

文章编号:1000-8055(2011)03-0670-05

基于符号序列联合熵的航空发动机 健康监控方法

史永胜, 姜 颖, 宋云雪

(中国民航大学 航空工程学院, 天津 300300)

摘 要: 引入信息熵理论, 结合对多参数时间序列的符号化处理, 探讨其在航空发动机健康监控中的应用. 利用符号序列联合熵对航空发动机性能参数无序程度的描述, 分析航空发动机健康状态及其演化方向. 对两组实际飞行数据的实验分析表明, 该方法描述的航空发动机的健康发展趋势与实际情况相吻合, 能够较好地反映航空发动机的健康状态, 可以为航空发动机的健康监控提供一种新的方法.

关 键 词: 航空发动机; 符号序列; 联合熵; 飞行数据; 健康监控

中图分类号: V23

文献标识码: A

Aero-engine health monitoring method based on joint entropy of symbolic series

SHI Yong-sheng, JIANG Ying, SONG Yun-xue

(College of Aeronautical Engineering,
Civil Aviation University of China, Tianjin 300300, China)

Abstract: Information entropy theory was introduced into this paper for aero-engine health monitoring in combination with symbolized multiple parameters time series. The joint entropy of symbolic series, which depicts the disorder degree of aero-engine performance parameters, was used to analyze aero-engine health condition and evolutive trend. Experiments of two data sets were done, and the results show that the method described in the healthy evolution of aero-engine coincides with the actual situation. The method can reflect the health condition and also can be applied in aero-engine health monitoring.

Key words: aero-engine; symbolic series; joint entropy; flight data;
health monitoring

对航空发动机进行健康监控并对其进行状态评估和故障诊断, 对于推动视情维修、减少飞机运营与维修成本、提高飞机安全性具有重要的意义^[1]. 航空发动机生产商对发动机性能参数如排气温度(T_{eg})、燃油流量(W_f)、转子转速(N_1 和 N_2)、滑油压力(p_{oi})以及滑油温度(T_{oi})等设定了相应的阈值, 大部分发动机的健康监控都包括监

控这些参数是否低于各自的阈值, 以确定发动机是否处于健康状态. 但是性能参数的阈值是为警告最严重的情况而设定的, 即若参数达到或超过该值时, 表示发动机可能发生了故障, 必须进行相应的维修处理^[2]. 仅是将发动机性能参数与一系列阈值进行对比, 对于了解发动机健康状态的内在发展显然是不够的, 不能有效实现对发动机性

收稿日期: 2010-02-07; 修订日期: 2010-04-01

基金项目: 国家自然科学基金委员会与中国民用航空总局联合资助(60879017); 天津市自然科学基金(08JCYBJC11600)

作者简介: 史永胜(1965—), 男, 辽宁锦州人, 教授, 博士后, 研究方向: 结构智能化设计、故障诊断与结构修理、专家系统、知识表示.

能参数的渐变分析^[3-5].而且,航空发动机的性能衰退可以看作是其内部子系统多种因素综合作用的结果,独立分析各个参数也存在较大的不确定性^[6-8],不利于对其运行状态进行综合分析.

本文基于信息熵理论,利用联合熵能将多源信息综合分析表征信源总体特征的特点,结合符号时间序列分析方法,探讨其在航空发动机健康监控中的应用.实验分别以单台航空发动机以及同一民航飞机上对称的两台同型号航空发动机为研究对象,通过性能参数的联合熵变化情况分析发动机健康状态的演化过程.

1 信息熵的有关理论

信息熵是从平均意义上来表征信源总体信息的测度,即信源的平均不确定程度.信息熵越大,表示信息越不确定,信息的无序程度越大.

因此,在工程应用中,可以用多元联合熵来定量刻画一个复杂系统的综合不确定程度,也就是随机程度.一个系统如果随机性很大,表现在其性能参数值的随机性很大,秩序混乱,则其联合熵也很大.航空发动机作为由多个子系统组合而成的复杂系统,其内部无序程度的增加将直接导致其整体性能的衰退^[5].航空发动机在运行过程中,即使相关性能参数都还处于允许范围内,其潜在的无序性的发展仍然是不容忽视的.本文将多元联合熵引入发动机健康监控中,通过量化多个性能参数的随机程度,进行多参数数据综合分析.

