

# 基于希尔伯特-黄变换的激光云高仪后向 散射信号去噪方法

何俊峰 刘文清 张玉钧 陈臻懿 阮俊 王立明

(中国科学院安徽光学精密机械研究所环境光学与技术重点实验室, 安徽 合肥 230031)

**摘要** 基于希尔伯特-黄变换(HHT)和 Savitzky-Golay 滤波,提出了一种结合 HHT 时频分析的半导体激光云高仪后向散射信号去噪方法。该方法在对云高仪后向散射信号的噪声特性及传统去噪方法的缺陷进行研究的基础上,选择截止分量阶一定的自适应时变时空滤波器组(TFB)结构并辅助 Savitzky-Golay 滤波抑制噪声,最后利用 HHT 时频分析辅助云层参数及垂直能见度的反演并验证去噪效果。经大量不同天气条件下的实验数据验证及反演结果对比,发现该方法在保留有用信息特征、有效抑制各种平稳或非平稳噪声的基础上,能够帮助云底云高识别及能见度反演,并且降低了云高漏报率、提高了云底云高的细节识别能力和能见度反演的精度。

**关键词** 大气光学;希尔伯特-黄变换;激光云高仪;后向散射信号;Savitzky-Golay 滤波

**中图分类号** TN957.52 **文献标识码** A **doi**: 10.3788/AOS201131.0201001

## A Denoising Method for Backscatter Signal of Laser Ceilometer Based on Hilbert-Huang Transform

He Junfeng Liu Wenqing Zhang Yujun Chen Zhenyi Ruan Jun Wang Liming

(Key Laboral of Environmental Optics and Technology, Anhui Institute of Optics and Fine Mechanics,  
Chinese Academy of Science, Hefei, Anhui 230031, China)

**Abstract** A denoising method combined with Hilbert-Huang transform (HHT) time-frequency analysis for backscatter signal of laser diode ceilometer based on HHT and Savitzky-Golay filtering is proposed. Considering characteristics of ceilometer backscatter signal proper structure of adaptive time-variant time-space filtering bank (TFB) assisted by Savitzky-Golay filtering is selected to suppress noise. Then, HHT time-frequency analysis is used to observe signal features and test denoising effect. It is proved by plenty of test data under various weather conditions that the method can help to suppress noise without deteriorating signal features and improve the ability of cloud base and height identification, and finally obtain precise retrieving of vertical visibility.

**Key words** atmospheric optics; Hilbert-Huang transform; laser ceilometer; backscatter signal; Savitzky-Golay filtering

**OCIS codes** 010.0010; 140.0140; 280.0280; 290.0290

## 1 引 言

云在天气系统发展、降水形成和大气辐射传输等物理过程中扮演着极其重要的角色。正确观测、分析云的变化,准确、及时地获取云的信息,是了解

认识大气物理状况,掌握天气变化规律的一个重要因素。进一步发展和完善云的宏、微观特征遥感系统,倍受世界各国的关注。随着光电技术的发展,基于激光技术的云层信息主动光学探测技术水平得到

**收稿日期**: 2010-04-25; **收到修改稿日期**: 2010-06-21

**基金项目**: 公益性行业(气象)科研专项基金(GYHY200706023)资助课题。

**作者简介**: 何俊峰(1980—),男,博士研究生,主要从事测云激光雷达系统及信号处理等方面的研究。

E-mail: hejunfeng@aiofm.ac.cn

**导师简介**: 刘文清(1954—),男,研究员,博士生导师,主要从事环境光学特性、有害痕量气体光学与光谱学监测方法、新型环境监测技术等方面的研究。E-mail: wqliu@aiofm.ac.cn

本文电子版彩色效果详见中国光学期刊网(<http://www.opticsjournal.net>)

了很大的提高,基于米氏散射理论的激光雷达目前已经得到广泛应用<sup>[1~3]</sup>。

米氏散射是激光辐射与大气相互作用的一种非常重要的物理过程。米氏散射激光云高仪是激光雷达的一种,其能够根据米氏散射原理探测气溶胶粒子的垂直分布,进而反演出与气溶胶相关的各种大气参数,如气溶胶后向散射系数和消光系数、云层信息、能见度信息等。但是由于云高仪系统内外存在的各种噪声影响,其性能受到很大制约,激光云高仪、尤其是功率较低的半导体激光云高仪后向散射信号的去噪一直是个该领域的难题。传统的信号处理方法,大多是基于线性和稳定的假设<sup>[4]</sup>。只是在最近几年才引入了新的分析非平稳和非线性数据的方法<sup>[5]</sup>,例如,小波分析和 Wagner-Ville 分布可处理线性非平稳数据。另外,也有些为非线性平稳的确定性系统而设计的非线性时间序列分析方法。但是,在大多数自然或人为的实际系统中,信号都像激光云高仪获取的后向散射信号、最有可能既非线性又非平稳,分析这样的系统数据一直是个难题。直到黄锷等提出著名的希尔伯特-黄变换(HHT),才为类似激光后向散射信号的去噪问题提供了比较有效的解决方案。

本文研究基于 HHT 变换和 Savitzky-Golay 滤波的激光云高仪后向散射信号噪声抑制方法,该方法在 HHT 变换过程中引入 Savitzky-Golay 滤波,有效利用 Savitzky-Golay 滤波和 HHT 对非线性非平稳信号处理的优势,经验证对激光云高仪的后向散射信号去噪及特征提取有明显效果。

## 2 半导体激光云高仪工作原理与传统去噪方法

基于米氏散射原理的半导体激光云高仪是光、机、电及计算机控制等技术的集成,常规的半导体激光云高仪一般包含激光发射系统、光学接收系统、激光功率监测系统及控制与信号处理系统四大部分,研究采用的激光云高仪系统 Vaisala CL31 的结构及性能参数分别如图 1 和表 1 所示。

半导体激光器向空中发射一激光脉冲,该激光脉冲在向上传播的过程中,不断与大气中的气溶胶粒子相互作用。当激光从大气进入云雾时,在云雾中、尤其是云的边界处将会产生较强的后向散射,散射信号的大小与激光功率、云底高度、云雾的类型和细节、大气特性和激光云高仪性能参数有关,这样,

