

一种新阈值函数的小波包语音增强算法*

任永梅, 张雪英[†], 贾海蓉

(太原理工大学 信息工程学院, 太原 030024)

摘要: 针对传统软、硬阈值函数去噪方法增强的语音存在失真的问题, 提出一种新阈值函数的小波包语音增强算法, 同时给出了新阈值函数和新的 Bark 尺度小波包分解结构。新阈值函数在小波包系数绝对值大于给定阈值的区间内, 灵活地结合了软、硬阈值函数; 在小波包系数绝对值小于给定阈值的区间内, 用一种非线性函数代替传统阈值函数中的简单置零, 实现了阈值函数的平缓过渡; 新的 60 个频带 Bark 尺度小波包分解结构能更好地模拟人耳的听觉感知特性。仿真实验结果表明, 在高斯白噪声和有色噪声背景下, 与传统软、硬阈值函数去噪方法相比, 新算法有效提高了增强语音信噪比和分段信噪比, 减少了语音失真, 具有更好的去噪效果。

关键词: 语音增强; 小波包; 阈值函数; Bark 尺度小波包分解结构; 信噪比; 分段信噪比

中图分类号: TP391 **文献标志码:** A **文章编号:** 1001-3695(2013)01-0114-03

doi:10.3969/j.issn.1001-3695.2013.01.027

New threshold function of wavelet packet speech enhancement algorithm

REN Yong-mei, ZHANG Xue-ying[†], JIA Hai-rong

(College of Information Engineering, Taiyuan University of Technology, Taiyuan 030024, China)

Abstract: Aiming at enhanced speech by using the traditional soft and hard threshold functions methods had the problem of distortion, this paper proposed a new threshold function of wavelet packet speech enhancement algorithm, and put forward the new threshold function and the new Bark scale wavelet packet decomposition structure. The new threshold function flexibly combined the soft and hard threshold functions when wavelet packet coefficients were greater than the given threshold. When wavelet packet coefficients were less than the given threshold, it used a nonlinear function to instead of setting wavelet packet coefficients to zero simply in the traditional threshold function, and it achieved the smooth transition of the threshold function. The new 60 sub-bands Bark scale wavelet packet decomposition structure could achieve better simulation of the human ear auditory perception characteristics. Simulation experimental results show that comparing with traditional threshold function methods, the new proposed algorithm improves the SNR and SSNR of the enhanced speech significantly and reduces the speech distortion in white Gaussian noise and colored noise background, and has better denoising effect.

Key words: speech enhancement; wavelet packet; threshold function; Bark scale wavelet packet decomposition structure; signal-to-noise ratio(SNR); segmental signal-to-noise ratio(SSNR)

0 引言

实际生活中的语音不可避免地受到周围环境和传输媒介等噪声的污染, 导致语音质量下降, 使语音处理系统性能急剧恶化^[1]。因此, 语音增强技术成为了国内外语音信号处理研究的热点。语音增强的目的有两个: a) 抑制背景噪声, 改善语音质量, 使听者乐于接受; b) 提高语音的可懂度, 方便听者理解^[2]。语音增强方法可分为时域方法和频域方法, 时域方法的计算复杂度高, 而频域方法中的谱减法虽然计算量小, 但后者增强的语音存在严重的音乐噪声, 达不到满意的去噪效果。

小波分析方法对信号的时域和频域都进行分析, 它具有多分辨率的特性并且适用于处理时变语音信号, 不仅仅局限于描述语音信号的局部, 还描述其细节^[3]。小波包分析方法^[4,5]是对小波分析的进一步延伸, 它对信号的低频和高频部分同时进

行分解, 具有更强的频带划分能力, 是分解和重构语音信号更加细致的方法。

小波去噪算法中应用最广的是小波阈值去噪算法, 其中硬阈值函数和软阈值函数是两种最为常用的阈值函数。由于硬阈值函数不连续, 导致重构的信号具有振荡性, 且含有明显的噪声; 软阈值函数虽然连续, 但经其增强之后的语音小波包系数和原始小波包系数之间存在恒定的偏差, 有过度平滑的现象。为了解决上述问题, 本文提出一种新的阈值函数和 Bark 尺度小波包分解结构。新阈值函数灵活地结合了软、硬阈值函数, 并对小波包系数绝对值小于给定阈值的区间进行了改进, 具有连续性好、便于调节的优点; 新的 60 个子频带 Bark 尺度分解结构对语音信号频率进行了更加细致的划分, 对人耳听觉感知特性的模拟程度更好。实验结果表明, 新算法在高斯白噪声和有色噪声背景下均能更好地抑制噪声, 减少语音信号的失真, 达到了更好的去噪效果。

