

# SDICA 方法在单通道信号故障分类中的研究<sup>\*</sup>

陈建国<sup>1</sup>, 王 珍<sup>1</sup>, 李宏坤<sup>2</sup>

(1. 大连大学机械工程学院 大连, 116622) (2. 大连理工大学振动工程研究所 大连, 116023)

**摘要** 提出了一种针对工程单通道信号的子带分解独立分量分析(subband decomposition independent component analysis, 简称 SDICA)故障分类方法。利用经验模态分解方法(empirical mode decomposition, 简称 EMD)得到的多个基本模式分量作为子带信号, 对子带信号进行独立分量分析(independent component analysis, 简称 ICA), 在 ICA 方法过程中提取了分离过程特征中产生的残余互信息值, 在估计子带信号中计算各自的近似熵值, 并把残余互信息和近似熵值作为特征参数, 输入广义回归神经网络实现故障分类。SDICA 方法在单通道信号故障分类中引入了 ICA 理论, 成功实现了工程单通道轴承信号 3 种故障高精度的识别, 验证了具有良好表征故障能力的残余互信息值和估计子带近似熵能够成为故障分类的重要参数。

**关键词** 故障分类; 特征提取; 广义回归神经网络; 子带分解独立分量分析

**中图分类号** TN912; TH133

## 引 言

工程振动信号大多都为单通道信号, 当故障较小或者干扰较大的情况下, 故障分类精确度不高。如果对故障信号分离得到故障成分丰富的故障源信号, 从而可以获得很高的分类精度。目前国内外针对单通道信号的盲源分离研究方法有动态嵌入方法构造延迟向量<sup>[1]</sup>、基函数子空间方法<sup>[2]</sup>、稀疏分量分析方法<sup>[3]</sup>等。前两种方法人为干预太多, 而稀疏分解高效成熟的算法不多, 并要求信号稀疏性较高。近年来国内外学者开始研究从 ICA 进一步扩展得到的 SDICA 方法, 并在图像和医学信号中得到成功的应用<sup>[4-7]</sup>。SDICA 对于单通道信号的故障源成分提取提供了一个思路。笔者利用 SDICA 方法在单通道信号故障分类方面开展研究, 通过统计分析, 选定基于 SDICA 的分离过程和分离结果的统计信息为特征向量, 利用广义回归神经网络获得精确的分类效果。

## 1 SDICA 方法

### 1.1 SDICA 理论

线性瞬时无噪 ICA 的模型可以描述为

$$\mathbf{X} = \mathbf{A}\mathbf{S} \quad (1)$$

其中:  $\mathbf{X}$  为  $m$  维的观测混合信号;  $\mathbf{S}$  为  $n$  维的源信号;  $\mathbf{A}_{m \times n}$  为混合矩阵。

若  $m < n$  时被称为欠定 ICA<sup>[8]</sup>。当  $m < n$  时, 从数学角度上由矩阵  $\mathbf{A}$  无法计算矩阵的逆矩阵  $\mathbf{W}$ , 只能把观测混合信号  $\mathbf{X}$  转化为一组子带  $\mathbf{X}_k(t)$ , 将  $\mathbf{X}_k(t)$  作为 ICA 算法的输入从而获取解混矩阵  $\mathbf{W}^{[5]}$

$$\mathbf{Y}_k(t) = \mathbf{W}\mathbf{X}_k(t) \quad (2)$$

其中:  $\mathbf{Y}_k(t)$  为估计源信号。

为了获得  $\mathbf{X}_k(t)$ , 构造一个分解或滤波算子  $\mathbf{T}_k$ , 用  $\mathbf{T}_k$  来提取子带成分

$$\mathbf{S}_k(t) = \mathbf{T}_k[\mathbf{S}(t)] \quad (3)$$

把式(2)、式(3)和  $\mathbf{T}_k$  带入式(1)可得

$$\mathbf{X}_k(t) = \mathbf{T}_k[\mathbf{A}\mathbf{S}(t)] = \mathbf{A}\mathbf{T}_k[\mathbf{S}(t)] = \mathbf{A}\mathbf{S}_k(t) \quad (4)$$

