进行反复的扫视，以确定加工，最终做出判断。

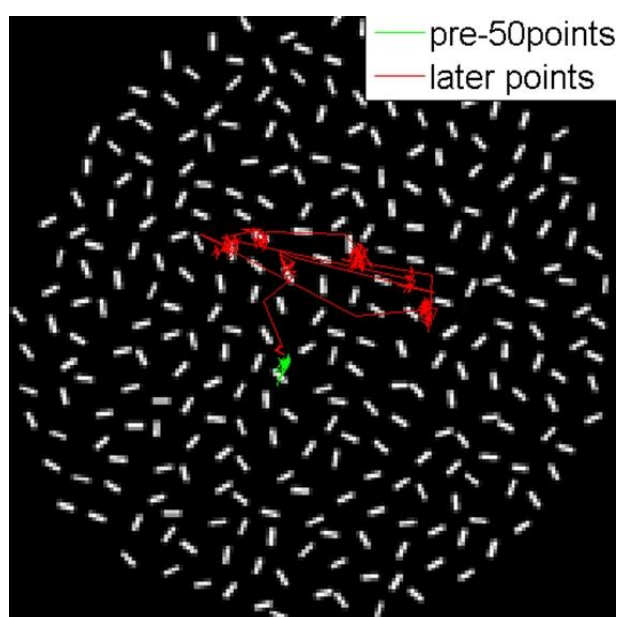


图 3-8 有显著轮廓情况下的受试者判断正确的眼动图（绿色部分代表前 50 个时间点的眼动轨迹，红色部分是剩余时间点的眼动轨迹图，为了区分开轨迹的先后，我们同样以 50 个时间点作为界线）

还有一种情况，如图 3-9 所示。该情况下被试不能快速的定位轮廓所处的位置，被试的眼动开始是在刺激中大范围的扫视，一旦发现类似轮廓的线段，便会在疑似线段处，来回扫视，确认每个线段的朝向与相邻线段的关系，最终做出抉择。

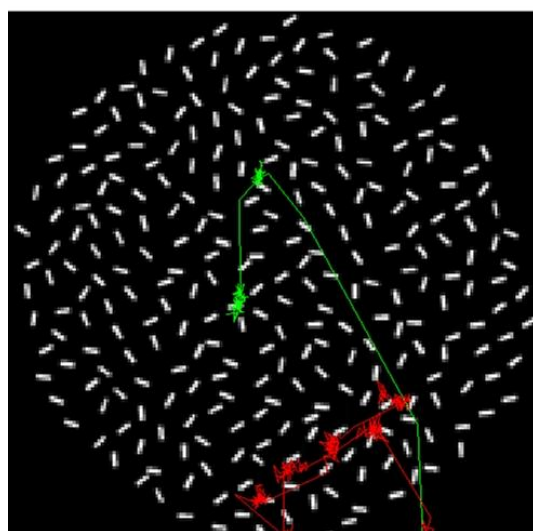


图 3-9 不显著轮廓情况下的受试者判断正确的眼动示例图（绿色部分代表前 50 个时间点的眼动轨迹，红色部分是剩余时间点的眼动轨迹图，为了区分开轨迹的先后，我们同样以 50 个时间点作为界线）

图 3-9 所示的情况通常是存在不显著轮廓刺激，真实状态情况更为复杂，前面的扫视点更加多，更加杂乱。

3.3 本章小结

本章首先从行为眼动数据得到了不同轮廓刺激在行为层面的表现，得到不同被试在观看这些刺激时所表现出的共性。我们对这三种轮廓的眼动注意控制也有了更深的认识。

针对第一种情况，即无轮廓的刺激，被试反应时间较长，正确率较高，同时眼动轨迹在整个刺激图片上散布，说明受试者进行轮廓检测时会控制眼动在整个刺激范围内进行查询，但是最终没能发现目标轮廓，从而没有形成整合过程。

第二种情况，有显著轮廓的刺激，被试反应时间最短，且正确率极高，从眼动轨迹图看到，被试在观看此类刺激时，能够快速定位轮廓位置，并且控制眼动在疑似位置处扫视以确认。

第三种刺激情况，存在不显著的轮廓，被试反应时间较长，反应正确率也低于 50%（chance level），任务难度较高，在大多数情况下，被试第一次判断都是错误的，正确率只有百分之三十左右。我们对其眼动情况作了分析发现，当判断正确时，如图 3-9，先是有个全局的检测，然后捕捉细节重点。而判断错误的情况下的眼动就如同无轮廓的（如图 3-7），尽管刺激存在轮廓，但是受试者一直在探索，最终也没有形成轮廓意识。

通过眼动实验结果的分析，我们对三种刺激在判断正确和错误下的行为表现得到了清楚的认识，这有助于帮助我们设计磁共振实验，同时为我们在磁共振数据分析的情况划分和被试对刺激处理过程的分析打下了基础。

第四章 磁共振实验

4.1 实验目的

轮廓整合作为视觉感知的基础，其最初的起源，以及最终调节作用形成感知的区域是发生在初级视皮层还是高级视皮层，这个问题一直是人们研究的焦点。Marina Shpaner 等人在 2013 年发表在《NeuroImage》上的文章，利用脑电 ERP 的方法，初步探究了轮廓整合中 V1 和 LOC 区所扮演的角色，研究了人类在 gabor 线段轮廓整合时，有无轮廓的差异。鉴于 ERP 的 C1 信号被认为是最早的视觉相关电位成分，是 V1 皮层活动的制造者，因此他们首先分析了 C1 成分来验证早期视皮层的轮廓整合是否在 V1 区建立，结果发现，虽然 V1 区对刺激有响应，但是有轮廓与无轮廓之间是不存在显著性差异的，也就是说整个处理过程中出现轮廓整合并非发生在初级视皮层。而真正的有轮廓与无轮廓的 ERP 显著性差异是发生在 NC1 处，那就说明整合是在此处开始的。他们的数据支持，轮廓的整合最早是依赖于汇聚到高级视觉皮层（例如 LOC）的整合汇聚信号的。

BogdanMijovic 等人在 2014 年《NeuroImage》上发表的文章“The dynamics of contour integration: A simultaneous EEG - fMRI study”，将脑电和 fMRI 技术结合起来，得到了在不同轮廓刺激情况下的知觉处理时空差异^[34]。他们采用的四种不同轮廓刺激如图 4-1 所示：

LG（Local and global）：所有轮廓元素沿嵌入图形轮廓对齐

GO（Global only）：只有一半元素沿嵌入图形对齐，其他垂直于该方向（把轮廓连接切断了全局形状只能依靠更高的层次形态出现，被试通过跨越更大的元素距离进行耦合）

NC（No counter）：嵌入图像的轮廓元素平行于其他，即质地均匀

D 图：为确保每个被试可以捕捉到轮廓，不参与分析，只是吸引注意并作按键反应

结果发现，LG 和 GO 两种刺激，只有在 LOC 区与 NC 刺激不同。而功能磁共振信号分析中 LG 和 GO 两者没有差别。但是在 ERP 中 200ms 处却有巨大差异，这表明两者的处理机制是不一样的，且不同仅仅是时间上的不同。同时在功能磁共振的分析中显示 LG 和 NC 差异比 GO 和 NC 差异大，特别是 LOC 区。上述研究强调了两点，一是强调了轮廓整合最早是发生在 LOC 等高级视觉皮层，二是发现无论轮廓刺激的难度如何，其整合机制是一样的，与无轮廓比较只是在 LOC 区存在差异，他们认为 LG 和 GO 两者的处理机制仅仅是时间的不同。

