

# IC卡在线检测的码号定位分割方法研究

郑胜林 潘保昌 徐杜 蒋永平

(广东工业大学信息工程学院数字图像技术研究所, 广州 510643)

**摘要** 针对IC卡的版面复杂性及版面质量,对IC卡在线检测中的码号定位分割问题进行了研究,提出了一种IC卡码号在线检测的定位分割方法。该方法采用预测样卡定位参数,结合初定位、宽区间定位、区间缩小精确定位对码号区定位,解决了不同版面、不同位置码号的IC卡定位问题。对定位方法进行了详细论述,采用求多投影直方图综合分割法对码号字符进行分割,有效地解决了IC卡在线检测中码号的定位分割问题。实验结果表明了方法的可行性与实用性。

**关键词** 定位 分割 IC卡

**中图分类号**: TP391.41 **文献标识码**: A **文章编号**: 1006-8961(2004)12-1430-07

## An on-line Method of Locating and Segmenting Code for IC Card

ZHENG Sheng-lin, PAN Bao-chang, XU Du, JIANG Yong-ping

(Institute of Digital Image Technology, Faculty of Information Engineering,  
Guangdong University of Technology, Guangzhou 510643)

**Abstract** Aiming at the complexity and the quality of IC cards, the location of code region and the segmentation of characters on IC card are studied. An on-line method of locating code region and segmentating characters on IC card are presented in the paper. It is based on the combination of the pre-testing location, the original position locating, the coarse region locating and the exact one. The location for different kinds of IC cards and different position of code regions on IC card are solved. The location method is discussed in detail. The multi-projection histograms are presented to segment the code. The experiment shows the result of the method is good.

**Keywords** locating, segmentation, IC card

## 1 引言

IC卡作为一种信息载体,近年来在银行、医疗、保险、交通、通信等领域得到广泛应用。通常,IC卡版面上有一些特殊信息,如卡号及其他号码,将这些号码统称码号。码号必须和IC卡芯片内对应的号码一致,否则将出现严重错误。IC卡表面上的码号一般是由计算机控制喷码机喷上去的。由于种种原因,在IC卡喷码时,可能出现错码、断码、漏码和重码,或喷码质量不好等(统称次卡)情况。一旦出现这些情况,该IC卡将不能使用。在IC卡制作过程中,需要将次卡检出。传统的方法是靠人工抽样进行,但人

工方法很难保证质量。因此在生产过程中,迫切需要对IC卡码号实现在线检测。在线检测的工作过程是:IC卡在生产线上喷码后,对IC卡进行图像扫描,并对图像中IC卡的码号字符区进行定位分割,对分割后的字符进行识别,并将识别结果与原喷码数据进行比对处理,然后根据比对结果自动地分成卡,剔除次卡,实现IC卡质量的在线检测。

IC卡的版面样式通常都不一样,每种卡的码号的数量不一。号码出现的位置、字体、尺寸大小也不完全相同,而且数码与数码间常出现粘连。码号周边通常都有其他信息,再加上噪声的影响,这使得IC卡码号的在线定位分割难度相当大<sup>[1~3]</sup>。图1示出了两张不同IC卡的图像实例。针对众多的版面和不



图 1 不同 IC 卡实例图片

同喷码质量对 IC 卡码号定位分割带来的难度,以及定位分割中存在的问题,提出了一种 IC 卡在线检测码号定位分割方法。该方法采用 IC 卡预测样卡定位参数技术,将初定位、宽区间定位、区间缩小精确定位结合,综合对码号区定位,以及采用多投影直方图综合分割法,有效地解决了 IC 卡在线检测中的定位分割问题。

## 2 IC 卡码号区定位方法

IC 卡在线检测的码号定位分割方法首先通过预测样卡码号的定位参数,然后根据预测定位参数对码号区进行初定位,根据初定位进行宽区间定位,在宽区间定位的基础上再进行区间缩小精确定位,从而准确找出 IC 卡实际码号位置。

