

基于非线性流形学习的喘振监测技术研究

张熠卓¹, 徐光华², 梁霖²

(1. 西安交通大学机械工程学院, 710049, 西安; 2. 西安交通大学机械制造系统工程国家重点实验室, 710049, 西安)

摘要: 为了提取压缩机喘振发作时表现出的非线性特性, 引入了一种新的喘振特征提取方法. 首先对原始信号进行多元统计分析, 构造高维特征空间, 然后利用局部切空间排列的流形学习方法提取出一维主流形, 进而通过主流形几何结构的变化来反映系统的非线性变化. 分析结果表明, 与相关积分方法相比, 该方法可以提前 1 s 识别出喘振特征, 并且能够降低误报率, 因此在喘振监测中具有良好的应用前景.

关键词: 喘振; 流形学习; 局部切空间排列

中图分类号: TP206.3 **文献标志码:** A **文章编号:** 0253-987X(2009)07-0044-05

Surge Monitoring Method Based on Nonlinear Manifold Learning

ZHANG Yizhuo¹, XU Guanghua², LIANG Lin²

(1. School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049, China; 2. State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: To extract the nonlinear characteristic of compressor surge, a new monitoring method based on manifold learning is proposed. High dimension space of raw pressure signals is constructed by multivariate statistical analysis, then the local target space alignment (LTSA) algorithm is employed for extracting one dimension principal manifold, to reflect the non-linear dynamic characteristics of compressor system. Experimental results and industrial measurements show that this approach, compared with the correlation integral method, forecasts surge about 1s in advance with lower false prediction rate.

Keywords: surge; manifold learning; local tangent space alignment

透平压缩机已被广泛应用于众多关系到国民经济命脉的基础制造业中. 喘振是透平压缩机不可避免的故障之一, 喘振一旦发生必将导致转子和固定部件的严重损坏. 喘振信号本质上是非平稳非线性信号, 用时频分析可以有效地捕获失稳过程中的非平稳信息, 但却不能反映系统的非线性演化. 文献[1-2]指出, 过渡阶段无论从时域幅度和频域能量都与平稳阶段十分接近, 因此直接从时域或频域判断喘振是十分困难的. 目前, 相关积分法是判断喘振的比较有效的方法^[3-5], 但却存在着计算复杂度高、对信号的野点较为敏感、对喘振先兆的辨识容易产生误报等缺点^[6]. 文献[7]将经验模态分解(Empirical

Model Decomposition, 简称 EMD)与相关积分方法相结合, 用于离心风机的喘振信号处理, 然而 EMD 方法所带来的端点效应和频率混叠等问题则会影响监测的结果.

近年来, 流形学习技术作为一种新的非线性维数约简方法, 在数据挖掘和特征提取等领域中被广泛应用^[8]. 本文将流形学习及非线性维数约简引入压缩机喘振信号的特征提取, 首先对原始信号的时域和频域指标进行多元统计, 构造出高维特征空间, 其次利用局部切空间排列(Local Tangent Space Alignment, 简称 LTSA)的流形学习算法对高维空间进行维数约简, 并提取出一维特征曲线. 根据一维特

征曲线几何结构的变化来判断喘振前兆,在此基础上提出了一种针对结构变化的余弦距离报警指标,经过实际测量数据的分析,验证了方法的有效性。

1 流形学习算法概述

1.1 流形学习

传统的维数约简方法如局部投影方法、奇异值分解等,其主要思路是将高维数据用线性的方法来逼近系统的空间流形。但是,工业实际应用中的数据一般为非线性的,这些方法的线性本质使其无法揭示复杂的非线性流形结构。流形学习旨在发现高维数据集分布的内在规律性,其基本思想是:高维观测空间中的点由少数独立变量的共同作用在观测空间张成一个流形,如果能有效地展开观测空间卷曲的流形或发现内在的主要变量,就可以对该数据集进行降维。

