

基于小波变换的语音信号去噪^{*}

李蕴华

(南通工学院 信息工程系, 江苏 南通 226007)

摘要: 讨论了离散小波变换在语音去噪中的应用。根据语音中浊音段和清音段的特点, 采用了不同的阈值方案, 可以保证在失真较小的前提下, 获得更好的去噪效果。

关键词: 小波变换; 语音去噪; 阈值

中图分类号: TN850.5

文献标识码: A

文章编号: 1671-5322(2002)02-0032-

04

语音信号的去噪是语音处理的一个重要课题。当噪声与语音的频谱相似时, 传统的单纯时域滤波或频域滤波往往无法达到很好的效果。离散小波变换是一种时-频分析法, 在时频两域都能表征信号的局部特性。利用小波变换实现信号的去噪, 具有较好的效果。

在本文中, 笔者将语音信号划分成浊音和清音两部分, 根据它们各自的特点, 在采用小波变换进行去噪时, 选择了不同的阈值方案进行了性能测试。

1 离散小波变换理论

有限序列 $s(n)$ 的离散小波变换(DWT)定义为:

$$\text{DWT}(s, 2^j, k2^j) = D_{j,k} = \sum_{n \in Z} s(n) h_1(n - 2^j k),$$

$$C_{j,k} = \sum_{n \in Z} s(n) h_0(n - 2^j k)$$

$$(j, k, m \in Z)$$

反变换 IDWT 定义为:

$$s(n) = \sum_{j=1}^J \sum_{k \in Z} D_{j,k} g_1^j(n - 2^j k) +$$

$$\sum_{n \in Z} C_{j,k} g_0^j(n - 2^j k)$$

式中 $h_0(n)$ 和 $g_0(n)$ 分别被称为尺度序列和对偶尺度序列, $h_1(n)$ 和 $g_1(n)$ 分别被称为小波序列和对偶小波序列。低通滤波器 h_0 及带通

滤波器 h_1 形成了一对镜像滤波器组 $h_1(n) = (-1)^{1-n} h_0(1-n)$ 。 $C_{j,k}$ 和 $D_{j,k}$ 分别被称为 2^{-j} 分辨率下的离散逼近信号(低频系数)和离散细节信号(高频系数)。

$D_{j,k}$ 和 $C_{j,k}$ 可由 Mallat 塔式算法算出, 运算过程为:

输入: $C_{0,k}$ (输入序列 $s(n)$) J (分解层次)

运算过程: for $j = 1$ to J

$$\{D_{j,k} = \sum_{n \in Z} h_1(n - 2k) C_{j-1,m};$$

$$C_{j,k} = \sum_{n \in Z} h_0(n - 2k) C_{j-1,m};$$

}

输出: $C_{j,k}, D_{j,k}$ 。 ($C_{j,k}, D_{j,k}$ 分别为第 j 个分解层次上第 k 个点的逼近信号和细节信号)

通过滤波器 h_1 和 h_0 , 将信号分解成低频空间 C_1 和低频空间 D_1 , 然后每次只对低频空间进一步分解, 对每层的高频系数不再继续分解。

对于小波各层高频系数可采用一定的阈值进行量化处理, 以去除噪声。然后根据各层低频系数及量化后的高频系数, 实现信号的重构。

2 小波去噪中阈值的选择

在对小波系数进行量化时, 选用何种阈值规则, 会影响到最终去噪效果的好坏。在保证失真小的前提下, 尽可能提高信噪比是选择合适阈值的原则。阈值方案一般有两种: 硬阈值法及软阈

* 收稿日期: 2001-03-18

作者简介: 李蕴华(1965-), 女, 江苏南通市人, 南通工学院讲师, 硕士, 研究方向为数字信号处理。

值法。前者主要保留信号的低频成分,衰减了大部分高频成分,使得主要表现为高频信号的噪声部分能被有效去除,但语音中的大量高频部分(如清音)也被去掉,产生失真。故我们采用软阈值处理。软阈值方法有(1)采用 Stein 的无偏似然估计原理(SURE)进行阈值选择。对一个给定的阈值 t ,得到它的似然估计,再将非似然 t 最小化,就得到了所选的阈值(2)用极大极小原理(Minimaxi)选择阈值。它以处理后信号与原始信号的最大概率逼近为约束条件(3)启发式阈值选择。它是一种最优预测变量阈值选择方法(4)固定阈值形式,其值为 $\sqrt{2\lg(\text{length}(s))}$,其中 $\text{length}(s)$ 为信号的长度。在这4种方法中,前两种阈值选取规则比较保守(它只将部分系数置0),当信号的高频信息有一部分在噪声范围内时,采用这两种阈值,可以将弱小信号提取出来;后两种阈值在去除噪声时,显得更为有效,但有可能把有用的高频特征当作噪声去除。

