

文章编号:1001-9081(2010)04-0950-03

## 基于形态学的文档图像径向校正算法

常 骏<sup>1</sup>, 苗立刚<sup>2</sup>

(1. 内蒙古自治区气象局 信息中心, 呼和浩特 010051; 2. 东北大学秦皇岛分校 电子信息系, 河北 秦皇岛 066004)

(changj@sina.com)

**摘 要:**手持相机拍摄的文档图像存在不同程度的镜头失真。根据文档图像的文本行信息,提出了一种基于数学形态学的镜头校正算法。首先利用自适应阈值方法分割文档图像,并通过形态学闭运算把连通体聚类为文本行。然后利用二次多项式模型拟合文本行的中心线,并建立径向失真校正的目标函数。该目标函数把中心线对应的曲线映射为直线,从而求出文档图像的镜头失真参数。实验结果表明,该校正算法可以有效地校正文档图像各种程度的径向失真。

**关键词:**文档图像;径向失真;数学形态学;闭运算

**中图分类号:** TP391.4 **文献标志码:** A

## Correcting radial distortion of document images by morphology

CHANG Jun<sup>1</sup>, MIAO Li-gang<sup>2</sup>

(1. Information Center, Inner Mongolia Meteorological Administration, Hohhot Nei Mongol 010051, China;

2. Electronic Information Department, Northeastern University at Qinhuangdao, Qinhuangdao Hebei 066004, China)

**Abstract:** Document image captured by hand-held camera suffers from lens distortion more or less, and a mathematic morphology based lens correction algorithm was proposed according to the text line of document images. First, it used adaptive thresholding algorithm to segment the document image, and clustered connected components into text lines with morphology closing algorithm. Then, it used second order polynomial to fit the central line of text line, and constructed the object function of lens distortion. It warped the curved text line to the straight text line, thus, to solve the distortion parameters of document image. The experimental results show this method can effectively correct lens distortion of document image with various degrees.

**Key words:** document image; radial distortion; mathematical morphology; closing algorithm

## 0 引言

近年来手持数字成像设备的普及程度越来越高,如数码相机、数码摄像机、PC摄像头、PDA以及可拍照手机等。利用手持成像设备可以方便地获取文档图像,并通过光学字符识别(Optical Character Recognition, OCR)技术将其转化为可编辑的文本文件。这些成像设备具有成像简单、快速以及非接触的特点,目前基于手持相机的文档图像处理技术已经成为研究的热点,并且在实际中得到了广泛的应用。

手持相机拍摄的文档图像都存在一定程度的径向失真,对文档图像的OCR识别率造成不利的影响。目前的镜头校正方法主要分为三类。1)标定模板方法<sup>[1-3]</sup>:它利用已知的标定网格坐标与图像中对应的特征点位置,求解计算相机的外部参数和内部参数。这种方法的精度较高,但是普通用户难以实用。2)铅垂线方法<sup>[4]</sup>:它仅需要单幅图像即可进行失真标定,它需要场景中存在线状结构,如城市场景、室内场景,航空场景中包含的建筑物或人造结构。3)特征点对应方法<sup>[5]</sup>:它利用相机在不同视角拍摄两幅或两幅图像,并利用特征点对和几何约束(基础矩阵或单应矩阵)校正镜头的径向失真。

针对手持相机拍摄的文档图像,本文提出了一种基于文本行聚类的镜头校正方法。在许多情况下文档图像缺少比较长的边缘,但是文本行的中心线通常可以认为是一条直线,径

向失真的影响使文档图像中的文本行的中心线变成扭曲的曲线。本文采用连通体分析方法对相同文本行的各个连通体进行聚类,并利用文本行的中心线建立镜头失真的目标函数,从而得到镜头失真参数。

## 1 镜头失真校正

### 1.1 文档图像分割

手持相机拍摄的文档图像受光照变化影响较大,本文采用自适应阈值分割方法。Niblack方法在文档图像分割中有较多的应用<sup>[6]</sup>,其基本思想是根据图像的局部均值 $m(x,y)$ 和标准差 $s(x,y)$ ,在一个局部邻域内计算每个像素的阈值 $T(x,y)$ :

$$T(x,y) = m(x,y) + k \times s(x,y) \quad (1)$$

其中 $k$ 是一个用户自定义的参数。

本文选择高斯低通滤波器对图像进行滤波,并把平滑结果作为局部均值图像。它能够适应文档字符大小的变化,并有效减少字符的粘连和断裂,对多数分辨率的文档图像都可以取得较好的分割效果。