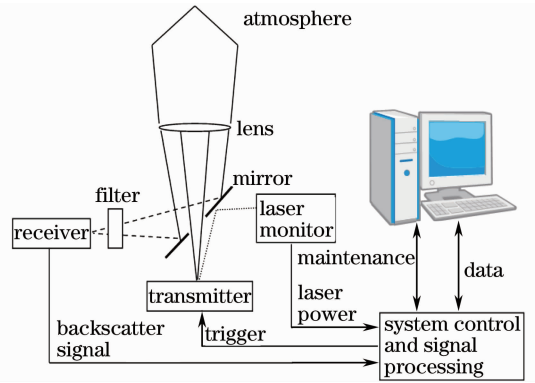


图 1 半导体激光测云雷达系统

Fig. 1 Ceilometer system scheme

表 1 Vaisala CL31 半导体激光云高仪性能参数

Table 1 Ceilometer Vaisala CL31 parameters

Parameters	Value
Beam divergence /mrad	$\pm 0.4 \times \pm 0.7$
Field-of-view divergence /mrad	0.83
Effective Lens diameter /mm	96
Optics focus length /mm	300
Measurement range /m	7500
Range resolution /m	5 or 10
Report interval /s	2~20
Laser type	InGaAs diode
Laser wavelength /nm	905
Operating mode	pulsed
Pulse properties	110 ns, 1.2 $\mu$ J/pulse
Repetition rate /kHz	10
Detector type	APD
Receiver bandwidth /MHz	30

通过分析激光回波的变化,可以判别云底位置、计算云高和能见度、分析大气结构的细节。一旦向上传播的激光脉冲进入接收望远镜的视场,则激光与大气物质相互作用所产生的光回波信号将被接收望远镜接收,经信号检测和处理后即可得到激光回波信号。一方面,激光雷达的发射光和回波光在大气中来回传播将受到大气的衰减;另一方面,激光雷达的回波信号强度还与激光雷达配置和技术应用方面的因素有关。在单波长的条件下,米氏散射型激光雷达方程如下<sup>[3]</sup>:

$$P(z) = LP_0 \frac{M}{z^2} \Delta z \beta(z) \exp \left[ -2 \int_0^z \sigma(z) dz \right], \quad (1)$$

将(1)式中右边的分母项  $z^2$  移到左边,可得

$$X(z) = P(z) \cdot z^2 = LP_0 M \Delta z \beta(z) \times \exp \left[ - \int_0^z 2\sigma(z) dz \right] = C \beta(z) \exp \left[ -2 \int_0^z \sigma(z) dz \right], \quad (2)$$

式中  $X(z)$  称为雷达回波距离修正信号,  $C = LP_0M\Delta z$  只与系统本身有关(包括激光雷达的校正常数  $L$ , 激光发射功率  $P_0$ , 接收望远镜面积  $M$ , 探测的高度分辨  $\Delta z$ ), 称为系统常数;  $\beta(z)$  和  $\sigma(z)$  分别为距离  $z$  处大气总的后向散射系数和消光系数, 可以通过此雷达方程反演得到。

在总体上激光雷达信号呈现与距离平方反比的衰减趋势; 而在局部, 则由于大气不稳定性以及噪声(光电探测器的各种噪声以及天空背景辐射等)的影响引起一定的随机不可预测的起伏, 因此该系统是比较典型的非线性非平稳系统。上述云底云高、垂直能见度及云雾细节等信息都可以或强或弱在激光后向散射信号的特征中得到反映。所有测量过程都会引入噪声, 累积和滑动平均等数字滤波是消除各种平稳噪声的常用方法。

基于安全、经济和系统小型化考虑, 很多云高仪采用功率较低的半导体激光器, 探测器直接接收到的后向散射信号中的有用信息基本被环境光背景和电噪声淹没, 而且这些噪声强度往往难以预测。累积平均的特点就是让激光器发射大量的周期性脉冲、在接收端对回波信号按照该周期进行分段累加。假定对被测信号  $P(z)$  [由有用信号  $S(z)$  和零均值、标准差为  $\sigma$  的高斯白噪声  $n(z)$  组成] 做  $m$  次同步累积, 那么  $x(t)$  的累积平均结果为

$$\hat{P}(z) = \frac{1}{m} \sum_{k=0}^{m-1} P(z+kZ) = \frac{1}{m} \sum_{k=0}^{m-1} S(z+kZ) + \frac{1}{m} \sum_{k=0}^{m-1} n(z+kZ) = S(z) + \frac{\sqrt{\sigma^2}}{\sqrt{m}}, \quad (3)$$

式中  $Z$  为脉冲周期与光速的乘积。这样有用信号会因累加而增强, 而其中的随机噪声则会被削弱, 累积平均后的信号信噪比将得到  $\sqrt{m}$  倍提高。但是由于天气条件随机变化(例如, 云层移动)等因素, 信号信噪比的提高受到限制, 往往难以满足云层和能见度信息反演的需要。而且, 累积平均的还有一个明显缺陷, 为了获得较高的信噪比往往需要采集大量的数据样本、运算量大、影响测量的实时性。Vaisala CL31 采用脉冲工作方式, 激光脉冲宽度 110 ns、重复频率 10 kHz, 接收端对后向散射信号进行了 16384 次累积平均, 提高了信噪比。但是数据分析结果显示, 累积平均后反演得到的后向散射系数廓线中噪声背景仍然较强, 影响各种信息的提取。

滑动平均滤波就是相当于有一个固定长度为  $m$  的滑动窗口沿离散时间序列滑动, 每滑动一个采样间隔, 窗口前面进入一个新的数据, 窗口后面去掉一

个旧的数据, 这样在窗口中始终有  $m$  个“最新”的数据。只要每次在滑动后把窗口中的  $L$  个数据进行加权算术平均, 就可得到一组经过滑动平均滤波的新序列:

$$\hat{P}(z) = \frac{1}{m} \sum_{k=0}^{m-1} c_k \cdot P(z - k \cdot \Delta Z), \quad (4)$$

式中  $\Delta Z$  为采样周期,  $c_k$  为各采样点权重。合理设置窗口宽度, 滑动平均滤波能有效地提高脉冲信号的信噪比, 但这种提高是有限度的, 对于非常微弱的信号还需要结合其他方式, 才能更加有效地提取出有用信号。滑动平均通常会削平输入信号的极值点, 而且对于包含脉冲的信号, 滑动平均也会使脉冲宽度展宽。在激光雷达等应用中, 激光的后向散射信号中的脉冲幅度和宽度都是包含重要物理含义的参数, 要求在信号去噪过程中尽量保留这些信号特征。根据滑动平均的实现原理, 其结果的零阶矩和一阶矩上特点都可以被保留, 而二阶矩特点将被忽视。可见滑动平均忽略了相邻数据样点之间的相关性, 其获得高信噪比的代价是弱化了信号细节。

