改进的 EM 算法在分块灰度图像二值化中的应用

王红霞 程艳芬

(武汉理工大学计算机学院 武汉 430063)

摘要:识别扫描得到的文本图像首先要进行二值化处理,以去掉冗余的信息.针对二值化过程中采取固定阈值会导致图像的背景和文本不能得到有效区分的问题,提出改进的分块 EM 算法,并应用在文本图像的二值化过程中,有效地分离了图像中的文本和背景信息,使识别的后续阶段提取的特征能够更好地代表要识别的模式.

关键词:模式识别;OCR;图像二值化;EM 算法;分块图像

中图法分类号:TP391.41

DOI: 10. 3963/j. issn. 1006-2823. 2011. 04. 021

0 引 言

文本图像的识别是模式识别领域的一个重要应用分支,而在整个 OCR(optical character recognition)领域中,最为困难的就是脱机自由手写字符的识别.

一幅通过扫描得到的手写体文本图像,无论 是彩色图像还是灰度图像,各像素因为颜色深浅 不一,于是在取值范围内会取不同的值.例如一个 彩色图像中某个像素(R,G,B),其R,G,B的取 值范围均在(0,255)之间,这种情况给文本图像的 处理带来了很大的困难. 因为文本图像的处理只 需要知道哪是文字,哪是背景,以及文字是什么, 不需要知道它们对应像素点颜色上的差异. 只有 通过对这种文本图像实施二值化后,系统根据像 素的灰度值相应地处理成黑、白两种颜色,图像中 各像素点的值由 0~255 灰度值转化为 0,1 两个 值中的一个,图像信息于是由灰度转化为黑白,其 信息量大大减少,从而使得文本的特征更集中,便 于图像处理[1]. 本文以脱机手写体阿拉伯字符的 识别研究为背景,将改进后的 EM 算法应用在分 块的脱机阿拉伯手写体文本图像识别的预处理过 程中,实验结果表明效果良好.

1 手写体文本图像的二值化

预处理是文字识别的第一步,在实际识别系统中是一个很重要的阶段,该阶段与特征抽取阶段也是紧密相连的.良好的预处理可以有效地保持图像信息,二值化后的文本图像内仅含黑、白二色的信息,在它们之间不存在其他的灰度变化,从而更清楚地反映文本图像中字符的本质特征,使得后续阶段提取的特征能够更好的代表要识别的模式[2].

由于脱机手写体文本图像的识别只需要处理 图像中的字型信息,对颜色等信息不作处理,所以 对扫描得到的文本图像要进行二值化(Binarization)处理,以去掉冗余的信息.和灰度化相似,图 像的二值化也有很多成熟的技术,但却没有一种 方法能对任何目标对象都普遍适用,必须根据具 体的处理对象而定.

二值化的方法根据其运算的范围不同,可分为全局阈值法和局部比较法.全局阈值法根据文本图像的直方图或灰度的空间分布确定一阈值,并根据此阈值实现灰度文本图像到二值化文本图像的转化.全局的阈值选取是根据整幅图像确定一个阈值,对输入图像的量化噪声或不均匀光照等情况抵抗能力差,应用受到极大的限制.局部阈

值选取方法是将图像划分为若干子图像,根据每个子图像确定相应的阈值,这种方法通过定义考察点的邻域,并由邻域计算模板来实现考察点灰度与邻域的比较,较全局方法有更广泛的应用.其中比较典型的方法有 Kamel-Zhao 算法和 Bernsen 算法.但局部比较法也存在缺点和问题,如实现伪影(ghost)等现象(即在背景区域受噪声干扰得到笔划结果).全局阈值选取方法对噪音比较敏感,因此应用中一般采用局部阈值选取方法.但是这二者并无本质的不同[3].

设文本图像中像素点(x,y)的灰度值为 f(x,y),f(x,y)的取值范围是 $0\sim255$,阈值采用下式确定 $\lambda=(f(x,y)_{\max}+f(x,y)_{\min})/3$,且使

$$f(x,y) = \begin{cases} 0 & f(x,y) < \lambda \\ 255 & f(x,y) \geqslant \lambda \end{cases}$$
 (1)

其中为表达方便,255的值一般用"1"值表示[4].

