

基于复合感受野的轮廓检测算法

杜晓凤^① 李翠华^② 李晶^③

^①(厦门大学自动化系 厦门 361005)

^②(厦门大学计算机系 厦门 361005)

^③(北京邮电大学软件学院 北京 100876)

摘要: 自然场景自动检测目标轮廓是非常困难的。该文提出了一种基于复合感受野的图像轮廓检测仿生模型。在传统的基于抑制模型的基础上引入去抑制机制,在减少背景纹理干扰的同时,保护较弱的轮廓,从而尽量保持轮廓的完整性。实验结果表明,这种仿生模型可以有效地抑制纹理边缘,减少对轮廓的破坏,提高了自然背景中轮廓检测的性能。

关键词: 目标识别; 轮廓检测; 初级视皮层; 复合感受野

中图分类号: TP391.41

文献标识码: A

文章编号: 1009-5896(2009)07-1630-05

Contour Detection Based on Compound Receptive Field

Du Xiao-feng^① Li Cui-hua^② Li Jing^③

^①(Department of Automation, Xiamen University, Xiamen 361005, China)

^②(Department of Computer Science, Xiamen University, Xiamen 361005, China)

^③(School of Software Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: Contour detection from natural scenes is a tough task in computer vision. This paper proposes a visual model of contour detection based on compound receptive field. By inducing dis-inhibition mechanism to traditional model that was based on inhibition mechanism, inferences are suppressed while weak edges are preserved, so that the integrity of contour is improved. Experiment results show that this approach can suppress edges due to texture, protect object contour and improve the performance of contour detection from nature scene.

Key words: Target identification; Contour detection; Primary visual cortex; Compound Receptive Field(CRF)

1 引言

视觉系统是一个多层次、动态的图像信息处理系统,其任务是提取环境中的关键视觉信息。探测物体的边界轮廓是视觉系统的一项基本任务,轮廓感知的生理基础是由于感受野具有方向选择性。人们通常认为高等哺乳动物的视觉方位、方向选择性主要起源于高层次的视皮层区,同时初级视皮层V1区的细胞在其中扮演着重要的角色。对V1区神经元的经典感受野及其周边的研究已有很多年,普遍认为每个感受野都有它自己的一个小的兴奋区域,而其周边则是其相邻单元所产生的抑制区。神经元的活动被那些邻近单元所施加的抑制所减弱,这种侧抑制作用是相互影响的。视觉感知中的这种侧抑制机制在机器视觉中受到越来越多的关注。Grigorescu

等人^[1]利用非经典感受野的抑制特性进行轮廓检测,减少了环境纹理的影响,在文献[2, 3]中进而通过多尺度、去噪声等技术,使得算法更加鲁棒。Ursino等^[4,5]在考虑侧抑制的同时,引入层级注意机制,可以得到不同精细度的图像轮廓。桑农等^[6]建立了基于侧抑制区的蝶形模型,提出一个迭代求解的轮廓检测算法。侧抑制机制的作用使得孤立的边缘要比群体边缘更为显著,基于侧抑制区的仿生模型,较好地去除了背景纹理产生的边缘。但抑制机制的引入,在一定程度上降低了边缘的强度,当部分弱轮廓的强度受到抑制而低于二值化设定的阈值时,轮廓中会形成断裂,破坏了轮廓的完整性。

近年来的生物实验研究表明^[7-10],在经典感受野以外范围存在的大周边区域也具有独立的方位选择性。这一方位选择性的平均强度与经典感受野中心相近。大周边与中心的方位选择性可能不同,对整个大范围的复合感受野(CRF)的方位选择性贡献也不同,并且大周边刺激的存在将调制中心的方位选择性,反之亦然。文献[7]中将复合感受野分为经

2008-09-16 收到, 2009-03-12 改回

国家 973 计划项目(2007CB311005), 国家 863 计划项目(2006AA01Z129), 福建省自然科学基金(A0710020)和厦门大学 985 二期信息创新平台项目资助课题

典感受野、经典外周区和去抑制区3个亚单元。Kapadia^[11]指出, V1区神经元的外周对方位的敏感性对中心区有着不同的影响, 而不仅仅是抑制作用。正是由于神经元这种复杂的方位感知功能, 才使得视觉系统可以从复杂的背景中迅速而完整地提取物体边界。