总之, 对于激光后向散射信号去噪, 只采用传统的滤波器往往收效甚微, 为了抑制噪声, 有时甚至会造成后向散射信号中有用的窄脉冲被严重削弱、展宽, 不利于云层信息等的获取。

### 3 HHT 原理及基于 HHT 的滤波器组

任何数学方法能够用以表述非线性和非平稳数据的必要条件是要有一个自适应的基。自适应性说明基的定义必须有数据依赖性、是后验定义的, 这与傅里叶变换等传统数据分析方法完全不同。HHT 由两部分组成: 经验模态分解(EMD)和希尔伯特谱分析(HSA), 这种方法对于非线性和非平稳数据分析是潜在可行的, 尤其是时频能量描述, 已得到全面测试和验证<sup>[6]</sup>。在应用案例中, HHT 给出的结果比用于时频能量描述传统分析方法的结果更加准确, 而且能够揭示许多被分析数据的真实物理意义。有实际意义的描述类似系统的方式可以从瞬时频率着手, 最简单的计算瞬时频率的方法是使用希尔伯特变换。不过, 直到黄锷提出经验模态分解方法后, 希尔伯特变换的真正优势才越来越明显。根据 HHT 的数据分析特点, 本文将研究其对激光云高仪后向散射信号去噪及特征提取的具体方案, 使其在保留包含真实物理意义的数据特征信息的前提下抑制噪声。

### 3.1 EMD 方法筛分过程

正如黄锷等所探讨,经验模态分解法非常适用于非平稳和非线性过程的数据处理<sup>[4]</sup>。相较于传统方法,这种新方法是自适应的、且更为直观和直接,其后验基通过分解方法基于数据、取自于数据。整个经验模态分解过程基于任何数据由不同的、简单的固有模态振荡组成的假设。每个固有模态,无论线性或非线性,都是一个简单的振荡过程,有同样数目的零点和极值点。所以把数据分解成本征函数并分析其有特定物理含义的瞬时频率,必须先定义一个必要条件:函数对称于局部零均值,且有相同的极值和过零点。

最后原始数据可以表示为<sup>[4]</sup>

$$X(t) = \sum_{j=1}^n C_j(t) + R_n(t). \quad (5)$$

将一个非稳态数据分解成  $n$  个本征函数或称为经验模态  $C_n(t)$  和一个趋势项  $R_n(t)$ , 其中  $C_n(t)$  或残留量  $R_n(t)$  比预先确定的有意义的值小,或残留量  $R_n(t)$  成为单调函数、从中没法再提取新的本征函数(IMF)。由于每一个 IMF 分量是代表一组特征尺度的数据序列,因此“筛分”过程实际上是将原始信号序列分解为各种不同特征波形的叠加。需要说明的是,每一个 IMF 分量既可以是线性也可以是非线性的。数据按经验模态分解到  $n$  重,并得到残留量  $R_n(t)$ ,  $R_n(t)$  可以有平均量的趋势或是个常量。IMF 拥有一个简单的振荡模式,但它显得更为一般:不同于一个简单的谐波成分中恒定的幅度和频率,IMF 有一个可变幅度和频率的时间函数。由于 EMD 的特征尺度由物理数据定义,其包含的各个成分通常具有物理意义,这已经得到证明,最近所做的大量研究,进一步确立了 IMF 的统计学意义<sup>[6~8]</sup>。因此,IMF 成分中是否包含重要的信息、或者仅仅包含噪声,可以得到检验。

### 3.2 希尔伯特谱分析

在取得 IMF 各个分量后,可以很容易对每一个分量应用希尔伯特变换、并计算瞬时频率。对每个 IMF 分量执行希尔伯特变换之后,原始数据可以表示为以下实部形式:

$$x(t) = \text{Re} \left\{ \sum_{j=1}^n a_j(t) \exp \left[ i \int \omega_j(t) dt \right] \right\}, \quad (6)$$

式中,残留量  $R_n(t)$  通常被有意从希尔伯特谱分析过程中排除,因为  $R_n(t)$  是一个单调函数或常量。虽然希尔伯特变换可以将单调趋势看作较长周期振荡的一部分,残留量中包含的能量表示较大的一个

平均偏移量。考虑到对较长趋势的不确定性,而且为了获取其他有清晰振荡的低能量分量中包含的信息,最后的非 IMF 分量可以被排除在外。不过,如果得到了实际物理因素的认可,也可以将  $R_n(t)$  包括在内。实际上,激光云高仪后向散射信号中筛分出的  $R_n(t)$  就包含了反映大气变化趋势的物理意义,根据需要可以保留。(5)式以时间函数的形式给出了每个分量的幅度和频率。同样的数据如果用傅里叶展开,将是

$$x(t) = \text{Re} \left\{ \sum_{j=1}^n a_j \exp[i\omega_j(t)t] \right\}, \quad (7)$$

式中  $a_j$  和  $\omega_j$  为常量。(5)式和(6)式之间的对比清晰:IMF 表示一个广义傅里叶展开。可变的幅度和瞬时频率,不仅大大提高了展开的效率,而且也使展开方法适应非线性和非平稳数据。随着 IMF 展开,幅度和频率调制也得以清楚地分离。因此,HHT 克服了傅里叶展开固定幅度和频率的限制,改用可变振幅和频率描述数据。这种时频的振幅分布被称为希尔伯特谱  $H(\omega, t)$ 。如果幅度平方是更首选的来表示能量密度的方法,那么幅度平方也可以用于取代伯尔布特能量谱。利用已有的伯尔希特谱定义,还可以定义边际谱  $h(\omega)$  为

$$h(\omega) = \int_0^T H(\omega, t) dt. \quad (8)$$

边际谱可以描述每个频率值上的总幅度(或能量)分布。该谱从概率意义上表示整个数据跨度内幅度的积累。HHT 没有必要用杂散谐波来表示非线性波形失真;对于基于先验基的卷积对,HHT 也没有时间或频率分辨率不确定性原理限制,能在时频能空间给出用于特征提取的结果。

