利用相空间重构时间延迟特性进行语音音质分析

杨云升† 刘志华 李博章†
(63888部队 河南省济源市 454650)

摘要 在简要分析相空间重构时间延迟计算方法的基础上,利用复自相关时间延迟计算方法进行了语音信号相空间重构时间延迟的计算分析,通过对比不同信噪比下复自相关-时间延迟曲线,提出了利用时间延迟参数进行语音音质评估的方法,利用BP网络对大量试验数据进行建模,通过统计分析认为该方法具有一定的可行性。

关键词 音质评估,时间延迟,相空间重构,复自相关

Evaluating speech quality based on the time delay of phase space reconstruction

YANG Yun-Sheng LIU Zhi-Hua LI Bo-Zhang
(*Unit 63888 of PLA, Jiyuan* 454650)

Abstract A method to obtain the time delay of phase space reconstruction is introduced, and the time delay in phase space reconstruction of speech signals are computed. A new way of evaluating the speech quality is introduced through contrasting the complex self-correlations and time delay curves of signals of various SNR, which is proven viable by statistically analyzing a large number of examination data.

Key words Speech quality evaluation, Time delay, Phase space reconstruction, Complex self-correlation

1 引言

语音是一种最常用的、最直接的信息传播方式。对语音音质进行快速准确的评价是语音信号处理很重要的一个方面,现在对语音音质的评价还主要集中在语音的主观评价(通过人耳进行直接听判)上,这种评价方法费时、费力,对人员要求比较高,且一致性比较差[1]。客观评价方法(利用参数进行音质分析)主要包括基于输入-输出和基于输出两种模式,基于输入-输出的客观评价方法[2]近年来发展比较快,形成了多种算法,且得到了一定程度的应用,但这种模式的评估方法主要是基于受扰前后的语音参数变化进行音质评估,需要进

2007-06-12 收稿; 2007-10-17 定稿

作者简介:杨云升(1973-)男,河南沁阳市人,硕士研究生、工程师、研究方向:语音通信及干扰。
刘志华(1975-)男,硕士研究生、工程师。 李博章(1979-)男,硕士研究生、工程师。
† 通信联系人 E-mail:yunshengyang@126.com

(C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

行严格的输入输出同步,对测评语音要求苛刻,严重限制了该方法在实际工作中的进一步应用;基于输出模式的客观评估方法[2]通过直接计算受扰语音参数进行音质评估,这种模式的评估算法近年来虽也得到了很大程度的重视,却还限于理论研究,应用中比较少见,但该模式的评估方法实现起来比较简单,具有良好的应用前景。本文通过分析语音信号的混沌参数——时间延迟特性来建立一种基于相空间时间延迟的语音音质评估方法,该方法是一种基于输出模式的语音音质评估方法。

2 相空间重构中时间延迟的选取

混沌是对确定性系统中非确定性的一种描述,混沌现象具有较强的类随机行为,在自然界中普遍存在,通常情况下混沌系统的系统变量是不可知或不完全可知的,因而不同于常规系统。进行混沌系统研究时主要是通过观测时间变量对系统进行整体特性的研究,其中常用的方法就是混沌吸引子的状态空间重构方法,状态空间重构的基本思想[3]是,通过一个观测变量所得到的时间序列中包含了未观测到的状态变量的信息,状态空间重构的目的就是要通过重构技术使得时间序列中包含的状态变量的信息更清晰地显现出来。针对时间序列的状态空间重建,最常用的方法是时间延迟法[3],它基于这样一个概念(Takens 的嵌入定理):用时间延迟的状态矢量重建得到的动态系统在某种意义上等效于原始动态系统。延迟法应用延迟坐标重构相空间矢量,即

$$X_m(n)=[x_n,x_{n+\iota},x_{n+2\iota},\Lambda,x_{n+(m-1)\iota}] \quad (1)$$

式中:x_n 是已离散化的待研究系统在 n 时刻的观测值,ι 称时间延迟,m 称重构维。

如何选择适当的重构维和时间延迟是相空间重构的主要研究内容。现有的时间延迟的选择方法一般都基于如下的两个准则之一[4]:

(1)序列相关法。让 X_m 内元素之间的相关性减弱,同时确保 X_m 包含的原动力学系统的信息不丢失;

