

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2018)02-0182-12

论文引用格式: Guo J P, Fan Y L, Wu W, Zhu Y P. Contour detection method based on the response and fusion model of visual pathway [J]. Journal of Image and Graphics, 2018, 23(2): 0182-0193. [郭金朋, 范影乐, 武薇, 朱亚萍. 引入视通路视觉响应与融合模型的轮廓检测 [J]. 中国图象图形学报, 2018, 23(2): 0182-0193.] [DOI: 10.11834/jig.170313]

引入视通路视觉响应与融合模型的轮廓检测

郭金朋, 范影乐, 武薇, 朱亚萍

杭州电子科技大学自动化学院, 杭州 310018

摘要: 目的 为了提高轮廓检测的综合性能,特别是增强弱轮廓边缘的提取能力,在结合视觉机制的基础上提出了本文方法。方法 模拟视觉信息在视通路中的传递和处理过程,首先根据神经节细胞的中心周边拮抗机制,实现初级轮廓信息的快速提取;接着利用高斯函数与高斯差函数之间的差异性来模拟外膝体非经典感受野的调制作用,实现纹理背景的抑制;然后构建了一种 V1 区多朝向简单细胞感受野模型,提出了一种基于负值效应的 DOG (difference of Gaussians) 响应改进评价模式;最后考虑 V1 区复杂细胞在表征视觉高级特征的能力,给出了一种基于并行处理的视通路视觉响应融合模型,实现目标轮廓的检测与增强。结果 为了验证本文方法对自然场景图像的轮廓检测具备有效性,本文选取 RuG 轮廓检测数据库中的 40 幅自然场景图进行轮廓检测实验,并与二维高斯导函数模型(DG)、组合感受野模型(CORF)和空间稀疏约束纹理抑制模型(SSC)等3种典型的自然图像轮廓检测方法进行了分析比较。结果表明,本文方法检测提取到的主体轮廓更加完整,具有较高的图像纯净度,整体上反映了本文所提轮廓检测方法所具备的生物智能性。本文方法的平均 P 指标为 0.45,相较于对比方法具有更好的轮廓检测性能。结论 本文方法具有较好的自然轮廓检测提取能力,尤其对于图像包含部分弱轮廓边缘的检测。本文构建的新模型将有助于对视通路中各层级功能和内在机制的理解,也将为基于视觉机制的图像分析和理解提供一种新的思路。

关键词: 轮廓检测; 多感受野; DOG 负值效应; 多视通路; 视觉机制

Contour detection method based on the response and fusion model of visual pathway

Guo Jinpeng, Fan Yingle, Wu Wei, Zhu Yaping

School of Automation, Hangzhou Dianzi University, Hangzhou 310018, China

Abstract: Objective Visual information is the main source of human perception of outside information. The visual system of the human brain, as the most important means of obtaining information from the outside world, has a near-perfect information processing capability, which is far superior to existing computer vision systems in all aspects. The model description of the visual information processing mechanism can provide a novel way of solving engineering application problems, such as image analysis and understanding. Therefore, the study of the visual information processing mechanism has become an im-

收稿日期: 2017-06-26; 修回日期: 2017-10-11; 预印本日期: 2017-10-18

基金项目: 国家自然科学基金项目(61501154)

第一作者简介: 郭金朋(1991—),男,现为杭州电子科技大学自动化学院控制科学与工程专业硕士研究生,主要研究方向为模式识别与数字图像处理。E-mail: g-4018@qq.com

通信作者: 范影乐,教授,E-mail: fan@hdu.edu.cn

Supported by: National Natural Science Foundation of China (61501154)

portant direction in brain and cognitive science research. The complexity of the visual system lies in the complexity of information transmission between neurons; multiple information paths exist, and high cortical information demonstrates feedback regulation. Visual computing is an important means of studying visual information processing mechanisms and promoting the development of computer vision-related applications. Researchers can study the coding and processing of visual information from different ranges, such as microcosmic to macroscopic and molecular to behavioral, with the continuous improvement of the technical means of visual mechanism research. However, only the experimental data, which are organized organically from different levels and angles, can help determine the laws and mechanisms of nature. Contour detection is crucial to understanding the function and application of a high cortex visual perception. **Method** This study considers the process of visual information transmission by taking the entire visual paths as the object in studying the visual response characteristics of different mechanisms in the paths and constructing the visual fusion model of multiple visual pathways. The response model of the antagonistic mechanism of ganglion cells in the pathway is improved, the negative value of the primary contour response is preserved, and several features of the non-classical receptive field of the lateral geniculate nucleus are enhanced. We designed and implemented a multi-oriented simple cell receptive field model based on the DOG negative effect and constructed a visual fusion model of the complex cells of the primary visual cortex to suppress texture and enhance contours through the visual information differences among various visual pathways. We simulated the transmission and processing of visual information in the visual pathway. First, we realized the rapid extraction of primary contour information according to the antagonistic mechanism of ganglion cells. Then, the difference between the Gaussian function and the DOG function was constructed to simulate the modulation of the non-classical receptive field of the LGN, which could suppress the background texture. A multi-oriented receptive field model of a simple cell in the V1 region was constructed, and an improved evaluation model that considers the negative effect of DOG was proposed. A visual response fusion method based on parallel processing was provided to enhance target contour given the capability of the complex cells in the V1 region to represent advanced visual features. **Result** Visual test and quantitative calculation results show that the method has an enhanced contour detection capability in a natural image with complex background and can detect certain weak contour edge information in the image. The miss rate of the DG method is low, but the error rate is high. The CORF method reduces the error rate but increases the miss rate. The mistake rate of the CORF method is improved compared with the DG method, but the miss rate is increased. The overall performance of the CORF method remains low, although the overall performance is higher in the CORF method than in the DG method. The SSC method strengthens the texture suppression while retaining the main contour and achieving improved detection results. However, the SSC method produces additional burrs at the periphery of the main contour, thereby resulting in insufficiently smooth contour lines. The proposed method has clear background and contours, thereby effectively suppressing the background of the texture and enhancing the contour of the subject. The method achieves a certain balance between the error and miss rates, thus improving the overall performance. In addition, the method can effectively suppress the texture background of the adjacent area of the subject contour, in which the extracted contour lines are smooth, and the burr phenomenon in the SSC method is removed. This study selected 40 natural scene graphs from the RuG contour detection database for contour detection experiments and compared them using three typical methods of natural image contour detection, namely, DG, CORF, and SSC, to verify the effectiveness of the proposed method in the contour detection of natural scene images. Results show that the main contours detected by the method proposed in this paper are complete, and the image purity is high. Overall, results reflect the biological intelligence of the proposed contour detection method. The average P index of the proposed method is 0.45, which indicates a better contour detection performance than the contrast methods. **Conclusion** In this paper, we improved the classical receptive field responses of ganglion cells in the visual pathway. We studied the enhancement effect of the LGN cells in the visual pathway by considering the visual pathway the main body of the non-classical receptive field regulation mechanism. We focused on the negative effects of the DOG produced by the antagonistic mechanism of ganglion cells and designed a multi-oriented simple cell receptive field model. We introduced a parallel mechanism in multiple visual pathways for visual information processing, and the visual pathway was divided into the main and vice paths. We proposed a method that uses the visual information difference of the different visual pathways for suppressing the texture and enhancing the contours. The proposed method has enhanced natural contour detection and extraction capabilities, especially in detecting several weak contour edges in images. The new model constructed

in this study will help in elucidating the function and internal mechanisms of the visual pathway and provide a new approach for image analysis and understanding based on visual mechanism.