### 3.3 基于 HHT 的滤波器组

正如传统的滤波器参数选择基于频域考虑、滤除有用信号频带外的频率分量,基于 HHT 的滤波器可以根据信号特征对 IMF 进行取舍。Flandrin<sup>[5]</sup> 提出了应用 EMD 构造一种滤波器组的思想。该滤波器的特点是其通带以及截止频率都是随着输入信号特点的变化而改变、是自适应的,这和传统滤波器的固定截止频率有所不同。当然这也就意味着,基于 HHT 的滤波器的参数选择需要遵循一定的准则。假设残留量为最后一级 IMF,定义自适应滤波器的截止分量阶  $l, k$ ,则可以构造一种自适应的、时变滤波器组(TFB)<sup>[7~9]</sup>,表示如下:

$$x(t) = \sum_{j=l}^k C_j, \quad (9)$$

式中  $l, k \in [1, N]$ ,  $N$  为包括残留项的 IMF 总数,  $x(t)$  为滤波器组输出信号。通常情况下, 当  $l > 1$ ,  $k = N$  时, TFB 是低通滤波器;  $l = 1, k < N$  时, TFB 是高通滤波器;  $1 < l \leq k < N$  时, TFB 是带通滤波器; 当  $1 < k < l < N$ , TFB 为带阻滤波器。

直接应用(9)式对信号进行滤波处理经验证是可行的<sup>[1,7~9]</sup>, 郑发泰等<sup>[1]</sup>提出了基于 EMD 的激光雷达大气后向散射信号降噪信号处理方法, 经验证, 有一定去噪效果。但是文献中没有论及该滤波方法的滤波器结构及其对去噪性能的影响, 没有将 HSA 应用于辅助去噪性能分析和后向散射信号特征提取。另外, 经大量数据处理验证, 直接应用(9)式存在以下问题: 一方面某些 IMF 分量中可能同时含有信号成分和噪声成分, 简单地用原信号减去经 EMD 分解的一个或多个 IMF 分量而实现滤波, 这样可能导致将所减去的 IMF 分量所包含的有用信息与噪声一起被滤除; 另一方面 EMD 分解过程完全自适应, 不能确定 IMF 分量数, 且噪声成分具体分布在哪些 IMF 分量(即  $k$  和  $l$  的确定)与具体信号相关。因此, 不仅需要确立 TFB 结构, 而且需要根据信号特点选择其他辅助滤波方法。

## 4 结合 Savitzky-Golay 的 TFB 滤波过程

滑动平均和 Savitzky-Golay 滤波都是有限脉冲响应(FIR)数字滤波的特例, 它们对消除电路中存在的高斯白噪声效果比较明显。滑动平均用递归方式对一定数量的数据采样点取平均值; Savitzky-Golay 滤波对每一数据点的一个邻域(长度固定的窗口)内各点的数据, 用一个高阶多项式拟合, 此多项式的系数可根据最小二乘法准则使拟合误差最小来确定, 由此得出滑动窗口内中心点的最佳拟合值, 即为降噪处理后的值<sup>[10]</sup>。滑动数据窗口依次沿着每一点滑动, 从而实现了平滑处理。

Savitzky-Golay 滤波滑动窗口内的权重函数由已知  $n$  阶多项式的未加权线性最小二乘拟合所产生, 设  $f_i = f(t_i)$  表示数据序列中的第  $i$  个点, 其中  $t_i = t_0 + i\Delta$ ,  $\Delta$  是数据点的采样间隔, 则有

$$g_i = \sum_{n=-n_L}^{n_R} c_n f_{i+n}, \quad (10)$$

式中  $f_i$  是输入信号,  $g_i$  是滤波后输出信号, 是计算过程中数据采样点左边所要用的点数, 是计算过程中数据采样点右边所要用的点数。上式表明  $g_i$

是  $f_i$  及其相邻点的线性组合, 如果式中  $c_n = 1/(n_L + n_R + 1)$  是常数, 上式就是滑动平均。若系数  $c_n$  是常数或随时间线性变化, 滤波结果中将不能引入偏移, 平均窗口一端幅度较大采样点会被窗口另一端幅度较小的采样点平衡; 若描述  $c_n$  的函数的二阶导数非零, 滤波结果中将引入偏移。Savitzky-Golay 滤波算法的出发点是寻找能够保留高阶矩特点的系数  $c_n$ , 也就是说, 其思想是在滑动窗口内用高阶多项式(研究采用 4 阶)而不是常数来逼近原信号。对每个数据点  $f_i$ , 在宽度为  $n_L + n_R + 1$  的滑动窗口内的所有相邻点进行最小二乘多项式拟合, 取多项式在  $i$  处的值作为  $g_i$ , 同时舍弃此多项式上的其他点; 然后对下一个数据点  $f_{i+1}$  在新的窗口内再次使用最小二乘拟合。如果每一个数据点都要重复使用最小二乘拟合, 那么整个滤波过程将显得十分繁琐; 好在最小二乘拟合涉及的只是一个线性的矩阵变换, 拟合多项式的系数与已知数据点本身呈线性关系。可以利用 QR 分解法实现最小二乘问题的求解, 确定 Savitzky-Golay 系数<sup>[10]</sup>。

Savitzky-Golay 滤波假设不同远近的数据点之间存在特定大小的冗余, 该冗余的特定属性表明数据能够用多项式来实现局部拟合, 从而在不造成分辨率损失的前提下抑制噪声, 而且能够保留脉冲峰值和宽度等数据特征。滑动平均也可以有效抑制噪声, 但是会将窄脉冲信号幅度降低、脉宽展宽。通常, 多项式阶数越高(高于 4 阶)、拟合点数越多, Savitzky-Golay 滤波的效果越好、特征保留能力越强<sup>[11]</sup>。研究中的 Savitzky-Golay 滤波采用四阶多项式进行对称的 33 样点拟合, 拟合结果如图 2 所示。

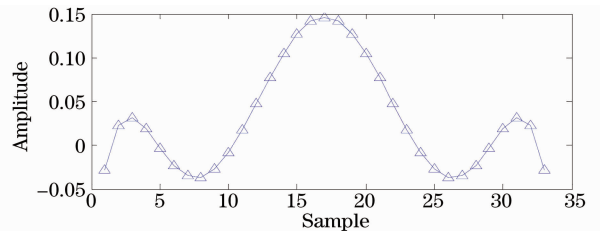


图 2 33 点 Savitzky-Golay 系数的四阶多项式拟合  
Fig. 2 4-order polynomial fit of 33 sample Savitzky-Golay coefficients

结合 Savitzky-Golay 的 TFB 滤波过程如下:

- 1) 对云高仪获取的后向散射系数廓线做 EMD;
- 2) 根据信号特点选择 TFB 滤波结构, 初步剔除高频非平稳噪声;
- 3) 对 HHT 滤波后信号做 Savitzky-Golay 滤波, 保留信号特征、进一步抑制平稳噪声;

