

# 基于视通路多感受野朝向性关联的轮廓检测方法

李康群 范影乐\* 甘海涛 武 薇

(杭州电子科技大学模式识别与图像处理实验室, 杭州 310018)

**摘要:** 基于视通路中各层次的感受野特性以及层次间感受野的关联特性, 提出一种图像轮廓检测新方法。根据神经元的放射状树突接收特性, 构建检测空间差异性信息的视网膜神经网络, 对轮廓信息进行前级编码; 建立 LGN 细胞的非经典感受野调节机制, 结合感受野的层次变换特性, 对前级编码结果进行全局调节; 提出简单细胞的多感受野朝向性关联模型, 模拟初级视皮层简单细胞的方向选择特性; 融合多个方向上的轮廓响应, 经过非极大值抑制和阈值处理, 得到轮廓检测结果。以 RuG 图库 40 幅图片为实验对象, 检测结果与基准轮廓图的平均  $P$  指标为 0.43, 结果显示该方法能够有效凸显主体轮廓, 强化轮廓与纹理边缘的区分度。利用多感受野的层次关联特性, 实现图像轮廓信息的编码与检测应用, 为研究高级视皮层的图像理解和视觉认知提供新的思路。

**关键词:** 轮廓检测; 多感受野; 前级编码; 全局调节; 朝向性关联

**中图分类号:** R318      **文献标志码:** A      **文章编号:** 0258-8021(2017) 01-0001-11

## A Contour Detection Method Based on Correlation of Orientation for Multiple Receptive Fields in Visual Pathway

Li Kangqun Fan Yingle\* Gan Haitao Wu Wei

(Laboratory of Pattern Recognition and Image Processing, Hangzhou Dianzi University, Hangzhou 310018)

**Abstract:** Based on characteristics of receptive fields on different layers of human visual path and their correlation between every adjacent layer, a new method of image contour detection was proposed in this paper. Using the receptive feature of neurons with radial dendrite, a retina neuron network that has the capability to detect spatial difference was constructed for pre-coding of contour information. Taking the variety between adjacent layers into account, the adjustment mechanism for non-classical receptive field of LGN neurons was built to make a global adjustment on the pre-coding result. Next, a model of simple cell on primary visual cortex named orientation correlation of multiple receptive fields was proposed to realize the orientation selection of simple cells. At last, the fused contour response of all orientations was processed with non-maximum suppression and hysteresis thresholding to acquire the contour detecting result. In this paper, 40 images in the RuG library were taken for processing, and the mean value of measure  $P$  between detecting result and ground truth was 0.43, indicating that method in this paper could highlight principal contour and enhance the discrimination between contour and textured region. In this study, the correlate characteristics of multiple receptive fields were utilized to code and detect image contour, providing a new idea for the image processing and visual cognition of high-level visual cortex.

**Key words:** contour detection; multiple receptive fields; pre-coding; global adjustment; orientative correlation

### 引言

轮廓是图像像素值发生跃变的地方, 不同于背

景纹理边缘的无序状态, 主体目标的轮廓表现出一定的方向性和连续性, 包括主体目标与背景、平面与平面的边界。轮廓检测是机器视觉和图像分析

doi: 10.3969/j.issn.0258-8021.2017.01.001

收稿日期: 2016-06-08, 录用日期: 2016-07-11

基金项目: 国家自然科学基金(61501154)

\* 通信作者 (Corresponding author), E-mail: fan@hdu.edu.cn

的基础内容,图像中主体目标轮廓的准确检测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<sup>[1]</sup>。

包括 Canny、Sobel 等算子在内的传统边缘轮廓检测方法,一般是利用数学模型来刻画局部邻域像素在特定方向上的灰度变化<sup>[2-3]</sup>,具有简洁快速的特点,但并不能有效区分主体轮廓和背景纹理边缘。尽管后续又出现了多尺度信息<sup>[4-5]</sup>、多局部特征<sup>[6]</sup>、上下文全局分析<sup>[7]</sup>等检测方法,一定程度上改善了轮廓检测效果,但是与人类视觉系统在复杂环境下快速准确捕捉主体目标的能力相比,这些方法在生物智能性和计算高效性上均有不足。因此,基于人类视觉感知特性,Grigorescu 等利用二维 Gabor 能量模型,模拟初级视皮层经典感受野的方向选择特性,并且依靠非经典感受野的抑制特性进行轮廓检测,有效减少了背景纹理的影响<sup>[8]</sup>。在此基础上,又出现了许多改进的方法:桑农等提出了蝶形域抑制方法,在邻域神经元的抑制作用上考虑得更加详尽,较好地去除了背景纹理产生的边缘<sup>[9]</sup>;杜晓凤等提出了基于复合感受野的轮廓检测方法,在减少背景纹理干扰的同时,保护了较弱的轮廓<sup>[10]</sup>;Yang 等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<sup>[11]</sup>。但是,上述方法大都忽略了视觉神经系统中的电生理活动,即数以千亿计神经元电活动与视觉认知的内在关联性,也没有考虑视觉信息在视觉通路加工和处理过程中的多感受野特性、感受野的层

次关联特性,以及这些特性所依赖的生理学结构基础,所以轮廓检测结果与主观的视觉感知存在着一定的差异。

本研究考虑了视通路中各层次的感受野尺度特性以及层次间感受野的相互关联性,模拟视通路中视网膜神经元的脉冲编码过程,以及外侧膝状体(lateral geniculate nucleus, LGN)细胞非经典感受野的全局调节特性,同时有朝向性地将多个 LGN 细胞感受野进行关联,用于模拟初级视皮层(V1)简单细胞的方向选择特性,提出了一种基于视通路多感受野朝向性关联的轮廓检测方法,充分发挥上述视觉特性在轮廓检测中的重要作用。

## 1 材料和方法

### 1.1 基本原理

#### 1.1.1 视网膜神经网络的前级编码

视网膜中感光细胞接受外界刺激,触发下游神经元产生动作电位,使刺激信息以脉冲发放的形式传递至后级神经元以及视觉皮层进行处理,直至形成视觉感知或识别。在神经信息编码过程中,神经元的树突接收刺激并将冲动传入细胞体。树突具有多种多样的空间表现形式,是视网膜对外界视觉信息进行精细汇聚的结构基础<sup>[12]</sup>。

