

适于移动终端字符识别环境的自适应多阈值二值化方法

朱德利¹ 杨德刚¹ 胡 蓉² 万 辉¹

(重庆师范大学计算机与信息科学学院 重庆 401331)¹

(西南大学计算机与信息科学学院 重庆 400715)²

摘 要 为了解决移动终端字符识别应用中光照不均匀、环境不可控而导致的图像二值化效果不佳的问题,提出一种基于积分图快速计算的多阈值自适应二值化方法。该方法首先以待求点为中心设置一个特定尺寸的滑窗,计算该滑窗内所有点的均值,再根据高斯函数加权计算当前滑窗的两个前置滑窗的均值。设置均值松弛因子来衡量当前点的光照情况。像素点的松弛阈值依据该点的松弛因子和光照情况的评价综合计算获得。以 Lenovo ZUK Z2 Pro 作为实验设备,在 Android 操作系统中编写程序,进行文字识别精度的测试。所提算法对前景划分的平均召回率为 95.5%,平均准确率为 91%。调用 Tesseract 4.0 的原生 OCR 识别引擎进行验证,在不规则阴影、多层次光照、线性光线变化等环境下,算法的文字识别准确率分别为 96.8%,98.2%和 93.2%,高于其他预处理算法。所提算法具有较强的鲁棒性和自适应能力,能满足移动终端字符识别应用的图像预处理要求。

关键词 移动终端,图像处理,积分图像,自适应二值化,字符识别

中图分类号 TP391 文献标识码 A DOI 10.11896/j.issn.1002-137X.2019.08.052

Adaptive Multi-level Threshold Binaryzation Method for Optical Character Recognition in Mobile Environment

ZHU De-li¹ YANG De-gang¹ HU Rong² WAN Hui¹

(College of Computer and Information Science, Chongqing Normal University, Chongqing 401331, China)¹

(College of Computer and Information Science, Southwest University, Chongqing 400715, China)²

Abstract In order to solve the problem of poor binaryzation quality caused by uneven illumination and uncontrollable environment in OCR applications of mobile terminals, this paper proposed an adaptive multi-level threshold binaryzation method based on integral graph. First, a specific sliding window is set by focusing on the points to be calculated. The normal threshold is the mean value of the sliding window where the current point is located. The two front sliding windows are weighted according to the Gauss function, and then the relaxation factor is obtained according to the weights. The relaxation threshold of pixels are obtained based on the evaluation of the relaxation factor and illumination condition. Experiments were carried out in typical mobile environments such as irregular shadows, multi-level illumination and linear light changes. Lenovo ZUK Z2 Pro is used as the test equipment. The average recall of the algorithm is 95.5% and the average accuracy is 91%. The recognition accuracy of this algorithm is 96.8%, 98.2% and 93.2% respectively in the environment of irregular shadow, multilevel illumination and linear light change. The result shows that the proposed algorithm has strong robustness and adaptability, and can meet the requirement of image preprocessing in the OCR application of mobile terminal.

Keywords Mobile terminal, Image processing, Integral graph, Adaptive binaryzation, Optical character recognition

1 引言

随着移动互联网在社会生活各个层面的快速推进,对移动终端应用快速而又准确地进行字符识别的需求越来越多。

二值化是字符识别前不可或缺的预处理过程。二值化的效果直接决定了字符分割、特征提取和模式识别的准确性。然而,移动终端应用环境存在较大的不确定性,在移动环境下获取图像时通常会不可避免地受到灯光、人影或其他不同物体的

到稿日期:2018-07-04 返修日期:2018-10-01 本文受重庆市教委科学技术研究项目(KJQN201800536, KJQN201800521, KJ1600322),重庆市科委基础研究与前沿探索计划(cstc2018jcyjAX0470)资助。

朱德利(1979—),男,副教授,主要研究方向为机器学习与机器视觉应用, E-mail: 463453339@qq.com; 杨德刚(1976—),男,博士,教授, CCF 会员,主要研究方向为神经网络、图像处理, E-mail: ydg42@163.com(通信作者); 胡 蓉(1980—),女,博士,讲师,主要研究方向为人机交互、移动视觉搜索; 万 辉(1980—),男,硕士,讲师,主要研究方向为机器学习、图像处理。

