

# 小波变换在语音信号处理上的应用

唐玲<sup>1,2</sup>, 程建<sup>2</sup>

(1. 四川理工学院自动化与电子信息学院, 四川 自贡 643000; 2. 电子科技大学电子工程学院, 成都 610054)

**摘要:**利用小波的多分辨率分析, 以及其良好的空间域和频率域局部化特点, 针对语音信号特征, 选取适当的小波算法进行去噪和增强语音, 压缩编码, 提取语音信号特征等处理。通过 Matlab 仿真分析, 得到增强后的信号图和压缩后的压缩比参数、能量保留参数、零系数比例系数, 提取的信号特征。结果表明, 基于小波变换的语音信号处理表现出良好的特性。

**关键词:**语音信号处理; 小波变换; 去噪; 增强; 压缩编码; 特征提取

**中图分类号:** TB115

**文献标识码:** A

小波分析是近十几年发展起来的一种新的时频分析方法, 它是泛函数、Fourier 分析、调和分析、数值分析的最完美的结晶; 在应用领域, 特别是在信号处理、图像处理、语音处理以及众多非线性科学领域, 它被认为是继 Fourier 分析之后的又一有效的时频分析方法。小波变换是传统傅里叶变换的继承和发展。由于小波的多分辨率分析具有良好的空间域和频率域局部化特性, 对高频采用逐渐精细的时域和空域步长, 可以聚焦到分析对象的任意细节, 因此, 特别适合于信号非平稳信源的处理, 并已成为一种信息处理的新手段。目前, 小波分析已成功应用于语音信号处理。

## 1 小波理论

### 1.1 连续小波变换

设  $\Psi(t)$  为一平方可积函数, 即  $\Psi(t) \in L^2(R)$ , 若其傅里叶变换  $\Psi(\omega)$  满足

$$\text{条件: } C_{\Psi} = \int_R \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < \infty, \text{ 则称 } \Psi(t)$$

为一基本小波或小波母函数, 称上式为小波函数的容许条件。将任意  $L^2(R)$  空间中的函数  $f(t)$  在小波基下展开, 函数  $f(t)$  连续小波变换 (continue wavelet transform, 简称 CWT) 表达式为<sup>[1]</sup>:

$$\begin{aligned} WT_f(a, \tau) &= \langle f(t), \Psi_{a, \tau}(t) \rangle = \\ &= \frac{1}{\sqrt{a}} \int_R f(t) \Psi\left(\frac{t-\tau}{a}\right) dt \end{aligned} \quad (1)$$

其重构公式(逆变换)为:

$$f(t) = \frac{1}{C_{\Psi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{|a|} W_f(a, b) \Psi\left(\frac{t-b}{a}\right) da db \quad (2)$$

### 1.2 离散小波变换

为了减小小波变换系数冗余度, 将小波基函数  $\Psi_{a, \tau}(t) = \frac{1}{\sqrt{a}} \Psi\left(\frac{t-\tau}{a}\right)$  的  $a, \tau$  限定在一些离散点上的取值。在离散小波变换中, 待分析信号  $f(t)$  和分析小波  $\Psi_{a, k\tau_0}(t)$  中的时间变换  $t$  并没有被离散化, 只是离散  $a, \tau$  栅格下的小波变换。

#### 1.2.1 尺度离散化

对尺度进行幂数级离散化即取  $a = a_0^m (a_0 > 0, m \in Z)$ , 此时对应的小波函数是:

$$a_0^{\frac{j}{2}} \Psi[a_0^{-j}(t-\tau)] \quad j = 0, 1, 2, \dots \quad (3)$$

#### 1.2.2 位移离散化

对  $\tau$  进行均匀离散取值, 以覆盖整个时间轴。在尺度  $j$  下由于  $\Psi(a_0^{-j}t)$  的宽度是  $\Psi(t)$  的  $a_0^j$  倍, 采样间隔可以扩大  $a_0^j$ , 但信息也不会丢失。这样,  $\Psi_{a, \tau}(t)$  就改为

$$\Psi[a_0^{-j}(t - ka_0\tau_0)] = a_0^{\frac{j}{2}} \Psi[a_0^{-j}(t - k\tau_0)] \quad (4)$$