(2)相空间扩展法。重构相空间轨迹应从相空间的主对角线(ι 很小时)尽可能地扩展,但又不出现折叠。

例如,自相关法、互信息量法、高阶相关法等属于第一类准则;填充因子法、摆动量法、平均位移法、SVF 法等属于第二类准则。

其中平均位移法是对相空间的一种几何度量方法,概念清晰,计算简洁,受到普遍重视[5],但文献[5]认为该方法在求取时间延迟参数时具有一定的随意性,仅仅根据实验结果得到,其理论根据不强。因而文献[6]在平均位移法方法的基础上结合自相关法提出了一种新的时间延迟计算方法——复自相关法,在忽略边缘点带来的差别的情况下由平均位移法和自相关法法推导出了复自相关法的公式表示形式:

$$R_{xx}^m(\iota)=\frac{1}{N}\sum_{i=0}^{N-1}\sum_{j=1}^{m-1}x_i x_{i+j\iota} \quad (2)$$

选取 $R_{xx}^m(\iota)$ 的第一零点为时间延迟 τ,复自相关法继承了平均位移法的几何意义,同时克服了自相关法的缺点,实现了构建向量内部的最大不相关,文献[6]讨论了复自相关法的高维特性、抗噪能力和对数据长度需求,其物理意义更符合相空间重构的轨迹扩展;由于复自相关法取函数的过零点为特征值,在求法上也比平均位移法的求函数斜率简单。下面我们将用复自相关法对语音信号进行具体分析。

3 语音信号的相空间时间延迟特性

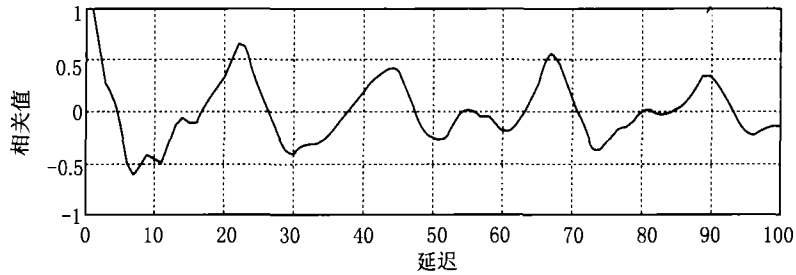
语音是声带振动发出的一种声音,具有特定的时域和频域结构,严格的声学和空气

(C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

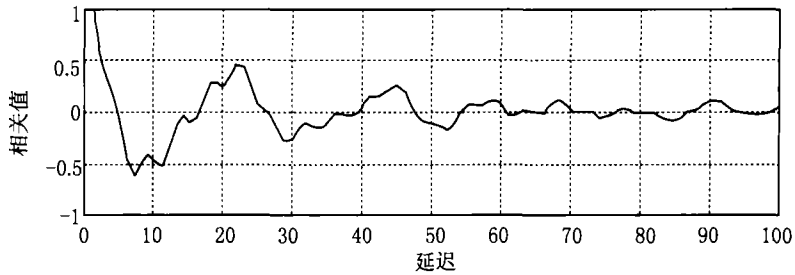
动力学业已证明:语音信号中存在着混沌机理[7],因而通过分析语音信号的混沌特性是近年来进行语音特性分析的一种全新方法,国内外涌现出了大量的利用混沌方法进行语音识别、端点检测、语音增强等方面的相关方法[8-9]。

3.1 不同延迟下语音信号的复自相关特性

下面我们利用复自相关法具体计算一段语音信号的时间延迟情况,计算公式如式2。首先看一下未受干扰语音的复自相关性随时间延迟的变化情况。



(a)语音长度为1s时的复自相关变化情况



(b)语音长度为10s时的复自相关变化情况

图1　不同长度下语音的复自相关特性随时间延迟的变化情况

图1中所有语音均为8kHz采样,并且去除了语音中间的静音部分。图1中(a)图为一段长度为1s的语音相关性随时间延迟的变化情况,可以看到:由于语音中存在着明显的基音倍周期,其相关性在不同时间延迟上也表现出了一定的周期现象,第一过零点是其最小的一个不相关点,以此来确定建立相空间的时间延迟是具有一定道理的。但本文中主要讨论的不是时间延迟参数的确立,而是通过时间延迟的变化进行语音音质分析,下面再看一下图1中(b)图,该图是长度为10s的语音相关性随维数的变化情况,相较于(a)图,该图的周期性明显变弱,但两图在第一个周期(尤其在前半个周期)内具有较强的一致性,因而在下文的讨论中主要采用前半个周期的数据进行处理。

