

## 经验模态分解中多种边界处理方法的比较研究

胡维平<sup>①</sup> 莫家玲<sup>②</sup> 龚英姬<sup>②</sup> 赵方伟<sup>③</sup> 杜明辉<sup>①</sup>

<sup>①</sup>(华南理工大学电子与信息学院 广州 510641)

<sup>②</sup>(广西师范大学物理与信息学院 桂林 541004)

<sup>③</sup>(西澳大学机械工程学院 珀斯 澳大利亚 WA 6009)

**摘要:** 经验模态分解(EMD)的一个关键问题是处理边界效应。尽管目前除了 Huang 申请的 NASA 专利的边界处理方法, 仍没有一个最终的解决方案, 但工程上已经提出了多种处理方法。本文实现了工程上常用的 5 种 EMD 边界处理方法: 线性外延, 多项式拟合, 镜像法, 径向基(RBF)神经网络预测和 AR 预测方法, 设计了一套消除了 EMD 处理中信号的相互作用及模式混淆影响的测试方法, 并利用准周期信号和随机信号对它们的边界效应处理结果进行了定量测试。结果表明镜像法是目前相对最优的 EMD 边界处理方法。

**关键词:** 信号处理; 经验模态分解; 边界效应; 模式混淆

中图分类号: TN911.7

文献标识码: A

文章编号: 1009-5896(2007)06-1394-05

## Methods for Mitigation of End Effect in Empirical Mode Decomposition: A Quantitative Comparison

Hu Wei-ping<sup>①</sup> Mo Jia-ling<sup>②</sup> Gong Ying-ji<sup>②</sup> Zhao Fang-wei<sup>③</sup> Du Ming-hui<sup>①</sup>

<sup>①</sup>(Department of Electronics and Communication Engineering, South China University of Technology, Guangzhou 510641, China)

<sup>②</sup>(College of Physics and Information Technology, Guangxi Normal University, Guilin 541004, China)

<sup>③</sup>(School of Mechanical Engineering, University of Western Australia, Perth, WA 6009, Australia)

**Abstract:** One of the most important problems in Empirical Mode Decomposition (EMD) applications is mitigation of the end effect. Except Huang's patented approach several methods have been proposed. However, a final solution for this problem is yet to be found. In this paper five common end effect mitigation methods of EMD have been investigated, including linear extending method, polynomial fitting extending method, mirror extrema extending method, RBF neural network prediction method and AR prediction method. With a quasi-periodical signal and a stochastic signal as the test bed a quantitative test method was proposed for elimination of the mode confusion effect of EMD. The five end effect mitigation methods were quantitatively evaluated and the comparison shows that mirror extrema extending method is the best option among the five methods.

**Key words:** Signal processing; EMD; End effect; Mode confusion

### 1 引言

1998年, 美国工程院院士Huang等人提出了一种全新的信号时频分析方法——HHT(Hilbert-Huang Transform)<sup>[1]</sup>。HHT包含两个过程: 首先对信号进行非线性的自适应分解, 称为经验模态分解(Empirical Mode Decomposition, EMD), 它把复杂信号分解成有限个本征模态函数(Intrinsic Mode Function, IMF)之和, 然后对各个IMF进行Hilbert变换去研究信号的时频能量分布。HHT除了不能事先确定分解的层数(这里我们借用小波分解的概念), 与小波分析相比, 不仅具有小波分析的全部优点, 而且具有更清晰准确的谱结构,

因而引起了信号处理相关领域的极大兴趣。

HHT分析的核心是EMD分解。因为只有对经过EMD分解后的IMF进行Hilbert变换, 得到的瞬时频率才具有物理意义。每一个IMF都需要多次筛选过程(sifting process)来实现, 而每一次筛选过程必须找到局部极大值构成的上包络和局部极小值构成的下包络并用3次样条插值来分别计算出上下包络的插值及其平均值。由于信号的端点一般不会同时是局部极值点, 因此必须对信号或极值序列进行外延, 以确保包络线能够抵达端点。包络线外延不准确会引起误差, 而任何由这种边界效应带来的误差会随着IMF分解层数的增加而向内传播, 继而“污染”整个数据序列, 使得最后的分解结果失去意义。尽管Huang已经宣称基本解决了边界问题并将该方法申请了美国宇航局专利<sup>[2]</sup>, 但他同时又指出, EMD方法面临的边界样条问题还没有完全解决, 因此解决边界端点

