

推扫式高光谱显微成像系统设计与实验

齐美捷¹, 刘立新^{1,2*}, 李艳茹¹, 刘玉杰¹, 张周锋², 屈军乐³¹西安电子科技大学光电工程学院, 陕西 西安 710071;²中国科学院西安光学精密机械研究所光谱成像技术重点实验室, 陕西 西安 710119;³深圳大学物理与光电工程学院光电子器件与系统教育部/广东省重点实验室, 广东 深圳 518060

摘要 高光谱显微成像(HMI)是一种新型无损光学诊断技术,其光谱数据能够反映样本的内部微环境变化,图像数据可以反映样本空间结构信息,因此可以作为癌症诊断工具,在未来具有广阔的应用前景。但 HMI 数据量大且数据结构复杂,将其应用于癌症诊断领域需要进行系统详细的数据解译。设计并搭建了一套推扫式 HMI 系统,并编写了系统控制、数据采集和数据分析软件,可提供多种基于机器学习的数据处理方法。基于 MATLAB 编制了具有图形化用户界面的 HMI 数据采集和数据分析软件,该软件可给出分析结果,为医生病理诊断提供了便利。利用该系统和软件进行皮肤癌的分类与分期研究,验证了系统的性能。HMI 系统的光谱范围为 465.5~905.1 nm,光谱分辨率约为 3 nm,视场尺寸为 400.18 μm ×192.47 μm ,放大倍率为 28.15,实际分辨率范围为 1.10~1.38 μm 。分别采集基底细胞癌、鳞状细胞癌和恶性黑色素瘤组织的 HMI 数据,利用图像数据实现了三种皮肤癌的分类,准确率为 85%;利用光谱数据实现了鳞状细胞癌的分期鉴别,准确率达到 96.4%。

关键词 医用光学; 高光谱显微成像; 皮肤癌; 图形用户界面; 癌症诊断

中图分类号 TH742

文献标志码 A

DOI: 10.3788/CJL202249.2007105

1 引言

高光谱成像技术起源于 20 世纪 80 年代^[1],是在多光谱遥感技术的基础上发展而成的。高光谱成像是将传统的成像技术与光谱技术相结合,既可以获取样本的图像信息,又可以获取光谱信息,具有“图谱合一”的特点,目前已在食品检测^[2-3]、遥感探测^[4-6]和文物鉴定^[7-8]等领域中获得广泛应用。高光谱显微成像(HMI)技术是将高光谱成像与显微成像相结合,可以实现对组织切片、细胞或微生物的检测^[9-11],因此被广泛应用于医学诊断领域。HMI 数据提供的丰富图像信息和光谱信息有助于提高疾病诊断的准确率、灵敏度和特异性,为一些疾病的早期诊断提供了更为有效的手段。

近年来国内外学者在 HMI 系统方面开展了大量研究,并取得了一定成果。2020 年,Zhang 等^[12]基于声光可调谐滤波器构建了 HMI 双滤波系统,波长范围为 450~650 nm;从光谱和成像的角度对优化后的系统进行了分析,在可见光范围内,单晶双滤波的光谱分辨率相比于单滤波提高了 37.08%~59.95%。2020 年,Xu 等^[13]开发了一个多模态 HMI 系统,集成了透射成像、反射成像以及荧光成像三种模式;高光谱成像仪

采用推扫式扫描,可以获得空间分辨率为 5 μm ,波长范围为 400~700 nm,光谱分辨率为 3 nm 的 HMI 数据立方;系统的照明和探测采用同轴结构,确保了视场中的均匀照明,消除了焦面上可能出现的阴影。2021 年,Yu 等^[14]提出了一个微透镜阵列式快照 HMI 系统,该系统分为两条光路,一条用于执行具有高空间分辨率的显微成像,另一条以探测器的帧速率执行 HMI 数据立方的高速采集;系统的空间分辨率为 2.5 μm ,波长范围为 400~800 nm,光谱分辨率约为 0.56 nm。

此外,HMI 技术能够在组织或细胞发生异常改变的早期获取其光谱信息的差异,可用于鉴别正常组织与病变组织,通过与人工智能技术相结合,有望实现疾病的早期诊断,例如,宫颈癌、乳腺癌、胃癌和皮肤癌等^[15-20]。2018 年,杜剑等^[17]探究了 HMI 技术在胃癌组织病理诊断中的应用,通过建立卷积神经网络,对正常/癌变组织进行了分类,准确率为 96.53%,灵敏度为 94.29%,特异性为 97.14%。2020 年,Chen 等^[18]利用高光谱显微镜采集了 H&E 染色的肝癌组织细胞核的透射光谱并进行了肝癌鉴别,基于支持向量机模型,对癌细胞鉴定的敏感性和特异性分别为 99%和 98%。2020 年,王健生等^[19]利用 HMI 技术结合基于

收稿日期: 2022-04-13; 修回日期: 2022-06-06; 录用日期: 2022-06-22

基金项目: 国家自然科学基金(62075177)、高等学校学科创新引智计划(B17035)、中国科学院光谱成像重点实验室开放基金(LSIT202005W)

通信作者: *lxliu@xidian.edu.cn

特征谱监督的最小二乘支持向量机算法,研究了皮肤黑色素瘤浅表扩散深度,该方法对恶性黑色素细胞的分割准确率大于 85%。2021 年, Deng 等^[20]利用 HMI 系统获取了胆管癌图谱数据,并基于 ResNet-50 深度学习网络对胆管癌进行了识别,准确率为 82.4%。

当前 HMI 系统的开发及其在癌症诊断中的应用均取得了较大进展,但仍存在一些问题。HMI 系统的成像光谱范围大多限制于可见光波段,空间分辨率较低,成像视场较小,无法观察癌变组织的整体形貌。此外, HMI 技术在癌症诊断中的应用仍处于初步研究阶段,样本的采集、数据预处理和建模分析等缺乏系统规范化的操作流程,这对癌症诊断结果有一定的影响。针对这些问题,本文设计并搭建了一套推扫式 HMI 系统,编制了基于图形化界面的 HMI 系统控制、数据采集与分析软件,可以对 HMI 数据采集过程中的参数进行调整并对采集到的 HMI 数据进行快速解析,操作简单直观。为验证所搭建的系统性能,分别对基底细胞癌、鳞状细

