

压缩投影法手写数字识别系统的研究

尤韦彦 张功镛 虞荣法 敖国利

摘 要

本文提出一种对手写数字采用压缩投影的模式识别方案,使信号处理与特征抽取方便,识别简单,迅速,设备容易研制。整个系统包括摄像机,信号处理控制器,专用接口,微型计算机,抽纸机等设备。本系统通过对二十五万手写数字识别的实用化试验,结果良好。

一、引 言

高校招生考试试卷成绩的计算与统计是一项工作量很大,很繁琐的脑力劳动。近年来使用计算机对考生的成绩进行统计排队在很大程度上减轻了这项劳动。但要使计算机得出满意的统计结果,还需人工将卷面上的成绩分数一个个的通过穿分数纸带正确无误地输入计算机。这仍是一项工作量很大,很繁琐的脑力劳动。如何实现该项工作的自动化?如何使计算机模拟人的识字能力,直接利用写在纸上的数字信息?这就是手写数字模式识别的研究课题。为此,我们设计了一种压缩投影的模式识别方案,并研制了整个识别系统。本系统在一九八二年八月初对二十五万字的卡片正式进行了试验,识别率达99.87%,拒识率是0.12%、误识率是0.01%。

二、书写限制与压缩投影

一个手写数字可看作是定位在 $M \times N$ 点阵内的一幅图象,图象信号经过二值化处理,使有笔迹的部分的值为1,无笔迹的部分的值为0。把这样一幅 $M \times N$ 的二值矩阵信息直接输入计算机,由计算机识别其所属类别。识别的方法很多,不同的方法具有不同的识别速度与识别率,并造成设备的不同复杂程度。对手写数字作出一定的书写限制和规定,往往可以收到事半功倍的效果。美,日等国OCR(Optical Character Reader)有多种规定的字体,有的书写限制极严。



图 1

本文于1982年9月21日收到

范之刚、陆伟宣、楼旦鸣、石开建同志曾参加过本项目中的部分工作。抽纸机由师院校办工厂制作。

本方案采用带有一定书写限制的常用手写体数字。标准书写字样如图 1 所示，规定书写需在框格内环绕二个导引点进行。

从许多人实际试验的大量书写样品来看，不需专门的训练都可以达到本书写规定的要求。实际书写字样抽样如图 2。从原理上说二个导引点是很小的点，但为了保证有较高的识别率，避免书写的数字在位置上有过大的偏移和旋转，将导引点适当扩大成两个小方块，既使书写方便又保证不让字体偏离标准太多。框格和导引点用与手写数字不同的颜色印刷，在光电转换时用滤色镜把框格与导引点滤去。



图 2

通过导引点可以把框格内 $M \times N$ 个点分割成左、右、上、中、下五个相互交迭的区域，见图 3。所谓压缩投影是把落在这五个相互交迭区域内的笔划信息分别按指定的方向合并相加。以数字“8”为例，通过压缩投影以后就得到图 4 所示的数字“8”的输入模式。

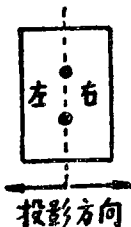
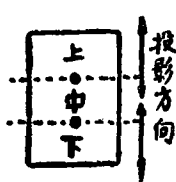


图 3

图 4

从投影过程可以看出此法的优点是使信号处理和特征抽取方便，计算机需处理的信息大大压缩，使识别过程简单，迅速。

对信号处理来说，首先，此法不须对笔划作常规的细化处理，因为笔划的粗细不直接影响数字模式的空间拓扑结构，不管一个数字写得粗或者细，压缩投影后可得到相同的输入模式。

其次，此法具有一定的抗干扰能力，对于孤立的黑点不必进行专门的去除处理，压缩投



图 5

影可以消除这种噪声的影响，使其被笔划信息所复盖，如图 5 所示。对于不能被投影所掩盖的孤立黑点，在抽取特征时也可以消除。

压缩投影实质是一种从观察空间到模式空间的变换，它使输入手写数字的种种书写变形规格化，便于反映输入数字的本质特征。图 6 示出四种“8”的书写形状，通过投影以后，都得到共同的模式特征。