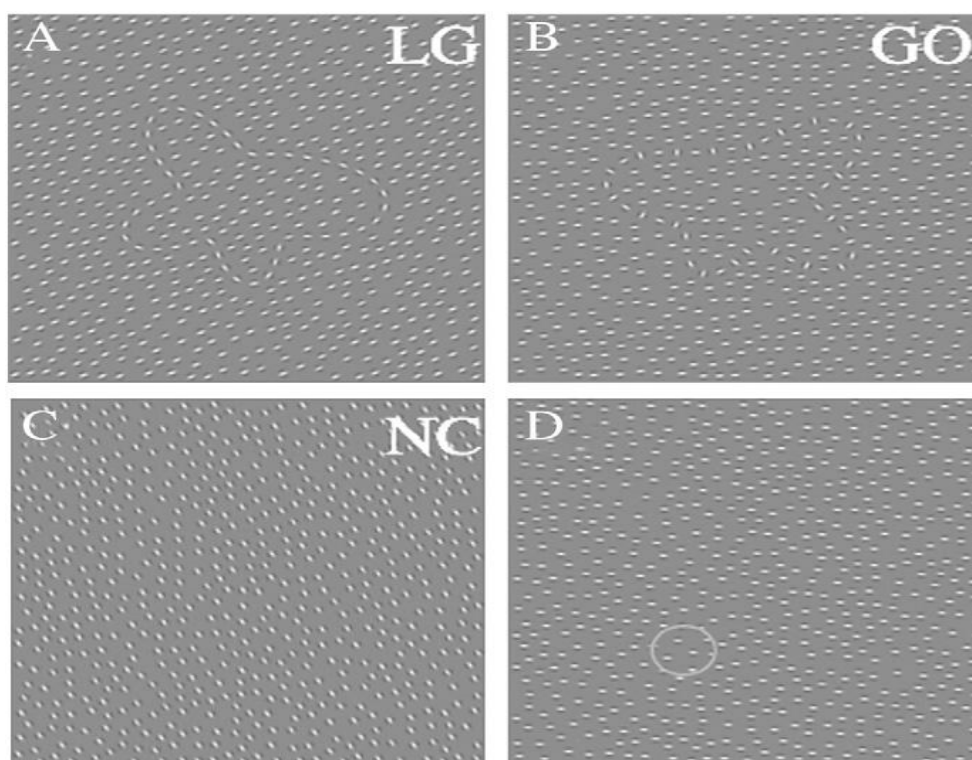


图 4-1 BogdanMijovic 实验刺激分类^[34], A 图既有全局信息又有局部信息, B 图只有全局信息, C 图无轮廓, D 图不参与分析, 只是吸引注意并作按键反应

早期的研究还发现, 在虚拟图形、虚拟数字甚至是背景和图形的感知任务涉及到自上而下的作用影响时, 都会有低级视觉皮层 (V1 或 V2) 调节现象的出现, 人们普遍认可这种调节是高级视觉皮层的一种反馈调节, 但是, 调节的信息流方向以及调节是抑制还是易化的, 却始终存在争议。有人从局部元素以及全局图形的角度出发, 发现低级视觉皮层的调节是对局部元素的易化, 对全局信息的抑制; 也有的研究者 (Peter Kok, 2014, *current biology*) 巧妙利用了虚拟图形具有屏蔽不期望的局部元素, 同时却能够形成期望的图形的特性, 以此来研究初级视觉皮层上的神经元, 发现在早期视觉皮层的反馈可能是抑制也可能是增强。这取决于反馈的信号是否与自下而上的输入信息适当叠合一致, 换句话说, 低层感官区域的神经元活动是取决于感官刺激以及这个刺激是否被反馈的预测信号预测到而决定的^[35]。

尽管 Bogdan 等人分析了 LG 和 GO 这两种情况下脑区激活情况是相同的, 但是他们没有针对任务难度以及是否形成感知去分析, 同时尽管 LG 和 GO 两者激活的脑区相同但是 GO 这种情形下其激活程度是否比 LG 情形下的激活程度增强或者抑制, 他们也没有去分析, 而这部分研究恰恰在复杂背景下的轮廓识别有重要意

义。

我们采用自己形成的刺激形式，在更加接近自然轮廓的前提下，同时将条件变量分成了背景（无轮廓）与图形（有显著轮廓和有不显著轮廓），形成知觉（被试判断正确）与未形成知觉（轮廓存在，但被试判断错误）的情况。很好的了解被试在背景感知，图形感知，以及认知形成前后初级视觉皮层的表现差异，分析人类在轮廓感知时的调节机制和信息传输情况。不仅如此，我们还在被试第一次判断错误，或者未检测到轮廓时，给被试一个快速的反馈，相当于间接的给被试 top-down 的信息，接着，刺激再次呈现在被试面前让其判断，从中我们得到了其低级视觉皮层以及中间层受先验知识调节的表现情况。

实验最终目的是得到有无轮廓，是否形成知觉的大脑皮层脑区差异，同时探究依靠先验知识的自上而下调节所产生的源头以及其反馈作用的低级或中级调节区位于何处。从宏观上初步推测出轮廓检测和整合的脑认知信息传输机制。

4.2 实验设计

4.2.1 被试

10 名被试（9 男 1 女，年龄 23.7 ± 2.41 , mean \pm SD），全部为右利手，有正常视力或者矫正后视力正常。

4.2.2 刺激

实验刺激使用 e-prime2.0 编写，借助背投屏幕 lc-xl100 投影机（1024×768 分辨率，60 Hz），调用的图片仍为眼动时生成的刺激图片，每种类型的图片由原来的 40 张减少为 30 张，并从剔除的图片中挑选几张让被试熟悉轮廓类型。刺激图片呈现在屏幕中央，大小像素点数为（658*661）。为防止白屏影响，屏幕周围的背景色设置为黑色，刺激流程在眼动实验上稍作改进，每个 Trial 固定长度 25s，90 幅图片随机呈现，随机顺序是进行伪随机，根据眼动在反应时的表现，我们将刺激呈现时间限定在 5s，然后 3s 内进行按键判断，超过 3s 则认为无响应处理，按键则设置为有向左的轮廓刺激（左边在上）按 1 键，向右的轮廓刺激(轮廓右边在上)按 2 键。为了不误按，无轮廓情况的按键设置为 4 键，为了使每个被试不受上一个刺激图片画面的影响，每个试次间隔之间用黑屏处理。

特别注意的是，当被试没有按键或者按键错误时，40ms 反馈提示与前面的原刺激一一对应，如果原刺激是无轮廓的情况，则反馈给被试文本信息“刺激为无轮廓”。若有轮廓被试判断错误，则呈现相应带有红色标注的反馈图片。