### 1.2 文本行聚类

除了径向失真之外,由于拍摄视角的变化,文档图像在获取过程中还会产生透视失真。尽管有着两种镜头失真的影响,但相邻字符连通体的宽度或高度仅产生微小的变化。因此,可以通过连通体分析检测出若干个高度或宽度都比较接

收稿日期:2009-10-15;修回日期:2009-12-09。

作者简介:常骏(1969-),男,山西忻州人,博士研究生,主要研究方向:图像处理、计算机网络;苗立刚(1974-),男,河南驻马店人,博士,主要研究方向:图像处理、模式识别。

近的相邻连通体,并将这些连通体构成一个最近邻链,从而检测出所有的文本行。最近邻聚类方法最早用来估计文档图像的偏斜角度,文献[7]在此基础上提出了一种最近邻链的方法,使偏斜角度的估计具有较高的精度,并且可以处理多种语言的文档图像。由于相机的径向失真通常不会太大,因而可以用最近邻方法把各个连通体聚类成本文行。

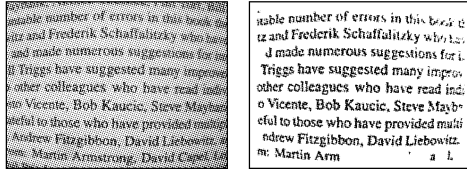


图1 利用自适应阈值方法分割文档图像

图1(a)中的连通体的外接矩形如图2(a)所示,用 $M$ 表示二值图像中连通体的数量,连通体 $C_i (i = 1, \dots, M)$ 的中心坐标表示为 $(x_{c_i}, y_{c_i})$ ,其外接矩形的高度和宽度分别用 $h_{c_i}$ 和 $w_{c_i}$ 表示。对于一个包含 $K$ 个连通体的字串 $[C_1, C_2, \dots, C_K]$ ,如果 $C_{i+1} (i = 1, \dots, K-1)$ 是 $C_i$ 的最近邻连通体,则称这 $K$ 个连通体组成一个长度为 $K$ 的最近邻链。由图2(a)生成的最近邻链如图2(b)所示,本文用 $S_n (n = 1, \dots, N)$ 表示第 $n$ 个长度为 $K_n$ 的最近邻链,其平均宽度 $\bar{w}_n$ 可定义为:

$$\bar{w}_n = \frac{1}{K_n} \sum_{c_i \in S_n} w_{c_i} \quad (2)$$

本文对最近邻链 $S_n$ 中的各个连通体中心点进行线性拟合,并计算 $S_n$ 的斜率 $\alpha_n$ 。为了提高斜率的精度,通常只对 $K_n \geq 3$ 的那些最近邻链进行计算。然后,对分割得到的二值图像 $A$ 进行方向平滑的闭运算(Directional Smoothing Closing Algorithm, DSCA):

$$DSCA_{\alpha_n}(A) = A \oplus \beta_n L_{\alpha_n} \ominus \beta_n L_{\alpha_n} \quad (3)$$

其中, $L_{\alpha_n}$ 表示斜率为 $\alpha_n$ 的单位长度的线状结构元, $\beta_n$ 表示结构元的长度,一般可取为 $\bar{w}_n$ 的2~4倍。

