

Research of Fault Classification in Aero-Engine Based on Multi-Dimension Sphere Information Fusion

Qingzeng Xue¹, Guoliang Fu², Mingzhao Ji³, Dong Li³, Tao Sun²

¹Engine Military Representatives Office of Navy in Shenyang, Shenyang Liaoning

²Naval Aviation Academy, Yantai Shandong

³The Unit 91899 of Army, Huludao Liaoning

Email: happyli.dong@163.com

Received: Jun. 7th, 2018; accepted: Jun. 22nd, 2018; published: Jun. 29th, 2018

Abstract

Aiming at the problem that fault parameters are coupled and classification is difficult to distinguish, distance information fusion on multi-geometrical sphere was introduced to solve the classification problem for multi-faults of aero-engine. This method used self-adaptive model to produce corresponding performance parameter; these parameters were composed of fault sample; the preview state of the sample was taken as classification label; exceptional sample was eliminated, every kind of fault was looked upon as a multi-dimension geometrical sphere, the core and radius of sphere were determined by self-adaptive particle swarm optimization with dynamic neighborhood, thus the sphere can classify fault well, and distance information of spheres was fused, overcoming the defects that classified simply by distance between sample point and core. This method was applied to fault classification in the aero-engine. The result indicates the accuracy of classification reaches 96.67%, and computation time is shorter than other methods.

Keywords

Aero-Engine, Multi-Dimension Geometrical Sphere, Distance Information Fusion, Fault Classification, Improved Particle Swarm Algorithm

基于多维球体信息融合的航空发动机故障分类研究

薛庆增¹, 付国梁², 嵇明钊³, 李冬³, 孙涛²

¹海军驻沈阳地区发动机专业军事代表室, 辽宁 沈阳

²海军航空大学, 山东 烟台

³91899部队, 辽宁 葫芦岛
Email: happyli.dong@163.com

收稿日期: 2018年6月7日; 录用日期: 2018年6月22日; 发布日期: 2018年6月29日

摘要

针对航空发动机各类故障表征参数存在耦合且故障类别不易区分的问题, 引入多维几何球体距离信息融合的方法, 旨在解决航空发动机多故障模式分类的问题。基于发动机自适应模型生成对应的性能参数和测量参数组成故障样本, 利用样本的先验状态作为类别标签。同时去掉异常故障样本, 并将每一类故障数据看成是一个多维几何球体, 利用带动态邻域的自适应粒子群算法确定最优的几何球体的中心以及半径, 得到的球体能更好地区分不同的故障类别。同时融合多个几何球体距离的信息, 克服了以往仅仅依据样本点与故障类别中心的距离进行分类的弊端。将该方法应用到某航空发动机故障分类中。结果表明, 故障分类精度达到96.67%, 且计算时间比其他故障分类方法短。

关键词

航空发动机, 多维几何球体, 距离信息融合, 故障分类, 改进粒子群算法

Copyright © 2018 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

航空发动机是一个非常复杂的热力学系统, 部件参数之间存在着关联和耦合, 各类故障的测量参数存在着交叉和部分重叠, 因此很难对发动机故障进行分类。传统建立的解析模型很难达到理想的精度, 故造成故障分类精度不高; 采用模式识别的方法, 需要大量的训练样本, 同时要求样本表征故障类型全面、特征明显, 由于很难获取大量、准确的样本数据, 加之不同故障参数之间的耦合, 同样造成故障分类精度不高。

国内外许多学者对此类问题进行了积极的尝试与探索。文献[1]提出一种基于几何模式识别技术的发动机传感器故障诊断方法, 但几何距离是定义在原始空间上的, 表征不同故障之间的部分参数存在交叉, 没有充分考虑不同类别的差异, 使得识别效果不理想。文献[2]采用支持向量描述球体的方法, 对模拟线路的故障进行诊断。但此种方法需要对原始数据进行非线性变换, 将非线性不可分的原始数据变换到高维线性可分的空间中, 通过在高维空间上的距离构造真实度和虚假度函数, 并以这两个函数为基础进行分类, 取得较好的分类结果, 但运算时间较长。文献[3]采用粗糙集和支持向量描述的方法研究了发动机视情维修问题, 首先利用粗糙集对原始数据进行处理, 去除异常样本, 而后利用支持向量数据描述方法确定发动机不同维修等级的时机。但支持向量描述方法只根据一类数据的信息, 没有充分融合不同时间的数据。文献[4]基于聚类和多尺度优化的超球体核距离的方法评估发动机性能衰退情况, 综合多个参数的信息, 并采用粒子群算法对各参数尺度进行优化, 凸显不同重要程度的参数, 取得了较好的结果。