4) 计算滤波后信号的 HHT 时频分布和边际谱, 并辅助常规反演方法提取信号特征。

图 3 演示了对 Vaisala CL31 云高仪所获取的后向散射系数廓线进行 EMD 的分解过程。经过筛选获得的从高频到低频有序排列的若干 IMF 体现了自适应的滤波特性。由于后向散射信号中含有高

斯白噪声(距离归一化前), 分解后得到的前一两个高频 IMF 分量通常为噪声, 而其中残留项一般可以表示后向散射信号的均值或趋势; 对比测量结果, 也发现其他每个 IMF 分量都具有明显的物理意义且包含了一定范围的特征尺度, 这一点可以根据 HHT 时频分布加以验证。

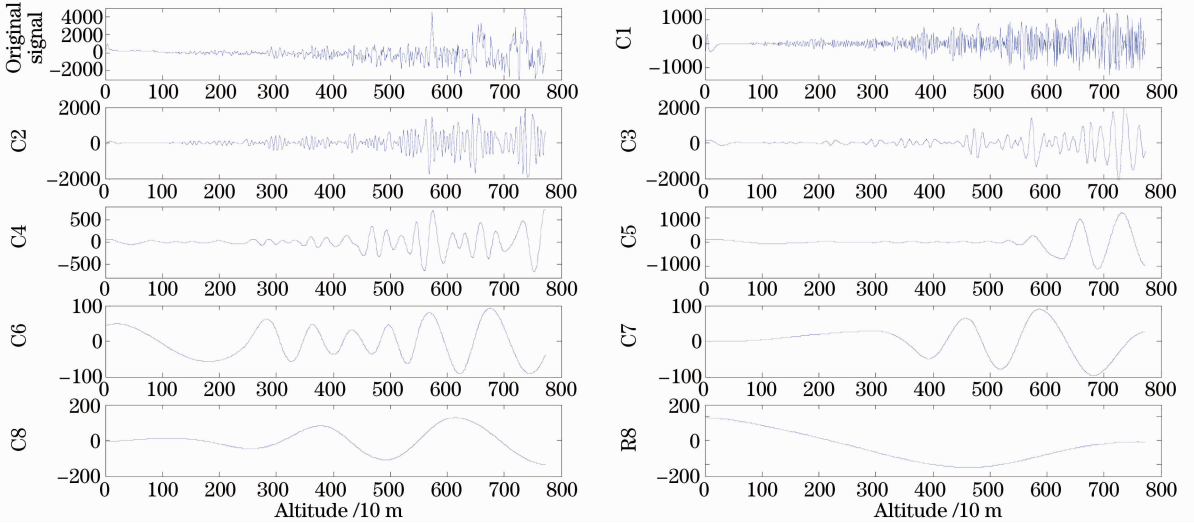


图 3 激光云高仪后向散射系数廓线的 EMD 过程

Fig. 3 EMD process of laser ceilometer backscatter coefficient profile

图 4 显示了不同截止分量阶  $l, k$  情况下 TFB 的滤波效果(采用的数据与图 3 中所采用的数据相同), 其中标志 SG 表明在 TFB 的基础上应用了 Savitzky-Golay 滤波。经过数据对比可见, 大气中的大尺度变化和小尺度变化分别在高阶 IMF 分量和低阶 IMF 分量中得到反映; 采用连续均方误差准则(CMSE)<sup>[9]</sup> 往往需要剔除较多低阶 IMF 分量, 这时可以获得较高的信噪比, 但是由于激光云高仪后

向散射信号的许多信息细节包含在高频低阶 IMF 分量中, 反而不利于许多有用信息的分辨。如图中当  $l > 3$ , 滤波之后的数据明显丢失了大量小尺度细节, 不利于云底云高的分辨; 当  $k < 9$ , 滤波之后的数据会丢失大尺度细节, 不利于能见度的反演。如图 4 中第二行, 如果在选择  $l \leq 2$  的基础上再应用 Savitzky-Golay 滤波, 不仅信噪比得到大幅提高, 而且后向散射信号的细节得到保留。

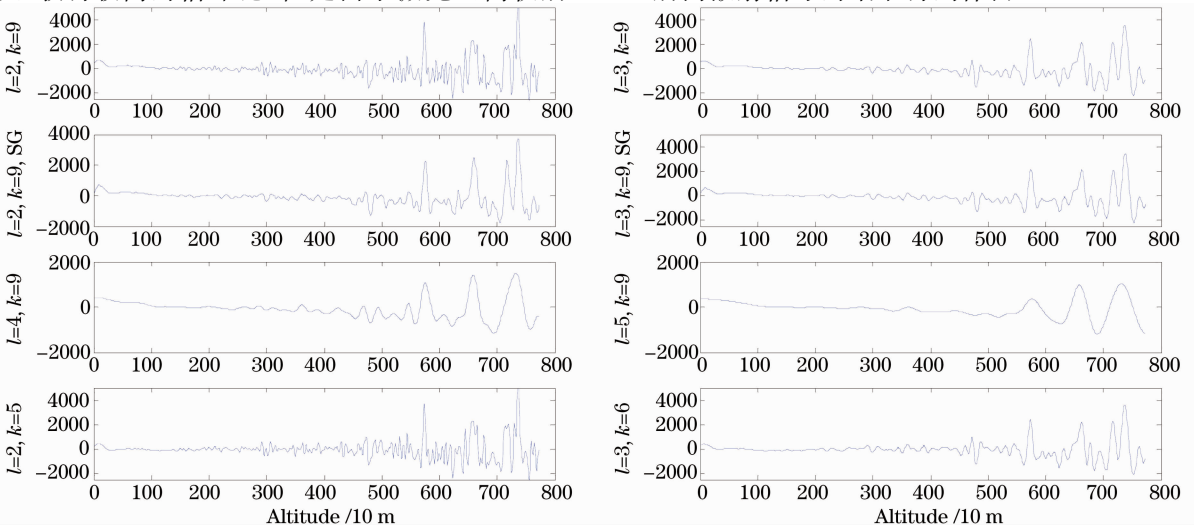


图 4 不同截止分量阶  $l, k$  情况下 TFB 的滤波效果

Fig. 4 TFB filtering effect of different  $l, k$

根据一般时频分布过程,后向散射信号中不同的信号特征一般对应不同的频率分量。时频分布中瞬时频率较高的点反映对应位置的后向散射信号变化较快的情况,如云底;幅度较高的点反映对应位置的后向散射信号幅度同样较强,如云高;频率较低且幅度变化平滑的区域对应后向散射信号中的变化趋势,如云厚或雾。可以根据后向散射信号的以上时频特点,进一步利用 HHT 时频分布对滤波前后的后向散射信号进行观察、比较,一方面,可以比较滤波效果,另一方面,可以据此提取后向散射信号中的信号特征。

