

二维 CWT 方法在时间序列分析中的应用研究

贾瑞玉, 王 亮, 王会颖

(安徽大学 计算机科学与技术学院, 安徽 合肥 230039)

摘 要:利用二维连续小波变换(CWT),结合时间序列分析的特点,对时间序列数据进行分析。从对多频率信号、突变信号以及噪声数据的容错性三个方面的分析来对利用二维连续小波变换方法在时间序列分析中的应用进行了研究,主要分析方式以趋势分析为主,最后通过对实验结果的分析得到良好的效果。

关键词:连续小波变换;时间序列;趋势分析

中图分类号:TP301.6

文献标识码:A

文章编号:1673-629X(2007)05-0160-03

Research for Applications of 2D-CWT Approach in Time Series Analysis

JIA Rui-yu, WANG Liang, WANG Hui-ying

(School of Computer Science and Technology, Anhui University, Hefei 230039, China)

Abstract: Based on the characteristic of time series analysis, data of time series was analyzed by the approach of 2D-CWT. The analysis of three aspects such as multi-frequency signals, jumping signals and fault tolerance of noisy data were made for the purpose of the research for applications of 2D-CWT in time series analysis. Trend analysis is major means. Finally, the experimentation result was discussed.

Key words: continuous wavelet transform; time series; trend analysis

0 引言

时间序列数据是按照时间顺序取得的一系列观测值^[1],一般相邻数值间的时间间隔是相等的。时间序列的一个本质特征就是相邻观测值的依赖性,对时间序列数据的数据挖掘所讨论的就是对这种依赖性的分析,通常这要求生成一系列随机动态模型^[2,3]。目前已有诸如:自回归综合移动平均模型(ARIMA)、混合自回归移动平均模型(ARMA)、卡尔曼(Kalman)过滤等,在多方面得到广泛的应用。但是由于时间序列数据在取得时由于客观条件影响往往存在扰动现象,传统方法难以奏效,同时传统方法通常只对频率域或时间域单独分析,存在分析不全面的问题。文中所采用的小波分析中的二维连续小波变换的方法具有多分辨率特性,同时在时域和频域都具有表征信号局部特征的能力,能够很好地处理非线性性和扰动的时间序列问题^[4]。

文中对采用小波分析中二维连续小波变换的方法做了一些探索,并通过一系列实验讨论得到了良好的

结果,为以后相关领域的研究提供了有益的参考。

1 连续小波变换

设 $f(x)$ 为平方可积函数(即 $f(x) \in L^2(R)$), $\Psi(t)$ 为小波母函数,且满足条件:

$$\int_{\mathbb{R}} \frac{|\Psi(\omega)|^2}{\omega} d\omega < \infty$$
, 则称将 $f(x)$ 在 $\Psi(t)$ 下的展开为函数 $f(x)$ 的连续小波变换(CWT),其表达式为^[5]:

$$WT_f(a, \tau) = \langle f(t), \Psi_{a,\tau}(t) \rangle = \frac{1}{\sqrt{a}} \int_{\mathbb{R}} f(t) \overline{\Psi\left(\frac{t-\tau}{a}\right)} dt \quad (1)$$

式(1)中 $WT_f(a, \tau)$ 为小波变换系数, a 为尺度因子, τ 为平移因子。函数 $f(x)$ 在尺度 a 、平移点 τ 上的小波变换系数实质上表示的是在 τ 处,时间段 $a\Delta t$ 上包含在中心频率为 $\frac{\omega_0}{a}$ 、带宽为 $\frac{\Delta\omega}{a}$ 频窗内的频率分量大小,

随着尺度 a 的变化,对应窗口的 $\frac{\omega_0}{a}$ 和 $\frac{\Delta\omega}{a}$ 也发生变化,因此与快速傅里叶变换不同,小波变换是一种多分辨率的时频联合分析方法。

与傅里叶基不同的是,尺度和位移均连续变化的连续小波基函数形成了一组非正交的过渡完全基,也即其任意函数的小波展开系数之间有一个相关关系。设 $K_{\Psi}(a, \tau; a', \tau')$ 为两个基函数 $\Psi(a, \tau)$ 和 $\Psi(a', \tau')$,

收稿日期:2006-07-14

基金项目:安徽省教育厅自然科学基金资助项目(2005kj056)

作者简介:贾瑞玉(1965-),女,河南浙川人,副教授,硕士生导师,研究方向为可视化信息处理、智能计算。

τ' 的冗余量(相关度),则

$$K_{\Psi}(a, \tau; a', \tau') = C_{\Psi}^{-1} \int_{\mathbb{R}} \Psi_{a, \tau}(t) \cdot \Psi_{a', \tau'}(t) dt \quad (2)$$

式(2)中 K_{Ψ} 表示连续尺度、时移平面的两个不同点之间的 CWT 系数的相关关系,也称再生核或重建核。

2 基于二维连续小波变换的时间序列分析

文中采用小波分析中的二维连续小波变换对时间序列数据进行分析,主要在以下三个方面进行了分析和讨论^[6]:①对多频率信号的分析;②对突变信号的分析;③对噪声数据的容错性分析。为了简化编程的复杂度,采用了 MathWorks® 公司的 MATLAB7.0 平台进行实验。

