

# 支票容错识别系统预处理的设计与实现\*

赵斌 马康坚 夏绍玮

(清华大学自动化系 北京 100084)

**摘要** 文章在支票容错识别系统的基础上,介绍了预处理算法的设计与实现,包括软件定位、各单元子图的提取与背景去除、倾斜校正处理、方格去除和字符串分割等。其中的重点是软件定位和倾斜校正两个预处理算法。

**关键词** 预处理,模式识别,表格处理,文档分析。

**中图法分类号** TP391

支票的自动识别与处理是银行自动化的一个重要研究课题,它不仅技术含量高,而且具有重大的实用意义与商业价值。国内外不少研究者已经建立了一些支票处理系统,对支票的识别已获得一定的成果<sup>[1]</sup>。

图1是带有动态密码的银行支票样张,通常有6个部分需要进行处理。

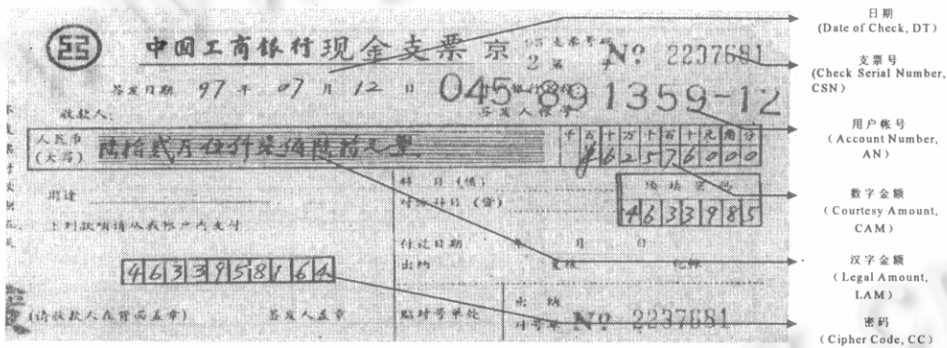


图1 带有动态密码的银行支票样张

针对带有动态密码的银行支票,我们设计了能自动处理银行支票的支票容错识别系统(fault-tolerant bankcheck recognition system,简称FBRs)。系统由预处理器、字符识别器和后处理器组成。其中预处理器处于系统最前端,它对字符识别器的影响很大,其重要性是显而易见的。

我们已经实现可以处理支票日期、数字金额和密码3个部分的FBRs原型系统<sup>[2]</sup>。本文将基于原型系统研究预处理算法的设计与实现,首先对FBRs的预处理器进行总体介绍,然后重点叙述预处理算法中软件定位和图像倾斜校正两个部分,最后是一个预处理实例与结论。

## 1 FBRs的预处理器

FBRs的预处理器由软件定位、各单元子图提取及背景去除、倾斜校正、方格去除、灰度子图二值化、子图字符串分割等模块组成<sup>[2]</sup>,如图2所示。

\* 本文研究得到国家自然科学基金资助。作者赵斌,1973年生,博士生,主要研究领域为模式识别,人工神经网络,图像处理,系统工程。马康坚,1974年生,博士生,主要研究领域为模式识别,人工神经网络,系统工程。夏绍玮,女,1932年生,教授,博士生导师,主要研究领域为神经网络,模式识别,决策支持系统,系统科学,系统工程。

本文通讯联系人:赵斌,北京100084,清华大学自动化系

本文1998-04-09收到原稿,1998-06-05收到修改稿

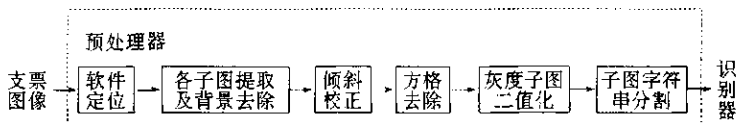


图2 FBRs的预处理器构成

### 1.1 软件定位

尽管系统在硬件基础上进行了特定处理,使录入的整张支票图像具有大体一致的扫描区域和位置,但很多后续算法对各部分图像位置很敏感,在进行其他处理前,必须对图像进行精确到像素级的定位处理。

