

基于模式识别的激光雷达遥感灰霾组分识别模型

王治飞 刘东 成中涛 杨甬英 罗敬 沈亦兵 白剑 汪凯巍

(浙江大学光电信息工程学系现代光学仪器国家重点实验室, 浙江 杭州 310027)

摘要 提出了一种基于贝叶斯模式识别的激光雷达大气遥感灰霾组分识别的方法。介绍了灰霾组分模式识别模型的建立过程, 并利用具体的贝叶斯判别函数作为灰霾粒子光学特征向量的选择依据对灰霾粒子进行识别分类。采用计算机仿真实现了该灰霾组分模式识别模型, 并通过两种自验证方法检验了模型的正确性和稳定性。讨论了该模型对现有大气遥感激光雷达的适用性, 凸显了偏振高光谱分辨率激光雷达(HSRL)的优势。

关键词 遥感; 激光雷达; 灰霾; 模式识别; 气溶胶; 高光谱分辨率激光雷达

中图分类号 TN958.98 **文献标识码** A **doi**: 10.3788/CJL201441.1113001

Pattern Recognition Model for Haze Identification with Atmospheric Backscatter Lidars

Wang Zhifei Liu Dong Cheng Zhongtao Yang Yongying Luo Jing
Shen Yibing Bai Jian Wang Kaiwei

(State Key Laboratory of Modern Optical Instrumentation, Department of Optical Engineering,
Zhejiang University, Hangzhou, Zhejiang 310027, China)

Abstract A pattern recognition model for haze identification with atmospheric backscatter lidars is proposed. The process of building the characteristics sample database for haze pattern recognition is described in detail. The classification of haze particles by using Bayesian discriminant function, as the selection basis of haze optical characteristics vector, is presented. Computer simulation for the proposed pattern recognition model of haze identification is carried out. Two self-calibration approaches are employed to check the validity and stability of the model. By analyzing the applicability of this model for atmospheric lidars, the advantage of polarized high spectral resolution lidar (HSRL) is highlighted.

Key words remote sensing; lidar; haze; pattern recognition; aerosol; high spectral resolution lidar

OCIS codes 010.1290; 010.3640; 010.1350; 010.1110; 280.1100

1 引言

霾, 又称灰霾、阴霾、烟霞, 是大量极细微的干尘粒等均匀地浮游在空中, 使水平能见度小于 10 km 的空气普遍浑浊现象^[1]。由于经济规模的迅速扩大和城市化进程的加快, 人为排放的大量气溶胶已使空气质量严重恶化, 致使霾的产生已不完全是一种

自然现象, 大气污染日益严峻^[2-3]。因此, 有效控制灰霾将对人们的生产生活、社会经济发展以及人类赖以生存的环境都具有非常重要的意义。而准确检测和识别灰霾的种类和组分是对其特征、产生机理及危害等方面进行全面了解、掌握并有效控制的前提。

收稿日期: 2014-04-14; **收到修改稿日期**: 2014-06-01

基金项目: 国家自然科学基金(41305014)、教育部博士点基金(20130101120133)、浙江省教育厅科研项目(Y201329660)、浙江省“仪器科学与技术”重中之重学科开放基金(JL130113)、中央高校基本科研业务费专项资金(2013QNA5006)、现代光学仪器国家重点实验室创新基金(MOI201208)

作者简介: 王治飞(1992—), 男, 本科生, 目前参与激光雷达大气遥感灰霾组分识别方面的研究。

E-mail: wangzhifeifly@163.com

导师简介: 刘东(1982—), 男, 博士, 副教授, 博士生导师, 主要从事光学传感与信息处理技术等方面的研究。

E-mail: liudongopt@zju.edu.cn(通信联系人)

本文电子版彩色效果请详见中国光学期刊网 www.opticsjournal.net

传统的大气颗粒物测量方法等离线分析手段虽然可以对样品进行全面、准确地分析,但这种方法费时费力、化学成分易受破坏且无法真正反映原始气溶胶的特征。单颗粒气溶胶质谱技术等实时在线测量技术以及气溶胶吸湿性谱仪(HTDMA)、光腔衰荡光谱仪(CRDS)等原位检测分析仪器^[4-5]仅能在有限区域内进行单点采样,应用受到一定的限制。

如何快速、大范围地对灰霾进行实时监测是目前急需开展的研究。遥感因其能够快速获得大范围的观测数据,且无需置身灰霾之中便可对大气污染、霾分布与强度状况进行快速的监测,具有非常好的发展前景^[6-7]。特别是大气遥感激光雷达,由于其具有较高的时间和空间分辨率,可以应用于车载、机载甚至是无人职守的地基等多种运行方式,适用于灰霾的遥感^[8-9]。

本文提出一种基于模式识别的激光雷达大气遥感灰霾组分的方法。对该方法的基本原理及实现过程进行了详细地介绍。计算机仿真实验表明,该方法对不同种类激光雷达可以实现不同效果。其中双波长的偏振高光谱分辨率激光雷达(HSRL)的灰霾识别率远远优于其他种类的激光雷达。该研究对快速、大范围、高精度地监测灰霾,进而对其进行有效控制,改善地球气候环境,提升人们的生活质量都具有非常重要的意义。

2 灰霾的光学属性

灰霾本质上是大气气溶胶系统,所以其具有气溶胶的物理和光学属性,如尺度谱分布、光学厚度(AOD)、消光系数、后向散射系数、激光雷达比等。这些参数对基于模式识别的激光雷达遥感灰霾的识别和分类具有重要的参考意义。

2.1 后向散射系数

不同气溶胶成分的后向散射系数与消光系数并没有很一般规律的差异,故在气溶胶探测与分类中常采用后向散射系数衍生的 Angstrom 指数进行研究,即

$$a_{\beta} = -\frac{\ln(\beta_a^{532}/\beta_a^{1064})}{\ln(532/1064)}, \quad (1)$$

式中 β_a 表示在波长 i 的气溶胶散射系数, $C_a = \beta_a^{532}/\beta_a^{1064}$ 表示散射颜色比。利用气溶胶后向散射系数还可以得到另外两个重要的量,即退偏比

$$\delta_a = \ln(\beta_a^{\perp}/\beta_a^{\parallel}), \quad (2)$$

和谱退偏比

$$R_a = \delta_{1064}/\delta_{532}. \quad (3)$$

2.2 消光系数

晴空背景和黑体目标情况下对白光而言,大气水平能见度 V_H (单位为 km^{-1}) 与消光系数 α (单位为 km^{-1}) 和识别目标对比阈值 $\epsilon \approx 0.02$ 的关系为^[10]

$$V_H = \frac{1}{\alpha} \ln \frac{1}{\epsilon} \approx \frac{3.912}{\alpha}. \quad (4)$$