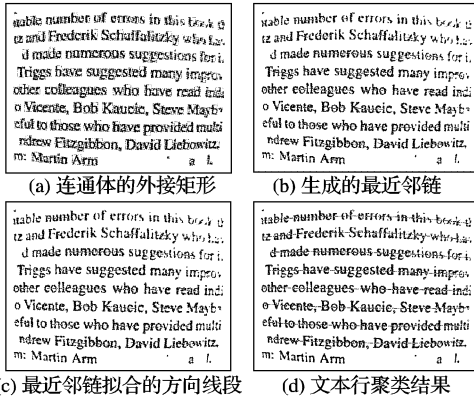


图2 文本行的聚类过程

对于方向平滑后的二值图像,标记出所有与 $S_n$ 连接的连通体,并把它们与 $S_n$ 中的连通体结合起来,生成一个新的较长的最近邻链 $S'_i$ (图2(c))。通常 $S'_i$ 包含属于同一文本行的大部分或全部连通体。如图2(d)所示,本文对 $S'_i$ 中的所有连通体的中心点,采用二次多项式模型拟合其中心线,即:

$$y = a_0 + a_1 x + a_2 x^2 \quad (4)$$

用 $K'_i$ 表示 $S'_i$ 的长度,将 $S'_i$ 所包含的所有连通体的中心点坐标 $(x_{c_i}, y_{c_i}) (i = 1, 2, \dots, K'_i)$ 代入式(4),即可得到关于二次多项式系数 $a_0, a_1, a_2$ 的 $K'_i$ 个线性方程组。一般情况下,

由于 $K'_i \geq 3$ ,该线性方程组为超定方程组,采用最小二乘法即可求解出三个多项式系数,从而得到每个文本行的中心线。本文用这些中心线来建立径向失真校正的目标函数,可以取得较好的效果。

### 1.3 镜头校正模型

摄像机的成像过程一般可以分解为两个步骤:1) 透视投影过程,它把空间中的三维点投影到理想图像平面上一个未失真的图像坐标 $\mathbf{x}_u = (x_u, y_u)^T$ ;2) 镜头失真过程,它把 $\mathbf{x}_u$ 沿径向平移到失真图像的坐标点 $\mathbf{x}_d = (x_d, y_d)^T$ 。

镜头失真主要分为径向失真和切向失真两部分。径向失真导致图像产生桶形或枕形失真<sup>[8]</sup>,切向失真是目标像素沿矢径切向的非线性变形。由于普通相机的切向失真通常比较小,本文仅考虑径向失真的校正。比较常用的镜头失真模型是多项式模型,径向失真模型通常把 $\mathbf{x}_u$ 表示为 $\mathbf{x}_d$ 的函数:

$$\begin{cases} x_u = x_d + (x_d - c_x)(k_1 r_d^2 + k_2 r_d^4 + \dots) \\ y_u = y_d + (y_d - c_y)(k_1 r_d^2 + k_2 r_d^4 + \dots) \\ r_d = \sqrt{\left(\frac{x_d - c_x}{s_x}\right)^2 + (y_d - c_y)^2} \end{cases} \quad (5)$$

其中: $k_1, k_2$ 表示径向失真系数; $\mathbf{e} = (c_x, c_y)^T$ 表示畸变中心的坐标; $r_d$ 表示以 $\mathbf{e}$ 为中心的失真半径, $s_x$ 表示纵横比。

在实际应用时,引入过多的非线性参数并不能显著提高径向失真的校正精度,反而会将增加求解过程的复杂度。对于普通相机,该模型的第一项径向畸变参数 $k_1$ 已能足够描述非线性畸变。为了校正文档图像的镜头失真,需要恢复 $\{k_1, c_x, c_y, s_x\}$ 共四个参数。相机符合针孔模型当且仅当空间中的直线投影到相机仍然为一条直线,如果找到一个变换使得空间中的每条直线在图像中也表示为一条直线,就可以校正径向失真。为此,本文对每条边缘的像素点拟合一条直线,并用所有像素点到该直线的距离之和作为该条曲线的失真度量,则镜头失真的目标函数可以表示为:

$$\min_{k, c_x, c_y, s_x} \sum_i \sum_j \frac{A_i x_{ij} + B_i y_{ij} + C_i}{\sqrt{A_i^2 + B_i^2}} \quad (6)$$

其中: $A_i, B_i, C_i$ 表示第 $i$ 条边缘拟合的直线的系数; $(x_{ij}, y_{ij})$ 表示第 $i$ 条中心线的第 $j$ 个像素的坐标。

本文对失真半径采用规范化坐标,也就是将 $r_d$ 和 $r'_d$ 除以 $\sqrt{(w/2)^2 + (h/2)^2}$ ,使得规范化失真半径总是在0~1变化,它可以提高失真参数的估计精度。本文利用Nelder-Mead提出的下山单纯形方法求解式(6)的最优解,从而可以快速求得所有镜头失真参数。

