

基于高斯混合模型的逆变器故障诊断方法研究

张万星, 王 炜, 白立辉

(天津工业大学 电气工程与自动化学院, 天津 300382)

摘要: 逆变器广泛应用于工业生产中的诸多领域; 逆变器在工作过程中会出现元器件性能退化或损坏, 造成经济损失甚至人员伤亡, 为了提高逆变器工作的可靠性, 识别出逆变器出现故障时的故障类型, 提出了基于高斯混合模型的逆变器故障诊断方法; 以谐振型逆变器的为例, 分析了逆变器的四种典型故障, 提取出逆变器在不同故障下输出电压的时域特征波峰系数和频域特征小波能量熵; 使用时频特征数据训练高斯混合模型, 并结合 EM 算法计算输入数据属于各种故障类型的概率, 建立逆变器的故障诊断模型; 仿真实验通过 Simulink 建立了并联谐振型逆变电路的模型, 模拟出四种典型的故障状态, 并用这些数据训练故障诊断模型; 仿真结果验证了基于高斯混合模型的故障诊断方法的有效性和准确性; 该故障诊断方法具有较高准确率, 对于四种类型故障的总的识别率达到 93.2%, 可以应用于工业现场逆变器的故障诊断及其它领域。

关键词: 逆变器; 故障诊断; 高斯混合过程; 并联谐振

Research on Fault Diagnosis Method of Inverter Based on Gaussian Mixture Model

Zhang Wanxing, Wang Wei, Bai Lihui

(School of Electrical Engineering and Automation, Tianjin Polytechnic University, Tianjin 300382, China)

Abstract: Inverters are widely used in many fields of industrial production. The performance degradation or malfunction of components would occur during running, resulting in economic losses and even casualties. In order to improve the reliability of the inverter and identify the fault type when the inverter fails, a fault diagnosis method based on Gaussian mixture model is proposed. Taking the resonant inverter power supply as an example, several typical faults of the inverter are analyzed, and the time-frequency characteristics of the faults are extracted. Then the Gaussian mixture model is trained using fault characteristics, and EM algorithm is used to predict the output of the test data. The model of the parallel resonant inverter circuit was established in Simulink, and several typical power failure states were simulated. Simulation experiments verify the effectiveness and accuracy of the proposed fault diagnosis method. The fault diagnosis method is easy to implement and the predictive accuracy of the model is as high as 93.2%. It can be applied to the fault diagnosis of industrial inverters and other fields.

Keywords: inverter; fault diagnosis; Gaussian mixing model; parallel resonance

0 引言

逆变器是把直流电转变成交流电设备, 广泛应用于工业生产的诸多领域, 如新能源汽车, UPS 电源等。目前国内使用的逆变器主要为以功率半导体为主开关元件的谐振型逆变电路^[1]。在电路的工作过程中, 会出现元器件性能退化或损坏, 导致系统设备连续运行的性能不佳或发生重大故障。如果不能及时发现并定位故障, 有可能会造成经济损失甚至危及人身安全。因此逆变器的故障诊断具有十分重要的意义。

故障诊断按照工作任务可以分为故障存在检测和故障类型识别。在实际的工作过程中, 为了提高逆变器的可靠性和稳定性, 不仅要判断出是否存在故障, 还要及时定位

故障出现的位置, 识别故障类型。目前, 国内外对于逆变器故障诊断的研究方法大致可以分为三类^[2]: 基于数学模型的故障诊断; 该方法通过建立逆变器的数学模型来对电路的输出状态进行评估, 从而诊断出故障, 但是逆变电路结构复杂, 建立准确数学模型的难度比较大^[3]。例如刘攀等人提出的基于相电压积分的逆变器故障诊断, 建立了门极信号和相电压的积分模型^[4], 根据相电压的积分作为故障诊断的判据, 来区分正常状态和故障状态。基于信号处理的故障诊断。此类方法对逆变器的输出信号进行变换处理, 提取电压或电流信号的时频特征进行故障诊断。例如杨忠林提出的基于小波包分解的逆变器开路故障诊断方法, 将直流侧电流信号进行小波包分解, 提取分解系数的均方根值, 作为开路故障诊断的判据。基于机器学习的故障诊断方法。该方法利用故障的特征数据, 通过监督学习的方法, 训练出故障诊断模型, 进行故障检测和分类。如文献 [5] 提出的基于隐藏层神经网络的逆变器故障诊断方法。该方法是一种简化的神经网络方法, 训练起来比较容易, 但是对于训练数据要求较高。

