

杭州电子科技大学

硕士学位论文

题目：基于视觉感知的神经计算模型及其应用研究

研究生 王典

专业 模式识别与智能系统

指导教师 范影乐 教授

完成日期 2015 年 12 月

杭州电子科技大学硕士学位论文

基于视觉感知的神经计算模型
及其应用研究

研究生：王 典

指导教师：范影乐 教授

2015 年 12 月

Dissertation Submitted to Hangzhou Dianzi University

for the Degree of Master

Research on Neural Computing Model and Its Applications Based on Visual Perception

Candidate: Wang Dian

Supervisor: Prof. Fan Yingle

December, 2015

杭州电子科技大学

学位论文原创性声明和使用授权说明

原创性声明

本人郑重声明： 所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不含任何其他个人或集体已经发表或撰写过的作品或成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。

申请学位论文与资料若有不实之处，本人承担一切相关责任。

论文作者签名： 日期： 年 月 日

学位论文使用授权说明

本人完全了解杭州电子科技大学关于保留和使用学位论文的规定，即：研究生在校攻读学位期间论文工作的知识产权单位属杭州电子科技大学。本人保证毕业离校后，发表论文或使用论文工作成果时署各单位仍然为杭州电子科技大学。学校有权保留送交论文的复印件，允许查阅和借阅论文；学校可以公布论文的全部或部分内容，可以允许采用影印、缩印或其它复制手段保存论文。（保密论文在解密后遵守此规定）

论文作者签名： 日期： 年 月 日

指导教师签名： 日期： 年 月 日

摘要

视觉感知是计算神经和神经工程的热点研究问题之一，一方面它将有助于揭示视觉系统具有惊人信息感知能力的内在机理，另一方面它将给以机器视觉为代表的智能信息处理方法带来新的思路。本文基于人类视觉感知机制构建神经元网络，利用具有时空特性的神经脉冲发放，探索在视觉激励下的神经信息加工过程，并尝试面向图像处理的实际应用。本文首先以突触连接的形式构建新的神经元感受野模型，通过模型对明暗刺激的响应，阐释新模型在实现视觉机制中所起的重要作用；然后考虑视觉通路上的多级方位敏感特性，构建视皮层下-初级视皮层的分级处理模型，利用对低对比度图像边缘的检测效果，解释了方位敏感机制在视觉轮廓感知中所起的关键作用；最后引入视觉信息稀疏编码规则和视觉注意机制，构建具有空间信息融合特性的神经元网络，通过显著性图像对边界响应的调制，提供了一种基于视觉信息反馈调控的图像处理应用新思路。本文主要工作及研究成果如下：

(1) 以突触连接的形式构建新的神经元感受野模型，实现方位敏感、侧向抑制以及选择注意机制，充分发挥其在感知对比刺激以及空间变化中的作用。将融合了感受野区域信息的图像特征输入神经元模型，将视觉刺激转化为脉冲序列；再从脉冲序列的时间编码信息中解析视觉刺激所包含的特征。神经元响应表明该感受野模型具备了方位选择以及明暗边缘检测等初步的视觉功能。

(2) 考虑了视觉通路上的多级方位敏感特性，提出一种基于突触连接视通路方位敏感的图像分级边缘检测新方法。针对具有多强度边缘或者丰富细节的图像，利用神经节细胞以及外膝体神经元感受野向心分布的生理结构特性，构建具有多方向敏感特性的视皮层下功能层，提取边缘敏感图像；并优化侧向抑制作用方式，形成初级视皮层功能层，去除冗余信息。实验结果显示，新方法在完整检测图像边缘的同时，并不引入纹理噪声，有着明显的优势。

(3) 针对视通路模型不具有区分主体轮廓和纹理边缘的能力，引入神经信息稀疏编码规则和视觉注意机制，提出一种基于显著性信息稀疏表示的图像轮廓检测新方法。首先针对初级视皮层 V1 对图像边界的响应结果，构建具有空间信息融合特性的稀疏神经元网络；其中，网络输入引入视觉注意机制形成的空间差异性特征，网络输出显著性图像；最后利用该显著性信息调控 V1 边界响应，得到图像轮廓。在自然图像的轮廓检测任务中，新方法能有效抑制纹理边缘的响应，而保留尽可能多的真实轮廓，具有较好的稳定性。

关键词：视觉感知，神经计算，感受野，视觉通路，轮廓检测

ABSTRACT

Visual perception is one of the hot research problems in computational neuroscience and neural engineering. On the one hand, it helps to reveal the inner mechanism of the amazing ability of the visual system to perceive information. On the other hand it will bring new ideas to the machine vision and other intelligent information processing methods. In this paper, neural networks based on human visual perception mechanism were constructed to explore the neural information processing using the spatial-temporal characteristics of firing spikes, and to try to deal with practical applications in image processing. Firstly, the proposed method designed new neurons receptive field model in the form of synaptic connections. Through the model response to the light and shade stimulus, we explained the important role of the new model in achieving the vision mechanism. Then, a under visual cortex - primary visual cortex hierarchical processing model was constructed considering the multi-stage directional sensitive nature of the visual pathway. We used the results of edge detection in low-contrast images to explain the key role of the orientation sensitive mechanism plays in the contour perception. Finally, sparse coding rules and visual attention mechanism were introduced to build neural network which had the characteristics of spatial-temporal information fusion. With the modulation of boundary response by saliency image, we provided a new way of image processing applications based on visual feedback control. The main tasks and research results are listed as follows:

(1) New receptive field models were designed in the form of synaptic connections to implement orientation sensitivity, lateral inhibition and selective attention mechanisms. These models could give full play to their roles in the perception of contrast and spatial variation in visual stimulus. Took the image feature which was integrated with the spatial information of receptive field as input of neuron model to convert visual stimuli into a pulse sequence. And characteristics of visual stimuli were encoded in the exact time the spike fired. Neuron responses indicated that these new models could perceive direction, detect shading edges, and could implement other preliminary visual function.

(2) Considering the multi-stage directional sensitive nature of the visual pathway and dynamic characteristics of synaptic connections, a hierarchical image edge detection method was presented. Using the physical structure feature of ganglion cells and LGN neurons receptive field distributing centripetal, a sub-cortex multi-direction sensitive function layer was constructed to get the edge sensitive strength of an image with fuzzy hierarchy and rich details. Then a primary visual cortex

function layer was formed with optimized lateral inhibition to remove redundant information. Experimental results showed that this method could completely detect image edge and effectively filter out texture noise.

(3) For visual pathway model didn't have the ability to distinguish between the body contour and texture edges, a new contour detection method base on sparse coding of salient region was proposed by introducing the neural sparse coding rules and visual attention mechanism. For the orientation perception of V1, a sparse neural network with spatial information was instituted to suppress the response of texture edges. The difference characteristics of classical receptive field and the non-classical receptive field was calculated and introduced to the network input, and the network output saliency image. Finally, the V1 contour response was obtained by modulating the feedback of saliency image with its orientation perception. In the contour detection task of natural images, the new method was proved to be can effectively inhibit the response of textured edges, while retaining as much of the real contour, and had superior stability.

Keywords: visual perception, neural computing, receptive field, visual pathway, contour detection

目录

摘要.....	I
ABSTRACT.....	II
目录.....	IV
第 1 章 绪论.....	1
1.1 研究目的与意义.....	1
1.2 国内外研究现状.....	2
1.3 本文主要工作及创新点.....	4
1.4 论文的组织结构.....	5
第 2 章 视觉感受野机制及脉冲编码.....	7
2.1 视觉通路概述.....	7
2.2 视觉感受野及其计算模型.....	8
2.2.1 神经节细胞与 LGN 神经元感受野	8
2.2.2 初级视皮层细胞感受野.....	9
2.2.3 非经典感受野.....	10
2.3 神经元模型.....	10
2.3.1 Hodgkin-Huxley 神经元模型.....	10
2.3.2 Integrate-and-Fire 神经元模型.....	11
2.4 脉冲编码.....	12
2.4.1 频率编码.....	13
2.4.2 次序编码.....	13
2.4.3 时间编码.....	14
2.5 小结.....	14
第 3 章 基于感受野特性的视觉机制建模.....	16
3.1 概述.....	16
3.2 方位敏感机制建模.....	16
3.3 侧向抑制机制建模.....	18
3.4 视觉注意机制建模.....	20
3.5 小结.....	22
第 4 章 基于突触连接视通路方位敏感的图像分级边缘检测.....	23
4.1 概述.....	23

4.2 视通路分级模型.....	23
4.2.1 视皮层下方位敏感.....	23
4.2.2 初级视皮层侧向抑制.....	25
4.3 算法及算法验证.....	26
4.3.1 算法.....	26
4.3.2 算法验证.....	28
4.4 实验结果.....	29
4.5 讨论与分析.....	32
4.6 小结.....	33
第 5 章 基于显著性信息稀疏表示的图像轮廓检测.....	34
5.1 概述.....	34
5.2 轮廓检测模型.....	34
5.2.1 V1 神经元感受野边界响应.....	34
5.2.2 稀疏编码神经元网络.....	35
5.2.3 视觉注意机制.....	36
5.2.4 高级视皮层反馈控制作用.....	37
5.3 实验结果.....	37
5.4 讨论与分析.....	41
5.5 小结.....	42
第 6 章 总结与展望.....	43
6.1 总结.....	43
6.2 展望.....	44
致谢.....	45
参考文献.....	46
附录:.....	51

第 1 章 绪论

1.1 研究目的与意义

神经科学和信息科学作为当今发展最为迅猛的科学前沿，已并不仅仅只是停留在各自的领域，而是互相借鉴交叉发展，具体表现为两个主要的研究方向：其一，结合计算机辅助和数学分析等方法探讨神经系统生理运作规律和信息处理的相关机制；其二，借鉴神经科学已有的科学研究成果，发展新的计算模型和信息处理方式，从而实现信息处理智能化。而对视觉感知机制的研究是神经科学重要的科研内容之一。一方面，通过不同的实验手段获得的大量生理数据需要借助数学模型的简化和分析才能发现其中的规律，实验过程中相关参数的选取和设定只有依靠计算机辅助计算才能进一步优化和精确；另一方面，生物视觉近乎完美的感知能力给计算机信息处理系统传递着一些启示，告知人们有可能存在与生物视觉信息处理方式一样高效的新的信息技术，激发越来越来多的信息科学研究人员试图通过建立仿生的计算模型来解决信息处理领域中的难题^[1]。尽管由于脑以及视觉系统的极度复杂性，导致人们对其相关机制的描述并不准确和完整，对其认识暂未完全，然而科学家们对视觉系统从眼睛到大脑皮层的研究依然取得了大量成果，而且积累下来了海量数据，这对于我们了解生物机制并将其运用于信息处理领域意义重大，尤其在机器视觉领域，视觉机制的相关理论研究成果已被广泛应用于图像处理、目标跟踪、行为理解等科研和实际应用之中。

在图像处理和计算机视觉中，图像或者视频信息的冗余常常给特征提取、目标识别、图像理解以及信息压缩、传输和存储带来较大的障碍，如何从海量数据中提取有效信息是数字图像处理最基本却又极具难度和深度的技术手段。Marr^[2]提出的视觉计算理论认为，在自底向上的图像特征分析过程中，边缘和轮廓是人类识别目标所依赖的主要特征之一，这主要是由于边缘和轮廓定义为图像中局部区域或者平面的交界处，是图像信息发生突变的地方，往往不确定性最大，所蕴含的信息更丰富，而同时还保留了图像中目标的形状信息。由此可见，边缘和轮廓特征提取将为图像理解或者运动目标行为分析提供重要的降维信息。其中，在准确定位边缘的前提下不发生漏检，以及避免出现虚假边缘是对边缘特征描述最基本的两项要求。传统边缘检测方法，通常试图利用数学模型来刻画局部邻域像素之间在特定方向上的灰度变化^[3]。例如 Canny 边缘检测方法^[4]，具有运算效率高和实时性较好的特点，但由于其从图像梯度出发设定阈值，面对具有多强度边缘或者丰富细节的图像时，很难同时满足上述两项要求。而轮廓与边缘又有不同，它主要指图像中能够反映主体目标的轮廓信息，包括主体目标与背景、平面与平面的边界，能够为后续基于形状的目标识别与图像理解等高级视觉行为提供更有用的信息^[5]。因此这就要求检测方法能够对平面内的纹理性边缘具有一定的抑制作用，凸显目标形状，这是非传统边缘检测方法所能达到的。

结合神经科学领域对视觉信息处理机制的相关理论研究成果，例如初级视皮层对输入刺激的方向选择特性，使得视觉系统具有优秀的轮廓感知能力，进而为后续高智能行为奠定了基础；以及具有生理特性的神经信息计算模型，例如神经元膜动作电位模型及由其组建的神经网络，能将视觉刺激转化为脉冲信号并进行编码，从而提取图像特征。因此如何从人类视觉感知机制出发，开发神经计算模型，对图像进行边缘和轮廓检测，具有非常重要的研究意义。

1.2 国内外研究现状

近几十年来，众多学者在传统方法基础上，通过引入其他优化算法提出大量边缘及轮廓检测方法^[5]。例如，在不影响运行速率的情况下，基于分布式图像块的 Canny 改进算法，能有效适应噪声环境^[6]；利用蚁群算法并行和智能处理能力，对图像梯度的幅度和方向进行线性加权从而提取图像显著性轮廓^[7]；引入了马尔科夫随机场能量函数的水平集算法，整合局部空间信息对医学图像进行分割，提高了传统水平集分割算法的检测速率和鲁棒性^[8]，等等。这些算法有效改善了边缘及轮廓检测的精度和效率，但相对于人类视觉系统从极度复杂场景中能选择性注意主体目标，识别场景中有效轮廓信息的能力而言，所表现出的智能性还远远不足。

另一方面，随着生理实验手段的发展以及数据分析技术在神经科学中的运用，对于视觉信息处理的脑机制的研究已取得丰硕成果^[9]，启发信息科学家和数理科学家建立相应的数学模型和计算方法并运用于图像处理和机器视觉领域，以期改善现有图像处理算法的不足之处。例如，基于初级视皮层 V1 简单细胞感受野具有方位敏感性，Daugman 提出用 Gabor 滤波器描述其响应特性^[10]。Gabor 滤波器具有方向选择、带通等优良性质，能用来描述简单细胞感受野对明暗对比边缘的敏感性响应。然而在面对复杂场景时，Gabor 函数虽能提取边缘信息，却不能抑制背景和纹理噪声的响应，对主体轮廓的提取造成很大的干扰。于是，许多研究者开始将注意力转移到感受野之外的被称之为非经典感受野(Non-Classical Receptive Field, NCRF)的周边区域，对该区域的直接闪光刺激虽然不能引起细胞反应，但却能影响该细胞对感受野内刺激所产生的反应，即调制该细胞的反应性，且这种调制作用大多数情况下是抑制的^[11, 12]。与 NCRF 相对，将中心传统意义的感受野区域称为经典感受野(Classical Receptive Field, CRF)。Grigorescu 等提出先用 Gabor 能量表征初级视皮层 CRF 对边缘线段的响应，再再用高斯差函数(Difference of Gaussian, DoG)经半波整流后的模型(DoG⁺)表征 NCRF，通过 NCRF 的侧向调控，抑制纹理边缘的能量响应凸显主体轮廓^[13]。Grigorescu 同时指出^[14]，根据是否考虑 CRF 跟 NCRF 最优方位的差异性，侧向抑制分为同步和异步两种方式，同步抑制作用强，对杂乱纹理具有较好的抑制作用，而异步抑制作用较弱，但对规律性纹理抑制作用更好，各有所长。近年来，Chen^[15, 16]以及 Xu^[17]等人对 NCRF 作了大量研究，发现 V1 通过检测 NCRF 与 CRF 空间差异来调制神经元响应，且 NCRF 不仅仅具有抑制作用，而且还存在去抑制区，这对于提高 CRF 边缘响应的连续性起到了重要作用，进而提高了视觉系统处理大范围复杂信

息的能力。于是,众多学者又提出对 Grigorescu 模型的改进方法。Zeng 则提出蝴蝶形状的 NCRF 抑制方式^[18, 19],根据方向或尺度特征自适应调整 NCRF 的抑制区和去抑制区,有效提高了所提取轮廓的平滑性。Yang 则认为 NCRF 对 CRF 响应的抑制是同步和异步方式共存的,且提出了两种模型,证实了将多方式组合后的模型能获得更好的边缘提取效果^[20]。这些改进虽提升了轮廓提取能力,但都仅在原有模型基础上考虑更多的视觉特性,并未引入其他图像特征, Yang 等则进一步通过提取自然图像的方向、光强和局部光强对比等多种图像特征,对感受野输出图像施以大范围非经典感受野抑制来获取图像轮廓,检测效果相对单一特征的边缘检测方法优势较大^[21, 22],进而他们又提出利用视锥细胞颜色信息处理的拮抗机制对彩色图像进行边缘检测,同样获得了不错的效果^[23]。自 Hartline 发现并提出神经信息处理侧抑制原理,神经元突触间的抑制作用得到不少学者的研究。Isaacson 通过实验分析并阐述了神经元间的抑制作用对于附近神经元维持自制性和独立性创造了条件,这进一步提高了神经系统的稳定性^[24]。Zi 等运用侧抑制网络增强对比度、突出空间变化的特点,对红外图像进行预处理,有效突出了弱小目标形状信息并保留目标细节,为后续目标检测创造了条件^[25]。史漫丽等通过计算图像局部特征来构建不同尺度和方向的高斯滤波器自适应确定侧抑制网络参数,以检测红外弱小目标,该方法能有效提高弱小目标检测算法的抗噪能力^[26]。还有研究人员通过抑制性突触作用方式组建神经元群网络对图像进行边缘检测,该方法综合了侧向抑制、视觉注意等多种视觉机制,相对其他单一视觉机制模型具有一定优势^[27]。

由此可见,上述方法大多数通过创建和改进符合视觉机制的数学模型对图像进行滤波处理;再者,结合更多的图像特征、改善特征融合办法、优化参数设置方式来提高算法性能,但却忽略了光刺激经视网膜神经节细胞转换为细胞膜电位,直至视神经中枢均以动作电位的形式进行传递这一基本的生理基础,必然丢失脉冲编码信息。Hodgkin 和 Huxley 在 1952 年首次以数理科学知识和方法建立了神经元动作电位产生和传导的模型,即 Hodgkin-Huxley 模型(简称 HH 模型),奠定了神经电生理学的基础。基于动作电位发放的神经元模型,研究人员又开始探究组建神经元网络来进行信息处理的新思路。20 世纪 90 年代初, Eckhorn 提出一种基于同步脉冲发放的神经网络模型——脉冲耦合神经网络(Pulse Coupled Neural Network, PCNN)^[28]。相对传统神经网络,PCNN 能直接从复杂背景中提取所需信息,而不需要经过训练,因此被广泛应用于快速图像分割、目标识别等。毕英伟等提出一种基于图像空间信息的自适应 PCNN 参数简化算法,实现图像的自动分割^[29]。Gao 等结合贝叶斯聚类算法改进 PCNN 迭代特性,实现图像自动分割^[30]。PCNN 模型虽然具有神经元脉冲发放和传递特性,但仅是对视觉神经网络进行黑箱模拟,无法实现其他的视觉机制,缺乏一定的灵活性,且模型参数众多寻优不易,稳定性较差。最近,研究人员又对以脉冲序列作为网络输出的神经网络产生很大的兴趣,这就是所谓的脉冲神经网络(Spiking Neural Network, SNN)。基于该网络良好的灵活性和信息编码能力,学者们展开了大量面向图像处理应用的研究。Kerr 等提出一种基于六边形感受野模型的 SNN 对图像进行边缘检测,与传统的矩形感受野进行比较,更符合生物

特性,检测效果明显高于后者^[31]。Wu 等通过三层 SNN 模型实现视觉注意机制,提取图像底层方向、颜色、纹理等特征,经过视觉 ON/OFF 通道将图像分解为不同部分,再利用自顶向下的控制信号识别场景中的感兴趣区域^[32]。另外,也有研究人员利用突触放电时间依赖的可塑性(spike-timing-dependent plasticity, STDP)训练非监督的神经网络,模拟视皮层对颜色信息的聚类特性^[33]。再者,用 SNN 将图像信息转化为脉冲序列,对脉冲序列进行时空编码,再从学习-分类的角度进行图像模式识别,探索人脑在进行目标识别、图像理解时的神经信息处理过程^[34-37]。

综上,可以看出国内外目前有关视觉信息处理机制与图像处理结合的发展趋势,一方面通过建立仿生的数学模型实现视觉机制,然后将其运用于图像处理,忽视了视觉信息传递过程中的电生理基础,即会丢失脉冲发放的时间编码信息;另一方面组建大规模神经网络,却局限于某种单一的视觉神经特性,注重实现神经网络自身的组织结构,忽略了与视觉机制的结合。所以本文研究基于视觉感知的神经计算模型,注重视觉通路上不同结构对视觉信息的分工处理,简化和优化视觉机制模型,通过脉冲编码的方式实现图像边缘和轮廓检测,以期获得比传统方法更好的结果,或者能提供一种基于视觉机制的图像处理和理解新思路。

1.3 本文主要工作及创新点

本文研究的主要内容是围绕基于视觉感知的神经计算模型,探索其在图像处理上的应用。一方面侧重于构建简化和优化的神经网络模型,实现视觉感知相关机制,将图像特征以动作电位的形式呈现,另一方面注重于视觉信号在视通路中传递时所依赖的生理结构及其功能流,实现图像信息的分级或者回馈等多种不同的处理方式。首先,基于神经生理发现的视觉感知机制,研究用神经网络的方式实现相应的功能,并以简单的感受野明暗刺激为例验证网络模型的有效性;然后,考虑了视通路上不同的组织结构对边缘信息有着不同层次的感知能力,构建视皮层下-初级视皮层的分级处理模型,结合动态突触连接、侧向抑制等机制对低对比度图像边缘进行了检测,取得不错的效果¹;最后,利用脉冲神经元不应期特性,构建具有空间信息融合特性的 leaky integrate-and-fire(LIF)神经网络,引入 CRF 和 NCRF 差异性特征实现显著性图像的稀疏表示,将其反向投递到初级视皮层,调控初级视皮层对边界的响应并提取图像轮廓²。