胞癌和恶性黑色素瘤组织进行成像,并利用所编制的数据分析软件对皮肤癌进行了分类与分期鉴别。

2 系统硬件实现

2.1 系统原理

本文所搭建的推扫式 HMI 系统原理图如图 1 所示。该系统主要元器件包括:卤钨灯光源、一维高精度步进电机型自动位移平台、二维手动位移平台、样品台、物镜、透镜组、反射镜和高光谱相机。卤钨灯光源照明位于样品台上的样本,携带了样本信息的透射光被物镜收集,依次经过反射镜、透镜组后到达高光谱相机。利用高精度步进电机型自动位移平台控制样品台,实现样品在 x 方向的移动,步长为 $1\ \mu\text{m}$,最终推扫出样本的 HMI 数据立方。通过调整二维高精度手动位移平台,可以实现系统的视场选择与聚焦状态调节,由电脑控制实现高精度步进电机型自动位移平台参数的调整以及三维 HMI 数据立方的采集。

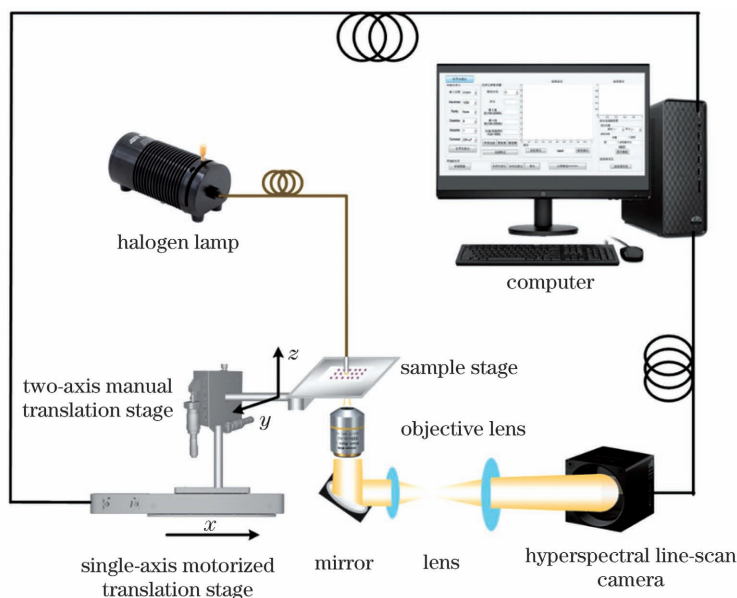


图 1 推扫式 HMI 系统的原理图

Fig. 1 Schematic of push-broom HMI system

2.2 系统器件

实验采用的高光谱相机的扫描方式为线扫描,能以每秒 170 个立方体的速度捕捉到高光谱图像。将有滤波片覆盖的传感器称为有源区域,则该传感器包含两个有源区域:有源区域 1 的尺寸为 $2048\ \text{pixel} \times 320\ \text{pixel}$,对应 64 个在可见光范围内的波段;有源区域 2 的尺寸为 $2048\ \text{pixel} \times 640\ \text{pixel}$,对应 128 个在近红外范围内的波段。该高光谱相机的工作波段共有 192 个,其中至少有 150 个波段用于数据采集,其余波段用于质量检查。两个有源区域是非连续的,相距 $120\ \text{pixel}$ 。经过校准后有源区域 1 中共对应 48 个波段,有源区域 2 中共对应 103 个波段。光源采用卤钨灯光源,波长范围为 $360\sim 2400\ \text{nm}$,覆盖可见光到近红外波段。高精度步进电机型自动平台的行程为

$85\ \text{mm}$,载重为 $78.4\ \text{N}$ 。物镜数值孔径为 0.45 ,工作距离为 $4\ \text{mm}$ 。

基于高光谱相机的传感器结构和线扫描的成像方式对样本进行观察时,为了获取完整的 HMI 数据,需要使样本相对高光谱相机在一个方向上平移,确保样本上每个点的数据均被采集。本文利用高精度步进电机型自动平台驱动样品台,实现相对位移,从而获得 HMI 数据立方。获取该 HMI 数据的过程简述如下:假设现有一虚拟线扫描传感器,该传感器含有两个滤波器,其中第一个滤波器(尺寸为 $26\ \text{pixel} \times 10\ \text{pixel}$)对应两个波段,第二个滤波器(尺寸为 $26\ \text{pixel} \times 5\ \text{pixel}$)对应一个波段,两个滤波器相隔 $5\ \text{pixel}$ 。为了获得样本完整的 HMI 数据立方,需要使样本的每个点依次通过上述三个波段,这一步骤通过控制高精度

