

# 基于 EEMD 和 ICA 的单通道列车信号盲分离\*

谷千伟, 金炜东, 余志斌

(西南交通大学 电气工程学院, 成都 610031)

**摘要:** 针对列车混合故障的诊断, 提出了一种基于集合平均经验分解(EEMD)和独立分量分析(ICA)的盲分离诊断方法。通过 EEMD 算法将混合信号分解为包含不同源信号特征的本征模态函数(IMF), 组成新的多维信号; 用主成分分析准确估计源信号个数, 解决了单通道信号盲分离的欠定问题; 利用快速独立分量分析(FastICA)算法实现了信号的盲分离。实验信号分别采用仿真信号和列车实验信号进行实验, 实验结果表明, 该算法可以有效地分离出列车的单故障信号。

**关键词:** 盲源分离; 单通道; 列车故障; 经验模态分解; 独立分量分析

**中图分类号:** TP391      **文献标志码:** A      **文章编号:** 1001-3695(2014)05-1551-03

**doi:** 10.3969/j.issn.1001-3695.2014.05.065

## Blind source separation of single-channel train signal based on EEMD and ICA

GU Qian-wei, JIN Wei-dong, YU Zhi-bin

(School of Electrical Engineering, Southwest Jiaotong University, Chengdu 610031, China)

**Abstract:** Blind source separation is an effective method for multiple fault diagnosis. This paper proposed a new blind source separation algorithm based on ensemble empirical mode decomposition (EEMD) for fault diagnosis of train signal. Nonlinear mixed signal filtered wave was decomposed into intrinsic mode function (IMF) containing different source signal characteristics, they became the new multidimensional signals. The application of principal component analysis (PCA) could accurately estimate the number of source signals to solve the underdetermined problem of single-channel blind signal separation. At last fast independent component analysis algorithm (FastICA) realized the blind separation of signals. The experimental signal used the simulation signal and train mixed fault signal. Experimental results show that this algorithm can effectively analyze the characteristics of train single fault and has important practical value.

**Key words:** blind source separation; single channel; train fault; EMD; independent component analysis

盲源分离技术(BSS)是20世纪80年代中后期迅速发展起来的一种功能强大的信号处理方法, 现已在众多领域中得以研究与发展, 其著名的应用就是“鸡尾酒会”问题。盲源分离是指仅从若干观测到的混合信号中分离出无法直接观测的各个原始信号的过程, “盲”是指源信号个数的不确定性和混合系统特性事先未知两个方面。对于多通道盲源分离, 目前已有成熟的解决方法<sup>[1-6]</sup>, 对于单通道盲源分离由于先验信息的缺乏一直没有得到很好的解决。文献[7]对小波分解与重构之后的细节信号组未加任何处理, 直接作为ICA算法的输入信号, 但如果考虑信道噪声和干扰信号的影响, 该方法无法得到理想的分离效果。后来张纯等人<sup>[8]</sup>提出了一种基于小波变换、奇异值分解和独立分量分析的单通道盲分离算法, 但是小波基的选择不同会导致分离结果差异很大。刘佳等人利用虚拟通道<sup>[9]</sup>、EEMD<sup>[10]</sup>和间隔抽采样<sup>[11]</sup>扩展通道, 然后结合PCA和独立成分分析实现了水声信号的单通道分离。后来研究人员在EEMD的基础上提出了基于功率谱密度函数<sup>[12]</sup>、贝叶斯算法<sup>[13]</sup>、二阶统计量奇异值分解<sup>[14]</sup>等估计源数方法, 最后用独立成分分析实现了信号分离。

本文在现有方法的基础上, 结合列车混合故障信号的特点, 把EEMD和Fastica应用到列车混合故障盲分离上尚属首例。

## 1 理论和方法

### 1.1 集合平均经验模态分解的原理<sup>[15]</sup>

集合平均经验分解(EEMD)算法是针对经验模态分解(EMD)方法的不足提出的一种噪声辅助数据分析方法。EEMD算法是针对EMD方法在对抗混叠效应效果较差这个缺点提出的, 主要利用了添加辅助噪声的数据分析方法。EEMD分解原理是: 在进行振动实验前, 都希望利用降噪技术将噪声从信号数据中去除, 不过在有些情况下可以通过加入噪声的方法来进行辅助分析, 这种方法就称为噪声辅助信号处理, EEMD算法就是利用了这一原理。首先, 附加的白噪声的频谱是均匀分布的, 当额外的噪声均匀地分布在整个时频空间时, 不同尺度的信号将自动映射到与背景白噪声相关的参考尺度上去。其次, 这些白噪声都属于零均值的噪声, 使用足够次数的平均值处理后, 噪声就会相互抵消。最后, 全体集成的均值作为最终的真正结果。

