

Detection of Fiber Contents of Cotton and Terylene Mixture Textile by Near Infrared Spectroscopy Combined with BP Neural Network*

Li Liu, Li Yan, Yaocheng Xie, Songzhan Li, Jie Xu, Weilin Xu

Institute of Material Science and Engineering, Wuhan Textile University, Wuhan
Email: li_liu2006@yahoo.com.cn

Received: Aug. 15th, 2012; revised: Sep. 7th, 2012; accepted: Sep. 22nd, 2012

Abstract: The prediction of fiber contents of mixture textile by near infrared spectroscopy (NIR) combined with back propagation (BP) neural network was investigated. The near infrared spectrum of samples with different cotton and terylene contents were obtained in the range of 4000 cm^{-1} - $10,000\text{ cm}^{-1}$. Wavelet transform (WT) was used for spectra data de-noise and compression. The correction model of cotton and terylene content based on BP neural network and reconstruction spectral signals was established. The number of hidden neurons, learning rate, momentum and epochs were optimized and decomposition levels of WT was discussed. Data procession was greatly reduced after the spectra signals were compressed by WT. When the compression level and the number of hidden neurons are 3 and 17 respectively, the prediction accuracy is the best. Correlation coefficients (R_p) of prediction set for the correction model of cotton and terylene content both are 0.998, and the root-mean-square error (RMSE) is 1.260% and 1.860% correspondingly. Experimental results have shown that this approach by Fourier transform NIR based on the BP neural network to predict the cotton and terylene content of textile mixture can satisfy the requirement of quantitative analysis and is also suitable to other fiber contents measurement of textile mixture.

Keywords: Textile Mixture; Near Infrared Spectroscopy; Fiber Content; BP Neural Network; Wavelet Transform

近红外光谱结合 BP 神经网络测定棉涤混纺面料的纤维含量*

刘 莉, 颜 丽, 谢尧城, 李颂战, 许 杰, 徐卫林

武汉纺织大学材料科学与工程学院, 武汉
Email: li_liu2006@yahoo.com.cn

收稿日期: 2012 年 8 月 15 日; 修回日期: 2012 年 9 月 7 日; 录用日期: 2012 年 9 月 22 日

摘 要: 通过近红外光谱结合误差反向传播的人工神经网络来检测棉涤混纺面料中纤维含量。测量了 4000 cm^{-1} ~ $10,000\text{ cm}^{-1}$ 范围内棉涤混纺面料样品的近红外吸收光谱。利用小波变换滤波技术对吸收光谱数据进行压缩和去噪处理, 结合滤波后重构光谱信号建立了棉涤混纺面料中棉和涤纶含量的 BP 神经网络校正模型。优化了隐含层神经元的节点数、学习率、动量因子和学习次数。对小波变换中的小波基和压缩尺度进行了详细的讨论。棉涤混纺样品的近红外光谱经过小波压缩, 可以大大降低数据运算量。在小波尺度为 3、隐含层神经元节点数为 17 时, 模型的预测精度最高。所建立的棉和涤纶含量校正模型的预测集相关系数(R_p)均为 0.998, 预测均方根误差为 1.260%和 1.860%。实验结果表明, 应用傅里叶变换近红外光谱和 BP 神经网络技术来预测棉涤混纺面料纤维含量, 可以满足定量分析的要求, 该方法也适合于其他混纺面料纤维含量的快速测定。

关键词: 混纺面料; 近红外光谱; 纤维含量; BP 神经网络; 小波变换

*资助信息: 国家自然科学基金(611108033、51003082)和湖北省自然科学基金(2009CBD295)资助。

1. 引言

纺织品是人类生活的必需品。随着纺织工业迅猛地发展壮大以及人们生活水平的不断提高,纯纺、混纺、交织物品种日益增多,在生产、科研及贸易各领域中对纤维分析的要求也越来越高。纺织品中各种纤维成分的含量,是衡量纺织品品质的重要指标,如何对其精确地检测,在纺织品生产及产品贸易中具有十分重要的意义。传统的纤维成分的定量检测方法主要是化学溶解法,这种方法测试时间长,不适宜大批样本的分析,也不适用于在线监测,而且在操作中要配置一系列的溶剂来溶解纤维,会产生大量有害气体污染工作环境^[1]。

