

杭州电子科技大学

硕士学位论文

题 目： 基于非经典感受野机制的
视觉处理模型及其应用研究

研 究 生 郭 笑 天

专 业 控制科学与工程

指导教师 范 影 乐 教 授

完成日期 2016 年 12 月

杭州电子科技大学硕士学位论文

**基于非经典感受野机制的视觉处理模型
及其应用研究**

研 究 生：郭 笑 天

指导教师：范 影 乐 教 授

2016 年 12 月

Dissertation Submitted to Hangzhou Dianzi University

for the Degree of Master

**Visual Processing Model
and Its Research on Application
Based on Non Classical Receptive Field**

Candidate: Guo Xiaotian

Supervisor: Prof. Fan Yingle

December, 2016

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单位名称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

边缘或轮廓作为图像的关键特征信息,对于分析或理解视野中的目标对象具有重要意义。目前边缘或轮廓检测方法大多依据像素亮度等属性在数学描述上的跳变特性,因此从图像空间上来看,这种跳变通常反映局部区域的像素属性变化趋势。但视觉系统在提取图像边缘或轮廓时,更多的是一种全局或大范围区域整体特征指导下的局部细节精细化,非经典感受野大外周所存在的去抑制区域正是上述视觉机制实现的基础。因此本文围绕着非经典感受野特性和应用展开研究,首先对传统三高斯模型进行改进,利用非经典感受野结构实现对输入的自适应加权响应,从而增强图像中的边缘信息;其次本文改善了传统的同质抑制模型,利用颜色空间信息实现背景异质成分的抑制,获得抗干扰能力更强的边缘细节;最后本文模拟视觉显著性信息提取与反馈,并结合纹理同质抑制模型,凸显主体轮廓。本文主要工作及研究成果如下:

(1)建立三种基于非经典感受野机制的视觉处理模型。首先根据非经典感受野的三高斯模型,提出一种自适应加权响应模型,使检测结果具有更完整的边缘特性。其次改善传统的经典感受野同质抑制模型,使用三高斯函数代替 DOG 函数来模拟非经典感受野,采用 Lab 颜色空间和片相似性分别描述图像像素点的颜色以及纹理相似性,获得图像背景纹理及异质成分。最后提出显著性信息提取模型,根据经典和非经典感受野存在的去抑制区域的差异性产生突触抑制电流,并通过幂指数归一化的方式进行显著性提取。

(2)提出一种基于非经典感受野机制的图像边缘检测新方法。首先通过 Log-Gabor 滤波器对图像进行方向选择,采用 izhikevich 神经网络结合次序编码模拟视觉通道获得初步边缘图像,其次通过自适应加权响应模型对边缘信息进行增强,最后采用颜色同质抑制模型去除图像的异质成分,得到最终的边缘图像。经过 ROC 和信息熵方法评价,本文方法十一幅图 ROC 曲线下面积平均值为 0.850,平均信息熵为 0.395,均优于传统检测方法。

(3)提出一种基于视觉显著性信息的图像轮廓检测方法。首先通过高斯导函数获得初级视皮层(V1)的多方向梯度响应,同时基于非经典感受野机制设置 Leaky integrate-and-fire 神经网络的突触兴奋和抑制电流,利用神经元脉冲频率信息获得视觉显著性信息,反馈并调控 V1 的视觉响应,并结合纹理同质抑制模型实现对背景纹理的去除,最终得到图像轮廓。对 RuG 图像库中的图像进行轮廓检测,本文方法获得的综合评价指标 P 平均值为 0.50,与 ISO 等方法相比在性能上具有一定的优势。

关键词: 非经典感受野, 神经网络, 边缘检测, 显著性信息, 轮廓检测

ABSTRACT

Edge or contour as the key features in images, which have a great significance for the analysis or understanding of the target object in visual field. Recently, the edge or contour detection methods are mostly based on the jump characteristics of the pixel brightness and other attributes by mathematical description. So from the image space point of view, the jump characteristics usually reflect the change tendency of the pixel properties in the local area. But when the visual system extracts the edge or contour of the image, it always refines the local details of global or large-scale regional features, and the disinhibition region of non-classical receptive field is the basis for achieving the visual mechanism. Therefore, this paper focused on the characteristics and applications of non-classical receptive field. Firstly, the traditional three Gaussian function model was improved, we used non-classical receptive field structure to realize the adaptive weighted response of the input, in order to enhance the edge information of image. Secondly, we improved the traditional homogeneous inhibition model, and used color space information to inhibit the heterogeneous components, then we obtained the edge details, which had the better anti-interference ability. Finally, we simulated the extraction and feedback of visual salience information, and then we combined with the texture homogeneous inhibition model, so as to highlight the main contour. The main tasks and research results are listed as follows:

(1) Three visual processing models based on non-classical receptive field mechanism were established. Firstly, an adaptive weighted response model was proposed, which based on the three Gaussian function model of non-classical receptive field, in order to get more complete edge characteristics. Secondly, we improved the traditional homogeneous inhibition model, which based on classical receptive field, the three Gauss function was introduced to replace the DOG function to simulate the non-classical receptive field. Lab color space and slice similarity were used to describe the color and texture similarity of image pixels, in order to obtain background texture and heterogeneous components of images. Finally, a salience information extraction model was proposed. According to the difference between classical and non-classical receptive field, the synaptic inhibition current was generated, and the exponential normalization was used to extract the salience information.

(2) A new edge detection method based on non-classical receptive field mechanism was proposed. Firstly, we used the Log-Gabor filter to select the direction of the image information, we adopted izhikevich neural network and sequential coding to simulate the visual channel, so as to

obtain the preliminary edge image. Secondly, the edge information was enhanced by the adaptive weighted response model. Finally, the color homogeneity inhibition model was adopted to remove the heterogeneous components, and the final edge image was obtained. The proposed method was evaluated as following, the average area under the ROC curve was 0.850, and the average value of information entropy was 0.395, both of them were superior to the traditional detection methods.

(3) An image contour detection method based on visual salience information was proposed. Firstly, the derivative of Gaussian transform was introduced to get the gradient responses of primary visual cortex (V1) in multiple directions. The synaptic excitatory and inhibitory currents were specified based on receptive field mechanism in the LIF-neuron network. Then visual salience information was obtained by coding the spike frequency. And then the responses of neurons in V1 were modulated by feedback to get the image contour response. Finally, the image contour response combined with texture homogeneity suppression model to get the final contour. The pictures used for this experiment were selected from the RuG library. In this paper, the mean value of measure P was 0.50, which had a significantly better performance than the traditional methods including ISO.

Keywords: non-classical receptive field, neural network, edge detection, significant information, 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 神经元模型及视觉机制特性.....	7
2.1 神经元模型.....	7
2.1.1 Hodgkin-Huxley 神经元模型.....	7
2.1.2 IF 和 LIF 神经元模型.....	8
2.2 视觉感受野.....	9
2.2.1 经典感受野响应.....	10
2.2.2 感受野的 DOG 模型.....	10
2.2.3 非经典感受野.....	11
2.3 方向选择.....	11
2.4 神经编码.....	11
2.4.1 相位编码.....	12
2.4.2 频率编码.....	12
2.4.3 时间编码.....	12
2.4.4 次序编码.....	13
2.5 小结.....	13
第 3 章 基于非经典感受野机制的视觉处理模型.....	14
3.1 基于三高斯模型的自适应加权响应模型.....	14
3.2 基于非经典感受野的同质抑制模型.....	16
3.2.1 颜色同质模型.....	17
3.2.2 纹理同质模型.....	17
3.3 显著性信息提取模型.....	19
3.4 小结.....	20
第 4 章 基于非经典感受野机制的图像边缘检测.....	21
4.1 概述.....	21
4.2 原理.....	21
4.2.1 方向选择.....	21
4.2.2 izehkevich 神经元模型结合次序编码.....	22
4.2.3 基于三高斯模型的自适应加权响应模型.....	22
4.2.4 非经典感受野颜色同质抑制.....	23
4.3 边缘检测步骤.....	23

4.4 实验结果.....	25
4.5 分析与讨论.....	27
4.6 小结.....	29
第 5 章 基于视觉显著性信息的图像轮廓检测.....	30
5.1 概述.....	30
5.2 基本原理.....	30
5.2.1 改进后的 LIF 神经元网络模型.....	30
5.2.2 初级视皮层多方向梯度响应.....	31
5.2.3 信息流的显著性提取.....	31
5.2.4 信息流的显著性反馈.....	32
5.2.5 非经典感受野纹理同质抑制.....	32
5.3 轮廓检测步骤.....	33
5.4 实验结果.....	33
5.5 分析与讨论.....	36
5.6 小结.....	37
第 6 章 总结与展望.....	38
6.1 总结.....	38
6.2 展望.....	39
致谢.....	40
参考文献.....	41
附录.....	46

第 1 章 绪论

1.1 研究目的与意义

视觉是我们感知世界的重要途径。尤其是在适者生存，弱者淘汰的原始时代，远距离非接触式的视觉感知是人类查看天象、搜寻猎物以及躲避威胁的主要途径。因此对视觉感知机制^[1]及其应用研究，也是一种从视觉层面认识大脑、理解大脑以及创造大脑的过程，具有重要的研究意义。

伴随着工业信息化和智慧城市建设的浪潮，作为人工智能与图像处理技术^[2]结合的计算机视觉需求日益迫切，例如智慧交通中的智能电子车道、工业产品表面缺陷自动检测以及基于视频分析的公共场所异常行为预警等具体应用。应该指出的是，虽然计算机视觉应用越来越广泛，但依然处于方兴未艾的阶段，存在很多不完善的地方，还无法媲美生物视觉的感知能力，所以对它的研究还在持续深入中。

计算机视觉应用的基本对象是图像或视频，它们是一批庞大的数据矩阵或矩阵向量。以图像为例，从生物视觉系统的高效性来看，需要对数据矩阵进行一种类似稀疏分解的操作。被保留下来的元素从视觉感知的角度来评价，对应于图像的边缘或轮廓点。因此边缘和轮廓检测一直是模式识别和图像处理领域的研究热点之一。对它们的准确检测，意味着将大大提高后续图像分析或理解过程的性能。边缘与轮廓特征是识别图像信息的关键，因为边缘与轮廓是图像中关键区域与背景区域的交界，也是图像像素信息发生剧变的地方，其中蕴含的信息最多，通过边缘和轮廓的描述，可以对整幅图有基本的了解，为之后的图像目标识别和检测等奠定了基础。

边缘检测^[3]大幅削减了图像的数据量，去除了很多无关紧要的背景信息，而保留图像中相对关键的结果。传统的边缘检测算法中比较常用的一阶算法有 Sobel 算子^[4]，可以分别对水平边缘与垂直边缘进行检测，通过对像素位置的加权处理，可降低边缘的模糊程度，但是缺少对人类视觉特性的模拟，所以在一些复杂的图像中目标与背景的区分不是很明显；Canny 算子^[5]是使用比较广泛的二阶算法，有两个阈值对检测结果进行调控，检测效率较高，适用场合较广，但与 Sobel 算子相同，面对信息丰富的图像检测效率偏低。轮廓与边缘不同，主要反映主体部分的轮廓信息，是边缘信息中的一部分，为后续基于形状的目标检测提供有用信息。轮廓检测需要对图像背景纹理的边缘信息进行抑制，主要凸出主体部分的形状特征，所以传统的边缘检测算法并不能够进行有效的轮廓检测。

正如前述，上述检测方法是建立在对图像像素亮度或颜色的梯度变化等数学描述基础上，虽然从输入输出之间的关系来看，具有一定的合理性。但毕竟最卓越的边缘或轮廓检测系统是人类自身的视觉感知系统，因此抛开黑箱化的检测模型，而是从研究视觉系统本身入手，

尤其是视觉神经计算中较为关键的非经典感受野模型，淡化传统方法过于关注像素点局部邻域的做法，强调感受野全局和局部相结合的特性。不仅将有利于我们深入了解视觉系统内在的机制和原理，还将为包括边缘和轮廓检测在内的图像处理应用带来一些全新的思路。

1.2 国内外研究现状

目前国内外对于边缘以及轮廓的检测众多，有的学者通过对传统检测方法进行算法上的优化，以求获得更好的检测效果；有的学者深入研究生物视觉系统，对生物视觉信息层层分析，将一系列的视觉机制用于图像特征的提取，显著信息的加强，对算法的性能进行大幅提高，取得了较好的检测效果。有的学者通过神经科学的电生理实验以及数据分析，建立对应的神经电生理模型以及数学计算方法，将此类仿生模型运用于图像处理中并结合视觉机制特性，在边缘以及轮廓检测中效果显著。

目前对传统检测方法进行算法优化的方式多样，比如利用各向异性非线性扩散滤波技术的 Canny 自适应边缘检测算法^[6]，在保持 Canny 算法原有优点的基础上，提高了对细节信息的提取和假边缘的抑制；通过引入马尔科夫随机能量函数的水平集算法^[7]，整合局部空间信息进行医学图像分割，提高传统算法的鲁棒性和检测效率；通过对梯度图像进行频域分割以去除无效边缘，之后采用空域阈值分割方法进行轮廓检测^[8]。

通过对神经科学中视觉信息处理方法的深入研究，进一步推动图像处理的发展。例如研究发现初级视皮层的感受野对刺激具有方向选择特性，Daugman 利用 Gabor 滤波器模拟感受野的特性。Gabor 滤波器^[9]的优点在于对刺激的方向敏感性，能对边缘信息进行较完整的提取。然而图像过于复杂时，其抑制环境中纹理边缘的效果较一般，会对边缘的提取造成干扰。随着对感受野的深入研究发现，在传统感受野外周还存在一块区域对经典感受野起着复杂的调节作用，即是非经典感受野(Non-Classical Receptive Field, NCRF)的区域。Grigorescu 通过 Gabor 滤波器的视觉特性模拟经典感受野对边缘的响应^[10]，同时使用经过半波整流后的高斯差函数代表非经典感受野，通过侧向调控的作用对纹理边缘进行抑制，得到效果凸出的图像轮廓。Xu^[11]和 Chen 通过对非经典感受野和经典感受野的空间差异对神经元响应的影响发现^[12,13]，在非经典感受野中存在去抑制区域，会加强中心区域的刺激响应。其去抑制性^[14,15]一定程度上减少了经典感受野检测过程中造成的低空间频率信息的损失，拓宽了对视觉细胞的认识，从而对图像的检测提供了更加充分的神经信息基础。在 Grigorescu 研究的基础上，Du 等提出了一种基于复合感受野模型的图像轮廓检测方法^[16]，引入复合的非经典感受野模型进行纹理抑制的同时保护图像较弱的轮廓。Zeng 等也提出了类似复合感受野的蝶形非经典感受野抑制方式^[17,18]，通过不同方位的刺激使得非经典感受野分别表现出抑制性和去抑制性，检测到的轮廓更为平滑。吴贺等建立复杂感受野的响应函数，提出一种基于朝向和对比度特性的非经典感受野抑制模型^[19]，方法指出当经典和非经典感受野的朝向一致时无朝向差，抑制作用最大，朝向垂直时则抑制作用最弱，用指数函数定义抑制权重，并用经典和非经典感受野相对位置的变化代表对比度特性，亦通过权值体现抑制作用的强弱，提高主体轮廓检测的准确度。

Yang 等^[20,21]通过提取图像的方向、光强等多种特征,对经典感受野输出的响应图像进行大范围的去抑制作用,得到的轮廓检测结果相比单一特征结果有明显优势,并通过视锥细胞颜色拮抗的处理,对彩色图像进行边缘检测,得到的边缘图像对背景纹理的抑制效果不错。除了以经典和非经典感受野为主的图像检测,还有很多方法利用其他视觉机制进行图像的处理。Zaghloul 等^[22]提出一种模式视网膜细胞的数学模型,该模型可以对图像的亮度和对比度进行调节。李林雨等研究视觉频率多通道对整体视觉进行认知^[23],陈勇等利用此发现,结合彩色图像的视觉感知特性用于彩图的评估^[24]。Kuang 等人认为视网膜的固视微动特性能进行冗余信息去除及对信息进行初步的提取^[25]。桑农等提出了一种基于初级视皮层抑制的轮廓检测方法^[26],神经元受到周围环境同方向刺激的抑制,方向变化的地方抑制作用减弱,使边缘信息得到显著性的表现。

上述方法中都运用到视觉感知机制,细胞感受野朝以及细胞的方向选择等等,建立合适的数学模型对图像进行处理,并通过提取特征、优化参数等方法提高检测的准确度,虽取得了不错的效果,但这些研究还未曾考虑到具体的神经元之间信息传递的方式,比如神经编码机制。HH 神经元模型的提出为学者了解神经信息传递奠定了重要的基础,人们开始运用神经网络对图像进行分析和处理。Raudies 等研究方向选择与 V1 等区域的神经元群体放电特性,提出相应的神经视觉机制模型对目标进行检测^[27]。Webb 等研究运动信号与神经元群体放电特性的关系,认为其不同于一般的图像统计,存在特殊的视觉解码对运动信号进行传递与分析^[28]。很多学者又对神经网络产生的脉冲序列进行研究,他们认为研究脉冲信号能对图像有更加清晰的认识。Masud 等分析了多个脉冲序列连接对信息处理功能的影响,估计了从参考脉冲序列到目标脉冲序列的影响权重矢量^[29]。Sacerdote 等利用数学工具研究了神经元之间产生脉冲序列的依赖性,为神经网络的研究提供了充分的实验支持^[30]。金小贤等对视网膜上神经元信息的传递进行深入研究,提出一种仿生图像处理方法^[31]。Dingran Lu 等人构建人工神经网络处理医学方面图像^[32],相比 Sobel 算子的图像处理效果更好。Waldemark 等利用 PCNN 对预处理后的卫星图像进行边缘检测^[33],将边缘信息从图像中分割出来,但由于神经网络的参数固定,并未真正从视觉编码的角度出发,与视觉系统中的神经网络存在较大差距。而在 PCNN 基础上,Kerr 提出一种基于六边形感受野的脉冲神经网络(Spiking Neural Network, SNN)^[34],六边形感受野不同于传统的矩形感受野,对图像的检测效果有进一步提升。Cui 等通过视觉注意机制处理^[35],结合图像细节以及其轮廓特征对图像质量进行评估。Liao 等提出的基于视觉编码机制的图像弱边缘检测方法,通过符合生物视觉特性的视觉感知机制结合神经网络实现图像边缘检测^[36],相对于其他简单运用神经网络的方法具有一定的优势。

