

DOI: 10.16289/j.cnki.1002-0837.2018.03.012

基于主视通路层级响应模型的轮廓检测方法

周 涛, 范影乐, 朱亚萍, 武 薇*

(杭州电子科技大学自动化学院模式识别与图像处理实验室, 浙江杭州 310018)

摘要: 目的 提出一种主视通路信息流层级传递和响应的新模型用于检测图像轮廓的新方法。方法 以 RuG 图库 40 幅图片为实验对象, 利用非下采样轮廓波变换模拟外侧膝状体(lateral geniculate nucleus, LGN)对视觉信息的频域分离作用; 构建 LIF 神经网络模型来表达视觉神经系统中的电生理活动, 通过 CRF 机制整合空间信息; 同时, 利用局部半波整流的高斯差函数来模拟 nCRF 的全局调节机制。有朝向性地将多个 LGN 细胞感受野进行关联。同时构建皮层下视通路来模拟它对于主视通路进行视觉信息处理的协同作用。然后, 经过非极大值抑制和阈值处理, 得到本文轮廓检测结果。最后将本文检测结果与 3 种经典方法(Nonin, SSC, ISO)的检测结果进行对比。结果 本方法的检测结果与基准轮廓图的平均 P 指标为 0.46, 大于经典的 3 种检测方法(P 指标分别为 0.36、0.40、0.42)。结论 本文算法不仅对纹理和背景具有抑制的作用, 而且能有效区分纹理强边缘和主体轮廓, 获得较佳的效果。

关键词: 轮廓检测; 非下采样轮廓波变换; 多感受野; 前级编码; 全局调节; 朝向性关联

中图分类号: R319 文献标识码: A 文章编号: 1002-0837(2018)03-0371-09

Fast Contour Detection Method Based on Hierarchical Response Model of Primary Visual Pathway. Zhou Tao, Fan Yingle, Zhu Yaping, Wu Wei. Space Medicine & Medical Engineering 2018, 31(3): 371-379

Abstract: **Objective** To propose a new model of hierarchical transfer and response of the information flow in the cortical visual pathway to detect the image contour. **Methods** Forty images in the RuG library were selected for processing. The non-subsampled contourlet transform was used to simulate the frequency-domain separation of visual information in lateral geniculate nucleus(LGN). The LIF neural network model was constructed to express the electrophysiological activities in the visual nervous system, and then the spatial information was integrated via CRF mechanism. At the same time, in order to simulate the global adjustment mechanism of the nCRF, the Gaussian difference function of the local half-wave rectification was utilized. After that, multiple LGN cells were associated into the field. Meanwhile, subcortical visual pathways were constructed to simulate the synergistic effect of visual information on the cortical pathways. The non-maximum suppression and threshold processing were used for processing and the contour detection results were obtained. In the end, the results were compared with that of other three classical methods (Nonin, SSC, ISO). **Results** The mean value of P between the detected result and the ground truth was 0.46, which was better than the other three classical methods (P values were 0.36, 0.40, 0.42 respectively). **Conclusion** The method not only has a deterrent effect on texture and background, but also can effectively distinguish the edge of the texture and the main contour, which can get better results.

Key words: contour detection; non-subsampled contourlet transform; multiple receptive fields; pre-coding; global adjustment; orientative correlation

Address correspondence and reprint requests to: Wu Wei. Laboratory of Pattern Recognition and Image Processing, Hangzhou Dianzi University, Hangzhou Zhejiang 310018, China

轮廓是图像目标描述中的低维度关键信息, 对于图像分析或理解具有重要作用, 轮廓检测的准确性将直接影响到后续特征提取过程的性能^[1]。

以 Canny 算子为代表的传统轮廓提取方法主要基于图像像素灰度或颜色值的局部跳变特性, 利用一阶或二阶梯度运算, 同时考虑到轮廓在方向上的连续性, 利用针对局部点的非极大值抑制

收稿日期: 2017-06-02 修回日期: 2018-03-18

基金项目: 国家自然科学基金(61501154)

第一作者: 周 涛, 男, 硕士研究生, 研究方向为视觉神经计算与图像处理。E-mail: arsenal_henry14@126.com

* 通讯作者: 武 薇, 女, 博士, 研究方向为医学图像处理。E-mail: ww@hdu.edu.cn

实现边缘信息的提取。它们更多的是从抽象的数学模型上来刻画轮廓特征,并没有借鉴视觉系统在信息流传递和处理过程中的内在机制。

近年来,随着神经工程实验手段和神经计算仿真能力的提高,模拟生物视觉信息处理的轮廓感知模型被陆续提出。有研究利用主视皮层神经元的非经典感受野(non-classical receptive field, nCRF)特性,能够在复杂背景下有效检测物体轮廓^[2],由于涉及到较大尺寸的大外周区域,因此主体轮廓具有一定的模糊化现象;也有研究利用感受野的侧抑制机制,并同时结合层级注意机制,可以得到多个层级精细程度的图像轮廓^[3]。另外 Tang 等^[4-5]根据 Gestalt 原则,提出了基于共圆约束的空间易化理论,并采用递归兴奋增强和抑制建立动态的轮廓检测调制模型,但存在轮廓片段化严重的问题。上述模型虽然考虑到视觉通路对信息进行加工和处理的机理,但大多采用主视通路和单一的感受野特性,忽视了皮层下视通路以及感受野之间的层级关联特性,并且忽略了视觉神经系统中的电生理活动,所以所获得的结果并不能真正充分反映视觉机制的作用。

