

文章编号: 1001-0920(2006)09-1033-04

基于仿生小波变换和自适应阈值的语音增强方法

杨玺, 樊晓平

(中南大学信息科学与工程学院, 长沙 410075)

摘要: 提出一种基于仿生小波变换以及自适应阈值的语音增强方法。含噪语音通过仿生小波变换后, 针对不同的尺度采用不同的阈值函数进行去噪。由于在小波变换过程中考虑了人耳的听觉特性, 所提出方法优于基本的小波语音增强方法。实验表明, 该方法在多种噪声条件下均具有较好的语音增强效果。

关键词: 语音增强; 仿生小波变换; 自适应阈值

中图分类号: TN 912

文献标识码: A

Speech Enhancement Based on Bionic Wavelet Transform and Adaptive Threshold Functions

YANG Xi, FAN Xiaoping

(College of Information Science and Engineering, Central South University, Changsha 410075, China)

Correspondent: YANG Xi, E-mail: xyyoung@163.com

Abstract: Based on the bionic wavelet transform and adaptive threshold functions, an improved wavelet-based speech enhancement method is presented. Due to the integration of human auditory system model into the wavelet transform, the main advantage of the proposed method is that the over thresholding of speech segments which is usually occurred in conventional wavelet-based speech enhancement schemes can be avoided. As a consequence, the enhanced speech quality of the proposed method is increased substantially from those of conventional approaches.

Key words: Speech enhancement; Bionic wavelet transform (BWT); Adaptive threshold function

1 引言

人们在语音通信过程中不可避免地会受到来自周围环境、传输媒介等引入的噪声以及通信设备内部的电噪声乃至其他讲话者的干扰。这些干扰最终将使接收者收到的语音已非纯净的原始信号。语音增强的一个主要目标就是从含噪语音中提取尽可能纯净的原始语音, 从而改进语音质量, 提高语音可懂度。迄今, 人们已提出了许多语音增强的算法, 但由于语音信号本身所具有的复杂性和非平稳性, 加之噪声的多样性, 这些算法尚不尽如人意。而小波变换作为一种新的多分辨分析方法, 可同时进行时域和频域分析, 具有时频局部化和变分辨特性, 特别适于处理非平稳信号。

近年来, 人们在将小波技术应用于语音增强方面做了大量工作, 取得了很好的去噪效果^[1~3]。由于语音的非平稳特性, Pinter 等^[1]在去噪过程中引入了人耳的听觉掩蔽特性, 将临界带宽与小波变换相结合, 改善了语音增强的效果。Mohammed 等^[2]提出了基于时间自适应的小波阈值语音增强方法, 其时间自适应是通过引入 Teager 能量算子实现的, 但当语音信号受低噪声干扰时仍存在过度阈值的问题。为了消除小波去噪过程中产生的音乐噪声, Hu 等^[3]提出了基于小波去噪的低方差的谱估计器, 改进了语音增强的效果。但上述语音增强方法对于含有有色噪声的语音信号, 去噪效果仍然欠佳, 而且仍存在少量的音乐噪声。针对语音的非平稳特性, 为了

收稿日期: 2005-06-06; 修回日期: 2005-08-31

基金项目: 国家自然科学基金项目(69975003)。

作者简介: 杨玺(1981—), 女, 江西新余人, 博士生, 从事语音处理、信号处理等研究; 樊晓平(1961—), 男, 浙江绍兴人, 教授, 博士生导师, 博士, 从事智能控制、机器人控制、虚拟现实等研究。

消除去噪过程中的音乐噪声, 本文将文献[4]中的仿生小波变换应用于单声道语音增强系统, 并在小波去噪的同时根据语音的特性采用多样的阈值函数

2 仿生小波变换

人耳是相当完美的声音信息采集和处理器, 它由外耳、中耳、内耳组成, 其中外耳和中耳的主要功能是对外界传来的声音进行增强和补偿, 而对声音的接收和频率分解则主要在内耳耳蜗的基底膜上进行. 基底膜对于信号的分辨能力随着频率的不同而发生变化, 对低频的分辨力要高于高频的分辨力. 在人耳对声音信息的采集处理过程中, 声音的幅度信息比相位信息更加重要. 因此, 本文在对语音信号进行处理的过程中一定要考虑人耳的听觉特性. 以往的方法都是在小波分解过程中结合人耳的临界带宽, 并不能很好地反应人耳对信号幅度以及频率的听觉特性, 所以本文将仿生小波变换引入语音增强处理过程.

