

手写数字识别中的动态字符切分方法

A new method for the character segment of the recognition of handwritten numerals

吕知辛 (山东大学计算机科学与技术学院 250061)

摘要:概括了目前数字字符识别中常用的切分方法,并对于影响手写数字识别精确性的切分这一关键步骤,提出了一种新颖的解决思路,使得可以适用于各种不同的书写方式和习惯,解决了目前绝大多数识别系统不能解决的问题,极大地拓宽了手写数字字符识别的应用范围;且该方法同样适用于其他字符的切分识别中。

关键词:字符识别 切分

1 计算机字符识别的现状

几十年来,人们在字符识别的处理方面投入了大量的精力和时间,取得了一些辉煌的成果,如,印刷体的识别和联机手写字符的识别方面,其识别技术已经趋于成熟,基本可以满足实用的需求,并且已经有许多实际的应用系统,像 TH-OCR、尚书 OCR 等是目前国内比较有代表性的识别产品。但是,作为手写数字,由于主要由阿拉伯数字及少量的符号组成,是一些统计报表的基本构成,而且这类报表在实际使用过程中是大量存在的,如果通过手写数字的识别技术可以解决这类数字信息的自动录入,将产生极大地社会和经济效益。

手写数字识别作为模式识别的一个重要应用领域,其应用前景非常广泛,研究的理论价值也是非常重要的,主要体现在:

(1) 数字不象其他的字符一样,它是世界各国的通用符号;

(2) 数字的类别较少,有助于进行深入的分析 and 验证一些新的识别理论;

(3) 手写数字的笔画较少,但是其变化却是非常多样的,与多笔画汉字的识别比较,其特征的把握更加困难;

(4) 对手写数字识别问题的研究,可以很方便的推广到其他的如英文字母、常用符号的识别上去。

2 手写数字识别中常用的切分方法及存在的缺陷

仅就数字字符本身而言,只有 0~9 十个种类,且每个数字的笔画数目又非常少,表面上看起来数字识别的问题应该比其他类型的字符简单得多,但是,笔画少在识别的过程中恰恰成为一个难点:由于数字书写笔画过少,书写的变化太多,我们对其识别时特征的选取十分困难,使得对其识别比其他字符如汉字的识别要艰难。还有,数字的前后单字符之间没有相互的对应关系,不象其他字符那样,可以从前后之间组成的词义上加以判断,这又使得数字的识别,尤其是单个数字字符的识别更加困难。还有就是,数字一般牵涉到财政、金融等重要领域,对识别的准确性要求特别的高,而且,在这些部门,数据的处理都是海量的,对识别的处理速度要求非常高,这是其他字符识别所难以达到的。

综上所述,手写数字识别是一项比其他字符的识别要求更为严格的识别问题,其技术的实现也远比其他字符的识别复杂得多。

我们知道,字符的切分好坏直接影响到识别的效果,因而,切分在识别中起着重要的作用。目前一些数字识别系统对数字书写的要求可以归纳如下:

(1) 将手写数字填写到预先印制好的带有方框的表格中,且表格中填写数字字符的方框中规定以描 8 的方式填写。并且,为了实现切分的方便,在表格的横向和纵向均加有定位的定位符,如图 1。

(2) 将手写数字填写到预先印制好的带有方框的表格中,但表格中数字字符的填写是自由书写,没有描8的限制,并且定位也不须在横向和纵向均有,而是只在少数的几个边角印制上表示方格高度和宽度的定位符即可,见图 2。

