

基于多源信息融合的飞行器部件剩余寿命预测

张秋雁¹, 杨忠^{2a}, 姜遇红^{2b}, 张启伦¹, 卢凯文^{2a}, 张辉斌^{2a}

(1. 贵州电网有限责任公司, 贵州 贵阳 550002;

2. 南京航空航天大学 a. 自动化学院, b. 无人机研究院, 江苏 南京 210016)

摘要:针对飞行器关键部件的多源变量数据统计信息,提出基于多源信息融合的相似性剩余寿命预测方法。介绍了相似性剩余寿命预测方法的基本思想和模型;提出一种使用BP神经网络融合多变量统计数据的方法;引入余弦相似度方法,将服役部件和参考部件退化模型进行模式匹配,确定与服役部件具有相同退化模式的参考部件,进而提高基于相似性剩余寿命预测方法的预测精度。通过NASA航空发动机数据集和相同评价指标下的对比分析,验证了该方法的有效性。

关键词:多源信息融合;相似性;神经网络;余弦相似度;剩余寿命预测

中图分类号:TP391.9 **文献标志码:**A **文章编号:**1671-5276(2020)01-0082-05

Residual Life Prediction of Aircraft Components Based on Multi-source Information Fusion

ZHANG Qiuyan¹, YANG Zhong^{2a}, JIANG Yuhong^{2b}, ZHANG Qilun¹, LU Kaiwen^{2a}, ZHANG Huibin^{2a}

(1. Guizhou Power Grid Co., Ltd., Guiyang 550002, China; 2. a.College of Automation Engineering

and b. Research Institute of UAV, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract: For the multi-source variable data statistical information of key components of aircraft, a similarity-based remaining life prediction method based on multi-source information fusion is proposed. The basic idea and model for the remaining life prediction are introduced, a method of using BP neural network to fuse multivariate statistical data is proposed. Then, the cosine similarity method is introduced to match the degradation models of service components and reference components, thus determining the reference components with the same degradation mode as the service components, so as to improve the prediction accuracy of similarity-based remaining life prediction method. The validity of this method is verified by the NASA aeroengine data set and comparison analysis under the same evaluation index.

Keywords: multi-source information fusion; similarity; neural network; cosine similarity; remaining life prediction

0 引言

飞行器内部关键部件(如发动机)的健康关系到飞行器系统的安全性、可靠性和稳定性。故障预测与健康管理系统(prognostic and health management, PHM)是一种基于视情维修的维修保障体系。部件剩余使用寿命(remaining useful life, RUL)预测技术研究是其关键技术之一^[1-2]。剩余使用寿命预测方法可以分为基于物理模型的方法和数据驱动的方法^[3-5]。

数据驱动的方法不需要构建精确的物理模型,依靠挖掘监测数据包含的特征及其变化趋势,显示部件或系统的性能变化情况,进一步达到预测部件或系统剩余使用寿命的目的。基于相似性的剩余寿命预测方法是其中的一个重要分支^[6]。WANG等^[7]在2008PHM数据挑战赛中采用基于相似性的剩余寿命预测方法取得不错的成绩。YOU等^[8]研究发现相似性剩余寿命预测方法具有较好的鲁棒性和长期预测能力。可见,基于相似性的剩余寿命预

测方法具有重要的研究价值和广阔的应用前景。

在工程实际中,部件的监测变量往往不只一个。同一部件在不同运行条件和环境影响下,呈现出不同的衰退趋势^[9]。在相似性剩余使用寿命预测方法研究中,雷从英等^[10]使用线性回归方法拟合多源监测数据,该方法需要提前选择具有线性趋势的监测变量且忽略了不同监测变量量纲的影响;李琪等^[11]提出一种变工况条件下基于相似性的剩余寿命预测方法,同样采用了线性回归方法建立模型。

