

初级视皮层机制启发下的特征提取

吕 超^{1,2}, 许悦雷¹✉, 马红强¹, 辛 鹏³, 李 帅¹, 秦占师², 田 雨², 朱源才²

(1. 空军工程大学航空工程学院, 西安, 710038; 2. 95876 部队, 甘肃山丹, 734100; 3. 93263 部队, 辽宁锦州, 121011)

摘要 针对传统计算机视觉方法在纹理细节提取和局部运动感知方面的局限性, 借鉴生物视觉系统的强大信息处理机制, 提出一种基于三维时空能量模型的初级视皮层(V1)视觉模型。模型首先利用三维 Gabor 滤波器初步模拟得到简单细胞感受野, 接着采用半波整流操作进一步得到简单细胞模型, 随后通过能量模型整合简单细胞感受野进行, 模拟得到复杂细胞模型。仿真结果表明: 所提模型在图像静态纹理特征和简单视频序列运动特征提取上取得了较好的效果。较传统模型而言, 模型将计算机视觉与生物视觉有效融合, 具有更强的动静态特征表达提取能力, 体现了 V1 细胞强大的信息处理机制, 为视觉脑启发计算提供新的思路。

关键词 初级视皮层; 特征提取; Gabor 滤波器; 能量模型; 感受野

DOI 10.3969/j.issn.1009-3516.2018.06.008

中图分类号 TP391.41 **文献标志码** A **文章编号** 1009-3516(2018)06-0048-05

A Feature Extraction Inspired by Primary Visual Cortex Mechanism

LYU Chao^{1,2}, XU Yuelei¹✉, MA Hongqiang¹, XIN Peng³, LI Shuai¹, QIN Zhanshi²,
TIAN Yu², ZHU Yuancai²

(1. Aeronautics Engineering College, Air Force Engineering University, Xi'an 710038, China;
2. Unit 95876, Shandan 734100, Gansu, China; 3. Unit 93263, Jinzhou 121011, Liaoning, China)

Abstract: In view of the limitation by the traditional methods in texture feature extraction and local motion perception, and referred to the powerful information processing mechanism of human visual system, a primary visual cortex(V1) vision perception model based on 3D spatio-temporal motion energy model is proposed. Firstly, simple cell receptive field is simulated by using 3D Gabor filter, then by half wave rectification operation, and further simple cell model is obtained. Finally, the simple cell receptive field is integrated by energy model, so that complex cell model is simulated. The simulation results show that the model possess good performance at extracting static texture feature of images and the local motion feature of simple video sequences. Compared with the traditional model, the model combines computer vision well with bio-vision, having stronger abilities at feature representation and extraction, and showing more powerful information processing mechanism of V1 cells. And this also provides new ideas for brain-inspired visual computing.

Key words: primary visual cortex; feature extraction; Gabor filter; motion energy; receptive field

收稿日期: 2018-03-13

基金项目: 国家自然科学基金(61701524)

作者简介: 吕 超(1991—), 男, 江苏连云港人, 硕士生, 主要从事图像处理研究。E-mail: lvchao1112@163.com

通信作者: 许悦雷(1975—), 男, 河北辛集人, 教授, 主要从事图像处理研究。E-mail: 719783186@qq.com

引用格式: 吕超, 许悦雷, 马红强, 等. 初级视皮层机制启发下的特征提取[J]. 空军工程大学学报(自然科学版), 2018, 19(6): 48-52. LYU Chao, XU Yuelei, MA Hongqiang, et al. A Feature Extraction Inspired by Primary Visual Cortex Mechanism[J]. Journal of Air Force Engineering University (Natural Science Edition), 2018, 19(6): 48-52.

