

一种基于颜色属性的通用目标检测方法

赵起超, 任明武

(南京理工大学 计算机科学与工程学院, 江苏 南京 210094)

摘 要: 通用目标检测是视频分析技术中目标定位、检测和跟踪的关键技术. 针对数字图像分析中的通用目标检测问题, 提出了一种基于颜色属性的通用目标检测算法, 用以提取图像中有可能包含目标的子区域并以外接矩形的形式输出. 该算法首先利用输入图像的颜色属性对其进行区域分割并计算划分后的区域相似度, 接着根据相邻区域间的相似度进行区域合并, 最后结合选择性搜索策略, 实现了通用目标的检测. 实验表明此算法检测出的候选框的精度可以满足通用目标检测的要求.

关键词: 颜色属性; 区域分割; 选择性搜索; 通用目标检测

中图分类号: TP391.41

文献标识码: A

文章编号: 1000-7180(2017)02-0128-04

A Method of Generic Object Detection Based on Image Color Names

ZHAO Qi-chao, REN Ming-wu

(School of Computer Science and Engineering, Nanjing University of
Science and Technology, Nanjing 210094, China)

Abstract: Generic object detection is the key to object localization, detection and tracking in video analyzing. In this paper, we propose a algorithm of generic object detection based on image color names, by which we can obtain sub-regions containing objects in the form of rectangle. In the beginning, an input image is segmented by using color names and the similarity of different regions is computed. Next, we merge the regions based on the similarity between the adjacent ones. Last but not least, selective search strategy is adopted to realize generic object detection in the images. The Experiments show that the candidate boxes detected by this method can meet the accuracy requirements of generic object detection.

Key words: color names; region segmentation; selective search; generic object detection

1 引言

目标检测、识别及跟踪是视频分析中的关键环节,而目标检测又是其他两个方面的基础. 在传统目标检测领域,如何快速有效地选取出目标的候选框尤为重要. 传统的目标检测中,采用的依然是穷举的搜索策略,即通过各种尺度下的滑动窗口来产生目标的候选框位置,但是该策略下需要搜索的窗口数目是巨大的,而且会产生很多不必要的重复候选窗口,为之后的目标识别、分类带来很多不必要的时间开销^[1].

为此,近年来提出了一种新的名为通用对象候

选区域检测算法^[2],这类算法的目的是提取出图像中有可能包含目标的区域并以外接矩形的形式予以输出. 通用对象检测算法应该满足在输出尽可能少的候选框的同时,还必须保证候选窗口中具有较高的检测目标的准确率. 但是由于不同图像中包含对象种类的多样性,以及图像来源的丰富性,使得设计一个高质量的通用对象检测算法异常困难.

文献[3]提出了一种基于图分割算法的选择性搜索策略,它首先采用文献[4]提出的基于图分割的算法对输入图像进行处理,进而可以得到分割后的子区域;其次分别将子区域作为起点,计算起点与其相邻子区域间的相似度,然后挑选与起点相似度最

大的子区域进行合并,把每次合并结果作为一次输出;之后把合并后的结果作为新的起点,重复上述过程迭代输出,直至输出结果为整幅图像.该算法的最大优点是可以检测出图像中不同尺度的对象,但是缺点是相邻子区域间相似度的计算在选择性搜索中反复迭代,造成时间开销较大^[5].

针对这一问题,本文提出了一种基于颜色属性的通用对象检测算法.实验结果表明本算法在保证质量的情况下产生的候选框的数量更少,可以满足我们的实际需求.

2 算法流程

通用目标检测算法一般可以分为分裂合并算法和窗口评分算法两大类.文献[6]对二者常见的几种算法作了比较,发现分裂合并算法在通用目标检测上具有召回率高、可重复性,以及检测速度快的优势.

本文结合分裂合并算法的优点,基于输入图像的颜色属性对其做了相关改进,其主要流程如图 1 所示.



