**高光谱在冬虫夏草含量及真伪鉴别中的应用**

赵俊华1，孙梅1\*，陈兴海2

(1.北京工商大学，北京 100048 2.北京卓立汉光仪器有限公司，北京 101102)

**摘要：**本文以冬虫夏草粉末为研究对象，利用高光谱成像技术建立对冬虫夏草粉末的真假鉴别及含量判断的无损检测模型：首先，通过光谱范围为1um~2.5um的高光谱分析仪提取真假及不同含量的冬虫夏草粉末的反射光谱信息。利用反射率校正法去除光谱的噪声与背景后，通过主成分分析（PCA）提取真假样本，再结合偏最小二乘法（PLSA）对样本含量进行分析判断。研究结果表明，基于高光成像技术可以实现对冬虫夏草粉的真假辨别，并可准确地判别出冬虫夏草粉末的有效含量。

**关键词：**高光谱成像技术；冬虫夏草粉末；真假鉴别；含量判断

**中图分类号：S513 文献标识码： A 文章编号：**

**Application of Hyperspectral Imaging Technique for Cordyceps Powder Content and** **True****and****False****Identification**

ZhAO Jun-hua 1，SUN Mei 1，CHEN Xing-hai2

(1. Beijing Technology and Business University, Beijing 100048，China.

2. ZOLIX INSTRUMENTS CO.,LTD,Beijing,101102,China)

**Abstract：**Based on the cordyceps powder, this paper uses the hyperspectral imaging technology to establish the nondestructive testing model which can identify true and false of cordyceps powder and judge the content: firstly, the spectrum in range of 1um to 2.5um was collected. The reflectance correction was adopted to eliminate the noise and background in original spectrum. The true and false was identified by principal component analysis method(PCA). The content of cordyceps powder was analyzed by the partial least squares (PLSA).The research results show that based on the hyperspectral imaging technology, it is convenient to identify the true and false of cordyceps, it is suitable for quality evaluation of traditional cordyceps powder.

**Key words**：Hyperspectral imaging; Cordyceps powder; True and false identification; Content measurement

基金项目：国家自然科学基金项目（61473009）、北京市自然科学基金项目（4132008）、北京市组织部优秀人才培养D类项目（2013D005003000008）联合资助

作者简介：赵俊华（1989-），女，硕士，主要从事高光谱图像与食品检测等研究。

通讯作者：孙梅(1976-)，博士，副教授

冬虫夏草最初载于《本草从新》[1]，主产于我国四川、云南、青海、西藏、甘肃等省区，是我国名贵中药材之一。与人参、鹿茸并称为补品“三宝”，具有良好的医疗保健作用。由于冬虫夏草人工抚育技术尚未突破，其来源仅靠野生采挖，产量较少，价格昂贵，少数不法分子为谋取暴利而采取以次充好、掺假充真等手段，导致市场上冬虫夏草的各种伪品层出不穷，严重危害消费者和患者的身体健康。因此，完善虫草的品种鉴定研究，准确地判定其来源，对药品的监督和检验显得尤为必要。目前常用的中药检测技术有色谱法、质谱法和光谱法[ 2] 等，其基本原理都是通过分离、提纯中药检品中的有效活性成分，并对其进行检测，以实现检品的定性、定量分析。这些技术虽然各有优势，但都难以快速、无损地判别中药粉末掺假的问题。

本文运用近红外高光谱成像技术对冬虫夏草粉末的真假鉴别及含量进行了无损检测的判断。高光谱成像技术把二维成像和光谱技术融为一体，其高光谱数据包含光谱和图像信息， 可以同时表征被测对象的外部特征和内部信息，近年来逐渐受到生物医学、精细农业、食品安全等许多领域的重视。例如，外在品质检测如水果、蔬菜表面损伤、淤痕[3-4];内部品质检测如水果的可溶性固体含量、水分含量、坚硬程度[5-6]，猪肉的嫩度[7-8]，鳕鱼的新鲜程度[9]，菠菜叶片硝酸盐含量等[10]；而食品安全检测主要指食品中是否含有可能损害或威胁人体健康的物质，如苹果、哈密瓜表面排泄物污染检测[11-13]，鸡肉排泄物污染检测[14]，玉米、辣椒等黄曲霉毒素检测等。

