

结合小波插值与 K-means 方法提取 彩色图像文字

张凯歌, 缪毅, 雷建坤, 彭可亮, 蒋慕蓉

(云南大学 信息学院 计算机科学与工程系, 云南 昆明 650091)

摘要: 彩色图像文字的提取和识别是人工智能与模式识别领域中的研究热点。现有方法在处理彩色图像中背景与文字相近或者字体较小的图像文字时会受到复杂背景的影响。为了解决彩色图像中文字与复杂背景颜色相近和模糊不清与排列不规则的文字提取与识别问题, 文中采取人工选取文字区域进行小波插值放大的方法使图像清晰化, 再利用 K-means 方法对文字色彩进行聚类得到单背景彩色文字图像, 然后通过二值化处理和文字分割后再结合 BP 神经网络进行文字识别。实验表明, 该方法不仅能处理大多数彩色图像文字的提取, 而且在识别低对比度彩色图像文字方面比目前常用的方法要好。

关键词: 文字提取与识别; 小波插值; K-means 方法; BP 神经网络

中图分类号: TP31

文献标识码: A

文章编号: 1673-629X(2013)03-0031-03

doi: 10.3969/j.issn.1673-629X.2013.03.008

Extraction of Color Image Texts Combining Wavelet Interpolation and K-means

ZHANG Kai-ge, MIAO Yi, LEI Jian-kun, PENG Ke-liang, JIANG Mu-rong

(Department of Information Science and Engineering, College of Information,

Yunnan University, Kunming 650091, China)

Abstract: The text extraction and recognition of color images is the research focus in the field of artificial intelligence and pattern recognition. In order to solve the problems of text recognition that the characters in color images are too small to see clearly and the color of the characters is similar as the complex background or the characters are arranged irregularly, use the method which combines the wavelet interpolation and K-means cluster to extract the color texts. First, choose the region including the color texts, then employ the wavelet interpolation to enlarge the color image, use K-means method to cluster the color to obtain a single background image, then change this image into the binary image and segment the text. At last, the trained BP neural network is used for the character recognition. Some experiments show that this method can process the most color text extraction easily, and more efficiently for recognizing the text with the lower contrast and complex background.

Key words: text extraction and recognition; wavelet interpolation; K-means method; BP neural network

0 引言

文字识别在多媒体技术、计算机视觉、模式识别、人工智能等领域有着广泛的应用。目前常用的文字识别方法大多基于图像块和纹理的特点, 先进行灰度转换, 再进行二值化, 然后通过神经网络对目标文字进行识别, 这类方法在处理彩色图像中背景与文字相近或者字体较小的图像文字时会受到复杂背景的影响。文

献[1]提供了统计识别法, 文献[2]使用神经网络在 Haar 小波解析特征空间中抽取文字块, 文献[3]中综合分析了空间差异和连通域, 文献[4]通过 Gabor 滤波空间方差分析提取文字区域, 上述算法具有一定的可操作性, 但这些算法对文字的字体风格比较敏感, 存在着定位不准和算法复杂度高等问题, 而且计算耗时。

在总结前人工作^[5~8]的基础上, 文中先对图像进行增强, 再对选取的目标文字进行小波插值放大, 之后采用改进的 K-means 聚类算法对文字颜色进行聚类, 然后对聚类后的文字进行分割, 最后运用 BP 神经网络对分割后的文字进行识别, 获得了目标文字的提取。实验验证, 该方法不仅能处理大多数彩色图像文字的

收稿日期: 2012-06-27; **修回日期:** 2012-09-29

基金项目: 国家自然科学基金资助项目(11161055); 云南省中青年学术和技术带头人项目(2008PY034)

作者简介: 张凯歌(1986-), 男, 硕士研究生, CCF 会员, 研究方向为图像分割、文字识别; 蒋慕蓉, 教授, 研究方向为分布式图像处理。

提取,而且在识别低对比度彩色图像文字方面比目前常用的方法要好。

1 利用小波插值与 K-means 聚类方法提取彩色图像文字

1.1 小波插值对目标图像进行放大处理

小波变换的技术原理是通过平移母小波获得信号的时间信息,通过缩放母小波的宽度获得信号的频率特性。对母小波的缩放和平移操作是为了计算小波系数,而这些系数代表小波与局部信号之间的相互关系。

