

DOI:10.19651/j.cnki.emt.2107706

基于高斯混合模型的平流层浮空器 RCS 分布拟合

李坤坤 曹锐 杨耀东 徐润田
(西北核技术研究所 西安 710024)

摘要: 雷达散射截面(RCS)是表征目标电磁散射特性的重要物理参数。针对典型起伏模型对平流层浮空器动态 RCS 分布特性描述精度较低的问题,采用高斯混合模型对浮空器动态 RCS 测量数据的起伏分布进行拟合并完成检验。首先,对浮空器体坐标系中的雷达视线角进行了解算;其次,采用多个高斯分布模型的混合叠加逼近浮空器 RCS 的概率密度分布,并引入期望最大化算法对各高斯分布分量的参数进行估计;最后,选取具有代表性的方位角度域内的浮空器 RCS 实测数据,分析了高斯混合模型对 RCS 概率分布的拟合效果并与典型起伏模型对比及检验拟合优度。数据分析结果表明,在最小二乘准则下高斯混合模型相比于典型起伏模型对 RCS 概率分布的拟合效果最高可提升 96.87%,验证了高斯混合模型的有效性。

关键词: 浮空器; RCS 分布; 雷达视线角; 高斯混合模型; 期望最大化; 拟合检验

中图分类号: P412.25 **文献标识码:** A **国家标准学科分类代码:** 510.70

Gaussian mixture model-based RCS distribution fitting for stratosphere aerostats

Li Kunkun Cao Rui Yang Yaodong Xu Runtian
(Northwest Institute of Nuclear Technology, Xi'an 710024, China)

Abstract: The radar cross section (RCS) is an important physical parameter to characterize the scattering characteristics of targets. This work aims at the problem that the typical fluctuation models have low precision in describing the RCS distribution characteristics of stratospheric aerostats, and the Gaussian mixture model is applied to fit the dynamic RCS distribution for aerostats. Firstly, the line-of-sight (LOS) angle for radar in aerostat's body coordinates is formulated. Secondly, a superposition of several Gaussian models is used to characterize the probability density distribution of aerostat's RCS, and the model parameters are approximated via the expectation maximization algorithm. Finally, the fitting effect of Gaussian mixture model on RCS distribution is studied and tested with some aerostat's RCS measurements and compared to typical fluctuation models. The results show that the Gaussian mixture model can improve the fitting effect of RCS probability distribution by up to 96.87% compared with other typical models under the least square criterion, thus demonstrating the effectiveness of Gaussian mixture model.

Keywords: aerostat; RCS distribution; radar LOS angle; Gaussian mixture model; expectation maximization; fitting and testing

0 引言

平流层浮空器是一种携带任务载荷并借助浮升气体升空的特种航空器,其通常在临近空间的平流层区域内驻空飞行,拥有续航时间长、覆盖区域广、携载运力大、使用效率比高等优点,可为预警探测、环境监测、通信中继等空地一体化任务需求提供解决方案,因而具有广泛的应用前景^[1]。

以往文献对于平流层浮空器的研究多集中在总体设计^[2]、动力学特性分析^[3-4]、航迹优化^[5-6]等方面。张礼学

等^[2]以平流层飞艇最大有效载荷质量为优化目标,在综合考虑能源平衡、材料强度等多方面约束条件下建立了飞艇的设计参数优化模型,相比于经验设计参数可增加飞艇的有效载荷质量。冯凯等^[3]设计了一种带辅助安定面的平流层飞艇气动布局,可以在不显著增加飞艇结构重量的前提下提高其航向静稳定性,为未来飞艇气动布局设计提供了理论参考。Lee 等^[5]提出了一种匹配轨迹优化方法来平衡 VIA-200 平流层飞艇的最优轨迹点质量模型和非线性跟踪六自由度模型之间的非匹配关系。尽管目前有关平流层浮

收稿日期:2021-08-26

空器的研究林林总总,但鲜有涉及其电磁散射特征参数的分析。平流层浮空器作为浮升体通常体积较为庞大,加上其囊体蒙皮材料的特殊性,使得其雷达散射截面(radar cross section, RCS)随浮空器飞行航迹及运动姿态变化呈现时变的起伏特性,而 RCS 作为电磁散射特征的重要组成部分,分析其起伏分布特性对雷达目标检测识别至关重要。相比于传统 RCS 仿真计算^[7]和静态测量^[8],动态测量 RCS 的分布特性由于综合了姿态扰动、背景杂波等各种不确定因素的影响,更能真实反映目标运动过程中的实际散射特征,现已作为目标动态特性分析的基础得到广泛研究。

