



论 文

# 视觉感知结合学习的自然图像轮廓检测

唐奇伶<sup>①\*</sup>, 桑农<sup>②</sup>, 刘海华<sup>①</sup>, 陈心浩<sup>①</sup>

<sup>①</sup> 中南民族大学生物医学工程学院, 武汉 430074

<sup>②</sup> 华中科技大学自动化学院, 武汉 430074

\* 通信作者. E-mail: tqtang@gmail.com

收稿日期: 2012-12-19; 接受日期: 2013-01-20

国家自然科学基金 (批准号: 60805006)、湖北省自然科学基金 (批准号: 2009CDB183) 和教育部科学技术研究重点项目 (批准号: 211211) 资助项目

**摘要** 本文结合非经典感受野的视觉特性与机器学习的方法, 提出了一种自然图像轮廓检测模型. 当非经典感受野中的刺激与感受野中心刺激形成一种精确的空间结构时, 将对中心产生一种增强效应; 另一方面非经典感受野中抑制作用会降低同质成分的响应, 我们将这两个机制分别用于增强光滑的轮廓和减少背景中与结构无关的干扰成分. 利用逻辑回归概率模型将感受野中的信息与来自非经典感受野中的信息进行有效融合, 并根据图像的手工标注数据库, 通过学习方法获得一组最优的模型参数. 自然图像的实验结果表明该轮廓检测方法能极大地抑制来自纹理的局部边缘, 减少虚假轮廓, 同时能增强具有一致空间结构的成分, 避免轮廓缺失. 最后利用 Berkeley 图像数据库定量地评价了我们方法的性能, 并与相关方法进行了比较. 该模型不仅为复杂场景中的轮廓检测提供了一个可行的策略, 并有助于对生理视觉机制的理解.

**关键词** 轮廓检测 视觉感知 上下文 学习 信息结合

## 1 引言

许多视觉感知任务都涉及到从自然场景中提取一维的轮廓结构, 这些线性特征描绘了物体的外表形状, 确定了区域之间的界线, 在视觉感知中具有重要的作用<sup>[1]</sup>, 它们对目标的识别、表面的重构以及图像的理解至关重要. 然而自动地提取自然图像的轮廓与边界被证实是一项非常困难的工作, 主要困难在于: 1) 排除大量由背景纹理所产生的局部边缘, 这些成分往往并不属于任何物体的轮廓和边界; 2) 低尺度下的物体, 其本身的本征信息丢失, 使得边界缺乏局部的有效证据, 需要结合场景中上下文的信息进行进一步判断. 这个挑战性的工作引起了视觉科学与计算机科学研究者的广泛注意.

视觉神经研究<sup>[2~5]</sup>表明: 由于视皮层的水平侧连接作用, 初级视皮层 (V1) 神经元对感受野中的刺激响应受到感受野外 (感受野周围) 上下文的影响. 根据视皮层的感知特性已经建立了各种结构特征检测模型. Grossberg 等人<sup>[6]</sup>提出了一个“边界轮廓系统 (boundary contour system)”作为皮层内和皮层之间相互作用的模型, 可以检测一些由视觉错觉所产生的虚假轮廓. Li<sup>[7]</sup>根据感受野内外刺激的方位进行累积传播, 从而增强具有共线结构的成分; 并通过定位同质性发生衰减的地方来检测不

**引用格式:** 唐奇伶, 桑农, 刘海华, 等. 视觉感知结合学习的自然图像轮廓检测. 中国科学: 信息科学, 2013, 43: 1124–1135, doi: 10.1360/112012-630

同区域的边界. 这些视觉模型主要是用于解释人类视觉系统是如何实现轮廓的结合以及边界的分割, 主要是针对人工合成图像, 很少用于自然图像的处理.

Grigorescu 等人<sup>[8]</sup>将初级视皮层非经典感受野抑制特性应用到了自然场景的轮廓检测, 通过各向同性抑制或各向异性抑制减少纹理边缘的影响, 突出区域的边界并保存孤立的轮廓. 但由于他们的模型采用的是一种环形的抑制区域, 有些轮廓也因抑制作用受到一定程度地破坏. 中科院的李朝义院士和同事在视皮层的感知特性以及视觉图像信息处理方面进行了广泛深入的研究<sup>[9,10]</sup>, 最近他们基于抑制作用的方位选择性提出了一种蝶形的抑制区域进行自然图像的轮廓检测, 能更有效地抑制由纹理产生的无实际物理意义的边缘成分, 同时尽量地保存了图像的轮廓与边界<sup>[10]</sup>.

图像轮廓的感知包含了认知理论采用的概率推理, 目前在计算机视觉领域基于学习的轮廓检测方法是主要发展趋势. Konishi 等人<sup>[11]</sup>首先将统计学习的方法引入边界检测, 通过对手工标注的图像学习, 然后根据统计特征进行边界检测, 对于具有纷乱背景的图像, 他们的结果比 Canny 边缘检测算子<sup>[12]</sup>有明显的优越性. 加州伯克利大学的 Martin 等人<sup>[13]</sup>将这个思想进一步的发展, 通过监督学习得到亮度、颜色、纹理多种低层特征的优化组合, 从而形成一个强大的局部边缘检测器, 在去除虚假边缘方面表现出更好的性能, 并且建立了一个手工标注的图像分割标注数据库<sup>[14]</sup>用于训练学习. Tu 与同事<sup>[15]</sup>将边界检测归结为一个二类分类问题, 他们构造了一个新的学习算法 probabilistic boosting tree (PBT), 将图像多个特征通过 PBT 结合从而提高检测的准确率. 高粱等人<sup>[16]</sup>采用模糊手段将图像的局部灰度信息、纹理信息和空间信息有机地融合起来, 求出每个像素所对应的模糊梯度值, 并由此构成模糊梯度特征向量, 通过分类器的学习进行边界检测.