磁共振实验的流程图见图 4-2。每个被试在进行正式实验前，进行 20 个试次的预实验，预实验的刺激数据与正式实验不存在交叉，同时也不参与数据分析。

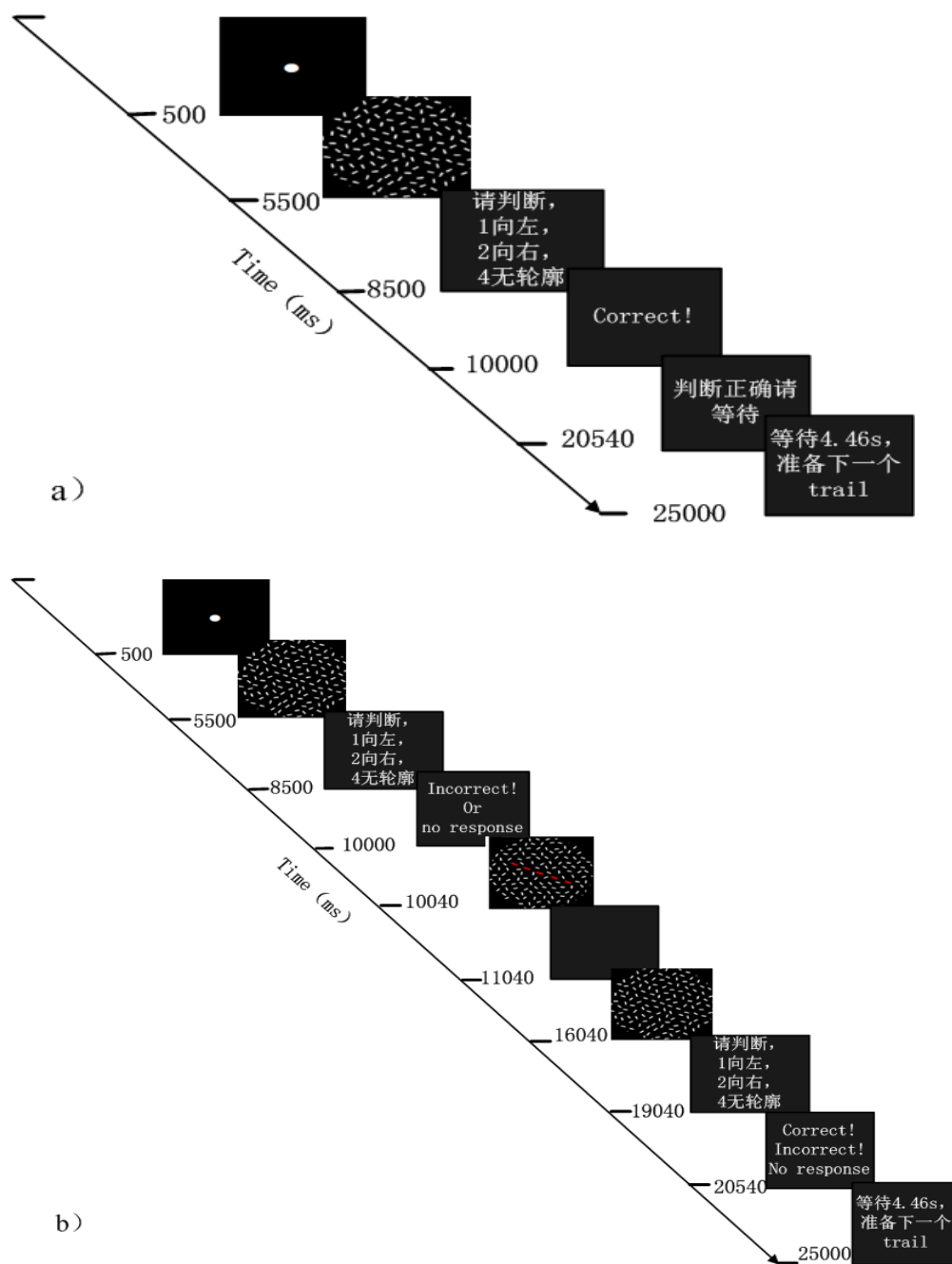


图 4-2 上面 a)、b) 两个图为磁共振实验的流程图，a) 图是被试观看刺激时判断正确的情况，反馈一个 Correct 和判断正确请等待的提示，最后等待进入下一个试次；b) 图是被试观看刺激时判断错误或者无反应的情况，随后依次呈现，40ms 的提示，1000ms 的 mask，5000ms 的原刺激，再次让被试判断，判断反馈，等待，进入下一个 trail。

4.3 实验结果

4.3.1 结果预处理

我们采用的序列：TR=2S,SLICES=43，隔层扫描，我们将 90 副图片分为三个 RUN，每个 RUN 有 30 个 Trails，每个 trail 总时长 25s，一共是 750s，一共有 375 个时间点。而静息态项有 255 个时间点。

实验对静息态结构像的预处理采用 DPARSF 软件，处理步骤包括：去除前十个时间点，以去除被试适应环境的不稳定影响；层级时间校正（slices timing）将各层之间的扫描时间同步；头动校正检验，包括三个方位的平动和转动，平动大于 2mm 或者转动大于 2 度的被试的数据被剔除掉，最终剩下七个被试数据；空间标准化，将每个被试的图像都配准到标准 EPI 模板，同时剩余的被试与 EPI 模板配准的效果都良好。

4.3.1.1 任务态实验数据的预处理

任务态实验数据的预处理同样使用 DPARSF 软件，做了头动校正和标准化处理。头动校正后，将受试者头动大于 1.5mm，1.5 度的数据剔除掉。

下一步，根据任务态的时间点，利用 SPM 将条件变量（condition 项）分为 5 种，分别是 NC,OB,UOR,UOW,UOWF，同时将被试判断正确的情况分类，比如 OB 情形下是第 2 4 6 个 trail，那么时间点就是 25.5 75.5 和 125.5s，结合五种条件变量进行个体统计。其中 NC 表示 no contour（无轮廓）情况，OB 表示 obvious contour（显著轮廓）情况，UOR 表示 unobvious contour reponse right（不显著轮廓被试判断正确）的情况，UOW 表示 unobvious contour reponse wrong（不显著轮廓被试判断错误）情况，UOWF 是指 unobvious contour reponse wrong then feedback（不显著轮廓被试判断错误并反馈）的情况。

4.3.2 实验结果

实验的结果分析是对五种不同情况两两相减进行 T-test，并采用了 AlphaSim 多重比较校正方法，分析它们在脑区激活上的表现是否存在显著差异。

4.3.2.1 OB（显著轮廓）与 NC（无轮廓）的差异

OB（有显著性轮廓）与 NC（没有轮廓）两者之间的差异，见图 4-3，OB-NC 是指用 OB 的激活区域减掉 NC 的激活区域进行的样本 T 检验，红色框圈起来的区域表示为存在正差异的正激活区，绿色框区域为负差异的负激活区，以下未说明

均属于同样的情况。

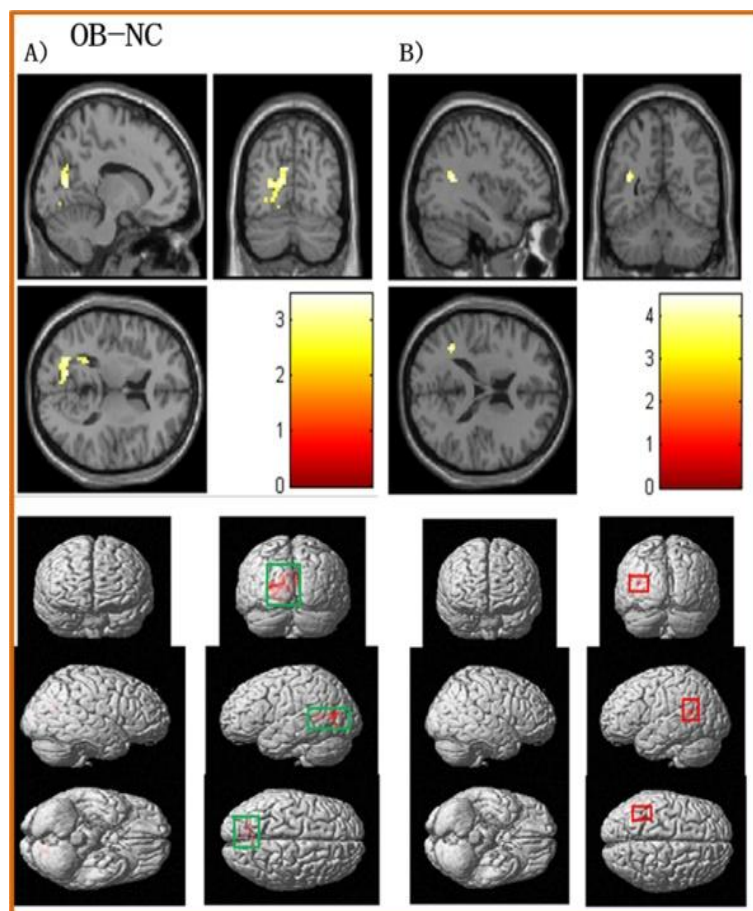


图 4-3 OB-NC，即被试判断正确情况下的显著轮廓激活脑区减去无轮廓激活脑区差异图，在立体图上用绿色框标记出来的脑区表示前者在该脑区激活强度较后者弱，其差异为负激活；反之，红色表示为正激活（AlphaSim correct, $p=0.05$, cluster size=10）

如 4-3（B）所示，显著轮廓激活脑区减去无轮廓的激活脑区差异在左颞上回（left temporal lobe superior temporal gyrus），即 BA22 对应的 V4 区有一个正差异（峰值坐标（-36 -54 18））。在右侧枕叶（Right occipital gyrus）的 V1 和 V2 区有一个负差异。相对于无轮廓的情况，当受试者观看显著轮廓时，人类较高级视觉皮层 V4 被明显激活，表现在图 B。不仅如此，我们还发现低级视觉皮层却出现一个抑制调节，表现在图 4-3（A）。

上述结果表明，OB 刺激本身有显著的视觉特征；NC 刺激元素则杂乱无章，无法形成轮廓意识，但受试者按任务需要进行检测有无轮廓，OB 较 NC 在 V4 区域显著激活，这就是人类 V4 区参与轮廓整合的强有力证据。间接地说明了显著轮廓在形成轮廓感知时自下而上的信息处理模式（即刺激驱动）使得信息流从低级