一般地,分析人员选择典型统计分布模型表征雷达目标 RCS 的起伏分布,如卡方分布、赖斯分布等,但由于模型参数数目及曲线特征的限制,传统单一模型难以真实地描述目标在实际复杂环境中的动态 RCS 分布特性。一些学者采用有限混合模型^[9-11]进行不同类型参量分布特性的统计建模,在这些有限混合模型中,高斯混合模型由于其表达形式的简洁性及实用性得到了广泛地运用。理论上来说,只要高斯分布个数足够多,高斯混合模型能够逼近并表征具有任意特点的统计分布。

本文为提高平流层浮空器动态 RCS 分布特性的描述精度,在雷达视线角解算的基础上采用高斯混合模型对浮空器动态 RCS 测量数据的起伏分布进行拟合,并引入期望最大化算法得到各模型成分中未知参数的估计,预期可取得良好的 RCS 分布拟合效果,对于平流层浮空器电磁散射特性的评估和雷达探测识别能力的提升具有重要的意义。

1 雷达视线角解算

雷达在对浮空器这类电大尺寸目标探测时一般产生多个强散射点,而各散射分布矢量在空间波数域合成时是对探测频率和姿态角敏感的,因而若探测频率一定,浮空器 RCS 将随不同探测视线角呈现出时变的起伏特性。下面基于浮空器运动姿态和坐标转换关系,对浮空器体坐标系中的雷达视线角进行解算。

如图 1 所示,定义雷达坐标系为:以天线收发阵面中心作为坐标原点 O_r , $O_r Z_r$ 铅垂向上, $O_r X_r$ 和 $O_r Y_r$ 位于水平面内,分别指向正东与正北。定义浮空器体坐标系为:以其质心作为坐标原点 O_b , $O_b X_b$ 沿体纵轴线指向头部, $O_b Z_b$ 位于体对称平面内且垂直于 $O_b X_b$ 指向上方, $O_b Y_b$ 垂直于体对称平面,由右手定则确定其指向。

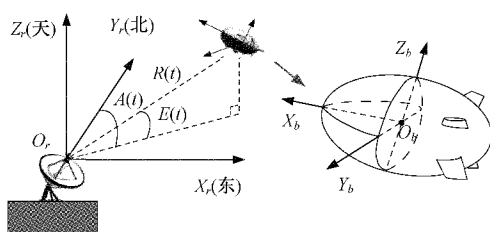


图 1 坐标系定义

采用俯仰角 $\varphi(t)$ 、偏航角 $\psi(t)$ 和滚转角 $\gamma(t)$ 这 3 个欧拉角来描述浮空器运动过程中的姿态变化,其中 $\varphi(t)$ 为 $O_b X_b$ 与雷达坐标系中 $O_r X_r Y_r$ 平面的夹角, $\psi(t)$ 为 $O_b X_b$ 在 $O_r X_r Y_r$ 平面上的投影线与 X_r 轴的夹角, $\gamma(t)$ 为 $O_b X_b Z_b$ 平面与包含 X_b 轴的铅垂面的夹角。

一般而言,雷达以径向距离 $R(t)$ 、方位角 $A(t)$ 和俯仰角 $E(t)$ 等球坐标形式给出浮空器位置,在雷达坐标系中它们可转化为等效位置观测,即:

$$\begin{cases} x_r(t) = R(t) \cos E(t) \sin A(t) \\ y_r(t) = R(t) \cos E(t) \cos A(t) \\ z_r(t) = R(t) \sin E(t) \end{cases} \quad (1)$$

浮空器飞行航迹及其运动姿态的变化都会造成雷达探测视线在浮空器体坐标系中不断变化。为便于分析,雷达探测视线需转换至浮空器体坐标系,即:

$$\begin{bmatrix} x_T(t) \\ y_T(t) \\ z_T(t) \end{bmatrix} = \mathbf{L} \cdot \begin{bmatrix} x - x_r(t) \\ y - y_r(t) \\ z - z_r(t) \end{bmatrix} \quad (2)$$

式中: (x, y, z) 为雷达在雷达坐标系中的坐标, $(x_T(t), y_T(t), z_T(t))$ 为雷达在浮空器体坐标系中的坐标, \mathbf{L} 为雷达坐标系至浮空器体坐标系的转换矩阵,矩阵各元素 l_{ij} 可表示为:

$$\begin{cases} l_{11} = \cos\psi(t) \cos\varphi(t) + \sin\psi(t) \sin\varphi(t) \sin\gamma(t) \\ l_{12} = \sin\psi(t) \cos\varphi(t) \\ l_{13} = -\cos\psi(t) \sin\varphi(t) + \sin\psi(t) \cos\varphi(t) \sin\gamma(t) \\ l_{21} = -\sin\psi(t) \cos\varphi(t) + \cos\psi(t) \sin\varphi(t) \sin\gamma(t) \\ l_{22} = \cos\psi(t) \cos\varphi(t) \\ l_{23} = \sin\psi(t) \sin\varphi(t) + \cos\psi(t) \cos\varphi(t) \sin\gamma(t) \\ l_{31} = \sin\varphi(t) \cos\gamma(t) \\ l_{32} = -\sin\gamma(t) \\ l_{33} = \cos\varphi(t) \cos\gamma(t) \end{cases} \quad (3)$$

