

杭州电子科技大学

硕士学位论文

题目：基于视觉感知的图像轮廓检测技术研究

研究生 张梦楠

专业 控制科学与工程

指导教师 范影乐 教授

完成日期 2018 年 3 月

杭州电子科技大学硕士学位论文

基于视觉感知的图像轮廓检测技术研究

研 究 生： 张 梦 楠

指导教师： 范 影 乐 教 授

2018 年 3 月

Dissertation Submitted to Hangzhou Dianzi University
for the Degree of Master

Research on Image Contour Detection Based on Visual Perception

Candidate: Zhang Mengnan

Supervisor: Prof. Fan Yingle

March, 2018

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

轮廓是视觉注意目标与干扰背景之间的分割边界，一方面它能够对待识别目标的几何外形进行精简且充分的有效描述，这对于提高后续图像分析和理解的效率和准确性具有重要的意义；另一方面轮廓感知是生物视觉系统在与外界环境交互时的关键能力，这对于视觉机制的理解和应用是一个较好的切入点，因此轮廓检测是视觉感知领域的研究热点之一。传统轮廓检测方法主要是基于例如梯度运算之类的数学模型，它们只是简单地利用轮廓与背景在灰度或颜色上的变化，并没有考虑到生物视觉在目标识别中的内在机制，因此对于背景纹理干扰严重的情形，有效目标的识别将变得不可行，此时轮廓检测的性能将急剧下降。随着神经实验手段和计算能力的改善，目前也有一些基于视觉机制的轮廓检测方法被陆续提出，但多数只是一种神经网络性质的黑箱模拟，它们更关心当前激励与响应之间的匹配，而忽视了视觉信息流在视通路中传递和处理的机理。因此本文围绕着视通路层次结构、感受野特性以及神经编码等重要视觉感知机制及其应用展开研究工作，首先对初级视皮层细胞感受野特性进行了模拟，并对视觉信息流的高低频分量分别采取了冗余和稀疏编码，在考虑神经元脉冲传递时的抑制与前馈作用时，还充分利用了经典和非经典感受野响应的差异性。然后考虑视通路的方位敏感特性，提出并实现了神经元感受野非对称方位敏感特性在轮廓检测中的应用；最后研究了感受野尺度对图像显著信息表征的影响，构建一种基于感受野多尺度的显著性特征融合机制，利用尺度变化来实现轮廓整体和细节的分级提取。本文的主要研究工作及成果如下：

(1) 基于视通路中视觉信息传递和处理机制，提出了一种基于频域融合编码与前馈机制的图像轮廓检测新方法。对传入的视觉信息进行 NSCT 频域分解，高频信号采用冗余编码的方式强化其轮廓信息，低频信号采用稀疏编码方式提取有效轮廓信息，再对上述轮廓信息进行融合；利用高斯导函数模拟初级视皮层感受野特性，引入经典和非经典感受野响应的差异性实现抑制作用与前馈调节作用，实现图像轮廓的提取。选取 RuG 图像库的自然图像为检测对象，计算检测结果与基准轮廓图的 P 指标，结果表明，新方法不仅能够有效提取轮廓，而且表现出对背景纹理噪声良好的抑制作用，在整体轮廓检测效果上优势明显。

(2) 基于视觉感知的多方位敏感特性，提出了一种基于神经元感受野多方位选择机制的图像轮廓检测方法。首先根据初级视皮层的神经元细胞响应特性对图像进行稀疏编码，并对感受野的方位选择机制进行改进，构建具有方向权值占比的响应边界；同时模拟初级视皮层感受野特性，基于非对称方位的敏感性感受野特性对图像进行感受野大小的自适应选择并进行有效方位响应筛选处理获取图像的有效轮廓。经计算综合评价指标 P ，结果表明，新方法获得的主体轮廓更为完整，背景纹理噪声得到有效抑制，轮廓凸显效果较好。

(3) 考虑到单一的感受野尺度无法完整体现轮廓的整体与细节特征，提出了一种基于显著信息多尺度融合的轮廓检测方法。首先在图像感受野的非对称方位敏感特性研究基础上，深入探讨其敏感特性在不同感受野尺寸下的显著信息表征特性，构建一种显著信息的融合机制，然后进行前述的 NSCT 频域分解，再引入改进机制下的初级皮层边界响应及感受野差异特性，完成图像轮廓的检测。实验结果表明，新方法得到的整体轮廓图像更为干净，背景纹理噪声的抑制效果更好，主体轮廓的完整度更高。

关键词：视觉感知，感受野，视觉机制，神经编码，突触特性，轮廓检测

ABSTRACT

Contour is the segmentation boundary between the target and background. On the one hand, it can simplify and fully describe the geometric appearance of the recognition target, which is of great significance for improving the efficiency and accuracy of subsequent image analysis and understanding. On the other hand, contour sensing is the key ability of biological visual system in interacting with the external environment, which is a good point for the understanding and application of visual mechanism. Therefore, contour detection is one of the hot topics in the field of visual perception. Traditional contour detection methods are mainly based on mathematical models such as gradient operation, they just simply take advantage of the changes in the gray or color of the contour and the background, and do not take into account the inherent mechanism of biological vision in target recognition. Therefore, the recognition of the correct target will become infeasible for the case of serious background texture interference, at the same time, the performance of the contour detection will drop sharply. With the improvement of neuroscience experiments and computing power, some methods based on visual mechanism have been proposed, but most of them like a black box simulation of neural network nature. They are more concerned with the matching of current incentives and responses, and ignore the mechanism of the transmission and processing of visual information flow in the visual path. Therefore, this dissertation focuses on the visual perception mechanism and its application, such as visual path hierarchy, receptive field characteristics and neural coding. Firstly, the receptive field characteristics of primary visual cortex cells were simulated, and the high and low frequency components of visual information flow were respectively taken redundant and sparse coding. Considering the inhibitory and feed-forward effects of neuronal pulse transmission, the differences between the classical and non-classical receptive field response are also fully utilized. Secondly, we consider the azimuthal sensitivity of visual pathway and propose the asymmetric azimuthal sensitivity of neuron receptive field in contour detection. Finally, the influence of different scales of receptive fields on the representation of significant information is studied, and a salient features fusion mechanism based on receptive field and multi-scale is constructed. The scale change is used to realize the classification of the whole and detail of the contour. The main research work and achievements of this paper are as follows:

(1) Based on the visual information transmission and processing mechanism in the visual pathway, we propose a new method of contour detection based on the fusion of frequency domain and feedforward mechanism. The incoming visual information is decomposed by NSCT in

frequency domain, the high frequency signal is enhanced by redundant coding, the low frequency signal is extracted by sparse coding to extract the effective contour information, and then the above contour information is fused. Gaussian derivative is used to simulate the receptive field of primary visual cortex and introducing the difference characteristics of classical and non-classical receptive field response, the inhibition and feedforward regulation were realized to extract the contour of the image. The dataset RuG40 is selected as the test object to calculate the comprehensive evaluation index P . The results show that the new method not only can effectively extract the contour, but also has a good inhibitory effect on background texture noise.

(2) Based on the multi-orientation sensitive features of visual perception, we propose a method of image contour detection based on multi azimuth selection mechanism of neuron receptive field. Firstly, the image is sparsely coded according to the response characteristics of neurons in the primary visual cortex, and the orientation selection mechanism of receptive fields is improved to construct the response boundary with the proportion of directional weights. At the same time, the receptive field characteristics of the primary visual cortex are simulated based on the non-Symmetric orientation of the sensitive receptive field characteristics of the image receptive field size adaptive selection and effective azimuth response filtering process to obtain the effective contour of the image. After calculating the comprehensive evaluation index P , the result shows that the main body contour obtained by the new method is more complete, the background texture noise is effectively suppressed, and the contour highlighting effect is better.

(3) Considering that the single sensorial scale cannot fully reflect the whole and detail features of the contour, we propose a contour detection method based on saliency information multi-scale fusion. Firstly, based on the asymmetric azimuth sensitivity of image receptive fields, we explore the significant information characterization characteristics of sensitive features under different receptive field sizes, constructs a fusion mechanism of significant information, then performs the aforementioned NSCT frequency domain decomposition, and the boundary of the primary cortical boundary and the difference of receptive field were introduced under the improved mechanism to complete the detection of the image contour. The experimental results show that the new method obtains a cleaner outline image, a better suppression for background texture noise and a higher integrity of the main contour.

Keywords: visual perception, receptive field, visual mechanism, neural coding, synapse character, contour detection

目录

摘要.....	I
ABSTRACT.....	III
目录.....	V
第 1 章 绪论.....	1
1.1 研究目的与意义.....	1
1.2 国内外研究现状.....	2
1.3 本文主要工作及创新点.....	4
1.4 论文的组织结构.....	5
第 2 章 视觉系统的感知特性概述.....	7
2.1 视觉通路.....	8
2.1.1 视网膜.....	8
2.1.2 外膝状体.....	8
2.1.3 视皮层.....	9
2.2 突触特性.....	10
2.3 视觉感受野.....	11
2.3.1 经典感受野.....	11
2.3.2 非经典感受野.....	11
2.3.3 拮抗式感受野.....	11
2.4 神经元模型及编码.....	12
2.4.1 Hodgkin-Huxley 神经元模型.....	13
2.4.2 Integrate-and-Fire 神经元模型.....	14
2.4.3 Leaky Integrate-and-Fire 神经元模型.....	15
2.4.4 脉冲编码.....	16
2.5 小结.....	16
第 3 章 基于频域融合编码与前馈机制的图像轮廓检测.....	17
3.1 基本原理.....	17
3.1.1 初级轮廓感知响应.....	18
3.1.2 视觉信息分频编码.....	19
3.1.3 视觉前馈机制.....	22
3.2 算法流程与验证.....	22
3.2.1 算法流程.....	22
3.2.2 算法验证.....	24
3.3 实验结果.....	25
3.4 讨论与分析.....	29
3.5 小结.....	29
第 4 章 基于神经元感受野多方位选择机制的图像轮廓检测.....	31
4.1 基本原理.....	31
4.1.1 改进的方向选择机制.....	31
4.1.2 非对称的方位敏感性感受野.....	33

4.2 算法流程与验证.....	34
4.2.1 算法流程.....	34
4.2.2 算法验证.....	36
4.3 实验结果.....	37
4.4 讨论与分析.....	41
4.5 小结.....	42
第 5 章 基于显著信息多尺度融合的轮廓检测.....	43
5.1 轮廓检测模型.....	43
5.1.1 神经元感受野边界响应.....	44
5.1.2 多尺度表征融合机制.....	45
5.1.3 轮廓检测流程.....	46
5.2 实验结果.....	49
5.3 讨论与分析.....	52
5.4 小结.....	52
第 6 章 总结与展望.....	53
6.1 总结.....	53
6.2 展望.....	54
致谢.....	55
参考文献.....	57
附录.....	61

第 1 章 绪论

1.1 研究目的与意义

视野中的浩瀚星空，是无边无际的宇宙烙印在人们脑海中的直观印象；而脑海里数千亿计的神经元就如同那相映成辉的繁星，构成了宇宙中最富有想象力的大脑。对大脑工作机制的研究，是我们认识自我和完善自我的必要途径。由于在大脑所接收的信息中，绝大部分来源于视觉系统，因此可以认为视觉感知在大脑功能的正常维持中扮演了重要的角色。随着神经电生理和光学实验手段的提高，目前视觉信息的主要传递路径已经较为清晰，即外界光刺激诱发视网膜上视锥和视杆细胞产生脉冲电位，通过电信号或化学媒介传递的方式，经过神经节和外膝体等路径中间节点的编码和传递，最终传入复杂的视觉感知皮层进行处理。目前神经工程和临床医学领域的关注点，更多的在于揭示视觉感知系统及其功能区域的各种神经特性，并不关心这些特性是否具有工程应用价值；而以深度学习为代表的人工智能虽然是目前工程技术领域研究的热点，却没有真正涉及到感知系统的内在机制，只是从神经编码和突触学习的角度去构建卷积神经网络，试图采取一种黑箱式的方式来模拟感知系统。所以从工程应用的视角来研究视觉感知机制，并建立合理的神经计算模型，尝试开展应用研究，一方面将改变传统工程技术手段的局限性；另一方面计算仿真的结果也将促进人们对大脑工作机制的进一步理解。

随着成像传感设备和计算处理能力的提升，基于图像处理的应用已经变得越来越普及，甚至连无人驾驶这类曾经遥不可及的应用也已经行进在路上。视觉神经生理和心理实验表明，视觉系统在对目标进行感知时，由于视觉感受野具有方向性、对比度敏感以及整体尺度等特性，外形轮廓将构成目标的主要特征，是视觉感知系统实现快速识别的基础。而对于被广泛应用的计算视觉技术而言，也通常会将分析区域定位于目标轮廓内，将极大地简化识别难度，显著提高运算效率和识别性能。因此，图像轮廓的自动提取对于图像内容的分析理解是至关重要的。图像轮廓上信息极为丰富，在信息最为集中的地方其轮廓表现最为陡峭或曲率最大，在信息较为多余的地方其轮廓方向表现相似或一致。物体的轮廓不同于边缘，它包含更多的信息，借助轮廓，人们可识别大量物体，获得重要精简的特征信息。

研究者们对视觉系统的相关理论及研究成果表明，视觉系统之所以有优越的轮廓感知能力，其中一个很重要的因素在于视觉感知系统存在着包括感受野特性、侧抑制、视通路融合、动态突触学习以及神经编码等机制，这为智能化行为的更深入研究提供了基础；由众多神经元及突触互连而形成的神经元网络，能够将视觉输入转化为脉冲信号编码，从生物学角度出发，对视觉感知进行研究，结合神经元模型理论并将其高效地应用于图像轮廓的检测，在计算机视觉研究领域具有极其重要的意义。

1.2 国内外研究现状

数字图像处理中，传统轮廓边缘检测方法中常用的检测算子有 Canny 算子^[1]、Sobel 算子^[2]、Prewitt 算子^[3]、Laplace 算子^[4]、Roberts 算子^[5]等，它们对简单的图像能够取得不错的检测效果，但当图像纹理背景较为复杂时，其检测效果就会急剧降低。经典轮廓提取技术大多基于微分运算，也有大批学者基于上下文分析、多分辨率分析等方法提出了大量边缘轮廓检测方法^[6]，相对于传统方法而言提高了检测精度及效率，但这些计算模型在当今的智能领域上的表现远达不到人们的期望。轮廓检测的初始概念是基于人类视觉对认识世界识别物体的整体感知而形成的，对于此，很多研究人员从人类的视觉感知系统出发^[7]，研究人类的视觉机制，建立数学仿真模型，或通过构建神经网络，考虑方式也从图像局部简单的亮度梯度特征向多个局部特征发展来实现图像轮廓的检测。Pb 算法就是一个计算多个局部特征的轮廓检测算法^[8]，它通过分析自然图像中多个局部特征的特点以及寻找这些局部特征的优化组合方式，最终给出比 Sobel 算子和 Canny 算子更好的检测方法。CRF 算法第一次将图像中物体全局结构融合进局部特征中，它通过构建 CDT 图和全局连续性模型完成 Pb 算法中断裂边界线的连接工作，使得轮廓检测的结果进一步准确^[9]。min-cover 算法从区域覆盖的角度分析，将寻找边界线问题转向寻找图像的最小区域覆盖问题^[10]，它提出在整幅图像中图像不同物体或背景区域覆盖的总和即是图像本身，则不同区域之间的交线即是边界线的观点。也有的研究人员^[11]对传统检测方法进行了优化改进，例如有研究针对在噪声干扰的情况下，如何更好地实现轮廓的定位，提出了 Mallat 小波^[12]；为了提高传统分割算法的鲁棒性、加快检测速率，从局部空间信息整合方向考虑，引入了马尔科夫随机场能量的水平集算法^[13]等。相对于传统图像轮廓检测方法，这些算法性能有了很大的改进和提高，但应用于基于人类视觉系统的复杂场景图像检测，提取图像主体轮廓时，其表现出的检测效率及准确度与期望效果还相差甚远。

21 世纪被科学家们喻为“脑的世纪”，生物医学和神经科学发展迅速，实验技术水平也有了极大提高，一大批研究人员已从纯数学角度对图像处理的研究转变为从生物角度进行研究，并在视觉神经系统的信息处理方面取得丰硕成果，在基于视觉机制结合计算机方法建立相应的数学模型，从根本上改善传统图像处理算法的不足。

研究者们基于视觉神经系统及神经元的研究，对视觉特性的理解和应用取得了更深入的成果。例如针对大鼠视网膜的实验研究，发现传统观点所认为的感受野周围神经元对刺激具有不敏感性的观点并不准确，更有可能是由于兴奋的中心神经元通过突触连接对周围神经元产生了侧向抑制^[14]。还有研究观察到了神经元响应对于不同方向激励的特异性，从而构建出能够实现运动检测的视觉神经计算模型^[15]。也有研究从神经元突触互连的角度入手，发现动态突触可塑性与方向选择性之间有着密切的联系，也就是模型响应与激励方向之间是一种动态调整的关系。也有观点认为相较于对单一神经元发放模式，无论是从编码信息量还是鲁棒性来说，神经元群的编码模式更具可行性^[14]，因此所提出的基于概率发放神经元模型的解码

方式^[15]，能够较好地解释编码信息。在视皮层的研究中，科学家们证实了初级视皮层 V1 层细胞感受野对于视觉刺激的方位表现出了明显的敏感性。由于 Gabor 滤波器的方向表达性跟视觉系统较为类似，因此通常利用 Gabor 滤波器来描述 V1 层细胞感受野的响应特性^{[18][19]}，可以很好地描述单细胞的感受野对边缘明暗敏感性的响应程度。对于复杂度较高的场景，存在大量背景纹理噪声，Gabor 函数不能对其进行抑制，所以 Gabor 滤波器虽然能够提取边缘信息，但对于图像主体轮廓的提取并不理想。基于以上存在问题，Grigorescu 等^{[20][21]}提出采用 Gabor 函数和高斯差函数(Difference of Gaussian, DoG)结合的方式，先把初级视皮层经典感受野(CRF)的边缘响应用 Gabor 能量进行表征，再对高斯差函数结果附加半波整流获得新函数(DoG⁺)，通常可以用 DoG⁺来表达非经典感受野(NCRF)所具有的大外周响应和侧向调控特性，从而实现背景纹理的抑制，凸显目标轮廓。近年来，Chen^{[22][23]}以及 Xu^[24]等人对非经典感受野作了大量研究，发现初级视皮层 V1 层的神经元响应可以通过检测经典感受野和非经典感受野的空间差异来进行调制，NCRF 在具备抑制作用的同时，还存在着去抑制区，这对提高 CRF 边缘响应的连续性作用极大，使得在视觉系统中复杂信息处理方面的研究更进一步。

随着研究的深入，基于初级视皮层的重要性，神经编码的研究越来越受到研究人员的青睐。Thorpe 等人将时序编码方式代替发放率编码方式^[25]，两者通过减少模型计算的时间开销，加强神经编码的实用性。Kayser 等人对空间信息的编码研究中^[26]，提出多距离尺度的综合性空间编码应用于视皮层能够对噪声干扰起到抑制作用。Cevikbas 以及黄丽鸿等人分别将神经元群编码模式应用于字符识别^[27]及运动检测^[28]；Luvizotto 等人则将视皮层编码理论应用于人脸识别^[29]，取得了较好的效果。但编码模式多种多样，虽然取得一定成果，但对于复杂的图像，依然任重而道远。为能够更准确形象地从视觉角度实现对图像信息进行处理，研究人员们试图构建更完美的神经网络模型。基于动作电位发放的神经元模型，Eckhorn 提出一种脉冲耦合神经网络(Pulse Coupled Neural Network, PCNN)^[30]，这是一种基于同步脉冲发放的神经网络模型，能够从背景复杂度较高的图像中提取有效信息，且无需进行训练，所以在图像的快速识别分割等处理中被广泛使用。虽然 PCNN 模型具备神经元脉冲发放特性及传递特性，但其实质只是对于视觉神经网络的黑箱模拟，并不能表征其它视觉机制的运作，此外，PCNN 模型对于不同图像参数波动范围大，稳定性并不理想。Kerr 等提出一种基于六边形感受野模型的 SNN(Spiking Neural Network)的图像边缘检测方法^[31]，相对于传统的矩形感受野，SNN 模型更加符合生物特性，效果也更为理想。

综上，从生物视觉角度看待图像结合数学模型的图像处理研究已成为当今图像视觉处理的发展趋势。基于视觉机制的研究已积累了丰富的资料并取得了很多成果，一部分研究人员专注于神经编码机制的研究，却忽略了视觉神经网络自身所具有的组织结构；还有一些研究专注于神经网络模型的构建，关注于模型的学习收敛性，而忽视了视觉神经系统应该所具有的内在特性。因此本文试图在计算模型构建过程中体现视觉感知的某些重要特性，并将其应

用于图像轮廓检测。本文将从生物视觉感知角度出发，关注视觉通路及感受野对视觉信息的处理方式，探索构建更符合视觉神经特性的视觉感知计算模型，为后续图像分析和理解的研究提供一些生物视觉机制应用方面的新思路。

1.3 本文主要工作及创新点

本文主要围绕视觉感知机制及其在图像轮廓检测中的应用展开研究，一方面基于神经网络模型的优化和简化，以图像输入转换为动作电位的工作模式，来实现视觉感知的相关机制，专注于视觉信号在传递过程中的传递方式、载体、影响因子及所依赖的生理结构，针对图像特点，使用图像信息的多层次处理或反馈等不同的视觉信息处理方法来实现轮廓的提取。首先，基于视通路的感受野特性及视通路中对视觉信息的感知机制，研究用频域分离处理，并行采用不同的神经编码，最后进行融合的模式，引入前馈作用，在轮廓检测中取得了较好的效果；然后，考虑神经元感受野的敏感特性，进一步探究其敏感特性表现出的非对称性，作用于方向权值占比调整，并基于感受野多方位特性的研究，构建一种感受野自适应大小选择方法，结合有效响应的筛选处理，实现轮廓的检测；最后，在以上研究基础上，专注于研究神经元敏感特性在不同感受野尺度下的信息显著性表征，运用感受野自适应选择方法获得最优感受野尺度，考虑到神经元细胞工作时的复杂性，据图像的复杂度对最优感受野尺度进行放大和缩小得到多个感受野尺度，作用于频域处理图像，分别计算各自的最佳特征时的轮廓响应点，最后进行融合以增强轮廓响应的高效性与完整性，并引入经典感受野与非经典感受野差异性处理，提取图像的主体轮廓。

(1) 以视通路中视觉信息传递的规则与特性为切入点，基于神经元电生理模型及其感受野特性进行研究，提出了一种基于频域融合编码与前馈机制的图像轮廓检测方法。一方面利用高斯导函数模拟初级视皮层的响应特性，得到视皮层初级轮廓响应图；另一方面通过 NSCT 对视觉输入进行高低频分量分离，然后对高频分量图像进行冗余编码处理，以加强主体轮廓的凸显性；对低频分量图像采取稀疏编码的方式，在去冗余的基础上达到有效轮廓提取的高效性；再将冗余编码后的图像与稀疏编码后放大的图像进行融合，使得图像轮廓部分响应点更为强烈，利用经典感受野与非经典感受野的差异特性，对融合后的轮廓进行纹理抑制以增强主体轮廓的凸显性，得到融合编码图像轮廓响应图像；最后将融合编码图像轮廓响应图像前馈作用于初级皮层视通路的末端，实现轮廓的检测。实验结果表明，本方法不仅能够有效凸显图像主体轮廓，在轮廓细节检测方面表现良好，而且对背景纹理产生了较佳的抑制效果，整体检测效果更为理想。

(2) 深入研究了神经元的多方位敏感特性，提出一种基于神经元感受野多方位选择机制的图像轮廓检测方法。对检测图像进行稀疏编码，得到特定边界响应轮廓，再根据感受野的方向响应特性构建神经网络以提取初级轮廓响应图像，再调整响应图像的方向权值占比以强化凸显轮廓的连续性；并基于图像复杂度对感受野大小进行自适应选择，根据感受野的非对称方位敏感特性对图像有效响应方位进行筛选，然后与响应图像融合，经阈值处理后实现

图像轮廓的检测。实验结果表明，本方法获得的图像轮廓具有较好的完整性和连续性，不仅背景及纹理边缘得到有效抑制，而且具有较高的图像纯净度和较好的稳定性，轮廓凸显效果有着绝对优势，同时也表明了本文轮廓检测方法所具备的生物合理性。

(3) 探究感受野的尺度与轮廓信息的表征存在影响关系，提出一种基于显著信息多尺度融合的轮廓检测方法。考虑到视通路中神经元细胞种类的复杂性及其感受野尺寸对图像轮廓凸显的影响，在感受野大小自适应选择的基础上，依据其图像复杂度对感受野尺寸进行放大和缩小，构建感受野多尺度表征融合机制，融合不同尺寸感受野下的显著信息以提高轮廓边缘的响应质量，采用上采样插值放大得到新的轮廓响应，再与经过方向权值占比调整处理过的初级轮廓响应融合，最后通过非极大值抑制及滞后阈值处理得到最终的轮廓检测结果。以自然场景下的图像为例，实验结果显示，本方法得到的图像轮廓较为完整，具有较好的纹理抑制效果，轮廓细节表达的较为出色。

1.4 论文的组织结构

本文组织结构及各章节内容简述如下：

第一章绪论，首先叙述了计算机科学与神经科学相互融合发展的趋势与必然，引出本文在研究在及图像处理领域计算机视觉领域的重要意义；然后对国内外视觉感知机制及其在图像处理中应用的研究现状进行阐述，总结相关研究具备的优势和存在的不足；从而引出本文的研究内容，最后对论文的组织结构进行了介绍。