如图 1 所示,结合神经元的放射状树突结构与感受野特性,将感受野内所有树突接收的外界激励综合作为模型输入,同时用包含轮廓强度信息的脉冲发放频率对外界视觉激励进行编码。

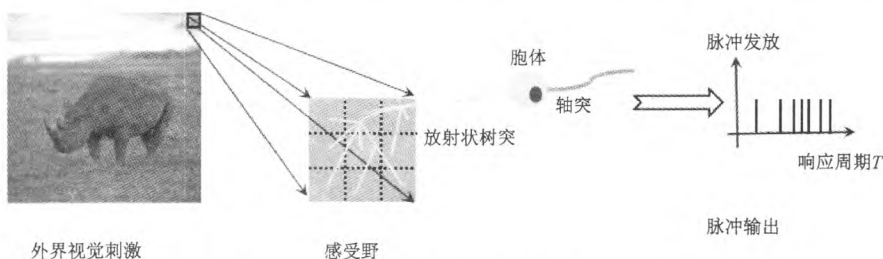


图 1 神经元利用放射状树突接收刺激并产生脉冲输出

Fig. 1 Neurons receive stimuli with radial dendrite and produce spikes

本研究从神经元模型的生物合理性与计算便捷性出发,选取仿真性能相对优越的漏积分放电(leaky integrate-and-fire, LIF)<sup>[13]</sup>模型来构建神经网络,对视觉输入的差异性进行检测,并且引入神经元放电过程中的绝对不应期,改进后的 LIF 模型如下:

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

式中: $x, y$  分别是视觉刺激的横纵坐标; $v$  是神经元膜电压; $c_m$  是膜电容; $g_l$  是漏电导; $I_{xy}$  是与视觉刺

激 $(x, y)$ 对应的神经元树突激励电流,即输入图像像素值的归一化结果;RF表示神经元的感受野; $\omega_{xy,\mu}$ 是与视觉刺激 $(x, y)$ 相对应的神经元树突接收权值,下标 $\mu$ 表示感受野的极性,取值为on或off,前者表示兴奋型感受野,后者表示抑制型感受野; $v_{th}$ 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;max表示取较大值运算,实现对激励电流的非负处理。

当 $v > v_{th}$ 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。

视网膜神经元对外界刺激的前级编码能够充分反映空间的差异性信息<sup>[14]</sup>。本研究在神经网络中引入了放射状树突的接收权值,同时为了编码外界输入的空间差异性信息,采用一种基于空间距离变化的双高斯差函数,对接收权值的变化规律进行模拟。感受野内的树突接收权值随着距离增大逐步减小,之后随着距离进一步增大出现负效应(接收权值为负数),最终因距离过大而趋向于零,

具体如下:

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

式中, $\sigma$ 为双高斯差函数的标准差。

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

$$\omega_{xy,off} = -\omega_{xy,on} \quad (3)$$

在本研究中,统计了在一个响应周期 $T$ 内每个神经元的脉冲发放次数,并将脉冲发放频率映射为灰度图像,从而得到视网膜神经网络对图像差异性信息的前级编码结果,如图2所示。显然,视网膜神经元只能对图像的差异性信息进行粗略检测,并不能有效区分主体轮廓和背景纹理边缘,但这一过程将为视通路后续层次的处理提供原始数据。

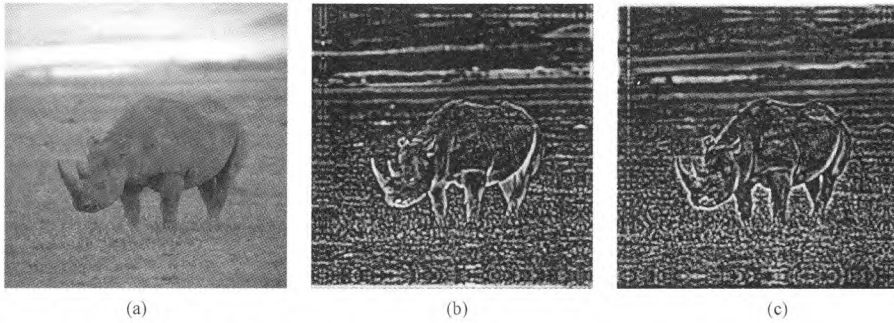


图2 视网膜神经网络的编码结果。(a) 原灰度图;(b) on型编码结果;(c) off型编码结果

Fig.2 Coding result of retina neuron network. (a) The original gray image; (b) The coding result of on-type; (c) the coding result of off-type

### 1.1.2 LGN 神经元感受野的调节机制

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

LGN细胞的感受野与视网膜神经元的感受野相比,结构相似但却具有更大的空间范围,能够综合处理来自双眼的信息<sup>[15]</sup>。不同于传统的感受野模型,仅仅关注图像中以像素点为处理单元的局部细节。在视网膜到LGN的信息投影过程中,由于感受野呈现由局部到全局的“点-块”变换特性,因此提出了以图像块为处理单元的LGN细胞感受野模型。将视网膜神经元编码结果做高斯平滑处理后,在空间上以图像块为基本单元做全局分解;以一个

目标点 $(x, y)$ 为中心构建局部窗口,计算窗口覆盖区域的平均值,选取覆盖区域中与平均值相差最大的像素点作为该区域的全局特征点 $WI_\mu(x, y)$ ,并将其作为LGN细胞的输入。

首先,计算窗口覆盖区域的平均值为

$$ave(x, y) = \frac{1}{ws^2} \sum_{1 \leq i, j \leq ws} GC_\mu((x-1) \cdot step + i, (y-1) \cdot step + j) \quad (4)$$

