

文章编号: 1001-3806(2016)06-0848-05

一种基于吸收峰特征的高光谱曲线匹配方法

石俊峰¹, 郭宝峰^{1*}, 沈宏海², 杨名字²

(1. 杭州电子科技大学 自动化学院, 杭州 310018; 2. 中国科学院 长春光学精密机械与物理研究所 航测部, 长春 130033)

摘要: 为了解决在应用传统高光谱地物识别方法时, 由于吸收峰个数不同, 造成的光谱匹配误差较大的问题, 采用了一种基于高光谱吸收峰特征的选择方法, 根据选择后的吸收峰特征进行光谱曲线匹配。该算法首先对高光谱曲线进行包络线消除并提取光谱特征参量矩阵, 然后根据标准特征参量矩阵与待测特征参量矩阵的每个向量的余弦距离-欧氏距离来逐一寻找吸收峰的匹配向量, 之后根据选择的吸收峰特征参量矩阵进行了理论分析和实验验证。结果表明, 该算法可以搜寻到最佳的特征参量向量, 从而实现吸收峰的选择, 用选择后的吸收峰的特征参量矩阵进行高光谱匹配。这一结果对降低匹配的误差是有帮助的。

关键词: 光谱学; 光谱匹配; 吸收峰选择; 向量距离

中图分类号: O433; P237

文献标志码: A

doi:10.7510/jgjs.issn.1001-3806.2016.06.016

Hyperspectral matching method based on absorption features

SHI Junfeng¹, GUO Baofeng¹, SHEN Honghai², YANG Mingyu²

(1. School of Automation, Hangzhou Dianzi University, Hangzhou 310018, China; 2. Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China)

Abstract: When adopting traditional hyperspectral ground objects identification method, the error of spectral matching became big because of the difference of absorption peak number. In order to solve the problem, an optional algorithm based on hyperspectral absorption peak characteristics was brought out, and then spectral matching according to the selected absorption was carried out. At beginning, the continuum removal in spectral curve and the extraction of spectral characteristic parameters matrix were made. And then the matching vector of absorption peak was searched gradually according to the cosine distance-Euclidian distance of every vector from the standard characteristic parameter matrix and the to-be-detected characteristic parameter matrix. After theoretical analysis and experimental verification of the selected absorption peak characteristic parameters matrix, the results show that this algorithm can get the best characteristic parameters vector, to realize the selection of absorption peak and make the hyperspectra matching with the selected absorption peak characteristic parameters matrix. The study is helpful for the decrease of the error of spectral matching.

Key words: spectroscopy; spectral matching; absorption peak selection; vector distance

引言

在进行基于高光谱曲线的物质识别时, 往往出现同一种物质, 提取的光谱曲线会不尽相同, 主要表现为吸收峰个数不同^[1]。如果在不经过前期处理而用传统的匹配方法时, 会导致匹配的误差比较大, 甚至出现错误匹配的现象^[2]。在不影响匹配结果准确性的基础上, 如何对光谱曲线吸收峰进行选择, 寻找匹配度最高的吸收峰, 是解决这个问题的一个思路。

吸收峰是物质本质特性的表现。总体来说, 不同的物质, 其吸收峰个数不一样, 吸收峰的位置、深度、宽度、对称度、面积等光谱吸收指数 (spectral absorption index, SAI) 都会呈现不同的数值^[3]。因此, 基于高光谱曲线的物质识别可以首先提取吸收峰的特征参量, 再根据这些特征参量来进行光谱的吸收峰选择^[4], 然后根据选择后的光谱特征进行匹配即光谱特征参量匹配 (characteristic parameter matrix matching, CPMM)。

作者提出了一种基于向量最小距离的矩阵选择方法, 以基于向量的夹角余弦距离-欧氏距离双重指数作为判定标准^[5]; 把维数低的矩阵作为基准逐一寻找目标矩阵每个向量的最佳匹配向量, 直到维数少的矩阵的所有向量被提取出来, 找到每个向量最匹配的向量为止。在吸收峰选择以后, 再根据光谱特征参量矩阵匹配方法进行光谱匹配。最后通过与常用的光谱角匹

