

一种基于运动图像的在线手写汉字识别方法

高 学¹, 徐 睿¹, 金连文¹, 尹俊勋¹, 镇立新²

(11 华南理工大学电子与通信工程系, 广东广州 510641; 21 Motorola China Research Center, 上海 200002)

摘 要: 在分析传统联机与脱机手写汉字识别系统基础上, 本文提出了一种基于运动图像的在线手写汉字识别新方法. 该方法利用普通摄像头, 实时地采集手写汉字墨迹的运动图像, 通过图像处理技术提取手写汉字, 并进行在线识别. 文中给出了笔尖匹配的快速算法和基于数学形态学的汉字提取方法. 最后, 通过实验验证了本文方法的有效性.

关键词: 手写汉字识别; 运动图像; 数学形态学

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2004) 08-1273-04

A Video2Based Online Handwritten Chinese Character Recognition Method

GAO Xue¹, XU Rui¹, JIN Lianwen¹, YIN Junxun¹, ZHEN Lixin²

(11 Dept1 of Electronics and Communication Engineering, South China University of Technology, Guangzhou, Guangdong 510641, China;
21 Motorola China Research Center, Shanghai 200002, China)

Abstract: On analysis of online and offline handwritten Chinese character recognition system, a video2based Online Handwri2ten Chinese Character Recognition Method is proposed. Through a video camera, the handwriting on the standard paper is captured in real time. The handwritten Chinese characters are extracted by applying image processing technique, and recognized online. We present a fast pen2tip matching algorithm and a mathematical morphology based handwritten Chinese character extraction approach. The exper2mental results show the efficacy of the proposed approach.

Key words: handwritten Chinese character recognition; video image; mathematical morphology

1 引言

智能信息处理是当今信息化社会发展的关键技术, 文字作为人类交流的基本媒介之一, 文字信息是一种重要的信息来源. 手写汉字识别的研究目标就是要解决汉字的自动录入问题. 目前的手写汉字识别系统可分为脱机和联机两种工作方式. 脱机手写汉字识别通过采集(如使用扫描仪)汉字文本图像, 然后, 以脱机的方式进行文本图像处理和汉字识别. 该方法缺乏实时处理能力, 而且, 由于难以提取笔画的时序信息, 目前, 这类系统的识别率仍未达到较为满意的水平. 联机手写汉字识别则采用专用的书写板和笔, 以便实时地采集手写汉字图像以及笔画时序信息. 尽管联机手写汉字识别研究已经取得了比较满意的识别结果, 但由于需要专用的书写设备, 而且书写不如普通的纸和笔自然, 使得联机手写汉字识别系统的应用受到了一定限制. 基于以上分析, 本文尝试了一种介于脱机和联机之间的在线手写汉字识别新方法, 即基于运动图像的在线手写汉字识别. 该方法利用普通摄像头, 实时地采集人在普通纸上书写汉字的墨迹图像, 通过图像处理技术, 分割和提取手写汉字, 并进行实时识别. 该方法不仅可以实时地进

行汉字输入, 而且所采集的图像序列中包含了汉字笔画的时序信息, 因而有望能提高汉字的分割和识别性能. 随着电子技术的发展, 越来越多的智能信息处理设备和终端配备了摄像设备. 该系统的研制将能实现一种全新的、更为自然的文字信息录入方法, 并在各种智能信息处理设备中有着广阔的应用前景.