随着美国、欧盟于 2013 年相继实施脑计划,大脑逐渐成为科学界的研究热点^[1]。我国于 2016 年正式开展以“脑科学和脑启发智能”为主题的脑计划,旨在推动以脑科学基础研究、生物学及人工智能这“一体两翼”的发展^[2]。视觉系统作为人类最重要的感觉系统,处理了人脑所接受到的 80% 的外界信息^[3]。探索人脑的视觉信息处理机制,构建更加接近人脑的视觉感知模型,对于智能机器人、无人作战飞机的研发等有着重要意义^[4]。大量生物研究表明,初级视皮层(Primary Visual, V1)作为视皮层的起点,是腹侧通路和背侧通路的共同区域,有着强大的静态细节感知与运动特征提取能力^[5]。探索 V1 细胞的感受野特性,并应用于视频图像的感知领域,是脑启发视觉的热点之一。

自从 1962 年 Hubel 和 Wiesel 对生物视觉系统 V1 进行了实验研究^[6],诸多 V1 模型被接连提出,这也使得其成为学者们研究最深入的视皮层区域。V1 对传入的视觉刺激进行一定的处理,可提取出轮廓、纹理等静态细节信息和速度、方向等运动信息^[7],并分别传入负责外观特征感知的腹侧通路和负责运动感知的背侧通路^[8-9]。学者们常用二维 Gabor 滤波器和小波变换等操作模拟简单细胞的感受野,并通过 MAX 模型和学习模型等方法整合得到复杂细胞的感受野^[10]。这些模型一定程度上实现了 V1 细胞的静态感知能力,但对纹理细节信息和运动信息的提取能力仍显不足。Simoncelli^[11]、Rust^[12]和邹洪中^[13]等人先后用高斯时空滤波器实现了对 V1 运动感知机制的模拟,但模型很难提取静态细节信息;整体看来,上述方法很难在静态感知和动态感知上取得统一,与 V1 细胞强大的动静态特征感知能力仍有很大差距。

鉴于此,本文充分探索 V1 细胞的信息处理机制,根据 V1 简单细胞和复杂细胞的感受野特性,提出一种基于三维时空能量模型的 V1 视觉模型。

1 三维时空能量模型

Gabor 滤波器由 Dennis Gabor 于 1946 年提出,是一类线性的带通滤波器^[14]。Gabor 滤波器具有方向选择和尺度选择特性,且对光照变化有良好的适应性,加之其与 V1 细胞感受野的相似性,Gabor 滤波器在图像处理领域应用广泛。同 V1 细胞相似,不同的 Gabor 滤波器可以感知不同区域不同方向的特征。单个 Gabor 滤波器所提取的图像特征是单一的,将一定数量的 Gabor 滤波器组的响应整合,可以得到较为完整的图像特征。本文模型的结构原理见图 1。

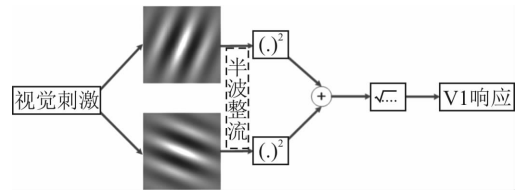


图 1 本文模型流程图

Fig. 1 The steps of the model proposed

根据图 1,模型首先采用三维 Gabor 滤波器^[15]来表示简单细胞的感受野 $g_{v,\theta,\varphi}(x,y,t)$,其表达式为:

$$g_{v,\theta,\varphi}(x,y,t) = \frac{\gamma}{2\pi\sigma^2} \exp\left[-\frac{((\bar{x}+v_c t)^2 + \gamma^2 \bar{y}^2)}{2\sigma^2}\right] \cdot \cos\left(\frac{2\pi}{\lambda}(\bar{x}+vt) + \varphi\right) \frac{1}{\sqrt{2\pi\tau}} \exp\left[-\frac{(t-\mu)^2}{2\tau^2}\right] h(t) \quad (1)$$

式中: $\bar{x}=x\cos(\theta)+y\sin(\theta)$, $\bar{y}=-x\sin(\theta)+y\cos(\theta)$, $(x,y,t) \in \Omega \subset R^3$,且满足:

$$h(t) = \begin{cases} 1, & \text{if } t \geq 0 \\ 0, & \text{if } t < 0 \end{cases} \quad (2)$$