从本质上说这些轮廓与边界的学习方法主要是根据图像局部特征进行判断, 然而在复杂的自然场景中, 仅仅根据局部信息进行轮廓判断往往不足够充分. 例如对一些低对比度区域, 从局部看可能根本不存在边界, 但从一个大范围的上下文推测却可以确定这个局部是边界的一部分, 因为视觉利用了上下文的信息. 形状信息在图像的一个大的距离范围都能产生相互影响, 这一点已被 Gestalt 心理学家充分证实<sup>[17]</sup>.

在我们早期的工作中<sup>[18,19]</sup>, 根据初级视皮层神经元的长程作用机制构建了一种自底向上的轮廓检测模型, 通过兴奋作用增强神经元对光滑轮廓的响应, 通过同方位的抑制作用减弱纷乱边缘的响应. 本文对抑制的属性、增强的区域和信息融合的方法进行新的定义: 1) 由单一的同方位抑制扩展到同质性抑制 (包含了同质纹理和颜色的抑制), 能更有效地减少纹理边缘, 突出区域的边界; 2) 对共圆增强的区域进行了有效地约束, 减少了场景中与结构无关成分的干扰, 增强了光滑轮廓的显著性; 3) 利用逻辑回归分类器将非经典感受野中的增强、抑制信息进行有效融合, 不同于文献<sup>[13]</sup>利用亮度、颜色、纹理等局部低层特征进行边缘检测, 我们的模型基于初级视皮层的感知机理, 捕获了更广泛的上下文信息. 最后通过自然图像检验了该模型的性能, 并与相关工作进行了定量的性能比较.

## 2 非经典感受野的增强与抑制作用

人类大脑视皮层的初级视皮层 (V1 区) 神经元对视野中特定区域的刺激会产生响应, 描述这些神经元响应的区域称为感受野, 感受野外部的周边区域, 都存在着一个对细胞反应起调制作用的外区域, 其面积比传统感受野大得多, 此区域内的小光点刺激并不能直接引起神经节细胞的反应, 但可以易化或抑制其反应, 调节由感受野产生的原有的效果, 称为非经典感受野. 一个细胞不仅仅对一个特定的特征敏感, 同时也受邻近的相似特征的影响. 非经典感受野把视皮层神经元接受特征信息输入的有效空

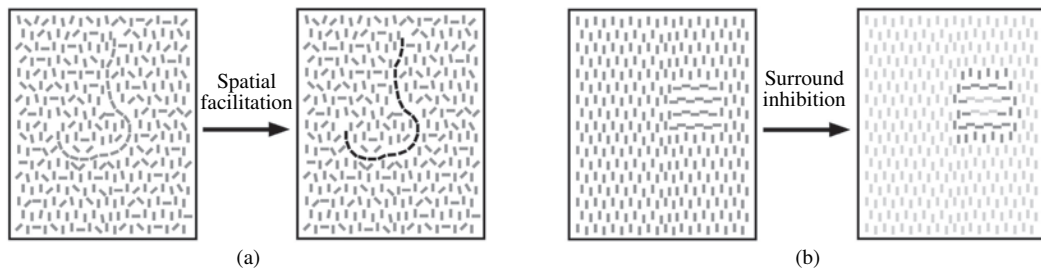


图 1 非经典感受野的增强与抑制作用对显著轮廓的感知

**Figure 1** Role of facilitatory and inhibitory interactions for salient contours. (a) Spatial facilitation, standing out coherent configurations; (b) surround inhibition, emerging out boundaries of texture regions

间范围扩大了数十倍<sup>[9]</sup>, 允许在复杂任务的早期处理中结合视野中大范围的信息 (如轮廓的结合, 表面的分割)。

当光栅和纹理呈现在环境中时通常会产生明显的抑制作用, 侧抑制减少了同质成分的神经元响应, 因此同质性发生变化的地方 (如区域边界) 具有相对更高的显著性; 然而, 当非经典感受野中的刺激与感受野中刺激能排列形成光滑的排列时, 又会增强 V1 神经元对中心刺激的响应, 并且当在环境中有多条线段与中心线段排列形成一个光滑的曲线时, 响应更为强烈<sup>[20]</sup>。因此我们的视觉能容易地从随机散布的边缘线段的背景中识别出由多个的边缘段组成的具有好的空间一致性的结构。我们将利用非经典感受野产生的增强作用实现轮廓结合和显著性突出, 利用抑制作用进行表面和纹理的分割。图 1 显示了非经典感受野的影响对轮廓结合与边界突出的作用。