论文具体研究工作包括以下几个方面:

(1) 本文以突触连接的形式构建新的神经元感受野模型,实现方位敏感、侧向抑制以及视觉注意等视觉感知机制,充分发挥其在感知明暗光对比刺激以及空间变化中的作用。首先将融合了感受野区域信息的图像特征输入神经元模型,通过神经元模型将视觉刺激转化为脉冲序列;再从脉冲序列的时间编码信息中解析视觉刺激所包含的特征。以明暗刺激施加于该感受野模型,通过神经元的响应证实了本文提出的方位敏感模型能有效提取视觉刺激中不同方

¹基于突触连接视通路方位敏感的图像分级边缘检测方法,中国生物医学工程学报,一级期刊,第一作者。

²基于显著性信息稀疏表示的图像轮廓检测,电子与信息学报,已投稿,审稿中。

位的亮暗对比边缘，侧向抑制模型能改善由于噪声和模糊造成的斜坡边缘对真实边缘定位的影响，视觉注意模型能显著整合空间对比信息并抑制纹理噪声对轮廓提取的干扰，实现了初步的视觉感知相关功能。

(2) 考虑视觉通路上不同生理结构对视觉信息的分工处理，针对具有多强度边缘或者丰富细节的图像，优化侧向抑制方式，从而提出一种基于突触连接视通路方位敏感的图像分级边缘检测新方法。首先考虑视皮层下视网膜神经节细胞和外膝体的方位敏感特性，探讨其感受野最优方位在视网膜上的分布规律，利用神经元突触间的动态连接特性构建视皮层下多方位敏感层；再结合初级视皮层更强烈的特殊方位条形光刺激感知能力，对上一功能层输出的脉冲序列编码，并对首发脉冲时间信息加以侧向抑制和阈值处理，得到图像边缘；对层次模糊而细节丰富的菌落图像进行处理，并使用重构相似度和边缘置信度对检测结果进行有效性评价，评价结果显示新方法能在边缘漏检和噪声误检之间达到较好的平衡。

(3) 针对视通路模型不具有区分主体轮廓和纹理边缘的能力，构建一种新的稀疏编码网络模型，并结合视觉注意机制在改善轮廓感知中的作用，从而提出一种基于显著性信息稀疏表示的图像轮廓提取方法，不仅能有效提取图像轮廓，而且能有效抑制纹理边缘。首先用二维高斯导函数实现 V1 的方位敏感特性，提取边界响应；该边界响应中包含大量纹理性边缘，故针对该边界响应，构建具有空间信息融合特性的稀疏神经网络；在脉冲神经元模型中加入不应期使得在时间维度上满足稀疏编码规则，同时引入 NCRF 和 CRF 差异性特征，实现视觉注意机制，将网络输出的显著性图像以高级视皮层反馈控制的形式作用于 V1 边界响应，得到最终的图像轮廓。以 RuG 图像库为例，检测结果显示该方法能有效抑制纹理边缘的响应，而保留尽可能多的真实轮廓；另外，计算最佳检测结果与基准轮廓图的 P 值指标，并利用盒须图 (Box-and-Whisker) 统计分析，结果表明，相对其他对比方法，本方法具有更好的稳定性。

1.4 论文的组织结构

本文组织结构及各章节内容简述如下：

第一章绪论，首先从神经科学与信息科学的交叉发展引出本文研究在图像处理和计算机视觉领域的研究目的与意义；对国内外视觉信息处理机制以及相关应用的研究现状进行了概述，阐述了视觉机制以及神经计算模型在边缘检测、图像分割等的应用及其局限性；针对现有研究思路的局限性，给出本文的主要研究内容和创新点；最后列出了论文的组织结构和各章节的内容安排。

第二章对人类视觉系统从视网膜到视皮层的视觉通路进行了简单介绍；对视通路中各级机构所具有的感受野特性及其计算模型予以回顾；另外，对 HH、IF 以及本文中所使用的 LIF 神经元模型进行了数学表述，同时展示其动作电位发放特性；最后对视觉信息编码的普遍规则及在计算模型中常用的编码方式进行了详细介绍。

第三章针对数字图像处理中普遍存在的噪声、模糊边缘以及轮廓提取等困难，结合生物

视觉系统方位敏感机制、侧向抑制机制以及视觉注意机制，通过组建能实现视觉感知功能的感受野模型，并对明暗边缘、斜坡边缘和纹理边缘等视觉刺激进行仿真处理，以神经元脉冲响应来表征模型在解决上述图像处理困难中的表现，结果表明所提模型在方位感知、边缘定位以及纹理抑制中能够实现初步的视觉功能，为后续基于视觉机制在图像处理上的应用奠定基础。

第四章提出一种基于突触连接视通路方位敏感机制的分级轮廓检测方法，讨论视觉通路上神经节细胞、外膝体以及初级视皮层在轮廓特征提取过程中的具体分工，探讨视皮层下细胞感受野分布与最优敏感方位之间的关系，同时引入初级视皮层在信息筛选过程的脉冲编码和侧向抑制机制，摒弃采用黑箱形式的数学模型对视觉机制进行模拟的思路，提出一种基于突触连接皮层下和皮层上分级处理的图像边缘检测方法，并与 Canny、Gabor 滤波器以及 MCI 等方法进行分析比较，表明本文提出的方法检测出的图像边缘细节更加丰富且不引入纹理噪声。

第五章提出一种基于显著性信息稀疏表示的图像轮廓提取方法。通过构建具有时空稀疏特性的神经网络，以较为新颖的方式实现显著性图像的稀疏表示，更多的考虑了该特性存在的生理基础，而非从矩阵分解等数学分析方式对图像进行降维表示；另外通过引入 CRF 与 NCRF 的差异特征实现视觉注意机制，并以视通路高级皮层反馈调控的作用方式，形成最终的轮廓响应。针对 RuG 图像库中的自然图像，通过与不加抑制的方法（Noninh）、空间稀疏限制的纹理抑制方法（SSC），以及 NCRF 同步抑制方法（ISO）所提取结果进行比较，表明该算法不仅对纹理和背景具有抑制的作用，而且能有效区分纹理强边缘和主体轮廓，获得较佳的效果；另外用盒须图（Box-and-Whisker）对算法在不同参数下的表现进行统计分析，得出算法的统计性能，表明本方法稳定性更高。

第六章对本文的研究工作及其成果进行了总结，然后对后续模型改进、应用创新等进行了展望。

第 2 章 视觉感受野机制及脉冲编码

人类视觉信息处理从视网膜到视皮层的视觉通路上，不同的组织结构对视觉信息有着不同功能层次的感知和分析能力。随着对相关视觉感知机制的研究不断深入，视通路上各个组织结构所扮演的功能陆续被揭示。而这些视觉机制的实现，很大程度上是依靠视觉感受野来整合空间分布信息，提取信息特征的。因此信息科学家和数理科学家通过建立相应的计算模型来模拟感受野机制，进而实现视觉感知特性，并希望设计出类似于人类视觉系统一样优秀和智能的计算机信息处理系统。另一方面，神经元是神经信息处理的基本单元，随着实验手段的逐渐发展，生理科学家从功能层面逐步进入到对分子、细胞层面的研究，电生理实验的研究成果促使能描述动作电位产生和传导的神经元模型被建立起来，例如 HH 模型、IF 模型等。神经元通过动作电位的发放将视觉信息进行编码，于是研究人员试图通过设计不同的编码方式来解读脉冲序列所包含的图像特征，例如频率编码(rate code)、次序编码(order code)、时间编码(temporal code)等。

2.1 视觉通路概述

人类通过视觉感知获取超过 80% 的外界信息，经大脑皮层的判断与理解从而认识世界。视觉感知过程包括目标定位，对目标形状和颜色等静态特征进行辨识，对移动目标进行运动检测，以及后续特征匹配、目标识别、行为理解等，包含了复杂的神经机制。从生理学和解剖学的角度来说，视觉过程始于眼睛，止于大脑皮层，图像刺激从视网膜沿着一定的视觉通路传至大脑视皮层，这个过程主要有两条通路：一条由视网膜神经节细胞经外膝体(lateral geniculate nucleus, LGN)达到视皮层，称为视觉的第一视通路；另一条由神经节细胞经上丘达到丘脑枕，再传至视皮层，称为第二视通路。由于人类视觉感知机制主要与第一视通路相关，故本文主要介绍第一通路的信息传递流程，如图 2.1 所示。

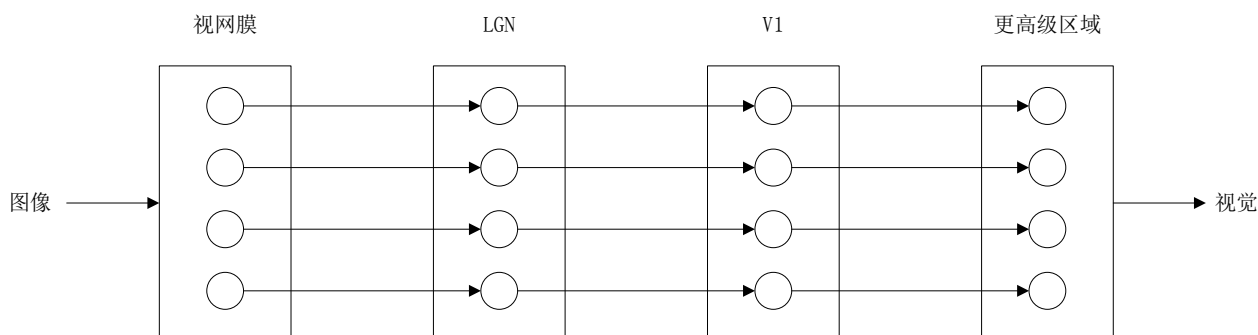


图 2.1 视觉通路示意图

图像刺激在视网膜中已进行了很好的初步处理，通过神经节细胞以动作电位的形式对视觉刺激进行编码，然后传递到外膝体。视网膜是大脑的一部分，因此视网膜并不仅仅将视觉

刺激呈现为清晰的图像，然后转换成电脉冲信号，而且对图像各方面信息进行有效提取和加工处理，以便大脑高级皮层对这些信息进行分析 and 解释。

外膝体包含六个细胞层，依次命名为 1~6 层，分两条通路接收神经节细胞的脉冲输出：**M** 通路包含前两层（1 层和 2 层），接受来自视网膜 **M** 型神经节细胞的投射，处理图像信息中的亮度、低频信息；**P** 通路包括剩下的四层（3、4、5 和 6 层），接受来自视网膜 **P** 型神经节细胞的投射，具有颜色拮抗特性并处理高频信息。目前普遍认为在 **LGN** 各层之间存在第三条通路（**K** 通路），该通路同样具有颜色拮抗的特性。

视觉信息自外膝体的突触传出至初级视皮层 **V1**（也称 17 区或纹状皮层），进行方位、颜色等其他特征的提取和处理。外膝体神经元会汇聚成群共同输入到同一个视皮层细胞，其感受野互相重叠并排列成一行，使得视皮层细胞对与其所排列方位相同或相似的条形光刺激有着强烈的反应，也就是 **V1** 细胞具有方位选择性。而视皮层也有一些细胞的感受野呈圆形，并具有颜色拮抗的中心-外周特性，这类细胞虽不具备方位选择性，但能对物体颜色进行感知。**V1** 区向更高层次的视皮层进行投递，作进一步有关运动信息、物体辨认等分析和处理。

另外，在视通路上，视觉信息的处理并不只存在图 2.1 所示单向的前馈通路，高级结构也会向下级结构反馈调控信号，形成错综复杂的神经网络，这有利于信号的进一步整合。而且，视觉信息处理的过程中存在普遍的规则和特性，例如稀疏编码理论，能够实现从大量复杂信号中提取出有效信息，达到压缩信息的目的，这也符合视通路从低级到高级、由简单到复杂的生理结构和感受野等级特性。由此可见，视觉感知的形成是一个极度复杂的过程，要完全弄清人脑是如何通过视觉系统认识世界的，对视觉感知机制的研究还有很长的路要走。

2.2 视觉感受野及其计算模型

在视网膜上一定的区域，对该区域施加给光刺激能改变神经元的放电活动，这个区域被称为视觉感受野。很多视觉机制正是由于神经元整合了感受野空间特征才得以表现出来。沿视觉通路，不同层次的生理结构具有不同特性的视觉感受野，整体上表现为，越高层次的神经元感受野范围越大、结构越复杂。于是，视觉信息沿着视通路被相应层次结构的感受野提取特征，并不断去除冗余信息，保留有效信息，从而形成视觉感知。目前，各层次神经元感受野已得到充分的研究和证实，获得大量重要的成果。

2.2.1 神经节细胞与 LGN 神经元感受野

1953 年，Kuffler 在对猫视网膜神经节细胞感受野的研究观察中发现中心-外周拮抗式感受野特性，分为 **ON**-中心型和 **OFF**-中心型两种，如图 2.2 中各左图所示，其中“+”表示兴奋区域，“-”表示抑制区域。对于 **ON**-中心型感受野，当光点刺激其感受野中心区域时神经元放电频率增加，而光点刺激环绕中心的外周区域时放电频率减少；对于 **OFF**-中心型感受野，则刚好相反。1965 年，Rodieck 又对此同心圆拮抗式感受野特性进行了量化描述，提出用双高斯函数之差（**DoG**）描述视觉刺激和神经元响应之间的关系^[38]，如图 2.2 中各右图所示，其中实线表示神经元感受野响应曲线，虚线和点线分别表示兴奋区域和抑制区域响应曲线。

同视网膜神经节细胞一样，LGN 神经元的感受野亦可表示为中心-外周拮抗式结构，但相对于神经节细胞感受野，其范围更大，外周作用更强。

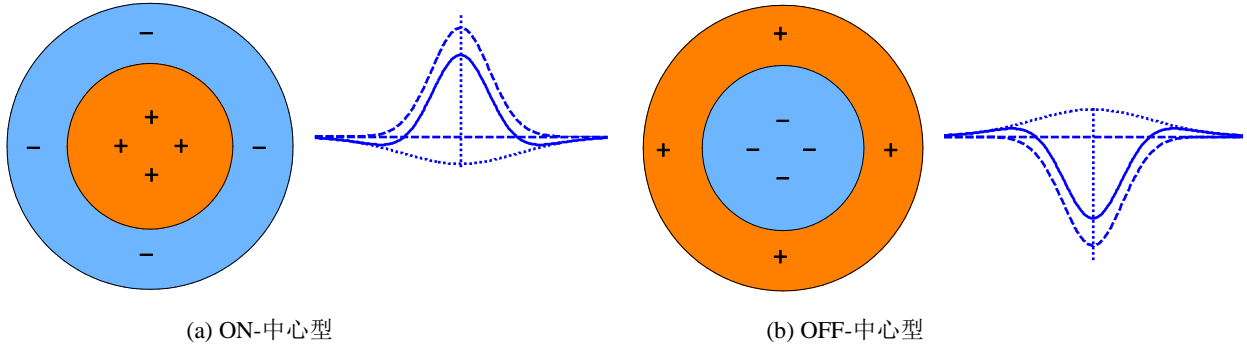


图 2.2 神经节细胞感受野及 DoG 计算模型示意图

视网膜神经节细胞和 LGN 神经元感受野相似的结构决定了其具类似的功能。空间拮抗的感受野作用方式，提供了视觉系统感知空间对比度分辨的能力，对于明暗对比边缘特别敏感，显然这为形成边界和形状知觉准备了条件。

2.2.2 初级视皮层细胞感受野

20 世纪 60 年代，Hubel 和 Wiesel 发现，初级视觉皮层（V1）细胞具有朝向敏感性，即由于外膝体神经元感受野重叠并排列成一行的生理结构，使得 V1 不同神经元对特定方位的视觉刺激会产生更强烈的响应，而对于其他方位反应较弱。由此引发了大量有关感受野模型的研究工作^[39]。其中，Daugman 基于简单细胞感受野中心 ON 区域为狭长形，提出用 Gabor 滤波器描述其响应特性^[10]。Gabor 滤波器具有方向选择、带通等优良性质，能用来描述简单细胞感受野对明暗对比边缘的敏感性响应；Chan 则用 Gabor 能量函数来表征复杂细胞感受野的相位一致性，在图像纹理特征提取中取得较好的效果^[40]；Young 则提出用二维高斯导函数（Gaussian Derivative, GD）模拟 V1 细胞方位敏感性感受野特性，用来提取图像刺激多方向边界信息，并能用来估计运动场景下运动方向与速度^[41]，其数学表达式如式(2.1)所示：

$$\begin{cases} \text{RF}(x, y; \theta) = \frac{\partial g(\tilde{x}, \tilde{y}; \theta)}{\partial \tilde{x}} \\ g(\tilde{x}, \tilde{y}; \theta) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\tilde{x}^2 + \gamma^2 \tilde{y}^2}{2\sigma^2}\right) \end{cases} \quad (2.1)$$

其中， $(\tilde{x}, \tilde{y}) = (x\cos(\theta) + y\sin(\theta), -x\sin(\theta) + y\cos(\theta))$ ， $\theta \in [0, \pi)$ 为神经元最优选择方位。在式(2.1)中 γ 表示滤波器的椭圆率，本文中取 $\gamma = 0.5$ ， σ 决定了 V1 细胞感受野大小。对于某一特定的方位 $\theta_i \in [0, \pi)$ ，计算如式(2.2)所示，其中 N_θ 表示有 N_θ 个用于提取方位信息的滤波器方向。

$$\theta_i = \frac{(i-1)\pi}{N_\theta}, i = 1, 2, \dots, N_\theta \quad (2.2)$$

2.2.3 非经典感受野

随着电生理实验研究的不断深入,许多研究者开始观察感受野之外的周边区域,对该区域的直接闪光刺激虽然不能引起细胞反应,但却能影响该细胞对感受野内刺激所产生的反应,即调制该细胞的反应性。且研究表明,这种调制作用大多数情况下是抑制的,抑制强度与距中心神经元的距离成反比,这就是非经典感受野(NCRF)。与此相对,将2.2.2节所述中心区域的感受野称为经典感受野(CRF)。Grigorescu等提出用DoG函数经半波整流后的模型(DoG⁺)表征NCRF区域^[13]。DoG⁺模型描述如式(2.3):

$$\text{DoG}^+(x, y) = \max\left\{\frac{1}{2\pi(\rho\sigma)^2} \exp\left(-\frac{x^2 + y^2}{2(\rho\sigma)^2}\right) - \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right), 0\right\} \quad (2.3)$$

式中, σ 决定感受野大小, ρ 表示NCRF的大小相对于CRF的比值,根据生理研究结果,NCRF直径为CRF的2至5倍,本文中设置 $\rho=4$ 。于是,抑制强度可用式(2.4)来表达。

$$W_{\text{nrcf}}(x, y) = \frac{\text{DoG}^+(x, y)}{\|\text{DoG}^+(x, y)\|_1} \quad (2.4)$$

2.3 神经元模型

神经元是神经系统信息处理的基本单元,而动作电位是神经元之间信息传递的“语言”,所以对动作电位的产生和传导过程建模是神经电生理学的基本问题。电生理实验研究表明,动作电位是神经细胞膜内外电位之差。随着跨膜电位和时间的变化,膜上钾、钠等离子通道的通透性会随之变化,并进一步对跨膜电位产生影响。基于这样的实验事实,不少科学家相继提出对神经细胞膜电位变化进行描述的数学模型。HH模型当属最为经典的神经元模型,它高度符合神经电生理特性,从而为理解神经机制奠定了理论基础。而后续又有研究人员提出HH模型的简化模型,在保证基本神经电生理特性的基础上,尽量减少数学计算,从使得的利用真实神经元模型构建大规模神经网络成为可能,比如IF模型。

2.3.1 Hodgkin-Huxley 神经元模型

Hodgkin-Huxley 神经元模型,是Hodgkin和Huxley两位科学家运用数理科学的方法建立的神经元动作电位产生和传导模型。HH模型通过对神经元膜电位与外界激励之间关系的描述,重现了神经元的脉冲发放过程,这其中又涉及钠、钾等离子通道的通透性对膜电位的变化所起的关键作用^[42],所以关系异常复杂。然而,该模型为神经系统细胞层次的研究奠定了数学基础,同时为更高层次的研究创造了条件。HH神经元模型的具体描述如(2.5)所示。其中, I_{ext} 为神经细胞外刺激电流; V 为跨膜电位; C_m 为膜电容; g_{Na} 、 g_{K} 、 g_{L} 为常数,分别表示各离子通道电导的最大值, V_{Na} 、 V_{K} 、 V_{L} 分别表示相应通道的逆向电势; m 和 h 为钠离子通道的通透性描述变量, n 为钾离子通道的通透性描述变量; $m_{\infty}(V)$ 、 $h_{\infty}(V)$ 、 $n_{\infty}(V)$ 和 $\tau_m(V)$ 、 $\tau_h(V)$ 、 $\tau_n(V)$ 分别表示 m 、 h 、 n 的饱和值以及时间常数。图2.3是HH模型描述的细胞膜电位变化示意图,其中一个尖脉冲表示产生了一个动作电位。

$$\begin{cases} C_m \frac{dV}{dt} = I_{\text{ext}} - g_{\text{Na}} m^3 h (V - V_{\text{Na}}) - g_{\text{K}} n^4 (V - V_{\text{K}}) - g_{\text{L}} (V - V_{\text{L}}) \\ \frac{dm}{dt} = \frac{(m_{\infty}(V) - m)}{\tau_m(V)} \\ \frac{dh}{dt} = \frac{(h_{\infty}(V) - h)}{\tau_h(V)} \\ \frac{dn}{dt} = \frac{(n_{\infty}(V) - n)}{\tau_n(V)} \end{cases} \quad (2.5)$$