其次,选取与平均值相差最大的像素点作为全局特征点,有

$$WI_\mu(x, y) = GC_\mu((x-1) \cdot step + p, (y-1) \cdot step + q) - \left| GC_\mu((x-1) \cdot step + p, (y-1) \cdot step + q) - ave_\mu(x, y) \right| =$$

$$\max \left( \left| GC_{\mu}((x-1) \cdot \text{step} + i, (y-1) \cdot \text{step} + j) - \text{ave}_{\mu}(x, y) \right| \right) \quad (\forall 1 \leq i, j \leq \text{ws}) \quad (6)$$

式(4)~(6)中:ws是局部方窗的窗长;step为移动步长,设置为与窗长相等;|·|表示取绝对值运算;max表示取最大值运算;(p,q)表示局部方窗覆盖区域内与平均值相差最大的像素点所对应的坐标索引;GC<sub>μ</sub>(x,y)是高斯平滑后的视网膜编码结果,下标μ与树突接收权值ω<sub>xy,μ</sub>的下标相对应,表示感受野的极性。

相关研究指出,在经典拮抗型感受野的大外周存在一块特殊的区域,即非经典感受野(non-classical receptive field, nCRF),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<sup>[16]</sup>:对全局特征图进行nCRF调节,不仅可以有效抑制图像中的背景纹理边缘,同时在一定程度上能够增强主体轮廓。具体而言,在亮度差异性较大的主体轮廓区域,采取增强性调节;在亮度差异性较小的背景纹理边缘区域,则是借助全局特征图的空间尺度变化,使背景纹理边缘得到有效剔除。本研究对传统的三高斯差函数进行微调,引入了局部半波整流,有

$$\text{TDog}(x, y) = \text{gau}_{\text{sur}}(x, y) - \max(\text{gau}_{\text{inh}}(x, y) - \text{gau}_{\text{ext}}(x, y), 0) \quad (7)$$

式中:max表示取较大值运算,用于实现半波整流;gau<sub>sur</sub>(x,y)、gau<sub>inh</sub>(x,y)、gau<sub>exc</sub>(x,y)是3个高斯函数,分别表示LGN细胞非经典感受野的大外周区、

周边抑制区与中心兴奋区。

周边抑制区的高斯函数标准差通常为中心兴奋区的4倍,且大外周区高斯函数标准差与中心兴奋区相同<sup>[17]</sup>。以微调后的三高斯差函数TDog(x,y)为度量因子,计算全局特征图WI<sub>μ</sub>(x,y)中的亮度差异性LD<sub>μ</sub>(x,y),有

$$LD_{\mu}(x, y) = \max \left( \sum_{(p, q)} WI_{\mu}(x + p, y + q) \cdot \text{TDog}(p, q), 0 \right) \quad (8)$$

式中,(p,q)表示像素点的坐标增量。

进一步将LD<sub>μ</sub>(x,y)的负指数归一化结果作为非经典感受野的调节量nCRF<sub>ajμ</sub>(x,y),把调节量与全局特征图按权值相融合,再通过双线性插值方法将融合结果调整为原输入图像的大小,将其作为LGN细胞的输出LGN<sub>μ</sub>(x,y),具体如下:

$$\text{nCRF}_{\text{aj}_{\mu}}(x, y) = -\exp \left( -\frac{LD_{\mu}(x, y)^2}{2\sigma^2} \right) \quad (9)$$

$$\text{LGN}_{\mu}(x, y) = \text{resize}(WI_{\mu}(x, y) + \alpha \cdot \text{nCRF}_{\text{aj}_{\mu}}(x, y)) \quad (10)$$

式中:δ是负指数化的强度参数,设置为0.2;resize表示双线性插值放大运算;α是融合权值,反映了全局调节的作用强度。

图3给出了上述LGN细胞感受野调节过程中的on型结果,而off型结果可通过类似的调节过程得到。从部分on型的调节结果中可以看出,LGN非经典感受野的全局调节能够在一定程度上抑制背景纹理,同时对视网膜前级编码结果中的主体轮廓进行增强。

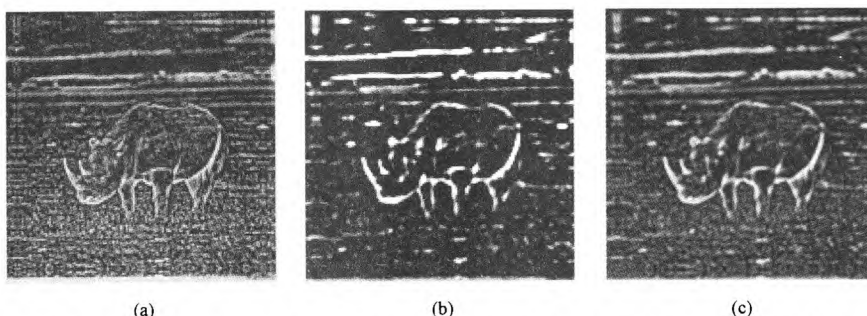


图3 LGN细胞感受野调节过程中的部分结果。(a) on型编码结果的全局特征;(b) on型全局特征的nCRF调节量;(c)经过双线性插值放大的on型调节结果,即LGN细胞的on型输出

Fig. 3 Some results of adjustment with receptive field of LGN cells. (a) Global feature map of on-type coding result; (b) Adjustment value of on-type global feature map; (c) On-type adjustment result after bilinear interpolation zooming, that is the on-type output of LGN cells

### 1.1.3 V1神经元的多感受野朝向性关联模型

V1简单细胞的方向选择特性是准确捕获目标轮廓信息的关键因素,因此在图像处理领域的相关

研究中,大都选择具备方向选择特性的二维Gabor能量函数<sup>[18]</sup>作为初级视皮层简单细胞的计算模型。尽管Gabor函数看上去十分简洁有效,但实际上用



它对简单细胞建模只是一种粗略的黑箱模拟,并没有考虑简单细胞方向选择特性的形成机制。

上面提到,视网膜的视觉信息通过膝状体-皮质投射为皮质提供输入,可见 V1 简单细胞的感受野特性与上一层 LGN 细胞的感受野有着千丝万缕的联系。Hubel 和 Wiesel 认为,初级视皮层的简单细胞可以由多个相邻 LGN 细胞同心圆式感受野的输出相连而成,而简单细胞的方向选择特性与多个

