

基于双边滤波的信号边界特征匹配与延拓

曹晓初 金 弟 鲁银涛 王宗仁 王启迪

(中国石油杭州地质研究院 杭州 310023)

摘 要 经验模态分解过程中的端点效应会直接影响到信号的分解效果。针对此问题,结合邻域双边滤波的思想,提出了基于双边滤波的信号边界特征点匹配的端点延拓方法。该方法有效抑制了解析过程中端点效应的影响,提高了信号分量的精度。边界子波的延拓在遵循原始信号内部特征规律的同时,还考虑了局部数据差异性的影响,因此子波扩展具有自适应性。仿真实验和地震资料应用对比分析表明,该方法能有效解决信号分解过程中的端点效应问题。

关键词 双边滤波,特征匹配,端点延拓,经验模态分解

中图分类号 TP391 文献标识码 A

Signal Boundary Characteristic Matching Extension Based on Bilateral Filtering

CAO Xiao-chu JIN Di LU Yin-tao WANG Zong-ren WANG Qi-di

(PetroChina Hangzhou Research Institute of Geology, Hangzhou 310023, China)

Abstract The end effects of empirical mode decomposition affect the quality of signal decomposition. A method of signal boundary characteristic matching extension based on bilateral filtering was proposed to overcome it. The influence of the end effect is effectively restrained and the accuracy of signal components is improved. Both inherent laws of source signal and difference between local data are considered to boundary wave extension. The simulation and seismic data analysis demonstrate that it is an effective method for signal decomposition.

Keywords Bilateral filtering, Characteristic matching, Endpoint extension, Empirical mode decomposition

随着油气田的大规模开发,大多储量分散、层薄、埋藏深、孔渗条件复杂的油气藏难于被发现。要找到这些储藏复杂的隐蔽性油气藏,就需要对采集到的信号做更准确更精细的处理,以便获得更详尽的地层信息。

地震信号通过经验模态分解方法可以提取出从高到低不同频率段的成份,它们随信号本身而发生变化,而储层信息可能就隐含在某些频段的成份之中。特别是在高频段成份,可能会发现更多与薄层相关的储层信息。

1 经验模态分解

美国学者 N. E. Huang 在 1996 年提出了一种全新的经验模态分解(Empirical Mode Decomposition, EMD)时频分析方法^[1-3],与传统的非线性非稳定处理方法相比,EMD 方法能够很好地表述出信号的局部特征,从信号自身进行分解而不需要核函数,具有高分辨率和自适应性,在很多研究领域得到应用^[4-6]。

1.1 EMD

经验模态分解是将信号中不同尺度的波动或趋势逐级分解,产生一系列具有不同特征尺度的数据序列,每一个序列为一个本征模函数(Intrinsic Mode Function, IMF),最低频率的 IMF 分量通常情况下代表原始信号的趋势或均值。

对信号 $x(t)$ 进行 EMD 分解,产生 n 个 IMF 分量 $c_i(t)$ 和

趋势项 $r_n(t)$ 。即

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

每个 IMF 分量可以看成近似单一模态的平稳信号。根据经验模态分解方法的过程,给出具体算法:

(1)初始化。 $r_1(t) = x(t), i=1, k=0$ 。

(2)分解出第 i 阶 IMF 分量 $c_i(t)$: ①设置 $h_1(t) = r_1(t)$; ②查找出 $h_k(t)$ 中所有极大值 MAX_k 和极小值 MIN_k ; ③利用三次样条插值函数对极大值 MAX_k 和极小值 MIN_k 进行拟合,求出上下包络线 $e_+(t)$ 和 $e_-(t)$; ④计算上下包络线均值 $m_+(t)$; ⑤ $h_{k+1}(t) = h_k(t) - m_k(t)$; ⑥计算两次 $h_k(t)$ 序列之间的总方差 $\sigma = \frac{\sum [h_{k+1}(t) - h_k(t)]^2}{\sum [h_k(t)]^2}$ 。若 $\sigma \leq \epsilon$ (ϵ 为筛分门限),

