**基于混沌分形理论的脑卒中患者声音的非线性特性研究**



**论文作者签名:**

**指导教师签名:**

论文评阅人1：

评阅人2：

评阅人3：

评阅人4：

评阅人5：

答辩委员会主席：

委员1：

委员2：

委员3：

委员4：

委员5：

答辩日期：**英文论文题目**（16ptTime New Roman，Bold）



**Author’s signature:**

**Supervisor’ s signature:**

External Reviewers:

Examining Committee Chairperson:

Examining Committee Members:

Date of oral defence：

浙江大学研究生学位论文独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得**浙江大学**或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文作者签名：签字日期：年月日

学位论文版权使用授权书

本学位论文作者完全了解**浙江大学**有权保留并向国家有关部门或机构送交本论文的复印件和磁盘，允许论文被查阅和借阅。本人授权**浙江大学**可以将学位论文的全部或部分内容编入有关数据库进行检索和传播，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

（保密的学位论文在解密后适用本授权书）

学位论文作者签名：导师签名：

签字日期：年月日签字日期：年月日

# 致谢

最诚挚地感谢我的导师李江老师。

在我攻读浙江大学控制系硕士学位的三年时间里，李老师给予了极大的关爱和细心指导。在选择课题、研究方向、研究方法和设计实验方面，李老师给我提供了很多十分有益的指导和建议。导师渊博的知识、治学严谨和工作敬业的精神使我受益匪浅。李老师和蔼没有架子，跟我们在一起讨论问题，为我们营造了很好的学术氛围，给予我们最大的关怀和帮助

感谢浙江大学工业控制技术研究所给我提供一个舒适的学术研究环境；在这两年半的时间中让我们不仅在学业上学有所成，并且交到了一群志同道合的朋友。感谢给我授课过的所有控制系老师，他们独特的授课方式让我学到了扎实的专业知识，他们孜孜不倦的科研工作精神深深地影响着我。

另外，感谢实验室的师兄李海、李智鹏、张同心、马镇均、苏帅，师妹王晓倩、李晓玉。他们在学习和生活中给予我很多帮助，我们之间建立了厚重的同门情谊。

最后，特别感谢父母在生活和学习上给了我最无私的关爱、理解和支持。

感谢所有关心和支持我的人！

# 摘要

脑卒中是一类高发病率，高致死率的疾病。在预测脑卒中发生和脑卒中患者康复观察的过程中，没有很好的客观评价方法，只能通过医生的临床经验。所以本论文通过结合人体声音产生的生理学特点，利用非线性动力学方法分析声音时间序列，提取特征量分析脑卒中患者的大脑的损伤状态，尝试可以找到度量脑卒中患者大脑状态的特征量。为脑卒中患者康复及预防提供客观评价。

本论文开篇介绍了研究大脑状态的多种方式，比如脑电波，脑CT，声音等等。之后介绍了大脑和声音产生的生理学关系，从语音产生的神经机制和脑成像机制的角度对利用声音分析大脑状态提供了理论上的支撑。最后介绍了本文中用到的声音样本的来源及采集方式，并且介绍了最合适的采集音节。

针对声音时间序列的混沌特性，本文引入了非线性动力学来分析声音时间序列，首先阐述了非线性混沌理论和混沌的特征，之后介绍了非线性时间序列的相空间重构理论，分别用互信息法得到的时间延迟和用改进伪最近邻法得到的CAO方法得到的嵌入维数重构相空间和吸引子。最后用小数据法计算声音时间序列的最大Lyapunov指数，提取得到的这些混沌特征量都证明了声音时间序列具有混沌特性。

接下来又介绍了用于验证时间序列非线性特性的间接方法之替代数据法，介绍了用到的零假设及其相应的替代数据生成算法。通过计算得到声音时间序列和其替代数据在关联维数的方差上有很大差异，而正常人的此差异和中风病人的此差异存在明显不同。所以将替代数据法和关联维数相结合得到了新的非线性特征量即归一化方差检测量，这种特征量比单纯得用关联维数更好得显示了脑卒中病人声音时间序列和健康人声音时间序列的差异。

最后用混沌非线性方法和改进的替代数据法对所有的样本包括脑卒中病人的声音样本和健康人的声音样本提取非线性的特征量包括互信息图的第一个最小值，关联维数，最大Lyapunov指数以及归一化方差检测量，对这些特征量按照健康人和脑卒中病人进行统计对比分析。从对比图和表格中可以看出脑卒中病人的声音时间序列特征量明显区别与健康人的声音时间序列的参数。最后通过K近邻分类算法分别对两组非线性特征量的组合进行分类。从分类结果中得到在对健康人声音和脑卒中病人声音的分类方面归一化方差检测量明显优于关联维数。

通过以上对健康人和脑卒中病人的声音时间序列的分析和处理，可以看出通过声音确实可以反映大脑损伤的状态，这也为以后用声音时间序列分析度量大脑状态提供了研究方向和基础。

关键词：混沌，非线性，声音分析，脑卒中，替代数据法，相空间重构

# ABSTRACT

Through combining the physiological characteristics of the process of human vocal ,this paper uses the method of nonlinear dynamic to analyze the voice, extract relevant feature to analyze the state of the broken brain. Here choosing the stroke people to research. In this way, this paper try to find the feature that could measure the state of our brain.

This paper firstly introduces several method of measuring the state of brain, such as the brain waves etc. Then, the paper proposed a new method of measuring the state of broken brain by analyzing the voice time series. Secondly, the paper introduces the physiological relationship between brain and human vocal. Here provides information about the neural mechanisms of human vocal and the brain image to support using voice to analyze the state of brain. Lastly, this paper introduces the source of voice samples and the syllable of the voice.

For the chaotic characteristics of the voice time series, this paper uses the nonlinear dynamic to analyze this time series. Firstly, here introduces the nonlinear chaotic theory and the feature of chaos. Then, this paper introduces the theory of phase space reconstruction on nonlinear time series. Using the mutual information method to calculate the time delay and using the CAO method to calculate the embedded dimension to reconstruct space and attractor. Then using the G-P algorithm to get the correlation dimension and using the small data algorithm to get the first max Lyapunov index. All the mentioned chaotic features prove the chaotic characteristic of the voice time series.

Then, this paper introduces the surrogate data method which is a method to prove the nonlinear of a time seriesnormally. Introducing the zero hypothesis of the surrogate method and the surrogate data produced relying on different zero hypothesis. Then, combining the surrogate method and the correlation dimension to get a new feature—normalized sigma variability. This new feature gives a better description between the healthy people’s voice sample and the stroke people’s voice sample.

Lastly, using the nonlinear method to analyze all the samples which include the voice of stroke people and that of the healthy people.Then, calculating the first minimal value of the mutual information, the correlation dimension, the first max Lyapunov index and the normalized sigma variability of the healthy people’s voice and that of the stroke people. The author draws the comparison chart and table of voice samples. From comparison chart and table, we can see the obvious difference between the voice of healthy people and that of stroke people.

Through the analysis and discussion of the voice time series of healthy people and stroke people, the paper proves that the voice can really describe the state of our brain which also provide the direction and basement of using voice to measure the state of our brain.

Keywords: chaos, nonlinear, voice analysis, stroke, surrogate data method, phase space reconstruction

# 目录

[致谢 4](#_Toc379635231)

[摘要 5](#_Toc379635232)

[ABSTRACT 7](#_Toc379635233)

[目录 9](#_Toc379635234)

[图目录 11](#_Toc379635235)

[表目录 13](#_Toc379635236)

[第1章 绪论 14](#_Toc379635237)

[1.1 课题的研究背景及研究意义 14](#_Toc379635238)

[1.2 课题研究现状 15](#_Toc379635239)

[1.2.1 脑功能监测方法 15](#_Toc379635240)

[1.2.2 声音信号分析和特征提取方法 17](#_Toc379635241)

[1.2.3 声音分析诊断流程图 22](#_Toc379635242)

[1.2.4 国内外研究现状 23](#_Toc379635243)

[1.3 论文的主要内容及各章节简介 24](#_Toc379635244)

[第2章 脑与声音分析和脑卒中患者声样采集处理 26](#_Toc379635245)

[2.1 大脑对语言功能的影响 26](#_Toc379635246)

[2.1.1语言功能的神经机制[31] 26](#_Toc379635247)

[2.1.2 语言机制的脑功能成像[31] 28](#_Toc379635248)

[2.2 大脑的非线性 29](#_Toc379635249)

[2.3声样的采集 31](#_Toc379635250)

[2.4本课题的采集样本 33](#_Toc379635251)

[2.4.1 样本采集装置 33](#_Toc379635252)

[2.5 声音信号增强处理 35](#_Toc379635253)

[2.5.1 噪声的特性 36](#_Toc379635254)

[2.5.2 声音去噪方法之小波去噪 37](#_Toc379635255)

[2.6本章小结 39](#_Toc379635256)

[第3章 非线性时间序列分析及其在嗓音信号分析中的应用 41](#_Toc379635257)

[3.1 混沌理论概述 41](#_Toc379635258)

[3.2 相空间重构理论 44](#_Toc379635259)

[3.3 延迟时间的计算和应用 45](#_Toc379635260)

[3.4嵌入维数的计算和应用 50](#_Toc379635261)

[3.4.1 计算嵌入维数的方法 50](#_Toc379635262)

[3.4.2 时间序列计算嵌入维数 52](#_Toc379635263)

[3.4.3 根据延迟时间和嵌入维数重构相空间 53](#_Toc379635264)

[3.5李雅普诺夫指数的计算和应用 55](#_Toc379635265)

[3.5.1 李雅普诺夫指数的定义 55](#_Toc379635266)

[3.5.1 李雅普诺夫指数的计算 56](#_Toc379635267)

[3.6.2 时间序列的李雅普诺夫指数 59](#_Toc379635268)

[3.7本章小结 60](#_Toc379635269)

[第四章 改进的替代数据法在嗓音分析中的应用 61](#_Toc379635270)

[4.1 替代数据法 61](#_Toc379635271)

[4.2 关联维数的计算和应用 63](#_Toc379635272)

[4.2.1 计算关联维数 63](#_Toc379635273)

[4.2.2 对时间序列计算关联维数 65](#_Toc379635274)

[4.3 改进替代数据法得到归一化检测量 68](#_Toc379635275)

[4.3.1 基本概念 68](#_Toc379635276)

[4.3.2归一化方差的检测量的提出和计算 68](#_Toc379635277)

[4.3.3 声音时间序列的归一化方差的检测量计算 70](#_Toc379635278)

[4.4 本章小节 71](#_Toc379635279)

[第5章 脑卒中病人和健康人声音样本分析和分类 72](#_Toc379635280)

[5.1 非线性参数在脑卒中病人和健康人样本中的对比 72](#_Toc379635281)

[5.1.1 互信息图的第一个最小值对比 72](#_Toc379635282)

[5.1.2 重构相空间的吸引子对比 73](#_Toc379635283)

[5.1.3 最大Lyapunov指数对比 74](#_Toc379635284)

[5.1.4 归一化方差的检测量对比 75](#_Toc379635285)

[5.2 k近邻分类算法 76](#_Toc379635286)

[5.2.1 什么是K近邻算法 76](#_Toc379635287)

[5.2.2 k近邻的距离度量表示法 78](#_Toc379635288)

[5.2.3 K值的选择 78](#_Toc379635289)

[5.3 脑卒中病人和健康人声音样本分类 79](#_Toc379635290)

[5.4 本章小节 80](#_Toc379635291)

[第6章总结与展望 81](#_Toc379635292)

[6.1 研究工作总结 81](#_Toc379635293)

[6.2 进一步工作展望 82](#_Toc379635294)

[参考文献 83](#_Toc379635295)

[附录攻读硕士学位期间的研究成果 88](#_Toc379635296)

# 图目录

图2. 1大脑分区 28

图2. 2语言过程的脑激活模型 30

图2. 3ECM999的频率响应 34

图2. 4声音采集软件 36

图2. 5真实声音采集图 36

图2. 6语音增强流程框图 38

图2. 7 noisbump信号的小波阈值去噪图 39

图2. 8健康人声音信号的小波阈值去噪图 40

图3. 1非线性时间序列分析框架图 42

图3. 2考毕兹振荡器时间序列求取时间延迟 50

图3. 3健康人声音时间序列求取时间延迟 50

图3. 4考毕兹振荡器时间序列求取嵌入维数 53

图3. 5健康人声音时间序列求取嵌入维数 54

图3. 6考毕兹振荡器时间序列吸引子 55

图3. 7健康人声音时间序列吸引子 55

图3. 8计算Lyapunov指数流程图 59

图3. 9考毕兹振荡时间序列计算Lyapunov指数 60

图3. 10健康人声音时间序列计算Lyapunov指数 61

图4. 1求关联维数流程图 66

图4. 2考毕兹振荡器时间序列的关联维数 67

图4. 3考毕兹振荡器时间序列嵌入维数—关联维数图 67

图4. 4健康人声音时间序列的关联维数 68

图4. 5健康人声音时间序列嵌入维数—关联维数图 68

图4. 6洛伦兹混沌时间序列 69

图4. 7加入SNR=10白色噪声下的嵌入维数—关联维数图 70

图4. 8加入SNR=100白色噪声下的嵌入维数—关联维数图 70

图4. 9原数据和替代数据嵌入维数—关联维数图 71

图5. 1互信息图第一个最小值对比图 73

图5. 2互信息图的第一个最小值统计图 74

图5. 3二维空间吸引子对比图 75

图5. 4最大Lyapunov指数对比图 75

图5. 5 k近邻算法示意图 78

# 表目录

表2. 1采样脑卒中病人详细病历 33

表3. 1几种常见的运动性态的特征量 59

表5. 1脑卒中病人的替代数据分析表 75

表5. 2脑卒中病人的替代数据分析表 75

表5. 3 第一组特征量组合分类结果 79

表5. 4第一组特征量组合分类结果 80

# 第1章 绪论

## 1.1 课题的研究背景及研究意义

脑卒中是中医学对急性脑血管疾病的统称。脑卒中分为两种类型：缺血性脑卒中和出血性脑卒中。它是以猝然昏倒，不省人事，伴发口角歪斜、语言不利而出现[半身不遂](http://baike.baidu.com/view/112427.htm)为主要症状的一类疾病。由于本病发病率高、死亡率高、致残率高、复发率高以及并发症多的特点，所以医学界把它同[冠心病](http://baike.baidu.com/view/29208.htm)、[癌症](http://baike.baidu.com/view/3942.htm)并列为威胁人类健康的三大疾病之一。

因发病急骤，症见多端，病情变化迅速，与风之善行数变特点相似，故名中风、[卒中](http://baike.baidu.com/view/370849.htm)。本病常留有后遗症，发病年龄也趋向年轻化，因此，是威胁[人类](http://baike.baidu.com/view/14713.htm)生命和[生活](http://baike.baidu.com/view/5747.htm)质量的重大疾患。

由于大脑功能受损则会引起语言功能丧失或受损，至少三分之一以上的脑卒中患可产生各种言语障碍。中风患者度过危险期后，由于脑部血管的损伤，大多留有半身不遂、言语不利等后遗症。中风病人康复和观察过程中用到了一些简单的判断其发音质量的方法。而这些方法的应用以来于专家医师的主观意见。

声音的产生不仅仅同发声器有关，很大程度上跟人脑的中枢神经系统和大脑皮质也有关联[1]，所以大脑功能受损往往会导致言语障碍。脑卒中病人中有很多因为大脑中部分功能受损或者障碍而导致言语不利，失语等症状也证明了这一点。很多学者研究表明[2,3]，人分别在疲劳状态、紧张状态、正常状态下发出的语音信号各不相同，从很大程度上证明了大脑的状态对人所发出的声音有很大影响，而脑卒中病人也正是这种情况下的极端情况。

声音分析的手段也是一种非侵入性，无创的辅助诊断方法，减轻了病人的痛苦，设备简单，速度快，价格低。在病人康复过程中，可以作为一种有效辅助手段评估病情。

近年来信号的时频分析等处理技术为客观化语音提供了基础，而随着现代非平稳信号处理技术的进步，改变了过去人为的将时间序列假定平稳或分段平稳化的做法，使得在处理信号时更贴近于信号自身的真实物理特性，非线性时间序列尤其是混沌时间序列分析随之提出并且得到广泛应用。声学及空气动力学理论证明，语音信号时一个复杂的非线性过程，其中也存在着产生混沌的机制，所以引入用非线性方法来分析声音信号得到结果的实际应用性也更强。

