

降低 DOAS 系统探测限的新型反演算法研究

麻金继^{1,2} 李素文²

(¹ 安徽师范大学物理与电子信息学院, 芜湖 安徽 241000; ² 中国科学院安徽光学精密机械研究所, 合肥 安徽 230031)

摘要 差分吸收光谱法(DOAS)是一种高灵敏测量大气痕量气体成分含量的有效的光学遥感方法,该方法基于最小二乘拟合模型,利用获得的痕量气体的差分吸收光学密度与标准的吸收截面进行拟合,反演待测气体的浓度。建立了基于径向基(RBF)神经网络的痕量气体浓度反演的新模型,对网络的隐层参数采用改进最近邻聚类学习算法训练,对输出层权值的训练采用梯度下降算法,使得网络收敛快,能更好地实时、在线反演测量光谱。并针对 DOAS 技术的特点,把拟合残差输入网络集中训练,使得 RBF 网络在反演真实痕量气体吸收时,效果更佳。实验结果表明该新型反演方法提高了 DOAS 系统的反演精度,降低了 DOAS 系统的探测限。

关键词 差分吸收光谱; 反演; 径向基函数神经网络; 最小二乘; 探测限

中图分类号 O433.4 文献标识码 A doi: 10.3788/AOS20092909.2351

Modified Retrieval Algorithm on Decreasing Detection Limits in Differential Optical Absorption Spectroscopy

Ma Jinji^{1,2} Li Suwen²

(¹ College of Physics and Electronic Information, Anhui Normal University, Wuhu, Anhui 241000, China)
(² Anhui Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Hefei, Anhui 230031, China)

Abstract Differential optical absorption spectroscopy (DOAS) has become a widely used method to measure trace gases in the atmosphere. It can identify trace gases through narrow-band molecular absorption. Concentration of trace gases is retrieved using method of least-squares fits of reference spectra to the measurement spectra. A novel retrieval method based on radial basis function (RBF) neural network was developed to retrieve the concentration of trace gases in DOAS system. The coefficient of the hidden layer was trained by modified nearest neighbor clustering algorithm, and that of the output layer was trained by gradient descent algorithm. These result in a fast speed of convergence of network. At last, there was comparison between the novel retrieval method and the conventional least-squares fitting. The experimental results show that the reliability and accuracy of DOAS are improved and detection limits are decreased by using the novel retrieval method.

Key words differential optical absorption spectroscopy; radial-basis-function neural network; retrieval; least squares; detection limits

1 引 言

差分吸收光谱(DOAS)技术已成为大气污染监测的常用方法之一,DOAS 方法具有一些传统监测方法所无法比拟的优点。DOAS 系统的监测范围广,测量结果比点测量仪器更具有代表性;该方法采用非接触方式,没有器壁损失;在线、实时测量,时间分辨率高;探测限低,可以监测浓度(体积分数)为 10^{-9} 量级,甚至 10^{-12} 量级的痕量气体^[1~4]。

痕量气体吸收遵从 Beer-Lambert 定律^[1,2],其吸收截面可写为窄带(快变)和宽带(慢变)部分。而在通常的 DOAS 系统数据处理中,一般是先用多项式拟合来去除慢变的部分(包括气体的宽带吸收瑞利散射和米氏散射等),剩下的光谱表达式经过低通去噪,再与痕量气体的标准吸收截面进行最小二乘(LSM)拟合,得到待测气体的浓度^[3~6]。

LSM 是使用最多和效果较好的一种反演方法,

收稿日期: 2009-01-15; 收到修改稿日期: 2009-03-13

基金项目: 国家自然科学基金(40701132)、教育部科学技术研究重点项目(209057)、安徽省自然科学基金(070412042)、安徽高校自然科学基金(KJ2008A114)和安徽师范大学博士启动基金资助课题。

