

## 噪声环境下语音分形特征的提取和分析

包永强 赵力 邹采荣  
(东南大学无线电工程系 南京 210096)

**摘要:** 该文针对目前的分维计算方法——盒维、关联维等精度虽高,但计算复杂,Katz 维计算简单、抗噪性能好、但精度不高的现状,提出了一种改进的基于波形的算法——IBW-FD,分析了对分形布朗曲线、含噪语音(高斯白噪声,三种非平稳噪声)的性能。理论分析和实验结果表明:IBW-FD 算法具有更强区分高斯白噪声和语音信号的能力;IBW-FD 算法抗平稳和非平稳噪声能力要普遍好于盒维和 Katz 维。结果表明 IBW-FD 算法在复杂度、精确度和抗噪性能方面均优于现有的分维算法,是一种比较好的分维计算方法,不仅可以应用在语音处理中,而且也可应用于其它信号处理中。

**关键词:** 语音信号处理;分形维;语言增强

中图分类号: TN912.3

文献标识码: A

文章编号: 1009-5896(2007)03-0585-04

## The Abstraction and Analysis of Fractal Characteristic of Noisy Speech

Bao Yong-qiang Zhao Li Zou Cai-rong

(Department of Radio Engineering, Southeast University, Nanjing 210096, China)

**Abstract:** According to simple computation, good anti-noise ability and low precision of Katz algorithm and complex computation and good precision of box-counting dimension and correlation dimension, an Improvement fractal algorithm Based on Wave (IBW) is presented and analyzed through the fractal Brown curve and noisy speech according to the characteristic of the box dimension and Katz dimension. The theory analyse and experiment show that IBW-FD has lower computation and higher precision than Katz dimension and box-counting dimension. IBW-FD also has stronger ability of anti-noise and distinguish Gaussian noise and speech than the others. It shows that IBW-FD is the good speech fractal algorithm because of low complexity, good precision and nice anti-noise ability.

**Key words:** Speech signal processing; Fractal Dimension(FD); Speech enhancement

### 1 引言

90 年代以来,混沌开始在语音研究领域中得到了应用<sup>[1,2]</sup>。Berry 等人<sup>[1,2]</sup>从语音的生成模型着手,研究声带振动的力学机理,在声带振动轨迹的相图中发现了分岔和混沌现象,并利用相图、频谱等手段指出浊音信号具有低维混沌的特征。

分形维理论<sup>[3]</sup>是描述混沌信号的一种手段,盒维、关联维等常用来对噪声环境下的语音信号<sup>[4-9]</sup>进行分析,目前分形理论主要应用于语音分割、端点检测和语音识别或说话人识别的特征提取等方面。Boshoff<sup>[4]</sup>将盒维应用于语音分割,Pitsikalis<sup>[5]</sup>等人将关联维作为语音识别的参数来进行识别;2003 年 Gu 等人<sup>[6]</sup>利用混沌和分形理论进行噪声环境中的语音端点检测, Yang 等<sup>[7]</sup>在盒维的基础上提出了一种分维算法用作端点检测; Jungpa Seo 等人<sup>[8,9]</sup>将元音的相关维作为说话人识别的参数。

无论是用作特征提取还是端点检测,应该说由于分形维

的引入,性能有了很大的提高,但是目前常用的分维算法中,计算量是限制其应用的重要原因。对于盒维来说,由于随着系统维数的增加,需要的盒子数量指数增加,需要用于计算的数据长度也相应增加,使得高维系统的收敛非常困难。关联维能够比较有效地计算语音的分形维数,区分出噪声和语音,但是它严重依赖于两个参数:最小延迟时间和最小嵌入维:如果最小延迟时间太大,达到吸引子大小,则所有矢量都是相关的,太小则计算出的相关维就是噪声的相关维。只有当嵌入维足够大时,计算出的嵌入空间的关联维才与吸引子的关联维相同,但是在实际环境中,噪声的影响会导致关联维曲线的发散,从而导致得不到准确的嵌入维。