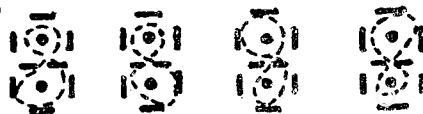


图 6

三、特征抽取与模式分类

本方案用八个特征度量，从输入的数字模式的八个投影线段上抽取（如图 7 所示）。当投影线段上有连续“1”出现，此特征度量就为“1”，否则就为“0”，通过灵活的特征编码，可得到输入待识数字模式的八位特征码。用

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_8 \end{pmatrix}$$

表示在八维特征空间里的特征矢量，此空间共有 $2^8 = 256$ 个离散的特征矢量。通过对大量手写数字的分析试验，考虑到书写中可能出现的断笔与变形，以及输入信息中可能出现的丢失与噪声，可把 256 个特征矢量分为三大类：一是可识类，二是待判类，三是拒识类，其中可识类又可区分成互不相混的十类，即十个数字类别。

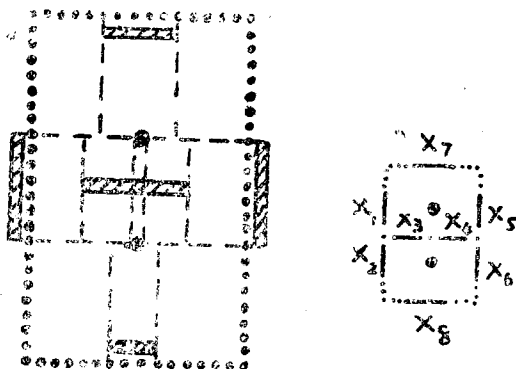


图 7

识别过程采用简单、迅速的查表法。所谓“表”，就是人工设计编纂的一本识别字典。字典中的每一个地址对应特征空间的一个矢量，因此字典总共有 256 个地址。在 Z-80 微机中，字典地址的低八位就是从输入模式中抽取的八位特征码，而每一地址的内容就是这八位特征码所代表的类别。

整个识别过程由两次判定构成，识别流程框图如图 8 所示。

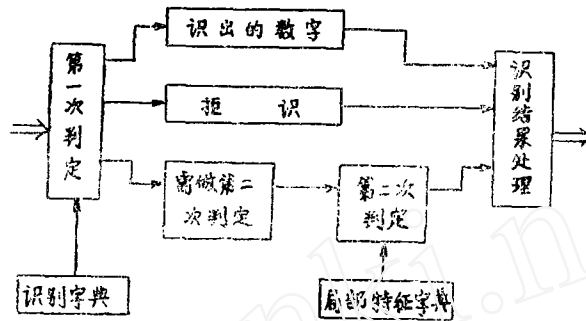


图 8

书写对象如能加以配合，按照书写规定和标准字样，笔划匀称地、清楚地书写，那绝大多数手写数字输入模式的特征码都在可识类里面，通过第一次判定就能得出肯定的识别结果。仅对于少量手写数字，其特征码在待识类里，通过第一次判定还不能完全确定其数字类别，需对其输入的原始的信息抽取最具有代表性的局部特征，并按这些局部特征进行第二次判定，得出识别结果。

这里需要指出的是，与特征空间可识类特征矢量相对应的输入模式（在模式空间里的矢量）是非常多的，与之相对应的手写数字原型（在观察空间里的矢量）更是极为丰富的。

随着大量实验的进行，识别字典的内容会不断丰富和更新，利用微计算机自动编纂字典的程序，只要求输入所设计的八位特征码及与其相对应的数字，程序就会自动修改特征码相对应的地址与内容，很方便地得到已更新过的新字典。