则 $c_i(t) = h_k(t)$; 否则,令 $k = k + 1$, 转②迭代。

(3) $r_{k+1}(t) = r_k(t) - c_{k+1}(t)$, 判断余量 $r_{k+1}(t)$ 是否为单调函数或是常量。如果满足条件,则整个 EMD 分解过程结束。

1.2 端点效应

端点效应是影响 EMD 分解的一个重要因素。由于地震信号序列的有限性,信号两端无法提供样条插值函数所需要的条件。分解过程中,包络线会在端点处产生发散现象,并且这种发散结果会逐渐向内“污染”整个数据序列而使所得结果严重失真。数据越短,其危害性越大。因此对端点采用合理

曹晓初(1978—),男,硕士,主要研究方向为分布式网络、信号处理;金 弟 男,硕士,高级工程师,主要研究方向为信息存储、高性能计算、系统集成;鲁银涛 男,硕士,工程师,主要研究方向为海域油气勘探与综合评价;王宗仁 男,助理工程师,主要研究方向为网络管理;王启迪 男,助理工程师,主要研究方向为网络管理。

的方式进行处理是必要的。

针对这一问题,目前提出了特征波法、斜率法、极值延拓法、极值镜像延拓法、多项式拟和法、波形镜像法、神经网络法、自回归模型法和支持矢量回归机等改进方法。文献[8]对上述方法进行了分析,同时给出了波型匹配的端点延拓方法,能较好地利用信号自身局部特征实现端点的边界延拓。但该方法存在以下3方面问题:(1)匹配子波与边界子波在纵向上可能存在一定的偏移,直接延拓可能会导致端点处延拓信号的突跳现象,使分解信号出现严重变形;(2)对内在规律不强的信号,虽然可以找到最小匹配偏差的子波,但未对偏差范围做限制,可能存在匹配子波与边界子波误差太大的现象,影响延拓子波的精度;(3)匹配特征子波后,边界子波根据匹配子波的数据直接延拓,没有考虑子波之间的局部差异性。

2 基于双边滤波的边界特征子波匹配与延拓

本文结合邻域滤波方法中双边滤波^[7]的思想,提出了基于双边滤波的边界特征匹配与延拓方法。对子波匹配的偏差范围进行控制,保证子波匹配的精度。以新的极值点作为边界延拓子波的特征点来确定波形趋势,以原始信号自身采样特点确定波形的插值长度,构建具有局部特征的延拓子波。

2.1 双边滤波

双边滤波是一种非线性、非迭代的局部简单滤波算法。它不仅考虑到距离上的邻近关系,同时还考虑数值上的相似性。

对信号序列进行双边滤波可以表示为:

$$x^{filter}(t) = \frac{1}{W_p} \sum_{t_i \in \Omega} x(t_i) f_r(|x(t_i) - x(t)|) g_s(|t_i - t|)$$

其中, $x^{filter}(t)$ 为 t 点处的滤波数据, $x(t_i)$ 为 t_i 处的采样数据, f_r 为基于值相似度的内核函数, g_s 为基于空间距离向量的内核函数, Ω 为邻域范围, W_p 为加权系数和。

$$W_p = \sum_{t_i \in \Omega} f_r(|x(t_i) - x(t)|) g_s(|t_i - t|)$$

式中,

$$\begin{cases} f_r(|x(t_i) - x(t)|) = e^{-\frac{|x(t_i) - x(t)|^2}{2\sigma_r^2}} \\ g_s(|t_i - t|) = e^{-\frac{|t_i - t|^2}{2\sigma_d^2}} \end{cases}$$

在子波匹配过程中,边界子波与匹配子波的特征点按顺序对应。因此,匹配成功后计算新的边界延拓数据时,空间距离相关的内核函数取值为1,即:

$$W_p = \sum_{t_i \in \Omega} f_r(|x(t_i) - x(t)|)$$

2.2 边界特征子波匹配与延拓

2.2.1 匹配波偏移量

特征子波的匹配主要实现波形的匹配。匹配出的子波虽然从波形特征上接近边界子波,但从波形衔接的连续性上来看,纵向方向可能存在一定的偏移。如果直接延拓,可能会出现突跳现象,严重影响分解结果。因此,在进行数据延拓时,需要将子波间的偏移量考虑进去。