1985年，Babloyantz等人将非线性动力学和混沌理论应用到EEG信号分析中，并得到了人类睡眠脑电是混沌信号这一结论[4]。随着对大脑功能研究的不断深入，大量科学研究已经证明脑电信号起源于高度的非线性系统，脑电活动具有确定性的混沌特征，人脑是一个复杂的自组织非线性动力学系统。

因此利用非线性动力学方法进行声音信号分析，可以更好的反映大脑的复杂水平，帮助我们更加深入的了解大脑的混沌特性。本文将脑卒中病人的声音分析同大脑的损伤程度结合而对其康复程度和大脑状态进行深入探讨，这不仅为我们揭开大脑的混沌特性，而且也为脑卒中病人康复程度提供参考。

## 1.2 课题研究现状

### 脑功能监测方法

人类最初研究大脑的方法是通过解剖方式来获取大脑的结构和功能知识，然而，由于这种方法对大脑的极大损害性和对医师经验的高要求性，导致许多更进一步的知识难以获得。随着一些辅助设备的发明，如脑电仪(EEG)、核磁共振仪(MRJ)、计算机断层成像仪(CT)、脑磁仪(MEG)，正电子发射断层扫描仪(PET)等，为进行无损伤地测量人脑内部结构、诊断病变区域提供了新的手段。尤其是脑电的发现，使得人们对大脑的研究变得更加方便和经济。

脑卒中病人临床诊断多数是侵入性的诊断手段，而近年来也出现了很多非侵入性的分析手段，例如对脑电信号(EEG)和声音信号的分析辅助诊断脑功能损伤状况也取得了很多阶段性成果。

1. 临床诊断

(1)脑CT扫描：脑CT对出血性脑卒中诊断率为100%，对缺血性脑卒中的诊断率在85%以上。但需要注意的是缺血性脑卒中在发病24小时以后才能显示清楚。所以如果不是十分必要，最好等脑卒中发病24小时后再查CT。

(2)血糖、尿糖测定：脑卒中病人的治疗和预后，与其有无[糖尿病](http://www.baiyun120.com/html/1tnb/)关系很大。脑卒中急性期空腹血糖超过11.12mmol/L，即提示其下丘脑损害严重，死亡率较高。所以，即使以前没有[糖尿病](http://www.baiyun120.com/html/1tnb/)病史，也应监[测血糖](http://www.baiyun120.com/html/gaoxuetang/)和尿糖。血糖尿糖的高低是医生确定治疗方案的重要依据。这是因为中风以后可以出现一过性的血糖升高，况且有的病人对自己过去有无糖尿病并不清楚。

(3)数字减影血管造影(DSA)：对缺血性血管病、动脉瘤、动一静脉畸形、烟雾病的诊断有重要意义。虽然是一种有创性检查，但在直观显示血管结构的同时，还可以进行介人性治疗，估计以后随着医学的发展，DSA会更加普及。

(4)腰穿脑脊液检查：腰穿检查脑脊液对于明确中风的性质极具有重要的意义。但随着更先进的检查手段如CT、MRI等的应用，脑脊液检查就不再成为脑卒中检测的必要诊断手段了。

(5)磁共振(MRI)：MRI 与CT相比，MRI无X线的辐射效应，因此对人体没有明显损害，且图像层次清楚，分辨率高。其缺点，一是价格昂贵，相应的检查费用也高;二是安装心脏起搏器的病人不宜使用;三是成像时间长，危重病人粮难接受长达40分钟的头部扫描。鉴于以上原因，我们认为，MRI对于急性中风病人并非首选的检查项目。

1. 新型诊断方法
2. 脑电图(EEG)

各种各样的生物均存在着生命活动的基本特征即生物电活动，人体中同样广泛存在着生物电现象。人身体上都有磁场，当人在思考的时候，大脑中的磁场会发生改变，就会形成一种生物电流，我们把它称为“脑电波”，我们的大脑无时无刻不在产生脑电波。脑电图（electroencephalogram, EEG）是通过电极记录下来的脑细胞群中的这种脑电波绘制而成。脑电图检查对许多大脑损伤和神经系统疾病的诊断、病情监测及疗效观察有十分重要的意义。有时其它检查(如CT)没有发现异常，而脑电图却发现有异常表现。因此，脑电图检查己广泛应用于多种疾病。

脑电图分析最先应用于癫痫病得分析诊断中，癫痫脑电特征波主要指的是棘波和尖波(有时把它们统称为棘波)，棘波的识别与提取研究已经发展了近三十年，其间各国的学者提出了很多处理方法。之后清华大学的李双双等人对新生猪脑损伤的振幅整合脑电图分析模拟新生儿的脑部状况也取得了不错的效果。使振幅整合脑电图对脑功能的评价更加直观。

1. 声音分析

大脑是一个复杂的神经系统，而其中的语言中枢神经和运动神经都控制着声音的产生。随着应用基础学科的发展，计算机及相应的各类专业软件的发展，为正常及病理性声音特征的多维分析提供了坚实的基础。声音学参数:包括基频(F0)、频率微扰(Jitter)、振幅微扰(Shimmer)、规范化噪声能量(NNE)、基频标准差(SDF0)、共振峰(Formant) 、声强等多种参数[5,6]，它们在病态声音检测中各具有不同的作用。同时除了这些基本的声音学参数之外，对声音信号进行小波变换，谱分析以及对声音信号进行非线性分析都得到了更多的分析声音信号的参数[7-9]。

现今应用于嗓音检查的主要仪器可分为3类：一是语图仪,二是窄带频谱分析仪，第三类则具有语图及频谱的双重功能。每类仪器内含有电子计算机装置，可进行快速的A/D转换和傅立叶变换。计算机技术及统计模式识别的迅猛发展，建立在模型统计基础上的病态嗓音自动检测也迅速成为了研究热点，经过几十年的发展，产生了很多的声学分析、参数与非参数特征提取、自动模式识别与统计方法。

一直以来，声音信号分析技术并没有得到更好的医学应用，原因是声音信号是一种非平稳信号(所谓的非平稳信号，是指信号的统计特性，包括时域统计特性和频域统计特性，是随着时间的变化而变化的)，如果应用常规的信号分析方法进行分析，提取声音信号特征将会遇到很大的难度，从而影响声音诊断的效果。如何在现有的声音信号采集技术上，对所采集到的声音信号进行恰当的信号分析处理，使得处理过后的声音信号能够明显地代表大脑状态的特征信息，是目前声音分析技术中急需要解决的一个难题。从某种意义上说，声音信号分析技术的发展受到了非平稳信号处理技术阻碍。当前，还原声音信号产生的物理特征，找寻一个先进的非平稳信号处理技术成为声音信号分析技术亟待解决的问题，这同时也是本论文研究内容之一。

### 声音信号分析和特征提取方法

对采集到的脑卒中病人的声音信号进行分析和处理，最基本的方法就是由有经验的专家直接观察，但是这个要具备丰富的领域知识和临床经验，而且主观性强。

在计算机技术和语音信号分析的不断发展下，许多自动分析和可视化方法为分析识别声音信号提供了良好的辅助工具。根据目前的研究成果，声音信号分为线性分析和非线性分析，如图1.1所示。



图1. 1声音信号分析方法

下面介绍声音分析中的几类方法：时域频域分析、谱分析、非线性分析。这些方法从不同角度研究声音信号。

1. 时频分析

时频分析即时频联合域分析（Joint Time-Frequency Analysis）的简称，作为分析时变非平稳信号的有力工具，成为现代信号处理研究的一个热点，它作为一种新兴的信号处理方法，近年来在各个信号分析领域得到越来越多的应用。时频分析方法提供了时间域与频率域的联合分布信息，清楚地描述了信号频率随时间变化的关系。时频分析的基本思想是：设计时间和频率的联合函数，用它同时描述信号在不同时间和频率的能量密度或强度。时频新分析按照导出方式主要分为四类：一是直接对傅里叶变换的基函数进行变换改造，如小波变换。二是先由信号得到一个双线性函数再进行傅里叶变换，如Winger-Ville分布等。三是对加窗信号进行傅里叶变换，如短时傅里叶变换(STFT)等。四是自适应参数时频分析方法，诸如匹配追踪算法等。一般在声音分析应用中短时傅里叶变换和小波变换应用最广泛[10,11]。

短时傅里叶变换的基本思想[12,13]是利用一个同等大小的滑动的窗函数对信号进行分析，并假定信号在窗函数以内是平稳的，以达到时域上的平稳性，这是一种线性的时频分析工具，避免了使用二次型分析工具时交叉项的干扰。传统的短时傅里叶变换[14]可由式1.1表示：

(1.1)

式中：为待分析的信号，为窗函数时域表示。

将式(1)的时域卷积用频域滤波器组来实现，可以将其改写为式1.2的形式：

(1.2)

式中：为待分析信号的傅里叶变换，为时间窗的傅里叶变换。

由式(2)表明短时傅里叶变换可以理解为信号的加窗谱的傅里叶反变换，解释为信号通过频率响应为的滤波器的输出乘以得到的结果。由式(1)、(2)分别确定了位于时域和频域的2个窗口，中心分别为和，窗口的宽度为和，由不同的窗函数来决定。可由式1.3表示为：

(1.3)

其中和表示了时频分析中的分辨率，从上面的分析可知，窗口越窄则分辨率越高。理想的情况是希望和都非常小，但根据海森堡(Heisenberg)不确定性原理[15]，和是互相制约的，不能同时任意小。

虽然短时傅里叶变换在一定程度上解决了标准傅里叶变换所不具有的局部分析能力，但是短时傅里叶变化也存在本身的局限。短时傅里叶变换是一种加窗变换，而窗函数一旦确定了，时间分辨率和频率分辨率也就确定了，，这意味着短时傅里叶变换是一种单一分辨率的分析方法，如果要改变分辨率的大小，则务必要选择一个新的窗函数。同时傅里叶变换也受到不确定性准则的约束，加窗变换同有的混叠现象必然也会在短时傅里叶变换中存在。

因为短时傅里叶变换的局限性，所以出现了小波变换[16]，为非平稳信号分析提供了很好的前景。小波变换是一种新的变换分析方法[17]，它继承和发展了短时傅立叶变换局部化的思想，同时又克服了窗口大小不随频率变化等缺点，能够提供一个随频率改变的时间—频率窗口，是进行信号时频分析和处理的理想工具。它的主要特点是通过变换能够充分突出问题某些方面的特征，因此，小波变换在许多领域都得到了成功的应用，特别是小波变换的离散数字算法已被广泛用于许多问题的变换研究中[18-20]。从此，小波变换越来越引起人们的重视，其应用领域来越来越广泛。

小波分析是时间—[尺度分析](http://baike.baidu.com/view/3814867.htm)和[多分辨分析](http://baike.baidu.com/view/207823.htm)的一种新技术，所谓多分辨率分析是指原始信号按一定的尺度被分解成几个具有不同分辨率的分量，而且由这些分量能够不失真地重建原始信号。所有闭子空间都是由同一个尺度函数伸缩后的平移系列张成的尺度空间。根据多分辨率分析的基本思想，利用尺度函数与小波函数对信号进行分解。

虽然小波分析采用可变窗口对信号进行分析，弥补了短时傅里叶变换的时间和频率分辨率的缺陷。但是小波分析实际上也是窗口可调的傅里叶变换，并没有从根本上解决傅里叶变换的局限问题[21]。再者，小波变换要根据不同的分析信号选择不同的小波基函数，所以小波变换的自适应性较差。

时频分析的应用是多种多样的，除了STFT、小波变换等线性时频分析方法之外，还包括诸如Wigner分布和Choi-Williams分布等一些常用的非线性视频分析方法。以Wigner分布及其改进型分布为代表的时频分析方法在一定程度上解决了STFT的不足，并且具有较高的时频分辨率以及其他优良特性。但是它们本身都有二次型交叉项的局限问题。

1. 谱分析

谱分析方法分为以傅里叶分析为基础的经典谱分析方法和现代谱分析方法（也成为参数模型法）。传统谱分析法中人们研究出两种基于傅里叶变换运算的谱估计技术。一种是间接的方法，它对自相关函数进行估计（一般都需要窗函数将自相关值加权，以减小自相关序列截断的影响），然后再作傅里叶变换得到功率谱估计值。这种谱估计法是1958年由Blackman和Tukey推出的，简称为BTPSD估计法。另一种是直接法，它是将观察到的有限个样本数据利用FFT算法作傅里叶变换再进行功率谱估计（不通过自相关函数的估计），这种方法称为周期图的方法。这种用周期图（包括平滑后的周期图）作为功率谱估计的方法可利用FFT进行计算，因而有计算效率高的优点，在谱分辨率要求不高的地方常用这种周期图法进行谱估计。周期图的方法对于样本，特别是短样本，求得离散谱的精度不高，而且这种方法计算出来的功率谱只是真实功率谱的渐近无偏估计，不是一致估计，样本增大，方差并不同步减小，所以传统的经典谱估计是有着其原理上的缺陷。

现代谱估计的提出主要是针对经典谱估计（周期图和自相关法）的分辨率和方差性能不好的问题。1967 年，Burg提出的最大熵谱估计，即是朝着高分辨率谱估计所作的最有意义的努力。现代谱估计法是以参数模型为基础的方法，大致可以分为参数模型谱估计和非参数模型谱估计，前者有自回归模型(AR)、滑动平均模型(MA)和自回归滑动平均模型(ARMA)等；后者有最小方差方法、多分量的MUSIC方法等。

参数模型法是主要方法，它把具有多个变量的复杂过程简化为只用少量参数就可以表示的简单过程。有理参数模型用有理系统函数表示，求出连续的功率谱，即有理谱。参数谱分析是建立在时序模型参数和阶次的估计上，对数据长度的要求不高，对短时序序列的谱估计更有效，更有利于对颤振信号的实时处理。另外，参数谱也是一种连续谱，谱线光滑，谱峰清晰，谱图的频率分辨率可以更高，且直观性好。由于具有这一系列良好的性能，因此被研究最多也得到最广泛的应用。

以参数模型为基础的谱估计方法一般按下列3步进行：

(1)选择合适的信号模型。

(2)根据已知的有限个观测数据，或者它的有限个自相关函数估计值，估计模型的参数；

(3)用估得的模型参数计算输出功率谱。

参数模型谱估计方法利用信号的信息对被窗函数截取的有限信号以外的信息进行预测或外推，提高了谱估计的分辨率和真实程度，故其也称为高分辨率谱估计。

非参数模型中应用最广泛的是空间谱分析，空间谱分析采用了类似时域谱估计中的非线性处理，从而产生了一些特殊的算法，如多重信号分类法(MUSIC)、旋转不变技术的参数估计法(ESPRIT)、最小内积法(JVW)、投影矩阵法和矩阵分解法等。这些方法都巧妙地利用接收信号的相关矩阵的特征结构，其中，MUSIC算法将阵列输出数据的自相关矩阵进行特征分解，然后利用这两个子空间的正交性来估计信号的参数。这一算法的提出开创了空间谱估计算法研究的新时代，成为空间谱估计理论体系中的标志性算法。采用这种空间谱估计技术不仅可以提高测向精度和分辨力，而且加快了信号处理的运算速度。

1. 非线性分析

非线性分析方法与前两种分析方法的出发点完全不一样，前两种分析方法均假设声音信号是随即非平稳信号，都是线性分析方法。而非线性分析方法假设声音信号是某个未知的非线性系统的一个变量或多个变量组合的输出信号，声音信号所表现出来的随机性是由于附加在上面的其他信号所引起的，声音信号背后的系统本身是确定的。而声学和空气动力学理论表明，声音信号是一个复杂的非线性过程，其中存在着产生混沌的机制[22,23]。混沌是指发生在确定性系统中的貌似随即的不规则运动。一个确定性混沌描述的系统，其行为却表现为不确定性，即不可重复和不可预测，这就是混沌现象。混沌对初始状态十分敏感，是非线性动力系统的固有特性与普遍存在的现象。对于非线性确定的系统所产生的信号，用线性的方法进行分析显然是不能达到最终目的，甚至可能产生一些错误的解释。然而，对于一个未知系统的时间序列，在用非线性分析方法分析之前，必须知道该时间序列是否是确定的，因此非线性分析方法一般分成两个阶段：首先检验系统的非线性确定性，然而进行非线性分析。

