

基于图像技术的 EAN-13 条码识别系统设计*

李 慧, 欧阳鑫玉, 张 锋

(辽宁科技大学 电子与信息工程学院, 辽宁 鞍山 114051)

摘 要: 条形码技术作为一种数据采集和自动识别技术应用非常广泛。为了更好地满足应用场景区的自动化功能需求, 利用图像处理相关技术, 对 EAN-13 条形码识别的相关算法进行了研究。首先, 针对采集到的条形码图像进行图像预处理, 包括图像灰度化、基于 Hough 变换的图像倾角矫正、目标分割和图像滤波; 然后根据条形码编码的既定规则, 采用平均值法和校验码验证法对条形码进行译码; 最后, 利用 MATLAB 仿真软件, 设计了一个条形码识别系统。通过实验验证, 所设计的系统识别率可以达到 99% 以上, 取得了很好的效果。

关键词: 条码识别; 图像技术; 系统设计; EAN-13

中图分类号: TP311

文献标识码: A

DOI: 10.19358/j.issn.2096-5133.2020.03.014

引用格式: 李慧, 欧阳鑫玉, 张锋. 基于图像技术的 EAN-13 条码识别系统设计[J]. 信息技术与网络安全, 2020, 39(3): 73-77.

Design of EAN-13 barcode recognition system based on image technology

Li Hui, Ouyang Xinyu, Zhang Feng

(School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan 114051, China)

Abstract: As a common data acquisition and automatic recognition technology, barcode technology has a wide range of application. In order to better meet the automation requirements of application scenarios, this paper studies the algorithms of EAN-13 barcode recognition by using image processing technology. Firstly, the collected barcode image is preprocessed, including image graying, image inclination correction based on Hough transform, target segmentation and image filtering; secondly, according to the coding rules of barcode, the barcode is decoded by mean value method and check code error correction processing; finally, based on MATLAB software, a barcode recognition system is designed. Through the experimental verification, the recognition rate of the designed system has reached more than 99%, and achieved good results.

Key words: barcode recognition; image technology; system design; EAN-13

0 引言

随着社会经济的快速发展, 条形码技术已成为一种常用的数据采集和自动识别技术^[1-3], 由于它具备输入速度快、准确率高、投入成本低、简单实用、自由度大、采用设备相对简单、容易制作等特点, 因此能够普遍应用在商品流通、工业生产、交通运输、智慧物流、医疗卫生等领域。

对于条码的识别, 通常是基于光电技术进行识别, 一般的光电识读设备通过线性扫描的方式读取到条码信息后, 再利用硬件进行条形码的识读。由

于条码可能存在印刷不清或者由于脏污形成噪声等情况, 会导致识别率降低; 此外, 识读设备成本过高, 采用人工手持条形码识别装置工作效率低, 难以适应现代智能化、自动化的应用场景功能需求。随着数字图像处理技术的发展, 采用图像处理技术和相关的智能算法对条形码进行识别已成为新的趋势, 它可以有效解决利用硬件识别带来的技术和成本问题。

本文正是在此背景之下, 根据实际应用需求, 利用图像处理相关技术, 对 EAN-13 一维条形码的识别进行了研究, 并利用 MATLAB 软件设计出识别系统, 克服了基于光电技术的条码识别带来的局限性。

* 基金项目: 辽宁科技大学大学生创新创业训练计划项目 (201910146117)

1 EAN-13 条形码

EAN-13 条形码是欧洲物品编码的一种终端产品条形码协议和标准,它的左侧是空白区,然后依次是起始符、左侧数据符、中间分隔符,接着是右侧数据符、校验符、终止符以及右侧空白区和供人识别的字符(如图 1 所示),应用比较广泛。

EAN-13 条形码是一种 (7,2) 码^[4-5],包含了 13 位数据字符,其中第 1~3 位字符代表国家代码,第 4~7 位字符代表厂商代码,第 8~12 位字符代表产品代码,第 13 位字符是校验码;每个字符由两个条和两个空交替组成,每一条或空由 1~4 个模块组成,每一数字总宽度为 7 个模块,分别用“0”和“1”表示条形码的空和条的模块。



