

# 基于 EMD 的前后置滤波语音增强算法

潘晴, 冉福星, 李雅昆

(广东工业大学 信息工程学院, 广州 510006)

**摘要:**针对低信噪比下语音增强困难的问题及经验模态分解(empirical mode decomposition, EMD)在语音增强中的有效应用,提出了一种基于 EMD 的前后置滤波语音增强算法.将含噪语音信号进行卡尔曼滤波后做 EMD 分解,采用能量阈值法判断出含有残余噪声的本征模态函数(IMF),对其小波变换后与其余 IMF 重构为最终的语音增强信号.实验表明,在输入信噪比为 $-10\sim 5$  dB 的不同噪声环境下,以时域分段信噪比及语音质量感知评测 PESQ 作为评价指标,效果均优于单独的 EMD、小波软阈值法及卡尔曼滤波算法,是一种有效的语音增强算法.

**关键词:**卡尔曼滤波;经验模态分解;能量阈值;小波变换

**中图分类号:**TN912.3

**文献标志码:**A

语音信号在产生、传输、接收的过程中不可避免地受到噪声信号的干扰,语音增强技术随之为人们所重视.近年来,信号处理方法经验模态分解(EMD)理论逐渐成熟,EMD 理论是将信号分解为有限个本征模函数(IMF),这些本征模函数包含了原始信号不同时间尺度的局部特征,可以更加清晰地分析信号.EMD 理论可以应用到任何时间序列信号上,对于非线性和非平稳信号有着独特的优势,EMD 已经应用到电力系统领域<sup>[1-2]</sup>.语音增强领域也得到应用,文献[3-4]基于 EMD 通过采用复杂的阈值算法直接去除噪声 IMF 分量的方法在低信噪比环境下语音增强效果明显,但语音信号仍残留部分噪声.文献[5]采用短时傅里叶变换、改进最小控制递归平均(IMCRA)及 EMD 综合算法对强噪声环境下的语音进行语音增强,但去噪的关键在于 IMCRA 估计噪声及语音估计增益函数计算的精确度,并且噪声估计存在延迟.文献[6]对含噪语音信号按清音与浊音进行识别分类然后分别进行 EMD 分解,再对噪声 IMF 去噪,最后重构达到语音增强的目的,语音增强效果进一步提升,缺点在于清浊音识别的正确与否会影响语音增强效果,特别是清音识别困难.因此,在 EMD 的基础上结合相关滤波算法对含噪语音信号具有更好的语音增强效果.卡尔曼滤波是一种在统计信号处理领域具有广泛应用的方法,语音增强<sup>[7]</sup>也是其应用的一个重要方面.文献[8]对卡尔曼滤波原理进行了详细的分析,并具体给出其在语音增强和说话人跟踪方面的应用,语音增强后的语谱图显示,可以去除绝大部分噪声.文献[9]利用卡尔曼滤波及递归期望最大化算法对语音在线混响去除研究,结果表明能够有效去除在线混响并提高了语音质量,卡尔曼滤波及最大期望值算法都属于迭代算法,运算量较大.小波变换理论在多个领域应用广泛,文献[10]小波分析用于电力系统暂态复合扰动信号检测.文献[11]将基于 Teager 能量  $t$  模型构造定制阈值函数用于感知小波包系数选取的语音增强算法,算法采用 teager 能量算子构造阈值函数,但算法较为复杂,运算量大.卡尔曼滤波语音增强效果好但存在残留周期噪声的问题,EMD 单独去噪存在大量残留噪声,而小波变换存在小波基选取困难的问题.因此,本文结合卡尔曼滤波、EMD 及小波变换的优缺点,充分利用各自的优点将 3 种方法结合在一起应用,即在 EMD 方法基础上采用前置卡尔曼滤波,然后运用 EMD 分解出噪声 IMF 分量,对其采用常用小波基再次语音增强,最后整体重构即为语音增强信号,既解决了卡尔曼滤波产生的周期噪声问题也避免了小波基选取困难的问题.

