

经验模式分解端点效应抑制方法的研究

郭 迪,沈洋洋,尹 兵

(南京邮电大学 电子科学与工程学院,江苏 南京 210003)

摘 要:针对经验模式分解法(Empirical Mode Decomposition, EMD)中的端点效应严重影响算法精度的情况,为了减小端点效应,文中提出一种新的先匹配后镜像延拓方法。该方法借鉴镜像对称延拓与波形特征匹配延拓,将信号先进行波形特征匹配延拓再进行镜像对称延拓。利用仿真数据对先匹配后镜像延拓的方法进行了验证,并与原有几种延拓方法在不同评价指标下进行了对比。结果表明,先匹配再镜像延拓后,信号的包络线发生畸变最小,同时新方法分解得到的 IMF 精度较高,正交性好。先匹配后镜像延拓方法提高了经验模式分解的精度,能更有效地抑制经验模式分解法中的端点效应。

关键词:经验模式分解;延拓;端点效应;镜像法;波形特征匹配法

中图分类号:TP301

文献标识码:A

文章编号:1673-629X(2016)03-0089-04

doi:10.3969/j.issn.1673-629X.2016.03.021

Research on Method for End Effects Reduction of Empirical Mode Decomposition

GUO Di, SHEN Yang-yang, YIN Bing

(College of Electronic Science and Engineering, Nanjing University of Posts and Telecommunications,
Nanjing 210003, China)

Abstract: The precision of Empirical Mode Decomposition (EMD) is reduced greatly by its end effects, so a new end extending method combining the waveform feature matching extending method and mirror extending method is presented. The waveform feature matching extending method and mirror extending are used for reference in this new method, the signal is extended by waveform feature matching method at first and then extended by mirror method. A simulation signal is applied to test the performance of the new method, and a comparison under different evaluating indicators between the new method and old methods is made and analyzed. The result shows that the signal envelope has minimum distortion, at the same time, the IMF has high precision and good orthogonality decomposed by new method. The proposed improved extending method can improve the precision of EMD and restrain the end effect effectively.

Key words: EMD; extending; end effect; mirror method; waveform feature matching method

0 引言

经验模式分解法(Empirical Mode Decomposition, EMD)是由 N. E. Huang^[1]等提出的新型自适应信号时频处理方法,它把复杂信号分解为模态函数(Intrinsic Mode Function, IMF)之和,同时通过希尔伯特变换对相位微分求解得到瞬时频率,赋予了瞬时频率实际的物理意义。由于分解基于信号的局部特征时间尺度,使得分解具有自适应性与高效性,特别适用于分析非线性、非平稳信号。在 EMD 分解过程中,需要根据信号提供的极值点信息进行三次样条插值,以便画出信号的上、下包络线,信号的左右两个端点可能不是极

值点,这样会使拟合时产生误差。所得到的上、下包络线在信号的边界附近会产生大幅摆动,并且这种拟合误差会随着分解逐渐向信号内部传播,导致得到的 IMF 分量的波形严重失真。

文中在对 EMD 分解端点效应深入分析的基础上,提出一种波形匹配法和镜像法相结合的延拓算法,并通过实验仿真证明了其有效性。

1 EMD 基本原理与端点效应

1.1 EMD 基本原理

EMD 分解最基本的步骤是将非线性、非平稳的信

号分解为表征信号特征时间尺度的 IMF, 它满足单分量信号的物理解释。任何 IMF 必须满足两个条件:

(1) 整个数据的极值点的个数和零点的个数相等或是相差一个。

(2) 在任意时刻, 由局部极值点形成的上、下包络线的平均值为零, 即上、下包络线关于时间轴局部对称。

EMD 分解步骤如下: 找出信号全部局部极大值点与局部极小值点的位置坐标, 然后用三次样条插值法分别得到局部极大值包络线与局部极小值包络线, 求上、下包络线坐标的平均值, 得到一条新曲线 (均值线) 并记为 m_1 , 记其与原始信号 $x(t)$ 的差值为 h_1 。

$$h_1 = x(t) - m_1 \quad (1)$$

判断 h_1 是否满足 IMF 的条件, 如果满足, 那么 h_1 就是 $x(t)$ 的第一个 IMF 分量; 如果不满足, 用 h_1 代替原始数据, 如果不满足, 则重复前述步骤直到 $h_{1k} = h_{1(k-1)} - m_{1k}$ 满足 IMF 的条件。 h_{1k} 为信号 $x(t)$ 的第一个本征模态函数。记 $c_1 = h_{1k}$, c_1 为输入信号中频率最高的分量。再对余量 $r_1 = x(t) - c_1$ 进行分解, 得到 c_2 , 重复上述筛选过程直到:

(1) 当分量 c_n 或残余分量 r_n 比预定值小时停止循环;

(2) 当 r_n 变为单调函数, 从中不能再筛选出 IMF 分量为止。

分解停止后, 得到信号 $x(t)$ 的全部本征模态函数。这样就有:

$$x(t) = \sum_{i=1}^n c_i + r_n \quad (2)$$

以式(3)信号为例, 采样频率取 $f_s = 2\,048$ Hz。

$$x = 5\sin(2\pi * 5t) + 5\sin(2\pi * 8t), t \in [1, 2] \quad (3)$$

按照如上步骤, 对信号进行经验模态分解, 得到 3 个 IMF 分量和一个 res 分量。如图 1 所示, IMF₁ 至 IMF₃ 分别代表从原始信号分解出来的第 1 至第 3 个本征模态函数, 其频率由高到低排列; res 代表残余分量。

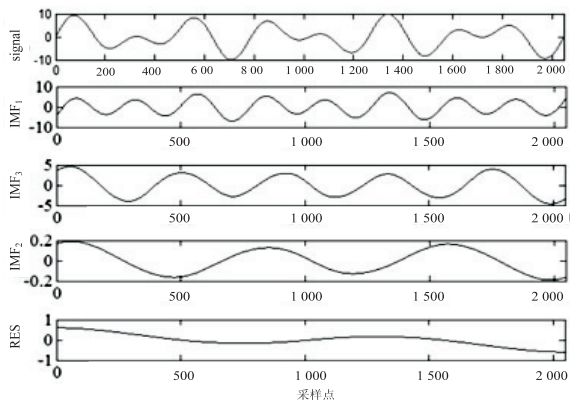


图 1 原始信号及各个 IMF 分量

然后求每个本征模态函数的希尔伯特谱, 并将它们相叠加, 得到原始信号的希尔伯特谱。如图 2 所示, 图中的 3 条曲线从上到下分别为 IMF₁ 至 IMF₃ 的希尔伯特谱。

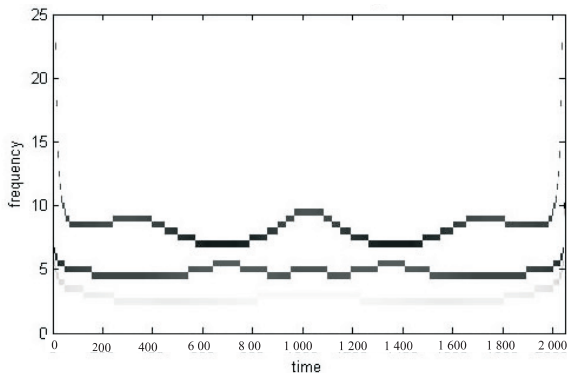


图 2 原始信号的希尔伯特谱

1.2 端点效应

经验模态分解过程中, 需要根据信号提供的极值点信息进行三次样条插值以便画出信号的上、下包络线, 由于不确定信号的左、右两个端点是否为极值点, 会使拟合时产生误差。所得到的上包络线 and 下包络线在信号的边界附近产生大幅摆动, 并且这种拟合误差会随着分解逐渐向信号内部传播, 导致得到的 IMF 分量的波形严重失真, 尤其是原始信号长度比较短时, 会对经验模态分解的效果造成严重影响。

图 1 为原始信号经过分解后得到的各个 IMF 分量, 可见 IMF₂ 的两端有明显的震荡。这是 HHT 的第一个端点效应^[2-5]。而且在对各个 IMF 分量进行 Hilbert 谱变换时, 由端点效应引起的误差从数据的中心向外逐渐扩散, 希尔伯特谱的两端也产生突变和振荡, 如图 2 所示。这是 HHT 中的第二个端点效应。端点效应的存在不利于信号的分析 and 处理。