依据子波匹配的特征点,定义偏移量为:

$$\Delta d = \min\{(M_1 - M_i), (N_1 - N_i), (S_1 - S_i)\}$$

即定义匹配子波间3个特征点纵向偏移量的最小值为匹配子波的偏移量。

2.2.2 匹配偏差控制

对于内在规律不强的信号,找出的偏差最小的特征子波未必能与边界子波真正相匹配。因此,通过设定匹配偏差阈值,来提高子波的匹配精度。匹配偏差阈值定义为:

$$\epsilon = \alpha * (|M_1| + |N_1| + |S_1| - 3 * \Delta d)$$

在考虑偏移量的情况下,将匹配偏差定义在边界子波特征点长度的一定比例范围内。 α 为比例系数,该值越小,则匹配精度越高;反之,则匹配精度越低。

2.2.3 子波匹配

边界特征子波的匹配过程如图1所示。边界子波特征点由端点 S_1 、极大值点 M_1 和极小值点 N_1 组成。遍历信号,寻找与边界子波特征匹配的子波,即匹配偏差最小的子波。

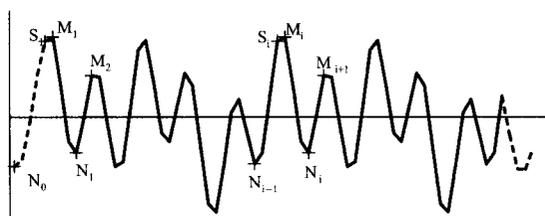


图1 基于邻域双边滤波的端点延拓示意图

假设匹配子波相对应的极大值 M_i 和极小值点 N_i 及 S_i , 与 S_1 相对应的 S_i 可能没有对应的采样点,通过线性插值可以计算出 t_{si} 和 S_i :

$$\begin{cases} t_{si} = t_{mi} + (t_{s1} - t_{m1}) \\ S_i = M_i + \frac{t_{si} - t_{m+1}}{t_{m+2} - t_{m+1}} (x(t_{m+1}) - M_i) \end{cases}$$

(1)计算匹配子波间的偏差:

$$e = (|M_1 - M_i| + |N_1 - N_i| + |S_1 - S_i| + |M_2 - M_{i+1}| - 4 * \Delta d)$$

式中,最后一项 M_2 和 M_{i+1} 为匹配波形的趋势项,反映了特征波形在曲线中相对极值点的位置。

(2)若最小波形的偏差小于匹配偏差阈值,即 $e_{\min} \leq \epsilon$,则认为对应子波为匹配波形。若存在多个相等的偏差子波时,则取与起始点距离最近的波形为匹配子波。

(3)若没有找到相匹配的子波,则取待延拓边界邻近的3个极大值之和与极小值之和的均值作为边界延拓后子波波形的极大极小值点。

(4)子波匹配成功后,参照匹配子波的特征,利用邻域双边滤波的思想计算边界延拓子波的极大值和极小值。

2.2.4 端点延拓

波形的极值点确定了波形的总体趋势,这与经验模态分解的原理是一致的。在计算出延拓子波的极值后,通过B样条插值的方法实现边界子波的延拓。为了尽量保持波形的内在局部特征,相邻极大值与极小值之间的插值点数依据信号自身周期采样数来确定。遍历原始信号序列,取相邻极大值与极小值点之间数据的点数最大值作为延拓子波极大值与极小值之间插值的数据点数。

完成边界特征子波匹配后,需要参照匹配子波计算边界子波的延拓极值点。以左边界为例,左边界子波由特征点 $S_1 - M_1 - N_1 - M_2$ 表征,对应的匹配特征子波则为 $S_i - M_i - N_i - M_{i+1}$ 。