收稿日期: 2019-07-02; 修回日期: 2019-08-29。

作者简介: 张万星(1990-), 男, 河北唐山人, 硕士研究生, 主要从事逆变电源, 逆变器故障诊断方向的研究。

王 炜(1963-), 男, 吉林德惠县人, 硕士生导师, 副教授, 主要从事电子设计应用、自动化、传动控制方向的研究。

本文提出了一种基于数据处理和机器学习相结合的逆变器故障诊断方法。首先进行数据采集,采集逆变电路中的两个开关器件和谐振电阻开路故障时逆变器输出的电压信号,然后使用信号处理的方法提取电压信号的时域特征和频域特征。利用时频特征训练高斯混合模型,确定高斯混合模型的参数。训练好的模型可以用来进行故障诊断。最后,使用 simulink 搭建了逆变电路的模型,进行仿真实验。仿真结果表明该模型的运行速度快且具有较高的诊断精度。本文提出的故障诊断模型,易于实现,可以应用于工业现场的逆变器故障诊断,也可以应用于其他领域的故障诊断。

1 逆变电源的故障分析

逆变电路的故障主要可以分为两类,突变型故障和渐变型故障。突变型故障是指逆变器的一些元器件短路或开路会导致系统瞬间无法运行;而渐变型故障发生时,系统的性能退化,如果长时间运行于这种状态下,会使系统部件或负载产生损坏。渐变型故障更加难以发现,因此,本文主要针对渐变型故障进行分析和诊断。

并联谐振型逆变电源的逆变回路电路原理图如图 1 所示。逆变回路部分包括一个直流源 DC,四个开关管 VT1—VT4,谐振电阻 R,谐振电感 L,谐振电容 C 等元器件。并联谐振型的逆变电路的故障类型有单开关管开路,双开关管开路,三开关管开路,四开关管开路以及谐振器件开路等。一般,双管、三管、四管开路的概率比较小^[6]。因此,本文选择单开关管开路故障,以及谐振器件故障作为研究对象进行故障分类识别。本文把正常状态作为一种特殊故障状态来处理,研究 4 种故障状态:正常状态,VT1 开路,VT2 开路和谐振电阻 R 开路。

如果每个故障处设一个故障检测点,检测电压或电流信号,能最直接准确地识别故障的类型和位置。但是这样会使故障诊断系统的复杂度大大增加,尤其是在逆变器结构复杂,故障类型较多的情况下。这种方法并不适合于实际生产。考虑到不同的故障类型会对电路的输出电压产生不同的影响。因此,测得逆变器的输出电压的特征就可以推测出故障的类型。测量输出电压操作简单,易于实现,适合于进行故障诊断。但是故障诊断之前,需要采集大量不同故障类型的数据,才能找到较为准确的故障类型和输出电压的关系。

2 故障特征提取

少中提取有用信息,去除冗余信息,从而使数据维度降低的过程^[7]。进行故障诊断需要采集逆变器运行过程中的数据,而电信号采集一般采样频率较高,如果直接用原始数据进行故障诊断,会占用大量存储和计算资源^[8-9]。因此,在故障诊断过程中,特征提取是十分必要的。数据的特征一般分为时域特征和频域特征^[10-11]。时域特征描述的是信号随时间变化的关系,时域特征的特点是计算简单,可以直观描述信号。常用的时域特征有平均值,均方根误

