

# 印鉴识别系统中印鉴录入的研究

刘文娟 何怡刚

(湖南大学电气与信息工程学院 长沙 410082)

**摘要** 通过对印鉴录入系统的研究,本文讲述了印鉴录入的过程,重点给出了圆章、椭圆章和方章的精确查找方法,在 VC++6.0 的平台上实现并证明了该方法的实用性,为银行验印系统的印鉴录入部分提供了参考。

**关键词** 直线探测,圆探测,椭圆探测,Hough 变换

## Study on the Seal Input of the Automatic Seal-identifying System

LIU Wen-juan HE Yi-gang

(College of Electrical and Information Engineering, Hunan University, Changsha 410082, China)

**Abstract** This paper describes the process of the seal input part in automatic seal-identifying system. Specially it gives a detailed method for finding the circle seal, ellipse seal, square seal. Such method has got quite good results in practice. Also it affords a reference for the design of seal imprint identification system in bank.

**Keywords** Line detection, Circle detection, Ellipse detection, Hough transform

### 1 引言

印鉴识别系统就是利用计算机算法来实现印鉴的自动识别。它的基本流程是:通过扫描仪或摄像机等图像采集设备,将预留印鉴图像采集到计算机内,经过特定图像处理算法形成电子标准印鉴。印鉴核实时,待测凭证印鉴(现金支票、汇帐单、凭证)通过图像采集设备被采集入计算机。经过图像处理和模式识别相关算法,计算机自动来判决印鉴的真伪,以此来代替传统的手工折角核对方法。预留标准印鉴作为验印的标准,它的质量直接影响到验印的通过率和识别率。

### 2 图像预处理

本文均以红色印泥的印鉴为例来对印鉴作处理。首先用 HSI 方法对图像(现金支票、汇帐单、凭证)如图 1 所示的印鉴卡进行分解,取出图片的红色信息,再对红色信息图片作自适应二值化处理,使印鉴的红色部分为 1,其它部分为 0。

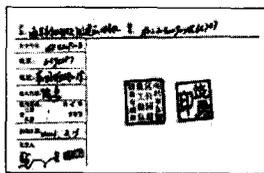


图 1 用于印鉴录入印鉴卡

割法;2. 印鉴骨架轮廓查找法。

(1)直接图片分割法:图片  $S$ , 宽为  $w$  像素, 高为  $h$  像素, 将图片  $S$  以每  $t$  ( $0 < t < w$ ) 像素为单位进行  $x$  和  $y$  方向的划分, 则  $x$  方向为  $m = w/t$ ,  $y$  方向为  $n = h/t$ , 设置域值  $k$  ( $0 < k < 1$ ), 若  $S_{ij}$  ( $0 < i < m, 0 < j < n$ ) 中有大于  $k$  的红色像素占有率, 则  $S_{ij}$  为 1, 否则为 0, 用相邻检测法(如果一个单位为 1, 若它相邻的 8 个块中为 1, 则它们在同一区域)得到多个区域的范围。优点:速度快。缺点:对两个距离较近的印鉴难作处理。

(2)印鉴骨架轮廓查找法:对二值化的图片作细化处理, 得到图形骨架。用相邻检测法得到多个印鉴区域的范围。优点:能较好处理印鉴距离较近问题。缺点:速度较直接图片分割法慢。

(3)查找印鉴范围方法选择:对印鉴有规定区域, 印鉴间有较大距离的情况选择直接图片分割法, 否则选择印鉴骨架轮廓查找法。

(4)图 1 所示的印鉴卡经过印鉴范围查找后, 可得到三个印章, 图 2 是其中的一个图。



图 2 印鉴范围查找所得图之一

### 3 印鉴范围查找

验印系统一般将图片(如现金支票)通过扫描仪或者摄像头等硬件把影像转变成数字图片, 图片大小一般在  $640 * 480$  以上, 如果直接对整个图片进行印鉴查找处理, 将直接影响图片录入运行时间。所以先查找和保存印鉴的范围, 再在这个范围内精确查找印鉴。查找印鉴范围的方法:1. 直接图片分