到较高级是加强的，向上汇集在 V4 完成轮廓的整合感知；而 NC 情况下，观察者的初级视觉皮层则表现为增强，这可能与 V4 向下的反馈调节有关。

4.3.2.2 UOW（不显著轮廓被试判断错误）与 NC（无轮廓）的差异

通过将 UOW 与 NC 相减做 T-test，发现两者出现显著性差异，主要集中在上/中额叶（Superior/middleFrontal Gyrus，即 BA8/BA10）以及其他与视觉不相关的位置（ $p=0.05$ ， $T=2.0547$ ， $df=2816$ ），两种情况在研究范围内不存在显著性差异，这与我们预期的结果相一致，原因是尽管两种刺激本身存在差异，但是 UOW 和 NC 两种条件在认知意识上是等效的，受试者都没有完成轮廓视觉特征检测整合。

4.3.2.3 UOR（不显著轮廓观察者判断正确）与 OB（显著轮廓）的差异

UOR-OB 我们得到如图 4-4 的结果

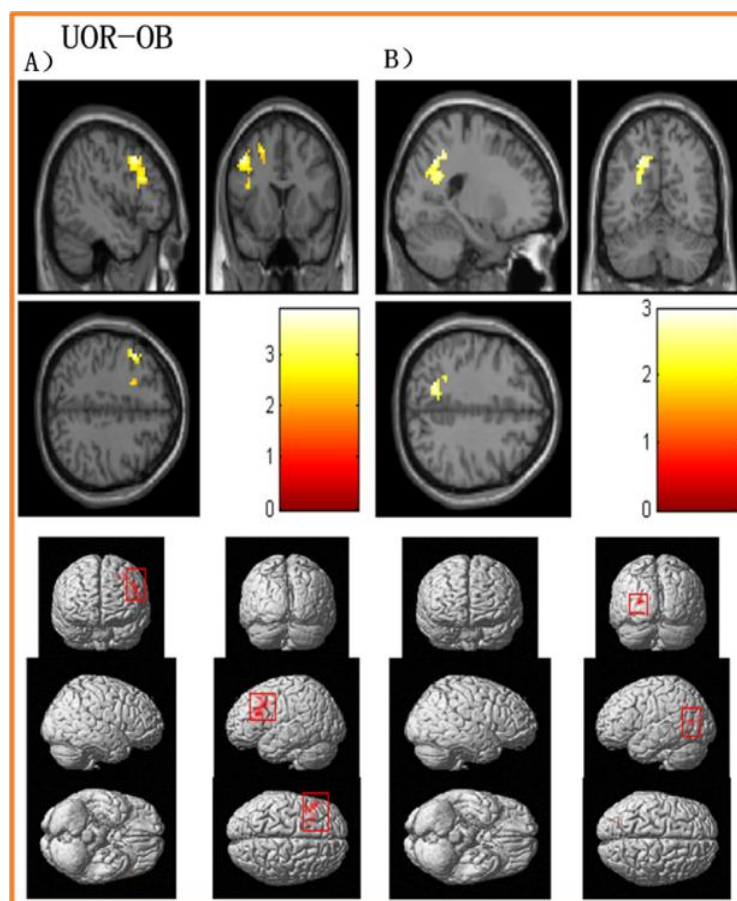


图 4-4 UOR-OB，即被试判断正确情况下的不显著轮廓激活脑区减去显著轮廓激活脑区差异图
（AlphaSim correct, $p=0.05$, cluster size=10）

两种情况都是被试完成了检测和整合，顺利完成了认知感知，但是不同的是

轮廓感知的任务难度不同。行为实验部分，我们知道所有被试在显著轮廓上的表现都达到了 90% 以上的正确率，被试对于不显著轮廓判断正确情况下的反应时间也较显著轮廓要长一些，眼动轨迹图也显示出在找到确定轮廓前，被试进行了全局检索。

两种情况下其脑区激活存在显著性差异，主要出现在了左枕中回，左侧额中回，楔前叶。集中在 BA6、8 以及 BA18 和 BA19 区组成的视觉联合皮层，还有与记忆有关的 BA30，其中 BA8 负责眼球的运动，这可能是与不显著轮廓下的前期探索活动有关。

我们的研究结果与 Bogdan 等人研究的 LG 和 GO 轮廓刺激发现不同，他们得到 LG 和 GO 两者脑区上不存在差异，仅仅是时间上的差异。我们则发现两种情况在 FEF（Frontal Eye Fields）存在差异，FEF 在控制视觉注意和眼动方面扮演重要的角色。受试者自主的进行 top-down 控制轮廓搜索时，FEF 很可能是其反馈主要来源。或者是自下而上汇集到 FEF 形成感知。但是无论是腹侧流还是背侧流的中间环节都没有发现显著性差异。说明两者的信息流处理机制是一致的，只是在眼动控制和高级意识上存在差异。

4.3.2.4 UOWF（不显著判断错误接受反馈后）与 UOW（不显著判断错误）的差异

当轮廓不显著时，其受试者判断正确率较低，当被试判断错误时，我们给被试一个快速的提示，持续 45ms，那么反馈后（top-down）和反馈前（bottom-up），与轮廓检测和整合相关的脑区神经元是如何调节的呢？如图 4-5，我们得到不显著判断错误接受反馈后与不显著判断错误的脑区激活差异。

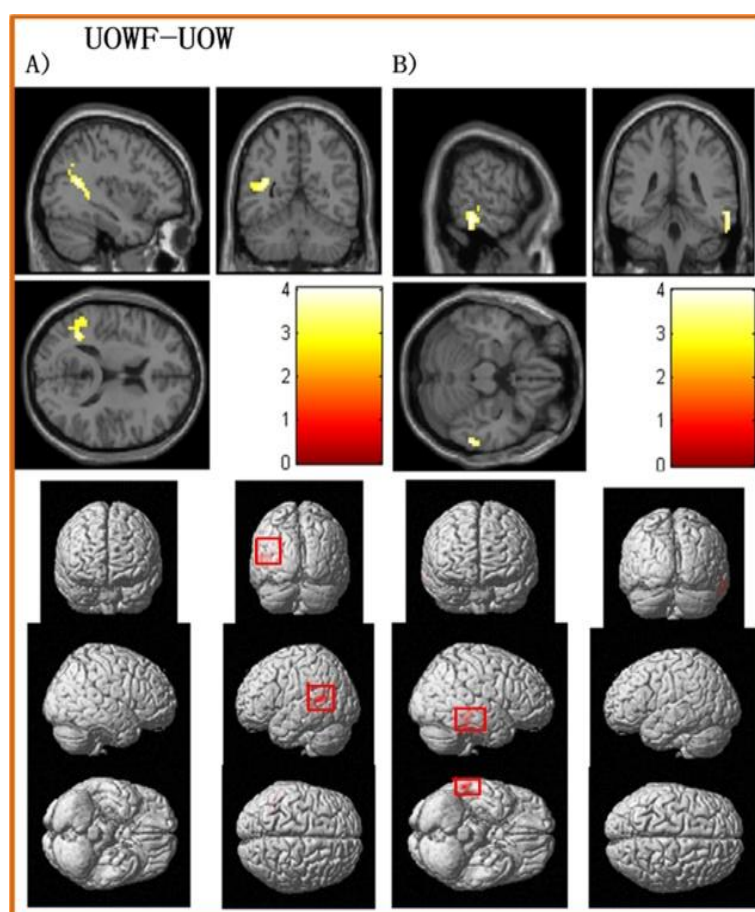


图 4-5 UOWF-UOW，即不显著轮廓呈现被试判断错误接受反馈后减去不显著轮廓判断错误情况在颞叶脑区的差异（AlphaSim correct, $p=0.05$, cluster size=10）

UOWF 和 UOW 的差别，就是多了自上而下的反馈，我们进行分析得到，接受反馈信息后的受试者大脑的激活区强度加强的地方有大脑右侧 BA20(颞下后回部分 TEO)、BA22(V4)等，如图 4-5（AB）。同时，在图 4-5（CD）中，接收反馈后，低级视觉皮层相比接受反馈前出现了一个抑制。

