

基于蚁群优化算法的复杂背景图像文字检测方法

李敏花, 柏 猛

(山东科技大学 电气信息系, 济南 250031)

(minhuali09@163.com)

摘 要:针对复杂背景图像中的文字检测问题,提出一种基于蚁群优化算法的复杂背景图像文字检测方法。该方法首先采用蚁群优化算法提取图像边缘;然后在边缘图像上提取特征,采取由粗到精多级检测、验证的策略进行文字检测。与基于Soble算子、Canny算子等方法的对比实验结果表明,所提出的基于蚁群优化算法的文字检测方法可有效地实现复杂背景图像中的文字检测。

关键词:文字检测;边缘检测;蚁群算法;复杂背景

中图分类号:TP391 **文献标志码:**A

Text detection from images with complex background by ant colony optimization algorithm

LI Min-hua, BAI Meng

(Department of Electrical and Information Engineering, Shandong University of Science and Technology, Jinan Shandong 250031, China)

Abstract: To detect text from images with different backgrounds, a text detection method with ant colony optimization algorithm was proposed. Before text detection, an ant colony optimization algorithm was adopted to detect image edges, and then features were extracted from the edge image. Afterwards, a coarse-to-fine strategy was applied to detect the text lines from image. Finally, the experimental results show that the proposed method achieves more precise detection than Sobel-based, Canny-based and other two detection methods.

Key words: text detection; edge detection; ant colony algorithm; complex background

0 引言

图像和视频中的文字包含丰富的信息,在视频内容分析、检索、图片内容理解等领域具有重要作用。因此,让计算机自动识别图像中的文字,即文字的自动检测、识别技术近年来逐渐引起人们的关注。

要将图像中的文字提取出来,首先需要将图像中的文字区域从复杂背景中准确定位出来,即对图像中的文字进行检测。现有文字检测技术大体可分为以下几类:基于边缘的方法^[1-2]、基于纹理的方法^[3-4]、基于颜色的方法^[5-6]和基于连通域分析的方法^[7-8]。其中,基于边缘的方法通过采用边缘检测算子,如Sobel算子、Canny算子等,检测速度快,并可以获得高的检测准确率,对于噪声较少的图像检测效果较好。但是以上边缘检测算子没有考虑人眼视觉特性,测得的边缘比较粗糙,容易出现断裂现象,尤其当图像较复杂或含有噪声时更是如此。而实际图像,尤其是复杂背景图像往往都带有噪声,这就限制了基于边缘检测方法的实用性。

鉴于此,本文提出一种基于蚁群优化算法的复杂背景图像文字检测方法。蚁群算法能够智能搜索、全局优化,而且具有鲁棒性、正反馈、分布式计算、易与其他算法结合等特点。将蚁群算法用于边缘提取,具有抗噪声能力强,边缘不易破裂等特点。在本方法中,首先采用蚁群优化算法提取图像边缘,然后在边缘图像上采取由粗到精多级检测、验证的策略进行文字检测。实验结果表明,该方法可以有效地检测出复杂背景中的文字信息。

1 基于蚁群优化算法的文字检测

基于蚁群优化算法的文字检测算法主要包括如下几部分:蚁群算法边缘检测、文本粗检测和文本验证。算法流程如图1所示。

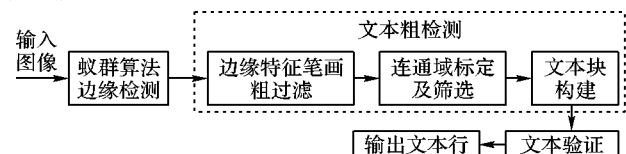


图1 基于蚁群优化算法的文字检测流程

1.1 基于蚁群优化算法的边缘检测

蚁群算法^[9]用于边缘检测的原理是:将图像抽象成一个无向图,然后将一群蚂蚁随机放置到图中,各蚂蚁将根据图像中像素点信息素强度和启发式引导函数求得概率,从其邻域像素点中选择一个,作为下一步转移的目的位置。通过多次循环迭代,最终使大多数蚂蚁聚集到图像边缘。

设要检测的图像为灰度图像 $I(i,j)$,大小为 $M \times N$,图像可看做一个具有 N 个顶点 E 条边的无向图 $G = \langle N, E \rangle$ 。顶点表示蚂蚁所在的具体位置,即图像像素点,边表示蚂蚁要走的路径。算法的目标是使蚂蚁觅食的最优路径位于图像边缘上。基于蚁群优化算法的边缘检测具体过程可概括为如下4个步骤:

1) 初始化。在初始化中,在图像 I 中随机放置 m 只蚂蚁,初始化各参数变量和像素上的信息激素初值,其中, $m \leq M \times N$ 。

收稿日期:2011-01-30;修回日期:2011-03-01。

基金项目:国家自然科学基金资助项目(61071087);山东科技大学春蕾计划资助项目(2010AZZ050;2010AZZ049)。

作者简介:李敏花(1981-),女,山东济宁人,讲师,博士,主要研究方向:图像理解与分析;柏猛(1981-),男,山东济宁人,讲师,博士,主要研究方向:人工智能、机器视觉。

N , 信息素的初值设为一个趋于0的正值。

2) 路径选择。对于二维图像, 蚂蚁从像素点 (i, j) 移动到其邻域像素点 (l, m) 的转移概率可表示为

$$p_{(i,j),(l,m)}(t) = \frac{(\tau_{l,m}(t-1))^\alpha (\eta_{l,m})^\beta}{\sum_{(l,m) \in \Omega_{(i,j)}} (\tau_{l,m}(t-1))^\alpha (\eta_{l,m})^\beta} \quad (1)$$

其中: $\tau_{l,m}(t-1)$ 为 $t-1$ 次周游迭代时像素点 (l, m) 处的信息素强度; $\Omega_{(i,j)}$ 为蚂蚁在像素点 (i, j) 处时的可选路径集合, 本文取像素点八邻域像素集合, 可选路径集合不包括蚂蚁已经走过的路径; $\eta_{l,m}$ 为节点 (l, m) 处的启发引导函数; 常数 α, β 分别代表信息素强度和启发式引导函数对蚂蚁路径选择的权重因子。

在本文中, 启发式引导函数取为像素 (l, m) 8 邻域内灰度梯度的最大值, 如式(2)所示。由启发式引导函数的选择可知, 图像边缘上的启发引导函数值比较大, 所以蚂蚁选择边缘的概率比较大。

$$\eta_{l,m} = \frac{1}{I_{\max}} \times \max[|I(l-1, m-1) - I(l+1, m+1)|, |I(l-1, m+1) - I(l+1, m-1)|, |I(l, m-1) - I(l, m+1)|, |I(l-1, m) - I(l+1, m+1)|] \quad (2)$$

3) 信息素更新。在本文中, 信息素要进行局部和全局两次更新。其中, 在局部更新时, 设在 t 次周游迭代过程中, 蚂蚁 k 每走一步, 信息素更新一次, 则更新规则可表示为

$$\tau_{i,j}(t) = \begin{cases} (1-\rho) \cdot \tau_{i,j}(t-1) + \rho \cdot \Delta\tau_{i,j}^k, & (i,j) \text{ 被蚂蚁 } k \text{ 访问} \\ \tau_{i,j}(t-1), & \text{其他} \end{cases} \quad (3)$$

其中: ρ 为信息素挥发系数; $\Delta\tau_{i,j}^k$ 的定义与启发引导函数有关, 本文取 $\Delta\tau_{i,j}^k = \eta_{i,j}$ 。

在所有蚂蚁都周游迭代完成一次后进行全局更新, 其更新规则可表示为

$$\tau(t) = (1-\varphi) \cdot \tau(t-1) + \varphi \cdot \Delta\tau \quad (4)$$

其中 φ 为信息素衰减系数。

4) 边缘点的确定。当蚂蚁走完规定的周游迭代次数时, 通过设定一阈值 T , 将每个像素点上的最终信息素值 $\tau_{i,j}$ 与 T 进行比较, 确定该像素点是边缘点或非边缘点。若像素点上信息素数值大于阈值 T , 则判断为边缘点, 反之, 为非边缘点。边缘点的判定可用下式表示:

$$E_{i,j} = \begin{cases} 1, & \tau_{i,j} > T \\ 0, & \text{其他} \end{cases} \quad (5)$$

综上所述, 基于蚁群算法优化的边缘检测算法流程可用图2表示。

1.2 文本粗检测

1) 文字笔画粗过滤。字符区域通常包含很多笔画, 相应的, 这些笔画会产生大量边缘。从纹理分析的角度看, 与大多数背景区域相比, 文字区域的边缘会更密集。基于此, 本文采用滑动窗口在边缘图像上抽取边缘密度特征 (Edge Density Feature, EDF) 对字符笔画进行粗过滤。边缘密度特征抽取方法可表示为:

$$D_e = \frac{1}{H \cdot W} \sum_{i=1}^H \sum_{j=1}^W f(i, j) \quad (6)$$

其中: H, W 分别为滑动窗口的长度和宽度; $f(i, j)$ 为边缘图像上像素的灰度值, 由于边缘密度特征在二值图像上抽取, 故

$f(i, j) = 0$ 或 255 。

2) 连通域标定及筛选。为了将过滤后的边缘图像中的像素点组合成具有逻辑意义的文字区, 需要进行连通域标定。对于文字检测来说, 4-邻域连通标定和 8-邻域连通标定, 对最后的检测结果影响很小。为了提高算法的运行效率, 本文采用 4-邻域连通域标定法进行连通域标定。进行连通域标定后, 通过筛选去掉那些与字符连通元有明显不同的背景物体连通元, 以减少后续处理中连通元数目。本文采用几何特征、边缘特征和形状特征^[10]等进行连通元筛选。

