

针对特定目标提取的二值化算法

李良华 罗彬杰

(西南财经大学成人(网络)教育学院 成都 610074)

摘要 根据中文支票识别的预处理过程中提取特定目标的需要,研究了多种二值化算法在预处理中的效果。通过分析 2000 张中文支票灰度图像的直方图,找到了可以用于图像分割的直方图梯度值信息,基于该梯度值信息,提出了一种用于提取支票图像中金额栏外围框线的二值化算法,使得支票灰度图像中的待提取框线更加清晰、凸出,更加易于定位。在中文支票预处理环境下,所提出的二值化算法在与其它多种二值化算法的对比测试中,表现出了更好的效果和更高的效率。

关键词 中文支票识别,二值化,直方图,梯度值

New Binarization Algorithm for Given Object Extraction

LI Liang-hua LUO Bin-jie

(School of Adult(Network) Education, Southwestern University of Finance and Economics, Chengdu 610074, China)

Abstract In order to meet the requirements of preprocessing of Chinese cheque image recognition, this paper studied several binarization algorithms. After analyzing the histograms of gray-level images of 2000 cheques, we found a clue, maximum gradient, in histogram of gray-level cheque image, which can be used for segmenting image and extracting given object parallel-line of amount field. With the clue, we developed a new binarization algorithm which makes the parallel-line of amount field more obvious and easier to be located in binarization image. Comparing with several other commonly used binarization algorithms, the algorithm proposed by the paper has been proven to be more feasible and advanced in the simulating tests.

Keywords Chinese cheque image recognition, Binarization, Histogram, Gradient

二值化是图像分割中的一类常用技术。有关图像分割技术的更多信息可以在文献[1,2]中找到。二值化的本质是使用一个最优化的阈值对图像中的像素点进行分类。当前已知的二值化算法可以分为以下两类:基于全局信息的算法和基于局部信息的算法。基于全局信息的算法使用单一阈值对图像进行分割。基于局部信息的算法通过基于局部信息的计算为每个像素或几组像素给出不同的阈值。在文献[3]中,作者对 20 种基于全局信息的算法进行了比较,利用一致性和形状两个定量指标给出了各个算法的测试结果。在文献[4]和文献[11]中可以看到有关 11 种基于局部信息算法的对比测试。文献[11]提出了一种“面向目标”的二值化算法评价准则。在不同的应用环境中,处理对象和二值化的目的都不尽相同,因此,在特定应用环境中,工程人员需要选择最适合的二值化算法。根据文献[3,4,11-13]的测试结果来看,在不同的条件下,最优的二值化算法一般是不同的。

在支票识别应用中,识别的目标之一就是金额栏中的数字。金额栏的定位、提取,作为获取待识别数字的基础,依赖于对金额栏外围平行框线的检测、定位,而检测效率的高低、定位的准确度对于后期的处理都是极为重要的。因此,一个好的二值化算法应当尽可能地使得金额栏的外围平行框线在二值化结果中凸显出来,同时,尽量减少其它无关物体在二值

化结果中的像素数量。通过对 2000 张中国现行支票灰度直方图的统计分析,本文提出了一种针对特定目标提取的算法,对该算法的详细描述见本文第 2 节,对比测试结果请见本文第 3 节。

1 针对特定目标的二值化

1.1 问题描述

提取金额栏由两步完成:二值化和平行直线检测。待定位、提取的目标物体就是金额栏的平行直线。在一张支票图像中还包括很多其它的物体(印章、手写体数字等),这些物体不是本阶段所需要处理的目标,所以可以将这些物体视为背景物体。在二值化图像中,这些物体所属像素点应当赋值为 0。金额栏的平行直线,作为目标物体,应当被看作前景物体,在二值图像中赋值为 1。

在整个图像的二值化过程中,将过多的物体标记为前景物体,会降低直线检测的效率,甚至导致无法检测出金额栏的平行直线。与此相反,如果将过多的物体标记为背景物体,金额栏的平行直线有可能部分缺失,无论直线检测算法的鲁棒性如何,严重的缺失同样会导致直线检测的失败。关于支票的整体图像及金额栏的平行直线在支票图像中的具体位置如图 1 所示;提取出的金额栏如图 2 所示。