本文考虑了视觉信息流传递或处理中视通路的各环节特点。首先,在外侧膝状体(lateral geniculate nucleus, LGN)环节,不仅引入了其对视觉信息的频域分离作用,还研究了外侧膝状体细胞的脉冲编码过程以及非经典感受野全局调节特性。然后,对于大脑纹状皮层(17区),即主视皮层(V1),提出一种有朝向性的将多个 LGN 细胞感受野进行关联的方法,用于模拟 V1 简单细胞的方向选择性。最后提出了一种新的基于视觉机制的自然场景轮廓提取方案,结合皮层下视通路对于主视通路进行视觉信息处理的协同作用,刻画了视觉通路中各个层次的感受野尺度以及层次间感受野的相互关联性。

1 基本原理

1.1 视觉信息流的频域分离

普遍认为 LGN 并不是一个简单的视觉信息中转站,它有着更加高级的功能和作用。外侧膝状体的前 2 层和后 4 层分别对应于视网膜神经节 M 细胞通路和 P 细胞通路的输入,分别负责亮度低频信息和颜色高频信息。有研究证实, LGN 输出

在向主视皮层(V1区)传递时仍然保持着信号的分离^[6],视觉信号的融合要等到视觉信号传递到视皮层才能进行。一般来说,图像信号的低频分量主要是对整幅图像强度的综合度量,描述整体性的均值信息,而高频分量主要是对图像轮廓的度量,描述局部的波动细节。在灰度图像中,亮度变化小的区域主要是低频成分,而亮度变化剧烈的区域(比如物体的边缘)对应于高频成分。考虑到非下采样轮廓波变换(nonsampled contourlet transform, NSCT)不仅具有平移不变性的缺陷,还拥有各向异性特征和捕获图像边缘等具有几何结构特征的能力,同时对图像中的边缘信息能够进行更稀疏的描述^[7]。因此本文利用 NSCT 分解来模拟 LGN 在视觉信息处理中的频域分离作用。

NSCT 交替使用非下采样的金字塔分解(nonsampled pyramid decomposition, NSP)和非下采样方向滤波器组(nonsampled directional filter banks, NSDFB)来处理原始图像,最后得到多方向多尺度的二维分解特征图^[8-9]。实现 NSCT 的结构框架及其频域划分示意图如图 1 所示。由图可知, NSCT 首先采用 NSP 对输入图像进行多尺度分解,得到低频子带图像和高频子带图像;然后再将得到的高频子带图像传递给 NSDFB,实现对该图像的多方向分解;同时将得到的低频子带图像作为新的输入图像,再依次使用 NSP 进行多尺度分解和 NSDFB 进行多方向分解。重复上述步骤,实现对图像多尺度多方向且平移不变的频域分解。

NSP 和 NSDFB 的滤波结构如图 2(a) 和 (b) 所示。其中, NSP 主要实现对图像的多尺度分解,而 NSDFB 主要实现对高频子带图像的多方向分解,它们均为一组双通道的滤波器组。图中 $H_0(z)$ 和 $H_1(z)$, $U_0(z)$ 和 $U_1(z)$ ($-\pi, \pi$) 均为分解滤波器, $G_0(z)$ 和 $G_1(z)$, $V_0(z)$ 和 $V_1(z)$ 则表示合成滤波器。NSP 和 NSDFB 能够完全重构的条件是分解滤波器和合成滤波器满足 Bezout 恒等式,分别对应于式(1)和(2)。