预测样卡码号分割定位参数是在检测一批相同类别 IC 卡之前,选取一张样卡,通过预先获取样卡码号位置的初始值,以及卡片上各个码号段的属性,为在线检测分割时提供参考数据。这是因为在 IC 卡的生产过程中,一般是批量生产,批量检测。在对 IC 卡分割处理时,考虑 IC 卡的版面样式通常都不一样,每种卡的码号也可能出现在不同的位置。采用预测样卡码号定位分割参数的方式可以很方便处理不同版面、不同位置码号的 IC 卡。

初定位是根据预测的定位分割参数,对扫描的 IC 卡图像码号区域进行定位。该定位区域是一个参考值。通常初定位区不一定能完全将码号包含在区间内。

宽区间定位则是在初定位基础上增加区域的偏移量,使区域加大,以保证码号落在整个定位区域内。宽区间定位偏移量的选取非常重要,太大将造成过多的干扰信息落入区间而给定位带来困难,太小则码号可能不能完全落入区间内,而造成定位分割时不能找出完整的码号。

区间缩小精确定位则是在宽区间定位确定的区域内,进一步分析码号和干扰,而将干扰排除掉,准确找出实际码号所在的区间,完成码号区的定位分割。

考虑到 IC 卡的码号及文字信息是按 X 方向喷印的,先求出码号区 Y 方向区域较先求 X 方向区域难度要小。因此本算法按先找 Y 方向码号区,再找 X 方向码号区的方式进行。

### 2.1 预测样卡码号定位参数

预测样卡码号定位参数包括确定样卡中码号区的坐标以及自动获取定位参数等工作。预测样卡码号定位参数方法如下:

(1) 预定码号区域位置 在预定码号位置坐标时,先从测试样本中取出一幅质量好的 IC 卡样卡图像,在图像中采用人机交互的方式定出每个码号区域的范围作为预定码号区域位置。

(2) 码号区域数目测定 由于码号区域不止一个,所以系统要根据所预定的区域进行自动统计,计算出码号区域个数,作为要处理的码号区域数。

(3) 码号区域坐标测定 码号区域的坐标是定位的主要参考数据,系统根据预定码号区域位置检测出码号区域的坐标,将所有码号区的坐标存放在结构数组  $rect * CodeCoor$  中,这个数组空间与码号区域数相同。

(4) 码号区域属性参数测定 码号区域的属性参数包括每个码号区的总宽度  $CodeWidth$ 、单个字符的宽度  $ChaWidth$ 、单个字符的高度、码号区域的字符总数。这些参数可以根据码号区域的数目和码号区域的坐标  $CodeCoor$ ,采用图像处理和字符分割技术自动测定计算。

系统预测样卡码号定位参数后,将所测定的参数即预测参数放入一个参数结构中,供在线检测使用,从而解决了同一批 IC 卡的定位属性问题。如果

要生产另一批卡片,只需对另一批卡片的样本进行测定即可。

### 2.2 Y 方向定位分割方法

Y 方向码号定位分割按初定位、宽区间定位、区间缩小精确定位进行。

在定位分割中,设预定码号区域坐标符号为:  $X01, X02, Y01, Y02$ , 码号区域精确定位坐标符号为  $X11, X12, Y11, Y12$ 。

#### 2.2.1 初定位

在 IC 卡批量检测过程中,预测参数只是定位分割的参考值。在线检测时,根据这些参考值,求出卡片的每个码号区初始定位坐标。初始定位坐标包括 X 和 Y 两个方向。IC 卡扫描图像坐标系统的原点在图像左上角,横轴为 X 轴,纵轴为 Y 轴。

对于区域的初定位,起始点坐标是定位的关键。系统的定位方法是先求出每个码号区的起点坐标,即左上角坐标  $X11, Y11$  和左下角坐标  $X12, Y12$ , 然后根据预测参数计算出区域的另外两个坐标

