

电子科技大学

UNIVERSITY OF ELECTRONIC SCIENCE AND TECHNOLOGY OF CHINA

专业学位硕士学位论文

MASTER THESIS FOR PROFESSIONAL DEGREE



论文题目 利用声源信息的多目标定位与跟踪研究

专业学位类别 工 程 硕 士

学 号 201122020677

作 者 姓 名 梁上烈

指 导 教 师 唐 续 副教授

分类号_____密级_____

UDC^{注1}_____

学 位 论 文

利用声源信息的多目标定位与跟踪研究

(题名和副题名)

梁上烈

(作者姓名)

指导教师_____唐 续_____副教授_____

电子科技大学_____成 都_____

(姓名、职称、单位名称)

申请学位级别 硕士 专业学位类别 工 程 硕 士

工程领域名称 电子与通信工程

提交论文日期 2016.03.28 论文答辩日期 2016.05.03

学位授予单位和日期 电子科技大学 2016 年 6 月

答辩委员会主席_____

评阅人_____

注 1：注明《国际十进分类法 UDC》的类号。

Study of Multi Target Location and Tracking Using the Sound Source Information

A Master Thesis Submitted to
University of Electronic Science and Technology of China

Major: **Master of Engineering**

Author: **Liang Shanglie**

Supervisor: **A.P. Tang Xu**

school: **School of Electronic Engineering**

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。据我所知，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，也不包含为获得电子科技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谢意。

作者签名：_____ 日期：_____ 年 _____ 月 _____ 日

论文使用授权

本学位论文作者完全了解电子科技大学有关保留、使用学位论文的规定，有权保留并向国家有关部门或机构送交论文的复印件和磁盘，允许论文被查阅和借阅。本人授权电子科技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，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应遵守此规定）

作者签名：_____ 导师签名：_____

日期：_____ 年 _____ 月 _____ 日

摘 要

利用声源信息的定位与跟踪技术已广泛应用于电话与视频会议、人机交互、说话人跟踪与监视、智能机器人、反狙击系统等。但由于声源信息易受环境及噪声的影响，如何精确而快速地对声源目标的定位与跟踪意义重大。在真实的应用场景中，目标往往是多个，而且个数会随着时间的变化，这就使得声源定位与跟踪问题变得更加复杂。基于随机集理论的概率假设密度（probability hypothesis density, PHD）滤波多目标跟踪算法，突破了传统的数据关联方法，具有严格的数学理论基础，大大降低了计算量，开拓了多目标跟踪的研究领域。本文在此基础上研究了一种多检测概率假设密度（MD-PHD）滤波多目标跟踪算法。该算法利用传播过程中的多径信号实现对目标的定位与跟踪。以下是本文的主要研究内容：

第一，研究了真实环境下声源信号的特征及传播模型。在此基础上建立了声源信号的传播模型，并通过仿真分析了房间内声音信号的脉冲响应。为了提高后续处理的高效性，通过仿真分析了声源信号的预处理过程。

第二，研究了时延估计算法中的多种广义互相关法。通过仿真实验比较了各种广义互相关法的优缺点，并重点研究了 PHAT-GCC 时延算法。针对标准的 PHAT-GCC 性能易受噪声影响的问题，研究了一种改进的 PHAT-GCC 算法。通过仿真验证了改进的 PHAT-GCC 算法性能更优。

第三，基于随机有限集的理论，建立了递归多目标跟踪模型，并研究分析了概率假设密度（PHD）滤波多目标跟踪算法及其粒子滤波实现。通过理论分析及实验仿真表明，粒子滤波概率假设密度（SMC-PHD）滤波多目标跟踪算法能够有效实现对目标的定位与跟踪。

第四，针对室内应用场景中，接收信号中除了直达波信号还包含多径信号的特点，本文研究了一种利用这些多径信息的多检测概率假设密度滤波多目标跟踪算法。该算法不仅能有效实现对声源目标的定位与跟踪，还减少了麦克风的数量。通过实验仿真，验证了该算法在单目标场景下和多目标场景下的有效性，评估了其在不同信噪比下的跟踪性能。