### 5 激光云高仪后向散射信号去噪结果分析

大气中云层信息和垂直能见度等参数的获取将依据激光云高仪后向散射系数廓线中的不同特征,而这些信号特征在不同的天气条件下也相差甚远。根据经验,比较常见的单层云的后向散射信号特征相对容易提取;但是当遇到多层云且云层之间距离

较近、云层处于背景噪声较高的远距离或云层较厚且需要观察厚云层中的细节时,传统滤波方法处理能力显得捉襟见肘;而大气垂直能见度的获取不仅要求保留信号特征,还要求尽量提高信噪比(噪声导致的大幅抖动会影响能见度判决精度)。本文提出的方法与传统的信号处理方法相比不仅能够有效抑制随机噪声,而且能够在不破坏原始信号时频特性的前提下保留有用的信号特征。为了验证 HHT 滤波去噪效果,利用该方法对不同天气条件下的激光云高仪后向散射信号做了处理,并按照前节提出的滤波过程对不同天气条件下的实现结果加以比较。

下面先对激光云高仪在不同天气条件下获取的单帧后向散射系数廓线利用 HHT 进行滤波处理和时频分析,并以此为依据比较该方法滤波及特征提取的性能。这些数据样本取自于中国科学院安徽光学精密机械研究所 2007~2009 年间在合肥市西郊董铺岛利用 Vaisala CL31 半导体激光云高仪进行长期连续监测得到的原始数据,具有一般性。

图5中浅色背景为原始后向散射系数廓线,深

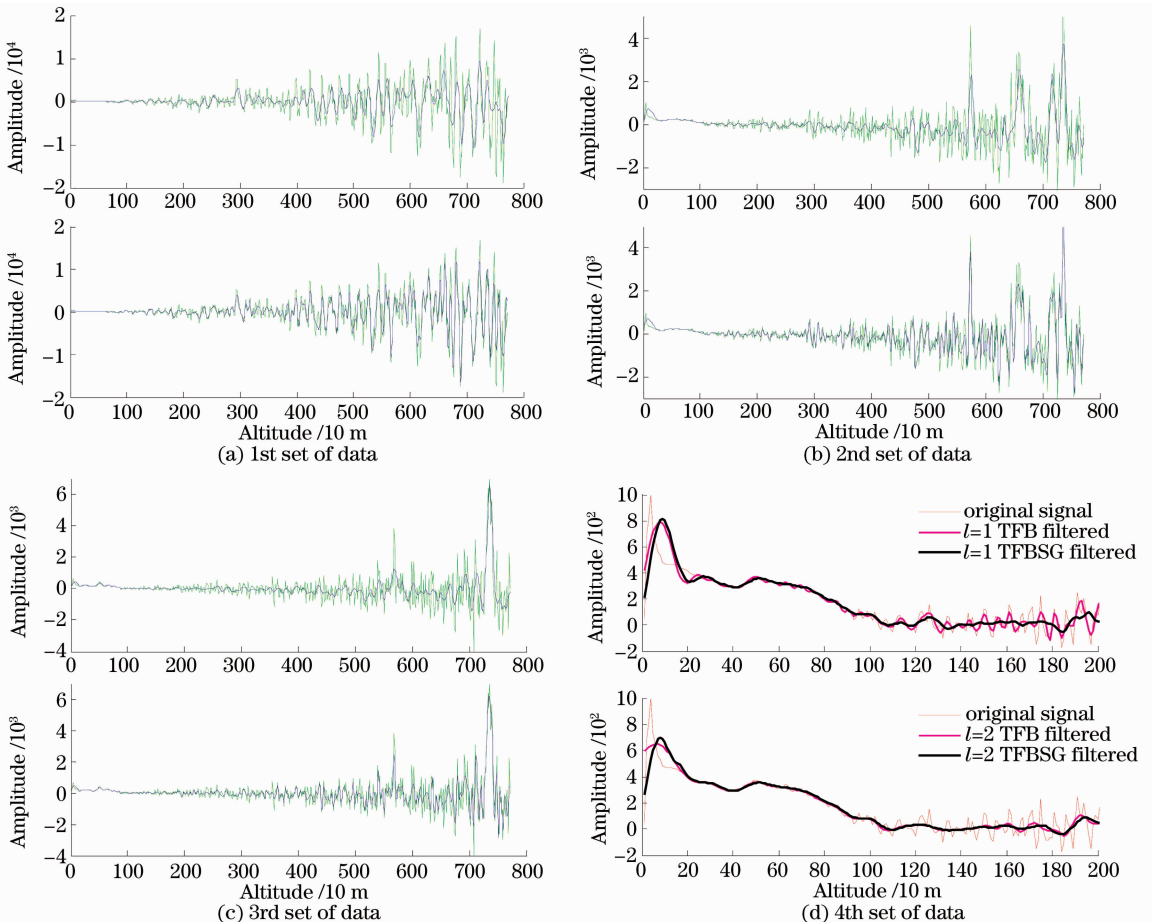


图5 Savitzky-Golay 辅助 TFB 滤波效果

Fig. 5 TFB assisted by Savitzky-Golay filtering effect

色上层为 Savitzky-Golay 辅助 TFB 滤波、下层为未加 Savitzky-Golay 辅助的 TFB 滤波。第一组数据测自 2007 年 4 月 24 日 12:36:57, 能见度较好, 经目测天空有薄云, 但 Vaisala CL31 未能测出云高。直接观察后向散射系数廓线, 发现信号背景噪声很强。经 Savitzky-Golay 辅助 TFB 滤波之后, 使远处的背景噪声被大大抑制的同时保留信号细节, 最后可以有效反演出一层云, 高 2960 m, 符合目测情况。第二组数据测自 2007 年 9 月 25 日 19:15:16, 能见度较差, 无法目测到云层, Vaisala CL31 测出能见度 890 m、未能测出云高。经滤波之后, 不仅可以利用垂直距离较远处的大尺度信号反演出同样的能见度, 而且保留了小尺度细节的、有效的四层云, 高度分别为 5750, 6580, 7170 和 7360 m。第三组数据测自 2008 年 9 月 12 日 17:14:02, 能见度较好, 目测可见高空厚云背景之下有薄云飘动, Vaisala CL31 测出云高 7350 m。经滤波之后, 反演结果除 7350 m 高处有云层之外, 5740 m 高处还有一层薄云, 符合目测情况。