目前, 基于运动图像的在线手写字符识别研究^[1~5]尚处于初步探索阶段. 文献[1, 2]通过相邻帧图像的差分得到笔画的时序信息. 由于笔和手及其阴影的影响, 该方法提取时序信息必须待字符书写完成之后, 需要存储大量的中间帧图像, 文献中未涉及识别问题. 文献[3, 4]通过跟踪笔尖来重建笔尖的运动轨迹, 然后, 根据起笔和落笔的判断得到笔画时序信息, 并应用于英文签名识别中. 该方法的缺点是对笔尖的形状敏感. 如果不满足其假定的笔尖形状, 起笔和落笔的判断将会产生较大误差. 文献[5]的方法类似于文献[2], 文中给出了一种减少笔与手的阴影影响的算法. 该方法的前提只书写一个汉字, 未涉及多个汉字与段落的情况, 以及相应的识别问题. 本文提出的方法通过实时地处理手写汉字的墨迹图像, 并提取和分割手写汉字. 然后, 利用脱机手写汉字识别的方法进行在

线识别. 实验结果表明, 本文的方法是行之有效的. 为了提高系统的实时性, 文中提出了笔尖跟踪的快速算法, 以缩小待处理区域的大小. 本文以下首先给出系统的基本组成, 然后介绍笔尖匹配的快速算法, 以及基于数学形态学的汉字提取方法. 最后给出试验结果以验证本文方法的有效性.

2 系统原理

识别系统主要由笔尖模板的初始化、笔尖匹配、图像处理区域的确定、手写汉字的提取和识别五个部分组成, 如图 1 所示.

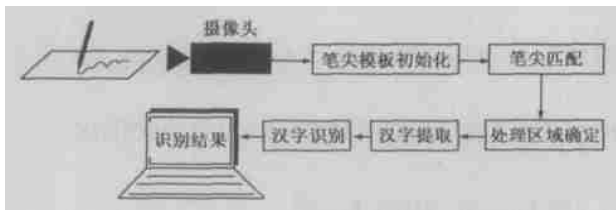


图1 系统流程图

对于实时采集的每帧图像, 系统首先通过相关匹配确定笔尖的位置. 设 $g(i, j)$ $i, j = 1, \dots, \text{pensize}$ 为笔尖模板, pensize 为笔尖模板的大小, $f_t(i, j)$ $i, j = 1, \dots, M; j = 1, \dots, N$ 为 t 时刻采集的灰度图像. 点 (i, j) 处与笔尖模板的相关匹配度量定义为:

$$c(i, j) = \frac{\sum_{m=1}^{\text{pensize}} \sum_{n=1}^{\text{pensize}} f_t(i+m, j+n) g(m, n)}{\left\{ \sum_{m=1}^{\text{pensize}} f_t^2(i+m, j+n) \right\}^{1/2} \left\{ \sum_{m=1}^{\text{pensize}} g^2(m, n) \right\}^{1/2}} \quad (1)$$

则有

$$c(i, j) \leq 1 \quad (2)$$

可以证明, 当且仅当满足式(3)时, 式(2)取等号.

$$f_t(i+m, j+n) = S g(m, n), (m, n) \in D_g \quad (3)$$

笔尖位置的确定采用以下准则:

如果点 (m, n) 满足条件

$$c(m, n) = \max_{(i, j) \in D_f} c(i, j) \quad (4)$$

$$c(m, n) > T_c \quad (5)$$

则点 (m, n) 为确定的笔尖位置. 否则, 不处理该帧图像. 式中 T_c 为设定的相关匹配度量的阈值.

对于每个确定的笔尖位置, 以该点为中心确定一个大小为 $W_h @ W_v$ 的图像区域为待处理区域. 其中 W_h 和 W_v 分别为处理区域的宽与高, 其值根据前一帧图像所提取汉字的大小进行在线更新, 分别取为汉字宽与高的 k 倍 ($k = 2 \sim 3$). 然后, 应用本文提出的基于数学形态学的方法提取手写汉字, 并进行在线识别. 为了避免手写汉字重复输出, 需要根据汉字的位置信息判断该汉字是否已经识别输出. 如果已经输出, 则更新前一个汉字的识别结果. 否则, 直接输出.

为了使系统不敏感于特定的笔尖形状, 在进行笔尖匹配和汉字提取之前, 需要进行笔尖模板的初始化, 以自适应地获取当前的笔尖匹配的模板. 根据相邻帧的差分图像, 可以判断笔尖是否已经进入图像区域, 待笔尖稳定在图像中部位置一定时间, 则利用系统内置的典型笔尖模板, 根据式(4)和(5)确

定笔尖位置. 然后, 以笔尖位置为坐标原点采集相应 $\text{pensize} @ \text{pensize}$ 大小的图像作为匹配的笔尖模板.