### 1.2 子带获取方法

获取子带的方法有固定滤波的方法<sup>[9]</sup>、自适应预滤波方法<sup>[4]</sup>、小波包分解方法<sup>[5-6]</sup>和线性滤波方法<sup>[10]</sup>等。根据对以上文献的研究发现, 子带获取方法应该根据信号先验知识、自身的特点、研究目的来恰当地选取子带。对于机械设备故障来说, 分析信号的目的是找出故障的特征信息。特别对于轴承故障, 若利用滤波器获取子带, 则需要知道故障特征频率、中心频率、频宽等难以获得的先验知识; 若利用

\* 国家自然科学基金资助项目(51175057); 辽宁省教育厅一般项目基金资助项目(L2013477)

收稿日期: 2015-03-17; 修回日期: 2015-09-06

小波分解获取子带,则小波基和分解层数难以确定。所以本研究采用 EMD 方法:a. EMD 方法是自适应性分解方法,能较好地保留原信号的信息成分;b. EMD 方法有处理非平稳信号的能力,可以更加有效地分解工程信号<sup>[11]</sup>。假定单通道观测信号  $\mathbf{X}$ ,经过 EMD 分解后得到

$$\mathbf{X} = \sum_{i=1}^N \mathbf{X}_{\text{IMF}_i} + \mathbf{r}_N \quad (5)$$

其中: $\mathbf{X}_{\text{IMF}_i}$  为第  $i$  个基本模式分量 (intrinsic mode function, 简称 IMF);  $\mathbf{r}_N$  为趋势项。

将  $\text{IMF}_i$  作为第  $i$  个子带信号。运用 SDICA 方法得到估计子带信号为

$$\hat{\mathbf{X}}_{\text{IMF}} = \{\hat{\mathbf{X}}_{\text{IMF}_1}, \hat{\mathbf{X}}_{\text{IMF}_2}, \dots, \hat{\mathbf{X}}_{\text{IMF}_N}\} = \mathbf{W}\mathbf{X}_{\text{IMF}} = \mathbf{W}\{\mathbf{X}_{\text{IMF}_1}, \mathbf{X}_{\text{IMF}_2}, \dots, \mathbf{X}_{\text{IMF}_N}\} \quad (6)$$

## 2 故障分类的特征参数选取

当单通道故障信号进行 SDICA 方法分析时,在相同初始值条件下,不同故障进行 SDICA 分离时的分离矩阵统计信息、分离算法的收敛性能应有所差异。SDICA 对单通道故障信号处理后,故障主要成分将集中到某个估计子带信号中,不同故障的估计子带的峭度、近似熵、时频能量等统计信息应有较大的差别。

### 2.1 分离过程特征信息

通过研究发现,SDICA 分离矩阵内数值波动较大,易受外部干扰,而收敛信息比较稳定,所以进一步研究表征分离算法收敛性能的估计子带互信息残差信息 (residual mutual information, 简称 RMI)<sup>[12]</sup>,记为  $\phi_{\text{RMI}}$ 。其具体公式为

$$\phi_{\text{RMI}} = \begin{bmatrix} \varphi_{11} & \varphi_{12} & \dots & \varphi_{1N} \\ \varphi_{21} & \varphi_{22} & \dots & \varphi_{2N} \\ \vdots & \vdots & & \vdots \\ \varphi_{N1} & \varphi_{N2} & \dots & \varphi_{NN} \end{bmatrix} \quad (7)$$

$$\varphi_{ij} \approx \frac{1}{4} \sum_{i \neq j} \text{cum}^2(\mathbf{x}_i, \mathbf{x}_j) + \frac{1}{12} \sum_{i \neq j} (\text{cum}^2(\mathbf{x}_i, \mathbf{x}_i, \mathbf{x}_j) + \text{cum}^2(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_j)) + \frac{1}{48} \sum_{i \neq j} (\text{cum}^2(\mathbf{x}_i, \mathbf{x}_i, \mathbf{x}_i, \mathbf{x}_j) + \text{cum}^2(\mathbf{x}_i, \mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_j) + \text{cum}^2(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_j, \mathbf{x}_j)) \quad (8)$$

