

文章编号 1004-924X(2014)01-0160-09

## 图像通用目标的无监督检测

宋修锐<sup>1,2</sup>, 吴志勇<sup>1\*</sup>

- (1. 中国科学院 长春光学精密机械与物理研究所, 吉林 长春 130033;
2. 中国科学院大学, 北京 100049)

**摘要:** 为了实现对图像中多种类目标的检测, 缩短目标搜索时间, 本文基于图像目标的 3 个显著性线索(显著性检测, 颜色对比, 超像素跨越), 构建了一种改进的通用无监督目标检测模型。通过机器学习 center-surrounding 比例参数, 计算各个线索的显著度得分, 并在朴素贝叶斯框架下对这 3 个目标显著性线索进行融合, 以最终确定窗口中包含图像目标的概率。实验参数在 PASCAL VOC 2007 图像库进行检测, 检测率为 28.94%, 击中率达 96.99%; 在 MSRC 图片库进行检测, 检测率为 80.64%, 击中率达 99.10%; 得到的结果证明了本文模型的通用性。另外, 该模型对单幅图像的处理时间较 Bogdan 的检测模型提高了 40%, 改进了目标检测效率。本文模型可为后续的目标识别, 图像分割提供更快、更准确的先验位置信息。

**关键词:** 无监督检测; 显著性检测; 颜色对比; 超像素跨越; 贝叶斯融合

**中图分类号:** TP391.4 **文献标识码:** A **doi:** 10.3788/OPE.20142201.0160

## Unsupervised detection of image object with any class

SONG Xiu-rui<sup>1,2</sup>, WU Zhi-yong<sup>1\*</sup>

- (1. *Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China;*
  2. *University of Chinese Academy of Sciences, Beijing 100049, China)*
- \* *Corresponding author, E-mail: wzy0607@yahoo.com.cn*

**Abstract:** To measure a variety of objects of an image and to reduce the detection time, an unsupervised object detection model was established to provide location priors. The model was mainly based on three image cues of a object, and they are saliency detection, color contrast and superpixel straddling. To determine the likelihood of image object contained in a window, the saliency scores of the three cues were calculated, and the saliency cues of the three objects were fused in a simple Bayesian framework by a machine learning center-surrounding proportion parameter. In experiments on the challenging PASCAL VOC 07 dataset, it shows that the detection rate is 28.94%, the hit rate is 96.99% and the combined measuring result is better than any cue alone. In experiments on MSRC dataset, it shows that the proposed model is generic and efficient, the detection rate is 80.64%, the hit rate is 99.10% and the average processing time is 40% less than that of Bogdan's model. These results from extensive field tests suggest that proposed model can provide better location priors to the

收稿日期: 2013-02-11; 修订日期: 2013-03-15.

基金项目: 国家 863 高技术研究发展计划资助项目 (No. 2008AA7034320B)

object recognition and image segmentation where the location of object is unknown.

**Key words:** unsupervised detection; saliency detection; color contrast; superpixel straddling; Naive Bayes fusion

## 1 引 言

图像目标检测是图像分割的初级阶段,它提供图像中最可能覆盖目标的小数目的窗口作为目标位置先验信息,广泛应用于图像检索、目标识别、辅助驾驶系统以及视觉监控等,近年来已成为图像处理一个重要的研究方向。在此类研究中,目标的外观建模研究<sup>[1-2]</sup>推动了目标标记的发展;Li Jiali 和 Li Feifei 利用增量图模型以及半监督学习方法搜集定位前景目标<sup>[3]</sup>;C. Gu 通过相似度学习形成分类器,判断区域标记<sup>[4]</sup>;中科院模式识别重点实验室利用多特征 Boost 结合上下文信息对目标进行建模;以上这些方法仅在目标可能区域作出标记框,局限在于标记框内不可避免地包含了部分的背景信息。单类目标定位中还有一类常用的方法是将目标识别与图像分割相结合<sup>[5-6]</sup>,然而这类方法受限于目标的外观,轮廓,形状变化,受遮挡、光线、视角变换、远近程度等因素影响很大。为解决这一问题,Fulkerson 利用超像素代替目标区域进行标记预测<sup>[7]</sup>,进一步借助条件随机场推导超像素之间的空间关系,最终将图像目标提取出来。然而该方法的局限性在于其性能依赖于超像素的邻域参数的选择,而这又是一个更加有难度的问题。

