

$S(x_1, x_2, \dots, x_n)$, 故障类集为 $W(\vec{w}_1, \vec{w}_2, \dots, \vec{w}_m)$, 其中 $\vec{w}_i = (x_1, x_2, \dots, x_n)$, 就是要寻找低维特征 $s(x_1, x_2, \dots, x_k) \in S, k < n$, 其中 k 为设定常数。以便对相应的故障类集 $W(\vec{w}_1, \vec{w}_2, \dots, \vec{w}_m)$ ($\vec{w}_i = (x_1, x_2, \dots, x_k)$), 使得:

$$A = \min \left\{ \arccos \frac{[\vec{w}_i, \vec{w}_j]}{||\vec{w}_i|| \cdot ||\vec{w}_j||}, i, j = 1, 2, \dots, k \right\}$$

在 C_n^k 种可能的组合中最大。但这种组合优化在 n 大于几十时, 所拥有的可能集就十分巨大, 采用穷尽搜索寻找最优集是不现实的。我们采用了一种简单贪婪算法和一种基于互信息的贪婪算法来实现这一过程。

2.1 简单贪婪算法的具体步骤

(1) 给定原始特征集 $S(x_1, x_2, \dots, x_n)$; (2) 对所有特征 $x_i \in S (i=1, 2, \dots, n)$, $S - x_i \rightarrow s$. 计算 S 所有子集 s 的 A . 找到特征 x_i , 有 $A_i = \max(A)$; 且所有故障样本分类正确; (3) $S - x_i \rightarrow s, n = n - 1$; (4) 如果 $n \neq k$, 转向 (2), 否则结束。

该算法的实质是根据分类准则, 单个淘汰特征, 只考虑了单个特征对分类的作用, 没有考虑多个特征的组合对分类的作用。但在考虑单个或多个回溯过程时, 就很容易陷入局部最优区出不来。因此需要考虑另外的方法。

2.2 基于互信息的特征选择

对于发动机故障诊断中所涉及到的故障模式识别问题, 由文献 [2] 知, 如果假设所有样本共涉及到 W_c 种故障模式, 当各故障类样本以一定的概率随机出现时, 判断样本所归属的类别存在一定的不确定性, 可用熵 $H(C)$ 来表示这种不确定性的程度。用于故障分类的原始特征集 $S(x_1, x_2, \dots, x_n)$ 中的所有特征对于每一个故障样本构成一个 n 维的矢量 $\vec{r}$, 全体故障样本构成一故障矢量集 R , R 有 N_r 种取值的可能。在已知矢量 $\vec{r}$ 的条件下, 类别 C 的不确定性用条件熵 $H(C|R)$ 表示。一般已知 $\vec{r}$ 的条件下 C 的不确定性小于 C 本身的不确定性, 即 $H(C|R) < H(C)$, 两者之间的差异用 $I(C; R)$ 表示:

$$I(C; R) = H(C) - H(C|R) = \sum_{c=1}^{W_c} \sum_{r=1}^{N_r} P_{c,r}(c, \vec{r}) \log \frac{P_{c,r}(c, \vec{r})}{P_c(c)P_r(\vec{r})}$$

$I(C; R)$ 代表了 R 中包含的 C 的信息, 它与 C 中包含的 R 的信息相等, 称之为两者之间的互信息。因为故障样本 $\vec{r}$ 是 n 维的矢量, 具体计算时, 将 C 与 $\vec{r}$ 中的每一维特征之间的互信息分别考虑, 设 x_i 是其中某一特征, 共有 N_{x_i} 种离散取值, 则 x_i 与 C 的互信息为:

$$I(C; x_i) = \sum_{c=1}^{W_c} \sum_{x_i=1}^{N_{x_i}} P_{c,x_i}(c, x_i) \log \frac{P_{c,x_i}(c, x_i)}{P_c(c)P_{x_i}(x_i)}$$

因为不同的特征 x_i 包含的分类信息是有差异的, 如果从所有特征中挑选对分类贡献最大的特征子集, 一般是选择那些 $I(C, x_i)$ 取值高的特征。但由于特征之间不是独立的, 设 x_i 和 x_j 是两个不同的特征, $I(x_i, x_j)$ 反映了两者的独立程度:

$$I(x_i; x_j) = \sum_{x_i=1}^{N_{x_i}} \sum_{x_j=1}^{N_{x_j}} P_{x_i,x_j}(x_i, x_j) \log \frac{P_{x_i,x_j}(x_i, x_j)}{P_{x_i}(x_i)P_{x_j}(x_j)}$$