### 3 特征提取

首先将原始 RGB 图像转换到与视觉感知一致的 CIE L\*a\*b 颜色空间, 从而得到 L 亮度通道和 a, b 颜色通道。二维 Gabor 函数能有效地描述哺乳动物视皮层简单细胞感受野剖面<sup>[21]</sup>, 通过奇偶对简单感受野滤波器的反应模, 能很好地模拟人类的运动与特征定位的检测, 它能捕捉到典型复杂细胞的基本特性。这些复杂细胞可以想象成一个局部方位能量算子, 可以用复杂细胞活动的最大值对图形边与线进行精确定位, 我们通过 Gabor 能量模拟复杂细胞对图像 L 亮度通道的响应。

二维 Gabor 函数是一个被复数正弦函数调制的高斯函数, 其表达式如下:

$$G(x, y, \theta) = g(x', y') \exp(j2\pi\omega x'), \quad (1)$$

式中:  $(x', y') = (x \cos \theta + y \sin \theta, -x \sin \theta + y \cos \theta)$ ,  $\theta$  表示滤波器的方位, 通过在  $x-y$  平面旋转可使滤波器达到任何所期望的方位。  $\omega$  表示中心频率, 它决定了滤波器带通区域中心在频域上的位置。  $g(x, y)$  为如下高斯函数:

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

式中  $\sigma_x$  和  $\sigma_y$  分别是沿着  $x$  和  $y$  轴的高斯包迹的空间常量, 由频率带宽  $B_f$  和方位带宽  $B_\theta$  决定,

$$\sigma_x = \sqrt{\frac{\ln 2}{2} \frac{1}{\pi\omega} \frac{2^{B_f} + 1}{2^{B_f} - 1}}, \quad (3)$$

$$\sigma_y = \sqrt{\frac{\ln 2}{2} \frac{1}{\pi\omega} \frac{1}{\tan(B_\theta/2)}}. \quad (4)$$

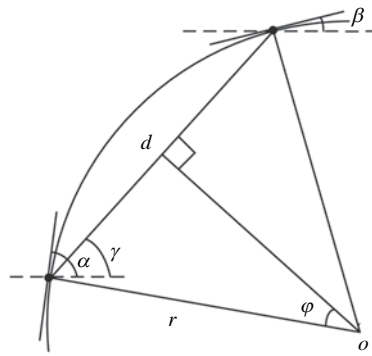


图 2 共圆几何关系

Figure 2 The cocircular geometric relationships

Gabor 函数的实部和虚部分别称为偶对称 Gabor 滤波器和奇对称 Gabor 滤波器, 奇、偶对称 Gabor 滤波器对图像响应的模, 称为 Gabor 能量, 表示为

$$GE(x, y, \theta) = \sqrt{[G_e(x, y, \theta) * I(x, y)]^2 + [G_o(x, y, \theta) * I(x, y)]^2}, \quad (5)$$

式中 \* 表示卷积,  $I$  表示输入图像,  $G_e$ ,  $G_o$  分别为偶对称和奇对称 Gabor 滤波器.

#### 4 增强作用及其上下文选择

Geisler 等人<sup>[22]</sup>通过对自然图像的边缘共生统计表明: 共圆准则能为自然场景中轮廓结合提供一个最大似然局部编组函数 (local grouping function). 共圆几何关系如图 2 所示, 如果中心位置方位为  $\alpha$  ( $0 \leq \alpha < \pi$ ), 那么其环境中与之形成共圆几何关系的方位  $\beta$  应满足

$$\beta = \begin{cases} 2\gamma - \alpha + \pi, & 2\gamma - \alpha < 0, \\ 2\gamma - \alpha, & 0 \leq 2\gamma - \alpha < \pi, \\ 2\gamma - \alpha - \pi, & \pi \leq 2\gamma - \alpha, \end{cases} \quad (6)$$

$\gamma$  为中心与环境成分连线的方位且  $0 \leq \gamma < \pi$ .

曲率是决定自然轮廓可探测性的一个重要因素, 低曲率轮廓的显著性通常更高, 共圆曲率  $k$  计算如下:

$$k = \frac{1}{r} = \frac{2}{d} \sin(\theta) = \begin{cases} \frac{2}{d} \sin \left| \frac{\beta - \alpha}{2} \right|, & 0 \leq 2\gamma - \alpha < \pi, \\ \frac{2}{d} \cos \left| \frac{\beta - \alpha}{2} \right|, & 2\gamma - \alpha < 0 \text{ or } 2\gamma - \alpha \geq \pi, \end{cases} \quad (7)$$

$d$  为中心与环境成分之间的欧氏距离.

