

基于高阶微分的 EMD 均值计算方法

黎 恒¹, 李 智^{2,3}

(1. 西安电子科技大学机电工程学院, 陕西西安 710071; 2. 桂林航天工业学院, 广西桂林 541004;
3. 桂林电子科技大学电子工程与自动化学院, 广西桂林 541004)

摘 要: 为了改善经验模态分解的分离性能, 提出一种基于信号高阶微分的分解算法. 本文首先讨论了经验模态实现模态分离的必要条件, 并证明对输入信号进行偶数阶数值微分可以提高模态分离性能. 然后在此基础上提出一种以偶数阶微分的过零点为特征的均值计算方法. 最后对仿真信号的分解进行了实验研究. 结果表明, 本文方法可以改善分离性能, 性能提高的程度与理论分析结果符合; 与经验模态分解相比, 本文方法具有更高的分解精度.

关键词: 经验模态分解; 数值微分; 时间序列分析

中图分类号: TN911.72

文献标识码: A

文章编号: 0372-2112 (2015)06-1073-05

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2015.06.005

Local Mean Computation for Empirical Mode Decomposition Based on Higher Order Derivative

LI Heng¹, LI Zhi^{2,3}

(1. School of Mechano-Electronic Engineering, Xidian University, Xi'an, Shaanxi 710071, China;

2. Guilin University of Aerospace Technology, Guilin, Guangxi 541004, China;

3. School of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin, Guangxi 541004, China)

Abstract: In order to improve the separation performance of empirical mode decomposition, an algorithm based on higher order derivative is put forward. First, a necessary condition for separability is studied. It is proved that separability can be improved by even order derivative of the input signal. An algorithm that builds local mean from zero-crossings of the even order derivative is then proposed. Finally, decomposition performance of the proposed method is demonstrated by application to synthetic signals. Numerical experiments show that the proposed method has enhanced separation quality which agrees well with the theoretical analysis. In compared with the empirical mode decomposition, the proposed method exhibits higher accuracy.

Key words: empirical mode decomposition; numerical derivative; time sequence analysis

1 引言

经验模态分解 (Empirical Mode Decomposition, EMD) 是美国科学院院士 Huang^[1] 提出的信号分解方法. EMD 通过一种被称为筛分的方法, 将多分量输入信号分解成多个本征模态函数 (Intrinsic Mode Function, IMF). 筛分的目的是不断地将低频分量从信号中分离, 最终得到关于时间轴对称的 IMF. 与其他方法相比, EMD 具有自适应、后验性等特点, 适用于平稳信号和非平稳信号. 使该方法一经提出便在各个领域得到了广泛的应用^[2~7].

EMD 方法以极值点作为信号的局部震荡尺度构造上下包络, 取其平均进行筛分. 当输入信号中存在频率接近、幅度差异较大的分量时, 极值点往往与真实震荡

尺度有偏差. 如果仍选择极值点作为特征进行分解, 将无法分离这些信号分量. 近年来, 大量学者针对这个问题对 EMD 进行了优化和改进. 文献[8]使用升余弦滤波器插值代替原方法的三次样条插值; 文献[9]采用优化算法寻找最佳插值点计算上下包络; 文献[10]通过计算局部积分均值点进行插值直接计算均值曲线. 这些方法对提高分解性能都有一定的效果, 但是性能提高程度有限, 而且大多没有给出性能提高程度的依据.

本文从信号的偶数阶微分入手, 研究高阶微分的特性与分解性能的关系, 以偶数阶微分的过零点时刻作为特征进行信号分解, 并建立了理论体系. 仿真结果表明, 该方法性能提升程度可预知、可控, 能显著地提高 EMD 方法的分离性能.

2 问题的提出

EMD 迭代过程的核心在于均值信号的计算, 均值信号的准确性直接决定着分解质量. EMD 方法^[1]认为, 信号的极值点分布包含了其自身局部震荡的时间尺度信息. EMD 方法通过对极大值和极小值点进行三次样条插值分别得到上包络和下包络, 取上下包络的算术平均值作为均值信号. 这种均值计算方法有其难以解决的缺点, 当信号中存在频率上较为接近的模式分量时, EMD 方法无法分离这些分量, 甚至在分解过程中会产生杂散分量. 如图 1(a) 所示含 15Hz 和 20Hz 两种频率的信号 $s(t) = \cos(15 \times 2\pi t) + \cos(20 \times 2\pi t)$, $t \in [0, 1.5]$, 给出了 EMD 方法首次迭代产生的上下包络和均值信号. 可以看出, 均值曲线明显与 15Hz 的低频分量相异. 对均值信号进行频谱分析结果如图 1(b) 所示, 有多个频点出现了较大的杂散分量, 影响了分解结果的准确性. 因此, 在这种情况下, 使用极值点作为特征进行分解已经不能满足要求, 有必要选取不同的特征进行分解.

