

# 基于混合智能算法的激光雷达数据拼接技术

段绿林, 刘 东, 张与鹏, 成中涛, 罗 敬, 杨甬英, 沈亦兵, 白 剑

浙江大学光电科学与工程学院现代光学仪器国家重点实验室, 浙江 杭州 310027

**摘要** 双通道数据拼接技术能有效地增大激光雷达系统的动态探测范围,改善大气回波信号的信噪比。提出了一种基于非支配排序遗传算法 II(NSGA-II)与邻域粗糙集(NRS)的大气遥感激光雷达数据拼接智能算法,即NRSWNSGA-II,提高了数据拼接的准确性和稳定性。该算法以三个判断双通道数据拟合优度的评价函数为优化目标,通过NSGA-II获得评价函数的Pareto最优解集,进而利用NRS训练数据样本得到的权重进行线性规划,实现了最优拟合范围内的全局随机搜索。实验结果表明,所提算法拼接效果良好并在全天性数据拼接工作中有较好的稳定性。

**关键词** 大气光学; 激光雷达; 数据拼接; 非支配排序遗传算法 II; 邻域粗糙集; 评价函数

**中图分类号** TN958.98 **文献标识码** A

**doi:** 10.3788/AOS201737.0601002

## Lidar Data Gluing Technology Based on Hybrid Intelligent Algorithm

Duan Lulin, Liu Dong, Zhang Yupeng, Cheng Zhongtao, Luo Jing,  
Yang Yongying, Shen Yibing, Bai Jian

*State Key Laboratory of Modern Optical Instrumentation, College of Optical Science and Engineering,  
Zhejiang University, Hangzhou, Zhejiang 310027, China*

**Abstract** Two-channel data gluing technology can extend the dynamic detection range of lidar system effectively and improve the signal to noise ratio of atmospheric echo signal. A lidar data gluing technology based on the non-dominated sorting genetic algorithm II (NSGA-II) and neighborhood rough set (NRS) is proposed to improve the accuracy and stability of data gluing. Three evaluation functions determining goodness of fitting are taken as objective functions and the Pareto optimal solution set is acquired by NSGA-II. The weight which is acquired by NRS is applied to accomplishing the global random searching for the best fitting range. The experimental results show that the proposed algorithm has a favorable gluing result and works steadily in gluing tasks all day.

**Key words** atmospheric optics; lidar; data gluing; non-dominated sorting genetic algorithm II; neighborhood rough set; evaluation function

**OCIS codes** 010.0280; 280.3640; 010.3640; 200.4560

## 1 引 言

作为一种主动遥感技术,激光雷达具有高时空分辨率、可连续监测等优势<sup>[1]</sup>,已经广泛地应用于对大气的探测和研究。激光雷达通过发射激光束与大气中微粒发生散射、退偏等相互作用得到大气回波信号,经过反演就可以获得大气的消光系数、后向散射系数等重要光学参数。大气中低空边界层的回波信号非常强,而更高的对流层、平流层的回波信号由于探测距离增加而逐渐减弱。为了扩展系统探测动态范围,实现对高低空的同时探测,双通道信号数据拼接技术被提出并得到广泛应用<sup>[2-4]</sup>。该技术将高低空探测任务分成双通道进行,其中低空通道可防止底层大气强信号导致的饱和(非线性)现象,而高空通道可实现对高层大气微弱信