为了解决以上问题,本文首先阐述相似性剩余寿命预测算法,进一步使用BP神经网络、余弦相似度等方法来改进算法,最后以NASA涡扇发动机仿真数据集FD001为例验证了算法有效性。

1 基于相似性的剩余寿命预测方法

1.1 概述

基于相似性的剩余寿命预测方法的基本思想是,若两

基金项目:国家自然科学基金项目(61473144);中国南方电网有限责任公司科技项目(066600KK52170074);航空科学基金(重点实验室类)(20162852031);江苏高校优势学科建设工程资助项目

第一作者简介:张秋雁(1965—),男,贵州人,教授级高级工程师,工学学士,主要从事测控技术研究。

个样本状态具有相似性,则其剩余寿命之间也具相似性。基于相似性的剩余寿命预测方法不仅能达到早期预测的目的,还避免了部件或系统建模的负担^[6]。

相似性方法需要将待预测部件与已失效历史部件之间进行相似性度量,往往将前者称为服役部件而后者称为参考部件。据此,基于相似性的剩余寿命预测方法可以更准确地表述为:如果一个服役部件的健康状态量与参考部件在某段时间段内的健康状态量之间存在较强的相似性,则可以认为该服役部件与某个或某些参考部件的剩余使用寿命也存在相似关系。然后由参考部件实际剩余寿命加权处理后达到估计服役部件剩余寿命的目的。而且,服役部件和参考部件相似度越高,其在加权求剩余寿命过程中所占的权重越大。

1.2 预测流程

以 $x_m(p)$ 表示参考部件 $m(m=1,2,\dots,M)$ 在运行周期 p 时的健康状态量, M 为参考部件总数目; $x_0(q)$ 表示服役部件监测点 q 处的健康状态量。

1) 确定时间范围 ΔT

为了更加准确地估计服役部件在某个运行周期的 RUL , 往往选择包含该运行周期在内的健康状态序列, 记该序列为:

$$x_0(q, H) = [x_0(q-H), \dots, x_0(q-1), x_0(q)] \quad (1)$$

式中 H 为非负整数, 则 $\Delta T = H + 1$ 。

显然, 参考部件应当与服役部件选择数量一致的健康状态量, 记为:

$$x_m(p, H) = [x_m(p-H), \dots, x_m(p-1), x_m(p)]$$

针对参考部件的每一个运行周期 $p(p \geq \Delta T)$, 逐个计算时间范围 ΔT 内 $x_m(p, H)$ 与 $x_0(q, H)$ 间的相似性。实践中, ΔT 可以依据操作经验确定。

2) 度量相似性

一般可选取欧几里得距离函数作为相似性测度函数, 在实际应用过程中, 为了估计服役部件在监测点 q 处的 RUL , 应当使越靠近 q 处健康状态量的权重越高, 故采用近邻增强的方式, 则相似性测度函数为^[12]:

$$d_m(p, \alpha) = \sum_{v=0}^H \alpha^v [x_0(q-v) - x_m(p-v)]^2 \quad (2)$$

式中: $\Delta T \leq p \leq T_m$, T_m 是参考部件 m 的失效时间; $\alpha (0 \leq \alpha \leq 1)$ 控制不同远近采样点对相似性度量的贡献, 称其为贡献度因子。

3) 估计服役部件剩余寿命

每个参考部件与服役部件有个 $(T_m - \Delta T + 1)$ 相似性度量值, 从中选择相似性程度最高的一个作为该参考部件的最佳匹配值, 用 d_m^{best} 表示, 则有

$$d_m^{\text{best}} = \min d_m(p, \alpha) \quad (3)$$

记此时得到的参考部件运行周期为 p_{best} , 同时可以得到参考部件此时的实际剩余寿命:

$$ARL_m(p_{\text{best}}) = T_m - p_{\text{best}} \quad (4)$$

根据欧氏距离越大其相似程度低的原理, 故可以使用指数函数来表示每个参考部件在剩余使用寿命计算中所占的权重, 最终服役部件在监测点 q 的剩余寿命 $PRL_0(q)$ 计算公式可表示为:

$$\begin{cases} \omega_m = \exp(-d_m^{\text{best}}(q)) \\ PRL_0(q) = \frac{\sum_{m=1}^M \omega_m \cdot ARL_m(p_{\text{best}})}{\sum_{m=1}^M \omega_m} \end{cases} \quad (5)$$