第二章介绍了视觉系统的感知特性。描述了突触功能及特性与视通路中不同层级的感受野特性；介绍了 HH 神经元模型、IF 神经元模型及 LIF 神经元模型，并进行了数学表述；最后简要阐述了经典的神经编码方法，为后续的机制构建及检测方法的提出提供了理论基础。

第三章提出一种基于频域融合编码与前馈机制的图像轮廓检测方法。考虑视通路中信号传递和处理的规则和特性，结合上文对视觉通路、神经元模型及编码等研究，构建具备神经电生理特性的神经网络，实现对图像轮廓的提取。对 RuG 图库进行轮廓检测，并计算检测结果与图像基准轮廓图的 P 指标，实验结果显示，本文方法能够有效抑制背景纹理，较好的凸显图像主体轮廓。

第四章提出一种基于神经元感受野多方位选择机制的图像轮廓检测方法。研究神经网络电生理特性及初级视皮层的感受野特性，考虑到图像轮廓的有效响应，引入神经元稀疏编码；并基于方位敏感性感受野所具备的位置及方位的严格选择性，改进感受野方位选择机制，构建一种感受野自适应大小的选择方法，再经过融合处理等实现轮廓的提取。最后将新方法应用于 RuG 图像库中，取得了不错的轮廓检测效果。

第五章基于前文对感受野特性的深入研究，探讨了不同尺度感受野选择下对图像轮廓显著信息表现的影响，提出了一种基于显著信息的多尺度选择机制的轮廓检测。利用感受野自适应选择方法获得最优感受野尺度，并对最优感受野尺度进行放大和缩小得到多个感受野尺度，并行计算具备不同显著特征的图像轮廓点响应矩阵，基于特征的影响因子进行融合，再

与初级皮层边界响应建立融合关系，以提高图像的检测效果。以自然场景下的图像为例，对本文方法的可行性进行了验证试验。

第六章对本文研究工作进行了回顾，并针对研究方法、实验过程以及成果方面的不足，对可后续进一步开展的研究工作进行了展望。

第 2 章 视觉系统的感知特性概述

包括人类在内的每一个生物个体，都不可避免的与外界环境有所交互，这种交互主要依赖于包括视觉和听觉等功能在内的感觉系统。例如在品尝美食的时候，我们会以色香味俱全的角度来对菜肴进行评价，此时可以清楚地看到感觉系统具有协同合作的特点，但各个感觉系统之间具有显著的差异性。例如我们需要通过与味蕾完全接触的方式品尝到菜肴的味道，也可以通过一定距离的气味来感知菜肴的香味，更可以远距离以视觉的方式来品味菜肴的精美。这种非接触远距离感知外界环境的能力，是大多数生物种类能够经受自然选择而生存下来的关键因素。作为感知信息主要来源的视觉系统，是人类感知外界环境的重要途径之一，因此视觉系统是神经生物学家所重点关注的研究对象。在视觉系统^[32]中，视网膜上按一定规则分布着视锥和视杆细胞，它们负责将包括亮度和颜色等特性在内的光刺激转换成神经电信号，随后通过后续包括外膝体等功能结构在内的视觉通路进行传递，之后再加工处理过的神经编码信息传入到视觉皮层进行处理，从而实现外界客观物体在大脑中的感知。正如前述的那样，这种感知通常是以一种交互的方式在进行，也就是说不仅生物个体感知到了外界环境的状态，这种感知信息将会参与到其他高级脑区的行为控制与反馈上。视觉系统结构示意图，如图 2-1 所示，结构上主要包括眼球、外侧膝状体以及视觉皮层等。

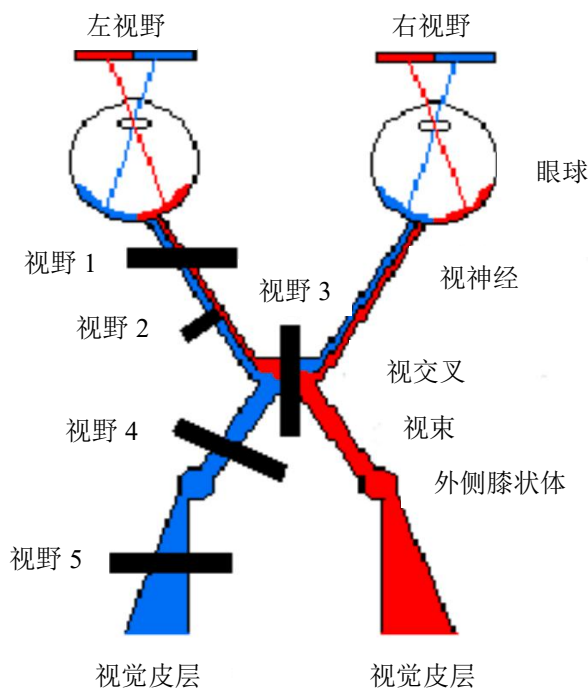


图 2-1 视觉系统结构示意图

2.1 视觉通路

光刺激经过外界物体的反射或透射作用后，聚焦于视网膜上特定区域内的视觉感受器，它们能够将光信号转变为电信号，再经过双极细胞和神经节细胞的处理，经视神经传入外侧膝状体。这里需要说明的是，包括视锥和视杆细胞在内的视觉感受器，在数量上将远超外侧膝状体的神经元数目。前述的视觉通路^[33]，主要是指视觉信息流在视网膜、外侧膝状体以及视觉皮层的传递通路，如图 2-2 所示。但也应该指出，目前越来越多的研究表明，视觉通路可能并非唯一的。在视觉感知过程中，视觉信息可能是以并行的方式在传递，因为只有如此，才有可能满足感知效率和性能的要求。

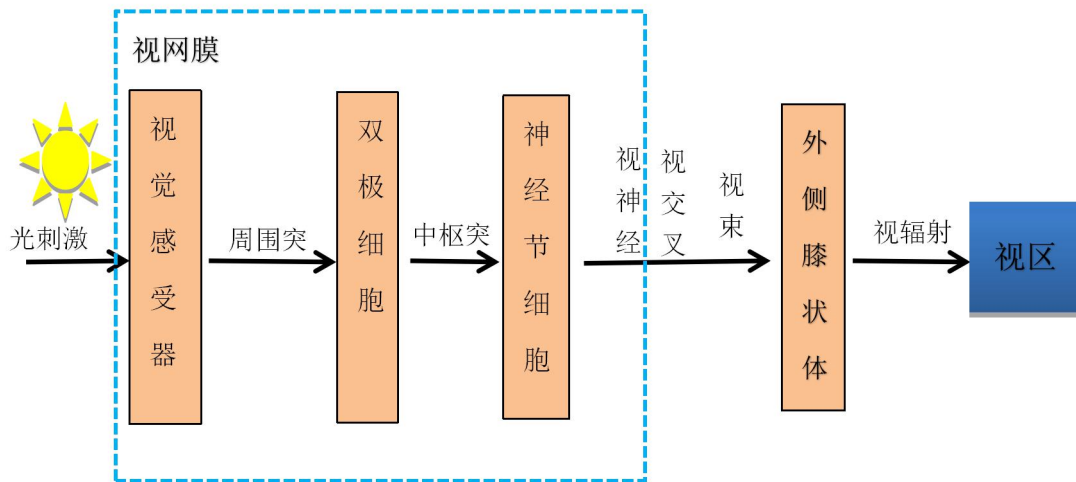


图 2-2 视觉通路

2.1.1 视网膜

视网膜位于眼球的后方，大脑皮层将接收到来自视网膜感光细胞的信息。如图 2-2 所示，视网膜主要由三层细胞构成，分别为视觉感受器、双极细胞和神经节细胞。视觉感受器主要负责将光刺激转换为脉冲电信号，然后脉冲电信号将传递至双极细胞和神经节细胞，信号经过加工整合后，将通过视神经传到外侧膝状体。而视网膜上还存在的水平细胞和无长突细胞，在光电转换以及电信号传递过程中，起到调节作用。

对于每一个个体而言，视觉感受器既具有共同性，也具有差异性。共同性表现在对于同一种光刺激的感知，个体的感受差异通常并不会很大。而差异性则主要是由视锥和视杆细胞的分布和比例所引起，因为视锥细胞数量相对较少，分布在中央凹附近，主要来感知亮环境中的颜色细节；而视杆细胞数量相对较多，均匀分布在远离中央凹的区域，主要来感知暗环境中的亮度变化。

2.1.2 外膝状体

外膝状体（lateral geniculate nucleus, LGN），也被称为外侧膝状体或外膝体，是视觉信

息传递的中继站，在视网膜到视皮层信息传递中起重要作用。外膝体 80% 的细胞的电活动具有以注视点为中心呈向心分布的方位敏感性。

研究表明，视网膜内存在多条平行通道对信息进行传递和处理。外膝体作为中继站，需要将不同种类细胞的空间频率通道、方位信息通道、颜色信息通道等分别传递到其相应的处理信息的皮层神经元。

生物研究显示，人类和猕猴的外膝体细胞共有 6 层结构，包含两层大细胞层，4 层小细胞层。正由于其结构的多层次性，外侧膝状体除了具备视觉信息的中继站的功能外，还具备视觉信息频域处理等更高级别的功能和作用，与视网膜的 M 细胞（大细胞）和 P 细胞（小细胞）两类神经节细胞共同作用于更复杂的信息处理。M 细胞将来自视锥细胞的信号叠加投射到外膝体的大细胞层，P 细胞将颜色信号处理转换至外膝体的小细胞层，外膝体的 6 层细胞综合处理后将信号输送到视皮质的不同的平行信息处理通道。有研究表明^{[31][34]}，外侧膝状体可以通过越级的方式对复杂细胞直接提供输入，其不仅仅局限于单通道的信息处理，而是以一种并行或多路径的方式对信息进行处理。尤其是在高级视觉活动中，之所以能够迅速高效地识别和理解图像信息，正是由于其视区的并行作用及视区之间的相互作用。

2.1.3 视皮层

人类具有学习记忆、视觉感知、语言行为和意识思考等能力，这些能力的发挥离不开大脑皮层的作用。研究表明，越高等的生物，其皮层的结构和功能越发达。视觉皮层在视觉感知方面有着至关重要的作用，被称为视觉系统的中央处理器。

视皮层从功能或位置来看，人们通常关注位于皮层 17 区的初级视皮层，即前述的 V1，它接收来自外膝状体的视觉信息，然后将处理后的结果传递至纹外视觉皮层及更高级的脑区^[31]。需要说明的是，这种传递的路径并不是显式串联的方式进行的，视皮层各个分区之间并没有明确的区域分隔，它们更是以一种并行计算的方式在完成视觉识别任务。例如有研究认为视皮层通过两条平行通路共同完成视觉信息处理，其中腹侧视觉通路负责处理颜色形状信息，而背侧视觉通路负责处理感知运动信息。

20 世纪 50 年代末科学家 Hubel 和 Wiesel 研究视皮层细胞对光刺激的反应特征^[35]，发现位于初级视皮层的大多数细胞具有强烈的方向选择特性。当刺激源表现为大面积弥散光时，这些细胞无反应；当刺激源表现为光棒、暗棒或具有朝向的明暗边时，这些细胞产生强烈反应；当刺激源从细胞反应最优方位偏离时，这些细胞的反应呈现骤减或停止状态。Hubel 和 Wiesel 对视皮层这些细胞进一步研究^[36]，发现了包括简单细胞和复杂细胞在内的神经元种类。

简单细胞对于处在拮抗边缘的具有一定方位和宽度的条形刺激会产生强烈反应，而对于大面积弥散光没有反应，其表现出对边缘方位的强烈选择特性，用于具有明暗对比的边缘检测。研究表明，每一个简单细胞都对应于一个最优方位，在这个最优方位上其反应最为强烈。复杂细胞对于具有一定方位和宽度的条形刺激同样会产生强烈反应，表现出强烈的方位选择特性，从形态学上来说，锥体细胞与复杂细胞特性极为相似。超复杂细胞对条形刺激的反应

与复杂细胞相似，但对条形刺激的长度有一定要求。因为在超复杂细胞感受野的一端或两端存在很强的抑制区，当条形刺激过长时就会产生抑制，其反应将会减少或消失。

2.2 突触特性

如图 2-3 所示，相邻神经元之间，是通过一种突触结构来传递信息，通过突触前膜的电位变化，引起突触后膜的离子通道开通或闭合^[37]从而实现后膜电位的触发。按突触传递的性质，可以分为兴奋性突触和抑制性突触^[38]。前者对应于兴奋性突触后电位（EPSP），将会引起突触后膜进入去极化状态；而后者对应于抑制性突触后电位（IPSP），将抑制连接神经元动作电位的发放。

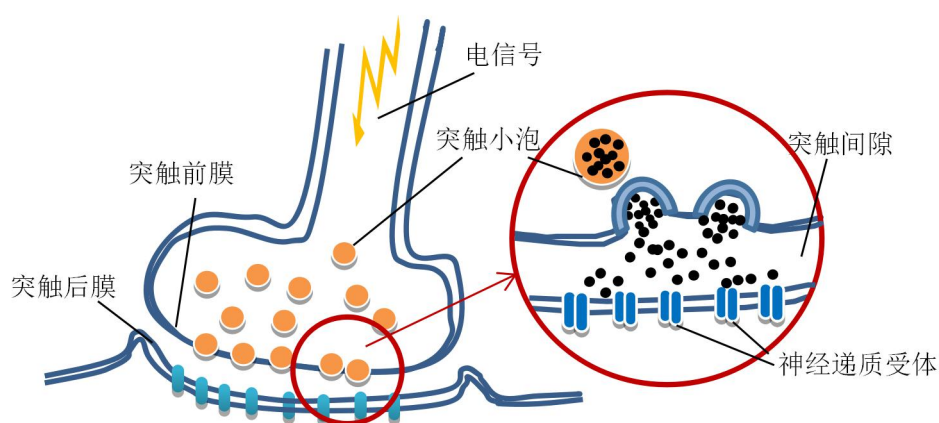


图 2-3 突触结构及信号传递示意图

神经递质的活动表现为释放于突触前膜，经突触间隙到达突触后膜，在无论在结构上还是功能上，突触在神经系统中的作用都至关重要，神经系统的可塑性（学习和记忆、机能代偿等）主要机制表现在突触可塑性（Synaptic plasticity）上。在神经科学中，突触可塑性分为短期突触可塑性和长期突触可塑性，短期突触可塑性主要包括易化、抑制和增强，长期突触可塑性主要表现为长期增强和长期抑制。

长期增强作用（Long-term potentiation, LTP），又称为长时程增强作用或长期增益效应，是一种两个神经元同时受到刺激而发生在信号传输中的一种持久增强的现象。LTP 是 1966 年泰耶·勒莫在兔海马体中发现的，科学家们认为记忆是基于突触的强度变化来编码的，所以 LTP 被认为是学习记忆基础的主要分子机制之一。LTP 具有输入专一性、关联性、协同性和持久性，一直都是研究者们感兴趣的研究主题，由于 LTP 与记忆行为等有着极大关系，其被作为医学领域的临床研究主题，如阿兹海默病和成瘾医学领域；一些研究者试图更深入了解其生物学原理，以开发提高记忆和改善学习的方法。

长期抑制作用（Long-term depression, LTD），又称为长时程抑制作用、长期抑势，通常指突触的抑制行为持续几小时甚至几天。其抑制产生原因包括前突触释放物的改变，强烈的突触刺激或长期的弱突触刺激导致后突触接受体的密度改变。小脑细胞的长期抑制被认为

有利于运动神经的学习，海马体的长期抑制被认为与过去记忆的清除有关。

2.3 视觉感受野

2.3.1 经典感受野

人类的视觉系统中，感知识别等任务的实现是视觉系统在感受野的逐级提取过程中逐步提取关键信息而实现的。在视网膜上，当每一个神经元某区域受到光刺激时，能够激活影响神经细胞的活动，我们称此特定区域范围为被影响神经细胞的感受野，又叫经典感受野(CRF)^[39]。感受野是视觉系统中信息处理的基本单元，神经细胞层次不同其感受野特性表现出的特性也有所差异。一般来说，神经元的层次与其感受野的结构复杂性正相关。

1953 年 Kuffler 等^[40]在研究猫的视网膜神经节细胞的响应特性时，发现有一类神经节细胞，在光斑照射其感受野中心时，放电反应最强烈；在光斑照射在感受野中心外周时，放电反应减弱甚至更低。这类具有以上特性的神经细胞的感受野称作“ON”型中心感受野，由中心兴奋区和外周抑制区组成，在空间呈同心圆分布，功能表现为“中心-外周”的拮抗作用。另一类神经节细胞，在光斑照射感受野中心时，放电反应很弱；在光斑照射中心外周时，放电反应反而很强烈。此类具有以上特性的神经细胞的感受野称作“OFF”型中心感受野，其由中心抑制区和外周兴奋区组成，在空间分布与功能作用上类似于“ON”型中心感受野。

外侧膝状体细胞的感受野与神经节细胞的感受野相似，分为 ON 型中心和 OFF 型中心感受野，功能表现为“中心-外周”的拮抗作用。皮层细胞的感受野呈狭长形，是由若干个外侧膝状体细胞汇集到一个皮层细胞上而形成的视网膜上一个更大区域。皮层细胞的感受野具有开区、关区两个对抗区域，对光棒或线条具有强烈反应。

2.3.2 非经典感受野

李朝义等人^[41]在测量视网膜神经节细胞和外侧膝状体细胞的反应特性时，发现用光斑照射感受野中心时，细胞放电反应很强烈，光斑面积增大时，细胞的放电反应随之增强，显然，光斑的照射区域为感受野的中心兴奋区域；照射面积继续增大，细胞放电反应到一定程度时开始逐渐减弱，也就是说，光斑此时的照射面积已经超过兴奋区域到达了外周抑制区，抑制区产生的抑制作用致使放电反应的减弱；照射面积继续增大，细胞的放电反应又开始增强。为了解释这个现象，经过大量实验研究，发现在感受野的外周抑制区之外还存在一个更大范围的对经典感受野具有去抑制作用的区域，并将这个区域称为非经典感受野(NCRF)，也叫做整合野(Integrating Field, IF)。

2.3.3 拮抗式感受野

研究已表明视网膜神经元存在着两种不同类型—兴奋性(ON 型)和抑制性(OFF 型)，ON 型神经元对与光刺激会产生兴奋，而 OFF 型神经元对光刺激则会产生抑制。ON 型和 OFF 型神经元在空间上分布均匀，在一定范围内相互作用构成了两种类型的拮抗式感受野—中心 ON 周围 OFF 和中心 OFF 周围 ON。

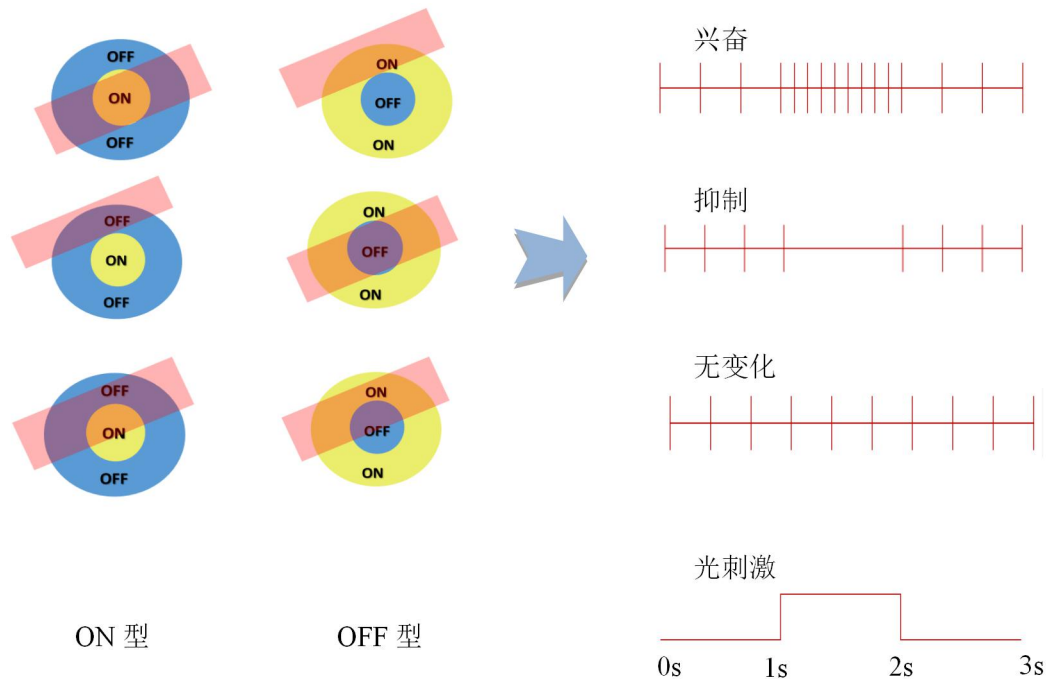


图 2-4 感受野特征性反应

在视觉感知系统中，感受野的结构使得神经元的脉冲发放必然受到刺激源与视野空间信息的综合作用。根据感受野的脉冲发放特点可以对神经元类型进行判断，如图 2-4 所示，在感受野中心受到刺激时，给定的时间窗内计算神经元的脉冲发放次数及全网络的平均发放次数，当神经元的脉冲发放次数大于全网络的平均发放次数时，表示神经元处在兴奋状态，即为 ON 型；当神经元的脉冲发放次数小于网络的平均发放次数时，表示神经元处在抑制状态，即为 OFF 型。据此给出了神经元激励电流的权值调整模型^[42]，表明神经元的输入刺激除了直接输入外，还受到周围神经元感受野的类型影响。研究表明，在图像处理中感受野特性不仅能够对噪声产生抑制作用，在弱边缘区域也能够对虚假边缘产生抑制，从而增强了对弱边缘信息的凸显，使得图像的轮廓检测效果更好。

2.4 神经元模型及编码

神经系统的结构和功能^[43]单位是神经细胞，即神经元。神经元机能最大特点是具有传递信息的绝缘性和极性，信息的传递主要依靠于神经元之间的突触进行以电信号的形式来完成的，轴突为其输入端，树突为输出端，电信号则以脉冲的形式来传导。可见神经系统的信息传递、编码与整合是通过放电活动来完成的，其放电脉冲形式与神经元受到的外界刺激有着极大联系，表明了神经系统有极为丰富的非线性动力学行为。

动作电位是神经元对刺激反应的一种表现形式。生物电的产生依赖于细胞膜对化学离子严格选择性的通透性及其在不同条件下的变化。细胞生物电现象主要有静息电位和动作电位两种表现形式。神经细胞的静息电位是细胞在安静时，存在于细胞膜内外两侧的电位差，表

现为膜内侧带负电，膜外侧带正电；当表现为膜内侧带负电，膜外侧带正电时，称为极化现象。神经细胞的动作电位是神经元细胞发生兴奋是所具有的特征性表现，也就是说动作电位的产生意味着神经冲动，即兴奋。

2.4.1 Hodgkin-Huxley 神经元模型

Hodgkin-Huxley 模型^{[44][45]}具备神经电生理特性，能够较好地描述神经元细胞膜电位与膜电流的变化过程。Hodgkin-Huxley 脑细胞神经元具有可激励性，当神经元细胞受到小的扰动时，能够迅速恢复到静息状态，但当神经元细胞所受扰动超过某一个阈值时，会产生一个快速的响应，转变为激发状态。HH 模型以上特性与动作电位的产生和传递规律相符合，描述了细胞膜上离子通道的开闭情况，所以其数学模型可以较好地用来模拟视觉系统的神经元的活动。

HH 数学模型是一组可以用来描述神经元细胞膜的电生理现象的非线性微分方程，其表达式如式(2.1)所示：

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

式中， I_{ext} 为外部电流输入； V 为膜电压； C_m 为膜电容； g_{Na}, g_K, g_L 分别代表钠离子、钾离子和漏电流通道电导最大值， V_{Na}, V_K, V_L 分别代表相应通道的逆转电位； m 和 n 为钠离子通道的门控变量， h 为钾离子通道的门控变量； $m_\infty(V), h_\infty(V), n_\infty(V)$ 和 $\tau_m(V), \tau_h(V), \tau_n(V)$ 分别为 m, h, n 的饱和值和时间常数；其模型描述的神经元细胞膜电位变化如图 2-5 所示，图中一个尖脉冲的发放对应于一个动作电位的产生。

$$I_{ext} = I_0 + S(t) + \xi(t) \quad (2.2)$$

式中， I_0 为恒定刺激，用来模拟一种总和输入； $S(t)$ 是检测对象； $\xi(t)$ 为高斯白噪声，满足 $\langle \xi(t) \rangle = 0, \langle \xi(t_1) \xi(t_2) \rangle = 2D\delta(t_1 - t_2)$ ， D 为噪声强度。

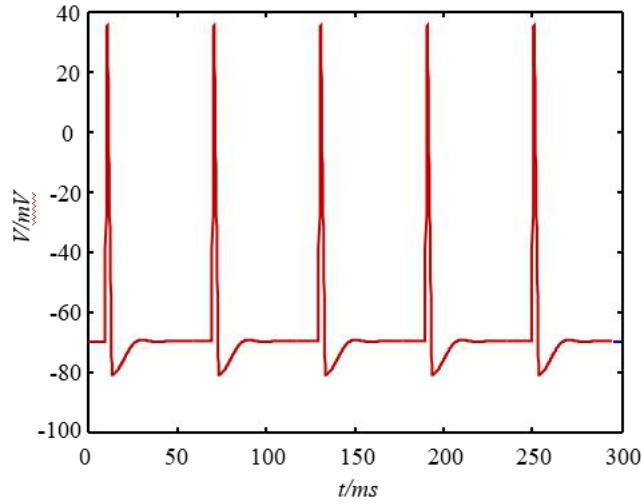


图 2-5 HH 神经元细胞膜电位变化示意图

2.4.2 Integrate-and-Fire 神经元模型

IF 神经元模型^[46]描述了电位的产生是基于神经细胞膜的电容特性，当输入电流渐渐累积到一定程度时就会放电。该模型通过设定一个电压阈值来控制神经元细胞放电脉冲，当膜电位超过设定阈值便产生脉冲，之后快速恢复到静息状态。IF 模型不仅能够形象模拟描述神经元动作电位的放电特性，而且其数学表达式较为简洁，在实际研究工作得到了广泛的应用。IF 模型的数学表达式如式(2.3)所示。