2.3 光学厚度

大气气溶胶光学厚度是指大气介质的消光系数在垂直方向上的积分,描述大气气溶胶对光的衰减作用。大气气溶胶的光学厚度与波长具有如下关系^[11]:

$$\tau_a(\lambda) = B\lambda^{-A}, \quad (5)$$

式中 B 为气溶胶混浊度系数, A 为 Angstrom 波长指数,它是衡量气溶胶粒子大小的一个重要光学参数,其值越大,说明粒子越小,反之亦然。

2.4 激光雷达比

普通的偏振 Mie 散射的激光雷达在某波段上为求解含有三个未知数(β_a^{\parallel} 、 β_a^{\perp} 、 α)的激光雷达方程常预先假设气溶胶消光系数与后向散射系数之比,即激光雷达比,其定义为气溶胶消光系数和后向散射系数之比,即

$$S_a = \alpha_a/\beta_a. \quad (6)$$

大气气溶胶和云雾粒子的退偏比、激光雷达比等光学参数和粒子的物理性质密切相关,不同种类粒子的退偏比、激光雷达比等光学参数不会完全相同,并且不同波长上的光学参数的相互关系也不相同。例如在退偏比方面,研究发现纯灰尘的典型特征是大的 532 nm 气溶胶退偏比能达到 30%~35%^[12];而灰尘和球状颗粒的混合物的退偏比值大约为 8%~10%^[13];晶状海盐和在大气中存在一段时间后的生物燃烧物以及火山气溶胶可由更小的退偏比来表征^[14-15]。在激光雷达比方面,水溶性气溶胶的激光雷达比随相对湿度增大而增大。烟灰和粗粒态矿物质的激光雷达比很大;烟灰和包含聚积态海盐类粒子的气溶胶在两个激光波长上有较大差异,1064 nm 波长的雷达比约为 532 nm 的 2 倍,而粗粒态矿物质刚好相反;各种卷云雾粒子的激光雷达比与光学厚度有关,当光学厚度较小时,激光雷达比很低,而光学厚度较大时,雷达比也较大^[16-18]。所以,可以发现退偏比、激光雷达比等光学参数将成为区分灰霾不可或缺的参量。

另外需要指出的一点是对于以上气溶胶光学参数,由于普通 Mie 散射激光雷达需要预先假设出气

溶胶粒子的激光雷达比,导致使用单波长偏振 Mie 散射激光雷达仅能准确得到退偏比 δ_a ,在此基础上进一步发展的双波长偏振 Mie 激光雷达也只能准确得到退偏比 δ_a 、谱退偏比 R_a 。而在 Mie 散射激光雷达的基本结构加上原子/分子滤光器或者干涉仪滤光器而形成的 HSRL,其通过滤光器将气溶胶的 Mie 散射和大气分子的瑞利散射在光谱上分离开来以得到一个额外的激光雷达方程,从而实现不依赖于先验假设而精确求解出气溶胶粒子的光学参数(气溶胶粒子的消光系数 α_a 及后向散射系数 β_a)^[19-22]。这使得单波长的偏振 HSRL 能精确获得退偏比 δ_a 和雷达比 S_a 的值(如 532 nm),而双波长的偏振 HSRL 由于可在两个波段上进行精确测量,故能准确得到退偏比 δ_a 、谱退偏比 R_a 、散射颜色比 C_a 和雷达比 S_a 这四个光学参量。

3 灰霾的激光雷达大气遥感模式识别模型

Burton 等^[22]对多波长偏振激光雷达在全球范围内获取的数据进行分析发现,冰晶粒子和灰尘均具有较高的气溶胶退偏比。但一般来说,冰晶粒子的激光雷达比要低于灰尘。而对于灰尘,在纯灰尘中混有少量其他成分时,它们的雷达比没有显著差异,但是退偏比却不同。在对相对洁净的海洋上空的气溶胶进行探测时发现,其 532 nm 激光雷达比、532 nm 退偏比、后向散射颜色比以及谱退偏比四个特征都比较小。但对有一定污染的海洋上空的气溶胶进行探测时发现它们的上述四个特征都居于海洋气溶胶和污染物气溶胶之间。城市的气溶胶和生物燃烧产生的气溶胶具有的共同特点是退偏比低,后向散射颜色比高,532 nm 的激光雷达比也很

高,但它们的谱退偏比却有一些差异。所以,可以充分利用大气遥感激光雷达以及上述灰霾的光学属性,构建一套基于模式识别的大气遥感激光雷达大气遥感灰霾组分的方法,如图 1 所示。总体来说,该方法的主要过程就是通过将模式识别特征样本库中的数据作为训练数据构建出灰霾的激光雷达大气遥感模式识别模型,得到模式识别的判别函数;再将待分类的灰霾遥感数据与建立的模式识别特征样本库进行比较,根据判别函数生成该灰霾属于各种类型灰霾的相对概率值,选择相对概率最大且大于判断阈值的灰霾类型为该待分类灰霾所属类型。

3.1 模式识别特征样本库

从图 1 中的原理框图不难看出,模式识别特征样本库的建立是该方法得以实现的关键。图 2 表示该模式识别特征样本库的建立过程,即针对搜集到的灰霾光学特性样本库 H_0 ,根据已有的灰霾分类进行光学特性分析,从而得到初步的光学特性特征向量 \mathbf{X}_0 ;为提高模型的运行效率,减小实际应用的复杂度,再通过因子分析对 \mathbf{X}_0 进行降维,最终获到模型中的光学特性特征向量 \mathbf{X} ;根据最终光学特性特征向量 \mathbf{X} 对最初搜集的灰霾光学特性样本库 H_0 进行整理,最终得到可用于模式识别的特征样本库 H 。

其中降维分析主要根据原灰霾光学特性样本库中的数据计算各种光学特性两两之间的相关系数。若其值趋近于 1,则说明这两种灰霾光学特性极为相似,即其中一种光学特性已能很好地代表另一种而实现特征空间的降维;而如果相关系数趋近于 0,则说明两种光学特性差别很大,应保留这两种特性。

综上,模式识别特征样本库 H 相对于最初搜集到的灰霾光学特性样本库 H_0 。在数据数量以及数据的维度上都得到了精简,剔除了冗余的灰霾光学特征,从而可以更准确高效地进行模式识别。

3.2 判别函数

在统计模式识别理论中,常用的判别策略有决策树和贝叶斯方法。决策树在进行多维大数据处理时效率不高。而贝叶斯方法以贝叶斯理论为基础,理论上与其他分类方法相比具有最小的误差率。