近红外光谱(NIR)是波长在 780~2526 nm 范围内的电磁波,主要是由分子基团振动的倍频和合频吸收产生的,谱带宽、信息量大、吸收弱,利用化学计量学方法解决光谱信息提取和背景干扰,可获得良好的分析检测效果。近红外光谱技术和化学计量学结合,更适合于做定量分析,是一种准确、快速、无损、无污染、低成本的分析技术,已广泛应用于医药、食品、农业、化工、环境监测等领域^[2-8]。近年来在纺织材料和纺织工业中也得到了越来越多的重视和运用^[1,9]。近红外光谱法用于测定纺织品中纤维成分已有报道^[10-12],但大多数需要采用光谱预处理手段对光谱数据预处理,然后采用变量筛选法筛选变量,再建立偏最小二乘法(PLS)模型,步骤繁琐。

本文在测量 51 个棉涤混纺面料样品的近红外吸收光谱基础上,利用小波变换直接对光谱数据进行压缩和消噪,通过误差反向传播的人工神经网络建立了棉涤混纺面料中棉和涤纶含量的预测模型,预测精度较高,模型简单、可靠,适合于混纺面料纤维含量的

快速测定。

2. 材料与方法

利用近红外光谱对棉涤混纺面料纤维含量进行检测,其流程如图 1 所示,包括棉涤样品制备、样品近红外光谱数据采集、光谱数据的预处理、建立纤维含量的 BP 神经网络预测模型以及未知样品纤维含量检测等 5 个步骤。

2.1. 样品的制备

样品由武汉纺织大学纺织材料与科学学院提供,用纯棉、纯涤纶的布料进行粉碎,按照不同的比例混合,制成各种不同棉、涤纶含量的棉-涤混合样品。棉涤样品 51 个,棉含量以 2% 的梯度增加,涤纶含量以 2% 梯度减少,总质量均为 0.05 g,从 51 个样品中随意抽取了编号为 5、8、14、17、25、29、33、36、42、47 的 10 个样品作为预测集,其余 41 个样品用来建立校正模型。另外制备 5 个未知样品用于纤维含量检测。

2.2. 近红外光谱采集

制备的样品在进行光谱采集前,先利用 DHG-9030A 型的上海索谱仪器有限公司生产的电热恒温鼓风干燥烘箱,以 60℃ 的温度烘干,避免纤维含水量过高,影响实验效果。

使用美国 Thermo Nicolet 仪器公司生产的 ANTARIS II 型傅立叶近红外分析仪,光谱范围是 4000~10,000 cm^{-1} ,采用积分球和旋转台测定样品近红外漫反射光谱。采样的分辨率 8 cm^{-1} ,采集点以 10 cm^{-1} 为单位间隔采集光谱数据,光源为 6 W/10V 的卤钨灯,InGaAs 探测器,扫描次数 32 次,均以空气为

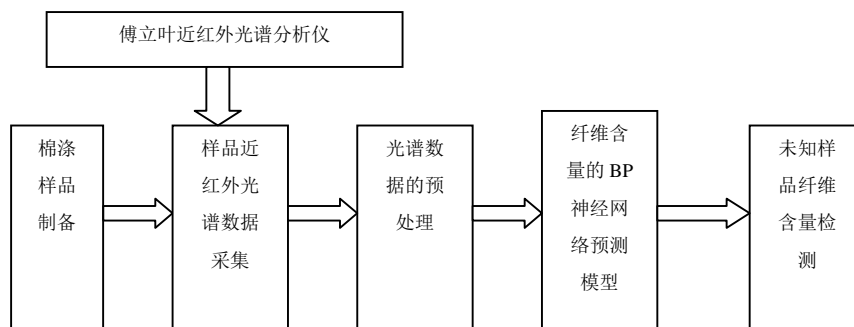


Figure 1. Detection flow of components for textile mixture by near infrared spectrum
图 1. 样品纤维含量的近红外光谱检测流程图

本底。工作环境为温度 23.3℃，相对湿度 70%。漫反射载样器件每次转动 120°，对样品在三个不同的角度各测定 1 次，即每个样品测量 3 次，取 3 次测量数据的平均值。图 2 给出了棉涤混纺面料的近红外光谱。