Key words: contour detection; multiple receptive fields; negative value effect of DOG; multiple visual pathways; visual mechanism

0 引言

边缘是指在图像像素与其局部邻域之间在特定方向上的颜色或灰度的跃变^[1],连续体现图像中主体目标的边缘则是轮廓。在机器视觉和图像分析中,轮廓检测属于基础性工作,准确检测图像中的主体目标轮廓,将为后续目标识别以及图像理解等高级视觉行为提供丰富的关键信息^[2]。

传统的边缘检测方法包括 Roberts 算子和 Canny 算子等利用数学模型描述局部邻域像素在特定方向上灰度梯度变化^[2-3]的方法,这些方法虽然简洁高效,但难以有效区分主体轮廓和纹理背景。后续出现的多尺度信息^[4-5]、多局部特征^[6]、上下文全局分析^[7]等检测方法,对轮廓的检测效果有所改善,但仍远远落后于人类视觉系统对于复杂背景下快速准确提取轮廓信息的能力。因此,Grigorescu 等人^[8]模拟人类视觉机制特性,采用 2 维 Gabor 能量模型模拟 V1 区简单细胞的方向选择特性,并提出了非经典感受野的抑制特性进行轮廓检测,有效减少了纹理背景的影响。此外,还有很多基于视觉机制的感受野改进型轮廓检测方法。例如,桑农等人^[9]提出了蝶形感受野抑制模型,较好地去除了纹理背景;Du 等人^[10]提出了复合感受野模型,既减少了纹理背景干扰,又保护了弱轮廓;Yang 等人^[11]利用视锥细胞颜色双拮抗机制和空间稀疏特性进行纹理抑制,提取彩色图像轮廓;而另一些科研人员则专注于对底层神经元建模^[12],构建图像处理网络。需要注意的是,上述方法大多是基于局部区域描述算子构建的方法,而忽视了视觉信息在视通路中传递和处理的过程性,也没有考虑到视通路视觉信息融合的层面。这种只对视通路中的局部视觉机制进行模拟的方法,会忽略视通路上不同机制之间和不同视通路之间的相互影响,因此导致轮廓检测的误检率和漏检率相对较高。

基于以上分析,本文充分考虑了视觉信息传递的过程性,将整个视通路作为对象进行研究,研究了

视通路中不同机制的视觉响应特征,构建了多视通路视觉融合模型。本文对视通路中神经节细胞的中心周边拮抗机制响应模型进行了改进,保留了初级轮廓响应中的负值,改进了外膝体非经典感受野调节的部分特性,设计并实现了一种基于 DOG 负值效应的多朝向简单细胞感受野模型,构建了初级视皮层复杂细胞的视觉融合模型,通过不同视通路间的视觉信息差异来抑制纹理和增强轮廓。

1 原理与实现

1.1 基本原理

1.1.1 神经节细胞层响应模型

视觉系统中,视网膜部分呈 3 层分级结构,即光感受器细胞层、连接细胞层和神经节细胞层。光感受器细胞层通过连接细胞层汇聚连接到神经节细胞层。相关研究表明,大部分神经节细胞表现出中心周边拮抗机制,而其输入的汇聚连接结构则是这一机制的生理学基础,汇聚连接的光感受器细胞的数量即为神经节细胞的感受野大小^[13]。本文采用基于空间距离变化的 2 维双高斯差函数对神经节细胞的中心周边拮抗机制进行模拟。高斯函数及其参数形式为

$$gaus(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (1)$$

式中, x 和 y 为 2 维空间坐标, σ 为标准差。

本文对神经节细胞层的响应模型进行了改进,舍弃了使用极大值函数将响应负值简单归零的操作,从而保留初级轮廓响应的负值,以增强视通路中向后传递视觉信息的有效性,改进后的神经节细胞层响应模型为

$$Id(x, y) = I(x, y) * [gaus(x, y, k \times \sigma) - gaus(x, y, \sigma)] \quad (2)$$

式中, k 为两个高斯函数的尺度比率, $I(x, y)$ 表示光感受器细胞层的图像亮度感知编码信号。

视网膜部分神经元只是对输入图像的空间差异性信息进行粗略检测,还不能有效区分主体轮廓和

纹理背景,但这一视觉机制的响应将为视通路上更高层次的视觉处理提供丰富的原始数据。

1.1.2 外膝体细胞的非经典感受野调节机制

视网膜中光感受器细胞的总数是神经节细胞的数百倍之多,而外膝体细胞则与神经节细胞数目几乎相等。因此本文认为外膝体细胞层是视觉信息传输处理的一个中继站,对神经节细胞输出的空间差异性图像信息具有调制增强作用。相关研究指出,外膝体细胞能够综合双眼输入信息^[14],而且在外膝体细胞经典感受野的大外周存在一块特殊的区域,即非经典感受野。单独刺激非经典感受野区域并不会引发神经元的放电活动,但其对经典拮抗型感受野的响应具有一种调节效应^[15]。对神经节细胞输出的中心周边拮抗响应特征图进行非经典感受野调

节,能够抑制空间差异性图像信息中非稀疏性的纹理,同时也能在一定程度上增强稀疏性的主体轮廓。采用高斯函数与高斯差函数之间的差异性来模拟非经典感受野的调制作用,即

$$TD(x, y) = \text{gaus}(x, y, \sigma) - \text{dog}^+(x, y, \Delta\sigma) \quad (3)$$

$$\text{dog}^+(x, y, \Delta\sigma) = \max \left\{ \frac{\text{gaus}(x, y, 4\sigma) - \text{gaus}(x, y, \sigma)}{\text{gaus}(x, y, \sigma)}, 0 \right\} \quad (4)$$