图 1 改进目标检测算法流程

上述过程的详细流程如下:

- (1)利用颜色分割算法将输入的 RGB 三色图转化为 11 色图;
- (2)将 11 色图中每个像素点作为图的顶点,利用文献[4]中的算法,得到划分后的初始子区域.
- (3)以得到的初始子区域作为起点,计算与起点相邻的子区域的颜色和纹理直方图,利用颜色纹理间的汉明距离作为评价相邻区域间相似度,选择与

- 起点相似度最高的子区域与之合并;
- (4)每合并一次,将合并结果作为一次目标检测候选区域的一次输出,同时把合并后的区域作为新的起点重复上一步骤;
- (5)当合并后的结果为整幅图像时,算法结束;
- (6)将得到的候选区域以外接矩形形式予以输出.

3 基于颜色属性的区域划分算法

3.1 图像的颜色属性

颜色属性(Color Names)是人类为自然界中存在的颜色赋予的语言标签^[7].在语言学领域中,文献[8]提出了常见的 11 种语言学的基本颜色:红色、绿色、蓝色、白色、灰色、黑色、紫色、粉色、黄色、棕色以及橙色. Color Names 的目的是给出图像中每个像素对应的语言学颜色.文献[9]通过使用机器学习算法中概率潜在语义分析模型(PLSA)来学习训练真实世界图像中的颜色,给出了图像中 RGB 值到语言学常见的 11 种颜色的映射关系.

本文将 RGB 三色图即常见的 2^{24} 种颜色简化为常见的上述 11 种语言学颜色的组合,图像中每一个像素用 11 色中的一种表示,进而可将图像划分为不同的区域.基于文献[9]机器学习后的结论,本文将 RGB 彩色图转化为 11 色图的基本原理如下:

$$\text{Color} = \text{LUT}[F(R,G,B)] \tag{1}$$

式中,数组 LUT 为文献中通过机器学习得到的颜色查找表, $F(R,G,B)$ 为 RGB 三色图到 11 色图的映射关系.

3.2 基于 11 色图的区域划分

图像分割就是将图像分为若干个特定的、具有特殊性质的区域并从中提取感兴趣目标的技术和过程.它是由图像处理到图像分析的关键步骤.

基于分裂合并的通用目标检测是在图像分割的基础上进行的,文献[4]中所提出的基于图的图像分割算法是目前在图像分割领域效率最高的算法之一.该算法将图像用加权图抽象化表示,初始时将每个像素点都视为图中一个区域,把两个顶点间的欧式距离作为两个区域边的相似度,当该相似度大于给定的阈值时,将二者合二为一.迭代合并,最终形成一个个区域.但是该算法在计算相似度时,采用欧式距离往往容易受到噪声的影响,而且对于阈值较为敏感;为此本文采用 11 色图代替 RGB 三色图来消除噪声的影响,而且在计算相似度时采用整形移位的汉明距离来代替简单的欧式距离,来消除浮

点型运算的时间开销,进而简化计算.

利用上述算法,处理后的结果如图 2、图 3 所示,可见处理后的 11 色图与人类的主观观测基本保持一致,但将颜色由 2^{24} 种减少为 11 种常见基本色,为之后的分割减少了大量数据量的计算.

此外,本文还将改进后的区域划分算法和文献[4]做了比较,对得到的区域用不同的颜色表示,通过统计输出的初始子区域数目,可以看出改进算法具有较好的变化.

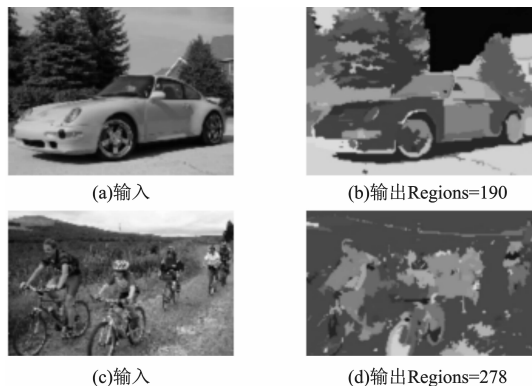


图 2 文献[4]分割算法的输入与输出

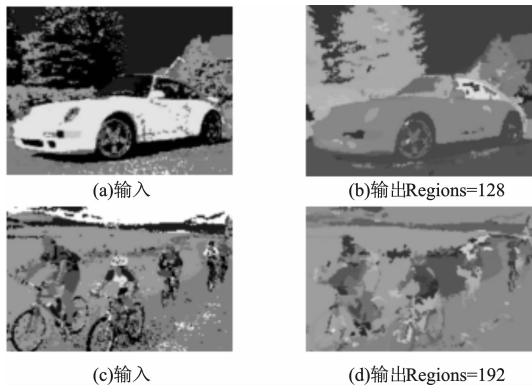


图 3 改进分割算法的输入与输出

4 区域间相似度计算

为了评价不同区域的相似度,本文根据不同区域间颜色和纹理特征的相异性,综合颜色、纹理相似度,提出了可以快速计算的综合相似度.