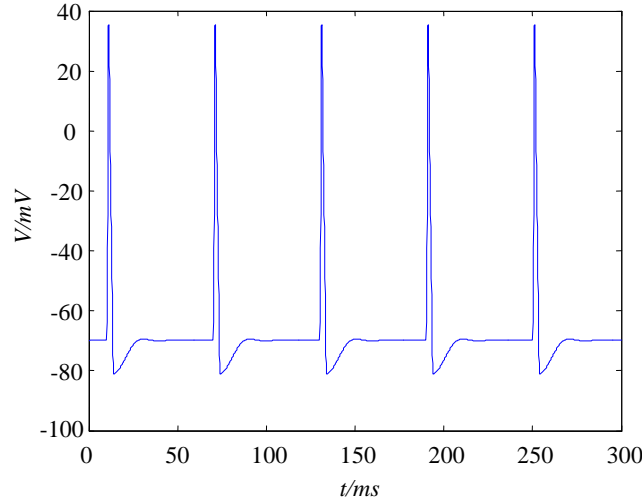


图 2.3 HH 神经细胞膜电位变化示意图

2.3.2 Integrate-and-Fire 神经元模型

IF 神经元模型是以 RC 电路为原型而提出的，它简化了引起细胞膜电位变化的离子运动等诸因素，而仅考虑神经元脉冲发放的特性，通过积分电路为电容蓄电，当达到一定阈值时，动作电位便产生，其数学模型如式(2.6)所示：

$$c_m \frac{dv}{dt} = I \quad (2.6)$$

式中， v 是神经元膜电压， c_m 是膜电容， I 是刺激电流。图 2.4 是 IF 模型描述的细胞膜电位变化示意图，当电位开始下降的前一时刻表示产生了一个动作电位。

基于 IF 模型的基本原理，后来又有学者考虑了更多的神经生理特性，比如漏电流、不应期等，从而提出了 Leaky Integrate-and-Fire 神经元模型，简称 LIF 模型^[43]。LIF 模型能更好地描述神经元放电活动，而且可以根据需要是否加入不应期特性，其数学描述如式(2.7)所示：

$$\begin{cases} c_m \frac{dv}{dt} = -g_l v + I \\ v = v_{\text{reset}} \quad (v > v_{\text{th}}) \end{cases} \quad (2.7)$$

式中， v 、 c_m 和 I 与式(2.6)描述一致， g_l 是漏电导， v_{th} 为脉冲发放阈值， v_{reset} 为静态电势。

当 v 大于 v_{th} 时,神经元脉冲发放,同时 v 被瞬间重置为 v_{reset} 。图 2.5 是 LIF 模型描述的细胞膜电位变化示意图,其中图 2.5(a)为不考虑不应期的情形,图 2.5(b)为考虑了绝对不应期的情形。

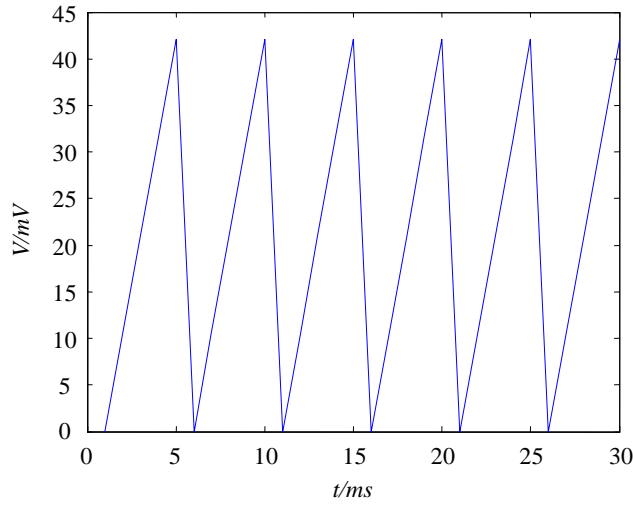


图 2.4 IF 神经细胞膜电位变化示意图

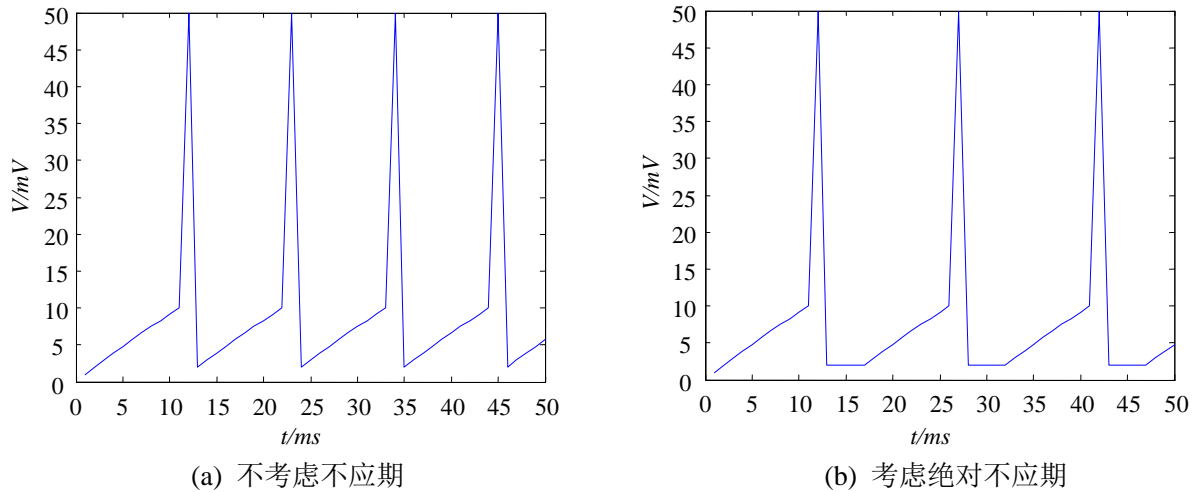


图 2.5 LIF 神经细胞膜电位变化示意图

2.4 脉冲编码

神经编码是神经系统对外界刺激响应的量化描述,神经编码规律的揭示对于大脑信息处理机制及其功能的研究具有极其重要的价值。目前,神经编码生理电实验的研究表明,神经信息处理过程中存在着一些普遍的规则和特性。例如神经元群体编码理论^[44],认为神经元个体对刺激信号的反应具有一定的随机性和不稳定性,只有组织成神经元集群,协同完成信息的编码,才能实现功能。这在一定程度上符合神经细胞感受野机制,也是组建神经元群网络进行智能信息处理的生理基础。再如稀疏编码原则^[45],即同一刺激作用于某一神经元群体,只有部分神经元同时被激活,它们构成了功能神经元集群,并实现从大量复杂信号中提取出有效信息。在视觉通路中,刺激信号经历从低级到高级、简单到复杂层次结构处理过程也是遵循这一规则,从而实现信息的压缩^[46]。在实际应用中,稀疏编码方案能实现大数据的存储和传输,而且计算简单高效,具有记忆功能^[47]。

在视觉系统中，视觉信息经视网膜神经节细胞处理即编码为动作电位，然后经神经纤维向视皮层进行传递形成视觉感知，而信息的传递方式主要为动作电位的传导。一个动作电位表现为一个快速的尖脉冲，它在沿轴突传递的过程中幅值和形状都能保持不变，这种特性保证其能精确保持并传递原始信息，神经元之间就是通过若干个动作电位进行信息的传递，所以动作电位序列中蕴藏着丰富的信息。脉冲编码就是指神经元通过动作电位的发放来传递和表达外界刺激的一系列规则和机制，这对于揭示视觉感知机制以及脑功能具有重要的研究意义。对于脉冲编码的研究，可以设计生物实验来进行记录，然而这仅能考察神经元对特定刺激的放电特性；而更重要的是通过数学和统计方法对实验结果进行分析，才能从此放电特性中解码出神经元所传递的信息。

2.4.1 频率编码

神经元动作电位发放频率是较早被运用于脉冲编码的，而且传统的神经电生理工作，也是通过对神经元施加特定刺激，通过其放电频率的改变来分析其反应特性的^[48]。频率编码理论认为，输入刺激的全部信息被包含于脉冲平均发放率之中，如式(2.8)所示：

$$Q = \frac{1}{T} \sum_{t=1}^T h(t) \quad (2.8)$$

其中， T 为放电活动时间片段，在 t 时刻，神经元发生放电活动，则函数 $h(t)=1$ ，否则 $h(t)=0$ 。

由此可见，频率编码最大的优点是计算和解码方式简便，然而由于该方式对脉冲序列编码的效率很大程度上依赖于时间片段 T ，因而解码速度受到一定的限制，而且在多个参数刺激情形下编码缺乏灵活性。

2.4.2 次序编码

人类视觉系统对视野景象具有快速的反应能力，实验表明视觉刺激从进入人的眼球开始，到视皮层完成解析所需时间不超过 150ms，在如此短的时间内，视觉信号需要经过视通路上多个层次和逐步复杂的结构处理，依靠频率编码这一单一的编码方式恐怕无法完成。另外，视觉感受野内的神经元之间是相互影响的，并通过突触连接形成复杂的神经网络，单个神经元的响应将损失空间特性。于是，有研究人员提出神经系统是以神经元群体为单位进行编码的，而且群体中单个神经元仅发放一次动作电位，以增强神经系统的快速响应能力，所有动作电位发放次序构成了对输入刺激的编码，这就是次序编码理论^[49]。

如图 2.5 所示，6 个神经元 A, B, C, D, E, F，依首次脉冲达到先后顺序构成序列：B>A>F>C>E>D，成为了对此次刺激的编码。而且，对于由 n 个神经元组成的神经元集群，可以展现出 $n!$ 种不同编码模式。相对于频率编码，次序编码方式拥有更快速的编解码速度，而且能表征输入刺激的时空特性，因此能被运用于图像特征的快速提取。

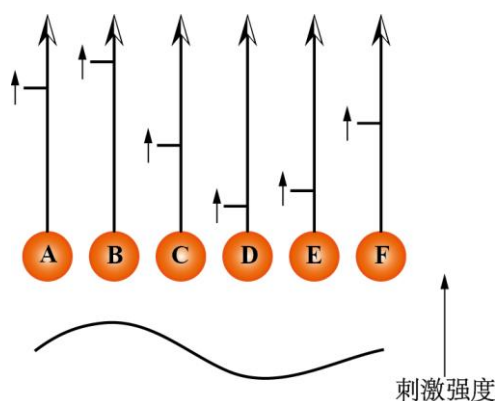


图 2.6 次序编码示意图

2.4.3 时间编码

虽然次序编码考虑了神经元群体活动的时间和空间特性，但是仅以放电次序作为视觉信息的表征未免又损失了大量信息。近年来，大量的研究工作表明，神经元脉冲发放的精确时间模式和相关性，对神经信息编码有着很大的贡献，这就是时间编码理论。例如，若有 3 个神经元: (n_1, n_2, n_3) ，其单次脉冲发放时间为 (t_1, t_2, t_3) 。根据时间编码方式，则 $(t_1 = 0.1, t_2 = 0.3, t_3 = 0.9)$ 和 $(t_1 = 0.29, t_2 = 0.3, t_3 = 0.32)$ 属于两种时间模式，然而次序编码方式则认为这两种模式没有任何区别。

尽管有关神经元动作电位的形成和传播，其细胞膜离子机制和突触连接的方式，生理科学家已给出明确的阐述，但神经脉冲编码的具体方式，至今尚未形成确定的认识，而且编码方式也可能并不唯一。但这却没能阻止人们对其作进一步研究和探索。例如在时间编码理论中，有研究人员提出通过神经元首次脉冲发放的时间来编码输入信息^[50]，弥补了次序编码在精确时间模式上的信息丢失，而又不失时效性。也有研究人员提出用熵编码的方式，考虑前后脉冲之间的时间间隔，进一步增强了时间特征在编码过程中的作用，而且还能表示脉冲序列所携带的信息量^[45]。所以，对脉冲序列编码方式的探究，任重而道远。

2.5 小结

本章主要介绍了视通路上不同生理结构的感受野机制以及基于神经元模型的脉冲编码方式。人类和动物的视觉信息处理主要与第一视觉通路相关，视觉信息沿这条通路由视网膜神经节细胞经 LGN 达到视皮层。神经节细胞和 LGN 神经元的感受野呈同心圆拮抗式，分为 ON-中心和 OFF-中心两种形式，一般认为该感受野结构可用不同尺度的高斯函数之差 (DoG) 来计算模拟。神经节细胞和 LGN 神经元感受野中心-外周的拮抗作用方式，提供了视觉系统感知空间对比度分辨的能力，为形成边界和形状知觉准备了条件。初级视皮层 CRF 呈长形，中心狭长的 ON 区域能对条形光刺激有更强烈的反应，故被认为具有方位敏感特性，可以用二维高斯导函数进行描述；而在 CRF 的外周更大的区域，即 NCRF，对该区域的给光刺激能对中心神经元的反应起到调制的作用，可用 DoG^+ 模型进行表述，目前对该感受野机制的研究大多集中于利用其对 CRF 反应的抑制，提高了 V1 神经元对大范围的空间信息的整合能力。而

神经元是神经信息处理的基本单元，本章主要介绍了 HH 模型和 IF 模型。HH 模型对动作电位的产生进行了完整的描述，但计算复杂，不适合工程应用；而 IF 模型对 HH 模型进行了简化，在保留神经元基本生理特性的基础上，尽量减少计算，可用于组建大规模神经元网络。最后，对视觉信息以脉冲序列在视通路上传导的编码方式进行了介绍。频率编码方式简单，但解码速度慢，不够灵活；次序编码方式考虑了神经元群体活动的单次脉冲，速度快，却丢失了时间信息；时间编码模式多样，主要依靠脉冲发放的精确时间，具体实现方式有首次脉冲时间编码、熵编码等。

第3章 基于感受野特性的视觉机制建模

如绪论所述，视觉信息处理的相关机制得到不断研究和揭示，例如方位敏感机制对视觉轮廓及形状知觉的形成提供了重要的特征信息，侧向抑制机制能检测空间变化加强反差，视觉注意机制能够迅速将注视点转移到视野中感兴趣的区域，弱化无用信息，强化有用信息，等。计算科学家在此基础上开发出大量新的计算原理和信息智能化处理的方法，但这些原理和方法大多是通过黑箱形式的数学模型对视觉刺激进行滤波处理，忽略了视觉信息在视通路中以脉冲序列进行传递的生理基础，往往会丢失脉冲发放的时间编码信息，所以本章以突触连接的形式构建新的感受野模型，通过神经元将视觉刺激转化为脉冲序列，通过对脉冲序列的编码解析视信息所包含的特征。以明暗刺激施加于本章提出的神经元感受野模型，为实现通过神经元的响应表征刺激特征、实现相应的视觉机制提供了新的思路。

3.1 概述

本章针对生物视觉系统普遍存在的方位敏感机制、侧向抑制机制以及视觉注意机制，考虑了其在视觉信息处理过程中的功能，分别以图 3.1 所示的视觉刺激施加于视觉感受野，该视觉刺激尺寸为 100×100 。首先构建同等规模的神经元阵列，每个神经元与其感受野范围内的其他神经元通过突触连接，连接权重构成了该神经元感受野的突触连接权值矩阵，而权值大小以及作用方式随视觉机制的不同而不同，通过对神经元输出脉冲频率、时间等编码实现初步的视觉功能。这将为后续基于视觉机制在图像处理上的应用奠定基础。

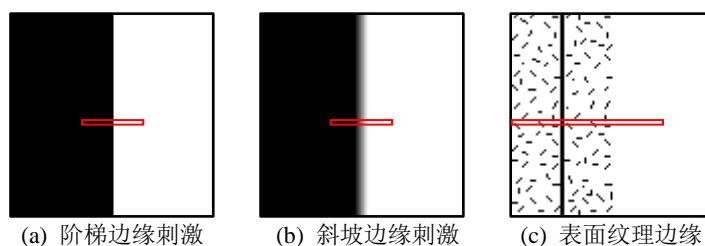


图 3.1 视觉刺激

3.2 方位敏感机制建模

感受野是视觉系统分辨空间对比度、感知明暗变化的神经生理基础，不管是神经节细胞的同心圆拮抗式感受野，还是初级视皮层 V1 狭长形简单细胞，对于视觉感知边界方位、轮廓线条以及目标形状等提供了重要的信息。本节以图 3.1(a)所示阶梯边缘刺激施加于感受野上，针对此方位的边缘，设计如图 3.2 所示的感受野模型，图中蓝色表示该神经元与中心神经元为抑制性连接，黄色表示该神经元与中心神经元为兴奋性连接，连接强度为距离的高斯分布函数，如式(3.1)所示：

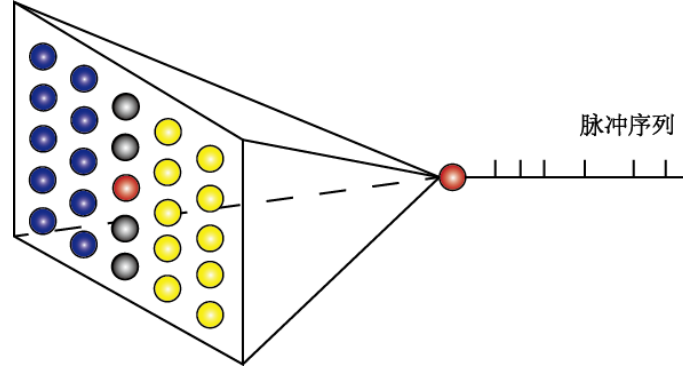


图 3.2 感受野模型

$$w_{\text{syn}} = w_{\text{max}} \exp\left(-\frac{\Delta d^2}{2R_{\text{Fr}}^2}\right) \quad (3.1)$$

式中， w_{max} 为作用强度的最大值， Δd 为两神经元间空间距离， R_{Fr} 是感受野半径。设置 $w_{\text{max}} = 1.0$ ，于是由式(3.1)计算得出的感受野突触连接矩阵 \mathbf{W}_{syn} 如式(3.2)所示。

$$\mathbf{W}_{\text{syn}} = \begin{bmatrix} -0.37 & -0.54 & 0 & 0.54 & 0.37 \\ -0.54 & -0.78 & 0 & 0.78 & 0.54 \\ -0.61 & -0.88 & 0 & 0.88 & 0.61 \\ -0.54 & -0.78 & 0 & 0.78 & 0.54 \\ -0.37 & -0.54 & 0 & 0.54 & 0.37 \end{bmatrix} \quad (3.2)$$

中心神经元接受感受野内其他神经元的突触影响，将融合后的信息输入式(2.7)所示的 LIF 神经元模型，记录其在一定时间内的脉冲响应。为便于观察感受野对亮暗边界的响应情况，记录高度为 50 自左向右序号为 40~60 共 20 个像素点（如图 3.1(a)中红色标识所示），所对应神经元响应的脉冲序列，如图 3.3 所示，其中横坐标为刺激呈现时间，纵坐标为自左向右的神经元序号，一个圆点表示一次脉冲发放。从图中可以看出，处于亮暗边界处的神经元对视觉刺激变化有着强烈的感知能力。将神经元脉冲发放频率映射为灰度值，可得到感受野对亮暗边界的感知结果，如图 3.4 所示。

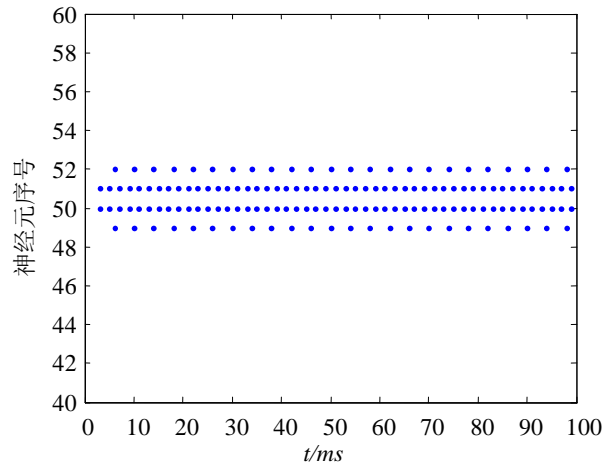


图 3.3 阶梯边缘脉冲响应光栅图

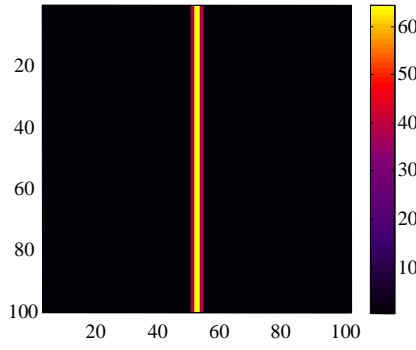


图 3.4 脉冲发放频率图

从图 3.3 与图 3.4 神经元脉冲响应情况可以看出，图 3.2 所示的感受野模型能感知方位为左暗右明的视觉刺激边界，而将该模型中兴奋与抑制性突触连接对换，得到感受野突触连接矩阵 \mathbf{W}_{syn} 如式(3.3)所示。

$$\mathbf{W}_{\text{syn}} = \begin{bmatrix} 0.37 & 0.54 & 0 & -0.54 & -0.37 \\ 0.54 & 0.78 & 0 & -0.78 & -0.54 \\ 0.61 & 0.88 & 0 & -0.88 & -0.61 \\ 0.54 & 0.78 & 0 & -0.78 & -0.54 \\ 0.37 & 0.54 & 0 & -0.54 & -0.37 \end{bmatrix} \quad (3.3)$$

同样，将突触作用后的中心神经元电流输入 LIF 模型，记录两种模型作用下明暗边界处坐标为(50,50)的神经元膜电位变化，如图 3.5 所示，其中膜电位超过 10mV 时，神经元发放一次脉冲。读图可以看出，对换后的感受野模型不具备感知左暗右明的边缘刺激，说明虽然同样是垂直边缘，但对于不同的方位，设计随方位变化的感受野模型能有效提取视觉刺激中的亮暗对比边缘，这对于后续通过方位敏感性感受野模型实现图像边缘提取的应用打下了基础。