其中:设估计子带信号  $\hat{\mathbf{X}}_{\text{IMF}_i} = \mathbf{x}_i$ ,  $\hat{\mathbf{X}}_{\text{IMF}_j} = \mathbf{x}_j$ ;  $\varphi_{ij}$  为  $\mathbf{x}_i, \mathbf{x}_j$  间的互信息值 (mutual information, 简称 MI);  $\text{cum}(\cdot)$  表示累积量。

### 2.2 估计子带特征信息

近似熵 (approximate entropy, 简称 ApEn) 从时间序列复杂性的角度上度量信号中产生新模式的概率大小。产生新模式的概率越大,序列的复杂性越大,相应的近似熵也越大<sup>[13]</sup>。不同故障信号的模式不同,那么近似熵的区分能力也应该较为明显。其计算方法如下。

设采集信号序列为  $\{\mathbf{x}(i), i=0, 1, \dots, N\}$ 。

1) 给定模式维数  $m$ , 原数据组成  $m$  维矢量

$$\mathbf{X}(i) = [\mathbf{x}(i), \mathbf{x}(i+1), \dots, \mathbf{x}(i+m-1)] \quad (9)$$

其中:  $i=1 \sim N-m+1$ 。

2) 定义  $\mathbf{X}(i), \mathbf{X}(j)$  之间的距离

$$d(\mathbf{X}(i), \mathbf{X}(j)) = \max_{k=0 \sim m-1} [|\mathbf{x}(i+k) - \mathbf{x}(j+k)|] \quad (10)$$

3) 给定阈值  $r$  ( $r=0.2 \sim 0.3$ ), 对每一个  $i$  值统计  $d(\mathbf{X}(i), \mathbf{X}(j))$  小于  $r$  的数目及此数目与距离总数  $N-m+1$  的比值, 记作  $c_i^m(r)$ 。

4) 对  $c_i^m(r)$  取对数, 再求其对所有的  $i$  的平均值, 记作

$$\Phi^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \ln c_i^m(r) \quad (11)$$

5) 把维数  $m$  加 1, 变成  $m+1$ , 重复 1~4 步, 得  $\Phi^{m+1}(r)$ 。

6) 理论上, 此序列的近似熵为

$$\text{ApEn}(m, r) = \lim_{N \rightarrow \infty} [\Phi^m(r) - \Phi^{m+1}(r)] \quad (12)$$

## 3 单通道轴承故障信号特征参数分析

### 3.1 特征参数计算

选用美国凯斯西储大学的滚动轴承故障模拟实验台的实验数据进行研究, 轴承位于驱动端, 型号为 SKF6205, SDICA 方法提取特征参数步骤如下:

1) 共选择 36 组信号, 1~12 组为内环故障信号, 13~24 为滚动体故障信号, 25~36 为外环故障信号;

2) 对每组信号 EMD 分解, 取前 5 个 IMF 子带信号, 利用 FastICA 算法进行振源分离, 得到 5 个估计 IMF 子带;

3) 根据 IMF 的频带宽度逐渐减小特性, 对估计 IMF 子带进行重新排序;

4) 计算 SDICA 残余互信息矩阵, 分析其在不同故障信号中的区分度;