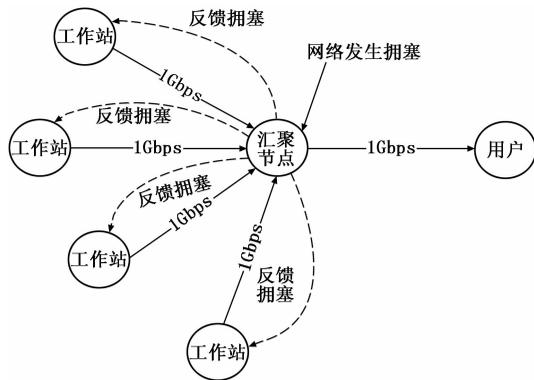


图1 逆变回路原理图

差,波形因子,裕度系数,波峰系数等;频域特征反映信号的动态特性,常用的提取频域特征的方法有傅里叶变换,小波变换,希尔伯特变换等。本研究分别提取故障数据的波峰系数和小波能量熵作为时域和频域特征,组成二维特征向量,作为故障诊断的依据。

波峰系数是信号峰值与均方根值的比值,是信号平稳度的一种度量。交流电是正弦波,其波峰系数约为 3。当系统发生渐变型故障,波形发生变化,其波峰系数也会相应改变。因此,本研究选择波峰系数作为故障的时域特征。小波变换是近年来发展起来的一种多尺度信号分析方法。该方法通过平移和缩放运算把信号分成低频和高频部分,从而有效提取信号的时频信息。熵是指体系的混乱程度,用来表示能量在空间中的分布均匀程度。小波能量熵是小波各个成分的能量在子空间分布无序程度的一种度量。本文使用小波能量熵提取故障数据的频域特征。

原始信号经过小波变换之后,可以获得不同尺度下的子信号。每个尺度下的信号都有一个小波系数。我们定义单一尺度下的小波系数的平方之和为该尺度下的小波能量。

$$E_j = \sum_{k=1}^n C_j(k)^2 (j = 1, 2, \dots, n)$$

其中: J 是小波变换分解的层数, $C_j(k)$ 是小波系数。

将小波能量标准化,则有:

$$p_j = \frac{E_j}{E_{\text{total}}}$$

其中: E_{total} 是所有小波能量的和。根据熵的定义,可以获得以下小波熵:

$$W_E = - \sum_j p_j \ln p_j$$

如果小波能量熵较大,说明原始信号在各个子频带上较分散;如果小波能量熵较小,说明原始信号集中分布在某些子频带,它能够作为频域特征来区分不同类型的数据。

3 用于诊断的高斯混合模型

经过数据采集和特征提取,最后需要一个分类器来进行故障分类识别。高斯混合模型(Gaussian mixture model, GMM)是一种非常基础的概率统计模型,由若干个基于高斯概率密度的子模型组合而成。高斯混合模型的优点是它的

结果是基于概率值的。我们可以计算经验置信区间，然后根据这些信息，在某个感兴趣的区域重新拟合，进行预测。使用高斯混合模型时，还可以根据需要指定不同的核函数。而且高斯混合模型可以通过增加子模型的数量来拟合较复杂的数据分布，灵活性好，能够更好地描述数据结构。

高斯混合模型的定义如下：

在数理统计领域中，一维高斯分布的概率密度函数如下所示：

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(x-u)^2}{2\sigma^2}\right)$$

其中： μ 是高斯分布的均值， σ^2 是高斯分布的方差。

对于多维变量 $X = (x_1, x_2, \dots, x_n)$ ，假设其服从联合高斯分布，则其概率密度函数为：

$$f(X) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(X-u)^T \Sigma^{-1}(X-u)\right]$$

其中： u 是多维变量 X 的均值， Σ 是多维变量 X 的协方差矩阵。

当 $n=2$ 时，二维变量可以拟合出二维坐标平面中椭圆形数据的分布。但是实际工程中的数据分布往往比较复杂，单一的高斯模型很难准确地拟合出数据的分布概率。这时，通过求解多个高斯模型，赋予每个模型一个权重，就形成了高斯混合模型。不同的权重决定了每个高斯函数在 GMM 中所起的作用。 K 个高斯模型组成的高斯混合模型的概率密度如下所示：

$$\rho(x) = \sum_{k=1}^K \rho(k) \rho(x | k) = \sum_{k=1}^K \pi_k N(x | u_k, \Sigma_k)$$

