

基于改进 K-means 聚类与分水岭的木材横截面管孔分割

程昱之¹, 钟丽辉², 何鑫², 王远¹, 李朝岚¹

(1. 西南林业大学 机械与交通学院, 云南 昆明 650224; 2. 西南林业大学 大数据与智能工程学院, 云南 昆明 650224)

摘要: 【目的】管孔是木材识别方面的重要特征之一。针对管孔随机分布、大小不一导致管孔分割鲁棒性不高, 木纤维、木射线以及轴向薄壁组织等噪声区域对管孔分割效果影响较大的问题, 本研究提出了一种改进 K-means 聚类与分水岭的木材横截面管孔分割算法。【方法】采用改进 K-means 聚类对管孔区域进行粗分割, 有效区分管孔区域与木纤维、木射线以及轴向薄壁组织等噪声区域。再对粗分割结果采用改进分水岭算法进行精分割, 分割出的管孔与实际管孔基本吻合。【结果】平均每张木材横截面微观图像有 97.1% 的管孔被准确有效地分割出来。本研究提出的改进分割算法与其他算法相比, 分割效果显著提升, 在大小不一且随机分布的管孔分割过程中鲁棒性高, 具有良好的分割性能。【结论】该算法有效解决了传统 K-means 聚类算法在图像分割时受噪声影响大和初始聚类中心随机性问题, 为阔叶材管孔特征提取和定量分析奠定了坚实基础。图 7 表 1 参 16

关键词: K-means 聚类; 分水岭算法; 管孔; 图像分割

中图分类号: S781.1 文献标志码: A 文章编号: 2095-0756(2022)01-0173-07

Segmentation of wood cross-section pores based on improved K-means clustering and watershed

CHENG Yuzhi¹, ZHONG Lihui², HE Xin², WANG Yuan¹, LI Chaolan¹

(1. College of Machinery and Transportation, Southwest Forestry University, Kunming 650224, Yunnan, China;

2. College of Big Data and Intelligent Engineering, Southwest Forestry University, Kunming 650224, Yunnan, China)

Abstract: [Objective] Pore is one of the important characteristics of wood identification. Aiming at the problems of poor robustness of pore segmentation due to random distribution and different sizes of pores and the impact of noise such as wood fiber, wood ray and axial parenchyma on the segmentation effect of pores, this study attempts to propose an improved K-means clustering and watershed algorithm for segmentation of wood cross-section pores. [Method] The improved K-means clustering was used to segment the pore area, which could distinguish the pore area from noise areas such as wood fiber, wood ray and axial parenchyma. Then, an improved watershed algorithm was used to segment the rough segmentation results, and the segmented pores were basically consistent with the actual pores. [Result] On average, 97.1% of the pores were segmented in each microscopic image of wood cross section. Compared with other algorithms, the improved algorithm had a significantly improved segmentation effect, high robustness and good segmentation performance in the process of pore segmentation with different sizes and random distribution. [Conclusion] This algorithm can effectively solve the problems of noise impact and randomness of initial clustering center in traditional K-means clustering

收稿日期: 2021-03-11; 修回日期: 2021-07-27

基金项目: 云南省农业基础研究联合专项 (2018FG001-108)

作者简介: 程昱之 (ORCID: 0000-0001-5036-1873), 从事木材图像处理研究。E-mail: 2848514993@qq.com。通信作者: 钟丽辉 (ORCID: 0000-0001-9630-1622), 从事木材图像处理研究。E-mail: zhongsimple@sina.com

algorithm in image segmentation, and lays a solid foundation for the feature extraction and quantitative analysis of hardwood pores. [Ch, 7 fig. 1 tab. 16 ref.]