5) 求取 5 个估计子带信号的近似熵,分析其在不同故障信号中的区分度,并对比 SDICA 估计 IMF 子带与原 IMF 子带在不同故障中区分度的强弱。

### 3.2 残余互信息 MI 的分析

由于 RMI 是一矩阵,故用 RMI 的行向量之和  $\varphi_1$  或 RMI 矩阵总值  $\varphi_s$  来表征不同故障信号,其公式为

$$\varphi_1 = \sum_{j=1, i \neq j}^5 \varphi_{ij} \quad (13)$$

$$\varphi_s = \sum_{i=1}^5 \sum_{j=1, j \neq i}^5 \varphi_{ij} \quad (14)$$

$\varphi_1$  和  $\varphi_s$  在 3 种轴承故障中的区分度如图 1 所示,横坐标为 36 组信号,纵坐标为残余 MI 值。

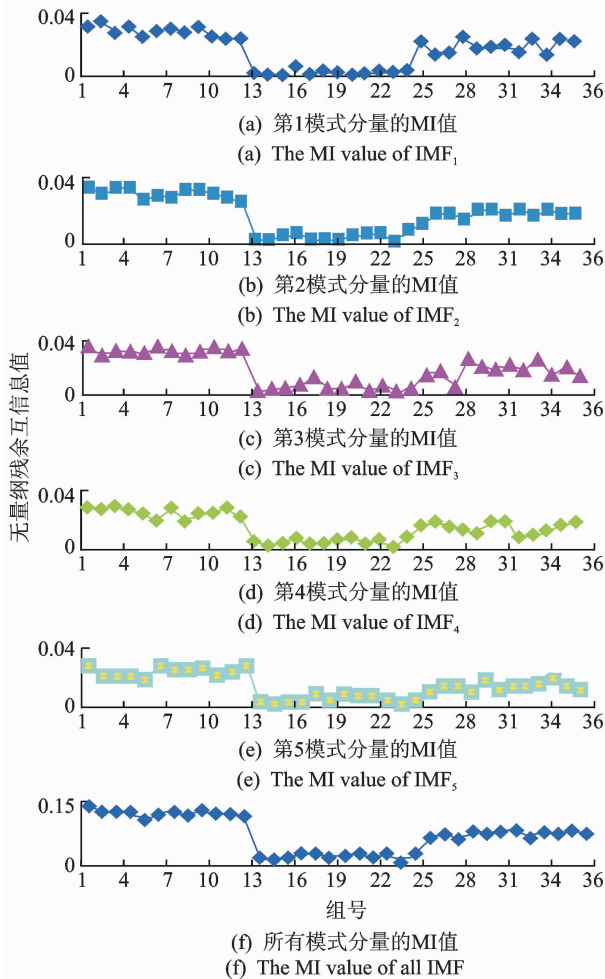


图 1 SDICA 残余互信息在 3 种故障中的区分度  
Fig. 1 The difference of SDICA residual mutual information in three fault type

由图可知,每个 IMF 的 MI 值  $\varphi_1$  在 3 种故障中都有较好的区分度。但是它在同类故障中有较大波动并有奇异点的存在,这会给故障分类带来一定影

响。而所有 IMF 的 MI 总值  $\varphi_s$  在一定程度上抑制了  $\varphi_1$  的波动,减少了奇异点的出现,能够明显地区分 3 种故障。因此,  $\varphi_s$  信息可以用于识别不同故障类型。

### 3.3 估计子带的近似熵

对于 SDICA 得到的 5 个估计子带信号分别作它们的近似熵值,如图 2 所示,横坐标为 36 组信号,纵坐标为估计子带的近似熵值。由图 2(b)可知,第 2 个估计子带的近似熵对 3 种轴承不同故障的区分度十分明显。而其他估计分量的近似熵对 3 种轴承故障的区分度不大。SDICA 所运用的子带由 EMD 得到,EMD 的自适应性分解造成了结果不可预知性,从而需要验证分解的稳定性问题。那么 IMF 的稳定性与估计 IMF 的稳定性是一致的,而估计 IMF<sub>2</sub> 的近似熵值在内环同组的 12 故障的标准偏差为 0.050 49,在滚动体同组 12 故障的标准偏差为 0.026 08,在外环同组 12 故障的标准偏差为 0.039 7。由于同组标准偏差较小,可以证明 EMD 对于滚动轴承实验台信号分解是稳定的。