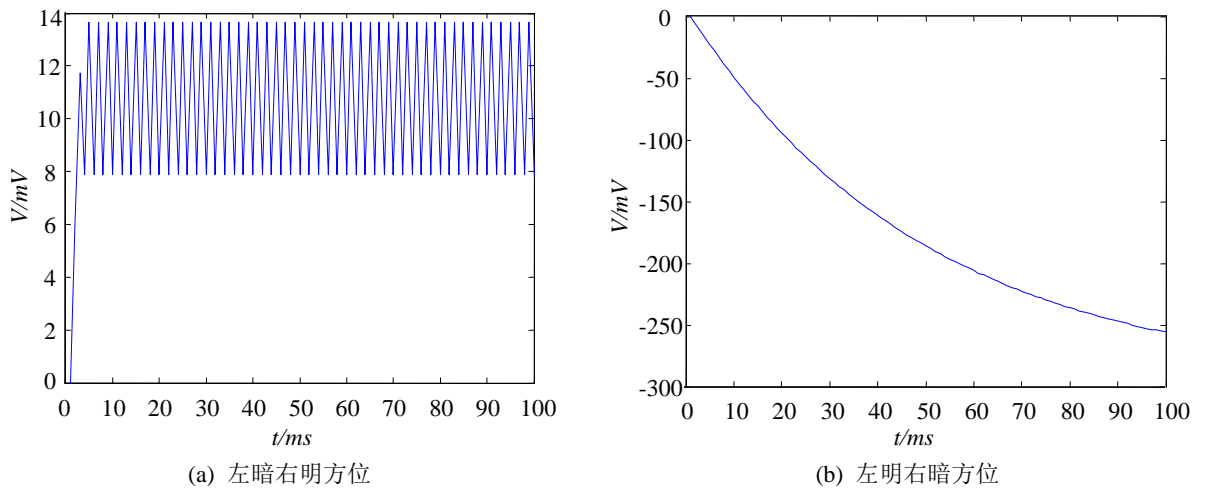


图 3.5 两种方位突触连接矩阵作用下神经元膜电位变化

3.3 侧向抑制机制建模

实际中的数字图像边缘大都存在模糊和噪声，并不如图 3.1(a)所示那般黑白分明，在这种

情况下，视觉刺激可表征为图 3.1(b)所示的斜坡边缘。用图 3.2 所示方位性感受野模型处理该刺激，记录高度为 50 自左向右序号为 40~60 共 20 个像素点（如图 3.1(b)中红色标识所示），所对应神经元响应的脉冲序列，如图 3.6 所示。由图可以看出，斜坡边缘区域均存在刺激的变化，神经元脉冲响应剧烈，容易出现粗边缘，如图 3.7(a)所示脉冲发放频率图，这在一定程度上给边缘准确定位带来麻烦。

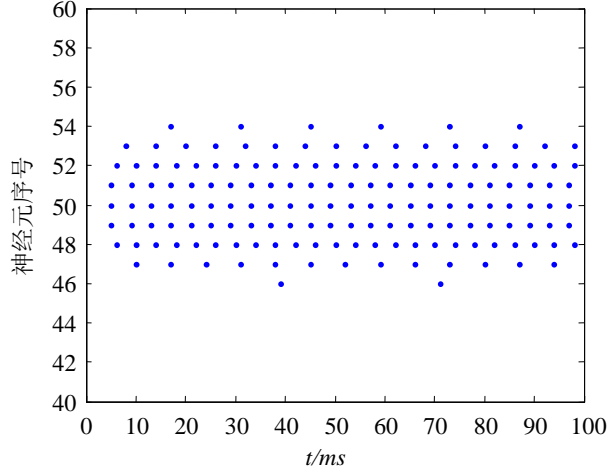


图 3.6 斜坡边缘脉冲响应光栅图

考虑到侧向抑制机制在加强对比度反差，突出边界的视觉感知作用，即感受野内的神经元处于一种竞争状态，脉冲发放时间早的神经元会抑制发放晚的神经元，相当于会对其产生一种延迟脉冲发放效应^[24]。本节试图对侧向抑制机制进行简化建模，以解决模糊边缘带来的真实边缘定位难的问题。

为简化计算，以脉冲发放频率作为神经元活动量，忽略侧向抑制递归的作用形式，仅考虑感受野内活动量比中心神经元大的神经元抑制作用。一般认为侧向抑制作用大小受神经元空间距离、神经元放电活动两方面要素影响。当神经元空间距离越大，突触连接强度越小，侧向抑制量也越小；而神经元放电频率差越小，表明该处存在冗余边缘的可能越大，故抑制量也越大。以此定义侧向抑制量计算表达式如式(3.4)所示：

$$w_{inh} = \exp\left(-\frac{\Delta d^2}{2R_{Fr}^2} - \frac{1}{4} \sqrt{\frac{r_{cen}}{r_{sur}}}\right) \quad (3.4)$$

其中，式中， Δd 为两神经元间空间距离， R_{Fr} 是感受野半径， r_{cen} 和 r_{sur} 分别是中心和周围神经元脉冲发放频率。经侧向抑制调节后的脉冲发放频率图如图 3.7(b)所示。从图中可以看出，根据公式(3.4)所定义的侧抑制网络模型能有效去除冗余边缘响应，精细并准确定位真实边缘，这有利于后续通过侧向抑制机制去除噪声和模糊造成的虚假边缘，实现弱边缘检测。

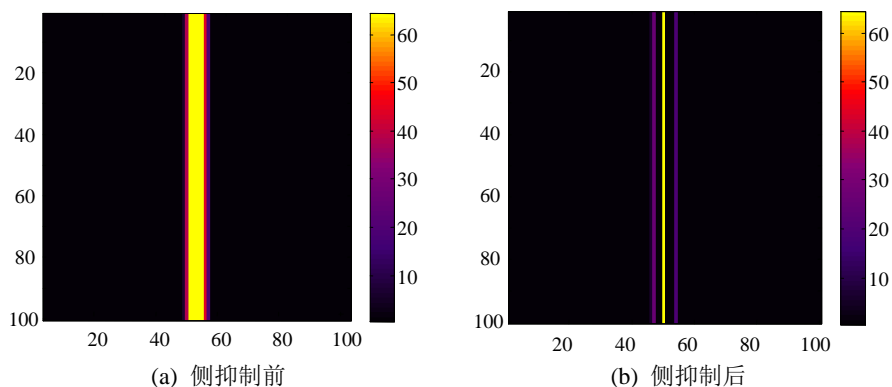


图 3.7 侧抑制作用前后脉冲发放频率图

3.4 视觉注意机制建模

轮廓以及形状感知是视觉系统中比明暗对比刺激感知更加高级的视觉任务，这要求视觉系统不仅能感知亮暗刺激的方位变化，而且能根据时空信息辨识视觉刺激中的表面纹理和主体轮廓，这有助于视觉系统从复杂环境中锁定目标。用图 3.1(c)所示的带有表面纹理的边缘表征该情形，同样经图 3.2 所示方位性感受野模型处理该刺激，记录高度为 50 自左向右序号为 0~80 共 80 个像素点（如图 3.1(c)中红色标识所示），所对应神经元响应的脉冲序列，如图 3.8 所示。从图中可以看出，带有表面纹理的区域与平面区域的边界被埋没在纹理本身的响应之中，无法分辨，这对于感知图像轮廓，即主体目标与背景或平面与平面的边界，造成一定的障碍。

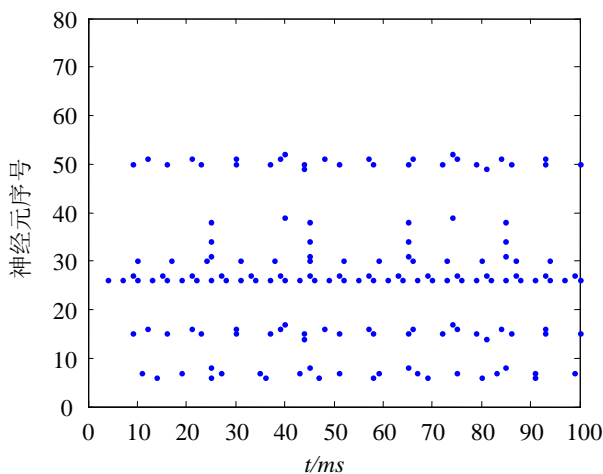


图 3.8 表面纹理边缘脉冲响应光栅图

在面对复杂的视野场景时，人的大脑会将注意力会更多的关注到相对重要的目标或者区域，视觉系统正是通过这样的注意机制来提高感知相关区域和环境特征神经元的活动，进而提取有效信息，这同时也是大脑能迅速反应眼前事件、优化判断并做出正确行为活动的前提^[51]。对于视觉注意机制的研究主要从两方面展开：一是通过自底向上（bottom-up）的方法，提取场景的方位、颜色、纹理及其他底层属性融合得到显著性图像^[32]；二是自顶向下（top-down）

的方法，通过高级脑区的反馈控制，有意识地将注意力转移到显著性区域^[52, 53]。本节通过计算局部空间特征，自底向上提取底层图像属性。考虑到非经典感受野对经典感受野中心神经元活动的调节作用，从视觉感知特性出发，计算 CRF 和 NCRF 区域内视觉输入的差异性，从而形成显著性图像。

本文在空间对比度特征计算时仅考虑了中心 ON 周围 OFF 的情形，即 CRF 对中心神经元作用为增强，而 NCRF 作用为抑制，计算方法为式(3.5)所示：

$$\Delta\phi(x, y) = \lambda \times \max\{\bar{\phi}_C(x, y) - \bar{\phi}_S(x, y), 0\} \quad (3.5)$$

式中， λ 定义为 $\bar{\phi}_C(x, y) / \bar{\phi}_S(x, y)$ ，用于进一步凸显差异， $\bar{\phi}_C(x, y)$ 和 $\bar{\phi}_S(x, y)$ 分别为 CRF 和 NCRF 的空间视觉输入，该输入分别以高斯函数和 DoG⁺模型（公式(2.3)）为距离因素进行量化：

$$\bar{\phi}_C(x, y) = \sum_{(m,n)} I(x+m, y+n) \times g(m, n) \quad (3.6)$$

$$\bar{\phi}_S(x, y) = \sum_{(m,n)} I(x+m, y+n) \times \text{DoG}^+(m, n) \quad (3.7)$$

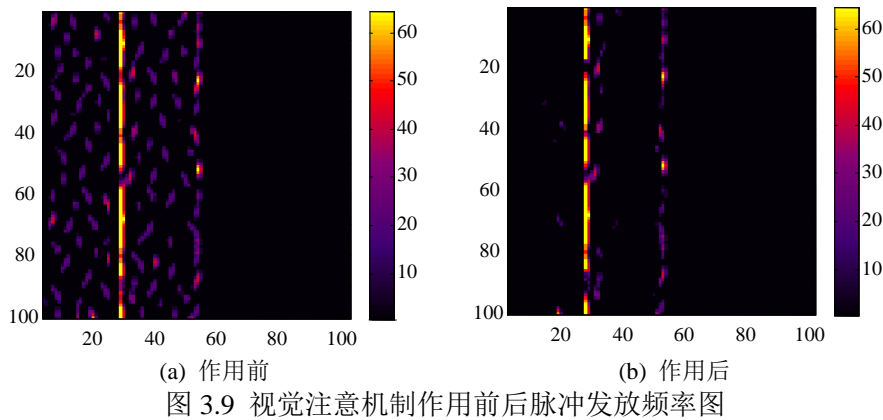
其中， $I(x, y)$ 表示视觉刺激， $g(m, n)$ 表示高斯函数。

最后，视觉输入的差异性经过幂指数函数归一化，形成的显著性图像，计算如式(3.8)所示：

$$S(x, y) = \exp\left(-\frac{\Delta\phi(x, y)^2}{2\delta^2}\right) \quad (3.8)$$

式中，参数 δ 为抑制强度调节参数，设置为 0.2。

如图 3.9 所示，为自底向上视觉注意模型作用前后脉冲发放频率图。从视觉注意图像可以看出，纹理边缘得到较好的抑制，而强弱边缘得到保留，符合视觉认知过程中将注视转移到有明显特征变化的区域的特性，同时对于后续通过计算空间差异性信息来提取图像主体轮廓的应用奠定了基础。



3.5 小结

本章针对数字图像处理中普遍存在的噪声、模糊边缘以及轮廓提取等困难，结合生物视觉系统方位敏感机制、侧向抑制机制以及视觉注意机制，通过组建能实现其在视觉信息处理之功能的网络模型，并以简单的明暗边缘、斜坡边缘和纹理边缘等视觉刺激进行仿真处理，以神经元脉冲响应来表征模型在解决上述图像处理困难中的表现，结果表明本章所提模型在方位感知、边缘定位以及纹理抑制中能达到预期的效果，实现了视觉感知的初步功能，为后续基于视觉感知计算模型在图像处理上的应用奠定基础。

第4章 基于突触连接视通路方位敏感的图像分级边缘检测

人类视觉信息处理通路上不同的组织结构对边缘信息有着不同层次的感知能力。视皮层下视网膜神经节细胞以及外膝体对物体轮廓线段的感知是大脑形状知觉的神经基础；初级视皮层对特殊朝向的条形光刺激有着更为强烈的反应，同时为后续目标识别与行为理解等高智能行为进行有效的信息筛选。另外，根据前一章对方位敏感机制、侧向抑制机制的建模研究，表明该模型基本能实现相应视觉机制在边缘信息感知中所起的作用。因此，本章从视觉方位敏感机制出发，构建具有神经电生理特性的神经元网络，引入更具有针对性的突触动态连接方式，以及改进的侧向抑制机制，实现图像的边缘检测。以层次模糊而细节丰富的菌落图像为例，通过重构相似度和边缘置信度对检测结果进行有效性评价，评价结果显示本文方法能在边缘漏检和噪声误检之间达到较好的平衡，对于传统检测方法和其他视觉感知机制建模方法具有明显的优势^[54]。

4.1 概述

本章提出一种基于突触连接视通路方位敏感机制的分级轮廓检测新方法，首先考虑视皮层下视网膜神经节细胞和外膝体的方位敏感特性，探讨其感受野最优方位在视网膜上的分布规律，利用神经元突触间的动态连接特性构建视皮层下多方位敏感层；再结合初级视皮层更强烈的特殊方位条形光刺激感知能力，对上一功能层输出的脉冲序列编码，并对首发脉冲时间信息加以侧向抑制和阈值处理，得到图像边缘。

4.2 视通路分级模型

视信号自视网膜成像形成电脉冲信号开始，直至到达视皮层完成对信号的感知与识别，视觉通道上不同的组织结构对图像轮廓特征的形成均有着不同层次的贡献。视皮层下神经节细胞和外膝体具有方位敏感特性，能感知像素梯度变化，而初级视皮层的侧向抑制作用能对脉冲信息进行有效筛选。通常认为，视网膜神经节细胞和外膝体神经元的方向敏感性感受野有着相似的结构。这里为了简化模型，本文将神经节细胞层和外膝体层的功能合并为一层（统称为视皮层下）来对视觉输入进行处理。于是可以将视通路分为视皮层下和初级视皮层两层功能作用区，构建具有边缘特征提取能力的分级模型，如图 4.1 所示。

4.2.1 视皮层下方位敏感

视网膜神经节细胞以及外膝体的方位敏感特性对视皮层精细的轮廓线段感知起着重要的前级安排作用^[55, 56]。利用神经元个体之间的动态突触连接特性，构建特定方位敏感的感受野模型，最终形成具有神经元群体边缘响应效应的网络，并对视觉输入进行差异性检测，将视觉激励映射为携带有边缘强度信息的脉冲序列。

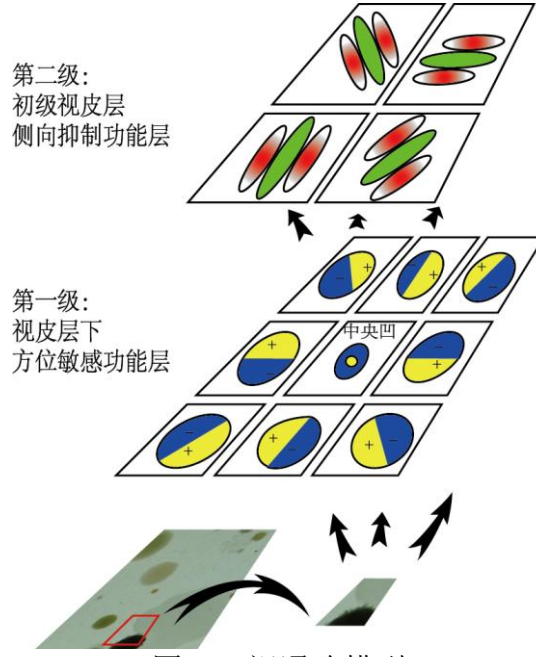


图 4.1 视通路模型

4.2.1.1 动态突触连接

神经元是神经网络的基本组成单元，而两个神经元发生相互作用直至形成神经元群体的互连结构称为突触。本研究选择既具备神经元电生理特性，并且计算简单高效的基于漏积分放电(LIF)模型来构建神经网络。结合视觉系统感受野特性，将感受野内通过突触相连的外界激励综合作为模型输入，同时，由于后期对脉冲的编码只考虑了首次脉冲到达时间，所以去除绝对不应期。改进后的模型如式(4.1)所示：

$$\begin{cases} c_m \frac{dv}{dt} = -g_l v + \sum_{(x,y) \in \mathbf{RF}} w_{x,y} I_{x,y} \\ v = v_{\text{reset}} \quad (v > v_{\text{th}}) \end{cases} \quad (4.1)$$

式中， \mathbf{RF} 是神经元感受野， $w_{x,y}$ 是突触连接权值， $I_{x,y}$ 是激励电流，对应图像像素 (x, y) 处的归一化值， v 、 c_m 、 g_l 、 v_{th} 、 v_{reset} 以及脉冲发放原理与公式(2.7)中所述一致。

突触在神经元与神经元之间起着信息传递的作用，神经生理学表明，这种突触传递的连接权值是动态变化的，这在一定程度上提高了神经系统对突变刺激的响应效率，而且还能充分体现空间差异信息^[57]。本文在神经网络中引入动态突触作用时，改进公式(3.1)仅考虑距离因素的方式，而引入输入激励差异对突触连接强度的影响，具体表现为感受野内突触连接强度不仅会随着距离的增大而减弱，且随着激励差异性增大而增强，具体如式(4.2)所示：

$$|w_{x,y}| = w_{\text{max}} \exp \left[- \left(\frac{|P_{x,y} - P_{x_c, y_c}|^2}{2\mathbf{R}\mathbf{F}r^2} - \frac{|I_{x,y} - I_{x_c, y_c}|^2}{2\Delta I^2} \right) \right] \quad (4.2)$$

式中， $|P_{x,y} - P_{x_c, y_c}|$ 与 $|I_{x,y} - I_{x_c, y_c}|$ 分别是周围神经元与中心神经元空间距离和激励差异， $\mathbf{R}\mathbf{F}r$ 是感受野半径， ΔI 是感受野内最大刺激与最小刺激之差， w_{max} 为权值调整常数。同时，根据突触连接功能特性，上述权值会形成兴奋与抑制两种状态，如式(4.3)所示。从而突触连接的动

态性表现为根据感受野的方位不同而呈现出兴奋与抑制两种状态之间的切换。

$$\begin{cases} w_{x,y}^{\text{ex}} = |w_{x,y}| \\ w_{x,y}^{\text{ih}} = -|w_{x,y}| \end{cases} \quad (4.3)$$

4.2.1.2 方位敏感性感受野

一直以来，人们都认为只有视皮层才具有方位以及方向选择特性。而近年来神经生理学研究揭示，视皮层下视网膜神经节细胞以及外膝体同样具有此特性^[58]，而且大多数神经元的最优方位平行于其感受野中心和视网膜中央区的连线，即呈现向心分布规律^[59]，如图 1 的第一级所示。根据此分布规律，选择围绕视网膜中央凹 8 个方位 ($0^\circ, 45^\circ, 90^\circ, 135^\circ, 180^\circ, 225^\circ, 270^\circ, 315^\circ$) 上的感受野，形成四类边缘（中心水平、中心垂直、正、负对角线），并以最优方位为界，将感受野分为两个分别具有兴奋和抑制突触连接作用的区域，以实现视皮层下细胞对特定方位亮度对比线段的选择作用。以中心水平边缘（即 0° 和 180° ）为例，它会产生响应方向为向上（下暗上亮）和向下（下亮上暗）两种类型的感受野，其数学模型如式(4.4)和(4.5)所示：

$$w_{x,y}^0 = \begin{cases} w_{x,y}^{\text{ex}} & ((y - y_c) > 0) \\ 0 & ((y - y_c) = 0) \\ w_{x,y}^{\text{ih}} & ((y - y_c) < 0) \end{cases} \quad (4.4)$$

$$w_{x,y}^{180} = \begin{cases} w_{x,y}^{\text{ih}} & ((y - y_c) > 0) \\ 0 & ((y - y_c) = 0) \\ w_{x,y}^{\text{ex}} & ((y - y_c) < 0) \end{cases} \quad (4.5)$$

其中 (x_c, y_c) 是感受野 **RF** 中心神经元坐标， $(x, y) \in \mathbf{RF}$ 。

依次类推可得到其他类型边缘的感受野数学模型。

4.2.2 初级视皮层侧向抑制

视觉输入经由视皮层下结构的处理，将视觉激励转化为脉冲序列供后续视觉结构处理。初级视皮层细胞接受前级结构生成的脉冲序列，以一定的形式对其进行编码。考虑到信息传递的有效性，本文选择以首次脉冲发放的时间作为编码方式^[50]，结合侧向抑制机制达到边缘的精确提取。

前级结构以首次脉冲时间作为输出，并映射成灰度等级，得到一幅边缘敏感图像；对边缘敏感图像上每一点 (x, y) ，选取 8 方位上最大响应所对应的边缘类别，角度记为 β ，例如边缘方位为中心水平，则 $\beta = 0^\circ$ ， $\beta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ ，作为该点的最优方位，即：

$$D_{x,y} = \beta \quad (4.6)$$

从而得到与该敏感图像对应的最优方位矩阵。

依据最优方位矩阵，对敏感图像引入侧向抑制作用，这有利于准确定位真实边缘、增加边缘精细程度^[25]。另外对 3.3 节所述侧向抑制模型的作用范围加以改进，采用如图 4.2 所示的

去最优方位侧向抑制方式。以 3×3 大小神经元感受野 \mathbf{RF} 为例，将首次脉冲时间比感受野中心早的神经元定义为先发状态，反之定义为后发状态。侧向抑制的方式为：最大响应方位（垂直于最优方位 \mathbf{OF} ）上的神经元若均处于后发状态（如图 4.2 中右分支所示），那么该神经元正常输出；反之，最大响应方位上只要存在一个先发状态的神经元（如图 4.2 中左分支所示），则该神经元要受到去最优方位感受野内所有处于先发状态神经元的抑制作用。