基金项目: 国家自然科学基金资助项目 (61375011); 浙江省自然科学基金资助项目 (LY13F030015)

作者简介: 石俊峰 (1988-), 男, 硕士研究生, 现从事高光谱地物识别的研究。

* 通讯联系人。E-mail: gbf@hdu.edu.cn

收稿日期: 2015-10-29; 收到修改稿日期: 2015-12-16

配法和最小距离匹配法进行比较,验证本文中所提选择方法和匹配方法的有效性。

1 基于光谱特征参量的光谱匹配

在所介绍的方法中,首先是对光谱反射曲线进行包络线消除^[6]。从高光谱数据中提取待识别物质的某个像素光谱曲线后,进行归一化和包络线消除。由于原始信号存在噪声现象比较明显,会造成光谱反射曲线波动,对光谱特征参量的提取以及后续的认识造成不利影响,因此需对其进行降噪处理,在此用小波降噪方法。

高光谱遥感的物光谱曲线反映了地物的吸收和反射特征。不同的地物具有相异的光谱吸收特征。因此,在高光谱遥感识别各种地物成分和成图研究中,其重要任务就包括从许多光谱吸收特征参量(spectral absorption feature parameter, SAFF)中提取各种地物的定性、定量信息。常用的描述光谱吸收特征的参量包括吸收波峰位置 P 、吸收深度 H 、吸收宽度 W 、面积 R 、吸收对称度 S 及变化率 V 等^[7]。

根据上述所求的光谱特征参量,具体算法步骤如下:(1) 从实验光谱库中提取某物质的光谱反射曲线,进行包络线去除和归一化处理^[8]; (2) 提取待测光谱反射曲线的特征参量组成光谱特征参量矩阵 $M = [P, H, W, R, S, V]$, M 的维数是光谱吸收峰的个数; (3) 提取标准光谱库的光谱特征参量矩阵 N , 计算待测物质光谱特征参量矩阵 M 和标准物质矩阵 N 的匹配程度,可以用矩阵的相似度计算矩阵的匹配。

矩阵内积反映了两个矩阵的夹角,表征了两个矩阵相似程度。设 $C_{m \times n}$ 表示 $m \times n$ 矩阵全体,若 $A, B \in C_{m \times n}$, 定义矩阵内积为: $\langle A, B \rangle = \text{tr}(B^T A)$, 由此内积导出范数 $\| \cdot \|$ 为:

$$\| A \| = \langle A, A \rangle^{1/2} \quad (1)$$

式中, $\text{tr}(\cdot)$ 表示矩阵主对角线元素之和。

因 A, B 为实数矩阵,则符合 Cauchy-Schwartz 不等式^[9],即下式:

$$|\langle A, B \rangle| \leq \| A \| \cdot \| B \| \quad (2)$$

当且仅当 A 与 B 完全线性相关,等式 $|\langle A, B \rangle| = \| A \| \cdot \| B \|$ 成立。定义下式:

$$\cos \theta = \frac{\langle A, B \rangle}{\| A \| \cdot \| B \|} \quad (3)$$

式中, θ 定义为两个矩阵的夹角, $\cos \theta$ 可以作为衡量两个矩阵 A, B 相似性依据,其值域为 $[-1, 1]$ 。若设 $r = \cos \theta$, $\theta = 90^\circ$ 时, $r = 0$, 两个矩阵没有相关性; 当 $\theta = 0^\circ$ 时, $r = 1$, 此时两个矩阵相似性最好^[10]。

2 对光谱特征参量矩阵的选择算法

由于欧氏距离能够体现个体数值特征的绝对差异,而夹角余弦距离更多是从方向上区分差异,而对绝对的数值不敏感。因此采用欧氏距离 P_e 和夹角余弦距离的结合,可以提高准确性。设特征参量矩阵的向量 a, b 的联合距离公式如下所示^[5]:

$$D_{ab} = \left(1 - \frac{m_a \cdot n_b}{|m_a| \cdot |n_b|} \right) \cdot P_e(m_a \cdot n_b) \quad (4)$$

