

基于多分辨率小波阈值去噪的响度补偿方法

张宝琳, 张玲华, 林志敏

(南京邮电大学, 江苏 南京 210003)

摘要: 进行多通道响度补偿时构建数字滤波器比较繁琐, 并且产生的噪声不易消除, 在增强了语音的同时, 噪声也被放大。为了解决这些问题, 文中提出利用小波阈值去噪和多分辨率分析的方法在小波域对消噪后信号进行响度补偿, 并且对补偿后的声强进行控制, 最后利用小波的完美重构特性对信号进行重构。一方面小波多分辨率分析能够满足多通道响度补偿对频带的非等宽划分, 并且小波的阈值去噪可以有效防止噪声的放大; 另一方面直接在小波域对补偿后声强进行控制, 可以使声强始终处于听觉的舒适域内。经实验仿真, 达到理想的效果。

关键词: 多分辨率分解与重构; 小波阈值去噪; 响度补偿; 增益控制

中图分类号: TP31

文献标识码: A

文章编号: 1763-629X(2012)12-0083-04

A Loudness Compensation Method Based on Multiresolution Wavelet Threshold Denoising

ZHANG Bao-lin, ZHANG Ling-hua, LIN Zhi-min

(Nanjing University of Posts & Telecommunications, Nanjing 210003, China)

Abstract: It is complex to construct digital filter to realize multi-bands loudness compensation. Also, the noise will be amplified when enhance person's voice because the noise is hard to eliminate. In this paper the method of using wavelet threshold denoising and multi-resolution analysis will be used to compensate the loudness of voice and control the gain of intensity of voice. At last, reconstruct the signal using new wavelet coefficients. On the one hand, multi-resolution analysis of wavelet can meet the requirement of dividing different bandwidth easily, also, it can avoid the noise amplification effectively by setting a thread of wavelet. On the other hand, using wavelet can control the gain of voice in wavelet domain directly and make the intensity of voice comfortable to listeners. According to experiment and simulation, it can achieve the ideal effectively.

Key words: multi-resolution analysis and reconstruction; wavelet threshold denoising; loudness compensation; gain control

0 引言

由于生活环境变得越来越复杂, 噪声污染也越来越严重, 人们在生活中许多的场所都需要对声音进行响度补偿。目前, 关于响度补偿算法中, 大多数算法都是在等频带划分基础上进行响度补偿, 例如等宽带滤波器组划分法^[1]。也有研究人员提出非等频带划分的响度补偿算法, 例如在 Bark 域分割子带频率间隔非等宽多通道补偿^[2-4]等方法。这些方法虽然对衰减的信号进行补偿, 但要将补偿后的语音控制在听者的最适听阈内, 都需要依赖外加的增益控制电路。小波的多分辨率分析能够对信号的频带进行多层次的非等宽划

分, 而这种划分方法又与实际中对听者进行听力测试完全一致。另外, 小波技术的成熟以及小波工具箱提供了大量的小波分析函数, 利用这些函数可以方便地对信号进行各种处理。因此, 文中利用小波变换的方法在小波域对语音信号进行多层分解并在每层设置适当的阈值^[5], 然后利用推导出的补偿算法对衰减信号进行响度补偿的和增益控制, 既有效地防止在增益时噪声被放大, 又使信号始终在听者的听觉的舒适范围内, 以保护听者的听力。通过 matlab 实验仿真, 利用小波变换的方法对语音信号进行响度补偿可以有效对声音进行增强, 达到理想效果。

1 响度补偿及增益控制原理

当声音的强弱小到人的耳朵刚刚可以听见时, 称为“听阈”。如果加大声音的强度, 使它大到人的耳朵感到疼痛, 这个阈值称为“痛阈”。人耳对不同频率上声音的强度的感觉是不同的^[6]。以不同频率对人们进

收稿日期: 2012-04-14; 修回日期: 2012-07-21

基金项目: 国家自然科学基金资助项目(60872105); 江苏省“青蓝工程”中青年学术带头人课题; 江苏省高校优势学科建设工程(PAPD)资助项目

作者简介: 张宝琳(1986-), 男, 硕士研究生, 研究方向为语音信号处理; 张玲华, 博士, 教授, 研究方向为语音信号处理。

行声压级不同的声强的测量可以分别得到听觉的“听阈—频率”曲线图。文中利用一个听障患者的“听阈—频率”曲线来模拟一个强度受到损失的声音,如图1所示。这里假设最一般的情况,设痛阈也随着频率变化。图1中,实线表示随频率的不同声音损失的大小,虚线表示随频率的不同人耳感觉到疼痛时阈值的变化。

