

学校代码: 10286

分类号:

密 级: 公开

U D C: _____

学 号: 130717



SOUTHEAST UNIVERSITY

东南大学

硕士学位论文

数字助听器的自适应验配技术的研究

研究生姓名: 冯超

导师姓名: 赵力 教授

申请学位类别 工学硕士

学位授予单位 东南大学

一级学科名称 信息与通信工程

论文答辩日期 20 年 月 日

二级学科名称 信号与信息处理

学位授予日期 20 年 月 日

答辩委员会主席

评 阅 人

20 年 月 日

東南大學

硕士学位论文

数字助听器的自适应验配技术的研究

专业名称: 信息与通信工程

研究生姓名: 冯超

导师姓名: 赵力

RESEARCH ON SELF-FITTING TECHNOLOGY OF DIGITAL HEARING AIDS

A Thesis Submitted to

Southeast University

For the Academic Degree of Master of

Science

BY

FengChao

Supervised by

ZhaoLi

School of Information Science and Engineering

Southeast University

April 2016

东南大学学位论文独创性声明

本人声明所呈交的学位论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，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，也不包含为获得东南大学或其它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了谢意。

研究生签名：_____日期：_____

东南大学学位论文使用授权声明

东南大学、中国科学技术信息研究所、国家图书馆有权保留本人所送交学位论文的复印件和电子文档，可以采用影印、缩印或其他复制手段保存论文。本人电子文档的内容和纸质论文的内容相一致。除在保密期内的保密论文外，允许论文被查阅和借阅，可以公布（包括以电子信息形式刊登）论文的全部内容或中、英文摘要等部分内容。论文的公布（包括以电子信息形式刊登）授权东南大学研究生院办理。

研究生签名：_____导师签名：_____日期：_____

摘要

听觉是人们日常生活中进行沟通交流的一大重要途径,听力的损失将严重影响个人的生活质量。然而,听力损失人群的比重却在不断上升。在现有医疗水平条件下,为矫正听损患者的听力状况,助听器的使用已经成为治疗的主要手段。传统的助听器验配过程相对繁琐,并且高度依赖听力专家的验配水平。近年来,助听器的自适应验配逐渐被大家提起。助听器的自适应验配是指由患者通过与相关软件的交互操作实现患者自行验配,中间无需听力专家的干预,打破了传统验配的束缚。

本文重点研究交互式遗传算法以及高斯过程回归方法在数字助听器自适应验配中的应用,在总结了前人已有工作成果的基础之上,提出了完整的算法实现流程,初步实现了数字助听器的自适应验配,取得了一些研究成果,主要研究工作如下:

(1) 介绍了人耳听觉系统、听力损失产生的原因、听力测试及治疗方法,分析了传统验配过程和自适应验配过程两种验配方式。

(2) 研究了基于交互式遗传算法的自适应验配技术。对各遗传算子所要进行的操作作出了具体的规定,包括染色体编码、选择操作、交叉操作以及针对听损类型的变异操作。在此基础之上提出了优化辅助模块,利用知识库中的现有数据辅助提升算法搜索效率。

(3) 研究了基于高斯过程回归的自适应验配技术。本文将人的听觉感知由不能直接测量的内部响应函数来表示,并将该过程用高斯过程来建模。在研究了高斯过程回归相关理论基础之上,本文进行了一系列理论推导,给出了用于预测验配者对当前参数喜好程度的公式。在此基础之上,引入了基于期望提高值(EI)的序贯设计,以找到当前“最有意义”的验配参数提供给验配者进行标记,从而形成一个交互式的循环验配流程。

(4) 针对本文提出的两种自适应验配方法,设计了相应的人机交互模块,包括纯音测试以及语音测试两个部分。为配合以上算法的研究,设计了16通道6阶IIR滤波器组,实验测试证明了该滤波器组性能良好,可以作为自适应验配中的语音信号处理模块。最后的语音识别率以及语音质量的测试再次证明了上述两种验配方法的良好性能。

关键词: 数字助听器, 自适应验配, 交互式遗传算法, 高斯过程回归, 多通道滤波器组

Abstract

Hearing is one of the most important way for people to communicate in daily life, but hearing loss will seriously affect the individual's normal life. However, the proportion of people with hearing loss is on the rise. Under the existing level of medical conditions, the use of hearing aids has become the primary means of treatment for hearing loss. The traditional fitting process of hearing aid is relatively cumbersome and highly dependent on hearing expert's experience. In recent years, self-fitting of hearing aids has been lifted gradually. Self-fitting means that patients with hearing loss complete the fitting process by themselves through interaction with related software without hearing expert's participation, which breaks the shackles of traditional fitting process.

This paper focuses the self-fitting technology based on interactive genetic algorithm and Gaussian process regression. Based on predecessors' achievement, two complete algorithm processes have been proposed which realizes self-fitting of hearing aids preliminarily. Some research results are concluded as follows:

(1) Human auditory system, the cause of hearing loss, hearing tests and treatments are studied in this paper. Two methods of fitting, traditional fitting and self-fitting, are discussed in detail.

(2) A self-fitting technology based on interactive genetic algorithm is elaborated. Specific provisions of each genetic operator's operation are made, including chromosome encoding, select operation, genetic crossover and mutation related to the type of hearing loss. Besides, a auxiliary module aimed at accelerating the evolution process is proposed, which takes advantage of the existing data in knowledge base to improve search efficiency.

(3) A self-fitting technology based on Gaussian process regression is proposed. In this paper, human auditory perception which can not be measured directly is expressed as a internal response function, and it is modeled by a Gaussian process. After studying the Gaussian process regression theory, a series of theoretical derivations are conducted, and a formula which can be used to predict one's preference over current fitting parameters is finally obtained. In addition, a sequential design based on EI (expected improvement) is done to get the "most meaningful" fitting parameters for one to mark, so as to form an interactive fitting loop.

(4) Appropriate interactive modules are designed corresponding to fitting methods mentioned above, including pure tone testing and speech testing. A 16-channel 6-order IIR filter bank is designed to tie in with the above algorithm research. Experimental tests show that this filter bank is good in performance which can be used as the signal processing module of a self-fitting system. Finally, the speech recognition experiment and voice quality test prove that the two methods this paper proposed can be up to self-fitting of hearing aids.

Keywords: digital hearing aids, self-fitting, Gaussian process regression, interactive genetic algorithm, multi-channel filter bank

目录

摘要.....	I
Abstract	III
目录.....	V
第一章 绪论.....	1
1.1 研究背景及意义	1
1.2 助听器技术和验配的发展研究现状	2
1.3 数字助听器基本原理	4
1.4 论文的研究内容和结构	5
第二章 助听器验配理论基础	7
2.1 人耳听觉系统	7
2.2 听力损失测试及治疗方法	8
2.2.1 听力损失	8
2.2.2 听力测试方法	10
2.2.3 听力损失治疗方法	12
2.3 传统助听器验配流程	13
2.4 助听器自适应验配原理	15
2.5 本章小结	17
第三章 基于交互式遗传算法的自适应验配方法	19
3.1 交互式遗传算法原理及自适应验配一般流程	19
3.1.1 交互式遗传算法原理	19
3.1.2 自适应验配流程	20
3.2 遗传算子的设计	22
3.2.1 染色体编码及选择操作	22
3.2.2 交叉操作	24
3.2.3 变异操作	26
3.3 辅助优化模块	28
3.3.1 用户偏好匹配	28
3.3.2 种群进化程度评估	30
3.4 实验仿真与主观测听评估	31
3.4.1 实验设置	31
3.4.2 优化变异仿真实验	32
3.4.3 主观测听实验	34
3.5 本章小结	36
第四章 基于高斯过程回归的自适应验配方法	37

4.1 高斯过程回归预测理论基础	37
4.1.1 高斯过程回归预测	37
4.1.2 高斯模型的训练	40
4.2 自适应验配的高斯过程模型	42
4.2.1 听觉心理响应函数的高斯建模	42
4.2.2 超参数的求解	45
4.2.3 预测输出	46
4.3 序贯设计及本章算法流程	47
4.3.1 基于期望增益的序贯设计	47
4.3.2 本章算法流程	48
4.4 主观测听评估	49
4.5 本章小结	52
第五章 自适应验配软件系统的设计及实验测试	53
5.1 自适应验配总体架构	53
5.2 人机交互模块设计	53
5.2.1 纯音测试	54
5.2.2 语音测试	54
5.3 语音信号处理模块	56
5.3.1 滤波器组的设计	56
5.3.2 测试与仿真	58
5.4 言语测听实验	61
5.4.1 可懂度测听实验	61
5.4.2 语音质量测听实验	63
5.5 本章小结	64
第六章 总结与展望	65
6.1 本文已取得的研究成果	65
6.2 展望	65
致谢	67
参考文献	69

第一章 绪论

1.1 研究背景及意义

听觉是人类赖以生存的重要感官之一，是人与人、人与世界沟通的桥梁。正是因为有了听觉，人们可以倾听亲朋好友的欢声笑语，可以欣赏美妙动听的歌声音乐。如果没有了听觉，我们会失去多姿多彩的生活，甚至会对周围环境感觉陌生而缺乏安全感。因此，听觉的重要性对于每个人而言都是不言而喻的。然而，随着科学技术的不断进步，人类文明的不断发展，产生了越来越多对人们听力有害的因素（如工人使用电钻施工的声音），伴随着人口老龄化以及耳朵疾病的影响，听力损失人群的比重不断攀升^[1]。据第二次全国残疾人抽样调查显示，我国的听力残疾人口已经超过 2000 万，占残疾人总数的 34%，在五类常见的身体残疾人士中（视力、听力、智力、肢体和精神），听力残疾人的数量占据首位。如何改善这些听力损失人群的听力状况已经成为迫在眉睫的重要课题，对我国经济社会的稳定发展以及和谐社会都具有重大的意义^[2]。

在现有医疗水平条件之下，助听器的使用已经成为治疗矫正听损患者听力状况的主要手段。助听器的发展已经有了上百年的历史，助听器也从模拟时代进入到了数字时代。现在的助听器采用 DSP 芯片对语音信号进行实时处理，具有体积小、延时低、低功耗以及高性能的优点。理论上，目前的助听器性能已经可以满足大部分具有残余听力的听损患者。然而，调查显示，国内听损患者的助听器普及率不到 10%，即使是在发达国家，听损患者使用助听器的比例也只有五分之一左右^[3]。影响发展中国家助听器使用率的因素是多种多样的，其中最突出的原因是以下两点：1）助听器价格过高。世界卫生组织建议一个国家的助听器价格应当和该国人民的收入相符合，这样可以保证助听器的价格对于当地普通家庭是可以承担的；2）专业听力人才的缺乏^[4]。传统的助听器验配过程是由听力专家进行一系列的检查、询问以及测试，然后将得到的信息转化为助听器中合适的相关参数，最后按照这些参数对助听器进行配置。这无疑对听力专家是具有挑战性的，需要丰富的临床经验才能完成。而且，患者也会对助听器的验配过程产生影响，由于患者的紧张、焦虑等情绪的干扰，患者的表述有可能和实际情况出现偏差，这会对听力专家的判断产生很大干扰。此外，在使用助听器的过程中一旦出现任何问题，即使问题很小，听损患者首先会想到找听力专家解决问题，这无疑浪费了听力专家大量的宝贵时间。不难想象，发展中国家和发达国家中合格的听力学家的人数都是严重不足的。但发展中国家还要面临听力人才流失的情况，这进一步制约了发展中国家助听器的发展，也间接导致了发展中国家的助听器附加成本的提高。

针对听力专家严重缺乏的问题，人们进行了相关的研究和探索，发现有如下几个解决办法。1）各国加大听力人才培养的力度。这种解决方式看上去是一个不错的解决方法。然而真正实施起来却困难重重。听力专家的培养不是一朝一夕，需要数年甚至数十年的学习、临床，才能符合一名合格的听力专家的标准。因此，该方法可以和其他方法配合进行，不能指

望短期内有所变化。2) 远程听力学^[5]。该技术依托于当前高速发展的信息通讯技术, 通过相关远程技术, 可以进行听力诊断测试、助听器的咨询和调节等。目前, 该方法可以克服一些传统方法的缺点, 如助听器附加成本高、患者与听力专家距离远、偏远地区缺乏专业听力医疗机构等, 为助听器服务增加了一条新的途径, 目前已经获得了一定程度的应用^[6]。然而, 该方法还是存在诸多问题。例如听损患者和听力专家来自不同的地区甚至是不同的国家, 那么从业资格、医疗责任的认定以及报销途径等等都可能出现混乱。这些因素使得远程听力医疗事业的发展变得非常缓慢。随着互联网和智能手机的迅猛发展, 远程听力医疗必然会成为治疗听力损失的一个重要手段。但是, 单靠该技术而言, 还是无法满足现有听损患者的需求。

上述两种方法目前来看都不是最佳解决方案, 从设计理念上来说, 助听器的自适应验配才是解决听力专家缺乏的最有效办法^[4]。助听器的自适应验配是指助听器的装配、验配以及使用管理都由听损患者本人完成, 整个过程不需要听力专家的帮助, 也无需任何其他设备。该方法的好处有以下几点 1) 价格低。助听器价格居高不下一直是制约助听器普及的重要原因, 这点在发展中国家尤为明显。传统助听器验配过程既复杂又漫长, 从前期的听力测试、检查, 中期的参数验配、调试, 再到后期的用户反馈、问卷调查。整个过程花费了大量的人力物力, 无形中增加的助听器的成本。而助听器自适应验配从理念上来讲是不需要听力专家参与的, 这使得助听器的附加服务费大大降低, 有利于助听器的推广使用。2) 环境适应性强。环境的变化会影响到信号的成分, 从而影响了助听器中使用算法的效果。例如, 对于降噪算法而言, 不同的环境所使用的算法参数是不一样的。目前的解决办法是场景分类, 即在助听器上设置不同的旋钮, 每个旋钮代表对应的某个特定场景。自适应验配的好处在于当用户对当前助听器效果不满意时, 可对助听器参数进行调整, 参数调整的依据是对真实声音的反馈, 这更能体现出用户的喜好, 极大地利用了助听器的性能。3) 增强患者的心理所有权意识。心理所有权是一种占有感, 把占有物视为自己的延伸。让用户自己操作验配助听器有利于加强患者的占有感, 激发用户对助听器的研究热情, 从而针对用户自身到达最佳的验配效果。基于上述原因, 本文将重点研究助听器的自适应验配。

1.2 助听器技术和验配的发展研究现状

助听器硬件的发展是助听器验配技术发展的主要推动力, 每当有新一代助听器产品问世的时候, 助听器验配技术也会革新。助听器的产生可以追溯到 19 世纪末 20 世纪初^[7], 但是, 专业的助听器验配大约始于 20 世纪 40 年代。通过最大舒适级 (maximum comfortable level, MCL) 确定助听器增益和频响特征的方法、二分之一增益法等这些如今被大家熟知的验配方法均产生于 20 世纪 40 年代。第二次世界大战后, 听力学成为了一门独立的学科, 至此以后, 助听器验配和助听器硬件技术基本保持着同步发展的态势。

在助听器还处于模拟时代的时候, 其发展周期大概是 10 到 20 年。这这段时期内, 助听器选配、验证方式以及功能都有重大突破和改进。19 世纪五六十年代, 随着对中耳声学特征的研究取得了突破性成果, 声阻抗测试仪器被开发出来并在临床上迅速推广。用于助听器

测试的标准耦合腔也是在该时期被美国国家标准局认定,这对于助听器验配质量的提升尤为关键。20 世纪 60 年代以后,耳蜗机制和心理声学的研究成为了听力学重点研究目标,人们对耳蜗的认识实现了从线性到非线性的跨越。助听器的自动增益控制功能理论基础就是随着对耳蜗认识的不断深入而成功开发。20 世纪 80 和 90 年代是集成技术、电子真空管、半导体等技术高速发展时期,得益于这些技术的研究成果,助听器也变得越来越小,从原来的耳背式助听器发展到了定制机。20 世纪 90 年代以后,数字助听器逐渐取代了模拟助听器,数字助听器的硬件水平已经到达了一个新的高度。非线性验配已经成为了当今助听器验配的主流,其中,DSL_{i/o} 和 NAL-NL 验配方法是非线性验配的典型代表,其市场覆盖率已经超过了 90%。进入 20 世纪以后,各助听器厂家有了配套的专业软件用于验配,这使得验配环节变得更加方便。由于使用方便、功能齐全,这样的配套产品深受广大听力专家、验配师的喜爱。

近几十年来,助听器验配方法已经渐渐打破原来固有方法的局限,朝着多元化的角度发展。部分研究人员利用如今助听器的硬件技术来提高助听器验配的效果。例如在控制声反馈的基础上为听损患者进行验配,这样可以保证助听器的清晰度。也有研究人员利用统计学工具,分析实际语音信号中诸如频率、强度之类的声学特征,然后根据患者需求,对特定频段的语音信号进行放大使得患者对助听器的效果感到满意。还有的研究人员提出了自适应验配的概念,即让患者自己成为验配的主体,根据自身对所听到声音(经过助听效果后)的满意度进行反馈,最后完成助听器参数的验配。1984 年, Köpke, Wiener 等人在专利中提出,用户的听阈值可以通过助听器内部的纯音生成器产生的纯音信号进行测量,由此产生一个合适的传递函数来实现助听器配置。这是助听器自适应验配的最基本概念,遗憾的是直到今天符合这些设计理念的助听器设备为没有实现。但是,对于不需要听力专家参与这点而言,已经有一些初级的产品问世,例如可编程助听器。

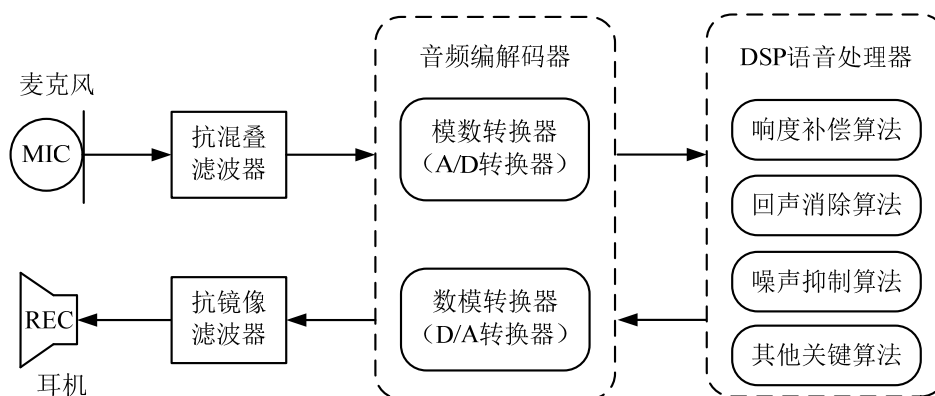
随着助听器验配技术的不断发展和进步,与此相对应的各种评估方法也有了较大的提高和改善,其中比较常见的验配效果评估方法有^[8]: 1) 言语信号测试。言语信号测试是指在声场中评价助听器的验配效果。佩戴助听器最根本的目的是增加患者对言语的理解能力,因此,患者佩戴助听器后对言语信号的辨识提高程度是检验助听器验配效果最直接的方法。可用于言语信号测试的材料有:连续言语、单音节词、无语义声音等。2) 非言语信号测试。非言语信号测试,顾名思义是一种使用非言语信号进行测试的评估方法。通常用于该方法的激励信号有:纯音、窄带噪声、调频音、调幅音等。3) 自我评估测试。自我评估测试是指由佩戴助听器的本人来完成评估测试。助听器的验配效果是由患者本人主观感受的,因此,一些客观的测试方法不一定能表达患者的主观感受。目前比较流行的自我评估测试有十余种,其中,最常用的方法是测定助听器效果缩略简表 (APHAB)^{[9][10]}。APHAB 量化了听力残障和佩戴助听器后听力残障的改变情况。通过该方法,可以比较不同调试方式的调试效果、预测助听器的长期佩戴情况。除了上述评估方法外,还有其他一些检验助听器验配效果的方法,如五音法和七音法等,不再赘述。这些评估方法可以相互结合使用,更加全面有效地完

成对验配效果的评估。

目前，我国出台了一些相关法律规定对助听器市场进行了一定程度的监管^[11]，然而，不同省市对于助听器企业的验收要求存在较大区别，仅有部分省对助听器企业的听力设备、场地面积、售后服务等作出了明确要求。2008年1月2日，劳动和社会保障部出台了《助听器验配师国家职业标准（试行）》，但是，对于助听器具体的验配实施环节并未作出具体的说明。目前国内助听器企业水平参差不齐，相关服务差别很大，一定程度上影响了我国助听器行业的健康发展。就国外助听器相关政策而言，不同国家的政策法规也有很大区别，相比于欧盟和澳大利亚，美国和加拿大的政策要更加严格。相比于国内助听器验配流程，国外的最明显不同点在于更加重视听力学诊断以及对助听器效果的评估。

1.3 数字助听器基本原理

数字助听器和模拟助听器有着很大差别，显而易见的区别就是模拟助听器处理的是连续的模拟信号，而数字助听器处理离散的数字信号。模拟助听器的核心是放大电路，功能是对连续语音信号的实时放大；数字助听器的核心是数字信号处理器（Digital Signal Processor, DSP）。模拟助听器一般包括以下三个主要组成部分：麦克风、放大器以及受话器^[12]。然而模拟助听器经常会受到使用者的一些抱怨，如：小声音听不见、大声音听着难受、能听见但是听不清等。数字助听器的问世带来了听损患者不同的听力感受。数字助听器通常由以下几个主要模块组成：麦克风、抗混叠滤波器、模数转换器、数字信号处理器、数模转换器、抗镜像滤波器、受话器，其结构图如图 1-1 所示。



图错误！未找到引用源。-1 数字助听器系统框图#

从上图可以看出数字助听器处理语音信号的整个过程，外界语音首先经过麦克风转化为电信号，通过抗混叠滤波器滤除发生混叠的频段，然后经过模数转换器转换为数字信号，接着由数字助听器的核心部分 DSP 对该信号进行处理，DSP 会执行一系列算法以达到补偿听损患者听力的目的，随后经过数模转换器重新将数字信号转换为模拟信号，再经过抗镜像滤波器，最后通过受话器将声音传入佩戴者的耳中。麦克风（microphone），学名扬声器，该装置的作用是采集外界语音信号，并将其转换为电信号，优质的麦克风可以将外界语音信号进行线性采集，完好保持原有语音信号的各个特征。早期的麦克风使用的是单个的全向性麦

克风,对各个方向语音信号都进行采集。显然,全向性麦克风在某些场合是不合适的,例如只有两个人聊天的情况。然后产生了方向性麦克风,该类麦克风降低了来自后方和两侧的声音,从而提高了噪音环境下语言的可听度和可懂度。数字助听器的输出设备是受话器,也有人称它为耳机,这是一种利用电磁学原理将电信号转换为语音信号的设备。在受话器的选择方面,由于受到很多主观因素的影响,应尽量选择性价比高的受话器。

数字助听器和模拟助听器都使用麦克风和受话器作为换能装置,但是数字助听器最突出也是最核心的模块是 **DSP**,即数字信号处理器。**DSP** 是一种可以对数字信号进行处理、运算的微处理器,可以方便快速的完成各类数字信号处理算法。此外,若需要调整助听器某方面的性能,只需要编程修改相关算法或者参数,这有利于挖掘助听器的性能。它具有强大的数据处理能力和极高的运算速度,极大地提高了助听器的性能。目前,利用 **DSP** 完成的助听器相关算法主要有:1) 回声消除算法。现在的数字助听器一般都做的比较小巧,这样显得更加美观。但是,这也导致了麦克风和受话器距离很近。当受话器输出处理过后的语音到人耳的时候,不可避免会泄露一部分到麦克风中去,这样就形成了正反馈循环,佩戴者会有听到回音的感觉,严重时甚至会产生啸叫,对耳朵形成了极大的伤害。因此,回声消除算法对于当前的数字助听器而言是必备的。2) 响度补偿算法。响度补偿算法是数字助听器中最关键的技术。不同听损患者的听损类型和听损程度是各不相同的,响度补偿算法需要根据各个佩戴者的实际情况,对原始信号进行适当的放大压缩,使得经过处理的语音信号符合听损患者的剩余听力动态范围,从而使患者能够正常的听到各类声音。3) 噪声抑制算法。由于听损患者或多或少都损失一部分听力能力,他们相对于正常人而言听力辨别能力是比较低的,在嘈杂的环境中影响更加明显。助听器对于噪声和正常的语音信号进行了同样的处理,但是可能输出的语音信号的信噪比会变得更低,助听器佩戴者可能会感觉到声音听得清但是听不懂,因为信噪比的降低干扰了听损患者对言语的判断。因此,噪声抑制算法显得尤为重要,也是目前数字助听器中的必须部分。除了上述三类算法之外,还有一些其他关键算法,如:场景分类、声源定位与跟踪等。这些技术还未普及,没有形成大规模的商业应用。但是这些算法同样在某些方面对助听器性能产生关键的影响,因此也同样重要,这些都是未来助听器算法的研究方向。

1.4 论文的研究内容和结构

目前而言,助听器的验配工作都是由专业的听力专家帮助听损患者完成,数字助听器的自适应验配技术尚处于研究阶段,还没有成型的商业化产品出现。本文旨在研究助听器的自适应验配技术,使得在无需听力专家干预的情况下,最大可能地通过听损患者自己和相关软件的人机交互,最终找到最适合该听损患者的助听器参数。本文各章节主要内有安排如下:

第一章绪论,介绍了助听器自适应验配的背景和意义,阐述了国内外验配技术的发展历史和研究现状。然后分析了数字助听器的工作原理,包括其组成结构以各部分功能作用。最后指出本文的研究方向,对本文的工作内容和文章结构进行了简要概述。

第二章首先介绍了人耳的听觉系统，包括其生理构造以及听觉形成的机理。然后阐述了听力损失的概念，并介绍了国际和国内对于听力损失划分的标准。接着总结了听力测试方法，罗列了几种治疗听力损失的常规方法。最后，对比了传统验配过程和自适应验配过程两种验配方式，阐明了助听器自适应验配的原理。

第三章研究了基于交互式遗传算法的自适应验配技术。本章首先阐述了交互式遗传算法理论基础，然后重点研究了优化算法中各遗传算子的设计，包括染色体编码、选择操作、交叉以及变异操作。接着提出了优化辅助策略以提升算法搜索效率。最后进行了仿真实验以及主观测听实验，结果表明本章提出的优化变异步骤是按照预期的方式进行有效操作，经过本章提出的自适应验配方法进行验配后，可以明显提升语音质量。

第四章研究了基于高斯过程回归的自适应验配技术。本章首先介绍了高斯过程回归用于预测的理论基础，然后进行了相关理论推导，给出了用于预测验配者对当前参数喜好程度的公式；在模型的训练方面，采用遗传算法对超参数进行优化并给出了具体操作步骤，避免了用共轭梯度法优化超参数的缺点；接下来引入了基于期望提高值（EI）的序贯设计，以找到当前“最有意义”的验配参数提供给验配者进行标记。最后，给出了基于高斯过程回归的自适应验配算法流程，并进行了主观测听评估，实验结果证明了该算法的有效性和可行性。

第五章本章从一般自适应验配总体架构出发，展示了本文设计自适应验配系统所采用的人机交互模块，包括纯音测试以及语音测试两个部分，其中采用不同验配算法所对应的语音测试界面稍有不同；阐述了信号处理模块所采用滤波器组的设计过程并进行了相应的仿真与测试，结果表明本文设计的滤波器组完全可以胜任自验配系统中的语音信号处理功能；最后对本文提出的两种验配方法进行了可懂度以及语音质量的对比实验。实验结果表明，本文提出的两种验配方法对语音的识别率以及语音质量都有着显著的提升。