图 1 EAN-13 码的结构

EAN-13 条形码字符集包括 0~9 共 10 个数字字符,条码的二进制表示形式由 A、B 和 C 三个子集(见表 1)确定。条形码中的前置码不用字符表示;右侧的数据符和最后的校验符全部用 C 子集来表示。前置码的数值决定左侧数据符是选用 A 子集还是选用 B 子集来表示,表 2 列出了左侧数据符的字符集选择规则。

表 1 条形码字符集

数字字符	A 子集	B 子集	C 子集
0	0001101	0100111	1110010
1	0011001	0110011	1100110
2	0010011	0011011	1101100
3	0111101	0100001	1000010
4	0100011	0011101	1011100
5	0110001	0111001	1001110
6	0101111	0000101	1010000
7	0111011	0010001	1000100
8	0110111	0001001	1001000
9	0001011	0010111	1110100

表 2 左侧数据符的字符集选择规则

前置码	左 1	左 2	左 3	左 4	左 5	左 6
0	A	A	A	A	A	A
1	A	A	B	A	B	B
2	A	A	B	B	A	B
3	A	A	B	B	B	A
4	A	B	A	A	B	B
5	A	B	B	A	A	B
6	A	B	B	B	A	A
7	A	B	A	B	A	B
8	A	B	A	B	B	A
9	A	B	B	A	B	A

校验符是用来检验前面的字符数字是否正确,计算公式如下:

$$K = \sum_{i=1}^6 T_{2i-1} + 3 \sum_{i=1}^6 T_{2i} = a_2 \times 10^2 + a_1 \times 10 + a_0 \tag{1}$$

$$L = \begin{cases} 10 - a_0 & a_0 \neq 0 \\ a_0 & a_0 = 0 \end{cases} \tag{2}$$

其中, T_i ($i = 1, 2, \dots, 12$) 表示条形码的第 1~12 位数字值, L 为校验符的值。

以 EAN-13 条形码 9787512426399 为例,其前置码为 9,根据表 2 可知,其左侧数据符的字符集为 ABBABA。根据校验码的计算方法,可知 $K = 131$, $a_0 = 9$,因此校验码是 9。又根据表 1 可知,该条形码后 12 位数的二进制码分别为:0111011、0001001、0010001、0110001、0110011、0010011、1011100、1101100、1010000、1000010、1110100、1110100。

2 条码图像识别方法

条码图像识别处理流程如图 2 所示。

条形码图像采集后,需要对图像进行预处理,即进行图像灰度化、倾角矫正、目标分割、滤波等操作,然后针对分割的条码图像,根据译码规则进行译码识别。

2.1 条形码图像预处理

(1) 条形码图像的灰度化

通过各种摄像设备采集到的条形码图像,大多是彩色图像,采用的是 RGB 颜色模式,即用红、绿、蓝三种原色来表示,各个颜色分量的取值范围都是

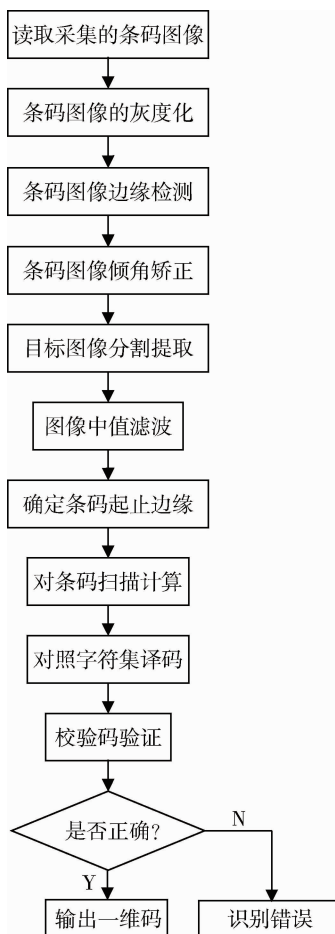


图2 条形码图像识别流程

0~255, 因此, 对于一个 RGB 图像, 在计算机中需要用一个三维矩阵来表示, 但实际上 RGB 并不能反映图像的形态特征, 只是从光学的原理上进行了颜色的调配。在进行图像处理时, 通常为了减少所要处理的运算量, 一般采用将 RGB 图像进行灰度化的方法。图像灰度化处理方法通常有分量法、最大值法、平均值法和加权平均法等。本文采用了平均值法进行灰度化, 即将分量 R、G、B 的值求平均得到一个灰度值, 处理效果如图 3 所示。图像灰度化之后, 在计算机中只需要用一个一维矩阵来表示, 大大提高了后续处理速度。