四、光电转换与图象二值化

本系统采用工业电视摄像机作为光电转换部件，原因是其信噪比，分辨率和惰性等技术指标基本上可满足要求，而一个突出优点是设备简单，成本低廉。采用电视扫描制式获得高速的码率可以与计算机的快速处理相配合。

按照高考成绩登记的要求，登分单设计成如图 9 所示的格式。实际上，只要书写框格位置不变，表格内容与格式可以变化。按每个书写数字占据 22×30 的点阵，可以把登分单的有效象面分割成 160×260 点阵。按电视扫描标准，实际只要每行输出 160 个点，每场输出 260 行的信息，每秒输出 50 场。计算机只要接收其中一场信息便可进行识别，因此可获得较高的识字速度。但考虑到机械换纸速度，摄像管的惰性，以及处理数据及判决识别需要时间，决定每秒识别一张。系统有效识字速度为 100 字/秒。由于在字与字，行与行之间都留有相当空白，实际识字速度为 42 字/秒。

要求光电转换设备输出一个信噪比高，幅度稳定，本底平坦的视频信号，因此对摄像机输出信号要进行黑切割放大，自动增益控制和黑斑补偿。

图象二值化是由一个具有自动阈值控制的高速 A/D 转换器来完成的。图象二值化的质量直接影响计算机的输入数据信息，与识别率关系极大。作为一个实用的文字识别系统，图象二值化的阈值自动调整是必定不可少的。因为纸张洁白度不同对光线的不同反射会影响背

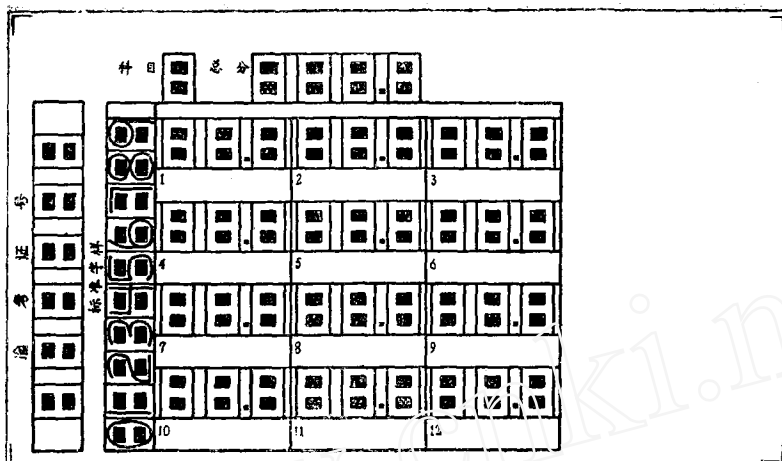


图 9

景白电平的起伏，而书写墨水的深淡与笔划的粗细会造成黑色信号的幅度差别，再加上背景白噪声的影响，使得很难有一个固定的量化阈值。为了应付这种复杂的情况，对于不同的纸张，不同粗细浓淡的笔划均能获得一幅符合要求的二值图象，本系统采用跟踪背景白电平和笔迹黑电平包络的方法来获得一个自动调整的阈值电平。如图10所示。

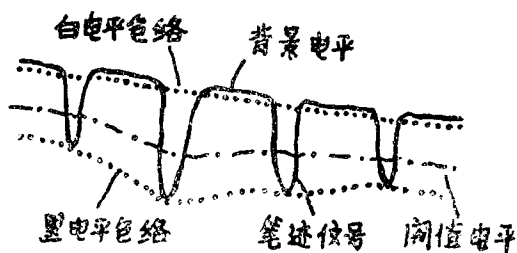


图 10

二值化的数据信息由经过定位的时钟脉冲来控制。以确保信息在点阵中的正确位置，使得计算机存储器中得到正确的输入模式。由于登分纸张上无特定的套色定位标志，只能直接利用纸张的边缘来定位，这就要求登分单有良好的印刷质量，以保证定位的正确。光电转换与信号处理控制部分的原理方框如图 11 所示。