2005-10-24 收到, 2006-07-20 改回  
广东自然科学基金(05006593)和广西自然科学基金(0448035)资助课题

延拓问题对于HHT理论研究和应用研究都具有非常重要的意义。

国内外学者对EMD边界处理方法作了许多研究。目前工程上常用的EMD边界处理方法有镜像法<sup>[3]</sup>,神经网络预测<sup>[4]</sup>和AR预测<sup>[5]</sup>,多项式外延<sup>[6]</sup>方法等等。每种方法都对特定信号进行了定性讨论,都可以成功完成EMD分解的任务。但是,在我们能够得到Huang的可能的最佳解决方案之前,在目前其他众多的可行的EMD边界处理方法中,哪一种方法的边界效应带来的误差相对最小从而比较适合工程应用的?本文设计了一套测试方法对多种常用的EMD边界处理方法进行了定量评价测试,得到了对工程应用非常有指导意义的结果。

## 2 HHT分析及特性

### 2.1 经验模态分解(EMD)

如果定义 $x(t)$ 为原信号序列,  $\text{imf}_i(t)$ 为经EMD得到的本征模态函数(IMF),  $r_n(t)$ 为余量,则原始信号可以表示为所有的IMF及余量之和:

$$x(t) = \sum_{i=1}^n \text{imf}_i(t) + r_n(t) \quad (1)$$

对于一个实信号  $x(t)$ , 其 Hilbert 变换可由柯西主值积分(Principal Value, PV)来表达:

$$y(t) = \frac{1}{\pi} \text{PV} \int_{-\infty}^{\infty} \frac{x(\tau)}{t-\tau} d\tau \quad (2)$$

由此生成的解析函数  $z(t)$ 可表示为

$$z(t) = x(t) + iy(t) = a(t)e^{i\theta(t)} \quad (3)$$

其中解析函数  $z(t)$ 的幅值函数和相位函数分别表示为:

$$a(t) = [x(t)^2 + y(t)^2]^{1/2}, \quad \theta(t) = \arctan \left( \frac{y(t)}{x(t)} \right) \quad (4)$$

这样可以得到瞬时频率  $\omega(t)$ 的表达式:

$$\omega(t) = d\theta(t)/dt \quad (5)$$

如果对所有的 IMF 分别进行 Hilbert 变换, 原始信号  $x(t)$ 可以重新表达为

$$x(t) = \sum_{j=1}^n a_j(t) \exp \left( i \int \omega_j(t) dt \right) \quad (6)$$

### 2.2 EMD 及 IMF 的性质

Huang 通过大量的数值实验方法证实 IMF 是正交的<sup>[1]</sup>。Flandrin<sup>[7]</sup>和Huang<sup>[8]</sup>分别独立地对EMD的滤波器组特性进行了研究。他们通过 5000 次重复对独立的高斯噪声进行EMD分解,然后分别对各个IMF进行傅里叶变换并求出其功率谱,累计得到各个IMF对应的平均归一化频率-功率谱曲线如图 1 所示。图 1 中的序号 1,2,...,7 分别代表 $\text{imf}_1, \text{imf}_2, \dots, \text{imf}_7$ 对应的频率-功率谱曲线。其结果可以解释为当输入是高斯噪声时,EMD等效于一个频域的二进滤波器组的输出。这样,他们从实验上证实了EMD方法具有和二进

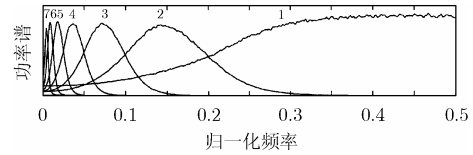


图1 EMD的滤波器组特性

离散小波分解完全类似的二进滤波器组结构的特性。

## 3 实验方法

作为HHT分析的核心步骤,EMD方法的优劣起着关键的作用。最近有学者提出了用B样条插值方法取代三次样条插值的EMD新方法,也取得了非常好的效果<sup>[9]</sup>。但它同样也面临边界效应问题。因此,EMD方法的边界处理问题是不可回避的。

