

基于数学形态学的联机手写字符识别去噪方法

孙 嫣^{1,2} 刘瀚猛¹ 芮建武¹ 吴 健¹

(中国科学院软件研究所基础软件国家工程研究中心 北京 100190)¹

(中国科学院研究生院 北京 100049)²

摘 要 手写输入时由于笔尖抖动等原因产生了大量噪声,有效地去除噪声是手写识别的前提和关键。根据联机手写识别中手写体字符形态的特性,分析了手写时由于各种原因而产生的噪声,运用数学形态学中膨胀、腐蚀、细化等基本运算,提出了一种将数学形态学应用于联机手写识别预处理的方法,该方法可以有效地消除大量的冗余信息。测试结果表明,提出的方法可行,具有很好的鲁棒性,可以配合其他方案应用于各种联机手写字符识别中。

关键词 数学形态学,联机手写识别,噪声消除,细化

中图分类号 TP391 **文献标识码** A

De-noising Approach for Online Handwriting Character Recognition Based on Mathematical Morphology

SUN Yan^{1,2} LIU Han-meng¹ RUI Jian-wu¹ WU Jian¹

(National Engineering Research Center of Fundamental Software, Institute of Software, The Chinese Academy of Sciences, Beijing 100190, China)¹

(Graduate University, The Chinese Academy of Sciences, Beijing 100049, China)²

Abstract There are lots of various noises when users are writing characters on the tablet. It's significant to eliminate these noises in order to make these characters be recognized accurately. According to the features of online handwriting characters, we analyzed all sorts of noises generated during writing on the tablet. By applying erosion, dilation, thinning operations of the mathematical morphology into pre-processing of online handwriting recognition, a de-noising approach was proposed in this article. Using appropriate structure elements, we can eliminate large amounts of noises in the glyphs of characters. The experiment shows that this solution is valid and robust for online handwriting character recognition.

Keywords Mathematical morphology, Online handwriting recognition, De-noising, Thinning

1 问题的提出

字符识别是模式识别的一个重要分支,它涉及图像处理、人工智能、组合论等多个学科,是一门综合性技术^[1],在办公室自动化、机器翻译等高新技术领域,都有着重要的实用价值和理论意义。其中,联机手写字符识别在近年来得到了广泛的发展和大量的应用。

联机手写识别是完全以平常的习惯,把要输入的字写在书写板设备上,运用预处理、特征提取和分类器等核心方法将字符识别出来的过程。其中,预处理是保证特征提取和识别算法的有效性的前提。预处理的主要目的是将同一个字符的不同手写样本之间的各种变形和噪声进行弱化或者去除。通常,联机手写字符的预处理主要包括线性归一化、增加虚拟点、非线性归一化、笔画间的等距离重采样以及柔化等操作^[2]。

非线性归一化是基于手写样本的点或者笔画等信息来重

新计算手写样本的结构^[3]。它减小了同一字符的不同手写样本间在字形上的变异,削弱偏长的笔画,放大过小的局部结构,使联机手写字符在整个字形上更加饱满。对于头重脚轻、上窄下宽、左右不均的手写字形,非线性归一化有着良好的修正效果。

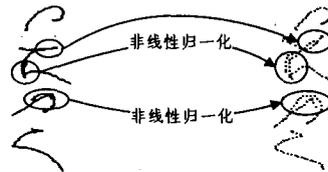


图1 非线性归一化的图像噪声放大效应

在分析手写样本时发现,由于在书写板书写时笔尖抖动产生的“噪声”,或者是书写者“涂改”所引起的多余或重复的笔画,导致了字符形态的失真。非线性归一化对形状很敏感,经过非线性归一化之后,会夸大在书写过程中由于抖动或者“涂改”带来的冗余信息或者噪声,从而增加识别的难度,图1