3) 文本块构建。在文本块构建中, 本文主要采用简单的几何约束来组建文字区域, 这些约束包括连通元尺寸、相邻连通元的距离和重叠程度等。

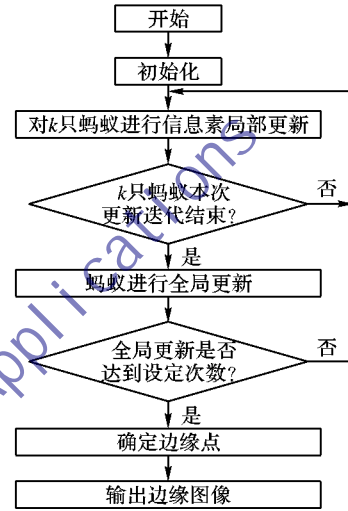


图2 基于蚁群优化算法的边缘检测流程

1.3 基于 SVM 的文本块验证

粗检测结束后, 本文采用文本块验证的方法去除误检文本块, 以提高检测准确率。在文本块验证阶段, 本文主要采用基于支持向量机 (Support Vector Machine, SVM) 的多特征验证策略对在粗检测阶段得到的候选文本块进行验证。用到的特征主要有边缘特征、投影特征、直方图特征和共生矩阵特征等^[11]。算法中用到的 SVM 分类器采用 LIBSVM, 核函数采用多项式核函数。

2 实验及结果分析

为了验证本文所提出方法的有效性, 本文设计了两部分实验: 一是文字图像边缘检测比较实验, 二是文字检测实验。

2.1 边缘检测比较实验

在该实验中, 将基于蚁群优化算法的边缘检测方法分别与 Sobel 边缘检测方法、Canny 边缘检测方法进行比较。其中, 基于蚁群优化算法的边缘检测中各参数设置如下:

蚂蚁数量 $m = \sqrt{M \times N}$, M, N 分别为图像的高度和宽度; 参数 $\alpha = 2, \beta = 0.6$; 初始信息素 $\tau = 0.0001$; 信息素挥发系数 $\rho = 0.1$; 信息素衰减系数 $\varphi = 0.05$; 最大循环周游迭代次数 $n = 4$ 。

图3所示为三种不同的方法在文字图像上的实验结果。由实验结果看出, 基于 Sobel 算子和 Canny 算子的边缘检测算法, 对噪声比较敏感, 在保留了文字边缘的同时, 还保留了很多噪声边缘, 而且造成过多的边缘破裂。而基于蚁群优化的边缘检测算法, 可较完整地保留文字边缘, 且能去除大部分背景边缘。由此可见, 基于蚁群优化算法提取的图像边缘, 更适

合复杂背景中的文字检测。

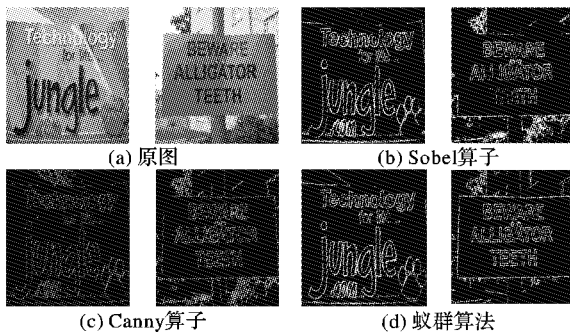


图3 不同方法边缘检测比较

2.2 文字检测实验

为了检验蚁群算法用于文本检测的有效性,本文在ICDAR2005 竞赛提供的数据集上进行文本检测实验。文字检测采用召回率(recall rate)和准确率(precision rate)作为评价指标,其定义如下。

$$\text{召回率} = \frac{\text{正确检测出的文本行个数}}{\text{文本行总个数}}$$

$$\text{准确率} = \frac{\text{正确检测出的文本行个数}}{\text{检测出的文本行总个数}}$$

图4列出了采用本文提出的方法进行文本检测的一些结果示例。可以看出,本文提出的方法能够比较准确地定位出复杂背景中的文字。

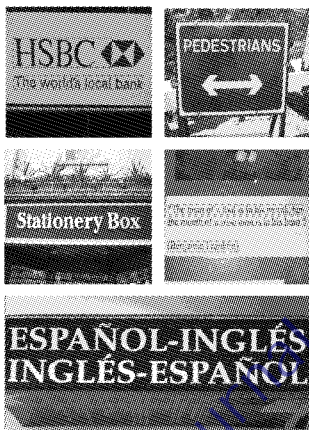


图4 文字检测结果示例

表1列出了本文提出的文字检测方法分别与基于Sobel算子、Canny算子的方法进行文字检测比较,并且与在ICDAR2005 文字定位竞赛中获得第一名的Hinnerk 和第二名的AlexChen^[12]的方法进行比较。