基于Giguere等于1994年提出的听觉外周系统的完整电路模型, Yao等^[4]将主动耳蜗机理与小波变换相结合, 提出了仿生小波变换(BW T), 它是一种基于生物模型的自适应时频分析方法. 与基本的小波变换相比, 仿生小波变换的优势在于它在时频域的尺度不仅可以调节信号的频率进行调节, 而且可以随着信号的瞬时幅度以及一阶微分系数自适应调节. 在基本的小波变换中, 母小波函数 $h(t)$ 必须满足容许条件, 所以可用其包络函数 $\tilde{h}(t)$ 来表示 $h(t)$ 以及信号 $f(t)$ 的小波变换

$$h(t) = \frac{1}{\sqrt{a}} \tilde{h}(t) \exp(j\omega t), \quad (1)$$

$$(W T_f)(\tau, a) =$$

$$\frac{1}{T\sqrt{a}} \int f(t) \tilde{h}\left(\frac{t-\tau}{a}\right) \exp(-j\omega\left(\frac{t-\tau}{a}\right)) dt \quad (2)$$

为了模拟主动耳蜗机理, 仿生小波变换在基本小波变换的母小波函数中引入了一个新的参数 T (其中 $T > 0$, 是被变换信号的幅度及其一阶差分的函数). 由于参数 T 的引入, 仿生小波变换中母小波函数的包络可根据目标信号的特性在时域上自适应地进行调节. 其对应的信号 $f(t)$ 的仿生小波变换以及参数 T 可表示为

$$(B W T_{\tau f})(\tau, a) =$$

$$\frac{1}{T\sqrt{a}} \int f(t) \tilde{h}\left(\frac{t-\tau}{aT}\right) \exp(-j\omega\left(\frac{t-\tau}{a}\right)) dt, \quad (3)$$

$$T(\tau + \Delta\tau) =$$

$$\left(1 - G_1 \frac{B W T_s}{B W T_s + |B W T_f(\tau, a)|}\right)^{-1} \left(1 + G_2 |B W T_f(\tau, a)/\Delta\tau|\right)^{-1} \quad (4)$$

根据Giguere的听觉外周系统的完整电路模型, G_1, G_2 和 $B W T_s$ 分别为主动因子和饱和常量, $B W T_f$ 则为 τ 时刻的仿生小波变换的系数. 尽管仿生小波变换与基本小波变换相比具有很多优点, 但在早期, 由于计算十分费时而使其应用受到限制. 为此, 对仿生小波变换的计算作了进一步的改进, 从而发现对于Morlet母小波, 仿生小波变换与基本小波变换存在以下关系:

$$B W T(a, \tau) = \frac{1}{C} \cdot \frac{1.772.5T_0}{\sqrt{1+T^2}} \cdot W T(a, \tau). \quad (5)$$

由于有了快速计算方法, 仿生小波在语音信号处理领域得到了应用^[5]. 本文将进一步将其应用于单声道语音增强系统中.

3 基于BW T和多样阈值函数的语音增强方法

目前, 小波变换已成为消除信号或图像中噪声的强大工具, 基于阈值的小波域语音增强算法是一种应用十分广泛而且简单有效的去噪方法. 本文在常规小波阈值去噪方法的基础上, 提出了一种基于仿生小波变换以及多阈值的语音增强方法. 本文的算法步骤如下:

Step1 含噪语音信号的仿生小波变换

对含噪语音信号进行仿生小波变换, 得到含噪信号的小波系数. 假设

$$y(i) = x(i) + e(i), \quad i = 0, 1, \dots, N-1. \quad (6)$$

其中: y 为含噪语音信号, x 为纯净语音信号, e 为加性噪声, N 为信号 x 的长度. 将仿生小波变换应用于含噪语音信号后, 可以得到

$$B W T_y = B W T\{y\}, \quad (7)$$

其中 $B W T_y$ 为含噪语音信号经小波变换后的系数.