阴影的影响。在这种环境下采集到的图像与实际图像之间存在较大的差异。亮度不可控对待识别图像的二值化效果有较大的副作用,直接影响了字符识别的准确度。

二值化是以一定的规则将灰度图像 $f(x,y)$ 分成背景和前景两个部分。二值化方法在字符识别应用中极为重要,得到了国内外学者的广泛研究。学者们由此提出了许多方法,常用的方法有全局阈值法和局部阈值法。其中,全局阈值法是整幅图像只有一个阈值,包括直方图双峰法、最大类间方差法、最小误差阈值法等^[1];局部阈值方法允许一张图像有多个阈值,包括阈值差值法、水线阈值法、基于小波的多阈值方法等^[2]。为解决相关应用问题,近年来一些新的方法被提出。段锁林等^[3]提出了一种多窗口图像二值化算法,采用多窗口模板分别对图像进行简单二值化处理,对不同窗口处理下获得的二值图进行综合判断,最终确定每个像素点的值。李真等^[4]使用改进的同态滤波和自适应灰度变换法相结合的方法对图像进行增强处理,然后使用 OTSU 算法对图像进行二值化。熊炜等^[5]利用最小均值法灰度化和双边滤波去噪进行预处理,然后通过形态学闭操作估计出图像背景,最后采用增广路算法求得最优二值图像。吴锐等^[6]提出了一种基于谱聚类的图像二值化方法。潘梅森等^[7]根据 SOFM 的聚类确定图像第一阈值作为循环迭代的初始值,对整幅图像进行循环迭代得到第二阈值,然后使用第二阈值对原始图像进行二值化。

国外对字符识别中的二值化也有较多的研究。Vo 等^[8]用回归方法解决强噪声和非均匀背景下的图像二值化问题; Sezgin 等^[9]对图像选定某一固定阈值,通过将此阈值与各像素值进行比较来获得二值化图像。日本学者大津^[10]提出如果前景和背景之间的类间方差越大,则构成图像的两个部分之间的差别越大,当部分目标被错分为背景或部分背景被错分为目标时,都会导致两部分的差别变小,当所取阈值的分割使类间方差最大时,就意味着错分概率最小。然而,以上这些固定阈值的算法往往不能很好地适应光照的变化。因此,为处理文档光照不均的问题,学者们提出了一些自适应的二值化方法。Zhang 等^[11]对大津法进行了改进,改进方法对文档对象的二值化效果较好; Michalak 等^[12]针对退化文档,使用对比度增强和区域方法进行二值化; Nasri 等^[13]使用基于全局和局部阈值的组合,获得了较好的效果; Ntirogiannis 等^[14]在综合研究退化文档的基础上,使用各种方法的组合对手写文档进行了二值化和字符识别的研究; Howe^[15]和 Su^[16]等则对各种条件下的文档使用自动化参数自适应地调整二值化效果,使结果的鲁棒性较好; Lu 等^[17]基于对比度增强方法对退化文档图像的二值化进行研究,取得了较好的效果。其中,Bradley 等^[18]首先把积分图方法引入到二值化计算中,提出以 $W * W$ 为模板的矩形区域的二维平滑值来代替一维加权值,该方法对光照不均情况下图像的二值化取得了较好的效果。

移动终端应用中除了光照环境多变外,还要求有较高的运行效率和较好的字符识别效果。本文在前人研究的基础上,以一定区域的二维模板计算为基础,引入阈值松弛因子,在评估与相邻区域的亮度关系的基础上,获得该区域的二值分割阈值;在计算二维模板像素的和值和均值时,改进了现有

的积分图计算方式,加快了算法的整体运行速度,期望算法能在效果和运行速度方面适于移动终端字符识别环境。

2 自适应多阈值二值化的快速算法

2.1 自适应多阈值二值化算法的基本原理

二值化算法的最终目标是确定一个像素点属于背景还是前景,如果是背景,则将该点赋值为 0,反之则将其赋值为 255 (或 1)。本文算法首先以待求点 P 为中心设置一个尺寸为 $H * H$ 的滑窗,计算该滑窗内所有点的均值,再根据高斯函数计算当前滑窗的 2 个前置滑窗的均值加权。设置均值松弛因子 ξ 来衡量当前点 P 的光照情况。如果当前点比前置点的整体光照情况更亮,则提升正常阈值 ξ ; 反之,则在正常阈值的基础上降低 ξ ,以适应相对更暗的光照环境。把基于松弛因子计算的新阈值作为每个区域二值判断的标准。