表1 ICDAR2003 图像集上的检测结果

算法	准确率	召回率	运行时间/s
本文方法	0.69	0.66	1.87
基于Canny算子方法	0.58	0.65	1.66
基于Sobel算子方法	0.65	0.62	1.52
Hinnerk方法	0.62	0.67	14.40
AlexChen方法	0.60	0.60	0.35

从表1的实验结果可看出,本文提出的文字检测方法,准确率和召回率均高于基于Canny算子和Sobel算子的文字检测方法。虽然本文提出的方法召回率比ICDAR2005 竞赛最好的结果低了1个百分点,准确率却高出7个百分点。可见,采用本文所提出的基于蚁群优化的文字检测方法可较好地实现复杂背景中的文字检测。本文提出的方法单幅图像的平均处理

时间为1.87 s,还不能满足实时应用的要求。下一步我们需要优化算法,提高文字检测速度。

3 结语

针对复杂背景图像中的文字检测问题,本文提出了一种基于蚁群优化算法的复杂背景图像文字检测方法。通过将本文方法与其他边缘检测方法进行边缘检测比较,实验结果表明,采用蚁群优化算法进行边缘检测,能够较完整地保留文字边缘,同时去除大部分背景边缘;与其他文字检测方法进行比较的实验结果表明,本文所提出的方法能够获得较好的文字检测准确率和召回率,有效地检测出复杂背景中的文字。

参考文献:

- [1] CHEN D, BOURLARD H, THIRAN J P. Text identification in complex background using SVM [C]// Proceedings of the 2001 IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2001, 2: 621 - 626.
- [2] PALAIAHNAKOTE S, HUANG W, LIM T. An efficient edge based technique for text detection in video frames [C]// The Eighth IAPR Workshop on Document Analysis Systems. Washington, DC: IEEE Computer Society, 2008: 307 - 314.
- [3] WU V, MANMATHA R, RISEMAN E M. Text finder: An automatic system to detect and recognize text in images[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999, 21(11): 1224 - 1229.
- [4] GLAVATA J, EWERTH R, FREISLEBEN B. Text detection in images based on unsupervised classification of high-frequency wavelet coefficients [C]// Proceedings of 17th International Conference on Pattern Recognition. Washington, DC: IEEE Computer Society, 2004: 425 - 428.
- [5] SRIVASTAVA A, KUMAR J. Text detection in scene images using stroke width and nearest-neighbor constraints [C]// Proceedings of IEEE TENCON 2008. Washington, DC: IEEE Computer Society, 2008: 1 - 5.
- [6] KIM P-K. Automatic text location in complex color images using local color quantization [C]// TENCON 99: Proceedings of the IEEE Region 10 Conference. Washington, DC: IEEE Computer Society, 1999: 629 - 632.
- [7] MANCAS-THILLOU C, GOSSELIN B. Spatial and color spaces combination for natural scene text extraction [C]// Proceedings of 2006 International Conference on Image Processing. Washington, DC: IEEE Computer Society, 2006, 1: 985 - 988.
- [8] WANG K, KANGAS J. Character Location in scene images from digital camera [J]. Pattern Recognition, 2003, 36(10): 2287 - 2299.
- [9] LEE M, KIM S, CHO W, et al. Segmentation of brain MR images using an ant colony optimization algorithm [C]// Proceedings of the 9th IEEE International Conference on Bioinformatics and Bioengineering. Washington, DC: IEEE Computer Society, 2009: 366 - 369.
- [10] ZHU KAIHUA, QI FEIHU, JIANG RENJIE, et al. Automatic character detection and segmentation in natural scene images[J]. Journal of Zhejiang University: Science A, 2007, 8(1): 63 - 71.
- [11] YE QIXIANG, HUANG QINGMING. A new text detection algorithm in images/video frames [C]// PCM 2004: 5th Pacific Rim Conference on Multimedia, LNCS 3332. Berlin: Springer-Verlag, 2004: 858 - 865.
- [12] SIMON M. ICDAR2005 Text locating competition results[C]// ICDAR2005: Proceedings of the 8th International Conference on Document Analysis and Recognition. Washington, DC: IEEE Computer Society, 2005, 1: 80 - 84.