2.1 对多频率信号的分析

为了分析的需要,自定了下面的分段函数,下面的样本数据由该函数产生,其表达式为:

$$s(t) = \begin{cases} \sin(0.02t), & 1 \leq t \leq 400 \\ \sin(0.1t), & 401 \leq t \leq 800 \end{cases}$$

该函数的时频图见图 1。

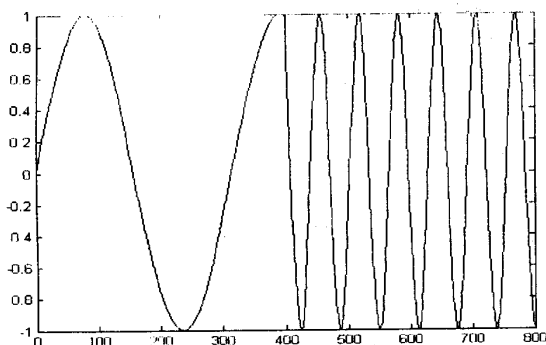


图 1 样本数据的时频图

将该数据经过二维连续小波变换后得到了一系列小波变换系数,这里使用的是 db2 小波,图 2 和图 3 分别为变换后的二维和三维小波系数图。

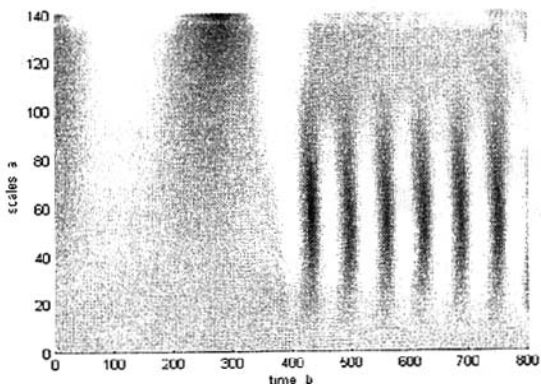


图 2 样本数据经二维连续小波变换后的小波系数图(2D)

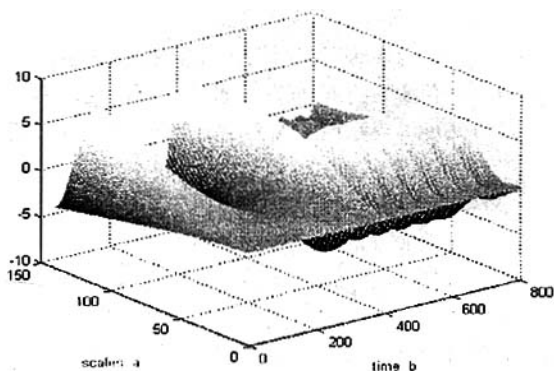


图 3 样本数据经二维连续小波变换后的小波系数图(3D)

其中在图 2 中的高亮区域表示小波系数值较大,低亮度区域表示小波系数值较小。图 3 与图 2 的本质是一样的,只是图 3 旋转了 90 度,这样只是为了能更加清楚地看到实验结果。从两个图中可以清楚地看出,当尺度因子由小变大时对变换后小波系数的影响,以及原时间信号在不同时域内的演化特征。

2.2 对突变信号的分析

在时间序列数据中常会出现由于外部条件或其他原因引起的观测值的突变,而及时准确地发现这些突变信息在对时间序列的数据挖掘中是十分有意义的。一般传统方法并不能准确及时地得到突变信息,造成信息的丢失,而小波变换的方法则可以较好地发现突变信息,从而及时适时调整模型,提高了预测的精度。

笔者将样本数据在不同尺度下进行变换,并将不同尺度下得到的结果进行比较,这里使用的样本数据为 $s(t) = \begin{cases} \sin(0.03t), & 1 \leq t \leq 500 \\ \sin(0.3t), & 501 \leq t \leq 1000 \end{cases}$ 分段函数产生的数据,尺度分别为: $a = 1, a = 20, a = 50, a = 100$ 。如图 4 所示,分别表示了尺度因子 a 在上述 4 种不同情况下的小波变换图,这里采用的是 db16 小波。

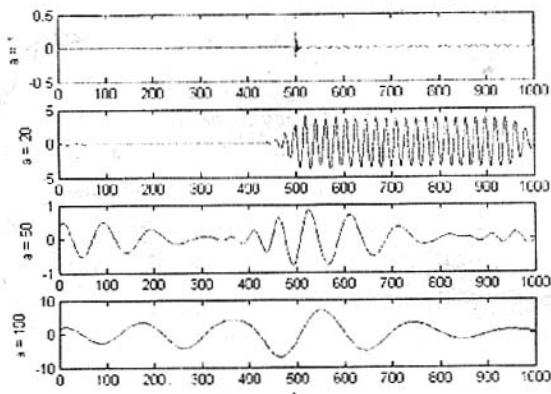


图 4 小波变换多尺度比较图

从图中可知,输入的样本数据在 $t = 500$ 时频率

发生了变化,尤其在尺度因子 $a = 1$ 时更为明显。因此在由多种频率构成的序列数据的分析中,采用小波分析可以获得较高的精度,同时通过比较可以发现小波变换的方法可以较好地发现突变信息,从而为后续预测的准确性提供了有力的保障。