生理学研究表明,非经典感受野的直径大约为经典感受野的 4 倍^[16]。本文将这种差异性进行了全局整型并通过指数函数归一化为非经典感受野的调节量 $aj(x, y)$, 将调节量与前级神经节细胞响应图按权值相融合,然后将其作为外膝体细胞的输出 $Ig(x, y)$, 具体为

$$aj(x, y) = \exp \left(- \frac{(\max \{ \sum_{(\Delta x, \Delta y)} Id(x + \Delta x, y + \Delta y) \times TD(\Delta x, \Delta y), 0 \})^2}{2\delta^2} \right) \quad (5)$$

$$Ig(x, y) = Id(x, y) - \alpha \times aj(x, y) \quad (6)$$

式中, δ 是抑制强度调节参数,其范围通常为 0.1 ~ 0.5, α 是非经典感受野调节系数,反映了非经典感受野调节的作用强度, $(\Delta x, \Delta y)$ 表示像素点的增量,使用 \max 函数进行全局整型。

外膝体细胞非经典感受野的调节能够在一定程度上抑制纹理背景,同时也在一定程度上对前级神经节细胞响应结果中的主体轮廓进行增强。

1.1.3 V1 区基于 DOG 负值效应的多朝向简单细胞感受野模型

V1 区简单细胞具有方向选择性,不同的简单细胞具有不同的偏好方向。某一 V1 区简单细胞对其偏好方向的刺激产生最大响应,而对其他方向的刺激响应减弱或消失。有研究认为神经节细胞具有某种方位选择性^[17],但更多生理实验表明 V1 区简单细胞才是方位选择的主体。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 2 维 Gabor 能量函数作为 V1 区简单细胞感受野的计算模型。但是 Gabor 能量函数实际上只是对简单细胞的一种黑箱模拟,没有考虑简单细胞方向选择特性的形成机制,更没有充分考虑前级细胞层 DOG 响应形成的负值信息。

V1 和外膝体细胞之间存在反馈突触连接,因此 V1 区简单细胞的感受野特性与前级外膝体细胞的感受野有着复杂的联系。Hubel 和 Wiesel 认为初级

视皮层可以由多个相邻的外侧膝状体细胞的同心圆式感受野相连而成,简单细胞的方向选择性与多个外侧膝状体细胞感受野的关联方式密切相关^[18]。

本文重点考察了神经节细胞的中心周边拮抗机制在轮廓边缘的响应特征,提出了 DOG 边界响应的负值效应的概念。由于对神经节细胞层的响应模型进行了改进,保留了响应的负值,故式(2)中 $Id(x, y)$ 值可正可负。图像在亮区和暗区之间的阶跃变化形成边界,神经节细胞感受野由边界亮区向暗区移动时,周边感受野亮度刺激减弱,总体感受野响应呈现兴奋且逐渐增强, $Id(x, y)$ 为正值;当中心感受野到达边界暗区时亮度刺激陡降,总体感受野响应呈现抑制且逐渐减弱, $Id(x, y)$ 为负值。反之,神经节细胞由边界暗区向亮区移动时,会出现相反且类似的结果。整个过程中总体感受野的响应呈现出一种类似心电图的上下波动形式。本文将感受野总体响应呈现抑制的状态,即 $Id(x, y)$ 为负值,定义为 DOG 响应的负值效应。在基于视觉机制的图像信息处理中,人们几乎都忽略了 DOG 响应的负值效应,对 DOG 产生的负值简单整合为 0,以保持图像的亮度值处于 $[0, 1]$ 或者 $[0, 255]$ 之间的合理范围。但事实上 DOG 响应的负值效应并非冗余信息,简单剔除 DOG 负值数据的轮廓检测,其响应特征往往不够突出,这对于弱轮廓图或者图中的弱轮廓部分来说,将会增大轮廓检测的漏检率,造成轮廓不完整,

也给后续的阈值处理部分造成困难。本文认为特定点处的边界响应与其邻域各点相关,如果通过对 DOG 的负值进行处理,使其也产生正向的边界响应并累加到原来的边界响应中,再经过负值归零之后,那么这时边界处将会出现增强的响应特征。为了充分利用感受野在图像边界处表现出的这种负值效应,本文将构建一个基于 DOG 负值效应的多朝向简单细胞感受野模型,用于进行轮廓检测。

为了处理方便,用尺寸为 $s \times s$ 的矩型感受野作为基础感受野,记为 $rmd(x, y)$, 取 s 为奇数以便计算。 $x_0 = y_0 = \frac{(s+1)}{2}$ 为 $rmd(x, y)$ 的中心坐标, 则 $r_0 = rmd(x_0, y_0)$ 为中心点,以 r_0 点为坐标原点建立直角坐标系。在 $rmd(x, y)$ 上建立一条具有特定朝向并通过原点的直线,用来表征 V1 区简单细胞的朝向偏好,定义为简单细胞最优朝向线,即

$$y = x \times \tan(\theta) \quad (7)$$

$$x, y \in rmd, \theta \in [0, 2\pi)$$

式中, θ 为 V1 简单细胞的偏好朝向角度, x 和 y 为 rmd 直角坐标系中的坐标。将简单细胞最优朝向线所在区域锯齿化,即

$$Z = y - x \times \tan(\theta) \quad (8)$$

$$rmd(x, y, \theta_i) = \begin{cases} 1 & Z = 0 \text{ 且 } \text{mod}(x, 2) = 0 \\ -1 & Z = 0 \text{ 且 } \text{mod}(x, 2) \neq 0 \end{cases} \quad (9)$$

式中, mod 为模运算。特别当式(8)中的 θ 为 90° 和 270° 时, $\tan(\theta)$ 分别为正无穷和负无穷, Z 依据 x 的

正负表示正无穷或者负无穷,而与 y 无关。 Z 为 0 表示 rmd 区域的最优朝向线,将其依据奇偶进行锯齿化。为了充分利用 DOG 响应的负值效应,需要将 V1 简单细胞最优朝向线的两侧区域进行差异化,即

$$rmd(x, y, \theta_i) = \begin{cases} 1 & Z > 0 \\ -1 & Z < 0 \end{cases} \quad (10)$$

由于细胞连接表现出空间距离的差异性,因此本文引入 2 维高斯函数对矩形基础感受野进行加权,这样做既使简单细胞的感受野模型产生空间距离调制,又使矩形感受野表现出圆形感受野的特性,具体形式如式(11)所示。单方向的响应如式(12)所示。

$$md(x, y, \theta_i) = rmd(x, y, \theta_i) \times \text{gaus}(x, y, \sigma) \quad (11)$$

$$lmd(x, y, \theta_i) = lg(x, y) * md(x, y, \theta_i) \quad (12)$$

图 1 为多朝向简单细胞感受野模型的轮廓检测示意图,从图 1(b)(c) 局部轮廓及其响应的对比结果可以明显看出,本文所提出的基于 DOG 负值效应多朝向简单细胞感受野模型能够很好的增强主体轮廓,并抑制纹理。图 2 给出了多朝向简单细胞感受野模型的部分方向选择结果。其中,图 2(a) 为各方向图像的最佳响应融合后的输出结果,主体轮廓已被增强凸显,纹理背景受到抑制。图 2(b) — (f) 为部分特定偏好方向的简单细胞选择结果,由于综合考虑了方向与尺度因素,因此特定方向的稀疏轮廓被选择增强,杂乱方向的非稀疏纹理被抑制。

