

# 手机文字识别中的暗角修正技术

应 骏 袁豪磊 李 莉

(上海师范大学数理信息学院, 上海 200234)

**摘 要** 为了提高手机图像的识别率,从手机文字扫描识别应用问题出发,在分析手机普遍存在的暗角退化特征的基础上,提出了一种从含有文字信息的图像中提取图像的暗角退化模型的方法;然后根据该暗角模型对输入图像进行复原,最终达到修正手机摄像模组暗角退化的目的。

**关键词** 文字识别 智能手机 二值化 退化模型 暗角修正

中图分类号: TP391.43 文献标识码: A 文章编号: 1006-8961(2007)10-1853-04

## Research on the Dark-corner Correcting in Cell Phone's Character Reorganization

YING Jun, YUAN Hao-lei, LI Li

(Mathematics & Science College, Shanghai Normal University, Shanghai 200234)

**Abstract** From the point of word image recognition application in a cell phone. We analyzed the tunnel-effect feature which popularly existed in a cell phone and proposed a method which can abstract image from the word-contained images. The established model can rejuvenate the input image so as to finally resist vignetting from degradation of camera pattern in a cell phone.

**Keywords** character recognize, smart phone, binarization, regression model, dark-corner correcting

## 1 引 言

手机的普及和发展赋予了手机更广泛的使用功能,硬件性能的不断提高,使得在手机上进行图像识别得以实现,其中最典型的就是利用具有摄像功能的手机实现文字扫描和识别,这样就可将纸面文字快速便捷地输入到手机之中。如今具备这些功能的手机大多属于高档配置,其均配有优良的光学镜头。由于手机摄像模组的专用镜头生产成本较高,国内大多数中低端配置的手机大多没有额外增配一组光学透镜,从而导致其拍摄的图像存在较为严重的暗角现象,这不仅影响了文字识别算法对图像中各个文字的定位,而且识别率大大降低。本文针对这一问题进行了分析研究,提出了一种对抗暗角的文字提取方法。

## 2 图像暗角及其影响

(1) 暗角的产生 为便于生产集成,目前几乎所有的手手机都采用传感器、光学镜头、配套集成电路一体化的模块,称为手机摄像模组(CMOS camera modules, CCM)。为了减小手机的整体尺寸,一方面各厂家将 CCM 的体积都做得很小,因此镜头直径也就很小;另一方面,作为消费类电子产品,要求 CCM 具有低功耗、低成本的特点,手机 CCM 普遍不采用专用图像处理专用镜头和光源。以上两个原因造成了手机 CCM 存在较为严重的光学退化,其中最明显的就是“暗角”现象<sup>[1]</sup>——图像的边角处显得灰暗(图1)。在摄影术语中,“暗角”又称为“失光”<sup>[2]</sup>。

(2) 暗角带来影响 在对文字进行识别之前必

基金项目:上海市教委基金项目(06D2007)

收稿日期:2007-05-30;改回日期:2007-07-01

第一作者简介:应骏(1973~),男,讲师,2000年获浙江大学通信与信息系统专业博士学位。主要研究方向为嵌入式系统、图像处理、多媒体通信。E-mail:jyuning@shnu.edu.cn



图 1 暗角现象

Fig. 1 Tunnel effect

须对图像进行二值化,以提取文字的笔段信息。但暗角污染导致了图像各个位置的亮度不均匀,由于这影响了从图像中提取文字的过程,从而严重影响了文字识别的效率。如果不采用任何措施,仅仅使用全局阈值二值化的处理方法是无法提取出文字内容的,图 2 显示了采用统计的图像灰度值的峰值的 50% 为阈值时,进行全局二值化后的效果。

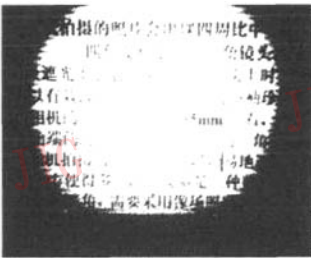


图 2 用单一阈值进行全局二值化

Fig. 2 Global binarization with one threshold

图 2 列出了一幅单一阈值二值化的图像,对比多幅采用不同阈值的全局阈值二值化的结果可以得出结论:由于暗角退化造成了图像的中心和四周之间的平均灰度严重不同,因此无论采用何种大小的阈值,全局阈值二值化都无法兼顾图像的中心区域和边缘区域之间的灰度,而且会造成大量信息的丢失。