定义 1 正常阈值 T 为 $P(x,y)$ 点所在滑窗的所有像素点的像素值的均值。以 $3 * 3$ 的滑窗为例,计算图像中每个点的正常阈值 T 的示意图如图 1 所示。 P 点的正常阈值用式 (1) 计算。对于不能被滑窗完全覆盖的点,如图像中边缘和角落的点,其正常阈值为距离该点最近的滑窗覆盖范围的均值。

$$T = (P + \sum_{i=1}^8 P_i) / 9 \quad (1)$$

	P_1	P_2	P_3	
	P_8	P	P_4	
	P_7	P_6	P_5	

图 1 用 $3 * 3$ 的滑窗计算图像中每个点的正常阈值 T
Fig. 1 Calculation of normal threshold value T for each pixel by $3 * 3$ sliding window

定义 2 松弛因子 ξ 是当前点 P 的正常阈值与前置点的正常阈值按照高斯分布分配权重后的比值。可以用松弛因子来衡量当前点的光照变换情况。

(1) 如果当前点为图像左侧第 1 个点,即左侧邻近像素数为 0,则 $\xi = 1$;

(2) 如果当前点为图像左侧第 2 个点,即左侧邻近像素数为 1,则 $\xi = C_1 / C$;

(3) 如果当前点为图像左侧第 3 个点,即左侧邻近像素数为 2,则其松弛因子按照式 (2) 进行计算:

$$\xi = \frac{\lambda_1 C_1 + \lambda_2 C_2}{C} \quad (2)$$

(4) 其余像素的松弛因子按照式 (3) 进行计算:

$$\xi = \frac{\lambda_1 C_1 + \lambda_2 C_2 + \lambda_3 C_3}{C} \quad (3)$$

以上公式中, C 为当前点的正常阈值, C_i 为图像左侧第 i 个点的正常阈值, λ_i 为 i 点的阈值权重。

图像是连续的,越靠近当前处理的像素的点集,其光照相关性越强。为了通过松弛因子表现图像的光照变化,本算法中阈值权重 λ_i 采用高斯函数来确定。

高斯函数的一维形式如式(4)所示:

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty < x < \infty \quad (4)$$

其中, μ 是 x 的均值, σ 是 x 的方差。本文方法需要使用前置点的正常阈值来评估光照变化情况, 并获取当前处理的像素点的松弛阈值。计算平均值时, 中心点就是原点, μ 等于 0, σ 按照标准正态分布假设取 1。权值分布的函数为:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (5)$$

计算松弛阈值时, 只需要将当前计算点作为原点, 前置点按照其在正态曲线上的位置分配权重, 就可以得到一个加权平均值。如图 2 所示, 由于算法使用的是待计算点的 3 个前置像素点, 使用式(5)计算 3 个前置像素点的权值, 分别为 68.26%, 27.18% 和 4.28%; 考虑计算的简便性, 实际计算的权值分别为 68%, 27% 和 5%。

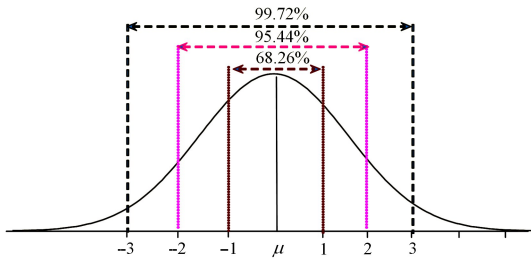


图 2 高斯函数决定的权值分布

Fig. 2 Weight distribution determined by Gauss function

在加权值的基础上, 图 3 给出当前点的松弛因子的计算方法。

	有 1 个前置点的权重分配方法 $\lambda_1 = 1, \xi = C_1 / C$
	有 2 个前置点的权重分配方法 $\lambda_1 = 0.7, \lambda_2 = 0.3, \xi = \frac{\lambda_1 C_1 + \lambda_2 C_2}{C}$
	有 3 个前置点的权重分配方法 $\lambda_1 = 0.68, \lambda_2 = 0.27, \lambda_3 = 0.05,$ $\xi = \frac{\lambda_1 C_1 + \lambda_2 C_2 + \lambda_3 C_3}{C}$