视觉输入刺激的强度分布函数为 $I(x,y,t)$,通过一个以 $g_{v,\theta,\varphi}(x,y,t)$ 线性滤波器,得到线性响应为:

$$\gamma_{v,\theta,\varphi}(x,y,t) = I(x,y,t)g_{v,\theta,\varphi}(x,y,t) \quad (3)$$

之后将响应值进行非线性的半平方处理。由于细胞的响应值是正值,而式(3)计算出的值有正有负,在不改变 V1 细胞方向调谐特性的前提下,采用半平方操作对响应进行半波整流^[10],即对于 t 时刻中心处于坐标 (x,y) 位置的简单细胞而言,其响应由上述线性响应进行半波整流操作得到:

$$s_{v,\theta,\varphi}^2(x,y,t) = \max[0, \gamma_{v,\theta,\varphi}(x,y,t)]^2 \quad (4)$$

复杂细胞与简单细胞相比,它的感受野尺寸要比简单细胞的大,并且复杂细胞对刺激方位也有着相似的偏好,根据 Hubel 等提出的神经元细胞的感受野等级假说,复杂细胞的感受野是由简单细胞的感受野以某种特定方式整合而成,因此复杂细胞接受简单细胞的输出信息作为输入。本文采用能量模型^[16],即复杂细胞的响应 $E_{v,\theta}(x,y,t)$ 可通过 2 个相位相互正交的简单细胞响应的平方和表示,即:

$$E_{v,\theta}(x,y,t) = \sqrt{s_{v,\theta,0}^2(x,y,t) + s_{v,\theta,\pi/2}^2(x,y,t)} \quad (5)$$

在整个模型中, Ω 表示全体细胞感受野的作用范围,即空间范围的“定义域”,其以 $(0,0,0)$ 为中心, γ 表征空间域中椭圆状高斯包络的纵横比。标准偏差 σ 决定了单个细胞感受野的大小。参数 v_c 是空间高斯包络线的中心沿 x 轴移动的速度;当 $v_c = 0$ 时,高斯包络线的中心是静止的,此时刺激 $I(x,y,t)$ 为静态图像。参数 λ 是空间波长, $1/\lambda$ 的余弦因子的空间频率。 $\theta \in [0, 2\pi]$ 为滤波器的偏好方向。参数 v 是余弦因子的相速度,决定了偏好运动速度。

式中期望为 μ , 标准差为 τ 的高斯分布函数, 表征感受野的兴奋或抑制程度随的时变特性; 式(2)中的阶跃函数 $h(t)$ 则确保了模型的因果特性。

在参数设置上, 本文模型与已有模型和生物实验数据保持最大程度的一致性。取 $\gamma=0.5$, 表征细胞感受野沿 y 轴方向延展, σ/γ 的值确定了感受野的空间带宽以及对刺激兴奋和抑制的程度比。其表达式为:

$$\frac{\sigma}{\gamma} = \frac{1}{\pi} \sqrt{\frac{\ln 2 \cdot 2^b + 1}{2 \cdot 2^b - 1}} \quad (6)$$

式中: b 为半响应空间频率带宽。选取 $\sigma/\lambda=0.56$, 由此确定 b 的取值。通过设置 $v_c=0$ 和 $v_c=v$ 可以分别得到一个具有静止包络或移动包络的线性滤波器。同时, 空间偏好波长 λ 和偏好速度 v 满足 $\lambda=\lambda_0 \sqrt{1+v^2}$, 其中 λ_0 为滤波器的时空周期, $\lambda_0=2$, 是数字图像序列最小的时空周期。以上确保了可以得到时空周期为 λ_0 的滤波器组, 同时也表明, 具有更高偏好速度的滤波器意味着更大尺度的细胞感受野。参数 v 和 θ 分别表示滤波器的偏好速度和偏好方向。由 $\lambda=\lambda_0 \sqrt{1+v^2}$ 和 $\sigma/\lambda=0.56$ 得, v 的取值决定偏好波长, 进而决定细胞感受野的尺度; θ 描述了滤波器的空间偏好方向。

