

基于初级视通路视觉感知机制的轮廓检测方法

张明琦¹ 范影乐¹ 武薇¹

摘要 考虑到初级视通路中视觉信息传递和处理过程中的特点, 本文提出了一种基于视觉感知机制的轮廓检测新方法. 构建视觉信息局部细节检测与整体轮廓感知的不同路径. 利用高斯导函数提取初级轮廓响应; 构建神经网络, 利用时空编码提高主体轮廓对比度; 然后, 利用非经典感受野的侧抑制作用抑制纹理背景; 另外, 针对轮廓信息强化以及检测鲁棒性的要求, 在视辐射区提出了一种信息冗余度增强编码机制; 最后, 将初级轮廓直接前馈至初级视皮层, 以达到轮廓响应的快速调节和完整性融合. 以 RuG40 图库为实验对象, 经过非极大值抑制和阈值处理, 得到的轮廓二值图与基准轮廓图比较, 在整个数据集中的最优平均 P 指标和每张图的最优平均 P 指标分别为 0.48 和 0.55, 并且 FPS 达到了 1/2. 结果表明本文方法能有效突出主体轮廓并抑制纹理背景, 为后续图像理解和分析提供了一种新的思路.

关键词 轮廓检测, 时空编码, 信息冗余度增强编码, 前馈

引用格式 张明琦, 范影乐, 武薇. 基于初级视通路视觉感知机制的轮廓检测方法. 自动化学报, 2020, 46(2): 264–273

DOI 10.16383/j.aas.2018.c170688

A Contour Detection Method Based on Visual Perception Mechanism in Primary Visual Pathway

ZHANG Ming-Qi¹ FAN Ying-Le¹ WU Wei¹

Abstract A new method of contour detection is proposed in this paper according to the characteristics of visual information transmission and processing which is based on visual perception mechanism in human primary visual pathway. Firstly, the spatio-temporal coding of neural information is realized on the basis of the Gaussian derivative function, and then the texture and background are suppressed by the non-classical receptive field. In addition, we propose a neural coding that can strengthen the redundancy of visual information in the optic radiation area to enhance the contour information and robustness of detecting. Finally, primary contour is fed forward to the primary visual cortex to achieve rapid adjustment of the contour response. In this paper, the RuG40 library is taken for processing, acquiring the contour detecting result through the non-maximal suppression and hysteresis threshold, and mean values P of optimal parameters for whole dataset and each image are 0.48 and 0.55, respectively, compared with the ground truth. In addition, the FPS of algorithm reaches 1/2. The results indicate the new method in this paper can effectively highlight the principal contour and suppress texture region which will provide a new idea for subsequent image comprehension and analysis.

Key words Contour detection, spatio-temporal coding, enhancement coding of information redundancy, feed forward

Citation Zhang Ming-Qi, Fan Ying-Le, Wu Wei. A contour detection method based on visual perception mechanism in primary visual pathway. *Acta Automatica Sinica*, 2020, 46(2): 264–273

轮廓特征是图像目标方向性的一种低维度表征, 如何在大量冗余背景中实现低对比度轮廓细节的快速准确提取, 对于降低后续图像识别或理解的复杂度, 并进一步提高系统识别性能具有重要的研究意义^[1].

传统轮廓检测方法主要包括有基于局部邻域梯度信息的数学运算^[2–3]、全局与局部信息的相关

性分析^[4] 以及相位一致性的提取方法^[5], 它们大多是从轮廓像素灰度跳变的数学特性上来考虑的. 随着视觉神经实验的发展和计算能力的提高, 人们开始逐渐了解并尝试应用视觉系统在复杂环境中快速感知目标的内在机制, 例如有研究通过高斯导函数来表征经典感受野机制^[6]; 有研究建立了外膝体 (Lateral geniculate nucleus, LGN) 感受野组合模型^[7], 在纹理背景边缘的抑制中强化了其与初级视皮层信息流的传递关系; 也有研究利用二维 Gabor 能量函数表征初级视皮层经典感受野轮廓方向的选择性^[8–9], 并引入非经典感受野的侧抑制作用, 分别通过各向异性和各向同性两种抑制作用, 能够在有效减少背景纹理的同时, 凸显主体轮廓. 另外还有融合方位、光强、对比度多种属性后再做周边抑制, 进

收稿日期 2017-12-05 录用日期 2018-05-18
Manuscript received December 5, 2017; accepted May 18, 2018
国家自然科学基金 (61501154) 资助
Supported by National Natural Science Foundation of China (61501154)
本文责任编辑 刘跃虎
Recommended by Associate Editor LIU Yue-Hu
1. 杭州电子科技大学模式识别与图像处理实验室 杭州 310018
1. Laboratory of Pattern Recognition and Image Processing, Hangzhou Dianzi University, Hangzhou 310018

进一步提升轮廓的检测效果^[10]. 还有研究从神经编码的角度出发, 利用多层神经网络中突触的动态可塑性, 实现图像轮廓检测^[11]. 但上述方法通常只关注于某一种特定的视觉机制, 例如分别考虑经典或非经典感受野在视觉感知中的作用, 或者是神经编码方法对于轮廓信息提取和纹理背景抑制的有效性. 而且需要指出的是, 一方面越来越多的实验证实初级视通路上存在着不同的感受野特性, 它们扮演着从局部细节检测到整体轮廓感知的角色; 另一方面如何充分挖掘视觉通路上神经电信号传递和处理能力, 尤其是前馈式的检测方式对于轮廓快速检测的作用, 将有助于视觉系统内在机制理解及其在轮廓检测中的具体应用.

因此本文基于视觉信息传递的神经元电生理模型, 经典感受野的方向选择特性, 以及非经典感受野的侧抑制作用, 重点在初级视通路中采用了串行式的神经编码方式, 利用时空编码来细化纹理背景突出主体轮廓, 考虑到初级视皮层处理信息的能力远远超过外膝体, 因此在其前级进行图像信息的冗余度增强编码, 强化轮廓信息. 另外利用跨层级的初级轮廓响应前馈机制, 在初级视皮层上实现轮廓信息的快速响应和完整性融合.

