

RBF 模糊神经网络用于 NIR 鉴别羊绒和羊毛的可行性研究

郭飞¹ 刘净净² 罗霄¹ 刘刚³

¹北京服装学院信息工程学院, 北京 100029
²北京服装学院图书馆, 北京 100029
³河南工业大学信息科学与工程学院, 河南 郑州 450001

摘要 为了实现羊绒、羊毛纤维的快速、无损检测,建立了羊绒、羊毛近红外光谱数据库,包括 228 组各地羊绒、羊毛数据,并应用于羊绒、羊毛的定性检测上。首先介绍了羊绒、羊毛近红外光谱检测的数据库建立过程;然后,在对羊绒、羊毛原始近红外光谱进行预处理的基础上,对数据进行主成分分析,选出 12 种主成分,并结合改进的 RBF 模糊神经网络,建立羊绒、羊毛检测模型。通过与主成分分析-马氏距离建模方法的对比分析实验表明,建立近红外光谱数据库,并结合主成分分析和改进的 RBF 模糊神经网络的方法是一种有效的无损检测羊绒、羊毛的方法,可快速建立高精度的羊绒、羊毛纤维检测模型。

关键词 光谱学;近红外光谱学;RBF 模糊神经网络;羊绒;羊毛;主成分分析

中图分类号 TP183;O657.33 **文献标识码** A **doi**: 10.3788/LOP49.083001

Feasibility Study of RBF Fuzzy Neural Network in Cashmere and Wool Identification Based on Near Infrared Spectroscopy

Guo Fei¹ Liu Jingjing² Luo Xiao¹ Liu Gang³

¹ School of Information Engineering, Beijing Institute of Fashion Technology, Beijing 100029, China
² Library, Beijing Institute of Fashion Technology, Beijing 100029, China
³ College of Information Science and Engineering, Henan University of Technology, Zhengzhou, Henan 450001, China

Abstract In order to realize the fast and nondestructive detection, the cashmere and wool near infrared spectroscopy database is created which includes the data of 228 groups of cashmere and wool from various districts, and it is applied to the qualitative detection of cashmere and wool. First the process of database creation in the cashmere and wool detection based on near infrared spectroscopy is introduced. Then on the base of the near-infrared spectroscopy original data preprocessing of cashmere and wool, the principal components of the data are analyzed, and 12 kinds of principal components are chosen. The detection model of cashmere and wool with radial basis function (RBF) fuzzy neural network is build. The comparative analysis experiments with PCA-MD modeling demonstrate that the method combining near infrared spectroscopy database, principal components analysis (PCA) and improved RBF fuzzy neural network is an effective and nondestructive detection method for cashmere and wool, and it can rapidly build high-accuracy detection models of cashmere and wool fiber.

Key words spectroscopy; near-infrared spectroscopy; radial basis function fuzzy neural network; cashmere; wool; principal components analysis

OCIS codes 300.6340; 200.4260; 200.4560

收稿日期: 2012-01-19; **收到修改稿日期**: 2012-03-26; **网络出版日期**: 2012-05-05

基金项目: 国家自然科学基金(61174056),北京市优秀人才培养资助项目(2009D005001000003)和北京市教委基金(KM201110012010)资助课题。

作者简介: 郭飞(1978—),女,博士,讲师,主要从事智能识别与控制、计算机应用、纺织自动化等方面的研究。

E-mail: lunwen95@163.com

1 引言

山羊绒即通常所说的羊绒,是一种非常珍贵的动物纤维,和羊毛同属天然蛋白质纤维,二者在表面形态、结构和化学性能上都非常相似,且越细的羊毛在形态结构上越接近羊绒,越难鉴别。在纺织服装领域,对于羊绒、羊毛的成分检测一直都是研究的难点。目前羊绒、羊毛鉴别的主要方法有:形态特征分析法、化学方法、计算机图像分析法和生物芯片法等^[1,2]。这些方法存在准确性差、过多依赖于检测人员经验和成本高等缺点,影响了其推广使用。