本文在上述基础上, 研究提高发动机故障分类精度的方法。首先利用发动机自适应模型生成发动机性能参数, 然后将测量参数和性能参数合并, 组成发动机故障样本, 利用参数的先验状态作为类别标签,

将每一类故障看成是一个多维几何空间的球体，以改进的带动态邻域的自适应粒子群算法优化几何球体的中心和半径，同时融合多个几何球体的距离信息，依此为依据对故障样本进行分类。

2. 发动机自适应模型

准确的发动机自适应模型是开展发动机故障分类的基础和前提。所谓的自适应模型，是指以发动机稳态模型为基础[5][6]，能够通过正向计算获取发动机的测量参数；并且依据测量参数通过反向推算，估计出部件的性能参数的模型。所建立的自适应模型主要具备两个功能：其一，基于发动机稳态模型，在输入量一定的情况下，以部件性能修正因子(压气机、涡轮效率和流量)作为已知量，将修改的部件特性代入到发动机稳态模型中生成一组发动机测量参数；其二在发动机稳态模型基础上，通过给定输入量，将部件性能参数修正因子看成待估参数，将其带到发动机稳态模型中去，通过修改部件特性，产生一组测量参数，与实际测量值对比求得相对误差，联立共同工作方程组作为代价函数，通过一系列优化算法估计出最佳的部件性能修正因子。以测量参数为基础进行故障分类，由于测量参数和性能参数之间的非线性程度非常高，因此有必要对故障样本进行扩充，将性能参数加入到故障样本中去，使得故障样本更好地表征故障类别。

基于自适应模型产生部件性能修正因子的计算如图1所示。

利用发动机自适应模型，经发动机参数相关性分析[7]，并结合外场实际获取的测量参数，最终确定测量参数为低压转子转速 n_l 、高压转子转速 n_h 、低压压气机后压力 p_5 、尾喷口排气温度 T_8 ；性能参数为低压压气机流量 m_{lc} 、高压压气机流量 m_{hc} 、高压涡轮效率 η_{ht} 、低压涡轮效率 η_{lt} ；进而确定故障样本数据为 $\mathbf{p} = [n_l, n_h, p_5, T_8, m_{lc}, m_{hc}, \eta_{ht}, \eta_{lt}]$ 。故障样本的数据取相对量，如下：

$$\mathbf{P} = \frac{\mathbf{P} - \mathbf{P}_0}{\mathbf{P}_0} \quad (1)$$

其中， \mathbf{P}_0 为正常状态下的数值。为了表示方便，相对量表示仍为 \mathbf{P} ，以此作为故障分类的数据基础，引入几何球体距离，进行故障分类。但是在分类之前，构建的故障样本化简后得到的诊断规则可能存在明显的意义相悖或者相近，需要利用粗糙集[8]对数据样本进行筛选，确定最佳的分类样本。

3. 几何球体距离信息融合

以故障样本为基础，依据几何距离进行故障分类，利用故障样本的类别状态作为先验信息，将每一个类用一个几何球体表示，如果待测样本落入到某一类的故障几何球体中，则认为待测样本属于这一故障类别。假设存在 n 个故障类别， d 表征待测样本与故障类别中心的距离， r 表征故障类别几何球体的半径。考虑三种情况：