Step2 小波系数阈值处理

在进行小波系数阈值处理时, 最为关键的是阈值以及阈值函数的选取. 语音信号大多只存在于较低频率的子带中, 通过设定适当的阈值对小波系数进行阈值处理便可以滤除噪声. 小波阈值方法的关键在于阈值的选定. Donoho和Johnstone提出了一种用于消除加性白噪声的统一阈值 λ . 本文考虑到实际噪声在小波域不同尺度上的差别, 采用基于尺度的阈值方法, 阈值定义为

$$\lambda = \sigma_j \sqrt{2 \log N}, \quad (8)$$

其中 σ_j 为噪声级别. 噪声级别估计如下:

$$\sigma_j = M A D_j / 0.6745, \quad (9)$$

其中 $M A D_j$ 为尺度 j 上的中值标准差.

常用的阈值函数有硬阈值函数和软阈值函数. 硬阈值函数将低于阈值的信号忽略不计, 高于阈值

的信号保留; 软阈值函数同样将低于阈值的信号忽略不计, 高于阈值的信号只保留阈值以上的部分. 实验表明, 硬阈值函数对噪声清除不够干净, 而且在一些地方会产生突变, 使处理后的语音混有类似音乐的噪声; 软阈值函数虽然噪声去除很彻底, 但对原始语音的损害较大, 使语音的清晰度大大降低. 综合上述两种方法, 发现二者都存在时 - 频不连续的问题, 因而采用文献[6]提出的修正自定义阈值函数, 该函数近似为软阈值函数与硬阈值函数的线性组合, 它能在阈值点 λ 处保持阈值函数的连续性, 即

$$T(x) = \begin{cases} x - \text{sgn}(x)(1 - \alpha)\lambda & |x| > \lambda \\ 0 & |x| \leq \lambda \\ \alpha \left(\frac{|x| - \lambda}{\lambda} \right)^2 \left\{ (\alpha - 3) \left(\frac{|x| - \lambda}{\lambda} \right) + 4 - \alpha \right\} & \text{其他} \end{cases} \quad (10)$$

其中: $0 < \lambda < \infty$, $0 < \alpha < 1$. 当 $\alpha = 0$ 时, $T(x)$ 近似为软阈值函数; 当 $\alpha = 1$ 时, $T(x)$ 近似于硬阈值函数. λ 为截取值, 使得所有小于 λ 的系数均设为零, 从而消除了部分微小的干扰噪声. 由于语音大多存在于较低频率的子带中, 对于语音信号的低频部分希望在去除噪声的同时尽量保留语音信号. 为此, 需根据语音信号的特点来确定阈值函数中的 α 值. 对于同一尺度上的每个频率子带, 每个子带的噪声级别均不相同. 这里让 α 值根据噪声级别自动调节, 使其与噪声标准差成正比, 即

$$\alpha = \alpha_0 + \alpha_1 \sigma \quad (11)$$

其中: σ 为 σ 归一化处理后的参数值; α_0 和 α_1 分别为待定常数, 确保 α 值在 $0 \sim 1$ 之间.

这样, 对于噪声级别较低的信号所使用的阈值函数便近似于硬阈值函数, 即选取较大的 α 值, 可以较好地保留原始语音; 而对于噪声级别较高的信号所使用的阈值函数近似于软阈值函数, 即选取较小的 α 值, 可以较好地去除噪声. 因此, 无需对信号信噪比进行先验估计, 而且消除了以往去噪方法中产生的音乐噪声.