$X21, Y21$  和  $X22, Y22$ 。

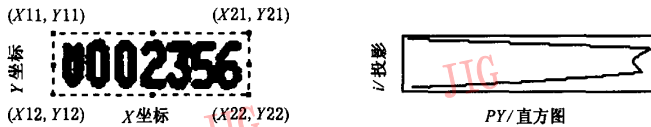
求起点 Y 方向的定位坐标时,将预测参数码号区域的坐标作为定位初值坐标  $X01, Y01$  和  $X02, Y02$ 。然后按一定宽度向 X 方向投影,求出  $Y11, Y12$ 。投影直方图公式为

$$PY_i = \sum_{j=X_1}^{X_n} F_{ij}, i = Y_1, Y_2, \dots, Y_n \quad (1)$$

其中,  $F_{ij}$  是卡片上的扫描图像数据。

$$\begin{aligned} X_1 &= X01, X_2 = X01 + 1, \dots, \\ X_i &= X01 + i - 1, \dots, X_n = X02; \\ Y_1 &= Y01, Y_2 = Y01 + 1, \dots, \\ Y_i &= Y01 + i - 1, \dots, Y_n = Y02 \end{aligned}$$

图 2 为码号区初定位的实例。图 2(a) 是理想情况下的码号区字符和投影直方图 PY, 其中码号区字符的虚线部分是预测参数确定的区域  $X01, Y01$  和  $X02, Y02$ , 从投影图可以看出,在理想情况下的字符区的  $Y11$  和  $Y12$  很容易从投影直方图求得。图 2(b) 是出现偏移情况的初定位实例。



(a) 理想情况码号区初定位和 Y 方向投影图



(b) 出现偏移的初定位实例

图 2 码号区初定位

#### 2.2.2 宽区间定位

实际在线处理时,由于 IC 卡的设计、印刷质量、系统硬件、数据传输、扫描传送的综合误差等原因,使码号字符区域并不一定完全在初定位的坐标区域之内。例如,在图 2(b)中,虚线方框将部分字符切去。这时,就需要对参数区域进行修改,使码号区域字符全部包含在虚线方框内。处理的方法是采取宽区间定位。宽区间定位是扩大投影区域,即扩大虚线方框。宽区间定位要保证码号全部落入定位区间。对投影区域参数  $X01, Y01$  和  $X02, Y02$  进行修改如下:

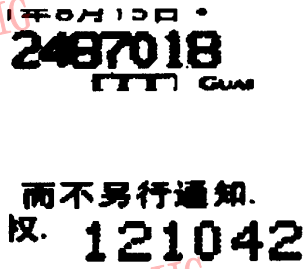
$$\begin{aligned} X01 &= X01 - \Delta X, X02 = X02 + \Delta X; \\ Y01 &= Y01 - \Delta Y, Y02 = Y02 + \Delta Y \end{aligned} \quad (2)$$

其中,  $\Delta X, \Delta Y$  是偏移量,其大小与卡片、码号区及系统性能有关。

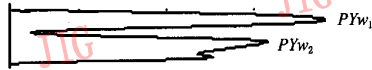
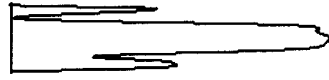
确定投影区域后,可按式(1)进行处理。这时由于投影区域扩大,虚线已将字符区全部包含,但也使投影区域包含了一些干扰信息,如图 3(a)所示。

#### 2.2.3 区间缩小精确定位

为了让投影区全部包含码号区,采用宽区间定位。宽区间定位可能使定位区域内含有其他干扰信号。要进行准确的定位,必须排除这些干扰,求出码



(a) 定位区域



(b) x 方向直方图

图 3 带有干扰的定位区域和 X 方向投影直方图

号的准确坐标。为此,采用区间缩小定位方法,以排除落入宽区间内的干扰信号,从而实现码号区的精确定位。

按式(1)对这些不同情况进行计算,得到 X 方向的投影直方图 PY,如图 3(b)所示。在 PY 中,存在 k 个结果区域,如图中标出的 PYw<sub>1</sub>,PYw<sub>2</sub> 两个结果区域(k=2 的情况)。区间缩小精确定位就是要从这些区域中找出真正的字符区,将宽区间缩小到实际码号字符区间。

