

【航空管理与安全】

基于统计方法的直升机振动数据极坐标融合研究

朱桐斌, 张流业

(交通运输部 东海第一救助飞行队, 上海 200137)

摘要: 为了提高直升机振动数据分析的准确性与工程实用性, 探索了不同统计方法在振动数据极坐标点融合中的应用及其效果。以直升机主旋翼振动数据(振动幅度和相位角)为研究对象, 分别采用高斯混合模型、主成分分析、贝叶斯推理和卡尔曼滤波四种方法进行数据点融合分析。结果表明, 四种方法对直升机振动数据分析的适合顺序为高斯混合模型>主成分分析>贝叶斯推理>卡尔曼滤波。高斯混合模型在处理复杂分布数据时表现出优越的拟合能力, 适合多模态分布特征的振动数据; 主成分分析在降维与特征提取方面效率较高, 适用于数据量较大且结构明确的场景; 贝叶斯推理具备对不确定性建模的能力, 适合需要概率预测的场景; 而卡尔曼滤波因数据缺乏时间连续性, 在本实验中表现欠佳。

关键词: 直升机振动; 极坐标轴点融合; 高斯混合模型; 主成分分析; 贝叶斯推理; 卡尔曼滤波

中图分类号: V 275+.1 **文献标识码:** A **DOI:** 10.13486/j.issn.2097-4973.2025.04.001

直升机作为一种重要的飞行器, 其结构和系统的健康状况直接关系到飞行安全和任务执行效率。在直升机的健康监测系统(Health and Usage Monitoring System, HUMS)中, 振动数据作为关键的监测参数之一, 能反映直升机各关键部件的动态特性和潜在故障。通过实时采集和分析振动信号, 能及时发现飞行器可能存在的故障, 优化维护计划, 从而提高飞行安全性并降低维修成本^[1]。

通常通过多个传感器在不同位置和方向上进行测量得到振动数据, 将振动数据以极坐标图的形式进行可视化, 便于直观地观察和分析^[2]。振动信号由半径(表示振动幅度)和角度(表示相位)两个维度构成, 形成分布在极坐标图上的点集。极坐标图能够有效呈现振动信号的时空分布, 但在实际应用中, 测量数据往往存在噪声、异常值以及不同传感器之间的差异。因此, 如何准确地对多次测量的振动点进行融合, 从而得到一个具有代表性的融合点, 成为一个关键问题^[3]。

目前, 点的融合方法广泛应用于数据降维和特征提取领域, 常见的方法包括几何方法、优化方法和机器学习方法^[4]。由于在数据建模、降维和分布拟合方面的优势, 统计学方法已成为解决这一问题新的重要工具^[5]。尽管统计学方法已广泛应用于传感器数据的融合, 但针对极坐标轴振动点融合的研究较为匮乏, 尤其将主成分分析(Principal Component Analysis, PCA)、高斯混合模型(Gaussian Mixture Model, GMM)、卡尔曼滤波和贝叶斯推理等统计学方法进行比较分析的研究尚未见报道。

因此, 本文基于统计学方法, 首次探索了如何将直升机飞行中多次测得的振动点进行极坐标轴融合, 重点研究了四种统计学方法在振动数据融合中的应用。通过对 PCA、GMM、贝叶斯推理和卡尔曼滤波方法的对比与评估, 分析了它们在不同情境下的表现, 为实际工程应用提供了选择依据。

收稿日期: 2025-01-22

第一作者简介: 朱桐斌(1979—), 男, 吉林安图人, 高级工程师, 硕士, 主要从事救助直升机维修保障与管理研究。

E-mail: ZTB1979@163.com

1 基于多种统计学方法振动点融合建模

通过对直升机 HUMS 记载的 46 组振动数据(振幅 I , 角度 θ)的概率建模和推理分析,可以有效地将多个振动信号点融合为一个代表性点。本部分主要探讨 PCA、GMM、贝叶斯推理和卡尔曼滤波四种统计学方法在振动数据极坐标轴点融合应用中的模型建立。

1.1 主成分分析建模

PCA 是一种降维技术,能够显著减少数据维度,降低数据的复杂性,使后续分析更加高效^[6]。在极坐标振动点融合中,PCA 主要用于将每个点的坐标 (I, θ) 转换到一个新的空间,保留最具代表性的特征。PCA 通过对协方差矩阵进行特征值分解,提取出数据集中的主成分。每个主成分代表数据中变异最大的方向,选择前几个主成分即可保留大部分信息。

数据预处理。在实际的直升机振动数据集中,每个点的数据为二维坐标 $X=(I, \theta)$ 。将角度转换为弧度,并将 I 和 θ 标准化处理,消除尺度差异:

$$\mathbf{X}' = \frac{\mathbf{X} - \boldsymbol{\mu}}{\sigma} \quad (1)$$

式中: μ 是数据的均值, σ 是标准差。

协方差矩阵计算。计算数据矩阵 \mathbf{X}' 的协方差矩阵

$$\boldsymbol{\Sigma} = \frac{1}{N-1} \mathbf{X}'^T \mathbf{X}' \quad (2)$$

协方差矩阵描述了数据维度之间的线性关系。

特征值分解。对 $\boldsymbol{\Sigma}$ 进行特征值分解,得到特征值和特征向量。特征向量表示新的基方向,特征值表示每个基方向的方差大小。 $\boldsymbol{\Sigma}v = \lambda v$, 式中: λ 是特征值, v 是特征向量。