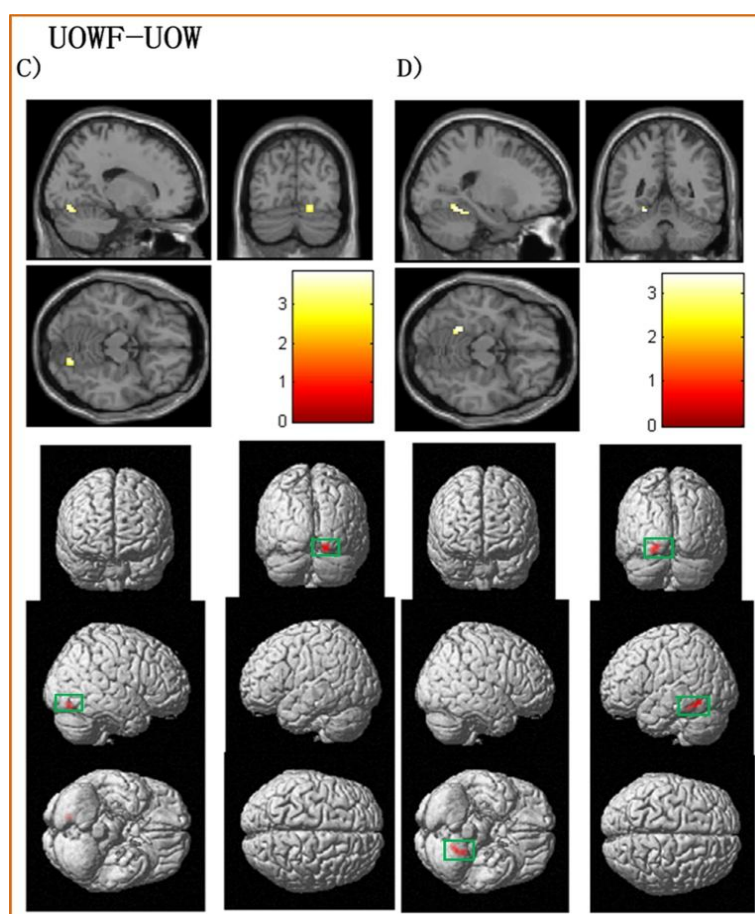


图 4-5 UOWF-UOW，即不显著判断错误接受反馈后减去不显著判断错误在早期视觉皮层脑区的差异（AlphaSim correct, $p=0.05$, cluster size=10）

我们知道，腹侧通路是从低级视觉皮层，进入颞叶下部（BA20、BA37）负责物体、文字、面孔的识别，即“是什么”。背侧通路 - 则延伸至中央后回（BA39、BA40）感知空间关系，即“在哪里”。

上述结果图 4-5（A）是 BA20 区被加强，所以表明人类在轮廓检测时，倘若任务难度较高，需要接收来自外界的提示信息或者以往的先验知识时，其视觉腹侧通路的高级视觉皮层 IT 区神经元活动便会加强。不仅如此，背侧通路的延伸区（BA39）中央后回也会出现活动加强。当被试接受短暂的 45ms 提示后，受试者再进行轮廓检测整合任务，便相当于是同时存在来自 top-down 的反馈信息（包括 what 通道的有轮廓信息和 where 通道的轮廓在什么空间位置信息）以及 bottom-up 机制下来自刺激本身的信息，而受试者接受反馈前则是仅仅凭借刺激本身的特征来完成感知任务的，所以两者最后得到的显著区域便是轮廓检测和整合相关的 top-down 信息通路，以及其神经元的调节情况。我们的结果表明，轮廓检测和整合相关的中间视觉皮层信息可能是作用于 BA20（IT）的。在腹侧通路和背侧通路

的反馈是否都是易化增强的呢？我们发现，尽管在通路的顶层，脑区的活动都是增强的，但是在视觉联合皮层 BA18 和 BA19 还是出现了不同程度的活动抑制。

同时，如 4-5 (E) 图所示，我们还在中央前回前端 (BA6) 发现一个较大程度的激活强度降低，也就是 FEF 区所在的地方，说明当被试接受外界提示而进行反馈时，负责视觉注意的 FEF 激活强度瞬间下降，同时，反馈信息从额叶回流作用于颞下回 (BA20) 及 V4 等区域，V4 (BA22 区和 39 区) 作为由低级视觉皮层通向额叶的门户，占据了重要的位置，当然也扮演着重要的角色。

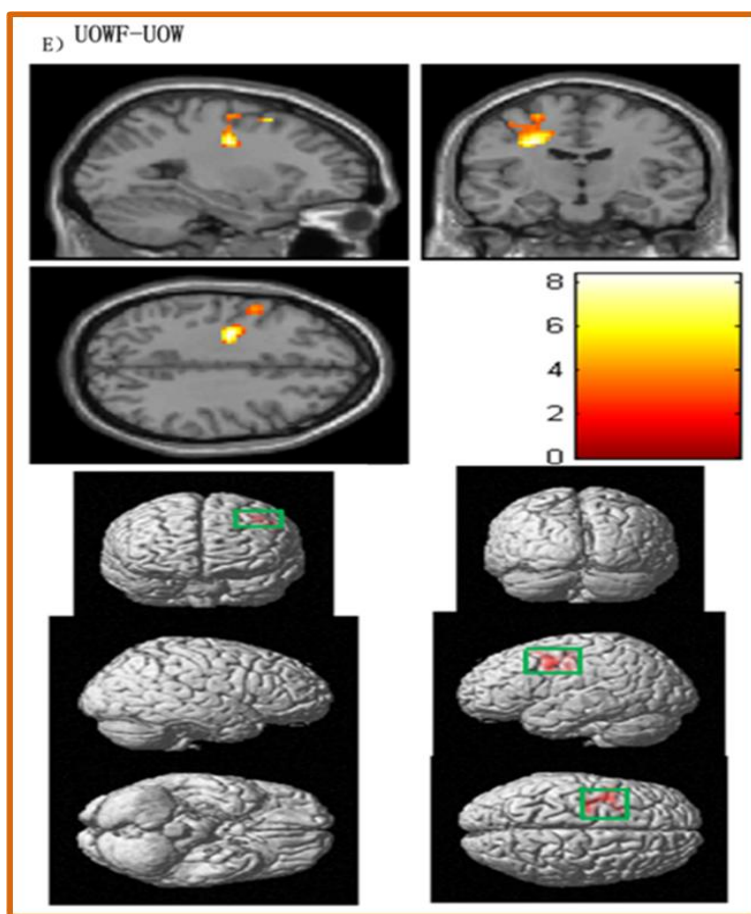


图 4-5 UOWF-UOW，即不显著轮廓刺激呈现判断错误接受反馈后减去不显著判断错误在前额眼动区的差异 (AlphaSim correct, $p=0.05$, cluster size=10)

4.3.2.5 UOWF (不显著判断错误接受反馈后) 与 OB (显著轮廓) 的差异

我们知道 UOWF 这种情况可以理解为我们设定的记忆驱动或者先验知识驱动的自上而下的反馈，而 OB 则是典型的基于轮廓刺激本身的自下而上驱动的任务范畴，这两种进行对比，我们可以如第四种对比一样来验证这两种认知机理的脑区差异和信息传输机制的流向。结果如图 4-6 所示：