2 纹理特征提取

纹理存在于各类事物的表面, 是一种低层次的视觉特征, 是识别物体的关键特征之一^[17]。虽然现有的图象图形学领域并未对纹理进行统一公认的定义, 但由于其包含丰富多样的细节信息, 使得各种纹理特征分类与提取方法层出不穷^[18]。有效的纹理提取, 可以为医学图像判读与诊断提供科学的参考, 为无人飞行器的自主导航与侦察监视提供有价值的资料, 为后续进一步的图像处理奠定基础。当前, 经典的纹理特征提取方法主要有结构法、模型法、频谱法和统计法。虽然这些方法在一定条件下有着良好的提取效果, 但在适用性上仍有较大不足。它们大多基于某些特定图像本身的纹理结构特点, 通过设计特定的算法来解决该条件下的纹理提取问题, 与人类视觉系统的普适性有较大差距, 在提取效果上也远远达不到人类视觉系统的水平。大量的生物实验发现, 人脑视皮层 V1 区对静态边缘、纹理等细节信息有着较强的敏感性, 可以高效鲁棒地提取各类图像的纹理信息。

鉴于此, 为验证模型的有效性, 以自然图像作为式(3)中的视觉输入刺激, 取 $v_c=0$, 并对不同参数下的纹理提取效果进行分析。在输入图像的选择

上, 同文献[14]一样, 选取纹理细节明显的 Middlebury 数据库中的典型图像。基于控制变量思维, 固定其他参数, 将 θ 在 $[0, 2\pi]$ 范围内进行 n 等分, $n \in \{4, 8, 12, 16, \dots\}$, 分别代表不同数量不同偏好方向的滤波器组。以 Schefflera 为例, 不同情况下纹理提取效果见图 2。

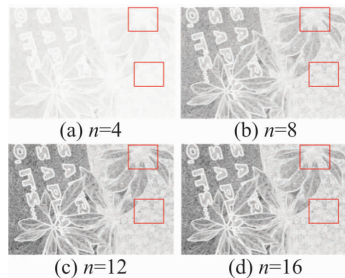


图 2 模型在不同条件下的纹理提取效果

Fig. 2 Texture extraction under different conditions

从图 2 可以看出, 当 $n=4$ 时, 模型未能提取完整的纹理, 丢失较多细节信息, 较难从中识别物体。这说明仅仅依靠 4 种偏好方向的细胞响应整合, 很难得到完整的视觉认知; 当 $n=8$ 时, 提取的细节信息更加丰富, 已基本可以辨识出物体的基本纹理结构, 但仍有个别区域细节信息丢失较多; 当 $n=12$ 或 $n=16$ 时, 模型可以提取出较为完整的纹理细节信息, 且纹理提取性能趋于一致。为在提取出完整纹理信息的同时减小计算量, 本文模型最终取 $n=12$ 。

为进一步检验本文模型的性能, 选取 4 幅纹理信息丰富的图像 Schefflera、Teddy、Backyard、Urban, 与文献[18]的模型进行对比, 结果见图 3。

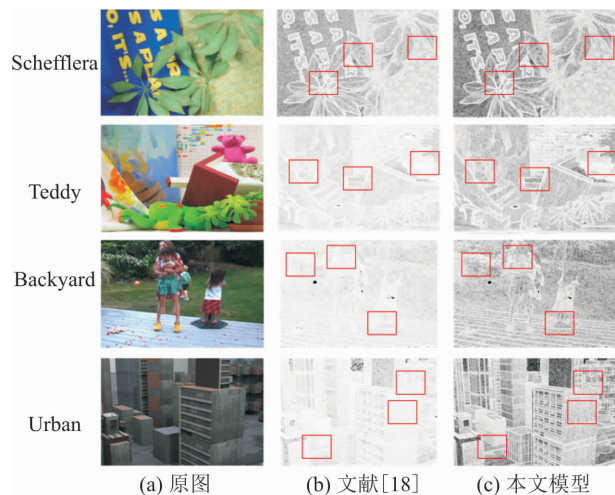


图 3 纹理提取效果对比

Fig. 3 Contrast of texture extraction effects

从图 3 可以看出, 本文模型较文献[18]而言, 能够提取更加丰富的纹理特征, 感知更全面的细节信息。本文模型采用三维时空滤波器和能量模型, 较文献[18]而言, 更加接近 V1 细胞感受野特性, 故可以提取更加丰富的细节信息。