收稿日期: 2012-05-01; **修回日期:** 2012-06-15 **基金项目:** 国家自然科学基金资助项目(61072087); 山西省留学回国人员科研资助项目(2011-035)

作者简介: 任永梅(1988-), 女, 陕西宝鸡人, 硕士研究生, 主要研究方向为语音信号处理; 张雪英(1964-), 女(通信作者), 教授, 博导, 博士, 主要研究方向为语音信号处理、DSP 应用、多媒体通信、物联网技术(ty Zhangxy@163.com); 贾海蓉(1977-), 女, 讲师, 博士, 主要研究方向为语音信号处理。

1 小波包变换及其去噪原理

1.1 小波包变换

令观察信号为

$$s(n) = x(n) + u(n) \quad (1)$$

其中: $x(n)$ 是原始信号, $u(n)$ 是噪声信号。假定 $u(n)$ 是零均值且服从高斯分布的随机序列,即服从 $N(0, \sigma_u^2)$ 分布^[6],小波包去噪就是从带噪信号 $s(n)$ 中去除噪声 $u(n)$ 。

在小波包的分解中,节点 $(j+1, p)$ 处的小波包系数^[7]为

$$d_{j+1}^p(k) = \sum_{m=-\infty}^{\infty} d_j^p(m) h_0(m-2k) \quad (2)$$

$$d_{j+1}^{p+1}(k) = \sum_{m=-\infty}^{\infty} d_j^p(m) h_1(m-2k) \quad (3)$$

其中: $d_j^p(k)$ 表示节点 (j, p) 即第 j 层第 p 个节点处的小波包系数; $h_0(k)$ 是通带在 $0 \sim \frac{\pi}{2}$ 之间的低通滤波器; $h_1(k)$ 为通带在

$\frac{\pi}{2} \sim \pi$ 之间的高通滤波器。

在节点 (j, p) 处的小波包系数 $d_j^p(k)$ 可由式(4)重建:

$$d_j^p(k) = \hat{d}_{j+1}^{2p}(k) * h_0(k) + \hat{d}_{j+1}^{2p+1}(k) * h_1(k) \quad (4)$$

其中, $\hat{d}_{j+1}^{2p}(k)$ 和 $\hat{d}_{j+1}^{2p+1}(k)$ 分别是 $d_{j+1}^{2p}(k)$ 和 $d_{j+1}^{2p+1}(k)$ 每两个点插入一个0后所得到的序列。

1.2 小波包去噪原理

经小波包变换后,语音信号的小波包系数会大于噪声的小波包系数。小波包变换具有稀疏性及去除信号的相关性,通过小波包变换将真实信号的能量集中在有限且幅度相对较大的小波包系数上;而高斯白噪声的小波包变换仍是高斯白噪声,其能量广泛分布在所有的小波包系数上,对应的小波包系数幅度不是很大。因此,可以设定一个阈值对小波包系数进行阈值量化处理,只保留大于阈值的系数,用大于阈值的系数来估计语音。小波包去噪原理如图1所示。



图1 小波包去噪原理

小波包去噪可以按照如下三个步骤进行:a)选择合适的小波包分解结构对带噪语音进行小波包分解,得到小波包分解系数;b)对高频的小波包分解系数选择适当的阈值进行阈值量化处理,得出估计的小波包系数;c)用小波包逆变换重构估计的小波包系数,得到增强的语音信号。

2 阈值的选取

小波阈值估计是小波去噪中的另一种关键技术。阈值选取^[8]的优劣直接影响着降噪的效果:阈值选取过大将会把语音信号当做噪声滤除,造成语音失真;阈值选取过小则会把噪声信号当做语音保留,达不到较好的去噪效果。