图3 条码图像灰度化前后对比

(2) 条形码图像的倾角矫正

对于采集到的图像, 由于条形码和摄像头的角度问题, 使得条形码图像和直角坐标系有一定的倾斜角度, 为了方便后续处理, 通常需要先计算出图像的倾斜角度, 然后通过旋转相应角度对图像进行矫正。

矫正图像倾斜角度的方法有两点法、最小二乘法、Hough 变换法以及 Radon 变换法等^[6-8]。两点法、最小二乘法计算量小, 但抗干扰能力差, 不能计算有断点的直线倾角。Hough 变换法和 Radon 变换法具有很好的鲁棒性, 对于抗噪声干扰能力强, 精度比较高, 但是运算量比较大。根据这种情况, 本文采用 Hough 变换法。

Hough 变换法是一种变换域直线提取法, 由 Paul Hough 于 1962 年首次提出。Hough 变换的基本原理是: 对于直角坐标系中的直线方程 $y = kx + b$, 可以在极坐标中用 $\rho = x \cos \theta + y \sin \theta$ 来表示, 因而这条直线在极坐标中表示的是一点 (ρ, θ) 。对于直角坐标系下呈现直线的所有点, 它们的斜率和截距均相同, 因而它们在极坐标系下对应于同一个点, 这样就把直线的提取转换成了点的计数而使问题得到解决。

本文在采用 Hough 变换法时, 进行了适当改进, 即考虑到条码图像中有大量的平行直线, 在处理时只选取含有条码的小块进行计算处理, 然后根据得到的倾斜角度, 再将整个图像旋转相应角度进行矫正, 这样既保证了精度, 也减少了运算量。利用 Hough 变换对条码图像进行倾角矫正后的效果如图 4 所示。



图4 利用 Hough 变换进行图像倾角矫正前后对比

(3) 条形码图像的分割及滤波

针对矫正后的条码图像, 采用水平和垂直两种投影方法, 检测出条形码图像区域, 并将其从原图像中分割出来。

同时考虑到图像本身或受污染可能存在噪声的情况, 需要对条码图像先进行去噪声处理。常用

的图像去噪方法有高斯滤波、中值滤波、维纳滤波、小波去噪、P-M 方程法以及 TV 法等。本文考虑条形码图像中存在孤立噪声多的状况,选择中值滤波法进行去噪。

中值滤波器作为一种非线性平滑滤波器经常被使用,它的基本原理是把数字图像中的某点的值用该点的一个邻域中各点值的中值替换,这样能够让四周的像素灰度值差比较大的像素改取与周围的像素值近似的值,从而达到去除孤立噪声点的目的,因此中值滤波方法对于滤除图像中的孤立噪声十分有效。

对条形码图像进行目标分割和中值滤波后的效果如图 5 所示。

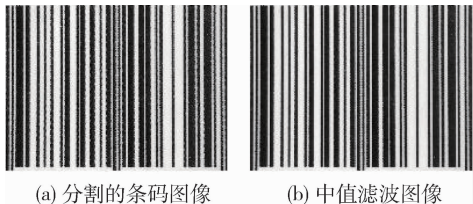


图 5 分割的条码图像和中值滤波效果

(4) 条形码图像的二值化

二值化就是将灰度图像上的各点的灰度值转换成为 0 或 255 两个值,使整个图像呈现出明显的黑白两色。对于灰度图,像素点有 256 个灰度等级,经过选择适当的阈值将灰度两极化后,二值化图像依然可以很好地反映图像的特征。

根据上述所得的条码图像的滤波图像,设定一个阈值 R ,若某点像素灰度值大于 R ,则令该点像素值为 255;若某点像素灰度值小于 R ,则令该点像素值为 0,即可将图像二值化。假如图像像素点的灰度值已经完成归一化,则在 0~1 之间确定阈值。通过反复测试,本文选择阈值为 0.7~0.74。针对所得条形码图像的滤波图像进行二值化后,得到的二值图像如图 6 所示,能够看出,目标图像和背景更加分明。