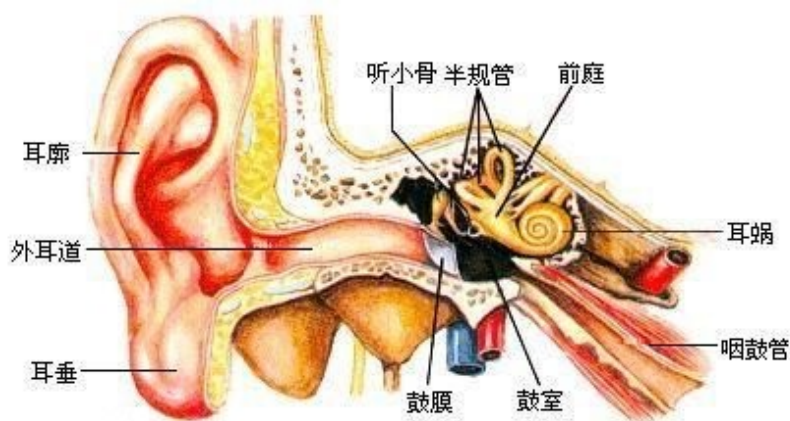
第六章总结全文，指出了数字助听器自适应验配技术的未来研究方向。

第二章 助听器验配理论基础

听觉对每个人的重要性是毋庸置疑的，它是人们进行沟通的最重要途径之一。然而，由于各种各样影响人们听力状况的因素出现，如许多场所充斥着高强度的噪音、各种各样的耳朵疾病等等，我们身边出现了越来越多患有听力损失的人。本章首先介绍了人耳的听觉系统，阐述了如何测试听力损失以及如何对听损患者进行治疗的方法，接着对比了传统助听器验配方法和助听器自适应验配的区别，研究了当前助听器验配发展的新的趋势。

2.1 人耳听觉系统

人耳听觉系统的主要功能是把外界空气震动产生的声波转换为大脑可以感知的神经脉冲信号，从而使人感知到外界信息。在现如今这个信息时代，各类仪器的精密性令我们叹为观止，然而我们人体内部隐藏着比任何精密仪器还要复杂的构造，人耳听觉系统就是其中之一。对于正常健康年轻人的耳朵而言，不仅可以感受到 $2 \times 10^{-5} \text{ Pa}$ 这样微小的声音，也可以承受 $2 \times 10^4 \text{ Pa}$ 这样的超强声音^[13]，这都得益于神奇的人耳生理构造。



图错误！未找到引用源。-1 人耳生理结构图#

人耳的生理构造十分复杂，大致可以分为以下三个部分：外耳、中耳以及内耳，人耳结构图如图 2-1 所示。外耳由耳壳（也称耳廓）和外耳道组成。耳壳就是大家平常所说的耳朵，它的作用主要是收集声音以及匹配耳道与空气之间的阻抗。外耳道是一端连接鼓膜的圆状管道，直径约为 0.5 厘米，长约 2.5 厘米，它是声音传进鼓室的通道。中耳由鼓膜和三块听小骨组成。鼓室内充满空气，通过欧氏管和鼻腔相连。中耳的作用是完成声能到机械能的转换。内耳是听觉系统的主要组成部分，包括耳蜗、神经纤维以及半规管等。其中，半规管的作用是保持人体平衡，与听觉无关。耳蜗由于形似蜗牛外壳而得名，耳蜗的内部充满了淋巴液。耳蜗沿长度方向被基底膜分为了两部分：前庭阶和鼓阶。基底膜上分布了多达 3 万根神经末梢元，这些神经纤维可以感知不同频率的声音，当听骨通过不同频率的声音引起振动，淋巴

液会传递到基底膜，使基底膜上的毛细胞产生共振，从而产生电脉冲传到大脑皮层，经过神经中枢进行进一步分析后，人就可以感知到声音了^[14]。

声音传入人耳通常有两种途径：一种是通过空气进入人耳，再由听觉器官将声音信息传入大脑听觉中枢；另一种是声源通过颅骨传到听觉器官，再传到听觉中枢。人们通常会有这种感觉，听自己讲话和听自己的录音感觉不一样，而听别人讲话和别人的录音是一致的，这是因为听别人的声音有两种途径，而听自己的声音只有一种途径，即颅骨传导声音。还有一个有趣的现象是：在一个嘈杂的环境中，人们可以将自己的听力集中到一个人的声音上面。这是因为人耳听觉系统可以分析出声音进入两耳的时间差从而分辨出声源的方向，从而选取想要获得的声音信号，这就是著名的“鸡尾酒会效应”。

2.2 听力损失测试及治疗方法

2.2.1 听力损失

听力损失（Hearing Loss），又被称为听力级（Hearing Level）或者聋度（Deafness），一般指人耳的听阈在某一频率比正常听阈高出的分贝数。上世纪 80 年代，世界卫生组织（WHO）公布了听力损失程度划分的 5 级标准^[15]，等级划分的依据是 500Hz, 1KHz, 2KHz 以及 4KHz 的气导平均阈值，具体划分标准见表 2-1。

表 2-1 听力损失程度分级标准

分级	听力损失程度	听阈均值（dB HL）
I	轻度（mild）	26~40
II	中度（moderate）	41~55
III	中重度（moderately severe）	56~70
IV	重度（severe）	71~90
V	极重度（profound）	91 以上

听阈均值在 25dB 以下的为正常听力，对一般的语言都可以分辨清楚；患有轻度听力损失的人对于大声讲话可以分辨，但是听不清细小的声音，如风吹树叶声、小声讲话等；对于中度听力损失的人群而言，日常生活中的语言交流已经变得困难，经常对他人说话感到模糊、难以理解，开始需要借助助听器的帮忙才能正常交流；如若一个人的听力损失到达了中重度的级别，那么即使是别人大声讲话也难以听清，佩戴助听器后可以很大程度上改善其听力状况；患有重度听力损失的人对于一般的声音已经感受不到，只有对类似于喊声、鼓声这类高强度的声音才有所反应，佩戴助听器后也可以有很大帮助；对于极重度听力损失的人而言，基本上对任何声音都很难察觉到，只有在助听器的帮助下才能对外界声音有所感应。

我国第二次残疾人抽样调查所采用的听力损失划分标准与国际标准接轨，从环境与支

持、活动与参与、结构以及功能等方面，根据听力损失程度的不同，将听力残疾进行了四级划分。

(1) 听力残疾一级：双耳全聋，较好耳的平均听力损失高于 90dB，听力系统的结构和功能有极重度的损伤。在没有听力辅助设备的帮助下，几乎听不到任何声音，在语言理解、交流方面存在严重障碍。

(2) 听力残疾二级：一耳全聋，较好耳的平均听力损失在 80dB 以上，听力系统的结构和功能有重度的损伤。在没有听力辅助设备的帮助下只能听见强度较大的声音，如鼓声、鞭炮声等，在社会活动交际方面存在严重障碍。

(3) 听力残疾三级：较好耳的平均听力损失在 61 到 80dB 之间，听力系统的结构和功能存在中重度的损伤。在没有听力辅助设备的帮助下，只能听懂一部分词语以及简单的句子，在社交方面有中等程度的障碍。

(4) 听力残疾四级：较好耳的听力损失在 41 到 60dB 之间，听力系统的结构和功能存在中等程度的损伤。在没有听力辅助设备的帮助下，对于言语声可以听见，但是辨别不清，在日常活动交流上存在轻度障碍。

按照听力损失的可恢复性可将其分为暂时性听力损失和永久性听力损失。人耳处于强噪声环境下一定时间后，会产生耳鸣甚至听力下降，如果听阈升高 10dB 且脱离噪声环境后数分钟后听力恢复，这种现象被称为听觉适应^[16]；如果人耳处于噪声环境中的时间较长，听阈升高 15 到 30dB，脱离噪声环境后需要几小时甚至数天才能恢复正常听力，这种现象被称作听觉疲劳，也称噪声所致暂时性阈移，这就是我们通常所说的暂时性听力损失。人耳长期处于噪声环境下可能会对听力系统产生不可逆转的损伤，这种情况被称为永久性听力阈移。永久性听力阈移一般由暂时性听力阈移慢慢发展而来，在听力图上的 3KHz 到 6KHz 处呈 V 字形或者 U 字形听力曲线，一般被称为 4KHz 听谷。

按照遗传因素、工作场所的噪声、疾病、临床表现特征等综合因素，则可以将听力损失划分为以下 5 种：

(1) 传导性听力损失：上一小节提到过，人听到声音的一条途径是声音由外耳经过中耳最后再传到内耳，如果这条途径中的任一器官（如听骨，鼓膜等）发生问题都会产生听力损失，这种情况引起的听力损失被称为传导性听力损失。传导性听力损失患者可通过药物或手术方法进行治疗，一般治疗以后患者听力水平会有明显提高，甚至会恢复到正常听力水平。

(2) 感音神经性听力损失：该类型听力损失通常伴随着内耳退化的情况，内耳退化可能是正常衰老的结果，也可能是由药物中毒引起的耳蜗、听觉神经乃至听觉中枢神经系统的退化。

(3) 混合性听力损失：如果既存在传导性听力损失，又存在感音神经性听力损失，则把这种听力损失称为混合性听力损失。该类型听损患者的治疗方式是药物治疗配合助听器辅助的方法。

(4) 中枢性听力损失：该类型听力损失通常表现为耳蜗核平面以上的损伤，具体表现

为脑干与大脑、蜗神经核及其中枢传导通路的某一处或多处病变，听力损失以低频为主。

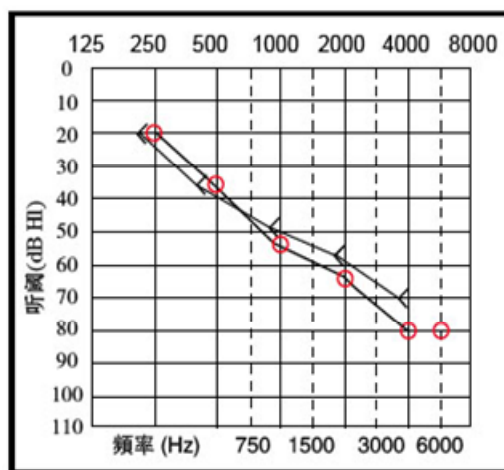
(5) 功能性听力损失：该类型听力损失一般由非器官性病变引起，通常是由某些心理因素造成的，不应使用常规治疗方法。

2.2.2 听力测试方法

听力测试的主要目的是对听损患者的听力状况进行定性或定量的诊断，只有在全面了解了当前听损患者的实际听力情况的基础之上，才能对下一步治疗方案做出正确判断。目前，临床听力测试方法有很多，大致分为两类：一类是利用各种声音刺激后，通过观察听损患者主观的反应，称为主观测听法；另一类通过客观方法测试听损患者的听力情况，称为客观测听法。主观测听法主要有以下几种：音叉检查、纯音测听、言语测听以及声场测听等；而客观测听法包括：听性脑干反应（ABR）、多频稳态诱发电位测试、声导抗测听以及耳声发射等。主观测听方法投资较小，可以较为准确的反映出听功能情况，但是操作费时；客观测听方法的特点是简单、便捷，可分别测得两耳的听力情况，但是投资较大，对操作人员的要求较高。鉴于以后章节会用到到纯音测听和言语测听，本小节将对这两种方法作进一步阐述。上述其他听力测试方法由于涉及医学临床知识且不是本文研究重点，将不再详述。

1. 纯音测试

纯音测试，又叫纯音听阈测试（pure tone audiometry, PTA），包括气导和骨导听阈测试。纯音测试一般过程为：对测试者采用特定频率的纯音信号，受测试者有一半次数以上能引起听觉反应的最小声压级即为受测试者在该频点的听阈值。目前，纯音测试方法不仅是临床上使用最多的听力测试方法，同时也是配合助听器使用最多的听力测试方法。纯音测试的结果最后会以听力图的形式体现^[17]，图 2-2 即为一个听损患者的听力图。



图错误！未找到引用源。-2 听损患者纯音听阈测试结果图

从上图可以看到，听力图的横坐标表示所听测试信号的频率（单位 Hz），横坐标采用倍频程标记；听力图的纵坐标表示听阈（单位为听力级，dB HL），即听损患者能感受到的

最小声压级。图**错误！未找到引用源。**-2 是一个典型的感音神经性听损患者的听力图，从该图中可以看出，该听损患者随着频率的升高，其听阈不断增大。

2. 言语测听

语言是人们用来相互交流，表达思想的工具，是人类文明得以发展的重要基础。言语是频率、时间、强度连续不断变化的声音，是口头说出的语言。由于听觉功能的一点点损伤都能对言语的理解产生影响，因此言语可以用来作为判断听觉功能的指标。言语测听就是这样的测听方法，具体指利用言语作为样本测试信号，用听力计来检查受试者听力状况的测听方法。言语测试使用不同的材料作为测试信号，如单音节词，双音节词语，短句甚至音乐。言语测试的常用指标如下：

(1) 言语察觉阈 (speech detection/awareness threshold, SDT 或 SAT)：对受试者采用指定言语信号测试，能察觉（不一定要听清）到 50% 所听词汇的最低声压级即为言语察觉阈。

(2) 言语接受阈 (speech reception threshold, SRT)：也称言语识别阈 (speech recognition threshold, SRT)，指的是理解言语所需要的最低声压级。该方法指标的测量过程为：以双音节词汇（一般采用扬扬格词，即具有相同重音的双音节词）作为听力测试材料，测试的声强逐渐增强，直到能够准确听清并复述所听词汇的 50% 以上，该声压级即为言语接受阈。

(3) 不适阈 (uncomfortable level, UCL)：当声强超过某一特定值的时候，受试者会感觉到不舒服甚至感到疼痛。听力正常人的不适阈大概在 130dB SPL（或者说 110dB HL）左右。

(4) 言语识别率 (speech recognition score, SRS)：对于特定的声压级，能够正确听懂并且复述测听所用词汇的百分比。

(5) 最大辨别率值 (maximum discrimination score, MDS)：对于同一个测听信号而言，一般加大声强后，受试者感受到的清晰度也会随之提高，然而，当信号增加到某一声压级后，清晰度不再提高，该声压级即为最大辨别率值。

(6) 最适响度级 (most comfortable level, MCL)：受试者能够听懂测听词汇且感觉舒适的言语声级。一般正常人的最适响度级为 65dB SPL。

(7) 清晰度区域 (articulation area, AA)：又称听阈距测定，指的是言语接受阈和言语不适阈之间的差距。听力正常人的听阈距为 50 到 80dB，感音神经性听损患者的清晰度区域小于 50dB，耳蜗性听损患者的清晰度区域小于 35dB。

(8) 噪声言语识别率 (speech recognition score in noise, SRSIN)：将测试言语识别率时的测听信号换成含噪信号后，重复 (3) 中方法所得到的结果即为噪声言语识别率。对于听力正常的人而言，语音中混合一定程度噪声后对言语的理解是没有影响的，但是对于听力损失的人来说，一点点噪声都可能影响到言语内容的理解，该指标对于听损患者的后期治疗有较大帮助。

2.2.3 听力损失治疗方法

听觉是人们的重要感官之一，如果一个人的听力损失已经到了严重干扰日常生活的程度，那就有必要寻求解决办法了。听力损失不像视觉衰退，现在治疗视觉衰退的方法比较成熟，针对不同的情况，有激光治疗、佩戴眼镜、移植眼角膜等方法。通常，通过相关治疗之后，视觉失常的患者是可以恢复或者矫正到正常水平的。然而，对于听损患者的治疗方法，虽然手段繁多，但是也只是尽可能恢复或矫正到正常水平。一般治疗听损患者的方法有：药物治疗、佩戴助听器、移植电子耳蜗、手术治疗、饮食治疗等。

1. 药物治疗

药物治疗包括中药治疗和西药治疗两种方法。中药治疗听力损失的思路一般是活血祛瘀、通络开窍，通过服用中药使得耳部经络阻塞症状缓解，从而达到恢复听力的目的。有时中医治疗还会采用针灸疗法予以辅助治疗。西药治疗一般采用静脉给药的方法，结合肌肉注射，将药物渗透进入人体。虽然目前关于中西药治疗听力损失的报道并不少见，但是真正具有权威机构认证并经过大范围临床应用检验的药物治疗方法却没有出现，而且该方法还可能伴有副作用^{[18][19]}。

2. 佩戴助听器

目前而言，佩戴助听器是一种治疗听力损失比较快捷并且相对有效的方法^[20]。通过对听力损失患者选配合适的助听器并将助听器参数调到合适的值，听损患者的听力程度一般会有很大程度上的改变。听损患者佩戴助听器后的感受通常都是仿佛耳朵获得新生一般，与他人的交流变得没有障碍，生活质量得到了极大地提升。但是，助听器治疗是一种治标不治本的手段。而且一般而言，很难像配戴眼镜一样，佩戴助听器后听力状况会有所改善，但没有达到跟正常人一样听觉感受的程度。

3. 移植电子耳蜗

电子耳蜗是一种人工耳蜗技术，该方法对极重度听力损失患者具有相对明显的治疗效果^{[21][22]}。对于双耳重度或者极重度感音神经性听损患者而言，即使佩戴助听器之后，他们也无法对声音做到有效的感知，人工耳蜗移植技术是当前国际公认的针对该类人群唯一有效的治疗方法。然而，该方法缺点也很明显，目前电子耳蜗的市场售价高达数十万元，这是一般家庭所承受不起的。此外，电子耳蜗移植也具有一般手术的风险性。

4. 手术治疗

手术治疗是在长期耳聋临床研究基础上发展起来的治疗方法，目前已经形成比较成熟的体系。它包括：水针疗法、光波治疗、脉冲电磁场治疗等。该治疗方式虽然效果明显，但是具有风险性，且治疗费用高昂，在治疗的过程中患者需要承受一定的痛苦。

5. 饮食治疗

饮食治疗是治疗听力损失的一种辅助方式。该方法是要控制脂肪的摄入量，补充蛋白质和维生素，多吃含锌食品。该方法虽然对人体没有危害，也无需额外治疗费用，但是效果不甚明显，最好是配合其他治疗方法进行同步治疗。

2.3 传统助听器验配流程

助听器与一般的商品不同，它是帮助听损患者改善听力状况的专业设备，助听器验配必须由具有从业资格的专业人员操作完成，以确保患者佩戴助听器后不会进一步损伤残余听力。此外，验配过程必须在专业设备的辅助下以及特定的声学环境中完成。从事助听器验配工作的资质要求应符合《助听器验配人员执业资格准入管理办法》，设备和场所要求均参照《助听器验配机构基本设置推荐规范》。助听器的验配过程如图 2-3 所示，可以分为以下六个阶段：前期准备、助听器预选、定制耳膜、验配及适应性训练、效果评估以及随访跟踪评估^[23]。

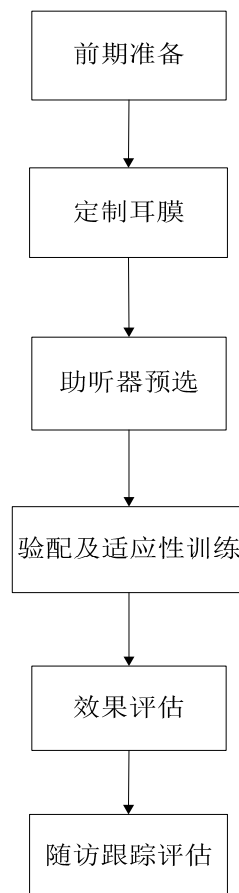


图 2-3 传统助听器验配流程图

1. 前期准备工作

(1) 咨询病史

咨询病史看似无关紧要，实际上对于整个助听器验配有基础性作用，只有做了详尽的病史采集，才能对患者随访、统计分析等后续环节做到心中有数。病史的咨询所要了解的信息较多，通常需要了解的问题有：听力损失产生的时间、在生活中对各种声音的反应以及家族史等。一些发达国家在这方面的档案信息会非常详细，甚至会记录患者家庭成员的性格、

态度等。

（2）常规耳科检查

在该环节应重点检查鼻咽、中耳腔以及咽鼓管这几个部位有没有发生病变，有些听损患者就是由于其中某处发生病变导致听力受损。此外，还要仔细检查外耳道和鼓膜，临床上有部分听损患者就是这方面出现了问题。常规耳科检查需要从业者亲自完成，发现助听器佩戴禁忌或需要耳科专家进一步确证时，应通知患者到耳科医生处做进一步检查。

（3）听力检测

听力测试的方法在 2.2.2 小节已经作出说明，在验配环节中需特别注意的是，要根据听损患者的年龄选择相应的测听方法。6 岁以下儿童可以使用的测听方法有：行为测听法、视觉强化测听法、游戏测听法，6 岁以上即可使用纯音测听方法。大龄儿童以及成年听损患者需用气导法测试听阈，为了方便后续的助听器选择，还应配合检查骨导听阈以及不适阈，此外还可以进行声导抗和声场测试。在条件允许的情况下，可以用听力计、助听器试听仪或者标准助听器等设备，让患者进行试听，以获得听损患者所需增益、斜率以及最大输出等相关参数。

2.定制耳膜

该步骤是可选步骤，对于选择标准助听器的用户而言可以跳过该步骤。但是，若听损患者选择耳背式或者盒式助听器，则必须制作与使用者相对应的专用耳膜。

耳膜与助听器相连，可以很方便的戴在听损患者的耳廓后面。耳膜不仅有舒适、美观的特点，对于助听器的使用效果也有很大影响。耳膜不仅可以固定助听器，还可以将经过助听器处理后的声音有效导入进入外耳道。此外，它还可以封闭外耳道，起到防止反馈啸叫的作用，一定程度上改善了助听器的声学效果。

3.助听器预选

在经过了助听器验配前期准备工作并确定验配耳之后，需要进行助听器预选。该环节对助听器验配从业人员的要求较高，不仅要求从业者有专业的听力学知识，还需要对各类助听器的性能有很好的掌握。

根据听力检查的结果选配合适的助听器，选配助听器的原则是：助听器的输出和频响曲线符合之前的测听结果，所选助听器的频响曲线要尽量吻合听损患者的实际听力状况。对于具备测试仪器的机构而言，可以使用耦合增益法进行预选。在没有相应设备的情况下，也可根据听损患者的听力损失程度进行助听器选配，该方法的原则是使得听力损失与助听器的最大输出声相匹配。根据听力损失选配助听器的标准见表 2-2。

表错误！未找到引用源。-2 助听器选配标准

听力残疾等级	听力损失程度	助听器最大输出声(dB SPL)
I	轻度(mild)	<105
II	中度(moderate)	115-124
III	重度(severe)	125-135
IV	极重度(profound)	>135

4.助听器验配及适应性训练

(1) 助听器验配及初步评估

目前常用于助听器验配的方法是真耳介入增益以及助听听阈测试法。将音量调节到某一适当的位置后,通过调节音调、增益等助听器控制旋钮,使得助听器效果达到患者满意的程度。若无论怎么调节,始终得不到满意效果,可以考虑更换另一种品牌或者型号的助听器。

(2) 适应性训练

对于长期患有听力损失的人而言,刚佩戴助听器的时候可能会有一段时间感到不适应。这是正常现象,人类的感知系统是敏感的,听力状态的突然改变会使人对周围环境产生不适感,通常经过一段时间后就会完全适应。刚开始佩戴助听器时,应尽量将音量调低,使得听损患者对周围环境的适应是过渡性的而非突发性。佩戴时间也应从刚开始的每天两到三个小时逐步增加到全天都佩戴助听器。适应期一般会有两到三周,儿童及老年可能会花费更长的时间来适应。

5.效果评估

助听器效果评估可以帮助验配人员了解听损患者佩戴助听器后对其听力状况的改善情况,从而判断是否可以通过进一步调整助听器参数来优化助听器效果。此外,助听器效果评估还有利于听损患者的下一步听力康复计划。具体效果评估方法在 2.2.2 小节已作介绍,有条件的情况下最好进行声场助听效果和真耳测试。

6.随访跟踪评估

助听器验配不是一次验配就结束的,使用数字助听器的听损患者在短期内(一般为数周)常常需要重新编程及调整。一般情况下,听力患者在初次佩戴助听器之后第一年每个季度要进行一次复查,之后每半年复查一次。复查一般要解决以下几个问题:助听器音质和音量如何、助听器参数是否需要调整、有没有听力方面困惑需要咨询等。

2.4 助听器自适应选配原理

在助听器验配领域中,分为传统选配方法(自下而上)以及新型选配方法(自上而下)。对于传统选配方法而言,选配的过程是由验配师根据对患者测得的声学特征(如患者的听阈、痛阈等)之后为患者选择合适的助听器以及相应的参数,接着通过与患者的交流来调整参数,

然后根据最终参数结合相应的验配处方公式，以求为其找到“最合适”的助听器^[24]。而新型的自适应验配方法是患者通过与相关软件交互操作，直观反映对助听器佩戴效果的感受，根据相关算法求得最适合患者的助听器参数。显然传统验配是依赖验配师主观经验的方法，而自适应验配方法甚至不需要验配师的参与，通过患者的直观感受就可以进行助听器的验配。由于自适应验配是由患者自己进行，该方法极大方便了听力患者的验配。而且，自适应验配总能为听力患者找到“最适合”自己的助听器。

传统验配方法示意图如图 2-4（a）所示，需要听力学专家或者专业人员的帮助。处方公式是听力学专家对患者进行助听器验配时常用的方法，根据患者的听力图以及处方公式，可以计算得到患者所需要的压缩比、增益值等助听器的最主要参数，从而得到患者所需的输入输出曲线（I/O 曲线），听力学专家尽量将助听器的频响设置在理想的状态。

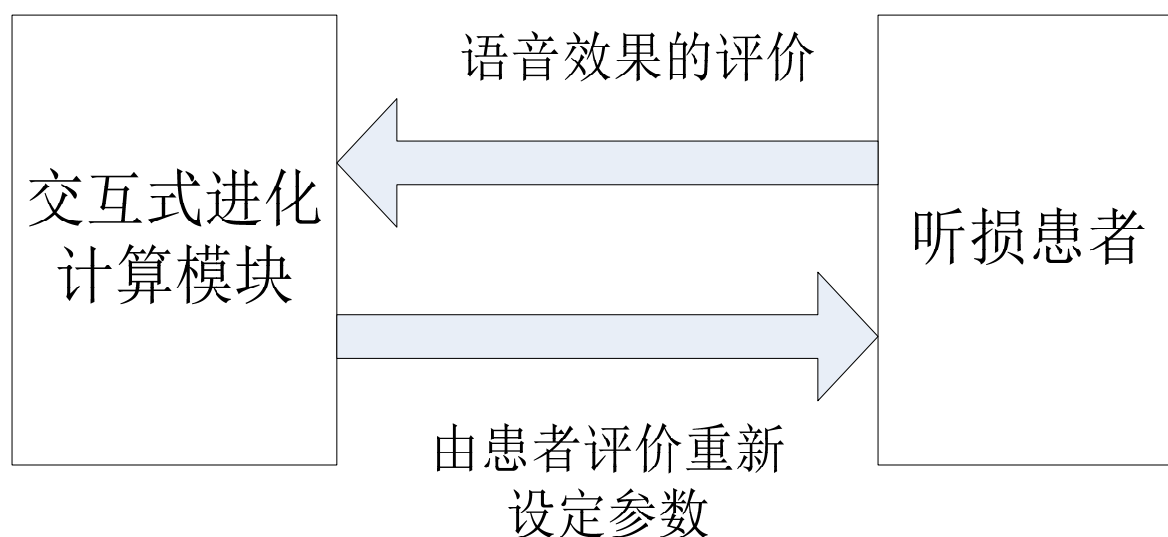


图 2-4 传统验配与自适应验配对比图

虽然传统验配方法是目前助听器验配的主流，但该方法具有很大的局限性，主要体现在下面几个方面。第一：需要听力学专家的参与；第二：需要测量患者耳朵的声学特性；第三：初次验配完成之后还需多次微调。由传统验配自身的特点可知，上述局限性均是不可避免的，因此，自适应验配是目前一个热门的研究方向，可以实现更加人性化的验配。自适应验配方法如图 2-4（b）所示。从图中可以看出，自适应验配交互软件代替了听力学专家的作用，此外，也无需选择相关验配公式。患者通过不断对交互接口传来的语音信号进行声音效果好坏的判断，自适应验配算法收到患者的反馈后对相关参数进行修正，再形成语音传送给患者。通过不断循环上述过程，最终得到合适的助听器参数以供患者使用。