4.1 颜色相似度

本文首先计算每个区域每个颜色通道的直方图,为了简化计算,本文将每个通道的颜色直方图归一化为 25 段,这样每个通道的颜色直方图可以用一个 25 维的向量来表示,进而一个区域可以用一个 75 维的向量来表示. 我们不妨计区域 r_i 的颜色直方图为 $C_i = \{c_i^1, \dots, c_i^n\}$,区域间颜色相似度可以通过下面的公式计算:

$$S_{\text{color}}(r_i, r_j) = \sum_{k=1}^n H(c_i^k, c_j^k) \quad (2)$$

式中, $H(c_i^k, c_j^k)$ 表示 c_i^k 和 c_j^k 的汉明距离.

当区域 r_i 的颜色分布直方图分布为 $F_{r_i}(g)$,则区域 r_i 和区域 r_j 合并成区域 r_t 的计算公式为

$$F_{r_t}(g) = \frac{F_{r_i}(g) \times \text{size}(r_i) + F_{r_j}(g) \times \text{size}(r_j)}{\text{size}(r_i) + \text{size}(r_j)} \quad (3)$$

4.2 纹理相似度

本文定义纹理相似度用 SIFT 特征,具体做法是对每个通道分别用八个方向,方差为 1 的高斯滤波器进行处理,之后对每个方向像素值的分布归一化为 10 段,进而统计每段的频率作为其纹理直方图. 这样一个区域的纹理直方图 T_i 可以用一个 240 维的向量来表示,即 $T_i = \{t_i^1, \dots, t_i^n\}$. 区域之间的纹理相似度的计算公式和颜色相似度的计算公式类似,如下所示.

$$S_{\text{texture}}(r_i, r_j) = \sum_{k=1}^n H(t_i^k, t_j^k) \quad (4)$$

合并后的纹理分布直方图,与上述颜色分布直方图计算方式相同,不再累述.

4.3 综合相似度

本文最后将上述两个相似度的计算方式结合起来,作为一个区域的综合相似度,其计算公式如下:

$$S(r_i, r_j) = a_1 S_{\text{color}}(r_i, r_j) + a_2 S_{\text{texture}}(r_i, r_j) \quad (5)$$

5 区域合并算法

通过上述过程可以得到相对较小的初始区域,接下来是将子区域逐一合并,我们采用贪心策略来迭代产生合并区域,并把合并后的区域作为目标的候选区域,以外接矩形的形式予以输出. 针对区域合并问题,本文该算法的详细流程如下所示.

Step1: 输入为上一步骤得到初始化区域集合 $R = \{r_1, r_2 \dots r_n\}$, 计算两两区域间的相似性 $S(r_i, r_j)$, 放到集合 S 里面存储.

Step2: 对于集合 R 中的区域 r_i , 根据计算后的相似性集合 S , 选择集合 R 中与 r_i 相似性最大的区域 r_j , 满足 $S(r_i, r_j) = \max(S(r_i, r^*))$.

Step3: 将区域 r_i 与 r_j 进行合并得到 r_t , 即 $r_t = r_i \cup r_j$.

Step4: 从集合 S 中, 移走所有与 r_i 和 r_j 相关的数据.

Step5: 计算新区域 r_t 与他的相邻的区域的相似

性 $S(r_i, r^*)$).

Step6: 将 r_i 加入区域集合 R 中 即 $R = R \cup r_i$.
重复 Step2 ~ 6, 直到集合 S 为空.

Step7: 对于最终的区域集合 R 中的每个区域, 给出其外接矩形表示.

6 实验结果与分析

为了验证算法的有效性和鲁棒性, 本文采用图像识别与物体分类挑战赛中的 VOC2007 数据集进行试验, 该测试数据集中给出了 20 个物体类别共计 4 952 张不同场景下获取的图像, 众多的目标数量、丰富的目标类别、不同的视角、尺度、位置、场景, 使得这个数据集适合去寻找图像中的目标.

