

基于支持向量机的“内外结合”中药质量控制新模式探索^Δ

陶梦琳^{1*}, 顾文涛¹, 侯珂惠¹, 汪子青¹, 崔书盛¹, 秦娜¹, 兰泽伦², 周霞^{1#}(1.西南交通大学生命科学与工程学院/电气工程学院, 成都 610031; 2.国家中医药管理局中药炮制技术重点实验室, 成都 610036)

中图分类号 R288 文献标志码 A 文章编号 1001-0408(2016)01-0118-04

DOI 10.6039/j.issn.1001-0408.2016.01.38

摘要 目的:探索基于支持向量机(SVM)的“内外结合”中药质量控制新模式。方法:利用仿生嗅觉、视觉及味觉系统采集中药材及其饮片的性状、颜色、气味、味道等外部信息,再利用SVM与中药材中化学成分的含量、水分、灰分、浸出物等内部信息建立关联数据库,建立基于SVM的中药整体质量控制新模式。选用来自四川、重庆、湖北和云南的黄连作为分析对象,分别采集其内外部和产地信息后对该系统进行有效性验证。结果与结论:内部特征、外部特征、“内外结合”全部特征这3种识别方法的识别率均大于90%,且“内外结合”全部特征的识别率较单一内部或外部特征高。选用SVM法对得到的内外数据建立关联和特征模型,可使传统“辨状论质”经验得以现代“察外知内”的再现,从而在传统经验的基础上建立更加规范、准确、客观的中药质量整体控制新模式。

关键词 仿生感官系统;中药;质量控制;支持向量机

Exploration of the “Internal and External Combined” New TCM Quality Control Mode Based on Support Vector Machines

TAO Menglin¹, GU Wentao¹, HOU Kehui¹, WANG Ziqing¹, CUI Shusheng¹, QIN Na¹, LAN Zelun², ZHOU Xia¹(1. School of Life Science & Engineering/School of Electrical Engineering, Southwest Jiaotong University, Chengdu 610031, China; 2. Key Research Laboratory of TCM Processing Technology, State Administration of Traditional Chinese Medicine, Chengdu 610036, China)

ABSTRACT OBJECTIVE: To explore “internal and external combined” new traditional Chinese medicine (TCM) quality control mode based on support vector machines (SVM). METHODS: External data were collected with bionic olfaction, vision and taste system, such as property, color, smell, taste; internal data as the content of effective ingredients, moisture, ash content and extracts were collected to establish relational database with SVM; new TCM quality control mode was established on the basis of SVM. *Coptis chinensis* from Sichuan, Chongqing, Hubei and Yunnan were selected as subject. External and internal data and producing area information were collected to validate the effectiveness of the system. RESULTS & CONCLUSIONS: The recognition rate of 3 methods are all higher than 90%, including internal characteristics, external characteristics and whole internal and external characteristics; the recognition rate of whole internal and external characteristics is higher than that of internal characteristics or external characteristics. Relational and characteristics model of internal and external data are established with SVM method, and realize traditional experience of “identifying quality from appearance” through modern method to “knowing internal quality from external characteristics”; so, more standardized, accurate and objective new TCM quality control mode will be established on the basis of traditional experience.

KEYWORDS Bionic sensory system; Traditional Chinese medicine; Quality control; Support vector machines

中药的质量控制和评价是中医药研究的热点和难点^[1-2]。中药材及其饮片都有其自身特殊的颜色、形状、气味和味道,这与其所含化学成分有一定相关性。传统的外观鉴别是依赖经验,通过眼看、口尝和鼻闻,有较大的局限性和主观差异,给中药质量控制带来较大困难。而现行中药质量控制的模式是在参照国外植物药的质量控制方法的基础上,借鉴了化学药品质量控制模式,以化学定性鉴别与指标成分检测为主要手段^[3]。

但中药是一个多组分的复杂体系,单一的指标成分难以真正控制中药功效,也无法表征中药饮片的物质基础和化学成分群的整体性和复杂性^[4]。

随着仿生技术的发展,利用仿生视觉、嗅觉及味觉系统采集中药材及其饮片的形状、颜色、气味、味道等外部信息,再利用支持向量机(SVM)与其内在化学成分含量、水分、灰分、浸出物等数据建立关联数据库,从而得到“察外知内”的中药鉴别模式识别系统,最终实现中药质量控制新模式。有研究表