2 基于多源信息融合的相似性剩余寿命预测方法

2.1 方法概述

多源信息融合下的相似性剩余寿命预测方法是先从多源统计数据中提取反映部件健康状态的一维状态量, 称为健康因子 (health index, HI), 比如前文中的 $x_m(p)$ 和 $x_0(q)$ 。

目前, 在基于相似性的剩余寿命预测方法中, 主要使用线性拟合的方法对数据进行处理。该方法需要从所有监测传感器中选择具有线性变化趋势的传感器, 故局限性很大。本文采用基于 PCA 和 BP 神经网络的方法对多源传感器数据进行信息融合, 该方法对传感器数据的要求相对宽松, 适用性更广。

2.2 基于 PCA 和 BP 神经网络的信息融合方法

主成分分析法 (principal component analysis, PCA) 是以参量间的相互关系为基础, 在信息损失很少的情况下将多个参量转换为少数的综合参量, 这些综合参量可以称为主成分^[13]。PCA 通过线性变换的方式将原始数据变换为一组各维线性无关的数据分量, 可以达到降维的目的。

本文首先使用 PCA 对元数数据做降维处理, 再利用 BP 神经网络的非线性拟合能力获取参考部件数据的 HI 值, 最后使用获得的 PCA 参数和训练好的 BP 神经网络模型获取服役部件的 HI 。

BP 神经网络模型是一种多层前馈网络, 由 MCCCELLAND 等人于 1996 年提出^[14], 其学习过程主要包含信息正向传播和误差反向传播两部分。BP 神经网络模型的拓扑结构由输入层、隐含层和输出层构成, 每一层网络又由不同个数的节点组成。其具体算法为:

1) 随机给输入层与隐含层间的权重矩阵 \mathbf{V} 、隐含层与输出层之间的权重矩阵 \mathbf{W} 、隐含层阈值 θ 和输出层阈值 σ 赋初值, 向输入层输入给定值 $x_1, \dots, x_k, \dots, x_{n_1}$ 。

2) 记输出层期望输出向量为 $\mathbf{Y}^{\text{real}} = [y_1^{\text{real}}, \dots, y_j^{\text{real}}, \dots, y_{n_3}^{\text{real}}]$, 则总误差为 $E = \frac{1}{2} \sum_{j=1}^{n_3} (y_j^{\text{real}} - y_j)^2$ 。

3) 通过信息正向传播, 可以得到网络模型输出值, 记 h_i 和 y_j 分别为隐含层和输出层节点的函数值, 则:

$$h_i = f_1 \left(\sum_{k=1}^{n_1} v_{ki} x_k + \theta_i \right) \quad (6)$$

$$y_j = f_2 \left(\sum_{i=1}^{n_2} w_{ij} h_i + \sigma_j \right) \quad (7)$$

其中 f_1 和 f_2 为激活函数。

4) 根据梯度下降法,输出层与隐含层之间权值 w_{ij} 和输出层阈值 σ_j 调整规则为:

$$w_{ij}^{\text{new}} = w_{ij} + \beta(y_j^{\text{real}} - y_j) f_2'(s_j) h_i \quad (8)$$

$$\sigma_j^{\text{new}} = \sigma_j + \beta(y_j^{\text{real}} - y_j) f_2'(s_j) \quad (9)$$