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

$$V_0(z)U_0(z) + V_1(z)U_1(z) = 1 \quad (2)$$

1.2 视觉信息流的脉冲编码

视觉信息流在 LGN 中进行传输和处理时,由

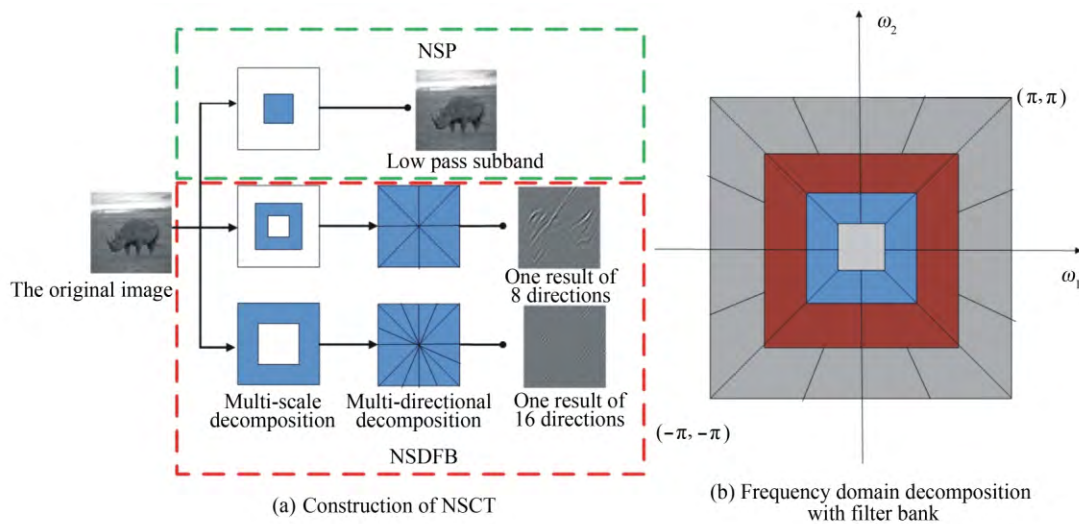


图1 非下采样轮廓波变换示意图
Fig. 1 Nonsubsampled contourlet transforms

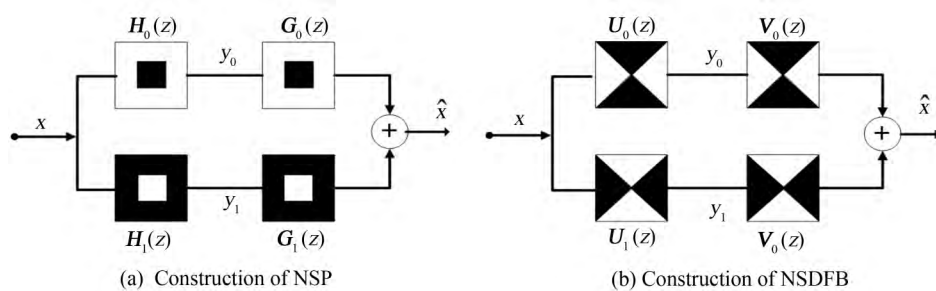


图2 NSP和NSDFB结构图
Fig. 2 Construction of NSP and NSDFB

于LGN与前级相比,在神经元数量级上将显著减少。通常认为在LGN中,视觉系统利用感受野对视觉信息进行空间和时间整合和特征信息的压缩^[10],从而剔除视觉信息中的部分冗余量,实现视觉特征的高效编码。

本文为了充分反映视觉系统中的神经电生理基础,构建LIF模型^[11]来组建神经网络,利用其脉冲发放特性,以脉冲发放频率对视觉特征进行脉冲编码。同时考虑到目标轮廓以及背景纹理是一种空间分布的结构信息^[12]。因此在视觉信息的脉冲编码过程,通过经典感受野(classical receptive field, CRF)机制整合空间信息^[13],对于特定注视点处的脉冲响应,将以之为中心的局部窗口信息作为神经元的突触电流输入,同时以一定的步长移动局部窗,不仅减小了神经网络规模,而且反映了LGN环节对于视觉空间去冗余的处理机制。LIF(Leaky Integrate and Fire neurons)模型如式(3)所示。

$$\begin{cases} v = v_{\text{reset}} & \text{ref} = \text{ref} - 1 \quad (\text{if } \text{ref} \neq 0) \\ c_m \frac{dv}{dt} = -g_1 v + \max\left(\sum_{(x,y) \in \text{CRF}} \omega_{xy}^{\mu} I_{xy} \rho\right) & (\text{if } \text{ref} = 0) \\ v = v_g & \text{ref} = \text{const} \quad (\text{if } v > v_{th}) \end{cases} \quad (3)$$

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_1 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; ω_{xy}^{μ} 是与视觉刺激 (x, y) 相对应的神经元树突接受权值,上标 μ 表示感受野的极性,其取值为 on 或 off,其中 on 表示兴奋型感受野,off 表示抑制型感受野; v_{th} 为脉冲出发阈值; v_{reset} 为静态电势; ref 为绝对不应期; v_g 为脉冲发放峰值; \max 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 又瞬间被重置为 v_{reset} 入绝对不应期,神经元的失活状态将与激励值无关,直到不应期结束。

为了更好地突出图像的细节特征,首先将

NSCT 分解得到的各个高频子带逐一进行高斯平滑,然后在空间上对以图像块为基本单位做一个全局分解:以目标点 (x, y) 式(4)计算窗口覆盖区域的平均值,然后选取与平均值相差最大的像素点作为该区域的全局特征点 $WI_{\mu}(x, y)$,如公式(5)所示。

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

$$WI_{\mu}(x, y) = \max(|GC_{\mu}((x-1) \cdot step + p, (y-1) \cdot step + q) - ave_{\mu}(x, y)|) \quad (5)$$

其中 ws 是窗口的边长, $step$ 为移动步长,一般设置为 $ws = step$; (p, q) 表示全局特征点所在的坐标; $GC_{\mu}(x, y)$ 是高斯平滑处理后的结果,上标 μ 表示感受野的极性。

研究表明,LGN 还存在着 nCRF 感受野类型,因此本文引入局部半波整流的三高斯差函数来模拟 nCRF 感受野的全局调节机制^[14],如式(6)所示。

$$TDoG(x, y) = gau_{sur}(x, y) - \max(gau_{inh}(x, y), gau_{ext}(x, y) \cdot \rho) \quad (6)$$

其中 \max 用于实现半波整流;高斯函数 $gau_{sur}(x, y)$, $gau_{inh}(x, y)$, $gau_{ext}(x, y)$ 分别表示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是中心兴奋区的 4 倍,且大外周区高斯函数与中心兴奋区域相同。本文以三高斯差函数 $TDoG(x, y)$ 作为度量因子,计算全局特征图 $WI_{\mu}(x, y)$ 中的亮度性差异性 $LD_{\mu}(x, y)$,如式(7)所示。

$$LD_{\mu}(x, y) = \max\left(\sum_{(p, q)} WI_{\mu}(x+p, y+q) \cdot TDoG(p, q) \cdot \rho\right) \quad (7)$$

其中 (p, q) 表示像素点的坐标增量。

最后,将 $LD(x, y)$ 的负指数归一化,将其作为 nCRF 调节的结果 $nCRF_aj_{\mu}(x, y)$,并把它与全局特征图按照一定的比例进行融合,使用双线性插值法恢复至原图尺寸,并将其作为 LGN 的输出 $LGN^{\mu}(x, y)$,如式(8)和(9)所示。

$$nCRF_aj_{\mu}(x, y) = -\exp\left(-\frac{LD_{\mu}(x, y)^2}{2\sigma^2}\right) \quad (8)$$

$$LGN^{\mu}(x, y) = \text{resize}\left(WI_{\mu}(x, y) + \alpha \cdot nCRF_aj_{\mu}(x, y)\right) \quad (9)$$