常用的小波插值放大方法有两种:一种是利用二进制小波变换实现目标图像放大,先将待放大的图像进行二进制小波变换得到四个频率分量的系数,之后将其中的高频分量用插值方法放大到一定倍数,再将低频分量用原始图像替换之后做小波逆变换重构图像;另一种方法是先将待放大图像用普通插值方法放大后,再对放大后的图像进行小波分解,将分解后的高频系数取出,再用原图像替换低频分量后与其高频分量做小波逆变换得到重构的放大图像。上述两种方法都存在一个很大的缺点,就是放大后图像与原图像的亮度不一样,放大后图像的灰度值明显变暗。为此,在文献[9~11]的基础上采用先对原始图像进行二维小波分解,取其低频分量进行内差值放大后代替原图像作为重构的低频分量,再将原图像进行插值放大后取其水平、垂直、对角线上的三个高频分量参与重构,实现图像放大^[12]。

利用小波插值进行图像放大,结果如图 1 所示。



(a) 原图

(b) 原图



(c) 截取并放大文字区域

图 1 小波放大图像

1.2 利用 K-means 方法去除彩色图像背景

K-means 方法是一种应用最广泛的聚类方法,其基本思想是在预先不知道样本数据有多少种类的情况下,以某种度量作为相似性准则,通过不断划分与聚集,将所有数据分配到不同的类别中,使得在这种分类情况下,同一种类中的数据相似度最大,不同种类间的数据差异性最大。

在彩色图像中,文字的色彩基本是相同的,因此,

可以利用 K-means 方法对文字色彩进行聚类处理。

步骤如下:

第一步:将原彩色图像的 RGB 空间转换到 $L * a * b$ 空间,用以测量不同色彩间的欧几里得距离。

第二步:使用如下 K-means 聚类算法对 $L * a * b$ 空间的色彩进行分类,将目标分割成 N 个簇,步骤如下:

(1) 首先取任意 k 个属性值向量为初始簇的中心,设这 k 个值为 $Y_1(l), Y_2(l), \dots, Y_k(l)$, 循环处理次数为 n , 其初始值为 1;

(2) 所有的属性值向量 X 由式(1)进行分类,使向量集 $\{X\}$ 中的向量分别属于簇的中心 $Y_1(n), Y_2(n), \dots, Y_k(n)$, 相应的子集为: $S_1(n), S_2(n), \dots, S_k(n)$ 。

$$d_i = \min \{d_j\} \rightarrow X \in S_i(n), \quad j \in \{1, 2, \dots, K\} \quad (1)$$

(3) 定义 X 与 $Y_j(n)$ 之间的距离 d_j 为:

$$d_j = \|X - Y_j(n)\| \quad (2)$$

(4) 计算任意两个样本间的距离,找到集合 Y_k 中距离最近的两个点形成集合 A_m , 在 Y_k 中找到距离集合 A_m 最近的点,将其加入集合 A_m 并从集合 Y_k 中删除;

(5) 由式(3)计算出各子集 $S_l(n) (l \in \{1, 2, 3, \dots, K\})$ 的新簇的中心 $Y_l(n+1)$

$$Y_l(n+1) = \frac{1}{N_l} \sum_{X \in S_l(n)} X \quad (3)$$

其中: N_l 是集合 $S_l(n)$ 的元素个数。(3) 式中 $Y_l(n+1)$ 是属于 S_l 的 X 的平均值。

(6) 将最终形成的 k 个集合中的样本点分别进行算术平均,从而形成 k 个初始聚类中心,并形成最终聚类。

第三步:用聚类结果对图像像素进行标记,对图像进行色彩分割,得到分割的彩色文字区域,结果如图 2 所示。



图 2 去除彩色背景图

1.3 文字分割

将去除彩色背景图进行二值化,再采用 13 点特征提取方法^[10]进行汉字分割得到单个字符(如图 3 所示)。



图 3 汉字分割

1.4 文字识别

利用神经网络对分割后的文字进行训练,截取数字图像像素值为 0(黑)的最大矩形区域,将此区域的图像经过集合变换,使之变成 16 * 16 的二值图像(如果保存时采用的是 24 位图格式或其他图像,则须使用图像转换命令转换为二值图像或者是灰度图像),然后将二值图像进行反色处理,得到的像素值为 0,1 的图像并作为神经网络的输入向量。所有训练样本图像都经过这样的处理。然后将小波放大后的图像进行上述方法的聚类处理,将得到分割后的文字采用上述神经网络的方法进行文字识别。

2 实验流程与结果分析

实验是在 Matlab 2011 环境下进行的,实验用图是网络上截取的中央电视台《新闻联播》的画面。实验流程如图 4 所示:

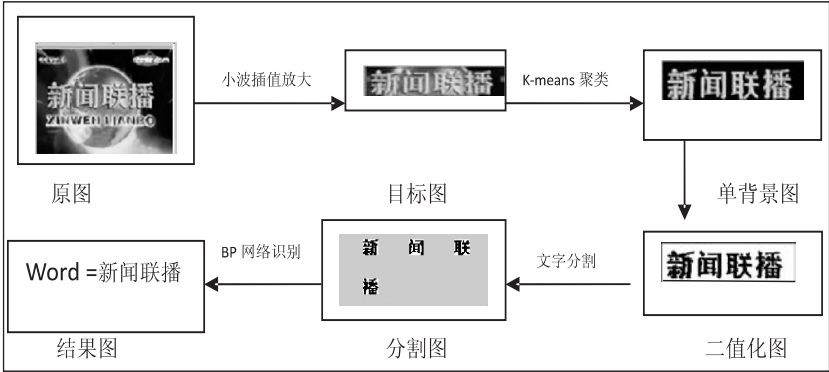


图 4 文字识别流程图

首先对截取的图像区域采用小波插值放大获取较清晰的文字处理区域,然后进行颜色聚类处理,得到单一背景的目标图像,重复聚类 5 次,以避免局部最小值。

然后将得到的单一背景目标文字图像进行灰度化、二值化等处理得到最终的图像,然后进行文字的分割

表 1 实验结果

	原图	文中方法获取目标文字
图 1		
图 2		
图 3		
图 4		
图 5		
图 6		

割,再采用 BP 神经网络进行识别。

为验证方法的有效性,还对其它多种类型的图片进行了实验,部分结果如表 1 所示。

由实验结果可以看出,在彩色图像汉字识别的过程中,最关键的是对二值化前图像的处理,通常有几类情况:第一类是普通的单背景彩色图像,直接可以聚类提取;第二类是与背景色彩近似如表 1 中图(2),但色差比较明显,可以直接进行文字提取;第三类是部分文字近似度太高无法聚类识别如表 1 中图(1);第四类是多种背景色但对图像中文字颜色影响较小如表 1 中图(5)和(6);第五类对彩色图像只有外轮廓的彩色文字,可以采用先提取图像文字轮廓,再进行图像文字的内部颜色填充,进而得到文字区域如表 1 图(4)。针对上述不同的情况,文中算法都取得了比较好的文字提取结果。

3 结束语

文中运用小波插值和 K-means 聚类算法对彩色目标图像进行放大和彩色图像聚类文字提取,通过处理各种形式的文字区域,表明该方法有一定的通用性。由于彩色图像背景相对比较复杂,如果直接二值化,会将文字和背景融合到一起,无法进行文字的分割,也很难进行识别,所以需要根

据实际图像选取不同处理方式。

由于静态图像中文字检测和提取方法各自具有其局限性和应用范围,在实际应用中为了实现有效地提取文字的目的,还应该结合其它多种方法,使得算法通用性更强。

参考文献:

[1] 丁晓青. 汉字识别研究的回顾[J]. 电子学报, 2002(9):1363-1367.

[2] Li Huiping, Doermann D. Automatic text detection and tracking in digital video [R]. USA: Maryland University, 1998.

[3] Zhong Y, Karu K, Jain A K. Locating text in complex color images [J]. Pattern Recognition, 1995, 28(10):1523-1536.

[4] Jain A K, Bhattacharjee S. Text segmentation using Gabor filters for automatic document processing [J]. Machine Vision and Applications, 1992, 5(3):169-184.

[5] Otsu N. A threshold selection method from

6 实验结果及分析

文中从 CASIA2 虹膜数据库选取了 900 多幅虹膜图像作为实验样本,并依照文中提出的算法对其进行了虹膜定位实验,定位的准确率达到了 94.5%,与 Hough 变换及 Canny 检测相比,有着更高的准确率和更快的速度,Hough 变换在此样本中的定位平均耗时达 14.7s,Canny 检测的定位耗时也达 11.8s,而文中的定位算法只要 2.2s,能够满足实时性的要求。在虹膜识别算法测试时,同类样本之间和不同类样本之间的匹配结果统计如图 6 所示。

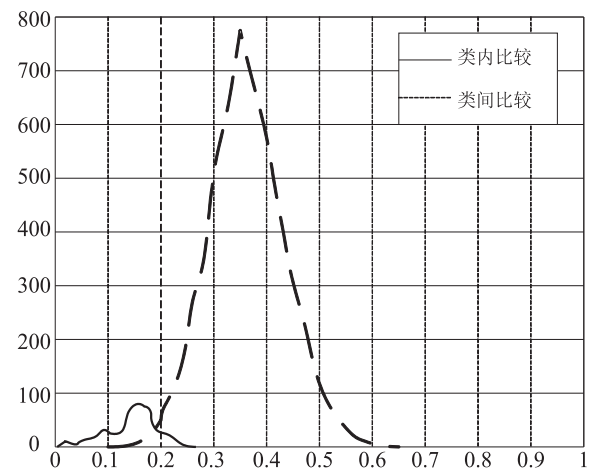


图 6 Hamming 距离匹配结果

从图中可以看出同类样本和不同类样本的匹配结果有着很大的差异,利用虹膜图像的纹理特征进行海明距离的求解可以非常有效地进行虹膜图像的识别。

7 结束语

实验结果表明,文中的提出的改进的定位算法优于 Hough 变换及 Canny 检测定位算法,表现为定位速度更快,准确率更高,提出的基于线段提取的虹膜归一化方法,较完整地保留了虹膜的纹理特征信息,更有利