1.2 局部切空间排列算法

在众多的流形学习算法中,LTSA 在参数的鲁棒性和主流形重构方面有很多优点,因此本文选择该算法来提取喘振特征的变化趋势。LTSA 的基本思想是利用样本点的切空间来表示局部的几何性质,然后将这些局部空间排列起来构造流形的全局坐标^[9]。对一个给定的高维数据集 $X=[x_1, x_2, \dots, x_N]$, $x_i \in \mathbf{R}^m$, 从中提取出一个 d 维 ($m > d$) 的主流形,LTSA 特征提取算法的步骤如下。

步骤 1 提取局部信息:对每个数据点 $x_i, i=1, 2, \dots, N$, 用欧式距离确定其 k_i 个邻近点组成的邻域 $X_i=[x_{i_1}, x_{i_2}, x_{i_3}, \dots, x_{i_{k_i}}]$ 。

步骤 2 局部线性拟合:在数据点 x_i 的邻域内选择一组正交基 Q_i 构成 x_i 的 d 维切空间,计算邻域中每个点 $x_{ij} (j=1, 2, \dots, k_i)$ 到切空间上的正交投影 $\theta_j^{(i)}=Q_i^T(x_{ij}-\bar{x}_i)$ 。其中, $\bar{x}_i$ 为邻域数据的均值。 x_i 的邻域数据在切空间的投影所构成的局部坐标 $\Theta_i=[\theta_1^{(i)}, \theta_2^{(i)}, \dots, \theta_{k_i}^{(i)}]$ 描述了 x_i 邻域内最主要的几何结构信息。

步骤 3 局部坐标全局排列:设 x_i 由 Θ_i 转换得到的全局坐标为 $T_i=[t_{i_1}, t_{i_2}, \dots, t_{i_{k_i}}]$, 则邻域数据的转换误差为

$$E_i = T_i(I - (1/k)\mathbf{e}\mathbf{e}^T) - L_i\Theta_i \quad (1)$$

式中: I 为单位矩阵; $\mathbf{e}$ 为单位矢量; k 为邻域点数; L_i 为转换矩阵。为了找到 t_i 和 L_i 使转换误差最小,则

$$L_i = T_i(I - (1/k)\mathbf{e}\mathbf{e}^T)\Theta_i^+ \quad (2)$$

$$E_i = T_i(I - (1/k)\mathbf{e}\mathbf{e}^T)(I - \Theta_i^+\Theta_i) \quad (3)$$

式中: Θ_i^+ 是 Θ_i 的广义 Moor-Penrose 逆。设

$$B = SWW^TS^T \quad (4)$$

令 $S=[S_1, S_2, \dots, S_N]$, $TS_i=T_i$, S_i 为 0-1 选择矩阵, T 为全局坐标,其权值矩阵为

$$W = \text{diag}(W_1, W_2, \dots, W_N) \quad (5)$$

且有

$$W_i = (I - (1/k)\mathbf{e}\mathbf{e}^T)(I - \Theta_i^+\Theta_i) \quad (6)$$

引入约束

$$TT^T = I_d \quad (7)$$

步骤 4 低维流形特征的提取:由于 $\mathbf{e}$ 是矩阵 B 的 0 特征值对应的特征矢量,所以由 B 的第 2~第 $(d+1)$ 个最小的特征值对应的特征矢量所组成的矩阵就是所求的 T , T 即为描述了高维数据集 X 中非线性主流形的低维全局坐标映射。

2 算法参数的选取

为了满足实际工程的需要,本文选择在时频域中较为常见的特征参数来覆盖信号在时域和频域内的完备特征^[10],然后利用基于 LTSA 的流形学习算法对高维空间进行维数约简,以反映喘振发展的变化情况。

2.1 高维特征集的选择

本文分别选取 8 个时域统计指标和 8 个频域指标来构造高维特征空间(见表 1),其中时域统计指标分别选取均值、最大峰值、均方根、根均方、方差、峭度、峰值因子和形状因子,频域选取分频段的频带能量作为统计指标,设 F_s 为采样频率,每一频段的宽度为 $F_s/16$, 则频带能量

$$F_i = \int_{(i-1)F_s/16}^{iF_s/16} s(f)df / \int_0^{F_s/2} s(f)df \quad (8)$$

表 1 特征维数和统计特征的选取

特征维数	时域统计指标	特征维数	频域统计指标
1	均值	9	F_1
2	最大峰值	10	F_2
3	均方根	11	F_3
4	根均方	12	F_4
5	方差	13	F_5
6	峭度	14	F_6
7	峰值因子	15	F_7
8	形状因子	16	F_8