助听器自适应验配软件需至少包含以下几个模块：优化模块、语音信号处理模块以及人机交互接口。优化模块的作用是搜索“最佳”的补偿参数，所谓“最佳”的标准由患者主观确定，即患者认为语音效果好则参数为优。语音信号处理模块根据优化模块传递过来的参数，对语音信号进行处理，最终得到处理后的语音交付给人机交互接口。人机交互接口顾名思义，

即直接跟患者交互的界面，患者从该接口得到处理后的语音，就可将自己对语音的评价反馈给人机交互接口。

2.5 本章小结

本章第一小节介绍了人耳的听觉系统。第二小节首先阐述了听力损失的概念，并介绍了国际和国内对于听力损失划分的标准，然后总结了听力测试方法，并对纯音测试和言语测听两种常用测听方法作了进一步阐述，最后罗列了几种治疗听力损失的常规方法。第三小节归纳了传统助听器验配的一般流程并对该过程的具体步骤进行了解释说明。第四节分析对比了传统验配过程和自适应验配过程两种验配方式，阐明了自适应验配的原理。

第三章 基于交互式遗传算法的自适应验配方法

3.1 交互式遗传算法原理及自适应验配一般流程

3.1.1 交互式遗传算法原理

自然界是人类赖以生存的物质世界，也是人类思想起源的地方，人们从自然界中探索世界、找寻规律。自然界的一些自适应优化现象给予了人类启发，人们从中总结出了一类优化方法，称为智能优化算法。所谓智能优化算法，即一类按照某种规则或思想进行的搜索过程，用以得到满足用户所求问题的解^[25]。智能优化算法对各类复杂的问题有很强的适应性，应用这些算法可以解决一些其他算法难以解决的优化问题。常见的智能优化算法有：遗传算法，蚁群算法，模拟退火算法等。遗传算法是一种随机进行迭代、进化的搜索算法，它将生物界中的自然选择和种群遗传学原理引入到搜索过程中，使其具有广泛的适用性^{[26][27]}。遗传算法将自然界中的适应环境者生存这一自然规律作为主旋律，并且结合了随机信息交换思想，适者生存这一思想消除了迭代过程中的不适应因素，随机信息交换这一思想利用了原本解自身的知识，从而提高了搜索速度^{[28][29]}。

传统遗传算法一般包含以下几个操作步骤。首先，对问题的解进行编码，即用染色体来表示问题的潜在解，并形成编码后的初始种群；适应度函数需要借鉴优化问题的目标函数来确定；然后，根据适应度函数值的大小来挑选个体参与遗传操作；最后，按照优胜劣汰和适者生存的生物学原理进行衍化，直到得到问题的最优解或者近似最优解^{[30][31]}。每个个体在种群的衍化过程中都要进行优劣评价并计算其适应度函数值，个体在选择、交叉及变异算子的作用下不断向更高的适应度进化，从而达到寻找优化问题最优解的目标。

然而，传统遗传算法还是有一些不足之处，最明显的一点就是必须有一个事先指定的适应度函数作为评价标准^[32]。但是对于很多问题而言，无法用一个明确的显式函数来表示适应度。例如音乐的创作、服装设计、基于用户满意度检索等研究，结果的好坏评判是以人为主观的判断为基础，这一类问题都是传统遗传算法无法解决的。如果将传统遗传算法的进化过程与中间过程结果的人为主观评价结合起来，就可以解决之前无法用显式表示价值函数的优化问题^{[33][34]}。1986年，Dawkins首先提了交互式遗传算法的思想，并将其应用于生物图形形成^[35]。2001年，Takagi对交互式遗传算法进行了综述，对该优化方法的理论基础以及应用研究提出了许多启发性的见解，这对于后来交互式进化计算领域的研究具有重大意义，该文章也被多次引用。目前，交互式遗传算法的研究越来越深刻，其应用也越来越广泛，在工业设计、图像识别、艺术创作、电路板布局优化等各个领域都有良好的应用^{[36][37][38]}。

本章提出的数字助听器的自适应验配方法采用的智能优化方法就是交互式遗传算法。由于自适应验配的完成是由听损患者本人完成，也就是说验配过程中声音好坏的评价是由听损患者主观判断而没有显式的费用函数（中间结果的评价没有听力专家的干预）。这个研究课题的背景与交互式遗传算法的特点是完全符合的。在2007年，Hideyuki Takagi, Miho Ohsaki

就提出了一种基于交互式进化计算的自适应验配方法^[39]。该方法设计了一种二维高斯分布的寻优模型，以求找到合适的参数来建立一个声压-频率-增益三维模型，一旦最优模型确定下来，对于每帧语音信号，都能根据其频率和声压特性找到对应的增益值来补偿听损患者的听力。该三维模型如图 3-1 所示：

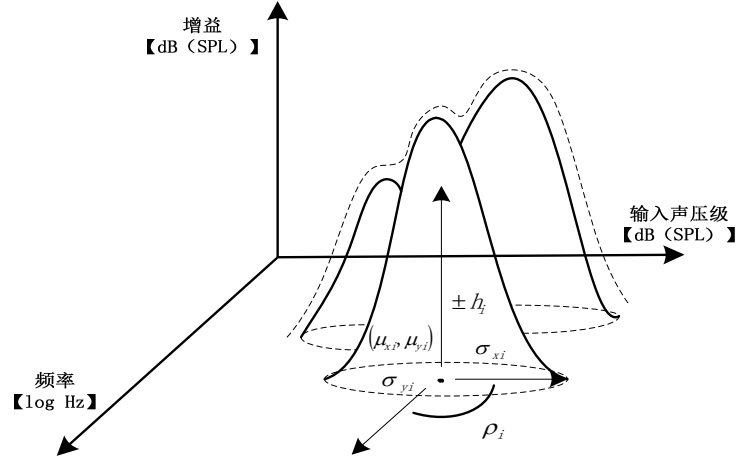


图 3-1 高斯验配模型示意图

X 轴为语音信号的频率 (Hz)，Y 轴为语音信号的声压级 (dB SPL)，Z 轴为听力补偿 (dB SPL)。对于某一帧语音信号的特定频段子带信号，计算其声压级后，通过响度曲面三维模型即可确定所需补偿的声压级。响度曲面采用二维高斯模型，其公式如下：

$$AL(x, y) = \sum_{i=1}^5 Gauss_i(\mu_{xi}, \sigma_{xi}, \mu_{yi}, \sigma_{yi}, \rho_i, h_i, \pm) \quad (3-1)$$

其中， x 为子带语音信号的中心频率 (Hz)， y 为该子带语音信号的声压级 (dB SPL)， $AL(x, y)$ 为该子带语音信号的听力补偿 (dB SPL)。然而，由于该方法要求种群较大（种群数设为 20），并且是在全局三维空间搜索，其收敛时间必然较大，在实际应用过程中很难满足患者的需求。本章提出的自适应验配方法与 Hideyuki Takagi 和 Miho Ohsaki 方法的区别在于：Hideyuki Takagi 和 Miho Ohsaki 是通过求解一个最优三维模型，然后采用响度补偿的算法直接依据该模型数值对听损患者进行听力补偿；本章提出的自适应算法则是希望通过智能优化算法找到最适合听损患者的相关助听器控制参数，如音量、频响、最大输出、自动增益调节等等。

3.1.2 自适应验配流程

一般而言，自适应验配系统需要包含至少三个部分组成，分别为寻优模块、信号处理模块以及人机交互模块。系统首先初始化一组用于遗传算法的参数，该组参数将影响最终的输出语音。原始语音信号以及优化模块的参数将会输入信号处理模块，经过语音信号处理模块的处理后得到一组新的语音信号。人机交互模块从信号处理模块获得经过处理后的语音信号并反馈给听损患者（即验配主体），验配者根据对每段语音信号的满意程度作出判断，给予

相应的满意度评级，人机交互模块把验配者作出的反馈传递给寻优模块以进行下一轮验配，直至验配者获得满意的听力体验或者优化算法的迭代次数达到设定的上限值。

由于本章重点研究交互式遗传算法在自适应验配中的应用，数字信号处理模块和人机交互接口两个模块将在第 5 章的软件设计部分再作介绍。优化模块是助听器自验配的核心部分，该部分使得验配者可以在脱离听力专家的情况下自行验配。优化模块的设计目的是为了获得最佳的参数，信号处理模块的语音合成过程以寻优模块获得的参数为依据。本章提出的自适应验配方法采用的是交互式遗传算法，算法流程图如图 3-2 所示：

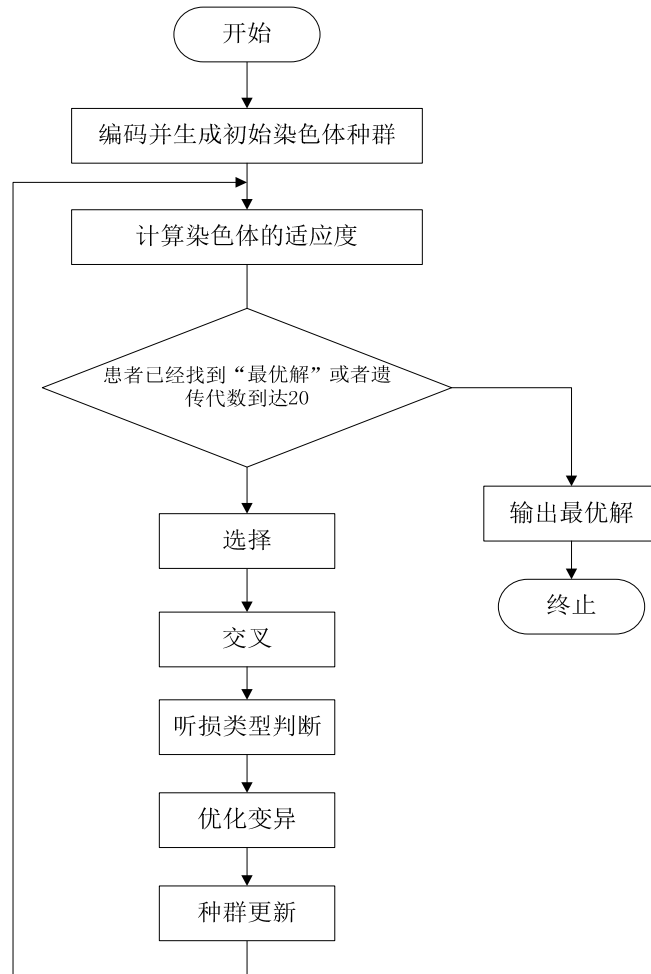


图 3-2 优化算法流程图

本章算法将 0 到 8000Hz 划分为 11 个频带，每一个染色体包含相应 11 个频段的增益值。算法包含基本的选择、交叉、变异步骤。但是自适应验配有其相应的特点，对算法的要求更高，如果按照传统遗传算法操作，用户的疲劳度过大可能会导致自适应验配失败。因此，本文对变异步骤作了相应的优化。变异步骤的作用是局部搜索并增加种群的多样性，传统的变异步骤一般是对染色体的某一或某几个基因进行随机变异，但是由于变异操作是完全随机的，变异操作可能会使得种群向着优化的方向发展，也可能向着糟糕的方向变异。本文的优化变异步骤是根据听损的类型进行有针对性的变异，使得变异步骤总体向着优化的方向发展，从而达到加速收敛的目的。

3.2 遗传算子的设计

考虑到交互性，种群的多样性以及进化特性，本章采取的优化算法是交互式遗传算法。在上一小节已经对交互式遗传算法（IGA）进行了简单的介绍。然而，自适应验配是 IGA 算法较为特殊的应用，必须要考虑到用户的主观感受，并且要尽量降低用户的疲劳度。此外，由于语音信号的一些固有特性，不能像传统遗传方法那样处理语音信号。因此，本节将主要阐述针对自适应验配的遗传算子的设计。

3.2.1 染色体编码及选择操作

编码就是用一种码来表示优化问题的解，从而使遗传算法的码空间和问题的状态空间对应起来，这在很大程度上要依赖于优化问题的性质，并将极大地影响遗传操作的设计。遗传算法的一个显著特点就是它在编码空间和解码空间中交替的工作，它在编码空间对染色体进行遗传运算，而它在解空间则对解进行适应度评估和选择。由此可以看出，遗传算法的优化过程不是直接问题参数本身产生作用，而是在一定编码机制所对应的码空间上进行的，因此编码的选择会对算法的性能和效率产生重要影响。常见的编码技术有：二进制编码、十进制编码、实数编码、格雷编码。

本章将语音信号从 0~8KHz 划分为 11 的频段，也即形成了 11 个通道。对于不同的患者，每个通道的增益是不一样的。因此，本文将每个通道的增益数值（dB）作为一个基因，所有通道的增益值组成一个数组形成染色体。采用的编码方式是二进制编码，由于正常的增益数值范围大概在 0~100dB，因此对于一个通道而言，7 比特足以表示增益的范围。每个染色体包含的比特数为 77 比特（7 比特*11）。

对于如何从父代选择合适的染色体进入子代，遗传算法的原则是“适者生存”，即适应度高的染色体遗传到下一代的可能性也就高。然而，对于本文的研究（助听器的自适应验配）而言，没有一个明确的目标函数或者费用函数，因此，适应度的选取不能像传统遗传方法一样，根据染色体和目标函数计算得出。考虑到人对语音的分辨能力以及患者的评价疲劳度^[40]，本章采取的方式是将适应度分为 5 个等级，分别对应 5 个不同的具体数值，如 1,4,9,16,25。这五个等级分别对应 5 个不同的评价等级，即“劣”，“差”，“中”，“良”，“优”。对于每组特定参数形成的语音，患者听取该段语音后，根据自己的主观感受进行 5 个等级的评价，若患者做出的评价为“良”，则该组语音信号所对应的适应度值即为“良”所对应的数值，该数值将作为该组参数遗传到下一代的概率依据。

具体将五个等级对应到哪五个不同的数值是一个比较复杂的问题，该数值的选定将直接影响算法性能的好坏。本文按照超几何算子的方法选取适应度值^[41]，即对于评价等级 N（本文选择为 5），第 n 个等级对应的遗传概率为：

$$P_n = \frac{1}{\gamma * n^q} \quad (3-2)$$

其中, q 为常量, 本章取值 0.5。 γ 的计算方式如公式 (3-3) :

$$\gamma = \sum_{k=1}^N \frac{1}{k^q} \quad (3-3)$$

用超几何算子形成不同等级的遗传概率, 代替了原来的适应度值, 使得遗传算法的进行变得更加平稳, 不会出现前几代竞争过于激烈, 而算法进行到中后期时竞争不够的情况。此外, 该方法还可以一定程度上避免“早熟”现象。具体进行“选择”操作的时候, 采用的是“轮盘赌”算法, 轮盘赌示意图如图 3-3:

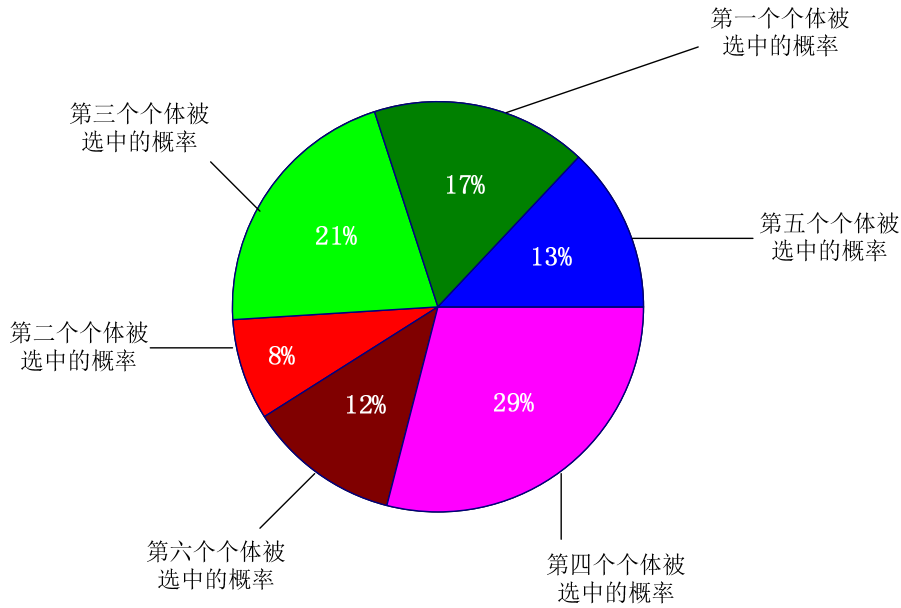


图 3-3 盘赌示意图

轮盘赌是一种常用的随机选择算法, 类似于博彩游戏中的轮盘赌, 首先将个体适应度转化为选择的概率, 按个体所占比值在一圆盘上进行划分, 每次转动圆盘后待圆盘停止时指针所指向的位置即为被选中的个体。显然, 个体选择概率越大, 其在圆盘上所占的面积也就越大, 被选中的机会也就越大。

具体实现方法如下: 首先对染色体进行排序, 排序依据是根据染色体的适应度值由低到高, 依次计算从序号 1 到 N 的前部累计概率, 如图 3-4 所示。然后利用随机函数产生一个 $[0,1]$ 之间的随机数, 判断该随机数落在哪一个概率区间内, 若落在第二个区间, 则选取第二个染色体遗传到下一代。例如, 随机产生的数字为 0.48, 显然该数字落在序号为 3 的区间内, 则应选择序号为 3 的染色体遗传到下一代。

染色体序号	0	1	2	3	...	n	
概率	0.08	0.1	0.14	0.21	...		
累计概率	0	0.08	0.18	0.32	0.54	...	1.0

图 3-4 轮盘赌实现示意图

为了保证算法的收敛性，本章采用精英保留策略^[42]。将上一代的最优染色体直接输出到下一代最终形成的染色体种群，期间不经过交叉、变异等步骤。这样做的意义在于保护了最优解不被下一轮的遗传操作破坏，以免阻碍了获得最优参数的进程，造成许多时间、操作上的重复与浪费。

3.2.2 交叉操作

遗传算法中的交叉操作是指两个将相互配对的个体按照某种方式相互交换各自的部分基因，从而形成两个新的个体。实际上，交叉就是两个染色体之间随机的交换信息。在自然界的生物进化过程中，其核心作用的是生物遗传基因的重组（加上变异）^[43]。常见的几种交叉方法有：单点交叉、两点交叉、多点交叉。

单点交叉的具体操作是：在染色体中随机地选择一个交叉点，进行交叉时，交换两个染色体交叉点前或后的染色体部分，从而生成两个新的个体。单点交叉操作的过程如图 3-5 所示：

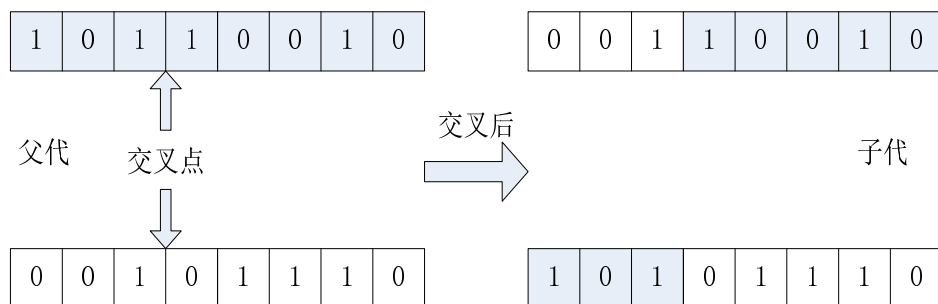


图 3-5 单点交叉操作

两点交叉的操作方法与单点交叉相类似，只是将单点交叉中的随机设定一个交叉点改为随机设定两个交叉点，然后将两个交叉点之间的基因结构进行互换。两点交叉操作的过程如图 3-6 所示：

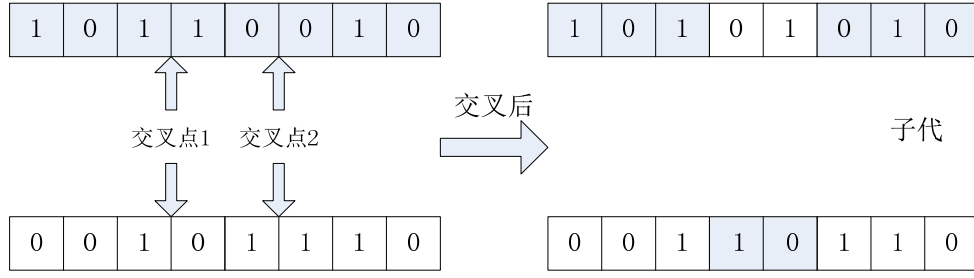


图 3-6 两点交叉操作

将上述两种交叉方法进行推广就可得到多点交叉的概念。多点交叉是指染色体的编码串中随机的设置多个（大于 2）交叉点，然后进行基因互换。多点交叉的操作过程与单点交叉和两点交叉相类似，但通常情况下很少使用多点交叉，因为该交叉方式不能有效的保存重要基因模式，会严重影响遗传算法的性能。

本文采用的交叉方式是单点交叉。对于单点交叉而言，需要解决两个问题：染色体配对以及交叉点的选择。染色体配对的主要是要避免两个相近的染色体发生交叉，因为两个相近的染色体交叉后基本等于它们原来本身，这样的交叉操作一般达不到搜索的效果。交叉点的选择对算法的性能有着很大的影响，同样的交叉对，不同的交叉点会产生完全不同的效果，有时甚至会出现无效的交叉操作，即交叉完成之后产生的新的染色体对与原染色体对完全相同。

染色体的配对首先需要分析两个染色体之间的相关性，对于染色体 $x_i = \{g_{i1}, g_{i2} \dots g_{iN}\}$ 以及染色体 $x_j = \{g_{j1}, g_{j2} \dots g_{jN}\}$ ，定义 x_i 和 x_j 的不相关指数为：

$$r(x_i, x_j) = \sum_{k=1}^N g_{ik} \oplus g_{jk} \quad (3-4)$$

配对染色体的选定过程如下：首先随机选定一个染色体 x ，染色体的配对池为 $\{y_1, y_2 \dots y_S\}$ （配对池为当前种群中还未进行配对的所有染色体），要在配对池中选定其中一个染色体和染色体 x 进行交叉操作。在标准遗传算法中，配对池中所有个体具有相同的被选概率。当种群的多样性较小，个体和其配对池中其他个体的总体差异不很明显时，交叉操作的效率会受到很大的影响。本文采取非等概率配对策略，给配对池中不相关指数较大的个体赋予较大的被选概率，配对池中个体 y_i 被选择与待配对染色体 x 进行配对交叉的概率

$P(y_i/x)$ 定义如下：

$$P(y_i/x) = \frac{1}{S} \left(1 + \lambda \frac{r(x, y_i) - r_{avg}}{r_{max} - r_{min}} \right), i = 1, 2, \dots, S \quad (3-5)$$

式中， λ 为常数， $0 \leq \lambda \leq 1$ ， S 为配对池中的染色体个数， $r_{avg} = \frac{1}{S} \sum_{i=1}^S r(x, y_i)$ ，

$$r_{\max} = \max\{r(x, y_i), i = 1, 2, \dots, S\}, \quad r_{\min} = \min\{r(x, y_i), i = 1, 2, \dots, S\}。$$

本文的交叉点的选方法为：首先确定有效交叉点区域，然后在有效交叉区域中随机选择一个位置作为交叉点，交叉区域为 (n_{\min}, n_{\max}) ， n_{\min} 以及 n_{\max} 由式（3-6）确定：

$$\begin{cases} n_{\min} = \min\{k/g_{ik} \neq g_{jk}, k=1, 2, \dots, N\} \\ n_{\max} = \max\{k/g_{ik} \neq g_{jk}, k=1, 2, \dots, N\} \end{cases} \quad (3-6)$$

3.2.3 变异操作

变异操作是指基因的某些组成部分发生突变，从而产生一些新的特性。但是对于突变的结果，无法预料是朝着有利于基因生存的方向还是使得基因的适应度比原来更低。换言之，变异操作实际上是将个体的染色体编码串中的某些基因上的值用其他等位基因来替换，从而形成新个体的过程。变异按照编码的方式可以分为实值变异和二进制变异，以0~1二进制编码为例，变异操作就是把某些基因上的值取反，即0变成1、1变成0。变异操作能够一定程度上克服交叉操作过程中的早熟收敛现象，增加了遗传种群的多样性。基本变异操作如图3-7所示：

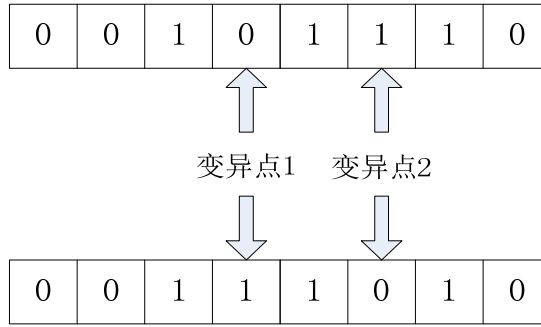


图 3-7 基本变异操作

除了上述基本的变异方法之外，还有一些其他的常用变异方法，如复制、删除、插入。换位等等^[44]。但是，由于自适应验配是遗传算法的一种特殊应用，本章采用优化的变异步骤。不同听损患者的听损类型本是不一样的，因此本章提出的变异步骤与听损类型直接相关，即针对听损类型进行定性变异。对于什么类型的听损就会对应什么类型的变异操作，这样的变异操作更有针对性，使得变异操作大体向着好的方向发展。由于前面的交叉操作可能会破坏种群中染色体的原有形状，对于某一特定听损类型的患者而言，被破坏了的参数值势必会影响患者的验配效果。因此，本章提出的优化变异步骤将破坏了的染色体变异为原来听损形状，这样的变异操作在大多数情况下是有益于种群收敛的。本章提出了三种类型的变异算子，分别对应三种听损类型：渐降型听损、陡降型听损以及盆型听损。

（1）渐降型听损变异算子

染色体 $x_i = \{g_{i1}, g_{i2}, \dots, g_{iN}\}$ ，分别对应频率[125, 250, 500, 750, 1k, 1.5k, 2k, 3k, 4k, 6k, 8k], (单位 Hz)。若以频率为横坐标，增益值（即染色体中某一基因数值）为纵坐标，可以构成一张增