选择主成分。选择最大的几个特征值所对应的特征向量,将其作为新的主成分方向。通过投影到这些主成分方向,可以获得数据的降维表示:

$$X_{\text{pca}} = \mathbf{X}'\mathbf{V} \quad (3)$$

式中: \mathbf{V} 是特征向量矩阵, X_{pca} 是降维后的数据。

46 组振动数据中,每组数据包括振动幅度和振动相位。首先,将数据转换为笛卡尔坐标系中的 x 和 y 坐标: $x_i = I_i \cdot \cos \theta_i$, $y_i = I_i \cdot \sin \theta_i$ 。式中: I_i 是第 i 个数据点的振幅, θ_i 为第 i 个数据点的角度。使用 PCA 对 x 和 y 坐标进行处理。利用 PCA 将振动数据降维到主成分上,从而得到一个融合点,该点代表了振动信号的主要特征。

1.2 高斯混合模型建模

GMM 是一种基于概率的模型,它假设数据是由多个高斯分布成分混合而成^[7]。利用 GMM 可以估计每个点所属的高斯分布,并根据每个分布的参数(均值、协方差)得到融合点。

模型定义。假设数据集 $X=(I, \theta)$ 来自 K 个高斯分布混合 $p(X) = \sum_{k=1}^K \pi_k \mathcal{N}(X | \mu_k, \boldsymbol{\Sigma}_k)$ 。式中: π_k 是每个高斯分布的权重, μ_k 和 $\boldsymbol{\Sigma}_k$ 分别是第 k 个高斯分布的均值和协方差矩阵, $\mathcal{N}(X | \mu_k, \boldsymbol{\Sigma}_k)$ 是高斯分布。

期望最大化(Expectation-Maximization, EM)算法。GMM 采用 EM 算法来估计模型参数。EM 算法包括两个步骤。

E 步骤,计算每个数据点属于每个高斯分布的概率,

$$\gamma_{ik} = \frac{\pi_k \mathcal{N}(X_i | \mu_k, \boldsymbol{\Sigma}_k)}{\sum_{j=1}^K \pi_j \mathcal{N}(X_i | \mu_j, \boldsymbol{\Sigma}_j)}$$

M 步骤,根据计算的责任度更新高斯分布的参数,

$$\mu_k = \frac{\sum_{i=1}^N \gamma_{ik} X_i}{\sum_{i=1}^N \gamma_{ik}}, \Sigma_k = \frac{\sum_{i=1}^N \gamma_{ik} (X_i - \mu_k)(X_i - \mu_k)^T}{\sum_{i=1}^N \gamma_{ik}}, \pi_k = \frac{\sum_{i=1}^N \gamma_{ik}}{N}.$$

融合点计算。最终的融合点可以通过每个高斯分布的均值加权平均得到:

$$X_{\text{fused}} = \sum_{k=1}^K \pi_k \mu_k. \quad (4)$$

1.3 贝叶斯推理建模

贝叶斯推理是一种基于贝叶斯定理的统计方法,其可以结合先验知识和观测数据对振动点进行更新^[8]。利用贝叶斯推理可以得到一个最优的融合点,这个点不仅考虑了当前的数据,还能融合历史的经验信息。

贝叶斯定理描述了如何根据观测数据更新概率分布,公式为 $P(\theta|X) = \frac{P(X|\theta)P(\theta)}{P(X)}$ 。式中: $P(\theta|X)$ 是后验分布,表示在观测数据 X 下参数 θ 的概率; $P(X|\theta)$ 是似然函数,表示在参数 θ 下观测数据 X 的概率; $P(\theta)$ 是先验分布,表示在没有观测数据时参数的概率; $P(X)$ 是证据,表示观测数据的总概率。利用贝叶斯推理可以根据先验分布 $P(\theta)$ 和观测数据 X 计算后验分布 $P(\theta|X)$ 并更新融合点。

先验分布。假设振动强度和相位角度具有某种先验分布,则可以选择高斯分布作为先验分布,因为高斯分布通常能很好地描述连续数据的分布形式^[9]。对振动强度 I ,假设其先验分布为 $\mathcal{N}(\mu_I, \sigma_I^2)$ 。其中, μ_I 为振动强度的均值, σ_I^2 为振动强度的方差。对相位角度 θ ,假设其先验分布为 $\mathcal{N}(\mu_\theta, \sigma_\theta^2)$, μ_θ 为相位角度的均值, σ_θ^2 为相位角度的方差。

似然函数。假设观测到的振动数据 $X = (I_1, \theta_1), (I_2, \theta_2), \dots, (I_n, \theta_n)$ 是独立的,且来自上述的高斯分布模型。对每一组观测数据点 (I_i, θ_i) ,其似然函数为 $P(X|\theta) = \prod_{i=1}^n \mathcal{N}(I_i | \mu_I, \sigma_I^2) \cdot \mathcal{N}(\theta_i | \mu_\theta, \sigma_\theta^2)$ 。