目前很多共圆增强模型都是从一个圆形区域中寻找与中心像素形成共圆结构的成分. 但是图像轮廓是一种线性结构特征, 如果从区域上下文考虑其增强作用, 会很大程度受到区域中大量的非结构成分的影响, 这往往导致结构特征被周边的非结构特征湮没. 本文将为轮廓特征选择一种合适的空间支持区域作为其增强作用的上下文, 从而大量地减少了非结构成分对增强作用的干扰. 具体实现如下:

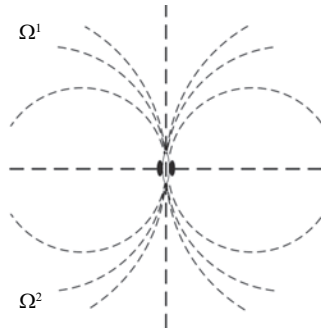


图 3 不同曲率共圆路径示意图

Figure 3 Schematic diagram of cocircular curves

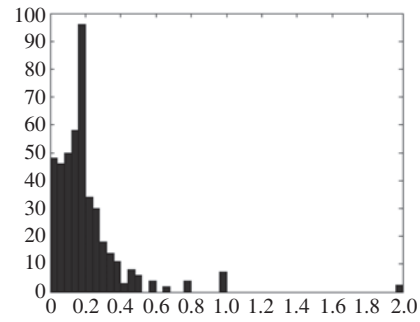


图 4 共圆曲率直方图

Figure 4 The histograms of cocircular curvatures

通过 Gabor 能量滤波器进行特征提取得到图像不同方位的响应, 并将每个像素的最大响应方位作为该位置的主方位. 在以每个像素为中心的局部区域内, 计算所有与中心像素的主方位能形成共圆结构的像素曲率, 图 3 给出了与垂直方位形成共圆结构的不同曲率的路径. 根据曲率计算直方图, 图 4 给出了中心为水平方位在 12 个像素半径的范围内, 与之形成共圆结构像素的曲率直方图统计结果 (横坐标为共圆曲率, 纵坐标为对应直方图 bin 中的像素数目).

沿着中心的法线方向将检测区域分成两部分  $\Omega^1$  和  $\Omega^2$ , 如图 3 所示. 视觉对轮廓的灵敏度随路径的长度和平直度增加, 反映了视觉对低曲率轮廓的偏好, 因此分别对  $\Omega^1$  和  $\Omega^2$  区域中直方图的每个 bin, 计算其曲率加权的平均 Gabor 能量, 表示如下:

$$BE_n^i(x, y, \alpha) = \frac{1}{N(B_n^i) \cdot (1+k)^p} \sum_{(x', y') \in B_n^i} GE(x', y', \beta), \quad i = 1, 2, \quad (8)$$

式中  $N(B_n^i)$  表示在  $\Omega^i$  区域中第  $n$  个直方图 bin 中像素的数目,  $GE$  是对应像素的 Gabor 能量,  $k$  是共圆曲率, 参数  $p$  控制着曲率的影响强度, 其值越大, 表示曲率对直方图 bin 选择的影响就越大, 本文根据实验取值  $p = 2$ , 其值小的波动, 对结果影响不大. 在每个  $\Omega^i$  区域中找到具有最大能量响应的 bin, 记为  $B_{n^*}^i$ ,

$$n^* = \arg \max_n BE_n^i(x, y, \alpha), \quad i = 1, 2. \quad (9)$$

我们将  $B_{n^*}^1 \cup B_{n^*}^2$  定义为对应像素的轮廓支持区域. 从图 5 的示例可以看出通过该方法捕获到的空间支持区域, 能自动适应局部形状的变化, 并极大的缩小了共圆成分的搜索区域. 轮廓支持区域所包含的共圆成分与中心位置形成光滑的曲线, 因此对中心响应产生一种增强的调制作用. 轮廓支持区域所包含的共圆成分产生的增强响应 FE 为

$$FE(x, y) = BE_{n^*}^1(x, y, \alpha) + BE_{n^*}^2(x, y, \alpha). \quad (10)$$

## 5 同质抑制

V1 区神经元对感受野中刺激的响应受到环境的抑制, 其抑制程度取决于中心与环境的对比度. 当周围环境的刺激与中心完全同质时, 抑制程度最强; 抑制效果随中心与环境刺激的距离以及它们的对比度增大而减小. 抑制作用使得具有相同方位、相似颜色、同质纹理成分的响应减弱, 从而使得那些与背景具有不同属性的成分感知突出, 如异质成分、纹理边界.



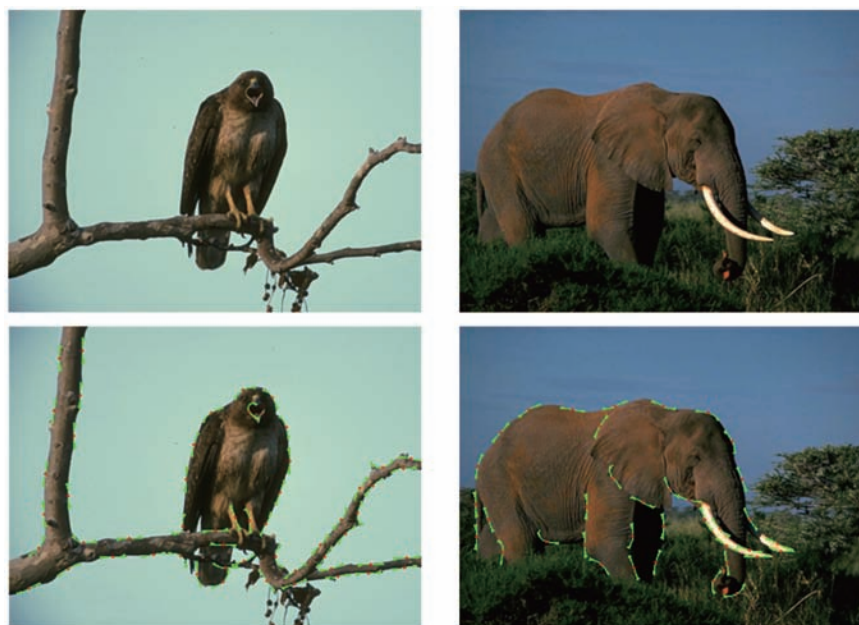


图 5 轮廓的空间支持区域示例. 红色中心为采样位置, 绿色成分表示对应的轮廓支持区域

**Figure 5** Examples of spatial support regions. Red is the sampling positions, and green denotes the corresponding support regions of contours