记为  $\Psi_{a_0, k\tau_0}(t)$ 。

离散小波变换定义为<sup>[1]</sup>:

$$WT_f(a_0^j, k\tau_0) = \int f(t) \Psi_{a_0^j, k\tau_0}(t) dt \quad (5)$$

$j = 0, 1, 2, \dots, k \in Z$

其重构公式为:

$$f(t) = c \sum_{-\infty}^{\infty} \sum_{-\infty}^{\infty} c_{j,k} \Psi_{j,k}(t) \tag{6}$$

其中,  $c$  是一个与信号无关的常数。

### 1.3 多分辨率分析

定义函数:

$$\varphi_{j,k}(t) = 2^{\frac{j}{2}} \varphi(2^{-j}t - k), j, k \in Z \tag{7}$$

则函数系  $\langle \varphi_{j,k}(t) |_{k \in Z} \rangle$  为规范正交。三层多分辨率分析树结构如图 1 所示。设  $V_j$  代表图中的低频部分  $A_j$ ;  $W_j$  代表分解中的高频部分  $D_j$ , 则  $W_j$  是  $V_j$  和  $V_{j+1}$  中的正交补, 即:  $V_j \oplus W_j = V_{j+1}, j \in Z$ 。显然:

$$V_j \oplus W_j \oplus W_{j+1} \oplus \dots \oplus W_{j+m} = V_{j+m} \tag{8}$$

则多分辨率的子空间  $V_0$  可以用有限子空间来逼近。即有

$$V_0 = V_1 \oplus W_1 = V_2 \oplus W_2 \oplus W_1 = \dots = V_N \oplus W_{N-1} \oplus \dots \oplus W_2 \oplus W_1 \tag{9}$$

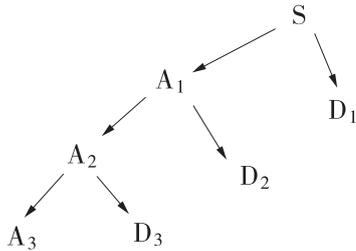


图 1 三层多分辨率分析树结构图

分解的最终目的是构造一个在频率上高度逼近  $L^2(R)$  空间的正交小波基, 频率分辨率不同正交小波基相当于带宽各异的带通滤波器。从图 1 看出, 多分辨率只对低频空间进行进一步分解, 频率分辨率变得越来越高。

## 2 小波变换在语音增强中的应用

在语音的产生、处理和传输过程中, 不可避免地会受到环境噪声的干扰, 使得语音信号处理系统, 如语音编码和语音识别系统的性能大大降低。为了改善语音质量, 提高语音的可懂度, 人们根据语音和噪声的特点, 采取各种语音增强方法抑制背景噪声。小波去噪的基本原理首先将经过预处理的含噪信号进行多尺度小波分解, 然后在各尺度下尽可能提取出信号的小波系数而去除属于噪声的小波系数, 最后用逆小波变换重构信号, 达到去噪的目的。小波去噪的方法有很多种, 概括起来主要有屏蔽去噪法、阈值去噪法、模板极大值检测去噪法等。下面简单介绍阈值去噪法<sup>[2]</sup>。它的思想是利用阈值对经过变换后的较低尺度上的小波系数进行处理, 将那些对应于高斯噪声的小波系数置零, 而保留那些对应于信号的小波系数, 再由处理后的小波系数重构原信号, 获得对有效信号的最优估计。阈值化处理的关键问题是选择合适的阈值。如果阈值太小, 去噪后的信号仍然有噪声存在; 如果阈值太大, 重要信号特征

将被滤掉, 引起偏差。从直观上, 对于给定小波系数, 噪声越大, 阈值就越大。大多数阈值选择过程是根据本组小波系数的统计特性, 计算出一个阈值。本文后边的 matlab 中采用 db3 母小波函数对语音波形进行 5 层分解, 并选用全阈值进行信号去噪增强处理。