以二元联合熵为例,若 X, Y 是一对离散型随机变量 $X, Y \sim p(x, y)$, 则 X, Y 的二元联合熵 $H(X, Y)$ 定义为

$$H(X, Y) = - \sum_{x \in X} \sum_{y \in Y} p(x, y) \log p(x, y) \quad (1)$$

其中对数运算的基底并没有明确规定,以 2 为基底的单位称为 bit(binary digits),而以 e 为基底的单位称为 nats(natural digits).

若 X_i 是 n 个离散型随机变量 $X_1, \dots, X_n \sim p(x_1, \dots, x_n)$, $i = 1, 2, \dots, n$, 则定义 $X_1, \dots, X_n$ 的多元联合熵 $H(X_1, X_2, \dots, X_n)$ 为

$$H(X_1, X_2, \dots, X_n) = - \sum_{x \in X_1} \dots \sum_{x \in X_n} p(x_1, x_2, \dots, x_n) \log p(x_1, x_2, \dots, x_n) \quad (2)$$

从多元联合熵的定义可以知道,联合熵值的大小与数据的具体数值无关,只与每个联合状态出现的频率有关.因此,计算联合熵,关键是要对

原始数据进行恰当的划分.时间序列符号化方法可以保留信号特征与时间状态点之间的良好对应关系,且方法简便,所以本文对发动机监控时序数据进行符号化处理,通过符号时间序列的联合熵分析发动机健康状态的演化.

2 符号时间序列分析

2.1 时间序列符号化

符号时间序列分析是从符号动力学理论、混沌时间序列分析和信息理论发展起来的一种分析方法,其实质是对时间序列的符号化.数据符号化的基本思想就是把时间序列的许多不同的值离散成几个可能值.数据符号化是一个“粗粒化”的过程,这一“粗粒化”过程是从高分辨的数据中生成低分辨的数据,即只捕获大尺度的特征,放弃众多在实际应用中并不关心的特征,留下表达系统关键特征的信息,并且降低了动力学噪声和测量噪声的影响^[9-10].

根据给定数据的值的大小,引入划分 $P = \{P_1, P_2, \dots, P_q\}$, 将原始数据状态空间划分为一系列区间,每一个区间分配不同的符号,构成符号集 $S\{s_i, i=1, \dots, n\}$. q 是划分个数, $q=1$ 是最简单的二进制划分. n 为符号集大小,表示用来把数据符号化的可用符号个数.显然, $n=q+1$.

对任意符号集大小 n , 如何划分数据空间影响到对数据的符号化描述的特征.常用的符号化方法有二进制符号化法^[10]、概率统计符号化法^[11]、小波空间法及伪近邻节点符号化法^[12]等.为了表征发动机性能参数监控数据的随机程度,有效监测发动机的异常状态,本文采用等间距法对原始数据进行划分.

对于原始数据序列 $\{x_m\}$, 等间距法是以数据中最大值和最小值为端点,将其平均划分为 n 个区间,给每个区间分配互不相同的符号,便得到了包含 n 个符号的符号区间 $S\{s_1, s_2, \dots, s_n\}$.将数据序列按照该符号区间的映射关系如式(3)进行数据符号化转换.

$$S = \begin{cases} s_1, & x_{\min} \leq x(k) \leq P_1 \\ s_2, & P_1 < x(k) \leq P_2 \\ \vdots, & \vdots \\ s_n, & P_n < x(k) \leq x_{\max} \end{cases} \quad (3)$$

其中 $x(k)$ 为原始数据序列的数据, P_i 为等间距划分的临界值.这样,时间序列 $\{x_m\}$ 就被转换为符号序列 $\{S_m\}$.

2.2 用联合熵分析符号序列

以二元联合熵为例,采用上面介绍的符号化方法,将2个发动机参数的原始数据序列 $\{x_m^1\}$, $\{x_m^2\}$ 分别转换为包含 n 个符号的符号序列 $\{S_m^1\}$, $\{S_m^2\}$.定义这2个参数的联合熵为

$$H(x^1, x^2) = - \sum_{i=1}^n \sum_{j=1}^n p_{ij} \log p_{ij} \quad (4)$$

式中 $p_{ij} = p(x_i^1) \cdot p(x_j^2 | x_i^1)$,为 (s_i^1, s_j^2) 的联合概率,即序列 S^1 中的第 i 个符号与序列 S^2 中的第 j 个符号组成的符号组合在整个联合符号序列的组合中所占的比重.在实际计算中,一般不会有 $n \times n$ 种组合全部出现的情况,计算时只需统计出现的组合的发生频率,则

$$p_{ij} = k_{ij} / N$$

$$(i = 1, 2, \dots, n; j = 1, 2, \dots, n) \quad (5)$$

其中 k_{ij} 为符号组合 (s_i^1, s_j^2) 出现的总数.