对静止雷达而言, (x, y, z) 取 $(0, 0, 0)$, 则式(2)可写为:

$$\begin{bmatrix} x_T(t) \\ y_T(t) \\ z_T(t) \end{bmatrix} = -\mathbf{L} \cdot \begin{bmatrix} x_r(t) \\ y_r(t) \\ z_r(t) \end{bmatrix} \quad (4)$$

将此直角坐标转换为极坐标形式,即:

$$\begin{cases} r(t) = \sqrt{x_T^2(t) + y_T^2(t) + z_T^2(t)} \\ \beta(t) = \arctan\left(\frac{y_T(t)}{x_T(t)}\right) \\ \epsilon(t) = \arctan\left(\frac{z_T(t)}{\sqrt{x_T^2(t) + y_T^2(t)}}\right) \end{cases} \quad (5)$$

式中: $r(t)$ 为浮空器体坐标系中的径向距离, $\beta(t)$ 为同坐标系中雷达视线方位角, $\epsilon(t)$ 则为雷达视线高低角。

2 高斯混合模型与期望最大化算法

基于有限混合模型来建模某参量统计分布,关键在于选择模型分量概率函数和确定其有关参数。因此,采用高

斯混合模型与期望最大化算法对平流层浮空器 RCS 分布的拟合分为 2 步:首先,以高斯混合模型表征浮空器 RCS 分布并构造有关数学模型;其次,以 RCS 样本数据为基础,根据期望最大化算法估计各高斯分布分量参数。

2.1 利用 GMM 建模 RCS 概率分布

高斯混合模型^[10](Gaussian mixture model, GMM)采用有限个高斯分布的加权和来表征或量化某一对象的统计分布。对于尚未确定的浮空器 RCS 概率密度函数 $p(\boldsymbol{\sigma})$, 其可表示为 J 个高斯分布的线性组合, 即:

$$p(\boldsymbol{\sigma}) = p(\boldsymbol{\sigma} | \boldsymbol{\theta}) = \sum_{j=1}^J \alpha_j p_j(\boldsymbol{\sigma} | \vartheta_j) \quad (6)$$

式中: $\boldsymbol{\sigma} = \{\sigma_i\}_{i=1}^M$ 为离散时刻雷达测量获取的浮空器一维 RCS 样本数据, 由于其具有较大的起伏随机性, 这里将其作为随机变量, M 为样本容量; $\boldsymbol{\theta} = \{\alpha_j, \vartheta_j\}_{j=1}^J$ 为待估参数, J 为高斯分布分量的个数; α_j 为第 j 个高斯分布分量的概率, 并满足 $\sum_{j=1}^J \alpha_j = 1$; $\vartheta_j = \{\mu_j, \delta_j\}$ 为第 j 个高斯分布分量的参数, 即均值 μ_j 和标准差 δ_j ; $p_j(\boldsymbol{\sigma} | \vartheta_j)$ 表示第 j 个服从高斯分布 $N(\mu_j, \delta_j^2)$ 的概率密度函数, 即:

$$p_j(\boldsymbol{\sigma} | \vartheta_j) = \frac{1}{\sqrt{2\pi}\delta_j} \exp\left[-\frac{(\boldsymbol{\sigma} - \mu_j)^2}{2\delta_j^2}\right] \quad (7)$$

由以上两个表达式可见高斯混合模型的未知量除待估参数 $\boldsymbol{\theta} = \{\alpha_j, \vartheta_j\}_{j=1}^J$ 外, 还包括模型个数 J 。理论上, J 愈大混合模型愈加逼近 $p(\boldsymbol{\sigma} | \boldsymbol{\theta})$, 但同时也可能造成待估参数过多而导致的估计失准和计算负担问题。一般来说, 3~4 个高斯分布模型便可以较好地拟合大部分复杂对象的概率密度分布。因此, 在综合考虑情况下本文选取模型数 $J = 3$ 来对浮空器 RCS 的概率密度分布进行逼近, 而对待估参数 $\boldsymbol{\theta}$ 使用期望最大化算法完成确定。

2.2 利用 EM 估计 GMM 参数

由上节利用 GMM 建立的浮空器 RCS 概率密度函数 $p(\boldsymbol{\sigma} | \boldsymbol{\theta})$ 模型, 可设想 RCS 观测数据集合 $\boldsymbol{\sigma} = \{\sigma_i\}_{i=1}^M$ 的产生过程为: 首先依概率 α_j 选择第 j 个高斯分布分量 $p_j(\boldsymbol{\sigma} | \vartheta_j)$, 然后依第 j 个高斯分布分量的概率分布 $p_j(\boldsymbol{\sigma} | \vartheta_j)$ 生成观测数据 σ_i , 此时, 观测数据 σ_i 是已知的, 而反映观测数据 σ_i 来自第 j 个高斯分布分量是未知的或不可观测的, 该事件可用隐变量 Γ_{ij} 表示, 其定义为:

$$\Gamma_{ij} \triangleq y_i \in \{j | j = 1, 2, \dots, J\}, i = 1, 2, \dots, M \quad (8)$$