2.3 对噪声数据的容错性分析

由于在现实中的数据采集常受错误和误差的影响,也即噪声,因此降低噪声数据对时间序列分析的影响就显的相当重要,其不仅可以缩小数据集的大小,从而提高分析的效率,还可以减小误差,提高分析预测的准确性。

文中以白噪声为例来分析小波变换对噪声数据的容错性效果。白噪声通常是指自相关函数为 $R(t_1, t_2) = I(t_1)\delta(t_1 - t_2)$ 的随机过程 $x(t)$ ^[6]。易见当 $t_1 \neq t_2$ 时, $R(t_1, t_2) = 0$ 。采用函数 $s(x) = \cos(0.03\pi t)e^{-0.1t} + b(t)$ 来产生测试样本数据,其中 $b(t)$ 为白噪声。用二维连续小波变换分别对未加白噪声的数据和加了白噪声的数据进行处理,这里使用 `coif1` 小波。图 5 所示为加噪声数据的时频图,图 6 所示为未加噪声与加噪声的数据经二维 CWT 变换后的时频图,通过变换后的波形比较,可以看到,加了白噪声后的时频图较之未加白噪声时的时频图变化很小。由此可得二维连续小波变换对于噪声有良好的抑制作用。

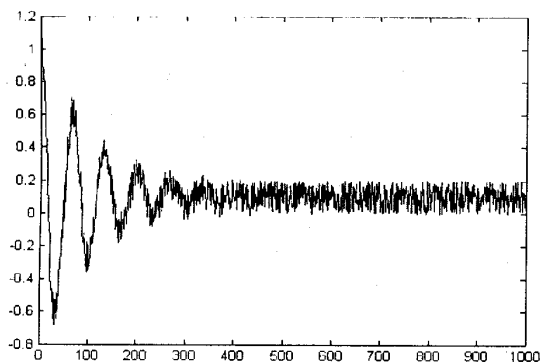


图 5 数据加噪声后的时频图

2.4 进一步的讨论

由上述分析可知,小波分析在对多频率信号和突变信号的分析上具有其他方法无法替代的优势,尤其小波分析在对噪声数据处理上,显示了较强的抗噪性。同时由于连续小波变换的系数具有较大的冗余量,其计算量较大,但该冗余量却对实现数据去噪和数据恢复十分有利。比如可以利用小波系数的一个子集来重

构原始信号,这样可极大地降低噪声对信号的影响,从而提高时间序列数据分析的准确性。因此如果能在减小计算量与去噪之间取得一个平衡,那么连续小波变换系数的冗余性就将是连续小波变换的一大优势。

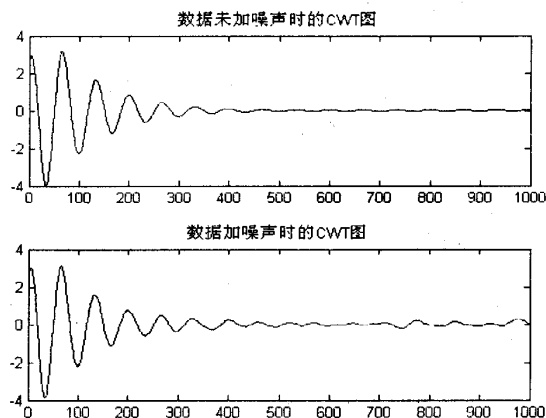


图 6 未加噪声与加噪声的数据经 CWT 变换后的时频图比较

3 结束语

二维连续小波变换具有时频联合分析、多分辨率及有效的抗噪性等特点,利用之可对时间序列数据中的非线性和带扰动的数据进行有效分析,尤其可显著提高数据抗噪声能力。若与一系列时间序列模型结合使用,则可大大提高预测的效率和精度,是未来极有价值的研究方向之一,有着广阔的应用前景。

参考文献:

- [1] Box P, Jenkins M, Reinsel C. 时间序列分析: 预测与控制 [M]. 顾 岚, 范金城译. 北京: 中国统计出版社, 1997.
- [2] Nason G P, von Sachs R. Wavelet in Time-series Analysis [J]. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 1999, 357(1760): 2511-2526.
- [3] Chiann C, Morettin P A. A wavelet analysis for time series [J]. Journal of Nonparametric Statistics, 1999, 10(1): 1-46.
- [4] Percival D, Walden A T. Wavelet Methods for Time Series Analysis [M]. Cambridge: Cambridge University Press, 2000.
- [5] 彭玉华. 小波变换与工程应用 [M]. 北京: 科学出版社, 2000: 13-27.
- [6] 孟 波. 基于小波分析的数据挖掘方法的研究 [D]. 南京: 南京大学, 2000.