

污染气体红外光谱特征的快速提取与识别

刘美娟, 冯巍巍, 史丰荣, 王学勤, 张 骏*

烟台大学光电信息学院, 山东 烟台 264005

摘 要 利用小波变换的多尺度分析对污染气体红外光谱数据进行处理, 并使用神经网络对红外光谱数据进行分类识别。实验结果表明: 小波变换与神经网络的有机结合, 有利于污染气体红外光谱的快速特征提取和识别, 并具有较高的识别率, 是一个有效的识别系统。

主题词 小波变换; 神经网络; 光谱识别

中图分类号: O657.3 **文献标识码:** A

文章编号: 1000-0593(2006)10-1854-04

引 言

光谱是表征物质物理属性的主要手段, 利用光谱特征可实现污染气体的遥感监测与快速报警^[1]。由于受到各种干扰的影响, 要从获取的微弱信号中提取目标的光谱特征是十分困难的, 因此研究人员尝试了不同的方法来减小这种影响或是增强目标的光谱特征^[2-5]。1989年, Mallat等在前人大量工作的基础上提出多尺度分析的概念和基于多尺度分析的小波基构造方法, 将小波正交基的构造纳入统一的框架之中, 使小波分析成为一种实用的信号分析工具^[6]。与傅里叶变换相似, 小波变换是一种同时具有时-频二维分辨率的变换。其优于傅氏变换之处在于它具有时域和频域“变焦距”特性, 十分有利于信号的精细分析^[7]。

小波分析以其优美的数学理论、高的分辨质量、卓有成效的应用实例为许多领域的科技工作者所信服, 因而被广泛应用于模式识别、特征提取、弱信号检测和机械故障诊断等领域。而人工神经网络是模仿真实人脑神经网络的结构和功能, 用若干基本特性的某种理论抽象、简化和模拟而形成的一种信息处理系统。由于神经网络是由大量神经元通过极其丰富和完善的连接而构成的自适应非线性动态系统^[8], 因此利用人工神经网络来实现特征的分类与识别越来越多地用于模式识别的研究中^[3, 4]。

建立快速、高效光谱识别器的困难在于: 污染气体红外光谱特征的分布非常复杂^[9]。由于光谱数据全部在野外实地获取, 各种干扰(如烟雾、尘土等)的影响和周围环境(如大气、温度等)的变化使得原本弱小的目标特性更加复杂和微弱, 有的特征淹没在噪声之中, 这给目标光谱的分类和识别

造成了很大困难。为了更有效地对远距离污染气体红外光谱进行监测与快速识别, 我们将小波变换中多尺度分析和神经网络的特点有机地结合起来, 发展了一套小波变换和神经网络组成的污染气体红外光谱特征快速提取与识别系统。实验结果表明: 该系统不仅有效地去除干扰物的影响, 增强了目标的光谱特征, 提高了整机系统的识别率, 同时为污染气体红外光谱多目标识别系统的研究奠定了基础。

1 小波方法的理论基础

小波变换是把信号 $x(t)$ 表示为一簇函数的加权和, 而这簇函数由基本小波 $\Psi(t)$ 经过伸缩和平移而形成的。伸缩尺度为 a 、时间平移为 b 的小波 $\Psi_{a,b}$ 定义为

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \Psi\left(\frac{t-b}{a}\right), \quad a \neq 0 \quad (1)$$

连续小波变换 $W_x(a,b)$ 定义为

$$W_x(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} x(t) \Psi\left(\frac{t-b}{a}\right) dt \quad (2)$$

将连续小波离散化, 设 $a = 2^j$, $b = k2^j$, 则二进离散小波为

$$\Psi_{j,k}(t) = 2^{-\frac{j}{2}} \Psi(2^{-j}t - k) \quad (3)$$

在此基础上, Mallat^[6]给出了二进小波变换的快速算法, 即 Mallat 算法。

信号的分解:

$$\begin{aligned} c_{j,k} &= \sum_n h_{n-2k} c_{j-1,n} \\ d_{j,k} &= \sum_n g_{n-2k} c_{j-1,n} \end{aligned}, \quad j = 1, 2, \dots, J \quad (4)$$

信号的合成

收稿日期: 2005-07-08, 修订日期: 2005-10-18

基金项目: 国家自然科学基金项目(69807003)和教育部高等学校骨干教师项目资助

作者简介: 刘美娟, 女, 1979年生, 烟台大学光电信息学院在读研究生 * 通讯联系人

$$c_{j-1,k} = \sum [k_{h-2n}c_{j,n} + g_{h-2n}d_{j,n}], j = J, J-1, \dots, 1$$

(5)