式中: $s(f)$ 为时域信号经快速傅里叶变换所得到的频谱; i 为频段序号。

在实际测量数据中,如果截取一段长度为 t_1 的信号,特征空间的构造方法先是利用延迟重构方法将数据划分为 N 个子数据向量,再分别计算每个子数据向量的16维统计指标,这样就构成了一个 $16 \times N$ 的矩阵,其中 N 称为特征空间的样本点数。

2.2 邻域点数选取

邻域点数的选取是流形学习方法面对的一个主要问题,当 k 值过大时,会使局部线性化的程度不够,损失了局部几何结构,而 k 过小时则会使切空间非满秩,计算结果出错。本文选用的邻域点数的选取方式对应每个样本点,选择欧氏距离下最近的 k 个点作为邻域点。根据参考文献[11]中的取值, $k_{\min} = m_{\min}$,即邻域点数不得小于系统的维数。对于一个16维的特征空间,本文选取的 k 值略大于空间维数,使得系统在满足参数局部线性化的条件下,同时减少计算的复杂度。

2.3 滑动时间窗参数的选取

为了反映压力信号从平稳阶段到喘振完全发作的变化情况,需要采用滑动时间窗的方式来动态地计算某一时间段内的主流形结构。当 t_1 较短时,LTSA计算结果的差异性就较小,因而无法判别奇异性;当 t_1 较大时,机器计算所花费的时间会很长,不能满足实时监测的需求。从实测数据截取一段长度为20 s的数据段,计算不同 t_1 的数据段所对应的约简特征集并记录所需的时间。在主频为2.8 GHz、内存为2 GB容量的计算机上测试的结果如图1所示。从图中可以得到,当 $t_1=4$ s时,计算所需时间为0.489 s,本文将 t_1 选择为4 s,滑动步长为0.5 s,因此在满足实时性要求的条件下,可以有效地反映系统的变化情况。

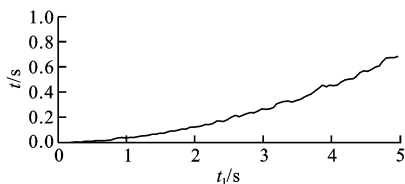


图1 维数约简时间随数据长度的变化曲线

3 特征提取效果对比

3.1 离线分析对比

实验数据取自某台250 kW的单级离心压气机,设 $F_s=150$ Hz,分别采用LTSA方法和相关积分方法对比分析喘振发生的动态信号数据。以一组

缓慢节流数据的分析为例,如图2所示,选择 $k=20$,可以看出相关积分曲线 c 在平稳工况下呈波浪状,并在约8.5 s处开始迅速下降,直到约10 s左右时下降停止,此时喘振已经完全发作,而与其相比,一维特征曲线 T_1 在平稳阶段更加平滑,并且在约7.5 s左右出现缓慢上升的趋势,由于平稳工况下流形非常稳定,所以微变量可以被识别出来。对于压缩系统,容腔内的能量由初始稳定阶段积累到诱发喘振,其值是非线性增长的, T_1 曲线反映出了系统由平稳到喘振完全发作过程的非线性变化,因此本文选择 T_1 作为喘振发展的特征函数。与相关积分法相比,LTSA方法可以提前1 s识别出喘振先兆。