1 材料和方法

早期视通路中的视觉处理对物体轮廓信息的获取以及对高级视皮层感知的形成有着重要的作用^[12]. 从 LGN 到初级视皮层的早期视通路上分布

着不同的神经元感受野^[13], 对前级神经元的刺激响应以及对后级神经元的信息传递有着不同的编码方式. 本文重点考虑视通路中不同的神经编码, 利用串行的组合神经编码方式提取特征, 并结合非经典感受野的侧抑制作用对特征图进行纹理背景抑制, 通过信息冗余度增强编码丰富特征. 最后基于整体轮廓的前馈机制, 实现对局部细节的完整性融合, 以此快速提取图像轮廓. 具体检测框架如图 1 所示.

1.1 基于多方向响应的初级轮廓感知

考虑到视通路对特定方向的光刺激有高度选择性, 例如在 LGN 中就已存在具有方向选择性的细胞, 而有些生物包括小鼠、家兔等甚至在视网膜中就存在方向选择性神经节细胞 (Direction-selective ganglion cell, DSGC)^[14]. 因此本文先引入高斯导函数来模拟视通路对图像多方位上的激励响应, 如式 (1) 所示:

$$\begin{cases} e_i(x, y; \theta_i, \sigma) = \left| I(x, y) * \frac{\partial g(\tilde{x}, \tilde{y}; \theta, \sigma)}{\partial \tilde{x}} \right| \\ g(\tilde{x}, \tilde{y}; \theta, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\tilde{x}^2 + \gamma^2 \tilde{y}^2}{2\sigma^2}\right) \end{cases} \quad (1)$$

式中, $I(x, y)$ 表示输入图像, “*” 表示卷积运算, $(\tilde{x}, \tilde{y}) = (x \cos(\theta) + y \sin(\theta), -x \sin(\theta) + y \cos(\theta))$, θ 表示不同的选择方位, 标准差 σ 和空间大小比率 γ 分别决定了非同心圆感受野的尺寸和椭圆率.

后采用 Winner-take-all 策略, 选取 N_θ 个方位中的响应极大值形成初级轮廓响应, 如式 (2) 所示:

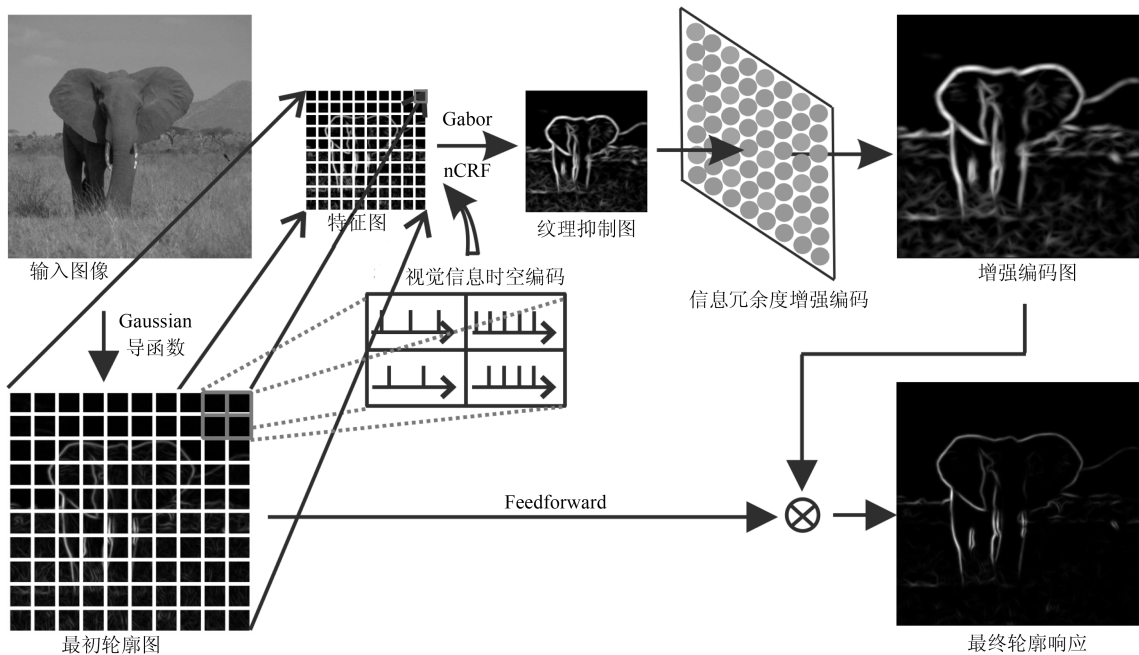


图 1 轮廓检测框架

Fig. 1 The framework of contour detection

$$E(x, y; \sigma) = \max(e_i(x, y; \theta_i, \sigma) | i = 1, 2, \dots, N_\theta) \quad (2)$$

1.2 视觉信息流的时空编码

通常在以视觉信息流传递和处理为模型的图像处理任务中只选取某一种神经编码方式^[15],但在初级视通路的不同节点或环节中,可能同时存在着多种神经信息流的编码方式,例如有发放率编码、稀疏编码、时间编码、相位一致性编码以及群体编码等^[16].不同的神经编码方式可能在激励信息的特征提取、去冗余化,以及在提高编码效率等功能上,发挥着各自不可或缺的作用.文献[17]中利用单峰时间编码将输入激励编码为时空峰值模式并通过时间学习规则来区分.文献[18]提出了时空记忆模型(Spatio-temporal memory, STM),利用时空编码和STDP(Spike-timing-dependent plasticity)学习法则来模拟记忆形成的动态过程.

但上述方法侧重于对神经编码和学习方法的探索,由于模型的复杂度,加大了算法在时间成本上的开销,无法满足视觉信息处理过程中快速响应的要求.本文考虑了在视通路 LGN 区,神经数量减少的空间特性以及神经元放电过程的时间特性,提出了一种对视觉信息的时空编码方式.不同于文献[17-18]中的时空编码,本文选择在计算简单且应用广泛的 LIF(Leaky integrate-and-fire)神经元模型^[19]的基础上,统计在一定时间内每个神经元脉冲发放的个数作为响应值,该编码方式满足了神经元编码模式的非线性以及神经系统在轮廓检测任务中的快速性.LIF神经元模型如式(3)所示,同时引入绝对不应期:

$$c_m \frac{dv}{dt} = -g_l v + I_{in} \quad (3)$$

式中, v 是膜电压, c_m 和 g_l 分别对应于神经元的膜电容和漏电导, I_{in} 对应于前述的初级轮廓响应 $E(x, y; \sigma)$.当膜电压 v 在电流 I_{in} 的充电作用下超过发放阈值时,膜电压将表现为一个脉冲上升过程,瞬间达到峰值又回落至静息电位,同时进入绝对不应期,在此期间将不再对任何激励进行响应.根据式(3)对初级轮廓响应进行编码,从而得到脉冲输出图像 I_{LIF} .