另外，对公式(3.4)所定义的侧向抑制量引入最优方位的影响，当感受野内最优方位越相似时抑制量越大，故改进后的侧向抑制量定义如下：

$$w_{x,y}^{vc} = \begin{cases} \exp \left[- \left(\frac{\Delta d^2}{2R_{Fr}^2} + \frac{1}{4} \sqrt{\frac{T_{x_c, y_c}}{T_{x, y}}} + \frac{\theta^2}{2\delta^2} \right) \right] & ((x, y) \in (\mathbf{RF} - \mathbf{OF})) \\ 0 & ((x, y) \in \mathbf{OF}) \end{cases} \quad (4.7)$$

式中， Δd 和 R_{Fr} 依然表示神经元间空间距离和感受野半径， $T_{x, y}$ 和 T_{x_c, y_c} 分别是周围和中心神经元首次脉冲到达时间， δ 是最优方位最小差，等于 45° ， θ 是两神经元最优边缘方位的角度差，定义为：

$$\theta = \min \{ |\sigma_{x, y} - \sigma_{x_c, y_c}|, 180^\circ - |\sigma_{x, y} - \sigma_{x_c, y_c}| \} \quad (4.8)$$

其中， $\sigma_{x, y}$ 和 σ_{x_c, y_c} 分别是周围和中心神经元最优边缘方位所对应的角度。

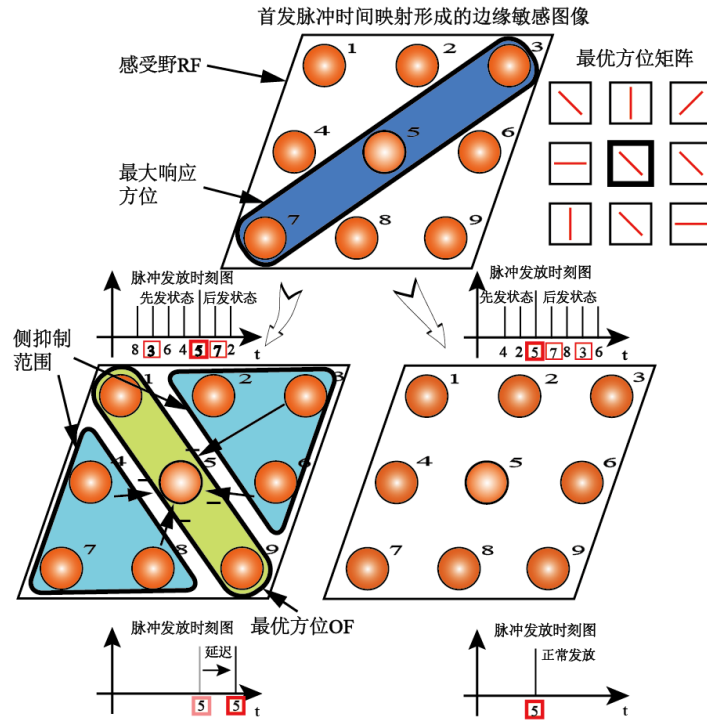


图 4.2 去最优方位侧向抑制示意图

4.3 算法及算法验证

4.3.1 算法

利用动态突触构建的多方位敏感型感受野对输入图像进行边缘强度检测，以首发脉冲时间进行编码，映射形成边缘敏感图像；然后引入改进的侧向抑制方法，对敏感图像进行有选

择性的抑制作用，再经过阈值处理形成最终边缘。程序执行流程如图 4.3 所示，具体检测过程如下所述。

步骤 1：根据原始图像 $I_o(x, y)$ ($x=1, 2, \dots, M; y=1, 2, \dots, N$) 大小构建相同尺寸的神经网络 $GC(x, y)$ ，各个神经细胞在其感受野 $RF_{x,y}^{gc}$ (5×5) 内以突触的形式与周围神经元互连，依式(4.2)计算突触连接权值 $w_{x,y}$ 的大小。

步骤 2：根据式(4.4)、(4.5)以及其他方位边缘的感受野数学模型，将步骤 1 中得到的权值 $w_{x,y}$ 赋予兴奋和抑制功能，形成对 8 个对不同方位敏感的感受野内突触连接权值，记为 $w_{x,y}^\alpha$ ， $\alpha = 0, 45, 90, 135, 180, 225, 270, 315$ 。

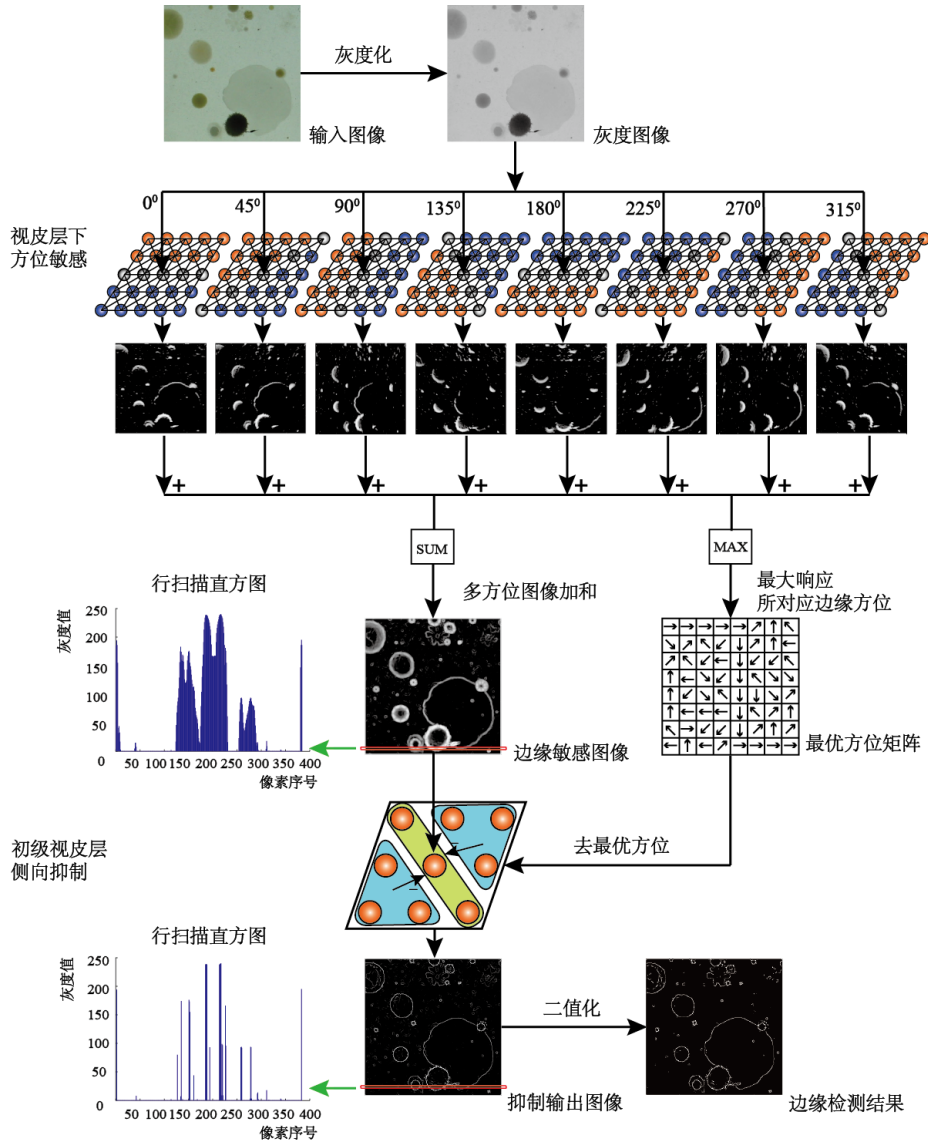


图 4.3 边缘检测算法流程图

步骤 3：对每个神经细胞 (x, y) 采用式(4.1)所示的 LIF 神经元模型，并将原始图像像素值归一化，以激励电流的形式输入神经元，分别将步骤 2 获得的 8 个突触连接权值代入该模型。

步骤 4：记录神经网络 $GC(x, y)$ 在不同类型的方位敏感性感受野作用下，各个神经元首次脉冲达到时间，从而得到 8 个时间矩阵 $T_{1st}^\alpha(x, y)$ 。

步骤 5: 利用式(4.9)融合多方向上的时间矩阵, 并映射成灰度等级, 得到视网膜前级安排作用下的边缘敏感图像:

$$\mathbf{IS}(x, y) = (1 - \frac{\sum_{\alpha} T_{1st}^{\alpha}(x, y)}{T_{max}}) \times 255 \quad (4.9)$$

式中, T_{max} 是融合后的时间矩阵 $\sum_{\alpha} T_{1st}^{\alpha}(x, y)$ 的最大值。另外, 根据式(4.6)计算得到与原始图像尺寸相同的最优方位矩阵 $\mathbf{D}(x, y)$ 。

步骤 6: 再次构造初级视皮层细胞感受野 $\mathbf{RF}_{x,y}^{vc}(3 \times 3)$, 由图 4.2 所示规则确定神经元是否受侧抑制作用及其在受到侧向抑制时的作用范围, 再根据式(4.7)和式(4.8)计算侧抑制系数 $w_{x,y}^{vc}$, 然后由式(4.10)得到视皮层精细边缘提取结果 $\mathbf{IT}(x, y)$:

$$\mathbf{IT}_{x,y} = \mathbf{IS}_{x,y} - \sum_{(x,y) \in \mathbf{RF}_{x,y}^{vc}} w_{x,y}^{vc} \times \mathbf{IS}_{x,y} \quad (4.10)$$

其中 $\mathbf{IT}_{x,y}$ 和 $\mathbf{IS}_{x,y}$ 分别是 $\mathbf{IT}(x, y)$ 和 $\mathbf{IS}(x, y)$ 在 (x, y) 处的值。

步骤 7: 对步骤 6 的结果进行阈值处理得到最终边缘图 $\mathbf{IE}(x, y)$:

$$\mathbf{IE}_{x,y} = \begin{cases} 1, & \mathbf{IT}_{x,y} > I_{th} \\ 0, & \mathbf{IT}_{x,y} \leq I_{th} \end{cases} \quad (4.11)$$

其中, $\mathbf{IE}_{x,y}$ 是 $\mathbf{IE}(x, y)$ 在 (x, y) 处的值, I_{th} 是设定阈值。

4.3.2 算法验证

菌落图像的边缘检测是用图像方法分析菌落的基本步骤之一, 而菌落图像所具有的细节丰富、层次模糊、形态各异等特征使得对其进行边缘检测存在一定难度。本文算法有着准确定位强弱边缘、抑制纹理噪声以及保护弱小细节等优点, 从原理上能较好地克服上述困难, 选取本实验室常规微生物实验需要而采集的 12 幅典型菌落图像作为实验用图, 部分图像如图 4.4 所示。另外, 选取 Canny、具有方向选择特性的 Gabor 滤波器方法以及文献[22]的方法(MCI)与所提出算法进行效果比较, 验证本算法的有效性。其中 Gabor 滤波器选取了 16 方向, MCI 方法中 Gaussian 函数设置为 8 方向。

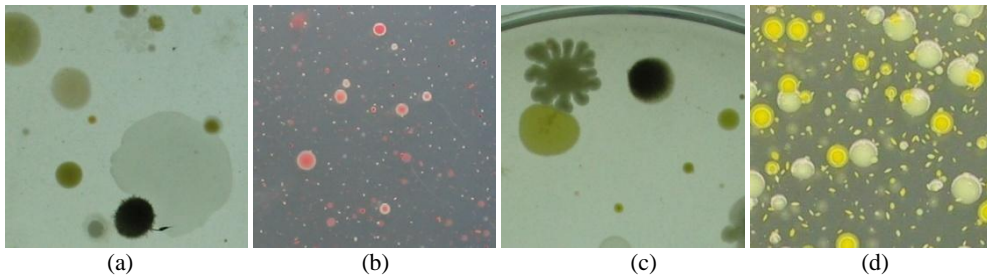


图 4.4 菌落示意图。(a) 多强度边缘菌落; (b) 外光环深层菌落; (c) 形态丰富菌落; (d) 粘连菌落

为客观的对检测结果进行评价, 利用边缘检测结果对原图进行重构^[60], 再计算重构图像相似度指标 $\mathbf{MSSIM}^{[61]}$ 和边缘置信度指标 $\mathbf{BI}dx^{[62]}$, 从而定量验证本文算法的有效性。同时, 考虑到 \mathbf{MSSIM} 和 $\mathbf{BI}dx$ 在一定程度上可以形成互补关系, 且二者均依赖于边缘像素点数量 $\mathbf{EdgeNum}$, 所以, 以两者的加权和作为综合指 $\mathbf{EI}dx^{[63]}$, 评价不同算法的优劣程度, 即为:

$$\begin{cases} EIdx = \omega_s \times MSSIM + \omega_b \times BIdx \\ \omega_s = k_s \times (1 - \frac{EdgeNum}{M \times N}) \\ \omega_b = 1 - \omega_s \end{cases} \quad (4.12)$$

式中， $M \times N$ 是图像大小，由于本文菌落图像中弱边缘以及细节信息较多，所以 $0.5 < k_s < 1$ ，这里取 $k_s = 0.6$ 。

最后，对所选取的 12 幅菌落图像的 EIdx 指标作 t 检验，将上述 3 种对比方法分别与本文方法进行比较，P 值小于 0.05 时被认为具有显著性差异。

4.4 实验结果

图 4.5~图 4.8 分别是对图 4.4 所示 4 幅菌落图进行边缘检测后的结果。观察这些结果图，可以发现本方法对图示菌落图像具有更好的边缘信息提取能力。各子图(b)所示的 Canny 方法检测结果，当菌落图像的丰富细节信息得到很好的检测时，会出现大量由纹理噪声形成的伪边缘，例如图 4.5(b)、图 4.6(b)和图 4.7(b)均出现这样的现象，说明 Canny 方法很难在弱边缘和纹理噪声之间达到平衡；如各子图(c)所示的 Gabor 滤波器方法检测结果对小的菌落会存在边缘断裂、不完整现象，这在一定程度上又增加了虚假边缘，如在图 4.6(c)和图 4.8(c)中均有所体现；再如各子图(d)所示的 MCI 方法检测结果，虽然纹理噪声得到了很好的抑制，但基本上只检测出了菌落的大致轮廓，细节丢失较为严重，如图 4.7(d)触角形态的边缘很多都连在了一起；相比较之下，各子图(e)所示的本方法对各类菌落图像都具有相对完整的检测效果，不仅能在不造成虚假边缘的条件下对多强度边缘图像（如图 4.5(e)）提取有效边缘轮廓，而且对包含大量细节信息图像（如图 4.6(e)、图 4.7(e)）的边缘提取效果也优于其他方法。

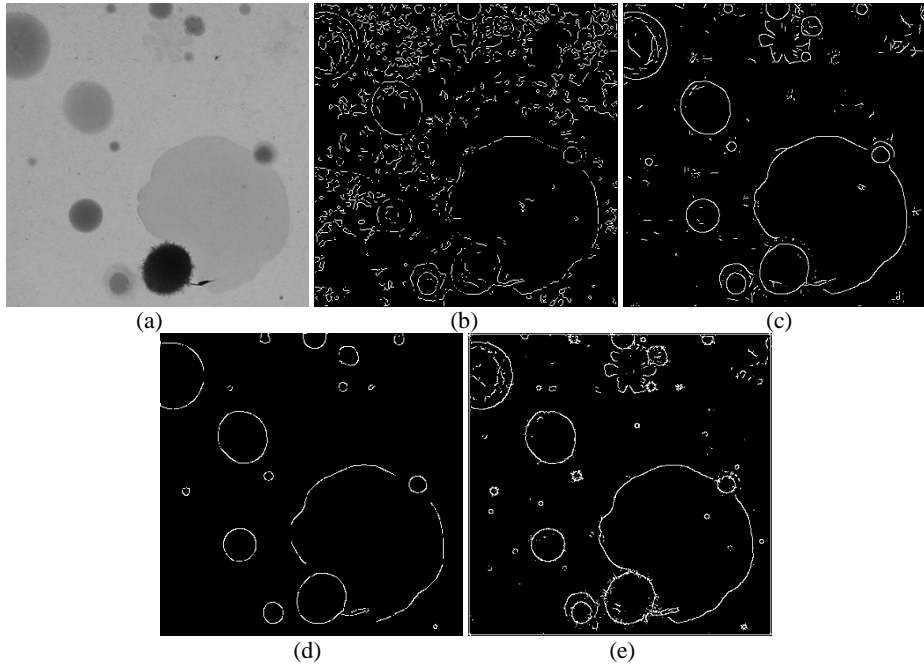


图 4.5 图 4.4(a)边缘检测结果。(a) 图 4.4(a)灰度图；(b) Canny 方法；(c) Gabor 滤波器方法；(d) MCI 方法；(e) 本文方法

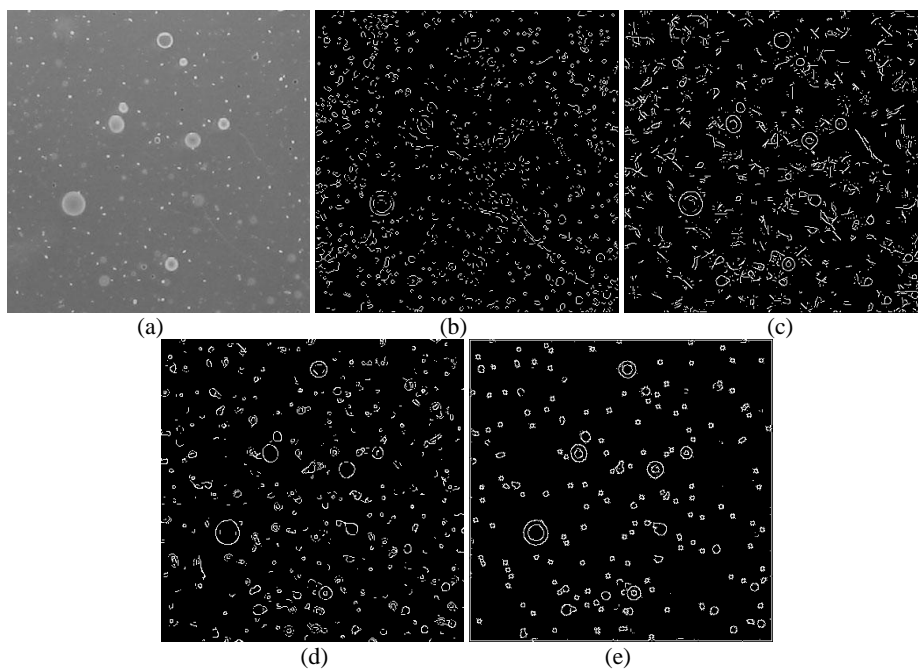


图 4.6 图 4.4(b)边缘检测结果。(a) 图 4.4(b)灰度图；(b) Canny 方法；
(c) Gabor 滤波器方法；(d) MCI 方法；(e) 本文方法

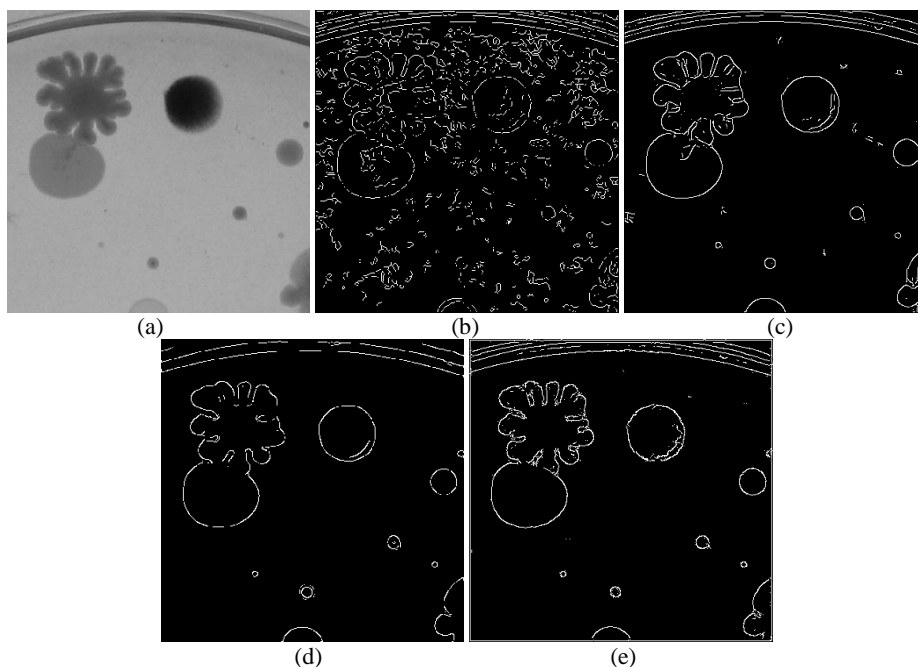


图 4.7 图 4.4(c)边缘检测结果。(a) 图 4.4(c)灰度图；(b) Canny 方法；
(c) Gabor 滤波器方法；(d) MCI 方法；(e) 本文方法

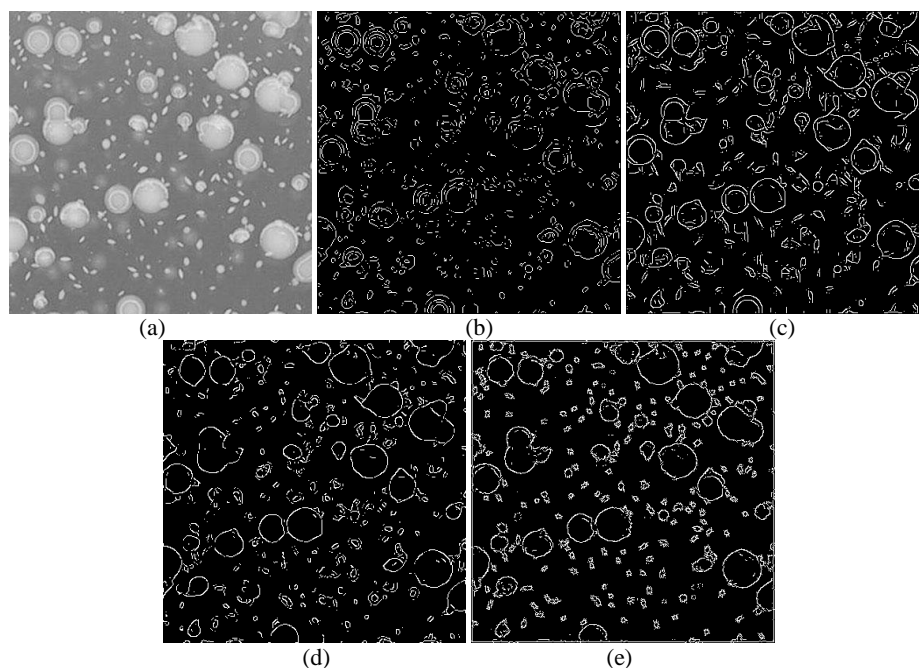


图 4.8 图 4.4(d)边缘检测结果。(a) 图 4.4(d)灰度图；(b) Canny 方法；(c) Gabor 滤波器方法；(d) MCI 方法；(e) 本文方法

