

文章编号: 1672-8785(2015)07-0038-07

人体头发中重金属元素铬含量的高光谱检测

夏芷玥 刘 浩 林志恒 梅 根 毛

(武汉大学遥感信息工程学院, 武汉 430079)

摘 要: 头发中的重金属含量可以反映出人体健康的变化。提出了运用高光谱数据检测头发中重金属元素铬含量的方法。对头发的透射率波长曲线进行了包络线消除、吸收特征参量化等处理。以化学检测的铬含量作为标准数据, 化学检测精度可达 90% 以上。然后训练人工神经网络, 通过调节网络的隐含层层数、结点数和激活函数来优化模型。实验计算表明, 隐含层层数为 1, 结点数为 7 或 9 的人工神经网络的预测效果较好。利用统计实验结果对人工神经网络的内部精度和外部精度进行评价。人体头发中铬的敏感波段为 1380 nm~1550 nm、1880 nm~2100 nm、2120 nm~2210 nm; 训练后的神经网络预测的均方根误差为 13%, 精度达 87%。实验结果表明, 应用高光谱技术可以快速无损地检测人体头发中的重金属元素铬的含量。

关键词: 高光谱; 头发; 重金属; 铬; BP 神经网络

中图分类号: TP79 **文献标志码:** A **DOI:** 10.3969/j.issn.1672-8785.2015.07.008

Estimation of Heavy Metal Chromium Content in Human Hair Based on Hyperspectral Data

XIA Zhi-yue, LIU Hao, LIN Zhi-heng, MEI Gen, MAO Zhe

(School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China)

Abstract: The heavy metal content in human hair can reflect the health status of a human body. A method for using hyperspectral data to detect the heavy metal element chromium content in human hair is put forward. The processing such as continuum removing and absorption characteristic parametrization is performed on the transmittance wavelength curves of some hair samples. Taking the chromium content detected by a chemical method as the standard data, the accuracy of chemical detection can reach above 90%. Then, the BP artificial neural network is trained and the model is optimized by adjusting the network number of hidden layer nodes, the number of nodes and the activated function. The experimental calculation results show that the artificial network which has one hidden layer and seven or nine nodes has a better prediction result. Both internal precision and external precision of the artificial network are evaluated by using the statistical experimental results. The Cr metal in human hair is sensitive in the wavebands from 1380 nm to 1550 nm, from 1880 nm to 2100 nm and from 2120 nm to 2210 nm. The trained BP artificial neural network has a prediction error RMS of 13% and an accuracy of 87%.

收稿日期: 2015-06-03

基金项目: 国家级大学生创新训练计划项目 (201410486068)

作者简介: 夏芷玥 (1993-), 女, 湖北武汉人, 在籍本科生, 主要从事定量遥感方面的研究。

E-mail: summerdarling@foxmail.com

The experimental results show that the hyperspectral technology can be used to measure the chromium content in human hair nondestructively and quickly.

Key words: hyperspectral; hair; mental; Cr; BP artificial neural network

0 引言

近年来, 环境污染引发的健康问题成了人们日益关注的焦点, 其中问题之一就是重金属污染^[1]。重金属元素在人体内的含量非常微小, 但它却与人的健康密切相关。体内重金属超标或者缺乏都将不同程度地引起人体生理异常, 甚至导致疾病^[2]。铬元素参与人体内糖和脂肪的代谢, 是人体必需的微量元素。人体中铬的缺乏会引起肥胖甚至糖尿病, 而铬含量超标则会引起人体内循环紊乱, 影响人体健康^[3]。铬可存在于人的头发和血液中, 头发中的铬含量不仅和人体健康相关, 还可以反映一段时间之内人体中铬含量的变化。

目前对头发中微量元素含量的检测方法越来越多元化, 从一份样本中可以连续检测出多种元素。现在应用广泛的检测方法有: (1) 原子吸收法: 石墨炉法、火焰法和氢化物吸收法。(2) 电感耦合等离子体发射光谱法, 可以连续测定头发中的微量元素; (3) 反相离子对高效液相色谱法: 用来测定头发中尼古丁、可天宁的含量^[4-6]。这些检测方法都是化学方法, 检测前需对头发样本进行前处理。目前, 检测前的处理多采用湿消化, 该方法需要破坏样本, 使待测的无机成分释放出来, 再进行测定。