对联合熵进行归一化处理,得到修正的联合熵为

$$H_s(x^1, x^2) = - \frac{1}{\log N} \sum_{i=1}^n \sum_{j=1}^n p_{ij} \log p_{ij} \quad (6)$$

其中 N 为出现的所有符号组合的个数.

3 实验与结果分析

以某航空公司两架飞机的巡航飞行数据为例分别进行实验.第一组以某飞机半年巡航飞行数据为研究对象,提取其一台PW4000发动机排气温度裕度 M_{egt} 、低压转子转速 N_1 、高压转子转速 N_2 及燃油流量 W_f 等参数.第二组以某双发飞机5个月的巡航飞行数据为研究对象,提取其PW4077两台发动机排气温度 T_{eg} , N_1 , N_2 及 W_f 等参数.每组参数均为同一时刻采样.

3.1 原始数据的符号化处理

对原始数据进行符号化转换时,符号集大小 n 不能太小,也不能太大.如果 n 太小,会造成符号区间过少,对数据的变化不敏感; n 太大,则会对数据的正常波动也过于敏感,不能合理刻画数据的随机程度,探测系统的异常状态.另外, n 太大还会增加计算的复杂度.

以第一组数据为例.从图1可以看到,在有限范围内,随着符号集大小 n 的增大,联合熵是不断增加的.而且当 n 为2~5时,联合熵变化较大.

为选择合适的符号集大小 n ,用 $H(n)$ 代表每个参数经过 n 个符号分区后得到的符号序列的归一化修正联合熵,定义

$$\Delta H(n) = H(n) - H(n-1), \quad n \geq 2 \quad (7)$$

图2为 $\Delta H(n)$ 随符号集大小变化曲线.从图中可以看到,当 n 取10和11时, $\Delta H(n)$ 最小且趋于稳定,但符号区间较多.而当 n 取7时, $\Delta H(n)$ 达到波谷值,是较优的选择.所以,本组实验最终将符号序列划分为7个区间.

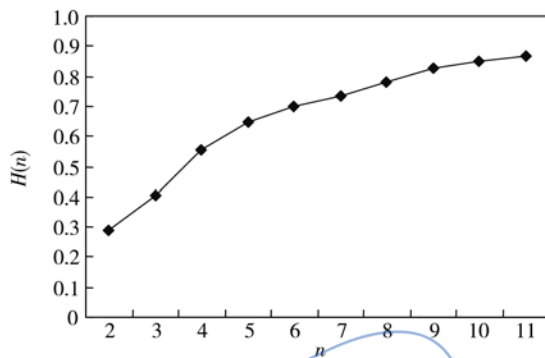


图1 联合熵随符号集大小变化曲线

Fig. 1 Joint entropy versus the size of symbol set

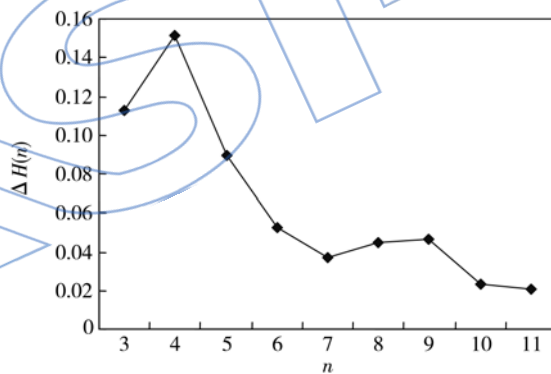


图2 $\Delta H(n)$ 随符号集大小变化曲线

Fig. 2 $\Delta H(n)$ versus the size of symbol set

3.2 发动机健康状态趋势分析

为表征发动机健康状态发展趋势,对经过符号化处理的发动机巡航数据,将每月数据分别按式(6)进行计算,对数运算统一以2作为底数.