Key words: K-means clustering; watershed algorithm; pore; image segmentation

阔叶材横截面微观图像由管孔、木射线、生长轮、木纤维和轴向薄壁组织构成,其中管孔是木材研究最重要的特征之一^[1-3]。管孔不仅是分析考古木材样品的构造特征与保护状态的研究要素^[3],也是研究某一地区若干年内木材生长量和环境气候变化相关性的依据^[4]。管孔的分割是微观木材图像中提取管孔特征量和进行木材树种自动识别的关键步骤^[5-6],因此,微观木材图像中的管孔分割有重要的研究意义。近些年来应用于管孔分割的算法主要有数学形态学、区域生长、目标遗传和水平集等^[7-9]。张广群等^[9]首先运用多目标遗传算法成功地分割出图像中的大部分管孔,但会出现复管孔或管孔团的漏分,同时部分轴向薄壁组织会连同其附近的管孔一起被分割出来。汪杭军^[10]运用局部水平集算法,以局部灰度分布的均值作为统计信息,加强了管孔区域和背景区域的对比度,该算法虽然克服了漏分问题,但出现了噪声的误分割,一部分噪声为一些轮廓较大的木射线组织,另一部分为形状大小与管孔相似的木纤维组织。祁亨年等^[11]运用改变结构元素和形态尺度的数学形态学算法对管孔进行分割,该算法克服了木射线和木纤维组织等噪声的影响,但出现了图像边缘细节丢失,如较小的管孔被当作噪声去除或是较大的管孔部分组织缺失。SHEN等^[12]通过检测闭合区域的平均面积得出最适合的结构元素,并运用形态学算法对管孔进行分割,但还是会出现小孔漏分,大孔误分的情况。由于木材横截面微观图像中管孔数量众多、形态各异且随机分布,木纤维和轴向薄壁组织横截面也呈现一个个相对较小的孔洞,从而对管孔进行分割较为困难。综上所述,微观木材图像中管孔的分割依然是目前的研究热点和难点。本研究提出了一种改进 K-means 聚类与分水岭的木材横截面管孔分割算法,旨在解决管孔漏分、误分和噪声去除的问题并快速分割出管孔区域。

1 数据与方法

1.1 数据来源

日本森林数据库 (<https://db.ffpri.go.jp/WoodDB/JWDB/home.php>)。该数据库成立于 1928 年,现有木材标本 29 000 余号,隶属 270 科 2 050 属 8 500 种;切片标本约 90 000 号^[13]。本研究用到的木材树种为大红叶枫 *Acer amoenum*(15 帧图像)、尖齿槭 *A. argutum*(5 帧)、细柄槭 *A. capillipes*(4 帧)、茶条槭 *A. ginnala*(1 帧)、桑叶槭 *A. morifolium*(3 帧)、毛果槭 *A. nikoense*(3 帧)、日本槭 *A. nipponicum*(5 帧)、赛波德槭 *A. sicboldianum*(24 帧)、花楷槭 *A. ukurunduense*(6 帧)、红脉槭 *A. rufinerve*(25 帧)、薄叶槭 *A. tenuifolium*(3 帧)和褐枝猕猴桃 *Actinidia rufa*(7 帧)。

1.2 改进的 K-means 聚类算法

传统的 K-means 聚类算法以距离作为相似性的评价指标,基本思想是按照距离将样本聚成不同的簇,两点距离越近,相似度就越大,以得到紧凑且独立的簇作为聚类目标^[14-15]。传统的 K-means 聚类算法将样本映射到 RGB 颜色特征空间,构造新的映射关系,能够打破聚类只能是球状的限制,但聚类中心的选择会较大程度上影响分类效果,尤其是当聚类中心的选取全部位于管孔以外时,管孔区域会被整体归为噪声。此外,由于初始聚类中心的随机生成,可能会导致分类结果不一致或将木纤维或轴向薄壁组织被整体分割出来。基于此,针对以上问题进行算法改进,实现聚类中心的自适应选取。在 RGB 色彩空间下提取彩色木材微观图像的三通道分量。结合木材横截面原图,能够看出生长轮和木射线区域颜色较深,颜色值较小,所以其对应颜色值较小的峰值附近区域;管孔区域颜色较浅,颜色值较大,所以其对应颜色值较大的峰值附近区域;轴向薄壁组织和木纤维区域颜色较为平均,颜色值介于前两者之间,所以其对应两峰之间的波谷附近区域。分别取 3 个分量的均值组成第 1 个中心点,代表颜色直方图波谷附近区域,该点作为轴向薄壁组织和木纤维区域的簇类中心;遍历三通道分量,对于大于各自均值的所有颜色值取中值作为第 2 个中心点,由于该点对应颜色值较大,所以该点作为管孔的簇类中心;小

