

基于阈值和区域生长法的原油 TLC 图谱分割法*

史健婷¹, 邹艳荣², 刘辉¹, 陈荣丽¹

(1. 黑龙江科技学院计算机与信息工程学院, 哈尔滨 150027

2. 中国科学院广州地球化学研究所有机地球化学国家重点实验室, 广州 510640)

摘要: 图像分割是数字图像处理中的一项关键技术。原油组分薄层色谱图的精确分割是后续各个组分定量计算与分析的基础。提出一种基于阈值和区域生长法相结合的分割方法, 对预处理后的薄层色谱图进行分割。通过与单一阈值法的比较, 结果表明该方法在提取目标的同时, 能够有效地分离背景像素, 提高分割精度。同时, 该方法运算简单, 能进一步提高分割的效率。

关键词: 阈值法; 区域生长法; 薄层色谱图; 定量计算

0 引言

图像分割是数字图像处理中的一项关键技术。它是指把图像分成各具特性区域的技术和过程。这里所说的特性可以是灰度、颜色、纹理等, 而目标可以对应单个区域, 也可以对应多个区域。通过图像分割, 可减少图像分析与识别的数据量, 不改变图像的结构特征。只有在分割的基础上才能对目标进行特征提取。图像分割结果的好坏直接影响对图像的理解。因此, 图像分割是由图像处理到图像分析的关键步骤, 在图像工程中占据重要的位置。

自 20 世纪 70 年代至今, 已提出上千种的分割算法。例如: 门限法^[1]、匹配法^[2]、区域生长法^[3]、马尔可夫随机场模型法^[4]、小波分析法^[5]等。常用的分割方法有灰度阈值分割法^[6]、区域生长法和分裂合并法^[7]。

原油组分的薄层色谱图(Thin Layer Chromatography, TLC)的精确分割是后续各个组分定量计算与分析的基础。本文提出一种基于阈值分割法和区域生长法

相结合的方法进行分割, 避免了过分割或欠分割问题。同时, 该方法运算简单, 能进一步提高分割的效率。

1 阈值分割法

阈值分割法是最常见的一种直接检测区域的分割方法^[8]。其思想是: 设定不同的特征阈值, 对图像像素点分类。由原始灰度或彩色值变换得到的特征和直接来自原始图像的灰度或彩色特征是其常用的两个特征。设原始图像为 $f(x, y)$, 找到特征值 T , 按照一定的准则将图像分割为两个部分, 分割后的图像为:

$$g(x, y) = \begin{cases} b_0 f(x, y) < T \\ b_1 f(x, y) \geq T \end{cases} \quad (1)$$

这里特征阈值的获取是关键步骤。最常用并且性能评估最好的一种方法叫做最大类间方差法(Otsu法)^[9]。其基本思路是对一幅图像 I , 特征值 T 为前景与背景的分割阈值, 前景点数占图像比例为 ω_0 , 平均灰度为 u_0 ; 背景点数占图像比例为 ω_1 , 平均灰度为 u_1 。图像的总平均灰度为: $u_T = \omega_0 \times u_0 + \omega_1 \times u_1$ 。从最小灰度到最大灰

* 基金项目: 国家自然科学基金项目(No.41273059)、黑龙江省教育厅科学技术研究项目(No.12513077)、黑龙江省教育厅面上项目(No.12521475)

收稿日期: 2013-02-01 修稿日期: 2013-02-04

作者简介: 史健婷(1981-), 女, 黑龙江鸡西人, 讲师, 硕士, 研究方向为图像处理与模式识别

度值遍历 T , 当 T 使得方差值 $\sigma^2 = \omega_0 \times (u_0 - u_T)^2 + \omega_1 \times (u_1 - u_T)^2$ 最大时, T 为分割的最佳阈值。方差是灰度分布均匀性的一种度量, 方差值越大, 说明构成图像的两部分差别越大, 也意味着错分概论最小。

