

用于信号特征提取和重建的脊提取算法¹

张晓冬 王 桥 吴乐南

(东南大学无线电工程系 南京 210096)

摘 要 该文论述了以信号在连续小波变换域上形成的脊作为信号特征进行信号的特征提取和信号重建的方法,并且将该理论应用到语音信号的处理中。仿真实验表明,对于最高频率为 4kHz 的语音信号来说,提取 3~5 条脊即可以很好地刻划信号的全部信息,重建的信号在主观感觉上达到了较好的效果。

关键词 连续小波变换,脊提取,信号重建

中图分类号 TN911.72

1 引 言

信号的特征分析与提取是信号处理中一个比较重要的方面,应用的范围很广,如雷达、声呐等系统对目标的检测、识别,语音信号处理中的语音识别等,都需要准确提取信号的特征。如果提取的特征能够完整地反映信号的全部信息,并且能够将信号重建,那么在传输的时候可以通过传输提取到的信号特征来代替对信号本身的传输,从而达到信号压缩的目的。

本文中应用的算法最初来源于对信号连续小波变换相位信息的研究。在信号与选取的小波满足一定条件的前提下,信号的连续小波变换可以用一个近似公式求出。信号连续小波变换平面上,呈现出一些局部的峰值点,类似于地形图上山脊的形态。进一步的分析表明,对应于“脊”所在的位置上的信息已经可以反映信号的全部特征,并且可以完全重构信号。

这样的结果是可以理解的,连续小波变换本身是将信号在冗余度很大的一组基上进行投影,可以有多种选择方式来选取其中的一部分基来完全重建信号。因为每个信号在连续小波平面上所呈现的脊的位置均是和信号相关的,所以提取信号对应的脊可以看作是对每个信号分别选取一组与信号相关的基来进行表示。脊的概念也同样适用于其它的信号线性表示,如匹配追踪算法^[1]等。

本文将此算法应用于语音信号的特征提取和信号重建。实验结果表明,语音信号经过连续小波变换后,脊的特征是很明显的,而且对于最高频率为 4kHz 的语音信号来说,往往只需要提取 3~5 条脊就可以将信号的全部特征提取完全,重建后语音的主观感觉良好。

2 脊重建信号的原理

设待处理的信号为渐近信号的形式,即

$$f(x) = A_f(x) \cos[\varphi_f(x)] \quad (1)$$

这里对信号的要求可以解释为,信号的振幅变化率相对于信号频率的变化来说是缓慢变化的,信号从整体上看是一个调频信号。满足(1)式的 A_f, φ_f 并不唯一,这里特指满足解析信号概念下的一组幅度、频率参数,即定义信号 $f(x)$ 对应的解析信号 $Z_f(x) = f(x) + if^H(x)$ 。式中 Z_f 为信号 $f(x)$ 对应的 Hilbert 变换, $f^H(x) = \frac{1}{\pi} P \cdot \int \frac{f(y)}{x-y} dy$, 则对(1)式为渐近信号的要求,意味着信号需要同时满足如下条件: $A_f(x) = |Z_f(x)|$, $\varphi_f(x) = \arg Z_f(x) = \arctan[f^H(x)/f(x)]$ 。

¹ 2002-01-17 收到, 2002-06-28 改回

选择满足容许条件的渐进小波函数 $g(x) = A_g(x) \exp[i\varphi_g(x)]$, 则信号 $f(x)$ 的连续小波变换可表示为

$$T_f(b, a) = \langle f, g_{(b,a)} \rangle = \frac{1}{a} \int_{-\infty}^{\infty} f(x) \overline{g\left(\frac{x-b}{a}\right)} dx \quad (2)$$

(2) 式中被积函数的相位记作: $\Phi_{(b,a)}(x) = \varphi_f(x) - \varphi_g[(x-b)/a]$. 此外, 当小波函数同时满足 $\xi < 0$ 时, $\hat{g}(\xi) = 0$, 则有下式:

$$T_f(b, a) = \langle f, g_{(b,a)} \rangle = (1/2) \langle Z_f, g_{(b,a)} \rangle$$

