

基于支持向量机的显微图像聚焦区域选取算法^①

吕美妮^{②**} 黄玉健^{③*} 王奎奎^{*}

(* 梧州学院广西高校图像处理与智能信息系统重点实验室 梧州 543002)

(** 桂林电子科技大学信息与通信学院 桂林 541004)

摘 要 聚焦窗口的选取易受噪声、镜头杂质和玻片杂质的影响,本文针对玻片杂质引起聚焦曲线失去理想曲线特性的问题,提出一种基于支持向量机(SVM)的显微图像聚焦区域选取算法。首先,获取当前聚焦平面的图像,通过对比度和邻域相关性系数结合的方法检测是否存在模糊区域,以此作为是否存在玻片杂质的标准;其次,检测出模糊区域后,将当前图片进行色彩模式(CMYK)转化,提取图像的 K 分量图,对像素点进行处理输出分量图;再次,大致将分量图分为 2 类,计算子块的梯度和像素值总和进行坐标表示,通过支持向量机的训练,找出最佳分割线,识别并标定杂质区域;最后,剔除玻片杂质区域,重新选取聚焦窗口并进行二次聚焦。实验结果表明,该算法能有效剔除玻片杂质的影响,使得聚焦曲线保持较好的单峰性,且解决了玻片杂质引起焦平面误判的问题,确保选取到内容丰富的子块作为聚焦窗口,提高了聚焦的准确性。

关键词 图像处理;自动聚焦;聚焦窗口;支持向量机(SVM);玻片杂质

0 引 言

聚焦是成像系统中的重要环节,通过聚焦使光学系统对目标准确聚焦,从而获得目标清晰图像^[1,2]。自动聚焦技术主要分为主动式和被动式 2 种^[3,4],主动式自动聚焦通过传感器或测量工具完成自动聚焦;而被动式自动聚焦是一种基于数字图像处理的方法,不需要额外的传感器,近年来得到了广泛的应用。其中,聚焦区域的选择是被动式自动聚焦的核心问题之一,它直接影响聚焦算法的计算量、复杂度和精确度^[5,6]。聚焦区域的选择易受镜头杂质、玻片杂质的干扰而导致聚焦的失败。镜头杂质像素在聚焦的过程中灰度值基本不变,对聚焦区域的选择影响不大,而玻片杂质的存在会使得聚

焦评价函数曲线出现“双峰”或是“多峰”。玻片杂质主要分布在盖玻片表面,容易被误判为标本区域,导致聚焦提前结束,无法获取理想聚焦平面^[7]。

针对玻片杂质引起聚焦评价函数失去理想曲线特性的问题,文献[7]中使用局部标准差与锐利边缘像素数量联合检测焦平面图像的失焦模糊区域,排除玻片杂质造成的焦平面误判。该方法通过检测出焦平面图像是否存在模糊区域来判断是否进行二次聚焦,但并未注意到焦平面图像如果是玻片杂质清晰而目标模糊,该方法能检测出模糊区域并进行二次聚焦,但在二次聚焦结束后焦平面图像有可能是目标清晰而玻片杂质模糊,在进行检测的过程中,该方法同样也能检测出存在模糊区域,并进行再次聚焦,这将导致聚焦陷入死循环,无法获取理想的聚焦平面。针对上述方法的不足,本文提出一种基于

① 国家自然科学基金(61562074),梧州市公安视频图像大数据分析处理平台建设与关键技术研发(桂科 AB16380273),光学显微镜设备智能化关键技术研发及产业化(桂科 AA181180361)和广西高校中青年教師科研基础能力提升(2019KY0675)资助项目。

② 女,1991 年生,硕士;研究方向:数字图像处理;E-mail: miny_lvmeini@163.com

③ 通信作者,E-mail: jianer2001@163.com

(收稿日期:2019-01-15)

支持向量机(support vector machine,SVM)的显微图像聚焦区域选取算法,该算法能有效地解决玻片杂质引起焦平面误判的问题。

1 理论知识

1.1 聚焦失败的因素

在自动聚焦技术中,被动式自动聚焦主要分为2类:聚焦深度法(depth from focus,DFF)和离焦深度法(depth from defocus,DFD)^[8-11]。聚焦深度法主要包括3个核心问题:聚焦评价函数的选取、聚焦窗口的选取和搜索算法的设置。其中造成聚焦失败的因素除了各个核心算法本身的不足之外,还有外部的因素。而外部因素主要分为3大类:随机噪声、镜头杂质和玻片杂质。聚焦评价函数易受到随机噪声的影响而造成聚焦的不准确性,通过对原有算法的改进,可减弱随机噪声的影响。而聚焦区域的选取易受到镜头杂质和玻片杂质的影响,引起聚焦失败。在聚焦的过程中,镜头杂质的位置和内容变化不大,故其对聚焦的影响不大,可忽略不计。而玻片杂质分布在盖玻片表面,无法与目标同时达到最清晰的状态,导致聚焦提前结束而无法获取最佳聚焦平面。在自动聚焦的过程中,玻片杂质会造成聚焦函数曲线失去理想聚焦曲线特性,使得聚焦曲线出现“双峰”或“多峰”的现象,如图1所示。