阈值分割法虽然是常用的一种分割方法, 但其不足在于会在分割出的目标中留下和目标灰度接近的背景区域, 从而引入噪声。

2 区域生长法

区域生长法是一种将处理过程分解为多个顺序步骤的串行方法^[10]。它的基本思想是以选定的生长点作为目标的起始像素; 选择搜索策略, 从该生长点出发, 根据一定机理依次检测生长点附近的像素, 判断是否满足一致性判别条件; 设定终止条件, 完成目标的分割。

如采用灰度均值作为判别依据的话, 一致性判别条件见式(2)。

$$T_2 = (1 - \frac{\sigma}{m}) \times T_c \quad (2)$$

生长准则见式(3):

$$|\bar{f}(k, l) - m| \leq T_2 \quad (3)$$

其中 $\bar{f}(k, l)$ 为待测生长点邻域内的灰度均值。

例如矩阵(a)是需要分割的图像矩阵, 假设已知有两个种子像素(用不同颜色标记), 现在要进行区域生长。现采用的判断标准是: 如果所考虑的像素与种子像素灰度值差的绝对值小于某个门限 T , 则将该像素包括进种子像素所在的区域。示例如图 1 所示。矩阵(a)为待分割的图像, 矩阵(b)给出了 $T=3$ 的区域生长结果, 可以看到有些像素无法判定; 矩阵(c)给出 $T=4$ 时的区域生长结果, 可以看到整幅图被较好地分成了两个区域; 矩阵(d)给出 $T=6$ 的区域生长结果, 整幅图都

被分在一个区域中了。如图 1 所示。由此可知, 实现区域生长的关键是生长点的选取, 生长规则的确定和区域生长停止条件的设定。

利用区域生长法进行图像分割可减少噪声的干扰。但是也存在着不足, 即会造成欠分割和过分割。欠分割是由于目标区灰度分布不均匀, 生长提前终止引起的; 过分割是由于目标边界模糊, 生长终止延后造成的。

两种方法在单独使用时均难以取得令人满意的分割效果。因此, 考虑将二者结合起来进行图像的分割。

3 基于阈值和区域生长法的图像分割算法

将两种方法结合起来进行原油组分薄层色谱图像分割的算法描述如下:

(1) 用最大类间方差法获取最佳灰度阈值 T 。直接应用最大类间方差法计算量较大, 因此, 在程序实现时采用了等价的公式 $\sigma^2 = \omega_0 \times \omega_1 (u_0 - u_1)^2$ 。由前面所述的思路求得使 σ^2 最大的 T 值即为最佳灰度阈值。

(2) 作为约束条件, 将该阈值应用到区域生长法的生长准则中。令种子区域为 R ; 其均值、标准方差 m, σ , 作为衡量灰度一致性的特征量。

$$m = \frac{1}{n} \sum_{(k, l) \in R} f(k, l),$$

$$\sigma = \sqrt{\frac{1}{n} \sum_{(k, l) \in R} |f(k, l) - m|^2} \quad (4)$$

n 为区域的像素数。由灰度阈值变化定义生长准则, 即:

$$\text{生长准则} \begin{cases} |\bar{f}(k, l) - m| \leq T(k, l) \\ \bar{f}(k, l) \leq T(\text{约束条件}) \end{cases} \quad (5)$$