于各自均值的所有颜色值取中值作为第 3 个中心点, 由于该点对应颜色值较小, 所以作为生长轮和木射线区域的簇类中心, 分别代表颜色直方图两峰峰值附近区域。设通道分量为 M_j , $M_j \in (r, g, b)$, X_n 处的颜色值为 I_n , $0 < I \leq 255$ 。计算每个通道分量的均值 $\overline{M}_j$, 均值两侧的中值为 M_1 和 M_2 , Meidan 表示取目标数组的中值函数。

$$\overline{M}_j = \frac{1}{n} \sum_{n=1}^n I_n; \quad (1)$$

$$M_1 = \text{Meidan}(0, \overline{M}_j); \quad (2)$$

$$M_2 = \text{Meidan}(\overline{M}_j, 255)。 \quad (3)$$

根据图像中各组织部分对应的颜色峰值或波谷得出这一组织部分的中心点, 再以这 3 个点作为初始聚类中心点, 运用 K-means 算法实现管孔的粗分割。

1.3 改进的分水岭算法

虽然 K-means 聚类算法能够将木材微观图像中的管孔区域和背景区域区分开, 但两者的颜色梯度差异较小, 这可能会导致分水岭变换的过度分割。为了增加管孔区域和背景区域的颜色梯度差异, 使分水岭变换在梯度运算判断极大值点时更加准确, 更加有效地实现管孔的精分割, 按照聚类的结果将木材微观图像进行 3 类颜色编码。第 1 类轴向薄壁组织细胞壁和木纤维区域进行红色编码, 第 2 类管孔区域进行绿色编码, 第 3 类生长轮和木射线区域进行蓝色编码。最后, 通过形态学操作和分水岭变换实现管孔的精分割。

1.4 管孔分割

为实现木材横截面微观图像的管孔分割, 具体步骤如下 (图 1): ①提取木材横截面微观图像的 RGB 分量并绘制颜色直方图; ②根据木材微观图像中的 3 类目标区域, 即轴向薄壁组织和木纤维、管孔、生长轮和木射线, 确定聚类初始中心点; ③根据像素点与聚类中心点之间的欧式距离求出新的簇类中心并判断中心点是否改变, 若不一致则继续迭代, 一致则进行下一步; ④对划分出的 3 类区域进行颜色编码; ⑤最后通过形态学操作去除面积较小的噪声, 运用分水岭算法分割出管孔。

1.5 验证

以大红叶枫横截面原图 (997×1 418) 为例, 将本研究算法和现有算法 (阈值分割、分水岭、K-means 聚类与阈值分割) 的分割效果进行对比分析, 并以平均管孔有效分割率 (F) 作为算法分割性能的评价指标进行评价。

$$P_i = \frac{T_i}{N_i}; \quad (4)$$

$$F = \frac{1}{n} \sum_{i=1}^n P_i。 \quad (5)$$

式 (4)~(5) 中, T_i 表示每张图像的管孔实际分割总数, N_i 表示每张图像中的管孔总数, P_i 表示每张图像的管孔有效分割率, 重复次数 n 的取值为 101,

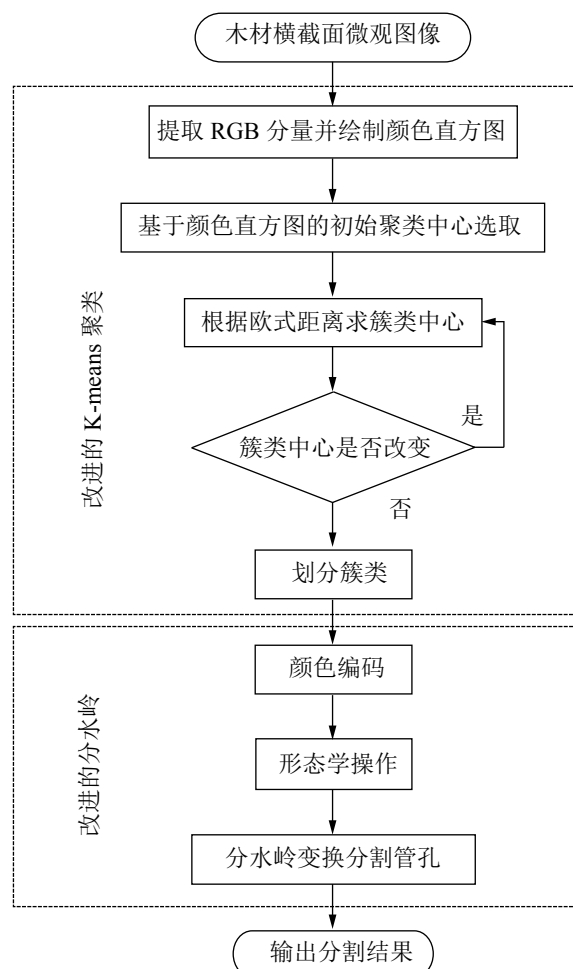


图 1 改进 K-means 与分水岭算法流程

Figure 1 Improved K-means and watershed algorithm flow

$i \in n$, F 表示平均管孔有效分割率。

2 结果与分析

2.1 算法实现

随机选择 3 种木材微观横截面原图 (图 2)。在 RGB 色彩空间下提取彩色木材微观图像的三通道分量 (图 3)。可以看出 3 个分量均有 2 个峰和 1 个波谷组成, 其中 2 个峰分别代表管孔区域和生长轮、木射线区域, 波谷代表轴向薄壁组织和木纤维区域。

