

# 近红外光谱云计算分析系统构架与实现<sup>\*</sup>

黄华 祝诗平 刘碧贞

(西南大学工程技术学院,重庆 400716)

**摘要:**传统的近红外光谱分析系统是单机版,建模工作难度大。为了在近红外光谱分析中实现对现有光谱模型资源共享,提出利用云计算中心的高性能服务器代替单机版的主机,然后在云服务器上开发近红外光谱软件分析系统,并详细分析了近红外光谱云分析系统的构架与设计步骤。该系统可以实现近红外光谱数据的预处理、定量分析、定性分析以及光谱模型查找和光谱模型转移等功能。最后以泔水油近红外光谱定性鉴别为例,分析比较了单机版和基于云计算的近红外光谱分析结果。在近红外光谱云分析系统中,对预测集 58 个样品进行判别,总体鉴别正确率为 86.21%,这一结果与在单机环境下的分析结果完全一致。实验结果表明近红外光谱云分析系统具有成本低、建模方便、接入方式灵活、可实现资源共享和远程访问等优点。

**关键词:**近红外光谱 云计算 光谱建模

**中图分类号:** TP393.09; O657.33

**文献标识码:** A

**文章编号:** 1000-1298(2014)08-0294-05

## 引言

近红外光谱分析(Near infrared spectroscopy, NIR)技术具有快速、准确、不破坏样品等特点。目前已经广泛应用于食品<sup>[1-5]</sup>、药品<sup>[6-9]</sup>、农产品<sup>[10-11]</sup>等领域的品质检测中。如测定谷物中的水分、淀粉、蛋白质、脂肪、糖分及食用纤维含量等<sup>[12-13]</sup>;检测苹果、葡萄、梨、草莓等果汁加工过程中可溶性和总固形物的含量变化。NIR还广泛应用于纺织、造纸、生物化工、天体学及地理、地质学等领域。近红外光谱分析得到如此广泛应用主要归因于两个方面:一方面,随着计算机技术的发展,NIR逐渐实现了数字化,同时NIR光谱仪扫描速度、光谱分辨率也在不断提高。另一方面,由于化学计量学及其相应软件的发展,使得NIR定量和定性分析的精度得到了提高。无论NIR分析技术应用于何种行业,都需要建立一个完整的模型。在光谱建模过程中样品收集和理化检测是整个建模工作中费用高、耗时最多的。不同的研究者或不同的科研单位,对同一样品进行NIR分析时都需要重复进行样品收集和理化检测,从而造成硬件与软件的重复建设、资源利用率不高。

云计算(Cloud computing)是分布式处理、并行处理和网格计算的新发展<sup>[14]</sup>。本文结合云计算的

成本低、规模大、虚拟化和良好的可扩展性等优点,提出将近红外光谱分析系统移植到云计算中心的构思,使得不同的研究者或不同的科研单位共享已有的近红外光谱模型、共享云计算基础设施平台和NIR云计算软件分析系统,实现NIR分析的硬件和软件资源共享、节省成本投入。

## 1 云计算及其特点

云计算是计算机与网络技术发展相融合的产物,也是今后计算机与网络技术的发展趋势。目前对于云计算的定义还没有统一的标准<sup>[15-21]</sup>。根据美国国家标准与技术研究院(NIST)的定义:云计算是一种利用互联网实现随时随地、按需、便捷地访问共享资源池(如计算设施、存储设备、应用程序等)的计算模式。云计算包括了资源池、虚拟化、动态配置、效用计算、按需服务和网络服务传递等概念,使得应用程序开发更加灵活。云计算核心服务通常可以分为3个子层:基础设施即服务层(Infrastructure as a service, IaaS)、平台即服务层(Platform as a service, PaaS)、软件即服务层(Software as a service, SaaS)。云计算的结构如图1所示。

基础设施即服务层:主要包括计算资源和存储资源,整个基础设施也可以作为一种服务向用户提供。平台即服务层:在基础设施之上的平台层可以

收稿日期:2013-07-21 修回日期:2013-08-07

<sup>\*</sup>国家自然科学基金资助项目(31071319)和中央高校基本科研业务费资助项目(XDJK2013C107)