3.2 不同受扰程度下复自相关特性随延迟的变化

噪声是一种随机信号,不具有特定的时频域特性,数据之间不具有相关性,因而其复自相关图不具有周期性。当语音受扰干扰后,在时频域上分布的语音特性会减弱,同时一定程度的具有噪声特性,利用这种现象可以进行语音的音质分析。

根据公式2的计算方法,图2给出了信噪比分别为20dB、2dB、-2dB和-6dB四个不同等级下的语音复自相关性随时间延迟的变化情况,由图中可以看出随着信噪比变化复自相关性是逐渐变化的,研究中主要通过第一拐点位置和其相对值来表征这种变化,

(C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

可以看出:随着信噪比变小第一拐点位置会逐渐前移,其相对值(相对于第一个位移点的大小)也会逐渐变小,并且拐点的变化具有阶跃性,为了更清晰地表示这种变化,表 1 给出一组逐渐变化受扰文件的复自相关特征值变化情况。

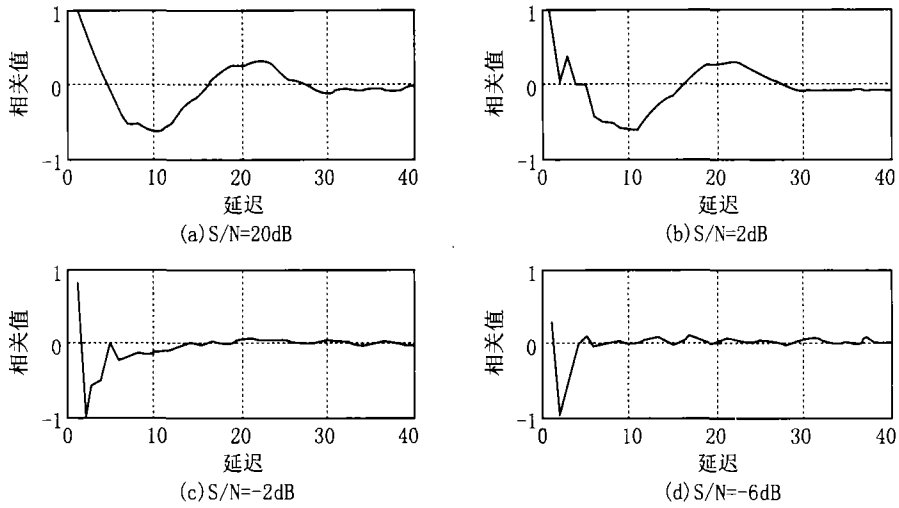


图 2 不同信噪比下信号复自相关随延迟的变化

表 1 不同信噪比下语音信号复自相关特征随延迟的变化情况

信噪比/dB	第一拐点位置	第一拐点相对值	信噪比/dB	第一拐点位置	第一拐点相对值
10	10	-0.62845	0	2	-0.74162
8	10	-0.62983	-1	2	-1.024766
6	10	-0.6306	-2	2	-1.189222
4	10	-0.63735	-3	2	-1.411773
3	10	-0.6358	-4	2	-1.685008
2	2	0.04938	-5	2	-2.280881
1	2	-0.34892	-6	2	-3.840216

表 1 中所有文件为一个标准语音文件在不同受扰程度下得到,信号均采用 8kHz 采样、16bits 量化。由表中可以看出:第一拐点位置开始是 10,信噪比变差后第一拐点位置变为 2,在拐点位置一定的情况下,随着信噪比的变化第一拐点相对值逐渐变小。

大量的数据分析中可以发现拐点位置并非只有 10 和 2 两个,具体在受扰语音中节点位置可能为 2 ~ 12 中的任何值,但信噪比高时节点位置靠后、信噪比低时节点位置靠前的趋势是一致的,利用这种变化关系,可以一定程度地的进行语音音质的分析。