### 1.2 盲源信号分离

#### 1.2.1 源数目估计<sup>[16]</sup>

为实现单通道源信号的盲分离, 首先要求估计系统的源信号数。在EEMD基础上运用主成分分析估计的源数。主成分

收稿日期: 2013-07-01; 修回日期: 2013-08-25      基金项目: 国家自然科学基金资助项目(61134002)

作者简介: 谷千伟(1986-), 男, 河南周口人, 硕士研究生, 主要研究方向为盲源分离(1060146935@qq.com); 金炜东(1959-), 男, 教授, 博士, 主要研究方向为智能信息处理、系统仿真与优化方法等; 余志斌(1977-), 男, 副教授, 博士, 主要研究方向为信号与信息处理、模式识别等。

分析步骤如下:

a) 计算样本数据的协方差矩阵  $S = (s_{ij})_{p \times p}$ , 其中  $s_{ij} = \frac{1}{n-1} \times \sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)$ ,  $i, j = 1, 2, \dots, p$ 。

b) 求出协方差矩阵  $S$  的特征值  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p > 0$  及相应的正交化单位特征向量:

$$a_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{p1} \end{pmatrix}, a_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{p2} \end{pmatrix}, \dots, a_p = \begin{pmatrix} a_{1p} \\ a_{2p} \\ \vdots \\ a_{pp} \end{pmatrix}$$

则  $X$  的第  $i$  个主成分为  $F_i = a_i X, i = 1, 2, \dots, p$ 。

c) 在已确定的全部  $p$  个主成分中合理选择  $r$  来实现最终的评价分析。一般用方差贡献率  $\alpha_i = \lambda_i / \sum_{j=1}^p \lambda_j$  解释主成分  $F_i$  所反映的信息量的大小,  $r$  的确定以累计贡献率  $G(r) = \sum_{j=1}^r \lambda_j / \sum_{j=1}^p \lambda_j$  达到足够大(一般在 90% 以上)为原则。

d) 计算  $n$  个样品在  $r$  个主成分上的得分:

$$F_i = a_{i1}X_1 + a_{i2}X_2 + \dots + a_{ip}X_p \quad i = 1, 2, \dots, r$$

PCA 主要用于数据降维, 对于一系列样本特征组成的多维向量, 其中的某些元素本身没有区分性, 如某个元素在所有的样本中都为 1, 或者与 1 差距不大, 那么这个元素本身就没有区分性, 用它作特征来区分, 贡献会非常小。所以本文的目的是找那些变化大的元素, 即方差大的那些维, 而去除掉那些变化不大的维, 从而使特征留下的都是“精品”, 而且计算量也变小了。

### 1.2.2 盲源信号分离

盲源分离是在源信号与传输通道参数均未知的情况下, 根据输入源信号的统计特性, 仅由观测信号恢复出源信号独立成分的过程<sup>[17,18]</sup>。现以盲源分离的瞬时混合模型为例, 设  $s(t) = (s_1(t), s_2(t), \dots, s_N(t))^T$ , 其中  $s_1(t), s_2(t), \dots, s_N(t)$  表示  $N$  个零均值而且统计独立的源信号,  $x(t) = (x_1(t), x_2(t), \dots, x_M(t))^T$ , 其中  $x_M(t)$  表示  $M$  个观测信号, 且  $M \geq N$ 。令  $A$  为  $M \times N$  的未知混合矩阵, 则瞬时混合模型为  $x(t) = As(t)$ , 盲源分离就是根据观测信号  $x(t) = (x_1(t), x_2(t), \dots, x_M(t))^T$  寻找  $N \times M$  的分离矩阵  $W$ , 并恢复源信号  $s(t)$ 。

