

分布式光纤传感器应变读数异常的自适应后处理算法

梁智洪¹, 邓凯文¹, 马云龙², 王明华¹, 刘德博², 吴会强², 王奕首^{1*}¹厦门大学航空航天学院, 福建 厦门 361005;²北京宇航系统工程研究所, 北京 100076

摘要 针对分布式光纤传感器在航空航天领域应用中存在的应变读数异常现象,提出一种可检测和快速清除分布式光纤传感器应变读数异常的自适应后处理算法,以提高传感器监测精度与数据可靠性。该算法采用K均值聚类方法进行自适应定义阈值,以区分不同结构特征及不同服役环境导致的数据分布与噪声响应的差异性。在此基础上,对扭曲的应变曲线实行连续的几何偏置来消除应变读数异常。最后以航天燃料贮箱压力循环实验中采集的分布式光纤传感数据处理为例验证所提方法的有效性,使用Pearson相关系数来表征后处理曲线与无异常曲线的相关性,并与其他后处理算法进行对比。结果表明,针对2种主要类型、8种典型案例的应变异常现象,所提方法均能获得最佳的后处理结果,与无异常曲线的相关性系数不低于0.917。

关键词 光纤光学; 瑞利散射; 分布式光纤传感; K均值聚类; 自适应阈值; 复合材料贮箱

中图分类号 O439; V19

文献标志码 A

DOI: 10.3788/AOS231457

1 引言

光纤传感器由于体小质轻、耐腐蚀、抗辐射、可组网和多参量同步测量等优势,被广泛应用于航空航天飞行器状态感知与损伤监测^[1-3]。目前常用于航空航天领域结构健康监测的光纤传感器主要有两类:一是布拉格光纤光栅传感器(FBG),二是分布式光纤传感器(DOFS)。FBG因对温度及应力敏感,被广泛用于空间环境监测^[4]、结构形状重构及载荷反演等^[5-6],但FBG中可复用的光栅数量有限,本质上仍然属于点式传感器,受限于解调设备的通道数量,无法有效将结构的局部损伤与整体状态关联^[7]。DOFS能够提供毫米级的空间分辨率,实现结构的大面积、高密度感知与测量,被认为是飞行器最理想的实时在线监测技术^[8-9]。

面对航空航天中恶劣的服役环境,DOFS测点位置选择以及DOFS封装工艺等因素导致的测量不确定性,分布式光纤传感器测量数据会产生应变读数异常(SRAs)现象,产生原因包括:监测对象的非均匀性、传感器的局部损伤、传感器的退化以及解调算法的局部故障等^[10-15]。这些因素均会导致监测结果在局部区域或某一时刻出现较大的应变峰值,使得DOFS无法提供正确的物理量测量值,测量结果解释难度增加。为

减少SRAs带来的负面效应,研究者采取了一系列数据处理方法,如多项式拟合方法^[16]、Savitzky-Golay滤波法^[17]、移动平均法、FFT滤波和小波分析法等^[18-20]。尽管这些数据处理方法可有效地降低测量数据中的随机误差,但并不能完全去除SRAs现象。Luna公司提出利用测量光谱与基线光谱之间的相关性水平来衡量测量读数的可靠性,即谱移质量(SSQ)法^[21],并被众多学者用于SRAs现象的消除^[22-24]。但SSQ的计算值与异常值之间并不是完全相关的,仍然存在将高可靠测量读数删除的风险^[25]。因此,部分学者提出一些新的先进后处理算法,如基于Hampel标识符的过滤算法^[26]、几何阈值方法,以及基于多项式拟合的应变插值算法等^[25]。大部分方法的本质特征是将标准差大于平均值的数据标记为异常值,但如果异常值在一定时间范围内持续存在,异常值也会影响标准差与平均值。因此,这些方法不能满足检测SRAs的必要要求,因为它不能有效区分SRAs与裂纹突然发生、冲击或载荷增加引起的应变峰值。同时,上述方法仍采用固定阈值方法检测和确定异常值,固定阈值的确定依赖于人工经验,检测效率低、虚警率高,限制了其在复杂的服役环境中应用。

针对上述问题,本文提出一种检测和快速清除

收稿日期: 2023-08-21; 修回日期: 2023-10-01; 录用日期: 2023-10-21; 网络首发日期: 2023-12-12

基金项目: 国家重点研发计划(2018YFA070144)、国防基础科研计划重大项目(JCKY2019203A003)、基础加强计划重点项目(JCJQZD-203)、国家自然科学基金(U2141245)

通信作者: *wangys@xmu.edu.cn

DOFS的应变读数异常的智能自适应后处理方法。该方法通过对扭曲的应变曲线实行连续的几何阈值偏置来消除应变读数异常现象。该方法考虑到不同测点之间数据分布和噪声响应的差异,引入K均值聚类算法自适应确定阈值,提高了SRAs的检测效率。为了验证所提方法的有效性,在航天燃料贮箱上进行压力循环实验,然后将该方法的后处理结果与现有两种方法的处理结果进行比较。

2 SRAs分析与解决方法

文献[25]定义了两种类型SRAs:无害的应变读数异常(HL-SRAs)和有害的应变读数异常(HF-SRAs)。HL-SRAs是指在特定的位置或时刻存在应变突变,应变发生突变但可以快速恢复,突变在局部稀疏分布,不影响对正确响应趋势的判断。HF-SRAs是指在应变出现突变后,后续的应变测量值都遵循突变后的值,存在“传染性”效应,在一段时间内不可恢复或频繁突变,从而严重影响对正确响应趋势

的判断。HL-SRAs分布较为稀疏时,并不会影响对正确趋势的判断,随着SRAs现象的密集发展,HL-SRAs会逐渐演变成HF-SRAs。因此,在频繁突变的SRAs情况下,HL-SRAs与HF-SRAs没有明确的界限。以图1所示阶梯式加载实验中光纤测量应变说明SRAs现象。图1(a)展示了阶梯式加载实验中压力与时间的变化关系,而图1(b)~(d)展示了光纤各个测点由压力引起的应变响应情况。一般地,一根分布式光纤应变测量数据可以表示为 $n \times m$ 维矩阵 S ,其中 n 为时间维度, m 为空间维度(即测点数量)。图1(b)中不存在SRAs现象的正常应变响应曲线;图1(c)展示了HL-SRAs,即某一时刻会发生短暂突变,但不影响工程人员对于正确值的判断,可直接剔除应变异常值。图1(d)展示的是HF-SRAs现象,首先在一定时间范围内频繁发生突变,最后形成不可恢复的形式,而“频繁突变”至“不可恢复”对应时间范围内的应变响应已经不可信,会严重影响工程人员对于正常值的判断。

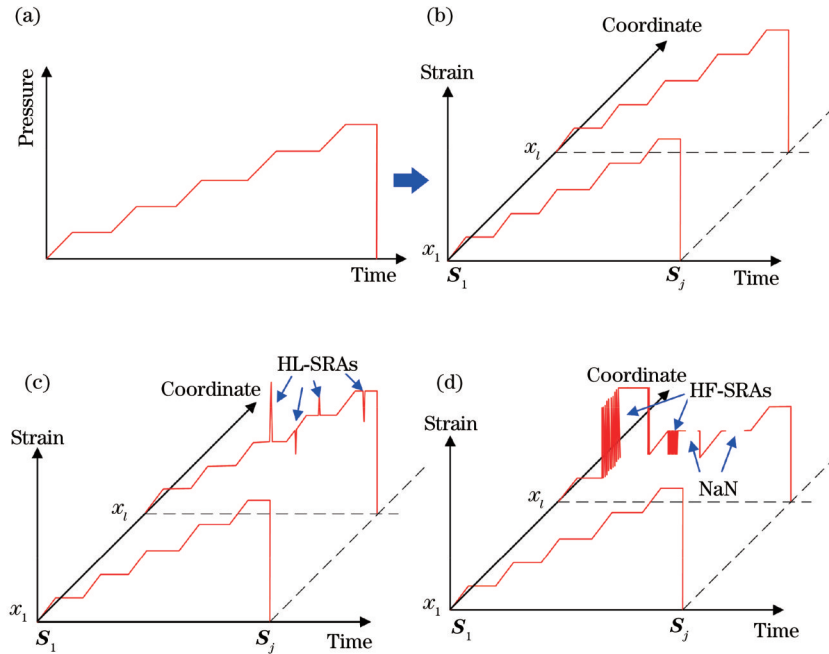


图 1 DOFS的SRAs的类型。(a)压力-时间图;(b)正常应变响应;(c)HL-SRAs;(d)HF-SRAs

Fig. 1 Types of SRAs for DOFS. (a) Pressure vs. time; (b) normal strain response; (c) HL-SRAs; (d) HF-SRAs

