

EMD 模态混叠消除及其在语音基音提取中的应用研究

全海燕¹, 刘增力¹, 吴庆畅²

(1. 昆明理工大学 信息工程与自动化学院, 云南 昆明 650051;

2. 云南民族大学 物理与电子电气信息学院, 云南 昆明 650031)

摘要: 针对信号固有奇异模态对 EMD 筛选构成模态混叠的问题, 在提取语音浊音信号的基音成分的分析方法中, 用 FIR 低通滤波器对信号进行预处理, 使随后进行的 EMD 处理筛选出有物理意义的模式分量, 即浊音的基音成分. 并利用 Hilbert 变换得到了基音的时频、时能分布图.

关键词: EMD; FIR 滤波器; 时频分布; 语音; 基音 Hilbert 变换

中图分类号: TN91 **文献标识码:** A **文章编号:** 1007-855X(2006)04-0043-03

Analytical Research on the Pitch of Speech Signal Based on EMD

QUAN Hai-yan¹, LIU Zeng-li¹, WU Qing-chang²

(1. Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650051, China;

2. School of Physics & Electronic and Electric Engineering, Yunnan Nationality University, Kunming 650031, China)

Abstract: In EMD method, the abrupt mode function of the signal will destroy sifting process, therefore, the mode function with physical concept cannot be achieved. Firstly the sonant in speech signal is processed with FIR low pass filter to filter out the abrupt mode function. Then it is sifted with EMD. In this way, the mode function with physical concept, the pitch component, can be obtained. At last Hilbert transform is used to get the instant time frequency distribution and time magnitude distribution of the pitch.

Key words: EMD; FIR filter; time frequency distribution; speech signal; pitch Hilbert transform

0 引言

1998年, Nordeng E. Huang 提出了一种对非平稳信号进行自适应分解的方案: 首先对信号进行经验模态分解(Empirical Mode Decomposition, 简称 EMD), 筛选出一系列的单频窄带模态分量(可利用它们通过线性叠加的方式构造出原信号), 然后对各个单频窄带模态分量进行 Hilbert 变换, 得到信号的瞬时时频分布和时能分布. 该方案被称为 Hilbert-Huang 变换. 该方法存在一个突出的问题, 若信号中存在奇异点, 则 EMD 筛选出的第一模态分量将包含该奇异点, 而正常的信号分量将被推后到下一模态分量, 逐级影响, 使原本具有完整物理意义的信号分量被筛选到不同的模态分量中. 该问题无统一的解决方法, 要根据所分析信号的特点具体处理. 本文采用 FIR 滤波对信号进行预处理, 消除高次谐波及噪音引入的奇异点, 使具有完整物理意义的信号分量以线性叠加的方式凸现出来. 然后再用 EMD 筛选出我们所需的有物理意义的模态信号. 作为应用, 文中用该方法完整提取出语音信号的浊音基音成分及其瞬时时频分布和时能分布.

1 单频窄带信号的 EMD 方法和 Hilbert 谱分析

对单频窄带信号 $x(t)$:

$$x(t) = a(t) \cos(\omega_0 t + \varphi(t)) \quad (1)$$

收稿日期: 2005-04-06.

第一作者简介: 全海燕(1970-), 男, 工程师. 主要研究方向: 信号处理. E-mail: quanhaiyan@163.com

式中, $a(t)$ 是一个低频信号, 其最高频率 ω_c 远小于正弦信号的频率 ω_0 , $\varphi(t)$ 是正弦信号的相位函数. 可以证明^[2]: $x(t)$ 的解析信号 $z(t)$ 为:

$$z(t) = a(t) \cos(\omega_0 t + \varphi(t)) + ja(t) \sin(\omega_0 t + \varphi(t)) = a(t) e^{j(\omega_0 t + \varphi(t))} \quad (2)$$

根据(1)式和(2)式可知: 实单频窄带信号 $x(t)$ 的瞬时幅度等于解析信号 $z(t)$ 的瞬时模值, 瞬时相位等于解析信号 $z(t)$ 的瞬时相位. 若定义 $x(t)$ 的瞬时频率为:

$$f(t) = \frac{1}{2\pi} \frac{d(\omega_0 t + \varphi(t))}{dt} \quad (3)$$

则实信号 $x(t)$ 的 $f(t)$ 的求取可这样进行: 首先应用 Hilbert 变换求取 $x(t)$ 的解析信号 $z(t)$, 求其相位后, 对其求导, 再除以 2π 即得. 同时, 若求取 $z(t)$ 的模值, 则其为实信号 $x(t)$ 的瞬时幅度. 这为单频窄带信号进行精确的时频分析提供了可实施的有效方法.