令 $\mathbf{y} = \{y_i\}_{i=1}^M$ 表示不可观测数据集合, 而 $y_i = j$ ($y_i \neq j$) 表示 RCS 观测数据 σ_i 是(非)由第 j 个高斯分布分量产生, 则 Γ_{ij} 为 0~1 随机变量, 即:

$$\Gamma_{ij} = \begin{cases} 1, & y_i = j \\ 0, & y_i \neq j \end{cases}, i = 1, 2, \dots, M; j = 1, 2, \dots, J \quad (9)$$

观测数据集合 $\boldsymbol{\sigma}$ 与不可观测数据集合 \mathbf{y} 可组成完全数据集 $\mathbf{z} = \{\boldsymbol{\sigma}, \mathbf{y}\}$, 其对数似然函数可表示为:

$$\ln p(\mathbf{z} | \boldsymbol{\theta}) = \ln p(\boldsymbol{\sigma}, \mathbf{y} | \boldsymbol{\theta}) = \ln \prod_{i=1}^M p(\sigma_i, y_i | \boldsymbol{\theta}) = \sum_{i=1}^M \ln \alpha_{y_i} p_{y_i}(\sigma_i | \vartheta_{y_i}) \quad (10)$$

式(10)的求解可通过机器学习方法中经典的期望最大化(expectation maximization, EM)算法^[12]完成, 该算法是一种处理含有隐变量或缺失数据的概率模型参数估计问题的优化方法, 其通常以极大似然估计理论为基础给出未知参数的迭代收敛估计, 每次迭代由期望步(E步)和极大化步(M步)交替组成。根据 EM 算法, 未知参数 $\boldsymbol{\theta}$ 的估计步骤如下。

步骤 1): 给定待估参数 $\boldsymbol{\theta}$ 的初始值 $\hat{\boldsymbol{\theta}}_{k=0}$ (第 $k = 1$ 步迭代前)和收敛精度 ε , 然后进行迭代。

步骤 2): 期望步(E步)。求完全数据集 \mathbf{z} 的对数似然函数估计, 可通过求数学期望实现, 为:

$$Q(\boldsymbol{\theta} | \hat{\boldsymbol{\theta}}_k) = E_{\mathbf{y} | \boldsymbol{\sigma}, \hat{\boldsymbol{\theta}}_k} \ln p(\boldsymbol{\sigma}, \mathbf{y} | \boldsymbol{\theta}) \quad (11)$$

在给定已知条件 $\hat{\boldsymbol{\theta}}_k$ 情况下, 有:

$$p(y_i | \sigma_i, \hat{\boldsymbol{\theta}}_k) = p(y_i = j | \sigma_i, \hat{\boldsymbol{\theta}}_k) = \frac{p(\Gamma_{ij} = 1 | \sigma_i, \hat{\boldsymbol{\theta}}_k)}{\sum_{s=1}^M p(\Gamma_{is} = 1 | \sigma_i, \hat{\boldsymbol{\theta}}_k)} =$$

$$\frac{p(\sigma_i | \Gamma_{ij} = 1, \hat{\boldsymbol{\theta}}_k) p(\Gamma_{ij} = 1 | \hat{\boldsymbol{\theta}}_k)}{\sum_{s=1}^M p(\sigma_s | \Gamma_{sj} = 1, \hat{\boldsymbol{\theta}}_k) p(\Gamma_{sj} = 1 | \hat{\boldsymbol{\theta}}_k)} = \frac{\hat{\alpha}_{y_i, k} p_{y_i}(\sigma_i | \vartheta_{y_i, k})}{\sum_{s=1}^M \hat{\alpha}_{y_s, k} p_{y_s}(\sigma_s | \vartheta_{y_s, k})}, k \in N \quad (12)$$

式中: N 为迭代步数上限。同时:

$$p(\mathbf{y} | \boldsymbol{\sigma}, \hat{\boldsymbol{\theta}}_k) = \prod_{i=1}^M p(y_i | \sigma_i, \hat{\boldsymbol{\theta}}_k) \quad (13)$$