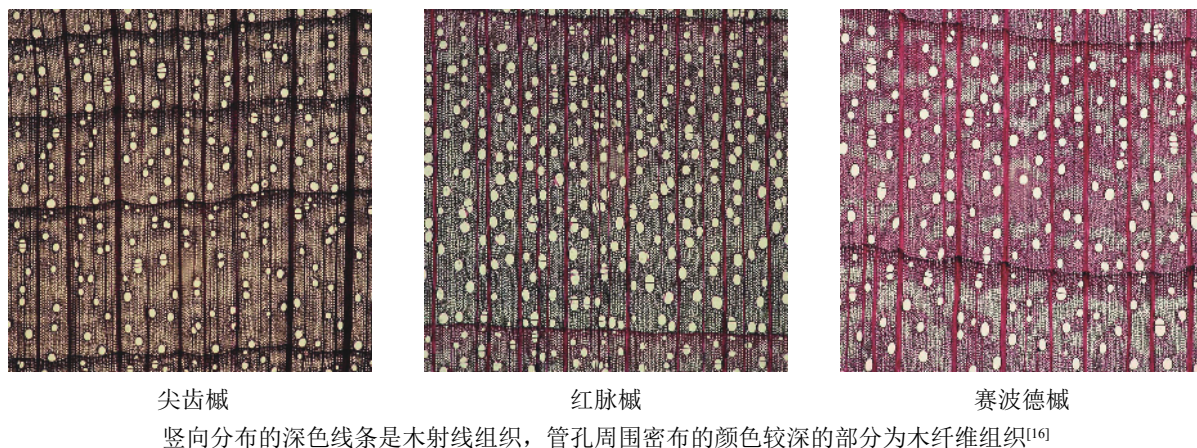


图 2 木材横截面原图

Figure 2 Original cross-section of wood

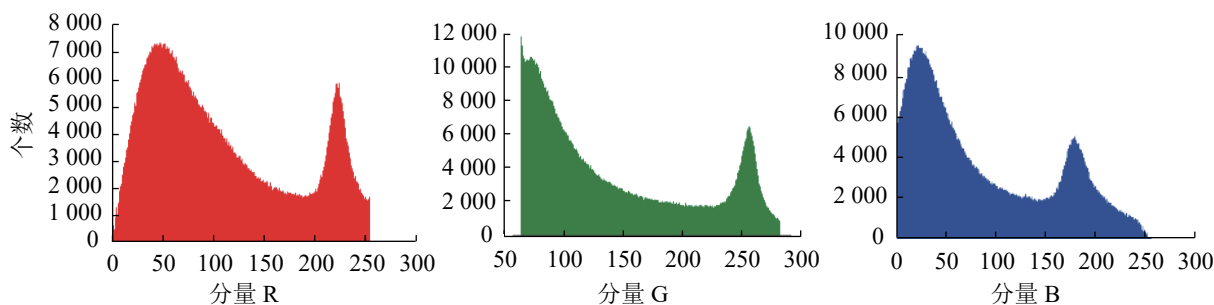


图 3 红脉槭颜色直方图

Figure 3 Color histogram of *A. rufinerve*

按照颜色直方图得到初始聚类中心点。尖齿槭颜色值为 (130.0, 144.0, 169.0), (20.0, 22.0, 49.0), (68.0, 78.4, 104.8); 红脉槭颜色值为 (84.6, 89.2, 110.4), (167.0, 184.0, 193.0), (32.0, 28.0, 56.0); 赛波德槭颜色值为 (181.0, 173.0, 207.0), (64.0, 40.0, 110.0), (113.0, 99.4, 154.9)。运用改进 K-means 聚类算法划分簇类 (图 4)。