2 抑制端点效应的常用方法

2.1 镜像对称延拓

镜像延拓的思路是在原始信号的两端某位置各放一面镜子, 得到原始信号在镜中的像, 镜子中原始信号的像关于镜子与原数据对称。

镜像对称延拓法关于镜子位置的选择, 分为以下几种情况:

对采样起始点:

(1) 第一个极值点是极大值时, 如果第一个采样点大于第一个极小值点, 以第一个极大值为对称中心; 如果第一个采样点小于第一个极小值点, 则将其作为极小值点, 并将其作为对称中心。

(2) 第一个极值点是极小值时, 如果第一个采样点小于第一个极大值点, 以第一个极小值为对称中心;

如果第一个采样点的值大于第一个极大值点,则将其作为极大值点,并将其作为对称中心。

- 对采样结束点:
- (1)最后一个极值点是极小值时,如果最后一个采样点的值小于最后一个极大值,将其作为对称中心;如果最后一个采样点的值大于最后一个极大值,则将其作为极大值点,并将其作为对称中心。
- (2)最后一个极值是极大值时,如果最后一个采样点的值大于最后一个极小值,则把最后一个极大值点作为对称中心;如果最后一个采样点的值小于最后一个极小值,则将其作为极小值点,并将其作为对称中心。

2.2 波形特征匹配延拓法

波形特征匹配延拓法基于波形匹配的思想,其基本思路是从信号内部寻找一个与边界处特征最相近的波形作为延拓数据的起点,按照所需延拓长度从信号内部取出匹配波形数据前面(或后面)相应长度的数据连接到边界处,完成延拓。通过波形特征匹配延拓得到的数据符合信号的自然趋势,能有效抑制 Hilbert 变换的端点效应。

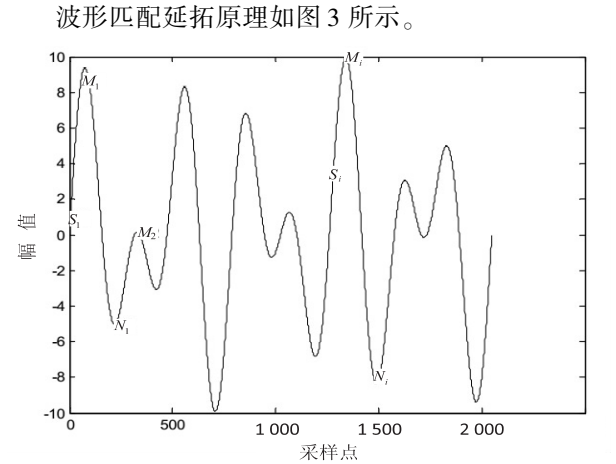


图 3 波形特征匹配延拓示意图

以信号左边界第一个极值点为极大值为例, M_i , $N_i(i = 1, 2, \cdots)$ 分别为波形的极大值、极小值点,其时间点为 tm_i, tn_i , S_1 为第一个采样点,波形特征匹配延拓法将 $S_1 - M_1 - N_1$ 作为特征波形,在全部波形中找到与 $S_1 - M_1 - N_1$ 构成的三角形最接近的波形为匹配波形 $S_k - M_k - N_k$,从匹配波形 $S_k - M_k - N_k$ 开始,向前对波形数据进行延拓。

- 具体步骤如下:
- (1)根据 S_1 与边界处极值点的时间相对位置寻找对应点 S_i ,其时间坐标为:

$$ts_i = \frac{tm_1 \cdot tn_i - tn_1 \cdot tm_i}{tm_1 - tn_1} \tag{4}$$

如果 ts_i 不在采样点上,可以用线性插值法计算其

- 精确值 s_i 。时间点和采样点可以互相转化。
- (2)计算匹配误差。
- $$E_i = |S_i - S_1| + |N_i - N_1| + |M_i - M_1| + |M_{i+1} - M_2| \tag{5}$$
- 式中, $|M_{i+1} - M_2|$ 为匹配波形的趋势项,它反映了特征波形相对后一个极值点的位置。
- (3)选取 E_i 值最小的波形为匹配波形,如果有多个 E_i 相等且为最小值,则取与左端点距离最远的波形为匹配波形。
- (4)从匹配波形的右端的前一点开始,将实际波形复制到 S_1 前,根据需要选择延拓点数,如果信号中 S_i 前数据点个数少于需要延拓的点数,可反复延拓此段波形。
- (5)采用同样的原理对右边界进行延拓。