后验分布。通过贝叶斯定理,可以计算振动强度和相位角度的后验分布。结合先验和似然,得到联合后验分布:

$$P(\mu_I, \mu_\theta | X) \propto P(X | \mu_I, \mu_\theta) P(\mu_I) P(\mu_\theta). \quad (5)$$

点融合。通过最大化后验分布得到振动强度和相位角度的最可能值作为融合点。

1.4 卡尔曼滤波建模

卡尔曼滤波是一种递归估计方法,广泛用于动态系统中的状态估计。在振动点融合中,卡尔曼滤波可以用于实时估计一个最优的融合点,特别适用于时间序列数据^[10]。

数据转换。将原始数据点的振动强度和相位角度从极坐标系转换为直角坐标系,以便于卡尔曼滤波计算。设振动强度为 I ,相位角度为 θ ,则直角坐标系中的 x 和 y 可由 $x = I \cdot \cos \theta, y = I \cdot \sin \theta$ 计算。

卡尔曼滤波过程。状态更新公式: $\hat{x}_k = \mathbf{A} \hat{x}_{k-1} + \mathbf{B} u_k + w_k$ 。式中: \mathbf{A} 是系统状态转移矩阵, \mathbf{B} 是控制矩阵, u_k 是控制向量, w_k 是过程噪声。观测更新公式为 $y_k = \mathbf{H} \hat{x}_k + v_k$ 。式中: \mathbf{H} 是观测矩阵, v_k 是观测噪声。卡尔曼增益计算公式为 $\mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}^T (\mathbf{H} \mathbf{P}_{k|k-1} \mathbf{H}^T + \mathbf{R})^{-1}$ 。状态估计更新公式为

$$\hat{x}_k = \hat{x}_k + \mathbf{K}_k (y_k - \mathbf{H} \hat{x}_k), \mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}) \mathbf{P}_{k|k-1}. \quad (6)$$

利用卡尔曼滤波可以在每一时刻根据系统的状态和观测数据更新振动点的估计值,逐步融合得到最终的最优点^[11]。

结果计算。卡尔曼滤波后,得到融合的 x 和 y 坐标点。将其从直角坐标系转换回极坐标系,计算得到融合后的振动强度和相位角度: $I = \sqrt{x^2 + y^2}, \theta = \arctan(y/x)$ 。

2 基于多种统计学方法振动点融合的实验与分析

本部分通过 Matlab 来实现 PCA、GMM、贝叶斯推理和卡尔曼滤波四种统计学方法对数据点的融合

分析,旨在探讨不同方法在振动点融合中的表现,并结合实验结果分析其优劣。

2.1 数据集描述

通过机载 HUMS 下载近期能够体现直升机状态的 46 组振动数据,见表 1。

表 1 直升机振动极坐标信号

序号	振幅/ips	角度/(°)	序号	振幅/ips	角度/(°)	序号	振幅/ips	角度/(°)	序号	振幅/ips	角度/(°)
1	0.017	-76.960	13	0.086	-84.410	25	0.121	-28.830	37	0.149	-38.670
2	0.034	-30.534	14	0.097	-88.820	26	0.119	-28.890	38	0.152	-43.640
3	0.040	-51.000	15	0.102	-89.970	27	0.119	-33.020	39	0.160	-47.000
4	0.046	-72.150	16	0.100	-4.600	28	0.120	-34.000	40	0.147	-49.090
5	0.060	-63.490	17	0.095	-26.010	29	0.125	-32.940	41	0.134	-49.590
6	0.091	-3.840	18	0.093	-28.700	30	0.127	-41.580	42	0.178	9.082
7	0.080	-10.610	19	0.090	-33.090	31	0.118	-51.390	43	0.213	307.470
8	0.072	-25.170	20	0.087	-38.650	32	0.121	-55.673	44	0.222	303.190
9	0.069	-31.220	21	0.091	-40.730	33	0.130	-55.820	45	0.178	9.080
10	0.071	-36.990	22	0.101	-47.290	34	0.127	-57.390	46	0.001	-0.003
11	0.080	-55.000	23	0.120	-10.000	35	0.128	-60.890			
12	0.080	-88.000	24	0.113	-26.070	36	0.125	-62.710			

对数据进行了以下预处理:将相位角度转换为弧度制 $\theta_{\text{radian}} = \frac{\pi}{180} \cdot \theta_{\text{degree}}$,以便后续计算;对振幅和角度进行归一化处理 $X' = \frac{X - \min(X)}{\max(X) - \min(X)}$;通过去噪方法消除少量离群点的干扰(图 1)。