关键词：声源信息，定位与跟踪，广义互相关，多检测概率假设密度，多径信息

ABSTRACT

The sound source information for locating and tracking technology has been widely used in telephone and video conference, human-computer interaction, the speaker tracking and monitoring, intelligent robot, the sniper system, etc. However, due to the sound source information susceptible to environmental noise, how to accurately and quickly achieve the sound source locating and tracking is of great significance. In the real application scenario, the target number often is not a single, but multiple, and the number constantly will change over time, this makes the sound source locating and tracking algorithm become more complex. Under random set theory, probability hypothesis density (probability hypothesis density, PHD) filter for multiple target tracking algorithm has a strict mathematical theory, breaks through the traditional method of data association, reduces the amount of calculation, opens up a multiple target tracking area of research. This paper studies multi-detection probability hypothesis density (MD-PHD) multiple target tracking algorithm that uses multipath signal propagation to locate and track target, the following is the main research content of this article:

First, study the characteristics of sound signal in a real environment and propagation model. Then establish a sound source signal propagation model. Through simulation, analyze the voice signal pulse response of the room. In order to improve the efficiency of subsequent processing, simulation of the sound source signal preprocessing was analyzed.

Second, study the variety of generalized cross-correlation method for time delay. Through the simulation experiment, compare the advantages and disadvantages of various kinds of generalized cross correlation, the PHAT-GCC delay algorithm is studied mainly. Because the standard PHAT-GCC performance is affected easily by noise, research an improved PHAT-GCC algorithm. Improved PHAT-GCC has better performance by the simulation the simulation results.

Third, on the concept of stochastic finite set, the recursive multiple target tracking model is established, then study and analyze the probability hypothesis density (PHD) filtering multiple target tracking algorithm and the particle filter. Through theoretical analysis and experimental simulation show that the particle filter probability hypothesis

density filter for multiple target tracking algorithm (SMC -PHD) can achieve the target locating and tracking effectively.

Fourth, in actual indoor application scenarios, receiving signal contains the direct signals and multipath signals, this paper studies Multi-detection probability hypothesis density multiple target tracking algorithm that exploits the multipath signals. The algorithm not only can achieve the target locating and tracking effectively, also reduce the volume of the microphone. Through the simulation experiments verify the effectiveness of the proposed algorithm in single-object scenario and multi-object scenario, and under different signal-to-noise ratio, the tracking performance is evaluated.

Keywords: sound source information, locating and tracking, generalized cross correlation, multi-detection probability hypothesis density, multi-path information

目 录

第一章 绪论	1
1.1 论文的研究背景及意义	1
1.2 国内外研究现状	2
1.3 麦克风阵列声源定位方法概述	4
1.3.1 最大输出功率的可控波束形成方法	4
1.3.2 基于高分辨率谱估计方法	4
1.3.3 基于时延估计方法	5
1.4 论文主要创新点及内容安排	6
第二章 声源信号分析与预处理	8
2.1 声源信号概述	8
2.2 声源信号的预处理	8
2.2.1 预滤波	8
2.2.2 分帧加窗处理	9
2.2.3 端点检测	11
2.3 麦克风阵列信号模型	12
2.4 本章小结	14
第三章 时延估计算法研究	15
3.1 时延估计技术	15
3.1.1 互相关时延估计法	15
3.1.2 广义互相关时延估计法	16
3.1.3 改进的 PHAT-GCC 定位算法	19
3.2 仿真	21
3.2.1 仿真环境	21
3.2.2 实验结果与分析	21
3.3 本章小结	24
第四章 多目标跟踪滤波算法研究与分析	25
4.1 随机集的多目标跟踪理论	25
4.1.1 随机集理论基础	26
4.1.2 随机集多目标跟踪模型	28
4.2 PHD 的粒子滤波实现	31