Katz 在 1988 年提出了一种基于波形长度的计算分维的算法<sup>[10]</sup>,该算法将波形长度和第一点抽样值与序列中其他点的距离之最大值的的关系作为分维值,与盒维和关联维相比,Katz 维不需要计算极限和对参数依赖性不强,Rosana<sup>[11]</sup>认为该算法在抗噪声能力上优于其他算法。但是 Katz 维存在着精度不高的缺点。

本文基于 Katz 算法利用波形长度来提取分维的思想,提出了一种新的分维计算算法——改进的基于波形的算法

(IBW-FD)。

## 2 一种改进的基于波形长度的盒维算法

Katz 算法虽然精度不高,但是它在抗噪性能上却优于其他算法<sup>[12]</sup>, Rosana 指出,与其他分维计算方法相比,利用波形的长度可以获得较好的抗噪性能<sup>[12]</sup>。因此,为了获得一种抗噪性能比较好,精度比较好,并且计算量不高的算法,本章进行如下推导:

由分形维定义可得:

$$D = \lim_{\delta \rightarrow 0} - \frac{\ln N_{\delta}(F)}{\ln \delta} \quad (1)$$

其中  $N_{\delta}(F)$  是尺度为  $\delta$  的盒覆盖曲线的最小个数。令

$$L(\delta) = \sum_{i=1}^{N_{\delta}(F)} l_i(\delta) \quad (2)$$

$l_i(\delta)$  是边长为  $\delta$  的第  $i$  盒内曲线长度。对于连续信号而言,当  $\delta$  趋向于一个非常小的数时,  $L(\delta)$  越接近于波形的实际长度,

当  $\delta$  小于等于某一个值  $\delta_1$  时,可认为:当  $\delta \leq \delta_1$  时,对于任意  $1 \leq i \leq N_{\delta}(F)$ , 都有  $l_i(\delta) = l(\delta)$ , 则有

$$L(\delta) = N_{\delta}(F)l(\delta), \quad \forall \delta \leq \delta_1 \quad (3)$$

则式(1)可变为

$$D = \lim_{\delta \rightarrow 0} - \left( \frac{\ln(L(\delta)/l(\delta))}{\ln \delta} \right) = \lim_{\delta \rightarrow 0} \left( \frac{\ln l(\delta)}{\ln \delta} - \frac{\ln L(\delta)}{\ln \delta} \right)$$

设  $0 < \delta \leq \delta_0$ , 用泰勒级数展开 ( $\delta_0$  为接近 0 的非常小的值,  $0 < \delta_2 < \delta$ ):

$$D = \left( \frac{\ln l(\delta_0)}{\ln(\delta_0)} - \frac{\ln L(\delta_0)}{\ln(\delta_0)} \right) + \lim_{\delta \rightarrow 0} \left( \frac{\ln l(\delta)}{\ln(\delta)} - \frac{\ln L(\delta)}{\ln(\delta)} \right) \Bigg|_{\delta=\delta_2} \cdot (\delta - \delta_0) \quad (4)$$

对上式第 2 部分进行计算得

$$\begin{aligned} & \lim_{\delta \rightarrow 0} \left( \frac{\ln l(\delta)}{\ln(\delta)} - \frac{\ln L(\delta)}{\ln(\delta)} \right) \Bigg|_{\delta=\delta_2} (\delta - \delta_0) \\ &= \lim_{\delta \rightarrow 0} \left( \frac{l'(\delta_2)(\delta - \delta_0)}{l(\delta_2) \ln \delta_2} - \frac{L'(\delta_2)(\delta - \delta_0)}{L(\delta_2) \ln \delta_2} \right. \\ & \quad \left. - \frac{\ln l(\delta_2)(\delta - \delta_0)}{\delta_2 \ln^2 \delta_2} + \frac{\ln L(\delta_2)(\delta - \delta_0)}{\delta_2 \ln^2 \delta_2} \right) \end{aligned} \quad (5)$$