Step3 仿生小波逆变换, 获取增强的语音信号

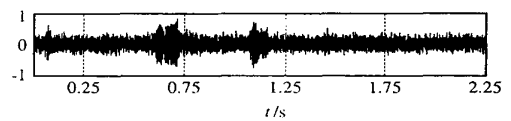
利用经阈值处理后的小波系数进行原始语音的小波重建, 即对处理后的小波系数进行仿生小波逆变换, 从而获取增强后的语音信号

$$\hat{x} = BW T^{-1}(T(BW T y)) \quad (12)$$

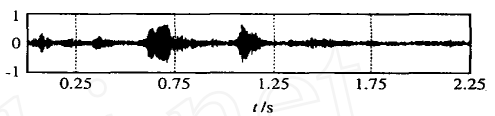
4 仿真结果和分析

根据本文提出的语音增强算法, 采用 T M IT 语音库中的多种纯净语音样本进行反复试验, 选取不同的噪声类型, 按不同的信噪比合成各种实验所需的含噪语音信号. 仿真时针对所选取的含噪语音信

号, 首先采用仿生小波变换, 其母小波函数选取 Morlet 函数, 变换尺度为 22, 噪声分别采用白噪声、F16 噪声、pink 噪声和 factory 噪声, 并在不同的信噪比 ($-10 \sim 10$) 下进行反复试验. 在阈值处理过程中, 考虑到语音信号的连续性和频率的变化, 选取 $\gamma = \lambda/2$, α_0 和 α_1 分别取 0.84 和 0.1, 从而使得去噪效果更佳. 图 1 给出了信噪比为 0, 噪声类型为 pink 噪声时的实验结果.



(a) 含噪语音信号 (pink 噪声, SNR = 0)



(b) 增强后的语音信号

图 1 信噪比为 0 的语音增强仿真结果

在不同的信噪比情况下, 采用本文提出的增强方法对多种含噪语音的仿真结果如图 2 所示.

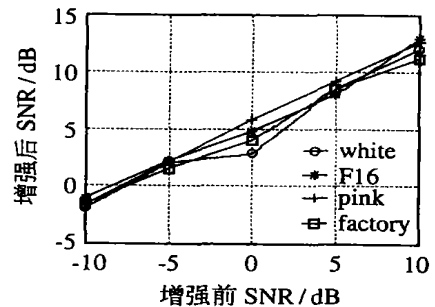


图 2 不同噪声的含噪语音的语音增强仿真结果

当含噪语音信噪比为 0 时, 分别采用本文方法、常规小波阈值去噪法以及单纯的仿生小波去噪法进行语音增强的仿真结果如表 1 所示.

表 1 信噪比为 0 的语音增强后的 SNR 值 / dB

增强方法	噪声类型			
	white	F16	pink	factory
本文方法	2.875 2	4.792	5.850 1	4.047 1
仿生小波去噪	2.548 7	2.513 2	4.513 2	2.813 2
常规小波去噪	1.076 2	1.365 0	1.416 3	0.515 7

通过对反复实验的数据进行分析, 从改善信噪比的角度看, 本文方法比常规小波去噪以及单纯的仿生小波去噪法效果好, 尤其在低信噪比时增强效果相对较好, 而且降低了音乐噪声. 针对各种不同的噪声环境, 本文方法对于白色噪声的增强效果相对欠佳. 以往的常规小波去噪方法中, 所有小波系数均采用相同的阈值, 从而忽略了噪声的特色, 本文方法

则能根据小波系数的尺度和噪声的特性动态地调整阈值,从而更好地抑制了噪声。此外,考虑到非平稳噪声的存在,通过采用仿生小波变换和阈值函数的变化,对语音信号进行相应的修正和处理,从而更好地保留了原始纯净语音。

5 结 语

本文提出的基于仿生小波变换以及多阈值的方法,由于考虑了人耳特性,而且对不同的尺度采用了不同的阈值方法,对于不同的尺度,其阈值函数也自适应地发生改变。因此,该方法在多种噪声环境中,尤其对于有色噪声和非平稳噪声具有较好的去噪效果。