其中 σ 是负指数化的强度参数,一般设置范围为 $0.2 \sim 0.4$; resize 表示双线性插值, α 为融合权值,反映了全局调节的作用强度。

1.3 主视皮层神经元的多感受野朝向性关联模型

对视皮层 V1 区的大量研究表明,V1 区主要负责朝向和颜色等信息的提取,其神经元感受野具有朝向选择性^[15]。研究表明,主视皮层的简单细胞可以由多个相邻的 LGN 细胞感受野的输出相连而成,而简单细胞的方向选择与其相连的 LGN 细胞感受野密切相关。故本文提出了一种 V1 简单细胞多感受野朝向性关联模型,接受来自 LGN 细胞 on 型感受野和 off 型感受野的输出。考虑到离散空间计算的方便性,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;将相邻的 N 个 on 型或 off 型矩形感受野的输出关联到同一个 V1 简单细胞,其中 $1 \leq N \leq \text{round}(\min(m, n)/w)$, m, n 分别为输入图像的宽和高。本文用一个三元素的元胞集合 RecRFSet 表示多个矩形感受野的朝向性分布,具体如式(10)所示:

$$\text{RecRFSet} = \{(\mu_i, r_i, \varphi_i) \mid i = 1, 2, \dots, N\} \quad (10)$$

式中 μ_i 为第 i 个感受野的极性, 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的距离, φ_i 为第 i 个感受野中心位置和 V1 感受野中心位置的连线相对于水平方向的夹角。

根据 (r_i, φ_i) 确定不同感受野的相对位置,可以实现多个感受野的朝向性关联,此时矩形感受野的输出为:

$$\text{RecRF}_{\mu_i, r_i, \varphi_i}(x, y) = LGN_{\mu_i}(x - \Delta x_i, y - \Delta y_i) \quad (11)$$

其中 $\Delta x_i = r_i \cos \varphi_i$ 和 $\Delta y_i = -r_i \sin \varphi_i$ 表示极坐标转换成直角坐标系后的位置偏移量。

由于当且仅当各个矩形感受野都处于响应状态时,多朝向性关联模型才会被激活,因此本文计算所有矩形感受野输出的几何平均作为 V1 简单细胞的方向选择结果 $os_{\theta}(x, y)$,有:

$$os_{\theta}(x, y) = \left(\prod_{i=1}^{2N} \text{RecRF}_{\mu_i, r_i, \varphi_i}(x, y)\right)^{1/(2N)} \quad (12)$$

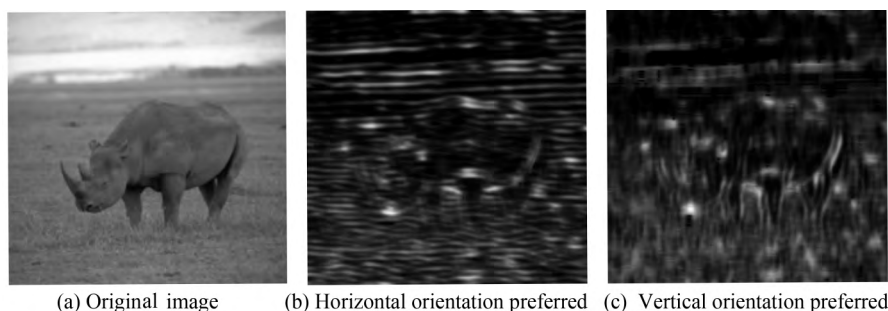


图3 多感受野朝向性关联模型的部分方向选择结果

Fig. 3 Orientation selective results using orientative correlation of multiple receptive fields model

式中 θ 表示 V1 简单细胞的方向偏好。

调整矩形感受野集合 RecRFSet 中每个元胞的极坐标方向 φ_i , 就可以实现其他朝向的关联方式, 从而可以获得其他 V1 细胞在其偏好方向上的方向选择结果。图 3 给出了多感受野朝向性关联模型的部分选择结果。

1.4 皮层下视通路的协同作用

大脑中存在皮层和皮层下两条视通路^[16], 皮层视通路主要负责对外界刺激的精细加工^[17]和智能识别^[18]; 皮层下视通路 (superior colliculus-amygdala, SC-A 通路) 是包括上丘 (superior colliculus, SC), 丘脑枕 (pulvinar, Pulv), 杏仁核 (amygdala, AMG) 的一条视通路。皮层下通路能够对视觉信息的加工具有粗略、快速和自动加工等特点^[19-20]。皮层视通路和皮层下视通路之间存在着复杂的联系, 两者协同完成复杂、完美的视觉系统。

本文引入了皮层下视通路来协同皮层视通路进行视觉信号的处理。又因为被研究者们广泛使用的 DoG 建模函数本身就是一种小波函数, 故其具有一个显著的性质: 多分辨率, 可以利用不同的观察尺度来探测和表征不同尺寸的特征。这与皮层下视通路的特点相似, 故本文通过 DoG 建模函数来模拟皮层下视通路对视觉信息的处理。其基本的数学表达式如式 (13) 所示。

$$\text{DoG}(x, y) = \frac{1}{2\pi(k\varepsilon)^2} \exp\left(-\frac{x^2 + y^2}{2(k\varepsilon)^2}\right) - \frac{1}{2\pi\varepsilon^2} \exp\left(-\frac{x^2 + y^2}{2\varepsilon^2}\right) \quad (13)$$

式中 x 和 y 表示空间变量; ε 决定感受野的大小, k 决定了 nCRF 区域相对于 CRF 区域的大小。研究显示, nCRF 的大小约为 CRF 大小的 2 ~ 5 倍。