近年来，生物医学中的心电、脑电、心音等信号都存在着混沌特征，被证实为混沌信号，混沌理论被广泛应用于这些信号的分析处理中[24-27]。美国和加拿大的研究人员以混沌理论为工具，通过对心电图的定量分析和测量发现，正常和病态的心脏都有着显著的动力学特征，他们尝试用非线性动力学的混沌分形方法进行研究，表明传统的心脏病学对于不规则的心搏做出了不正确的概括，从而对传统的医学原理提出了挑战。混沌理论的最新研究表明，人脑的神经网络是一个混沌系统，它对信息的处理过程是一种混沌现象。Soong等人研究了人脑的α节律，表明α节律是由确定性动力系统决定的混沌运动[28,29]。清醒时，大脑的混沌程度较高，信息处理较快，并能做出更多的响应。另一方面，癫痫病人在癫痫发作期间的脑电图有周期性，如果用接近于无规则脑电波的声波刺激病人的大脑，病人的脑电图会恢复到健康的混沌状态。混沌和分形理论还可以用于模拟生物医学工程领域的相似结构，例如循环系统中的血管-毛细血管网以及呼吸系统中的气管—支气管网的形成。混沌理论还应用于对呼吸机能障碍的研究以及对红细胞和白细胞浓度控制的反馈机制研究。此外，研究还表明，生物反馈系统、细胞代谢、血小板生成系统以及免疫系统中都存在着混沌现象。当人衰老时，这些系统的混沌程度降低，从而呈现简单的规则变化。

### 声音分析诊断流程图

基于声音分析的脑功能损伤病人分析诊断流程图如图1.2所示。脑损伤病人发出声音（一般为元音），通过声音采集设备采集声音信号转换为数字信号记录到电脑上，对这些数字信号进行语音加强处理，去除多余噪声，再经过特征提取按照脑损伤和脑部正常进行分类，并将分类结果显示出来。



图1. 2信号分析诊断流程图

### 国内外研究现状

国内外研究人员通过声音分析对大脑状态进行评估始于21世纪初，普遍认为人体大脑的状态在一定程度上改变了所发出声音信号的非线性参数，从而开启了应用声音分析这种非侵入的手段研究大脑状态的篇章。

波兰学者Damian.K等人对脑卒中病人的声音进行了线性分析[1]，从功率谱和共振峰频谱上发现了频谱随脑卒中病人康复程度的变化。研究发现脑卒中病人和健康人发出的声音的基音的语谱图和共振峰的语谱图有所区别。然后又对脑卒中病人嗓音进行FFT变换，MUSIC分析，时频分析等。证明根据病人健康程度的不同基本特性发生变化，这些依赖同脑干损伤程度成比例。在MUSIC分析的频谱图（功率-频率）中健康人声音分析的曲线轮廓较缓和，脑卒中病人声音频谱图中的曲线则有较大幅度变化。而随着病人住院时间的增长，曲线轮廓慢慢变缓和。从中也证实了语音的生成不仅仅由发生器直接作用，还有很大一部分程度上取决于中枢神经的控制，分布于大脑皮层，脑干的神经及其神经突触所受到的刺激。

日本学者Kakuichi.S等人通过声音对人的大脑活动进行度量[2]，分析了人在疲劳时候发音和非疲劳时发音的对比，他们的研究指出人类发出声音的第一李亚普诺夫指数会随着人的状态发生变化，但是第一李亚普诺夫指数不足以评估人类声音的时域性质。之后提出了脑指数的概念，脑指数用来评估具有混沌特质和其功率谱有一个顶峰的时域信号的特性，比如说人类的声音。并且Kakuichi.S设计了一套算法叫SiCECA（shiomi’s cerebral exponent calculation algorithm）来计算脑指数。

## 1.3 论文的主要内容及各章节简介

经过上面几节的分析，本文选择用声音分析来反映脑卒中病人脑损伤的状态，即将脑卒中病人发出的声音作为脑损伤状态的检测量。脑卒中病人声音分析包括几个方面：脑卒中病人声音采集、脑卒中病人声音信号分析处理、故障诊断特征量的构造和结果分析。其中脑卒中病人声音信号处理分析和故障诊断特征量的构造是有待解决的核心问题。如何能够提取出相应的反映大脑损伤状态的特征量是反映脑卒中病人脑损伤状况的研究热点。

本文撇弃了国内外通过脑电波来研究大脑状态的方法，提出了用语音分析大脑状态，并且通过采集脑卒中病人声音，脑卒中病人中因为脑损伤普遍引起了失语症和言语障碍，通过分析处理这类声音信号更能够真实反映出大脑状态。主要的工作如下：

1)研究了脑卒中病人声音采集最佳样本

脑卒中病人中严重者，有些已经不能清楚得说出句子，所以本文采集了脑卒中病人的单个音节的发音。在选取哪些音节时，笔者通过各种文献研究得到元音/a/更加适合作为采集样本。

2)研究了声音信号的非线性处理方法

脑卒中病人声音信号分析是大脑状态诊断的关键之一，目前声音信号分析很多处理方法是线性的例如短时傅里叶变换和小波变换。但是声音信号本身是一种非线性信号，所以通过非线性的分析更加能够反映出声音信号本质的特性。本文采用的信号分析手段就是对声音信号进行非线性分析处理，提取声音非线性时间序列的基本特征量，例如关联维数、最大Lyapunov指数、互信息图第一个最小值等进行统计分析。

3)研究了非线性特征量和替代数据结合的新的声音分析特征量

提出了一种用非线性特征量——关联维数和替代数据相结合得到的一种归一化方差检测量。这种检测量经过证实相比于关联维数能够更好得反映脑卒中病人声音时间序列和健康人声音时间序列之间的差别。

全文具体章节安排如下：

第一章绪论。首先简明阐述了本篇论文的研究背景和国内外对人体大脑状态研究的现状；然后介绍了对大脑状态诊断分析的常用方法；之后介绍了声音分析常用的线性及非线性方法；最后说明了声音的产生受大脑控制并且能够用来反映大脑的状态，由此引出了本文所做的工作及研究意义。

第二章脑与声音。本章中重点讨论了大脑对声音产生的影响机制。先从语音功能的神经机制方面阐述了大脑控制声音的产生，再从语言功能的脑成像方面说明了声音产生同大脑的关系。之后介绍了本文分析的声音时间序列的来源和采集方法，最后介绍了采样声音样本的预处理去噪等。

第三章非线性时间序列分析及其在嗓音信号分析中的应用。首先介绍了非线性混沌的基本概念和混沌信号的主要特征。之后介绍了重构相空间理论以及重构相空间必须的两个参数即时间延迟和嵌入维数，建议使用互信息法计算时间延迟和CAO方法计算嵌入维数。并且用这些方法对考毕兹振荡器混沌电路产生的混沌时间序列进行计算并且重构相空间，画出了三维空间中的吸引子，验证了方法的可靠性。接下来又介绍了最大Lyapunov指数等非线性特征量的计算方法。

第四章替代数据法在嗓音分析中的应用。这一章介绍了替代数据法这种检验时间序列非线性的间接检测方法，介绍了几种基本的零假设及其替代数据产生方式。之后又将替代数据法同非线性基本特征量之关联维数结合在一起提出了新的特征量之归一化方差检测量。

第五章脑卒中病人和健康人声音样本分析。本章中应用第三章和第四章的非线性分析方法对采集到的所有的脑卒中病人声音信号和健康人声音信号提取了非线性基本特征量，并且绘制了对比统计图。将所有的声音时间序列的归一化特征量、关联维数等绘制成了对比表格。最后将得到的特征量组合对两类声音进行分类。

第六章总结和展望。对本文的所有研究工作进行总结，并且进一步研究的内容和方向。

# 第2章 脑与声音分析和脑卒中患者声样采集处理

## 2.1 大脑对语言功能的影响

语言（language）是人类用来交流思想、情感和需要的按一定规则组成的符号系统。一百多年前，学者对脑损伤语言障碍患者的研究就提出了“我们用左脑说话”（Broca, 1865）。随后众多言语和读写障碍的研究进一步探索了语言障碍和损伤脑区的关系，提出了负责语言功能的关键脑区[30]。近年来脑功能成像技术的出现，更加促进了语言神经机制研究的发展。目前研究不仅深入揭示了左半球语言网络的分布和作用性质，而且丰富了右半球在语言中的作用，并对人类不同语言之间脑机制的共性和个性进行广泛探索。

### 2.1.1语言功能的神经机制[31]

人们要进行正常的言语交流需要与之相关的脑神经环路的功能正常，各种累及重要言语中枢的脑损害都会导致言语功能异常，如说话不流利、语法功能受损；内容杂乱、理解缺陷；复述障碍；命名不能等，从而形成形形色色的失语症（即获得性脑损伤后原有口头或书面语言表达和/或理解功能受损或丧失）。脑损伤－语言障碍的研究模式无疑为揭示正常言语功能的神经机制做出了巨大贡献。

言语的产生是指把思想、意图、情感或知觉等用说话即语音的形式表达出来的过程。言语的产生需要以下几个过程。首先，确定要说的内容，这些概念性信息主要由后部大脑皮质即枕叶、颞叶、顶叶参与；其次，把这些信息转化为对应的词句，要选择适当的词语、遵循合理的句法组合成句子；然后，形成把这些词句进行语音输出的计划；最后执行这个计划，即通过发音的过程把词句说出来。言语产生的关键过程主要借助左侧额叶得以完成[32]。

1861年，法国医生、解剖学家Paul Broca发现1例严重的言语表达障碍患者，该患者唯一能说的话仅是“tan”这个音。尸解发现损伤病灶在左半球的额下回。结合更多的类似病例研究，Broca在1865年向全世界宣布“我们用左脑说话”。后人把左额下回后部相当于Brodmann分区的（一种根据细胞构筑特点进行大脑皮质分区的经典方法，Brodmann为德国神经病学家）第44、45区，称为Broca区，见图2.1。

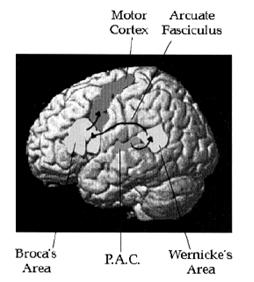
[](http://photo.blog.sina.com.cn/showpic.html#blogid=a59e41c4010159t3&url=http://s8.sinaimg.cn/orignal/4a8a448ena0ceab33a577)

图2. 1大脑分区

    Broca区周围损伤患者常会导致Broca失语，也称为表达性失语（expressive aphasia）或运动性失语(motor aphasia)。关键症状就是言语产生障碍：言语不流利，发音困难、说话费力、速度缓慢、有发音错误即语音性错语（phonemic paraphasia）（如把“电灯”说成“电当”）。

Broca失语患者发音困难、费力、错语等和Broca区及周围神经环路的作用紧密相关。Broca区储存了发音必需的有关肌肉运动程序或顺序的记忆，包括控制舌头、口唇、下颌以及声带等发音器官的肌肉运动程序。发音过程中要将语言运动信息转变为运动冲动、经锥体束至运动神经核团，支配构音器官。同时椎体外系也有纤维支配这些核团从而影响控制发音肌肉的肌张力和共济运动，以保证声音的音调和音色。这种运动程序必需有序和协调地传向初级运动皮质的口面部对应区，从而发放下行冲动，通过外周神经支配发音器官协调运动，完成言语产生过程如图2.1所示。如果Broca区损害，这种快速、有序、协调的发音运动被破坏，就会出现发音困难、发音错误等言语障碍。倘若脑损害仅导致患者出现口语障碍，而听理解、读写、智力等正常，则称为纯词哑（pure word dumbness），或称言语失用（apraxia of speech），即无法产生快速、有序、协调的发音运动导致的单纯性言语障碍。

一般认为，失语症的发生与大脑的Broca区、Wernicke区以及连接两者的弓状束及角回周围脑区有明显关系。其中人脑的Broca区负责语音产生，Wernicke区负责语音感知和理解，Broca和Wernicke区域又是相互连接[33]。如果你损伤了Broca区，那么语音产生受到严重损伤，但具有良好的语言理解。如果你损伤了Wernicke区，那么患者具有非常差的语音理解，但有良好的语音产生能力。因此，没有Broca区域，就不能产生语音；没有Wernicke区域，就不能理解语音。除上述的大脑皮质语言区损伤导致失语外，单纯的皮质下结构如基底节区（nucleus basalis）、内囊（internal capsule）、丘脑（thalamus）等损伤也会导致失语症，统称为皮质下失语症（subcortical aphasia）。

### 2.1.2 语言机制的脑功能成像[31]

上一节中从语言障碍角度谈论了语言功能的神经机制。这一节将从脑功能成像研究出发，简单介绍新近的语言神经机制的研究。

Cathy J. Price小组利用PET（positron emission tomography，即正电子发射计算机断层显像）对各种语言功能进行系列研究。发现了和失语症研究结果类似的语言脑区。例如，听觉词汇首先到左颞上回的听皮质加工，然后传送到Wernicke区进行语音识别。视觉词汇在舌回和梭状回后部视皮质加工后，传到左颞枕交界腹侧的梭状回中部进行词形识别。视、听词汇识别后传向广泛的语义加工区如左角回、左颞中/下回、前颞叶等。如果进行言语产生如复述或朗读，则由Broca区和左前岛叶制定发音计划，最后由口面部对应的运动感觉区皮质执行如图2.2所示。 随着技术的发现，现在人们还可以利用功能磁共振成像对脑部进行扫描，侦测出大脑磁场中产生的无线电频信号，以呈现血量、血流速以及氧消耗量等大脑神经活动的非直接结果，并画出大脑解剖图以指示大脑活动的变化。可见，脑功能成像技术的发展将为进一步揭示语言的神经机制做出独特贡献。

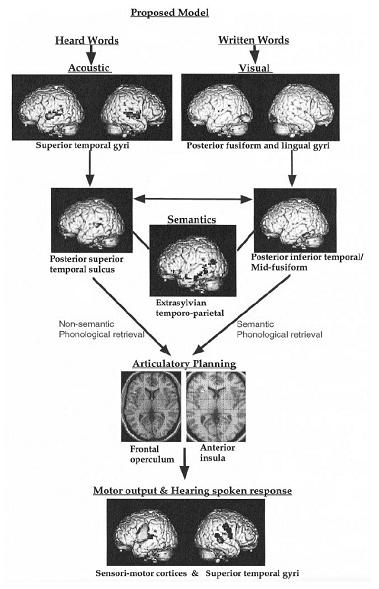
[](http://photo.blog.sina.com.cn/showpic.html#blogid=a59e41c4010159t7&url=http://s9.sinaimg.cn/orignal/4a8a448et76ba08b33858)

图2. 2语言过程的脑激活模型

 脑损伤语言障碍的研究以及正常人语言的脑功能成像研究，揭示了语言优势半球相关的神经网络，而且同时也提示我们非优势半球如右半球仍具有语言作用，左右半球互相协调才能完成正常的语言交流。脑卒中病人的脑损伤在一定程度上都影响到了其语言表达，所以通过病人的语音信号更能很好得反映其脑损伤情况。

## 2.2 大脑的非线性

人类大脑是由巨量的神经元组成，大脑与外界的信息交流具有很大的不确定性，大脑内部的生存状态也具有不确定性，这一切都是大脑神经系统注定是一个非线性动力系统。

大脑是人体内功能和结构最复杂的器官，大脑就像一个“黑匣子”一样摆在人类面前，它如何能支配人类进行各类思维活动，如何产生各种思想，如何认知外部世界，它的工作原理是什么却知之甚少。多少年来，人们从神经解剖、神经生理、神经病理、神经免疫、神经生化以及神经影像学等方面对大脑进行研究和探索，取得了不少关于动物大脑和人类大脑的实验数据和研究成果。20世纪90年代，神经科学、心理学、生物医学等领域出现了一场脑研究的热潮，这十年因此被称为“脑的十年”。对大脑的探索，对人类自身的认识是人类永恒的课题和挑战。

从系统论的角度看，大脑认知过程绝对不是单一和静止的，而是动态的，演变的，非线性的和混沌的，不同时点、不同脑区、不同频段的工作环路的暂态、同步、整合等机制构成了一个复杂的大脑认知功能系统。比利时著名化学家、物理学家伊利亚·普里高津曾说过，“众所周知，心脏在很大程度上是规则的，否则你必死无疑。但大脑却近乎是〖HTH〗不规则的，不然你就患有癫痫。这表明不规则性、混沌性导致了复杂系统，它并不是全然的无秩序。相反，我要说正是混沌使生命和智慧成为可能。大脑已被选择得如此不稳定，极小的作用就会导致秩序的形成。”也就是说，大脑是人类非线性进化的非线性产物。

早在1987年，《科学美国人》杂志的一篇文章总结了目前神经生理学关于记忆的研究成果。该文报告了神经科学家们在具有相互联系反馈环的六个脑区（感觉区、杏仁、海马、间脑、前额叶皮质和基前脑）描记视觉记忆通路。这是在整个大脑许多尺度上都存在的非线性类型的一个大比例尺图解。