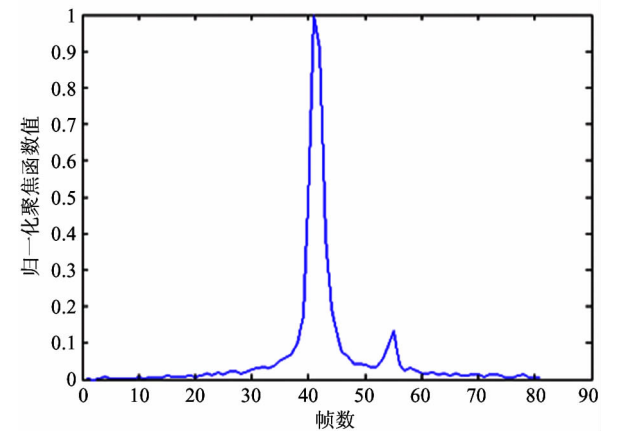


图1 玻片杂质序列的聚焦曲线

在搜索的过程中,每个峰值点都可以被判为最佳聚焦平面。如图2所示,玻片杂质与目标无法同时达到最清晰的状态,不论哪个聚焦平面都存在模

糊区域,故通过检测焦平面图像是否存在模糊区域来进行判断是否存在玻片杂质。

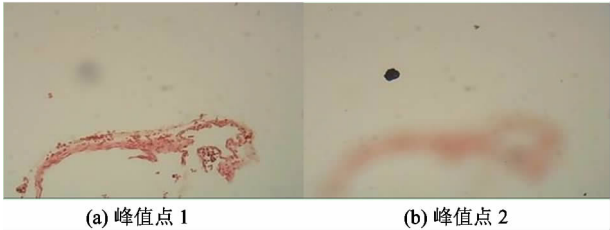


图2 焦平面图像

1.2 玻片杂质特征

玻片杂质主要分布在盖玻片表面,其聚焦平面与目标聚焦平面并不在同一位置上。如图3所示,玻片杂质焦平面图像的子块主要分为目标子块、背景子块和玻片杂质子块(其中因镜头杂质对聚焦的影响不大,镜头杂质则忽略不计)。从图3可以看出,标本经过染色,显微镜观察到的目标区域都是彩色的,而玻片杂质未经过染色,显微镜观察到的玻片杂质区域都是黑色。利用这个特性,从颜色特征上可以区分玻片杂质和目标。

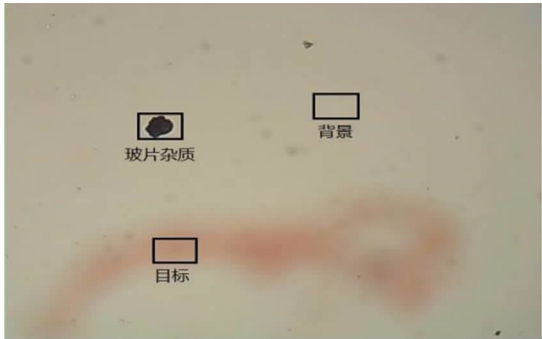


图3 玻片杂质聚焦平面

2 玻片杂质分析

2.1 颜色特征参数分割

对彩色图像进行分割,选取颜色特征 R 作为分割参数,通过比较非玻片杂质区域和玻片杂质区域的颜色特征参数,找到差别明显的参数做为阈值,以正确区分两者。对各焦平面的图像进行玻片杂质识别如图4、图5所示,其中边框中白色区域为识别出的玻片杂质。