综上为国内外有关图像处理中结合视觉信息处理机制的研究趋势,一方面注重对视觉机制的深入研究,例如非经典感受野部分,发现其不仅存在抑制作用,同样存在去抑制的效果。通过对其朝向和对比度的具体研究,还有其模型的选择,在图像的边缘和轮廓检测的信息处理过程中起到重要的作用。又如视觉显著性、方向选择等视觉机制特性同样对图像处理起着

巨大的帮助。另一方面研究注重于神经元网络的建立以及神经编码的拓展。研究表明,生物视觉信息的传递离不开神经元信息的传递,而其传递以神经编码的方式完成^[37]。综合以上两方面研究趋势,本文构建符合视觉特性的神经元网络,以对非经典感受野特性的研究为重点,结合多种视觉机制特性,对图像进行边缘以及轮廓的检测,以期获得好于传统方法的研究结果,提供了一种基于生物视觉特性的图像处理新思路。

1.3 本文主要工作及创新点

本文研究的内容主要是以非经典感受野特性为重点,结合多种视觉机制以及建立神经网络模型进行处理,进行图像边缘以及轮廓的检测。一方面研究单个神经元之间信息传递的方式以及神经网络的特性,另一方面研究非经典感受野等视觉机制在图像处理中起到的具体作用。首先对神经编码方式以及神经网络模型进行分析,其次考虑视觉机制中的方向选择、显著性信息,以及非经典感受野的多种特性。在边缘检测中,利用 izhikevich 神经元模型以及次序编码对方向选择图像进行边缘提取,结合基于非经典感受野三高斯模型的自适应加权响应模型以及颜色同质抑制模型,使最终得到的边缘图像拥有不错的效果;在轮廓检测中,利用二维高斯导函数获得初级视皮层(V1)的多方向梯度响应,同时基于非经典感受野机制设置 Leaky integrate-and-fire 神经网络的突触兴奋和抑制电流,通过神经元脉冲频率信息,获得视觉显著性信息,反馈并调控 V1 的视觉响应,得到轮廓响应。在对该轮廓响应进行纹理同质抑制处理得到最终的轮廓检测结果。

论文具体研究工作包括以下几个方面:

(1)建立了三种基于非经典感受野特性的视觉处理模型。首先根据非经典感受野的三高斯模型,提出一种自适应加权响应模型,利用像素的局部对比度实现对非经典感受野兴奋、抑制以及去抑制区域参数的自适应动态调整,通过带权值感受野窗口对图像进行遍历,实现对边缘信息的增强作用,使结果具有更完整的边缘特性。其次改善传统的经典感受野同质抑制模型,使用三高斯函数代替 DOG 模型来模拟非经典感受野,采用 Lab 颜色空间描述图像像素点的颜色相似度,根据片相似性描述图像像素点之间的纹理相似度,经过颜色和纹理同质间相互抑制的方式,将纹理及异质成分显示出来并去除。通过实验对比发现经过同质抑制后的图像,背景纹理以及异质成分减少。最后提出显著性信息提取模型,非经典感受野的去抑制作用将有助于局部空间特征的提取,所以根据经典和非经典感受野存在的去抑制区域的差异性产生突触抑制电流,并通过幂指数归一化的方式进行显著性提取。此三种模型的研究分析为之后对于边缘以及轮廓的检测奠定了基础。

(2)提出一种基于非经典感受野机制的边缘检测新方法。首先使用具有视觉感知特性的 Log-Gabor 滤波器对图像进行方向选择,采用 izhikevich 神经元单元构建神经网络,利用次序编码方式对视觉信息流进行定量的分析。考虑到 Log-Gabor 滤波后的方向选择图像边缘信息的缺失,根据非经典感受野的三高斯模型特性,提出自适应加权响应模型对图像边缘信息进行增强处理,同时考虑到图像存在较多的纹理,对彩色图像进行 Lab 颜色空间信息的提取,

通过两点颜色的差异性进行非经典感受野颜色同质抑制处理。以含有丰富细节的菌落图像边缘检测为例，结果表明，获得的边缘信息更完整，背景纹理更少。经过 ROC 和信息熵方法评价，本文方法要优于传统检测方法。

(3)提出了一种基于视觉显著性信息的图像轮廓检测方法。首先模拟低级视皮层提取视觉激励的多方向梯度信息，并通过降低冗余度的方法来弱化纹理特征，关注于具有整体性和连续性的轮廓细节；利用局部移动窗获取 V1 层面的多方向梯度响应，结合一种基于突触电流调制的经典和非经典感受野机制，采用高级视皮层的神经元脉冲发放频率作为图像的显著性提取结果，并通过模拟高级视皮层对视觉信息流的反馈调控作用，将来自高级视皮层的显著性信息和来自低级视皮层的多方向梯度信息进行融合，获得初步轮廓响应；最后将该轮廓响应通过纹理同质抑制模型处理，去除背景纹理，从而获得最终的轮廓检测结果。对 RuG 图像库中的图像进行轮廓检测，与 ISO 等传统方法相比，本文方法的轮廓图与基准轮廓图比较获得的综合评价指标 P 值最高，性能上具有一定的优势。

1.4 论文的组织结构

论文的组织结构以及各章节的内容如下：

第一章绪论，首先提出视觉感知机制及其应用研究的重要意义，引出对图像进行边缘以及轮廓检测的重要性，并对传统的边缘以及轮廓检测算法进行了介绍并分析其不足；再对国内外在视觉感受野等视觉机制方面的研究及用途进行介绍；亦对神经网络模型的发展进行总结；给出了本文的主要研究内容与创新，并进行分析与说明；最后列出了本文的组织结构以及各章节的内容。

第二章首先对神经元模型进行分析，如最经典的 HH 神经元模型、IF 神经元模型以及 LIF 神经元模型。再对视觉机制进行分析，其中以视觉感受野的发展为重点，从经典感受野的研究发现再到非经典感受野的提出，以及视觉机制中具备的方向选择特性。最后对各种神经编码方式进行详细介绍，指出神经编码对图像处理的重要作用。

第三章着重于研究视觉机制中的关键部分，即非经典感受野。通过建立三种模型，基于非经典感受野三高斯模型的自适应加权响应模型、改善后的同质抑制模型以及显著性信息提取模型分别对其在视觉系统中起到的作用进行深入研究。通过初步的实验以及仿真处理，从结果看出基于非经典感受野特性的模型在图像边缘增强、背景纹理抑制等方面取得不错的效果，为后续的基于非经典感受野等视觉机制以及神经网络结合的图像处理应用奠定了基础。

第四章提出一种基于非经典感受野机制的边缘检测新方法。首先使用 Log-Gabor 方向选择，并通过 izhikevich 神经网络模型结合次序编码的处理，获取神经元首次放电的信息，得到初步的边缘图像，再结合基于三高斯模型的自适应加权响应模型进行边缘增强，并通过颜色同质抑制使得图像背景异质成分凸显并去除，得到最终的边缘检测结果。通过绘制 ROC 曲线以及信息熵的评定得出本文方法好于传统边缘检测方法，获得的边缘细节也好于其他对比方法。

第五章提出了一种基于视觉显著性信息的图像轮廓检测新方法。模拟低级视皮层提取视觉激励的多方向梯度信息，注重于降低冗余度的方法来弱化纹理特征，结合一种基于突触电流调制的经典和非经典感受野机制，获取图像的显著性信息，将其反馈并调控低级视皮层，得到轮廓响应。最后将该轮廓响应通过纹理同质抑制模型处理，去除背景纹理，从而获得最终的轮廓检测结果。对 RuG 图像库中图像进行轮廓检测，本文方法在错检和漏检指标上达到一种均衡，从综合评价指标 P 看出，相比其他方法具有更理想的轮廓检测性能。

第六章对本文的研究工作以及成果做了详细的总结，并对后续的研究进行展望。

第2章 神经元模型及视觉机制特性

人类获取的大部分外界信息都来自视觉感知，强大的视觉感知系统能获取外界中各个目标的形状、颜色、纹理，并对目标后续的识别与匹配提供信息。所以将人类的视觉感知系统运用于对机器视觉中图像的理解与分析有着重大的意义。随着对视觉感知系统的深入研究，发现神经网络在视觉信息的获取、分析、识别中起着基础而重要的作用。研究已知视网膜获取到外界的信息，通过神经网络的分析处理将这些信息传递到视皮层形成清晰的图像。而神经网络具体的处理方式正在学者们深入的探索中逐步揭晓。生物科学家研究发现神经元作为神经网络中最基本的一部分，通过突触之间的连接，相互作用形成了高效稳定的神经网络，当接收到外界信息时，通过神经元放电对电信号进行编码，然后通过突触的作用将信息传递给周围的神经元，而不同的神经元类型显示出不同的视觉编码特性，多种视觉编码特性经过神经网络的处理，在大脑视皮层呈现出视觉图像。由此可见，研究机器视觉的关键之一就是探索神经元以及神经编码特性的作用方式。

2.1 神经元模型

神经元是神经网络构建的基础，神经元通过动作电位传递信息。研究发现，细胞膜内外的钠、钾离子会根据神经元收到的刺激而发生改变从而引发动作电位。基于这类研究结果，科学家们分别提出了自己的数学模型用来表示神经细胞膜的电位变化情况。其中最经典的模型之一是 HH(Hodgkin-Huxley)神经元模型，精确的数学建模充分体现了神经电生理特性，因此也被广泛的学者使用；后来学者对 HH 神经元模型进行简化，提出一种 IF 神经元模型，因考虑了主要的神经元放电特性，所以计算量上大大减少；学者们考虑到不应期等神经生理特性在 IF 神经元模型基础上提出了 LIF 神经元模型，进行更好的仿生模拟。

2.1.1 Hodgkin-Huxley 神经元模型

HH 神经元模型^[38]是 Hodgkin 和 Huxley 两位科学家通过研究神经元放电特性而建立的，描述了神经元细胞膜电流、电压和时间的变化关系。HH 神经元模型为后续对神经编码机制的研究奠定了稳固的基础，式(2.1)是 HH 神经元模型的具体描述：

$$\begin{cases} C_m \frac{dV}{dt} = I_{ext} - g_{Na} m^3 h (V - V_{Na}) - g_K n^4 (V - V_K) - g_L (V - V_L) \\ \frac{dm}{dt} = \frac{(m_\infty(V) - m)}{\tau_m(V)} \\ \frac{dh}{dt} = \frac{(h_\infty(V) - h)}{\tau_h(V)} \\ \frac{dn}{dt} = \frac{(n_\infty(V) - n)}{\tau_n(V)} \end{cases} \quad (2.1)$$

式中, C_m 和 V 分别为细胞膜电容和膜电压; g_{Na} 、 g_K 、 g_L 表示各个离子通道电导的最大值, 对应的逆导电位为 V_{Na} 、 V_K 、 V_L ; m 和 h 表示钠离子的通透性, n 表示钾离子的通透性; $m_\infty(V)$ 、 $h_\infty(V)$ 、 $n_\infty(V)$ 和 $\tau_m(V)$ 、 $\tau_h(V)$ 、 $\tau_n(V)$ 分别表示 m 、 h 、 n 的饱和值以及对应的时间常数。图 2.1 对 HH 模型进行描述。

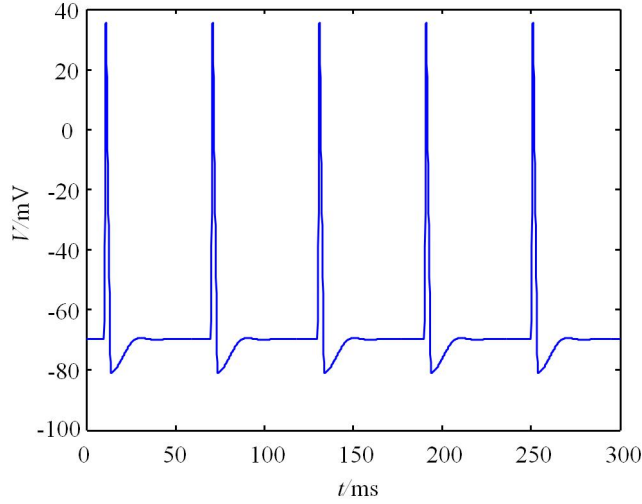


图 2.1 HH 神经细胞膜电位变化示意图

HH 神经元模型能够对神经元电生理特性进行精确的描述, 但是由于其信息的传递的数学描述过程过于复杂, 计算量过大导致不能适用于规模较大的神经网络集群。

2.1.2 IF 和 LIF 神经元模型

IF(integrate-and-fire)神经元模型是以 RC 电路为模型建立的, 考虑了神经元脉冲放特性, 简化了细胞膜离子运动, 并忽略了不应期等神经元特性。其数学公示非常简洁, 广泛应用于神经电生理的研究工作中, IF 神经元模型如式(2.2)所示:

$$C_m \frac{dv}{dt} = I \quad (2.2)$$

式中, C_m 、 v 、 I 分别为神经元膜电容、膜电压、刺激电流。图 2.2 为 IF 模型细胞膜电位变化示意图。

根据 IF 模型的基本特性, 学者们考虑到不应期等神经生理特性而提出了 Leaky integrate-and-fire 模型, 简称为 LIF 模型^[39]。LIF 模型通过添加不应期、漏电导等特性能更好地描述神经元的放电过程, 其数学模型如式(2.3)所示:

$$\begin{cases} C_m \frac{dv}{dt} = -g_l v + I \\ v = v_{reset} \quad (v > v_{th}) \end{cases} \quad (2.3)$$

式中, C_m 、 v 、 I 与式(2.2)描述相同, g_l 、 v_{reset} 、 v_{th} 分别为漏电导、静息电位和脉冲发放阈值电位。当膜电位 v 达到脉冲发放阈值电位时, v 将瞬时被置为脉冲峰值电位, 同时神经元将进入绝对不应期, 不再响应任何激励。图 2.3 为 LIF 神经元模型的细胞膜电位变化示意图, (a)

不考虑不应期, (b)考虑不应期。

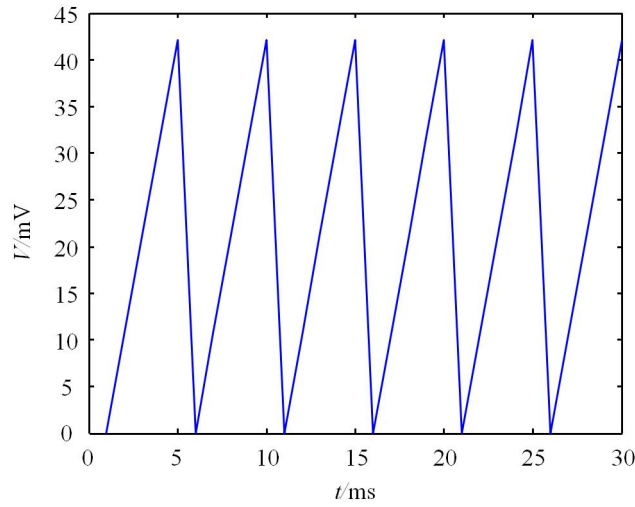


图 2.2 IF 神经细胞膜电位变化示意图

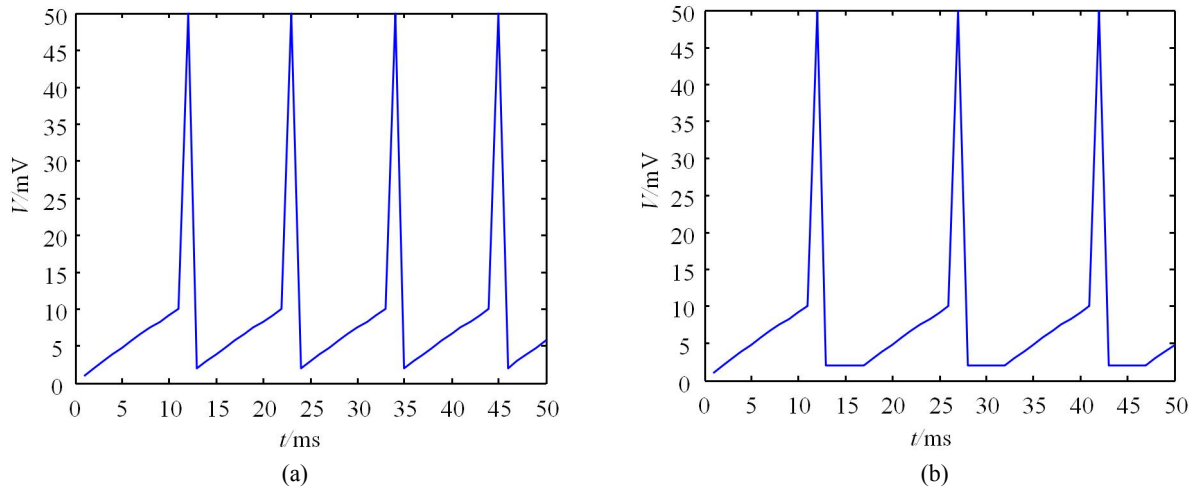


图 2.3 LIF 神经细胞膜电位变化示意图
(a) 不考虑不应期; (b) 考虑绝对不应期

IF 神经元模型只考虑了神经元脉冲发放特性, 所以计算非常简便, 适用于大规模的神经网络中, 但是其简单的数学模型中丢失了相当一部分的神经元特性, 例如脉冲发放中的不应期缺失, 所以对神经元电生理特性的描述不够准确。而 LIF 模型通过添加不应期、漏电导等特性能更好地描述神经元的放电过程, 对神经元生理描述更为准确。

2.2 视觉感受野

Marr 和 Hildreth 两位科学家根据对动物做的视觉实验结果, 提出了一种双高斯模型 (DOG), 并且能很好地解释人类视觉行为; 汪九云等在研究发现视觉感受野时空特性, 并用模型进行描述; Munder 等通过视觉感受野提取图像的特征进行研究; 李朝义等在对猫的视网膜神经节细胞进行电生理实验发现, 细胞感受野外还存在很大的一个区域, 该区域通过减弱经典感受野抑制区域对中心区域的抑制作用, 从而减少因抑制作用导致的图像低空间频率信息的损失, 这个区域称为非经典感受野。于是李朝义等提出非经典感受野模型能有效地对图