**收稿日期:** 2016-11-30; **收到修改稿日期:** 2017-02-13

**基金项目:** 国家重点研发计划(2016YFC0200700)、国家自然科学基金(41305014,61475141)、浙江省公益技术研究项目(2016C33004)、现代光学仪器国家重点实验室创新基金(MOI2015QN01)

**作者简介:** 段绿林(1989—),男,硕士研究生,主要从事激光雷达大气遥感方面的研究。E-mail: matthew\_wei@126.com

**导师简介:** 刘 东(1982—),男,博士,副教授,主要从事光电检测与遥感技术方面的研究。E-mail: liudongopt@zju.edu.cn (通信联系人)

号的探测,从而增大了系统探测距离,提高了反演精度<sup>[5]</sup>。探测到的两组信号数据经拼接算法拼接成动态范围大、信噪比良好的回波信号。

目前主要有两种方式可以实现双通道探测,第一种是分光技术,利用两个探测器分别完成高、低空探测任务,并由双望远镜接收回波信号<sup>[2]</sup>或者利用分光片匹配门控技术实现分光;第二种是电路分离技术<sup>[6-7]</sup>,对一个探测器的电信号进行分离并分别采用光子计数(PC)和模数采集(AD),PC技术能够有效抑制探测器噪声,实现高空微弱信号的探测。数据拼接需要从双通道信号数据中选择合适的数据,利用线性回归建立两者的关系,一般为高空通道信号向低空通道信号拟合回归。选择对应的拼接点<sup>[3]</sup>或拼接范围<sup>[8]</sup>,拼接低空通道数据和拟合后的高空通道数据。可以看出,参与拟合的高低空数据对的选取方法是拼接算法的核心问题,这也是国内外研究的主要内容。目前有两种约束进行数据对的选取,即强度约束和高度约束。为避免统计偏差,强度约束对象应为参与拟合的自变量<sup>[3]</sup>。强度上限通常为常数,其值由探测器的非线性程度决定;强度下限可由低空通道的信噪比程度得到<sup>[9]</sup>,或者利用背景噪声进行动态补偿<sup>[3,10]</sup>。Zhang等<sup>[10]</sup>采用强度约束时,发现当参与拟合数据对的相关系数下降时,回归将非常不稳定。2016年,D'Amico等<sup>[4]</sup>利用高度约束依次对回归模型进行回归系数、稳定性判定,以特定步长逐渐缩小高度约束范围,获得了更加可信的拟合范围。黄立峰等<sup>[8]</sup>则在由经验给定的高度范围内,以固定长度的拟合区域进行位置遍历(滑窗步长约为100 m),对不同位置得到的相关系数等参数进行线性规划,获得最优的区间位置。除了线性回归方法,利用卤灯标定回归系数<sup>[11]</sup>或统计学原理<sup>[12]</sup>也能够实现数据拼接,但实现方法相对复杂或者缺乏实验基础。

面对如何选取合适的拟合范围的问题,研究人员之前只是利用背景噪声补偿<sup>[3,10]</sup>的强度约束或者以固定步长对高度约束范围进行操作<sup>[4,8]</sup>来选取拟合数据,并且衡量拟合范围优良度的标准并不全面,导致这些算法对拟合范围参数与步长较为敏感,拼接结果并不稳定。本文基于邻域粗糙集(NRS)<sup>[13]</sup>权值确定的非支配排序遗传算法 II(NSGA-II)<sup>[14]</sup>,提出了一种混合智能算法,即 NRSWNSGA-II 算法,能够实现自适应全局随机搜索,综合考虑回归相关系数  $R$ 、回归方程稳定性参数  $S$  以及通道增益比偏差  $T$ ,避免了上述问题,提高了数据拟合与拼接结果的准确性和稳定性。

## 2 NRSWNSGA-II 原理与流程

### 2.1 总算法流程

图 1 为所提 NRSWNSGA-II 算法的原理简图。该算法总体思路是先利用 NSGA-II 全局随机搜索出以回归相关系数  $R$ 、回归方程稳定性参数  $S$  为目标函数的拟合范围内 Pareto 最优解集  $X_p$ ,再利用  $R$ 、 $S$  与通道增益比偏差  $T$  的线性加权来寻找最优拟合约束范围。其中,权重  $\omega$  是利用 NRS 训练样本数据预先得到的,能全面判定拼接解的优良程度。NSGA-II、 $R$ 、 $S$  与  $X_p$  将在 2.2 节进行介绍,NRS 与  $T$  将在 2.3 节进行介绍。

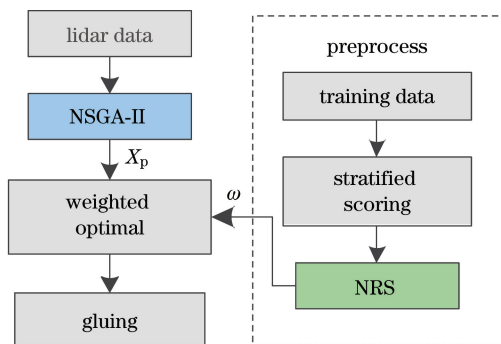


图 1 NRSWNSGA-II 拼接算法简图

Fig. 1 Schematic of NRSWNSGA-II gluing algorithm

### 2.2 NSGA-II 多目标优化

NSGA-II 是由 Deb 等<sup>[14]</sup>于 2002 年提出的一种遗传算法,是应用最为广泛、最成功的多目标优化算法。NSGA-II 的基本思想是在给定的强度范围内随机产生待评价拟合范围的上下限,每个待评价拟合范围称为

个体,个体数量称为种群数目。分别对每个个体计算  $R$ 、 $S$ ,从中选择合适的父代,通过交叉、变异的遗传操作产生子代,再淘汰子代种群中  $R$ 、 $S$  评价差的个体,经过不断的交叉、变异与淘汰,个体  $R$ 、 $S$  将逐渐变优,在满足终止条件时算法结束,从终止种群中选择一部分最优秀的个体作为 NSGA-II 得到的 Pareto 最优解集。

### 2.2.1 种群初始化与评价函数

对种群个体进行一定的约束,主要包括高空通道强度约束、参与拟合数据点数量约束以及回归相关系数约束,即

$$\begin{cases} H_{\min} < H_{n,k} < H_{\max} \\ P_{\min} < P < P_{\max} \\ R_{\min} < R \end{cases}, \quad (1)$$

式中  $H_{n,k}$  为高空通道数据,下标  $n$  和  $k$  表示高度和时间; $H_{\min}$ 、 $H_{\max}$  为强度约束条件,需要考虑背景噪声和探测器的非线性<sup>[3,10]</sup>;  $P$  为参与拟合数据对数量; $P_{\min}$ 、 $P_{\max}$  和  $R_{\min}$  为对应的约束条件,在不满足个体的评价函数中加入惩罚因子实现,在算法结束后评价函数将还原。

