

视觉感知的目标识别算法

高庆一¹, 逯鹏², 袁道任²

(1. 北京航空航天大学 计算机学院, 100191 北京; 2. 郑州大学 电气工程学院, 450001 郑州)

摘要: 针对图像的目标识别问题,采用视觉感知的方法,模拟感受野的分层信息处理机制,并引入神经元间的侧抑制机制,对神经元响应进行筛选,通过检测视觉基本特征的方式识别图像中的目标. 算法首先在简单细胞的感受野中对图像进行预处理;其次,在复杂细胞的感受野中,将简单细胞的感受野刺激进一步拓扑特征提取,得到感受野刺激响应;最后,通过侧抑制机制对响应神经元筛选,找出对刺激响应较强烈的神经元,将其输出作为目标识别的参数标准. 实验结果表明,基于视觉感知的算法可以用少量样本解决大量图像中的目标识别问题,识别率高于边缘检测和图像分割方法,算法的目标识别率达到 95.56%.

关键词: 感受野;视觉感知;侧抑制;目标识别

中图分类号: TP2

文献标志码: A

文章编号: 0367-6234(2013)01-0101-05

Target recognition algorithm based on visual perception

GAO Qingyi¹, LU Peng², YUAN Daoren²

(1. School of Computer Science and Engineering, Beihang University, 100191 Beijing, China;

2. School of Electrical Engineering, Zhengzhou University, 450001 Zhengzhou, China)

Abstract: Aiming at the problem of target recognition in image processing, a novel algorithm based on visual perception with better information processing mechanism was proposed for target recognition, in which the hierarchical information processing mechanism of receptive field were simulated and the lateral inhibition mechanism which was used to filter response of neurons was introduced. Firstly, in receptive field of simple cells, the images are initially processed. Secondly, in receptive field of complex cells, stimulation of receptive field from former layer is further feature extracted, and then a new receptive field is obtained. And finally, the neurons which are strongly responded to the stimulation are found out by lateral inhibition mechanism, and their correspondent content can be output as the standard of aerial target recognition, the recognition rate of algorithm is 95.56%. Experimental results show that this novel algorithm has high efficiency, and can achieve a great number of target recognitions with few samples, and the recognition rate is higher than that of edge detection and image segmentation method.

Key words: receptive field; visual perception; lateral inhibition; target recognition

目标识别是图像处理中的研究热点之一. 与基于像素的边缘检测^[1]、图像分割^[2]、模板匹配^[3]等传统方法不同,人类视觉是通过感知信息的检测、提取、整合等过程完成目标识别. 其原理和方法与计算机方法有着本质的不同. 作为感知

信息检测中最重要的机制,模拟感受野的信息处理模式是解决基于视觉的目标识别^[4-5]的关键问题. 目前典型的模型是二维 Gabor 函数^[6-7]. 二维 Gabor 函数拟合所得到的结果与视觉的非线性处理机制是冲突的^[8]. 主要原因是其只模拟简单细胞,缺少对复杂细胞感受野机制的模拟.

根据感受野特征的复杂度逐渐增高的特点,以感受野分层特性及神经元响应筛选的必要性为前提,以实现目标识别为目的,模拟复杂细胞感受野并利用侧抑制机制筛选神经元响应,进而设计

收稿日期: 2011-10-21.

基金项目: 国家自然科学基金资助项目(60971110,61172152);
郑州市科技攻关资助项目(112PPTGY219-8).

作者简介: 高庆一(1979—),男,博士研究生.

通信作者: 逯鹏, lupeng@zzu.edu.cn.

算法并验证模型的有效性. 从而建立新的视觉模拟计算模型, 该模型首先去除图像二阶相关性, 减少冗余; 然后对简单细胞感受野的刺激进行拓扑特征提取, 得到复杂细胞的感受野刺激响应; 进而模拟侧抑制机制, 解决非经典感受野对经典感受野的抑制作用易受制约的问题^[9]. 实验结果表明, 算法能利用少样本实现大量图像的目标识别.

1 视觉计算模型

模拟视觉的计算模型如图 1 所示.