由于  $\delta \rightarrow 0$  和  $\delta_2 < \delta$ , 可得  $\delta_2 \rightarrow 0$ , 则式(5)可变为

$$\begin{aligned} & \lim_{\delta \rightarrow 0} \left( \frac{\ln l(\delta)}{\ln(\delta)} - \frac{\ln L(\delta)}{\ln(\delta)} \right) \Bigg|_{\delta=\delta_2} (\delta - \delta_0) \\ & \approx \lim_{\delta \rightarrow 0} \left( \frac{l'(\delta)(\delta - \delta_0)}{l(\delta) \ln \delta} - \frac{L'(\delta)(\delta - \delta_0)}{L(\delta) \ln \delta} \right. \\ & \quad \left. - \frac{\ln l(\delta)(\delta - \delta_0)}{\delta \ln^2 \delta} + \frac{\ln L(\delta)(\delta - \delta_0)}{\delta \ln^2 \delta} \right) \end{aligned} \quad (6)$$

为简化分析起见, 设  $l(\delta) = k\delta$ , 则式(6)括号中第 1 项可变为

$$\begin{aligned} \lim_{\delta \rightarrow 0} \frac{l'(\delta)(\delta - \delta_0)}{l(\delta) \ln \delta} &= \lim_{\delta \rightarrow 0} \frac{k(\delta - \delta_0)}{k\delta \ln \delta} = \lim_{\delta \rightarrow 0} \left( \frac{1}{\ln \delta} - \frac{\delta_0}{\delta \ln \delta} \right) \\ &= - \lim_{\delta \rightarrow 0} \frac{\delta_0}{\delta \ln \delta} \end{aligned} \quad (7a)$$

由于  $L(\delta)$  趋向于曲线实际值, 其导数  $L'(\delta)$  趋向于极小值, 因此(6)式括号中第 2 项分子趋向于 0, 分母趋向  $-\infty$ , 第 2 项可为 0。

式(6)括号中第 3 项可变为

$$\begin{aligned} & \lim_{\delta \rightarrow 0} \left( \frac{\ln k}{\ln^2 \delta} + \frac{1}{\ln \delta} - \frac{\ln(k\delta)\delta_0}{\delta \ln^2 \delta} \right) \\ & \Rightarrow \lim_{\delta \rightarrow 0} \left( - \frac{\ln k \times \delta_0}{\delta \ln^2 \delta} - \frac{\delta_0}{\delta \ln \delta} \right) \end{aligned} \quad (7b)$$

式(6)括号中第 4 项可变为

$$\lim_{\delta \rightarrow 0} \frac{\ln(L(\delta)) \times (\delta - \delta_0)}{\delta \ln^2 \delta} = \lim_{\delta \rightarrow 0} \left( \frac{\ln L(\delta)}{\ln^2 \delta} - \frac{\ln L(\delta) \times \delta_0}{\delta \ln^2 \delta} \right) \quad (7c)$$

因此, 式(6)可变为

$$\begin{aligned} & \lim_{\delta \rightarrow 0} \left( \frac{\ln l(\delta)}{\ln(\delta)} - \frac{\ln L(\delta)}{\ln(\delta)} \right) \Bigg|_{\delta=\delta_2} (\delta - \delta_0) \\ & \Rightarrow \lim_{\delta \rightarrow 0} \left( - \frac{2}{\delta \ln \delta} - \frac{\ln k \times \delta_0}{\delta \ln^2 \delta} \right) + \lim_{\delta \rightarrow 0} \left( \frac{\ln L(\delta)}{\ln^2 \delta} \left( 1 - \frac{\delta_0}{\delta} \right) \right) \end{aligned} \quad (8)$$

对于离散时间序列来说,  $\delta_0$ ,  $\delta$  不能无限小, 它们的最小值为离散时间序列的抽样间隔。

为了简化分析起见, 令

$$\delta_0 = m\delta, \quad \beta = 1 - \frac{\delta_0}{\delta} \quad (9)$$

将式(8)、式(9)代入式(6)并化简整理可得

$$D \approx a - \frac{\ln L(m\delta)}{\ln(m\delta)} + \beta \frac{\ln L(\delta)}{\ln^2 \delta} \quad (10)$$

式(10)中:

$$a = \lim_{\delta \rightarrow 0} \left( \frac{\ln(km\delta)}{\ln(m\delta)} - \frac{2}{\delta \ln \delta} - \frac{\ln k \times m\delta}{\delta \ln^2 \delta} \right) \quad (11)$$