从式(1)可以看出,上述确定阈值方法一个明显的 弊端是,单纯地以 2 个点的灰度值: $f(x,y)_{max}$, $f(x,y)_{min}$ 作为代表来求整个图像的阈值过于片面,于是提出将期望值最大法(expectation-maximization,EM)算法^[5-6]的思想用在灰度图像的二值化中,可以有效地避免这种以偏概全导致的误差.

2 用改进的 EM 算法实现分块灰度 图像的二值化

文本图像二值化的关键是要找到合适的阈值 T来区分目标和背景. 阈值判定法利用了图像中 所要提取的目标物(即前景)与其背景在灰度特性 上的差异,把图像视为具有不同灰度级的两个区域的组合,通过选取阈值,将目标区域从背景中分 离出来. 所以阈值 T 的选择是关键,选得好,可以 很好地将图像中的文字和背景分离开;选的不好, 可能会造成待识别文字信息的丢失,导致误识率 增高.

如图 1 所示,其中图 a)显示的是一幅待识别的阿拉伯手写体文本图像,图 b)是图像中一个块的直方图,从直方图可以看出,图像的前景(即文本)和背景的像素点形成了 2 个类,Background Peak 和 Text Peak 分别是背景和目标的"峰",th是这 2 个类的"谷".由此,灰度图像二值化对 EM算法来说就是一种比较特殊的情形,那就是类的数目 2 是已知的:代表背景的类(用 0 表示)和代

表目标的类(用1表示).

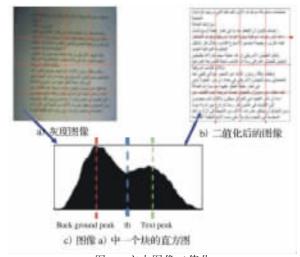


图 1 文本图像二值化

其具体算法如下.

- 1) 数据准备 对于一个宽为 W, 高为 H 的 文本图像 I(x,y), 0 < x < W, 0 < y < H, 用一维数组 $D[W \times H]$ 来表示, 并且 $0 \le D[i] < 255$, $0 < i < W \times H$.
- 2) 初始化 给两个类的期望赋初值为 E[1], E[2]. 可以随机赋值,但是初值的好坏对收不收敛以及收敛的速度都有很大的影响. 初始值的选取也会影响算法的稳定性,如果初始值选取的比较好,算法相对稳定. 于是设 D_{\min} , D_{\max} 和 D_{ave} 分别是数组 D的最小值、最大值和平均值,并且取

$$E[1] = \frac{D_{\min} + D_{\text{ave}}}{2}$$
 $E[2] = \frac{D_{\max} + D_{\text{ave}}}{2}$

 $Error \lceil 0 \rceil = 0;$

3) 对数组 D 中的每个元素D[i]计算和 E[1],E[2]的距离为

$$d[i] = \arg\min_{k} |D[i] - E[k]|$$

其中:k=1,2;和 E[1], E[2]哪个值近就归为哪一类,由此得到两个一维数组 D_1 , D_2 ,同时统计出 D_1 , D_2 2 数组的大小 L_1 , L_2 .

$$\mathit{Error}[1] = \frac{\sum\limits_{i=1}^{W imes H} d[i]}{W imes H}$$

- 4) $\varepsilon = |Error[1] Error[0]|$,若 ε 符合精度要求,转第 6 步,否则 Error[0] = Error[1].
 - 5) 重新计算类 0 和类 1 的期望E[1], E[2]:

$$E[1] = rac{\sum\limits_{i} D_1[i]}{L_1} \qquad E[2] = rac{\sum\limits_{i} D_2[i]}{L_2}$$

释放数组 D_1,D_2 ,转第三步(俗称重新洗牌).

6) 如此图像中的所有像素点分到 D_1,D_2 两

类中,然后将 D_1 中像素点的灰度值全都置为 0, D_2 中像素点的灰度值全都置为 1. 这样就实现了图像的二值化.

3 实验结果

图 2 为文本灰度图像,图 3 是其直方图,可以看出直方图处于整个灰度值的低端范围,这说明图像整体偏暗. 先用文中式(1)所描述的方法,得到阈值 th=(0+122)/3≈40.7,二值化的结果如图 4 所示,可以看出效果不是很好,图像周围应是白色的地方被二值化到黑色. 以图 2 中圈出来的像素点(591,36)为例,从图中看出该点应归为"背景"类,也就是白色,现在已知它的灰度值 28,按照上面计算出来的阈值 40.7 来划分,该点的灰度值小于 40.7,被划分到"目标"类中,于是变成了黑色. 这是因为整个图像光线不均匀,导致中间偏亮,四周偏暗,由此可以得出,一个全局的固定阈值不适用于二值化类似图 2 这样的整张图像.