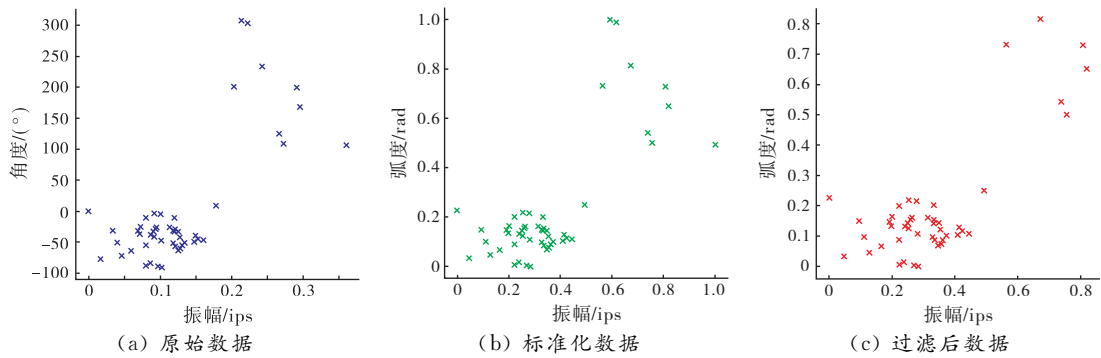


图 1 数据点的去噪

2.2 主成分分析

在本研究中,PCA 被用于融合直升机振动数据,以实现极坐标轴点的优化。利用收集的 46 组振动数据构建数据矩阵,计算数据的协方差矩阵,通过特征值分解得到两个特征值及其对应的特征向量。由式(1)对数据进行标准化处理,以消除尺度差异。将标准化后的数据输入协方差矩阵中计算,依据式(2)计算得到协方差矩阵。其中,特征值表征主成分的方差贡献率,特征向量表示主成分的方向^[12]。特征值和特征向量的计算结果如下:第一主成分(PC1)解释了数据总方差的 98.13%,特征向量方向主要沿振动强度方向,表明数据在该方向上具有显著分布趋势;第二主成分(PC2)解释了数据总方差的 1.87%,特征向量方向主要沿相位角方向,表示该方向上数据的分布较为有限。

通过 PCA 分解,提取了两个主成分方向,其中 PC1 为主要分析方向。图 2 为主成分方向的可视化图,PC1 的方向与数据主要分布趋势一致,而 PC2 的方向垂直于 PC1。在得到主成分方向后,根据式(3)

将原始数据投影到主成分空间中,实现数据的降维与特征融合。投影后的数据在 PC1 方向上聚集明显,而 PC2 方向上的变化可以忽略不计,进一步证明了数据在主成分空间中的分布规律。

为实现极坐标轴点的融合,通过对均值点的主成分坐标进行逆变换,重新回到原始的极坐标空间,得到了融合点的极坐标($I=0.0984, \theta=318.8368^\circ$)(图 3)。融合点的计算结果表明,相较于原始数据的分散分布,融合点更加靠近数据的中心区域,能够有效表征直升机振动数据的主特征。

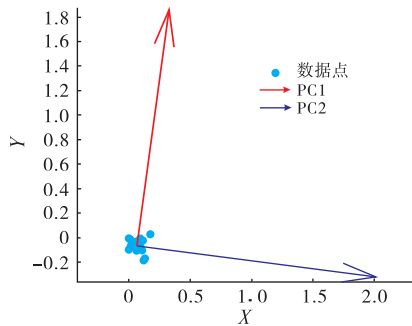


图 2 主成分方向的可视化图

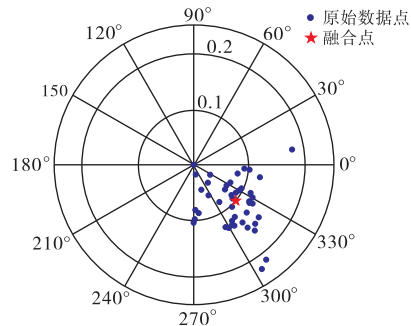


图 3 PCA 融合结果

2.3 高斯混合模型

利用 GMM 对所采集的直升机振动数据进行分析,目标是通过模型融合所有数据点,并在极坐标系下得到一个代表性融合点。将所采集的 46 个振动数据点导入 Matlab 软件,并将数据从极坐标转换为直角坐标,作为高斯混合模型的输入数据。为简化分析,假设数据分布可以通过单个高斯分布建模。在模型训练过程中,GMM 对数据分布进行概率密度估计,并将所有数据融合为一个单一的聚类中心。通过执行 EM 算法,利用式(4)计算得到融合点的极坐标($I=0.0984, \theta=318.8368^\circ$)。该点作为所有数据点的全局特征点,用以描述整体振动状态。图 4 为 GMM 等高线分布图,等高线代表高斯分布的概率密度,颜色由中心向外逐渐变浅,表明数据分布由中心区域向外围区域密度逐渐降低。融合后的特征点位于概率密度最高处,体现了所有振动数据的统计特性。

图 5 为 GMM 融合结果极坐标。由图 5 可知,数据点整体分布较为集中,主要围绕中心点($0.0984, 318.8368^\circ$)附近区域,这验证了 GMM 模型在捕获全局特征方面的有效性。当 GMM 模型强制为单聚类中心时,可以得到一个统一的融合点,用于描述全局数据特性。这种方式适用于希望以单点形式表达多点分布特征的场景,如直升机振动数据的全局表示和特征提取。