图3给出了第一组发动机以 M_{egt} , N_1 , N_2 , W_f 为评估参数的联合熵变化曲线.从图3可以看到,随着飞行时间的增加,联合熵也在逐渐增加.说明随着使用时间的增加,发动机监控数据的无序程度在增大,规律性不断减少,发动机系统性能在逐步地衰退.这与实际情况是相吻合的.尤其是第6个月,联合熵出现较大幅度增加,说明发动机健康状况恶化明显.

图4给出了第二组发动机以 T_{eg} , N_1 , N_2 及 W_f 为评估参数的联合熵变化曲线.从图中可以看

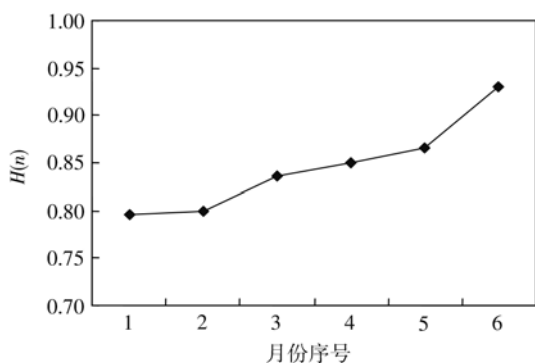


图3 PW4000 发动机联合熵随时间变化曲线

Fig. 3 Joint entropy of PW4000 engine versus time

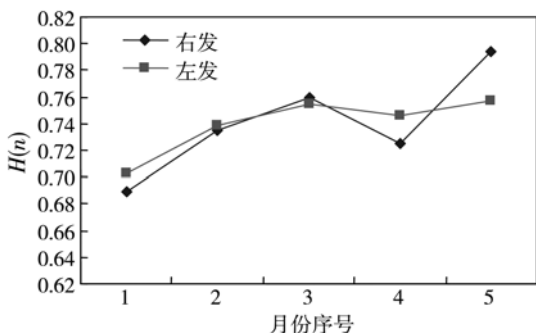


图4 PW4077 发动机联合熵随时间变化曲线

Fig. 4 Joint entropy of PW4077 engine versus time

到,两台对称的发动机健康程度总体呈现衰退趋势,但是在第4个月时出现了波动.这说明发动机局部缺陷在运行过程中没有迅速扩展,对发动机整体性能呈现周期性影响,表现在性能参数含有规律性数据,联合熵值有所降低.这种现象是发动机性能整体衰退、有时波动的一种体现.值得注意的是这两台发动机此后两个月联合熵值呈现较大差异,两台发动机的健康状态发展已经不再同步,需要重点进行相关监控及检测,进行进一步的诊断,排除故障,避免由于两台发动机负载不平衡等原因,导致加剧其中一台发动机的性能恶化,缩短发动机的使用寿命.

作为一种先验性的辨识手段,航空发动机联合熵是研究发动机整体健康状况的指标,而不是研究发动机中具体部件是否发生性能衰退.对于同一型号的发动机,在一定时间段内,采用相同的监控参数,并进行相同的符号化处理后,对于特定的健康状态,确定联合熵值的范围,即熵带,可以作为该健康状态的参考阈值.

为了进一步分析联合熵的变化趋势,将以上三台发动机的联合熵值分别进行指数曲线拟合.其中,指数曲线拟合函数为

$$Y = ae^{bx} \quad (8)$$

拟合得到各自联合熵变化的趋势线如图5所示.联合熵趋势线更加显著地表征了三台发动机均呈现性能衰退的趋势.对于两台PW4077发动机,右发联合熵趋势线变化率大于左发的,说明右发联合熵变化程度较大,性能衰退状况发展较快.从联合熵值拟合曲线可以发现,航空发动机性能衰退的快慢程度与 b 成正比,该值可作为航空发动机的健康风险评价的重要指标.