到稿日期:2009-01-20 返修日期:2009-03-20

李良华(1963-),男,博士生,副研究员,主要研究方向为电子商务、支付系统、图像处理,E-mail:lilianghua@swufe.edu.cn;罗彬杰(1984-),男,博士生,主要研究方向为电子商务、图像处理、数据挖掘。



图1 金额栏的平行直线



图2 提取后的金额栏

1.2 问题分析

1)如何评价二值化效果

文献[4]提出了以下5条准则来评价二值化结果的好坏:

- a)断裂直线结构;
- b)断裂物体、文字等;
- c)模糊直线、物体及文字;
- d)物体整体缺失;
- e)同类区域中的噪音。

平行直线的检测算法是基于霍夫变换的,而霍夫变换本身能够对虚线、断裂直线进行检测、定位。为了给直线检测提供一个较好的二值化基础,有些物体应当尽量被标记为背景,而这与准则 d)发生冲突。因此,本文使用了文献[6]提出的评价方法。

在支票数字化图像中,肉眼所见的一条直线通常是由多条单像素直线所构成的,如图3所示。为了更好地描述直线检测结果,有必要从多条单像素直线中确定一个“主线”作为其唯一的代表,现给出主线的定义:

给定指数 I ,其值由 $I=ML+DL$ 计算,其中 ML 表示一条单像素直线所有像素点的平均灰度值, DL 表示一条单像素直线所有像素点的灰度值的标准差。具有最小 I 值的单像素直线是主线。

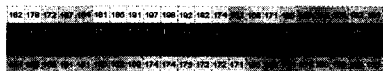


图3 灰度数字图像中的一条直线

基于“主线”的定义,现给出二值化结果的评测指标 R ,由式(1)所定义:

$$\frac{PL}{PO} \times 100\% = R \quad (1)$$

其中, PL 表示二值化结果中像素值为1的主线像素数量, PO 表示二值化结果中像素值为1的所有像素点的数量。

2)对其它二值化算法的分析

相对于基于全局信息的二值化算法,基于局部信息的二值化算法一般能够提供更好的二值化效果,但是,基于局部信息的二值化算法通常也需要更多的处理时间。在文献[4]中,作者证明了 Yanowitz-Bruckstein's 算法[8]相对于其它二值化算法处理效果更好,尤其是在光照不均的情况下。但是,在相同的实验条件下,Yanowitz-Bruckstein's 算法[8]所需的处理时间大约是 Otsu 算法[5]的 327 倍。在中文支票识别应用中,支票图像是由专用的扫描仪生成的,因此一般不存在光照不均的情况。Iterative Multimodel Subimag (IMS)算法[9],是一种针对支票图像提取手写体数字的二值化算法,但是,它所基于的支票图像与我国现行支票有很大的不同,票面情况如图4所示,同时,它所提取的对象也与本文所要提取的对象不同。考虑到以上因素,在本文的应用环境中,一种快速的、有

效的基于全局信息的二值化算法是较为合适的选择。



图4 文献[9]中的支票图像样例

2 本文提出的算法

2.1 2000 张中文支票图像的直方图分析

我们从北京邮电大学提供的支票图像数据库中,选择了2000张支票图像(95 dpi,256 gray-level)。使用文献[14]提出的算法对每张支票的灰度直方图进行了“双峰”检测,结果如表1所列。

表1 “双峰”检测结果

测试图像数量	“双峰”直方图数量	非“双峰”直方图数量
2000	1	1999

从表1的测试结果来看,大部分的直方图并不符合“双峰”假设,因此,使用基于“双峰”假设的全局二值化算法,例如 Otsu 算法[5],是不恰当的。在本文第3节的算法对比测试中,我们选择了并不基于“双峰”假设的两种算法,Minimum Square Error(MSE)算法[6]和 Histogram Concavity Analysis(HCA)算法[7];同时,由于 Maximum Entropy(ME)算法[10]作为 Entropy 一类二值化算法的代表也参与了算法对比测试,更多的关于 Entropy 一类算法的讨论可以在文献[13]中找到。现定义:

$$m(i) = \frac{\sum_{n=1}^{2000} h_n(i)}{2000} \text{ 和 } s(i) = \sqrt{\frac{\sum_{n=1}^{2000} (h_n(i) - m(i))^2}{2000 - 1}}$$