于虹膜特征信息的提取,采用基于二维 Log-Gabo 滤波的方法较完整地提取了虹膜的相位信息,比传统的二维 Gabo 滤波有着更小的等错误率和相同接收率下更小的错误拒绝率,可以更好地进行虹膜识别。

参考文献:

[1] 万里光. 虹膜识别技术的应用与发展[J]. 船电技术,2008, 28(5):308-311.

[2] Richard P W. Iris recognition:an emerging biometric technology[J]. Proceedings of IEEE,1997,85(9):1348-1363.

[3] Daugman J G. How iris recognition works[J]. IEEE Transaction on Circuits and Systems for Video Technology,2004,14 (1):21-30.

[4] 王风华,韩九强. 基于 2D Log-Gabor 滤波器的虹膜识别研究[J]. 系统仿真学报,2008,20(7):1808-1811.

[5] 李忠科. 虹膜识别系统的若干算法研究[D]. 长沙:长沙理工大学,2011.

[6] Ma Li,Tan Tieniu,Wang Yunhong. Efficient iris recognition by characterizing key local variations[J]. IEEE Transactions on Image Processing,2004,13(6):738-749.

[7] 苑玮琦,冯琪,白晓光. 基于 2D-Gabor 滤波器的虹膜噪声检测方法[J]. 光子学报,2010,39(2):369-373.

[8] 袁晓燕. 虹膜定位、形变及特征提取研究[D]. 上海:上海交通大学,2008.

[9] Yuan Weiqi,Xu Lu,Lin Zhonghua. Iris localization algorithm based on gray distribution features of eye images[J]. Journal of Optoelectronics Laser,2006,17(2):226-230.

[10] 唐崇武,李会方,何俊婷,等. 改进的直方图均衡图像增强算法及实现[J]. 信息安全与通信保密,2009,12(3):72-74.

[11] 唐荣年,韩九强,张新曼. 一种 Log-Gabor 滤波器结合多分辨率分析的虹膜识别方法[J]. 西安交通大学学报,2009 (4):30-33.

[12] 郑沁锋,李晓宁. 基于 Gabor 滤波的快速虹膜识别算法[J]. 计算机工程与设计,2011(3):976-979.

(上接第 33 页)

gray level histograms[J]. IEEE Transactions on Systems,Man and Cybernetics,1979(1):62-66.

[6] Ye Q, Gao W, Huang Q. Automatic text segmentation from complex background[C]//IEEE International Conference on Image Processing. Singapore:IEEE Press,2004:2905-2908.

[7] Chaddha N. Text segmentation in mixed mode images[C]//28th Asilomar Conference on Signals, Systems and Computers. Pacific Grove, CA, USA:IEEE Press,1994.

[8] Fu Ping,Li Meng,Yin Hongtao. Gabor Filter Based Text Extraction from Digital Document Images[J]. 电子学报,2006 (z1):2387-2390.

[9] Cai X C,Keyes D E,Melvin R G,et al. Parallel Newton-Krylov-Schwarz algorithms for the transonic full potential equation [J]. SIAM Journal on Scientific Computing,1998,19(1):246-265.

[10] 庞茜,侯正信,杨爱萍,等. 图像修复的邻域差值扩散模型[J]. 计量学报,2009(5):409-412.

[11] 谢安明,邵美珍,李弼程. 基于小波的图像放大模型分析[J]. 信息工程学院学报,2000(2):57-60.

[12] 钟铭华. 小波插值与自蛇模型在放大图像清晰化中的应用研究[D]. 昆明:云南大学,2012.

结合小波插值与K-means方法提取彩色图像文字

作者: [张凯歌](#), [缪毅](#), [雷建坤](#), [彭可亮](#), [蒋慕蓉](#)
作者单位: [云南大学 信息学院 计算机科学与工程系, 云南 昆明650091](#)
刊名: [计算机技术与发展](#)
英文刊名: [Computer Technology and Development](#)
年, 卷(期): 2013(3)

本文链接: http://d.g.wanfangdata.com.cn/Periodical_wjfz201303010.aspx