该式表明: 特征 x_i 和 x_j , 若对于同样的样本, 它们取值的分布处于同样的区段时, 两者之间的互信息就大, 反之就小。它的直观解释如图1和图2所示, 图中方形和椭圆形代表两个特征。由此得到

确定特征子集的准则为类-特征互信息量大而同时子集中特征-特征之间的互信息量小。相应的特征选择算法为^[3]：(1) 原始的特征集 S 包含所有的 n 个特征；特征子集 S' ， $S' \leftarrow \phi$ ；(2) 对于每一个特征 $x \in S$ ，计算 $I(C, x)$ ，设其中 $x_{\max}$ 使 $I(C, x)$ 达到最大，令 $S \leftarrow S - \{x_{\max}\}$ ， $S' \leftarrow \{x_{\max}\}$ ；(3) 对所有的 $x \in S$ 和 $s \in S'$ ，计算 $I(C, s)$ ；(4) 定义综合互信息指标： $I(C; x) - \sum_{s \in S'} I(s; f)$ ，选择 S 中的特征 x 使上式最大，并令： $S \leftarrow S - \{x\}$ ， $S' \leftarrow \{x\}$ ；(5) 如果 $|S'| = N$ ，输出 S' ，否则转向 (3)。 N 为设定的特征子集 S' 中的特征的目标个数。

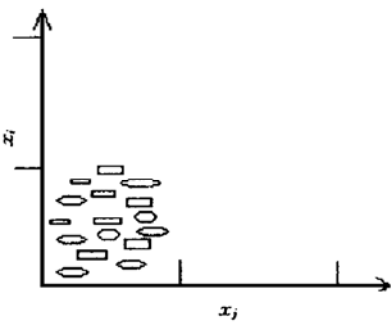


Fig. 1 Schematic diagram with large mutual information

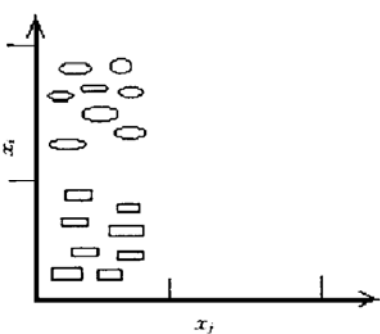


Fig. 2 Schematic diagram with small mutual information

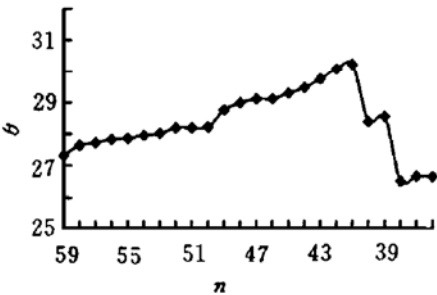


Fig. 3 Feature number of optimal feature set

性不但没有减少还有所增加。采用方法2比方法1的效果要好得多,我们将每个特征量在考虑所有故障时的变化范围划分为7个区间的情况下，最小的特征数目为17个。详细结果见表1，2。

3 故障特征选择结果及分析

对文献 [1] 中的故障样本集，利用方法1对原始特征进行特征选择的计算表明（见图3），因原始特征之间的相关性而导致的分类信息冗余，使得 $\max(A)$ 随特征个数的减少而增加。最小的特征数目为36个，比原始特征的60个减少了近一半，所有的

测试样本都实现了正确分类，且故障类之间的可分性

Table 1 The minimum angle between class with different method

The source of feature subspace	The number of feature	The minimum angle between class
Method 1	36	26.62
Method 2	17	31.25
The original space of feature	60	26.82

Table 2 The max angle between class of sixteen kinds of fault with different feature set

The code of fault	Method 1	Method 2	The original space of feature
1	17.610	14.808	18.663
2	11.736	15.919	12.381
3	42.726	27.481	38.288
4	5.078	5.686	5.204
5	26.745	30.456	30.633
6	2.581	8.572	10.819
7	11.800	11.954	10.036
8	2.278	4.809	3.926
9	15.268	14.486	17.256
10	40.402	25.927	47.108
11	3.713	3.555	3.419
12	36.380	20.544	40.420
13	30.975	22.244	30.187
14	17.965	21.086	23.274
15	34.059	21.683	29.914
16	2.577	5.304	7.360

