

# 杭州电子科技大学

## 硕士学位论文

题 目： 基于视觉感知机制的图像轮廓  
检测技术研究

研 究 生 李 康 群

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

基于视觉感知机制的图像轮廓  
检测技术研究

研 究 生：李 康 群

指导教师：范 影 乐 教 授

2016 年 12 月

Dissertation Submitted to Hangzhou Dianzi University  
for the Degree of Master

# **Research on Image Contour Detection Technique Based on Visual Perception Mechanism**

**Candidate: Li Kangqun**

**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

轮廓信息是图像降维后的关键信息，它对于后续准确高效的目标识别和图像理解等应用具有重要意义。传统轮廓检测方法对轮廓的方向性和连续性进行数学建模，在剔除背景纹理边缘、区分主体轮廓和非轮廓边缘方面具有一定的有效性。但必须指出的是，传统检测方法并没有引入或参考视觉感知机制在轮廓提取中的信息流传递处理过程，更多的只是一种抽象化的数学描述。因此本文关注于视觉感知系统在轮廓检测或提取中的机理，重点研究神经网络的脉冲编码以及视觉通路不同层次对视觉信息的响应特性，并尝试图像轮廓提取等处理的具体应用。首先构建视通路多感受野朝向性关联模型，突出方向选择特性的形成机制以及视通路各个层次在轮廓提取过程中的作用；其次给出了一种颜色拮抗与群体编码效应的实现手段，并应用于彩色图像的轮廓检测；最后引入双层神经网络的脉冲调控编码，加强高级视皮层对视觉信息的调控甄选作用，能够在轮廓检测中有效剔除背景纹理。本文主要工作及研究成果如下：

(1)提出了一种基于视通路多感受野朝向性关联的轮廓检测方法。根据视通路各层次的感受野特性、层次之间感受野的变换特性，模拟实现视皮层的方向选择特性。利用视网膜神经元网络的前级编码作用、LGN 细胞的非经典感受野调节机制以及视皮层方向选择特性对图像轮廓进行检测。以 RuG 图库为实验对象，实验结果的平均  $P$  指标为 0.43，表明本文方法具备较好的轮廓检测能力。

(2)由于视皮层下神经元的单拮抗感受野具有动态调节机制，能够响应颜色边界与亮度边界；而方位敏感的初级视皮层双拮抗神经元，能够借助动态突触连接，实现轮廓信息的群体编码。于是，本文提出基于神经元颜色拮抗与群体编码的轮廓检测方法，以 BSDS500 图库为实验对象，其中验证集全部图像实验结果的平均  $P$  指标约为 0.55，相比而言本文方法更加贴近真实的视觉感知。

(3)针对颜色拮抗与群体编码模型无法显著地区分主体轮廓和纹理边缘，本文提出一种基于双层神经网络脉冲调控编码的轮廓检测方法，根据高级视觉皮层的反馈调控机制和选择注意机制，在双拮抗神经网络的基础上，再次构建高级视皮层神经网络用于获取选择注意图像，并将其反馈投递至初级视皮层，得到最终轮廓响应。BSDS500 验证集图像的实验结果平均  $P$  指标约为 0.58，说明本文方法能够进一步区分主体轮廓与纹理边缘。

**关键词：**视觉感知，轮廓检测，感受野，群体编码，反馈调控

## ABSTRACT

Contour information is the crucial part of image after dimension-reduction, which is significant to image applications in subsequent procedure, such as precise and efficient object recognition or image understanding. Traditional contour detection methods that describe the directivity and continuity of contour with mathematical models are usually capable to eliminate unwanted edge produced by background texture and discriminate major contour with non-contour edge. However, in contour detection task, traditional methods always ignore the transmitting and processing procedure of information stream based on visual perception mechanism, which is more like a kind of abstract mathematical description. In this paper, we paid attention to the mechanism of visual perception system in contour detection or extraction, and focused on spikes coding of neural networks and the response characteristics to specific visual stimulus over different levels in visual path, then we attempted to involve concrete applications of image process, such as image contour detection. Firstly, a model named orientative correlation of multiple receptive fields in visual path was constructed, which interpreted the formation mechanism of orientation selectivity and the function of different levels of visual path when carrying on contour detection; secondary, we proposed a method to accomplish the color opponent and population coding of neurons, and applied the method to contour detection for color images; finally, the double-level neural networks with spikes regulative coding was introduced to enhance the modulatory and selective effect for visual information of high-level visual cortex, which could efficiently get rid of the texture over background. The main tasks and research results are listed as follows:

(1) The contour detection method called orientative correlation of multiple receptive fields in visual path was proposed in this paper. Then the orientation selectivity of primary cortex was modeled on the characteristics of receptive field at each level in visual path and the transformational characteristics between adjacent levels. The method based on the pre-coding effect of neural network on retina, the adjustment mechanism for non-classical receptive field of LGN neurons, and the orientation selectivity of primary cortex, was utilized to contour detection. Images in RuG library were taken for processing, and the mean value of measure  $P$  was 0.43, indicating that method in this paper has favorable contour detection ability.

(2) Neurons over sub-cortex usually have the single color-opponent receptive field which can dynamically respond to visual stimulus such as color boundary and lightness boundary; and double color-opponent neurons which are orientation-sensitive on primary cortex can connect with each

other by dynamic synapse to code contour information in a population way. In this paper, the contour detection method based on color-opponent and population coding of neurons was proposed to detect images in BSDS500 library, and the mean value of measure  $P$  of validation set was 0.55, which indicated the new method is much closer to visual perception system.

(3) As the contour detection method based on color-opponent and population coding cannot significantly distinguish the major contour from texture edge, a new contour detection method based on spikes regulative coding mechanism of double-level neural networks was proposed. According to the feedback-regulative principle and selective attention effect of high-level cortex, another neural network of high-level cortex was constructed to capture attention map based on double-opponent neural network at primary cortex. Feeding the attention map back to the response of primary cortex, the final contour was obtained. The mean value of measure  $P$  of validation set in BSDS500 was 0.58, indicating this method can distinguish the major contour from texture edge.

**Keywords :** visual perception, contour detection, receptive field, population coding, feedback-regulation

## 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.....                          | 6  |
| 第 2 章 神经元模型及视觉感知机制的初步建模.....              | 7  |
| 2.1 神经元模型与脉冲编码.....                       | 8  |
| 2.1.1 Hodgkin-Huxley 神经元模型.....           | 8  |
| 2.1.2 Integrate-and-Fire 神经元模型.....       | 9  |
| 2.1.3 Leaky Integrate-and-Fire 神经元模型..... | 10 |
| 2.1.4 神经元的脉冲编码.....                       | 11 |
| 2.2 视觉感受野.....                            | 11 |
| 2.2.1 经典感受野.....                          | 11 |
| 2.2.2 非经典感受野.....                         | 12 |
| 2.3 方向选择特性.....                           | 13 |
| 2.4 选择注意机制.....                           | 14 |
| 第 3 章 基于视通路多感受野朝向性关联的轮廓检测.....            | 15 |
| 3.1 概述.....                               | 15 |
| 3.2 基本原理.....                             | 15 |
| 3.2.1 视网膜神经网络的前级编码.....                   | 15 |
| 3.2.2 LGN 神经元感受野的调节机制.....                | 17 |
| 3.2.3 V1 神经元的多感受野朝向性关联模型.....             | 19 |
| 3.3 算法流程与验证.....                          | 20 |
| 3.3.1 算法流程.....                           | 20 |
| 3.3.2 算法验证.....                           | 22 |
| 3.4 实验结果.....                             | 23 |
| 3.5 讨论与分析.....                            | 25 |



|                                |    |
|--------------------------------|----|
| 3.6 小结.....                    | 26 |
| 第 4 章 基于神经元颜色拮抗与群体编码的轮廓检测..... | 27 |
| 4.1 概述.....                    | 27 |
| 4.2 基本原理.....                  | 27 |
| 4.2.1 视皮层下神经元的颜色拮抗特性.....      | 27 |
| 4.2.2 初级视皮层双拮抗神经网络的群体编码.....   | 31 |
| 4.3 轮廓检测系统.....                | 34 |
| 4.4 实验结果.....                  | 36 |
| 4.5 讨论与分析.....                 | 39 |
| 4.6 小结.....                    | 41 |
| 第 5 章 基于双层神经网络脉冲调控编码的轮廓检测..... | 42 |
| 5.1 概述.....                    | 42 |
| 5.2 轮廓检测模型.....                | 42 |
| 5.2.1 高级视皮层点阵神经元.....          | 44 |
| 5.2.2 选择注意机制的脉冲频率编码.....       | 45 |
| 5.2.3 高级视皮层的反馈调控.....          | 47 |
| 5.3 实验结果.....                  | 47 |
| 5.4 讨论与分析.....                 | 49 |
| 5.5 小结.....                    | 50 |
| 第 6 章 总结与展望.....               | 51 |
| 6.1 总结.....                    | 51 |
| 6.2 展望.....                    | 52 |
| 致谢.....                        | 53 |
| 参考文献.....                      | 54 |
| 附录.....                        | 59 |

## 第 1 章 绪论

### 1.1 研究目的与意义

一直以来, 认知意识与客观事物的关系都被认为是人类科学事业面临的巨大挑战之一。古代哲学家与思想家对这一问题的探索仅仅停留在猜想阶段, 直到 19 世纪, 研究者开始关注大脑感知的形成机制, 并在系统神经科学和认知神经心理学的基础上逐步形成了现代认知神经科学, 进一步阐述了知觉或者意识的产生机理。视觉感知机制是认知神经科学的重要组成部分, 长期受到科研人员的关注。如今, 视觉感知机制已经发展成为近三十年来最热门的研究领域之一, 这个最早由人类好奇心驱使而诞生的科学领域不再局限于实验室阶段的理论探索, 而是渴望在新时代的互联网浪潮之中发挥其作为认知神经科学关键要素的“英雄本色”。一方面随着计算机系统技术和信息技术的不断发展, 利用计算机系统的仿真手段和信息处理的数学方法可以对视觉生理学实验进行有效的辅助分析, 尤其是计算机系统对海量实验数据的高效处理能力以及对实验结果的图形化展现形式, 能够在微观细胞层面上更加高效、直观地揭示视觉感知机制的电生理学基础, 例如细胞膜离子通道的极化特性<sup>[1]</sup>、神经元的脉冲发放特性<sup>[2]</sup>、神经元对刺激的编码方式<sup>[3]</sup>、神经元群体的脉冲整合机制<sup>[4]</sup>及神经网络中各神经元的突触连接特性<sup>[5]</sup>; 另一方面, 在视觉感知机制现有理论趋于成熟、未知领域逐渐明朗的背景下, 研究人员开始注意到视觉系统几乎无可挑剔的感知能力, 并且希望相关研究成果能够反作用于实际的工程应用中, 例如视通路不同层次感受野特性的联动机制<sup>[6]</sup>, 方向选择特性的形成机制和响应机制<sup>[7]</sup>, 高级视皮层神经元对初级皮层编码结果的反馈调控作用<sup>[8]</sup>, 在日常生活和工业生产领域都有十分广泛的应用前景。尽管目前对视觉感知机制的认识并不完整, 但这并不妨碍科研人员利用已知的实验数据和视觉感知理论去引领图像处理、机器视觉和人工智能的发展方向。

在图像处理和机器视觉的具体任务中, 通过有效提取主体目标的轮廓能够快速确定图像或者视频中的关键信息, 同时避免冗余数据的干扰, 有助于提高基于形状信息的特征提取、目标识别、图像理解以及图像数据压缩存储等具体任务的执行效率。因此, 如何从图像大数据中快速准确地提取关键性信息, 是当前面临的一个技术难题。图像的轮廓是图像像素值发生跃变的地方, 不同于背景纹理边缘的无序状态, 主体目标的轮廓表现出一定的方向性和连续性, 包括主体目标与背景、平面与平面的边界<sup>[9]</sup>, 是基于图像特定语意分割的信息筛选方法, 侧重于对主体目标形状的连续性描述, 被认为是人类识别目标过程重要的中层视觉特征之一<sup>[10]</sup>。由此可见, 轮廓提取是视觉系统对视觉信息进行精细汇聚的重要手段之一。在具体的轮廓检测过程中, 背景纹理产生的边缘实际上并不是主体目标的真实轮廓, 相反还会对真实轮廓造成干扰, 因此如何在复杂场景中有效剔除背景纹理边缘, 同时突显主体目标的真实

轮廓，是轮廓检测的本质问题。传统边缘检测方法，如 Canny<sup>[11]</sup>、Sobel<sup>[12]</sup>等算子，通常利用数学模型来刻画局部邻域像素在特定方向上的灰度变化，具有简洁快速的特点，但并不能有效区分主体轮廓和背景纹理边缘。而在视觉系统中存在多种局部视觉感知机制，能够对多层视觉信息进行筛选融合，从而高效地解决这一难题。

因此，借助认知神经科学对视觉感知机制的研究成果，例如视通路不同层次神经元的感受野特性，在视觉信息精细汇聚中有着重要的协同作用；视皮层方向选择特性的形成机制和响应机制，与神经元网络脉冲编码相结合，能够检测不同方位的轮廓信息，进一步提取完整的图像轮廓。可见，将视觉感知机制的计算模型应用于图像轮廓检测或提取具有重要的意义。

## 1.2 国内外研究现状

在数字图像处理领域，传统边缘检测方法基本上是从纯数学的角度构建微分算子来度量图像像素值的阶跃响应，常见的边缘微分算子有 Prewitt 算子<sup>[13]</sup>、Roberts 算子<sup>[14]</sup>、Sobel 算子和 Laplace 算子等，这些算子通常对噪声敏感，对简单场景的边缘检测效果尚可，但其在轮廓检测任务中还远不能满足要求；在此基础上，为了降低噪声的影响，Marr 提出了 LoG<sup>[15]</sup>高斯拉普拉斯算子，在边缘检测之前进行高斯滤波以削弱图像噪声，Canny 将边缘检测问题归纳为寻找最优检测算子的问题，引入的非极大值抑制和双门限阈值处理方法至今有着广泛应用；近年来，为满足轮廓检测精度要求和性能要求，在传统边缘检测方法的基础上许多学者做了诸多改进，例如 Mallat 小波<sup>[16]</sup>考虑了轮廓检测算子的多尺度特性，通过选取合适的滤波器，减少特征之间的相互干扰，有效解决了图像噪声和轮廓定位问题，而 Arbeláez P<sup>[17]</sup>以谱聚类分析为基本手段，通过全局框架内的多种局部特征融合，对图像进行轮廓检测，取得了十分不错的检测效果。尽管传统的轮廓检测方法相比边缘检测方法能够在一定程度上区分出主体轮廓与非轮廓的边缘，但大都没有考虑视觉感知机制在轮廓提取中的信息流传递处理过程，与视觉感知系统在复杂环境下快速、准确捕捉主体目标的能力相比，这些方法在生物智能性和计算高效性上均有不足。

随着计算机科学和认知神经科学的发展，视觉神经系统对外界信息的优秀处理能力受到了研究人员的重点关注。基于人类视觉感知特性，Grigorescu<sup>[18]</sup>等人利用二维Gabor能量模型模拟初级视皮层经典感受野的方向选择特性，并且依靠非经典感受野的抑制特性进行轮廓检测，有效减少了背景纹理的影响。在此基础上又出现了许多改进的方法：桑农<sup>[19]</sup>等提出了蝶形域抑制方法，在邻域神经元的抑制作用上考虑得更加详尽，较好地去除了背景纹理产生的边缘；杜晓凤<sup>[20]</sup>等提出了基于复合感受野的轮廓检测方法，在减少背景纹理干扰的同时，保护了较弱的轮廓；Yang<sup>[21]</sup>等利用神经元的颜色双拮抗机制和空间稀疏特性进行纹理抑制，构建外膝体神经元的颜色单拮抗和初级视皮层的颜色双拮抗模型，并利用二维高斯导函数检测特定方位的轮廓信息，最后通过空间稀疏抑制去除纹理边缘，在彩色图像的轮廓提取中取得了较好的效果。在此基础上，Yang K等人又通过融合包括方向、亮度、亮度对比度在内的多种视觉特征，以调节非经典感受野的周边抑制强度，最终的轮廓检测效果相比单一视觉要素

的模型有了显著提升<sup>[22]</sup>。此外，Ichida等人根据初级视皮层感受野周边抑制的响应机制<sup>[23]</sup>，提出了基于非经典感受野的轮廓检测模型，与经典感受野相比，检测效果有了较大改进；Zeng C等人发现初级视皮层中心-外周型感受野的边区和端区对视觉信息有着不同的响应机制，并在轮廓检测模型中引入自适应机制<sup>[24]</sup>；Shen Z等通过研究清醒猕猴初级视皮层神经元的中心-外周非连续性，发现非经典感受野中心区和外周区的视觉输入差异性决定了周边抑制的强度，提出了图像特征不变性检测模型<sup>[25]</sup>。Azzopardi<sup>[26]</sup>等根据Hubel与Wiesel对初级视皮层方向选择特性的推断，将多个外侧膝状体神经元感受野进行特定朝向的排列组合，实现了对不同方位视觉刺激的响应，最后将相关模型应用于轮廓检测之中，结果表明相比于Gabor函数的黑箱模拟，组合感受野的检测结果更加符合视觉感知特性。

可见，上述方法基本上是利用数学模型对相应视觉机制进行模拟，以满足这些机制在人类视觉系统中的作用，或者是将多种视觉特征进行融合以提升模型的信息处理能力，不管是哪种方式这些方法大都忽略了神经元脉冲放电的生理学基础作用，即数以千亿计神经元电活动与视觉认知的内在关联性<sup>[27]</sup>，因此势必会造成视觉信息的丢失或者冗余。直至1952年，Hodgkin和Huxley根据神经元细胞膜离子通道特性的研究成果，首次提出了神经元动作电位发放的Hodgkin-Huxley数学模型(HH模型)<sup>[28]</sup>，为基于动作电位的神经计算模型拉开了崭新的序幕。以动作电位发放的神经元模型为基本单元，研究人员又关注到视觉信号传递过程中的非线性信息可能与神经元的互连结构——神经突触息息相关，于是针对神经突触的可塑性，Carver等人的研究表明突触连接的强度差异会造成神经元对特定方位刺激的响应增强或者抑制<sup>[29]</sup>。实际上，视觉信息在视通路中的快速传递直至在大脑皮层形成知觉，都需要借助于局部视觉感知机制对信息的高效编码和解码。因此，为了进一步探索神经元的编码机制，Pillow等人的研究表明：仅仅依靠单个神经元的脉冲频率编码或者时间编码并不能完全表征神经元脉冲序列中蕴含的视觉信息，而神经元群体中多个神经元通过同步放电模式可以显著增加视觉信息的编码量<sup>[30]</sup>，他们还进一步利用概率统计学的理论给出了神经元群体的解码模型<sup>[31]</sup>，进一步完善了神经元群体的编码理论；同时，也有学者提出单纯依赖于神经元之间的突触作用并不能完全解释视觉神经系统对信息的封装和整合方式，他们认为应该从更加宏观的角度来研究神经元群体脉冲发放模式的波动<sup>[32]</sup>。此外，许多学者针对神经元动作电位蕴含的时间信息和空间信息展开研究：Reich等人通过对视皮层刺激响应模式的研究，发现视觉系统能够将视觉刺激的空间特征编码为神经元脉冲发放的时间特征<sup>[33]</sup>，为多种视觉特征的混合编码和解码提供了理论支持，Kayser等人针对视皮层神经元空间信息的编码机制展开研究，结果表明神经元会根据不同距离尺度对指定范围内的视觉刺激按权重进行综合编码<sup>[34]</sup>，这实际上为初级视皮层神经元感受野的黑箱模型，如Gabor函数模型、高斯导函数模型等提供了电生理学解释。另外，随着神经元动作电位模型和脉冲响应特性被逐步揭示，一部分理论开始在实际应用中发挥关键作用：例如Zhou Liang等利用脉冲耦合神经元网络模型(PCNN)，通过自适应的参数调节，在彩色图像的边缘检测任务中取得了不错的效果<sup>[35]</sup>；而Delorme等人从神经元脉

冲编码的实用性和高效性出发, 改进了脉冲编码方式, 以神经元放电的次序编码方式取代传统的发放率编码方式<sup>[36]</sup>, 大大减少了神经元的编码时间, 在实际仿真和具体应用中更加贴近真实的视觉感知系统; 另外, 脉冲神经网络(SNN)在新时期受到了许多学者的青睐, Kerr<sup>[37]</sup>等人将六边形感受野与SNN相结合对图像边缘进行检测, 发现特定尺度的六边形感受野相比矩形感受野更加符合视觉特性。

综上所述, 可以发现国内外的图像轮廓检测研究开始脱离传统数学模型, 逐步融入视觉感知与生物智能。一方面, 侧重于从宏观角度对人类视觉特性进行黑箱模拟或直接应用, 如利用双高斯差函数模拟神经元经典感受野并应用于边缘检测, 或以三高斯差函数模拟非经典感受野并应用于图像增强, 或以Gabor函数模拟初级视皮层神经元的方向选择特性, 最终应用于轮廓提取。另一方面, 主要侧重于微观神经元动作电位模型和脉冲编码机制的研究, 如神经元的脉冲序列模式分析、神经元突触连接作用以及神经元的群体编码方式等, 并且在依赖于神经网络的具体应用中往往会忽略了潜在的视觉感知机制, 如视觉感受野、方向选择、群体编码以及高级皮层调控机制等。所以, 本文研究基于视觉感知机制的图像轮廓检测技术, 在神经网络脉冲编码过程中结合视觉通路中多种重要的视觉感知机制, 对图像轮廓进行检测, 期望获得比其他方法更加优秀的结果, 或者为后续基于视觉机制的研究和应用提供一些启示。

### 1.3 本文主要工作及创新点

本文的研究内容以基于视觉感知机制的图像轮廓检测技术为中心展开。一方面, 从宏观角度出发, 对视通路各个层次的视觉感知机制以及层次间的联动机制进行研究和建模, 实现视通路不同结构之间的信息前馈传递或者反馈调控; 另一方面, 从微观神经网络脉冲编码机制出发, 研究神经元单细胞的脉冲编码以及神经元群体编码特性, 并在神经网络脉冲编码过程中结合多种视觉感知机制, 对图像轮廓进行检测。首先, 对视通路中普遍存在的视觉特性如方向选择等进行论述、初步分析与数学建模; 其次结合神经网络的前级编码、非经典感受野的调节特性以及方向选择特性的形成机制, 构建视通路多感受野朝向性关联模型, 对自然场景图像进行轮廓提取, 获得了不错的效果; 然后, 考虑到颜色信息对视觉编码的辅助作用, 提出视皮层下神经元的颜色拮抗动态响应机制, 在此基础上考虑神经元树突极性分布和动态突触连接, 提出双拮抗神经网络群体编码模型, 对彩色图像进行轮廓检测; 最后, 引入高级视觉皮层的反馈调控机制和选择注意机制, 在双拮抗神经网络的基础上再次构建高级视皮层神经网络用于获取选择注意图像, 并对初级视皮层的轮廓响应进行调控, 其轮廓检测效果进一步提高。

论文主要研究内容包括以下几个方面:

(1) 本文对神经元模型和视觉通路中不同层次的视觉感知机制, 包括方向选择、感受野、颜色拮抗、群体编码、选择注意进行了描述, 充分发挥其在复杂场景下感知亮度变化和颜色刺激的作用。考虑神经元模型的计算高效性和生物合理性, 选取漏积分放电(Leaky

Integrate-And-Fire, LIF)模型作为神经元的基本计算原型,用于构建神经网络;轮述了视觉通路不同层次的感受野特性,方向选择特性的黑箱模型与形成机制;研究神经元对颜色信息的响应能力,描述了神经元的颜色拮抗特性;介绍了神经元单细胞的脉冲编码特性,以及神经元群体的编码效应;最后为抑制图像中的纹理边缘,对高级皮层的选择注意机制和反馈调控机制进行叙述与建模。

(2)本文充分探讨了视觉通路中不同层次的神经元感受野特性,通过定位各层次感受野在轮廓提取过程中的具体任务分工,先后引入了视网膜神经元的轮廓信息脉冲编码机制、外膝体感受野的全局调节特性以及外膝体-视皮层信息传递过程中的层次关联特性,摒弃了传统上采用单感受野黑箱模型模拟方向选择特性的做法,提出一种基于视通路多感受野朝向性关联的轮廓检测方法,对复杂场景下的自然图像(来自 RuG 图库)进行轮廓检测,并与主流轮廓检测方法进行对比分析,结果表明本文方法能够有效提取主体目标的真实轮廓,在一定程度上具备生物智能性。

(3)由于基于多感受野朝向性关联的轮廓检测方法并未考虑颜色特征在轮廓检测中的关键作用,并且仅仅依靠点阵神经元的单细胞脉冲编码无法全面高效地捕捉轮廓信息。因此,提出一种基于神经元颜色拮抗和群体编码的轮廓检测方法。首先,利用视网膜锥细胞分离三个颜色通道和亮度通道;其次,根据视皮层下神经元的颜色单拮抗动态特性,对颜色边界和亮度边界进行自适应调节;然后按照单细胞神经元树突分布极性对方向选择特性响应机制的生理学解释,构建双拮抗神经网络对特定方位的轮廓信息进行检测,以脉冲发放频率作为单细胞的轮廓编码结果;最后,在群体感受野内引入神经元的动态突触连接,将单细胞的脉冲编码结果同步整合为神经元群体编码结果。将 4 个拮抗通道在 8 个方位上的轮廓响应融合为最终的轮廓响应。对 BSDS500 中的部分彩色图像进行轮廓检测,经过误检率、漏检率与综合指标的定量评价,表明本文方法可以有效检测彩色图像的轮廓。

(4)考虑到颜色拮抗与群体编码模型不能显著地抑制纹理边缘、突显真实轮廓,另外高级视皮层的选择注意机制在快速定位真实轮廓中有着重要作用,且初级视皮层的编码结果往往需要高级视皮层参与调控。因此,提出一种基于双层神经网络脉冲调控编码的轮廓检测新方法。构建初级视皮层双拮抗神经网络和高级视皮层点阵神经网络,以彩色图像作为视觉激励,经过视皮层下颜色单拮抗感受野的自适应调节,输入到双拮抗神经网络,以脉冲频率的同步整合值作为初级视皮层神经元的群体编码响应;随后以初级视皮层的响应值作为高级视皮层神经元的输入,将其脉冲频率输出映射为选择注意图像,并利用选择注意图对初级视皮层轮廓响应进行反馈调控,进一步增强主体轮廓、抑制纹理边缘。对 BSDS500 中的部分彩色图像进行轮廓提取,检测结果表明该方法能高效剔除背景纹理边缘,保留完整连续的真实轮廓;另外,计算最佳检测结果与基准轮廓图的  $P$  值指标,表明本文方法具备更加智能的轮廓检测效果。