步进电机型自动平台沿 x 方向移动样品台来实现。数据采集步骤如图 2 所示,其中五角星表示待观测样本。整个高光谱数据立方的采集过程可以分为三个部

分:启动阶段[图 2(a)~(d)]、稳态阶段[图 2(e)]和关闭阶段[图 2(f)~(h)]。样本扫描结束后形成的 HMI 数据立方如图 2(i)所示。

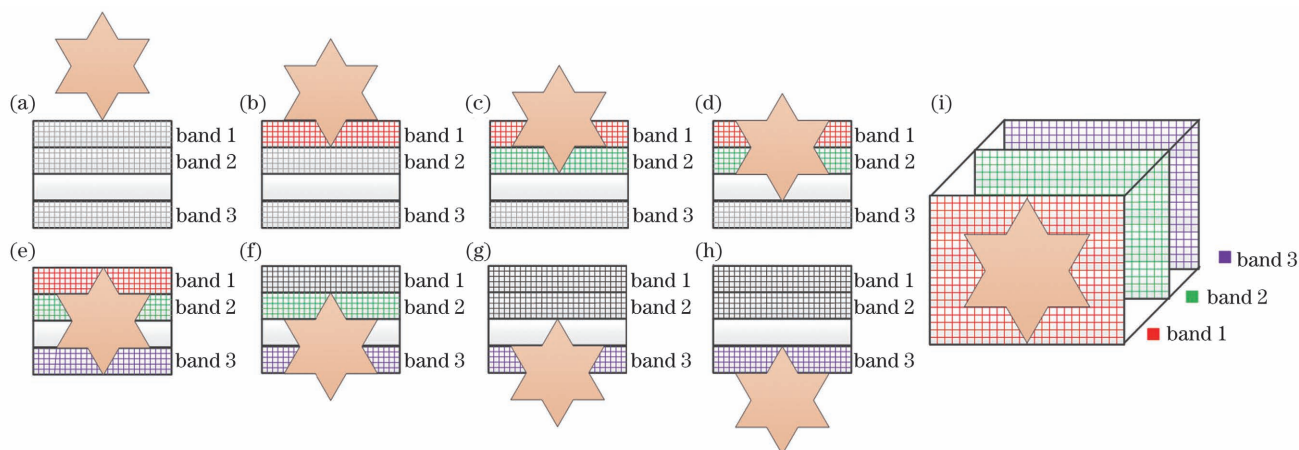


图 2 HMI 系统中的样本扫描示意图。(a)~(d)启动阶段;(e)稳态阶段;(f)~(h)关闭阶段;(i)扫描结果

Fig. 2 Schematics of sample scanning in HMI system. (a)–(d) Start-up phase; (e) steady-state phase; (f)–(h) shutdown phase; (i) scanning result

2.3 系统技术指标

2.3.1 波段数及光谱分辨率

实验采集到的 HMI 数据波段数取决于高光谱相机的传感器,所使用的传感器的工作波段共有 192 个,根据厂家提供的校准方法对高光谱相机传感器进行校准后,确定共有 151 个波段,覆盖波长范围为 465.5~905.1 nm。在相同光谱范围内,系统采集的波段数越多,光谱分辨率越高,可以识别的样本细节越丰富。光谱分辨率(σ)的计算公式为

$$\sigma = \frac{\lambda_{\max} - \lambda_{\min}}{n_u}, \quad (1)$$

式中: λ_{\max} 为最大波长; λ_{\min} 为最小波长; n_u 为波段数,通过计算可得本文搭建的系统的光谱分辨率约为 3 nm。

2.3.2 空间分辨率

在传统宽场显微成像中,光学系统的衍射极限分辨率(δ)为

$$\delta = \frac{0.61\lambda}{NA}, \quad (2)$$

式中: λ 为波长; NA 为物镜数值孔径。光学系统分辨率与波长成正比,与数值孔径成反比。系统中所用物镜的 $NA=0.45$,分别代入最大波长 $\lambda_{\max}=905.1$ nm,最小波长 $\lambda_{\min}=465.5$ nm,可得横向光学分辨率范围为 0.63~1.23 μm 。

相机的图像分辨率是由相机的像素大小决定的,当芯片大小相同时,像素尺寸越小,相机分辨率越高。根据奈奎斯特采样定律可知,当相机像素尺寸小于等于所需分辨样品尺寸的 1/2 时,才能准确分辨。因此相机的图像分辨率(δ')的计算公式为

$$\delta' = \frac{2 \times P}{M}, \quad (3)$$

式中: M 为系统放大率; P 为像素大小。经计算,相机的图像分辨率为 0.39 μm 。比较图像分辨率和光学分辨率,可知系统的分辨率受限于光学分辨率,理论值为 0.63~1.23 μm 。

为验证系统的实际分辨率,本文对分辨率板进行推扫成像。图 3 所示为分辨率板的单波段推扫图样,

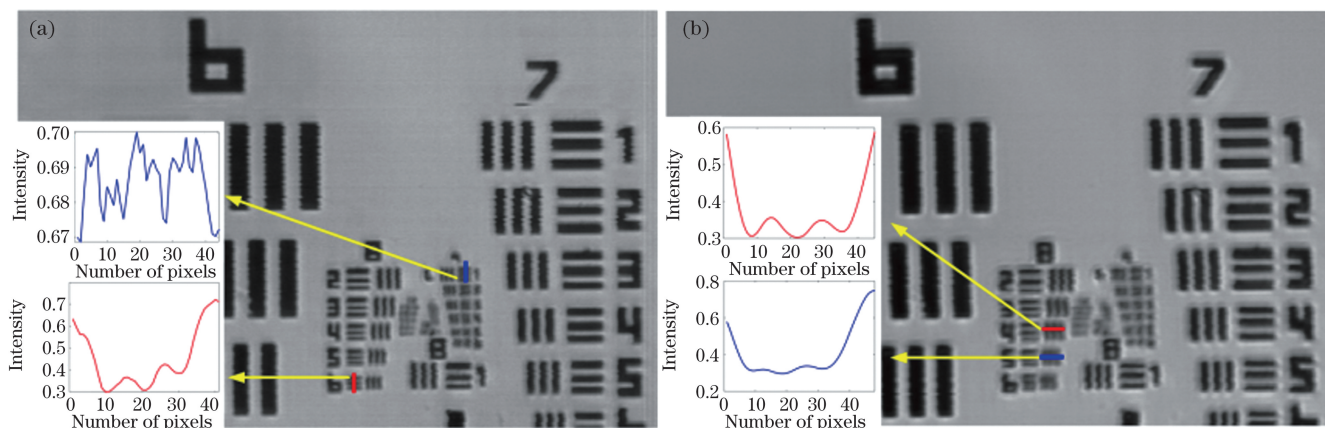


图 3 分辨率板的单波段扫描图像。(a)波长为 465.5 nm 时;(b)波长为 905.1 nm 时

Fig. 3 Single band scanning images of resolution test panel. (a) When wavelength is 465.5 nm; (b) when wavelength is 905.1 nm