LGN 细胞感受野的关联方式密切相关<sup>[19]</sup>。

本研究以外侧膝状体和初级视皮层的层次关联特性为切入点,综合考虑简单细胞方向选择特性的生理学形成机制以及计算模型的简洁高效性,构建如图 4 所示的 V1 简单细胞多感受野朝向性关联模型。其中,绿色系图标代表 LGN 细胞的 on 型感受野,橘色系图标代表 LGN 细胞的 off 型感受野,蓝色系图标代表 V1 简单细胞的感受野。

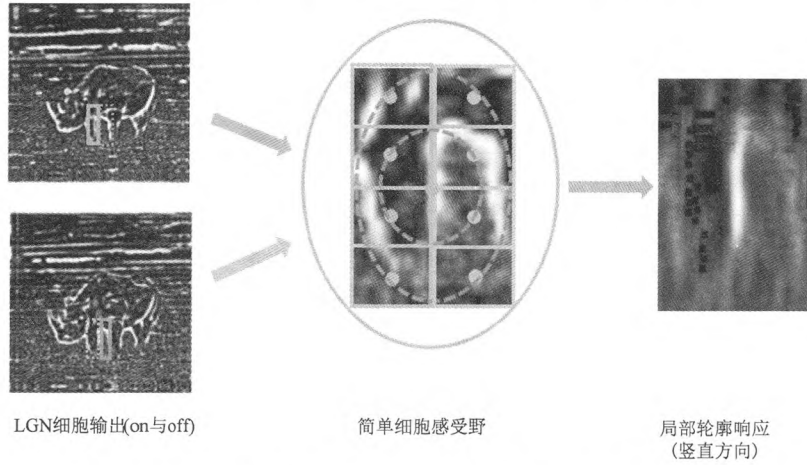


图 4 多感受野朝向性关联模型的示意图

Fig. 4 Sketch map of orientative correlation of multiple receptive fields model

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

用一个三元素的元胞集合  $\text{RecRFSet}$  表示多个矩形感受野的朝向性分布,具体如下:

$$\text{RecRFSet} = \{(\mu_i, r_i, \varphi_i) \mid i = 1, \dots, 8\} \quad (11)$$

式中,  $\mu_i$  为第  $i$  个矩形感受野的极性,  $r_i$  为第  $i$  个矩形感受野中心位置到 V1 简单细胞感受野中心位置的距离,  $\varphi_i$  为第  $i$  个矩形感受野中心位置和 V1 简单细胞感受野中心位置的连线相对于水平方向的夹角。

显然,极坐标  $(r_i, \varphi_i)$  确定了  $2N$  个矩形感受野的相对位置,从而实现多个感受野的朝向性关联。朝向型关联之后,矩形感受野输出为

$$\text{RecRF}_{\mu_i, r_i, \varphi_i}(x, y) = \text{LGN}_{\mu_i}(x - \Delta x_i, y - \Delta y_i) \quad (12)$$

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

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

$$os_\theta(x, y) = \left( \prod_{i=1}^{2N} \text{RecRF}_{\mu_i, r_i, \varphi_i}(x, y) \right)^{1/(2N)} \quad (13)$$

式中,  $\theta$  表示 V1 简单细胞的偏好方向,由集合  $\text{RecRFSet}$  决定。

在初级视皮层中,不同偏好方向的简单细胞的感受野结构相似,而输入来源不同<sup>[12]</sup>。针对不同偏好方向的简单细胞的多感受野输入,采用不同朝向的关联方式,具体如下:

$$os_{\theta_k}(x, y) = \left( \prod_{i=1}^{2N} \text{RecRF}_{\mu_i, r_i, \varphi_i + \Delta \varphi_k}(x, y) \right)^{1/(2N)} \quad (14)$$

式中:  $\theta_k$  表示其他偏好方向,  $k = 1, \dots, 24$  表示不同偏好方向的索引;  $\Delta \varphi_k$  为朝向性关联时的角度增量。

调整矩形感受野集合  $\text{RecRFSet}$  中每个元胞的

极坐标方向  $\varphi_i$ , 即可实现其他朝向的关联方式, 从而进一步获取其他 V1 简单细胞在其偏好方向上的

方向选择结果。图 5 给出了多感受野朝向性关联模型的部分方向选择结果。

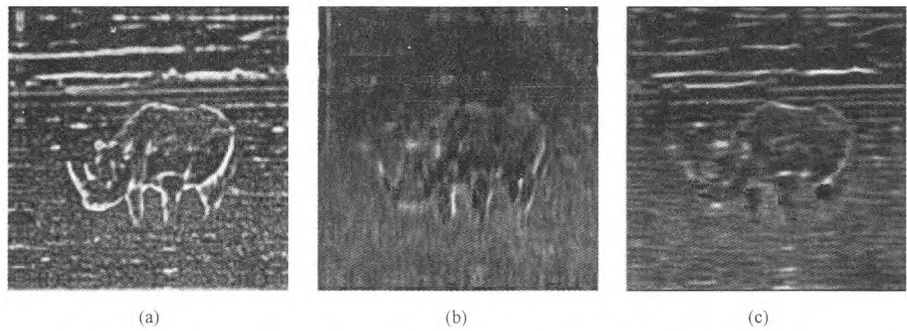


图 5 多感受野朝向性关联模型的部分方向选择结果。(a) LGN 细胞 on 型输出; (b) 偏好方向为竖直方向的简单细胞方向选择结果; (c) 偏好方向为水平方向的简单细胞方向选择结果

Fig.5 some orientation selective results using orientative correlation of multiple receptive fields model. (a) The on-type output of LGN cells; (b) Orientation selective result of simple cells with vertical orientation preferred; (c) Orientation selective result of simple cells with horizontal orientation preferred

1.2 算法流程

按早期视觉通路的信息处理流程,对轮廓检测过程做如下叙述:

- 1) 根据感受野内神经元的放射状树突接收特性,构建检测空间差异性信息的视网膜神经网络,对轮廓信息进行前级编码;
- 2) 基于感受野的层次变换特性,引入 LGN 细胞的非经典感受野调节机制,对前级编码结果中的主体轮廓信息和背景纹理边缘进行全局调节;
- 3) 根据 LGN 与初级视皮层的层次关联特性,以及初级视皮层简单细胞的方向选择特性形成机制,构建简单细胞的多感受野朝向性关联模型,实现对特定方向的轮廓检测;
- 4) 综合全部方向上的轮廓响应,经过非极大值抑制和阈值处理,得到轮廓二值图。

算法执行流程如图 6 所示(见下页)。

具体检测流程如下:

步骤 1: 根据输入灰度图像  $IG(x,y)$  ( $x = 1, 2, \dots, m; y = 1, 2, \dots, n$ ) 大小构建相同尺寸的视网膜神经网络  $GN(x,y)$ , 神经元接收感受野  $RF(7 \times 7)$  内的外界综合刺激,按照式(2)、(3)计算树突接收权值  $\omega_{xy,\mu}$  的大小。

步骤 2: 网络中的每个神经元都采用式(1)所示的 LIF 神经元改进模型,将灰度图像的像素归一化为神经元的激励电流,统计一个响应周期内各神经元的脉冲发放频率,作为视网膜神经元对轮廓信息的前级编码结果。

步骤 3: 视网膜神经元的编码结果经过高斯平

滑得到  $GC_{\mu}(x,y)$ , 按式(4)~(6)进行全局分解,得到全局特征图  $WI_{\mu}(x,y)$ , 将其作为 LGN 细胞的输入。

步骤 4: 按照式(7)~(10),模拟 LGN 神经元非经典感受野对全局亮度的调节机制,对全局特征图  $WI_{\mu}(x,y)$  的主体轮廓和背景纹理进行调节,得到 LGN 细胞的输出  $LGN_{\mu}(x,y)$ 。

步骤 5: 根据式(11)提供的矩形感受野空间位置信息以及式(12)给定的矩形感受野输出值,对各个 LGN 细胞矩形感受野进行朝向性关联,最后根据式(13)计算各个矩形感受野的几何平均值,作为初级视皮层简单细胞在其偏好方向上的方向选择结果。另外,利用式(14)可以得到其他偏好方向的简单细胞的方向选择结果。

步骤 6: 融合各偏好方向简单细胞的方向选择结果,计算初级视皮层轮廓响应  $RV1(x,y;\sigma)$  以及

与灰度图像尺寸相同的最优方向矩阵  $BO(x,y)$ , 有

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

(15)

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

(16)

步骤 7: 对步骤 6 的轮廓响应进行非极大值抑制处理和滞后阈值处理,得到轮廓响应的二值图  $BC(x,y)$ , 具体计算如下:

$$thin(x,y) = \begin{cases} RV1(x,y) & (p_2 \geq RV1(x,y) \geq p_1) \\ 0 & (other) \end{cases}$$

(17)

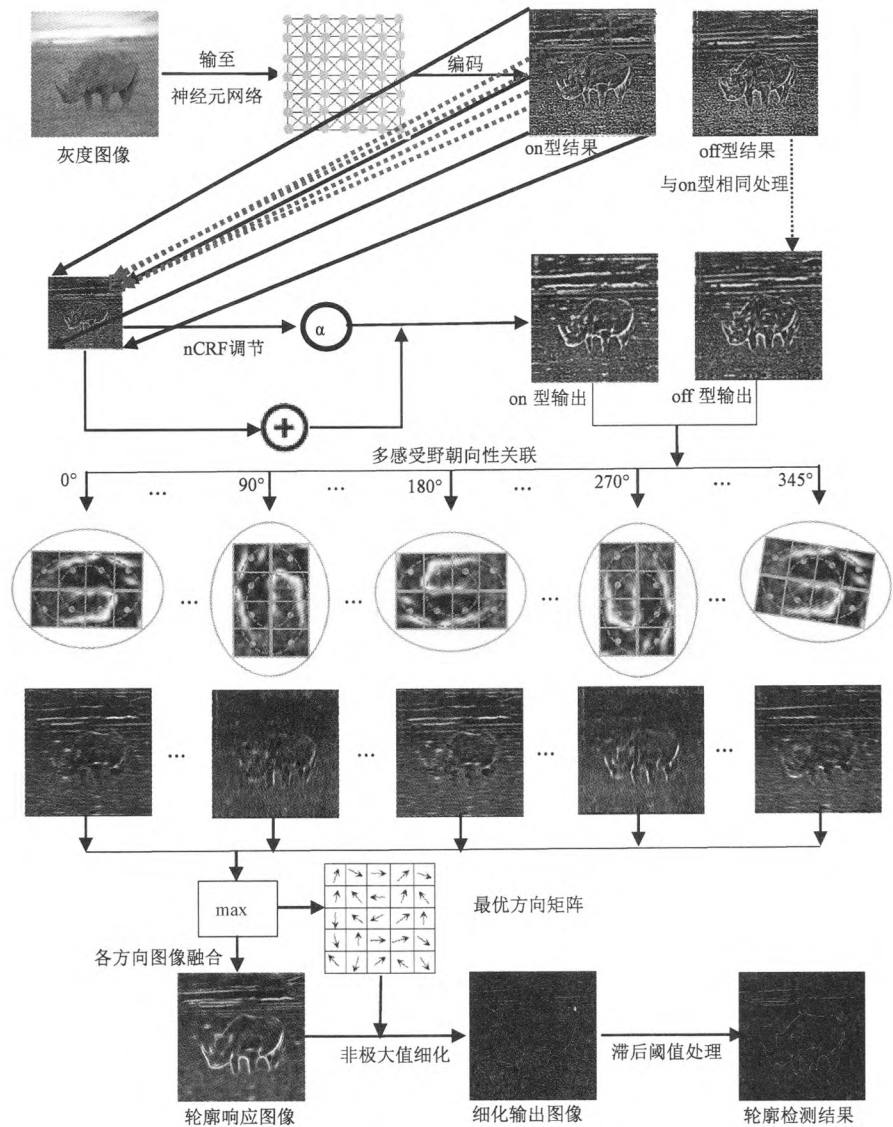


图 6 轮廓检测算法流程  
Fig. 6 procedure of contour detect algorithm

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

(18)

