

基于视觉显著性的人脸文字区域提取

何贵阳*

(重庆城市管理职业学院文化产业管理学院,重庆 401331)

摘要:基于 Matlab 图像文字区域检测新方法对图文进行分离和显著性操作,并采用显著性度量对传统人脸文字区域检测的结果做进一步的显著性融合,实现了基于视觉显著性的人脸文字区域提取。实验结果表明,这种基于视觉显著性的人脸文字区域提取解决方案具有比较高的查全率和精确度。

关键词:OpenCV;显著性提取;图文分离;C++

中图分类号:TP301.6 **文献标识码:**A **doi:**10.3969/j.issn.1006-6055.2014.03.011

Regional Extraction of Human Faces and Text Based on Visual Significant

HE Guiyang*

(Institute of Culture Industry Management, Chongqing City Management College, Chongqing 401331)

Abstract:Based on the Matlab image text region detection method, a new method combining graphic separation and significant operation was proposed. Further significant fusion about traditional face text area detection results was realized based on visual significant face text region extraction demonstration system. The experimental results show that the proposed solution has higher recall ratio and precision.

Key words:OpenCV; significant extraction; graphic separation; C++

1 引言

近年来,对于彩色图像文字区域提取的研究已经取得了大量研究成果,薛卫^[1]提出了基于离散傅里叶变换特征、多分辨率处理及支持向量机分类技术的视频文字大小自适应提取算法;Agnihotri 通过边缘滤波与形态学运算结合处理图像,利用启发性规则确定文字区域;程豪^[2]等提出了综合利用视频文字时空、笔画、颜色和几何特征的文字分割算法,快速准确地从复杂背景的视频图像中分割出文字;谭伟等^[3]提出了基于形态学和窗口特征的快速文本提取算法;杨俊^[4]利用文字条角点特征提取出文字块等。由于图像存在背景纹理复杂,文字尺寸不一致,字体多变,颜色多样,排列方式多变等的复杂性,这些算法仍存在文字区域提取的准确率不高,精确度不高,效率不高等问题^[5-7]。目前对人脸、文字等图片重要目标的提取的相关研究也取得了不错的成果。但是将视觉显著性和传统的目标检测方法相结合来增强检测效果的研究还比较少,这也是本文要解决的主要问题之一。

本文主要借鉴 Ming-Ming Cheng^[8]自顶向下的视觉注意力模型,采用 Adaboost 实时人脸检测算

法,深入研究任务驱动的图像显著性算法及低层特征提取方法,实现基于视觉显著性和传统目标检测方法相结合的人脸、文字等兴趣区域提取。基于 Matlab 图像的文字区域检测新方法,结合图文分离和显著性操作,并采用 C++ 加以实现,对传统人脸文字区域检测的结果采用显著性度量进行了进一步的显著性融合。因此最终的检测结果融合了视觉显著性。本文利用 RC 显著性图和二值化的 RC 显著性图,对已提检测出的人脸区域进行显著性属性方面的考察,来最终确定检测结果。

2 区域提取系统设计

基于视觉显著性的人脸和文字区域提取系统主要分为人脸区域提取、文字区域提取、显著性检测、显著性融合、结果演示等五个模块,其系统流程如图 1。该视觉显著性提取系统主要由人脸区域提取、文字区域提取、显著性检测、显著性融合、结果演示等五个模块组成。其各个模块的功能说明详见表 1。

表 1 视觉显著性提取系统模块化说明

Table 1 Modular specification of visual significance extraction system

模块序号	模块名称	模块功能	备注
1	人脸区域提取	人脸检测,输出人脸矩形区域	Adaboost 算法
2	文字区域提取	文字区域提取,输出文字矩形区域	自定义
3	显著性检测	RC 显著图生成	RC 算法
4	显著性融合	利用 RC 显著图对 1、2 步骤中的人脸、文字区域进行显著性融合	自定义
5	结果演示	界面设计,效果展示	无