目前的图像目标检测方法主要有有监督学习、弱监督学习和无监督学习 3 类。有监督学习需要标注先验信息(如目标物体位置、种类等)以对图像目标进行分类识别,在某些无法获知先验信息的情景下不适用;弱监督学习利用图像数据的内在分布特点,需要标定少量先验信息以得到更准确的分类模型,在已标注的训练信息不足而未标注的信息却大量存在的情况下效果较好;无监督学习不需要任何的先验信息,训练学习后自主进行图像数据分析,最终得到目标检测结果。本文进行了无监督学习的目标检测。

受人类视觉系统的启发,图像显著性体现人眼对于图像中目标物体的重视程度。虽然目前已

有的目标检测分类方法已经很多,但是目前的图像目标检测都主要针对某种特定的目标类进行检测识别。无关于目标类别,本文构建了基于图像目标显著性特征的通用目标检测模型,并在 PASCAL VOC 2007 图像库<sup>[16]</sup>和 MSRC 图像库进行了测试。通用目标检测可在目标位置未知的情况下,快速计算目标候选区域,这在遥感图像处理,视频监控系统中进行目标跟踪定位中具有重大意义,也可为进一步的图像分割,图像自适应压缩,内容敏感图像编辑,图像检索提供目标定位。

## 2 基于目标显著性的图像线索

图像目标有颜色、边缘、显著性等特征,Bogdan 提出所有的目标都至少含有以上 3 条特征中的一条,根据以上图像目标特征,总结出 4 个用于量化目标特征的图像线索:多尺度显著性,颜色对比,超像素跨越,边缘密度<sup>[8]</sup>。本文的目标检测采用了前 3 个线索联合作用,重点对第一和第三个线索进行了优化改进,并对最终结果进行了合理整合。

### 2.1 显著性检测(SD)

图像显著性是图像目标重要的视觉特征,体现了人眼对图像的目标区域的重视程度。目前存在大量的显著性映射方法,鉴于本文对于目标整体性的要求,采用了整体性较好的频率调整的显著区域检测(IG 算法)方法<sup>[9]</sup>。

频率调整的显著区域检测的算法思想如图 1 所示。首先将图像转换为 Lab 色彩空间,进行高斯滤波,然后计算各个通道的均值,再计算两者的欧式距离,最终得到图像的显著性图。

文献[8]采用了文献[10]中提出的基于空间频域分析的剩余谱算法(SR 算法),这种算法方法简单易实现,但是效果粗糙,整体性差,甚至可能出现误判。图 2 是本文显著性图获取方法和文献[10]提出方法对比,实验结果比较发现,本文方法整体性更好,且不需要多尺度,大大减少了处理时间,提高了处理效率。

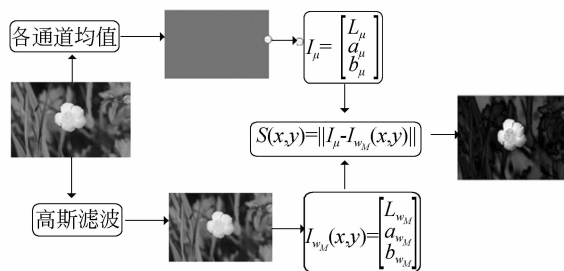


图 1 频率调整的显著区域检测算法流程图

Fig. 1 Flow of frequency-tuned salient region detection

使用文献[9]得到显著性图定义每一个像素  $p$  的显著性。基于显著性图,窗口  $w$  的显著性得分定义如式(1)。

$$SD(W, \theta_{SD}) = \frac{\sum_{\{p \in w \mid I_{SD}(p) \geq \theta_{SD}\}} I_{SD}(p) \times |\{p \in w \mid I_{SD}(p) \geq \theta_{SD}\}|}{|w|}, \quad (1)$$

图 2 频率调整的显著区域检测<sup>[9]</sup>和基于空间频域分析的剩余谱算法<sup>[10]</sup>

Fig. 2 Frequency-tuned salient region detection maps vs saliency detection maps of spectral residual approach

其中: 阈值  $\theta_{SD}$  是后文学习得到的参数,  $|w|$  表示

窗口中的像素数。窗口中显著的像素点密集度越高, 整体显著性就越高, 则得分会倾向于面积较大的窗口; 如果只比较密度, 则含有几个离散高显著点的窗口会得分偏高。而本文的目的在于使整个高显著块附近的相关窗口得分最高, 这样是对整个目标进行作用, 所以将其映射到显著性图。图 3 显示了 SD 线索单独作用的图像目标检测结果。