3 视频序列的运动特征提取

作为视皮层的初始区域, V1 可以对运动特征进行有效的预提取。V1 的视野主要集中于局部区域, 可感知局部运动信息, 后经 MT 整合, 可获取全局运动信息^[19]。Movshon^[20]和 Rust^[11]等人通过实验对 V1 的运动感知特性进行研究, 发现 V1 有效提取光栅和光格(由 2 个方向的光栅组成)刺激的运动

信息。整体上对光栅的响应呈单峰特性, 对光格的响应则呈双峰特性(见图 4)。为检验本文模型对运动信息的感知特性, 我们与文献[11]和[13]的模型进行对比, 且参数设置与视觉刺激严格与上述文献保持一致。选取偏好方向为零、偏好速率为 1(像素/帧)的 V1 细胞模型, 即取 $\theta=0, v=1$ 。如图 4(a)所示, 光栅和光格刺激的运动速率为 1(像素/帧), 在空间均匀分布的 36 个方向运动, 箭头指向仅表示其中之一。图 4 即为各模型响应极坐标图。

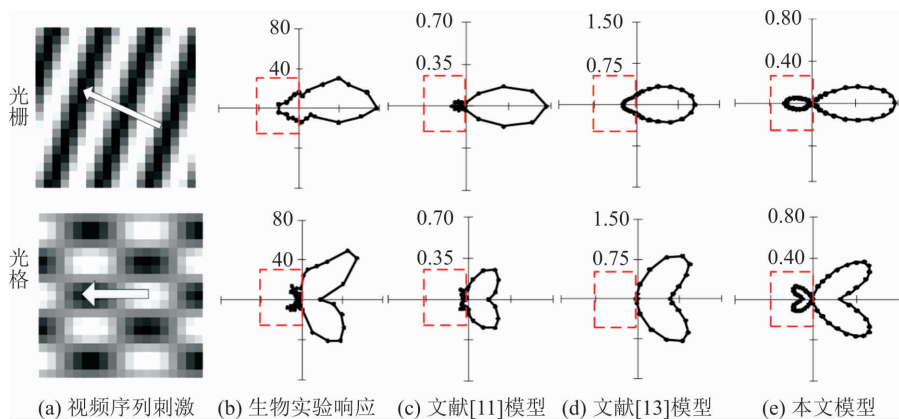


图 4 本文模型与各模型在运动感知提取性能对比

Fig. 4 Contrast of different models on motion perception

对比图 4 的仿真结果, 可以得到文献[11]和[13]的模型可以较好地模拟真实 V1 复杂细胞对 2 类刺激(光栅和光格)的响应特性。整体上, 对光栅和光格刺激分别呈单峰和双峰特性, 吻合了生物现象。然而, 当视觉刺激为光栅时, 上述 2 个文献的模型在非偏好方向的模拟输出响应较小, 几乎为零, 与生物实际存在较大差别。而本文模型在非偏好方向的输出响应与实验数据吻合度较高。当刺激为光格时, 本文模型在非偏好方向均存在响应, 整体响应曲线较好地拟合了生物实验数据, 同时, 模型对光格刺激包含的 2 组光栅具有较好的区分能力。本文模型较上述文献的 2 个模型而言, 在运动感知方面与生物实验数据具有更高的一致性, 更加真实“类脑”。

型能够像人脑视觉系统一样, 捕捉到“最感兴趣”的有用信息, 在运动感知方面具有良好的鲁棒性。

上述实验所用刺激均属无噪视频序列, 而我们的视觉系统常能从含噪刺激中提取有价值的信息。为进一步检验模型对于噪声的鲁棒性, 选取以 $\pi/2$ 为运动方向, 运动速度为 1 像素/帧的条状刺激为视觉输入, 加以随机高斯噪声, 对模型的输出响应进行分析。当加入均值为零、方差为 0.1 随机高斯噪声后, 模型的输出响应见图 5。