图 3 松弛因子权重的计算方法

Fig. 3 Calculation method of the weight of relaxation factor

定义 3 松弛阈值 TS 是像素点的正常阈值 T 与松弛因子 ξ 的乘积。

$$TS = T * \xi \quad (6)$$

由松弛因子 ξ 的定义可知, ξ 能够衡量当前点与前置点的相对光照情况。如果当前点比前置点的整体光照情况更亮, 则松弛阈值比正常阈值大 ξ ; 反之, 则在正常阈值的基础上降低 ξ , 以适应相对更暗的光照环境。

松弛阈值是算法在每个像素点真正使用的阈值。如果当前点的值大于松弛阈值, 则该像素点设置为 255, 否则为 0。式(7)是在松弛阈值的基础上得到最终二值化图像的方法。

$$P = \begin{cases} 255, & \text{原像素值大于松弛阈值} \\ 0, & \text{原像素值小于或等于松弛阈值} \end{cases} \quad (7)$$

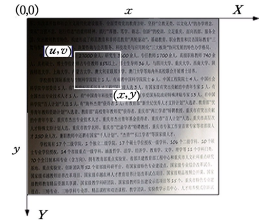
2.2 基于积分图的区域均值快速计算方法

在以上算法的实现过程中, 一个关键问题是要多次计算不同的 $H * H$ 区域的均值, 这需要大量计算图像不同区域的像素值的和。如果使用直接求和再求均值的方法, 会有很多重复计算, 并且计算速度很慢。根据文献[18]的方法, 本文引入积分图来解决这个问题, 使得算法只需要遍历一次图像就能获取任意区域的像素值的和; 同时根据算法的特点, 改进积分图的获取方法, 以便更快地获取到图像的积分图。

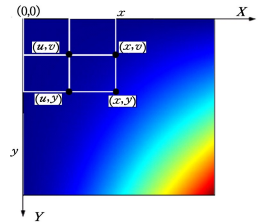
2.2.1 用积分图计算特定区域像素值的和

积分图像是在原图像的基础上衍生出的新图像。新图像的像素值是取原图像左上侧的全部像素值的累加和。整体思路是先计算整幅图像的累加表, 然后以此累加表为辅助, 以简单的加减运算获得任意区域的累加和。如图 4 所示, 原图像 M 上的 (x, y) 点在积分图像上的 (x, y) 位置的值通过式(8)进行计算。

$$I(x_n, y_n) = \sum_{x=1}^{x_n} \sum_{y=1}^{y_n} image(x, y) \quad (8)$$



(a) 原图像 M



(b) 积分图像 N

图 4 图像 M 的积分图像 N

Fig. 4 Integral graph N of M

在积分图的辅助下, 可以不用再次扫描图像即可计算任意区域的像素值的和。如图 5 所示, 设点 $(u, v), (x, v), (u, y), (x, y)$ 4 个点的积分图值分别为 $I(u, v), I(x, v), I(u, y), I(x, y)$, 则区域 D 的像素值的和用式(9)进行计算。

$$\begin{aligned} Sum(D) &= [Sum(A) + Sum(B) + Sum(C) + Sum(D)] - \\ &= [Sum(A) + Sum(B)] - [(Sum(A) + Sum(C)) + Sum(A)] \\ &= I(u, v) - I(x, v) - I(u, y) + I(x, y) \end{aligned} \quad (9)$$