近年来，许多研究人员收集了大脑是一个非线性反馈装置的实验证据，取得了一定的成果。

加州大学洛杉矶分校的研究人员唐·沃尔特(Don Walter)和艾伦·加芬科(Alan Garfinkel)建立了模拟神经元发放模式的方程式。在他们的模型中，把三个神经元联在一起，产生了具有隐秩序的低水平神经混沌的证据。

加州大学伯克利分校的沃尔特·弗里曼(Walter Freeman)和克里斯廷·斯卡达(Christine S karda)，通过在动物的大脑上做的研究，说明了这种神经元混沌发放是如何转换为秩序的。两位科学家将多达64个微电极植入兔子嗅球中，当给兔子嗅几个分子的不同气味时，监测脑电波模式。两位研究人员发现，一种气味被检测出时，这个闻嗅的部分大脑中的低水平浑沌顷刻自组织，也就是说，个体神经元的整个嗅球发放以集体方式耦合在一起。

普里高津曾指出的，如果大脑秩序长时期过于规则，就麻烦了。斯坦福大学一位神经科学家罗伊·金(Roy King)，通过一种叫多巴胺的神经递质与精神分裂症症状(如幻觉和思维紊乱)之间的联系，描述了这个麻烦。金和斯坦福大学的同事们用关于多巴胺活动的已知数据建立了一个数学模型，把它放在计算机上试验。这个模型表明，精神分裂症的症结在于大脑中多巴胺释放的速率。该研究提出对于大脑，浑沌完全正常，但由过多秩序引发的混沌是有害的。对于精神分裂症患者来说，他们发病时源起于大脑过多的秩序——受陷秩序——引发了过多的混沌，使病人遭受混沌的大规模攻击。

大脑精致地维持着其混沌和秩序之间的平衡，给人们带来很大启发。随着科技发展，脑电图给科学家们研究大脑提供了方便的工具。科学家们详细地分析患者脑电图上那些极不规则的图形。一些研究人员根据这些脑电图寻找大脑的奇怪吸引子。布鲁塞尔自由大学的巴布洛延茨(A.Babloyantz)发现这些复杂的脑电图与分形很相象，于是决定测量睡眠期间大脑产生的奇怪吸引子的分维。对于清醒的大脑，神经元发放的混沌活动处于低水平。但随着大脑渐沉入睡，混沌变得愈益显著。然而，在快速眼运动期间(即做梦时)，背景混沌的量值减少。巴布洛延茨认为，大脑奇怪吸引子的分维能提供各种睡眠阶段不同深度的量度。

洛斯阿拉莫斯非线性研究中心的科学家们计算了与不同麻醉深度有关的奇怪吸引子的分维。这个小组还认为，将有可能实现脑电图的计算机分析，以刻划不同种类的症状。其他科学家还想就高层次思维甚或创造性的分形特征，探究复杂的脑电波。

脑科学前沿是一个广阔的领域，研究人员们已经开始闯入这片荒野。大脑模型以摇滚舞星般的频率在普及中变化，很可能在一个世纪的时间里，目前的神经生理学景象图就象16世纪的美洲海图那样离奇。但是地图必须从某处起步，在制图者们当中，越来越多的科学家正试图在这幅大画面中描绘一个非线性轮廓。而经过数年的研究，也证实了人的大脑是一个非线性动力系统。其中存在着非常深奥复杂的混沌机制[34]。

## 2.3声样的采集

发声过程中由声带振动激励发出的声音（浊音）一般为元音；由气息摩擦撞击声道产生湍流发出的声音（清音）则一般为辅音. 在日常发音中，发音间隙短、连读、省略、儿化等情况，特别是大量存在的噪声干扰，给辅音的准确提取和识别带来很大的困难. 相对而言，元音具有有基频、能量大、过零率低等特点，更易于检测。因此，在语音信号处理的研究中，通过确定元音位置指导辅音提取、音节端点检测，利用元音辅音差异进行语音识别、说话人识别等成为一种重要的思路。

研究人员们已经提出了语音产生的理论及物理模型，普遍认为语音的产生可分为三个部分，声源产生、声道传递和口鼻辐射[35,36]。语音的产生与声源、声道和口鼻辐射都有密切关系。言语的产生由两个基本的机械作用构成：发声和发音。发声即产生声学信号；发音包括说话入对声学信号的调音过程，主要用舌、嘴唇和软腭，以及口腔和鼻腔的共鸣作用。音素分为浊辅音、清辅音和元音。

因此，本课题采集的声音样本是脑卒中病人和正常人发出的元音信号，避免采集与口腔包括唇齿舌鄂发生任何摩擦而产生的辅音，从而排除影响声音非线性性质的摩擦噪音等因素，这样的声音信号更能够反应大脑的控制状态。但是，采用哪一个元音尚有很多不同看法。

元音的产生有3个步骤：声带振动调制呼吸气流，产生一复合声；声门与口唇之间的声道对复合声中各分音进行选择传递；经调变后的声音从口唇发出。以往有关元音的研究主要集中在共振和构语器官上，认为在产生元音的3个步骤中，仅声道的选择性传递对各个元音的区别具有重要意义，即形成不同的共振峰。元音声道的缩窄区均近于声门，但[a]音声道缩窄程度较[ai]音声道缩窄程度更明显[37]。此外，在构语步骤上，研究认为元音的共振峰频率与舌位相关。近年来有很多文章提出不同的元音影响到基本的声音学参数。

1982年，Horili对成年男性发8种英语元音作嗓音声学分析，发现不同的元音的频率微扰(Jitter)和振幅微扰(Shimmer)有所不同。Linville等通过测试28名老年人发三种元音[a]、[i]、[u]，发现[u]音的频率微扰(Jitter)和振幅微扰(Shimmer)较[a]、[i]低，[a]和[i]差异不明显[38]。

魏春生发现发[ai]和[a]音的基频(F0)和声音的扰动值无显著性差异，但发[a]音时规范化噪声能量(NNE)值明显低于[ai]音，推测发[a]音时声带呈较为自然的状态，张力相对较小，声门闭合的程度相对更完全[39]。

侯丽珍等通过分析认为发[i]音时声带张力较大，声门下压较大，易受代偿能力影响，掩盖真实状况，基频高，振动微扰值小，谐噪比最大，[u]音噪声能量值太大，采样时有可能影响其它声学参数的准确性；研究提出[ai]和[a]是较好的检测声样，二者相比，[a]音在声音分析检测中更为适合[40]。

因此基于以上分析，本课题采用了[a]音作为采样元音，使声音能够不受声道调节的影响，不与口腔包括唇齿舌鄂发生任何摩擦，还原最原始从大脑控制系统发出的声音信号。

## 2.4本课题的采集样本

### 2.4.1 样本采集装置

本课题选用Superlux ECM999背驻极电容式麦克风，适用于室内音频分析系统，以及各类测试、测音和录音。其频率响应20-20000hz拥有线性频率响应（如下图3），可以真实的记录声音。电容式麦克风也有很多优点，电容式音头极为轻薄的振动膜，具有极快速的瞬时响应特性。电容式麦克风的振动膜比较轻，先天上就具有超低触摸杂音的绝佳特点。同时电容式麦克风具有超高的灵敏度，可以感应极微弱的声波，输出最清晰、细腻及精准的原音。本课题中采集声音时的采样频率为22050hz，16bit。

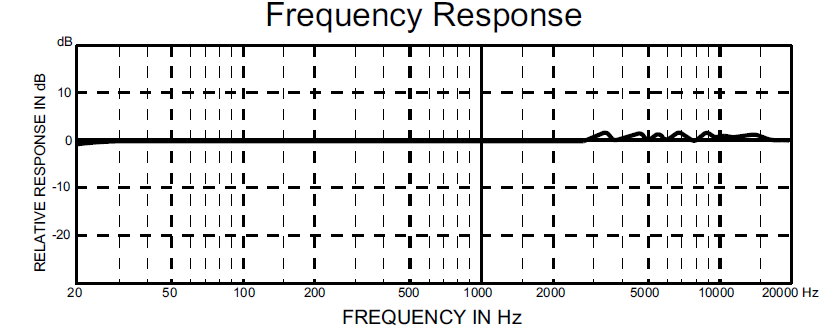


图2. 3ECM999的频率响应

脑卒中研究对象为2012年10月—12月连续收住在杭州市中医院针灸康复科的全部脑卒中患者，在患者自愿配合采样的情况下，采集了当时住院35例脑卒中患者中的17例患者。这些患者均为脑梗塞或者脑出血患者，由临床医师确诊，并有CT、MRI证实。脑卒中患者详细病例见表2.1。脑卒中研究对象年龄层为50-80左右。用于对比的研究对象为同年龄段的健康人，共计20例。

表2. 1采样脑卒中病人详细病历

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 年龄 | | | 性别 | | 诊断病症 |
| 69 | 男 | | | 脑梗塞,右侧脑室枕叶损伤,言语含糊,左侧肢体障碍 |
| 74 | | 女 | | 脑梗塞,左侧脑颞顶枕叶及基底节区损伤,运动型失语 | |
| 80 | | 男 | | 脑出血,右侧额顶叶损伤,左侧肢体障碍 | |
| 64 | | 男 | | 脑梗塞，左侧基底节区，左肢体障碍 | |
| 67 | | 女 | | 脑梗塞，左侧侧脑室旁梗塞 | |
| 78 | | 男 | | 脑出血，两侧侧脑室旁基底节区和脑干出血，言语不利 | |
| 68 | | 女 | | 脑出血，右侧基底节区出血，左侧肢体障碍，言语不利 | |
| 82 | | 女 | | 脑梗塞，右侧肢体障碍 | |
| 74 | | 女 | | 脑梗塞，左侧额，顶，枕叶，右侧肢体障碍，言语不利 | |
| 75 | | 男 | | 脑梗塞，左侧脑桥，双侧基底节区梗塞，右侧肢体功能障碍 | |
| 65 | | 男 | | 脑梗塞，两侧脑室旁，基底节区，右侧丘脑及脑干，双侧肢体障碍 | |
| 81 | | 男 | | 脑梗塞，左侧颞顶叶，右侧肢体障碍，言语不利 | |
| 74 | | 男 | | 脑梗塞，左侧肢体障碍 | |
| 59 | | 男 | | 脑出血，右侧基底节区出血，左侧偏瘫 | |
| 85 | | 男 | | 脑出血，左侧硬膜下血肿，左侧肢体障碍，言语含糊 | |
| 71 | | 男 | | 脑梗塞，右侧脑室旁及脑干梗塞，左侧肢体障碍 | |
| 80 | | 男 | | 脑梗塞，右侧额顶叶梗塞 | |

采集声音样本时的环境要求在安静的室内进行；采所有声音样本是以25kHz的采样率和16位的分辨率采集的，声音采集软件如图2. 4所示，时间约为1.5 s-3 s，受试声样为汉语元音/a/。受试者检测前先作短时的发声训练。发音要求平稳(以舒适发音为好)，数据采集应尽可能多，重复采集了三次。取声样中平稳的部分进行分析。口与麦克风的距离要保持15-20cm。如图2. 5所示。

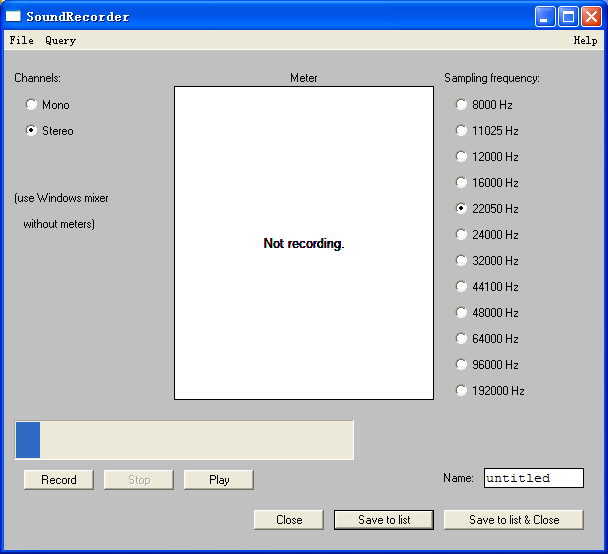


图2. 4声音采集软件



图2. 5真实声音采集图

## 2.5 声音信号增强处理

在采集人体发出声音的过程中，虽然是在安静的环境中采集，但还是会有一些噪声的存在，这些噪声会直接影响到采集声音信号的质量，从而影响接下来的声音信号处理过程。所以本节会对声音信号的噪声来源作出简要分析，并且应用相应算法进行语音增强和降噪处理。

### 2.5.1 噪声的特性

噪声的来源多种多样，特性变化无穷，基本取决于当时的采样环境。根据噪声同采集语音信号的关系，可以将噪声分为加性噪声和非加性噪声两种。非加性噪声可以通过同态滤波等变换转换为加性噪声。某些量化噪声也可以通过变换转为独立的加性噪声。加性噪声有周期性噪声，宽带噪声，冲激噪声，语音干扰噪声等等。非加性噪声主要是传送网络中的电路噪声。

1. 周期性噪声

周期性噪声一般见于电动机、电风扇等周期地运作的机械发出的噪声，交流电的工频干扰也是一种周期性的噪声。这种噪声的特点是具有离散线谱的频谱图，这些线谱分量是时变的，且位置也不定。这些噪声一般采用自适应滤波法降噪。

1. 宽带噪声

宽带噪声有多种来源，说话时呼吸气流的噪声，热噪声以及多种随机噪声声源，量化噪声都是宽带噪声。在实际应用中，一般将这类噪声近似为白噪声或者Gauss噪声。其噪声消除比较困难，一般采用非线性方法消除。

1. 冲激噪声

冲激噪声来源于爆炸，打火，撞击等干扰。这类噪声时间域上有尖端窄脉冲。一般采用在时域中消除该类噪声，有两种方法：对带有噪声的声音信号设置幅度阈值，该阈值由幅度平均值确定，过滤掉超出此阈值的信号以达到降噪目的；当窄脉冲不密集时，可以用内插法对声音信号进行平滑降噪。

1. 语音干扰噪声

该类噪声是因为在同一信道中同时包括了两种不同的语音。一般利用这两种语音基音不同的特点，用梳状滤波器提取基音和谐波再恢复原语音信号，以达到消除语音干扰噪音的目的。

通过语音增强降噪技术提高语音质量的过程如下图2. 6所示。



图2. 6语音增强流程框图

### 2.5.2 声音去噪方法之小波去噪

由于噪声有多种种类，所以语音去噪的方法也有所不同。多年来，人们针对种种噪声研究出了各种方法，主要分为几类：参数法，非参数法，统计法和其他方法。

参数法主要依赖于使用语音生成模型比如AR模型，需要提取相应模型的参数。如果采用滤波器模式，主要有维纳滤波器、卡尔曼滤波器等。

非参数法不需要依赖于模型参数，因此得到了广泛应用。这类方法包括改进自适应滤波的自适应噪声抵消法，该方法有一个参考噪声作为辅助输入，因此起到了很好的降噪效果。减谱法假设噪声是独立与语音的，从带噪声语音的功率谱中减去噪声的功率谱，起去噪声目的。

统计方法是一类利用语音和噪声统计特性来去除噪声的方法。

其他方法包括小波变换，人工神经网络等等近年来兴起的方法。

本文中采用的对采集声音信号的去噪方法是小波去噪的方法。在第一章中提到过用小波分析对信号进行分析处理，小波变换其实就是由一系列滤波器组成。其中包括一个低通滤波器和一组带通滤波器。通过这些滤波器小波变换可以将信号在多个尺度上进行分解提取，在这些尺度上分析得到的小波系数代表了不同分辨率上面的信息。语音信号的纯净语音和噪声在小波系数上有着很大的不同，所以可以通过这个区别点将两者分开。因此小波变换越来越广泛得应用于去噪中。

下面就将小波变换应用于嗓音去噪中来进行分析。先分析噪声在小波分解下的特性，将噪声看成一个普通的信号，并对它进行分析，有三个主要特征需要注意，即相关性、频谱和频率分布。总体上说，对于一个一维离散的信号，小波分解的高频第一层依赖于该信号的高频部分，该信号的低频部分影响的是小波分解的最深层及其低频层。如果对一个只是白噪声组成的信号进行小波分解，则可以看成：高频系数的幅值随着分解层的增加而很快地衰减，并且，高频系数的方差也很快地衰减。