益补偿图（其中横坐标采用倍频）。假定直线都经过（125hz,20db），当斜率为 13 的时候，8000hz 处的最高增益达到 100，已经达到增益上限。因此斜率的变异范围为 0.5~13。由于变异操作是局部搜索，斜率变异范围不宜过大，因此每次变异的值从[-2, 2]中选取一个随机数作用到原来的直线斜率上。由于经过前几轮算法之后，染色体上的值不可能很有规律地排列在一条直线上下附近，所以在变异操作之前，还需求得每条染色体的线性回归方程 $y = b_i * x + a_i$ ，其中：

$$b_i = \frac{\sum_{k=1}^N X_k g_{ik} - N \bar{X} \bar{g}_i}{\sum_{k=1}^N X_k^2 - N(\bar{X})^2} \quad (3-7)$$

$$a_i = \bar{g}_i - b_i \bar{X} \quad (3-8)$$

\bar{g}_i 为 $x_i = \{g_{i1}, g_{i2}, \dots, g_{iN}\}$ 的各通道值的均值，得到每组染色体的线性回归方程后，由变异后的斜率再加上固定点（125hz,20db），则可以得到一条新的增益直线。保持原来的点到线性回归方程的距离与新的点到新的增益直线的距离相等，即可得到一组新的增益值。其次，对直线局部进行变异。由于单个频率点的频率对语言理解帮助较小，所以随机产生两个变异点，对两点间的所有增益值进行 10db 以内的变异。最后，由于是渐降型听损，所以对每组增益值进行排序，最终得到新的种群。

（2）陡降型听损变异算子

陡降型听损的特点是存在某一区间，在该区间内听阈是急剧下降的。因此，针对该类型的听损进行变异时，需要考虑陡降区间以及除陡降区间之外的两端区间。对于陡降区间的变异，需要重新确立两个变异点 M1, M2，则区间[M1,M2]则代表新的陡降区间，对于该区间内的各点应进行线性变异。

由于陡降型听损在这两个区间本就是相对平坦的，不是完全随机的，所以保持“相对平坦”的特性是必要的。本文对于陡降型听损的两端区间的变异操作如下：对于变异点 M1 之前的区间，首先对该区间求均值 E1，将各点值都变异到原值与均值 E1 间的某一随机位置，并在其基础上加上[-5dB,+5dB]的随机值；M2 后的区间同上。这样的变异操作不仅会使得两端区间区域趋于平坦，同时保持了两端区间的局部搜索能力，不会随着搜索代数的提高而变得绝对平坦。

（3）盆型听损变异算子

盆型听损的特点是存在某一拐点，在该点的听损值最大，两端的听损值较小，首端端点到拐点的区间内端点值成抛物线下降趋势（凹形），拐点到末端端点的区间内端点值成抛物线上升趋势（凹形）。因此，针对该类型的听损进行变异时，需要考虑的因素有：两端端点、拐点，首端端点到拐点区间以及拐点到末端端点的区间。

盆型听损变异包括如下四个步骤。①在 1000,1500,2000hz 中选一个位置作为拐点，即“盆

形”的最低点。②将拐点位置的增益值变异到 60-90dB。③将两端端点值变异到 20 至 40dB。④计算第一个点到拐点的直线方程与抛物线方程，将之间的点变异到两个方程之间的位置。拐点到末端之间的区间变异方法同上。

3.3 辅助优化模块

辅助优化模块的作用是加速收敛，缩小搜索步长，提高搜索的准确性。该模块包含两个部分：用户偏好匹配以及种群多样性分析。用户偏好匹配是利用现有的匹配库，在分析用户偏好的同时在匹配库中寻找和当前用户偏好最相似的染色体，以该染色体替换当前种群中适应度最差的染色体。匹配库中各染色体为不同用户已完成验配最终搜索到的染色体。种群多样性分析的目的是防止早熟现象，算法每进行一代搜索，都要进行多样性分析，若当前种群多样性过低则提高交叉、变异概率，加大搜索步长，从而使得当前种群含有丰富的信息以供进一步进化操作。辅助优化模块结构图如图 3-8 所示：

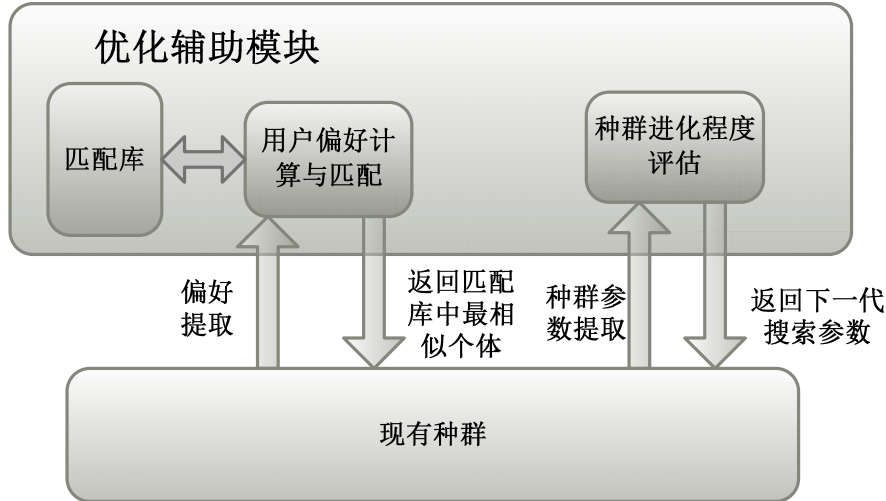


图 3-8 优化辅助模块结构图

3.3.1 用户偏好匹配

在交互式遗传算法中，用户对个体的评价体现出用户对于该染色体的偏好。因此，可以从用户对染色体的评价计算得到用户的偏好。在某个用户进行助听器自验配的过程中，可以不断的匹配库中用户的偏好，找出与当前验配用户偏好最相似的染色体，并将该染色体加入到现有进化种群中去，并且去除掉当前种群中适应度最差的染色体，以达到加速进化的目的，提高优化解的质量。

本章的染色体由不同频段的增益值构成，对于种群中的第 i 个染色体 $x_i = \{g_{i1}, g_{i2}, \dots, g_{iN}\}$ ， N 表示通道数，取值为 11，其中 g_{ik} 表示种群中第 i 个染色体的第 k 个通道（频段）的增益。本章对增益值采用的是二进制编码，每个增益值包含 7 比特，因此， g_{ik}^j 表示种群中第 i 个染色体的第 k 个通道的第 j 个比特值， g_{ik}^j 取值 0 或 1。

记用户已评价个体的集合为 $D = \{(x_i, f(x_i)), i \in \{1, 2, \dots, M\}\}$ ， M 为种群中染色体个数。记 $\alpha(*, *)$ 为比特位 g_{ik}^j 的开关量，若 g_{ik}^j 比特位值为 1，则 $\alpha(g_{ik}^j, x_i) = 1$ ，否则， $\alpha(g_{ik}^j, x_i) = 0$ 。对于基因意义单元 g_k 的第 j 个比特位 g_k^j （基因 g_k 的第 j 个等位基因意义单元），其“适应度” $F(g_k^j)$ 可由公式（3-9）计算得出：

$$F(g_k^j) = \frac{\sum_{i=1}^M \alpha(g_{ik}^j, x_i) f(x_i)}{\sum_{i=1}^M (1 - \alpha(g_{ik}^j, x_i)) * f(x_i)} \quad (3-9)$$

因为对染色体的适应度评价是人为主观评价，因此，在评价的过程中不可避免会发生“评价准则”的改变，也即人的评价会随着进化过程的推移而改变。在助听器自验配刚开始进行的时候，用户对染色体质量好坏的认知是模糊的，所以此时用户对染色体的好坏判断是要打折扣的；随着评价过程的深入，用户对染色体的好坏评价的心理标准逐渐统一，此时，患者的评价是可信的。所以，本文引入置信度函数 $R(N)$ 来模拟人主观评价的可信程度：

$$R(x) = \begin{cases} 1 + e^{-a*N_s} - e^{-a*x}, & x < N_s \\ 1, & x \geq N_s \end{cases} \quad (3-10)$$

式（3-10）中， x 表示算法进化的代数， a 为置信度系数。当患者在使用自适应验配系统的过程中，若对语音质量好坏的评估掌握的相对缓慢时， a 取值较小， N_s 取值较大；反之， a 取值较大， N_s 取值较小。本章 N_s 取值为 5， a 取值 0.15。用户对等位基因意义单元 g_k^j 的偏好 $P(g_k^j)$ 为经过置信度函数调整后的“适应度”， $P(g_k^j)$ 的计算方法由式（3-11）确定：

$$P(g_k^j) = R(x) * F(g_k^j) \quad (3-11)$$

由于基因意义单元实际意义为某一频段的增益值，采用的是二进制编码，因此不同的等位基因意义单元 g_k^j 所占的权重是不一样的。在与库中染色体匹配的时候，不同比特位所占的权重也应该是不一样的，用户对各个比特位的偏好应该乘以系数 2^j ，其中， j 为比特位 g_{ik}^j 对应的二进制位置。现仍然考虑基因意义单元 g_k ，记 $\sigma_k(u_1, u_2)$ 为用户 u_1 和 u_2 对基因意义单元 g_k 的偏好相似度， $\sigma_k(u_1, u_2)$ 的计算方法如式（3-12）：

$$\sigma_k(u_1, u_2) = \sqrt{\frac{1}{L} \sum_{j=0}^{L-1} ((P_{u1}(g_k^j) - P_{u2}(g_k^j)) * 2^j)^2} \quad (3-12)$$

式 (3-12) 中, L 表示等位基因意义单元数目, 本文取值为 7。不妨设 u_1 为当前用户, u_2 为匹配库中曾经完成验配的某一用户, 记 $\sigma(u_1, u_2)$ 为用户 u_1 与匹配用户 u_2 的相似度, 其计算方法如式 (3-13) :

$$\sigma(u_1, u_2) = \sum_{k=1}^N \sigma_k(u_1, u_2) \quad (3-13)$$

显然, $\sigma(u_1, u_2)$ 越小, 表示用户 u_1 与匹配用户 u_2 对于所有基因意义单元的相似度越高; 反之, 则表示用户 u_1 与匹配用户 u_2 对于所有基因意义单元的相似度低。在种群进化的过程中, 每进化一代, 都会在匹配库中寻找与当前用户偏好最相似的用户, 并将该匹配用户的最终验配结果 (即最优染色体) 加入到当前进化种群中, 并取代当前种群中适应度最差的个体, 加速当前种群的进化, 降低用户的疲劳度。

3.3.2 种群进化程度评估

遗传算法的主要步骤包含以下几个方面: 编码、选择、适应度评价、交叉、变异。对于传统遗传算法而言, 进化所需参数都是固定的, 不会随着进化的深入而改变, 因此会出现早熟、局部收敛等现象。本章算法在种群进化的过程中, 不断评估当前种群的进化程度, 并采用自适应交叉、变异方法, 从而显著提高搜索效率, 跳出局部最优解。

交叉操作作为遗传算法提供了一种粗粒度、大步伐的搜索策略。这种大步伐搜索策略虽然有利于 GA 全局搜索, 但却对 GA 的局部搜索性能产生了颇为不利的影响; 变异操作的作用是弥补交叉操作的不利影响, 维护种群的多样性。但是, 传统遗传算法采用的固定变异概率在算法中后期阶段性能是较差的。基于上述原因, 本文采用自适应交叉变异算法^[45]。当种群中各个体适应度趋于一致或者趋于局部最优解的时候, 增大交叉、变异概率, 用以跳出局部最优解; 而当群体适应度比较分散的时候, 降低交叉、变异概率, 以保留优良染色体。本文采用的自适应交叉变异算法是 HD-AGA 算法, 自适应交叉公式如式 (3-14) 所示, 变异概率公式如式 (3-15)。

$$P_c = \begin{cases} P_{c \max} = \left\{ \exp(-0.382) * \frac{f - f_{avg}}{f_{\max} - f_{avg}} \right\}, f > f_{avg} \\ P_{c \max}, f \leq f_{avg} \end{cases} \quad (3-14)$$

$$P_m = \begin{cases} P_{m \max} = \left\{ \exp(-0.618) * \frac{f_{\max} - f}{f_{\max} - f_{avg}} \right\}, f > f_{avg} \\ P_{m \max}, f \leq f_{avg} \end{cases} \quad (3-15)$$

其中, $P_{c \max}$ 为交叉概率上限, $P_{m \max}$ 为变异概率上限, f 为待交叉的两个染色体中适

应度较大个体的适应度值, f_{\max} 为种群中的最大适应度值, f_{avg} 为种群的平均适应度值。

上述自适应交叉变异的做法是从种群的适应度来判断当前种群的进化情况, 除此之外, 种群的多样性对优化算法的性能也有着重大的影响。种群的多样性是衡量遗传算法进化状态的重要标志, 当算法找寻到一个存在极值的区域时(不管是局部的还是全局的), 种群中的个体会不断向这一区域集中, 出现很多相同或相似的个体, 使种群的多样性变差, 从而影响遗传算法操作的效率和探索其它极值区域的能力。种群的多样性定义如下:

$$d = 1 - \frac{1}{0.5 * N * L} \sum_{k=1}^N \sum_{j=1}^L |q_k^j - 0.5| \quad (3-16)$$

式中, N 为通道数, L 为每个通道的比特位数, q_k^i 为第 k 个通道第 i 位基因取 1 的概率,

取值 $q_k^i = \frac{1}{M} \sum_{j=1}^M g_{ik}^j$, M 为种群中染色体个数。这里, d 的取值范围为 $0 \leq d \leq 1$, d 越大,

表示种群的多样性越好。当多样性 d 小于阈值 0.4 时, 交叉概率在式 (3-14) 的基础上加上 5%, 上限为 0.95, 变异概率在式 (3-15) 基础上增加 0.5%, 上限为 0.3; 否则, 分别按照式 (3-14)、(3-15) 返回交叉概率及变异概率。

3.4 实验仿真与主观测听评估

3.4.1 实验设置

为取得良好的实验结果, 首先要确定算法中的相关参数, 经过多次测试后, 本章最终选取参数如表 3-1。表 3-1 中给出了交叉, 变异概率的最大值, 因为采取的是自适应交叉变异, 实际验配过程中采用的交叉、变异概率一般小于给定的最大交叉、变异概率。此外, 从表中可以看出, 变异概率的取值相对于一般的遗传算法的应用是偏大的, 因为本章采取了优化变异操作, 变异步骤不同于传统算法。传统遗传算法的变异是为了局部搜索, 而本章的变异操作不仅有局部搜索的功能, 还可以调整被交叉操作破坏了结构的染色体, 该变异操作对于全局搜索也是有益的, 因此, 本章采取的变异概率较一般算法而言取值较大。测试信号来自于江苏省人民医院制作的助听器验配测试语音库, 语音库对应的文本由男声和女声各朗读一遍, 信号的采用频率为 44.1kHz, 在处理前先降采样为 8kHz。语音信号的声压级调整为 55 dB、65 dB、75 dB, 分别代表日常生活中讲话的小声、中声、大声。测试在消音室进行, 测试语音采取随机播放的方式, 通过头戴式耳机森海塞尔 HD650 单耳给出。

表 3-1 GA 参数的选择

优化算法	GA
编码方式	二进制编码
选择方法	轮盘赌、精英保留
最大交叉概率	0.95
最大变异概率	0.3
初始化	随机初始化
染色体数目	11
种群大小	8
最大遗传代数	20
适应度函数	主观 5 级评估、超几何算子
多样性阈值	0.4

3.4.2 优化变异仿真实验

本章提出的优化变异操作包含三个不同的类型，分别针对渐降型听损、陡降型听损以及盆型听损。本小节变异仿真实验是为了检验 3.2.3 小节提出的优化变异操作能否达到预期功能。为了能直观体现各个变异算子的性能，仿真实验首先提取了算法进化过程某一种群的各个染色体，在此基础上直接进行变异，跳过了交叉步骤，增大变异概率到 0.7，这样可以方便的看出变异前后的染色体变化，图 3-9 是变异操作前后种群的染色体变化对比图：

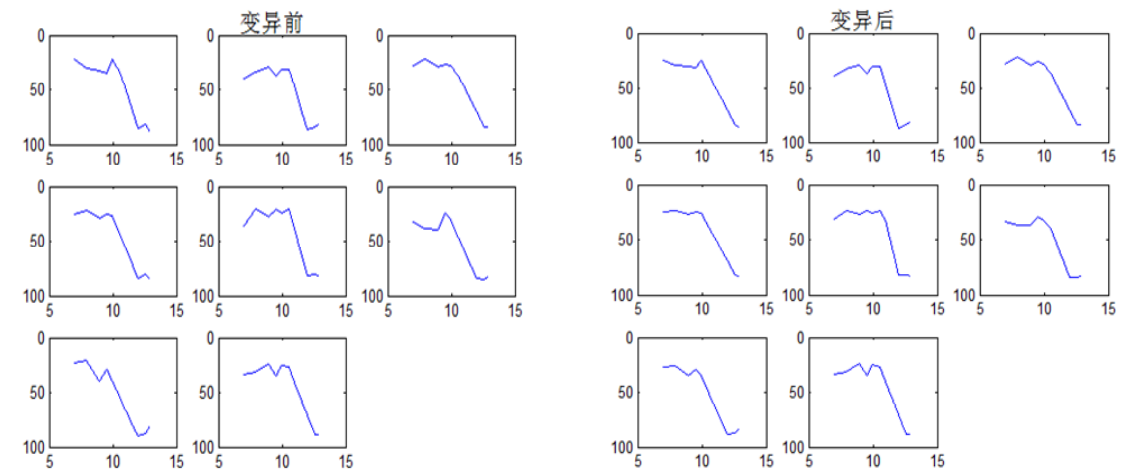


图 3-9 变异前后种群对比图（陡降型听损）

通过对比可以看出，除第 2、第 3 和第 8 个染色体外，其余五个染色体均产生了变异，变异的趋势是两端向均值靠拢，且有随机波动。通过对比可以看出，优化变异操作基本实现了预期功能。对渐降型变异算子以及盆型变异算子采用同样的测试方法，仿真结果分别如图 3-10、3-11 所示：

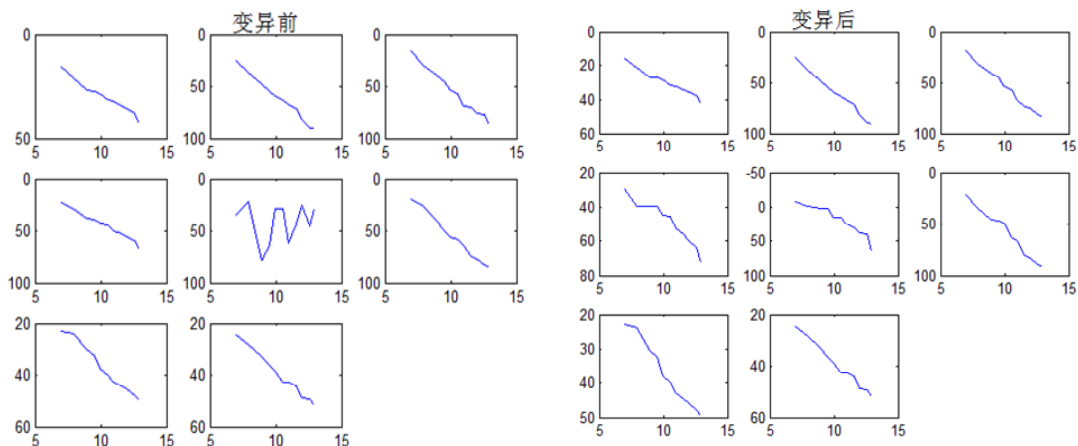


图 3-10: 变异前后种群对比图 (渐降型听损)

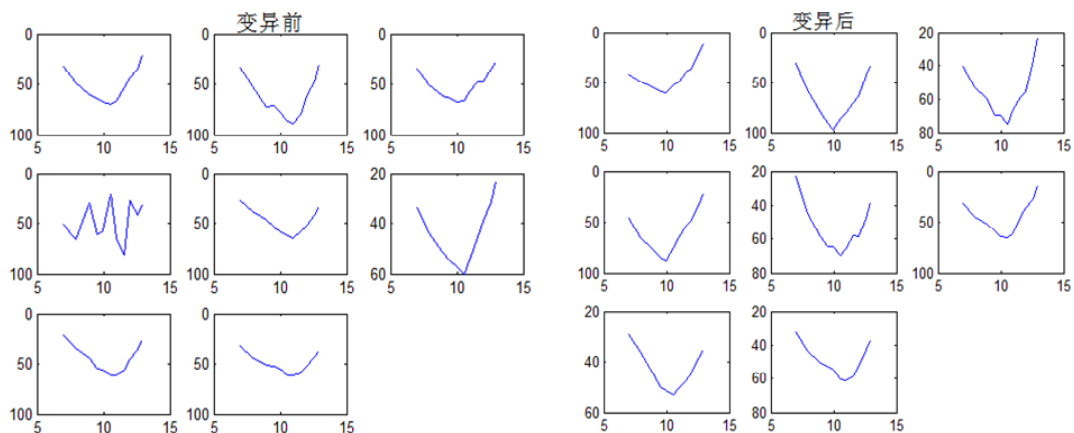


图 3-11: 变异前后种群对比图 (盆型听损)

从图 3-10 可以看出,除了第 2 和第 8 个染色体,其余六个染色体发生了变异。各染色体依然保持了“渐降”的趋势,这对全局优化是有益处的。注意到种群中的第五个染色体在进化过程中遭到了破坏,但是经过变异操作后,该染色体被修复成正常渐降类型的染色体。图 3-11 是盆型染色体变异前后的种群对比图,可以看到种群中的所有染色体均发生变异。变异操作将一个“盆形”曲线变异到另外一个“盆形”曲线,两个曲线没有关联。可以看到种群中第四个染色体在进化过程中也遭到了破坏,经过变异操作后成为了新的“盆型”染色体。

本章提出的优化变异操作的总体思想是:对应于某一种特定类型的听损,变异操作就相应调整针对该类型,使得变异后的染色体大体上符合这样的听损类型,同时在变异过程中保持一定的随机性。一般而言,经过优化变异操作后,染色体从某一特定类型随机变异成为该类型的另一个染色体,从这个意义上讲变异前后染色体质量好坏是不可得知的。但是,对于被破坏的染色体而言,变异操作总能将该染色体还原成相应听损类型的染色体。因此,总体而言,本章提出的优化变异操作对于种群进化是有益的。

3.4.3 主观测听实验

对于数字助听器的自适应验配，最佳实验方法为：找一些听损患者进行实际测试，看他们在与软件交互过程中，是否可以在相对满意的时间内得到相应的听力补偿。然而，限于实验条件关系，无法找到相应的志愿者参与到实验中来。现反向行之，人为改变不同频段的增益，使一段语音在正常人耳中听起来变得不正常，即模拟听损患者耳中的语音效果，通过不断的人机交互，看最终能否获得相应增益以补偿原来的听损。

由于很难直接判定验配的“绝对”效果，所以本文采取相对方法比较“自适应验配”对于语音质量的提升。实验的设计如表 3-2。实验的进行分为两个阶段。第一阶段：用本章设计的验配方法进行验配，得到“自适应验配”的相关验配参数；第二阶段：选定五位实验对象（无听损）以及男女各 10 段语音，对于每一位实验对象，随机让其听取自适应验配处理过后的语音以及未经处理的（含听损）语音，比较二者的语音质量优劣，并以语音质量的“提升”和“下降”进行评估。

表 3-2 实验设计

阶段一（验配阶段）	A: 自适应验配
	B: 未经验配
阶段二（评估阶段）	语音质量评估：比较 A、B 两种语音质量

以测试者 A 为例，在实验阶段一，对于 16 通道的人为听损设定为固定值 [20,23,25,25,27,31,35,40,44,44,50,50,56,56,56,62]，然后通过不断的人机交互，可以得到每一代的相关参数，将其保留到相关文件中，格式如表 3-3 所示，从表 3-3 中即可看出每代算法所得到的增益参数。由于将整个遗传算法的参数全部展示出来会显得过于冗长，现直接给出验配结果。从第五代以后的最优解 [18,25,28,28,31,36,41,42,45,58,58,62,62,62,65] 一直保持到 12 代，虽然从数值上与实验设置听损有一些差距，但该组增益值所对应的听觉效果已经极为接近正常的语音，该组参数值将作为实验阶段一的最终结果，并以该结果作为实验阶段二的依据。

表 3-3 算法每代参数记录

语音文件名称	参数位置	参数数值	评价值
tapestry_05_30_14.wav	0 代 1 组	20,25,25,25,30,30,30,40,45,45,50,50,55,55,55,60	良
tapestry_05_30_14.wav	0 代 2 组	16,35,44,44,46,53,62,71,75,75,78,78,90,90,90,95	差
tapestry_05_30_14.wav	0 代 3 组	22,23,31,31,40,41,44,46,47,47,55,55,59,59,59,59	良
tapestry_05_30_14.wav	0 代 4 组	22,28,41,41,42,50,51,57,60,60,65,65,78,78,78,81	中
tapestry_05_30_14.wav	0 代 5 组	19,24,27,27,29,33,35,37,38,38,40,40,44,44,44,46	中
tapestry_05_30_14.wav	0 代 6 组	17,31,41,41,41,45,55,63,63,63,69,69,72,72,72,81	差
tapestry_05_30_14.wav	0 代 7 组	20,25,28,28,29,30,30,32,33,33,36,36,39,39,39,43	中
tapestry_05_30_14.wav	0 代 8 组	21,23,24,24,25,30,31,31,36,36,37,37,38,38,38,43	中
tapestry_05_30_14.wav	1 代 1 组	18,20,23,23,25,26,29,30,32,32,32,32,35,35,35,36	中
tapestry_05_30_14.wav	1 代 2 组	16,22,23,23,29,42,47,48,51,51,53,53,54,54,54,55	差
tapestry_05_30_14.wav	1 代 3 组	22,22,22,22,25,31,37,38,42,42,57,57,62,62,62,66	良
tapestry_05_30_14.wav	1 代 4 组	20,25,28,28,29,30,30,32,33,33,36,36,39,39,39,43	差
tapestry_05_30_14.wav	1 代 5 组	21,23,24,24,25,30,31,31,36,36,37,37,38,38,38,43	差
tapestry_05_30_14.wav	1 代 6 组	18,34,37,37,42,60,62,69,70,70,77,77,79,79,79,89	劣
tapestry_05_30_14.wav	1 代 7 组	18,37,45,45,46,50,61,70,71,71,78,78,80,80,80,91	差
tapestry_05_30_14.wav	1 代 8 组	20,25,25,25,30,30,30,40,45,45,50,50,55,55,55,60	良

在实验第二阶段,测试者会听到相同语音源的两种语音,一种是经过自适应验配方法处理后的语音,另一种是未经处理的语音(只含听损)。为消除测试者的先验心理,两种语音随机播放。每次听完一段对话的两种语音后,测试者被要求选定二者中相对语音质量好的一种,也可选择两者质量“相差不大”。在所有五位测试者听完所有 20 段对话后,实验结束并统计实验结果。实验结果如表 3-4 所示。

表 3-4 实验阶段二测试结果统计

测试者编号	语音类型(男/女)	验配效果优于未经处理效果	验配效果劣于未经处理效果
A	男	9/10	0/10
	女	10/10	0/10
B	男	10/10	0/10
	女	7/10	0/10
C	男	9/10	0/10
	女	10/10	0/10
D	男	7/10	1/10
	女	8/10	0/10
E	男	9/10	0/10
	女	10/10	0/10
统计结果		89/100	1/100