图 2 一幅待识别的灰度图像

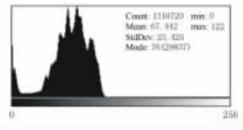


图 3 直方图

如果使用上面阐述的分块 EM 算法,对图像每块中的阈值动态聚类,就能有效地解决这个问题.

还是针对图 2 中的文本灰度图像,对输入的该图像 I,先分块,在每块上使用改进后的 EM 算法,经过几次循环,反复修正 E[1], E[2],将图像 I 中所有像素正确聚类到黑、白二个类,二值化结果如图 5 所示,这个结果显然好得多.



图 4 采用式 1 确定阈值二值化的结果



图 5 EM 算法二值化的结果

4 结束语

手写体的文本识别一直是一个非常活跃的研究领域文中采用改进的 EM 算法对分块的阿拉伯手写体文本图像进行了二值化处理,实验效果良好,但是在实际应用中,对文本图像分块上不能一概而论.例如,若扫描过程中由于光线分布不均得特别厉害,则分块数势必要增加才能很好地将背景和文本分离,但是增加图像分块的同时也加大了计算量,降低了识别系统的效率,这二者之间的平衡还需要通过实验来获得一个较佳的终值.

参考文献

- [1] 章毓晋. 图像处理和分析[M]. 北京:清华大学出版 社,2001.
- [2] Lorigo L M, Govindaraju V. Offline arabic handwriting recognition: a survey, pattern analysis and machine intelligence [J]. IEEE Transactions, 2006, 28 (5):712-724.
- [3] 庄 军,李弼程,陈 刚. 一种有效的文本图像二值 化方法[J]. 微计算机信息,2005,21(8):56-57.
- [4] 杨 玲. 脱机手写体汉字识别研究[D]. 成都: 西华 大学数学与计算机学院, 2008.
- [5] Al-Shaher A A, Hancock E R. Learning mixtures of point distribution models with the EM algorithm.

pattern recognition [J]. Pattern Recognition, 2003, (36):2805-2818.

[6] Xu L , Jordan M I. On convergence properties of the

em algorithm for Gaussian mixtures. Neural Computation [J]. Neural Computation, 1996(8):129-151.

A Novel Block Image Binarization Method Based on EM Algorithm

Wang Hongxia Cheng Yanfen

(School of Computer Science and Technology, WUT, Wuhan 430063, China)

Abstract: Image binarization is an important preprocess step for a document processing system which is used to remove redundant information from a color or gray scale document. In this paper, we implemented an advanced block image binarization method based on the EM algorithm to binarize documents with uneven lighting. It can solve the problem that a fixed threshold will not effectively distinguish the backgrounds from the text. In this research we applied two techniques to effectively solve this problem: 1) We cut the gray sale image into small blocks, and 2) to each block, we use the advanced EM algorithm to cluster the background and text pixels into two groups, therefore, the background and text pixels can be easily separated within each block. The binarized clean image can be used further for feature extraction and other high level process.

Key words: pattern recognition; OCR; image binarization; EM algorithm; block image

(上接第734页)

Prediction Method of Cabled-Stayed in Construction Based on Wavelet Neural Network

Li Yuansong¹⁾ Li Xinping²⁾ Huong Minshui¹⁾

(School of Environmental and Civil Engineering, Wuhan Institute of Technology, Wuhan 430073, China)¹⁾

(School of Civil Engineering and Architecture, WUT, Wuhan 430070, China)²⁾

Abstract: The basic theory and implementing steps of wavelet neural network are introduced. The internal relations between monitoring data and the wavelet neural network are discussed. A predicting model is set up base on wavelet neural network. With engineering practice example for background, the elevation of cantilever installation and initial cable force of Tianxinzhou cable-stayed bridge are simulated and predicted by means of the wavelet network. The results show that the predicting the value by wavelet network nearly accords with measuring value, the predicting method and the precision satisfy the request of the engineering and construction control.

Key words: bridge engineering; construction monitor; wavelet neural network; cable-stayed bridge; prediction method