

独立分量分析基网络应用于旋转机械故障特征抽取与分类*

杨世锡 焦卫东 吴昭同
(浙江大学机械工程及自动化系 杭州 310027)

摘要: 提出了一种新颖的、基于独立分量分析(ICA)的多层神经网络,用于旋转机械不同模式(如正常及轴承故障等)的特征抽取,随后利用多层感知器(MLP)实施最终的模式分类。借助独立分量分析方法,隐藏于多通道振动观测中的不变特征得到有效提取,从而建立起稳定的MLP分类器。试验所获得的成功分类结果表明,所建议的新的旋转机械健康状况监测方法具有较大的应用潜力。

关键词: 独立分量分析 互信息 主分量分析 多层感知器

中图分类号: TN912.3

0 前言

电动机、发电机及透平等均属旋转机械,作为很多工业部门的关键设备,其状态监测与故障诊断工作受到相当的重视^[1]。理论上,机器内部任何力与(或)运动(例如由转子不平衡及轴承等故障所导致的碰撞冲击引起)的波动都可以视为一个有意义的源^[2],即本底振源,其中包含着故障诊断的有用信息。本质上,旋转机械状态监测与故障诊断可视为一种特殊的模式识别,其中涉及到从机器不同运行模式(由传感器所测的多维振动表达)中有效地提取特征。但是,与故障直接相关的机器振动,常由于机器自身的结构振动以及来自周围干扰机器的噪声而受到破坏。另外,由于传感测量中的空间冗余与分集效应,置放于机器外壳的振动传感器,其所测的信号只能是多个本底振源的混合,这使得许多与故障有关的谱峰同时出现于多个传感观测之中^[3],从而导致故障分类中有效特征抽取的困难。主分量分析在特征抽取中有一定应用^[4]。PCA只利用信号的二阶统计信息,在正交基约束下对数据加以描述,这对于高斯型分布的数据分析来说一般是足够了。不过在旋转机械中,初始故障在时间序列上往往表现出较强的脉冲冲击特性。例如,轴承故障以及基座松动故障会引发振动观测的非高斯特性(即峭度增加),这同时产生了很多外加的高阶统计信息^[5]。因此,机械故障诊断领域迫切需要新的数据分析技术与方法,以提高故障诊断的准确性。

近年来,独立分量分析方法及其在语音、图像和生物医学等信号的源分离及特征抽取方面的应用引起人们很大兴趣^[6~11]。特别令人感兴趣的是它在旋转机械源分离方面的应用^[3~5]。事实上,ICA也是一种强有力的数据分析、特别是非高斯数据分析工具。作为标准PCA的高阶扩展,ICA在无正交限制下抽取信号的统计独立分量。从统计意义上看,ICA相当于一种冗余取消技术,它使得各输出之间的互信息(MI)——一个基于信号所有高阶统计量构建的统计测度为零^[11],这使得ICA在非高斯数据分析方面具有独特优势。试验结果表明,ICA在旋转机械健康状况自动监测中颇具潜力。

1 信息最大化 ICA 算法

ICA通常假设:观测信号(这里指多维振动测量)由几个本底源经线性混叠而成^[12]。在源信号 s 统计独立的假设下,ICA寻求线性变换 W ,使得

$$y = \hat{s} = Wx \quad (1)$$

这里观测矢量 x 已知,而其矢量矩阵 W (源分离问题中称为分离矩阵)和坐标系 s (源分离问题中称为源矢量)均未知。估计过程中,需要大量已知类别的观测样本 x 。作为基于所有高阶统计量的统计独立测度,互信息

$$I(s) = \int f(s) \log \frac{f(s)}{\prod_i f_i(s_i)} ds \quad (2)$$

或它的某些变形被用作估计问题的准则函数。这里 $f(s)$ 为联合概率分布,通常具有较高维数。边缘分布 $f_i(s_i)$ 与基矢量 b_i (形成基矢量矩阵 W)有关。当