假设待确定样本的特征向量为 \mathbf{x} ,其属于 n 种灰霾中的第 i 种灰霾类型 ω_i 的概率为

$$P(\omega_i | \mathbf{x}) = \frac{P(\mathbf{x} | \omega_i)P(\omega_i)}{\sum_{j=1}^n P(\mathbf{x} | \omega_j)P(\omega_j)}, \quad i = 1, 2, 3, \dots, n, \quad (7)$$

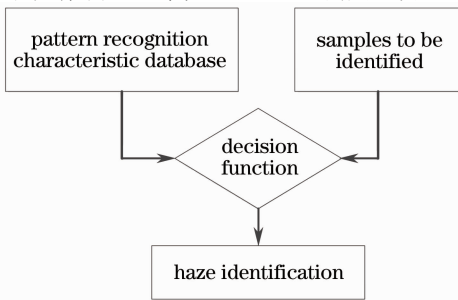


图 1 基于模式识别的大气遥感激光雷达大气遥感灰霾组分的原理框图

Fig. 1 Principle diagram of haze components recognition using atmospheric remote sensing lidar based on pattern recognition

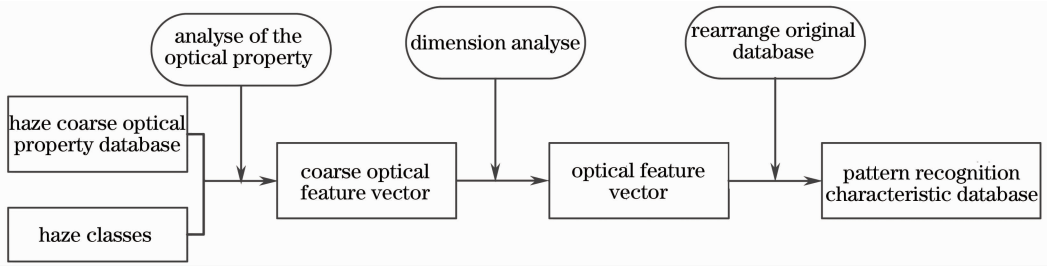


图 2 模式识别样本库的建立过程

Fig. 2 Establishment of the pattern recognition sample library

式中 $P(\omega_i)$ 是特征样本库中每种灰霾类型样本占总样本库的比例, $P(\mathbf{x} | \omega_i)$ 指若待确定样本属于灰霾类型 ω_i 时其出现的概率。Cattrall 等^[17,22] 都已指出各种类灰霾的特征样本库都可认为是多维高斯分布, 即

$$P(\mathbf{x} | \omega_i) = \frac{1}{(2\pi)^{d/2} |\boldsymbol{\Sigma}_i|^{1/2}} \exp\left\{-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}_i)^\top \boldsymbol{\Sigma}_i^{-1} (\mathbf{x} - \boldsymbol{\mu}_i)\right\}, \quad (8)$$

式中 $\mathbf{x} = (x_1, x_2, \dots, x_d)^\top$ 为 d 维特征向量, $\boldsymbol{\mu}_i = (\mu_1, \mu_2, \dots, \mu_d)^\top$ 为第 i 类的 d 维均值向量, $\boldsymbol{\Sigma}_i$ 是第 i 类的 $d \times d$ 维协方差矩阵, 定义为 $\boldsymbol{\Sigma}_i = E\{(\mathbf{x} - \boldsymbol{\mu}_i)(\mathbf{x} - \boldsymbol{\mu}_i)^\top\}$, 而 $\boldsymbol{\Sigma}_i^{-1}$ 是 $\boldsymbol{\Sigma}_i$ 的逆矩阵。可以看出, 只需通过实验数据估计出上述多维高斯分布的参数, 即可代入(7)式计算出 $P(\omega_i | \mathbf{x})$, 即待确定样本属于第 i 种灰霾的相对概率。决策规则可以有两种, 1) 设定一个判断阈值, 若待确定样本属于第 i 种灰霾的相对概率超过该阈值则认为待确定样本即为该种类灰霾; 2) 选取最大相对概率的灰霾种类作为待确定样本的识别结果, 即若 $P_i(\mathbf{x}) = \max_{i=1,2,\dots,n} P(\omega_i | \mathbf{x})$, 则 \mathbf{x} 属于 ω_i 类型。

4 计算机仿真分析

根据灰霾光学特性样本库选择出灰霾的光学特征向量, 构建出灰霾的模式识别特性样本库; 通过模式识别特征样本库数据作为训练数据进行模型学习, 得到贝叶斯判别模型的模型参数; 再将待分类灰霾数据输入到模式识别模型, 生成该实例属于各种类灰霾的相对概率, 选出其中的最大值或与判断阈值比较从而得到该待分类灰霾所属的灰霾类型。为检验模式识别模型的可用性及稳定性, 利用训练样本库即灰霾的模式识别特性样本库以及重新生成的测试样本数据库对灰霾的模式识别模型进行了两种自验证校验, 最后还讨论了在灰霾特征向量降维的情况下激光雷达对模式识别结果的影响。

4.1 模型的建立

4.1.1 灰霾的分类

要对灰霾组分进行识别, 首先需要对其进行分类。Burton 等^[22-23] 在大量实地测量数据的基础上进行数据分析, 通过聚类分析等方法进行归纳, 获得了较为合理的气溶胶种类划分。由于灰霾粒子本质上是气溶胶粒子, 故可采纳 Burton 等的划分。将灰霾粒子分为以下 8 类: 冰晶粒子、纯灰尘、灰尘混合物、海洋气溶胶、海洋污染物、城市气溶胶、生物燃烧气溶胶、新鲜烟尘。

4.1.2 特征向量的确定

在对灰霾进行分类的过程中, 可以发现由后向散射系数的运算而得到的量, 如退偏比 δ_a 、谱退偏比 R_a 、散射颜色比 C_a 等, 以及由后向散射系数和消光系数共同决定的雷达比 S_a 等并不会随灰霾浓度的变化而变化, 且不同种类的灰霾粒子在这四个特性中至少有一种特性存在较为明显的区别, 这对于区分不同种类的灰霾气溶胶有重大意义。