其中 $c_{j,k}$ 是原始信号的平滑信号, $d_{j,k}$ 是细节信号, g_n 是与小波函数相关的带通滤波器的脉冲响应, h_n 是与尺度函数相关的低通滤波器的脉冲响应。

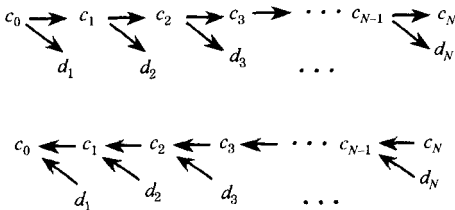


Fig 1 Sketch map Mallat arithmetic

从图 1 所示的信号分解与合成可知, 通过 Mallat 算法, 原始信号 $c_0(n)$ 可以分解为某一尺度下的平滑信号和细节信号。通过小波变换, 可以很好地去噪声的影响, 提取污染气体红外光谱特征, 再利用神经网络的模式识别手段来实现污染气体红外光谱特征的快速识别和报警。

2 BP 神经网络识别系统

1986 年 Rumelhart 等完整而简明地提出一种误差反向传播训练算法, 简称 BP 算法^[8]。误差反传算法的主要思想是把学习过程分为两个阶段: 第一阶段(正向传播过程), 给出输入信息通过输入层经隐含层逐层处理并计算每个单元的实际输出值; 第二阶段(反向过程), 若在输出层未能得到期望的输出值, 则逐层递归的计算实际输出与期望输出之差(即误差), 以便根据此差调节权值。本文中的 BP 网络由输入层、输出层和一个隐含层构成(见图 2)。

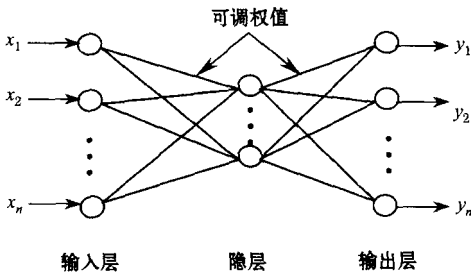


Fig 2 Diagram of a BP neural network

假设输入层有 N 个节点, 输出层有 M 个节点, 隐含层节点若干。第 P 个训练样本 $X_p = [x_{p1}, x_{p2}, \dots, x_{pN}]$, 对应的理想输出为 $D_p = [d_{p1}, d_{p2}, \dots, d_{pM}]$ 。网络中当前层第 j 个神经元为 u , 其输入为 s_{pj}

$$s_{pj} = \sum_i \omega_{ji} o_{pi} + \theta_j$$

(6)

其中 o_{pi} 是上一层第 i 个神经元的输出, ω_{ji} 是上一层第 i 个神经元的输出与本层第 j 个神经元的输入之间的连接权。神经元 u_j 的输出 o_{pj} 为

$$o_{pj} = f(s_{pj})$$

(7)

这里 $f()$ 是非线性 S 型函数:

$$f(x) = \frac{1}{1 + e^{-x}}$$

(8)

误差反向传播:

$$\text{隐层 } \delta_{pj} = o_{pj} (1 - o_{pj}) \sum_k \delta_{pk} \omega_{jk}$$

(9)

$$\text{输出层 } \delta_{pj} = o_{pj} (1 - o_{pj}) (d_{pj} - o_{pj})$$

(10)

修改权和阈值

$$\omega_{ji}(t+1) = \omega_{ji}(t) + \mu \delta_{ji} o_{pi}$$

(11)

3 实验结果与分析

将收集到的红外光谱分为训练和测试两大序列。训练序列共有 420 条光谱数据。样本的类型包括: 目标云团 (DMMP)、尘土、煤灰以及 DMMP 和尘土干扰物的混合谱。为了检验识别系统的识别率, 测试序列除了包含训练序列所列的样品种类外, 还加入了训练序列中没有的样本。如 DMMP 和干扰物的混合谱。本文中神经网络的输出层设有四个节点, 分别代表 DMMP、尘土、煤灰和其他。在理想条件下, 四个节点对应的输出值如表 1 所示的假设。

