

基于视觉认知的禁令交通标志检测

胡晓光^{1,2,†} 程承旗² 李德仁³

1. 中国人民公安大学刑事科学技术学院, 北京 100038; 2. 北京大学工学院, 北京 100871; 3. 武汉大学测绘遥感信息工程国家重点实验室, 武汉 430079; † E-mail: hxgnol@126.com

摘要 根据交通标志的反差特性会强烈吸引人类视觉注意的设计原则, 结合生物视网膜会强烈响应场景中反差视觉刺激的机理, 将基于视觉反差的层次结构的显著性分析框架引入交通标志的检测问题, 提出一种适合现实街景中交通标志检测的多线索视觉注意模型, 将对交通标志的检测定位转变为对显著目标的发现与分割问题。实验表明, 所提方法优于典型的显著性方法在面对现实街景时的目标检测性能。

关键词 交通标志检测; 街景; 显著性分析; 视觉反差

中图分类号 TP391

Prohibition Traffic Signs Detection Based on Visual Cognition

HU Xiaoguang^{1,2,†}, CHENG Chengqi², LI Deren³

1. School of Criminal Science and Technology, People's Public Security University of China, Beijing 100038; 2. College of Engineering, Peking University, Beijing 100871; 3. State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079; † E-mail: hxgnol@126.com

Abstract Considering the design principles that traffic signs is to strongly attract the human visual attention, combining the phenomenon that the retina strongly responds to large contrast visual stimulation, a hierarchy saliency analytic framework based on visual contrast is introduced. The authors propose a multi-cue visual attention model for traffic sign detection in street scene, so traffic sign detection and segmentation problem is converted to saliency object discovery and location problem. Experimental results show that the proposed method is better than typical saliency methods.

Key words traffic sign detection; street scene; saliency analysis; visual contrast

随着汽车拥有量的不断增加, 针对智能交通领域的研究逐渐兴起并成为当前热点, 其中车辆无人驾驶和辅助驾驶等课题, 因其成果易应用于其他机器视觉领域, 近年来受到研究人员的极大重视。交通标志作为一种对于行车安全与秩序具有重要作用的道路信息, 其自动检测与识别是实现无人驾驶需解决的首要问题之一。同时, 作为一个集影像处理、模式识别等多学科交叉应用的研究方向, 也具有重要的理论研究价值。传统的交通标志检测多从颜色^[1]、形状^[2-3]或者综合使用颜色和形状等特征^[4-8]着手。其中, 基于色彩的方法比较容易实现,

如基于 RGB 和 HSL 的方法等, 但这种方法容易受到外部光线、色温、光照角度以及获取角度的影响, 性能不稳定。形状在光照变化下能够得到完整的保留, 但容易受到遮挡物和视角的影响。单一特征的不足, 使多特征的融合成为目前研究的主流, 但不同特征自身的局限依然影响着最终检测能力的提高。这是因为当前方法多从机器认知的角度出发, 当不满足其中的某项前提时, 便无法取得最佳效果。

认知神经科学的研究表明, 复杂环境中存在的多种干扰因素使人眼捕获的影像不能完全反映场景

的本来信息,但人类依然具有排除干扰,准确识别出其中某一目标的能力。对这种生物能力的模拟近几年成为提高机器对外界认知的一种思路。我们发现交通标志的形状和颜色与周围环境存在较大的反差,这使其能够在街景中强烈地吸引驾驶员的视觉注意。本文将这种设计原则与生物学上视网膜会强烈响应大反差视觉刺激的机理相结合,提出一种新的基于视觉认知的交通标志检测方法。

1 基于多线索视觉注意模型的标志检测

1.1 多线索视觉注意模型

Sun 等^[9]发现,在对复杂场景的认知过程中,人类并不只使用显著性机制这种底层的认知手段,而是与高层的知识引导一起完成某种目标的检测任务,机器视觉的应用反馈也说明了这一点。据此,本文提出如图 1 所示的基于多线索视觉注意模型的交通标志检测流程,由粗到细,分为获取视觉注意区域、分析视觉注意区域两个阶段。

在视觉注意初期,使用多线索的视觉注意机制,在场景中预先找到可能包含标志的感兴趣区域(即候选区域),以缩小检索范围,其中的多线索指来自上层的先验指导和底层的显著性分析。先验指导即 Top-Down 阶段,我们选用由 Asakura 等^[10]提出的简单向量滤波器,该滤波器能够高效地提取指定颜色和去除轮廓,并对红、蓝和黄色有较好的分割效果,非常适合颜色交通标志的提取。显著性分析即 Bottom-Up 阶段,我们对场景的反差信息进行显著性分析,根据标志牌与背景的大反差线索来提取标志牌,这符合人眼的视觉机理。在获得两条线索的

