

红外光谱相似谱及其在中药鉴别中的应用

徐永群¹, 陈小康^{1*}, 陈勇², 黄冬兰¹, 陈灶鑫², 卢文贯¹

1. 韶关学院化学系, 广东 韶关 512005

2. 广东工业大学轻工化工学院, 广东 广州 510006

摘要 红外光谱相似谱能描述两红外光谱相似的程度, 在波数 $\tilde{\nu}_i$ 处相关谱的值 r_i 能反映两红外光谱在 $[\tilde{\nu}_i - n, \tilde{\nu}_i + n]$ 区间内线性相关程度的大小, 值在 $-1 \sim 1$ 之间, $r_i > 0$ 意味着两光谱正相关, $r_i < 0$ 意味着负相关, $r_i = 0$ 则没有线性关系。在偏移量 n 一定的条件下, 依相似谱形状的差异可辅助红外光谱的解析, 用相似谱或相似谱提供的信息提取的光谱特征值可用于模式识别。计算了18个黄芪、红兰芪和党参样本在 $1\ 600 \sim 700\ \text{cm}^{-1}$ 范围内的红外光谱相似谱, 结果表明各类药材均有各自的特征相似谱, 凭借这些可实现对它们的鉴别; 依灵芝红外光谱相似谱提供的信息, 计算了 $1\ 560 \sim 1\ 502$, $1\ 460 \sim 1\ 421$ 和 $1\ 319 \sim 1\ 260\ \text{cm}^{-1}$ 三个波数区间内93个灵芝样本红外光谱的相关系数, 在无先验知识的情况下, K-means聚类分析法成功地将其分为了赤芝和紫芝类、黑芝类、青芝类和黄芝类, 这一分类结果与形态分类法得到的结果基本一致。

关键词 红外光谱相似谱; 谱图解析; 中药鉴别; 聚类分析

中图分类号: O657.3 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2012)08-2131-04

引言

红外光谱与分子结构密切相关, 是研究和表征分子结构的一种手段, 常用于物质的鉴别。但大分子或混合物体系的谱图是比较复杂的, 谱峰难以指认, 解谱困难。近十多年来, 人们一直在探索智能化的解谱方法, 建立了各种检索系统、专家系统和模式识别系统, 这些对某些特定问题的解决取得了较好的成效, 但均有一定的局限性。为了进一步推动计算机辅助解谱技术的发展, 寻求新的解谱思路仍是一项既有意义又有价值的工作^[1,2]。中药材是一个组分极其复杂的混合物系统, 谱图具有指纹性, 但复杂, 难以由人工直接解析^[3]。作者设计了计算红外光谱相似谱的软件, 该软件能提供两红外光谱差异最为突出的波数范围或位置的信息, 这些信息可辅助谱图解析, 可实现对差异较小光谱的识别, 为谱图的解析提供了一种新的分析思路和新的分析方法, 将其与模式识别技术结合起来, 可推动计算机智能化解谱技术进一步发展。

1 实验部分

1.1 仪器设备和参数设置

Thermo Nicolet380 傅里叶红外光谱仪(Thermo Electron

公司), 测定范围 $4\ 000 \sim 400\ \text{cm}^{-1}$, 扫描次数16次, 分辨率 $2\ \text{cm}^{-1}$ 。

1.2 样品来源和处理

样品来源: 灵芝由韶关天仁堂野生灵芝店收集, 韶关学院英东生命科学学院鉴定; 黄芪、党参和红兰芪均由丽珠集团利民制药厂采集并鉴定。

样品处理: 原始药材(灵芝只取菌盖)在 $65\ ^\circ\text{C}$ 下烘4h, 粉碎, 过200目筛。取约3mg灵芝粉末和约250mg KBr粉末一起研磨均匀, 压片测定红外光谱。