作者简介:黄华,讲师,博士生,主要从事系统检测研究,E-mail:hh317991@126.com

通讯作者:祝诗平,教授,博士生导师,主要从事系统检测研究,E-mail:zpswu@126.com

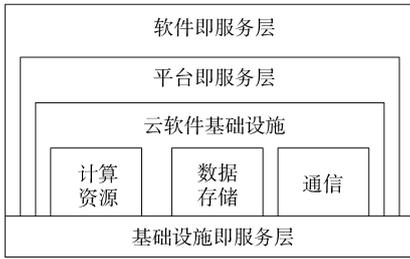


图 1 云计算结构

Fig. 1 Cloud computing architecture

认为是整个云计算系统的核心层,主要包括并行程序设计和开发环境、结构化海量数据的分布式存储管理系统、海量数据分布式文件系统以及实现云计算的其他系统管理工具。软件即服务层:它是面向用户提供简单的软件应用服务以及用户交互接口等。

## 2 近红外光谱云分析系统

在传统的近红外光谱分析中建立光谱分析模型难度大,主要是因为不同的研究者或科研单位相互独立,对同一样品进行近红外光谱分析时都需要重复进行相同的化学浓度测试实验,没有实现共享已有的光谱分析模型,造成硬件和软件重复建设、投入成本高、仪器设备利用率低、项目研发周期长。传统近红外光谱分析系统的完整流程如图 2 所示。

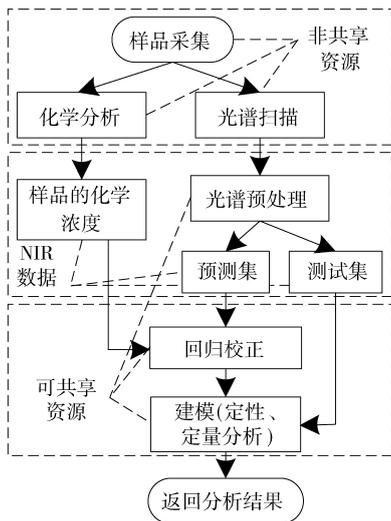


图 2 NIR 完整分析流程图

Fig. 2 Complete analysis process of NIR

从图 2 中可以看出,从光谱采集到得到光谱分析结果的过程中,同一样品的化学浓度可以在不同的研究者或不同的科研单位之间实现资源共享,同时光谱分析中所使用的相关软件及算法(如光谱分析的预处理、回归校正、定量和定性分析等)也可以在不同样品的光谱分析中实现算法移植与共享,并且在传统的近红外光谱分析系统中,运行这些算法的硬件资源,如高性能多核处理器、大容量存储器相

互独立,没有实现资源共享。如果能建立一个公共的近红外光谱分析平台,就可以实现近红外光谱分析中硬件和软件资源上的共享。

近红外光谱云分析系统的核心思想是:以云计算为平台,构建一个公共的、开放的近红外光谱分析系统,在网络上实现平台共用和资源共享。在云计算中心构建一个高性能的分布式服务器,然后在服务器上开发近红外光谱的预处理算法、定量分析算法以及定性分析算法等,并存储已有的光谱模型。使得任何人、在任何地方、任何时候都可以通过网页的方式访问该系统,利用已有的光谱分析模型,用户只需要上传样品的光谱数据,就能返回用户想要的分析结果。实现光谱模型和近红外分析系统在硬件和软件上的资源共享。避免近红外光谱分析硬件和软件上的重复建设,从而节省近红外光谱分析成本,缩短近红外光谱分析的研发周期,有利于推动近红外光谱分析技术的应用。