从表 3-4 可以看出，对于所有 100 段语音（ $5*10*2$ ），有 89 段语音被认为效果优于未经验配，1 段被认为效果劣于未经验配的语音质量，其余 10 段语音被认为两者相差不大。因此，可以得出结论，自适应验配可以明显提升语音质量。而且，验配的效果对于男女声源几乎没有差别。

3.5 本章小结

本章第一小节首先阐述了交互式遗传算法理论基础。第二小节重点讨论了优化算法中各遗传算子的设计，包括染色体编码、选择操作、交叉以及变异操作。第三小节提出了辅助优化模块，其中包括用户偏好匹配以及种群进化程度评估，该模块对算法的搜索效率有了极大地提升，大大缩短了收敛时间。第四小节进行了仿真实验以及主观测听实验，仿真实验的结果表明本章提出的优化变异步骤是按照预期的方式进行有效操作，提高了搜索的准确性；主观测听实验的结果显示：经过本章提出的基于交互式遗传算法的自适应验配方法进行验配后，可以明显提升语音质量。

第四章 基于高斯过程回归的自适应验配方法

上一章提到的交互式遗传算法是人工智能的一种,这是一种受到了自然界的启发而开发出来的优化算法。而机器学习是人工智能的另外一个重要分支,机器学习算法指的是利用现有数据分析得到规律,从而对新的数据实现预测的一类算法。近年来,机器学习相关理论研究得到了迅速的发展,在模型泛化能力的研究、基于机器学习的数据挖掘应用于商业、入侵系统检测等诸多方面都有了长足的进步^[46]。随着机器学习算法的普及,高斯过程模型在该领域得到了广泛的应用^{[47][48][49]}。高斯过程模型包括分类模型(离散输出)以及回归模型(连续输出),其中,高斯过程回归方法是一种监督学习中的回归方法^[50]。传统的监督学习无法满足本课题的研究,因其不具备交互性。本章中,采用基于主动学习的高斯回归方法,让听损患者对待估计参数进行一系列的标签,最终获得最适合受试者的助听器参数配置。

4.1 高斯过程回归预测理论基础

4.1.1 高斯过程回归预测

对于一个输入集合 $\mathbf{X} = \{\mathbf{x}_i \in R^d \mid i = 1, \dots, n\}$, 有观测集合 $\mathbf{y} = \{y_i \mid i = 1, 2, \dots, n\}$, 输入集合 \mathbf{X} 与观测集合 \mathbf{y} 构成训练集 $D = \{\mathbf{X}, \mathbf{y}\}$ 。回归任务的本质是找到一个能反映出输入 \mathbf{X} 与输出 \mathbf{y} 之间的函数映射关系: $(f(\cdot): R^d \mapsto R)$, 然后对新的输入 \mathbf{x}_* 预测出最有可能出现的 $f(\mathbf{x}_*)$ 。在监督学习中,有两类方法可以用来确定上述函数映射关系: 参数回归法以及贝叶斯回归法^[51]。1) 参数回归法指的是输出由输入 \mathbf{x} 与参数 \mathbf{w} 共同定义: $f(\mathbf{x}; \mathbf{w})$ 。该方法最核心的问题是如何找到一组“最优参数” \mathbf{w} , 使得该模型可以对将来的输入值进行完美的预测。通常可以采用最小化损失函数 $L(\mathbf{w})$ 的方法解决上述问题, 最小二乘多项式回归是其中常用的一种损失函数。2) 贝叶斯回归法需要定义一个函数分布并赋予每种可能的函数一个先验概率, 可能性大的函数对应较高的先验概率。但是, 函数的选择有无限种可能, 使用高斯过程回归可以有效解决上述问题。由于高斯过程回归具有容易实现、输出具有概率意义、对非线性问题具有良好适应性等优点, 本章采用该方法来实现助听器的自适应验配。

高斯过程是任意有限个随机变量的集合, 这有限个随机变量符合联合高斯分布。一个高斯过程的性质完全由其均值和协方差函数决定, 均值 $m(\mathbf{x})$ 以及协方差 $k(\mathbf{x}, \mathbf{x}')$ 的函数定义如下:

$$m(\mathbf{x}) = E[f(\mathbf{x})] \quad (4-1)$$

$$k(\mathbf{x}, \mathbf{x}') = E[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))] \quad (4-2)$$

则高斯过程可以定义为： $f(\mathbf{x}) \sim GP(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}'))$ 。一般会对数据进行预处理，使均值为 0，这样可以使得符号简洁。假设观测过程受到加性噪声的污染，则观测结果 y 的数学形式如下：

$$y = f(\mathbf{x}) + \varepsilon \quad (4-3)$$

其中， \mathbf{x} 为输出向量， ε 为加性噪声污染。通常，我们假设加性噪声 ε 符合均值为 0，方差为 σ_n^2 的正态分布，即 $\varepsilon \sim N(0, \sigma_n^2)$ 。有了上述定义以及假设后，我们可以得到观察值 \mathbf{y} 的先验分布：

$$\mathbf{y} \sim N(0, \text{var}(\mathbf{y}) + \sigma_n^2 \mathbf{I}) \quad (4-4)$$

以及观测值 \mathbf{y} 和预测值 f_* 的联合先验分布：

$$\begin{bmatrix} \mathbf{y} \\ f_* \end{bmatrix} \sim N\left(0, \begin{bmatrix} \mathbf{K} & \mathbf{K}_*^T \\ \mathbf{K}_* & K_{**} \end{bmatrix}\right) \quad (4-5)$$

\mathbf{K} 、 \mathbf{K}_* 以及 K_{**} 的定义如下：

$$\mathbf{K} = \begin{bmatrix} k(x_1, x_1) & k(x_1, x_2) & \dots & k(x_1, x_n) \\ k(x_2, x_1) & k(x_2, x_2) & \dots & k(x_2, x_n) \\ \dots & \dots & \dots & \dots \\ k(x_n, x_1) & k(x_n, x_2) & \dots & k(x_n, x_n) \end{bmatrix} \quad (4-6)$$

$$\mathbf{K}_* = [k(x_*, x_1) \quad k(x_*, x_2) \quad \dots \quad k(x_*, x_n)] \quad (4-7)$$

$$K_{**} = k(x_*, x_*) \quad (4-8)$$

根据高斯过程的性质可以得出， $f_* | \mathbf{X}, \mathbf{y}, x_*$ 同样服从高斯分布：

$$f_* | \mathbf{X}, \mathbf{y}, x_* \sim N(\mu_*, \Sigma_*) \quad (4-9)$$

其中， $\mu_* = \mathbf{K}_* \mathbf{K}^{-1} \mathbf{y}$ ， $\Sigma_* = K_{**} - \mathbf{K}_* \mathbf{K}^{-1} \mathbf{K}_*^T$ 。 μ_* 即为对测试输出点 \mathbf{x}_* 预测的输出值， Σ_* 为预测输出值的方差^[52]。下面用一个简单的例子展示高斯过程回归预测的整个过程。给定一个带噪声的样本，如图 4-1 所示，样本包含 6 个点，在这样的条件下要对位置输入点 $\mathbf{x}_* = 0.2$ 预测输出。

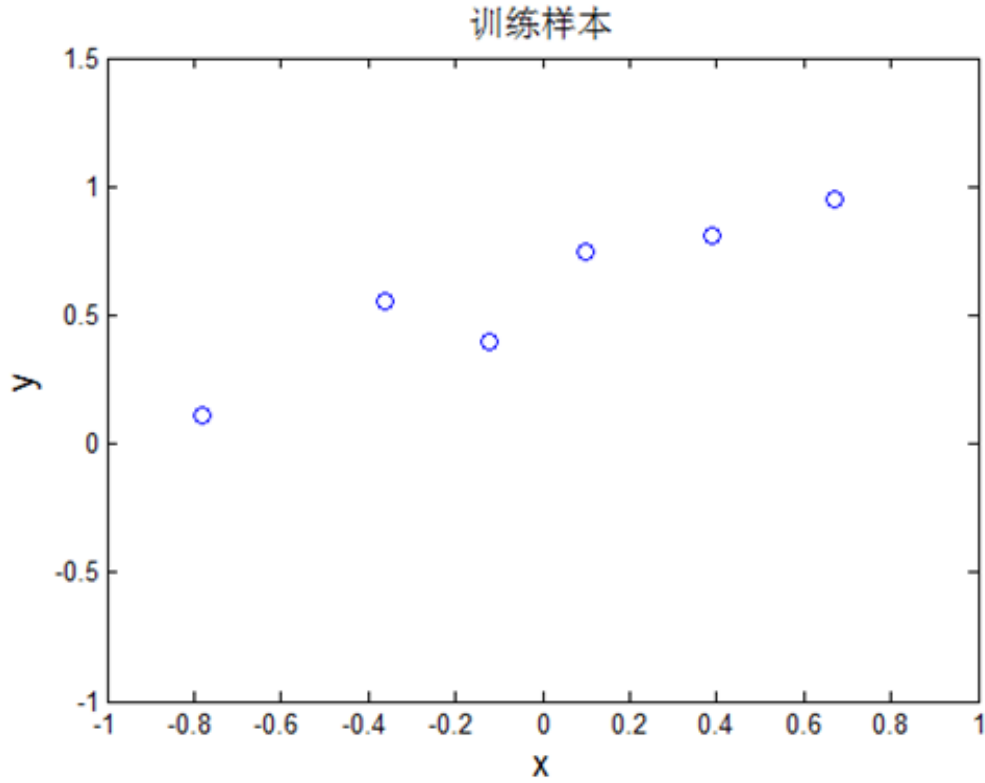


图 4-1 训练样本

如图 4-1 所示, 现有训练样本: $\mathbf{x}_* = [-0.78 \ -0.36 \ -0.12 \ 0.10 \ 0.39 \ 0.67]$, 对应的观察值为 $\mathbf{y} = [0.11 \ 0.55 \ 0.40 \ 0.75 \ 0.81 \ 0.95]$, 假设我们选取的协方差函数为:

$$k(x, x') = \sigma_f^2 \exp \left[\frac{-(x - x')^2}{2l^2} \right] + \sigma_n^2 \delta(x, x') \quad (4-10)$$

其中, $\delta(x, x')$ 为克罗内克函数。在 $\sigma_n = 0.25$, $\sigma_f = 1.38$, $l = 1$ 的条件下, 由公式(4-6)、(4-7)、(4-8) 可以计算出:

$$\mathbf{K} = \begin{bmatrix} 1.97 & 1.74 & 1.53 & 1.29 & 0.96 & 0.67 \\ 1.74 & 1.97 & 1.85 & 1.71 & 1.44 & 1.12 \\ 1.53 & 1.85 & 1.97 & 1.86 & 1.67 & 1.39 \\ 1.29 & 1.71 & 1.86 & 1.97 & 1.83 & 1.62 \\ 0.96 & 1.44 & 1.67 & 1.83 & 1.97 & 1.83 \\ 0.67 & 1.12 & 1.39 & 1.62 & 1.83 & 1.97 \end{bmatrix}$$

$$\mathbf{K}_* = [1.18 \ 1.63 \ 1.81 \ 1.89 \ 1.87 \ 1.71]$$

$$K_{**} = 1.97$$

然后根据公式 (4-9) 的预测分布, 可以计算出 $\mu_* = 0.74$, $\Sigma_* = 0.08$ 。最后, 我们对于输入 $\mathbf{x}_* = 0.2$ 的预测为: 0.74。

4.1.2 高斯模型的训练

上一小节我们在选定协方差函数为公式 (4-10), 以及给定超参数 $\sigma_n = 0.3$, $\sigma_f = 1.27$, $l = 1$ 的条件下, 对输入 \mathbf{x}_* 进行了预测。然而, 利用高斯过程回归方法进行预测的一般都是较为复杂的实际问题。不同的协方差函数形式的选择产生的预测结果有较大差异, 进而对模型预测的准确定产生较大影响, 协方差函数中的参数的选择也是同样如此。因此, 在解决实际问题的过程中, 我们通常需要解决两方面的问题: 1) 协方差函数形式的选择。2) 协方差函数中超参数的确定。这两方面综合起来也就是所谓的高斯模型的训练。这一小节将对这两方面进行简单介绍。

1. 几种常见的协方差函数形式

(1) 平方指数协方差函数

平方指数协方差函数形式如下:

$$k_{SE}(r) = \exp\left(-\frac{r^2}{2l^2}\right) \quad (4-11)$$

其中, $r = |\mathbf{x} - \mathbf{x}'|$, l 定义了协方差函数的平滑度。从式 (4-11) 可以看出, 平方指数形式的协方差函数具有无限可微的性质, 这也就意味着由其定义的高斯过程是任意阶均方可微的。平方指数协方差函数是在该领域应用最广泛的协方差函数形式^[53]。

(2) γ 指数协方差函数

γ 指数协方差函数形式如下:

$$k(r) = \exp\left(-(r/l)^\gamma\right), \quad 0 < \gamma \leq 2 \quad (4-12)$$

从式 (4-12) 中可以看出, 该类协方差函数既包含平方指数协方差函数也具有指数协方差函数的形式。显然, 只有当 $\gamma = 2$ 的时候, γ 指数协方差函数才是无限可微的。因此, 该类协方差函数形式缺少了一点灵活性。

(3) Matern 类协方差函数

Matern 类协方差函数形式如下:

$$k_{Matern}(r) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu}r}{l}\right)^\nu K_\nu\left(\frac{\sqrt{2\nu}r}{l}\right) \quad (4-13)$$

上式中, l 和 ν 均为正数, K_ν 是修正贝塞尔函数。从上式中不难发现, 当 $\nu \rightarrow \infty$ 的时候, Matern 类协方差函数退化为平方指数协方差函数。当且仅当 $k < \nu$ 时, 由 Matern 类协

方差函数确定的随机过程是 k 次均方可微的。

(4) 有理二次型协方差函数

有理二次型协方差函数形式为：

$$k_{RQ}(r) = \left(1 + \frac{r^2}{2\alpha l^2}\right)^{-\alpha} \quad (4-14)$$

上式中的 α 和 l 都要求大于零，同样的，当 $\alpha \rightarrow \infty$ 的时候，有理二次型协方差函数也会趋近于平方指数协方差函数。有理二次型协方差函数的可微性由参数 α 确定，若对于某一个确定的 α 值， $k_{RQ}(r)$ 是无穷可微的，则由其形成的随机过程是无穷阶均方可微的。

上面讲述了四类常见的协方差函数，此外还有不少其他形式的协方差函数，如修正贝塞尔、t 分布等，这里不再赘述。在选择何种形式协方差函数的时候，可以利用现有样本对不同形式的协方差函数进行测试，取其中预测结果较好的一个。

2. 超参数的确定

平方指数协方差是一类最典型的协方差函数形式，本节中关于如何确定超参数的讨论将以该类协方差形式为例。首先将式 (4-10) 推广到多维：

$$k(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{x}')^T \mathbf{M}^{-1}(\mathbf{x} - \mathbf{x}')\right) + \sigma_n^2 \delta(\mathbf{x}, \mathbf{x}') \quad (4-15)$$

其中， $\mathbf{M} = \text{diag}[l_1^2, \dots, l_n^2]$ ，为 n 阶对角矩阵。由 $l_r^2 (r=1, \dots, n)$ 、 σ_f^2 、 σ_n^2 组成了一个超参数集合 $\theta = \{l_1^2, \dots, l_n^2, \sigma_f^2, \sigma_n^2\}$ 。求取超参数集合 θ 的一般方法是求 θ 的最大后验估计。由贝叶斯原理可知：

$$p(\theta | \mathbf{y}, \mathbf{X}) = \frac{p(\mathbf{y} | \mathbf{X}, \theta)p(\theta)}{p(\mathbf{y} | \mathbf{X})} \quad (4-16)$$

对上式求最大值的結果即为 θ 的最大后验估计。但是，在绝大多数情况下 θ 的先验分布是无法获知的。为此采取的一般办法是假设 θ 是均匀分布的，即 θ 等可能的取任意值，因此可以得到以下近似推导：

$$\arg \max(p(\theta | \mathbf{y}, \mathbf{X})) \approx \arg \max(p(\mathbf{y} | \mathbf{X}, \theta)) \quad (4-17)$$

则， $\hat{\theta} = \arg \max(p(\mathbf{y} | \mathbf{X}, \theta))$ ， $p(\mathbf{y} | \mathbf{X}, \theta)$ 即为我们通常所说的边缘似然。所以，对 θ 求最大后验估计等价于求极大似然估计。对边缘似然 $p(\mathbf{y} | \mathbf{X}, \theta)$ 取负对数得到：

$$L(\theta) = \log p(\mathbf{y} | \mathbf{x}, \theta) = \frac{1}{2} \mathbf{y}^T \mathbf{K}^{-1} \mathbf{y} + \frac{1}{2} \log |\mathbf{K}| + \frac{n}{2} \log 2\pi \quad (4-18)$$

上式对超参数 θ 求偏导：

$$\frac{\partial L(\theta)}{\partial \theta_i} = \frac{1}{2} \text{trace} \left((\alpha \alpha^T - \mathbf{K}^{-1}) \frac{\partial \mathbf{K}}{\partial \theta_i} \right) \quad (4-19)$$

其中， \mathbf{K} 为协方差矩阵， $\alpha = \mathbf{K}^{-1} \mathbf{y}$ 。然后利用相关优化算法（如牛顿法、共轭梯度法等）求式（4-18）的最小值，即可得到 θ 的近似最大后验估计。一旦超参数 θ 确定下来，就可利用式（4-9）分布中的均值预测对将来的输入进行预测。利用极大似然近似法求解得到 θ 不仅容易实现，且得到的超参数形成的高斯模型具有很高的准确性。因此，该方法也是求解高斯过程中的超参数最常用的方法。

4.2 自适应验配的高斯过程模型

本章研究了一种基于机器学习的助听器自适应验配方法（MLHAF）。MLHAF 通过听损患者对输出声音的自我感知来优化助听器参数，而不是由听力专家通过患者的模糊表述来实现助听器验配。通过听损患者对不同参数配置的助听器输出语音的直观听觉感受以及比较，MLHAF 可以找到患者对输出声音的偏好从而确定最佳助听器参数。听损患者的主动参与在心理学层面有更好的所有权意识，因此对于听力的康复有着更好的效果^[54]。

在 MLHAF 中，我们假定听损患者的听觉感知可由不能直接测量的听觉心理响应函数（PRF, Psychoacoustic response function）的形式刻画。当听损患者比较两个不同的激励信号时，PRF 的输出量可以决定听损患者偏向两个激励信号中的哪一个信号。PRF 不能够被直接测量，由于各种不可控因素都可影响听损患者的直观听觉感受，因此我们认为 PRF 是随机的。显然，PRF 只能通过听损患者对特定激励信号的评估来估计。特定的助听器参数 \mathbf{x}_i ，直接决定了助听器的声学输出特性。因此，PRF 是一个由 d 维助听器参数组成向量 \mathbf{x} 决定的函数 $f(\mathbf{x})$ ， $\mathbf{x} = [x_1, x_2, \dots, x_d]^T$ 。注意，本章中的 MLHAF 既可以用于优化元参数，也可用于优化算法中的参数，对于这两种不同的参数不加以区别。在 MLHAF 中，PRF 的模型选择用高斯过程（GP）来刻画^{[55][56]}。患者对于助听器参数的评估采取的是成对比较的方式，通过这种方式，听损患者不需要记住之前评估结果，因此减轻了听损患者的记忆负担。为了尽量减少估计 PRF 所需要的评估次数，听损患者不仅要评估出两组测试信号中哪一个更加满意，还要给出一组参数优于另一组参数的程度。经过一系列这样的观察评估后，听损患者的 PRF 分布将被更新，使得该 PRF 分布输出最大值的参数设置即为最佳的助听器参数配置。从模型的角度上而言，PRF 由 d 维助听器参数 \mathbf{x} 决定，主要任务就是完成基于 PRF 的全局多参数优化。

4.2.1 听觉心理响应函数的高斯建模

听损患者的 PRF 基于高斯过程建模，本小节将具体阐述基于偏好程度估计的非标准高斯过程的建模步骤。高斯过程回归是贝叶斯无参数框架的核方法^[57]，对任意输入向量

$\mathbf{x}, \mathbf{x}' \in R^d$ ，高斯过程可以定义为：

$$f(\mathbf{x}) \sim GP(0, k(\mathbf{x}, \mathbf{x}')_{\theta_c}) \quad (4-20)$$

其中， $k(\cdot, \cdot)_{\theta_c}$ 是关于参数 θ_c 的协方差函数。一般而言，协方差函数决定了分布的平滑程度。各向同性均方指数核函数是经常使用的协方差函数形式，其解析形式如下：

$$k_{SE}(\mathbf{x}, \mathbf{x}') = \sigma_f \exp\left(-\frac{1}{2\lambda}(\mathbf{x} - \mathbf{x}')^T(\mathbf{x} - \mathbf{x}')\right) \quad (4-21)$$

不同平滑参数 λ 对应的高斯过程先验分布如图 4-2 所示：

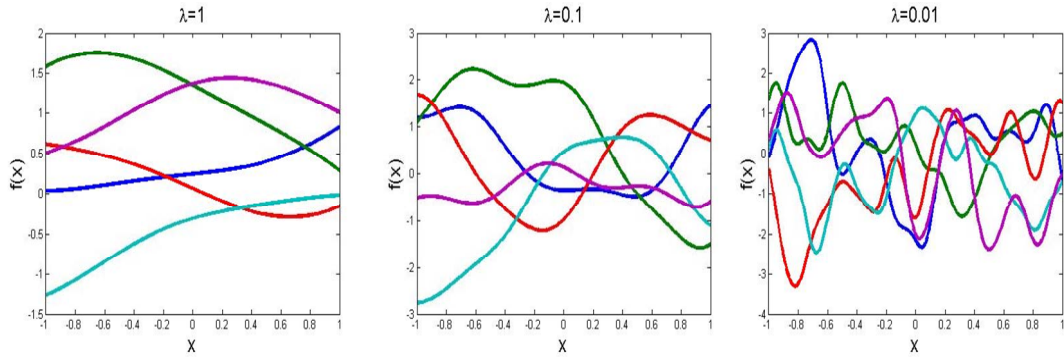


图 4-2 三种平滑参数 λ 对应的高斯过程先验分布对比图

从图 4-2 中的不同平滑参数的对比中可以看出， λ 取值越大，构建的高斯模型就越平滑。本章采用如下形式的协方差函数：

$$k(\mathbf{x}, \mathbf{x}') = \sigma_f \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{x}')^T \mathbf{P}^{-1}(\mathbf{x} - \mathbf{x}')\right) \quad (4-22)$$

其中， $\mathbf{P} = \text{diag}([\lambda_1, \dots, \lambda_d]^T)$ ，为平滑系数对角矩阵。协方差形式确立后，超参数 θ_c 也相应确定下来： $\theta_c = \{\sigma_f, \lambda_1, \lambda_2, \dots, \lambda_d\}$ 。

高斯过程是一组随机变量的集合，任意有限个高斯变量服从联合高斯分布。对于一个输入集合 $\mathbf{X} = \{\mathbf{x}_i \in R^d \mid i = 1, \dots, n\}$ ，随机向量 $\mathbf{f} = [f(\mathbf{x}_1), \dots, f(\mathbf{x}_n)]^T$ 服从联合高斯分布，其解析形式如下：

$$p(\mathbf{f} | \mathbf{X}, \theta_c) = N(\mathbf{f} | \mathbf{0}, \mathbf{K}) \quad (4-23)$$

其中， \mathbf{K} 是关于参数 θ_c 的协方差矩阵， $[\mathbf{K}]_{i,j} = k(\mathbf{x}_i, \mathbf{x}_j)_{\theta_c}$ ， $N(\mathbf{z} | \mu, \Sigma)$ 表示向量 \mathbf{z} 服从均值为 μ ，方差为 Σ 的标准正态分布。

给定观测值集合 \mathbf{y} 后，一旦确定了概率密度 $p(\mathbf{y} | \mathbf{f}, \theta_c)$ ，对于有限高斯变量组成的向量

\mathbf{f} 而言，其后验分布可以通过贝叶斯公式推导得出：

$$p(\mathbf{f} | \mathbf{y}, \mathbf{X}, \theta) = \frac{p(\mathbf{y} | \mathbf{f}, \theta_\ell) p(\mathbf{f} | \mathbf{X}, \theta_c)}{p(\mathbf{y} | \mathbf{X}, \theta)} \quad (4-24)$$

$$= \frac{p(\mathbf{y} | \mathbf{f}, \theta_\ell) p(\mathbf{f} | \mathbf{X}, \theta_c)}{\int p(\mathbf{y} | \mathbf{f}, \theta_\ell) p(\mathbf{f} | \mathbf{X}, \theta_c) d\mathbf{f}} \quad (4-25)$$

其中，超参数 $\theta = \{\theta_\ell, \theta_c\}$ ，包含似然参数和协方差参数。

连续有界响应的建模可以用基于 β 分布的似然函数来实现，该方法针对成对比较观察评估^{[58][59]}。这与本章提出的自适应验配方法是一致的，因此，该方法可以应用到目前的研究中。

对于任意两个用于评估的相互独立的输入变量 $\mathbf{x}_{u_k}, \mathbf{x}_{v_k} \in \mathbf{X}$ ，其中 $u_k, v_k \in \{1, \dots, n\}$ ，有成对观察结果集合 $\mathbf{y} = \{y_k \in (0, 1) | k = 1, \dots, m\}$ 。若听损患者偏向输入变量 \mathbf{x}_{u_k} ，在观察结果上表现为 $y_k \rightarrow 0$ ；若听损患者偏向于输入变量 \mathbf{x}_{v_k} ，在观察结果上则表现为 $y_k \rightarrow 1$ ；若听损患者觉得对于两个输入变量 $\mathbf{x}_{u_k}, \mathbf{x}_{v_k} \in \mathbf{X}$ ，输出信号的满意程度是相当的，则 $y_k = 0.5$ 。对于向量 $\mathbf{f}_k = [f(\mathbf{x}_{u_k}), f(\mathbf{x}_{v_k})]^T$ ， $p(y_k | \mathbf{f}_k)$ 可由下式给出：

$$p(y_k | \mathbf{f}_k) = \text{Beta}(y_k | \alpha(\mathbf{f}_k), \beta(\mathbf{f}_k)) \quad (4-26)$$

其中， $\alpha(\mathbf{f}_k) = \nu \mu(\mathbf{f}_k)$ ， $\beta(\mathbf{f}_k) = \nu(1 - \mu(\mathbf{f}_k))$ ， ν 是影响函数 $\mu(\cdot)$ 的离散参数，函数 $\mu(\cdot)$ 定义如下：

$$\mu(\mathbf{f}_k, \sigma) = \Phi\left(\frac{f(\mathbf{x}_{v_k}) - f(\mathbf{x}_{u_k})}{\sqrt{2}\sigma}\right) \quad (4-27)$$

上式中， $\Phi(\cdot)$ 是标准高斯分布（均值为 0，方差为 1）的累积密度函数， σ 是斜率参数。

综上可得：

$$p(y_k | \mathbf{f}_k, \theta_\ell) = \text{Beta}(y_k | \nu \mu(\mathbf{f}_k, \sigma), \nu \mu(1 - (\mathbf{f}_k, \sigma))) \quad (4-28)$$

其中， $\theta_\ell = \{\nu, \sigma\}$ 。

假设各个观察值之间是相互独立的，则 $p(\mathbf{y} | \mathbf{f}, \theta_\ell)$ 有如下形式：

$$p(\mathbf{y} | \mathbf{f}, \theta_\ell) = \prod_{k=1}^m p(y_k | \mathbf{f}_k, \theta_\ell) \quad (4-29)$$