为了增加管孔区域和背景区域两者的颜色梯度差异, 对聚类粗分割结果进行色彩编码 (图 5), 第 1 类对轴向薄壁组织细胞壁和木射线区域进行红色编码, 颜色值为 (202, 12, 22); 第 2 类对管孔区域进行绿色编码, 颜色值为 (53, 106, 195); 第 3 类对生长轮以及木射线区域归进行蓝色编码, 颜色值为 (29, 209, 107)。

经过灰度转换、开闭运算等图像形态学处理后, 消除了绝大部分的木射线、木纤维和轴向薄壁组织等面积较小的噪声。此外运用分水岭变换对已分类的木材微观图像进行目标分割, 即管孔分割 (图 6), 将管孔快速准确地分割。从最终的分割结果来看, 除了极个别的细小管孔未被检测到之外, 其余管孔均被完整地分割, 分割出的轮廓也与实际轮廓基本吻合。

2.2 验证结果

如图 7 所示: 阈值分割算法每张木材横截面微观图像平均有 88.1% 的管孔被准确有效地分割出来;

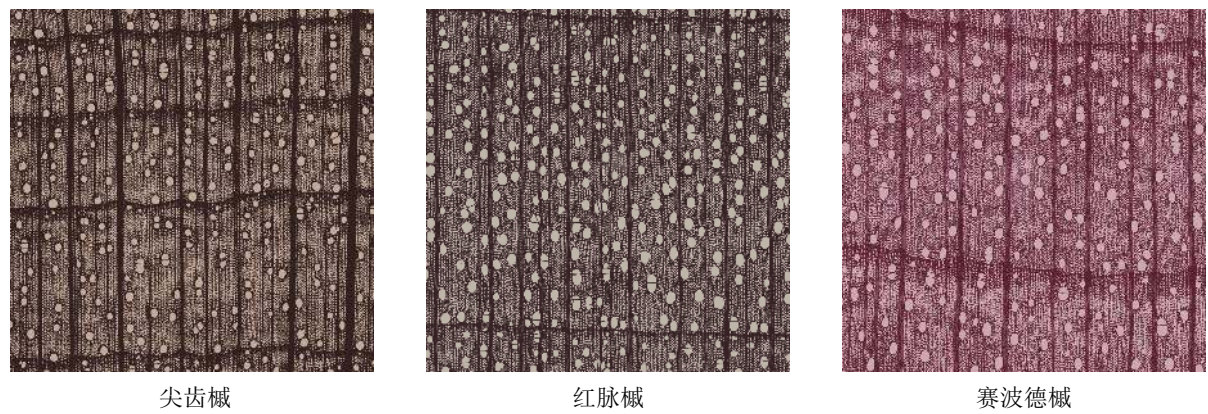


图 4 管孔粗分割结果

Figure 4 Rough segmentation results of pores

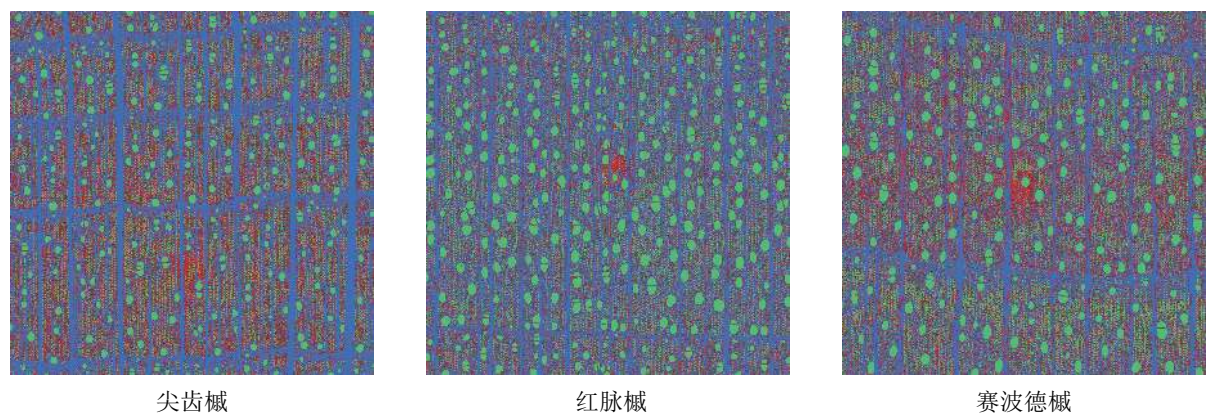


图 5 颜色编码结果

Figure 5 Color coding results

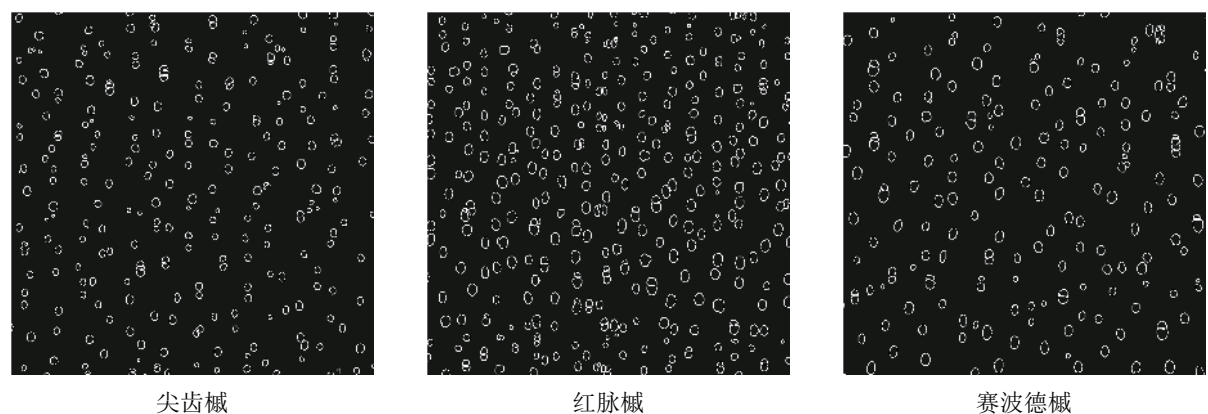
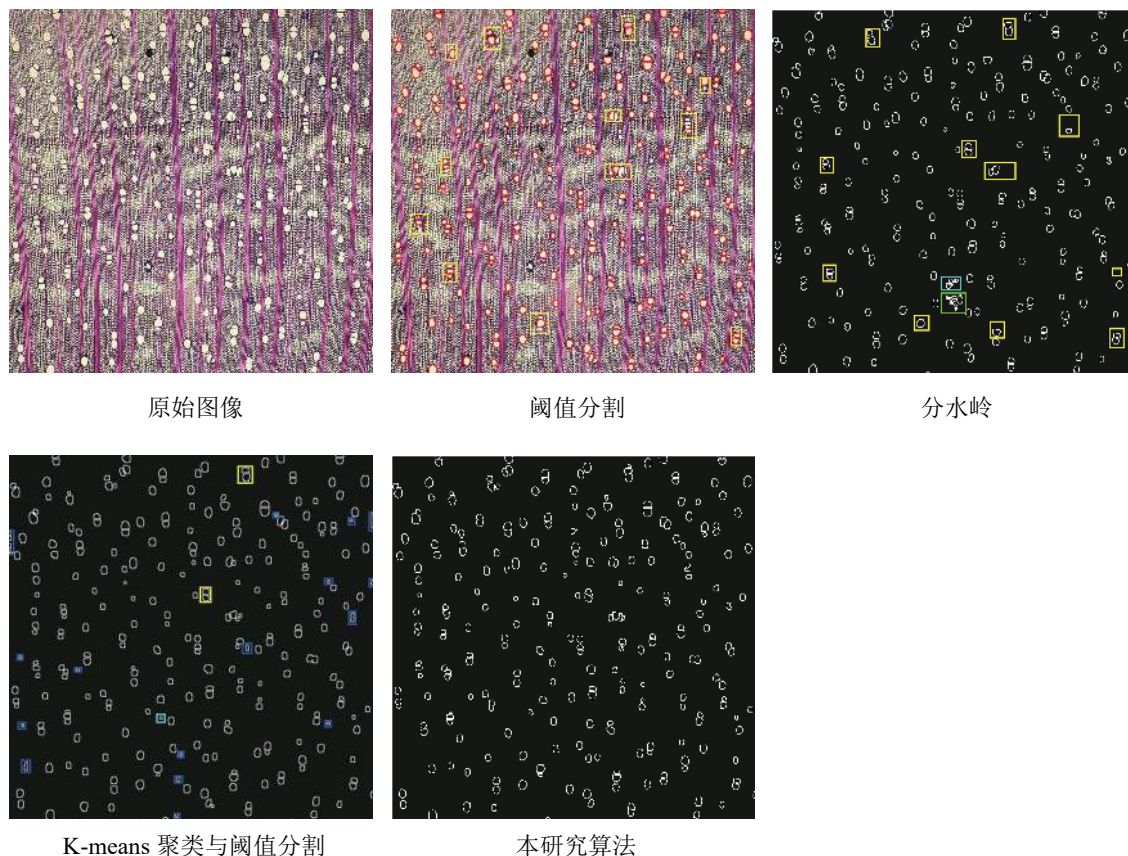