先求出直方图中 PY 投影区的平均高度

$$PY_{zh} = \frac{1}{r} \sum_{i=1}^r PY_i \quad (PY_i \neq 0) \quad (3)$$

其中,r 为 PY 中不为零的行数。

再根据 PY 求出每个区域的宽度 PYw<sub>i</sub>(i=1~k)和每个区域的平均高度 PYh<sub>i</sub>(i=1~k)。图 3(b)示出其中一个码号区的 PYw<sub>i</sub>(i=1~k),这里区域数 k=2。

根据两个条件来判断求出 Y11 和 Y12 的坐标。

(1) PYw<sub>i</sub> 中取出最大值 PYw<sub>m</sub> 作为候选字符区,即区域宽度最大值;

(2) 求出 PYw<sub>m</sub> 对应的区域平均高度 PYh<sub>m</sub> 要满足:PYh<sub>m</sub>>T;

其中,取阈值 T=PYzh/3。因为每个码号段的特征是独立的,用这种方式求得的阈值是浮动的,适用于各种情况的判断。

从满足条件的码号段 PYw<sub>m</sub> 中求出其起点和终点,即得到码号区定位坐标 Y11 和 Y12。

经区间缩小精确定位处理后,Y 方向的干扰全部被排除。

### 2.3 X 方向定位

X 方向的定位同样需要进行初定位、宽区间定位和区间缩小精确定位。不同的是,X 方向的初定

位和宽区间定位已在确定 Y 方向坐标时完成。这里,X 方向的定位实际上是完成 X 方向的区间缩小精确定位。

X 方向定位是求出起始点 X 方向左上角和左下角坐标 X11 和 X12。这里,X11=X12,所以只要确定 X11 即可。

求 X 方向定位坐标时,Y 方向的准确坐标 Y11 和 Y12 已得到,其定位初值坐标取 X01、Y11 和 X02、Y12,然后按一定宽度向 Y 方向投影,求 Y 方向投影直方图。再根据直方图求出 X11。投影计算公式为

$$PX_j = \sum_{i=Y_1}^{Y_n} F_{ij} \quad (j = X_1, X_2, \dots, X_n) \quad (4)$$

其中,F<sub>ij</sub>是卡片上的扫描图像数据。

$$X_1 = X01, X_2 = X01 + 1, \dots, X_i = X01 + i - 1, \dots, X_n = X02$$

$$Y_1 = Y01, Y_2 = Y01 + 1, \dots, Y_i = Y01 + i - 1, \dots, Y_n = Y02$$

码号区域字符和投影直方图 PX 如图 4 所示。图 4 示出了不同情况下码号区字符的 3 种情况实例。图 4(a)和图 4(b)中,PX 的第一个字符区的 X 起点坐标就是码号的定位区域坐标。图 4(c)中 PX 的第一字符区也有可能是干扰。所以对于求出的 PX,还需要进一步确定第一区域段是字符还是干扰。

为此,采用以下算法对其进行处理:

(1) 求出投影直方图 PX 的平均高度

$$PX_{zh} = \frac{1}{c} \sum_{i=1}^c PX_i \quad (PX_i \neq 0) \quad (5)$$

其中,c 为 PX 中不为零的列数。

(2) 确定候选字符区域 将 PX 中的前面 k(k≤3)个字符区域作为起始字符区的候选区(这里考虑字符段前可能有两个干扰的情况),求出候选区的宽度 PXw<sub>i</sub>及每个区域的平均高度 PXh<sub>i</sub>。每个候选字符区要满足以下条件:

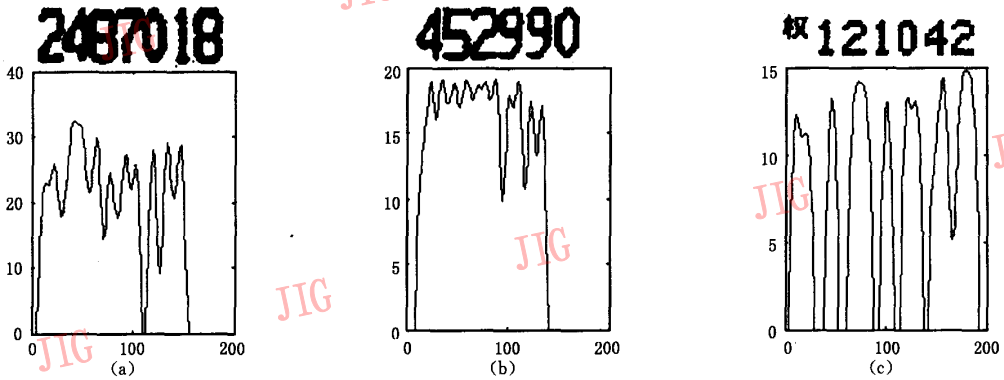


图 4 确定 Y 方向定位坐标后的图像及 X 方向投影直方图

- 区域宽度:  $PXw_i > ChaWidth/2$
- 区域平均高度:  $PXh_i > PXzh/2 \quad (i=1 \sim k)$  (6)

ChaWidth 是在预测参数中所获取的码号区的字符平均宽度。判断阈值取 ChaWidth/2 是为了保证不丢掉码号中的有效信息。图 4(a)和图 4(b)无干扰信号,只有一个候选字符区。式(6)的两个判断条件可消除一些毛刺噪声,如图 5 所示。

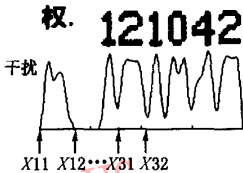


图 5 坐标位置示意图

确定了候选字符区域后,在 PX 的投影直方图中从左到右搜索,将每个候选区域的起始点坐标找出,起始点坐标为

$$X_{i,j} \quad (i = 1 \sim k; j = 1, 2)$$

其中,  $i$  表示候选字符区序号,  $j=1$  表示每个选字字符区的起始坐标,  $j=2$  为结束坐标。

图 5 示出了其中一个候选区参数的示意图 ( $k=3$  的情况),可以看出,字符区域宽度为

$$PXw_i = X_{i,2} - X_{i,1}$$

(3) 排除干扰求出定位字符区

进一步对候选字符区  $PXw_i$  进行判断,确认求出的 X 起点坐标,其判断步骤为

- ① 对  $PXw_i$  进行判断,如果  $k=1$ ,则将  $PXw_i$  作为码号的第一字符。
- ② 当  $k > 1$  时,求出第  $i$  个候选区的有关参数:  
 区域宽度:  $W_i = PXw_i$ ;  
 区域间隙宽度:  $CW_i = X_{i+1,1} - X_{i,2}$ ;

- 区域总宽度:  $SW_i = W_i + CW_i$ ;
  - 判断是否干扰的条件为  
 条件 1: 区域间隙宽度大于区域宽度:  $CW_i > W_i$ ;  
 条件 2: 区域总宽度大于某个阈值:  $SW_i > T$ ;
- 第一个条件说明间隙太大,可能是干扰信号。第二个条件是为了排除码号的第一个字符为“1”的情况。当码号的第一个字符为“1”时,可能会满足第一种情况,但总宽度不会太大。这种判断方式,既可以排除干扰,又可以不丢掉有效信息,从而可准确地找到定位区的起点。

最后将确定的定位区起点字符段的  $X_{i,1}$  作为码号的 X 方向的起始坐标:  $X11 = X12 = X_{i,1}$ 。

2.4 码号区右下边界定位

当确定了码号区的起始坐标后,由于同一种 IC 卡的码号区的相对位置是基本不变的,根据预测参数所得的码号区域的属性参数、码号区的总宽度 CodeWidth 可计算出区域的另外两个坐标 X21, Y21 和 X22, Y22。

