

# 基于视觉显著性和超像素融合的物体定位方法

邵明正\*, 齐剑锋, 王希武, 王 路

(军械工程学院 信息工程系, 石家庄 050003)

(\*通信作者电子邮箱 7213182@163.com)

**摘 要:**针对选择性搜索算法所需定位窗口数量过多的问题,提出了一种基于视觉显著性和超像素融合的改进方法。首先,利用视觉显著性图像粗略估计物体的位置;然后,从这些初始位置开始,根据图像的表现特征融合相邻超像素,并引入一种背景分析方法以避免过度融合;最后,利用贪心算法将融合后的区域再进行组合,并生成最终的定位窗口。在 Pascal VOC 2007 数据集上的实验结果表明,与选择性搜索方法相比,在同样的检测标准下(查全率为 0.91),改进后的方法所使用的窗口数量减少了 20%,而重叠率达到了 0.77。该方法由粗到细地进行物体定位,在定位窗口数量较少的情况下仍能保持较高的重叠率和查全率。

**关键词:**物体定位;视觉显著性;超像素;滑动窗口;物体识别

**中图分类号:** TP391.413 **文献标志码:** A

## Object localization method based on fusion of visual saliency and superpixels

SHAO Mingzheng\*, QI Jianfeng, WANG Xiwu, WANG Lu

(Information Engineering Department, Ordnance Engineering College, Shijiazhuang Hebei 050003, China)

**Abstract:** Considering the weakness of the selective search method that needs a large number of windows to localize objects, a novel object localization method based on fusion of visual saliency and superpixels was proposed in this paper. Firstly, the visual saliency map was used to coarsely localize the objects, and then the adjacent superpixels could be merged according to the appearance features of image, starting from the above coarse positions. Furthermore, the method employed a simple background detector to avoid the over-merge. Finally, a greedy algorithm was used to iteratively combine the merged regions and generate the final bounding boxes. The experimental results on Pascal VOC 2007 show that the proposed method leads to a 20% reduction in the number of the bounding boxes on the same detection rate (recall of 0.91) compared to the selective search algorithm, and its overlap rate reaches 0.77. The presented method can keep higher overlap rate and recall scores with fewer windows because of its coarse-to-fine process.

**Key words:** object localization; visual saliency; superpixel; sliding window; object recognition

## 0 引言

近年来,物体识别成为了计算机视觉领域中一个重要的研究方向,并且取得了长足的进步。物体识别一般可以分为两个阶段:定位和识别。在较为出色的物体识别算法中,大多数定位方法主要是滑动窗口的方法<sup>[1-3]</sup>。滑动窗口算法的最大优势就在于它使用简便,不需要在定位阶段耗费大量的计算资源。然而,一般滑动窗口会产生多达几十万、上百万的搜索窗口并且对窗口长宽比有限制,而这会严重影响后期识别阶段的快速性和准确性<sup>[4]</sup>,因此在后续识别中,为了计算速度上的考虑,它们往往会采用弱特征加分类器级联的方式进行识别,不宜使用那些性能优秀的复杂特征和分类器,降低了识别的准确度。

另外一种与滑动窗口不同的定位方法是通过一些图像特征更为直接地去搜索可能的物体位置。这类方法将待处理的窗口数量降低到几千甚至几百个,为后续处理过程中使用更为强大的特征和分类器创造了条件。Wang 等<sup>[5]</sup>在其物体识别算法中就采用了一种选择性搜索办法,大大提高了物体识

别的准确度,但他同时也指出这种搜索办法也是计算的一个瓶颈。有目的的物体定位在图像处理中有广泛的应用前景,本文的目的是寻求定位数量、定位精度和计算时间上的平衡,力争以较少的定位窗口数量实现较高质量的定位。