* 国家自然科学基金(50205025)和浙江省自然科学基金(5001004)资助项目。20030324日收到初稿,20030923收到修改稿

$f(s) = \prod_i f_i(s_i)$ ，即各坐标分量 s_i 统计独立时，等式(2)取最小值。等式(1)有解的条件是数据分布必须是非高斯的(在高斯分布情形，PCA 可给出所需要的解)。的确，线性方法(1)往往导致最大化恢复分量非高斯性的算法，亦即 ICA 使得观测数据在 ICA 基矢量上的投影，尽可能的远离标准高斯分布。如果一个数据集中源独立假设不成立，此时 ICA 等价于一种探测投影追求方法(EPP)，它通过构造线性投影的非线性函数，探测观测空间中“有趣的”数据结构，同时实现数据维数削减。

此处要解决的基本问题是：对输入 x ，如何使得一个神经网络的输出 y 的互信息最小？定义

$$I(y, x) = H(y) - H(y|x) \quad (3)$$

式中 $H(y)$ ——神经网络的输出熵

$H(y|x)$ ——条件熵

Bell 和 Sejnowski^[12]提出一种 ICA 算法，该算法最大化线性变换的联合熵 $H[g(Wx)]$ ，其中所用的线性变换受制于一个神经网络激活函数 $g(\cdot)$ ，即欲抽取信号的累积量密度函数(c.d.f)。对于高峭度分布的数据，通过如下对数激活函数的采用，即使不知道待抽取信号的 c.d.f，亦可获得良好的源分离。

$$g(u) = [1 + \exp(-u)]^{-1} \quad (4)$$

$$u = Wx + w_0$$

式中 W, w_0 ——ICA 网络的权矩阵和偏移矢量

通过权矩阵的渐进更新，输出熵得以最大化，从而产生如下的学习规则

$$\Delta W \propto [W^T]^{-1} + (1 - 2y)x^T \quad (5)$$

$$\Delta w_0 \propto 1 - 2y \quad (6)$$

对于单个权参数 w_{ij} ，学习规则可写作

$$\Delta w_{ij} \propto \frac{\text{cofw}_{ij}}{\det W} + x_j(1 - 2y_i) \quad (7)$$

cofw_{ij} 为 w_{ij} 的代数余子式。

2 故障特征抽取与模式分类

新的故障分类系统的整体架构如图 1 所示。它在原理上与参考文献[11]中的方法很相似，该方法用于医学上分析心纤维颤震的电描记图以诊断心脏疾病。研究中发现，当属于某一类的多维振动观测信号进入 ICA 网络，而这个网络已经过来自于同一模式类的观测样本加以训练，此时网络输出间的残余相关，较之于通过其他 ICA 网络(经其他不同类的样本训练)所获得的残余相关更小。试验中，分别使用了两种不同模式类，即潜水泵正常运行及轴承故障状态的振动观测。从每一种观测中选取数据样本

构建训练集，用于批训练 ICA 网络，一种模式样本训练一个 ICA 网络。当一个数据段输入时，两个 ICA 网络输出的“总体相关”形成该数据段的特征矢量，该矢量随后用于训练 MLP 模式分类器。

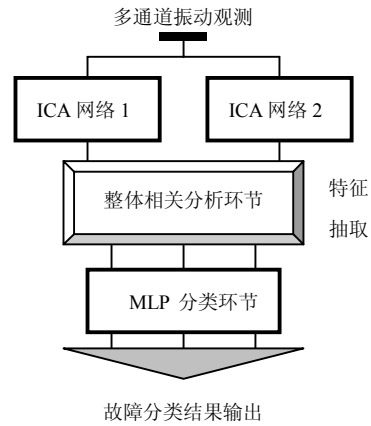


图 1 复合 ICA-MLP 旋转机械故障诊断系统

根据参考文献[11]，一个 ICA 网络输出的总体相关 TC 可定义如下

$$TC = \sqrt{\sum_{i,j=1}^N CC_{ij}^2} \quad (8)$$

式中 CC_{ij} ——ICA 网络的第 i 个与第 j 个输出通道间的相关参数

为测试这一算法，将测试集中的每一数据段分别作为两个 ICA 网络的输入，MLP 的输出指出哪一个 ICA 网络输出的残余相关最小，结果反过来表明输入振动样本所属的类别。ICA 特征矢量抽取的整个过程可表示为如图 2 所示。其中 \circ 和 \blacksquare 分别代表用于寻找数据独立特征的训练集及经过训练的 ICA 网络。 x 为待分类的数据段。最后，由每一个数据段，借助对应模式类样本训练的 ICA 网络，可获得一个两维的特征矢量 (f_1, f_2) 。