注意区域后做交运算,得到共同的注意结果。

在分析和检测阶段,根据形状特征对这些候选区域进行分析,以获得所需目标区域。本文选用对交通安全具有重要影响的禁令标志作为实验对象。利用其先验信息(如禁令标志的形状为圆形等),使用一种中央双十字形状度量的方法^[11]来分析感兴趣区域的形状特性,进一步排除无关干扰因素,得到最终结果。

1.2 基于反差信息的显著性分析

基于视觉反差的显著图生成与目标检测,是对视网膜在接受外部场景光线刺激时会强烈响应其中高反差目标的现象和初级视觉皮层上视觉信息的生成机制进行的模拟,模拟方法由两层计算单元组成: S 和 C, 分别代表初级视觉皮层中的简单细胞和复杂细胞。

S 单元: 人眼感受野对于大反差视觉信息的输入有强烈的反映,这种生物特性可以借助高分辨率的中央层和低分辨率的周边层之间的差操作^[12]来模拟。灵长类的初级视觉皮层包含简单细胞和复杂细胞。简单细胞的感受野只包括视野中的一小部分,只有当这些局部的单元信息被汇集成具有全局意义的特征后,视觉系统才能感知其中的目标。Edward^[13]指出复杂细胞是简单细胞集群的一种非线性时空整合。本文视反差为局部特征,将视网膜输出的反差信息作为初级视觉皮层中简单细胞所感受到的局部单元信息,当这些局部的反差信息集成在一起形成场景中的反差区域时,就构成复杂细胞所感受的全局信息。

C 单元: C 单元由 S 单元汇集而成,而计算池模型是沟通灵长类视觉皮层中简单细胞和复杂细胞的

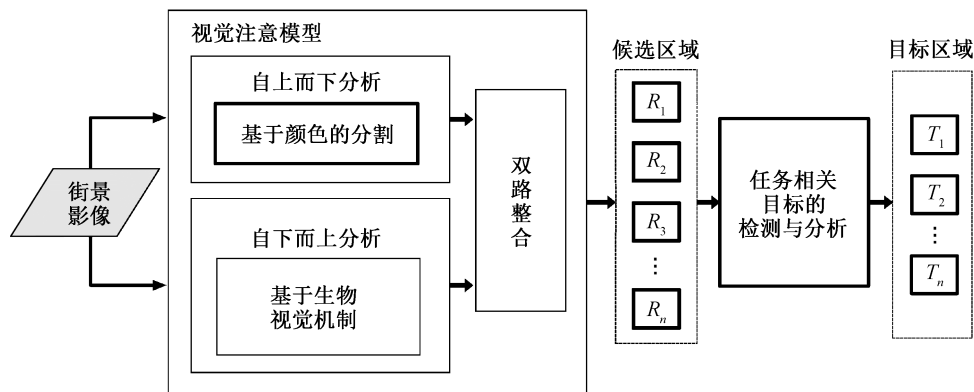


图 1 交通标志的检测流程
Fig. 1 Detection flowchart of traffic sign

桥梁。文献[14]中提到 3 种基本的计算池模型, 包括 Maximum model, Energy model 和 Half-wave model。研究显示, 最大化的计算池模型是简单细胞向复杂细胞汇集的方式^[15-17]。据此, 我们选用最大化模型作为整合的方法。

图 2 是在亮度通道进行的标志反差显著性分析流程。经过 S 单元和 C 单元对反差信息的分析, 影像右上角具有大反差特征的交通标志信息获得较高的显著度, 这一结果符合生物对外界光刺激的响应机制。

标志包含多种可用的显著性特征, 通过对其颜色、亮度和方位特征以及它们之间组合的显著性进行实验分析, 我们得到: 方位特征并不适合交通标志的显著性检测, 亮度特征的效果居中, 颜色特征的效果最好。为了最大化标志的显著性, 我们选用亮度、红绿及蓝黄对比色作为交通标志反差显著性分析的显著特征。基于视觉反差显著性的交通标志检测步骤如下。

步骤 1 生成影像金字塔。输入一幅彩色影像后, 首先使用式(1)计算其亮度图。使用式(2)和(3)计算其红绿和蓝黄对比图, 得到标志最为显著的特征通道。Red, Green 和 Blue 分别表示一幅彩色影像的红绿蓝 3 种光谱波段。

$$M_I = \text{Max}(\text{Red}, \text{Green}, \text{Blue}), \quad (1)$$

$$M_{RG} = |\text{Red} - \text{Green}|, \quad (2)$$

$$M_{BY} = |\text{Blue} - \min(\text{Red}, \text{Green})|. \quad (3)$$

使用所得到的亮度图 M_I 、红绿对比图 M_{RG} 和蓝黄对比图 M_{BY} , 建立对应的高斯金字塔 $M_I(\delta)$, $M_{RG}(\delta)$ 和 $M_{BY}(\delta)$ 。根据实验观察, 我们发现金字