图 6 示出了最终定位结果的实例图。图中结果已经排除各种干扰,准确地找出码号区的坐标位置。



图 6 IC 卡码号区域定位结果实例图

3 多投影直方图综合分割法

确定码号定位区后,即可进行码号区单字符的分割。对于质量很好,字符无粘连的码号,分割可用密度投影直方图实现。由图 7 可以看出,在字符区比较理想的情况下,直方图中每个字符的边界非常明显,搜索边界即可将单个字符分割出。但在对 IC 卡进行实

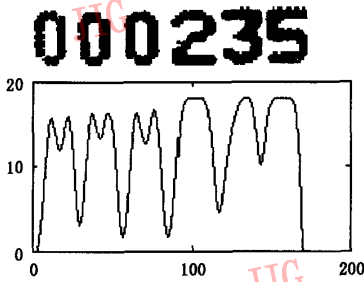


图 7 定位码号区投影密度直方图

际处理时,很多码号区的字符存在粘连,所以仅用密度投影直方图不能完成 IC 卡的字符分割。为此,提出采用多投影直方图综合分割法,即将密度投影直方图、上边沿轮廓投影直方图、下边沿轮廓投影直方图、双边沿投影直方图综合进行码号区的字符分割。

采用多种投影直方图相结合的字符分割方法对

码号区进行字符分割,其方法如下:

先按式(4)求出 Y 方向密度投影直方图  $Ph_1$ ;

再求出 Y 方向上边沿轮廓投影直方图  $Ph_2$ :

$$Ph_2 = (Ph_{2_1}, Ph_{2_2}, \dots, Ph_{2_i}, \dots, Ph_{2_m})$$

其中,  $m$  为码号准确定位区宽度;  $Ph_{2_i}$  为从定位区起点到列向的第一个非零点即上边沿点。

同样可以从下边沿求出下边沿轮廓投影直方图  $Ph_3$ :

$$Ph_3 = (Ph_{3_1}, Ph_{3_2}, \dots, Ph_{3_i}, \dots, Ph_{3_m})$$

根据  $Ph_2$ 、 $Ph_3$  求出双边沿轮廓投影直方图

$Ph_4$ :

$$Ph_4 = Ph_2 + Ph_3$$

这样可求得码号定位区内的 4 种不同的直方图。图 8 示出了一幅粘连码号区的图像和 4 种直方图。

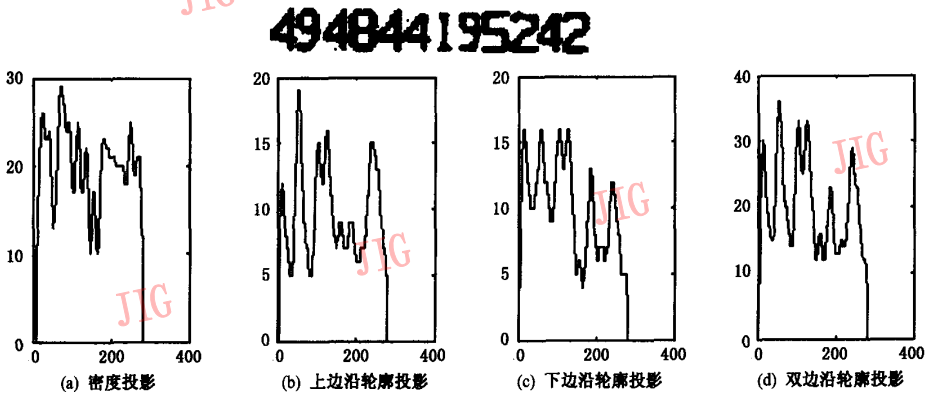


图 8 粘连码号区的不同方式的投影直方图

当分割字符存在粘连时,先根据预测参数求出要分割字符的区间,进行区间定位,然后在定位区间中求 4 条直方图的谷值点,求直方图的谷值点可以借助求曲线极小值点的方法。设用  $Ph_i(X)$  代表同一码号区内的 4 种直方图曲线。那么这些直方图的极小值点应满足:

$$\begin{aligned} \frac{\partial Ph_i(X)}{\partial X} &= 0 \\ \frac{\partial^2 Ph_i(X)}{\partial X^2} &> 0 \end{aligned} \quad (7)$$