成立. 应用稳定定位相法的近似公式, 可以得到信号连续小波变换的另一种表示形式^[2]

$$T_f(b, a) = \frac{Z_f(t_s) \overline{g[(t_s - b)/a]}}{\text{Corr}(b, a)} + O(\varepsilon) \quad (3)$$

其中

$$\text{Corr}(b, a) = \sqrt{2/\pi} \sqrt{a^2 |\Phi''_{(b,a)}(t_s)|} \exp[(-i\pi/4) \text{sgn}(\Phi''_{(b,a)}(t_s))] \quad (4)$$

定义相位导数为零的位置为脊所在的位置, (3), (4) 式中的 t_s 为相位导数取到零的时候 x 的取值. 在连续小波变换取模后的平面上, 脊所在的位置对应着局部极大值所在的点. 由 (3) 式可以得到, 限制在脊上的连续小波变换为

$$T_f(x, a_r(x)) \approx Z_f(x) \overline{g(0)} / \text{Corr}(x, a_r(x))$$

进一步的分析表明, 限制在脊上的小波变换反映了所分析的解析信号的性质, 而脊以外位置上的数值只是与小波本身有关的量, 与信号无关. 脊上的小波变换数值可以用来完全重建信号. 重建公式为

$$Z_f(x) \approx T_f(x, a_r(x)) \cdot \text{Corr}(x, a_r(x)) / \overline{g(0)}$$

3 脊的提取算法

有效地提取脊所在的位置是本文的算法在应用时需要解决的首要问题.

从脊的定义可以看出, 直观的方法可以通过信号连续小波变换的相位来确定脊所在的位置. 但是在信号有噪声的情况下, 相位信息所受的影响很大, 由此计算出来的脊的位置会出现很大的偏差. 另一种方法是在连续小波变换后的平面上找模极大值对应的位置. 本文选取后一种方法.

对于信号只含有一条脊的情况, 已经提出很多的算法, 如直接搜索方法、Snake 方法等^[3]. 这几种提取脊的算法核心的想法都是: 假定脊在连续小波变换平面上为一条连续缓慢变化的曲线. 选择一个惩罚函数, 然后在整个连续小波变换平面上搜索一条使得惩罚函数取得最小数值的曲线, 作为信号对应的脊. 考虑到信号可能含有很强的噪声, 搜索得到的曲线可能停留在某个局部极大值的位置上不动, 为了得到惩罚函数全局的极小值, 还可以在搜索时采用模拟退火算法中的一些技巧, 从而使脊提取算法对噪声的稳定性得到增强.

上述使用惩罚函数的方法只适用于有单一脊的情况, 当信号中含有多条脊的时候, 则完全不再适用. 而常见的如语音信号等, 都含有多条脊的成份, 在提取脊的时候, 需要采用相应的同时提取多条脊的算法^[4].

本文采用 Crazy Climber 算法同时提取多条脊. 该算法的核心想法是, 在要提取脊的平面上均匀放置若干个点 (Climber), 每个点都按同样的规则进行移动. 选择合适的规则, 使得每

个点的运动方向趋向于附近区域内的模极大值点。在移动了适当的次数以后，对点的分布情况进行测量。平面上聚集点较多的位置连接成的几条曲线认为是信号的脊对应的位置所在。具体的点初始分布和运动规则如下：在连续小波变换平面上，设沿小波平移参数 b 和尺度变化参数 a 构成一个 $M = B \cdot A$ 的网格。在网格的格点上均匀放置 N 个点 (Climber)，并且设每个点的质量为 $1/N$ 。在沿着小波平移参数 b 的方向上，各点每一次各按 0.5 的概率向左右随机地移动一个网格点；在沿着小波尺度参数 a 的方向上，各点每一次选择按 0.5 的概率向上或向下找到一个可能的移动点，并且计算该点对应的数值与当前所在点的数值之差 Δm 。如果 $\Delta m \geq 0$ ，则运动到该点；如果 $\Delta m < 0$ ，则按概率 P 运动到该点，按概率 $1 - P$ 仍停留在当前点。平面上的 N 个点均按照同样的规则移动一次，完成一次循环。这之后，调整 P 的大小，进入下一次循环，直到循环到预先设定的次数或 P 值达到预先设定的数值为止。循环结束后，计算平面上各点的质量分布。质量大于某个设定的门限的点，认为是对应于脊所在的位置。理想情况下，大多数点应该聚集在几条曲线的附近，从而得到对脊的初步估算。