## 3 近红外光谱云分析系统构架

以云计算为中心,构建一个公共的、开放的近红外光谱云分析系统,如图 3 所示。

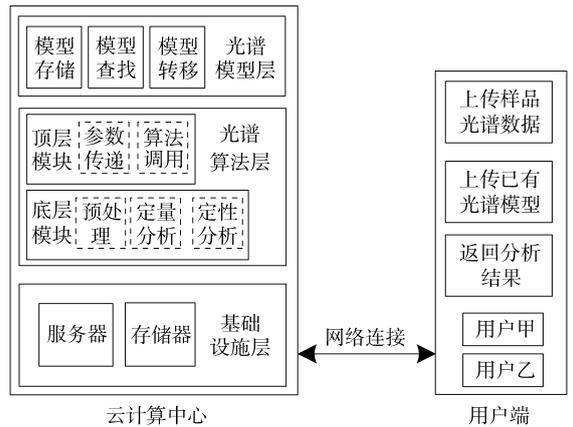


图 3 NIR 云分析系统构架

Fig. 3 Architecture of NIR analysis system

在 NIR 云分析系统中,整个系统由云计算中心和用户端两部分构成。云计算中心分成 3 个层:基础设施层、光谱算法层、光谱模型层。在基础设施层中,根据系统对处理器运算能力和存储容量的需求,在云计算中心按需租用一个相应的服务器,以提供光谱处理所需要的基础设施,如服务器和存储器等。基础设施层为光谱算法层和光谱模型层提供服务。在光谱算法层中,采用自下而上的设计思路,将底层处理模块与顶层控制模块分开设计。底层模块完成各种算法处理,一个模块实现一个算法。在底层模块算法中,包括光谱预处理算法如中值滤波、基线校正、正交信号校正(OSC)以及小波去噪等;光谱定

量、定性分析算法如主成分回归(PCR)、偏最小二乘回归(PLS)、人工神经网络(ANN)、支持向量机(SVM)等;在控制模块中实现各个底层模块的调用和参数传递。在光谱模型层中,实现对已有光谱模型的查找、新光谱模型的存储以及光谱模型的转换。

远程用户端通过有线或无线的网络连接登录近红外光谱云分析系统,远程用户只需要上传样品的NIR光谱数据,然后在近红外光谱云分析系统中选择相应的功能模块即可实现近红外光谱的远程分析,并返回分析结果至用户端。用户也可以上传新的光谱模型,让其他的研究者共享,进而丰富和扩大近红外光谱云分析系统。使用云计算的另一个优点是在这一分析过程中用户不必关心处理器的维护与升级,这些将由云计算中心提供服务。节省用户的硬件和软件的投入与维护成本。

### 4 NIR 云分析系统设计方案

目前国外的云计算服务提供商主要有 Amazon、Google、Microsoft、IBM 和 Dell 等;国内主要有阿里云、盛大云等。Amazon 的云计算平台提供可快速部署的虚拟服务器,实现了基础设施的按需分配。MapReduce 等新型并行编程框架简化了海量数据处理模型。Google 的 App Engine 云计算开发平台为应用服务提供商开发和部署云计算服务提供接口<sup>[14]</sup>。阿里云基于先进的虚拟化、分布式存储等云计算技术,同基础资源整合在一起。以 Web 的方式提供计算能力服务。

云计算主机由云计算服务提供商维护和升级。终端用户仅需要通过有线或无线的方式接入因特网,就能实现远程分析,节省了终端设备的成本。与传统的近红外光谱分析系统相比,近红外光谱云分析系统是将完成近红外光谱分析的主机搬到云计算中心,用户不需要构建主机所需要的硬件和软件成本花费,只需要按计算量来付费或租用主机,如同使用水、电、气一样方便。近红外光谱云分析系统实现步骤如图 4 所示。

从图 4 中可以看出,实现基于云计算的近红外光谱分析系统需要 6 个步骤:①对构建的系统进行需求分析。例如在云服务提供商处租用一个云服务器,采用 Xeon 2.26 GHz 双核处理器,数据存储器的容量为 100 GB,操作系统使用 Windows Server 2003。②云服务器上安装 Matlab 应用程序和 Web 开发软件包(SDK)。③安装云服务器主机需要的相关软件,建立一个完整的 NIR 主机。④以 Matlab 为开发工具,编写近红外光谱预处理、定量分析、定性分析的各种算法子程序。⑤在云服务器上发布近红外光