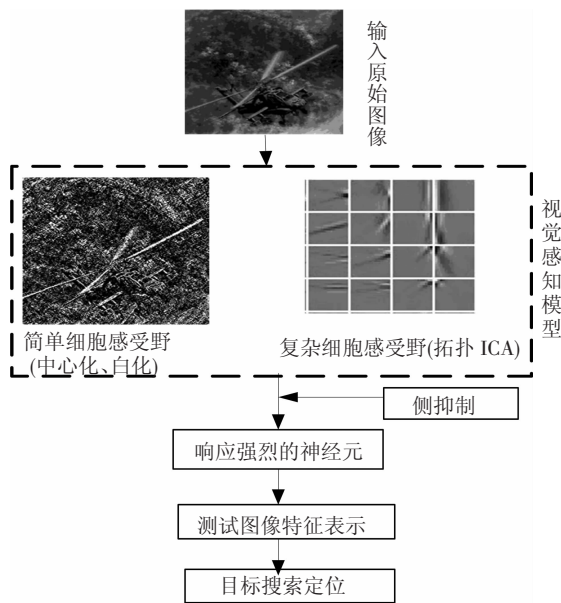


图 1 模拟视觉计算模型图

1.1 基函数训练

图像拓扑特征提取采用拓扑 ICA^[10] 方法, 算法思想是允许小邻域 $h(i, j)$ 内分量间相关邻域外分量尽量独立, 基函数的相关性是采用能量相关定义的.

$$h(i, j) = \begin{cases} 1, & \text{if } |i - j| \leq m; \\ 0, & \text{otherwise.} \end{cases}$$

式中: m 为邻域的宽度. 当 $m = 2$ 时, 则定义了一个 3×3 的邻域.

基函数的相关性用能量相关定义为

$$\text{cov}(s_i^2, s_j^2) = E\{s_i^2 s_j^2\} - E\{s_i^2\} E\{s_j^2\}.$$

式中: $\text{cov}(s_i^2, s_j^2) \neq 0$ 为两个神经元的响应 s_i 和 s_j 是拓扑相邻的, 其优化目标函数为

$$\log L(\mathbf{W}) = \sum_{i=1}^T \sum_{j=1}^n G\left(\sum_{i=1}^n h(i, j)(w_i^T x(t))^2\right) + T \log |\det \mathbf{W}| + \text{constant}.$$

式中: $G(\cdot)$ 为非线性调节函数; w_i 为分离基向量; $x(t)$ 为样本; 学习算法采用自然梯度下降方法, 更新规则为

$$\Delta \mathbf{W} = -\eta \frac{\partial J}{\partial \mathbf{W}} \mathbf{W}^T \mathbf{W} = \eta [I - \varphi(X(k) u^T(k)) \mathbf{W}^T] \mathbf{W} = \eta [I - \varphi(X(k) x^T(k))] \mathbf{W}. \quad (1)$$

式中: $\varphi(X(k) u^T(k))$ 为批处理中的均值.

最后, 将训练的拓扑基函数集合与图像结合得到图像的拓扑基函数响应.

为筛选并确定集合中响应强烈的神经元, 首先根据目标特征设定阈值为

$$\theta = w_{\max} - \lambda(w_{\max} - w_{\min})/160.$$

式中: $w_{\max}$ 为最大响应系数; $w_{\min}$ 为最小响应系数; λ 为常数.

算法 1 拓扑基函数学习算法.

输入: 16 幅样本图像.

输出: 响应矩阵 $\mathbf{W}$ 及相应的拓扑基函数 $\mathbf{A}$.

步骤:

- 1) 随机采样样本图像, 得到训练样本数据;
- 2) 中心化和白化处理样本数据, 作为算法输入数据;
- 3) 基于 ICA 模型 $S = \mathbf{W}X$ 计算 X , 根据式(1)对 $\mathbf{W}$ 进行更新, 每个基函数归一化成为单位向量;
- 4) 如果 $\Delta \mathbf{W} \leq \theta$ 时, 则停止迭代, 否则回到步骤 3);
- 5) 停止学习, 输出结果响应矩阵 $\mathbf{W}$ 及原图像空间基函数矩阵 $\mathbf{A}$.