^Δ 基金项目:四川省中医药管理局科研项目(No.2014-F-075);全国大学生科研创新训练计划项目(No.201410613052)

* 本科生。研究方向:中药制剂。E-mail:tml0702@foxmail.com

通信作者:讲师,博士。研究方向:中药炮制与制剂。E-mail:wwangyidi@126.com

本栏目协办

四川博文网络科技有限责任公司

地址:四川省遂宁市射洪县滨江花园C栋
电话:0825-6698000 邮编:629200

明,仿生技术结合SVM建立的中药材分类模式对药材种类鉴别的准确率可达到90%以上^[5]。本文旨在结合仿生技术和化学定量方法,立足于SVM统计学习理论,探究中药整体质量控制新模式。

1 资料来源

以“电子鼻”“电子舌”“电子眼”“中药”“Electronic tongue”“Electronic nose”“Electronic nose eye”“Herb”等为关键词,组合查询1996—2015年中国知网(CNKI)和西南交通大学科学引文索引(SCI)等数据库中文献。结果,查询到相关文献1 826篇,其中有效文献50篇。在文献基础上对仿生系统及SVM的原理及其应用进行简要介绍,并拟将两者结合起来,建立“内外结合”的中药质量控制新模式。

2 仿生(感官)系统概述及其应用

随着微制造工艺和微电子机械加工技术的发展,各种新型传感器不断被研制出来,仿生视觉、嗅觉与味觉系统发展迅速,并且在各领域应用广泛,前景深远^[6-11]。仿生系统能客观评价物体的形、色、气味,并将采集到的信息翻译成对应形、色、气味的数据记录,其检测结果具有高灵敏度、可靠性、重复性。这使得此方法在客观表达和控制中药外观性状方面表现出巨大的应用潜力^[12]。现将其主要的原理及在中药行业的应用概述如下:

2.1 仿生视觉系统

仿生视觉系统又名机器视觉系统,是利用计算机技术模拟人的视觉功能,一般被称为电子眼(Electronic eye)。其利用机器来模拟人的视觉功能,采用图像摄取装置抓取图像,通过运算来提取目标特征(如尺寸、性状、颜色等),进而作出判别。电子眼已在工业生产、军事、医学等方面有广泛的应用,而且有对各类产品进行检验、监视的作用^[13-18]。陈梁等^[19]利用色彩色差仪客观量化不同种植方式的防风饮片的外观颜色差异,为防风饮片分级的建立奠定了基础。黄学思等^[20]提出将电子眼技术应用在中药炮制中,利用色彩装置设备对炮制后的药材进行颜色自动分析,采用CIE L*a*b确定获取颜色的基本信息。

2.2 仿生嗅觉系统

仿生嗅觉系统也称为电子鼻(Electronic nose),是能够感知和识别气味的电子系统。其将仿生学、传感技术、信号处理、模式识别和计算机科学等多种学科融于一体,模仿生物感官——鼻子的功能,既能检测特定的气体,又能评价混合气体或挥发性化学成分^[21]。由于电子鼻灵敏度高、鉴别速度快,已广泛应用于医学、食品工业、环境保护和军事等方面^[22-25]。Zhou X等^[26]利用电子鼻技术成功地应用于识别半夏硫熏品和自然干燥品,获得了较高的区分度。Abdullah MZ等^[27]利用电子鼻对超过100个马来西亚人参样品气味进行分类检测,其正确区分率大于90%。常健丽^[5]证实电子鼻对白前、徐长卿、五加皮等常见中药材质量有良好的辨识度。刘红秀等^[28]就电子鼻在中药材种类识别、道地中药材鉴别和中药百草油质量分析方面的应用作了详细介绍。孔繁瑶等^[29]应用电子鼻方法成功区分了不同产地的豨莶草。

2.3 仿生味觉系统

仿生嗅觉系统也称为电子舌(Electronic tongue),是20世纪后期发展起来的一种可分析、识别液体成分的智能仪器,主要由传感器阵列、信号处理和模式识别系统组成。传感器阵