其中:图 3(a)对应波段 1,波长为 465.5 nm;图 3(b)对应波段 151,波长为 905.1 nm。由图 3(a)可以看出,在波长为 465.5 nm 的波段下,该系统可以清晰分辨分辨率板第 8 组第 6 号线对,但无法清晰分析第 9 组第 1 号线对,对应的系统最高分辨率为 1.10 μm 。由图 3(b)可以看出,在 905.1 nm 波段下,该系统可以清晰分辨分辨率板中的第 8 组第 4 号线对,但无法清晰分辨第 8 组第 5 号线对,对应的系统最低分辨率为 1.38 μm 。因此该 HMI 系统的实际分辨率范围为 1.10~1.38 μm 。

2.3.3 视场及放大率

本文采集到的单波段图像大小为 2048 pixel \times 985 pixel,根据分辨率板对应的线对线宽参数,计算各个线宽在像素中所占的比例,最终确定推扫式 HMI 系统的视场为 400.18 μm \times 192.47 μm 。系统放大率计算公式为

$$M = \frac{a \times b}{l}, \quad (4)$$

式中: a 为横向像素个数; b 为像素大小; l 为横向视场长度。经过计算,放大率 M 为 28.15。

3 系统软件实现

3.1 软件环境

HMI 系统的数据采集与数据分析软件的编写均基于 MATLAB 语言,其中软件平台为 MATLAB R2019a (9.6.0.1072779) 64-bit(win64)。操作系统为 Windows 10,运行内存为 16.0 GB,计算机处理器为 Intel(R)Core(TM)i5-10400CPU@2.90 GHz。

3.2 HMI 数据采集软件

HMI 数据采集软件的图形化界面如图 4 所示,主要包括三个模块。HMI 系统控制及数据采集模块[图 4(a)]可以实现高光谱相机和高精度步进电机型自动平台的控制、HMI 数据采集、光源背景校正和频域滤波处理。HMI 数据显示与处理模块[图 4(b)]可对 HMI 数据立方和单波段图像进行展示或裁剪,以及计算各个波段之间的相关性。数据保存与软件退出模块[图 4(c)]可实现高光谱数据处理结果的保存和采集软件的退出,也可以通过点击“分类算法”按钮实现分析软件的跳转。

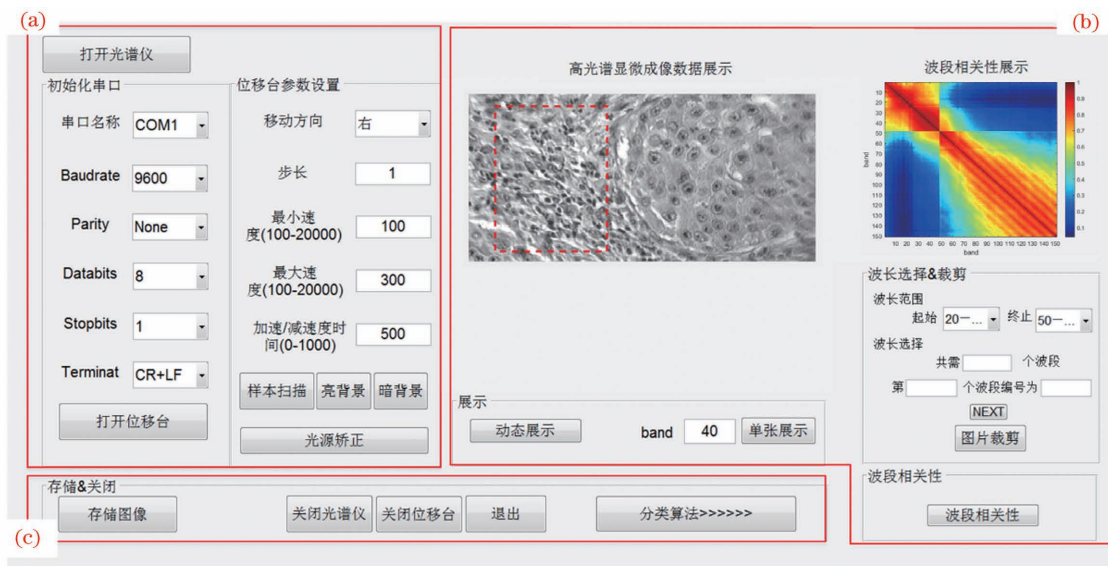


图 4 HMI 数据采集软件的界面

Fig. 4 Interface of HMI data acquisition software

3.3 HMI 数据分析软件

HMI 数据分析软件是对采集到的 HMI 数据立方进行分析,图 5 所示为 HMI 数据分析软件的图形化界面,可分为两个模块。模块 1[图 5(a)]为数据提取与查看模块,可以实现 HMI 图像展示、图中任意区域光谱查看、将三维 HMI 数据转为二维光谱数据、以任意三个单波段图像合成红绿蓝(RGB)图像等功能。模块 2[图 5(b)]为高光谱显微成像数据处理模块,可以分别进行图像数据和光谱数据的分析,实现基于机器学习的样本分类。图像数据处理流程如图 5(b)中虚线框所示,首先选择图像数据所在文件夹,其次选择降维方法[主成分分析(PCA)和偏最小二乘(PLS)]和需要提取的图像特征(颜色矩、HSV 颜色空间、方向梯

度直方图、灰度共生矩阵和局部二值模式),然后选择样本集划分方法(随机划分、自助法和留出法)划分训练集和测试集,最后选择分类模型[极限学习机(ELM)、支持向量机(SVM)、决策树(DT)和随机森林(RF)]。光谱数据处理流程如图 5(b)中实线框所示,可提供四种光谱数据预处理方法[一阶导数法(FD)、二阶导数法(SD)、多元散射校正(MSC)和标准正态变量变换(SNV)],对光谱数据进行预处理和降维后,再依次进行划分数据集、选择分类模型等操作。界面右侧区域可依次展示主成分贡献率图、算法结果图、准确率和 KAPPA 值。