近红外光谱(NIR)检测技术作为一种快速、无损的检测技术,可利用全谱数据进行定性、定量分析,近年来越来越多的受到国内外学者的关注,已经被越来越多的应用于各种行业,主要包括农业、轻工食品、医药、石油化工和环境保护^[3-5]等方面,已成为一种重要的检测技术。在毛纺织领域,近红外光谱技术也取得了一定的可信成果并有着广阔的应用前景,主要被用于羊毛直径、含水量和二氯甲烷可溶性物质测定等,以及混纺毛织品中羊毛含量的测定^[6,7],羊绒、羊毛的定性、定量检测等领域。其中在羊绒、羊毛纤维的定性检测方面主要采用主成分分析-马氏距离(PCA-MD)或者主成分分析结合 BP(Back-propagation)神经网络建模的方法^[8,9]。

本文首先建立了羊绒、羊毛近红外光谱数据库,并针对羊绒、羊毛近红外光谱数据的特点,采用主成分分析结合改进型 RBF(Radial basis function)模糊神经网络的方法建立其检测模型。并将其应用于基于近红外光谱的羊绒、羊毛纤维检测中,实现了羊绒、羊毛纤维的快速、无损、准确的鉴别。改进型 RBF 模糊神经网络的建模方法具有收敛速度快、泛化能力强的特点,避免了 BP 神经网络建模时易陷入局部最小的缺点,有很好的全局逼近能力。

2 羊绒、羊毛近红外光谱检测模型的数据库建立过程

羊绒、羊毛近红外光谱检测模型的数据库建立过程如图 1 所示,其主要处理过程如下:

1) 数据库的建立

为了能够更方便有效的分析处理数据,本文用微软的 SQL SERVER 2005 建立羊绒、羊毛近红外光谱数据库。数据库包括近红外光谱原始谱图表(NIR)、二阶谱图表(SNIR)、光谱主成分表(PCA)、模糊神经网络学习参数表(FANN)等。

羊绒、羊毛近红外光谱数据的采集,采用美国 FOSS 公司生产的 6500 光栅型近红外光谱仪,Pbs 检测器,测定范围为 1100~2500 nm,光谱采集时间为每秒 1.8,分辨率 2 nm,软件采用 FOSS 公司 VISION3.4。

采用不同产地的羊绒、羊毛纤维。产地有内蒙古自治区鄂尔多斯、内蒙古锡林浩特、辽宁本溪、辽宁丹东、阿尔巴斯、阿拉善盟、鄂托克旗和甘肃等。

2) 数据选择

在波长 1100~2500 nm 范围内,采用反射模式采集样品光谱,所有样品均重复采集 4 次光谱,建立原始谱图数据库。

3) 数据预处理

扫描所得羊绒、羊毛纤维的 NIR 原始谱图如图 2 所示。

由于羊绒与羊毛的物质组成与化学结构几乎相同,且其内部的化学基团对近红外光的吸光度相近,因此羊绒与羊毛的原始谱图看起来非常相似,直接通过原始谱图很难鉴别羊绒与羊毛。本文采取对原始谱图求