### 4 印鉴的精确查找

以常用的印鉴(方章、圆章、椭圆章)为例。预备处理:采用 Hilditch 算法进行细化处理, 对细化印鉴作相对边缘连续采集, 存储在  $P[size]$  的数组中。图 3 为二值化后的印鉴图,

刘文娟 硕士研究生, 研究方向为神经网络理论; 何怡刚 教授, 博士生导师, 研究方向为智能测试与诊断、神经网络理论、集成电路设计、电力规划与设计等。

图 4 为细化后的印鉴图。

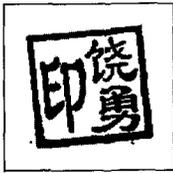


图 3 二值化后的印鉴图



图 4 细化后的印鉴

#### 4.1 方章的精确查找

##### 4.1.1 Hough 变换找直线

如果对整张图片作 Hough 变换来探测直线,将消耗大量的时间。这里只对  $P[size]$  作 Hough 变换,可高速得到方章的 4 条直线。交点分析:得到的 4 条直线  $\theta$  相近为一组,分成两组,两组间两两相交得到 4 个交点坐标  $Pti[4]$ 。

(1)得到印鉴矩形框、倾斜角(锐角)、中心坐标:对  $Pti[4]$  求得  $\min X, \min Y, \max X, \max Y$ , 得到包围印鉴的最大矩形框。对 4 条直线的倾斜角  $\theta$ , 如果  $\theta > 90^\circ, \theta = 180^\circ - \theta$ , 求得  $\min \theta$  作为印鉴的倾斜角。中心坐标  $P_x = (\min X + \max X) / 2, P_y = (\min Y + \max Y) / 2$ 。

(2)验证是否为方章:在 4 条支链中,若  $\theta[i] < 45^\circ$ , 在  $x$

$$\begin{cases} x^2 + y^2 + dx_1 + ey_1 + f = 0 \\ x^2 + y^2 + dx_2 + ey_2 + f = 0 \\ x^2 + y^2 + dx_3 + ey_3 + f = 0 \end{cases} \Rightarrow \begin{cases} d = \frac{(y_1 - y_2) * (x_3^2 + y_3^2 - x_2^2 - y_2^2) + (y_3 - y_2) * (x_2^2 + y_2^2 - x_1^2 - y_1^2)}{(y_1 - y_2) * (x_2 - x_3) + (y_3 - y_2) * (x_1 - x_2)} \\ e = \frac{(x_1 - x_2) * (x_3^2 + y_3^2 - x_2^2 - y_2^2) + (x_3 - x_2) * (x_2^2 + y_2^2 - x_1^2 - y_1^2)}{(x_1 - x_2) * (y_2 - y_3) + (x_3 - x_2) * (y_1 - y_2)} \\ f = -x_1^2 - y_1^2 - d * x_1 - e * y_1 \end{cases}$$

把印鉴细化后的连续边缘点  $P[size]$  分成 3 等份  $s = size / 3$ 。

分别将  $P[i], P[s+i], P[2 * s+i] (0 < i < s)$  这三点代入方程组求得  $d[i], e[i], f[i]$ 。半径  $r[i] =$

$$\frac{\sqrt{d[i]^2 + e[i]^2 - 4f[i]}}{2}, \text{中心坐标 } cx[i] = \frac{-d[i]}{2}, cy[i] = \frac{-e[i]}{2}, \text{由此得到 } s \text{ 组 } r, cx, cy \text{ 的值。最后对 } r[i] (0 < i < s) \text{ 求}$$

方差,设一个阈值  $f$ , 如果方差大于  $f$ , 舍弃  $r[i]$  中偏离平均值最远的一个,继续对剩下的  $r[i]$  求方差,直到方差值小于  $f$ 。最后将剩下的  $r$  求平均值,从而得到一个较准确的  $r$ 。  $cx, cy$  的求法同  $r$  的求法。

##### 4.2.2 验证是否为圆章

以  $x$  方向从  $cx - r$  递增到  $cx + r$ , 代入方程  $x^2 + y^2 + dx + ey + f = 0$ , 求得  $y_1, y_2$ , 设立一个计数器  $n$ , 如果坐标  $(x, y_1)$  或  $(x, y_2)$  处在二值化图中为 1, 计数器  $n$  加 1, 设一个阈值  $k$ , 如果  $n / (2r * 2) > k$ , 则圆存在, 否则圆不存在。