像的边缘信息加强, 不仅具有对图像的视觉感知特性, 而且通过恢复低空间频率信息保证边缘信息的完整性, 经典和非经典感受野模型如图 2.4 所示。

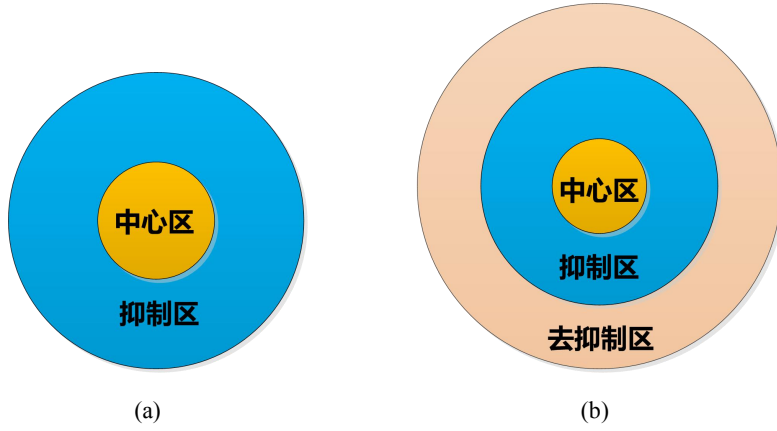


图 2.4 感受野模型示意图
(a) 经典感受野模型; (b) 非经典感受野模型

2.2.1 经典感受野响应

研究发现 Gabor 能量算子能很好的描述复杂细胞对刺激的响应, 所以很多学者用 Gabor 函数的数学模型来描述经典感受野的响应, 如式(2.4)所示:

$$g(x, y, \theta) = \exp\left(-\frac{\tilde{x}^2 + \gamma^2 \tilde{y}^2}{2\sigma^2}\right) \cdot \cos\left(2\pi \frac{\tilde{x}}{\lambda} + \varphi\right) \quad (2.4)$$

式中, $\tilde{x} = x \cos \theta + y \sin \theta$, $\tilde{y} = -x \sin \theta + y \cos \theta$, γ 、 σ 、 λ 分别表示感受野的椭圆率、空间大小以及波长参数; θ 表示朝向角度, 当输入的刺激朝向与感受野的最优朝向角度相同时, 刺激响应也就达到最大值; φ 是相位偏移量, 同理朝向角度, 当输入的刺激相位信息与感受野最优相位相同时, 输出值最大。

2.2.2 感受野的 DOG 模型

1965 年 Rodieck 通过研究提出了一种双高斯差模型(DOG), 对细胞经典感受野的输入输出特性进行定量描述。DOG 模型通过模拟经典感受野的同心圆结构和特性, 用数学式很好的表现其性质, 如式(2.5)所示:

$$DOG(x, y) = Ae^{-(x^2+y^2)/\sigma_1^2} - Be^{-(x^2+y^2)/\sigma_2^2} \quad (2.5)$$

式中, A 代表感受野中心区域的兴奋程度, B 代表感受野抑制区域的抑制程度, x 和 y 代表感受野上空间变量, σ_1 代表中心区域的半径, σ_2 代表抑制区域的半径, 通常抑制区域半径大于中心区域半径。

在 DOG 模型中, 中心区域与抑制区域被描述为同心圆结构, 通过二维高斯函数来对两部分刺激分布进行计算, 从而得到感受野的输入输出数值。DOG 模型能够有效的描述经典感受野的特性, 将其运用于图像处理中能有效获取边缘信息, 但同样由于抑制区域的作用导致

图像的亮度和低空间频率信息受到损失。

2.2.3 非经典感受野

通过感受野同心圆结构的提出以及 DOG 模型的建立,已经能对感受野的特性做出基本的说明。但是经过学者的进一步研究发现,在传统的经典感受野之外还存在一大块区域,此区域同样对中心区域的兴奋程度起到调节作用,称为非经典感受野^[40,41]。McIlwain 等研究得到,在经典感受野之外的区域移动刺激光点,细胞的响应会随之增强,在当时学者们并没有太过关注这个发现。然而近几十年来对经典感受野外区域的研究越来越多,更为复杂的视觉现象可能通过对这块区域的研究得到合理的解释。基于此,李朝义等对猫的外膝体细胞以及视网膜神经节细胞做了电生理实验,发现在经典感受野外存在大片区域能够抵消抑制区域对中心兴奋区域的抑制效果,从而起到增强中心兴奋度的作用,这片区域被称为去抑制区。同样研究得到去抑制区不仅不会削弱传统感受野模型对图像边缘信息的增强效果,还能保留更多的低空间频率信息,从而使得边缘信息更为全面。

2.3 方向选择

研究得到视皮层中神经元对于图像不同方向的响应程度并不一致,对于特定的某一方向会达到一个最大程度的响应^[42],这说明神经元感受野中对于输入的刺激进行了一定的选择响应。通过方向选择,视皮层不同神经元的功能得到体现,对不同的图像进行处理获得最佳的效果。

1946 年 Gabor 在前人的基础上提出了一种方法可在频域和时域内获得最佳的局部效果,称为 Gabor 变换。Gabor 函数具有很好的频率选择和方向选择特性,经过研究发现 Gabor 函数特性与视觉神经细胞的工作原理比较接近,所以将 Gabor 滤波器广泛运用于图像处理当中。Gabor 函数的窗函数为高斯函数,如式(2.6)所示:

$$h(x, y) = \frac{1}{2\pi} \exp\left(-\frac{x^2 + y^2}{\sigma^2}\right) \exp[j2\pi\omega_0(x \cos \theta + y \sin \theta)] \quad (2.6)$$

式中, θ 、 ω_0 分别为正弦函数的方向角和空间频率, σ 表示空间常数。Gabor 函数有实部和虚部两部分组成,实部表现为纹理几何信息,虚部则是纹理边缘信息。

2.4 神经编码

神经系统在接收到外界刺激后,通过一定的方式进行信息的传输和处理,称为神经编码。通过神经元生理电实验,可以发现神经元处理信息确实有一定的规律和特征,由于神经编码复杂而且多样,对于选取编码的识别和分析非常重要,有助于分析脑皮层对刺激信息的处理与整合。有研究认为神经元是根据单位时间内脉冲发放次数对信息进行编码。稀疏编码理论^[43]认为,刺激某一块区域的神经元,并不是所有的神经元都会被激活而传递信息,实际上只有一小部分神经元完成对信息的编码和传递。而神经元群体编码理论^[44]认为,单个神经元对信息的处理存在不稳定性,所以神经元通过组成一个群体对信息进行编码,本文也是通过构建

神经网络来模拟视觉神经的处理过程。

人类视网膜接收到外界的信息后通过神经元将其编码为动作电位，再进行神经元之间的传递在大脑视皮层形成清晰的图像，所以信息的传递也正是动作电位的传导。研究神经脉冲编码的关键在于清楚动作电位的形成和传递过程。研究认为一个尖脉冲即为一个动作电位，在通过细胞间突触传递中动作电位的形状和幅度保持稳定，使得信息不会因为传递而导致丢失或者差错，所以研究清楚动作电位的传递过程就能清楚了解视觉系统获取信息的方式，对机器视觉研究至关重要。而目前对脉冲编码的研究也分很多方向，如频率编码、相位编码、次序编码、时间编码等等，研究通过数学的方式将其表现出来，才能运用在计算机视觉系统的处理中，以下是对各种方法的介绍。

2.4.1 相位编码

在视觉神经以及其他的脑部神经中，输入信号会进行周期性的震荡，神经元在输入基准信号的不同相位上触发脉冲，根据相位的差异进行编码称为相位编码。相位编码对光的利用率接近百分之百，在图像处理中对噪声也有很好的鲁棒性。相位编码虽能用于表示大脑皮层间信息的传递，但是也在一定程度上忽略了脉冲神经元的信息。

2.4.2 频率编码

频率编码是根据神经元动作电位发放频率而得到的一种脉冲编码方式，首先通过不同的刺激作用于神经元，再根据其频率变化的情况进行分析^[45]。所以频率编码认为脉冲平均发放率中包罗了所有的刺激信息，如式(2.7)所示：

$$Q = \frac{1}{T} \sum_{t=1}^T h(t) \quad (2.7)$$

式中， T 代表一段脉冲发放时间，如果 t 时刻有脉冲电位产生，则 $h(t)=1$ ，否则 $h(t)=0$ 。

频率编码相对于其他的脉冲编码方式比较简单且容易计算，而脉冲放电时间 T 的取值十分重要，影响到这段时间时候会有脉冲电位的存在，所以灵活性不足。同样深入研究频率编码发现，激发神经元动作电位产生需要一个阈值，当刺激信号的强度值达到这个阈值时，即进行动作电位发放。不同的频率编码所需要的刺激阈值不同，其中包含的信息还需要更深入的研究。

2.4.3 时间编码

任何与刺激有关的信息都会存在脉冲发放时间内，跟频率编码相比，时间编码的每个神经元中包含的信息更多。研究发现时间编码能表达频率编码中无法表达的脉冲发放信息，比如神经元脉冲发放的次序、神经元首次脉冲的潜伏时间。而时间编码的局限性在于对神经元的刺激反应太过于敏感容易导致可靠性下降，并且对刺激开始时间的把握也不够精准，所以有一定的试用范围。

2.4.4 次序编码

人的视觉系统对外界信息获取的速度极快，实验发现从视网膜神经细胞接受刺激到在大脑视皮层形成图像的时间只需要短短 150ms，所以在视觉系统中，过于简单的脉冲编码方式未必能在短时间完成复杂的信息传递，而且神经元互相作用，通过突触连接构成一个庞大的神经网络，所以脉冲编码的方式更加值得深究^[46]。基于脉冲编码的快速有效性，有研究者提出神经系统是通过合成神经元集群来编码的，而且神经元集群的各个神经元只进行一次电位发放以达到外界刺激信息的快速传递，也就是说神经元首次放电的先后次序已经构成了对外界信息的编码^[47]，这种编码方式被称成为次序编码。

次序编码也是时间编码的一种形式，跟频率编码、相位编码相比，次序编码的编码速度更快，通过神经元集群编码体现信息的空间特性，广泛应用于快速提取图像特性信息。本文也采用次序编码对图像信息进行处理，根据感受野中神经元响应的先后顺序设立对应的权值，体现次序编码中对刺激响应先后顺序的差异性。

科学家对于神经元动作电位的产生和神经元间刺激的传输做了很多研究和解释，但是神经脉冲编码到底是以怎样的方式工作，至今没有明确统一的定义。除了上述的神经脉冲编码理论外，Wang 等也提出了一种新的能量编码理论，希望通过能量来表达神经编码的基本规律，同时对神经信息传递通过能量的方式进行转化。总之每种脉冲编码理论都有一定的适用范围，还需要研究者进一步的探究。

2.5 小结

本章主要对传统的神经元模型和神经编码方式，以及经典和非经典感受野、方向选择等视觉机制特性进行介绍。HH 神经元模型具备很好的电生理特性，但不适用于较大规模的神经网络集群中；IF 神经元模型具有高效的计算性，但是其模型中丢失了相当一部分的神经元特性；LIF 模型能更好地描述神经元的放电过程。神经编码中相位编码具有较好的鲁棒性，但是忽略了脉冲神经元的信息；频率编码相对于其他的脉冲编码方式比较简单且容易计算，但是灵活性不足；时间编码对刺激开始时间的把握也不够精准，所以有一定的试用范围；而次序编码相比其他神经编码方式速度更快，通过神经元集群编码体现信息的时空特性，广泛应用于快速提取图像特性信息。介绍了视觉机制的方向选择特性的 Gabor 函数，经典感受野的提出以及非经典感受野的发现，使得对于视觉信息处理的研究越来越丰富，为后续本文的研究奠定基础。

第3章 基于非经典感受野机制的视觉处理模型

自从非经典感受野提出以来，学者们对它建立多种模型为探究其在生物视觉信息传递中起到的具体作用。本章针对非经典感受野的特性，建立三种模型对非经典感受野的作用进行模拟，三种模型分别为自适应加权响应模型、同质抑制模型以及显著性信息提取模型，通过本章的研究为本文后续的边缘以及轮廓检测奠定基础。

3.1 基于三高斯模型的自适应加权响应模型

视觉感受野是生物视网膜神经节以及外膝体细胞对接收到的外界刺激进行分析处理的关键部分，随着对其深入研究发现，经典感受野之外还存在大块区域，能够抵消抑制区域对中心区域的抑制作用，这片区域称为去抑制区域，也就是非经典感受野的发现过程。随着研究深入，李朝义等经过研究并建立了基于非经典感受野的三高斯(TDOG)模型，相比较 DOG 模型更好得满足生物合理性的要求，去抑制区作用后不仅不会削弱 DOG 模型对图像边缘信息的增强效果，还能保留更多的低空间频率信息，从而使边缘信息更为全面，数学模型如式(3.1)所示：

$$\begin{cases} TDOG(x, y) = Ae^{-(x^2+y^2)/\sigma_1^2} - (Be^{-(x^2+y^2)/\sigma_2^2} - Ce^{-(x^2+y^2)/\sigma_3^2}) \\ A = \frac{1}{2\pi\sigma_1^2} \\ B = \frac{1}{2\pi\sigma_2^2} \\ C = \frac{1}{2\pi\sigma_3^2} \end{cases} \quad (3.1)$$

式中， $TDOG(x, y)$ 表示神经元感受野内某点的兴奋响应； A 、 B 、 C 分别代表感受野中心区域的兴奋程度、抑制区域的抑制程度和去抑制区域的去抑制程度； σ_1 、 σ_2 、 σ_3 分别代表中心兴奋区域、抑制区域和去抑制区域的半径大小。通过计算每个点兴奋响应的代数和得到整个神经元的响应程度。

然而研究发现非经典感受野中三块区域的处于一种动态变化的状态，输入刺激的局部对比度提高会使得非经典感受野抑制区域的半径缩小而使得抑制强度的减少，相对的中心区域与去抑制区域的强度有所增强，所以在本文通过改变三块区域半径 σ_1 、 σ_2 和 σ_3 的大小对三高斯模型进行自适应调控。

输入刺激局部对比度的计算方式：中心像素点设为 $templet(5)$ ，以此为中心的 3×3 邻域进行升序排列记为 $Templet(i)$, $i = 1, 2, \dots, 9$ ，通过计算标准差得到局部对比度 LC ，计算方式如式(3.2)所示：

$$\begin{cases} average = (\sum_{i=1}^9 templet(i)) / 9 \\ LC = (\sum_{i=1}^9 (templet(i) - average)^2)^{\frac{1}{2}} \end{cases} \quad (3.2)$$

研究发现 Sigmoid 函数能对兴奋细胞进行很好的模拟, 本文通过改进其函数实现对 σ_1 、 σ_2 和 σ_3 的动态调整, 当局部对比度 LC 增加, 抑制作用减弱, 所以 σ_2 的减小, 而中心兴奋度以及去抑制作用得到增强, σ_1 和 σ_3 增大, 具体实现方式如式(3.3)所示:

$$\begin{cases} \sigma_1 = 0.4 + \frac{1}{50 + e^{-LC}} \\ \sigma_2 = 1 + \frac{1}{10 + e^{10 \times LC}} \\ \sigma_3 = 4 + \frac{1}{10 + e^{-LC}} \end{cases} \quad (3.3)$$

式中, 看出 LC 即使变化很大时, 不会对 σ_1 、 σ_2 和 σ_3 产生剧烈影响, 防止由于数值变化过大导致图像纹理噪声的增加。

图 3.1 中以 5×5 的非经典感受野窗口为例, 具体三层区域对图像处理的方式如图所示, 其中黄色区域代表非经典感受野的中心区域, 红色为抑制区域, 淡黄色为去抑制区域。

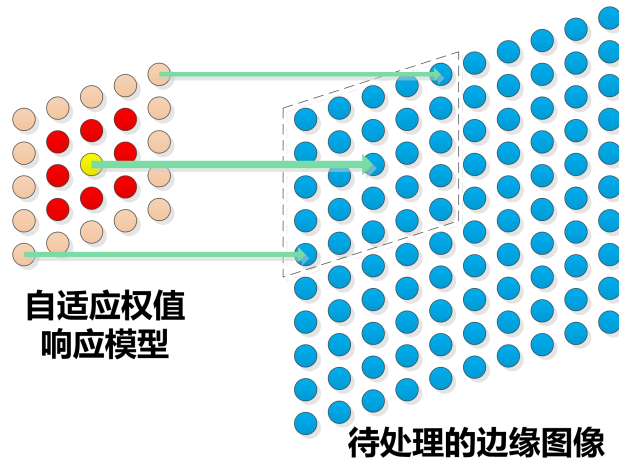


图 3.1 5×5 窗口为例的自适应加权响应模型示意图

三高斯数学模型中 σ_1 、 σ_2 和 σ_3 的值根据式(3.3)得到, 根据 σ_1 、 σ_2 和 σ_3 的值可以确定一个 5×5 具备权值的窗口, 并对该窗口进行升序排列, 记为 $weight(i), i = 1, 2, \dots, 25$, 与神经网络处理后的时间信息图像一一对应, 中心点设为 $image(13)$ 的 5×5 窗口中的时间信息按照升序排序记为 $image(k), k = 1, 2, \dots, 25$, 具体取值根据式(3.1)得到, 首先将 5×5 的感受野窗口的中心点与时间信息图像第一个点为中心的 5×5 邻域对应, 各个元素对应相乘再加和, 得到中心点在外界刺激作用后在兴奋、抑制以及去抑制区域动态作用下的响应强度, 如式(3.4)所示:

$$sum = \sum_{i=1, k=1}^{25} (weight(i) \times image(k)) \quad (3.4)$$

将 sum 值赋予图像中心点 $image(13)$ ，即得到该点经过非经典感受野自适应处理后的响应强度。

以上述方式为例，对每个点进行上述处理，每个点根据邻域得到的局部对比度不同， σ_2 和 σ_3 的值也各不相同，所以感受野窗口权值实现自适应变化，利用式(3.4)得到该点的响应强度，对整幅图像从左到右，从上到下进行遍历得到全新边缘信息图像，本文自适应加权响应模型处理前后的边缘信息如图 3.2 所示，经过处理后的边缘相比较未处理之前更加完整。