4.3 PHD 的粒子滤波仿真	32
4.4 本章小结.....	35
第五章 基于声源信息的多目标定位与跟踪研究与实现	36
5.1 单声源目标跟踪模型	36
5.2 多声源目标跟踪模型	37
5.2.1 多声源目标状态空间模型	37
5.2.2 多声源目标跟踪实现算法研究	38
5.3 多传感器 PHD 滤波算法研究与仿真	40
5.4 多检测 PHD 滤波算法研究	43
5.4.1 多检测目标跟踪背景.....	43
5.4.2 多检测 PHD (MD-PHD) 滤波器.....	45
5.5 实验仿真及结果分析	47
5.5.1 单声源目标定位与跟踪实现	47
5.5.2 多声源目标定位与跟踪实现	50
5.6 本章小结.....	53
第六章 总结与展望	55
6.1 总结	55
6.2 工作展望.....	56
致 谢.....	57
参考文献.....	58

第一章 绪论

1.1 论文的研究背景及意义

众所周知，定位与跟踪技术已经广泛用于通讯、雷达、医学、声纳、和航空航天等诸多领域，而且相当成熟，但是将定位与跟踪技术应用于声音领域还比较少，所取得的研究成果就更少。利用声源信息对目标进行定位和跟踪有着广泛的应用前景，可应用于视频会议^[1]、智能机器人^[2]、声源监控^[3]、反狙击步枪^[4]等领域。正是由于其广泛的应用前景，近年来，利用声源信息对目标进行定位和跟踪已经成为一门新的研究热点。在传统的视频会议系统中，通常情况下，采集视频的设备（比如摄像头）都是固定在室内的某个地方，并由专门的人员根据说话者位置的变化而人工地移动摄像头，虽然人工的方法执行起来简单，但是人工的方法即耗时耗力又不能快速的将摄像头对准的说话者，而利用声源定位和跟踪技术可以直接定位出提供说话人的方位，使摄像头自动发现说话人的方位，对准说话人。同时，对目标是非静止的，摄像头能跟踪行动的目标，摄像头随着目标位置的改变而改变。在军事领域，隐蔽的狙击手，从来都是步兵的最大敌人，特别是在城市巷战中，当敌方的狙击手潜伏在城市的建筑物里面时，我方很难发现其踪迹。因此，如何快速地定位出狙击手的位置，并迅速地歼灭狙击手减少我方人员的伤亡，将成为影响战争态势的一个重要因素。

利用声源定位与跟踪技术较雷达、图像等其他探测与跟踪技术而言，它具有以下几点优势：第一，声音不受视线的制约，声波波长比较长，可以传播到障碍物后面去，因此声音探测装置就可以定位出障碍物后的目标，而无线电、激光、图像等方式就没有这样的优势；第二，隐蔽性好，声音由声源发出，被动声探测系统接收，自身不向外发出任何形式的信号，因此可以避免电磁波的干扰，隐蔽性强，不易被发现；第三，外界环境影响对系统性能影响较小，声探测系统可以在各种复杂的环境（如阴天、雾天、天上、水下）下正常使用，因此受外界环境影响较小；第四，探测器成本低，声探测器（如麦克风）价格都不贵，因此可以通过大量的使用声探测器达到高性能的跟踪结果的效果。在实际的研究和应用领域，一般选择麦克风作为声探测系统装置。

然而，利用声源信息进行定位与跟踪也不是十分完善，自身也存在许多问题。与传统的探测技术相比，利用声源信息对目标进行定位与跟踪存在以下难度和挑战：第一，声源信号频率分布范围较宽，而每个发声体的中心频率又各不相同，所以声源面临的就是一个复杂的多频宽带信号；第二，声源接收装置处理的声源

信号通常是时变的且又是非平稳信号，因此传统的处理方法在这里不再适用；第三，声源接收器处的声学环境由于受到如噪声、接收器位置误差、房间混响等等因素的影响，这从而使得跟踪性能受到较大的影响^[5]。