小波去噪的方法一般有三种：第一类方法是基于小波变换模极大值原理，第二类方法是对含噪信号作小波变换之后，计算相邻尺度间各点小波系数的相关性，根据相关性的大小区别小波系数的类型，从而进行取舍，然后直接重构信号。第三类方法是Donoho 提出的阈值方法。小波阈值去噪的步骤是（1）一维信号的小波分解。（2）小波分解高频系数的阀值量化，对第1 到第N 层的每一层高频系数，选择一个阀值进行阀值量化处理，得出估计小波系数。（3）一维小波的重构，根据小波分解的第N 层的低频系数和经过量化处理后的第1 层到第N 层的高频估计系数进行一维信号小波重构。

图2. 7中对noisbump含噪声信号进行小波阈值去噪，其中第二张图中是用Birge-Massart法估计阈值进行小波去噪，其中的第三张图是用缺省阈值法进行的小波去噪。



图2. 7 noisbump信号的小波阈值去噪图

图2. 8中对采集到的声音信号进行小波阈值去噪，其中第二张图中是用Birge-Massart法估计阈值进行小波去噪，其中的第三张图是用缺省阈值法进行的小波去噪。



图2. 8健康人声音信号的小波阈值去噪图

从图2. 7和图2. 8中可以看出对于noisbump信号，两种阈值估计方法都可以起到很好的小波去噪的作用，但是对于健康人声音信号，Birge-Massart阈值估计法则使得声音信号缺失了很大一部分信息，而缺省阈值法则可以很好地去噪。这也说明了对于小波阈值去噪来说，选取一个合理的阈值对信号的去噪和恢复有很大的影响，因为这个阈值门限直接影响了对信号去噪的效果和恢复信号的失真度。

## 2.6本章小结

本章首先介绍了语言功能的神经机制和脑成像机制，详细介绍了语言功能跟大脑状态的关系，以及大脑损伤状态会导致的语言功能障碍等等。之后介绍了大脑的非线性研究，众多研究者经过多年研究证实了大脑的非线性。接下来介绍了本文中的声样数据的采集设备，采集元音选择标准以及声音采集软件和真实的采集声音的环境设备等等。之后又列举了采集声样的脑卒中病人的详细病例。最后本章介绍了对采集到的声音信号进行去噪和增强，讨论了各种噪音的来源和性质，及其相应的去噪方法，之后介绍了应用于本文采样数据的小波阈值去噪方法。

# 第3章 非线性时间序列分析及其在嗓音信号分析中的应用

非线性时间序列分析方法的研究对象是无法建立数学模型或未知数学模型的复杂系统，研究过程中仅有的信息是通过观测或实验手段在这个复杂系统中获得的单变量或多变量时间序列，研究的目标就是为了寻找能够反映该复杂系统本质特征的不变量，分析复杂系统的特征及内在变化规律，最终在实际问题中可以指导控制具体复杂系统。具体流程如图3. 1所示。



图3. 1非线性时间序列分析框架图

非线性动力学中的混沌理论与生物医学工程的结合，极大推动了生理学、病理学和脑科学等生物医学问题的深入研究[41,42]。同时也为各种复杂生物信号的分析提供了新的手段和方法。声音信号近年来经空气动力学和声学研究证明是非线性信号，所以本章对声音时间序列采用非线性时间序列分析的方法，通过对声音信号的分析以窥其控制系统—大脑控制系统的状态。

## 3.1 混沌理论概述

混沌(Chaos)是确定性系统中出现的极其复杂的、类似随机的现象。这里，“确定性系统”是指混沌系统由确定的动力学方程所描述。“随机”是指混沌本身具有内随机性，表现为系统长期行为的不可预测性。混沌现象表明了确定性与随机性两者是相通的，体现了两者即对立又统一的关系，即确定性内在地包含随机性，随机性隐含着确定性。混沌是有序中产生的无序运动状态，无序来自有序，无序中蕴含着有序。混沌不等于混乱，是一种貌似无序的复杂有序现象。混沌系统的最大特点就是在于系统的演化对初始条件十分敏感，因此从长期意义上讲，系统的未来行为是不可预测的[43]。

迄今为止，对于混沌还没有严格并且被普遍接受的数学定义。最早给出混沌定义的是由J.A.Yorke和李天岩在1975年发表的一篇《周期3意味着混沌》的论文中提出的，因此该定义成为Li-Yorke定义。

Li-Yorke混沌定义：设f:JJ是闭区间JR上的连续映射，如果满足下列条件：

1. f周期点的周期无上界；
2. 存在不可列集SJ，S不含周期点，且满足
3. 对任意x,yS，xy，有

(3.1)

1. 对任意x,yS，有

(3.2)

1. 对任意xS及周期点yJ，有

(3.3)

式中，表示对函数的n次迭代，即

(3.4)

则称：JJ在S上是混沌的。

根据Li-Yorke定义，Day指出了混沌系统应该具有的三个性质是：

1. 存在所有阶的周期轨道。
2. 存在一个不可数集合，该集合值含有混沌轨道，且任意两个轨道既不趋向远离也不趋向接近，而是两种状态交替出现。同时，任一轨道不趋于任一周期轨道，即该集合不存在渐近周期轨道。

混沌轨道具有高度的不稳定性。

1989年，Davancy R.L，提出了一个影响比较大的混沌定义，为多数人所接受。在给出这个定义之前要先定义两个术语：

术语1 设J为拓扑空间，映射：JJ。如果对任意的开集UJ，VJ，存在k>0，kZ，使

(3.5)

其中，表示对函数的次迭代，则称：JJ为拓扑传递的。

术语2 如果对给定的，对任何xJ和在x的任何邻域U上，都存在yU和，，使得

(3.6)

则称：JJ对初始条件具有敏感依赖性。

直观上，映射具有对初始条件的敏感依赖性是指，对于任意接近的点，在的有限次迭代之后，它和的分离程度可大于任意给定的。这里强调，并非要求附近的所有点都在迭代下与分离，而是要求在的每一个邻域中必须存在这样的点。

Devancy R.L混沌定义：设V为一集合，如果满足下列三个条件，则称：VV在V上是混沌的：

1. 有对初始条件的敏感依赖性；
2. 是拓扑传递的；
3. 状态点在V中是稠密的。

在Devany R.L的定义说明混沌映射具有三个要素，即不可预测性，不可分解性和规律性。

混沌现象具有其特有的特征，其中包括[44]：

1)混沌对初始条件有极端敏感性，所以不可预测，这个是混沌区别于其他运动形态的本质特征。混沌系统长期行为是不可预测的，其短期行为是可以预测和确定的。

2)混沌是有界的，它的运动轨迹始终局限于一个确定的混沌吸引域中，所以从总体来说混沌系统是相对稳定的。

3)混沌具有确定性，其确定性分为两个方面：首先，混沌系统是确定的物理系统；其次混沌并不是真正随即的，而是貌似随机的。系统下一时刻的状态受当前状态的影响。

4)混沌具有内随机性，这种随机性是确定性系统的内在随机性，这种随机性说明，混沌系统是局部不稳定的。

5)混沌具有分维性，混沌的运动状态具有自相似结构，混沌的这种行为特征可以用分维性来表示。

6)混沌有某种统计确定性。例如李雅普诺夫指数是正的，这个也是判定是否是混沌时间序列的特征之一，而且还具有连续的功率谱，同样是判据之一。

## 3.2 相空间重构理论

20世纪70年代，混沌学作为一门新的科学正是诞生。到了80年代，混沌的理论得到了进一步发展。刻画混沌的基本概念例如Lyapunov指数，分数维和吸引子等被确定下来。1980年Packard等人提出了由一维可观察量重构一个“等价”的相空间，来重现系统的动力学特征[45]。同时Takes等人提出了根据时间序列重构系统动力学结构的延迟坐标法[46]，为混沌时间序列的建模预测提供了数学理论基础。相空间重构的基本思想是：由于混沌系统产生的轨迹经过一定时间的变化后，会最终做一种有规律的运动，产生一种规则的，有形的轨迹(混沌吸引子)，相空间重构法虽然是用一个变量在不用时刻的值构成相空间，但是系统中任意分量的演化都是由与之相互作用的其他分量所决定，相关分量的信息隐含在任一分量的发展过程中，即用系统的一个观察量可以重构出原动力系统模型。因此重构的相空间的轨线也反映系统状态的演化规律。其重构原理如下：

设有一时间序列重构维相空间，得到一组相空间的矢量

(3.7)

在脑卒中病人的声音分析中即为原始时间序列，其中是重构的相空间的嵌入维数，是时间延迟。Takens证明了可以找到一个合适的嵌入维，即如果延迟坐标的维数,是动力系统的维数。在这个嵌入维空间里可以把有规律的轨迹（吸引子）恢复出来。所以按照Takens定理可以在拓扑等价的意义下恢复吸引子的动力学特征。

一个比较适用的方法也是普遍应用的方法是有Grassberger和Procaccia提出的G-P算法[58]，其只要步骤如下：

1. 利用时间序列，先给出一个较小的值，对应一个何从够的相空间。
2. 计算关联函数

(3.8)

其中表示相点和之间的距离，是Heaviside函数，是一个累积分布函数，表示相空间中吸引子上两点之间的距离小于的概率。

1. 对于的某个适当范围，吸引子的维数与累积分布函数应满足对数线性关系，即。从而由拟合求出对应于的关联维数估计值。
2. 增加嵌入位数，重复计算步骤2）和3），直到相应的维数估计值不再随着的增长，停留在一定的误差范围之内。此时的即为吸引子的关联维数。反之，如果不能随着的增长停留在一个稳定值，而是一直增长，则说明所分析的系统是产生的时间序列是一个随机时间序列。

由Takens定理知，在没有噪声的无限长精确数据情况下，可以任意选择延迟时间，但是在实际实验分析中，被测的时间序列是有限长的，且一般都有噪声污染，因此在重构相空间过程中的选取也起着非常重要的作用。公式中的和的选取具有十分重要的意义，一般只能根据经验来选择，选取时的基本思想是使得和具有某种程度的独立但又不完全无关，以便它们能够在重构相空间中作为独立的坐标处理。在实际的应用中，和的选取一般有两种观点：一种是认为两者都是互不相关的，他们的选取可以独立进行，Takens认为和是互不相关的。另外一种认为两者是相关的，和的选取是相互依赖的，基于这种观点的选取方法有1996年Kugiumtzis提出的时间窗口法，1999年Kim等人基于嵌入窗法的思想提出的C-C方法等。以上提到的方法都具有一定的主观性，目前并没有一种适合分析各种非线性时间序列的统一通用方法，需要结合实际情况选取合适的重构方法。与此同时，也有各种新的重构方法不断被人们提出。

目前求取时间延迟的方法主要有自相关法、复相关法和互信息法等。寻找最小嵌入维数的方法主要有最大特征不变量法、几何不变量法、伪邻近点法以及对其改进的CAO方法等。

## 3.3 延迟时间的计算和应用

由上节中的Takens定理可以知道，延迟时间的选值很重要，如果选择的太小，则可能使得两个采样点间隔太近以至于不能区分两者，从而不能提供两个独立点，导致吸引子的轨迹过于靠近相空间的对角线，重构的相空间杂乱无规则的；而如果选择太大，那么两个点可能会不想关，吸引子轨迹会映射到两个完全不想关的方向上，从而使得相空间中的轨线失真。下面就介绍一些比较主流的方法。

3.3.1 计算延迟时间方法

1、自相关法

自相关函数法是一个非常成熟的计算延迟时间的方法，它主要是提取时间序列之间的线性相关性[47]。一般都可以对一个混沌时间序列写出其自相关函数，之后做出自相关函数关于时间的函数图像。根据数值分析结果得出，当混沌序列的自相关函数下降到初始值的时，所得的时间就是重构相空间的时间延迟。

对于连续变量，其自相关函数如下：

(3.9)

其中是时间的延迟值，表示两时刻（和）运动或随机过程的相互关联或相似的程度。当的幅值一定时，越大，则意味着和越相似，从而越大。反之，越大，则和的差别可能越来越大，最后以至和完全无关，而愈来愈小直至趋于0。如果带越宽，越窄，对于理想的白噪声，变成函数。

自相关函数的性质：

1. 是实偶函数，即；
2. 是的最大值；
3. 等于的均方（平均平方）值；
4. 若是周期函数，则也是周期函数。此时不仅在处有极大值，而且在（为周期）处都会有极大值；
5. 在有些情况下，运动并不是规则的，但也不是完全随机的，而是如混沌运动和噪声等。设如果包含有随机过程和规则运动两部分，那么有，则的自相关函数等于这两部分各自的自相关函数之和

(3.10)

对于混沌时间序列，为了处理问题方便，这里设总点数为，序列的时间跨度为的自相关函数为

(3.11)

由此可固定，作出自相关函数关于时间延迟（）的函数图像，则自相关函数下降到初始值的时，所得的时间就是重构相空间的最佳时间延迟。

但是，自相关函数法仅仅能提取时间序列的线性相关性。根据这种方法得到的时间延迟可以分别使和及和之间不相关，但是和之前的相关性可能会很强，这一点也限制了将通过自相关法得到的时延迟推广到高维。

2、互信息法

由于互相关函数法无法推广到高维，并且只能度量两个变量的线性依赖性的局限，Shaw首先提出了以互信息第一次达到最小时的滞时作为相空间重构的时间延迟，Fraser给出了互信息计算的递归算法[48]。互信息函数法改进了互相关函数法的局限，可以度量两个变量的总体依赖性。在大量数值实验中可以发现，互信息法对应第一个极小值的互信息量较小，折叠的开端能够比较清楚地区分，因而重构的相空间可以定量和定性分析吸引子的动力学特征。

互信息的基本概念如下：

设有离散变量，，其状态数分别为，。则最普通的信息函数——熵的定义为：

(3.12)

其中是变量在状态出现的概率。而变量和的联合熵定义为

(3.13)

这里的是变量在状态且变量在状态时出现的概率。所以，由两个变量，的熵和联合熵的定义可得到两者的互信息为

(3.14)

互信息作为一个总体关联量相似于线性相关系数，但互信息对其他关系也非常敏感，不仅仅只是线性独立性。

互信息计算的递归算法：

设混沌时间序列，时间延迟为,嵌入维数，重构相空间得到

(3.15)

其中。则系统对变量的平均信息量为系统的熵

(3.16)

设，考虑一个总的耦合系统，假定已知为，则的不定性为

(3.17)

其中是条件概率。

设在时刻时已知，则在时刻的平均不定性为

(3.18)

这里

(3.19)

其中是孤立的的不确定性，是已知的的不定性。所以的已知减少了的不确定性，其互信息为：

(3.20)

互信息不是单个变量或者的函数，而是联合概率分布的函数，它是分布的总体度量。

关于互信息函数的估计，当这两个序列一个是，一个是时，借助数据分布的直方图，是时间序列落在直方图中第个盒子里的概率，是在第个盒子并且在第个盒子中的联合概率。

则互信息函数的估计如下[49]：

(3.21)

我们取出现第一个极小值时的延迟值作为最佳的延迟时间。

互信息法虽然有很大的计算量，但是其包含了时间序列的非线性特征，改进了自相关法的局限，是一种普遍使用的方法。

3.3.2时间序列提取延迟时间

通过上一节计算时间延迟的方法的讨论，本文采取的时间延迟的计算方法是互信息函数法，分别对考毕兹振荡器（Colpittsoscillator）混沌电路产生的时间序列和采集的健康人的声音信号计算时间延迟。

图3. 2是采用互信息函数法对考毕兹振荡器（Colpittsoscillator）混沌电路产生的时间序列进行延迟时间间隔的求取。



图3. 2考毕兹振荡器时间序列求取时间延迟

图3. 3是采用互信息函数法对采集到的健康人的声音时间序列进行延迟时间间隔的求取。



图3. 3健康人声音时间序列求取时间延迟

图3. 2和图3. 3中的曲线表示平均互信息函数随时延的变化趋势，箭头所指的点为函数曲线的第一个局部极小值。即为混沌时间序列的最佳的时间延迟。

在两个图中还标示出了第一个最小极值对应的值。这个值表示了原始声音序列和延迟序列之间的差异，如果差异越大说明原始序列越不规律。

## 3.4嵌入维数的计算和应用

### 3.4.1 计算嵌入维数的方法

根据Packard和Takens所提出的基于拓扑嵌入理论的重构等价相空间法，嵌入维数的合适选取才能使得吸引子被正确地恢复。在实际的应用中常用以下的方法选取嵌入维数。

1. 伪最近邻法

伪邻近点法是一种从几何观点出发较容易实现的方法，其基本思想：当维数从变成时，考察轨线邻点中的真实邻近点和伪邻近点，当没有伪邻近点时，可以认为几何结构完全打开。

设是的最近的邻接点，它们之间的距离为；当维数增加到，它们之间的距离为。若的值比的值大很多的时候，可以认为是由于高维吸引子中两个不相邻的点，当投影到低维轨道上时变成了相邻的两点造成的，因此，这样的邻近点是虚假的，伪的。若