#### 4.3 椭圆章查找

##### 4.3.1 分段

椭圆的一般方程  $ax^2 + 2bxy + cy^2 + dx + ey + f = 0$  中有 6 个未知数, 其中  $2bxy$  项是由旋转产生,  $dx$  和  $ey$  项是平移

$$T \cdot D \cdot T^T = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} A & B \\ B & C \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} = \begin{pmatrix} A \cos^2 \theta - 2B \sin \theta \cos \theta + C \sin^2 \theta & (A - C) \sin \theta \cos \theta + B \cos(2\theta) \\ (A - C) \sin \theta \cos \theta + B \cos(2\theta) & A \sin^2 \theta + 2B \sin \theta \cos \theta + C \cos^2 \theta \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

$(A - C) \sin \theta \cos \theta + B \cos(2\theta) = 0$ , 根据二倍角公式, 可知  $(A - C) \sin(2\theta) + 2B \cos(2\theta) = 0$ , 当  $\theta \neq 45^\circ$  时, 有  $\tan(2\theta) = 2B / (C - A)$ , 最终得到  $\theta = \frac{1}{2} \arctan(2B / (C - A))$ 。对  $D$  求特征值, 根据特征值的定义  $|\lambda E - D| = 0$  其中  $E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ , 代入已知数据

$$\begin{vmatrix} \lambda - A & -B \\ -B & \lambda - C \end{vmatrix} = 0, \text{由此可得方程 } (\lambda - A)(\lambda - C) - B^2 = 0,$$

方向从  $\min X$  以 1 像素单位递增到  $\max X$ , 得到  $y_i = x_i * \tan \theta[i] + b_i$ , 设立一个计数器  $n$ , 如果坐标  $(x_i, y_i)$  对应的像素点的值在二值化图中为 1, 则计数器  $n$  加 1。设阈值为  $k$ , 如果  $k > \{n / (\max X - \min X)\}$ , 则直线存在, 否则直线不存在。若  $\theta[i] > 45^\circ$ , 则在  $y$  方向从  $\min Y$  以 1 像素单位递增到  $\max Y$ , 得到  $x_i = y_i * \tan(90^\circ - \theta[i]) + b_i$ , 设立一个计数器  $n$ , 如果坐标  $(x_i, y_i)$  处在二值化图中为 1, 计数器  $n$  加 1。人工设一个阈值  $k$ , 如果  $k > \{n / (\max Y - \min Y)\}$ , 则直线存在, 否则直线不存在。如果有 3 条以上的直线存在(可能方章有一条边不清楚), 则找到一个方章, 否则没有找到方章。

(3)旋转  $\min \theta$  为印鉴的倾斜角,  $(x' \ y') =$

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}, x, y \text{ 为原始印鉴坐标, 旋转 } \min \theta \text{ 后得到新坐标。}$$

#### 4.2 圆章的精确查找

通常的做法是用 Hough 变换检测得到圆心, 但考虑到 Hough 变换很消耗内存和时间, 本文将提出一种新的算法。

##### 4.2.1 求圆章的中心坐标和半径 $r$

圆的一般方程  $x^2 + y^2 + dx + ey + f = 0$  中有 3 个未知数, 求三元一次方程组, 可得到中心坐标、半径  $r$ 。

产生。根据椭圆的中心对称性, 用 Hough 变换求椭圆的中心(参考了文献[5], 设置一个存储空间  $B[width * height]$ , 对边缘连续点  $P[size]$  中每一点  $P[i] (0 < i < size)$ , 求得该点和除该点以外所有点的中点坐标  $cx, cy$  保存在  $B[cx * height + cy]$  空间中, 然后对  $B$  空间进行排序得到空间中最大值项  $B_{\max}[n]$ , 其中  $n = cx * height + cy$ , 得到椭圆中心  $centerX$  和  $centerY$ 。对图片上所有点作相对移动, 使椭圆的中心为原点。所以方程中的  $dx$  和  $ey$  项消除, 得到新方程  $Ax^2 + 2Bxy + Cy^2 = 1$ , 方程中有 3 个未知数, 把  $P[size]$  分成 3 等份  $s = size / 3$ 。