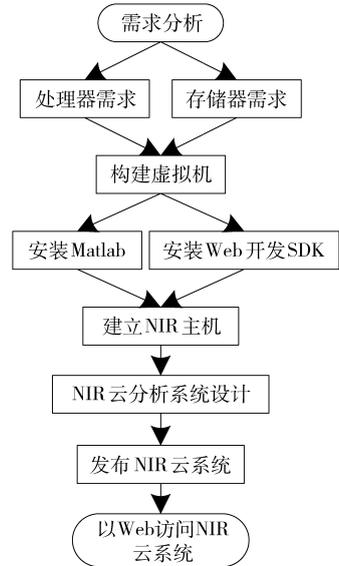


图 4 近红外光谱分析的云计算实现步骤

Fig. 4 Steps for NIR analysis system based on cloud computing

谱云分析系统。⑥用户通过网页访问云服务器主机和近红外光谱分析系统。

近红外光谱云分析系统主要功能如图 5 所示。

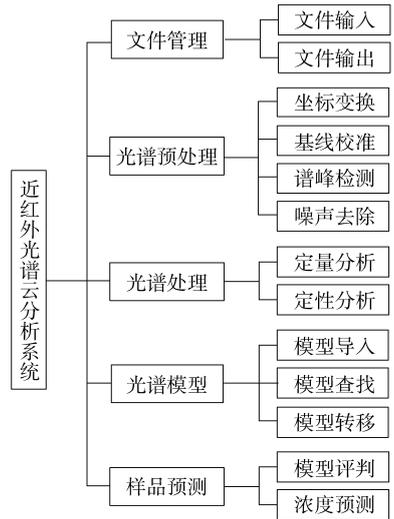


图 5 近红外光谱云计算分析系统功能框图

Fig. 5 Cloud computing analysis system of NIR

近红外光谱云计算分析系统主要功能包括光谱文件的输入、输出;各种光谱的预处理算法,如中值滤波、基线校正、正交信号校正以及小波去噪等;光谱定量、定性分析算法如主成分回归、偏最小二乘回归、人工神经网络和支持向量机等;光谱模型的导入、查找与转移;以及对未知样品的浓度的预测。近红外光谱云分析原型系统主界面如图 6 所示。

### 5 基于云计算的近红外光谱实例分析

#### 5.1 光谱数据来源

从火锅店、街边小吃店、酒店、超市、油脂公司油

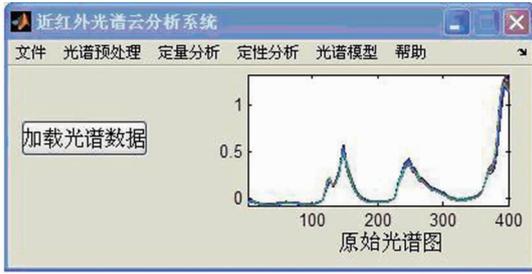


图 6 NIR 云分析原型系统主界面

Fig. 6 Main interface of NIR cloud system

脂储炼厂、废弃油脂回收冶炼厂、生物燃料公司等收集 82 份滴水油和合格食用植物油样品。采用 BRUKER 公司 MATRIX-F 型傅里叶变换近红外光谱仪对每份样品进行近红外光谱扫描,扫描范围  $11\ 996 \sim 3\ 999\ \text{cm}^{-1}$ ,扫描次数 32 次,分辨率  $8\ \text{cm}^{-1}$ ,光谱点数为 2 074<sup>[4]</sup>。

## 5.2 实验方法

将 82 份样品分成 2 组,第 1 组 24 份样品(其中合格油样品 13 份,编号范围 1~13,滴水油样品 11 份,编号范围 14~24),第 2 组 58 份样品(其中合格油样品 32 份,编号范围 1~32,滴水油样品 26 份,编号范围 33~58)。用“1”类代表合格油的类型,用“2”类代表滴水油的类型。

以第 1 组 24 份样品作为校正集,建立 DPLS 模型(光谱未进行预处理),鉴别第 2 组 58 份样品(预测集),将基于 Matlab 开发的云计算近红外光谱分析系统分别在本地计算机(主频 2.4 GHz,内存 4 GB)和云服务器(主频 2.26 GHz 双核处理器,内存 8 GB)上运行。两种情况下的鉴别结果完全相同,如图 7 所示。