其中, $h_n(i)$ 表示第 n 张支票的灰度值为 i 的像素点数量。

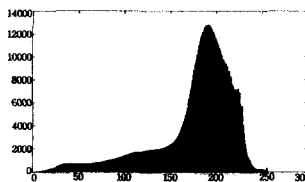


图5 $m(i)$ 的直方图

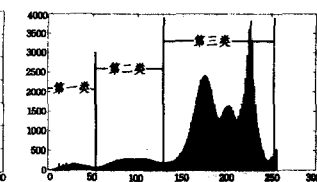


图6 $s(i)$ 的直方图

通过对图5的观察我们很难发现任何有价值的信息,但是,从图6可以直观地发现以下几点:

- a)支票图像的直方图大致由3类物体构成(已经标记在图6中);
- b)某些灰度值的像素数量较为稳定,例如,图6中第一类物体的和第二类物体的;
- c)金额栏的平行直线主要分布在第二类物体中。

现在假设图6中的每一类物体的像素数量都服从正态分布,利用矩估计获得正态分布的参数,分别对图5按照最小方差准则进行正态分布拟合,拟合图像如图7、图8、图9所示(拟合曲线由“+”线表示)。

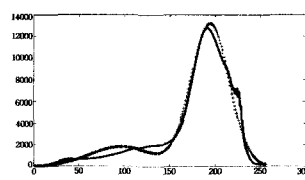


图7 两峰正态分布拟合结果

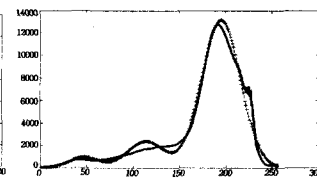


图8 三峰正态分布拟合结果

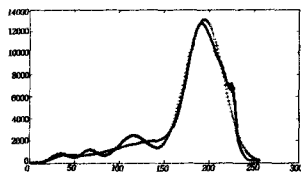


图9 四峰正态分布拟合结果

从表2所列的拟合残差平方和来看,三峰正态分布的拟合的残差平方和最小,证明了图6的分类假设是合理的。

表2 正态分布拟合结果

两峰拟合残差平方和	三峰拟合残差平方和	四峰拟合残差平方和
101071808.787449	86565669.8686561	87679000

2.2 本文提出的算法

基于本文2.1节中的分析,我们认为理想的中文支票的灰度直方图应当如图10所示,A点和B点之间的高度就是直方图的最大梯度值,而利用最大梯度值则可以确定分割阈值A点所对应的灰度值。

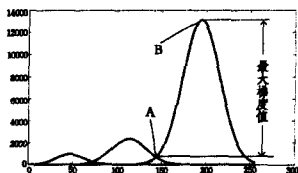


图10 直方图最大梯度值示意图

本文提出的算法描述如下:

$$mt(i) = \max_{j=1,2,3,\dots,ms} [h(i+j) - h(i)] \quad (2)$$

$$T^* = \operatorname{argmax}_{i=1,2,\dots,255-ms} mt(i) \quad (3)$$

其中, $h(i)$ 表示灰度值为 i 的像素点数目, $ms = \left\lfloor \frac{256}{3} \right\rfloor$ 是根据经验确定的最大阈值, T^* 就是最佳分割阈值。

3 算法对比测试与分析

从支票数据库中选择了2000张不需要旋转矫正的支票作为算法对比测试的样本。

在不同的二值化基础上,进行金额栏平行直线检测的精度如表3所列。表3中的“线间距”表示检测结果与主线(由1.2节所定义)之间的距离(一个像素作为单位距离)。

表3 直线检测精度测试结果

线间距	Otsu ^[5]	MSE ^[6]	HCA ^[7]	ME ^[10]	本文算法
0	638	584	703	643	643
1	923	832	891	859	988
2	325	507	283	396	369
3	105	70	85	92	0
≥ 4	9	7	38	10	0