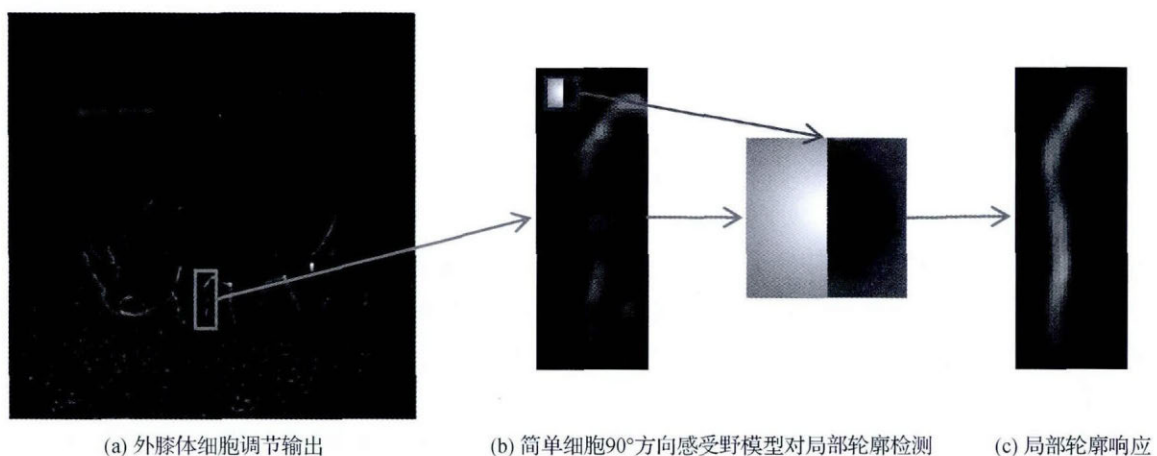


图 1 多朝向简单细胞感受野模型的轮廓检测示意图

Fig. 1 Sketch map of contour detection using multi-oriented simple cell receptive fields model. ((a) regulated output of LNG cells; (b) 90 degree direction receptive field model of simple cells for local contour detection; (c) local contour response)

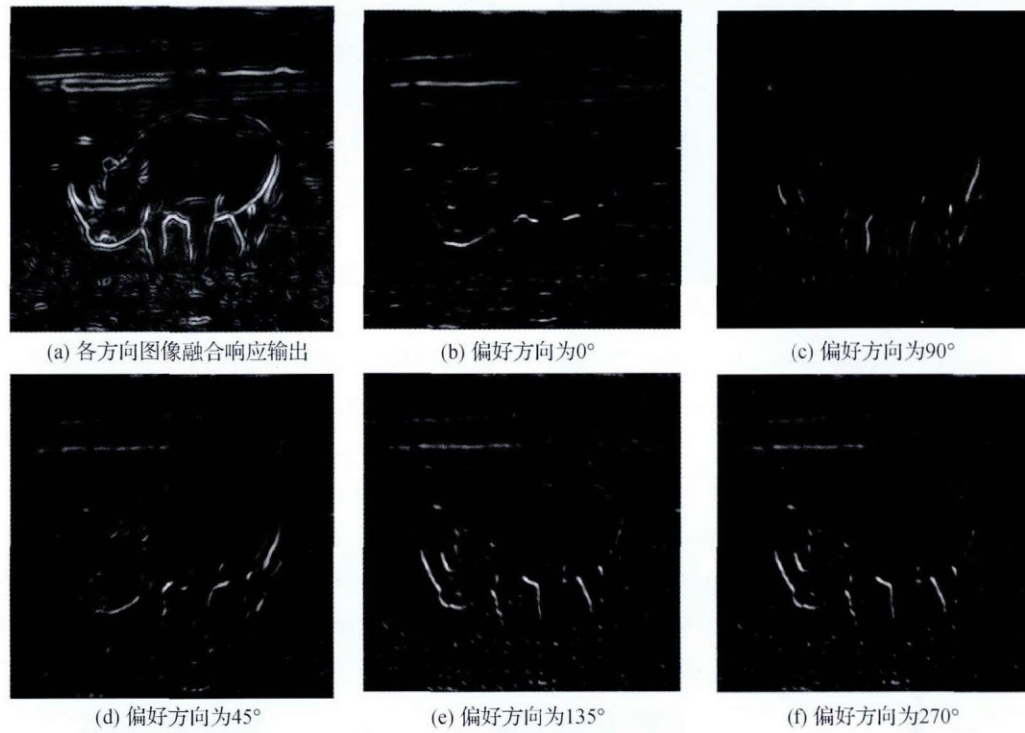


图2 多朝向简单细胞感受野模型的部分方向选择结果

Fig. 2 Some orientation selective results using multi-oriented simple cell receptive fields model ((a) integrated output of all orientations; (b) — (f) orientation selective result of simple cells when the preferred orientations are 0° , 90° , 45° , 135° and 270°)

1.1.4 V1区复杂细胞的多通路视觉信息融合

相关研究已经表明,从视网膜到视皮层存在多条并行的视通路,视觉信息在不同的视通路上分级并行处理。本文充分考虑了这种多通路并行视觉机制,认为在视觉信息处理上存在一条主视通路,并将其其他的视通路统归为副视通路。V1区不仅存在简单细胞,还存在复杂细胞,复杂细胞的感受野相对较大,接受简单细胞的输出信号,能够整合双眼的信息输入,产生融合的图像信息响应。复杂细胞能够对主视通路和副视通路上的图像信息进行整合,利用不同视通路间视觉信息的差异性来抑制纹理和增强轮廓。副视通路感受野对特定朝向条形光刺激的响应具有选择性这一生理特性是副视通路对轮廓及形状感知的基础,考虑到2维高斯导函数能够提取图像多方向边界信息,采用该函数来模拟副视通路上的图像信息处理,其表达式为

$$RF(x, y; \theta) = \frac{\partial g(\tilde{x}, \tilde{y}; \theta, \sigma)}{\partial \tilde{x}} \quad (13)$$

$$g(\tilde{x}, \tilde{y}; \theta, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\tilde{x}^2 + \gamma^2 \times \tilde{y}^2}{2\sigma^2}\right) \quad (14)$$

$$\tilde{x} = x \times \cos(\theta) + y \times \sin(\theta) \quad (15)$$

$$\tilde{y} = -x \times \sin(\theta) + y \times \cos(\theta) \quad (16)$$

式中, $\theta \in [0, \pi)$, γ 表示滤波器的椭圆率,其取值范围通常为 $0 \sim 1$ 。引入响应强度系数对其进行调节,即

$$e_i(x, y; \theta_i) = |I(x, y) * RF(x, y; \theta)| \quad (17)$$

$$\mu(x, y) = \frac{\max\{e_i(x, y; \theta_i)\}}{\sum_i e_i(x, y; \theta_i)} \quad (18)$$

$$\theta_i = \frac{(i-1)\pi}{N_\theta}; \quad i = 1, 2, \dots, N_\theta \quad (19)$$