从图4、图5可以看出,在自动聚焦中,如果当前的焦平面图像是玻片杂质清晰,通过颜色特征 R

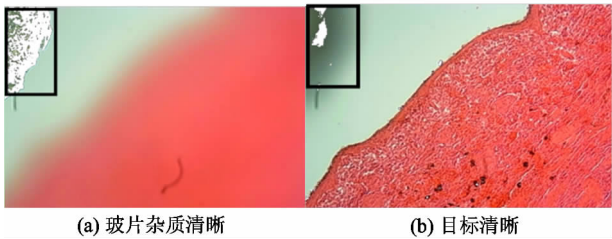


图 4 纸纤维

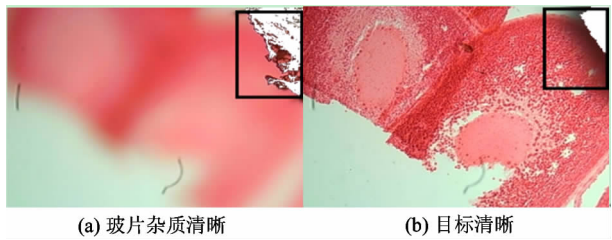


图 5 塑料纤维

做为分割参数是有一定的效果的,但如果当前焦平面图像是目标清晰而玻片杂质模糊,利用颜色特征 R 做为分割参数,不能很明显地区分玻片杂质区域和非玻片杂质区域。而在自动聚焦的过程中,搜索到的焦平面并不能保证是玻片杂质清晰的图像,故该方法不适用于玻片杂质的排除。

2.2 K 分量分割

玻片杂质在显微镜的观察下呈黑色,主要针对采集图片提取黑色分量来进行区分。将彩色图像转换为 CMYK 彩色模型^[12],提取 CMYK 模型的黑色 K 分量,如图 6 所示。

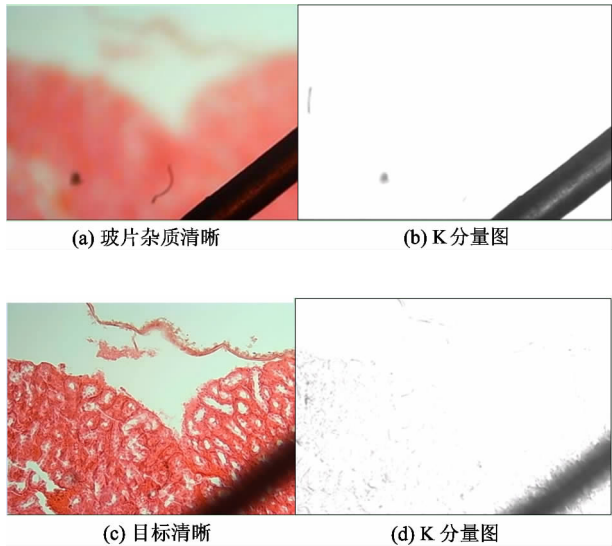


图 6 聚焦图像的 K 分量图

从图 6 中的分量图可以看出,聚焦图像是玻片杂质清晰而目标区域模糊的情况下,提取 K 分量能很好地把玻片杂质分割出来;但如果聚焦图像是目标清晰而玻片杂质模糊的情况下,提取 K 分量能把玻片杂质分割出来,同时也能把部分目标区域分割出来,这是因为当目标区域清晰的情况,颜色特征比较明显,其中包含的黑色分量也被分割出来。通过提取聚焦图像的 K 分量,虽然能将大部分玻片杂质分割出来,但同时也会把部分内容目标也分割出来,后续可利用 K 分量图像进行进一步分析。

2.3 特征值分析

利用聚焦图像(包含玻片杂质清晰和目标清晰 2 种)和 K 分量图像进行判定识别,将 K 分量图像分为杂质子块(后续的杂质子块都代表玻片杂质子块)和非杂质子块。如图 7 所示,为了区分出杂质

