

# 基于模糊熵与方向相似度的液体火箭发动机故障检测\*

谢涛 张育林

(国防科技大学航天技术系 长沙 410073)

**摘要** 先定义系统的广义模糊熵及其计算公式,再给出滑动数据窗口中采样数据矢量方向分布中心的离散度定义及其与模糊熵之间的负指数解析关系,并基于离散度概念对自适应窗口滤波器进行了改进,提高了滤波器对噪声的敏感度以及故障检测算法对强干扰噪声环境的鲁棒性。同时,根据滑动数据窗口中二模糊聚类数据中心矢量方向相似度的变化来监视发动机系统故障的发展趋势。本文基于受强噪声环境干扰的实际试车数据并用自组织模糊聚类算法作为滑动数据窗口上的聚类算法进行了数字仿真试验。仿真结果表明:基于模糊熵的故障检测算法具有对强噪声环境的鲁棒性,是低信噪比环境下的一种客观的故障检测算法。

**关键词** 液体火箭发动机, 故障检测, 滑动数据窗口, 模糊熵

**分类号** V434

## Fuzzy Entropy and Directional Similarity Based-Fault Detection and Its Applications to Liquid Rocket Engine

Xie Tao Zhang Yulin

(Department of Astronautical Technology, NUDT, Changsha, 410073)

**Abstract** This paper proposes a definition for the fuzzy entropy of clustered result, i. e., the fuzziness measure of a fuzzy set, and based on this definition, the definition for a measure of the distribution divergence of the sampled data vectors is also given in this paper for the first time. The divergence definition offers an ideal and objective criterion about whether there exist heavily interfered data vectors in the sliding data window for the Adaptive Windowing Filter. Thus, the noise discrimination ability has been much improved, and this increases the robustness of the fault detection algorithm to the strong noises and the sensitivity to subtle fault transients. In addition, the directional similarity between the two clustered centroid vectors (corresponding to the sliding data window) has been employed to detect faults and distinguish the intermittent faults from the failure. The practical testing data with strong noises have been used to verify its efficiency and objectivity in the fault detection and discrimination from the noised data vectors.

**Key words** liquid rocket engine, fault detection, sliding data window, fuzzy entropy.

对于复杂系统的故障监测,可以基于功能主义的“黑箱”理论,应用数理统计技术中的聚类方法、模式识别方法,从表征系统行为的输入输出测量信号中提取系统的正常与故障特征信息,作为诊断系统的知识依据<sup>[1-3]</sup>。文[2]介绍了一种适用于离线或在线系统故障检测与诊断算法框架,该算法框架的内核是自适应窗口滤波器,故障检测的准确率取决于自适应窗口滤波器对噪声的敏感度。文[2]中假定滑动数据窗口上的聚类结果使得较少数据矢量的个数少于或等于人为设定个数的聚类组中的数据是受噪声干扰的异常数据,这种对滑动数据窗口中采样数据矢量方向的离散度定义是硬边界的,因此,不适应于强噪声环境下使用自适应窗口滤波器对噪声的滤波。本文基于滑动数据窗中数据矢量方向分

\* 1997年5月20日修订

第一作者: 谢涛,男,1966年生,博士

布中心离散度的模糊熵定义, 改进了检测算法的流程图。

## 1 滑动数据窗口中数据矢量方向分布中心离散度的模糊熵定义及其计算公式

一般地, 液体火箭发动机地面试车是一个强噪声环境, 监测参数受到相互耦合的强噪声干扰, 试车数据统计分析表明: 相互耦合的强噪声在幅值上呈均值不为零的正态分布而在时序上的随机性服从一阶或二阶 Markov 过程。图 1 说明了一次正常试车过程中某些参数的随机波动情况。因此, 滑动数据窗口中的采样数据矢量在欧氏空间中的矢量方向几何