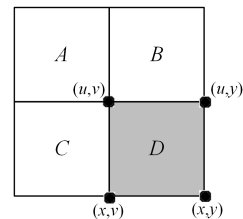


图 5 用积分图直接计算 D 区域像素值的和的方法

Fig. 5 Method of calculating sum in region D by integral graph

使用式(9)能获取任意区域的像素值的和,在和的基础上除以像素个数即可获得图像在该区域的像素均值。

2.2.2 获取积分图的快速算法

基于原图像的一个储存了每一个像素点左上角像素值的和的衍生数表就是积分图。计算区域均值的前提条件就是先有这个数表储存在内存中,故积分图计算的复杂性是影响整个二值化算法计算复杂性的重要因素。现有计算积分图的方法一般是按行扫描每一个像素,然后相加。这种方法没有充分利用已经计算的积分结果,时间复杂度比较高。文献[18]使用了待求点的左上点积分值再配合右边一点和上面一点的积分值来计算,但该方法要额外对第一行和第一列进行特殊处理,引入了新的计算复杂性。本文在原积分图数表上新增值为0的一行和一列,同时利用已经计算的积分值来获取新的积分值,通过储存空间的少许增加,换取了时间上的较高效率。

如图6所示,设点(x,y)的像素值为Image(x,y),把(x,y)点看作图5中的区域D;将该点所在行的左边部分看作区域C;将该点所在列的上边部分看作区域B。由式(9)可得:

Image(x,y) = I(x-1,y-1) - I(x,y-1) - I(x-1,y) + I(x,y)

把待求的(x,y)点的积分值移到等式的左边,得到式(11):

I(x,y) = I(x,y-1) + I(x-1,y) - I(x-1,y-1) + Image(x,y)

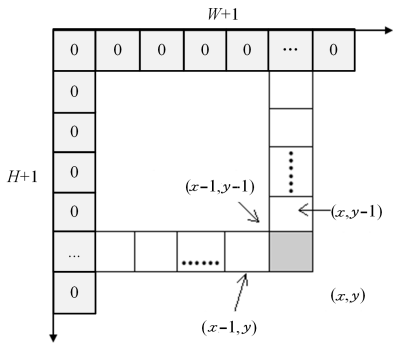


图6 积分图的快速计算方法示意图

Fig. 6 Quick calculation method of integral graph

由图6可以发现,用新的积分图计算方法可以避免每一行和每一列像素值的分别求和,充分利用了待求像素点的前置点的积分值。以上计算方法所用算法的伪代码表示如算法1所示。

算法1 积分图加速计算算法

```
Procedure AccelerateIntegral(InImage, integralimage) {
Begin
w ← InImage.width
h ← InImage.height
integralimage[w+1, h+1] {创建一个尺寸为 w+1 和 h+1 的新容器存放积分值}
for i=0 to w+1 do
integralimage[i, 1] ← 0
for j=0 to h+1 do
integralimage[1, j] ← 0 {初始化积分图的第一行和第一列}
```

```
for i=1 to w+1 do
for j=1 to h+1 do
integralimage[i, j] ← integralimage[x, y-1] + integralimage[x-1, y] - integralimage[x-1, y-1] + srcImage[x, y]
End
```

3 实验结果与分析

为了证明本文算法在移动终端环境下对字符识别的效果,使用Lenovo ZUK Z2 Pro的移动终端设备在不规则阴影、多层次光照、线性光线变化等实际应用环境下采集图像。在Android操作系统中编写程序进行文字识别精度和速度的验证。图像的二值化处理算法采用NDK配合JNI模式,用C++语言进行实现。统一调用Tesseract OCR识别引擎的API获得文字识别的结果。Lenovo ZUK Z2 Pro配置的CPU为四核高通骁龙820(MSM8996),主频是2.15GHz,带高通Adreno530的GPU, RAM容量为6GB。