- 1) 待测样本没有落到任何一种故障类别几何球体中，即

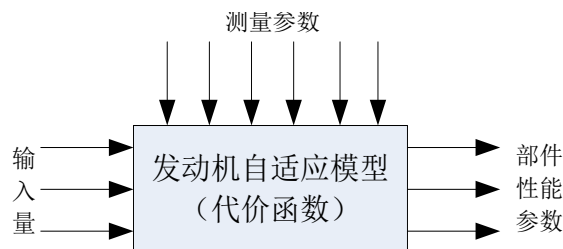


Figure 1. Component performance parameter produced by self-adaptive model

图1. 基于自适应模型产生部件性能参数

$$d > r_i, i = 1, 2, \dots, n;$$

则认为该故障样本不属于任何一种故障类别，属于一种新的故障类别；

2) 待测样本只落入到某一个故障类别几何球体中，不妨设落入到第 l 个故障类别中，即

$$d \leq r_l \text{ 并且 } d > r_i, i = 1, 2, \dots, l-1, l+1, \dots, n;$$

则认为该故障样本属于第 l 个故障类别；

3) 待测样本落入到几个重叠的故障类别几何球体中去，不妨设落入到第 $m, m+1, \dots, n$ 故障类别中去，即 $d \leq r_j$ ，并且 $d > r_i$ ， $i = 1, 2, \dots, m-1$ ， $j = m, m+1, \dots, n$ 。

第(3)种情况判断故障样本所属类别通常的做法是分别计算待测样本与落入到的故障类别几何球体之间的距离，距离最小者对应的故障类别即为所属故障类别。本文在坚持几何距离判断标准的基础上，充分考虑待测样本与落入的其他几何球体的距离，重新设计新的判断依据，将“距离优势”充分放大，直到判断出所属故障样本。基于此种设想，首先引入两个函数，如下：

$$t_j = \begin{cases} 0 & d_j > r_j \\ \frac{1}{p_j}, p_j = \frac{d_j}{\sum_{s=1}^n d_s} & d_j \leq r_j \end{cases} \quad (2)$$

$$f_j = \begin{cases} 0 & d_h > r_h \\ \sum_{h=m+1, m+2, \dots, n, h \neq j} \frac{1}{p_h} & d_h \leq r_h \end{cases} \quad (3)$$

式中， t_j 表示待测样本属于第 j 个故障类别几何球体的可信度， f_j 表示不属于第 j 个故障类别几何球体的可信度。由分析可知，对某个待测故障样本而言， t_j 越大， f_j 越小，表示待测样本属于第 j 个故障类别的概率越大。因此构造故障分类的判别函数如下：

$$T = \frac{t_j}{f_j} \quad (4)$$

由上述分析可知， T 越大，表示待测样本属于第 j 个故障类别的概率越大。构建的判别函数 T 将“距离优势”进一步扩大，即待测样本到所属样本中心的距离小，而到其他故障类别的距离大。式(4)的比值将进一步扩大。融合不同球体距离的信息，依据式(4)即可完成对样本的正确分类。

4. 故障类别几何球体及故障分类

1) 故障类别几何球体

当待测样本判断所属故障类别的函数确定后，下一步确定最优的几何球体中心与半径，以便更好地将故障样本聚类，进行故障分类。

最优几何球体的中心，并不一定是类中所有故障样本的几何中心，这样并不能最佳表征样本的实际意义，某些维数对于分类的作用更加明显，而某些维数对于分类的作用不大，为了凸显不同维数在故障分类中的作用，本文采用优化算法确定几何球体的中心。鉴于粒子群算法[9] [10] [11]具有很强的全局优化能力，能较快地收敛于可接受解，控制参数少，算法简单，效率高的特点，因此采用粒子群算法确定几何球体的中心。在构建各个故障类别样本的过程中，以欧式距离作为判别故障类别的标准。