当实验过程中出现SRAs现象时,SRAs所对应的谱移质量系数(QF)会下降。QF由分布式解调仪制造商美国Luna公司提出,用于描述单个测点的谱移与物理量之间的互相关强度,它的范围是0~1,其中1表示完全相关,0表示不相关。同时,Luna公司还提出另一项评价指标——质量阈值(QT)。QT是QF的阈值,默认设置为0.38,Luna公司认为,QF高于0.38表示该读数的应变/温度测量结果具有良好的相关性和准确性,而所有低于QT的QF所对应的测量值都用非数值(NaN)代替^[21]。QT可根据实际场景进行调整,但

阈值过高可能存在将高可靠测量读数删除的风险^[25]。虽然通过QT能够去除一部分SRAs,但QT并不能过滤所有SRAs,当面临一些特殊的测量场景时,仍有大量测点存在SRAs。

为了解决QT无法完全过滤SRAs的问题,研究者提出几何阈值法(GTM)与多项式插值比较法(PICM)来检测和消除SRAs,以获得更准确、可靠的应变测量值^[25]。针对不同类型的SRAs现象,无论何种方法均不需要提前判定SRAs的类型,但不同的算法对于不同类型SRAs的适用性不同。GTM是通过判断相邻

时刻下的应变增量是否超过阈值来检测应变异常值,当应变增量超出所定义的阈值时,采用 NaN 替代异常值。PICM 通过对比多项式曲线与原始应变曲线之间的差异来判断应变异常值。当二者的差异值大于所定义的阈值,则判定为异常值,异常值采用对应时刻的多项式曲线值进行替代。对于同一序列的原始应变曲线, PICM 可选取任意段进行拟合,分段越多,效率越低。因此,拟合段数、拟合阶数以及阈值的选择均是后处理结果的主要影响因素。

从上述两种方法看, SRAs 处理目标是检测和剔除原始应变测量值的异常值并获得一组更可靠的新数据,并且尽可能不受任何 SRAs 引起的畸变的影响^[25]。核心处理流程包括:步骤 1,自动检测应变异常值;步骤 2,移除应变异常值;步骤 3,填充应变异常值。第一步中,如何定义阈值是判断应变异常值的关键。然而, GTM 和 PICM 方法均依赖于人工经验,检测效率低、虚警率高。 GTM 和 PICM 方法的本质特征是将标准差大于平均值的数据标记为异常值,但如果异常值在一定时间范围内持续存在,异常值也会影响标准差与平均值。因此,两种方法仅能用于符合上述特征的 SRAs 现象,即 HL-SRAs。同时,这些方法难以区分 SRAs 与裂纹突然发生、冲击或载荷增加引起的应变峰值。本文针对上述方法的不足,结合基于瑞利散射的分布式光纤传感特点,提出一种基于 K 均值聚类自适应算法的几何阈值偏移方法 (AGTOM) 来处理 SRAs。

3 基于 AGTOM 的 SRAs 自适应后处理算法

如前所述,利用固定阈值判别方法无法对应应变异常值进行自适应判别,对不同的受载状态、复杂服役环境的适应性较差,自动化效率较低。所提 AGTOM 的基本思想是通过将对数组异常时刻到最终时刻的应变值进行整体连续的偏移来消除异常值的影响,因此,无论是 HL-SRAs 还是 HF-SRAs 均不需要进行差异性处理。这种方法的优点是:若 SRAs 存在连续振荡、突变, AGTOM 可通过整体连续的偏置手段对数据进行修正;若 SRAs 在突变后不可恢复, AGTOM 也只是针对突变时刻进行偏置。此外,利用 K 均值聚类将应变增量数据集划分为不相交的子集,具有簇内相似度高、簇间相似度低的特点,从而实现对应应变异常值进行自适应判断。图 2 所示为 AGTOM 实现流程,具体步骤如下:

步骤 1:输入原始应变矩阵 S 并计算相邻时刻下的应变增量绝对值矩阵 S_l ,应变增量矩阵 S_l 是一个 $(n-1) \times m$ 阶矩阵。通常来说,任一测点的应变曲线通常由 n 个应变值构成,而相邻时刻的应变增量绝对值有 $n-1$ 个。

步骤 2:将 l 定义为光纤上的某一测点, $l \in [1, m], l$

的起始值为 1。判断 l 是否在 $[1, m]$ 范围内。若满足,则进入下一步;否则,结束。

步骤 3:通过 K 均值聚类算法进行阈值自适应判定。

步骤 3-1:设置样本类簇划分的数量 k , k 的范围设置为 $2 \leq k \leq 20$,从应变增量的数据集 S_l 中随机选取任意个样本点划分为 k 个类簇,计算每个类簇的簇心,得到 $\mu^{(l)} = [\mu_1^{(l)}, \mu_2^{(l)}, \dots, \mu_k^{(l)}]$ 。

步骤 3-2:分别计算每个样本 $S_{i,jl}$ 到各个簇心 $\mu_a^{(l)}$ 的距离 $D_{ja}^{(l)}$,其计算公式为

$$D_{ja}^{(l)} = \|S_{i,jl} - \mu_a^{(l)}\|_2. \quad (1)$$

然后,根据就近原则将 $S_{i,jl}$ 划分到最近的类别 $\lambda_a^{(l)}$,公式为

$$\lambda_a^{(l)} = \operatorname{argmin}_{a \in [1, k]} D_{ja}^{(l)}. \quad (2)$$

步骤 3-3:重新计算每个类别 $C_a^{(l)}$ 的簇心 $\mu_a^{(l)}$,公式为

$$\mu_a^{(l)} = \frac{1}{|C_a^{(l)}|} \sum_{S_i \in C_a^{(l)}} S_i. \quad (3)$$

重复步骤 3-2 和 3-3,直到簇心不再变化即可终止。

步骤 3-4:由于测点位置的应变增量规律并不确定,因此采用聚类算法的内部评价指标 Davies-Bouldin 指数 (DBI) 来估计聚类数量。DBI 是一种用于 K 均值聚类的性能指标,它考虑了簇内的紧密度和簇间的分离度,可以帮助评估不同 k 值下的聚类性能,并选择最优的 k 值。首先,分别计算各个簇与其他簇之间的平均距离与簇内的点到其中心的平均距离的比值;然后,选择最大值,对所有簇求平均,以此实现对聚类结果的评估。当簇内紧密度高、簇间分离明显时, DBI 值较小;当簇内紧密度低、簇间分离不明显时, DBI 值较大。DBI 的取值范围为 $[0, +\infty]$, DBI 越小,表示聚类效果越好,因此选择最小 DBI 值对应的 k 值为最优 k 值。通过使用 DBI 值来确定最优 k 值,能够在不依赖先验知识的情况下,基于数据的聚类性能指标来选择最优 k 值。DBI 的计算公式为

$$R_{\text{DBI}} = \frac{1}{k} \sum_{a=1}^k \max_{a \neq b, 1 \leq b \leq k} \left[\frac{\operatorname{avg}(C_a^{(l)}) + \operatorname{avg}(C_b^{(l)})}{d_c(u_a, u_b)} \right], \quad (4)$$

式中: $d_c(u_a, u_b)$ 为簇 a, b 的质心 u_a, u_b 的欧氏距离; $\operatorname{avg}(C_a^{(l)})$ 与 $\operatorname{avg}(C_b^{(l)})$ 分别为簇 a, b 的平均距离,表示各个簇内每一对数据点之间的平均距离,用于衡量簇内数据之间的紧密性。

$$\operatorname{avg}(C_a^{(l)}) = \frac{1}{n_a(n_a - 1)} \sum_{p, q \in C_a^{(l)}} d_e(p, q), \quad (5)$$