在利用机器学习解决分类问题时,通常采用一些评价指标对分类模型的性能进行评估,其中准确率和

KAPPA 值是较为常见的评估标准。模型准确率(p_0)的计算公式为

$$p_0 = \frac{\sum_i x_{ii}}{N}, \quad (5)$$

式中: N 为样本数; x_{ii} 为第 i 类正确分类样本数。

用 p_e 表示模型的期望精度, 计算公式为

$$p_e = \frac{\sum_{i=1}^n (\sum_{j=1}^n x_{ij} \times \sum_{j=1}^n x_{ji})}{N^2}, \quad (6)$$

式中: x_{ij} 为第 i 类样本被预测为第 j 类样本的数量; n 为类别个数。式(6)右端分子表示所有类别对应的实际数量乘积与预测数量乘积的总和。

KAPPA 系数(K_A)的计算公式为

$$K_A = \frac{p_0 - p_e}{1 - p_e}. \quad (7)$$

KAPPA 系数用于模型的一致性检验, 通常情况下系数的计算结果分布范围为 0~1。将此区间划分为 5 组, 分别表示不同的一致性等级, 当 KAPPA 值大于 0.4 时对应的模型具有参考意义。

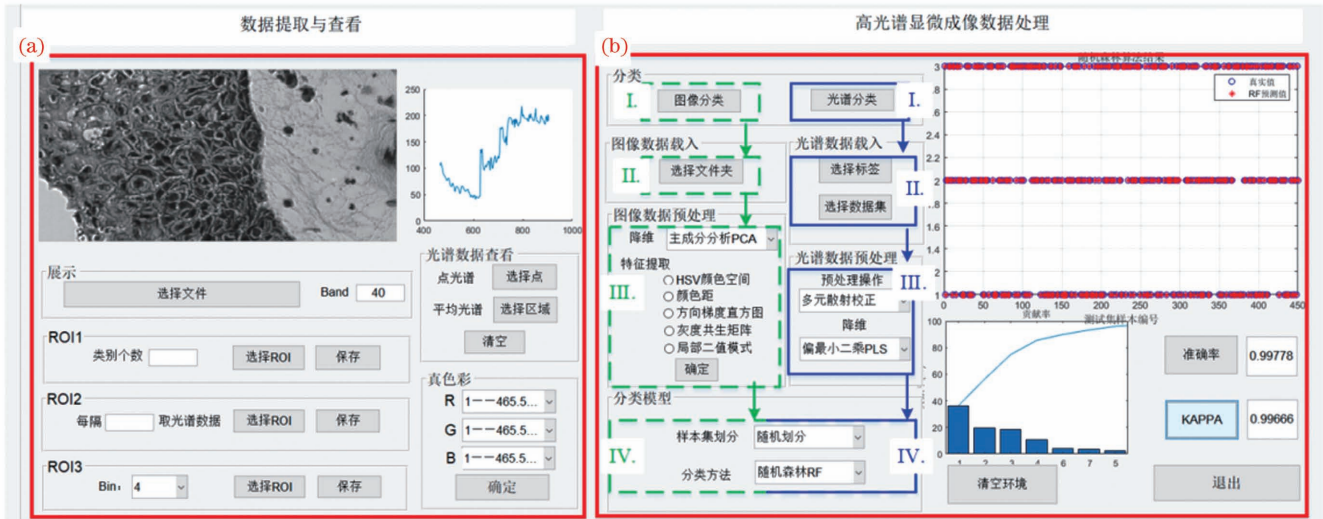


图 5 HMI 数据分析软件界面
Fig. 5 Interface of HMI data analysis software

4 实验测试

为了评估所搭建的推扫式 HMI 系统的性能, 本文对基底细胞癌、鳞状细胞癌及恶性黑色素瘤组织进行 HMI 数据采集与处理。实验样本为两组 H&E 染色皮肤癌组织微阵列, 分别购于西安百思达生物科技有限公司和中科光华(西安)智能生物科技有限公司, 其中包括基底细胞癌 34 例(I 期 24 例, II 期 10 例, III 期 0 例), 鳞状细胞癌 63 例(I 期 13 例, II 期 37 例, III 期 13 例), 恶性黑色素瘤 39 例(I 期 2 例, II 期 30 例, III 期 7 例)。

4.1 皮肤癌的 HMI 数据采集

利用 HMI 系统对皮肤癌组织进行数据采集, 首先用卤钨灯光源预热 10 min, 将皮肤癌组织微阵列置于样品台后, 打开 HMI 数据采集软件, 设置高精度步进电机自动平台移动步长为 $1 \mu\text{m}$ 。依次采集皮肤癌样本的 HMI 数据、暗背景 HMI 数据与亮背景 HMI 数据, 并进行背景光源校正和频域滤波。

图 6 所示为鳞状细胞癌组织的高光谱显微图像, 其中图 6(a)为光源噪声校正和频域滤波后的三维高光谱显微图像, 图 6(b)为图 6(a)中标记点的光谱分布。传统彩色图像数据包含红光(R)、绿光(G)、蓝光(B)三通道, 而 HMI 系统采集的数据包含

151 个波段。根据国际照明委员会确定的 RGB 三基色为波长为 700.0 nm 的红光、波长为 546.1 nm 的绿光和波长为 435.8 nm 的蓝光, 本文选取波段 74 (波长为 700.1 nm)、波段 23(波长为 545.3 nm)和波段 1(波长为 465.5 nm), 分别获取单波段成像 [图 6(c)~(e)], 并进行 RGB 图像合成, 结果如图 6(f) 所示。