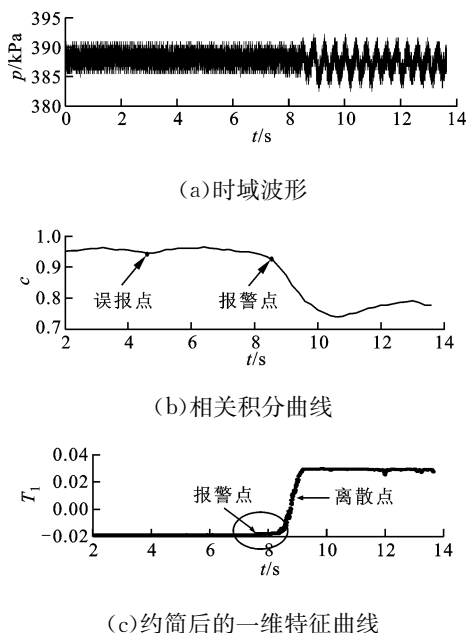
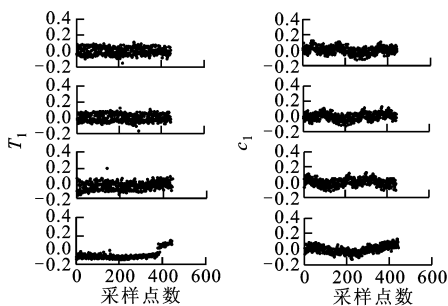


图2 缓慢节流压力信号的时域波形和相关积分曲线

3.2 在线分析对比

利用LTSA计算缓慢节流的压力信号,从平稳工况到喘振发作连续4个时间段的 T_1 值的变化情况如图3a所示。在相对应的时间段中,利用核主分量分析方法(KPCA)计算得到的一维主分量 c_1 ,其时间段顺序由上至下,如图3b所示。在图3a中,前3个时间段 T_1 趋于稳定,而第4个时间段内流形有增长趋势并且出现明显的断层,此时系统已表现出喘振先兆,而相对应的KPCA分析结果却无法识别。另外,对比图3b的前3个时间段,平稳阶段的 c_1 曲线在结构上有较明显的上下波动,聚类效果不如LTSA方法,容易出现误报。由此可见,对比传统的非线性降维方法,LTSA方法比KPCA更加有效。



(a)LTSA方法 (b)KPCA方法

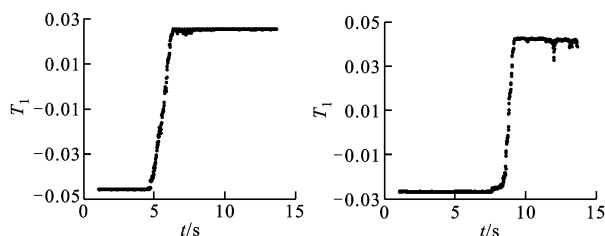
图3 不同时间段2种方法的分析效果

4 预警方法

对于不同的节流类型,利用 LTSA 方法计算出的 T_1 曲线的幅值变化有可能具有差异,但是在喘振发展的各个阶段,曲线所表现出的几何结构变化情况却是相似的,如图 4 所示。根据这一结果,本文提出利用余弦距离作为监测喘振的度量指标,即

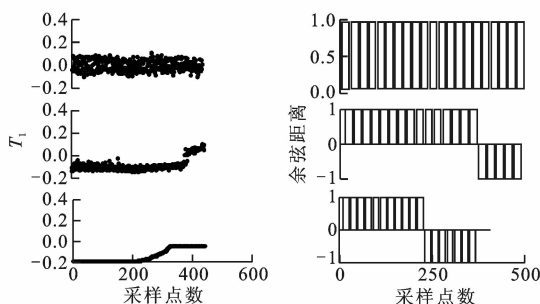
$$d_c(u, v) = \frac{u \cdot v}{\|u\| \|v\|} \quad (7)$$

式中: u, v 为特征向量,向量的余弦距离可以反映出高维空间中任意 2 个向量在方向上的差异。选择 T_1 稳定阶段的一组 20 个样本点的平均值作为参考向量,图 5 为过渡阶段和喘振发作阶段的余弦指标条形图。其中,每一个样条代表 20 个样本点,平稳阶段的余弦距离均为 1,而过渡阶段则出现了 -1 的值,



