

Research on Image Pretreatment Technique in Recognition of Handwritten Number

Cuifang Zhang, Guowei Yang

Qingdao University, Qingdao Shandong
Email: zcf_hy@163.com

Received: May 26th, 2016; accepted: Jun. 13th, 2016; published: Jun. 17th, 2016

Copyright © 2016 by authors and Hans Publishers Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY).
<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Pretreatment is a vital step in recognition of handwritten numerals. The merits of the processing directly affect the result of recognition and the recognition performance of the algorithm. In this paper, pretreatment process for image is discussed and it shows a good result on the database of MNIST.

Keywords

Recognition of Handwritten, Pretreatment, Image Recognition

手写数字识别中图像预处理技术的研究

张翠芳, 杨国为

青岛大学, 山东 青岛
Email: zcf_hy@163.com

收稿日期: 2016年5月26日; 录用日期: 2016年6月13日; 发布日期: 2016年6月17日

摘要

预处理是手写数字识别中重要的一步, 处理结果的优劣直接影响到识别结果的正确与否, 直接影响到识别算法的性能。本文主要讨论手写体数字图像的预处理过程, 在所用的MNIST库上进行实验, 效果良好。

关键词

手写数字识别, 预处理, 图像识别

1. 引言

手写数字识别(Handwritten Numeral Recognition) 是光学字符识别(Optical Character Recognition, 简称OCR)的一个分支[1], 主要研究利用计算机自动辨认人手写于纸张上或电子设备的 10 个阿拉伯数字[2]。手写数字识别技术有很强的综合性, 涉及到图像处理, 模式识别, 人工智能, 模糊数学, 统计决策理论, 模糊数学, 综合数学, 信息论, 计算机等学科, 并且在数据统计、邮政系统、文献检索、票据处理、办公自动化等方面有着非常广阔的应用前景。

手写数字识别虽然只有 0~9 十个字符, 书写和笔画结构都比较简单, 但识别困难在于不同人的书写样式不同, 而且数字的书写一般没有上下文的关联。脱机手写字符的识别一般从纸质载体上采集字符, 因而图像预处理的结果将直接影响识别系统的识别率, 本文主要研究手写数字识别的预处理过程。

字符识别系统一般分为图像预处理、特征提取、字符识别等模块。预处理技术是手写数字字符识别中非常重要的一步, 原始的待识别图像在预处理时被转换成计算机所能接受的二进制形式。减少图像冗余信息、滤出图像中的噪声为图像预处理的主要目的, 即在图像预处理时, 不但与图像无关的信息要去掉, 原图像的特征还要尽量保留完全。

图像预处理时得到的图像信息越好越完整, 后期进行特征提取和字符识别时得到的结果越理想, 如果图像预处理的结果不理想, 不仅会降低系统的识别率, 而且也会影响判别的速度, 从而降低整个系统的性能。对一般手写体数字识别技术, 预处理一般包括二值化、平滑去噪、细化等步骤。

2. 图像的预处理

2.1. 二值化

图像的二值化处理是指将手写体数字图像中所有像素点的像素值处理成仅有黑(像素值 = 0)和白(像素值 = 1)两个灰度值的图像的过程。进行二值化后的图像信息量大大减少, 系统的处理速度得到很大的提升, 同时进行特征提取和字符识别时也更加方便快捷。现有的图像二值化的方法大致可以分为 3 大类: 整体阈值二值化法、局部阈值二值化法和动态阈值二值化法。

整体阈值二值化法, 指图像的阈值是根据已有的经验或灰度图像直方图的分布来确定的。人工设定整体阈值法要根据不同要本图像的实际情况来设定图像的整体阈值, 设 $f(x, y)$ 为灰度图像中 (x, y) 点的灰度值, θ 为二值化算法给定的阈值, 则二值化后的图像 $g(x, y)$ 为:

$$g(x, y) = \begin{cases} 255 & f(x, y) \geq \theta \\ 0 & f(x, y) < \theta \end{cases}$$