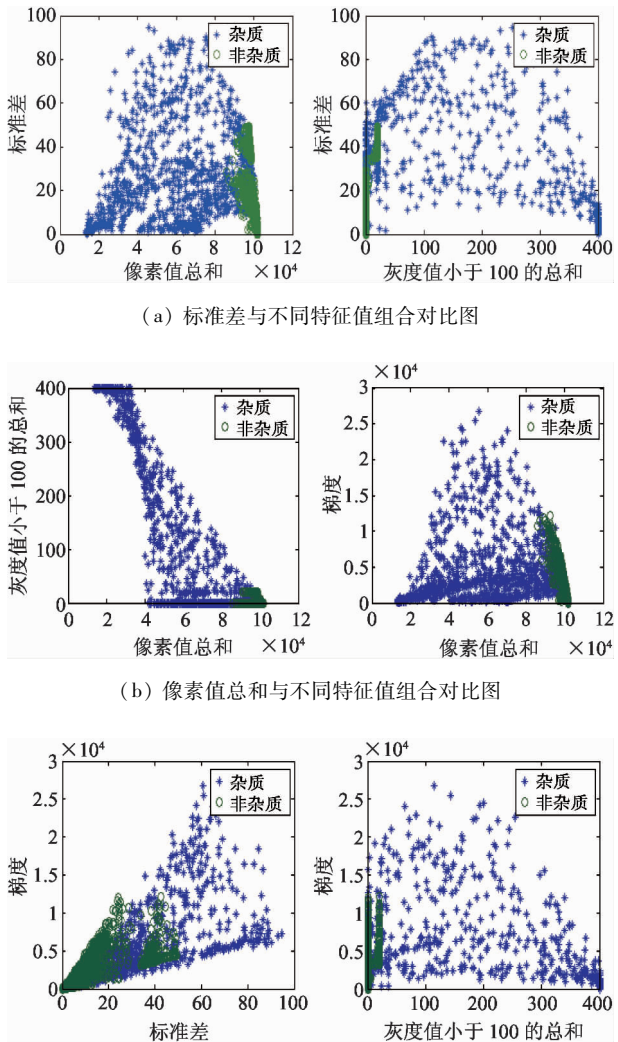
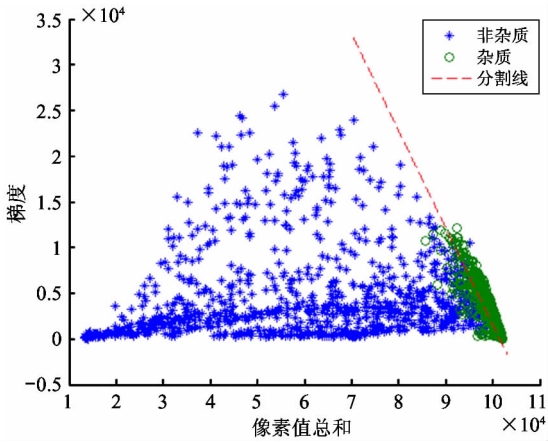


图 7 图像块各特征值对比图

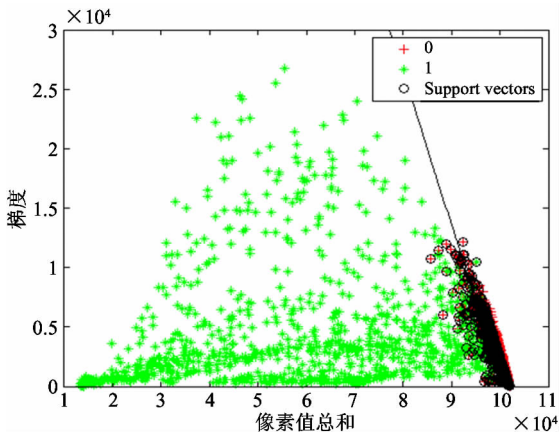
与非杂质,分别对含有杂质的图像块和不含杂质的图像块提取多种特征进行对比。

从图 7 可以看出,利用梯度值和像素值总和特征值进行子块区分,杂质子块和非杂质子块分布比较集中,且分界比较明显,而其他特征值的对比中,子块分布比较零散且没有明确的分界。

如图 8 所示,通过利用梯度值和像素值总和来进行最优平面线性划分,判断玻片杂质的位置,剔除玻片杂质的影响并重新选择二次聚焦的窗口。在分割线下方还有一些误判的子块,这些误判的子块一些包含杂质边缘,包含的杂质内容过少,另一部分则是提取 K 分量图的过程中,目标处于清晰的情况下,目标颜色分量明显,一些目标也会被提取出 K 分量,但这部分的 K 分量较分散,且不集中分布,故子块的特征值与玻片杂质的特征值并不是很集中,通过分割线能消除这部分的影响。



(a) 平面线性划分



(b) 最优平面线性划分

图 8 子块分割线

3 基于支持向量机的显微图像聚焦区域选取算法

搜索过程结束后,提取聚焦平面图像,判断图像是否存在模糊区域。将图像 f 分割为 $w \cdot w$ 大小的 $m \cdot n$ 个子块,求子块的对比度:

$$C(i, j) = \sum_{\delta} \delta(k, l)^2 p_{\delta}(k, l) \quad (1)$$

其中 (i, j) 表示图像子块坐标,相邻像素间灰度差 $\delta(k, l) = |k - l|$, $p_{\delta}(k, l)$ 为相邻像素间灰度差为 δ 的像素分布概率。

模糊子块和零内容子块对比度小,而清晰子块的对比度大,通过设定阈值,可区分出清晰子块。利用阈值函数 T_1 剔除清晰子块。

$$g(i, j) = \begin{cases} 0 & C(i, j) > T_1 \\ |f_{cor}(i, j) - f'_{cor}(i, j)| & C(i, j) < T_1 \end{cases} \quad (2)$$