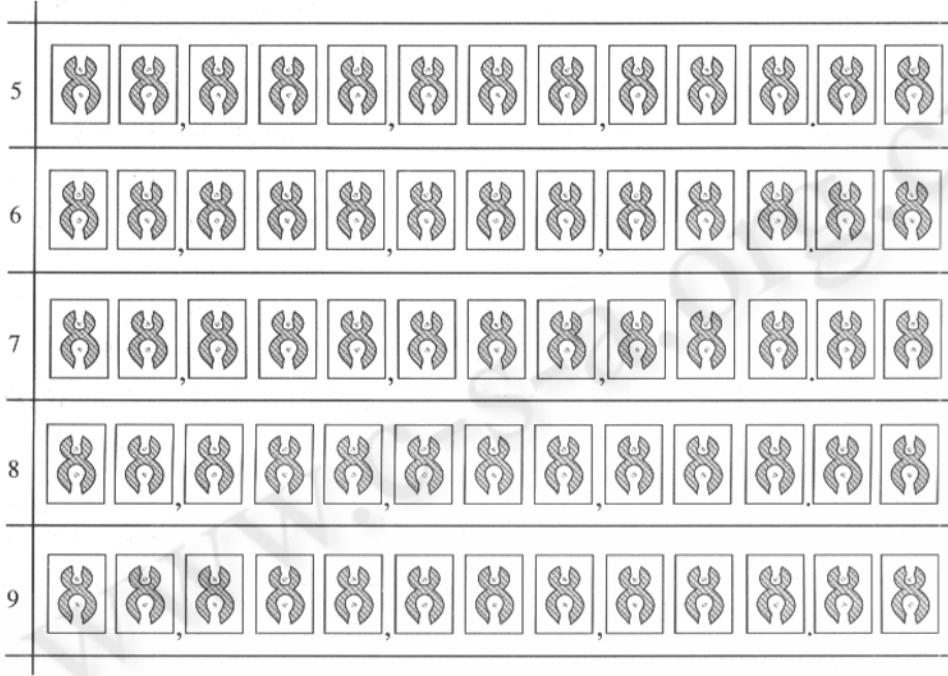


图 1

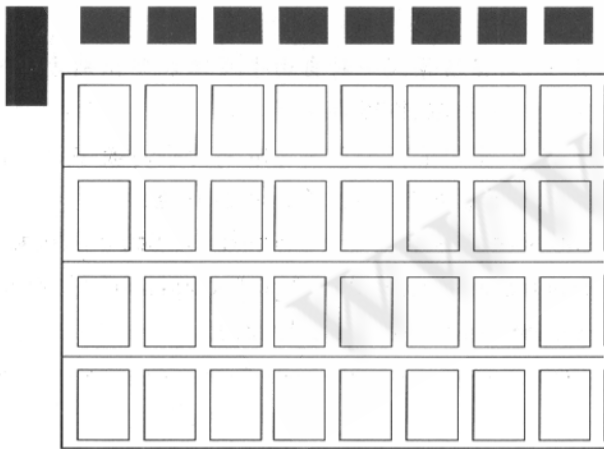


图 2

(3) 无需将数字字符填写到方格中,只要按行列填写即可。

对于第一种情况,由于限制了书写的格式和笔画,

所以,无论切分还是识别都是很好实现的;第二种情况,限制了书写的位置,但书写的方式是在一定范围内的自由方式,所以,其切分实现较简单,但数字的识别较复杂;第三种情况,几乎没有任何的书写限制,数字字符的识别与第二种类似,但是最主要的

关键问题是这种自由书写格式中数字字符的切分。

目前,切分一般首先进行行方向的切分,即先进行行划分,对划分好的一个个的行,再进行列方向的切分,将其切分成一个个的单个数字字符。即首先将整篇字符图像做Y方向的投影,分出行;在将分出的行做X方向的投影,分割出单个的字符^{[2][3]}。

这种切分方式的最大缺陷是,对于书写不是很规范的,尤其是在行的方向上不在一个水平线上的情况(如图 3),将无法进行切分。但是,目前的切分大都没有做到与数字字符的书写相结合。