1.3 软件

基于 Visual Basic 6.0 自设计了中药材红外光谱相似谱分析软件。基于 Matlab 7.0 设计了 k-means 三维等责分析程序。

2 红外光谱相似谱

2.1 相似谱的计算

设 A 和 B 两物质的中红外光谱分别为 $A(x_{4\ 000}, x_{3\ 999}, \dots, x_i, \dots, x_{400})$, $B(y_{4\ 000}, y_{3\ 999}, \dots, y_i, \dots, y_{400})$, 则二者的相似谱 R 为 $R(r_{4\ 000}, r_{3\ 999}, \dots, r_i, \dots, r_{400})$, 其中波数 i 处的值 r_i 可用相关系数计算公式(1)式计算, 该值能反映两光谱强度在 $[\tilde{\nu}_i - n, \tilde{\nu}_i + n]$ 波数区间内线性相关程度的大小,

收稿日期: 2012-01-06, 修订日期: 2012-04-30

基金项目: 国家自然科学基金项目(21071099)和广东省科技计划项目(2006B35630010)资助

作者简介: 徐永群, 1956年生, 韶关学院化学系教授 e-mail: xyqun@sina.com

* 通讯联系人 e-mail: chk@sgu.edu.cn

相关程度大, 谱图的相似性就大, 相关程度小, 谱图的相似性就小, 故称 R 为相似谱。

$$r_i = \frac{\sum_{i=\tilde{\nu}_i-n}^{\tilde{\nu}_i+n} (x_i - \bar{x})(y_i - \bar{y})}{\left[\sum_{i=\tilde{\nu}_i-n}^{\tilde{\nu}_i+n} (x_i - \bar{x})^2 \sum_{i=\tilde{\nu}_i-n}^{\tilde{\nu}_i+n} (y_i - \bar{y})^2 \right]^{\frac{1}{2}}} \quad (1)$$

式中: n 为对中心波数 $\tilde{\nu}_i$ 的偏移量; \bar{x} 和 \bar{y} 分别为 A 和 B 两光谱强度在 $[\tilde{\nu}_i - n, \tilde{\nu}_i + n]$ 波数区间内的平均值。

2.2 相似谱的性质

红外光谱相似谱中 r_i 的值在 $-1 \sim 1$ 之间, 它能反映在中心波数 $\tilde{\nu}_i$ 处两光谱相似程度的大小, 如果两光谱强度大小变化方向一致, 则 $r_i > 0$, 两光谱正相关; 如果变化相反, 则 $r_i < 0$, 两光谱负相关; 如果两光谱强度之间没有线性关系, 相关不密切, 则 $r_i = 0$ ^[4]。图 1 中 R1 和 R2 为中药黄芪和红兰芪(黄芪伪品)的红外光谱相似谱图, 其中 R1 的偏移量 $n = 50$, R2 的偏移量 $n = 20$, 从图中可看出, $r \approx 1$ 的位置, 两药材的谱图极其相似, r 较小位置谱图的差异性较大, 如: $2\ 928\ \text{cm}^{-1}$ 处, 两光谱吸收峰的形状极其相似, 故相似谱 R2 在此处的值基本为 1; 在特征波段 $1\ 727 \sim 865\ \text{cm}^{-1}$ 的区间内, 两药材的红外光谱既相似但又有差异, 抓住该特征段红外光谱的分析和对比, 可以实现对黄芪正品和伪品的鉴别。

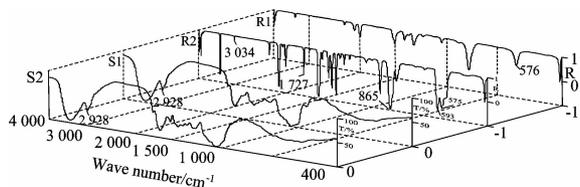


Fig. 1 FTIR spectra and similarity spectra of astragalus and red-blue stragalus

S1, FTIR of stragalus; S2, FTIR of red-blue stragalus; R1, similarity spectrum ($n=50$); R2, similarity spectrum ($n=20$)