式中:  $(p_1,p_2) = \text{intpol}(BO(x,y),RV1(x,y))$  表示  $RV1(x,y)$  在其 8 邻域中根据  $BO(x,y)$  进行梯度插值,  $p_1、p_2$  为  $RV1(x,y)$  在其最优方向  $BO(x,y)$  上的插值结果;  $t$  为滞后阈值,  $\text{edge}$  表示已经确定的轮廓点,  $\rightarrow$  表示 8 邻域相连接。

1.3 算法验证

自然场景图像的轮廓检测是利用图像方法分析和识别主体目标的基础内容,但自然场景图像由于轮廓纹理相似性、噪声干扰等特征,对其进行轮

廓检测有较高的难度。为验证本方法对自然场景图像的轮廓检测具备有效性,选取 RuG 图库中的 40 幅自然场景图进行轮廓检测实验,图库中每张图片都有一张手工制作的基准轮廓图,用于计算轮廓检测的定量评价指标,选取图库中部分有代表性的图片作为实验对象。

首先,选取直接模拟初级视皮层轮廓响应的数学模型作为对比方法,该模型就是与 Gabor 能量模型类似的二维高斯导函数模型(DG),体现本方法中 V1 多感受野朝向关联模型的生物合理性;其次,选取文献[20]提出的仅考虑了 LGN 细胞排列组合的轮廓检测模型(CORF),证明本方法中视网膜前级编码和 LGN 感受野调节的有用性;最后,考虑纹理抑制特性,选择文献[11]中去除颜色双拮抗机制

并适用于单通道灰度图像的空间稀疏限制方法 (SSC) 作为对比方法。实验中, 选取 10 个不同的 LGN 非经典感受野调节强度, 对应参数为  $\alpha = [0.1: 0.1: 1.0]$ 。另外, 为了与图像库提供的二值标准轮廓图进行比较, 需要对提取的轮廓响应图进行非极大值抑制处理和滞后阈值处理<sup>[3,8]</sup>。在滞后阈值处理中, 取上限阈值  $t = [0.1: 0.05: 0.8]$ , 因此对于本方法和 SSC 方法可得到 150 组测试参数, 而由于 DG、CORF 方法并不涉及全局系数  $\alpha$ , 故只进行 15 组参数测试。

为客观地对检测结果进行评价, 采用文献[8]中的定量指标计算公式对本方法和对比方法做定量分析。显然, 最终得到轮廓二值图像 BC 的轮廓点坐标一般不可能与基准轮廓图 GT 中的轮廓点坐标完全吻合, 相反如果检测结果中的各个轮廓点出现在基准轮廓图对应像素点的某个邻域内, 则会显得更加合理。因此, 本实验设置了与轮廓检测对比算法相同的邻域参数: 若检测到的轮廓点出现在基准轮廓点的  $5 \times 5$  邻域内, 就认为该轮廓点检测正确。定义  $E_{BC}$  和  $E_{GT}$  分别为算法检测到的轮廓像素集和基准轮廓像素集, 则算法正确检测到的像素集为  $E = E_{BC} \cap (E_{GT} \oplus T)$  ( $\oplus$  为膨胀操作,  $T$  为  $5 \times 5$  的结构单元); 错误检测到的像素集为  $E_{FP} = E_{BC} - E$ , 漏检的像素集为  $E_{FN} = E_{GT} - (E_{GT} \cap (E_{BC} \oplus T))$ 。算法性能评价指标误检率  $e_{FP}$ 、漏检率  $e_{FN}$  和整体性能指标  $P$  可通过如下公式计算得到, 有

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

(19)

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

(20)

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

(21)

式中,  $\text{card}(X)$  代表求集合  $X$  中元素的个数。

显然, 误检率漏检  $e_{FP}$  越小, 表明检测结果具有更少的背景纹理; 漏检率漏检  $e_{FN}$  越小, 说明检测结果保留了更多的主体轮廓。因此, 当误检率和漏检率趋向于 0 时,  $P$  趋向于 1, 即  $P$  越大则表明算法检测效果更优, 检测到的轮廓图更接近于基准轮廓图。

2 结果

图 7 给出了部分图片的轮廓检测对比结果(见下页): 第 1 行为测试对象, 即自然场景图; 第 2 行为图库提供的基准轮廓图; 第 3 行为 DG 方法检测结果; 第 4 行为 CORF 方法检测结果; 第 5 行为 SSC 方法检测结果; 第 6 行为本方法检测结果。从结果中

可以看出, DG 方法直接用数学公式模拟初级视皮层的轮廓响应, 然后利用阈值化来保留一定强度的轮廓, 可检测出了绝大多数的轮廓细节, 但阈值处理并不能有效抑制背景纹理, 导致结果中出现较多的伪轮廓; CORF 方法考虑了 LGN 感受野的空间组合特性, 能一定程度上减少背景纹理干扰, 但由于缺乏视通路其他层次的编码与调节, 容易因为过抑制和阈值处理不当而造成真实轮廓丢失; SSC 方法由于考虑了轮廓的显著性与稀疏特性, 同时引入了纹理抑制, 在保留主体轮廓和抑制背景纹理上取得了不错的平衡结果, 但在主体轮廓的周边区域对纹理边缘的抑制作用不够明显, 形成较多的毛刺。综合而言, 本方法保留了尽可能多的主体轮廓信息, 同时去除了大部分的背景纹理边缘, 在图像纯净度上要优于 SSC 方法。

按上述定量方法, 计算图 7 中各算法最佳结果的定量指标, 如表 1 所示。

表 1 图 7 所示图像对应的参数设置及性能指标

Tab. 1 Parameter set and performance index of detect results in Fig. 7

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

不难看出, 表中计算结果基本与上文的定性分析结果相吻合。DG 方法的漏检率很小, 说明目标轮廓非常完整, 但误检率  $e_{FP}$  很大, 整体效果  $P$  值不太理想; 例如 Rino 的结果, 漏检率仅为 0.1, 而误检率却达到了 1.79, 显然较小的漏检率难以弥补极高