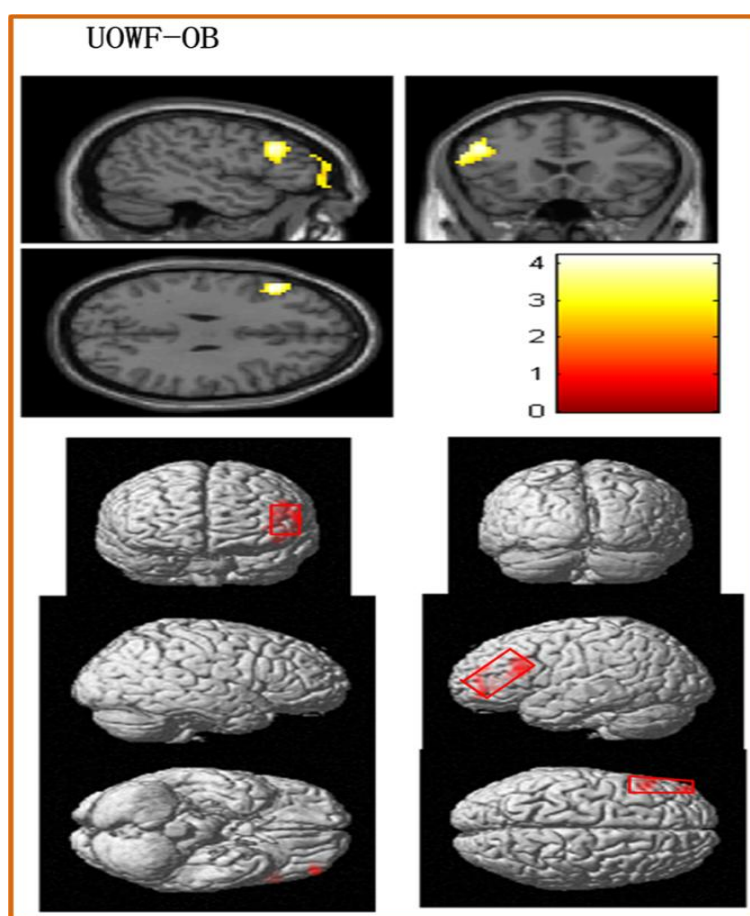


图 4-6 UOWF-OB，即不显著轮廓判断错误接受反馈后减去显著轮廓判断正确在前额叶脑区的差异（AlphaSim correct, $p=0.05$, cluster size=10）

不显著轮廓判断错误接受反馈后减去显著判断正确的脑区激活差异，其峰值出现在 MNI 坐标（-48 21 30）处，表现为正激活，位于左前额叶皮层区（Left Prefrontal Lobe）。

在 2005 年 Neuron 发表的文章“Searching for “the Top” in Top-Down Control”上讲到了 PFC（The prefrontal cortex）在许多理论模型中被认为是认知控制过程中自上而下的主要来源，其中一种进程就是工作记忆，这是一套必要的维护操作和信息处理在线服务行为，许多人类和动物的神经生理学研究发现 PFC 是活动在工作记忆（working Memory, WM）的阶段^[36]，很多的辩论都在研究这种信号的功能性贡献到底是什么，研究表明，PFC 可以根据当前任务调控引导自下而上驱动区域的自上而下的信号，从而有助于认知控制^[37, 38]。在此之前没有可能的测试和优化的模型揭露证实了 PFC 在自上而下的信号感知和记忆控制^[39, 40]。而由我们研究的第二种和第五种比较情况，进一步证实，记忆甚至是期望或者先验知识的信息

处理阶段都会有 PFC 区的活动增强，这也与 Buschman 在《科学》杂志上面发表的结果相一致^[41]。

综上，我们整理了基于条件变量之间的脑区激活差异，归纳到表 4-1 中，分别作了变量间的差值，比如 OB-NC，即显著轮廓情况下比无轮廓情况有明显激活的脑区。同时还列出了激活脑区团块的峰峰值坐标，所处的解剖结构位置，对应的布罗德曼分区，以及是正激活还是负激活。如下所示

表 4-1 五种情况间的脑区差异表

Conditions	Peak MNI coordinate			Anatomical structure	BA	+/-
OB-NC	-36	-54	18	Left superior temporal gyrus	BA22	+
	-12	-72	12	Right occipital gyrus	BA17、18	-
UOW-NC	---	---	---	---	---	---
UOR-OB	-45	15	42	Left middle frontal gyrus	BA6、8	+
	-24	-72	12	Left Middle Occipital Gyrus	BA18、19	+
UOWF-UOW	63	-39	-21	Right Inferior Temporal Gyrus	BA20	+
	-36	-54	15	Left Superior temporal gyrus	BA22	+
	-27	-12	36	Left middle frontal gyrus	BA6	-
	18	-75	-15	Right occipital lobe lingual gyrus/ Lingual Gyrus	BA19	-
	-21	-48	-15	left occipital lobe	BA19	-
UOWF-OB	42	-69	3	Right middle Temporal Gyrus and occipital	BA19	+
	-48	-60	12	Leftsuperior/middle temporal gyrus	BA22、39、19	+
	-48	21	30	Left Prefrontal Lobe	BA10.11 .9	+

结果：Top-down 控制与注意在复杂场景的轮廓提取中具有不可或缺的作用，它的影响主要是中间皮层的增强以及初级皮层的抑制。任务相对简单，刺激特征明显时，仅仅凭借刺激驱动的前馈完成整合。空间注意可以通过这种与轮廓刺激相关的行为任务，分析脑区特定部分，随之增强的响应区域，包括 V4, MT 区(the medial temporal area) 和 IT 区 (the inferior temporal area), 还与刺激相关的错误的信息之间的竞争的选择相关。早期的研究表明，高级视觉皮层比初级皮层更容易受注意的影响，而影响的程度与刺激配置还有任务属性相关。

Top-down 可以影响空间中心位置，还可以改变神经元启动率，逆向流携带的信息比来自于注意点的信息更多。Top-down 影响皮层的早期，改变他们的信息特性，这不仅仅是目标控制，也包括改变神经元以携带更多的与任务相关的信息，

对目标的期望成分有选择性，且在眼动情况下保持一个稳定的表达，与任务相关的信息一部分改变神经元的调优，另一部分改变神经元群之间的结构，不同的 **top-down** 出现在不同的区，且与该区的功能特性相关。

我们的研究还发现，**V1** 缺乏注意的效应而 **V4** 和 **MT** 有较强的效应，表明沿着视觉通道有渐强的注意影响。

我们还对 **bottom-up** 和 **top-down** 模型进行了总体评估，发现感受野中的目标竞争计算资源，有且仅有一个获胜方，要么是显著性上（**bottom-up** 注意）要么是行为相关性上（**top-down** 控制）。而行为相关性上的控制在我们的实验中表现为腹侧通道调制较背侧通道占优，这可能与我们的提示信息位置重要性较高有关，表现在反馈上，则背侧 **where** 通道没有太大差异。我们的实验发现自上而下的影响脑区不一致，这就是说不同的任务可能会选择不同的再进入调节内部连接，从而达到一个合理的预期，而 **V1** 水平连接存在的原因就是允许在进入网络连接的存在。而受试者在不同条件变量下的时间进程，则是网络调节选择所需的时间。这种自上而下和自下而上结合的模式提供了轮廓整合和显著性的机制，与前人的研究相一致。

4.4 本章小结

我们的结果发现，当轮廓特征比较显著，任务相对简单的时候，轮廓感知在高级视觉皮层的表现仅仅在 **V4** 便可完成。而当任务难度较高但是最终人们根据任务期望，先验知识等判断正确时，则负责眼动和视觉注意的 **FEF** 区可能参与做出抉择前的轮廓探索。眼动实验部分，我们已经得到在三种不同刺激时，其眼动轨迹是存在显著性的差异的，因此，在磁共振实验中得到与眼动控制相关的 **FEF** 区，也可能不是与 **top-down** 控制相关的脑区，而仅仅是在轮廓检测的前期探索扮演着重要角色。还有一种情况，人们在复杂场景中无法识别目标轮廓时，外界给予一定提示后，被人们普遍认为与认知控制相关的 **PFC** 激活，同时 **top-down** 控制机制的反馈信息流则抑制 **FEF**，回流作用于 **IT** 区后间接调节 **V4/V2/V1** 等视觉皮层。

我们的研究支持，不同的环境刺激，不同任务难度的轮廓感知，是存在不同的脑认知机理和信息传输机制的。我们根据几种典型的 **conditions** 之间的比较，初步推测得出这样三条可能与轮廓检测相关的脑区激活及信息传输方向。如图 4-7 所示