目前最常用的阈值是根据 Donoho 等人^[9]给出的关于阈值估计风险定理确定的阈值:

$$\lambda = \sqrt{2 \ln N} \sigma_u \quad (5)$$

其中: N 为噪声信号的长度; σ_u 为噪声信号的标准差,即

$$\sigma_u = \frac{\text{median}(|d_j(k)|)}{0.6745} \quad (6)$$

其中: j 是小波包分解尺度; $d_j(k)$ 表示第 j 层第 k 个小波包分解

系数;median 是 MATLAB 中求中值运算的命令。

正交小波包变换有去除相关和使能量集中的性质。因此,经过小波包变换之后,其能量集中分布在少数小波包系数上。增加数据点数,就增加了信息的冗余,但代表信号的小波包系数个数不会增加,仅仅是增加了幅度。允许阈值有所增加的前提下,式(5)中的阈值随信号长度 N 的增大缓慢增加,因此本文选取式(5)作为阈值。

3 新阈值函数和新的 Bark 尺度小波包分解结构

3.1 传统阈值函数

小波阈值去噪算法中传统的阈值函数^[10]主要有:

a)硬阈值函数。与硬阈值函数不同的是,当小波包系数的绝对值大于给定的阈值时,则令其保持不变,即

$$\hat{w}_{j,k} = \begin{cases} w_{j,k} & |w_{j,k}| \geq \lambda \\ 0 & |w_{j,k}| < \lambda \end{cases} \quad (7)$$

b)软阈值函数。与硬阈值函数不同的是,当小波包系数的绝对值大于给定的阈值时,令其减去阈值,即

$$\hat{w}_{j,k} = \begin{cases} [\text{sign}(w_{j,k})] (|w_{j,k}| - \lambda) & |w_{j,k}| \geq \lambda \\ 0 & |w_{j,k}| < \lambda \end{cases} \quad (8)$$

其中: $w_{j,k}$ 是小波包系数的大小; $\hat{w}_{j,k}$ 是阈值量化后的小波包系数的大小; λ 是阈值。

3.2 新阈值函数

硬阈值函数处理之后的语音往往有很大的振荡性,含有很多残留的噪声;软阈值函数处理之后的 $\hat{w}_{j,k}$ 和 $w_{j,k}$ 之间存在恒定的偏差,损失了部分有用的高频分量,严重影响重构语音的质量,造成很大的语音失真。为了克服软、硬阈值函数的缺点,本文提出一种新的阈值函数如式(9):

$$\hat{w}_{j,k} = \begin{cases} \gamma w_{j,k} + (1 - \gamma) \text{sign}(w_{j,k}) (|w_{j,k}| - \lambda) & |w_{j,k}| > \lambda \\ \text{sign}(w_{j,k}) \frac{w_{j,k}^4}{\lambda^3} & |w_{j,k}| \leq \lambda \end{cases} \quad (9)$$

其中: $\gamma = 1 - e^{-\alpha \sqrt{|w_{j,k}| - \lambda}}$ ($\alpha > 0$), γ 是调节因子,介于0和1

之间。当 $|w_{j,k}| = \lambda$ 时 $w_{j,k} = \lambda$,当 $w_{j,k} \rightarrow \lambda$ 时 $w_{j,k} \rightarrow \lambda$,所以 $w_{j,k}$ 在 λ 处是连续的,克服硬阈值函数的不连续性;当 $|w_{j,k}| \rightarrow \infty$ 时, $\gamma \rightarrow 1$, $w_{j,k} \rightarrow w_{j,k}$,因此随着 $|w_{j,k}|$ 的逐渐增大, $w_{j,k}$ 逐渐逼近 $w_{j,k}$,克服了软阈值函数在处理小波包系数时出现的恒定偏差。

在 $|w_{j,k}| > \lambda$ 时,可以通过灵活地调节 α ,使新阈值函数在软、硬阈值函数之间灵活选择;在 $|w_{j,k}| \leq \lambda$ 时,并没有简单地置零,而是采用一种非线性函数,对阈值进行了压缩,保证了新阈值函数的连续性,从而抑制了由于直接截断可能引起的振荡,减少了重构信号和原始信号之间的恒定误差。因此,新阈值函数较传统的软、硬阈值函数具有明显的优势。为了获得较好的去噪效果,在具体应用中需要根据实验结果选出 α 的最优值,本文取 $\alpha = 25$ 。