在真实应用过程中，目标不止一个而是多个，而目标的个数和位置也不是固定不变的。在实际的环境中，多目标已经不足为奇了，所以通常在设计声源定位与跟踪系统时也要考虑多个声源的问题。而在实践过程中，发现要完成对多声源目标定位与跟踪是相当艰难的，该技术不仅要抑制声源之间的想到影响，还要克服背景噪声和混响的对结果的影响。所以，如何快速地有效地定位多个声源位置，并对多个声源进行实时跟踪具有重要的应用意义，近年来对多声源定位与跟踪算法的探索已经成为一门新的研究课题。

1.2 国内外研究现状

在声源的这一领域，国外很早就开始研究，技术也相对比较成熟，其研究成果也已经用于实际应用中。以美国为首的一些国家，早在上世纪中旬，开始重视利用声音作有一种供军事领域的新的作战力量。目前已经研究出的声源定位系统有，如美国研制出的 PALS 声源定位装置、英国研制出的 HALO 火炮定位武器、瑞典研制出的 Soras6 炮兵声测装置、以色列研制出的 IGLOO 等，作用范围可达 20~40km。还有就是智能雷弹系统，它利用侦测到的声音信号实现目标的定位，然后引导攻击武器实现对目标的效打击。知道雷达易受强电磁波的干扰，容易丢失目标，而且对超低空飞行的目标也不能实现有效的跟踪，声源定位系统相比之下，所受干扰较少，可以侦测特定特征的声音实现对目标的识别与定位跟踪。到目前为止，许多国家都在使用这种新的被动声探测武器^[6]。这些探测与定位系统不仅原理简单，利于便携，生产成本低，而且还有很高，很精准的探测性能。如 2003 年的美伊战争，为了应对复杂的城市巷战环境，美军在战场上使用了狙击手定位系统^[7]。该系统由 7 个小麦克风阵列组成，工作单元被安放在车辆上，这 7 个麦克风组成的阵列能够通过枪口发出的冲击波和子弹与空气相遇生成的超声波定位和测量目标。7 个麦克风会收集到接信号的时间差，然后系统会在很短的时间内计算出目标的方位。它在战场上能及时地探测出狙击手的位置，减少了美军人员的伤亡。

声源信息进行定位和跟踪技术的开发和利用进展非常的快，该技术不仅仅在军事上有用，在民用方面也有较好的前景。如 Flanagan 在 1985 年开发的语音增强系统中就应用到了传声器阵列技术；其后 Flanagan 将该技术应用于截获发声体发出的声音信号；后来 Flanagan 又将其应用于说话人识别^[8]；Brandstein 和 Silverman

利用麦克风阵列实现了对目标的实时跟踪；加拿大研制出了一种由 8 个麦克风组成的机器人，可以实现避障，实现了对多个目标的跟踪。经过多年的发展，现有的成果已经在很多实际运用中体现作用了，如用于声音识别、电视电话视频会议、声音获取、助听装置等等。目前在国外，对声源定位和跟踪技术进行研究与探索的公司有许多，如 Intel、宝利通、东芝等。但在国内，还没有同等规模的研究机构或部门对该领域进行开发，以至于也没有相应的自主知识产权的产品，所以加强在该领域的研究意义重大。

我国目标在利用声源信号方面开展的工作还相对比较少，相对国外存在差距较大，所取得的成果和进展也相对比较少，主要是一些企业和研究所在做。通过近几年的努力，我国也取得了一些成绩。但是开发出来的产品还比较少，功能比较单一，与国外相比，还处于一个开始阶段。例如，陈可、汪增福等人采用声压幅度比来确定声音目标的位置。从电压信号幅度与待测声源目标距离关系就可以得到一系列约束条件表达式，从这些表达式中就能准确定位出目标声源的位置；后一年，林静然等人改进了先前的定位和分离技术，提出了另一种方法，该方法利用双波束二维定位算法和近场最小方差波束形成技术，通过一定的方法将二者联合起来，在基于阵列条件下完成了对声源目标的分离和定位。与 2D-MUSIC 不同，该方法运行起来计算小，而且效果更好，主要运用于近场环境下。居大亮等人采用了一种基于麦克风阵列的子阵算法，通常被告称为三维定位子阵算法。后来，哈工大研究所的工作人员，又将模糊推理用于机器人研究，帮助机器人实现自能定位。该方法首先分析室内环境下噪声的分布，应用所分析的结果对麦克风阵列进行分成两组，通过调节子阵的位置，使收到的声音信号一样，而方向噪声最小，最后运用互相关矩阵求出目标的位置。这能够较好地抑制干扰、混响等噪声，从而改善了系统对目标的跟踪功能。蔡卫平等提出了粒子滤波（PF）的混合声源目标跟踪算法^[9-10]：首先利用可控响应功率可以得出每个时间段的粒子状态观测值，再得出其方差，通过方差的大小对信号帧进行划分：高信噪比和低信噪比。