$$d_{li} = \sqrt{\sum_{d=1}^8 (\mathbf{x}_{ld} - \mathbf{c}_{ld})^2} \quad (5)$$

各类故障样本到类别中心的距离应该是最短的，其中， d 表示欧式距离， \mathbf{x} 表示故障样本， \mathbf{c} 表示类

别中心。为了克服粒子群算法固有的缺点，增强解的多样性和全局收敛速度，对基本粒子群算法进行改进，提出带动态邻域的自适应粒子群算法[12]。首先对所有故障样本归一化处理，消除不同单位制的影响。按照基本粒子群的算法，种群当前迭代到第 $iter$ 代，如果适应度数值 $fit(Pop_{g(iter)}) - fit(Pop_{g(iter-\alpha)})$ 小于预先规定的阈值，其中： $fit(Pop_{g(iter)})$ 为第 $iter$ 代的全局最优解，则认为种群陷入了局部最优，失去了多样性，需要进一步对种群进行更新；进一步将种群中的各粒子按照适应度值从大到小排序，其中： $fit(Pop_1)$ 是适应度最小的个体， $fit(Pop_N)$ 是适应度最大的个体。如果有 $fit(Pop_1) = fit(Pop_N)$ ，则需要对种群中所有个体 ($1 \sim N$) 更新，并且采取随机形式重新生成种群个体；如果有 $fit(Pop_1) \neq fit(Pop_N)$ ，则需要对种群中部分 ($1 \sim D$) 个体进行更新。计算公式如下：

$$\begin{cases} D = \text{ceil}(N \cdot \beta_{iter}) \\ \beta_{iter} = (\beta_e - \beta_s)(iter/iter_M) + \beta_s \end{cases} \quad (6)$$

其中， $\text{ceil}()$ 为向正无穷方向取整， β 为动态因子， D 为动态邻域， N 为种群个体， $iter_M$ 为最大迭代次数。如果差值大于规定的阈值，则说明种群中的个体正在向最优值收敛，此时不需要更新个体。依据上述约定，规定每隔 α 代检查适应度的变化，如果满足更新条件，对种群进行更新，增强种群的多样性，尽快向最优值收敛。

适应度函数 $fit(Pop)$ 的设计同时考虑类内最小距离总和 $d_{\text{intra-class}}$ ，类间最大距离总和 $d_{\text{inter-class}}$ 以及样本的训练精度 acc ，计算公式如下：

$$d_{\text{intra-class}} = \sum_{i=1}^n \sum_{l=1}^{ns_i} \left(\sqrt{\sum_{d=1}^8 (\mathbf{x}_{id} - \mathbf{c}_{id})^2} \right) \quad (7)$$

$$d_{\text{inter-class}} = \sum_{i=1}^n \sum_{l=i}^n \left(\sqrt{\sum_{d=1}^8 (\mathbf{c}_{id} - \mathbf{c}_{ld})^2} \right) \quad (8)$$

$$acc = \frac{nr = nr + 1 \mid \text{if}(F(x_i) = \text{label}(x_i))}{N} \quad (9)$$

$$\max fit(Pop) = 1/[d_{\text{intra-class}} - d_{\text{inter-class}}] + acc \quad (10)$$

其中，式(7)为类间距离总和，式(8)为类内距离总和，式(9)为样本训练精度， nr 为正确分类个数， $F()$ 利用本方法得到的样本分类结果， $\text{label}()$ 表征样本的先验状态对应的类别标签。有了样本的先验状态，在利用带动态邻域的粒子群优化过程中，尽量使各个样本都做到正确的归类，确保最大的分类精度，以便对待测样本做出正确分类。

确定完故障类别几何球体的中心后，继续确定故障球体的半径，计算属于各类别的样本与中心的距离，认为各类别中的半径最大者认为是球体半径，这样就完成了各个故障类别几何球体的构建。

2) 故障分类步骤

构造完故障类别几何球体后，对待测故障样本进行分类。本文确定的故障分类流程如图2所示。

1) 对于一组故障样本(发动机测量数据)，利用发动机自适应模型，得到对应的性能参数，这种性能数据和测量参数一起组成完整的故障样本；

2) 利用粗糙集理论，去掉异常样本，同时利用带动态邻域的自适应粒子群算法，以 $fit()$ 为适应度函数，确定各类样本的几何球体中心及半径，这样便完成了各个故障所属类别几何球体的建立；