3 动态字符切分方法

3.1 方法的提出

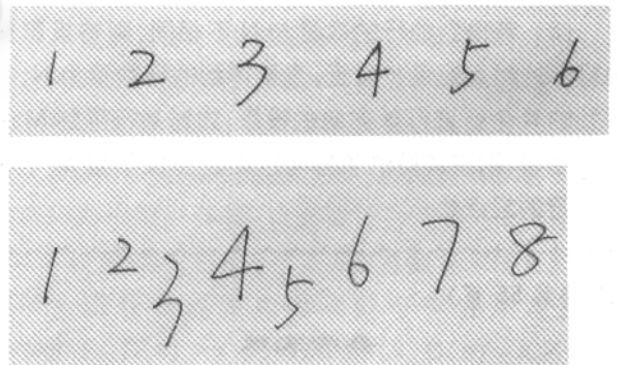


图 3

对于手写数字字符而言,无论书写的习惯怎样,通常情况下,在不会出现断笔的现象的前提下,我们可以

以此为切分的依据,而不是仅仅从宽度上进行切分;另外,对于有些书写不是很规范的情况,或者将扫描图像经过处理后,如去噪、细化等,出现的异常断笔笔画处,首先按照插值多项式的方法进行断点的续连,使之更加符合原来的书写模式,并且得到的字符的图像与相应的数字字符更加符合。

3.2 方法的实现

我们先来看一下人的阅读习惯。大多数人在阅读时是一个字符、一个字符进行阅读的,即阅读完一个字符时,接着阅读与之相邻的下一个字符。当一个字符与上一个字符的位置从水平方向和垂直方向均很接近时,一般将其视为是上一个字符的下紧邻。也就是说,我们在进行字符的切分时,不是按照先行后列的切分步骤,而是首先确定每一个书写行的首字符的位置,以及它的高度、宽度,然后,将水平方向的步长定为首字符的宽度的 2 至 3 倍,切分下一个字符,再将下一个水平方向的步长的起点选定在下一个即将进行切分的字符的起点处,重复上述步骤,直至本行结束。这样的切分,可以从根本上解决原来的方法中,不能对于书写倾斜的字符进行识别的问题,从而,使得识别的适应范围更加拓宽。

也可以这样进行变动:

(1) 首先按每一行的首字符进行 Y 方向的投影,计算出行的个数 N;

(2) 再进行全图像的 Y 方向投影,若得到的行的数量与 N 相同,则接着进行 X 方向投影,完成对单个数字字符的切分;

(3) 若得到的行的数量与 N 不相同,则将投影面向投影的起点处逐步拉近,直至得到的行数量为 N,对得到的 N 个行进行 X 方向的投影,如对某行得到 M 个切分字符,则将该行的 Y 方向投影起点位于第 M 个切分字符的起点处;

(4) 将投影的起点位置选定,重复步骤 3 的过程,直至所有行的切分全部完成。

3.3 进一步的设想

在动态切分完成后,还应对切分得到的每一个方框内的数字字符进行处理,一般的常规处理在此不多叙述,但是,有些人的书写习惯上喜欢将尾笔拉长,这样带来的直接后果就是影响到位于下面一行的同一位置上的数字字符的识别。我们可以在切分后,将不必要的赘笔去除,令其不要超出切分的单元格,将超出部分进行自动擦除,这样,将使得字符之间的相互影响降到最低,更进一步的提高了识别的准确性,使误识率降低。

利用上面介绍的切分方法,我们可以将一些书写不规范的,尤其是一些行倾斜的书写,进行数字字符的切分,从而保证识别的正常进行。而且,以上介绍的切分方法不仅适用于数字字符的切分,也可以适用于其它字符的切分过程,如英文、汉字等的切分过程。

对于连续的手写数字的切分与识别,在上面介绍的各种切分方法的基础上,结合笔划跟踪[1]的方法,进一步确定切分的准确性。

参考文献

- 1 黄为民、荣钢、边肇祺,《由手写体图像恢复笔划的一种方法》,清华大学学报(自然科学版),1995 年第 5 期。
- 2 张平、潘保昌等,《一种自由手写体数字识别方法研究》,光电工程,1995 年第 3 期。
- 3 毛艳华、张新荣,《一个高速准确的手写数字识别系统》,计算机应用研究,2001 年第 11 期。
- 4 Laurence C Agba, Rari Shankar. A Septon Feature Scheme in Handwritten Digit Recognition [J]. Proc.