

方向选择性模型 及其在图像处理中的应用

任远 著

清华大学出版社



方向选择性模型及其在图像 处理中的应用

任 远



清华大学出版社

内 容 简 介

本书论述了视皮层中简单细胞的建模及其在图像理解中的几点应用,包含了人工智能、神经科学及认知科学等领域的内容,属于交叉学科的研究成果总结。本书在给简单细胞的方向选择性现象建立数学模型后,逐步将其用于图像的方向检测与表征、高层语义的识别与理解、场景的三维信息获取以及几何视错觉的成因分析。书中除了用严格的数学理论证明模型的正确性外,还用大量的数值仿真实验、真实图像实验以及统计实验验证算法的可行性。

本书为计算机视觉、神经视觉及认知科学的跨学科研究提供了一个新的视角,具有一定的学术性,可以为人工智能及认知科学研究者提供参考。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

方向选择性模型及其在图像处理中的应用/任远著. —北京:清华大学出版社, 2015

ISBN 978-7-302-37329-2

I. ①方… II. ①任… III. ①计算机网络—网络模型—研究 IV. ①TP393.02

中国版本图书馆 CIP 数据核字(2014)第 160794 号

责任编辑:秦 甲

封面设计:杨玉兰

责任校对:周剑云

责任印制:宋 林

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址: 北京清华大学学研大厦 A 座 邮 编: 100084

社 总 机: 010-62770175

邮 购: 010-62786544

投稿与读者服务: 010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈: 010-62772015, zhiliang@tup.tsinghua.edu.cn

