

基于 Matlab GUI 的维吾尔文字符识别系统的设计

伊力哈木·亚尔买买提, 哈力旦·A

(新疆大学 电气工程学院, 新疆 乌鲁木齐 830008)

摘要: 基于建立对视频中维吾尔文字符的识别提取系统, 以视频中维吾尔字符为对象, 首先对维吾尔文字符进行了字符检测, 以便确定其维吾尔文字符在视频中的大小、位置, 其次对所确定的维吾尔文字符进行字符定位以便作进一步的处理, 然后运用 Canny 算子将其提取出的维吾尔文字符进行彩色图像转换为灰度化的边缘图像, 最后利用 Matlab GUI 设计了维吾尔文字符的识别系统, 实现了视频中维吾尔文字符的检测、字符定位、图像抖动、对比度调整、灰度化、边缘检测、提取字符等功能。实验结果均显示出该算法的优良性能, 并证明了该识别系统性能的稳定性和极其良好的扩展性, 为维吾尔文字符的识别研究提供了一个简单有效的仿真平台。

关键词: 视频; 维吾尔文字符; 识别; 提取字符; 算法

中图分类号: TP391.43

文献标识码: A

文章编号: 1673-629X(2010)11-0092-03

Designs of Uyghur Character Identify System Based on Matlab GUI

Yilihamu Yaermaimaiti, Halidan A

(College of Electric Engineering, Xinjiang University, Urumqi 830008, China)

Abstract: Based on the establishment of the video extracts Uyghur character recognition system, with Uyghur characters in the video as object, firstly characters on the character Uyghur testing to determine its Uyghur characters in the video size, location, followed by the right set the Uyghur character for character positioning for further processing, and then use Canny operator to extract the Uyghur characters of color image is converted to gray scale edge image, finally using Matlab GUI Uyghur character recognition system to realize the video detection Uyghur character, character positioning, image dithering, contrast adjustment, gray level, edge detection, extraction of characters and so on. Experimental results show excellent performance of the algorithm and prove the performance of the recognition system the stability and very good scalability, as the Uyghur character recognition provides a simple and effective simulation platform.

Key words: video; Uyghur character; identify; character distill; algorithm

0 引言

在视频中包含众多的文字信息, 这些信息与内容丰富的视频是密切相联系的, 经常被用来标记为基于内容的视频检索视频语义线索^[1]。为了方便在网络上搜索并访问视频中的维吾尔文字符并改善维吾尔文字符信息的传播和利用率, 在文中以维吾尔语视频为例, 给定一个视频序列中维吾尔文字符检测与获取方法, 为实现对视频检索系统奠定了理论基础。

科学计算标准软件 Matlab 以其丰富的数据类型和结构、友善的面向对象、快速的图形可视、广博的应用开发工具在国际控制界得到了广泛的应用^[2]。利用

Matlab 强大的图形对象属性设置技术及 GUI 图形用户界面而制作技术为维吾尔文字符的识别技术提供了良好的人机界面^[3]。

1 系统设计思路

1.1 维吾尔文字符的特点

维吾尔语属于阿尔泰语系突厥语组, 被新疆地区的七百六十万人使用。维吾尔语言以及在新疆地区使用的哈萨克、柯尔克孜等语言都借用了部分阿拉伯文和波斯文字母书写。维吾尔字符由 32 个字母组成, 而且有 120 多个文字形式, 但笔画简单, 与由统一大小方形字组成的汉字和由拉丁字母拼写而成的英文有着明显差异, 这些差异说明维吾尔文字符的识别不能直接应用汉字和英文的识别技术^[4]。

1.2 设计思路的提出

图形文字可分为两类: 现场文字和图像文字。现场文本出现在现场中, 是在视频的演化过程中获得的

收稿日期: 2010-03-23; 修回日期: 2010-06-21

基金项目: 国家自然科学基金(60865001)

作者简介: 伊力哈木·亚尔买买提(1978-), 男(维吾尔族), 讲师, 硕士, 研究方向为图像文字处理、模式识别; 哈力旦·A, 教授, 硕士生导师, 研究方向为多媒体通信、数字图像处理。