感受野在受到外界刺激时, 不但受到周边的抑制, 同时也受到大周边的去抑制区的影响。在轮廓检测中, 如果只考虑传统的侧抑制机制, 轮廓会受到周边背景纹理的抑制, 则有可能使得检测到的轮廓断裂, 影响后续的目标识别; 在引入去抑制机制后, 弱轮廓可得到加强, 从而可以保持轮廓的完整性, 以提高目标检测的性能。本文主要研究基于复合感受野的轮廓检测算法。文章的组织结构如下: 第2节是基于复合感受野模型的设计, 第3节介绍轮廓检测算法, 第4节为实验及结论, 最后一部分是结束语。

2 基于复合感受野的轮廓检测

2.1 Gabor能量

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, 可以用来表达典型复杂细胞的基本特性, 通常利用奇偶对简单感受野滤波器的反应模来模拟视觉系统对运动与特征定位的检测。而复杂细胞可以看成局部方位能量算子, 通过Gabor能量来模拟复杂细胞的响应^[1]。二维Gabor函数表达式如下:

$$g_{\lambda, \sigma, \theta, \varphi}(x, y) = \exp\left[-\frac{\tilde{x}^2 + \gamma \tilde{y}^2}{2\sigma^2}\right] \cos\left(2\pi \frac{\tilde{x}}{\lambda} + \varphi\right) \quad (1)$$

其中 $\tilde{x} = x \cos \theta + y \sin \theta$, $\tilde{y} = -x \sin \theta + y \cos \theta$ 。 γ 为一个表示椭圆形感受野长短轴比例的常数, 参数 λ 为波长, $1/\lambda$ 为余弦函数的空间频率, σ/λ 决定了空间频率的带宽, φ 是相角参数。

对输入的图像, 简单细胞的响应 $r_{\lambda, \sigma, \theta, \varphi}$ 为感受野函数 $g_{\lambda, \sigma, \theta, \varphi}$ 与图像 I 的卷积“*”。

$$r_{\lambda, \sigma, \theta, \varphi}(x, y) = (I * g_{\lambda, \sigma, \theta, \varphi})(x, y) \quad (2)$$

复杂细胞响应定义为一对奇偶对(相角相差 $\pi/2$)简单细胞响应的组合, 计算方法如下:

$$R_{\lambda, \sigma, \theta}(x, y) = \sqrt{r_{\lambda, \sigma, \theta, 0}^2(x, y) + r_{\lambda, \sigma, \theta, \pi/2}^2(x, y)} \quad (3)$$

2.2 基于侧抑制机制的感受野模型

传统的基于侧抑制的感受野模型一般定义为: 感受野中心 (x, y) 对刺激的反应强度 $E_{\text{center}}(x, y)$ 减掉抑制成分, 抑制成分为区域权重函数 w_{in} 与抑制区响应 E_{in} 的卷积:

$$E(x, y) = aE_{\text{center}}(x, y) - b(w_{\text{in}} * E_{\text{in}})(x, y) \quad (4)$$

式中 a 和 b 为常数。

文献[1]建立了原始的抑制区域模型, 即中心与周边的圆环形的抑制区域。进一步, 学者们提出一些改进算法, 其目的是为了避免抑制成分对轮廓完整性的破坏。文献[2]中将抑制区变为两个半圆环形区域以避免轮廓本身的自抑制, 文献[6]中提出了一个蝶形的抑制区域以避免轮廓成分之间的相互抑制。

2.3 基于复合感受野模型

文献[7]中将同类刺激定义为那些出现在外周区域并与中心区最优响应方向相同的刺激, 并指出方位选择性感受野的形成主要是由椭圆形感受野长轴和短轴方向上的空间频率响应不同造成的, 而出现在不同区域的同类刺激对中心区的影响是不同的。Kapadia^[11]指出抑制区域位于与最优方位轴向垂直方向的两侧, 这一区域中的同类刺激会对中心响应产生抑制作用, 而沿感受野的最优方位轴向的同类刺激却会增强神经元的响应。图1为复合感受野各区域的示意图。沿最优方位轴向的区域为去抑制区 A_{dis} , 轴向垂直方向的两侧区域为抑制区 A_{in} 。在空间位置上, 外周区域(包括抑制区 A_{in} 与去抑制区 A_{dis})对中心的影响程度是与到感受野中心的距离成反比的。在考虑抑制机制的同时, 引入去抑制机制, 定义复合感受野中心 (x, y) 的响应强度如下:

$$E(x, y) = aE_{\text{center}}(x, y) - b(w_{\text{in}} * E_{\text{in}})(x, y) + c(w_{\text{dis}} * E_{\text{dis}})(x, y) \quad (5)$$