## 1.4 论文的组织结构

本文组织结构及章节内容简述如下：

第一章绪论，首先介绍了认知神经科学的发展历程、视觉感知机制在互联网时代的应用前景以及图像轮廓检测的价值与难点，进而引出本文研究在图像处理领域的研究目的和意义；然后对国内外相关领域的研究现状进行阐述分析，总结视觉感知机制或者神经元脉冲编码应用于图像处理领域的优势和不足；针对现有图像轮廓检测方法的局限性，有针对性地给出了本文的主要研究内容以及创新点；最后对论文的组织结构和章节安排进行简要说明。

第二章首先介绍了几种神经元模型及其动作电位的发放特点，然后简要描述了视觉通路各个层次的视觉感知机制及其计算模型，主要包括方向选择、视觉感受野、脉冲编码和选择注意机制。为后续计算模型的改进以及轮廓检测模型的构建奠定了理论基础。

第三章提出了一种基于视通路多感受野朝向性关联的轮廓检测方法，探讨了视网膜神经节细胞、外侧膝状体以及初级视皮层在视觉通路中的信息加工机制；根据不同层次感受野的特征提出了相应的感受野计算模型，同时重点研究了视网膜-外膝体以及外膝体-视皮层的信息投递过程中感受野的层次关联特性，最终通过外膝体感受野的朝向性关联实现初级视皮层感受野的方向选择特性。将模型应用于自然场景图像的轮廓检测之中，经过与标准轮廓图的  $P$  值指标分析，结果表明本文方法能够有效提取轮廓，相比其他方法更加贴近真实视觉感知。

第四章考虑到颜色特征在轮廓检测中的关键作用，以及神经元群体编码对轮廓信息的整合汇聚，构造神经元的颜色拮抗机制与群体编码网络，对彩色图像进行轮廓提取。经视网膜对彩色图像进行颜色、亮度通道分离之后，根据神经元的颜色单拮抗特性对轮廓信息进行动态调节；随后在双拮抗神经元网络中考虑神经元树突极性分布，并实现方向选择特性；最后结合神经元群体编码效应，将4个拮抗通道在8个方位上的轮廓响应融合为最终的轮廓响应。对彩色图像进行轮廓检测，经过误检率、漏检率与综合指标的定量评价，表明本文方法在彩色图像轮廓检测方面具备明显的优势。

第五章提出一种基于双层神经网络脉冲调控编码的轮廓检测新方法。针对双拮抗神经网络不能显著抑制背景纹理边缘、突显主体轮廓，在初级视皮层双拮抗神经网络的基础上构建高级视皮层点阵神经网络，结合选择注意机制对初级视皮层的群体编码响应进行反馈调控。其中双拮抗神经网络的作用机制与第四章保持一致，即彩色图像经过颜色单拮抗感受野调节之后，输入到双拮抗神经网络，然后利用动态突触模型获取神经元的群体编码响应；在此基础上，构建高级视皮层点阵神经网络，以编码实现选择注意机制，并利用选择注意图像对初级视皮层四个颜色拮抗通道的响应结果进行反馈调控，得到最终的图像轮廓。对 BSDS500 的图像进行轮廓检测，从定性结果看该方法能够有效地剔除背景纹理边缘，拥有更高的生物智能性，另外计算最佳检测结果与基准轮廓图的  $P$  值指标，要优于其他对比方法。

第六章对本文的研究内容和成果进行总结、归纳，并对后续可能的研究方向进行了展望。

## 第 2 章 神经元模型及视觉感知机制的初步建模

大千世界五彩缤纷的客观事物承载着各式各样的视觉信息，它们或移动或静止、或远或近、或鲜艳或暗淡，不管视觉刺激借助何种载体、呈现何种形式，人类都能通过自身的视觉系统(HVS) 快速、准确地捕捉外界的各种视觉变化，小到纹理差异、光照波动、距离变化，大到行为理解、意识联想、视觉记忆。视觉系统近乎完美的信息处理能力，依赖于视觉系统中数以千亿计神经元的电生理学活动，并且视觉系统特定区域的神经元通过神经突触相互连接形成神经元网络，以动作电位的方式为视觉系统响应视觉刺激提供最基本的功能保障，同时为视觉通路中各种视觉感知机制的形成奠定了重要的生理学基础。神经元作为视觉系统的基本单位，经过生理学实验对其细胞膜离子通道特性的分析，产生了多种描述动作电位发放和传播的神经元响应模型，例如 HH 模型与 LIF 模型。此外，为了对不同形式的视觉特征进行描述，神经元具备不止一种脉冲编码方式：例如脉冲频率编码、次序编码等，而且针对神经元之间的相互作用，还需要考虑神经元的群体编码方式。

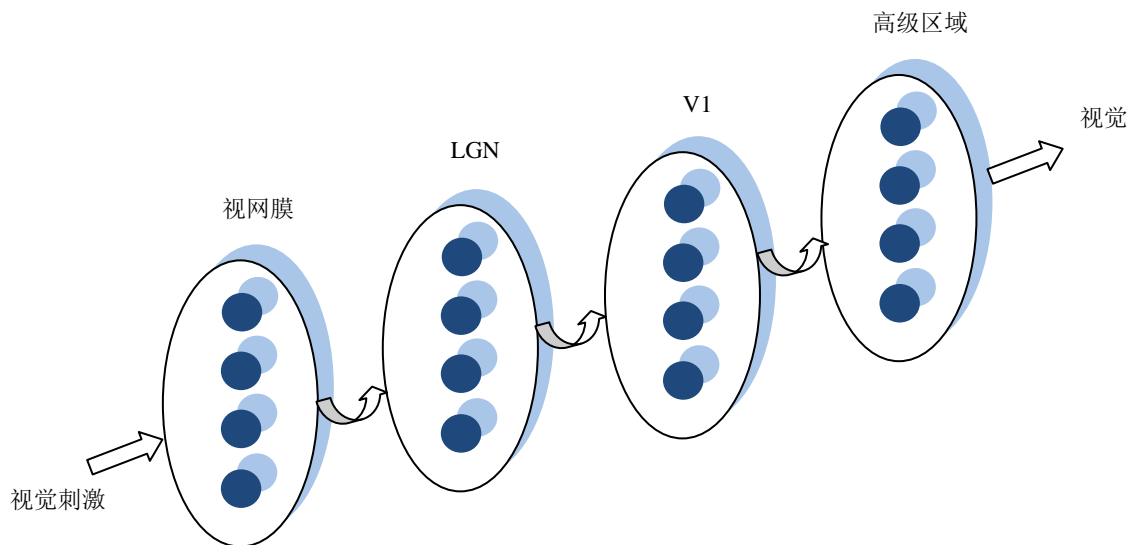


图 2.1 视觉系统的基本信息处理通路

另一方面，从全局角度来看，人类视觉系统对视觉信息的加工处理存在一个基本的前馈通路，如图 2.1 所示：视觉信息经过眼角膜与瞳孔到达眼球，经过视网膜感光细胞的光电变换和神经节细胞对信息的精细汇聚后，传入丘脑的外侧膝状体 (Lateral Geniculate Nucleus, LGN)，然后通过外膝体-皮质投射为视皮层提供输入，最后经过视觉皮层和其他大脑结构的处理，形成视觉感知<sup>[38]</sup>。在视觉通路的不同层次，神经元表现出不同的视觉特性与感知机制，例如视网膜神经节细胞的经典感受野呈现同心圆式拮抗结构，LGN 细胞的非经典感受野具有更大的空间作用范围，而初级视皮层简单细胞的感受野具有方向选择特性，能够检测特定



方位的视觉刺激。可见，视觉通路上不同层次的视觉感知机制在视觉信息处理过程中承担着不同的任务分工，而且相邻层次之间关联紧密，在信息投影过程中扮演着关键角色。当然，除了早期视觉通路，视觉系统同样依赖于高级视觉结构对视觉信息的甄选作用，并且不同于早期视觉通路的前馈传递方式，高级视觉结构通常以反馈调控的方式作用于初级视觉响应，形成更加完整的信息交互网络。

因此，越来越多的研究人员考虑在视觉神经计算模型中融入各种视觉感知机制，以期获得实验结果更加符合真实的视觉感知。

## 2.1 神经元模型与脉冲编码

神经元作为视觉系统神经网络的基本单元和视觉感知机制形成的基本素材，具有重要的基础作用。目前，研究人员提出了许多神经元计算模型，它们通常都是根据神经元细胞膜通透性和离子通道特性，对细胞膜内外电位的变化规律进行分析，然后进一步利用微分电路模型描述神经元动作电位的产生。最为经典的神经元模型就是 Hodgkin 和 Huxley 提出的 HH 模型，该模型充分考虑了神经元动作电位的产生原理，为基于神经元动作电位的编码机制奠定了数学基础；另一方面，考虑到神经元模型在实际应用中的效率问题，诸多学者在 HH 模型的基础上提出了许多改进模型，希望在构建大规模神经元网络的时候具有更加高效的计算性能，例如由 HH 模型简化而来的 IF 模型。当然，也有很多研究人员为了平衡神经元模型生物合理性与计算高效性，进一步提出了既不丢失生理学特性，同时在实际应用中能快速响应的神经元模型，比较典型的是漏积分放电 (LIF) 模型。下面对常用的神经元模型进行简单介绍。

### 2.1.1 Hodgkin-Huxley 神经元模型

Hodgkin-Huxley 模型，简称 HH 模型，是两位科学家对神经元细胞膜内外电位差异以及离子通道特性进行研究的基础上，运用数理方法建立的神经元响应模型，两人因此获得了 1963 年的诺贝尔生理医学奖。该模型准确描述了神经元膜电位、电流在刺激发生后随时间的变化关系，比较精确地反映了单个神经元的基本生理活动，为后续大规模神经网络的构建以及基于神经元放电活动的视觉感知机制奠定了数学基础。HH 神经元模型的基本数学表达式<sup>[39]</sup>如式(2.1)所示。

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

其中， $I_{\text{ext}}$  是外部刺激电流输入； $V$  为细胞膜内外电位差，即膜电压； $C_m$  是膜电容； $g_{\text{Na}}$ 、 $g_{\text{K}}$ 、

$g_L$  分别表示钠离子、钾离子和漏电流通道的最大电导值；相应地， $V_{Na}$ 、 $V_K$ 、 $V_L$  分别表示对应离子通道的逆转电位； $m$ 、 $h$ 、 $n$  分别描述各离子通道的通透性，其中  $m$  与  $h$  为钠离子通道的门控变量， $n$  为钾离子通道的门控变量； $m_\infty(V)$ 、 $h_\infty(V)$ 、 $n_\infty(V)$  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。HH 神经元模型动作电位发放示意图如图 2.2 所示，其中动作电位以尖峰脉冲的形式呈现。

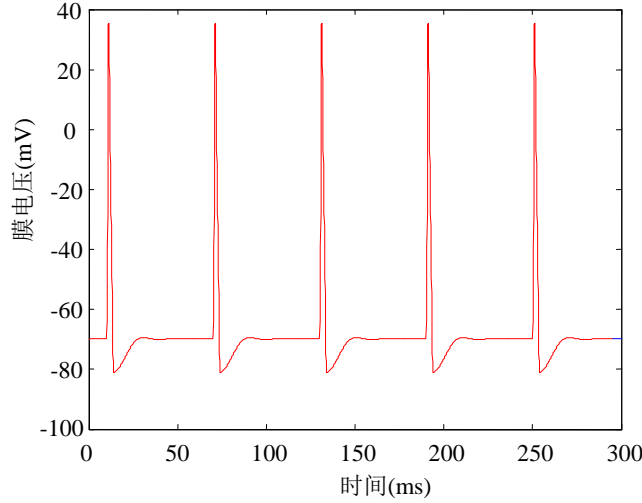


图 2.2 HH 神经元模型细胞膜电位变化示意图

### 2.1.2 Integrate-and-Fire 神经元模型

积分放电(Integrate-and-Fire, IF)神经元模型是描述神经元膜电位产生与演化的常用数学模型，一般情况下包含了神经元在噪声、外部刺激或者膜电位阈值变化时的电位发放与衰减特性<sup>[40]</sup>。相对于真实的神经元，尽管部分神经元特性被简化忽略，比如绝对不应期和相对不应期等，但是由于基本的神经元放电特性，如阈值放电特性等得以保留，在具备神经元基本生理特性的基础上，数学表达式的整体描述变得更加简洁，弥补了 HH 神经元模型计算复杂、效率偏低的重要问题，因此在实际应用场景或者构建神经网络的时候，得到了广泛的使用。IF 模型的基本数学表达式，如式 2.2 所示，首先利用简单的 RC 电路建立 IF 神经元模型，然后在刺激电流的作用下，RC 积分电路不断地为电容充电，当电容电压达到一定阈值时，神经元产生脉冲发放。

$$\begin{cases} c_m \frac{dv}{dt} = I_{\text{ext}} \\ \text{if } v \geq v_{\text{thresh}}, \text{ then } v = \text{const} \end{cases} \quad (2.2)$$

式中， $v$  是神经元膜电压， $c_m$  是膜电容， $I_{\text{ext}}$  是刺激电流， $v_{\text{thresh}}$  为电压阈值， $\text{const}$  为电压常数。IF 神经元模型表示的细胞膜电位产生与变化，如图 2.3 所示。另外，由于简化了一部分神经元的发放特性，IF 模型的神经元膜电压不会出现类似 HH 模型的负值情况。

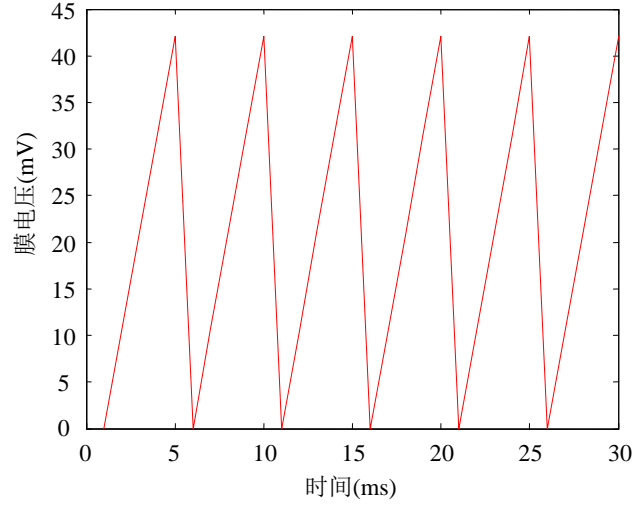


图 2.3 IF 神经元模型细胞膜电位变化示意图

### 2.1.3 Leaky Integrate-and-Fire 神经元模型

漏积分放电(Leaky Integrate-and-Fire, LIF)神经元模型<sup>[41]</sup>是在 IF 模型的基础上, 对模型的生物合理性和计算实用性进行了合理地平衡化, 即将神经元的漏电流特性和不应期特性重新加入到模型当中, 后来又有学者在 LIF 模型中考虑了更多的神经电生理特性。LIF 模型的数学表达式如式(2.3)所示:

$$\begin{cases} c_m \frac{dv}{dt} = -g_l v + I_{\text{ext}} \\ \text{if } v \geq v_{\text{thresh}}, \text{ then } v = \text{const} \end{cases} \quad (2.3)$$

式中,  $v$ 、 $c_m$ 、 $I_{\text{ext}}$ 、 $v_{\text{thresh}}$  和  $\text{const}$  的含义与式(2.2)保持一致,  $g_l$  是漏电导, 当  $v$  大于  $v_{\text{thresh}}$  时, 神经元脉冲发放, 同时  $v$  被瞬间重置为  $\text{const}$ 。图 2.4 是 LIF 模型的细胞膜电位变化示意图。

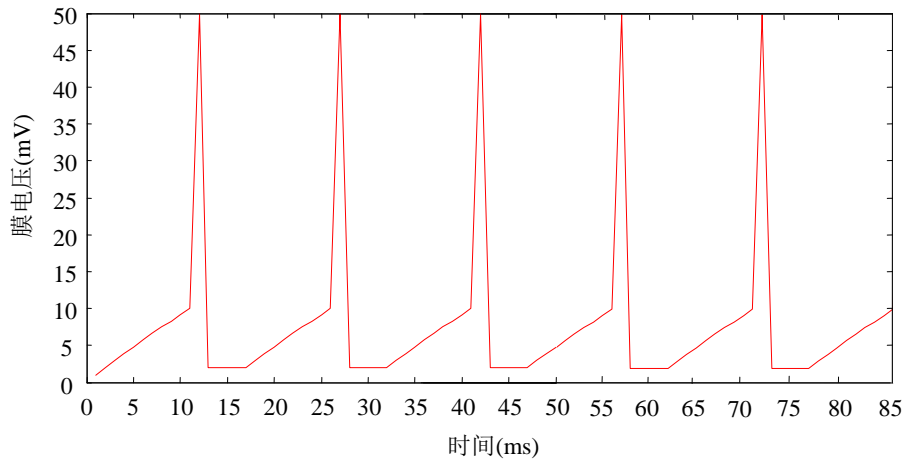


图 2.4 LIF 神经元模型细胞膜电位变化示意图 (考虑绝对不应期)

### 2.1.4 神经元的脉冲编码

随着神经元计算模型的提出和改进, 研究人员开始关注神经元放电序列中蕴含的神经编码信息, 事实上神经元的脉冲编码是将外界视觉进行数量化的过程, 有助于从视觉系统的最底层去解释相关视觉特性的形成机制, 有着十分重要的研究价值。目前, 学术界普遍支持神经元脉冲放电有着潜在的规律和特性, 是神经元对外界视觉信息进行编码的基本形式。例如, 神经元动作电位的发放频率在神经元编码机制中得到了广泛的使用, 即神经元脉冲发放频率中反映了外界视觉刺激的某种特征, 具有重要的编码价值<sup>[42]</sup>。此外, 考虑到视觉系统对外界信息的快速反应能力, 基于脉冲发放频率的编码方式难以满足视觉系统实时性的要求, 并且通过单一神经元的脉冲发放频率表征复杂的视觉信息有较大的难度, 因此有学者考虑以若干神经元的先后放电次序对视觉信息进行编码, 即次序编码机制<sup>[43]</sup>, 以满足视觉系统响应复杂视觉刺激的实时性需求。另一方面, 神经元的群体编码理论得到了研究人员的广泛认可, 即神经元通过彼此间的相互作用组成群体网络, 从而更加全面、高效地对视觉信息进行编码, 并且群体编码机制的时间特性与神经元感受野的空间特性有着千丝万缕的联系。

## 2.2 视觉感受野

在人类视觉系统中, 视觉感受野(Receptive Field, RF)是指在受到一定视觉刺激之后, 神经元放电活动发生变化的视觉区域。上述视觉区域除了自身在响应视觉刺激中的作用之外, 还是其他许多视觉特性或者感知机制的组成素材。在视觉信息处理通路中, 不同层次的神经元感受野表现出不同的视觉特性, 一般来说从视觉通路的低层次到高层次, 感受野的大小逐步增大、结构变得更加复杂。此外, 相关研究进一步表明: 感受野具备动态变化特性, 即感受野的大小、相对位置、空间频率与兴奋强度等随着外界视觉刺激情况的改变而实时变化<sup>[44]</sup>; 在经典兴奋-抑制型感受野的外周区域, 还存在一个大外周区, 这个区域对视觉刺激没有直接响应, 而是对感受野中心的响应会产生特定的调制作用<sup>[45]</sup>, 从而引出了非经典感受野的概念。

### 2.2.1 经典感受野

经典感受野通常具备中心-外周的同心圆结构, 包括两部分区域: 兴奋性区域和抑制性区域。如图 2.5 所示, 经典感受野分为中心兴奋型(ON)和中心抑制型(OFF), 其中“+”表示兴奋性区域, “-”表示抑制性区域。利用中心-外周的差异化结构, 经典感受野被认为可以检测视觉差异性信息。在具体实现时, 双高斯差函数(DoG)数学模型<sup>[47]</sup>经常被用来定量描述经典感受野的输入输出特性, 并且能够很好地反映经典感受野的同心圆结构, 其基本的数学表达式如式(2.4)所示。

$$\text{DoG}(x, y) = A \exp\left(-\frac{x^2 + y^2}{\sigma_1^2}\right) - B \exp\left(-\frac{x^2 + y^2}{\sigma_2^2}\right) \quad (2.4)$$

其中,  $x$  和  $y$  表示空间变量;  $A$  和  $B$  分别表示感受野兴奋性区域和抑制性区域的敏感度;  $\sigma_1$  和  $\sigma_2$  分别代表感受野兴奋性区域和抑制性区域的尺寸大小。图 2.6 是 DoG 模型的三维示意

图与响应示意图。

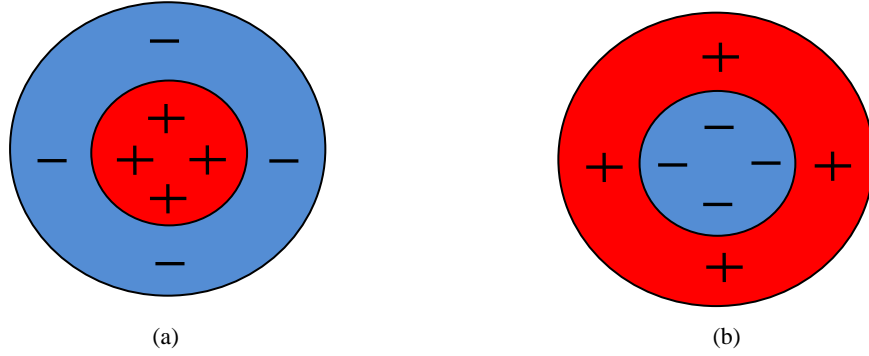


图 2.5 经典感受野示意图。(a)ON 型经典感受野；(b)OFF 型经典感受野

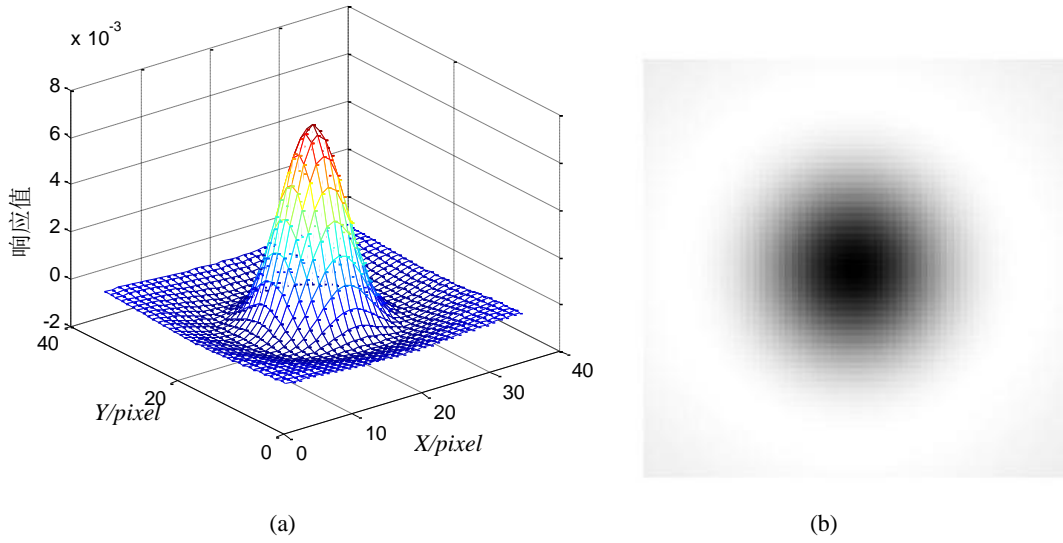


图 2.6 DoG 模型。(a)三维示意图；(b)响应示意图

### 2.2.2 非经典感受野

随着对感受野的深入研究，大量电生理学实验表明在经典感受野的外周存在一个对经典感受野中心区域响应进行调节的特殊区域，这一调节效应称为外周效应<sup>[48]</sup>，对应的区域称为非经典感受野的去抑制区域。于是，李朝义<sup>[49]</sup>等根据对猫视网膜神经节细胞和外膝体细胞的电生理学实验结论，提出了三高斯模型来模拟非经典感受野的特性，其数学表达式如式(2.5)所示。

$$TG(x, y) = A \exp\left(-\frac{x^2 + y^2}{\sigma_1^2}\right) - [B \exp\left(-\frac{x^2 + y^2}{\sigma_2^2}\right) - C \exp\left(-\frac{x^2 + y^2}{\sigma_3^2}\right)] \quad (2.5)$$

在式(2.5)中， $x$  和  $y$  表示空间变量； $A$ 、 $B$  和  $C$  分别表示非经典感受野兴奋性区域、抑制性区域与去抑制区域的敏感度； $\sigma_1$ 、 $\sigma_2$  和  $\sigma_3$  分别代表非经典感受野兴奋性区域、抑制性区域和去抑制区域的尺寸大小。非经典感受野的示意图和三高斯模型的响应示意图如图 2.7 所示。

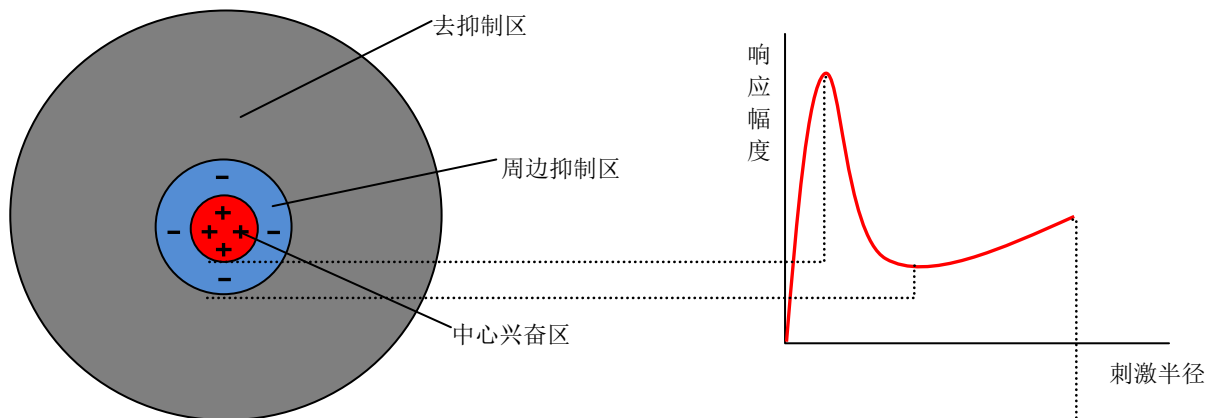


图 2.7 非经典感受野与三高斯模型响应的示意图

## 2.3 方向选择特性

研究表明，初级视皮层简单细胞的感受野对特定方位的视觉刺激相比于其他方位，具有更加强烈的响应，并且一般认为是由于外膝体细胞的感受野按照特定朝向进行排列组合，使得初级视皮层神经元感受野具备方向选择特性。方向选择特性对于检测特定方位的弱小刺激有着关键作用，因此很多学者对其展开了研究，提出了许多方向选择特性的计算模型。其中，Gabor 滤波器<sup>[50]</sup>由于具有良好的带通特性与方向敏感特性，并且能够检测明暗对比的边缘轮廓信息，在方向选择特性的建模之中得到了广泛使用；另外，利用二维高斯函数的导数模拟初级视皮层神经元的方向选择特性，能够较为快速地响应各个方位的视觉刺激，同时检测视觉输入的差异性信息，其方向选择特性的图形化结果如图 2.8 所示。