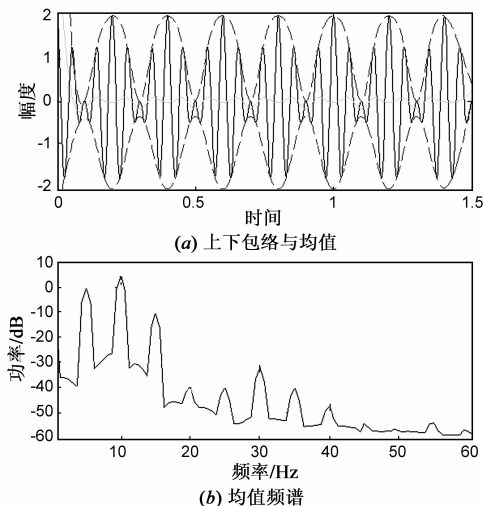


图1 EMD方法首次迭代

3 基于高阶数值微分的经验模态分解

3.1 理论模型

首先研究信号实现分离的条件, 考虑式(1)的双音信号模型:

$$s(t) = \cos 2\pi t + a \cos(2\pi f t + \varphi) \quad (1)$$

其中 a 为幅度比, f 为频率比, φ 为相位差. 当 $f \in [0, 1)$ 时, $\cos 2\pi t$ 为频率较高的信号分量. 为得到低频信号, 本文从信号的高阶导数特征进行分析.

定理 1 式(1)所表示的信号, 其中 $a > 0, f \in (0, 1)$. 记 $s^{(n)}(t)$ 为 $s(t)$ 的 n 阶导数, $n = 0, 1, 2, \dots$, $s^{(2n)}(t)$ 过零点为 t_0 , 则有: $s(t_0) = a(1 - f^{2n}) \cos(2\pi f t_0 + \varphi)$.

证明 对 $s(t)$ 求 $2n$ 阶导数得

$$s^{2n}(t) \propto \cos 2\pi t + a f^{2n} \cos(2\pi f t + \varphi) \quad (2)$$

在 $s^{(2n)}(t)$ 的过零点时刻 t_0 有

$$\cos 2\pi t_0 + a f^{2n} \cos(2\pi f t_0 + \varphi) = 0 \quad (3)$$

将式(3)代入 $s(t_0)$ 得证. 定理 1 表明了只要求出信号偶数阶导数的过零点时刻 t_0 , 就能求出与低频信号同频同相, 幅度为 $a(1 - f^{2n})$ 的信号. 容易验证, 该规律对多分量组成的信号仍然是成立的.

下面讨论实现分解的必要条件. 在计算过程中, 低频分量可通过对 $s(t_0)$ 插值重建, 如使用 B 样条插值或 sinc 插值等. 对插值而言, 周期内提供插值的点数越多, 波形越准确. 对 $s^{(2n)}(t)$ 而言, t_0 个数的上限(下限)分别为高频(低频)分量的过零点个数. 当 t_0 的个数和低频分量的过零点个数一致时, 单个周期内插值点只有两个, 不可能恢复该信号. 当 t_0 的个数与高频分量过零点个数相同时, 才最有可能通过插值恢复出低频信号. 下面分析双音信号在什么情况下, 其偶数阶导数的过零点个数与高频分量相同.

定理 2 式(1)所表示的信号, 其中 $a > 0, f \in (0, 1)$. 记 $s^{(n)}(t)$ 为 $s(t)$ 的 n 阶导数, 其中 $n = 0, 1, 2, \dots$, 当 $a f^{2n+1} < 1$ 时, $s^{(2n)}(t)$ 的过零点个数与 $\cos 2\pi t$ 相同.