**收稿日期:**2017-09-27;**修回日期:**2018-01-03.

**基金项目:**国家自然科学基金(61372173)

**作者简介(通信作者):**潘晴(1975-),男,江苏宜兴人,广东工业大学副教授,博士,主要研究方向为模式识别、信号处理,  
E-mail:panqing@gdut.edu.cn.

# 1 算法原理分析

## 1.1 AR 模型参数估计

自回归模型(AR 模型)是现代参数功率谱估计模型,任何具有功率谱密度的随机信号  $x(n)$  都可以看作由白噪声  $u(n)$  激励-物理网络所形成.定义

$$x(n) = \sum_{l=1}^p a_l x(n-l) + u(n), \quad (1)$$

(1) 式为  $p$  阶自回归模型,简称 AR 模型,其中,  $a$  为回归系数,  $u(n)$  为零均值,方差为  $\sigma_u^2$  的高斯白噪声.常用的估计 AR 模型参数的方法有相关函数算法、反射系数算法、最小二乘算法.最小二乘算法估计 AR 模型参数方法如下.

设时间信号  $x(n)$ , ( $n=1, 2, \dots, N$ ), 记为

$$\begin{aligned} \mathbf{X} &= [x(p+1) \quad x(p+2) \quad \cdots \quad x(p+N)]^T, \\ \boldsymbol{\varepsilon} &= [u(p+1) \quad u(p+2) \quad \cdots \quad u(N)]^T, \boldsymbol{\gamma} = [a_1 \quad a_2 \quad \cdots \quad a_p]^T, \\ \mathbf{B} &= \begin{bmatrix} x(p) & x(p-1) & \cdots & x(1) \\ x(p+1) & x(p) & \cdots & x(2) \\ \vdots & \vdots & & \vdots \\ x(N-1) & x(N-2) & \cdots & x(N-p) \end{bmatrix}. \end{aligned}$$

由最小二乘算法原理得回归系数

$$\boldsymbol{\gamma} = (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \mathbf{X}. \quad (2)$$

噪声估计值

$$u(n) = x(n) - a_1 x(n-1) - \cdots - a_p x(n-p), (n=p+1, p+2, \dots, N), \quad (3)$$

噪声方差

$$\sigma_u^2 = \frac{1}{N-p} \sum_{i=p+1}^N u^2(i) = \frac{1}{N-p} \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon}. \quad (4)$$

## 1.2 卡尔曼滤波

卡尔曼滤波器源于匈牙利数学家 Rudolf Emil Kalman 的博士论文,是一个最优化自回归数据处理算法.卡尔曼滤波是以最小均方误差为估计的最佳准则,来寻求一套递推估计的算法.其基本思想是采用信号与噪声的状态空间模型,利用前一时刻地估计值和现时刻的观测值来更新对状态变量的估计,求出当前时刻的估计值.在不考虑控制系统的条件下,假设线性离散系统方程为

$$\mathbf{X}(n) = \mathbf{H}\mathbf{X}(n-1) + \mathbf{G}\mathbf{U}(n), \quad (5)$$

$$\mathbf{Y}(n) = \mathbf{A}\mathbf{Y}(n) + \mathbf{w}(n), \quad (6)$$

其中,  $\mathbf{X}(n)$  是系统  $N$  维被估计的状态变量,  $\mathbf{H}$  为状态转移矩阵,  $\mathbf{G}$  是系统噪声输入矩阵,  $\mathbf{U}(n)$  是系统过程噪声,  $\mathbf{Y}(n)$  是测量变量,  $\mathbf{A}$  是测量矩阵,  $\mathbf{w}(n)$  是测量噪声.

系统噪声与测量噪声的统计特性,假设为:

$$\begin{cases} E(\mathbf{U}(n)) = \mathbf{0}, \text{cov}(U_k, U_j) = Q\delta_{ij}, \\ E(\mathbf{w}(n)) = \mathbf{0}, \text{cov}(w_k, w_j) = R\delta_{ij}, \\ \text{cov}(U_k, w_j) = 0 \end{cases} \quad (7)$$

其中,  $\mathbf{Q} = \sigma_u^2 \mathbf{I}$  过程噪声协方差矩阵,  $\mathbf{R} = \sigma_w^2 \mathbf{I}$  测量噪声协方差矩阵,  $\mathbf{I}$  为单位矩阵.状态矢量

$$\hat{\mathbf{X}}(n) = \mathbf{H}\hat{\mathbf{X}}(n-1) + \mathbf{K}(n)[\mathbf{Y}(n) - \mathbf{A}\mathbf{H}\hat{\mathbf{X}}(n-1)], \quad (8)$$