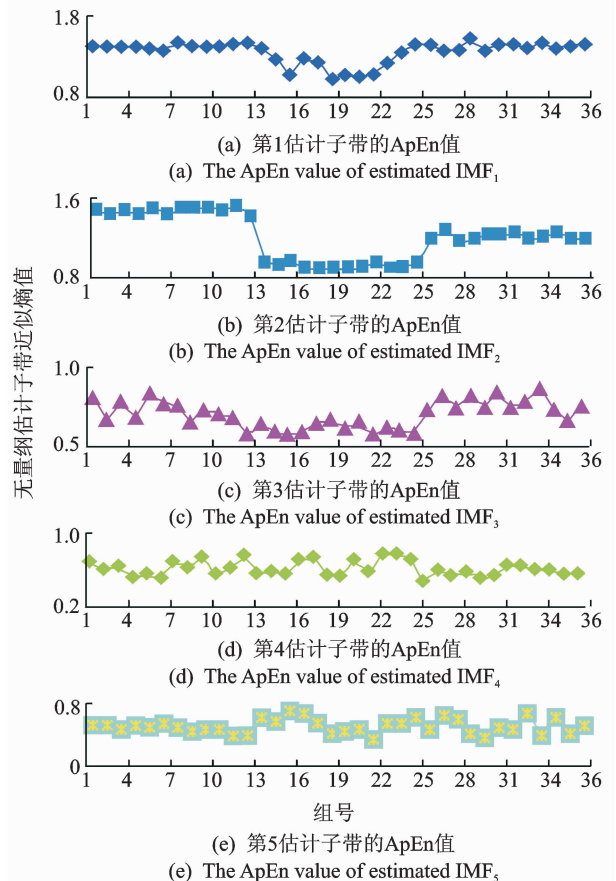


图 2 估计 EMD 子带近似熵在 3 种故障的区分度  
Fig. 2 The difference of EMD estimated subband approximate entropy in three fault type

为了说明 SDICA 在故障分类中能够提取有效的故障特征参数,下面计算 EMD 分解后得到的前 5 个原基本模式分量的近似熵值,如图 3 所示,横坐标为 36 组信号,纵坐标为 IMF 的近似熵值。对比图 2 和图 3 可知,图 2(b)在 3 种故障中有明显的区分度,这说明 SDICA 方法有效的将大多故障信号成分集中到第 2 估计子带,从而使其近似熵值在 3 种不同故障中呈现很好的区分度。图 3(a)~(e) 5 个 IMF 信号的近似熵值在 3 种故障中区分度很差,这说明 EMD 直接分解得到的 5 个 IMF 中 3 种故障信号成分分布较广,在每个 IMF 中占有微弱比重,从而使近似熵值不能区分 3 种故障。