其中, f 表示焦平面图像, f' 表示显微镜从焦平面位置沿 Z 轴移动一小段距离的图像, f_{cor} 和 f'_{cor} 表示子块的领域相关函数值,具体的公式如下所示:

$$\begin{cases} f_{cor}(i, j) = \frac{\sum_{x=1}^w \sum_{y=1}^w f(x, y) \times \bar{f}}{\sum_{x=1}^w \sum_{y=1}^w f(x, y)^2} \\ \bar{f} = \frac{1}{4} \times [f(x-1, y) + f(x+1, y) \\ + f(x, y-1) + f(x, y+1)] \end{cases} \quad (3)$$

由式(3)可知,在聚焦的过程中,目标子块和杂质子块的细节变化明显,背景子块基本不变或小幅变化^[13]。故在判定模糊子块的过程中,可利用邻域相关性的变化差来判别是否存在模糊子块。对邻域相关性系数变化差 $g(i, j)$ 进行判定是否存在模糊子块,先定义 $S = 0$,通过式(4)计算 $g(i, j)$ 中大于 0 的个数 S :

$$S = \begin{cases} S & g(i, j) = 0 \\ S + 1 & g(i, j) > 0 \end{cases} \quad (4)$$

判断 S 中个数总和是否大于 T_2 , $g(i, j)$ 中清晰子块和背景子块都为 0,而背景子块在聚焦的过程中虽然基本不变或在一定的范围波动,但式中利

用的是移动一小段距离,搜索步长较小,故背景子块变化可忽略不计,所以背景子块的 $g(i, j)$ 也为 0。但在具体的应用中,背景子块的变化也有部分不为 0,故需要阈值 T_2 来判断 S 中是否存在模糊子块的个数,进而排除背景子块的影响。如果 S 小于 T_2 ,则表示当前聚焦平面不存在模糊子块,且该焦平面为最佳聚焦平面;如果 S 大于 T_2 ,则表示当前聚焦平面存在模糊子块,说明该平面不是最佳聚焦平面,需要进一步处理。将焦平面图像 f 转化成 CMYK 彩色模型,提取图像的 K 分量图。转化 CMYK 模型方程^[14]如下:

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 255 \\ 255 \\ 255 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$
$$K = \min(\min(C, M), Y) \tag{5}$$

提取聚焦图像的 K 分量图进行判别:

$$k(i, j) = \begin{cases} K(i, j) \geq 255 & 255 \\ K(i, j) < 0 & 0 \end{cases} \tag{6}$$

其中, $K(i, j)$ 表示 K 分量图上点 (i, j) 的像素值, k 表示判别后的输出图像。

将判别后的 K 分量图像(图像大小为 $M \cdot N$)分为 $w \cdot w$ 大小的 $m \cdot n$ 个子块,其中 $m = M/w, n = N/w, i = m \cdot n$ 利用 K 分量图像进行判定识别,将子块大致分为杂质子块和非杂质子块,计算子块的像素值总和 z_i :

$$z_i = \sum_r^m \sum_s^n k(r, s) \tag{7}$$

计算子块的像素值总和后,再计算子块的梯度 g_i :

$$g_i = \sum_r^m \sum_s^n |k(r, s) - k(r - 1, s)| + |k(r, s) - k(r, s - 1)| \tag{8}$$

则子块可以用坐标表示,坐标的表达式为 $x_i = (z_i, g_i)$,在标定的过程中,令杂质子块为正例,标记符为 +1,,而非杂质子块为负例,标记符为 -1。训练集样本表示为 $T = \{(x_1, y_1), (x_2, y_2), \cdots, (x_i, y_i)\}$,其中, y_i 为 x_i 的标记类, $y_i \in \{+1, -1\}$ 。对训练集 T 通过线性可分支持向量机找到最优分割平面,具体如下^[15]。

假设直线 $w \cdot x + b = 0$ 为最优分割线,样本点到分割线的几何距离为

$$\frac{y_i \cdot y(x_i)}{\|w\|} = \frac{y_i \cdot (w^T \cdot \Phi(x_i) + b)}{\|w\|} \tag{9}$$

分割线满足支持向量到其最小距离最大,即 $\max[\text{支持向量机到分割线的最小距离}]$,只要算出支持向量机到分割线的距离就能满足条件,则支持向量机到分割线的最小距离为

$$\min_{x_i} \left[\frac{w \cdot x_i + b}{\|w\|} \right] = \min_{x_i} \left[\frac{y_i(w \cdot x_i + b)}{\|w\|} \right] \tag{10}$$

最终优化的公式为

$$\max_{w, b} \left[\min_{x_i} \frac{y_i(w \cdot x_i + b)}{\|w\|} \right] \tag{11}$$