设皮层视通路和皮层下视通路所携带的信息分别为 vp_1 和 vp_2 , 则输入到 V1 区的信息 $V1_input$ 为:

$$V1_input = a \cdot vp_1 + b \cdot vp_2 \quad (14)$$

其中 a 和 b 分别为两条视通路信息融合的权重, 两者的取值范围都为 $[0, 1]$ 。

此外, 由于视皮层 V1 区的简单细胞对具有一定朝向和宽度的输入反应最强烈, 从而达到特定特征提取作用。由于 V1 区简单细胞输出都是特定的特征信息, 故这些信息将进行融合。根据目前的研究推测, 复杂细胞可以看作是一系列的简单细胞的感受野叠加而来^[10]。故本文将简单细胞的输出的某个特定特征作为复杂细胞的输入。设共有 n 个特征输入复杂细胞, 则最后的总特征 $total_feature$ 为:

$$total_feature = \sum_{i=1}^n (\beta_i \cdot V1_output(i)) \quad (15)$$

其中 β_i 为每个特征的融合权重, 其取值范围为 $[0, 1]$, $V1_output(i)$ 为经 V1 区处理的第 i 个特征图谱。

2 轮廓检测算法流程

按照视觉通路的信息处理流程, 对轮廓检测过程做如下描述:

1) 根据 LGN 的 M 通路和 P 通路特点性质, 按式 (1)、(2) 来选定合适的滤波器来构建 NSCT 分解, 然后模拟两个通道对视觉信息的处理, 然后根据式 (3) 对处理后的视觉信息进行稀疏编码。

2) 基于感受野的层次变换特性, 引入 LGN 细胞的 nCRF 调节机制, 按式 (4) 和 (5) 对编码后的视觉信息进行全局调节, 突出轮廓信息, 抑制背景纹理信息。

3) 根据 LGN 与主视皮层的层次关联特性,以及主视皮层简单细胞的方向选择特性形成机制,按式(7)~(12)来构建简单细胞的多感受野朝向性关联模型,实现特定方向的轮廓检测。

4) 按式(13)模拟构建皮层下视通路,然后将皮层下视通路的视觉信息传入 V1 区。

5) 按式(14)和式(15)综合全部方向的轮廓信息,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 4 所示。

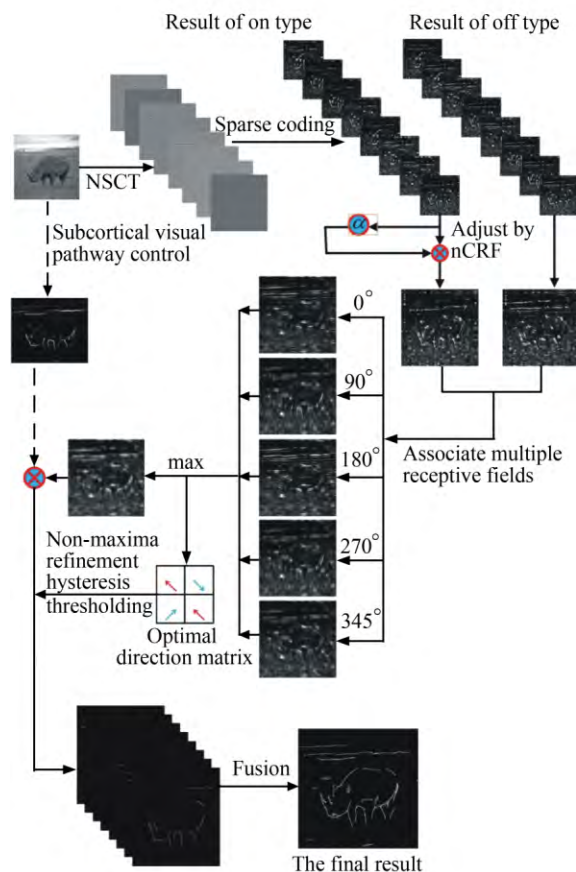


图 4 轮廓检测流程图

Fig. 4 Procedure of contour detect algorithm

The solid lines represent the cortical visual pathway, and the dotted lines represents the subcortical visual pathway.

3 结果

为验证本方法的有效性,本文选取具有强轮廓、纹理相似性、强噪声干扰等特征的自然场景图像进行有效性测试。本文测试用的所有数据均来自 RuG 图库 (<http://www.cs.rug.nl/~imaging/>),该图库中每张图片都有一张手工绘制

的基准轮廓图,用于来定量评价算法有效性。

同时为了说明方法的真实有效性,本文首先选择了主视皮层神经元感受野未经任何后续处理的结果 $E(x, y)$ (Noninh) 作为对比,体现本文后续算法的可行性;另外,选择文献[22]中的空间稀疏限制方法(SSC)和文献[21]中非经典感受野同步抑制方法(ISO)对 $E(x, y)$ 进行相应的纹理抑制处理,得到该方法下的实验结果,并与本文算法进行比较。实验中,对于 NSCT 分解,本文选取的参数是 [2, 3, 3, 3],最后选取第二次分解所得的所有高频信号子带图像分别进行神经网络稀疏编码。同时选取 10 个不同的 LGN 的 nCRF 调节强度,对应的参数为 $\alpha = [0.1 : 0.1 : 1.0]$ 。根据生理研究结果,本文将 nCRF 和 CRF 的直径之比设置为 4。而在多感受野朝向性关联模型中,设定将输入的视觉信息分解为 12 个方向。在多方向信息融合时,权衡两条视觉传输路径在性能准确性和快速性、局部和整体等方面的特性,本文将信息融合权重 a 和 b 之比设为 2:1。此外,为了与图库提供的二值标准轮廓进行比较,需要对提取的轮廓响应进行非极大值抑制处理和滞后阈值处理。滞后门限方法中,取上限阈值 $t = [0.1 : 0.02 : 0.8]$ 。对于 ISO 方法和本文方法可得到 350 组不同参数进行测试,而由于 Noninh 和 SSC 方法并不涉及全局系数,故只进行 10 组参数测试。