(1)方程求解, 得到中心坐标、短轴  $n$ 、长轴  $m$  和倾斜角  $\theta$ 。取 3 点  $P[i], P[i+s], P[i+2s] (0 < i < s)$  代入方程  $Ax^2 + 2Bxy + Cy^2 = 1$  中求得  $A, B, C$ 。把  $Ax^2 + 2Bxy + Cy^2 = 1$  写成矩阵的形式:

$$\begin{pmatrix} x & y \end{pmatrix} \times \begin{pmatrix} A & B \\ B & C \end{pmatrix} \times \begin{pmatrix} x \\ y \end{pmatrix} = 1$$

$$\text{设 } D = \begin{pmatrix} A & B \\ B & C \end{pmatrix}, \text{ 旋转矩阵 } T = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, D' =$$

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \text{ 存在 } T \cdot D \cdot T^T = D', \text{ 代入得:}$$

即  $\lambda^2 - (A + C)\lambda + AC - B^2 = 0$ , 令  $bb = -(A + C), cc = AC - B^2$ , 则  $\lambda^2 + bb\lambda + cc = 0$ , 若椭圆的长短轴为  $n, m$ , 解方程得到两个解为  $1/n^2$  和  $1/m^2$ 。

##### 4.3.2 验证

以  $x$  方向从  $cx - r$  递增到  $cx + r$ , 代入方程  $ax^2 + 2bxy + cy^2 + dx + ey + f = 0$ , 求得  $y_1, y_2$ , 设立一个计数器  $n$ , 如果坐

(下转第 153 页)

训练时间最低,但获得的分类效果最差。本文提出的算法由于采用了改进的 PCM 算法来计算样本的隶属度值,使得部分样本(实际上是很少部分样本)可能会被训练多次,导致了时间开销上比最小包围球方法稍大。相对于标准 SVM 而言,在数据集 Sonar 上,时间开销基本相当,训练时间差不多,但在其它几个数据集上,时间优势还是比较明显的。

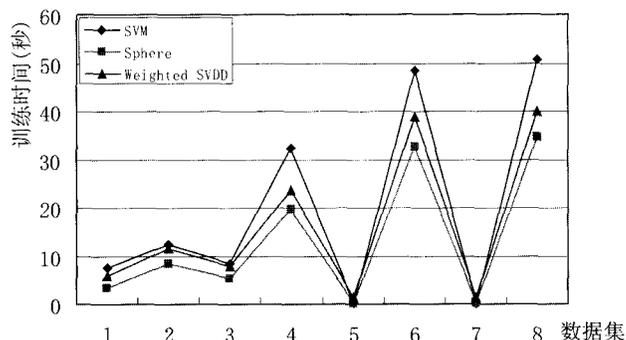


图 1 训练时间对比图

**结束语** 针对 1-类分类问题中的支持向量数据描述 (SVDD), 根据其特点的相关分析, 本文将其成功地应用到多类分类问题中。另外, 为了更好地处理训练数据中的噪声和异常点, 在算法中引入了模糊技术, 并采用一种改进的 PCM 方法来计算样本的模糊隶属度值, 从而提出了一种基于改进 PCM 方法的加权支持向量数据描述多类分类算法。经理论证明, 该算法所提出的分类决策函数满足贝叶斯优化决策理论。最后将该算法与其它的方法进行了实验比较, 得到了较满意的结果。

### 参考文献

- [1] Platt J C, Cristianini N, Shawe-Taylor J. Large margin DAGs for multiclass classification. *Advances in Neural Information Processing Systems*, MIT Press, Cambridge, MA, 2000; 547-553
- [2] Hsu C W, Lin C J. A comparison of methods for multi-class support vector machines. *IEEE Trans. on Neural Networks*, 2002, 13(2): 415-425
- [3] Weston J, Watkins C. Multi-class support vector machines. Technical Report, SD2TR298204. Department of Computer Science, Royal Holloway University of London, 1998
- [4] Zhang Y, Chi Z X, Liu X D, et al. A novel fuzzy compensation multi-class support vector machines. *Applied Intelligence*, 2007, 27(1): 21-28
- [5] Zhu M L, Chen S F, Liu X D. Sphere-structured support vector machines for multi-class pattern recognition. *Lecture Notes in*