4.2 皮肤癌的 HMI 数据分析

首先, 根据推扫式 HMI 系统采集到的图像数据, 对基底细胞癌、鳞状细胞癌和恶性黑色素瘤进行分类。按照图 5 所示的图像数据处理流程, 提取颜色矩、灰度共生矩阵和局部二值模式特征, 利用偏最小二乘方法进行降维, 通过留出法划分出样本集和训练集, 采用支持向量机分类模型进行分类, 最终实现对三类皮肤癌的分类, 准确率为 85%, KAPPA 值为 0.77。

然后, 根据采集到的鳞状细胞癌 I 期、II 期和 III 期的光谱数据, 按照图 5 所示的光谱数据处理流程, 对光谱数据进行标准正态变量变换预处理操作, 并利用偏最小二乘方法进行光谱数据的降维, 采用留出法划分出训练集和测试集, 分别输入不同的分期模型进行分析, 得到的鳞状细胞癌的分期结果如表 1 所示。可以看出, 由随机森林模型所得到的分期结果的准确率和 KAPPA 值最高, 分别为 96.4% 和 0.95。

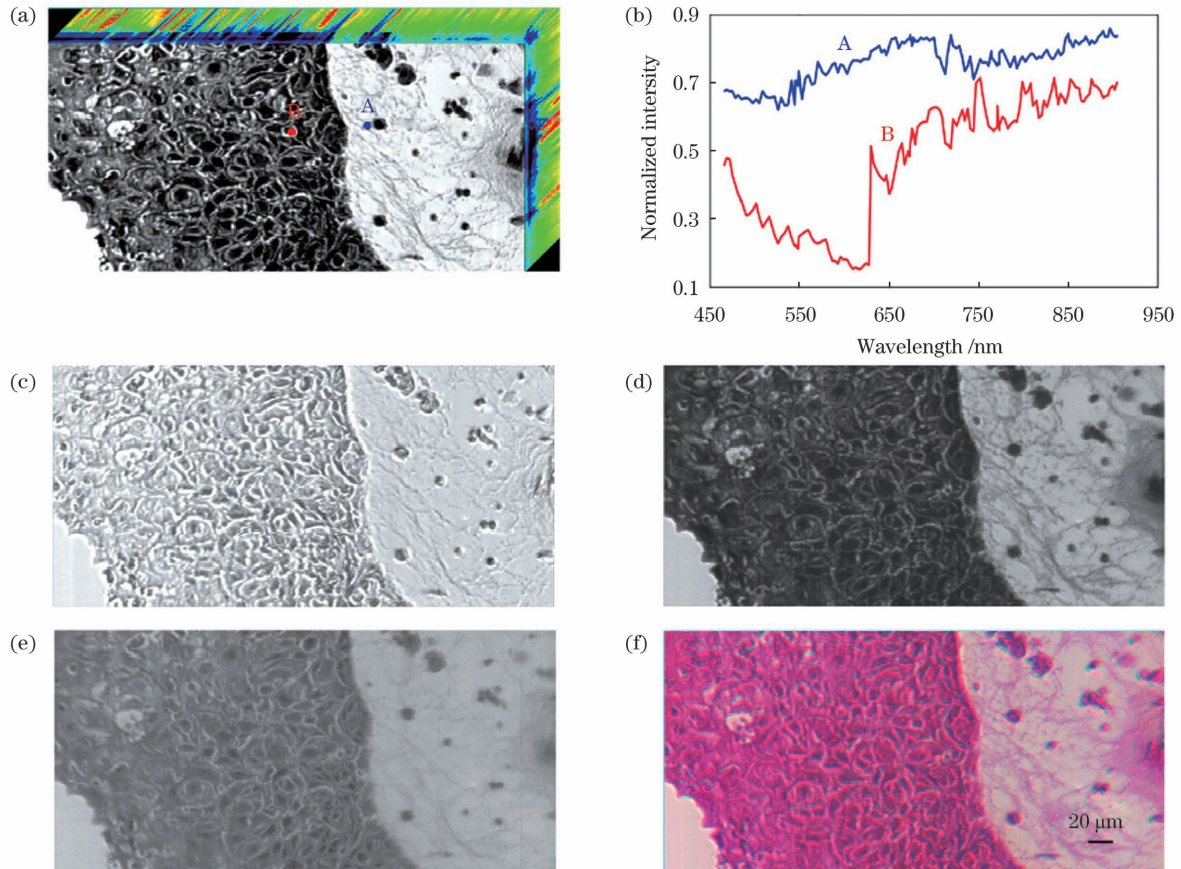


图 6 鳞状细胞癌组织的 HMI 数据分析结果。(a)三维 HMI 数据;(b)图 6(a)中标志点的光谱;(c)单波段图像,波长为 700.1 nm;(d)单波段图像,波长为 545.3 nm;(e)单波段图像,波长为 465.5 nm;(f)RGB 图像

Fig. 6 HMI data analysis results of squamous cell carcinoma tissue. (a) Three-dimensional HMI data; (b) spectra at marked points in Fig. 6(a); (c) single band image with wavelength of 700.1 nm; (d) single band image with wavelength of 545.3 nm; (e) single band image with wavelength of 465.5 nm; (f) RGB image

表 1 鳞状细胞癌的分期结果

Table 1 Staging results of squamous cell carcinoma

Classification model	Accuracy	KAPPA value
ELM	95.8%	0.937
SVM	94.2%	0.913
DT	88.9%	0.833
RF	96.4%	0.947

以上实验结果证明,利用所搭建的 HMI 系统和编写的数据分析软件,可实现对皮肤癌的分类和分期。该方法在癌症诊断领域中具有很大的应用潜力。

5 结 论

利用高精度步进电机型自动平台、高光谱相机及物镜等器件,搭建了一套推扫式 HMI 系统,光谱范围为 465.5~905.1 nm,光谱分辨率约为 3 nm,共有 151 个工作波段,成像视场为 $400.18 \mu\text{m} \times 192.47 \mu\text{m}$,放大率为 28.15,空间分辨率为 $1.10 \sim 1.38 \mu\text{m}$ 。同时利用 MATLAB 软件编制了具有图形化用户界面的 HMI 数据采集和数据分析软件,可实现高光谱图像数据的采集与处理一体化,操作便捷、直观性强,能够帮助医生更高效地进行病理诊断。利用该系统对皮肤癌