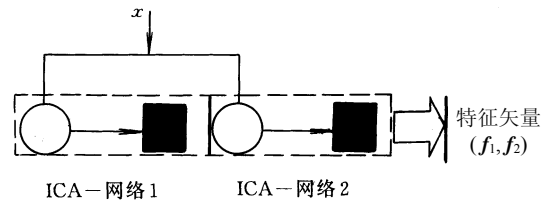


图 2 用多个 ICA 网络抽取旋转机械模式类的特征

3 试验验证

为全面清楚地评估多 ICA 网络的特征抽取能力，利用 5 通道潜水泵振动观测数据^[8]，对应两个模式类(即正常及轴承故障模式)，每一模式类的观测由 20 480 个数据采样组成。有关试验装置及数据

采集的详细信息，见参考文献[9]。利用信息最大化ICA方法^[11]及Netlab工具箱对数据进行分析。每种模式中的前512、1024、2048、4096、8192和10240个样本，分别用于构成6个独立的训练集，批训练ICA特征抽取网络，每类样本训练一个网络，得到如图3所示的训练结果。

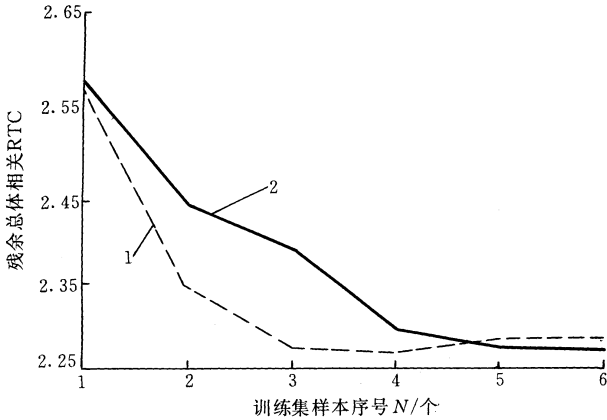
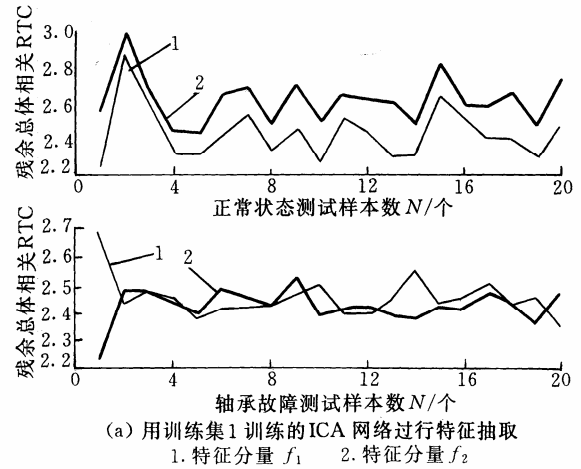