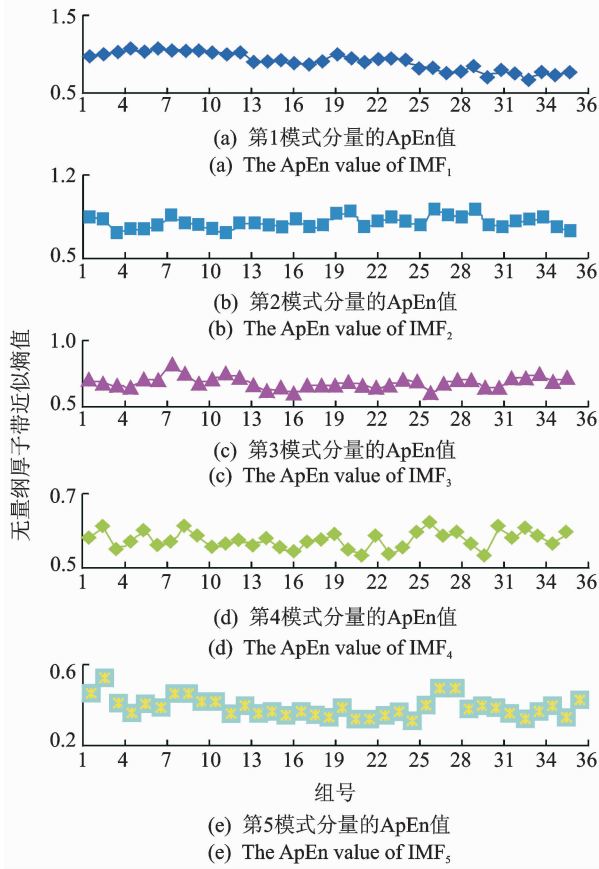


图 3 原 EMD 子带近似熵在 3 种故障中区分度

Fig. 3 The difference of EMD subband approximate entropy in three fault type

### 4 基于 SDICA 与 GRNN 的故障分类

现探讨利用 SDICA 的特征信息和广义回归神经网络 (general regression neural network, 简称 GRNN) 实现轴承故障的智能分类。GRNN 是基于数理统计基础上的神经网络,能够根据样本数据逼近其中隐含的映射关系,即使样本数据稀少,网络的

输出结果也能够收敛于最优回归表面。由于它具有结构自适应、输出与初始权值无关等优良性能,在多维面拟合、自由曲面重构、函数逼近等领域有比较多的应用。基于 GRNN 的上述优点,将 GRNN 实现滚动轴承的故障自动分类。

分别截取滚动轴承数据库中 3 种单通道轴承故障的 100 组信号。利用前 80 组信号作为训练样本,选取后 20 组作为测试样本,选用所有 IMF 的 MI 值  $\varphi_i$  和第 2 估计子带近似熵值作为特征向量对 GRNN 神经网络进行训练,从而实现其智能分类结果如表 1 所示。由表 1 可知,3 种故障的 20 个样本都得到正确的分类,滚动体故障的分类效果最好分类系数达到 100%,内环故障、外环故障的分类系数达到了 95% 和 92% 以上。由此表明,SDICA 方法处理后提取的特征信息不但全部正确分类,而且其分类系数远远大于基本分类系数 50%。

表 1 轴承 3 种故障经 GRNN 分类结果

Tab. 1 The classification result of bearing three fault type applied by GRNN

样本	内环故障		滚动体故障		外环故障				
1	0.99	0	0.01	0	1	0	0.02	0	0.98
2	0.96	0	0.04	0	1	0	0.06	0	0.94
3	0.97	0	0.03	0	1	0	0.08	0	0.92
4	1.00	0	0.00	0	1	0	0.04	0	0.96
5	0.97	0	0.03	0	1	0	0.03	0	0.97
6	0.96	0	0.04	0	1	0	0.10	0	0.90
7	0.97	0	0.03	0	1	0	0.03	0	0.97
8	0.98	0	0.02	0	1	0	0.03	0	0.97
9	0.97	0	0.03	0	1	0	0.03	0	0.97
10	0.97	0	0.03	0	1	0	0.03	0	0.97
11	0.96	0	0.04	0	1	0	0.09	0	0.91
12	1.00	0	0.00	0	1	0	0.01	0	0.99
13	0.97	0	0.03	0	1	0	0.03	0	0.97
14	0.97	0	0.03	0	1	0	0.03	0	0.97
15	0.98	0	0.02	0	1	0	0.02	0	0.98
16	0.97	0	0.03	0	1	0	0.01	0	0.99
17	0.98	0	0.02	0	1	0	0.01	0	0.99
18	0.96	0	0.04	0	1	0	0.07	0	0.93
19	0.96	0	0.04	0	1	0	0.06	0	0.94
20	0.97	0	0.03	0	1	0	0.03	0	0.97

## 5 结束语

笔者开展了SDICA方法在单通道轴承信号故障分类中的研究。通过对滚动轴承故障精确分类实验表明:SDICA在滚动轴承故障分类中成功地提取出包含故障信息丰富的估计子带信号;估计子带信号的SDICA过程特征信息及估计子带近似熵信息在3种故障分类中差异性最明显,可以成为滚动轴承故障分类的重要依据。