为了解决吸收峰个数不一样而导致的特征参量矩阵匹配时维数不等的问题,综合欧氏距离与向量夹角余弦在判定向量的距离时的优缺点,本文中提出了一种新的吸收峰选择方法。根据向量之间的夹角余弦距离-欧氏距离的双重指数即作为相似度判定标准和吸收峰特征参量和,设计了一种特征参量选择方法,算法流程图图 1 所示。

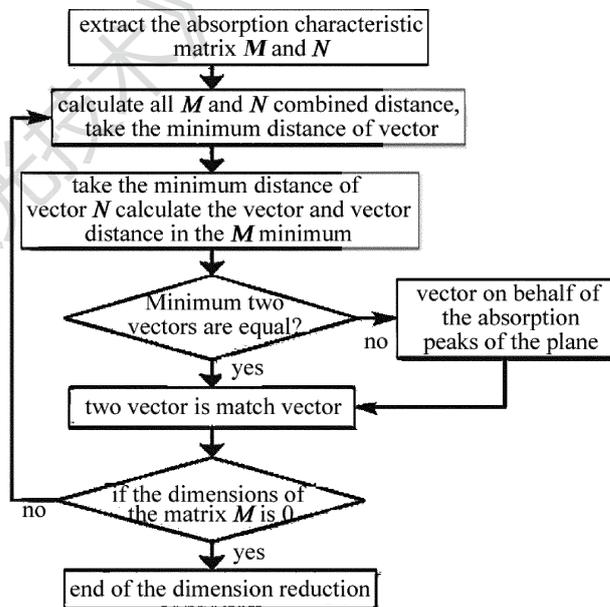


Fig. 1 Flow chart of selection algorithm

该方法的具体实现步骤如下。

(1) 设标准光谱特征参量矩阵 $M = [m_1, m_2, \dots, m_i]^T$, 待测光谱特征参量矩阵 $N = [n_1, n_2, \dots, n_j]^T$, i 为标准光谱吸收峰个数, j 为待测光谱吸收峰个数。吸收峰位置表明了该种物质在特定波长的吸收情况,是最能表征光谱的特征,将其当作匹配的重要参量。

(2) 取矩阵维数小的矩阵的第 1 行向量(假设 M 矩阵的维数小) m_1 , 分别计算 m_1 与待测光谱特征矩阵 N 的向量 n_1, n_2, \dots, n_j 的夹角余弦和欧氏距离的联合距离 $D_{11}, D_{12}, \dots, D_{1j}$; 向量联合距离最小值为 D_{1k} , 即 m_1 和 n_k 的距离最接近。

$$D_{1k} = \left(1 - \frac{\mathbf{m}_1 \cdot \mathbf{n}_k}{|\mathbf{m}_1| \cdot |\mathbf{n}_k|} \right) \cdot P_e(\mathbf{m}_1 \cdot \mathbf{n}_k) \quad (5)$$

(3) 由步骤(2)求出的第 k 个向量作为与 N 中最匹配的向量, 现计算 \mathbf{n}_k 与标准光谱特征矩阵 M 的向量 $\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_i$ 的向量联合距离, 分别记为 $D_{k1}, D_{k2}, \dots, D_{ki}$, 取其中的最小值的 D_{kh} , 即 \mathbf{n}_k 和 \mathbf{m}_h 的距离最小。

$$D_{kh} = \left(1 - \frac{\mathbf{m}_h \cdot \mathbf{n}_k}{|\mathbf{m}_h| \cdot |\mathbf{n}_k|} \right) \cdot P_e(\mathbf{m}_h \cdot \mathbf{n}_k) \quad (6)$$