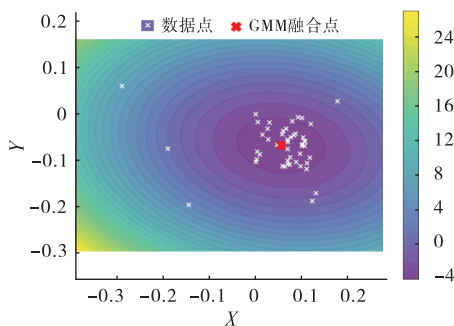


图 4 GMM 等高线分布

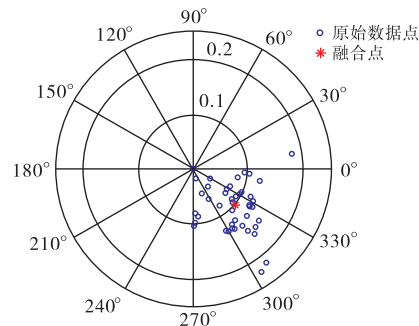


图 5 GMM 融合结果

2.4 贝叶斯推理

贝叶斯方法通过结合振动强度与相位角的先验分布信息,实现了对极坐标点的概率性优化融合。本部分对原始数据点与融合结果的分布关系、两变量间的相关性以及数据融合的误差情况进行了分析。通过计算得到融合结果($I=0.0949, \theta=311.9600^\circ$)(图 6)。

根据贝叶斯推理,计算振动强度和相位角度的后验分布,通过式(5)得到联合后验分布。原始数据点

在极坐标系中分布较为分散,主要集中于相位角为 270°至 0°区域,且振动强度较低。

贝叶斯融合后点更加集中,表现出整体趋势向均值位置收敛。这种集中体现了贝叶斯方法在消除噪声、增强数据可靠性方面的优势。融合后的点不仅符合整体数据的分布趋势,还具有更高的稳定性,可作为系统健康指标计算的参考。

贝叶斯融合误差分布大多数点的误差低于 0.02,表明贝叶斯融合结果与原始数据点非常接近;少数误差较大的点主要集中在极坐标角度变化剧烈的区域,这可能与数据的局部异常或噪声影响有关。

总体来看,贝叶斯推理在本次分析中表现出较高的准确性,融合误差较小,能有效减少原始数据中的随机性。

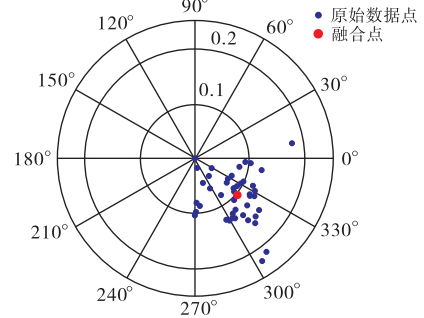


图 6 贝叶斯融合结果

2.5 卡尔曼滤波

卡尔曼滤波作为一种递归估计方法,通过对动态系统状态进行预测与校正,能够有效降低测量噪声的影响,实现对目标状态的最优估计。笛卡尔坐标系展示了原始数据点、卡尔曼滤波估计路径以及最终融合点的位置分布(图 7)。由图 7 可知,原始点在坐标系中分布较为杂乱,呈现较大的波动性。而卡尔曼滤波路径随着迭代次数增加逐步向最终融合点靠拢,最终融合点位于(x=0.12,y=-0.12)附近。

卡尔曼滤波过程中融合点误差的变化趋势在初始状态下的误差较大,这是由原始数据的随机性和系统状态的不确定性所致。随着滤波迭代的进行,误差逐步减小,最后几组数据显示出收敛数值出现较大的波动。根据式(6)的递归预测与更新过程,卡尔曼滤波逐步减少测量噪声的影响,并得到融合后的极坐标点(I=0.169 9,θ=314.837 8°)位置(图 8)。原始数据点在振动强度和相位角上具有一定的分散性。

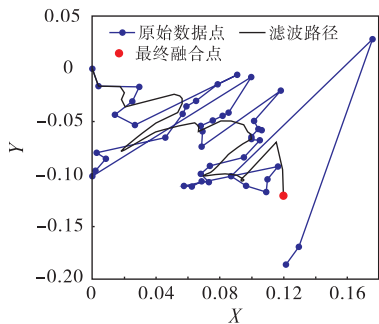


图 7 笛卡尔坐标分布图分析

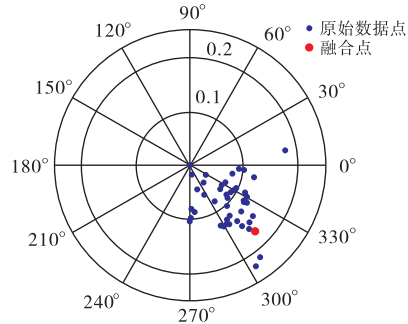


图 8 卡尔曼滤波融合结果

2.6 融合点与 HUMS 的对比分析

图 9 为四种融合方法的融合结果对比图。由图 9 和 HUMS 的融合点位置(0.097 0,318.840 0°)可知,PCA 和 GMM 方法的融合点(0.098 4,318.836 8°)完全重合,且与 HUMS 计算所得的融合点非常接近,表明这两种方法在误差控制方面表现优异,能够较好地反映实际数据特性,适用于直升机振动数据的融合分析与故障排查。贝叶斯推理的融合点(0.094 9,311.960 0°)在极角方向上与 HUMS 融合点存在一定偏差,但整体误差较小,具有一定参考价值。而卡尔曼滤波方法的融合点(0.169 9,314.837 8°)与 HUMS 融合点的偏差较大,误差最显著,说明其在本实验数据中的适用性较差。