其中 $s_j = \sum_{i=1}^{n_2} w_{ij} h_i + \sigma_j$

5) 进一步,输入层和隐含层之间权值 v_{ki} 和隐含层阈值 θ_i 调整规则为:

$$v_{ki}^{\text{new}} = v_{ki} + \beta x_k \delta_{ki} \quad (10)$$

$$\theta_i^{\text{new}} = \theta_i + \beta \delta_{ki} \quad (11)$$

其中 $\delta_{ki} = f_1' \left(\sum_{k=1}^{n_1} v_{ki} x_k + \theta_i \right) \sum_{j=1}^{n_3} w_{ij}^{\text{new}} (y_j^{\text{real}} - y_j) f_2'(s_j)$ 。

6) 按照上面的学习步骤,转到步骤 2) 反复学习,直到满足终止条件为止。终止条件常设置为,神经网络模型的输出减少至可接受的范围或者达到了预先设置的学习次数。

2.3 基于多源信息融合的相似性预测步骤

本文以 NASA 涡扇发动机仿真数据集 FD001 为例^[15],具体介绍基于多源信息融合的相似性剩余寿命预测方法。数据集包含已知全寿命周期的 100 个样本,每个样本含 21 个传感器监测信息。

为了对同一发动机进行连续预测,选择前 99 组发动机单元作为参考部件,最后一组作为服役部件。结合泛化相似性剩余寿命预测模型,具体步骤为:

1) 数据预处理。使用 PCA 方法对参考部件的 21 个传感器数据进行降维处理,根据累积方差贡献率^[16]大小,选取前 6 个主元作为神经网络模型的输入。

2) 多源信息融合。设计 BP 神经网络,输入层、隐含层和输出层节点个数分别设置为 6、12 和 1,采用自适应学习率梯度下降法进行训练。

3) 提取退化轨迹。上一步得到的健康因子衰退曲线波动剧烈,难以反映部件的退化轨迹,采用核回归平滑方法进一步提取各参考部件退化模型(图 1)。从图 1 中可以看出,退化模型具有明显的退化趋势,有利于相似性度量,为下一步服役部件的估计奠定了基础。

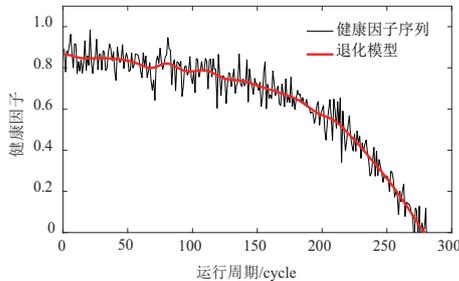


图 1 某参考部件健康因子衰退曲线与退化模型对比

4) 服役部件预处理。根据步骤 1) 和步骤 2) 获取 BP 神经网络模型计算服役部件的健康因子并提取退化模型。

5) 设置表 1 所示的预测参数。从预测起始点开始,每 10 个运行周期对服役部件进行一次预测,直到设定的预测终点,共得到 15 个预测结果。

表 1 预测参数

参数	含义	值
t_p/cycle	预测起始点	50
t_{Eop}/cycle	预测终点	190
ΔT	时间范围	10
α	贡献度因子	0.5

6) 估计服役部件剩余寿命。由式(2)依次度量服役部件与参考部件之间的相似性,统计所有参考部件的权重和剩余寿命,再由式(7)估计预测点 q 时服役部件剩余使用寿命。

2.4 基于余弦相似度的退化模式匹配方法

考虑到发动机受运行环境等影响,不同发动机的退化模型具有不同的退化模式,图 2 显示了服役部件在某个预测点的时间序列和不同参考部件的退化轨迹。显然,参考部件具有不同的性能衰退模式,如果能够在所有的参考部件中找到与服役部件最相近的退化曲线,那么根据相似性预测方法估计的服役部件剩余寿命更加精确^[17]。