其中： $N(x | u_k, \Sigma_k)$ 是第 k 个高斯模型的概率密度函数。 π_k 是第 k 个模型被选择的概率。

如果高斯模型的数量足够多，通过调整每个高斯模型的均值方差和每个高斯模型的权重，这个高斯混合模型可以准确拟合任意的样本分布。

高斯混合模型的参数为 (μ, Σ, π) ，最大期望算法 (Expectation-Maximization algorithm, EM) 通常被用来求解高斯混合模型的参数。EM 算法一种通过迭代进行极大似然估计的优化算法。通常用于对包含隐变量或缺失数据的概率模型进行参数估计。EM 的标准计算框架由 E 步 (Expectation-step) 和 M 步 (Maximization step) 组成。算法具有收敛性，可以迭代求得最优值。EM 算法的步骤如下所示：

- 1) 设置每个模型的初始值 π_k, μ_k 和 Σ_k 。
- 2) E step。

根据当前的 π_k, μ_k 和 Σ_k 计算每个数据 x_i 由第 k 个模型生成的后验概率：

$$\gamma(i, k) = \frac{\pi_k N(x_i | \mu_k, \Sigma_k)}{\sum_{j=1}^K \pi_j N(x_i | \mu_j, \Sigma_j)}$$

- 3) M step。

根据 E step 中计算的 $\gamma(i, k)$ 再重新计算 π_k, μ_k 和 Σ_k ：

$$\mu_k^{new} = \frac{1}{N_k} \sum_{i=1}^N \gamma(i, k) x_i$$

$$\Sigma_k^{new} = \frac{1}{N_k} \sum_{i=1}^N \gamma(i, k) (x_i - \mu_k^{new})(x_i - \mu_k^{new})^T$$

$$\pi_k^{new} = \frac{N_k}{N}$$

其中： $N_k = \sum_{i=1}^N \gamma(i, k)$ 。

4) 计算对数似然函数。

$$\ln p(x | \mu, \Sigma) = \sum_{i=1}^N \ln \sum_{k=1}^K \pi_k N(x_i | \mu_k, \Sigma_k)$$

5) 重复 2) ~4) 步，直到对数似然函数收敛为止。

上述迭代方法一定会使对数似然函数收敛于局部最大值。也可以证明该局部最大值就是全局最大值。

在求解 GMM 模型的过程中，用到了全概率公式和朴素贝叶斯公式和极大似然算法。GMM 算法所得到的对数似然概率不仅包含了数据点的标签，也包含了数据点属于这一标签的概率。这一点对于故障的识别与判断具有重要的意义。

4 实验仿真

为了验证本文提出的基于高斯混合模型的并联谐振型逆变器的故障诊断方法，在 simulink 中搭建了并联谐振逆变器的仿真模型，进行故障诊断。故障诊断的主要步骤如下所示。

4.1 数据采集

不同的故障类型会产生不同的数据，这些数据都能从逆变器输出电压反映出来。因此，仿真实验采集的数据为输出电压。在 simulink 模型中模拟出正常状态，VT1 开路，VT2 开路和谐振电阻 R 开路这四种工作状态并记录输出电压。

4.2 数据预处理

数据预处理是对数据提前进行处理和修正。采集的原始信号可能会出现异常点或缺失点，在数据预处理时需要进行删除或补全，以避免对后续计算产生干扰。数据预处理还包括标准化、归一化、特征转换等。

4.3 特征提取

实验采集的电压信号数据量大，含有较多冗余信息，需要进行特征提取。本实验提取的特征是波峰系数和小波能量熵。它们组成二维特征向量，作为高斯混合模型的输入。

4.4 模型训练

模型训练是使用提取的特征数据来确定高斯混合模型参数 (μ, Σ, π) 的过程。本仿真实验使用二维时频特征训练高斯混合模型，混合模型的数量 $n = 7$ 。

4.5 模型测试

模型测试是使用从原始数据单独留出的样本，对训练好模型的故障识别能力进行评估的过程。模型测试可以得到模型的准确率，评价模型的好坏，作为模型调整的依据。如果故障诊断模型在训练时具有较高的准确率，而在测试