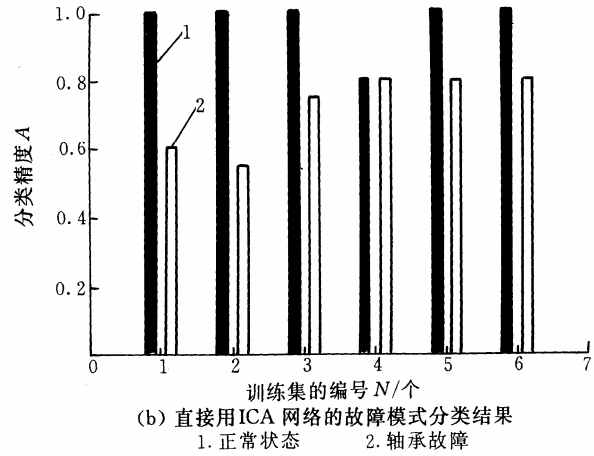
图3 用6个训练集训练ICA特征抽取网络的结果输出
1.正常状态 2.轴承故障

根据前述思想，测试样本进入已经过来自于同一模式类样本训练的ICA网络，网络输出间的残余总体相关较之与通过其他ICA网络(未经同类样本训练)所获得的残余相关更小。为实施特征抽取，将两模式观测的前10240个样本分割为20个数据段，每个数据段包含512个数据采样。将这些数据段输入已经训练的两个ICA网络，即可得图4a。从图4a中可以看到，当将20个数据段作为输入，通过已经512个样本(即图3中的训练集1)训练的两个ICA网络(正常和轴承故障)时，特征矢量的各分量大小基本呈现出前述的变化趋势，不过在图4a的第二个子图中标记为第2、5~12和20的数据段处还是出现了例外，此时特征分量 f_1 小于特征分量 f_2 ，这违背了前述的规律。从而，训练集1所得的ICA网络，对轴承故障样本的直接分类精度只能达到60%。另外，从1到6所有训练集训练的ICA网络的直接分类精度在图4b中给出。由图中很容易发现小训练样本集(如512和1024等)对ICA网络分类结果的影响，这进一步证实了采用足够样本训练ICA特征抽取网络的重要性。在理想情况下，ICA网络的输出应该只有两种模式，分别对应两个模式。但在实际中，某些可使得ICA网络输出残余相关最小的振动样本并未用于网络训练。而且，不同模式类间不可避免地存在着重叠。因此，即便ICA网络已经同类样本训练，当该故障类的测试集样本通过该网络时，也不能保

证总是获得具有最小残余相关的输出。所以，有必要采用一个多层感知器(MLP)网络进一步分类故障模式。



(a) 用训练集1训练的ICA网络过行特征抽取
1.特征分量 f_1 2.特征分量 f_2



(b) 直接用ICA网络的故障模式分类结果
1.正常状态 2.轴承故障

图4 ICA网络特征抽取与机械故障模式分类

将每种模式剩余的10240个样本分割为20个测试数据段，每段512个采样点，用于测试MLP分类器的分类能力。首先用训练集(由最初的10240个样本分割而成)抽取出的各特征矢量，采用均匀变化步长的梯度算法，训练一个MLP分类器。该MLP的隐层节点数为10，激活函数取为对数S函数(Log-Sigmoid)。复合ICA-MLP网络分类器的分类结果如图5和表所示。可以看到，当采用较大的样本集训练ICA特征抽取网络时，可以得到比较好的故障模式分类结果。ICA-MLP分类器的分类精度(图5)，与单纯的ICA网络直接分类精度(图4b)相比有比较明显的改进。

表 ICA-MLP网络的分类测试结果 %

	正 常	轴 承 故 障
训练集1训练的ICA-MLP	85	100
训练集2训练的ICA-MLP	70	90
训练集3训练的ICA-MLP	75	100
训练集4训练的ICA-MLP	80	85
训练集5训练的ICA-MLP	90	90
训练集6训练的ICA-MLP	95	100

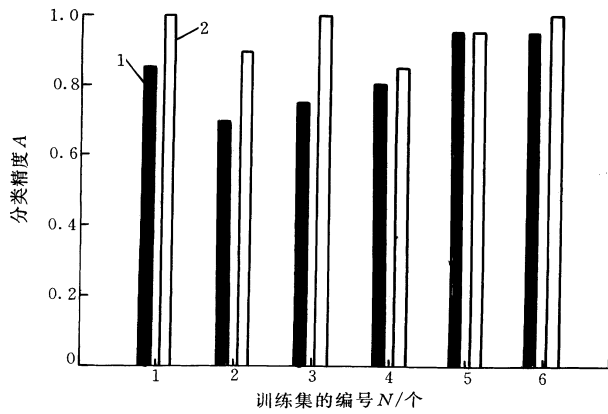


图5 用ICA-MLP网络进行模式分类

4 结论

用复合ICA-MLP网络,进行了旋转机械健康状况监测的应用研究。理论上,ICA通过寻找数据中有意义结构的方向,在数据的独立性与非高斯性限制下挖掘数据的高阶特征。结果表明:基于ICA特征抽取的MLP分类器可实现潜水泵正常及轴承故障模式的较高精度的识别,这意味着ICA在探求隐藏于数据中的非高斯性特征的能力。