2.1 笔尖匹配的快速算法

为了减少笔尖匹配的计算量, 提高系统的实时性, 本文给出了一种基于粗匹配和精确匹配的两级笔尖快速匹配搜索算法. 该方法首先利用粗匹配确定笔尖位置的候选集合, 然后在候选集合中, 利用式(1~5)进行笔尖的精确定位. 粗匹配采用以下两种策略:

策略1: 相关匹配度量的计算只在笔尖模板像素点的子集进行, 如图2所示, 笔尖模板的子集仅由图2中的黑像素点组成.

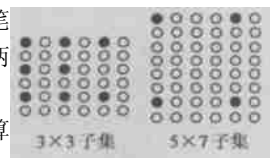


图2 粗匹配的笔尖模板子集示意图

策略2: 将笔尖模板和当前帧图像的相应匹配区域分别向水平轴作投影, 则可以得到两个维数为 pensize 的一维投影向量. 然后, 计算相应向量的相关匹配度量, 并根据匹配度量的大小确定笔尖位置的候选集合. 图3所示分别为笔尖所在行图像的水平投影的一部分与笔尖模板的投影. 有图中可以看出, 笔尖所在位置图像投影与笔尖模板的投影具有较好的相似性.

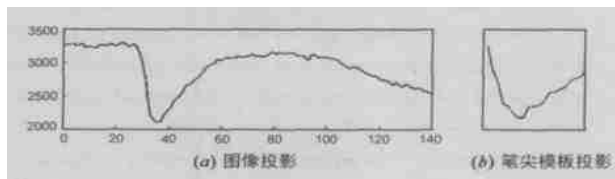


图3 笔尖所在行图像和笔尖模板的水平投影曲线

2.2 汉字提取

由于环境光线条件的影响, 以及笔和手的存在, 采集的图像中不仅包含手写汉字墨迹, 而且还存在笔与手及其阴影等噪声因素, 如图4(a)所示. 根据手写汉字笔画的结构特点, 本文提出了一种基于数学形态学的手写汉字提取方法. 该方法可以有效地去除噪声因素的影响, 从而提取正确的手写汉字.

数学形态学作为一种新的数字图像处理理论和方法, 对于图像的几何特征或基本结构具有较强的描述能力^[6,7]. 数学形态学的两种最基本运算为膨胀和腐蚀. 根据待处理图像为二值图像或灰度图像, 运算也具有不同的表达形式, 本文的处理对象为灰度图像.

设 $f(s, t)$ 为输入灰度图像, $b(m, n)$ 为形态学的结构元素. 则有

$$(f \odot b)(s, t) = \max_{(m, n) \in D_b} \{f(s-m, t-n) + b(m, n)\} \quad (6)$$

$$(f \oslash b)(s, t) = \min_{(m, n) \in D_b} \{f(s+m, t+n) - b(m, n)\} \quad (7)$$

其中 $(s, t) \in D_f$, $(m, n) \in D_b$, D_f 、 D_b 分别为输入图像函数和结构元素的定义域, \odot 、 \oslash 分别表示膨胀和腐蚀运算.

可以看出, 膨胀和腐蚀运算分别为在结构元素所定义的范围取最大、最小值. 经过膨胀运算, 图像中的黑色细节将减弱, 而腐蚀运算则使图像中的明亮细节减弱. 细节的减弱程度取决于图像的灰度值分布与结构元素的形状和幅值. 由于

手写汉字笔画具有显著区别于手与笔等噪声因素的几何形状,且笔画宽度远小于噪声因素的尺寸,如果设计适当的形态学结构元素,经过膨胀运算,将可以有效地从图像中去掉汉字笔画,同时保留噪声因素.这样,经过与原图像的相减运算,就能够从图像中分离出手写汉字.然而,膨胀运算不仅能去掉手写汉字笔画,也同样使图像中手和笔的面积减小.这样,经过图像的相减运算,手和笔的轮廓将会出现在图像中.为了消除手和笔等的轮廓,膨胀后的图像再以相同的结构元素做腐蚀运算,则手和笔的面积就能够恢复到原始图像的大小.从而,相减得到图像中将只保留手写汉字的笔画.