在定量分析方面, 对已知种类的灰霾进行大量测量, 获得各不同类别灰霾的光学属性, 建立样本数据库作为样本训练的集合。再根据前面的分析, 通过对这个灰霾光学特性样本库进行降维分析后发现, 其余的灰霾光学特性可由这四个灰霾光学特征来表示, 故将灰霾类别划分的特征向量表示为

$$\mathbf{x} = (\delta_a, S_a, C_a, R_a)^\top, \quad (9)$$

从而构建出灰霾粒子的四维特征空间。这样, 在建立样本集时就可以有针对性地提取已知灰霾成分的这些特征量, 并存储为数据库供以后使用。

4.1.3 模式识别特征样本库

基于 Burton 等在北美洲进行的野外任务得到的实验数据^[20], 模拟构建出本计算机仿真实验的灰霾模式特征样本库。数据库中有冰晶粒子样本 10000 个, 纯灰尘 3000 个, 灰尘混合物 75000 个, 海洋气溶胶 13000 个, 海洋污染物 11000 个, 城市气溶胶 85000 个, 生物燃烧气溶胶 22000 个, 新鲜烟尘

39000 个,四维灰霾模式识别特征空间在各空间中的投影分布如图 3 所示。图 3(a)~(c)为在 532 nm 退偏比-532 nm 雷达比空间、532 nm 退偏比-后向散射颜色空间、532 nm 退偏比-谱退偏比空间投影

情况;(d)~(f)分别是在 532 nm 激光雷达比-后向散射颜色比空间、532nm 激光雷达比-谱退偏比空间、后向散射颜色比-谱退偏比空间的投影情况。

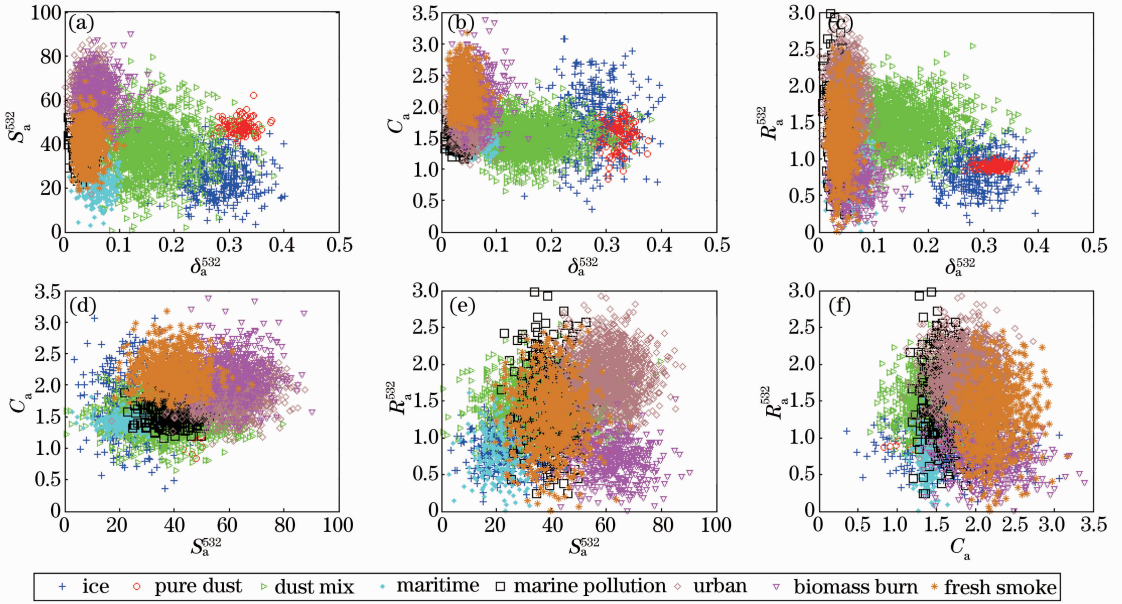


图 3 仿真中所使用的灰霾数据库。(a) 在 532 nm 退偏比-532 nm 雷达比空间投影情况;(b) 在谱退偏比-后向散射颜色空间投影情况;(c) 在 532 nm 退偏比-谱退偏比空间投影情况;(d) 在 532 nm 激光雷达比-后向散射颜色比空间投影情况;(e) 在 532 nm 激光雷达比-谱退偏比空间投影情况;(f) 在后向散射颜色比-谱退偏比空间投影情况

Fig. 3 Haze database used in simulation. (a) Projection in 532 nm depolarization to 532 nm lidar ratio space; (b) projection in depolarization spectral ratio to backscatter color ratio space; (c) projection in 532 nm depolarization to depolarization spectral ratio space; (d) projection in 532 nm lidar ratio to backscatter color ratio space; (e) projection in 532 nm lidar ratio to depolarization spectral ratio space; (f) projection in backscatter color ratio to depolarization spectral ratio space

4.2 自验证的准确率

为验证激光雷达大气遥感灰霾组分模式识别模型的实用性和稳定性,在计算机仿真实验中用两种自验证的方法检验仿真实验的正确性。

采用严格的自验证方法,即将实验中灰霾数据库中的各种类灰霾的数据视作待确定样本作为仿真的输入,计算出属于该种类灰霾类型的相对概率,取 66.7% 为判断阈值,即若一待确定样本属于某种类灰霾相对概率在 66.7% 以上就可以认为这个待确定样本属于该种灰霾。8 种灰霾自验证的准确率如表 1 所示。图 4 表示的是 8 种灰霾自验证的详细分析结果,图中纵轴代表进行自验证的灰霾类型对象,不同的颜色及其横轴范围代表对某种灰霾进行自验证时被分类成的灰霾类型及其比例。例如冰晶粒子横方条蓝色部分代表冰晶粒子样本在自验证时仍被分类为冰晶粒子占总冰晶粒子样本点数的比例,横方条红色部分代表冰晶粒子被误分类为纯灰尘的样

本点占有所有冰晶粒子样本的比例,依此类推。

表 1 八种灰霾严格自验证的准确率

Table 1 Accuracy rates of eight kinds of hazes self-calibration

| Haze class | Accuracy of self-verification |
|------------------|-------------------------------|
| Ice | 91.75% |
| Pure dust | 97.47% |
| Mix dust | 98.66% |
| Maritime | 95.21% |
| Marine pollution | 73.24% |
| Urban | 91.52% |
| Biomass burn | 85.39% |
| Fresh smoke | 86.09% |

采用广义的自验证方法,即重新产生与某种类灰霾的分布一致但取值不同的相同规模的测试样本库,将其作为仿真输入,计算这些样本点属于该种类灰霾类型的相对概率。仍然取 66.7% 为判断阈值,8 种灰霾的广义自验证的准确率如表 2 所示,这里进行了四次仿真实验并取平均值。