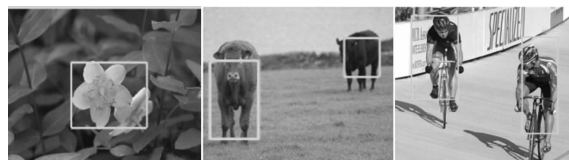


图 3 显著性检测线索单独作用图

Fig. 3 Results of saliency detection

SD 线索体现了目标的显著性特征, 对图像目标进行了初步定位。如果 SD 单独作用, 对于某些显著性较高的图像效果较好, 但是对于某些背景比较复杂的图像会效果差些。

## 2.2 颜色对比(CC)

颜色对比的图像线索来源于图像目标与周围环境的颜色差异特征。Lab 直方图描述的是不同色彩在整幅图像中所占的比例, 而并不关心每种色彩所处的空间位置。由于图像上的前景物体和背景颜色分布差异很大, 从而在直方图上会出现双峰特征, 但背景和前景颜色较为接近的图像不具有这个特征。

窗口  $w$  的周边  $Surr(w, \theta_{CC})$  是将窗口在各个方向上按照 center-surrounding 参数  $\theta_{CC}$  (学习得到) 等倍数扩大得到的矩形环。所以:

$$\frac{|Surr(w, \theta_{CC})|}{|w|} = \theta_{CC}^2 - 1. \quad (2)$$

一个窗口与其周边环境的颜色对比度 CC 如公式(3), 由窗口与其周围环境 Lab 直方图的卡方距离度量。

$$CC(w, \theta_{CC}) = \chi^2(h(w), h(Surr(w, \theta_{CC}))). \quad (3)$$

颜色对比线索中, 如果窗口中颜色与周围环境颜色差异越大, 则窗口中含目标物体的可能性越大。图 4 显示了颜色对比线索单独作用效果图。

颜色对比仅对图像目标与背景颜色差异大的图像起作用, 如果图像目标自身颜色多变, 或者与周围环境相似, 则此时颜色对比线索失效。



图 4 颜色差异线索单独作用效果图

Fig. 4 Results of color contrast

### 2.3 超像素跨越(SS)

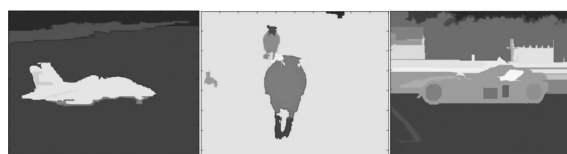
获取目标闭合边界的特征的方法是使用超像素作为线索。超像素可以将图像分割成不同颜色或纹理的小区域,它的一个重要的特征是保护目标边界。一个超像素中的所有像素点都属于同一个目标,即使一个目标物体被过分割为几个超像素,也不会有一个跨越它的边界。基于超像素图像分割的这个特性,得到目标检测的最后一个线索—超像素跨越。

$$SS(\omega, \theta_{SS}) = 1 - \sum_{s \in S(\theta_{SS}) \cap \{|\omega| < 0.6|I|\}} \frac{\min(|s/\omega|, |s \cap \omega|)}{|\omega|}, \quad (4)$$

其中: $S(\theta_{SS})$ 是通过文献[7]得到的超像素分割图中的超像素集。对于每一个超像素  $S$ ,计算此超像素与待测窗口交集  $|s \cap \omega|$ ,以及超像素位于待测窗口外部的区域  $|s/\omega|$ 。两者的最小值即为此超像素跨越窗口  $\omega$  的程度,然后计算所有超像素跨越程度的总和。