塔所抽取的层数并不是越大越好, 尺寸过小的层对最终显著图的贡献不大, 反而会消耗系统的计算资源, 因此, 在我们的流程中, 只经验性地抽取 4 层金字塔, 所以 $\delta \in [1, \dots, 4]$, 第 1 层是原始影像尺寸的 1/4, 往后依次减半, 如第 4 层影像与第 1 层影像之比为 1/8。

步骤 2 计算简单单元 S 的信息。在亮度图和颜色图的图像金字塔上进行中央周边差操作, 即精细尺度与粗略尺度影像的差操作, 中央尺度 $c \in \{1, 2, 3\}$, 周边尺度 $s \in \{2, 3, 4\} (s=c+\delta, \delta=1)$ 。将周边层 $s \in \{2, 3, 4\}$ 插值后放大到与中央层 $c \in \{1, 2, 3\}$ 相同的尺度, 然后按照式(4)~(6)计算得到 S 单元中的多尺度反差显著图, 3 个特征通道共包含 $3 \times 3=9$ 幅影像。 Θ 表示跨尺度的减操作, $N(\cdot)$ 表示不同层影像的归一化操作。

$$S_{I,c,s} = N(M_I(c) \Theta M_I(s)), \quad (4)$$

$$S_{RG,c,s} = N(M_{RG}(c) \Theta M_{RG}(s)), \quad (5)$$

$$S_{BY,c,s} = N(M_{BY}(c) \Theta M_{BY}(s)). \quad (6)$$

步骤 3 计算复杂单元 C 的信息。将来自 S 单元的多尺度反差显著图插值放大到与层数 $\delta=1$ 相同的大小, 然后根据 Max 模型, 使用式(7)~(9)对各特征通道中 S 单元的影像进行计算, 得到亮度、RG 和 BY 通道各自的整体反差显著图 C_I , C_{RG} 和 C_{BY} 。

$$C_I = N(\text{Max}(S_I)), \quad (7)$$

$$C_{RG} = N(\text{Max}(S_{RG})), \quad (8)$$

$$C_{BY} = N(\text{Max}(S_{BY})). \quad (9)$$

步骤 4 生成总显著图 C_{total} 。对 C_I , C_{RG} 和 C_{BY} 进行 Max 和 $N(\cdot)$ 操作, 得到总显著图 C_{total} 。

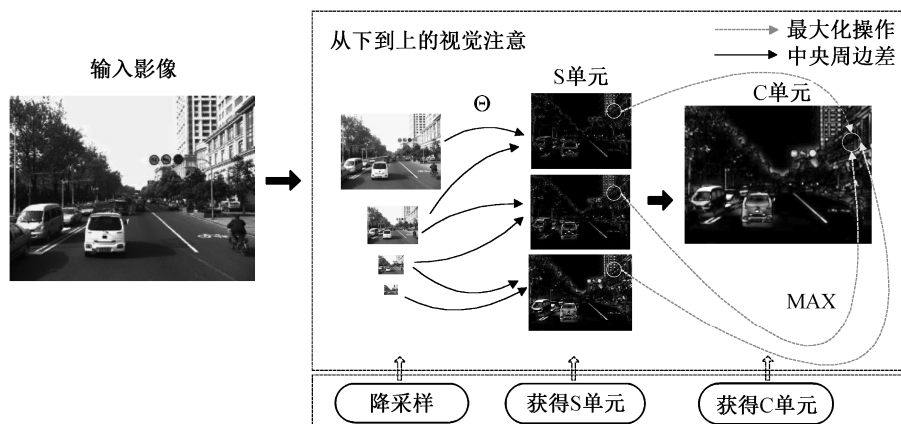


图 2 标志的反差显著性分析流程

Fig. 2 Flowchart of contrast saliency analysis

$$C_{total} = N(\text{Max}(C_I, C_{BY}, C_{RG}))。 \quad (10)$$

为了突出显著区域, 获得更好的视觉效果, 使用式(11)对总显著图 C_{total} 进行高斯滤波, 得到最终的总显著图 $S(\text{Image})$ 。本文的高斯滤波算子 h_g 大小为 9×9 , 方差为 8。

$$S(\text{Image}) = h_g \times C_{total}。 \quad (11)$$

步骤 5 得到注意结果图。首先使用式(12)对显著图进行阈值分割, 得到二值化的显著图, $E(S(\text{Image}))$ 为显著图的平均强度, 经验阈值 $\text{threshold} = E(S(\text{Image})) \times 3$ 。

$$S(\text{Image}) = \begin{cases} 1, & S(\text{Image}) > \text{阈值}, \\ 0, & S(\text{Image}) \leq \text{阈值}. \end{cases} \quad (12)$$