作者简介: 麻金继(1975—),男,博士,副教授,主要从事遥感信息处理方面的研究。E-mail: jinji_ma@yahoo.com.cn

但是由于大气中各种气体之间复杂的非线性关系,以及光源的波动,CCD分辨率等干扰因素存在^[1,2],影响了反演精度。近年来,随着人工神经网络模型的成熟和相关理论的完善,神经网络在分类、非线性模型识别方面已经被证明是有效的方法^[7~10]。特别是神经网络法的自学习和非线性逼近能力,使得网络映射函数具有很强的柔软性,可以任意精度逼近任意形状的函数^[11~13]。径向基函数(RBF)网络是以函数逼近理论为基础而构造的一类前向网络,这类网络的学习等价于在多维空间中寻找训练数据的最佳拟合平面,比多项式函数更为灵活,能有效解决的拟合精度问题。本文针对差分吸收光谱技术的特点,建立了基于RBF神经网络来反演大气痕量气体浓度的新型反演模型。

2 RBF 网络

RBF网络是以函数逼近理论为基础而构造的一类前向网络,这类网络的学习等价于在多维空间中寻找训练数据的最佳拟合平面。RBF网络的每个隐层神经元的函数都构成了拟合平面的一个基函数,网络也因此得名。RBF网络是一种局部逼近网络,即对于输入空间的某一个局部区域只存在少数的神经元用于决定网络的输出,因而将它应用于复杂序列预测会取得很好的效果。特别是用RBF网络来逼近非线性连续函数具有最佳逼近性质。它适合于多变量函数的逼近,只要中心选择得当,只需很少的神经元就可获得很好的逼近效果^[8~10]。

RBF神经网络的工作原理分为2个阶段:1)学习阶段,选定充足和高质量的训练样本训练RBF神经网络,学习结果以权值的形式存储在网络结构之中;2)工作阶段,当测试样本输入RBF神经网络时,训练好且具有一定泛化性的网络将进行内插和外推等方式自适应完成特征匹配过程,给出客观的评价结果。

2.1 RBF 网络结构

RBF网络由三层组成,其结构如图1所示。输入层节点只传递输入信号到隐层,隐层节点由像高斯函数那样的辐射状RBF构成,而输出节点是简单的线性函数。最常用的基函数是高斯函数:

$$R_i(x) = \exp\left(-\frac{\|x - c_i\|^2}{2\sigma_i^2}\right), \quad i = 1, 2, \dots, K \quad (1)$$

式中 x 是 n 维输入向量, c_i 是第 i 个基函数的中心, σ_i 是第 i 个感知的变量,它决定了该基函数中心点的宽度; I 是感知单元的个数, $\|x - c_i\|$ 表示 x 和 c_i 之

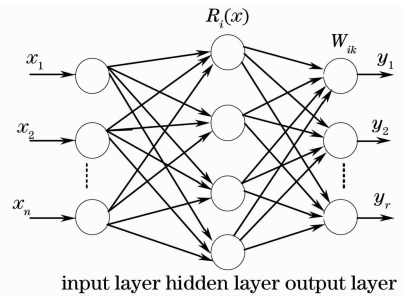


图1 RBF神经网络

Fig. 1 RBF neural networks

间的欧几里得距离。

输入层实现从 $x \rightarrow R_i(x)$ 的非线性映射,输出层实现从 $R_i(x) \rightarrow y_k$ 的线性映射,即

$$y_k = \sum_{i=1}^I w_{ik} R_i(x), \quad k = 1, 2, \dots, r \quad (2)$$

式中 r 是输出节点数,输出层权值为 w_{ik} 。为了提高RBF神经网络的收敛速度,将隐含层参数 c_i, σ_i 和输出层权值 w_{ik} 分开进行训练。对隐含层参数训练采用一种新的聚类算法,即改进最近邻聚类学习算法,此算法具有学习时间短、计算量小、网络性能优良等优点,对输出层权值的训练采用梯度下降算法。

2.2 改进最近邻聚类算法过程

1)在训练样本中,共有 I 个条件;采用计数器统计属于各类的样本个数。

2)对第1个条件的所有特征参数向量,进行自适应最近邻聚类,设聚成 L_1 类,且第 l_1 个计数器的值最大,则令第 l_1 个聚类中心为 $c_1, l_1 \in [1, 2, \dots, L_1]$,计算第1个条件的所有特征参数向量与 c_1 的范数,则令其中最大的范数为 σ_1 。