副视通路的边界响应定义为

$$E(x, y) = \mu(x, y) \times \max\{e_i(x, y; \theta_i) \mid i = 1, 2, \dots, N_\theta\} \quad (20)$$

复杂细胞层利用主视通路与副视通路图像信息响应信号的差异性进行相互融合增强,即

$$I_o = E(x, y) \times Imd(x, y) \quad (21)$$

式中, I_o 为主副视通路融合响应。图3为复杂细胞相应的响应结果,从图3(a)(b)对比结果来看,主副视通路信息的融合能够很好的抑制纹理,凸显轮廓。

得到最终的轮廓响应后,还需要进行后续处理

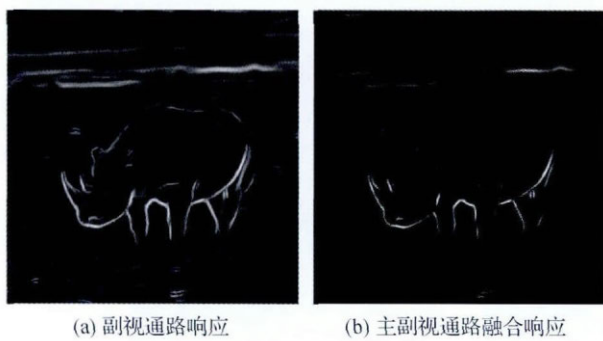


图3 复杂细胞响应

Fig. 3 The response of complex cells ((a) response of second visual pathway; (b) integrated response of main visual pathway and second visual pathway)

以产生单像素轮廓,以便于对算法性能进行量化评价。采用 Canny 算法中的非极大值抑制和滞后阈值处理方法进行后续处理,非极大值抑制处理产生单像素轮廓,滞后阈值处理产生二值轮廓图。

1.2 算法流程

轮廓检测算法流程如图4所示,具体轮廓检测流程如下:

- 1) 按式(1)一式(2)构建神经节细胞响应模型,并对输入灰度图像进行处理。
- 2) 按式(3)一式(6)构建外膝体细胞非经典感受野调节机制,对前级输入进行调节。
- 3) 按式(7)一式(12)构建 V1 区基于 DOG 负值效应的多朝向简单细胞感受野模型,对轮廓进行多方向检测。
- 4) 按式(13)一式(20)构建副视通路图像信息响应,并按式(21)模拟 V1 区复杂细胞的多通道视觉融合机制,输出最终的轮廓响应。
- 5) 对最终轮廓响应进行后续处理,进行非极大值抑制产生单像素轮廓,然后进行滞后阈值处理,得到二值轮廓图。

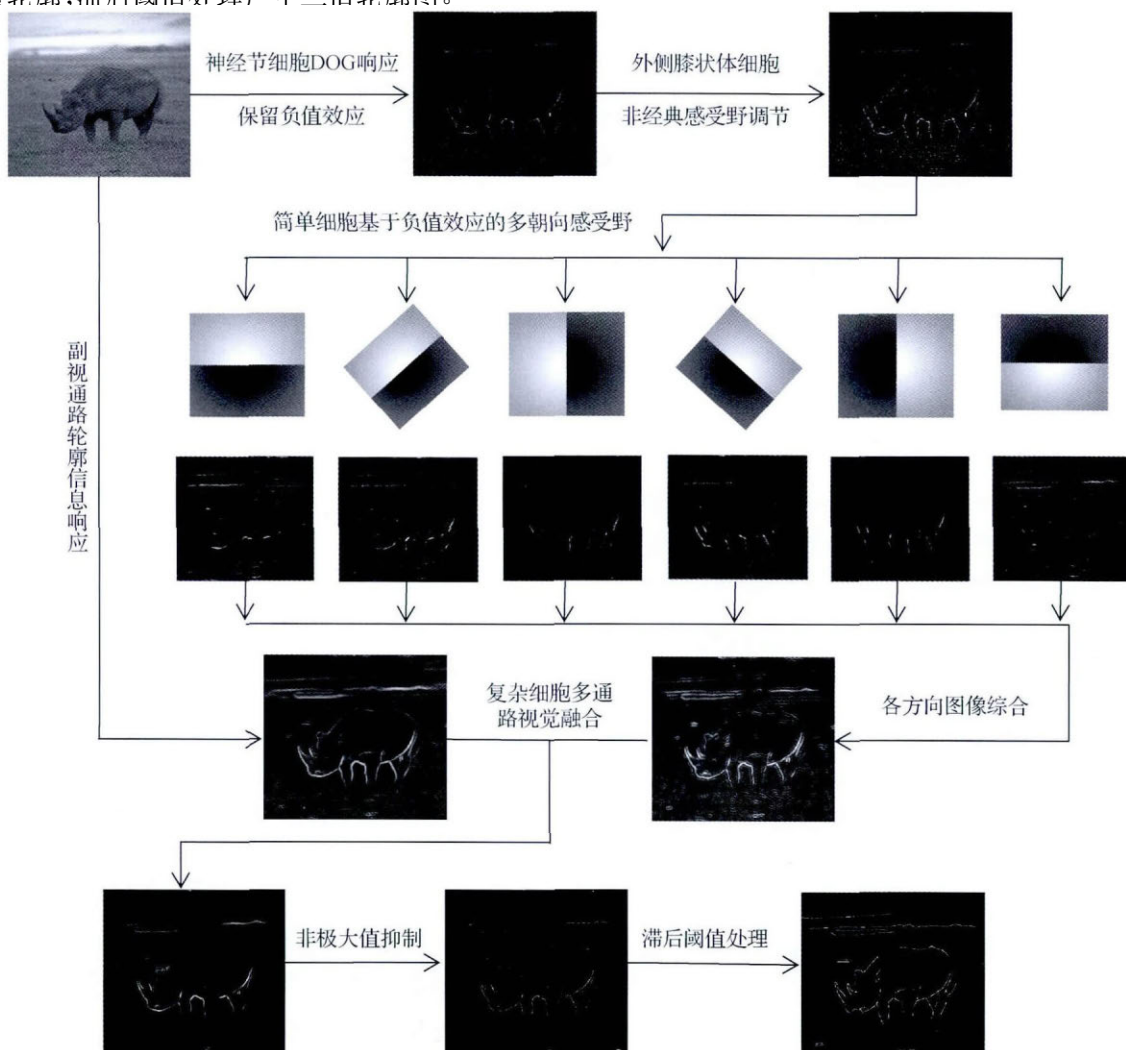


图4 轮廓检测算法流程

Fig. 4 The procedure of contour detect algorithm

1.3 算法验证

本文选取 RuG 轮廓检测数据库中的 40 幅自然场景图进行实验,以验证本文方法对自然场景图像的轮廓检测具备有效性。每幅图像对应有一张手工绘制的基准轮廓图,本文选取其中部分有代表性的图像作为实验对象,如图 5 所示。