2014-03-06 收稿,2014-04-14 接受

* 通讯作者, E-mail: oldsomething@163.com; Tel: 18623506275

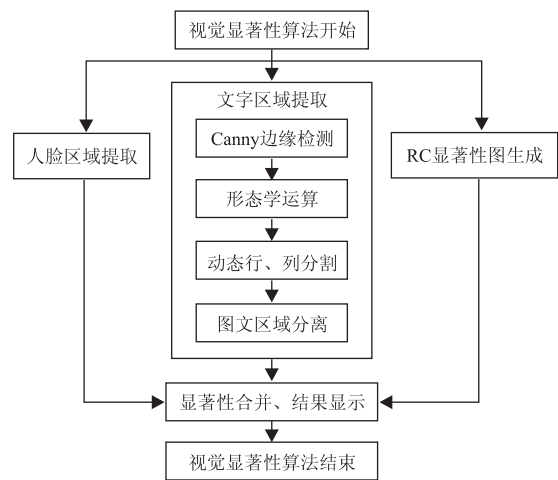


图 1 基于视觉显著性的人脸和文字区域提取系统流程
Figure 1 Extraction processes of face and text area based on the visual distinctiveness

3 基于视觉显著性区域提取实现

3.1 基于视觉显著性的人脸和文字区域提取算法

基于视觉显著性的人脸和文字区域提取在此主要是要将两种并行的视觉特性有机结合起来。将底层的数据驱动自下而上分析产生的数字图像显著性区域记为 X。由上层任务驱动自上而下分析产生的数字图像人脸或者文字检测区域记为 Y。然后采用权值法、集合分析法等办法来将 X 和 Y 进行有机结合,形成最后的检测区域,从而实现基于视觉显著性的人脸和文字区域提取。具体显著性融合过程中,其具体算法如下:

1)RC 显著性图中已检测人脸、文字区域矩形集合元素的平均像素,计算后和指定阈值比较

RC 算法首先通过一个图像分割算法把输入图像进行分割,然后建立关于每个区域颜色直方图。我们采用每个区域与其他区域的颜色对比来定义其区域显著性:

$$S(rk) = \sum_{rk \neq ri} w(ri) Dr(rk,ri) \tag{1}$$

其中, $w(ri)$ 是区域 r_i 的权重; $Dr(\cdot, \cdot)$ 是两区域之间的颜色距离度量。在此我们采用区域 r_i 的像素数作为权值 $w(r_i)$ 来突出大区域的颜色对比。其中区域 r_1 和 r_2 之间的颜色距离定义如下:

$$Dr(r_1,r_2) = \sum_{i=1}^{n1} \sum_{j=1}^{n2} f(c_1,i)f(c_2,j) D(c_1,i,c_2,j) \tag{2}$$

其中, $f(c_k,i)$ 是第 k 个区域 r_k 中 $n_k(k=1,2)$ 个颜色中第 i 个颜色 $c_{k,i}$ 的概率。请注意我们使用区域概率密度函数(即归一化颜色直方图)来定义颜色的

概率以强调主色之间的颜色差异。存储和计算每个区域的矩形式的直方图是低效的,因为每个区域仅仅包含整个图像颜色的少数部分。相反,我们使用稀疏直方图来实现高效的存储和计算表示。

2)二值化的 RC 显著性图中人脸、文字区域矩形集合元素的有效像素和计算后,指定阈值比较

其中,在第一步对 RC 显著性图 I 进行处理时,图像 I 中第 r 个矩形的平均像素和 $pixel_ratio_{I,r}$ 定义如下:

$$pixel_ratio_{I,r} = \frac{sum(R_r)}{(R_r.width * R_r.height)} \tag{3}$$

式中, R_r 表示图像 I 中当前考察的矩形区域; $R_r.width$ 表示矩形的宽度; $R_r.height$ 表示矩形的高度, $Sum(R_r)$ 表示 I 图像中矩形 R_r 内的像素和。对于人脸和文字区域,要分别考察。当 $pixel_ratio_{I,r}$ 大于固定阈值 T_0 时为人脸区域且为非文字区域。反之则为文字区域且为非人脸区域。当采用自适应阈值 T_1 时, T_1 取各人脸或者文字区域矩形集合的平均像素和的平均值。即自适应阈值 T_1 定义如下:

$$T_1 = \sum_{i=0}^N pixel_ratio_{I,i} / N \tag{4}$$

一个区域与其周边区域的对比度总是高于与离它远的区域的对比度。由于计算像素级的对比度时直接引进的空间关系的计算成本太高,引入了对比分析的方法,即区域对比(RC),融入空间关系到区域计算中。在 RC 中,首先分割输入图像,然后计算在区域一级的颜色对比度,并用每个区域与其他其他区域的颜色对比度的加权求和来定义该区域的显著性。权重设置时较远空间距离的区域赋予较小的权重。

3.2 区域空间加权

进一步通过在式(1)加入空间权值来引入空间信息,可以增加较近区域的影响并减少更远区域的影响。具体来说,对于区域 r_k ,基于空间加权的区域对比显著性被定义为:

$$S(rk) = \sum_{rk \neq ri} \exp(-DS(rk,ri)/\delta_s^2) w(ri) Dr(rk,ri) \tag{5}$$

其中, $DS(rk,ri)$ 是区域 rk 和 ri 的空间距离, δ_s 控制空间加权的强度,较大的 δ_s 能够减少空间加权的影响从而使得较远的区域能够对当前区域贡献较大的对比度。两个区域之间的空间距离定义为区域中心点之间的几何距离。在我们的实现当中,我们采用 $\delta_s^2 = 0.4$ 并归一化像素坐标到 $[0,1]$ 。

4 实例验证

针对视觉显著性理论、自底向上和自顶向下的视觉注意力模型,视觉显著性主要从生理学,心理学,以及图像处理等角度出发,理解视觉显著性原理,建立模型来模拟,并采用基本的图像处理算法,在计算机视觉方面来加以实现。自底向上的数据驱动视觉特性主要来自于图像本身颜色、强度、方向、纹理等特征的对比。所以此类算法核心就是对比。对边的强烈程度就反映了显著性区域的显著性强弱。而自上而下的任务驱动视觉特性的基本思想是:先根据期望目标的先验信息选择某种相对独特的图像特征,将具有该特征的图像区域检测出来,再对这些区域是否包含期望目标进行证实。

针对单张图像处理,首先对源图像进行人脸检测处理,其检测效果如图 2 所示。然后对源图像进行 RC 显著性检测并生成如图 3a 所示的 RC 显著性图,并对 RC 显著性图进行自适应二值化得到如图 3b 所示的 RC 显著性图的二值化图。接下来进行文字检测,其中中间步骤比较多。最后对检测出人脸区域、文字区域集合进行显著性融合,就得到了最终的检测结果。

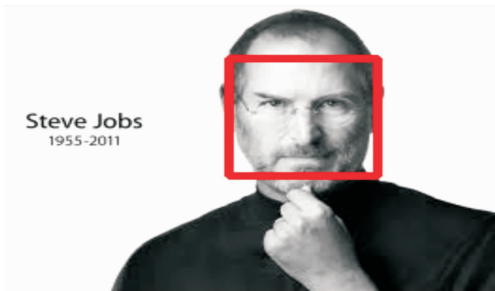


图 2 人脸检测效果图
Figure 2 Face detection effect

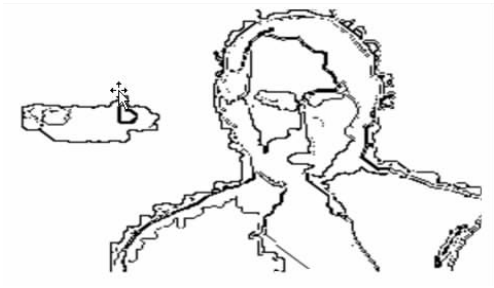
对自适应阈值二值化后进行 Canny 边缘检测和形态闭运算闭运算,迭代次数 89 后得效果图图 4。