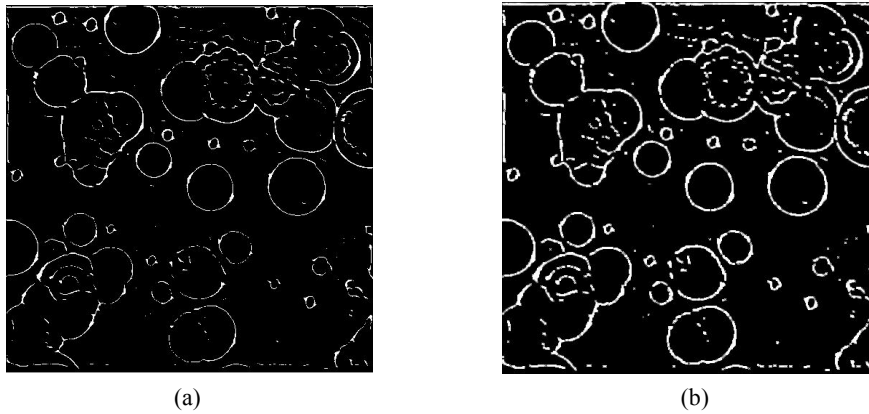


图 3.2 自适应加权响应模型处理前后结果对比图

(a) 自适应加权响应模型处理前；(b) 自适应加权响应模型处理后

3.2 基于非经典感受野的同质抑制模型

研究发现通过兴奋作用可以增强初级视皮层神经元对边缘的响应，而对同方向的刺激进行抑制可以减弱对背景边缘的响应。

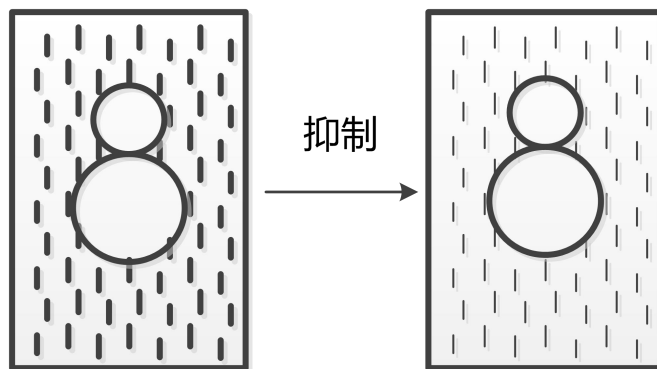


图 3.3 非经典感受野抑制作用对显著边缘的影响

从图 3.3 中可以看出，在感受野的处理中，纹理和光栅会对刺激响应产生抑制作用，抑制作用导致同质部分的神经元响应减弱，由于同质部分发生变化的区域通常是边缘位置，抑制作用减弱使得神经元响应增强，检测到的边缘更为显著。而当感受野中刺激方位一致时，对中心区域的刺激响应会增强。所以人的视觉系统能够容易得从复杂的环境中找到每个目标

的边缘。

本文改善传统的经典感受野同质抑制模型，使用三高斯函数代替 DOG 模型来模拟非经典感受野的结构，采用 Lab 颜色空间描述图像像素点的颜色相似度，根据片相似性描述图像像素点之间的纹理相似度，经过颜色和纹理同质间相互抑制的方式，将纹理及异质成分提取并去除。颜色同质抑制适用于彩色图像，而纹理同质抑制适用于复杂而具有较多纹理特征的图像。

3.2.1 颜色同质模型

Lab 颜色空间是一种符合生物生理特征的颜色系统，其中的 L 分量代表像素的亮度，从纯黑到纯白，取值范围[0,100]， a 分量代表从红色到绿色，取值范围为[127,-128]， b 代表的是从黄色到蓝色，取值范围为[127,-128]。Lab 的色彩空间大于 RGB 图像，而且通过数学方法将 RGB 图像转化为 Lab 颜色空间时，可自动弥补滤色片等器件造成的误差。本文使用更加符合生物视觉特性的 Lab 颜色空间对图像的像素点进行颜色相似性分析。

图像任意两像素点分别为 $(x, y) = (l, a, b)$ ， $(x', y') = (l', a', b')$ ，两点之间的颜色相似性度量函数如式(3.5)所示：

$$d((x, y), (x', y')) = \left((l - l')^2 + \tau(a - a')^2 + \tau(b - b')^2 \right)^{\frac{1}{2}} \quad (3.5)$$

式中， τ 代表颜色差异性权值，本文 $\tau = 1$ ，颜色相似度模型如图 3.4 所示。

图 3.4 图像任意两像素点 Lab 颜色相似度模型

3.2.2 纹理同质模型

复杂图像中蕴涵着丰富的纹理信息，怎样区分背景纹理并将主体目标凸显出来是计算机视觉及图像处理的一个关键部分。

本方法利用片相似性来描述纹理同质性，对输入图像 $I(x, y)$ 进行处理。首先令 $P(x, y)$ 为图像 $I(x, y)$ 上以像素点 (x, y) 为中心的 3×3 的领域中所有像素灰度值的有序向量，将片 $P(x, y)$ 排列成 9 维的向量 $P(x, y) = (I(x-3, y-3), \dots, I(x+3, y+3))^T$ ，即图像片都是通过将图像 3×3 局部内的像素点按照行或列排列后得到的一个向量。片 $P(x, y)$ 代表像素点 (x, y) 的纹理特性，片 $P(x', y')$ 代表像素点 (x', y') 的纹理特性：

$$d(P_l(x, y), P_l(x', y')) = \frac{1}{9} \left(\sum_{l=1}^9 (P^l(x, y) - P^l(x', y'))^2 \right)^{\frac{1}{2}} \quad (3.6)$$

式中, $d(P_l(x, y), P_l(x', y'))$ 表示由图像构造的片空间上的相似性度量函数, 体现该图像中任意两点 (x, y) 与 (x', y') 的纹理相似程度, 纹理相似度模型如图 3.5 所示。

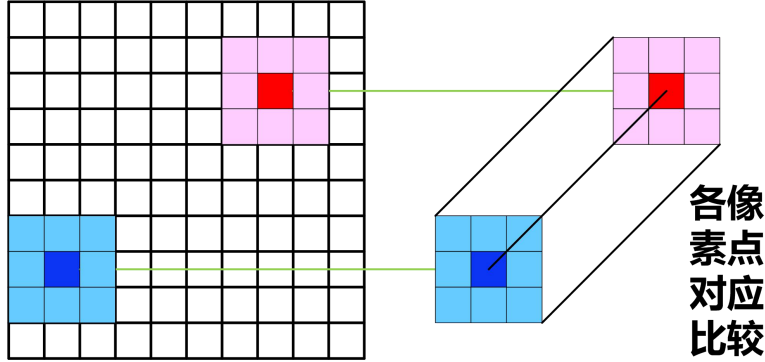


图 3.5 图像任意两像素点纹理相似度模型

本方法利用体现非经典感受野的三高斯函数 $TDOG(x, y)$ 式(3.1)来模拟抑制环境。像素点相似性度量对抑制作用的影响表示如式(3.7)所示:

$$\eta(x, y) = \frac{N(TDOG(x, y))}{\|N(TDOG)\|_1} \quad (3.7)$$

$$N(TDOG(x, y)) = \begin{cases} TDOG(x, y), & TDOG(x, y) > 0, \\ 0, & else. \end{cases} \quad (3.8)$$

式中, $\|N(TDOG)\|_1$ 表示三高斯函数结果的范数表示, 用于调控抑制作用的强弱。

颜色相似性响应计算如式(3.9)所示:

$$CE(x, y) = \sum_{(x', y')} \eta(x', y') \times e^{-d((x, y), (x', y'))} \times D(x', y') \quad (3.9)$$

式中, $D(x', y')$ 代表待处理的边缘图像, $\eta(x', y')$ 代表像素点颜色相似性度量对抑制作用的影响。

纹理相似性响应计算如式(3.10)所示:

$$WE(x, y) = \sum_{(x', y')} \eta(x', y') \times e^{-d(P_l(x, y), P_l(x', y'))} \times SE(x', y') \quad (3.10)$$

式中, $SE(x', y')$ 代表待处理的轮廓响应图像, $\eta(x', y')$ 代表像素点纹理颜色相似性度量对抑制作用的影响。

本文对 Gabor 能量算子的响应结果分别进行颜色、纹理同质抑制处理。

从图 3.6 中可以看出经过颜色或者纹理同质抑制的处理, 图像的纹理以及异质成分得到较好的去除。

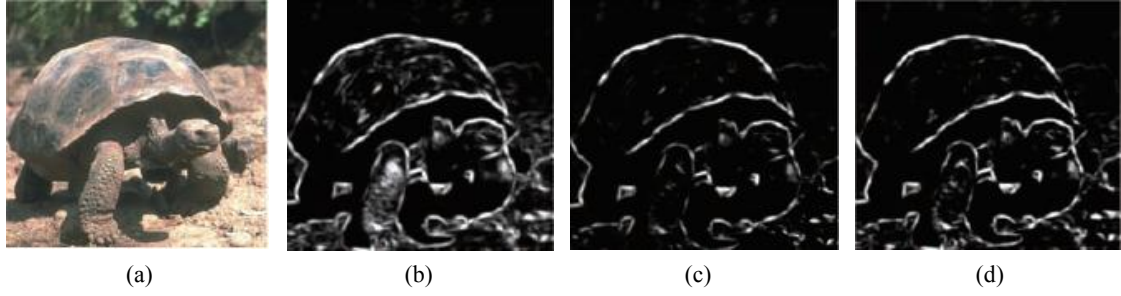


图 3.6 同质抑制模型处理示意图

(a) 原始图像；(b) Gabor 能量算子；(c) 颜色同质抑制结果；(d) 纹理同质抑制结果

颜色和纹理同质抑制模型是根据图像上任意两点的颜色和纹理相似度来对其同质成分进行判断，同质性越高得到同质抑制响应越小，同质性越低则同质抑制响应越大，通过此来判断图像的异质成分和背景纹理，并将其去除，从而凸显主体轮廓信息。需要说明的是，在本文第四章的图像边缘中，检测对象为彩色菌落图，并且菌落图中的纹理信息较少，所以采用颜色同质抑制模型；而第五章中的轮廓检测中，检测对象为灰度自然图像，而纹理信息丰富，遂采用纹理同质抑制模型进行图像处理。

3.3 显著性信息提取模型

在视觉信息流的传递和处理过程中，存在自底向上和自上向下两种通路，目前显著性信息获得方式主要通过自底向上的方法，通过提取场景中的纹理、颜色、方位等属性得到。

研究表明，非经典感受野的去抑制作用将有助于局部空间特征的提取。因此本文将充分利用经典和非经典感受野对于视觉激励的不同响应，利用两者的差异性来设置突触抑制电流。考虑到经典感受野区域外还有更大范围的非经典感受野起到去抑制的效果，所以在研究局部空间特征时，计算经典和非经典感受野对视觉响应的差异性，即是去抑制区域的差异性，由三高斯模型式(3.1)可得到去抑制区域的表达方式，如式(3.11)所示：

$$\Delta\phi(x, y) = \max\left\{\sum_{(m,n)} I_{syn}(x+m, y+n) \cdot Ce^{-(m^2+n^2)/\sigma_3^2}, 0\right\} \quad (3.11)$$

式中， $I_{syn}(x, y)$ 代表总的突触兴奋电流， $Ce^{-(m^2+n^2)/\sigma_3^2}$ 代表去抑制区域，取 $C = 1/2\pi\sigma_3^2$ ，非经典感受野的尺寸通常取为经典感受野尺寸的 4 倍，所以 $\sigma_3 = 4\sigma$ 。

将式(3.11)所获得的感受野差异性，经过式(3.12)所示的幂指数归一化操作，获得显著性信息图像：

$$S(x, y) = \exp\left(-\frac{\Delta\phi(x, y)^2}{\varepsilon}\right) \quad (3.12)$$

式中，参数 ε 为抑制强度调节参数，设为 0.08。

如图 3.7 为显著性信息提取作用前后脉冲发放频率图。显著性信息提取后的图像在保留边缘信息完整性的基础上对纹理进行较好的抑制，为后续通过计算空间差异性信息来突出图像轮廓的提供了新的思路。

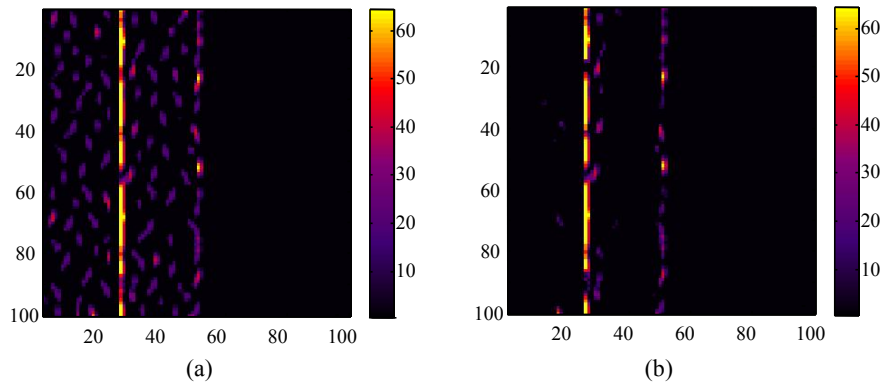


图 3.7 显著性信息提取机制作用前后脉冲发放频率图

(a) 作用前; (b) 作用后

3.4 小结

本章着重于研究视觉机制中的关键部分，即非经典感受野。通过建立三种模型，自适应加权响应模型、同质抑制模型以及显著性信息提取模型分别对其在视觉系统中起到的作用进行深入研究。通过初步的实验以及仿真处理，从结果看出基于非经典感受野机制的模型在图像边缘增强、背景纹理抑制等方面取的不错的效果，为后续更多基于视觉机制以及神经网络结合的图像处理应用奠定基础。

第 4 章 基于非经典感受野机制的图像边缘检测

边缘检测是机器视觉处理图像的重要过程，对目标边缘检测的准确率直接影响到机器对客观对象的理解。目前除了一些较为传统的检测方法，例如 canny、sobel 算法，还有一些利用生物神经系统对图像进行处理的方法，但是大部分方法只是简单的利用神经网络进行边缘检测，而具体对视觉机制中的非经典感受野、脉冲编码、方向选择等特性的综合考虑还是非常少，并且很多方法难以得到定量的检测数据作为评定好坏的参考。所以本文提出了一种基于非经典感受野机制的图像边缘检测方法，利用脉冲编码中较为新颖的次序编码思路，通过建立神经网络记录神经元首次放电时间，结合非经典感受野的特性以及同质抑制作用对图像进行处理。对显微镜拍摄到的菌落图进行图像边缘检测，结果显示边缘相比较传统方法更为清晰，边缘信息更加丰富，并且通过定量的指标验证结果的有效性。

4.1 概述

考虑到人类视觉系统对图像处理效果的显著性，本文利用神经元网络对图像进行信息的提取，并考虑到计算的高效性以及神经电生理特性，采用 izhikevich 神经元模型构建神经网络。首先经过符合人类视觉特性的 Log-Gabor 滤波器，进行多方向选择，重构得到边缘图像；然后将图像经过神经网络处理，根据每个神经元首次脉冲动作电位的时间信息得到新的时间信息边缘图像；并使用基于非经典感受野三高斯模型的自适应加权响应模型进行处理，得到新的边缘图像；将该边缘图像经过颜色同质抑制处理，最后通过二值化处理得到结果图像。

4.2 原理

本文使用 Log-Gabor 滤波器对原始图像进行处理，采用 izhikevich 神经网络结合次序编码模拟视觉通道获得初步边缘图像，同时由于 Log-Gabor 滤波处理后会损失一部分的边缘信息，考虑到非经典感受野对图像边缘信息的增强作用，引入自适应加权响应模型，并考虑到边缘图像有背景异质成分的产生，引入非经典感受野颜色同质抑制模型对异质成分进行去除，最终得到符合视觉特性的边缘检测图像。

4.2.1 方向选择

研究外界刺激在视觉系统中的传递，我们可以得知初级视皮层中的不同神经元具有不同功能，同时每种神经元对刺激特定方向的刺激响应会强于其他方向的刺激响应，称之为方向选择。所以本文边缘检测方法首先对初级视皮层的方向选择特性进行模拟。传统的 Gabor 函数具有很好的方向选择特性，但是由于函数本身带宽的限制，所以当频谱覆盖增大时，只能通过增加 Gabor 滤波器的数量进行处理，然而使得计算效率大大下降。而本文使用的 Log-Gabor 函数不仅具备 Gabor 函数的方向选择特性，同时拥有更多的优点：人类的视觉系统是非线性的，而 Log-Gabor 函数的对数体现了非线性特性，函数的尾拖与自然图像的幅度

谱衰落相似，符合视觉特性；Gabor 滤波器会产生非零的直流分量，使得图像处理受到亮度的影响，而 Log-Gabor 不会产生直流分量；同时 Log-Gabor 函数对带宽无限制，使得计算量比 Gabor 函数小很多，而且精度更高。综合上述原因，为充分模拟人类的视觉特性，本文采用 Log-Gabor 滤波器对图像初级视皮层进行方向选择，数学模型如式(4.1)所示：

$$H(\omega, \theta) = \exp\left\{\frac{-(\log(\omega/\omega_0))^2}{2(\log(\sigma_\omega/\omega_0))^2}\right\} \exp\left\{\frac{-(\theta-\theta_0)^2}{2\sigma_\theta^2}\right\} \quad (4.1)$$

式中， ω 、 θ 分别为待滤波图像的频率和方向， ω_0 、 θ_0 分别为中心频率和滤波器的方向， σ_ω 、 σ_θ 分别决定了带宽的大小和方向。

对灰度化后的图像进行初级视皮层多方向选择，得到八个方向选择结果（角度分别为 $0^\circ, 22.5^\circ, 45^\circ, 67.5^\circ, 90^\circ, 112.5^\circ, 135^\circ, 157.5^\circ$ ），通过叠加的方式得到整体的方向选择响应。