不妨设

$$\omega_1 = f_r(|M_1 - M_i|) = e^{-\frac{|M_1 - M_i - \Delta d|^2}{2\sigma_r^2}}$$

$$\omega_2 = f_r(|N_1 - N_i|) = e^{-\frac{|N_1 - N_i - \Delta d|^2}{2\sigma_r^2}}$$

ω_1 和 ω_2 为匹配子波间对应极值的相似程度。因此, $(1-\omega_1)$ 和 $(1-\omega_2)$ 可以理解为在计算新的延拓极值时, 两匹配子波对应极值点对生成延拓极值的偏差贡献权重。左边界新的延拓极小值计算:

$$N_0 = \frac{(1-\omega_1)M_i + (1-\omega_2)N_i + N_{i-1}}{3-\omega_1-\omega_2}$$

在延拓极值 N_0 的基础上再计算出左边界的极大值 M_0 。确认边界延拓的极大值和极小值后, 采用 B 样条插值的方法, 依照信号自身的周期采样点数完成一个周期的子波延拓。同理完成右边界的端点延拓。

2.3 分量评判

为了检验分解信号的精确程度, 定义分量与原始子信号之间的偏差大小来进行评判。

将分解信号与原始信号偏差平方求和后取均方根, 作为评判分解信号准确性的标准:

$$\delta = \sqrt{\frac{1}{N} \sum_{i=1}^N |s_i^{dec} - s_i^{org}|^2}$$

其中, s_i^{dec} 为分离出来的信号, s_i^{org} 为对应的原始信号。 δ 值越小, 说明分解出来的分量信号越接近于原始子信号。

3 仿真实验及应用

给定 3 个子信号的叠加混合信号, 通过仿真实验检验算法分解的实际效果。混合信号为:

$$f(t) = \sin(4\pi t) + 1.5\cos(8\pi t) + 2.8\sin(20\pi t + \pi)$$

经验模态分解对短信号是有效的。实验中信号的数据采样点取 60 个, 极大值与极小值点之间最大采样点数为 4, 在完成左右边界扩展后, 对新的扩展信号进行分量信号的提取。

为了验证效果, 本方法与文献[8]特征波匹配的方法进行了对比, 图 2 为仿真实验的结果。从分解图上可以看出, 前两个分量信号与原始子信号基本接近, 靠近右边缘部分分解出来的信号有些发散, 通过扩散影响到了数据中间的分解效果。第 3 个分量信号对比明显, 但改进算法得到的信号效果优于特征波匹配方法得到的信号。

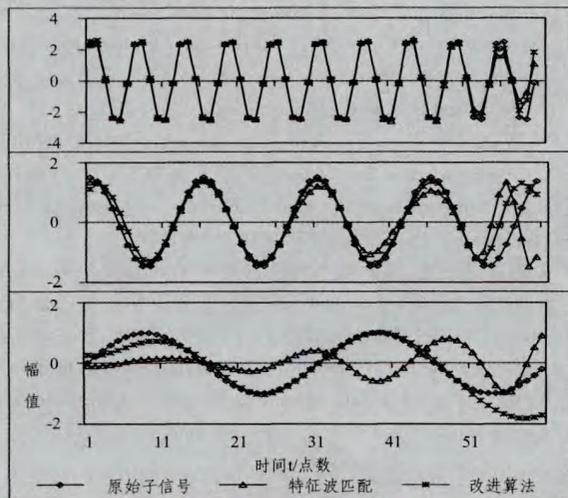


图 2 分量信号对比

对提取的各分量信号计算信号偏差, 如表 1 所列。从分量信号的偏差表中看出, 改进算法得到的分量信号精度整体优于文献[8]所得到的分量信号。这是因为改进算法在匹配子波完成后, 将匹配波形的特征偏差考虑了进来, 重新确定边界延拓子波的极值特征点, 通过 B 样条插值完成边界子波的延拓。

表 1 分量信号偏差对比

算法	分量		
	分量一	分量二	分量三
端点未扩展	2.0898	0.3257	3.9598
特征波匹配	0.8084	0.6794	0.3567
改进算法	0.4013	0.3466	0.3194