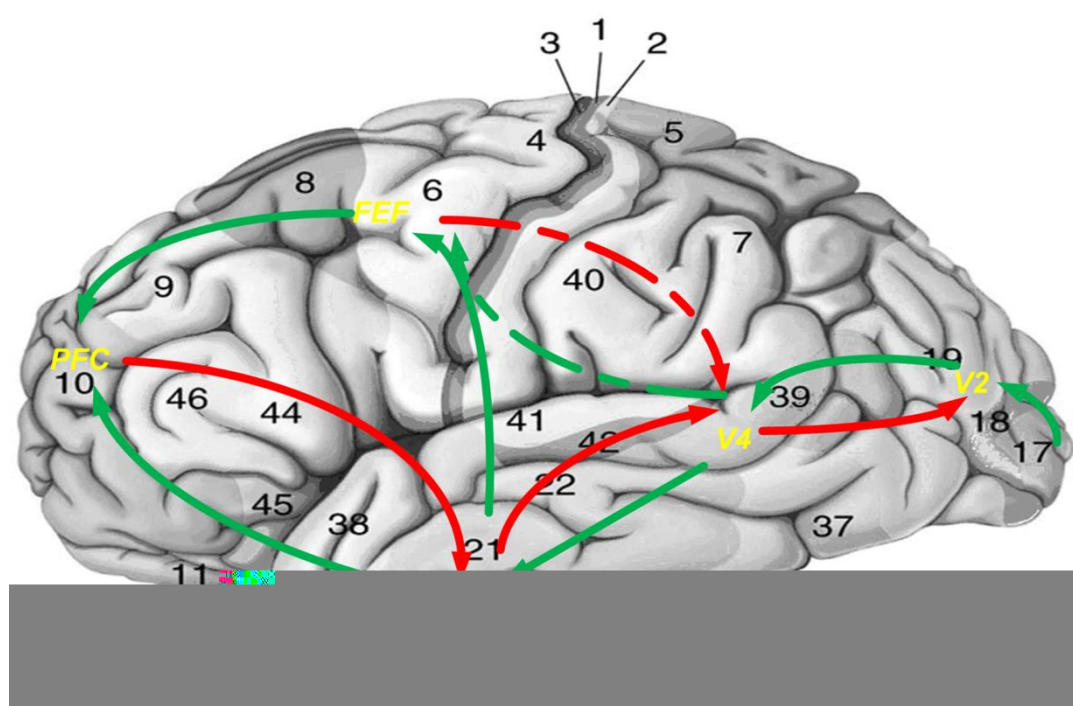


图 4-7 简化的轮廓检测和整合信息传输机制图（蓝色箭头为正反馈、红色箭头为负反馈）

1) $V1 \Leftrightarrow V2 \Leftrightarrow V4$ (简单轮廓的整合), 轮廓整合任务较简单时, 信息流向可能是自 $V1$ 进行前期提取, 经过 $V2$, 将来自刺激的信息汇集到 $V4$ 完成整合。

2) $V1 \Leftrightarrow V2 \Leftrightarrow V4 \Leftrightarrow FEF$ 。我们认为这条信息流是轮廓整合任务难度较大时, 受试者需借助先验知识, 或者需依赖更多的眼动控制进行轮廓检测时, 信息传输是在 1) 的基础上, 继续汇集到 FEF 区, 同时也存在着向下的信息互动, 来完成轮廓整合, 但是这条通路并非必须的, 在接受外界明确提示后可能被抑制, 因此在图中用虚线表示。

3) 仅凭借先验知识的第 2) 条信息流无法完成感知时, 人类的脑机制就会由外界的提示或者回忆而形成反馈信息流: $PFC \Leftrightarrow IT \Leftrightarrow V4 \Leftrightarrow V2$ 。我们的研究发现, 外界提示作用的皮层位于 PFC 区, 向下作用于颞叶, 最终反馈信息调节枕叶视觉感知相关的脑区, 通过正反馈信息流协同合作完成轮廓整合。

第五章 全文总结与展望

5.1 全文总结

本文以物体的识别和感知为研究背景，主要采用了眼动和任务态磁共振成像结合的方法，对轮廓检测和整合的脑认知机制和信息传输机理这一分支进行研究。

行为眼动实验是在客观分析了人类视觉系统对轮廓特征检测与整合的影响因素和轮廓与背景相互作用的基础上设计出来的，根据人类视觉系统的神经元朝向选择性以及局部轮廓间的共线组织特点设计了刺激图片，又由格式塔原则的接近性原则，拉大局部元素间的距离，使任务难度增加，对轮廓情况进行细分。结果表明：人类在进行轮廓的检测和整合时，不仅受有无轮廓的差异影响，还受任务难度，先验知识的影响。当轮廓检测任务简单时，受试者的感知过程也相应的简单，前期的检测过程迅速，感知意识能够快速完成。当任务难度较高时，轮廓检测整合前期的探索轮廓过程变得更为复杂，必要的话还需要先验知识或者外界提示，才能够完成。

磁共振实验则分析了轮廓感知皮层间的 bottom-up 和 top-down 的大脑机制，实验中随机呈现有显著轮廓，不显著轮廓和无轮廓的刺激，同时，通过加入反馈提示的方法，直观的得到 top-down 机制在皮层上的表现。我们的结果发现，受试者观看任务简单，轮廓明显的刺激时，大脑皮层的 V4 区显著激活，尽管先前的研究发现初级视觉皮层就可以对简单的共线轮廓形成感知，但我们的实验倾向于 V4 在轮廓意识形成上的作用。

然而，受试者在观看任务难度较高的轮廓刺激时，自己独立完成轮廓整合和接受反馈后完成轮廓感知，这两者是存在认知机制差异的，差异表现在激活脑区位置和从高级视觉皮层往低级视觉区域的反馈信息流向两个方面。受试者观看轮廓刺激不借助外界的提示时，top-down 的 top 或者说 bottom-up 的 up 是位于 FEF 的，同时 top-down 的 down 或者说 bottom-up 的 bottom 是位于 V4 区域的，这由 UOR-OB 两者的差异得到。当屏幕呈现的是不显著轮廓刺激，被试判断错误，接受我们提供的提示后，大脑皮层的 IT、V4 区域有了显著的增强，同时相比于简单的显著轮廓情况，top-down 调节机制在 PFC 有显著强度的激活。

5.2 后续工作展望

轮廓检测和整合的算法研究近几年发展迅速，生理皮层上的研究也渐渐明朗，

但是在脑认知机制和信息传输导向上，却一直缺少成熟的理论来指导人类视觉机制算法的研究，在本文研究工作的基础上，仍有以下方向值得进一步深究和探讨：

轮廓检测和整合在人类实际认知过程中的情况分类是千变万化的，任务，期望，上下文影响，场景等可变因素众多，每一种因素又可以演变出很多其他的变量，所以给轮廓检测整合的脑认知机制研究带来了巨大的困难。同时，我们的研究设计也存在着变量因素覆盖不全，设计流程有待完善的问题，同时在数据处理中也没有很好的构建任务态的磁共振激活脑区的因果分析，在信息流的动态互连上欠缺足够的说服力，以后的研究方向可以在这些个方面多加考虑，同时在验证轮廓检测和整合的脑认知机理时，可以尝试采用 TMS 的方法，相信能够得到更多的可靠性结果，验证得到更加具有说服力的证据。此外，数据分析结果上面，还可以尝试用格兰杰因果(Granger Causality, GC)等方法来分析各脑区间的功能连接和有效连接，更加清楚的得到提示轮廓检测与轮廓整合时相关脑区间的有向信息传输机制^[42]。

相信人们在生理、行为、认知机制等层面的共同研究下，人类视觉机制在处理目标物体轮廓的过程会渐渐的明朗，更好的促进人类在计算科学，认知科学的发展。

由于本人在认知专业领域的时间和水平有限，研究课题论文中难免会出现一些不可预知的问题，请评审和读者能够多加批评和指正，为人类的脑科学发展共同努力。

致 谢

在攻读硕士学位期间，首先衷心感谢我的导师颜红梅教授，是她将我领进了认知领域的大门，在我迷茫的时候给我指引方向，帮忙解决了我在课题开题，技术攻坚，数据处理等多方面遇到的问题，她在科研技术水平上严谨的态度，必将影响我以后的工作和学习。其次，也十分感谢实验室李永杰，王玲，陈科等老师在论文内容和实验结果处理的建议。

我还要感谢跟我一起设计和完成实验的师兄师姐，我的论文能够顺利完成离不开陈诚师姐对我耐心的指导，也离不开陈伟强师弟对我实验任务的分担。在此，对他们表示最诚挚的谢意和最美好的祝福，祝愿他们能够顺利完成学业。