虽然我国已经对声源定位与跟踪技术已经有部分研究成果了，但是基于声源定位与跟踪系统的研究成果目前在我国还是很匮乏的，在这个计算机和信息并存的时代，拥有一款自主研发的高科技产品无论对军事还是民用，都极其重要，引起了大量学者的科研兴趣^[11]。因而声源定位与跟踪系统这方面的研究也日渐增多，但是挑战与机遇并存，我们还面临着许多还未解决的工作。

1.3 麦克风阵列声源定位方法概述

目前, 利用麦克风阵列对声源目标进行定位的技术有很多种, 下面主要讨论其中的三种, 这三种比较有代表性, 其它的都是在它们的基础上进行的延伸: 一是最大输出功率的可控波束形成 (Steered Beam Forming, SBF) 方法^[12-13], 该技术对传声器收到的声音信号做滤波、加权求和处理, 然后对可能出现目标的位置进行逐一搜索, 然后求波束输出功率, 其中最大值所对应的位置就是目标的位置; 二是高分辨率谱的定位方法, 如最小方差谱估计 (MV)^[14]、自回归模型 (AR)^[15]、多重信号分类 (MUSIC) 方法^[16-17]等等, 它们的基本思路都是利用 EVD 对数据的协方差矩阵做分解, 得到两个独立的子空间: 信号子空间和噪声子空间, 然后求得与噪声子空间正交的方向矢量就是目标声源的可能方向; 三是时延估计 (TDOA) 的定位方法, 通常情况, 分为两步: 首先计算出信号到达不同麦克风相互之间的到达时间差, 然后将该差值乘以声音传播速度就可以得到距离差, 利用多对麦克风求得的距离差就可以求得目标在空间中的位置。

1.3.1 最大输出功率的可控波束形成方法

最大输出功率的可控波束形成 (Steered Beam Forming, SBF) 方法, 该方法首先对传声器收到的声音信号做滤波、加权求和处理, 然后对可能出现目标的位置进行逐一搜索, 然后求波束输出功率, 其中最大值所对应的位置就是目标的位置。基于可控波束形成的定位算法有多种, 主要研究下面两种: 延迟求和波束形成方法 (delay-and-sum beamforming, DSB) 和自适应滤波求和波束形成方法 (filter-and-sum beamforming, FSB)。延迟求和波束形成方法是最容易实现的方法: 延迟传声器所接收到的多路声源信号, 弥补声源到传声器的传播延迟时间, 然后通过求和来得到输出信号, 最后计算输出能量。该方法原理简单、计算量小, 能够较逼真地还原声源信号, 但是在有噪声的情况下, 得到的效果就不好。自适应波束算法是一种相对复杂的算法, 即补偿时延的同时, 也在对信号作滤波处理。自适应滤波器算法由于加入了滤波器, 所以比上一种方法计算量大, 而且信号失真大, 但它只需要较少阵元就能得到较好的效果, 抗混响效果好。

1.3.2 基于高分辨率谱估计方法

高分辨率谱估计方法, 并且在理论条件下在一定的限制条件下可实现任意准确度的定位, 不受信号采样频率的约束, 正是由于这一特性才使得它在阵列信号处理中获得了成功应用。基于高分辨率谱估计方法的定位算法主要分为特征子空间类算法、空间拟合类算法、最小方差 (MV) 法、自回归 (AR) 法、MEM 法、