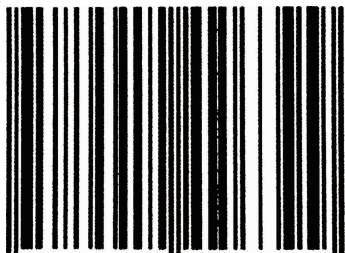


图 6 二值化后的条码图像

2.2 条形码译码

条形码的译码过程主要分为边缘确定、译码、校验和显示结果。

(1) 边缘确定

对于条码图像,首先要确定条形码扫描方向,这里采用从左至右进行扫描,然后对图像进行垂直投影,根据投影结果确定起止条码的左右边缘。

(2) 译码

对于条形码的译码,这里采用平均值法,即依据起止边缘,计算出条码符号图像中整个条码的宽度,然后用这个宽度除以总模块数 95,就可以得到每个模块总像素的宽度。再依据每个字符含有的模块数,参考前述表 1 的字符集进行译码。

(3) 校验

条码字符译码完成后,需要确定得到的译码是否正确,利用前述的校验码计算方法,可以对其校验验证。

(4) 显示结果

通过校验码验证后,如果正确,则输出识别到的条码数字;否则,输出“识别错误”。

3 EAN-13 条形码识别系统实现

基于 MATLAB 仿真软件,利用其图形用户界面工具包 GUI 进行了 EAN-13 条形码识别系统设计,得到的系统界面如图 7 所示。



图 7 EAN-13 条形码识别系统界面

在本系统中,用户需要先将采集的条形码图像存放到一个文件夹中。在运行程序后,用户可以对采集的图像进行选择和导入,利用该系统的相关控件,可以显示原图以及部分过程处理效果图,并能将最终的识别结果显示出来。

经过对采集的 620 张条形码图片的处理实验,正确识别 614 张,其他 6 张因为图片污染严重或图

像采集时造成大的扭曲未能正确识别,系统识别准确率达到了 99% 以上,取得很好的效果,达到了工业应用所需的识别率要求。该系统也可以利用其他高级语言进行移植后供实际应用。

4 结论

本文针对一维条形码的识别问题,基于图像处理技术,提出了从图像预处理到条形码译码的系列方法,并运用 MATLAB 仿真软件,设计出了一个条形码的识别系统,克服了基于光电技术的条码识别带来的局限性,实验验证了本文采用的处理方法和设计的系统的有效性,为条码自动识别的实际应用提供了基础和参考。

参考文献

- [1] 刘艳洋,曹玉东. EAN-13 条形码图像的识别[J]. 计算机技术与发展,2015,25(6):202-205.
- [2] 赵琪,高美欢. 基于图像的工业在线条码识别算法研究[J]. 软件导刊,2018,17(8):125-128.
- [3] 张志凌,梁潇,杨天宇,等. 对图像中条形码进行定位识别的应用[J]. 电脑知识与技术,2019,15(14):140-142.
- [4] 龚亚欢,王超,胡晨. EAN-13 码的图像识别系统设计与

实现[J]. 现代电子技术,2009,32(22):106-109.

- [5] 王雅静. 基于图像处理的 EAN-13 条码识别算法[J]. 山东理工大学学报(自然科学版),2005,19(4):14-18.
- [6] 赵楠楠,欧阳鑫玉,吴庆洪. 线性回归与 Hough 变换在计算图像倾斜角度中的应用[J]. 鞍山科技大学学报,2003,26(6):457-460.
- [7] 梁添才,皮佑国,彭晶,等. 基于 Hough 列车客运票图像倾斜校正[J]. 华南理工大学学报,2007,35(5):35-39.
- [8] 陈超,杨柳. 基于数字图像处理的条形码识别预处理算法研究[J]. 河南工程学院学报(自然科学版),2015,27(3):69-73.

(收稿日期:2019-12-31)

作者简介:

李慧(1998-),女,本科生,主要研究方向:图像处理。

欧阳鑫玉(1974-),通信作者,男,博士,教授,硕士生导师,主要研究方向:智能控制、机器学习。E-mail:13392862@qq.com。

张锋(1996-),男,硕士研究生,主要研究方向:图像处理。

版权声明

经作者授权，本论文版权和信息网络传播权归属于《信息技术与网络安全》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《信息技术与网络安全》编辑部
中国电子信息产业集团有限公司第六研究所