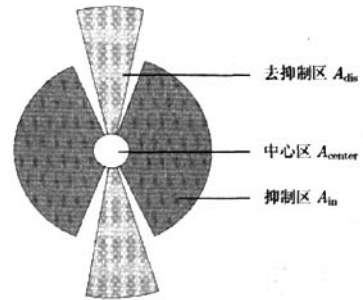


图1 复合感受野各区域

即中心感受野的响应强度 E_{center} 减掉抑制区的响应 E_{in} 与权值函数 w_{in} 的卷积, 然后加上抑制区的响应 E_{dis} 与权值函数 w_{dis} 的卷积。 a , b 和 c 为常数。

下面介绍区域能量的计算模型, 为了描述简单, 这里用 A 代表抑制区 A_{in} 或去抑制区 A_{dis} , 设 σ_A 为 A 区域的带宽, σ 为中心区的带宽, 定义:

$$\text{DoG}(x, y, \sigma_A) = \frac{1}{2\pi\sigma_A^2} \exp\left(-\frac{x^2 + y^2}{2\sigma_A^2}\right) - \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (6)$$

$$w_d(x, y, \sigma_A) = \frac{N(\text{DoG}(x, y, \sigma_A))}{\|N(\text{DoG})\|_1} \quad (7)$$

$\| \cdot \|_1$ 为一阶 (L^1) 范数。

$$N(\text{DoG}(x, y)) = \begin{cases} \text{DoG}(x, y), & \text{DoG}(x, y) > 0, (x, y) \in A \\ 0, & \text{其他} \end{cases} \quad (8)$$

最终区域A的能量是区域内所有复杂细胞的最优响应的总和, 定义为

$$E_A = \sum_{(x, y) \in A} R_{\lambda, \sigma, \theta}(x, y) \quad (9)$$

其中 $\Theta(x, y) = \theta_k$, $k = \arg \max \{R_{\lambda, \sigma, \theta_i}(x, y) | i = 1, 2, \dots, N_\theta\}$ 。

$\Theta(x, y)$ 表示感受野中心 (x, y) 的最优响应方位, $R_{\lambda, \sigma, \theta}(x, y)$ 为复杂细胞的最大响应。在实际的轮廓检测中, λ , σ 和 N_θ 是常数。

2.3.1 抑制区权值函数模型 抑制区对中心响应的抑制程度, 随感受野中心的距离增大而减小, 随感受野最优方位的夹角增大而减小^[11], 并且抑制的程度与刺激的类型相关。我们定义抑制区的权值函数为

$$w_{in}(x, y) = w_d(x, y, \sigma_{A_{in}}) \cdot w_r \cdot w_e \quad (10)$$

其中 w_d 为距离权值, $\sigma_{A_{in}}$ 为抑制区的带宽, 见式(7)。 w_r 为方位权值。设 α 为感受野响应的最优方位, β 为抑制区中同类刺激出现的方位, 定义 w_r 如下:

$$w_r = \exp\left(-\frac{\theta^2}{2\sigma_r^2}\right), \theta = \min(|\beta - \alpha|, \pi - |\beta - \alpha|) \quad (11)$$

w_e 为相似度权值。设 θ 为抑制区中对刺激的响应的最优方位, 也就是刺激的方向。定义 w_e 如下:

$$w_e = \exp\left(-\frac{\varphi^2}{2\sigma_e^2}\right), \varphi = \min(|\theta - \alpha|, \pi - |\theta - \alpha|) \quad (12)$$

图2中a, b和c分别是出现在抑制区不同方位的同类刺激(线段的方向与出现在中心区的线段方向相同), d处为非同类刺激(线段的方向与出现在中心区的线段方向不同), a, b和d到中心的距离相等, 另外, b, d与中心最优方位轴向的夹角相同。从距离上讲, 在远离中心区的c处的刺激带来的抑制程度, 要小于a处与b处。而从方位上看, 虽然a, b两个位置与中心区距离相同, 但在与感受野最优方位垂直的a处的刺激所造成的抑制, 大于出现在b处的刺激。最后, 由相似度上相比, 在d位置出现的是非同类刺激, 由式(12)可知, 刺激类型越不相同, 也