从相似谱 R2 可以看出, 相似谱的峰位置 (r 极小处) 可以辅助分析存在差异的位置。相似谱峰的形状、位置和数目与偏移量有关, 偏移量较大时, 统计的波数范围较宽, 峰数较少, 峰值较大, 峰形较钝; 偏移量较小时, 统计的波数范围较窄, 峰数较多, 峰值较小, 峰形尖锐, 偏移量小有利于两红外光谱细节的比对。如: R1 在 $576\ \text{cm}^{-1}$ 处只出现了一个较宽的峰, 而在 R2 中出现了 593 和 $575\ \text{cm}^{-1}$ 两个相对应的峰, 显然, 偏移量小分辨率高。在红外光谱分析中, 可根据实际情况设置一个比较合理的偏移量, 以达到最佳辅助分析目的。

2.3 相似谱的用途

中药和食品等天然产物主要由淀粉和纤维素等物质构成, 由于主体成分相似, 故谱图较为相像, 但又因为它们各自的组成不同, 故又有一定的差异, 其特点是谱峰难以指认, 差异较小的谱图, 人工难以辨认。红外光谱相似谱能突显两红外光谱的差异, 抓住特征波段相似谱谱峰位置、大小和形状的差异, 可辅助谱图的解析。本研究以山西浑源特等黄芪(0 号样)的红外光谱为公共比较对象, 计算了 18 个中药样本在特征波段 $1\ 600 \sim 700\ \text{cm}^{-1}$ 范围内的红外光谱相似谱, 其样本编号及信息见表 1, 相似谱见图 2。从红外光谱图可看出, 黄芪、红兰芪和党参的红外光谱图极其相似, 但又有明显的差异, 党参在 $900 \sim 750\ \text{cm}^{-1}$ 之间有 3 个明显的吸收峰, 黄芪的吸收峰较小, 而红兰芪就更小, 凭借这些可实现对三类天然产物的辨认。但相似谱图能提供更为明显和详细的谱图差异的信息, 对黄芪类药材来说, 山西浑源、天镇和内蒙的相关谱图又有明显的差异, 内蒙黄芪(6—8 号样)在 $1\ 528\ \text{cm}^{-1}$ 处有一个相关系数最小值接近于 -0.2 的宽峰; 山西天镇黄芪(9—12 号样)在 $1\ 538$ 和 $1\ 523\ \text{cm}^{-1}$ 处产生了 2 个连续的峰; 山西浑源黄芪(0—4 号样)与内蒙黄芪的差异小于山西浑源与山西天镇黄芪之间的差异; 红兰芪谱图与黄芪谱图有较大差异, 依此可实现黄芪伪品的鉴别。对党参来说, 在

Table 1 The information of astragalus, red-blue astragalus and codonopsis

名称	编号、产地及等级等信息
黄芪	0(山西浑源, 特等)、1(山西浑源, 一等)、2(山西浑源, 二等)、3(山西浑源, 三等)、4(山西浑源, 四等)、6(内蒙, 大棒子)、7(内蒙, 二棒子)、8(内蒙, 小棒子)、9(山西天镇, 一等)、10(山西天镇, 二等)、11(山西天镇, 三等)、12(山西天镇, 四等)
红兰芪	5(山西浑源, 黄芪伪品)
党参	白条党参: 13(甘肃, 二等)、14(甘肃, 三等)、15(甘肃, 四等) 纹党参: 16(甘肃, 二等)、17(甘肃, 三等)、18(甘肃, 四等)