图7为不规则阴影图像在不同算法中的二值化效果;图8为多层次光照环境中采集的图像二值化效果;图9是线性光线变化环境下的图像二值化效果。

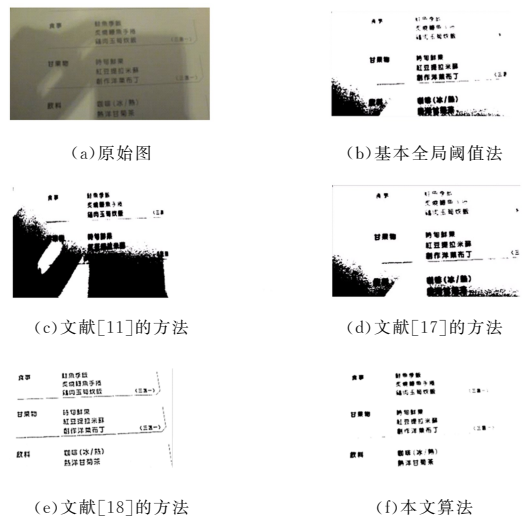


图7 不规则阴影图像的处理效果

Fig. 7 Processing results of image with irregular shadow

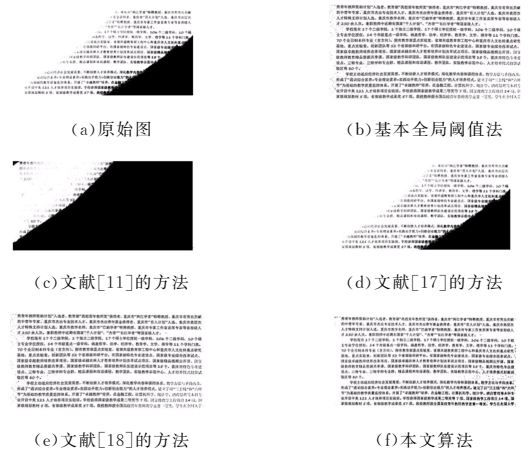


图8 多层次光照图像处理效果

Fig. 8 Processing results of image with multilevel illumination

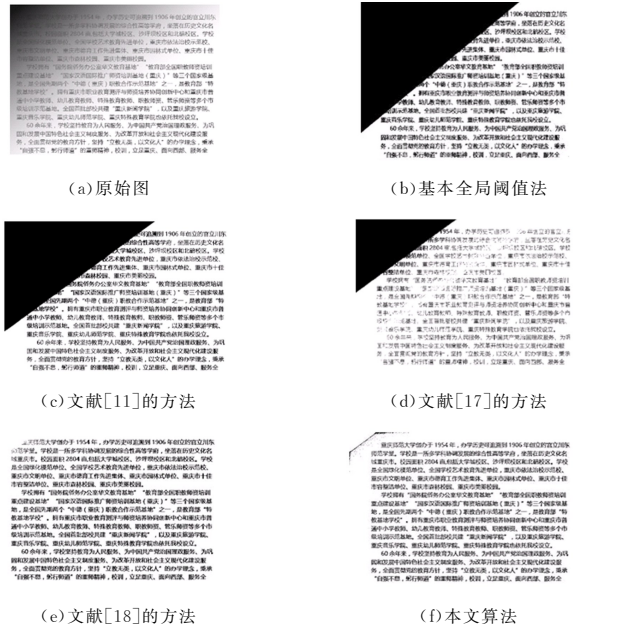


图 9 线性光线变化图像的处理效果

Fig. 9 Processing results of image with Linear ray change

从处理结果的视觉效果方面来看,由于移动环境下光线存在各种变化,使用常规的二值化方法都不能获得满意的效果。如在图 7 用手机拍摄菜单进行字符识别的应用场景下,基本全局阈值法、文献[11]改进的大津法、文献[17]基于对比度增强的退化图像二值化方法等都是针对图像全局使用的单阈值,由于光线不规则,这些方法都不可避免地吧不规则阴影当成前景,或者把有用的文字图像信息过滤掉。文献[18]的方法和本文的方法由于都是使用的多阈值,能较好地提取各种环境下的文字信息。但由于文献[18]获取区域阈值时没有考虑周边像素点的亮度变化情况,处理结果中的文字部分容易出现细微的断笔,尤其是对图 8 所示的多层次光照和图 9 所示的线性光线变化环境下的文字信息的提取不够精细,多层次光照环境下对背景噪音的去除不是很理想。本文的方法使用了多阈值来自适应不同的应用场景,同时考虑到了相连像素点的光线变化关系,获得了比以上这些方法更好的效果。