Table 1 Hypothesis for four nodes in the ideal condition

样本类型	输出节点 1	输出节点 2	输出节点 3	输出节点 4
DMMP	1	0	0	0
尘土	0	1	0	0
煤灰	0	0	1	0
其他	0	0	0	1

污染气体 DMMP 在 8~12 μm 有 3 个特征峰^[1] (见图 3 和图 4)。图 3(a) 中所示的原始光谱由于各种干扰的存在, 很难直观地看出 DMMP 的 3 个特征峰。图 3(b) 是经过 FFT 滤波的结果, 由于快速 Fourier 变换不具有局域变换的特性, 图中仍不能很好地将 DMMP 的特征显现出来。图 3(c) 是经过小波多尺度分析后给出的结果, 图中结果显示: 多尺度分析方法不仅有效去除了噪声的影响, 而且对增强 DMMP 的 3 个特征峰非常有利, 这一处理为光谱特征的准确提取提供了保障。

图 4(a) 给出的是一条尘土和 DMMP 的混合光谱。由于尘土的成份比较复杂, 加之尘土对 DMMP 红外光谱具有一定的屏蔽作用, 使得原本比较微弱的 DMMP 的光谱特征就更加的小。虽经过快速 Fourier 滤波, 但光谱的特征还是比较小, 如图 4(b) 所示。采用小波多尺度变换处理后, DMMP 的 3 个特征峰值能显现出来, 这对光谱特征的提取是非常有利的。但图中仍保留了许多与 DMMP 光谱特征无关的小峰值。

表 2 列出了小波处理后的数据和其他方法处理的数据, 由表 2 可以看出, 虽然训练次数相同, 但是使用小波方法对污染气体红外光谱进行处理后, 由于光谱特征得到增强, 神经网络能识别参与训练的光谱数据, 如表中的 DMMP、尘土和煤灰等, 且识别率高; 而用其他方法处理的数据经神经网络系统识别, 虽然能识别 DMMP, 但同时对于煤灰和尘土等干扰物产生误判, 如表中的尘土, 神经网络将它错判成 DMMP, 其结果将直接影响监测报警工作。为了检验所建立

的特征提取与识别系统的识别率,使用了多个未参加训练的光谱作为测试样品。测试的结果表明:该系统对含有 DMMP 的光谱非常敏感,实验的结果进一步证实了多尺度分析与神