## 3 小波变换在语音压缩编码中的应用

语音信号的压缩编码是研究如何降低语音信号编码速率的问题。数码率低于 64Kbit/s 的语音编码方法称为语音压缩编码技术。语音的压缩编码方法归纳起来可以分为三大类: 波形编码、参数编码和混合编码。波形编码比较简单, 失真最小, 方法简单, 但数码率比较高。参数编码的编码速率可以很低, 但音质较差, 其次是复杂度高。混合编码吸收了波形编码和参数编码的优点, 从而在较低的比特率上获得较高的语音质量。

小波变换编码属于变换域编码方式。变换域编码<sup>[3]</sup>是将输入信号用时间窗分成短区间数据块, 然后进行正交变换, 并对变换系数逐个地进行量化并传送的方法。它将时域信号变换到时频域中小波分析结合矢量量化编码, 可进一步地压缩编码速率, 图 2 为典型编码框图。

这里符号  $2 \downarrow$  表示小波分析的分解, 原信号经作用得到平滑版本和细节版本, 平滑版本再经分解, 直至在整个变换域中得到全部重叠正交变换。矢量量化可定义为  $K$  维欧氏空间到有限子集的映射, 编码中将输入矢量映射成码矢量, 为降低编码的复杂度常采用二叉树的结构。

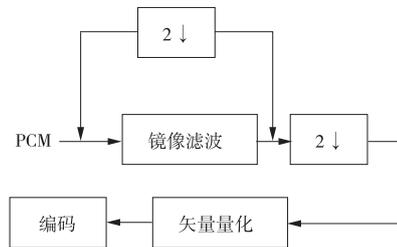


图 2 典型编码框图

目前较为完善的小波变换宽带音频编码方案是采用离散小波包变换 (Discrete Wavelet Packet Transform)。其基本特点是, 用离散小波包取代传统的滤波器组或正交变换基, 由心理声学模型计算掩蔽域门限, 直接控制自适应小波基的优化和比特分配。基于小波变换的宽带音频编码器在 64 kb/s 的速率可以达到透明质量, 从技术前景来看, 它还可在 10 kb/s 速率上合成音乐信号。但现行的小波变换的音频压缩编码方案还存在着计算复杂度高和时延长的缺陷一些优秀的音频特性, 如高保真度、动态音频范围及丰富的音质都需要高采样率的保证。同时, 在许多互联网和多媒体系统应用中, 由于带

宽限制,使这种高采样率技术不适用<sup>[4]</sup>。现利用小波变换的方法来降低数据的传输比特率<sup>[5]</sup>。其处理过程主要有以下四步:(1)小波的选择。设计高质量的语音编码器时,母小波的选择非常重要,其目标是 minimized 重构残差的方差以及最大化信噪比。多次试验表明,针对语音信号特点,选用小波母函数 db5 效果较好。所以,本次仿真试验选用此函数来实现。(2)小波分解。基于小波压缩的思想,用少量的逼近系数以及某些细节参数准确表示原始语音。分解级数的决定取决于语音信号的类型和其它一些参数。本实验中,对语音信号进行了 5 层分解,压缩参数取  $1.5^{[6-8]}$ 。(3)系数的裁剪。对原始语音分解后,压缩过程可以通过裁减小于某一阈值小波系数来实现。这是依据人耳的生理特性和人对声音的心理特征去除人耳朵不能感知的声音成分的原理,进行压缩裁减。使用 Birge - Massart 策略进行语音波形的一维小波阈值计算,即用函数 wdcbm 进行硬阈值压缩信号。此函数可返回尺度相关阈值和系数的个数,尺度相关阈值使用基于 Birge - Massart 策略的小波系数选择规则得到。(4)系数编码。最后进行剩下的小波系数编码,有利于传输。其中,编码压缩后原语音信号的长度比编码后的长。