事实上,某些高端品牌的手机通过为 CCM 加装一组透镜镜片,虽可从根本上减小暗角的产生,但是由于研磨 CCM 透镜镜片的工艺要求较高,生产成本阻碍了其普及,因此很多中低端手机出于对成本的考虑,不再额外为 CCM 加装透镜,或者仅将一块平镜安装在 CCM 的前面起到抗磨损、保护 CCM 的作用。由于这类手机所拍摄图像的暗角现象是十分严重的,因此为这些中低端手机设计的文字识别软件就必须特别考虑到暗角问题。此外,暗角的中心

位置和衰减程度会随着拍摄角度、距离、光源等因素而发生较大的变化<sup>[3]</sup>。为了取得很高的识别率,对于每一幅输入图像的暗角问题必须单独地考虑。

### 3 动态建立暗角模型

本文方法是首先通过从原始图像中提取背景参数;然后重建背景的暗角退化模型,进而去除图像中的暗角污染。

根据退化模型来对图像进行复原的原理是:由于光学畸变和信号噪声引起的拍摄到的图像与实物之间存在差别,因而产生了图像退化<sup>[4]</sup>。退化可以被抽象为一个退化函数,即输入图像  $f(x, y)$  在退化函数  $H$  的作用下产生退化图像  $g(x, y)$ 。图像复原的目的就是根据退化函数  $H$  的一些知识,尽可能好地求出关于原始图像的近似估计  $\hat{f}(x, y)$ <sup>[5]</sup>。

一旦找到了暗角的退化模型就可以很好地恢复受暗角影响的图像,但问题是如何从原图中提取相关的参数,用来建立暗角退化的模型。一般而言,文字之间存在一定的行间距,只要纸面没有用图像做衬底,那么行间空白就是该区域的背景图像。也就是说,通过观察行间空白就可以得到暗角退化的特征参数,进而即可建立每幅图像特定的暗角模型。

#### 3.1 滤波的方法建立暗角模型

为了建立暗角模型,需要找到区分行间空白和文字行的方法。区分方法可以对图像中的各行进行水平扫描,如果扫描到的是行间空白区域,则由于这些区域的灰度是缓慢变化的,其扫描的波形如图 3(a)所示(图 3 为不同扫描行的灰度值曲线);如果扫描到的是文字所在的行,而且由于这些区域存在很多文字的笔画,灰度变化剧烈,其扫描的波形如图 3(b)

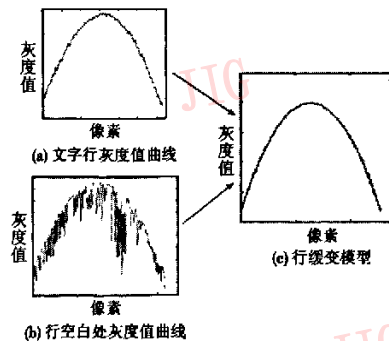


图 3 用滤波的方法建立暗角模型

Fig. 3 Model built by filter

所示,因此只要采用求差分之类的一些简单运算就可以将两者区分开来。由于图像存在行间相关性,因此图3(a)和图3(b)的曲线会各自成簇地出现,簇与簇之间的交接线就是文字行和空白行之间的交界线,这一信息可以被用于后期的文字分割操作中。

从图3(b)中可以看出,文字的黑色笔画使得扫描波形产生了大量的波动,其中波谷代表图像中的前景图像(文字),波峰代表了图像中的背景图像(纸面)。如果对图3(a)和图3(b)进行滤波运算,则可以通过去除波动带来的影响而得到如图3(c)所示的波形,这样也就是去除了图像中的黑色文本,保留了图像的背景信息。根据这一思路建立的退化模型的方法步骤如下:

(1) 用一个低通滤波器对图3(a)和图3(b)进行滤波,取出暗角的缓变曲线(图3(c));

(2) 根据步骤(1)的结果画出背景的暗角图像,首先建立退化模型(图4(d)),然后将模型中的所有像素乘以一个二值化系数 $L(0 < L < 1)$ ;