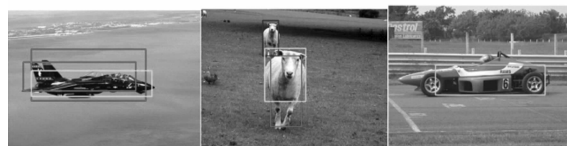
超像素完全位于窗口内部或外部对于(4)式中加和的贡献值为 0,所以 SS 将得分较高。对于一个跨越的超像素,当它大部分位于窗口内部或者外部,对于式(4)中加和的贡献也是比较低的。所以,SS 线索对于紧紧包围目标的窗口得分最高。但是这里存在一个问题,当窗口大小趋近于整张图像的大小时,所有的超像素均位于窗口内部,不存在超像素跨越窗口的问题,则所有的差集趋向于 0,所以 SS 得分趋近于满分 1。这对于本文的 SS 得分计算造成了干扰。为了避免过大窗口的出现,进行了多次验证,验证显示图像目标一般不会超过整张图像的 60%,所以对于大于图像 60%的窗口,做置零处理。这种处理方法可能会舍去某些较优窗口,但总体上效果会更好些。经过窗口面积筛选处理后的 SS 线索,在各个线索单独作用时效果最好。SS 单独作用实验结果如图 5 所示。

综上所述,以上 3 个基于图像目标特征的显著性线索对于目标的获取都是有一定效果的,但是由于各图像库图像的复杂性和差异性,3 个线索单独作用都有可能在某些情况下出现失效。SD 线索作用于成块的高显著区域,可以得到目标的大致位置;CC 线索提供更加准确的窗口,但有时会丢失整个与周围环境颜色差异小的目标;SS 线索虽然总体效果较好,但它严重依赖于超像素分割图像,且对于面积比较小的目标分割效果较差。为了得到更好的效果,本文尝试将这几个线索进行融合以达到最佳效果。



(a) 超像素分割图

(a) Superpixel segmentation images



(b) 超像素跨越单独作用结果

(b) Superpixel straddling detection results

图 5 超像素跨越单独作用目标检测结果

Fig. 5 Superpixel straddling detection results

### 2.4 基于贝叶斯框架的融合

以上 3 个图像线索是完全独立且不相互影响,那么同时使用其中若干个线索有可能提高目标检测效果。如果简单线性组合各个线索,由于结合线索过分单纯化,适用性较差。为了有效结合以上的 3 个线索,本文训练了一个贝叶斯分类器来区分 3 个线索的正样例与负样例。其中正样例与负样例定义如下:

正样例(positive examples):一个窗口与 Ground Truth 中目标窗口的交并比大于 0.5,则定义为正样例。简记为  $W^{obj}$ ;

负样例(negative examples):一个窗口与 Ground Truth 中目标窗口的交并比小于等于 0.5,则定义为负样例。简记为  $W^{bg}$ 。

在朴素贝叶斯模型中,各个线索是独立的。训练过程包含估测联合概率中正样例概率  $p(obj)$ ,负样例概率  $p(bg)$ 以及每个线索的先验概率  $p(cue|c)$ ( $c$ 包括  $obj$ 和  $bg$ )。

在测试图像集中对每一张测试图片首先根据

SD 线索得到任意数目的样本窗口,之后计算这些窗口在各个线索下的得分。一个窗口包含目标可能性的后验概率满足公式(5)。

$$p(obj | C) = \frac{p(obj)p(C | obj)}{p(C)}$$

$$p(obj | C) = \frac{p(obj)\prod_{cue \in C} p(cue | obj)}{\sum_{c \in \{obj, bg\}} p(c)\prod_{cue \in C} p(cue | c)}, \quad (5)$$

其中: $C$ 为线索集。式(5)的结果即为最终窗口包含目标的概率,它可以将任意线索自由组合而不需要重复计算概率。

经过贝叶斯融合后,得到各个窗口的得分,取前十名窗口发现,窗口均分布在目标物体附近,最高得分基本可以紧紧地包围目标。但是某些图像的检测效果较差,即使最高得分的窗口也不能紧紧地框住目标。在此本文做了整合处理,鉴于前十名窗口均分布在目标附近,已经基本将目标定位,故将前十名的窗口取均值,最终整合为一个窗口,此方法对于单目标图像检测效果极佳,很好地达到图像目标检测的目的。对于多目标图像,仍然选取前十个窗口进行定位,实验结果发现,前十名的窗口基本可以将所有目标覆盖,达到了多目标检测的目的。