这种压缩方法是一种有损压缩,但可用最少的数据量来表示音频信号,保持良好的保真度。

#### 4 基于小波变换的语音信号的特征提取

小波分析在信号特征的提取上有很多应用,以语音信号基频的提取为例说明它的应用。语音的能量来源于正常呼吸时肺部呼出的稳定气流,而通过声带的开启和闭合使气流形成一系列的脉冲,每开启和闭合一次的时间称为基音周期,其倒数称为基音频率,简称基频 ( $F_0$ )。考虑到基音频率多分布在 50 ~ 500 Hz 之间,而多分辨率分析有把频率逐级对分的特点,可用 Mallat 算法<sup>[9]</sup>直接对语音信号进行分解,用低频小波系数对信号低频部分重构,然后用小波变换进行信号奇异点检测,通过奇异点确定得到基音周期。通过求倒数得到基音频率,再计算出  $F_0$  平均值、 $F_0$  最大值和  $F_0$  变化范围 (分别记为  $F_0$ 、 $F_{0_{max}}$  和  $F_{0_{range}}$ )。

对于基音周期的提取,采用了小波变换奇异点检测算法。选取具有紧支集正则的二次样条小波  $U(x)$ 。则二次样条小波的相应滤波器传递函数为:

$$H(X) = e^{jX^2} [\cos(X/2)]^3 \quad (10)$$

$$G(X) = 4je^{jX^2} \sin(X/2) \quad (11)$$

在每一尺度  $s = 2^j$  上,信号分解为低带半带  $S_2^{j-1}f$  和高带半带  $W_2^{j-1}f$ :

$$W_2^{j-1}f = S_2^{j-1}f * G(n) \quad (12)$$

$$S_2^{j-1}f = S_2^{j-1}f * G(n) \quad (13)$$

当  $j=0$  时,  $S_2^j f = S_2 f$ , 即  $f(x)$  的采样值,式 (12) 和式 (13) 即为小波变换。计算  $W_j f(x)$ , 并求大于  $0.8W_{smax}$  的极值点。在同一尺度下,相邻极值点的时间间隔就是基音周期  $T$ , 然后求基音周期  $T$  的倒数就是基音频率  $F_0$ 。

#### 5 实验结果与总结

原始语音信号波形如图 3 所示。其语音内容为 A 到 G 七个英文字母的发音。可以看见语音信号中含有一定的噪声,为使效果更加明显,首先对原始语音追加噪声<sup>[10-11]</sup>, 然后对其利用小波变换进行分解,最后重构。由于本文所采用的信号较弱,所以没有做加噪处理。在 Matlab 中使用小波变换的图形方式 (GUI), 而得到如图 4 所示的增强语音波形信号。从图中明显看出,增强后的语音信号很光滑,提高原语音信号的信噪比。

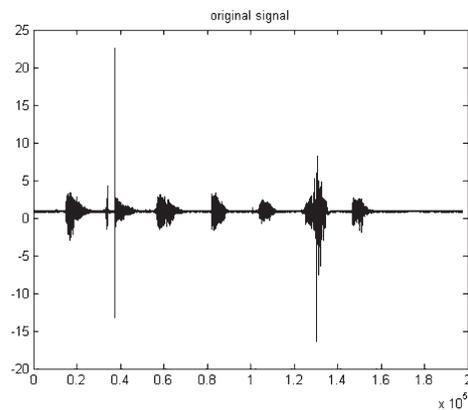


图 3 原始语音信号

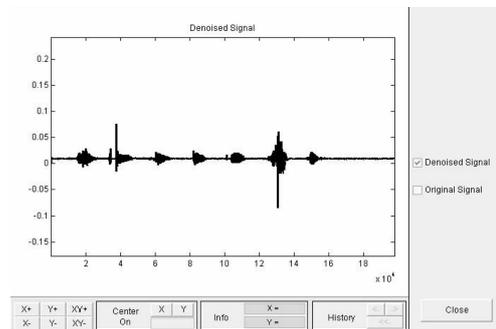


图 4 去噪增强后语音信号

按照本文上述的语音压缩算法,将原始语音信号波形 (见图 3) 进行压缩编码,得到如下结果。零系数的比例:  $pz = 96.87$ ; 保留能量大小:  $saveenergy = 88.48$ 。Matlab 实现如图 5 所示。

按照上文关于语音信号基频提取的分析用 Matlab 编写程序,运行后调用 matlab 结果得  $F_0 = 124.2$  Hz,  $F_{0_{max}} = 143.7$  Hz,  $F_{0_{range}} = 79.7$  Hz。