在上述过程中,普遍的观点认为应该存在着某种对视觉信息的特征提取和去冗余的编码机制^[20].因此本文以一定大小的局部窗口表示神经元集群,根据神经元的竞争机制,将窗口内脉冲响应最大的神经元作为输出,并以窗口大小作为移动步长,达到视觉信息在空间上的稀疏性表达,如式(4)所示:

$$E_{ft}(x, y) = \max \left(\bigcup_{1 \leq w, h \leq step} I_{LIF}((x-1) \times step + w, (y-1) \times step + h) \right) \quad (4)$$

其中, E_{ft} 为通过时空编码后得到的特征图像, $step$ 表示移动步长, \cup 表示局部窗口内神经元的集合.

1.3 背景信息的各向同性抑制

LGN 细胞感受野的周边抑制能力比神经节细胞感受野更强,再加上经典感受野外的大外周抑制区对其的调控能力,使 LGN 细胞对提高对比敏感度以及抑制周围噪声方面有很强的作用^[21-22].因此在轮廓提取任务中能够提高主体轮廓的对比度以及抑制周边细小边缘和纹理噪声.本文引入二维 Gabor 能量函数^[8]来模拟 LGN 细胞对轮廓进行处理,如式(5)和(6)所示:

$$G_{\lambda, \sigma, \theta}(x, y) = \sqrt{R_{\lambda, \sigma, \theta, 0}^2(x, y) + R_{\lambda, \sigma, \theta, -\frac{\pi}{2}}^2(x, y)} \quad (5)$$

$$R_{\lambda, \sigma, \theta, \varphi}(x, y) = \exp\left(-\frac{\tilde{x}^2 + \gamma^2 \tilde{y}^2}{2\sigma^2}\right) \cos\left(2\pi \frac{\tilde{x}}{\lambda} + \varphi\right) * E_{ft}(x, y) \quad (6)$$

式中, $(\tilde{x}, \tilde{y}) = (x \cos(\theta) + y \sin(\theta), -x \sin(\theta) + y \cos(\theta))$, γ 代表了感受野的椭圆率, σ 决定了感受野的大小, $1/\lambda$ 表示 cosine 因子的空间频率, θ 决定 Gabor 滤波器最优选择方位, $\varphi = 0$ 和 $\varphi = \pi$ 表示 Gabor 滤波器是对称的, $\varphi = -\pi/2$ 和 $\varphi = \pi/2$ 表示 Gabor 滤波器是非对称的, R 表示图像的边缘响应.

利用各向同性方法抑制周边纹理背景边缘,首先计算 Gabor 能量函数在各个方位上的最大值响应:

$$\hat{G}_{\lambda, \sigma}(x, y) = \max(G_{\lambda, \sigma, \theta_i}(x, y) | i = 1, 2, \dots, N_\theta) \quad (7)$$

对于抑制项 $b_{\lambda, \sigma}(x, y)$ 及抑制权重 $w_\sigma(x, y)$,其表达式为(8)所示:

$$\begin{cases} b_{\lambda, \sigma}(x, y) = \hat{G}_{\lambda, \sigma}(x, y) * w_\sigma(x, y) \\ w_\sigma(x, y) = \frac{1}{\|H(\text{DOG}_\sigma)\|_1} H(\text{DOG}_\sigma(x, y)) \\ \text{DOG}_\sigma(x, y) = \frac{1}{2\pi(4\sigma^2)} \exp\left(-\frac{x^2 + y^2}{2(4\sigma^2)}\right) - \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \\ H(z) = \begin{cases} 0, & z < 0 \\ z, & z \geq 0 \end{cases} \end{cases} \quad (8)$$

式中, 抑制项 $b_{\lambda,\sigma}(x,y)$ 通过最大 Gabor 能量函数与抑制权重 $w_{\sigma}(x,y)$ 的卷积得到, $\|\cdot\|_1$ 表示 L_1 范数, DOG_{σ} 表示高斯差函数, $H(z)$ 函数对应于半波整流作用. 最后得到各向同性抑制后的轮廓响应为:

$$f_{\lambda,\sigma}^{\alpha}(x,y) = H(\hat{G}_{\lambda,\sigma}(x,y) - \alpha b_{\lambda,\sigma}(x,y)) \quad (9)$$

式中, H 的定义如式 (8), α 为非经典感受野抑制作用的强度系数.

1.4 视辐射区信息冗余度增强编码

初级视皮层内分布着许多简单细胞、复杂细胞及超复杂细胞, 其感受野不仅能感知物体形状, 还能对颜色、运动和深度视觉信息进行处理, 通过各种细胞的协同作用, 使初级视皮层对视觉信息流的处理能力大大加强^[23]. 因此在初级视皮层内的图像信息越丰富, 视觉系统对图像的分析理解就越完整. 而对生理解剖研究发现, 视辐射区神经元数目急剧增加, 能为初级视皮层提供更多的特征信息. 因此, 本文提出了视辐射区信息冗余度增强编码, 通过提高神经元的数量来增加特征信息的冗余度, 强化轮廓信息. 将提取的轮廓信息通过冗余度增强编码, 使边缘像素点的对比度增大, 轮廓加粗高亮, 提高了初级视皮层对轮廓提取的容错率及鲁棒性.