### 1.2 各单元子图提取及背景去除

根据支票上每个待处理部分在整张支票图像中的区域,将各部分图像提取出来,去除背景噪声和干扰,形成各自区域的 256 级灰度图像,称为子图。设所提取子图的矩形点阵大小为  $W \times H$ ,  $W$  为点阵宽度(列数),  $H$  为点阵高度(行数)。设每个像素点元素为  $D(i, j)$ , 其中  $i$  为像素点的水平坐标,  $j$  为垂直坐标, 满足

$$0 \leq D(i, j) \leq 255, \quad i=0, 1, \dots, W-1; \quad j=0, 1, \dots, H-1. \quad (1)$$

### 1.3 各单元子图的倾斜校正

扫描录入的图像难免有一定的倾斜度。系统对于特别小的倾斜度,大约在  $\pm 2^\circ$  以内,可以自动适应,无需校正处理。一般而言,对于倾斜度在  $\pm 10^\circ$  范围内,就要系统进行倾斜校正处理了。对更大的倾斜度,由于倾斜校正运算量大,会产生较大的误差和失真,建议将支票图像重新扫描录入。

### 1.4 数字金额子图和密码子图的方格去除

支票上的数字金额和密码都书写在一定大小和数目的方格串里。FBRs 要把数字金额和密码子图的限制方格去掉。经过倾斜校正的图像,直接利用它的垂直投影和水平投影轮廓线就可以得到方格线的位置和粗细。在提取的单元子图上,按照方格的位置和大小,把方格线去掉,便得到删除了方格的支票金额和密码图像。

### 1.5 灰度子图二值化

二值化处理是把灰度图像信号变换成二值(0,1)的数字信号,点阵像素经过二值化处理后为  $B(i, j)$ :

$$B(i, j) = \begin{cases} 1, & D(i, j) \text{ 为前景像素;} \\ 0, & D(i, j) \text{ 为背景像素.} \end{cases} \quad (2)$$

其中  $i=0, 1, \dots, W-1; j=0, 1, \dots, H-1$ .  $W$  为点阵图像宽,  $H$  为点阵图像高。

二值化方法通常有整体阈值法和自适应的动态阈值法<sup>[5]</sup>。实际处理的图像比较复杂,如从银行支票上提取手写字符时,需要抑制污点和噪声背景。为了能够更好地适应书写质量差或背景情况复杂的图像,我们使用了动态阈值法。

### 1.6 子图字符串分割

子图字符串分割是把各个单元子图中的字符串分成单字符图像的处理过程。支票的金额子图和密码子图经过方格去除处理后,和日期的子图一样,可以运用字符数目未知的无约束手写体数字的分割算法来进行字符串分割。字符串分割的算法比较复杂,这里不详细介绍,可参见文献[4]。

## 2 软件定位

### 2.1 基本概念

在整张支票图像中(如图 1 所示),支票数字金额和汉字金额被几乎横贯整张支票的矩形方格框起来,该矩形称为定位矩形,这是整张支票图像中最为突出的标志之一。定位矩形的左上角顶点称为左上定位点,定位矩形的右上角顶点称为右上定位点。

分别找到定位矩形的左上定位点和右上定位点,再根据对该类型支票上定位矩形和其他各个待处理部分相互的几何位置关系,把其他部分的区域推算出来,就可完成整张支票图像的定位。为此,我们分别构造了如图 3 所示大小为  $(2 \times K + 1) \times (2 \times K + 1)$  的定位模板  $T_1$  和定位模板  $T_2$ 。