课件下载: <http://www.tup.com.cn>, 010-62791865

印 装 者: 三河市中晟雅豪印务有限公司

经 销: 全国新华书店

开 本: 148mm×210mm 印 张: 5.25 字 数: 112 千字

版 次: 2015 年 4 月第 1 版 印 次: 2015 年 4 月第 1 次印刷

定 价: 25.00 元

产品编号: 059154-01

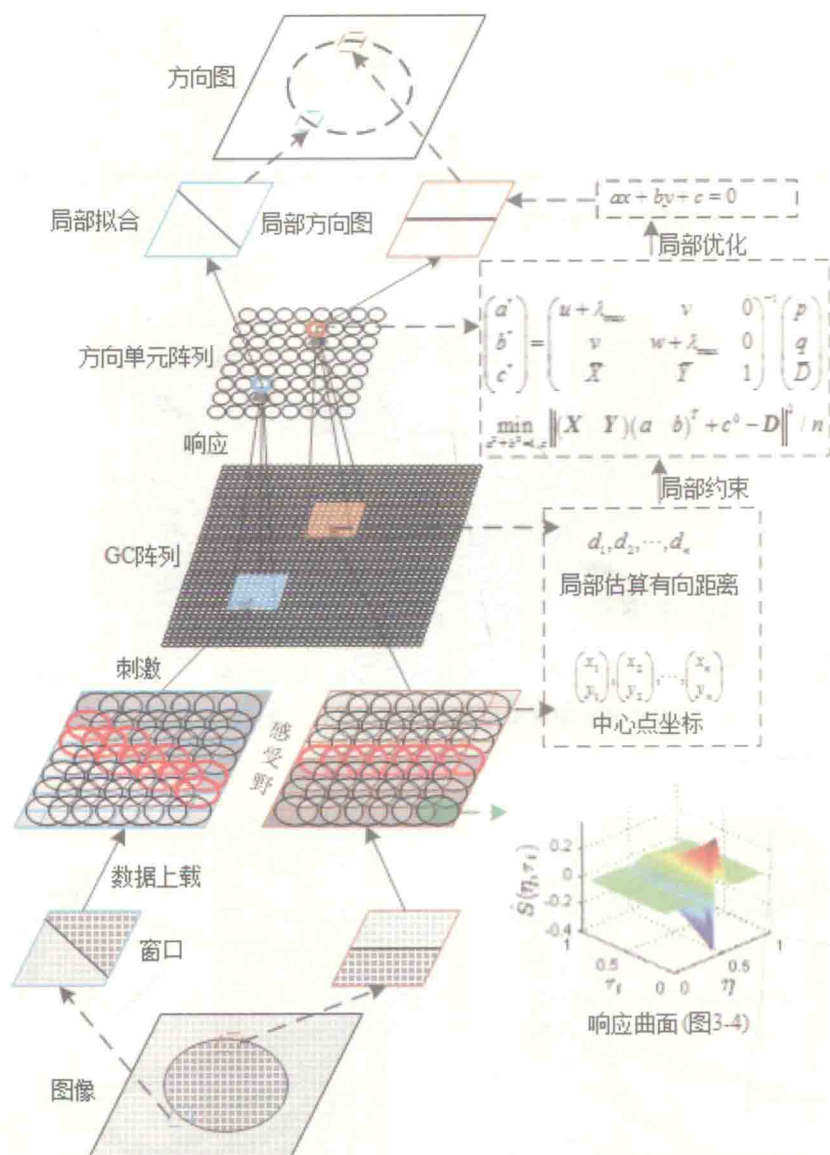


图 6-1 方向检测计算流程图，图中变量的意义参见4.2节(p.37) 和4.3节(p.40)



图 7-3 检测到的方向(蓝色)和实际方向(黑色)存在偏差,使得干扰线(红色)和被检线段(黑色)的锐角夹角变大



图 7-5 Zöllner 错觉的解释^[123]。局部方向图中,检测出的方向用红色粗线突出,蓝色框表明和真实方向比较偏差较小,绿色框表明偏差较大

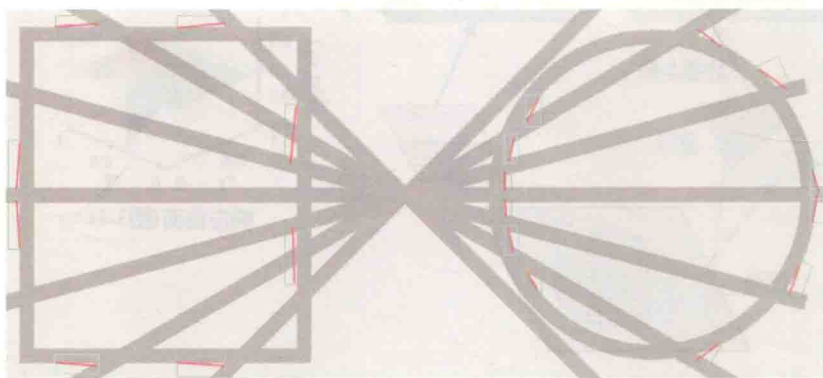


图 7-8 Orbison 错觉的解释^[123]

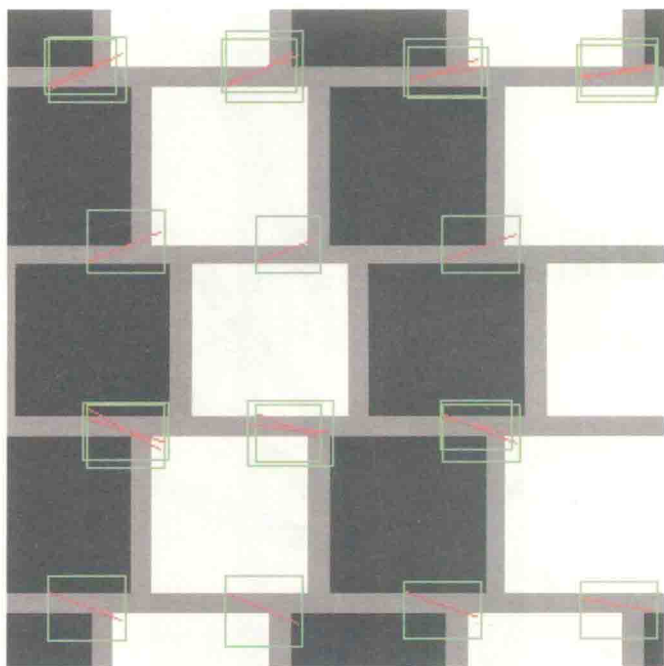


图 7-10 Café Wall 错觉的解释^[123]