1.2 基于侧抑制的神经元筛选

神经系统的侧抑制^[11] 是某个神经元产生兴奋时, 对相邻神经元有抑制作用. 神经元侧抑制模型^[12] 如图 2 所示.

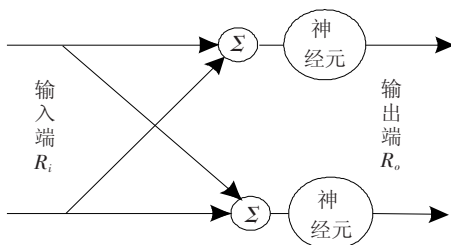


图 2 神经元侧抑制模型

设定 R_o 为神经元响应输出, R_j 为输入, R_{ij} 为抑制阈值, n 为神经元细胞的抑制的范围, K_j 为抑制系数, 则侧抑制网络的数学模型为

$$R_o = R_i - \sum_{j \neq i}^n K_j (R_j - R_{ij}).$$

两神经元之间的距离定义为欧几里德距离 d . 侧抑制系数 K_j 可表示为

$$K_j = \begin{cases} \frac{\alpha}{d}, & d \geq 1; \\ 0, & \text{otherwise.} \end{cases}$$

式中 α 为常数.

侧抑制网络的数学模型可以表示为

$$R_o = \begin{cases} R_i - \sum_{j \neq i}^n \frac{\alpha}{d} (R_j - R_{ij}), & d \geq 1; \\ R_i, & \text{otherwise.} \end{cases}$$

通过对抑制阈值 R_{ij} 和抑制系数 K_j 的控制, 达到对神经元响应的筛选. 设筛选完成之后的输出为 Y , 若 $R_j \geq R_{ij}$, 将 Y 置 1; 否则, 将 Y 置 0.

图像刺激只是引起少部分神经元响应强烈, 大部分神经元在基值 0 附近, 模拟侧抑制机制, 筛选出响应强烈的神经元.

算法 2 基于侧抑制的筛选算法.

输入: 含有识别目标的图像和待测试图像.

输出: 响应强烈神经元的侧抑制特征

步骤:

- 1) 对样本图像顺序采样;
- 2) 用主分量分析方法进行白化, 得到预处理后的输入图像对应矩阵;
- 3) 设定抑制阈值 R_{ij} 和抑制系数 K_j , 将其输入与 R_j 比较;
- 4) 将比较后的对应内容输出, 即为响应强烈神经元对应特征.

2 目标识别算法

基于上述模型, 目标识别算法 TRA (Target Recognition Algorithm) 基本思想是: 首先从样本图像中学习并提取拓扑基函数; 然后利用侧抑制机制对这些基函数进行筛选; 最后用含有目标和待测试图像对神经元进行模拟刺激, 通过对比差异识别目标.

算法 3 目标识别算法 TRA.

输入: 含有识别目标的图像和待测试图像.

输出: 感受野感受到的相应刺激内容.

步骤:

- 1) 对采样的图像块进行白化处理;
- 2) 基于 ICA 模型 $S = WX$ 计算 S ;
- 3) 利用侧抑制机制对 S 进行筛选, 找出最佳神经元响应;
- 4) 输出该神经元所表示的感受野内容;
- 5) 对有差异的图像神经元所表示的感受野内容进行对比, 定位差异区域.

3 实 验

选取实验图像 10 幅 256×256 的规则的自然图像作为样本图像进行训练.

3.1 数据采样及预处理

算法在 Matlab 7.1 环境下运行, 操作系统为 Windows XP, CPU 为 Intel (R) Core (TM) 2 Duo T5870 + @ 2.0 GHz, 2.0 GHz@2 GB.

采用 16×16 的滑动空间子窗口对每幅图像进行随机采样, 这样可以消除像素之间的统计特性误差, 10 幅图像共得到 $50\,000 \times 256$ 的输入数据集合; 然后在简单细胞的感受野中将集合中心化和白化后并降至 160 维, 得到的样本矩阵作为基函数学习的输入; 最后在复杂细胞感受野中对输入数据进行训练, 得到视神经的拓扑基函数, 如图 3 所示.