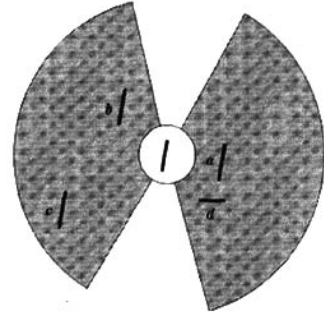


图2 抑制区

就是说刺激之间的夹角越大, 抑制程度越小, 因此, 虽然b与d处于距感受野中心区同样距离, 同样的方位夹角, 即有相同的 w_d 和 w_r , 但它们的相似度权值 w_e 不同, 对中心区域抑制程度也是不同的。这里采用文献[6]中的蝶形区域, A_{in} 位于感受野的两侧以减少轮廓之间的相互抑制, 单个扇状区域的圆心角为 $5\pi/6$ 。

2.3.2 去抑制区权值函数模型 完整真实的轮廓往往是连续的, 局部平滑的, 也就是说, 在沿着轮廓片断的方向上, 往往存在着相同方向的边缘。这样, 就可以根据边缘上下文来提高轮廓检测的性能。如图3所示, 复合感受野的去抑制区域一般位于感受野中心最优响应方向的两端, 其对中心响应的影响与抑制区的恰好相反, 起到增强的作用, 从而可以保护轮廓。权值函数同样与三个方面相关, 分别是到感受野中心的距离, 与感受野最优方位的夹角及与中心响应的相似度。定义去抑制区的权值函数为

$$w_{dis}(x, y) = w_d(x, y, \sigma_{A_{dis}}) \cdot w_r \cdot w_e \quad (13)$$

其中 w_{dis} 为距离权值, $\sigma_{A_{dis}}$ 为去抑制区的带宽, 定义如式(7)。另外, w_r 为方位权值, w_e 为相似度权值, 其定义与式(11), 式(12)相同。如图3所示, 去抑制区同样是一个蝶形区域, 以距离权值为例, 图3中a, b两处出现的同类刺激对中心响应的增强程

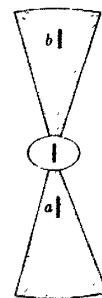


图3 去抑制区

度与到感受野中心的距离成反比, 即 a 处的权重大于 b 处。同样, 方位权值与相似度权值的分布与抑制区分布相似。这里, 单个扇状区域的圆心角为 $\pi/9$ 。

3 轮廓检测

设待检测图像为 I , 对其中每一点 $I(x, y)$:

- (1) 利用式(3)进行Gabor滤波, 得到滤波结果 $E_{center}(x, y)$;
- (2) 利用式(9)与式(10)得到点 (x, y) 处的抑制成分 $(w_{in} * E_{in})(x, y)$;
- (3) 利用式(9)与式(13)得到点 (x, y) 处的去抑制成分 $(w_{dis} * E_{dis})(x, y)$;
- (4) 由式(5), 得到响应结果 $E(x, y)$ 。

后续处理, 采用了Canny算法中的非最大抑制与滞后门限的方法对结果进行二值化处理并去除短小的噪声边缘。

4 实验及结论

首先对合成图像进行轮廓检测来验证模型的性能。图4(a)是一幅合成图像, 在垂直线的中部的两侧有同类的竖线; 图4(b)为图像的Gabor滤波后的能量图; 图4(c)为利用抑制模型得到的结果, 同类刺激造成的抑制不仅出现在纹理区, 也对长竖线的中部产生了抑制; 图4(d)为使用去抑制区模型得到的结果, 很明显, 长竖线整体都有较强的响应; 图4(e)为利用式(4)得到的结果; 图4(f)为仅利用式(5)得到的结果。由图4(e)与图4(f)的结果可以看出, 去抑制模型有利于保护长边缘的连续性, 避免了单纯的抑制模型造成的边缘的断裂。

在对比实验中, 采用的Gabor滤波器的参数与文献[1]一致, $\sigma/\lambda = 0.56$, $\gamma = 0.5$, $N_\theta = 12$, DoG函数的方差 $\sigma = 3$, $\sigma_{A_{in}} = 4\sigma$, $\sigma_{A_{dis}} = 5\sigma$, 抑制区的