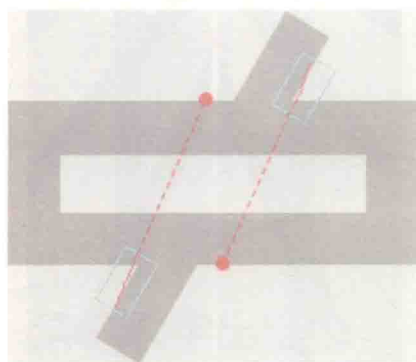


图 7-12 Poggendorff 错觉的解释

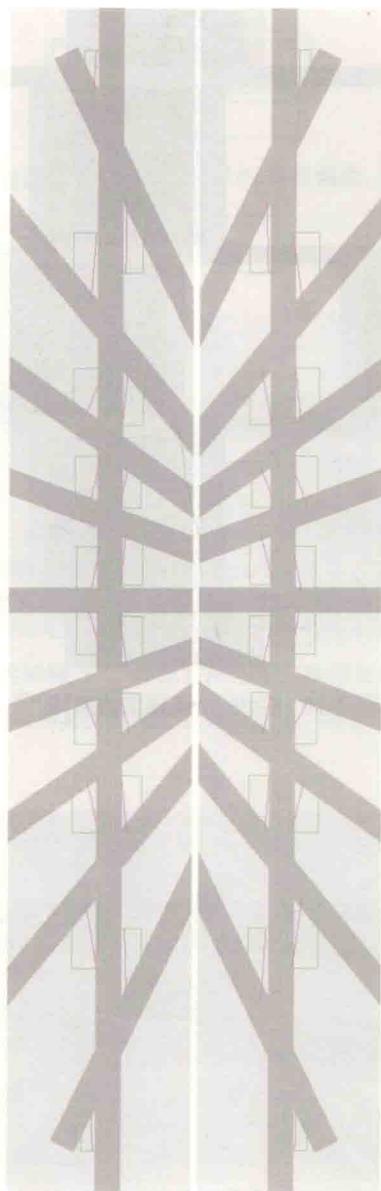


图 7-14 Hering 错觉的解释^[124]

前言

图像理解是计算机视觉领域中的一个核心问题。对图像理解而言,关键的一步是提取图像中构成场景语义结构的特征并将其表征为高层任务容易利用的形式。一种传统的特征提取方法为先寻找边缘,然后形成直线或组合轮廓线。这些特征为图像高层语义的理解提供了基础。

传统的特征提取方法在一定程度上取得了很大进展,目前有大量针对边缘检测、直线检测以及轮廓检测的算法。然而,由于这些算法本质上都是依赖边缘检测的,所以都存在一些显著的问题,如对边缘检测的结果依赖性过高及需要人工设定的参数过多等。一般来说,这些算法都是单纯从计算的角度尝试定义问题,而往往忽略问题的本质,即生物视觉系统如何完成边缘—直线—轮廓检测的生理机制,所以难以从根本上解决这些问题,更难以解决基于这些结果的更高层的图像理解任务。

对人眼而言,视觉场景中最显著的特征就是不同颜色形成的边缘,以及边缘形成的直线、曲线和更复杂的形状。本质上,我们看到的直线及曲线都可以看作相同或不同方向的短直线的分段组合。因而,本书将方向(短直线)看作对图像理解最有意义的一种特征。显然,方向这一初级特征高于像素颜色、灰度等底层特征,又低于轮廓等高层语义特征。作为基于神经科学的人工智能和认知科学的跨学科研究,本书尝试利用生物视觉系统检测方向的神经机制,设



计一个仿生学的方法,从根源上实现特征提取及表征这一基本问题的解决,进而为更复杂的任务提供生理基础。

在视觉神经科学中,简单细胞的方向选择性一直是一个热点问题。诺贝尔生理学奖获得者 David Hubel 和 Torsten Wiesel 提出过一个简单细胞的感受野模型(以下简称 Hubel-Wiesel 神经模型),该模型用几何形式的约束解释了简单细胞对条形刺激的方向选择性。一方面,该模型具有简单朴素的形式;另一方面,该模型也有不少的缺陷,因而受到了许多挑战。尽管如此,目前没有严格的证据直接支持或者否定这个经典的模型。

本书在 Hubel-Wiesel 神经模型的基础上,提出一个方向计算的双层网络模型(以下简称本书模型)。与 Hubel-Wiesel 神经模型相比,本书模型对底层神经元的限制更少,对刺激形式的要求更低,且实现了每一层计算的细节,因而灵活度更高。数值仿真实验表明,本书模型能够很好地模拟简单细胞的方向选择性。

利用方向计算模型,本书进一步提出一个图像的方向检测算法(以下简称本书算法)。在合成图像及自然图像上得到的实验数据表明,该算法能够从复杂图像中提取满意的方向图。与传统算法得到的边缘图、直线图等相比,方向图突出显著语义特征而抑制琐碎干扰信息,更接近物体的真实轮廓图,因而促进了分割及识别等高层任务更好更快地完成。此外,本书算法对参数调节的依赖更低、可并行计算程度高。