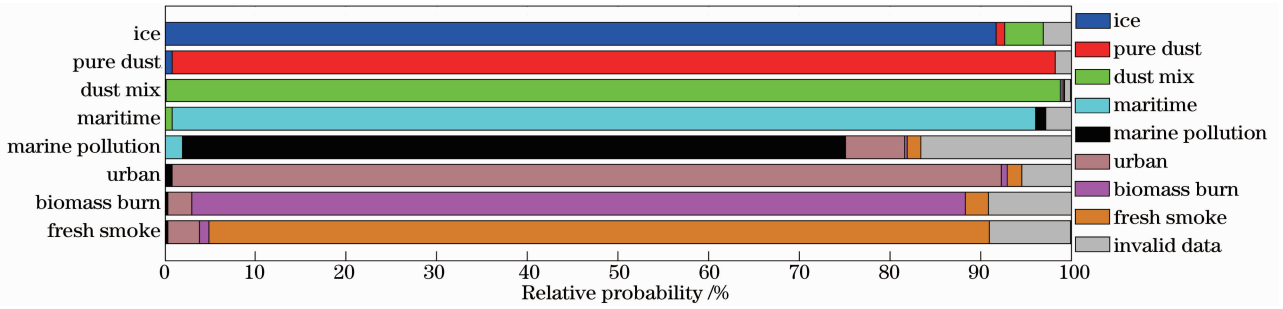


图 4 八种灰霾自验证的详细分析结果

Fig. 4 Analysis results of eight kinds of hazes self-calibration

表 2 八种灰霾广义自验证的准确率

Table 2 Accuracy rates of eight kinds of hazes generalized self-calibration

| Haze type | No. | 1 | 2 | 3 | 4 | Mean result |
|------------------|-----|--------|--------|--------|--------|-------------|
| Ice | | 91.46% | 92.04% | 92.31% | 91.19% | 91.75% |
| Pure dust | | 96.90% | 96.77% | 97.17% | 96.50% | 96.84% |
| Mix dust | | 98.66% | 98.67% | 98.67% | 98.67% | 98.67% |
| Maritime | | 95.40% | 94.89% | 94.95% | 95.22% | 95.12% |
| Marine pollution | | 73.68% | 73.58% | 74.35% | 74.25% | 73.97% |
| Urban | | 91.55% | 91.44% | 91.70% | 91.67% | 91.59% |
| Biomass burn | | 85.19% | 85.05% | 85.69% | 85.63% | 85.39% |
| Fresh smoke | | 86.05% | 85.96% | 86.07% | 86.05% | 86.03% |

由上述两种自验证检验结果可以看出,海洋污染物样本的再识别最为困难,只有 74%左右的样本点能被正确分类。这一方面是由于灰霾数据库中的海洋污染物样本数比较少,另一方面也是因为海洋污染物样本在特征空间中与其他种类的灰霾重叠区域较大,尤其是在 532 nm 退偏比-532 nm 雷达比空间和 532 nm 退偏比-谱退偏比空间的投影中海洋污染物与其他种类的灰霾几乎完全混叠无法分开。同时,生物燃烧气溶胶和新鲜烟尘再识别的正确率同样较低,只有 86%左右,这完全是因为在特征空间中这两种灰霾粒子的特征向量与城市气溶胶的特征向量三者之间相互混叠的原因。而冰晶粒子、灰尘混合物等其他五种灰霾粒子自验证的准确率均基本达到了 90%及以上,尤其是灰尘混合物几乎不会出现错误。

另外,从两种自验证分析可以发现,严格自验证检验和广义自验证检验的结果很接近,说明灰霾模型数据库中的样本点数已足够多,自验证检验结果已趋于稳定状态,在相同分布取值的情况下随机改变样本点的值并不会严重影响模式识别结果的正确性。故之后的分析中只列举出严格自验证结果。

进一步分析发现,冰晶粒子样本点中有 4.30%的样本点被误分类为灰尘混合物,海洋污染物、城市气溶胶、生物燃烧气溶胶和新鲜烟尘四种灰霾的串扰相对严重,尤其是有 6.47%的海洋污染物样本点会被误分类到城市气溶胶,被误分类为城市气溶胶和新鲜烟雾的生物燃烧气溶胶样本点都超过了 2%,还有大约 3.47%的新鲜烟雾的样本被误分为城市气溶胶,除此之外的其余串扰的相对概率都不大于 1.6%。不过也由此可看出,许多待判断的样

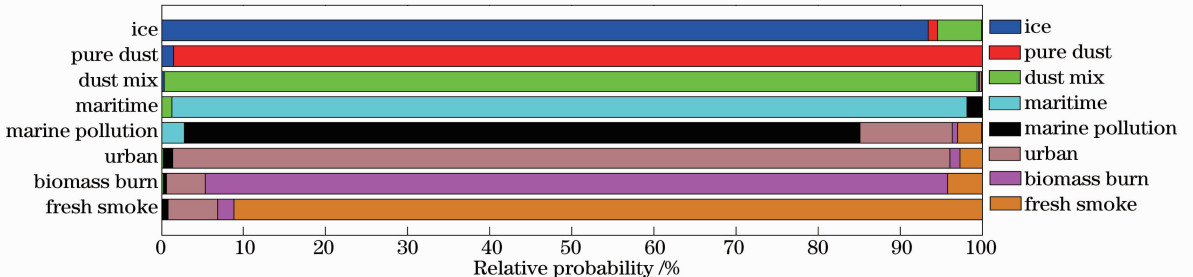


图 5 不设判断阈值时八种灰霾自验证的详细分析结果

Fig. 5 Analysis results of eight kinds of hazes self-calibration without threshold

本点会被错误地划归为无效数据点。这可以通过降低判断阈值甚至取消判断阈值来实现,但如此做将会增大不同灰霾种类之间的串扰。图 5 为不设判断阈值时八种灰霾自验证的详细分析结果。与图 4 进行对比可以发现,不设判断阈值虽然在一定程度上提高了某些灰霾类型的自验证准确度,但各种灰霾之间串扰更加严重,故综合来看确定一个判断阈值还是有必要的,但需根据多次实验和实际需要优化出一个最佳的判断阈值。

然而需要指出的是,要想获得四个灰霾的特征参数最理想的是利用双波长偏振 HSRL,但是由于研究深度和技术制约等因素,双波长偏振 HSRL 还未被广泛接受和使用。目前为获得这些光学特征参数通常的是使用单波长偏振 HSRL(532 nm 波段)和其他大气遥感设备如单波长偏振 Mie 散射激光雷达(1064 nm 波段)等联合测量实现。