## 1 相关研究

在物体定位方法上,存在两大研究方向:一种是针对滑动窗口的优化;一种是探寻更为直接的物体定位算法。在滑动窗口优化研究中,Lampert 等<sup>[6]</sup>提出了分支界定技术,使用表观特征模型指导搜索,直接在图像中选出最优的物体定位,但是这种方法在非线性分类器中依然会产生 100 000 个搜索窗口。Felzenszwalb 等<sup>[2]</sup>提出的基于部件的物体定位方法也取得了较大的成功,该方法的效果对物体和其部件的组合有较强的依赖。

上述方法都是针对特定的物体类别进行的优化,Alexe 等<sup>[7]</sup>提出了一种方法来定量地衡量一个窗口中有多大的可能性会包含有一个物体,而且这个物体可能属于任何一种物体类别,即它是类独立的。该方法首先学习正确物体窗口的

收稿日期:2014-07-31;修回日期:2014-09-19。

**作者简介:**邵明正(1986-),男,山东济宁人,硕士研究生,主要研究方向:计算机视觉、图像处理;齐剑锋(1969-),男,河北石家庄人,副教授,博士,主要研究方向:机器学习、计算机视觉;王希武(1966-),男,河北衡水人,副教授,博士,主要研究方向:数据挖掘。

特征,训练出分类器之后,将其用在滑动窗口所产生的大量窗口上进行筛选,窗口的数量大大降低。Vedaldi 等<sup>[8]</sup>提出了跳跃窗口技术,使用词袋模型中的视觉词去预测物体的位置。Endres 等<sup>[9]</sup>利用性能出色的边缘检测算子<sup>[10]</sup>对一幅图像进行多次分割,根据之前训练的特征计算一个分割区域包含物体的概率,然后进行排序以得到物体可能位置;但是该方法需要耗费极大的计算资源,在物体识别的实际应用中尚不能有效利用。

以上方法多是通过正确事例的学习而训练出分类器,将其用于位置预测。而 van de Sande 等<sup>[4]</sup>所提出的选择性搜索(Selective Search)方法则利用图像的表现特征,采用不同的组合策略,将图像中的各个区域进行趋于完整物体的多次组合。该方法在定位精度上达到了新的高度,但是其定位窗口数量仍然有减少的空间。

人类的视觉系统可以在不知道图像中究竟有什么物体的前提下,就可以将图像中的各个元素组成一个可能的物体区域<sup>[11]</sup>。人在观察一幅图像的时候,往往是先对图像中的物体产生总体映像,然后才会根据需要进行定位到具体位置,再仔细地观察物体的颜色、轮廓等细节特征。在这个过程中,人眼对于其他的位置并不会过多地注意。基于以上考虑,本文提出一种基于视觉显著性和超像素融合的方法,先粗略定位图像中的物体;然后逐步细化,由粗到细地确定物体可能的位置。

## 2 本文的物体定位方法

### 2.1 总体框架

视觉显著性检测在人的视觉系统中是一个选择处理过程,目的是定位图像中重要的区域或物体<sup>[12]</sup>。被检测出来的图像显著部分很有可能包含了一个物体或者物体的某一部分,本文利用这一点实现物体的粗略定位。图1是利用文献[13]的方法实现的一个视觉显著性检测的例子,从图1中可以看出,图1(a)中的“汽车”在图1(c)中可以找到大概位置。但是仅仅知道大概定位是不够的,还需要知道物体的具体范围,也就是包围框。图1(d)是文献[13]利用显著性图像所进行的物体检测,可以看到其检测出的物体并不完全,这在物体识别应用中是要避免的。