高光谱成像技术的定义是在多光谱成像的基础上，在从紫外到近红外（200-2500nm）的光谱范围内，利用成像光谱仪，在光谱覆盖范围内的数十或数百条光谱波段对目标物体连续成像。在获得物体空间特征成像的同时，也获得了被测物体的光谱信息。高光谱成像技术具有超多波段（上百个波段）、高的光谱分辨率（几个nm）、波段窄（≤10-2λ）、光谱范围广（200-2500nm）和图谱合一等特点。优势在于采集到的图像信息量丰富，识别度较高和数据描述模型多。由于物体的反射光谱具有“指纹”效应，不同物不同谱，同物一定同谱的原理来分辨不同的物质信息，它不追求单一成分的控制，其整体性和模糊性可以提供丰富的中药信息，能够更加有效地体现中药成分的综合作用，从而更好地鉴别中药真伪，评价中药质量。

1 材料与方法

1.1 仪器与设备

高光谱图像系统根据三维数据块的获取方式不同可分为两种：第一种是基于滤波器（或滤波片）的高光谱图像系统，它获取高光谱图像数据的方式是通过连续采集在一系列波长下的鱼肉样品的二维图像来得到；第二种方法是基于图像光谱仪的高光谱图像系统，它是采用“推扫式”成像的方法来获得高光谱图像数据。高光谱成像数据采集采用北京卓立汉光仪器有限公司的GaiaSorter高光谱分选仪系统。该系统主要由高光谱成像仪，CCD相机、光源、暗箱、计算机组成，如图1所示。



**图1 GaiaSorter高光谱分析仪系统**

**Fig.1 GaiaSorter hyperspectral imager system**

其中，实验仪器参数设置如表1所示。

**表1 高光谱设备参数设置**

**Tab.1 Parameter setting of hyperspectral system**

1.2实验样本

|  |  |  |
| --- | --- | --- |
| 名称 | 数量 | 说明 |
| GaiaSorter高光谱分选仪 | 1台 | 技术参数：1.光谱扫描范围: 350-1000nm 2.光谱分辨率：2.8 nm3.采样间隔：1.9nm4.测定速度: 每个样品＜ 1分钟主要特点：1.快速准确；2.不需要任何化学试剂3.生肉及熟肉产品均可检测。 |

样本是由青海唐古拉药业公司提供的10种虫草粉样本。其中伪品两个（9号与10号），标准样品三个（6号、7号与8号），以及将8号与10号真假样本以不同比例混合的样品（1~5号），样品的配置如图2所示。



图2 冬虫夏草的实验样品

**Fig.2 Cordyceps experimental samples**

1.3 实验步骤

在进行图像采集前，为了保证图像的清晰程度，应该根据光源的照度预先对光谱相机摄像头的曝光时间进行设定，同时为了避免图像空间上的分辨率失真，对于输送装置速度的调整也是非常必要的。并且为了克服光强分布较弱的波段所存在的图像噪声和暗电流的影响，首先需要扫描标准白板以采集反射率为1的全白标定图像Dw，而后盖上摄像头的盖子以采集反射率为0的全黑标定图像Dd，进行过黑白校正后，再进行样本光谱图像Ds的采集，由此可以根据公式1得到的黑白校正后的相对样品光谱图像的感兴趣像素区域c或波段i处的反射率R：

 （1）

在数据采集过程中，为了获取到所要扫面空间上每个像素点在整个光谱区域上的光谱数据，在光学焦平面的垂直方向上，需要线性探测器对其进行横向扫描。与此同时，置于输送装置上的实验样本作垂直于摄像机的纵向移动，从而完成了对整个实验样本图像的采集过程。上述过程采集到的鱼肉样本图像数据块，不仅含有特定像素光谱信息，同时也含有特征波段下的图像信息。具体操作步骤如下：

（1）打开GaiaSorter系列高光谱分选仪并启动计算机，运行spectraSENS高光谱数据采集软件，在软件界面进行仪器连接检测和预热；

（2）并将上述实验样品分别置于载物台上，放入GaiaSorter系列高光谱分选仪的载物台上；

（3）设置完测试参数后，点击“开始扫描”，约40s即可采集到一分样品的光谱信息，即样品光谱图，每个样品分别扫描10次取其平均值作为相应样品的光谱反射率值。

（4）每个样品光谱数据扫描完成后，将载物台上的残留物质清理干净，以此确保每次扫描得到的光谱准确性。

**2 结果与讨论**

由于主成分图像都是由原始数据中的各个波段下的图像经过线性组合而成，根据 （其中，为第m个主成分，为该主成分的权重系数，为单个波段的原始图像）。比较该线性组合的权重系数，如果权重系数绝对值越大，对主成分图像贡献就越大。最大权重系数所对应的波长下的图像为最佳特征波长图像。为了选取最佳的波长组合，必须保证它们的权重系数绝对值尽可能大，同时还要保证它们之间有一定的波长间隔。将测试原始数据进行降噪处理后，通过主成分分析变换（PCA）进行背景扣除后，再次进行主成分分析变换（PCA）后结果如图3所示。