回归相关系数  $R$  是衡量参与拟合的数据之间的相关程度和判定拟合优度的参数,也是诸多研究者观察或判断拼接结果优劣的参数之一<sup>[8,10]</sup>,其定义为

$$R = \frac{\sum_k (H_{n,k} - \bar{H}_n)(L_{n,k} - \bar{L}_n)}{\sqrt{\sum_k (H_{n,k} - \bar{H}_n)^2 \sum_k (L_{n,k} - \bar{L}_n)^2}}, \quad (2)$$

式中  $L_{n,k}$  为低空通道数据, $\bar{H}_n$ 、 $\bar{L}_n$  分别为参与拟合的高低空通道数据的平均值。回归相关系数的下降表明数据之间偏离线性,拟合将导致较大偏差。

回归模型检测是回归分析的重要内容,采用的回归方程稳定性参数  $S$  为 D'Amico 等<sup>[4]</sup> 检测回归模型系数时采用的等级因子。将待检测数据对等分成两个区间,分别将这两个区间的双通道数据进行线性拟合,拟合得到的残差  $R_{es}$  再分别与距离  $z$  进行拟合,即  $R_{es} = \rho z$ ,求得拟合的斜率  $\rho_1$ 、 $\rho_2$  和标准差  $\Delta\rho_1$ 、 $\Delta\rho_2$ ,则回归方程稳定性参数为

$$S = \frac{|\rho_1 - \rho_2|}{\sqrt{\Delta\rho_1^2 + \Delta\rho_2^2}}. \quad (3)$$

对于高斯分布,当  $S$  增大为 1 时,在统计学意义上  $\rho_1$ 、 $\rho_2$  只有 32% 的概率是吻合的<sup>[4]</sup>。

### 2.2.2 非支配序值与拥挤距离

基于  $R$ 、 $S$  值对每个个体进行优劣排序。对于多个评价函数而言,个体的优劣性由非支配序值表征,非支配序值是对种群的分层。支配是多目标优化问题中的概念,其含义如下,假设  $X_A$  和  $X_B$  为满足约束条件的可行域的两个解,对于评价函数  $f(x)$ ,当且仅当

$$\forall i=1,2,\dots,I, f_i(X_A) \leq f_i(X_B) \wedge \exists j=1,2,\dots,I, f_j(X_A) < f_j(X_B), \quad (4)$$

称  $X_A$  支配  $X_B$ ,记做  $X_A > X_B$ 。当一个解不被其他任何解所支配时,称此解为非支配解。非支配序值是通过非支配排序得到的,其步骤为:找寻得到种群中的非支配解,得到第一层非支配解,序值为 1;将序值为 1 的解从种群中除去,再对种群进行第二轮非支配解找寻,得到第二层非支配解,序值为 2;如此反复,直到种群中所有的解都已被标记。对同一层的解,用拥挤距离判定优劣。可依据评价函数的  $R$  对个体进行大小排序,前后个体的  $R$  之差除以种群中的最大、最小值之差  $R'_{\max} - R'_{\min}$  即为该个体关于  $R$  的拥挤程度,最首、最末个体的拥挤程度为无穷大。 $R$ 、 $S$  拥挤程度之和为个体总拥挤程度,其值越大个体评价越优。

### 2.2.3 遗传操作与 Pareto 最优解集

算法采用的选择机制为锦标赛原则,选择随机提取的两个体中更加优良的个体为父代进行交叉和变异操作。交叉算子选择中间交叉,子代的拟合强度上下限将随机生成在两父代之间。变异算子选择自适应变异,父代强度上下限以随机方向与随机步长产生子代强度上下限,并且子代满足边界约束。进行选择、交叉和变异后,种群数量增长为原来的两倍。此时需要对种群进行修剪,淘汰较差的解,使种群数量恢复原始数量。

通过不断地进行选择、交叉、变异和修剪,NSGA-II 将不断逼近以  $R$ 、 $S$  为目标函数的理论 Pareto 最优解集。Pareto 最优解集定义为可行域内非支配解的集合,在本文中可行域为边界约束。当满足终止条件(最大进化代数)时算法终止并输出最优个体群,其数目由种群数目比值和最优前端个体系数决定。

### 2.3 NRS 的权重训练

NRS 通过挖掘样本数据的属性来确定属性特征值权重<sup>[13]</sup>。随机从不同日期激光雷达信号的 Pareto 最优解集中选取  $q$  个训练样本  $U = \{x_1, x_2, \dots, x_q\}$ , 取  $q = 2100$ , 得到样本评价函数值矩阵  $\mathbf{A}$ , 其包含  $R$ 、 $S$  和  $T$ 。通道增益比偏差  $T$  定义为

$$T = \frac{|s - s_c|}{s_c}, \quad (5)$$

式中  $s$  为样本的增益比(拟合斜率),  $s_c$  为标定的增益比常数,理论上高低空通道增益比在全天性工作时应该保持稳定<sup>[10]</sup>。对三个条件属性分别进行评价计分,具体划分如表 1 所示,例如  $R$  大于 99% 时有关  $R$  的得分为 0, 而当  $R$  为 98.6% 时有关  $R$  的得分为 1, 其他两个评价函数与之类似,得分越小其解越优。