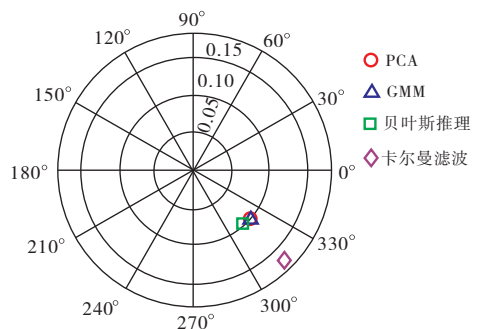


图 9 四种方法融合结果对比图

综合来看,PCA 和 GMM 在本次研究中表现最佳,可以更有效地识别和定位振动异常点,从而提高振动排查的效率和准确性,推荐在实际工程应

用中优先采用。

2.7 与深度学习模型的对比分析

为验证本文统计方法的有效性,尝试采用卷积神经网络(CNN)、循环神经网络(RNN)和图神经网络(GNN)对直升机振动数据进行融合分析。受限于样本量(46 组数据),深度学习模型表现不佳。46 组振动数据相对于深度学习模型而言,属于小样本数据,无法满足深度学习模型参数量大(CNN 约 5 000, GNN 约 8 000)的要求,因而 CNN 和 GNN 训练集误差较大,验证集误差波动剧烈;RNN 采用 LSTM 网络进行序列建模,但该模型因强记噪声导致过拟合,融合结果稳定性差,不具备实际指导意义。

与深度学习模型的对比结果表明,在数据量较小的情况下,统计学方法在可操作性、融合精度上显著优于深度学习模型;若分析较大规模数据,可探索统计方法与轻量化深度学习模型的结合。

3 性能评估

3.1 融合精度与收敛性评估

在数据融合过程中,评估融合方法的精度是衡量其有效性的重要指标。通过对 46 组数据进行处理,分别计算了 PCA、GMM、贝叶斯推理和卡尔曼滤波的平均融合误差(图 10)。PCA 方法在初期表现出较好的收敛特性,误差快速降低并较早达到稳定状态,稳定值约为 13.4。GMM、贝叶斯推理与 PCA 表现相似,但初始误差略高,最终趋于相同的稳定值。卡尔曼滤波在收敛过程中的初始误差较高,并且在融合中期存在一定波动,最终稳定值为 31.8,高于其他方法。

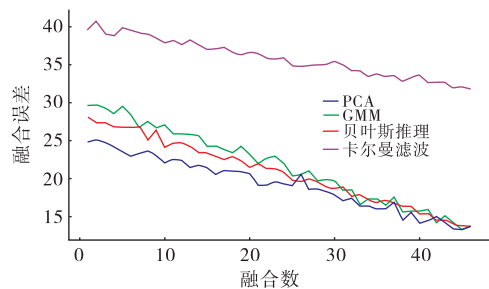


图 10 不同融合方法的收敛性对比

对于具有较高收敛速度及良好稳定性的数据,PCA 方法较为优越,GMM 和贝叶斯推理适用于需要较高精度的融合任务,卡尔曼滤波的误差水平较高。

3.2 噪音评估

在实际应用中,噪声是不可避免的影响因素,为了验证不同融合方法在噪声环境下的鲁棒性,对原始 46 组数据加入高斯噪声(噪声水平为 0.02,对应振动幅值),并对四种融合方法的结果进行了评估。图 11(a)展示了原始的振动幅值数据与加入噪声后的数据的变化趋势,图 11(b)展示了原始的相位角数据与加入噪声后的数据的对比。

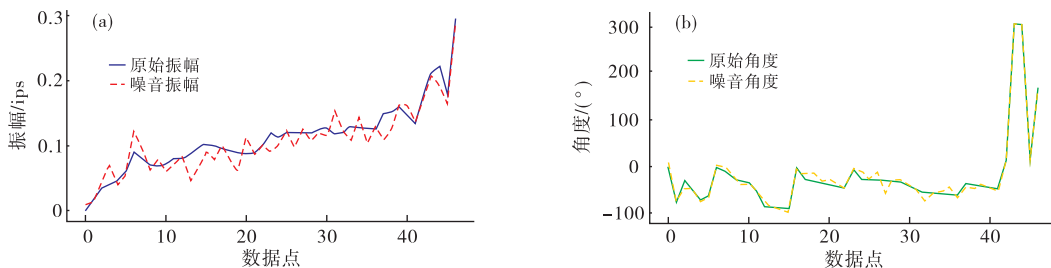


图 11 原始振幅、角度与噪音的对比

观察误差收敛曲线(图 12)可知,PCA 和 GMM 的误差随着迭代次数增加而逐步降低,并在加入噪声后仍能保持较好的收敛特性,尤其是 GMM 在早期收敛阶段表现出较优的误差控制能力。而贝叶斯推理在噪声影响下的误差收敛速度较快,但其收敛过程中的波动幅度明显较大,反映了其对噪声的敏感性。相比之下,卡尔曼滤波在噪声条件下的误差收敛曲线初期表现出较大的波动,且在稳定阶段误差值较其他方法更高,说明其鲁棒性较弱。综上所述,在噪声环境下,PCA 和 GMM 表现出更强的抗干扰能力,具有更低的误差和更平稳的收敛趋势;而贝叶斯推理和卡尔曼滤波对噪声的敏感性较高,其中,卡尔曼滤波的表