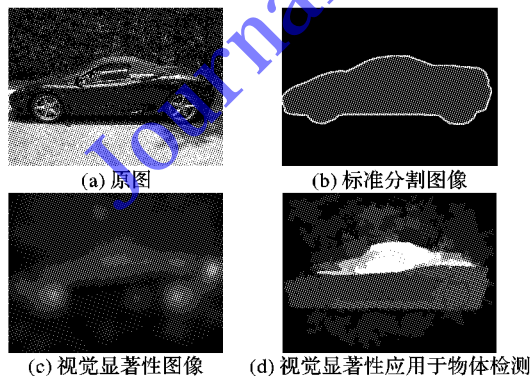


图1 视觉显著性图像应用于物体定位

因此在利用视觉显著性图像确定了粗略位置后,还要进行物体范围的细化。这里本文选择超像素融合的方法,虽然生成超像素本身也要消耗一定的资源,但是超像素包含有更多的局部信息,而且可以避免对单个像素的冗余操作<sup>[14]</sup>。结合图像的表现特征,本文方法从显著性图像确定的粗略位置

开始超像素融合,将超像素尽量向一个完整的物体进行融合。同时,为了避免融合过程中“泄漏”现象的发生,本文采用了一种简单的背景判断方法来抑制超像素的过度扩散。最后,通过多种组合来解决多尺度的问题并确定出物体的最终位置。整个处理过程如图2所示。

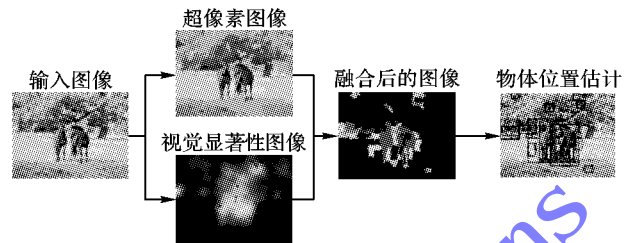


图2 本文方法的处理流程

### 2.2 物体粗略位置估计

为了实现粗略定位,本文所提的办法需要先从前视觉显著性图像中获得若干初始定位点。显然,显著性图像中的区域性表示并不能直接使用,但是图上的灰度等级表示了显著程度,通过设置灰度值门限来筛选显著部分;同时为了避免重复采样,本文通过洪泛算法来连通一定灰度值范围内的像素<sup>[13]</sup>,每个连通域确定一个初始定位点  $S(i)$ ,最后生成初始点集  $S$  作为后续超像素融合的初始位置。在算法实现中,本文选用了文献[13]的基于布尔图的显著性检测方法中 eye fixation prediction 模式,该方法简单快捷,是人眼关注点预测方法中较为出色的一个。

### 2.3 超像素融合方法

#### 2.3.1 超像素分割

超像素是指具有相似纹理、颜色、亮度等特征的相邻像素构成的图像块<sup>[15]</sup>。进行超像素分割的算法有很多,比如基于图论的方法<sup>[16]</sup>、规范图割方法<sup>[17]</sup>等,考虑到后续处理过程中对各融合因素的使用,本文选择 SLIC (Simple Linear Iterative Clustering)<sup>[18]</sup> 分割方法作为前期过分割算法,主要是因为 SLIC 算法生成的超像素形状规则、紧凑,有较好的边界查全率和较低的欠分割错误,它的计算有效性(算法复杂度是  $O(N)$ ) 对于本文方法的前期处理是较为有利的。

考虑 SLIC 算法的参数设置。本文对边界查全率有较高的要求,所以希望过分割越细越好,但同时过小的超像素会失去局部特征的优势并影响后续计算效率。综合分析,每个超像素的大小设为 300 个像素,紧凑度因子(compactness)设为 15。

#### 2.3.2 多点出发的超像素融合

由视觉显著性确定的物体概略位置确定融合的初始超像素,将一个物体内的不同超像素融合在一起。通过分析图像特征来确定哪些相邻超像素可以进行融合,本文使用颜色、纹理和边界强度这 3 个因素。