将属性得分相加得到总评价得分即决策向量(决策属性)  $\mathbf{E}$ , 得分越小其解越优。利用最大最小值法将  $\mathbf{A}$  进行归一化, 即

$$\tilde{\mathbf{A}}_{u,v} = \frac{\mathbf{A}_{u,v} - \min(\mathbf{A}_v)}{\max(\mathbf{A}_v) - \min(\mathbf{A}_v)}, \quad (6)$$

式中  $\max(\mathbf{A}_v)$ 、 $\min(\mathbf{A}_v)$  为对应评价函数的最大、最小值;  $v = 1, 2, 3$ ;  $u = 1, 2, \dots, q$ 。评价函数  $a_v \in \{R, S, T\}$  的邻域半径为

$$\delta_{a_v} = S_{td}(\tilde{\mathbf{A}}_v) / \lambda, \quad (7)$$

式中  $S_{td}(\tilde{\mathbf{A}}_v)$  为对应  $a_v$  的标准差;  $\lambda$  为特征参数, 可以调整邻域分类精度, 取  $\lambda = 3$ 。

表 1 单因素分层计分标准

Table 1 Criterion of single factor stratified scoring

Property point	$R$ / %	$S$	$T$ / %
0	[99, 100]	[0, 0.1]	[0, 1]
1	[98, 99)	(0.1, 0.2]	(1, 2.5]
2	[96, 98)	(0.4, 0.2]	(2.5, 4]
3	[92, 96)	(0.6, 0.4]	(4, 6]
4	Else	Else	Else

由邻域半径可以得到样本  $x_u$  各评价函数邻域集合为

$$\delta_{a_v}(x_u) = \{x_h \mid x_h \in U, |\tilde{\mathbf{A}}_{h,v} - \tilde{\mathbf{A}}_{u,v}| < \delta_{a_v}\}. \quad (8)$$

对于类似  $B_{12} = \{R, S\}$  评价集的邻域集合为  $\delta_{B_{12}}(x_u) = \delta_R(x_u) \cap \delta_S(x_u)$ 。根据不同的得分,  $\mathbf{E}$  将  $U$  划分成  $N$  个等价类  $Y_1, Y_2, \dots, Y_N$ 。可以求得对  $\forall B \subseteq \{R, S, T\}$  关于  $\mathbf{E}$  的下近似(正域)为

$$\underline{N}_B \mathbf{E} = \bigcup_{m=1}^N \underline{N}_B Y_m, \quad (9)$$

式中

$$\underline{N}_B Y_m = \{x_u \mid \delta_B(x_u) \subseteq Y_m, x_u \in U\}. \quad (10)$$

进一步可以得到决策属性  $\mathbf{E}$  对  $B$  的依赖度为

$$\gamma_B(\mathbf{E}) = \text{card}(\underline{N}_B \mathbf{E}) / \text{card}(U), \quad (11)$$

式中  $\text{card}$  为集合元素个数的算符。对任一个评价函数  $a_v$ , 其属性重要度为

$$S_{\text{ig}}(a_v, \mathbf{E}) = \gamma_{\{R, S, T\}}(\mathbf{E}) - \gamma_{\{R, S, T\} - a_v}(\mathbf{E}). \quad (12)$$

由此, 可以求得每个评价函数的权值为

$$\omega(a_v) = \frac{S_{\text{ig}}(a_v, \mathbf{E})}{\sum_{v=1}^3 S_{\text{ig}}(a_v, \mathbf{E})}. \quad (13)$$



采用 NRS 训练的权重能全面评价样本的优良度,并能再精细分离同样得分的样本。

### 2.4 最终解的确定与拼接

利用加权的权重对 Pareto 最优解集进行收敛,通过最小化

$$D(X) = \omega(R)\tilde{R}(X) + \omega(S)\tilde{S}(X) + \omega(T)\tilde{T}(X) \quad (14)$$

筛选出合适的最优解,式中  $\tilde{R}(X)$ 、 $\tilde{S}(X)$  与  $\tilde{T}(X)$  为解集评价函数的线性归一化值。拼接处理采用的是基于分界点将拼接信号分成两部分的方法,减小了其他处理形成的数据失真对激光雷达数据反演的影响。采用的拼接点为最优解中强度上限  $H_m$ ,最终拼接信号为

$$G_{lu} = \begin{cases} \hat{H}_{n,k} & \text{for } H_{n,k} < H_m, \\ L_{n,k} & \text{for } H_{n,k} \geq H_m \end{cases}, \quad (15)$$

式中  $\hat{H}_{n,k}$  为拟合之后的远场信号,可通过将最优解得到的拟合系数和截距作用于整个远场信号得到。 $L_{n,k}$  避免了近场饱和现象, $\hat{H}_{n,k}$  在高空仍然信噪比良好。