则式(11)可展开为:

$$Q(\boldsymbol{\theta} | \hat{\boldsymbol{\theta}}_k) = E_{\mathbf{y} | \boldsymbol{\sigma}, \hat{\boldsymbol{\theta}}_k} \ln p(\boldsymbol{\sigma}, \mathbf{y} | \boldsymbol{\theta}) =$$

$$\sum_{\mathbf{y} \in \Gamma} [\ln p(\boldsymbol{\sigma}, \mathbf{y} | \boldsymbol{\theta})] p(\mathbf{y} | \boldsymbol{\sigma}, \hat{\boldsymbol{\theta}}_k) = \sum_{j=1}^J \left[\sum_{i=1}^M \ln \alpha_j p_j(\sigma_i | \vartheta_j) \right] p(j | \sigma_i, \hat{\boldsymbol{\theta}}_k) = \sum_{j=1}^J \sum_{i=1}^M [\ln(\alpha_j)] p(j | \sigma_i, \hat{\boldsymbol{\theta}}_k) + \sum_{j=1}^J \sum_{i=1}^M [\ln(p_j(\sigma_i | \vartheta_j))] p(j | \sigma_i, \hat{\boldsymbol{\theta}}_k) \quad (14)$$

其中,

$$p(j | \sigma_i, \hat{\boldsymbol{\theta}}_k) = \frac{\hat{\alpha}_{j, k} p_j(\sigma_i | \vartheta_{j, k})}{\sum_{s=1}^M \hat{\alpha}_{s, k} p_s(\sigma_i | \vartheta_{s, k})}, j = 1, 2, \dots, J \quad (15)$$

$$p_j(\sigma_i | \vartheta_j) = \frac{1}{\sqrt{2\pi}\delta_j} \exp\left[-\frac{(\sigma_i - \mu_j)^2}{2\delta_j^2}\right] \quad (16)$$

$$p_j(\sigma_i | \vartheta_{j, k}) = \frac{1}{\sqrt{2\pi}\delta_{j, k}} \exp\left[-\frac{(\sigma_i - \mu_{j, k})^2}{2\delta_{j, k}^2}\right] \quad (17)$$

步骤 3): 极大化步(M步)。在得到 $Q(\boldsymbol{\theta} | \hat{\boldsymbol{\theta}}_k)$ 后, 极大化其期望值, 得到第 $k + 1$ 步迭代的参数估计 $\hat{\boldsymbol{\theta}}_{k+1}$ 为:

$$\hat{\theta}_{k+1} = \underset{\theta}{\operatorname{argmax}} Q(\theta | \hat{\theta}_k), \text{ s. t. } \sum_{j=1}^J \alpha_j = 1 \quad (18)$$

由一阶 KKT 条件 (Karush-Kuhn-Tucker condition), 可得到第 j 个高斯分布分量参数 $\vartheta_j = \{\mu_j, \delta_j\}$ 的估计为:

$$\hat{\mu}_{j,k+1} = \frac{\sum_{i=1}^M \sigma_i p(j | \sigma_i, \hat{\theta}_k)}{\sum_{i=1}^M p(j | \sigma_i, \hat{\theta}_k)} \quad (19)$$

$$\hat{\delta}_{j,k+1}^2 = \frac{\sum_{i=1}^M p(j | \sigma_i, \hat{\theta}_k) (\sigma_i - \hat{\mu}_{j,k+1})^2}{\sum_{i=1}^M p(j | \sigma_i, \hat{\theta}_k)} \quad (20)$$

然后, 利用 Lagrange 乘子法并结合约束条件 $\sum_{j=1}^J \alpha_j = 1$, 可求得第 j 个高斯分布分量概率 α_j 的估计 $\hat{\alpha}_{j,k+1}$ 为:

$$\hat{\alpha}_{j,k+1} = \frac{\sum_{i=1}^M p(j | \sigma_i, \hat{\theta}_k)}{M} \quad (21)$$

步骤 4): 将求得的 $\hat{\theta}_{k+1}$ 作为下步迭代初始值, 重复步骤 2) 和 3), 直至对数似然函数值 $Q(\theta | \hat{\theta}_{k \rightarrow N})$ 收敛, 即 $|Q(\theta | \hat{\theta}_{k+1}) - Q(\theta | \hat{\theta}_k)| < \epsilon$, 或达到迭代步数上限 N 。

至此, 利用 EM 算法完成 GMM 建模浮空器 RCS 概率分布时各高斯分布分量的参数估计。

3 实测数据分布拟合

采用某驻空平流层浮空器的一段雷达动态实测 RCS 数据展开分析, 如图 2 所示。雷达工作频率为单脉冲 6.5 GHz, 采用 VV 极化方式, 数据采样周期 T_{rad} 为 0.05 s, 浮空器惯导装置记录的运动姿态采样周期 T_{inc} 为 1 s、1.5 s 不等。为便于分析, 预处理过程中, 对 RCS 异常数据进行剔除, 并采用样条插值方法将惯导装置记录数据率统一至雷达采样数据率。