为了区分杂质子块与非杂质子块,将 $y(w \cdot x + b) \geq 1$ 带入式(11)得出新的优化公式后,再进一步将公式优化,优化最大化分数,转化为优化最小化分母,优化后的公式为

$$\begin{cases} \min_{w, b} \frac{1}{2} \|w\|^2 \\ \text{s. t. } y_i(w^T \cdot \Phi(x_i) + b) \geq 1, i = 1, 2, 3, \cdots, m \times n \end{cases} \tag{12}$$

利用式(12)求出最优解 w 和 b ,得出最优分割线 $w \cdot x + b = 0$,将分割线上的判定为杂质并标记,将分割线下面的判定为非杂质子块。剔除杂质子块后需重新确定聚焦窗口的位置并进行二次聚焦。而聚焦区域的选取则通过将当前非杂质子块根据图像子块的梯度幅值总和来确定聚焦窗口的位置,进而通过文献[16]提出的改进算法来计算图像序列的聚焦评价函数值,并进行二次聚焦。

4 玻片杂质处理实验与结果分析

本文实验环境为 CPU, Intel (R) Core (TM) i7-6700 3.4GHz; RAM, 8.00GB; MATLAB (R2014a)。

本文实验图片(480 × 640 像素)均通过梧州奥卡光学仪器公司研发的光学显微镜(XSZ-700)进行采集。

实验中选用 3 组图像大小为 480 × 640 的显微图像序列(含有玻片杂质),聚焦图像如图 9(a) ~ 图 9(c)所示。图 9(a)、(c)所示的图像序列,最佳聚焦位置在第 41 帧,图 9(b)所示的图像序列,最佳聚焦位置在第 56 帧。

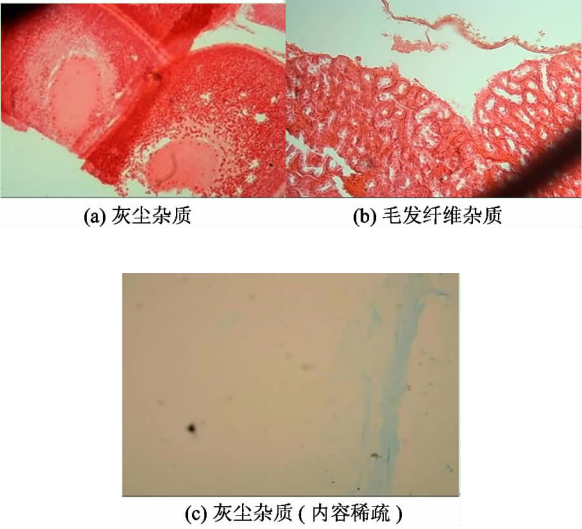
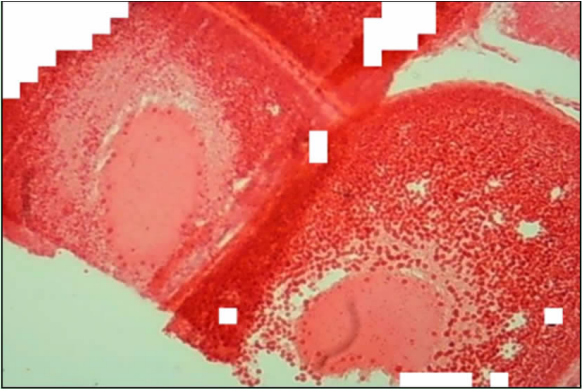
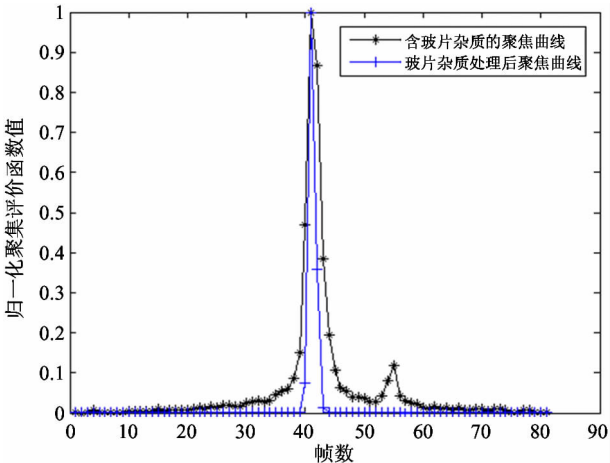