3 波形匹配法和镜像法相结合的新算法

3.1 算法步骤

对以上两种延拓方法进行理论研究,可发现:镜像法虽然对端点是否为极值点进行了判断,但在以内部极值点为对称中心对信号进行延拓时,需要截断原始数据,这对短数据序列是不适用的;波形匹配法在出现信号中 S_i 前数据点个数小于需要延拓的点数这种情况时,给出的解决方案是反复延拓此段波形。这会造成信号幅值的突变。为此,文中提出一种新的波形延拓方法:波形特征匹配法和镜像法相结合的新算法。

- 算法步骤如下:
- (1)先对原始信号 x_1 的两端进行波形匹配延拓,得到信号 x_2 。延拓的长度为端点到第 2 个极值点的长度。
- (2)对 x_2 进行镜像延拓,得到信号 x_3 。镜子的位置选在 x_2 最外端的极值点处。
- 下文中将波形匹配法和镜像法相结合的算法称为先匹配后镜像延拓算法。
- 将式(3)中的原始信号进行先匹配后镜像延拓,得到的延拓信号如图 4 所示。

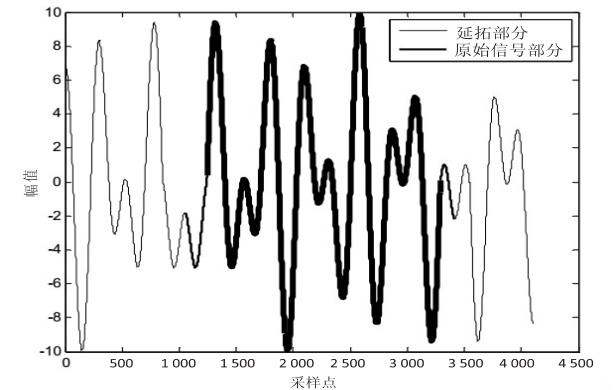


图 4 先匹配后镜像延拓效果图

3.2 端点效应评价

一些学者对于端点效应的评价标准做了研究。文献[6]提出通过计算 EMD 后信号的总能量来评估端点效应的影响程度;文献[7]提出利用 IMF 与对应信号分量的相关系数来评价 EMD 得到的 IMF 精度;文献[8]提出考虑比较不同算法的运行时间。

文中采用下述 4 个指标来评价多种端点效应处理方法的效果:

(1)从能量的角度,端点效应的实质是信号的包络线发生畸变,而产生了一些虚假成分,各 IMF 总能量会相应增加,因此,可以比较 EMD 分解前后的能量来评估端点效应的影响程度^[9-14]。

$$\theta = \frac{\left| \sqrt{\sum_{i=1}^n \text{RMS}_i} - \text{RMS}_{\text{original}} \right|}{\text{RMS}_{\text{original}}} \quad (6)$$

其中: $\text{RMS}_{\text{original}}$ 为原信号有效值; RMS_i 为第 i 个 IMF 的有效值; n 为 IMF 总个数(包括 EMD 的残余分量); θ 的值越小,表示端点效应的影响越小。

表 1 端点处理方法性能比较

延拓方法	θ	ρ_{IMF_1}	ρ_{IMF_2}	ρ_{IMF_3}	t/s	ORT
镜像延拓	0.070 5	0.887 3	0.696 9	0.070 1	0.863	0.071
波形匹配延拓	0.020 9	0.889 8	0.755 2	0.025 6	0.327	0.103 9
先匹配再镜像延拓	0.017 9	0.886 6	0.720 9	0.091 5	0.566	0.075 7

对式(3)的原始信号使用先匹配再镜像的延拓方法,然后进行 EMD 分解,并根据采样点数截取延拓组 IMF 分量中对应着原始信号的部分,得到的结果如图 5 所示。

从图中可看出,先匹配再镜像延拓法分解信号的效果非常好,很好地抑制了 EMD 分解中出现的端点效应,把端点效应造成的误差限制在了信号端点的外侧,得到的 IMF 分量比较理想。