1) 颜色距离  $d_c$ 。衡量两个超像素之间的颜色差别程度。本文使用颜色直方图衡量颜色差别,在  $L^*a^*b^*$  颜色空间上,分别统计出每一个超像素在 3 个颜色通道的颜色直方图,然后将 3 个颜色直方图合并成一个。为了使用直方图相交法比较两个超像素的距离,将直方图作归一化处理,即  $c_i = \{c_i^a, c_i^b, c_i^c\} = \{c_i^1, c_i^2, \dots, c_i^n\}$ , 且  $\sum_{k=1}^n c_i^k = 1$ 。则,相邻两个超像素  $sp_i$  和  $sp_j$  的颜色距离为:

$$d_c(i, j) = 1 - \sum_{k=1}^n \min(c_i^k, c_j^k) \quad (1)$$

显然,  $d_c \in [0, 1]$ , 其值越小表示相似程度越高。根据实验对比, 设置直方图箱子的个数: L 通道取 20, a、b 通道各取 32, 于是  $n = 84$ 。

2) 纹理距离  $d_T$ 。衡量两个超像素之间的纹理差别程度。本文使用局部二值模式 (Local Binary Pattern, LBP)<sup>[19]</sup> 来衡量一个超像素的纹理特征。LBP 纹理特征简单有效, 经扩展的 LBP 具有旋转不变性。具体来说, 在 L 通道上使用统一形式的 LBP 生成具有 59 个箱子的 LBP 直方图来代表超像素的纹理特征; 同颜色特征一样, 将直方图作归一化处理, 即  $\sum_{k=1}^n t_i^k = 1$ 。在比较直方图方法上, 仍然选用直方图相交法, 如式(2):

$$d_T(i, j) = 1 - \sum_{k=1}^n \min(t_i^k, t_j^k) \quad (2)$$

其值域范围与颜色距离一致。

3) 边缘距离  $d_B$ 。边缘距离指的是两个超像素在分界线处的平均边缘强度。假设两个超像素  $sp_i, sp_j$  之间分界线的像素个数为  $n_b(i, j)$ , 用边缘检测算子提取两个超像素的边缘强度, 对于每一个分界线像素  $k$  其边缘强度为  $pb_k(i, j)$ , 则两个相邻超像素的边缘距离为:

$$d_B(i, j) = \left( \sum_{k=1}^{n_b(i, j)} pb_k(i, j) \right) / n_b(i, j) \quad (3)$$

遗憾的是, 带有强度计算的边缘检测算子计算量都是非常大的, 比如文献[10]算法, 虽然是在较小的超像素区域上进行计算, 但是极高的复杂度依然不符合本文对算法有效性的需求。所以, 简单起见, 本文用被检测出的边缘像素数目表示边缘强度, 如果用  $n_e(i, j)$  表示由检测算子检测出的边缘像素数目, 则两个超像素的边缘距离为:

$$d_B(i, j) = n_e(i, j) / n_b(i, j) \quad (4)$$

经实验, Laplacian of Gaussian 边缘检测算子在本文方法中的应用效果较好, 算子使用默认参数。同上,  $d_B \in [0, 1]$ ,  $d_B$  越小表示两个超像素之间的边缘越不明显,  $d_B = 0$  表示两个超像素之间没有边缘。在具体实现中, 由于超像素分割算法并不能保证两个超像素的分界线一定落在真实边缘上, 所以在搜索边缘像素时, 对每一个分界线像素, 在其周围的  $3 \times 3$  的区域内搜索最强边缘像素。

4) 距离组合。计算出各特征距离之后, 将各特征距离进行整合, 以确定出一对超像素之间最终的距离。按式(5)计算:

$$D(i, j) = \omega_c d_c(i, j) / \theta_c + \omega_T d_T(i, j) / \theta_T + \omega_B d_B(i, j) / \theta_B \quad (5)$$

其中:  $\theta_c, \theta_T, \theta_B$  分别为颜色、纹理和边界特征各自衡量相似程度的门限;  $\omega_c, \omega_T, \omega_B$  分别为 3 种特征在距离计算中所占的权重。颜色相近, 纹理相近, 边缘较弱的区域属于同一个物体。根据权重设置关系, 总距离门限  $\theta$  可以设为 1,  $D(i, j) < \theta$  说明两个超像素可以进行融合。