卡尔曼增益

$$\mathbf{K}(n) = \mathbf{P}(n | n-1) \mathbf{A}^T [\mathbf{R} + \mathbf{A} \mathbf{P}(n | n-1) \mathbf{A}^T]^{-1}. \quad (9)$$

预测误差协方差矩阵

$$\mathbf{P}(n | n-1) = \mathbf{H}\mathbf{P}(n-1)\mathbf{H}^T + \mathbf{G}\mathbf{Q}\mathbf{G}^T. \quad (10)$$

估计误差协方差矩阵

$$\mathbf{P}(n) = [\mathbf{I} - \mathbf{K}(n)\mathbf{A}]\mathbf{P}(n | n - 1). \quad (11)$$

状态最优估计值

$$\hat{\mathbf{x}}(n) = \mathbf{A}\hat{\mathbf{X}}(n). \quad (12)$$

设定初值  $\hat{\mathbf{X}}(0)$  及  $\mathbf{P}(0)$  即可求得  $n$  时刻的最优估计值  $\hat{\mathbf{x}}(n)$ .

### 1.3 经验模态分解

EMD 通过迭代过程把信号  $\hat{x}(n)$  分解为有限个频率单一的 IMF 和残余分量. 每个 IMF 可代表信号  $\hat{x}(n)$  一固定范围的频段. 这种方法的本质是通过数据的特征时间尺度来获得本征波动模式, 然后分解数据. 每个 IMF 必须满足 2 个条件:

- (a) 信号的极值点数目和过零点数目相差不多于一个;
- (b) 由信号局部极大值构成的包络及局部极小值构成的包络的平均值为 0.

EMD 分解方法基于数据的时间尺度特征进行信号分解, 属于局部平稳化, 不需要预先设计基函数. 信号  $\hat{x}(n)$  至少含有 2 个极值(一个极大值和一个极小值). 通过 EMD 得到本征模函数分量的算法过程如下:

- 1) 找出信号  $\hat{x}(n)$  的所有极值点.
- 2) 用 3 次样条插值函数拟合形成信号  $\hat{x}(n)$  的包络线, 计算包络线的均值记作  $p(n)$ .
- 3) 用原信号  $\hat{x}(n)$  减去  $p(n)$  得到新序列  $h(n)$ , 即  $\hat{x}(n) - p(n) = h(n)$ .
- 4) 令  $h(n)$  作为新的信号重复以上步骤.

直到  $h(n)$  满足 IMF 的条件(a)和(b), 则信号序列  $h(n)$  即是信号  $\hat{x}(n)$  第 1 个本征模函数分量记为  $c_1$ . 然后用  $\hat{x}(n)$  减去  $c_1$  得到一个新数据序列  $r_1$ , 对  $r_1$  再进行上述步骤得到第 2 个本征模函数分量  $c_2$ , 如此重复下去直到最后一个数据序列  $p(n)$  不可分解. 最终  $\hat{x}(n)$  可以被分解为有限个 IMF 分量和残余分量  $r_M(n)$ , 可以用

$$\hat{x}(n) = \sum_{i=1}^M c_i(n) + r_M(n) \quad (13)$$

表示.

### 1.4 小波变换

小波变换理论是近年来发展起来的一种时频局域化分析方法. 它被应用于非平稳信号处理<sup>[10]</sup>; 与其他方法结合应用于语音增强<sup>[11]</sup>. 小波去噪的难点在于选取小波基. 复杂算法耗时, 给实时处理带来不便.

基本的小波变换可根据小波  $\Psi(n)$  和尺度函数  $\varphi(n)$  为信号函数  $f(n)$  定义小波序列

$$f(n) = \sum_k c_{j_0}(k) \varphi_{j_0,k}(n) + \sum_{j=j_0}^{\infty} \sum_k d_j(k) \psi_{j,k}(n), \quad (14)$$

其中,  $j_0$  是任意开始尺度,  $c_{j_0}$  和  $d_j(k)$  分别为尺度系数及小波系数. 通过阈值算法保留较大尺度系数和小波系数重构信号达到去除噪声的目的.

小波去噪方法主要有: 模板大值检测法、阈值去噪法、相关去噪法. 小波阈值去噪法的主要过程: 1) 对待处理信号进行小波分解, 得到各尺度的小波系数; 2) 对各尺度的小波系数做阈值处理; 3) 小波重构, 得到去噪后的信号.