其中,  $i=1 \sim 4$  代表 4 种直方图曲线;  $X=1, 2, \dots, m$ ;  $m=X_{21} - X_{11}$  为码号区宽度。

分割字符时,可以先对直方图进行平滑滤波,再根据式(7)求出 4 条曲线在字符坐标区间对应的极小值点<sup>[4]</sup>。在实际分割处理时,由于字符的粘连,有

时可能从一条曲线中不能确切找到极小值点,因而需综合考虑 4 条直方图的结果求字符分割点。

当求到  $t(t \leq 4)$  个极小值点  $Xk_i(i=1 \sim t)$  后,用平均的方法求出字符的分割点

$$Xk = \frac{1}{t} \sum_{i=1}^t (Xk_i) \quad (8)$$

采用多直方图搜索区域谷值的方法进行码号区字符的分割,有效地解决了粘连字符的分割问题。

### 4 测试结果

系统对定位分割方法进行了反复的测试。测试样本为 500 张卡片,每张卡片 3 条码号共 1500 条码号,结果全部正确。

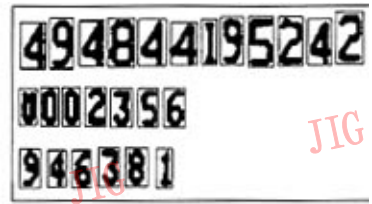
图 9 给一张 IC 卡扫描处理及分割结果示例。



(a) 原始 IC 卡图



(b) 码号定位分割结果



(c) 码号区字符分割结果

图 9 定位分割实例

### 5 结 论

IC 卡的码号检测是 IC 卡质量保证的重要环节之一。而在 IC 卡的码号检测过程中，码号的定位分割直接影响码号识别结果。IC 卡的版面复杂性和版面设计的特殊性及版面质量又给 IC 卡的定位分割带来难度。本文针对 IC 卡的定位问题进行研究，采用预测样卡定位参数的方法确定定位分割参数作为在线检测定位的参考数据，解决了不同 IC 卡中不同位置码号的初定位。将初定位、宽区间定位、区间缩小精确定位综合对 IC 卡码号区进行定位，解决了不同的 IC 卡码号的定位问题。提出多投影直方图综合字符分割法，解决了码号区字符分割问题。这些方法的综合使用，使 IC 卡的定位分割问题得到较好的解决，实验表明定位分割效果良好。

### 参 考 文 献

- 1 杨超峰, 季振洲. “金卡”系统的信息安全性[J]. 电子技术应用, 1996, (3): 18~20.
- 2 张世栋, 郑永清. IC 卡系统数据一致性研究[J]. 计算机应用, 2002, 22(8): 12~13, 16.
- 3 郑南宁, 刘健勤. 基于区域特征的自适应图像分割方法[J]. 电子学报, 1995, 23(7): 98~101.
- 4 章毓晋. 图像处理与分析[M]. 北京: 清华大学出版社, 1999: 195~204.



郑胜林 1951 年生。副教授、系主任。1983 年于重庆大学获光电精密仪器专业硕士学位。主要研究方向为图像信息技术、模式识别。  
E-mail: zsliner@21cn.com



潘保昌 1949 年生。教授、博士生导师、所长。1990 年于重庆大学获光电精密仪器专业博士学位。主要研究方向为图像信息技术、模式识别。



徐 杜 1958 年生。教授、院长。1988 年于重庆大学获光电精密仪器专业硕士学位。研究方向为计算机技术与智能控制。



蒋永平 1956 年生。高级工程师。1985 年于重庆大学工业自动化专业本科毕业。研究方向为计算机技术与应用。