提取后的结果仍然是大致分布在几条曲线上的一些点，还需要将这些点连接成对应的曲线。在连接之前，需要对脊的特征进行一些设定。首先假设脊本身是比较平滑的，没有很大的突变，其次，假设脊延伸的方向是沿着水平方向，也就是小波平移参数方向变化的。将 Crazy Climber 算法的结果按此规则连接成几条曲线，作为提取到的信号的多条脊所在的位置。

由脊的提取算法可以看到，在决定每个点的运动规则时，用到了模拟退火算法中的思想，主要目的在于当搜索平面上存在由于噪声等因素产生的局部极大值点时，各个点 (Climber) 能够以一定的概率脱离该点，从而得到惩罚函数的全局极小值，增强了噪声条件下算法的稳定性。

4 信号重建的算法

在提取了信号的脊以后，考虑如何利用脊上的信息，对信号进行有效的重构。考虑到连续小波变换是线性变换，可以分别对每一条脊所对应的信号进行重构，然后将各条脊重构的信号叠加起来，得到完整的重构信号。这样做可以简化运算，将问题转化成为如何在单一脊的情况下重构信号。

从已经提取的某一条脊上选择若干个观察点，作为重构该条脊对应的信号时要用到的全部信息，即

$$z_j = T_f(b_j, a_j), \quad j = 1, 2, \dots, n \quad (5)$$

(5) 式作为信号重构的限制条件，即要求重构得到的信号在做连续小波变换以后，在相应的点上得到的数值与限制条件相符合。光滑样条方法可以用来求解满足一定的限制条件的重构信号^[5]。

信号的重构问题可以归纳为求解一个最优化问题，即

$$\hat{f} = \arg \min_f \left[\frac{1}{n} \left| \Sigma^{-1/2} [Z - T_f(\cdot, \cdot)] \right|^2 + \mu \langle Qf, f \rangle \right] \quad (6)$$

其中 $\langle Qf, f \rangle$ 为待求函数 $f(x)$ 的二次方形式， Z 表示所取 n 个观察样点上的数值。这个优化问题的解可以表示为： $\hat{f}(x) = \sum_{j=1}^n \lambda_j \tilde{\varphi}_j(x)$ 。其中的对偶小波定义为： $\tilde{\varphi}_j = Q^{-1} \varphi_j$ 。对偶小波前的参数可以用拉格朗日乘数法求解，表示为： $\Lambda = A^{-1} \Sigma^{-1/2} Z$ 。其中 Λ 为参数构成的列矩阵，式中 $A = n\mu I + \tilde{\Sigma}$ ， $\tilde{\Sigma}$ 可以表示为： $\tilde{\Sigma}_{j,k} = \langle \tilde{\varphi}_j, \tilde{\varphi}_k \rangle$ 。

考虑最简单的一种情况，即惩罚函数只限定重构后的信号符合限制条件即可，而不附加其它的条件，同时忽略量化等步骤中引入的噪声等因素，可以得到重构信号的最简单的一种形式，即

$$\hat{f}(x) = \sum_{j=1}^n \lambda_j \varphi_j(x)$$

在这种情况下, 重构信号所用的小波与对信号分析所用的小波相同, 重构时各小波对应的系数就是在观察点的数值。将所有观察点上对应的小波叠加即得到重构后的信号。

$$A = \bar{\Sigma}^{-1} Z = Z$$

这种重构信号的算法因为只保证了重构信号与原信号在观察点上的一致性, 而没有能对脊本身的平滑性做适当的约束, 所以在信号信噪比较小的情况下, 信号重构的效果不是很理想。进一步的改进可以根据具体情况下所处理的信号类型, 在 (6) 式的目标函数中增加对脊的平滑性等限制条件, 相应的信号重构算法变得略复杂些。