3.3 新的 Bark 尺度小波包分解结构

根据小波包的函数空间正交剖分理论,可以构造与人耳 bark 域频率描述相似的小波包分解结构,称为 Bark 尺度小波包分解^[11]。人耳 Bark 域在 20 Hz ~ 16 kHz 的频率范围内包括 24 个 Bark^[12,13],可以用小波包分频来模拟人耳的 24 个 Bark 频率群。对于采样率为 8 kHz 的语音,选取 1 ~ 17 个 Bark 频率

群,得到的小波包子带的中心频率相差为1 Bark。Bark域频率 z 和赫兹(Hertz)域频率 f 的转换关系为

$$z = \begin{cases} 0.01f & 0 < f < 500 \\ 0.007f + 1.5 & 500 \leq f < 1220 \\ 6 \ln f - 32.6 & f \geq 1220 \end{cases} \quad (10)$$

多次实验证实^[14],如果对Bark域作进一步分解,可以更好地模拟人耳听觉感知特性,更加细致地描述语音,计算量也不大。本文对低频信号和高频信号进行更加细致的划分之后,使部分小波包子带中心频率相差1/4 Bark,提出如图2所示的60个子频带的Bark尺度小波包分解结构。图3为60个子频带小波包分解结构(图2)的子带中心频率和人耳实际频率特性的对比曲线。

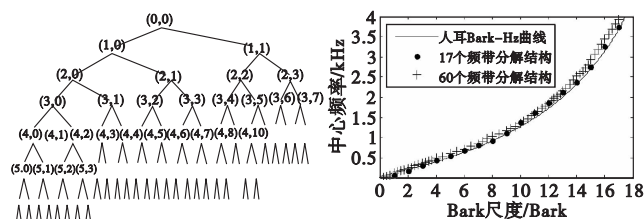


图2 60个频带Bark小波包分解结构图

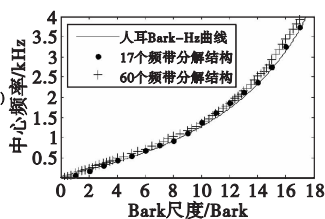


图3 小波包分解结构对人耳Bark域频率的模拟

由图2可以看出本文的分解结构是不完全小波包分解,降低了计算复杂度。从图3可以看出,本文新提出的60个频带小波包分解结构比常规17个频带对语音频率分解得更加细致,更加接近人耳实际频率特性曲线,能够更好地模拟人耳的听觉感知特性;经过本文小波包分解结构去噪后的语音的频谱连续性和整体的人耳听觉感知的自然度均不错。

4 仿真实验与分析

用MATLAB 7.4对本文算法进行了实验仿真,并和软、硬阈值函数方法进行了对比,验证了本文算法的有效性。具体实验中,纯净语音为语音数据库NOI ZEUS中的英语男声短句“The birch canoe slid on the smooth planks”,噪声来源于Noise × 92数据库的高斯白噪声和火车、汽车噪声。语音和噪声的采样率均为8 kHz,用13阶Daubechies小波对带噪语音进行7层60个频带Bark尺度小波包分解,分别在-5 dB、0 dB、5 dB、10 dB、15 dB信噪比下对高斯白噪声信号和0 dB、5 dB、10 dB信噪比下对火车噪声和汽车噪声信号进行了仿真实验。图4和5分别为5 dB高斯白噪声和汽车噪声背景下,软、硬阈值函数去噪方法和本文新阈值函数算法增强的语音波形图。