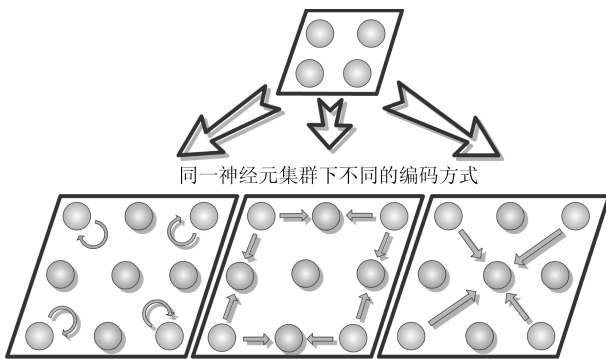


图2 三种局部信息冗余度增强编码

Fig. 2 Three local enhancement codings of information redundancy

考虑到突触连接与神经元之间的距离关系和各神经元的位置分布, 采用了三种增强编码方式. 该编码方式在一定程度上模拟了局部神经元之间的竞争机制, 并且探究了突触信息传递的非线性过程, 通过给激励大的神经元分配更大的权重系数, 使兴奋的神经元在小范围内抑制周边的神经元响应. 图2为本文所定义的三种局部增强编码方式, 图中小箭头代表神经元的激励方向. 例如左边的神经元由其单独激励, 而中间、右边中的神经元分别由其两边或四周的神经元共同作用, 权重系数的大小如式 (10) 所

示:

$$\alpha(x,y;\theta) = \frac{I(x,y;\theta,u)}{\sum_{\theta \in \Omega} I(x,y;\theta,u)} \quad (10)$$

$$n \in \{1, 2, 4\}, u \in \{0, 1\}$$

式中, $I(x,y;\theta,u)$ 表示神经元的激励大小. θ 表示不同的突触连接方向, Ω 表示神经元的单邻域, 4邻域以及 D 邻域范围, 分别对应于图2中的左、中、右三种模式. n 和 u 表示在位于 (x,y) 位置周围与其共同作用的神经元个数以及偏移量. 以 $f_{\lambda,\sigma}^{\alpha}(x,y)$ 中各像素的值作为编码前神经元的激励 I , 编码后各个神经元的响应通过式 (11) 得出:

$$\tilde{I}(i,j) = \sum_{\theta \in \Omega} \alpha(x,y;\theta) I(x,y;\theta,u) \quad (11)$$

经过信息冗余度增强编码后, 图像在空间尺寸上和初级轮廓提取结果保持一致.

1.5 视通路多路径前馈融合机制

传统基于视觉机制的信息流处理方法或应用, 通常忽略了系统响应快速性的要求, 只从表征激励与响应之间的关联性来探讨各种神经编码的可行性. 必须要考虑到的是, 基于统计学意义上的编码方式如何能够满足对于响应快速性的需求. 神经电生理实验和计算仿真结果均表明, 由于神经元产生脉冲的过程涉及到细胞膜上离子通道的开合, 浓度差诱导的离子跨膜运动、突触延迟以及脉冲发放存在着不应期等因素, 激活状态下的神经元脉冲发放频率大致在数十 Hz 左右, 因此如果说前后级神经元的关联是依赖于脉冲发放时间、相位或频率的统计特性, 那么显然将会带来信息传递效率较低的问题, 根本无法与生物视觉系统的高效性相匹配.

目前视觉生理或心理实验, 均从不同侧面验证了信息传递上的视通路多路径融合, 传统基于串级神经编码的传递路径可能对应于视觉信息的精细化局部感知, 而视通路上跨层级的传递路径则有利于实现视觉信息的粗略化整体检测. 因此本文提出一种基于视觉信息流多路径融合的前馈机制轮廓检测方法. 对于提取的初级轮廓图, 一方面将在初级视通路中进行前述的串级神经编码; 而另一方面将通过前馈机制, 直接传递到初级视通路的末端, 有助于提高视觉系统对于整体轮廓感知的快速性, 并对整体细节的准确感知具有一种完整性融合作用. 最终的轮廓响应如式 (12) 所示:

$$R(x,y) = E(x,y;\sigma) \tilde{I}(x,y) \quad (12)$$

式中 $E(x,y;\sigma)$ 表示初级轮廓图, $\tilde{I}(x,y)$ 表示信息冗余度增强后的编码图, $R(x,y)$ 表示最终在初级视

皮层形成的轮廓响应, 通过像素点乘方式来达到图像的快速融合.

2 结果

图3为视通路神经编码及前馈融合示意图. 输入图像3(a)经过多方向的高斯导函数初步提取轮廓信息, 得到轮廓响应图3(b), 可以看到高斯导函数提取的轮廓虽然细节丰富, 但存在较多纹理背景边缘. 之后根据神经元的空间特性以及电生理特性, 取与之对应的时空编码来提取初级轮廓的有效特征信息, 得到图3(c), 相对于图3(b), 细化了纹理背景边缘, 提高了主体轮廓对比度. 根据LGN非经典感受野抑制作用, 利用Gabor能量函数及各向同性的方法对纹理进行抑制, 得到图3(d), 可以明显看出图像主体轮廓呈高亮状态, 且边缘部分加粗. 接着经过信息冗余度增强编码后得到图3(e), 将图像尺寸放大的同时强化图像轮廓信息并传递给处理能力更强的初级视皮层. 最后引入另一条视通路前馈路径, 快速调节图3(e)从而得到最终轮廓响应图3(f), 可以清楚地看到与图3(b)相比, 在保留主体轮廓的同时, 纹理及背景被大量的消除, 轮廓提取的效果明显.

本文使用RuG40自然场景图像数据库作为实验用图, 包括了40幅自然图像, 每张图像都有一张手工描绘的基准轮廓图与之对应. 不失一般性, 随机选取部分图片作为实验对象, 如图4第一行所示. 最后对整个数据库进行定量评价, 体现本文算法综合性性能.

首先用高斯导函数(Gaussian derivative, GD)^[6]处理得到的轮廓响应图作为与后续算法的对比. 其次选择文献[7]提出的LGN感受野组合模型(Combination of receptive field, CORF), 文献[8]提出的非经典感受野抑制模型(Isotropic surround inhibition, ISO), 文献[10]提出的多特征周边抑制模型(Multiple-cue inhibition, MCI), 文献[11]提出的多层神经元脉冲编码突触抑制模型(Inhibitory synapse spike coding, ISSC), 以及在本文模型(Multiple-neural coding, MNC)中去掉时空编码以及信息冗余度编码后的模型(Non-neural coding, NNC), 来和MNC作比较. 通过与前面5种方法比较来体现MNC的效果性能, 通过与NNC的对比来说明引入神经编码生理机制后对模型的性能提升是否有帮助. 本文中 $\sigma = 2$, $\lambda = 3.5$, $\gamma = 0.5$. 由于LGN非经典感受野的抑制作用大小对后续的