本文采用常见的 VOC2007 测试集的 4 952 张图片进行实验, 程序在内存为 4 GB 的 PC 上在 VS2012 环境下运行. 为了更加直观地观测检测出的效果, 本文针对测试集中的每帧图像, 将与图像中目标重叠率最高的候选框以红色外接矩形输出, 并将人工标注的真实位置予以绿色外接矩形输出. 由于本文是基于文献[2]提出的框架, 为此本文针对其结果与文献[2]进行了性能比较, 结果如表 1 所示.

表 1 两种算法的结果对比表

算法	图片数量	平均窗口 数量/帧	M-ABO/ 帧	Pref X1000/帧
文献[2]算法	1 000	302	0.680 1	2.252
	2 000	303	0.685 7	2.262
	3 000	304	0.680 1	2.237
	4 952	303	0.672 6	2.225
本文算法	1 000	151	0.678 3	4.492
	2 000	153	0.680 1	4.433
	3 000	153	0.676 7	4.423
	4 952	154	0.674 3	4.379

从表 1 中的数据可以看出, 改进后的算法和文献[2]相比, 在测试集每帧平均最大目标重叠率基本不变的情况下, 候选框的数量减少了一半. 究其原因, 是因为基于 11 色图像的图分割算法确定的初始子区域数量较原算法数量上有了较大的减少. 此外, 本文算法在保证通用目标检测准确率的基础上, 由于初始子区域数量的较少, 极大地提高了运算效率. 因此, 本文算法非常适合特定图像目标检测的实际应用.

7 结束语

本文提出一种基于颜色属性的通用目标检测算法, 不用于绝大多数通用检测算法, 首先基于颜色属性对输入图像进行了 RGB 三色图到 11 色图的转化, 得到了较少的初始子区域, 之后在计算区域间相

似度时采用整形的位运算计算汉明距离来代替浮点型的欧式距离, 极大地降低了算法的执行效率. 此外, 数据量的减少使得本算法输出的候选窗口数量基本维持在一个较小的数量级, 但是通用目标检测率并没有因此而降低, 而是稳定在一个较高的水平, 这极大地减少了后续的目标识别的搜索空间, 提高了算法的应用性. 实验结果表明, 本文算法可以在保持较高的目标检测率的情况下用较少的候选框来描述目标的候选区域, 具有较好的通用性.

参考文献:

[1] Hosang J , Benenson J , Dollar R. What makes for effective detection proposals [J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2015, 99:1.

[2] Uijlings J R R , K E A. van de Sande , Gevers T, et al. Selective Search for object Recognition[J]. International Journal of Computer Vision, 2013, 104: 154-171.

[3] Koen E A van de Sande, Jsaper R R Uijlings, Theo Gevers, et al. Segmentation as selective search for object recognition[R]. ICCV, 2011:1879-1886.

[4] Felzenszwalb P F, Huttenlocher D P. Efficient graph-based image degmentation[J]. International Journal of Computer Vision, 2004, 59:167-181.

[5] 闾跃龙, 张石清, 郭文平, 等. 基于感知哈希和自适应搜索的通用对象检测[J]. 光电子·激光, 2016, 27 (2):231-238 .

[6] Hosang J, Benenson R, Schiele B. How good arre de-tection proposals, really? [C]// Proceedings British Machine Vision Conference 2014. Nottingham , 2014.

[7] Fahad Shahbaz Khan, Rao Muhammad Anwer, Joost van de Weijeer. Color Attributes for Object Detection [C]// Proceedings CVPR 2012. CVPR, 2012.

[8] Xu Dan, Tang Zhenmin, Xu Wei. Combining color names with spatial information for salient object detection[J]. Journal of Image and Graphics, 2014, 19(4):541-548.

[9] Joost van de Weijer, Cordelia Schmid, Jakob Verbeek. Learning Color Names from Real-World Images[C]// IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis:IEEE, 2007:1-8.

作者简介:

赵起超 男, (1991-), 硕士研究生. 研究方向为图像处理、模式识别与智能系统.

任明武(通讯作者) 男, (1969-), 博士生导师. 研究方向为图像处理、模式识别与智能系统、机器视觉、智能公共安全.

E-mail: renmingwu@njust. edu. cn.