文字信息。图像文字是人为添加到字符视频信息中。图像文本阅读简单,符合一般人的阅读习惯,并在单一视频中,文本动态变化效果已经消失,只受复杂背景的影响^[5]。维吾尔文字符在 MTV 视频中表现的特点十分复杂,主要表现在以下几点:首先维吾尔文字符笔画复杂、繁琐,在识别中容易把维吾尔文字符某些笔画当成符号而来辨认;其次在视频中,维吾尔文字符的尺寸大小不同,当维吾尔文字符尺寸太大时,文本会遮住其后面的图像内容,而当维吾尔文字符的尺寸太小时,对维吾尔文字符的定位造成了困难。然后视频中绝大多数的维吾尔文字符与背景之间具有较强的边缘,通过检测边缘方法可提高维吾尔文字符提取的性能。最后由视频中多个维吾尔文字符构成的文本具有一种独特的纹理,文本区域内亮度的方差比较大,这对维吾尔文字符的识别提供了较好的条件^[6]。

鉴于此,笔者研究了一种从视频流的关键帧中提取维吾尔文字符的区域并且分离出字符的新方法,并通过 Matlab 视频图像处理工具箱进行了仿真和实现,为在复杂背景下提取维吾尔文字符提供了一个良好的解决方案。

2 维吾尔文字符识别算法的提出

首先对维吾尔文字符进行了字符检测,以便确定其维吾尔文字符在视频中的大小、位置,其次对所确定的维吾尔文字符进行字符定位以便作进一步的处理,然后运用 Canny 算子将其提取出的维吾尔文字符进行彩色图像转换为灰度化的边缘图像^[7],最后对图像中的维吾尔文字符进行精确的提取文本信息,实验结果表明了该算法的有效性。

2.1 算法的总体框架

视频中的文字有以下特征^[8]:

- (1) 文字大多数为单向,沿着水平方向排列形式融合在一起;
- (2) 文字大多数为单一颜色,与背景对比明显;
- (3) 字符的大小有一定的轮廓;
- (4) 文字的笔画多样,其边缘轮廓信息比背景复杂且丰富。

鉴于以上特征,研究出的检测算法框架如图 1 所示。

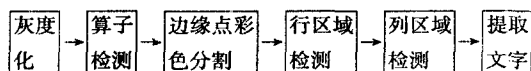


图 1 检测算法框架图

2.2 算法中关键技术的实现

2.2.1 Canny 算法检测

Canny 算法临界检测先计算出 x 和 y 方向的梯度

最大幅值平方和^[9],其局部梯度最大值超过给定值的像素点检测为临界边缘。其保留最大值局部改变最大点的进程称为非极大值抑制。Canny 算子的临界检测有如下 3 个指标:①对每一个边缘点有单独的响应,得到一个像素宽的边缘;②发生错误极低,既要少丢失真正的边缘点,也要少将其非边缘点裁定为边缘点;③定位精度高,边缘检测应当是真正的边界。为此,Canny 相应的总结了 3 个准则函数以表达对这些指标的限制:a.单边缘反应准则函数;b.定位精确度准则函数;c.信噪比准则函数。将 3 个准则相联系可以检测得最好的边缘^[10]。

Canny 算子中有 3 个数据参数, σ 是高斯函数的分散函数,它掌控平滑度、高阈值(TG)及其低阈值(TL)。实验中 σ 设定为 1, $TL=0.4 \times TG$ 。即在 Canny 算子中仅仅保存了单个参数高阈值(TG)。在实际调用中,TG 取值仅仅需要满足达到两个条件^[11]:①尽可能地计算检测出文本字符的边缘点,这大致限制了 TG 取值的最大上限;②假边缘尽可能的少,这限制了 TG 取值的最大下限。由于文本字符和背景对比度明显,TG 有一个很大的取值范围^[12],在实验中取 $TG=0.2$ 。Canny 算子边缘检测结果如图 2 所示。