为了对检测结果进行定量评价，利用边缘检测结果对原图进行重构，重构结果如图 4.9 所示，并计算该图与原图的 $MSSIM$ 指标，计算结果示于表 4.1 左侧，再计算 $BIdx$ 指标，计算结果示于表 4.1 右侧。从表 4.1 可以看出，本方法的重构相似度均显著大于其他方法，这说明其对于原图边缘的完整性有很好的保留，边缘信息丰富；本方法的边缘置信度也基本较前 3 种方法好，表明其检测出来的边缘质量好，真实度高。另外，Canny 方法的重构指标明显要优于 Gabor 滤波器方法和 MCI 方法，而其边缘置信度却不如后两种方法，所以有必要使用二者的综合指标 $EIdx$ 对结果进行评价。

表 4.1 各边缘检测方法的重构图像相似度和边缘置信度

图像	重构图像相似度 ($MSSIM$)				边缘置信度 ($BIdx$)			
	Canny	Gabor	MCI	本方法	Canny	Gabor	MCI	本方法
图 4.4(a)	0.781 4	0.559 9	0.487 6	0.930 0	0.076 6	0.169 1	0.261 8	0.313 8
图 4.4(b)	0.792 4	0.777 2	0.647 4	0.932 7	0.218 0	0.266 9	0.379 3	0.659 6
图 4.4(c)	0.704 4	0.542 0	0.524 2	0.911 2	0.175 8	0.413 5	0.502 2	0.506 6
图 4.4(d)	0.780 1	0.784 0	0.724 8	0.848 9	0.515 4	0.574 0	0.623 5	0.722 9

对 12 幅菌落图像（其中前 4 幅如图 4.4 所示）计算其在不同方法下边缘检测的 $EIdx$ 指标，并统计该指标均值及标准差，统计结果示于表 4.2。从表 4.2 可以看出，Canny、Gabor 以及 MCI 方法的综合指标均值分别为 0.520 8、0.522 5 和 0.524 4，而本方法所对应的数据为 0.746 8。这表明，本方法从整体上来看也是具有较好的边缘检测效果，不仅完整检测出边缘细节信息，而且真实度高。另外，对这 12 幅菌落图像的综合指标作 t 检验，结果显示，3 种对比方法与本方法数据的 P 值均小于 0.01，表明本方法与对比方法的综合指标具有显著性差异。

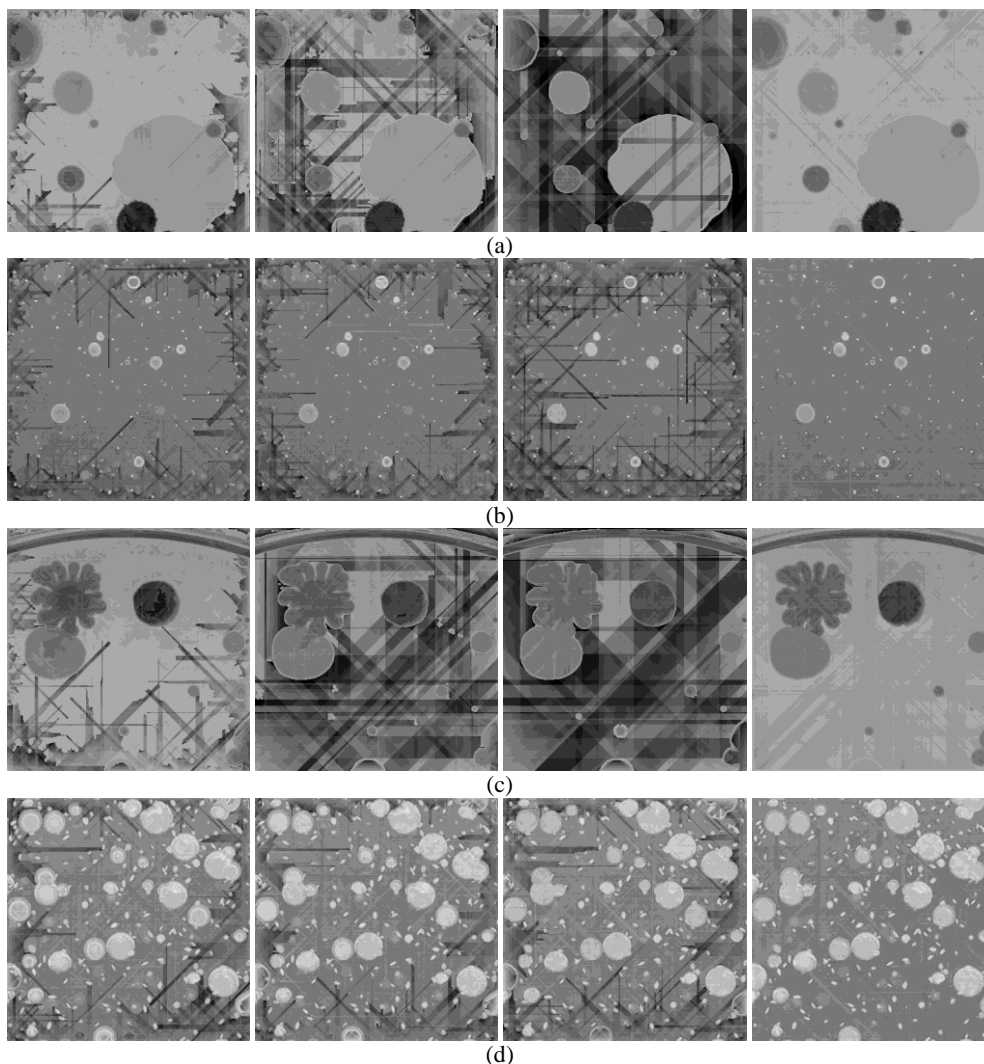


图 4.9 基于菌落边缘轮廓重构图像（每行从左至右分别对应基于 Canny、Gabor、MCI 方法以及本方法所获得的边缘重构得到的图像）。(a)~(d)分别对应于图 4.4 中的(a)~(d)

表 4.2 各边缘检测方法综合指标的均值及标准差（均值 \pm 标准差）

方法	Canny	Gabor	MCI	本方法
综合指标(EIdx)	$0.520\ 8 \pm 0.089\ 2$	$0.522\ 5 \pm 0.116\ 0$	$0.524\ 4 \pm 0.110\ 0$	$0.746\ 8 \pm 0.062\ 4$

4.5 讨论与分析

通过以上实验结果不难看出，本文方法具有较好的边缘提取能力，特别是针对弱对比度边缘，而且具有丰富细节信息的图像。从表 4.1 所展示的重构图像相似度和边缘置信度两项指标能反推本文算法的优势所在。

首先分析表 4.1 左侧重构图像相似度评价指标。重构相似度是基于亮度、对比度和结构三者综合，满足人类视觉感官的图像相似度指标。本文在构建视皮层下方位敏感功能层时，考虑了视皮层下视网膜神经节细胞以及外膝体神经元最优方位的向心分布规律，利用突触动态连接特性，形成多方位敏感的感受野模型，能更大程度的保留图像边缘细节信息；另外，对相同类型的边缘，例如水平边缘，本文计算响应方向为向上和向下时的边缘响应，分别对应下暗上亮和下亮上暗两种情形的图像边缘，更进一步地保证图像轮廓的完整性。从相似度

评价指标的结果来看,本文算法数据均在 0.8 以上,优于其他对比方法,同时,通过图 4.9 所示图像重构结果,也不难看出该评价结果与定性的观测结果是吻合的。

再分析表 4.1 右侧边缘置信度评价指标。边缘置信度是利用邻域相对标准差来说明所检测边缘的真实可靠性。本文在构建初级视皮层侧向抑制功能层时,利用了前层神经元脉冲输出的先后次序设定抑制规则。图像边缘通常会形成像素梯度变化的区域,而该变化中总会有一个最大的变化即为真实边缘所在。在最大响应方位上有选择的进行侧向抑制不仅有利于去除伪边缘和纹理噪声,而且能保护真实边缘不受破坏,进而能准确定位;对抑制区域去除最优方位能够避免边缘之间自抑制,保护弱小细节,还原更真实的图像轮廓。故本文方法能在拥有高相似度指标的情况下,还能保证边缘置信度指标也优于对比方法。与之相反的是,Canny 算法虽然也拥有较好的相似度指标,但边缘可靠性不高,引入大量纹理噪声,故置信度指标明显低于其他方法。分析其原因,这是由于重构图像的质量会随算法检测出的边缘数量增加而提高,Canny 方法检测出来的边缘数量较多,其中不仅包括真实边缘,也包含大量虚假边缘,从检测结果图 4.5~图 4.8 的子图(b)也很容易看出其误检较多。

综合来看,Canny 算法通过高低阈值选取确定最终的边缘,不能有效区分细节信息和纹理噪声,稳定性较差;MCI 方法虽然考虑了非经典感受野抑制作用,使得重构相似度和边缘置信度有个很好的平衡,表现出较好的稳定性,但大范围的非经典感受野抑制使得主要菌落边缘更清晰的同时丢失大量细节;Gabor 滤波器方法虽然也考虑了方向敏感特性,但没有神经元之间的侧向抑制作用,导致大量伪边缘,而且边缘连续性上表现较差,表明这三种方法都不适合于检测层次模糊而细节丰富的图像边缘。相反,本文方法不仅能准确在纹理噪声中区分出弱边缘和细节信息,保证边缘完整性,而且还不失真性,优势较大。

4.6 小结

本章通过讨论视觉通路上神经节细胞、外膝体以及初级视皮层在轮廓特征提取过程中的具体分工,探讨视皮层下细胞感受野分布与最优敏感方位之间的关系,同时引入初级视皮层在信息筛选过程的脉冲编码和侧向抑制机制,摒弃采用黑箱形式的数学模型对视觉机制进行模拟的思路,提出一种基于突触连接皮层下和皮层上分级处理的图像边缘检测方法,并与 Canny、Gabor 滤波器以及 MCI 等方法进行分析比较,表明本文提出的方法检测出的图像边缘细节更加丰富且不引入纹理噪声。菌落图像的边缘检测实验结果表明,本文通过视皮层下功能层凸显边缘强度信息,再通过初级视皮层功能层进行有效信息筛选,并对微弱信息加强,不但能降低纹理和噪声对真实轮廓的干扰,还能加强细节信息,从而获得更好的整体效果。

第 5 章 基于显著性信息稀疏表示的图像轮廓检测

轮廓与边缘不同，它主要指图像中能够反映主体目标的轮廓信息，因此轮廓能够为后续目标识别与图像理解等高级视觉行为提供更有用的信息。尽管视皮层下结构以及 V1 区对朝向、位置等信息有着强烈的反应，这对于物体轮廓的感知起到了一定的作用，但如何区分纹理边缘和主体轮廓，直至形成大脑形状知觉，势必需要更高级视皮层的参与。

虽然人脑是如何认识世界的还没有统一说法，但大量生理电实验证明，神经信息处理的过程中存在普遍的规则和特性，例如稀疏编码理论，即同一刺激作用于某一神经元群体，只有部分神经元同时被激活，它们构成了功能神经元集群，协同完成信息的编码，能够实现从大量复杂信号中提取出有效信息。再如前述视觉选择注意特性，能够迅速将注意力转移到视野中感兴趣的区域，弱化无用信息，强化有用信息，这对于轮廓感知的意义尤为明显。另外在视通路上，视觉信息的处理并不只存在单向的前馈通路，高级结构也会向下级结构反馈调控信号，形成错综复杂的神经网络，这有利于信号的进一步整合。因此，本章结合视信号处理过程中的这些规则及机制，构建具有神经电生理特性的神经网络，实现图像轮廓提取。以 RuG 图像库为例，计算最佳检测结果与基准轮廓图的 P 指标，计算结果显示本文方法能有效抑制纹理边缘的响应，而保留尽可能多的真实轮廓；另外，盒须图（Box-and-Whisker）统计的结果表明，相对其他对比方法，具有更好的稳定性。

5.1 概述

本章提出一种基于神经信息传递稀疏编码特性的神经网络对图像进行轮廓提取，用二维高斯导函数提取 V1 边界响应结果，将其输入到具有时空编码特性的稀疏神经网络，期间引入 NCRF 和 CRF 差异实现视觉注意机制，提取显著性图像，并以高级视皮层反馈控制的形式作用于 V1 边界响应图像，得到最终的轮廓响应。

5.2 轮廓检测模型

初级视皮层对朝向、位置等信息有着更为强烈的反应，这对于物体轮廓的感知起到了一定的作用，但如何区分纹理边缘和主体轮廓，直至形成大脑形状知觉，势必需要更高级视皮层的参与。本文通过构建高级视皮层稀疏神经网络，提取 V1 边界响应结果的显著性图像，并将显著性信息反馈到 V1，以此提取图像轮廓。具体检测框架如图 5.1 所示。

5.2.1 V1 神经元感受野边界响应

初级视皮层细胞对特定朝向的条形光刺激有高度选择性反应，这是轮廓及形状感知的生理基础。引入公式(2.1)所示二维高斯导函数用于模拟初级视皮层细胞感受野特性。对于某一特定的方位 $\theta_i \in [0, \pi)$ ，输入图像 $I(x, y)$ 在初级视皮层方位敏感性感受野作用下的响应为：

$$e_i(x, y; \theta_i) = |I(x, y) * \text{RF}(x, y; \theta)| \quad (5.1)$$

其中*为卷积运算， θ_i 如式(2.2)所示。

最后，采用 winner-take-all 策略，即选取各个方位下响应的最大值作为边界输出，另外考虑到方位响应的响应强度对最终结果的影响，引入响应强度系数对其进行调节，定义为：

$$\mu(x, y) = \frac{\max_i \{e_i(x, y; \theta_i)\}}{\sum_i e_i(x, y; \theta_i)}, i = 1, 2, \dots, N_\theta \quad (5.2)$$

从而初级视觉层最终的边界响应定义为式(5.3)所示：

$$E(x, y) = \mu(x, y) \times \max \{e_i(x, y; \theta_i) | i = 1, 2, \dots, N_\theta\} \quad (5.3)$$

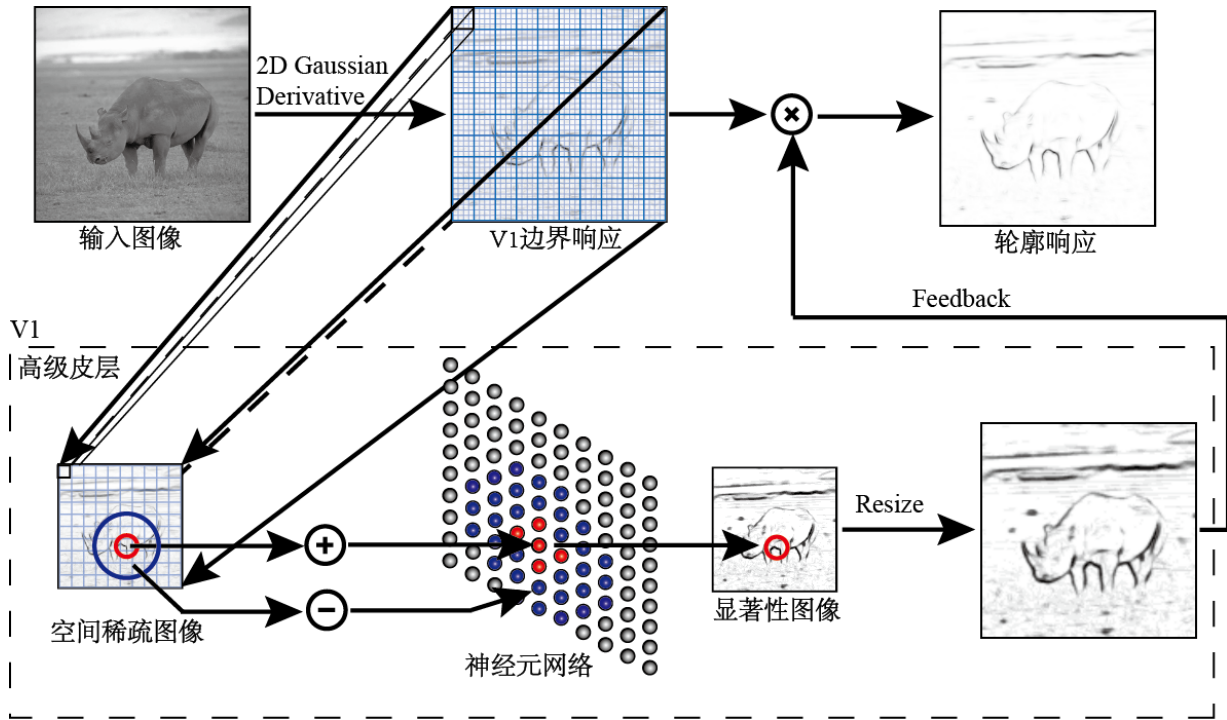


图 5.1 轮廓检测框架

5.2.2 稀疏编码神经网络

神经元通过脉冲发放的方式将外界刺激传递给大脑，信息在传送和表达过程中所采用的一系列规则和机制叫做神经信息编码^[64, 65]。稀疏编码理论指出，当同一刺激作用于某一神经元群体时，只有部分神经元同时被激活，这些被激活的神经元通过协同作用共同完成了信息的编码。这有利于完成信息在向高级神经中枢传递过程的逐步去冗余化，从大量复杂刺激中提取有效信息。神经元是神经信息编码的生理结构基础，神经元群体活动对视觉刺激的编码将影响生物体的行为任务、学习记忆等^[45]。本文选择既具备脉冲发放特性，并且计算上简单高效的 LIF 模型组建神经网络，同时为符合稀疏编码规则，引入绝对不应期，以有限数量的脉冲个数表示图像有用信息。具备绝对不应期特性的 LIF 模型如式(5.4)所示。式中， v 是神经元膜电压， c_m 是膜电容， g_l 是漏电导， I_{syn} 是突触输入电流， v_{th} 为脉冲发放阈值， v_{reset} 为静态电势， ref 为绝对不应期， v_G 为脉冲发放峰值。当 v 大于 v_{th} 时，神经元脉冲发放达到 v_G ，

同时 v 又被瞬间重置为 v_{reset} ，开始进入绝对不应期，在绝对不应期内不管刺激电流多大，神经元都不能被激活，直至不应期结束。

$$\begin{cases} c_m \frac{dv}{dt} = -g_l v + I_{\text{syn}} (\text{ref} = 0) \\ v = v_{\text{reset}}, \text{ref} = \text{ref} - 1 (\text{ref} \neq 0) \\ v = v_G, \text{ref} = \text{const} (v > v_{\text{th}}) \end{cases} \quad (5.4)$$

另外，图像信息的稀疏编码将整合空间信息，对于包含大量纹理边缘的初级视皮层方位选择结果，物体轮廓所在的区域被认为是稀疏的，而纹理边缘所在区域对于不同的方向和尺度都有着相当的响应，故被认为是非稀疏的。基于此，对于一个特定点处的响应，将以之为中心的局部窗的空间信息作为神经元的突触电流输入，同时以一定的步长移动局部窗，从而减小神经元网络规模，从空间上满足稀疏编码规则。选取局部窗尺寸与移动步长相等，故神经元网络的突触输入电流可描述为：

$$I_{\text{syn}}(x, y) = \text{mean} \left(\sum_{1 \leq w, h \leq \text{step}} E[(x-1) \times \text{step} + w, (y-1) \times \text{step} + h] \right) \quad (5.5)$$

式中， step 为局部窗移动步长。

5.2.3 视觉注意机制

本文首先通过自底向上的方法，提取初级皮层响应的局部空间特征，以突触抑制电流的形式输入到基于稀疏编码的神经元网络，以网络输出得到显著性图像，并且通过神经元脉冲发放量来表征感兴趣区域的显著性强度。然而本文并非以单向的前馈通路进行后续处理，而是借鉴自顶向下的方法，将显著性强度信息反馈到前级皮层，减少视觉冗余信息，从而将更多的主体目标信息传递到更高的神经中枢作进一步的计算和判断。

以突触输入电流 $I_{\text{syn}}(x, y)$ 作为 3.4 节模型输入，经视觉注意的自底向上方法，如式(3.8)

所示，计算得到的突触抑制电流用 $I_{\text{inh}}(x, y)$ 表示。从而最终的突触输入电流：

$$I_{\text{syn}}(x, y) = I_{\text{syn}}(x, y) - \alpha \times I_{\text{inh}}(x, y) \quad (5.6)$$

式中，参数 α 为可调节突触抑制系数，反映了突触抑制电流作用强度。

在视通路上，视觉信息的处理并不仅仅存在单向的前馈通路，高级结构也会向下级结构反馈调控信号，形成错综复杂的神经网络。因此，神经元之间除了与同级结构内的其他神经元形成互连的局部网络，还要接受前级结构的信号输入以及高层结构的反向投递信息^[66]。这在一定程度上解释了视觉注意机制自顶向下形成显著性图像存在的生理基础。

于是，本文将公式(5.6)获得的突触电流输入到式(5.4)描述的 LIF 神经元模型，以脉冲发放频率作为显著性图像输出，表示为 $\mu_r(x, y)$ 。再将其反馈回初级视皮层，以控制系数的形式作用于具有大量纹理边缘的方位选择结果，最终的轮廓结果如式(5.7)所示。

$$\text{SE}(x, y) = \mu_r(x, y) \times E(x, y) \quad (5.7)$$

值得注意的是，由于 V1 边缘响应图像整合了空间信息后尺寸减小，相应神经网络规模以及网络输出的显著性图像尺寸与原图不一致，这里的 $\mu_t(x, y)$ 为将显著性图像经双线性插值算法放大为原图大小后的图像。