### 1.2.3 盲源分离算法步骤<sup>[16,19]</sup>

a) 单通道观测信号  $x(t)$  的 EEMD 分解。单通道观测信号的 EEMD 分解将得到 IMF 分量  $x_{\text{imf}} = (c_1, c_2, \dots, c_{1n}, r_{1n})^T$ 。

b) 源数估计。单通道信号  $x(t)$  和 IMF 组合成为新的多维信号  $x_{\text{imf}} = (x_1, c_1, \dots, c_{1n}, r_{1n})^T$ , 其相关矩阵  $R_x = E[x_{\text{imf}}(t)x_{\text{imf}}^H(t)]$ , 并奇异值分解, 根据其特征值估计源信号数目。

c) 合成新的多维信号。将 IMF 组合成为新的多维信号  $x_1 = (c_1, c_2, \dots, c_n)^T$ , 并使其维数等于估计的源信号数。

d) 盲信号分离。针对新的多维信号, 应用 ICA 相关算法实现盲源分离, 得到分离后的源信号  $y$ 。

## 2 实验结果和分析

### 2.1 仿真信号实验

为了验证列车信号分离的有效性, 这里使用两个模拟非平稳信号进行仿真实验, 取样频率为 100 Hz, 取样点数为 1 000 个, 信号分别为  $s_1 = \cos(2\pi \times 20 \times t)$ ,  $s_2 = \sin(2\pi \times 5 \times t)$ 。

设混合信号模型  $s = as_1(t) + bs_2(t) + n(t)$ , 这里取  $a = 1, b = 1, n(t)$  为随机白噪声信号。这样观测信号为 1 个, 源信号数为 2 个, 使得观测信号数小于源信号数, 这就是盲源分离的欠定问题。仿真信号  $s_1, s_2$  及混合信号  $s$  的频谱如图 1 所示。

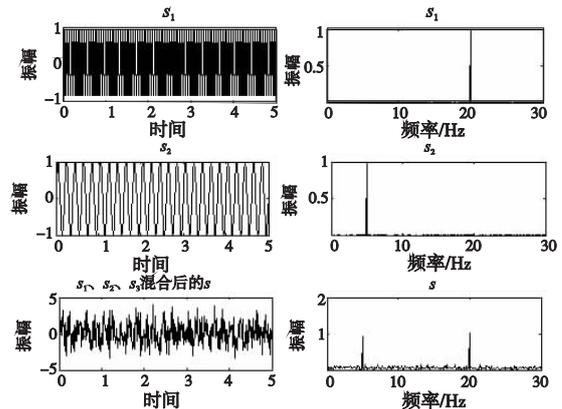


图1 源信号及混合信号的时域和频域波形

使用 EEMD 的盲源分离方法, 并用 PCA 进行源信号数的估计, 选择了涵盖源信号信息的 95% 为依据, 判定源信号为 2 个, 最后用常用的 Fastica 盲源分离方法分离。从图 2 可以看出,  $s_1, s_2$  的图谱已经完全分离, 信号得到了很好的分离效果。

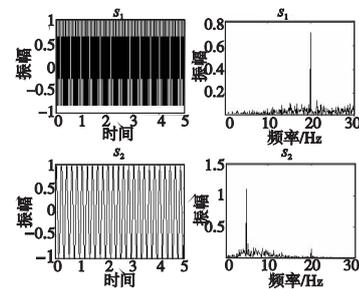


图2 恢复的源信号时域和频域波形

### 2.2 实际信号实验

#### 2.2.1 列车信号特征分析

实验采用单通道盲源分离, 因为单通道的观测是最特殊的一种欠定分离情况, 对单通道信息的研究可以有助于列车混合故障的盲源分离, 并且可以更好地展现 EEMD、PCA 和 ICA 的优越性。数据采集来自列车空气弹簧失气和横向减振器失效的混合故障, 空气弹簧失气属于垂向故障, 横向减振器失效属于列车水平方向的故障。空气弹簧失气一般由于二系悬挂发生问题引起, 往往使构架的振动加速度加大, 造成舒适度指标下降, 影响行车安全, 其特征是构架垂向加速度增大, 主振动能量趋于低频段, 但是大小变化不大。横向减振器失效导致车辆的横向运动失稳, 其失稳情况还与列车速度有关。随着速度的增加, 发生失稳的概率加大, 其特征是转向架有持续的较大的横向加速度, 有较窄频宽的能量分布突出点, 即周期性成分增加。并且构架横向加速度及其频谱变化可以成为此故障判断条件之一, 根据时频域的分析, 此种故障情况下, 在构架上其故障频率在高频和低频下都有分布, 此特征随着速度的增加变得明显<sup>[20]</sup>。