本文选择了 3 种典型的自然图像轮廓检测方法作为对比,以说明本文方法的有效性。首先选取以数学模型直接模拟初级视皮层轮廓响应的 2 维高斯导函数模型(DG)进行对比,体现本文方法中 V1 区基于 DOG 负值效应多朝向简单细胞感受野模型的

生物合理性。再选取以 LGN 细胞排列组合的轮廓检测模型^[19](CORF),以证明本文方法中 V1 区复杂细胞多通路视觉融合机制的合理性和有用性。最后选择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^[11](SSC)作为对比方法。对比实验中,选取 11 个不同的非经典感受野调节系数,对应参数为 $\alpha = [0.5 : 0.1 : 1.5]$,在滞后阈值处理中,取上限阈值 $t = [0.1 : 0.05 : 0.8]$ 。对于本文方法和 SSC 方法可得到 150 组测试参数,而由于 DG, CORF 方法并不涉及系数 α ,故只进行 15 组测试参数。

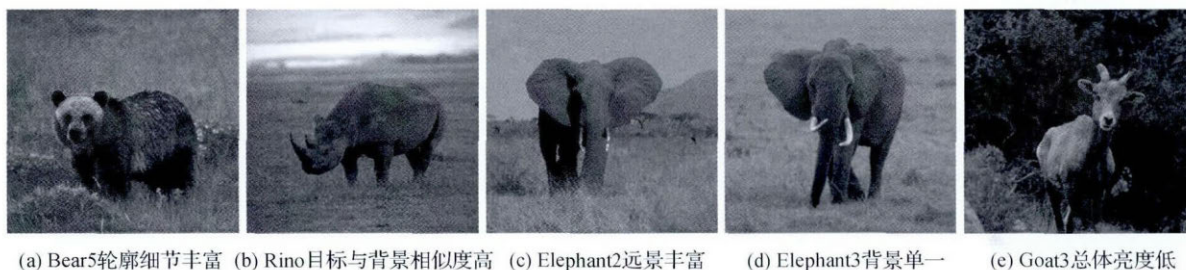


图 5 RuG 图库中的部分图片

Fig. 5 Some images in RuG image library((a) Bear5 with rich contour details; (b) Rino with high similarity to background; (c) Elephant2 with rich details on deep view; (d) Elephant3 with single background; (e) Goat3 with low overall brightness)

基于 RuG 图库中基准轮廓图,采用 Grigorescu 等人^[8]提出的定量检测轮廓评价指标,并设定基准轮廓点的 5×5 邻域内出现的轮廓点为正确轮廓点。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 5×5 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 分别为

$$e_{FP} = \frac{\text{card}(E_{FP})}{\text{card}(E)} \quad (22)$$

$$e_{FN} = \frac{\text{card}(E_{FN})}{\text{card}(E_{GT})} \quad (23)$$

$$P = \frac{\text{card}(E)}{\text{card}(E) + \text{card}(E_{FP}) + \text{card}(E_{FN})} \quad (24)$$

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

2 结 果

用轮廓检测定量评价指标对本文方法及对比方法的检测结果进行计算,计算结果列于表 1,各方法在最优参数下检测到的轮廓如图 6 所示。

从图 6 的直观检测结果和表 1 的量化计算结果中可以看出,本文方法对于复杂背景下的自然图像具有更好的轮廓检测提取能力,能够检测提取到图像中的部分弱轮廓边缘信息。从图 6(c) DG 方法的直观检测结果看,其优点在于检测出的轮廓细节多,漏检率较低,但其不足在于不能有效抑制纹理背景,误检率很高。这一点,从表 1 中 DG 方法的量化计算结果中也可得到印证,以 Bear5 为例,虽然其有 0.13 的极低漏检率,但由于其误检率高达 3.30,故整体性能指标偏低。因此,其直观结果表现为:背景杂乱,难以凸显主体轮廓。图 6(d) CORF 方法能在一定程度上减少背景纹理干扰,降低了误检率,但图像中的主体轮廓也缺失严重,造成漏检率较高。例