方位权值函数与相似度权值函数方差 $\sigma_r = \sigma_e = \pi/6$; 去抑制区方位权值函数方差 $\sigma_r = \pi/6$, $\sigma_e = \pi/3$; $a = 1, b$ 一般在 $[2, 3.5]$ 之间, c 在 $[0.5, 1]$ 之间; 对于图像大小为 512×512 , 噪声边缘定义为长度小于 20 个像素的边缘。

采用文献[1]中的评测方法。评测公式如下:

$$P = \frac{\text{card}(C)}{\text{card}(C) + \text{card}(C_{FP}) + \text{card}(C_{FN})} \tag{14}$$

评测标准 P 在 $[0, 1]$ 之间。式中 $\text{card}(X)$ 表示集合 X 中成员的数目; C , C_{FP} 和 C_{FN} 分别表示正确检测的轮廓, 虚假轮廓与遗漏的轮廓。如果所有的真实的轮廓都正确地检测了出来, 并且没有背景边缘被错检为轮廓像素, 则 $P = 1$ 。错检(漏检)越多时, P 越接近 0。图 5 的各种检测结果的性能指标如表 1, 测试图像来自文献[1]。基于抑制机制的方法主要是从控制抑制区域的大小上来减轻对轮廓的影响, 而没有考虑轮廓上下文之间的相互作用。这样在后续的二值化中, 弱的轮廓由于受到抑制而被破坏, 造成了漏检。复合感受野模型中的去抑制成分对中心响应起到的是增强作用, 使得弱的轮廓强度得以增大, 从而保护了轮廓。

表 1 评测标准 P 对比数据

	各向同性抑制 ^[1]	各向异性抑制 ^[1]	本文方法
elephant 2	0.42	0.40	0.51
goat 3	0.34	0.32	0.44
hyena	0.55	0.51	0.60

5 结束语

轮廓检测是基于形状的目标识别任务的关键。只有尽可能地提取完整的目标轮廓, 才能使目标识别更加准确。人类的视觉系统可以快速地背景中辨识出目标的轮廓, 本文提出的模型, 模拟了视觉系统的抑制机制与去抑制(增强)机制, 在考虑背景抑制的同时, 也考虑了目标轮廓上下文之间的增强作用。在后续处理这一步骤中, 可以与其它阈值选择算法和细化算法相结合, 进一步提高轮廓检测的精度。另外, 视觉系统是一个迭代反馈的过程, 通过重复的迭代, 不断加强对最优方位的响应, 如何利用这些生物视觉上的发现, 建立合理的模型, 尚需要开展进一步的工作。