然后将二值图像作为模板与实际影像进行叠加, 获得注意结果图 **Result**:

$$\text{Result} = \text{Image} \times S(\text{Image})。 \quad (13)$$

2 实验结果与分析

本研究针对禁令标志, 采集的户外街景数据分为两类: 一类来自移动测量系统自动采集, 另一类来自人工采集。由于自动化采集的时间间隔、采集角度和环境往往是确定的, 不能完全体现实际街景中遇到的各种复杂情况。为了对方法的适应性进行检验, 我们在使用移动测量系统获取数据的基础上, 通过人工对多种标志在不同角度和不同环境中进行补充采集。自动采集数据由移动摄影测量车在湖北省武汉市汉口沿江大道采集(武汉立得公司提供), 采集过程使用移动测量系统中4个视向的 CCD 摄像机, 但未使用其提供的定位信息; 人工采集数据在武汉大学附近采集。实验区域的信息类型包括标志牌、道路、建筑、车辆、行人、地面标志线、树木和阴影等。

街景是一种典型的复杂场景, 距离远近、倾角、户外的光线条件、人造光源和人工地物等都会极大影响显著性分析的结果, 原因在于: 1) 场景中有多种干扰因素存在, 一些目标的光谱、形状和纹理与交通标志的显著特征相似; 2) 同一目标在不同的角度、天气和光线下会有不同的表现, 待识别标志的底层特征并不唯一。

我们对基于视觉反差机制的显著性分析方法在街景中的视觉注意能力进行实验。由于视觉注意能力在本文方法中的重要性, 所以需要与同属视觉注意机制的方法进行对比, 以便体现方法的差异。对

比的方法为基于空间域的 Itti 显著区域提取模型^[12]和基于频率域的 Hou 波谱残差模型^[18]。图3为在不同天气、角度、复杂度的场景中, 不同显著性方法的视觉注意结果。由于对比方法未以从视网膜到初级视觉皮层这条视觉通路上的信息处理机制为依据, 所以在图 3(b)和(c)中, 注意区域没有排除非标志区域的干扰而范围过大, 或者细节损失过多而无法保持标志的完整性, 因此不能为形状分析提供良好的基础。图 3(d)是基于反差信息的注意结果, 可以看到, 该方法能更好地保持场景中各类目标的细节, 在突出交通标志显著性的同时压低场景中的无关信息, 利于进一步处理。

图 4 为本文方法在几种典型场景中的标志检测结果, 场景的选取考虑了标志在低照度、大倾角、尺度不同、有强干扰等情况下的被检测性, 检测到的标志在图中叠加显示。图 4(a)~(c)显示本文方法对于小尺度目标和由于倾角而造成的标志变形的检测能力, 结果显示本文方法具有较好的尺度和形变的适应性。图 4(d)反映本文方法在光照不佳情况下的性能。图 4(e)~(f)显示本文方法在面对场景中的人工地物和人造光源时排除干扰的能力, 表明本文方法能够排除大部分的干扰, 但对于具有相近显著特征的目标(如红色的交通信号灯具有较高的亮度显著性, 圆形的蓝色指示标志具有相同的圆形特征等)还不具备较好的辨别能力。这是本文方法的不足, 也是虚警率偏高的一个重要原因。

最后, 对几种方法在大数据集下的性能进行对比, 结果如表 1 所示。实验数据集共包括 673 幅街景影像, 其中含 259 幅交通标志。为了保证场景的复杂性, 图像不做任何切割, 图像的尺寸为 640×480 。实验在软件平台为 MATLAB 2011b, 硬件平台为 AMD 速龙 II X3 425 2.7 G, 2 G 内存的条件下进行。分析表 1 可知, 本文方法与 Itti 方法和 Hou 方法相比, 检测率由 62.2% 和 44.8% 提高到 83.8%, 但仍存在虚警率偏高的不足。

表 1 不同方法的交通标志检测结果对比
Table 1 Comparison of detection results of different methods

方法	检测率/%	虚警率/%
Itti 方法 ^[12] 的多线索检测+区域分析	62.2	24.1
Hou 方法 ^[18] 的多线索检测+区域分析	44.8	31.5
本文方法的多线索检测+区域分析	83.8	30.6