证明 要证明 $s^{(2n)}(t)$ 的过零点个数与 $\cos 2\pi t$ 相同, 即证明 $s^{2n}(t)$ 极值点出现频率与 $\cos 2\pi t$ 相同. 假设 $s^{(2n)}(t)$ 的一个极值点在 $t = t_0$ 时刻出现, 即

$$s^{2n+1}(t)|_{t=t_0} \propto \sin 2\pi t_0 + a f^{2n+1} \sin(2\pi f t_0 + \varphi) = 0 \quad (4)$$

$s(t)$ 的 $2n+2$ 阶导数可以表示为

$$s^{2n+2}(t) \propto \cos 2\pi t + a f^{2n+2} \cos(2\pi f t + \varphi) \quad (5)$$

令

$$\gamma = |\cos 2\pi t_0|^2 - |a f^{2n+2} \cos(2\pi f t_0 + \varphi)|^2 \quad (6)$$

将式(4)代入式(6), 得

$$\gamma = 1 - (a f^{2n+1})^2 \sin^2(2\pi f t_0 + \varphi) - (a f^{2n+2})^2 \cos^2(2\pi f t_0 + \varphi) \quad (7)$$

当 $a f^{2n+1} < 1$ 时, 有 $(a f^{2n+2})^2 < (a f^{2n+1})^2 < 1$, 从而有 $\gamma > 0$, 由式(6)可以得到

$$|\cos 2\pi t_0| > |a f^{2n+2} \cos(2\pi f t_0 + \varphi)| \quad (8)$$

将式(8)与式(5)比较可以看出 $\cos 2\pi t_0 + a f^{2n+2} \cos(2\pi f t_0 + \varphi)$ 的符号与 $\cos 2\pi t_0$ 一致, 即 $s^{2n+2}(t_0)$ 的符号与 $\cos 2\pi t_0$ 一致. 容易得知, $s^{2n+2}(t_0)$ 的符号对应着 $s^{2n}(t_0)$ 的凹凸性和极值点类型, 当 $s^{2n+2}(t_0) > 0$ 时, $s^{2n}(t_0)$ 位于凹曲线内, 为极小值点; 而当 $s^{2n+2}(t_0) < 0$ 时, $s^{2n}(t_0)$ 位于凸曲线内, 为极大值点. 由此可以得到, $s^{2n}(t)$ 的凹凸性变化频率与高频分量 $\cos 2\pi t$ 的极值点出现频率一致. 由式(4), 当 $a f^{2n+1} < 1$ 时, $s^{2n+1}(t)$ 的符号与 $\sin 2\pi t$ 一致, 因此 $s^{2n+1}(t)$ 的极值点数与 $\sin 2\pi t$ 一致, $s^{2n+2}(t)$ 的过零点个数与 $\cos 2\pi t$ 一致. 从而 $s^{2n}(t)$ 极值点出现频率

与高频分量 $\cos 2\pi t$ 的极值点出现频率相同,得证。

定理 2 分析了本文方法所能达到的分解性能,即对输入信号进行 $2n$ 阶数值微分有可能分离出频率和幅度比满足 $af^{2n+1} < 1$ 的信号分量。

3.2 算法实现

通过 3.1 节分析可知,对信号进行偶数阶微分能估计出均值信号分量。在应用中,模数转换器和处理器的数字有效位数有限,输入信号不可避免地存在量化误差、截断误差或舍入误差^[11]。这些误差在数值微分的过程中会被迅速放大,严重影响微分的结果。为降低微分误差,本文使用高阶 B 样条拟合^[12]实现数值微分的近似。B 样条拟合是对信号进行平滑的过程,本身具有低滤波器的性质,能抑制细微扰动、失真和间断对微分结果的影响。首先对输入信号进行等节点间隔 B 样条拟合(本文节点间隔取信号相邻两极值点距离最小值的四分之一,且至少取 4),再对拟合曲线进行数值微分。考虑到 n 阶 B 样条函数具有 $n-1$ 阶连续导数,为保证精度, B 样条阶数应至少比微分阶数高 3 阶以上。将本文提出方法称为微分经验模态分解(Derivative Empirical Mode Decomposition, DEMD)。设输入信号为 $s(t)$, DEMD 的均值计算过程按如下步骤进行:

- (1) 预设微分阶数 n , 迭代终止参数 σ , 最大迭代次数(本文取 10), 插值阶数 u ;
- (2) 计算 $s(t)$ 的 $2n$ 阶数值微分, 得到 $s^{(2n)}(t)$;
- (3) 找到 $s^{(2n)}(t)$ 过零点对应的时刻 t_i , 即 $t_i | s^{(2n)}(t_i) = 0 (i = 1, 2, \dots)$;
- (4) 对 $(t_i, s(t_i))$ 进行 u 阶 B 样条插值, 得到均值曲线 $m(t)$;
- (5) 令 $h(t) = s(t) - m(t)$, 将 $h(t)$ 看做新的 $s(t)$ 并重复步骤 2~5, 直到满足 $\sum_i \frac{|m(t_i)|}{|h(t_i)|} < \sigma$ 或达到预设的最大迭代次数为止, σ 通常取小于 0.3 的正值;
- (6) 对步骤 5 得到的 $h(t)$ 作为 IMF 输出。

在 EMD 方法中, 信号样点率对分解精度有较大的影响^[13], 本文方法也存在同样的问题。当样点率较低时, 对极值点/过零点时刻的估计将存在较大误差, 这些误差会导致插值样点的不准确。为保证精度, 信号的采样频率应远大于其奈奎斯特频率。如采样频率较低, 则应在分解前进行适当的插值处理。

4 实验分析

4.1 分解精度

对双音信号 $s(t) = \cos 2\pi t + \cos(2\pi ft)$ 进行分解, 研究在不同频率比 f 和不同插值阶数下, 分解结果第一阶 IMF 的均方误差 (Mean Square Error, MSE)。

$$\text{MSE} = \frac{\sum_{t=1}^N (\text{imf}_1(t) - \cos 2\pi t)^2}{N} \quad (9)$$

其中 N 为样点总数。DEMD 的微分阶数为 8; 迭代终止参数为 0.15。分别对频率比 f 为 0.1~0.8 的信号进行分解, 分解结果如图 2(a) 所示, 可以看出 f 越小, 越容易得到更高的分解精度。提高插值阶数可以在一定程度上提高精度。随着插值阶数的提高, 精度的提升趋于平缓。一般认为对实际信号, 阶数选择 8~10 可满足需求。当信号中低频分量幅度较大时, 可适当提高插值阶数以便于把被淹没的高频分量提取出来。

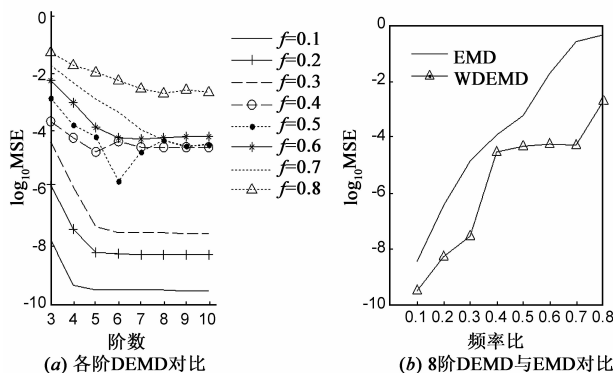


图2 信号分解精度

将插值阶数为 8 的分解结果与 EMD 方法进行精度对比, 结果见图 2(b)。比较得出, 对所有频率比 f 的信号, DEMD 方法的精度明显优于 EMD, 精度至少高 0.5 个数量级。在 $f > 0.6$ 以后, DEMD 的精度优势更明显, 因为在这个时候, EMD 已经无法正确分离频率如此接近的分量。

4.2 分离性能

研究双音信号 $s(t) = \cos 2\pi t + a \cos(2\pi ft)$ 在不同微分阶数下的分解。使用文献[14]中定义的分离指标来衡量性能

$$C(f, a, \varphi) = \frac{\|h - \cos 2\pi t\|_{L_2}}{\|a \cos 2\pi ft + \varphi\|_{L_2}} \quad (10)$$

其中 h 为经过迭代后分解出来的第一阶 IMF 分量。考察在不同幅度比 a 和频率比 f 和相位 φ 下的分解结果。设定与该文献相同的阈值, 当 $c > 0.5$ 时认为分解成功。DEMD 配置情况如下: 插值阶数为 10; 迭代终止参数为 0.15。通常认为相位 φ 对性能影响不大, 实验从 $0^\circ \sim 360^\circ$ 均匀抽取 5 个相位, 并将实验结果取平均。为消除边缘效应对分解结果的干扰, 对分解结果进行了首尾截断处理, 只取中间 80% 长度的数据做分析。