## 3 实验与讨论

采用的数据来源于威斯康辛大学激光雷达组设于新加坡的 AHSRL(Arctic High-Spectral Resolution Lidar)系统采集于 2013 年 7 月的数据,AHSRL 系统位于东经 103.78°、北纬 1.29°,当地时间(LT)为 UTC 时间加 8(下文采用 UTC 时间),AHSRL 采用 532 nm 波长的激光,采集的双通道数据都是 PC 类型,空间分辨率为 7.5 m,时间分辨率为 30 s,最大探测高度为 30 km。

### 3.1 算法参数的确定

在进行数据处理前需要确定算法参数,算法运行参数如表 2 所示。参与拟合的高空通道数据的强度范围为优化目标(变量)。最优前沿个体系数的取值范围为 0~1,与种群数量共同决定最终 Pareto 最优解集个数。不断测试算法的稳定性与收敛速度后,种群个体数、最大进化代数和前沿个体系数分别设为 100、70 和 0.3。参与拟合数据点的数量综合考虑拟合斜率稳定度,防止算法破坏并防止高空微弱信号干扰,选择了 60 与 1000。 $R$  阈值选为 90%。研究中采用无云晚上累积的双通道增益比 0.0194 为标准增益比。得到 Pareto 最优解集之后,进行最后的加权评价判断,经过 NRS 权重训练得到  $R$ 、 $S$ 、 $T$  的加权系数分别为 0.428、0.188 和 0.384。

表 2 算法运行参数

Table 2 Operating parameters of the algorithm

Parameter	Value	Parameter	Value
Population size	100	Number of generation	70
Pareto front population fraction	0.3	Lower limit of $R$ / %	90
Lower limit of fitting intensity / MHz	0.06	Upper limit of fitting intensity / MHz	10
Lower limit of fitting number	60	Upper limit of fitting number	1000

### 3.2 实验结果与讨论

#### 3.2.1 双通道拼接实验

威斯康辛大学激光雷达组的 AHSRL 系统于 2013 年 7 月 11 日 9:40(UTC)在新加坡测得的回波信号如图 2(a)所示。图中横坐标为海拔高度,纵坐标为回波光光子数。绿色曲线代表高空通道信号,橙色曲线代表低空通道信号,双通道数据都进行了脉冲叠堆校正<sup>[3]</sup>。

图 2(a)显示在 300 m 附近信号出现极大值,高空通道出现了信号饱和现象,因为从标准增益比换算的高空通道信号应该在  $10^{10}$  Hz 以上。在 6 km 高度处可能存在云层,双通道数据的轮廓具有良好的一致性,在 6.4 km 后低空通道信噪比严重下降,导致信号无法使用,但高空通道的高层信号信噪比良好。

经过 NSGA-II 的多目标优化后,得到的 Pareto 最优解集如图 2(b)所示,横坐标是回归相关系数  $R$ ,左纵坐标是稳定性参数  $S$ ,右纵坐标是增益比偏差  $T$ ,橙色曲线表示 Pareto 前沿(即 Pareto 最优解集的函数

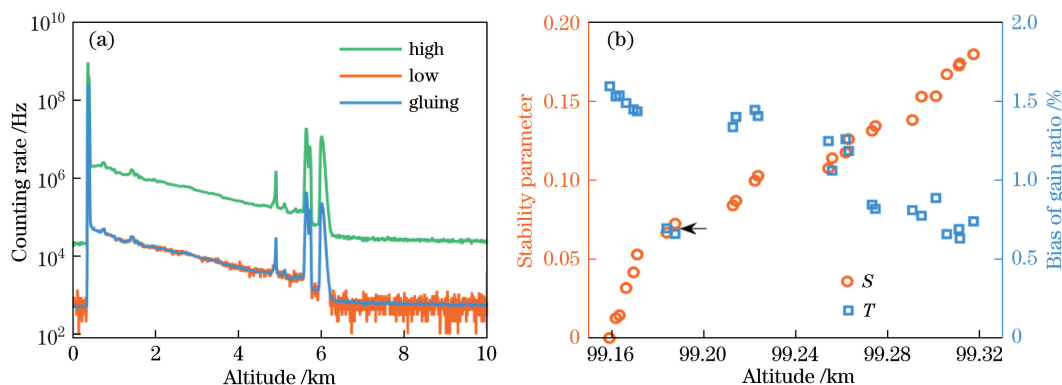


图 2 拼接信息图。(a) AHSRL 系统双通道信号;(b) Pareto 最优解集

Fig. 2 Figure of gluing information. (a) Two-channel signal of AHSRL; (b) Pareto optimal solution set

值),其分布较为均匀,说明算法模型设计良好。可以看出得到的 Pareto 最优解集的回归相关系数在 99% 以上,稳定性参数在 0.2 以下,对应的增益比偏差在 2% 以下。利用 NRS 训练的加权系数对得到的 Pareto 最优解集进行最后的线性寻优,得到的最优解如图 2(b)中箭头所指,其  $R$ 、 $S$ 、 $T$  分别为 99.19%、0.073 和 0.66%,拟合强度区域为 0.327~1.960 MHz,增益比为 0.0193。对比图 2(b)中其他 Pareto 解,最优解的  $R$  值 99.19% 并不突出,此现象与表 1 计分标准有关。所有解的  $R$  都超过了 99%,为 0 分,而  $S$ 、 $T$  有 0 分、1 分两个得分层。最优解的  $R$ 、 $S$ 、 $T$  均属于第一得分层(0 分),且相比临近点最优解, $R$ 、 $T$  表现更好。得到的拼接信号如图 2 中蓝色曲线所示,强信号的轮廓吻合,高空信噪比得到增强。对原始数据重复做 30 次拼接实验,发现拼接点位置一致,且增益比变化量小于 0.6%。