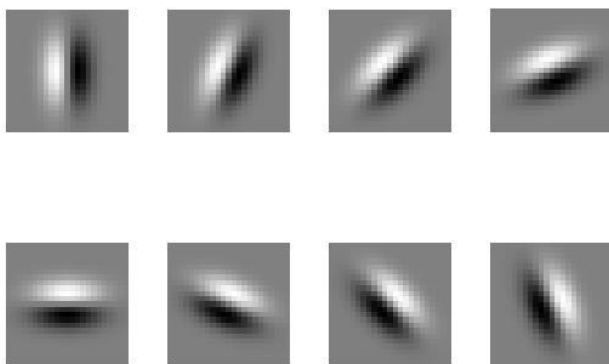


图 2.8 二维高斯导函数的方向选择特性示意图

实际上，不论 Gabor 滤波器还是高斯导函数都忽略了方向选择特性的形成机制，即若干外膝体感受野的特定几何排列方式对于方向选择特性的生理学解释。本文将在第 3 章对这一问题进行详细分析。

## 2.4 选择注意机制

选择注意指的是人眼观测外界特定场景时，对场景中的若干特定对象表现出更加浓厚的兴趣并选择其作为注意力中心，而忽略场景中其他次要元素的现象<sup>[51]</sup>。在选择注意区域，视觉系统通常以更加精细的处理方式对视觉信息进行响应，并且其产生的响应对整个场景的视觉整合具有十分重要的调节作用。通常，可以按照自底而上(bottom-up)与自顶而下(top-down)两种方式对选择注意机制进行数学建模，自底而上指的是视觉刺激中的显著性信息决定了视觉系统对注意区域的选择；而自顶而下则强调了主观意识下根据具体的视觉任务选择场景中任务相关的目标，将其定义为选择注意区域。关于选择注意机制的建模，有学者从视觉感知特性出发，通过计算经典感受野和非经典感受野的视觉输入差异性，最后形成显著性图像<sup>[52]</sup>。其具体数学表达式如式(2.6)~(2.8)所示。

$$\Delta\phi(x, y) = \lambda \times \max\{\bar{\phi}_C(x, y) - \bar{\phi}_S(x, y), 0\} \quad (2.6)$$

$$\bar{\phi}_C(x, y) = \sum_{(m,n)} I(x+m, y+n) \times g(m, n) \quad (2.7)$$

$$\bar{\phi}_S(x, y) = \sum_{(m,n)} I(x+m, y+n) \times \text{DoG}^+(m, n) \quad (2.8)$$

其中， $\Delta\phi(x, y)$ 为通过灰度逆映射得到的显著性图像； $\lambda$ 定义为 $\bar{\phi}_C(x, y) / \bar{\phi}_S(x, y)$ ， $\bar{\phi}_C(x, y)$ 和 $\bar{\phi}_S(x, y)$ 分别表示以高斯函数和 DoG 函数为距离因素来度量经典感受野和非经典感受野空间视觉输入； $I(x, y)$ 表示视觉刺激， $g(m, n)$ 表示高斯函数。

## 第3章 基于视通路多感受野朝向性关联的轮廓检测

如前文所述，基于视觉特性或者视觉感知机制的轮廓检测方法大都忽略了视觉神经系统的电生理活动，即数以千亿计神经元电活动与视觉认知的内在关联性，另外也没有考虑视觉信息在视觉通路加工和处理过程中的多感受野特性、感受野的层次关联特性，以及这些特性所依赖的生理学结构基础，所以轮廓检测结果与主观的视觉感知存在一定的差异。于是，本章考虑了视通路中各层次的感受野尺度特性以及层次间感受野的相互关联性，模拟视通路中视网膜神经元的脉冲编码过程，以及外侧膝状体细胞非经典感受野的全局调节特性，同时有朝向性地将多个 LGN 细胞感受野进行关联，用于模拟初级视皮层(V1)简单细胞的方向选择特性，提出了一种基于视通路多感受野朝向性关联的轮廓检测方法，充分发挥上述视觉特性在轮廓检测中的重要作用。

### 3.1 概述

本章首先考虑了视网膜神经节细胞对视觉信息的前级编码作用，根据神经元放射状树突的接收特性构建检测空间差异性信息的神经元网络，利用脉冲频率对轮廓信息进行初步编码；然后考虑视网膜-外膝体信息投影过程中感受野的变换特性以及外膝体的非经典感受野特性对前级编码结果进行调节，初步剔除背景纹理的干扰信息；最后，根据初级视皮层简单细胞方向选择特性的形成机制，将多个 LGN 细胞感受野进行朝向性关联，有效实现了方向选择特性，在轮廓检测中发挥了关键作用。

### 3.2 基本原理

在视觉通路中考虑各个层次的感受野特性与层次变换特性，依次构建视网膜神经元网络、外膝体 LGN 非经典感受野调节机制以及 LGN 感受野朝向关联形成初级视皮层方向选择特性，通过视通路三个层次的协同作用完成图像轮廓信息的获取。

#### 3.2.1 视网膜神经网络的前级编码

视网膜中感光细胞接受外界刺激，触发下游神经元产生动作电位，使刺激信息以脉冲发放的形式传递至后级神经元以及视觉皮层进行处理，直至形成视觉感知或识别。在神经信息编码过程中，神经元的树突接收刺激并将冲动传入细胞体。树突多种多样的空间表现形式，是视网膜对外界视觉信息进行精细汇聚的结构基础<sup>[53]</sup>。如图 3.1 所示，结合神经元的放射状树突结构与感受野特性，将感受野内所有树突接收的外界激励综合作为模型输入，同时以包含轮廓强度信息的脉冲发放频率对外界视觉激励进行编码。

本文从神经元模型的生物合理性与计算便捷性出发，选取仿真性能相对优越的漏积分放电(LIF)模型来构建神经元网络，对视觉输入的差异性进行检测。并且引入神经元放电过程中



的绝对不应期，改进后的 LIF 模型如式(3.1)所示。

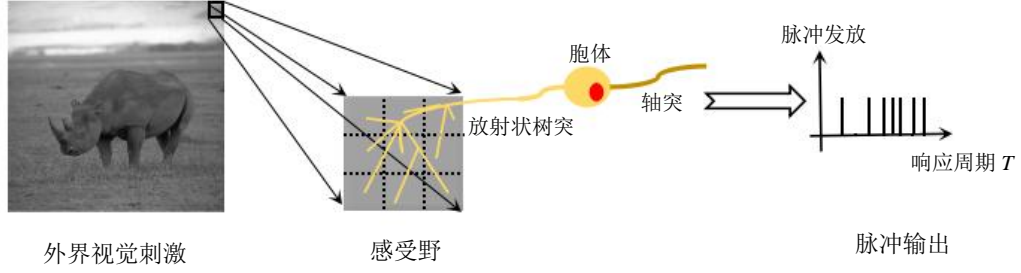


图 3.1 神经元利用放射状树突接收刺激并产生脉冲输出

$$\begin{cases} v = v_{\text{reset}}, \text{ref} = \text{ref} - 1, \text{if } \text{ref} \neq 0 \\ c_m \frac{dv}{dt} = -g_l v + \max(\sum_{(x,y) \in \text{RF}} \omega_{xy}^{\mu} I_{xy}, 0), \text{if } \text{ref} = 0 \\ v = v_G, \text{ref} = \text{const}, \text{if } v > v_{\text{th}} \end{cases} \quad (3.1)$$

在式(3.1)中， $x, y$  分别是视觉刺激的横、纵坐标， $v$  是神经元膜电压， $c_m$  是膜电容， $g_l$  是漏电导， $I_{xy}$  是与视觉刺激  $(x, y)$  对应的神经元树突激励电流，即输入图像像素值的归一化结果， $\text{RF}$  表示神经元的感受野， $\omega_{xy}^{\mu}$  是与视觉刺激  $(x, y)$  相对应的神经元树突接收权值，上标  $\mu$  表示感受野的极性，取值为 on 或 off，前者表示兴奋型感受野，后者表示抑制型感受野； $v_{\text{th}}$  为脉冲发放阈值， $v_{\text{reset}}$  为静态电势， $\text{ref}$  为绝对不应期， $v_G$  为脉冲发放峰值。当  $v$  大于  $v_{\text{th}}$  时，神经元脉冲发放达到  $v_G$ ，同时  $v$  又被瞬间重置为  $v_{\text{reset}}$ ，开始进入绝对不应期。 $\max$  表示取较大值运算，实现激励电流的非负处理。

视网膜神经元对外界刺激的前级编码能够充分反映空间的差异性信息。本文在神经元网络中引入了放射状树突的接收权值，同时为了编码外界输入的空间差异性信息，采用一种基于空间距离变化的双高斯差函数，对接收权值的变化规律进行模拟：感受野内的树突接收权值随着距离增大逐步减小，之后随着距离进一步增大出现负效应（接收权值为负数），最终因距离过大而趋向于零，具体如式(3.2)所示。

$$\omega_{xy}^{\text{on}} = \frac{1}{2\pi(0.5\sigma)^2} \exp\left(-\frac{x^2 + y^2}{2(0.5\sigma)^2}\right) - \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (3.2)$$

其中  $\sigma$  为双高斯差函数的标准差。

由于神经元感受野会呈现出兴奋或抑制两种状态，以两种类型的树突接收权值表示这种特性，于是有：

$$\omega_{xy}^{\text{off}} = -\omega_{xy}^{\text{on}} \quad (3.3)$$

本文统计了在一个响应周期  $T$  内，每个神经元的脉冲发放次数，并将脉冲发放频率映射为灰度图像，从而得到视网膜神经元网络对图像差异性信息的前级编码结果，如图 3.2 所示。

显然，视网膜神经元只能对图像的差异性信息进行粗略检测，并不能有效区分主体轮廓和背景纹理边缘，但这一过程将为视通路后续层次的处理提供原始数据。

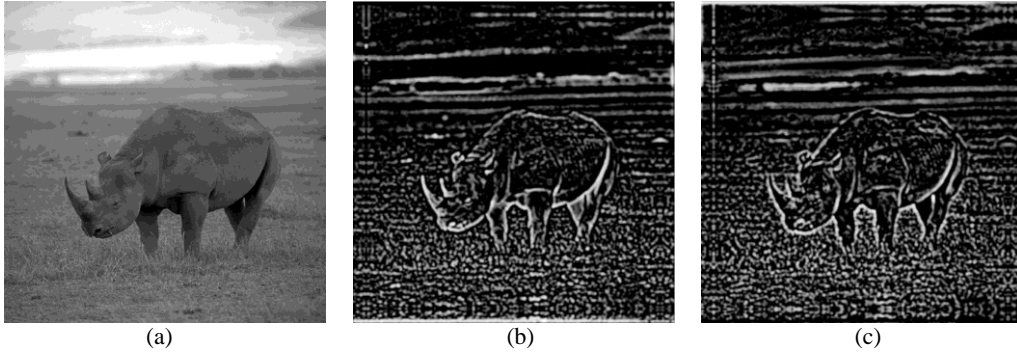


图 3.2 视网膜神经元网络的编码结果。(a) 原灰度图；(b) on 型编码结果；(c) off 型编码结果

### 3.2.2 LGN 神经元感受野的调节机制

早期视觉通路中，从视网膜到 LGN 的投射路径，包含了超过 90% 的视神经轴突，并且通过膝状体-皮质投射为皮质提供输入。本节主要关注的是视网膜-膝状体通路中感受野的层次特性和调节机制。

一方面，考虑到 LGN 细胞的感受野与视网膜神经元的感受野相比，结构相似但却具有更大的空间范围，能够综合处理来自双眼的信息。不同于传统的感受野模型，仅仅关注图像中以像素点为处理单元的局部细节；本文考虑了在视网膜到 LGN 的信息投影过程中，感受野呈现的由局部到全局的“点-块”变换特性，因此提出了以图像块为处理单元的 LGN 细胞感受野模型——将视网膜神经元编码结果做高斯平滑处理后，在空间上以图像块为基本单元作全局分解：即以一个目标点  $(x, y)$  为中心构建局部窗口，计算窗口覆盖区域的平均值，选取覆盖区域中与平均值相差最大的像素点作为该区域的全局特征点  $WI^u(x, y)$ ，并将其作为 LGN 细胞的输入。首先计算窗口覆盖区域的平均值如式(3.4)所示。

$$ave(x, y) = \frac{1}{ws^2} \sum_{1 \leq i, j \leq ws} GC^u((x-1) \cdot step + i, (y-1) \cdot step + j) \quad (3.4)$$

其中， $ws$  是局部方窗的窗长； $step$  为移动步长，设置为与窗长相等； $(x, y)$  是图像块的坐标， $(i, j)$  图像块内部的像素值坐标。

其次，选取与平均值相差最大的像素点作为全局特征点，如式(3.5)~(3.6)所示。

$$WI^u(x, y) = GC^u((x-1) \cdot step + p, (y-1) \cdot step + q) \quad (3.5)$$

$$\begin{aligned} & |GC^u((x-1) \cdot step + p, (y-1) \cdot step + q) - ave^u(x, y)| = \\ & \max(|GC^u((x-1) \cdot step + i, (y-1) \cdot step + j) - ave^u(x, y)|) \quad \forall 1 \leq i, j \leq ws \end{aligned} \quad (3.6)$$

在式(3.5)~(3.6)中， $||$  表示取绝对值运算； $\max$  表示取最大值运算； $(p, q)$  表示局部方窗覆盖区域内与平均值相差最大的像素点所对应的坐标索引； $GC^u(x, y)$  是高斯平滑后的视

网膜编码结果，上标  $\mu$  与树突接收权值  $\omega_{xy}^\mu$  的上标相对应，表示感受野的极性。

另一方面，利用外膝体的非经典感受野(non-Classical Receptive Field, nCRF)特性对全局特征图进行 nCRF 调节，可以有效抑制图像中的背景纹理边缘，同时在一定程度上能够增强主体轮廓。具体而言，在亮度差异性较大的主体轮廓区域呈增强性调节；在亮度差异性较小的背景纹理边缘区域，借助前文所述的全局特征图的空间尺度变化，使背景纹理边缘得到有效剔除。本文对传统的三高斯差函数进行微调，引入了局部半波整流，如式(3.7)所示。

$$\text{TDoG}(x, y) = \text{gau}^{\text{sur}}(x, y) - \max(\text{gau}^{\text{inh}}(x, y) - \text{gau}^{\text{exc}}(x, y), 0) \quad (3.7)$$

其中， $\max$  表示取较大值运算，用于实现半波整流； $\text{gau}^{\text{sur}}(x, y)$ 、 $\text{gau}^{\text{inh}}(x, y)$ 、 $\text{gau}^{\text{exc}}(x, y)$  是三个高斯函数，分别表示 LGN 细胞非经典感受野的大外周区、周边抑制区与中心兴奋区。周边抑制区的高斯函数标准差通常为中心兴奋区的 4 倍，且大外周区高斯函数标准差与中心兴奋区相同。并以微调后的三高斯差函数  $\text{TDoG}(x, y)$  为度量因子，计算全局特征图  $\text{WI}^\mu(x, y)$  中的亮度差异性  $\text{LD}^\mu(x, y)$ ，如式(3.8)所示。

$$\text{LD}^\mu(x, y) = \max(\sum_{(p, q)} \text{WI}^\mu(x + p, y + q) \cdot \text{TDoG}(p, q), 0) \quad (3.8)$$

式中， $(p, q)$  表示像素点的坐标增量；进一步将亮度差异性  $\text{LD}^\mu(x, y)$  的负指数归一化结果作为非经典感受野的调节量  $\text{nCRF\_aj}^\mu(x, y)$ ，把调节量与全局特征图按权值相融合后，通过双线性插值方法将融合结果调整为原输入图像的大小，将其作为 LGN 细胞的输出  $\text{LGN}^\mu(x, y)$ ，具体如式(3.9) ~ (3.10)所示。

$$\text{nCRF\_aj}^\mu(x, y) = -\exp\left(-\frac{\text{LD}^\mu(x, y)^2}{2\sigma^2}\right) \quad (3.9)$$

$$\text{LGN}^\mu(x, y) = \text{resize}(\text{WI}^\mu(x, y) + \alpha \cdot \text{nCRF\_aj}^\mu(x, y)) \quad (3.10)$$

其中， $\delta$  是负指数化的强度参数，设置为 0.2； $\text{resize}$  表示双线性插值放大运算； $\alpha$  是融合权值，反映了全局调节的作用强度。图 3.3 给出了上述 LGN 细胞感受野调节过程中的 on 型结果，而 off 型结果可通过类似的调节过程得到。从部分 on 型调节结果中可以看出，LGN 非经典感受野的全局调节能够在一定程度上抑制背景纹理，同时对视网膜前级编码结果中的主体轮廓进行增强。

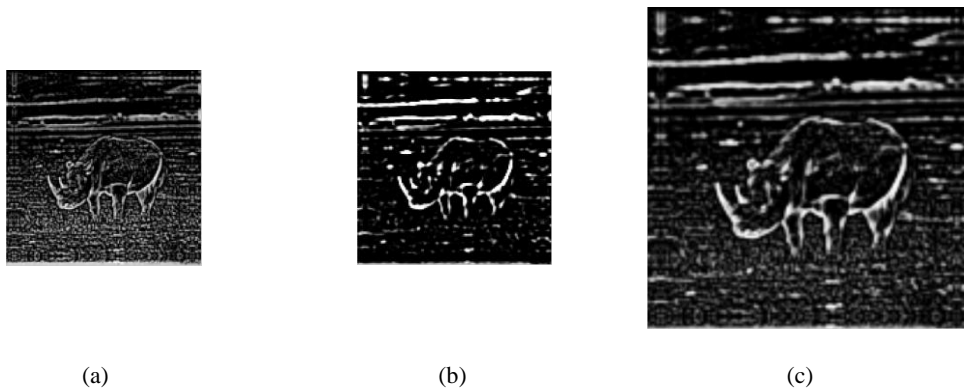


图 3.3 LGN 细胞感受野调节过程中的部分结果。(a) on 型编码结果的全局特征图；(b) on 型全局特征图的 nCRF 调节量；(c) 经过双线性插值放大的 on 型调节结果，即 LGN 细胞的 on 型输出

### 3.2.3 V1 神经元的多感受野朝向性关联模型

如前文所述，在图像处理领域的相关研究中，大都选择具备方向选择特性的二维 Gabor 函数或者二维高斯导函数作为初级视皮层简单细胞的计算模型。事实上，以这些函数对简单细胞建模只是一种粗略的黑箱模拟，并没有考虑到简单细胞方向选择特性的形成机制。

上文提到视网膜的视觉信息通过膝状体-皮质投射为皮质提供输入，可见 V1 简单细胞的感受野特性与上一层 LGN 细胞的感受野有着千丝万缕的联系：Hubel 和 Wiesel 认为初级视皮层的简单细胞可以由多个相邻 LGN 细胞同心圆式感受野的输出相连而成，而简单细胞的方向选择特性与多个 LGN 细胞感受野的关联方式密切相关<sup>[54]</sup>。

本文以外侧膝状体和初级视皮层的层次关联特性为切入点，综合考虑简单细胞方向选择特性的生理学形成机制以及计算模型的简洁高效性，构建如图 3.4 所示的 V1 简单细胞多感受野朝向性关联模型，其中绿色系图标代表 LGN 细胞的 on 型感受野，橘色系图标代表 LGN 细胞的 off 型感受野，蓝色系图标代表 V1 简单细胞的感受野。

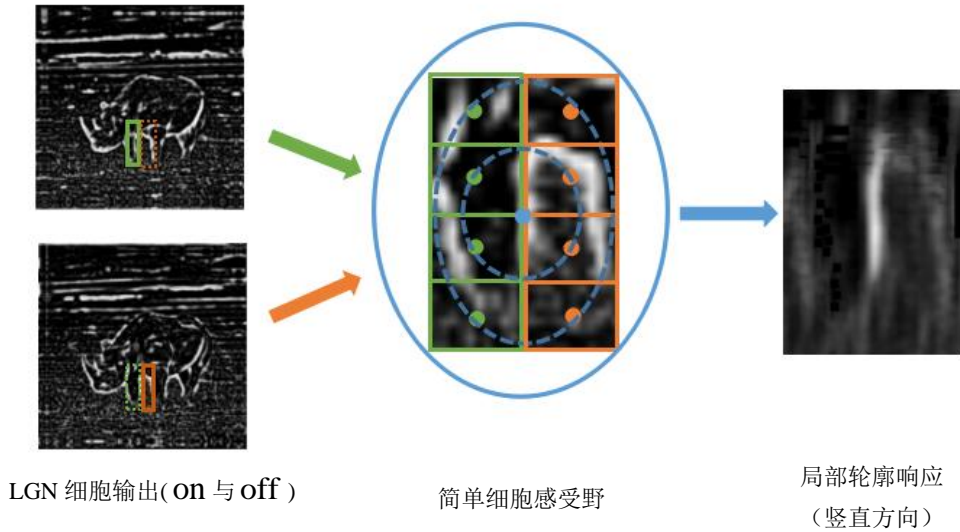


图 3.4 多感受野朝向性关联模型的示意图

以尺寸为  $w \times w$  的矩形感受野代替 LGN 细胞的同心圆式感受野，且矩形感受野的响应取自 LGN 细胞的输出；将相邻的  $N$  个 on 型与  $N$  个 off 型，共  $2N$  个矩形感受野的输出有朝向性地关联到同一个 V1 简单细胞。并且对矩形感受野的个数  $N$  进行如下限定： $1 \leq N \leq \text{round}(\min(m, n)/w)$ ， $m$ 、 $n$  分别是输入图像的宽和高， $\text{round}$  为取整运算， $\min$  表示取较小值。

用一个三元素的元胞集合 **RecRFSet** 表示多个矩形感受野的朝向性分布，具体如式(3.11)所示：

$$\text{RecRFSet} = \{(\mu_i, r_i, \varphi_i) | i = 1, \dots, 8\} \quad (3.11)$$

其中， $\mu_i$  为第  $i$  个矩形感受野的极性， $r_i$  为第  $i$  个矩形感受野中心位置到 V1 简单细胞感受野中心位置的距离， $\varphi_i$  为第  $i$  个矩形感受野中心位置和 V1 简单细胞感受野中心位置的连线相对

于水平方向的夹角；显然极坐标 $(r_i, \varphi_i)$ 确定了 $2N$ 个矩形感受野的相对位置，从而实现多个感受野的朝向性关联。朝向型关联之后的矩形感受野输出如式(3.12)所示：

$$\text{RecRF}_{\mu_i, r_i, \varphi_i}(x, y) = \text{LGN}^{\mu_i}(x - \Delta x_i, y - \Delta y_i) \quad (3.12)$$

其中， $\Delta x_i = -r_i \cos \varphi_i$ ， $\Delta y_i = -r_i \sin \varphi_i$ ，表示极坐标转换到直角坐标后各矩形感受野的位置偏移量。

当且仅当各个矩形感受野全部处于响应状态，多感受朝向性关联模型才会被激活，因此计算 $2N$ 个矩形感受野输出的几何平均值作为 V1 简单细胞的方向选择结果 $\text{os}_\theta(x, y)$ ：

$$\text{os}_\theta(x, y) = \left( \prod_{i=1}^{2N} \text{RecRF}_{\mu_i, r_i, \varphi_i}(x, y) \right)^{1/(2N)} \quad (3.13)$$

其中， $\theta$ 表示 V1 简单细胞的偏好方向，由集合 **RecRFSet** 决定。根据初级视皮层中不同偏好方向的简单细胞的感受野结构相似而输入来源不同，针对不同偏好方向的简单细胞的多感受野输入，采用不同朝向的关联方式，具体如式(3.14)所示。调整矩形感受野集合 **RecRFSet** 中每个元胞的极坐标方向 $\varphi_i$ ，即可实现其他朝向的关联方式，进一步获取其他 V1 简单细胞在其偏好方向上的方向选择结果。

$$\text{os}_{\theta_k}(x, y) = \left( \prod_{i=1}^{2N} \text{RecRF}_{\mu_i, r_i, \varphi_i + \Delta \varphi_k}(x, y) \right)^{1/(2N)} \quad (3.14)$$

其中， $\theta_k$ 表示其他偏好方向， $k=1, \dots, 24$ 表示不同偏好方向的索引； $\Delta \varphi_k$ 为朝向性关联时的角度增量。图 3.5 给出了多感受野朝向性关联模型的部分方向选择结果。

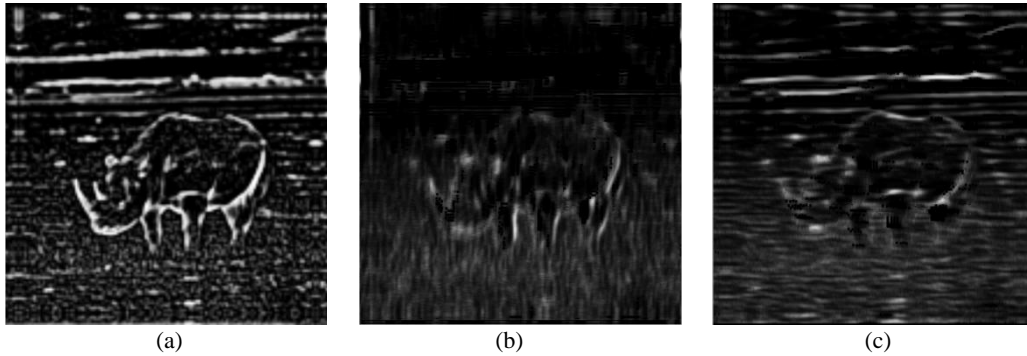


图 3.5 多感受野朝向性关联模型的部分方向选择结果。(a)LGN 细胞 on 型输出；(b)偏好方向为竖直方向的简单细胞的方向选择结果；(c)偏好方向为水平方向的简单细胞的方向选择结果