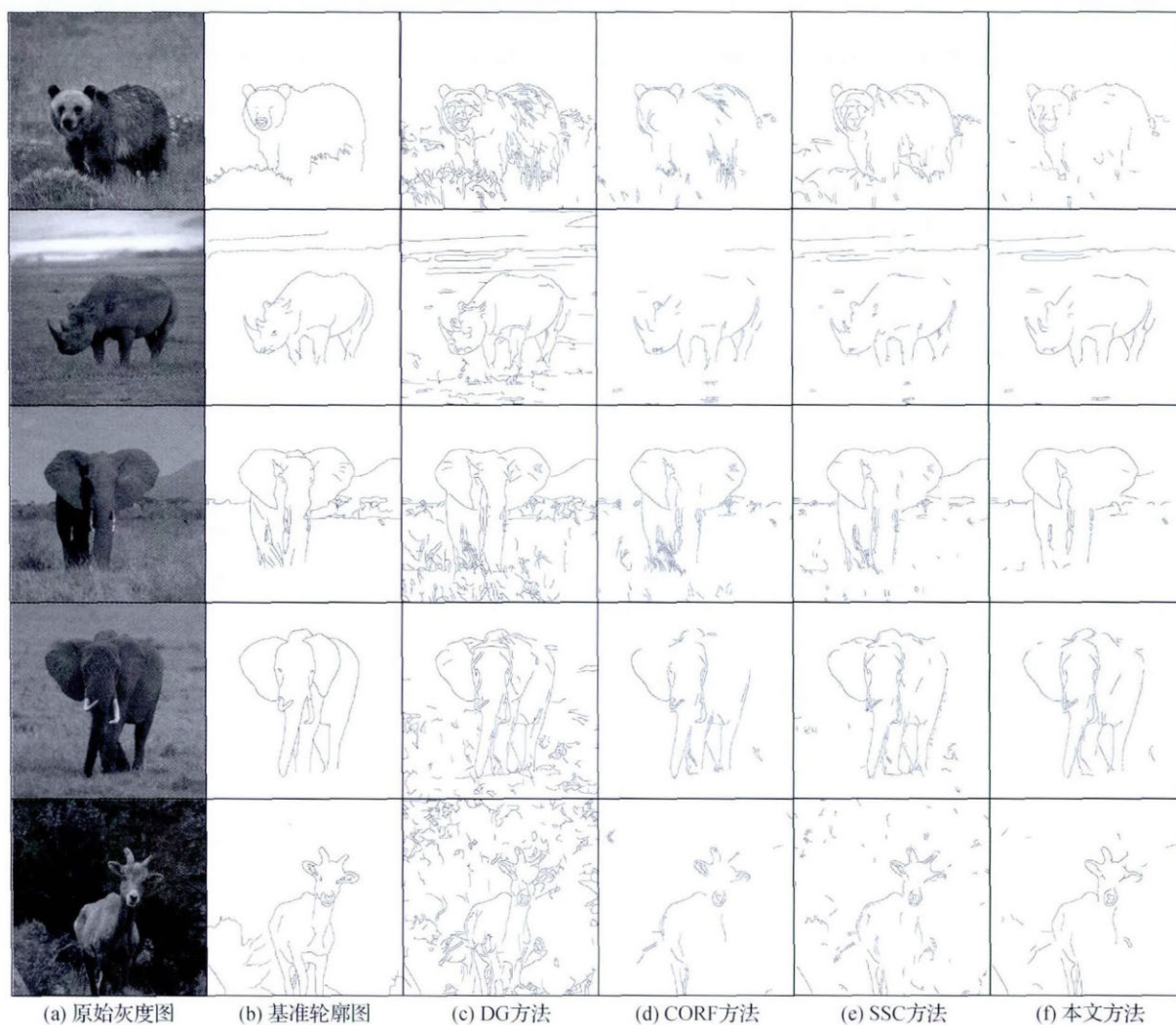


图6 图库中部分图像轮廓检测结果及不同方法的比对

Fig. 6 Contour detection results of some images in image library and the comparison of different methods

((a) original gray images; (b) reference contour images; (c) DG method; (d) CORF method; (e) SSC method; (f) ours)

如,该方法 Rino 检测结果中,犀牛的背部和后腿部轮廓断裂严重,天空中云的线条轮廓几乎全部漏检。从表 1 中量化计算结果看,CORF 方法的误检率较 DG 方法有所改善,但漏检率有所上升,虽然整体性能较 DG 方法有所提升,但仍然偏低。图 6(e) SSC 方法在保留主体轮廓的同时,加强了对纹理的抑制,取得了较好的检测结果,但在主体轮廓的周边产生较多毛刺,造成轮廓线条不够顺滑。相较于 CORF 方法,SSC 方法在漏检率和误检率上取得了一定程度的平衡,提高了整体检测性能,但对于图像中的部分弱轮廓边缘的检测仍然不够有效。图 6(f) 本文方法取得了背景干净、轮廓清晰的检测结果,有效抑制了纹理背景,增强了主体轮廓,在漏检率和误检率上同样取得了一定程度的平衡,使得整体性能表现

地更好。从表 1 可以看出,本文方法和 SSC 方法在漏检率和误检率上取得了不同程度的平衡,使得整体检测性能提高,而且本文方法要优于 SSC 方法。本文方法对于图像中部分弱轮廓边缘的检测提取具有相对较好的能力,从 Rino 的轮廓直观检测结果中可以看出,本文方法对于犀牛后腿部分的弱轮廓边缘检测要优于 SSC 方法。此外本文方法能够有效抑制主体轮廓邻接区域的纹理背景,使得检测提取的轮廓线条顺滑,去除了 SSC 方法中出现的毛刺现象。

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纹理背景边缘,在图像纯净度和轮廓清晰度上表现更好,具有更佳的自然图像轮廓检测能力。

表1 图6所示图像对应的参数设置及性能指标
Table 1 Parameter set and performance index of detect results in fig. 6

图像	算法	参数		性能指标		
		α	t	e_{FP}	e_{FN}	P
Bear5	DG		0.10	3.30	0.13	0.22
	CORF		0.50	1.63	0.47	0.28
	SSC	0.7	0.10	1.83	0.33	0.30
	本文方法	1.1	0.50	0.72	0.46	0.37
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
	本文方法	1.0	0.40	0.42	0.27	0.56
Elephant2	DG		0.10	1.46	0.11	0.38
	CORF		0.20	0.48	0.30	0.53
	SSC	1.1	0.10	0.23	0.36	0.54
	本文方法	1.3	0.30	0.09	0.42	0.54
Elephant3	DG		0.10	2.0	0.07	0.33
	CORF		0.20	0.32	0.27	0.59
	SSC	1.0	0.10	0.62	0.17	0.54
	本文方法	1.0	0.30	0.07	0.41	0.56
Goat3	DG		0.10	2.77	0.15	0.25
	CORF		0.40	0.32	0.57	0.37
	SSC	0.7	0.10	0.79	0.37	0.42
	本文方法	1.1	0.50	0.23	0.46	0.48

注: 加粗字体为本文最优结果。

3 讨论

以上实验结果已经论证本文方法具有更好的自然轮廓检测提取能力,尤其对于图像包含部分弱轮廓边缘的检测。从图6的直观检测结果和表1的量化计算结果中,都可以直接论证这一结论。

DG方法使用数学方法来模拟初级视皮层的轮廓响应过程,再用阈值检测出一定强度的轮廓。该方法对于轮廓细节的检测表现良好,但是使用阈值处理来抑制纹理背景,效果不佳,因此导致检测结果中伪轮廓较多,误检率急剧升高,降低了整体检测性能。CORF方法利用了外膝体细胞感受野的空间检

测特性,因此在一定程度上抑制了纹理背景的干扰,但没有考虑视通路上各层级细胞的分级加工和相互影响,这将导致该方法因为阈值处理不当或者过抑制而出现主体轮廓严重断裂。从图6(d)检测结果中可以看出,CORF方法的主要问题就在于主体轮廓缺失严重。从表1中CORF的结果也可以看出,CORF方法的漏检率普遍较高。SSC方法综合考虑了目标轮廓的显著特性和稀疏特性,在增强主体轮廓的同时,抑制了纹理背景的干扰,对误检率和漏检率两个指标进行了平衡,提升了整体检测性能,取得了较好的检测结果。但由于对主体轮廓邻近区域的纹理抑制效果不够突出,因此在富含纹理背景的主体轮廓上形成很多毛刺,致使轮廓线条不够顺滑,这