本文选取动车组 CHR3 构架上传感器的混合数据, 对它进行分离以验证单通道盲源分离的有效性。实验的采样频率为 243 Hz, 采样点数选取 14 580, 并选取 200 km/h。发生单个故障时两种故障的频谱如图 3 所示。

混合信号的时域和频域波形如图 4 所示。对于该 1 路混

合信号,首先进行EEMD分解,得到其IMF分量。

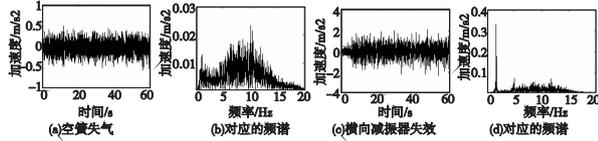


图3 源信号的时域和频域波形

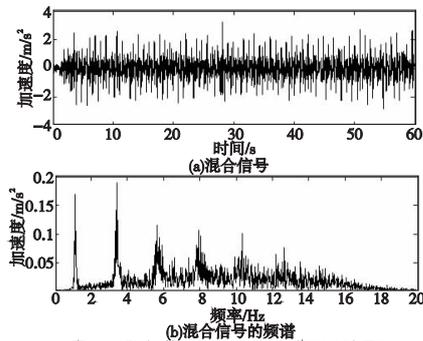


图4 混合信号的时域和频域波形

### 2.2.2 源个数的估计和盲源分离

单通道观测信号 $x(t)$ 和它的本征模函数 $x_{imf} = (c_1, c_2, \dots, c_{13}, r_{13})^T$ 组合成新的多维信号, $x_{imf} = (x_1, c_1, \dots, c_{13}, r_{13})^T$ ,其相关矩阵 $R_x = E(x_{imf}(t)x_{imf}^H(t))$ 。

得到特征值矢量 $\Lambda = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_{14}\}$ ,如表1所示,易知 $\lambda_1$ 和 $\lambda_2$ 已经包含了信号的主要信息。

表1 多维信号 $x_{imf}$ 的特征值

特征值	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	$\lambda_5$
数值	7615.2	2044.5	593.1	466.7	44.7
特征值	$\lambda_6$	$\lambda_7$	$\lambda_8$	$\lambda_9$	$\lambda_{10}$
数值	28.1	9.4	2.3	2.1	0.9
特征值	$\lambda_{11}$	$\lambda_{12}$	$\lambda_{13}$	$\lambda_{14}$	
数值	0.7	0.4	0.2	0.1	

对比图3和5可得:图5所示的两个分离信号的时域波形和频谱分布与图3中原始信号基本相同,列车源信号得到了很好的分离。图5的两个分离信号的时域和频域幅值发生了变化,这恰好说明了盲源分离的不确定性。由此可以看出,利用EEMD,不仅可以单通道机械信号升维,而且可以深入挖掘隐藏在单通道列车故障信号内部的深层信息,从而便于应用相应的盲源分离算法,并将源信号分离。

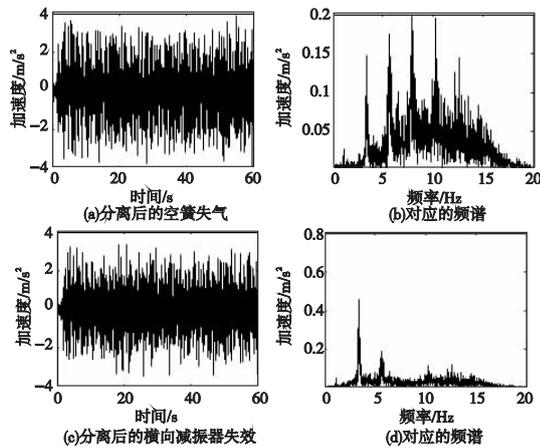


图5 恢复的源信号时域和频域波形

该方法应用于列车其他混合故障的分离,也可以正确地分离出相应的源信号,表明该方法是可以应用于列车故障的盲信号诊断实践的。

### 3 结束语

综上所述,此算法先通过EEMD算法把信号分解成具有内在特征的多个本征函数,根据PCA方法得到的特征值分布情况能获得源信号数目估计,再进行重新组合,把欠定问题变成适定问题,然后利用盲源分离算法进行故障特征的提取。仿真研究表明,在利用EEMD将列车故障的非线性、非平稳信号分解为一系列线性、平稳的IMF分离的基础上,能正确估计列车混合故障的源信号个数,从而为列车故障盲源分离研究和应用提供了一种新方法。