由上面的分析可知: 若能将要分析的信号分解为单频窄带信号, 则可实现对信号的瞬时时频分析. Nordeng E. Huang 等人认为^[2]: 任何信号均可分解为单频窄带的模态信号 (IMF) 的线性之和, 并给出了固有模态信号的筛选方法: 经验模态分解方法 (EMD). 通过 EMD 筛分, 信号 $s(t)$ 可分解为:

$$s(t) = \sum_{i=1}^n c_i(t) + r_n(t) \quad (6)$$

其中, $c_i(t)$ 为单频窄带的模态信号, 可通过上面的分析方法求得其瞬时时频分布和时能分布. $r_n(t)$ 为最终残余函数, 表示信号的平均趋势.

EMD 实现了对信号的自适应、高分辨时频解析, 但它也有一些问题, 最为突出的是端点问题和奇异模态分量 (信号中有实际意义但未能分布于整个信号时段的模态分量) 问题. 端点问题由于 Zhao and Huang 的努力, 通过镜像延拓法得到了较好的解决, 本文下面的实验均采用这种方法解决端点问题. 奇异模态分量的存在将使 EMD 筛选出的模态分量产生混叠. 该问题至今未有一般性的解决办法, 本文采用对信号进行 FIR (finite impulse response filter: 有限冲激响应滤波器) 滤波的预处理手段避免该问题.

2 语音浊音信号的基音频率提取

语音信号是一个很复杂的信号, 根据发音时声带是否振动可分为清音和浊音. 浊音由于声带的振动, 所以是一个准周期性信号, 含有基波及各次谐波成分. 其中, 基音信号的频率是语音信号中相对平稳的成分, 故它是一个重要的参数, 对它的估计是语音信号分析中的一个方面. 在基音频率估计的方法中, 常用的方法是相关法和小波分析法. 这些方法都是估计出语音信号一个时间段的基音频率, 且稳定性不好.

为了利用 EMD 方法从语音信号中提取其基音成分, 我们采用以下方法: 首先对其进行预处理, 利用 FIR 低通滤波器进行滤波, 得到浊音信号中的基音成分及其低次谐波成分. 这样做是考虑到浊音信号中的基音成分及其低次谐波成分稳定地分布在整个浊音时段, 而被滤除的高次谐波和随机干扰成分是奇异模态, 从而避免了随后 EMD 筛选中因奇异模态影响造成的模态混叠, 以便在被筛选出的模态信号中得到完整的基音信号. 进一步可通过 Hilbert 变换求得模态信号的解析信号, 根据解析信号即可求得模态信号 (即基音信号) 的瞬时时频分布和时能分布.

图 1 是“到 (DAO)”这个字的双元音部分“AO”的语音波形, 利用上面的方法处理, 结果得到的基音波形及其时频、时能分布如图 2 所示.

图 3 是“谈 (TAN)”这个字的元音部分“AN”的语音波形, 利用上面的方法处理, 结果得到的基音波形及其时频、时能分布如图 4 所示.

从上面两个例子的分解结果可看到, 首先利用 FIR 滤波器对语音信号的浊音成分进行预处理, 然后再进行 EMD 筛选, 在筛选出的模态信号中可得到完整的基音信号. 对该基音信号进行 Hilbert 变换, 可得到其解析信号. 由于被筛选出的基音信号是一单频窄带信号, 根据上节中的结论和(3)式, 可得其瞬时时频分布和时能分布. 由瞬时时频分布和时能分布, 清晰地揭示了基音的频率和能量随时间的分布情况.