这种二值化方法速度最快最简单。在图像数字清晰、轮廓明显、干扰很小的情况下, 该方法的效率特别高。人工设定整体阈值法的缺点是: ① 不能根据每幅图像的具体情况设定最佳的阈值; ② 阈值 θ 确定后, 在外界环境条件改变时, 阈值不能随之改变。

由灰度值直方图确定整体阈值法, 图像整体的阈值是根据图像和背景的灰度值自动确定的, 图像灰度值的整体分布情况由灰度直方图显示出来。设规范化灰度值 g 的范围为 $0 \leq g \leq 1$, $g = 0$ 为最黑, $g = 1$ 为最白。 M 为灰度级数目, $p(g_k)$ 为第 k 级灰度的概率, n_k 是在图像中出现的灰度级为 k 的次数, n 为图

像中象数的总数。则有

$$p(g_k) = \frac{n_k}{n} \quad 0 \leq g_k \leq 255$$

灰度直方图就是以 $p(g_k)$ 为纵坐标, 以 g_k 为横坐标作图得到的图像。手写数字图像的灰度直方图一般有两个峰值, 图像的笔画部分对应一个峰值, 背景部分对应另一个峰值。阈值选取两个峰值处的波谷, 波谷越深越陡, 图像二值化的效果越好。

2.2. 平滑去噪

图像平滑的主要目的是减少图像中的噪声, 在图像处理中噪声是不可避免的问题, 图像会因为噪声的存在而变的不清晰, 特征不明显, 特征提取和识别将会受到严重的影响, 因此高效的处理噪声问题在图像处理中是至关重要的。噪声的种类十分复杂, 对图像的影响也各不相同, 所以要根据噪声的性质选取合适的除噪方法, 才能达到预期的效果。

平滑处理, 就是将一个 $n*n$ 的像素窗口, 依次在二值化字符图像的每个像素点上移动, 利用逻辑表达式来消除孤立像素的一种技术[3]。设定如下图所示的平滑窗口

?	1	?	?	1	?	0	0	?	0	0	0
1	0	1	1	0	1	0	1	1	0	1	1
?	1	?	0	0	1	0	0	1	0	0	?
(1)			(2)			(3)			(4)		

同样包括将上述这些窗口旋转 90° , 180° 和 270° 以后得到的新的平滑窗口。

中值滤波是一种基于排序统计理论的非线性信号细化处理技术, 中值滤波可以有效地解决噪声所带来的图像质量下降问题[4], 该技术不仅能有效的抑制图像中的噪声, 而且还可以有效的保持图像的边缘特性, 图像不会产生显著的模糊。中值滤波的基本思想是, 将一点邻域中的像素值按灰度等级进行排序, 当前点的灰度值用该邻域的灰度的中值代替, 使周围点的像素值更接近真实值, 从而消除图像中孤立的噪声点。中值滤波步骤为: 采用某种结构的二维滑动模板在图像中行走, 并且将模板的中心与图像中某像素点重合; 将模板中各个像素对应的灰度值读出, 并将这些灰度值从小到大或从大到小进行排序; 取出该列数据的中间值, 将模板中心位置的像素值用该值代替。中值滤波的输出为 $g(x, y) = med\{f(x-k, y-l), (k, l \in W)\}$, 其中, $f(x, y)$ 、 $g(x, y)$ 分别为原始图像和中值滤波后的图像, W 为二维模板, 通常为 $3*3$, $5*5$ 区域。因为图像是二维信号, 所以其窗口的形状和尺寸对滤波效果的影响特别大, 往往根据不同的图像和不同的应用要求选择窗口的形状和大小。

2.3. 细化

细化就是连续移除图像的最外层像素, 直到连通的、单位像素的图像骨架产生, 图形用该骨架所取代。中轴线是图像细化后最理想的状态, 好的细化算法一般要满足: (1) 算法的收敛性; (2) 细化后细线的连通性; (3) 原图像的基本形状保持不变; (4) 减少笔画相交处的畸变; (4) 中轴线为细化后的结果; (5) 细化的快速性; (6) 迭代次数少。