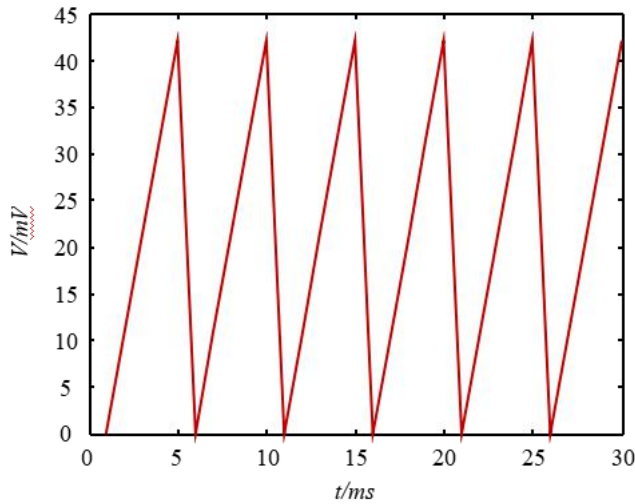


图 2-6 IF 神经细胞膜电位变化示意图

$$\begin{cases} dv/dt = I_{ext} + a - bv \\ \text{if}(v \geq v_{thresh}), \text{then}, v \leftarrow c \end{cases} \quad (2.3)$$

式中， v 是细胞膜电位， I_{ext} 是输入刺激， a, b, c 为模型参数， v_{thresh} 为阈值。如果 v 大于等于 v_{thresh} ，神经元发放脉冲，同时 v 被重新设定为 c ，如果 v 未超过阈值 v_{thresh} 那么神经元不

发放 spike; IF 模型描述的神经细胞膜电位变化如图 2-6 所示。

2.4.3 Leaky Integrate-and-Fire 神经元模型

漏积分放电(Leaky Integrate-and-Fire, LIF)神经元模型^[47]是在 IF 模型的基础上,从模型的计算实用角度和生物特性角度进行了平衡化,即在模型中重新加入神经元细胞的漏电流特性和不应期特性,其数学表达式如式(2.4)所示:

$$\begin{cases} c_m \frac{dv}{dt} = -g_l v + I_{ext} \\ \text{if } v \geq v_{thresh}, \text{ then } v = \text{const} \end{cases} \quad (2.4)$$

式中, v 表示神经元膜电压, c_m 为膜电容, g_l 为漏电导, I_{ext} 表示突触输入电流, 当 v 大于 v_{thresh} 时, 神经元脉冲发放, 同时 v 被瞬间重置为 const 。LIF 模型描述的细胞膜电位变化如图 2-7 所示。

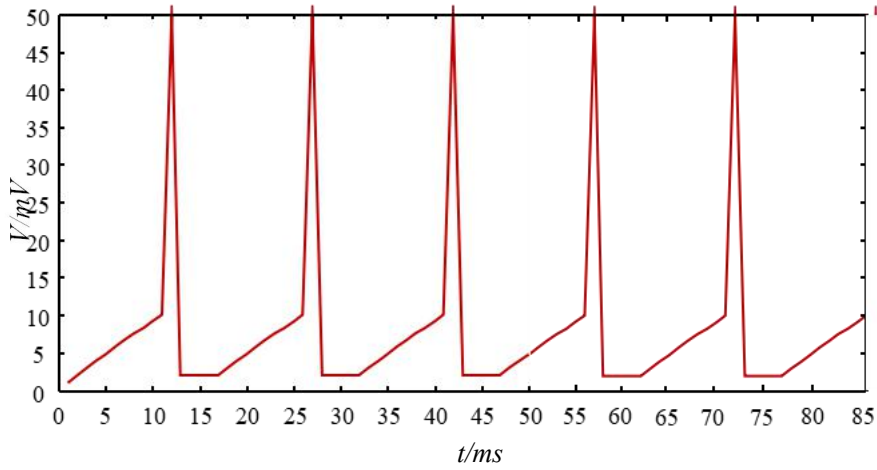


图 2-7 LIF 模型的细胞膜电位变化示意图

考虑了到更多的神经元电生理特性,在模型中引入绝对不应期,在具备脉冲发放特性、计算简单高效的 LIF 模型上进行改进,改进后 LIF 模型如式 (2.5) 所示。

$$\begin{cases} c_m \frac{dv}{dt} = -g_l v + I_{ext} (ref = 0) \\ v = v_{reset}, ref = ref - 1 (ref \neq 0) \\ v = v_G, ref = \text{const} (v > v_{th}) \end{cases} \quad (2.5)$$

式中, v 、 c_m 、 I_{ext} 、 v_{thresh} 的含义与式(2.4)一致, v_{th} 为脉冲发放阈值, v_{reset} 为静态电势, ref 为绝对不应期, v_G 为脉冲发放峰值。当 v 大于 v_{th} 时, 神经元脉冲发放达到 v_G , 同时 v 又被瞬间重置为 v_{reset} , 开始进入绝对不应期, 并且在绝对不应期内无论刺激电流多大, 神经元都不能被激活, 直至不应期结束。

2.4.4 脉冲编码

神经元在外界刺激和内在耦合下产生放电活动，并完成神经元细胞间的信息传递，同时表现出了其特定的空间模式和时间模式。神经元模型对其放电活动进行了模拟描述，而其特定的时间与空间模式表现的是视觉感光层对于图像的特征（颜色、边缘、纹理、轮廓等），为了描述这些特征，需要将其模拟量转化为神经元的动作电位序列，这就需要通过神经元的编码机制^[48]来实现。

在大脑皮层，连续的动作时间序列没有特定规律性，需要根据图像特征采取不同的编码方式。常用的方法有以下两种，一是根据神经元的脉冲发放，将其脉冲的首次发放时间或其脉冲发放的平均频率作为编码特征；二是依据外界刺激移动达到一定时长后，神经元脉冲发放因此而产生的差异性，计算神经元细胞响应的波动系数，再基于其波动系数大小进行编码。

2.5 小结

本章围绕视觉系统的感知特性，主要介绍了视觉通路、突触特性、视觉感受野、神经元模型及神经编码特性。视觉通路是视觉信息传递的依赖，本章从视网膜、外侧膝状体 LGN 及视皮层三个模块的结构与功能方面进行了阐述；视觉信息能够得以顺利传递，突触起到了关键作用，分别从结构与特性两个方面进行了描述；在信息传递过程中信息的有效性与感受野息息相关，所以对感受野的工作机制及特性进行了详细说明；视觉信息的提取离不开神经编码的作用，对神经编码特性及神经元模型展开了介绍。其中，神经元模型主要介绍了 HH 神经元模型、IF 神经元模型和 LIF 神经元模型，HH 神经元模型能够完整精确地描述神经元动作电位产生的全过程，但其计算量大且过程复杂，不具备普遍适用性；LIF 神经元模型是 IF 神经元模型的平衡化模型，相对于 HH 神经元模型，丢失了很多神经元特性，但保留了神经元的基本特性，并且具备较好的计算性能，因此被广泛应用于大型神经元网络的组建。

第3章 基于频域融合编码与前馈机制的图像轮廓检测

正如前述，视觉感知系统具有清晰的层次结构，视网膜上视椎和视杆细胞将外界光刺激转换为电信号，然后以压缩映射的方式传递至外膝体进行处理，最后视觉信息流进入视皮层，并与大脑其他功能区域联合完成视觉的感知任务。传统观点认为视觉信息流就是按这样的方式进行逐级编码，由信息流传递中的各个节点完成特定的编码任务。但必须要指出的是，由于神经编码的基础在于神经元的动作电位发放，表现为一种脉冲序列的形式。而动作电位的产生与神经元膜上的钾钠等离子通道闭合有直接关系，神经元在脉冲发放时还存在着固定的不应期，另外神经元之间是通过突触传递电信号或者化学信号。所以显而易见的是，脉冲序列的产生和传递都是需要消耗大量的时间。研究表明单个神经元的脉冲发放频率仅为几十赫兹，如果考虑到多级传递的因素，整个视觉感知系统对于视觉信息的处理效率将会变得非常低下，这与人类视觉系统的高效性能并不吻合。

越来越多的视觉神经实验表明，视觉信息流在传递的时候，视通路既有层次特点，同时还可能存在多条路径，其中部分路径能够将视觉目标的整体信息快速前馈至视皮层，也有路径能够精细化地提取视觉信息的细节部分，最后在更高级的皮层进行融合。这样对于细节部分而言，可能需要消耗一定的识别时间，但在细节部分缺失的情况下，整体信息已经能够在一定程度上反映视觉目标的特征，此时视觉系统的快速识别或检测将变得可能。因此，本章将试图来探索这种视觉机制在图像处理应用中的可行性，以图像的轮廓检测为例，研究如何利用感受野特性来快速获取初级视皮层的轮廓响应，同时重点研究外膝体节点在视觉信息流处理中的分频作用，然后针对分别表达信息局部和全局特性的高频和低频信号，研究如何对分频后的信号采取相应的编码机制？最后本章还将研究如何对不同路径传递的视觉信息进行前馈方式的融合。

3.1 基本原理

初级视皮层中具有快速获取图像轮廓整体信息的路径，类似于一种轮廓粗检测的过程，本章将这种检测结果称为初级轮廓感知响应。考虑到初级轮廓感知响应体现了神经节细胞感受野对于外界光刺激的一种方向选择性，而二维高斯导函数既能反映感受野尺寸与检测性能之间的关联性，又能体现对特定方向的选择特性，所以本章将利用二维高斯导函数与视觉信息的卷积来获取初级轮廓感知响应。对于视通路中外膝体节点类似于对视觉信息流的分频特性，本章将采用非下采样轮廓波变换的分频能力来模拟外侧膝状体的频域分离性能，提取视觉信息的高低频分量来分别重点反映轮廓的细节和整体特征。针对高频细节，提出采用冗余编码方式进行轮廓的强化，针对低频整体，采用稀疏编码方式进行处理，减弱背景纹理噪声对于轮廓识别的影响，最后融合高低频信息的编码结果，并将其作为初级轮廓感知响应的细

节补充，前馈到初级视皮层，实现图像轮廓的提取。

3.1.1 初级轮廓感知响应

初级视皮层神经节细胞感受野对特定方向的光刺激会产生剧烈反应并表现出选择性，利用二维高斯导函数 RF 与图像 $I(x, y)$ 之间的卷积来模拟初级视皮层神经节细胞对图像多方位的响应特性 $e(x, y; \theta_i, \sigma)$ ，如图 3-1 所示，然后选取 N_θ 个方位中的响应极大值形成视皮层初级轮廓响应 $E(x, y; \sigma)$ ，如式(3.1)和式(3.2)所示。

$$e(x, y; \theta_i, \sigma) = |I(x, y) * RF(x, y; \theta_i, \sigma)| \quad (3.1)$$

$$E(x, y; \sigma) = \max\{e(x, y; \theta_i, \sigma) | i = 1, 2, 3, \dots, N_\theta\} \quad (3.2)$$

式中， $*$ 表示卷积运算， σ 为相对应的感受野尺寸， N_θ 为方向选择中遍历角度 θ_i 的个数。二维高斯导函数 RF 数学表达式如式(3.3)所示：

$$\begin{cases} RF(x, y; \theta) = \partial g(\tilde{x}, \tilde{y}; \theta) / \partial \tilde{x} \\ g(\tilde{x}, \tilde{y}; \theta) = 1/(2\pi\sigma^2) \cdot \exp(-(\tilde{x}^2 + \tilde{y}^2)/2\sigma^2) \\ \tilde{x} = x \cos(\theta) + y \sin(\theta), \theta \in [0, \pi) \\ \tilde{y} = -x \sin(\theta) + y \cos(\theta), \theta \in [0, \pi) \end{cases} \quad (3.3)$$

式中， θ 表示不同的选择方位， σ 为标准差，决定非同心圆感受野的尺寸， γ 为空间大小比率，决定非同心圆感受野的椭圆率。对于某一特定方位 $\theta_i \in [0, \pi)$ ，计算如式 (3.4)。

$$\theta_i = (i-1)\pi / N_\theta, i = 1, 2, 3, \dots, N_\theta \quad (3.4)$$

式中， N_θ 表示提取方位信息的滤波器个数。

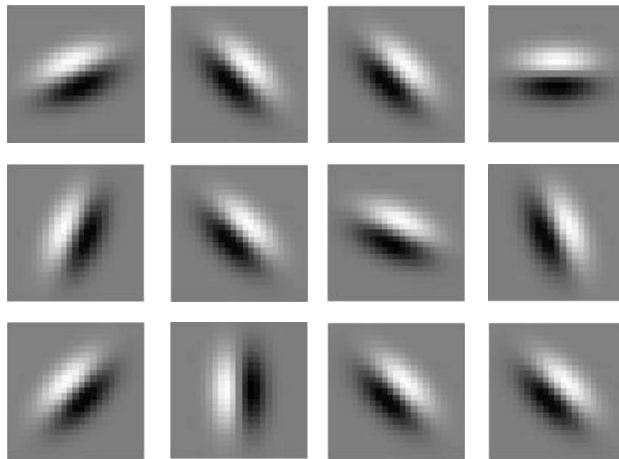


图 3-1 二维高斯导函数方向响应示意图

3.1.2 视觉信息分频编码

大量有关视觉感知机制的研究表明，外侧膝状体除了具备视觉信息的中转站的功能外，还具备视觉信息频域处理等更高级别的功能和作用。外侧膝状体结构中，前两层对应视网膜神经节 M 细胞，作用于亮度低频信息；后四层对应视网膜神经节 P 细胞，作用于颜色高频信息。有研究表明，视觉信号从外侧膝状体到初级视皮层的传递过程中，其信号一直处于并行传递状态，直到视皮层才进行融合。在图像处理中，图像的低频信号和高频信号也叫做低频分量和高频分量，用以度量图像中各个位置间的强度变化，其中，低频分量主要用于对整幅图像的综合度量，高频分量主要用于对图像的边缘度量和轮廓度量。在灰度图像中，低频分量对应于像大片色区块这样变化平缓的地方，高频分量对应于灰度或亮度变化剧烈的地方。考虑到非下采样轮廓波变换(Nonsubsampled contourlet transform, NSCT)^[49]具有平移不变性、多方向特性和捕获图像边缘纹理等具备几何结构特征的能力，因此利用 NSCT 对图像信号进行分解来模拟 LGN 在视觉信息处理中的频域分离作用。

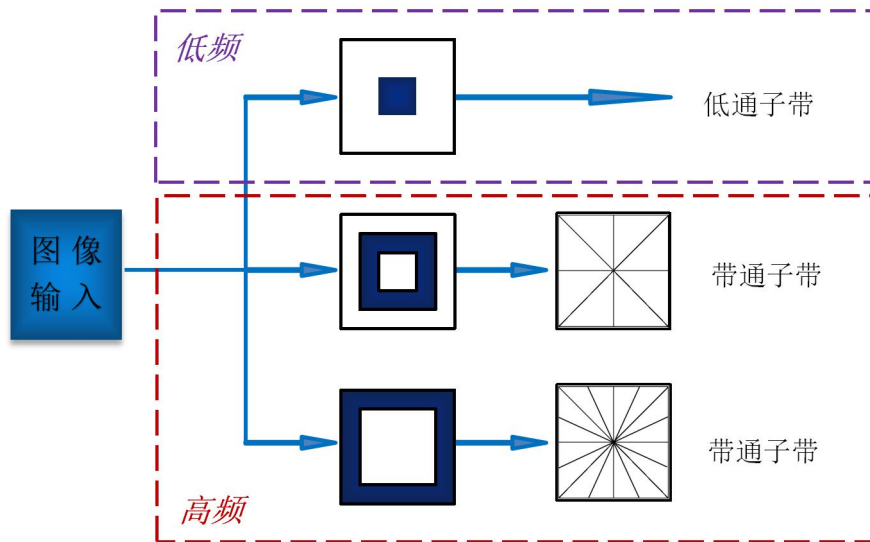


图 3-2 NSCT 频域分离示意图

NSCT 首先采用非下采样金字塔滤波器组(NonsubSampled Pyramid Filter Bank, NSPFB)对输入图像 $I(x, y)$ 进行多尺度分解，分解为高频子带图像 $h_1(x, y)$ 和低频子带图像 $b_1(x, y)$ ，即高频分量和低频分量，图 3-2 中的带通子带和低通子带；然后将得到的高频子带图像 $h_1(x, y)$ 传递给非下采样方向滤波器组(NonSubsampled Directional Filter Banks, NSDFB)分解为多方向子带图像 $h_2(x, y)$ ，再将低频子带图像 $b_1(x, y)$ 作为输入图像，重复 n 次上述过程，将所有得到的高频子带图像加和得到最终的高频分量图像 $h(x, y)$ ，低频子带图像加和得到最终的低频分量图像 $b(x, y)$ ，从而实现多尺度多方向的频域分解。NSPFB 和 NSDFB 的滤波器结构如图 (3-2) 所示，均为一组双通道的滤波器组，NSPFB 主要用于实现图像的多尺度分解，NSDFB 主要用于实现高频子带图像的多方向分解。

如图 3-3， $H_0(Z)$, $H_1(Z)$ 与 $U_0(Z)$, $U_1(Z)$ 表示分解滤波器， $G_0(Z)$, $G_1(Z)$ 与 $V_0(Z)$, $V_1(Z)$ 表示合成滤波器，它们满足 Bezout 恒等式，如式(3.5)和(3.6)。

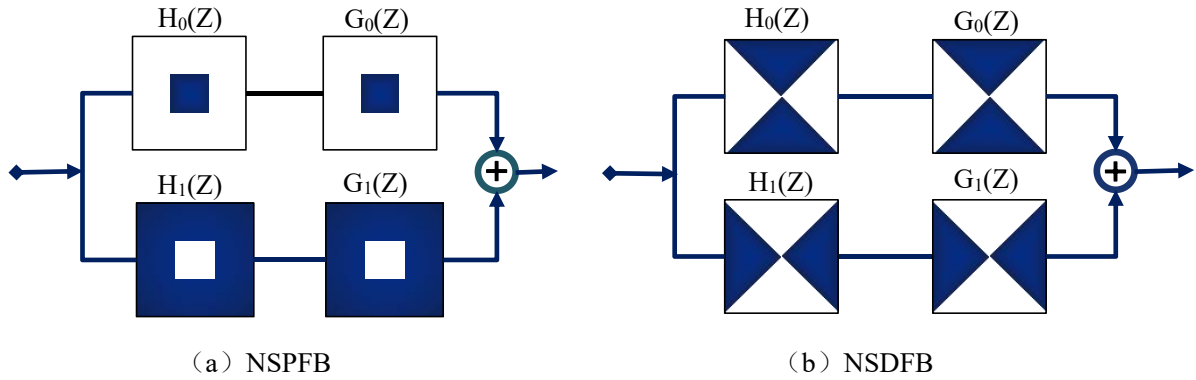


图 3-3 NSPFB 和 NSDFB 的滤波器结构示意图

$$H_0(Z)G_0(Z) + H_1(Z)G_1(Z) = 1 \quad (3.5)$$

$$V_0(Z)U_0(Z) + V_1(Z)U_1(Z) = 1 \quad (3.6)$$

在视通路中，信息的传递是在一系列的神经编码作用下完成的，不同的编码方式，在视觉信息的处理中发挥着不同的作用。因此，将经过 NSCT 分解得到的高频分量图像 $h(x, y)$ 和低频分量图像 $b(x, y)$ 分两通道并行编码处理。

考虑到稀疏编码能够整合空间信息，并且在信息传递处理中表现出去冗余化功能，能够从繁多的刺激中提取有效信息，而 NSCT 分解后得到的低频分量图像 $b(x, y)$ 信息复杂，所以采用稀疏编码对其进行处理，以达到去冗余化，提取有效轮廓的效果。

为符合稀疏编码规则，引入绝对不应期，引用具备脉冲发放特性、计算简单高效的 LIF 模型（如式(2.5)所示）构建具备稀疏特性的神经网络，对低频分量图像 $b(x, y)$ 进行编码，以图像某区域中某特定点的响应为中心，用局部窗滑动，以窗口空间信息作为神经元突触电流的输入，使得局部窗尺寸与步长相等，从而达到减小神经网络规模的目的，使其从空间上满足稀疏编码规则。令局部窗尺寸与移动步长相等，神经网络的突触输入电流 $i(x, y)$ 描述为：

$$i(x, y) = \text{mean}\left(\sum b[(x-1) \times \text{step} + w, (y-1) \times \text{step} + h]\right) \quad (3.7)$$

式中， step 为局部窗移动步长， w, h 分别表示窗口的宽和高。

再将突触电流 $i(x, y)$ 输入到式（2.5）所述的 LIF 神经网络模型，以脉冲发放频率为稀疏编码后的图像输出，记为 $L(x, y)$ 。

对于高频分量图像 $h(x, y)$ ，为了提高对图像轮廓检测的容错率和鲁棒性，采用冗余编码提高神经元响应数量来使得图像边缘对比度增大，从而使得特征信息冗余度增加，达到强化轮廓信息的作用，其编码特征表现为图像中边缘像素点的对比度增大，轮廓呈现加粗高亮，因此，冗余编码后得到的图像的空间大小与输入图像一致。

将高频分量图像 $h(x, y)$ 其作为输入，冗余编码后的图像记为 $H(x, y)$ ，如式（3.8）所示。

$$\begin{cases} H(2i-1, 2j-1) = h(i, j) \\ H(2i-1, 2j) = \left(\sum_{n=0}^1 h^2(i, j+n) \right) / \left(\sum_{n=0}^1 h(i, j+n) \right) \\ H(2i, 2j-1) = \left(\sum_{n=0}^1 h^2(i+n, j) \right) / \left(\sum_{n=0}^1 h(i+n, j) \right) \\ H(2i, 2j) = \left(\sum_{m=0}^1 \sum_{n=0}^1 h^2(i+m, j+n) \right) / \left(\sum_{m=0}^1 \sum_{n=0}^1 h(i+m, j+n) \right) \end{cases} \quad (3.8)$$

式中， i, j 分别表示图像 $h(x, y)$ 的长和宽。

利用双线性插值法将低频分量编码后的响应图像 $L(x, y)$ 放大到原图尺寸，与高频分量编码后的响应图像 $H(x, y)$ 进行融合，得到融合的轮廓特征响应图像 $S(x, y)$ ，如式 (3.9) 所示。

$$S(x, y) = H(x, y) \oplus \text{resize}(L(x, y)) \quad (3.9)$$

式中， resize 表示双线性插值放大， \oplus 表示高低频响应图像融合。

融合后的轮廓响应特征图像虽然轮廓响应明显，但仍存在背景纹理干扰，考虑到经典感受野 (CRF) 的增强作用和非经典感受野 (NCRF) 的纹理抑制作用，引入经典感受野与非经典感受野的差异性，对融合后的轮廓特征响应图像 $S(x, y)$ 进行抑制，如式 (3.10) 计算。

$$\begin{cases} \Delta\phi(x, y) = \lambda \times \max\{\bar{\phi}_C(x, y) - \bar{\phi}_N(x, y), 0\} \\ \lambda = \bar{\phi}_C(x, y) / \bar{\phi}_N(x, y) \end{cases} \quad (3.10)$$

式中， λ 为差异凸显系数， $\bar{\phi}_C(x, y)$ 为 CRF 空间视觉输入， $\bar{\phi}_N(x, y)$ 为 NCRF 空间视觉输入，分别用高斯函数 $G(m, n)$ 和 DOG^+ 模型为距离因素进行量化，如式 (3.11)。

$$\begin{cases} \bar{\phi}_C(x, y) = \sum_{(m, n)} S(x+m, y+n) \times G(m, n) \\ \bar{\phi}_N(x, y) = \sum_{(m, n)} S(x+m, y+n) \times DOG^+(m, n) \end{cases} \quad (3.11)$$

DOG^+ 模型，用以表征非经典感受野 (NCRF) 区域，是 DOG 函数^[50]经半波整流后的模型，描述如式(3.12)所示，

$$DOG^+(m, n) = \max\left\{ \frac{1}{2\pi(\rho\sigma)^2} \exp\left(-\frac{m^2+n^2}{2(\rho\sigma)^2}\right) - \frac{1}{2\pi\sigma^2} \exp\left(-\frac{m^2+n^2}{2\sigma^2}\right), 0 \right\} \quad (3.12)$$