为了进一步降低计算量, 假定由较小的尺寸获得的曲线长度基本不变, 即  $L(m\delta)$  近似等于  $L(\delta)$ , 由此带来的误差可由参数  $\alpha, \beta, m$  得以调整, 则式(10)简化为

$$D_{\text{IBW}} = \alpha - \frac{\ln L(\delta)}{\ln(m\delta)} + \beta \frac{\ln L(\delta)}{\ln^2 \delta} \quad (12)$$

本文将式(12)称为改进的基于波形的分形维算法(IBW-FD), 上式中  $\delta$  可为最小抽样间隔的整数倍,  $L(\delta)$  为对应的曲线长度。  $\alpha, \beta, m$  的取值需要在实验中确定。

## 3 IBW 短时分形维算法的性能分析

### 3.1 IBW-FD 算法的 $\alpha, \beta, m$ 参数的确定

IBW-FD 的  $\alpha, \beta, m$  参数采用分形布朗曲线来确定。实验中, 令分形布朗曲线的分维值在 1.1~1.9 中变化, 即分形布

朗曲线的分维差为  $1.9-1.1=0.8$ 。实验的目的是寻找那些使得 IBW 分维的波动范围在 0.8 左右的参数值。图 1, 图 2 中的 IBW-FDD 为 IBW 分维差 (IBW Fractal Dimension Diference, IBW-FDD)。

(1) $\beta$  的讨论: 图 1 给出了  $\alpha, m$  一定时, 不同  $\beta$  对应的 IBW-FD 算法得出的分维差值。从图中可看出,  $\beta$  与 IBW-FD 分维差值呈线性关系, 当  $0.6 < \beta < 1.2$  时, IBW-FD 分维差值  $\approx 0.8$ , 此时对应的 IBW-FD 分维值没有明显的压缩和拉长。

(2) $m$  的讨论: 图 2 给出了  $\alpha, \beta$  一定时,  $m$  不同对应的 IBW-FD 算法得出的分维差值。从图中可看出, 当  $8 \leq m \leq 11$  时, IBW-FD 分维差值  $\approx 0.8$ , 此时对应的 IBW-FD 分维值没有明显的压缩和拉长。

(3) $\alpha$  的讨论:  $\alpha$  的值只影响 IBW-FD 分维的绝对值, 确定了  $\beta, m$  的值, 也就确定了  $\alpha$  的值。由上面的分析可知,  $m$  在 8~11 之间,  $\beta$  在 0.6~1.2 之间,  $\alpha$  的值随之就可确定。

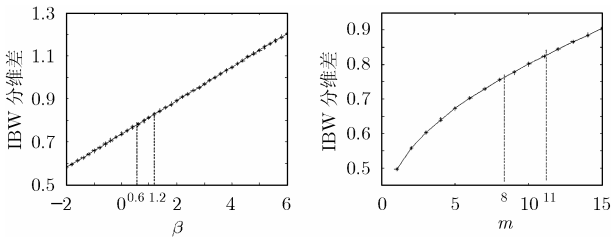


图 1  $\alpha, m$  一定时不同  $\beta$  对应的 IBW-FD 算法的分维数差值 ( $\alpha=0.8, m=10$ )

图 2  $\alpha, \beta$  一定时不同  $m$  对应的 IBW-FD 算法的分维数差值 ( $\alpha=0.8, \beta=0.9$ )

### 3.2 分形布朗曲线下的 IBW 短时分形维算法的性能

表 1 中 FBC 为分形布朗曲线 (Fractal Brownian Curve, FBC), 为了比较 IBW-FD 算法的性能, 本小节将之与盒维和 Katz 维进行比较, 表 1 的分形布朗曲线是计算机通过 15 次迭代产生的, 共 32786 个点, 程序将该曲线分帧, 每帧 512 个点, 计算每帧的分形维。表中的数值是将各帧分维取平均得到的。其中 IBW-FD 算法的参数值取为  $m=8, \beta=0.9, \alpha=0.85$ 。表 1 给出了盒维, Katz 维, IBW-FD 的分维均值。IBW-FD 算法的分维均值最接近于理论值, 其次是盒维, 最后才是 Katz 算法。