阈值施加方法及对阈值估计对小波变换的效果至关重要, 常用的阈值施加方法有硬阈值和软阈值方法以及在此基础上的改进方法. 因此, 小波变换去噪不仅需要选择合适的小波基同时阈值施加方法也是一个重要的问题.

## 2 EMD 前后置滤波算法

本文在 EMD 方法的基础上采用前后置滤波算法在语音最小失真的前提下最大限度地去除噪声信号. 首先, 卡尔曼滤波对含噪语音信号进行预处理增强, 原始含噪语音信号

$$y(n) = x(n) + w(n), \quad (15)$$

其中,  $x(n)$  为纯净的语音信号;  $w(n)$  为高斯白噪声, 均值为 0, 方差为  $\sigma_w^2$ . 通过最小统计法(MS)<sup>[12]</sup> 估计得

出.语音信号在较短时间内(10 ~ 30 ms)可近似看作平稳信号,纯净语音  $x(n)$  可以认为是根据  $p$  阶 AR 模型产生的,表示为(1)式,其中回归系数  $a$  及系统过程白噪声  $u(n)$  的值及方差  $\sigma_u^2$  可通过最小二乘算法求得.将(1)和(15)式构造成卡尔曼滤波的离散系统方程(5)和(6),其中  $\mathbf{X}^T = [x(n-p+1), x(n-p+2),$

$$x(n-p+3), \dots, x(n)], \mathbf{H} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ a(p) & a(p-1) & a(p-2) & \cdots & a(1) \end{bmatrix}_{p \times p}$$

为状态转移矩阵,  $\mathbf{A} = \mathbf{G}^T =$

$[0 \ 0 \ \cdots \ 0 \ 1]_{1 \times p}$  分别为测量矩阵和系统噪声输入矩阵.设置初值  $\hat{X}(0)$  及  $P(0)$  均为 0 即可求得  $n$  时刻的卡尔曼滤波最优估计值  $\hat{x}(n)$ .

经过卡尔曼滤波后的语音增强信号

$$\hat{x}(n) = x(n) + v(n), \tag{16}$$

其中,  $x(n)$  为纯净的语音信号,  $v(n)$  为残留高斯白噪声信号.为了进一步去除卡尔曼滤波后的残留噪声而又减少语音信号失真,通过 EMD 方法分解出有限个 IMF 分量,通过对噪声 IMF 分量单独处理,其余语音信号 IMF 分量不做任何处理达到去除残留噪声减少语音信号失真的目的.对卡尔曼滤波后的语音信号  $\hat{x}(n)$  做 EMD 分解得到有限个 IMF 分量和残余分量,EMD 分解如图 1 所示.