式中， σ 为感受野大小， ρ 表示 NCRF 与 CRF 的大小比值，根据生理研究结果，NCRF 直径为 CRF 的 2 至 5 倍。

将引入的差异性做幂指数函数归一化处理，得到最终的融合编码轮廓响应特征图像 $Z(x, y)$ ，如式 (3.13) 所示。

$$Z(x, y) = \exp(-\Delta\phi(x, y)^2) / 2\xi^2 \quad (3.13)$$

式中, ξ 为抑制强度系数, 调节抑制强度, 取值范围一般为 0.2~0.8。

3.1.3 视觉前馈机制

根据大量的神经电生理特性研究及其仿真实验结果, 可知神经元脉冲的产生是一个受诸多因素影响的比较复杂的过程, 其影响因素包括细胞膜上离子通道的开合、浓度差诱导的离子跨膜运动、突触延迟效应以及脉冲发放存在不应期等, 此外, 神经元在激活状态下脉冲的发放频率达数十赫兹。而在真实视觉感知中, 对于视觉信息的处理是快速而高效的, 所以前后级神经元并不是只存在单一通路的联系, 而是存在多通路多形式的联系。在视觉生理和心理实验中, 都从侧面验证了视通路在视觉信息的传递过程中呈现多路径的特点。因而, 视通路上跨层级的信息传递路径有利于视觉信息的整体粗略检测的实现, 所以, 为了提高视觉系统对轮廓感知的准确性, 通过前馈机制将神经编码融合后的信息直接传递到初级视通路的末端, 得到最终的轮廓响应 $f(x, y)$, 如式(3.14)所示。

$$f(x, y) = E(x, y; \sigma) \times Z(x, y) \quad (3.14)$$

式中, $E(x, y; \sigma)$ 表示初级轮廓响应, $Z(x, y)$ 表示最终的融合编码轮廓响应特征图像。

最后对最终的轮廓响应 $f(x, y)$ 进行非极大值抑制和滞后阈值处理, 得到最终的图像的轮廓检测结果。

3.2 算法流程与验证

3.2.1 算法流程

本章通过视觉信息在是视通路中的传递方式和特性的研究, 模拟视通路对信息的处理形式, 提出基于频域融合编码与前馈机制的图像轮廓检测方法。具体检测步骤表述如下:

- (1) 引入二维高斯导函数 RF , 模拟视皮层是响应特性, 计算视皮层初级轮廓响应图 $E(x, y; \sigma)$;
- (2) 通过 NSCT 对灰度图像进行频域分离, 得到低频分量分离图像 $b(x, y)$ 和高频分量图像 $h(x, y)$, 然后采取并行编码的方式进行处理;
- (3) 针对低频分量图像 $b(x, y)$, 构建稀疏编码模型, 得到其稀疏特征响应图像 $L(x, y)$;
- (4) 针对高频分量图像 $h(x, y)$, 采用冗余编码的方式, 得到冗余特征响应图像 $H(x, y)$; 采用双线性插值法将稀疏特征图像 $L(x, y)$ 放大至原图大小, 然后与冗余编码后的图像进行融合, 得到新的轮廓特征响应图像 $S(x, y)$;
- (5) 针对得到的新轮廓响应图像 $S(x, y)$ 存在的背景纹理, 引入经典感受野与非经典感受野的差异特性进行抑制, 去除纹理干扰, 得到最后的融合编码图像轮廓响应图像 $Z(x, y)$;
- (6) 最后引入前馈机制, 将融合编码图像轮廓响应图像 $Z(x, y)$ 作用于初级皮层视通路的末端, 结合初级视皮层响应 $E(x, y; \sigma)$ 得到最终的轮廓响应 $f(x, y)$, 并对 $f(x, y)$ 进行非极大值抑制和滞后阈值处理, 提取图像主体轮廓。

轮廓检测流程如图 3-4 所示:

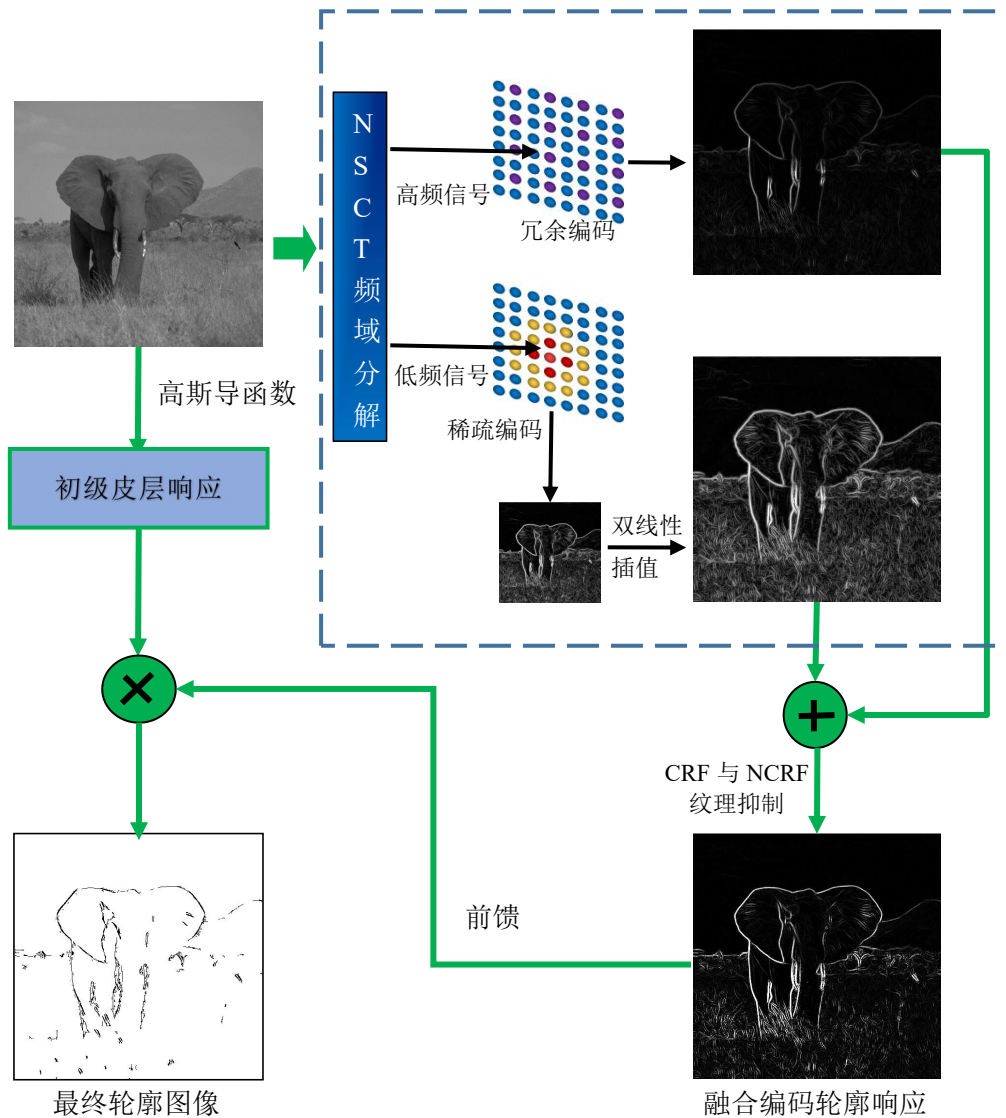


图 3-4 轮廓检测流程图

运用以上算法对自然图像 **Elephant** 进行轮廓检测提取，过程图如图 3-5 所示。输入图像如图 3-5 (a) 所示，利用多方向的高斯导函数实现初步轮廓信息的提取，得到初级皮层轮廓响应图 (d)，可看到高斯导函数提取的图像轮廓响应细节丰富的同时存在很多纹理边缘。考虑视觉信号从外侧膝状体到初级视皮层的信号传递过程中存在并行处理状态，对图像进行分频处理后，对高频分量图像冗余编码后得到 (b)，可看到高频信号冗余编码后主体轮廓边缘凸出明显，但其细节部分存在丢失；对低频分量图像稀疏编码后得到 (e)，可看到低频信号稀疏编码后的图像对轮廓边缘响应强烈，弱边缘部分也得到了响应；将主体轮廓响应强烈但细节丢失的 (b) 与主体轮廓与细节边缘皆有响应的 (e) 融合后，利用感受野差异性做抑制处理，得到图像 (c)，从中可看出融合并抑制处理后的图像，主体轮廓响应加强，背景纹理干扰得到了较好的抑制。将 (c) 前馈作用于初级皮层轮廓响应环节末端，得到最终轮廓图像

(f)。将经过编码融合及抑制处理的图像(c)与初级皮层响应图像(d)对比,从图中可以清楚观察到在主体轮廓的凸显上,(c)的轮廓凸显更完整,轮廓细节部分也得以凸显;从最终提取结果来看,经过前馈作用后的图像(f)在轮廓较为完整凸显的同时,背景纹理抑制效果明显。

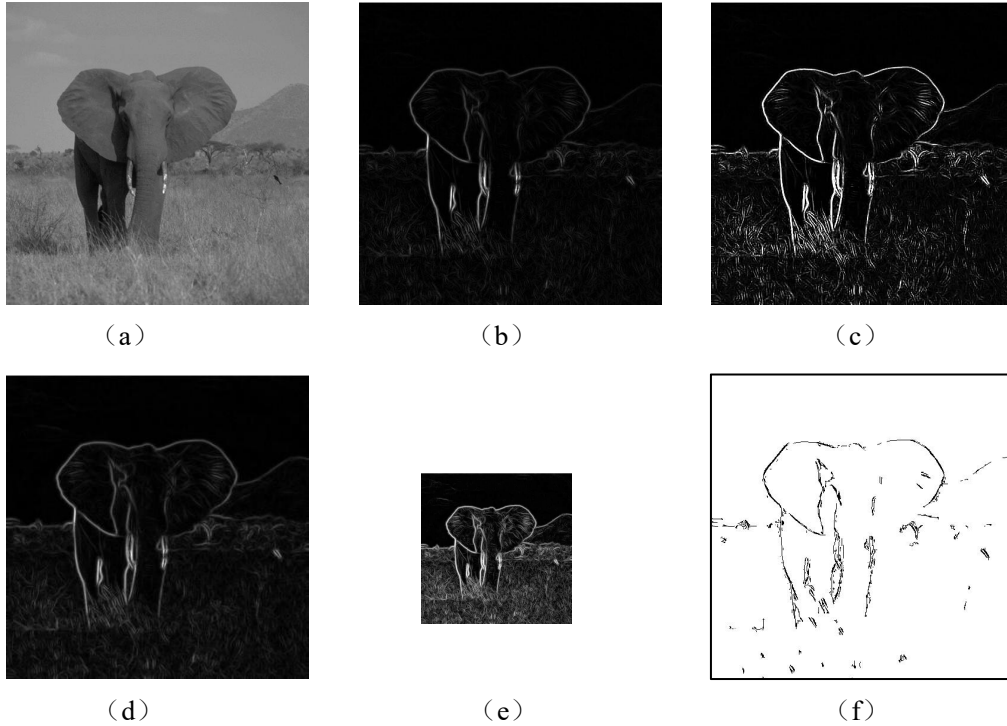


图 3-5 视觉信息处理示意图

(a)输入图像;(b)高频分量冗余编码图像;(c)融合编码图像;(d)初级轮廓响应图像;
(e)低频分量稀疏编码图像;(f)最终轮廓提取图像。

3.2.2 算法验证

为说明本方法在检测图像主体轮廓上所具备的优越性,本文选取了几种近年来基于生物视觉机制研究方面比较广泛应用的轮廓检测方法进行对比。首先,为体现本章方法在方向选择上符合视觉注意机制,选取基于外侧膝状体具备方向选择的轮廓检测模型(CORF)作为对比模型^[51];其次,为了体现本章方法在神经网络及稀疏编码效应的良好表现,选取空间稀疏限制方法轮廓检测模型(SSC)作为第二对比模型^[52];另外,为说明本章方法在纹理抑制方面具有一定优势,选取非经典感受野同步抑制方法(ISO)第三对比模型进行比较^[53]。对于本章方法,设置感受野尺寸为2;基于局部与全局的处理,在局部方向角矩阵处理之后,选取了改进的12个不同方位的二维高斯导函数($N_\theta = 12$);选取了10个不同强度的抑制系数 $\alpha = [1.1:0.1:2.0]$,滞后阈值处理部分,轮廓保留比例参数为 $p = [0.1:0.1:1.0]$,可得到100组不同参数进行测试。对于CORF模型,选取二维高斯导函数的标准差 σ 作为测试参数,取 $\sigma = [2.0:0.5:4.0]$;在滞后阈值处理部分,轮廓保留比例参数为 p ,可得到50组测试参数。对于SSC方法,因其不涉及全局系数,故选择10组参数进行测试。对于ISO方法选取与本

章方法一样的参数 α 和 p ，亦可得到 100 组参数测试。

为了客观的评价轮廓检测结果，本文采用 Grigorescu 等提出的轮廓检测定量指标^[52]来对以上三种轮廓检测方法以及本文轮廓检测方法进行定量分析。因为检测出的轮廓点与基准轮廓是不可能完全重合的，所以设定判定规则为：检测轮廓点若在基准轮廓点的 5×5 邻域内，则判定为正确，否则，判定为错误检测轮廓点。记轮廓像素集为 E_{BC} ，基准轮廓集为 E_{GT} ，定义正检集（正确检测轮廓点到的像素集） E 、误检集（错误检测轮廓点到的像素集） E_{FP} 、漏检集（漏检测轮廓点到的像素集） E_{FN} ，如式（3.15）所示。

$$\begin{cases} E = E_{BC} \cap (E_{GT} \oplus T) \\ E_{FP} = E_{BC} - E \\ E_{FN} = E_{GT} - E_{GT} \cap (E_{BC} \oplus T) \end{cases} \quad (3.15)$$

式中， \oplus 为膨胀操作， T 为 5×5 的结构单元。

引入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 ，误检率 e_{FP} 越小表明背景及纹理边缘越少，漏检率 e_{FN} 越小表明主体轮廓越完整， P 值越大表示检测出的轮廓图与基准轮廓图越接近。评价指标值按如式（3.16）~式（3.18）进行计算。

$$e_{FP} = \text{card}(E_{FP}) / \text{card}(E) \quad (3.16)$$

$$e_{FN} = \text{card}(E_{FN}) / \text{card}(E) \quad (3.17)$$

$$P = \text{card}(E) / [\text{card}(E) + \text{card}(E_{FP}) + \text{card}(E_{FN})] \quad (3.18)$$

式中， $\text{card}(S)$ 为集合 S 中的元素数量。当错误检测轮廓点越少，同时漏检轮廓点也少的情况下，轮廓的整体的检测效果是最理想的。从以上公式亦可以看出，误检率 e_{FP} 和漏检率 e_{FN} 同时越小，用来表示整体性能的指标 P 值就越大，检测的整体效果就越好。

3.3 实验结果

自然场景具有复杂的纹理噪声，存在光线多变等情况，因此，自然场景图片因其复杂度较高、综合背景因素强，识别难度高，成为图像轮廓检测的难点之一，此外，因其具备生物环境特性，普适性高，是图像轮廓检测的主要应用之一。基于以上考虑，为了证明本章方法的有效性，选取 RuG 数据库¹中的图片作为检测对象，该数据库中的所有图片，皆有手工绘制的轮廓基准图，所以能够客观地评价图像的轮廓检测效果。引入轮廓检测定量指标，对图像检测结果进行定量分析，选取数据库中部分图像轮廓的检测结果进行说明，部分检测结果如图 3-6 所示。

¹ 该图像库被广泛用于评价轮廓检测算法性能，可从网站 <http://www.cs.rug.nl/~imaging/> 获得。

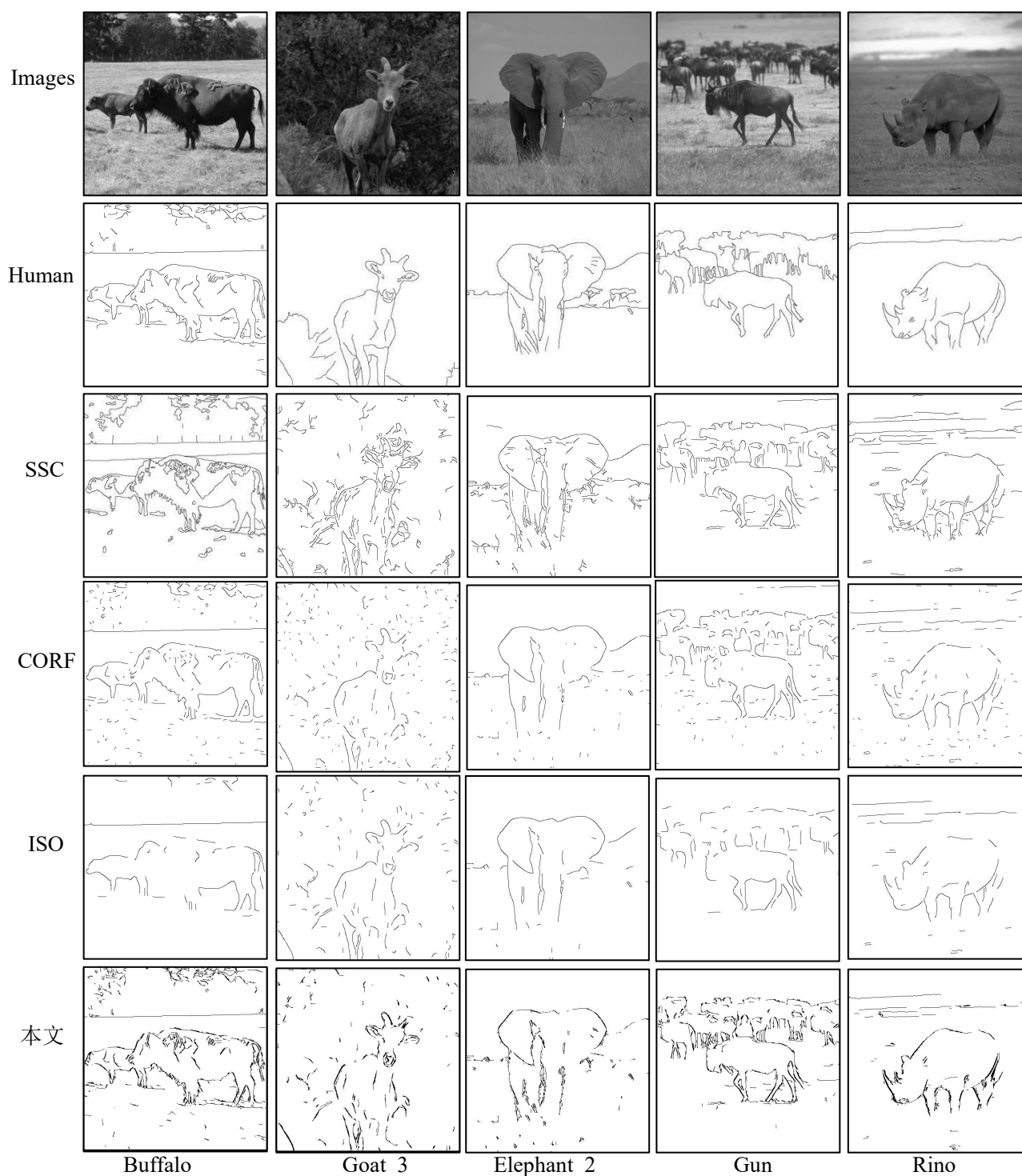


图 3-6 轮廓检测结果

运用以上算法性能评价指标的计算方法，对图 3-6 第一行图像中各算法轮廓检测最优结果进行计算，并记录检测结果最优时相对应的参数值，统计结果如表 3-1 所示。由表中数据可得，定量的性能指标可以表现检测图中的定性描述。为了更直观对检测结果指标进行比较，分别对图 3-4 中各图像的误检率 e_{FP} 、漏检率 e_{FN} 和综合评价指标作直方图统计，表示为图 3-7~图 3-9。

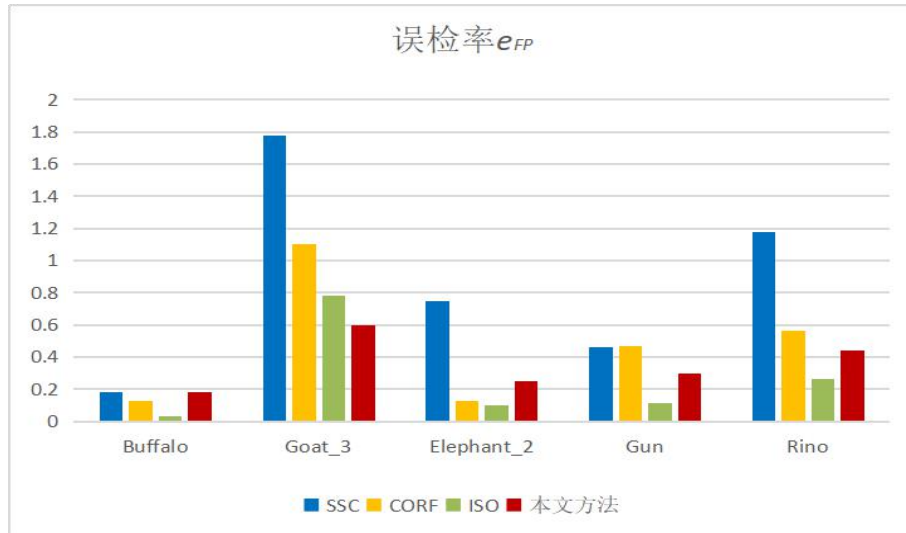


图 3-7 误检率

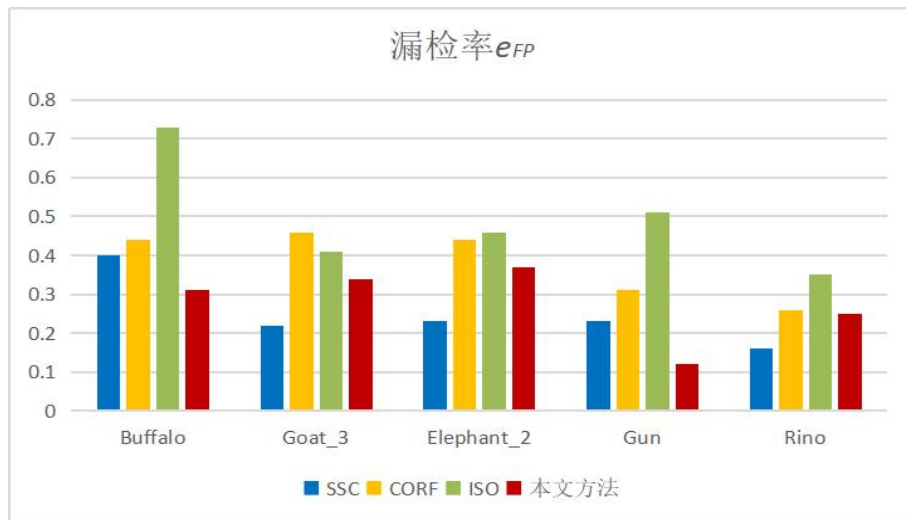


图 3-8 漏检率

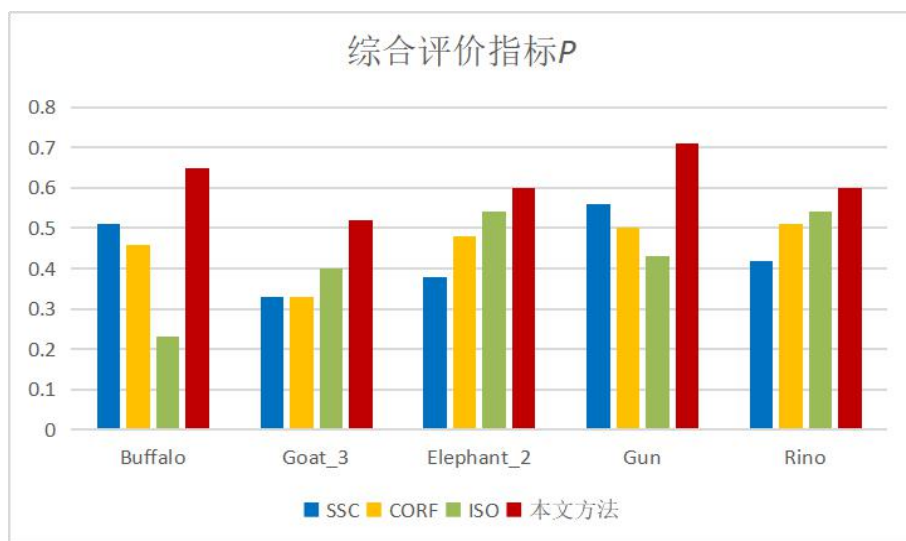


图 3-9 整体性能评价指标

以上方法中，在漏检方面 SSC 方法表现较好，其漏检率 e_{FN} 相对来说最低，表明该方法检测出了轮廓完整性好，但该方法的误检率 e_{FP} 相对来说表现为最高，说明该方法检测出的轮廓存在大量纹理干扰，结合检测图像，可以看出该方法在纹理抑制上表现并不理想，虽然主体轮廓完整性较高，但纹理干扰了其轮廓的凸显，使得检测效果并不理想。在误检方面，ISO 方法表现相对最好，表明该方法背景纹理抑制上表现较好，但其漏检率 e_{FN} 很高，说明出现了纹理过抑制，结合检测图像可以看出，图像轮廓细节部分丢失严重。相对于 SSC 方法与 ISO 方法来说，CORF 方法与本章方法在漏检和误检方面的平衡度上具备绝对优势。与 CORF 方法相比较，在误检方面，两种方法表现差不多；在漏检方面，本章方法表现较为优秀，表明本章方法在与 CORF 方法在检测轮廓完整度相当的情况下，本章方法对背景纹理的抑制表现