表 1 IBW-FD 算法和盒维、Katz 维的分维均值比较

曲线	分维	盒维	Katz	IBW-FD	曲线	分维	盒维	Katz	IBW-FD
直线	1	0.9531	0.9993	0.9982	FBC	1.5	1.4575	1.4735	1.5272
FBC	1.1	1.1744	1.1617	1.0767	FBC	1.6	1.5264	1.5629	1.6380
FBC	1.2	1.2226	1.1825	1.2079	FBC	1.7	1.6084	1.6488	1.7193
FBC	1.3	1.2890	1.2470	1.2909	FBC	1.8	1.6714	1.7558	1.7897
FBC	1.4	1.3834	1.3494	1.4181	FBC	1.9	1.7044	1.8287	1.8619

## 4 含噪语音的短时分形维算法的性能分析

### 4.1 高斯白噪声环境下短时语音的分维分析

在实验中, 语音数据为在实验室内录制的语音, 采样频率是 8kHz, 8bit 采样。在纯净语音上叠加高斯白噪声和非平稳噪声 (噪声源由英国 TNO 感知学会所属的荷兰 RSRE 语音研究中心提供)。实验中, 数据每帧长  $N=256$ , 50% 的帧重叠。

图 3 给出了含噪语音 (0dB 高斯白噪声) 和纯净语音的短时分维曲线对比, 纵坐标为分维值 (FD), 横坐标为帧数 (frame  $n$ ), 从图中可看出: (1) 含噪语音中, 对于 IBW-FD 维来说, 维数大于 1.8 的基本上是噪声帧, 小于 1.8 的基本上为语音; (2) 含噪语音的语音段的 IBW-FD 和 Katz 维明显要小于噪声段, 而盒维则两者差别不大, 这体现了 IBW-FD 算法和 Katz 维有更强的区分出噪声和语音的能力; (3) 含噪语音的语音段与纯净语音的 IBW-FD 维差要小于 Katz 维差。图 3 说明, 在强噪声情况 (高斯白噪声) 下, IBW-FD 算法仍然能够区分出噪声部分和语音部分, 并且语音帧的 IBW-FD 分形维比 Katz 维和盒维更接近于纯净语音的分维。

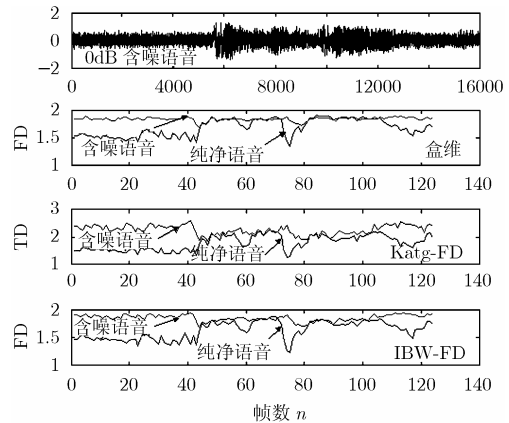


图 3 IBW-FD 算法与 Katz 分维、盒维的对比

### 4.2 多种噪声影响下语音信号的分维方差分析

表 2 给出了语音帧在高斯白噪声、粉红色噪声 (pink)、汽车发动机噪声 (volvo)、车间噪声 (factory) 下各分形维与纯净语音时的方差。表 2 说明, IBW-FD 受高斯白噪声、粉红色噪声 (pink) 和车间噪声 (factory) 的影响要比盒维和 Katz 维来的更小, 受汽车发动机噪声 (volvo) 的影响要大于盒维。

表 2 各种噪声下语音帧的 IBW-FD 维、盒维、Katz 维与纯净语音的方差对比

		白噪声	Pink	volvo	Factory
0dB	盒维	0.4035	0.3620	1.8954	0.3671
	Katz 维	6.9154	1.8301	9.5039	1.7432
	IBW-FD	0.3774	0.2951	2.7352	0.3255
10dB	盒维	0.3031	0.1072	0.7897	0.0914
	Katz 维	5.0829	1.0674	5.0369	0.6875
	IBW-FD	0.2333	0.0908	1.1374	0.0849
20dB	盒维	0.2254	0.0755	0.2277	0.0306
	Katz 维	3.8773	0.9066	2.0623	0.3109
	IBW-FD	0.1647	0.0543	0.2951	0.0244