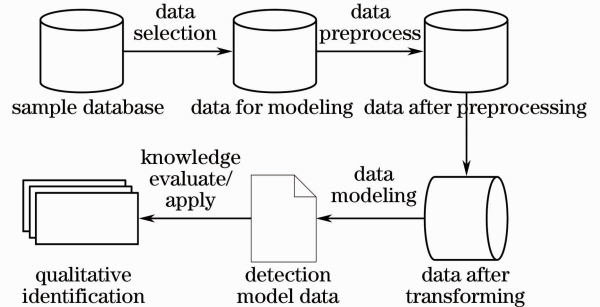


图 1 羊绒、羊毛检测模型数据库建立过程

Fig. 1 Database creation processes of cashmere and wool detection model

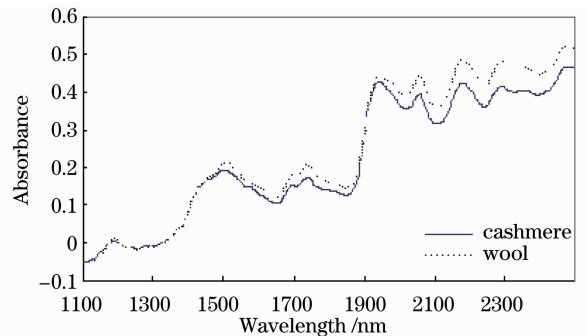


图 2 羊绒、羊毛纤维的 NIR 原始谱图

Fig. 2 Original NIR spectrum of cashmere and wool

二阶导数的方法,以降低基线的漂移,强化光谱信号^[9]。图3为经过二阶导数处理后的羊绒、羊毛光谱图。

4) 数据变换

实验采集的羊绒、羊毛光谱从1100~2500 nm 共有700个数据点,数据量大,冗余信息多,需要对数据进行降维处理,以更好的适应模糊神经网络处理数据的要求。

主成分分析(PCA),是多元统计中的一种数据挖掘技术,其目的就是为了降维,构造少数几个综合主成分变量,在尽可能多地表征原变量的数据特征的同时消除众多信息共存中相互重叠的信息部分^[10]。

对羊绒、羊毛光谱数据进行主成分分析,经过计算前12个主成分的累积贡献率已经达到99.713%(如表1所示),说明这12个变量很好的表征了原光谱数据特征。

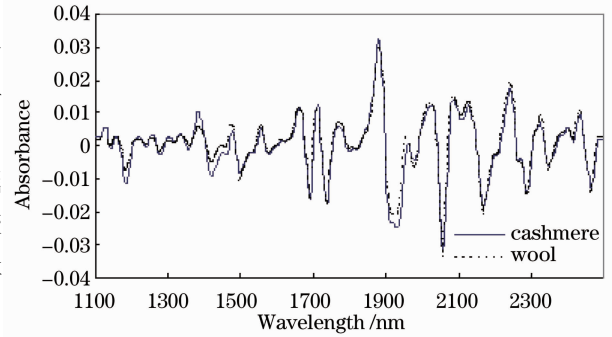


图3 二阶导数处理后羊绒、羊毛 NIR 光谱图
Fig.3 NIR spectrum of cashmere and wool by second derivative

表1 主成分贡献率

Table 1 Contribution rate of principal components

PCs	Contribution rate /%	PCs	Contribution rate /%	PCs	Contribution rate /%
PC1	78.693	PC5	94.82	PC9	98.392
PC2	87.652	PC6	96.153	PC10	98.991
PC3	90.784	PC7	97.053	PC11	99.49
PC4	93.019	PC8	97.764	PC12	99.713

5) 数据建模

化学计量学中建模的主要方法有偏最小二乘法(PLS)、K-最邻近分类(KNN)、SIMCA、主成分分析-马氏距离(PCA-MD)和人工神经网络等方法^[11~13]。本文根据羊绒、羊毛近红外光谱数据的特点,采用改进的RBF模糊神经网络方法。

6) 知识评价和应用

将建立的羊绒、羊毛检测模型,用于羊绒、羊毛的定性检测,实现了羊绒、羊毛纤维的快速、无损、准确的鉴别,后面将详细介绍。

3 改进型 RBF 模糊神经网络建模方法

模糊神经网络是模糊理论与神经网络结合的产物,它结合二者的优点,集学习、识别、联想和信息处理等于一体^[14~16]。多输入单输出 RBF 模糊神经网络结构如图4所示,该网络由前件网络和后件网络两部分组成,分别匹配模糊规则的前件和后件^[17]。

3.1 前件网络

前件网络进行模糊推理,计算每条规则的适应度 α_l ,并进行归一化处理,数学描述如下

$$\alpha_l = \min(\mu_1^i, \mu_2^j, \dots, \mu_n^i), \quad l = 1, 2, \dots, m, \quad m = \prod_{i=1}^n m_i, \quad j = 1, 2, \dots, m_i, \quad \bar{\alpha}_l = \frac{\alpha_l}{\sum_{l=1}^m \alpha_l}, \quad (1)$$

式中 $\mu_i^j = \exp\left[-\frac{(x_i - c_{ij})^2}{\sigma_{ij}^2}\right]$ ($i = 1, 2, \dots, n; j = 1, 2, \dots, m_i$) 是输入分量隶属于各语言变量值模糊集合的隶属度; m_i 是 x_i 的模糊分割数, c_{ij} 和 σ_{ij} 分别表示隶属度函数的中心和宽度。