4.2.2 izhikevich 神经元模型结合次序编码

izhikevich 神经元模型^[48]与第二章中介绍的 HH 以及 IF 神经元相比属于折中的神经元模型，不仅能保留较多的神经元特性，又能相应地减少计算量，因此该模型在一定程度上具有 HH 神经元模型精确表示神经元电生理特性的能力，又拥有 IF 神经元模型具备的高效计算能力，数学模型如式(4.2)所示：

$$\begin{cases} \frac{dv}{dt} = 0.04v^2 + 5v + 140 - u + I_{ext} \\ \frac{du}{dt} = a(bv - u) \\ \text{if } v \geq v_{thresh}, \text{ then } v \leftarrow c \quad u \leftarrow u + d \end{cases} \quad (4.2)$$

式中， v 、 u 、 I_{ext} 、 v_{thresh} 分别为细胞膜电位、细胞膜恢复变量、刺激电流和阈值电位； a 、 b 、 c 、 d 为该模型的参数。当 $v \geq v_{thresh}$ 时，神经元进行脉冲发放，同时重置 v 为参数 c ， u 被设置为 $u + d$ ；如果 $v < v_{thresh}$ 时则神经元不进行脉冲发放。

神经元模型不仅在构建大规模神经网络时具有计算的高效性，更兼顾生物合理性，所以本文采用 izhikevich 神经元模型构建神经网络。同时考虑到视觉系统处理信息高效的能力，研究认为神经元在脉冲首次发放时已经对信息进行了编码并传递到下一个神经元中，这种编码方式称为次序编码。本文采用 izhikevich 神经元模型结合次序编码的方式，图像每个像素点对应一个神经元，组成复杂的神经网络，并通过设立阈值 v_{thresh} ，模拟神经元首次脉冲发放，并记录下相应的时间信息，通过时间信息体现神经网络处理后的图像边缘响应。izhikevich 神经元模型与次序编码的结合应用既体现视觉信息传递的高效性，又满足生物生理特性。

4.2.3 基于三高斯模型的自适应加权响应模型

随着非经典感受野的不断研究，学者研究得到非经典感受野的三高斯(TDOG)模型，用于替代传统的 DOG 模型，其加入去抑制区作用后不仅不会削弱传统感受野模型对图像边缘信

息的增强效果，还能保留更多的低空间频率信息，从而使得边缘信息更为全面。本文研究该三高斯模型，根据以像素点邻域内局部对比度的变化，提出一种自适应加权响应模型，通过中心兴奋区域半径 σ_1 、抑制区域半径 σ_2 以及去抑制区域半径 σ_3 的变化调整感受野窗口的权值，响应强度由式(3.4)得到，作用方式如图 3.1 所示。

4.2.4 非经典感受野颜色同质抑制

本文选择彩色菌落图进行边缘检测，由于图像颜色信息丰富，而纹理信息偏少，所以本文采用改善后的颜色同质抑制模型对其进行处理。首先将原 RGB 图像转化 Lab 颜色空间图像，得到图像的亮度通道 L 以及颜色通道 a 和 b ，任意图像两点为 $(x, y) = (l, a, b)$ ， $(x', y') = (l', a', b')$ 。具体式见(3.5)、(3.7)~(3.9)。颜色同质抑制模型是根据图像上任意两点的颜色相似度来对其同质成分进行判断，同质性越高得到颜色同质抑制响应 $CE(x, y)$ 越小，同质性越低则颜色同质抑制响应 $CE(x, y)$ 越大，通过此来判断图像的异质成分，并将其去除，从而凸显主体轮廓。

4.3 边缘检测步骤

步骤 1：首先将原始图像 $I(x, y)(x = 1, 2, \dots, M; y = 1, 2, \dots, N)$ 灰度化之后，进行 Log-Gabor 滤波处理，得到八个方向选择结果（角度分别为 $0^\circ, 22.5^\circ, 45^\circ, 67.5^\circ, 90^\circ, 112.5^\circ, 135^\circ, 157.5^\circ$ ），记为 $f_k(x, y)(k = 0, 1, \dots, 7)$ ，用式(4.3)对八个方向进行叠加得到重构边缘图像：

$$result(x, y) = \sum_{k=0}^7 f_k(x, y)(x = 1, 2, \dots, M; y = 1, 2, \dots, N) \quad (4.3)$$

步骤 2：首先构造点阵神经网络 $Neuron(x, y)(x = 1, 2, \dots, M; y = 1, 2, \dots, N)$ ，每个神经元模型如式(2.3)所示。将重构边缘图像结果 $result(x, y)$ 输入到点阵神经网络 $Neuron(x, y)$ 中，其中重构边缘图像每个像素点的值分别对应神经元输入刺激 I_{ext} ，再记录下每个神经元 $Neuron(x, y)(x = 1, 2, \dots, M; y = 1, 2, \dots, N)$ 首次脉冲动作电位的发放时间，得到新的时间信息图像为 $T(x, y)(x = 1, 2, \dots, M; y = 1, 2, \dots, N)$ 。

步骤 3：以 $image(13)$ 为中心点，三高斯模型中 σ_1 、 σ_2 和 σ_3 的值由式(3.2)和(3.3)得到，根据 σ_1 、 σ_2 和 σ_3 的值可以确定一个 5×5 带有权值的窗口，并对该窗口进行升序排列，记为 $weight(i), i = 1, 2, \dots, 25$ ，与神经网络处理后的时间信息图像一一对应，中心点设为 $image(13)$ 的 5×5 窗口中的时间信息按照升序排序记为 $image(k), k = 1, 2, \dots, 25$ ，具体取值根据式(3.1)得到，首先将 5×5 的权值窗口的中心点与时间信息图像第一个点为中心的 5×5 对应，各个对应元素相乘再加和，表现中心点在外界刺激作用后在兴奋、抑制以及去抑制区域动态作用下响应强度 sum ，如式(3.4)所示，并将 sum 值赋予图像中心点 $image(13)$ ，即得到该点经过非经典感受野动态处理后的响应强度。通过图像从左到右，从上到下进行如此遍历，即得到经过非经典感受野动态处理后的边缘图像 $D(x, y)(x = 1, 2, \dots, M; y = 1, 2, \dots, N)$ 。

步骤 4：颜色同质抑制处理过程：首先将原始菌落图像的 RGB 图像转换为符合视觉感知特性的 Lab 颜色空间，得到图像的亮度通道 L 以及颜色通道 a 和 b ，所以图像点 $(x, y) = (l, a, b)$ ，点 $(x', y') = (l', a', b')$ ，根据颜色同质抑制模型以及式(3.5)、(3.7)~(3.9)得到颜色相似性响应结

果 $CE(x, y)$ 。

去除颜色同质抑制后的边缘图像结果如式(4.4)所示：

$$B(x, y) = D(x, y) - CE(x, y) \quad (4.4)$$

步骤 5：对 $B(x, y)(x = 1, 2, \dots, M; y = 1, 2, \dots, N)$ 进行二值化处理得到最终的边缘检测结果 $F(x, y)(x = 1, 2, \dots, M; y = 1, 2, \dots, N)$ 。

图 4.1 为本方法的边缘检测步骤图。

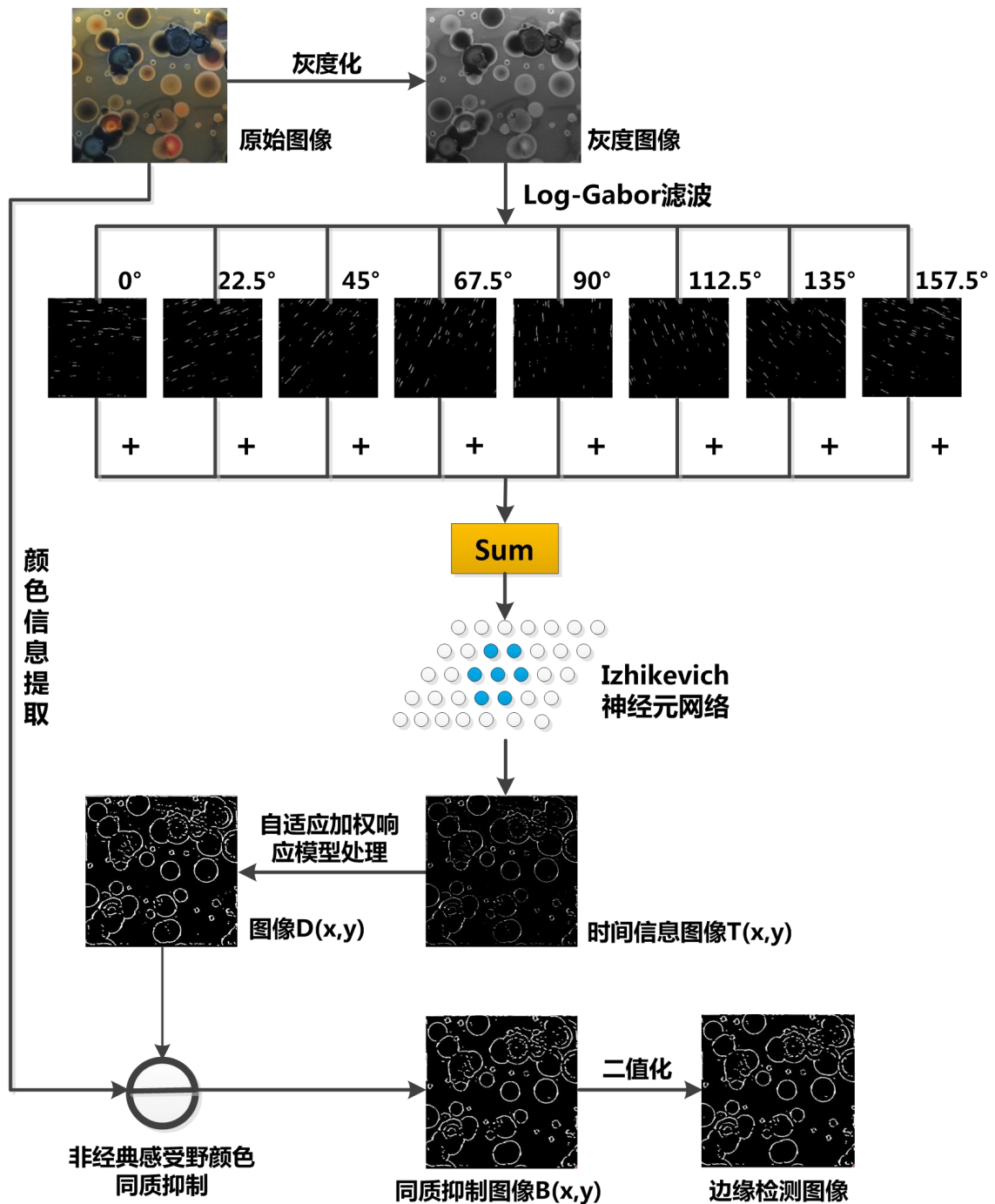


图 4.1 边缘检测步骤图

4.4 实验结果

本文使用的菌落图均为实验专用而采集的，采用仿自然光悬浮式暗视野系，对培养皿上的菌落进行千万级成像。选取比较有代表性的 Sobel 算法、Canny 算法以及 DOG 算法与本文方法比较，采用定量的 ROC 指标^[49-52]和信息熵^[53,54]对各个算法的优劣进行比较。

图 4.3~图 4.5 分别是各种方法对图 4.2 中三幅菌落图的检测效果。图(a)为三幅菌落原图的灰度化结果，图(b)、图(c)、图(d)和图(e)分别是使用 Sobel 算法、Canny 算法、DOG 算法和本文方法得到的边缘检测结果。