### 3.1 参与测试的 EMD 边界处理方法

本文在 Matlab 下实现了 5 种工程上常用的 EMD 边界处理方法,并用于比较测试,它们分别为

- (1) 线性外延方法 利用最靠近边界上的两个极值点的线性外延以确定边界上的等效极值点。
- (2) 端点镜像方法<sup>[3]</sup> 以两端的边界为对称将原信号镜像,从而得到完整的上下包络插值曲线。
- (3) AR预测方法<sup>[5]</sup> 用AR预测方法对数据序列的两端分别进行逐点预测,并逐步外延以确定边界之外的极值点。
- (4) 多项式拟合方法<sup>[6]</sup> 分别利用最邻近端点的 3 个极值点进行二次三项式拟合以确定在边界上的极值点位置。
- (5) 神经网络预测方法<sup>[4]</sup> 用径向基函数(RBF)神经网络对数据序列的两端分别进行逐点预测,并逐步外延以确定边界之外的极值点。

EMD方法中的三次样条插值需要上下(局部极大极小)包络的数据,故在两端边界上需要分别确定 2 个极大和 2 个极小位置。在一般涉及预测的方法中,都会利用视数据的端点为一个极值点的处理方法,这样的方法一般被认为是一种较为鲁棒的选择<sup>[10]</sup>。但是,当最邻近端点的极值点和端点非常靠近时,上述的处理方法就可能带来较大的边界效应。针对这个问题,本文在所有涉及该问题的程序里作了修正,即当最邻近端点的极值点和端点非常靠近时(小于局部极值序列极值点平均距离的 40%),用相应的方法对数据进行外延以确定端点上的极值位置,否则还是利用该端点作为边界上的一个极值点。实验证明,这一处理方法可以较好地改善边界效应。

### 3.2 测试信号及评价方法

本文分别用准周期信号和随机信号对这 5 种方法进行了比较研究。

3.2.1 准周期信号 评价EMD方法的关键在于研究其IMF的分解特性。因为EMD方法的最终目的是提供准确的幅度和频率谱分析,因此研究其经EMD得到的IMF的幅度和频率谱畸变就可以正确评价EMD。考虑到IMF的正交性和EMD二进滤波器组结构的特性,当两信号频率足够靠近时(两信号

频率比值 $f_2/f_1 > 0.5$ , 在 $f_2 < f_1$ 的约束条件下), EMD将会把这两个信号分解到同一个IMF中, 而如果考虑到EMD的模式混淆影响, 落入单个IMF的信号频率带可能超出上述的范围(两信号频率比值 $f_2/f_1 < 0.5$ )<sup>[11]</sup>。因此为尽可能正确地评价每种EMD方法的边界效应, 就必须尽可能地排除实验中EMD的信号相互作用及模式混淆的影响。再考虑到HHT分析的对象的部分可以看成准周期信号, 例如器械振动信号和语音信号, 因此本文使用信号中幅度/频率成线性变化的AM/FM组合信号来模拟准周期信号, 如果设计的AM和FM信号的频率合适且其频域距离足够大, 理论上就可以避开信号的相互作用和模式混淆影响, 并可以用EMD分解得到 $imf_1$ 和 $imf_2$ 基本不失真地去分别恢复AM和FM的原信号。通过研究EMD分解后恢复的AM/FM信号中线性成分的畸变, 即可对不同的EMD边界处理方法带来对IMF信号的影响给出定量的评估:

AM 信号:

$$Y_1(n) = 4 \cdot \sin\left(\frac{\pi}{3600}n\right) \cdot \sin\left(\frac{b \cdot \pi}{60}n\right), \quad b \in [8, 15] \quad (7)$$