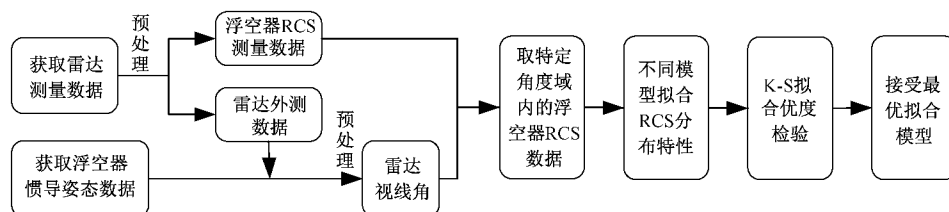


图 2 浮空器 RCS 分布拟合分析过程

驻空段平流层浮空器惯导装置记录的俯仰角 $\varphi(t)$ 、偏航角 $\psi(t)$ 和滚转角 $\gamma(t)$ 随相对时间 t 的变化曲线如图 3 所示。可以看出, 驻空阶段浮空器俯仰角、滚转角变化范围很小, 而偏航角由于平流层水平风及飞行航向控制缘故变化范围较大。对应地, 浮空器体坐标系中的雷达视线方位角 $\beta(t)$ 变化范围较大, 而视线高低角 $\epsilon(t)$ 变化范围很小, 如图 4 所示, 说明该阶段雷达视线方位角的变化占据

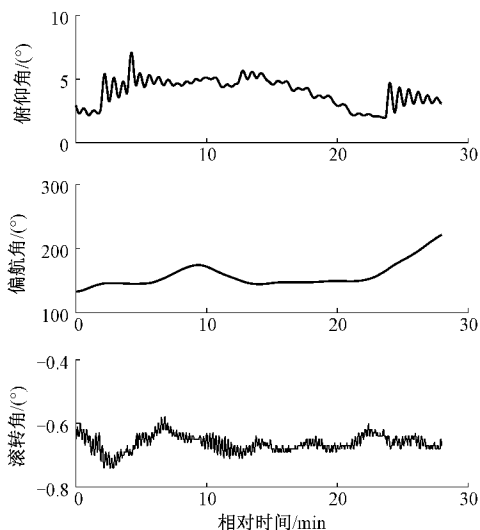


图 3 浮空器惯导姿态角变化

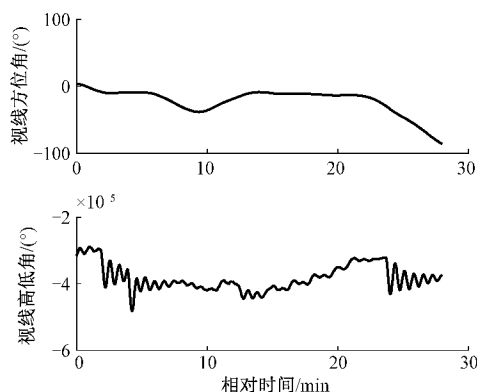


图 4 浮空器体坐标系中的雷达视线角变化

主导地位, 并主要由浮空器的偏航姿态角变化引起。驻空段浮空器动态 RCS 与雷达视线方位角的变化如图 5 所示, 可以看到 RCS 起伏与体视线方位角变化存在较强的对应关系。选取相对时间 14~28 min 范围内的浮空器 RCS 数据片段, 如图 6 所示, 该段对应的雷达视线高低角很小, 而视线方位角变化范围较大, 由 -8.5° 变化至 -82° , 近似对应于雷达水平入射波从目标头部开始探测并逐步变化至其侧面。由于浮空器囊体结构的旋转对称性, 这种近似 $1/4$ 扫角范围的方位角度域测量可以满足平流层浮空器的动态 RCS 分布特性分析要求。