高光谱技术在某些特定波段对于微量元素的探测敏感性高^[7]。相比于化学检测方法, 高光谱的优点在于快速且无损。当利用光谱仪测试样本的光谱曲线时, 通过代入模型进行计算, 在检测的瞬间就可以计算得到头发中的铬含量。这种检测方法无需剪下头发, 也无需消解样本, 可直接测定头发的光谱曲线。

1 概述

实验中, 采集不同性质的样本, 利用波谱仪测定所有样本的光谱透射率和反射率。从所有

样本的光谱曲线中提取光谱吸收参数, 将三分之二的样本数据作为神经网络的建模训练样本, 另外三分之一的样本数据作为检验样本, 测试分析神经网络模型的稳定性和敏感性。具体流程如图 1 所示。

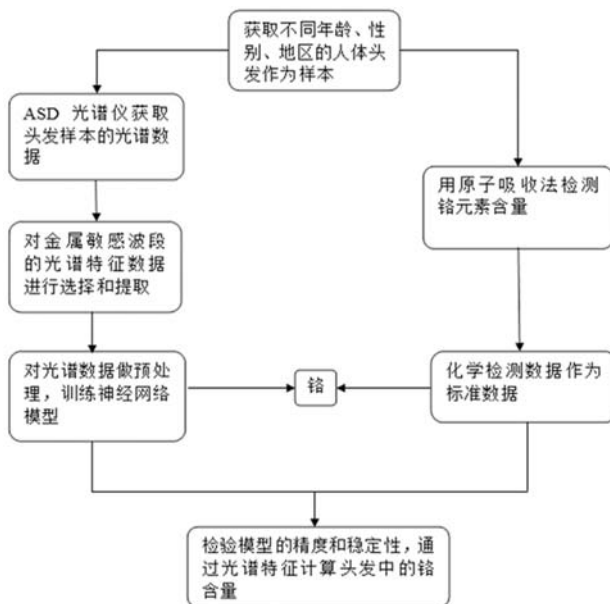


图 1 总流程图

2 数据获取及处理

2.1 样本采集

共采集 18 份不同的样本, 每份样本的重量均大于 0.2 g, 样本包括不同性别、年龄、地区人的头发。从 18 个样本中随机抽取 12 个样本用于训练神经网络, 编号为 1-1~1-12, 6 个样本用于检测模型的稳定性和精度, 其编号为 2-1~2-6。

2.2 高光谱数据获取

获取高光谱数据时选用的仪器是美国 ASD 公司生产的 ASD FieldSpec 4 Pro 光谱仪, 它是一种为地面遥感数据采集专门设计的仪器, 其波长范围为 350 nm~2 500 nm。仪器的外观和测量头发光谱的探头叶片夹如图 2 所示。测定头发光谱前, 先对波谱仪进行白板和黑板的校对, 以保

证测定的准确性。将头发样本铺满叶片夹,以避免外来光源如太阳光、日光灯等对头发光谱造成影响。对样本的透射率和反射率各测量五次,计算取 5 次测量的平均值,采用样本的透射率作为光谱特征曲线。用 ASD 仪器配套的 ViewSpecPro 软件存储并导出得到标准光谱数据。



图 2 ASD FieldSpec 4 Pro 光谱仪

2.3 铬含量的化学检测

实验中的所有样本均在武汉大学化学检测中心接受铬含量测定。该中心技术先进,利用原子吸收法所测得的头发中的铬含量误差可控制在 10% 以内,精度达 90% 以上。实验中,我们将该中心检测得到的铬含量作为标准数据与我们计算得到的铬含量进行了比对。

3 光谱特征参量化

对高光谱曲线特征进行定量表达,用数值化形式来描述反射率随波长的变化特征。首先对光谱曲线进行平滑预处理,去除噪声,再采用包络线消除法先对原始光谱曲线做归一化处理。根据经验法获取重金属的敏感波段,并在这些波段提取光谱吸收特征参数。