### 3.3 算法流程与验证

#### 3.3.1 算法流程

按早期视觉通路的信息处理流程对轮廓检测过程作如下叙述：(1)根据感受野内神经元的放射状树突接收特性，构建检测空间差异性信息的视网膜神经元网络，对轮廓信息进行前级编码；(2)基于感受野的层次变换特性，引入 LGN 细胞的非经典感受野调节机制，对前级编



码结果中的主体轮廓信息和背景纹理边缘进行全局调节；(3)考虑 LGN 与初级视皮层的层次关联特性以及初级视皮层简单细胞的方向选择特性形成机制，构建简单细胞的多感受野朝向性关联模型，实现对特定方向的轮廓检测；(4)综合全部方向的轮廓响应，经过非极大值抑制和阈值处理，得到轮廓二值图。算法执行流程如图 3.6 所示。

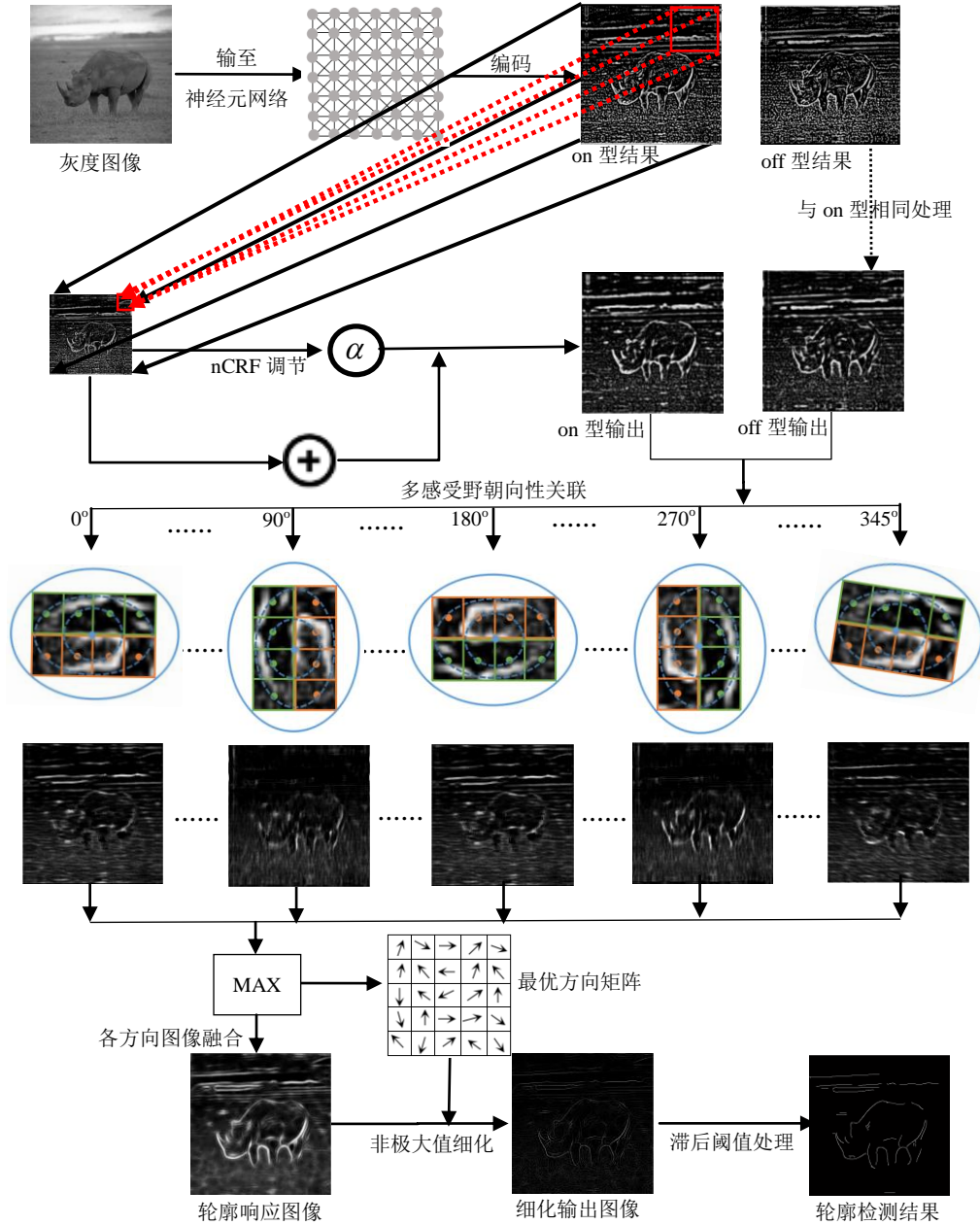


图 3.6 轮廓检测算法流程

具体检测流程如下：

步骤 1：根据输入灰度图像  $IG(x, y)$  ( $x=1, 2, \dots, m; y=1, 2, \dots, n$ ) 大小构建相同尺寸的视网膜神经网络  $GN(x, y)$ ，神经元接收感受野  $RF(7 \times 7)$  内的外界综合刺激，按式(3.2)、(3.3)计算树突接收权值  $\omega_{xy}^{\mu}$  的大小。

步骤 2：网络中的每个神经元都采用式(3.1)所示的 LIF 神经元改进模型，将灰度图像像

素归一化为神经元的激励电流，统计一个响应周期内各神经元的脉冲发放频率，作为视网膜神经元对轮廓信息的前级编码结果。

步骤 3：视网膜神经元编码结果经过高斯平滑得到  $GC''(x, y)$ ，按式(3.4)、(3.5)和(3.6)进行全局分解，得到全局特征图  $WI''(x, y)$ ，将其作为 LGN 细胞的输入。

步骤 4：按式(3.7)~(3.10)，模拟 LGN 神经元非经典感受野对全局亮度的调节机制，对全局特征图  $WI''(x, y)$  的主体轮廓和背景纹理进行调节，得到 LGN 细胞的输出  $LGN''(x, y)$ 。

步骤 5：根据式(3.11)提供的矩形感受野空间位置信息以及式(3.12)给定的矩形感受野输出值，对各个 LGN 细胞矩形感受野进行朝向性关联，最后计算各个矩形感受野的几何平均值作为初级视皮层简单细胞在其偏好方向上的方向选择结果，如式(3.13)所示。另外，利用式(3.14)可以得到其他偏好方向的简单细胞的方向选择结果。

步骤 6：利用式(3.15)融合各偏好方向简单细胞的方向选择结果，得到初级视皮层轮廓响应  $RV1(x, y)$ ，另外根据式(3.16)得到与灰度图像尺寸相同的最优方向矩阵  $BO(x, y)$ 。

$$RV1(x, y) = \max\{os_{\theta_k}(x, y) | \forall k = 1, \dots, 24\} \quad (3.15)$$

$$BO(x, y) = \theta_j, os_{\theta_j}(x, y) = \max\{os_{\theta_k}(x, y) | \forall k = 1, \dots, 24\} \quad (3.16)$$

步骤 7：按式(3.17)对步骤 6 的轮廓响应进行非极大值抑制处理，紧接着利用式(3.18)进行滞后阈值处理，得到轮廓响应的二值图  $BC(x, y)$ 。

$$thin(x, y) = \begin{cases} RV1(x, y), & \text{if } (p2 \geq RV1(x, y) \geq p1) \\ 0, & \text{other} \end{cases} \quad (3.17)$$

$$BC(x, y) = \begin{cases} 1, & \text{if } (thin(x, y) \geq t) \\ 1, & \text{if } (0.5t \leq thin(x, y) < t \ \& \ thin(x, y) \rightarrow \text{edge}) \\ 0, & \text{other} \end{cases} \quad (3.18)$$

其中  $(p1, p2) = \text{intpol}(BO(x, y), RV1(x, y))$ ，表示  $RV1(x, y)$  在其 8 邻域中根据  $BO(x, y)$  进行梯度插值， $p1$ 、 $p2$  为  $RV1(x, y)$  在其最优方向  $BO(x, y)$  上的插值结果。 $t$  为滞后阈值， $\text{edge}$  表示已经确定的轮廓点， $\rightarrow$  表示 8 邻域相连接。

### 3.3.2 算法验证

自然场景图像的轮廓检测是利用图像方法分析和识别主体目标的基础内容，但自然场景图像由于轮廓纹理相似性、噪声干扰等特征，对其进行轮廓检测有较高的难度。为验证本文方法对自然场景图像的轮廓检测具备有效性，选取 RuG 图库中的 40 幅自然场景图进行轮廓检测实验，图库中每张图片都有一张手工制作的基准轮廓图，用于计算轮廓检测的定量评价指标，选取图库中部分有代表性特点的图片作为实验对象，如图 3.7 第一行所示。

本文首先选取直接模拟初级视皮层轮廓响应的数学模型作为对比方法，即与 Gabor 能量模型类似的二维高斯导函数模型(DG)，体现本文方法中 V1 多感受野朝向关联模型的生物合理性；另外，再选取文献[26]提出的仅考虑了 LGN 细胞排列组合的轮廓检测模型(CORF)，证

明本文方法中视网膜前级编码和 LGN 感受野调节的有用性；最后考虑纹理抑制特性，选择文献[21]中去除颜色双拮抗机制并适用于单通道灰度图像的空间稀疏限制方法(SSC)作为对比方法。实验中，选取 10 个不同的 LGN 非经典感受野调节强度，对应参数为  $\alpha = [0.1:0.1:1.0]$ ，另外为了与图像库提供的二值标准轮廓图进行比较，需要对提取的轮廓响应图进行非极大值抑制处理和滞后阈值处理<sup>[11,18]</sup>。在滞后阈值处理中，取上限阈值  $t = [0.1:0.05:0.8]$ ，因此对于本文方法和 SSC 方法可得到 150 组测试参数，而由于 DG, CORF 方法并不涉及全局系数  $\alpha$ ，故只进行 15 组测试参数。

为客观的对检测结果进行评价，采用文献[18]中的定量指标计算公式，对本文方法和对比方法做定量分析。显然最终得到轮廓二值图像 BC 的轮廓点坐标一般不可能与基准轮廓图 GT 中的轮廓点坐标完全吻合，相反如果检测结果中的各个轮廓点出现在基准轮廓图对应像素点的某个邻域内，就会显得更加合理。因此本实验设置了与轮廓检测对比算法相同的邻域参数：若检测到的轮廓点出现在基准轮廓点的  $5 \times 5$  邻域内，就认为该轮廓点检测正确。定义  $E_{BC}$  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，则算法正确检测到的像素集为  $E = E_{BC} \cap (E_{GT} \oplus T)$  ( $\oplus$  为膨胀操作， $T$  为  $5 \times 5$  的结构单元)；错误检测到的像素集为  $E_{FP} = E_{BC} - E$ ；漏检的像素集为  $E_{FN} = E_{GT} - (E_{GT} \cap (E_{BC} \oplus T))$ 。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(3.19)~(3.21)计算得到。

$$e_{FP} = \text{card}(E_{FP}) / \text{card}(E) \quad (3.19)$$

$$e_{FN} = \text{card}(E_{FN}) / \text{card}(E_{GT}) \quad (3.20)$$

$$P = \frac{\text{card}(E)}{\text{card}(E) + \text{card}(E_{FP}) + \text{card}(E_{FN})} \quad (3.21)$$

其中， $\text{card}(X)$ 代表求集合  $X$  中的元素的个数。显然，误检率漏检  $e_{FP}$  越小，表明检测结果具有更少的背景纹理；漏检率漏检  $e_{FN}$  越小，说明检测结果保留了更多的主体轮廓，因此当误检率和漏检率趋向于 0， $P$  趋向于 1，即  $P$  越大则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。

### 3.4 实验结果

图 3.7 给出了部分图片的轮廓检测对比结果，第一行为测试对象，即自然场景图；第二行为图库提供的基准轮廓图；第三行为 DG 方法检测结果；第四行为 CORF 方法检测结果；第五行为 SSC 方法检测结果；第六行为本文方法检测结果。从结果中可以看出，DG 方法直接用数学公式模拟初级视皮层的轮廓响应，然后利用阈值化来保留一定强度的轮廓，优点是检测出了绝大多数的轮廓细节，但阈值处理并不能有效抑制背景纹理，导致中结果出现较多的伪轮廓；相反 CORF 方法考虑了 LGN 感受野的空间组合特性，能一定程度上减少背景纹理干扰，但由于缺乏视通路其他层次的编码与调节，容易因为过抑制和阈值处理不当造成真实轮廓丢失；而 SSC 方法由于考虑了轮廓的显著性与稀疏特性，同时引入了纹理抑制，在保



留主体轮廓和抑制背景纹理上取得了不错的平衡结果，不过在主体轮廓的周边区域，对纹理边缘的抑制作用不够明显，形成较多的毛刺。综合而言，本文方法保留了尽可能多的主体轮廓信息，同时去除了大部分的背景纹理边缘，在图像纯净度上要优于 SSC 方法。

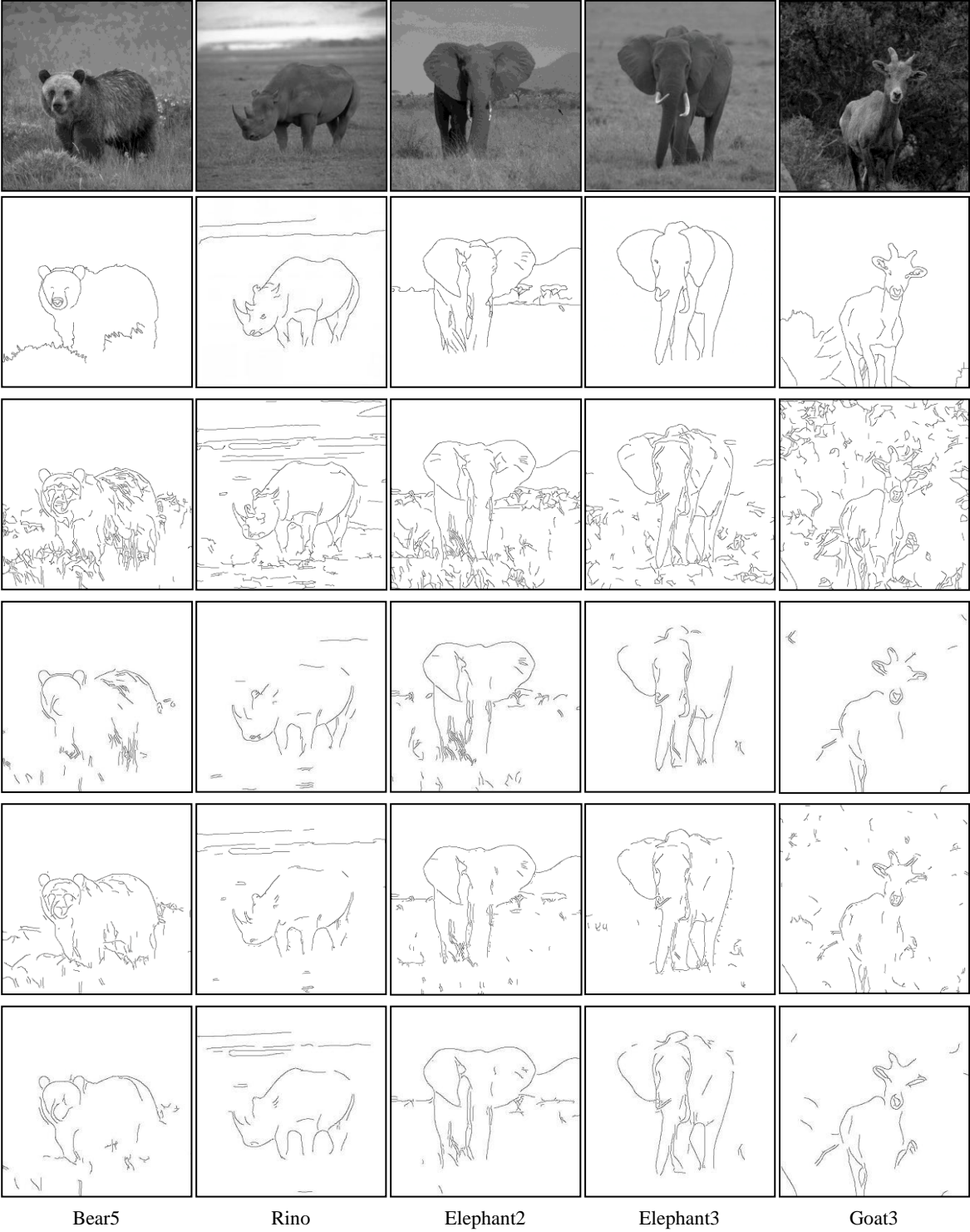


图 3.7 图库中的部分检测结果

按上述方法计算图 3.7 中各算法最佳结果的定量指标，如表 3.1 所示。不难看出，表中计

算结果基本与上文的定性分析相吻合。DG 方法的漏检率很小, 说明目标轮廓非常完整, 但误检率  $e_{FP}$  很大, 整体效果  $P$  值不太理想, 例如 Rino 的结果, 漏检率仅为 0.1, 而误检率却达到了 1.79, 显然较小的漏检率难以弥补极高的误检率所带来的负面影响, 导致整体效果不佳; CORF 方法则显著减小了误检率, 但漏检率控制的不太理想, 造成主体轮廓丢失, 整体效果  $P$  值的改善程度有限, 从 Rino、Elephant2 等结果都可以表明这一点; 而 SSC 方法与本文方法在误检率和漏检率上取了不同程度的平衡, 例如 Elephant2 的结果, 本文方法和 SSC 的漏检率分别是 0.32 与 0.24, 误检率分别是 0.23 与 0.42, 而整体效果  $P$  值保持一致, 均为 0.57, 总体来说本文方法误检率更小, 具有较高的图像纯净度, 而漏检率也处于一个较优水平。由表中  $P$  值可以看出本文方法的整体检测效果优于对比方法。另外, 计算全部 40 幅图片检测结果与对应基准轮廓图的平均  $P$  指标, DG、CORF、SSC 分别为 0.36、0.40、0.42, 而本文方法为 0.43, 可见本文方法在总体上具有更优秀的性能。

表 3.1 图 3.7 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     |          |             |
|-----------|------|----------|------|----------|----------|-------------|
|           |      | $\alpha$ | $t$  | $e_{FP}$ | $e_{FN}$ | $P$         |
| Bear5     | DG   |          | 0.10 | 3.30     | 0.13     | 0.22        |
|           | CORF |          | 0.5  | 1.63     | 0.47     | 0.28        |
|           | SSC  | 0.70     | 0.10 | 1.83     | 0.33     | 0.30        |
|           | 本方法  | 0.80     | 0.55 | 0.98     | 0.49     | <b>0.33</b> |
| Rino      | DG   |          | 0.10 | 1.79     | 0.10     | 0.34        |
|           | CORF |          | 0.40 | 0.34     | 0.48     | 0.44        |
|           | SSC  | 0.6      | 0.10 | 0.44     | 0.29     | 0.53        |
|           | 本方法  | 0.8      | 0.60 | 0.33     | 0.33     | <b>0.54</b> |
| Elephant2 | DG   |          | 0.10 | 1.46     | 0.11     | 0.38        |
|           | CORF |          | 0.20 | 0.48     | 0.30     | 0.53        |
|           | SSC  | 0.60     | 0.30 | 0.42     | 0.24     | 0.57        |
|           | 本方法  | 0.50     | 0.45 | 0.23     | 0.32     | <b>0.57</b> |
| Elephant3 | DG   |          | 0.10 | 2.0      | 0.07     | 0.33        |
|           | CORF |          | 0.20 | 0.32     | 0.27     | 0.59        |
|           | SS   | 1.0      | 0.10 | 0.62     | 0.17     | 0.54        |
|           | 本方法  | 0.4      | 0.45 | 0.33     | 0.23     | <b>0.61</b> |
| Goat3     | DG   |          | 0.10 | 2.77     | 0.15     | 0.25        |
|           | CORF |          | 0.40 | 0.32     | 0.57     | 0.37        |
|           | SSC  | 0.70     | 0.10 | 0.79     | 0.37     | 0.42        |
|           | 本方法  | 0.50     | 0.65 | 0.42     | 0.50     | <b>0.41</b> |

### 3.5 讨论与分析

从以上实验结果不难看出, 本文方法具有较强的轮廓提取能力, 特别是针对主体轮廓具备丰富细节, 并且背景纹理与主体轮廓相似性很强的自然场景图像, 从图 3.7 的定性结果以

及表 3.1 的整体性能指标都可以反推出这一结论。

相对于其他方法, CORF 方法虽然与本文方法的最终环节具有相似的视觉神经理论基础, 即都考虑了 LGN 感受野的空间组合特性, 但 CORF 方法理论支持的覆盖面有限, 仅仅考虑了视觉通路中 LGN 感受野与初级视皮层感受野的相互关系, 并且直接将待检测图像作为 LGN 感受野组合模型的输入, 而忽视了前级视网膜以及 LGN 非经典感受野的视觉信息处理能力, 一定程度上丧失了对早期视觉通路的完整性描述, 故而在检测结果中表现为纹理抑制不足或者真实轮廓丢失; 而本文方法不仅考虑 LGN 感受野与初级视皮层感受野的层次关联特性, 而且引入了视网膜对视觉信息的前级编码以及 LGN 非经典感受野的调节功能, 是对早期视觉通路的完整描述, 因此其检测结果可以比较完整的保留主体轮廓, 同时抑制背景纹理。另外 SSC 方法虽然与本文方法有着相近的检测水平, 但是 SSC 方法通过构建数学描述定义空间稀疏特性, 用于区分纹理区域和主体轮廓区域, 缺乏相应的生理学理论支撑, 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在主体轮廓的周边区域形成较多毛刺; 而本文方法根据 LGN 非经典感受野的调节特性, 实现背景纹理抑制以及主体轮廓增强, 具有更加充分的生物合理性, 并且在检测结果中有效区分了背景纹理和主体轮廓, 具有较高的图像纯净度。

本文所做工作及其创新性如下。(1)提出了一种基于视通路多感受野朝向性关联的视觉轮廓信息处理新方法。研究了视网膜、外侧膝状体以及初级视皮层的神经元感受野特性, 不同层次间感受野的空间尺度变换在视觉信息精细汇聚中的作用, 以及感受野的朝向性关联对方向选择形成机制的生理学解释, 突出了生物视觉系统的信息处理能力的应用研究。(2)实现了一种模拟视觉通路信息加工机制的图像轮廓检测具体应用。根据神经元的放射状树突接收特性, 构建检测空间差异性信息的视网膜神经元网络, 对轮廓信息进行前级编码; 建立 LGN 细胞的非经典感受野调节机制, 结合感受野的层次变换特性, 对前级编码结果进行全局调节; 提出简单细胞的多感受野朝向性关联模型, 模拟初级视皮层简单细胞的方向选择特性; 融合多个方向上的轮廓响应, 经过非极大值抑制和阈值处理, 得到轮廓检测结果。(3)为如何利用视觉通路多个层次的感受野关联特性进行图像轮廓检测, 或者是后续的目标识别与分析提供了新思路。

### 3.6 小结

视通路中神经元感受野在视网膜、外侧膝状体以及初级视皮层这三个逐级层次上呈现出不同的结构组织特性与刺激响应机制, 本文通过探讨不同层次神经元感受野在轮廓提取过程中的具体任务分工, 引入视网膜的轮廓信息脉冲编码、外膝体感受野的全局调节特性以及感受野的层次关联特性, 摒弃传统上采用单感受野数学模型模拟视觉机制的做法, 提出一种基于视通路多感受野朝向性关联的轮廓检测方法, 与 DG、CORF 以及 SSC 方法进行比较分析, 结果表明本文方法检测到的自然图像轮廓具有更加完整且连续, 同时抑制了大部分的背景纹理边缘, 具有较高的图像纯净度, 整体上反映了本文轮廓检测方法所具备的生物智能性。

## 第4章 基于神经元颜色拮抗与群体编码的轮廓检测

正如绪论中介绍，在基于边缘、轮廓等关键特征的工程应用中，颜色特征在彩色图像的轮廓检测和图像分割项目中是不可或缺的<sup>[55]</sup>。实际上，颜色信息在彩色图像轮廓检测任务中的作用不亚于灰度信息，而且神经元颜色单拮抗感受野对颜色边界和亮度边界的动态调节机制有助于进一步发挥两者在轮廓检测中的作用；另一方面神经元通过动态突触相互连接，为神经元的群体编码提供了生理学结构基础，为实现轮廓信息的同步整合起到了关键作用。因此本章针对彩色图像的轮廓检测任务提出了基于神经元颜色拮抗与群体编码的轮廓检测方法。以 BSDS500 彩色图像库中的自然场景图像为实验对象，通过  $P$  指标评价，结果表明本文方法能够有效检测主体轮廓，相比较其他基于视觉机制的轮廓检测方法具有一定的优势。

### 4.1 概述

本章提出一种基于神经元颜色拮抗与群体编码的轮廓检测方法，首先考虑视皮层下神经元的颜色单拮抗感受野对颜色信息和亮度信息的动态调节，在此基础上利用初级视皮层神经元的放射状树突极性分布构建方位敏感的双拮抗神经元网络对亮度边界和颜色边界进行脉冲频率编码，最后根据神经元动态突触连接特性将单细胞的编码结果进行同步整合，得到神经元群体编码结果，经非极大值抑制和阈值化处理，得到图像轮廓。

### 4.2 基本原理

在视觉通路的处理流程中，利用视皮层下神经元对颜色信息的单拮抗特性，建立颜色-亮度的动态调节模型；根据神经元放射状树突的极性分布，实现双拮抗神经元网络的方向选择特性；结合神经元动态突触特性，形成神经元群体，对脉冲发放频率进行同步整合，最终完成对图像轮廓信息的群体编码。

#### 4.2.1 视皮层下神经元的颜色拮抗特性

在颜色视觉形成的早期阶段，视觉系统对颜色的感知由视网膜的锥细胞起始，随着锥细胞光感受器将外界刺激转换为电信号，颜色信息流经过神经节细胞和外膝体神经元的颜色拮抗处理后得到进一步的剥离，并且在多个对立的颜色信道中进行传输。其中，神经节细胞和外膝体神经元作为初级视皮层的前级信息处理环节，具有十分相似的感受野结构与颜色拮抗机制，因此将两者统一归入视皮层下。

一方面，在视觉信息进入视皮层下之前，视网膜锥细胞根据刺激的波长分离视觉信息中的颜色通道。视网膜锥细胞对不同波长光谱的敏感程度不同，如图 4.1 所示，短波（蓝）锥细胞对波长 430nm 的光有最大反应，而中波（绿）和长波（红）锥细胞的感受性峰值平移到了更长的波长。因此，可以将锥细胞分为 L（长）、M（中）、S（短）三种类型，分别对应三原色 R（红）、G（绿）、B（蓝）。