3.2 后件网络

后件网络共有三层,由 r 个结构相同的并列子网组成,每个子网产生一个输出量。后件网络用于计算每

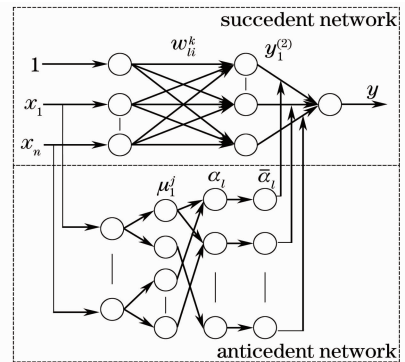


图4 RBF 模糊神经网络结构

Fig.4 RBF fuzzy neural network structure

条规则的后件,每个节点代表一条规则,数学表述如(2)式所示:

$$y_l^{(2)} = w_{l0}^i + w_{l1}^i x_1 + \dots + w_{ln}^i x_n = \sum_{k=0}^n w_{lk}^i x_k \quad (i = 1, 2, \dots, r; l = 1, 2, \dots, m) \quad (2)$$

系统输出为 $y = \sum_{l=1}^m \bar{\alpha}_l y_l^{(2)}$, 即为各规则后件的加权和, 加权系数为各模糊规则经归一化的适应度, 也就是前件网络的输出用于后件网络第三层连接权值。

3.3 参数修正

假设各输入量的模糊等级数是预先确定的, 那么要学习的主要参数就是后件网络的连接权 w_{li}^k 以及前件网络的第二层各结点隶属度函数的中心值 c_{ij} 和宽度 σ_{ij} 。RBF 模糊神经网络的结构本质上也是一种多层前馈网络, 可采用 BP 网络的误差反传的方法来设计调整参数的学习算法。

3.4 算法改进

为了加快模糊神经网络的学习速度, 在后件网络中采用了权值的二阶学习, 加重了权值的作用, 即规则的加权系数为权值的平方, 由线性的加权变为非线性加权, 公式如下:

$$y_l^{(2)} = (w_{l1}^i)^2 x_1 + \dots + (w_{ln}^i)^2 x_n = \sum_{k=0}^n (w_{lk}^i)^2 x_k \quad (3)$$

同时, 采用梯度算法修正连接权 w_{li}^k 、中心值 c_{ij} 和宽度 σ_{ij} 时, 分别采用不同的学习速率 $\lambda_w, \lambda_c, \lambda_\sigma$, 且各自的学习速率采用自适应寻优策略, 在学习过程中自动调整学习速率。以连接权 w_{li}^k 的学习速率 λ_w 为例, 调整策略如下:

$$\lambda_w(t) = 1.1^d \lambda_w(t-1), \quad d = \text{sgn}[D(t)D(t-1)], \quad (4)$$

式中 $D(t)$ 和 $D(t-1)$ 分别为权系数 w_{li}^k 在 t 和 $t-1$ 时刻的负梯度。即当连续两次迭代梯度方向相同时, 表明收敛慢, 可加快学习速率; 当连续两次迭代梯度方向相反时, 表明收敛过快, 需减小学习速率。通过学习速率的自适应学习, 避免了因学习速率不合适而产生的收敛振荡和收敛速度过慢问题。

4 羊绒、羊毛近红外光谱检测模型实例分析

采用 60 组羊绒谱图数据和 20 组羊毛谱图数据作为训练集数据。前 12 个主成分作为 RBF 模糊神经网络的输入, 输出节点为 1 (值 1 为羊绒, 值 2 为羊毛), 目标误差为 0.003。

用已建立的结合主成分分析和 RBF 模糊神经网络模型对 10 组未知样本 (羊绒、羊毛各 5 组) 进行定性预测, 其中 1 号样本和 6 号样本分别为含水量较大的羊绒和羊毛样本, 预测结果全部正确, 如表 2 所示。

表 2 未知样本的 RBF 模糊神经网络预测结果

Table 2 Prediction results for unknown samples by RBF fuzzy neural network

Samples	Real values	Prediction values	Samples	Real values	Prediction values
1	1	1.003	6	2	1.998
2	1	1	7	2	1.999
3	1	0.998	8	2	2
4	1	1	9	2	2.002
5	1	1.001	10	2	2.001

