

模式识别技术在机械设备故障 诊断中的应用*

颜玉玲 徐尹格

(南京航空学院) (北京 交通管理干部学院)

摘 要

本文从统计意义上分析了机械设备在不同运行状态下振动信号的特征;选择信号幅值的概率分布的前 n 阶矩作为特征向量来进行状态信息凝聚。在此,将设备的状态分为“完好”和“故障”两种,应用模式识别技术进行状态分类。最后通过对试验数据的分析,证实了特征参数的稳定性和对故障的敏感性。结果表明,分类判据是有效的。

关键词 模式识别 状态信息凝聚 散度指标 类内距离 类间距离

一、引 言

设备故障的识别实质上是状态分类问题,它和模式识别具有密切的联系。选择什么特征参量来凝聚状态信息,以及采用什么识别准则是正确地进行分类的关键。

通常,设备在发生故障前,都具有一些前兆,如振动强烈、温度升高等。在机器运行中,振动是普遍存在的,它作为一种信息载体,直接预示着机器运行的正常与否。振动监测是目前使用最多,也是最有效的方法之一。

由于振动信号的随机性,研究其统计特征是有效的。在正常运行状态下,信号幅值应遵循高斯分布;而当故障出现时,就会产生强烈振动,使信号模式由平缓变化型向脉冲型过渡。信号幅值的分布情况将随之而变化,尤其表现在概率密度函数的尾部形状变化。一般常用一些无量纲因子,如峭度因子(四阶分布矩用二阶矩进行规一化所得)来描述这种变化,并作为故障识别的指标。许多实验证明,仅用单一特征作为识别指标是不可靠的。可以预见,应用模式识别技术,综合各状态特征参数进行故障诊断是发展所趋。

二、随机变量的特征函数

设 ξ 是定义在概率空间的连续型随机变量,其分布密度函数为 $p(x)$, ξ 的特征函数定义为:

* 1989年6月10日收到。1988年6月10日第一次收到。
MMM-Ⅱ会议论文,庐山,1989。

$$f_{\xi}(t) = E(e^{jt\xi})$$

根据复随机变量的数学期望定义, 有

$$\begin{aligned} f_{\xi}(t) &= E(\cos t\xi) + jE(\sin t\xi) \\ &= \int_{-\infty}^{+\infty} \cos tx \cdot p(x) \cdot dx + j \int_{-\infty}^{+\infty} \sin tx \cdot p(x) dx \\ &= \int_{-\infty}^{+\infty} e^{jt\xi} \cdot p(x) \cdot dx \end{aligned}$$

易知, $f_{\xi}(t)$ 为 $p(x)$ 的付里叶积分。式中, t 具有频率的量纲。所以, $f_{\xi}(t)$ 表示了随机变量的概率密度分布在频域的结构。不难证得, $f_{\xi}(t)$ 与 ξ 的分布矩之间具有如下关系

$$f_{\xi}^{(k)}(0) = j^k \cdot E\xi^k \quad (k \leq n)$$

上式成立的前提是 ξ 的 n 阶矩存在 ($n=1, 2, 3, \dots, n$)。将 $f_{\xi}(t)$ 在零点展开为泰勒级数可得:

$$f_{\xi}(t) = \operatorname{Re}[f_{\xi}(t)] + j\operatorname{Im}[f_{\xi}(t)]$$

$$\operatorname{Re}[f_{\xi}(t)] = 1 - \frac{1}{2!} E(\xi^2)t^2 + \frac{1}{4!} E(\xi^4)t^4 + \dots + \frac{(-1)^{i_1}}{(2i_1)!} E(\xi^{2i_1})t^{2i_1}$$

$$\begin{aligned} \operatorname{Im}[f_{\xi}(t)] &= E(\xi) \cdot t - \frac{1}{3!} E(\xi^3) \cdot t^3 + \frac{1}{5!} E(\xi^5) \cdot t^5 + \dots \\ &\quad + \frac{(-1)^{i_2-1}}{(2i_2-1)!} E(\xi^{2i_2-1})t^{2i_2-1} \end{aligned}$$

$$\text{其中, } \begin{cases} i_1 = \operatorname{Int}(n/2), \\ i_2 = \operatorname{Int}((n+1)/2) \end{cases} \quad \operatorname{Int}(\cdot) \text{ 表示取整}$$

从式中可见, 特征函数 $f_{\xi}(t)$ 由 ξ 的各阶矩 $E(\xi^n)$ 唯一确定, 反之亦然。

三、状态的分类识别

我们以“正常”和“故障”表示设备的两个状态, 故障诊断的问题即为两类识别问题。分类的第一步是选择适当的状态特征参数。一般, 所选取的特征参数应具备下列条件:

1. 易于度量, 提取和存贮,
2. 稳定。表现在类内距离小, 类间距离大, 即具有一定的散度指标。

综上, 通过识别幅值域分布的变化可以判断设备的运行状态正常与否。为了使识别定量化, 在此, 我们以特征函数的实部特征——前 $2i$ 阶分布矩的综合, 作为状态的特征向量, 并考虑到量纲的统一, 取为:

$$\mathbf{V} = \{(Ex^2)^{1/2}, (Ex^4)^{1/4}, (Ex^6)^{1/6}, \dots, (Ex^{2i})^{1/2i}\}$$

在选取特征向量构成待识别的模式后, 应给出与参考模式间的差异程度。对于同样性质的模式, 应存在一个紧密性假设, 即如果两个模式是相似的, 它们的特征向量在向量空间必然相近的。表征这种近似程度的度量有 Minkowski 距离, Euclidean 距离是它的一个特例, 定义为:

$$d[\mathbf{V}_i, \mathbf{V}_j] = \|\mathbf{V}_i - \mathbf{V}_j\|_2 = \sqrt{(\mathbf{V}_i - \mathbf{V}_j)^T \cdot (\mathbf{V}_i - \mathbf{V}_j)}$$

式中, $\|\cdot\|_2$ 表示“2”范数, $\mathbf{V}_i, \mathbf{V}_j$ 分别表示 i, j 状态的特征向量

根据最邻近判别原则, 本文给出下列判据:

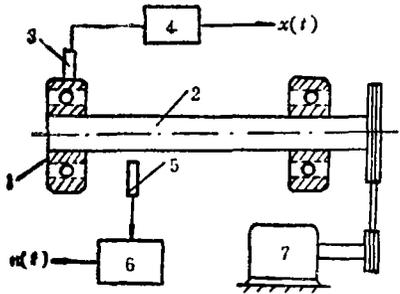
若 $d[V_k, V_i] > d[V_k, V_j]$, 则 $k \in i$

若 $d[V_k, V_j] > d[V_k, V_i]$, 则 $k \in j$

其中, V_k 为待识别状态 k 的特征向量, V_i, V_j 是供比较的参考样本, 分别表征状态 i 和状态 j .

四、试验数据分析

图1为轴承试验台装置, 加速度计3安装在轴承座上, 用来测量轴承振动的加速度信号。涡流传感器6用来测量转速信号, 将每次采样控制在主轴转动的整周期内进行。这样, 每次测得的数据所凝聚的信息具有等效性。分析过程见图2。在试验过程中, 我们分别测得完好和故障状态下六组信号数据(图3中绘出了其中一组测量信号), 分析结果见表1。表中特征参数 E_2, E_4, E_6, E_8 分别表示 $(Ex^2)^{1/2}, (Ex^4)^{1/4}, (Ex^6)^{1/6}$ 和 $(Ex^8)^{1/8}$ 。



1: 滚动轴承, 2: 主轴, 3: 加速度计 4: 电荷放大器, 5: 感应头, 6: 涡流传感器, 7: 电机, $n(t)$: 转速信号, $x(t)$: 加速度信号