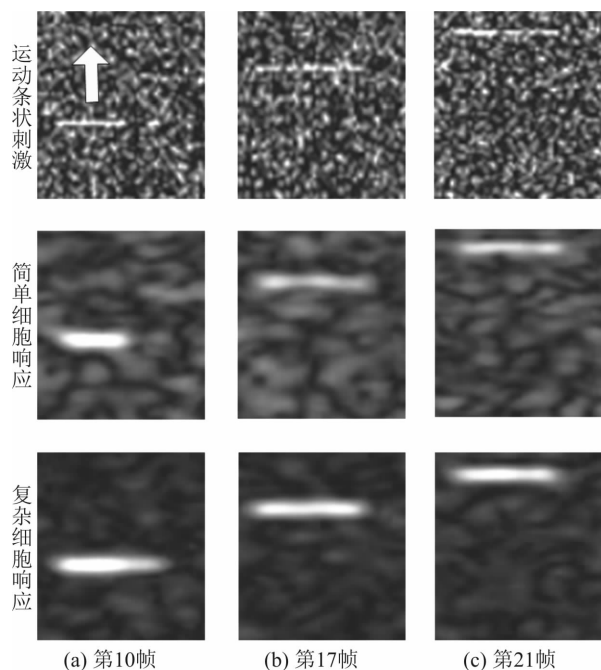


图 5 模型对加噪条状运动刺激的响应

Fig. 5 Responses towards moving bars with noise

从图 5 可以看出, 加入随机高斯噪声以后, 简单细胞和复杂细胞均在一定程度上抑制了噪声。其中, 复杂细胞对噪声几乎“无兴趣”, 仅存在细微响应。通过响应的数据分析, 噪声强度被相对抑制了 70% 左右; 同样, 在加入椒盐噪声后, 所得结论相同。本文模

4 结语

通过深入剖析 V1 的三维感受野特性, 提出了一种基于时空能量模型的 V1 感知模型, 有效地将静态

的细节纹理特征提取与动态的运动特征感知融合于同一模型框架的处理,较好地模拟了真实 V1 细胞的信息提取能力。通过三维 Gabor 时空滤波器和半波整流操作模拟 V1 简单细胞,再通过能量模型整合简单细胞的响应,以实现复杂细胞的特征感知机制。在纹理提取与动态感知两方面分别与现有模型作对比分析,有力验证了所提时空能量模型的有效性。

本文模型在充分考虑静态信息提取的基础上,同时探索了神经细胞模型在局部运动感知上的应用。相较于 V1 细胞,腹侧通路 V4 具有更强大的静态细节信息提取能力;背侧通路 MT 细胞视野更开阔,感受到的运动信息更全面。下一步将拓延模型,将模型拓展至 V4 及 MT 区域,并探索双通路模型的交叉映射机制,提高对视觉目标动态特征的深层次表达能力。

参考文献(References):