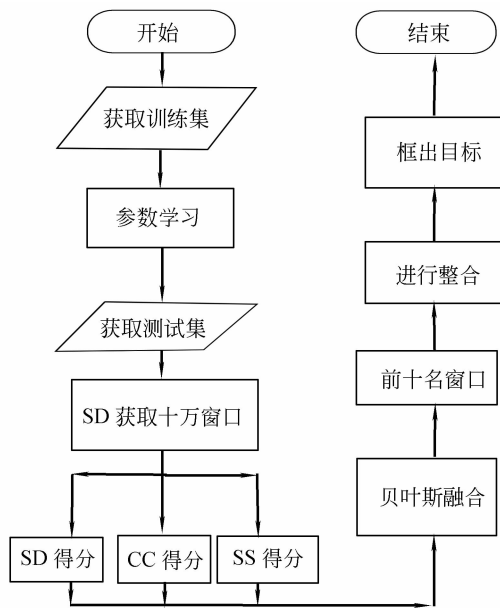


图 6 算法框架流程图

Fig. 6 Algorithm framework

### 3 算法框架

本文首先使用 PASCAL VOC 2007 图像库进行图像目标检测。VOC 2007 图像库中图像的目标都在 Ground Truth 中有矩形框数据记录,本文的窗口结果很适合与之进行比较。在一幅图像中可以形成上亿的随机窗口,但计算每一个窗口的得分时间复杂度很高,所以本文先进行初步定位,在最可能是目标的区域通过滑动窗口机制得到均匀覆盖在目标上的  $1 \times 10^5$  个窗口。通过反复实验发现,  $1 \times 10^5$  个窗口是最佳数目,足以覆盖目标且较省时。本文首先通过 SD 线索在目标区域附近得到  $1 \times 10^5$  个窗口,再计算各个线索关于这  $1 \times 10^5$  个窗口的得分,然后进行贝叶斯融合。最终结果输出前十名得分的框以及十个框整合形成的最终目标窗口。实验流程图如图 6 所示。

### 4 实验过程

本文使用 PASCAL VOC 2007 训练库学习参数。其中包括 5 011 张图像,7 380 个目标,共分为 20 类。使用 PASCAL VOC 2007 测试库(4 952 张图像,8 000 多个目标)及 MSRC 图像库(501 张图像,600 多个目标)进行测试。测试结果与 GROUND TRUTH 中标注目标进行比较。

在学习过程中,本文提取测试库中的目标窗口,而不是它们的类标,因为本文的检测方法是超越目标类别通用的。本文统计两组联合概率及 3 组先验概率,学习 3 个参数  $\theta_{SD}, \theta_{CC}, \theta_{SS}$ 。

正样例概率  $p(obj)$  由训练样本统计数据得到。将滑动窗口机制产生的窗口与 GROUND TRUTH 中的目标窗口进行比较,如果两者的交并比大于 0.5,则定义为  $W^{obj}$ ,否则定义为  $W^{bj}$ 。正样例概率及负样例概率定义如式(6)。

$$\begin{cases} p(obj) = |W^{obj}| / (|W^{obj}| + |W^{bj}|), \\ p(bg) = 1 - p(obj) \end{cases} \quad (6)$$

3 组先验概率采用直方图密度估计(density estimation)的方法得到。按照一定的步长将窗口得分数据值覆盖的数据区间分成几个等子区间;统计落在各个子区间中由各个线索决定的正样例的个数,得到 3 组先验概率。

#### 4.1 SD 线索参数学习

最优的阈值应使窗口与 VOC2007 库中训练图像的 Ground Truth 中提供的目标窗口交并比达到最大值。

$$\theta_{SD}^* = \operatorname{argmax}_{\theta} \sum_{o \in O} \max \frac{|w \cap o|}{|w \cup o|}. \quad (7)$$

使用(7)式确定的阈值可以使得显著性高的块尽量突出,便于目标较准确地初步定位。

#### 4.2 CC,SS 线索参数学习

本文使用贝叶斯公式学习 CC, SS 线索的 2 个参数。由于这两个参数使用相同的方法进行学习,所以以  $\theta = \theta_{CC}$  为例,其他两个参数的训练过程依此类推。对于每一张训练集中的图像,使用滑动窗口算法产生均匀分布在整张图像的十万个窗口。

最优的参数值应使窗口成为正样例的可能性最大。所以应满足(8)式:

$$\theta^* = \operatorname{argmax}_{\theta} \prod_{w \in W^{obj}} p_{\theta}(obj | CC(w, \theta))$$

$$\theta^* = \operatorname{argmax}_{\theta} \prod_{w \in W^{obj}} \frac{p(obj) \cdot p_{\theta}(CC(w, \theta) | obj)}{\sum_{c \in \{obj, bg\}} p_{\theta}(CC(w, \theta) | c) \cdot p(c)} \quad (8)$$