FM 信号:

$$Y_2(n) = 2 \cdot \cos((n/50)^2 \cdot \pi/5) \quad (8)$$

测试信号  $Y(n)$  为 AM 和 FM 信号的叠加<sup>[12]</sup>:

$$Y(n) = Y_1(n) + Y_2(n) \quad (9)$$

假定原信号为式(9)表示的信号  $Y(n)$ , 其EMD分解后的IMF之和可以表达为式(1), 按式(3)对 $imf_i$ 进行Hilbert变换, 分别得到瞬时频率  $\omega_i(n)$  和幅度函数  $a_i(n)$ ; 而对式(9)表示的信号  $Y_1(n)$  和  $Y_2(n)$  直接做Hilbert变换(两个信号本身满足IMF条件), 瞬时频率和幅度函数分别记为  $\omega_{yi}(n)$ ,  $a_{yi}(n)$ , 其中  $i=1$  对应于AM信号,  $i=2$  时对应于FM信号。它们的理想的Hilbert谱特征是水平线或斜线, 则误差评价公式为

$$\begin{aligned} \text{err}_{-a_i} &= \frac{\sum_n (a_i(n) - a_{yi}(n))^2}{\sum_n (a_i(n))^2} \\ \text{err}_{-\omega_i} &= \frac{\sum_n (\omega_i(n) - \omega_{yi}(n))^2}{\sum_n (\omega_i(n))^2} \end{aligned} \quad (10)$$

综合评价公式可表示为

$$\text{ev} = \sqrt{(\text{er}_{-a_1})^2 + (\text{er}_{-\omega_1})^2 + (\text{er}_{-a_2})^2 + (\text{er}_{-\omega_2})^2} \quad (11)$$

3.2.2 随机信号 为全面评价 EMD 方法的边界效应, 本文同时采用了 200 次随机信号分别对上述 5 种 EMD 方法进行了比较研究。具体为: 原始数据长度为 1512, 取分解后的 1512 长度数据的中部 512 点(定义为  $imf_{yi}(n)$ )作为无边界效应的对照组与中段 512 点数据的直接 EMD 分解的结果(定义为  $imf_i(n)$ )进行比较。考虑到单个随机信号的频域分布没有规律, 故不进行 Hilbert 变换而在时域进行比较, 其相对均方误差评价公式为

$$\text{err}_{-imf_i} = \frac{\sum_n (imf_i(n) - imf_{yi}(n))^2}{\sum_n (imf_{yi}(n))^2} \quad (12)$$

考虑到EMD分解的二进滤波器组结构的特性, 其通道能量可以视为随分解序号  $i$  ( $imf_i$ ) 的增加按二进方式递减, 并且考虑到EMD分解过程中的误差是继承的并随  $i(imf_i)$  的增加被放大, 因此本文对不同的  $imf_i$  的误差予以不同的权重, 故其综合误差评价公式定义为

$$\text{ev} = \sqrt{4 \cdot (\text{er}_{-imf_1})^2 + 2 \cdot (\text{er}_{-imf_2})^2 + 1 \cdot (\text{er}_{-imf_3})^2} \quad (13)$$

### 4 实验结果和讨论

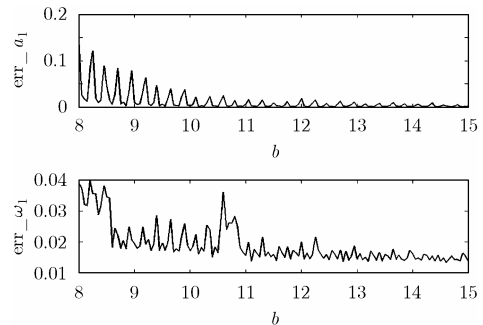
应用上述定义的测试信号, 分别对前面提到的 5 种 EMD 边界处理方法进行了测试。