图 6 从左到右分别对应 64 种纹理元, 输入图像, 纹理元映射

**Figure 6** The left, middle and right panels correspond to 64 textons, input image and texton map, respectively

我们利用颜色与纹理元 (包含了方位信息) 来描述图像的同质性. 纹理元是一些滤波器组简单的线性结合, 被广泛的用于纹理特征分析, 如文献 [23] 利用结构和纹理特征进行场景识别. 本文根据文献 [13,24] 的纹理元获取方法, 采用 13 个不同尺度、方位和形状的滤波器组计算图像特征响应, 通过 K 均值方法从伯克利数据库 200 幅训练集图像聚类得到 64 种常见纹理元, 然后将图像的每个像素映射到最近的纹理元, 如图 6 所示.

两个纹理元  $i$  与  $i'$  之间的距离  $d_t$  计算如下:

$$d_t(i, i') = \|\text{texon}(i) - \text{texon}(i')\|_1, \quad (11)$$

式中  $\|\cdot\|_1$  表示  $L_1$  范数, 其值越大表示纹理元之间的差别越大, 值越小表示纹理元之间的差别越小, 相似性越高.

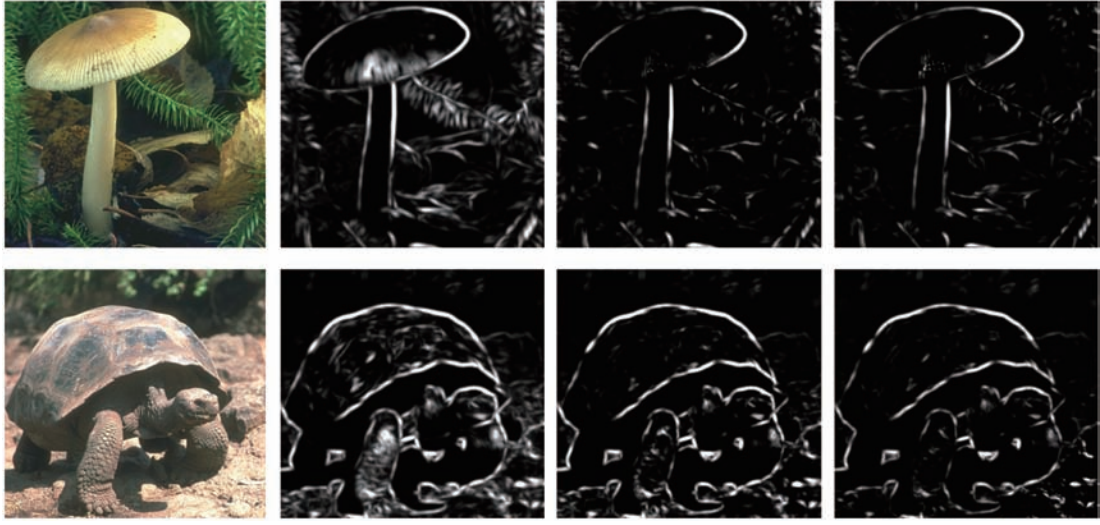


图 7 同质抑制效果, 从左到右分别对应输入图像、主方位的 Gabor 能量响应、纹理抑制结果、颜色抑制结果  
**Figure 7** The effects of homogeneous inhibition. Panels from left to right correspond to input image, Gabor energy, texture inhibition and color inhibition, respectively

两个颜色  $i = (l, a, b)$  和  $i' = (l', a', b')$  之间的距离  $d_c$  计算如下:

$$d_c(i, i') = \sqrt{(l - l')^2 + \tau(a - a')^2 + \tau(b - b')^2}, \quad (12)$$

通常 L\*ab 颜色空间  $\tau = 1$ , 灰度图像对应于特殊情况  $\tau = 0$ .