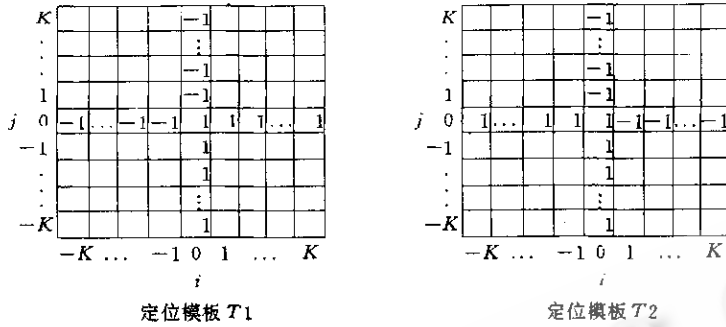


图3 定位模板定义

设模板中的元素为  $T(i, j)$ ,  $i = -K, \dots, -1, 0, 1, \dots, K$ ;  $j = -K, \dots, -1, 0, 1, \dots, K$ , 其中  $i$  为水平坐标,  $j$  为垂直坐标. 定位模板 T1 中各个元素取值定义如下:

$$T1(i, j) = \begin{cases} 1, & j=0 \text{ 且 } 0 < i \leq K; \text{ 或 } i=0 \text{ 且 } -K \leq j < 0; \text{ 或 } i=j=0; \\ -1, & j=0 \text{ 且 } -K \leq i < 0; \text{ 或 } i=0 \text{ 且 } 0 < j \leq K; \\ 0, & \text{其他.} \end{cases} \quad (3)$$

类似定位模板 1, 对于定位模板 T2 的各元素, 有

$$T2(i, j) = \begin{cases} 1, & j=0 \text{ 且 } -K \leq i < 0; \text{ 或 } i=0 \text{ 且 } -K \leq j < 0; \text{ 或 } i=j=0; \\ -1, & j=0 \text{ 且 } 0 < i \leq K; \text{ 或 } i=0 \text{ 且 } 0 < j \leq K; \\ 0, & \text{其他.} \end{cases} \quad (4)$$

定义在模板中相应于  $i=j=0$  的元素称为该模板的模板中心点.

### 2.2 左上定位点的基本算法

在整张支票图像上初始化一个包含左上定位点的搜索区域  $D$ , 该区域大小为  $W \times H$ ,  $W$  和  $H$  分别为搜索区域  $D$  的宽度和高度. 如果从左上定位点出发, 分别沿水平和垂直方向各作两条方向相反的射线, 形成水平和垂直交叉的“+”字形状的 4 条射线. 在这 4 条射线上, 被左方和上方这两条射线所覆盖的图像点绝大部分由背景点构成, 而右方和下方这两条射线上所覆盖的图像点绝大部分由矩形方框前景点构成. 据此构造并定义量化指标  $q(x, y)$ :

$$q(x, y) = \sum_{i=-K}^K \sum_{j=-K}^K D(x+i, y+j) * T1(i, j). \quad (5)$$

对在所初始化的搜索区域内的每个像素点  $P(x, y)$  进行遍历, 先让模板中心点与该点  $P$  重合, 然后在搜索范围内应用 T1 按式(5)进行卷积运算, 得到点  $P(x, y)$  的量化指标  $q(x, y)$ . 搜索区域内具有最大量化指标  $q^*$  的像素点  $P^*(x^*, y^*)$ , 就是我们所要寻找的左上定位点. 设点  $P^*$  的横坐标为  $x^*$ , 纵坐标为  $y^*$ , 量化指标为整个搜索区域中的最大量化指标  $q^*$ , 则遍历搜索算法可以表示为

$$q^*(x^*, y^*) = \max_{\substack{0 \leq x \leq W \\ 0 \leq y \leq H}} q(i, j). \quad (6)$$

### 2.3 改进算法

定位矩形在整张支票图像中所具有的全局特性, 不会因为图像分辨率的改变而发生较大的改变. 由此得到提示: 在搜索算法中, 采用隔点采样, 可以使数据量和运算量大大减少, 同时提高计算速度.

对基本算法的改进分成两个部分. 首先, 对量化指标  $q(x, y)$  的计算进行改进:

$$q(x, y) = \sum_{i=-K'}^{K'} \sum_{j=-K'}^{K'} D(x+2*i, y+2*j) * T1(2*i, 2*j), \quad (7)$$

其中  $K' = \lfloor K/2 \rfloor$ , 符号  $\lfloor \cdot \rfloor$  表示取整数运算. 改进后, 计算量将减少一半.