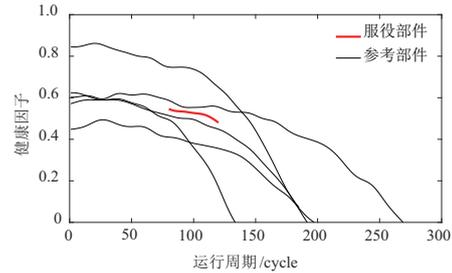


图 2 服役部件与参考部件退化模型匹配

在相关性计算方法中,余弦相似度是将个体的指标数据映射到向量空间,用向量空间中两个向量夹角的余弦值作为衡量两个个体的差异^[18]。余弦相似度的值能反映两个向量方向或走势之间的相似性,其能反映健康因子随时间变化的趋向。服役部件与参考部件退化模式之间的匹配程度可定义为:

$$\cos_m(p) = \frac{\sum_{v=0}^q x_0(q-v) \cdot x_m(p-v)}{\sqrt{\sum_{v=0}^q x_0(q-v)^2} \cdot \sqrt{\sum_{v=0}^q x_m(p-v)^2}} \quad (12)$$

从式(12)中可以看出,将服役部件所有运行周期对应的健康因子用于与参考部件退化模型的模式匹配,且随着服役部件运行周期 q 的增加,运行周期总数目也增加,与参考部件退化模型之间的匹配也更贴合。相应地,预测结果也更加准确。

进一步,在服役部件 q 不变情况下,改变 p 取值,计算参考部件的所有余弦相似度值,得到该参考部件下最佳的余弦相似度值:

$$\cos_m^{\text{best}} = \max(\cos_m(p)) \quad (13)$$

以此类推,计算服役部件与所有参考部件进行余弦相似度,将 \cos_m^{best} 从大到小排序,选择其中前 M_{new} 个参考部

件来构建模型库。

3 算例分析

3.1 评价指标

为了评价不同预测方法的预测结果,首先给出本文采用的6种评价指标,各个指标的简要含义如表2所示,其值越小越好。

表2 评价指标

名称	含义
SCORE	PHM2008 竞赛得分
FNR/(%)	滞后预测率
FPR/(%)	超前预测率
MAPE/(%)	平均绝对百分比误差
MAE	平均绝对误差
MSE	均方误差

其中,MAPE、MAE和MSE是从算法角度出发,而SCORE、FNR和FPR是从PHM领域预测效果角度定义^[8,19]。

3.2 实验一

首先对比两种信息融合方法对相似性剩余寿命预测方法的影响,得到图3所示的预测结果。

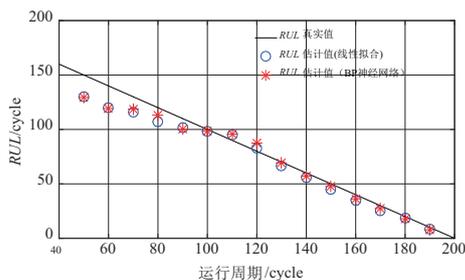


图3 不同信息融合方法预测结果

为了进一步比较两者的区别,得到表3所示的结果,其中,线性拟合预测实验中选择7个具有线性趋势的传感器编号,具体为2,3,4,7,11,12,15^[7]。

表3 不同信息融合方法预测结果

评价指标	线性拟合	BP 神经网络
SCORE	15.283 3	14.531 4
FNR/(%)	13.33	13.33
FPR/(%)	86.67	86.67
MAPE/(%)	9.313 0	8.818 3
MAE	7.338 6	6.726 9
MSE	91.913 6	84.508 6

从表3中可以看出,本文采用的基于BP神经网络信息融合方法的评价指标SCORE、MAPE、MAE和MSE均好于线性拟合方法。两种方法的FPR均远大于FNR。针对发动机这类关键部件,超前预测率高于滞后预测率表明预测结果满足需要,因为发动机一旦失效将造成严重故障,故相比于滞后预测,超前预测能够避免更大的损失。

3.3 实验二

为了验证基于退化模式一致性的泛化相似性预测方法,采用BP神经网络构建健康因子,在此基础上比较常规相似性剩余寿命预测方法和改进相似性剩余寿命预测方法在航空发动机数据集上的实验效果。将使用退化模式匹配的相似性剩余寿命预测方法称为方法1,基于退化模式匹配的相似性预测方法称为方法2,得到图4所示的对比结果。