图 5 中第四组原始数据测自 2008 年 12 月 11 日 09:50:38, Vaisala CL31 测出垂直能见度为

820 m, 但是比较当时连续监测的数据, 发现样本数据前后 2 min 的时间跨度内, CL31 所测能见度出现 760~1040 m 的随机抖动。显然该抖动源自近垂直距离处的噪声。由于后向散射信号的 IMF 分量中, 低阶项一般具有高频、交流的特点, 而一般认为大气垂直能见度只与正常人视力识别目标的对比度阈值和大气消光系数的积分相关<sup>[11]</sup>, 因此, 为了帮助反演垂直能见度, 同样可以采用 Savitzky-Golay 辅助 TFB 滤波。经 Savitzky-Golay 辅助 TFB 滤波( $l=1$ )之后, 样本数据的垂直能见度分别反演为 850 m, 样本数据前后 2 min 的跨度内, 能见度的抖动范围为 840~870 m, 抖动幅度明显降低。根据统计理论, 这样相当于降低了能见度测量值的标准差, 提高了反演精度。

观察图 6 中对第一至三组数据的时频分布描述(依次为原始数据、TFB 滤波和 Savitzky-Golay 辅助 TFB 滤波数据), 可见原始数据的时频分布明显比较散乱, 很难从中提取特征信息; 经过 TFB 滤波之后的数据时频分布高频噪声分量明显减少, 特征也得以凸显, 但是若根据此时频分布进行云底云高识别, 准确率仍然不高; 经过 Savitzky-Golay 辅助 TFB 滤波, 数据时频分布中特征信息更加明显, 可

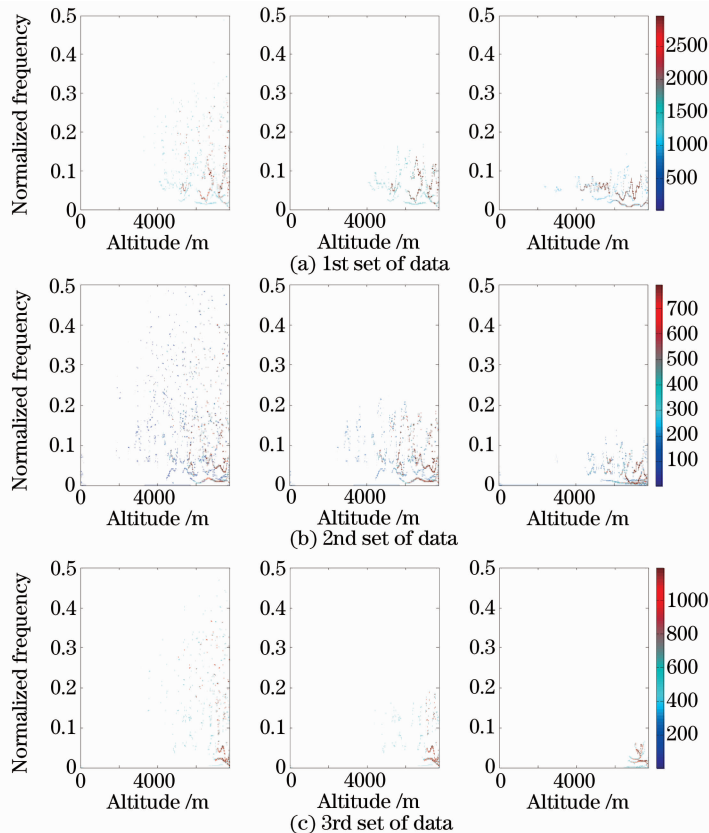


图 6 各组数据对应的时频分布图

Fig. 6 Time-frequency distribution of preceding data



以根据前文提出的后向散射信号时频分布特点帮助实现云底云高信息的提取(图中颜色的深浅反映信号幅度大小,频率高低反映信号变化的快慢)。对比前文已经识别的云层信息和图6中显示的每组数据的第三个时频分布图,可见得信号特征非常明显。

对于第四组数据的垂直能见度反演情况,可以利用 HHT 边际谱来比较。根据垂直能见度反演原则,需要剔除数据中高频交流振荡、保留低频趋势,从图7容易看出  $l=1$  的 TFB 可以明显抑制高频,而辅助 Savitzky-Golay 滤波后,高频分量进一步减少。 $l=2$  时效果则更加显著。

为了进一步验证本文研究方法对后向散射信号处理的可靠性,下面将 Vaisala 分别于 2009 年 7 月

11 日和 13 日直接连续获取的后向散射信号时空分布及其反演结果与该方法处理结果进行对比。2009 年 7 月 11 日北京时间 17:00~18:00,能见度良好,目测可见高空持续有卷云,从获取的时空分布可见云层轮廓[图 8(a)中白色部分]。比较两种云层连续反演结果,发现 Vaisala 云高漏报率约 24% [图 8(a)底部三角记录虚报]、结果偶尔不能反映实际情况;经滤波处理后的反演结果没有出现云高漏报、云高基本反映实际情况。2009 年 7 月 13 日北京时间 06:00~07:00,有薄雾,能见度较差,目测水平能见度不到 1 km。比较两种垂直能见度反演结果,发现 Vaisala 结果在 930~1160 m 之间跳变严重;滤波处理后的结果在 950~1040 m 间变化平缓。

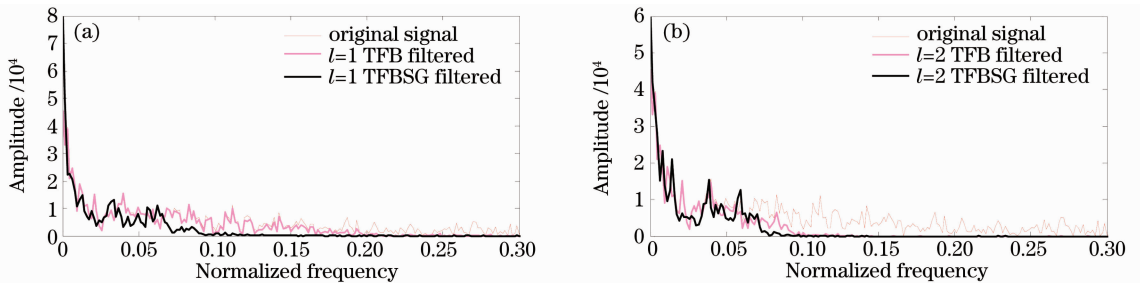


图7 第四组数据(2008年12月11日)的边际 HHT 谱

Fig. 7 HHT Marginal spectra of the 4th set of data (Dec 11, 2008)