5.2.4 高级视皮层反馈控制作用

高级视皮层反馈控制作用如图 5.2 所示。输入图像（图 5.2(a)）经多方位的二维高斯导函数滤波获得初级视皮层边界响应（图 5.2(b)）。该边界响应描述了 V1 感受野对落入其中的视觉刺激的响应，其对纹理边缘同样有着强烈的响应。为了抑制纹理、凸显轮廓，引入高级视皮层选择注意信息的反馈控制，以达到抑制纹理和背景区域的响应。可以看出，自底向上的局部空间特征形成的突触抑制电流（图 5.2(c)），在主体轮廓上的抑制小，而纹理和背景区域的抑制较大，从而使得注意力有选择地转移到轮廓所在区域，弱化对纹理边缘的关注。从神经网络输出的显著性图像（图 5.2(d)、图 5.2(e)）也能看出，视觉注意在轮廓区域的强度远大于纹理区域，而且，经神经元的稀疏作用，大量背景以及一些弱纹理区域的响应直接被去除，从而得到较好的轮廓响应结果（图 5.2(f)）。从该轮廓结果可以看出，将视觉注意信息反馈回初级视皮层，能有效抑制视皮层对纹理边缘的响应，凸显主要目标的主体轮廓和图像中有价值的背景轮廓（天地交汇处）。由此可见，通过高级视皮层反馈控制所获得的轮廓结果更符合人类视觉注意机制，更接近人类轮廓和形状感知的结果。

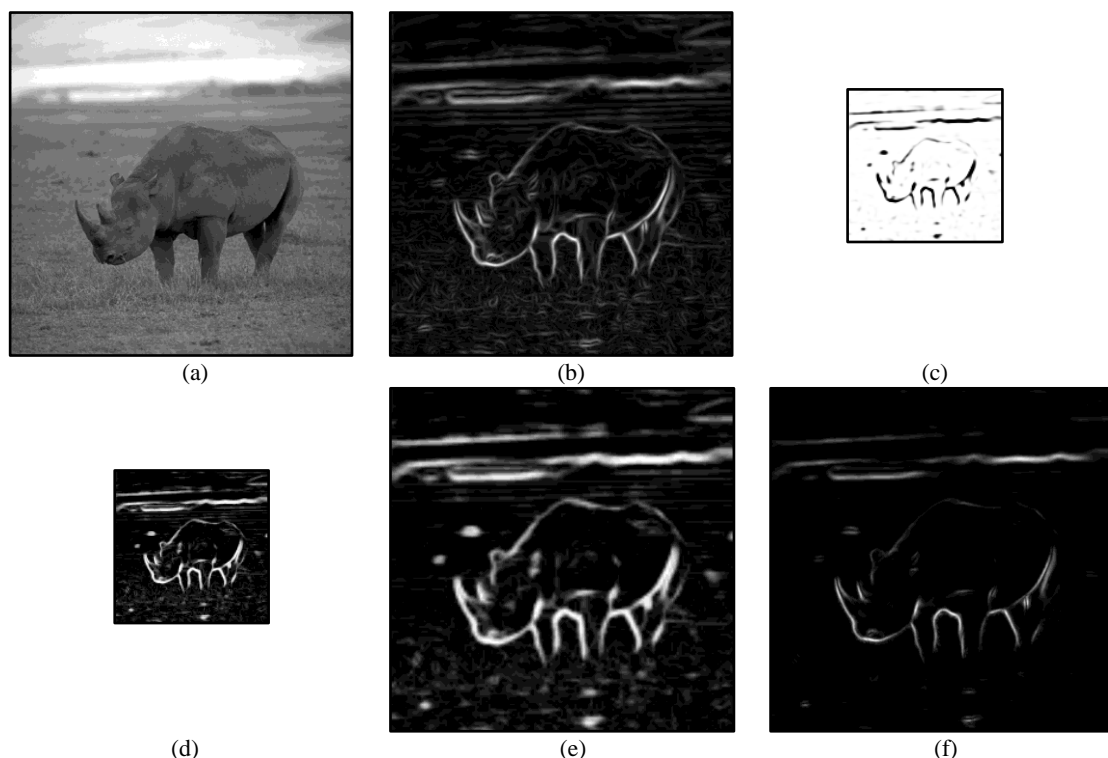


图 5.2 高级视皮层反馈控制作用。(a) 输入图像；(b) 初级视皮层边界响应；(c) 突触抑制电流；(d) 显著性图像；(e) 图(d)经双线性插值放大至原图大小；(f) 图(b)经反馈控制后的轮廓结果

5.3 实验结果

本文使用 RuG 轮廓检测数据库作为实验用图，该数据库共包含 40 幅自然图像，每幅图

像对应一幅手工描绘的基准轮廓图（多名被试者手工描绘轮廓图的平均结果），部分图像如图 5.3 第一行所示，所对应基准轮廓图如图 5.3 第二行所示。该图像库在评价轮廓检测算法性能时被广泛使用，可从网站 <http://www.cs.rug.nl/~imaging/> 获得。

为说明本文方法的有效性，首先选择了初级视皮层神经元感受野未经任何后续处理的结果 $E(x, y)$ (Noninh) 作为对比，体现本文后续算法的有用性；另外，选择文献[23]中的空间稀疏限制方法 (SSC) 和文献[13]中非经典感受野同步抑制方法 (ISO) 对 $E(x, y)$ 进行相应的纹理抑制处理，得到该方法下的实验结果，并与本文算法进行比较。实验中，对于初级视皮层神经元感受野，我们选取了 8 个不同的方位的二维高斯导函数，即 $N_\theta = 8$ ，并设置感受野尺寸 $\sigma = 2$ ；另外，由于突触抑制系数的大小反应了突触抑制电流强弱，我们选取了 10 个不同的强度，对应参数为 $\alpha = [0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5]$ 。为了与图像库提供的二值标准轮廓图进行比较，对本文方法以及对比方法提取的轮廓灰度图进行标准非极大值抑制提取单像素轮廓，然后用滞后门限方法进行二值化处理，得到最终的结果图。滞后门限方法中，选择保留的轮廓像素比例参数 $p = [0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0]$ ，对于 ISO 方法和本文方法可得到 100 组不同参数进行测试，而由于 Noninh 和 SSC 方法并不涉及全局系数 α ，故只进行 10 组参数测试。

图 5.3 为部分自然图像经各测试方法得出的结果。第一行为图像库原图，第二行为图像库中相应的基准轮廓图；第三行为 Noninh 方法检测结果；第四行为 SSC 方法检测结果；第五行为 ISO 方法检测结果；第六行为本文方法检测结果。从图中可以看出，进行了纹理边缘抑制的方法 (SSC、ISO 和本文方法) 相对于只进行方位性检测方法，能有效抑制背景及纹理边缘对主体轮廓的干扰，提高主体轮廓凸显能力。Noninh 方法是对全图进行多方位边缘线段检测，细化处理后仅选取一定百分比的强边缘作为最终效果，特点是主体轮廓比较完整，边缘漏检较少，在图像中主体轮廓像素较多时会获得比较不错的效果，如图 Buffalo，但背景区域变大、纹理更丰富时，区分能力明显下降，如其他子图所示，纹理及背景形成的误检较多。再看 SSC 方法的检测结果，虽然背景区域的误检较少了，但在主体轮廓附近的区域，对纹理边缘的抑制作用不够明显，形成大量毛刺，各子图均有体现。ISO 方法在三种对比方法中表现最好，大多数纹理都能有效去除，主体轮廓突出明显，但在图像各处仍存在大量散列的强纹理边缘。综合来看，本文方法获得的检测结果主体轮廓相对连续而完整，对强纹理边缘的抑制作用相比 ISO 方法有着一定的优势，整体检测效果最优。

基于 RuG 数据库中的基准轮廓图，利用 Grigorescu 等提出的轮廓检测定量评价指标，对轮廓检测算法效果进行定量评价，计算方法如下。首先，由于检测到的轮廓像素并不是总与基准轮廓图像像素精确对应，因此只要检测到的像素出现在基准轮廓像素的 5×5 邻域内，就认为该像素检测正确。定义 E_D 和 E_{GT} 分别为算法检测到的轮廓像素集和基准轮廓像素集，那么算法正确检测到的像素集为 $E = E_D \cap (E_{GT} \oplus T)$ (\oplus 为膨胀操作， T 为 5×5 的结构元)；错误检测到的像素集为 $E_{FP} = E_D - E$ ；漏检的像素集为 $E_{FN} = E_{GT} - (E_{GT} \cap (E_D \oplus T))$ 。这样，算法性

能评价指标误检率 e_{FP} 、漏检率 e_{FN} 和整体性能指标 P 可通过公式(5.8)~(5.10)计算得到。其中， $\text{card}(S)$ 为集合 S 中的元素数量。显然，错误检测和漏检到的像素同时越少，即误检率 e_{FP} 、漏检率 e_{FN} 越小，表明背景及纹理边缘少、主体轮廓明显，整体性能 P 值就越大。所以，以 P 值作为评价算法轮廓检测性能指标，该值越大表明算法检测效果更优，检测到的轮廓图更接近于基准轮廓图。

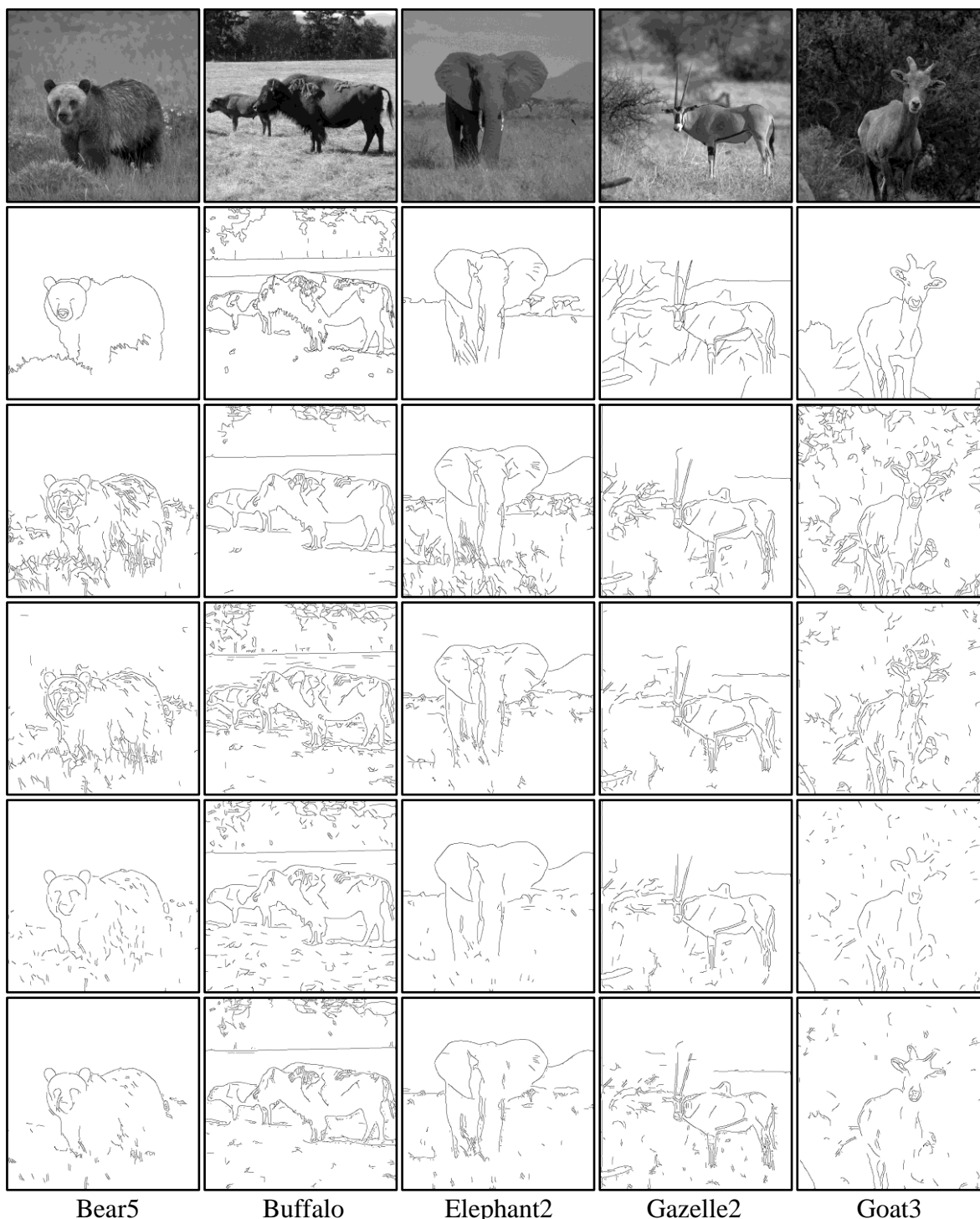


图 5.3 各方法在图像库上的测试结果

$$e_{FP} = \text{card}(E_{FP}) / \text{card}(E) \quad (5.8)$$

$$e_{FN} = \text{card}(E_{FN}) / \text{card}(E_{GT}) \quad (5.9)$$

$$P = \frac{\text{card}(E)}{\text{card}(E) + \text{card}(E_{FP}) + \text{card}(E_{FN})} \quad (5.10)$$

运用上述公式对图 5.3 中各算法获得的最优效果计算其性能指标, 以及获得最优效果时的参数设置, 计算结果如表 5.1 所示。读表可知, 定量的性能指标与前述对图 5.3 的定性描述大致吻合。Noninh 方法的漏检率 e_{FN} 较小, 表明获得的轮廓比较完整, 但误检率 e_{FP} 较大, 相应得到的 P 值不太理想。但也有例外, 如图像 Buffalo, 该方法获得的 P 值为 0.58, 优于其他对比方法, 其原因大概与前文所述一致, 是由于图像中主体轮廓像素较多所致。SSC 方法对比 Noninh 方法有了一定改善, 但整体仍表现误检较大、漏检较小, P 值虽有所改善, 但效果不佳。而 ISO 和本文方法在误检和漏检之间取得一定的平衡, 整体性能改善较大, 但由表中 P 值可以看出, 本文所提方法整体检测效果都优于其他对比方法。

表 5.1 图 5.3 所示图像对应的参数设置及性能指标

图像	算法	参数		性能指标		
		α	P	e_{FP}	e_{FN}	P
Bear5	Noninh		0.10	3.30	0.13	0.22
	SSC		0.10	2.56	0.19	0.26
	ISO	1.20	0.10	1.48	0.40	0.31
	本方法	0.70	0.10	1.14	0.43	0.34
Buffalo	Noninh		0.10	0.15	0.33	0.58
	SSC		0.20	0.62	0.21	0.52
	ISO	.080	0.20	0.39	0.26	0.56
	本方法	0.60	0.30	0.29	0.26	0.59
Elephant2	Noninh		0.10	1.46	0.11	0.38
	SSC		0.10	0.75	0.23	0.48
	ISO	1.10	0.10	0.23	0.36	0.54
	本方法	0.60	0.30	0.42	0.24	0.57
Gazelle2	Noninh		0.10	0.79	0.33	0.43
	SSC		0.10	0.66	0.39	0.43
	ISO	0.60	0.10	0.57	0.38	0.44
	本方法	0.70	0.30	0.67	0.35	0.45
Goat3	Noninh		0.10	2.77	0.15	0.25
	SSC		0.10	1.78	0.22	0.33
	ISO	1.30	0.10	0.78	0.40	0.40
	本方法	0.70	0.10	0.79	0.37	0.42

为证明本文方法的一般性, 运用统计的方法评价本文算法在更多图片、不同参数下的性能以及稳定性。图 5.4 所示为图库中图像检测结果的 P 值的盒须图 (Box-and-Whisker) 统计结果。盒须图的最顶端横线表示算法在 100 组参数中取得的最优结果, 即最大 P 值, 而中间的水平红线表示不同参数下 P 值的中值, 另外, 盒须图越短表明算法对不同参数越不敏感, 稳定性更高。从统计结果可以看出, 本文算法相对其他对比方法在整体上是呈现更优的趋势。

大部分图像本文算法都能获得最优性能（最大 P 值），虽然也存在部分图像本文方法最佳表现不如 ISO 方法，如图 Hyena，但本文算法盒须图更短，鲁棒性更好，在稳定性的表现上反而优于 ISO 方法。另外由于 Noninh 方法和 SSC 除了 p 参数外并无其他可调节参数，只测试了 10 组参数，故在盒须图的表现上会呈现比其他方法更优的表象。

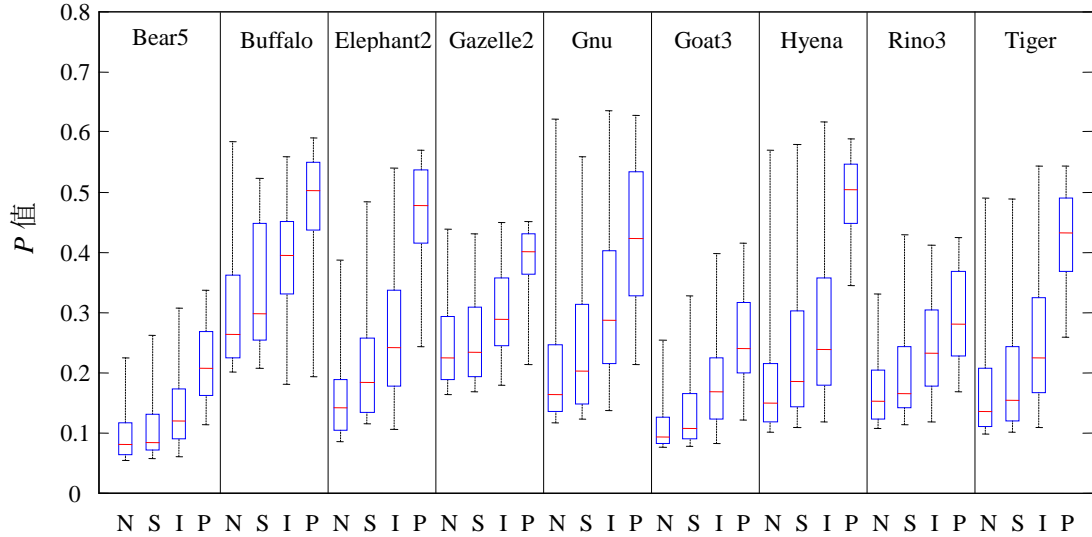


图 5.4 各方法在多组参数下的检测结果的 P 值盒须图 (Box-and-Whisker) 统计。其中 N 表示 Noninh 方法，S 表示 SSC 方法，I 表示 ISO 方法，P 表示本文方法

5.4 讨论与分析

初级视皮层细胞对视觉刺激中的方位、位置等信息具有高度选择作用，这对于物体轮廓的感知起到了一定的作用，本文使用二维高斯导函数对这样的方位敏感作用进行了模拟，结果显示细胞对主要的边缘，包括主体轮廓和纹理边缘，均有强烈的反应。于是，抑制背景和纹理边缘、凸显主体轮廓是本文工作的出发点。我们考虑到，大脑形状知觉的形成，除了初级视皮层的方位感知作用，势必需要更高级视皮层的参与，提出一种将神经信息的稀疏编码特性与脉冲神经网络相结合的轮廓检测方法，并引入视觉注意机制，通过神经元脉冲发放活动表征图像显著性信息。

本文所做工作及其创新性如下。(1) 提出了一种基于视觉信息处理机制的图像轮廓感知新方法。针对初级视皮层对朝向选择的处理结果，考虑高级视皮层在信息处理过程中的稀疏编码规则，对前级结果进行稀疏表示，过程中考虑视觉选择注意机制，进一步筛选有效信息，然后将高级皮层处理的结果进行反向投射，与 V1 的朝向选择结果融合作为最终的轮廓感知结果。突出了生物视觉系统信息处理能力的应用研究。(2) 实现了一种基于稀疏编码网络模型的轮廓检测方法在自然图像中的具体应用。针对自然图像空间信息的重要性，摒弃神经网络规模与输入图像相当的做法，以一定步长将图像分块，并以块信息作为神经元输入，体现神经元集群编码特性，另外考虑脉冲神经元不应期特性，以较少的脉冲数量表示图像有用信息；过程中进一步考虑神经元经典感受野与非经典感受野输入信息的差异性，实现视觉选择注意机制，将最终提取的选择注意信息作用于 V1 朝向选择的结果，能有效抑制纹理边缘，

凸显主体轮廓。(3) 为如何利用脉冲神经元特性实现图像的稀疏表示以及选择注意等视觉机制进行图像处理，或者是后续的图像分析和理解提供了新思路。脉冲神经元的不应期特性使得神经元网络在视觉信息处理的极短时间内用有限的脉冲数量表示视觉输入，符合神经信息处理过程中的稀疏编码规则；而针对自然图像中纹理边缘会对主体目标造成极大干扰的情形，视觉系统要从中提取有效轮廓，势必要考虑更大范围的空间信息，这里将大范围的非经典感受野与经典感受野内视觉输入的差异以突触权值的形式来实现选择注意，能有效提取目标轮廓所在的区域，而舍弃纹理性边缘。本文以视通路反向信息投递为生理基础，构建了具有时空特性的稀疏脉冲神经网络，整合 V1 和高级视皮层的处理的结果，为图像轮廓提取及后续形状感知与行为理解提供了新的思路。