图 5 为部分图片的轮廓检测对比图。图 5(a)为图像库原图,图 5(b)为图像库中相应的基准轮廓图,图 5(c)为 Noninh 方法检测结果,图 5(d)为 SSC 方法检测结果,图 5(e)为 ISO 方法检测结果,图 5(f)为本文方法检测结果。

从图 5 可以看出,进行了纹理边缘抑制的方法(SSC、ISO 和本文方法)能有效抑制背景及纹理边缘对主体轮廓的干扰,显著提高主体轮廓凸显能力。Noninh 方法是对全图进行多方位边缘线段检测,细化处理后仅选取一定百分比的强边缘作为最终效果,特点是主体轮廓比较完整,边缘漏检较少,在图像中主体轮廓像素较多时会获得比较不错的效果,如图 Rino,但当背景区域变大、纹理更丰富时,区分能力明显下降。对于 SSC 方法的检测结果,其纹理边缘的抑制作用不够明显,毛刺现象严重。ISO 方法在 3 种对比方法中表现最好,大多数纹理都能有效去除,主体轮廓突出明显,但

在图像各处仍存在大量散列的强纹理边缘。

综合来看,本文方法获得的检测结果主体轮廓相对连续而完整,对强纹理边缘的抑制作用相比ISO方法有着一定的优势,整体检测效果良好。但是,由于本文在模拟皮层视觉通路在对视觉信息进行处理的时候采取了一次双线性插值操作,

这导致了两条视通路所检测出的轮廓信息在空间上存在差异性。此外,皮层视觉通路和皮下层视觉通路存在边缘检测的尺度差异,导致两者检测结果也在空间上存在一定的差异。这两种可能最终导致出现了局部检测结果存在毛刺现象。

基于RuG数据库中的基准轮廓图,利用

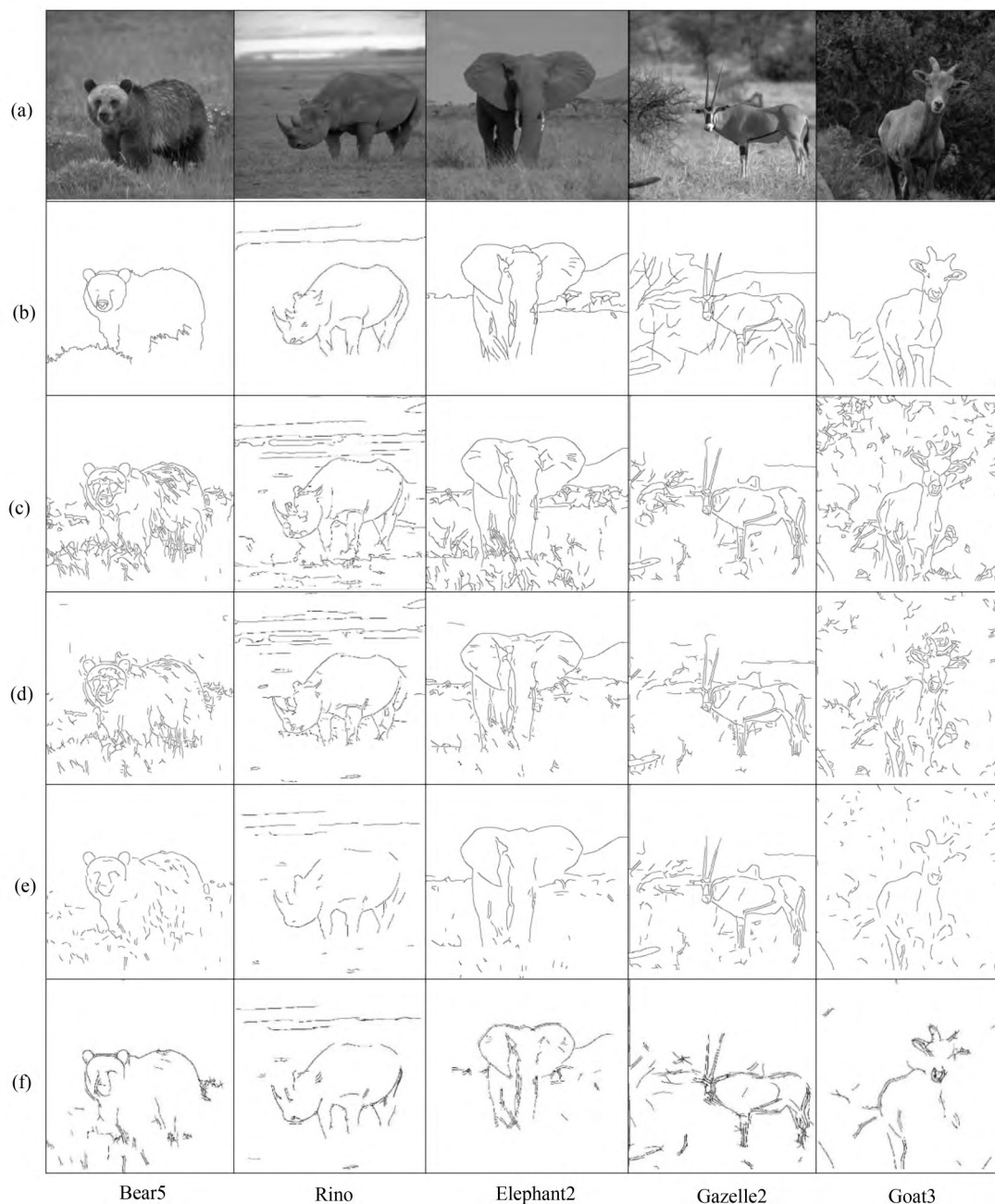


图5 RuG 图库中典型图片及其轮廓检测结果

Fig.5 Some typical images in RuG image library and their contour detection results

The first row is natural images for testing; the second row is the ground truth; the third row is result of Noninh; the fourth row is result of SSC; the fifth row is result of ISO; the last row is the result of the proposed method.