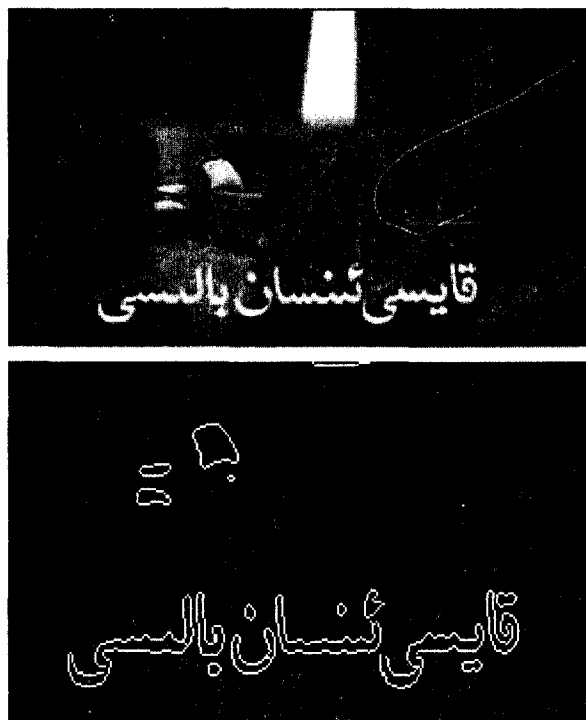


图 2 Canny 算法检测的边缘

2.2.2 文字临界点彩色分割

文本图像有一些特制的特点,文本图像的结构更复杂,边缘像素点较多,且更多集中分布。受光线条件和不同的文字颜色对降解不同的影响,文本文字局部特征如亮度、对比度等更易干扰,但是文字的边缘分布

点有较强的抗干扰性能。Canny 算子检测到的临界点在 HIS 色彩空间 S 分量上进行文字临界点分割^[13]。

$$BW_Word(i, j) = \begin{cases} 1 & BW(i, j) = 1 \\ 0 & \text{其他} \end{cases}$$

式中, $BW(i, j)$ 是 Canny 算子检测的边缘, $BW_Word(i, j)$ 是文字的临界点。

3 系统软件设计

3.1 系统软件设计框图

在此所研究的基于彩色图像的文本提取系统中, 一幅彩色图像(可以是 BMP, JPEG, GIF, TIF 等格式)输入系统后, 经过字符检测、字符定位、图像抖动、对比度调整、灰度化、边缘检测、提取字符等这一系列过程可以用如图 3 的系统框图详细地描述。

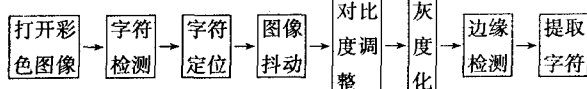


图 3 系统框图

3.2 程序运行结果演示

本系统在 Windows 环境下用 Matlab 实现, 经过系统界面的设计, 通过对其视频中彩色图像维吾尔文

字符检测、字符定位、提取字符等过程的实现, 可以演示其最终的系统程序结果。系统的整体界面及其演示结果如图 4 所示。

3.3 实验结果分析

为了测试不同情况下的维吾尔文字效果, 选择了不同的复杂情形, 运用了视频中维吾尔 MTV 来测试和提取。用 Media Player 收集了很多图像作为检测数据, 这些样本数据包括不同复杂背景情况下不同字符、不同大小的字符特征区域, 拥有广泛的代表性。所使用达到帧的速率平均为 24.99 的图像进行了检测试验, 在进行检测之后其提取文字的准确性达到了满意效果。

4 结束语

从以上对维吾尔文字的识别系统中, 可以看出 Matlab GUI 的系统界面具有直观、方便、简洁的特点, 对于研究维吾尔文字的识别提供了良好的人机界面。相信随着 Matlab 版本的升高, 其功能的不断完善, 会使得 Matlab 在维吾尔文字识别应用中发挥更大的作用。