感谢所有参加我实验的校友被试们。也感谢从未谋面的吉尔伯特，李武教授，构成本论文核心部分的想法和观点是他们在该领域的科研成果的基础上，影响下演变而成的。

最后，我要对自己的爸爸妈妈表示感谢，多年来，他们一直陪伴我，他们在生活，学习上的鼓励和支持是此论文课题付诸实现的关键。

参考文献

- [1] Graf T, Therrien R. A method to discretize non - planar fractures for 3D subsurface flow and transport simulations [J]. International journal for numerical methods in fluids, 2008, 56(11): 2069-2090.
- [2] Geisler W S, Perry J S, Super B J, et al. Edge co-occurrence in natural images predicts contour grouping performance [J]. Vision Research, 2001, 41(6): 711-724.
- [3] Palmer S, Rock I. Rethinking perceptual organization: The role of uniform connectedness [J]. Psychonomic Bulletin & Review, 1994, 1(1): 29-55.
- [4] Hansen T, Sepp W, Neumann H. Recurrent long-range interactions in early vision [M]. Emergent neural computational architectures based on neuroscience. Springer. 2001: 127-138.
- [5] Chen M, Yan Y, Gong X, et al. Incremental integration of global contours through interplay between visual cortical areas [J]. Neuron, 2014, 82(3): 682-694.
- [6] Li W, Thier P, Wehrhahn C. Contextual influence on orientation discrimination of humans and responses of neurons in V1 of alert monkeys [J]. Journal of Neurophysiology, 2000, 83(2): 941-954.
- [7] Murray S O, Wojciulik E. Attention increases neural selectivity in the human lateral occipital complex [J]. Nature neuroscience, 2004, 7(1): 70-74.
- [8] Gilbert C D, Sigman M. Brain states: top-down influences in sensory processing [J]. Neuron, 2007, 54(5): 677-696.
- [9] Oram M W, Földiák P, Perrett D I, et al. The Ideal Homunculus': decoding neural population signals [J]. Trends in neurosciences, 1998, 21(6): 259-265.
- [10] Gilbert C D, Li W. Top-down influences on visual processing [J]. Nature Reviews Neuroscience, 2013, 14(5): 350-363.
- [11] Hubel D H, Wiesel T N. Receptive fields and functional architecture of monkey striate cortex [J]. Journal of Physiology, 1968, 195(1): 215-243.
- [12] Li C Y. Integration fields beyond the classical receptive field: Organization and functional properties [J]. Physiology, 1996, 11(4): 181-186.
- [13] Huang R Y M, Feng X. Studies on solvent evaporation and polymer precipitation pertinent to the formation of asymmetric polyetherimide membranes [J]. Journal of Applied Polymer Science, 1995, 57(5): 613-621.
- [14] Chen G, Dan Y, Li C Y. Stimulation of non - classical receptive field enhances orientation

- p selectivity in the cat [J].
- Journal of Physiology*
- , 2005, 564(1): 233-243.
- [15] Stanley D A, Nava R. fMRI activation in response to illusory contours and salient regions in the human lateral occipital complex [J]. *Neuron*, 2003, 37(2): 323-331.
- [16] Wu L, Gilbert C D. Global contour saliency and local colinear interactions [J]. *Journal of Neurophysiology*, 2002, 88(5): 2846-2856.
- [17] Wu L, Valentin P C, Gilbert C D. Learning to link visual contours [J]. *Neuron*, 2008, 57(3): 442-451.
- [18] Shpaner M, Molholm S, Forde E, et al. Disambiguating the roles of area V1 and the lateral occipital complex (LOC) in contour integration [J]. *Neuroimage*, 2012, 69(4): 146–156.
- [19] Altmann C F, Bülthoff H H, Zoe K. Perceptual organization of local elements into global shapes in the human visual cortex [J]. *Curr Biol*, 2003, 13(4): 342–349.
- [20] Nathalie V H, Nadine S, Frouke H, et al. The role of eye movements in a contour detection task [J]. *Journal of Vision*, 2013, 13(14): 1-19.
- [21] Mcmanus J N J, Wu L, Gilbert C D. Adaptive shape processing in primary visual cortex [J]. *Proc Natl Acad Sci U S A*, 2011, 108(24): 9739-9746.
- [22] Bushnell B N, Harding P J, Yoshito K, et al. Partial occlusion modulates contour-based shape encoding in primate area V4 [J]. *Journal of Neuroscience the Official Journal of the Society for Neuroscience*, 2011, 31(31): 4012-4024.
- [23] Mcmanus J N, Li W, Gilbert C D. Adaptive shape processing in primary visual cortex [J]. *Proceedings of the National Academy of Sciences*, 2011, 108(24): 9739-9746.
- [24] Brigell M, Uhlarik J, Goldhorn P. Contextual influence on judgments of linear extent [J]. *Journal of Experimental Psychology: Human Perception and Performance*, 1977, 3(1): 105.
- [25] Ullman S. Object recognition and segmentation by a fragment-based hierarchy [J]. *Trends in cognitive sciences*, 2007, 11(2): 58-64.
- [26] Gölçü D, Gilbert C D. Perceptual learning of object shape [J]. *The Journal of Neuroscience*, 2009, 29(43): 13621-13629.
- [27] Cohen M R, Maunsell J H. Attention improves performance primarily by reducing interneuronal correlations [J]. *Nature neuroscience*, 2009, 12(12): 1594-1600.
- [28] Miller B T, D'Esposito M. Searching for “the top” in top-down control [J]. *Neuron*, 2005, 48(4): 535-538.
- [29] Zeki S. A vision of the brain [J]. *Tex Dent J*, 16(12):533-534. 1993,
- [30] Sugase Y, Yamane S, Ueno S, et al. Global and fine information coded by single neurons in the temporal visual cortex [J]. *Nature*, 1999, 400(6747): 869-873.

-
- [31] Middlebrooks P G, Sommer M A. Neuronal correlates of metacognition in primate frontal cortex [J]. *Neuron*, 2012, 75(3): 517-530.
- [32] Kobatake E, Tanaka K. Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex [J]. *Journal of neurophysiology*, 1994, 71(3): 856-867.
- [33] Young M P, Yamane S. Sparse population coding of faces in the inferotemporal cortex [J]. *Science*, 1992, 256(5061): 1327-1331.
- [34] Mijović B, De Vos M, Vanderperren K, et al. The dynamics of contour integration: A simultaneous EEG-fMRI study [J]. *NeuroImage*, 2014, 88(10-21).
- [35] Kok P, De Lange F P. Shape perception simultaneously up-and downregulates neural activity in the primary visual cortex [J]. *Curr Biol*, 2014, 24(13): 1531-1535.
- [36] Heide W, Blankenburg M, Zimmermann E, et al. Cortical control of double - step saccades: implications for spatial orientation [J]. *Annals of neurology*, 1995, 38(5): 739-748.
- [37] Paus T, Collins D, Evans A, et al. Maturation of white matter in the human brain: a review of magnetic resonance studies [J]. *Brain research bulletin*, 2001, 54(3): 255-266.
- [38] Heide W, Kömpf D. Combined deficits of saccades and visuo-spatial orientation after cortical lesions [J]. *Experimental Brain Research*, 1998, 123(1-2): 164-171.
- [39] Thulborn K R, Martin C, Voyvodic J T. Functional MR imaging using a visually guided saccade paradigm for comparing activation patterns in patients with probable Alzheimer's disease and in cognitively able elderly volunteers [J]. *American Journal of Neuroradiology*, 2000, 21(3): 524-531.
- [40] Cutsuridis V. A cognitive model of saliency, attention, and picture scanning [J]. *Cognitive Computation*, 2009, 1(4): 292-299.
- [41] Buschman T J, Miller E K. Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices [J]. *Science*, 2007, 315(5820): 1860-1862.
- [42] Hiemstra C, Jones J D. Testing for linear and nonlinear Granger causality in the stock price - volume relation [J]. *The Journal of Finance*, 1994, 49(5): 1639-1664.

攻读硕士学位期间取得的成果

- [1] Cheng Chen,Xianghui Chen.Contextual influence on the tilt after-effect in foveal and para-foveal vision[J]. Neuroscience Bulletin,June 2015, Volume 31,Issue 3, pp 307-316