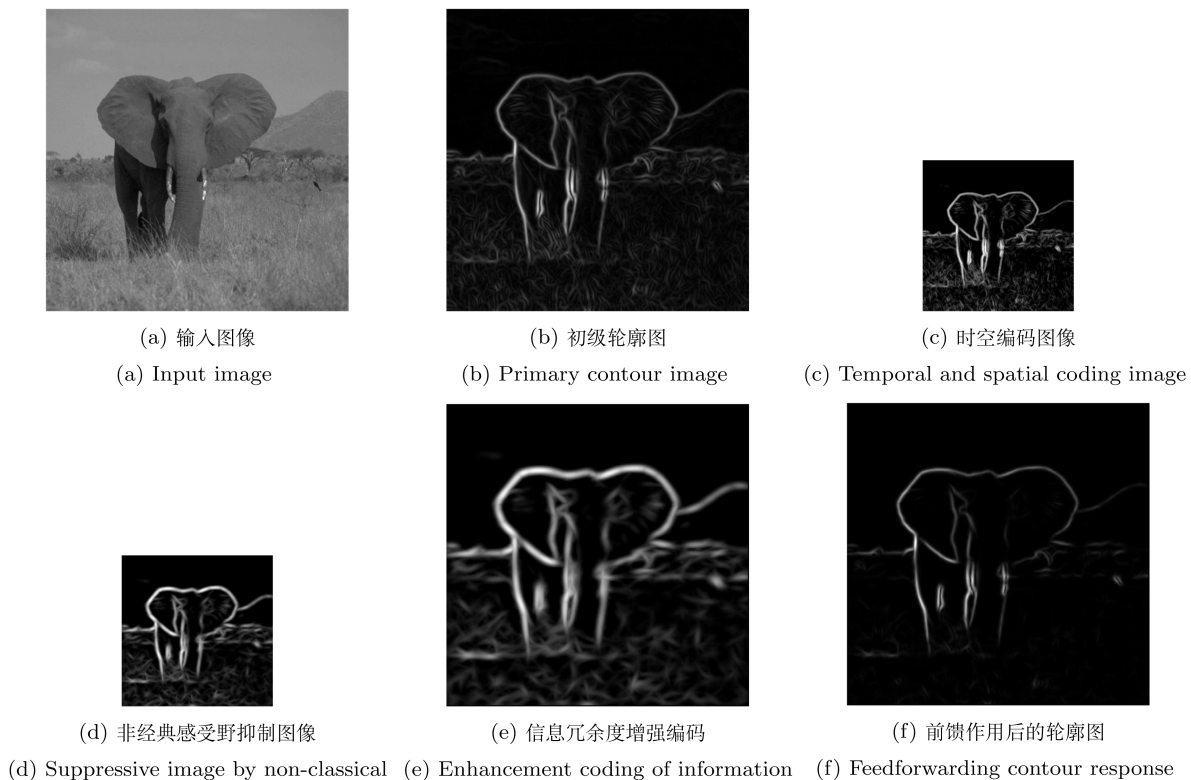


图3 视通路神经编码及前馈融合示意图

Fig. 3 Schematic diagram of visual pathway with neural coding and feedforward fusion



图4 RuG40 图库中的轮廓测试结果 (第1行为用于测试的自然场景图; 第2行为基准轮廓图; 第3行为GD检测结果; 第4行为CORF检测结果; 第5行为ISO检测结果; 第6行为MCI检测结果; 第7行为ISSC检测结果; 第8行为NNC检测结果; 第9行为MNC检测结果)

Fig. 4 Contour test results of the RuG40 gallery (The first line is the input of natural scene images; The second line is the contour baselines; The third line is the results of GD; The fourth line is the results of CORF; The fifth line is the results of ISO; The sixth line is the results of MCI; The seventh line is the results of ISSC; The eighth line is the results of NNC; The ninth line is the results of MNC)

神经编码结果有一定影响,因此选取 10 个不同的抑制系数强度参数 $\alpha = [0.1 : 0.1 : 1]$. 另外为了做定量的指标计算,对最终的轮廓响应图做非极大值抑制和滞后阈值进行二值化处理. 在滞后阈值方法中,取上限阈值因素 $p = [0.1 : 0.1 : 1]$,对于 ISO、MCI、NNC 和 MNC,每张图像可以得到 100 组测试参数,而 GD、CORF、ISSC 方法没有参数 α ,故只有 10 组测试参数.

从图 4 结果中可以看出, GD 算法虽然对轮廓的提取比较完整,但存在相当多的纹理背景边缘. CORF 方法考虑了 LGN 细胞感受野的组合,跟 GD 方法对比,很大程度上减少了纹理背景的干扰,但对轮廓周边的纹理却没有太大效果,有很多密集的毛刺. 而 ISO 方法采用各向同性和各向异性方式模拟非经典感受野的抑制作用,纹理背景抑制明显,却将部分轮廓信息也一同抑制,很难达到理想的平衡. MCI 结合多特征的信息,在抑制纹理的同时也较好的检测出了轮廓. ISSC 利用多层神经元脉冲编码,主体轮廓相对完整,但虚假纹理背景抑制效果不足. 综合来看, MNC 与前面几种模型相比,在保留主体轮廓信息的同时,尽可能地抑制周边纹理背景的影响,图像效果更接近于基准轮廓图,并且与 NNC 相比,可以看到在增加了神经编码后,轮廓的提取效果更加明显.

为了评价检测结果,采用文献 [8] 中提出的方法对结果进行定量评价,检测到正确的像素集为 E ,错误的像素集为 E_{FP} ,漏检的像素集为 E_{FN} 和整体性能指标 P .

$$e_{FP} = \frac{\text{card}(E_{FP})}{\text{card}(E)} \quad (13)$$

$$e_{FN} = \frac{\text{card}(E_{FN})}{\text{card}(E_B)} \quad (14)$$

$$P = \frac{\text{card}(E)}{\text{card}(E) + \text{card}(E_{FP}) + \text{card}(E_{FN})} \quad (15)$$

式中, $\text{card}(C)$ 表示集合 C 中的元素数量. 误检率 e_{FP} 的大小代表了算法对纹理背景的敏感程度,漏检率 e_{FN} 表示算法对主体轮廓提取的能力,漏检率越小,说明轮廓提取的越完整. P 作为整体的性能指标,随着 e_{FP} 和 e_{FN} 减小而增大,表明算法对轮廓的检测效果越好. 通过式 (13)~(15) 对图 4 的二值轮廓图进行定量的指标分析,得到算法最佳结果,如表 1 所示. 显然,表 1 中的计算结果与图 4 的定性描述基本一致. GD 方法的漏检率 e_{FN} 很小,而误检率 e_{FP} 非常大,说明该算法对轮廓的提取完整,但纹理背景抑制差,整体的 P 值最小. CORF 方法误检率显著减少,但始终高于漏检率,说明去除了一些纹理背景边缘,但由于轮廓周边毛刺的存在, P 值的