式中: n_a 为簇 a 中的数据点数量; p 和 q 为簇 a 中的数据点; $d_e(p, q)$ 为 p 和 q 之间的距离,一般采用欧氏距离度量方法。

步骤 3-5: 聚类的目的是将不同变化规律、不同响应情况的应变增量进行分隔, 从而区分正常的应变增量以及异常的应变增量。聚类结果的第一类别通常为正常的应变增量, 因此阈值的设置必须大于第一类别中的应变增量边界值 $\max_{S_{i,p} \in C_1^{(l)}} S_{i,p}$ 。考虑到最佳阈值介于应变增量的第一类别边界值与第二类别边界值之间, 因此, 最佳阈值可为 $S_T = \frac{1}{2} \left(\max_{S_{i,p} \in C_1^{(l)}} S_{i,p} + \min_{S_{i,q} \in C_2^{(l)}} S_{i,q} \right)$ 。

步骤 4: 用最近的非缺失值替代原始应变数组 S 中

的 NaN。

步骤 5: 判断 j 是否属于 $[2, n]$, l 的起始值为 2。若属于, 则进入下一步; 否则, 回到步骤 2。

步骤 6: 判断应变增量值 $S_{i,jl}$ 是否大于阈值 S_T 。若 $S_{i,jl} > S_T$, 则判断 $S_{jl} > S_{il}$; 否则, 令 $i = i + 1, j = j + 1$, 回到步骤 5。

步骤 7: 判断 S_{jl} 是否大于 S_{il} 。若 $S_{jl} > S_{il}$, 则 $S' = [S_{1l}, \dots, S_{il}, S_{il} - S_{i,jl}, \dots, S_{nl} - S_{i,jl}]$; 否则, $S' = [S_{1l}, \dots, S_{il}, S_{il} + S_{i,jl}, \dots, S_{nl} + S_{i,jl}]$ 。

步骤 8: 令 $S = S', i = i + 1, j = j + 1$, 回到步骤 5。

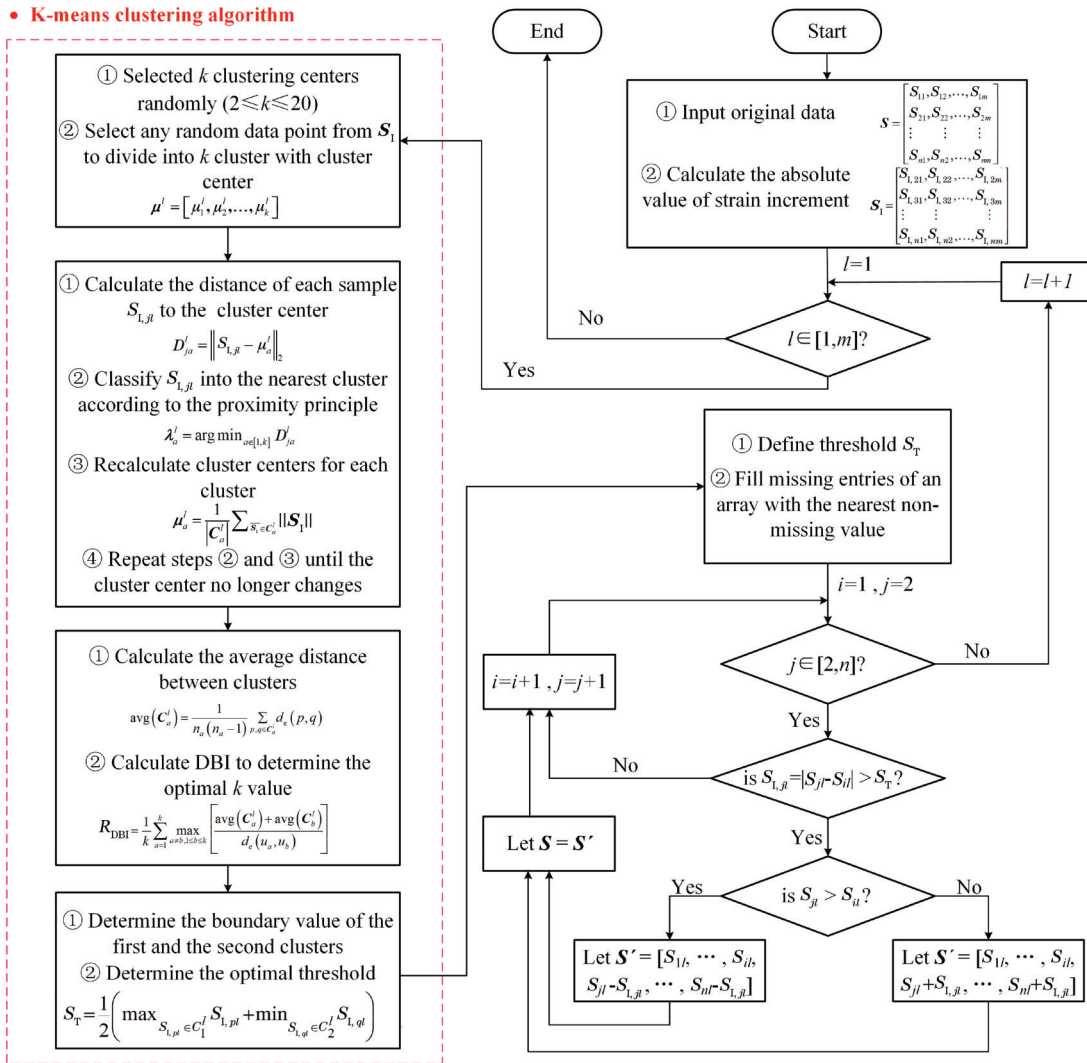


图 2 基于 AGTOM 的 SRAs 处理方法
Fig. 2 SRAs processing method with AGTOM

4 案例研究

4.1 实验对象

实验对象为加筋复合材料贮箱, 由北京宇航系统工程研究所设计制造。贮箱由 T800 碳纤维缠绕制备而成, 内部利用复合材料筋条形成周期性栅格区域, 贮箱中间存在一条宽度为 60 mm 的金属环作为内部加

强筋栅格的连接加强环。加筋复合材料贮箱总高度为 874 mm, 中部筒段环形区域的总高度为 449 mm, 外径尺寸为 800 mm, 内径为 780 mm, 其缠绕层厚度为 10 mm, 贮箱上下顶面均有一圈内径为 250 mm 的金属法兰, 如图 3 所示。

4.2 实验方案设计

通过开展贮箱的压力循环实验, 以实现贮箱在设

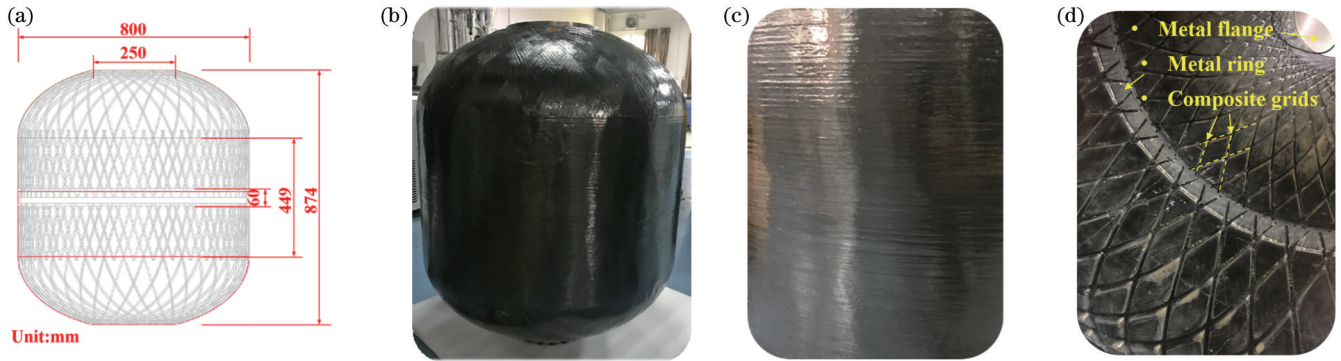


图 3 加筋复合材料贮箱。(a)几何尺寸;(b)结构外形;(c)结构表面;(d)结构内部

Fig. 3 Reinforced composite tank. (a) Geometric size; (b) structural shape; (c) structural surface; (d) structural interior

计考核阶段的性能评估与实验验证。设计常温条件下的压力循环实验,以模拟加筋复合材料贮箱在服役中的燃料充放过程。在压力循环实验过程中,首先将贮箱注满水,然后密封贮箱,最后通过注入氮气来加压。本次压力循环实验共计测试 10 个循环,每一个循环包含增压、保压、泄压 3 个阶段。增压阶段的压强范围为 0.1~1.0 MPa,保压阶段则维持 1.0 MPa,持续约 10 min,泄压阶段的压强范围为 1.0~0.4 MPa,实验流程如图 4 所示。另外,在压力循环前后,贮箱进行氦质谱检测,以对比压力循环前后的渗漏情况,氦质谱实验结果在本文不再详述。