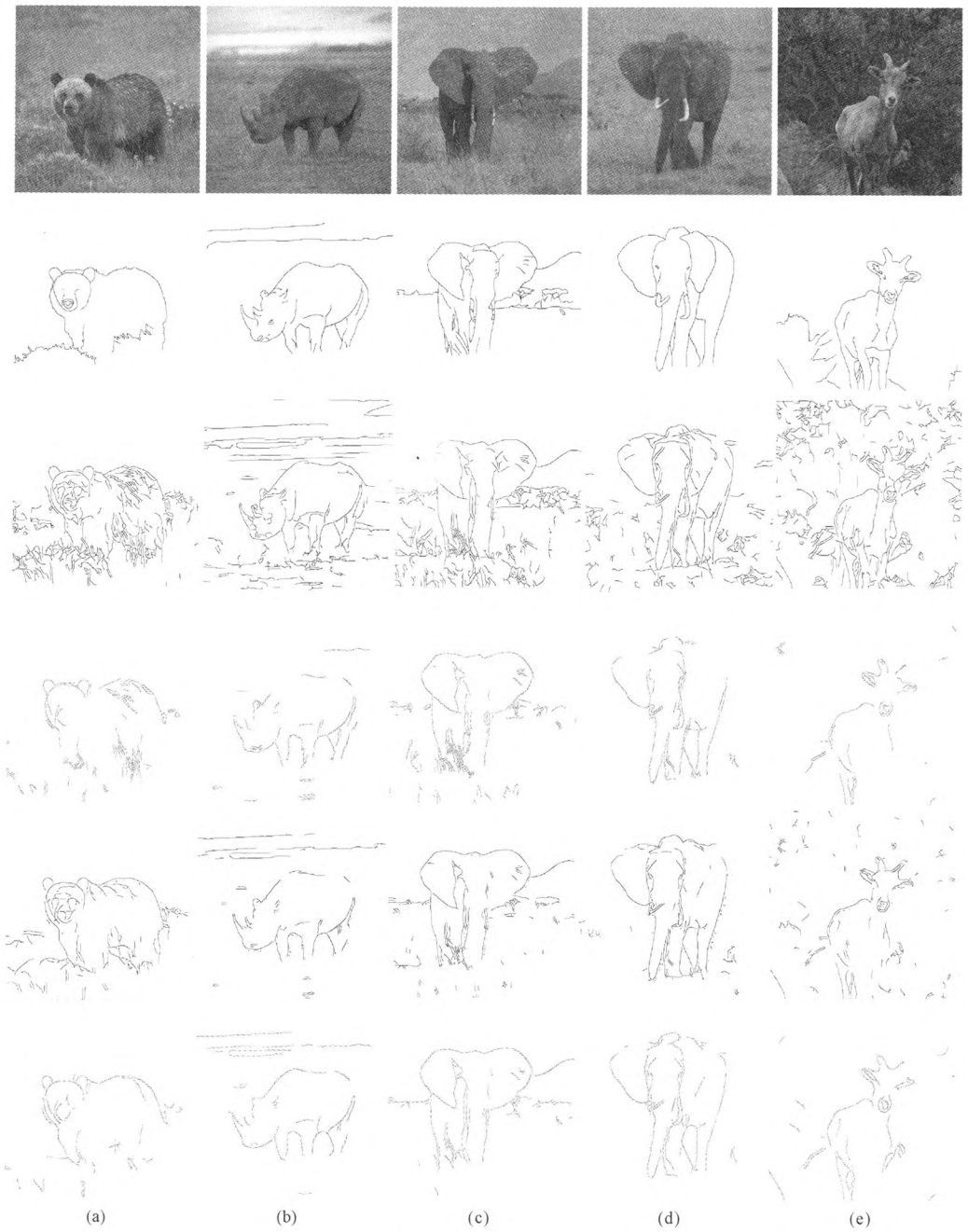


图 7 RuG 图库中的典型图片及其轮廓检测结果(第 1 行为用于测试的自然场景,具有轮廓与纹理相似的基本特征;第 2 行为图库提供的基准轮廓;第 3 行为 DG 方法的检测结果;第 4 行为 CORF 方法的检测结果;第 5 行为 SSC 方法的检测结果;第 6 行为本方法的检测结果)。(a) Bear5 轮廓细节丰富;(b) Rino 目标与背景相似度高;(c) Elephant2 远景丰富;(d) Elephant3 背景单一;(e) Goat3 总体亮度低

Fig.7 Some typical images in RuG image library and their contour detect results ( The first row is natural images for testing, whose contour is basically similar with texture; The second row is the ground truth; The third row is result of DG; The fourth row is result of CORF; The fifth row is result of SSC; The sixth row is result of proposed method ). ( 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

的误检率所带来的负面影响,导致整体效果不佳。  
CORF 方法则显著减小了误检率,但漏检率控制得

不太理想,造成主体轮廓丢失,整体效果  $P$  值的改善程度有限,从 Rino、Elephant2 等的结果看都表明

了这一点。SSC 方法与本方法在误检率和漏检率上取了不同程度的平衡,例如 Elephant2 的结果,本方法和 SSC 的漏检率分别是 0.32 与 0.24,误检率分别是 0.23 与 0.42,而整体效果  $P$  值保持一致,均为 0.57,总体来说本方法误检率更小,具有较高的图像纯净度,且漏检率也处于一个较优水平。由表 1 中  $P$  值可以看出,本方法的整体检测效果优于对比方法。另外,计算全部 40 幅图片的检测结果与对应基准轮廓图的平均  $P$  指标, DG、CORF、SSC 分别为 0.36、0.40、0.42,而本方法为 0.43,可见本方法在总体上具有更优秀的性能。

### 3 讨论

从以上实验结果不难看出,本方法具有较强的轮廓提取能力,特别是针对主体轮廓具备丰富细节,并且背景纹理是与主体轮廓相似性很强的自然场景图像,从图 7 的定性结果以及表 1 的整体性能指标都可以反推出这一结论。