#### 4.1 准周期信号测试

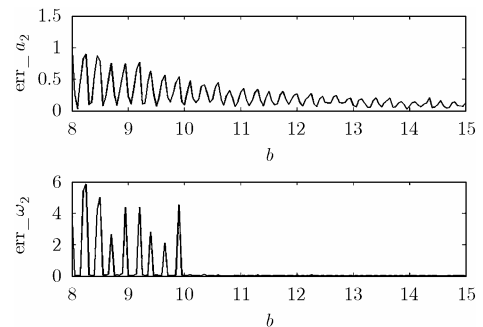
4.1.1 误差随  $b$  值变化 5 种 EMD 程序皆在 Matlab 下用三次样条插值实现, 采样相同的停止条件  $SD=0.1$ <sup>[1]</sup>; 对式(9)定义的信号进行了 EMD 分解, 其中式(7)中随频率变化参数  $b$  取值范围 8 至 15, 步长 0.05, 每一个方法都得到的  $\text{err}_{-a_1}$ ,  $\text{err}_{-\omega_1}$ ,  $\text{err}_{-a_2}$ ,  $\text{err}_{-\omega_2}$  的变化曲线。限于篇幅, 这里仅给出了方法 2、方法 5 的误差随  $b$  值变化的曲线, 如图 2, 图 3 所示。

从图 2、图 3 的趋势(其余的几种方法也有完全类似的趋势)可以清晰看出, 当频率变化参数  $b > 11$  时, 各误差趋向平稳的较小的水平, 此时可以视为两信号间相互影响和模式混淆的影响已降为最低。因此后续的实验中文献仅将频率变化参数  $b=11-15$  范围的误差结果考虑到误差平均值的计算中。

4.1.2 EMD 的信号相互影响及模式混淆 4.1.1 节的结果说明经 EMD 恢复信号的失真与 AM, FM 两信号频谱在频域



(a) 方法 2 恢复的 AM 信号误差曲线



(b) 方法 2 恢复的 FM 信号误差曲线

图 2

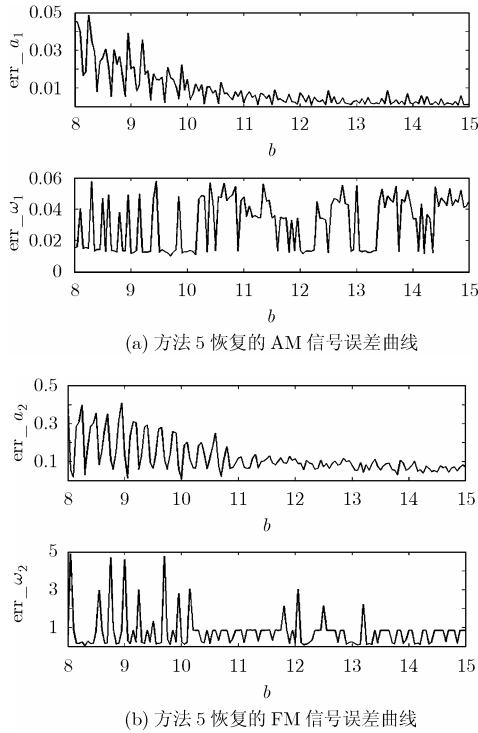


图 3

上的距离有关。为直观地给出说明，本文选择了 EMD 方法 1 的程序分别对频率变化参数  $b=6.5; 7.5; 9$  时的 3 种情况对式 (9) 定义的信号进行了 EMD 分解，给出信号恢复比较结果如图 4 所示。

图 4 给出了当  $b=6.5; 7.5; 9$  变化时，经 EMD 方法恢复 AM/FM 信号时的对应及比较结果。图 4 也以图示的方法证实了 4.1.1 节的误差趋向。结合 4.1.1 节的结果，图 4 充分说