图 3(a)~(d) 给出了 DEMD 在 2 阶、4 阶、6 阶和 8 阶微分下的分解结果。同时图 3(e) 给出 EMD 的分解结果。黑色区域表示分离成功, 白色区域表示分离失败。通过对比可知, 通过微分, DEMD 的分离性能与 EMD 相

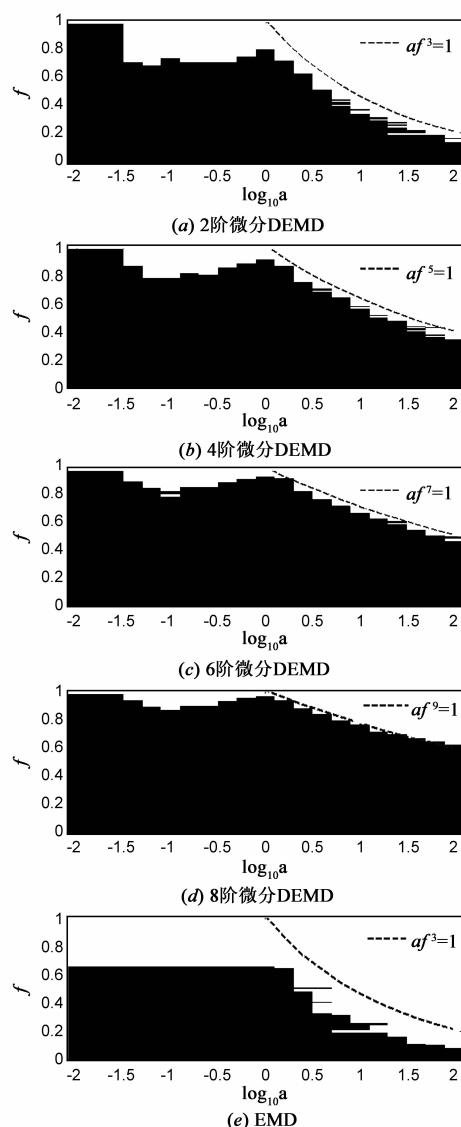


图3 双音信号分离性能

比得到了很大提高,微分阶数越高,分离性越好.图中分别画出了 $af^{2k+1}=1, k=1,2,\dots,4$ 曲线.回顾定理 2,双音信号进行 $2n$ 阶微分后,如果满足 $af^{2n+1}<1$,则为成功分解的必要条件.可以看出,DEMD 的分离性过渡区轮廓曲线与理论曲线 $af^{2k+1}=1$ 基本重合.当 $af^{2n+1}>1$ 时,对应图中 $af^{2k+1}=1$ 曲线上的白色区域,这时 DEMD 无法正确分解.当 $af^{2n+1}<1$ 时,才可能做出正确的分解.当 f 接近 1 时,表示分量的频率非常接近,这时 DEMD 也无法分离这两个模态.但是微分阶数越高,DEMD 分离能力越强.当微分阶数为 8 时,分离能力达到了 $f=0.9$,大大优于 EMD(约 0.65).

4.3 多分量信号分解

将合成信号的模态个数选为 4 个.设输入信号为

$$s(t) = \cos(2\pi \times 50 \times t + \sin(2\pi \times 35 \times t)) + 1.5\cos(2\pi \times 28 \times t) + 0.4\cos(2\pi \times 15 \times t) \quad (11)$$

其中 $t \in [0,1]$, DEMD 配置插值阶数为 10;微分阶数为 8.图 4 为 DEMD 和 EMD 的分解结果,限于篇幅,图中只给出分解的前 4 阶 IMF.从 DEMD 的可以看出,四个模态清晰可辨,体现出该方法分离性高的优势.而 EMD 方法不但没有实现模态的分离,还产生了模态间的混叠,造成严重的失真,使其分解结果完全失去物理意义.