分布虽然具有随机性, 但总体上具有不同的分布中心, 且受强噪声干扰的采样数据矢量在欧氏空间中的几何分布会较明显地偏离未受噪声干扰的采样数据矢量并呈正态分布, 一般地, 滑动数据窗口中受强噪声干扰的数据矢量大大少于未受干扰的数据矢量。高维欧氏空间上的数据矢量几何分布的离散程度可以用欧氏距离或矢量方向的离散程度来描述, 鉴于监控数据矢量方向在故障检测中的实用性。本文仍使用数据矢量方向的几何分布离散度。由于稳态工况下系统状态及测量参数的稳定性, 未受噪声干扰的数据矢量在高维欧氏空间中几何分布均匀且高度集中, 其矢量伸展方向高度一致, 因此可以通过对滑动数据窗口上的数据矢量进行聚类来确认当前数据窗口中是否存在受强噪声干扰的数据矢量, 以便滤去受强噪声干扰的数据矢量。正确识别当前数据窗口中是否存在异常数据矢量的关键问题是如何定义滑动数据窗口中数据矢量方向分布中心的离散度。

如果当前滑动数据窗口中不存在受强噪声干扰的数据矢量, 则对该窗口中的数据矢量通过 FKCN 聚类算法得到的两组数据矢量在个数上基本一致, 且数据矢量的模糊隶属度 (两组数据矢量的聚类中心矢量构成二互补模糊集合) 均接近 0.5, 即每一数据矢量对二互补模糊集合的隶属性极其模糊。相反, 如果当前滑动数据窗口中存在受强噪声干扰的数据矢量, 则对该窗口中的数据矢量通过 FKCN 聚类算法得到的两组数据矢量在个数上相差显著, 且数据矢量的模糊隶属度为 0 或 1 (不属于第一组就属于第二组矢量), 即每一数据矢量对二互补模糊集合的隶属性极其清晰。样本对模糊集合的隶属度的模糊性定义为模糊集合的模糊度。

如果当前滑动数据窗口中不存在受强噪声干扰的数据矢量, 则对该窗口中的数据矢量通过 FKCN 聚类算法得到的两组数据矢量在个数上基本一致, 且数据矢量的模糊隶属度 (两组数据矢量的聚类中心矢量构成二互补模糊集合) 均接近 0.5, 即每一数据矢量对二互补模糊集合的隶属性极其模糊。相反, 如果当前滑动数据窗口中存在受强噪声干扰的数据矢量, 则对该窗口中的数据矢量通过 FKCN 聚类算法得到的两组数据矢量在个数上相差显著, 且数据矢量的模糊隶属度为 0 或 1 (不属于第一组就属于第二组矢量), 即每一数据矢量对二互补模糊集合的隶属性极其清晰。样本对模糊集合的隶属度的模糊性定义为模糊集合的模糊度。

设:  $V = \{v_1, v_2, \dots, v_T\}$ , 其中  $v_i$  ( $1 \leq i \leq T$ ) 为  $d$  维采样数据矢量;

$O = \{c_1, c_2, \dots, c_c\}$ , 其中  $c_i$  ( $1 \leq i \leq c$ ) 为通过 FKCN 聚类算法得到的  $d$  维聚类中心数据矢量;

$u_{ij} \in [0, 1]$ ,  $u_{ij} = \frac{1}{T} \sum_{i=1}^T u_{ij}$ ,  $1 \leq j \leq T, 1 \leq j \leq c$ , 为通过 FKCN 聚类算法得到的数据矢量  $v_j$  对以  $d$  维聚类中心数据矢量  $c_i$  所代表的模糊集合的隶属度。