(4) 若 $h=1$, 即联合距离为 D_{k1} , 由于 $D_{k1} = D_{1k}$; 所以与 \mathbf{n}_k 距离最小的向量为 \mathbf{m}_1 ; 若 $h \neq 1$, 由吸收位置为第一权重特征参量分析^[11], 下一步比较 \mathbf{m}_1 和 \mathbf{m}_h 的吸收位置与 \mathbf{n}_k 的距离 $|d_{m_1} - d_{n_k}|$ 和 $|d_{m_h} - d_{n_k}|$ 的大小。把与 \mathbf{n}_k 的距离最接近的吸收峰为最匹配的吸收峰。

$$d_1 = \{ |d_{m_1} - d_{n_k}|, |d_{m_h} - d_{n_k}| \}_{\min} \quad (7)$$

(5) 取出 M 和 N 中的已匹配的向量, M 和 N 剩余的矩阵为:

当 \mathbf{m}_1 和 \mathbf{n}_k 匹配时,

$$\begin{cases} M_1 = (\mathbf{m}_2, \mathbf{m}_3, \dots, \mathbf{m}_i) \\ N_1 = (\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_{k-1}, \mathbf{n}_{k+1}, \dots, \mathbf{n}_j) \end{cases} \quad (8)$$

当 \mathbf{m}_h 和 \mathbf{n}_k 匹配时,

$$\begin{cases} M_1 = (\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_{h-1}, \mathbf{m}_{h+1}, \dots, \mathbf{m}_i) \\ N_1 = (\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_{k-1}, \mathbf{n}_{k+1}, \dots, \mathbf{n}_j) \end{cases} \quad (9)$$

(6) 重复步骤(2)~步骤(4), 直到维数小的矩阵的所有向量在维数大的向量中得到最匹配的向量。按照匹配的顺序组成新的矩阵则为降维后的矩阵 M, N 。矩阵维数为 δ :

$$\delta = (i, j)_{\min} \quad (10)$$

3 实验与分析

3.1 实验数据

实验中所用数据是拍摄于 1995 年 10 月的 Urban 数据^[12], 图像大小为 307×307 , 包含 210 个波段, 其第 60 个波段的灰度图如图 2 所示。选取图像中的沥青、屋顶、草地和树木 4 种物质进行实验。



Fig. 2 Picture of the 60th bands

3.2 光谱曲线预处理与特征参量提取

实验之前, 去除了低信噪比和水蒸气吸收的波段 (1~4, 76, 87, 101~111, 136~153 和 198~210), 剩下

162 个波段用于本实验, 采用了经过去除噪声和改善图像视觉效果^[9] 的 Urban 数据。在进行光谱匹配之前, 需要提取标准光谱曲线以备下述的光谱匹配。现提取土壤的高光谱曲线, 每种物质提取出其中的 100 条光谱曲线, 对这 100 条的光谱曲线求均值, 所求得的均值作为光谱匹配时的标准光谱曲线。随后再每种物质提取 100 条光谱曲线作为待测光谱曲线, 用来计算与待测光谱的匹配度。在得到土壤的标准光谱曲线后, 为了进行下一步分析的方便, 需要对所有的光谱曲线进行预处理, 预处理包括滤波、包络线消除、归一化处理^[13]。在此首先进行光谱曲线的小波平滑滤波处理。在平滑滤波后, 将标准光谱曲线和待测光谱曲线