相对于其他对比方法，SSC 方法虽与本文所提算法有着相似的神经信息理论基础，都考虑了视信号处理过程所遵循的稀疏编码机制，但 SSC 方法通过数学算子描述空间稀疏特性，用于区分纹理区域和非纹理区域，只体现了空间稀疏特性，并且丧失了稀疏编码的神经元脉冲发放的生理基础，故而在检测结果中表现为主体轮廓附近出现大量误检；而本文方法不仅考虑了空间稀疏特性，还通过提取的空间差异性信息以突触抑制电流的形式对其纹理区域进行抑制，再通过考虑了不应期的神经元处理，从时间维度体现稀疏特性，进一步将轮廓和纹理进行区分。另外，ISO 方法对方位检测结果进行不加区分的大范围非经典感受野抑制作用，尽管部分图像取得较好效果，但当图像中包含主体包括丰富细节时表现不佳，这是由于当背景中出现较强的纹理边缘时，经过抑制作用与主体轮廓仍然无法区分，而增大抑制系数时，主体轮廓又会出现断裂，漏检率升高，造成整体算法稳定性不强；本文方法虽然也考虑了非经典感受野对中心神经元的调控作用，但是以其与经典感受野的对比性差异形成突触抑制电流，而计算所得对比性差异对主体轮廓与纹理背景具有较好的区分性，同时本文方法并不以抑制后的神经元输出作为最终的结果，而是以反向投递的方式作用于前级视皮层处理的结果，使最终结果尽量保存更多细节，这使得后续通过调节参数和阈值来进一步改善效果、得到最优结果成为可能。

5.5 小结

本章通过构建具有神经电生理特性的神经元网络，以较为新颖的方式从时间、空间维度实现稀疏编码特性，更多的考虑了该特性存在的生理基础，而非从矩阵分解等数学分析方式对图像进行降维表示；另外通过引入经典感受野与非经典感受野的对比性差异实现视觉注意机制，并以视通路高级皮层反馈的形式形成最终轮廓检测。针对 RuG 图像库中的自然图像，通过与不加抑制的 Noninh 方法、空间稀疏限制的纹理抑制方法 SSC，以及非经典感受野同步抑制方法 ISO 所提取结果进行比较，表明本文算法不仅对纹理和背景具有抑制的作用，而且能有效区分纹理强边缘和主体轮廓，获得较佳的效果，尽管本文算法在部分图像上的表现并非最佳，但如同前述统计结果显示的一样，稳定性更高。

第6章 总结与展望

6.1 总结

目前,国内外有关视觉信息处理机制与图像处理应用相结合的发展趋势,一方面通过建立仿生的数学模型实现视觉机制,然后将其运用于图像处理,忽视了视觉信息传递过程中的电生理基础,即会丢失脉冲发放的时间编码信息;另一方面组建大规模神经元网络,却局限于某种单一的视觉神经特性,注重实现神经网络自身的组织结构,忽略了与视觉机制的结合。本文研究基于视觉感知的神经计算模型,注重视觉通路上不同结构对视觉信息的分工处理,简化和优化视觉机制模型,通过脉冲编码的方式实现图像边缘和轮廓检测,检测结果具有一定优势,更重要地,提供了一种基于视觉感知机制的图像处理和理解新思路。

对于视觉感知机制在应用上的实现方法,大多是通过黑箱形式的数学模型对视觉刺激进行滤波处理,但如何利用突触连接特性构建感受野模型,再利用神经元模型对视觉信息进行脉冲编码,还没有深入研究。本文基于神经生理发现的视觉机制及其在图像特征感知中的作用,研究用神经元网络的方式实现相应的功能,并以简单的感受野明暗刺激验证了所提网络模型的有效性,为后续应用奠定了基础。

尽管基于人类视觉感知机制的图像边缘检测已取得了不少成果,并且相对传统边缘检测方法具有视觉机理上的优势,但它们忽略了视觉通路上不同结构对视觉信息的分工处理。本文考虑了视通路上不同的组织结构对边缘信息有着不同层次的感知能力,构建视皮层下-初级视皮层的分级处理模型,结合动态突触连接、侧向抑制等机制对低对比度图像边缘进行了检测,取得不错的效果。

另外,神经元不应期特性使得视觉系统在视觉信息处理的极短时间内,用有限的脉冲数量表示视觉刺激,符合神经信息处理过程中的稀疏编码规则;而且在视觉通路上,视信息的处理并不只存在单向的前馈通路,高级结构也会向下级结构反馈调控信号,形成错综复杂的神经网络,这有利于信号的进一步整合。本文利用神经元不应期特性,构建具有空间信息融合特性的 LIF 神经元网络,并引入 CRF 和 NCRF 差异性特征实现显著性图像的稀疏表示,最后将其反向投递到初级视皮层,调控其对边界的响应,提出一种基于显著性信息稀疏表示的图像轮廓检测新方法。

论文具体研究工作总结如下:

(1) 本文以突触连接的形式构建新的神经元感受野模型,实现方位敏感、侧向抑制以及选择注意等视觉机制,充分发挥其在感知明暗光对比刺激以及空间变化中的作用。首先将融合了感受野区域信息的图像特征输入神经元模型,通过神经元模型将视觉刺激转化为脉冲序列;再从脉冲序列的时间编码信息中解析视觉刺激所包含的特征。以明暗刺激施加于该感受

野模型，通过神经元的响应证实了本文提出的模型能实现相应视觉机制的功能，为后续进一步的应用打下了基础。

(2) 考虑视觉通路上不同生理结构对视觉信息的分工处理，针对具有多强度边缘或者丰富细节的图像，优化侧向抑制方式，从而提出一种基于突触连接视通路方位敏感的图像分级边缘检测新方法。首先考虑视皮层下视网膜神经节细胞和外膝体的方位敏感特性，探讨其感受野最优方位在视网膜上的分布规律，利用神经元突触间的动态连接特性构建视皮层下多方位敏感层；再结合初级视皮层更强烈的特殊方位条形光刺激感知能力，对上一功能层输出的脉冲序列编码，并对首发脉冲时间信息加以侧向抑制和阈值处理，得到图像边缘。实验结果表明，该方法检测出的图像边缘细节更加丰富且不引入纹理噪声。

(3) 通过构建一种新的稀疏编码网络模型，并结合视觉注意机制在改善轮廓感知中的作用，提出一种基于显著性信息稀疏表示的图像轮廓提取方法，不仅能有效提取图像轮廓，且能抑制纹理边缘。首先用二维高斯导函数实现 V1 的方位敏感特性，提取边界响应；针对该边界响应，构建具有空间信息融合特性的稀疏神经元网络，在神经元模型中加入不应期使得在时间维度上满足稀疏编码规则；同时引入 NCRF 和 CRF 差异性特征，实现视觉注意机制，将网络输出的显著性图像以高级视皮层反馈控制的形式作用于 V1 边界响应，得到最终的图像轮廓。实验结果表明，该方法能有效抑制纹理边缘的响应，而保留尽可能多的真实轮廓，具有较好的稳定性。

6.2 展望

本文将基于视觉感知机制的神经计算模型应用于图像处理，证实了该方法的可行性和有效性。对于神经计算网络模型，可以进一步引入突触连接及学习机制，比如 STDP 机制，增强网络模型对于视觉特征的辨识能力，并且提高网络输出的多变性；对于脉冲编码方式，可以更多的考虑动作电位发放时间、空间特征，多维度考虑脉冲序列所携带的图像信息。

另外，视觉系统对视觉刺激的感知能力，还有更多方面的体现，例如颜色拮抗机制，对于彩色图像处理具有很好的借鉴意义；还如视觉适应性，这有利于视觉系统过滤冗余的信息，提高神经网络信息处理效率和反应能力，等等。后续可将更多的视觉机制与图像处理应用相结合，从而实现模型的推广。

最后，本文所构建稀疏神经元网络模型，对于空间差异性具有一定的识别能力，轮廓提取只是对该网络应用的初步研究，后续可以将其扩展到人脸识别、身份认证等领域。

致谢

两年半硕士生涯匆匆而过，回首过去，几多欢喜几多忧愁也是如鱼饮水，冷暖自知。虽不是硕果累累，但也是一种经历，一种成长。

值此学业即将完成之际，回顾这些年里接触的人和事，最真挚的感谢应该给予我的导师范影乐教授。范老师是一个真正从学生的立场出发，再给予帮助和指导的好老师！还记得刚开学时范老师就跟我们说，两年半时间说长其实很快就过去了，重点是过程一定要充实而愉悦。现在看来，范老师对我学业上的指导和生活上的关心确实是以这样一条准则而进行着。科研上，范老师治学严谨但从刻板，总是以启发的方式来引导我打开思路，完成论文的撰写和发表工作。范老师工作繁忙，但从不懈怠对我们的关心，总是利用自己的休息时间帮我们修改论文；暑期，鼓励我们参加各种科研竞赛，从初赛时的创意思路、文档撰写、修改、排版，到决赛的答辩，甚至我们在外地的住宿问题都会帮我们考虑，从这点看范老师于我是亦师亦友。

感谢武薇老师，虽然跟她接触的时间不多，但依然记得研一上数字图像处理课时，她告诉我们文献阅读一定要注意方法，并发给我们《怎样读文献》的 word 文档，受用至今！另外，感谢在档案室（我们称之为“小黑屋”）已毕业的李丹菁师兄、陆晓娟师姐、王梦蕾师姐、小明哥、健哥，在我初到实验室时给予的学习和生活上的帮助，档案室的生活是愉悦而温馨的！感谢师兄陈金龙、廖进文、罗佳骏给予我科研上无私的帮助；感谢同窗好友王迪、方芳，大家一起讨论交流、互帮互助；感谢师弟谢尚周、李康群、郭笑天、李嘉祥、王强、郭金鹏、胡钧皓，师妹张梦楠带来实验室的多多欢乐。

最后，我要感谢我的家人，谢谢父母这么多年无私的付出、支持与鼓励，你们的关怀与理解是我学习上最大的动力！

下沙.杭州电子科技大学
二零一五年十二月二十日

参考文献

- [1] 顾凡及, 梁培基. 神经信息处理[M]. 北京: 北京工业大学出版社, 2007.
- [2] Marr D, Hildreth E. Theory of edge detection[J]. Proceedings of the Royal Society of London B: Biological Sciences, 1980, 207(1167): 187-217.
- [3] Maini R, Aggarwal H. Study and comparison of various image edge detection techniques[J]. International journal of image processing (IJIP), 2009, 3(1): 1-11.
- [4] Canny J. A computational approach to edge detection[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 1986, PAMI-8(6): 679-698.
- [5] Papari G, Petkov N. Edge and line oriented contour detection: State of the art[J]. Image and Vision Computing, 2011, 29(2): 79-103.
- [6] Xu Q, Varadarajan S, Chakrabarti C, et al. A distributed canny edge detector: Algorithm and fpga implementation[J]. Image Processing, IEEE Transactions on, 2014, 23(7): 2944-2960.
- [7] 张志龙, 杨卫平, 李吉成. 一种基于蚁群优化的显著边缘检测算法[J]. 电子与信息学报, 2014, 36(9): 2061-2067.
- [8] Yang X, Gao X, Tao D, et al. An efficient MRF embedded level set method for image segmentation[J]. Image Processing, IEEE Transactions on, 2015, 24(1): 9-21.
- [9] 寿天德. 视觉信息处理的脑机制[M]. 第2版. 合肥: 中国科学技术大学出版社, 2010.
- [10] Daugman J G. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters[J]. JOSA A, 1985, 2(7): 1160-1169.
- [11] Shou T, Wang W, Yu H. Orientation biased extended surround of the receptive field of cat retinal ganglion cells[J]. Neuroscience, 2000, 98(2): 207-212.
- [12] Sun C, Chen X, Huang L, et al. Orientation bias of the extraclassical receptive field of the relay cells in the cat's dorsal lateral geniculate nucleus[J]. Neuroscience, 2004, 125(2): 495-505.
- [13] Grigorescu C, Petkov N, Westenberg M. Contour detection based on nonclassical receptive field inhibition[J]. Image Processing, IEEE Transactions on, 2003, 12(7): 729-739.
- [14] Grigorescu C, Petkov N, Westenberg M A. Contour and boundary detection improved by surround suppression of texture edges[J]. Image and Vision Computing, 2004, 22(8): 609-622.

- [15] Chen K, Song X, Li C. Contrast-dependent variations in the excitatory classical receptive field and suppressive nonclassical receptive field of cat primary visual cortex[J]. Cerebral Cortex, 2012, 23(2): 283-292.
- [16] Chen K, Song X, Dai Z, et al. The spatial summation characteristics of three categories of V1 neurons differing in non-classical receptive field modulation properties[J]. Vision research, 2014, 96: 87-95.
- [17] Xu T, Wang L, Song X, et al. The detection of orientation continuity and discontinuity by cat V1 neurons[J]. Plos One, 2013, 8(11): e79723.
- [18] Zeng C, Li Y, Li C. Center – surround interaction with adaptive inhibition: A computational model for contour detection[J]. NeuroImage, 2011, 55(1): 49-66.
- [19] Zeng C, Li Y, Yang K, et al. Contour detection based on a non-classical receptive field model with butterfly-shaped inhibition subregions[J]. Neurocomputing, 2011, 74(10): 1527-1534.
- [20] Yang K, Li C, Li Y. Potential roles of the interaction between model V1 neurons with orientation-selective and non-selective surround inhibition in contour detection[J]. Frontiers in neural circuits, 2015, 9(30): 1-16.
- [21] 杨开富. 基于多视觉特征的非经典感受野模型及应用研究[D]. 成都: 电子科技大学, 2012.
- [22] Yang K, Li C, Li Y. Multifeature-based surround inhibition improves contour detection in natural images[J]. Image Processing, IEEE Transactions on, 2014, 23(12): 5020-5032.
- [23] Yang K, Gao S, Guo C, et al. Boundary Detection Using Double-Opponency and Spatial Sparseness Constraint[J]. Image Processing, IEEE Transactions on, 2015, 24(8): 2565-2578.
- [24] Isaacson J S, Scanziani M. How inhibition shapes cortical activity[J]. Neuron, 2011, 72(2): 231-243.
- [25] Zi F, Zhao D, Zhang K. Image pre-processing algorithm based on lateral inhibition[C]. // Electronic Measurement and Instruments, 2007. ICEMI'07. 8th International Conference on. Xi'an: IEEE, 2007, 2: 701-705.
- [26] 史漫丽, 彭真明, 张启衡, 等. 基于自适应侧抑制网络的红外弱小目标检测[J]. 强激光与粒子束, 2011, 23(4): 906-910.
- [27] 廖进文. 基于视觉编码机制的图像弱边缘检测技术及其应用研究 [D]. 杭州: 杭州电子

科技大学, 2015.

- [28] Eckhorn R, Reitboeck H J, Arndt M, et al. Feature linking via synchronization among distributed assemblies: Simulations of results from cat visual cortex[J]. *Neural Computation*, 1990, 2(3): 293-307.
- [29] 毕英伟, 邱天爽. 一种基于简化 PCNN 的自适应图像分割方法[J]. *电子学报*, 2005, 33(4): 647-650.
- [30] Gao C, Zhou D, Guo Y. Automatic iterative algorithm for image segmentation using a modified pulse-coupled neural network[J]. *Neurocomputing*, 2013, 119: 332-338.
- [31] Kerr D, Coleman S, McGinnity M, et al. Biologically inspired edge detection[C]. // *Intelligent Systems Design and Applications (ISDA)*, 2011 11th International Conference on. Cordoba: IEEE, 2011, 802-807.
- [32] Wu Q, McGinnity T M, Maguire L, et al. A visual attention model based on hierarchical spiking neural networks[J]. *Neurocomputing*, 2013, 116: 3-12.
- [33] Eguchi A, Neymotin S A, Stringer S M. Color opponent receptive fields self-organize in a biophysical model of visual cortex via spike-timing dependent plasticity[J]. *Frontiers in neural circuits*, 2014, 8(16): 1-14.
- [34] Yu Q, Tang H, Tan K C, et al. Rapid feedforward computation by temporal encoding and learning with spiking neurons[J]. *Neural Networks and Learning Systems, IEEE Transactions on*, 2013, 24(10): 1539-1552.
- [35] Masquelier T. Relative spike time coding and STDP-based orientation selectivity in the early visual system in natural continuous and saccadic vision: a computational model[J]. *Journal of computational neuroscience*, 2012, 32(3): 425-441.
- [36] Wöhrer A, Kornprobst P. Virtual retina: a biological retina model and simulator, with contrast gain control[J]. *Journal of computational neuroscience*, 2009, 26(2): 219-249.
- [37] 黄丽鸿, 谌先敢, 刘海华. 模拟初级视皮层脉冲神经元的动作识别系统[J]. *自动化学报*, 2012, 38(12): 1975-1984.
- [38] Rodieck R W, Stone J. Analysis of receptive fields of cat retinal ganglion cells[J]. *Journal of Neurophysiology*, 1965, 28(5): 833-849.
- [39] Lindeberg T. A computational theory of visual receptive fields[J]. *Biological cybernetics*, 2013, 107(6): 589-635.

- [40] Chan W, Coghil G. Text analysis using local energy[J]. Pattern Recognition, 2001, 34(12): 2523-2532.
- [41] Young R A, Lesperance R M. The gaussian derivative model for spatial-temporal vision: II. Cortical data[J]. Spatial vision, 2001, 14(3): 321-389.
- [42] Wang H, Sun Y, Li Y, et al. Influence of autapse on mode-locking structure of a Hodgkin - Huxley neuron under sinusoidal stimulus[J]. Journal of theoretical biology, 2014, 358: 25-30.
- [43] Chouhan A S. An Analytical Study Of Leaky Integrate - And-Fire Neuron Model Using MATLAB Simulation[J]. Int J Eng Res Technol, 2013, 2(4): 2242-2245.
- [44] Tanabe S. Population codes in the visual cortex[J]. Neuroscience research, 2013, 76(3): 101-105.
- [45] 白文文. 神经元群体/集群电活动对行为任务的稀疏编码机制研究[D]. 天津: 天津医科大学, 2012.
- [46] Zylberberg J, Murphy J T, Deweese M R. A sparse coding model with synaptically local plasticity and spiking neurons can account for the diverse shapes of V1 simple cell receptive fields[J]. PLoS Comput Biol, 2011, 7(10): e1002250.
- [47] 张盛平. 基于稀疏编码的视觉目标跟踪方法研究[D]. 哈尔滨: 哈尔滨工业大学, 2012.
- [48] Okun M, Yger P, Marguet S L, et al. Population rate dynamics and multineuron firing patterns in sensory cortex[J]. The Journal of Neuroscience, 2012, 32(48): 17108-17119.
- [49] Vanrullen R, Thorpe S J. Surfing a spike wave down the ventral stream[J]. Vision research, 2002, 42(23): 2593-2615.
- [50] Gollisch T, Meister M. Rapid neural coding in the retina with relative spike latencies[J]. science, 2008, 319(5866): 1108-1111.
- [51] Anderson J R. Cognitive psychology and its implications[M]. Macmillan, 2005.
- [52] 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.
- [53] Zhang Z, Wu Q, Zhuo Z, et al. Wavelet transform and texture recognition based on spiking neural network for visual images[J]. Neurocomputing, 2015, 151: 985-995.
- [54] 王典, 范影乐, 张梦楠, 等. 基于突触连接视通路方位敏感的图像分级边缘检测方法[J]. 中国生物医学工程学报, 2015, 34(5): 522-532.
- [55] Rivlin-Etzion M, Wei W, Feller M B. Visual stimulation reverses the directional preference of

- direction-selective retinal ganglion cells[J]. *Neuron*, 2012, 76(3): 518-525.
- [56] Naito T, Okamoto M, Sadakane O, et al. Effects of stimulus spatial frequency, size, and luminance contrast on orientation tuning of neurons in the dorsal lateral geniculate nucleus of cat[J]. *Neuroscience research*, 2013, 77(3): 143-154.
- [57] Venkataramani S, Taylor W R. Orientation selectivity in rabbit retinal ganglion cells is mediated by presynaptic inhibition[J]. *The Journal of Neuroscience*, 2010, 30(46): 15664-15676.
- [58] Zhao X, Chen H, Liu X, et al. Orientation-selective responses in the mouse lateral geniculate nucleus[J]. *The Journal of Neuroscience*, 2013, 33(31): 12751-12763.
- [59] Suematsu N, Naito T, Miyoshi T, et al. Spatiotemporal receptive field structures in retinogeniculate connections of cat[J]. *Frontiers in systems neuroscience*, 2013, 7(103): 1-13.
- [60] Govindarajan B, Panetta K, Agaian S. Image reconstruction for quality assessment of edge detectors[C]. // *Systems, Man and Cybernetics*, 2008. SMC 2008. IEEE International Conference on. Singapore: IEEE, 2008, 691-696.
- [61] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: from error visibility to structural similarity[J]. *Image Processing, IEEE Transactions on*, 2004, 13(4): 600-612.
- [62] Meer P, Georgescu B. Edge detection with embedded confidence[J]. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2001, 23(12): 1351-1365.
- [63] 磨少清. 边缘检测及其评价方法的研究[D]. 天津: 天津大学, 2011.
- [64] Zhu M, Rozell C J. Visual nonclassical receptive field effects emerge from sparse coding in a dynamical system[J]. *PLoS Computational Biology*, 2013, 9(8): 1-15.
- [65] Spratling M W. Image segmentation using a sparse coding model of cortical area V1[J]. *Image Processing, IEEE Transactions on*, 2013, 22(4): 1631-1643.
- [66] Liang Z, Shen W, Shou T. Enhancement of oblique effect in the cat's primary visual cortex via orientation preference shifting induced by excitatory feedback from higher-order cortical area 21a[J]. *Neuroscience*, 2007, 145(1): 377-383.

附录:

作者在读期间发表的学术论文及参加的科研项目

发表(录用)学术论文:

1. 基于突触连接视通路方位敏感的图像分级边缘检测方法. 中国生物医学工程学报, 2015, 34(5):522-532. (一级期刊), 1/4.
2. 基于显著性信息稀疏表示的图像轮廓检测. 电子与信息学报, 已投稿, 审稿中.

授权(申请)专利:

发明专利

1. 一种基于视通路方位敏感的图像边缘分级检测方法, 已公开, 201510172244.2, 2/4.
2. 一种仿视觉感光层功能的图像多强度边缘检测方法, 已公开, 201410128822.8, 4/4.

参与的科研项目:

1. 基于双稳态随机共振非周期阈上响应的强噪声图像增强技术研究(60872090), 国家自然科学基金。
2. 基于眼前房角 OCT 影像的青光眼计算机辅助诊断研究 (Y201328513), 浙江省教育厅科研项目。
3. 自动化检测中的机器视觉开发模块研发(063113017), 横向课题。