5 语音信号的脊提取与信号重构

本文将前述信号的脊提取、重构算法应用于语音信号处理。语音信号在进行连续小波变换后, 在平移参数 b 和尺度变化参数 a 形成的平面上, 脊的特征还是非常明显的, 对于常用的通信线路上传输的最高频率为 4kHz 的语音信号, 可以提取到 3~5 条脊。针对语音信号的特点, 还需要对算法做一些相应的调整, 才能达到比较理想的结果。

语音信号的一个特点是, 信号各频率成份的动态范围较大, 高频部分的信号能量和低频部分相比, 有很明显的衰减。表现在连续小波变换的结果上, 随着尺度参数的减小, 计算得到的结果也明显减小, 使得在该平面上, 信号脊的特点变得不明显。

针对这一特点, 可以有两种方法进行补偿。首先, 可以选择在信号进行连续小波变换以前, 对信号进行 6dB/ 倍频程的预加重, 从而对信号的高频成份进行补偿。其次, 可以先对信号进行连续小波变换, 然后对得到的结果进行相应的加权处理, 使得沿尺度参数变化方向上各条脊在幅度上均比较突出, 从而有利于对脊所在位置的提取。

语音信号的另一个特点是, 在频域上有明显的共振峰结构。对于长途语音信号, 前 3 个或 4 个共振峰结构就已经能够很完整地刻划信号所含的全部信息。对应于脊提取算法提取到的脊, 在第 1 共振峰内的每 1 个峰值一般对应着 1 条脊, 而频率较高的第 2、3 共振峰, 基本上 1 条脊所提取到的信息就对应着同一共振峰内的多个峰值。

根据多次实验观察的结果, 对语音信号脊的特征做以下的假设:

(1) 信号的脊是沿小波平移参数 b 的方向延伸的; (2) 如果信号含有多条脊, 则脊是互不交叉, 彼此大体平行分布的; (3) 脊是连续平缓变化的。

在上述假设的前提下, 对长途语音信号进行脊的提取。结果表明, 一般来说, 3~5 条脊已经可以完整地描述信号的全部特征。图 1 为原始的语音信号, 图 2 为信号做连续小波变换后, 在连续小波域的等高线图。从图中可以看到, 脊的特征是很明显的。图 3 为提取到的脊, 图 4 为根据脊上的信息重建的信号。仿真实验中选择的小波函数为 Morlet 小波, 尺度取值为 $a = 2 \cdot 2^{n/10}$, $n = 1, 2, \dots, 50$ 。重建信号与原始信号相比, 基本上保留了全部的信息, 而且从主观感受上, 语音信号恢复的效果也比较理想。仍然存在的问题是, 对于图 1 所示的浊音, 脊的特征比较明显, 可以得到较好的特征提取与信号恢复效果, 但是对于清音, 在连续小波变换后的平面上, 脊的特征不是很明显, 还需要对算法进行改进。