学习 SD 线索中阈值参数使得图像显著性高的区域块更加突出;学习参数  $\theta_{CC}^*$  得到最优的 center-surrounding 比例参数;学习  $\theta_{SS}^*$  定义了最优的超像素分割的尺度。各个参数的学习的目的都是一致的,都在于提高图像窗口覆盖目标物体的可能性,以便在将来的实验中更准确地目标检测定位。

## 5 实验结果

### 5.1 PASCAL VOC 2007 图像库测试

本文首先在目前视觉领域公认的最具有挑战性的 PASCAL VOC 2007 图像库进行测试。通过 PASCAL VOC 2007 测试图像库,测试了各个线索单独作用以及组合作用的效果,文中使用检测率以及击中率两项指标衡量检测效果。

检测率:检测目标窗口与测试库提供的 Ground Truth 中目标所在矩形框进行交集与并集的计算。如果交并比大于 0.5,则认为检测成功;如果交并比小于等于 0.5,则认为检测失败。

实验结果如表 1 和表 2 所示。其中,单个线索的作用效果不如多线索融合作用的效果。从表 1 和表 2 可以比较得到某些类的检测成功率偏高一些,比如飞机,马,车等,而瓶子,椅子等的检测成功率偏低,主要原因是他们的显著性本身比较低。最终结果表明单线索作用时,SS 线索单独作用检测率较其他两个高,为 18.72%;多线索联合作用时,SD,CC,SS 3 个线索联合作用总的检测率最高,为 28.94%。

表 1 单线索作用检测率表  
Tab.1 Detection rates of signal cue (%)

class	SD	CC	SS
Airplane	16.52	17.38	24.90
Bicycle	5.20	11.99	14.36
Bird	15.13	13.21	23.08
Boat	14.46	12.40	15.80
Bottle	5.78	3.11	4.3
Bus	9.89	8.24	21.69
Car	23.83	22.64	22.62
Cat	7.99	6.35	21.46
Chair	11.34	10.87	11.05
Table	12.06	11.6	16.22
Dog	16.90	12.10	18.50
Horse	18.39	13.96	30.89
Bike	8.65	8.20	20.61
Person	10.30	10.23	16.19
Sheep	13.66	13.77	15.18
Sofa	12.70	12.20	21.23
Train	10.64	12.10	20.62

为多角度评价本文改进的目标检测模型,本文又在 VOC2007 库上测试了 SD+CC+SS 3 个线索联合作用的图像目标击中率。

击中率:得分前十名的窗口分别与测试库提供的 Ground Truth 中目标所在的矩形框求交集与并集,计算交集与并集的比值。至少有一个窗口计算得到的交并比大于 0.5,则认为击中;所有窗口计算得到的交并比小于等于 0.5,则认为未击中。

表 2 多个线索联合作用检测率表  
Tab. 2 Detection rates of several cues (%)

class	SD+CC	SD+SS	SD+CC+SS
Airplane	29.10	39.58	41.83
Bicycle	10.47	18.17	24.60
Bird	16.42	18.61	18.76
Boat	16.10	15.34	27.97
Bottle	10.96	11.00	22.58
Bus	15.42	20.27	20.57
Car	31.43	33.45	46.53
Cat	20.34	20.00	31.23
Chair	12.01	11.80	12.11
Table	12.54	15.43	19.47
Dog	10.64	20.00	26.21
Horse	19.21	29.34	33.85
Bike	12.35	21.20	23.66
Person	10.65	15.79	19.33
Sheep	16.34	16.82	21.59
Sofa	13.22	20.96	23.10
Train	18.54	22.45	34.92

击中率结果如表 3 所示,SD+CC+SS 的击中率高达 96.99%,此击中率显示出本文模型的高效性。

表 3 SD+CC+SS 联合作用击中率  
Tab. 3 Hit rate of SD+CC+SS (%)

class	Hit-rate
Aeroplane	97.97
bicycle	95.30
bird	97.19
boat	95.79
bottle	96.70
bus	96.96
Car	95.93
Cat	96.35
Chair	94.44
table	95.28
Dog	96.71
Horse	96.88
Motorbike	95.42
Person	95.77
Sheep	97.87
Sofa	97.51
Train	96.31