图1 试验装置简图

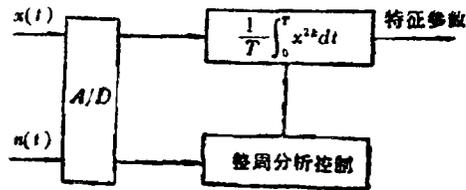


图2 分析过程示意图

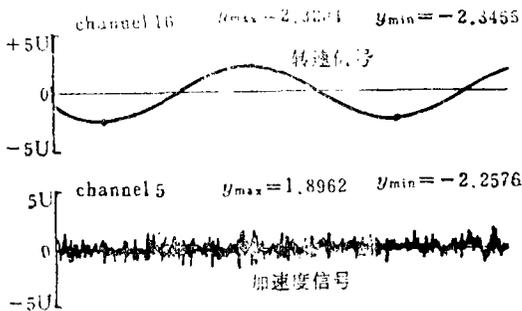


图3a 轴承完好时的信号

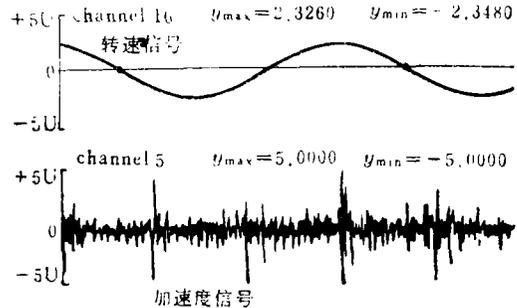


图3b 轴承出现故障时的信号

表1 分析结果

特征参数		测量序号						
		1	2	3	4	5	6	平均值
完好状态 (i)	E_2	0.6125	0.6368	0.5573	0.5233	0.5616	0.459	0.5584
	E_4	0.8059	0.8799	0.7461	0.717	0.7463	0.6381	0.7556
	E_6	0.9715	1.065	0.893	0.8472	0.8843	0.8059	0.9112
	E_8	1.121	1.203	1.005	0.9378	0.9932	0.9581	1.036
故障状态 (j)	E_2	1.090	1.148	1.085	1.030	1.171	1.090	1.102
	E_4	1.858	1.946	1.839	1.808	1.902	1.858	1.869
	E_6	2.515	2.564	2.464	2.450	2.471	2.515	2.497
	E_8	2.973	2.992	2.907	2.901	2.883	2.973	2.938

将 Euclidean 距离用特征维数 N 标准化, 记为

$$d_N[V_i, V_j] = N^{-1} \cdot \|V_i - V_j\|_2$$

在此, 特征维数 N 取为 4.

为了判断分类的好坏, 我们定义类内距离以及类间距离为:

$$\text{类内距离: } d(i) = E\{d_N[V_i^{(l)}, \mu_i]\}$$

$$\text{类间距离: } d(i, j) = \{d_N[\mu_i, \bar{\mu}] + d_N[\mu_j, \bar{\mu}]\} / 2$$

式中, i, j 表示类别; l 表示第 l 次测量; μ_i, μ_j 分别为状态 i 和状态 j 的特征向量的平均; $\bar{\mu}$ 表示 μ_i 和 μ_j 的平均.

$$\text{散度指标: } \text{div} = 2d(i, j) / (d(i) + d(j))$$

计算结果见表 2.

表 2 分类的散度指标

测 量 序 数		1	2	3	4	5	6	平均值
$d(i)(\times 10^{-3})$	完好状态	1.01	2.138	0.2941	1.014	0.4071	1.5977	1.077
$d(j)(\times 10^{-3})$	故障状态	0.3367	0.9827	0.450	0.8833	0.773	0.3367	0.6271
$d(i, j)$		0.6923						
div		812.511						

五、结 论

本文方法是针对诊断中常见的问题, 即采用单一参数作为判别故障的依据往往会误断或漏断而提出的。故障的诊断必须基于与正常状态的比较, 它实质上是模式分类问题。通常为两类识别(“正常”或“故障”), 也可以用多类识别给出更细的状态分类。在分类中, 选择适当的特征量非常关键。本文中选取了信号幅值的统计特征, 试验结果表明, 两状态间有较大的散度, 因而分类判据是有效的。特征量亦可以在频域或者其它域内选取, 有关研究有待进一步进行。

此外, 在采集样本信号或待识别信号时, 应保证每次测量的等效性。例如, 对于旋转机械, 信号具有周期特征, 应保证在主轴回转的整周期内采样、分析。

参 考 文 献

- [1] Julius, T. Tou, *Pattern Recognition Principles*, Addison-Wesley Pub. Comp. Inc., Massachusetts (1974).
- [2] 王飞龙, 《模式识别基础》, 湖北科技出版社 (1986).
- [3] Kashyap, R. L., Optimal feature selection and decision rules in classification problems with time series, *IEEE Trans. Inform. Theory*, **IT-24** (3) (1978), 281—288.
- [4] 中山大学数学力学系, 《概率论与数理统计》(上), 人民教育出版社 (1980).

The Application of Pattern Recognition Techniques in Fault Diagnosis of Machinery Equipment

Yan Yu-ling

(Nanjing Aeronautical Institute, Nanjing)

Xu Yin-ge

(Beijing Jiaotong Manager College, Beijing)

Abstract

In this paper, the characteristics of vibration signal of machinery in different running conditions are statistically analysed, and some moments of statistical distribution of signals are selected as the eigenvector to condense the state information. Here, we divide the states of machinery into two: 'good' and 'faulty', the pattern recognition techniques are used to classify the running conditions of machinery. At the end of this paper, the authors present some test data, and from the results obtained, it's verified that the eigenvector selected is reliable and sensible to faults. And the results also show the effectiveness of classification rule.

Key words pattern recognition, condensed state information, divergence index, inter-object distance, intra-object distance