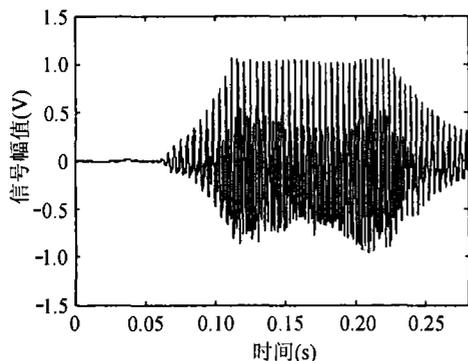


图1 原始信号时域波形

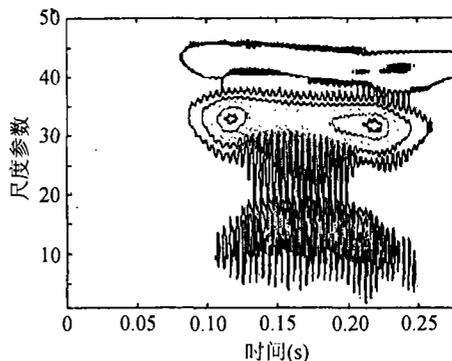


图2 连续小波变换后的等高线图

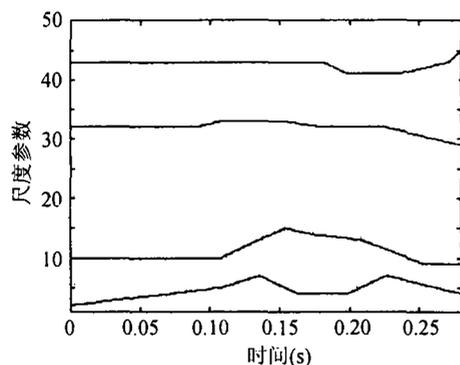


图3 提取的脊所在位置

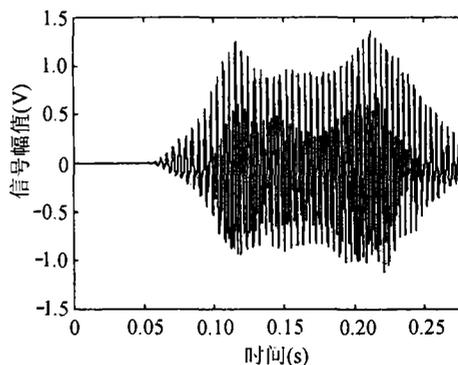


图4 重建信号时域波形

6 总结

信号在进行连续小波变换后,在连续小波变换域上的结果呈现出明显的“山脊”特征。通过脊所在位置上的信号的信息来提取信号的特征,并进行信号重建是近年来出现的较新的信号分析方法。本文将这一理论应用到对长途语音信号的分析 and 处理上,针对语音信号多条脊的分布特点,对脊的提取算法进行了研究。实际的处理结果表明,信号的特征可以被很好地提取出来,重建的信号在主观感受上达到了较好的效果。

参 考 文 献

- [1] S. Mallat, S. Zhang, Matching pursuit with time-frequency dictionaries, *IEEE Trans. on Signal Processing*, 1993, SP-41(12), 3397-3415.
- [2] N. Delprat, *et al.*, Asymptotic wavelet and Gabor analysis: Extraction of instantaneous frequencies, *IEEE Trans. on Info. Theory*, 1992, IT-38(2), 644-664.
- [3] R. Carmona, W. L. Hwang, B. Torresani, Characterization of signals by the ridges of their wavelet transforms, *IEEE Trans. on Signal Processing*, 1997, SP-45(10), 2586-2590.
- [4] R. Carmona, W. L. Hwang, B. Torresani, Multiridge detection and time-frequency reconstruction, *IEEE Trans. on Signal Processing*, 1999, SP-47(2), 480-492.

- [5] R. Carmona, W. L. Hwang, B. Torresani, Practical Time-Frequency Analysis, San Diego, Academic Press, 1998, Chapter 8.

CHARACTERIZATION EXTRACTION AND
RECONSTRUCTION OF SIGNALS BY
THE RIDGES OF CONTINUOUS WAVELET TRANSFORM

Zhang Xiaodong Wang Qiao Wu Lenan

(Dept. of Radio Engineering, Southeast University, Nanjing 210096, China)

Abstract The theory of detecting ridges in the modulus of the continuous wavelet transform is presented as well as reconstructing signal by using information on ridges. Application of the theory on speech signal is discussed in detail. Experimental result shows that as for signals with highest frequency at 4kHz, 3~5 ridges is enough to describe all of the information contained in the signal. The reconstructed signal reserves most of the necessary information of the original signal.

Key words Continuous wavelet transform, Ridge extraction, Signal reconstruction

张晓冬: 男, 1974 年生, 博士生, 研究方向为信号分析和信号表示.

王 桥: 男, 1966 年生, 博士, 教授, 《IEEE 信息论会刊》、《Zentralblatt Math.》等杂志评审员, 现主要从事微局部分析及其在信号分析与多媒体通信中的应用研究.

吴乐南: 男, 1952 年生, 博士, 教授, 现主要从事多媒体信号处理, 数字广播技术等方面的研究.