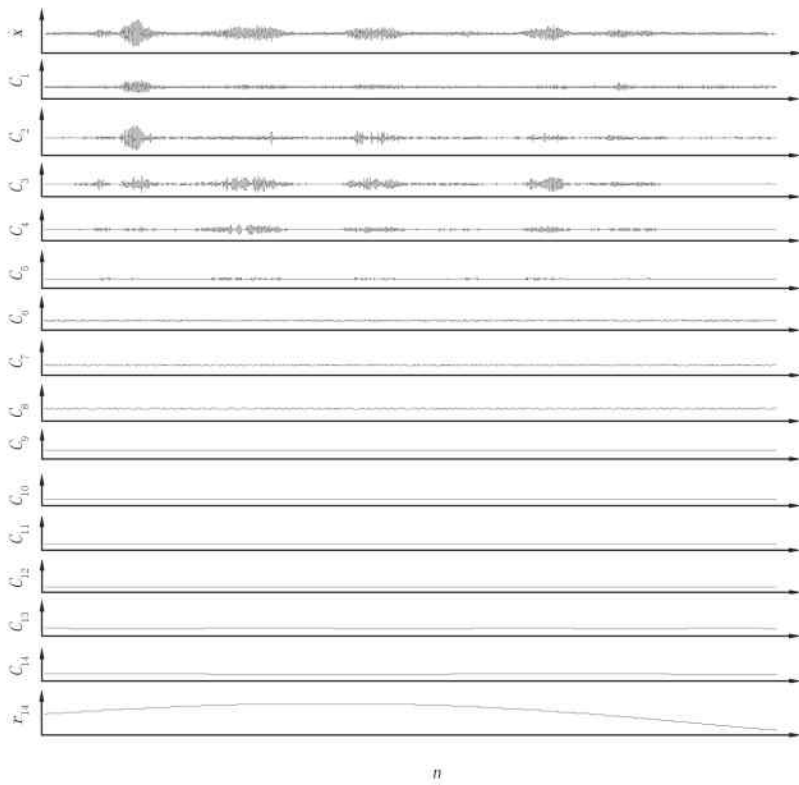


图 1 EMD 分解得到 IMF 分量及余量 (输入信噪比 -5 dB)

从 EMD 分解可知噪声主要集中在前面的 IMF 分量,因此后续小波法只对噪声 IMF 分量进行处理即可.噪声 IMF 分量选取采用一种自适应的能量阈值判定方法.公式如下

$$L = \operatorname{argmax}(E_k), 1 \leq k \leq M, \tag{17}$$

其中,  $L$  为能量最大 IMF 分量的索引值,即第  $L$  个 IMF 的能量值最大,IMF 能量值

$$E_k \cong \frac{1}{N} \sum_{n=1}^N [c_k(n)]^2, \quad (18)$$

式中  $E_k$  对应于第  $k$  个 IMF 分量的能量.因此,  $L$  一旦确定则  $i < L$  的 IMF 分量即可选取出来进行后续去噪处理,而  $i \geq L$  的 IMF 分量则不需要做任何处理.

噪声 IMF 分量确定后,采用小波变换对其进行处理.小波对单独的噪声 IMF 分量处理效果不理想.因为单独的 IMF 分量属于低信噪比信号,小波对低信噪比信号处理效果不好,所以,采用先对选定的噪声 IMF 分量重构再进行小波变换处理.此时的信号相当于间接提高了信噪比,同时小波变换提高了运算速度减少了耗时.噪声 IMF 重构表示

$$g(n) = \sum_{i=1}^{L-1} c_i(n). \quad (19)$$

采用常用的小波基 Sym8 小波,分解层数为 5 层,采用 rigrsure 准则,软阈值.Sym8 小波函数具有紧支集,有良好的对称性和连续性,适合对连续性较好的信号去噪.分解层数对于去噪影响较大,通常分解层数过多会造成信号丢失严重,信噪比反而下降,同时导致运算量增大使处理速度变慢.分解层数过少则达不到消噪的目的.通过实验证明分解层数为 5 层可以达到较好的去噪且运算量不大,降低了算法运算时间.采用 rigrsure 准则,软阈值,更好的保留原始语音信号.表示如下

$$f(n) = \sum_k c_{j_0}(k) \varphi_{j_0,k}(n) + \sum_{j=j_0}^{\infty} \sum_k d_j(k) \psi_{j,k}(n),$$

其中,  $c_{j_0} = \langle g(n), \varphi_{j_0,k}(n) \rangle$ ,  $d_j(k) = \langle g(n), \psi_{j,k}(n) \rangle$ .

原始含噪信号  $y(n)$  经过卡尔曼滤波语音增强后得到信号  $\hat{x}(n)$ ,再通过经验模态分解(EMD)分解为有限个 IMF 分量,并通过能量阈值的方法确定噪声 IMF 分量并重构得到  $g(n)$  进一步采用小波变换去噪处理.最后,将小波变换去噪处理后的信号  $f(n)$  与 EMD 分解的其余语音信号 IMF 分量及残余分量进行重构即为最终的语音增强信号

$$s(n) = f(n) + \sum_{i=L}^M c_i(n) + r_M(n). \quad (20)$$

算法流程:

- 1) 语音信号假设为 AR 模型并估计模型参数;
- 2) 含噪语音信号卡尔曼滤波得到  $\hat{x}(n)$ ;
- 3)  $\hat{x}(n)$  做 EMD 分解得到有限个 IMF 分量;
- 4) 采用能量阈值确定噪声 IMF 分量并重构;
- 5) 对重构的噪声 IMF 分量进行小波变换处理得到  $f(n)$ ;
- 6) 将  $f(n)$ 、其余语音信号 IMF 分量及残余分量重构即为最终语音增强信号  $s(n)$ .