表 3 对非双波长偏振 HSRL 八种灰霾自验证的准确率

Table 3 Accuracy rates of eight kinds of hazes self-calibration for non-double-wavelength polarized HSRL

| Haze type | Lidar type | Dual-wavelength polarized | Single-wavelength | Single-wavelength polarized |
|------------------|-----------------------|---------------------------|-----------------------|-----------------------------|
| | Mie backscatter lidar | polarized HSRL | Mie backscatter lidar | |
| Ice | 71.63% | 84.02% | 58.68% | |
| Pure dust | 32.73% | 93.77% | 0 | |
| Mix dust | 97.39% | 97.24% | 96.31% | |
| Maritime | 0 | 74.25% | 0 | |
| Marine pollution | 0.59% | 0 | 0 | |
| Urban | 71.03% | 78.57% | 0 | |
| Biomass burn | 14.61% | 0.89% | 0 | |
| Fresh smoke | 0 | 43.76% | 0 | |

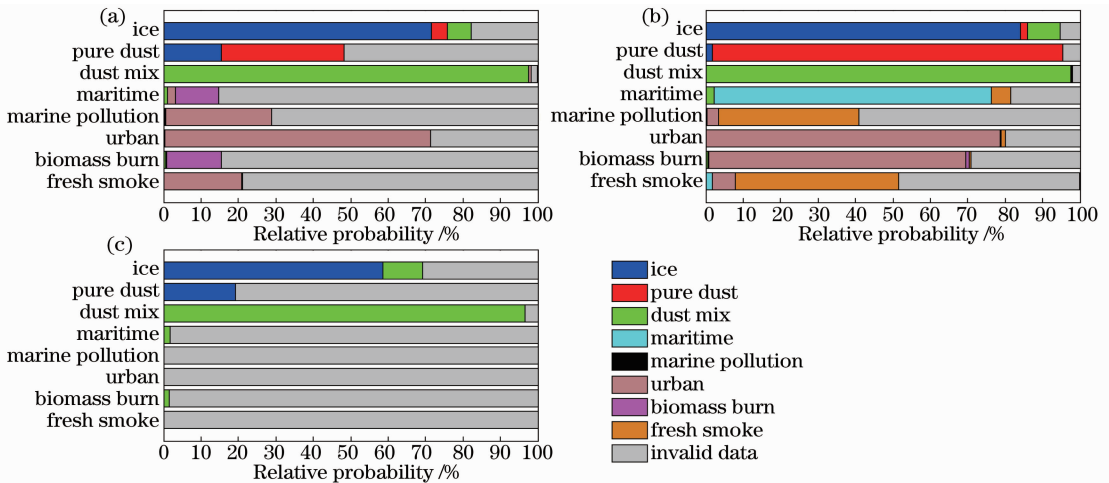


图 6 对非双波长偏振 HSRL 八种灰霾自验证的详细分析结果。(a) 双波长偏振 Mie 散射激光雷达八种灰霾自验证的准确率;(b) 单波长偏振 HSRL 八种灰霾自验证的准确率;(c) 单波长偏振 Mie 散射激光雷达八种灰霾自验证的准确率

Fig. 6 Analysis results of eight kinds of hazes self-calibration for non-double-wavelength polarized HSRL. (a) Accuracy rates of eight kinds of hazes self-calibration for double-wavelength polarized Mie scattering lidar; (b) accuracy rates of eight kinds of hazes self-calibration for single-wavelength polarized HSRL lidar; (c) accuracy rates of eight kinds of hazes self-calibration for single-wavelength polarized Mie scattering lidar

由此可以看出,尽管单波长偏振 HSRL 只能获得退偏比 δ_0 和雷达比 S_0 两个灰霾特征参量,但由于各种类灰霾在这两维空间中分布较为独立,故还是能够有较大把握判断出灰霾所属种类,特别是对冰晶粒子、纯灰尘、灰尘混合物、海洋气溶胶这四种灰霾,有超过 74% 的概率能准确分类,尤其是对灰尘混合物有超过 97% 的正确率;尽管对城市气溶胶再识别的准确性也能达到 79% 左右,但其他种类的灰霾与其串扰很严重,尤其是生物燃烧气溶胶有接近 69% 的样本点会被错误分类到城市气溶胶,这样即使待分类灰霾被判定为城市气溶胶其置信率也很低。

相对于单波长偏振 HSRL,双波长偏振 Mie 散射激光雷达对灰霾组分识别的能力会相对更弱一

些。其只能比较好地对冰晶粒子和灰尘混合物进行再识别;纯灰尘虽然严格自验证的正确率也能接近三分之一,但也有 15% 左右的样本点被误分类为冰晶粒子,超过一半的样本点被认为是无效数据;而其余类型的灰霾则很难相互区分。

单波长偏振 Mie 散射激光雷达的识别能力更差。其能准确实现再识别的也仅有冰晶粒子和灰尘混合物,对其余种类灰霾都不能很好地进行分类,特别是会将绝大部分的海洋气溶胶、海洋污染物、城市气溶胶、生物燃烧气溶胶和新鲜烟雾样本点判为无效数据。

当不设判断阈值,将决策规则修改为选取相对概率最大的灰霾类型为待判断灰霾样本的组分识别结果,则各种类灰霾自验证结果如表 4 和图 7 所示。

表 4 不设置判断阈值时对非双波长偏振 HSRL 八种灰霾自验证的准确率

Table 4 Accuracy rates of eight kinds of hazes self-calibration for non-double-wavelength polarized HSRL without threshold

| Haze type | Lidar type | Dual-wavelength polarized | Single-wavelength | Single-wavelength polarized |
|------------------|-----------------------|---------------------------|-----------------------|-----------------------------|
| | Mie backscatter lidar | polarized HSRL | Mie backscatter lidar | |
| Ice | 80.31% | 86.66% | 76.54% | |
| Pure dust | 69.17% | 95.53% | 38.97% | |
| Mix dust | 98.40% | 98.39% | 98.30% | |
| Maritime | 1.03% | 83.69% | 0 | |
| Marine pollution | 1.20% | 0 | 0 | |
| Urban | 90.60% | 92.39% | 98.92% | |
| Biomass burn | 78.61% | 6.82% | 2.05% | |
| Fresh smoke | 45.57% | 81.97% | 0 | |