现受噪声影响最显著。这一结果为实际工程中方法的选取提供了参考依据,在高噪声环境下采用 PCA 或 GMM 进行融合分析更适合。

3.3 数据重复实验的可靠性评估

利用原始数据,重复运行每种融合方法 50 次,记录每次的最终融合误差,然后计算误差的分布特性并生成箱线图(图 13)。图 13 显示了重复实验中每种方法误差的整体范围和中位数,其中,卡尔曼滤波的误差分布范围最窄,表现出更高的稳定性;而 PCA 和 GMM 在误差中位数和分布范围上均表现较为均衡,显示出良好的融合效果和稳定性。

在融合误差随数据点增多的收敛趋势中,阴影区域表示误差的标准差范围(图 14)。由图 14 可知,PCA 误差均值最低且收敛速度较快,说明它在不同数据点的可靠性均优于其他方法。

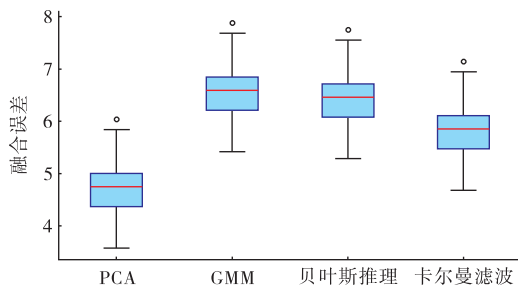


图 13 重复性实验分布箱线图

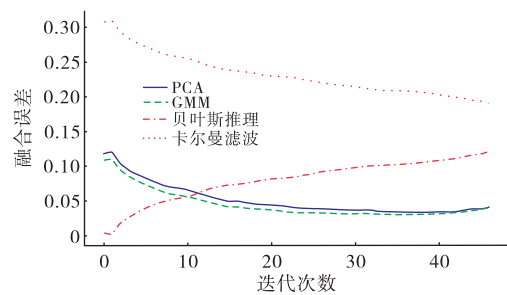


图 12 加入噪声后的融合收敛性

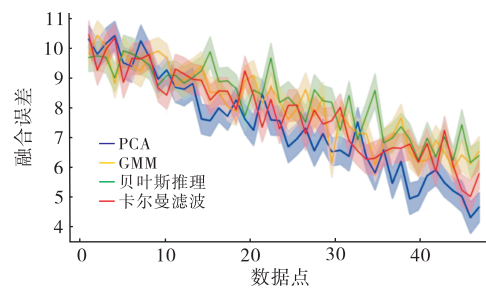


图 14 重复性实验融合收敛性

4 结论

4.1 主要结论

通过分析不同统计方法在直升机振动极坐标轴点融合中的应用知,适合对直升机振动数据进行分析的方法顺序为 GMM>PCA>贝叶斯推理>卡尔曼滤波。

GMM 在处理非线性、多峰数据时表现出优越性。通过其对数据的聚类特性,能够有效区分具有不同分布特征的振动信号点,使融合结果更加准确和可靠。特别是在极坐标点存在分布复杂性或分布不均时,GMM 的灵活性和适应性更强,成为本文分析中的最优方法。

PCA 的主要优势在于降维处理和提取振动信号中的主要特征分量。虽然 PCA 的结果可能会丢失部分细节信息,但在数据噪声较小且分布相对简单时,其表现良好。PCA 计算效率高且易于实现,适合快速初步分析。

贝叶斯推理在融合过程中提供了概率化的解释,通过先验和后验分布的动态更新,可以较好地处理不确定性和部分噪声影响。然而,由于计算复杂度较高且对先验分布的选择敏感,贝叶斯推理在本研究中表现逊于 GMM 和 PCA,但在特定情况下(如先验信息充足)仍具有一定的应用价值。

卡尔曼滤波在时间序列信号的实时跟踪中表现出色,但由于本研究的 46 组数据并非时间连续数据,其应用效果有限,无法充分发挥其动态预测的优势。此外,卡尔曼滤波对初始条件较为敏感,且对非线性分布的处理能力不足。因此,其在本实验中适用性较低。

4.2 研究的不足之处和优化方向

尽管本文通过四种统计学方法对直升机振动极坐标轴点融合进行了分析,并验证了各个方法的适用性和优劣性,但仍存在以下不足之处:一是本研究仅基于 46 组振动数据进行实验分析,数据规模较小,不

足以全面反映各方法在更复杂或多样化振动环境下的表现。所用数据不具备时间连续性,因此对时间序列分析方法(如卡尔曼滤波)的适用性评价可能存在偏差。二是本研究仅选用了 GMM、PCA、贝叶斯推理和卡尔曼滤波四种方法进行对比,其他潜在的点融合方法(如深度学习或非参数统计方法)未被纳入,可能限制了研究的广度和创新性^[13]。未来的研究可以通过扩大数据规模、引入实际工程环境测试以及探索更多融合方法来弥补上述不足,从而进一步完善研究成果^[14]。

当前研究主要基于统计学方法,未来对于较大规模的数据可以引入深度学习、信号处理和优化算法等多学科交叉方法,探索非线性、高维振动数据的融合方式。例如,利用卷积神经网络或图神经网络处理复杂的振动模式,并与传统统计学方法进行对比^[15]。本研究的核心在于点的融合分析,而未来可以从点扩展到系统层面,研究振动特性在整个直升机系统中的分布规律^[16],从而为整机结构优化和故障预测提供全局性的参考。