改善效果有限. ISO 方法的误检率与漏检率在总体上达到了一定的平衡,但实际的效果图由于跟像素点数量也有关系,导致有些图片的误检率和漏检率的差异较大,如 lions 的结果,误检率比漏检率小很多,并从图 4 看, lions 的二值轮廓图的主体轮廓缺失较严重. MCI 由于结合了多种特征的属性,误检率和漏检率在大多数情况下都要比前面几种方法更小,最终的性能指标也仅次于 MNC. ISSC 的漏检率相对较小,说明主体轮廓比较完整,但误检率却比较大,抑制效果不理想. 从 NNC 的性能指标看,虽然比 MCI 模型要差,但结合了神经编码的 MNC 从整体都要优于 MCI 和前面的对比方法.

另外,对各算法的运行速度作比较. 以 FPS 作为评价指标, FPS 指算法每秒所处理的图像数量, FPS 越大,意味着模型运行速度越快,能够满足轮廓检测快速性的需求. 由于算法对每张图的处理速度基本一致,因此本文挑选 buffalo 图像作为实验对象,以各模型在单张图片单组参数的情况下进行对比. 计算机硬件 CPU 为 i7 6700K,软件平台为 MATLAB 2015. 从表 1 中可以发现,本文模型 MNC 的 FPS 虽然比 GD 和 ISO 低很多,但在性能指标上有很大的优势,并且比 MCI 的 FPS 大 10 倍多. 因此,从模型的整体上考虑, MNC 在算法的性能以及运行速度上达到了很好的平衡.

图 5 为各模型在整个 RuG40 数据集中的定量分析,实心柱代表在整个数据集中的最优平均 P 值,条纹柱代表每张图片的最优平均 P 值. 本文方法 MNC 在这两项对比中都表现出了卓越的效果. 除此之外,为验证其一般性,采用统计的方法评价各模型在不同参数下的性能及稳定性. 由于 GD、CORF 的参数少于其他模型,且从图 5 的 P 值来看,效果明显差很多,因此不做统计评价. 盒须图的最顶端表示不同参数下取得的最优值,中间的横线表示不同参数下的 P 中值,并且盒须图越短表明对参数的敏感度越低,稳定性更好. 从统计结果可以看出, MNC 模型在整体上呈现更优的趋势.

3 讨论

从上述实验结果可以看出,本文方法在有效提取轮廓的同时,能较强地去除纹理背景. 对于轮廓细节丰富,纹理背景相似的自然场景图有很强的轮廓提取能力,二值轮廓图更接近于基准轮廓图. CORF、ISO、MCI、ISSC 都从生物视觉机制出发. CORF 将 LGN 感受野在空间上组合起来,并考虑了其与其与初级视皮层之间的联系. 但未处理的原始图像直接作为 LGN 感受野的输入,忽视了 LGN 前级和后级视觉区对于信息的处理,方法过于单一,故导致检测的图像纹理抑制不足轮廓主体周围产生较

表 1 图 4 中不同方法的参数设置, 性能指标及运行速度.

Table 1 Parameters, speed and performance of the different methods in Fig. 4

| 图像 | 算法 | α | p | e_{FP} | e_{FN} | P | FPS |
|------------|------|-------------|-------------|-------------|-------------|-------------|------|
| buffalo | GD | | 0.20 | 1.50 | 0.13 | 0.30 | 4 |
| | CORF | | 0.20 | 0.42 | 0.30 | 0.50 | 3/5 |
| | ISO | 0.80 | 0.20 | 0.25 | 0.38 | 0.55 | 3 |
| | MCI | 0.70 | 0.30 | 0.18 | 0.29 | 0.59 | 1/22 |
| | ISSC | | 0.10 | 0.22 | 0.27 | 0.58 | 1/8 |
| | NNC | 0.20 | 0.10 | 0.22 | 0.32 | 0.54 | 1 |
| | MNC | 0.10 | 0.30 | 0.21 | 0.27 | 0.64 | 1/2 |
| elephant 2 | GD | | 0.10 | 1.29 | 0.25 | 0.39 | |
| | CORF | | 0.20 | 0.32 | 0.29 | 0.52 | |
| | ISO | 1.00 | 0.10 | 0.37 | 0.32 | 0.53 | |
| | MCI | 0.40 | 0.30 | 0.22 | 0.30 | 0.58 | |
| | ISSC | | 0.10 | 0.25 | 0.30 | 0.56 | |
| | NNC | 0.20 | 0.30 | 0.28 | 0.30 | 0.56 | |
| | MNC | 0.10 | 0.20 | 0.13 | 0.36 | 0.62 | |
| golf cart | GD | | 0.25 | 1.33 | 0.18 | 0.45 | |
| | CORF | | 0.20 | 0.41 | 0.29 | 0.50 | |
| | ISO | 0.60 | 0.20 | 0.29 | 0.27 | 0.58 | |
| | MCI | 0.90 | 0.40 | 0.23 | 0.28 | 0.62 | |
| | ISSC | | 0.3 | 0.30 | 0.30 | 0.55 | |
| | NNC | 0.10 | 0.30 | 0.27 | 0.29 | 0.57 | |
| | MNC | 0.10 | 0.40 | 0.26 | 0.24 | 0.62 | |
| hyena | GD | | 0.25 | 1.33 | 0.18 | 0.45 | |
| | CORF | | 0.30 | 0.55 | 0.14 | 0.58 | |
| | ISO | 0.90 | 0.10 | 0.21 | 0.27 | 0.61 | |
| | MCI | 0.60 | 0.50 | 0.19 | 0.22 | 0.65 | |
| | ISSC | | 0.20 | 0.28 | 0.24 | 0.59 | |
| | NNC | 0.20 | 0.20 | 0.30 | 0.24 | 0.59 | |
| | MNC | 0.10 | 0.20 | 0.21 | 0.27 | 0.64 | |
| lions | GD | | 0.20 | 1.77 | 0.15 | 0.25 | |
| | CORF | | 0.30 | 0.67 | 0.32 | 0.45 | |
| | ISO | 0.80 | 0.20 | 0.32 | 0.51 | 0.47 | |
| | MCI | 1.00 | 0.50 | 0.48 | 0.29 | 0.50 | |
| | ISSC | | 0.3 | 0.45 | 0.30 | 0.48 | |
| | NNC | 0.30 | 0.30 | 0.45 | 0.31 | 0.47 | |
| | MNC | 0.70 | 0.40 | 0.45 | 0.28 | 0.51 | |