3.1 光谱曲线平滑和包络性消除

实际测得的光谱曲线包括三种误差:(1)探头在光谱仪不同波段引起的台阶;(2)随机误差;(3)水气吸收带存在不规则的跳动^[8]。因此,为了反映真实的光谱曲线,必须对光谱曲线进行平滑处理。采用中值滤波法对原始光谱曲线进行处理。

样本透射率的大小受样本自身的许多物理化学参数影响,不便于计算。因此,为了突出光

谱的吸收特征,对光谱进行包络线消除处理^[9]。包络线是与原始光谱曲线相切或相离的曲线,包络线消除法的具体公式为

$$\rho' = \frac{\rho}{\rho_{curve}} \quad (1)$$

式中, ρ' 为包络线消除后的曲线; ρ 为原始光谱曲线; ρ_{curve} 为包络线。

经过滤波处理和包络线消除法处理后的光谱曲线,吸收特征更加明显,便于计算光谱的吸收特征参数。图 3 为样本 1-7 的原始光谱曲线,即波长与透射率的关系图。图 4 为样本 1-7 经平滑处理后的光谱曲线和包络线消除之后的光谱曲线,图中箭头指向处是明显的吸收峰。

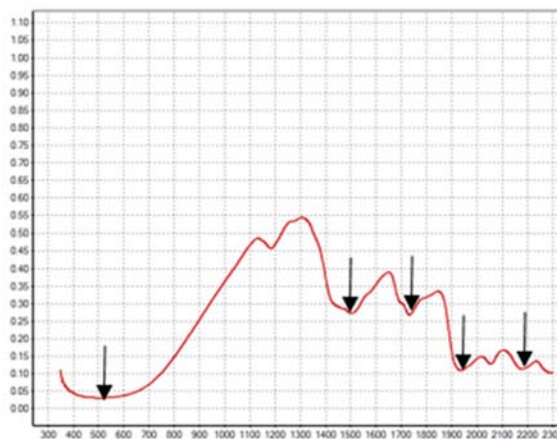


图 3 原始光谱曲线

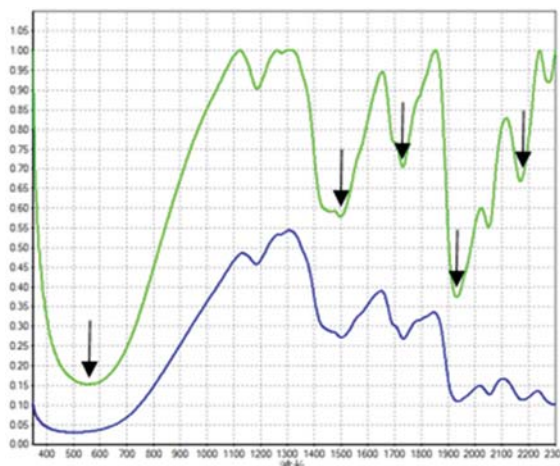


图 4 处理后的光谱曲线

从图 3 中可以看出,原始光谱曲线中存在吸收带,但是吸收特征表现不明显,不利于计算提

取。从经过包络线消除后的光谱曲线中可以更清楚地看到, 头发光谱曲线中的几个吸收峰分别位于 550 nm、1500 nm、1700 nm、1920 nm 和 2170 nm 附近。

3.2 光谱吸收参数提取

铬是一种重金属元素, 近年来, 已有大量用近距离高光谱遥感定量研究重金属的报道。2011 年南京大学地理与海洋科学学院的研究表明, 土壤中的铬含量与原始波段高度相关的波段为 2160 nm 和 1918 nm^[10]。2014 年一项对亚热带土壤铬元素引起高光谱的响应研究中给出土壤中铬元素的敏感波段为 520~530 nm、1400~1590 nm 和 2230~2249 nm, 500 nm 左右处的吸收带主要由铁锰氧化物引起^[11]。高光谱数据是庞大的高维数据, 高维信息在降维过程中存在信息损失的风险, 提取光谱的吸收特征能够有效地分析物质的成分^[12]。从经验法选取头发中铬元素的敏感波段。本次实验中我们提取头发样本光谱曲线中 1500 nm、1930 nm 和 2170 nm 三个吸收位置的六个光谱特征。吸收特征参数量化方法如下:

吸收位置 (P): 吸收峰中光谱最小值对应的波长。

吸收深度 (D): 吸收峰的相对深度。

吸收宽度 (W): 吸收带的最大波长和最小波长之差。

吸收面积 (A): 吸收峰面积的绝对值。

吸收峰的对称性 (S): $S = A_1/A$ 。其中 A_1 为吸收峰的左半部分面积, A 为吸收峰总面积。

吸收指数 (SAI) 为

$$SAI = [S \times R_s + (1 - S) \times R_e] / R_p \quad (2)$$

式中, R_e 为吸收终点处理后的透射相对值, R_s 为吸收起点的透射相对值, R_p 为吸收点 P 的透射相对值。六个吸收参数的具体算法参照参考文献 [13]。

4 基于神经网络模型的铬含量预测

人头发中铬的含量会引起光谱特性曲线某些波段的细微变化, 但是除了铬含量的不同, 头发中的其他物理化学性质也会对相同的波段造成影响。因此, 头发样本的光谱特征不仅仅由铬元素含量引起, 两者之间不是简单的线性关系^[14]。对于复杂的非线性关系建模, 常规的线性回归方法准确性可能不高。人工神经网络具有较强大的自适应性非线性处理能力, 能够有效解决非线性建模问题。本文采用 BP 神经网络 (误差反向传播神经网络) 有效模拟了样本光谱曲线和铬元素含量的关系。

4.1 BP 神经网络对头发中铬含量检测的模型

BP 神经网络包括输入层、隐含层和输出层。输入层和隐含层间的键结值、隐含层和输出层之间的键结值可在训练中计算得到。通过调节隐含层的层数、各隐含层的节点数目和各层之间的激活函数可优化模型的模拟运算效果^[15]。隐含层数量可为一层或两层甚至更多, 但研究表明, 对于生物特性的模拟, 隐含层的数量为一层时效果较其他情况更佳^[16]。隐藏层节点的选取应为上一层结点数的 $3/7 \sim 3/4$ 之间, 并且在优化时选取节点数大于一般经验给出的隐含层节点数来训练网络结构^[17]。

本实验的数据统计处理应用 EXCEL、SPSS、MATLAB 三个软件来统计分析和绘图等。实验选择单隐层的三层神经网络, 输入端为样本光谱曲线在吸收位置 1500 nm 处的吸收参数 D1、W1、A1、S1 和 SAI1, 在吸收位置 1930 nm 的吸收参数 D2、W2、A2、S2 和 SAI2, 在吸收位置 2170 nm 处的吸收参数 D3、W3、A3、S3 和 SAI3, 共 15 个输入端数据。神经网络的隐含层结点数选用 6~10 个, 输入端和隐含层之间的激活函数为双曲正切函数和 sigmoid 函数, 隐含层和输出端的激活函数为恒等函数、双曲正切函数和 sigmoid 函数。比较建立的神经网络模型的内部精度和外部检验精度, 得到拟合度最佳的模型。

实验中, 输入端和隐含层之间的激活函数为双曲正切函数。隐含层和输出端的激活函数为恒等函数时, 内部拟合效果较其他模型好, 所以我们在此情况下来比较隐含层结点数为 6、7、8、9 和 10 时模型的内部精度和外部精度。