表 3-1 图 3-5 所示图像对应的参数设置及性能指标

图像	算法	参 数		性能指标		
		α	p	e_{FP}	e_{FN}	P
Buffalo	SSC	3.0	0.1	0.18	0.40	0.51
	CORF	3.0	0.9	0.13	0.44	0.46
	ISO	1.3	0.1	0.03	0.73	0.23
	本方法	1.6	0.2	0.18	0.31	0.65
Goat_3	SSC	3.0	0.1	1.78	0.22	0.33
	CORF	3.0	0.9	1.10	0.46	0.33
	ISO		0.1	0.78	0.41	0.40
	本方法	1.5	0.4	0.60	0.34	0.52
Elephant_2	SSC	3.0	0.1	0.75	0.23	0.38
	CORF	3.0	0.9	0.13	0.44	0.48
	ISO		0.1	0.10	0.46	0.54
	本方法	2.0	0.4	0.25	0.37	0.60
Gun	SSC	3.0	0.1	0.46	0.23	0.56
	CORF	3.0	0.90	0.47	0.31	0.50
	ISO	1.3	0.10	0.11	0.51	0.43
	本方法	1.6	0.1	0.30	0.12	0.71
Rino	SSC	3.0	0.1	1.18	0.16	0.42
	CORF	3.0	0.9	0.56	0.26	0.51
	ISO		0.1	0.26	0.35	0.54
	本方法	1.1	0.1	0.44	0.25	0.60

要更好。相应地，在同一图像的轮廓检测中，本章方法漏检率 e_{FN} 与误检率之间差值相对更小，说明本章方法的在漏检率 e_{FN} 与误检率 e_{FP} 的平衡上略胜一筹。结合检测图像亦可以看出，本章方法在背景纹理抑制方面确实表现更好。

3.4 讨论与分析

上述轮廓检测结果显示，本章方法不仅能够有效提取轮廓，而且在背景纹理抑制上表现优秀，对于轮廓细节丰富的自然场景图像具有很强的轮廓检测提取能力。

SSC, CORF 和 ISO 都从生物视觉机制出发，考虑外侧膝状体与初级皮层间的联系，但它们在并没有对输入前的灰度图像做预处理，而是将灰度图像直接作为外侧膝状体感受野的输入，忽略了外侧膝状体前后级间的联系，因而使得图像检测结果出现背景纹理抑制不好或过抑制，主体轮廓不能得到较好凸显。SSC 方法采用了一种基于数学模型构造的稀疏方法，主体轮廓特征明显，但其并没有真正引入视觉机制特性，因而在纹理抑制方面表现较差，检测出的图像轮廓存在背景纹理干扰，影响了图像主体轮廓的凸显。ISO 方法是利用 Gabor 函数能够提取多方位边缘特性来模拟视觉神经元的经典感受野提取图像轮廓，并结合各向同性与各向异性方法模拟视觉神经元的非经典感受野对背景纹理抑制，相对于 SSC 方法与 CORF 方法，其具有更强的纹理抑制作用，但却不能很好的区分背景纹理与轮廓细节，使得检测出的轮廓存在部分丢失，轮廓连续性不理想。CORF 方法基于视通路中 LGN 感受野模型，考虑了感受野的空间组合特性，虽然轮廓检测效果由于 ISO 方法和 CORF 方法，但其对背景纹理的抑制的表现还是有所欠缺。上述方法都是基于视觉特性提出的，都引入了神经元感受野特性，但却忽视了视觉信息传递过程中的神经编码方式，使得图像整体轮廓的检测效果不理想。本文方法从视觉感知角度出发，研究视觉信息在视通路中的传递形式及编码方式，有效提取在背景纹理抑制之前，对图像分频处理并分别引入冗余编码与稀疏编码两种不同的编码方式，使得对图像主体轮廓凸显加强的同时，能够有效提取轮廓细节部分；最后引入前馈作用方式，将通过背景纹理抑制后的轮廓响应图像前馈于初级视皮层轮廓响应部分，检测结果表明，本章方法不仅能够有效凸显图像的主体轮廓，而且能够对背景纹理抑制效果好，更符合人类视觉感知信息处理的生物机制。

3.5 小结

本章通过研究视觉系统视通路中视觉信号的传递以及处理的规则和特性，考虑外侧膝状体具有对视觉信息分类的特性，对图像进行分频处理，同时构建具备神经电生理特性的神经元网络，并根据生物机制特性采取不同的编码方式。首先通过高斯导函数对初级视皮层细胞感受野进行模拟，然后采取频域分解方法对高低频分量图像分别编码并融合，提高轮廓响应的准确性与有效性，根据外侧膝状体感受野特性，引入经典感受野与非经典感受野的差异特性，对融合响应图像进行纹理抑制，考虑到前馈作用对后续环节处理的重要作用，将经过融合抑制后的图像前馈作用于初级皮层视通路末端，得到最终的图像轮廓响应。使用 RuG 图像库，对自然图像进行检测，并通过与空间稀疏限制的纹理抑制 SSC 方法，基于非经典感受野

同步抑制 ISO 方法以及基于视通路中 LGN 感受野模型的 CORF 方法所获得的轮廓检测结果相比较,表明本章方法不仅能够有效凸显图像主体轮廓,在轮廓细节检测方面表现良好,而且对背景纹理产生了较佳的抑制效果,整体检测效果更为理想。

第4章 基于神经元感受野多方位选择机制的图像轮廓检测

如前文所述，初级视皮层细胞对有一定方位或朝向的亮暗对比边，或光棒、暗棒有强烈反应，即对边缘的位置和方位有强烈的选择性，这在视觉机制中对物体轮廓边缘的感知具有重要作用。但对物体的感知识别势必不仅仅于此，其间还有视觉通路上感受野的共同作用，有研究表明，感受野特性能够对图像纹理噪声及弱边缘区域的虚假边缘产生抑制作用，可见感受野对图像识别及理解的重要作用。视觉感知的实现是一系列的组织结构共同作用的结果，显然，感受野在视觉机制中扮演的角色不是单一的，而是具备复杂性和多面性的，还有很多潜在的功能作用值得去探索研究。因此，本章从人类视觉感知特性出发，基于视觉机制研究感受野的非对称方位敏感特性对轮廓信息提取的影响，探索一种方向选择机制的改进方法，并探究如何基于感受野的特性来实现其大小的自适应选择，使得轮廓的提取更智能化？最后结合方向响应加权处理方法及感受野抑制作用，凸显图像轮廓特征。

4.1 基本原理

基于以上考虑，深入研究了视觉系统中神经网络电生理特性及初级视皮层的感受野特性，本章提出一种基于神经元感受野多方位选择机制的图像轮廓检测方法，一方面通过对稀疏神经元进行编码，构建神经网络以提取初级视皮层边界响应图像，调整响应图像的方向权值占比以强化凸显轮廓的连续性；另一方面根据图像复杂度对感受野大小进行自适应选择，基于感受野的非对称方位敏感特性，对图像有效响应方位进行筛选，然后与调整后的响应图像融合，经过非极大值抑制及滞后阈值处理，得到最终图像轮廓。

4.1.1 改进的方向选择机制

视通路上的神经节细胞对视觉激励表现出一定的方向选择特性，也就是特定神经元的兴奋状态只与某一个方向的外界激励之间呈现出较高的相关性。传统上通常利用二维高斯导函数 RF 与图像 $I(x, y)$ 之间的卷积，并通过求取所谓的最佳方向响应 $E(x, y; \sigma)$ 来实现神经节细胞的方向选择机制，如式(4.1)和(4.2)所示。

$$e(x, y; \theta_i, \sigma) = |I(x, y) * RF(x, y; \theta_i, \sigma)| \quad (4.1)$$

$$E(x, y; \sigma) = \max(e(x, y; \theta_i, \sigma) | i = 1, 2, 3, \dots, N_\theta) \quad (4.2)$$

式中， $*$ 表示卷积运算， σ 对应于感受野的尺寸， N_θ 为方向选择中遍历角度 θ_i 的个数，一般为了获得精细的方向细节， N_θ 会选取得足够大，例如 8 或 12。

虽然上述卷积过程考虑了 RF 作用域内像素对于方向信息的加权累积，而参数 σ 则体现了局部与全局之间的平衡关系，因此对于当前像素点 (x, y) 而言，总能找到一个方向响应的极值。但需要指出的是，对于轮廓而言，最重要的应该是相邻像素之间的方向关联性或连续性（如

图 4-1 (a) 所示, 图像不同局部区域处于轮廓线附近的像素方向, 其方向朝向性分布有一定的规律; 如图 4-1(b)所示, 区域 A 中的轮廓线大致方向为左上到右下, 区域 B 中的大致为一条直线。), 而在上述过程中并没有反映这种特性, 所以, 图像中的纹理背景被误判为轮廓点是导致轮廓过检测的主要因素。因此本章将首先进行局部邻域内某个方位响应的一种累加统计, 并进而筛选出可能的轮廓细节点, 然后再进入下一轮的最优方向搜索。

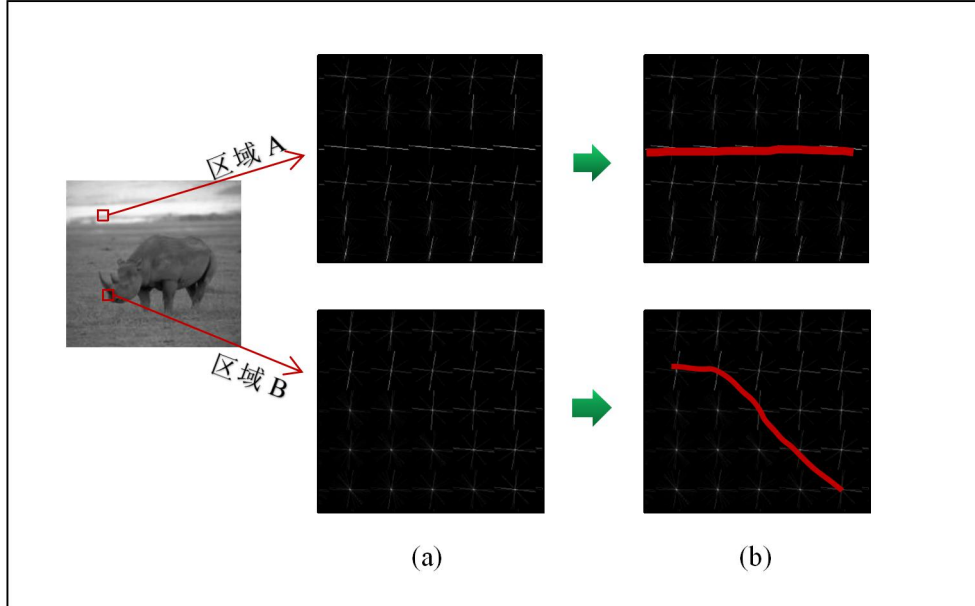


图 4-1 不同的局部区域像素点方向关联性及连续性示意图

采用如式(4.1)和(4.2)所示的操作, 但感受野尺寸 σ 将选取的相对较大, 这样, 既能从低分辨率的角度来过滤图像的背景细节, 又可以提高方向选择的快速性。感受野尺寸 σ 的选取, 本文根据不同图像的场景复杂度 P , 按式(4.4)计算, 来进行自适应选择, 若检测图像环境比较复杂, 选择覆盖至图像 $I(x, y)$ 的 P_1 (一般取 1/16), 以达到更好的纹理抑制效果; 若检测环境复杂度较低, 则选择覆盖图像 $I(x, y)$ 的 P_2 (一般取 1/25), 以达到检测轮廓更完整的目的。取样本数据库中场景图片按式 (4.4) 进行计算, 得到一个综合均值作场景复杂度的比较值, 记为 P_0 , 感受野尺寸 σ 按式 (4.3) 进行选择, 式中, $|I(x, y)|$ 为图像大小。

$$\begin{cases} \sigma = |I(x, y)| \cdot p_1, p > p_0 \\ \sigma = |I(x, y)| \cdot p_2, p < p_0 \end{cases} \quad (4.3)$$

检测图像的场景复杂度 P 按式 (4.4) 进行计算:

$$p = \sum_{i=1}^{N_\theta} A_{\theta_i} [e_i(x, y; \theta_i, \sigma)] / N[I(x, y)], \quad i = 1, 2, 3, \dots, N_\theta \quad (4.4)$$

式中, A_{θ_i} 为图像 θ_i 方向上的有方向响应的像素点个数, N_θ 为方向选择中遍历角度 θ_i 的个数, $N[I(x, y)]$ 为整幅图像的像素点个数总和, P 为有方向响应的像素点在整幅图像中的占比。

设经过感受野尺寸 σ 选择后的局部邻域为 $R^{m \times n}$ ， m 和 n 分别表示以当前像素为中心的矩形邻域宽度和长度。以一维矩阵 Z 作为方位响应的累加统计结果，其中 Z 的维数取决于 N_θ ，在筛选可能轮廓细节点时，为了提高角度区间的命中率，方位个数一般设置较少，例如 $N_\theta = 4$ 。按式(4.5)和式(4.6)进行方位响应的累加统计。

$$e(x, y; \theta_j, \sigma) > e(x, y; \theta_i, \sigma), \quad i = 1, 2, 3, \dots, N_\theta, \quad i \neq j \quad (4.5)$$

$$Z[j] = Z[j] + e(x, y; \theta_j, \sigma) \quad (4.6)$$

如果在某个方位上的最优响应累加值较为显著时，可通过其与所有累加值之和的比值来判断，如式(4.7)所示的条件成立，那么可将 $R^{m \times n}$ 的中心点判断为可能的轮廓细节点，其中 μ 为可控的比值参数。通常目标的轮廓细节在方向上具有一定的连续性，因此在图像中的小邻域会表现出这种方位的显著性；而对于背景中的纹理干扰，它们方向上较为随机，一般会表现出一种方位分布上的均匀性。

$$Z[j] > (\sum_j Z[j]) \times \mu \quad j = 1, 2, 3, \dots, N_\theta \quad (4.7)$$

当对图像 $I(x, y)$ 所有的像素点完成上述筛选后，将其中可能的轮廓细节点继续按传统方向选择机制进行方位上的精细检测，如式(4.1)和(4.2)所示，此时式(4.1)中的 $I(x, y)$ 将是筛选后的有效轮廓候选点。

4.1.2 非对称的方位敏感性感受野

在视通路中，视皮层的神经元树突分布形式也起着方位敏感性的作用。以单细胞的感受野为例，传统上为了计算简便的需要，会假设所有树突的接收权值都等于1。虽然这样的假设体现了不同方位感受野的朝向性，并能表现出感受野的兴奋性树突接收区域和抑制性树突区域，如图1所示的0度方位情形，红色、蓝色、白色区域分别表示兴奋性、抑制性以及零接收树突。对应的计算方式如式(4.8)所示，其中 y_c 对应于零接收树突区域的纵向位置。

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

但值得注意的是，以上仅考虑了单细胞感受野的朝向性及突触接收域的兴奋抑制性，而忽略了感受野之间的关联性及其突触接收域之间的相互影响。所以图像轮廓检测中将轮廓点判为背景纹理而去除成为造成轮廓欠检测的因素。虽然上述树突极性分析计算中考虑了不同敏感方位的单细胞感受野对神经元树突的接收极性具有不同方位的朝向性分布，可以形成多方位的单细胞感受野。但考虑到在像素点邻域内，兴奋性树突接收区域与抑制性树突接收区域存在多个单细胞感受野，当感受野一起作用时，彼此间会出现竞争从而削弱了区域响应强度。

因此本文对图像某一像素方向响应邻域内，以敏感方位为界限，根据树突接收性质赋予不同的权值，以放大像素点邻域内对轮廓点方向的兴奋性响应作用，而减弱其抑制作用，实

现一种非对称性的方位敏感感受野，从而达到对轮廓点更好的检测效果。因此记兴奋性接收树突的权值为 λ ($\lambda > 1$)，而其他情形树突的权值保持不变。此时树突的综合作用如式(4.9)所示，其中 $\delta_{x,y}^\tau$ 表示在方位为 τ 的情况下，加权处理后的树突兴奋或抑制作用。而对应的轮廓边界响应 $g(x,y)$ 如式(4.11)所示。

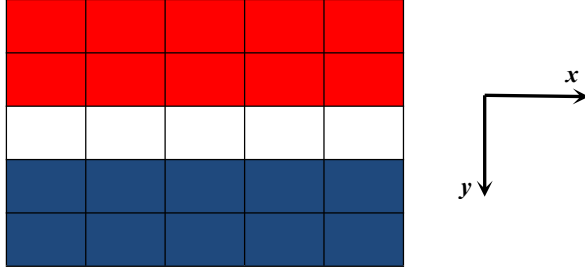


图 4-2 树突极性分布示意图(方位 0 度)

方位为 τ 时，记树突的兴奋或抑制接收区域为 (x,y) ，感受野的中心坐标为 (x_c, y_c) ，定义其最佳响应点权重为 $\delta_{x,y}^\tau$ ，如 (4.9) 所示：

$$\delta_{x,y}^\tau = d(\tau) \cdot \sum_{i=1} \delta(x_i - x_c, y_i - y_c; \lambda) \quad (4.9)$$

式中， (x_i, y_i) 表示接收区域第 i 个响应角所在的方位矩阵坐标， d 表示方位角数值化。

$$\delta_{x,y} = \sum_{\tau} \delta_{x,y}^\tau \quad (4.10)$$

$$g(x,y) = \delta_{x,y} \times E(x,y;\sigma) \quad (4.11)$$

将筛选后的轮廓候选点 $I(x,y)$ ，与最终轮廓边界响应融合后，得到轮廓检测图像 $f(x,y)$ ，如式(4.12)所示。

$$f(x,y) = g(x,y) \times I(x,y) \quad (4.12)$$

从多方位对轮廓检测图像 $f(x,y)$ 进行非极大值抑制，再经过滞后阈值处理，得到总体轮廓的最佳方向响应，从而得到最终轮廓图像检测结果。

4.2 算法流程与验证

4.2.1 算法流程

本章从视觉感知特性，对神经元感受野方位选择特性深入研究，提出了基于神将原感受野多方位选择机制的图像轮廓检测方法，具体轮廓检测步骤如下：

(1) 采取双通道的处理模型，一通道对输入图像进行稀疏编码，在默认设置的感受野尺度下，得到稀疏编码下的特定轮廓响应图像；

(2) 对于特定轮廓边界响应图像，引入二维高斯导函数得到最佳方位响应图像，同时引入感受野的非对称方位敏感特性，针对响应的方向特性调整响应方向的权值占比，以更好地

凸出主体轮廓；

(3) 另一通道基于图像复杂度构建感受野大小自适应选择机制，计算图像复杂度，调整选择合适感受野尺度，计算得到有效响应候选轮廓点；

(4) 最后将有效响应候选轮廓点与经过响应方向的权值占比调整后轮廓相融合，经过非极大值抑制及滞后阈值处理，得到最终的图像轮廓检测结果图。

具体轮廓检测流程如图（图 4-3）所示：

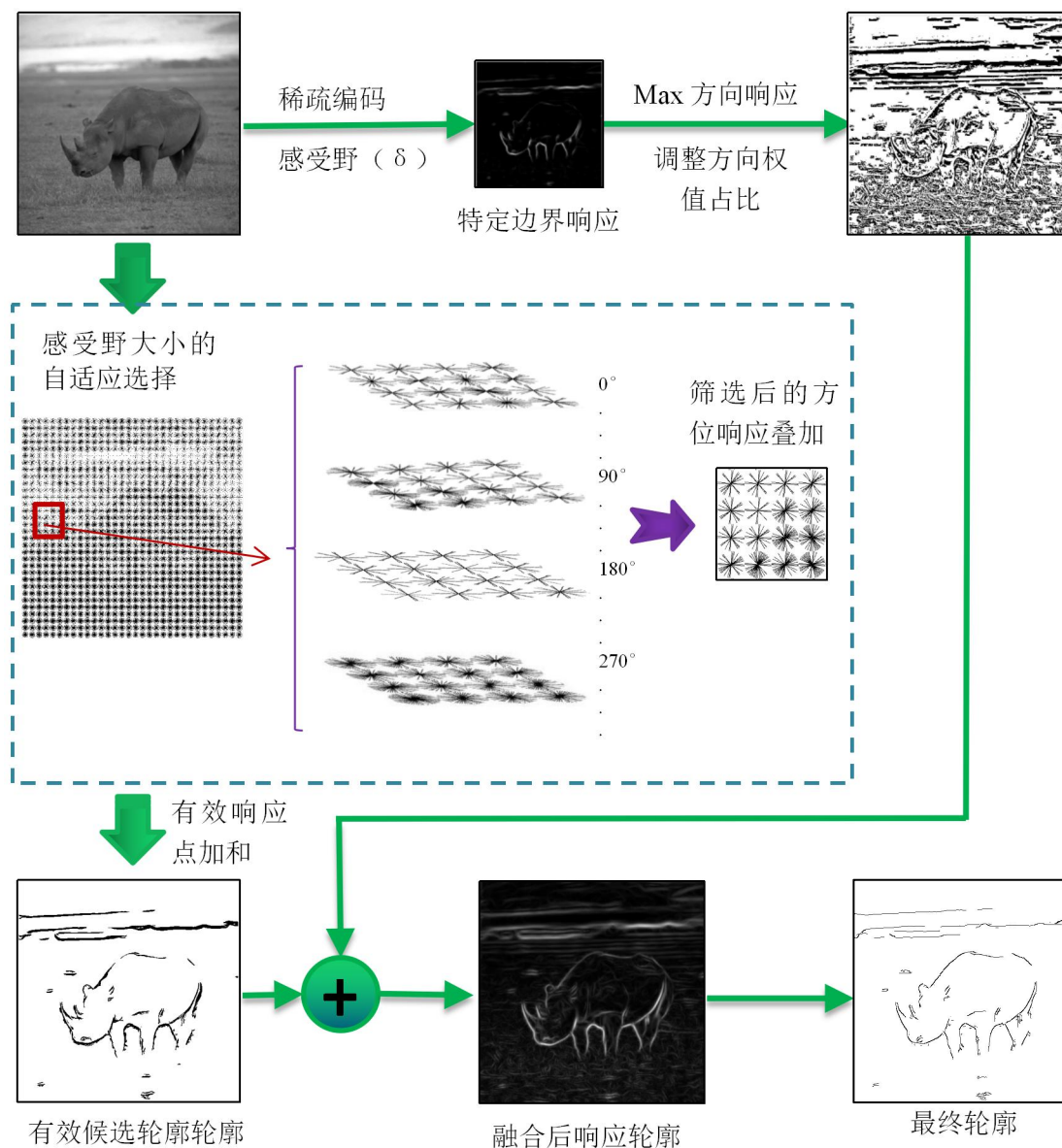


图 4-3 轮廓检测原理流程图

输入图像如图 4-4(a)所示。输入图像经过多方位的二维高斯导函数与图像卷积得到一个特定边界响应，该边界对纹理边缘有着强烈的方位响应；根据该边界响应强度，引入模拟高

级视皮层神经元稀疏编码，抑制纹理干扰，得到图像的初级轮廓响应（图 4-4（b）），从图中可看出，经由神经编码得到的响应图像很好的凸出了图像轮廓边缘；考虑到图像不同方位感受野朝向性及突触接收域间的影响，为了能够更好凸显轮廓边缘，本文对图像像素点响应的树突接收权值进行了加权处理（图 4-4（c）），可观察到图像的整体轮廓得以凸显；此外，考虑到局部与全局的关系及非对称方位敏感性感受野特性，为能够更好的去除纹理边缘影响，提取更加完整的图像轮廓，本文从方位角度有效响应轮廓点角度出发，对图像进行了局部到整体的自适应检测筛选，来去除杂乱纹理的干扰；再结合感受野特性在多方位响应角度的树突接收权值加权结果进行更深层次研究改进得到融合图像（图 4-4（d）），图中清楚的显示了轮廓的连续性，为图像轮廓的高效提取奠定了基础；把融合后的图像经过非极大值抑制及滞后阈值处理，得到最终检测轮廓图像（图 4-4（e））。轮廓检测结果图中，可观察到图像的整体轮廓得以凸显，杂草等构成的杂乱纹理边缘干扰得到了较好的抑制。由此说明，基于神经元网络与多方位敏感性感受野的图像轮廓检测符合人类视觉感知机制，可以在此方向上展开更深入的研究。

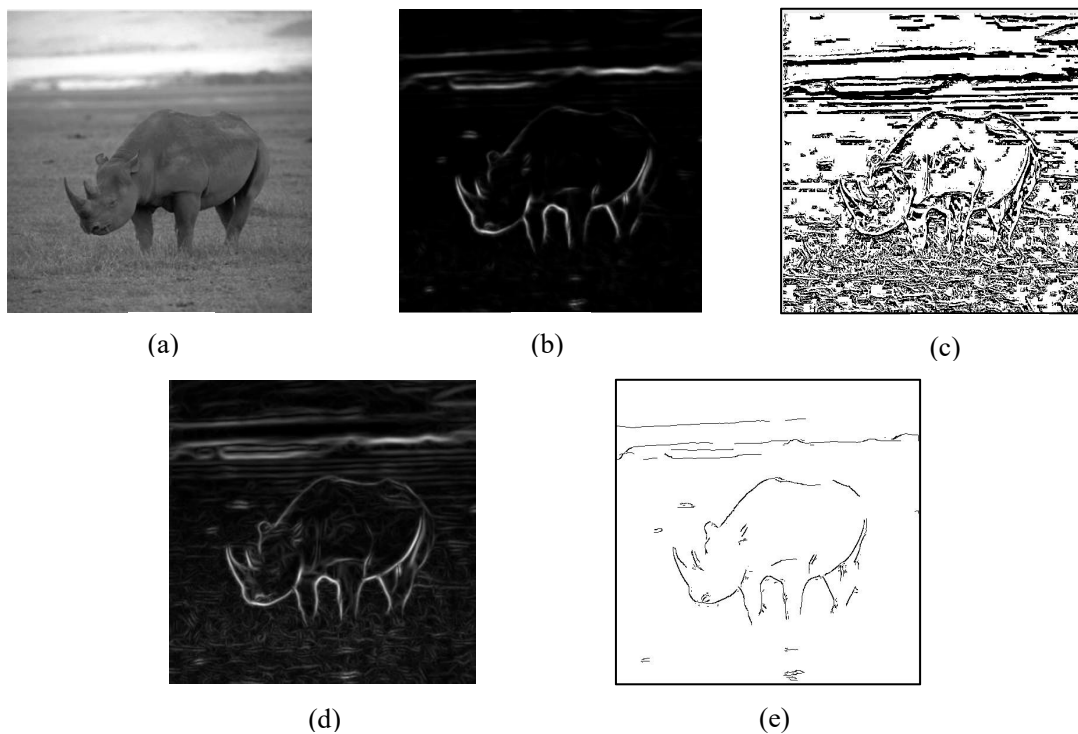


图 4-4 基于非对称方位选择机制检测效果

(a)输入图像；(b)特定方位轮廓响应；(c)初级轮廓响应加权结果；(d)非对称方位筛选与 c 融合图像结果；(e) d 经非极大值抑制及滞后阈值处理后的轮廓检测结果