### 3.2.2 全天性拼接实验

为了测试 NRSWNSGA-II 的稳定性,对全天性的数据进行了实验,并与文献[10]提出的算法进行了对比。采用的数据是 AHSRL 系统于 2013 年 7 月 5 日在新加坡采集的数据,当天双通道的背景噪声如图 3 所示,20:00 之后数据质量下降严重,因而弃用。

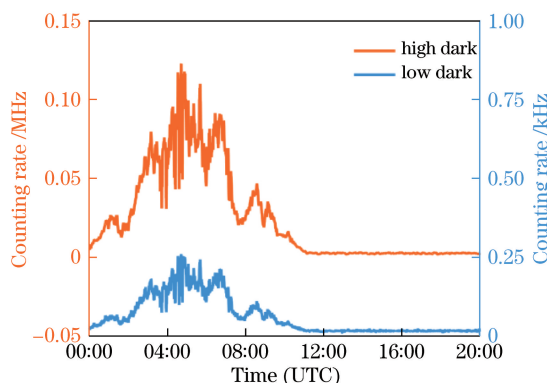


图 3 2013 年 7 月 5 日 AHSRL 在新加坡采集数据的背景噪声

Fig. 3 Background noise of data recorded by AHSRL on 5 July 2013, Singapore

由图 3 可以看出双通道噪声有很大相关性,且在 5:00 左右达到最大,在 10:00 之后逐渐减弱。根据文献[10]提出的算法设定,当高空通道信号作为拟合自变量时拟合强度范围为  $0.02 \text{ MHz} + 2H_k^b < H_{n,k} < 4.2 \text{ MHz}$ ,其中  $H_k^b$  为高空通道背景噪声;而当低空通道信号作为拟合自变量时拟合强度范围为  $0.005 \text{ MHz} + 2L_k^b < L_{n,k} < 0.21 \text{ MHz}$ ,其中  $L_k^b$  为低空通道背景噪声。

利用拟合系数、回归方程稳定性和增益比偏差三个参数,较全面地对比了两种拼接算法的直接效果,也在一定程度上间接对比了反演的大气光学特性的影响。两种算法的结果对比如图 4 所示,从左到右依次为评价函数的  $R$ 、 $S$ 、 $T$ ,需要注意  $R$  的坐标是倒置的。可以看出文献[10]提出的算法中  $R$  在 5:00 时背景光最为强烈,其拟合范围内的数据对已经失去相关性,大部分拟合系数低于 80%,个别甚至低至 40%,且起伏很大;而本文算法  $R$  波动很小,绝大部分在 90% 以上。两者的  $S$  都在 1 之内,文献[10]提出的算法参数