结果可知，对于发动机故障分类而言，所采用的特征数目并不是越多越好，由于发动机性能参数的内在联系，它们所包含的分类信息存在冗余，去掉这些冗余信息，能使分类器实现更好的分类效果。方法2由于比较全面地考虑了特征之间的独立性，不但特征数目少，类间夹角大，而且原来类内分散较大的，还变得紧凑了，增大了类间的可分性，因而效果要好得多。利用选择出来的特征子集形成的十六种故障的模板，对仿真出来的每个故障四十个样本，共640个样本的训练集能够如同采用原始特征集一样实现全部正确分类。对另外每个故障五个样本的测试集同样能实现正确分类。详细结果不再一一列出。

4 结 论

原始特征集中的某些特征之间，由于发动机系统部件功能上的内在联系而存在强烈的相关性，因而存在大量的信息冗余，从而影响所设计的分类器性能。因此采用基于分类目的特征选择方法实现充分利用原始特征集中的有效的故障分类信息，以提高分类器的性能是很必要的。本文采用的两种特征选择方法可以较好的实现这一目的。同时由于这些方法的对象就是实际的发动机的性能参数，如果依据所选择的结果来确定发动机监测参数，就能更好、更直接地获取故障分类信息，以便于事后的发动机故障分析。

参 考 文 献

1 刘冰，张育林. 液体火箭发动机故障诊断中的特征提取. 推进技术，1997，18（6）
2 Bati Roberto. Using mutual information for selecting features in supervised neural netlearning. IEEE TRANSACTIONS ON NEURAL NETWORK 1994，5（4）
3 刘伟权，王明会，钟义信. 利用遗传算法实现手写体数字识别中的特征维数的压缩. 模式识别与人工智能，1996，9（1）

整体式固体火箭冲压发动机研制^{*}

张家骅 胡顺楠 顾炎武 曹中俄

(航天工业总公司31所, 北京, 100074)

摘 要: 概要介绍了我国第一种整体式固体火箭冲压发动机试验研究样机的研制, 讨论了样机地面试验研究和飞行试验研究两大阶段中, 整机(含主级和助推级)、各部件及分系统的设计、试制和试验工作。样机结构合理, 主级比冲比国外同类型号导弹发动机有较大提高, 助推级综合性能和热防护技术也优于后者。

主题词: 整体式火箭冲压发动机, 发动机设计, 推进系统试验

分类号: V235.21

DEVELOPMENT OF AN INTEGRAL SOLID ROCKET-RAMJET ENGINE

Zhang Jiahua Hu Shunnan Gu Yanwu Cao Zhong'e

(The 31st Research Inst., Beijing, 100074)

Abstract: An integral solid rocket-ramjet engine in China is introduced. The design, manufacture and test of the whole engine (both sustainer and booster), components and sub-systems during ground and flight test phases were described. The test results show that this prototype was reasonable, the specific impulse of cruise-stage was higher than that similar to engines abroad, and the integrated booster performance and thermal protection technique were satisfactory.

Subject terms: Integral rocket-ramjet, Engine design, Propulsion system test

1 引 言

整体式固体火箭冲压发动机(简称固冲组合发动机)是冲压和火箭技术组合的一种二级组合发动机。它由贫氧固体推进剂燃气发生器(贫氧火箭)作为主级的燃料源及一次燃烧室;助推火箭发动机和主级冲压补燃共用一个燃烧室(称为助推补燃室),由转级系统实现转级过程;发动机和弹体一体化,常采用旁侧进气道。这种发动机的比冲高于火箭发动机,能在较长时间内提供超声续航推力,系统简单、工作可靠和使用维护方便。

70年代以来,国外十分重视这种发动机的发展,其技术日趋成熟。前苏联首先采用这种发动机,并已用于实战。西欧和美国的固冲导弹正从预研转向型号研制。80年代初,在预先研制的基础上,研制了样机。在整机、各部件及分系统的设计、试制和试验诸方面作了大量工作,经地面试验和飞行试验,发动机主级比冲比国外同类型号导弹发动机有较大提高,助推级综合性能和热防护技术也优于后者,具有良好的飞行性能。

2 发动机设计

样机由进气道、燃气发生器、助推补燃室、组合喷管、点火系统及转级控制系统组成。它是一

^{*} 收稿日期: 1997-08-02, 本项目获部科技进步一等奖和国家科技进步三等奖