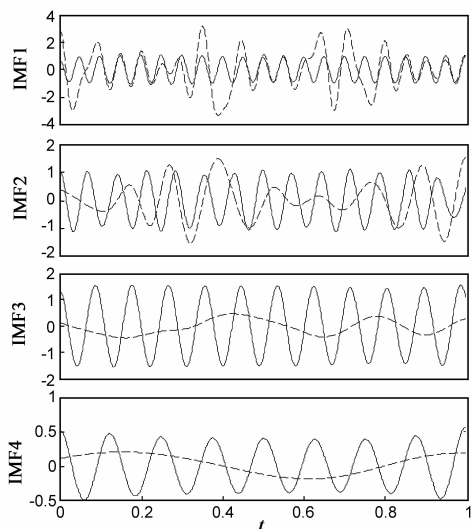


图4 多分量信号分解对比(实线:DEMD;虚线:EMD)

5 结论

本文对提高 EMD 分解性能的可行性进行了深入的讨论,提出一种基于偶数阶微分的 EMD 分解方法.该方法能显著提高分解的精度和模态分离能力.分离性实验结果表明,该方法适用于分解频率较为接近的模态,分离性能与理论分析符合,表明了方法的合理性和有效性.目前 EMD 方法及其改进方法普遍缺乏完整的理论体系,本文的研究是对 EMD 理论体系有价值的补充. DEMD 使用偶阶微分的过零点作为特征进行直接均值计算,是对 EMD 筛分过程有意义的推动.后续工作中,如何提高对含噪音和间断信号的分解质量是重点研究的内容.

参考文献

- [1] Huang N E, Shen Z, 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 of London. Series A: Mathematical, Physical and Engineering Sciences, 1998, 454(1971): 903 - 995.
- [2] Lei Y, Lin J, He Z, et al. A review on empirical mode decomposition in fault diagnosis of rotating machinery[J]. Mechanical Systems and Signal Processing, 2013, 35(1): 108 - 126.
- [3] Pei S C, Hsiao Y Z, Tzeng M, et al. Uneven illumination removal and image enhancement using empirical mode decomposition[J].

- sition[J]. Journal of Electronic Imaging, 2013, 22(4): 043037 – 043037.
- [4] Li F, Hu D, Ding C, et al. InSAR Phase Noise Reduction Based on Empirical Mode Decomposition[J]. IEEE Geoscience and Remote Sensing Letters, 2013, 10(5): 1180 – 1184.
- [5] He D, Li R, Zhu J. Plastic bearing fault diagnosis based on a two-step data mining approach[J]. IEEE Transactions on Industrial Electronics, 2013, 60(8): 3429 – 3440.
- [6] 田中大, 高宪文, 等. 基于 EMD 与 LS SVM 的网络控制系统时延预测方法[J]. 电子学报, 2014, 42(5): 868 – 874.
Tian Zhongda, Gao Xianwen, et al. Time-delay prediction method of networked control system based on EMD and LS-SVM[J]. Acta Electronica Sinica, 2014, 42(5): 868 – 874. (in Chinese)
- [7] 王文波, 张晓东, 等. 基于主成分分析的经验模态分解消噪方法[J]. 电子学报, 2013, 41(7): 1425 – 1430.
Wang Wenbo, Zhang Xiaodong, et al. Empirical mode decomposition de-noising method based on principal component analysis [J]. Acta Electronica Sinica, 2013, 41(7): 1425 – 1430. (in Chinese)
- [8] Roy A, Doherty J F. Raised cosine filter based empirical mode decomposition[J]. IET Signal Processing, 2011, 5(2): 121 – 129.
- [9] Kopsinis Y, McLaughlin S. Investigation and performance enhancement of the empirical mode decomposition method based on a heuristic search optimization approach[J]. IEEE Transactions on Signal Processing, 2008, 56(1): 1 – 13.
- [10] Hong H, Wang X, Tao Z. Local integral mean-based sifting for empirical mode decomposition[J]. IEEE Signal Processing Letters, 2009, 16(10): 841 – 844.
- [11] Schwartz L M. Round-off error[J]. Analytical Chemistry, 1980, 52(7): 1141 – 1147.
- [12] Unser M, Aldroubi A, Eden M. B-spline signal processing[J]. IEEE Transactions on Signal Processing, 1993, 41(2): 821 – 833.
- [13] Rilling G, Flandrin P, Goncalves P. On empirical mode decomposition and its algorithms[A]. IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing[C]. Grado: IEEE, 2003. 8 – 11.
- [14] Rilling G, Flandrin P. One or two frequencies? The empirical mode decomposition answers[J]. IEEE Transactions on Signal Processing, 2008, 56(1): 85 – 95.

作者简介



黎 恒 男, 1982 年生于广西玉林. 博士研究生. 研究方向为时频信号处理.

E-mail: wudaomana@gmail.com



李 智 男, 1965 年生于广西灵川. 教授, 博士生导师, 研究方向为智能仪器, 测试理论与技术.