## 2 实验结果

本文利用普通摄像头拍摄的文档图像进行镜头校正实验,其中实验图像大小为 $640 \times 480$ 。利用文本行聚类结果得到的失真参数为: $k_1 = 0.1366, c_x = 0.4741, c_y = 0.5900, s_x = 0.9055$ ,对应的校正结果如图3所示。

利用该校正参数分别对另外两幅图像的实验结果如图4所示。可以看出,图4(a)中的文档图像存在明显的镜头失真,其中文本行都变成一条扭曲的曲线。应用本文方法得到的这两幅图像的校正参数分别为 $\{0.1279, 0.4871, 0.5499, 0.8786\}$ 和 $\{0.1559, 0.5091, 0.5618, 0.9716\}$ 。校正后的两幅图像如图4(b)所示,其中文本行都大致为一条直线。实验

表明该校正方法具有较高的校正精度,完全可以满足实际的需要。

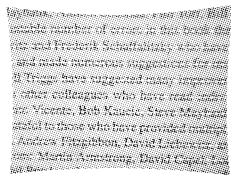
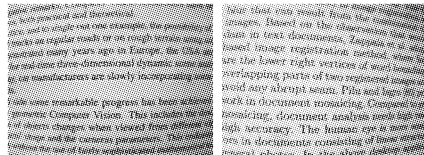
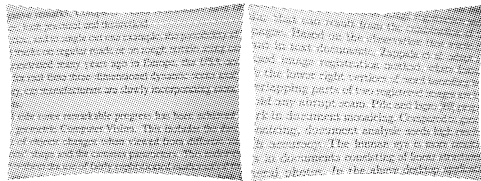


图3 文档图像的校正结果



(a) 从不同角度拍摄的两幅文档图像



(b) 校正结果

图4 两幅文档图像的校正结果

### 3 结语

本文利用文档图像中的文本行信息,提出了一种基于形态学的径向校正方法。该方法提出了方向平滑的闭运算把连通体聚类为文本行,把中心线对应的曲线映射为直线,可以精确地求出文档图像的径向失真参数。实验表明,该方法可以