在实际应用中, 选取了某区块实际地震数据剖面进行分析。同样将本方法与文献[8]特征波匹配的方法进行了对比。图 3 中上幅剖面图为特征波匹配提取, 下图为改进算法所提取。通过对比可以发现, 改进算法获得的剖面振幅较均一, 地震反射更加连续, 畸变现象明显减少, 有效地刻画出了地层的反射特征。

地震剖面的振幅谱显示, 边缘部分的突变频率基本消除, 分量频段进一步向中心频率靠拢。

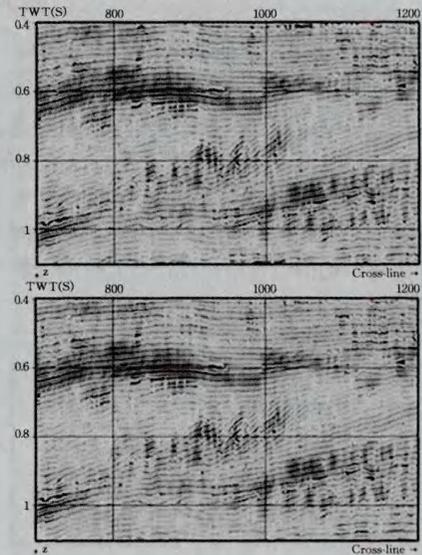


图 3 地震剖面分量对比

结束语 由于信号序列的有限性, 经验模态分解过程中的端点效应问题不可避免。在总结特征波匹配方法的基础上, 结合邻域双边滤波的思想, 提出了基于双边滤波的边界子波特征匹配及延拓方法, 即根据原始信号自身周期采样的特点, 确定边界延拓子波的长度; 利用双边滤波的方法计算延拓子波周期内的极大极小值, 再通过样条插值完成波形的延拓, 该方法具有自适应性。仿真实验及实际地震剖面分析验证了方法的有效性。通过对比可以看出, 改进算法相对于特征波匹配方法更有优势。

参考文献

- [1] Huang N E, Shen Z, Long S R. The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-stationary Time Series Analysis [J]. Proc. R. Soc. Ser. A, 1998, 454: 903-995

[2] Huang W, Shen Z, Huang N E, et al. Engineering Analysis of Biological Variables; an Example of Blood Pressure Over 1 Day [J]. Proc. Natl. Acad. Sci. USA, 1998, 95(9): 4816-4821

[3] Huang N E, Wu M L, Qu W D. Applications of Hilbert-Huang Transform to Non-stationary Financial Time Series Analysis [J]. Appl. Stochastic Models Bus. Indust, 2003, 19(3): 245-268

[4] Ditommaso R, Mucciarelli M, Parolai S, et al. Monitoring the Structural Dynamic Response of a Masonry Tower; Comparing Classical and Time-Frequency Analyses [J]. Bulletin of Earthquake Engineering, 2012, 10(4): 1221-1235

[5] Guang Y, Sun X, Zhang M, et al. Study on Ways to Restrain End

Effect of Hilbert-Huang Transform [J]. Journal of Computers, 2014; 25(3): 22-31

[6] Huang N E, Wu Z. A review on Hilbert-Huang transform; Method and its applications to geophysical studies [J]. Rev. Geophys, 2008, 46(2)

[7] Tomasi C, Manduchi R. Bilateral Filtering for Gray and Color Images [C] // Proceedings of the 1998 IEEE International Conference on Computer Vision. Bombay, India, 1998: 839-846

[8] 胡爱军, 安连锁, 唐贵基. HILBERT-HUANG 变换端点效应处理新方法 [J]. 机械工程学报, 2008, 44(4): 154-158

(上接第 295 页)