Grigorescu 等提出的轮廓检测定量评价指标^[1] , 对轮廓检测算法效果进行定量评价。首先值得注意的是 ,由于检测到的轮廓像素并不是总与基准轮廓图像像素精确对应 ,因此通常认为只要检测到的像素出现在基准轮廓像素的 5×5 邻域内 ,就认为该像素检测正确。定义 E_D 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 ,那么算法正确检测到的像素集为 $E = E_D \cap (E_{GT} \oplus T)$ (\oplus 为膨胀操作 , T 为 5×5 的结构元) ;错误检测到的像素集为 $E_{FP} = E_D - E$;漏检的像素集为 $E_{FN} = E_{GT} - (E_{GT} \cap (E_D \oplus T))$ 。这样 ,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 可通过式 (16) ~ (18) 计算得到。

$$e_{FP} = \text{card}(E_{FP}) / \text{card}(E_D) \tag{16}$$

$$e_{FN} = \text{card}(E_{FN}) / \text{card}(E_{GT}) \tag{17}$$

$$P = \frac{\text{card}(E)}{\text{card}(E) + \text{card}(E_{FP}) + \text{card}(E_{FN})} \tag{18}$$

其中 $\text{card}(S)$ 为集合 S 中的元素数量。运用公式 (16) ~ (18) 对图 5 中各算法获得的最优效果进行性能指标计算 ,以及获得最优效果时的参数设置 ,计算结果如表 1 所示。由表可知 ,定量的性能

指标与前述对图 5 的定性描述大致吻合。Noninh 方法的漏检率 e_{FN} 较小 ,表明获得的轮廓比较完整 ,但误检率 e_{FP} 较大 ,相应得到的 P 值不太理想。但也有例外 ,如图像 Rino ,该方法获得的 P 值为 0.59 ,优于其他对比方法 ,其原因可能与前文所述一致 ,是由于图像中主体轮廓像素较多所致。SSC 方法对比 Noninh 方法有了一定改善 ,但整体仍表现误检较大、漏检较小 , P 值虽有所改善 ,但效果不佳。而 ISO 和本文方法在误检和漏检之间取得一定的平衡 ,整体性能改善较大 ,但由表中 P 值可以看出 ,本文所提方法整体检测效果都优于其他对比方法。此外 ,计算全部 40 张图片的检测结果与对应基准轮廓图的平均 P 指标 ,本方法为 0.46 ,大于其他 3 种方法 ,可见本方法在总体上也具有更良好的性能。

综合图 5 和表 1 结果可知 ,本文方法具有较好的轮廓提取性能。考察其他对比方法 ,SSC 方法虽与本文方法有着相似的视觉信息处理基础 ,都考虑了视信号处理过程所遵循的稀疏编码机制 ,但 SSC 方法通过数学算子描述空间稀疏特性 ,用于区分纹理区域和非纹理区域 ,只体现了空间稀疏特性 ,并且丧失了稀疏编码的神经元脉冲发放

表 1 图 5 所示图像对应的参数设置及性能指标
Table 1 Parameter settings and performance indexes of images in Figure 5

Image	Algorithm	Parameter		Performance		
		α	t	e_{FP}	e_{FN}	P
Bear5	Noninh	—	0.10	3.30	0.13	0.22
	SSC	—	0.10	2.56	0.19	0.26
	ISO	1.20	0.10	1.48	0.40	0.31
	Our method	0.70	0.48	1.23	0.42	0.35
Rino	Noninh	—	0.10	0.15	0.33	0.58
	SSC	—	0.20	0.62	0.21	0.52
	ISO	0.80	0.20	0.39	0.26	0.56
	Our method	0.60	0.50	0.35	0.29	0.59
Elephant2	Noninh	—	0.10	1.46	0.11	0.38
	SSC	—	0.10	0.75	0.23	0.48
	ISO	1.10	0.10	0.23	0.36	0.54
	Our method	0.70	0.35	0.34	0.28	0.60
Gazelle2	Noninh	—	0.10	0.79	0.33	0.43
	SSC	—	0.10	0.66	0.39	0.43
	ISO	0.60	0.10	0.57	0.38	0.44
	Our method	0.70	0.40	0.61	0.42	0.45
Goat3	Noninh	—	0.10	2.77	0.15	0.25
	SSC	—	0.10	1.78	0.22	0.33
	ISO	1.30	0.10	0.78	0.40	0.40
	Our method	0.60	0.40	0.46	0.49	0.46

的生理基础,故而在检测结果中表现为主体轮廓附近出现大量误检;而本文方法不仅考虑了空间稀疏特性,还通过 LGN 非经典感受野的调节特性,实现了背景纹理的抑制,以及主体轮廓的增强,同时也与生物视觉具有更为相似的特性。另外,ISO 方法对方位检测结果进行不加区分的大范围非经典感受野抑制作用,尽管部分图像取得了较好效果,但当图像中包含主体包括丰富细节时表现并不理想。这是由于当背景中出现较强的纹理边缘时,经过抑制作用与主体轮廓仍然无法区分,而增大抑制系数时,主体轮廓又会出现断裂,漏检率升高,造成整体算法稳定性不强;而本文所提出的方法虽然也考虑了非经典感受野对中心神经元的调控作用,但是还通过神经元网络稀疏编码和皮层下视通路来协同作用,来尽可能地保持视觉信息的完整性,同时又可以起到增强轮廓信息,抑制纹理信息的作用。