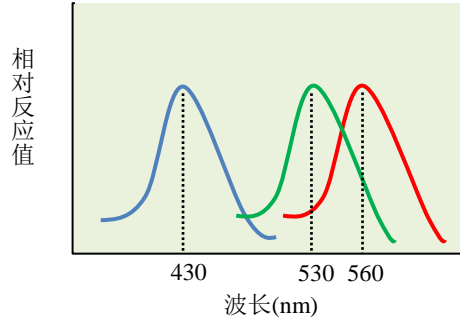


图 4.1 锥细胞的光敏感度函数

经过锥细胞对彩色图像颜色通道的分离，可以得到  $R(x, y)$ 、 $G(x, y)$ 、 $B(x, y)$ 、 $Y(x, y)$  四个颜色分量，其中  $Y(x, y) = (R(x, y) + G(x, y)) / 2$ 。考虑到图像亮度信息在视觉信息编码中的作用，以及锥细胞对亮度的高效捕捉能力，结合各颜色通道给出亮度通道的定义，如式(4.1)所示。

$$L(x, y) = 0.3R(x, y) + 0.59G(x, y) + 0.11B(x, y) \quad (4.1)$$

其中  $R(x, y)$ 、 $G(x, y)$ 、 $B(x, y)$  分别代表图像的红、绿、蓝三个颜色分量， $L(x, y)$  表示亮度信息。

另一方面，视皮层下神经节细胞与外膝体的输入来自视网膜锥细胞的颜色响应。而神经节细胞和外膝体具有结构相似的颜色拮抗感受野<sup>[56]</sup>，并且都呈现出单拮抗的输出方式。如图 4.2 所示，颜色单拮抗感受野包括中心-外周与单中心两种结构类型。

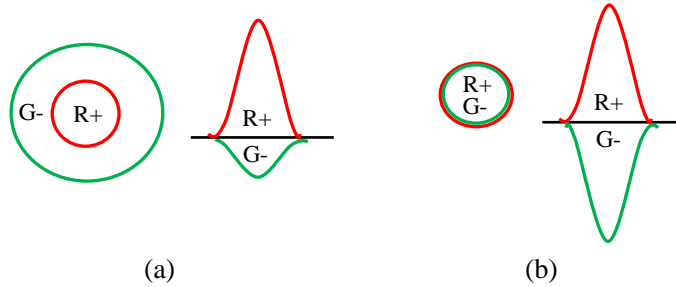


图4.2 视皮层下神经元的颜色单拮抗感受野。(a) 中心-外周型颜色单拮抗感受野；(b) 单中心型颜色单拮抗感受野

$$\begin{cases} SO_{RG}(x, y) = Gau1(x, y) * R(x, y) - Gau2(x, y) * G(x, y) \\ SO_{GR}(x, y) = Gau1(x, y) * G(x, y) - Gau2(x, y) * R(x, y) \\ SO_{BY}(x, y) = Gau1(x, y) * B(x, y) - Gau2(x, y) * Y(x, y) \\ SO_{YB}(x, y) = Gau1(x, y) * Y(x, y) - Gau2(x, y) * B(x, y) \end{cases} \quad (4.2)$$

本文以中心-外周型单拮抗感受野为建模对象，在颜色拮抗过程过融合了DoG（双高斯差）函数对空间距离的描述：利用两个不同尺度的高斯函数分别与对立的两个颜色通道做卷积运算，将卷积之差作为视皮层下神经元的颜色单拮抗输出，具体如式(4.2)所示。

其中， $SO_{RG}(x, y)$ 、 $SO_{GR}(x, y)$ 、 $SO_{BY}(x, y)$ 、 $SO_{YB}(x, y)$  依次代表  $R+/G-$ 、 $G+/R-$ 、 $B+/Y-$ 、

Y+/-四种类型神经元的颜色单拮抗响应； $R(x, y)$ 、 $G(x, y)$ 、 $B(x, y)$ 、 $Y(x, y)$  分别代表图像的红、绿、蓝、黄四个颜色分量； $Gau1(x, y)$ 与 $Gau2(x, y)$ 为两个不同尺度的高斯函数，其定义如式(4.3)所示。

$$\begin{cases} Gau1(x, y) = \frac{1}{2\pi\sigma_1^2} \exp(-\frac{x^2 + y^2}{2\sigma_1^2}) \\ Gau2(x, y) = \frac{1}{2\pi\sigma_2^2} \exp(-\frac{x^2 + y^2}{2\sigma_2^2}) \end{cases} \quad (4.3)$$

其中， $\sigma_1$ 与 $\sigma_2$ 分别代表中心-外周型颜色单拮抗感受野的中心区尺寸和外周区尺寸。

由于颜色信息的提取有助于分析图像的颜色边界，而亮度信息对图像亮度边界的获取同样有着重要意义，而根据颜色单拮抗感受野的模型分析，在分离图像亮度信息和颜色信息的过程中，当两个高斯函数定义的权值完全相等时，单拮抗模型只对颜色边界敏感，而对亮度边界不响应；当两个高斯权值的差异性增大时，单拮抗模型对亮度边界的响应得到加强，同时对颜色边界也有响应。

在此基础上提出单拮抗感受野的动态调节机制，具体表现为：在图像局部亮度较低的区域，减小两个高斯权值的差异性，从而发挥单拮抗模型对颜色边界的响应；在局部亮度较高的区域，增大两个高斯权值的差异性，从而使得亮度信息也参与到轮廓信息的编码过程之中。

在具体实现过程中，首先给定高斯权值差异性的初始定义： $k = \sigma_1 / \sigma_2$ ，并且初步限定  $k \in (0, 1]$ 。于是两个高斯权值的差异性完全由  $k$  的取值决定，当  $k = 1$  时，两个高斯权值相等，意味着单拮抗模型只对颜色边界敏感，此时颜色单拮抗感受野从中心-外周型退化为单中心型；当  $k \rightarrow 0$  时，高斯权值差异性增大，感受野外周区尺寸数倍于中心区尺寸，此时单拮抗模型对亮度边界也产生有效响应，同时也能捕捉颜色边界信息。

其次根据式(4.1)获取的亮度矩阵，统计局部亮度平均值，并将其归一化结果作为动态调节驱动因子  $L\_avg(x, y)$ ，如式(4.4)所示。

$$L\_avg(x, y) = (\text{mean}(L((x-1) \cdot \text{step} + i, (y-1) \cdot \text{step} + j))) / mlight \quad (4.4)$$

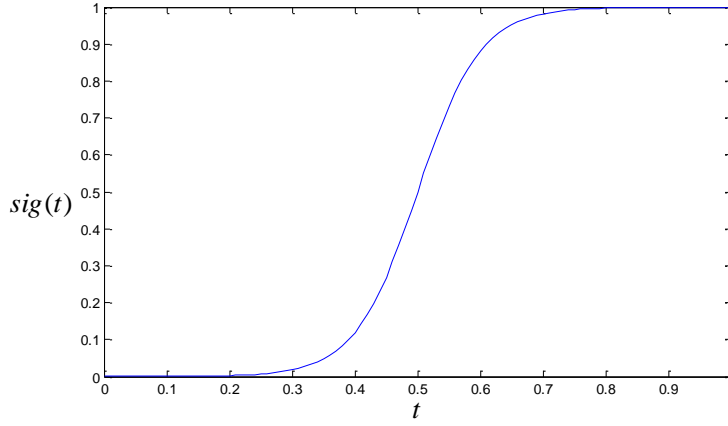
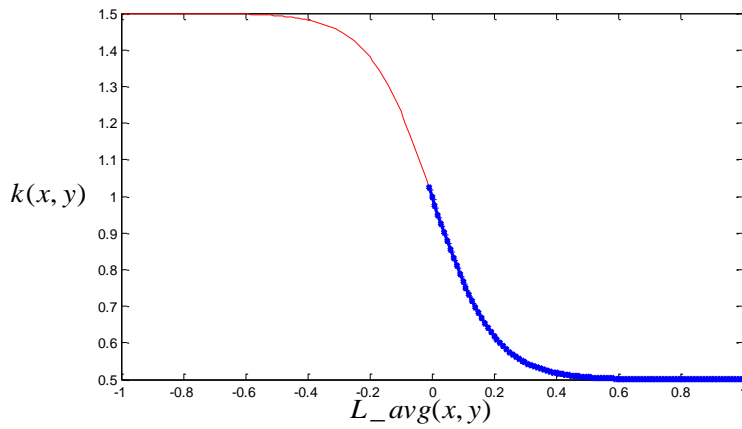
$1 \leq i, j \leq ws$

其中  $(x, y)$  表示局部窗之间的相对坐标， $(i, j)$  表示局部窗范围内的亮度值坐标； $ws$  是局部方窗的窗长，用于限定计算平均亮度的范围； $\text{step}$  为移动步长，设置为与窗长相等； $L\_avg(x, y)$  为局部亮度平均值； $mlight$  是亮度通道中最大的局部亮度均值。

为了使动态调节更加稳定、可控，本文通过一个改进的 sigmoid 函数对动态驱动因子进行约束，sigmoid 函数原型具体如式(4.5)所示。

$$\text{sig}(t) = 1 / (1 + \exp(-\alpha(t - \tau))) \quad (4.5)$$

其中，参数  $\alpha$  与  $\tau$  分别决定了函数的陡峭程度和水平位移，且当  $t \rightarrow \tau$  时函数输出有一个快速增长，其余情况函数总体变化缓慢，图 4.3 给出了其函数曲线走向图。


 图 4.3 sigmoid 函数曲线图( $\alpha = 20, \tau = 0.5$ )

 图 4.4 高斯权值差异性  $k(x, y)$  随  $L\_avg(x, y)$  的变化曲线

于是，给出高斯权值差异性  $k(x, y)$  随局部亮度  $L\_avg(x, y)$  动态变化的数学表达式，如式 (4.6) 所示。

$$k(x, y) = 0.5 + 1 / (1 + \exp(10 \times L\_avg(x, y))) \quad (4.6)$$

图 4.4 是高斯权值差异性  $k(x, y)$  随  $L\_avg(x, y)$  的变化曲线，其中蓝色曲线为实际有效区段。按照生理学对视觉皮层下神经元感受野的研究结论<sup>[57]</sup>，本文限定外周区感受野的最大半径约为中心区的两倍，即通过动态变化表达式将  $k(x, y)$  的最小值修正为 0.5。

通过视网膜锥细胞的颜色分离以及视觉皮层下神经元的拮抗感受野对颜色边界的动态响应，得到视觉皮层下颜色拮抗感受野的动态响应结果  $SO_{RG}(x, y)$ 、 $SO_{GR}(x, y)$ 、 $SO_{BY}(x, y)$ 、 $SO_{YB}(x, y)$ 。图 4.5 给出了各拮抗通道的动态响应示意图。



图 4.5 颜色拮抗通道的动态响应。(a)R+/G-; (b)G+/R-; (c)B+/Y-; (d)Y+/B-



#### 4.2.2 初级视皮层双拮抗神经元的群体编码

在颜色视觉形成的早期阶段，视网膜和视皮层下的神经元属性描述了颜色感知的一些基本计算过程，但颜色视觉的最终形成势必需要视通路皮质区域的参与。生理学研究表明，初级视皮层 V1 接收来自视皮层下神经元对外界刺激的响应；同时初级视皮层有相当数量的神经元具备颜色空间双拮抗特性<sup>[58]</sup>（具体形式如图 4.6 所示），能够同时响应颜色边界和亮度边界；而且人们普遍认为在初级视皮层定义了方位信息以及方向选择特性，即针对各个方位的视觉刺激，初级视皮层的特定神经元对其最优方位上的视觉刺激响应更加强烈，这种方向选择特性是准确捕获目标轮廓信息的关键因素。

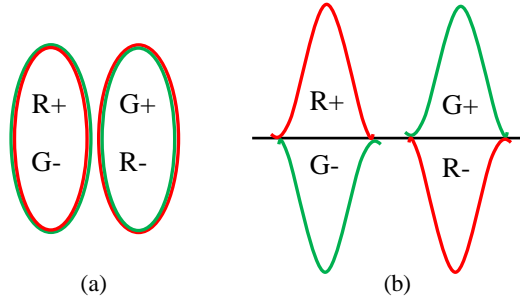


图 4.6 V1 神经元在输入权重相等时的双拮抗感受野。(a)双拮抗感受野示意图；(b) 双拮抗感受野响应示意图

本文从神经元视觉编码机制的底层出发，根据初级视皮层单细胞的树突接收特性，构建方位敏感的双拮抗神经网络，对视觉输入的颜色差异性和亮度差异性进行检测，最终再由单细胞形成神经元群体，并在群体感受野内对每个神经元的脉冲发放频率进行调制整合，以获取神经元群体对视觉激励的群体编码结果。

一方面，本文从全新角度解释了视皮层单细胞感受野的方位敏感特性。首先，单个神经元的计算模型是构建群体神经网络的基础单元，本文考虑了神经元模型的生物合理性和计算高效性，选取基于漏积分放电(LIF)模型来构建点阵神经网络，结合神经元的放射状树突结构与感受野特性，将单细胞感受野内所有树突接收的外界激励综合作为模型输入，同时以包含轮廓强度信息的脉冲发放频率对外界视觉激励进行编码。具体模型如式(4.7)所示。

$$\begin{cases} v = v_{\text{reset}}, \text{ref} = \text{ref} - 1, \text{if } \text{ref} \neq 0 \\ c_m \frac{dv}{dt} = -g_l v + \sum_{(x,y) \in \text{SRF}} \delta_{xy} I_{xy}, \text{if } \text{ref} = 0 \\ v = v_G, \text{ref} = \text{const}, \text{if } v > v_{\text{th}} \end{cases} \quad (4.7)$$

其中， $x, y$  分别是视觉刺激的横纵坐标， $v$  是神经元膜电压， $c_m$  是膜电容， $g_l$  是漏电导； $I_{xy}$  是与视觉刺激 $(x, y)$ 对应的神经元激励电流，本文将视皮层下颜色单拮抗感受野的动态响应结果进行归一化，作为初级视皮层神经元的视觉刺激； $\text{SRF}$  表示初级视皮层的单细胞感受野， $\delta_{xy}$  是感受野范围内的神经元树突接收极性，取值为 1、-1、0，分别表示兴奋性树突、抑制性树突以及零接收树突； $v_{\text{th}}$  为脉冲发放阈值， $v_{\text{reset}}$  为静态电势， $\text{ref}$  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。

其次，神经生理学相关研究对视皮层方向选择的解释为：若干外膝体神经元的感受野按特定方位进行排列分布，并且输入到同一个视皮层神经元，实现对平行于外膝体感受野排列方位的方向选择；而视皮层神经元通过放射状树突接收来自视皮层下外膝体的视觉响应。因此在外膝体到视皮层的信息投射过程中，外膝体感受野的排布规律和视皮层神经元树突的分布形式，分别从视通路的不同层次对方向选择特性做出解释，即视皮层下外膝体从宏观感受野的角度剖析了方向选择特性的形成机制，而视皮层神经元从微观细胞的角度解析了方向选择特性的响应机制。

于是，对应于第 3 章提出的多感受野朝向性组合，本文在不同敏感方位的单细胞感受野内，对神经元树突的接收极性进行不同方位的朝向性分布，并且为了与图 4.6 所示输入权重相等的双拮抗感受野保持一致，设定单细胞感受野内所有树突的接收权值都等于 1。以 0 度和 45 度敏感方位为例，在图 4.7 中给出感受野的树突极性分布示意，其具体计算方法如式(4.8)和式(4.9)所示。

$$\delta_{xy}^0 = \begin{cases} 1, & \text{if } (y - y_c) > 0 \\ 0, & \text{if } (y - y_c) = 0 \\ -1, & \text{if } (y - y_c) < 0 \end{cases} \quad (4.8)$$

$$\delta_{xy}^{45} = \begin{cases} 1, & \text{if } (y - x) > 0 \\ 0, & \text{if } (y - x) = 0 \\ -1, & \text{if } (y - x) < 0 \end{cases} \quad (4.9)$$

其中， $(x_c, y_c)$  是单细胞感受野 SRF 的中心坐标值， $(x, y) \in \text{SRF}$ 。同理可以得到其他方位的树突极性分布计算方法。



图 4.7 树突极性分布示意，其中红色、蓝色、灰色方块分别表示兴奋性树突、抑制性树突以及零接收树突。(a) 0 度方位的树突极性分布示意图；(b) 45 度方位的树突极性分布示意图

通过上述树突极性的分布，可以形成 0 度、45 度、90 度、135 度、180 度、225 度、270 度、315 度这 8 个方位的单细胞感受野，同时以敏感方位为界限将感受野分为兴奋性树突接收区域和抑制性树突区域，并且最优方位上的树突不接收视觉刺激。

另一方面，本文进一步提出了基于动态突触连接的神经元群体编码机制，首先将 4 个颜色单拮抗感受野的动态响应分别输入到构建的点阵神经网络，得到每个颜色拮抗通道在各个敏感方位的脉冲频率响应，然后通过最大值运算将 4 个拮抗通道在 8 个敏感方位上脉冲频

率响应映射为一幅灰度图像  $SinFire(x, y)$ ，如图 4.8 所示。

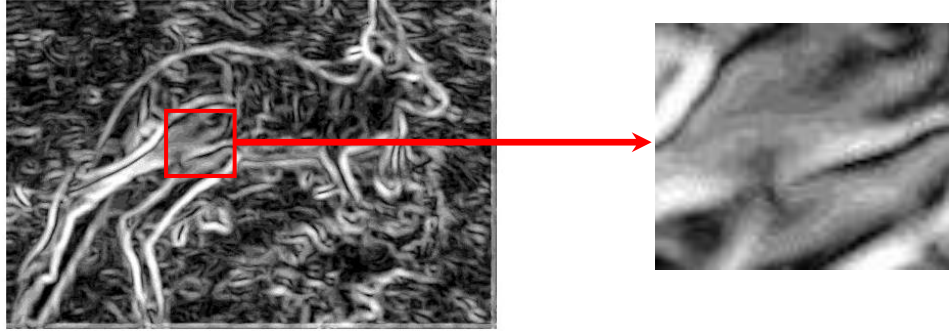


图 4.8 单细胞脉冲频率编码结果

显然，点阵神经网络有效捕捉了视觉刺激中的差异性信息，但是由于网络中的每个神经元都是以独立方式进行编码，而没有考虑神经元之间的相互连接作用，最终得到的轮廓编码结果呈现出非连续的层次断裂性。

而相关研究发现，仅仅依靠单个神经元的脉冲频率编码或者时间编码并不能完全表征神经元脉冲序列中蕴含的视觉信息，而神经元群体中多个神经元通过同步放电模式可以显著增加视觉信息的编码量。于是，本文在视皮层神经元的群体感受野范围内对单细胞的脉冲频率信息进行同步整合，最终形成具有神经元群体轮廓响应的互连网络，实现对轮廓信息的群体编码。

生理学研究发现绝大部分神经细胞群体的群体感受野面积小于群体内所有单细胞感受野面积的平均值，于是本文设定群体感受野的半径为单细胞感受野的五分之四；其次，神经元群体中每个单细胞之间的相互作用依赖于神经突触，突触是单细胞形成神经元群体的生理学结构基础，在神经元群体内部的信息交互和融合中发挥着重要作用，并且突触的连接作用是动态变化的，有助于提高神经系统对视觉信息的同步整合能力。本文在群体感受野范围内，分别从单细胞的空间距离和脉冲频率差异性两个方面给出动态突触的定义：突触连接作用随着单细胞之间距离的增大以及脉冲输出差异性的增大而减小，具体如式(4.10)所示。

$$synapse(x, y) = s_{const} \exp \left[ - \left( \frac{|P(x, y) - P(x_c, y_c)|^2}{2GRFr^2} + \frac{|F(x, y) - F(x_c, y_c)|^2}{2\Delta F^2} \right) \right] \quad (4.10)$$

其中， $synapse(x, y)$  是动态突触权值， $s_{const}$  是突触调节常数， $|P(x, y) - P(x_c, y_c)|$  表示群体感受野中周边神经元与中心神经元的距离， $|F(x, y) - F(x_c, y_c)|$  表示群体感受野周边神经元与中心神经元的脉冲频率差异， $GRFr$  是群体感受野的半径， $\Delta F$  是群体感受野内最大脉冲频率输出和最小脉冲频率输出的差值。

突触的作用方式为动态突触权值在群体感受野范围内与单细胞脉冲频率输出进行卷积运算，具体如式(4.11)所示。

$$GroFire_{ch}^{dir}(x, y) = synapse(x, y) * SinFire_{ch}^{dir}(x, y) \quad (4.11)$$

其中， $GroFire_{ch}^{dir}(x, y)$  是特定颜色拮抗通道在其特定敏感方位上的神经元群体编码结果，

$SinFire_{ch}^{dir}(x, y)$  是特定颜色拮抗通道在特定敏感方位的单细胞脉冲频率编码输出,  $dir = 1, 2, \dots, 8$  是敏感方位的索引值,  $ch \in \{RG, GR, BY, YB\}$  是颜色拮抗通道。

显然, 神经元群体通过动态突触的连接作用, 实现了单细胞脉冲频率输出的同步和调制, 在保留主体轮廓信息的同时, 弥补了单细胞点阵神经元编码结果中的层次断裂性, 有效实现了神经元群体基于脉冲频率的同步编码。神经元群体编码结果如图 4.9 所示。

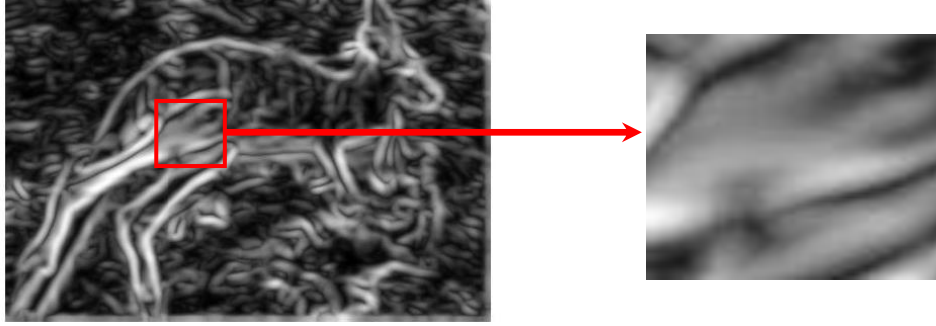


图 4.9 神经元群体脉冲频率编码结果

### 4.3 轮廓检测系统

本文轮廓检测的算法流程如图 4.10 所示, 部分环节以 R-G 颜色拮抗通道为例, 也同样适用于 B-Y 颜色拮抗通道。具体检测过程如下所述。

(1) 根据视网膜锥细胞的颜色敏感特性, 将输入彩色图像  $I_c(x, y)$  ( $x = 1, 2, \dots, m; y = 1, 2, \dots, n$ ) 分离为  $R(x, y)$ 、 $G(x, y)$ 、 $B(x, y)$ 、 $Y(x, y)$  四个颜色分量, 其中  $Y(x, y) = (R(x, y) + G(x, y)) / 2$ 。同时按式(4.1)计算锥细胞对亮度通道的分离结果  $L(x, y)$ 。

(2) 根据式(4.2)计算颜色单拮抗感受野的响应值, 并且按式(4.4)、(4.6)描述的单拮抗感受野动态调节方法, 对式(4.3)所示的高斯权值进行局部修正, 实时作用于颜色单拮抗感受野响应值的计算过程, 得到视皮层下颜色单拮抗感受野的动态响应结果  $SO_{RG}(x, y)$ 、 $SO_{GR}(x, y)$ 、 $SO_{BY}(x, y)$ 、 $SO_{YB}(x, y)$ 。

(3) 基于视皮层单细胞感受野的方向选择特性响应机制, 按式(4.8)、(4.9)以及其他树突极性分布模型, 每个树突赋予相同的接收权值, 形成 8 个不同敏感方位单细胞感受野的树突接收权值, 并以树突极性标记为  $\delta_{xy}^\theta$ ,  $\theta = 0, 45, 90, 135, 180, 225, 270, 315$ 。

(4) 根据原始图像  $I_c(x, y)$  的大小构建相同尺寸的点阵神经网络  $SN(x, y)$ , 每个神经元在其单细胞感受野 SRF( $11 \times 11$ ) 内通过放射状树突接收上一级神经元的视觉激励。其中, 每个神经元单细胞采用式(4.7)所示的 LIF 神经元模型, 分别将步骤(3)得到的 8 个单细胞感受野中的树突接收权值代入模型, 并将步骤(2)中 4 个颜色单拮抗感受野的动态响应结果进行归一化, 以激励电流的形式输入到点阵神经网络。

(5) 在一个响应周期内, 计算  $SN(x, y)$  每个神经元的脉冲发放频率, 得到 4 个拮抗通道在 8 个敏感方位上的单细胞脉冲频率输出  $SinFire_{ch}^{dir}(x, y)$ , 方位索引  $dir = 1, 2, \dots, 8$  与前文  $\theta = 0, 45, 90, 135, 180, 225, 270, 315$  相对应;  $ch \in \{RG, GR, BY, YB\}$ 。