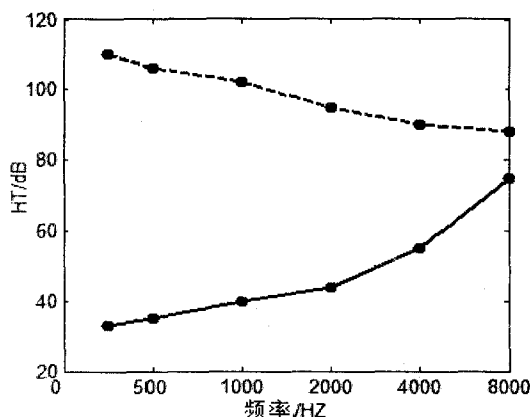


图1 声音损失及最大强度随频率变化曲线

由于不同频率段上的声音强度的损失不同,所以需要在不同的频带上进行不同的增益。二进小波变换可以实现信号的高频段粗分、低频段细分的要求,因此可以利用小波对语音信号进行四层分解,分解后小波系数 $CA_4, CD_4, CD_3, CD_2, CD_1$ 分别代表 $0 \sim 500\text{Hz}$, $500 \sim 1000\text{Hz}$, $1000 \sim 2000\text{Hz}$, $2000 \sim 4000\text{Hz}$, $4000 \sim 8000\text{Hz}$ 五个频段上的信号,然后可以对每个频段上的小波系数进行响度补偿,将人们听到的衰减信号恢复到信号原来的声音强度。有研究表明对高度衰减的高频信号进行高增益不但不能达到理想效果反而有负面影响^[7,8],所以文中仅采用小波变换对前四个频段($0 \sim 4000\text{Hz}$)进行响度补偿。

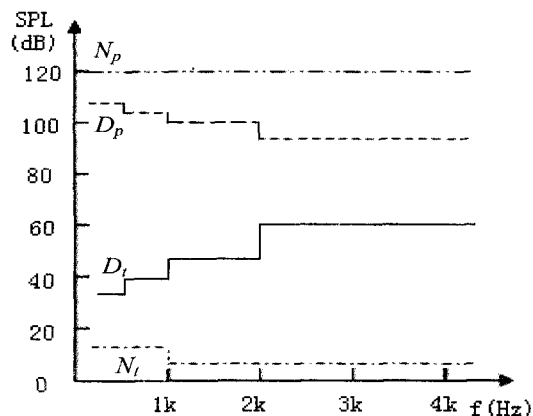


图2 各频段阈值近似曲线

对语音信号进行四层小波分解后,在四个不同的频段上分别对小波系数进行响度补偿,先将衰减的语音增强至衰减前的声音强度,但是由于声强的上限值也随着频率在变化,若将补偿后的信号直接送入听者

的耳中这必然会引起听者听觉的不适感,因此必须对补偿后的声音强度进行调整^[9,10]。由图1可以得到各频段阈值近似曲线(如图2所示),根据图2对声音的强度进行调整,调整的目标是:按一定的规则将低于 D_t 和高于 D_p 的声强调整到目标区间 D_t 和 D_p 之间。其中, N_t, D_t 分别表示调整前、调整后的听阈, N_p, D_p 分别表示调整前、调整后的痛阈。最后根据小波的完美重构特性,对处理后的小波系数进行重构,得到响度补偿和增益控制后的语音信号。

2 系统流程设计及小波响度补偿算法

2.1 算法流程设计

(1) 根据图2模拟产生一个在不同频段声强受到不同衰减的语音信号;

(2) 利用小波变换对该信号进行尺度为四的小波分解并由图2的声强增益关系图得到每个频段的响度补偿系数;

(3) 运用多分辨率小波阈值去噪的方法^[11,12],在各层设置不同阈值对小波系数进行处理,防止在增益语音信号同时噪声也被放大;

(4) 利用第二步获得的补偿系数对该小波系数进行响度补偿并根据图3进行增益控制,得到新的小波系数;