(3.22)

则是的虚假的最近的邻近点，这边的阈值可以在选取。对无限长精确的数据，用上述标准可获得较好的结果；对有限长具有噪声的数据，则要用以下的标准：

若

(3.23)

其中，，则此时是的虚假邻近点。

对实测时间序列，从2开始，取，计算伪邻近点的比例，然后增加到虚假邻近点的比例小于5%或虚假邻近点不再随着的增加而减少时，可以认为完全打开，此时的即为最小嵌入相空间维数。

1. CAO方法

伪邻近点法求取嵌入维数的方法对信号中含有的噪声非常敏感，有其局限性导致结果不准确。90年代Cao Liangyue 在伪最近邻法的基础上提出了改进方法即CAO方法[50]，该方法只需要延迟时间一个参数就能够很好地区分随机信号和确定性信号，在用较少的数据量的情况下求得嵌入维数。

CAO方法的基本思想是：由时间序列构成的维相空间：

(3.24)

定义：

(3.25)

其中为求最大值范数，即是的最近邻近点，邻近点的求法同伪最近邻法中的求法相同。

记所有关于的均值为

(3.26)

只与两个变量和相关，为了研究嵌入维数从变为时相空间的变化情况，定义

(3.27)

随着的增长，停止变化时记录下，则即为我们找的嵌入维数。

对于随机系统的时间序列，原则上，随着的增长，不会趋于稳定，但是实际计算中，当取值很大时，很难看出这时的是否趋于稳定。为了解决该问题，定义了量：

(3.28)

同样定义

(3.29)

对于随机的时间序列，由于两个时刻之间的值是独立的，对于任何，将始终为1，对于确定性系统的时间序列，不为常数。因此我们同时计算和来确定时间序列重构相空间的最小嵌入维数。

### 3.4.2 时间序列计算嵌入维数

通过上一节计算最小嵌入维数的方法的讨论，本文采取的时间延迟的计算方法是CAO方法，分别对考毕兹振荡器（Colpittsoscillator）混沌电路产生的时间序列和采集的健康人的声音信号计算最小嵌入维数。

图3. 4是采用互信息函数法对考毕兹振荡器（Colpittsoscillator）混沌电路产生的时间序列进行最小嵌入维数的求取。



图3. 4考毕兹振荡器时间序列求取嵌入维数

图3. 5是采用互信息函数法对采集到的健康人的声音时间序列进行最小嵌入维数的求取。



图3. 5健康人声音时间序列求取嵌入维数

图中点线表示曲线随着嵌入维数的增加而逐渐变化的趋势，可以看出随着增长，趋于稳定。图中三角形点线表示曲线随着嵌入维数的变化趋势。可以看到并不是对于所有都为一个常数，说明考毕兹振荡器（Colpittsoscillator）混沌电路产生的时间序列和健康人的声音时间序列都是确定性的时间序列。

### 3.4.3 根据延迟时间和嵌入维数重构相空间

根据前面几节的对时间延迟和嵌入维数的计算方法的讨论，分别用应用互信息法求得到时间延迟，应用CAO方法得到的嵌入维数，可以得到重构时间序列的相空间。一般来说高维的相空间不能直接作图，因此本文中都是重构的相空间在二维和三维空间的投影图。

图3. 6是对考毕兹振荡器（Colpittsoscillator）混沌电路产生的时间序列重构相空间，取延迟时间，取最小嵌入维数。重构相空间，得到映射到三维空间下的吸引子如下。



图3. 6考毕兹振荡器时间序列吸引子

图3. 7是对健康人的声音时间序列重构相空间，取延迟时间，取最小嵌入维数。重构相空间，得到映射到三维空间下的吸引子如下。



图3. 7健康人声音时间序列吸引子

## 3.5李雅普诺夫指数的计算和应用

### 3.5.1 李雅普诺夫指数的定义

混沌系统由相空间中的不规则轨道奇怪吸引子来描述。奇怪吸引子的一个明显特征就是吸引子邻近点的指数离析。由于相空间中的点表示着整个系统，所以邻近点的指数离析意味着初始状态完全确定的系统，在长时间的情况下也会发生变化。这种运动性态就是系统对初始条件具有敏感依赖性的反映。Lyapunov指数正是对这种敏感性的量度，它表征了系统在相空间中相邻轨道间平均收敛性或平均分散性的平均指数率。近二十年来，Lyapunov指数已广泛应用于判别系统的混沌行为，同时也应用到语音信号，心电信号，脑电信号等非线性信号的研究中，成为一种重要的判别混沌信号的工具之一。

对于维相空间中的连续动力学系统，定义Lyapunov如下：对于一个无穷小的维球面的长时间演化，由于流的局部变形特性，球面将变为维椭球面。将维椭球的个主轴按其长度顺序排列，那么将第个主轴的长度的增加速度定义为第个Lyapunov指数

(3.30)

有，成为Lyapunov指数谱，其中最大的称为最大Lyapunov指数。

Lyapunov指数的大小表明了相空间中相近轨道的平均收敛或发散的指数率。一般来说，具有正和零Lyapunov指数的方向，都对支撑起吸引子起作用，而负的Lyapunov指数对应着收缩的方向，这两种因素相互作用的结果就是伸缩和折叠操作，从而形成了奇怪吸引子的空间几何形状。因此，对于奇怪吸引子而言，其最大的Lyapunov指数越大，系统的混沌性越强；反之则越弱[43]。但是其中必然存在着负的Lyapunov指数，这样才能使得系统轨线收缩为吸引子。

对于耗散系统，Lyapunov指数谱不仅仅描述了各条轨道的性态，而且还描述了一个吸引子的吸引域出发的所有轨道的稳定性性态。

一维情形：吸引子只能是不动点，此时必有。

二维情形：吸引子为不动点或者极限环。对于稳定不动点，任意方向的相空间中两个靠近点之间的距离都要收缩，故这时两个Lyapunov指数都是负数，即。对于稳定的极限环，若相空间中两个靠近点之间的距离始终是垂直于环线方向的，它一定要收缩，此时的Lyapunov指数为负数；过相空间中两个靠近点之间的距离沿轨道的切线方向，则它既不增大也不缩小，此时的Lyapunov指数是零。所以极限环的。

三维情形：

，不动点；

，极限环；

，二维环面；

，不稳定极限环；

，不稳定二维环面；

，奇怪吸引子。

Lyapunov指数可以表征系统运动的特征，其沿某一方向取值的正负和大小，分别决定了系统的吸引子中相邻轨道沿该方向的平均发散或者收敛的快慢程度。最大Lyapunov指数决定了轨道发散直到覆盖整个吸引子的快慢，最小Lyapunov指数则决定了轨道收缩的快慢。所有的指数和大体表征轨道总的平均发散快慢。

因此经过上述分析，还可以看出如果是混沌系统，必须有一个Lyapunov指数是正的，这个就可以作为是否为混沌吸引子的判据。

### 3.5.1 李雅普诺夫指数的计算

Lyapunov指数的计算方法大体分属两大类：Wolf方法和Jocobian方法。其中Wolf方法适用于时间序列无噪声，切空间中小向量中的演变高度非线性[51]；Jocobian方法适用于时间序列噪声大，切空间中小向量的演变接近线性[52]。随后在这两种方法的基础上，G.Barana等提出了p-范数算法，但是这种方法计算很复杂。M.T.Rosenstein等人提出了一种小数据量的计算方法，该方法操作起来比较方便，而且计算量较小，对小数据组比较可靠。所以本文在这边要介绍应用的方法是小数据量方法[53]。

小数据量方法是一种计算混沌时间序列的最大Lyapunov指数的方法。在混沌研究和实际应用中，不必计算出所以的Lyapunov指数谱，只要计算得到最大Lyapunov只是就可以对系统的混沌性质做判断。

小数据量方法如下：

根据时间延迟和嵌入维数重构相空间；用快速傅里叶变换来计算序列的平均周期。找相空间中的每个点的最近邻点，并限制短暂分离，即

(3.31)

对相空间中的每个点，计算出该邻点对的个离散时间步后的距离

(3.32)

对每个，求出所有的平均即

(3.33)

其中是非零的数目，并用最小二乘法作出回归曲线，该直线的斜率就是最大Lyapunov指数。

下面图3. 8给出了计算最大Lyapunov指数的流程图：



图3. 8计算Lyapunov指数流程图

下面给出几种常见的运动形态的特征量，见表3. 1。

表3. 1几种常见的运动性态的特征量

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 吸引子 | 维数 | Kolmogorov熵 | Lyapunov指数 |
| 稳定定态 | 点 | 0 | 0 |  |
| 周期运动 | 闭曲线 | 1 | 0 |  |
| 随机运动 | 无 |  |  |  |
| 混沌运动 | 奇异 | 非整数 |  |  |

### 3.6.2 时间序列的李雅普诺夫指数

经过上一节讨论了计算最大Lyapunov指数的步骤，下面用最小数据量方法计算最大Lyapunov指数。

图3. 9为对考毕兹振荡器（Colpittsoscillator）混沌电路产生的时间序列计算其最大Lyapunov指数，以其吸引子的10000个时间采样点为例，利用最小数据法求取最大Lyapunov指数。图3.9中是随离散时间步长变化的曲线，取线性最好的一段进行拟合，计算得到斜率为0.05，此即为其吸引子的最大Lyapunov指数。



图3. 9考毕兹振荡时间序列计算Lyapunov指数

图3. 10为对健康人声音时间序列计算其最大Lyapunov指数，以其吸引子的10000个时间采样点为例，利用最小数据法求取最大Lyapunov指数。图3.10中是-变化的曲线，取线性最好的一段进行拟合，计算得到斜率为0.003，此即为声音时间序列吸引子的最大Lyapunov指数。



图3. 10健康人声音时间序列计算Lyapunov指数

通过计算时间序列的最大Lyapunov指数，可以得到考毕兹振荡器（Colpittsoscillator）混沌电路产生的时间序列和健康人声音时间序列的最大Lyapunov指数都是正的，说明这些信号都具有混沌性质。

## 3.7本章小结

本章首先介绍了混沌序列的概念、特征及其在非线性动力学方面的广泛应用。其次详细讨论了非线性重构相空间理论，为了重构相空间需要计算相关的延迟时间和嵌入维数这两个参数，并且列举了求取方法，最后建议使用互信息函数法求取延迟时间，利用改进伪邻近点法的CAO方法确定最佳嵌入维数。通过考毕兹吸引子的分析，验证了方法的有效性，并且应用到了健康人的声音时间序列中。对健康人的声音时间序列重构相空间并且画出了三维空间的奇怪吸引子图像。之后从Lyapunov指数出发，研究了得到最大Lyapunov指数的算法，建议用最小数据法求取最大Lyapunov指数。用最小数据法求得考毕兹吸引子和健康人的声音时间序列的最大Lyapunov指数，直接证明了两者时间序列都具有混沌特性。

# 第四章 改进的替代数据法在嗓音分析中的应用

## 4.1 替代数据法

在现实生活中，存在很多的时间序列，这些时间序列的形态呈现出复杂性和多样性。一般人们对这些时间序列采用两种方法进行处理：一种是以统计学为基础对其建立线性模型。另一种是对时间序列进行非线性动力学分析。对于随机序列一般采用统计学方法进行线性分析，而对于具有混沌性质的序列，必须要用非线性方法分析，这种情况通常有两种方法：一种是直接对时间序列提取混动动力学特征，另外一种是通过检测数据中的非线性成分，间接分析其混动特性。在直接方法中，已有很多有效的方法。比如分频采样法，庞加莱截面法，主分量分析法，关联维数法，功率谱法，最大Lyapunov指数法等等。间接方法有一种是1992年Theiler等人提出的替代数据法[54,55]，这种方法通过对比替代数据和原始数据对时间序列判断是否具有非线性性质。该方法在对于像天气、股票等复杂的实际观测时间序列中的非线性分析中比较常用[56]。替代数据法的主要步骤分为三部分，第一步为提出零假设，第二步为生成替代数据，第三步为检验统计量。提出零假设即对时间序列的特征进行一种假设性的猜想，不同的零假设会对应不同的产生替代数据的算法。之后再计算统计量，原则上零假设是可以与任何一种检测统计量一起使用来判断原始时间序列的性质。如果原始序列和其替代数据序列在统计量上相差很大说明该零假设被拒绝，从而表示原始序列和替代数据序列不同，反之则说明零假设被接受，原始序列和替代数据序列基本一致。

一般有一下几种零假设[55]：

零假设1 原始数据是分布相同但是相互独立的随即变量生成。

此种零假设的替代数据是用伪随即数发生器产生高斯白噪声，之后再对实验数据按照噪声序列的次数重新排列而成。这种方法曾经用于股票市场分析中。

零假设 2 原始数据是有Ornstein-Uhlenbeck过程生成。

此种零假设的替代数据是由原始数据的均值，方差和自相关函数拟合成的方程迭代得到的时间相关性序列。

零假设3 原始数据由具有原始数据的均值和方差的线性相关高斯过程产生。

零假设3可以用来检验原始数据中是否有非线性成分，它可以用以下自回归模型表示：

(4.1)

式中，为均值为0、方差为1的高斯白噪声。

此零假设的替代数据，一般可以由两种方法得到：一种是直接对上式进行不断的迭代产生替代数据，但是用产生的替代数据拟合式中的系数时会产生误差，导致结果容易发散，所以这种方法不稳定；另一种方法是由Theiler等人提出的用傅里叶变换将数据的相位随机化打散而得到替代数据的处理方法，这种方法比较稳定，也是常用的方法[57]。这种方法的基本思想是，替代数据是由原数据经过傅里叶变换后，经过相位随机化，然后再进行傅里叶反变换得到。这样替代数据同原数据保持了一样的功率谱、均值、方差等特征，但是有不同的幅值直方图。

生成替代数据的具体步骤如下：

第一步：对原始数据进行傅里叶变换，设原始数据为，则它的离散傅里叶变换为：

(4.2)

第二步：对得到的各个频率出的变换值进行随机化处理：

(4.3)

其中从区间随机选取，并且要满足斜对称条件，使得，以确保傅里叶逆变换的结果为实数。

第三步：进行傅里叶逆变换：

(4.4)

即为所求的替代数据。

零假设4 观测数据由线性相关的高斯噪声经静态非线性变换生成。

原则上，零假设可以同任何检测统计量一起使用，为了更好得比较原始数据和替代数据的检验统计量，一般采用sigma检验，即如下判据：

(4.5)

其中代表替代数据检验统计量的均值，代表替代数据检验统计量的均方差。一般取显著性水平为，则当时，表明原始数据和替代数据之间存在明显差别，则原始时间序列以95%的置信度为非线性时间序列。

## 4.2 关联维数的计算和应用

奇怪吸引子是轨道在相空间中经过无数次靠拢和分离，来回拉伸与折叠形成的几何图形，具有无穷层次的自相似结构。混沌体系是由这种奇怪吸引子来进行描述的，由于耗散系统在相空间的收缩，使得这种奇怪吸引子维数小于相空间的维数。奇怪吸引子的几何性质，一般通过研究它的空间维数来确定。奇怪吸引子的维数一般都是非整数，又因为它的组成部分与整体有某种方式的相似，所以这种不限于整数的维数称为分形维数，即分维数。分数维是20世纪70年代由曼德勃罗创立的。分维数也有很多中不同的定义：Hausdoff维数，关联维数，自相似维数，盒维数，Lyapnunov维数，信息维数等等。

### 4.2.1 计算关联维数

关联维数是最普遍的用来度量研究系统复杂性的一种维数量度。它刻画了相空间中点的分布，表示系统在多维空间的疏密程度，反映了系统中点的关联程度。系统越复杂，一般有较高的关联维数，在不确定系统中，关联维数无限制，但是在确定性系统中关联维数趋向于确定的值。1983年，Grassberger和Procaccia提出了从时间序列计算吸引子的G-P算法[58]。对单变量时间序列，构造维矢量，支撑起一个嵌入空间。只要嵌入维数足够高，就可以在拓扑等价的意义下恢复原来的动力学性态。

对于维重构的相空间，如下：

(4.6)

在构造好矢量后，定义其中任意两个矢量和的最大分量差作为距离 (4.7)

并且规定，凡是距离小于给定的正数的矢量，成为有关联的矢量。设重构相空间中有个点，计算其中的关联的矢量对数，它在一切可能的种配对中所占的比例为关联积分[59]

(4.8)