列对液体样品作出响应并输出信号,信号经计算机系统进行处理后得到反映样品味觉特征的结果^[30]。该系统可模仿生物器官——舌头的功能,对人的5种基本味觉(酸、甜、苦、辣、咸)可进行有效的识别,而且试验证明其对液体、胶状和固体物质的味觉的检测均有良好的适用性^[31-32]。目前,电子舌在水质量的评定、制药行业、食品工程、环境及生物医学检测方面均有成功的应用^[33-35]。杜旭^[36-37]对电子鼻和电子舌技术在生药质量鉴别中的应用作了有益的尝试:利用气味传感器成功区分人体嗅觉难以分辨的白色、无臭的茯苓、牡蛎、龙骨粉末;利用味觉检测仪对不同味道的生药进行了分类;对从生药中提取的小檗碱利用味觉检测及高效液相色谱法进行测定,发现味觉检测仪的描绘强度与性能评价的实测强度间有显著相关性。

3 SVM概述与应用

SVM是Borges CJC等^[38]根据统计学习理论中结构风险最小化原则提出的一种新的机器学习算法,以训练误差作为优化问题的约束条件,以置信范围值最小化作为优化目标,即采用结构风险最小化准则的学习方法,使模型具有良好的泛化能力^[39]。

在建立SVM算法时需要对其进行模型选择,即核函数类别选择、核函数参数选择和二次规划参数选择。在选用核函数时,尽管满足Mercer条件的函数在理论上都可作为核函数,但不同的核函数,其分类器性能却不同。另一方面,即使选择了某一类核函数,其相应的参数如多项式的阶次、径向基函数的宽度也存在如何选择的问题。另外,SVM中二次规划参数如C参数、V参数等对分类器的推广能力也有影响。

SVM是机器学习领域的重要成果,正在成为继模式识别和神经网络研究之后机器学习领域新的研究热点。SVM在模式识别领域应用广泛,包括手写、字符识别、计算机视觉、网页和文本自动分类、人脸检测识别、头部姿态识别、性别分类、基因分类、目标识别、遥感图象分析等。在函数回归、估计、函数逼近、密度估计、时间序列预测、非线性系统控制等问题中,都有SVM的成功应用^[40-42]。此外,在中药质量检测、药材及复方识别、中药制剂工艺参数优化、中药片剂包衣质量分析等方面,SVM也显示出了良好的性能^[5,43-45]。汪劲等^[46]将SVM用于中药材的傅里叶红外光谱的快速鉴别方法中,取得了较好的效果。蔡从中、Wang JF等^[47-48]应用SVM在中药药方的识别和新中药药方的预测方面进行了有益的探索。朱恒民等^[49]基于SVM从中药提取工艺的历史数据中挖掘确定提取参数,并建立了较可靠的预测模型。李军等^[50]基于SVM对滴丸剂生产工艺参数进行优化,在实际生产中取得了良好效果。通过理论分析和仿真研究证明,该方法学习速度快、跟踪性能好、泛化能力强、对样本的依赖程度低,有较好的推广能力。

4 “内外结合”中药质量控制新模式的思考

笔者在查阅和分析仿生系统和SVM在中药质量控制中的应用时,认为仿生视觉、嗅觉及味觉系统可对符合《中国药典》标准的药材及饮片性状相关的形状、颜色、气味、味道等难以量化的外部信息进行数据化处理,实现客观表达。再运用化学定量方法(如仪器分析等)对药材有效成分含量、水分、灰分、浸出物等内部信息进行数据采集。然后将得到的信息分为训练集(即建模所用的数据库)和测试集(即对需要辨识的未知样品),并进行相关性分析,在高度相关的情况下(前提是

以内部特征作为基础),达到未知样本只需检测外部特征即可快速、无损、准确地控制药材或饮片整体质量的目的。

该方法主要原理为仿生系统中的传感器阵列可将待测样品输入信息(包括有效成分含量、水分等内在数据特征和形、色、气味等外部数据特征)转换成电信号,并分别建立SVM1和SVM2的算法,形成对应的模型1和模型2,通过对模型1和模型2的识别率等指标的相关性比对,建立一个新的“内外结合”的质量控制新模式(见图1)。当外部数据特征的识别率与内部数据特征的识别率一致时(差别在±5%以内),可采用外部数据特征构建的模型2对待测样本进行分析检测。