在数学形态学处理方法中,不同的结构元素形状将直接影响到汉字提取的效果.结构元素的确定一般要根据图像中目标物体与背景中其他物体的形状差异,选择能够反映目标物体形状特征的结构元素.考虑到汉字主要由/横0、/竖0、/撇0、/捺0四种基本笔画组成,其宽度远小于手和笔等噪声因素的宽度,本文采用 $N@N$ 的正方形结构元素,其值均为 1,该结构元素能够适应不同方向的汉字笔画.图 4 所示为一幅图像的处理结果.

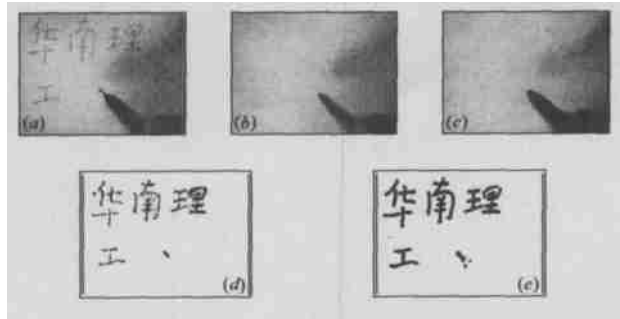


图 4 7@7模板的形态学处理结果.(a)原图;(b)膨胀后的图像;(c)腐蚀后的图像;(d)相减并取反后的图像;(e)二值化结果

由图 4(d)可以看出,经过相减运算后,有效地分离出了仅包含手写汉字的图像.图像经过二值化和行字分割,可以得到每个手写汉字.

文本的行字分割是字符识别研究中的另一个关键问题.由于本文方法的图像处理区域只是包含笔尖所在位置的一个有限邻域,行和字的倾斜对分割结果产生的影响很小,因而有利于改善行字分割的效果.本文所给出的为利用行和列投影方法进行分割的结果.而对于连笔字的分割问题,由于篇幅限制,文中未涉及.

3 实验结果

为了检验所提出方法的有效性,本文进行了实验测试.图像采集采用普通 USB 摄像头,图像分辨率为 320@240,采样率为 5 帧/秒,相对于普通的汉字书写速度(1~2 字/秒),采样率可以满足系统的实时性要求,主机 CPU 为 PIII700.表 1 给出了本文的两种笔尖匹配快速算法的测试结果,其中笔尖位置的候选集合的大小设定为 50.针对不同形状笔尖的实验表明,笔尖位置的候选集合大小为 50 可以满足笔尖匹配的精度要求.

表 1 两种笔尖匹配的快速算法的性能

匹配策略	策略 1			策略 2	全局匹配
	4@4子集	5@5子集	6@6子集		
时间(ms)	100	50	30	50	600

由表 1 可以看出,相对于直接进行全局匹配,本文的快速匹配算法较大地降低了笔尖匹配的运算时间.

表 2 给出了系统各模块的运算时间测试结果.其中图像待处理区域初始化为以笔尖为中心的一个 80@60 区域,k 取 216.可以看出,笔尖的快速匹配算法可以满足系统实时性的要求.

表 2 系统各模块的时间分配

	笔尖匹配	汉字提取	识别	图像显示	帧间隔
时间(ms)	50	20	20	7	200

为了检验本文的汉字提取方法与系统的整体识别性能,本文分别对 5 位不同书写者进行了测试,其中每人在线书写 10 个不同汉字.图 5 为其中一人的书写样本与汉字提,图(5) a 为在线书写时的最后一帧图像.