多的毛刺. ISO 先利用 Gabor 函数高效提取多方位边缘的特性来模拟经典感受野提取图像轮廓, 接着采用各向同性和各向异性的方法模拟大范围非经典感受野的抑制作用, 相比 CORF 能更强地抑制纹理背景边缘, 但主要问题还是部分纹理细节无法与主体轮廓区分, 当增大抑制系数时, 主体轮廓会出现间断现象, 导致漏检率升高. MCI 模型融合了方位, 光强和对比度三个属性来提取图片特征, 并考虑了周边抑制作用来提取轮廓, 虽然检测效果好于其他几种方法, 但主体轮廓的抑制问题还是没能很好解决.

ISSC 从神经编码的角度出发, 考虑了神经元的电生理特性, 利用多层神经网络中突触的动态可塑性, 很好地保留了主体轮廓, 但是并未考虑视通路中感受野细胞的抑制功能, 导致对背景纹理的抑制效果不理想. 上述 4 种方法以感受野的生理作用为基础进行改进, 或是采用了单一的神经编码, 却忽略了视觉信息处理过程中多种神经信息编码与感受野细胞的协同作用, 以及视通路多路径信息流的传递, 故整体检测效果并不满意. 本文方法 MNC 从视通路全局的角度出发, 重点根据视通路神经元的特性提出

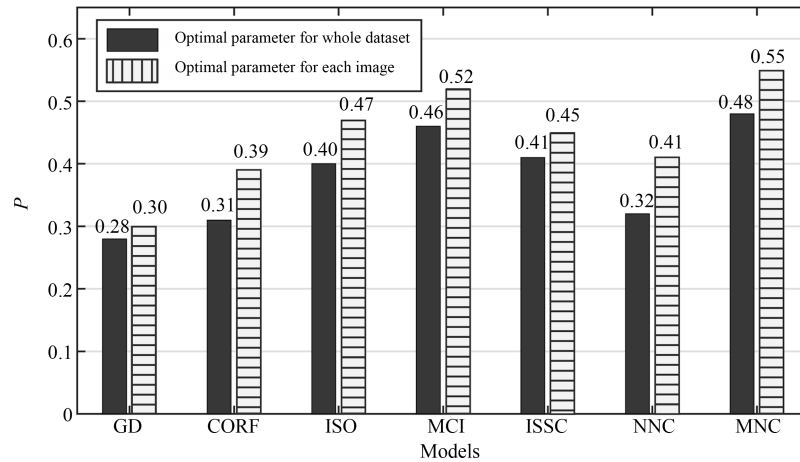


图5 各模型在整个数据集中的定量分析图

Fig. 5 Quantitative comparison of various models on the whole RuG40 dataset

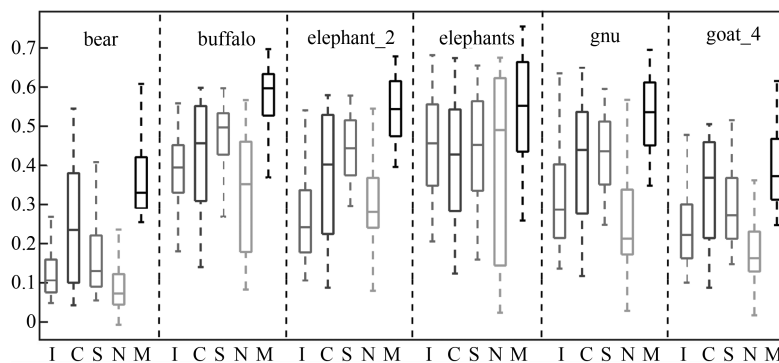


图6 随机选取的6组图像在多组参数下检测结果的P值盒须图统计(I表示ISO方法,C表示MCI方法,S表示ISSC方法,N表示NNC方法,M表示MNC方法)

Fig. 6 Box-and-Whisker plots of the performance of the ISO (denoted by I), the MCI (denoted by C), the ISSC (denoted by S), the NNC (denoted by N), and the MNC (denoted by M) for six random test images of multiparameters

视觉信息的时空编码,在纹理抑制前,先将图像做稀疏处理,弱化了纹理背景边缘的影响,并用时空编码凸显了主体轮廓,为后续的纹理背景抑制做了有效的预处理.然后对抑制了纹理背景的图像进行信息冗余度增强编码,强化图像特征信息,有助于初级视皮层对图像的完整分析和处理.最后通过多路径前馈融合来快速调节初级视皮层内的轮廓响应,在突出主体轮廓的同时,有效地抑制了纹理背景,更符合视通路处理视觉信息的生物机制.

4 结论

本文重点探讨视通路中神经元的电生理特性,在不同层级间根据生物机制采取不同的神经编码.首先,根据神经元特性提出了时空编码方式,提高神经元信息传递的效率,根据外膝体细胞感受野特性,引入非经典感受野的侧抑制作用;然后,在视辐射区提出信息冗余度增强编码,强化特征信息;最后,考虑到视觉系统的快速响应能力,提

出前馈环节,调节初级视皮层的轮廓信息.通过与GD、CORF、ISO、MCI、ISSC、NNC相比,本文方法MNC所提取的轮廓信息更加完整,同时很大程度上抑制了周边纹理背景边缘,从整体上反映了算法的智能性.