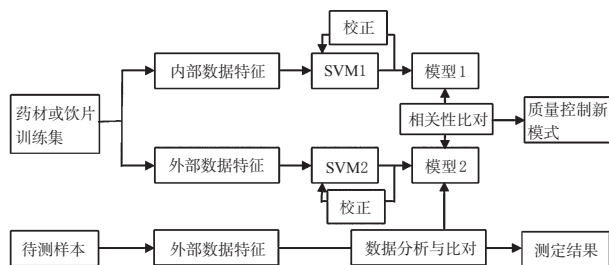


图1 基于支持向量机的中药质量控制新模式示意图

Fig 1 Diagrammatic sketch of new TCM quality control mode based on SVM

5 研究实例

试验选取来自四川、重庆、湖北和云南的黄连作为分析对象,为了保证样本的代表性和试验的客观性,在每个地区分别选取了4~12个不等的黄连样本,样本总数为36个,对每组样本提取其内、外部特征。其中,内部特征包括水分、总灰分、浸出物、表小檗碱、黄连碱、巴马汀、小檗碱含量共7维特征[参照2010年版《中国药典》(一部)相关项下的试验方法进行检验];外部特征分为3个部分进行提取,包括电子鼻提取的18维气味特征数据(由18根传感器LY2/LG、LY2/G、LY2/AA、LY2/GH、LY2/gCTL、LY2/gCT、T30/1、P10/1、P10/2、P40/1、P70/2、PA/2、P30/1、P40/2、P30/2、T40/2、T40/1、TA/2所采集到的数据,每根传感器的数据作为1维特征值),电子舌提取的7维味道特征数据(7根传感器ZZ、AB、GA、BB、CA、DA、JE所采集到的数据,每根传感器的数据作为1维特征值)和电子眼提取的5维外部形态特征数据(长度、宽度、厚度、外部颜色及内部颜色)。将以上所有试验得到的特征数据在MATLAB 7.0软件平台上降维及建立算法进行融合,每个黄连样本得到37维特征,总体的特征矩阵为36*37。特征矩阵经过SVM的学习、训练(指的是输入矩阵特征值后建立算法、构建模型的过程)和识别,最终输出结果为各个黄连样本的归属地。试验中特征数据降维处理、算法和建模过程在MATLAB 7.0软件上运算,识别过程如图2所示。

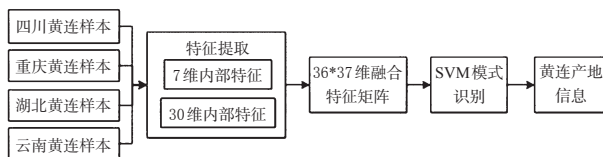


图2 黄连的特征提取和产地模式识别过程

Fig 2 Identification process of characteristics extraction and producing area mode of *C. chinensis*

为了对试验中提取到的内部特征与外部特征进行有效性验证,利用SVM针对不同的特征组合进行产地信息识别正确

率的对比试验:另收集4个产地的未知样本作为测试集,分别采集内外部特征数据,将所得数据输入到所构建的SVM模型中,进行结果的比对与相关性分析。结果,在不同特征组合黄连产地识别率中内部特征(7维)、外部特征(30维)、“内外结合”全部特征(37维)的识别率分别为93%、92%、96%。

结果表明:(1)内部特征、外部特征与“内外结合”全部特征这3种识别方法的识别率均大于90%(数据输入后由软件自动识别生成),能较好区分不同产地的黄连;(2)“内外结合”全部特征的识别率较单一的内部或外部特征高,说明“内外结合”的模式识别更能准确区分产地;(3)尽管外部特征比内部特征的识别率稍低,但其提取数据时对样本的无损、快速,与提取内部数据的耗时、破坏样品及耗费试剂资源等相比,具有一定优越性。

6 结语

综上所述,近年来关于仿生系统及SVM在中药研究方面均有一定应用,但将二者结合起来综合控制中药质量的却较少。笔者在分析相关文献基础上,提出基于SVM的“内外结合”中药质量控制新模式,即以传统经验表达转化为数字式的量化表达为基础的“内外结合”评价模式,使中药传统鉴别“辨状论质”经验得以现代“察外知内”的再现,且更加客观和准确,可用以补充和完善现有的中药质量控制手段和方法。