$p(\mathbf{f}|\mathbf{y}, \mathbf{X}, \theta)$ 的解析形式目前还无法得出，因为式 (4-25) 的分母部分是积分形式，该部分解析形式比较困难，所以我们无法采用分析的方式来直接得到得到 $p(\mathbf{f}|\mathbf{y}, \mathbf{X}, \theta)$ 的表达式。可以采用拉普拉斯近似法来得到 $p(\mathbf{f}|\mathbf{y}, \mathbf{X}, \theta)$ 的一个高斯逼近。拉普拉斯近似法的思想是用一个高斯分布 $q(\mathbf{f}|\mathbf{y}, \mathbf{X}, \theta)$ 来代替解析困难的后验概率 $p(\mathbf{f}|\mathbf{y}, \mathbf{X}, \theta)$ ：

$$p(\mathbf{f}|\mathbf{y}, \mathbf{X}, \theta) \approx q(\mathbf{f}|\mathbf{y}, \mathbf{X}, \theta) = N(\mathbf{f}|\mathbf{f}_{map}, \mathbf{A}^{-1}) \quad (4-30)$$

上式中 \mathbf{f}_{map} 为最大后验估计，关于矩阵 \mathbf{A} 的求解在诸多文章中都已提到^[60]，现直接给出其解析结果：

$$\mathbf{A} = \mathbf{W} + \mathbf{K}^{-1} \quad (4-31)$$

将上式代入式 (4-30) 中可得：

$$p(\mathbf{f}|\mathbf{y}, \mathbf{X}, \theta) \approx N(\mathbf{f}|\mathbf{f}_{map}, (\mathbf{W} + \mathbf{K}^{-1})^{-1}) \quad (4-32)$$

其中， $[\mathbf{W}]_{i,j} = -\sum_{k=1}^m \nabla \nabla_{i,j} \log p(y_k | \mathbf{f}_k, \theta_\ell)$ ，定义二阶偏导算子 $\nabla \nabla_{i,j} = \frac{\partial^2}{\partial f(\mathbf{x}_i) \partial f(\mathbf{x}_j)}$

4.2.2 超参数的求解

正如本章 4.1.2 小节所述的那样，通常求解超参数的方法是利用边缘似然函数来推导模型的参数，最终采用共轭梯度法获得边缘似然最大值来求解超参数。但是，共轭梯度法存在明显缺点，例如容易陷入局部最优解、对初始值依赖性强等。为了克服共轭梯度法学习超参数的缺点，很多学者利用已有的一些其他优化算法来学习超参数，如粒子群算法、人工蚁群算法以及遗传算法等^{[61][62][63]}。相比之下，遗传算法具有极强的自组性，对种群的初始值没有要求，在多目标隐性参数的优化求解上有着良好的应用。此外，遗传算法不需要利用一阶导数的信息，比共轭梯度法更加简单，也更容易操作。因此，本小节利用遗传算法来优化超参数。具体步骤如下：

- (1) 随机生成种群数量为 N 的超参数群体
- (2) 结合训练样本的输入，对应每个超参数染色体，通过预测模型可以得到其预测输出
- (3) 将各染色体的预测输出，通过适应度评价公式计算其适应度值
- (4) 判断迭代次数是否达到算法设定上限值，若已经达到，则返回当前适应度最高个体，该染色体对应的基因单元即为各超参数的学习结果；反之，执行操作 (5)
- (5) 根据计算得到的适应度值，执行相关遗传算法操作（选择、交叉、变异），形成新的种群，转入步骤 (2)
- (6) 重复步骤 (2) ~ (5)，直至到达算法设定进化代数的上限。最优染色体即为超参

数的学习结果。

染色体的编码方式采用 12 位二进制编码（最高位权值为 2^3 ，最低位权值为 2^{-8} ），采用固定概率交叉和变异操作，交叉概率为 0.9，变异概率为 0.1。选择操作采用轮盘赌的方式，如前一章节 3.2.1 小节所述。由于本小节利用遗传算法的目的是为了优化超参数，模型预测输出与训练样本所给输出越接近的染色体，其适应度应该越高。因此，本节考虑将均方误差用作适应度评价函数，具体见式（4-33）：

$$F(x_l) = \exp \left\{ -\alpha \times \sum_{i=1}^d (x_l - y_i)^2 \right\} \quad (4-33)$$

其中， $l \in [1, 20]$ ， d 为训练样本的数目。调整系数 α 使得 $F(x_l)$ 落在一个合理的概率区间范围内，取值 $\alpha = 0.2$ 。本节算法迭代次数上限为 50。综上，本小节利用遗传算法学习超参数的算法流程如图 4-3 所示：

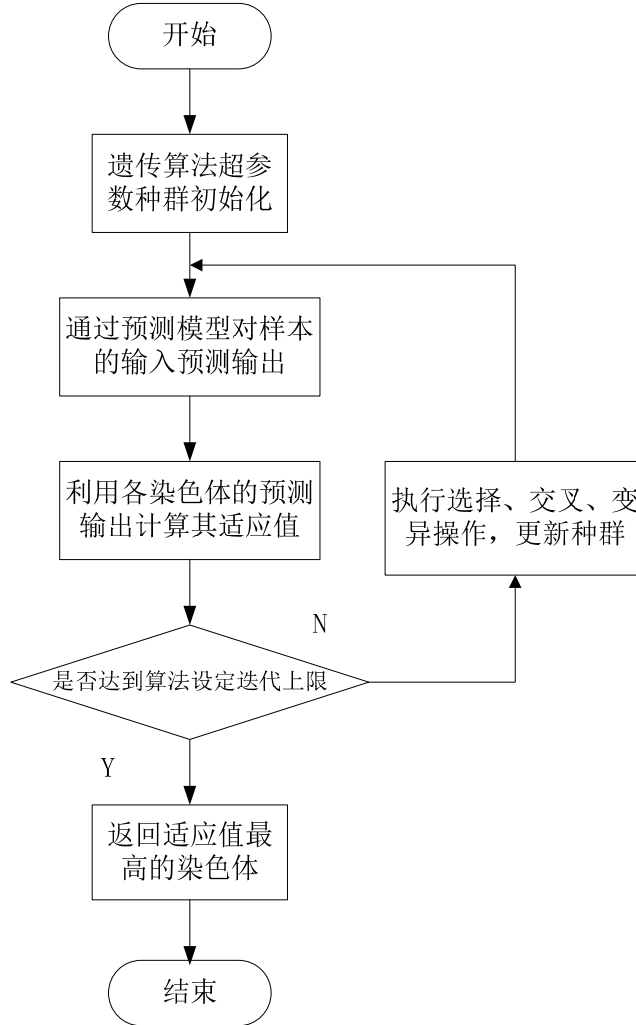


图 4-3 基于遗传算法的高斯过程回归超参数学习流程图

4.2.3 预测输出

通常而言，机器学习算法的目的是为了对将来的输入预测其输出。在本章所述高斯回归法中，待预测输入集合记为 $\mathbf{X}_* = \{\mathbf{x}_l^* \in R^d \mid l=1, \dots, o\}$ ，预测输出记为 $\mathbf{f}_* = [f(\mathbf{x}_1^*), \dots, f(\mathbf{x}_o^*)]^T$ 。在贝叶斯框架中，所有变量被认为是随机的。对 \mathbf{f}_* 的预测是建立在求解其后验分布 $p(\mathbf{f}_* | \mathbf{y}, \mathbf{X}_*)$ 的基础之上。一旦后验分布 $p(\mathbf{f}_* | \mathbf{y}, \mathbf{X}_*)$ 确定下来，相关的统计变量可以被估计出来，如均值估计 $\boldsymbol{\mu}_*$ ，以及方差估计 $\boldsymbol{\Sigma}_*$ 。

训练函数 \mathbf{f} 以及预测值 \mathbf{f}_* 的联合先验分布为：

$$\begin{bmatrix} \mathbf{f} \\ \mathbf{f}_* \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \mathbf{K} & \mathbf{K}_* \\ \mathbf{K}_*^T & \mathbf{K}_{**} \end{bmatrix} \right) \quad (4-34)$$

其中， $[\mathbf{K}_{**}]_{l,r} = k(\mathbf{x}_l^*, \mathbf{x}_r^*)$ ， $[\mathbf{K}_*]_{l,r} = k(\mathbf{x}_l, \mathbf{x}_r^*)$ 。显然，条件分布 $\mathbf{f}_* | \mathbf{f}$ 也服从高斯分布，因此：

$$p(\mathbf{f}_* | \mathbf{y}, \mathbf{X}, \mathbf{X}_*, \theta) = \int p(\mathbf{f}_* | \mathbf{f}, \mathbf{X}, \mathbf{X}_*, \theta) \times q(\mathbf{f}, \mathbf{y}, \mathbf{X}, \theta) d\mathbf{f} \quad (4-35)$$

$$= N(\mathbf{f}_* | \boldsymbol{\mu}_*, \boldsymbol{\Sigma}_*) \quad (4-36)$$

上式中 $p(\mathbf{f}_* | \mathbf{y}, \mathbf{X}, \mathbf{X}_*, \theta)$ 是两个高斯分布乘积的积分，因此， $p(\mathbf{f}_* | \mathbf{y}, \mathbf{X}, \mathbf{X}_*, \theta)$ 也服从高斯分布：

$$p(\mathbf{f}_* | \mathbf{y}, \mathbf{X}, \mathbf{X}_*, \theta) = N(\mathbf{f}_* | \boldsymbol{\mu}_*, \boldsymbol{\Sigma}_*) \quad (4-37)$$

参考相关文章^[64]，可以得到上式中均值和方差的推导结果：

$$\boldsymbol{\mu}_* = \mathbf{K}_*^T \mathbf{K}^{-1} \mathbf{f}_{map} \quad (4-38)$$

$$\boldsymbol{\Sigma}_* = \mathbf{K}_{**} - \mathbf{K}_*^T (\mathbf{K} + \mathbf{W}^{-1})^{-1} \mathbf{K}_* \quad (4-39)$$

$$= \mathbf{K}_{**} - \mathbf{K}_*^T (\mathbf{I} + \mathbf{W}\mathbf{K})^{-1} \mathbf{W}\mathbf{K}_* \quad (4-40)$$

式 (4-38) 是内部响应函数 PRF 的均值预测，式 (4-39) 得出的是 IRF 预测输出值的方差。 $\boldsymbol{\mu}_*$ 及 $\boldsymbol{\Sigma}_*$ 在下一小节介绍的主动学习过程中将被用到，可以用来形成第 $k+1$ 组比较对，以供听损患者进行下一轮比较估计。

4.3 序贯设计及本章算法流程

4.3.1 基于期望增益的序贯设计

在 MLHAF 中，采用主动学习的方法决定下一组由听损患者比较的两对参数，使得在尽

可能少的评估次数内完成全局优化。主动学习一般是用来减少训练模型所需要的样本数量，通常情况下，主动学习应用在对样本进行标签所需代价比较高的情形下。

对于大多数机器学习算法而言，旨在优化模型泛化能力的序贯设计通常都是基于一些特定的准则。对于包含了新的样本的后验香农熵而言，贝叶斯准则就是这样的优秀准则^[65]。对于当前的研究而言，模型的泛化能力并不是主要考虑的问题。如何找到一个使得未知 PRF 取得最大值的输入才是本节的研究重点。为此，本节采用期望增益（EI）的方法。EI 中的增益，在本章中指的是，相对于满足当前 IRF 输出最大的 $\hat{\mathbf{X}}_{\max}$ ，待测试点 \mathbf{X}_l^* 的提高值：

$$I = f(\mathbf{x}_l^*) - f(\hat{\mathbf{x}}_{\max}) \quad (4-41)$$

显然，变量 I 也服从正态分布：

$$I \sim N(I | \mu_I, \sigma_I^2) \quad (4-42)$$

在标准 EI 中， $f(\hat{\mathbf{X}}_{\max})$ 被视为非随机变量^[66]，变量 I 均值为 $\mu(I) = [\boldsymbol{\mu}_*]_l - \hat{\mathbf{f}}_{\max}$ ，方差为 $\sigma_I^2 = [\boldsymbol{\Sigma}_*]_{l,l}$ 。但在本文中，待测点 \mathbf{X}_l^* 和最优解 $\hat{\mathbf{X}}_{\max}$ 的联合分布被考虑在内，调整均值和方差为： $\mu(I) = [\boldsymbol{\mu}_*]_l - [\boldsymbol{\mu}_*]_{\max} = [\boldsymbol{\mu}_*]_l - f(\hat{\mathbf{x}}_{\max})$ ， $\sigma_I^2 = [\boldsymbol{\Sigma}_*]_{l,l} + [\boldsymbol{\Sigma}_*]_{\max,\max} - 2 \times [\boldsymbol{\Sigma}_*]_{l,\max}$ 。

最终的期望增益则是式（4-41）中增益 I 的期望值：

$$EI = E_{p(I)}\{\max(I, 0)\} = \int_0^\infty I p(I) dI \quad (4-43)$$

$$= \int_0^\infty I N(I | \mu_I, \sigma_I^2) dI \quad (4-44)$$

$$= \frac{\sigma_I}{\sqrt{2\pi}} \exp\left(-\frac{\mu_I^2}{2\sigma_I^2}\right) + \mu_I \Phi\left(\frac{\mu_I}{\sigma_I}\right) \quad (4-45)$$

上式中 $\Phi(\bullet)$ 表示标准正态分布的分布函数。由于听损患者的个人最佳参数配置是无法得知的，因此，本章提出的方法无法评估听损患者的真实最佳参数配置与 MLHAF 测出的最佳参数配置之间的差异。但是，我们可以用 EI 作为算法是否已经收敛的检验标准。当 EI 均值为 0 或者接近 0 的时候，我们认为新的测试样本中已经不会有超过当前最佳参数的点。

因此，当前的最佳参数配置 $\hat{\mathbf{X}}_{\max}$ 被认为是全局最佳参数配置。

4.3.2 本章算法流程

4.2.2 小节介绍了利用遗传算法进行超参数优化的方法，该方法对于超参数的初始化不敏感，全局优化能力强。超参数学习完成之后，PRF 模型相应确定下来，利用上一小节提出的序贯设计选取两组助听器参数，即可形成两组新的语音信号让受试者进行评估，从而形成

一个验配循环流程。对于主动学习部分的式(4-45)，当参数维数 d 比较小，如 $d=2$ 时，可以利用计算机将所有2维的点全部计算一遍，然后求得EI的最大值；当参数维数 d 较大时，可以采用BFGS梯度下降法或遗传算法求取EI的最大值。EI产生最大值所对应的参数配置与当前最优参数配置就是听损患者需要比较的下一组参数。综上，MLHAF循环执行以下三个步骤：1) 利用当前样本得出PRF分布，通过主动学习确定接下来需要比较的两组参数。2) 听损患者对1)中给出的两组参数进行比较评估。3) 结合所有评估过的样本，对PRF的分布重新估计。当算法收敛或者迭代次数到达设定上限时，使得该PRF分布输出最大值的参数设置即为最佳的助听器参数配置。MLHAF算法流程图如图4-4所示：

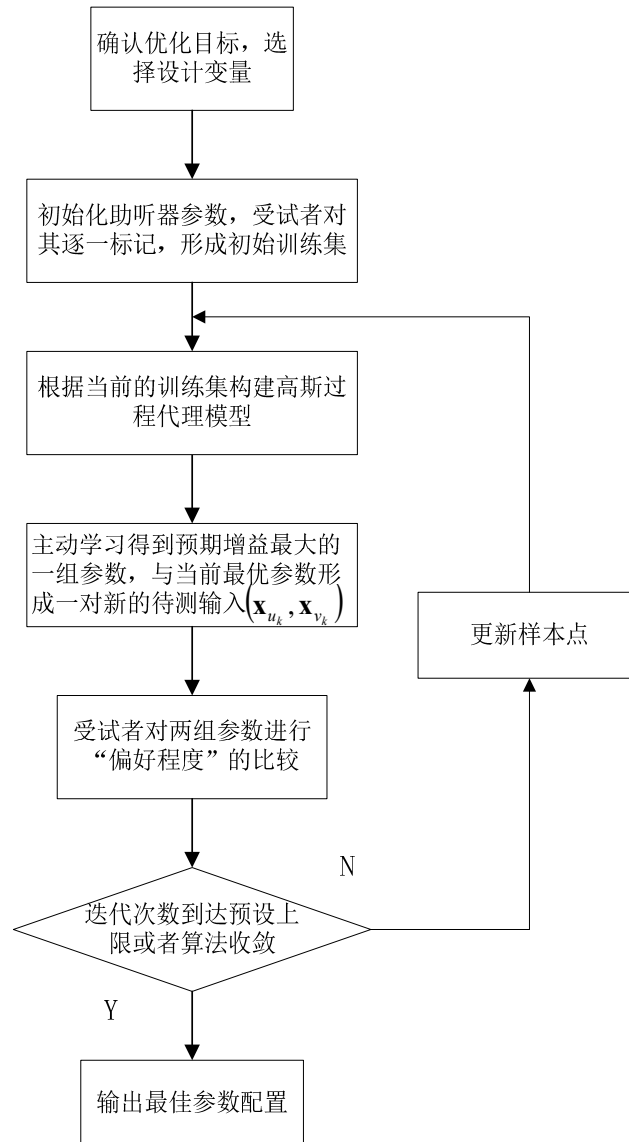


图 4-4 基于高斯过程回归的自适应验配算法流程图

4.4 主观测听评估

本节采用的主观测听评估方法与3.4.3小节中采用的方法类似，实验的进行同样分为两个阶段。第一阶段：用本章所述的基于高斯过程回归的自适应验配方法进行验配，得到“最

佳的”助听器相关参数；第二阶段：选定五位实验对象（无听损）以及男女各 10 段语音，对于每一位实验对象，随机让其听取自适应验配处理过后的语音以及未经处理的（含听损）语音，比较二者的语音质量优劣，并以语音质量的“提升”和“下降”进行评估。

在受试者进行验配的过程中，算法相关参数及评价结果同样会被保留到相应文件中（如表 4-1 所示），以方便后续分析和研究。表 4-1 显示的是测试者 A 在与验配软件交互过程中产生的中间结果：

表 4-1 算法每代参数记录

语音文件名称	算法代数	参数 A 数值	参数 B 数值	评价结果
Track 17_1.wav	第 1 代	36,50	51,62	A 优于 B0.7
Track 17_1.wav	第 2 代	36,50	48,52	A 优于 B0.4
Track 17_1.wav	第 3 代	68,45	36,50	B 优于 A0.7
Track 17_1.wav	第 4 代	36,50	37,42	B 优于 A0.3
Track 17_1.wav	第 5 代	57,46	37,42	B 优于 A0.6
Track 17_1.wav	第 6 代	37,42	49,61	A 优于 B0.8
Track 17_1.wav	第 7 代	43,50	37,42	B 优于 A0.3
Track 17_1.wav	第 8 代	37,42	28,48	A 优于 B0.1
Track 17_1.wav	第 9 代	37,42	43,27	A 优于 B0.5
Track 17_1.wav	第 10 代	37,42	23,38	B 优于 A0.3
Track 17_1.wav	第 11 代	23,38	42,18	A 优于 B0.4
Track 17_1.wav	第 12 代	23,38	21,18	B 优于 A0.4
Track 17_1.wav	第 13 代	37,28	21,18	B 优于 A0.6
Track 17_1.wav	第 14 代	21,18	25,40	A 优于 B0.4
Track 17_1.wav	第 15 代	25,30	21,18	A 优于 B0.2
Track 17_1.wav	第 16 代	16,28	25,30	B 优于 A0.3
Track 17_1.wav	第 17 代	25,30	32,17	A 优于 B0.5
Track 17_1.wav	第 18 代	25,30	20,24	B 优于 A0.6
Track 17_1.wav	第 19 代	20,24	18,31	A 优于 B0.4
Track 17_1.wav	第 20 代	20,24	31,10	A 优于 B0.7

对本章算法进行实验测试时选择的验配参数个数为 2，第一个参数值对应 5.3 小节滤波器组前八个通道的负增益值，第二个参数值对应后八个通道的负增益值。对测试者 A 所设定的听损值为 20dB、25dB，从表 4-1 中可以看出，验配进行到第 18 代的时候，已经获得比较满意的验配结果：20dB,24dB。五位测试者的验配结果如表 4-2 所示：

表 4-2 主观测听结果

受试者编号	预设听损（dB）	验配结果（dB）	验配时间（分钟）
A	20,25,	20,24	20
B	52,47	49,47	19
C	18,35	19,38	23
D	15,63	21,34（未收敛）	29
E	27,28	27,30	25

从表 4-2 中可以看出，编号为 D 的受测试人员没有得到最佳验配参数，其对应的 EI 值不趋近于 0。其他四位测试者得到了算法收敛的验配参数，验配时长在 20 分钟左右。我们可以看看两个典型的验配结果，如图 4-5 所示：

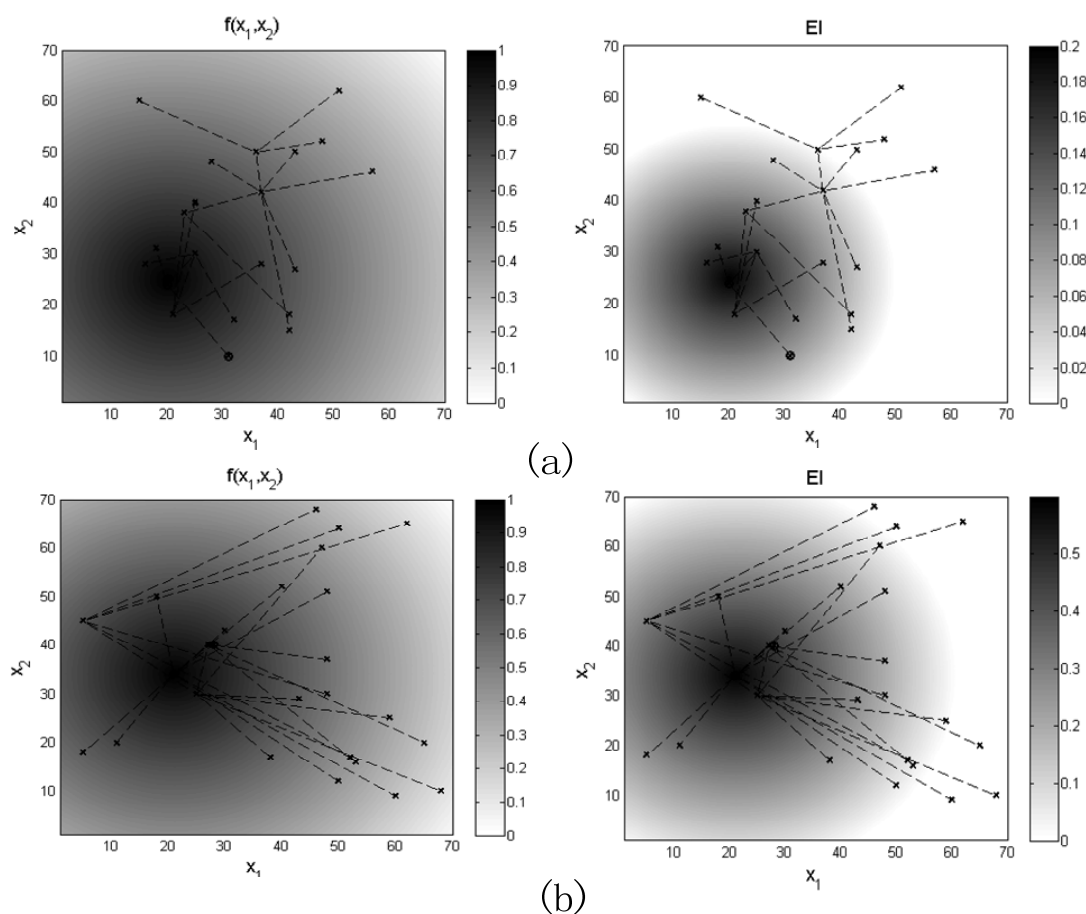


图 4-5 两个验配过程比较图（包含模型预测输出 $f(x_1, x_2)$ 与 EI 值）

图 4-5 中的 (a) 和 (b) 分别对应测试者 A 与 D 的验配过程，图中分别展示了两人在预测过程中 PRF 的预测输出 $f(x_1, x_2)$ 以及 EI 值。从表 4-2 的主观测听结果我们已经知道，测试者 A 整个验配过程顺利，最后获得了满意的验配值；而测试者 D 最后算法没有收敛，系统给出的验配结果也不是令人满意。这与图 4-5 中给出的验配过程是一致的，从图 4-5 中可以看出，测试者 A 在经过 20 代算法比较后，已经得到了比较满意的验配结果，其 EI 均值接近于零，算法收敛；而测试者 D 在经过 25 代（预设上限值）算法比较后，仍然没有得到理想的验配结果，其 EI 均值明显不趋近于 0。在验配过程完成后，从测试者 D 的描述中我们得知，测试者 D 并不能稳定他的评价标准，常常判断不出两个语音的好坏，而其他验配成功的人则正好相反。因此，维持稳定的评价标准是该算法运行成功的一个重要前提。

在实验第二阶段，测试者会听到相同语音源的两种语音，一种是经过本章自适应验配算法处理后的语音，另一种是未经处理的语音（只含听损）。两种语音随机播放以消除测试者的先验心理。每次听完一段对话的两种语音后，测试者被要求选定二者中相对语音质量好的

一种，也可选择两者质量“相差不大”。在所有五位测试者听完所有 20 段对话后，实验结束并统计实验结果。实验结果如表 4-3 所示。

表 4-3 实验阶段二测试结果统计

测试者编号	语音类型（男/女）	验配效果优于未经处理效果	验配效果劣于未经处理效果
A	男	9/10	0/10
	女	10/10	0/10
B	男	7/10	0/10
	女	8/10	0/10
C	男	9/10	0/10
	女	9/10	0/10
D	男	4/10	2/10
	女	6/10	3/10
E	男	10/10	0/10
	女	7/10	0/10
统计结果		79/100	5/100

测试结果统计数据如下：对于所有 100 段语音（5*10*2），有 79 段语音被认为效果优于未经验配的语音，5 段被认为效果劣于未经验配的语音质量，其余 16 段语音被认为两者相差不大。从表 4-3 中可以看到，只有测试者 D 对于部分语音认为效果比验配前差，其他测试者认为经过验配后绝大部分语音效果优于验配前。这与前面的分析也是一致的，测试者 D 并没有对系统给出的一些语音进行合理的评价致使最后算法没有收敛，系统给出的验配参数与最优参数差距较大。对于其他测试者而言，验配结果都比较理想。综上所述，本章提出的基于高斯过程回归的自适应验配算法可以有效的搜索数字助听器验配参数。此外，对于进行验配的人员也提出了一定要求，要有合理判别不同语音信号质量的能力，且整个验配过程中要保持评价标准的一致性。

4.5 本章小结

本章在第一小节首先介绍了高斯过程回归用于预测的理论基础，并用一个简单的例子演示了高斯过程回归预测的整个过程；第二小节进行了基于高斯过程回归的自适应验配算法相关理论推导，给出了用于预测验配者对当前参数喜好程度的公式。此外，该小节还利用遗传算法对超参数进行优化并给出了具体操作步骤，避免了利用共轭梯度法优化超参数的缺点；第三小节介绍了基于期望提高值（EI）的序贯设计，以找到当前“最有意义”的验配参数提供给验配者进行标记，并给出了基于高斯过程回归的自适应验配的完整算法流程；第四小节进行了主观测听评估。测试的结果表明，在验配者保持评价标准一致性的前提下，本章提出的自适应验配方法能有效搜索到适用于验配者的最优验配参数。