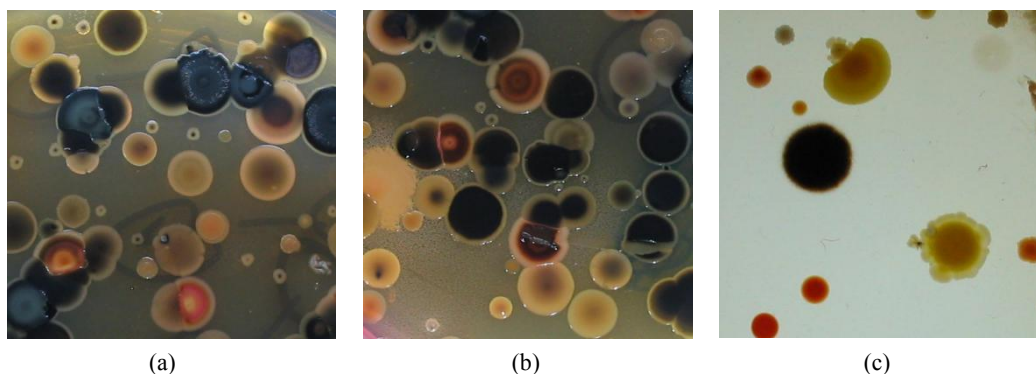


图 4.2 菌落示意图

(a) Picture_1; (b) Picture_2; (c) Picture_3

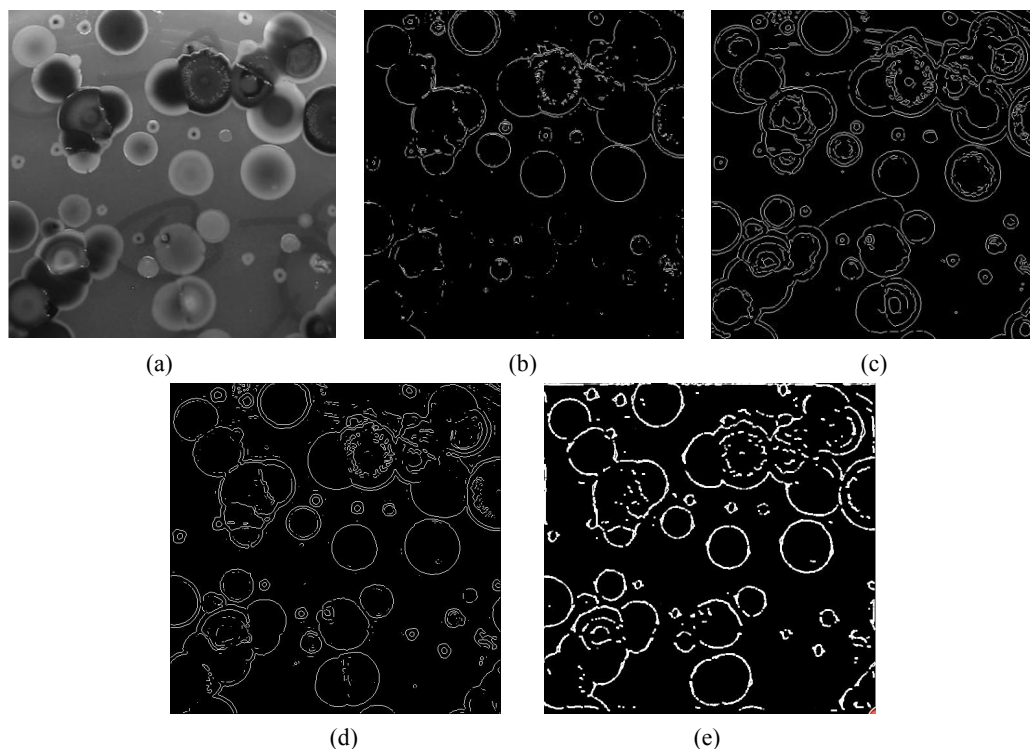


图 4.3 Picture_1 边缘检测结果

(a) Picture_1 灰度图; (b) Sobel 算法; (c) Canny 算法; (d) DOG 算法; (e) 本文方法

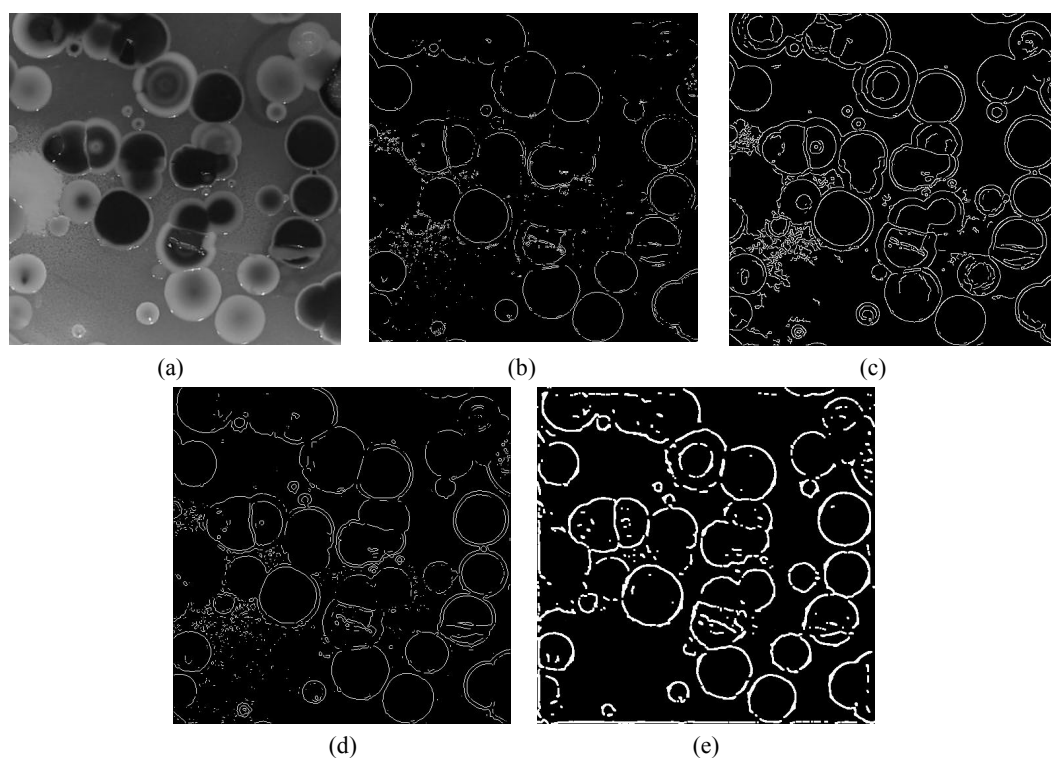


图 4.4 Picture_2 边缘检测结果

(a) Picture_2 灰度图; (b) Sobel 算法; (c) Canny 算法; (d) DOG 算法; (e) 本文方法

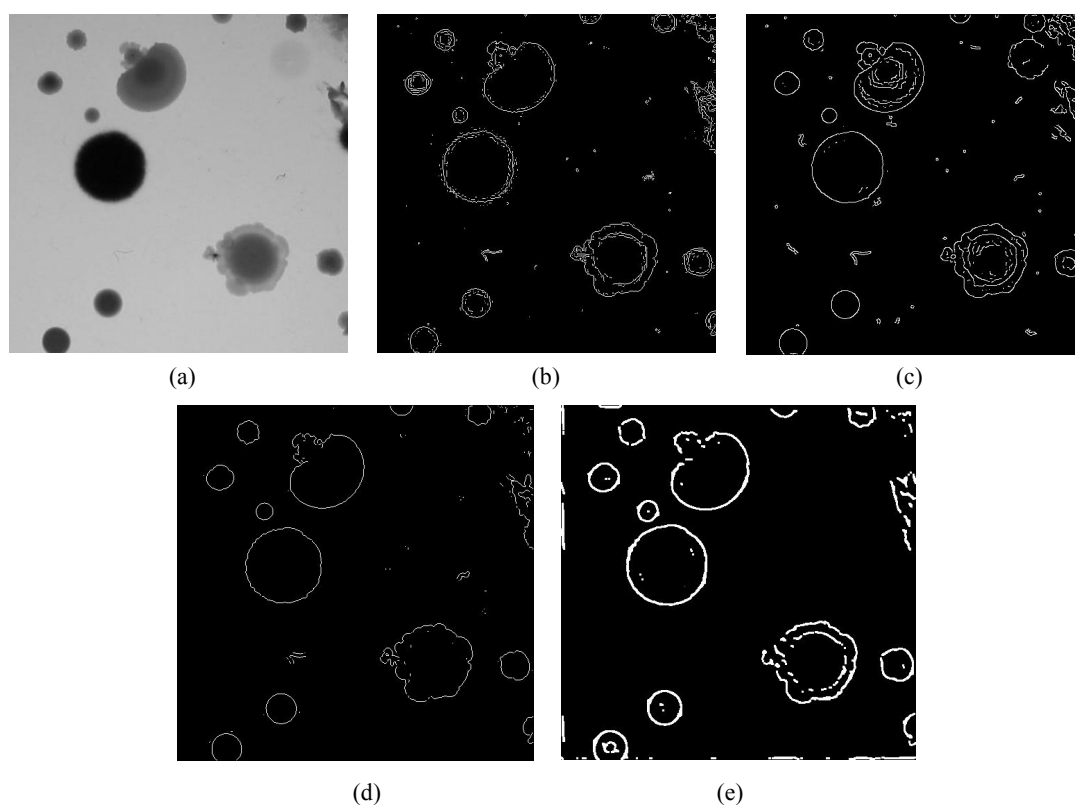


图 4.5 Picture_3 边缘检测结果

(a) Picture_2 灰度图; (b) Sobel 算法; (c) Canny 算法; (d) DOG 算法; (e) 本文方法

各幅图中的图(b)是 Sobel 算法的检测结果, Sobel 算法是基于—阶导数的边缘检测, 通过逼近导数检测边缘, 从图中可以看出对菌落的检测效果, 与 Canny 算法检测不同, 背景线条

并没有被误检测出来,但是存在检测不充分的情况,从图 4.3 中(b)发现正下方细胞边缘并没有被检测到,说明其检测当中也存在相当一部分边缘信息丢失的情况;各幅图中的图(c)是 Canny 算法的检测结果, Canny 算法是基于梯度幅值的双阈值法对图像进行检测,从图中可以看出对菌落的边缘达到过检测的效果,如图 4.3(c)的菌落检测出多层边缘,而且将菌落环境中的较暗的线条也进行了边缘检测,导致真正的边缘信息得不到很好的观察效果,图 4.4 中过检测情况更为严重,几乎无法分辨边缘;各幅图中的图(d)是 DOG 算法的边缘检测结果, DOG 算法使用高斯滤波进行平滑处理,再进行差分处理,从视觉机制的角度中看出是模拟经典感受野抑制区域对中心区域的作用,如图 4.4(d)有较多呈点状的纹理成分,但是边缘信息比较丰富,相对于上述的 Canny 算法和 Sobel 算法从肉眼效果上看有较大的提高;各幅图中的图(e)是本文的边缘检测算法,本文算法采用神经网络进行处理,再使用基于非经典感受野特性的自适应加权响应模型对图像边缘信息进行增强,并经过非经典感受野的颜色同质抑制作用,本文方法相比 DOG 方法得到的边缘细节更加全面,同样可以看出图中的背景异质成分也得到很好的去除。综合上述三种方法和本文方法,从人眼视觉观察的角度可以看出本文方法对菌落图边缘检测更加准确,效果比较突出。

4.5 分析与讨论

本文已从定性的角度对实验结果图像进行了分析,为体现本文方法的客观有效性,用定量的方法,对采集到的十幅菌落图进行处理,首先利用 ROC(receiver operating characteristic curve)指标进行评价,由于绘制 ROC 曲线时,各个方法的曲线容易发生交叉,影响效果的判别,所以本文通过计算 ROC 曲线下的面积 AUC(area under the curve)来体现本方法的有效性,由于 ROC 指标方法判定需要边缘基准图,而人工绘制基准图有太大的主观性而且效率不高,本文采用文献[50]中的客观方法来得到基准图,首先对实验结果中的三种对比方法与本文方法取多个阈值,从而得到 N 幅边缘检测图像,并统计图像相同位置为边缘点的个数,得到 N 幅备选的边缘基准图,最后根据 ROC 统计指标和诊断曲线(diagnosis line)选择效果最佳的图作为最终的边缘基准图。

表 4.1 为 Sobel 算法、Canny 算法、DOG 算法和本文算法的边缘图像 AUC 值。其中 Picture_1、Picture_2、Picture_3 分别为实验结果中的例图。从 AUC 值的检测结果中可以看出本文方法的效果明显好于 Canny 算法、Sobel 算法以及 DOG 算法,这与肉眼观测的结果保持一致。分析其原因, Sobel 算法利用像素点周围邻域的灰度加权,根据边缘点处达到极值这一方式进行检测,虽然误检边缘较少,但会存在边缘细节的丢失,因此不能得到较好的 AUC 值;而 Canny 算法通过高低阈值对检测结果进行调控,不能将有效的细节信息和背景纹理区分开,过检测导致虚假边缘过多,同样导致 AUC 值不高;而 DOG 方法和本文方法的特点都是利用到了视觉神经的一些特性,结果上看明显好于前两种方法,具体对比两种方法发现, Picture_1 和 Picture_2 中本文方法比较 DOG 方法也有较大提升,而 Picture_3 差别不大,可能根据菌落图的细胞分布有关, Picture_3 中细胞个数较少,而 Picture_1 和 Picture_2 中细胞分布较多,

黏连情况也比较普遍，当检测基数大时可以看出本文方法的效果比较明显，看出经过自适应加权响应模型的处理后，边缘细节信息得到增强，同时非经典感受野的颜色同质抑制效果也对背景异质成分进行了较好的去除，综合使得本文方法得到的边缘检测图像与边缘基准图 (Ground Truth) 匹配度最高。

表 4.1 不同检测算法下边缘图像的 AUC 值

Item	Sobel	Canny	DOG	本文方法
Picture_1	0.628	0.679	0.786	0.862
Picture_2	0.687	0.721	0.774	0.849
Picture_3	0.645	0.712	0.785	0.856
Picture_4	0.673	0.751	0.791	0.851
Picture_5	0.617	0.662	0.733	0.860
Picture_6	0.618	0.671	0.729	0.867
Picture_7	0.631	0.679	0.714	0.845
Picture_8	0.643	0.677	0.708	0.836
Picture_9	0.622	0.664	0.730	0.860
Picture_10	0.614	0.641	0.753	0.829
Picture_11	0.633	0.654	0.729	0.841

为了进一步体现本文方法的客观有效性，本文对十一幅菌落图边缘检测结果进行信息熵指标评价。

表 4.2 不同检测算法下边缘图像的信息熵值

Item	Sobel	Canny	DOG	本文方法
Picture_1	0.229	0.255	0.311	0.451
Picture_2	0.207	0.221	0.294	0.358
Picture_3	0.284	0.173	0.321	0.374
Picture_4	0.196	0.247	0.349	0.417
Picture_5	0.228	0.214	0.364	0.375
Picture_6	0.183	0.201	0.290	0.344
Picture_7	0.211	0.254	0.351	0.382
Picture_8	0.255	0.242	0.383	0.415
Picture_9	0.227	0.190	0.289	0.323
Picture_10	0.189	0.213	0.376	0.490
Picture_11	0.160	0.237	0.312	0.417

图像信息去除冗余后的平均信息称为图像的信息熵。信息熵越大表明其中包含的信息越丰富。从表 4.2 中可以看出与 Sobel 算法、Canny 算法以及 DOG 算法相比，本文方法所获得平均信息熵更高。可能是前两种方法缺少对视觉机制方面的考虑，只是从微分算子的角度考虑，所以两者的信息熵差别也不大；DOG 方法是简单的对经典感受野的一种模拟，属于视觉机制的基本运用，信息熵好于传统的 Canny 和 Sobel 算法；本文方法从非经典感受野等视觉

机制以及结合神经网络的角度充分模拟视觉系统,用其对图像做更深入的信息处理和传递,体现的边缘细节更加丰富,信息熵值相比 DOG 方法更高。综上所述,与其余对比方法相比,本文的信息熵值最大,说明本文的边缘检测图像保留的细节最丰富。

4.6 小结

本章依据人类视觉系统图像成像的原理引入方向选择、神经网络等视觉系统独有特点对图像信息的处理得到初步边缘图像,并通过非经典感受野的自适应加权响应模型以及颜色同质抑制模型处理,对图像边缘信息加强以及异质成分去除,得到最终的边缘结果,并通过与 Sobel 算法、Canny 算法以及 DOG 算法得到的边缘结果进行比较,突出本文方法在图像边缘检测中的优势。本文对菌落图进行边缘检测,检测效果比较其他方法得到的边缘更为清晰,并且边缘信息更加丰富,综合来看取得了不错的边缘检测效果。

第 5 章 基于视觉显著性信息的图像轮廓检测

轮廓信息是对图像进行空间冗余度消除后的常见特征，在后续图像的分析 and 理解起着关键的作用。而人类视觉系统在实现对周围环境或目标的感知时，对图像轮廓检测具有突出的敏感性和高效性。因此如何从生物视觉机制研究出发，研究视觉信息流的传递和处理过程，并建立以神经电生理为基础的视觉神经计算模型就显得非常有意义。目前仅有基于初级视皮层感受野的 Noninh 方法；也有研究利用空间稀疏特性抑制纹理，并基于颜色双拮抗机制来有效提取彩色图像轮廓信息的 SSC 方法[21]。另外还有研究基于 V1 的经典感受野和非经典感受野组合，在抑制纹理细节的基础上实现轮廓线段多方向感知的 ISO 方法^[55]。需要指出的是，虽然上述轮廓检测方法利用图像的多特征融合^[56,57]，考虑到了包括感受野在内的视觉机制作用，但对视觉信息流的神经脉冲时空编码过程进行了数学上的简化，从本质上仍然是一种基于邻域描述算子的滤波方法。本章主要结合视觉机制中非经典感受野和显著性信息等特性，并通过构建神经网络，实现轮廓的提取。本章选取在轮廓检测中常见的数据库(RUG)作为实验用图，计算本文方法检测结果以及 Noninh、SSC、ISO 等对比方法检测结果与基准轮廓图的 P 指标，结果表现本文方法能更准确保留真实轮廓，取得轮廓效果最佳。

5.1 概述

视觉神经实验表明，初级视皮层对视觉激励具有明显的方向性和位置选择性，虽然 V1 已经能够感知视觉目标中的线和点等初级梯度信息。但对于如何过滤隐含在梯度信息中的纹理边缘，并获得更高层次上的对于面和立体结构的理解，应该考虑到高级视皮层对视觉信息流的反馈调控作用。本文利用二维高斯导函数，提取视觉激励的多方向梯度信息，并降低多方向梯度信息的冗余度；构建 Leaky integrate-and-fire(LIF)神经网络模型，同时基于感受野机制设置网络突触的兴奋和抑制电流，提取显著性信息并反馈到 V1 以此提取图像轮廓响应，并继续对该轮廓响应进行非经典感受野纹理同质抑制处理，得到纹理抑制效果更佳的轮廓检测结果。

5.2 基本原理

本文提取的多方向梯度信息中存在较多的纹理成分，通过结合一种基于突触电流调制的经典和非经典感受野机制，将突触电流作用于构建的 LIF 神经网络模型中，采用神经元脉冲发放频率得到图像的显著性信息，并将其反馈到 V1 以对纹理进行有效抑制，再通过非经典感受野的纹理同质抑制作用，进一步去除背景纹理。

5.2.1 改进后的 LIF 神经网络模型

在视觉信息流传递与处理过程中，视皮层神经元通过脉冲放电的形式对视觉激励进行方向和位置等信息的编码，因此需构建具有高效计算性能的神经网络模型。考虑动作电位具

体波形对神经信息编码^[58]的冗余性，本文选择如式(5.1)所示的 LIF 神经元作为神经网络模型的基本单元，由于神经元脉冲发放存在不应期特性，因此引入时间参数 ref 来调整模型的动力学行为。改进后的 LIF 模型如式(4.1)所示：

$$\begin{cases} c_m \frac{dv}{dt} = -g_1 v + I_{syn} (ref = 0) \\ v = v_{reset}, ref = ref - 1 (ref \neq 0) \\ v = v_G, ref = const (v > v_{th}) \end{cases} \quad (5.1)$$

式中， v 、 v_{th} 、 v_{reset} 、 v_G 分别表示为神经元膜电位、脉冲发放阈值电位、静息电位以及脉冲峰值电位； c_m 和 g 分别为膜电容和等效漏电导； I_{syn} 对应于突触输入电流，是兴奋和抑制两部分电流有符号加和的结果。本文按如下方式添加绝对不应期的作用：当膜电位 v 达到脉冲发放阈值电位时， v 将瞬时被置为脉冲峰值电位，同时神经元将进入绝对不应期，不再响应任何激励，直到时间参数 ref 被重新置为 0。

5.2.2 初级视皮层多方向梯度响应

V1 神经元在感知来自膝状体的视觉信息时，关注于信息的相对梯度特征，表现出对方向 θ_i 和位置 (x, y) 的选择特性。设输入图像为 $I(x, y)$ ，将图像与特定尺度因子的高斯导函数进行卷积运算，结果取绝对值即为 V1 多方向梯度响应，如式(5.2)所示：

$$e_i(x, y; \theta_i) = \left| I(x, y) * \frac{\partial g(\tilde{x}, \tilde{y}; \theta_i)}{\partial \tilde{x}} \right| \quad (5.2)$$

式中， $\tilde{x} = x \cos(\theta_i) + y \sin(\theta_i)$ ， $\tilde{y} = -x \sin(\theta_i) + y \cos(\theta_i)$ ，函数 g 为高斯导函数。

考虑到方向信息的遍历性，当计算 n_θ 个方向梯度响应时，有：

$$\theta_i = \frac{(i-1)\pi}{n_\theta}, i = 1, 2, \dots, n_\theta \quad (5.3)$$

对于特定的位置 (x, y) 而言，本文选取方向梯度响应的极大值，并按响应比重进行加权处理，将结果 $E(x, y)$ 作为多方向梯度响应的输出，如式(5.4)和式(5.5)所示：

$$e_{\max} = \max \{e_i(x, y; \theta_i)\}, i = 1, 2, \dots, n_\theta \quad (5.4)$$

$$E(x, y) = e_{\max} \times \frac{e_{\max}}{\sum_i e_i(x, y; \theta_i)} \quad (5.5)$$

5.2.3 信息流的显著性提取

在视觉信息流的传递和处理过程中，存在自下向上和自上向下两种通路，前者利用点和线条的颜色或纹理等特征，通过冗余度去除来获取视野中的显著性信息^[59]，而后者是将来自高级视皮层的信息流逆向反馈至 V1，对多方向梯度响应进行调控。

本部分首先提取视觉激励的多方向梯度信息，经过移动局部窗操作，获得神经元突触的

兴奋电流；然后利用非经典感受野机制获得突触抑制电流，以 LIF 神经元的脉冲发放频率作为显著性输出，最后反馈并与初级视皮层的多方向梯度信息进行融合，获得轮廓检测结果。

视觉神经实验证实，对于特定的视觉激励而言，视觉神经系统只有部分神经元被激发，这将有利于提高神经编码效率。因此本文首先进行高斯导函数卷积来获得图像的多方向梯度信息，并更多关注于图像中目标的整体轮廓信息，减弱与背景纹理相对应的冗余信息描述。以图像某像素点 (x, y) 为例，构建以其为中心的局部方形窗 W ，按式(5.6)求窗内方向选择的均值作为突触兴奋电流，式中 $\text{mean}(\cdot)$ 为均值函数：

$$I_{syn}(x, y) = \text{mean}_{(\alpha, \beta) \in W}(E(\alpha, \beta)) \quad (5.6)$$

本文将充分利用经典和非经典感受野对于视觉激励的不同响应，利用体现两者的差异性的去抑制区域来设置突触抑制电流。利用第三章中的显著性信息提取模型以及式(3.11)和(3.12)获得突触抑制电流 I_{inh} 。