作为模型的应用之一,本书巧妙地将方向检测算法用于解释几种著名的几何视错觉。本书用定性计算的方法从局部探索错觉的产生机制,生成了与人眼观察理解相似的错觉现象,并将这种方法推广到一系列错觉的解释。作为更深层次的应用,本书最后尝试利用

单幅图像的方向图，恢复其中场景的部分三维信息，也取得了一定的成果。

本书正文部分安排如下：第1章介绍相关研究的现状与本研究的主要内容及特点；第2章介绍本书模型建立的生理基础与依据；第3章介绍模型的底层设计，讨论外侧膝状体细胞对刺激的响应函数；第4章介绍模型的顶层设计，讨论简单细胞利用传入外侧膝状体对刺激的响应确定边缘方位的几种计算方法；第5章用数值实验证明模型的正确性与可行性，确定模型中参数的优化取值，并使用仿真实验模拟形成简单细胞具有方向选择性的感受野；第6章介绍基于方向选择性模型的方向检测算法在图像表征、分割及匹配等方面的应用；第7章将方向检测用于解释一系列几何视错觉；第8章利用单幅图像的方向图部分恢复场景的三维信息；总结部分总结全书并展望后续研究课题。

作为我的第一本学术著作，本书包含了2007~2013年我在复旦大学攻读研究生期间的主要研究成果，也是我的博士学位论文的主要部分。许多人对本书的完成提供了不可或缺的帮助和启发。首先要感谢父母对我的养育之恩；感谢王晓梅师姐为本书第2章的写作提供了大量参考资料，感谢肖锦文师弟为本书第6章6.4.2节提供了实验数据；最要感谢我的导师危辉教授，从课题选择、数学建模、算法设计到实验分析各环节都凝聚着危老师的辛勤指导，是他的精心培养才使我顺利完成了学业。

本书既然是对之前研究成果的总结，自然不是终结，而是后续工作的起点。本书中悬而未决的问题与不尽理想的结果都是我正在研究的重要课题，也是我未来的研究方向。作为一名追求科学的年轻学者，这是我第一次撰写学术论著，书中难免包含不严谨之处甚



至错误，敬请学界前辈与同仁不吝赐教。若有任何宝贵建议，请发送至 renyuan@fudan.edu.cn，本人在此深表谢意！

任 远

目录

第1章 绪论	1
1.1 研究动机	1
1.2 相关研究	3
1.3 关于本书	4
1.4 符号说明	7
第2章 模型的生理基础	9
2.1 初级视觉通路	9
2.2 神经节及外膝体细胞的感受野	12
2.3 简单细胞的感受野	16
2.3.1 简单细胞的方向选择性	16
2.3.2 简单细胞感受野的经典模型	17
2.3.3 经典模型的问题及相关模型	18
第3章 外膝体对刺激响应	22
3.1 与对比度无关的响应	22
3.2 响应函数及其性质	27
3.3 响应曲线	31



第4章 简单细胞的方向选择性	34
4.1 基本最小二乘模型	34
4.2 非线性优化模型	37
4.2.1 底层神经元: 基于响应的有向距离估计	37
4.2.2 顶层神经元: 基于最优化决策的方向判定	38
4.3 模型求解及解的性质	40
4.3.1 带二次约束的最小二乘法	40
4.3.2 解的性质	41
4.3.3 数值解	43
4.4 误差分析	44
4.5 改进的非线性加权模型	46
4.6 经典模型的方向歧义性	47
第5章 实验及分析	51
5.1 方向检测方法	51
5.2 模型的选择	53
5.3 参数的确定	54
5.3.1 覆盖率 η 的选择	54
5.3.2 神经元密度	56
5.4 简单细胞感受野的模拟	57
5.5 刺激复杂度与计算误差	60
第6章 应用一: 图像方向检测	62
6.1 检测方法	62