使用 Gatos^[19]提出的二值化效果评价方法,在像素级别上对二值化前景和背景判断的准确性进行判断,使用召回率 R 和准确率 P 两个指标进行定量分析。用 C_{tp} 表示算法得到的前景像素点分割结果与真实前景对象相符的像素点个数; C_{fn} 表示本该为目标物体的像素被错分割为背景部分的像素点个数; C_{fp} 表示本该为背景部分而被错分割为目标物体的像素点个数。算法的召回率和准确率的计算公式分别用式(12)和式(13)进行计算。

$$R = C_{tp} / (C_{tp} + C_{fn}) \times 100\% \quad (12)$$

$$P = C_{tp} / (C_{tp} + C_{fp}) \times 100\% \quad (13)$$

针对本算法在不同环境下处理的结果图像进行分析,计算被正确判定的前景点和背景点个数,然后对不同图像的判定结果求取平均值,可以得到不同算法的召回率和准确率结果,如表 1 所列。

表 1 不同二值化算法的召回率和准确率比较

Table 1 Comparison of recall and precision of different algorithms

算法	召回率	准确率
基本全局阈值法	82.2	85
文献[11]的方法	75	88
文献[17]的方法	72	84.1
文献[18]的方法	90.2	86
本文算法	95.5	91

利用 5 种二值化算法处理在不同光照条件下获取的图像,并将结果图像送往 OCR 系统,然后将 OCR 识别的结果与原文字进行比较。对二值化后的图像统一调用 Tesseract 4.0 的原生 OCR 识别引擎,利用文字识别的准确率来分析不同算法在实际环境中的应用效果,如表 2 所列。

表 2 结果图像在 Tesseract OCR 引擎中的识别准确率

Table 2 Recognition accuracy with Tesseract OCR engine

使用环境	识别准确率				本文算法
	基本全局 阈值法	文献[11] 的方法	文献[17] 的方法	文献[18] 的方法	
不规则阴影	79.1	71.8	89.6	90.3	96.8
多层次光照	59.6	55.6	95.7	92.6	98.2
线性光线变化	80.9	83.5	85.5	90.5	93.2

从统计结果可以看出,在适应移动终端应用中的不同环境方面,本文方法的效果较优。文献[18]的结果与本文算法最为接近,但是文献[18]的结果图像中背景仍然存在部分噪声,尤其体现在处理结果会产生字迹笔画断笔、字迹浓淡不均等问题。而其他常规的二值化方法,尤其是全局阈值的二值化方法,对移动终端所在的复杂应用环境的适应能力不如本文所采用的算法。

结束语 本文提出一种基于积分图快速计算方法的多阈值自适应二值化方法,来解决移动终端字符识别应用中光照不均匀、环境不可控导致的识别效果不佳的问题。该方法通过一定尺度的滑窗来获取像素点的正常阈值,引入松弛因子来评价当前点的光照变化情况,并根据前置点的权重值计算综合确定前景划分的松弛阈值。通过对在不规则阴影、多层次光照、线性光线变化等典型的移动环境下采集的图像进行实验,采用 Lenovo ZUK Z2 Pro 作为实验设备,在 Android 操作系统中编写程序进行文字识别精度的实验。结果表明,本文算法的平均召回率为 95.5%,平均准确率为 91%。调用 Tesseract 4.0 的原生 OCR 识别引擎进行验证,在不规则阴影、多层次光照、线性光线变化等环境下,本文算法的文字识别准确率分别为 96.8%、98.2% 和 93.2%,高于其他预处理算法。这说明本文的算法在移动终端环境的复杂光照背景下能取得相对较好的字符识别效果,具有较强的环境自适应能力和鲁棒性,能满足移动终端字符识别应用的图像预处理要求。从实验数据上看,本文的研究在前景提取的召回率和准确率方面还有进一步提升的空间,未来的研究将进一步优化积分图计算方案,引入信息熵来评价和确定图像的前景及背景。