(a)缓慢节流 (b)快速节流

图4 不同节流方式的一维特征曲线结构



(a)缓慢节流 (b)快速节流

图5 喘振发作不同阶段的指标差异

发作阶段出现的值为 0。可以看出,余弦距离计算简单且易于实现,它作为监测指标十分直观,易于辨识,可早期发现喘振征兆。

5 结论

为了避免喘振误报现象,本文构建了一种新的监测方法,在构造出原始时间序列高维特征空间的基础上,利用非线性流形学习方法提取出一维主流形,从而反映了系统的非线性变化情况。实际分析结果表明,与相关积分法相比,本文方法的喘振先兆识别时间可以提前约 1 s,因此降低了误报率。在离线分析和在线分析的基础上,利用余弦距离可以有效地识别喘振发展的不同阶段。

参考文献:

- [1] 李美林,张春梅,王尚锦,等. 用小波和分维数方法处理喘振实验数据 [J]. 工程热物理学报,2002,23(2): 175-178.
LI Meilin, ZHANG Chunmei, WANG Shangjin, et al. Surge data processing based on wavelet analysis and fractal methods [J]. Journal of Engineering Thermophysics, 2002, 23(2): 175-178.
- [2] 蒋峻. 离心风机喘振在线预报和自适应调节的实验研究 [D]. 北京:中科院工程热物理所,2001.
- [3] BRIGHT M M, QAMMER H K, WEIG H J, et al. Stall precursor identification in high-speed compressor stages using chaotic time series analysis methods [J]. ASME Journal of Turbomachinery, 1997, 119(3): 491-499.
- [4] BRIGHT M M, HELEN K Q, WANG Leizhen. Investigation of pre-stall mode and pip inception in high speed compression through the use of correlation integral [J]. AEME Journal of Turbomachinery, 1999, 121(4): 743-750.
- [5] 吴云,李应红,张朴. 某涡喷发动机压气机气动失稳过程的非线性分析 [J]. 航空动力学报,2005,20(5): 890-894.
WU Yun, LI Yinghong, ZHANG Pu. Nonlinear analysis of the process of loosing compressor aerodynamic stability in a turbojet engine [J]. Journal of Aerospace Power, 2005, 20(5): 890-894.
- [6] 孙涛,徐光华,张春梅. 基于喘振频域特性的失稳预警技术研究 [J]. 西安交通大学学报,2007,41(11): 1321-1325.
SUN Tao, XU Guanghua, ZHANG Chunmei. Instability early warning based on frequency-domain characteristic of surge [J]. Journal of Xi'an Jiaotong Univer-

- sity, 2007, 41(11): 1321-1325.
- [7] 孙涛, 徐光华, 张四聪, 等. 经验模式分解法结合相关积分法的离心风机喘振先兆辨识 [J]. 动力工程, 2006, 26(6): 849-853.
- SUN Tao, XU Guanghua, ZHANG Sicong, et al. Surge symptom identification of centrifugal blowers by EMD in connection with correlation integrals [J]. Journal of Power Engineering, 2006, 26(6): 849-853.
- [8] 阳建宏, 徐金梧, 杨德斌, 等. 基于主流形识别的非线性时间序列降噪方法及其在故障诊断中的应用 [J]. 机械工程学报, 2006, 42(8): 154-158.
- YANG Jianhong, XU Jinwu, YANG Debin, et al. Noise reduction method for nonlinear time series based on principal manifold learning and its application to fault diagnosis [J]. Chinese Journal of Mechanical Engineering, 2006, 42(8): 154-158.
- [9] ZHANG Zhenyue, ZHA Hongyuan. Principal manifold and nonlinear dimension reduction via local tangent space alignment [J]. SIAM Journal on Scientific Computing, 2004, 26(1): 313-338.
- [10] HE Qingbo, KONG Fanrang, YAN Ruqiang. Subspace based gearbox condition monitoring by kernel principal component analysis [J]. Mechanical Systems and Signal Processing, 2007, 21(4): 1755-1772.
- [11] 阳建宏, 徐金梧, 杨德斌, 等. 邻域自适应选取的局部投影非线性降噪方法 [J]. 振动与冲击, 2006, 25(4): 64-67.
- YANG Jianhong, XU Jinwu, YANG Debin, et al. Nonlinear noise reduction method by local projection with adaptive neighborhood selection [J]. Journal of Vibration and Shock, 2006, 25(4): 64-67.

(编辑 管咏梅)