正如式(1)所示,传统ICA方法假设观测为几个独立本底源的线性混叠,且无外加噪声。事实上,这些限制与实际情况并不总是相符。许多旋转机械故障(如不平衡、油膜涡动及转轴裂纹等)常表现为强烈的非线性振动特征^[1]。令人欣慰的是,有关非线性源分离及有噪模型的非线性ICA方法已经出现^[13-15],为增强ICA的实际应用能力,必须采用新的ICA非线性混叠模型。借助于适当的非线性ICA方法,有望获取旋转机械多维振动观测中更有效的故障判别特征,这将是未来的研究工作。

参 考 文 献

- 徐敏. 机械故障诊断手册. 西安: 西安交通大学出版社, 1998
- Yon R H. Machinery noise and diagnostics. Boston: Butterworths, 1987
- Alexander Y, Amir L, Robert P W D. Blind separation of rotating machine sources: bilinear forms and convolutive mixtures. Neurocomputing-Special Issue on ICA/BSS, 2002
- Gelle G, Colas M, Delaunay G. Blind sources separation applied to rotating machines monitoring by acoustical and vibrations analysis. Mechanical Systems and Signal Processing, 2000, 14(3): 427~442
- Alexander Y, David M J T, Robert P W D. Robust ma-

chine fault detection with independent component analysis and support vector data description. Pattern Recognition Group, Dept. of Applied Physics, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands, 1999

- Pierre C. Independent component analysis, a new concept. Signal Processing, 1994, 36: 287~314
- Porrill J, Stone J V, Berwick J, et al. Analysis of Optical Imaging Data Using Weak Models and ICA. Based on a workshop held after the 1999 International Conference on Artificial Neural Networks (ICANN), 1999
- This dataset was acquired in the Delft "Machine diagnostics by neural networks"-project with help from * Landustrie Sneek b.v., The Netherlands * SKF Condition Monitoring b.v., The Netherlands, and can be downloaded freely at the following web-address: <http://www.ph.tn.tudelft.nl/~ypma/mechanical.html> (c) 12-08-99. Pattern Recognition Group, Delft University of Technology, 1998
- Ronald L, Robert P W D, Eduard E E F, et al. Machine diagnostics by neural networks, Experimental Setup, Pattern recognition group, Computational physics group, Delft University of Technology, 2000
- Chitroub S, Houacine A, Sansal B. Compound PCA-ICA neural network model for enhancement and feature extraction of multi-frequency polarimetric SAR imagery. Signal Processing Laboratory, Electronics Institute, University of Sciences and Technology of Houari Boumedienne, 2000
- Govindan A, Deng G. Independent component analysis applied to Electrogram classification during Atrial Fibrillation. School of Electronic Engineering, La University, Melbourne Vic-3083, Australia, 2001
- Bell A, Sejnowski T. An information-maximization approach to blind separation and blind deconvolution. Neural Computation, 1995(7): 1 129~1 159
- Lee T W. Nonlinear approaches to independent component analysis. The Salk institute, Institute for Neural Computation, San Diego, USA, University of California, 2001
- Luis B Almeida. ICA of linear and nonlinear mixtures based on mutual information. IST and INESC-ID, Lisbon, Portugal, 2000
- Chakra C, Allan J. Sparse PCA extracting multi-scale structure from data. Department of Computer Science, University of Toronto, 6 King's College Road Toronto, ON M5S 3H5, Canada, 1998

INDEPENDENT COMPONENT ANALYSIS BASED NETWORKS FOR FAULT

FEATURES EXTRACTION AND CLASSIFICATION OF ROTATING MACHINES

Yang Shixi Jiao Weidong Wu Zhaotong
(Zhejiang University)

Abstract: A novel multi-layer neural networks is proposed, which is based on independent component analysis (ICA) for feature extraction of different modes (for example normal and bearing fault etc.), followed by a multi-layer perceptron (MLP) which implements the final classification. By the use of ICA,

invariable features embedded in multi-channel vibration measurements can be captured. Thus, stable MLP classifier is constructed. The successful results achieved by the selected experiments indicated great potential of the new method in health condition monitoring of rotating machines.