到稿日期:2008-11-11 返修日期:2009-03-02 本文受信息产业部电子信息产业发展基金项目“跨平台的藏文办公套件研发”资助。

孙 嫣(1983-),女,硕士研究生,主要研究方向为系统软件与中文信息处理, E-mail: sunyan06@iscas.ac.cn;刘瀚猛(1983-),男,主要研究方向为系统软件与中文信息处理;芮建武(1972-),男,博士,主要研究方向为操作系统与中文信息处理;吴 健(1962-),男,研究员,主要研究方向为系统软件与中文信息处理。

列举了两个不同的字符,经过非线性归一化后,放大了冗余信息。因此,如何在不改变字符基本形态的基础上,去除冗余信息,规范字符形态,是预处理过程中主要需要解决的问题,也是保证特征提取和识别算法的有效性的前提。

数学形态学是分析几何形状和结构的数学方法,是用集合论方法定量描述几何结构的方法,为大量的图像处理问题提供了一种一致的有力方法^[4]。数学形态学的主要内容是设计一套变换算法,用来描述图像的基本特征,其中的膨胀、腐蚀等运算,在消除噪声、抹去毛刺和连接断线中具有显著的作用。

在研究图像处理的过程中发现,数学形态学中的膨胀操作可以根据设置结构元素的大小“吸收”边缘像素,从而将邻近区域的重复笔画连接在一起;腐蚀操作可以将膨胀后的原本没有重复笔画和冗余信息的笔画还原,则不会破坏通过膨胀而连接在一起的区域,再通过细化操作在不破坏连通性的前提下,将连通区域即重复笔画进行骨架的提取,使其形成一个正常的、有效的单一笔画,从而可以将大量因手写抖动和“涂改”造成的冗余信息消除。

本文在分析和研究数学形态学基本操作的基础上,提出了一种解决手写识别中去噪声和重复笔画从而平滑和规范字符的方法。实验表明,该方法具有一定的鲁棒性和有效性,并适用于多种文字识别。

2 相关研究

去除噪声是手写识别中预处理的一个目标。在众多的预处理方案中,去除噪声的方法相对于归一化方法等步骤更加单一而且简单。最简单且常用的方法是计算出笔画中包含的点数,如果这个点数小于一个阈值,或者点数与整个字符中包含的点数之比小于一个阈值,说明该笔划可能是由于笔尖无意接触了手写板而造成的冗余信息,则将该笔画删除。在这个过程中,可以处理噪声点,但是由于计算方法缺乏鲁棒性和灵活性,导致该方法无法去除重复笔画,因为通常重复笔画中所包含的点数与有效笔画相当。

近年来,数学形态学在图像处理和模式识别中得到了广泛的应用。细化算法在指纹识别中被大量采用,用来提取指纹的骨架特征,展现图像形态,提高获取特征点的精度^[5-7]。在脱机手写识别中,同样使用过细化算法对字符图像进行了细化,使笔画宽度小于一个阈值^[8],或者使笔画宽度限制在一个像素内^[9],开操作以及闭操作可以平滑字符的边缘^[10]。可见,在与图像处理相关的脱机识别中,数学形态学已经发挥了作用。然而在联机手写识别中,由于识别方法具有实时性,不同于脱机中的图像处理,到目前为止,还鲜有将数学形态学应用于联机手写识别中。

本文的去噪方法运用了数学形态学中基本操作,该方法不同于脱机识别中单纯的图像处理方法,它保存了联机中笔迹的方向信息和点的先后顺序,有效地去除了重复笔画和冗余信息,达到了较好的效果。

3 去噪处理过程

通过分析,本文所讨论的基于数学形态学的手写字符识别预处理过程,可以由传统的预处理过程转变为如图2所示的过程,其中加入的膨胀、腐蚀、细化操作是去除噪声的关键

步骤。

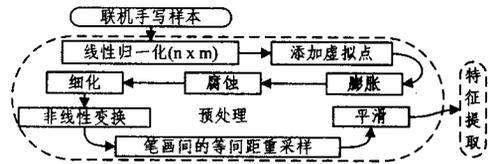


图2 预处理流程图

根据所处理图像类型的不同,数学形态学分为二值形态学和灰度形态学两种。本文是针对二值图像进行形态学处理的。

3.1 膨胀操作

二值形态学中运算对象是集合。设 B 为描述图像的二值点集合, S 为结构元素, 结构元素是任意大小、包含 0, 1 组合的一个区域。形态学操作就是用 S 对 B 进行操作。在结构元素上可任意指定一点作为原点。