4.2 内部精度分析

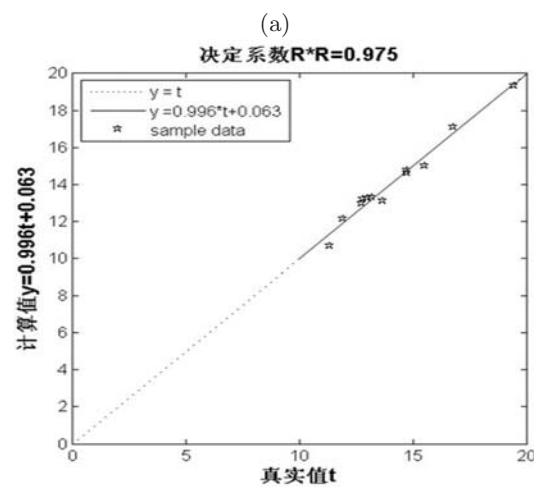
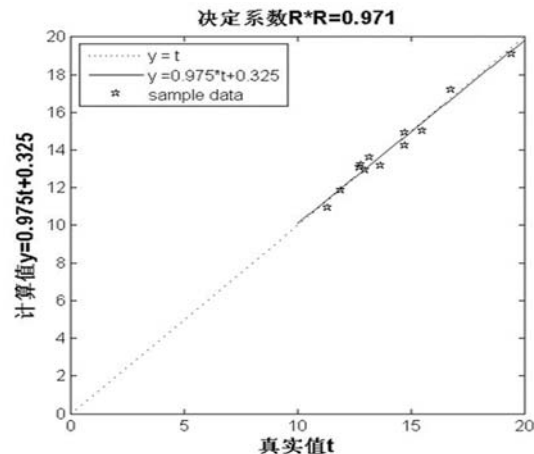
用样本 1-1~1-12 训练 BP 神经网络, 得到隐含层结点和输入端输出端之间的键结值, 将模型模拟得到的铬含量的值和检测得到的测定值进行比较, 计算得到 BP 神经网络输出端的因变量铬含量的相对错误率。将样本铬含量的测定值作为自变量, 并将神经网络计算的铬含量计算值作为因变量, 对其进行线性回归, 然后评价测定值和计算值之间的相关性。决定系数 R^2 越大, 拟合优度就越高。

统计训练过的 BP 神经网络的内部精度, 评估模型的精度。评价指标选用因变量铬的计算值相对错误 RE 、决定系数 R^2 和拟合曲线的斜率 $Slope$ ^[18]。 RE 反映模型的稳定性, 决定系数代表了模型的拟合优度, 拟合曲线的斜率则表现测定值和真实值之间的相关性, 通过这三个指标的综合评价来判断 BP 神经网络的适应性。

表 1 BP 神经网络内部精度评价

结点数	6	7	8	9	10
RE%	2.9%	0.9%	2.6%	2.9%	0.1%
R^2	0.973	0.992	0.975	0.971	0.999
Slope	1.013	1.016	0.996	0.975	0.994

从表 1 中可以看到, 五个神经网络模型的因变量铬含量计算值的错误率均小于 3%, 神经网络内部模拟准确, 计算较精准。铬含量的测定值和计算值的决定系数均大于 0.97, 拟合优度良好, 拟合曲线的斜率均接近于 1, 测定值和模型计算值之间的拟合度高。图 5(a) 和 5(b) 依次是结点数为 6 和 7 时的铬含量测定值, 即真实值 t 和计算值 y 的关系、拟合曲线和理想曲线。



(b)

图 5 不同 BP 神经网络的测定值与真实值关系曲线图

5 个不同结点数的三层 BP 神经网络模型的内部精度良好, 测定值 t 与计算值 y 的拟合曲线和理想曲线 $y = t$ 接近, 且离散数据点的分布在拟合曲线附近, 没有较大的偏差。

4.3 外部精度分析

用测试组的样本 2-1~2-6 来检测模型的预测精度, 将提取的测试样本的 15 个光谱特征参数输入到神经网络的输入端, 输出样本中铬含量的预测值。将预测值与测定值进行比对, 比较神经网络结点数为 6、7、8、9 和 10 时预测值和测定值的平均相对误差 RE 、均方根误差 $RMSE$ 和最大相对误差 MAX , 结果如图 6 所示。