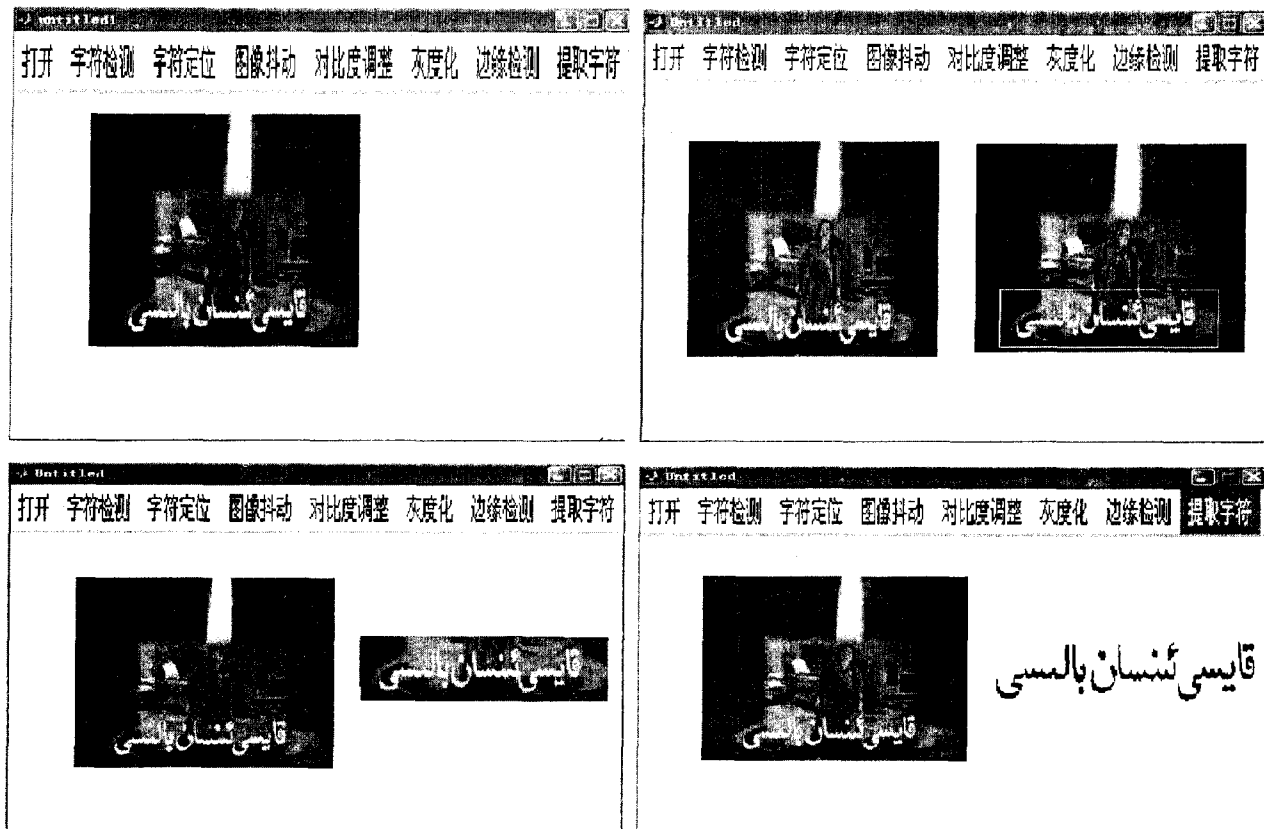


图 4 系统及演示结果

参考文献:

- [1] Pezner L, Hearst M. Acritique and improvement of anevalua-

tion metric for text segmentation[J]. Computational Linguis-

(下转第 99 页)

表 1 基本遗传算法与文中算法的比较

遗传算法	数据集	最大偏差	最小偏差	达到最优解的次数	达到最优解的平均迭代次数
基本遗传算法	student	79.0120	76.53114	13	4.3
	teacher	526.6243	330.35464	—	—
	parents	2.56420e5	2.198754e5	16	6.0
文中算法	student	75.6520	75.6510	14	7.8
	teacher	315.2185	315.2175	15	1.2
	parents	2.12051e5	2.12041e5	16	4.1

4 结束语

文中提出了一种基于加权二叉树的自适应遗传算法,通过构建遗传因子二叉树,利用满二叉树实现遗传因子的变异操作,并通过蚁群算法求出蚁群信息素对其进行加权,采用最短距离基因匹配的方法作为交叉操作的基础,引入自适应排序在后代中选择优秀个体进行交叉,选择最优解。实验结果表明该算法易操作、快速、准确、能够较好地克服局部最优的缺陷,大大提高了基本遗传算法的局部搜索能力和收敛速度,具有很好的应用前景。

参考文献:

- [1] Dorigo M, Maniezzo C A. The ant system: optimization by a colony of cooperating agents[J]. IEEE Trans. on System, Man and Cybernetics, 1996, 26(1): 1-13.
- [2] Dorigo M, Gambardella L M. Ant colony system: a cooperative learning approach to the traveling salesman problem[J]. IEEE Trans. on Evolutionary Computation, 1997, 1(1): 53-66.
- [3] 曹道友,程家兴. 基于改进的选择算子和交叉算子的遗传算法[J]. 计算机技术与发展, 2010, 20(2): 44-45.
- [4] Holland J. 遗传算法的基本理论与应用[M]. 李敏强, 译. 北京: 科学出版社, 2003.
- [5] Mantawy A H, Abdel - Magid Y L, Selim S Z. Integration genetic algorithm, tabu search, and simulated annealing for the unit commitment problem[J]. IEEE Transactions on Power Systems, 1999, 14(3): 829-836.
- [6] Murthy C A, Chowdhury N. In search of optimal clusters using genetic algorithms[J]. Pattern Recognition Lett., 1996, 29(3): 825-832.
- [7] 张应辉, 王志伟, 曾庆华. 基于蚁群信息素的遗传操作算法[J]. 计算机科学, 2007, 34(6): 170-173.
- [8] 冀俊忠, 黄 振, 刘椿年. 一种快速求解旅行商问题的蚁群算法[J]. 计算机研究与发展, 2009, 46(6): 968-978.
- [9] 刘芳华, 赵建民, 朱信忠. 基于改进遗传算法的物流配送路径优化的研究[J]. 计算机技术与发展, 2009, 19(7): 83-84.
- [10] 周洪伟, 原锦辉, 张来顺. 遗传算法早熟现象的改进策略[J]. 计算机工程, 2007, 33(19): 201-203.
- [11] 冯冬青, 王 非, 马 雁. 遗传算法中选择交叉策略的改进[J]. 计算机工程, 2008, 34(19): 189-190.
- [12] 陆林花, 王 波. 一种改进的遗传聚类算法[J]. 计算机工程与应用, 2007, 43(21): 171-172.
- [13] 李 媛, 卡米力·毛依丁. 维吾尔语笔迹鉴别方法研究[J]. 计算机技术与发展, 2008, 18(5): 9-11.
- [14] 力晓光, 李晓华, 沈兰荪. 基于纹理的图像字符自动定位技术对比研究[J]. 电路与系统学报, 2006, 11(2): 258-260.
- [15] 刘晓明, 仲元红, 欧静兰. 基于 DSP 的火灾图像识别系统设计及应用[J]. 计算机技术与发展, 2006, 16(6): 96-100.
- [16] Chen Datong, Jean - Marc O, Hervé B. Text detection and recognition in images and video frames[J]. Pattern Recognition, 2004, 37(3): 595-608.
- [17] 杨述斌, 张 阳. 复杂车辆图像中的车牌快速形态定位算法[J]. 计算机技术与发展, 2008, 18(6): 50-53.
- [18] 杨 胜, 钟玉琢. 一种从 MPEG 压缩视频流中提取关键帧的方法[J]. 中国图像图形学报, 2001, 6(3): 254-258.
- [19] Lienhart R, Wernicke A. Localizing and segmenting text in images and videos[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2002, 12(4): 256-268.
- [20] 段汉根, 汪继文. 基于邻域滤波的图像修复[J]. 计算机技术与发展, 2007, 17(10): 34-36.

(上接第 94 页)

- [1] Passoneaur, Litmand. Discourse segmentation by human and automated means [J]. Computational Linguistics, 1997, 23(1): 103-139.
- [2] 周晓兰, 张 杰. MATLAB 在通信系统仿真中的应用[J]. 计算机技术与发展, 2006, 16(9): 166-168.
- [3] 蔡 群, 周美莲, 段杰峰, 等. 基于 Matlab 分布式工具箱的流场计算及其可视化[J]. 计算机技术与发展, 2007, 17(9): 51-54.
- [4] 哈力旦·A, 伊力哈木·亚尔买买提, 库尔班·买提木沙. 复杂背景下维吾尔文字符的分割算法[J]. 计算机工程与应用, 2007(20): 163-165.