时准确率较低, 说明训练的模型可能存在过拟合现象, 需要对训练的模型进行调整, 需要加入正则化项或提前结束训练来减少过拟合现象。

仿真实验的仿真电路如图 2 所示。该电路是一个典型的并联谐振逆变电路, 这种电路基于并联谐振的原理, 把换流电容和负载电路并联, 广泛应用于中频逆变器上。逆变桥由四个臂组成, 每一个臂由一个晶闸管和限流电感串联而成。负载回路由电容 C 和感应线圈 L、谐振电阻 R 并联组成。模型中的参数为输入电压源为直流 250 VDC, 谐振电容 0.01 F, 谐振电感 0.001 H。针对以上分析的四种故障类型, 仿真得到四种电压波形如图 3 所示。仿真实验采集了 4 种故障类型的数据一共 3 000 组, 其中 2 000 组数据用来训练, 1 000 组数据用来测试。

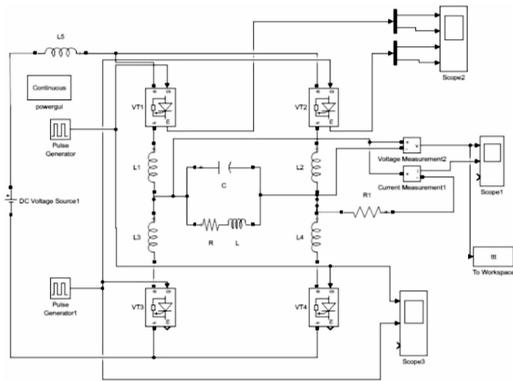


图 2 逆变器仿真原理图

对四种故障分别提取时域的波峰系数和频域的小波能量熵, 特征分布的部分结果如图 4 所示。从图中可以看出, VT2 故障较容易区分, VT1 故障、谐振电阻故障和正常状态的特征比较接近。使用 2 000 组特征数据训练高斯混合模型, 训练的参数为: 高斯模型 7 个, 协方差类型是 full, 迭代停止阈值是 $1e-3$ 。训练后得到特征数据的负对数分布如图 5 所示, 从图中可以看出, 越接近样本点, 概率值越大, 说明训练的模型准确有效。最后使用 1 000 组 4 种类型的数据输入模型进行测试。通过 EM 算法可以计算输入的测试数据属于每种故障类型的概率, 并把概率最大的那个故障类型判断为该条测试数据的故障类型。EM 算法预测的结果在表 1 中列出。该模型对于正常数据, 模型的识别率为 100%; 对 VT1 故障的识别率为 75.2%; 对 VT2 故障的识别率为 100%, 对谐振电阻 R 故障的识别率为 97.91%。总的故障识别率达到 93.2%。本文提出的逆变器故障诊断方法能较好地将故障识别。

表 1 四种故障诊断正确率

故障类型	测试数量	正确率	总正确率
正常	250	100%	93.2%
VT1 故障	250	75.2%	
VT2 故障	261	100%	
谐振电阻 R 故障	239	97.91%	