SDICA方法在单通道信号故障分类中的成功应用为机械设备故障分类提供了一种新的方法,同时也为单通道信号振源分离提供了一个思路。

### 参 考 文 献

- [1] Jang G J, Lee T W, Oh Y H. Single channel signal separation using MAP-based subspace decomposition [J]. *Electronics Letters*, 2003, 39(24):1766-1767.
- [2] James C J, Lowe D. Single channel analysis of electromagnetic brain signals through ICA in a dynamical systems framework[C]// *Proceedings of the 23rd annual international conference of the IEEE engineering in medicine and biology society (EMBS)*. Istanbul, Turkey: Buiding New Bridges at the Frontiers of Engineering and Medicine, 2001:1974-1977.
- [3] 王向红,胡宏伟,张志勇,等.微弱裂纹信号的稀疏编码提取[J]. *振动工程学报*, 2013, 26(3):311-317.  
Wang Xianghong, Hu Hongwei, Zhang Zhiyong, et al. Extraction of weak crack signals by sparse code [J]. *Journal of Vibration Engineering*, 2013, 26(3): 311-317. (in Chinese)
- [4] Zhang Kun, Chan Laiwan. An adaptive method for sub-band decomposition ICA [J]. *Neural Computation*, 2006, 18(1):191-223.
- [5] Kopriva I, Sersic D. Wavelet packets approach to blind separation of statistically dependent sources[J]. *Neurocomputing*, 2008 (71):1642-1655.
- [6] 胥永刚,孟志鹏,陆明.双树复小波包和ICA用于滚动轴承复合故障诊断[J]. *振动、测试与诊断*, 2015, 35(2):513-518.  
Xu Yonggang, Meng Zhipeng, Lu Ming. Compound fault diagnosis of rolling bearing based on dual-tree complex wavelet packet transform and ICA[J]. *Journal of Vibration, Measurement & Diagnosis*, 2015, 35(2):513-518. (in Chinese)

- [7] Naik G R, Kumar D K, Singh V P, et al. Hand gestures for HCI using ICA of EMG[C]// *Hcsnet Workshop on Use of Vision in Human-Computer Interaction*. Australia: Australian Computer Society, 2006: 67-72.
- [8] Lee T W, Lewicki M S, Girolami M, et al. Blind source separation of more sources than mixtures using overcomplete representations [J]. *IEEE Signal Processing Letters*, 1999, 6(4):87-90.
- [9] Kim S, Yoo C D. Underdetermined blind source separation based on subspace representation [J]. *IEEE Transactions on Signal Processing*, 2009, 57(7):2604-2614.
- [10] Yang Zuyuan, Zhou Guoxu, Wu Zongze, et al. New method for signal encryption using blind source separation based on subband decomposition [J]. *Progress in Natural Science: Materials International*, 2008, 18(6): 751-755.
- [11] Huang N E, Shen Zheng, Long S R, et al. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time Series analysis [J]. *Proceedings of the Royal Society A Mathematical Physical & Engineering Sciences*, 1998, 454(1971): 903-995.
- [12] 杨世锡,焦卫东,吴昭同.基于独立分量分析特征提取的复合神经网络故障诊断法[J]. *振动工程学报*, 2004, 17(4):438-442.  
Yang Shixi, Jiao Weidong, Wu Zhaotong. Multi-neural networks for faults diagnosis based on ICA feature extraction[J]. *Journal of Vibration Engineering*, 2004, 17(4):438-442. (in Chinese)
- [13] 张学清,梁军.基于EEMD-近似熵和储备池的风电功率混沌时间序列预测模型[J]. *物理学报*, 2013, 62(5): 76-85.  
Zhang Xueqing, Liang Jun. Chaotic time series prediction model of wind power based on ensemble empirical mode decomposition-approximate entropy and reservoir[J]. *Acta Physica Sinica*, 2013, 62(5):76-85. (in Chinese)



**第一作者简介:**陈建国,男,1977年1月生,博士、讲师。主要研究方向为机械振动信号分析。曾发表《Sub-Band ICA with selection criterion for BBS of dependent images》(《*Journal of Harbin Institute of Technology*》2011, Vol. 18, No. 4)等论文。

E-mail:jg\_chen@126.com