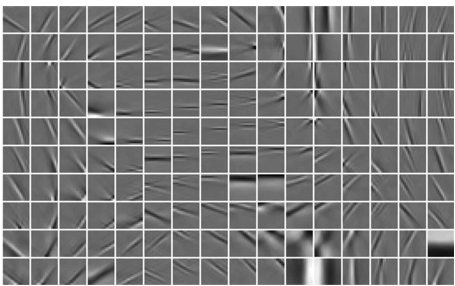
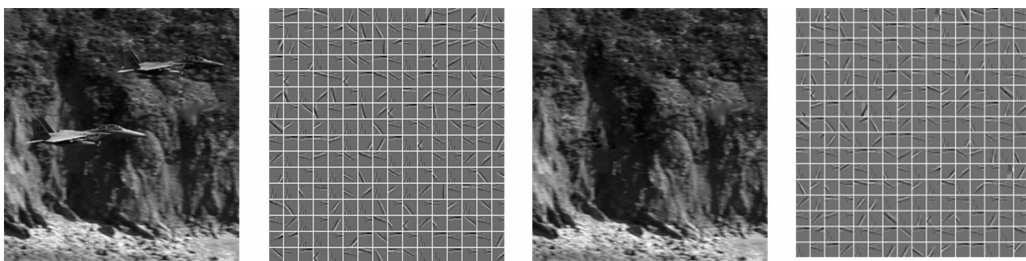


图 3 10 幅图像训练的拓扑基函数

3.2 目标识别

采用大小为 16×16 的像素滑动空间子窗口对图像进行顺序采样. 在实验中, 受到实验图像刺激后, 会有神经元产生响应, 而这些响应刺激形成视觉感知的感受野, 其感受野如图 4 所示.



(a) 含识别目标的图像 (b) 含识别目标图像的感受野 (c) 待测试图像 (d) 待测试图像的感受野

实验组(1)

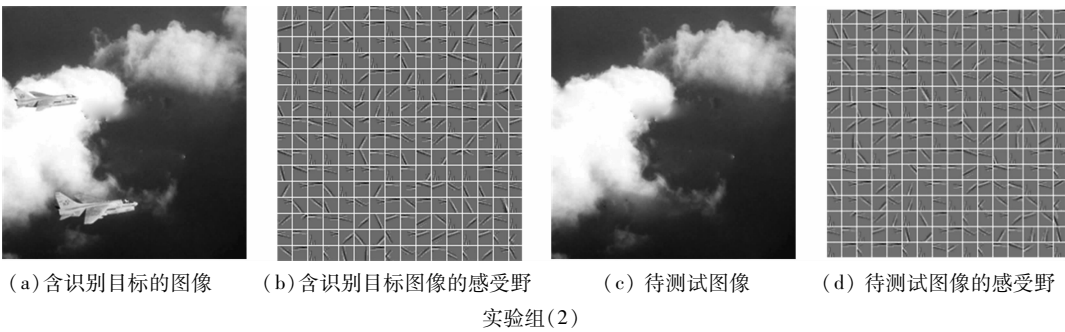


图 4 实验图像的感受野

目标识别过程首先利用侧抑制机制筛选出响应刺激较强的神经元,然后根据神经元的响应位置定位图像区域的目标,如图 5 所示.

3.3 实验对比

分别采用 TRA 算法、边缘检测和图像分割 3 种方法进行目标识别,其实验结果如图 6 所示.

由图 6 可知,在山峰背景下和云层背景下,TRA 算法能够准确识别目标区域位置;边缘检测算法不能检测出目标轮廓;图像分割算法能检测出模糊目标轮廓,目标不易识别. 90 幅有目标图像,3 种算法识别目标的结果如表 1 所示.