第五章 自适应验配软件系统的设计及实验测试

5.1 自适应验配总体架构

自适应验配总体架构如图 5-1 所示，共有三个部分组成，分别为优化模块、转换模块、用户接口。用户接口基本功能包含参数设置、信息显示和声音满意度评价，用户可以通过用户接口反馈声音质量的好坏，也可以通过用户接口上的相应菜单选项确定算法中所用参数。转换模块的作用是将优化模块搜索得到的参数结合原始语音信号，形成一组新的语音，以待用户听取并反馈，该模块应包含相应的数字信号处理功能。本文的转化模块采用的是分解滤波器组和综合滤波器组，优化模块的参数以不同频段的增益给出，转化模块则根据优化模块得到的增益形成新的语音信号。优化模块则是自适应验配的核心模块，该模块的任务是执行具体的优化算法，包含第三章与第四章提出的两种优化方法。此外，对于第三章提出的验配算法而言，还增加了优化辅助模块，该模块的作用是分析用户的偏好，计算种群的多样性等，以此形成相应的“知识”反馈给优化模块，以达到降低用户疲劳度的目的。

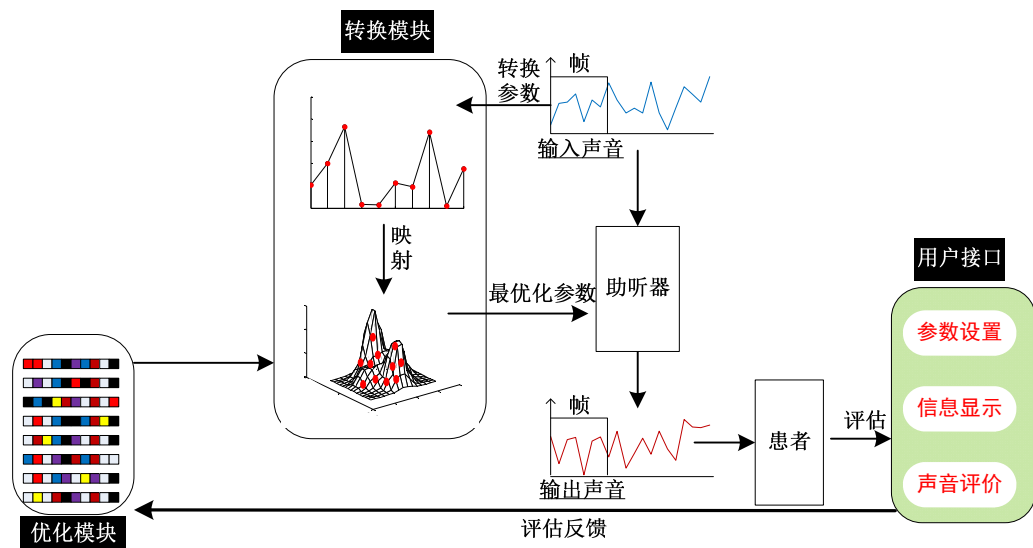


图 5-1 自适应验配总体架构图

5.2 人机交互模块设计

人机交互模块作为直接和用户进行交互的部分，主要分为两个部分：纯音测试与语音测试。纯音测试模块的功能是测试用户在 125Hz 到 8000Hz 共 11 个频点的听力情况，从而确定用户的听阈参数，而语音测试部分则是与用户交互的主体。用户首先进行纯音测试，然后进行语音测试，纯音测试的结果将作为语音测试的初始化参考依据。VS2010 是微软于 2010 年推出的集成开发环境，用于构建和部署功能强大而安全的联机.NET 软件。由于其操作的方便性以及跨平台设计特性，本文在 VS2010 平台上设计了人机交互模块。算法的研究与实

现在 MATLAB2012a 上进行，然后采用动态链接的方式，将在 MATLAB 上的研究成果结合到人机交互模块上面。

5.2.1 纯音测试

纯音测试的界面如图 5-2 所示，通过下拉选项选择测试的频点后，系统会发出相应频率的声音，若测试者听不到该频点声音，则可以通过点击声强按钮提高声音的强度，直至能听到对应频点的声音为止，系统播放声音的强度会在界面正中间的上方实时显示。点击“下一频段”，系统则会发出下一个待测试特征频点的声响，依次完成 11 个频点的测试后，系统就会得到测试者的 11 个频点听阈参数，并会在右下角的“参数”栏中显示出来。完成全部纯音测试后，点击“语音测试”按钮，即可进入人机交互的核心环节，即语音测试。若测试者中途因故要放弃测试，可以点击“退出”结束测试。



图 5-2 纯音测试界面

5.2.2 语音测试

用户可以不断的进行语音测试，凭主观感觉判断当前播放语音是否清楚，然后进行语音质量评定，直到最终获得最优参数为止。获得的最终参数即为最适合用户的听力验配参数。由于第三章和第四章采用了两种不同的优化算法，相应的交互界面也略有不同。图 5-3 所示的是基于交互式遗传算法的语音测试界面：

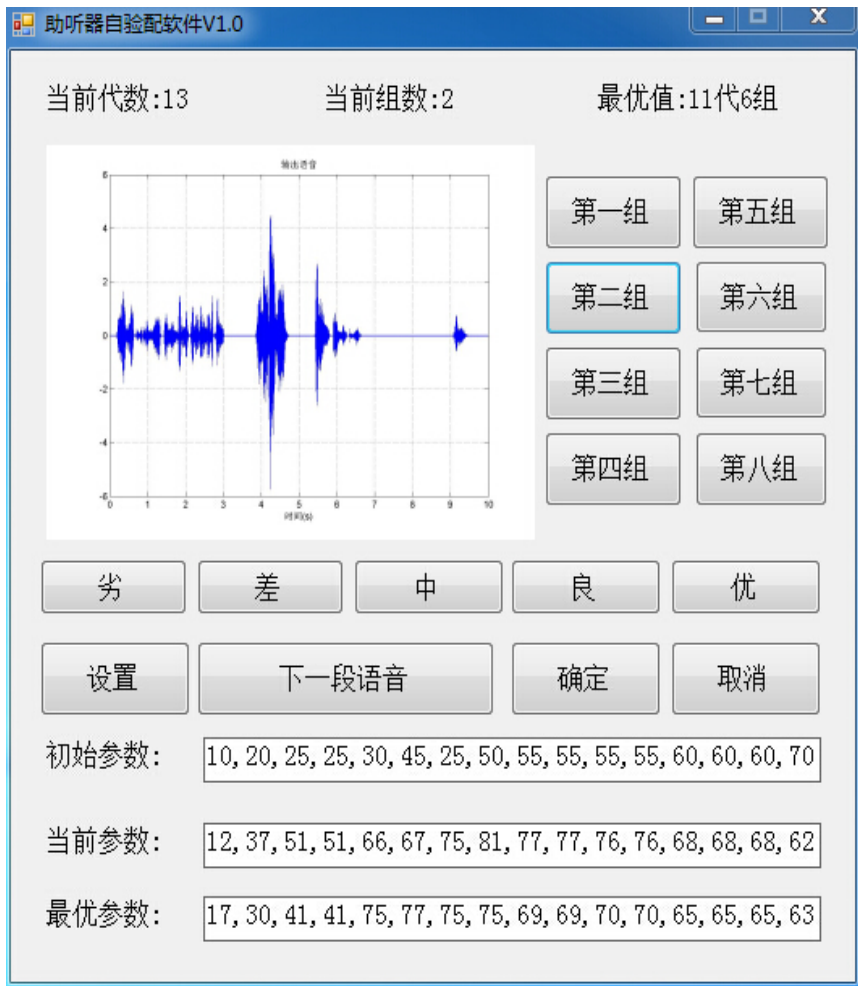


图 5-3 基于交互式遗传算法的语音测试界面

初始化完成后，点击“配置”按钮，即可进入“配置”界面。在配置界面，可以完成语音路径的更改，以及对“交叉概率”和“变异概率”参数的配置。配置完成后即可进行语音测试。点击“第一组”按钮后，系统会播放相应语音，当前播放语音的频谱图会在界面中更新显示。用户听完该段语音后，根据语音是否容易理解进行相应的评价，评价标准有“劣”、“差”、“中”、“良”、“优”五等（语音质量由差到好）。语音测试界面下方会实时显示“初始参数”、“当前参数”以及“最优参数”。依次完成 5 组测试后，点击“确定”按钮完成本组测试。点击“下一段语音”系统会将语音播放内容换为下一段测试语音。用户一直进行语音测试，直到获得最优参数为止。点击“取消”按钮，退出系统。

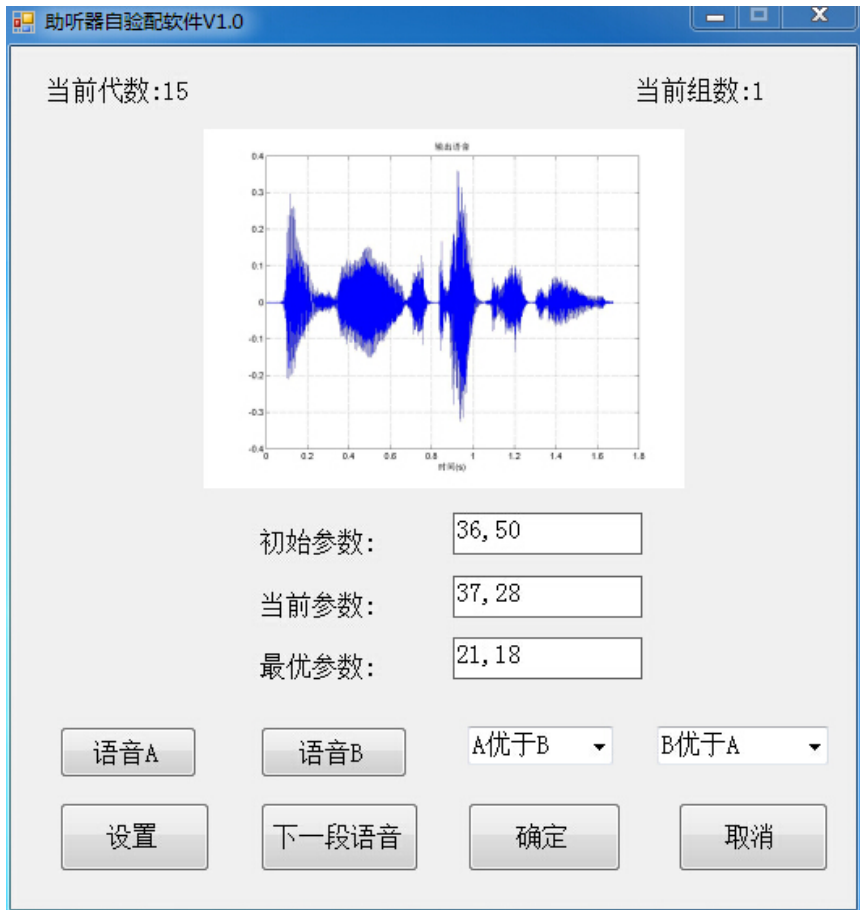


图 5-4 基于高斯过程回归的语音测试界面

基于高斯过程回归的语音测试界面如图 5-4 所示，从图中可以看出，算法每个轮回都会给出两组语音供测试者进行判别，对应于图中的“语音 A”和“语音 B”。测试者需听完这两组语音再作出评价，不仅要评价出两组语音的优劣，还要更具体的指出一种语音优于另外一个语音的程度。若测试者觉得语音 A 的效果比语音 B 好，则在“A 优于 B”对应的下拉菜单中选择 A 优于 B 的程度，下拉菜单中有 0 到 0.5 六个等级，间隔为 0.1，数值越小说明语音 A 优于 B 的程度越高（0.5 表示两者相当）。类似的，若测试者认为语音 B 优于语音 A，则在下拉菜单“B 优于 A”中进行相应程度的选择，下拉菜单中有 0.6 到 1.0 五个等级，间隔为 0.1，数值越大表明语音 B 优于 A 的程度越高。与交互式遗传算法的语音测试界面一样，语音频谱、初始参数、当前参数以及最佳参数也会在测试界面中实时显示。

5.3 语音信号处理模块

5.3.1 滤波器组的设计

语音信号处理模块的功能是将语音信号按频率划分为 16 个通道，便于对不同的频段进行不同的处理，最后将处理过后的各个单通道语音整合成为完整的语音信号反馈给测试者。由此可见，语音信号处理模块的主要是对分解滤波器组与综合滤波器组的设计，其常见模型如图 5-5 所示。

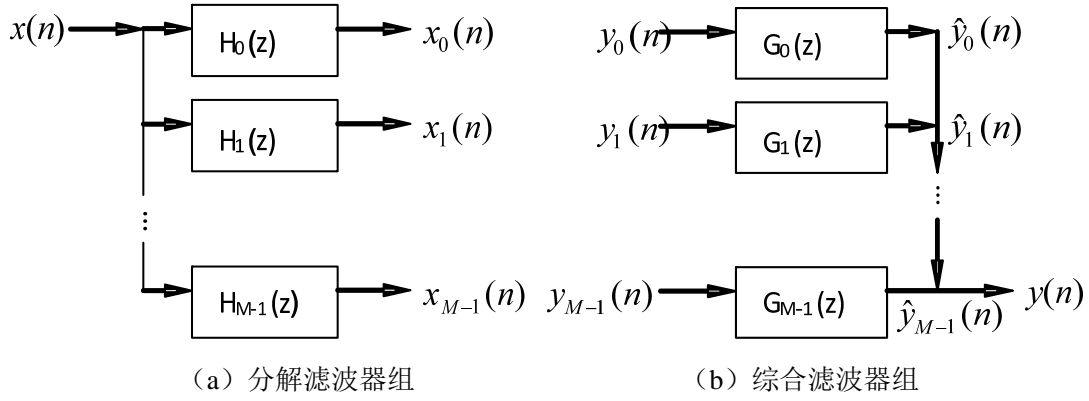


图 5-3 滤波器组模型

FIR 滤波器虽然能实现严格的线性相位，但是其阶数远远高于 **IIR** 滤波器。因此 **IIR** 滤波器在计算速度以及存储单位上相比于 **FIR** 滤波器都具有明显的优势。此外，由于人耳的特殊结构使得人对于语音信号的相位误差并不敏感。基于上述原因，本文的滤波器组选择采用 **IIR** 滤波器。对于分解滤波器组，本文采取双线性变换法进行设计。滤波器设计过程如下：

首先需要选择模拟低通滤波器的原型。本文要设计一个 6 阶 **IIR** 滤波器，对应的原型模拟滤波器则需要 3 阶。由于椭圆滤波器在通带和阻带内都是等波纹的，因此本文选择椭圆滤波器作为原型滤波器。查表可知，3 阶归一化椭圆低通滤波器（通带最大衰减为 0.5dB，阻带最小衰减不小于 20dB 的归一化系统函数）系统函数如下：

$$H_a^1(s) = \frac{0.3642s^2 + 0.9094}{s^3 + 1.2197s^2 + 1.487s + 0.9094} \quad (5-1)$$

得到了系统函数 $H_a^1(s)$ 后，将关系式 $s = \frac{s}{\Omega_c}$ 代入 $H_a^1(s)$ 后，即可得到反归一化系统函数：

数函数：

$$H_a(s) = H_a^1\left(\frac{s}{\Omega_c}\right) \quad (5-2)$$

其中， Ω_c 指的是模拟低通边界频率，其计算方式如下：

$$\Omega_c = \frac{\cos \omega_0 - \cos \omega_1}{\sin \omega_1} \quad (5-3)$$

式 (5-3) 中的 ω_1 表示所设计 **IIR** 带通滤波器的上边界角频率， ω_0 表示所设计 **IIR** 带通滤波器的中心频率。

由于各子滤波器均为带通滤波器，故从 s 域映射到 z 域的双线性变换表达式为：

$$s = \frac{(z - e^{j\omega_0})(z - e^{-j\omega_0})}{(z - 1)(z + 1)} = \frac{z^2 - 2z \cos \omega_0 + 1}{z^2 - 1} \quad (5-4)$$

将上式代入 $H_a(s)$ 后即可得到所设计的 IIR 滤波器的系统函数 $H(z)$ 。综上所述：本文设计的带通滤波器组将按照如下步骤生成：

- (1) 确定各带通 IIR 滤波器上下限频率 f_h 和 f_l （具体通带划分标准将在下一小节给出），采样频率 $f_s=16\text{kHz}$ ；
- (2) 按照公式 $\omega = \frac{2\pi f}{f_s}$ 计算 IIR 滤波器上下限角频率 ω_h 和 ω_l ，并计算其中心频率 ω_0 ；
- (3) 按照式 (5-3) 计算模拟低通边界频率 Ω_c ；
- (4) 将双线性变换式 (5-4) 代入 $H_a(s)$ ，即可生成本文所需的 6 阶 IIR 带通滤波器的系统函数 $H(z)$ 。

由于分解滤波器将语音信号分解到互不重叠的频带上，且相位是否线性不影响数字助听器输出语音的效果，本文的综合滤波器组选择与分解滤波器组同样的形式，即：

$$G_i(z) = H_i(z), i = 1, \dots, 16 \quad \text{式 (5-5)}$$

5.3.2 测试与仿真

本文利用各滤波器组的增益值作为验配参数，因此，通道的划分显得尤为重要，每个通道增益值相应的改变都应在受测试者的耳中产生明显影响。在进行纯音测试的时候，往往只会测试听损患者几个关键频点的听力情况，然后通过插值最终获得患者的完整听力图。受到纯音测试的启发，在设计分解滤波器的时候，应当尽量保证频率特征点落在各个带通滤波器的中心频率附近。常见人耳频率特征点有：

$[125\text{Hz}, 250\text{Hz}, 375\text{Hz}, 500\text{Hz}, 750\text{Hz}, 1\text{kHz}, 1.5\text{kHz}, 2\text{kHz}, 3\text{kHz}, 4\text{kHz}, 6\text{kHz}]$ 。此外，由于人耳在低频处比较敏感，因此在设计分解滤波器组的时候，要注意低频的通道数目适当增加，高频的通道数目适当减少。兼顾以上原则，本文最终设计的滤波器各通道情况如

表 5-1 所示。

表 5-1 通道划分结果表

通道	下限 f_l/Hz	上限 f_h/Hz	带宽/Hz	中心频率 f_c/Hz
1	1	90	89	46
2	90	187	97	139
3	187	312	125	250
4	312	437	125	375
5	437	625	188	531
6	625	875	250	750
7	875	1125	250	1000
8	1125	1375	250	1250
9	1375	1625	250	1500
10	1625	1875	250	1750
11	1875	2250	375	2063
12	2250	2750	500	2500
13	2750	3200	450	2975
14	3200	3750	550	3475
15	3750	5500	1750	4625
16	5500	7999	2499	6750

结合 5.3.1 小节的算法步骤以及表 5-1 的通带划分标准，滤波器组的设计工作已经全部完成，为检验所设计的 IIR 滤波器性能，本小节将对其进行仿真与测试。图 5-6 是本文设计的 16 通道 6 阶 IIR 滤波器组的频率响应曲线图。从图中可以看出，在 3dB 带宽处完全符合表 5-1 中提出的设计要求，且通带波动小于 0.5dB，滤波后的语音不会产生明显失真；阻带衰减均达到了 20dB，过渡带衰减快，没有出现相邻通道发生混叠的情况。

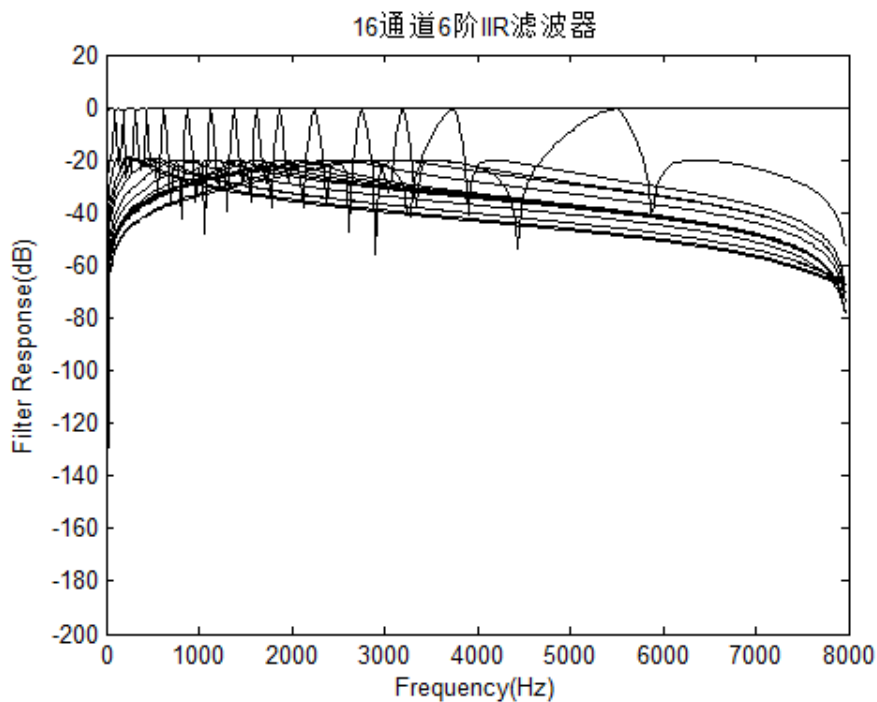


图 5-6 本文设计的 16 通道滤波器组频响图

接下来用语音信号对滤波器组进行进一步测试，测试语音信号是 TIMIT 语音库 dr2-fram1 类中的 si730.wav 文件，采样率为 16kHz，采样精度为 16bit，语音文件内容为 “Another memo for sightseers: bring your camera along to museums.”。测试方法为：将原始语音信号经过分解滤波器，然后将得到的信号直接送入综合滤波器进行处理，得到最终的输出语音信号，比较原始信号与输出语音信号的波形图与语谱图，以此评判本文设计的分解与综合滤波器的效果。

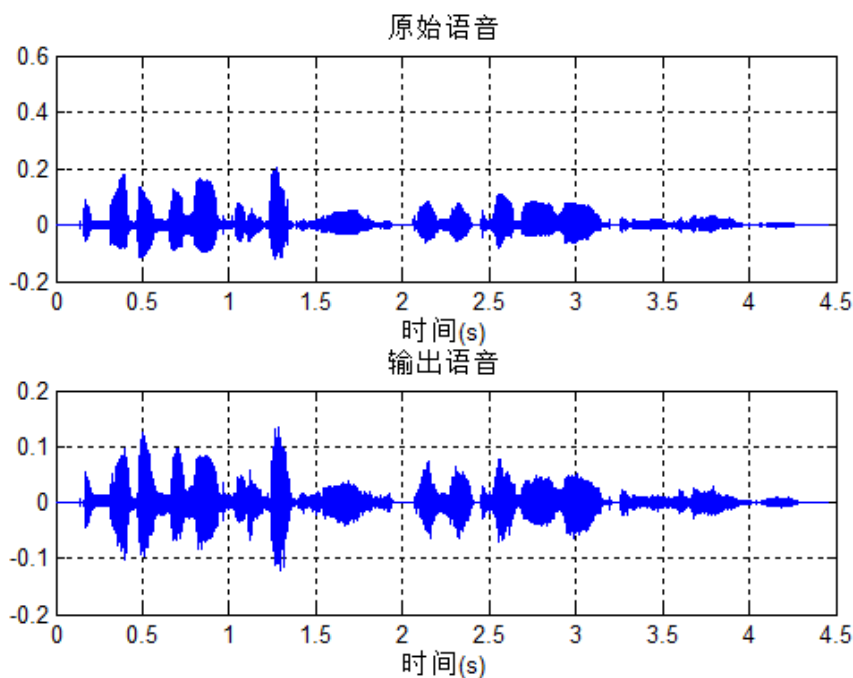


图 5-7 原始语音信号与本文设计的滤波器组输出波形比较图

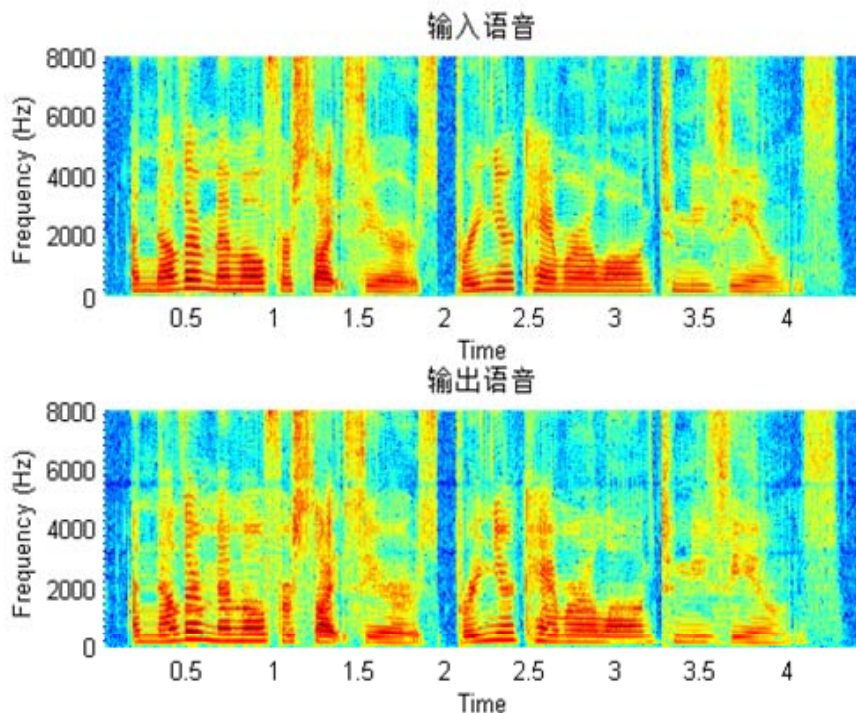


图 5-8 原始语音信号与本文设计的滤波器组输出语谱比较图

图图 5-上方的子图是原始语音信号的波形图，下方是语音信号经过本文设计的分解与综合滤波器组处理后的输出波形图；图 5-8 上方的子图是原始语音信号的语谱图，下方是经过本文设计的分解与综合滤波器组处理后的输出语谱图。通过对比图 5-7 可以发现，经过本文设计的滤波器组处理后的语音信号包络与原始信号基本一致，没有出现较大的失真，声音保真度较好；比较图 5-8 中的两个信号可知，处理后的语音信号在各个频率分量都完好的保留了原始语音信号，两者语谱图基本没有差异。将 TIMIT 语音库中的其他测试信号按照上面的方法进行测试后发现结果是类似的。通过上述仿真与测试可以证明，本文设计的 16 通道 6 阶 IIR 滤波器组性能良好，可以作为自适应验配中的语音信号处理模块。

5.4 言语测听实验

本小节的自验配言语测听实验包含两个部分：①比较未经验配和分别经过本文提出的两种自验配方法验配过后的语音可懂度的差异②比较未经验配和分别经过本文提出的两种自验配系统验配过后语音质量的差别。上述的“可懂度”指的是不同声压级下的言语识别率；而“语音质量”则为受测试者对当前播放语音信号的满意程度。

5.4.1 可懂度测听实验

十位志愿者（不含听损）参与了本次实验。《清晰度测试词表》以及《汉语测试词表》是临床上常用于测试听损患者言语识别率的词表，本小节采用的测听材料（见表 5-2）均摘自于这两个词表。