这里的为Heaviside单位函数：

(4.9)

已经知道，关联积分在时与存在以下

(4.10)

其中称为关联维数，恰当地选取使得能够描述混沌吸引子的自相似结构。计算关系式

(4.11)

在实际的运算中，通常给定一些具体的值（充分小），如果取得太小，已经低于环境噪声和测量误差造成的矢量差别，由计算关联维数的公式得到的就不是关联维数，而是嵌入维数。在实际的求解中，通常让从小到大递增，画出曲线，一般不考虑极小时的噪声区和极大时的饱和区，中间的直线部分的斜率便是关联维数。

图4.1中是求取关联维数的流程图：



图4. 1求关联维数流程图

### 4.2.2 对时间序列计算关联维数

根据上节中计算关联维数的G-P算法，对时间序列计算了其关联维数如下。

图4.2是对考毕兹振荡器（Colpittsoscillator）混沌电路产生的时间序列求取关联维数。图中的9条曲线是曲线在嵌入维数从2到10的曲线。取拟合图中的9条曲线形状相似的线性部分，可以从图中看出到的曲线，其线性部分的斜率基本保持相同。计算线性部分的斜率，就得到了关联维数。



图4. 2考毕兹振荡器时间序列的关联维数

图4.3中画出了不同嵌入维数下的关联维数的变化图，可以看出从之后关联维数趋于稳定。



图4. 3考毕兹振荡器时间序列嵌入维数—关联维数图

图4.4是对健康人的声音时间序列求取关联维数。图中的9条曲线是曲线在嵌入维数从2到10的曲线。取拟合图中的9条曲线形状相平行的直线部分，可以从图中看出到的曲线，其线性部分的斜率基本保持相同。计算线性部分的斜率，就得到了关联维数。



图4. 4健康人声音时间序列的关联维数

图4.5中画出了不同嵌入维数下的关联维数的变化图，可以看出从之后关联维数趋于稳定。



图4. 5健康人声音时间序列嵌入维数—关联维数图

关联维数是描述混沌特征的几何不变量之一，成功得提取关联维数对我们认识整个系统的混沌性质有着重要的意义。

## 4.3 改进替代数据法得到归一化检测量

### 4.3.1 基本概念

上节中已经介绍了替代数据法在非线性时间序列判别中的应用。混沌信号的特征同信号的相位和幅度密切相关。其中相位更加具有决定性作用，因为相位的一点点失真和畸变能够干扰到混沌信号的规律性。替代数据法也正是利用这个来获得区别线性信号和非线性信号的判别标准的方法。近年来，研究者将替代数据法同关联维数和Lyapunov指数等混沌特征量相结合，取得了很好的效果[60,61]。

本节中采用时间序列的关联维数作为检验统计量，Aboozar Taherkhani等人提出在混沌时间序列中加入噪声能够削减时间序列的确定性，使得时间序列的关联维数增大。替代数据法能够测量这种噪声的增加量。

### 4.3.2归一化方差的检测量的提出和计算

上节中提到加入噪声的时间序列的关联维数会增大，本节中用替代数据法计算该加入噪声时间序列的替代数据的关联维数，并且同计算加入噪声的时间序列的关联维数共同提取检测统计量，对加入不同比例的噪声查看该检测统计量的变化。用到的混沌时间序列是洛伦兹系统产生的时间序列，如图4.6所示。



图4. 6洛伦兹混沌时间序列

用洛伦兹系统产生的时间序列进行对比试验，分别加入SNR=10的白色噪声和SNR=100的白色噪声产生新的时间序列，再用替代数据法结合关联维数对新的两组时间序列分析，得到图4.7和图4.8如下，图中虚线表示了替代数据的关联维数曲线，实线在两个图中分别代表了增加SNR=10噪声和SNR=100噪声的洛伦兹时间序列的关联维数曲线。



图4. 7加入SNR=10白色噪声下的嵌入维数—关联维数图



图4. 8加入SNR=100白色噪声下的嵌入维数—关联维数图

从图中可以看出加入噪声的两组时间序列的关联维数，加入50%噪声的时间序列的关联维数明显高于加入20%噪声的时间序列的关联维数，而且跟替代数据的关联维数曲线的差别越来越小。由此也用了替代数据的sigma方法来刻画这个差别。将每一个嵌入维数下的替代数据和原始数据的关联维数值做差，之后求出均值，来刻画差别，称为归一化方差的检测量，其公式如下：

(4.12)

其中是重构的相空间嵌入维数的最大值。

### 4.3.3 声音时间序列的归一化方差的检测量计算

将归一化方差的检测量公式应用于健康人的声音时间序列和脑卒中病人的声音时间序列。脑卒中病人的声音时间序列相比于健康人的声音时间序列破坏了更多的系统中确定性的性质，因为大脑的受损，使得控制语言中枢的神经有所损伤，从而其发出的声音中加入了不确定的噪声。

图4.9中左侧图为健康人的声音时间序列和三组替代数据时间序列的关联维数曲线，右侧图为脑卒中病人的声音时间序列和三组替代数据时间序列的关联维数曲线。从图中可以看出脑卒中病人的原始声音时间序列的关联维数增大，这是因为脑卒中病人发出的声音呈现出明显的波动，在共振峰的位置、声门频率以及声音经由声道的传递上有了一定的改变，所以使得吸引子的几何特征变化，从而关联维数增大。



图4. 9原数据和替代数据嵌入维数—关联维数图

同样，用归一化方差的检测量分别对健康人声音时间序列和脑卒中病人时间序列计算，相比于这两个时间序列原始数据序列的关联维数曲线的差别，用归一化方差的检测量更加能够反映出差别。

## 4.4 本章小节

本章首先从非线性时间序列分析的基本问题——时间序列的非线性检验入手，引入了替代数据法这种间接的非线性验证方法，给出了4种常用的零假设及其相应的替代数据生成的方法。然后，从混沌吸引子具有的特殊的分形结构出发，研究了普遍使用的关联维数的求取算法，用G-P算法求取关联维数，对考毕兹吸引子和健康人的声音时间序列分别计算了关联维数，验证了信号的混沌性。接着提出了将替代数据中的检验统计量同非线性时间序列的特征量之关联维数相结合，得到了新的特征量即归一化方差的检测量，这个特征量计算了原始数据和替代数据在不同嵌入维数下的关联维数的均值方差。通过加入不同百分比噪声的洛伦兹吸引子时间序列进行了验证，说明这种方法能够反映出时间序列的混沌性质。之后对健康人和脑卒中病人的声音时间序列用新的特征量进行对比，得出了他们的声音时间序列之间确实存在的差异。

# 第5章 脑卒中病人和健康人声音样本分析和分类

在本文的第3章和第4章中介绍了非线性方法在分析混沌时间序列中的应用，这一章中，将第3章和第4章中的方法应用到采集到的脑卒中病人和健康人的声音时间序列中，对比了非线性序列的基本特征量如互信息图的第一个最小值、最大Lyapunov指数、关联维数、还有通过将替代数据法和关联维数结合得到的归一化方差的检测量。通过这些特征量的对比可以看出脑卒中病人因为脑损伤确实改变了发出的声音时间序列的非线性特性，从而使得通过声音时间序列分析得到大脑状态的度量成为可能。

## 5.1 非线性参数在脑卒中病人和健康人样本中的对比

### 5.1.1 互信息图的第一个最小值对比

图5. 1中是健康人和脑卒中病人声音时间序列用互信息法求延迟时间时求得的函数曲线对比图。在图中的箭头标示出了第一个最小极值对应的延迟时间和值。可以看出健康病人的声音信号和延迟序列的互信息函数的最小值比脑卒中病人的最小值大，说明健康病人的原始声音序列和延迟序列之间的差异比脑卒中病人要小，脑卒中病人的声音序列更加不规律。



图5. 1互信息图第一个最小值对比图

分别对采集到的17位脑卒中病人声音时间序列和20例健康人声音时间序列进行非线性分析提取的互信息法最小值如下图5. 2所示，其中倒三角形代表了健康人声音时间序列的数据，星号代表了脑卒中病人声音时间序列的数据。



图5. 2互信息图的第一个最小值统计图

从图中可以看出脑卒中病人声音信号的互信息图第一个最小值集中的区间是[1.03,1.51],健康人声音信号的互信息图第一个最小值集中的区间是[1.32,1.76],由此可知脑卒中病人的互信息图第一个最小值平均小于健康病人的，这意味着在声音序列和其延迟声音序列到达最大差异的时候，脑卒中病人的这个差异大于健康人，也说明了脑卒中病人的声音序列更加不规律。

### 5.1.2 重构相空间的吸引子对比

图5. 3中展示了健康人声音时间序列和脑卒中病人声音时间序列在重构的相空间中的吸引子，分别根据互信息法和CAO的方法求得最佳的嵌入维数，构造出相空间，得到的吸引子。从图中可以看出健康人声音的吸引子更加有规律，而脑卒中病人的声音信号的吸引子有很大的波动和起伏，更加杂乱。



图5. 3二维空间吸引子对比图

### 5.1.3 最大Lyapunov指数对比

分别对采集到的17位脑卒中病人声音时间序列和20例健康人声音时间序列进行非线性分析提取的最大Lyapunov指数如下图5. 4所示，其中倒三角形代表了健康人声音时间序列的数据，星号代表了脑卒中病人声音时间序列的数据。



图5. 4最大Lyapunov指数对比图

从图中可以看出健康人声音信号的Lyapunov指数集中区间为[0.002 1,0.003 9], 脑卒中病人声音信号的Lyapunov指数集中区间是[0.003 1,0.009 8],由此看出健康人声音信号的Lyapunov指数平均小于脑卒中病人声音信号的Lyapunov指数，说明脑卒中病人的声音的混沌性高于健康病人。以上分析结果也说明了脑卒中病人发出的声音更加复杂，并且波动性很高。

### 5.1.4 归一化方差的检测量对比

分别对采集到的17位脑卒中病人声音时间序列和20例健康人声音时间序列进行非线性分析提取归一化方差的检测量。统计结果如下表所示。

表5. 1中是17位脑卒中病人声音时间序列的关联维数，替代数据时间序列的关联维数均值和归一化方差的检测量。

表5. 2中是20位健康人声音时间序列的关联维数，替代数据时间序列的关联维数均值和归一化方差的检测量。

表5. 1脑卒中病人的替代数据分析表

|  |  |  |
| --- | --- | --- |
|  | <> | S |
| 2.528 5 | 3.292 2 | 18.624 8 |
| 2.638 6 | 3.506 9 | 10.533 5 |
| 2.440 0 | 3.565 7 | 24.558 4 |
| 2.990 0 | 3.407 0 | 6.051 4 |
| 2.540 6 | 3.557 4 | 12.476 5 |
| 2.623 8 | 3.486 8 | 14.362 6 |
| 3.581 0 | 3.107 0 | 14.224 3 |
| 2.945 5 | 3.379 8 | 6.632 2 |
| 2.843 7 | 3.280 5 | 13.245 5 |
| 2.815 3 | 3.293 7 | 7.743 8 |
| 2.616 9 | 3.393 0 | 5.490 7 |
| 2.954 8 | 3.087 5 | 3.505 5 |
| 2.856 9 | 3.508 8 | 21.62 4 |
| 3.066 8 | 3.157 6 | 14.597 0 |
| 2.927 5 | 3.607 1 | 17.576 1 |
| 2.892 6 | 3.077 8 | 3.380 1 |
| 3.057 6 | 4.253 0 | 27.139 3 |

表5. 2脑卒中病人的替代数据分析表

|  |  |  |
| --- | --- | --- |
|  | <> | S |
| 2.214 1 | 3.450 5 | 46.422 8 |
| 2.030 6 | 3.454 9 | 21.573 6 |
| 2.316 6 | 3.437 5 | 38.226 0 |
| 2.011 6 | 3.134 8 | 52.338 3 |
| 1.991 9 | 3.323 2 | 55.665 7 |
| 1.756 5 | 3.700 7 | 45.413 4 |
| 2.109 6 | 3.245 7 | 35.228 5 |
| 2.186 5 | 3.206 5 | 39.097 9 |
| 1.948 5 | 3.059 4 | 40.963 3 |
| 1.937 6 | 3.092 2 | 55.351 2 |
| 1.652 1 | 3.342 7 | 85.660 9 |
| 2.074 0 | 3.250 2 | 28.780 7 |
| 1.638 0 | 3.160 0 | 39.023 5 |
| 1.996 6 | 3.281 9 | 42.675 2 |
| 2.314 8 | 3.314 1 | 104.720 4 |
| 1.989 4 | 3.526 3 | 102.628 0 |
| 1.918 6 | 3.057 9 | 28.798 6 |
| 1.745 9 | 3.123 0 | 61.234 5 |
| 1.834 8 | 3.324 7 | 76.345 2 |
| 2.096 4 | 3.412 8 | 67.834 6 |

从两个表中可以看出，脑卒中病人的原始声音序列的关联维数高于健康人声音序列的关联维数，这是因为脑卒中病人发出的声音呈现出明显的波动，在共振峰的位置、声门频率以及声音经由声道的传递上有了一定的改变，所以使得吸引子的几何特征变化，从而关联维数增大。另外，可以看出脑卒中病人和健康人的的变化范围是1.7-3.5，而归一化方差的检测量的变化范围是3.5-104.7，所以归一化方差的检测量能够更好得区分脑卒中病人和健康人。通过实验结果和脑卒中病人病历的对比分析发现，在脑卒中病人中根据脑损伤的不同，发音所受影响也不尽相同，脑损伤越严重的病人在提取声音的各项指标中跟正常人有更加明显区别。

## 5.2 k近邻分类算法

### 5.2.1 什么是K近邻算法

何谓K近邻算法，即K-Nearest Neighbor algorithm，简称KNN算法，单从名字来解释，可以简单的认为是：K个最近的邻居，当K=1时，算法便成了最近邻算法，即寻找距离最近的那个邻居。为何要找到邻居？假如你来到一个陌生的村庄，你需要找到与你有相似特征的人群加入他们，所谓入伙。

官方解释来说，所谓K近邻算法，即是给定一个训练数据集，对新的输入数据实例，在训练数据集中找到与该数据实例最邻近的K个实例（也就是上面所说的K个邻居），这K个数据实例的多数属于某个类，就把该输入实例分类到这个类中。根据这个说法，来看下引自维基百科上的一幅图如图5.5所示：

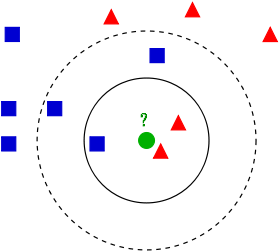


图5. 5 k近邻算法示意图

如图5.5所示，用蓝色的小正方形和红色的小三角形分别表示两类不同的样本数据集合，而图正中间绿色的圆所代表的数据则是待分类的数据。也就是说，我们不知道中间那个绿色的数据是归类于哪一类（蓝色正方形或是红色三角形），下面我们就要解决这个问题：给这个绿色的圆归类。

我们常说：物以类聚，人以群分。判断一个人是一个怎样品质特征的人，常常可以从其身边的朋友入手，所谓观其友，而识其人。那么在我们的归类中，如果要判别上图中那个绿色的圆是属于哪一类数据，就要从它的相邻数据集下手。但一次性看多少个相邻数据呢？从上图中，可以看到：

如果K=3，绿色圆点的最近的3个数据是2个红色小三角形和1个蓝色小正方形，基于统计中少数从属与多数的方法，判定这个绿色待分类点属于红色的三角形一类。

如果K=5，绿色圆点的最近的5个数据是2个红色三角形和3个蓝色的正方形，还是基于统计中少数从属于多数的方法，判定这个绿色待分类点属于蓝色的正方形一类。

于此我们看到，当无法判定当前待分类点是从属于已知数据集类别中的哪一类时，我们可以依据统计学的理论看它所处的位置特征，衡量它周围邻居的权重，而把它归类到权重更大的那一类。这就是K近邻算法的核心思想。

### 5.2.2 k近邻的距离度量表示法

上一小节中，我们看到，K近邻算法的核心在于找到实例数据点的邻居，这个时候，问题就接踵而至了，如何找到邻居，邻居的判定标准是什么，用什么来度量。这一系列问题便是下面要讲的距离度量表示法。

特征空间中两个实例数据点的距离反应出两个实例数据点之间的相似性程度。K近邻模型的特征空间一般是n维实数向量空间，使用的距离可以是欧式距离，也是可以是其它距离，本节中采用的是最基本的欧式距离如下：

欧氏距离是最常见的两点之间或多点之间的距离表示法，又称之为欧几里得度量，它定义于欧几里得空间中，如点和 之间的距离为：

(5.1)