图 7,图 8 分别为 PASCAL VOC2007 图像库中,SD+CC+SS 3 个线索联合作用下的图像单目标检测和多目标检测结果。从图中可以看出,单目标检测效果很好,整合后窗口能将图像中单目标紧紧框出。多目标检测效果比较好,能将图像中大多数的目标框出,定位基本准确。某些情况下将某些较深远的小目标遗漏,这依赖于后期阈值的选取。总之,对于 PASCAL VOC 2007 图像库,本文目标检测效果较好,本文在结论中与其他同类方法进行了比较。

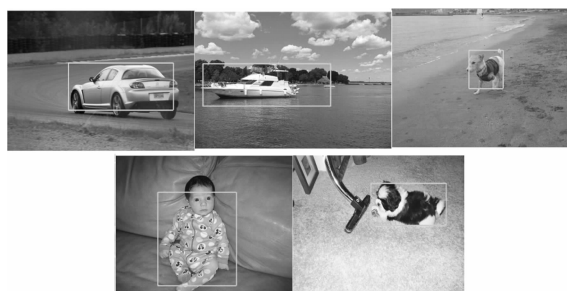


图 7 PASCAL VOC 2007 图像库单目标图像检测结果  
Fig. 7 Detection results of single object in PASCAL VOC 2007 dataset



图 8 PASCAL VOC 2007 图像库多目标图像检测结果  
Fig. 8 Detection results of several objects in PASCAL VOC 2007 dataset

## 5.2 MSRC 图像库检测结果

虽然本文实验参数是在 PASCAL VOC 2007 训练库进行训练的,但是本文改进后模型是具有通用性的,故又选择了 MSRC 图像库中的 13 类目标进行检测,测试结果如表 4 和图 10 所示。

表 4 MSRC 图像库目标检测率表

Tab. 4 Detection rates of MSRC image objects

class	飞机	牛	人脸	汽车	自行车	羊	花	飞禽	椅子	狗	船	人	猫
RATE	0.87	0.87	0.80	0.67	0.43	0.97	0.90	0.93	0.83	0.80	0.60	0.53	0.90

结果表明检测率达到 80.64%，而击中率高达 99.10%。由于 MSRC 库的图像背景比较单一，场景较简单，干扰因素较少，目标较明显，所以

检测成功率远远高于 VOC2007 的检测成功率。

图 9 和图 10 分别为本文方法在 MSRC 图像库单目标检测和多目标检测的结果。

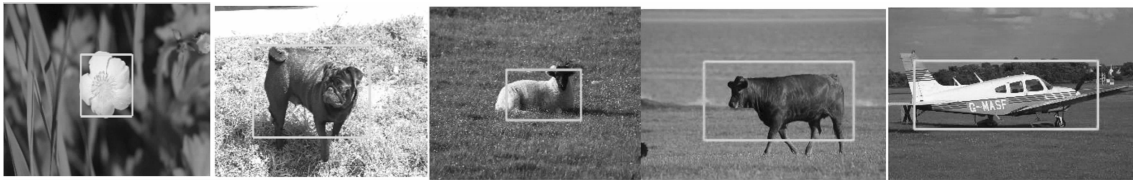


图 9 在 MSRC 图像库和单目标图像检测结果

Fig. 9 Detection results of single object in MSRC dataset

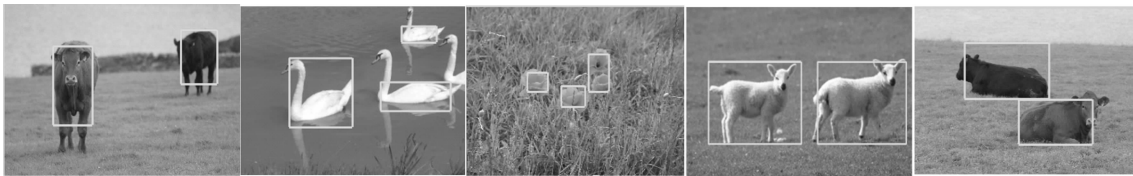


图 10 在 MSRC 图像库的多目标图像检测结果

Fig. 10 Detection results of several objects in MSRC dataset

## 6 结 论

本文提出的改进的无监督通用目标检测模型，基于显著性检测、颜色对比、超像素跨越 3 个图像线索，定位准确，在 VOC 2007 图像库和 MSRC 图像库的检测率都较目前其他方法高，图像目标检测的处理速度也较快。该模型作用于 VOC 2007 库的检测率为 28.94%，在一台主频 2.8 GHz，内存 4 G 的 PC 机上处理单幅图像的处理时间为 5 s 左右。与目前其他无监督目标检测方法作用于 VOC 2007 库相比，文献[13]基于模型的有区分训练部分的目标检测的平均检测率为