膨胀是数学形态学的基本操作之一。当结构元素 S 的原点移动到 (x, y) 位置时, 记为 S_{xy} , 表示以 (x, y) 为原点的结构元素。如果 S 与图像 B 有任何一点同时为 1, 则新图像 D 上相应点为 1; 如果 S 与 B 完全没有相交, 则新图像 D 上的点为 0。即:

$$D = B \oplus S = \{x, y | S_{xy} \cap B \neq \emptyset\}$$

通过膨胀, 可以将两个距离比较近的物体连通在一起, 对于填补图像分割后物体中的空洞十分有效。使用膨胀操作, 可以使手写笔画中由于“涂改”而产生的大量无用和重复笔画连通在一起。这里的结构元素可以根据具体测试结果来调整形态和大小。结构元素一般选择 3×3 大小, 而每一个像素的值可以根据需要设定, 如图3所示, (a)、(b) 分别为两个不同的结构元素, 其中 1 值代表黑色像素, 0 值代表白色像素。

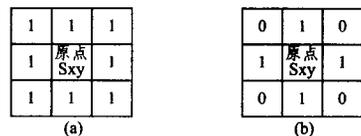


图3 结构元素举例

联机手写识别区别于脱机手写识别中最重要的因素之一, 便是联机手写识别中具有手写笔迹的方向信息以及笔迹点的先后顺序。所以, 在进行膨胀操作时, 不能像对普通二值图像一样, 从左至右从上到下地进行扫描, 而应该根据笔迹点的先后顺序扫描, 进行膨胀操作, 才能不丢失笔迹顺序等重要信息。

3.2 腐蚀操作

腐蚀是膨胀的对偶操作。当结构元素 S 的原点移动到 (x, y) 位置时, 如果 S 完全包含于图像 B 中, 则新图像 E 上相应点为 1, 否则为 0。即:

$$E = B \ominus S = \{x, y | S_{xy} \subseteq B\}$$

腐蚀可以去除小于结构元素的物体, 如果两个物体之间有细小连通, 当结构元素足够大时, 腐蚀可以将物体分开。不同的结构元素对结果影响很大。

在选择适当的结构元素对字符图像进行腐蚀时, 需要注意的是, 结构元素要根据膨胀操作时所使用的结构元素来选择。如果腐蚀的结构元素大于膨胀的结构元素, 则腐蚀操作就会将大部分的笔迹信息去除; 反之, 如果腐蚀的结构元素小于膨胀的结构元素, 则以前无重复以及噪声的正常笔画信息

便无法被还原,影响处理结果和识别效果。如图4所示,使用 3×3 的结构元素膨胀后,如果用 4×4 的结构元素腐蚀,则大部分笔迹信息将会去掉;如果使用 2×2 大小的结构元素,则无法还原原始图像中单一像素的笔画信息。因此,本文采用的是与膨胀操作中相同的结构元素。

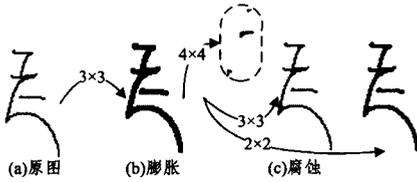


图4 使用不同的结构元素腐蚀后的图像

3.3 细化操作

经过了闭操作即先膨胀再腐蚀之后的图像,“涂改”区域呈现的是粗细不均匀的非单一像素的笔画,故需要对该部分进行细化,提取骨架,形成具有一个像素宽度的笔画,从而最终达到规范手写字符去除冗余信息的目的。

本文采用的细化算法是ZHANG和SUEN提出的并行细化算法^[11]。该方法的具体方案如下:

对于二值图像 B ,设 $P_1(i, j)$ 是图像中的任意一个像素,当 P_1 为黑色像素时值为1,为白色像素时值为0。则与其相邻的点为 $P_2(i-1, j), P_3(i-1, j+1), P_4(i, j+1), P_5(i+1, j+1), P_6(i+1, j), P_7(i+1, j-1), P_8(i, j-1), P_9(i-1, j-1)$,将这些点称为 P_1 的八邻域,如图5(a)所示。

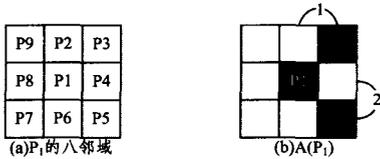


图5

该算法的核心思想是通过从4个方向不断移除部署图像骨架的像素达到细化目的。为了保持原图像的连续性,细化过程的每一次细分为2个子迭代过程,通过扫描图像的每一个黑色像素来执行。

在第一次迭代中,假设 P_1 是被检测的点,当 P_1 满足如下4个条件时,说明 P_1 不是骨架上的点,予以删除。

- a) $2 \leq B(P_1) \leq 6$
- b) $A(P_1) = 1$
- c) $P_2 \times P_4 \times P_6 = 0$
- d) $P_4 \times P_6 \times P_8 = 0$

其中, $A(P_1)$ 是指 P_1 的八邻域序列 P_2, P_3, \dots, P_9 中,由白色像素变为黑色像素的次数之和,如图5(b)所示。 $B(P_1)$ 是指 P_1 的八邻域中黑色像素的总和,即: $B(P_1) = \sum_{i=2}^9 P_i$ 。

在第二次迭代中,只有条件c)和d)变化了,其他条件不变,将满足条件的点删除:

- c') $P_2 \times P_4 \times P_6 = 0$
- d') $P_2 \times P_6 \times P_8 = 0$

通过不断地进行迭代,删除图像中标记的点,直到没有可以删除的点为止。

通过以上过程,可以发现条件a)要求 P_1 点的连接数在2和6之间,从而避免笔画端点被删除;条件b)要求 P_1 点的交叉数为1,从而使像素 P_1 的八邻域按黑白分别聚集,不会出现交错情形,保证了 P_1 的删除不会破坏笔画的连通性;条件c),d),c'),d')分别控制了删除的点为右下方、左上方、左下方和右上方的边缘像素,通过不断地从4个方向删除像素达到

细化的目的。

ZHANG和SUEN提出的并行细化算法存在一些缺陷,尤其在处理两像素宽度的斜线时会删除部分骨架信息^[12]。为了解决这个问题,在每次迭代之后删除像素之前,使用条件e)来判断是否删除像素 P_1 的八邻域中黑色像素个数不大于2的像素。如果大于2个则说明存在冗余像素,允许删除此像素,否则保留该像素。

- e) $B(P_1) \leq 2$

在经过了上述处理之后的字符图像中,保存了图像的骨架信息,但图像中仍然存在冗余的像素,如图6(a)所示,放大的图像中从一个端点到另一个端点的通路不止一条,此时会发生二义性。为了删除这些冗余像素,避免通路的二义性,这里按照手写笔迹点的先后顺序进行扫描:当扫描点 P_1 不满足条件e)时,即当前点的8个邻域内不止2个黑色像素,则计算当前点的八邻域中黑色像素点与八邻域外紧邻的点 P_{next} 的距离,保留与 P_{next} 距离最近的点,删除与 P_{next} 距离次近的点,不断迭代直到没有可以删除的点为止。此时点与点的连接情况称为 m 邻接^[4],如图6(b)所示。