参考文献

[1] ISMAIL S M, ABDULLAH S N H S, FAUZI F. Statistical bina-

- rization techniques for document image analysis[J]. *Journal of Computer Science*, 2018, 14(1): 23-36.
- [2] BATAINEH B, ABDULLAH S N, OMAR K. Adaptive binarization method for degraded document images based on surface contrast variation[J]. *Pattern Analysis & Applications*, 2015, 20: 1-14.
- [3] DUAN S L, ZHU F, YAN X. Research of Multi window binarization algorithm [J]. *Computer Engineering and Application*, 2017, 53(17): 212-217. (in Chinese)
段锁林, 朱方, 严翔. 多窗口图像二值化算法研究[J]. *计算机工程与应用*, 2017, 53(17): 212-217.
- [4] LI Z, LI G Y. Research on the binarization of power meter in complex lighting environment [J]. *microcomputers and applications*, 2017, 36(15): 45-48. (in Chinese)
李真, 李功燕. 复杂光照环境下电力仪表的二值化研究[J]. *微型机与应用*, 2017, 36(15): 45-48.
- [5] XIONG W, WANG X R, FENG C. Document image binarization based on background estimation and energy minimization [J]. *Computer Application* 2018, 38(3): 1-8. (in Chinese)
熊炜, 王鑫睿, 冯川. 基于背景估计和能量最小化的文档图像二值化[J]. *计算机应用* 2018, 38(3): 1-8.
- [6] WU R, HUANG J H, TANG J L. Binarization method of text image based on gray histogram and spectral clustering [J]. *Journal of Electronic and Information*, 2009, 31(10): 2460-2464. (in Chinese)
吴锐, 黄剑华, 唐降龙. 基于灰度直方图和谐聚类的文本图像二值化方法[J]. *电子与信息学报*, 2009, 31(10): 2460-2464.
- [7] PAN M S, RONG Q S. Image fusion binarization method based on SOFM neural network [J]. *Optical Precision Engineering*, 2007(3): 401-406. (in Chinese)
潘梅森, 荣秋生. 基于 SOFM 神经网络的图像融合二值化方法[J]. *光学精密工程*, 2007(3): 401-406.
- [8] VO G D, PARI C. Robust Regression For image binarization under heavy noise and nonuniform background[J]. *Pattern Recognition*, 2018, 81(2): 224-239.
- [9] SEZGIN M, SANKUR B. Survey over image thresholding techniques and quantitative performance evaluation[J]. *Journal of Electronic Imaging*, 2004, 13(1): 146-168.
- [10] OTSU N. A Threshold Selection Method from Gray-Level Histograms[J]. *IEEE Transactions on Systems, Man, and Cybernetics*, 1979, 9(1): 62-66.
- [11] ZHANG Y, WU L. Fast Document Image Binarization based on an improved adaptive Otsu's method and destination word accumulation[J]. *Journal of Computational Information Systems*, 2011, 7(6): 1886-1892.
- [12] MICHALAK H, OKARMA K. Fast Adaptive Image Binarization Using the Region Based Approach[C]// *Computer Science On-line Conference*. Cham: Springer, 2018: 79-90.
- [13] NASRI M, HOSSEIN-NEJAD Z, HOSSEINI-ZAHMATKESH P. Document Image Binarization Based on Combination of Global and Local Thresholding Methods[J]. *International Journal of Image & Graphics*, 2018, 18(2): 179-186.
- [14] NTIROGIANNIS K, GATOS B. Combined approach for the binarization of handwritten document images[J]. *Pattern Recognition Letters*, 2014, 35: 3-15.
- [15] HOWEN R. ERRATUM. Document binarization with automatic parameter tuning[J]. *International Journal on Document Analysis & Recognition*, 2013, 16(3): 247-258.
- [16] SU B, LU S, TAN C L. Robust document image binarization technique for degraded document images[J]. *IEEE Transactions on Image Processing*, 2013, 22(4): 1408-1417.
- [17] LU D, HUANG X, SUI L X. Binarization of degraded document images based on contrast enhancement[J]. *International Journal on Document Analysis & Recognition*, 2018, 21(1/2): 123-135.
- [18] BRADLEY D, ROTH G. Adaptive Thresholding using the integral graph[J]. *Journal of Graphics Gpu & Game Tools*, 2007, 12(2): 13-21.
- [19] GATOS B, PRATIKAKIS I, PERANTONIS S. Adaptive degraded document image binarization[J]. *Pattern Recognition*, 2006, 39(3): 317-327.