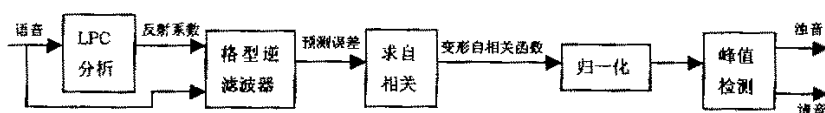


图1 清音和浊音的区分

Fig.1 Division of unvoiced and voiced speech

对于浊音来说,由于其具有准周期性,归一化变形自相关函数在基音周期可能存在的 2.5~15 ms 处会出现尖峰(峰点值大于 0.25),该尖峰位置其实就是基音周期判断的依据。而对于清音来说,则无明显峰起。用这种方法可以准确的区分出清浊音。

在整个过程中,分析窗的最小长度的选取,应是两倍的基音周期,太小无峰点出现,而太大则引入了不希望有的平滑。

另外,通过实验我们发现,利用该方法也能准确的区分含噪语音的清、浊音段。

4 去噪实验及结果

利用话筒将语音信号采集到计算机,并人为掺入高斯白噪声,生成数据文件供实验使用。

在去噪实验中,我们选用了 Biorthogonal 函数作为小波函数,它具有线形相位特性,主要用于信号的重构中。实验表明,采用这种函数能得到较好的去噪效果。小波分解层次 $J=6$ 。

在对高频小波系数量化中,对于浊音段及清音段我们比较了各种不同的软阈值方案,发现采

语音信号可以被分成浊音段及清音段两部分,这两部分的主要区别是:浊音呈现出准周期性,其周期即为该段的基音周期,且含有较多的低频成分,容易和低频噪声区别开。清音的信号波形类似于白噪声,与浊音相比,频率较高且无周期性。若语音中掺入了含高频成分的噪声,对浊音段和清音段应采用不同的阈值方案,才能获得最佳的去噪效果。由前面的分析可知,对于清音,可采用(1)(2)两种阈值规则,不至于损失过多的有用信号,而对于浊音,可采用(3)(4)两种阈值规则,更有效地去除噪声。

3 浊音段和清音段的划分

由于针对清、浊音我们采用不同的阈值方案,故准确划分清、浊音段是一个关键性的问题。逆滤波法可以有效的判别是否浊音段,并可确定基音周期。逆滤波法过程见图1:

用固定阈值 $\sqrt{2\lg(\text{length}(s))}$ 形式,由于去掉了较多的高频成分,引入了平滑,无论是对于清音还是浊音,都存在较大失真(如图2),故在语音去噪中我们不采用该阈值。

为得到较好的去噪效果,经实验发现,对于浊音段,可采用启发式阈值方案,而对于清音段,则采用 SURE 阈值方案。图3及图4是采取不同的阈值法下的浊音、清音的原始波形和去噪后波形。

从图3和图4可以看出,对于浊音,采用启发式阈值法比采用 SURE 阈值法可获取更强的去噪效果,但是对于频率成分与噪声很相似的清音,启发式阈值法会去掉一些有用的清音,产生失真,而应用 SURE 阈值则可保留这部分清音。由此可见,针对浊、清音不同特点,浊音部分选取启发式阈值方案,清音部分选取 SURE 阈值方案,可使整个语音段能达到最佳去噪效果。

我们用话筒录下语音“工学院”3个字,采样频率为 20 kHz,数据长 30 000 点。加入白噪声,构成含噪信号(信噪比 $RSN=1$ dB)。利用逆滤波法将含噪语音分成浊、清音段,浊音段约有 1 400 点,剩余的是清音段及无声段。在利用小波