搜索算法也同样需要进行改进:

$$q^*(x^*, y^*) = \max_{\substack{0 \leq x \leq W \\ 0 \leq y \leq H}} q(2*i, 2*j), \quad (8)$$

其中  $x' = [x/2], y' = [y/2]$ , 改进后的运算量将减少为原来的四分之一。

参考图 4 的改进算法定位模板  $T1'$ , 相应于  $T1'$ , 量化指标  $q(x, y)$  计算公式为

$$q(x, y) = \sum_{i=-2K'}^{2K'} \sum_{j=-2K'}^{2K'} D(x+i, y+j) \times T1'(i, j), \tag{9}$$

其中  $K'$  的定义同式(7)。

### 2.4 算法分析

由于右上定位点与左上定位点有相似特征, 前述对左上定位点的分析说明, 对右上定位点也有类似结论, 而且对右上定位点的基本算法也可以类似地加以改进。

文献[5]中简单综述了文档处理中 5 类常用的图像倾斜检测和校正的算法: (1) 运用投影轮廓的方法; (2) 使用 Hough 变换技术; (3) 基于 Fourier 变换的方法; (4) 用最近邻分类; (5) 用相关性的方法。

上述 5 种算法一般从整幅文本图像中定位一条直线来达到定位的目的。有的方法对噪声特别敏感; 有的是从图像的全局特征出发, 因而抗干扰性能好, 但是运算量通常很大。比较而言, 我们提供的定位算法具有计算量小、运算速度快、定位准确等优点。

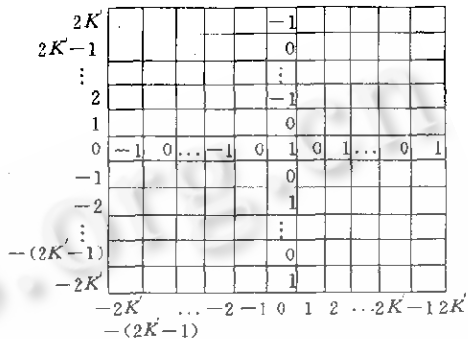


图4 改进算法的定位模板  $T1'$

### 3 各单元子图的倾斜校正

左上定位点和右上定位点所形成的直线的倾斜程度代表了整张支票图像的倾斜度, 系统按照这个倾斜角度对各个子图进行倾斜校正处理。

首先介绍计算机图形学中平移和旋转变换的基础知识<sup>[6]</sup>。

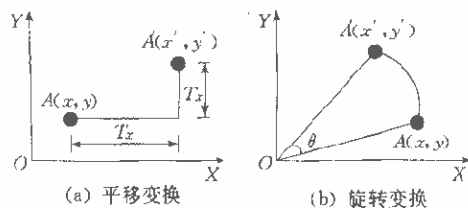


图5 平移和旋转变换图示

平移: 把平面直角坐标系上的点  $A(x, y)$  平移到点  $A'(x', y')$ , 如图 5(a) 所示, 用笛卡尔坐标表示为

$$\begin{cases} x' = x + T_x, \\ y' = y + T_y. \end{cases} \tag{10}$$

旋转: 把平面直角坐标系上的点  $A(x, y)$  绕原点  $O$  逆时针旋转  $\theta$  角变到  $A'(x', y')$ , 如图 5(b) 所示, 用笛卡尔坐标表示为

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \tag{11}$$

#### 3.1 算法介绍

文献[3]介绍了对任意倾斜角度的校正方法, 校正后的图像质量很高, 但不足之处是时间开销较大。对于小角度倾斜图像, 可以使用下面介绍的近似算法来完成校正功能。

设定位算法找到的左上定位点和右上定位点的坐标分别为  $P_1(x_1, y_1)$  和  $P_2(x_2, y_2)$ , 则整幅支票图像的倾斜角为

$$\theta = \arctan \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1}. \tag{12}$$