则模糊聚类结果 (FKCN 聚类算法所形成的以  $c_i$  为聚类中心矢量的  $c$  个模糊聚类集合) 的模糊度可如下计算:

$$F(O) = - \frac{1}{T} \sum_{j=1}^T \sum_{i=1}^c u_{ij} \ln(u_{ij});$$

用 Lagrange 乘子法容易证明: 当  $u_{1j} = u_{2j} = u_{3j} = \dots = u_{cj} = \frac{1}{c}, 1 \leq j \leq T$  时, 也即聚类结果及其模糊 (同一聚类中心矢量) 时, 模糊度  $F(O)$  取最大值  $\ln c$ ; 相反, 当  $u_{ij} \in \{0, 1\}$  时, 模糊度  $F(O)$  取最小值 0。因此, 聚类结果的广义模糊度可用来度量论域中样本分布中心的离散度。由于 De Luca A. 和 Termini S. 早在 1972 年发表了适应于两个互补模糊集合或语句存在的狭义模糊熵<sup>[7]</sup>, 且  $F(O)$  在形式上类似 Shannon 熵<sup>[8]</sup>, 因此称其为广义模糊熵。

容易看出, 滑动数据窗口中数据矢量在欧氏空间上的方向分布中心的离散程度与滑动数据窗口上聚类结果的模糊熵成反比, 数据矢量方向分布中心的离散程度越大, 聚类结果的广义模糊熵就越小。进

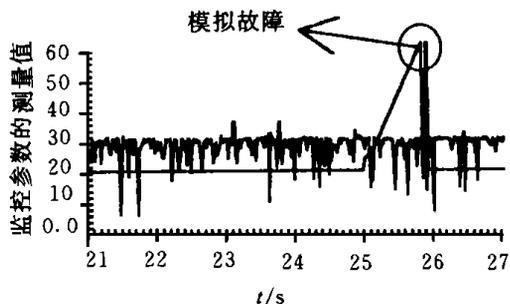


图 1 受强噪声干扰的监控参数的测量值

一步定义滑动数据窗口中数据矢量方向分布中心的离散度为广义模糊熵的负指数, 即:

$D(V, \hat{O}) = e^{-F(O^*)}$ , 其中  $O^*$  为最优聚类中心。易知,  $D(V, \hat{O})_{\max} = 1.0$ ,  $D(V, \hat{O})_{\min} = e^{-lnc} = \frac{1}{c}$ 。根据离散度  $D(V, \hat{O})$  值即可确定当前滑动数据窗口中是否存在受强噪声干扰的数据矢量。

取  $D(V, \hat{O}) = \frac{1}{c} + k_1$  为滤波门限值, 作为滑动数据窗口中是否存在异常数据矢量的判断准则, 其中  $0 < k_1 < 1 - \frac{1}{c}$ 。  $k_1$  越小, 自适应窗口滤波器对噪声的敏感度越高, 滤波效果越好; 相应地, 基于自适应窗口滤波器的故障检测算法对故障的敏感度越大。图 2 (a) 说明受随机噪声干扰的参数经自适应窗口滤波器处理后的结果 ( $k_1 = 0.45, c = 2$ ), 滑动数据矢量窗口上的模糊二分类聚类采用 FKCN 技术。

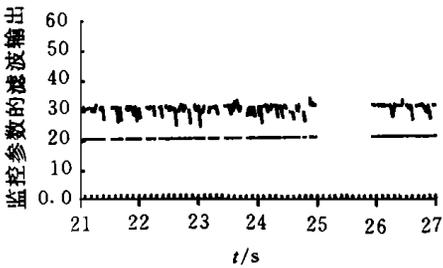


图 2 (a) 基于离散度模糊熵定义的自适应滤波器的滤波结果

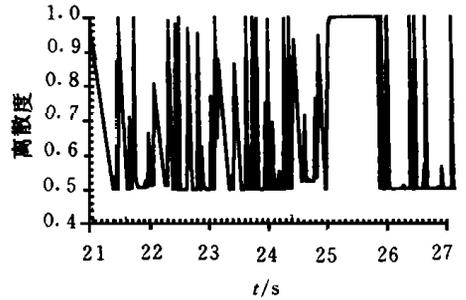


图 2 (b) 采用基于模糊熵的自适应窗口滤波器滤波时, 滑动数据窗口中数据矢量分布中心的离散度值 (基于模糊熵定义) 随时间的变化

## 2 模糊自组织聚类算法: FKCN 算法

FKCN 自组织聚类算法在聚类迭代过程中利用其自组织功能, 不断地修改其学习因子及待学习的相邻样本数目, 使聚类问题逐渐收敛到其全局最优解或局部最优解。FKCN 实质上是 FCM 聚类算法与 KCN 聚类算法的有机结合, 每一次 FKCN 迭代过程即是一次相关参数变化了的 FCM 聚类算法。因此从“自适应性”角度来说, FKCN 优于 FCM 聚类算法, 从而聚类结果的全局性及收敛性也优于 FCM 聚类算法<sup>[5,6]</sup>。