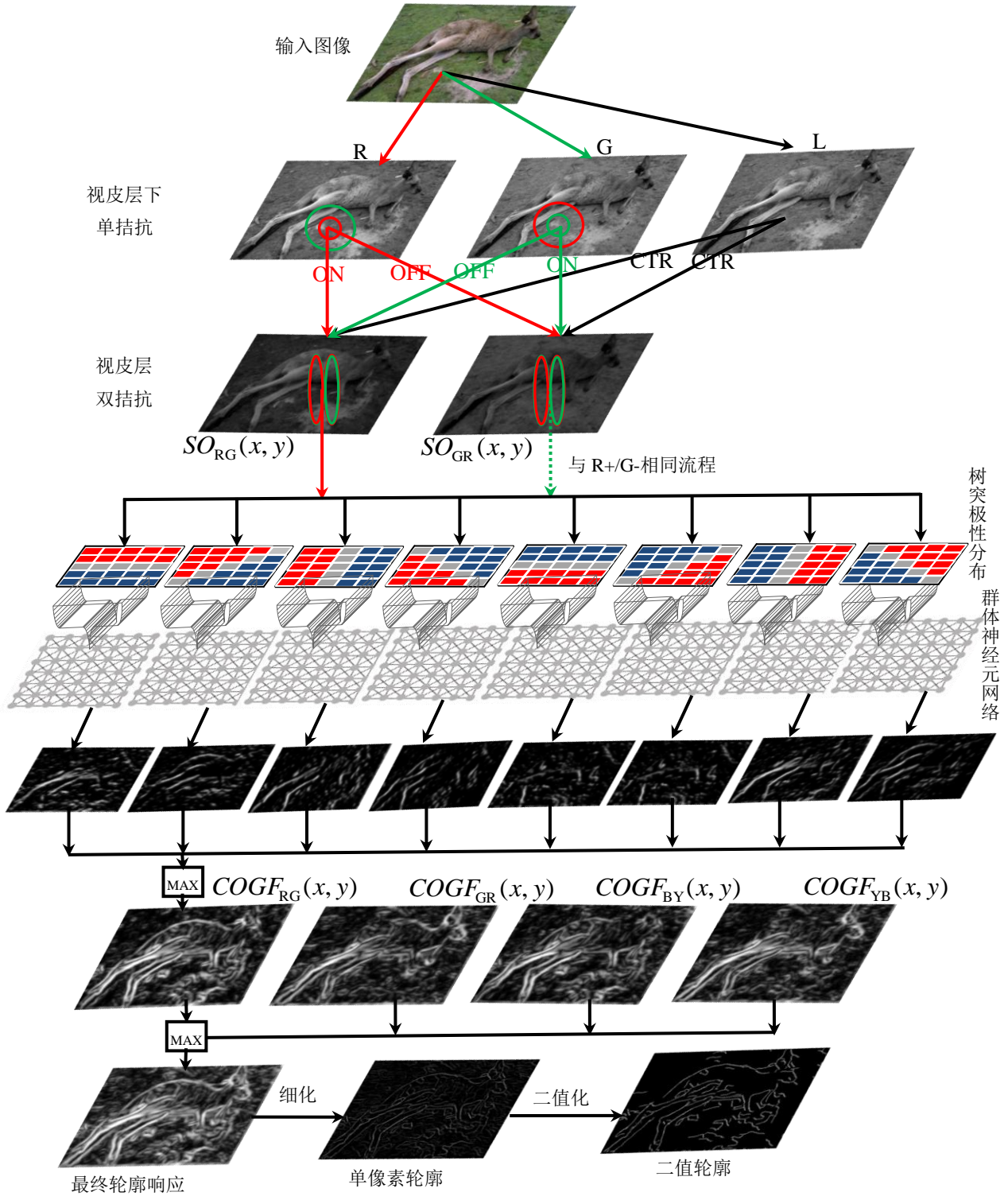


图 4.10 轮廓检测算法流程图

(6)再次构造神经元的群体感受野  $GRF(9 \times 9)$ ，并在群体感受野范围内利用式(4.10)所描述的动态突触作用，将点阵神经元中的单细胞相互连接，形成神经元群体。然后按式(4.11)将点阵神经元单细胞的脉冲频率输出  $SinFire_{ch}^{dir}(x, y)$  融合为互连神经元的群体编码结果  $GroFire_{ch}^{dir}(x, y)$ 。

(7)利用式(4.12)选取各个敏感方位群体编码结果的最大值作为该通道下的最终轮廓响应



$COGF_{ch}(x, y)$ ，利用式(4.13)选取各个颜色拮抗通道的最大值作为最终轮廓响应  $COGF(x, y)$ 。

$$COGF_{ch}(x, y) = \max\{GroFire_{ch}^{dir}(x, y) | dir = 1, \dots, 8\} \quad (4.12)$$

$$COGF(x, y) = \max\{GroFire_{ch}^{dir}(x, y) | ch \in \{RG, GR, BY, YB\}\} \quad (4.13)$$

(8)根式(3.17)~(3.18)，经过非极大值抑制和滞后阈值处理得到二值化的单像素轮廓结果。

#### 4.4 实验结果

本文以加利福尼亚大学伯克利分校提供的图像分割数据库(Berkley Segmentation Data Set, BSDS500)<sup>[59]</sup>为实验对象，图库中每一张图像具备若干张标准轮廓图(若干人工绘制结果)。因此，将本文算法的轮廓检测结果同每个人工绘制的标准轮廓进行比对，以多次比对结果的平均值来评估检测结果与真实轮廓的贴合程度，最终作为算法有效性的评价指标。

为了证明本文算法的优越性，选取目前主流的几种基于视觉机制的轮廓检测方法 with 本文算法结果进行对比分析。首先选取直接模拟初级视皮层方向选择特性的二维高斯导函数模型(DG)作为对比方法，体现本文方法神经网络具备的电生理学基础优势；其次，选取文献[26]提出的基于外膝体感受野排列组合的轮廓检测模型(CORF)，证明本文方法中视皮层下颜色动态拮抗特性对轮廓提取的有用性；最后为了体现本文方法的神经网络及其群体编码效应，选择文献[63]中考虑了颜色拮抗的二维高斯导函数模型(CODG)进行对比。由于四种方法的参数变化均涉及到高斯函数的标准差参数 $\sigma$ ，对于DG方法与CODG方法而言 $\sigma$ 代表了初级视皮层感受野尺寸；对CORF方法和本文方法而言 $\sigma$ 表示DoG感受野模型外周区的尺寸。于是，对DG和CORF方法选定 $\sigma=[3.0:0.5:5.0]$ 共5组参数，对CODG和本文方法选定 $\sigma=[1.0:0.5:3.0]$ 共5组参数，同时在滞后阈值处理中，取上限阈值 $t=[0.1:0.1:1.0]$ 共10组参数，并且设定下限阈值为 $0.5t$ 。因此每种方法都可得到50组测试参数。对每种轮廓检测算法均选择在其最佳参数条件下的轮廓检测结果进行对比，部分结果如图4.11~4.16所示。

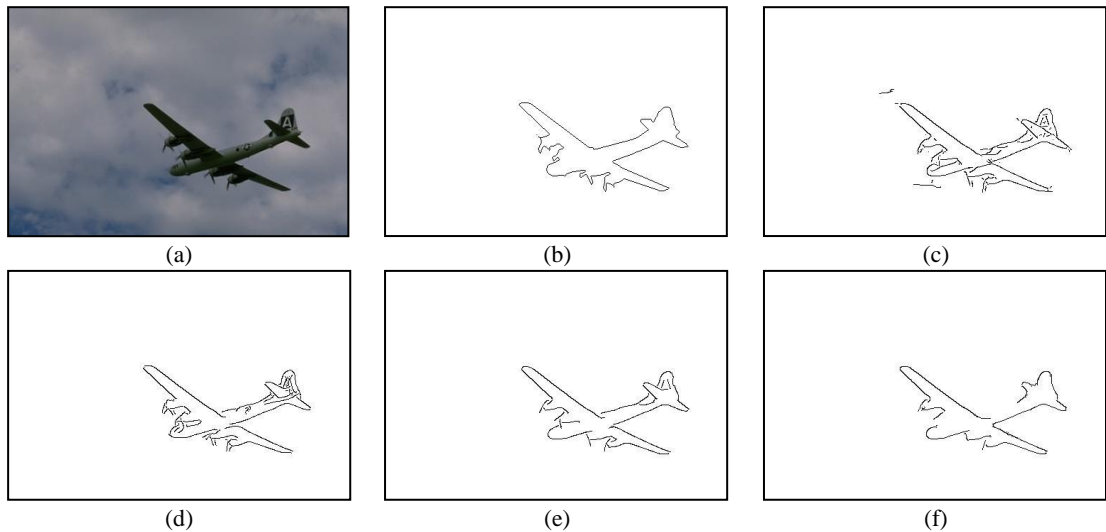


图 4.11 图 3096 处理结果。(a)原图；(b)标准轮廓(其一)；(c)DG 方法；(d)CORF 方法；(e)CODG 方法；(f)本文方法

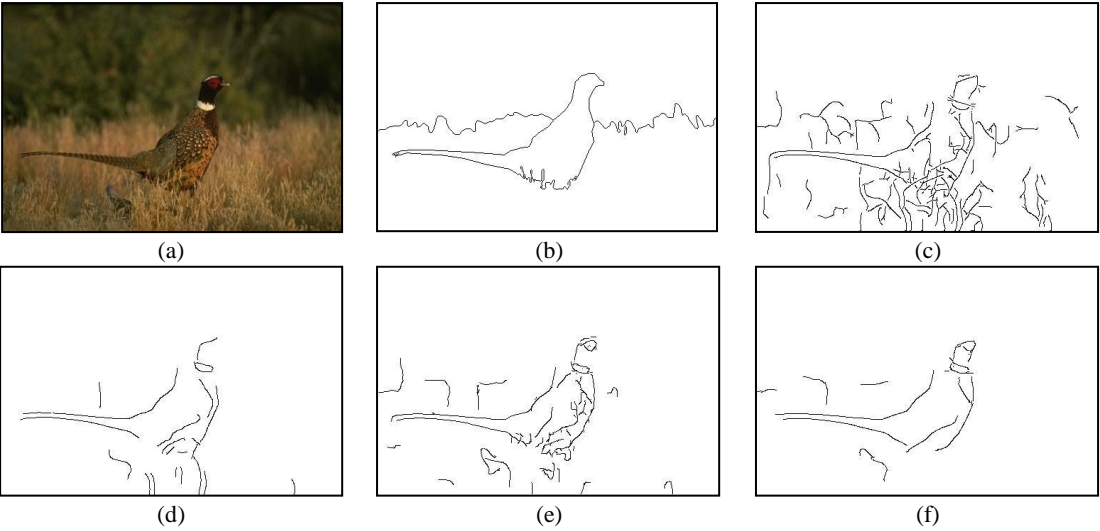


图 4.12 图 43074 处理结果。(a)原图；(b)标准轮廓(其一)；(c)DG 方法；  
(d)CORF 方法；(e)CODG 方法；(f)本文方法

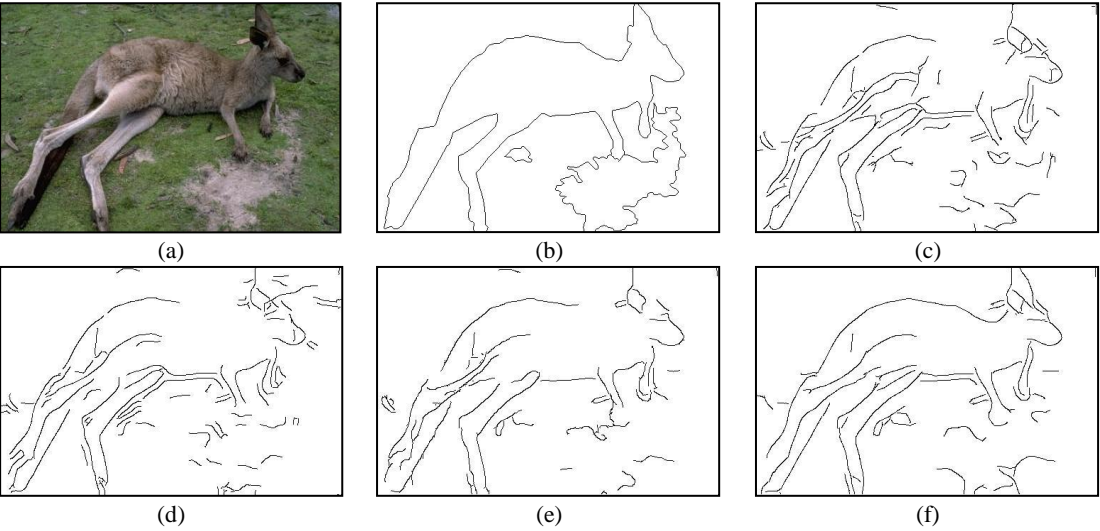


图 4.13 图 69020 处理结果。(a)原图；(b)标准轮廓(其一)；(c)DG 方法；  
(d)CORF 方法；(e)CODG 方法；(f)本文方法

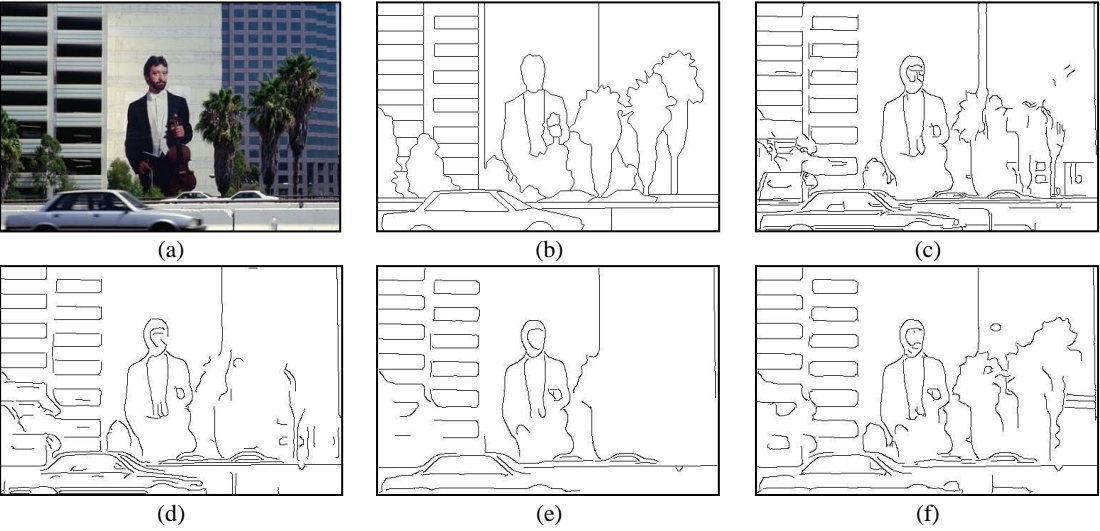


图 4.14 图 119082 处理结果。(a)原图；(b)标准轮廓(其一)；(c)DG 方法；  
(d)CORF 方法；(e)CODG 方法；(f)本文方法

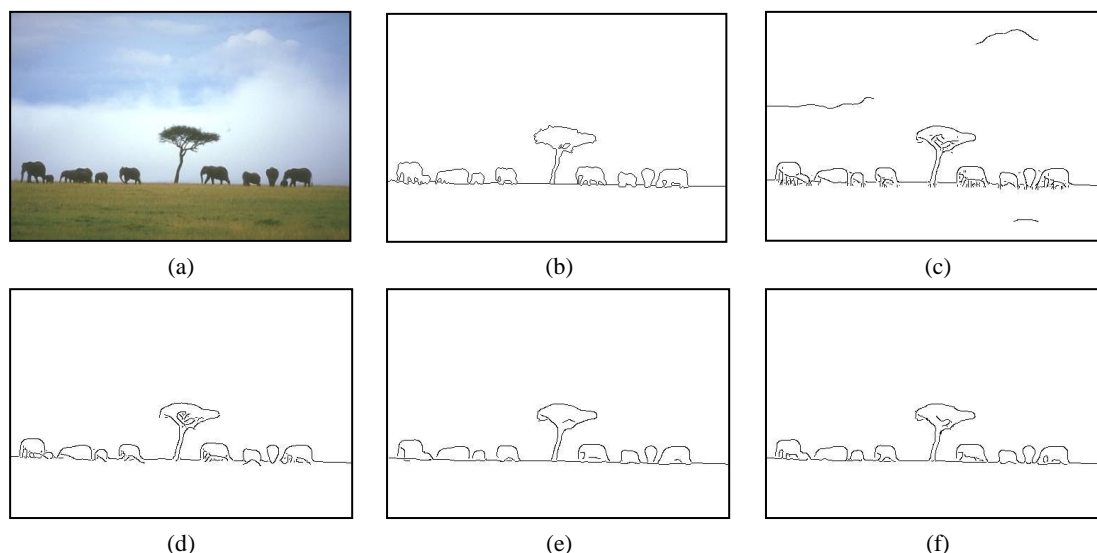


图 4.15 图 253036 处理结果。(a)原图；(b)标准轮廓(其一)；(c)DG 方法；  
(d)CORF 方法；(e)CODG 方法；(f)本文方法

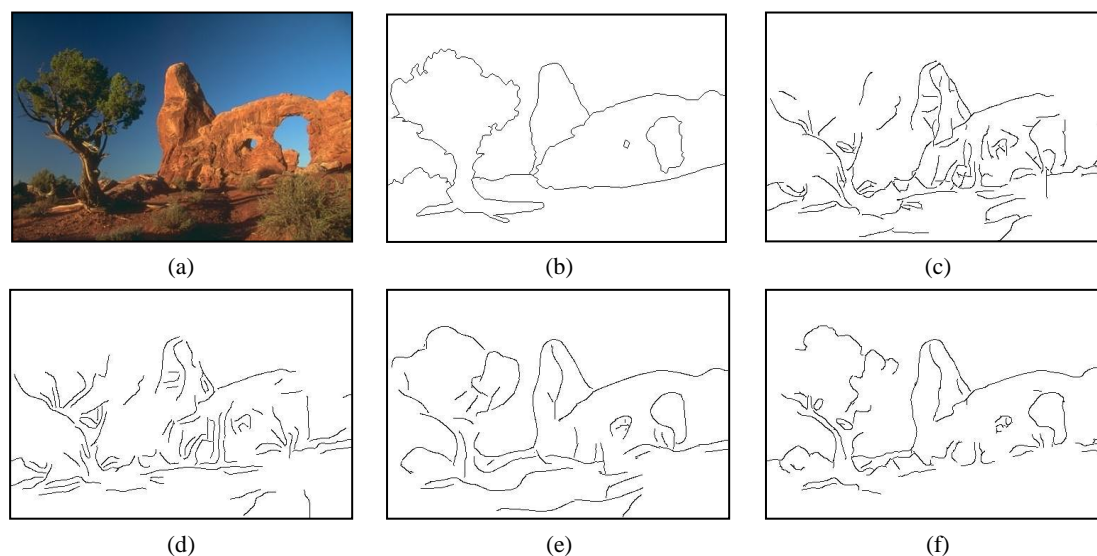


图 4.16 图 295087 处理结果。(a)原图；(b)标准轮廓(其一)；(c)DG 方法；  
(d)CORF 方法；(e)CODG 方法；(f)本文方法

通过观察图 4.11~4.16，可以发现本文方法能够检测出主体目标的基本轮廓，例如对于图 4.12 所示的具有复杂纹理背景的自然场景图以及图 4.13 所示的主体目标拥有完整轮廓的情形，本文方法检测的目标轮廓更加完整，图像背景更加纯净；而 DG 方法直接利用高斯导函数和阈值处理进行轮廓检测，无法有效区分主体目标轮廓和背景纹理边缘，同时阈值化处理难以有效维持误检率和漏检率的相对平衡，例如在图 4.12 的场景下出现过度检测导致纹理边缘偏多，而图 4.13 中为了更好的抑制纹理又导致了主体轮廓的局部丢失；CORF 方法虽然考虑了方向选择特性的组合形成机制，但是由于没有考虑视通路前后环节之间的相互作用，往往导致过度追求画面纯净度而导致主体轮廓不完整，如图 4.12 和图 4.16 所示，其算法总体性能与 DG 方法相当；相比而言，本文方法和 CODG 方法均考虑了颜色信息对彩色图像轮廓提取的关键作用，两种方法的检测效果基本上要优于 DG 和 CORF，但是 CODG 方法并没有考虑到

神经元群体编码对轮廓信息的电生理学基础作用，其对主体轮廓的认定存在不准确的情况，如图 4.12 所示的场景过多选择主体轮廓而忽略了背景纹理的干扰，相反在图 4.14 中却忽视了主体轮廓的基本组成部分，导致轮廓信息丢失，存在不稳定的现象；相对而言，由于本文方法不仅考虑了神经元的颜色拮抗特性，而且充分利用了方向选择特性的响应机制以及神经元群体编码效应。一方面，在背景单一主体轮廓清晰的场景下，本文方法可以尽可能有效提取目标的主要轮廓，同时使得画面保持相当高的纯净度，如图 4.11、4.15、4.16 所示；另一方面，针对背景复杂、纹理丰富的场景，本文方法有着更加明显的优势，不仅可以保留完整的主体目标轮廓，而且对背景纹理边缘的控制也处于较高水准，从图 4.12~4.14 可以看出这一点。

## 4.5 讨论与分析

为客观地对检测结果进行评价，同样采用文献[18]中的定量指标计算公式，对本文方法和对比方法做定量分析。因此，根据式(3.19)~(3.21)对图 4.11~4.16 的轮廓检测结果进行定量计算。如前所述，由于 BSDS500 图库中存在多张由不同人手工绘制的标准轮廓图，于是将每种方法的轮廓检测结果与多张标准轮廓进行定量计算，取与多张标准轮廓图比对的指标平均值来评估检测结果与真实轮廓的贴合程度，更加合理地避免了过多主观因素的干扰。图 4.11~4.16 的定量计算结果如表 4.1 所示。其中， $e_{FPa}$ 、 $e_{FNs}$ 、 $Pa$  分别表示检测结果与多张标准轮廓图定量计算结果的平均误检率、平均漏检率和平均  $P$  指标。

首先，DG 方法总体上拥有相对较小的漏检率，能够尽可能多地保留主体目标轮廓，但是付出的代价是无法有效去除背景纹理的干扰，其误检率往往达到数倍于其他方法，无疑拉低了算法的总体性能，这一特点在图 4.12(43074)和图 4.15(253036)表现得尤为明显，两者的误检率分别达到了 2.39 和 0.32；其次，CORF 方法在面对背景纹理单一、主体轮廓清晰的场景时相比 DG 具有更低的误检率，同时漏检率也处在较高水准，如图 4.11(3096)和图 4.15(253036)所示，但对于拥有复杂纹理的场景，如图 4.12(43074)、图 4.13(69020)和图 4.14(119082)所示，CORF 方法的漏检率过高，分别达到了 0.57、0.37 和 0.40，导致主体轮廓的断裂和不连续；而 CODG 方法总体上要优于 DG 和 CORF 方法，这可能是因为 CODG 方法考虑了神经元的颜色拮抗特性，其不足在于算法性能的不稳定性，一方面图 4.12(43074)和图 4.16(295087)中的漏检率分别为 0.42 和 0.38，要优于本文方法的 0.53 和 0.40，但是其误检率分别达到了 0.93 和 0.99，显著大于本方法的 0.39 和 0.85，导致其总体结果分别为 0.40 和 0.39 不如本文方法的 0.42 和 0.41，另一方面如图 4.13(69020)所示，其误检率为 0.85 显然要优于本文方法的 0.90，但是其漏检率为 0.34 不如本文方法的 0.27，总体性能 0.44 略低于本文方法的 0.46，综上可见 CODG 方法对于复杂背景纹理的场景难以有效定位真实轮廓以及主体目标轮廓的不连续性，造成算法性能不稳定的问题，这可能是缺少神经元脉冲群放电编码的辅助分析作用造成的；最后，本文提出的方法相比 CODG 方法在复杂场景下具有更加准确和稳定的轮廓检测效果，如前文对图 4.12(43074)和图 4.13(69020)的分析结果所示，另一方面，本文方法相对于 CODG 方法的优势在于对简单轮廓的提取结果具有更高的画面纯净度，如图



表 4.1 图 4.11~4.16 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             |                   |             |
|--------|------|----------|------|------------------|-------------------|-------------|
|        |      | $\sigma$ | $t$  | $e_{\text{FPa}}$ | $e_{\text{FN a}}$ | $Pa$        |
| 3096   | DG   | 5.00     | 0.10 | 0.49             | 0.26              | 0.55        |
|        | CORF | 3.00     | 0.20 | 0.30             | 0.23              | 0.65        |
|        | CODG | 1.00     | 0.50 | 0.20             | 0.30              | 0.63        |
|        | 本方法  | 1.50     | 0.60 | 0.05             | 0.36              | <b>0.66</b> |
| 43074  | DG   | 5.00     | 0.10 | 2.39             | 0.36              | 0.26        |
|        | CORF | 5.00     | 0.30 | 0.68             | 0.57              | 0.35        |
|        | CODG | 1.50     | 0.30 | 0.93             | 0.42              | 0.40        |
|        | 本方法  | 1.50     | 0.40 | 0.39             | 0.53              | <b>0.42</b> |
| 69020  | DG   | 5.00     | 0.10 | 1.56             | 0.36              | 0.33        |
|        | CORF | 4.50     | 0.20 | 1.51             | 0.37              | 0.34        |
|        | CODG | 1.50     | 0.40 | 0.85             | 0.34              | 0.44        |
|        | 本方法  | 1.50     | 0.40 | 0.90             | 0.27              | <b>0.46</b> |
| 119082 | DG   | 3.00     | 0.20 | 2.0              | 0.32              | 0.32        |
|        | CORF | 3.50     | 0.30 | 1.87             | 0.40              | 0.31        |
|        | CODG | 2.00     | 0.40 | 1.42             | 0.35              | 0.36        |
|        | 本方法  | 1.50     | 0.40 | 1.39             | 0.33              | <b>0.37</b> |
| 253036 | DG   | 4.50     | 0.10 | 0.32             | 0.08              | 0.71        |
|        | CORF | 3.50     | 0.20 | 0.09             | 0.11              | 0.82        |
|        | CODG | 1.00     | 0.60 | 0.03             | 0.15              | 0.83        |
|        | 本方法  | 1.00     | 0.50 | 0.03             | 0.10              | <b>0.89</b> |
| 295087 | DG   | 5.00     | 0.10 | 2.04             | 0.50              | 0.25        |
|        | CORF | 5.00     | 0.40 | 2.24             | 0.55              | 0.23        |
|        | CODG | 2.00     | 0.30 | 0.99             | 0.38              | 0.39        |
|        | 本方法  | 1.00     | 0.40 | 0.85             | 0.40              | <b>0.41</b> |

4.11(3096)和图 4.16(295087)所示, 本文的误检率分别为 0.05 和 0.85, 要低于 CODG 方法的 0.20 和 0.99, 而本文方法的漏检率为 0.36 和 0.40 也只是稍微高于 CODG 方法的 0.30 和 0.38, 总体性能上优于 CODG。另外计算图库验证集全部图像检测结果的平均  $P$  值, DG、CORF、CODG 分别为 0.46、0.48、0.53, 而本方法为 0.55, 总体上更具优势。

综上所述, 表 4.1 所示的定量计算结果与图 4.11~4.16 所示的轮廓检测结果基本保持一致。由表中  $P$  值平均结果  $Pa$  可以看出本文方法的整体检测效果优于对比方法。

相对于其他基于颜色拮抗或者视觉感知机制的轮廓检测方法。本文所做工作及其创新性如下: (1)提出了一种基于神经元颜色拮抗与群体编码的轮廓检测新方法。研究了视皮层下神经元颜色单拮抗感受野的拮抗机制与动态调节特性, 以神经元树突极性分布为基础的初级视皮层方向选择特性的响应机制, 以及基于动态突触的双拮抗神经网络群体编码特性, 突出了视觉感知机制在信息处理中的应用研究。(2)实现了一种融合视通路颜色拮抗特性与群体编码特性的彩色图像轮廓检测具体应用。根据视皮层下神经元的颜色单拮抗动态调节机制, 对轮廓信息进行早期分离; 同时考虑初级视皮层双拮抗神经元的方向选择特性, 利用其树突极