参考文献

- [1] 甘师俊,李振吉,邹健强.中药现代化发展战略[M].北京:科学技术文献出版社,1998:22.
- [2] 肖培根,肖小河.21世纪与中药现代化[J].中国中药杂志,2000,25(2):67.
- [3] 肖小河,金城,赵中振,等.论中药质量控制与评价模式的创新与发展[J].中国中药杂志,2007,32(14):1377.
- [4] 马梅芳,李洁.对中药质量控制的思考[J].中医药导报,2007,13(12):15.
- [5] 常健丽.基于机器嗅觉的中药材质量检测技术的研究[D].杭州:浙江师范大学,2012.
- [6] Röck F, Barsan N, Weimar U. Electronic nose: current status and future trends[J]. *Chem Res*, 2008, 108(2):205.
- [7] Vlasov Y, Legin A, Rudnitskaya A, et al. Nonspecific sensor arrays (“electronic tongue”) for chemical analysis of liquids: IUPAC technical report[J]. *Pure and Applied Chemistry*, 2005, 77(11):1965.
- [8] Abdullah MZ, Mohamad-Saleh J, Fathinul-Syahir AS, et al. Discrimination and classification of fresh-cut starfruits (*Averrhoa carambola* L.) using automated machine vision system[J]. *J Food Eng*, 2006, 76(4):506.
- [9] Rudnitskaya A, Legin A. Sensor systems, electronic tongues and electronic noses, for the monitoring of biotechnological processes[J]. *J Ind Microbiol Biotechnol*, 2008, 35(5):443.
- [10] 刘曙光,刘明远,何钺.机器视觉及其应用[J].河北科技大学学报,2000,21(4):11.
- [11] Yang SL, Xie SP, Xu M, et al. A novel method for rapid discrimination of bulbus of fritillaria by using electronic nose and electronic tongue technology[J]. *Analytical Methods*, 2015, 7(3):943.