### 3 仿真分析

实验语音材料选自 IEEE 库的 4 条语音,采样频率 8 000 Hz,时长均为 3 秒左右,并添加 NOISEX-92 数据库的白噪声、buccaneer 噪声及 hfchannel 噪声,全局信噪比为 -10 dB, -5 dB, 0 dB, 5 dB, 一共 48 条带噪语音作为实验样本.仿真实验在 MATLAB2014a 平台下进行.以时域分段信噪比(segSNR)及 ITU-T P.862 国际标准的输出语音质量感知评测 PESQ 作为衡量标准,时域分段信噪比(segSNR)反映去除噪声的程度,输出语音质量感知评测 PESQ 反映语音增强后的听觉质量,得分范围一般在 -0.5 到 4.5 之间.仿真实验分别运用 EMD 方法、小波软阈值法、卡尔曼滤波法及本文算法对含噪语音进行语音增强处理.小波基选取 Sym8 小波,分解层数为 5 层,采用 rigrsure 准则,软阈值.对纯净语音添加不同噪声使输入信噪比为 -10 dB 至 5 dB 范围进行实验.本文算法与其他方法的时域分段信噪比及输出语音质量感知评测 PESQ 的实验结果对比见表 1.

本文的语音增强方法在输入信噪比 -10dB 到 5dB 时时域分段信噪比及输出语音质量感知评测 PESQ 总体优于 3 种算法,在 hfchannel 噪声环境 EMD 方法在 -10dB 至 0dB 语音质量 PESQ 得分高,因为强噪声



环境 EMD 方法直接去除噪声分量,具有一定效果,但时域分段信噪比较低,说明 EMD 方法仍然残留大量噪声.本文算法去除噪声效果好,适用噪声范围更大,并且改善听觉质量,提高了语音可懂度.

表 1 时域分段信噪比及输出语音质量感知评测 PESQ

环境噪声类型	输入信噪比/dB	EMD		小波软阈值		卡尔曼滤波		本文算法	
		segSNR/dB	PESQ	segSNR/dB	PESQ	segSNR/dB	PESQ	segSNR/dB	PESQ
白噪声	-10	-6.38	1.41	-4.11	1.01	-2.80	1.38	-2.64	1.50
	-5	-5.04	1.53	-2.89	1.53	-1.73	1.72	-1.55	1.85
	0	-2.72	1.89	-1.07	1.85	-0.33	2.03	-0.16	2.13
	5	-0.92	2.11	1.26	2.16	1.26	2.31	1.54	2.39
buccaneer 噪声	-10	-6.69	1.54	-6.32	1.27	-4.84	1.30	-4.81	1.52
	-5	-5.07	1.70	-4.45	1.66	-3.04	1.81	-3.01	2.02
	0	-3.56	1.90	-2.65	1.82	-1.69	2.09	-1.67	2.22
	5	-1.74	2.11	-0.24	2.09	0.06	2.35	0.08	2.42
hfchannel 噪声	-10	-7.22	1.68	-7.86	1.37	-6.11	1.38	-5.72	1.43
	-5	-4.84	1.81	-4.62	1.52	-2.73	1.41	-2.43	1.51
	0	-2.33	2.06	-2.53	1.71	-1.28	1.79	-0.95	1.91
	5	-0.89	2.09	-0.01	1.98	0.33	2.11	0.59	2.23

语音增强对于实时性要求较高,算法运行时间可以间接反映算法复杂度的高低.在 MATLAB2014a 平台下随机抽取一条语音数据计算在白噪声环境下运用不同算法进行语音增强的程序运行时间.本文算法与 EMD 方法、卡尔曼滤波法、小波软阈值法对含噪语音信号进行语音增强处理运算时间对比,见表 2.

表 2 本文算法与其他几种方法运算时间对比

环境噪声类型	算法耗时/s			
	EMD	小波软阈值	卡尔曼滤波	本文算法
高斯白噪声	2.43	1.22	1.55	4.58

本文算法与其他语音增强算法运算时间对照结果表明,EMD 方法运算需要一定时间,卡尔曼滤波法也属于迭代算法需要耗时,本文算法是 3 种算法的综合,算法耗时高于其他 3 种算法,后期将会优化 EMD 及卡尔曼滤波过程中的迭代算法进一步提高实时性.