4.3 监测方案设计

采用总长度为 15 m 的分布式光纤传感器对贮箱进行传感布设,以实现加筋复合材料贮箱的高密度应变监测。本次实验采用美国 Luna 公司的分布式光纤解调仪 ODISI-B 进行监测,该设备采用基于后向瑞利散射技术的光频域反射技术。ODISI-B 解调仪在不同的测量模式下有不同的采集速率、最大监测长度以及测点间距^[21]。在实验采样频率为 10 Hz,长度为 15 m 的情况下,传感器相邻测点的间距为 2.61 mm,测点总数为 5852 个。将解调仪的 QT 值设置为默认的 0.38,低于 0.38 的 QF 则采用 NaN 替代。分布式光纤采用表贴式粘接至贮箱表面,其宽度控制为 10 mm,厚度为

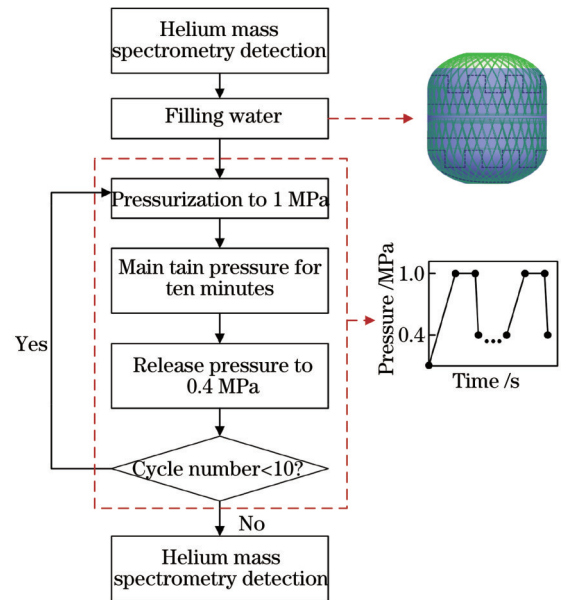


图 4 测试流程图

Fig. 4 Testing flowchart

1 mm 左右,胶黏剂采用上海华谊树脂有限公司提供的高性能低温结构胶 DW-1。图 5 展示了分布式光纤传感网络设计图,监测部位包括贮箱的前穹顶、筒段、后穹顶及底部法兰附近,覆盖了横向和纵向应变测量。

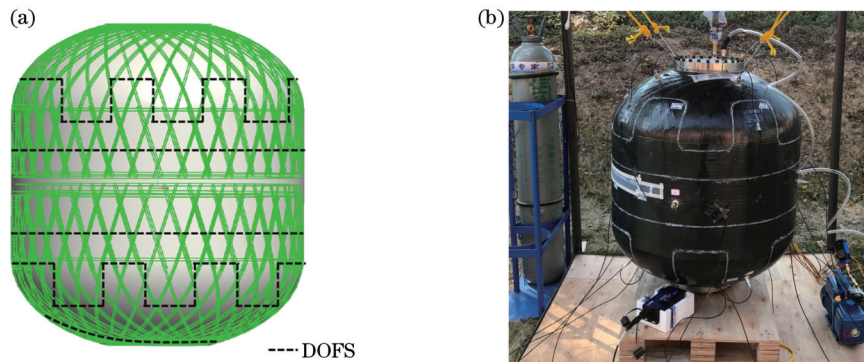


图 5 分布式光纤传感网络设计。(a) 传感网络设计图;(b) 实际布设图

Fig. 5 DOFS network design. (a) Design drawing of sensor network; (b) actual layout diagram

5 结果与讨论

5.1 压力循环实验结果

图 6 展示了复合材料贮箱压力循环实验的原始应变结果。从空间维度来看,原始应变结果存在大量的 SRAs 现象,含 SRAs 现象的应变测点占比较高;从时间维度来看,SRAs 类型复杂,HL-SRAs 与 HF-SRAs 相互耦合,原始应变结果无法真实地反映实验过程中贮箱的结构状态。从 SRAs 的响应情况来看,应变结果范围约为 $\pm 15000 \mu\epsilon$,超出了合理范围。SRAs 产生的原因推测是复合材料贮箱表面纤维导致的表面不均匀[即使打磨充足也不能消除表面的不均匀性,如图 3 (c) 所示],以及胶粘工艺的不均匀。为了进一步说明原始应变测量结果中是否有 SRAs 现象的差异性,抽取原始应变结果中不含 SRAs 现象的应变结果,如图 7 所示。图 7(a) 展示了 DOFS 中 0~400 mm 长度位置的数据结果,可以看到,原始应变结果(不含 SRAs)在空间维度上的呈现较为连续。由于不同测点所对应的结构特征不一致,因此不同测点之间的应变峰值并不

相同,其应变峰最大值不超过 $1500 \mu\epsilon$;任意选取 DOFS 位于 100 mm 长度位置的测点数据,其应变响应情况能较好地反映贮箱在实验中的压力循环情况,如图 7(b) 所示。类似地,在空间维度上,不同测点的应变响应情况类似,均能反映贮箱在实验过程中的结构状态。

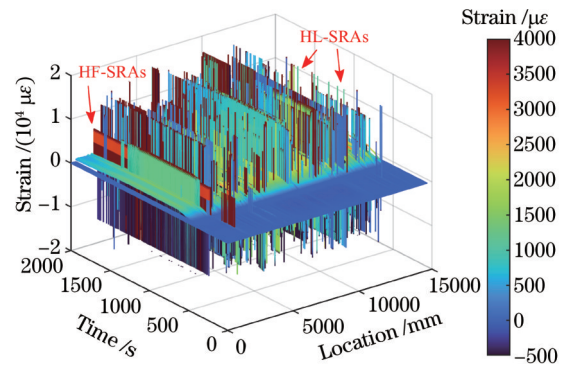


图 6 原始应变结果(含 SRAs)

Fig. 6 Original strain results (with SRAs)

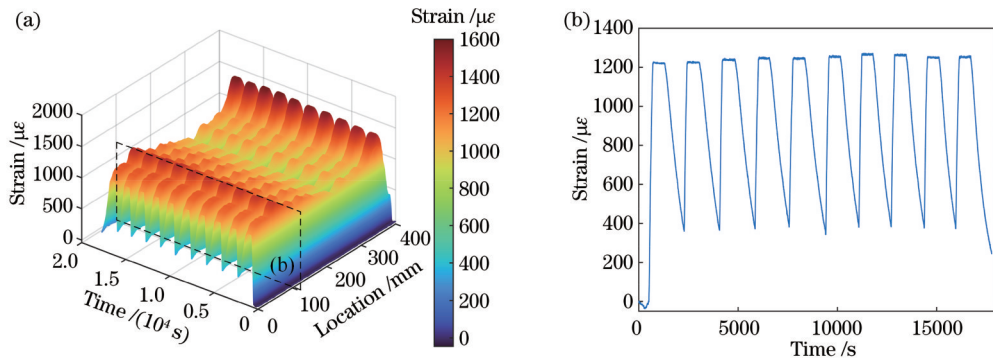


图 7 原始应变结果(不含 SRAs)。(a)长度位置 0~400 mm 的应变结果;(b)长度位置为 100 mm 的应变结果

Fig. 7 Original strain results (without SRAs). (a) Strain results in length position 0-400 mm; (b) strain results at length position of 100 mm

从含 SRAs 的原始应变测量结果中分别选取以 HL-SRAs 为主与以 HF-SRAs 为主的应变曲线,每种 SRAs 类型选取 4 个不同 SRAs 分布规律的典型案例,如图 8 所示。由图 8 可知,并非每种 SRAs 案例中只包含单个类型的 SRAs,而是特定案例中特定 SRAs 会占据主导地位,因为 SRAs 的产生和分布规律都是随机的,图 8 中用箭头标记 HF-SRAs 出现的位置。图 8(a) 所示的案例 I HL-SRAs 中,SRAs 在应变曲线中的分布较为稀疏。案例 II HL-SRAs 中,SRAs 在某一段时间区间内出现较为频繁,见图 8(b) 的 500~2000 s、2500~3500 s 以及 4000~5000 s 的保压阶段,其分布较为密集,特别是 2500~3500 s 段,为 HF-SRAs。案例 III HL-SRAs 中,SRAs 现象的出现会伴随一些 NaN 值,此外,还存在一处 HF-SRAs,见图 8(c) 的 10000 s 附近。案例 IV HL-SRAs 中,密集的 HL-SRAs 与稀疏的 HL-SRAs 并存,同时还存在 NaN 值,密集的 SRAs