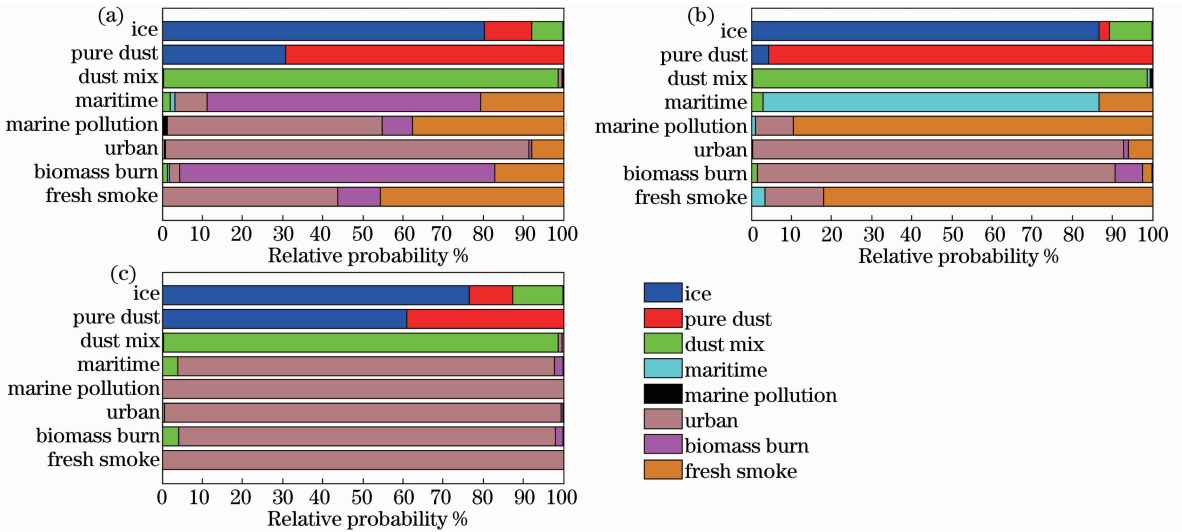


图 7 不设置判断阈值时对非双波长偏振 HSRL 八种灰霾自验证的详细分析结果。(a) 双波长偏振 Mie 散射激光雷达八种灰霾自验证的准确率;(b) 单波长偏振 HSRL 八种灰霾自验证的准确率;(c) 单波长偏振 Mie 散射激光雷达八种灰霾自验证的准确率

Fig. 7 Analysis results of eight kinds of hazes self-calibration for non-ouble-wavelength polarized HSRL without threshold. (a) Accuracy rates of eight kinds of hazes self-calibration for double-wavelength polarized Mie scattering lida without threshold; (b) accuracy rates of eight kinds of hazes self-calibration for single-wavelength polarized HSRL lidar without threshold; (c) accuracy rates of eight kinds of hazes self-calibration for single-wavelength polarized Mie scattering lidar without threshold

由仿真结果可以看出,当不设定判断阈值时双波长偏振 Mie 激光雷达和单波长偏振 HSRL 对各种类灰霾的自验证正确率都有了一定程度的提高,但各种类灰霾之间的串扰也都有更加严重的恶化,如前所述这样并不能很好地提高激光雷达大气遥感灰霾组分识别结果的正确性。但对单波长偏振 Mie 散射激光雷达来说,虽然各种类灰霾之间的串扰大大恶化,但其对纯灰尘类型识别的正确率有了明显的提高,故对单波长偏振 Mie 散射激光雷达而言或许不设定判断阈值更好。

4.4 讨论

需要注意的是,仿真实验里的数据库采用 Burton 等的研究成果并结合美国国家航空航天局兰利研究中心(NASA LaRC)的 18 次北美洲野外调查实验得到的数据库,该数据库主要着眼于全球范围内普遍存在的类型。而针对于某一种特定类型灰霾的再细分,如对城市气溶胶进一步细分为汽车尾气气溶胶、居民生活气溶胶、工业气溶胶等的研究将是下一部分的工作。同时该模型能否继续很好地运用于再细分后的灰霾亚类之间的判别也将有待于进一步验证。还需注意的是,仿真中所用数据库中数据皆为 532 nm 和 1064 nm 两个波长上的灰霾光学参数。但从仿真中可以发现,只要可以在两个波长上对灰霾气溶胶进行高精度遥感,得到退偏比、激光雷达比、谱退偏比和后向散射颜色比即可,至于是否一定是 532 nm 和 1064 nm 则没有限制。所以,原则上来说,355 nm 和 532 nm 的偏振 HSRL 也是可以实现准确的激光雷达遥感灰霾的识别分类的。

5 结论

提出了基于贝叶斯模式识别的激光雷达遥感灰霾识别分类模型,探讨了激光雷达大气遥感灰霾组分的可行性问题。该模型具有较好的泛化能力,可适用于各种不同雾霾数据库以及分类方式。具体论述了相关模式识别特征样本库、特征向量以及模式识别模型的构造,并进行了计算机仿真。通过相应的自验证检验发现在灰霾模式识别特征样本库的样本数足够的情况下该模型有良好的稳定性,同时指出了各种类灰霾粒子再识别的准确率和各种灰霾粒子之间的相互串扰情况,提出并证明了判断阈值在避免这种串扰中的作用。

此外还详细讨论了降维情况下灰霾识别模型的适用性。通过分析发现,单波长偏振 HSRL 对灰霾组分的识别能力相对最强,双波长偏振 Mie 散射激光雷

达次之,单波长偏振 Mie 散射激光雷达最弱。其中单波长偏振 HSRL 能对冰晶粒子、纯灰尘、灰尘混合物和海洋气溶胶进行较为准确的判断;双波长偏振 Mie 散射激光雷达也能较好地对冰晶粒子、纯灰尘和灰尘混合物进行识别,但单波长偏振 Mie 散射激光雷达只能较好地实现冰晶粒子和灰尘混合物的识别。在设定判断阈值的情况下会有较多的数据样本被判定为无效数据,但当不设定判断阈值时则会存在很严重的串扰。其中双波长偏振 Mie 散射激光雷达和单波长偏振 HSRL 可以通过选择适当的判断阈值来减少各种类灰霾之间的串扰,但对单波长偏振 Mie 散射激光雷达而言,可以通过较大程度地减小判断阈值甚至不设定判断阈值来提高纯灰尘识别的准确率。通过对这些降维情况进行分析,凸显了偏振 HSRL,尤其是双波长偏振 HSRL 在激光雷达大气遥感灰霾组分识别应用中的巨大优势。