从图 6(e) 的结果可以看出, Elephant3 的检测结果尤为明显。本文方法考虑了神经节细胞感受野的中心周边拮抗机制, 根据其轮廓响应的负值效应以及 V1 区简单细胞的选择性来增强轮廓, 提取信息, 考虑了多视通路视觉信息的差异性, 并以此来抑制纹理, 最终实现轮廓的检测提取。另外, 关于本文算法的运算速度, 需要说明的是, 本文为了简化研究, 采用串行计算来模拟生物视觉系统的并行计算, 对于 RuG 图库中的单幅图像处理时间大概在 2~3 s 左右, 和其他基于视觉机制的轮廓检测算法大致相同。最后, 正如前述, 本文方法验证时针对基准轮廓图, 使用了 150 组测试参数搜寻最优结果, 但其实大部分实际应用场景并不具有基准轮廓图, 此时属于一种无监督模式决策问题, 因此需改变传统的漏检率和误检率等评价方式, 直接从目标轮廓本身的连续性、封闭性以及光滑性等特性进行评价, 设置依赖于检测结果本身的优化函数, 从而实现参数的自适应选择。

总的对比来看, 本文方法更加有效地去除了纹理背景, 保留了主体轮廓, 在漏检率和误检率上取得了不错的平衡结果, 具有更好的整体检测性能。

4 结 论

通过对视通路中神经节细胞的经典感受野响应进行了改进, 研究了外膝体细胞在视通路上的中继增强作用, 将其作为非经典感受野调节机制的主体, 重点考察了神经节细胞的中心周边拮抗机制产生的 DOG 负值效应, 并设计实现了一种多朝向简单细胞感受野模型, 同时引入了视觉信息处理的多通路并行机制, 将视通路划分为主视通路和副视通路加以研究, 并提出了利用不同视通路中视觉信息差异来抑制纹理和增强轮廓的方法, 设计了 V1 区复杂细胞的多视通路视觉融合模型。通过实验验证, 本文所提方法取得了良好效果, 并与 DG、CORF 和 SSC 等方法进行了分析比较。结果表明, 本文方法检测提取到的自然图像主体轮廓更加完整, 同时抑制了大部分的纹理背景边缘, 具有较高的图像纯净度, 整体上反映了本文所提轮廓检测方法所具备的生物智能性。本文提出的引入视通路视觉响应与融合模型的轮廓检测方法, 为图像轮廓检测技术提供了新的研究思路, 后续研究可基于视通路更高皮层的视觉

特性, 探讨本方法提取的轮廓特征如何在高级视觉层的图像理解与视觉认知中发挥作用。

参考文献(References)

- [1] Martin D R, Fowlkes C C, Malik J. Learning to detect natural image boundaries using local brightness, color, and texture cues [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(5): 530-549. [DOI: 10.1109/TPAMI.2004.1273918]
- [2] Papari G, Petkov N. Edge and line oriented contour detection: state of the art [J]. Image and Vision Computing, 2011, 29(2-3): 79-103. [DOI: 10.1016/j.imavis.2010.08.009]
- [3] Tang Q L, Sang N, Liu H H. Contrast-dependent surround suppression models for contour detection [J]. Pattern Recognition, 2016, 60: 51-61. [DOI: 10.1016/j.patcog.2016.05.009]
- [4] Azzopardi G, Rodríguez-Sánchez A, Piater J, et al. A push-pull CORF model of a simple cell with antiphase inhibition improves SNR and contour detection [J]. PLoS One, 2014, 9(7): #e98424. [DOI: 10.1371/journal.pone.0098424]
- [5] Mignotte M. A biologically inspired framework for contour detection [J]. Pattern Analysis and Applications, 2017, 20(2): 365-381. [DOI: 10.1007/s10044-015-0494-y]
- [6] Yang K F, Li C Y, Li Y J. Multifeature-based surround inhibition improves contour detection in natural images [J]. IEEE Transactions on Image Processing, 2014, 23(12): 5020-5032. [DOI: 10.1109/TIP.2014.2361210]
- [7] Arbelaez P, Maire M, Fowlkes C, et al. Contour detection and hierarchical image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(5): 898-916. [DOI: 10.1109/TPAMI.2010.161]
- [8] Grigorescu C, Petkov N, Westenberg M A. Contour detection based on nonclassical receptive field inhibition [J]. IEEE Transactions on Image Processing, 2003, 12(7): 729-739. [DOI: 10.1109/TIP.2003.814250]
- [9] Sang N, Tang Q L, Zhang T X. Contour detection based on inhibition of primary visual cortex [J]. Journal of Infrared and Millimeter Waves, 2007, 26(1): 47-51, 60. [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, 60.] [DOI: 10.3321/j.issn:1001-9014.2007.01.011]
- [10] Du X F, Li C H, Li J. Contour detection based on compound receptive field [J]. Journal of Electronics & Information Technology, 2009, 31(7): 1630-1634. [DOI: 10.1109/CISP.2012.6469689]
- [11] Yang K F, Gao S B, Guo C F, et al. Boundary detection using double-opponency and spatial sparseness constraint [J]. IEEE Transactions on Image Processing, 2015, 24(8): 2565-2578.

- [DOI: 10.1109/TIP.2015.2425538]
- [12] Lansky P, Sacerdote L, Zucca C. The Gamma renewal process as an output of the diffusion leaky integrate-and-fire neuronal model [J]. *Biological Cybernetics*, 2016, 110 (2-3): 193-200. [DOI: 10.1007/s00422-016-0690-x]
- [13] Michael G, Richard B, George R. *Cognitive Neuroscience the Biology of the Mind* [M]. 5th ed. Beijing: China Light Industry Press, 2010: 27-35.
- [14] Li G, Zhu R, Chai H L. A contour detector with improved corner detection [J]. *Multimedia Tools and Applications*, 2017, 76(4): 5965-5984. [DOI: 10.1007/s11042-015-2809-9]
- [15] Pont-Tuset J, Arbelúez P, Barron J T, et al. Multiscale combinatorial grouping for image segmentation and object proposal generation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(1): 128-140. [DOI: 10.1109/TPAMI.2016.2537320]
- [16] Yang Y, Tong S, Huang S, et al. Log-Gabor energy based multimodal medical image fusion in NSCT domain [J]. *Computational and Mathematical Methods in Medicine*, 2014, 2014: #835481. [DOI: 10.1155/2014/835481]
- [17] Venkataramani S, Taylor W R. Orientation selectivity in rabbit retinal ganglion cells is mediated by presynaptic inhibition [J]. *The Journal of Neuroscience*, 2010, 30(46): 15664-15676. [DOI: 10.1523/JNEUROSCI.2081-10.2010]
- [18] Yang K F. Non classical receptive field model based on multiple visual features and its application [D]. Chengdu: University of Electronic Science and Technology of China, 2012. [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.]
- [19] Azzopardi G, Petkov N. A CORF computational model of a simple cell that relies on LGN input outperforms the Gabor function model [J]. *Biological Cybernetics*, 2012, 106(3): 177-189. [DOI: 10.1007/s00422-012-0486-6]