发展为 HF-SRAs,见图 8(d) 的 2000~4000 s,其余均为 HL-SRAs。

不难发现,当出现密集的 HL-SRAs 时,其特征极易演化为 HF-SRAs,二者的主要区别是 HF-SRAs 的连续化程度更高,SRAs 的振荡特性更明显,见图 8(e) 所示的案例 I HF-SRAs,基本为 HF-SRAs。当 HF-SRAs 的出现同时伴随大量的 NaN 值,可认为是案例 II HF-SRAs,见图 8(f) 的 6000 s、8000 s 以及 10000 s 附近。此外,在案例 III HF-SRAs 中,HF-SRAs 与稀疏分布的 HL-SRAs 出现耦合,见图 8(g) 的 6000 s 以及 7000 s 附近。最后,在图 8(h) 展示的案例 IV HF-SRAs 中,存在两处 HF-SRAs(如箭头所示),同时,HF-SRAs 与稀疏分布的 HL-SRAs 出现耦合,还存在大量的 NaN 值。

5.2 最佳阈值的确定

传统的后处理算法在处理不同典型案例的原始应

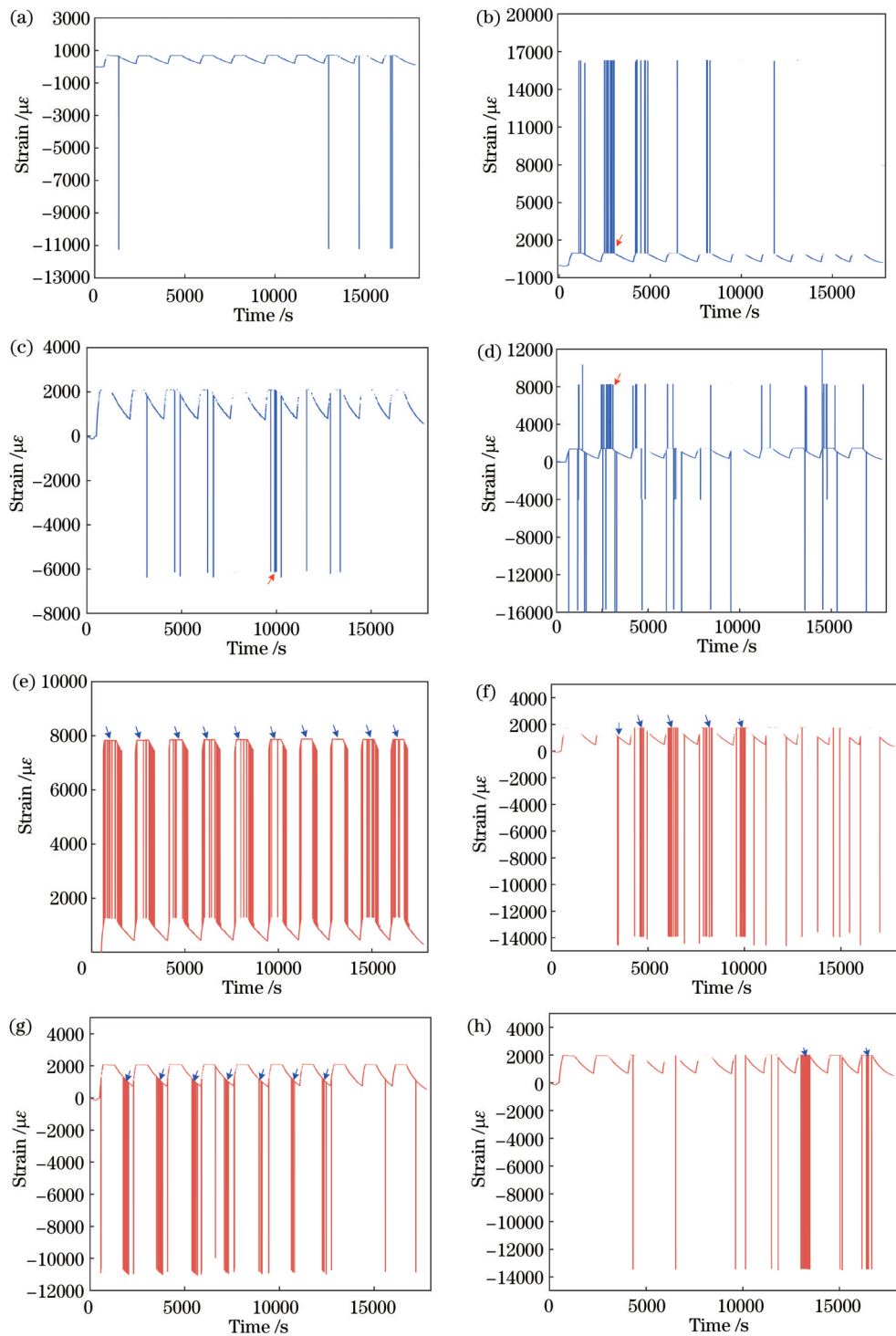


图 8 原始应变结果(含 SRAs)。(a)案例 I HL-SRAs;(b)案例 II HL-SRAs;(c)案例 III HL-SRAs;(d)案例 IV HL-SRAs;(e)案例 I HF-SRAs;(f)案例 II HF-SRAs;(g)案例 III HF-SRAs;(h)案例 IV HF-SRAs

Fig. 8 Original strain results (with SRAs). (a) Case I HL-SRAs; (b) case II HL-SRAs; (c) case III HL-SRAs; (d) case IV HL-SRAs; (e) case I HF-SRAs; (f) case II HF-SRAs; (g) case III HF-SRAs; (h) case IV HF-SRAs

变曲线之前,需要定义用于检测和判定 SRAs 的阈值。DOFS 由于传播距离长、监测范围广,常常被用于结构的大面积监测,因此不同测点面对的实验环境和所处的结构特征并不完全相同。采用固定阈值的传统人工选择方法不适于 DOFS 的不同测点之间的多特征数据分布。根据 AGTOM 算法的步骤 3 确定每个测点的参

数值(k 和 DBI),从而确定最佳阈值(S_T)。表 1 展示了 SRAs 不同典型案例所对应的 DBI 值、 k 值以及 S_T 值。

5.3 不同方法的处理结果对比

为了说明不同后处理方法在实际工程中检测和消除 SRAs 的有效性,分别采用 AGTOM、GTM 以及 PICM 对压力循环实验中不同测点的原始应变进行处

表 1 不同典型案例的自适应阈值

Table 1 Adaptive thresholds for different typical cases

Type	DBI	k	S_T
Case I HL-SRAs	0.0014	2	115.72
Case II HL-SRAs	0.0015	3	126.85
Case III HL-SRAs	0.0116	2	277.02
Case IV HL-SRAs	0.0296	10	56.72
Case I HF-SRAs	0.0027	2	263.85
Case II HF-SRAs	0.0081	2	120.65
Case III HF-SRAs	0.0026	2	478.16
Case IV HF-SRAs	0.0026	2	126.68

理。AGTOM 使用自适应阈值,而其他两种方法使用固定阈值方法(均设为 100)。此外,为了量化每种后处理方法检测和消除 SRAs 的性能,使用 Pearson 相关系数(PCC; R_{PCC})来计算后处理曲线与无异常值应变曲线之间的相关性。PCC 是两个随机变量(实值向量)之间的线性相关性的指标。无异常值应变曲线选取实验过程中无 SRAs 现象的测点数据如图 7(b)所示。PCC 的计算公式为

$$R_{PCC, ab} = \frac{\sum_{i=1}^n (a_i - \bar{a})(b_i - \bar{b})}{\sqrt{\sum_{i=1}^n (a_i - \bar{a})^2} \sqrt{\sum_{i=1}^n (b_i - \bar{b})^2}}, \quad (6)$$