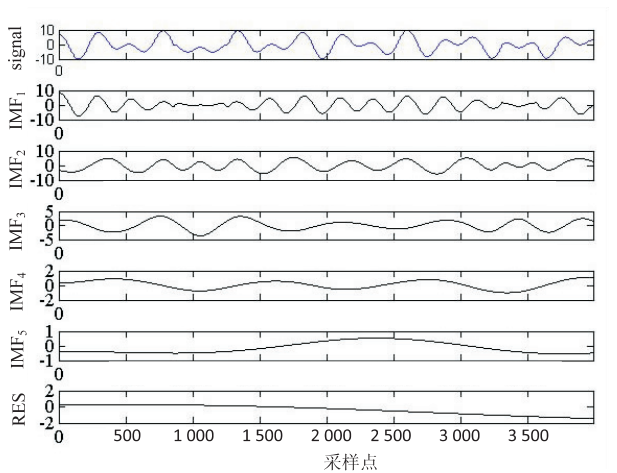


图 5 先匹配再镜像延拓信号的 IMF

由此可见,先匹配再镜像延拓方法是一种有效的延拓方法,当原始信号长度较短时,使用此方法对其进

(2)比较 EMD 分解得到的模态分量和原信号相应分量的相关系数 ρ 来评价 IMF 的精度。

(3)比较各个方法的运算时间 t (单位为 s),在保证处理效果的同时,要避免算法过于复杂。

(4)EMD 分解的正交性要满足实际意义,即:

$$x^2(t) = \sum_{i=1}^{n+1} c_i^2(t) + 2 \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} c_i(t) c_j(t) \quad (7)$$

如果分解的信号满足正交性,则式(7)右边第二项为零,因此可用正交性系数 ORT 来衡量分解精度,ORT 值越小表示时间序列的 IMF 分量正交性越好。

$$\text{ORT} = \sum_{i=0}^T \left(\sum_{j=1}^{n+1} \sum_{k=1}^{n+1} c_j(t) c_k(t) / x^2(t) \right), j \neq k \quad (8)$$

按照上述评价指标,求得的几种端点处理方法的评价指标如表 1 所示。

从表中可以看出,从能量角度,先匹配再镜像延拓后,信号的包络线发生畸变最小,同时新方法分解得到的 IMF 精度较高,正交性好。

行延拓,然后再使用 EMD 方法对其进行分解,可有效抑制端点效应的产生。

4 结束语

介绍了两种常用的延拓方法,并提出了一种先匹配再镜像延拓新方法,给出了具体操作步骤,并进行了实例仿真。

给出了端点效应评价的 4 种指标,对先匹配再镜像延拓进行了端点效应评价,并与原有的镜像匹配延拓与波形匹配延拓在不同评价指标下进行了详细比较。

通过比较可以证明,使用先匹配再镜像延拓的方法,增加了 EMD 分解得到的 IMF 的正交性,提高了 EMD 分解精度,改进了原有延拓方法的不足。

参考文献:

[1] Huang N E, Wu M C, Long S R, et al. A confidence limit for the empirical mode decomposition and the Hilbert spectral analysis[J]. Proc of Roc Soc London, Ser A, 2003, 459: 2317–2345.

[2] 高 强, 段晨东, 赵艳青, 等. 基于最大相关波形延拓的经验模式分解端点效应抑制方法[J]. 振动与冲击, 2013, 32

双投影归一化特征的识别时间长。

4 结束语

文中提出了一种基于多投影归一化特征的脱机手写藏文字符特征提取的方法。使用 KNN 分类器,对 1 747 个待测样本进行测试。实验结果表明,与双投影归一化特征提取相比,多投影归一化特征提取的识别效果更好。

参考文献:

- [1] 王建平,王梦泽. 三枝粗糙集和变粒度原理的手写体汉字识别[J]. 计算机工程与应用,2014,50(22):223-227.
- [2] 朱程辉,甘恒,王建平. 基于 FSVM 脱机手写体汉字分类识别研究[J]. 计算机工程与应用,2014,50(23):189-193.
- [3] 姜文,刘立康. 基于实值 Gabor 能量特征的手写体维文字符识别[J]. 计算机工程与设计,2013,34(9):3273-3278.
- [4] Huang Heming, Da Feipeng. Wavelet and moments based off-line handwritten Tibetan character recognition[J]. Journal of Information and Computational Science, 2013, 10(6):1855-1859.
- [5] Huang Heming, Da Feipeng, Han Xiaoxu. Wavelet transform and gradient direction based feature extraction method for off-line handwritten Tibetan letter recognition[J]. Journal of Southeast University (English Edition), 2014, 30(1):27-31.
- [6] 李永忠,王玉雷,刘真真. 藏文印刷体字符识别技术研究

[J]. 南京大学学报:自然科学,2012,48(1):55-62.