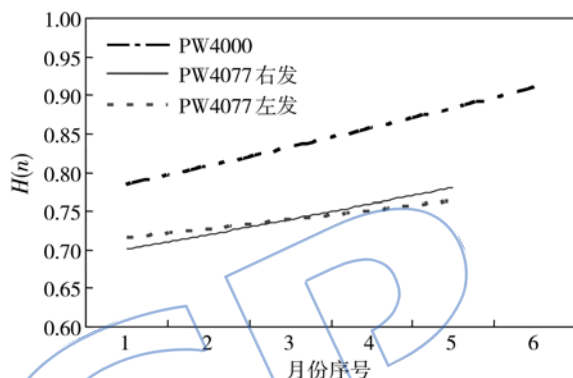


图5 联合熵的指数曲线拟合

Fig. 5 Exponential curve fitting of joint entropy

4 结论

结合符号时间序列分析方法,将联合熵引入航空发动机健康监控中,能够通过量化发动机多个性能参数的整体无序程度,在一定程度上反映发动机系统的健康状态.该方法实现简单,根据实际的需要,可以选择不同的参数进行分析.联合熵的变化情况与发动机的实际性能演化规律相符合,可以将联合熵作为发动机健康监控的指标,为分析发动机健康状态发展提供依据.同时,对变化的联合熵进行曲线拟合,有助于进一步了解发动机健康状态的变化程度.

参考文献:

- [1] Babbar A, Syrmos V L, Ortiz E M, et al. Advanced diagnostics and prognostics for engine health monitoring[C]// Proceedings of IEEE Aerospace Conference, Piscataway, NJ, United States: IEEE Computer Society, 2009: 1-10.
- [2] Ortiz E, Babbar A, Syrmos V L. Extreme value theory for engine health monitoring and diagnosis[C]// Proceedings of the IEEE International Conference on Control Applications, New York, United States: IEEE, 2009: 1069-1074.
- [3] Bronson R J, Depold H R, Rajamani R, et al. Data normalization for engine health monitoring[R]. ASME GT2005-68039, 2005.

- [4] Yedavalli R K. Robust estimation and fault diagnostics for aircraft engines with uncertain model data[C]// Proceedings of the American Control Conference, Piscataway, NJ, United States; IEEE, 2007: 2822-2827.
- [5] Littles J W, Jr., Pettit R G, Schirra J J, et al. Demonstration of an advanced prognostic health management approach for gas turbine engines[C]// Materials Damage Prognosis: Proceedings of a Symposium of the Materials Science and Technology 2004 Conference, New Orleans, LA, United States; Minerals, Metals and Materials Society, 2005: 191-196.
- [6] 李俊涛, 张凤鸣, 李永宾. 基于多元联合熵的航空发动机性能分析[J]. 微计算机信息(测控自动化), 2007, 23(1): 53-54.
LI Juntao, ZHANG Fengming, LI Yongbin. Performance analysis of aero-engine based on multi-joint entropy[J]. Control & Automation (Measurement & Control), 2007, 23(1): 53-54. (in Chinese)
- [7] Volponi A J, Brotherton T, Luppold R. Development of an information fusion system for engine diagnostics and health management[R]. NASA 2004-212924, 2004.
- [8] 张亮, 杜纯, 张凤鸣. 基于多元联合熵的复杂装备健康监控方法研究[J]. 传感器与微系统, 2009, 28(3): 45-47.
ZHANG Liang, DU Chun, ZHANG Fengming. Research on health monitoring method for complex equipment based on multi-joint entropy[J]. Transducer and Micro-system Technologies, 2009, 28(3): 45-47. (in Chinese)
- [9] LI Biqiong, CHEN Dingfang, ZHANG Yu. The theory and method of symbolic time series[C]// Proceedings of the First International Conference on Machine Learning and Cybernetics, Beijing: [s. n.], 2002, 4: 2201-2204.
- [10] 张雨, 任国峰, 王爱国. 基于符号序列 Shannon 熵提取机器的运行特征[J]. 中国机械工程, 2006, 17(15): 1595-1599.
ZHANG Yu, REN Guofeng, WANG Aiguo. Operational feature extraction of machine based on Shannon entropy of symbolic series[J]. China Mechanical Engineering, 2006, 17(15): 1595-1599. (in Chinese)
- [11] 马东玲, 陈晓平, 梁晓花. 基于符号序列 Shannon 熵的轴承故障趋势分析[J]. 微计算机信息(测控自动化), 2008, 24(31): 172-173.
MA Dongling, CHEN Xiaoping, LIANG Xiaohua. Analysis of fault trend of bearings based on Shannon entropy of symbolic series[J]. Control & Automation (Measurement & Control), 2008, 24(31): 172-173. (in Chinese)
- [12] Tolani D K, Yasar M, Ray A. Anomaly detection in aircraft gas turbine engines[J]. Journal of Aerospace Computing, Information, and Communication, 2006, 3(2): 44-51.