性分布解释了方向选择的响应机制；最后根据神经元群体的动态突触连接，在单细胞脉冲频率编码的基础上实现轮廓信息的群体编码；经过非极大值抑制和滞后阈值处理，得到单像素轮廓响应。(3)为如何利用视觉感知机制如颜色拮抗、方向选择、群体编码等实现更高层次的图像处理任务奠定了基础。

## 4.6 小结

本文引入了视皮层下神经元的颜色拮抗特性、基于树突极性分布的方向选择响应机制以及初级视皮层双拮抗神经元网络的群体编码机制，提出了一种基于神经元颜色拮抗与群体编码的轮廓检测方法，通过与同样模拟视觉机制的 DG、CORF、以及 CODG 等方法的定性与定量比较，表明本文提出的方法在图像轮廓检测上具有更加优越的性能。以多种场景的自然图像为例，本文方法的检测结果具有更加完整连续的主体目标轮廓，同时在一定程度上有效剔除了背景纹理边缘。

## 第 5 章 基于双层神经网络脉冲调控编码的轮廓检测

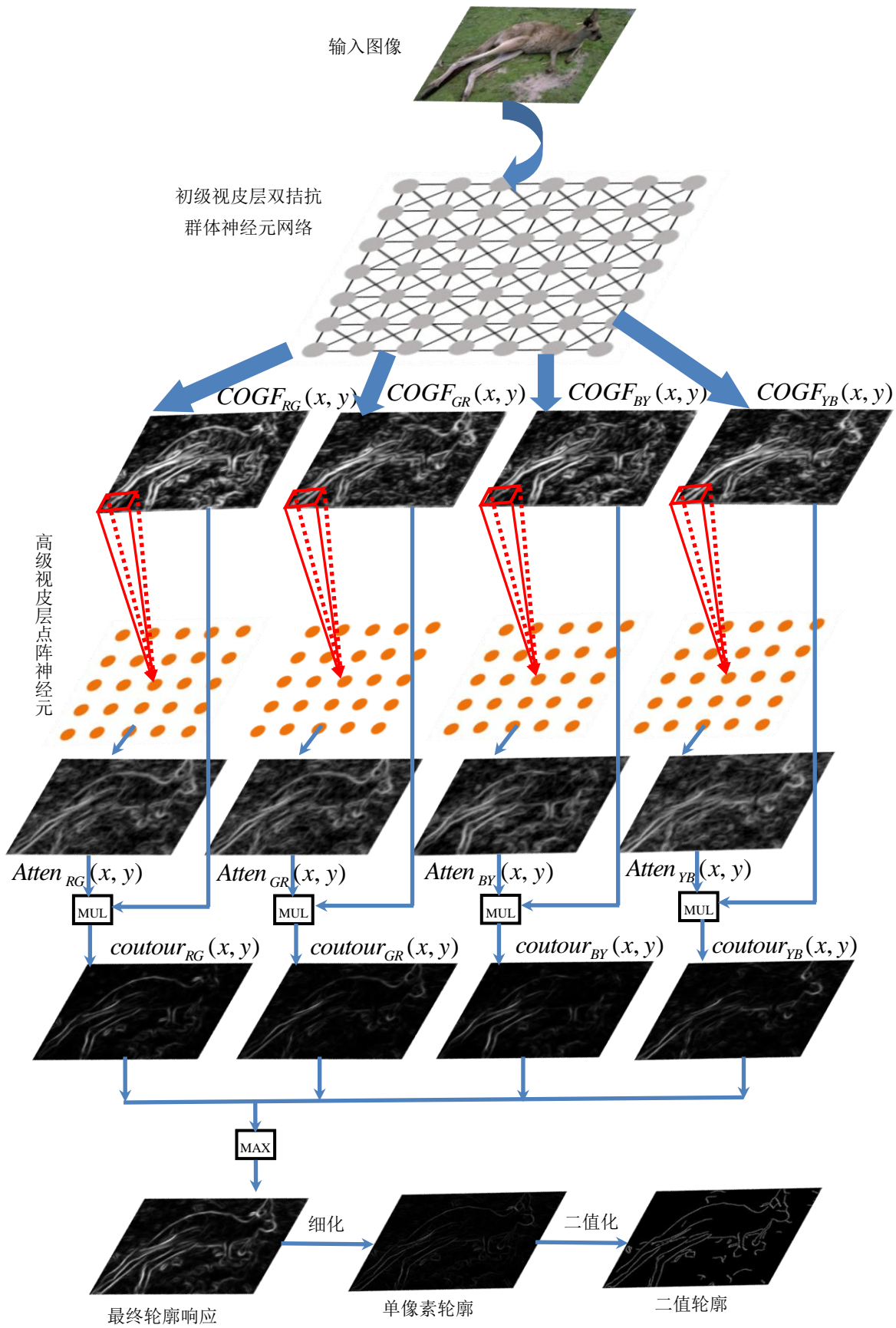
第 4 章提出的基于神经元颜色拮抗与群体编码的图像轮廓检测方法考虑了视通路不同层次神经元的颜色拮抗特性、方向选择特性的响应机制以及基于动态突触的神经元群体编码,相比其他方法获得的轮廓结果更加契合人类视觉系统,但是该方法仅仅在最后阶段依靠阈值处理来区分主体轮廓与纹理边缘,在轮廓响应阶段尚不具备有效剔除背景纹理而同时突显主体轮廓的智能性。可见,单凭早期视觉通路的感知机制无法进一步区分真实轮廓与虚假边缘,在视觉感知系统中势必存在更高层次的视觉皮层参与轮廓信息编码。因此本文提出基于双层神经网络脉冲调控编码的轮廓检测技术,在前文颜色拮抗神经网络的基础上,针对每个颜色拮抗通道再一次构建高级视皮层点阵神经网络,并且根据空间稀疏性实现高级视皮层神经元对注意性信息的自适应脉冲频率编码,最后将四个颜色拮抗通道的选择注意图像反馈到各自的初级视皮层轮廓响应,进一步融合全部颜色拮抗通道的调控结果得到最终轮廓响应。

### 5.1 概述

为了进一步发挥高级视觉皮层神经元在轮廓检测中的作用,本文从高级视觉皮层的感知机制出发提出了一种基于双层神经网络脉冲调控编码的轮廓检测方法,按照第 4 章构建具备方向选择特性的双拮抗神经网络,对视皮层下单拮抗响应进行群体编码,然后针对四个颜色拮抗通道的群体编码结果再次构建高级视皮层点阵神经网络,并且每个神经元的脉冲发放阈值随着局部刺激的空间稀疏性的改变而发生变化,以自适应脉冲编码的方式响应视觉刺激中的注意性信息,然后利用编码得到的选择注意图像对每个颜色拮抗通道的初级视皮层响应进行反馈调控,最后融合全部颜色拮抗通道的调控编码结果,得到最终的图像轮廓。结果显示,本文方法在轮廓响应阶段就能够抑制大量无用的纹理边缘,同时一定程度上能够增强主体目标的真实轮廓,相比其他方法有着显著优势。

### 5.2 轮廓检测模型

在颜色拮抗、方向选择和群体编码等早期视觉感知机制的基础上,本文着重构建第二层高级视皮层的点阵神经网络,利用神经元对局部空间稀疏性的自适应调节机制获取注意性信息的脉冲编码结果,并经过灰度逆映射得到选择注意图像,最后利用高级视皮层的反馈调控机制将选择注意图像作用于初级视皮层双拮抗神经网络的轮廓响应,充分发挥高级视皮层轮廓响应在轮廓信息精炼过程中的关键作用。本章轮廓检测方法的具体框架如图 5.1 所示,其中,由于第一层初级视皮层的双拮抗神经网络与第 4 章保持一致,图中仅仅给出了简单的黑箱示意。



### 5.2.1 高级视皮层点阵神经元

在早期视觉通路中，视觉系统对外界视觉刺激的编码起始于视网膜，并终止于初级视皮层(V1)<sup>[60]</sup>，在这样的信息处理流程中，视觉响应往往整体上包含了外界刺激的大部分信息，而通常缺乏对显著性视觉信息的关键描述，形成真正的视觉意识还需要更多高级层次的参与。

神经信息编码指出神经元通过脉冲发放的形式将视觉刺激按特定的规则与方法传递给视觉通路的各个层次，直到大脑形成意识感知<sup>[61]</sup>。可见，高级视皮层神经元同样具有脉冲编码特性，能够将初级视皮层的视觉响应进行精细汇聚，删除冗余因素，保留关键信息。高级视皮层的神经元是视皮层信息调控编码的生理结构基础，而神经元通常以集群方式编码视觉信息，有助于进一步形成后续的行为认知、学习记忆等<sup>[62]</sup>。本文以兼具生物合理性与计算高效性的 LIF 模型组建点阵神经元，同时引入绝对不应期，以一个响应周期内每个神经元的脉冲发放次数表征初级皮层响应中的视觉注意信息。不同于视网膜与初级视皮层的神经元，高级皮层神经元表现出更高的智能性，即面对不同强度的视觉刺激时，其脉冲发放阈值和静态电压随着特定需求而发生变化。因此，高级视皮层神经元的 LIF 模型如式(5.1)所示。

$$\begin{cases} c_m \frac{dv}{dt} = -g_l v + I(x, y), \text{if } \text{ref} = 0 \\ v = v_{\text{reset}}(x, y), \text{ref} = \text{ref} - 1, \text{if } \text{ref} \neq 0 \\ v = v_G, \text{ref} = \text{const}, \text{if } v > v_{\text{th}}(x, y) \end{cases} \quad (5.1)$$

其中， $v$  表示神经元细胞膜电压，随着时间  $t$  发生变化， $c_m$  是膜电容， $g_l$  是漏电导， $I(x, y)$  是神经元的电流激励， $v_{\text{th}}(x, y)$  与  $v_{\text{reset}}(x, y)$  分别是神经元的脉冲发放阈值和静态电压，随着不同空间位置  $(x, y)$  处的视觉刺激发生改变； $\text{ref}$  为绝对不应期， $v_G$  是脉冲发放峰值。针对每一个空间位置  $(x, y)$ ，都有一个与之对应的神经元，当其膜电压  $v$  大于对应阈值  $v_{\text{th}}(x, y)$  时，神经元脉冲发放达到  $v_G$ ，同时瞬间重置为对应的静态电压  $v_{\text{reset}}(x, y)$ ，进入绝对不应期之后，神经元不再进行脉冲发放，直至不应期结束。

同时，第 3 章提到外膝体细胞感受野相比视网膜神经元具有更大的空间作用范围<sup>[63]</sup>，而在初级视皮层到高级视皮层的信息投射过程中，这种感受野的层次变化更为明显。另一方面，随着视通路的层次变化，神经元数目不断减少，高级皮层一个神经元必须能够处理多个初级皮层神经元的综合响应。于是，本文以相对较小规模的点阵高级神经元群体对初级视皮层的每个颜色拮抗通道的轮廓响应  $COGF_{\text{ch}}(x, y)$  进行编码，其中初级皮层神经元的响应以局部区块的形式投射到单个高级皮层神经元，具体的实现方式与式(3.4)相同，即以初级皮层轮廓响应  $COGF_{\text{ch}}(x, y)$  的一个目标点  $(x, y)$  为中心建立局部窗口，以窗口中的平均响应值作为高级皮层神经元的激励输入，每个颜色拮抗通道的局部平均响应进行归一化后记为  $I_{\text{ch}}(x, y)$  (即神经元的激励电流)，以上论述中  $\text{ch} \in \{\text{RG}, \text{GR}, \text{BY}, \text{YB}\}$ ，具体实现如图 5.2 所示。

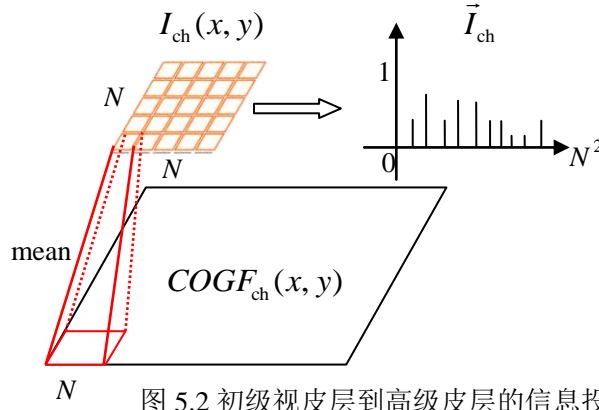


图 5.2 初级视皮层到高级皮层的信息投射示意

### 5.2.2 选择注意机制的脉冲频率编码

如 2.4 节所述，有学者通过自底而上的方法，将 2.4 节所述模型的显著性信息输入到空间稀疏的神经网络，以其脉冲频率编码输出作为选择注意的响应<sup>[52]</sup>，与真实视觉机制比较贴近。由此可见，视觉信息的空间稀疏性以及显著性是诱发选择注意机制响应的外界因素，因此本文利用局部空间的稀疏性定义，不断地修正神经元的脉冲响应阈值和静态电压，通过脉冲频率编码视觉刺激中的注意性信息。Alpert<sup>[63]</sup>等人认为主体目标显著性轮廓所在区域具备相当高的稀疏性，而背景纹理边缘区域是非稀疏的，并且基于局部梯度幅度直方图给出了稀疏性定义，如式(5.2)所示。

$$spar(x, y; \vec{h}) = \frac{1}{\sqrt{n-1}} \left( \sqrt{n} - \frac{\|\vec{h}(x, y)\|_1}{\|\vec{h}(x, y)\|_2} \right) \quad (5.2)$$

其中， $\vec{h}(x, y)$  表示以  $(x, y)$  为中心的局部区域处的梯度幅度直方图， $n$  是  $\vec{h}(x, y)$  的空间维数， $\|\vec{h}(x, y)\|_1$  和  $\|\vec{h}(x, y)\|_2$  分别表示  $\vec{h}(x, y)$  的 1 范数与 2 范数，即  $\|\vec{h}(x, y)\|_1 = \sum |h_i|$ ， $\|\vec{h}(x, y)\|_2 = \sqrt{\sum |h_i|^2}$ 。

一方面，由于基于直方图的分析将损失许多像素的空间分布信息，Yang 等人<sup>[21]</sup>在 Alpert 模型的基础上直接利用空间尺度变换和稀疏性调节因子实现对非稀疏纹理的抑制，并将调整后的稀疏性信息映射为显著性图像，得到选择注意机制的响应，具体如式(5.3)所示。其中， $\vec{H}(x, y)$  为前级轮廓响应经尺度变换后得到的局部幅值向量，向量中每个元素代表前级轮廓响应中一个局部区域的响应幅值， $\text{mean}(\vec{H}(x, y))$  表示取平均值， $\text{ScaleSpar}(x, y)$  表示尺度变换后的前级轮廓响应， $\min$  表示取较小值，反映了该模型能够在非稀疏区域抑制纹理，在其他区域维持前级轮廓响应不变。

$$\begin{cases} saliency(x, y) = \lambda(x, y) \cdot spar(x, y; \vec{H}) \\ \lambda(x, y) = \min(1, \frac{\text{ScaleSpar}(x, y)}{\text{mean}(\vec{H}(x, y))}) \end{cases} \quad (5.3)$$

另一方面, 尽管 Yang 的方法一定程度上抑制了背景纹理, 但是没有考虑到神经元脉冲在视觉注意性信息编码中的电生理学基础作用, 而且只能够对背景纹理进行抑制, 无法进一步增强主体目标轮廓。因此, 本文提出一种全新的选择注意机制响应模型: 结合式(5.1)提出的神经元模型, 根据初级皮层与高级皮层的感受野变换特性和神经元数目递减规律, 计算初级皮层每个颜色拮抗通道的局部平均响应作为高级皮层神经元的激励电流, 同时利用激励电流的局部稀疏性驱动高级皮层神经元对视觉注意性信息进行自适应脉冲编码, 从而得到选择注意图像。具体表现为: 在非稀疏或者稀疏性很小的纹理边缘区域, 高级皮层神经元反而具有较大的脉冲发放阈值, 对应的脉冲发放频率显著下降, 以抑制无用的纹理边缘; 在稀疏性较大的真实轮廓区域, 高级皮层神经元具有较小的脉冲发放阈值, 将产生更加强烈的脉冲发放, 以增加真实轮廓的强度。具体实现如式(5.4)所示:

$$\begin{cases} v_{th}(x, y) = v_{const}(1 - spar(x, y; \vec{I}_{ch})) \\ v_{reset}(x, y) = 0.2 \times v_{th}(x, y) \end{cases} \quad (5.4)$$

其中,  $v_{const}$  为阈值电压常数, 本文取值为 10;  $\vec{I}_{ch}$  为  $I_{ch}(x, y)$  在神经元感受野范围内形成的幅值向量, 而  $I_{ch}(x, y)$  是初级皮层局部响应平均值的归一化结果, 且  $ch \in \{RG, GR, BY, YB\}$  表示四个颜色拮抗通道, 具体如图 5.2 所示。

利用式(5.3)和(5.4)的计算模型, 将高级皮层神经元的脉冲发放频率映射成特定颜色拮抗通道的注意性图像, 需要注意的是高级皮层点阵神经元尺寸以及其自适应脉冲编码得到的注意图像尺寸都要小于原图, 此处利用双线性插值法将其放大为原始大小, 记为  $Atten_{ch}(x, y)$ ,  $ch \in \{RG, GR, BY, YB\}$ 。图 5.3 给出了 R+/G-通道单拮抗响应图像、R+/G-通道的初级皮层轮廓响应、Yang 方法获取的视觉显著性图像以及本文方法获取的选择注意图像, 显然本文提出的选择注意机制能够更加精确地定位主体轮廓区域与背景纹理区域。

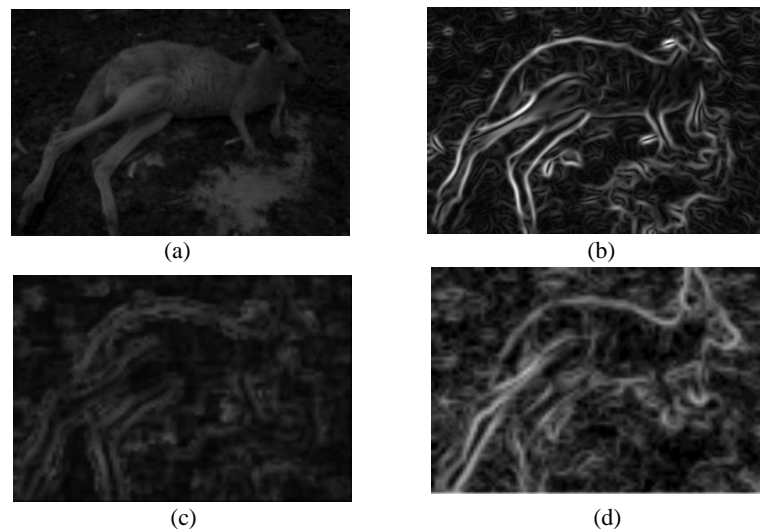


图 5.3 不同方法选择注意图像的对比。(a) R+/G-通道单拮抗响应图像; (b) R+/G-通道初级皮层轮廓响应; (c) Yang 方法的视觉显著性图像; (d) 本文方法的选择注意图像。

### 5.2.3 高级视皮层的反馈调控

早期视觉通路通常以前馈传递的基本方式对视觉信息进行加工处理，而视觉系统中的高级结构除了前馈通道还也存在着反馈调节通路，以一种自顶而下的全新调节机制控制初级结构的视觉信息编码结果。于是，本文将高级皮层点阵神经元获取的选择注意图像反馈投射到初级皮层，从而进一步减少冗余信息保留真实有用的主体轮廓信息，以便更高级神经系统作进一步的识别与感知。

故本文将每个颜色拮抗通道的选择注意图像  $Atten_{ch}(x, y)$  反馈至初级皮层，以调控系数的形式作用于对应通道的初级皮层轮廓编码结果  $COGF_{ch}(x, y)$ ，如式(5.5)所示，得到最终轮廓响应，其中  $ch \in \{RG, GR, BY, YB\}$ 。

$$contour_{ch}(x, y) = Atten_{ch}(x, y) \times COGF_{ch}(x, y) \quad (5.5)$$

最后按式(4.13)将四个颜色拮抗通道的轮廓响应进行融合，经过非极大值抑制与阈值处理后得到单像素轮廓响应。

## 5.3 实验结果

本文依然使用文献[59]中的 BSDS500 图像分割数据库为实验对象，该数据库包含训练集、测试集和验证集共 600 张自然场景图像，并且每幅图像对应 5 幅不同的手工绘制标准轮廓图。为了证明本文算法在抑制背景纹理，突显主体轮廓方面的智能性，针对性选取图库中背景纹理边缘丰富、主体轮廓难以突显的图像作为实验对象，部分图像如图 5.4 第一行所示，5 幅标准轮廓图的叠加结果如图 5.4 第二行所示。BSDS500 图像库被广泛用来评价彩色图像轮廓提取与图像分割的算法性能，图像库可以从加州大学伯克利分校的机器视觉研究组网站 (<https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/>) 下载。

在定性分析的时候，为体现本文算法的生物智能性优势，首先选择仅考虑初级视皮层方向选择特性的二维高斯导函数模型(DG)，以证明本文算法中初级视皮层颜色拮抗机制和高级皮层神经元调控编码的有效性；然后，为评价本文算法的纹理抑制效果，选择文献[21]中结合颜色拮抗与空间稀疏限制的方法(SCO)和文献[52]中基于显著性信息稀疏编码的 SCNN 方法对实验图像进行轮廓检测和纹理抑制，并且与本文算法进行比较。在实验中，除了本文方法，其他三种方法均选择二维高斯导函数模拟初级视皮层感受野，为了与本文初级视皮层模型保持一致，也同样设定 8 个敏感方位的高斯导函数，同时为减少过多参数之间相互干扰，设定感受野尺寸  $\sigma = 1.5$  (在本文方法中  $\sigma$  表示颜色单拮抗感受野模型中感受野外周区的尺寸)。另外，由于其他三种方法没有引入颜色拮抗的动态调节或者神经元的自适应编码，因此对 SCO 设置颜色拮抗通道权值  $w = [-0.1: -0.1: -1.0]$ ，SCNN 设置突触电流抑制系数  $\alpha = [0.6: 0.1: 1.5]$ ，而本文只需通过视觉刺激的变化来实时确定计算参数，至于 DG 方法没有考虑纹理抑制因此不涉及参数设置。为了与 BSDS500 图库中的二值标准轮廓图进行定量指标计算，本文方法和对比方法的轮廓响应结果需要经过非极大值抑制和滞后阈值处理，因此设定滞后阈值处理中



的高门限阈值  $t=[0.1:0.1:1.0]$ ，低门限阈值为  $0.5t$ 。于是，对于 DG 和本文方法只需设置 10 组测试数据，SCO 和 SCNN 方法各自需要设置 100 组测试数据。

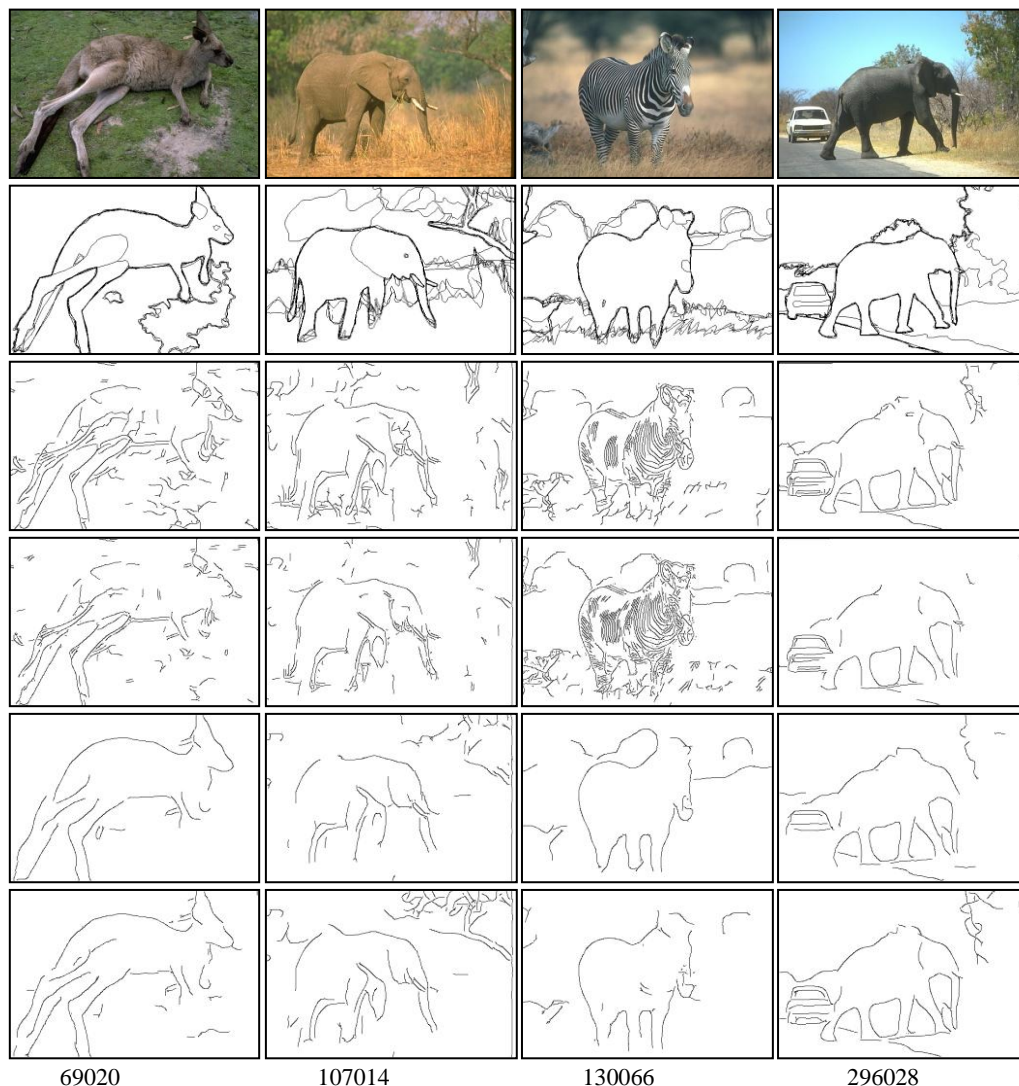


图 5.4 BSDS500 图库中部分图像检测结果

图 5.4 给出了本文方法以及对比方法的部分图像实验结果。从上到下依次是原始彩色图像、5 幅标准轮廓图叠加结果、DG 方法检测结果、SCNN 方法结果、SCO 方法检测结果以及本文方法结果。从总体效果上来看，考虑了背景纹理抑制的三种方法相对于只考虑初级视皮层方向选择的高斯导函数方法 DG，通常情况下都能对纹理边缘进行不同程度的抑制，从而增强了主体轮廓与背景纹理边缘的区分度。DG 方法是利用高斯导函数的方位敏感直接检测图像中的全部差异性信息，然后依靠滞后阈值方法在细化结果中选取较高强度的像素作为最终轮廓，其主要特点是能够很好地保证主体轮廓的连续性和完整性，而且面对背景单一、目标轮廓较为明显的场景时具有不错的效果，如图 296028 的检测结果显示，然而 DG 方法在面对大面积复杂背景纹理、主体目标不够突显的场景时，其轮廓辨别能力直线下降，将许多背景纹理与目标内部纹理误检为轮廓，如其余 3 幅图像所示。再看 SCNN 方法的检测结果，大