5) 背景因素。不同图像在相同的特征上的差异程度并不相同, 上述提出的门限方法并不能确保超像素只在同一物体范围内进行融合, 一旦超像素扩展到其他物体区域上, 则分割就会产生较大误差, 即发生融合时的“泄漏”。为了避免这样

的情况, 增加一个是否为背景的判断过程。

自然图像中, 物体大都集中在图像的中间区域<sup>[7]</sup>。在拍摄自然图像的时候, 人们往往会将突出的物体放置在图像的中间部分, 图像边缘通常是背景或背景的一部分。从自然图像的拍摄规律考虑, 本文认为图像四周边缘为背景的可能性较大。如图 3 所示, 在每个边缘上取 3 个超像素, 共获得 12 个超像素  $sp_{bg}$ 。将待融合的超像素  $sp_i$  与 12 个  $sp_{bg}$  进行比较, 如果有 2 个以上不在同一边缘上的  $sp_{bg}$  与  $sp_i$  相似, 则认为  $sp_i$  也为背景元素, 不进行融合。显然, 本文的背景因素考虑方式更适用于人主动拍摄的自然图像。



图3 背景元素提取示意图

另外, 为了保证算法的查全率 (recall), 对于待融合超像素  $sp_i$ , 如果判断其既不在融合门限内也不是背景元素, 则认为其可能是一个新的物体, 所以将该超像素加入融合初始集中进行新一轮的融合。依据以上分析, 从多个初始点出发的超像素融合算法描述如下。

算法 1 多点出发的超像素融合算法。

输入: 邻接表  $labelsList$ , 初始超像素队列  $S$ , 辅助队列  $Q$ , 融合标志数组  $merged[]$ 。  
输出: 融合结果链表  $Seg$ 。  
将  $Q$  设为空队列, 将数组  $merged$  置假  
repeat  
  将  $SP_o$  设为队列  $S$  的队首元素, 并从队列中移除队首元素  
  if  $SP_o$  没有被融合  
     $merged[SP_o]$  置真  
    将  $SP_o$  加入结果链表  $Seg$   
    将  $SP_o$  加入  $Q$  队尾  
  repeat  
    将  $SP_h$  设为队列  $Q$  的队首元素, 并从队列中移除队首元素  
    for 对于每一个  $SP_h$  的相邻超像素  $SP_i$  (由  $labelsList$  决定)  
      if  $SP_i$  没有被融合  
        if  $SP_i$  与  $SP_h$  的距离小于门限并且  $SP_i$  不是背景元素  
          将  $SP_i$  加入  $Seg$  中  
          将  $SP_i$  加入  $Q$  队尾  
           $merged[SP_i]$  置真  
        elseif  $SP_i$  不是背景元素  
          将  $SP_i$  加入队列  $S$  队尾  
      end if  
    end if  
  end for  
until 队列  $Q$  为空  
end if  
until 队列  $S$  为空

## 2.4 区域组合

过分割图像经本文所提出方法融合之后, 有意义的区域数量减少为几十至数百, 这极大减小了物体定位所需考虑的范围。为了从这些区域中找到可能的物体位置, 本文采用了



文献[4]所提出的贪心算法进行区域组合,在具体实现中,上述提到的边缘强度特征也应用到组合策略中,因为边缘强度在后续区域的组合中仍然能够有效区分不同的区域。

### 3 实验分析

用 Matlab R2013a 实现了本文所提算法,算法中视觉显著部提取、LBP 特征计算采用了相应文献的作者所提供的源代码。实验机器配置为 2.5 GHz CPU + 4 GB RAM。

#### 3.1 评价指标

1) 查全率 (Recall)。在物体定位中,查全率是衡量图像中的所有物体被正确定位出的比率。在 Pascal 重叠标准中,只要式(7)的值大于 0.5 就认为物体被定位出来。