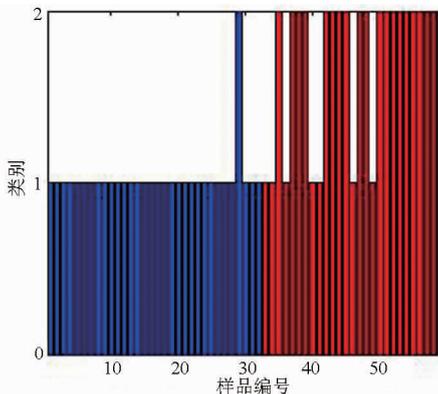


图 7 样品鉴别结果

Fig. 7 Sample identification results

由图 7 可知,用校正集 24 个样品建模,对预测集 58 个样品进行判别,光谱无预处理,总体有 8 个样品误判,总体鉴别正确率为 86.21%,有 1 个合格油样品被误判为滴水油,鉴别正确率为 96.88%,有

7 个滴水油样品被误判为合格油,鉴别正确率为 73.08%。这一结果与在单机环境下的分析结果完全一致<sup>[4]</sup>。

## 5.3 实验结果比较

通过对单机版和基于云计算的 NIR 分析系统的比较分析可知,基于云计算的 NIR 分析系统具有以下优势:①易于光谱建模。通过近红外光谱云分析系统的模型查找功能,获得相关光谱模型,不必重复已有的研究工作,节省了成本,缩短了项目研究周期。②按需配置,硬件扩容方便。在云计算中可以实现动态配置近红外光谱分析系统。当需要扩展服务器的计算能力或数据存储器的容量时,用户可以通过重新配置虚拟机来实现,操作简单。③接入方便。用户可以通过网页访问云服务器,使用近红外光谱云分析系统。可以实现远程在线操作。海量的光谱数据、现有的光谱模型及算法可以存储在云服务器上。④易于资源共享。任何人、在任何地方、任何时候都可以通过远程桌面使用近红外光谱云计算系统。实现了硬件基础设备和软件算法上的共享。单机版和云计算的 NIR 分析系统比较如表 1 所示。

表 1 单机版和云计算的 NIR 分析系统比较

Tab. 1 Comparison of NIR analysis system between stand-alone and cloud computing

| 比较项目   | 单机版 NIR 系统 | NIR 云分析系统 |
|--------|------------|-----------|
| 综合成本   | 高          | 低         |
| 硬件维护   | 需要         | 不需要       |
| 软件升级   | 需要         | 不需要       |
| 光谱模型共享 | 否          | 能         |
| 动态分配资源 | 否          | 能         |
| 接入方式   | 本地         | 远程        |

## 6 结束语

为了解决近红外光谱分析中资源共享的问题,提出了将近红外光谱分析系统移植到云计算中心的构思,给出了近红外光谱云分析系统的架构,详细分析了近红外光谱云分析系统的设计步骤。在云服务器上开发了近红外光谱分析系统,系统的主要功能包括光谱文件的输入、输出,各种光谱的预处理算法,光谱定量、定性分析算法,光谱模型的导入、查找与转移以及对未知样品的浓度预测。不同的研究者或不同的研究单位都可以通过网页访问和使用该系统。从而使得近红外光谱的建模工作简单、节省了硬件和软件成本,缩短了项目研究周期、实现了远程分析和资源共享。