图 6 管孔精分割结果

Figure 6 Fine segmentation results of pores

分水岭算法每张木材横截面微观图像平均有 91.4% 的管孔被准确有效地分割出来，但存在噪声误分的问题；K-means 聚类与阈值分割算法先运用 K-means 聚类算法将木材图像分为 3 类，再采用阈值分割和 canny 算子做管孔精分割，结果每张木材横截面微观图像平均有 96.3% 的管孔被准确有效地分割出来，但部分木射线和轴向薄壁组织也被分割出来，且存在轮廓不饱满的问题；本研究算法为改进 K-means 聚类与分水岭算法，结果每张木材横截面微观图像平均有 97.1% 的管孔被准确有效地分割出来，分割出的管孔与实际管孔基本吻合。

为验证本研究算法分割管孔的准确性，对 12 类树种的横截面微观图像在各算法下的分割结果进行



管孔漏分标记为黄色，木纤维误分标记为绿色，轴向薄壁组织误分标记为浅蓝色，木射线误分标记为深蓝色

图7 各算法分割效果

Figure 7 Segmentation effect diagram of each algorithm

了比较(表1),能够看出阈值分割不存在管孔误分,但因面积较小而漏分的管孔较多;分水岭算法不仅存在管孔漏分且会误分割出木射线、木纤维以及轴向薄壁组织;K-means聚类与阈值分割存在木射线和轴向薄壁组织的误分割;本研究算法基本不存在管孔的漏分和误分问题。

3 结论

本研究改进的K-means聚类与分水岭的管孔分割算法,解决了初始中心点的随机性,并能有效地区分管孔区域与木纤维、木射线以及轴向薄壁组织等噪声区域,对管孔分割更为准确。本研究算法采用取图像颜色直方图对应的峰值确定初始聚类中心点和色彩编码的方式,充分将管孔与绝大部分噪声区域区分开,更适合于木材横截面微观图像分割任务。通过形态学处理和分水岭算法,对管孔进行精分割,将两者结合从而得到更为准确的分割结果。与其他分割算法相比,本研究算法分割效果显著提升,在大小不一且随机分布的管孔分割过程中鲁棒性高,有效避免了管孔漏分问题,具有良好的分割性能。

4 参考文献

- [1] 黄鹏桂,赵璠,李晓平,等.基于计算机的木材特征提取和分类识别技术研究综述[J].世界林业研究,2020,33(1):44-48.

HUANG Penggui, ZHAO Fan, LI Xiaoping, *et al.* Review of computer-based wood feature extraction and identification [J].

表1 各算法分割效果对比

Table 1 Comparison of segmentation effects of various algorithms				
算法	管孔漏分	木射线误分	轴向薄壁组织误分	木纤维误分
阈值分割	是	否	否	否
分水岭	是	是	是	是
K-means聚类与阈值分割	否	是	是	否
本研究算法	否	否	否	否