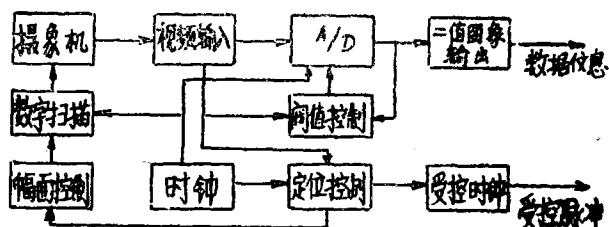


图 11

五、数字扫描与幅面控制

本识别方案是以框格和导引点位置为依据的，所以每个手写数字信息的正确定位是确保识别率的关键。当控制时钟根据纸张边沿起始定位以后，扫描器的线性是否良好，幅度是否正确就决定了所有框格与导引点的位置。因此，一个线性良好的，与控制时钟同步的，幅度可控的数字扫描器就必不可少。其原理框图如图 12 所示。

利用恒流源在时钟的每个节拍向摄像管偏转线圈内相应增注一股恒定电流，扫描电子束在每个节拍偏移一个相等的间隔，从有效幅面开始到有效幅面结束，行里正好偏移 160 点，场里正好偏移 260 行。要使这个关系稳定，扫描幅度必须进行反馈控制。若扫描幅度变大，图象有效幅面就变小，由时钟计数所得的标准参考幅面就大于图象有效幅面，用二者的幅面误差反馈控制扫描恒流源，使扫描幅度减小，重新达到稳定，反之，则使扫描幅度增大。行扫描，场扫描都采用这类数字扫描。另外，由于摄像机偏转线圈的线性精度不够，使得整个扫描线性不够理想，通过数字扫描可以分段调整，校正其非线性，更显示了数字扫描的优点。

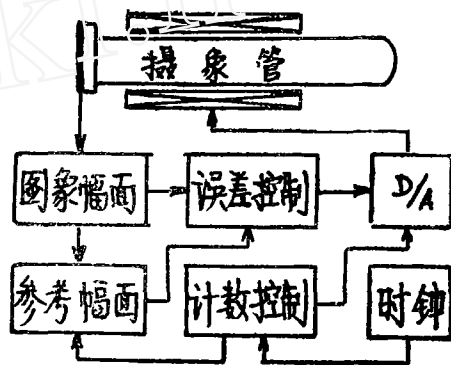


图 12

六、专用接口

通过二值化后的数据是每场为 160×260 的串行信息，在送入 Z-80 微机的通用接口之前还需要一个专用接口。专用接口的功能是：

1. 逐行串并变换。由受控时钟将串行数据信息存入一串行移位缓冲寄存器中，在一行时钟结束以后，将一行数据信息并行送入一批批输送寄存器。
2. 部分信息压缩。在串并变换的过程中，去除空白处的无用信息，减少送入计算机的信息量，同时将书写出格的边缘部分的信息压缩到框格内。
3. 逐个字节输出。在计算机的控制下，将有用信息按八位一字节逐个输出送入计算机，它属于有条件的程序传输方式。