2) 综合平均最佳重叠率 (Mean Average Best Overlap, MABO)。MABO 是文献[4]将平均最佳重叠率 (Average Best Overlap, ABO) 扩展到一般类上的衡量定位质量的指标。针对某一物体类  $c$ , 有:

$$ABO = \frac{1}{|G^c|} \sum_{g_i^c \in G^c} \max_{l_j \in L} \text{Overlap}(g_i^c, l_j) \quad (6)$$

其中:  $G^c$  为物体类  $c$  全部的实际位置集合,  $L$  为算法所估计的定位集合。Overlap 按式(7) 计算:

$$\text{Overlap}(g_i^c, l_j) = \frac{\text{area}(g_i^c) \cap \text{area}(l_j)}{\text{area}(g_i^c) \cup \text{area}(l_j)} \quad (7)$$

其中 area 表示包围盒的面积,即像素数目。

MABO 是将 ABO 在所有的物体类别上求平均。

#### 3.2 结果分析

为了评估本文所提方法的定位质量,本文在具有挑战性的 Pascal VOC 系列数据集上进行实验。

首先在 Pascal VOC 2007 数据集的测试集中,将本文方法与文献[4]方法、文献[7]方法和文献[11]方法进行比较实验。实验中所用程序是文章作者在網上公开的源代码,实验结果如表 1 和图 4 所示。

表 1 是对 4 种方法的 MABO 得分、窗口数量和运行时间的比较。其中:文献[7]方法中用的组合因素为 MS + CC + SS;文献[4]方法用了两种颜色空间、两种不同的初始分割门限及两种组合策略。从表 1 中可以看出,文献[4]方法的 MABO 得分最高(0.804),但同时它所用的定位窗口也是最多的(2 134 个),相比较而言,本文方法用 658 个窗口就达到了 0.789 的 MABO 得分,逼近了文献[4]方法的定位准确度。虽然,本文方法为了融合超像素而消耗了一定计算时间,但却减少了 60% 的定位窗口数量,这种平衡给予了用户更大的选择空间。另外,文献[11]方法实现了更少的定位窗口数量(517 个),但是它的检测时间却达到了 425 s,这在追求实时性的应用中是不可容忍的。

表 1 MABO、定位窗口数量和运行时间的对比

方法	MABO	窗口数量	运行时间/s
文献[7]方法	0.694	1 853	3.57
文献[4]方法	0.804	2 134	4.01
文献[11]方法	0.770	517	425.00
本文方法	0.789	658	8.12

图 4 显示的是 4 种方法的查全率和定位窗口数量之间的关系。从图 4 中可以看出,本文方法在窗口数量较少时查全

率要比其他方法高,比如定位窗口数量为 500 时,本文方法的查全率约为 0.920,而其他 3 种方法的查全率分别为 0.915, 0.870, 0.865,这与 MABO 检测结果是一致的。由于本文方法采用视觉显著性来估计物体粗略位置,所以当窗口数量较少时,包含有初始位置的窗口仍能捕捉到明显的物体位置。另外,由于本文方法在窗口数量上的限制,最终查全率没有文献[4]方法和文献[7]方法高,只能达到约 92% 的查全率,因此扩展窗口数量以适用更多的应用也是未来的研究工作。