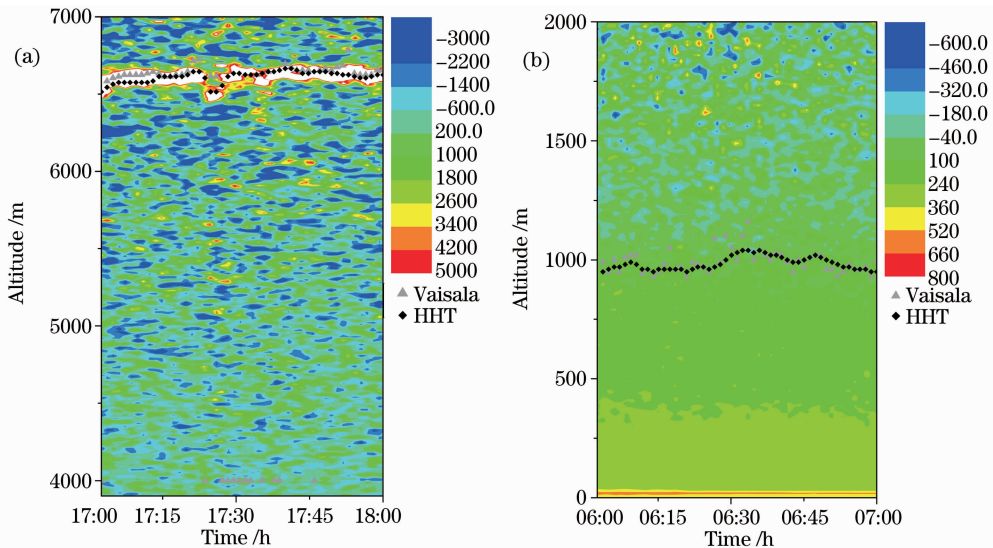


图8 Vaisala 直接反演和本文反演的云高和能见度比较

Fig. 8 Cloud heights and vertical visibility comparison of Vaisala-direct and HHT retrieving

## 6 结 论

基于米氏散射理论的半导体激光云高仪后向散射信号不仅背景噪声强、有用信号弱,而且背景噪声功率的随机起伏往往较大,难以预测。这决定了抑制其噪声并从中提取特征信号难度较大。经过大量

不同天气条件下的半导体激光云高仪所获取的实验数据验证,希尔伯特-黄变换对于具有非线性非平稳特点的激光云高仪后向散射信号有广泛的适应性,同时,应用该变换,通过特定的 Savitzky-Golay 辅助 TFB 结构进行滤波、并结合 HHT 时频分析能够在

保留有用信息特征、有效抑制各种平稳或非平稳噪声的基础上,帮助云底云高识别及能见度反演,并且降低了云高漏报率、提高了云底云高的细节识别能力和能见度反演的精度。

## 参 考 文 献

- 1 Zheng Fatai, Hua Dengxin, Zhou Awei. Empirical mode decomposition algorithm research & application of Mie lidar atmospheric backscatter signal[J]. *Chinese J. Lasers*, 2009, **36**(5): 1068~1074  
郑发泰, 华灯鑫, 周阿维. Mie 散射激光雷达大气回波信号经验模式分解算法的研究与应用[J]. *中国激光*, 2009, **36**(5): 1068~1074
- 2 Bo Guangyu, Zhong Zhiqing, Liu Bo *et al.*. Design of optical receiver and signal detection unit for combined rayleigh-Raman-Mie lidar[J]. *Chinese J. Lasers*, 2009, **36**(11): 3020~3025  
伯广宇, 钟志庆, 刘 博等. 瑞利-拉曼-米氏激光雷达光学接收和信号检测设计[J]. *中国激光*, 2009, **36**(11): 3020~3025
- 3 Dong Yunsheng, Liu Wenqing, Liu Jianguo *et al.*. Application study of lidar in urban traffic pollution[J]. *Acta Optica Sinica*, 2010, **30**(2): 315~320  
董云升, 刘文清, 刘建国等. 激光雷达在城市交通污染中应用研究[J]. *光学学报*, 2010, **30**(2): 315~320
- 4 Fan Guangqiang, Hu Shunxing, Xie Jun *et al.*. Application of the adaptive filter to the data processing for Raman lidar[J]. *Chinese J. Lasers*, 2009, **36**(11): 3002~3006  
范广强, 胡顺星, 谢 军等. 自适应滤波在拉曼激光雷达数据处理中的应用[J]. *中国激光*, 2009, **36**(11): 3002~3006
- 5 Su Liping, Zhao Weijiang, Ren Deming. Processing methods for laser scattering signals of bubbles[J]. *Chinese J. Lasers*, 2009, **36**(4): 814~818  
苏丽萍, 赵卫疆, 任德明等. 气泡激光散射信号处理方法[J]. *中国激光*, 2009, **36**(4): 814~818
- 6 Huang N. E., Shen Z., Long S. R. *et al.*. The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis[J]. *Proc. Royal Soc. Lond A*, 1998, **454**(1971): 903~995
- 7 Flandrin P., Rilling, G. Gonçalves P. Empirical mode decomposition as a filterbank[J]. *IEEE Signal Proc Lett.*, 2003, **11**(2): 112~114
- 8 Tan Shanwen, Qin Shuren, Tang Baoping. The filtering character of Hilbert-Huang transform and its application[J]. *Journal of Chongqing University*, 2004, **27**(2): 9~12  
谭善文, 秦树人, 汤宝平. Hilbert-Huang 变换的滤波特性及其应用[J]. *重庆大学学报*, 2004, **27**(2): 9~12
- 9 Boudraa A. O., Cexus J. C.. EMD based signal filtering[J]. *IEEE Transactions on Instrumentation and Measurement*, 2007, **56**(6): 2196~2202
- 10 José Luis Guinón, Emma Ortega, José García-Antón *et al.*. Moving average and Savitzki-Golay smoothing filters using mathcad[C]. Proceedings of the ICEE, 2007, September. 3~7
- 11 Han Daowen, Liu Wenqing, Liu Jianguo *et al.*. Retrieval method for aerosol mass concentration vertical distribution[J]. *Chinese J. Lasers*, 2006, **33**(11): 1567~1573  
韩道文, 刘文清, 刘建国等. 气溶胶质量浓度空间垂直分布的反演方法[J]. *中国激光*, 2006, **33**(11): 1567~1573