我们利用高斯差分 (DoG) 函数来模拟抑制环境:

$$\text{DoG}(x, y) = \frac{1}{2\pi(4\sigma)^2} \exp\left\{-\frac{x^2 + y^2}{2(4\sigma)^2}\right\} - \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{x^2 + y^2}{2\sigma^2}\right\}. \quad (13)$$

距离对抑制作用的影响表示如下:

$$w(x, y) = \frac{1}{\|N(\text{DoG})\|_1} N(\text{DoG}(x, y)), \quad (14)$$

$$N(\text{DoG}(x, y)) = \begin{cases} \text{DoG}(x, y), & \text{DoG}(x, y) > 0, \\ 0, & \text{else.} \end{cases} \quad (15)$$

本文将非经典感受野中的同质性抑制度量纹理相似性抑制  $\text{SE}_t$  与颜色相似性抑制  $\text{SE}_c$ , 分别计算如下:

$$\text{SE}_t(x, y) = \sum_{(x', y')} w(x', y') \cdot \exp\left\{-\frac{d_t^2(i, i')}{2\sigma_t^2}\right\} \cdot \text{GE}(x', y', \theta^*), \quad (16)$$

$$\text{SE}_c(x, y) = \sum_{(x', y')} w(x', y') \cdot \exp\left\{-\frac{d_c^2(i, i')}{2\sigma_c^2}\right\} \cdot \text{GE}(x', y', \theta^*), \quad (17)$$

式中  $i, i'$  分别对应于位置  $(x, y), (x', y')$  的纹理元/颜色,  $\theta^*$  对应当前位置的主方位, 即  $\theta^* = \arg \max_{\theta} \text{GE}_{\theta}$ .

图 7 显示了同质抑制的效果 (用每个像素主方位的响应分别减去其纹理同质抑制和颜色同质抑制), 通过抑制作用可以减少自然图像中大量纹理边缘和无实际物理意义的虚假边缘, 同时能很好地保存场景中的轮廓和边界.

## 6 信息结合

根据初级视皮层感知特性: 感受野对刺激的响应受到其非经典感受野的影响, 我们用 Gabor 能量描述感受野对刺激的响应, 同质性抑制和共圆增强描述来自非经典感受野中上下文的影响. 每个像素位置的响应由自身强度 GE, 及其上下文的调制信息组成 (包括共圆增强 FE、同质纹理抑制  $SE_t$  和同质颜色抑制  $SE_c$ ), 因此需要将这多种信息有效地结合成一个整体的检测器. 我们通过监督学习获得各种信息的最优结合规则, 和文献 [11,13,15] 学习多种低层局部图像特征的优化组合区别在于: 我们模型基于初级视皮层的感知机理, 学习的是各种中程上下文信息的结合方法, 其捕获的信息范围更广泛.

对于我们的模型, 逻辑回归 (logistic regression) 是一个简单有效的信息结合方法. 通过对输入和输出之间的关系建模, 训练一个逻辑回归分类器来实现多种上下文信息的有效融合. 逻辑回归函数表示如下:

$$P(y_i | X) = \frac{1}{1 + \exp\{-(w_0 + W_1^T f_i(X))\}}, \quad (18)$$

式中  $y_i$  是一个二值标记;  $f_i(X)$  表示位置  $i$  的特征向量, 由主方位的响应 GE、共圆增强 FE、同质抑制  $SE_t$ 、 $SE_c$  组成,  $W_1^T f_i(X)$  的物理意义为当前位置的 Gabor 能量经过上下文信息调制后所得的最终响应, 即非经典感受野的易化和抑制作用调节了感受野产生的原有效果;  $W = \{w_0, W_1\}$  是模型参数集, 其中  $w_0$  是一个偏置项, 参数集可通过极大似然估计方法训练获得. 选择逻辑回归分类器因为它设计简单, 计算迅速, 并能明确表达信息结合的物理含义. 我们利用伯克利图像分割数据库中的训练集样本 (200 幅图像) 进行学习, 从而得到公式 (18) 的最佳权值组合, 该数据库收集了几百幅涵盖各类场景的自然图像, 并为每一幅图像提供了 5~7 个手工标注的分割结果.

## 7 实验结果

皮层细胞的带宽主要在 1~2 倍频程范围, 选取 Gabor 函数的频率带宽  $B_f = 1.2$ , 采样方位的数目为 12, 方位带宽  $B_\theta = \pi/6$ , 纹理与颜色相似性抑制项中的方差分别为  $\sigma_t = 0.4$ ,  $\sigma_c = 12.8$ . 为了清晰的勾画目标轮廓, 我们采用了 Canny 在其经典论文 [12] 中使用的非最大抑制与滞后门限方法对结果进行了二值化处理.

通过伯克利数据库中的图像来检验模型进行轮廓检测的性能, 并与 Canny 边缘检测算子、Gabor 能量边缘检测算子进行了比较, 如图 8 所示. 从实验结果可以看出: Canny 算子与 Gabor 能量算子 (没有加入上下文调制作用), 由于不能区分物体轮廓与纹理边缘, 因此检测的结果中包含了大量的非物体结构的局部边缘片段. 我们的方法引入了上下文的相互作用, 通过同质抑制, 很大程度消除了局部纹理边缘的干扰, 减少了虚假轮廓, 同时通过共圆兴奋增强了具有一致空间结构成分的响应, 避免了轮廓缺失.