对于比较小的倾斜角度  $\theta$ , 有

$$\begin{cases} \sin\theta \approx \tan\theta = \frac{\Delta y}{\Delta x} = \frac{1}{K}, \\ \cos\theta \approx 1. \end{cases} \tag{13}$$

式中整数  $K = [1/\tan\theta]$ ,  $[ \cdot ]$  表示取整运算。所以有

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} \approx \begin{bmatrix} 1 & \frac{\Delta y}{\Delta x} \\ -\frac{\Delta y}{\Delta x} & 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}, \quad (14)$$

展开得

$$\begin{cases} x \approx x' + \frac{\Delta y}{\Delta x} \cdot y' = x' + \frac{y'}{K}, \\ y \approx -\frac{\Delta y}{\Delta x} \cdot x' + y' = y' - \frac{x'}{K}. \end{cases} \quad (15)$$

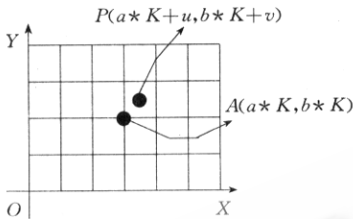


图6 录入图像的等分示意图

以  $K$  为步长来等分图像, 则原图像将被分成  $[W/K] * [H/K]$  个大小为  $K * K$  的正方形和在边缘上数目不定的小矩形, 如图 6 所示. 假设某个正方形(或者矩形)的图像块左下角点坐标为  $A(a * K, b * K)$ , 其中  $a$  的取值范围是  $0, 1, 2, \dots, [W/K] + 1$ ,  $b$  的取值范围是  $0, 1, 2, \dots, [H/K] + 1$ . 该图像块中任意一点坐标设为  $P(a * K + u, b * K + v)$ ,  $0 < u < K, 0 < v < K$ . 代入式(15), 并取整得

$$\begin{cases} x \approx x' + \frac{b * K + v}{K} \approx x' + b, \\ y \approx y' - \frac{a * K + u}{K} \approx y' - a. \end{cases} \quad (16)$$

对比式(10), 对于图像块中的每个点, 图像旋转相当于把它的坐标按式(16)进行平移变换, 这样就旋转变换转化成平移变换. 对同一个图像块中的所有像素点, 公式中的变量  $a$  和  $b$  都是相同的, 所以直接对正方形(或矩形)小图像块进行整体平移. 本算法的不足之处是, 平移处理会使相邻图像块的交接处出现一个像素的错位. 在实际支票图像中, 这种仅仅在一个像素范围内的错位对系统几乎没有任何影响.

对于倾斜角度  $\theta$  在  $-3^\circ \sim 3^\circ$  范围内的图像, 可以用上述小倾斜角度的改进算法予以校正. 大量实验表明, 对于 FBRs 提取和识别支票填写的内容来说, 这种方法已经能满足实用化的要求.

### 4 实例及结论

我们以图 1 的银行支票样张作为实例来总结本文所介绍的预处理内容和方法. 图 7 所示是软件定位处理结果. 图 8 是对日期子图的提取与分割. 图 9 是对金额和密码子图的预处理过程.



图7 软件定位

通过文中介绍的预处理方法, 完成对支票图像的软件定位、各单元子图的提取和背景去除、倾斜校正、方格去除、二值化和数字串分割, 可以得到单个数字字符的图像点阵, 直接送入单字符识别器进行识别. 在预处理器的基础上, 加上其他必要的识别功能构件, 就构成 FBRs 的原型系统. 这项工作在国家自然科学基金的支持下已基本完成.