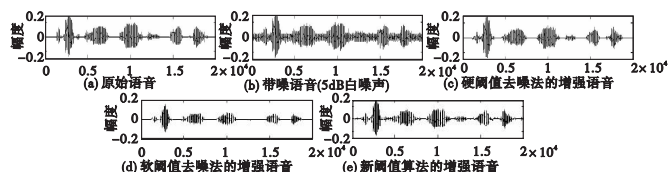


图4 高斯白噪声背景下三种算法增强语音的波形对比

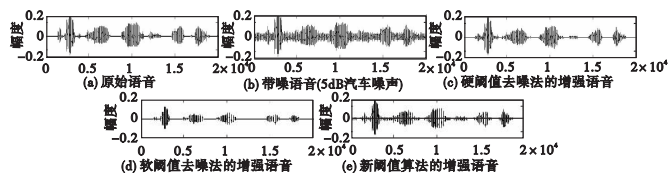


图5 汽车噪声背景下三种算法增强语音的波形对比

从图4和5可以看出,在高斯白噪声和有色噪声背景下,硬阈值函数滤除了大部分噪声,但也削弱了语音信号的细节,

造成了失真,这是硬阈值函数将绝对值小于阈值的系数直接置零所导致的结果;软阈值函数增强的语音和原始语音的幅度相差最多,失真最大;相反,经过本文新阈值函数增强的语音平滑度最好,语音的清音分量得到很好的保留,减少了语音失真,从直观上可以看出新阈值函数增强的语音最接近原始语音。

本文采用信噪比(SNR)和分段信噪比(SSNR)对实验后的增强语音质量进行客观评价,仿真实验结果如表1和2所示。

表1 三种算法增强语音的信噪比(SNR)比较 /dB

噪声类型	输入信噪比	输出信噪比		
		硬阈值函数去噪方法	软阈值函数去噪方法	新阈值函数去噪方法
白噪声	15	14.153 4	9.957 6	16.401 4
	10	11.150 1	7.737 7	13.338 3
	5	8.259 8	5.590 0	10.029 4
	0	5.836 7	3.773 2	6.843 4
	-5	3.338 8	2.129 8	3.522 1
火车噪声	9.3302	9.778 7	6.753 8	11.765 2
	4.330 1	6.797 4	4.6083	8.241 9
	-0.6698	3.774 7	3.101 1	4.244 0
	9.3302	9.269 6	6.763 9	11.091 7
	4.330 1	6.881 8	4.4152	8.749 9
	-0.669 8	4.692 8	2.7541	5.716 3

表2 三种算法增强语音的分段信噪比(SSNR)比较 /dB

噪声类型	输入信噪比	输入分段信噪比	输出分段信噪比		
			硬阈值函数去噪方法	软阈值函数去噪方法	新阈值函数去噪方法
白噪声	15	7.801 7	6.381 3	4.184 1	7.686 2
	10	5.230 3	4.854 7	3.085 8	6.032 1
	5	2.974 9	3.340 6	2.112 1	4.295 9
	0	1.240 1	2.356 3	1.386 3	2.915 6
	-5	0.171 8	1.455 9	0.796 7	1.721 8
火车噪声	9.330 2	5.220 3	4.130 5	2.650 5	5.310 3
	4.330 1	3.234 1	2.821 8	1.776 2	3.696 6
	-0.669 8	1.479 2	1.741 0	1.204 9	2.201 3
	9.330 2	5.356 6	4.057 1	2.653 9	5.337 6
	4.330 1	2.901 8	2.733 2	1.635 5	3.721 5
	-0.669 8	1.050 8	1.838 2	1.044 7	2.268 4

从表1和2可以看出,在相同输入信噪比和输入分段信噪比的高斯白噪声和有色噪声背景下,新阈值函数增强语音的信噪比和分段信噪比均比软、硬阈值函数增强的语音有明显提高,并很好地保留了原始信号的细节特征,减少了语音失真,说明本文算法是有效的。

5 结束语

本文提出的新阈值函数小波包变换语音增强算法,采用能更好地模拟人耳听觉感知特性的60个频带Bark小波包分解结构和一种新阈值函数进行语音增强。新阈值函数不仅是软、硬阈值函数的折中,而且对小波包系数绝对值小于给定阈值的区间用非线性函数压缩阈值,保证了函数的连续性。实验表明:本文算法对语音的增强效果在信噪比值和分段信噪比值上均优于传统软、硬阈值函数去噪方法;本文算法更好地保留了有用语音信息,有效地减少了语音信号的失真;新算法在高斯白噪声和有色噪声情况下均有效。(下转第137页)