3)对第 i 个条件的所有特征参数向量,进行自适应最近邻聚类,设聚成 L_i 类,且第 l_i 个计数器的值最大,则令第 l_i 个聚类中心为 $c_i, l_i \in [1, 2, \dots, L_i]$,计算第 i 个条件的所有特征参数向量与 c_i 的范数,则令其中最大的范数为 σ_i 其中 $i = 2, 3, \dots, I$ 。

4)最后得到的向量 (c_1, c_2, \dots, c_m) 及 $(\sigma_1, \sigma_2, \dots, \sigma_m)$,就是RBF神经网络隐含层参数 c_i 和 $\sigma_i, i = 1, 2, \dots, I$ 。

2.3 梯度下降算法过程

1)给 w_{ik} 赋随机初值, $k = 1, 2, \dots, r$;由改进最近邻聚类算法得到隐含层参数 c_i 和 σ_i 以及采用(1)式计算神经网络的输出 y 。

2)计算输出误差为

$$e_n = y_{\text{real}} - y_k = y_{\text{real}} - \sum_{i=1}^m w_{ik} R_i(x); \quad (3)$$

3)定义目标函数为

$$E(n) = \frac{1}{2} e_n^2; \quad (4)$$

4) 在 $n + 1$ 时刻, 输出层权值按如下的规则更新, 其中 α 为训练系数

$$w_{ik}(n+1) = w_{ik}(n) + \alpha \frac{\partial E(n)}{\partial w_{ik}}. \quad (5)$$

3 RBF 网络在 DOAS 中的应用

DOAS 技术是利用差分吸收的思想, 首先把测量的原始大气谱经过初步处理得到 DOAS, 然后基于 LSM 把 DOAS 和其中含有的待测痕量气体的标准截面进行拟合获得待测气体的浓度^[14~17]。在 300~360 nm 波段内, 有 SO₂, HCHO, NO₂ 和 O₃ 等气体的吸收在里面, 把获得的 DOAS 和 SO₂, HCHO, NO₂ 和 O₃ 的标准截面进行 LSM 拟合, 其拟合谱图和残差如图 2 所示。

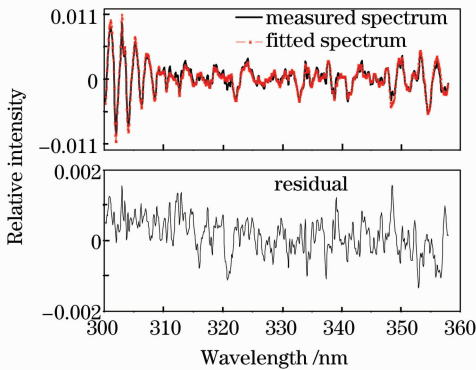


图 2 基于 LSM 的谱图拟合和残差

Fig. 2 Fitting and residual based on LSM

本文研究利用 RBF 网络来反演痕量气体的浓度。网络的隐含层参数训练采用改进最近邻聚类学习算法, 对输出层权值的训练采用梯度下降算法。选择波段从 300~360 nm, 通道从 50~950 研究其效果, 训练时不仅把 SO₂, HCHO, NO₂, O₃ 标准吸收截面参与训练, 还把各种干扰最终体现的残差一同训练网络, 其中为了使选取的残差具有代表性, 选取实际多条大气拟合残差的统计均值。网络输出为 SO₂, HCHO, NO₂ 和 O₃ 这四种气体的浓度, 其拟合谱图和残差如图 3 所示。为了有一个已知的标准, 采用已知浓度、不同比例的 SO₂, HCHO, NO₂, O₃ 的差分吸收和残差构成不同模拟光谱, 其中每条模拟光谱叠加的残差选自不同真实大气吸收谱拟合后的残余部分, 其中 20 条用来训练, 20 条用来测试网络, 来检验新方法的效果。图 2、图 3 中是一个例子, 参考实际大气中可能的含量, 所取 SO₂, NO₂, HCHO, O₃ 的体积分数分别为 25×10^{-9} , 15×10^{-9} , 30×10^{-9} 。