(5) 最后利用新的小波系数重构出响度补偿和增益控制后的语音信号。

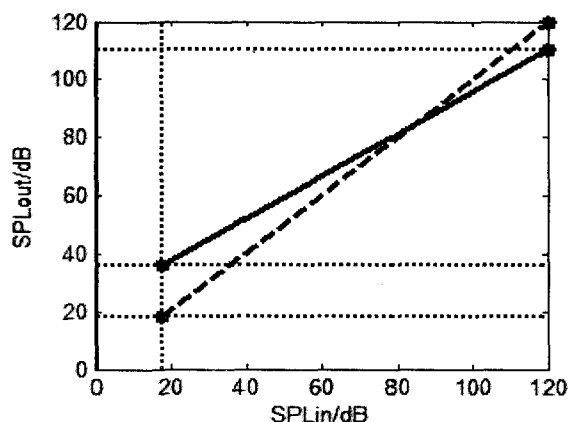


图3 声强调整前后的映射曲线

2.2 基于小波的响度补偿和增益控制算法

首先根据图2的 D_t 曲线由公式 $\gamma = 10^{\gamma/20}$ 计算出每个频段的补偿系数,其中 γ 为每频段的听阈值。然后对衰减后的语音信号进行四层小波分解得到每个频段的小波系数 $w_{i,k}$,然后设置适当的阈值 δ ,再对各频段的小波系数根据关系式 $w'_{i,k} = w_{i,k} \times \gamma$ 进行响度补偿。

经过响度补偿后,衰减的语音基本达到正常的声音强度,但是由于响度范围的变化,在进行补偿后还要

对声强进行调整。取 500 ~ 1000Hz 为例,画出声强调整前后的映射曲线,如图 3 所示,并根据图 3 推导出小波增益控制关系式如下:

设变换后的听阈为 D_i , 痛阈为 D_p 。变换前的听阈为 N_i , 痛阈为 N_p 。对应图 3 中所示的声强的映射关系,可设增益前后小波系数关系为:

$$W'_{j,k} = K_j W_{j,k} + C_j \quad (1)$$

其中, $W'_{j,k}$, $W_{j,k}$ 分别为增益后的小波系数和增益前的小波系数,单位为 dB。

$$K_j = \frac{D_p - D_i}{N_p - N_i} \quad (2)$$

$$C_j = D_i - K_j \times N_i \quad (3)$$

将关系式(3)代入(1)中可得

$$W'_{j,k} = D_i + K_j (W_{j,k} - N_i) \quad (4)$$

换元得

$$20 \lg w'_{j,k} = D_i + K_j (20 \lg w_{j,k} - N_i)$$

$$w'_{j,k} = 10^{(D_i + K_j (20 \lg w_{j,k} - N_i)) / 20} = 10^{(D_i - K_j N_i) / 20} \cdot w_{j,k}^{K_j} \quad (5)$$

关系式(5)就是某一频段小波系数的增益关系式。可以根据每一个频段的增益控制曲线得到 K_j, D_i, N_i , 利用关系式(5)就可以计算出增益后的小波系数,最后利用增益后的小波系数重构语音信号。

3 实验与仿真

首先读取原语音信号,然后根据图 2 模拟衰减后的语音,再将衰减的语音用小波响度补偿和增益控制算法进行响度补偿和增益控制得到增益后的语音。

图 4 中画出了原始语音,衰减后语音,以及增益控制后的语音的小波系数。图中横轴为小波点数,纵轴为小波幅值。

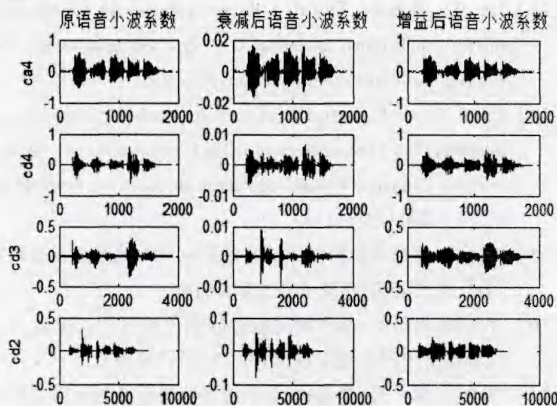


图 4 小波系数比较图

由于在高频的衰减极其严重,特别是 4000Hz 以上的频率衰减更是超过了 60dB,而痛阈又极其靠近听阈,并且人们 80% 的语音信息都包含在 4000Hz 以下,因此在进行响度补偿时舍弃代表清音的高频部分,只对表征代表 4000Hz 以下的语音信息的小波系数进行