对形态闭运算闭运算后的效果图的结果进行逐行连通域考察,连通性考察时,固定最小行高取 $Hang = 6$ 像素。若在每一行中,当前列的有效像素(二值化图中为 1 的像素)的个数大于行高的 80% (此处 $int(Hang * 80\%) = int(6 * 80\%) = 4$) 则标记为矩形区域,得到最原始目标区域矩形集合效果图 5 其中 $int(x)$ 表示对 x 向下取整),采用红色矩形标记出文字区域,对连通性考察结果进行动态列分割(相同行列合并),其效果图见图 5 所示。

采用红色矩形标记出文字区域,对与动态列分

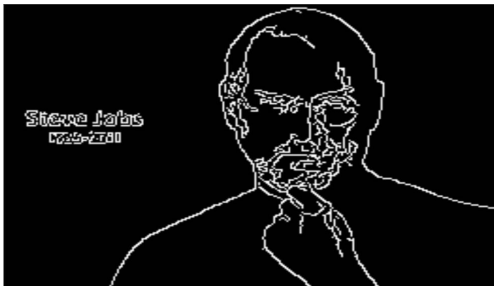


(a) RC 显著性图



(b) 自适应阈值二值化后效果图

图 3 RC 显著性与自适应阈值二值化效果图
Figure 3 Binarization rendering of RC significance and adaptive threshold



(a) Canny 边缘检测效果图



(b) 形态闭运算闭运算后的效果图

图 4 Canny 边缘检测与形态闭运算闭运算后的效果图
Figure 4 Effects after Canny edge detection and morphological closed operations

割后的结果进行文字区域非文字区域分离后得到的效果图,见图 6a:采用绿色矩形框标记出人脸区域,采用红色矩形标记出文字区域。显著性融合后的人脸文字综合检测效果图如图 6b 所示。

然后经过批处理实验,发现本文提出的方法有比较好的精确度和查全率。但是也存在一些误检



图 5 连通性考察与动态列分割后的效果图
Figure 5 Effects after the connectivity investigation and dynamic column split

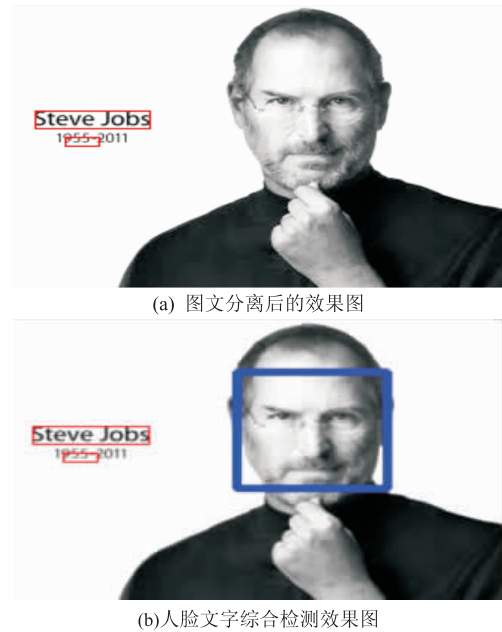


图 6 显著性融合后的人脸文字综合检测效果图
Figure 6 Face and text comprehensive detection effect after significant fusion

率,特别文字区域提取,其将非文字区域检测为文字区域的误检率要高于人脸区域提取的误检。本课题的主要任务就是学习视觉显著性的相关理论知识,理解自底向上和自顶向下的视觉注意力模型,深入研究任务驱动的图像显著性算法及底层特征提取方法,实现基于视觉显著性和传统目标检测方法相结合的人脸、文字等兴趣区域提取。

针对人脸和文字特征,必须提取其基本特性,基于这些特性能够良好的对图像中的人脸或者文字进行数学上的描述和表达。这些常见特性描述方式包括显著点、边缘、轮廓、区域、小波等数学工具。抽象出来的已有特征,就可以有意识的在图像中搜索这些相似区域,这就实现了自上而下的基于任务驱动的人脸识别和文字区域提取。由于底层的数据驱动自下而上分析可以产生数字图像的显著性区域 X,而由上层任务驱动的自上而下的分析又可以产生数字图像的人脸或者文字检测区域 Y,然后采用权值法、集合分析法等办法来对显著性区域 X 或者人脸或者文字检测区域 Y 进行有机结合,便形成了最后的检测区域,从而可以实现基于视觉显著性的人脸和文字区域提取。