4 利用语音的相空间时间延迟建立语音音质分析模型

研究过程中录制了 700 多个不同受扰等级的语音文件,通过对大量受扰语音的分析计算,利用 BP 网络进行建模处理。网络结构参数如下:网络由三层组成,第一层为输入层,节点数为 2(由输入参数个数决定),第二层为传输层,节点数为 10,传递函数为“tansig”型函

(C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

数,第三层为输出层,节点数为 1(由输出参数个数决定),传递函数为"logsig"型函数,网络训练方法采用"trainlm"方法,图 3 给出了网络拟合结果与主观实际 MOS 的对应关系。

网络训练结果如图 3 所示。由图中可以看出,经过网络拟合,结果基本上分布在主对角线附近,具有较好的一致性,经计算数据分布的标准差为 0.1940,拟合的相关系数为0.9809。

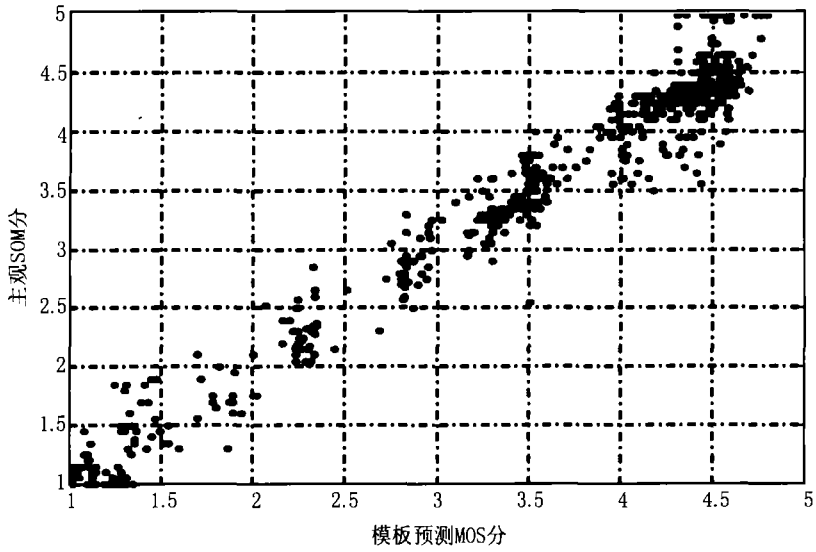


图 3　利用时间延迟参数建模训练结果

由上述计算分析可知时间延迟参数可以较好地表征信号受扰程度的强弱。

5　结束语

结合大量试验数据,在分析统计的基础上给出了一种基于输出的语音音质评估方法,具有一定的应用前景,但本文的分析主要针对噪声调频干扰方式进行了研究,至于别的干扰样式下的模板有待进一步的研究。

参　考　文　献

[1] 易克初,田斌,付强. 语音信号处理. 北京:国防工业出版社,2000:331 - 339

[2] 付强. 语音的参数表示和质量客观评价研究. 西安电子科技大学博士学位论文,2000.9

[3] 林嘉宇,黄芝平,王跃科,等. 语音信号相空间重构中嵌入维数的选择. 电子科学学刊,1999,21(6):735 - 742

[4] 马红光,李夕海,王国华,等. 相空间重构中嵌入维和时间延迟的选择. 西安交通大学学报,2004,38(4):335 - 338

[5] Resenstein M T, Collins J J, Luca C. Reconstruction expansion as a geometry-based framework for choosing proper delay time. Physica D, 1994, 73:82 - 98

[6] 林嘉宇,王跃科,黄芝平,等. 语音信号相图重构中时间延迟的选择——复自相关法. 信号处理. 1999,15(3):220 - 225

[7] Thompson C, Mulpur A, Mehta V. Transition to chaos in acoustically driven flow. J Acoust Soc Am. 1991,90(4):2097 - 2103.

[8] 陈亮,张雄伟. 基于分形维数实现语音分割和增强. 北京邮电大学学报,2003,26(6):112 - 114

[9] 陈国,胡修林,张蕴玉,等. 基于短时分形维数的汉语语音自动分段技术研究. 通信学报,2000,21(10):7 - 13

(C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net