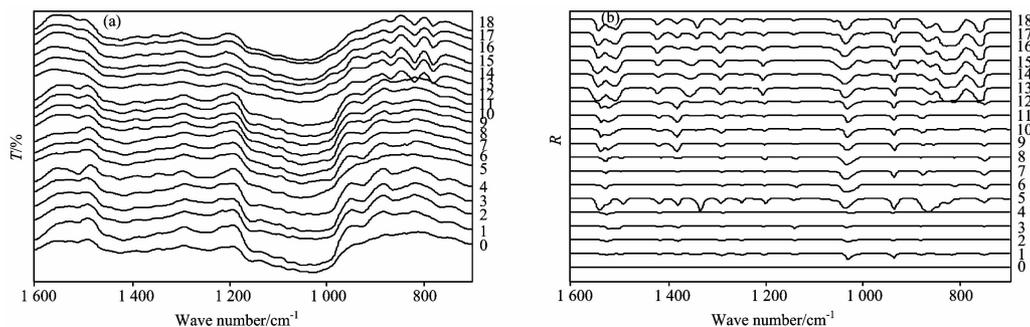


Fig. 2 FTIR spectra and its similarity spectra of astragalus, red-blue astragalus and codonopsis

(a): The infrared spectra; (b): The similarity spectra ($n=20$)

1 422~1 293 cm^{-1} 之间, 白条党参(13—15 号样)只有一个峰, 而纹党参(16—18 号样)有 2 个峰, 凭借这一差异, 可实现对 2 类党参的鉴别。

3 红外光谱相似谱在灵芝聚类分析中的应用

3.1 灵芝红外光谱相似谱

Table 2 The information of ganoderma lucidum

种类	编号、产地及名称
赤芝	1(云南野生赤芝)、2(广西野生赤芝)、3(韶关赤芝)、4(海南喜热灵芝)、5(海南无柄赤芝)、6(韶关泰山灵芝)、7(四川野生赤芝)、8(海南喜热灵芝)
紫芝	9(韶关无柄紫灵芝)、10(韶关紫芝)、11(贵州紫芝)、12(贵州野生紫芝)、13(韶关紫芝)
黑芝	14(海南野生黑芝)、15(韶关皱盖乌芝)、16(越南黑芝)、17(云南黑乌芝)、18(广西黑芝)
青芝	19(韶关珊瑚型平盖灵芝)、20(韶关野生平盖灵芝)、21(韶关平盖灵芝)、22(韶关飞碟型平盖灵芝)、23(韶关平盖灵芝)
黄芝	24(云南斑褐灵芝)、25(云南野生黄芝)、26(广西黄芝)、27(韶关黄芝)、28(海南野生斑褐灵芝)、29(云南野生斑褐灵芝)、30(韶关野生黄芝)、31(韶关本地黄芝)

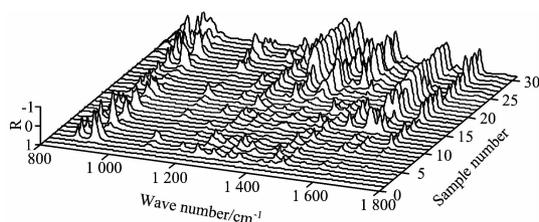


Fig. 3 The similarity spectra of 31 kinds of ganodermas ($n=20$)

3.2 红外光谱相似谱规律分析

图 3 为特征波段 800~1 800 cm^{-1} 区间内的相似谱, 其中出峰位置是相关系数较小的位置, 峰越高, r_i 的值就越小, 即样本与公共比较对象的红外光谱的差异就越大。图中五类灵芝的谱图均有特征性, 类与类之间有明显的差异: 赤芝(1—8 号)的峰较小, 但在 1 500~1 200 cm^{-1} 波数区间内有一系列较小的峰出现; 紫芝(9—13 号)在 1 700~1 450 cm^{-1} 波数区间内有些较小的峰出现; 黑芝(14—18 号)在 1 319~1 260 cm^{-1} 波数区间有较强的峰出现; 青芝(19—23 号)在 1 560~1 502 cm^{-1} 波数区间较强的峰出现; 黄芝(24—31 号)在 1 560~1 502 和 1 460~1 421 cm^{-1} 波数区间均有很强的峰出现。凭借这些差异, 可实现对五类灵芝的鉴别; 依相似谱总体差异的大小, 大致可判断出各类样本之间差异的大小。如: 赤芝与紫芝谱图的差异较小, 说明该二类灵芝的化学组成比较接近。传统中医认为赤芝和紫芝的疗效最佳,