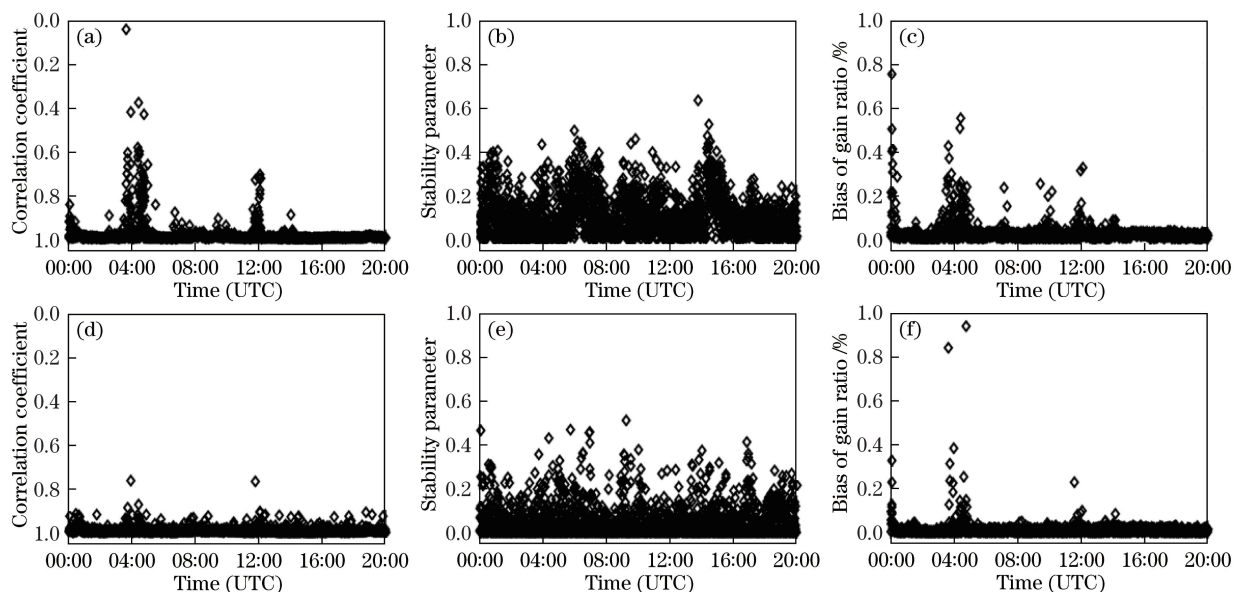


图 4 两种拼接算法结果对比。(a)(b)(c)文献[10]提出算法的回归相关系数、回归方程稳定性参数和增益比偏差；(d)(e)(f)本文算法的回归相关系数、回归方程稳定性参数和增益比偏差

Fig. 4 Comparison of the two gluing methods. (a)(b)(c) Correlation coefficient, stability parameter, bias of gain ratio of the gluing method in Ref. [10], respectively; (d)(e)(f) correlation coefficient, stability parameter, bias of gain ratio of the gluing method of NRSWNSGA-II, respectively

的均值为 0.1,本文算法参数的均值为 0.04,并且更加集中在小于 0.2 的范围之内,说明拟合时的回归模型偏差相对较小。

增益比偏差相对较为重要,虽然双方求得平均斜率都是 0.0194,但在增益比偏差小于 0.05% 和 1% 时,文献[10]方法满足要求的拼接数占全天数据总数的 10.1% 和 27.5%,而 NRSWNSGA-II 分别为 33.7% 和 60.5%,并且增益比偏差小于 3% 的拼接数占全天数据总数的 96%,NRSWNSGA-II 的全局寻优效果得到了体现。同时在 5:00 时刻,NRSWNSGA-II 算法的解对应的  $R$  和  $S$  良好,但得到的  $T$  发生较严重偏离。选取偏差最大的时间点进行个案研究,原始信号以及拼接结果如图 5 所示。

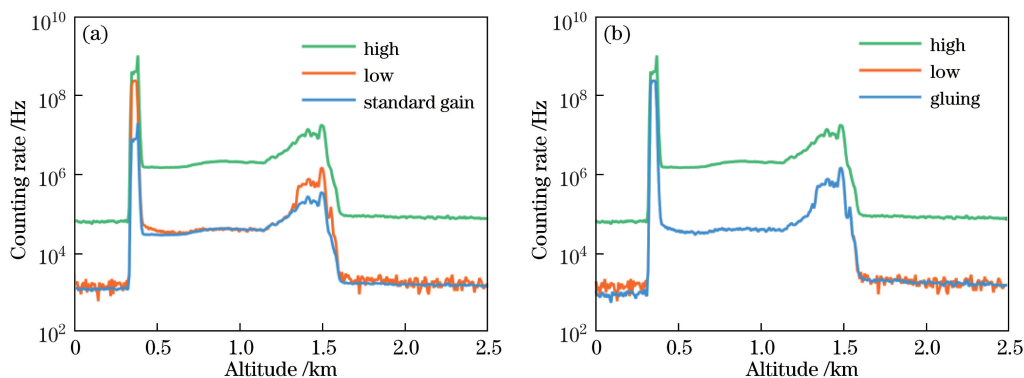


图 5 信号分析与拼接结果。(a) 原始信号和标准增益比拟合图;(b) NRSWNSGA-II 拼接结果  
Fig. 5 Signal analysis and gluing result. (a) Original data and fitting data applying standard gain ratio; (b) gluing result of NRSWNSGA-II

图 5(a)是双通道信号和乘以标准增益比系数的高通道信号,数据对在 0.5 km 和 1.5 km 处出现较严重偏差(在线性范围内),这是由基线作用与脉冲后效应导致的<sup>[15]</sup>。图 5(b)为 NRSWNSGA-II 拼接结果图,可以看出拼接结果良好,只在 400 m 以下发现偏离,采用的拟合范围避开了偏差较大的区域。在实际拼接中,由于数据的不匹配,有一种折中的拼接办法,在背景噪声最大的中午建议用系统标准的增益比实现强行拼接<sup>[3,10]</sup>。

此外,利用 3.6 GHz 主频的计算机对两种算法进行 50 次的单廓线拼接实验以测试算法效率,文献[10]所提算法的运行时间范围为 0.013~0.019 s,平均值为 0.014 s;NRSWNSGA-II 运行时间范围为 1.579~2.997 s,平均值为 2.191 s。NRSWNSGA-II 利用迭代进行寻优,在一定程度上增加了算法复杂度,但通过并行计算、优化程序、将所采用的 Matlab 平台转化为 C 或 Python 等方法进一步优化,NRSWNSGA-II 在大量激光雷达数据集的快速拼接与批量处理中也能获得可接受的算法效率。