参考文献(References)

- [1] Pinter Istvan. Perceptual Wavelet-representation of Speech Signals and Its Application to Speech Enhancement [J]. *Computer Speech and Language*, 1996, 10(1): 1-22
- [2] Mohammed Bahoura, Jean Rouat. Wavelet Speech Enhancement Based on the Teager Energy Operator [J]. *IEEE Signal Processing Letters*, 2001, 8(1): 10-12
- [3] Hu Y, Loizou, Philipos C. Speech Enhancement Based on Wavelet Thresholding the Multitaper Spectrum [J]. *IEEE Trans on Speech and Audio Processing*, 2004, 12(1): 59-67.
- [4] Yao J, Zhang Y T. Bionic Wavelet Transform: A New Time-frequency Method Based on an Auditory Model [J]. *IEEE Trans on Biomedical Engineering*, 2001, 48(8): 856-863
- [5] Yao J, Zhang Y T. The Application of Bionic Wavelet Transform to Speech Signal Processing in Cochlear Implants Using Neural Network Simulations [J]. *IEEE Trans on Biomedical Engineering*, 2002, 49(11): 1299-1309
- [6] Yoon Byung-Jun, Palghat P Vaidyanathan. Wavelet-Based Denoising by Customized Thresholding [A]. *Proc 29th Int Conf on Acoustics, Speech, and Signal Processing (ICASSP)* [C]. Montreal, 2004: 925-928

2006 中国控制与决策学术年会在津举行

本刊讯 2006 中国控制与决策学术年会(18th CDC)于7月27日至30日在天津召开。来自国内部分高等院校和科研院所的140多名专家和代表参加了本届年会。日本、韩国等国的学者出席会议,并参加了“电动汽车、运动控制和超级电容应用”专题组的研讨。

7月27日上午举行了大会开幕式。CDC年会组织委员会主席、河北工业大学党委副书记孙鹤旭教授主持会议, CDC年会程序委员会主席、东北大学校长助理王福利教授致开幕辞,河北工业大学副校长杨庆新教授致欢迎辞。河北省教育厅、中国兵工学会自动控制专业委员会等单位分别发来贺信。

本届年会邀请5位国内外知名学者做了大会专题学术报告。报告人和报告题目分别是:东北大学张嗣瀛院士:复杂系统的自聚集;日本东京大学堀洋一教授: Motion Control of Electric Vehicle — Future Electric Vehicle Powered by Super Capacitor; 上海交通大学席裕庚教授: 预测控制系统的稳定性分析; 新加坡南洋理工大学王邨维教授: Feed Forward Control Supplements Feedback Control: Iterative Approach; 北京科技大学孙一康教授: 提高热连轧板材厚度精度的措施。这些报告分别介绍了不同前

沿学科的最新研究成果和未来发展方向,受到与会代表的热烈欢迎和普遍好评。

会议期间发行了《2006 中国控制与决策学术年会论文集》,本届年会论文集共收入各方面论文383篇。在分组交流中,代表们宣读了论文,进行了广泛交流和深入研讨。经过民主推选和专家评议,本届年会评选出了9篇优秀学术论文。

在大会闭幕式上, CDC年会指导委员会主任委员张嗣瀛院士为优秀论文作者颁奖, CDC年会程序委员会副主席、东北大学信息科学与工程学院院长刘建昌教授致闭幕辞,江南大学通信与控制工程学院院长纪志成教授盛情邀请广大同行明年前往江南大学参加第19届CDC年会。

会议期间,鲁东大学科研处处长刘晓华教授正式提出申办第20届CDC年会。本会指导委员会、顾问委员会和程序委员会经过研究,同意鲁东大学的申办报告,决定第20届CDC年会在山东烟台举行。

本届年会由《控制与决策》编辑委员会主办,中国航空学会自动控制专业委员会、中国自动化学会应用专业委员会、中国系统仿真学会仿真方法与建模专业委员会、中国人工智能学会智能控制与管理专业委员会联合协办,河北工业大学具体承办。