表 5-2 言语测试词表

图书	活动	打算	黑色	风筝
应用	游戏	治疗	参考	收割
兄弟	司机	首都	工人	家乡
围巾	教师	农业	树枝	灰尘

词表 5-2 所对应的文本由男声和女声各朗读一遍，语音信号的声压级调整为为 55 dB、65 dB、75 dB，分别代表日常生活中讲话的小声、中声、大声。在上述测试信号中加入预设听损形成最终的测试信号。测试在消音室进行，测试语音采取随机播放的方式，通过头戴式耳机森海塞尔 HD650 单耳给出。被测试者随机听取上述含有预设听损的 120 个测试语音信号，每条测试语音播放一遍，间隔 5 秒，然后由受测试者进行复述，统计测试者能答对的个数。最后将 10 位测试者答对的词语个数综合起来，并做归一化处理，答对词语个数所占比例即为“可懂度”。完成上述实验后，让各测试者先经过第三章提出的自适应验配方法进行验配，再重复上述步骤，即可获得经过基于交互式遗传算法的自适应验配后的语音可懂度。基于高斯过程的自适应验配方法可懂度测试方法同上。

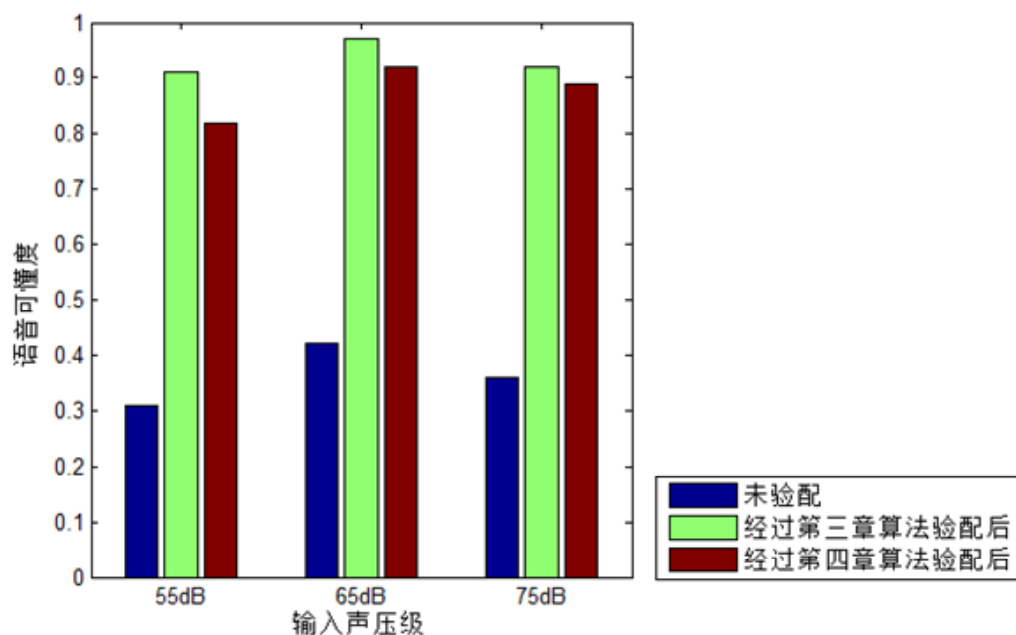


图 5-9 语音可懂度对比图

测试结果如图 5-9 所示，从上图可以看出，在未验配的情况下，测试者在各个声压级大概能够复述正确百分之三十到四十的词语，这是由于人为加入了听力损失（即改变各通道的增益值从而影响听力感官）的原因。经过第三章或者第四章算法进行验配后，测试者对于测试内容的识别率明显提高，基本在百分之九十左右。而第三章的算法效果比第四章效果略好，这是由于第三章的变异算子更具有针对性。但是测试过程中有一个细节值得注意，采用第三章的算法进行验配平均用时 40 分钟左右，而采用第四章算法只需 20 分钟左右即可完成。也就是说，第四章算法在验配效率方面要明显优于第三章算法。

5.4.2 语音质量测听实验

这一小节需要完成的是“语音质量的差异比较”的实验。首先挑选日常生活中经常出现的 5 段短句作为原始语音信号（测试文本内容见表 5-3），表 5-3 所对应的文本由男生和女生各朗读一遍，经过声压级处理后（55 dB、65 dB、75 dB）形成 30 个测试语音信号。在上述测试信号中加入预设听损形成最终的测试信号。测试环境与上一小节相同，在消音室中进行，测试信号通过头戴式耳机森海塞尔 HD650 单耳给出，测试语音随机播放。然后，被测试者随机听取含有预设听损的 30 个测试语音信号，每条测试语音播放三遍，间隔 15 秒，然后由受测试者进行打分，10 分为优，1 分为差。最后将 5 位测试者的评分综合起来，并做归一化处理，即得到最终的“语音信号的质量”。完成上述实验后，让各测试者先经过第三章提出的自适应验配方法进行验配，再重复上述步骤，即可获得经过基于交互式遗传算法的自适应验配后的语音质量。基于高斯过程回归的自适应验配方法验配后的语音质量测试方法同上。

表 5-3 语音质量实验测试文本

序列	文本
1	最近几周天气非常不错
2	他昨天不小心伤了脚踝
3	这个城市旅游业很发达
4	爸爸今天要加班到很晚
5	这身衣服真是好看极了

由于该小节提出的实验是基于测试者打分的主观实验，为证明实验的可行性，在测试进行之前，需要先测试同一评估者对于相同语音信号两次打分的相关性。该步骤为预实验，10 位受测试者随机听取 5 段测试语音（包括未经验配的含听损语音以及经过验配补偿后的语音信号），每段语音播放两遍，然后由测试者进行打分。预实验测试结果如下：

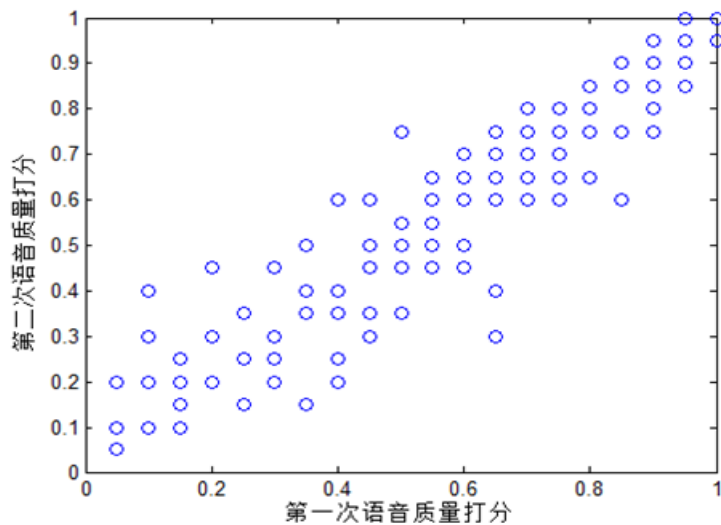


图 5-10 语音质量打分散点图

从图 5-10 的测试结果可以初步看出，两次打分的结果应该是强相关的，计算其 pearson 相关系数为 0.9123，相关强度属于极强相关。由此可知，由测试者对语音质量进行打分是可靠的，并且是可重复的。接下来进入正式测试阶段，测试结果如图 5-11 所示：

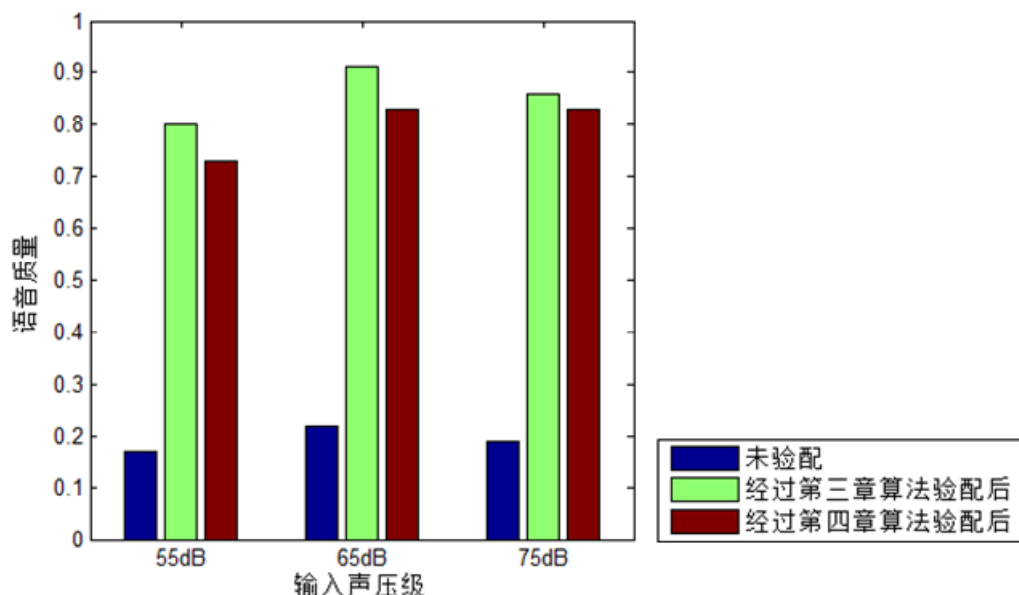


图 5-11 语音质量对比图

从图 5-11 可以看出，语音质量打分实验与上一小节的实验结果基本一致。经过本文的两种自适应验配算法进行验配后，测试者对于语音信号的质量都觉得有明显的提升，并且基于交互式遗传算法的自适应验配算法略优于基于高斯过程回归的自适应验配算法。此外，通过对比图 5-9 与图 5-11 可以看出，测试者对于语音质量的满意度是要低于语音可懂度的，这是符合常理的，测试者能够理解语音信号的内容，但不一定对该语音质量满意度很高。

5.5 本章小结

本章从一般自适应验配总体架构出发，首先在第一小节介绍了数字助听器自适应验配系统所需要的各个功能模块，明确了各模块的作用和意义；第二小节展示了本文设计自适应验配系统所采用的人机交互模块，包括纯音测试以及语音测试两个部分，其中采用不同验配算法所对应的语音模块稍有不同；第三节阐述了信号处理模块所用滤波器组的设计过程并进行了相应的仿真与测试，结果表明本文设计的滤波器组完全可以胜任自适应验配系统中的语音信号处理功能；最后一小节中，进行了三种语音信号（未验配音音以及经过本文两种验配方法验配后的语音信号）的可懂度以及语音质量的对比实验。实验结果表明，本文提出的两种验配方法对语音的识别率以及语音质量都有着显著的提升，在效果上，基于交互式遗传算法的验配算法略优于基于高斯过程回归的验配算法，而在验配效率上，后者要优于前者。

第六章 总结与展望

6.1 本文已取得的研究成果

佩戴助听器已经成为当下治疗听力损失的一种重要手段，同时也是其中最高效、便捷的方法。然而，助听器的验配流程过于繁琐，并且高度依赖验配师的经验水平，这对于助听器的推广是极其不利的，因此助听器的自适应验配成为了目前国际上研究的一个热点。但是，就国内而言，还没有看到相关研究成果出现。为弥补国内相关领域的空缺，本文重点研究了如何利用现有智能优化算法实现数字助听器的自适应验配，并设计了友好的人机交互界面。本文的工作内容具体如下：

(1) 介绍了人耳的听觉系统，包括其生理构造以及听觉形成的机理。然后阐述了听力损失的概念，并介绍了国际和国内对于听力损失划分的标准。总结了听力测试方法，罗列了几种治疗听力损失的常规方法。对比了传统验配过程和自适应验配过程两种验配方式，阐明了助听器自适应验配的原理。

(2) 研究了基于交互式遗传算法的自适应验配技术。交互式遗传算法的研究和应用已经十分广泛，这是一种结合人的主观评价的优化算法，其特点与本文的研究方向十分吻合，因此本文首先尝试将该算法应用到助听器的自验配中来。文章详细介绍了各遗传算子的设计，包括染色体编码、选择操作、交叉操作以及针对听损类型的变异操作。此外，还提出了优化辅助模块以提升算法搜索效率。仿真和实验结果证明了交互式遗传算法应用到助听器验配中的可行性。

(3) 研究了基于高斯过程回归的自适应验配技术。高斯过程回归是基于统计学发展起来的一种机器学习方法，在决策和分类领域也有着广泛的应用。在研究了高斯过程回归相关理论基础之上，本文进行了一系列理论推导，给出了用于预测验配者对当前参数喜好程度的公式。在模型的训练方面，采用遗传算法对超参数进行优化并给出了具体操作步骤。同时，引入了基于期望提高值(EI)的序贯设计，以找到当前“最有意义”的验配参数提供给验配者进行标记。主观测听评估的实验结果证明了该算法对于助听器验配的有效性。

(4) 针对本文所述两种自适应验配方法，设计了相应的人机交互模块，包括纯音测试以及语音测试两个部分。其中，两个算法的纯音测试界面相同，而语音测试界面稍有差别。为配合以上算法的研究，设计了16通道6阶IIR滤波器组，实验测试证明了该滤波器组性能良好，可以作为自适应验配中的语音信号处理模块。最后的语音识别率以及语音质量的测试再次证明了两种自适应验配方法的良好性能。

6.2 展望

由于助听器的自适应验配技术发展时期非常之短，该课题的研究人员没有详尽的资料可以参考，尚属研究的起步阶段。本文虽然对两种自适应验配技术进行了研究，但由于研究水

平、时间以及实验条件的限制，本文的研究内容还有许多不足之处以待完善，谨提出以下几点需要进一步研究的问题以供参考：

(1)就基于交互式遗传算法的自适应验配技术而言，平均验配时间都在 30 分钟以上(不包含教测试者使用验配系统的时间)，本文虽然设计了旨在降低用户疲劳度的优化辅助模块，但是用户在进行自适应验配的过程中还是会感到乏味。因此，如何进一步降低用户的疲劳度是一个值得深入的研究方向。此外，本文所设计的变异算子并不具备通用性，对于不同于本文的助听器验配参数，本文的变异算子是不适用的。因此，需要设计出更合理更通用的变异算子。

(2)对于基于高斯过程回归的自适应验配方法，本文虽然完成了整个算法的设计流程，但由于本人在机器学习领域的水平有限，没有对核函数以及超参数的选择进行更深一步的研究。此外，该方法对于不同个数的验配参数，所设计的超参数形式也需要调整。如何设定一个统一的标准，使得不同个数的验配参数可以有相应的超参数形式，这里还需要进行大量的工作。

(3)本文所设计的人机交互界面虽然初步实现了助听器自适应验配的相关功能，但是可以尝试设计更为完善的交互界面，使得用户有更好的操作体验。此外，由于研究条件限制，本文的实验基本都是在 MATLAB 平台上完成，并没有在真正的助听器开发平台上验证算法效果。有条件的情况下，可以在相应平台上完成算法效果测试。

致谢

时光匆匆如白驹过隙，转眼间硕士学习的阶段已经接近尾声。两年多的时间教会了我很多知识，也让我变得更加成熟。回首往昔，许多画面如电影片段般浮现在我的眼前，是老师的教诲、同学的扶持以及朋友的鼓励让我在科研的道路上拼搏奋进。值此论文即将完成之际，特向所有关心帮助过我的老师、同学和朋友表示由衷的感谢！

首先要感谢的是我的导师赵力教授，在攻读硕士学位期间，赵老师不仅在科研上给予了我无私的指导，在生活中更是给予了极大的关心和帮助。赵老师治学严谨、谦逊待人，其忘我的工作态度和对科研的执着追求深深感染了我们身边每一位人。在这里，我要向赵老师表示崇高的敬意和由衷的感谢！

感谢王青云博士、梁瑞宇博士，感谢他们无私分享研究成果。感谢我的师兄师姐伏志超、郝文杰、夏岱岱、陈文汐、高歌等人，在我科研上遇到困难的时候，他们总是无私的将自己的宝贵经验传授给我。感谢我的同门巨晓正、吕亮、荆丽、袁杰、朱从贤、郑丽武，是大家的努力让实验室的氛围变得积极向上、充满阳光，也是他们使得我的研究生生涯变得丰富多彩。

感谢我的父母，他们是我这么多年学习、生活的坚强后盾，是他们辛勤的工作让我的物质生活得到保障，也是他们正确的教育和引导给予了我健全的人格，他们的无私奉献让我在前进的路上不敢有丝毫懈怠。

感谢所有参与本文审阅工作的各位专家和教授，感谢你们提出的宝贵意见。最后，再一次向所有帮助过我的老师、同学和朋友表示由衷的感谢，祝愿你们身体健康，平安快乐！

参考文献

- [1] 孙喜斌, 于丽玫. 听力残疾研究现状及发展趋势[J]. 中国听力语言康复科学杂志, 2009 (2): 29-34.
- [2] 孙喜斌, 魏志云, 于丽玫, 等. 中国听力残疾人群现状及致残原因分析[J]. 中华流行病学杂志, 2008, 29(7): 643-646.
- [3] Gopinath B, Schneider J, Hartley D, et al. Incidence and predictors of hearing aid use and ownership among older adults with hearing loss[J]. Annals of epidemiology, 2011, 21(7): 497-506.
- [4] Convery E, Keidser G, Dillon H, et al. A self-fitting hearing aid: Need and concept[J]. Trends in amplification, 2011: 1084713811427707.
- [5] Swanepoel D W, Clark J L, Koekemoer D, et al. Telehealth in audiology: The need and potential to reach underserved communities[J]. International journal of audiology, 2010, 49(3): 195-202.
- [6] Jacobs P G, Saunders G H. New opportunities and challenges for teleaudiology within Department of Veterans Affairs[J]. Journal of rehabilitation research and development, 2014, 51(5): vii-xii.
- [7] 彭丰, 王杰. 助听器验配: 一场悄然兴起的技术革命[J]. 中国听力语言康复科学杂志, 2012, 10(2): 135-138.
- [8] 陈晓靖, 谷京城. 助听器验配效果的评估方法及进展[J]. 听力学及言语疾病杂志, 2004, 12(3): 191-193.
- [9] Cox R M, Alexander G C. The abbreviated profile of hearing aid benefit[J]. Ear and hearing, 1995, 16(2): 176-186.
- [10] Yun D H, Yoon T H, Lee K S. Subjective satisfaction in hearing aid users by APHAB[J]. Korean Journal of Otolaryngology-Head and Neck Surgery, 2000, 43(7): 698-702.
- [11] 段世梅. 助听器验配环节的发展策略探讨[J]. 中国医疗器械杂志, 2014, 06: 445-447.
- [12] 曹宏彦, 曹国强. 全数字助听器的基本原理与应用[J]. 电脑知识与技术, 2007 (24).
- [13] 王鹏. 数字助听器响度补偿算法研究及其系统实现[D]. 中北大学, 2013.
- [14] 唐燕. 数字助听器自适应声反馈消除的研究[D]. 南京邮电大学, 2013.
- [15] 侯非. 残疾人残疾分类和分级[J]. 标准科学, 2014 (010): 19-24.
- [16] 冯定香, 曾高滢. 噪声性听力损失[J]. 中国听力语言康复科学杂志, 2012, 10(1): 70-72.
- [17] Moore B C J, Glasberg B R, Stone M A. New version of the TEN test with calibrations in dB HL[J]. Ear and hearing, 2004, 25(5): 478-487.
- [18] 梁瑞宇. 数字助听器关键技术研究[D]. 东南大学, 2012.

- [19] Ahn J H, Kang H H, Kim T Y, et al. Lipoic acid rescues DBA mice from early-onset age-related hearing impairment[J]. *Neuroreport*, 2008, 19(13): 1265-1269.
- [20] Gopinath B, Schneider J, Hartley D, et al. Incidence and predictors of hearing aid use and ownership among older adults with hearing loss[J]. *Annals of epidemiology*, 2011, 21(7): 497-506.
- [21] Kiefer J, Pok M, Adunka O, et al. Combined electric and acoustic stimulation of the auditory system: results of a clinical study[J]. *Audiology and Neurotology*, 2005, 10(3): 134-144.
- [22] James C, Albegger K, Battmer R, et al. Preservation of residual hearing with cochlear implantation: how and why[J]. *Acta oto-laryngologica*, 2005, 125(5): 481-491.
- [23] 陈振声. 助听器验配方案——小儿助听器验配指导[J]. *听力学及言语疾病杂志*, 2004, 12(4): 265-267.
- [24] Gatehouse S, Naylor G, Elberling C. Benefits from hearing aids in relation to the interaction between the user and the environment[J]. *International Journal of Audiology*, 2003, 42(sup1): 77-85.
- [25] Wang L. Intelligent optimization algorithms with applications[J]. Tsinghua University & Springer Press, Beijing, 2001.
- [26] Altıparmak F, Gen M, Lin L, et al. A genetic algorithm approach for multi-objective optimization of supply chain networks[J]. *Computers & Industrial Engineering*, 2006, 51(1): 196-215.
- [27] Ho S Y, Shu L S, Chen J H. Intelligent evolutionary algorithms for large parameter optimization problems[J]. *Evolutionary Computation, IEEE Transactions on*, 2004, 8(6): 522-541.
- [28] Gonçalves J F, Mendes J J M, Resende M G C. A genetic algorithm for the resource constrained multi-project scheduling problem[J]. *European Journal of Operational Research*, 2008, 189(3): 1171-1190.
- [29] Fogel D B. What is evolutionary computation?[J]. *Spectrum, IEEE*, 2000, 37(2): 26, 28-32.
- [30] Goncalves J F, de Magalhães Mendes J J, Resende M G C. A hybrid genetic algorithm for the job shop scheduling problem[J]. *European journal of operational research*, 2005, 167(1): 77-95.
- [31] Tsai J T, Liu T K, Chou J H. Hybrid Taguchi-genetic algorithm for global numerical optimization[J]. *Evolutionary Computation, IEEE Transactions on*, 2004, 8(4): 365-377.
- [32] 巩敦卫, 郝国生, 严玉若. 交互式遗传算法基于用户认知不确定性的定向变异[J]. 2015.
- [33] Takagi H. Interactive evolutionary computation: Fusion of the capabilities of EC optimization and human evaluation[J]. *Proceedings of the IEEE*, 2001, 89(9): 1275-1296.
- [34] Kacprzyk J. *Studies in Computational Intelligence, Volume 143*[J]. 2008.

- [35] Dawkins R. The blind watchmaker: Why the evidence of evolution reveals a universe without design[M]. WW Norton & Company, 1986.
- [36] Hayashida N, Takagi H. Acceleration of EC convergence with landscape visualization and human intervention[J]. Applied Soft Computing, 2002, 1(4): 245-256.
- [37] Saez Y, Isasi P, Segovia J, et al. Reference chromosome to overcome user fatigue in IEC[J]. New Generation Computing, 2005, 23(2): 129-142.
- [38] Hsu F C, Huang P. Providing an appropriate search space to solve the fatigue problem in interactive evolutionary computation[J]. New Generation Computing, 2005, 23(2): 115-127.
- [39] Takagi H, Ohsaki M. IEC-based hearing aid fitting[C]//Systems, Man, and Cybernetics, 1999. IEEE SMC'99 Conference Proceedings. 1999 IEEE International Conference on. IEEE, 1999, 3: 657-662.
- [40] Wang S, Takagi H. Improving the Performance of Predicting Users' Subjective Evaluation Characteristics to Reduce Their Fatigue in IEC[J]. Journal of physiological anthropology and applied human science, 2005, 24(1): 81-85.
- [41] Durant E A. Hearing aid fitting with genetic algorithms[D]. The University of Michigan, 2002.
- [42] Deb K, Pratap A, Agarwal S, et al. A fast and elitist multiobjective genetic algorithm: NSGA-II[J]. Evolutionary Computation, IEEE Transactions on, 2002, 6(2): 182-197.
- [43] Whitley D. A genetic algorithm tutorial[J]. Statistics and computing, 1994, 4(2): 65-85.
- [44] 王小平, 曹立明. 遗传算法: 理论, 应用及软件设计[J]. 2002.
- [45] Srinivas M, Patnaik L M. Adaptive probabilities of crossover and mutation in genetic algorithms[J]. Systems, Man and Cybernetics, IEEE Transactions on, 1994, 24(4): 656-667.
- [46] 张寅. 用于回归预测的高斯过程模型研究[D]. 河北工业大学, 2014.
- [47] 苏国韶, 张研, 张小飞. 高斯过程机器学习方法在地下水位预测中的应用[J]. 中国农村水利水电, 2008 (12): 48-50.
- [48] 苏国韶, 宋咏春, 燕柳斌. 高斯过程机器学习在边坡稳定性评价中的应用[J]. 岩土力学, 2009, 30(3): 675-679.
- [49] 雷雨, 赵丹宁, 高玉平, 等. 基于高斯过程的日长变化预报[J]. 天文学报, 2015, 56(1): 53-62.
- [50] Deisenroth M P, Fox D, Rasmussen C E. Gaussian processes for data-efficient learning in robotics and control[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2015, 37(2): 408-423.
- [51] 何志昆, 刘光斌, 赵曦晶, 等. 高斯过程回归方法综述[J]. 控制与决策, 2013, 28(8): 1121-1129.
- [52] 庄超. 基于高斯过程回归的强化学习算法研究[D]. 苏州大学, 2014.

- [53] 周楠. 基于高斯过程机器学习方法的证券预测模型研究[D]. 电子科技大学, 2012.
- [54] Dillon H, Zakis J A, Mcdermott H, et al. The trainable hearing aid: What will it do for clients and clinicians?[J]. *Hearing Journal*, 2006, 59(4).
- [55] Rasmussen C E. Gaussian processes for machine learning[J]. 2006.
- [56] Nielsen J B B, Nielsen J, Larsen J. Perception-based personalization of hearing aids using Gaussian processes and active learning[J]. *Audio, Speech, and Language Processing, IEEE/ACM Transactions on*, 2015, 23(1): 162-173.
- [57] Vanhatalo J, Riihimäki J, Hartikainen J, et al. GPstuff: Bayesian modeling with Gaussian processes[J]. *The Journal of Machine Learning Research*, 2013, 14(1): 1175-1179.
- [58] Ferrari S, Cribari-Neto F. Beta regression for modelling rates and proportions[J]. *Journal of Applied Statistics*, 2004, 31(7): 799-815.
- [59] Jensen B S, Nielsen J B, Larsen J. Efficient preference learning with pairwise continuous observations and Gaussian Processes[C]//*Machine Learning for Signal Processing (MLSP), 2011 IEEE International Workshop on*. IEEE, 2011: 1-6.
- [60] 贺建军. 基于高斯过程模型的机器学习算法研究及应用[D]. 大连: 大连理工大学控制科学与工程学院, 2012.
- [61] 张研, 苏国韶, 燕柳斌. 基于粒子群优化与高斯过程的协同优化算法[J]. *系统工程与电子技术*, 2013, 35(6): 1342-1347.
- [62] 张乐, 刘忠, 张建强, 等. 基于人工蜂群算法优化的改进高斯过程模型[J]. *国防科技大学学报*, 2014, 36(1): 154-160.
- [63] 马俊燕, 廖小平, 夏薇, 等. 基于高斯过程机器学习的注塑过程建模及工艺参数优化设计[J]. *机械设计与制造*, 2013 (3): 17-19.
- [64] Chu W, Ghahramani Z. Preference learning with Gaussian processes[C]// *International Conference, Bonn, Germany, August*. ACM, 2005:137--144.
- [65] Houthby N, Huszár F, Ghahramani Z, et al. Bayesian active learning for classification and preference learning[J]. *arXiv preprint arXiv:1112.5745*, 2011.
- [66] Jones D R. A taxonomy of global optimization methods based on response surfaces[J]. *Journal of global optimization*, 2001, 21(4): 345-383.