- [7] Ma L L, Liu H D, Wu J. MRG-OHTC database for on-line handwritten Tibetan character recognition[C]//Proc of international conference on document analysis and recognition. Beijing, China:[s. n.], 2011:207-211.
- [8] Huang H M, Da F P. Sparse representation-based classification algorithm for optical Tibetan character recognition[J]. Optik-international Journal for Light and Electron Optics, 2014, 125(3):1034-1037.
- [9] 王维兰. 藏文基本字符识别算法研究[J]. 西北民族学院学报:自然科学版,1999,20(3):20-23.
- [10] 周治紧,李玉鑑. 基于投影归一化的字符特征提取方法[J]. 计算机工程,2006,32(2):197-199.
- [11] 盛业华,张卡,叶春,等. 基于灰度投影的数字近景摄影立体影像匹配[J]. 光学学报,2005,25(12):1623-1628.
- [12] 邵楠,张科. 基于投影熵特征的图像识别算法[J]. 计算机应用,2013,33(10):2874-2877.
- [13] Ma Long, Xiao Baohua, Wang Chunheng. Sparse representation based on K-nearest neighbor classifier for degraded Chinese character recognition[C]//Proc of PCM. Berlin: Springer, 2010:506-514.
- [14] Wang Mengzuo, Zhang David, Wang Kuanquan. On kernel difference-weighted k-nearest neighbor classification[J]. Pattern Analysis and Applications, 2008, 11(3):247-257.
- [15] 卜凡军,钱雪忠. 基于向量投影的 KNN 文本分类算法[J]. 计算机工程与设计,2009,30(21):4939-4941.

(上接第 92 页)

- (2):62-66.
- [3] 杨斌,陈桂明,刘建友. 基于标准互相关函数的经验模态分解端点效应处理方法[J]. 机械工程学报,2013,49(5):63-68.
- [4] 白春华,周宣赤,林大超,等. 消除 EMD 端点效应的 PSO-SVM 方法研究[J]. 系统工程理论与实践,2013,33(5):1298-1306.
- [5] 任达千,杨世锡,吴昭同,等. LMD 时频分析方法的端点效应在旋转机械故障诊断中的影响[J]. 中国机械工程,2012,23(8):951-956.
- [6] Benkherratt M, Bauguerra R, Choufa T. Removal of ocular artifacts from related evoked potentials using VSSLMS adaptive filter[C]//Proc of international conference on "computer as a tool". Warsaw, Poland:[s. n.], 2007:349-352.
- [7] Venkataramanan S, Kalpakam N V. A novel wavelet based technique for detection and de-noising of qcular artifact in normal and epileptic electroencephalogram[C]//Proc of IC-

CCAS. [s. l.]:[s. n.], 2004:27-29.

- [8] Kumar P S. Removal of ocular artifacts in the EEG through wavelet transform without using an EOG reference channel[J]. Int J Open Problems Compt Math, 2008, 1(2):188-200.
- [9] 孟宗,顾海燕,李姗姗. 基于神经网络集成的 B 样条经验模态分解端点效应抑制方法[J]. 机械工程学报,2013,49(9):106-112.
- [10] 曹冲锋. 基于 EMD 的机械振动分析与诊断方法研究[D]. 杭州:浙江大学,2009.
- [11] 黄先祥,李胜朝,谢建. 新型经验模式分解端点效应消除方法[J]. 机械工程学报,2008,44(9):1-5.
- [12] 朱晓军. HHT 变换及其在脑电信号处理中的应用研究[D]. 太原:太原理工大学,2012.
- [13] 杨贤昭. 基于经验模态分解的故障诊断方法研究[D]. 武汉:武汉科技大学,2012.
- [14] 张进林,张榆锋,张燕,等. 经验模态分解端点效应抑制的常用方法比较研究[J]. 云南大学学报:自然科学版,2010,32(4):406-412.