1#

10#

2#

4#

6#

8#

9#

7#

5#

3#

**图3 主成分分析的结果**

**Fig. 3 The results of principal component analysis**

通过主成分分析，可以明显将9#、10#样品与其它样品进行区分，由此可以判断9#与10#样品为伪样品。如图4为6#、9#、10#样品的典型光谱，红色为6#、蓝色和绿色分别为9#、10#。



**图4 6#、9#、10#样品典型光谱**

**Fig.4 Reflectance spectra of the 6#、9#、10# samples**

偏最小二乘法是一种数学优化的技术，它主要是通过最小化误差的平方和来找到一组数据的最佳函数匹配，然后用最简的方法求得一些绝对不可知的真值，而令误差平方之和为最小。偏最小二乘法就相当于将多元线性回归分析、典型的相关分析以及主成分分析方法融合在一起的数学方法。偏最小二乘法主要是通过主成分分析法将多为空间数据的曲线压缩到较低维的空间数据上，使其原曲线分解为多种主成分分析曲线，而不同的曲线的主成分分别代表不同的主分和因素间对曲线的贡献率，选取贡献率较大的主成分，去除有干扰组分和干扰因素的主成分，仅仅将贡献率较高的主成分与质量参数进行回归。因此，采用偏最小二乘法建立的模型与传统的多元线性回归模型相比，主要有如下特点：

（1）能够在存在严重的自变量多重相关性的条件下进行回归建模；

（2）在最终建立的模型中将包含所有的原有自变量；

（3）能够在变量个数多于样本点个数的条件下进行回归建模；

（4）模型比传统的多元线性回归模型更易于辨识系统的噪声与信息；

（5）在采用偏最小二乘回归所建模型中，每一个自变量的回归系数将更容易解释。

进一步将6#样品与8#样品做为样本，将7#样品做为未知区域进行偏最小二乘法变换（PLSA）实现分类判别，判定结果如表2所示。

**表2 分类判别结果**

**Fig.2 The result of Classification**

|  |  |
| --- | --- |
| Predicted as:  | # Predicted |
| 6# 8# Not Classified Total 3170 (100%)  | 122 (3.85%)2933 (92.5%)115(3.63%)3170 (100%) |

分析结果表明：7#样品为8#与6#的混合样品，混合比例约为3.85%，92.5%，其中有约3.63%的区域无法归类差别。将8#样品与10#样品做为样本，将1～5#样品作为未知区域进行偏最小二乘法变换（PLSA）实现分类判别，判断8#与10#样本的混合比例，结果如下：

**1# sample：**

|  |  |
| --- | --- |
| Predicted as:  | # Predicted |
| 8#  | 1841 (91.1%) |
| 10#  | 168 (8.31%) |
| Not Classified  | 12 (0.594%) |
| Total 2021 (100%)  | 2021 (100%) |

**3# sample：**

|  |  |
| --- | --- |
| Predicted as:  | # Predicted |
| 8#  | 1816 (90.5%) |
| 10#  | 182 (9.07%) |
| Not Classified  | 8 (0.399%)0  |
| Total 2006 (100%)  | 2006 (100%) |

 **4# sample：**

|  |  |
| --- | --- |
| Predicted as:  | # Predicted |
| 8#  | 1905 (97.5%) |
| 10#  | 6 (0.307%) |
| Not Classified  | 42 (2.15%)8  |
| Total 1953 (100%)  | 1953 (100%) |

 **5# sample：**

|  |  |
| --- | --- |
| Predicted as:  | # Predicted |
| **8#**   | **1866 (65.9%)** |
| **10#**  | **959 (33.9%)** |
| Not Classified  | 5 (0.177%)  |
| Total 2830 (100%)  | 1. 0%)
 |

3 结论

通过以上对冬虫夏草粉样品在近红外波段的反射高光谱图像采集，经过主成分分析，可有效的对冬虫夏草粉末的真伪进行鉴别。进一步通过偏最小二乘法分析对样本区域进行分析判断，可对样品的有效成分含量进行鉴别。由于目前样本采样数量较小，对于成分含量判断的准确性还需进一步实验验证。

实验初步验证了高光谱技术在虫草粉鉴别的可行性。进一步还需通过实验和分析判断出理想的特征波段，以降低数据采集量，并进一步优化数据分析模型与数据处理速度，从而达到在线检测的速度与准确性要求。

**参考文献:**

[1]张白翟，爱华.中药冬虫夏草的真伪鉴别[J].学员天地，2005，12（3）：58.