专用接口框图如图 13 所示。

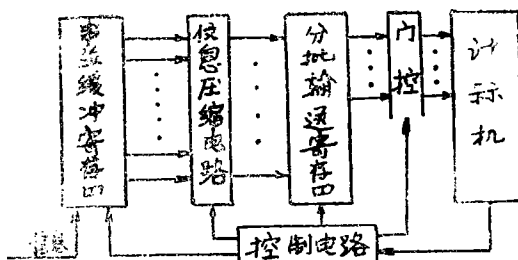


图 13

七、系统组成

本识别系统组成方框如图 14 所示。

登记考生成绩的登分单若干张（例如 1000 份）送入抽纸机。电视摄像机摄取第一张登分单的图象，通过信号处理控制器变换成数据信息。当计算机接收到由抽纸机发出的取数请求信号以后，立刻向信号处理控制器发出取数命令，信号处理控制器即向专用接口输出全部数据信息及受控时钟。计算机对接收到的信息进行识别。当识别结束之后，就发出一个识别完成脉冲，控制抽纸机把最上面的一张登分单抽走，摄像机就摄取下一张登分单的图象。如果被抽走的登分单上的手写数字全部可识，该登分单就被送到可识箱里。只要其中有一个手写数字被拒识，计算机在发出识别完成脉冲的同时，还发出一个拒识脉冲，让这张登分单被送往拒识箱里。当一张登分单被送入指定箱子时，抽纸机发出对第二张登分单的取数请求信号，这时新的一张登分单的图象早已稳定地转换成数据信息，等待计算机的接收。系统的工作进程就是如此循环。当全部登分单都识完之后，所有被识别的登分单上的成绩已全部存入计算机内存。而被拒识的登分单上的成绩，可以再次送入抽纸机进行识别，最后确实不能识出的少量登分单，可由人工通过键盘将数字输入。计算机将内存中的全部考生成绩存入软盘，等待最终的考生分数排队处理。

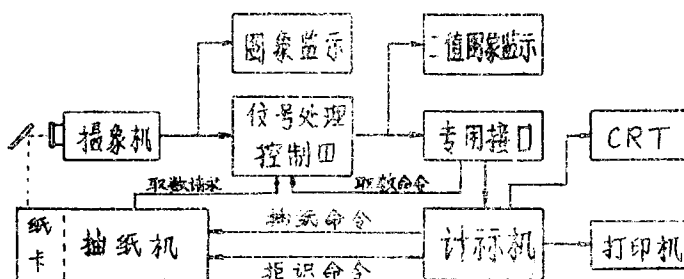


图 14

结束语：本系统采用压缩投影法识别手写数字，具有简单、快速识别的特点，系统设备研制容易，成本低廉。在书写对象给予配合的情况下，可以获得很高的识别率。对于高考成绩统计的“迅速，准确，保密”要求及自动化作业具有较大意义。

本系统可以用于其它需要识别手写数字的项目，由于采用电视扫描制式，本系统的一些主要设备也可用于小型图象处理及其它模式识别。

附：微计算机打印的登分单手写数字的二值图象（已缩小）及最后识别结果。（图15、图16）

参 考 文 献

- [1] 森健一 理論はどのように体系付けられているか日径エレクトロニクス No. 100 1975 特集パターン認識の現状と将来
- [2] 安田道夫 文字読み取り装置はどのように構成されているか日径エレクトロニクス No. 100 1975 (同上)
- [3] 活字並みにドット文字が読めるOCRを開発 日径エレクトロニクス No. 231 1980.2.4
- [4] 手書きOCR装置 National Technical Report vol 26 No. 1 1980.2
- [5] New Page Optical Character Reader OCR-V TOSHIBA REVIEW No. 91 1974
- [6] 倪光南等 一台手写文字识别机 计算机学报 1980.5 第二期
- [7] 复旦大学文字识别研究组 光学文字识别系统
- [8] "A flexible method of automatic reading of handwritten numerals" Philips Tech. Rev. vol. 33 pp. 89-101 and 130-137, 1973
- [9] Digit Pattern Recognition. Edited by K. S. Fu

The Research on the Recognition System of Handwritten Numerals Using Packing Projection Method

You Weiyang Zhang Gongdu Yu Rongfa Ao Guoli

Abstract

In this paper a pattern recognition scheme for the recognition of handwritten numeral using packing projection method is introduced. In this scheme the signal processing and the feature extracting are very convenient. The recognition is simple and fast, and it is quite easy to research and manufacture its devices as well. This system consists of a set of television camera, a signal processing controller a paper-extractor, and other devices. The result of the system is satisfactory after having been tested practically on 250 thousand handwritten numerals.