最终突触电流是如式(5.7)所示的突触兴奋电流和抑制电流的加权和，其中 α 为突触抑制系数：

$$I_{syn}(x, y) = I_{syn}(x, y) - \alpha \cdot I_{inh}(x, y) \quad (5.7)$$

当根据感受野机制获得突触电流后，将其作为式(5.1)所描述的神经元输入，以神经元的脉冲发放频率作为图像的显著性提取结果，表示为 $\mu_r(x, y)$ 。

5.2.4 信息流的显著性反馈

正如前述，在视觉通路中存在着自上向下的途径，同时高级视皮层基于对信息流的理解，将处理结果反馈至低级皮层，对后者的响应进行调控。因此本文将获得的显著性结果反馈到初级视皮层，与最终的多方向梯度信息进行融合，从而得到轮廓响应结果如式(5.8)所示：

$$SE(x, y) = \mu_r(x, y) \times E(x, y) \quad (5.8)$$

5.2.5 非经典感受野纹理同质抑制

对信息流的显著性反馈后的轮廓响应结果还存在一部分背景纹理，所以最后通过纹理同质抑制的方式去除纹理。本方法利用片相似性来描述纹理同质性，对方向选择得到的图像 $E(x, y)$ 进行处理。首先令 $P(x, y)$ 为图像 $E(x, y)$ 上以像素点 (x, y) 为中心的 3×3 的领域中所有像素灰度值的有序向量，将片 $P(x, y)$ 排列成 9 维的向量 $P(x, y) = (E(x-3, y-3), \dots, E(x+3, y+3))^T$ ，即图像片都是通过将图像 3×3 局部内的像素点按照行或列排列后得到的一个向量。片 $P(x, y)$ 代表像素点 (x, y) 的纹理特性，片 $P(x', y')$ 代表像素点 (x', y') 的纹理特性，经过纹理同质抑制模型以及式(3.6)~(3.8)、(3.10)得到纹理抑制响应图像。纹理同质抑制模型是根据图像上任意两点的纹理相似度来对其同质成分进行判断，同质性越高得到响应 $WE(x, y)$ 越大，同质性越低则纹理同质抑制响应 $WE(x, y)$ 越小， $WE(x, y)$ 代表纹理成分，通过式(5.9)将其去除，从而凸显主体轮廓信息：

$$FE(x, y) = SE(x, y) - WE(x, y) \quad (5.9)$$

5.3 轮廓检测步骤

轮廓检测步骤如图 5.1 所示。

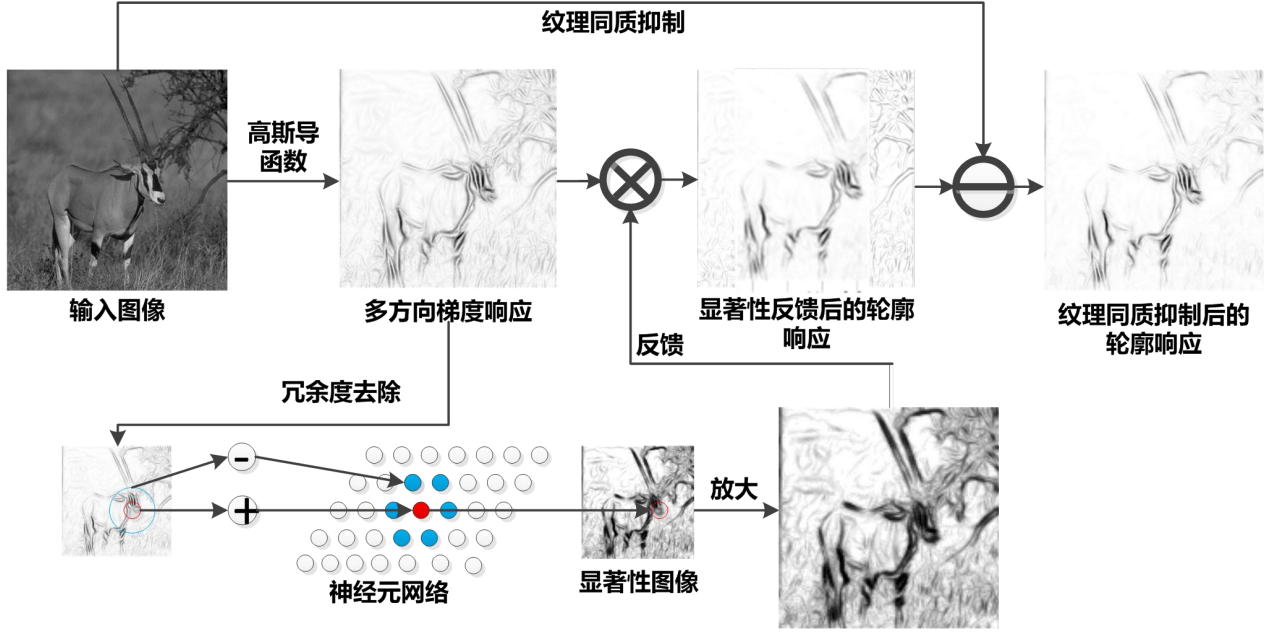


图 5.1 轮廓检测步骤图

步骤 1：首先将原始图像 $I(x, y)$ 与特定尺度因子的高斯导函数进行卷积运算，结果取绝对值得到 V1 多方向梯度响应 $e_i(x, y; \theta_i)$ ，并选取方向梯度响应的极大值，并按响应比重进行加权处理得到最终多方向梯度响应的输出 $E(x, y)$ 。

步骤 2：对多方向梯度响应图像 $E(x, y)$ 进行冗余度去除，通过式(5.6)得到突触兴奋电流，根据经典和非经典感受野机制得到的突触抑制电流，通过式(5.7)得到最终突触电流参数 I_{syn} ，将突触电流参数 I_{syn} 代入 LIF 神经网络，根据神经元脉冲发放频率得到显著性图像 $\mu_r(x, y)$ 。

步骤 3：将显著性信息 $\mu_r(x, y)$ 经过双线性差值放大并反馈到初级视皮层，与最终的多方向梯度信息 $E(x, y)$ 通过式(5.8)进行融合得到轮廓响应结果 $SE(x, y)$ 。

步骤 4：对轮廓响应结果 $SE(x, y)$ 进行纹理同质抑制，经过纹理同质抑制模型和以及式(3.6)~(3.8)、(3.10)，得到纹理图像 $WE(x, y)$ ，通过式(5.9)得到最终的轮廓响应结果 $FE(x, y)$ 。

步骤 5：对轮廓响应图像 $FE(x, y)$ 使用非极大值抑制处理方法得到单像素轮廓，并用滞后门限方法进行二值化，得到最终的轮廓图。

5.4 实验结果

本文选取在轮廓检测中常见的数据库(RUG)作为实验用图，RUG 中每一幅图都提供了由人工所绘制的基准轮廓图。本文以数据库中的部分图像为例来说明方法的有效性，如图 5.2 所示。同时本文选取了前述 Noninh、SSC 以及 ISO 等主流的轮廓检测方法进行比较。其中实验参数 $\sigma = 2$ ， $n_\theta = 8$ ，突触抑制系数 $\alpha = [0.6:0.1:1.5]$ 。

轮廓检测结果如图 5.3 所示。其中(a)~(c)分别为 Noninh、SSC 以及 ISO 方法的轮廓检测结果，(d)则为本文方法结果。

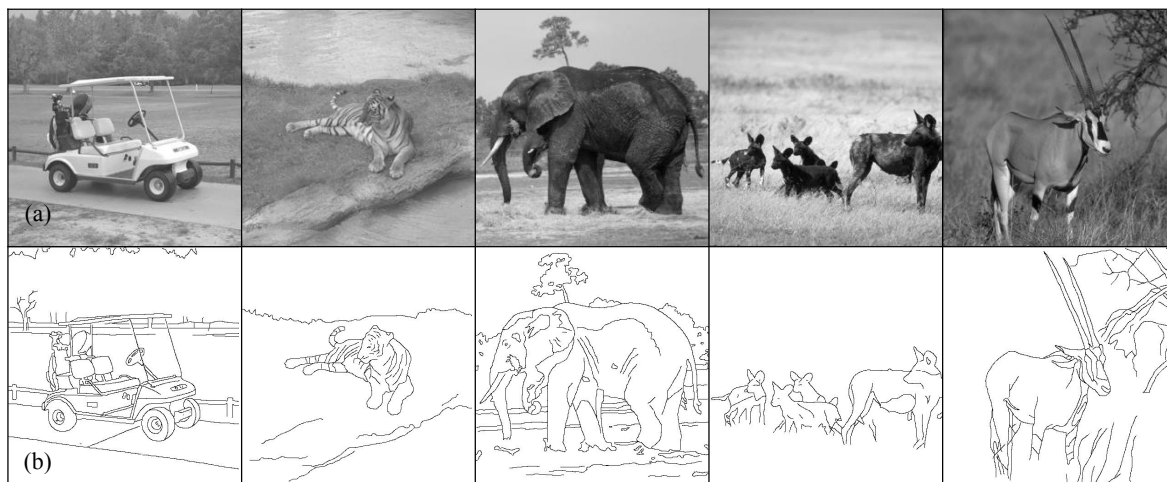


图 5.2 实验用图及其基准轮廓图

(a) 待处理图像，分别为 golfcart, tiger, elepelephant, hyena, gazelle; (b) 对应的基准轮廓图

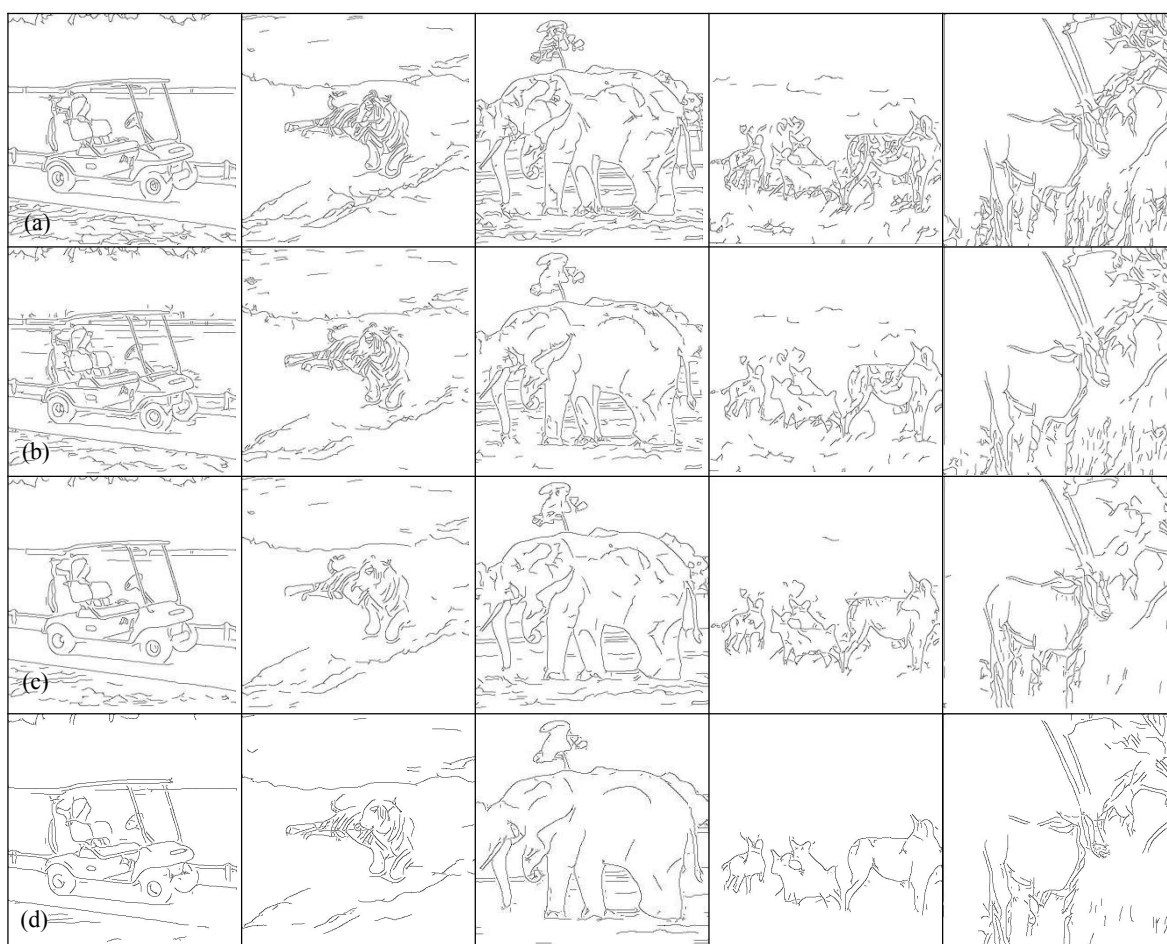


图 5.3 轮廓检测结果

(a) Noninh 方法; (b) SSC 方法; (c) ISO 方法; (d) 本文方法

本文利用轮廓检测中常见的评价指标^[56]进行定量比较。考虑到基准轮廓图由人工绘制，虽然是多人结果的平均，但仍具有一定的主观性，所以在定量评价时，在基准轮廓的 5×5 邻

域内的波动即可视为正检。轮廓检测的正确集 E 、错误集 E_{FP} 、漏检集 E_{FN} 、错检指标 e_{FP} 、漏检指标 e_{FN} 和综合指标 P 分别如式(5.4)~(5.9)所示:

$$E = E_D \cap (E_{GT} \oplus T) \quad (5.4)$$

$$E_{FP} = E_D - E \quad (5.5)$$

$$E_{FN} = E_{GT} - (E_{GT} \cap (E_D \oplus T)) \quad (5.6)$$

$$e_{FP} = \text{card}(E_{FP}) / \text{card}(E) \quad (5.7)$$

$$e_{FN} = \text{card}(E_{FN}) / \text{card}(E_{GT}) \quad (5.8)$$

$$P = \frac{\text{card}(E)}{\text{card}(E) + \text{card}(E_{FP}) + \text{card}(E_{FN})} \quad (5.9)$$

式中, E_D 和 E_{GT} 分别代表轮廓检测结果的像素集和基准轮廓像素集, \oplus 表示对像素进行膨胀操作, T 表示边长为 5 的方形邻域单元, 各方法检测的轮廓像素点在基准轮廓像素集的 5×5 范围内即视为正检。轮廓检测结果的定量数据如表 5.1 所示。

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	e_{FP}	e_{FN}	P
Golfcart	Noninh		0.47	0.20	0.53
	SSC		0.40	0.22	0.54
	ISO	1.20	0.29	0.27	0.58
	本文方法	0.70	0.23	0.24	0.64
Tiger	Noninh		0.91	0.15	0.46
	SSC		0.86	0.17	0.48
	ISO	0.80	0.46	0.23	0.56
	本文方法	0.60	0.47	0.19	0.60
Elephant	Noninh		0.56	0.15	0.56
	SSC		0.36	0.26	0.57
	ISO	1.10	0.25	0.24	0.60
	本文方法	0.60	0.26	0.21	0.62
Hyena	Noninh		0.86	0.20	0.43
	SSC		0.74	0.27	0.45
	ISO	0.80	0.33	0.29	0.50
	本文方法	0.60	0.31	0.26	0.54
Gazelle	Noninh		0.76	0.21	0.41
	SSC		0.65	0.24	0.45
	ISO	0.90	0.32	0.33	0.49
	本文方法	0.70	0.28	0.30	0.52

5.5 分析与讨论

从图 5.3 中可以看出, 基于初级视皮层感受野特性的 Noninh 方法, 但背景纹理和显著性信息无法区分开, 因此从结果看轮廓细节与背景纹理之间的区分度较弱, 错检率较高; 基于颜色双拮抗机制的 SSC 方法, 较好地利用了颜色成分在轮廓检测中的作用, 但从结果中可以看出在轮廓细节附近的纹理干扰较大; 而融合经典和非经典感受野的 ISO 方法, 考虑了局部细节与全局特征的关联性, 与前两者相比, 轮廓错检率有所下降。而本文提出的新方法, 由于体现了低级和高级视皮层之间的交互作用, 融合图像多方向梯度信息与显著性信息, 并结合纹理同质抑制模型, 因此从视觉上看整体轮廓较为清晰, 相对于前述方法具有明显的优势。

从表 5.1 的性能指标来看, Noninh 方法的漏检指标 e_{FN} 较低, 但错检指标 e_{FP} 显著高于其他方法, 表明该方法的轮廓检测结果混入了较多的纹理干扰, 因此综合指标 P 均为最低。观察 SSC 方法的实验结果, 虽然 e_{FP} 较前者有所下降, 但对应的 e_{FN} 却相应增大, 所以从综合指标 P 来评价, 性能并没有明显改善。而对于 ISO 和本文方法, 在保持漏检率基本不变的情况下, 在错检指标 e_{FP} 上有了明显的改善, 因此与 Noninh 和 SSC 方法相比较, 整体性能上有了较大提高。需要进一步指出的是, Noninh 方法和 SSC 方法的漏检指标均略低于本文方法, 但它们的错检指标远高于本文方法。以上两种方法均存在过检测的情况, 而导致纹理信息干扰大, 虽使得漏检指标较低, 但仍无法弥补较高错检率所导致的综合性能不佳。而 ISO 方法与本文方法的错检和漏检指标较为接近, 其中图 Tiger 和图 Elephant 所对应的错检指标稍低于本文方法, 但漏检指标偏大, 得到的综合评价指标低于本文方法; 而图 Golfcart、Hyena、Gazelle 对应的错检和漏检指标均高于本文方法结果。综合以上分析可知, 本文方法在错检和漏检指标上达到一种均衡, 从综合评价指标 P 看出, 相比其他方法具有更理想的轮廓检测性能。

综上所述, 由于 Noninh 没有考虑到纹理抑制, 因此从结果来看属于一种轮廓过检测的状态; SSC 方法考虑了图像的局部特征, 利用了视锥细胞的颜色双拮抗机制, 但由于这只是视觉通路上的一个前期环节, 因此从结果来看, 尤其是在轮廓细节点附近, 其对于背景纹理的干扰处理并不理想; 而 ISO 方法利用经典感受野和非经典感受野对于视觉激励的调制作用, 在一定程度上体现了局部细节和全局轮廓之间的关联性, 但仍然是一种单向的视觉信息流处理方式。而本文利用初级视皮层的多方向梯度响应, 同时降低细节冗余度, 关注于图像的整体特征, 然后利用生物视觉系统的感受野特性提取显著性信息, 并利用自上而下的视通路途径, 将显著性信息与低级视皮层所获取的多方向梯度信息进行融合, 获得较好的轮廓响应基础上再通过非经典感受野纹理同质抑制的作用, 进一步去除背景纹理成分, 最终得到更佳的轮廓响应结果。因此从理论分析和实验结果来看, 这种视觉显著性信息流传递和处理方式并结合纹理同质抑制作用, 既有效降低轮廓点的漏检, 同时又充分抑制纹理对检测的干扰。