*Computer Science*, 2003, 2639: 589-593

- [6] Wang J, Neskovic P, Cooper L N. Bayes classification based on minimum bounding spheres. *Neurocomputing*, 2007, 70: 801-808
- [7] Schölkopf B, Burges C, Vapnik V N. Extracting support data for a given task // *Proceedings of First International Conference on Knowledge Discovery and Data Mining*. 1995: 252-257
- [8] Schölkopf B, Platt J, Shawe-Taylor J, et al. Estimating the support of a high-dimensional distribution. *Neural Computation*, 2001, 13(7): 1443-1471
- [9] Tax D, Duin R. Support vector domain description. *Pattern Recognition Letter*, 1999, 20: 1191-1199
- [10] Tax D, Duin R. Support vector data description. *Machine Learning*, 2004, 54: 45-66
- [11] Lee D, Lee J. Domain described support vector classifier for multi-classification problems. *Pattern Recognition*, 2007, 40: 41-51
- [12] Song Q, Hu W J, Xie W F. Robust support vector machine with bullet hole image classification. *IEEE Trans. on Systems, Man and Cybernetics*, 2002, 32(4): 440-448
- [13] Hu W J, Song Q. An accelerated decomposition algorithm for robust support vector machines. *IEEE Trans. on Circuits and Systems*, 2004, 51(5): 234-240
- [14] Lin C F, Wang S D. Fuzzy support vector machines. *IEEE Trans. on Neural Networks*, 2002, 13(2): 464-471
- [15] Lin C F, Wang S D. Training algorithms for Fuzzy support vector machines with noisy data // *IEEE 8th Workshop on Neural Networks for Signal Processing*. 2003: 517-526
- [16] Krishnapuram R, Keller J M. A possibilistic approach to clustering. *IEEE Trans. on Fuzzy System*, 1993, 1(2): 98-110
- [17] Krishnapuram R, Keller J M. The possibilistic c-means algorithm: insights and recommendations. *IEEE Trans. on Fuzzy System*, 1996, 4(3): 385-393
- [18] Gardner A B, Krieger A M, Vachtsevanos G. One-class novelty detection for seizure analysis from intracranial EEG. *Journal of Machine Learning Research*, 2006(7): 1025-1044
- [19] Lee S W, Lee S W. SVDD-based illumination compensation for face recognition // *The 2nd International Conference on biometrics*. LNCS 4642, 2007: 154-162
- [20] Tao X M, Liu F R, Zhou T X. A novel approach to intrusion detection based on support vector data description // *The 30th Annual Conference of IEEE industrial Electronics Society*. 2004: 2016-2021
- [21] <http://www.ics.uci.edu/~mllearn/MLRepository.html>

(上接第 130 页)

标  $(x, y_1)$  或  $(x, y_2)$  处在二值化图中为 1, 计数器  $n$  加 1, 设一个阈值  $k$ , 如果  $(n/(2r * 2)) > k$ , 则椭圆存在, 否则椭圆不存在。

#### 4.3.3 旋转

**结束语** 本文介绍了印鉴识别系统中的印鉴录入过程, 在本文中提到的精确查找圆章、方章、椭圆章的方法在实验中效果不错, 并且已经应用到实际的印鉴识别系统中。

### 参考文献

- [1] 侯宇. 圆和椭圆边缘检测的快速方法. *中国计量学院学报*, 2000
- [2] 姜震, 胡钟山, 杨静宇. 支票自动处理系统中的图像处理及识别. *南京大学理工大学学报*, 1999
- [3] 求是科技. *Visual C++ 数字图像识别技术典型案例*. 北京: 人民邮电出版社, 2004
- [4] 王怀群. 二值图象的细化. *无锡轻工大学学报*, 2001
- [5] 屈稳太. 基于弦中点 Hough 变换的椭圆检测方法. *浙江大学学报(工学版)*, 2005