3) 利用多维几何球体信息融合方法，依据各个故障类别，以 $T()$ 为类别判断函数，确定待测样本的所属类别。

5. 算例分析及实例分析

1) 算例分析

为了验证本文所用方法的有效性,取UCI数据库中的Wine数据进行验证。Wine数据库是最常用、最典型的分类数据。Wine数据是 178×13 的数据,具有3个类别标签,为了避免随机性对三类样本分类精度的影响,在相同条件下进行3次连续试验检验分类效果。对Wine数据不进行维数扩展,以带动态邻域的自适应粒子群算法确定球体中心和半径,利用球体距离信息融合的方法分类。设定粒子群算法最大迭代次数 $iter_M = 500$,种群规模为 $N = 50$,阈值 $e = 0.01$,成熟期 $\alpha = 50$ 。取Wine数据中的前89组数据进行故障类别的建立,剩余的样本则用于参数分类。其中,第1类30组数据,第2类30组数据,第3类29组数据。分类结果如下。

Wine各类数据中即包含非线性成分,又包含线性成分,由表1可以看出,仅有2个样本出现错误,具有很高的分类精度,因此可以将本文提出的方法用于发动机故障分类。

2) 实例分析

某型航空发动机故障往往不是单一表现形式,表现为复合形式,不同维参数间具有较强的非线性,各类故障参数之间存在交叠。将本文提出的分类算法应用到发动机复合故障诊断中。根据实际获取的故障形式,确定复合故障包括5种形式[13]:无故障(NON);低压压气机+高压压气机故障(LP+HP);高压压气机+高压涡轮故障(HP+HT);高压涡轮+低压涡轮故障(HT+LT);低压涡轮+低压压气机故障(LT+LP)。数据通过有效途径获取,并结合发动机自适应模型,进一步获取故障样本数据。将本文提出的“自适应模型扩充数据维数+粒子群算法优化球体半径和球心+球体距离信息融合分类”的方法称为“方法1”,“不扩充数据维数+粒子群算法优化球体半径和球心+球体欧式距离分类”的方法称为“方法2”,同时为了对比,引入神经网络分类和支持向量机分类两种方法,后者数据样本维数不进行扩充。取训练样本60组数据,测试样本60组数据,分类结果见表2。

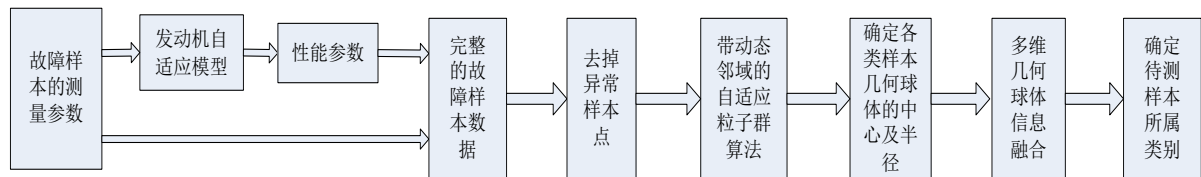


Figure 2. Classification flowchart of fault

图2. 设计的故障分类流程

Table 1. Classification of data for Wine

表1. Wine数据分类结果

分类结果	第1次分类	第2次分类	第3次分类
训练样本分类精度(%)	100	100	100
测试样本分类精度(%)	97.75	97.75	97.75
错分的故障样本	7:1 → 2; 35:2 → 3	7:1 → 2; 35:2 → 3	7:1 → 2; 35:2 → 3

Table 2. Classification of complex faults

表2. 发动机复合故障分类结果

分类结果	方法1	方法2	神经网络	支持向量机
训练样本分类精度(%)	100	95.00	100	98.33
测试样本分类精度(%)	96.67	70.00	90.00	86.66
计算时间(s)	121.24	80.12	152.35	170.56

由表 2 清晰的看到,“自适应模型扩充数据维数 + 粒子群算法优化球体半径和球心 + 球体距离信息融合分类”的方法分类精度最高,“不扩充数据维数 + 粒子群算法优化球体半径和球心 + 球体欧式距离分类”的方法由于采用欧式距离作为分类判据,没有融合多个球体的信息,数据有效维数不高,所以分类精度不高。神经网络的方法需要进行复杂的非线性运算,支持向量机的方法需要将原始分类数据映射到高维空间中,进行非线性变换,计算时间相对较长,并且只利用测量数据作为分类数据,表征的故障数据不全面,因而分类精度不高。