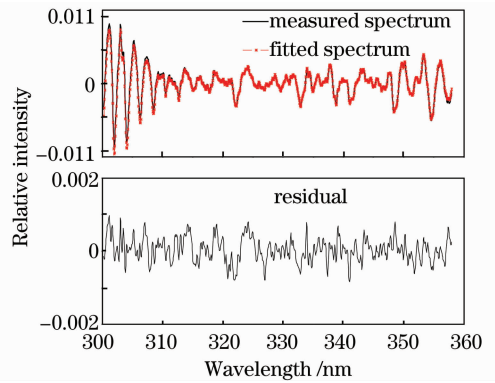


图 3 基于 RBF 网络的谱图拟合和残差

Fig. 3 Fitting and residual based on RBF

从图 2, 图 3 可以发现, 基于 RBF 网络的拟合效果较好, 残差也较小。表 1 为基于 RBF 网络的反演的四种测试气体产生的平均相对误差, 以及采用 LSM 拟合反演产生相对误差。对于 SO₂ 因为吸收强度大, 所以两种方法都得到了满意效果, 但是当吸收截面较小时, 尤其对于 O₃ 来说, 可以清晰发现基于 RBF 网络的反演方法好于传统的 LSM 反演的误差。由于在训练网络时, 把各种可能残差同时参与训练, 所以网络在识别弱信号时, 效果很好。从表 1 可以清晰发现基于 RBF 网络的反演方法好于传统的 LSM 反演, 该方法提高了系统的反演精度, 降低了系统误差, 因而也降低的 DOAS 系统的探测限。

表 1 平均相对误差比较

Table 1 Comparisons of average relative errors

Trace gases	Average relative errors of RBF / %	Average relative errors of LSM / %
SO ₂	0.3	0.3
NO ₂	2.4	3.3
HCHO	3.0	5.7
O ₃	3.2	8.3

4 结 论

本文提出基于 RBF 网络模型来反演大气痕量气体的浓度, 首先对网络的隐层参数采用改进最近邻聚类学习算法训练, 对输出层权值的训练采用梯度下降算法, 使得网络收敛快, 以任意设定的精度逼近待反演光谱。另外考虑到 DOAS 技术的特点, 把大气中和系统中各影响因素的最终体现—残差加入训练, 使得 RBF 网络在反演真实大气时, 效果更好, 模拟实验表明采用新型反演算法提高了 DOAS 系统的反演精度, 降低了反演误差, 因此也达到降低探测限的目标。