26.80%，本文模型比其高出 2 个百分点；文献[8]的检测率为 25.91%，单幅图像的处理时间为 8 s 左右，而本文模型检测率比其高出 3 个百分点，处理时间提高了 40%。在 MSRC 库上进行测试，文献[14]的检测率 58%，文献[15]的检测率为 64%，本文方法检测率均高于两者。此检测模型对于单目标的检测率较高，但是对于多目标的检测还存在不足。某些背景比较复杂的图像，可能会将小目标丢失；多个目标重叠时，可能只能将目标并集块框出，而不能将目标分别检测出来。本文进一步的工作将致力于对图像中的多个目标进行精确的检测定位。

## 参考文献:

[1] LEIBE B, LEONARDIS A, SCHIELE B. Combined object categorization and segmentation with an implicit

shape model [C]. *European Conference on Computer Vision, ECCV*, 2004: 17-32.

[2] FELZENSZWALB P F, GIRSHICK R B, MCALL-ESTER R B D, *et al.*. Object detection with dis-



- criminally trained part based models [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, TPAMI, 2010; 1627-1645.
- [3] LI L J, WANG G, LI F F. OPTIMOL: automatic online picture collection via incremental model learning [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, CVPR, 2007; 1-8.
- [4] GU C, LIM J J, ARBELAEZ P, *et al.*. Recognition using regions [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, CVPR, 2009; 1030-1037.
- [5] KUMAR M P, TORR P H S, ZISSERMAN A. OBJCUT: Efficient segmentation using top-down and bottom-up cues[C]. *IEEE Conference on Computer Vision and Pattern Recognition*, CVPR, 2005; 18-25.
- [6] LEIBE B, SCHIELE B. Interleaved object categorization and segmentation[C]. *British Machine Vision Conference*, BMVC, 2003; 264-271.
- [7] FULKERSON B, VEDALDI A, SOATTO S. Class segmentation and object localization with superpixel neighborhoods [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, CVPR, 2009; 670-677.
- [8] ALEXE B, DESELARES T, FERRARI V. What is an object? [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, CVPR, 2010; 73-80.
- [9] ACHANTA R, HEMAMI S, ESTRADA F, *et al.*. Frequency-tuned salient region detection[C]. *IEEE Conference on Computer Vision and Pattern Recognition*, CVPR, 2009; 1597-1604.
- [10] HOU X, ZHANG L. Saliency detection: a spectral residual approach[C]. *IEEE Conference on Computer Vision and Pattern Recognition*, CVPR, 2007; 1-8.
- [11] EVERINGHAM M, Van Gool L, WILLIAMS C K I, *et al.*. The PASCAL Visual Object Classes Challenge 2006[R]. 2006.
- [12] SHOTTON J, JOHNSON M, CIPOLLA R. Semantic texton forests for image categorization and segmentation [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, CVPR, 2009; 1-8.
- [13] 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*, TPAMI, 2010; 1627-1645.
- [14] SHOTTON J, WINN J, ROTHER C, CRIMINISI A. Textonboost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture[J]. *International Journal of Computer Vision*, IJCV, 2007; 2-23.
- [15] VERBEEK J, TRIGGS B. Region classification with markov field aspect models [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, CVPR, 2007.
- [16] EVERINGHAM M, VAN GOOL L, WILLIAMS C, *et al.*. Zisserman. The PASCAL Visual Object Classes Challenge 2007[R]. 2007.

#### 作者简介:



宋修锐(1983-),男,山东烟台人,博士研究生,2006年于吉林大学获得学士学位,主要从事图像目标检测识别技术方面的研究。E-mail: temp2013@163.com

#### 导师简介:



吴志勇(1965-),男,内蒙通辽人,研究员,博士生导师,主要从事光电测控设备总体技术和光纤通讯技术方面的研究。E-mail: wuzy@ciomp.ac.cn