为了进行定量的性能比较, 本文采用文献 [8] 的评价方法. 设  $E_{GT}$  与  $B_{GT}$  分别为 “ground truth” 中的轮廓像素集合与背景像素集合,  $E_D$  与  $B_D$  分别为算子检测结果中的轮廓像素集合与背景像素集合. 正确检测到的轮廓像素集合为  $E = E_D \cap E_{GT}$ . “ground truth” 中存在却被算子漏检的轮廓, 称之为 “错误负”, 通过集合  $E_{FN} = E_{GT} \cap B_D$  给出. “错误正” 即虚假轮廓, 通过集合  $E_{FP} = E_D \cap B_{GT}$  给出. 性能评价指标定义如下:

$$\mathbb{P} = \frac{\text{card}(E)}{\text{card}(E) + \text{card}(E_{FP}) + \text{card}(E_{FN})}. \quad (19)$$





图 8 自然场景中轮廓检测。从左往右分别为: 输入图像、Canny 边缘检测、Gabor 能量算子、本文的结果  
Figure 8 Contour detection in natural images. Panels from left to right correspond to input images, the contours with Canny operator, the contours with Gabor energy operator, the contours by our proposal

表 1 性能指标 P

Table 1 Performance measure P

Row	1st	2nd	3rd	4th	5th	6th	7th	8th
Canny edges	0.28	0.24	0.44	0.38	0.33	0.35	0.43	0.46
Gabor energy	0.24	0.19	0.37	0.36	0.28	0.37	0.39	0.41
Our approach	0.43	0.45	0.51	0.46	0.45	0.53	0.49	0.57

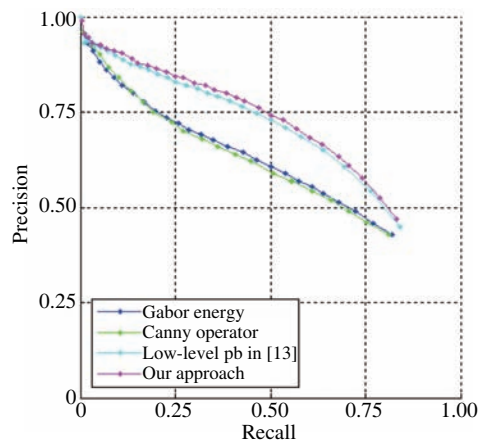


图 9 精度 — 召回率曲线分析

**Figure 9** Precision-recall curves. The performance of our detector compared to Canny operator in [12], Gabor energy edge detector without contextual influences, and low-level pb in [13]

式中  $\text{card}(X)$  表示集合  $X$  中成员的数目.  $\mathbb{P}$  的取值范围为  $[0, 1]$ , 如果所有的真实轮廓像素都被正确检测并且没有背景像素被错误检测作为轮廓像素, 那么  $\mathbb{P} = 1$ ; 如果  $\mathbb{P}$  值越接近 0 表示越多的像素被错误检测或者被轮廓算子漏检. 表 1 给出了图 8 的性能指标.

为了进一步评价算法的有效性, 我们使用精度 — 召回率 (precision-recall) 的评价标准给出了该模型对伯克利图像库检测结果的统计性能分析 (如图 9 所示), 精度 (precision) 是度量边缘检测算法检测到的结果的正确性, 召回率 (recall) 是度量通过算法检测到真实边缘的完整性. 我们的边缘检测性能明显优于局部灰度梯度特征的方法, 如 Canny 边缘算子、Gabor 能量边缘算子, 略微优于融合多种低层特征的边缘检测方法<sup>[13]</sup>.

## 8 结论

本文将非经典感受野的视觉特性引入自然图像轮廓检测, 通过抑制机制大量地减少了背景中由纹理所产生的局部有向边缘, 通过增强机制使得具有一致空间结构的成分从背景中突出, 并采用学习的方式将增强与抑制信息进行了有效融合. 最后的实验结果表明对复杂自然场景中的轮廓检测, 我们的方法比传统方法表现出更好的性能. 模型不仅为复杂场景中的轮廓检测提供了一个可行的策略, 并为人类视觉的感知建模提供了统计学习的途径.

自然场景中轮廓结合是一个非常复杂的问题, 空间频率在轮廓的结合中也起着重要作用, 根据图像的局部特性选择不同的频率将为局部成分的聚集提供更多的结合信息. 将空间频率融入图像轮廓检测是我们进一步工作的发展方向.