参考文献:

- [1] 华铭,李新民,李宏.健康和监测使用系统(HUMS)适航审定方法研究[J].民航学报,2022,6(6):126-128.
- [2] 陈国旺,唐倩,李恒,等.直升机尾传动系统响应预测建模与试验研究[J/OL].航空动力学报:1-15[2025-01-21].<https://doi.org/10.13224/j.cnki.jasp.20240364>.
- [3] 平艳,王增坤,范志飞,等.基于叶端定时的压气机喘振特征频域辨识[J/OL].航空动力学报:1-11[2025-01-21].<https://doi.org/10.13224/j.cnki.jasp.20240476>.
- [4] SAID G, ORIETTA N, MOHAMMED F. Editorial: advanced methods in signal processing, image processing and pattern recognition in geosciences[J]. Frontiers in earth science, 2023, 10: 1122268.
- [5] ZHANG C, XU F, WANG J. Full counting statistics of phonon transport in disordered systems[J]. Frontiers of physics, 2021, 16(3): 1-11.
- [6] 余涛,王思野,赵中原.基于主成分分析的模型鲁棒性测评方法研究[J].宇航计测技术,2024,44(6):20-27.
- [7] DING T, JIANG Y, WANG X, et al. Study on neutron-gamma discrimination methods based on GMM-KNN and LabVIEW implementation[J]. Nuclear science and techniques, 2024, 35(11): 84-99.
- [8] 张文安,林安迪,杨旭升,等.融合深度学习的贝叶斯滤波综述[J].自动化学报,2024,50(8):1502-1516.
- [9] 周凯,王凯,朱宇航,等.基于GMM的容器定制化调度策略[J/OL].计算机科学:1-13[2025-01-21].<http://kns.cnki.net/kcms/detail/50.1075.tp.20241225.1943.018.html>.
- [10] 李艳红,金晨,师浩杰,等.基于卡尔曼滤波算法的无人机定位控制[J].内江科技,2024,45(10):40-42.
- [11] 申富媛,李炜,蒋栋年,等.基于两阶段卡尔曼滤波的四旋翼无人机自主预测维护[J].控制理论与应用,2023,40(12):2265-2276.
- [12] 王肖锋,贾星宇,张晓敏,等.基于距离度量的鲁棒主成分分析低维表征算法研究(上)[J/OL].光子·激光:1-17[2025-01-21].<http://kns.cnki.net/kcms/detail/12.1182.o4.20250107.1658.004.html>.
- [13] 邓雨阳,李继超,刘景源,等.基于深度学习神经网络的轴流压气机失速预测[J/OL].推进技术:1-13[2025-01-21].<http://kns.cnki.net/kcms/detail/11.1813.V.20250106.1759.001.html>.
- [14] TAO C, ZHANG C, JI H, et al. Fatigue damage characterization for composite laminates using deep learning and laser ultrasonic[J]. Composites part B: engineering, 2021, 216: 108816.
- [15] LUO D, WANG T, HAN J, et al. Study of three-dimensional distribution of chloride in coral aggregate concrete: a CNN-BiGRU-attention data-intelligence model driven by beluga whale optimization

algorithm[J]. Construction and building materials, 2025, 458: 139740.

- [16] 康晖明, 吴伟. 基于数据驱动的直升机飞行控制律设计[J/OL]. 飞行力学: 1-7 [2025-01-21].
<https://doi.org/10.13645/j.cnki.f.d.20241127.001>.

Helicopter vibration data polar coordinate fusion based on statistical methods

ZHU Tongbin, ZHANG Liuye

(East China Sea First Rescue Flight Team, Ministry of Transport of the People's Republic of China,
Shanghai 200137, China)

Abstract: In order to improve the accuracy and engineering practicality of helicopter vibration data analysis, the application and effectiveness of different statistical methods in polar coordinate point fusion of vibration data was explored. Taking helicopter main rotor vibration data (including vibration amplitude and phase angle) as the research object, four methods including Gaussian mixture model (GMM), principal component analysis (PCA), Bayesian inference, and Kalman filtering were used for data point fusion analysis. The results indicate that the four approaches are suitable for analyzing helicopter vibration data in the order of GMM>PCA>Bayesian inference>Kalman filtering. The GMM exhibits superior fitting ability when dealing with complex distributed data and is suitable for analyzing vibration data with multimodal distribution characteristics; PCA is highly efficient in dimensionality reduction and feature extraction, and is fit for scenarios with large data volumes and clear structures; Bayesian Inference is prone to model uncertainty and thus suitable for scenarios that require probabilistic prediction. However, the Kalman filter performed poorly in the experiment owing to the lack of temporal continuity in the data.

Keywords: helicopter vibration; polar coordinate point fusion; Gaussian mixture model; principal component analysis; Bayesian inference; Kalman filtering

(责任编辑:王新亮)

- 引用格式 朱桐斌, 张流业. 基于统计方法的直升机振动数据极坐标融合研究[J]. 山东航空学院学报, 2025, 42(4): 1-10.
ZHU T B, ZHANG L Y. Helicopter vibration data polar coordinate fusion based on statistical methods[J].
Journal of Shandong University of Aeronautics, 2025, 42(4): 1-10.