(3) 用原图和步骤(2)的结果逐像素进行逻辑比较判断,即可得到去除暗角后的二值化输出图像。

虽然图3(b)经过低通滤波后可以得到类似图3(c)的暗角变化灰度曲线,进而可利用每行扫描线来建立整幅图像的暗角退化模型,但是当文字行的笔段非常密集或存在大片的黑色区域时,滤波的结果严重偏离期望的形状,从而影响所建立模型的精确性。

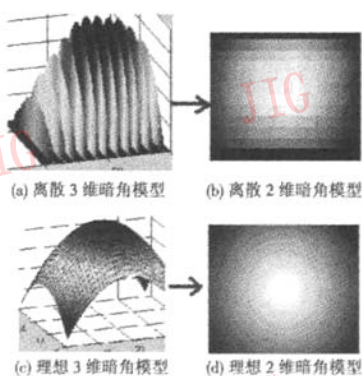


图4 用拟合的方法建立暗角模型

Fig. 4 Model built by approach method

### 3.2 拟合方法建立暗角模型

为了避免这个问题,可以不考虑那些文字行的扫描波形而采用近似拟合的方法计算出被去除行的内容:首先只从原图中提取行间空白作为退化模型

的建立依据,然后通过对各层空白行之间进行近似拟合来表达文字所在行的退化模型曲线。

取得行间空白的暗角退化参数之后,就可以确定图像中各行文字之间的行序,以确定各文字所在行。其结果见图4(b),该信息可以供给后期文字识别程序用来确定文字在图中的位置。(这里假设输入图像不存在严重的旋转或已经过配准)。其算法步骤如下:

(1) 首先用一个低通滤波器对图3(a)进行滤波,不必考虑图3(b)所在行的情况,然后取出空白行的暗角缓变曲线,画出背景的部分暗角图像并根据各行的形状求出拟合函数的基本形式(此时的图像呈斑马线形状);

(2) 从步骤(1)的结果中逐次取出一列,并将此列的数据代入基本拟合函数,再通过求出该列上的所有值来得到暗角模型(图4(d)),最后将模型中的所有像素乘以一个二值化系数 $L(0 < L < 1)$ ;

(3) 用原图和步骤(2)的结果逐像素进行逻辑比较来得到去除暗角后的二值化输出图像;

(4) 拟合函数形式的确定:由于注意到暗角现象呈现中心辐射的轴对称图形,因此拟合函数的形式和空白行的形式相同(图3(c))。

### 3.3 快速近似插值方法

虽然按照上述方法建立的暗角图像是最精细的,但这一算法需要统计出拟合函数的形式,并且要逐列地进行拟合。由于计算开销很大,不适合手机硬件的处理能力。鉴于此,可采用线性插值的方法来建立暗角模型,这就免去了拟合操作,计算量也大大降低,但缺点是最终二值化输出结果不如拟合法得到的结果来得精确。

图4(b)显示了采用对相邻空白行求平均的方法获得的暗角模型。由于采用的是最简单的插值方法(求平均),因此图像呈现明显的行间过渡特征。这样做的优点是大幅度地减少了计算量和存储空间开销,因而便于手机处理器的运算。为了便于说明,图4中分别列出了两种模型的3维扫描图<sup>[6]</sup>,根据3维图形可以对两种模型之间的差别做出另一种解释:图4(a)是图4(c)的一种离散化近似。

最后,利用图4(b)作为阈值对图1进行动态阈值二值化,就可以得到较清晰的二值化输出图像,同时利用之前已确定的行序信息就可以将文字逐行分割开来。

图5显示了图像中的某一行的处理结果。本例中的处理步骤如下:

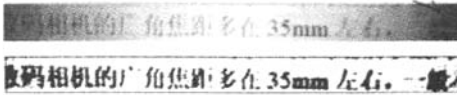


图 5 图像中某一行的处理效果

Fig. 5 Result of processing

(1) 根据行投影获取投影值的局部灰度峰值,在确定行(留白)位置信息之后再行分割,并确定两个相邻行留白的位置;

(2) 根据相邻留白位置,采用  $3 \times 1$  窗口(去噪考虑)确定文字行两端的白平衡峰值;

(3) 根据上述两个灰度峰值进行插值,并计算文字行中的峰值;