4.2.2 算法验证

为了能够直观体现本章方法在轮廓检测上取得的优秀检测效果，仍采取上一章中运用了视觉机制的基于外侧膝状体具备方向选择的轮廓检测模型（CORF）、结合了稀疏编码的空间

稀疏限制方法（SSC）和具备纹理抑制特性的非经典感受野同步抑制方法（ISO）作为对比方法。参数说明如下：对于本章方法，感受野尺寸为自适应选择（不作设置），仍选取改进的 12 个不同方位的二维高斯导函数（ $N_\theta = 12$ ），设定强度抑制系数 $\alpha = [0.5:0.1:1.4]$ ，轮廓保留比例参数 $p = [0.1:0.1:1.0]$ ，得到 100 组不同参数进行测试。对于 CORF 模型，仍选用高斯导函数标准差 σ 作为测试参数，取 $\sigma = [2.0:0.5:4.0]$ ，轮廓保留比例参数为 p ，得到 50 组测试参数。对于 SSC 方法，选择 10 组参数进行测试。对于 ISO 方法，选取与本章方法一样的参数设定，强度抑制系数 $\alpha = [0.5:0.1:1.4]$ ，轮廓保留比例参数 $p = [0.1:0.1:1.0]$ ，得到 100 组参数测试。

对于轮廓检测而言，误检率和漏检率是比较关键的评价参数，所以仍引入上一章中所采用的轮廓检测定量指标（误检率 e_{FP} 、漏检率 e_{FN} 和整体性能指标 P ）来进行评价，其计算如式（3.15）~式（3.18）。用误检率 e_{FP} 来衡量其背景纹理的抑制效果，用漏检率 e_{FN} 来衡量其主体轮廓的完整度，用整体性能 P 值来说明其综合检测效果。

4.3 实验结果

自然场景下的图像是图像轮廓检测的主要应用之一，其背景具有较强的纹理噪声干扰性，及光线强弱多变性，在目标对象的主体轮廓检测识别中具有一定的难度。为了证明本章方法的有效性，对 RuG 数据库中的 40 幅图片自然场景下的图片进行了测试，该数据库每幅图片都有其对应的手工绘制的轮廓基准图，用于评价轮廓检测效果的定性指标。这里选取图库中部分测试图像结果进行说明，部分检测结果如图 4-5~图 4-9 所示。

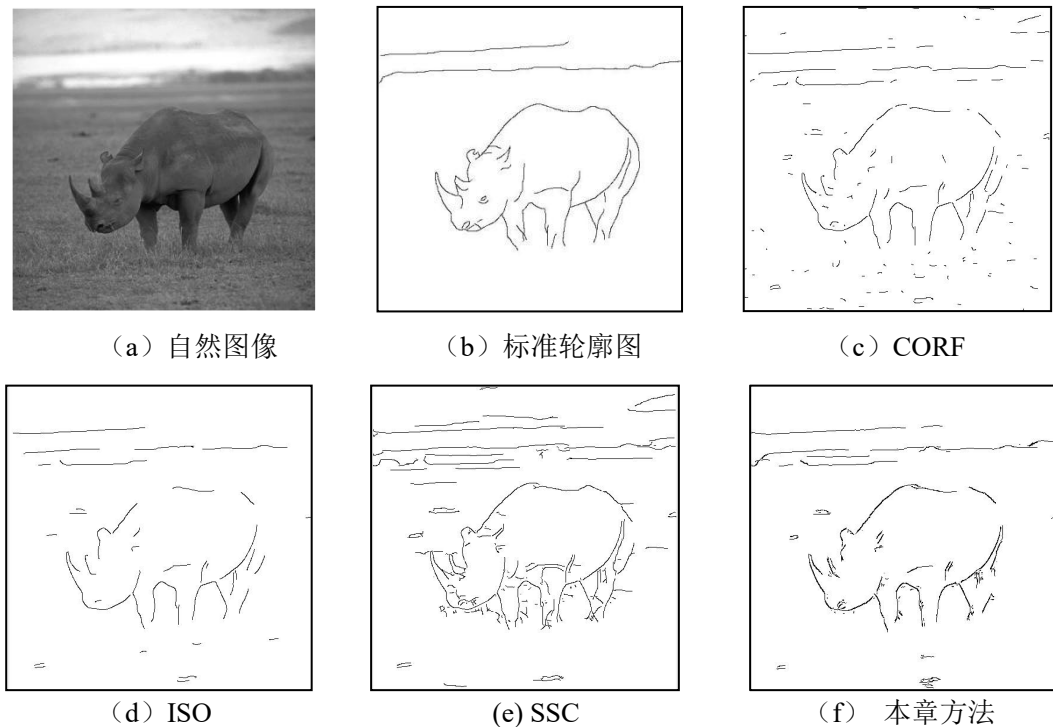


图 4-5 图像 rino 处理结果

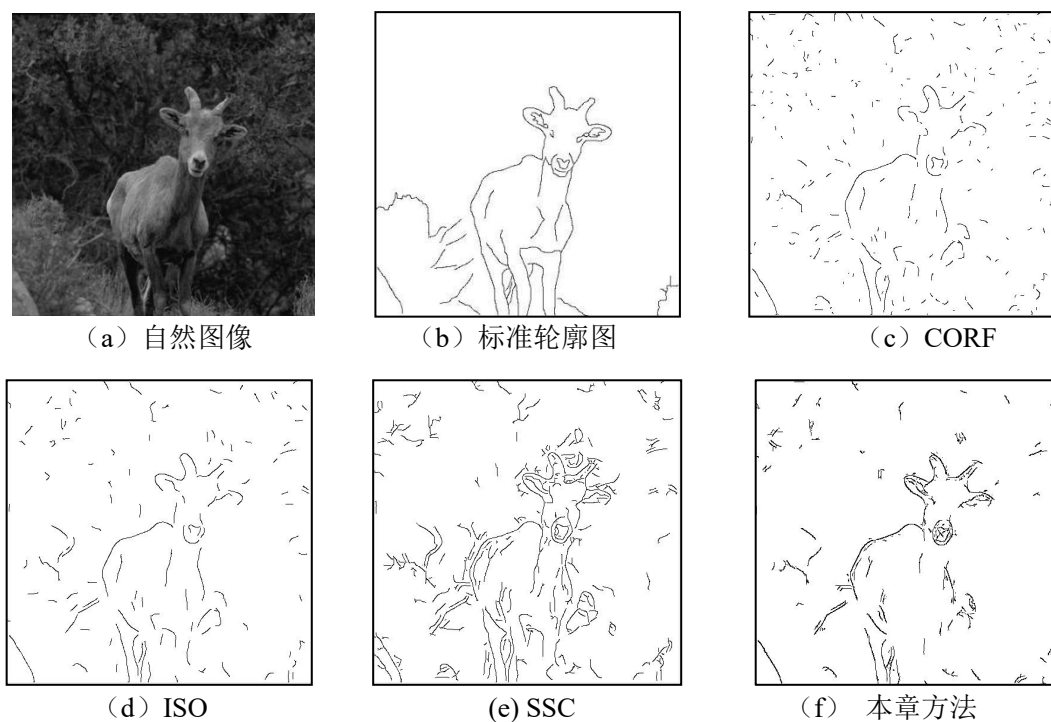


图 4-6 图像 goat_3 处理结果

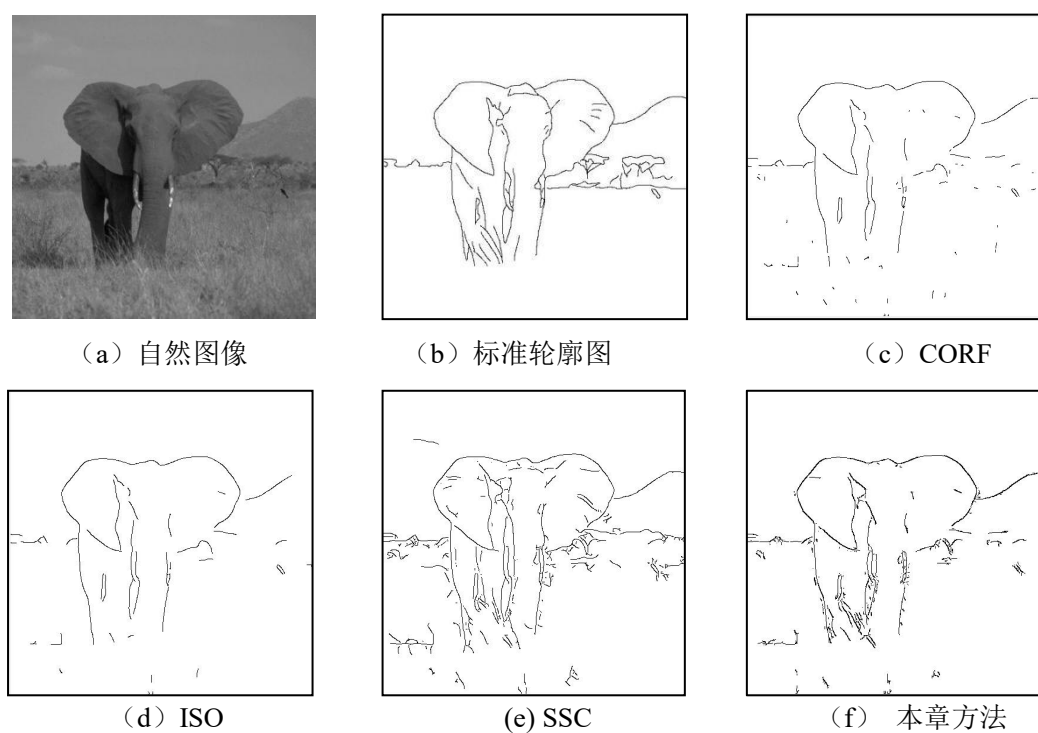


图 4-7 图像 elephant_2 处理结果

对图 4-5~图 4-9 中各算法获得的最优效果计算其性能指标，以及获得最优效果时的参数值，计算结果如表 1 所示。有表中数据可得，定量的性能指标可以表现检测图中的定性描述。以上方法中，在漏检方面 SSC 方法表现最好，其漏检率 e_{FN} 相对来说是最低的，这说明了该

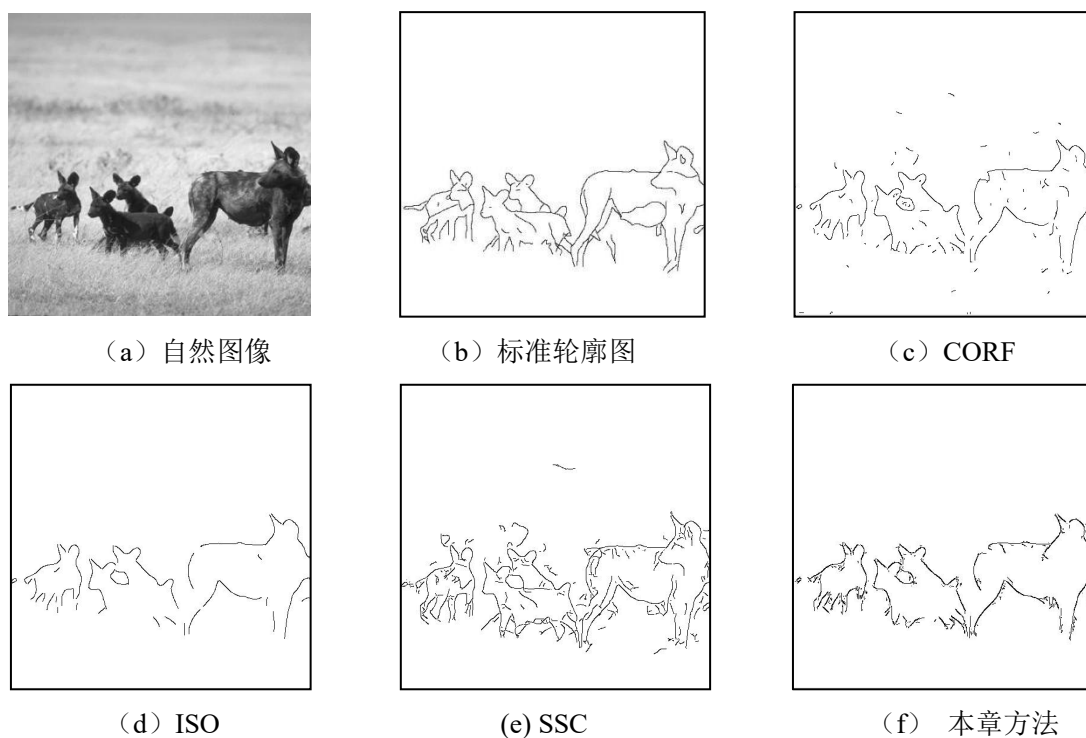


图 4-8 图像 hyena 处理结果

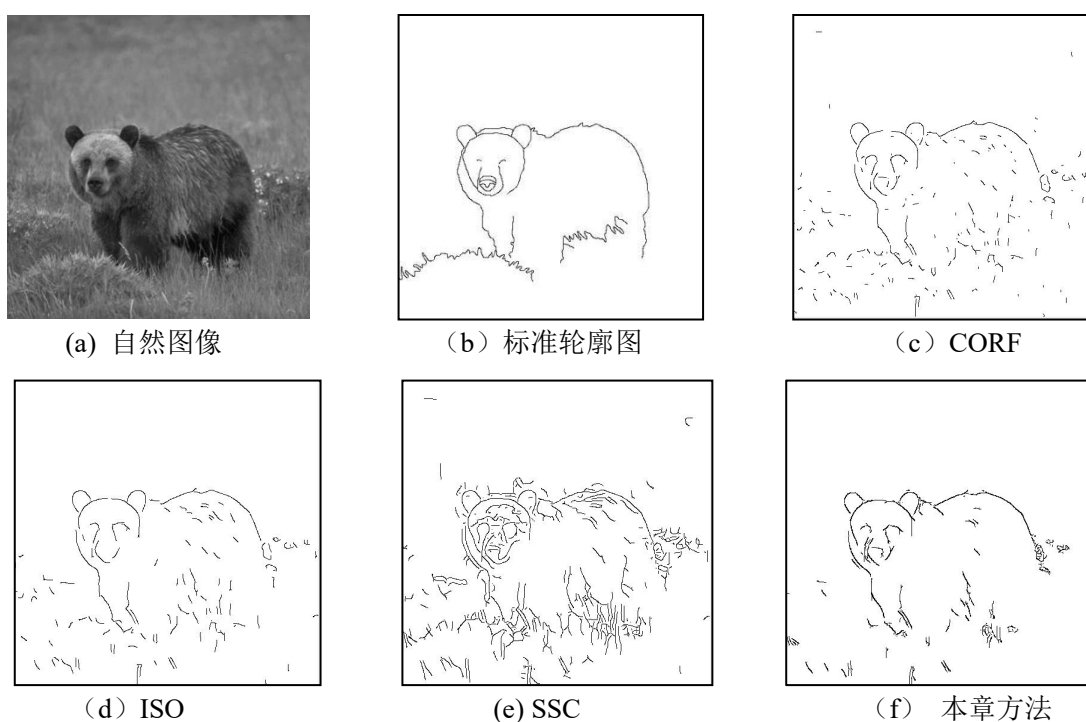


图 4-9 图像 bear_5 处理结果

方法检测出的轮廓完整度较好；但考虑其误检方面，则表现不太理想，其相对来说较大的误检率 e_{FN} 使其轮廓存在许多纹理边缘，干扰了其主体轮廓的表现。在误检方面则 ISO 方法表现总体较佳，其误检率 e_{FP} 相对来说较低，这说明了该方法在纹理抑制方面表现较好，但从图中可以看出，其在轮廓细节部分存在很大程度上的丢失，出现了纹理过抑制现象，看其计算

得出的漏检率 e_{FN} 也是相对有些高的, 由于其具备较低的误检率, 使得其在凸显轮廓方面上优于 SSC 方法。CORF 方法与本章方法, 在误检率和漏检率表现上, 相对取得一定的平衡, 总体来说在轮廓检测效果上具备相对优势; 但 CORF 检测在时间上是最不理想的, 从表中数据亦可看出, 在基于漏检和误检相对平衡的基础上, 本章方法相对于 CORF 方法, 误检率 e_{FP} 和漏检率 e_{FN} 都比较低, 具备绝对的检测优势; 从图中检测结果也可以观察到, 本章方法在对多余的纹理边缘抑制, 凸显图像主体轮廓综合表现最优。

表 4-1 图 4-5~图 4-9 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		
		t/s	α	p	e_{FP}	e_{FN}	P
Rino	SSC	1.76	3.0	0.1	1.18	0.16	0.42
	CORF	19.51	3.0	0.9	0.56	0.26	0.51
	ISO	1.63		0.1	0.26	0.35	0.54
	本方法	2.65	0.5	0.1	0.41	0.26	0.58
Goat_3	SSC	1.88	3.0	0.1	1.78	0.22	0.33
	CORF	19.35	3.0	0.9	1.10	0.46	0.33
	ISO	1.75		0.1	0.78	0.41	0.40
	本方法	2.62	0.7	0.1	0.72	0.38	0.46
Elephant_2	SSC	1.85	3.0	0.1	0.75	0.23	0.38
	CORF	19.76	3.0	0.9	0.13	0.44	0.48
	ISO	1.79		0.1	0.10	0.46	0.54
	本方法	2.54	0.5	0.2	0.32	0.30	0.59
Hyena	SSC	1.67	3.0	0.1	0.55	0.15	0.58
	CORF	20.11	3.0	0.9	0.21	0.34	0.56
	ISO	1.59		0.1	0.15	0.46	0.50
	本方法	2.67	0.5	0.2	0.18	0.34	0.62
Bear_5	SSC	1.96	3.0	0.1	2.56	0.19	0.26
	CORF	19.58	3.0	0.9	1.26	0.42	0.32
	ISO	1.69		0.1	1.29	0.46	0.30
	本方法	2.72	0.6	0.1	0.96	0.41	0.39

为了更直观的比较不同方法的整体检测性能, 以折线图的形式呈现图 (4-5)~图 (4-9)

中检测图像的轮廓检测性能指标 P 值。如图 4-10 所示：

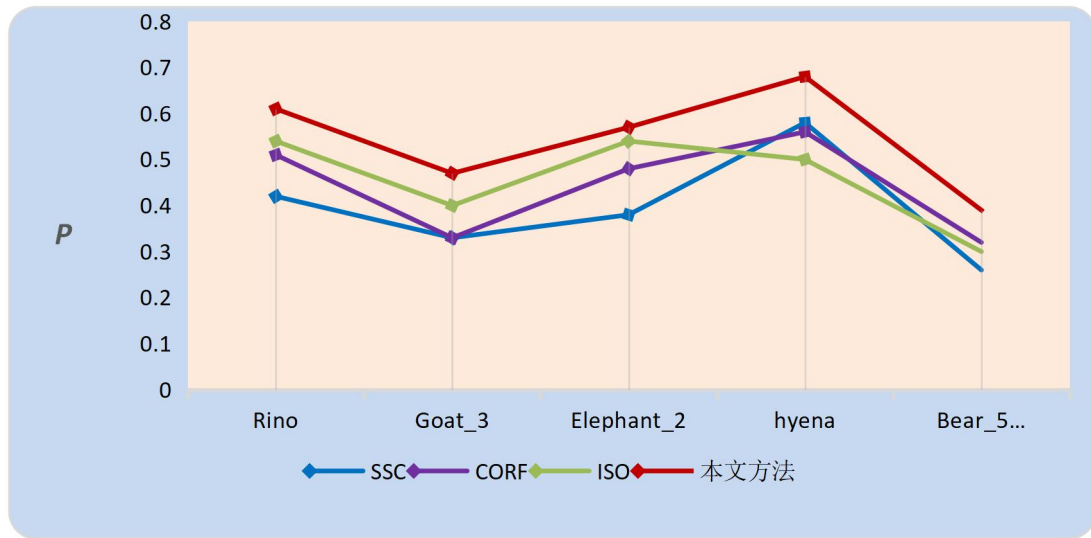


图 4-10 整体性能指标 P

4.4 讨论与分析

基于上文的不同方法的图像轮廓检测图对比，及指标参数的定性分析，综合各方面检测因数（漏检率、误检率、综合性能指标评价参数、检测时间等）可得出，在复杂的自然场景图像轮廓检测中，本章方法在抑制背景纹理凸显图像主体轮廓上表现良好，尤其是对于具备丰富细节的图像（如图 4-6，图 4-10）。CORF 方法与本章都考虑了感受野的空间组合特性，但 CORF 主要基于视通路中 LGN 感受野模型，本章在同时考虑方位感受野的敏感特性并融合了神经网络稀疏编码，同时结合非对称方位特性进行研究，因而在检测效果上具备相当的优势，对于具备丰富细节的图像，有选择性的对纹理进行抑制，从而避免了 ISO 方法那样过度抑制轮廓边缘纹理，造成有效细节轮廓的丢失。ISO 方法是基于非经典感受野的同步抑制方法，本章方法与 ISO 方法相比较，说明了在感受野特性考虑上本章方法的有效性及所构建模型符合人类注意视觉机制，且具备一定的优越性。SSC 方法是基于数学模型所构建的空间稀疏特性的方法，而本章方法是基于神经网络的稀疏编码，在检测的轮廓效果图上，本章方法的主体轮廓线虽与其有一定相似，但本章方法在纹理边缘的抑制上效果极佳，从而使得整体轮廓检测效果得到极大提高。综上分析，本章方法分别与以上三种基于视觉机制的主流轮廓检测方法相比较，都具备较好的检测优势，在实现背景纹理抑制增强图像主体轮廓效果上，图像纯净度较高，具备更加充分的生物合理性。

人类视觉机制下，初级视皮层细胞对处于拮抗区边缘具有一定方位和宽度的刺激反应强烈，并对其位置表现出严格的选择性，这在人类对物体的轮廓感知起到了重要作用。本文从人类视觉机制的生物特性出发，考虑方位敏感性感受野特性及神经网络特性，用二维高斯导函数对初级视皮层进行模拟，经过稀疏编码得到了清晰的图像主体轮廓响应，并对其进行加权处理以更好的强化凸显轮廓边缘。但背景及部分纹理边缘对图像的凸显存在干扰，为更

好检测出图像的主体轮廓，本章经分析研究，考虑到边缘纹理存在杂乱的方向响应问题，研究方向感受野的非对称特性，采取了局部到全局的自适应有效轮廓点筛选的方法对图像进行处理。

4.5 小结

本章从生物视觉角度出发，研究神经网络电生理特性及初级视皮层的感受野特性，考虑到图像轮廓的有效响应，引入神经元稀疏编码；并基于方位敏感性感受野所具备的位置及方位的严格选择性，改进感受野方位选择机制；摒弃传统上采用数学模型模拟视觉机制的做法，提出了一种基于神经元感受野多方位选择机制的图像轮廓检测方法，并选取 RuG 图像库中的自然图像为检测图像，与基于视觉机制角度提出的 CORF 方法、ISO 方法以及 SSC 方法进行比较分析。结果表明本章方法检测得到的图像轮廓具有较好的完整性和连续性，不仅背景及纹理边缘得到有效抑制，而且具有较高的图像纯净度和较好的稳定性，同时也表明了本章轮廓检测方法所具备的生物合理性。

第5章 基于显著信息多尺度融合的轮廓检测

第4章利用视觉系统神经元的响应特性以及感受野特性，提出了一种基于神经元感受野多方位选择机制的轮廓检测新方法，虽然在与传统方法的对比中，表现出了一定的检测优势。但研究发现，图像轮廓信息的显著表征与感受野尺寸密切相关，因此如何选择合适的感受野尺寸就显得尤为关键。必须要指出的是，所谓合适的感受野尺寸可能并不存在，因为图像轮廓的显著信息从空域的角度来看，轮廓对比度上呈现多级分布的特性，而从频域的角度来看，则具有高频跳跃和低频平缓的特性，因此自然会有如下两种思路：(1) 考虑到某些特定的感受野尺寸虽然不能使整体轮廓的获取达到最优，但对于部分局部轮廓的显著信息而言却是最佳的凸显，因此可针对不同的感受野尺度进行方位响应处理，由于轮廓的显著信息表征直接与有效方位响应相关联，所以本章根据所得到的有效方位响应特征矩阵，构建融合机制以得到更为完善的方位响应特征矩阵，然后结合初级视皮层轮廓响应，进一步融合调控处理得到最终轮廓响应。(2) 考虑到图像轮廓显著信息在频域上的分布特点，可针对高低频信息分别采用独立的感受野处理策略，例如对于高频跳跃信号，显然应该减小感受野尺寸，利用视觉机制更好地勾画出轮廓细节；而对于低频缓变信号，应适当增大感受野尺寸，既可以提高神经计算的执行效率，也能把握住图像轮廓的整体特性。