## 4 结 论

通过理论论证及实验仿真,本文提出的基于 EMD 的前后置滤波语音增强算法在低信噪比情况下相对于单独的 EMD、小波软阈值及卡尔曼滤波语音增强算法效果具有进一步提升,前置滤波加后置滤波双重语音增强,达到了较好效果.前置滤波的关键问题是保证去除大量噪声的同时保证语音尽量小的失真.EMD 优势在于后置滤波只针对噪声 IMF 分量处理,对含有大量语音信号的 IMF 分量不做处理,从而保证了语音信号的最小失真,虽然算法复杂度稍高于单独的语音增强方法,语音增强效果明显.后期将会进行算法优化降低算法复杂度.因此,本文算法可以解决低信噪比语音增强问题,具有时域分段信噪比高及输出语音质量感知评测 PESQ 高的优势,是一种有效的语音增强算法.

## 参 考 文 献

- [1] 史林军,周佳佳,温荣超,等.基于经验模态分解与小波分析相结合的风电功率平滑控制[J].电力系统保护与控制,2016,44(24):9-16.
- [2] 孙曙光,庞毅,刘建强.基于 Hilbert 频移的 EEMD 谐波检测方法[J].电力系统保护与控制,2017,45(15):85-91.
- [3] Zhang J C,Zhao L.A speech denoising method based on improved EMD[C]//2011 International Conference on Multimedia and Signal

- Processing. Washington DC: IEEE Computer Society, 2011: 305-309.
- [4] Zao L, Coelho R, Flandrin P. Speech enhancement with EMD and Hurst-Based mode selection[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2014, 22(5): 899-911.
- [5] Chatlani N, Soraghan J J. EMD-Based Filtering (EMDF) of low-frequency noise for speech enhancement[J]. IEEE Transactions on Audio, Speech, and Language Processing, 2012, 20(4): 1158-1166.
- [6] Khaldi K, Boudraa A O, Turki M. Voiced/unvoiced speech classification-based adaptive filtering of decomposed empirical modes for speech enhancement[J]. IET Signal Processing, 2016, 10(1): 69-80.
- [7] Mallat S H, Wang W L. Singularity detection and processing with wavelet[J]. IEEE Transaction on Information Theory, 1992, 38(2): 617-643.
- [8] Gannot S. Speech processing utilizing the Kalman filter[J]. IEEE Instrumentation & Measurement Magazine, 2012, 15(3): 10-14.
- [9] Schwartz B, Gannot S, Habets E A P. Online speech dereverberation using kalman filter and EM algorithm[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2015, 23(2): 394-406.
- [10] 曹玲芝, 李振杰. 基于 dq 变换与小波多分辨率分析的电力系统暂态复合扰动信号检测方法[J]. 电力系统保护与控制, 2016, 44(21): 138-143.
- [11] Islam M T, Shahnaz C, Zhu W P, et al. Speech enhancement based on student modeling of teager energy operated perceptual wavelet packet coefficients and a custom thresholding function[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2015, 23(11): 1800-1811.
- [12] Martin R. Noise power spectral density estimation based on optimal smoothing and minimum statistics[J]. IEEE Transactions on Speech and Audio Processing, 2001, 9(5): 504-512.

## Speech enhancement algorithm based on EMD for pre and post filter

Pan Qing, Ran Fuxing, Li Yakun

(School of Information Engineering, Guangdong University of Technology, Guangzhou 510006, China)

**Abstract:** To solve the problem of difficulty in speech enhancement under low SNR conditions, this paper proposes a pre- and post-filtered speech enhancement algorithm based on empirical mode decomposition. The algorithm first performs Kalman filter on the speech signal containing noise, and then performs empirical mode decomposition. And uses energy threshold algorithm to determine the intrinsic modal function with residual noise. After the wavelet transform, the intrinsic modal function is reconstructed with the rest of the intrinsic modal function to become the final speech enhancement signal. In the experiment, SNR of the noise environment was set within the range of  $-10\text{dB}$  to  $5\text{dB}$ , and the time domain segmented signal-to-noise ratio and speech quality perception evaluation were used as evaluation indices. Experimental results show that the effect of speech enhancement is better than that using EMD, wavelet soft thresholding, and Kalman filtering algorithm respectively.

**Keywords:** Kalman filter; empirical mode decomposition; energy threshold; wavelet transform

[责任编辑 杨浦]