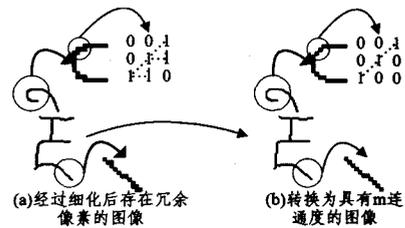


图6

4 实现结果

研究了当前的各种方法后,使用自己建立的藏文手写样本库对本文介绍的预处理方法进行了实验。在测试过程中,首先将每个字符图像归一化为 40×80 大小的图像,因为藏文本身字丁的宽高比例即为1:2,然后添加虚拟点,接着对藏文字丁进行膨胀和腐蚀操作,采用的结构元素大小为 3×3 ,然后细化字丁,非线性归一化,最后对字丁进行重采样和平滑。

经过形态学操作之后的字丁笔迹图像如图7所示,可以发现,在去除“涂改”等冗余信息时,膨胀、腐蚀和细化操作起到了至关重要的作用。

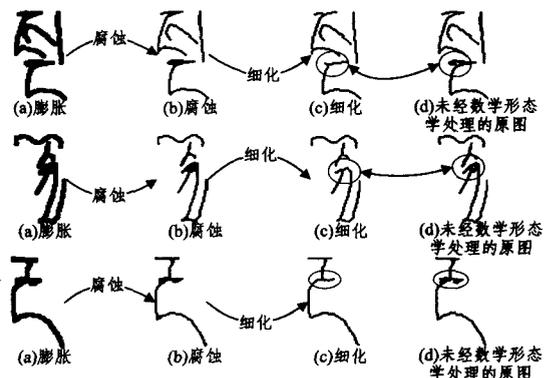


图7 数学形态学预处理对比图

全部预处理之后的字丁形态如图8所示。从图中可以看出,原始字丁中存在的冗余信息在经过了非线性归一化之后,

(下转第295页)

结束语 本文从避免图像放大时图像容易产生锯齿型边缘现象和马赛克现象入手,分析了传统的最近邻插值算法和曲面拟合算法存在的不足,提出了一种适用于图像缩放的“弹性模型”,给出了相应的算法。通过实验说明该算法所得结果图像的效果明显好于最近邻算法,比曲面拟合算法也要好一些。说明本文提出的“弹性模型”是合理的。另外,本文只给出了“弹性模型”在图像放大运算中的实验结果。如果按文中介绍的基于“弹性模型”的图像缩小算法做出图像缩小的实验结果,将更能说明本文提出的“弹性模型”的合理性。

参考文献

[1] 吴均,朱重光.带锐度保持的斜投影 B-样条图像缩放[J].遥感学报,2002,6(2):108-112
 [2] Zhao M, de H G. Content-adaptive video up scaling[C]//Proceedings of Application Specific Integrated Circuit. IEEE, 2003: 151-156

[3] Hentschel C. Generic method for 2D image resizing with simple, non-separable filters[C]// International Conference on Consumer Electronics. IEEE, 2003: 14-15, 17-19
 [4] 张俊华,陈建华,王道.基于边缘移动匹配法的图像插值[J].计算机工程与应用,2003,39(6):73-75
 [5] Nira S, Hagit A S, Ilan S, et al. Adaptive low complexity algorithm for image zooming at fractional scaling Ratio[C]// 21st Convention of the Electrical and Electronic Engineers. Israel, 2000: 253-256
 [6] 王效灵,陈涛,汪颖,等.边沿自适应图像缩放算法[J].浙江大学学报:工学版,2006,40(9):1507-1510
 [7] 肖建平,邹雪城,刘政林,等.基于 LCD 定标器的文本型图像缩放算法研究[J].华中科技大学学报:自然科学版,2005,33(5): 46-48
 [8] 何东健,耿楠,张义宽,等.数字图像处理[M].西安:西安电子科技大学出版社,2003:115-120

(上接第 239 页)