根据是否采用迭代运算可以分为非迭代算法和迭代算法。非迭代算法一次就可以产生骨架, 如游程长度编码细化、基于距离变换的方法等。迭代算法需要把图像边缘满足一定条件的像素删除掉, 使骨架变成单像素带宽。根据其检查像素的方法不同又可把迭代方法再分成串行算法和并行算法。串行算法在

是否删除像素在每次迭代的执行中是固定顺序的, 它取决于前次迭代的结果和本次迭代中已处理过的像素点。并行算法是像素点删除与否与像素图像中的顺序无关, 仅取决于前次迭代的结果。

目前已有不少细化算法, 如经典细化算法, Pavlidis 异步细化算, Deutseh 算法, Zhang 快速并行细化算法等。其中 Zhang 快速并行细化算法是一种实用的细化方法, 它具有速度快和保持细化后曲线的连通性等优点。

在 1984 年 Zhang 提出了 Zhang 快速并行细化算法。其最突出的优点是对于直线、T 行交叉和拐角, 经过算法细化后能够比较精确地和原图像保持一致[5]。设 1 标记已知的目标点, 用 0 标记背景点。本身标记为 1 而其 8-连通邻域中至少有 1 个点标记为 0 的点定义为边界点。算法对边界点进行如下操作:

考虑以边界点为中心的 8-邻域, 记中心点为 p_1 , 其邻域的 8 个点顺时针绕中心点分别记为 p_2, p_3, \dots, p_9 , 其中 p_2 在 p_1 上方。首先标记同时满足下列条件的边界点:

$$2 \leq N(p_1) \leq 6;$$

$$S(p_1) = 1;$$

$$P_2 * p_4 * p_6 = 0;$$

$$P_4 * p_6 * p_8 = 0;$$

同第(1)步, 仅将前面条件(1.3)改为(2.3) $p_2 * p_4 * p_8 = 0$; 条件(1.4)改为条件(2.4) $p_2 * p_6 * p_8 = 0$ 。同样当对所有边界点都检验完毕后, 将所有标记了的点除去。

一次迭代算法由以上两步操作构成。直至没有再满足标记条件的点, 剩下点即为运行细化算法后得到的图像的骨架。

3. 结论

图像预处理是手写体数字识别的基础, 特征提取和图像识别直接受到预处理结果好坏的影响。由于手写数字图像目标和背景像素在灰度值上有明显的差异, 采用基于直方图阈值的双峰法进行图像的二值化, 采用 Zhang 并行细化算法很好的保持了原图像的骨架、减少了像素的存储量。通过以上的预处理操作, 图像中与识别无关的信息点被去除, 减少了存储的信息量, 并且保留了原字符图像的特征, 有效的提高了数字识别网络的性能。

参考文献 (References)

- [1] 张猛, 余仲秋, 姚绍文. 手写体数字识别中图像预处理的研究[J]. 微计算机信息, 2008, 22(6-1): 256-258.
- [2] 张光建. 手写数字识别的图像预处理 GUI 设计[J]. 智能计算机与应用, 2015, 5(1): 98-100.
- [3] 瞿海金. 手写体数字识别方法的研究与实现[D]: [硕士学位论文]. 南京: 南京理工大学, 2005.
- [4] 王红君, 施楠, 赵辉, 岳有军. 改进中值滤波方法的图像预处理技术[J]. 计算机系统应用, 2015, 24(5): 237-240.
- [5] 陈梅, 王健, 图像并行细化算法的改进[J]. 中国印刷与包装研究, 2014, 1(6): 30-34.

再次投稿您将享受以下服务：

1. 投稿前咨询服务 (QQ、微信、邮箱皆可)
2. 为您匹配最合适的期刊
3. 24 小时以内解答您的所有疑问
4. 友好的在线投稿界面
5. 专业的同行评审
6. 知网检索
7. 全网络覆盖式推广您的研究

投稿请点击：<http://www.hanspub.org/Submission.aspx>