图3 不同显著性方法的分析结果
Fig. 3 Saliency analysis results of different methods



图4 部分实验影像的检测结果
Fig. 4 Part of the detection results

3 结论

本文根据交通标志在场景中与其他目标具有较大反差的设计原则,将基于视觉反差的显著性分析方法引入现实街景中的交通标志检测,并提出一种用于交通标志检测的多线索视觉注意模型。模型将视觉反差作为自底向上数据驱动的线索与由上而下知识引导的线索综合起来获取视觉注意区域,对该区域进行任务相关的目标分析后,排除无关区域,完成交通标志检测的任务。使用两处实验区域的测试样本对该方法进行验证,结果显示,相对于 Itti 模型和 Hou 模型,本文模型能更好地将注意力集中在街景中的标志目标,取得较高的正检率。但本文方法的不足是虚警率过高。这是由于目前的算法只专注于在视觉注意机制的框架内解决标志检测问题,未考虑大脑对视觉结果的反馈调节,面对实景目标提取这样的复杂性检测任务时,无法排除显著度相近目标的干扰。将来会在本文方法中融入学习与记忆能力,实现更加智能的交通标志检测模型。

参考文献

- [1] Fleyeh H, Davami E. Eigen-based traffic sign recognition. *IET Intelligent Transport Systems*, 2011, 5(3): 190–196
- [2] González Á, Garrido M, Llorca D. Automatic traffic signs and panels inspection system using computer vision. *Intelligent Transportation Systems, IEEE Transactions on*, 2011, 12(2): 485–499
- [3] Zhang K, Sheng Y, Li J. Automatic detection of road traffic signs from natural scene images based on pixel vector and central projected shape feature. *Intelligent Transport Systems*, 2012, 6(3): 282–291
- [4] Greenhalgh J, Mirmehdi M. Real-time detection and recognition of road traffic signs. *IEEE Transactions on Intelligent Transportation Systems*, 2012, 13(4): 1498–1506
- [5] 陈龙, 潘志敏, 毛庆洲, 等. 利用 HOG-LBP 自适应融合特征实现禁令交通标志检测. *武汉大学学报: 信息科学版*, 2013, 38(2): 191–194
- [6] Zaklouta F, Stanculescu B. Warning traffic sign recognition using a HOG-based Kd tree. *IEEE Intelligent Vehicles Symposium*, 2011, 32(14): 1019–1024
- [7] Khan J F, Bhuiyan S M, Adhami R R. Image segmentation and shape analysis for road-sign detection. *IEEE Transactions on Intelligent Transportation Systems*, 2011, 12(1): 83–96
- [8] Meuter M, Nunn C, Gormer S M. A decision fusion and reasoning module for a traffic sign recognition system. *IEEE Transactions on Intelligent Transportation Systems*, 2011, 12(4): 1126–1134
- [9] Sun Y, Fisher R. Object-based visual attention for computer vision. *Artificial Intelligence*, 2003, 146(1): 77–123
- [10] Asakura T, Aoyagi Y, Hirose O K. Real-time recognition of road traffic sign in moving scene image using new image filter // *SICE. Iizuka*, 2000: 13–18
- [11] Hu Xiaoguang, Zhu Xinyan, Li Hui. Traffic sign detection based on visual attention model // *Geoinformatics, 2011 19th International Conference on*. Shanghai, 2011: 1–5
- [12] Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998, 20(11): 1254–1259
- [13] Edward M C. Neural mechanisms for the generation of visual complex cells. *Neuron*, 2001, 32(3): 378–380
- [14] Sakai K, Tanaka S. Spatial pooling in the second-order spatial structure of cortical complex cells. *Vision Research*, 2000, 40(7): 855–871
- [15] Riesenhuber M, Poggio T. Hierarchical models of object recognition in cortex. *Nature Neuroscience*, 1999, 2(11): 1019–1025
- [16] Lampl I, Ferster D, Poggio T, et al. Intracellular measurements of spatial integration and the MAX operation in complex cells of the cat primary visual cortex. *Journal of Neurophysiology*, 2004, 92(5): 2704–2713
- [17] Gawne T J, Martin J M. Responses of primate visual cortical V4 neurons to simultaneously presented stimuli. *Journal of Neurophysiology*, 2002, 88(3): 1128–1135
- [18] Hou X, Zhang L. Saliency detection: a spectral residual approach // *Computer Vision and Pattern Recognition, IEEE Conference on*. Minneapolis, MN, 2007: 1–8