6.1.1	算法描述	62
6.1.2	并行优化	65
6.2	形状图像	65
6.3	自然图像	67
6.3.1	与边缘检测算法比较	68
6.3.2	与全局直线检测算法比较	69
6.3.3	与边缘检测及局部直线检测算法比较	72
6.3.4	与局部直线检测及轮廓检测算法比较	75
6.3.5	由粗到细的表征	78
6.4	对更高层处理的增强	80
6.4.1	分割效果的提升	80
6.4.2	匹配效果的提升	81
第7章	应用二：视错觉的解释	85
7.1	干扰导致的计算偏差	85
7.2	错觉的解释	86
7.2.1	Zöllner 错觉	87
7.2.2	Orbison 错觉	89
7.2.3	Café Wall 错觉	90
7.2.4	Poggendorff 错觉	92
7.2.5	Hering 错觉	93
7.2.6	其他相关错觉	94
第8章	应用三：三维场景理解	98
8.1	图像的三维信息获取	98



8.2	成像模型	101
8.2.1	坐标系的设置	101
8.2.2	坐标系的变换	102
8.3	基本图形的三维信息	104
8.3.1	矩形	105
8.3.2	圆形	107
8.3.3	特殊三角形	111
8.4	场景综合特征分析	113
8.5	实验	116
8.5.1	特征提取	116
8.5.2	朝向计算实验	116
8.5.3	三维构建实验	120
总结		123
参考文献		128

第1章 绪论

1.1 研究动机

机器视觉主要研究如何用计算机模拟生物特别是人类视觉的功能。对传统应用而言,照相机(录像机等)将视觉场景中的光学信号转换为电信号,再存储为数字信号,如图像和视频等,然后用计算机对数字信号进行处理。计算机视觉的一个重要研究目标是:让计算机理解图像或视频中记录的视觉场景,并且获取的场景语义要与观察者的认知一致。计算机视觉的应用范围很广泛,主要包括医学辅助诊断、智能交通系统、遥感图像分析、医学图像处理、机器人导航、多媒体技术、文字识别、地图绘制、运动分析、工业检测、精密加工、军事侦察和目标跟踪等众多领域。数字图像处理是计算机视觉的一个重要分支,是一种通过计算机对数字图像或视频进行去增强、复原、分类、识别等处理的方法和技术^[1]。

一方面,数字图像处理(此下简称为图像处理),作为计算机研究领域的一个重要分支,在其出现后的几十年来取得了非常显著的发展,如高清视频技术已经深入到日常生活。另一方面,图像处理技术仍然面临许多严峻的问题,如图像增强效果不理想、物体识别能力与效率偏低等。目前还没有一个完整的理论体系能够解决计算机视觉的全部问题。然而,对计算机视觉很困难的问题,人和许多

高级哺乳动物却可以用他们(或它们)的视觉系统轻易地解决。例如,人通常可以仅用一只眼睛就能从繁杂的环境中轻松辨识出物体,而且这种识别能力是相对稳定的,受各种变化如光照、位移、遮挡等的影响不大。高等哺乳动物的视觉系统在主动性、实时性和智能性等各方面都具有目前计算机视觉系统所无法比拟的优越性。因此,如果能用计算机模拟高等哺乳动物视觉系统的功能将是非常有意义的。事实上,越来越多的学者正在进行基于生物视觉机制的图像处理算法的研究。

随着神经科学的发展,科学家对高等哺乳动物视觉系统的研究取得了重要的进展,对视觉的脑机制的研究也逐渐取得显著成果^[2]。诺贝尔生理学奖获得者 Hubel 和 Wiesel 通过对猫的大脑纹状皮层的研究认为,可以用一个由简单细胞(simple cell)层(以下简称 S 层)和复杂细胞(complex cell)层(以下简称 C 层)交替组成且不断加深复杂化的网络来模拟大脑处理视觉信息的过程^[3,4]。这一理论为基于生物视觉的计算机视觉研究提供了神经基础,此后进而出现了许多受生物视觉启发的模式识别方法。1980年, Fukushima 提出了一种自组织式神经网络模型^[5],此模型运用了 Hubel 和 Wiesel 的理论,由许多 S 层和 C 层交替组成。其中, S 层逐步提高特征的复杂性, C 层逐步提升特征的位置不变性。文献[6, 7]利用对猴子大脑视皮层研究得到的成果,提出了新的分层结构式的方法。文献[8]将脉冲发放神经网络用于图像的边缘检测及分割。文献[9]提出一个基于神经视觉的手写字符识别算法。文献[10-14]等也都提出了基于生物视觉的模式识别模型。

神经系统对视觉信息处理是一个能动加工的过程,其首要的问题是将外界物理刺激表征为内部可以处理及存储的信息,表