#### 参考文献:

- [1] TICHAVSKY P, KOLDVSKY Z, OJA E. Performance analysis of the FastICA algorithm and Cramér-Rao bounds for linear independent component analysis[J]. IEEE Trans on Signal Processing, 2006, 54(4): 1189-1203.
- [2] TAOUFIKI M, ADIB A, ABOUTAJ-DINE D. Blind separation of any source distributions via high-order statistics[J]. Signal Processing, 2007, 87(8): 1882-1889.
- [3] ZBYNEK K, PETR T, OJA E. Efficient variant of algorithm FastICA for independent component analysis attaining the Cramér-Rao lower bound[J]. IEEE Trans on Neural Networks, 2006, 17(5): 1256-1277.
- [4] 孙世军, 彭承琳, 侯文生, 等. 基于时间序列变分贝叶斯理论的信号盲源分离[J]. 仪器仪表学报, 2009, 30(8): 1571-1576.
- [5] 任东晓, 叶茂, 殷英, 等. 基于互信息最小的非线性混合胎儿心电图信号提取方法[J]. 电子测量与仪器学报, 2010, 24(7): 680-685.
- [6] OU Shi-feng, GAO Ying, ZHAO Xiao-hui. Adaptive combination algorithm and its modified scheme for blind source separation[J]. Journal of Electronics and Information Technology, 2011, 33(5): 1243-1247.
- [7] SHAO Hua, SHI Xin-hua, LI Lin. Power signal separation in milling process based on wavelet transform and independent component analysis[J]. International Journal of Machine Tools and Manufacture, 2011, 51(9): 701-710.
- [8] 张纯, 杨俊安, 叶风, 等. 基于小波分量奇异值分解的单通道盲分离算法[J]. 电子测量与仪器学报, 2011, 25(11): 991-997.
- [9] 刘佳, 杨士莪, 朴胜春, 等. 单观测通道船舶辐射噪声盲源分离[J]. 声学学报, 2011, 36(3): 265-270.
- [10] 刘佳, 杨士莪, 朴胜春, 等. 基于EEMD的地声信号单通道盲源分离算法[J]. 哈尔滨工程大学学报, 2011, 32(2): 194-199.
- [11] 刘佳. 多途环境下的单通道水声信号盲源分离[J]. 振动与冲击, 2012, 31(6): 15-18.
- [12] 刘晓伟. 车辆结构振动与噪声源的盲分离技术研究[D]. 南京: 南京航空航天大学, 2012.
- [13] 赵志强, 颜学龙. 基于EMD和ICA的单通道语音盲源分离算法[J]. 电子科技, 2012, 25(7): 66-68, 75.
- [14] 董绍江, 汤宝平, 宋涛, 等. 基于最优匹配跟踪算法的单通道机械信号盲源分离[J]. 振动工程学报, 2012, 25(6): 724-731.
- [15] 王新立, 徐元博. 基于EEMD分解的盲源分离算法在振动筛轴承故障诊断中的应用[J]. 矿山机械, 2013, 41(4): 85-89.
- [16] MIJOVIC B, De VOS M, GLIGORJEVIC I, et al. Combining EMD with ICA for extracting independent sources from single channel and two-channel data [C]//Proc of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2010: 5387-5390.
- [17] COMON P. Independent component analysis: a new concept [J]. Signal Processing, 1994, 36(3): 287-314.
- [18] HYVARINEN A, KARHUNEN J, OJA E. Independent component analysis[M]. New York: Wiley, 2001.
- [19] 毋文峰, 陈小虎, 苏勋家. 基于经验模态分解的单通道机械信号盲源分离[J]. 机械工程学报, 2011, 47(7): 13-16.
- [20] 张兵. 列车关键部件安全监测理论与分析研究[D]. 成都: 西南交通大学, 2007.