## 参 考 文 献

- 1 Sánchez M T, La Haba D, Pérez-Marín D. Internal and external quality assessment of mandarins on-tree and at harvest using a portable NIR spectrophotometer [J]. *Computers and Electronics in Agriculture*, 2013, 92(3): 66–74.
- 2 Kemsley E K, Tapp H S, Binns R, et al. Feasibility study of NIR diffuse optical tomography on agricultural produce [J]. *Postharvest Biology and Technology*, 2008, 48(2): 223–230.
- 3 王多加, 周向阳, 金同铭, 等. 近红外光谱检测技术在农业和食品分析上的应用 [J]. *光谱学与光谱分析*, 2004, 24(4): 447–450.  
Wang Duoia, Zhou Xiangyang, Jin Tongming, et al. Application of near-infrared spectroscopy to agriculture and food analysis [J]. *Spectroscopy and Spectral Analysis*, 2004, 24(4): 447–450. (in Chinese)
- 4 祝诗平, 梁晶, 屠大伟, 等. 基于近红外光谱与 DPLS 的泔水油快速鉴别方法 [J]. *西南大学学报: 自然科学版*, 2012, 34(5): 1–6.  
Zhu Shiping, Liang Jing, Tu Dawei, et al. Study on rapid identification methods of waste edible oil based on near infrared spectroscopy and discriminant partial least squares [J]. *Journal of Southwest University: Natural Science Edition*, 2012, 34(5): 1–6. (in Chinese)
- 5 刘燕德, 应义斌. 基于 MATLAB 语言的苹果糖度近红外光谱定量分析 [J]. *浙江大学学报: 工学版*, 2004, 38(10): 1371–1374.  
Liu Yande, Ying Yibin. Quantitative analysis of near infrared spectra in apple sugar content based on MATLAB [J]. *Journal of Zhejiang University: Engineering Science*, 2004, 38(10): 1371–1374. (in Chinese)
- 6 Möltgen C V, Puchert T, Menezes J C, et al. A novel in-line NIR spectroscopy application for the monitoring of tablet film coating in an industrial scale process [J]. *Talanta*, 2012, 92(4): 26–37.
- 7 Scheibelhofer O, Balak N, Wahl P R, et al. Monitoring blending of pharmaceutical powders with multipoint NIR spectroscopy [J]. *AAPS PharmSciTech*, 2013, 14(1): 234–244.
- 8 Obregon L, Quinones L, Velazquez C. Model predictive control of a fluidized bed dryer with an inline NIR as moisture sensor [J]. *Control Engineering Practice*, 2013, 21(4): 509–517.
- 9 王思宏, 尹起范, 范艳玲, 等. 长白山地区几种红景天品种的傅里叶变换红外光谱法鉴别研究 [J]. *光谱学与光谱分析*, 2004, 24(8): 957–959.  
Wang Sihong, Yin Qifan, Fan Yanling, et al. Identification studies of wild and cultivated *Rhodiola saccharinensis* A. Bor and *Rhodiola angusta* Nakai taken from Changbai Mountains by Fourier transform infrared spectrometry [J]. *Spectroscopy and Spectral Analysis*, 2004, 24(8): 957–959. (in Chinese)
- 10 Jetter K, Depczynski U, Molt K, et al. Principles and applications of wavelet transformation to chemometrics [J]. *Analytica Chimica Acta*, 2000, 420(2): 169–180.
- 11 Rinnan Å, Berg F, Engelsen S B. Review of the most common pre-processing techniques for near-infrared spectra [J]. *TrAC Trends in Analytical Chemistry*, 2009, 28(10): 1201–1222.
- 12 马毅, 汪西原, 雍慧. 小波变换去噪应用于鲜枣糖度近红外光谱检测的研究 [J]. *计算机与应用化学*, 2011, 28(3): 303–306.  
Ma Yi, Wang Xiyuan, Yong Hui. Study on fresh jujube sugar content using near infrared spectroscopy based on wavelet transform denoising [J]. *Computers and Applied Chemistry*, 2011, 28(3): 303–306. (in Chinese)
- 13 Ferrari M, Quaresima V. A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application [J]. *Neuroimage*, 2012, 63(2): 921–935.
- 14 罗军舟, 金嘉晖, 宋爱波, 等. 云计算: 体系架构与关键技术 [J]. *通信学报*, 2011, 32(7): 3–21.  
Luo Junzhou, Jin Jiahui, Song Aibo, et al. Cloud computing: architecture and key technologies [J]. *Journal on Communications*, 2011, 32(7): 3–21. (in Chinese)
- 15 Garg S K, Versteeg S, Buyya R. A framework for ranking of cloud computing services [J]. *Future Generation Computer Systems*, 2013, 29(4): 1012–1023.
- 16 Calheiros R N, Ranjan R, Beloglazov A, et al. CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms [J]. *Software: Practice and Experience*, 2011, 41(1): 23–50.
- 17 Marston S, Li Z, Bandyopadhyay S, et al. Cloud computing—the business perspective [J]. *Decision Support Systems*, 2011, 51(1): 176–189.
- 18 杨辉华, 李灵巧, 覃炳达, 等. CloudChem—基于云计算的化学计量学软件服务 [J]. *计算机与应用化学*, 2011, 28(5): 523–526.
- 19 Armbrust M, Fox A, Griffith R, et al. A view of cloud computing [J]. *Communications of the ACM*, 2010, 53(4): 50–58.
- 20 郑博, 赵不贿, 景亮. 农业图像云平台中参数化图像处理系统设计方法 [J]. *农业机械学报*, 2012, 43(7): 168–173.  
Zhen Bo, Zhao Buhui, Jing Liang. Design method of parametric image processing system in agricultural image cloud [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2012, 43(7): 168–173. (in Chinese)
- 21 Buyya R, Yeo C S, Venugopal S, et al. Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility [J]. *Future Generation Computer Systems*, 2009, 25(6): 599–616.