2.3. 光谱数据的预处理

2.3.1. 光谱波段选择

混合纤维的近红外吸收光谱图包含了 601 个波数点上的光谱数据，倘若对 601 个数据全部进行分析并用于检测，则数据的输入维数很大，导致计算机处理起来很慢。从图 2 中可以看出，混合纤维的近红外光谱吸收特征峰主要集中在波数范围 4000~7200 cm^{-1} 上，而后面的波数段几乎没有特征峰或特征峰不明显，包含的有用信息很微弱。因此，选择波数范围 4000~7200 cm^{-1} 上的光谱数据作为后续对光谱分析的输入数据。

2.3.2. 光谱数据的消噪与压缩

小波分析是当前数学中一个迅速发展的新领域，应用范围十分广泛，可滤波、消噪、压缩、传递等。小波消噪与线性滑动平滑法消噪相比，有着明显的优势，尤其是在对一些信号的细节点或段上进行消噪时，小波消噪的优势更加明显^[13]。由于混纺面料成分的近红外吸收光谱图上有一些基团的特征吸收峰，并且这些特征峰变化比较快或者处在一些细节点或段上。直接利用小波分析，对原始数据进行小波消噪。本文用到 `ddencmp` 和 `wdencmp` 两个函数，`ddencmp` 函数的功能是自动生成小波消噪阈值选择方案，也就是寻找默认值，`wdencmp` 函数的功能是使用小波消噪。对一维信号进行压缩，主要步骤包括：信号的

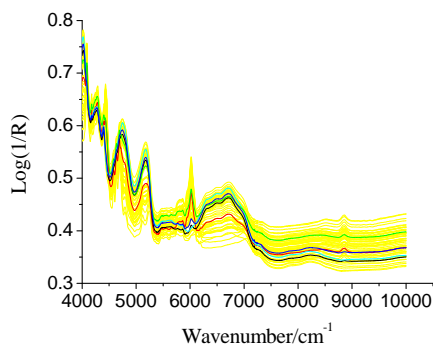


Figure 2. Near infrared spectrum of cotton and terylene textile mixture

图 2. 棉涤混纺面料的近红外光谱图

小波分解；对高频系数进行阈值量化处理；根据量化后的系数，进行一维小波重构。利用 MATLAB 软件中提供的 `wavedec` 和 `appcoef` 两个小波变换函数对光谱数据进行压缩，从而大大减少光谱数据量，提高模型的运算速度。其中，`wavedec` 为多尺度的一维小波分解函数，`appcoef` 为提取一维小波低频系数函数。

2.4. BP (Back Propagation)神经网络模型

近红外光谱分析中，对于测定的物质为非透明液体并且物质成分的百分含量较高时，样品的近红外吸收与其组分的百分含量的关系偏离了郎伯——比尔定律，即样品的红外吸收与其组分的百分含量成非线性关系。因此，线性回归模型在解决此类问题时已不能达到良好的效果，需要构建非线性模型，而神经网络就是一种在解决非线性问题上能够发挥良好作用的模型。神经网络具有很强的学习能力，能够实现输入与输出之间的高度非线性映射^[14]。本文采用的 BP(Back Propagation)神经网络，是误差反向传播学习算法，由信息的正向传播和误差的反向传播两个过程组成。输入层各神经元负责接收来自外界输入信息，并传递给中间层各神经元，中间层可以设计为单隐层或者多隐层结构，最后一个隐层传递到输出层各神经元的的信息，经进一步处理后，完成一次学习的正向传播处理过程，由输出层向外界输出信息处理结果。当实际输出与期望输出不符时，进入误差的反向传播阶段。误差通过输出层，按误差梯度下降的方式修正各层权值，向隐层、输入层逐层反传。周而复始的信息正向传播和误差反向传播过程，是各层权值不断调整的过程，也是神经网络学习训练的过程，此过程一直进行到网络输出的误差减少到可以接受的程度，或者预先设定的学习次数为止。