(4) 采用拟合曲线峰值的 50% 为阈值进行二值化。

图 5 为实验结果,由图 5 可以看到,不仅输出的图像较为清晰,而且杂波很少。

## 4 与传统方法的对比

利用相邻文字行间空白处来提取文字图像的暗角模型的快速算法,充分利用了文字识别中的留白行的信息。该算法通过进行动态插值来计算二值化阈值,是将全局建立的暗角模型问题转化为行列可分的问题。由于其至少降低了一个  $O(N)$  次数,从而降低了系统的整体运算复杂程度。

而对抗图像各区域平均灰度不均的经典方法是采用分块二值化<sup>[7]</sup>,虽然其在普通情况下有较好的处理结果,但在严重暗角效应的时候,效果不甚理想。图 6 显示了这种情况下采用不同大小的块时,某一分块策略对图像进行二值化后的效果,从图中可看出,选择不同的块内阈值,始终存在块效应与文字丢失的矛盾。

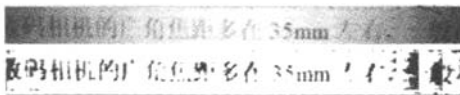


图 6 用不同的尺寸进行分块二值化

Fig. 6 Result of block-binarization

## 5 结论

暗角退化本质上是由手机摄像模组的光学结构

所产生的<sup>[8]</sup>,虽然通过对光学镜头的升级可以彻底解决这一问题,但是由于 CCM 镜头的工艺和生产成本限制,短时间内无法达到普及,因此手机上的暗角问题会继续存在相当的一段时间,这将成为各类手机图像识别应用研究的障碍。

限于手机处理能力的限制,本文提出的退化模型建模方法,是根据传统的图像复原方法,充分利用了文字识别的特征提出的快速改进算法。该方法既获得了较高的性能,又没有运算量的明显增加,整个方法简单、实用,完全适合手机的硬件特点。这就在资源受限的手机平台上,做到了良好的文字二值化。进而为进一步的文字切分、文字识别提供了良好的数据源。这不仅使得二值化之后的图像有很好的显示效果,同时信噪比提高了 3dB 以上。

## 参考文献 (References)

- Gonzalez Rafael C, Woods Richard E. Digital Image Processing[M]. Beijing: Publishing House of Electronics Industry, 2003. [Gonzalez Rafael C, Woods Richard E 著. 阮秋琦, 阮宇智等译. 数字图像处理[M]. 北京:电子工业出版社,2003.]
- Castleman K R. Digital Image Processing[M]. Beijing: Publishing House of Electronics Industry, 1998. [Castleman K R 著. 朱志刚, 石定机译. 数字图像处理[M]. 北京:电子工业出版社,1998.]
- Jia Yun-de. Machine Vision [M]. Beijing: Science Publishing House, 2000. [贾云得. 机器视觉[M]. 北京:科学出版社, 2000.]
- Gonzalez Rafael C, Woods Richard E. Digital Image Processing Using MATLAB[M]. Beijing: Publishing House of Electronics Industry, 2005. [Gonzalez Rafael C, Woods Richard E 著. 阮秋琦等译. 数字图像处理(MATLAB 版)[M]. 北京:电子工业出版社,2005.]
- Jain Anil K. Fundamentals of Digital Image Processing[M]. Cliffs Englewood, NJ, USA, Prentice Hall, 2006. [Jain Anil K 著. 韩博, 徐枫译. 数字图像处理基础[M]. 北京:清华大学出版社, 2006.]
- Zhang Hang, Luo Da-yong. Status and development of study on blind image restoration algorithm[J]. Journal of Image and Graphics, 2004, 9(10): 1145 ~ 1152. [张航, 罗大庸. 图像盲复原算法研究现状及展望[J]. 中国图象图形学报, 2004, 9(10): 1145 ~ 1152.]
- Zhang Hong-gang, Chen Guang, Guo Jun, et al. Image Processing and Recognition [M]. Beijing: Beijing University of Posts and Telecommunications Press, 2006. [张洪刚, 陈光, 郭军等. 图像处理与识别[M]. 北京:北京邮电大学出版社,2006.]
- Kiyoharu Aizawa, Katsuhiko Sakaue, et al. Image Processing Technologies—Algorithms, Sensors, and Applications [M]. New York: Marcel Dekker, 2004.