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

本研究所做的工作及其创新性如下:

1) 提出了一种基于视通路多感受野朝向性关

联的视觉轮廓信息处理新方法。研究了视网膜、外侧膝状体以及初级视皮层的神经元感受野特性,分析了不同层次间感受野的空间尺度变换在视觉信息精细汇聚中的作用,论述了感受野的朝向性关联对方向选择形成机制的生理学解释,突出了生物视觉系统的信息处理能力的应用研究。

2) 实现了一种模拟视觉通路信息加工机制的图像轮廓检测。根据神经元的放射状树突接收特性,构建检测空间差异性信息的视网膜神经元网络,对轮廓信息进行前级编码;建立 LGN 细胞的非经典感受野调节机制,结合感受野的层次变换特性,对前级编码结果进行全局调节;提出简单细胞的多感受野朝向性关联模型,模拟初级视皮层简单细胞的方向选择特性;融合多个方向上的轮廓响应,经过非极大值抑制和阈值处理,得到轮廓检测结果。

3) 为利用视觉通路多个层次的感受野关联性进行图像轮廓检测或者是后续的目标识别与分析提供了新思路。以具备轮廓纹理相似性、噪声干扰等特征的自然场景图像为例,突出视觉通路不同层次感受野及其层次间关联特性在轮廓感知过程中的任务分工,利用视网膜感受野的前级编码、外膝体感受野在层次变换中表现出的空间尺度特性、外膝体感受野与初级视皮层感受野的朝向关联特性,构建视通路多感受野朝向性关联的图像轮廓检测模型,不仅能够有效抑制背景纹理边缘,而且也能在一定程度上增强主体目标轮廓。

### 4 结论

视通路中神经元感受野在视网膜、外侧膝状体以及初级视皮层这 3 个逐级层次上呈现出不同的结构组织特性与刺激响应机制,笔者通过探讨不同层次神经元感受野在轮廓提取过程中的具体任务分工,引入视网膜的轮廓信息脉冲编码、外膝体感受野的全局调节特性以及感受野的层次关联特性,摒弃传统上采用单感受野数学模型模拟视觉机制的做法,提出一种基于视通路多感受野朝向性关联的轮廓检测方法。与 DG、CORF 以及 SSC 方法进行比较分析,结果表明本方法检测到的自然图像轮廓更加完整且连续,同时抑制了大部分的背景纹理边缘,具有较高的图像纯净度,整体上反映了本方法所具备的生物智能性。本研究提出的基于视通路多感受野朝向性关联的轮廓检测方法,为图像轮廓检测技术提供了新的思路。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皮层的图像理解与视觉认知中发挥作用。

参考文献

[ 1 ] Papari G, Petkov N. Edge and line oriented contour detection: State of the art [ J ]. Image and Vision Computing, 2011, 29 (2): 79-103.

[ 2 ] Martin DR, Fowlkes CC, Malik J. Learning to detect natural image boundaries using local brightness, color, and texture cues [ J ]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2004, 26(5): 530-549.

[ 3 ] John C. A computational approach to edge detection [ J ]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, 8(6): 67-69.

[ 4 ] Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis [ J ]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 1998, 20(11): 1254-1259.

[ 5 ] Sun Junxi, Gu Dongbing, Chen Yazhu, et al. A multiscale edge detection algorithm based on wavelet domain vector hidden Markov tree model [ J ]. Pattern Recognition, 2004, 37 (7): 1315-1324.

[ 6 ] Yang Kaifu, Li Chaoyi, Li Yongjie. Multifeature-based surround inhibition improves contour detection in natural images [ J ]. IEEE Transactions on Image Processing, 2014, 23(12): 5020-5032.

[ 7 ] Arbelaez P, Maire M, Fowlkes C, et al. Contour detection and hierarchical image segmentation [ J ]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2011, 33(5): 898-916.

[ 8 ] Grigorescu C, Petkov N, Westenberg M. Contour detection based on nonclassical receptive field inhibition [ J ]. IEEE Transactions on Image Processing, 2003, 12(7): 729-739.

[ 9 ] 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.

[ 10 ] Du Xiaofeng, Li Cuihua, Li Jing. Contour detection based on compound receptive field [ J ]. Journal of Electrics and Information Technology, 2009, 31(7): 1630-1634.

[ 11 ] Yang Kaifu, Gao Shaobing, Guo Cefeng, et al. Boundary detection using double-opponency and spatial sparseness constraint [ J ]. IEEE Transactions on Image Processing, 2015, 24(8): 2565-2578.

[ 12 ] Michael G, Richard B, George R. Cognitive Neuroscience the Biology of the Mind [ M ]. 3rd ed. Beijing: China Light Industry Press, 2010.

[ 13 ] Chouhan A. An analytical study of leaky integrate-and-fire neuron model using MATLAB simulation [ J ]. International Journal of Engineering Research & Technology, 2013, 2 (4): 2242-2245.

[ 14 ] Sowmya V, Rowland T. Orientation selectivity in rabbit retinal ganglion cells is mediated by presynaptic inhibition [ J ]. Journal of Neuroscience, 2010, 30(46): 664-676.

[ 15 ] 罗四维. 视觉感知系统信息处理理论 [ M ]. 北京: 电子工业出版社, 2006.

[ 16 ] Li Chaoyi. Integration fields beyond the classical receptive field: organization and functional properties [ J ]. Physiology, 1996, 11 (4): 181-186.

[ 17 ] Xu Xiangmin, Bonds A, Casagrande V. Modeling receptive-field structure of koniocellular, magnocellular, and parvocellular LGN cells in the owl monkey ( Aotustrivigatus ) [ J ]. Visual Neuroscience, 2002, 19(6): 703-711.

[ 18 ] Yang Yong, Tong Song, Huang Shuying, et al. Log-Gabor energy based multimodal medical image fusion in NSCT domain [ J ]. Computational & Mathematical Methods in Medicine, 2014, 2014(1): 481-493.

[ 19 ] 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.

[ 20 ] 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.