3. 结果与讨论

3.1. 光谱数据的小波压缩和小波系数选择

在进行小波压缩数据的时候，分解尺度取值越大，压缩后的数据量越小，使得输入数据的维数越小，校正模型的预测速度越快；但同时信号失真的越厉害，使得校正模型的预测精度降低。兼顾到校正模型的预测速度和精度，取分解尺度分别为 3, 4, 5，小波基函数为 `db1`，建立了混纺面料的近红外光谱模型，依

次将 321 个光谱数据压缩为 41、21、11 个，从而降低了校正模型的输入数据维数，便于分析计算。

图 3 为分解尺度为 3 时的光谱信号重建。将图 3 与棉涤混纺面料的近红外光谱图相比较可以看出，采用 41 小波系数重构得到的光谱信息，在很大程度上保留并凸显了特征峰的原有信息，从而不会影响校正模型的预测精度。

3.2. 网络参数的优化

3.2.1. 隐含层神经元节点数的确定

网络以波数范围 $4000\text{ cm}^{-1}\sim 7200\text{ cm}^{-1}$ 的吸光度为输入值，以小波尺度 3 为例，最终网络的输入层节点数为 41 个，输出层的节点数为 2，即棉涤样品的棉和涤纶含量值。隐含节点数的选择比较复杂，当隐含节点数较少时，网络中的权重不充分，误差下降到一定程度后会变化很小，不能够得到好的预测模型；当隐含节点数过多时，不仅网络训练时间长，还会出现过拟合问题，降低神经网络预测功能。本工作主要通过检验来确定，即由预测结果的好坏来确定隐含层神经元节点数目。隐含节点数在 15~22 之间取值，经过网络训练，根据实际得出的均方根误差(RMSE)、平均绝对误差(MAE)值做出折线图如图 4 示。RMSE、MAE 越小，表明棉涤校正模型的校正精度和预测精度越高。经过多次实验，当隐含层点数为 17 时，预测效果较好，精度较高，说明隐含层节点数取输入输出节点数的几何平均数的两倍左右是可行的，也避免神经网络由于隐含层节点数过多而出现拟合现象。小波尺度为 3 时，隐含层节点数取为 17。对于小波尺度为 4，网络模型结构为 21-13-2；小波尺度为 5，网络模型结构为 11-9-2。

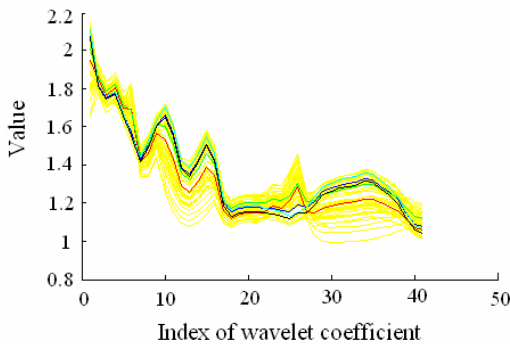


Figure 3. Reconstructed spectral signal
图 3. 重构光谱信号图

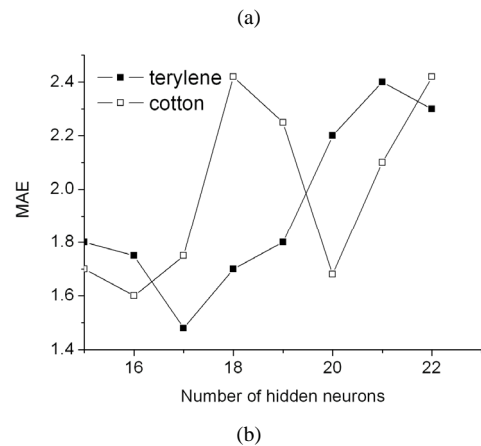
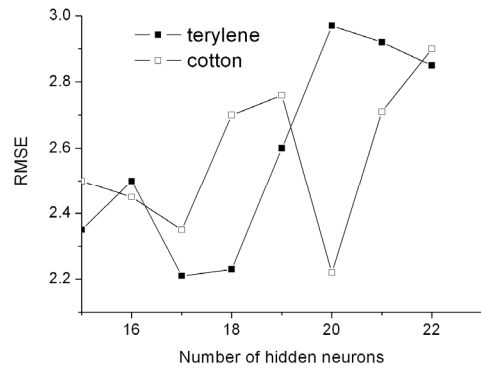