 ZHANG Bai-zhai，AI Hua. Identification of Traditional Chinese Medicine Cordyceps Sinensis [J]. Students of Heaven and Earth，2005，12（3）：58.

[2] 聂黎行，王钢力，李志猛等.近红外光谱法在中药生产过程分析中的应用[J].光学学报，2009，29 ( 2) : 541-547

 Nie Li-hang， WANG Gang， LI Zhi-meng et al. Near Infrared Spectroscopy in The Application of Traditional Chinese Medicine Production Process Analysis [J]. Journal of Optics， 2009，，29 (2) : 541-547

[3]Nicola B，L tze E，Peirs A. Non-Destructive Measurement of Bitter Pit in Apple Fruit Using NIR Hyperspectral Imaging[J].Postharvest Biology and Technology，2006，40: 1－6.

[4]Xing J，Bravo C，Jancsók P T，et al. Detecting Bruises on ‘Golden Delicious’ Apples Using Hyperspectral Imaging with Multiple Wavebands [J]. Biosystems Engineering，2005，90:27－36.

[5]Rajkumar P，Wang N，EImasry G. Studies on Banana Fruit Quality and Maturity Stages Using Hyperspectral Imaging[J].Journal of Food Engineering，2012，108: 194－200.

[6]吴建虎，彭彦昆，陈菁菁，等． 基于高光谱散射特征的牛肉品质参数的预测研究[J]． 光谱学与光谱分析，2010，30( 7) :1 815 - 1 819．

 WU Jian-hu ，PENG Yan-kun， CHEN Jing-jing， et al. Based on the Hyperspectral Scattering Characteristics of Beef Quality Parameters Prediction Research [J]. Spectroscopy and Spectral Analysis， 2010， 30 (7) : 1 815-1 815.

[7]Wu J H，Peng Y K，Li Y Y，et al． Prediction of Beef Quality Attributes Using VIS /NIR Hyperspectral Scattering Imaging Technique[J]． Journal of Food Engineering，2012，109( 2) : 267 - 273．

[8]Sivertsen A H，Kimiya T，Heia K. Automatic Freshness Assessment of Cod( Gadus Morhua) Fillets by Vis /Nir Spectroscopy [J].Journal of Food Engineering，2011，103: 317－323.

[9] 薛利红，杨林章. 基于可见近红外高光谱的菠菜硝酸盐快速无损测定研究[J].光谱学与光谱分析，2009，29(4)：926-930

XUE Li-hong，yang Lin-zhang. Based on the Visible Near Infrared Hyperspectral Spinach Nitrate Quick Nondestructive Determination of Study [J]. Spectroscopy and Spectral Analysis， 2009， 29 (4) : 926-930

[10]Vargas A M，Kim M S，Tao Y，et al. Detection of Fecal Contamination on Cantaloupes Using Hyperspectral Fluorescence Imagery[J].Journal of Food Science，2005，70: E471－E476.

[11]Lefcout A M，Kim M S.Technique for Normalizing Intensity Histograms of Images When Approximate Size of The Target is Known: Detection of Feces on Apples Using Fluorescence Imaging [J]. Computers and Electronics in Agriculture，2006，50:135－147.

[12]Lefcout A M，Kim M S. Systematic Approach for Using Hyperspectral Imaging Data to Develop Multispectral Imaging Systems: Detection of Feces on Apples [J]. Computers and Electronics in Agriculture，2006，54: 22－35.

[13]Park B，Windham W R，Lawrence K C，et al. Contaminant Classification of Poultry Hyperspectral Imagery Using A Spectral Angle Mapper Algorithm[J]. Biosystems Engineering，2007，96:323－333.

[14]Yao H，Hruska Z，Brown R L，et al. Hyperspectral Bright Greenish － Yellow Fluorescence ( BGYF ) Imaging of Aflatoxin Contaminated Corn Kernels[J]. Proceedings of the SPIE，2006，6381: 63810B1－63810B8.