6. 总结

本文针对航空发动机复合故障分类精度不高的问题,提出依据“自适应模型扩充数据维数 + 粒子群算法优化球体半径和球心 + 球体距离信息融合分类”的分类方法,取得了较好的分类效果。得到的主要结论如下:

- 1) 将每一类故障看成是一个几何球体,以带动态邻域的自适应粒子群算法优化得到几何球体的中心和半径,同时融合多个几何球体的距离信息,能准确判别故障的所属类别;
- 2) 对于待测样本,有 3 种可能的故障分类,一种是落入某一球体内,则认为属于该故障类别,一种没有落入任何球体,则认为不属于任何故障类别;一种落入几个球体的重叠区域,则进一步需要对其分析,判断所属的故障类别;
- 3) 本文提出的“自适应模型扩充数据维数 + 粒子群算法优化球体半径和球心 + 球体距离信息融合分类”方法是有效的,和其他分类方法相比,具有很高的精度,分类精度达到了 96.67%,并且计算时间相对较短。

致谢

感谢孙涛和李冬博士在论文撰写中所给予的技术和数据帮助支持。

参考文献

- [1] 黄向华, 丁毅. 基于几何模式识别的发动机传感器故障诊断[J]. 航空学报, 2006, 27(6): 1018-1022.
- [2] Luo, H., Wang, Y.-R. and Cui, J. (2011) A Vague Description Method for Analog Circuit Fault Diagnosis Based on Description Sphere. *Chinese Journal of Aeronautics*, **24**, 768-776. [https://doi.org/10.1016/S1000-9361\(11\)60091-5](https://doi.org/10.1016/S1000-9361(11)60091-5)
- [3] 李冬, 宋岩, 马力, 等. 基于粗糙集和支持向量数据描述的发动机视情维修研究[J]. 燃气轮机技术, 2013, 26(1): 46-50.
- [4] 李本威, 李冬, 王永华, 等. 基于聚类和多尺度优化的超球体核距离评估的航空发动机性能衰退[J]. 推进技术, 2013, 34(7): 977-983.
- [5] 王信德. 某型涡扇发动机自适应建模研究[D]: [硕士学位论文]. 南京: 南京航空航天大学, 2003.
- [6] 李冬, 李本威, 杨欣毅, 等. 自适应核模式分析方法及其在航空发动机部件性能衰退识别中的应用[J]. 推进技术, 2013, 34(9): 1272-1278.
- [7] 尹大伟. 某型发动机模型修正研究[D]: [硕士学位论文]. 烟台: 海军航空工程学院, 2007.
- [8] 薛庆增, 李冬, 孙涛, 等. 基于粗糙集理论的振动信号数据挖掘研究[J]. 航空计算技术, 2010, 25(7): 275-278.
- [9] 郑波. 基于 PSO-SVM 的民航发动机送修等级决策研究[J]. 推进技术, 2013, 34(5): 687-692.
- [10] 郑波, 高峰. 基于 IPSO-SVR 的航空发动机磨损预测研究[J]. 润滑与密封, 2014, 39(11): 81-87.
- [11] 杨欣毅, 刘剑锋, 张强, 等. 粒子群算法求解航空发动机模型的应用[J]. 航空计算技术, 2007, 37(6): 39-41.
- [12] 郑波, 高峰. 基于 S-PSO 分类算法的故障诊断方法[J]. 航空学报, 2015, 39(11): 3640-3651.
- [13] 马力, 李冬, 薛庆增, 等. 基于量子粒子群和流形学习的分类方法及其在发动机故障诊断中的应用[J]. 推进技术, 2015, 34(10): 1275-1283.

知网检索的两种方式：

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择：[ISSN]，输入期刊 ISSN：2161-8801，即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入，输入文章标题，即可查询

投稿请点击：<http://www.hanspub.org/Submission.aspx>

期刊邮箱：csa@hanspub.org