- World For Res*, 2020, **33**(1): 44 – 48.
- [2] ESPINOSA L F, HERRERA R J, POLANCOTAPIA C. Segmentation of anatomical elements in wood microscopic images using artificial vision techniques [J]. *Maderas Cienc Technol*, 2015, **17**: 106 – 108.
- [3] 黄鹏, 郑淇, 梁超. 图像分割方法综述[J]. 武汉大学学报 (理学版), 2020, **66**(6): 519 – 531.
HUANG Peng, ZHENG Qi, LIANG Chao. Overview of image segmentation methods [J]. *J Wuhan Univ Nat Sci Ed*, 2020, **66**(6): 519 – 531.
- [4] 韩刘杨, 田兴玲, 周海宾, 等. 饱水清代木材的构造腐朽程度及加固处理对其颜色的影响[J]. *西南林业大学学报 (自然科学)*, 2020, **40**(1): 132 – 138.
HAN Liuyang, TIAN Xingling, ZHOU Haibin, *et al.* The Influences of the anatomical structure and deterioration state of wood from a Qing Dynasty shipwreck on wood color after the consolidation treatment [J]. *J Southwest For Univ Nat Sci*, 2020, **40**(1): 132 – 138.
- [5] 杨建飞, 宁莉萍, 杨了, 等. 黑壳楠生长量及木材解剖特征的径向变异[J]. 南京林业大学学报 (自然科学版), 2018, **42**(1): 181 – 187.
YANG Jianfei, NING Liping, YANG Liao, *et al.* Variation in radial increment of *Lindera megaphylla* and its wood anatomical characteristics [J]. *J Nanjing For Univ Nat Sci Ed*, 2018, **42**(1): 181 – 187.
- [6] 中国林业科学研究院木材工业研究所. 一种木材识别方法及系统: CN201810830841.3[P]. 2019-01-03.
Research Institute of Wood Industry, Chinese Academy of Forestry. *A Wood Identification Method and System*: CN201810830841.3[P]. 2019-01-03.
- [7] 于海鹏, 刘一星, 崔永志, 等. 基于二值形态学的木材横切面显微构造特征测量[J]. *材料科学与工艺*, 2008, **16**(1): 107 – 111.
YU Haipeng, LIU Yixing, CUI Yongzhi, *et al.* Measurement of wood microstructural parameters on transverse section by binary morphology [J]. *Mater Sci Technol*, 2008, **16**(1): 107 – 111.
- [8] 计智伟, 汪杭军, 何涛, 等. 基于改进区域生长的木材导管形态特征提取方法[J]. 北京林业大学学报, 2011, **33**(3): 64 – 69.
JI Zhiwei, WANG Hangjun, HE Tao, *et al.* A morphological feature extraction method of wood pores based on an improved growing region algorithm [J]. *J Beijing For Univ*, 2011, **33**(3): 64 – 69.
- [9] 张广群, 汪杭军. 基于多目标遗传算法的管孔组合特征识别[J]. *林业科学*, 2012, **48**(4): 87 – 92.
ZHANG Guangqun, WANG Hangjun. Multi-objective genetic based pore combination recognition [J]. *Sci Silv Sin*, 2012, **48**(4): 87 – 92.
- [10] 汪杭军. 基于纹理的木材图像识别方法研究[D]. 合肥: 中国科学技术大学, 2013.
WANG Hangjun. *The Study of Recognition Methods of Texture-based Wood Images*[D]. Hefei: University of Science and Technology of China, 2013.
- [11] 祁亨年, 陈丰农, 方陆明, 等. 一种基于形态学的木材导管图像分割方法[J]. *北京林业大学学报*, 2008, **30**(4): 12 – 16.
QI Hengnian, CHEN Fengnong, FANG Luming, *et al.* A method for wood pore image segmentation based on mathematical morphology [J]. *J Beijing For Univ*, 2008, **30**(4): 12 – 16.
- [12] SHEN Pan, KUDO M. Segmentation of pores in wood microscopic images based on mathematical morphology with a variable structuring element [J]. *Comput Electron Agric*, 2010, **75**(2): 250 – 260.
- [13] 何拓, 高瑞清, 焦立超, 等. 世界木材标本馆现状与发展建议[J]. *木材工业*, 2020, **34**(3): 40 – 43, 55.
HE Tuo, GAO Ruiqing, JIAO Lichao, *et al.* Present status of global leading xylaria and suggestions for xylaria development in China [J]. *Chin J Wood Sci Technol*, 2020, **34**(3): 40 – 43, 55.
- [14] 李玉功. K-means 聚类算法的改进及其在彩色图像分割中的应用[D]. 合肥: 安徽大学, 2018.
LI Yugong. *Improvement of K-means Clustering Algorithm and Its Application in Color Image Segmentation*[D]. Hefei: Anhui University, 2018.
- [15] 卢印举, 段明义, 李祖照, 等. 一种改进的路面裂缝图像分割方法[J]. *科学技术与工程*, 2020, **20**(26): 10815 – 10821.
LU Yinju, DUAN Mingyi, LI Zuzhao, *et al.* An improved road crack image segmentation method [J]. *Sci Technol Eng*, 2020, **20**(26): 10815 – 10821.
- [16] 李平. 4 种槭属植物形态解剖学研究[D]. 长春: 吉林农业大学, 2011.
LI Ping. *Morphological and Anatomical Studies on Fours Species of Acer L.* [D]. Changchun: Jilin Agricultural University, 2011.