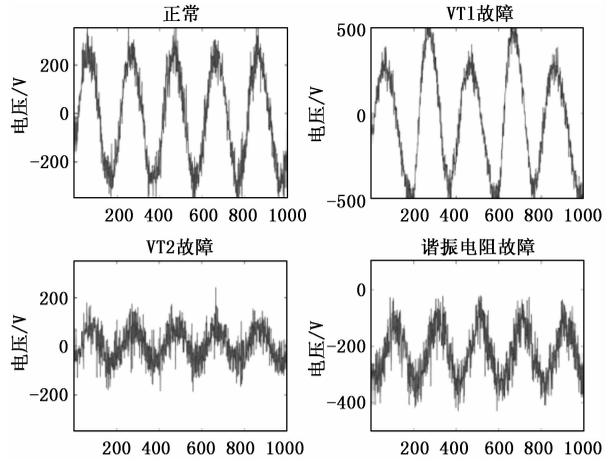


图 3 四种故障类型的输出电压数据

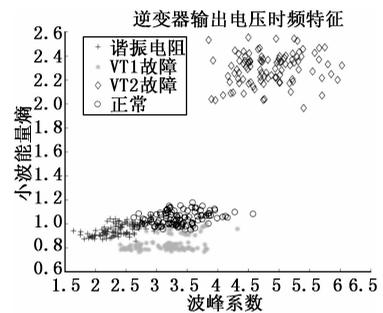


图 4 四种故障类型的特征分布

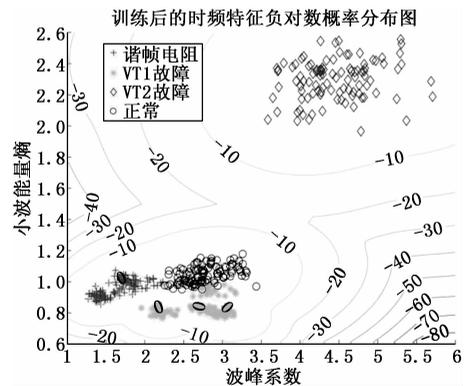


图 5 训练后的高斯混合模型特征概率分布

5 结论

针对并联谐振型逆变器, 本文提出了一种基于高斯混合模型的逆变器故障诊断方法。采集并分析逆变器输出电压数据, 提取电压数据的波峰系数和小波能量熵作为故障数据的特征。用特征数据训练高斯混合模型, 最后使用 EM 算法对输入数据进行识别分类, 确定故障的类型。为了验证所提出故障诊断方法的可行性和准确率, 使用 simulink 建立了谐振型逆变器的电路模型, 进行了仿真实验。仿真结果表明, 对于 4 种类型的电压数据, 该诊断方法的准确率可以达到 93.2%, 说明本文建立的模型具有可行性和较高的准确率, 能够满足实际应用需求。

本文提出的故障诊断方法，操作简单，容易实现，可以为其他类型的逆变器故障诊断提供参考依据，也可以推广于其他应用领域的故障诊断。

参考文献:

[1] 马晓辉, 毕雪芹, 张立广. 基于神经网络的谐振型逆变电源故障分析 [J]. 电子测量技术, 2018, 21 (41): 6.

[2] Yaghoubi M, et al. IGBT Open—Circuit Fault Diagnosis in a Quasi—Z—Source Inverter [J]. IEEE Transactions on Industrial Electronics, 2019, 66 (4): 2847—2856.

[3] Estima J O, Marques Cardoso A J. A New Approach for Real—Time Multiple Open—Circuit Fault Diagnosis in Voltage—Source Inverters [J]. IEEE Transactions on Industry Applications, 2011, 47 (6): 2487—2494.

[4] 刘 华, 王文江, 陈 轩. 基于相电压积分的逆变器开路故障诊断策略 [J]. 电子测量技术, 2017, 40 (10): 3.

[5] Huang Z, et al. Multi—Switches Fault Diagnosis Based on Small Low—Frequency Data for Voltage—Source Inverters of PMSM Drives [J]. IEEE Transactions on Power Electronics, 2019, 34 (7): 6845—6857.

[6] 赵丹阳, 董唯光, 高锋阳. 一种改进卷积神经网络的逆变器故障诊断 [J]. 电源学报, 2019, 1 (13).

[7] Uddin M P, Mamun M A, Hossain M A. Effective feature extraction through segmentation—based folded—PCA for hyperspectral image classification [J]. International Journal of Remote Sensing, 2019, 40 (18): 7190—7220.

[8] Talha M, Asghar F, Kim S H. A Novel Three—Phase Inverter Fault Diagnosis System Using Three—dimensional Feature Extraction and Neural Network [J]. Arabian Journal for Science and Engineering, 2018, 44 (3): 1809—1822.

[9] Froyland G, Rock C P, Sakellariou K. Sparse eigenbasis approximation: Multiple feature extraction across spatiotemporal scales with application to coherent set identification [J]. Communications in Nonlinear Science and Numerical Simulation, 2019, 77: 81—107.

[10] Ghasemzadeh H, et al. Detection of vocal disorders based on phase space parameters and Lyapunov spectrum [J]. Biomedical Signal Processing and Control, 2015, 22: 135—145.