变换和比较。其中 ca4 为 0 ~ 500Hz 小波系数,cd4 为 500 ~ 1000Hz 小波系数,cd3 为 1000 ~ 2000Hz 小波系数,cd2 为 2000 ~ 4000Hz 小波系数。

从图 4 中可以看出,低频系数代表的低频语音的幅值一般较大,因此变换后的小波系数稍有压缩,而代表高频信息的高频系数较小,经过响度补偿和增益控制后小波系数被放大。

图 5 中画出了原始语音,衰减后的语音以及增益后的语音时域波形图。图中的横坐标为时间样本,纵坐标为幅度的归一化值。

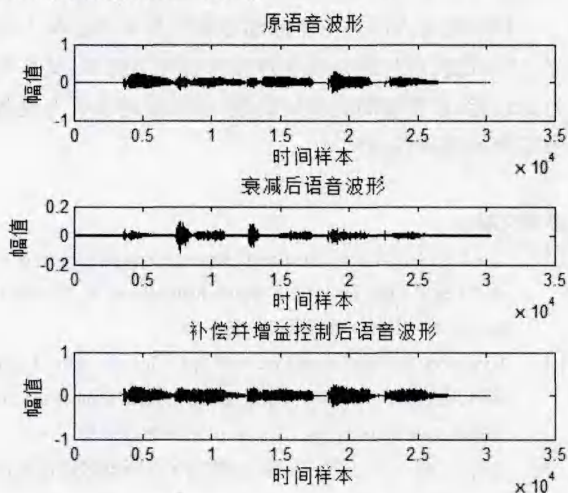


图 5 原语音、衰减语音、增益后语音时域波形图

从实验仿真的波形图上可以看出,经过在不同频段上衰减后,语音信号衰减严重,通过试听已经不能分辨出原语音。经过补偿和增益控制后的语音基本与原语音保持一致,但也可以看出增益后的语音与原语音有些偏差,这是对各频段进行增益控制所导致的。

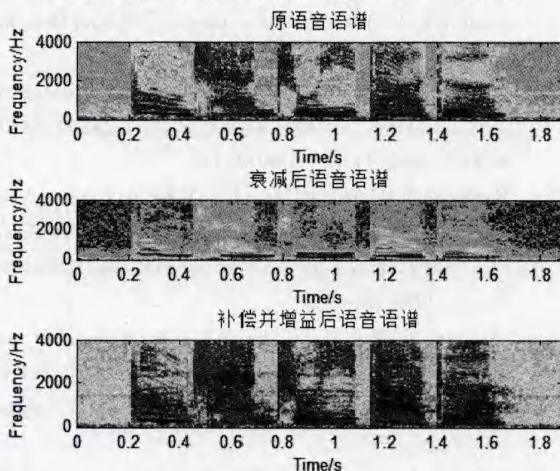


图 6 增益前后信号语谱比较图

图 6 画出了原语音,衰减语音及增益后语音的语谱图。为方便观察只画出了 4000Hz 以下频段的语谱图。可以看出经过补偿后,衰减掉的能量得到了补偿,语谱比较清楚,经过试听,声音强度跟原语音持平,达

到较理想的效果。

4 结束语

以上实验结果表明,通过小波变换在小波域进行响度补偿和增益控制,既符合实际需要又可以方便有效地对衰减信号进行响度补偿和增益控制。小波多分辨率分析既能够满足多通道响度补偿时对信号频带的非等宽划分,同时小波的阈值去噪的方法可以方便地对产生的噪声进行消除,有效防止噪声的放大。另外小波的完美重构特性使得重构后的语音失真很小。

仿真结果表明该方法能够显著提高听者的听力水平,达到比较理想的效果并且该算法较为简单,又具有好的性能,具有较高的实用价值,可以应用于听力设备及需要响度补偿的领域。

参考文献:

- [1] Lars S N, Jens S. Designing asynchronous circuits for low power: An IFIR filter bank for a digital hearing aid[J]. *Proceedings of the IEEE*, 1999, 87(2): 268-281.
- [2] Wyrsch S. Subband signal processing for hearing aids[C]// *Proceedings of the 1999 IEEE International Symposium on Circuits and Systems*. [s. l.]: [s. n.], 1999: 29-32.
- [3] 王青云, 赵力, 赵立业, 等. 一种数字助听器多通道响度补偿方法[J]. *电子与信息学报*, 2009, 31(4): 832-835.
- [4] 王青云, 赵力, 乔杰, 等. 符合人耳听觉特性的数字助听器子带响度补偿[J]. *应用科学学报*, 2008, 26(6): 580-585.
- [5] 李海东, 李青. 基于阈值法的小波去噪算法研究[J]. *计算机技术与发展*, 2009, 19(7): 56-58.
- [6] 乔杰. 利用人耳听觉特性的子带双声道回波抵消[J]. *应用科学学报*, 2007, 25(4): 331-336.
- [7] Miller-Hansen D R, Nelson P B, Widen J E, et al. Evaluating the Benefit of Speech Recording Hearing Aids in Children[J]. *American Journal of Audiology*, 2003, 12(2): 39-44.
- [8] Turner C. The limits of high frequency amplification[J]. *Hear Jour*, 1999, 52(2): 10-14.
- [9] 肖宪波, 胡广书, 刘春红. 基于形态学的数字助听器单频带压缩放大算法[J]. *清华大学学报(自然科学版)*, 2005, 45(12): 1680-1683.
- [10] Drake L A, Rutledge J, Cohen J. Wavelet analysis in recruitment of loudness compensation[J]. *Signal Processing*, 1993, 41(12): 3306-3312.
- [11] 曲天书, 戴逸松, 王树勋. 基于SURE无偏估计的自适应小波阈值去噪[J]. *电子学报*, 2002, 30(2): 266-268.
- [12] Pasti L, Walezak B, Massart D L. Optimization of signal denoising in discrete wavelet transform[J]. *Chemometrics and Intelligent Laboratory System*, 1999, 48(1): 21-34.
- [5] Stauffer C, Wel G. Adaptive background mixture models for real-time tracking[C]// *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Fort Collins, Colorado: [s. n.], 1996: 248-252.
- [6] Long W, Yang Y H. Stationary background generation: An alternative to the difference of two images[J]. *Pattern Recognition*, 1990, 23(12): 1351-1359.
- [7] Lipton A J, Fujiyoshi H, Patil R S. Moving target classification and tracking from real-time video[C]// *Proceeding of WACV 98*. Fort Collins: [s. n.], 1998: 8-14.
- [8] Horprasert T, Harwood D, Davis L S. A statistical approach for real-time robust background subtraction and shadow detection[C]// *IEEE Frame-rate Applications Workshop*. California: [s. n.], 1999: 728-812.
- [9] Anderson C, Burt P, van der W G. Change detection and tracking using pyramid transformation techniques[C]// *Proceeding of Intelligent Robots and Computer Vision IV*. California: [s. n.], 1985.
- [10] Elgammal A, Harwood D, Davis L. Non-parametric model for background subtraction[J]. *Computer Vision-ECCV*, 2000, 20(7): 751-767.
- [11] Wren C R, Azarbayejani A, Darrell T, et al. Pfister: Real-time tracking of the human body[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, 19(7): 780-785.
- [12] Power P W, Schoonees J A. Understanding background mixture models for foreground segmentation[C]// *Proceedings of Image and Vision Computing*. New Zealand: [s. n.], 2002: 79-83.
- [13] Lee D S, Hull J J, Erol B. A Bayesian framework for Gaussian mixture background modeling[C]// *Proceeding of Image Processing*. California: [s. n.], 2003: 974-976.
- [14] Gao X, Boulton T, Coetzee F, et al. Error analysis of background adaption[C]// *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Bethlehem: [s. n.], 2000: 503-510.
- [15] 曲丽丹. 交通视频监控中的背景提取与运动跟踪方法研究[D]. 哈尔滨: 哈尔滨工业大学, 2010.
- [16] 苏礼坤, 黄继红. 基于帧间差分的彩色视频背景提取[J]. *成都信息工程学院学报*, 2010, 25(2): 167-171.
- [17] 林洪文, 涂丹, 李国辉. 基于统计背景模型的运动目标检测方法[J]. *计算机工程*, 2003, 29(16): 97-99.
- [18] 王占全, 徐慧. 精通 Visual C++ 数字图像处理技术与工程案例[M]. 北京: 人民邮电出版社, 2009.
- [19] 魏宗舒. 概率论与数理统计教程[M]. 北京: 高等教育出版社, 2008.

(上接第82页)