Figure 4. Effect of number of hidden neurons on network: (a) RMSE; (b) MAE
图 4. 隐含节点对训练网络的作用: (a) 均方根误差; (b) 平均误差

3.2.2. 网络训练参数的选取

训练次数和网络的预报能力是考察人工神经网络优劣的主要因素。本研究需要的是能够对未知样本提供可靠预测结果的模型，时又不希望训练次数太多。为此，对 51 个实验样本，取其中的 10 个样本作为预测集用于检验，其余的 41 个样本作为训练集用于建模。学习速率对迭代次数影响较大，它的值一般取 0.01~0.8 之间，大的学习速率可能导致系统的不稳定，而小的学习速率又将导致较长的训练时间，收敛很慢。为提高网络的收敛速度，对动量项进行改进。经过多次调试确定 BP 神经网络的训练参数：学习速率为 0.6，量因子为 0.3，取收敛判据为总体误差小于或等于 0.0009，训练次数为 5000 次。

3.3. 结果与分析

表 1 为小波尺度为 3 预测样品的检验结果。表 2 为小波尺度为 3、4、5 时 BP 神经网络模型(分别记为 WT-ca3-BP、WT-ca4-BP、WT-ca5-BP)的预测结果以

Table 1. Prediction results of validation samples at compression level 3
表 1. 小波尺度 3 时, 预测集的检验结果

| 样本序号 | 白棉 | | | 涤纶 | | |
|------|-------|--------|--------|-------|--------|--------|
| | CV(%) | PV(%) | AE(%) | CV(%) | PV(%) | AE(%) |
| 5 | 8 | 6.808 | 1.191 | 92 | 93.465 | -1.465 |
| 8 | 14 | 13.224 | 0.775 | 86 | 86.815 | -0.815 |
| 14 | 26 | 26.082 | -0.082 | 74 | 73.944 | 0.056 |
| 17 | 32 | 33.858 | -1.858 | 68 | 66.124 | 1.876 |
| 25 | 48 | 44.066 | 3.934 | 52 | 56.094 | -4.094 |
| 29 | 56 | 58.886 | -2.886 | 44 | 41.302 | 2.698 |
| 33 | 64 | 64.032 | -0.032 | 36 | 36.077 | -0.077 |
| 36 | 70 | 69.94 | 0.060 | 30 | 30.133 | -0.133 |
| 42 | 82 | 81.466 | 0.534 | 18 | 18.532 | -0.532 |
| 47 | 92 | 90.747 | 1.253 | 8 | 9.1489 | -1.149 |

Table 2. Comparison of prediction results by BP model and PLS model
表 2. BP 神经网络模型与 PLS 模型预测结果的比较

| 模型 | 白棉 | | | 涤纶 | | |
|-----------|--------|---------|-------|--------|---------|-------|
| | MAE(%) | RMSE(%) | R_p | MAE(%) | RMSE(%) | R_p |
| WT-ca3-BP | 1.260 | 1.860 | 0.998 | 1.289 | 1.886 | 0.998 |
| WT-ca4-BP | 2.159 | 2.655 | 0.997 | 2.423 | 2.906 | 0.996 |
| WT-ca5-BP | 3.194 | 3.740 | 0.995 | 3.224 | 3.752 | 0.995 |
| PLS | 4.730 | 5.988 | 0.981 | 4.682 | 7.033 | 0.980 |

Table 3. Prediction values of unknown samples
表 3. 未知样品的预测值

| 样品序号 | 校正模型: WT-ca3-BP | |
|------|-----------------|----------|
| | 白棉 PV(%) | 涤纶 PV(%) |
| 52 | 36.7232 | 63.2759 |
| 53 | 77.5824 | 22.4182 |
| 54 | 47.2936 | 52.7058 |
| 55 | 93.4871 | 6.5132 |
| 56 | 89.5641 | 10.4364 |

及与 PLS 预测结果的比较。对于小波尺度为 3 的人工神经网络模型, 棉涤验证样品预测后的绝对误差(AE)在 4%以内, 棉的平均绝对误差(MAE)、均方根误差(RMSE)分别为 1.260%和 1.860%, 涤纶的 MAE 和 RMSE 分别为 1.289%和 1.886%, 棉和涤纶的实际值和预测值的相关系数 R_p 均为 0.998, 预测精度最高。比较 BP 人工神经网络模型和 PLS 法, 可以发现, 无论是小波尺度为 3、4 还是 5, 其模型的预测精度均好于 PLS 法, 这可归因于棉涤混纺面料的近红外光谱数据变量个数多, 并与样品组分含量呈非线性关系的缘