- [1] JORGENSON L A, NEWSOME W T, ANDERSON D J, et al. The BRAIN Initiative: Developing Technology to Catalyse Neuroscience Discovery[J]. *Philosophical Transactions of The Royal Society B, Biological Sciences*, 2015, 370(1668): 20140164.
- [2] POO M, DU J, IP N Y, et al. China Brain Project: Basic Neuroscience, Brain Diseases, and Brain-Inspired Computing[J]. *Neuron*, 2016, 92(3): 591-596.
- [3] 寿天德. 视觉的神经基础[J]. *自然杂志*, 2015, 37(1): 17-25.
SHOU T D. Neuronal Basis of Vision[J]. *Chinese Journal of Nature*, 2015, 37(1): 17-25. (in Chinese)
- [4] 郑南宁. 计算机视觉与模式识别[M]. 北京:国防工业出版社, 1998:49-73.
ZHENG N N. Computer Vision and Pattern Recognition[M]. Beijing: National Defense Industry Press, 1998:49-73. (in Chinese)
- [5] 辛鹏,许悦雷,马时平,等. 仿视皮层 V1 特性的 SAR 图像边缘检测[J]. *空军工程大学学报(自然科学版)*, 2017, 18(2): 49-53.
XIN P, XU Y L, MA S P, et al. A Method of SAR Image Edge Detection by Modeling the Properties of V1 in Cortex [J]. *Journal of Air Force Engineering University(Natural Science Edition)*, 2017, 18(2): 49-1653. (in Chinese)
- [6] HUBEL D H, WIESEL T N. Receptive Fields, Binocular Interaction and Functional Architecture in the Cat's Visual Cortex[J]. *The Journal of Physiology*, 1962, 160(1): 106-154.
- [7] SAPROO S, SERENCES J T. Attention Improves Communication between V1 and MT [J]. *Journal of Vision*, 2012, 12(9): 126.
- [8] BERGMANN J, GENC E, KOHLER A, et al. Neural Anatomy of Primary Visual Cortex Limits Visual Working Memory[J]. *Cerebral Cortex*, 2016, 26(1): 43-50.
- [9] AKSELROD M, HERZOG M H, OGMEN H, et al. Tracing Path-Guided Apparent Motion in Human Primary Visual Cortex V1[J]. *Scientific Reports*, 2014, 4(1): 6063.
- [10] RIESENHUBER M, POGGIO T. Hierarchical Models of Object Recognition in Cortex[J]. *Nature Neuroscience*, 1999, 2(11): 1019-1025.
- [11] SIMONCELLI E P, HEEGER D J. A Model of Neural Responses in Visual Area MT [J]. *Vision Research*, 1998, 38(5): 743-761.
- [12] RUST N C, MANTE V, SIMONCELLI E P, et al. How MT Cells Analyse the Motion of Visual Patterns [J]. *Nature Neuroscience*, 2006, 9(11): 1421-1431.
- [13] 邹洪中,许悦雷,马时平,等. 基于视皮层 V1 模型的随机点视频序列运动特征提取[J]. *计算机应用*, 2016, 36(6): 1677-1681.
ZOU H Z, XU Y L, MA S P, et al. Motion Feature Extraction of Random-Dot Video Sequences Based on V1 Model of Visual Cortex [J]. *Journal of Computer Applications*, 2016, 36(6): 1677-1681. (in Chinese)
- [14] PRAMOD R T, ARUN S P. Do Computational Models Differ Systematically from Human Object Perception[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE, 2016: 1601-1609.
- [15] WANG Y, CHUA C. Face Recognition from 2D and 3D Images Using 3D Gabor Filters[J]. *Image and Vision Computing*, 2005, 23(11): 1018-1028.
- [16] PETKOV N, SUBRAMANIAN E. Motion Detection, Noise Reduction, Texture Suppression, and Contour Enhancement by Spatiotemporal Gabor Filters with Surround Inhibition[J]. *Biological Cybernetics*, 2007, 97(5): 423-439.
- [17] HU Y, LIANG Z, SONG B, et al. Texture Feature Extraction and Analysis for Polyp Differentiation via Computed Tomography Colonography[J]. *IEEE Transactions on Medical Imaging*, 2016, 35(6): 1522-1531.
- [18] 邹洪中,许悦雷,马时平,等. 仿视皮层 V1 结构的纹理特征提取[J]. *西安电子科技大学学报*, 2017, 44(3): 96-100, 107.
ZOU H Z, XU Y L, MA S P, et al. Texture Extraction Simulating the Structure of V1 in the Visual Cortex[J]. *Journal of Xidian University*, 2017, 44(3): 96-100, 107. (in Chinese)
- [19] WU X F, WANG X, LU H Z, et al. Motion Feature Extraction for Stepped Frequency Radar Based on Hough Transform[J]. *IET Radar, Sonar and Navigation*, 2010, 4(1): 17-27.
- [20] ELSHAMAYLEH Y, MOVSHON J A. Neuronal Responses to Texture-Defined Form in Macaque Visual Area V2[J]. *The Journal of Neuroscience*, 2011, 31(23): 8543-8555.