References

- 1 Mignotte M. A biologically inspired framework for contour detection. *Pattern Analysis and Applications*, 2017, **20**(2): 365–381
- 2 Ding L J, Goshtasby A. On the Canny edge detector. *Pattern Recognition*, 2001, **34**(3): 721–725
- 3 Jiang Zhao-Hui, Wu Qiao-Qun, Gui Wei-Hua, Yang Chun-Hua, Xie Yong-Fang. Adaptive detection of blast furnace surface contour with fractional multi-directional differential operator. *Acta Automatica Sinica*, 2017, **43**(12): 2115–2126 (蒋朝辉, 吴巧群, 桂卫华, 阳春华, 谢永芳. 基于分数阶的多向微分算子的高炉料面轮廓自适应检测. *自动化学报*, 2017, **43**(12): 2115–2126)
- 4 Arbelaez P, Maire M, Fowlkes C, Malik J. Contour detection and hierarchical image segmentation. *IEEE Transac-*

- tions on *Pattern Analysis and Machine Intelligence*, 2011, **33**(5): 898–916
- 5 Martin D R, Fowlkes C C, Malik J. Learning to detect natural image boundaries using local brightness, color, and texture cues. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, **26**(5): 530–549
 - 6 Young R A, Lesperance R M. The Gaussian derivative model for spatial-temporal vision: II. Cortical data. *Spatial Vision*, 2001, **14**(3): 321–389
 - 7 Azzopardi G, Petkov N. A CORF computational model of a simple cell that relies on LGN input outperforms the Gabor function model. *Biological Cybernetics*, 2012, **106**(3): 177–189
 - 8 Grigorescu C, Petkov N, Westenberg M A. Contour detection based on nonclassical receptive field inhibition. *IEEE Transactions on Image Processing*, 2003, **12**(7): 729–739
 - 9 Zeng C, Li Y J, Li C Y. Center-surround interaction with adaptive inhibition: a computational model for contour detection. *NeuroImage*, 2011, **55**(1): 49–66
 - 10 Yang K F, Li C Y, Li Y J. Multifeature-based surround inhibition improves contour detection in natural images. *IEEE Transactions on Image Processing*, 2014, **23**(12): 5020–5032
 - 11 Liao Jin-Wen, Fan Ying-Le, Wu Wei, Gao Yun-Yuan, Li Yi. Image edge detection based on spike coding of multilayer neuronal population with inhibitory synapse. *Chinese Journal of Biomedical Engineering*, 2014, **33**(5): 513–524 (廖进文, 范影乐, 武薇, 高云园, 李轶. 基于抑制性突触多层神经元群放电编码的图像边缘检测. *中国生物医学工程学报*, 2014, **33**(5): 513–524)
 - 12 Chen M G, Yan Y, Gong X J, Gilbert C D, Liang H L, Li W. Incremental integration of global contours through interplay between visual cortical areas. *Neuron*, 2014, **82**(3): 682–694
 - 13 Freud E, Plaut D C, Behrmann M. ‘What’ is happening in the dorsal visual pathway. *Trends in Cognitive Sciences*, 2016, **20**(10): 773–784
 - 14 Taylor W R, Vaney D I. Diverse synaptic mechanisms generate direction selectivity in the rabbit retina. *The Journal of Neuroscience*, 2002, **22**(17): 7712–7720
 - 15 Bays P M. A signature of neural coding at human perceptual limits. *Journal of Vision*, 2016, **16**(11): 4
 - 16 Zhu M C, Rozell C J. Visual nonclassical receptive field effects emerge from sparse coding in a dynamical system. *PLoS Computational Biology*, 2013, **9**(8): e1003191
 - 17 Yu Q, Tang H J, Tan K C, Li H Z. Rapid feedforward computation by temporal encoding and learning with spiking neurons. *IEEE Transactions on Neural Networks and Learning Systems*, 2013, **24**(10): 1539–1552
 - 18 Hu J, Tang H J, Tan K C, Li H Z. How the brain formulates memory: a spatio-temporal model research frontier. *IEEE Computational Intelligence Magazine*, 2016, **11**(2): 56–68
 - 19 Buonocore A, Caputo L, Pirozzi E, Carfora M F. A leaky integrate-and-fire model with adaptation for the generation of a spike train. *Mathematical Biosciences and Engineering*, 2016, **13**(3): 483–493
 - 20 Muhle-Karbe P S, Duncan J, De Baene W, Mitchell D J, Brass M. Neural coding for instruction-based task sets in human frontoparietal and visual cortex. *Cerebral Cortex*, 2017, **27**(3): 1891–1905
 - 21 Tang J Y, Jimenez S C A, Chakraborty S, Schultz S R. Visual receptive field properties of neurons in the mouse lateral geniculate nucleus. *PLoS One*, 2016, **11**(1): e0146017
 - 22 Xie Zhao, Tong Hao-Hao, Sun Yong-Xuan, Wu Ke-Wei. Dynamic contour detection inspired by biological visual perception. *Acta Automatica Sinica*, 2015, **41**(10): 1814–1824 (谢昭, 童昊浩, 孙永宣, 吴克伟. 一种仿生物视觉感知的视频轮廓检测方法. *自动化学报*, 2015, **41**(10): 1814–1824)
 - 23 Li Zhi-Yong, He Shuang, Liu Jun-Min, Li Ren-Fa. Motion filtering by modelling R3 cell’s receptive field in frog eyes. *Acta Automatica Sinica*, 2015, **41**(5): 981–990 (李智勇, 何霜, 刘俊敏, 李仁发. 基于蛙眼 R3 细胞感受野模型的运动滤波方法. *自动化学报*, 2015, **41**(5): 981–990)



张明琦 杭州电子科技大学自动化学院硕士研究生. 2016 年获得杭州电子科技大学学士学位. 主要研究方向为计算机视觉, 图像处理.

E-mail: 161060075@hdu.edu.cn

(**ZHANG Ming-Qi** Master student at the College of Automation, Hangzhou Dianzi University. He received his bachelor degree from Hangzhou Dianzi University in 2016. His research interest covers computer vision and image processing.)



范影乐 杭州电子科技大学自动化学院教授. 2001 年获得浙江大学生物医学工程博士学位. 主要研究方向为神经信息学, 机器视觉, 机器学习. 本文通信作者.

E-mail: fan@hdu.edu.cn

(**FAN Ying-Le** Professor at the College of Automation, Hangzhou Dianzi University. He received his Ph.D. degree from Zhejiang University in 2001. His research interest covers neuroinformatics, machine vision, and machine learning. Corresponding author of this paper.)



武薇 杭州电子科技大学自动化学院讲师. 2012 年获得浙江大学生物医学工程系博士学位. 主要研究方向为医学信息学, 计算机图像处理.

E-mail: ww@hdu.edu.cn

(**WU Wei** Lecturer at the College of Automation, Hangzhou Dianzi University. She received her Ph.D. degree from Zhejiang University in 2012. Her research interest covers medical informatics, computer image processing.)