组织切片进行成像,可以获取高光谱分辨率和高空间分辨率的图像,基于机器学习算法以较高准确率实现了对不同皮肤癌的分类和鳞状细胞癌的分期鉴别。实验结果表明 HMI 技术和机器学习的结合在医学领域具有很大的应用潜力。

参 考 文 献

- [1] Goetz A F, Vane G, Solomon J E, et al. Imaging spectrometry for Earth remote sensing[J]. Science, 1985, 228(4704): 1147-1153.
- [2] Xie A G, Sun D W, Xu Z Y, et al. Rapid detection of frozen pork quality without thawing by Vis-NIR hyperspectral imaging technique[J]. Talanta, 2015, 139: 208-215.
- [3] 刘立新, 何迪, 李梦珠, 等. 基于高光谱技术与机器学习的新疆红枣品种鉴别[J]. 中国激光, 2020, 47(11): 1111002. Liu L X, He D, Li M Z, et al. Identification of Xinjiang jujube varieties based on hyperspectral technique and machine learning [J]. Chinese Journal of Lasers, 2020, 47(11): 1111002.
- [4] 姜杰, 于泉洲, 梁天全, 等. 星载成像高光谱的湿地景观光谱特征分析[J]. 光谱学与光谱分析, 2022, 42(2): 524-529. Jiang J, Yu Q Z, Liang T Q, et al. Analysis of spectral characteristics of different wetland landscapes based on EO-1 Hyperion [J]. Spectroscopy and Spectral Analysis, 2022, 42(2): 524-529.
- [5] Restaino R, Vivone G, Addesso P, et al. Hyperspectral sharpening approaches using satellite multiplatform data[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(1): 578-596.
- [6] 徐俊杰, 卜令兵, 刘继桥, 等. 机载高光谱分辨率激光雷达探测

- 大气气溶胶的研究[J]. 中国激光, 2020, 47(7): 0710003.
- Xu J J, Bu L B, Liu J Q, et al. Airborne high-spectral-resolution lidar for atmospheric aerosol detection[J]. Chinese Journal of Lasers, 2020, 47(7): 0710003.
- [7] Liang H D. Advances in multispectral and hyperspectral imaging for archaeology and art conservation[J]. Applied Physics A, 2012, 106(2): 309-323.
- [8] Cucci C, Delaney J K, Picollo M. Reflectance hyperspectral imaging for investigation of works of art: old master paintings and illuminated manuscripts[J]. Accounts of Chemical Research, 2016, 49(10): 2070-2079.
- [9] Siddiqi A M, Li H, Faruque F, et al. Use of hyperspectral imaging to distinguish normal, precancerous, and cancerous cells[J]. Cancer, 2008, 114(1): 13-21.
- [10] Baltussen E J M, Kok E N D, de Koning S G B, et al. Hyperspectral imaging for tissue classification, a way toward smart laparoscopic colorectal surgery[J]. Journal of Biomedical Optics, 2019, 24(1): 016002.
- [11] 刘立新, 李梦珠, 赵志刚, 等. 高光谱成像技术在生物医学中的应用进展[J]. 中国激光, 2018, 45(2): 0207017.
- Liu L X, Li M Z, Zhao Z G, et al. Recent advances of hyperspectral imaging application in biomedicine[J]. Chinese Journal of Lasers, 2018, 45(2): 0207017.
- [12] Zhang X F, Zhang C G, Wang Y, et al. High-resolution hyperspectral microscopic imaging with single acousto-optic tunable filter based on double filtering[J]. IEEE Access, 2020, 8: 11570-11576.
- [13] Xu Z P, Jiang Y M, He S L. Multi-mode microscopic hyperspectral imager for the sensing of biological samples[J]. Applied Sciences, 2020, 10(14): 4876.
- [14] Yu C B, Yang J, Song N, et al. Microlens array snapshot hyperspectral microscopy system for the biomedical domain[J]. Applied Optics, 2021, 60(7): 1896-1902.
- [15] Wang C J, Zheng W L, Bu Y G, et al. Multi-scale hyperspectral imaging of cervical neoplasia[J]. Archives of Gynecology and Obstetrics, 2016, 293(6): 1309-1317.
- [16] Aref M, Aboughaleb I H, Youssef A B, et al. Novel approach exploiting the hyperspectral imaging system for breast cancer therapy and diagnosis[J]. Archives of Breast Cancer, 2020, 7(4): 189-201.
- [17] 杜剑, 胡炳樑, 张周锋. 基于卷积神经网络与显微高光谱的胃癌组织分类方法研究[J]. 光学学报, 2018, 38(6): 0617001.
- Du J, Hu B L, Zhang Z F. Gastric carcinoma classification based on convolutional neural network and micro-hyperspectral imaging[J]. Acta Optica Sinica, 2018, 38(6): 0617001.
- [18] Chen Y, Zhu S, Fu S, et al. Classification of hyperspectral images for detection of hepatic carcinoma cells based on spectral-spatial features of nucleus (Article)[J]. Journal of Innovative Optical Health Sciences, 2020, 13(1): 2050002.
- [19] 王健生, 李庆利, 周梅, 等. 显微高光谱成像的皮肤黑色素瘤浅表扩散深度识别方法[J]. 红外与毫米波学报, 2020(6): 749-759.
- Wang J S, Li Q L, Zhou M, et al. Identification and measurement of cutaneous melanoma superficial spreading depth using microscopic hyperspectral imaging technology[J]. Journal of Infrared and Millimeter Waves, 2020(6): 749-759.
- [20] Deng Y J, Yin J T, Wang Y, et al. ResNet-50 based method for cholangiocarcinoma identification from microscopic hyperspectral pathology images[J]. Journal of Physics Conference Series, 2021, 1880(1): 012019.