若像素灰度值 $f(k, l)$ 与 m 接近, 同时满足生长准则, 则 (k, l) 可并入区域 R , 同时更新 m, σ 。

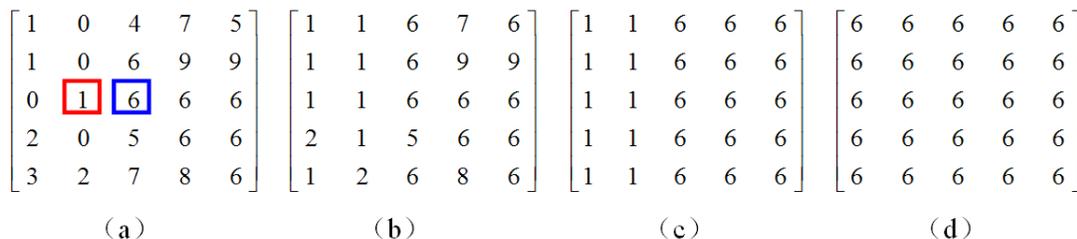


图 1 区域生长示例图

$$T(k,l) = (1 - \frac{\sigma}{m}) \times T_c \quad (6)$$

T_c 是人为设定的反映生长条件苛刻程度的量,受到 T 的约束,并在区域生长的过程中可动态调整。

(3) 继续检测新的邻域点,直至无法生长为止,完成对目标区域的分割。在进行区域生长时,新的测试像素是否并入生长区受到生长规则和灰度阈值 T 的双重制约,可以避免造成欠分割和过分割这两种现象。

传统生长方式中存在着冗余计算的问题,为了减少搜索过程的计算量,本文采用一种新的生长路径规则,可以提高分割速度。首先,从某一生长点 G 开始垂直向上搜索满足生长条件的像素点,并找到该方向上的上限点;之后再沿该生长点垂直向下搜索,找到该方向上的下限点;以该生长点为起点的垂直方向搜索完毕。然后水平向左移动取其左邻接点为新的生长点,继续沿垂直方向上下搜索。不断向左重复该过程,直至左侧无新的生长点出现。完成区域的左侧搜索以后,再以 G 为中心,水平向右移动生长点,取其右邻接点为新的生长点,沿垂直方向上下搜索,不断向右重复该过程,直至右侧无新的生长点出现。此时完成了区域的生长过程。

4 实验结果与分析

根据上述算法原理,利用 MatLab 与 VC++ 6.0 混合编程实现了基于阈值和区域生长法的图像分割算法。本实验对乌 55494、乌 39497、乌 39496、文昌凹陷、孤北、牙哈、塔中 62 井、轮南八个不同油田/油井的原油样品进行薄层色谱分离,薄层色谱的操作流程分为选择固定相、制板、点样、展开、显色定位,得到各个样品的薄层色谱图,图像的获取操作均在中国科学院广州

地球化学研究所有机地球化学国家重点实验室完成。以乌 39496 为例所得的原油样品薄层色谱图如图 2 所示。得到原始的图像之后,还需要利用计算机图像处理技术对其进行预处理,消除影响图像质量的噪声、背景,以免为后续的定量分析带来干扰。同时,为了获取图像中的有用部分,还需要将图像中的目标区域导出,并提取 RGB 图像中的任意通道(多选绿色通道)进行分析。得到去噪后的图像如图 3 所示。利用单一阈值法进行分割,得到的图像如图 4 所示。利用本文提出的算法进行分割,得到的图像如图 5 所示。

图 4 和图 5 分别是采用单一阈值法和本文算法分割后的结果。利用背景颜色与组分颜色的差异可以将图像分割为两个区域。图中白色的背景区域是一类物质(显色剂),黑色区域是另外一类物质(含有不同组分的原油)。为了显示白色背景区域,在两图中分别加入了蓝色边框。通过对比发现,单一阈值法仅根据像素的灰度特征进行图像分割,对于目标与背景的阈值范围存在交叉、重叠的图像而言,不可避免地将目标区域与背景错分,造成一定的分割误差;而本文算法不仅考虑了图像的灰度变化特性,还考虑了像素的位置,使区域生长尽可能接近区域的原始边界,从而获得更好的分割质量和精度。

对图 5 中的分割结果可以进一步的分类,把原油中的各个组分识别出来,已知各个原油组分在薄层色谱法中显色的速度不同,可知原油各个组分的迁移距离不同。在图 5 中体现的是各个黑色小斑点的高度不同,这样就可以利用四个黑色子区域距离图像底部的距离特征进行原油组分的分类,图 5 的进一步分类结果如图 6 所示,图中白色背景区域仍然是显色剂,蓝色