5.6 小结

本文考虑视通路中视觉信息流的双向传递和处理问题，首先模拟低级视皮层提取视觉激励的多方向梯度信息，并通过降低冗余度的方法来弱化纹理特征，关注于具有整体性和连续性的轮廓细节；然后基于经典感受野和非经典感受野特性提取图像的显著性信息，并通过模拟高级视皮层对视觉信息流的反馈调控作用，将来自高级视皮层的显著性信息和来自低级视皮层的方向梯度信息进行融合，获得较好的轮廓响应；最后通过纹理同质抑制方式得到背景纹理响应并去除，获得轮廓检测最终结果，实验结果验证了方法的有效性。本文新方法将为基于生物视觉机制的图像分析或理解应用研究提供一种新的途径和思路。

第 6 章 总结与展望

6.1 总结

目前运用神经计算在图像处理中的方式主要有两个方面，一方面是对神经元模型以及神经网络的研究，通过数学模型表现其神经元信息传递的方式；另一方面集中在各种视觉机制的具体运用以及创新，例如对视觉感受野的研究中，从早期的经典感受野特性，以同心圆的结构表现信息之间的传递过程，再到后来研究者发现非经典感受野，更进一步完善了对视觉感知的认识。本文方法是对上述两方面的结合，不仅注重于神经元编码的研究和神经网络的选择，而且充分利用多种视觉机制所发挥的特性，获得了较好的检测结果，也为后续基于生物视觉机制和生理特性的图像处理和理解提供了新的思路。

目前已有部分研究利用生物视觉特性对图像进行边缘和轮廓的检测，比传统的检测方法更符合生物生理特性，取得的效果也好于传统方法，但实际上它们对于视觉机制仍然属于一种黑箱模拟，另外还有研究试图利用某一种视觉机制进行研究，这对于复杂的生物视觉系统来说缺乏一定的说服力。本文方法考虑了视觉信息在视觉通路上的传递过程，利用多种视觉机制，并结合神经编码，构建非经典感受野特性的模型结构，开展了图像边缘和轮廓检测的应用研究。

本文首先对彩色菌落图进行边缘检测。构建神经网络模型，利用神经编码、方向选择等视觉机制对图像进行边缘初步提取，建立符合非经典感受野特性的自适应加权响应模型和颜色同质抑制模型，对图像进行边缘信息增强以及异质成分去除，获得优于传统方法的边缘检测结果。

另外本文还对 RuG 图像库中的自然图像进行轮廓检测。模拟低级视皮层提取视觉激励的多方向梯度信息，并注重于降低冗余度的方法来弱化纹理特征。通过结合一种基于突触电流调制的经典和非经典感受野机制，获取图像的显著性信息，将其反馈并调控低级视皮层，得到轮廓响应，并运用纹理同质抑制模型去除轮廓响应中的背景纹理，最终取得了优于另外三种具备视觉机制方法的检测结果。

论文具体研究工作总结如下：

(1) 本文首先建立了三种基于非经典感受野特性的视觉处理模型。首先基于非经典感受野三高斯模型，提出一种自适应加权响应模型，利用像素的局部对比度特性实现对非经典感受野兴奋、抑制以及去抑制区域参数的自适应动态调整，通过带权值感受野窗口对图像进行遍历，实现对边缘信息的增强作用，使结果具有更完整的边缘特性。其次改善传统的经典感受野同质抑制模型，使用三高斯函数代替 DOG 模型来模拟非经典感受野，采用 Lab 颜色空间描述图像像素点的颜色相似度，根据片相似性描述图像像素点之间的纹理相似度，经过颜色

和纹理同质间相互抑制的方式，将纹理及异质成分显示出来并去除。通过实验对比发现经过同质抑制后的图像，背景纹理以及异质成分减少。最后提出显著性信息提取模型，非经典感受野的去抑制作用将有助于局部空间特征的提取，所以根据经典和非经典感受野存在的去抑制区域的差异性产生突触抑制电流，并通过幂指数归一化的方式进行显著性提取。此三种模型的研究分析为之后对于边缘以及轮廓的检测奠定了基础。

(2)本文通过对非经典感受野机制的研究以及结合方向选择特性、神经编码方式、神经网络模型，提出一种基于非经典感受野机制的边缘检测新方法。首先使用具有视觉感知特性的 Log-Gabor 滤波器对图像进行方向选择，采用 izhikevich 神经元单元构建神经网络，利用次序编码方式对视觉信息流进行定量的分析。考虑到 Log-Gabor 滤波后的方向选择图像边缘信息的缺失，根据非经典感受野的三高斯模型特性，提出自适应加权响应模型对图像边缘信息进行增强处理，同时考虑到图像存在较多的纹理，对彩色图像进行 Lab 颜色空间信息的提取，通过两点颜色的差异性进行非经典感受野颜色同质抑制处理。以含有丰富细节的菌落图像边缘检测为例，结果表明，获得的边缘信息更完整，背景纹理更少。经过 ROC 和信息熵方法评价，本文方法要优于传统检测方法。

(3)基于非经典感受野机制的显著性信息提取模型以及纹理同质抑制模型的具体运用，提出了一种基于视觉显著性信息的图像轮廓检测方法。首先采用二维高斯导函数模拟低级视皮层提取视觉激励的多方向梯度信息，并通过降低冗余度的方法来弱化纹理特征，关注于具有整体性和连续性的轮廓细节；利用局部移动窗获取突触兴奋电流，并结合一种基于突触电流调制的经典和非经典感受野机制，采用高级视皮层的神经元脉冲发放频率作为图像的显著性提取结果，并通过模拟高级视皮层对视觉信息流的反馈调控作用，将来自高级视皮层的显著性信息和来自低级视皮层的方向梯度信息进行融合，获得初步轮廓响应；最后将该轮廓响应通过纹理同质抑制模型处理，去除背景纹理，从而获得最终的轮廓检测结果。对 RuG 图像库中的图像进行轮廓检测，与 ISO 等方法相比，本文方法的轮廓与基准轮廓图比较获得的综合评价指标 P 值最高，性能上具有一定的优势。

6.2 展望

本文研究建立了基于非经典感受野的三种模型，并结合多种视觉机制以及神经元模型模拟视觉信息的传递与处理过程，在边缘检测以及轮廓检测中去取得了不错的效果，证明了模拟模型的有效性。本文在图像边缘检测中运用了 Log-Gabor 滤波器处理，虽然其良好地模拟生物视皮层的方向选择特性，但也存在边缘信息丢失的问题，对此方面还需更多的研究。另外后续可以对能够反映非经典感受野机制的模型进行扩展或者融合。而且对于视觉神经机制，无论是在神经网络的结构和规模、突触动态可塑性、还是视皮层间信息流的前馈与反馈传递，都可进一步深入研究。同时针对图像处理应用的实际需求，构建具有高效性的模型实现框架，也是需要重视的方向。

致谢

读研的这两年半时间马上就要到了，回顾过去，对我来说是一段美好而短暂的时光，学到了很多的东西，不仅是知识方面，更多的是对人对事的处理方式以及价值观的提升。研一在听课学习专业知识中度过；研二则主要是科研方面，对论文的完善；研三更多则是准备工作应聘的事。其中有很多艰辛，但看到成果的那一刻，让我觉得这段时光会是我将来美好的回忆。

在这两年半的求学以及生活中，我最需要感谢的人就是我的导师范影乐教授。范老师是一个很随和的人，初次见面是考研复试前与他在办公室的交流，整个过程让我感觉非常的亲切。学习方面，范老师怕我们大四毕业后无聊，也为了让我们更好适应读研的生活，研一之前的暑假提早将我们几个同门聚在一起，和师姐一同参加智慧城市比赛，每个礼拜抽出自己的时间跟我们一起讨论，并指导我们努力的方向。在论文的写作中，范老师也给我指明了研究方向，使我不再迷惘，他的一句话令我印象非常深刻，就是“有问题随时来找我”。当然这指的不只是学习上的问题，生活中他也像长辈一样对我们无微不至。电脑坏了，椅子坏了，他都会第一时间给我们解决；有一次讨论的迟了，过了饭点，范老师就开着车带着大家去外面好好吃了一顿；我们喜欢收集纪念币，他也牢牢记着，每次有新的纪念币他就会第一时间告知我们。由于篇幅限制，还有很多的事我就不一一例举了，总之能在范老师门下学习我真的非常庆幸，每当和别的同学谈论自己导师的时候我都是非常自豪。

其次我要感谢同门谢尚周、李康群和李嘉祥，小谢是我的室友，寝室弄得干干净净，和他住一起非常舒心；小李在论文内容方面对我的帮助很大，有不懂的地方我就请教他；每次论文格式出问题了，嘉祥总是第一个过来帮我解决，和你们在一起学习生活让我变得更加开朗充实。感谢武薇师姐在大家一起郊游的时候带了一大堆的零食给我们，让我们能够畅玩；感谢廖进文、罗佳骏和李丹菁师兄以及陆晓娟和王梦蕾师姐，他们是我学习的榜样；感谢方芳师姐和王典、王迪师兄，对我的论文以及学习方面提供了巨大的帮助；同时也感谢师妹张梦楠，师弟王强、胡钧皓、郭金朋、张明琦、杜宇华、蒋涯和周涛，是你们使得实验室充满欢乐。

最后我要感谢我的家人，谢谢全家人对我考研读研的支持，父亲是我将来工作的榜样，用自己的行动教会了我很多待人处事的道理；母亲是对我生活上无微不至的照顾，让我更加有精力投入进学习中；每次回家爷爷奶奶总会对我嘘寒问暖，给我精神上的激励；他们永远都是我坚实的后盾，让我在今后的道路上能够砥砺前行。

下沙.杭州电子科技大学
二零一六年十二月十日

参考文献

- [1] 李雄,刘允. 视觉机制研究对机器视觉的启发示例[J]. 中国图象图形学报, 2013, 18(2): 152-156.
- [2] 段瑞玲,李庆祥,李玉和,等. 图像边缘检测方法研究综述[J]. 光学技术, 2005, 31(3): 415-419.
- [3] 张衡,谭晓阳,金鑫,等. 基于多视图聚类的自然图像边缘检测[J]. 模式识别与人工智能, 2016,1(2): 163-170.
- [4] 袁春兰,熊宗龙,周雪花,等. 基于 Sobel 算子的图像边缘检测研究[J]. 激光与红外, 2009, 39(1): 85-87.
- [5] Karen A, Sos S, Shahan C, et al. Shape-dependent canny edge detector[J]. Optical engineering, 2011, 50(8): 087008-1-087008-12.
- [6] 张玲艳. 基于 Canny 理论的自适应边缘检测方法研究[D]. 西北大学, 2009.
- [7] 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.
- [8] 曲智国,高颖慧,王平,等. 基于空频域联合阈值分割的轮廓检测方法[J]. 计算机科学, 2012, 39(10): 286-289.
- [9] Wang S, Xia Y, Liu Y, et al. Gabor feature based nonlocal means filter for textured image denoising[J]. Journal of visual communication & image representation, 2012, 23(7): 1008-1018.
- [10] Grigorescu C, Petkov N, Westenberg M, et al. Contour detection based on nonclassical receptive field inhibition[J]. Image processing, IEEE transactions on, 2003, 12(7): 729-739.
- [11] 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.
- [12] Chen K, Song X, Li C, et al. 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.
- [13] 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.
- [14] Shou T, Wang W, Yu H, et al. Orientation biased extended surround of the receptive field of cat retinal ganglion cells[J]. Neuroscience, 2000, 98(2): 207-212.
- [15] 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.
- [16] 杜晓凤,李翠华,李晶,等. 基于复合感受野的轮廓检测算法[J]. 电子与信息学报, 2009, 31(7): 1630-1634.
- [17] 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.
- [18] Zeng C, Li Y, Li C, et al. Center-surround interaction with adaptive inhibition: a computational model for contour detection[J]. NeuroImage, 2011, 55(1): 49-66.
- [19] 吴贺,胡玉兰. 非经典感受野的朝向和对比度特性在轮廓提取中的应用[J]. 沈阳理工大学学报, 2013, 32(6): 12-16.
- [20] Yang K, Li C, Li Y, et al. Multifeature-based surround inhibition improves contour detection in natural images[J]. Image processing, IEEE transactions on, 2014, 23(12): 5020-5032.
- [21] 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.
- [22] Zaghloul K A, Boachen K. Optical nerve signals in a neuromorphic chip : outer and inner retina model[J]. IEEE trans on bio-medical engineering, 2004, 51(4) : 657-666.
- [23] 李林雨. 基于空间频率平均能量谱的整体视觉认知理论研究[D]. 天津大学, 2011.
- [24] 陈勇,李愿,吕霞付,等. 视觉感知的彩色图像质量积极评价[J]. 光学精密工程, 2013, 21(3): 742-750.
- [25] Kuang X, Poletti M, Victor J D, et al. Temporal encoding of spatial information during active visual fixation[J]. Current biology, 2012, 22(6): 510-514.
- [26] 桑农,唐奇伶,张天序,等. 基于初级视皮层抑制的轮廓检测方法[J]. 红外与毫米波学报, 2007, 26(1): 47-51,60.
- [27] Raudies F, Neumann H. A model of neural mechanisms in monocular transparent motion perception[J]. Journal of physiology-paris, 2010, 104(1): 71-83.
- [28] Webb B S, Ledgeway T, McGraw P V, et al. Relating spatial and temporal orientation pooling

- to population decoding solutions in human vision[J]. Vision research, 2010, 50(22): 2274-2283.
- [29] Masud M S, Borisyuk R. Statistical technique for analysing functional connectivity of multiple spike trains[J]. Journal of neuroscience methods, 2011, 196(1): 201-219.
- [30] Sacerdote L, Tamborrino M, Zucca C, et al. Detecting dependencies between spike trains of pairs of neurons through copulas[J]. Brain research, 2012, 1434: 243-256.
- [31] 金小贤, 李卫军, 陈旭, 等. 一种基于视觉特性的仿生图像增强算法[J]. 计算机辅助设计与图形学学报, 2010, 22(3): 534-537, 544.
- [32] Lu D, Yu X, Jin X, et al. Neural network based edge detection for automated medical diagnosis[C]. // Information and automation (ICIA), 2011 IEEE international conference on. IEEE, 2011: 343-348.
- [33] Waldemark K, Lindblad T, Bečanović V, et al. Patterns from the sky: satellite image analysis using pulse coupled neural networks for pre-processing, segmentation and edge detection[J]. Pattern recognition letters, 2000, 21(3): 227-237.
- [34] Kerr D, Coleman S, McGinnity M, et al. Biologically inspired edge detection[C]. // Intelligent systems design and applications (ISDA), 2011 IEEE international conference on. IEEE, 2011, 802-807.
- [35] 崔力, 浩明. 基于视觉注意机制的图像质量评价[J]. 东南大学学报, 2012, 42(5): 854-858.
- [36] 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]. 杭州电子科技大学, 2015.
- [37] 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.
- [38] Han C, Wang J, Deng B. Fire patterns of modified HH neuron under external sinusoidal ELF stimulus[J]. Chaos, solitons & fractals, 2009, 41: 2045-2054.
- [39] Chouhan A S. An Analytical Study of leaky integrate-and-fire neuron model using matlab simulation[J]. International journal of engineering research and technology, 2013, 2(4): 2242-2245.
- [40] 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.
- [41] Tony L. A computational theory of visual receptive fields[J]. Biological cybernetics, 2013,

107(6): 589-635.

- [42]Sadagopan S, Ferster D. Feedforward origins of response variability underlying contrast invariant orientation tuning in cat visual cortex[J]. *Neuron*, 2012, 74(5): 911-923.
- [43]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]. 天津医科大学, 2012.
- [44]Tanabe S. Population codes in the visual cortex[J]. *Neuroscience research*, 2013, 76(3): 101-105.
- [45]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.
- [46]崔文博, 蔺想红, 徐满意, 等. 脉冲神经网络图像分割的编码方法[J]. *计算机工程*, 2012, 38(24): 196-199.
- [47]Thorpe S J. Spike-based image processing: can we reproduce biological vision in hardware[C]. *Computer vision-ECCV 2012. Workshops and demonstrations*. Springer berlin heidelberg, 2012: 516-521.
- [48]Izhikevich EM. Simple model of spiking neurons[J]. *Neural networks, IEEE transactions on*, 2003, 14(6): 1569-1572.
- [49]Bowyer K, Kranenburg C, Dougherty S, et al. Edge detector evaluation using empirical ROC curves[J]. *Computer vision and image understanding*, 2001, 84(1): 77-103.
- [50]Yitzhaky Y, Peli E. A method for objective edge detection evaluation and detector parameter selection[J]. *Pattern analysis and machine intelligence, IEEE transactions on*, 2003, 25(8): 1027-1033.
- [51]Fernández-García N L, Carmona-Poyato A, Medina-Carnicer R, et al. Automatic generation of consensus ground truth for the comparison of edge detection techniques[J]. *Image and vision computing*, 2008, 26(4): 496-511.
- [52]Fawcett T. An introduction to ROC analysis[J]. *Pattern recognition letters*, 2006, 27(8): 861-874.
- [53]Brink A D. Using spatial information as an aid to maximum entropy image threshold selection[J]. *Pattern recognition letters*, 1996, 17(1): 29-36.
- [54]Verma O P, Hanmandlu M, Kumar P, et al. A novel bacterial foraging technique for edge detection[J]. *Pattern recognition letters*, 2011, 32(8): 1187-1196.

- [55] Grigorescu C, Petkov N, Westenberg M, et al. Contour detection based on nonclassical receptive field inhibition[J]. Image processing, IEEE transactions on, 2003, 12(7): 729-739.
- [56] Yang K, Li C, Li Y, et al. Multifeature-based surround inhibition improves contour detection in natural images[J]. Image processing, IEEE transactions on, 2014, 23(12): 5020-5032.
- [57] Yang K, Li C, Li Y, et al. 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.
- [58] Zhu M, Rozell C J. Visual nonclassical receptive field effects emerge from sparse coding in a dynamical system[J]. Computational biology, 2013, 9(8): 1-15.
- [59] 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.

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. 基于视觉显著性信息的图像轮廓检测. 计算机应用与软件(核心期刊), 已录用, 待发表, 1/5。

参与的科研项目:

1. 基于双稳态随机共振非周期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 影像质量分级的原发性闭角型青光眼辅助自动诊断研究(61501154), 国家自然科学基金。
3. 自动化检测中的机器视觉开发模块研发(063113017), 横向课题。