式中: \bar{a} 是变量 a 的平均值; \bar{b} 为变量 b 的平均值。PCC 提供了两个随机变量 a 和 b 之间关系的线性强度指标,范围为 $[-1, 1]$ 。PCC 值接近 1 表示强正相关,0 表示无相关性, -1 表示强负相关。还应注意的是,即使是在同一次实验过程中,不同测点所测量的应变值也并不一致,这意味着后处理曲线与其他无异常值应变曲线之间具有不一致的响应规律。因此,在计算 PCC 之前,需要对原始应变数据进行归一化。例如,将变量 a 进行归一化处理,即

$$a_{nor} = \frac{\sum_{i=1}^n a_i - \max(a)}{\max(a) - \min(a)}, \quad (7)$$

式中: $\max(a)$ 是变量 a 的最大值, $\min(a)$ 是其最小值。

图 9 展示了 3 种后处理算法处理 HL-SRAs 不同典型案例的结果。表 2 也呈现了不同后处理曲线与无 SRAs 应变曲线的 PCC 结果,以表征后处理效果。由表 2 可知,存在 SRAs 的原始应变曲线,除了案例 I HF-SRAs 外,其余案例的 PCC 均不超过 0.441。针对存在稀疏分布 HL-SRAs 的应变曲线,GTOM 与 AGTOM 均能获得较好的处理结果,其 PCC 值均大于 0.994。PICM 在 2000 s 附近能获得较好的处理效果,但存在少量 NaN 值时,PICM 则受到较大干扰,如图 9(a)所示。同样地,即使针对存在密集分布 HL-SRAs 的应变曲线,GTOM 与 AGTOM 仍取得较好结果,其 PCC 值均不低于 0.982。相反地,PICM 的 PCC 仅为 0.512,如图 9(b)所示。从图 9(c)的结果发现,当 HL-SRAs 现象的出现伴随一些 NaN 值时,PICM 的效果

最差,GTOM 开始受到干扰,只有 AGTOM 依然获得较好结果,三者的 PCC 值分别为 0.218、0.798 以及 0.971。此外,当稀疏分布、密集分布以及 NaN 值三种特征同时存在时,同样只有 AGTOM 能获得较好结果,其 PCC 值最佳,为 0.965,其余两种算法均受到不同程度的干扰。

图 10 展示了不同后处理算法对 HF-SRAs 不同典型案例的处理结果。由图 10(a)可知,PICM 的曲线与含 SRAs 的原始应变曲线类似,而 GTOM 在 3000 s 以及 17000 s 附近出现明显失真,说明 PICM 与 GTOM 在处理 HF-SRAs 时易出现误判。仅 AGTOM 获得较好的处理结果,其 PCC 值为 0.933,高于其余二者的 0.884 (PICM) 以及 0.855 (GTOM)。需要注意的是,HF-SRAs 的第一类典型案例的原始应变曲线与 PICM 曲线的 PCC 值均大于 0.884,但实际上 PICM 和 GTOM 与 AGTOM 的应变响应值并不相同。这是因为本文计算 PCC 值时采用了归一化方法,曲线经过归一化后大部分分布在 0~1 的范围,其包络轮廓可能会重叠,因此会出现 PCC 值偏高的假象。此外,当 HF-SRAs 出现的同时伴随大量的 NaN 时,如图 10(b)所示,3 种方法中 AGTOM 的 PCC 值最高,其次为 GTOM,而 PICM 的结果呈现负相关。同样地,当 HF-SRAs 与稀疏分布的 HL-SRAs 出现耦合时,AGTOM 仍保持较高的 PCC 值,为 0.917,其次为 GTOM,为 0.832, PICM 的结果略有改善,上升至 0.776,如图 10(c)所示。最后,当 HF-SRAs 与稀疏分布的 HL-SRAs 出现耦合,同时还存在大量的 NaN 值时,同样仅有 AGTOM 保持较好的处理结果,其 PCC 值仍高于 0.95,其余两种方法仍存在曲线扭曲现象,其 PCC 值不大于 0.757。

综上所述,当针对 HL-SRAs 现象时:AGTOM 均能得到最佳的后处理结果,其 PCC 值均不低于 0.965;其次为 GTOM,其 PCC 值均不低于 0.798,GTOM 仅在 HL-SRAs 与 NaN 值的耦合时会发生干扰[见图 9(c)中的 12000~18000 s 与图 9(d)中的 5000~10000 s];PICM 仅在第一类典型案例(即稀疏分布的 HL-SRAs)中能获得较好的处理结果,而在其余三类典型案例中,PICM 仍然发生失真扭曲,其 PCC 值不大于 0.512。值得注意的是,无论是 GTOM 还是 PICM,当 HL-SRAs 与 NaN 值的耦合时,其后处理曲线会失真扭曲。

当针对 HF-SRAs 现象时:AGTOM 均能得到最佳的后处理结果,其 PCC 值均不低于 0.917;PICM 易发生误判,这恰好反映了 HF-SRAs 与 HL-SRAs 的不同,由于前者的主要特征是在应变值发生突变后,后续的应变值会遵循错误的应变响应或频繁突变,而 PICM 是基于拟合值与当前值的对比进行检测与消除 SRAs,难以判断前后应变增量的变化;与 PICM 相比,GTOM 考虑了应变增量的突变,因此,当 HF-SRAs 由密集突变的 SRAs 构成时,GTOM 的后处理结果有所改

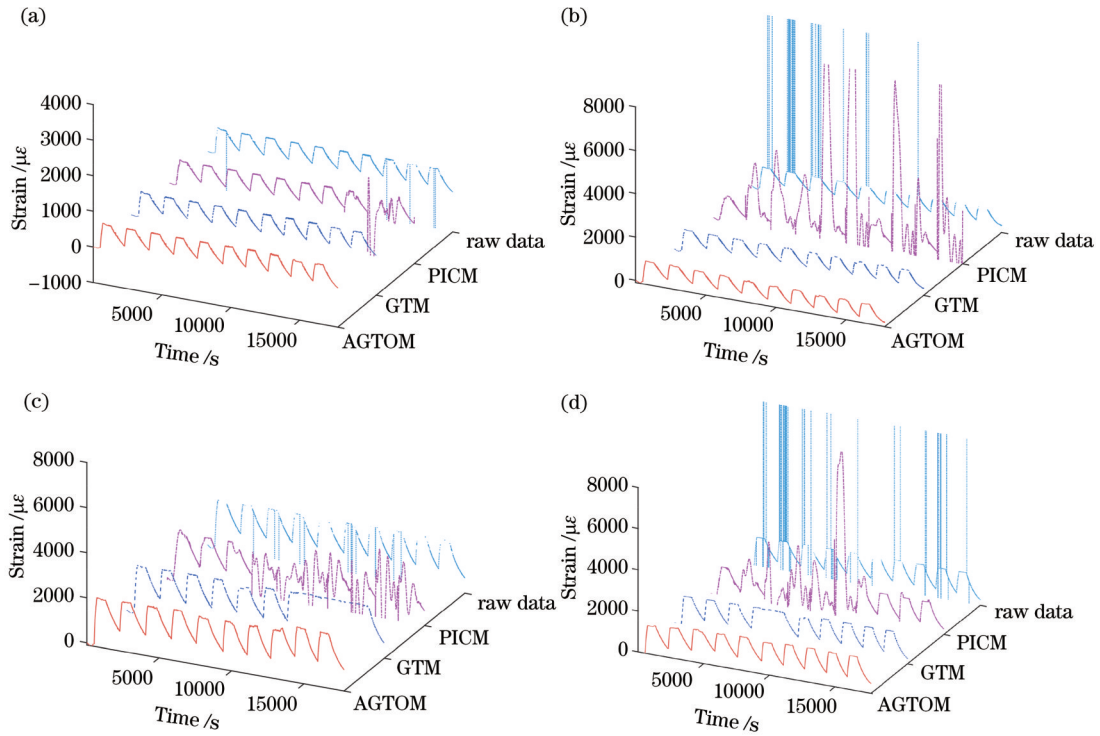


图 9 不同方法的后处理结果对比(HL-SRAs)。(a)案例 I HL-SRAs; (b)案例 II HL-SRAs; (c)案例 III HL-SRAs; (d)案例 IV HL-SRAs

Fig. 9 Comparison of post-processing results for different methods (HL-SRAs). (a) Case I HL-SRAs; (b) case II HL-SRAs; (c) case III HL-SRAs; (d) case IV HL-SRAs