其它用于评价结果的指标如表4所列。其中 R 值由1.2节中的式(1)所定义。

表4 其它评价指标

算法名称	Otsu ^[5]	MSE ^[6]	HCA ^[7]	ME ^[10]	本文算法
平均 R 值	1.48%	2.25%	1.34%	1.77%	2.59%
平均阈值	146	118	154	137	111
二值化平均所耗时间(秒)	0.078	0.075	0.076	0.832	0.087
直线检测平均所耗时间(秒)	21	15	23	17	12
平均线间距	0.96	1.04	3.27	0.98	0.86

根据表3和表4的检测结果,可以得出以下几点:

a)本文提出的算法取得了最大的平均 R 值,表明本文算法所提供的二值化基础中,金额栏平行直线所占比例是最大的,达到了凸出目标物体的效果。

b)本文提出的算法由于目标物体凸出,因此直线检测平均所需时间最少。

c)在本文算法所构造的二值化基础上进行直线检测的平均线间距最小,精度最高。

结束语 本文的主要贡献在于,通过对中文支票灰度直方图的统计分析,发现了目标物体在直方图中的大致位置,并以此设计了基于直方图最大梯度值的二值化算法。该算法在中文支票识别的应用环境中较好地保留了目标物体,减少了其它物体在二值化结果中的像素数量,为金额栏的定位、提取提供了一个良好的二值化基础。无论从效率上还是从效果上来看,本文的算法都较好地满足了支票识别应用中第一步的需要。

参考文献

- [1] Medioni G, Lee M S, Tang C K. A Computational Framework for Segmentation and Grouping[M]. Elsevier, 2000
- [2] Zhang Y J. New advancements in image segmentation for CBIR [M]. Encyclopedia of Information Science and Technology, Idea Group Reference, Mehdi Khosrow-Poured, 2005, 4: 2105-2109
- [3] Sahoo P K, Soltani S, Wong A K C, et al. A survey of thresholding techniques[J]. Comput. Vis. Graph and Image Processing, 1988, 41: 233-260
- [4] Trier Ø D, Taxt T. Evaluation of Binarization Methods for Document Images[J]. IEEE Trans. Pattern Analysis and Machine Intelligence, 1995, 17: 312-315
- [5] Otsu N. A threshold selection method from grey - level histograms[J]. IEEE-SMC, 1979, 9: 62-66
- [6] Ramesh N, YOO J-H, Sethi I K. Thresholding based on histogram approximation[J]. IEE Proc -VisImage Signal Process, 1995, 142(5): 271-279
- [7] Rosenfeld A, Torre P. Histogram concavity analysis as an aid in Threshold selection[J]. IEEE-SMC, 1983, 13: 231-235
- [8] Yanowitz S D, Bruckstein A M. A new method for image segmentation[J]. Computer Vision, Graphics and Image Processing, 1989, 46: 82-95
- [9] Dawoud A, Kamel M. Iterative Multimodel Subimage Binarization for Handwritten Character Segmentation[J]. IEEE Trans. Image Processing, 2004, 13: 1223-1230
- [10] Pun T. A new method for gray-level picture thresholding using the entropy of the histogram[J]. Signal Processing, 1980, 2: 223-237
- [11] Trier Ø D, Jain A K. Goal - Directed Evaluation of Binarization Methods[J]. IEEE Trans. Pattern Analysis and Machine Intelligence, 1995, 17(12): 1191-1201
- [12] Sezgin Z, Sankur B. Survey over image thresholding techniques and quantitative performance evaluation[J]. J. Electron. Imag., 2004, 13 (1): 146-165
- [13] Chang C-I, Du Y, Wang J, et al. Thoun Survey and comparative analysis of entropy and relative entropy thresholding techniques [J]. IEE Proc. -Vis. Image Signal Process., 2006, 153(6)
- [14] Chow C K, Kaneko T. Automatic detection of the left ventricle from cineangiograms[J]. Computers and Biomedical Reearch, 1972, 5: 388-410