故。因此, BP 人工神经网络方法较 PLS 法更适合对棉涤混纺面料纤维成分的定量检测。

模型 WT-ca3-BP 对校正集样品的预测达到了非常好的效果, 所以可用此模型对未知样品进行预测, 棉涤未知样品的预测结果如表 3 所示。

4. 结论

应用近红外光谱技术对棉涤混纺面料纤维含量的检测方法进行了研究。制备了 51 个样品, 随机抽取 10 个样品作为预测集, 其余 41 个样品用来建模。利用小波变换滤波技术对吸收光谱数据进行压缩和去噪处理, 在小波尺度为 3、4、5 情况下, 建立了棉涤混纺面料中棉和涤纶含量的 BP 神经网络预测模型, 优化了网络训练参数, 并将预测结果与 PLS 法进行了比较。与小波变换结合的 BP 神经网络模型具有模型简单稳定、数据量少、运算速度快、预测精度高的特点, 能够实现对混纺面料棉和涤纶含量的快速检测, 该方法也适合于其他混纺面料纤维含量的快速测定。

参考文献 (References)

- [1] 周莹, 徐惠荣, 应义斌. 近红外技术在自然纺织纤维品种鉴别及成分预测上的应用[J]. 光谱学与光谱分析, 2008, 28(12): 2804-2807.
- [2] 王徽蓉, 李卫军, 刘扬阳等. 基于遗传算法与线性鉴别的近红外光谱玉米品种鉴别研究[J]. 光谱学与光谱分析, 2011, 31(3): 669-672.
- [3] 孙德勇, 李云梅, 王桥等. 利用高光谱数据估算太湖水体 CDOM 浓度的神经网络模型[J]. 武汉大学学报·信息科学版, 2009, 4(7): 851-855.
- [4] T. A. Lestander, P. Geladi. NIR spectral information used to predict water content of pine seeds from multivariate calibration. *Canadian Journal of Forest Research*, 2005, 35(5): 1139-1248.
- [5] V. A. Saptari, T. K. Youcef and J. Zhang. NIR measurement of glucose in synthetic biological solutions using high-throughput angle tuned filter spectrometer. *The International Society for Optical Engineering*, 2004, 5325(1): 1-10.
- [6] 单扬, 朱向荣, 许青松等. 近红外光谱结合小波变换——径向基神经网络用于奶粉蛋白质与脂肪含量的测定[J]. 红外与毫米波学报, 2010, 29(2): 128-131.
- [7] 刘炜, 常庆瑞, 郭曼, 邢东兴等. 夏玉米可见/近红外小波主成分提取与氮素含量神经网络检测[J]. 红外与毫米波学报, 2011, 30(1): 48-54.
- [8] 张愿, 张录达, 白琪林等. 近红外光谱法快速无损识别普通、高油、超高油玉米籽粒[J]. 光谱学与光谱分析, 2009, 29(3): 686-689.
- [9] 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, 8(6): 1260-1263.
- [10] 陈斌, 崔广, 金尚忠等. 近红外光谱在快速检测棉制品中含棉量的应用[J]. 江苏大学学报(自然科学版), 2007, 28(3): 185-188.
- [11] 冯红年, 甘彬, 金尚忠. 棉涤混合纺织面料含量的近红外光谱检测[J]. 激光与红外, 2005, 35(10): 768-770.
- [12] 陈斌, 王小天, 倪凯. 相关分析法在 NIR 快速检测纺织原料真丝含量中的应用[J]. 光谱仪器与分析, 2006, Z1: 52-56.
- [13] 飞思科技产品研发中心. 小波分析理论与 MATLAB R2007 实现[M]. 北京: 电子工业出版社, 2005: 81.
- [14] 葛哲学, 孙志强. 神经网络理论与 MATLAB R2007 实现[M]. 北京: 电子工业出版社, 2007: 46.