同时采用常规定性建模方法主成分分析-马氏距离方法 (PCA-MD) 进行对比分析, 结果见表 3 (P 表示通过, F 表示未通过)。从表 3 中可以看出, 对于样本 1 和 样本 6 出现了预测错误的情况。这是因为马氏距离的聚类方法精度依赖于样本的协方差矩阵的计算, 而含水量较大的羊绒、羊毛样本

表 3 未知样本的 PCA-MD 预测结果

Table 3 Prediction results for unknown samples by PCA-MD

Samples	Prediction values	Cashmere	Wool	Samples	Prediction values	Cashmere	Wool
1	0.912	F	P	6	0.774	P	F
2	0.635	P	F	7	1.000	F	P
3	0.579	P	F	8	0.968	F	P
4	0.727	P	F	9	0.897	F	P
5	0.623	P	F	10	0.937	F	P

其近红外光谱的协方差矩阵和普通样本差值较大,使得传统的 PCA-MD 方法对于含水量较大的羊绒和羊毛样本预测效果较差。改进型 RBF 模糊神经网络,结合了 RBF 神经网络和模糊逻辑的优点,提高了神经网络学习的泛化性能,增强系统的容错性,具有全局逼近能力,因此采用基于 PCA 的改进型 RBF 模糊神经网络建模方法具有更准确的预测效果。

5 结 论

在纤维鉴别中,由于羊绒和羊毛的化学特性相差不大,仅在鳞片结构和纤维细度方面有些差异,所以羊绒、羊毛纤维的鉴别是公认的技术难度最大、最不易鉴别的纤维,很多科技人员长期以来都致力于羊绒羊毛检测方面的研究,但检测结果一直不能达到理想的效果。本文首先建立了羊绒、羊毛近红外光谱数据库,对数据采用主成分分析的方法进行了变换,并结合改进型的 RBF 模糊神经网络进行预测建模。与常规的主成分分析—马氏距离方法进行了对比定性分析实验,实验结果表明本文中的主成分分析-RBF 模糊神经网络方法达到了较好的预测效果。实验说明该模型对于羊绒、羊毛的检验有较高的准确率和快速的识别速度,更符合实际应用的需要,为开发简单、低成本、无损、快速的羊绒和羊毛检验设备奠定了基础。

参 考 文 献

- 1 Jiang Gaoping, Zhong Yueqi, Wang Rongwu. Identification of wool and cashmere based on characteristics of spectral line [J]. *Journal of Textile Research*, 2010, **31**(4): 15~19
蒋高平, 钟跃崎, 王荣武. 基于谱线特征的羊绒与羊毛的鉴别[J]. *纺织学报*, 2010, **31**(4): 15~19
- 2 Zhang Xiaoli, Chi Haitao, Zhang Jinghua *et al.*. Research progress in the detection technique of cashmere and wool[J]. *Wool Textile Journal*, 2009, **37**(3): 56~59
张小莉, 池海涛, 张经华 等. 羊绒和羊毛纤维检测技术研究发展[J]. *毛纺科技*, 2009, **37**(3): 56~59
- 3 Li Gang, Zhao Jing, Li Jiaying *et al.*. Visible-infrared reflectance spectroscopy applied in rapid screen of diseases[J]. *Acta Optica Sinica*, 2011, **31**(3): 191~196
李 刚, 赵 静, 李家星 等. 可见-近红外反射光谱用于疾病快速筛查[J]. *光学学报*, 2011, **31**(3): 191~196
- 4 McCaig T. N.. Extending the use of visible/near-infrared reflectance spectrophotometers to measure color of food and agricultural products[J]. *Food Res. Int.*, 2002, **35**(8): 731~736
- 5 Sun Guangming, Liu Fei, Zhang Fan *et al.*. Fast determination of proline in herbicide-stressed oilseed rape leaves based on near infrared spectroscopy[J]. *Acta Optica Sinica*, 2010, **30**(4): 1192~1196
孙光明, 刘 飞, 张 帆 等. 基于近红外光谱技术检测除草剂胁迫下油菜叶片中脯氨酸含量的方法[J]. *光学学报*, 2010, **30**(4): 1192~1196
- 6 Innocenti R., Zoccola M.. Near infrared reflectance spectroscopy as a tool for the determination of dichloromethane extractable matter and moisture content in combed wool slivers[J]. *J. Near Infrared Spec.*, 2003, **11**(5): 333~340
- 7 Yuan Hongfu, Chang Ruixue, Tian Lingling *et al.*. Study of nondestructive and fast identification of fabric fibers using near infrared spectroscopy[J]. *Spectrosc. & Spectral Anal.*, 2010, **30**(5): 1229~1233
袁洪福, 常瑞学, 田玲玲 等. 纺织纤维及其制品非破坏性快速鉴别的研究[J]. *光谱学与光谱分析*, 2010, **30**(5): 1229~1233
- 8 Lü Dan, Yu Chan, Zhao Guoling. Study on identification of cashmere and wool using near infrared spectroscopy[J]. *Journal of Beijing Institute of Clothing Technology (Natural Science Edition)*, 2010, **30**(2): 29~34
吕 丹, 于 婵, 赵国樑. 利用近红外光谱进行羊绒与羊毛的鉴别技术研究[J]. *北京服装学院学报(自然科学版)*, 2010, **30**(2): 29~34
- 9 Wu Guifang, Zhu Dengsheng, He Yong. Identification of fine wool and cashmere by using Vis/NIR spectroscopy technology[J]. *Spectrosc. & Spectral Analy.*, 2008, **28**(6): 1260~1263
吴桂芳, 朱登胜, 何 勇. 可见-近红外光谱用于鉴别山羊绒与细支绵羊毛的研究[J]. *光谱学与光谱分析*, 2008, **28**(6): 1260~1263
- 10 Lu Wanzhen. *Modern Near Infrared Spectroscopy Analytical Technology (Second Edition)* [M]. Beijing: China Petrochemical Press, 2007
陆婉珍. *现代近红外光谱分析技术(第二版)*[M]. 北京: 中国石化出版社, 2007
- 11 Chu Xiaoli. *Molecular Spectroscopy Analytical Technology Combined with Chemometrics and Its Applications*[M]. Beijing:

Chemical Industry Press, 2011

楮小立. 化学计量学方法与分子光谱分析技术[M]. 北京: 化学工业出版社, 2011

- 12 Li Yaping, Zhang Guangjun, Li Qingbo. Application of O2-PLS in experimental study on non-invasive measurement of blood glucose[J]. *Acta Optica Sinica*, 2010, **30**(3): 854~860
李亚萍, 张广军, 李庆波. 基于 O2-PLS 方法的血糖无损检测实验研究[J]. *光学学报*, 2010, **30**(3): 854~860
- 13 Zhang Haihong, Zhang Shujuan, Wang Fenghua *et al.*. Study on fast discrimination of seabuckthorn juice varieties using visible-nir spectroscopy[J]. *Acta Optica Sinica*, 2010, **30**(2): 574~578
张海红, 张淑娟, 王凤花 等. 应用可见-近红外光谱快速识别沙棘汁品牌[J]. *光学学报*, 2010, **30**(2): 574~578
- 14 Zhu Jing. Fuzzy Control Theory and System Principle[M]. Beijing: China Machine Press, 2005
诸 静. 模糊控制理论与系统原理[M]. 北京: 机械工业出版社, 2005
- 15 Zhao Jun, Chen Jianjun. Improving strategies on fuzzy neural network control for nonlinear object[J]. *Control Theory & Applications*, 2010, **27**(4): 466~472
赵 俊, 陈建军. 非线性系统模糊神经网络控制的改进策略[J]. *控制理论与应用*, 2010, **27**(4): 466~472
- 16 Lin Haijun, Teng Zhaosheng, Yang Jinbao. Error compensation for digital temperature sensor based on RBF neural network ensembles-fuzzy weighing output[J]. *Chin. J. Scientif. Instrum.*, 2011, **32**(7): 1675~1680
林海军, 滕召胜, 杨进宝. 基于 RBF 神经网络集成-模糊加权输出的数字温度传感器误差补偿[J]. *仪器仪表学报*, 2011, **32**(7): 1675~1680
- 17 Sun Zhengqi, Xu Hongbing. Fuzzy-neural network based on T-S model[J]. *Journal of Tsinghua University (Science and Technology)*, 1997, **37**(3): 76~80
孙增圻, 徐红兵. 基于 T-S 模型的模糊神经网络[J]. *清华大学学报(自然科学版)*, 1997, **37**(3): 76~80