参 考 文 献

- 1 Wu Dui. A discussion on difference between haze and fog and warning of ash haze weather[J]. *Meteorology*, 2005, 31(4): 3-7.
吴 兑. 关于霾与雾的区别和灰霾天气预警的讨论[J]. *气象*, 2005, 31(4): 3-7.
- 2 Cao Pan. The current situation and prospects of haze monitoring [J]. *Public Communication of Science & Technology*, 2013, (1): 88.
曹 攀. 灰霾监测研究现状及展望[J]. *科技传播*, 2013, (1): 88.
- 3 Mao Yuqing, Sun Yan, Jiang Aijun, *et al.*. An experimental study of haze prediction method in Nanjing [J]. *Climatic and Environmental Research*, 2011, 16(3): 273-279.
毛宇清, 孙 燕, 姜爱军, 等. 南京地区霾预报方法试验研究 [J]. *气候与环境研究*, 2011, 16(3): 273-279.
- 4 Huang Zhengxu, Gao Wei, Dong Junguo, *et al.*. Development of a real-time single particle aerosol time-of-flight mass spectrometer [J]. *Journal of Chinese Mass Spectrometry Society*, 2010, 31(6): 331-336.
黄正旭, 高 伟, 董俊国, 等. 实时在线单颗粒气溶胶飞行时间质谱仪的研制[J]. *质谱学报*, 2010, 31(6): 331-336.
- 5 Mi Yunping, Wang Xiaoping, Zhan Shuyue. Review on cavity ring down spectroscopy technology and its application [J]. *Optical Instruments*, 2007, 29(5): 85-89.
宓云平, 王晓萍, 詹舒越. 光腔衰荡光谱技术及其应用综述 [J]. *光学仪器*, 2007, 29(5): 85-89.
- 6 Zhang Wenyu, Xin Jinyuan, Yuan Jiuyi, *et al.*. Multi-band remote sensing study on aerosol optical depth in Tengger desert [J]. *Plateau Meteorology*, 2003, 22(6): 613-617.
张文煜, 辛金元, 袁九毅, 等. 腾格里沙漠气溶胶光学厚度多波段遥感研究 [J]. *高原气象*, 2003, 22(6): 613-617.
- 7 Zhou Jun, Yue Guming, Qi Fudi, *et al.*. Optical properties of aerosol derived from lidar measurements [J]. *Chinese Journal of Quantum Electronics*, 1998, 15(2): 140-148.
周 军, 岳古明, 戚福第, 等. 大气气溶胶光学特性激光雷达探测 [J]. *量子电子学报*, 1998, 15(2): 140-148.
- 8 Z Tao, D Liu, Z Zhong, *et al.*. Measurements of cirrus clouds with a three-wavelength lidar [J]. *Chin Opt Lett*, 2012, 10(5):

- 050101.
- 9 Ji Chengli, Tao Zongming, Hu Shunxing, *et al.*. Cirrus measurement using three-wavelength lidar in Hefei [J]. *Acta Optica Sinica*, 2014, 34(4): 0401001.
季承荔, 陶宗明, 胡顺星, 等. 三波长激光雷达探测合肥地区卷云特性[J]. *光学学报*, 2014, 34(4): 0401001.
- 10 Li Xuebin, Xu Qingshan, Wei Heli, *et al.*. Study on relationship between extinction coefficient and mass concentration [J]. *Acta Optica Sinica*, 2008, 28(9): 1655–1658.
李学彬, 徐青山, 魏合理, 等. 气溶胶消光系数与质量浓度的相关性研究[J]. *光学学报*, 2008, 28(9): 1655–1658.
- 11 A Ångström. The parameters of atmospheric turbidity [J]. *Tellus*, 1964, 16(1): 64–75.
- 12 N Sugimoto, C H Lee. Characteristics of dust aerosols inferred from lidar depolarization measurements at two wavelengths [J]. *Appl Opt*, 2006, 45(28): 7468–7474.
- 13 T Murayama, J M Masonis, J Redemann, *et al.*. An intercomparison of lidar-derived aerosol optical properties with airborne measurements near Tokyo during ACE-Asia [J]. *Journal of Geophysical Research*, 2003, 108(D23): 8651.
- 14 T Sakai, T Nagai, Y Zaizen, *et al.*. Backscattering linear depolarization ratio measurements of mineral, sea-salt, and ammonium sulfate particles simulated in a laboratory chamber [J]. *Appl Opt*, 2010, 49(23): 4441–4449.
- 15 T Murayama, H Okamoto, N Kaneyasu, *et al.*. Application of lidar depolarization measurement in the atmospheric boundary layer: Effects of dust and sea-salt particles [J]. *Journal of Geophysical Research*, 1999, 104(D24): 31781–31792.
- 16 J Ackermann. The extinction-to-backscatter ratio of tropospheric aerosol: A numerical study [J]. *Journal of Atmospheric & Oceanic Technology*, 1998, 15(4): 1043–1050.
- 17 C Catrall, J Reagan, K Thome, *et al.*. Variability of aerosol and spectral lidar and backscatter and extinction ratios of key aerosol types derived from selected Aerosol Robotic Network locations [J]. *Journal of Geophysical Research*, 2005, 110 (D10): D10S11.
- 18 D Müller, A Ansmann, I Mattis, *et al.*. Aerosol-type-dependent lidar ratios observed with Raman lidar [J]. *Journal of Geophysical Research*, 2007, 112(D16): D16202.
- 19 Huang Hanlu, Liu Dong, Yang Yongying, *et al.*. Design of the high spectral resolution lidar filter based on a field-widened Michelson interferometer [J]. *Chinese J Lasers*, 2014, 41(9): 0913003.
黄寒璐, 刘东, 杨甬英, 等. 基于视场展宽迈克尔逊干涉仪的高光谱分辨率激光雷达滤光器设计研究 [J]. *中国激光*, 2014, 41(9): 0913003.
- 20 Cheng Zhongtao, Liu Dong, Luo Jing, *et al.*. Influences analysis of the spectral filter transmission on the performance of high-spectral-resolution lidar [J]. *Acta Optica Sinica*, 2014, 34(8): 0801003.
成中涛, 刘东, 罗敬, 等. 光谱滤光器透射率参数对高光谱分辨率激光雷达的影响分析 [J]. *光学学报*, 2014, 34(8): 0801003.
- 21 D Liu, Y Yang, Z Cheng, *et al.*. Retrieval and analysis of a polarized high-spectral-resolution lidar for profiling aerosol optical properties [J]. *Opt Express*, 2013, 21(11): 13084–13093.
- 22 S P Burton, R A Ferrare, C A Hostetler, *et al.*. Aerosol classification using airborne High Spectral Resolution Lidar measurements—methodology and examples [J]. *Atmospheric Measurement Techniques*, 2012, 5(1): 73–98.
- 23 A H Omar, J G Won, D M Winker, *et al.*. Development of global aerosol models using cluster analysis of Aerosol Robotic Network (AERONET) measurements [J]. *Journal of Geophysical Research*, 2005, 110(D10): D10S14.

栏目编辑: 王晓琰