4 结论

本文通过引入外侧膝状体对视觉信息的分类功能,实现了增强轮廓信息抑制背景纹理信息。同时构建具有神经电生理特性的神经网络,以较为新颖的方式从时间、空间维度实现稀疏编码特性,更多地考虑了该特性存在的生理基础,而非从矩阵分解等数学分析方式对图像进行降维表示。通过引入外侧膝状体非经典感受野的调节作用以及视觉通路多感受野朝向性关联的特征,同时考虑了皮层下视通路的协同作用,实现了轮廓信息的提取,背景纹理的抑制。针对 RuG 图像库中的自然图像,通过与不加抑制的 Noninh 方法、空间稀疏限制的纹理抑制方法 SSC,以及非经典感受野同步抑制方法 ISO 所提取结果进行比较,表明本文算法不仅对纹理和背景具有抑制作用,而且能有效区分纹理强边缘和主体轮廓,获得较佳的效果。

[参考文献]

- [1] Yi S, Labate D, Easley GR, et al. A shearlet approach to edge analysis and detection[J]. IEEE Transactions on Image Processing, 2009, 18(5): 929-941.
- [2] Grigorescu C, Petkov N, Westenberg MA. Contour detection based on nonclassical receptive field inhibition[J]. IEEE Transactions on Image Processing, 2003, 12(7): 729-739.
- [3] Ursino M, La Cara GE. A model of contextual interactions and contour detection in primary visual cortex[J]. Neural Networks the Official Journal of the International Neural Network Society, 2004, 17(6): 719-735.
- [4] Tang Q, Sang N, Zhang T. Contour detection based on contextual influences[J]. Image & Vision Computing, 2007, 25(8): 1282-1290.
- [5] Tang Q, Sang N, Liu H, et al. Detecting natural image contours by combining visual perception and machine learning[J]. Scientia Sinica, 2013, 43(9): 1124.
- [6] 李孟寒. 仿生脑外侧膝状体模型及若干工作机理研究[D]. 长春: 吉林大学, 2011.
- [7] Li MH. Study on the model and several working mechanism of LGN in the bionic brain[D]. Changchun: Jilin University, 2011.
- [8] Da CA, Zhou J, Do MN. The nonsubsampling contourlet transform: theory, design, and applications[J]. IEEE Transactions on Image Processing, 2006, 15(10): 3089-3101.
- [9] Yan CM, Guo BL, Meng YI. Fast algorithm for nonsubsampling contourlet transform[J]. Acta Automatica Sinica, 2014, 40(4): 757-762.
- [10] 张强, 郭宝龙. 基于非采样 Contourlet 变换多传感器图像融合算法[J]. 自动化学报, 2008, 34(2): 135-141.
- [11] Zhang Q, Guo BL. Fusion of multi-sensor images based on the nonsubsampling contourlet transform[J]. Acta Automatica Sinica, 2008, 34(2): 135-141.
- [12] Hubel DH, Wiesel TN. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex[J]. Journal of Physiology, 1962, 160(1): 106.
- [13] Chouhan AS. An analytical study of leaky integrate-and-fire neuron model using MATLAB simulation[C]//International Journal of Engineering Research and Technology. ESRSA Publications, 2013, 2(4): 2242-2245.
- [14] Alpert S, Galun M, Brandt A, et al. Image segmentation by probabilistic bottom-up aggregation and cue integration[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2012, 34(2): 315-326.
- [15] Li C Y. Integration fields beyond the classical receptive field: Organization and functional properties[J]. News in Physiological Sciences, 1996, 11(4): 181-186.
- [16] Wei H, Zuo Q, Lang B. Multi-scale image analysis based on non-classical receptive field mechanism[C]//International Conference on Neural Information Processing, Springer Berlin Heidelberg, 2011: 601-610.
- [17] Yang K, Gao S, Li C, et al. Efficient color boundary detection with color-opponent mechanisms[C]//Computer Vision and Pattern Recognition, IEEE, 2013: 2810-2817.
- [18] 寿天德. 视觉信息处理的脑机制[M]. 北京: 中国科学技术大学出版社, 1997.
- [19] Shou TD. Brain Mechanism of Visual Information Processing[M]. Beijing: Press of University of Science and Technology of China, 1997.
- [20] Dicarlo J, Zoccolan D, Rust N. How does the brain solve visual object recognition? [J]. Neuron, 2012, 73(3): 415-434.
- [21] Hung CP, Kreiman G, Poggio T, et al. Fast readout of object identity from macaque inferior temporal cortex[J]. Science, 2005, 310(5749): 863.
- [22] Evans HM. The emotional brain: the mysterious underpinnings of emotional life[J]. The Quarterly Review of Biology, 1999, 43(4): 91-95.
- [23] Panksepp J. Affective neuroscience: The foundations of human and animal emotions[J]. American Journal of Psychiatry, 2000, 159(10): 1805.
- [24] Yang KF, Gao SB, Guo CF, et al. Boundary detection using double-opponency and spatial sparseness constraint[J]. IEEE Transactions on Image Processing, 2015, 24(8): 2565-2578.
- [25] Chen M, Yan Y, Gong X, et al. Incremental integration of global contours through interplay between visual cortical areas[J]. Neuron, 2014, 82(3): 682.

(责任编辑: 廖苹)