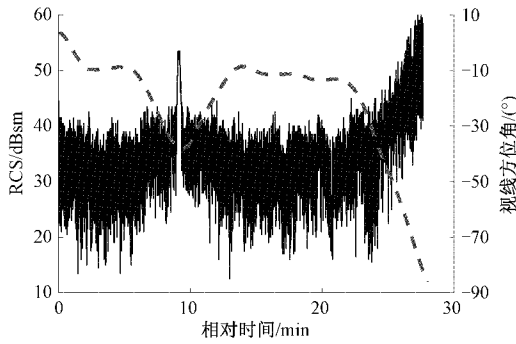


图 5 浮空器 RCS 与雷达视线方位角

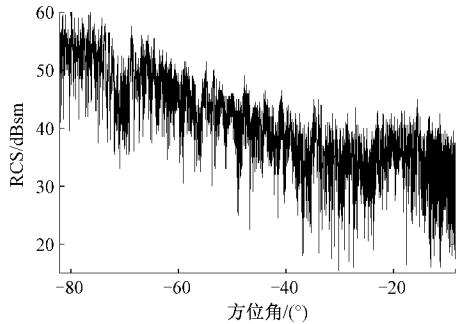


图 6 一定方位角度域内浮空器 RCS

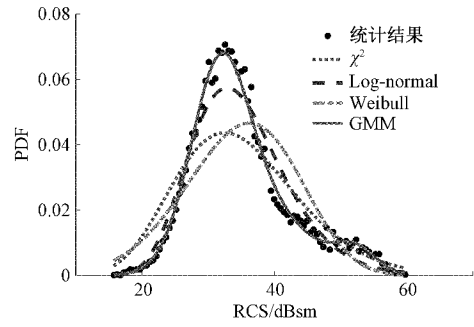


图 7 浮空器 RCS 分布 PDF 拟合结果

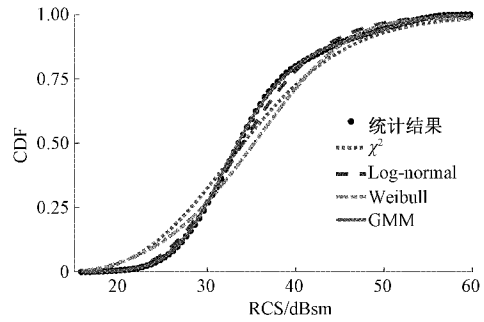


图 8 浮空器 RCS 分布 CDF 拟合结果

为了比较高斯混合模型的拟合效果,使用 χ^2 分布、Log-normal 分布及 Weibull 分布等 3 种典型 RCS 起伏分布模型^[13]同时拟合所选取 RCS 样本数据的分布。各模型的随机变量均为 RCS 值 σ , 因变量为 RCS 概率密度函数 PDF, 即 $\hat{P}(\sigma)$ 。其中, χ^2 分布模型参数为双自由度数值 k , 其可以不为整数; Log-normal 分布模型参数为 RCS 均值 σ 及标准差 σ_s ; Weibull 分布模型参数为尺度参数 λ 及形状参数 κ 。浮空器 RCS 概率密度函数 PDF、累积概率分布函数 CDF 的拟合结果分别如图 7、8 所示, 相比于几种典型单一起伏分布模型, 高斯混合模型实现了对实测 RCS 样本数据分布的良好拟合。事实上, 典型起伏分布模型属于一种参数化模型, 其拟合精度会受到有限个数的模型参数取值限制及各参数之间的耦合性影响, 无法较好地描述 RCS 分布的细节, 而高斯混合模型则集合了多个高斯分布函数, 各高斯分布分量可以最大程度地逼近 RCS 样本的一些细节特性, 因而拟合精度更高。

采用式(22)最小二乘误差公式^[13]衡量各模型的拟合误

差大小, 各拟合模型的参数及拟合误差对比结果如表 1 所示。表中, χ^2 分布模型的参数通过最小二乘准则确定; Log-normal 分布、Weibull 分布模型的参数则分别由 MATLAB 自带的 lognfit 函数和 wblfit 函数得到它们的最大似然估计; 高斯混合模型的参数通过 EM 算法估计得到, 但该算法对参数初值选取比较敏感, 其选取不佳会导致对数似然期望收敛至局部极值, 通常应对策略为采用多组不同参数初值开启迭代, 然后对各组最终得到的参数估值加以比较并从中择优^[14], 这种比较费时费力。考虑到 K 均值聚类算法对大型样本数据集的处理效率较高, 且可以给出样本参数的良好初始聚类结果^[15], 因此本文采用 MATLAB 自带的 K 均值聚类算法“k-means”函数处理浮空器 RCS 样本数据, 并获取各高斯分布分量的参数迭代初值。由各模型的最小二乘拟合误差可以计算得到, 高斯混合模型的拟合效果相比于 χ^2 分布、Log-normal 分布及 Weibull 分布等 3 种单一起伏分布模型, 分别提升 95.83%、85.71%、96.87%, 显示了高斯混合模型的优越性。

表 1 RCS 分布拟合参数及拟合误差对比

项目	χ^2 分布	Log-normal 分布	Weibull 分布	GMM
模型参数	$k = 13.81$	$\sigma = 34.26, \sigma_s = 1.23$	$\lambda = 38.10, \kappa = 4.71$	$\alpha_1 = 0.58, \mu_1 = 31.69, \delta_1 = 4.32$ $\alpha_2 = 0.36, \mu_2 = 37.77, \delta_2 = 7.27$ $\alpha_3 = 0.06, \mu_3 = 52.21, \delta_3 = 3.05$
拟合误差	0.002 4	0.000 7	0.003 2	0.000 1

$$E = \sum_{i=1}^K [P(\sigma_i) - \hat{P}(\sigma_i)]^2 \quad (22)$$

式中: $P(\sigma_i)$ 为 RCS 样本数据在第 i 个区间的 PDF 统计值; K 为统计区间数,一般取 $K = 1.87(M-1)^{0.4}$, M 为样本容量; $\hat{P}(\sigma_i)$ 为拟合模型在对应区间的 PDF 值; E 为拟合误差,该值越小代表拟合模型效果越好。