变换进行分解及信号重构中,对于浊音段,我们采用启发式阈值,清音及无声段则采用 SURE 阈值规则。去噪后的语音波形如图所示(见图 5)。从

图 5 可以看到,去噪后的信号中,不管是清音或浊音段,失真都较小,且其中噪声部分得到了较大程度的削减,测试结果其信噪比约提高了 8 dB。

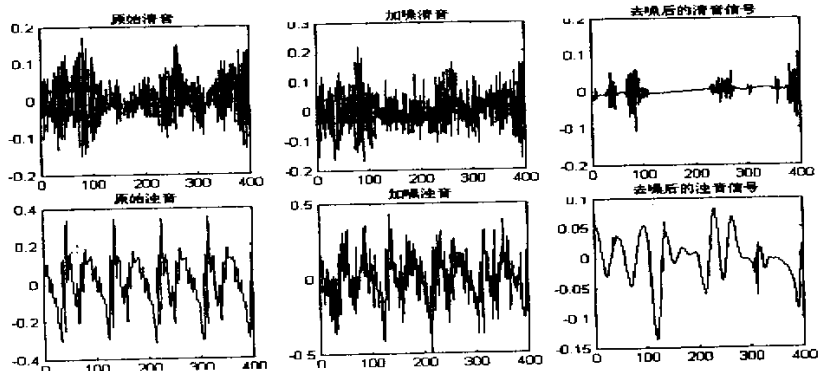


图 2 采用固定阈值时去噪效果

Fig.2 The effect of denoising by using stable threshold

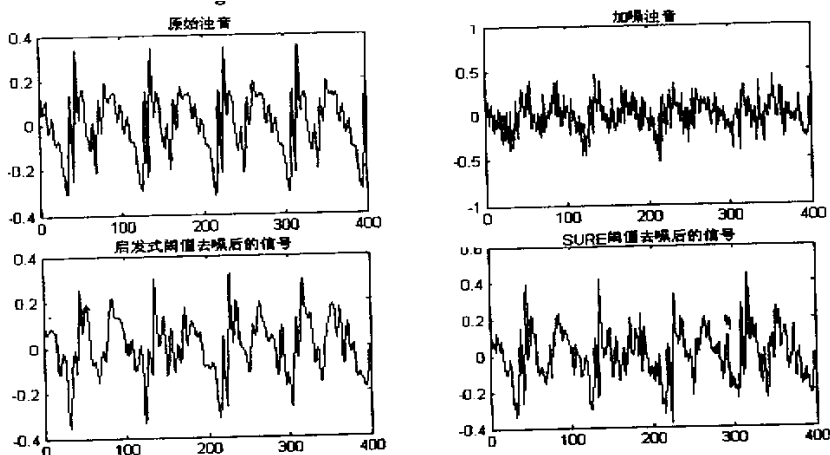


图 3 不同阈值时浊音段去噪分析

Fig.3 The denoising of voiced speech by using defferent threshold

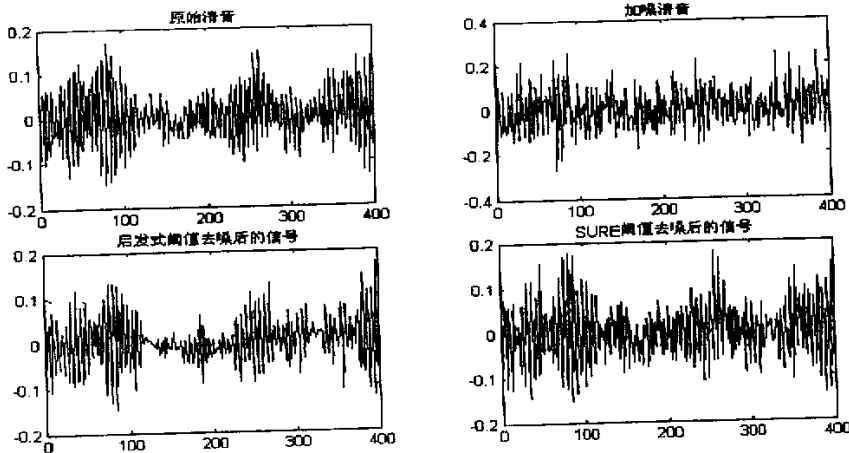


图 4 不同阈值时清音段去噪分析

Fig.4 The denoising of unvoiced speech by using defferent threshold

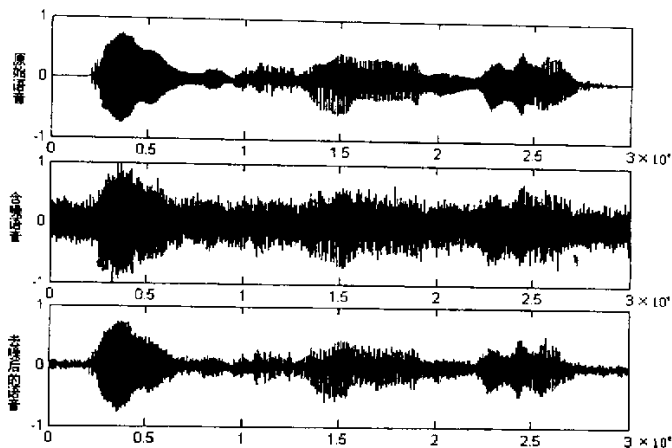


图 5 一段语音信号的去噪分析

Fig.5 The denoising analysis of a speech segment

5 结束语

采用了不同的阈值进行小波去噪 ,并获得了较好的去噪效果。

本文根据语音清音段及浊音段的不同特点 ,

参考文献 :

- [1] 秦前清.实用小波分析[M].西安 西安电子科技大学出版社 ,1998.
- [2] 拉宾那 L R.语音信号数字处理[M].北京 科学出版社 ,1984.
- [3] 张贤达.现代信号处理[M].北京 清华大学出版社 ,1998.
- [4] 李世雄.小波变换及其应用[M].北京 高等教育出版社 ,1997.

Speech Signal Denoising based on Wavelet Transform

LI Yun-hua

(Department of information Engineering of Nantong Institute of Technology Jiangsu Nantong 226007 China)

Abstract :In this paper , an application of wavelet transform for speech signal denoising is discussed. According to the properties of voiced segment and unvoiced segment in speech signal , we use different thresholds. And a better effect of denoising is derived with less distortion.

Keywords :wavelet transform ; speech signal denoising ; threshold