- 18 叶伯生, 郭显金, 熊烁. 计及关节属性的 6 轴工业机器人反解算法[J]. 华中科技大学学报: 自然科学版, 2013, 41(3): 68-71.  
Ye Bosheng, Guo Xianjin, Xiong Shuo. 6-axis industrial robot inverse kinematic solution algorithm based on joint attribute[J].  
Journal Huazhong University of Science and Technology: Natural Science Edition, 2013, 41(3): 68-71. (in Chinese)
- 19 胡俊峰, 张宪民, 朱大昌. 柔性并联机器人动力学建模[J]. 农业机械学报, 2011, 42(11): 208-213.  
Hu Junfeng, Zhang Xianmin, Zhu Dachang. Dynamic modeling of flexible parallel robot[J]. Transactions of the Chinese Society for  
Agricultural Machinery, 2011, 42(11): 208-213. (in Chinese)
- 20 车林仙. 6-CPS 正交并联机器人位置正解分[J]. 农业机械学报, 2009, 40(11): 212-218.  
Che Linxian. Positional analysis of 6-CPS orthogonal parallel manipulators[J]. Transactions of the Chinese Society for  
Agricultural Machinery, 2009, 40(11): 212-218. (in Chinese)

## Stiffness Model of a Machining Robot

Zhang Yonggui Liu Wenzhou Gao Jin'gang

(School of Mechanical and Electrical Engineering, Lanzhou University of Technology, Lanzhou 730050, China)

**Abstract:** Taking robot arm as research objective, this paper analyzes and researches the stiffness model of a machining robot. First, the size of the equivalent stiffness is calculated according to the robot transmission components, which are come down to elastic components and reduced to the corresponding joints. After obtaining the stiffness of each robot joint, the joint deflection performance in robot end is mapped out by using Jacobian matrix. In addition, the robot arm is regarded as a flexible rod, and the deflection deformation models of stretch, torsion and curve are built, respectively. And then the robot kinematics model is established, in which the staging parameters are contained. From calculating the operation end mapping, each arm rod deflection has micro variable. Finally, the joint deflection and arm rod deflection are stacked in the robot operation end, and then the robot general operating space stiffness is solved through stiffness definition.

**Key words:** Robot arm Joint stiffness Machining Arm stiffness Deflection at the end of robot arm

(上接第 298 页)

## Architecture and Implementation of NIR Analysis System Based on Cloud Computing

Huang Hua Zhu Shiping Liu Bizhen

(College of Engineering and Technology, Southwest University, Chongqing 400716, China)

**Abstract:** Traditional NIR systems are stand-alone and difficult to model. In order to share resource of existed NIR model, a NIR analysis system based on cloud computing is proposed. The NIR software analysis system is designed on a high performance server instead of the host of stand-alone version, and the architecture and design procedures of the NIR cloud analysis system are described in detail. The system has the functions of preprocessing of near infrared spectral data, quantitative analysis, qualitative analysis, spectral model search and spectral model transfer. The identification results of waste edible oil between NIR cloud analysis system and stand-alone version are compared. The overall rate of correct identification is 86.21% for the 50 samples of waste edible oil by the NIR cloud analysis system. This result is fully consistent with the analysis results in stand-alone. The experimental results show that the cloud NIR analysis system is low cost, easy modeling, access to flexible, enabling resource sharing and remote access, etc.

**Key words:** Near infrared spectroscopy Cloud computing Spectral modeling