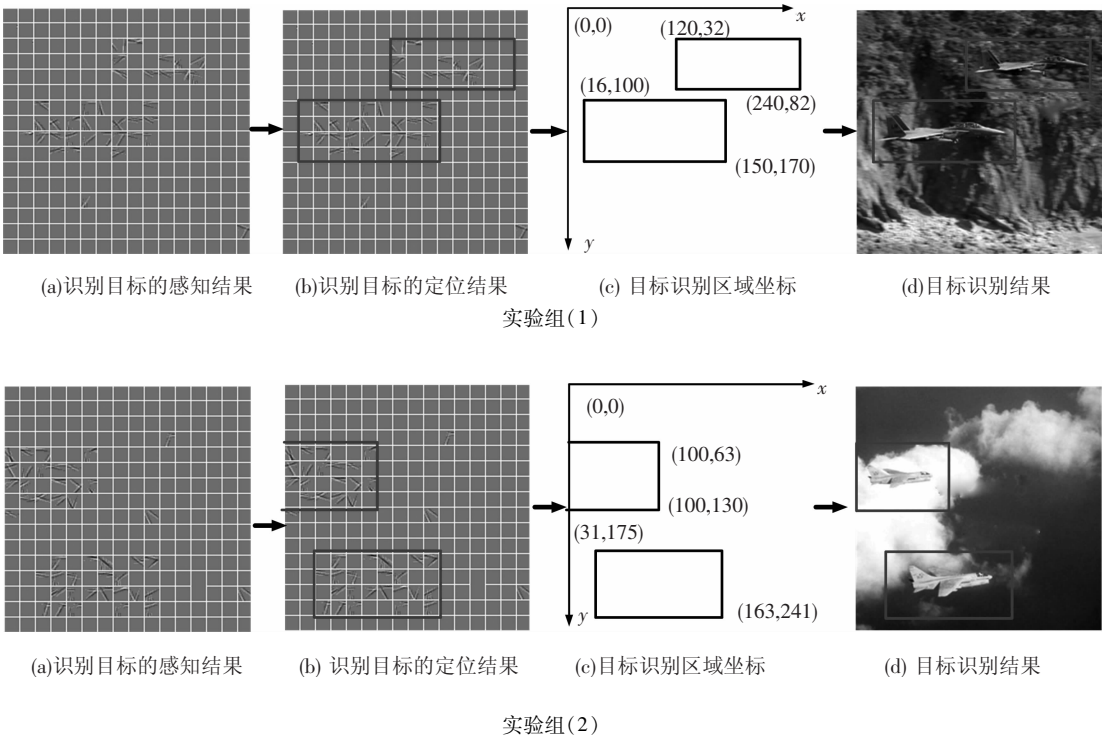
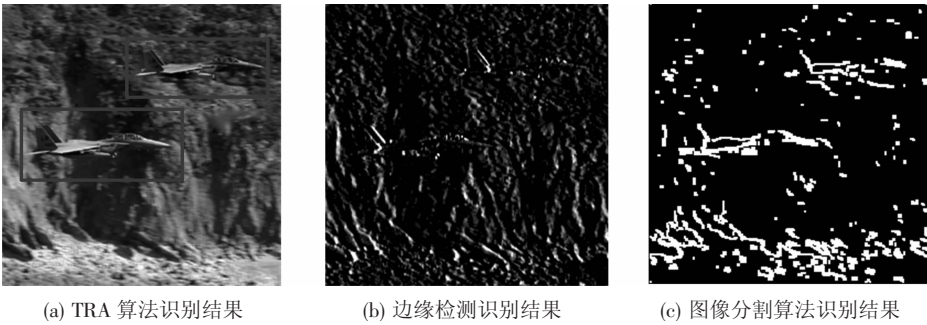


图 5 目标识别



实验组(1)