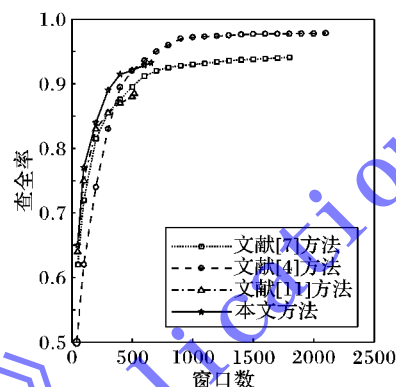


图 4 不同定位窗口数量下的查全率变化曲线

为了评估本文方法针对不同物体的通用性,从图像范围更为广泛的 Pascal VOC 2012 数据集的训练集中选取部分人、自行车、飞机、狗和鸟的图像构成本文第 2 个测试数据集,每种物体类别选取 200 张图像,共 1000 张图像组成测试集。实验将本文方法与文献[4]方法、文献[7]方法对于 MABO 得分、窗口数量和不同类别物体的平均最佳重叠率 (ABO) 作了对比,结果在表 2 和图 5 中表出。表 2 显示的结果与表 1 是一致的,用 612 个窗口达到了 0.769 的 MABO 得分和 0.919 的查全率,反映了本文方法对于减少定位窗口数量是有效的。由于 VOC2012 中的图像尺度变化较大、内容复杂,各方法检测出的结果与 VOC 2007 数据集上的实验结果相比均有所下降。值得注意的是文献[7]方法,其查全率达到了 0.933,但是其 MABO 得分只有 0.655。这是因为 Pascal 标准规定:只要窗口重叠率达到 50%,就认为物体被检测到。显然,文献[7]方法以并不是精确的定位实现了较高的查全率,这对于后续识别处理是不利的。而本文方法以图像分割为基础,定位窗口多以 70% 以上的重叠率检测出了物体位置,体现了较高的定位精度。

表 2 在 Pascal VOC 2012 上的定位质量对比

方法	Recall	MABO	窗口数量
文献[7]方法	0.933	0.655	1 971
文献[4]方法	0.971	0.783	2 204
本文方法	0.919	0.769	612

图 5 显示的是 3 种方法针对不同类别物体的平均最佳重叠率,从中可以看出,本文的方法在提取窗口数量较少的情况下,ABO 得分基本能达到其他两种优秀方法的水平,但是该方法在某些类上,如自行车类,表现不佳,应该是和自行车图像在表现特征上以及场景元素上比较复杂有关,不利于超像素的融合。

### 4 结语

本文提出了一种基于视觉显著性和超像素融合的物体定

位方法。由视觉显著性确定物体粗略位置;然后利用图像特征尽量将属于同一物体的超像素进行融合,避免后续无意义的组合。本文方法减少了物体定位窗口的数量,为物体识别阶段使用更为复杂精确的模型奠定了基础。但是,本文方法在处理较复杂场景和物体时定位不准确、查全率不高,这也是下步将要解决的问题。同时,更进一步地减少物体位置假设的数量是这个领域不断追求的目标。