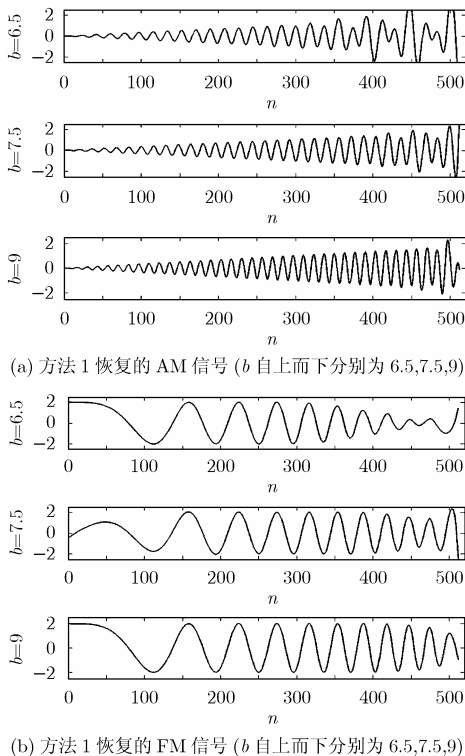


图 4

明了经 EMD 方法恢复的信号所产生的畸变，并不像有的学者认为的完全由边界效应引起<sup>[12]</sup>，合理的解释应该是，当欲处理的两个信号频率比较靠近时，考虑到 EMD 的频域滤波器组特性，两个信号在频域的谱线可能会落在 EMD 对应滤波器组的一个通道里，因此，在 IMF 分离时频率相对较低信号的相对高频部分能量就被强制划入先行分离的 IMF 的信号中，这就是图 4 中频率变化参数  $b=6.5$  时信号右端的情况。随着两个信号谱线距离的增加，被相对高频 IMF 信号强制划走的能量逐步减少(图 4 中  $b=7.5$  时的情况)，并随着两个信号谱线距离的继续加大((图 4 中  $b=9$  时的情况))，两信号间的相互影响逐步减少以致最终消失，这时经 EMD 恢复的 AM 和 FM 信号与原信号的畸变相对最小并且主要来自于边界效应而非 EMD 方法中的信号相互影响和模式混淆效应。

**4.1.3 5 种 EMD 方法的平均结果** 根据 4.1.1 节及 4.1.2 节的实验结果，设计了评估 EMD 边界处理方法在准周期信号下平均误差的计算方法。即对 5 种方法分别进行误差随频率变化参数  $b$  变化的分析，且仅计算每种方法的频率变化参数  $b=11\sim 15$  范围内的平均值和按式(11)的总误差  $ev$ ，这样得到的平均误差结果基本排除了信号间的相互影响和模式混淆，能够较真实地反映边界处理方法对 EMD 的影响。得到比较结果如表 1 所示。

表 1 5 种 EMD 边界处理方法的准周期信号误差平均结果

	err_ $a_1$	err_ $\omega_1$	err_ $a_2$	err_ $\omega_2$	ev
(1)	0.0036	0.0793	0.0308	0.5954	0.6015
(2)	<b>0.0038</b>	<b>0.1417</b>	<b>0.0157</b>	<b>0.0303</b>	<b>0.1458</b>
(3)	0.0038	0.0381	0.1037	1.0889	1.0945
(4)	0.0046	0.0977	0.0144	0.1791	0.2206
(5)	0.0033	0.0808	0.0342	0.6669	0.6726

4.2 随机信号测试

应用 3.2.2 节定义的测试信号和评价方法，本文分别对 5 种 EMD 边界处理方法进行了随机信号的测试。得到比较结果如表 2 所示。综合表 1 及表 2 的实验结果清晰地表明了镜像法(方法 2)在所有的的方法中误差最小而成为最佳选择。

表 2 5 种 EMD 边界处理方法的随机信号误差平均结果

	err_ imf1	err_ imf2	err_ imf3	ev
(1)	0.0153	0.1375	0.3239	0.2070
(2)	<b>0.0123</b>	<b>0.1255</b>	<b>0.2880</b>	<b>0.1568</b>
(3)	0.0168	0.1611	0.3782	0.2883
(4)	0.0244	0.2085	0.4443	0.4888
(5)	0.0153	0.1537	0.3541	0.2858

5 结束语

在充分考虑了 EMD 正交性和频域滤波器组特性的基础上，本文设计了 AM+FM 的复合信号和一套基本排除了 EMD 分解后信号相互影响和模式混淆影响的比较测试方法，