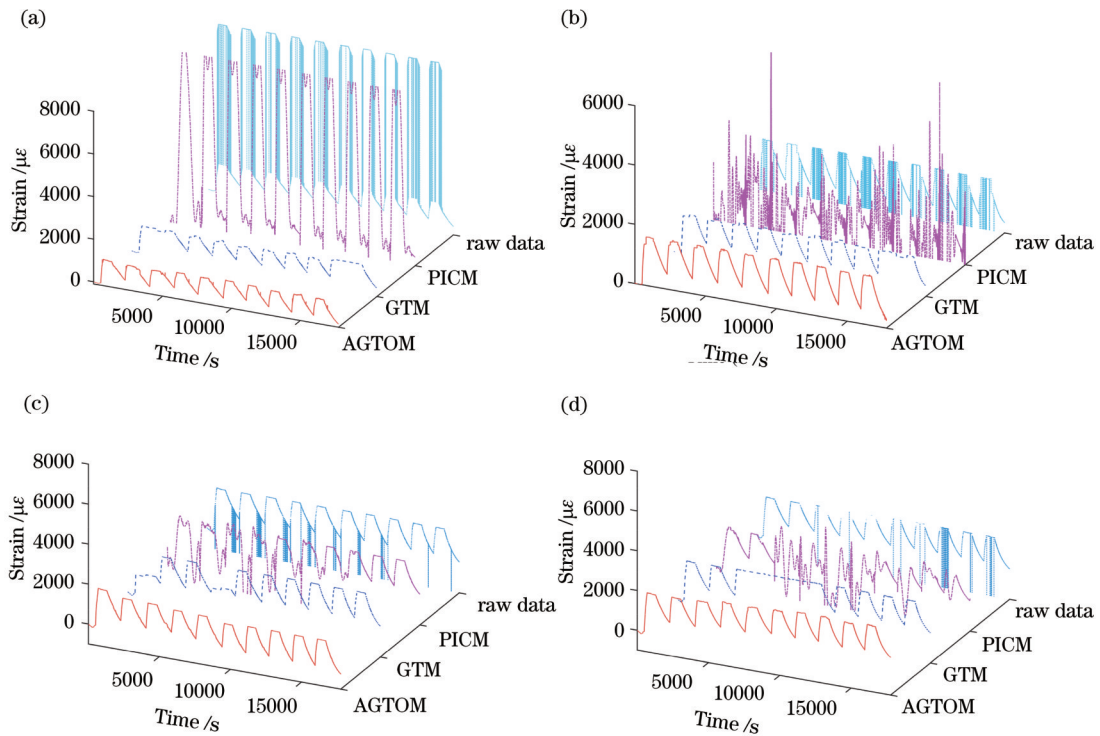


图 10 不同方法的后处理结果对比(HF-SRAs)。(a)案例 I HF-SRAs; (b)案例 II HF-SRAs; (c)案例 III HF-SRAs; (d)案例 IV HF-SRAs

Fig. 10 Comparison of post-processing results for different methods (HF-SRAs). (a) Case I HF-SRAs; (b) case II HF-SRAs; (c) case III HF-SRAs; (d) case IV HF-SRAs

表 2 后处理曲线与无 SRAs 应变曲线的 PCC 结果
Table 2 PCC results for post-processed curves and strain curves without SRAs

Type	Raw data	PICM	GTM	AGTOM
Case I HL-SRAs	0.413	0.383	0.996	0.994
Case II HL-SRAs	0.307	0.512	0.992	0.984
Case III HL-SRAs	0.437	0.218	0.798	0.971
Case IV HL-SRAs	0.340	0.454	0.920	0.965
Case I HF-SRAs	0.837	0.884	0.855	0.933
Case II HF-SRAs	-0.008	-0.224	0.883	0.929
Case III HF-SRAs	0.441	0.776	0.832	0.917
Case IV HF-SRAs	0.248	0.107	0.757	0.954

善。与 HL-SRAs 类似,无论是 GTM 还是 PICM,当 HF-SRAs 存在 NaN 值干扰时,后处理效果会下降[见图 10(b)中的 15000~18000 s 与图 10(d)中的 5000~10000 s],反映了 GTM 与 PICM 的算法鲁棒性低于 AGTOM。

6 结 论

提出一种检测和快速清除 DOFS 中应变读数异常的自适应后处理算法,并应用于航天器燃料贮箱压力循环实验的分布式光纤应变数据的后处理分析。通过对比分析现有算法(GTM 与 PICM)与本文算法(AGTOM)的后处理结果,得到如下结论:

1) K 均值聚类算法采用内部评价指标 DBI 来表征应变增量的聚类效果,从而得到最优的 k 值,以此获得自适应阈值,因此该算法能够适应不同结构特征及服役环境下的不同测点之间数据分布与噪声响应。

2) 针对 HL-SRAs 现象,GTM 与 AGTOM 方法均能得到不错的处理效果,而 PICM 面对密集分布、耦合分布的 HL-SRAs 易受到干扰,其后处理曲线存在严重失真扭曲。

3) 针对 HF-SRA 现象,除了 AGTOM,其他两种算法的后处理曲线均存在不同程度的失真扭曲,AGTOM 与无异常应变曲线相比,具有最佳的相似度。

4) 无论是 HL-SRAs 还是 HF-SRAs,当 SRAs 与 NaN 耦合时,GTM 与 PICM 均会受到干扰,说明二者的算法鲁棒性低于 AGTOM。

为了进一步验证 AGTOM 的有效性,未来仍需将该方法应用到不同的实验场景中。

参 考 文 献

- [1] Mizutani T, Takeda N, Takeya H. On-board strain measurement of a cryogenic composite tank mounted on a reusable rocket using FBG sensors[J]. *Structural Health Monitoring*, 2006, 5(3): 205-214.
- [2] Wang Y S, He M Y, Sun L, et al. Weighted adaptive Kalman filtering-based diverse information fusion for hole edge crack monitoring[J]. *Mechanical Systems and Signal Processing*,

2022, 167: 108534.

- [3] He M Y, Ramakrishnan K R, Wang Y S, et al. A combined global-local approach for delamination assessment of composites using vibrational frequencies and FBGs[J]. *Mechanical Systems and Signal Processing*, 2022, 167: 108577.
- [4] Hegde G, Prasad M V N, Asokan S. Temperature compensated diaphragm based Fiber Bragg Grating (FBG) sensor for high pressure measurement for space applications[J]. *Microelectronic Engineering*, 2021, 248: 111615.
- [5] 尹国路, 徐州, 蒋锐, 等. 基于光频域反射仪的光纤分布式三维形状传感技术[J]. *光学学报*, 2022, 42(1): 0106002.
Yin G L, Xu Z, Jiang R, et al. Optical fiber distributed three-dimensional shape sensing technology based on optical frequency-domain reflectometer[J]. *Acta Optica Sinica*, 2022, 42(1): 0106002.
- [6] 付彩玲, 彭振威, 李鹏飞, 等. OFDR 分布式光纤温度/应变/形状传感研究进展[J]. *激光与光电子学进展*, 2023, 60(11): 1106007.
Fu C L, Peng Z W, Li P F, et al. Research on distributed fiber temperature/strain/shape sensing based on OFDR[J]. *Laser & Optoelectronics Progress*, 2023, 60(11): 1106007.
- [7] Bado M F, Casas J R, Barrias A. Performance of rayleigh-based distributed optical fiber sensors bonded to reinforcing bars in bending[J]. *Sensors*, 2018, 18(9): 3125.
- [8] 尹国路, 徐州, 朱涛. 光纤形状传感器封装过程中残余应力的分布式实时监测[J]. *光学学报*, 2022, 42(16): 1606002.
Yin G L, Xu Z, Zhu T. Distributed real-time monitoring of residual stress during packaging process of optical fiber shape sensor[J]. *Acta Optica Sinica*, 2022, 42(16): 1606002.
- [9] 桂鑫, 李政颖, 王洪海, 等. 基于大规模光栅阵列光纤的分布式传感技术及应用综述[J]. *应用科学学报*, 2021, 39(5): 747-776.
Gui X, Li Z Y, Wang H H, et al. Review of distributed optical fiber sensing technology and application based on large-scale grating array fiber[J]. *Journal of Applied Sciences*, 2021, 39(5): 747-776.
- [10] Okabe Y, Tsuji R, Takeda N. Measurement of non-axisymmetric thermal residual strain in CFRP laminates using FBG sensors[J]. *Journal of the Japan Society for Composite Materials*, 2004, 30(5): 199-206.
- [11] Shiratsuchi T, Imai T. Development of fiber Bragg grating strain sensor with temperature compensation for measurement of cryogenic structures[J]. *Cryogenics*, 2021, 113: 103233.
- [12] Kahandawa G C, Epaarachchi J, Wang H, et al. Use of FBG sensors for SHM in aerospace structures[J]. *Photonic Sensors*, 2012, 2(3): 203-214.
- [13] Lupi C, Felli F, Brotzu A, et al. Improving FBG sensor sensitivity at cryogenic temperature by metal coating[J]. *IEEE Sensors Journal*, 2008, 8(7): 1299-1304.
- [14] Zhang N Z, Davis C, Chiu W K, et al. Fatigue performance of type I fibre Bragg grating strain sensors[J]. *Sensors*, 2019, 19(16): 3524.
- [15] Zhang N Z, Turk S, Davis C, et al. Fatigue performance of type I and type II fibre Bragg gratings fabricated by femtosecond laser inscription through the coating[J]. *Sensors*, 2022, 22(22): 8812.
- [16] Ding Y, Wang P, Yu S. A new method for deformation monitoring on H-pile in SMW based on BOTDA[J]. *Measurement*, 2015, 70: 156-168.
- [17] Mohamad H, Tee B P. Instrumented pile load testing with distributed optical fibre strain sensor[J]. *Jurnal Teknologi*, 2015, 77(11): 1-7.
- [18] Ding Y, Shi B, Zhang D. Data processing in BOTDR distributed strain measurement based on pattern recognition[J]. *Optik*, 2010, 121(24): 2234-2239.
- [19] Gao L, Han C, Xu Z Q, et al. Experimental study on deformation monitoring of bored pile based on BOTDR[J].