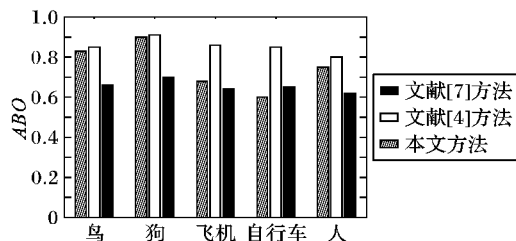


图5 不同类别物体的 ABO

#### 参考文献:

- [1] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection [C]// Proceedings of the 18th IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2005: 886–893.
- [2] FELZENSZWALB P F, GIRSHICK R B, McALLESTER D, *et al.* Object detection with discriminatively trained part based models [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(9): 1627–1645.
- [3] ZHU L, CHEN Y, YUILL E, *et al.* Latent hierarchical structural learning for object detection [C]// Proceedings of the 23rd IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2010: 1062–1069.
- [4] van de SANDE K E A, UIJLINGS J R R, GEVERS T, *et al.* Segmentation as selective search for object recognition [C]// Proceedings of the 13rd IEEE International Conference on Computer Vision. Piscataway: IEEE, 2011: 1879–1886.
- [5] WANG X, YANG M, ZHU S, *et al.* Regionlets for generic object detection [C]// Proceedings of the 14th IEEE International Conference on Computer Vision. Piscataway: IEEE, 2013: 17–24.
- [6] LAMPERT C H, BLASCHKO M B, HOFMANN T. Efficient sub-window search: a branch and bound framework for object localization [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(12): 2129–2142.
- [7] ALEXE B, DESELAERS T, FERRARI V. Measuring the objectness of image windows [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(11): 2189–2202.
- [8] VEDALDI A, GULSHAN V, VARMA M, *et al.* Multiple kernels for object detection [C]// Proceedings of the 12nd IEEE International Conference on Computer Vision. Piscataway: IEEE, 2009: 606–613.
- [9] ENDRES I, HOIEM D. Category independent object proposals [C]// Proceeding of the 11st European Conference on Computer Vision, LNCS 6315. Berlin: Springer, 2010: 575–588.
- [10] ARBELAEZ P, MAIRE M, FOWLKES C, *et al.* Contour detection and hierarchical image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(5): 898–916.
- [11] CARREIRA J, SMINCHISESCU C. Constrained parametric min-cuts for automatic object segmentation [C]// Proceedings of the 23rd IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2010: 3241–3248.
- [12] YAN Q, XU L, SHI J, *et al.* Hierarchical saliency detection [C]// Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2013: 1155–1162.
- [13] ZHANG J, SCLAROFF S. Saliency detection: a boolean map approach [C]// Proceedings of the 14th IEEE International Conference on Computer Vision. Piscataway: IEEE, 2013: 153–160.
- [14] FULKERSON B, VEDALDI A, SOATTO S. Class segmentation and object localization with superpixel neighborhoods [C]// Proceedings of the 12nd IEEE International Conference on Computer Vision. Piscataway: IEEE, 2009: 670–677.
- [15] WANG C, CHEN J, LI W. Superpixel segmentation algorithms review [J]. Application Research of Computers, 2014, 31(1): 6–12. (王春瑶, 陈俊周, 李炜. 超像素分割算法研究综述[J]. 计算机应用研究, 2014, 31(1): 6–12.)
- [16] FELZENSZWALB P F, HUTTENLOCHER D P. Efficient graph-based image segmentation [J]. International Journal of Computer Vision, 2004, 59(2): 167–181.
- [17] SHI J, MALIK J. Normalized cuts and image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(8): 888–905.
- [18] ACHANTA R, SHAJI A, SMITH K, *et al.* SLIC superpixels compared to state-of-the-art superpixel methods [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(11): 2274–2282.
- [19] OJALA T, PIETIKAINEN M, MAENPAA T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7): 971–987.
- [10] YUAN G, YE F, YANG L, *et al.* Study of medical image interactive 3-D measurement based on VTK [J]. Computer Engineering and Design, 2008, 29(13): 3549–3550. (袁皋, 叶峰, 杨玲, 等. 基于 VTK 的医学图像交互式三维测量技术研究[J]. 计算机工程与设计, 2008, 29(13): 3549–3550.)
- [11] YANG P. The spline curve study of highway horizontal alignment design [D]. Chongqing: Chongqing Jiaotong University, 2011. (杨鹏. 样条曲线在道路平面线形中的应用研究[D]. 重庆: 重庆交通大学, 2011.)
- [12] HU X. The study of segmentation algorithm of heart CT image [D]. Beijing: Beijing Jiaotong University, 2011. (胡晓清. CT 心脏图像分割方法研究[D]. 北京: 北京交通大学, 2011.)
- [13] YANG Q. How to play video at its original FPS in OpenCV [J]. China Science and Technology Information, 2012(21): 75–75. (杨青锦. OpenCV 下按视频帧率(FPS)播放视频[J]. 中国科技信息, 2012(21): 75–75.)
- [14] LYU X, ZHANG C, HOU H, *et al.* Multi-function rendering technology based on graphics process unit accelerated ray casting algorithm [J]. Journal of Computer Applications, 2014, 34(1): 135–138. (吕晓琪, 张传亭, 侯贺, 等. 基于图形处理器加速光线投射算法的多功能体绘制技术[J]. 计算机应用, 2014, 34(1): 135–138.)

(上接第 214 页)