灵芝为灵芝科灵芝属(*Ganoderma Karst*)真菌, 具有较高的药用价值, 我国药典将其分为赤芝、紫芝、黄芝、白芝、黑芝、青芝等六种。本研究收集到了赤芝、紫芝、黄芝、黑芝和青芝等五种灵芝的样品共 31 个, 样品编号、种类及产地等信息见表 2。以 1 号样云南野生赤芝的红外光谱为公共比较对象, 设置偏移量为 20, 计算了在 800~1 800 cm^{-1} 波数区间内各样品的红外光谱相似谱, 其谱图见图 3。

故一般以该二种灵芝入药^[5]。谱图呈现的结果与传统中医的经验基本一致, 可相互印证其中的道理。

3.3 红外光谱相似谱辅助聚类分析

聚类分析是依“物以类聚”的原理对样品进行分类的一种统计分析法, 它能按样品的特性进行合理地归类, 同一类中的对象具有一定的相似性, 不同类间的对象具有一定的差异, 利用归类结果可以实现对物质的分类和鉴别。图 3 在 1 560~1 502, 1 460~1 421 和 1 319~1 260 cm^{-1} 三个区间内, 五类灵芝各有不同的变化规律, 抓住这些, 分别计算了各样本在这三个区间内的红外光谱与 1 号样本之间的相关系数 R_1 , R_2 和 R_3 , 依这三个参数, 用 K-means 聚类分析法对 31 个样品(每个样品重复测定 3 次)测得的 93 个红外光谱的样本进行了三维聚类分析, 聚类结果如图 4。图中将 93 个样本分为了 4 大类, 即: 赤芝和紫芝类、黑芝类、青芝类和黄芝类。这一无先验知识分类的结果将赤芝与紫芝归为了一类, 它与目前占主导地位的形态分类法得到的结果基本一致。赵继鼎依菌盖表面的构造与漆样光泽的有无将灵芝属分为灵芝亚属、粗皮灵芝亚属和树舌亚属, 灵芝亚属由灵芝(赤芝)组和紫芝组构成, 即赤芝和紫芝同属于灵芝亚属^[6-9]。

4 结果与讨论

提出了红外光谱相似谱的概念, 给出了计算方法, 设计了计算程序。红外光谱相似谱能提供两红外光谱差异最为突

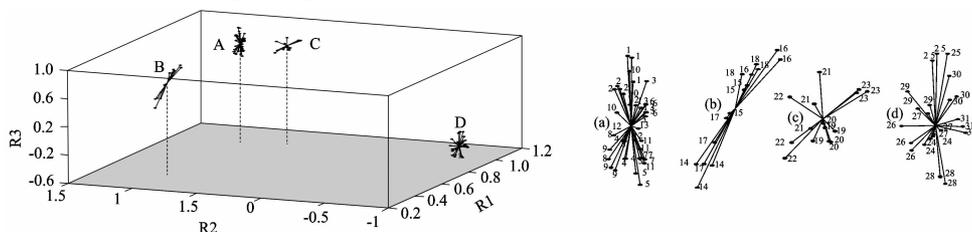


Fig. 4 The K-means 3d clustering analysis results of 93 samples of ganodermas

Left: The 3D figure; Right: the information figure of samples: A: *Ganoderma lucidum* and *Ganoderma sinensis*; B: *Ganoderma atrum*; C: *Ganoderma aoshiba*; D: *Ganoderma multipliatum*

出的波数范围或位置的信息,可比较两红外光谱差异的大小,可辅助红外光谱的解析和模式识别特征信息的提取。

描述了红外光谱相似谱的性质。相似谱的值在 $-1 \sim 1$ 之间,值大处两谱相似程度大,值小处两谱相似程度小。相似谱峰的形状、位置和数目与偏移量 n 的大小有关。依实际