- Applied Sciences, 2019, 9(12): 2435.
- [20] He J P, Zhou Z, Chen G D, et al. Measurement accuracy improvement of Brillouin signal using wavelet denoising method [J]. Proceedings of SPIE, 2009, 7293: 72930B.
- [21] Luna Technologies. ODISI-A optical distributed sensor interrogator (Users Guide) [EB/OL]. [2023-06-05]. <https://lunainc.com/sites/default/files/assets/files/resource-library/ODISI-A-Users-Guidev1.2.pdf>.
- [22] Barrias A, Casas J R, Villalba S. Embedded distributed optical fiber sensors in reinforced concrete structures: a case study[J]. Sensors, 2018, 18(4): 980.
- [23] Sweeney D C, Sweeney D M, Petrie C M. Graphical optimization of spectral shift reconstructions for optical backscatter reflectometry[J]. Sensors, 2021, 21(18): 6154.
- [24] Gebhart T M J, Seeberg M, Çelik H, et al. Advanced methods for the detection of measurement outliers in fibre optic strain measurements for the application of non-destructive monitoring of the structural health of composite over-wrapped pressure vessels of type-IV [EB/OL]. [2023-08-02]. <http://dx.doi.org/10.2139/ssrn.4515167>.
- [25] Bado M F, Casas J R, Gómez J. Post-processing algorithms for distributed optical fiber sensing in structural health monitoring applications[J]. Structural Health Monitoring, 2021, 20(2): 661-680.
- [26] Mustafa S, Sekiya H, Maeda I, et al. Identification of external load information using distributed optical fiber sensors embedded in an existing road pavement[J]. Optical Fiber Technology, 2021, 67: 102705.

An Adaptive Post-Processing Algorithm for Strain Reading Anomalies in Distributed Optical Fiber Sensors

Liang Zhihong¹, Deng Kaiwen¹, Ma Yunlong², Wang Minghua¹, Liu Debo², Wu Huiqiang², Wang Yishou^{1*}

¹*School of Aerospace Engineering, Xiamen University, Xiamen 361005, Fujian, China;*

²*Beijing Institute of Astronautical Systems Engineering, Beijing 100076, China*

Abstract

Objective Distributed optic fiber sensor (DOFS) is widely used for status monitoring and damage detection of aerospace vehicles due to its ability to achieve large-area and high-density sensing of structures. However, in the face of uncertainties caused by the harsh service environment of aerospace, a phenomenon of strain reading anomalies (SRAs) occurs in DOFS measurements. These SRAs result in significant strain peaks occurring in localized regions or at specific moments in time, thereby posing challenges for DOFS to accurately measure physical quantities and making it even more difficult to interpret these measurements. To minimize the negative effects of SRAs, some researchers have adopted a series of data processing methods, such as polynomial fitting method, spectral shift quality (SSQ) method, geometrical threshold method (GTM), and polynomial interpolation comparison method (PICM). Although these data processing methods are effective in reducing random errors in measurement data, they fail to completely remove the phenomenon of SRAs, and there is still a risk of removing highly reliable measurement readings. Meanwhile, the above methods still use the fixed threshold method to detect and determine the anomalies, and the determination of the fixed threshold relies on manual experience, which has low detection efficiency and a high false alarm rate, thus limiting its application in complex service environments. Therefore, we propose an intelligent adaptive post-processing method for detecting and quickly removing SRAs from DOFS.

Methods The proposed algorithm, namely the adaptive geometrical threshold offset method (AGTOM), adopts the K-means clustering method to adaptively determine thresholds for distinguishing differences of thresholds caused by various structural features and service conditions. A continuous geometric correction is implemented on the distorted strain curves to effectively eliminate SRAs. To verify the effectiveness of the proposed method, a case study is conducted on the processing of DOFS measurement data collected during the pressure cycling test of a fuel tank. The Pearson correlation coefficient (PCC) is utilized to evaluate the correlation between the post-processing curves and normal strain curves. Besides, a comparison is conducted with other post-processing algorithms (GTM and PICM) to highlight the advantages of the proposed method.

Results and Discussions Based on their different response characteristics, SRAs can be classified into two categories: harmless strain reading anomalies (HL-SRAs) and harmful strain reading anomalies (HF-SRAs). For the HL-SRAs, AGTOM consistently yields optimal post-processing results with PCC values not less than 0.965. It is followed by GTM, whose PCC values are all not less than 0.798. However, GTM interferes when HL-SRAs are coupled with NaN values. In addition, PICM achieves promising processing results only in the first typical case (i. e., sparsely distributed HL-

SRAs). In the remaining three typical cases, PICM still produces distortions with a PCC value not greater than 0.512. Importantly, both GTM and PICM exhibit distorted post-processing curves when HL-SRAs are coupled with NaN values. For HF-SRAs, AGTOM also yields the highest post-processing results, with no PCC value lower than 0.917. The susceptibility of PICM to curve distortion accurately reflects the difference between HF-SRAs and HL-SRAs because the main characteristic of the former is that strain values follow an erroneous strain response or frequent sudden changes. It is difficult to determine the change in strain increment using PICM because it detects and removes SRAs by comparing the fitted value with the original value. Compared with PICM, GTM takes into account the sudden changes of the strain increment, resulting in improved post-processing results when HF-SRAs consist of densely changed SRAs. However, similar to HL-SRAs, both GTM and PICM show worsened post-processing results when NaN values interfere with HF-SRAs, indicating lower algorithmic robustness for GTM and PICM compared to AGTOM.

Conclusions The proposed algorithm AGTOM is able to distinguish the differences in thresholds due to different structural characteristics and service environments. The K-mean clustering algorithm uses an internal evaluation metric, namely Davies-Bouldin index (DBI), to characterize the clustering effect of strain increments. The threshold is determined by obtaining the optimal k value. For the HL-SRAs, both GTM and AGTOM methods can achieve satisfactory processing results. However, PICM is susceptible to interference when facing densely distributed and coupled HL-SRAs, leading to serious distortions in its post-processing curves. For HF-SRA, the post-processing curves of the other two algorithms are distorted to varying degrees, except for AGTOM, which exhibits the highest PCC compared to the normal strain curve. For both HL-SRAs and HF-SRAs, GTM and PICM are interfered with when SRAs are coupled with NaN, indicating that the algorithmic robustness of both is lower than that of AGTOM. To further validate the effectiveness of AGTOM, it will still be necessary to test AGTOM by applying it to different experimental scenarios in the future.

Key words fiber optics; Rayleigh scattering; distributed optical fiber sensing; K-mean clustering; adaptive threshold; composite tank