图9 含玻片杂质的显微图像序列



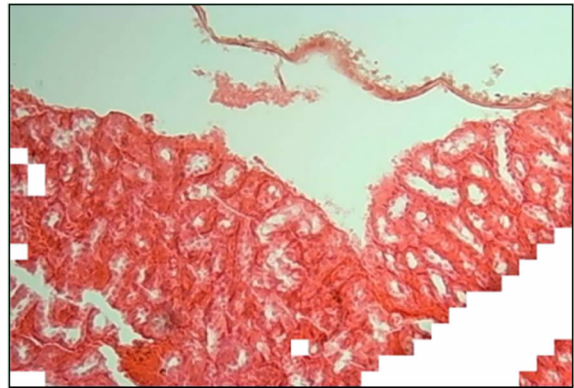
(a) 焦平面图像的杂质标定

此处需要说明的是,文献[7]方法有可能导致聚焦陷入死循环,无法获取理想的聚焦平面。故在实验中并未将此算法进行对比,且在现阶段的研究中,并未有新的算法来处理玻片杂质,所以在实验中,主要通过和原本含有玻片杂质的聚焦曲线进行对比。而在玻片杂质的识别和标定中,主要利用子块,而非单个像素,且在聚焦窗口的选取中,目前主要也是利用目标子块,而利用子块可减少计算量。仿真实验结果如图 10 ~ 图 12 所示。

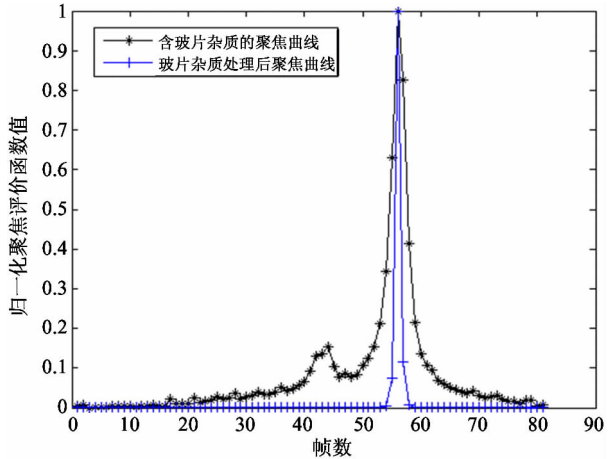


(b) 杂质处理前后的聚焦曲线对比图

图 10 纸纤维杂质处理的实验结果

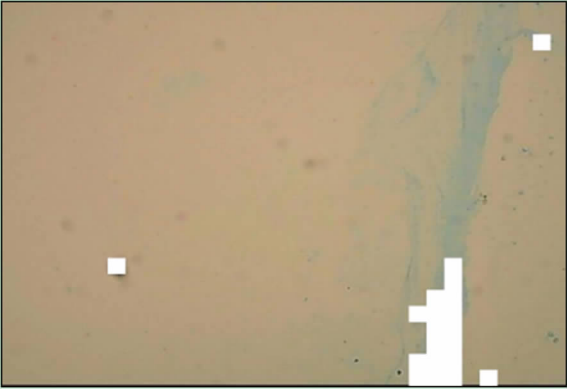


(a) 焦平面图像的杂质标定

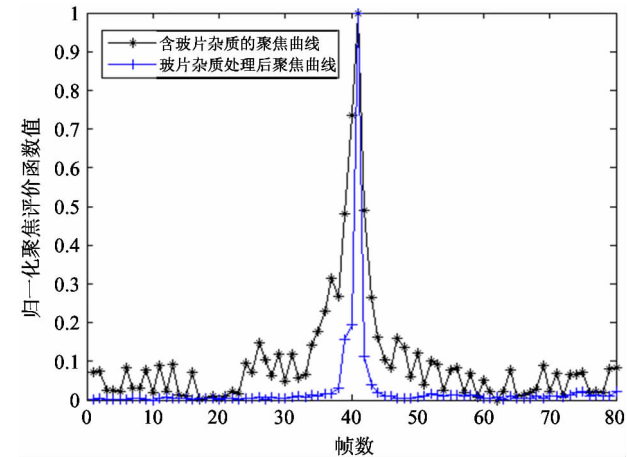


(b) 杂质处理前后的聚焦曲线对比图

图 11 毛发杂质处理的实验结果



(a) 焦平面图像的杂质标定



(b) 杂质处理前后的聚焦曲线对比图

图 12 灰尘杂质(内容稀疏)处理的实验结果

5 结 论

聚焦窗口是显微镜自动聚焦的核心问题之一，聚焦窗口决定后续算法的计算量、精度和准确性。在自动聚焦的选取中，聚焦窗口易受到玻片杂质的影响，而玻片杂质一般分布在盖玻片表面，玻片在载物台观察的过程中，由于风尘或擦拭布不干净等因素，使得玻片上遗留有杂质。而在自动聚焦的过程中，玻片杂质的存在，导致聚焦曲线存在“双峰”和“多峰”的情况，造成在搜索的过程中，使得搜索提前结束而无法获取最佳聚焦平面。而本文提出的算法，能有效剔除玻片杂质的影响，解决玻片杂质引起聚焦平面误判的问题，且能确保选取到内容丰富的子块作为聚焦窗口，提高聚焦的准确性。故该算法适用于显微成像系统。