为了更深入研究视通路神经元感受野在轮廓检测中的作用，同时考虑到轮廓的显著信息表征与感受野尺寸差异特性的关联性，于是，在基于感受野敏感特性的研究基础上，进一步探究不同感受野尺寸下显著信息的融合方法，提出了一种基于显著信息的多尺度融合的轮廓检测方法。一方面，基于初级视皮层的神经元细胞响应特性对图像进行稀疏编码，运用第4章改进的感受野方位选择机制，构建具备方向权值占比的响应边界；另一方面，根据感受野大小的自适应选择方法计算最优感受野，考虑到处理的高效性，引入第3章 NSCT 分频方法对图像进行频域处理，针对不同频域对感受野进行一定的放大和缩小；然后模拟初级视皮层感受野特性，分别对图像做有效方位响应筛选处理，以得到不同感受野尺度下不同的显著信息特征矩阵，构建特征融合机制，得到新的显著特征矩阵，然后经过上采样放大处理与初级皮层边界响应相融合，再经过非极大值抑制和滞后阈值处理，实现图像轮廓的提取。结果显示，本章方法在轮廓响应中能够有效抑制背景纹理，使得主体轮廓更为完整，更为突出，在轮廓检测中，相对于其他方法有着更优秀的轮廓提取效果。

5.1 轮廓检测模型

在神经元感受野特性及其非对称的方位敏感特性，感受野自适应选择等视觉机制的研究基础上，本章专注于探究神经元敏感特性的在不同感受野尺寸下的表现特征，利用感受野自适应选择方法获得最优感受野尺度，为使其更符合神经元细胞工作时的复杂性，对图像进行

NSCT 频域处理, 根据图像的复杂度对最优感受野尺度进行放大(高频分量图像)和缩小(低频分量图像)得到多个感受野尺度, 然后并行计算不同感受野尺度下的具备不同显著特征的图像轮廓点响应矩阵, 基于特征的影响因子进行融合后得到最佳特征的图像轮廓点响应矩阵及最佳响应轮廓点, 再与初级皮层边界响应建立融合关系, 充分发挥最佳响应轮廓点对轮廓信息提取的重要作用, 最后引入非极大值抑制和滞后阈值处理得到最终轮廓检测结果。

5.1.1 神经元感受野边界响应

在视通路中, 神经元感受野不仅具有朝向性, 而且能够在一定程度上抑制噪声, 起到凸显轮廓边缘信息的作用。非对称性的方位敏感性感受野基于图像某一像素的响应领域, 以敏感方位为界限, 根据树突接收性质赋予不同的权值, 以放大像素点邻域内对轮廓点方向的兴奋性响应作用, 而减弱其抑制作用, 对轮廓点的检测达到了不错的效果。在以上轮廓点检测过程中, 当感受野的尺度大小选取不同时, 图像轮廓边缘信息的凸显表现出差异性。

基于视通路上神经节细胞的感受野特性, 将输入图像 $I(x, y)$ 按照式 (5.1) 计算其感受野尺度大小, 并引入第 4 章所构建的感受野的自适应选择机制根据式 (4.3) 计算得最优感受野, 记为 σ , 再次根据式 (5.1) 的方式计算得到相应的神经节细胞方向选择机制的最佳方位响应为 $E(x, y; \sigma)$; 并对所选取的最优感受野 σ 进行一定程度的放大记为 σ_{a_n} , 按式 (5.1) 进行计算得到相应的神经节细胞方向选择机制的最佳方位响应记为 $E(x, y; \sigma_{a_n})$, 然后将最优感受野 σ 进行一定程度的缩小记为 σ_{b_n} , 按式 (5.1) 进行计算得到相应的神经节细胞方向选择机制的最佳方位响应记为 $E(x, y; \sigma_{b_n})$ 。

$$E(x, y; \sigma) = \max(|I(x, y) * RF(x, y; \theta_i, \sigma)|), i = 1, 2, 3, \dots, N_\theta \quad (5.1)$$

式中, $*$ 表示卷积运算, σ 为相对应的感受野尺度, 引入第 4 章的计算方法, 如式 (4.3) 和式 (4.4), N_θ 为方向角选择中遍历角度 θ_i 的个数, RF 为二维高斯导函数, 计算如式 (3.3)。

基于方位敏感性感受野的非对称特性, 运用以上改进的视觉机制, 在自适应选取得到最佳感受野尺度 σ 的同时, 分别对其进行放大和缩小并记为 σ_{a_1} , σ_{b_1} , 将经 NSCT 分频后的高低频分量图像和未分频图像在得到的三个不同尺寸的感受野下分别结合上一章检测方法进行轮廓检测, 结果如图 5-1 所示。可清楚观察到感受野大小不同, 图像轮廓显著信息的表现存在的差异性特征。

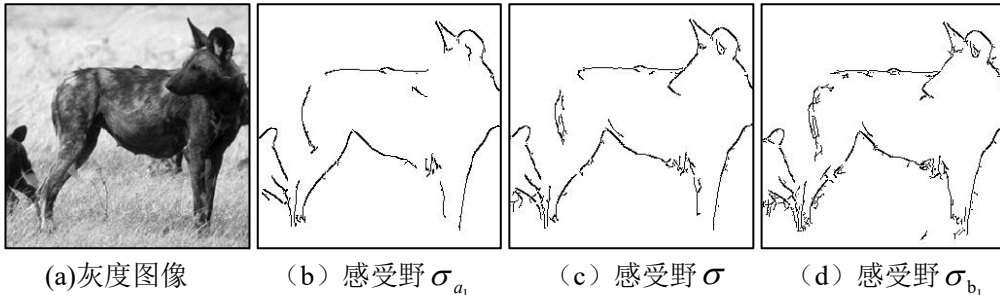


图 5-1 不同感受野尺寸下的轮廓检测结果

图 5-1 中, (a) 为待检测图像, (c) 是根据感受野自适应选取机制得到的最佳感受野

尺度 σ 下的检测结果，(b) 为最佳感受野尺度 σ 增大一定尺寸后的轮廓检测结果，(d) 为最佳感受野尺度 σ 减小一定尺寸后的轮廓检测结果。可以看出感受野尺度越小其轮廓是完整度越高，但其纹理抑制作用相应地减弱。图中 (a) 的显著信息特征表现为轮廓光滑干净，但其不连续部分最多；(c) 的显著信息特征表现为轮廓完整度最高，但其纹理也最多；(b) 的显著信息特征表现为轮廓相对完整，纹理干扰也很少，但其轮廓部分细节不如 (d)。

当然，干净、完整、没有纹理干扰的轮廓是我们最理想的轮廓提取结果，虽然这样理想的提取结果不可能达到，但可以通过改进提取方法，使得图像轮廓提取结果更接近我们的理想结果。基于以上分析，本章对如何更全面地提取图像轮廓显著信息特征进行了进一步研究。

5.1.2 多尺度表征融合机制

在视觉皮层中，神经元的感受野除了同心圆状这种简单型感受野外，还存在复杂型感受野和超复杂型感受野，相应的感受野响应特性也更为复杂。视觉刺激下，这些细胞的感受野共同作用，其兴奋性和抑制性的界限变得复杂模糊，难以描述。所以视觉机制作用时，感受野的表现具备多样性，因此，探究感受野不同尺度下的表现特性，构建一种能够综合其显著性信息的融合机制，符合人类的视觉特性。

依据感受野尺度不同所带来的图像轮廓信息的显著性特征的差异，在自适应最佳感受野尺度 σ 选取的基础上，计算基于最佳感受野尺度放大和缩小的感受野尺度 $\sigma_{a_n}, \sigma_{b_n}$ 的值，如式 (5.2) 所示。

$$\begin{cases} \sigma_{a_n} = \sigma + n \cdot \frac{|p - p_0|}{p_0} \cdot \sigma, a_n = a_1, a_2, a_3, \dots \\ \sigma_{b_n} = \sigma - n \cdot \frac{|p - p_0|}{p_0} \cdot \sigma, b_n = b_1, b_2, b_3, \dots \\ \sigma_{b_n} > 0 \end{cases} \quad (5.2)$$

式中， a_n 为第 n 次放大的感受野尺度， b_n 为第 n 次缩小的感受野尺度， $\sigma_{a_n}, \sigma_{b_n}$ 为相对应的经过第 n 次放大和第 n 次缩小的感受野尺度（一般 n 取 2-3 次）。

为了更好地描述图像特征，降低特征融合的复杂度，以及避免过拟合问题的出现，采取池化的方法，分别对基于不同尺寸的感受野的最佳方位响应 $E(x, y; \sigma)$ 进行池化（最佳方位响应 $E(x, y; \sigma)$ 的按式 (5.1) 计算）。为了让池化模型具备较高的稳定性，产生的误差尽量小，同时考虑到图像的普适性，因此采取中间值池化法。这是对平均池化法和最大池化法的一种折中算法，可以兼顾平均池化法和最大池化法获取到的特征值。

设输入最佳方位响应 $E(x, y; \sigma)$ ，设池化域为矩阵 $K[x, y]$ ，偏置为 b_2 ，移动步长为 $step$ ，采用中间值池化后特征矩阵 $S(x, y; \sigma)$ 计算如式 (5.3) 所示。

$$S(x, y; \sigma) = \frac{1}{2step^2} \left(\sum_{x=1}^c \sum_{y=1}^c E(x, y; \sigma) \right) + \frac{1}{2} \cdot \max_{x=1, y=1}^c \{E(x, y; \sigma)\} + b_2 \quad (5.3)$$

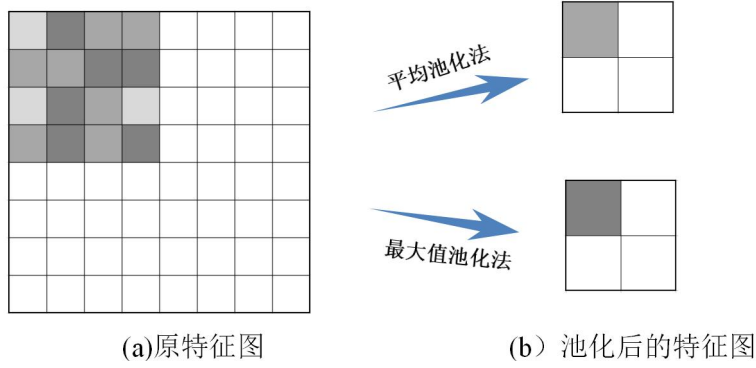


图 5-2 池化示意图

将图像进行 NSCT 频域处理，得到高频分量图像和低频分量图像，不同感受野尺度 $\sigma_{a_n}, \sigma, \sigma_{b_n}$ 对应的方位响应 $E(x, y; \sigma_{a_n}), E(x, y; \sigma), E(x, y; \sigma_{b_n})$ 分别采用中间值池化法按式 (5.3) 进行计算，可分别得到池化后的特征矩阵 $S(x, y; \sigma_{a_n}), S(x, y; \sigma), S(x, y; \sigma_{b_n})$ ，即不同感受野尺寸下的图像轮廓响应的显著特征矩阵。为了能够得到完整性更高、理干扰性更小的高质量轮廓，综合它们的显著特征矩阵，将其按照一定比例进行融合。记融合后新的显著性特征矩阵为 $F(x, y; \sigma)$ ， α, β, γ 为显著特征矩阵的融合系数，如式 (5.4) 所示。

$$F(x, y; \sigma) = \alpha \cdot \frac{1}{n} \sum_{n=1}^n S(x, y; \sigma_{a_n}) + \beta \cdot \frac{1}{n} \sum_{n=1}^n S(x, y; \sigma_{b_n}) + \gamma \cdot S(x, y; \sigma) \quad (5.4)$$

$F(x, y; \sigma)$ 经过上采样双线性插值法放大后，得到新的轮廓响应 $I(x, y)$ ，将其与对应的依据突触特性进行权值处理的最终轮廓边界响应 $g(x, y)$ 融合，得到轮廓检测图像 $f(x, y)$ ，如式 (5.5) 所示。

$$f(x, y) = I(x, y) \times (\delta_{x,y} \times E(x, y; \sigma)) \quad (5.5)$$

$$\delta_{x,y} = \sum_{\tau} (d(\tau) \cdot \sum_{i=1} \delta(x_i - x_c, y_i - y_c; \lambda)) \quad (5.6)$$

式中， $\sigma_{x,y}$ 表示最佳响应点权重，方位为 τ 表示方位角， d 表示方位角数值化。 (x_c, y_c) 表示感受野的中心坐标， (x_i, y_i) 表示接收区域第 i 个响应角所在的方位矩阵坐标。

最后再对轮廓检测图像 $f(x, y)$ ，进行非极大值抑制与滞后阈值处理，得到总体轮廓的最佳方向响应，实现图像轮廓的提取。

5.1.3 轮廓检测流程

本章基于上一章方法，考虑到视通路中神经元细胞种类的复杂性及其感受野尺度变化对图像轮廓显著信息表征的密切影响，提出基于显著信息多尺度融合的轮廓检测方法，轮廓检测流程如图 5-3 所示，具体检测步骤表述如下：

(1) 对于输入图像，一方面引入第 4 章方法，基于稀疏编码的方式得到初级轮廓响应图

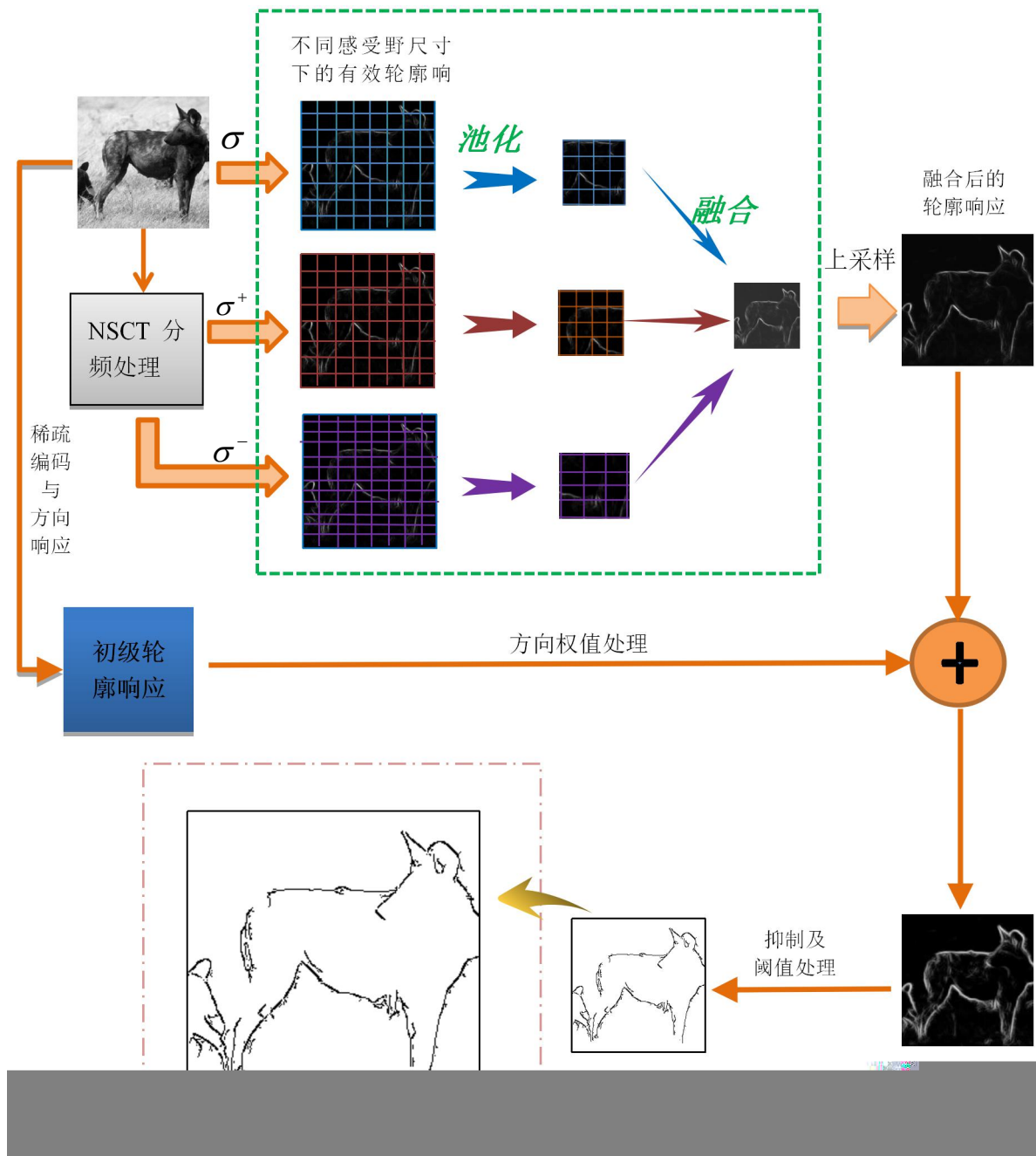


图 5-3 轮廓检测原理流程图

像 $E(x,y;\sigma)$ ，并引入感受野的非对称方位敏感特性，对其进行方向权值调整，得到权值调整图像 $g(x,y)$ ；

(2) 对于输入图像，另一方面引入第 3 章 NSCT 频域处理方法对图像进行分频处理，将图像分解为高频分量图像与低频分量图像；

(3) 引入第 4 章的感受野大小自适应选择机制，计算输入图像的最佳感受野大小为 σ ，然后对其进行 n 次放大并分别作用于高频分量图像，可得到 n 个放大尺度的感受野所对应的轮廓响应图像，同时，对其进行 n 次缩小并作用于低频分量图像，可得到 n 个缩小尺度的感

受野所对应的轮廓响应图像；

(4) 采用中间值池化法，对第(3)步骤所得到的 $(2n+1)$ 幅轮廓响应图分别进行池化，得到相对应的显著特征矩阵；

(5) 利用上文中所构建的融合机制，将第(4)步骤中得到的 $(2n+1)$ 幅显著特征图像进行融合，得到新的显著性特征矩阵 $F(x,y;\sigma)$ ，经过上采样插值放大后得到新的轮廓响应图像 $I(x,y)$ ；

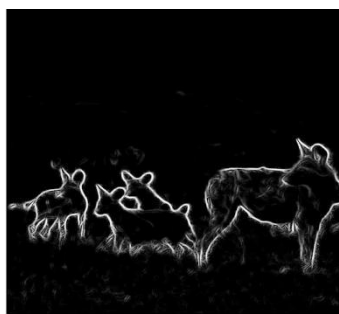
(6) 将新的轮廓响应图像 $I(x,y)$ 与第(1)步骤中的经过方向权值占比调整处理的图像 $g(x,y)$ 融合，最后通过非极大值抑制及滞后阈值处理得到最终的轮廓检测结果。

输入图像（如图 5-4 所示），按照以上流程对其轮廓进行检测，同时，运用第 4 章基于神经元感受野多方位选择机制的图像轮廓检测方法检测其轮廓并输出结果。



图 5-4 输入图像

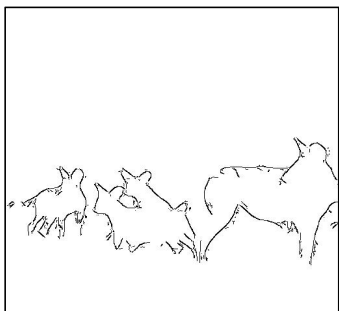
基于第三章感受野自适应选取方法，对图 5-4 图像进行感受野自适应选取，得到最佳感受野尺寸 σ 为 1.2，有效轮廓点响应图像如图 5-5 中 (a) 所示。据式 (5.4) 进行计算，这里



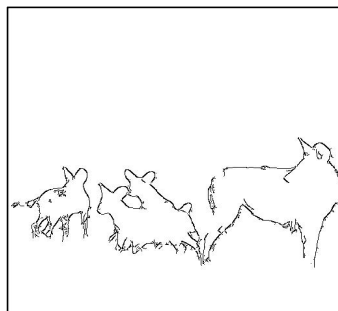
(a) 轮廓响应_Ch4



(b) 轮廓响应_本章



(c) 检测结果_Ch4



(d) 检测结果_本章

图 5-5 第 4 章与本章轮廓检测对比图

n 取 2（即分别进行了两次放大和缩小）得到放大和缩小后的感受野 $\sigma_a, \sigma_a, \sigma_b, \sigma_b$ 大小分别为 1.5, 1.8, 0.9, 0.5, 将它们相应的轮廓响应图像依据显著性表征机制融合后, 得到的轮廓响应图像如图 4.5 的 (b) 所示。对比 (a) (b) 可以看出, (a) 的轮廓边缘响应强度较为强烈, 部分纹理也有响应; (b) 的整体轮廓响应稍弱, 但纹理响应得到了更好的抑制。经过与初级轮廓响融合后, 最终处理后的检测结果, 如图中 (c) (d) 所示。对比 (c) (d) 可以看出, 相对于 (c) 来说, (d) 中的轮廓细节更为丰富, 同时纹理抑制效果也得到了提升。由此说明, 本章方法相对于上一章的轮廓检测方法, 最终轮廓检测效果得到了提升与优化。

5.2 实验结果

背景纹理噪声的抑制与主体轮廓的完整性一直是自然场景下的图像轮廓检测的主要关注点。为了证明本章方法的优越性, 再次对 RuG 数据库中的 40 幅图片自然场景下的图片进行了实验测试, 为了便于比较, 这里选取上一章中所呈现的测试图像结果为例进行说明, 如图 5-6 所示。

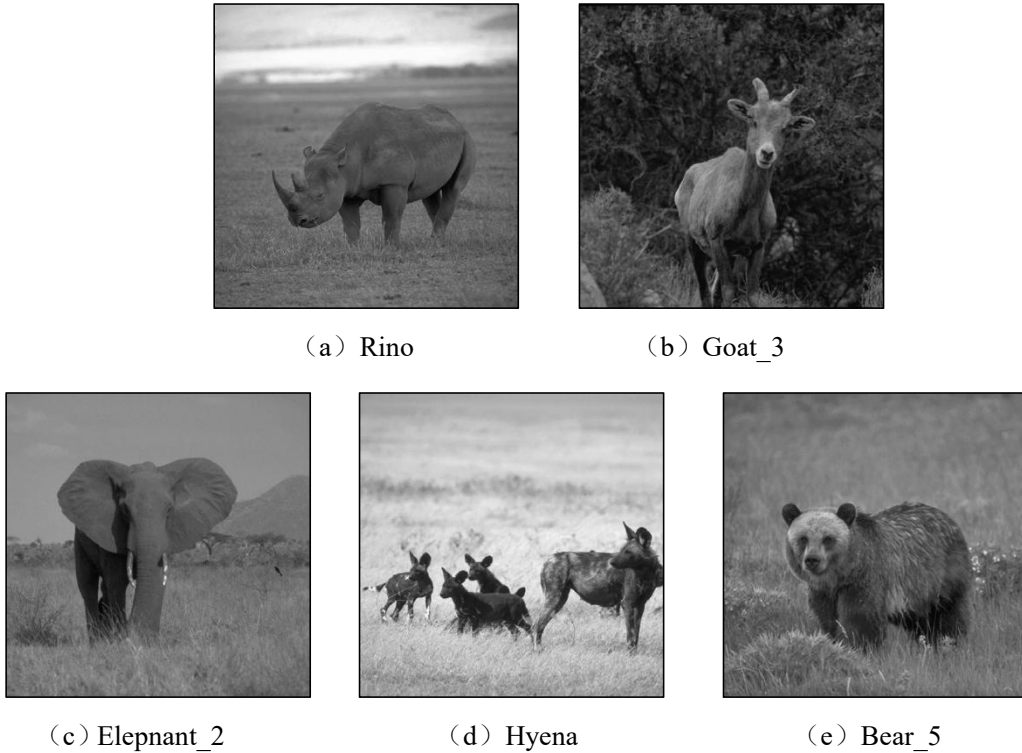


图 5-6 待检测自然图像

在第 3 章方法和第 4 章方法中, 已经证明其轮廓检测效果优于基于外侧膝状体 (LGN) 具备方向选择的轮廓检测模型 (CORF), 具有稀疏编码效应的空间稀疏限制方法 (SSC), 在纹理抑制方面具有一定优势的非经典感受野同步抑制方法 (ISO), 所以, 本章方法主要与第 3、4 章方法进行比较。本章方法中的部分仍采用第 4 章中基于局部与全局的处理的方法, 在局部方向角矩阵处理之后, 选取改进的 12 个不同方位的二维高斯导函数 ($N_\theta = 12$); 选取 10 个不同强度的抑制系数 $\alpha = [0.5:0.1:1.4]$, 轮廓保留比例参数为 $p = [0.1:0.1:1.0]$, 感受