情况,选择合适的偏移量,有助于红外光谱的解析。

不同中药具有不同的特征红外光谱相似谱,相似谱能突显光谱之间的差异,抓住这些差异,可实现对中药的分类和鉴别,利用这些差异,可较好地辅助谱图的解析。

References

- [1] QIN Shen-jun, WANG Jian-bin, YAO Hong-wei, et al(秦身钧,王建滨,姚宏伟,等). Journal of Hebei Normal University(河北师范大学学报), 2006, 30(3): 332.
- [2] CHEN Ling-ran, ZHANG Mao-sen(陈玲然,张懋森). Chemical Journal of Chinese Universities(高等学校化学学报), 1990, 11: 1346.
- [3] XU Yong-qun, HUANG Hao, ZHOU Qun, et al(徐永群,黄昊,周群,等). Chinese Journal of Analytical Chemistry(分析化学), 2003, 31: 6.
- [4] ZUO Zhen-zhao, ZHANG Yan-hong, YUAN Bo(左振钊,张艳红,袁博). Journal of Hebei North University(河北北方学院学报), 2005, 21(5): 11.
- [5] Pharmacopoeia Committee of Ministry of Health, the People's Republic of China(中华人民共和国卫生部药典委员会编). Pharmacopoeia of the People's Republic of China(中华人民共和国药典). Beijing: Chemical Industry Press(北京:化学工业出版社), 2005. 130.
- [6] Ibrahim Sönmez, Ali ümran Kömüştü. Theor. Appl. Climatol, 2011, 106: 500.
- [7] TANG Chuan-hong, SU Chun-li, ZHANG Jin-song, et al(唐传红,苏春丽,张劲松,等). Acta Edulis Fungi(食用菌学报), 2007, 14(3): 87.
- [8] JIA Ding-hong, ZHENG Lin-yong, ZHANG Xiao-ping, et al(贾定洪,郑林用,张小平,等). Southwest China Journal of Agricultural Sciences(西南农业学报), 2003, 16: 15.
- [9] GU Sheng-dong(谷胜东). Ganoderma Lucidum(灵芝). Tianjin: Tianjin Science and Technology Press(天津:天津科学技术出版社), 2005. 20.

Similar Spectrum of Infrared Spectrum and Its Application in Identification of Chinese Herbs

XU Yong-qun¹, CHEN Xiao-kang^{1*}, CHEN Yong², HUANG Dong-lan¹, CHEN Zao-xin², LU Wen-guan¹

1. Department of Chemistry, Shaoguan College, Shaoguan 512005, China

2. Faculty of Chemistry and Light Engineering, Guangdong University of Technology, Guangzhou 510006, China

Abstract In the present paper, the similar spectra of 18 samples, which include *Astragalus*, red-blue *Astragalus* and *Codonopsis*, were obtained in the range of $1\ 600 \sim 700\ \text{cm}^{-1}$. The result showed that all kinds of herbs have their own characteristic similar spectra, and 18 samples can be identified according to the characteristic similar spectra. Furthermore, three correlation coefficients of 93 ganoderma samples were calculated which is in the range of $1\ 560 \sim 1\ 502$, $1\ 460 \sim 1\ 421$ and $1\ 319 \sim 1\ 260\ \text{cm}^{-1}$ according to the information of similar spectrum of infrared spectrum of ganoderma. Without priori knowledge of the classification of these samples, the K-means cluster analysis can successfully divide them into four classes, i. e. *Ganoderma lucidum* and *Ganoderma sinensis*, *Ganoderma atrum*, *Ganoderma aoshiba*, *Ganoderma multiplicatum*. This result is consistent with the result of morphological classification.

Keywords Similar spectrum of infrared spectrum; Spectrum analysis; Identification of Chinese herbs; Cluster analysis

* Corresponding author

(Received Jan. 6, 2012; accepted Apr. 30, 2012)