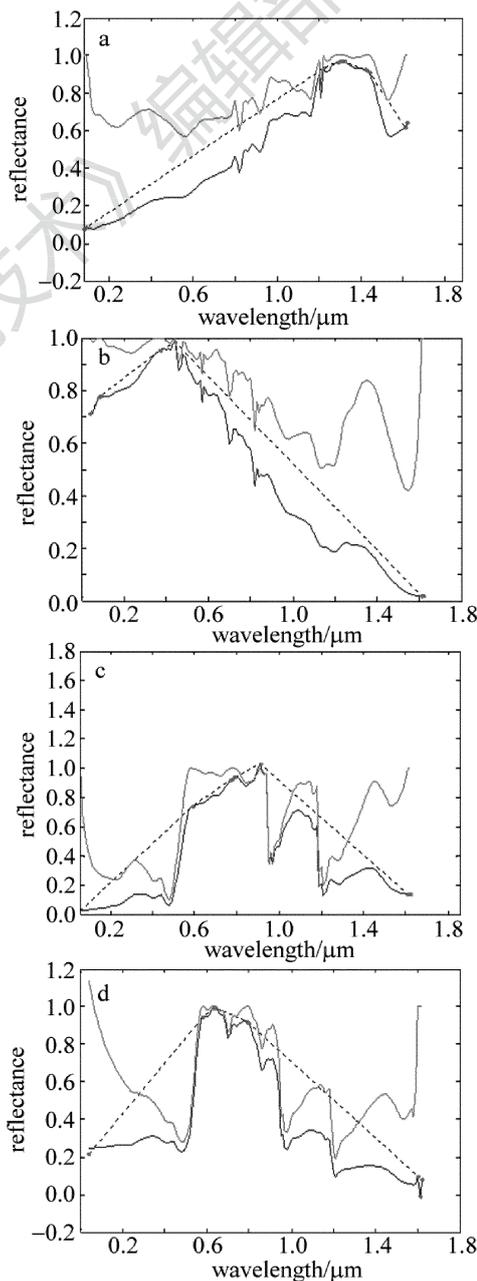


Fig. 3 Standard spectral curve of four materials and normalization processing of envelope eliminate

进行包络线消除处理和归一化处理,其余 3 种物质也做同样的处理。图 3 为待测的沥青、屋顶、草地和树木 4 种物质提取的标准光谱曲线以及进行包络线消除后的结果。

3.3 吸收峰的选择

首先根据上面提出来的包络线消除吸收峰选择方法对吸收峰进行选择。将光谱曲线每种物质提取 100 条曲线作为待测样本。表 1 和表 2 分别为降维前后 4 种物质、100 条光谱曲线中统计的吸收峰个数不同的曲线个数,其中包括光谱曲线条数 (sample size, SS), 吸收峰不同的曲线条数 (different sample size, DSS), 吸收峰不同曲线百分比 (different samples proportion, DSP)。由表 1 和表 2 可以看出,在经过选择算法后,特征参量矩阵的吸收峰个数相同,吸收峰个数不同的光谱曲线数目分别有了明显的下降,4 种物质中的吸收峰不同的光谱曲线所占的比率从 28%,42%,37%,23% 下降为 2%,5%,4%,4%。吸收峰选择算法效果得到验证。

Table 1 Spectral curve number of different absorption peaks before reducing dimensionality

	pitch	housetop	lawn	tree
SS	100	100	100	100
DSS	28	42	37	23
DSP	28%	42%	37%	23%

Table 2 Spectral curve number of different absorption peaks after reducing dimensionality

	pitch	housetop	lawn	tree
SS	100	100	100	100
DSS	2	5	4	4
DSP	2%	5%	4%	4%

3.4 根据选择后的吸收峰进行光谱匹配

在对光谱曲线吸收峰进行选择以后,需要根据选择后的特征参量矩阵进行光谱匹配。根据上面所提出的光谱匹配方法,计算出每条待测曲线与标准光谱曲线的匹配度。现选择常用的光谱匹配方法,即光谱角匹配 (spectral angle match, SAM)、最小距离匹配法 (minimum distance matching, MDM)^[14] 和 CPMM。SAM 采用角度相似系数 (夹角余弦) 描述光谱间的匹配度,是基于光谱度量匹配的高光谱分类方法中最具有代表的一种匹配方法。由于特征参量匹配方法是把光谱曲线的特征参量提取出来组成矩阵形式,求列向量的夹角余弦作为匹配的参量,再对多条光谱曲线的匹配度求均值。由于都是对光谱角余弦的比较,可以作为对比吸收峰选择前后光谱匹配的一种标准。