## 参考文献

- 1 Baddeley R J, Tatler B W. High frequency edges (but not contrast) predict where we fixate: a Bayesian system identification analysis. *Vision Res*, 2006, 46: 2824–2833
- 2 Bartolucci M, Smith A T. Attentional modulation in visual cortex is modified during perceptual learning. *Neuropsychologia*, 2011, 49: 3898–3907
- 3 Braddick O, Atkinson J. Development of human visual function. *Vision Res*, 2011, 51: 1588–1609
- 4 Herzog M H, Fahle M. Effects of grouping in contextual modulation. *Nature*, 2002, 415: 433–436
- 5 Willmore B D B, Bulstrode H, Tolhurst D J. Contrast normalization contributes to a biologically-plausible model of receptive-field development in primary visual cortex (V1). *Vision Res*, 2012, 54: 49–60
- 6 Grossberg S, Mingolla E, Ross W. Visual brain and visual perception: how does the cortex do perceptual grouping? *Trends Neurosci*, 1997, 20: 106–111
- 7 Li Z. A neural model of contour integration in the primary visual cortex. *Neural Comput*, 1998, 10: 903–940
- 8 Grigorescu C, Petkov N, Westenberg M A. Contour detection based on nonclassical receptive field inhibition. *IEEE Trans Image Process*, 2003, 12: 729–739
- 9 Li C. Integration fields beyond the classical receptive field: organization and functional properties. *News Physiol Sci*, 1996, 11: 181–186
- 10 Zeng C, Li Y, Yang K, et al. Contour detection based on a non-classical receptive field model with butterfly-shaped inhibition subregions. *Neurocomputing*, 2011, 74: 1527–1534
- 11 Konishi S, Yuille A L, Coughlan J M, et al. Statistical edge detection learning and evaluating edge cues. *IEEE Trans Pattern Anal*, 2003, 25: 57–74
- 12 Canny J F. A computational approach to edge detection. *IEEE Trans Pattern Anal*, 1986, 8: 679–698
- 13 Martin D, Fowlkes C, Malik J. Learning to detect natural image boundaries using local brightness, color and texture cues. *IEEE Trans Pattern Anal*, 2004, 26: 530–549
- 14 Fowlkes C, Martin D, Malik J. The berkeley segmentation dataset and benchmark (BSDb). <http://www.cs.berkeley.edu/projects/vision/grouping/segbench/>
- 15 Dollár P, Tu Z, Belongie S. Supervised learning of edges and object boundaries. In: *Proc. Computer Vision and Pattern Recognition*. New York: IEEE, 2006. 1544–1551
- 16 Gao L, Liu X, Liao Z, et al. Fuzzy boundary detection algorithm using multi-information fusion. *Chin J Sci Instrum*, 2011, 32: 2506–2514 [高梁, 刘晓云, 廖志武, 等. 一种基于多信息融合的模糊边界检测算. *仪器仪表学报*. 2011, 32: 2506–2514]
- 17 Kanizsa G. *Organization in Vision: Essays on Gestalt Perception*. New York: Praeger, 1979. 1–267
- 18 Tang Q, Sang N, Zhang T. Extraction of salient contours from cluttered scenes. *Pattern Recogn*, 2007, 40: 3100–3109
- 19 Tang Q, Sang N, Zhang T. Contour detection based on contextual influences. *Image Vision Comput*, 2007, 25: 1282–1290
- 20 Weliky M, Kandler K, Fitzpatrick D, et al. Patterns of excitation and inhibition evoked by horizontal connections in visual cortex share a common relationship to orientation columns. *Neuron*, 1995, 15: 541–552
- 21 Daugman J G. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *J Opt Soc Am A*, 1985, 2: 1160–1169
- 22 Geisler W S, Perry J S, Super B J, et al. Edge co-occurrence in natural images predicts contours grouping performance. *Vision Res*, 2001, 41: 711–724.
- 23 Zhou L, Hu D, Zhou Z. Scene recognition combining structural and textural features. *Sci Sin Inf*, 2012, 42: 687–702 [周莉, 胡德文, 周宗潭. 综合结构和纹理特征的场景识别. *中国科学: 信息科学*, 2012, 42: 687–702]
- 24 Malik J, Belongie S, Leung T, et al. Contour and texture analysis for image segmentation. *Int J Comput Vision*, 2001, 43: 7–27

# Detecting natural image contours by combining visual perception and machine learning

TANG QiLing<sup>1\*</sup>, SANG Nong<sup>2</sup>, LIU HaiHua<sup>1</sup> & CHEN XinHao<sup>1</sup>

<sup>1</sup> College of Biomedical Engineering, South Central University for Nationalities, Wuhan 430074, China;

<sup>2</sup> School of Automation, Huazhong University of Science and Technology, Wuhan 430074, China

\*E-mail: tqltang@gmail.com

**Abstract** A model for detecting contours in natural images is presented by combining the visual perceptual mechanisms of nonclassical receptive fields and machine learning. When the stimuli inside and outside the receptive field (RF) form a precise spatial configuration, the surround stimuli will enhance the response of the central stimulus, on the other hand, surround inhibition will reduce the responses to homogeneous elements. The two different visual mechanisms are used to enhance the well-organized structures and reduce the non-meaningful distractors engendering from texture fields, respectively. We approach the task of information combination as a supervised learning problem using the logistic regression model, where we will learn the combination rules from the ground truth data. Our experiments demonstrate that the model can dramatically suppress texture edges and reduce spurious contours, and meanwhile can enhance the responses to the elements with coherent spatial configurations and avoid ground-truth contours missed by the detector. Finally, we evaluate quantitatively our results with the Berkeley segmentation dataset and compare with related work. The proposed method not only provides a biologically feasible scheme for contour detection but also contributes to further understand visual mechanisms.

**Keywords** contour detection, visual perception, context, learning, cue integration



**TANG QiLing** was born in 1973. He received the Ph.D. degree in the pattern recognition and intelligent systems from the Huazhong University, Wuhan city, in 2007. Currently, he is an associate professor at South Central University for Nationalities. His research interests are in computer vision, image understanding, and perceptual modeling of human vision.



**SANG Nong** was born in 1968. He received the Ph.D. degree in pattern recognition and intelligent systems from Huazhong University of Science and Technology, Wuhan city, in 2001. Currently, he is a professor at Huazhong University of Science and Technology. His research interests include computationally modeling of biological vision perception and applications in computer vision, image analysis and object recognition based on statistical learning, medical image processing and analysis, interpretation of remote sensing images, and intelligent video surveillance.