Design and Experiment of Push-Broom Hyperspectral Microscopic Imaging System

Qi Meijie¹, Liu Lixin^{1,2*}, Li Yanru¹, Liu Yujie¹, Zhang Zhoufeng², Qu Junle³

¹School of Optoelectronic Engineering, Xidian University, Xi'an 710071, Shaanxi, China;

²Key Laboratory of Spectral Imaging Technology, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, Shaanxi, China;

³Key Laboratory of Optoelectronic Devices and Systems of Ministry of Education and Guangdong Province, College of Physics and Optoelectronic Engineering, Shenzhen University, Shenzhen 518060, Guangdong, China

Abstract

Objective Hyperspectral microscopic imaging (HMI) technology combines optical microscopy and hyperspectral imaging to obtain both image and spectral information, thereby revealing spatial distribution and physical and chemical properties of a sample simultaneously. HMI, a novel nondestructive optical imaging technology, can be used to diagnose normal/cancerous tissues with high accuracy, sensitivity, and specificity. However, HMIs have a large amount of data and a complex data structure; thus, systematic and detailed data interpretation is required in cancer diagnosis. In this study, a push-broom HMI system is designed and developed, and the graphical user interface (GUI)-based software for system control, data acquisition, and data analysis is programmed to aid doctors in pathological diagnosis. The classification and staging of skin cancers (basal cell carcinoma, squamous cell carcinoma, and malignant melanoma) are studied on the basis of HMI technology and machine learning algorithms to confirm the performance of the system software. We hope that our HMI system, GUI-based software, and experimental results will be useful in cancer diagnosis and have application potential in biomedicine.

Methods First, a push-broom HMI system consists of a halogen lamp, objective lens, sample stage, single-axis motorized translation stage, two-axis manual translation stage, hyperspectral line-scan camera, and other optical devices (Fig. 1). The halogen lamp illuminates the sample on the sample stage. The transmitted light is collected by the objective lens and directed to the hyperspectral camera after passing through the mirror and lens group in sequence to obtain one-dimensional (1D) spatial and spectral information. The motorized translation stage controls the sample stage to move in

the x -direction with a step size of $1\ \mu\text{m}$ for HMI data cube acquisition. The spectral resolution of the hyperspectral camera is calibrated and calculated based on the sensor configuration (Fig. 2). HMI system performance parameters, such as spatial resolution, field of view, and magnification, are obtained by imaging a resolution target. Second, the software with graphical user interfaces for system control, data acquisition, and data analysis is programmed using MATLAB. Several machine learning-based data processing methods are provided. Finally, the HMI data cubes of basal cell carcinoma, squamous cell carcinoma, and malignant melanoma tissues are obtained using the HMI system and data acquisition software; subsequently, the classification and staging of skin cancer are studied using data analysis software.

Results and Discussions The push-broom HMI system has a spectral range of 465.5–905.1 nm, with a spectral resolution of $\sim 3\ \text{nm}$, field of view of $400.18\ \mu\text{m} \times 192.47\ \mu\text{m}$, system magnification of 28.15, and actual spatial resolution of $1.10\text{--}1.38\ \mu\text{m}$ (Fig. 3); it can collect a data cube of $2048\ \text{pixel} \times 985\ \text{pixel} \times 151$. Additionally, GUI-based HMI data acquisition software and analysis software are designed and programmed using MATLAB. The data acquisition software includes the following three modules (Fig. 4): HMI system control and data acquisition module for controlling the hyperspectral camera and motorized translation stage, HMI data acquisition, light source background correction, and frequency domain filtering; HMI data display and processing module for displaying or cropping the HMI data cube and single-band image and calculating the correlation between each band; and save and exit module for saving the data processing results and exiting the acquisition software. The data analysis software consists of the following two modules (Fig. 5): a data extraction and viewing module that can realize HMI image display, spectrum viewing in the region of interest, converting 3D HMI data into 2D spectral data, and synthesizing RGB images with any three single-band images; and an HMI data processing module that can analyze image and spectral data and realize sample classification based on machine learning. HMI data from basal cell carcinoma, squamous cell carcinoma, and malignant melanoma are obtained to evaluate the performance of the system, and the machine learning is used to achieve the classification of three types of skin cancers and staging of squamous cell carcinoma. Spectral distribution, as well as 3D HMI, single-band, and RGB images, can be displayed (Fig. 6). The classification of three types of skin cancers based on image data is achieved using the data analysis software, and the highest classification accuracy of 85% and KAPPA value of 0.77 are obtained using color moment, gray-level co-occurrence matrix and local binary pattern as image features, partial least squares for dimensionality reduction, hold-out method for dividing the dataset, and a support vector machine models for classification. The optimal model for spectral data staging of squamous cell carcinoma corresponding to the standard normal variable transformation for spectral preprocessing, partial least squares for dimensionality reduction, hold-out method for dividing the dataset, and random forest for staging. The highest staging accuracy of 96.4% and a KAPPA value of 0.95 are obtained (Table 1).

Conclusions In this study, a push-broom HMI system that can simultaneously obtain image and spectral information is developed to reveal spatial distribution and physicochemical properties of the samples. The HMI system can provide a data cube of $2048\ \text{pixel} \times 985\ \text{pixel} \times 151$, a spectral resolution of $\sim 3\ \text{nm}$, and actual spatial resolution of $1.10\text{--}1.38\ \mu\text{m}$. The HMI data acquisition software and analysis software are programmed using MATLAB. The graphical user interface of the software can standardize experiment procedures, allows intuitive data collection and processing, and provides analysis results, all of which can assist doctors in pathological diagnosis. Using this HMI system to image skin cancer tissues, high spectral and spatial resolution images are obtained, and the classification of different skin cancers and staging of squamous cell carcinomas can be achieved with high accuracy using machine learning algorithms. Our study shows that the combination of HMI technology and machine learning has significant application potential in the field of biomedicine.

Key words medical optics; hyperspectral microscopic imaging; skin cancer; graphical user interface; cancer diagnosis