3 结论

EMD 及其 Hilbert - Huang 变换是一种直接针对信号本身进行自适应的分解方法, 它抛弃了传统方法

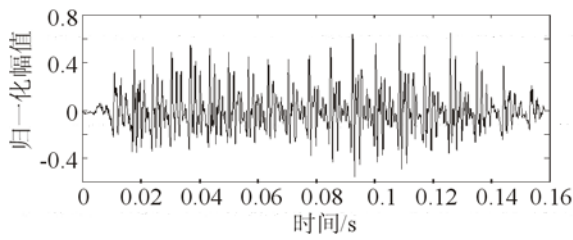


图1 “AO”元音的波形
Fig.1 Vowel(AO) wave

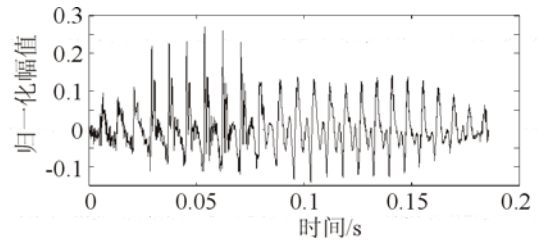


图3 “AN”元音的波形
Fig.3 Vowel(AN) wave

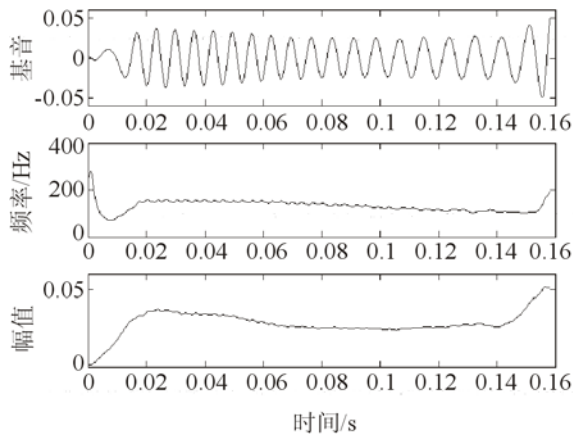


图2 提取的“AO”元音的基音波形
及其时频、时能分布
Fig.2 Time frequency distribution
of pitch of vowel(AO)

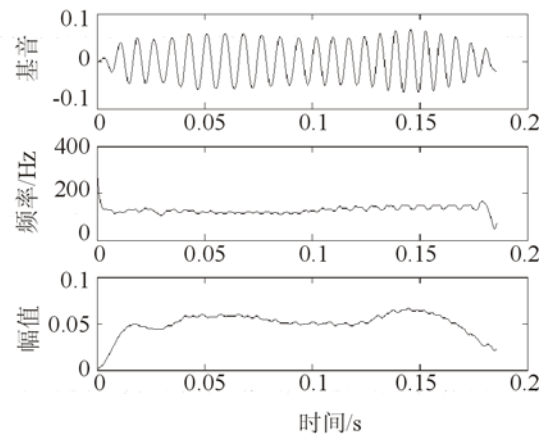


图4 提取的“AN”元音的基音波形
及其时频、时能分布
Fig.4 Time frequency distribution
of pitch of vowel(AN)

中采用固定基函数对信号进行分解的方法,是一种信号分析的创新.但该方法还处于探索阶段,仍有一些问题有待解决.如奇异模态的影响使分解出的各模态分量的物理意义难于解释.本文用 FIR 低通滤波器对信号进行预处理,消除奇异模态对 EMD 筛选的影响,分解出了有物理意义的语音浊音中的基音成分.进一步的方法有待研究以解决该问题.

在利用该方法对大量的语音浊音信号进行分析后表明,该方法可提取浊音中的基音成分,并能得到基音的时频分布和时能分布.在对双元音浊音信号的基音提取实验中,还可通过调节 FIR 低通滤波器的截止频率,使 EMD 筛选出的前后两个元音的基音成分分布在不同的模态分量中.这可用于双元音的定位.

可以看到:通过 EMD 方法和其它方法的结合,可充分发挥 EMD 的优点,将 EMD 方法有效地应用到信号分析处理中去.

参考文献:

- [1] 易克初,田斌,付强. 语音信号处理[M]. 第1版. 北京:国防工业出版社,2001:60-66.
- [2] Norden E Huang, Zheng Shen, Steven R Long, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[J]. Proc R Soc Lond A, 1998,454:903-995.
- [3] Zhao Jin-ping, HUANG Da-ji. Mirror extend and circular spline function for empirical mode decomposition method[J]. Journal of Zhejiang University (Science), 2001,2(3):247-252.
- [4] 赵进平. 异常事件对 EMD 方法的影响及其解决方法研究[J]. 青岛海洋大学学报,2001,31(6):805-814.
- [5] Zhao Jinping. Improvement of the Mirror Extending in Empirical Mode Decomposition Method and the Technology for Eliminating Frequency Mixing[J]. High Technology Letters,2002,8(3):40-47.
- [6] 黎明. 一类非线性发展方程的抛物线解各扭波解[J]. 昆明理工大学学报(理工版),2005,30(1):111-114.