在用于对比的经典方法中,最小距离 (minmum

distance, MD) 匹配法采取计算标准光谱与待测光谱之间的距离,距离越小,光谱相似性越大;本文中采用常用的欧氏距离 (Euclidian distance, ED)、马氏距离^[15]。另外一种是用光谱角匹配法计算光谱角余弦值大小。现取 100 条经过初步处理过的待测光谱,分别使用上述两种方法和本文中提出的方法进行光谱匹配计算;其中降维曲线匹配度 (dimension reduction, DR)、未降维的曲线匹配度 (no dimension reduction, NDR)、所有曲线匹配度 (all spectral curve, ASC) 的匹配结果如表 3 所示。

Table 3 Matching results of four kinds of material

		ED	MD	SAM	CPMM
pitch	DR	0.1402	0.3153	0.8145	0.9142
	NDR	0.1900	0.4013	0.7408	0.8201
	ASC	0.1735	0.3871	0.7717	0.7900
housetop	DR	0.0942	0.1641	0.7487	0.8592
	NDR	0.1632	0.2284	0.6503	0.7373
	ASC	0.1345	0.1931	0.6947	0.8123
lawn	DR	0.2312	0.1464	0.8329	0.9423
	NDR	0.3053	0.2358	0.7321	0.8256
	ASC	0.2740	0.2074	0.7634	0.9135
tree	DR	0.0532	0.2252	0.8124	0.9012
	NDR	0.1354	0.3421	0.7031	0.7343
	ASC	0.1026	0.3089	0.7412	0.8910

在和常用的光谱角匹配法和最小距离法进行比较之后发现,未经过选择的吸收峰的光谱曲线在用传统的光谱角匹配时的误差较大问题,通过本文中提出来的吸收峰选择方法,计算的光谱匹配度有所改观。在不同的匹配方法的横向比较时,如表 4 所示,4 种物质的光谱曲线在两种不同的匹配方法 SAM, MDM 和 CPMM 时,匹配结果之差要比未选择的匹配方法之差有所增大,也就是说,选择后拉大了光谱曲线之间匹配度的差。在原始光谱不变的情况下,说明选择后的匹配度有所增强。

Table 4 Analysis results of match degree difference between different methods

	pitch	housetop	lawn	tree
DR	0.0997	0.1105	0.1094	0.0888
NDR	0.0793	0.0870	0.0935	0.0312

比较相同匹配方法下的 4 种物质的匹配度,用 SAM 算法,MDM 算法和用 CPMM 算法的时候计算所得的匹配度,来分析降维前后的匹配度之差,如表 5 所示。

从表 5 分析得出,吸收峰选择后得到的矩阵在后续进行光谱匹配时,相同的匹配方法经过选择算法处理的特征参量矩阵比未进行选择的匹配度之差要大。

Table 5 Analysis results of Euclidean distance and Mahalanobis distance and of the same matching method

	pitch	housetop	lawn	tree
ED	0.0498	0.0690	0.0741	0.0822
MD	0.0860	0.0643	0.0894	0.1168
SAM	0.0737	0.0984	0.1008	0.1093
CPMM	0.0941	0.1219	0.1167	0.1669

而且在用欧式距离和马式距离进行分析时,经过选择的光谱曲线,其马式距离和欧式距离普遍要小于没有进行选择的。分析原因是因为在吸收峰的选择过程中,标准光谱的每个列向量,也就是每个吸收峰的特征参量向量,总是与待测光谱中的匹配度最高的向量进行匹配,每个与标准光谱的吸收峰特征向量都是对应的,都是与待测光谱中最匹配的吸收峰,把匹配度小的吸收峰在这一步直接排除掉,从而达到选择的目的。综上所述,该方法不仅达到了选择效果,而且提高了光谱的匹配效果。

4 结 论

主要介绍了对光谱特征参量矩阵进行选择的方法,以及根据选择后特征参量矩阵的光谱匹配方法。由于同一种物质的光谱曲线不同,造成的吸收峰个数不同,从而导致匹配的误差比较大。本文中根据每个吸收峰的特征参量向量求取与之距离最小的特征参量向量,并结合吸收位置是吸收峰最重要的特性,选择向量间联合距离最小的特征参量向量,从而达到吸收峰选择的目的,并根据选择后的特征参量矩阵进行光谱匹配。