通过在 Matlab 中仿真,可以看到一维小波变换可以

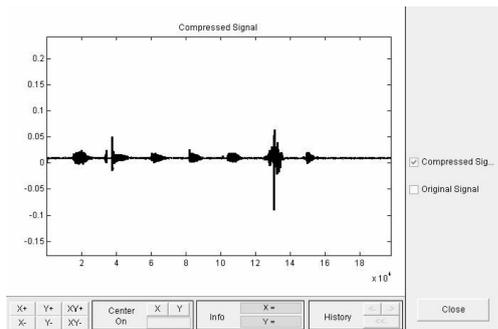


图5 压缩后的语音信号

对语音信号进行去噪、压缩,特征提取等处理,并表现出良好的参数特性。在增强语音信号中,选用小波母函数 db3 进行分解,对于语音信号而言,5 层分解已足矣,采用全局阈值进行信号增强处理,主要用 wavedec 函数来实现。在压缩语音编码时,选用 db5 小波函数进行 5 层分解,但要通过函数 wdcbm 计算出阈值,函数 wdcbm 是以 Birge - Massart 策略设置一维小波压缩阈值。而后,用硬阈值法压缩语音信号。这种处理保存了低频信号,而且灵活性很大<sup>[12]</sup>。

#### 参考文献:

- [1] 胡昌华,李国华,刘涛,等.基于 matlab 6.x 的系统分析与设计——小波分析[M].西安:西安电子科技大学出版社,2004.
- [2] 赵岩,王宁,张宝泉.基于小波变换的人声语音增强技术研究[J].齐齐哈尔大学学报,2008,24(1):44-47.
- [3] 景新幸.小波变换及其在宽带音频压缩编码中的应用[J].电声技术,1999(6):7-8.
- [4] 杨行峻,迟惠生.语音信号处理[M].北京:电子工业出版社,1995.
- [5] Panaga Adriane,Zaro Mil Ton A,Schuck JR,et al.Quan-

titative assessment of the use of continuous wavelet transform in the analysis of the fundamental frequency disturbance of the synthetic voice[J].Medical Engineering and Physics, September-October, 2002, 16 (2): 553-559.

- [6] Obaigdat M S, Brodzik Andy, Sadoun.A performance evaluation study of four wavelet algorithms for the pitch period estimation of speech signals[J].Information Sciences,1998,21(1):213-221.
- [7] Singh R,Davis K,Rao P V S,et al.Model-based speech recognition with intermediate wavelet transform domains [J].Computer Speech & Language,1997,20(3):253-273.
- [8] Namba Munehiro,Iahida Yoshihisa.Wavelet transform domain blind deconvolution[J].Signal Processing,1998, 24(2):119-124.
- [9] 陈峰.Blackfin 系列 DSP 原理与系统设计[M].北京:电子工业出版社,2004.
- [10] Lu Chingta,Wang Hsiao-Chuan.Enhancement of single channel speech based on masking property and wavelet transform[J].Speech Communication,2003,39(1):409-427.
- [11] Irino Toshio Patterson,Roy D.Segregating information about the size and shape of the vocal tract using a time-omain auditory model: The stabilised wavelet-Mellin transform[J]. Speech Communication March, 2002,29(4):181-203.
- [12] Pinte R,Istva N.Perceptual wavelet-representation of speech signals and its application to speech enhancement[J].Computer Speech& Language,1996,13(3):1-22.

## Wavelet Transformation Application in Speech Signal Processing

TANG Ling<sup>1,2</sup>, CHENG Jian<sup>2</sup>

- (1. School of Automation and Electronic Information, Sichuan University of Science & Engineering, Zigong 643000, China;  
2. School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu 610054, China)

**Abstract:** By the time-frequency analytic feature of wavelet transformation, the appropriate wavelet functions are selected to strengthen, to code, to compress and to extract signal features of speech according to the characteristics of the speech signals. Simulated by Matlab, the strengthened signals, the compression ratio parameter, the energy reservation parameter, the zero coefficient parameter and the speech signal features were obtained. The results show that the speech signal processing based on wavelet transformation exhibits good characteristics.

**Key words:** speech signal processing; wavelet transformation; strengthening; de-noising; feature extraction; condensation encode