## 3 数字仿真结果及其比较分析

数字仿真实验中的测试数据是由某液体火箭发动机正常试车参数的磁带记录经采样频率为 50Hz 的采样系统采样而形成的测试数据包, 选取 8 个参数进行故障检测的数字仿真试验, 其中包括受强噪声干扰最大的氢泵壳温度 (T2), 氢泵壳温度受到的强噪声干扰在幅值上有时高达额定值的 3 至 4 倍, 而且是双向分布的叠加噪声。整个地面试验过程持续 380.00s, 发动机起动段持续 2.70s 左右即进入稳定的主级运行段运行, 为验证故障检测算法的性能, 对 25.00s 到 25.82s 间的采样数据进行某种定向的微小修改, 如让 8 个参数测量值中某些参数持续减小, 而让某些参数测量值持续增大, 以模拟某些典型故障的过渡过程, 图 1 给出了其中一个参数的变化情况。

本文用 FKCN 聚类算法对滑动数据窗口上的数据矢量进行二分组聚类试验, 并与基于硬边界的数据矢量方向分布离散度定义的故障诊断结果进行了比较。数字仿真试验结果表明: 二分组聚类能准确地检测出故障, 并具有对干扰噪声的强鲁棒性, 基于硬边界离散度定义的故障检测算法有可能产生误报警 (由未被滤去的强噪声干扰所引起) 或漏报警 (系统故障被强噪声所淹没)。

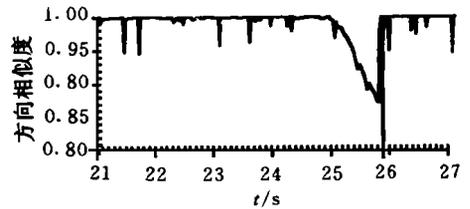


图 2 (c) 滑动数据窗口中二统计聚类中心矢量的方向相似度随时间的变化关系

图 2 (b) 说明基于数据矢量方向几何分布离散度的模糊熵定义并利用 FKCN 聚类算法, 对滑动数据窗口中的数据矢量进行模糊二分组聚类来检测故障时, 滑动数据窗口中数据矢量方向分布中心的离散度随采样时刻的变化趋势。容易看出, 当故障检测算法处理到试车数据包中的 25.00s 时刻后, 滑动数据窗口中数据矢量的方向分布离散度值持续超过  $(1/c + k_1)$  值。直到人为设置的故障过渡过程结束。同时, 滑动数据窗口中二模糊聚类数据中心矢量方向随采样时间发生基本定向的偏离, 且偏离度越来越大, 相应地, 二模糊聚类数据中心矢量方向相似度随采样时间越来越小, 图 2 (c) 即说明滑动数据窗口中二模糊聚类数据中心矢量方向相似度值随采样时间的变化特征。故障检测算法可根据滑动数据窗口中二模糊聚类数据中心矢量方向相似度的变化, 对发动机系统是否发生故障 (并区分间歇性故障与永久性故障) 作出判断。