- [12] 李文敏,吴纯洁,黄学思,等.基于电子鼻、电子舌技术实现中药性状气味客观表达的展望[J].中成药,2009,31(2):282.
- [13] 夏敏磊,郭斌,高峰.机器视觉中的采集技术研究[J].机电工程,2005,22(11):22.
- [14] 龚爱平.基于嵌入式机器视觉的信息采集与处理技术研究[D].杭州:浙江大学,2013.
- [15] 颜发根,刘建群,陈新,等.机器视觉及其在制造业中的应用[J].机械制造,2004,42(11):28.
- [16] 岑益科.基于机器视觉的鸡蛋品质检测方法研究[D].杭州:浙江大学,2006.
- [17] 冀瑜.基于机器视觉的高精度尺寸检测技术及应用研究[D].北京:中国计量科学研究院,2006.
- [18] 赵鹏.基于机器视觉的药品包装检测技术研究[D].长沙:湖南大学,2009.
- [19] 陈梁,李丽,肖永庆,等.仿野生与人工栽培防风饮片的色彩色差分析[J].中国实验方剂学杂志,2013,19(12):92.
- [20] 黄学思,吴纯洁,李文敏,等.基于色彩色差计和电子鼻的槟榔炒制火候判别及其指标量化研究[J].中国中药杂志,2009,34(14):1786.
- [21] 祝佳婧.电子鼻传感器技术的研制进展[J].北京生物工程,2002,21(4):298.
- [22] Westenbrink E, Arasaradnam RP, O'Connell N, et al. Development and application of a new electronic nose instrument for the detection of colorectal cancer[J]. *Biosens Bioelectron*, 2015, doi: 10.1016/j.bios.2014.10.044. Epub 2014 Oct 23.
- [23] Brudzewski K, Osowski S, Pawlowski W. Metal oxide sensor arrays for detection of explosives at sub-parts-per million concentration levels by the differential electronic nose[J]. *Sensor Actuat B-Chem*, 2012, 161(1):528.
- [24] Alam H, Saeed SH. Modern applications of electronic nose: a review[J]. *IAES International Journal of Electrical and Computer Engineering*, 2013, 3(1):52.
- [25] 吴守一,邹小波.电子鼻在食品行业中的应用研究进展[J].江苏理工大学学报:自然科学版,2000,21(6):13.
- [26] Zhou X, Wan J, Chu L, et al. Identification of sulfur fumed pinelliae rhizoma using an electronic nose[J]. *Pharmacogn Mag*, 2014, 10(37):135.
- [27] Abdullah MZ, Rahman ASA, Shakaff AYM, et al. Discrimination and classification of eurycoma longifolia jack in medicinal foods by means of a DSP-based electronic taste sensor[J]. *TI Meas Control*, 2004, 26(1):19.
- [28] 刘红秀,骆德汉,李卫东,等.基于电子鼻的中药材鉴别新方法研究[J].辽宁中医药大学学报,2011,13(10):32.
- [29] 孔繁瑶,邵露,胡慧华.应用电子鼻方法区分不同产地豨莶草[J].中国药房,2014,25(19):1793.
- [30] Legin AV, Rudnitskaya AM. The features of the electronic tongue in comparison with the characteristics of the discrete ion-selective sensors[J]. *Sensor Actuat B-Chem*, 1999, 58(1):464.
- [31] 胡洁,李蓉,王平.人工味觉系统:电子舌的研究[J].传感技术学报,2001(2):169.
- [32] 张爱霞.基于差别度检验的电子舌应用方法学研究[D].杭州:浙江工商大学,2012.
- [33] 宁珂.电子鼻与电子舌融合技术及其应用[D].吉林:东北电力大学,2014.
- [34] Ha D, Sun QY, Su KQ, et al. Recent achievements in electronic tongue and bioelectronic tongue as taste sensors [J]. *Sensor Actuat B-Chem*, 2015, doi: 10.1016/j.snb.2014.00.077.
- [35] Daikuzono CM, Dantas CAR, Volpati D, et al. Microfluidic electronic tongue[J]. *Sensor Actuat B-Chem*, 2015, doi: 10.1016/j.snb.2014.09.112.
- [36] 杜旭.味觉检测仪在生药气味评价及质量管理方面的应用[J].国际中医中药杂志,2006,28(2):109.
- [37] 杜旭.使用气味传感器对生药质量进行评价[J].国际中医中药杂志,2006,28(6):359.
- [38] Burges CJC. A tutorial on support vector machines for pattern recognition[J]. *Data Min Knowl Disc*, 1998, 2(2):121.
- [39] Hsu CW, Lin CJ. A comparison of methods for multiclass support vector machines[J]. *IEEE Trans Neural Networks*, 2002, 13(2):415.
- [40] 王国胜.支持向量机的理论与算法研究[D].北京:北京邮电大学,2007.
- [41] 常甜甜.支持向量机学习算法若干问题的研究[D].西安:西安电子科技大学,2010.
- [42] 叶磊,骆兴国.支持向量机应用概述[J].电脑知识与技术,2010,6(12):9815.
- [43] 张录达,苏时光,王来生.支持向量机(SVM)在傅里叶变换近红外光谱分析中的应用研究[J].光谱学与光谱分析,2005,25(1):33.
- [44] 孙燕,臧传新,任廷革.支持向量机分类器在中医方剂模式识别中的应用研究[J].中医药管理杂志,2006,14(11):25.
- [45] 孙燕,臧传新,任廷革,等.支持向量机方法在《伤寒论》方分类建模中的应用[J].中国中医药信息杂志,2007,14(1):101.
- [46] 汪劲,程存归.傅里叶变换红外光谱的SVM快速中药鉴别[J].仪器仪表学报,2005,26(8):710.
- [47] 蔡从中,袁前飞,肖汉光,等.中药组方的计算机辅助分类与识别[J].重庆大学学报:自然科学版,2007,29(10):42.
- [48] Wang JF, Cai CZ, Kong CY, et al. A computer method for validating traditional chinese medicine herbal prescriptions[J]. *Am J Chin Med*, 2005, 33(2):281.
- [49] 朱恒民,刘文杰,王宁生.数据挖掘技术在优化中药提取工艺中的应用[J].计算机与应用化学,2006,23(3):233.
- [50] 李军,黄海宽,曹琦.基于支持向量机的中药工艺参数优化研究[J].计算机工程与应用,2007,43(36):205.

(收稿日期:2015-03-15 修回日期:2015-05-21)

(编辑:林静)