用高光谱数据对本文中提出的光谱曲线特征参量矩阵选择方法进行仿真实验。实验结果表明,本文中的吸收峰的选择算法,不论待测光谱吸收峰个数大于或者小于标准光谱的吸收峰个数,都能达到选择的目的。并且在对选择后的特征参量矩阵进行光谱匹配的时候,由于本算法在对吸收峰选择的同时,把匹配度最低的吸收峰排除掉,在与传统的光谱角匹配方法结果进行横向和纵向比较时,可以看出,光谱匹配度都有一定程度的提升。

参 考 文 献

[1] TONG Q X, ZHANG B, ZHENG L F. Hyperspectral remote sensing

[M]. Beijing: Higher Education Press, 2006: 59-67 (in Chinese).

[2] LIAO W, PIZURICA A, PHILIPS W, *et al.* Semisupervised local discriminant analysis for feature extraction in hyperspectral images [J]. IEEE Transactions on Geo science and Remote Sensing, 2013, 51(1): 184-198.

[3] ZHANG Ch, CAI H J, LI Zh J. Hyperspectral characteristic parameter of winter wheat photosynthetic active radiation component estimation model [J]. Spectroscopy and Spectral Analysis, 2013, 51(1): 184-198 (in Chinese).

[4] ZHANG Y, ZHANG J L, ZHAO X Sh, *et al.* Extraction of mineral alteration information from core hyperspectral images based on weight of absorption peak [J]. Remote Sensing for Land and Resources, 2015, 27(2): 154-159 (in Chinese).

[5] SUN Y L, ZHANG X, SHUAI T, *et al.* 2015 Radiometric normalization of hyperspectral satellite images with spectral angle distance and Euclidean distance [J]. Journal of Remote Sensing, 2015, 19(4): 618-626 (in Chinese).

[6] FAN L, ZHAO W J, GONG Zh N, *et al.* Correspondence analysis of rock spectra based on continuum removing [J]. Journal of Jilin University (Earth Science Edition), 2012, 42(2): 575-582 (in Chinese).

[7] ZHANG J, SHEN Y T, WANG X J. Hyperspectral quantitative models for chlorophyll-a of algae based on spectral absorption feature parameters and spectral absorption index [J]. Journal of Agro Environment Science, 2011, 30(8): 1622-1629 (in Chinese).

[8] LI H J, XU Sh Y, YAN D J. Research of remote sensing image matching with sub-pixel accuracy [J]. Laser Technology, 2008, 31(5): 493-495 (in Chinese).

[9] SHU X W, ZHANG Y J, GENG H, *et al.* Study on phase shift anto-balanced laser absorption spectroscopy [J]. Laser Technology, 2011, 35(5): 618-621 (in Chinese).

[10] PADMA S, SANJEEVI S. Jeffries Matusita based mixed-measure for improved spectral matching in hyperspectral image analysis [J]. International Journal of Applied Earth Observations and Geoinformation, 2014, 32(4): 138-151.

[11] GAN R T, GUO Zh N, LIN J B. Spectral matching technology for light-emitting diode-based jaundice photodynamic therapy device [J]. Journal of Modern Optics, 2015, 62(3): 212-217.

[12] WU X W, GUO B F, CHEN Ch Zh, *et al.* RX anomaly detection based a weighted combination kernel and its parameters selection [J]. Laser Technology, 2015, 39(6): 745-751 (in Chinese).

[13] WANG K, QU H M. Anomaly detection method based on improved minimum noise fraction transformation [J]. Laser Technology, 2015, 39(3): 381-385 (in Chinese).

[14] HUANG T T, WEI Zh H, XIU L C, *et al.* The research of mineral spectral matching based on weighted absorption peaks [J]. Rock and Mineral Analysis, 2011, 30(5): 584-589 (in Chinese).

[15] LIN H J, ZHANG H F, GAO Y Q, *et al.* Mahalanobis distance based hyperspectral characteristic discrimination of leaves of different desert tree species [J]. Spectroscopy and Spectral Analysis, 2014, 34(12): 3358-3362 (in Chinese).