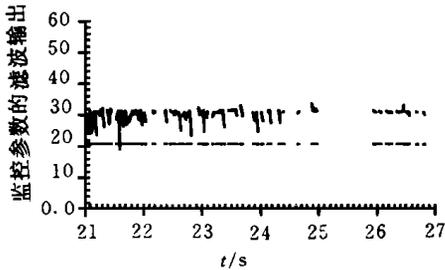


图 3(a) 基于硬边界离散度定义的自适应滤波器的滤波结果

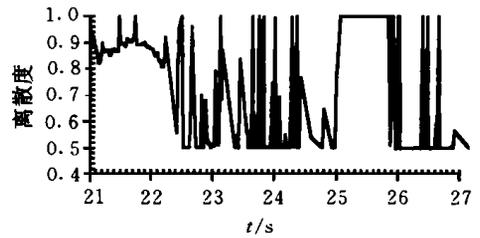


图 3(b) 采用硬边界离散度定义的自适应窗口滤波器滤波时, 基于模糊熵的负指数定义的数据矢量分布中心离散度值随时间的变化关系

图 3 (a) 说明基于硬边界的离散度定义并利用 FKCN 聚类算法, 对滑动数据窗口中的数据矢量进行模糊二分组聚类分析时的滤波结果。与图 2 (a) 比较可知, 基于硬边界的离散度定义的自适应窗口滤波器对噪声的敏感度低于基于离散度的模糊熵定义的自适应窗口滤波器。图 3 (b) 表明相应于图 3 (a) 的自适应窗口滤波过程中, 基于模糊熵定义的滑动数据窗口中数据矢量方向分布中心的离散度随采样时刻的变化趋势。容易看出, 当检测算法处理到 21.10s 时, 可能产生故障误报警 (由未被滤去的强噪声干扰所引起), 虽然也能检测出 25.00s 后人为所设置的轻度故障。

## 4 结束语

基于聚类结果模糊熵的负指数定义的离散度客观地反映了数据矢量聚类中心的离散程度, 用实际试车数据进行故障检测的数字仿真结果表明: 离散度的模糊熵定义可以成功地用于自适应窗口滤波器对数据矢量方向分布中心的离散程度进行客观描述, 提高了滤波器对噪声的识别能力, 因而增强了基于自适应窗口滤波器的故障检测算法对强噪声干扰的鲁棒性, 提高了算法对故障的敏感度。此外, 仿真结果还表明可根据滑动数据窗口中二模糊聚类数据中心矢量方向相似度的变化特征来对发动机系统是否发生故障作出判断, 并且可用二模糊聚类数据中心矢量方向相似度的变化来监视系统故障的发展趋势。

因为基于模糊熵的离散度定义依赖于滑动数据窗口上的模糊聚类结果, 因而较难用于故障的实时检测与诊断。如果已知所有监控参数所受随机干扰噪声的统计分布规律或噪声在幅值分布上的先验概率密度, 就可以基于 Shannon 概率意义上的统计信息熵来度量当前时刻采样数据矢量偏离系统正常工作的程度, 据此即可实时判断系统是否发生故障及故障的大小程度。随机干扰噪声在幅值分布上的先验概率密度可以根据现有试车数据通过统计分析得到, 因此, 基于 Shannon 统计熵的液体火箭发动机的实时故障检测算法可望进行工程应用开发。

## 参考文献

- 2 谢涛, 张育林. 基于滑动数据窗口聚类结果分析的复杂动力系统故障检测与诊断框架及其应用. CW/CICIA 97. 西安, 1997
- 3 谢涛, 张育林. 基于聚类数据中心矢量方向相似度的液体火箭发动机故障检测研究. 湖南省宇航学会推进年会, 1996
- 4 E. Nemeth And A. M. Norman, Jr.. Development Of A Health Monitoring Algorithm. AIAA/SAE/ASME/ASEE 26th Joint Propulsion Conference, July 16 ~ 18, 1990/Orlando, FL
- 5 C. Bezdek. Pattern Recognition With Fuzzy Objective Function Algorithms. Plenum. New York, 1981
- 6 Eric Chen-Kuo Tsoa, James C. Bezdek And Nikhil R. Pal. Fuzzy Kohonen Clustering Networks, Pattern Recognition. 1994, 27 (5): 757
- 7 De Luca A and Termini S. A definition of nonprobabilistic entropy in the setting of fuzzy sets. Infor. Contr., 1972, 20: 201
- 8 鲁晨光著. 广义信息论. 合肥: 中国科学技术大学出版社, 1993