部分情况下相比 DG 方法，由于引入了对显著性信息的编码，能够在一定程度上减少非显著或非稀疏的背景纹理，但是该方法纹理抑制强度有限，依然存在许多短线状毛刺，也牺牲了一些主体轮廓信息，图 69020 等子图均有体现，而且该方法对主体目标内部纹理的抑制能力十分有限，并不能有效抑制如子图 130066 所示的斑马条纹，整体而言算法性能不够稳定。SCO 方法的检测效果相对而言较为优秀，四幅子图中的大部分背景纹理边缘得到了有效的抑制，并且主体目标的轮廓尽可能多地被保留下来，整体上具有很高的画面纯净度，唯一的不足是损失了主体目标的轮廓细节，每个子图基本都有呈现。综合而言，本文方法的检测结果中主体轮廓完整而平滑，具有较好的连续性，并且保留了主体目标的大部分轮廓细节，在纹理抑制上也有良好的效果，整体上更胜一筹。

## 5.4 讨论与分析

基于 BSDS500 彩色图像数据库，依然采用文献[18]中的算法有效性评价指标对本文算法以及对比算法进行定量分析，具体计算公式已由式(3.19)~(3.21)给出。对图 5.4 所示的图像及其检测结果进行定量指标计算，其中各算法检测结果基于不同标准轮廓图的平均误检率、平均漏检率以及综合指标  $P$  的平均值如表 5.1 所示。

表 5.1 BSDS 图库中部分图像轮廓检测结果对应的参数设置及性能指标

| 图像     | 算法   | 参数          |      | 性能指标             |                  |             |
|--------|------|-------------|------|------------------|------------------|-------------|
|        |      | $-w/\alpha$ | $t$  | $e_{\text{FPa}}$ | $e_{\text{FNs}}$ | $Pa$        |
| 69020  | DG   |             | 0.10 | 1.99             | 0.33             | 0.30        |
|        | SCNN | 0.80        | 0.10 | 1.67             | 0.43             | 0.31        |
|        | SCO  | 0.90        | 0.70 | 0.34             | 0.40             | 0.54        |
|        | 本方法  |             | 0.40 | 0.41             | 0.36             | <b>0.55</b> |
| 107014 | DG   |             | 0.10 | 2.02             | 0.47             | 0.27        |
|        | SCNN | 1.30        | 0.10 | 1.36             | 0.53             | 0.31        |
|        | SCO  | 0.60        | 0.20 | 0.98             | 0.55             | 0.33        |
|        | 本方法  |             | 0.30 | 0.93             | 0.42             | <b>0.40</b> |
| 130066 | DG   |             | 0.10 | 3.23             | 0.37             | 0.21        |
|        | SCNN | 0.60        | 0.40 | 4.08             | 0.33             | 0.19        |
|        | SCO  | 1.00        | 0.40 | 0.54             | 0.51             | 0.40        |
|        | 本方法  |             | 0.20 | 0.43             | 0.53             | <b>0.41</b> |
| 296028 | DG   |             | 0.10 | 0.67             | 0.24             | 0.52        |
|        | SCNN | 1.10        | 0.10 | 0.47             | 0.41             | 0.49        |
|        | SCO  | 0.10        | 0.40 | 0.30             | 0.39             | 0.52        |
|        | 本方法  |             | 0.30 | 0.45             | 0.30             | <b>0.54</b> |

查表可知，图 5.4 所示图像的定量指标计算结果与前文的定性分析基本吻合。大多数情况下 DG 方法的误检率与漏检率处于两个极端，较小的漏检率说明该方法检测出了大多数的主体轮廓，付出的代价是较高的误检率，所以结果中存在许多无用的背景纹理边缘，但是对于图 296028 这种背景平滑单一、主体轮廓显著的场景，该方法综合指标  $P$  的平均值达到了 0.52，具备相对优秀的水准。SCNN 方法相比 DG 方法总体效果有一定提升，基本表现为误检

率有所减小而漏检率有所增大,整体效果的改善程度有限,而且无法抑制主体目标内部的无用纹理,图 130066 完全展现了这一点,可能的原因是该方法没有考虑颜色拮抗的作用,同时对显著性信息的稀疏性定义不够准确。而 SCO 与本文方法的效果十分接近,两者在漏检率与误检率之间有着不同的侧重点,具体表示为 SCO 方法的误检率低于本文方法,而漏检率高于本文方法,主要原因是 SCO 只考虑纹理抑制,而本文方法在考虑纹理抑制的同时,在一定程度上对主体轮廓进行增强,尽可能防止纹理抑制过程中主体轮廓细节的丢失。另外计算图库验证集全部图像检测结果的平均  $P$  值, DG、SCNN、SCO 分别为 0.46、0.50、0.55,而本方法为 0.58,总体上可以看出本文方法具有更加优秀的轮廓检测效果。

与其他三种轮廓检测方法相比,尽管 SCNN 方法也考虑了初级视皮层神经元的感受野作用机制,并在其基础上建立了显著性信息的定义与神经元稀疏编码的表征方式,但是该方法没有考虑到颜色信息在轮廓编码中的关键作用,同时也忽略了神经元脉冲放电与视觉显著性信息、注意性信息的关联,从而算法整体浮动性较大,在轮廓结果中表现为对纹理边缘抑制不够充分,甚至失去纹理抑制效果;而本文方法不仅考虑了神经元的颜色拮抗机制,而且在初级视皮层建立了群体编码神经网络,是对早期视觉通路轮廓响应过程的完整性描述,同时还引入了高级皮层神经元对初级皮层轮廓响应的自适应脉冲调控机制,从底层的神经元放电活动对轮廓感知过程进行约束,能够有效区分出主体轮廓与纹理边缘。而 SCO 方法虽然也考虑了颜色拮抗机制,但是忽略了神经元电生理学编码基础,而且直接采用数学算子对稀疏性和显著性进行定义,在纹理抑制过程中缺乏一定的智能性,同时也不能弥补纹理抑制过程中局部轮廓细节的丢失;相反,本文方法在初级视皮层和高级视皮层都考虑了神经元脉冲放电对视觉感知机制形成的生理学解释,特别是在高级皮层中引入神经元对注意性信息的自适应调控编码,通过设置基于空间稀疏性表达的动态阈值变化模型,实现轮廓提取过程中的智能性调节,即在进行纹理抑制的同时,有效防止局部轮廓细节的丢失,为后续通过阈值化处理获取最优结果提供了更加容易的参数设定方式。

## 5.5 小结

本章在第 4 章双拮抗神经网络轮廓响应的基础上,再次构建高级视皮层点阵神经元,以全新的方式实现选择注意机制,更多地融合了神经元的电生理学特性,而非直接通过数学模型对图像进行计算;另外,在选择注意机制的形成过程中,充分考虑了高级皮层神经元相对初级皮层神经元在感受野范围、神经元数目以及神经元自适应特性等方面的差别,最后将高级视皮层神经元的自适应调控脉冲编码响应反馈到初级视皮层,得到轮廓响应;对 BSDS500 数据库中的自然场景图像进行实验,通过与无纹理抑制的 DG 方法、显著性信息稀疏编码神经网络 SCNN,以及空间稀疏抑制颜色拮抗模型 SCO 所检测的结果进行比较,表明本文方法不仅能够有效抑制纹理边缘,同时一定程度上突显了主体目标的真实轮廓,具有不错的轮廓检测性能,与其他方法相比更加贴近视觉系统的真实感知。

## 第6章 总结与展望

### 6.1 总结

随着认知神经科学与视觉系统电生理学实验的不断深入，国内外的图像轮廓检测算法逐渐告别数学微分模型的时代，开始朝着视觉感知机制或者生物智能理论与图像轮廓检测相结合的方向发展。一方面，对视觉系统的局部感知机制进行黑箱模拟或者直接应用，而通常忽略了神经元脉冲发放的编码作用；另一方面，则是试图从神经元动作电位模型及其脉冲编码机制对视觉轮廓信息进行响应，但在神经元脉冲编码过程中，通常缺乏局部视觉感知机制的有效支撑。本文研究基于视觉感知机制的图像轮廓检测技术，在神经网络脉冲编码过程中结合视觉通路中多种重要的局部视觉感知机制，对图像轮廓进行检测，结果表现出一定的优势，而且能够为后续基于视觉机制的研究和应用提供新的启示。

虽然目前利用视觉机制的图像轮廓检测方法，相比传统方法有了长足的进步与提升，但是大都没有考虑到视通路各层次感受野的变化特性以及神经元对视觉信息的前级编码作用。本文研究了早期视觉通路各层次的感受野特性，在此基础上结合神经元前级编码与方向选择特性的形成机制，构建多感受野朝向性关联轮廓检测模型对图像轮廓进行检测，取得了不错效果。

另外，神经元感受野对颜色信息具有拮抗特性，视皮层下神经元能够动态调节颜色与亮度的关系，而初级视皮层神经元具有双拮抗特性，可以利用自身树突极性分布对特定方位的刺激进行响应；另外，神经突触的动态特性在单细胞脉冲编码形成神经元群体编码响应的过程中起到了关键作用。于是，提出一种基于神经元颜色拮抗与群体编码的轮廓检测方法。

高级视皮层神经元相比初级视皮层神经元在神经元数量、感受野大小以及信息处理模式上具有更高的智能性。因此，在初级视皮层轮廓响应的基础上，结合高级皮层神经元对空间稀疏性的自适应编码特性，构建高级皮层点阵神经网络，以脉冲频率发放的形式实现选择注意机制，最终通过反馈调控的方式作用于初级皮层轮廓响应，提出一种基于双层神经网络脉冲调控编码的轮廓检测模型，在彩色图像轮廓检测中取得了不错的效果。

论文的具体研究成果总结如下：

(1)本文充分探讨了视觉通路各层次神经元的感受野特性，引入了视网膜神经元的轮廓信息脉冲编码机制、外膝体感受野的全局调节特性以及外膝体-视皮层信息传递过程中的层次关联特性，摒弃方向选择机制的传统黑箱模型，提出一种基于视通路多感受野朝向性关联的轮廓检测方法，对自然场景图像进行轮廓检测，与主流轮廓检测方法作对比，结果表明本文方法能够有效检测主体目标的真实轮廓。

(2)考虑到颜色特征在轮廓检测中的关键作用，并且仅仅依靠点阵神经元的单细胞脉冲编

码无法全面、高效地捕捉轮廓信息。因此,提出一种基于神经元颜色拮抗与群体编码的轮廓检测方法,首先构建视皮层下单拮抗感受野的动态调节机制,以充分响应颜色边界和亮度边界;然后利用单细胞的树突极性分布,构建初级视皮层的双拮抗神经网络,实现对特定方位视觉刺激的响应;在神经元的群体感受野范围内,考虑神经元的动态突触连接,融合单细胞的脉冲频率响应,实现对轮廓信息的群体编码。以 BSDS500 部分图像为实验对象,经过误检率、漏检率与综合指标的定量评价,表明本文方法可以有效检测彩色图像的轮廓。

(3)高级视皮层的选择注意机制有助于快速定位真实轮廓,并且初级皮层的编码结果往往需要高级皮层参与调控。因此,提出一种基于双层神经网络脉冲调控编码的轮廓检测新方法。在初级视皮层轮廓响应的基础上,构建高级视皮层点阵神经网络,结合神经元脉冲发放阈值随空间稀疏信息动态变化的特性,将神经元脉冲发放频率映射为选择注意图像,然后利用选择注意图对初级视皮层轮廓响应进行反馈调控,进一步增强主体轮廓、抑制纹理边缘。依然对 BSDS500 中的图像进行轮廓检测,结果表明本文方法能较好地抑制纹理边缘,同时防止轮廓局部细节的丢失;通过  $P$  值指标计算,本文方法相比其他方法具有更高的生物智能性。

## 6.2 展望

本文在图像轮廓检测的实际任务中,充分融合视觉感知机制,证明了利用视觉感知机制进行图像处理与应用的有效性和可行性。对于神经网络模型,后续可以考虑神经元突触连接的可塑性机制,以及视通路各层次神经网络如何对视觉信息进行层层过滤;另外,由于基于视觉背侧或者腹侧通路的神经计算模型在计算效率上有待提高,后续研究还可以考虑视觉系统对敏感视觉信息的快速反应机制在提高计算效率方面的作用;最后,可以尝试将基于视觉感知机制的神经网络模型应用于更高层次的图像处理或者机器视觉任务中,例如目标检测、行为分析等。

## 致谢

时光荏苒，岁月如梭，当时间幻化成沙在指尖静静流走，不变的只是春去秋来、四季轮回，变幻着的却是我们的青春年华与岁月雕刻而渐渐成熟的脸庞。就这样，我们毕业了，简单充实，却又是两三年的光阴逝去。

此时，回忆着读研期间的点点滴滴，忽然发现有些事已经遗忘在脑海深处，唤不起一丝丝记忆，然而可以清晰地想起的是身边出现的一张张可爱的脸庞。印象最为深刻的就是“老范”，没错，就是我们的导师——范影乐教授，俊朗稳重的外表下，总是对学生怀揣着一份关心和爱护。一直以来，范老师在生活、科研和工作方面都给予我们巨大的帮助，总能让我们在轻松愉快的氛围里有条不紊地开展科研工作，直至顺利完成学业。生活上，从住宿环境到睡眠质量，范老师总能帮我们解决生活上的各种难题；科研上，不论是科研竞赛的文档撰写还是小论文和专利的发表，范老师都尽心尽力为我们修改文档，给出关键性的修改意见，甚至经常熬夜帮我们解决难题；当面临找实习、找工作的时候，范老师总能站在我们的立场上全力支持我们。在此，衷心感谢范老师在读研期间对我的关怀和指导！

此外，感谢廖进文师兄、罗佳骏师兄在研一期间对我的照顾和帮助，让我初到传感器实验室时很快地融入了师门大家庭；感谢“小黑屋”的李丹菁师兄、陆晓娟师姐、王梦蕾师姐，你们的热情、爽朗让我印象深刻；感谢方芳师姐在智慧城市、移动终端科研竞赛上的帮助与支持，北京之行是一段快乐难忘的时光；特别感谢王典师兄在小论文撰写方面的巨大支持以及王迪师兄在找工作方面提供的宝贵建议；感谢同门好友谢尚周、郭笑天、李嘉祥，一直以来总是相互合作、共同进步；感谢师弟王强、郭金鹏、胡钧皓、张明琦、杜宇华、蒋涯、周涛，师妹张梦楠，你们让实验室充满朝气与欢乐。

最后，我要感谢我的家人，感谢父母无私的付出和关爱，感谢哥哥在我求学生涯上的鼓励与支持，今后我将更加努力地工作，幸福地生活！

## 参考文献

- [1] Peterson J K. Calculus for cognitive scientists: partial differential equation models[M]. Springer publishing company, incorporated, 2016:401-484.
- [2] Cessac B, Paugammoisy H, Viéville T. Overview of facts and issues about neural coding by spikes[J]. Journal of physiology Paris, 2009, 104(1-2):5-18.
- [3] Eyherabide H G, Samengo I. The information transmitted by spike patterns in single neurons[J]. Journal of physiology Paris, 2009, 104(3-4):147-155.
- [4] Tanabe S. Population codes in the visual cortex[J]. Neuroscience research, 2013, 76(3):101-105.
- [5] Harvey C D, Svoboda K. Locally dynamic synaptic learning rules in pyramidal neuron dendrites[J]. Nature, 2007, 450(7173):1195-1200.
- [6] Haupt S S, Spengler F, Husemann R, et al. Receptive field scatter, topography and map variability in different layers of the hindpaw representation of rat somatosensory cortex[J]. Experimental brain research, 2004, 155(4):485-499.
- [7] Kondo S, Ohki K. Laminar differences in the orientation selectivity of geniculate afferents in mouse primary visual cortex[J]. Nature neuroscience, 2016, 19(2):316-319.
- [8] Thompson A D, Picard N, Min L, et al. Cortical feedback regulates feedforward retinogeniculate refinement[J]. Neuron, 2016, 91(5):1021-1033.
- [9] Papari G, Petkov N. Edge and line oriented contour detection: State of the art[J]. Image & vision computing, 2011, 29(2): 79-103.
- [10] Kiani, Roozbeh, Christopher , et al. Natural grouping of neural responses reveals spatially segregated clusters in prearcuate cortex[J]. Neuron, 2015, 85(6):1359-1373.
- [11] Bao P, Zhang L, Wu X. Canny edge detection enhancement by scale multiplication[J]. IEEE transactions on pattern analysis & machine intelligence, 2005, 27(9):1485-1490.
- [12] Kazakova N, Margala M, Durdle N G. Sobel edge detection processor for a real-time volume rendering system[C]. International symposium on circuits and systems, 2004:II-913-16 Vol.2.
- [13] Seif A, Salut M M, Marsono M N. A hardware architecture of Prewitt edge detection[C].

Sustainable utilization and development in engineering and technology, 2010:99-101.

- [14] Zhao D, Yang L, Wu X, et al. An improved roberts edge detection algorithm based on mean filter and wavelet denoising[M]. Advances in information technology and industry applications, Springer Berlin Heidelberg, 2012:299-305.
- [15] Xiao Z, Guo C, Yu M, et al. Research on log Gabor wavelet and its application in image edge detection[C]. International conference on signal processing proceedings, 2002:592-595 vol.1.
- [16] Guifang W U, Ke X U, Jinwu X U, et al. Application of morphological wavelet in surface crack defect detection of medium plates[J]. Journal of university of science & technology Beijing, 2006, 28(6):591-594.
- [17] Arbeláez P, Maire M, Fowlkes C, et al. Contour detection and hierarchical image segmentation[J]. IEEE transactions on pattern analysis & machine intelligence, 2011, 33(5):898-916.
- [18] Grigorescu C, Petkov N, Westenberg M. Contour detection based on nonclassical receptive field inhibition[J]. IEEE transactions on image processing a publication of the IEEE signal processing society, 2003, 12(7):729-739.
- [19] 桑农, 唐奇伶, 张天序. 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.
- [20] Du X, Li C, Li J. Contour detection based on compound receptive field[J]. Journal of electric and information technology, 2009, 31(7): 1630-1634.
- [21] Yang K, Gao S, Guo C, et al. Boundary detection using double-opponency and spatial sparseness constraint[J]. IEEE transactions on image processing, 2015, 24(8):2565-2578.
- [22] Yang K, Li C, Li Y. Multifeature-based surround inhibition improves contour detection in natural images[J]. IEEE transactions on image processing, 2014, 23(12): 5020-5032.
- [23] Ichida J M, Schwabe L, Bressloff P C, et al. Response facilitation from the "suppressive" receptive field surround of macaque V1 neurons[J]. Journal of neurophysiology, 2007, 98(4):2168-2181.
- [24] 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.
- [25] Shen Z, Xu W, Li C. Cue-invariant detection of centre-surround discontinuity by V1 neurons in awake macaque monkey[J]. Journal of physiology, 2007, 583(2):581-592.



- [26] Azzopardi G, Petkov N. A CORF computational model of a simple cell that relies on LGN input out performs the Gabor function model[J]. *Biological cybernetics*, 2012, 106(3):177-189.
- [27] Brown E N, Kass R E, Mitra P P. Multiple neural spike train data analysis: state-of-the-art and future challenges[J]. *Nature neuroscience*, 2004, 7(5):456-461.
- [28] Sun Y, Zhou D, Rangan A V, et al. Library-based numerical reduction of the Hodgkin-Huxley neuron for network simulation[J]. *Journal of computational neuroscience*, 2009, 27(3):369-390.
- [29] Carver S, Roth E, Cowan N J, et al. Synaptic plasticity can produce and enhance direction selectivity[J]. *PLoS computational biology*, 2008, 4(2): 199-204.
- [30] Pillow J W, Shlens J, Paninski L, et al. Spatio-temporal correlations and visual signalling in a complete neuronal population[J]. *Nature*, 2008, 454(7207): 995-999.
- [31] Pillow J W, Paninski L, Uzzell V J, et al. Prediction and decoding of retinal ganglion cell responses with a probabilistic spiking model[J]. *The journal of neuroscience*, 2005, 25(47): 11003-11013.
- [32] 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.
- [33] Reich D S, Mechler F, Purpura K P, et al. Interspike intervals, receptive fields, and information encoding in primary visual cortex[J]. *The journal of neuroscience*, 2000, 20(5): 1964-1974.
- [34] Kayser C, Montemurro M A, Logothetis N K, et al. Spike-phase coding boosts and stabilizes information carried by spatial and temporal spike patterns[J]. *Neuron*, 2009, 61(4): 597-608.
- [35] Zhou L, Sun Y, Zheng J. Automated color image edge detection using improved PCNN model[J]. *Wseas transactions on computers*, 2008, 7(4):184-189.
- [36] Delorme A, Perrinet L, Thorpe S J. Networks of integrate-and-fire neurons using rank order coding B: Spike timing dependent plasticity and emergence of orientation selectivity[J]. *Neurocomputing*, 2001, s38-40(01):539-545.
- [37] Kerr D, McGinnity T M, Coleman S, et al. A biologically inspired spiking model of visual processing for image feature detection[J]. *Neurocomputing*, 2015, 158:268-280.
- [38] Frintrop S, Rome E, Christensen H I. Computational visual attention systems and their cognitive foundations: a survey[J]. *ACM transactions on applied perception*, 2010, 7(1):6-6.
- [39] Gerstner W, Kistler W M. *Spiking Neuron Models*[M]. *Spiking neuron models*, Cambridge

University press, 2002:277–280.

- [40] Waław M, Justyna S. On the interspike-intervals of periodically-driven integrate-and-fire models[J]. Journal of mathematical analysis & applications, 2015, 423(1):456-479.
- [41] Nahmias M A, Shastri B J, Tait A N, et al. A leaky integrate-and-fire laser neuron for ultrafast cognitive computing[J]. IEEE journal of selected topics in quantum electronics, 2013, 19(5):1-12.
- [42] Schwabe L, Adorján P, Obermayer K. Spike-frequency adaptation as a mechanism for dynamic coding in V1[J]. Neurocomputing, 2001, 38-40(01):351-358.
- [43] Vanrullen R, Thorpe S J. Surfing a spike wave down the ventral stream[J]. Vision research, 2002, 42(23): 2593-2615.
- [44] Spiridon M, Gerstner W. Effect of lateral connections on the accuracy of the population code for a network of spiking neurons[M]. International journal of business and economics. Feng Chia University, 2002:409-421.
- [45] Lazareva N A, Shevelev I A, Novikova R V, et al. The disinhibitory zone of the striate neuron receptive field and its sensitivity to cross-like figures[J]. Neuroscience and behavioral physiology, 2002, 32(6):595-602.
- [46] Shapley R, Hawken M. Neural mechanisms for color perception in the primary visual cortex[J]. Current opinion in neurobiology, 2002, 12(4):426-432.
- [47] Zhen Z, Zhe C, Li L, et al. Adaptive background suppression method based on visual receptive field[J]. Chinese journal of scientific instrument, 2014, 35(1):191-199.
- [48] Nonclassical Receptive Fields[M]. Encyclopedia of computational neuroscience. Springer New York, 2015:2104-2104.
- [49] Li C Y, Xing P, Yi-Xiong Z, et al. Role of the extensive area outside the X-cell receptive field in brightness information transmission[J]. Vision research, 1991, 31(9):1529-1540.
- [50] Huang L L, Shimizu A, Kobatake H. Robust face detection using Gabor filter features[J]. Pattern recognition letters, 2005, 26(11):1641-1649.
- [51] Wolfe J M. Visual attention - seeing - chapter 8[J]. Seeing, 2000: 335–386.
- [52] 王典. 基于视觉感知的神经计算模型及其应用研究[D]. 杭州: 杭州电子科技大学, 2015.
- [53] Michael G, Richard B, George R. Cognitive neuroscience the biology of the mind(3rd edition)[M]. China light industry press, 2010:152-169.

- [54] 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.
- [55] Shevell S K, Kingdom F A. Color in complex scenes[J]. *Psychology*, 2008, 59(59):143-166.
- [56] Zhao X, Chen H, Liu X, et al. Orientation-selective response in the mouse lateral geniculate nucleus[J]. *The journal of neuroscience*, 2013, 33(31):12751-12763.
- [57] Xu X, Bonds A B, Casagrande V A. Modeling receptive-field structure of koniocellular, magnocellular, and parvocellular LGN cells in the owl monkey (*aotus trivigatus*)[J]. *Visual neuroscience*, 2002, 19(6):703-711.
- [58] Reskó B, Baranyi P, Korondi P. Orientation selective contour detection using oriented motion blur[C]. *International conference on artificial intelligence and pattern recognition*, Aipr-07, Orlando, Florida, USA, July. 2007:272-277.
- [59] Martin D, Fowlkes C C, Tal D, et al. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics[C]. *International conference on computer vision*, 2001: 416-423.
- [60] Hirsch J A. Synaptic physiology and receptive field structure in the early visual pathway of the cat[J]. *Cerebral cortex*, 2003, 13(1):63-69.
- [61] 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.
- [62] 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.
- [63] Yang K, Gao S, Li C, et al. Efficient color boundary detection with color-opponent mechanisms[C]. *Computer vision and pattern recognition*, 2013: 2810-2817.

## 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. 基于视通路多感受野朝向性关联的轮廓检测方法. 中国生物医学工程学报, 2017, 36(1):1-11.(一级期刊), 1/4.
2. 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, 3/5.

#### 授权(申请)专利:

##### 发明专利

1. 一种基于视通路多感受野朝向性关联的轮廓检测方法, 已公开, 2016104306848, 2/4.

#### 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. 基于多模态信息的人体运动功能障碍评估研究(2013C24016), 浙江省科技国际合作项目。
3. 自动化检测中的机器视觉开发模块研发(063113017), 横向课题。