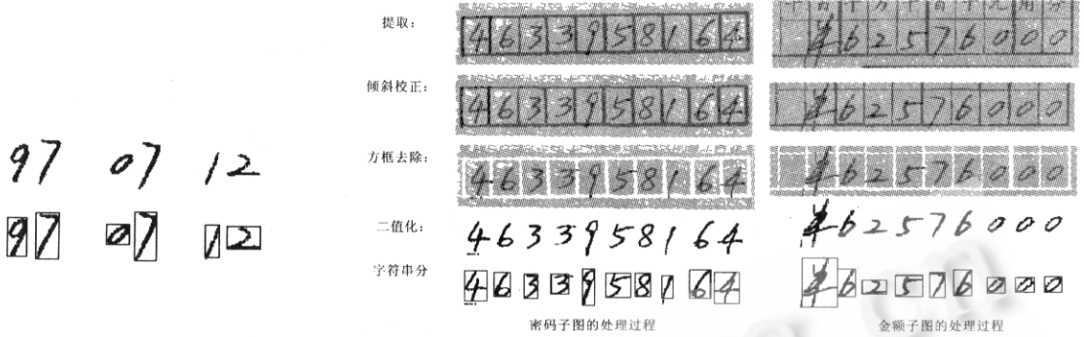


图 8 日期子图的提取与分割

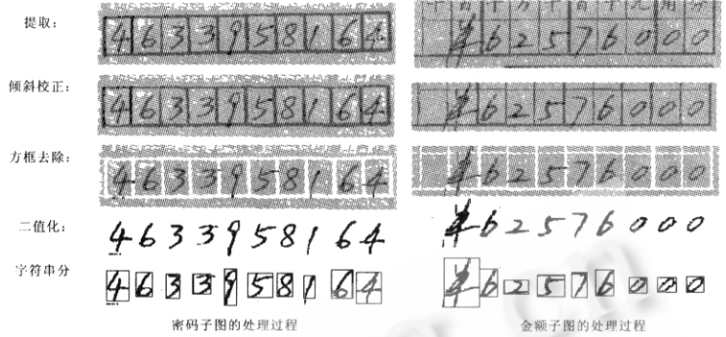


图 9 对支票密码和支票金额的预处理

致谢 本文的研究工作得到国家自然科学基金资助,此项目编号为 69775001.

参考文献

- 1 Impedovo S, Wang P S P, Bunke H. Automatic Bankcheck Processing. Singapore: World Scientific, 1997
- 2 Wang Song, Zhao Bin, Ma Feng *et al.* A fault tolerant Chinese bank check recognition system. In: Proceedings of 1997 IEEE International Conference on Intelligence Processing Systems. New York: IEEE, Inc. 1997. 1164~1168
- 3 胡家忠. 计算机文字识别技术. 北京: 气象出版社, 1994  
(Hu Jia-zhong. Computer Character Recognition Technologies. Beijing: Weather Press, 1994)
- 4 Zhao Bin, Su Hui, Xia Shao-wei. A new method for segmenting unconstrained handwritten numeral string. In: Proceedings of the 4th International Conference on Document Analysis and Recognition. Los Alamitos: IEEE Computer Society Press, 1997. 524~527
- 5 Sun Chang-ming, Si De-yi. Skew and slant correction for document images using gradient direction. In: Proceedings of the 4th International Conference on Document Analysis and Recognition. Los Alamitos: IEEE Computer Society Press, 1997. 142~146
- 6 卢传贤等. 实用计算机图形学. 峨眉山: 西南交通大学出版社, 1989  
(Lu Chuan-xian *et al.* Practical Computer Graphics. Emeishan: Southwest Communication University Press, 1989)

Design and Implementation of Preprocessing in Fault-tolerant Bankcheck Recognition System

ZHAO Bin MA Geng-jian XIA Shao-wei

(Department of Automation Tsinghua University Beijing 100084)

Abstract Based on the fault-tolerant bankcheck recognition system (FBRS), design and implementation of the preprocessing methods are proposed in this paper, which include software locating algorithm, extracting and erasing background of each sub-image, skew correction of each sub-image, deleting restricting rectangle frame of some sub-images and segmentation of characters' string. Among them, the software locating algorithm and skew correction are discussed in detail.

Key words Preprocessing, pattern recognition, form processing, document analysis.