MUSIC 方法等等。高分辨率谱估计方法是得用时间上平均来估计传声器接收到的信号之间的相关矩阵，但是前提是信号中的参数保持不变，同时，平均值的个数要足够大。然而在实际应用中，声音信号处理一般都位于密闭的室内环境中，室内环境中存在各种混响、反射和噪声，定会对声音信号产生意料不到的干扰。从原理可以知道，本方法只对远场模型有用，在近场模型下，定位误差比较大。同时，高精度谱估计技术通常处理的是窄带信号，而不能处理像声音信号一样的宽带信号，这就需要将宽带信号转化为窄带信号，这不仅会增加运算的复杂度降低算法的实时性，而且还会影响定位的精度，因此不利于实时和高精度的应用中。

1.3.3 基于时延估计方法

基于时间延迟估计的定位方法的思路是将到达两个不同传声器的时间差转变成距离差，从而可以建立一个双曲线面，有多少对麦克风对就有多少个双曲面，而它们的相交的位置就是目标的可能的位置。基于到 TDOA 定位算法步骤如下：一是计算出时延估计，估计出目标点到麦克风对的时间差（TDOA），有多少对麦克风就有多少个时间差，将通常应用到的求时差的方法进行了归纳有如下方法：基本互相关法、最小均方自适应滤波法、广义互相关法、互功率谱相位法^[18-20]等；第二步是位置估计，利用麦克风对求得的时间差，通过一定的方式转化成距离差，然后利用如最小二乘及线性逼近的方法^[21]和极大似然估计方法^[21-22]等方法就能估计出目标的位置。

针对以上三类声源定位方法进行分析研究得出其优缺点如下：

（1）最大输出功率的可控波束方法，需进行全局搜索，所以就增加了算法的计算量，一般不用于实时系统，因此，该算法不太适用于实时快速处理系统。

（2）高分辨率的谱定位方法，该方法运算量大的，不适用于近场信号。同时声音信号是时变信号，其相关估计参数都不是相对固定的，因此语音信号往往不能满足该方法要求信号具有平稳性的条件。

（3）时延估计的定位方法，该方法在实际应用过程中有许多优点，比如计算量小，求得的结果精度高，使得它的运用相当广泛。但是该方法也有许多缺点，主要缺点如下：一是第一阶段的误差将严重影响第二阶段的定位结果；二是实时性差，第二阶段是在第一阶段运算结束后才开始；三是对单一目标可以执行，但是当目标有多个时，由于多个目标相互干扰，因此产生的结果也就不准确；四是麦克风位置布局好坏对定位的精度有较大影响，在实际应用过程中，麦克风的位置和布局存在许多不确定性，这将对计算结果产生明显的影响。

综上所述，利用时延估计定位方法与前两种方法相比，不仅计算量小，而且

定位估计精度也较高，稳定性更强，因此本文将采用时延估计定位方法。

1.4 论文主要创新点及内容安排

本文主要创新点有：一是本文将不采用传统的电磁波，光线，图像等传统信号，而是利用声音信号实现对目标的定位与跟踪；二是在实现目标定位与跟踪的过程中，本文只使用一对麦克风，而不是使用大量的麦克风对，这样的好处是不仅减少了成本的投入，而且利于装备的铺设与快速投入使用；三是针对真实应用场景下存在多径信息，本文研究了一种利用多径信号的多检测概率假设密度滤波多目标跟踪算法，该算法以贝叶斯滤波和概率假设密度为基础，充分利用多径信息，最终通过仿真实验，验证了该算法的有效性。图 1-1 为系统的流程图，一共分为三步：第一步声音信号特性分析，第二步求时延，第三步跟踪。本文正是围绕着这三部分进行了研究。