经网络的有机结合,对污染气体红外光谱特征的快速提取非常有效,有利于该系统对污染气体的快速识别。

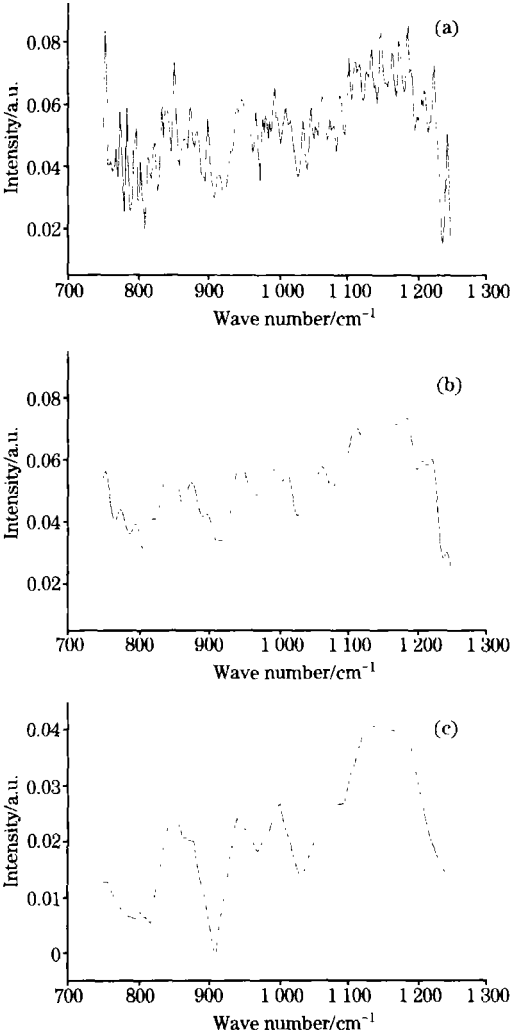


Fig. 3 Infrared spectra of target
(a): DMMP; (b): FFT 滤波; (c): 小波分解

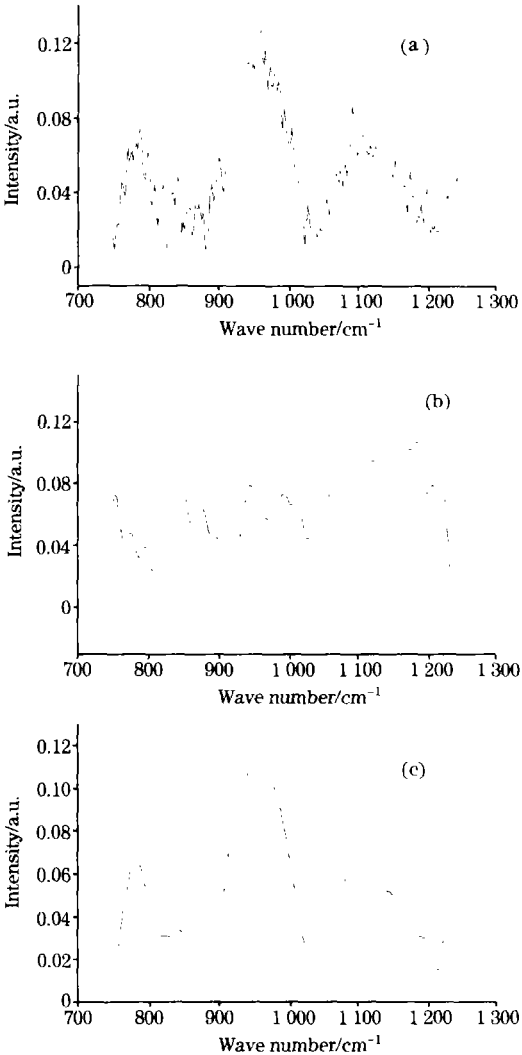


Fig. 4 Infrared spectra of target and dust
(a): DMMP+尘土; (b): FFT 滤波; (c): 小波分解

Table 2 Capability comparison of two method (training times= 10 000)

数据类型	本类型	分类输出结果样			
		节点 1	节点 2	节点 3	节点 4
小波处理后的数据	DMMP	1 000 000	0 000 000	0 000 000	0 000 000
	尘土	0 000 476	0 994 718	0 000 005	0 000 000
	煤灰	0 000 018	0 000 661	0 998 902	0 000 146
	DMMP+ 尘土	0 996 284	0 000 708	0 000 000	0 000 000
	其他	0 000 000	0 000 041	0 000 331	0 997 986
	测试样品	0 999 974	0 000 148	0 000 000	0 000 000
其他方法处理的数据	DMMP	0 999 985	0 000 000	0 000 000	0 000 004
	尘土(误判)	0 999 976	0 000 000 2	0 000 024	0 000 000
	煤灰(误判)	0 002 609	0 000 064	0 284 446	0 347 763
	DMMP+ 尘土	1 000 000	0 000 000	0 000 000	0 000 000
	其他	0 000 138	0 000 001	0 000 032	0 992 694

4 结 论

效的滤除了噪声的影响,增强了光谱特征,提高了识别率和运行,有利于污染气体红外光谱的快速特征提取和识别。利用该系统可以实现污染气体红外光谱的多目标识别。

结果表明,小波多尺度分析与神经网络有机的结合,有

参 考 文 献

- [1] Hoffland L D, Piffath R J, Bouck J B. Opt. Eng., 1985, 24(6): 982.
- [2] ZHANG Jun, XUN Yulong(张 骏, 荀毓龙). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 1998, 18(6): 649.
- [3] Mayfield H T, Eastwood D, Burggraf L W. Chemical and Biological Sensing, 2000, 4036: 54.
- [4] Tanabe K, Tamura T, Uesaka H. Applied Spectroscopy, 1992, 46: 807.
- [5] NIE Liang, ZHANG Jun(聂 亮, 张 骏). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2004, 24(8): 953.
- [6] Mallat S G. IEEE Trans. on Pattern Analysis Machine Intelligence, 1989, 11(7): 674.
- [7] Daubechies I. Ten lecture on Wavelets. Philadelphia: Capital City Press, 1992.
- [8] YU AN Zeng-ren(袁曾任). Artificial Neural Network and Its Application(人工神经网络及其应用). Beijing: Tsinghua University Press (北京: 清华大学出版社), 1999. 66.
- [9] ZHANG Jun, et al(张 骏, 等). J. of Infrared and Millimeter Waves(红外与毫米波学报), 1997, 16(6): 463.

Fast Algorithm for Feature Extraction and Identification of Infrared Spectra of Polluted Gases

LIU Mei-juan, FENG Wei-wei, SHI Feng-rong, WANG Xue-qin, ZHANG Jun*

Institute of Science and Technology for Optoelectronic Information, Yantai University, Yantai 264005, China

Abstract With the multi resolution analysis, features of infrared spectra of polluted gases were extracted. Then the data were trained or identified by a neural network system. The experimental results show that the combination of the wavelet transform and the neural network has the great ability of feature extracting. And the system is quite efficient for identifying infrared spectra.

Keywords Wavelet transform; Neural network; Spectral identification

(Received Jul. 8, 2005; accepted Oct. 18, 2005)

* Corresponding author