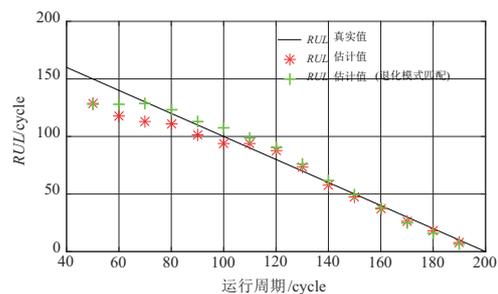


图4 不同方法预测结果

为了定性分析退化模式匹配对泛化相似性剩余寿命预测方法的影响,给出表4所示的评价指标对比结果。

表4 不同预测方法预测结果

评价指标	方法1	方法2
SCORE	14.531 4	13.476 1
FNR/(%)	13.33	46.67
FPR/(%)	86.67	53.33
MAPE/(%)	8.818 3	10.082 4
MAE	6.726 9	6.068 8
MSE	84.508 6	65.8576

从表4的评价指标SCORE、MSE和MAE可以看出,采用退化模式匹配后的相似性剩余寿命预测方法在整体上比常规相似性剩余寿命预测方法的预测结果更准确。但是,从图4中可以看出在预测早期方法2的滞后预测率高于方法1,但相比于真实剩余寿命,其误差在可接受范围内。MAPE中每个预测点的计算是绝对误差值对真实剩余寿命的百分比,图4中随着预测点越接近预测终点,方法2的预测误差略高于方法1,故表4中方法2的MAPE高于方法1,这也是今后需要进一步改进的地方。

4 结语

本文对多源统计数据下的相似性剩余使用寿命预测方法进行改进。为了充分利用多退化变量监测信息,建

立基于 PCA 和 BP 神经网络的智能模型进行信息融合,充分挖掘数据信息,通过发动机实例验证了所提方法的有效性。此外,为了提高相似性度量的可靠性,利用退化模式信息优化参考模型库的构建方法,进一步改进相似性度量的计算方法,提升了飞行器服役部件的整体预测效果。本文方法同样适用于机器人关键部件的剩余寿命预测。

参考文献:

- [1] SHEPPARD J W, KAUFMAN M A, WILMER T J. IEEE standards for prognostics and health management[J]. IEEE Aerospace & Electronic Systems Magazine, 2007, 24(9): 34-41.
- [2] LEE J, WU F, ZHAO W, et al. Prognostics and health management design for rotary machinery systems—reviews, methodology and applications[J]. Mechanical Systems & Signal Processing, 2014, 42(1/2): 314-334.
- [3] TSUI K L, CHEN N, ZHOU Q, et al. Prognostics and health management: A review on data driven approaches[J]. Mathematical Problems in Engineering, 2015(6): 1-17.
- [4] 彭宇, 刘大同. 数据驱动故障预测和健康管理综述[J]. 仪器仪表学报, 2014, 35(3): 481-495.
- [5] 朱晓栋, 陈则王. 基于维纳过程的电池剩余使用寿命预测[J]. 机械制造与自动化, 2018, 47(4): 197-200.
- [6] RAMASSO E, SAXENA A. Performance benchmarking and analysis of prognostic methods for CMAPSS Datasets[J]. International Journal of Prognostics & Health Management, 2014, 5(2): 1-15.
- [7] WANG T, YU J, SIEGEL D, et al. A similarity-based prognostics approach for remaining useful life estimation of engineered systems[C] // International Conference on Prognostics and Health Management, IEEE, 2008: 1-6.
- [8] YOU M Y, MEMG G. A Framework of similarity-based residual life prediction approaches using degradation histories with failure, preventive maintenance, and suspension events[J]. IEEE Transactions on Reliability, 2013, 62(1): 127-135.
- [9] ZHAO Z Q, LIANG B, WANG X Q, et al. Remaining useful life prediction of aircraft engine based on degradation pattern learning[J]. Reliability Engineering and System Safety, 2017, 164: 74-83.
- [10] 雷从英, 夏良华, 林智崧. 基于相似性的装备部件剩余寿命预测研究[J]. 火力与指挥控制, 2014, 39(4): 91-94.
- [11] 李琪, 高占宝, 李善营, 等. 变工况条件下基于相似性的剩余使用寿命预测方法[J]. 北京航空航天大学学报, 2016, 42(6): 1236-1243.
- [12] 尤明懿. 基于状态监测数据的产品寿命预测与预测维护规划方法研究[D]. 上海: 上海交通大学, 2012.
- [13] BISHOP C M. Neural networks for pattern recognition[M]. [S.I.]: Oxford University Press, 1995.
- [14] RUMELHART D E, MCCLELLAND J L, GROUP C P. Parallel distributed processing: explorations in the microstructure of cognition, vol. 2: psychological and biological models[J]. Language, 1986, 63(4): 45-76.
- [15] SAXENA A, GOEBEL K. Turbofan engine degradation simulation dataset [EB/OL] [2013-09-12]. <http://ti.arc.nasa.gov/project/prong-ostic-data-repository>.
- [16] CHINAKAY P, WONGSA S. A PCA-based fault monitoring of splitter nozzles in gas turbine combustion chamber using exhaust gas temperature[C] // International Conference on Instrumentation, Control and Automation, IEEE, 2017: 120-125.
- [17] ENRICO Z, FRANCESCO D M. A data-driven fuzzy approach for predicting the remaining useful life in dynamic failure scenarios of a nuclear system[J]. Reliability Engineering and System Safety, 2010, 95(1): 49-57.
- [18] DUBEY V K, SAXENA A K. A cosine-similarity mutual-information approach for feature selection on high dimensional datasets[J]. Journal of Information Technology Research, 2017, 10(1): 15-28.
- [19] SAXENA A, CELAYA J, SAHA B, et al. Evaluating algorithm performance metrics tailored for prognostics[C] // Aerospace Conference, IEEE, 2009: 1-13.

收稿日期: 2019-03-13

(上接第 46 页)

参考文献:

- [1] 罗少敏, 王亚平, 邹衍. 某型手枪击发机构装配公差分析与信息化集成研究[J]. 机械设计与制造, 2014(12): 132-136.
- [2] SALOMONS O W. Computer support in the design of mechanical Products, constraint specification and satisfaction in feature based design for manufacturing [D]. The Netherlands: University of Twente, 1995.
- [3] TAGUCHI G. Performance Analysis Design[J]. International Journal of Production Design, 1978(16): 521-530.
- [4] 陈飞, 徐学林. 尺寸链计算中对称度误差的处理方法[J]. 林业机械与木工设备, 2006, 34(1): 37-38.
- [5] 张伟. 具有形位误差的装配尺寸链的分析计算[J]. 甘肃农业大学学报, 1997, 32(4): 327-330.
- [6] 刘燕德, 蒋育华, 欧阳爱国. 考虑形位误差的装配尺寸链计算探讨[J]. 江西农业大学学报, 2003, 23(1): 158-160.
- [7] 同长虹, 黄建龙, 董世芳. 在尺寸链计算中如何考虑形位公差——公差原则在尺寸链计算中的应用[J]. 现代制造工程, 2008(1): 89-91.
- [8] 高磊. 基于蒙特卡洛法的公差分析及优化设计方法研究[D]. 哈尔滨: 哈尔滨理工大学, 2015.
- [9] 吴炳晖, 裘建军, 官娜. 基于蒙特卡洛法的结构可靠度预计[J]. 机械制造与自动化, 2015, 44(6): 68-70.

收稿日期: 2018-10-29