有效地校正文档图像的径向失真。

### 参考文献:

- [1] TSAI R Y. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses[J]. IEEE Journal of Robotics and Automation, 1987, 3(4): 323 - 344.
- [2] ZHANG Z Y. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330 - 1334.
- [3] 丁莹,范静涛,杨华民,等. 一种广角镜头光学畸变的数字校正算法[J]. 长春理工大学学报, 2009, 32(2): 184 - 188.
- [4] DEVERNAY F, FAUGERAS O. Straight lines have to be straight: Automatic calibration and removal of distortion from scenes of structured environments[J]. Machine Vision and Applications, 2001, 13(1): 14 - 24.
- [5] FITZGIBBON A W. Simultaneous linear estimation of multiple view geometry and lens distortion[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2001: 125 - 132.
- [6] SAUVOLA J, PIETAKSINEN M. Adaptive document image binarization[J]. Pattern Recognition, 2000, 33(2): 225 - 236.
- [7] LU Y, TAN C L. A nearest-neighbor chain based approach to skew estimation in document images[J]. Pattern Recognition Letters, 2003, 24(14): 2315 - 2323.
- [8] WENG J, COHEN P, HERNIOU M. Camera calibration with distortion models and accuracy evaluation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1992, 14(10): 965 - 980.
- [8] EFROS A A, LEUNG T K. Texture synthesis by non-parametric sampling[C]// Proceedings of the IEEE Computer Society International Conference on Computer Vision. Washington DC: IEEE Computer Society, 1999: 1033 - 1038.
- [9] HARALD G. A combined PDE and texture synthesis approach to inpainting[C]// Proceedings of 8th European Conference on Computer Vision. Prague: [s. n.], 2004: 214 - 224.
- [10] RANE S D, SAPIRO G, BERTALMIO M. Structure and texture filling-in of missing image blocks in wireless transmission and compression applications[J]. IEEE Transactions on Image Processing, 2003, 12(3): 296 - 303.
- [11] CRIMINISI A, PEREZ P, TOYAMA K. Region filling and object removal by exemplar-based image inpainting[J]. IEEE Transactions on Image Processing, 2004, 13(9): 1200 - 1212.
- [12] CRIMINISI A, PEREZ P, TOYAMA K. Object removal by exemplar-based inpainting[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. New York: ACM, 2003: 721 - 728.
- [13] 王维兰,唐仕喜. 唐卡图像复杂破损区域的修复[J]. 计算机工程与设计, 2007, 28(3): 703 - 705.
- [14] ZENZO S D. A note on the gradient of a multi-image[J]. Computer Vision, Graphics, and Image Processing, 1986, 33(1): 116 - 125.
- [15] GONZALEZ R C, WOODS R E. 数字图像处理[M]. 阮秋奇,阮宇智,译. 2版. 北京: 电子工业出版社, 2003: 265 - 266.
- [16] LIU HUA-MING, WANG WEI-LAN, XIE HUI. Thangka image inpainting using adjacent information of broken area[EB/OL]. [2009 - 07 - 01]. [http://www.iaeng.org/publication/IMECS2008/IMECS2008\\_pp646-649.pdf](http://www.iaeng.org/publication/IMECS2008/IMECS2008_pp646-649.pdf).
- [1] 唐卡图像介绍[EB/OL]. [2009 - 07 - 01]. <http://tangka.cn.com/>.
- [2] 白云飞,杜华. 如何修复破损唐卡[N]. 人民日报: 海外版, 2004 - 05 - 29(8).
- [3] BERTALMIO M, SAPIRO G, CASELLES V, et al. Image inpainting[C]// Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques. New York: ACM Press, 2000: 417 - 424.
- [4] RUDIN L, OSHER S, FATEMI E. Nonlinear total variation based noise removal algorithms[J]. Physical D, 1992, 60(1/4): 259 - 268.
- [5] CHAN T, SHEN J. Non-texture inpainting by curvature-driven diffusions (CDD) [J]. Journal of visual Communication and Image Representation, 2001, 12(4): 436 - 449.
- [6] 张红英,彭启琮. 数字图像修复技术综述[J]. 中国图象图形学报, 2007, 12(1): 1 - 10.
- [7] BERTALMIO M, VESE L, SAPIRO G, et al. Simultaneous structure and texture image inpainting[J]. IEEE Transactions on Image Processing, 2003, 12(8): 18 - 20.

(上接第946页)

增加优先权准确性和合理性,从而使破损区域填充顺序变得更加合理,修复图像的边缘的连通性和平滑性得到增强。而最佳样本块不唯一问题的解决使得修复效果进一步提高,消除了由于最佳样本块不唯一而引起的不平滑现象。但是对于破损区域比较复杂的唐卡图像修复效果不是很好,有时会被误修复,不能与人的视觉保持一致。因此,对于破损区域比较复杂的唐卡图像的修复还需要进一步的研究。

### 参考文献:

- [1] 唐卡图像介绍[EB/OL]. [2009 - 07 - 01]. <http://tangka.cn.com/>.
- [2] 白云飞,杜华. 如何修复破损唐卡[N]. 人民日报: 海外版, 2004 - 05 - 29(8).
- [3] BERTALMIO M, SAPIRO G, CASELLES V, et al. Image inpainting[C]// Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques. New York: ACM Press, 2000: 417 - 424.
- [4] RUDIN L, OSHER S, FATEMI E. Nonlinear total variation based noise removal algorithms[J]. Physical D, 1992, 60(1/4): 259 - 268.
- [5] CHAN T, SHEN J. Non-texture inpainting by curvature-driven diffusions (CDD) [J]. Journal of visual Communication and Image Representation, 2001, 12(4): 436 - 449.
- [6] 张红英,彭启琮. 数字图像修复技术综述[J]. 中国图象图形学报, 2007, 12(1): 1 - 10.
- [7] BERTALMIO M, VESE L, SAPIRO G, et al. Simultaneous structure and texture image inpainting[J]. IEEE Transactions on Image Processing, 2003, 12(8): 18 - 20.