由图 5 可以看出,本文方法有效地提取了每个手写汉字.

汉字识别特征采用 Gabor 特征,并应用 LDA 方法将特征压缩到 158 维^[8].分类器为欧式距离分类器,分类器的字典包括国标一级字库的 3755 个汉字.

分类器的训练采用国家 863 手写汉字样本库 HCL2000 作为训练样本,平均汉字识别率为 85%,证明本文方法是行之有效的.

然而,手写汉字识别系统作为一个整体,分类器性能的好坏直接影响系统的整体性能.由于受图像分辨率及纸张大小的限制,所提取汉字的分辨率要比 300DPI 的 HCL2000 低得多(图示的汉字样本分辨率约为 120DPI 左右),这是影响汉字识别率的一个重要因素.如果利用所提取较低分辨率的手写汉字样本训练分类器,改进分类方法,将会使系统的识别率得到进一步的改善.

4 结论

本文在分析传统的联机与脱机手写汉字识别系统的基础上,提出了一种基于运动图像的在线手写汉字识别新方法.该方法利用普通摄像头,实时地采集手写汉字墨迹的运动图像,通过图像处理技术提取手写汉字,并进行在线识别.本文的贡献在于提出了笔尖匹配的快速算法和基于数学形态学的手写汉字提取方法.由实验结果可以看出,所提出的笔尖快速匹配算法和手写汉字提取方法是行之有效的.然而,基于运动图像的在线手写汉字识别方法,既与传统的脱机手写汉字识别有相通之处,又具有自己的特点.如果针对其手写汉字分辨率较低的特点,设计相应的分类器,将会提高系统的识别率.与此同时,记录手写汉字墨迹的运动图像中包含了汉字笔画的时序信息,如果能够提取笔画的时序信息,并用于手写汉字的识



图 5 汉字提取结果 (a) 在线书写的最后一幅图像;(b) 提取的汉字

别,系统的识别性能将有望进一步提高. 这些方面的工作将是我们下一步的研究内容.

参考文献:

- [1] Yamasaki T, Hatton T. A new data tablet system for handwriting characters and drawing based on the image processing [A]. Proc. Int. 1 Conf. System, Man and Cybernetics [C]. Beijing, China, 1996. 428-431.
- [2] Bunke H, Siebenhal T V, Yamasaki T, Schenkel M. Online handwriting data acquisition using a video camera [A]. Proc. Int. 1 Conf. Document Analysis and Recognition [C]. Bangalore, India, 1999. 573- 576.
- [3] Muri dh ME, Perona P. Visual input for pen2based computers [J]. IEEE Trans. Pattern Analysis and Machine Intelligence, 2002, 24(3): 313-328.
- [4] Fink G A, Wienecke M, Sagerer G. Video2based on2line handwriting recognition [A]. Werner B. Proc. Int. 1 Conf. Document Analysis and Recognition [C]. Seattle, Washington, USA, 2001. 226- 230.
- [5] Tang X O, Lin F. Video2based handwritten character recognition [A]. Proc. Int. 1 Conf. Document Analysis and Recognition [C]. Orlando, Florida, USA, 2002. 3748- 3751.
- [6] 章毓晋. 图像分割 [M]. 北京: 科学出版社, 2001.

- [7] 阮秋琦. 数字图像处理 [M]. 北京: 电子工业出版社, 2001.
- [8] 覃剑钊, 金连文, 韦岗. 基于 Gabor 滤波器的手写汉字特征提取方法的研究 [A]. 赵明生. 第十二届全国神经计算学术会议 [C]. 北京: 中国电子学会, 2002. 465- 469.

作者简介:



高 学 男, 1967 年出生于河南, 华南理工大学电子与通信工程系讲师、博士, 主要研究方向: 汉字识别, 图像处理, 遗传算法.



徐 睿 男, 1980 年出生于山东, 华南理工大学通信与信息系统专业研究生, 研究方向: 机器视觉, 模式识别, 人机交互等.