参考文献

[1] 卢洋. 光学显微自动聚焦图像处理算法设计与系统研制[D]. 青岛:中国石油大学(华东)信息与控制工程学院, 2014:-6

[2] 张来线, 孙华燕, 郭惠超, 等. 基于图像灰度梯度最大值累加的自动调焦算法[J]. 光子学报, 2013, 42(5):605-610

[3] 蒋涛, 左昉, 王灵国. 严重离焦状态下的自动聚焦实现[J]. 激光技术, 2015, 39(6):794-797

[4] 林忠, 黄陈蓉, 卢阿丽. 基于离焦量差异定性分析的

自动对焦方法[J]. 计算机应用, 2015, 35(10):2969-2973

[5] 李惠光, 王帅, 沙晓鹏, 等. 显微视觉系统中自动聚焦技术的研究[J]. 光电工程, 2014, 41(8):1-9

[6] 尤玉虎, 刘通, 刘佳文. 基于图像处理的自动对焦技术综述[J]. 激光与红外, 2013, 43(2):132-136

[7] 彭国晋, 玉振明, 于健海. 光学显微镜自动聚焦取窗方法研究[J]. 应用光学, 2015, 36(4):550-558

[8] 商艳芝, 江旻珊. 光显微成像系统自动对焦技术的研究[J]. 光学仪器, 2016, 38(2):145-148

[9] 王家川. 基于图像处理的自动聚焦理论研究[D]. 重庆:重庆大学光电工程学院, 2012: 9-12

[10] 裴锡宇, 冯华君, 李奇, 等. 一种基于频谱分析的离焦深度自动对焦法[J]. 光电工程, 2003, 30(5):62-65

[11] 王彦芳. 自动聚焦系统中评价函数性能与动态区域选取的研究[D]. 济南:山东大学信息科学与工程学院, 2011: 11-12

[12] 于福洋, 张伟东, 刘佳. CMYK 色彩模式的数字化定义[J]. 消费导刊, 2008(17):182-182

[13] 翟永平, 周东翔, 刘云辉, 等. 聚焦函数性能评价指标设计及最优函数选取[J]. 光学学报, 2011, 31(4):234-244

[14] 胡维友. 颜色空间 RGB 与 CMYK 之间的变换[J]. 印刷世界, 2007(3):9-12

[15] Lisi1129. SVM 算法最通俗易懂讲解[EB/OL]. <http://blog.csdn.net>: CSDN, 2017

[16] 吕美妮, 玉振明. 基于 DCT 零系数和局部标准差的自动聚焦算法[J]. 激光技术, 2018, 42(1):66-71

An algorithm for selecting focus area of micro image based on support vector

Lv Meini^{* **}, Huang Yujian^{*}, Wang Kuikui^{*}

(^{*} Guangxi Colleges and Universities Key Laboratory of Image Processing and Intelligent
Information System, College of Wuzhou, Wuzhou 543002)

(^{**} College of Information and Communication, Guilin University of Electronic Technology, Guilin 541004)

Abstract

The selection of focusing window is easy to be affected by noise, lens impurity and glass impurity. A support vector machine (SVM) based micro-image focus area selection algorithm is proposed to solve the problem of the loss of ideal curve characteristics caused by glass impurity. Firstly, an image of a current focusing plane is acquired to detect whether a blurred region exists by a method of combining contrast and neighborhood correlation coefficient as a standard for judging whether a glass impurity is present. Secondly, after the blurred region is detected, the current picture is transformed by color mode (CMYK), the K component of the image is extracted and the pixel is processed and the component map is output. Thirdly, the component map is roughly divided into 2 categories. The gradient of the sub-block and the sum of the pixel values are calculated to represent the coordinates. The best segmentation line is found through training of support vector machines, and the impurity regions are identified and calibrated. Finally, the glass impurity region is removed, and the focus window is selected again for secondary focusing. The experimental results show that the algorithm can effectively eliminate the influence of glass impurities, keep the focusing curve unimodal, solve the problem of focal plane misjudgment caused by glass impurities, and ensure that the rich content sub-blocks are selected as the focus window, improving the accuracy of focusing.

Key words: image processing, automatic focusing, focusing window, support vector machine (SVM), glass impurities