对5种EMD边界处理方法进行了比较测试;为更全面地衡量各方法边界处理的性能,同时本文也设计了随机信号参与测试。结果表明,镜像法算法(方法(2))无论在准周期信号还是在随机信号中都以较明显的优势成为首选。因为无论对于长短周期信号,镜像法基本上都能够很好地保留边界上邻域的信号特征,在准周期信号甚至随机信号的处理中都具有较好的鲁棒性;在预测方法中,神经网络据认为是较好的选择<sup>[4]</sup>,在这里,本文选择了神经网络的径向基函数RBF方法而不是文献[4]中的单层神经网络,因为RBF是一种性能良好的前向网络,它具有快捷的全局收敛性质和最佳逼近性能,非常适合系统的实时辨识和预测。在本实验中神经网络RBF(方法(5))比传统的AR预测(方法(3))在准周期信号上确有较大改进,但在随机信号应用上仅仅有极微弱的优势,应该属同一水平,由于预测原信号序列方法的局限性(需要预测的长度和信号的周期属性相关),预测的准确度和预测的长度成反比,以及对随机信号准确预测的困难性,因此无论是处理渐变准周期信号还是随机信号时与镜像算法(方法(2))还是有一定的差距;而多项式延拓(方法(4))在处理准周期信号时尽管表现比较出色,但在随机信号上相对最差的表现说明使用多项式方法的EMD并不合适于无规律的信号处理。至于线性外延法(方法(1))尽管相对比较粗糙,但它在处理渐变准周期信号和随机信号时表现出了令人惊讶的稍优于2种预测方法的性能,鉴于它算法的简单高效,可望在实时处理的场合得到应用。

本文用准周期信号及随机噪声信号对5种改进的EMD边界处理方法进行了比较研究。结果表明:(1)镜像法(方法(2))无论在准周期信号还是在随机噪声信号应用方面以较大的优势成为首选;(2)因为本文的实验基本覆盖了传统的边界处理方法,因此在得到理论突破之前(例如Huang的特征波的EMD边界处理专利方法),继续研究传统的预测延拓等的并用特例来证实的“新方法”可能不会得到更有意义的结果;(3)有希望的边界处理研究方向应该是在针对EMD的分解特性下进行。

### 参 考 文 献

- [1] Huang N E, Shen Z, and Long S R. The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis [J]. *Proc. R. Soc. Lond. A*, 1998, 454: 903-995.
  - [2] Huang N E. Computer implicated empirical mode decomposition method, apparatus, and article of manufacture [P]. U.S. Patent Pending 1996.
  - [3] 黄大吉, 赵进平, 苏纪兰等. 希尔伯特-黄变换的端点延拓. *海洋学报*[J], 2003, 25(1): 1-11.
  - [4] 邓拥军, 王伟, 钱成春等. EMD方法及Hilbert变换中边界问题的处理[J]. *科学通报*, 2001, 46(3): 257-263.
  - [5] 张郁山, 梁建文, 胡埭贤. 应用自回归模型处理EMD方法中的边界问题[J]. *自然科学进展*, 2003, 13(10): 1054-1059.
  - [6] 刘慧婷, 张旻, 程家兴. 基于多项式拟合算法的EMD端点问题的处理[J]. *计算机工程与应用*. 2004, 40(16): 84-86.
  - [7] Flandrin P, Gabriel Rilling G, and Goncalves P. Empirical mode decomposition as a filter bank. *IEEE Signal Processing Letters*, 2004, 11(2): 112-114.
  - [8] Wu Z and Huang N E. A study of the characteristics of white noise using the empirical mode decomposition method. *Proc. R. Soc. London A*, 2002, submitted for publication.
  - [9] Chen Qiu-hui, Huang Norden, Riemenschneider S, and Xu Yuesheng. A B-spline approach for empirical mode decompositions. *Advances in Computational Mathematics*, 2006, 24(1-4): 171-195.
  - [10] 陈忠, 郑时雄. EMD信号分析方法边缘效应的分析[J]. *数据采集与处理*. 2003, 18(3): 114-118.
  - [11] Deering R and Kaiser J F. The use of a masking signal to improve empirical mode decomposition. *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'05)*. Philadelphia, Pennsylvania, USA, March 2005, Vol.4: 485-488.
  - [12] 程军圣, 于德介, 杨宇. 基于EMD的信号瞬时特征的小波分析方法[J]. *地震工程与工程振动*, 2004, 24(2): 181-186.
- 胡维平: 男, 1963年生, 博士生, 副教授, 研究方向为生物医学信号处理和图像处理等。
- 杜明辉: 男, 1964年生, 博士生导师, 教授, 研究方向为医学图像处理、信号处理、生物医学工程等。