## 5 计算复杂度分析

表 3 中, 假设曲线有  $N$  个点, 计算 IBW-FD、Katz 维的步长为 1, 计算盒维的最小盒长也为 1。从表 3 中可看出, 盒维的加法次数、乘法次数、非线性运算次数均要明显大于 Katz 维和 IBW-FD 方法; 而 IBW-FD 方法在加法、乘法、比较、非线性运算方面在三者中均处于最小。

表 3 IBW-FD、Katz 维和盒维的计算复杂度比较

	加	乘	比较	非线性
盒维	远大于 $N$	远大于 $N$	0	远大于 $N$
Katz 维	$3(N-1)$	4	$N-1$	$4(N-1)$
IBW-FD	$2N$	4	0	$N+4$

## 6 结束语

由于目前的分维计算方法普遍比较复杂, 本文根据 Katz 维的特点, 提出了一种改进的基于波形的算法——IBW-FD。通过分形布朗曲线验证了该算法比盒维和 Katz 维算法具有更高的精确度。并分析了在高斯白噪声情况下和 3 种非平稳噪声下的性能。

分析结果表明: (1)IBW-FD 的精确度要比 Katz 维和盒维高; (2)IBW-FD 具有更强区分高斯白噪声和语音信号的能力; (3)在高斯白噪声、粉红色噪声(pink)和车间噪声下 IBW-FD 的抗噪声能力要好于盒维和 Katz 维, 抗汽车发动机噪声能力则略低于盒维; (4)IBW-FD 分维的计算复杂度最小。

IBW-FD 算法的良好性能使得它不仅能够应用于噪声环境下语音信号的分析 and 特征提取方面, 而且也能应用于其他信号处理的分形维计算。

## 参考文献

- [1] Berry D A, Herzel H, Titze I R, and Krischer K. Interpretation of biomechanical simulations of normal and chaotic vocal fold oscillations with empirical eigenfunctions. *J. Acoust. Soc. Am.*, 1994, 95(6): 3594-3604.
- [2] Kumar A and Mullick S K. Nonlinear dynamical analysis of speech. *J. Acoust. Soc. Am.*, 1996, 100(1): 615-620.
- [3] Mandelbrot B B and Vess J W V. Fractional Brownian motion, fractional noise and application. *SIAM Review*, 1968, 10(4): 540-546.
- [4] Boshoff H F V. A fast box counting algorithm for determining the fractal dimension of sampled continuous functions. COMSIG'92., Proceedings of the 1992 South African Symposium on 11 Sept., 1992: 43-48.
- [5] Pitsikalis V and Maragos P. Speech analysis and feature extraction using chaotic models. ICASSP'2002, IEEE International Conference on, Orlando FL, 2002, 1: I533-I536.
- [6] Gu Lingyun, Gao Jianbo, and Harris G J. Endpoint detection in noisy environment using a pointcare recurrence metric. ICASSP'03, 2003 IEEE International Conference on, Hongkong, 2003: I428-I431.
- [7] Yang Su and Li Zong-ge. A fractal based voice activity detector for internet telephone, ICASSP 2003, IEEE International Conference on, Hongkong, 2003: I808-I811.
- [8] Seo Jungpa, Hong S, Kim M, Baek I, Kwon Y, and Lee K. New speaker recognition feature using correlation dimension. ISIE 2001, IEEE International Symposium on Pusan, 2001, 1: 505-507.
- [9] Petry A and Barone D A C. Fractal dimension to speaker identification. ICASSP'01, 2001 IEEE International Conference on Salt Lake City, UT, 2001, 1: 405-408.
- [10] Katz M. Fractals and the analysis of waveforms. *Comput. Biol. Med.*, 1988, 18(3): 145-156.
- [11] Esteller R and Vachtsevanos G. A comparison of waveform fractal dimension algorithms. *IEEE Trans. on Circuits and Systems*, 2001, 48(2): 177-183.

包永强: 男, 1973 年生, 博士生, 研究方向为语音信号处理。

赵力: 男, 1958 年生, 教授, 博士生导师, 研究方向为语音信号处理。

邹采荣: 男, 1963 年生, 教授, 博士生导师, 副校长, 研究方向为信号处理。