## 4 结 论

基于混合人工智能算法 NRSWNSGA-II 的激光雷达数据拼接技术进行多目标优化,实现了对最优拟合区域的自动搜索。NRSWNSGA-II 先利用优化的 NSGA-II 得到回归相关系数和回归方程稳定性参数的 Pareto 最优解集,得到的 Pareto 前沿分布均匀,数学模型的设计符合实际。而由邻域粗糙集训练得到的加权系数具有很好的区分度,能够选择出三个评价参数全面优异的解。对比文献[10]的拼接算法,NRSWNSGA-II 得到的拼接方案在回归相关系数、回归方程稳定度及增益比偏差方面更令人满意,实现了良好稳定的激光雷达数据拼接,为兼顾近地面气象和高空气象活动提供了良好的解决方法。进一步的研究可以尝试在双通道数据存在偏差的情况下,选择高度范围作为优化变量进行拟合拼接。将人工智能算法结合激光雷达的信号处理具有挑战且引人入胜,更多智能算法的尝试将为数据处理与分析提供全新的解决方法。

### 致谢

感谢威斯康辛大学激光雷达组所提供的 AHSRL 系统的双通道数据。

### 参 考 文 献

- [1] Liu Dong, Yang Yongying, Zhou Yudi, *et al.* High spectral resolution lidar for atmosphere remote sensing: A review[J]. *Infrared and Laser Engineering*, 2015, 44(9): 2535-2546.  
刘 东, 杨甬英, 周雨迪, 等. 大气遥感高光谱分辨率激光雷达研究进展[J]. *红外与激光工程*, 2015, 44(9): 2535-2546.
- [2] Li Jun, Gong Wei, Mao Feiyue, *et al.* Dual field of view lidar for observing atmospheric aerosols over Wuhan[J]. *Acta Optica Sinica*, 2013, 33(12): 1201001.  
李 俊, 龚 威, 毛飞跃, 等. 探测武汉上空大气气溶胶的双视场激光雷达[J]. *光学学报*, 2013, 33(12): 1201001.
- [3] Newsom R K, Turner D D, Mielke B, *et al.* Simultaneous analog and photon counting detection for Raman lidar[J]. *Applied Optics*, 2009, 48(20): 3903-3914.
- [4] D'Amico G, Amodeo A, Mattis I, *et al.* EARLINET single calculus chain-technical-part 1: Pre-processing of raw lidar data[J]. *Atmospheric Measurement Techniques*, 2016, 9(2): 491-507.
- [5] 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.
- [6] Bo Guangyu, Xie Chenbo, Wang Bangxin, *et al.* Observational study of aerosol distribution over Jing-Jin-Ji area using airborne lidar[J]. *Acta Optica Sinica*, 2015, 35(9): 0901007.  
伯广宇, 谢晨波, 王邦新, 等. 机载激光雷达探测京津冀地区气溶胶的空间分布[J]. *光学学报*, 2015, 35(9): 0901007.
- [7] Fu Yibin, Wang Yu, Zhang Tianshu, *et al.* Signal acquisition system with simultaneous analog and photon counting measurement for lidar[J]. *Chinese J Lasers*, 2015, 42(8): 0814001.  
付毅宾, 王 煜, 张天舒, 等. 模拟与光子计数融合的激光雷达信号采集系统设计[J]. *中国激光*, 2015, 42(8): 0814001.
- [8] Huang Lifeng, Gong Wei, Li Jun, *et al.* Signal splicing of dual-receiver Mie scattering lidar in atmospheric remote sensing[J]. *Journal of Remote Sensing*, 2012, 16(4): 705-719.  
黄立峰, 龚 威, 李 俊, 等. 大气遥感双视场米散射激光雷达信号拼接[J]. *遥感学报*, 2012, 16(4): 705-719.
- [9] Liu Z S, Li Z G, Liu B Y, *et al.* Analysis of saturation signal correction of the troposphere lidar[J]. *Chinese Optics*



- Letters, 2009, 7(11): 1051-1054.
- [10] Zhang Y P, Yi F, Kong W, *et al.* Slope characterization in combining analog and photon count data from atmospheric lidar measurements[J]. Applied Optics, 2014, 53(31): 7312-7320.
- [11] Walker M, Venable D, Whiteman D N. Gluing for Raman lidar systems using the lamp mapping technique[J]. Applied Optics, 2014, 53(36): 8535-8543.
- [12] Veberič D. Maximum-likelihood reconstruction of photon returns from simultaneous analog and photon-counting lidar measurements[J]. Applied Optics, 2012, 51(2): 139-147.
- [13] Hu Qinghua, Yu Daren, Xie Zongxia. Numerical attribute reduction based on neighborhood granulation and rough approximation[J]. Journal of Software, 2008, 19(3): 640-649.  
胡清华, 于达仁, 谢宗霞. 基于邻域粒化和粗糙逼近的数值属性约简[J]. 软件学报, 2008, 19(3): 640-649.
- [14] Deb K, Pratap A, Agarwal S, *et al.* A fast and elitist multiobjective genetic algorithm: NSGA-II[J]. IEEE Transactions on Evolutionary Computation, 2002, 6(2): 182-197.
- [15] Razenkov I. Characterization of a Geiger-mode avalanche photodiode detector for high spectral resolution lidar[D]. Madison: University of Wisconsin-Madison, 2010.