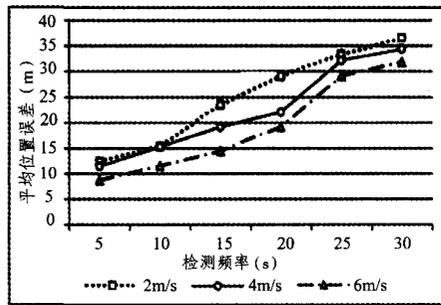


图 5 移动阅读器速度对平均位置误差的影响

最后一种场景研究移动阅读器数量对推动和拉引策略下定位延时的影响。定位间隔为 60s, 允许移动阅读器每 10s 检测周围的标签, 结果如图 6 所示。不难发现, 当移动阅读器数量较小时, 拉引策略的性能优于推动策略, 因为位置检索没有任何延误地被回复或转发, 而在推动策略下, 移动阅读器必须等待定位间隔将它们的位置信息推送给其他阅读器。当移动阅读器数量较多时, 定位延时趋于收敛。

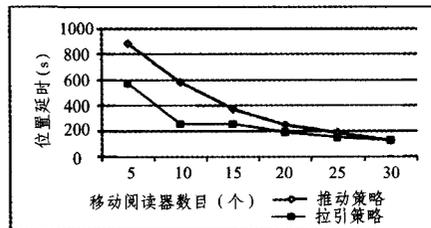


图 6 推拉策略的位置延时

结束语 RFID 技术对交通、航空运输和医疗等领域产生影响, 在多种物联网应用中发挥重大作用。RFID 技术的普及促使人们研究如何使用该技术来解决对象定位问题。虽然人们已经提出许多基于 RFID 的定位系统, 但是这些系统在对象的可扩展性、机动性和异构性方面无法应对物联网的挑战。本文提出的 LSMT-RFID 系统利用一组 Ad hoc 异构移动式 RFID 阅读器来实现移动标签的合作式定位, 并可确保移动标签位置的及时传播。系统具有较强的适应性, 其性能受到多个参数的控制: 检测间隔、定位间隔、移动阅读器数量和移动速度。这些参数视具体应用情况而定。下一步将研究这些参数的相互关系, 并考察移动阅读器按照事先确定好的路径移动时的系统性能。

参考文献

[1] 彭商谦, 李战怀, 李强, 等. RFID 数据流上多目标复杂事件检测 [J]. 计算机研究与发展, 2012, 49(9): 1910-1925

[2] La Porta T F, Maselli G, Petrioli C. Anticollision protocols for single-reader rfid systems; Temporal analysis and optimization [J]. IEEE Transactions on Mobile Computing, 2011, 10(2): 267-279

[3] Bu K, Xiao B, Xiao Q, et al. Efficient pinpointing of misplaced tags in large RFID systems [C] // 2011 8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad hoc Communications and Networks (SECON). IEEE, 2011: 287-295

[4] Sanpechuda T, Kovavisaruch L. A review of RFID localization: Applications and techniques [C] // 5th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, 2008 (ECTI-CON 2008). IEEE, 2008: 769-772

[5] Choi B S, Lee J W, Lee J J, et al. A hierarchical algorithm for indoor mobile robot localization using RFID sensor fusion [J]. IEEE Transactions on Industrial Electronics, 2011, 58(6): 2226-2235

[6] Han S, Lim H S, Lee J M. An efficient localization scheme for a differential-driving mobile robot based on RFID system [J]. IEEE Transactions on Industrial Electronics, 2007, 54(6): 3362-3369

[7] Asadzadeh P, Kulik L, Tanin E. Gesture recognition using RFID technology [J]. Personal and Ubiquitous Computing, 2012, 16(3): 225-234

[8] Papapostolou A, Chaouchi H. RFID-assisted indoor localization and the impact of interference on its performance [J]. Journal of Network and Computer Applications, 2011, 34(3): 902-913

[9] Ni W, Xiao W, Toh Y K, et al. Fingerprint-MDS based algorithm for indoor wireless localization [C] // 2010 IEEE 21st International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC). IEEE, 2010: 1972-1977

[10] Zhao Y, Liu Y, Ni L M. VIRE: Active RFID-based localization using virtual reference elimination [C] // International Conference on Parallel Processing, 2007 (ICPP 2007). IEEE, 2007: 56-66

[11] Bouet M, Pujolle G. L-VIRT: Range-free 3-D localization of RFID tags based on topological constraints [J]. Computer Communications, 2009, 32(13): 1485-1494