为进一步衡量各模型拟合效果,采用 K-S 拟合优度检验方法^[16] 分别开展检验,有关公式为:

$$D = \max_{1 \leq i \leq K} \{ |F(\sigma_i) - \hat{F}(\sigma_i)| \} \quad (23)$$

式中: $F(\sigma_i)$ 为 RCS 样本数据在 i 个区间的 CDF 统计值; $\hat{F}(\sigma_i)$ 为拟合模型在对应区间的 CDF 值; D 值为 K-S 检验参数,该值越小代表拟合模型越容易被接受。

在给定显著性水平 $\alpha = 0.05$ 下,由 K-S 检验临界值表可获取拒绝临界值 $D_{crit} = 0.0106$,拟合优度检验结果如表 2 所示。其中,高斯混合模型的检验参数最小且低于 D_{crit} ,其余模型的检验参数均高于 D_{crit} ,故接受高斯混合模型。综合拟合误差对比及检验结果可知,高斯混合模型相比于单一起伏分布模型可以更好地拟合驻空段平流层浮空器的 RCS 起伏分布。

表 2 K-S 拟合优度检验结果对比(显著性水平 $\alpha = 0.05$, 拒绝值 $D_{crit} = 0.0106$)

项目	χ^2 分布	Log-normal 分布	Weibull 分布	GMM
K-S 参数	0.0269	0.0132	0.0282	0.0091
检验结果	拒绝	拒绝	拒绝	接受

4 结 论

本文采用高斯混合模型实现了对平流层浮空器动态 RCS 分布的良好拟合,并使用期望最大化算法得到了模型未知参数的估计。与传统单一起伏分布模型相比较,高斯混合模型无论是拟合效果还是检验结果上均有很大地提升,显示了高斯混合模型应用于 RCS 概率密度分布拟合的有效性。与此同时,高斯混合模型也存在参数迭代收敛过程随样本容量增加而耗时较大的不足,且本文关于模型成分数目的选取带有一定的人为经验因素,这些问题都是下一步工作需要优化解决的方向。

参考文献

[1] 俞朝阳,李源源.浮空器学科发展研究[C].2014—2015

航空科学技术学科发展报告,2016:27-51.

- [2] 张礼学,王中伟,杨希祥.面向有效载荷质量的平流层飞艇设计参数优化[J].中南大学学报(自然科学版),2013,44(4):1367-1372.
- [3] 冯凯,赵达,孟小君,等.带辅助安定面平流层飞艇的航向静稳定性分析[J].宇航学报,2018,39(7):715-723.
- [4] 张泰华,姜鲁华,周江华.放飞过程中平流层飞艇运动与受力分析[J].北京航空航天大学学报,2018,44(4):691-699.
- [5] LEE S, JANG J, RYU H, et al. Matching trajectory optimization and nonlinear tracking control for HALE[J]. Advances in Space Research, 2014, 54(9): 1870-1887.
- [6] 邓小龙,丛伟轩,李魁,等.风场综合利用的新型平流层浮空器轨迹设计[J].宇航学报,2019,40(7):748-757.
- [7] 赵博,田宇.基于仿真技术的箔条云团 RCS 建模方法研究[J].国外电子测量技术,2015,34(10):49-53.
- [8] 贾宏进,韩红斌,李亨.电大复杂目标 RCS 缩比模型验证[J].电子测量技术,2015,38(5):33-36.
- [9] 李保珠,董云龙,李秀友,等.基于 t 分布混合模型的抗差关联算法[J].电子与信息学报,2017,39(7):1774-1778.
- [10] 谭博,郭静波,常广,等.基于正弦高斯混合模型的磁目标通用快速检测[J].仪器仪表学报,2019,40(2):1-10.
- [11] 李应求,宋玉华,李德如.基于混合对数正态分布的桥梁车辆荷载模型研究[J].公路与汽运,2018,10(1):145-148.
- [12] 廉文娟,史丹丹,安其立,等.基于最大期望的初始聚类中心选择算法[J].软件,2020,41(5):9-13.
- [13] 胡程,方琳琳,王锐,等.昆虫雷达散射截面特性分析[J].电子与信息学报,2020,42(1):140-153.
- [14] 李航.统计学习方法[M].北京:清华大学出版社,2019:181-183.
- [15] 胡叶,张爱军.基于 K-means 的室内行人导航步态识别方法[J].国外电子测量技术,2021,40(3):32-36.
- [16] 夏嘉斌,祝连庆,姚齐峰,等.基于 Kolmogorov-Smirnov 检验的远程拉曼光谱寻峰方法[J].仪器仪表学报,2018,39(3):141-147.

作者简介

李坤坤,硕士,助理工程师,主要研究方向为口标跟踪、雷达数据处理。

E-mail: lkk9419nudt@163.com