[11] Anklshan H, Inam S C. A hybrid measure for the discrimination of the acoustic signals: Feature matrix (FMx) [J]. Applied Acoustics, 2019, 152: 88—100.

(上接第 4 页)

标测量系数的完整对比表单结构 (已知指向性目标测定系数与航天器隐藏目标点测量结果保持反比关系), 如图 6 所示。

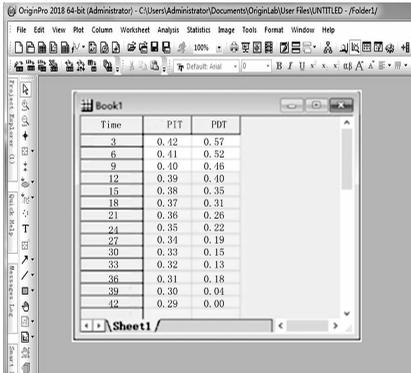


图 6 指向性目标测量系数变化图

分析图 6 可知, 随着实验时间的增加, PIT 系数保持均匀下降趋势, 每次的下降幅度均为 0.01, PDT 系数也保持持续下降的变化趋势, 但整体下降幅度明显大于 PIT 系数, 且最终实验数据结果达到 0。综上可知, 在 MEMS 激光雷达测量技术的作用下, 指向性目标测量系数出现明显的下降趋势, 可促进航天器隐藏目标点测量行为的顺利进行。

5 结语

MEMS 激光雷达测量是一种新型的航天轨迹追踪手段, 可通过整理上位机采集数据的方式, 描述相关性能指标, 并以此作为目标特征模型的建立条件, 对轨道内隐藏的特征值进行似然性估计。从实用性能力的角度来看, 应用这种方式建立的航天器隐藏目标点测量方法能够调节节点微动频率, 使其达到理想水平高度, 并抑制指向性目标系数

的不当攀升, 使激光雷达能够获取到更加多变的测量位置。整个方法搭建过程相对简单, 不需要对各项数值结果进行精准计算, 只需对其所属物理数值空间进行限定, 大大降低由数值取舍而造成的测量误差, 使提升航天器隐藏目标点测量准确性成为可能。

参考文献:

[1] 李光祚, 默 迪, 王 宁, 等. 一种新的高重频宽带相干激光雷达系统研究 [J]. 电子与信息学报, 2018, 40 (3): 525—531.

[2] 张亚亚, 徐明明, 刘 东, 等. 星载激光雷达滤波器组件热分析及参数优化 [J]. 应用光学, 2017, 38 (4): 660—664.

[3] 吕 婕, 吕永乐. 基于奇异值分解的雷达性能组合预测方法 [J]. 计算机测量与控制, 2019, 27 (1): 137—141.

[4] 李荣华, 温帅方, 肖余之, 等. 空间在轨激光成像雷达指标优选宏模型 [J]. 红外与激光工程, 2018, 3 (2): 155—162.

[5] 黄 艺, 贾英民. 非合作目标绕飞任务的航天器鲁棒姿轨耦合控制 [J]. 控制理论与应用, 2018, 35 (10): 1405—1414.

[6] 王兴龙, 周志成, 曲广吉. 目标捕获后航天器组合体的角动量转移与抑振规划 [J]. 宇航学报, 2018, 39 (3): 249—256.

[7] 黄允浒, 吐尔洪江·阿布都克力木, 刘芳园, 等. 基于特征值问题的正交小波构造方法 [J]. 电子设计工程, 2017, 25 (15): 1—5.

[8] 王 颖, 郁 梅, 应宏微, 等. 基于视差重映射的立体图像视觉舒适度提升 [J]. 中国图象图形学报, 2017, 22 (4): 452—462.

[9] 闫玲玲, 刘中印, 艾永乐, 等. 基于 0—1 整数规划算法的 PMU 量测点优化配置新方法 [J]. 电力系统保护与控制, 2017, 45 (12): 101—106.

[10] 陈升富, 常思江. 比例导引法中剩余飞行时间的计算方法 [J]. 弹道学报, 2017, 29 (3): 14—19.