1. 二维平面上两点与间的欧氏距离：

(5.2)

1. 三维空间两点与间的欧氏距离：

(5.3)

1. 两个n维向量与间的欧氏距离：

(5.4)

　也可以用表示成向量运算的形式：

(5.5)

### 5.2.3 K值的选择

上一节中讨论了如何定义邻居的问题，还有一个选择多少个邻居，即K值定义为多大的问题。K值选择问题对K近邻算法的结果会产生重大影响。如李航博士的一书《统计学习方法》上所说：

如果选择较小的K值，就相当于用较小的领域中的训练数据实例进行预测，学习的近似误差会减小，只有与输入数据实例较近或相似的训练数据实例才会对预测结果起作用，与此同时带来的问题是学习的估计误差会增大，换句话说，K值的减小会使整体模型变得复杂，容易发生过拟合；

如果选择较大的K值，就相当于用较大领域中的训练数据实例进行预测，其优缺点同选择较小的K值时相反，优点是可以减少学习的估计误差，缺点是学习的近似误差会增大。这时候，与输入数据实例较远（不相似的）的训练数据实例也会对预测器产生作用，使的预测产生错误，且K值的增大会使得整体的模型变简单。

将K值取为所有训练数据实例的个数，则是完全不可取的，因为此时无论输入实例是什么，都只是简单的预测它属于在训练实例中最多的那一类，模型过于简单，忽略了训练数据实例中大量有用信息。

在实际的分类中，K值一般取一个比较小的数值，例如采用[交叉验证](http://zh.wikipedia.org/zh/%E4%BA%A4%E5%8F%89%E9%A9%97%E8%AD%89)法（即一部分样本做训练集，一部分做测试集）来选择最优的K值。

## 5.3 脑卒中病人和健康人声音样本分类

对采集到的17例脑卒中病人和20例健康人的语音数据样本集中，随机取12例脑卒中病人声音数据和14例健康人声音数据作为训练样本集，剩余样本为测试样本集。本实验中采用互信息法第一个最小值，最大Lyapunov指数，关联维数和归一化方差检测量4维特征量作为分类特征量，用K近邻方法对数据集进行分类。

为了验证新的特征量即归一化检测量可以提高分类正确率，将非线性特征量分为两组特征组合以进行实验，第一组即互信息图第一个最小值，最大Lyapunov指数和关联维数，第二组即互信息图第一个最小值，最大Lyapunov指数和归一化方差检测量。

由于实验数据比较少，所以本文中采用3组循环训练样本集和测试样本集进行分类实验，及循环在总的样本集中抽取训练样本集和测试样本集以得到分类正确率。其分类结果如下。

表5. 3 第一组特征量组合分类结果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 样本组号 | 训练样本数 | 测试样本数 | 训练正确度 | 测试正确度 |
| 1 | 26 | 11 | 98.4% | 88.4% |
| 2 | 26 | 11 | 100% | 89.1% |
| 3 | 26 | 11 | 95.5% | 91.5% |

表5. 4第一组特征量组合分类结果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 样本组号 | 训练样本数 | 测试样本数 | 训练正确度 | 测试正确度 |
| 1 | 26 | 11 | 99.1% | 94.6% |
| 2 | 26 | 11 | 100% | 98.5% |
| 3 | 26 | 11 | 99.8% | 97.9% |

表5.3和表5.4分别为用第一组特征量组合（互信息图第一个最小值+Lyapunov指数+关联维数）和第二组特征量组合（互信息图第一个最小值+Lyapunov指数+归一化方差检测量）应用K近邻算法得到的分类结果。从两个表中对比可以看出用归一化方差检测量代替关联维数后其测试样本分类的正确度明显提高，说明这个归一化方差检测量能够很好得反映脑卒中患者声音特征和健康人声音特征的区别，提高分类正确度。

## 5.4 本章小节

本章用第3章和第4章给出的非线性方法运用到了采集到的17例脑卒中病人和20例健康人的声音时间序列中，提取了互信息法的最小值，最大Lyapunov指数，归一化方差检测量，并且进行了对比，可以清晰地看出脑卒中病人的声音时间序列和健康人的声音时间序列的不同，并从脑损伤方面对实验结果作出解释。之后介绍了K近邻分类算法，并且将该算法应用到区别脑卒中患者的声音和健康人的声音分类中。最后对分类结果进行分析得到了归一化方差检测量更能够区分两类声音。

# 第6章总结与展望

## 6.1 研究工作总结

在广泛研究阅读了国内外关于大脑状态分析与探讨和声音分析处理技术相关文献的基础上，提出了一种用非线性分析脑卒中病人声音信号的方法，以期能够用声音特征量度量人体大脑的状态。

本文主要分为以下几个部分：脑卒中病人声音采集、脑卒中病人声音信号分析处理、故障诊断特征量的构造和结果分析。本文首先介绍了课题的研究背景、现状和声音分析技术的优缺点以及大脑状态研究的现状，说明了应用声音分析大脑损伤状态的可行性和重要意义。之后第二章中详细介绍了大脑和声音产生的关系，并且从神经机制和脑成像方面证明了声音受大脑控制产生。接下来第三章介绍了非线性方法对声音信号的分析应用，对声音时间序列重构相空间和吸引子，提取相应非线性特征量。第四章提出了用替代数据法和关联维数相结合得到的归一化方差检测量。最后第五章中，将非线性分析的方法将应用到了所有采集到的脑卒中病人和健康人的声音时间序列中，进行统计分析对比。通过数据对比发现，脑卒中病人的声音时间序列提取的特征量反映出其序列比健康人声音时间序列更加杂乱无规律，并且给出了反映大脑状态的解释。

本文的主要研究结果和创新点如下：

1）对脑卒中病人采用声音分析的方法，为医疗工作者提供辅助诊断手段。

2）采用替代数据法结合混沌非线性方法对声音进行分析提出新的非线性特征量。

3）提供可以量化人体大脑功能障碍的参数，对脑功能损伤病人康复程度做出评估。

4）采用混沌非线性方法研究大脑神经系统的非线性，对非线性在生物医学领域的研究有重要意义。

5）通过声音信号分析大脑损伤状态，为以后用声音度量大脑状态提供了研究方向。

## 6.2 进一步工作展望

尽管本论文在用声音分析度量脑损伤方面进行了有益的研究和探索，但是由于本人水平和研究时间的限制，还是有很多工作没有完成，有待深入研究。从研究工作的延续和进一步完善的角度看，以下几个方面是未来可以继续工作和研究的方向。

1）本文的实验数据因为客观条件的限制，所采集到的脑卒中病人的样本偏少，这方面也涉及到了病人的隐私，所以不易采集，可以考虑和更多的医院合作，获取更多的声音样本数据。

2）本文中只对比了健康人和脑卒中病人的声音信号的特征量，如果声音分析可以真正度量到大脑状态，可以对某个脑卒中病人跟踪观察，从病人发作初期到病情渐渐稳定，跟踪其声音时间序列的变化，从而得到真正能够度量大脑状态的特征量，度量大脑状态。

# 参考文献

[1] Krzesimowski D, Ciota Z. Voice Signal Processing for Patients with Stroke Hospitalization [C]// 16th International Conference “Mixed Design of Integrated Circuits and Systems”, Poland,2009.

[2] Shiomi K. Voice Processing Technique for Human Cerebral Activity Measurement [C]// IEEE 1-4244-2384-2/08 2008.

[3] Shiomi K, Hirose S. Fatigue and Drowsiness Predictor for Pilots and Air Traffic Controllers [C]// Proc of 45th Annual ATCA Conference, 2000.10.

[4]Babloyantz A,Salazar J M,Nicolis C.Evidence of Chaotic dynamics of brain activity during the sleep cycle[J].Phys Lett,1985,IIIA:152-156.

[5]吴朝晖，杨莹春.说话人识别模型与方法[M].北京：清华大学出版社.2009.

[6]刘幺和，宋庭新.语音识别与控制应用技术[M].北京：科学出版社,2008.

[7]于燕平，胡维平.病态嗓音特征的小波变换提取及识别研究[J].计算机工程与应用.2009,45(22) 194-196.

[8]王修信，胡维平.基于小波变换的相对信噪比在喉疾病检测中的意义[J].临床研究.2002,10(4) 225-226.

[9]颜景斌.基于连续小波和支持向量机的病态嗓音检测[J].电脑与信息技术.2008,16(3) 21-22 39.

[10]刘晨轩,蓝贤桂. 语音信号短时分析算法研究与实现[J]. 价值工程, 2012,(12):191-192.

[11]何俊,李艳雄,贺前华. 变异特征加权的异常语音说话人识别算法[J].华南理工大学学报(自然科学版), 2012,40(3):106-111.

[12]科恩，白居宪．时一频分析：理论与应用[M]．西安：西安交通大学出版社，1998．

[13]Gabor D．Theory ofcommunication[J]．Part l：The analysis of information，ElectricalEngineers—Part III：Radio and Communication Engineering．Journal ofthe Institution of,1946，93(26)：429-441．

[14]吴湘淇．信号、系统与信号处理[M]．北京：电子工业出版社，1996．

[15]Cohen L．Generalized phase-space distribution functions[J]．Joumal of Mathematical Physics，1966，(7)．781-786.

[16] Daubechies I．Orthonormal bases ofcompactly supported wavelets II[J]．Variations on atheme，SIAM Journal on Mathematical Analysis，1 993；24(2)：499-519．

[17] Daubechies I．Ten lectures on wavelets[M]．SIAM，1992.

[18]杨福生．小波变换的工程分析与应用[M]．北京：科学出版社，1999.

[19] 彭玉华．小波变换与工程应用[M]．北京：科学出版社，1999.

[20] 高志，余啸海．Matlab小波分析与应用[M]．北京：国防工业出版社，2007.

[21] Morlet J，Areas G Fourgeau E，et a1．Wave propagation and sampling theory；Part I，Complexsignal and scatteringin multilayered media[J]．Geophysics，1982，47(2)：203-21．

[22]Dibazar A A, Park H O, Berger T W. Nonlinear dynamic modeling of impaired voice[C] //32nd Annual International Conference of the IEEE EMBS, Buenos Aires, Argentina, 2010.

[23]Orozco J R, Vargas J F, Alonso J B. Voice pathology detection in continuous speech using nonlinear dynamics [C]// Information Science, Signal Processing and their Applications (ISSPA), 2012.

[24]洪时中．非线性时间序列分析的最新进展及其在地球科学中的应用前景．1999，14(6)：559

[25] T．Schreiber．Interdisciplinary application of nonlinear time series methods．Physics Reports．1999，308：1-64

[26]王晨.基于非线性时间序列的胎儿心电信号分析与提取[J].

[27]Maria Eugenia Dajer, Jose Carlos Pereira ,Carlos Dias Maciel.Nonlinear Dynamical Analysis of Normal Voices[C].Proceedingsof the Seventh IEEE International Symposium on Multimedia,IEEE.2005.

[28]A.C.K.Soong,Z.J.Koles.Principal-Component Location of the Sources of the Background EEG.IEEE Transactions on Biomedical Engineering.1995,vol.42,59-67.

[29]Y.Wanga,P.Bergc,M.Scherga,Common Spatial Subspace Decomposition Applied to Analysis of Brain Responses under Multiple Task Conditions:A Simulation Study.Clinical Neurophysiology,1999,vol.110,604-614.

[30]汪铭虎，王琳.脑与语言认知[M].北京：清华大学出版社,2013.

[31]单春雷.生理心理学[M].北京师范大学出版社.

[32]M. F. Bear, B. W. Connors, M. A. Paradiso, Neuroscience: Exploring the Brain, 2nd ed[M]. 2001.

[33]Sun Tao, Walsh Christopher A. 2006. Molecular approaches to brain asymmetry and handedness[J]. Nature Reviews Neuroscience, 7: 655-662.

[34]刘曾荣，文铁桥，姚晓东.脑与非线性动力学[M].北京：科学出版社,2006.

[35]Haruhiko Terai，Masashi Shimahara.Evaluation of speech intelligibility after a secondary dehiscence operation using an artificial graft in patients with speech disorders after partial glossectomy[J].Br J Oral Maxillofac Surg．2004 Jun；42(3)：190-194.

[36]T.Sabesan，P.L.Ramchandani，V.llankovan.Sensory recovery of noninnervated free flap in oral and oropharyngeal reconstruction[J].Int J Oral Maxillofac Surg．2008Sep；37(9)：819-823.

[37]J．Cheol-Woo，K．Dae-Hyun.Classification ofpathological voice into normal／benign／malignantState[J].Proceedings ofEurospeech,Budapest，Hungary，1999(1)：571-574.

[38]S Linville，E Korabic，M Rosera，Intraproduction variability in jitter measures from elderlySpeakers[J].In：J．Voice，1990(4)l：45-51.

[39]魏春生，陈锋，王薇.国际音标/ai/和/a/的选择对声学测试参数的影响[J].临床耳鼻咽喉科杂志，1999(3)：102-103.

[40]侯丽珍，韩德民，徐文.嗓音检测中元音声样的选择[J].听力学及言语疾病杂志，2002，10(1)：16-18.

[41]邱天爽,唐洪. 统计信号处理:医学信号分析与处理[M].北京科学出版社.2012,ISBN: 978-7-03-033286-8.

[42]丁晓蓉,郭兴明,钟丽莎. 基于混沌理论的心音信号非线性动力学分析[J]. 振动与冲击, 2012,31(15):55-59.

[43]王兴元. 复杂非线性系统中的混沌[M].电子工业出版社.2003,ISBN:9787505387942.

[44]吕金虎,陆君安. 混沌时间序列及其应用[M].武汉大学出版社.2002.

[45] N．H．Packard，J．E Crutchfield，J．D．Farmer．Geometry from a time series[J]．Pbysical ReviewLetters．1980，45(9)：712-716.

[46]Floris T, Detecting strange attractors in turbulence [J].NewYork:Springer,1981,898:366-381.

[47]M．T．Rosenstein，J．J．Collins，C．J．De Luca．Reconstruction expansion as a geometry—basedframework for choosing proper delay times．Physica D．1994，73：82-98.

[48]杨志安，王光瑞，陈式刚．用等间距分格子法计算互信息函数确定延迟时间[J]．计算物理．1995，12(4)：442-447.

[49]Fraser A M, Swinney H L. Independent coordinates for strange attractors from mutual information [J]. American Physical Society,1986,33(2): 1134-1140.

[50] Cao L. Practical method for determining the minimum embedding dimension of a scalar time series [J]. Physica D: Nonlinear Phenomena,1997,1(10): 43-50.

[51]A．Wolf,J．B．Swift H．L．Swinney．Determining Lyapunov exponents from a time series．Physica 16D．1985，285-317.

[52]GBarana,I．Tsuda．A new method for computing Lyapunovexponents．Physics Letter A．1 993175：421-427.

[53]M．T．Rosenstein,J．J．Collins，C．J．De Luca．A practical method for calculating largestLyapunov exponents from small data sets．Physica D．1993，65：117-134.

[54]T．Schreiber,A．Schmitz．Improved surrogate data for nonlinearity tests．Physical ReviewLetters．1996，77(4)：635-638.

[55]J．Theiler,S．Eubank，A．Longtin．Testing for nonlinearity in time series：the method ofsurrogate data．Physica D．1992，58：77-94.

[56]J.A.Scheinkman,B.Lebaron.Nonlinear dynamics and stock returns.Journal of Business,1989,62(3),311-337.

[57]J.Theiler,S.Eubank,A.Longtin.Testing for nonlinearity in time series,the method of surrogate data.Physica D,1992,58,77-94.

[58]P．Grassberger,I．Procaccia．Measuring the strangeness of strange attractors[J]．Physica D．1983．9：189-208.

[59]Kantz H, Schreiber T. Nonlinear Time Series Analysis[C]// Cambridge,U.K:Cambridge Univ Press,1990.

[60]王立媛,刘玉萍,肖青,等. 胎儿心率信号的替代数据分析[J].长春理工大学学报(自然科学版), 2007,30(1):72-75.

[61]Taherkhani A, Seyyedsalehi S A, Mohammadi A. Nonlinear Signal Processing for Voice Disorder Detection by Using Modified GP Algorithm and Surrogate Data Analysis[C]// IEEE International Symposium on Signal Processing and Information Technology, 2007.

# 附录攻读硕士学位期间的研究成果

**攻读硕士学位期间录用论文：**

李江，赵雅琼，包晔华．基于混沌和替代数据法的脑卒中病人声音分析[J]．浙江大学学报（工学版）(录用时间为2013年11月，预计在2014年发表)