5 结论

本文提出的解决方案基于自上而下的模块化思想,对人脸区域提取、文字区域提取、显著性检测等方面分而治之。采用显著性融合来实现分析结果的融合,最终得到一个完整的基于视觉显著性的人脸文字区域提取的解决方案。本解决方案简单、有效、可行,能够对具有复杂背景的图像进行人脸、文字等感兴趣区域提取。

本解决方案虽然基本解决了基于视觉显著性的人脸文字区域提取问题,但也还有许多地方可以进一步改进,包括固定阈值的确定问题。本解决方案里面用到了许多固定阈值,如 Canny 边缘检测处理中的迭代次数,形态学闭运算的迭代次数、行分割的最小行高;图文分离处理中的平均像素和、文字区域长宽比、有效像素比等。这些阈值的选择都会反映在最终的检测效果中。其中有些阈值我们虽然提供了简单的自适应阈值算法,但我们的自适应算法仅仅采用当前图像的待考察区域的相关参数的平均值来作为阈值,这种处理办法还有待进一步改进提高。目前传统的人脸文字区域提取算法都已经比较成熟,将其检测结果和显著性检测时,如果采用“与”思想进行融合,势必对检测结果的查全率影响较大,这源于一个很明显的事实,即存在图像,人脸或者文字并不是里面的显著性区域。反之,如果采用“或”思想进行融合时,势必会对精确度造成一定影响,这会让非人脸、文字的显著性区域凸显出来。在本文中,我们暂时采用的是“与”思想,但在进行显著性

(下转第 276 页)

(4):1-5.

[9]SPECHT D F. A General Regression Neural Network[J]. Transactions on neural networks,1991,2(6):568-576.

[10]史峰,王小川,郁磊,等. MATLAB 神经网络 30 个案例分析[M]. 北京:北京航空航天大学出版社,2010.

[11]关昌余,裴玉龙. 基于 GM-GRNN 国家高速公路网规模预测研究[J]. 公路交通科技,2008,25(4):141-145.

作者简介

刘 勇(1971-),男,博士,副教授,主要研究方向:信息管理与信息系统,物流供应链管理;

王 萱(1989-),女,硕士在读,主要研究方向:信息管理与信息系统,物流供应链管理;

邹慧挺(1982-),男,硕士在读,主要研究方向:工业工程,物流供应链管理。

(上接第 271 页)

融合的时候采用两个层次的显著性考察,在一定程度上提高了查全率。

参考文献

[1]薛卫,都思丹,吴书凯. 基于视频文字大小自适应提取算法[J]. 计算机应用研究,2009,26(3):1 147-1 159.

[2]程豪,黄磊,刘金刚. 基于笔画提取和颜色模型的视频文字分割算法[J]. 计算机工程,2009,35(4):192-195.

[3]谭伟,方超,杜建洪. 基于形态学和窗口特征的快速文本提取[J]. 信息与电子工程,2010,8(3):535-539.

[4]秦宏伟,孙劲光,王强,等. 基于自适应预处理和 PCA 的人脸识别

方法的研究[J]. 计算机工程与应用,2011,47(14):186-188.

[5]杨俊. 图像数据的视觉显著性检测技术及其应用[D]. 长沙:国防科学技术大学信息与通信工程学院,2007.

[6]梁路宏,艾海舟,徐光档,等. 人脸检测研究综述[J]. 计算机学报,2002,25(5):449-458.

[7]李志强. 视觉显著性模型研究及其在影像处理中的应用[D]. 上海:上海交通大学电子信息与电气工程学院,2009.

[8]方昱春,王蕴红,谭铁牛. 融合人脸轮廓和区域信息改进人脸检测[J]. 计算机学报,2004,27(4):482-490.

[9]FELZENSZWALB P, HUTTENLOCHER D. Efficient graph – based image segmentation [J]. International Journal of Computer Vision, 2004,59(2):412- 413.

作者简介

何贵阳(1968-),男,学士,副教授,主要研究方向:摄影摄像技术。