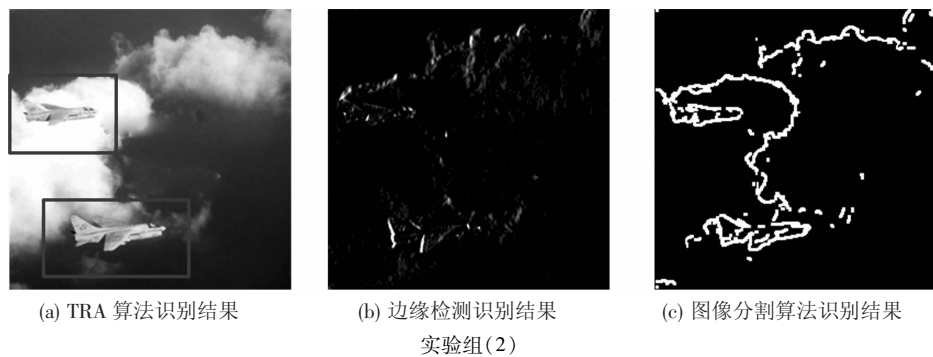


图 6 实验对比结果

表 1 90 幅含有识别目标的图像的实验统计结果

算法	图片/ 幅	正确识别/ 幅	遗漏识别/ 幅	错误识别/ 幅	正确率/ %
TRA 算法	90	86	4	0	95.56
边缘检测算法	90	44	35	11	48.89
图像分割算法	90	51	28	11	56.67

4 结 论

1)模拟复杂细胞感受野分层信息处理机制,并引入神经元之间的侧抑制机制来筛选神经元较强烈的响应,从而能模拟生物视觉的信息处理机制的视觉感知模型。

2)设计算法、高目标识别率验证了模型的有效性.实验结果表明,该方法能快速、准确地进行目标识别。

参考文献

[1] LEI Xiaogang, LI Ying, ZHAO Na, *et al.* Fast segmentation approach for SAR image based on simple Markov random field[J]. Journal of Systems Engineering and Electronics, 2010, 21 (1): 31 – 36.

[2] CHEN Weibin, ZHANG Xin. Based on an improved sobel edged etection operator in the application [C]// Proceedings of 2010 4th International Conference on Intelligent Information Technology Application(IITA). Hebei Province: IEEE, 2010: 439 – 442.

[3] OUYANG Yi, XING Jianguo. Based on adaptive weighted CAMSHIFT human motion tracking algorithm [C]//Proceedings of 2010 2nd International Conference on Intellectual Technology in Industrial Practice (ITIP2010). Hebei Province: IEEE, 2010: 85 – 87.

[4] 逯鹏,李永强,王治忠,等. 基于视觉感知的故障图像检测算法[J]. 仪器仪表学报,2010,31(9):1997 – 2002.

[5] 张菁,沈兰荪,David Dagan Feng. 基于视觉感知的图像检索的研究[J]. 电子学报,2008,36(3): 494 – 499.

[6] HARUTIUNIAN-KOZAK B A, KHACHVANKYAN D K, GRIGORYAN G G, *et al.* Dynamic spatial organization of receptive fields of neurons in the 21a cortical area [J]. Neurophysiology, 2010, 45 (3): 175 – 184.

[7] XU Wanying, HUANG Xinsheng, WANG Yiping, *et al.* A biologically motivated corner detection method based on the oriented receptive fields of simple cortical cells [C]//Proceedings of the 2010 International Conference on Biomedical Engineering and Computer Science (ICBECS). Washington, DC: IEEE Computer Society, 2010: 1 – 4.

[8] OLSON S J, GROSSBERG S. A neural network model for the development of simple and complex cell receptive fields within cortical maps of orientation and ocular dominance [J]. Neural Networks, 1998, 11(2):189 – 208.

[9] DOU Y, WANG L F, KONG L F. A model of nonclassical receptive field in visual cortex [J]. Journal of Yanshan University, 2009, 33(2): 109 – 113.

[10] LEE Gin-Chong, LOO Chu-Kiong, LETCHUMANAN C. Two-factor face authentication: topographic independent component analysis (TICA) and multispace random projection (MRP) [C]// Proceedings of the 2009 International Conference of Soft Computing and Pattern Recognition. Washington, DC: IEEE Computer Society, 2009: 756 – 761.

[11] RATLIFF F, HARTLINE H. The responses of the Limulus optic nerve fibers to patterns of illumination on the receptor mosaic [J]. J Gen Physiol, 1959, 42 (6): 1241 – 1255.

[12] CHEN Yuepeng, FU Hongwei. Study and application of lateral inhibition models in image's contour enhancement [C]//Proceedings of the 2010 International Conference on Computer Application and System Modeling (ICCASM). Washington, DC: IEEE Computer Society, 2010: 23 – 27.

(编辑 张 红)