Key words: Independent component analysis
Mutual information
Principal component analysis
Multi-layer perceptron

作者简介: 杨世锡, 男, 1968 年出生, 博士, 副教授。主要研究方向为智能检测与机械故障诊断等方面的研究。

《机械工程学报》被 EI 检索目录 (2003 年第 3 期)

- 基于特征的层次式公差信息表示模型及其实现 刘玉生 高曙明 吴昭同等 (1)
- 机床直行部件运动误差的在线测量 苏 恒 洪迈生 魏元雷等 (8)
- 行星分度凸轮机构凸轮廓线的设计 王 皓 张 策 杨玉虎 (13)
- 半碗六电极悬吊球形转子的静电场干扰力矩分析 贺晓霞 高钟毓 吴秋平 (17)
- 机械运动方案设计的状态空间方法 王德伦 张德珍 马雅丽 (23)
- 单级齿轮系统拍击特性研究 张锁怀 沈允文 董海军等 (28)
- 基于瞬时频率估计的旋转机械阶比跟踪 郭 瑜 秦树人 汤宝平等 (32)
- 蠕变损伤本构方程的优化特性 李柏林 林建国 姚 新 (37)
- 弹性连杆机构结构和噪声控制一体化设计 卢剑伟 张宪民 沈允文 (40)
- 修正齿形的复合齿轮泵流量特性 许贤良 刘小华 朱玉川等 (44)
- 主轴回转误差补偿机理和动力学模型研究 杜正春 杨建国 姚振强等 (48)
- 冗余度机械臂关节运动和末端运动的同步容错规划 赵 京 缪 萍 荆红梅 (53)
- 地面效应对汽车外部流动的影响 谢今明 张扬军 涂尚荣 (58)
- 电磁成形中板料厚度与变形深度的关系 初红艳 费仁元 陆 辛等 (62)
- 旋转周期为曲柄二整周的平面双曲柄四杆机构 陈辛波 岩附信行 林 严等 (66)
- 离心泵无分离条件下叶片型线方程研究 朱玉才 梁 冰 张永利等 (71)
- 定流网络二次调节液压系统 顾临怡 曹建伟 邱敏秀 (76)
- 数控机床热误差补偿模型在线修正方法研究 杨建国 任永强 朱卫斌等 (81)
- 基于复合型紊流润滑理论的径向滑动轴承紊流润滑性能的研究 王小静 苏 荭 张直明 (85)
- 多轴随机载荷下断裂平面的确定方法 王 雷 王德俊 (90)
- 共注成形充模流动的理论模型和模拟 周国发 柳和生 何成宏等 (94)
- 三维转子密封系统气流激振的研究 何立东 高金吉 金 琰等 (100)
- 利用品质功能展开进行产品优化设计 唐加福 庞士宗 汪定伟等 (105)
- Hill 非二次屈服准则次数 m 值识别及应变控制研究 吴建军 李 祥 (110)
- 虚拟机床结构的反问题方法 佟 维 卢碧红 (113)
- 振动式微机械陀螺的等效电学模型 车录锋 黄小振 熊 斌等 (117)
- Al—Fe 二元合金共晶组织的数值模拟 李荣德 周振平 李述军 (121)
- 大型锻件中疏松缺陷压实条件研究 任运来 聂绍珉 陈 颖等 (126)
- 支承电压与电压变化速率对静电轴承性能影响的研究 韩丰田 高钟毓 王永梁 (130)
- 多模式失效的机械零件可靠度计算新理论 喻天翔 孙玉秋 张祖明 (134)
- 锥台形壳体液压胀形过程的数值模拟 詹艳然 张哲武 王仲仁 (139)
- 纳米压痕技术理论基础 黎 明 温诗铸 (142)
- 离心力场下铸造充型行为数值模拟 徐达鸣 安 蕴 李 鑫 (146)
- 齿条刃边测头在位测量大型齿轮齿形误差的基本原理及误差与变异分析 李文龙 商向东 金嘉琦等 (151)
- 小型液压缓冲器的动态特性分析 郝鹏飞 张锡文 何 枫 (155)