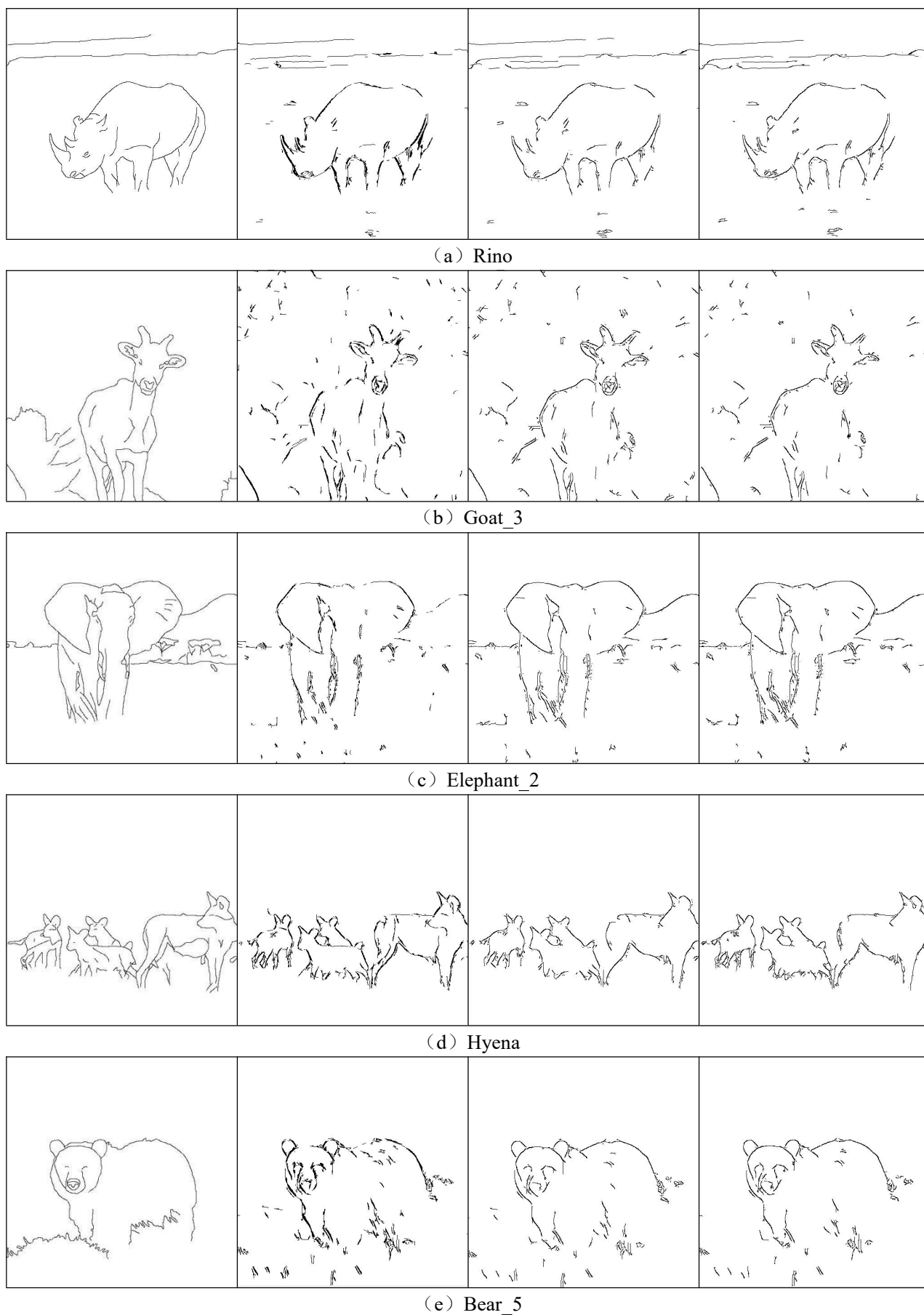


图 5-7 轮廓检测结果图 (列 1_标准轮廓图 , 列 2_第 3 章方法, 列 3_第 4 章方法, 列 4_本章方法)

野尺度自适应选择，得到 100 组不同参数进行测试。

图像相对应的测试结果如图 5-7 所示，第 1 列图为其对应的手工绘制的轮廓基准图，用于评价轮廓检测效果的定性指标，第 2 列图为第 3 章方法检测所得到的轮廓结果图，第 3 列图为第 4 章方法检测所得到的轮廓结果图，第 3 列图为本章方法检测所得到的轮廓结果图。

为了达到客观准确的评价效果，采用与第 3、4 章一致评价方法（误检率、漏检率与整体性能指标定量评价），评价指标计算方法保持也一致。运用式（3.15）~式子（3.18）对图 5-6 中(a)~(e)图像检测中三种算法获得的最优效果计算其性能指标，结果如表 5-1 所示。表中数据可以看出，本章方法误检率 e_{FP} 相对于第 3、4 章方法普遍下降，但漏检率 e_{FN} 有所上升，综合来看整体性能 P 值及轮廓提取质量得到提高。与 CORF 方法、ISO 方法以及 SSC 方法比较分析，第 3 章方法在对于主体轮廓提取的完整性方面表现最好，第 4 章方法在对多余的纹理边缘抑制，凸显图像主体轮廓综合表现已达到最优。本章方法相对于第 3 章和第 4 章方法来说，在基于漏检率与误检率平衡的基础上，本章方法表现更为优秀。结合检测结果图像可观察到，在本章方法中纹理边缘得到了更好的抑制，凸显的主体轮廓综合表现更为优秀。

表 5-1 各图像对应的性能指标

图像	算法	性能指标		
		e_{FP}	e_{FN}	P
Rino	第三章方法	0.44	0.25	0.60
	第四章方法	0.41	0.26	0.58
	本章方法	0.22	0.44	0.60
Goat_3	第三章方法	0.60	0.34	0.52
	第四章方法	0.72	0.38	0.46
	本章方法	0.43	0.55	0.48
Elephant_2	第三章方法	0.25	0.37	0.60
	第四章方法	0.32	0.30	0.59
	本章方法	0.27	0.33	0.61
Hyena	第三章方法	0.26	0.23	0.65
	第四章方法	0.18	0.34	0.62
	本章方法	0.28	0.17	0.67
Bear_5	第三章方法	1.25	0.38	0.37
	第四章方法	0.96	0.41	0.39
	本章方法	0.42	0.94	0.40

5.3 讨论与分析

基于上文中三种方法的轮廓检测结果图对比分析, 及图像轮廓的误检率、漏检率、综合评价指标等参数进行分析, 表明在复杂的自然场景图像检测中, 本章方法的综合表现更为优秀。

第 3 章和第 4 章所用的检测方法所构建的检测模型符合人类的视觉机制, 与基于视觉机制的主流轮廓检测的 CORF 方法、ISO 方法以及 SSC 方法相比较, 检测结果也表现了其所具备的优越性——图像主体轮廓都具有较高完整性, 背景纹理相对于其它方法得到了很好的抑制。对于前两章而言, 第 3 章方法中轮廓提取完整度最高, 但同时也保留了比第 4 章方法更多的纹理, 第 4 章方法中背景纹理抑制效果最好, 虽然主体轮廓完整度略低于第 3 章, 但综合表现优于第 3 章。第 4 章方法中所选取的感受野尺寸一致, 从真实生物视觉角度来说, 神经元细胞种类多样, 并且相互之间存在影响, 在刺激作用下感受野尺寸必然是极具复杂性的, 故本章基于感受野的多尺寸融合进行轮廓检测, 是基于真实生物视觉作用的一种探究方式。实验结果也显示, 本章方法在背景纹理抑制上表现更为优秀, 得到的图像轮廓检测结果纯净度更高, 对于细节部分及轮廓弱边缘的检测更为良好, 虽然其检测出的细节轮廓及弱边缘存在部分不连续性, 直接导致了漏检率的上升, 但对于图像整体轮廓的凸显, 综合来看本章方法表现更胜一筹。

在人类视觉感知系统中, 视觉感知系统能够实现感兴趣目标的快速识别与定位分析, 并选择性忽略其背景区域, 其中, 视觉注意机制具备主动选择感兴趣区域并处理信息、能够有效提高目标检索与识别等图像处理能力的视觉特性, 其特性离不开神经元细胞的作用, 神经元细胞感受野方位敏感特性具有极为重要的意义, 对边缘位置和方位的选择发挥了极大的作用。本章从生物特性出发, 在考虑方位敏感性感受野非对称方位特性的同时, 深入探究了感受野尺寸大小对图像显著信息表现特性的影响, 对不同感受野尺寸的轮廓响应图像进行显著特征池化融合, 从而加强了有效轮廓边缘的凸显, 弱化了背景纹理干扰, 因而在检测效果上更为理想。

5.4 小结

本章从人类视觉感知系统出发, 考虑视觉注意机制, 深入研究了不同尺度感受野选择下对图像轮廓显著信息表现的影响, 考虑到视通路神经元细胞的种类与复杂性及其感受野的复杂性, 进一步研究发现感受野尺度不同, 所对应的轮廓表征信息的显著性存在的差异性。视通路的神经元细胞相互作用, 其感受野之间相互影响, 视觉感知的实现是基于综合作用下的实现, 在视觉刺激下其感受野尺度呈现复杂性, 所以从感受野尺度的不同选取来构建融合机制, 分别作用于经 NSCT 频域处理后高低频分量图像, 符合真实生物视觉特性。以自然场景下的图像为例, 本章方法检测出的轮廓较第 3、4 章方法而言, 图像轮廓的凸显更具备完整性, 细节更为丰富, 纹理抑制效果更好, 整体轮廓的提取质量更高。

第 6 章 总结与展望

6.1 总结

在人工智能成为行业驱动力的大背景下,人工智能技术与计算机视觉技术已然密不可分。对视觉信息的感知处理将成为人工智能重要的技术创新点,并呈现出以图像识别、智能 3D、计算机视觉等为突破口,重点向 AI、AR 等领域布局与发展的趋势。图像轮廓检测的研究也呈现出与纯数学角度告别,与视觉感知及生物特性相结合发展的方向趋势。目前,对于图像轮廓检测的方法,很多集中于构建数学模型来实现对视觉机制的模拟,并将所构建模型应用在图像处理中,而忽略了信号在传递过程中的损耗与丢失,也就是说其忽略了视觉信号在传递时所具备的电生理特性;也有不少专注于神经元网络模型的构建,但对神经元特性的考虑片面化,关注点在于所构建网络的结构组织,而不是引入视觉机制的方式方法。因此本文研究从视觉感知角度出发,关注于视通路结构对信号的影响及处理方式,深入研究神经元感受野特性,改进及构建视觉机制模型来实现轮廓的检测,结果表明本文所用方法在检测效果上表现优秀,并且符合生物视觉特性。更为重要的是,本文方法为图像处理及理解在视觉信息的感知处理研究领域提供了一种新思路。

近年来,将视觉机制与图像处理相结合的方法被大量提出,但大部分都忽略了视通路神经元感受野的变化特性与视觉信息在传递处理中存在的前级编码影响。本文基于视觉感知作用下,探讨了视通路的神经元感受野特性以及视觉信息传递和处理的规则和方式,研究神经元的编码机制并引入一种分频思想,结合前馈作用处理,在轮廓特征提取上取得了良好效果。

虽然研究者们提出了大量的对视觉感知的应用实现方法,但大多数都是基于数学模型的黑箱模拟,在视觉感知下感受野对视觉信息的处理有着至关重要的作用,亟待深入研究。本文研究感受野的非对称方位敏感特性,对图像进行神经稀疏编码及自适应的感受野大小选择,并结合方向响应加权处理提取轮廓特征,取得了良好的检测效果。

此外,考虑到人类神经元细胞感受野的真实结构的复杂性,进一步探究了其敏感特性的在不同感受野尺寸下的表现特征,并基于特征的影响因子构建显著信息的融合机制,提出了一种基于显著信息多尺度融合的轮廓检测新方法。

论文具体研究工作总结如下:

(1) 本文探讨了视通路中信号传递和处理的规则和特性,结合神经元感受野特性,引入神经元电生理模型及神经编码机制,提出一种基于频域融合编码与前馈机制的图像轮廓检测方法。一方面通过高斯导函数对初级视皮层细胞感受野特性进行模拟,另一方面对图像进行频域分解得到高频分量图像与低频分量图像,并根据得到的图像特征差异性,采用冗余编码与稀疏编码两种不同的编码方式分别处理,最后对其融合处理并前馈作用于初级皮层。以 RuG

图库的自然图像作为实验对象，并与主流轮廓检测方法对比，经评价指标计算对比分析（漏检率、误检率、综合评价指标）表明该方法能够有效提取图像轮廓，特别在纹理抑制方面优势明显。

（2）考虑到神经元感受野特性与轮廓响应存在密切关系，深入研究了神经元感受野的非对称多方位敏感特性，提出一种基于神经元感受野多方位选择机制的图像轮廓检测方法。首先基于初级视皮层的神经元细胞响应特性对图像进行稀疏编码，改进感受野的方位选择机制，构建具有方向权值占比的响应边界；同时模拟初级视皮层感受野特性，并基于非对称的方向敏感性感受野特性，针对图像复杂度构建一种感受野大小自适应选择机制，并进行有效方位响应筛选处理，再引入感受野的差异特征，实现图像轮廓的提取。考虑到自然图像的复杂性，对 RuG 库图像进行轮廓检测，检测结果表明该方法得到的图像轮廓具有较好的完整性和连续性，不仅能够有效抑制背景及纹理边缘，而且具有较高的图像纯净度和较好的稳定性，具备生物合理性。

（3）进一步探究发现感受野的尺度与轮廓显著信息的表征存在影响关系，结合以上对视通路的结构及对信息处理方式的研究，提出一种基于显著信息多尺度融合的轮廓检测方法。在视通路神经元感受野的非对称方位敏感特性研究基础上，深入探讨其敏感特性的在不同感受野尺寸下的显著信息表征特性，构建一种显著信息的融合机制，然后引入基于第四章改进机制下的初级皮层边界响应边界及感受野差异特性，完成图像轮廓的检测。实验结果表明，该方法对图像背景纹理具有更好的抑制效果，得到的主体轮廓的完整性更高，整体轮廓提取效果更好。

6.2 展望

本文将视觉感知应用于图像轮廓检测并取得了不错的检测效果，证明了本文方法不仅具备可行性和有效性，而且符合生物视觉特性。对于多通道神经元编码方式，后续可以更深入研究视通路不同结构层级在视觉信息传递过程中的所表现出的特点及处理方式，分而治之；对于方位选择机制，后续可以更多考虑感受野之外的存在的生理结构影响及反馈作用；对于多尺度融合策略，后续可结合多维度引入空间特征进一步探究。

此外，可以引入更多的视觉编码特性，使其更趋近于生物的真实视觉，这有助于对视觉信息的感知、识别、理解和再创造，后续可扩展应用到机器学习、人脸识别、图像检索、无人机视觉等系统中。同时，5G 技术的发展加速消费升级，计算视觉产业将迎来更广阔的场景应用需求，基于视觉感知的轮廓检测技术的研究将极具挑战，也将极富意义。

致谢

数载好景君须记，最是橙黄橘绿时。雏鸟初长成，别离声声恋。

恰同学少年，指点江山，激扬文字，在杭电。逢疑惑困难，答疑解惑，关切相助，在杭电。道不尽的谢，数不清的情，于师，于友，于亲人，在我心。

从一叶浮萍随汪洋，到一心所愿是杭城，从云深不知处，到柳暗花明又一村，只因，一路有你们，感谢，一路有你们。相遇，是此生莫大的福气，相识，是此生最大的幸运。

尤念安克叔叔的留下，留下。

范影乐教授，由衷感谢您！时光之须臾，人生之转折。心之大，行致远。学贵以恒，研贵以专，思之以谨微，行之以高效。风雨中，当明，进度与进步，自知与知趣。仰望星空，脚踏实地，此情最铭记。

徐旭初教授，深深感谢您！生命之柔韧，前行之无畏。八千里路云和月，修之布施，持戒，忍辱，精进，禅定，般若。谨记在心，制度是人的制度，人是制度的人。一路的指引，一路的鼓励，此情最真挚。

武薇老师，感谢您如姐姐般的关怀，偶遇的四叶草，静静躺在书页上，愉悦美好。余善恩老师，感谢您的科研参赛指导，严谨负责，细心耐心之言教，收获颇多。

李丹菁师兄，感谢你一直以来的鼓励，一直以来的关心与照顾，慷慨自古英雄色，甘洒热血写青春！共勉。廖进文、罗佳骏师兄，王梦蕾师姐，感谢你们相伴，红豆悠哈漫步，劝勉当惜时，笑泪付风云。方芳师姐，王典、王迪师兄，感谢你们的指导，“这是‘芳芳’……”初见恍如昨日，却是岁不我与，科研生涯，予我以渔，路见崎岖，予我以力量，感激于心。李康群、谢尚周、郭笑天、李嘉祥师兄，感谢你们的帮助，学亦乐乎，玩亦乐乎，风华摇曳，思齐焉。

木欣欣以向荣，泉涓涓而始流。感谢同窗郭金朋、王强、胡钧皓，学术同研，生活互助，凌晨的杭电，凌晨的下沙，凌晨的身影，凌晨的收获，乘风破浪正当时。感谢师弟周涛对我科研工作的协助及论文撰写的帮助，山不在高，水不在深，君子以自强不息。感谢师弟张明琦对我论文工作的支持，勇往直前，待满月雕弓射天狼。感谢师弟杜宇华、蒋涯给我科研工作增添欢乐色彩，业精于勤，博观而约取，厚积而薄发。感谢房涛及师弟王阳辉、谭明明，师妹黄春燕、田纯带来热闹氛围，善听导师言，海阔凭鱼跃。感谢室友李夏如的陪伴，夏花之灿烂，戏蝶时时舞，娇莺恰恰啼，与我关爱多多，携手共成长。感谢学长董帅对我工作的指导与建议，如闻芷兰之香，如获海底之珍。

感谢我的朋友，一路相伴，之交若水。感谢我的亲人，爱我宠我支持我。感谢我的父母，予我生命，予我自由，予我爱与支持。最后，还要感谢我自己，上下求索。

日月之行，星汉灿烂，谢之于言，感之于心。凡心所向，素履所往，其犹未悔。怀炽热之心，习泰然之气，登高望远。忆往昔，值得；于当下，珍惜；谈明天，希冀。

Hello ,world !

Thanks, world !

下沙.杭州电子科技大学
二零一八年三月

参考文献

- [1] 贺强, 晏立. 基于 LOG 和 Canny 算子的边缘检测算法[J]. 计算机工程, 2011,37(3):210-212.
- [2] 王康泰, 戴文战. 一种基于 Sobel 算子和灰色关联度的图像边缘检测方法[J]. 计算机应用, 2006, 5.
- [3] 张永亮, 刘安心. 基于 Prewitt 算子的计算机数字图像边缘检测改进算法[J]. 解放军理工大学学报: 自然科学版, 2005, 6(1): 44-46.
- [4] 张洁. 数字图像边缘检测技术的研究[D]. 合肥工业大学学报, 2009.
- [5] 康牧, 许庆功, 王宝树. 一种 Roberts 自适应边缘检测方法[J]. 西安交通大学学报, 2008, 42(10):1240-1244.
- [6] Papari G, Petkov N. Edge and line oriented contour detection: State of the art[J]. Image and Vision Computing, 2011, 29(2): 79-103.
- [7] 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.
- [8] 柴志雷, 张圆蒲. 基于 GPU 的轮廓提取算法的并行计算方法研究[J]. 计算机应用研究, 2015(2):630-634.
- [9] Grigorescu C, Petkov N, Westenberg M A. Contour and boundary detection improved by surround suppression of texture edges[J]. Image and Vision Computing, 2004, 22(8): 609-622.
- [10] 王利. 基于局部特征与全局结构的轮廓检测算法研究[D]. 吉林大学, 2016.
- [11] 谢昭, 童昊浩, 孙永宣,等. 一种仿生物视觉感知的视频轮廓检测方法[J]. 自动化学报, 2015, 41(10):1814-1824.
- [12] 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.
- [13] 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.
- [14] Van Wyk M, Wässle H, Taylor W R. Receptive field properties of ON-and OFF-ganglion cells in the mouse retina[J]. Visual neuroscience, 2009, 26(03): 297-308.
- [15] 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.

- [16]Daugman J G. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters[J]. JOSA A, 1985, 2(7): 1160-1169.
- [17]张巧荣. 视觉注意计算模型及其关键技术研究[D]. 哈尔滨工程大学, 2011.
- [18]Grigorescu C, Petkov N, Westenberg M. Contour detection based on nonclassical receptive field inhibition[J]. Image Processing, IEEE Transactions on, 2003, 12(7): 729-739.
- [19]王典. 基于视觉感知的神经计算模型及其应用研究[D]. 杭州电子科技大学, 2016.
- [20]Chen K, Song X, Li C. Contrast-dependent variations in the excitatory classical receptive field and suppressive nonclassical receptive field of cat primary visual cortex[J]. Cerebral Cortex, 2012, 23(2): 283-292.
- [21]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.
- [22]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.
- [23]Thorpe S J, Gautrais J. Rapid visual processing using spike asynchrony[J]. Advances in neural information processing systems, 1997: 901-907.
- [24]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.
- [25]Cevikbas I, Yildirim T. Digit recognition in a simplified visual cortex model[C]//Innovations in Intelligent Systems and Applications (INISTA), 2013 IEEE International Symposium on. IEEE, 2013: 1-3.
- [26]黄丽鸿, 谌先敢, 刘海华. 模拟初级视皮层脉冲神经元的动作识别系统[J]. 自动化学报, 2012, 38(12): 1975-1984.
- [27]Luvizotto A, Rennó-Costa C, Pattacini U, et al. The encoding of complex visual stimuli by a canonical model of the primary visual cortex: Temporal population code for face recognition on the iCub robot[C]//Robotics and Biomimetics (ROBIO), 2011 IEEE International Conference on. IEEE, 2011: 313-318.
- [28]Eckhorn R, Reitboeck H J, Arndt M, et al. Feature linking via synchronization among distributed assemblies: Simulations of results from cat visual cortex[J]. Neural Computation, 1990, 2(3): 293-307.
- [29]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.

- [30]章海军. 视觉及其应用技术[M]. 浙江大学出版社, 2004.
- [31]罗四维. 视觉感知系统信息处理理论[M]. 电子工业出版社, 2006.
- [32]陈金龙. 面向图像处理应用的生物视觉感知计算模型研究[D]. 杭州电子科技大学, 2013.
- [33]De Valois R L, Albrecht D G, Thorell L G. Spatial frequency selectivity of cells in macaque visual cortex[J]. Vision Research, 1982, 22(5): 545-559.
- [34]ALindeberg T. A computational theory of visual receptive fields[J]. Biological cybernetics, 2013, 107(6): 589-635.
- [35]Nauhaus I, Benucci A, Carandini M, et al. Neuronal selectivity and local map structure in visual cortex[J]. Neuron, 2008, 57(5): 673-679.
- [36]杨利建. 从神经元到神经网络:随机动力学研究[D]. 华中师范大学, 2007.
- [37]张艳山. 基于视觉感受野模型的图像去噪算法研究[D]. 电子科技大学, 2015.
- [38]尧德中, 李朝义, 杜馨瑜,等. 一种基于视觉非经典感受野模型的图像去噪方法:, CN 101930592 A[P]. 2010.
- [39]李武, 李朝义. 视觉感受野外整合野研究的进展[J]. 神经科学, 1994(2):1-6.
- [40]罗佳骏, 武薇, 范影乐,等. 基于视觉感光层功能的菌落图像多强度边缘检测研究[J]. 中国生物医学工程学报, 2014, 33(6):677-686.
- [41]李康群, 范影乐, 甘海涛,等. 基于视通路多感受野朝向性关联的轮廓检测方法[J]. 中国生物医学工程学报, 2017, 36(1):1-11.
- [42]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.
- [43]陈可. 弱信号随机共振检测机制及其在图像增强中的应用研究[D]. 杭州电子科技大学, 2011.
- [44]Cofré R, Cessac B. Dynamics and Spike Trains Statistics in Conductance-based Integrate-and-Fire Neural Networks with Chemical and Electric Synapses[J]. BMC Neuroscience, 2013, 14(Suppl 1): P58.
- [45]罗佳骏. 基于视觉神经网络计算模型及其在图像边缘检测中的应用研究[D]. 杭州电子科技大学, 2015.
- [46]蔡荣太, 吴庆祥. 基于脉冲神经网络的边缘检测[J]. 微电子学与计算机, 2010, 27(10):178-181.
- [47]杜超本, 贾振红, 覃锡忠,等. 基于 NSCT 的遥感图像模糊增强算法[J]. 计算机工程, 2012, 38(4):188-190.

- [48] 窦燕, 王柳锋, 孔令富. 一种视皮层非经典感受野的模型[J]. 燕山大学学报, 2009, 33(2):109-113.
- [49] 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.
- [50] 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.
- [51] Grigorescu C, Petkov N, Westenberg M. Contour detection based on nonclassical receptive field inhibition[J]. Image Processing, IEEE Transactions on, 2003, 12(7): 729-739.
- [52] 顾凡及, 梁培基. 神经信息处理[M]. 北京: 北京工业大学出版社, 2007.
- [53] 寿天德. 视觉信息处理的脑机制[M]. 第2版. 合肥: 中国科学技术大学出版社, 2010.
- [54] 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.
- [55] 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.
- [56] Chan W, Coghill G. Text analysis using local energy[J]. Pattern Recognition, 2001, 34(12): 2523-2532.
- [57] Liang Z, Shen W, Shou T. Enhancement of oblique effect in the cat' s primary visual cortex via orientation preference shifting induced by excitatory feedback from higher-order cortical area 21a[J]. Neuroscience, 2007, 145(1): 377-383.
- [58] Vanrullen R, Thorpe S J. Surfing a spike wave down the ventral stream[J]. Vision research, 2002, 42(23): 2593-2615.
- [59] Anderson J R. Cognitive psychology and its implications[M]. Macmillan, 2005.
- [60] 廖进文. 基于视觉编码机制的图像弱边缘检测技术及其应用研究 [D]. 杭州: 杭州电子科技大学, 2015.
- [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.

附录

作者在读研期间获得的的学术成果以及参与的科研项目

发表(录用)学术论文：

1. 基于突触连接视通路方位敏感的图像分级边缘检测方法.中国生物医学工程学报(一级期刊), 2015, 34(5):522-532, 3/4.
2. 基于神经元突触可塑机制图像边缘检测方法.华中科技大学学报(EI 期刊), 2015(s1), 4/4.

授权(申请)专利：

发明专利

1. 一种基于多重随机共振机制的图像边缘检测方法, 已授权, 201310739932.3, 3/5.

获得的科研奖项：

1. 第三届全国研究生智慧城市智能技术挑战赛人脸识别项目组, 国家三等奖, 3/4, 2016 年.

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), 国家自然科学基金项目。
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。