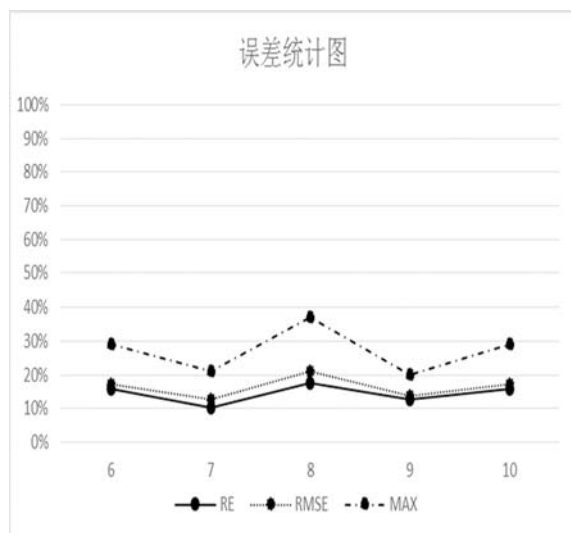
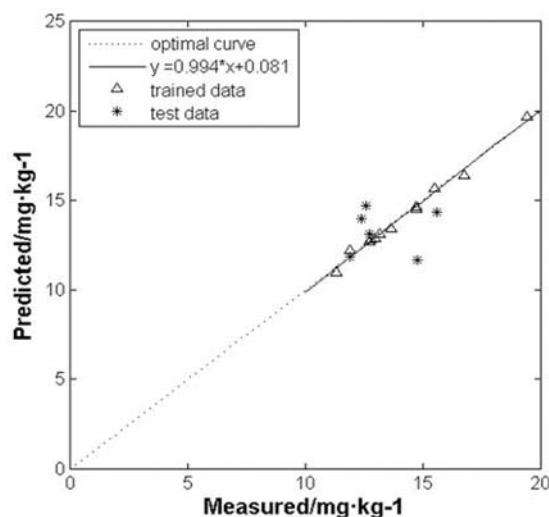
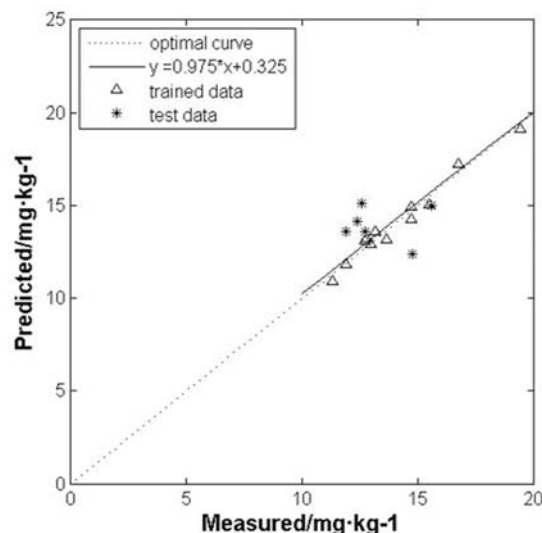


图 6 测试样本的预测误差统计

从误差统计图中可以看到, 三层 BP 神经网络可以预测头发中的铬含量。隐含层有 7 个结点和 9 个结点时, 神经网络表现良好, 最大误差不超过 20%, 标准误差 RMSE 为 13%, 相对平均误差为 10%。图 7 以测试样本的铬含量测定值为自变量, 并以预测值为因变量, 其分布靠近理想曲线。图 7(a) 为结点数为 7 时样本的预测结果, 图 7(b) 为结点数为 9 时样本的测定值和 BP 神经网络预测值的关系。



(a)



(b)

图 7 测试样本和训练样本的真实值与预测值的关系图

应用 BP 神经网络对头发的光谱特征和铬元素含量进行建模, 输入层为三个波段的光谱吸收特征参数, 共 15 个参数, 输出层为铬含量。拟合效果最好的 BP 神经网络的隐含层层数为一层, 输入层和隐含层的激活函数为双曲正切函数, 隐含层和输出层的激活函数为恒等函数。当隐含层的结点数为 7 和 9 时, BP 神经网络的外部精度最高, 均方根误差为 13%, 最大误差不超过 20%, 最低精度在 80% 以上, 整体精度达到 87%。

5 结论

本实验在室内光照几何条件一致的情况下, 采用性状差异明显的人体头发作为实验样本, 在研究其高光谱特征的基础上, 提取光谱曲线的光谱吸收特征参数, 以 BP 神经网络为建模技术方式, 初步确立了人体头发中的铬含量的定量预测模型。实验结果表明, 在 1380 nm~1550 nm、1880 nm~2100 nm 和 2120 nm~2210 nm 三个波段, 人体头发的透射率光谱与头发中的铬含量明显相关, 这三个波段是人体头发中铬元素的敏感波段。实验结果还表明, 用光谱吸收特征参数作为 BP 神经网络输入端, 可以得到拟合效