信息被放大了,导致了字形的过度杂乱,严重影响了识别效果;而通过对比可以得出,冗余信息去除后,非线性归一化得到了较好的形态,也达到了均匀笔迹密度的目的。

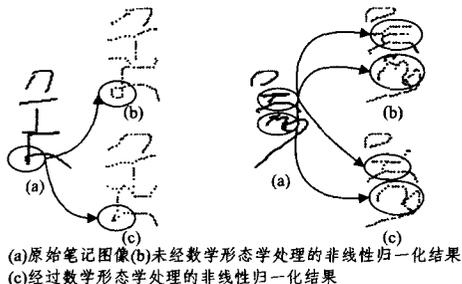


图 8 非线性归一化后对比图

因为该方法不是根据特定语言文字而进行的特殊处理,所以,该预处理方法不仅适用于藏文,也适用于其他语言的字符。通过调节腐蚀与膨胀的结构元素,可以达到不同的效果。因此,根据特定的字符形态,不断地进行试验与调节,与其它预处理方法配合使用,从而进一步规范字符形态,为特征提取打好基础。

在识别结果方面,不采用形态学处理时,封闭测试的识别率是 97.7226%,采用本文的方法后,识别率提高到 98.22%,说明该方法有效。由于是联机手写识别,实时性需要得到保证。在对 20259 个字进行特征提取时共用时 1291 秒,即每个字形的特征提取时间为 0.06 秒;对所有字进行模板匹配测试共用时 120 秒,则每个字形识别时间为 0.006 秒,因此对一个字形的识别总时间为 0.066 秒,满足实时性要求。

结束语 本文在研究数学形态学的基础上,提出了一种将数学形态学应用于手写体字符识别的预处理方向的方法,将图像处理与模式识别进行了有机结合。从实验可以看出,该方法通过对原始字符图像进行闭操作与细化,可以有效地降低和消除由于手写板笔尖“抖动”或者书写者本人“涂改”所造成的大量冗余信息,并且处理后保持了字符的原始形态,是一种可行的方法。通过配合其他方法和有效的识别器,可以达到较好的识别效果。

参考文献

[1] 边肇祺,张学工.模式识别(第二版)[M].北京:清华大学出版

社,2002
 [2] Bai Zhen-Long, Huo Qiang. A Study on the Use of 8-Directional Features for Online Handwritten Chinese Character Recognition [C]// Proceedings of the Eighth International Conference on Document Analysis and Recognition, 2005: 262-266
 [3] Zhen-Long B A I, Huo Q. A Study of Nonlinear Shape Normalization for Online Handwritten Chinese Character Recognition: Dot Density vs. Line Density Equalization [C]// Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06). 2006, 2: 921-924
 [4] 冈萨雷斯,等.数字图像处理(第二版)[M].阮秋琦,等译.北京:电子工业出版社,2006
 [5] He Yuliang, Tian Jie, Luo Xiping, et al. Image Enhancement and Minutia Matching in Fingerprint Verification[J]. Pattern Recognition Letters, 2003, 24(9-11): 1349-1360
 [6] Patil P M, Suralkar S R, Sheikh F B. Rotation Invariant Thinning Algorithm to Detect Ridge Bifurcations for Fingerprint Identification[C]// Proceedings of the 17th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'05). 2005
 [7] 叶茂,闵春平,李传光.基于 BP 神经网络图像的指纹细化算法 [C]// Proceedings of the 26th Chinese Control Conference. July 2007: 145-148
 [8] Wang Gwo-en, Wang Jhing-fa. A New Hierarchical Approach for Recognition of Unconstrained Handwritten Numerals[J]. IEEE Transactions on Consumer Electronics, 1994, 40(3): 428-436
 [9] Senior A W, Robinson A J. An Off-Line Cursive Handwriting Recognition System[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(3): 309-321
 [10] Jeong Cha-Sup, Jeong Dong-Seok. Hand-written Digit Recognition Using Fourier Descriptors and Contour Information[J]. IEEE Tencon, 1999: 1283-1286
 [11] Zhang T Y, Suen C Y. A Fast Parallel Algorithm for Thinning Digital Pattern[J]. Communication of the ACM, 1984, 27(6): 236-239
 [12] Wang P S P, Zhang Y Y. A Fast and Flexible Thinning Algorithm[J]. IEEE Transactions on Computers, 1989, 38(5): 741-745