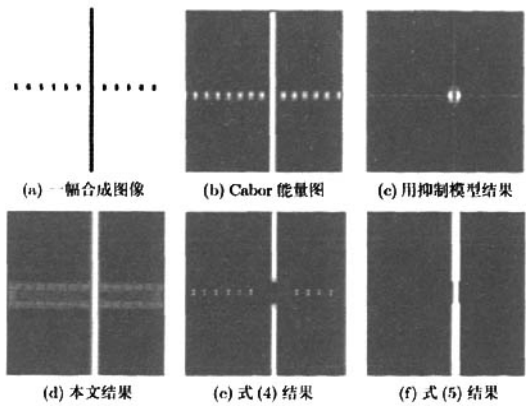


图4 合成图像

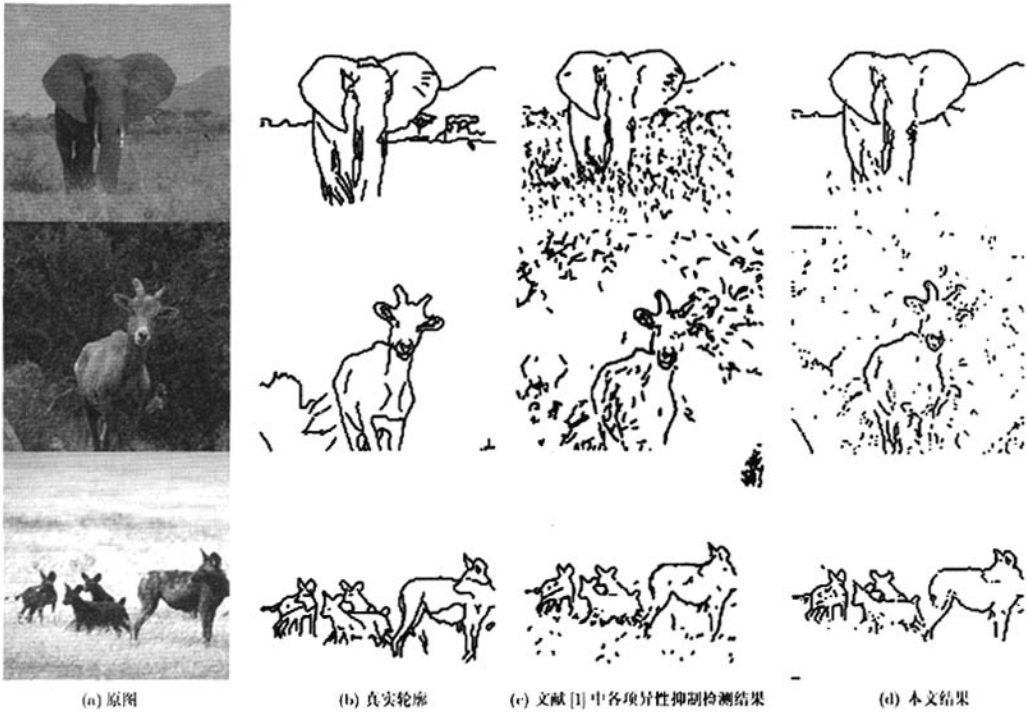


图 5 对比实验

参 考 文 献

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

[2] Grigorescu C, Petkov N, and Westenberg M A. Contour and boundary detection improved by surround suppression of texture edges [J]. *Image and Vision Computing*, 2004, 22(8): 609-622.

[3] Papari G, Campisi P, Petkov N, and Neri A. A biologically motivated multiresolution approach to contour detection [J]. *EURASIP Journal on Advances in Signal Processing*, 2007, 1(Article ID 71828): 119.

[4] Ursino M and La Cara G E. A model of contextual interactions and contour detection in primary visual cortex [J]. *Neural Networks Archive*, 2004, 17(5-6): 719-735.

[5] La Cara G E and Ursino M. A model of contour extraction including multiple scales, flexible inhibition and attention [J]. *Neural Networks*, 2008, 21(5): 759-773.

[6] 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.

Sang Nong, Tang Qi-ling, and Zhang Tian-xu. Contour detection based on inhibition of primary visual cortex [J]. *Journal of Infrared and Millimeter Waves*, 2007, 26(1): 47-51.

[7] 寿天德, 周逸峰. 视觉系统皮层下细胞的方位和方向敏感性

[J]. *生物物理学报*, 1996, 48(2): 105-112.

Shou Tian-de and Zhou Yi-feng. Orientation and direction sensitivity of cells in subcortical structures of the visual system [J]. *Acta Physiologica Sinica*, 1996, 48(2): 105-112.

[8] 操安, 周逸峰, 寿天德. 视皮层细胞方位选择性研究的新进展 [J]. *生理科学进展*, 1998, 29(1): 71-73.

Cao An, Zhou Yi-feng, and Shou Tian-de. A new progress of orientation sensitivity of cells in visual cortical [J]. *Progress in Physiological Sciences*, 1998, 29(1): 71-73.

[9] 邱志诚, 黎斌, 顾凡及等. 视网膜神经节细胞感受野的一种新模型 II [J]. *生物物理学报*, 2000, 16(2): 296-302.

Qiu Zhi-cheng, Li Zang, and Gu Fan-ji, et al. A new computational model of retinal ganglion cell receptive fields II [J]. *Acta Biophysica Sinica*, 2000, 16(2): 296-302.

[10] Cai Chao-feng, Liang Pei-ji, and Zhang Pu-ming. A simulation study on the encoding mechanism of retinal ganglion cell [J]. *Lecture Notes in Computer Science*, 2007, 4689: 470-479.

[11] Kapadia M K, Westheimer G, and Gilbert C D. Spatial distribution of contextual interactions in primary visual cortex and in visual perception [J]. *Neurophysiol*, 2000, 84(4): 2048-2062.

杜晓凤: 女, 1975 年生, 博士生, 研究方向为图像处理与模式识别。

李翠华: 男, 1960 年生, 博士, 教授, 博士生导师, 主要从事计算机视觉、图像处理与分析、小波变换理论及其应用等方面的研究工作。