参 考 文 献

- 1 Jochen Stutz, Ulrich Platt. Numerical analysis and estimation of the statistical error of differential optical absorption spectroscopy measurement with least-squares methods [J]. *Appl. Opt.*, 1996, **35**(30): 6041~6053
- 2 Martin Hausmann, Uwe Brandenburger, Theo Brauers *et al.*. Simple monte carlo methods to estimate the spectra evaluation error in differential optical absorption spectroscopy [J]. *Appl. Opt.*, 1999, **38**(3): 462~475
- 3 M. Hausmann, D. Ritz, U. Platt. A new coaxial "Long-Path-DOAS"-system; First application to BrO measurement in the arctic troposphere [C]. *SPIE*, 1992, **1715**: 341~352
- 4 Hao Nan, Zhou Bin, Chen Limin. Measurement of nitrous acid and retrieval of aerosol parameters with differential optical absorption spectroscopy [J]. *Acta Physica Sinica*, 2006, **55**(3): 1529~1534
郝楠, 周斌, 陈立民. 利用差分吸收光谱法测量亚硝酸和反演气溶胶参数[J]. *物理学报*, 2006, **55**(3): 1529~1534
- 5 A. Geyer. *The Role of the Nitrate Radical in the Boundary Layer: Observation and Modeling Studies* [D]. *Institute of Environmental Physics, University of Heidelberg*, 2000
- 6 Li Suwen, Xie Pinhua, Liu Wenqing *et al.*. A study of applicability of light emitting diodes in differential optical absorption spectroscopy measurements [J]. *Acta Physica Sinica*, 2008, **57**(3): 646~650
李素文, 谢品华, 刘文清等. 发光二极管在差分吸收光谱系统中的应用研究[J]. *物理学报*, 2008, **57**(3): 646~650
- 7 Cathy Clerbaux, Juliette Hadji-Lazaro, Sébastien Payan *et al.*. Retrieval of CO from nadir remote-sensing measurements in the infrared by use of four different inversion algorithms [J]. *Appl. Opt.*, 2002, **41**(33): 7068~7078
- 8 Mauro Forti. Some extensions of a new method to analyze complete stability of neural networks [J]. *IEEE Trans. Neural Netw.*, 2002, **13**(5): 1230~1238
- 9 Liang XueBin, Wang Jun. A recurrent neural network for nonlinear optimization with a continuously differentiable objective function and bound constraints [J]. *IEEE Trans. Neural Netw.*, 2000, **11**(6): 1251~1262
- 10 S. Chen, C. F. Cowan, P. M. Grant. Orthogonal least squares learning algorithm for radial basis function networks [J]. *IEEE Trans. Neural Netw.*, 1991, **2**(2): 302~309
- 11 Ryad Zemouri, Daniel Racocanu, Noureddine Zerhouni. Recurrent radial basis function network for time-series prediction [J]. *Engineering Applications of Artificial Intelligence*, 2003, **16**: 453~463
- 12 Ewa Skubalska-RafajLowicz. Random projection RBF nets for multidimensional density estimation [J]. *Int. J. Appl. Math. Comput. Sci.*, 2008, **18**(4): 455~464
- 13 Qi Feng, Liu Wenqing, Zhou Bin *et al.*. Improving DOAS system measurement precision with artificial neural network method [J]. *Acta Optica Sinica*, 2002, **22**(11): 1345~1349
齐锋, 刘文清, 周斌等. 利用人工神经网络方法提高差分光学吸收光谱系统测量精度研究 [J]. *光学学报*, 2002, **22**(11): 1345~1349
- 14 Xie Pinhua, Liu Jianguo, Qin Min *et al.*. Retrieval of monocyclic aromatic hydrocarbons with differential optical absorption spectroscopy [J]. *Spectroscopy and Spectral Analysis*, 2006, **26**(9): 1584~1588
谢品华, 刘建国, 秦敏等. 差分吸收光谱方法反演大气环境单环芳香烃有机物 [J]. *光谱学与光谱分析*, 2006, **26**(9): 1584~1588
- 15 Li Suwen, Xie Pinhua, Li Yujin *et al.*. A method of data processing based on wavelet transform in differential optical absorption spectroscopy [J]. *Acta Optica Sinica*, 2006, **26**(11): 1601~1604
李素文, 谢品华, 李玉金等. 基于小波变换的差分吸收光谱数据处理方法 [J]. *光学学报*, 2006, **26**(11): 1601~1604
- 16 Si Fuqi, Liu Jianguo, Xie Pinhua *et al.*. Determination of aerosol optical thickness and atmospheric visibility by differential optical absorption spectroscopy [J]. *Acta Optica Sinica*, 2006, **26**(7): 961~964
司福祺, 刘建国, 谢品华等. 差分吸收光谱技术监测气溶胶光学厚度及大气能见度的研究 [J]. *光学学报*, 2006, **26**(7): 961~964
- 17 Peng Fumin, Xie Pinhua, Zhang Yinghua. Effect of spectral resolution on measurement of trace gases in atmosphere by differential optical absorption spectroscopy [J]. *Acta Optica Sinica*, 2008, **28**(9): 1643~1648
彭夫敏, 谢品华, 张英华. 分辨率对大气中痕量污染气体的 DOAS 测量性能影响研究 [J]. *光学学报*, 2008, **28**(9): 1643~1648