果较好的预测模型,其光谱吸收特征参数的提取来自三个吸收峰,吸收位置分别为 1500 nm、1930 nm 和 2170 nm。从本次实验结果中分析得到,使用一个隐含层的神经网络,结点数设置为 7 或 9,拟合效果良好,神经网络模型内部的均方根误差 RE 低至 1%,精度可达 99%。运用训练过的神经网络预测样本铬含量时,最大误差不超过 20%,整体误差 RE 为 13%,所有样本精度均高于 80%,整体精度达到 87%。

研究中,基于高光谱曲线吸收特征参数和 BP 神经网络模型,探究了人体头发中重金属元素铬含量的预测方法。本次实验采用近距离定量遥感分析方法,得到了较可靠的实验结果。铬是人体不可或缺的微量元素,本研究能为快速无损地检测人体头发中的铬含量提供理论依据和技术支持,对医疗设备信息化和健康监测实时化等应用有积极的参考价值。

参考文献

- [1] 时圣刚. 重金属对环境与人体健康影响浅议 [J]. **安徽农业科学**, 2013(14):64256426.
- [2] 牛芸民, 杨天林. 若干重要微量金属元素的生物化学功能及其与人体健康的关系 [J]. **微量元素与健康研究**, 2014, 31(2):78-80.
- [3] 吴茂江. 铬与人体健康 [J]. **微量元素与健康研究**, 2014, 31(4):72-73.
- [4] 齐文信, 刘俊辉, 刘巍等. 原子吸收检测儿童青少年头发微量元素分析 [J]. **吉林医学**, 2013(8):1487-1488.
- [5] 王立立, 李永才, 曾庆方等. 电感耦合等离子体发射光谱法同时测定健康人头发中 10 种微量元素的含量 [J]. **药学服务与研究**, 2010, 10(4):305-306.
- [6] 胡雁, 杨俊, 谷勋刚, 等. 反相离子对高效液相色谱法同时测定头发中的尼古丁和可天宁 [J]. **分析测试学报**, 2006, 25(2):77-80.
- [7] 孙梅, 陈兴海, 张恒, 等. 高光谱成像技术的苹果品质无损检测 [J]. **红外与激光工程**, 2014(4):1272-1277.
- [8] 樊磊, 赵文吉, 宫兆宁, 等. 基于包络线消除法的岩石光谱对应分析 [J]. **吉林大学学报(地球科学版)**, 2012, 42(2):575-582.
- [9] 陈文霞, 陈安升, 蔡之华, 等. 基于高光谱吸收特征参数的分类研究 [J]. **计算机工程与应用**, 2008, 44(28):230-232, 241.
- [10] 徐明星, 吴绍华, 周生路, 等. 重金属含量的高光谱建模反演: 考古土壤中的应用 [J]. **红外与毫米波学报**, 2011, 30(2):109-114.
- [11] 吴明珠, 李小梅, 沙晋明, 等. 亚热带土壤铬元素的高光谱响应和反演模型 [J]. **光谱学与光谱分析**, 2014(6):1660-1666.
- [12] 蔡辉, 李娜, 赵慧洁, 等. 基于本征模函数的高光谱数据特征提取方法 [J]. **红外与激光工程**, 2013(12):3475-3480.
- [13] 童庆禧, 张兵, 郑芬兰. **高光谱遥感**[M]. 北京: 高等教育出版社: 2006.
- [14] 刘向阳. BP 神经网络建模与应用 [J]. **科技信息(学术版)**, 2007(16):163-165.
- [15] 张庆庆, 贺兴时. BP 神经网络结点数选取的改进方法及其应用 [J]. **西安工程大学学报**, 2008, 22(4):502-505.
- [16] 张宝菊, 雷晴, 李刚, 等. 基于 BP 神经网络的人体血液中红细胞浓度无创检测 [J]. **光谱学与光谱分析**, 2012, 32(9):2508-2511.
- [17] 王平, 刘湘南, 黄方, 等. 受污染胁迫玉米叶绿素含量微小变化的高光谱反演模型 [J]. **光谱学与光谱分析**, 2010, 30(1):197-201.
- [18] 朱西存, 赵庚星, 王凌, 等. 基于高光谱的苹果花氮素含量预测模型研究 [J]. **光谱学与光谱分析**, 2010, 30(2):416-420.