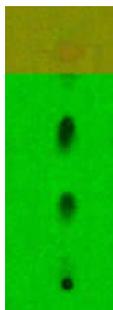


图 2 “乌 39496”原油薄层色谱图



图 3 预处理后的薄层色谱图

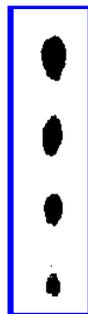


图 4 单一阈值法分割的结果

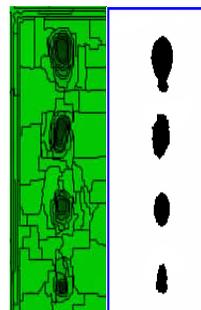


图 5 本文算法分割后的结果

区域是饱和烃,绿色区域是芳烃,红色区域是非烃,黑色区域是沥青质。

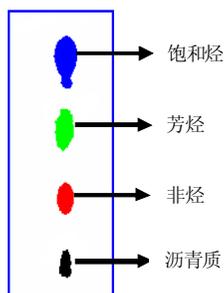


图 6 分割后的分类图像

5 结 语

本文根据阈值分割法和区域生长法的特点,提出了基于阈值分割和区域生长相结合的图像分割算法,对原油组分的薄层色谱图进行了分割,有效地解决了灰度分布存在交叉、重叠的组分目标与介质背景的分离问题,提高了分割精度。再根据分类结果确定各个组分的分析区域,并准确识别出原油中的各个组分,为后续的定量计算与分析打下了基础。文中所提的分割方法对于薄层色谱这类平面色谱图像的分割具有一定的实际应用价值。

参考文献

- [1]沈庭芝,方子文. 数字图像处理及模式识别[M]. 北京理工大学出版社,1998
- [2]李久权,王平,王永强等. CT 图像分割几种算法[J]. 微计算机信息,2006,22(21):240~242
- [3]罗文村. 基于阈值法和区域生长法综合集成的图像分割法[J]. 现代计算机,2001,5(25):13~14
- [4]Liu Bo, Cheng Heng-da, Huang Jian-hua. Fully Automatic and Segmentation-Robust Classification of Breast Tumors Based on Local Texture Analysis of Ultrasound Images [J]. Pattern Recognition, 2010(43):280~298
- [5]侯艳丽. 融合多特征的纹理图像分割算法[J]. 计算机技术与发展, 2012(5)
- [6]Cheng heng-da, Shan juan, Ju wen. Automated Breast Cancer Detection and Classification Using Ultrasound Images: A Survey[J]. Pattern Recognition, 2010(43): 299~317
- [7]何斌,马天予等. Visual C++数字图像处理(第二版)[M]. 人民邮电出版社,2002
- [8]张益针贞,刘帧. Visual C++实现 MPEG/JPEG 编码技术[M]. 人民邮电出版社,2002
- [9]张德丰. 详解 MatLab 数字图像处理[M]. 电子工业出版社, 2010
- [10]黄谊,任毅. 基于阈值法和区域生长法的图像分割算法研究[J]. 电子测试,2010,10(10):23~25

A Segmentation Algorithm for Thin Layer Chromatograph Image of Crude Oil Based on Threshold and Region Growing

SHI Jian-ting¹, ZOU Yan-rong², LIU Hui¹, CHEN Rong-li¹

- (1. College of Computer and Information Engineering, Heilongjiang Institute of Science and Technology, Harbin 150027;
2. State Key Laboratory of Organic Geochemistry, Guangzhou Institute of Geochemistry, Chinese Academy of Sciences, Guangzhou 510640)

Abstract: Image segmentation is a key technique of digital image processing. Analyses the crude oil group types and calculates the percentage of each group based on the accurate segmentation. Proposes a novel segmentation algorithm based on Threshold and Region Growing. Compared with the single threshold segmentation, the experiment indicated that the segmentation method presented has the advantage of extracting the target without any pixel of the background. In the meanwhile, with simple operation, the method can improve the efficiency of segmentation.

Keywords: Threshold Method; Region Growing Method; Thin Layer Chromatograph; Quantitative Calculation