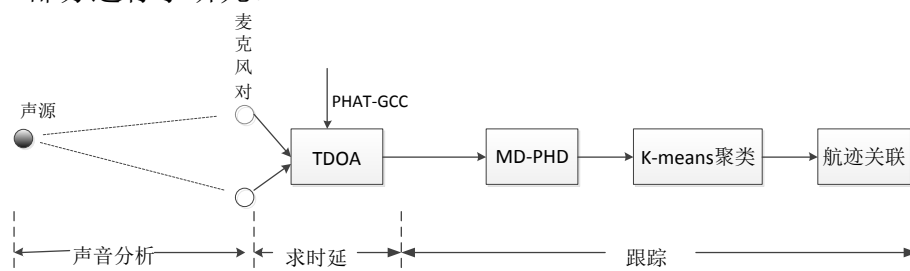


图 1-1 系统流程图

本文研究的主要内容有：

第一章主要介绍利用声源信息的目标定位与跟踪的研究背景及其研究意义，并对其国内外的研究现状进行了综述，最后归纳分析了传统的声源定位方法的优缺点。

第二章主要介绍了声源信号的特性以及声源信号的传播模型，为了提高系统处理信号的有效性，分析研究了声源信号预处理过程并进行了仿真。

第三章归纳分析了现有的各种广义互相关时延估计技术，并对它们各自的优缺点及适用场合进行了详细地探讨和分析，最终选择了 PHAT-GCC 作为时延估计的算法，并在此基础了研究了一种改进的 PHAT-GCC 算法。通过仿真，验证其性能更优。

第四章主要介绍了随机集理论及概论假设密度（PHD）的基本知识，并在此基础上建立了多目标跟踪模型。然后，针对 PHD 滤波算法中积分问题，研究了 SMC-PHD 算法。通过实验仿真验证了其可行性。

第五章首先建立了单声源和多声源目标跟踪模型，然后研究了多对麦克风情

况下的目标跟踪算法并进行了仿真。最后，针对真实应用场景下存在多径信息，研究了一种利用一对麦克风实现对多目标定位与跟踪的多检测概率假设密度滤波算法。最后，通过实验仿真验证了该算法的可靠性。

第六章为本文的结束语，主要对本文中的重点工作做了一个全面的总结，并指出了多检测概率假设密度滤波多目标跟踪算法实现对多目标的定位与跟踪还存在的问题，并对以后的工作进行了一个展望。

第二章 声源信号分析与预处理

2.1 声源信号概述

自然界的声音都是声源发生体通过振动发出的，一般情况下主要研究频率为 20Hz ~ 20kHz 的声音，因为这个范围内的声音能被人的耳朵听到。区别声音，通常用下面三个因素进行区别：响度、音高、音色，也就是通常说到的振幅、频率和相位。声音传播的是波，波通常有振幅，相位，频率和波长，它们通常决定着波的特性。声音的传播速度通常情况下由传播媒质（通常情况下是空气）决定，与声源本身的自我特征无关。空气的温度与湿度也会影响声音传播的速度，通常情况下假设声音的速度不变 $C = 340\text{m/s}$ 。本文主要研究人发出的声音，即主要频段在 100Hz ~ 4kHz 的声音。

2.2 声源信号的预处理

通常情况下，在接收器接收的声音信号中，不仅包含有想要的声音信号，还包含有不想要的噪声或者其它物体发出的声音，所以如果直接把接收器接收到的信号都认为是有用信号，那么系统的准确度将受到很大影响，所以这就需要在定位跟踪之前，对接收器接收到的声音信号进行一系列的预处理操作，主要包括：预滤波、加窗分帧、预加重、归一化和去噪等。

2.2.1 预滤波

对接收到的信号都要先对其进行滤波处理，把所需要的部分保留，不需要的部分去掉，麦克风接收到的信号中包含着大量不需要的信号以及大量的噪声，所以通过预滤波的方式去除这一部分信号。常用的滤波器有低通滤波器、高通滤波器和带通滤波器等等，图 2-1 是运用高通滤波器对声音信号进行滤波的结果，图 2-1(a)高通滤波器的幅频响应和相位响应，图 2-1(b)经高通滤波后的声音信号时域图，图 2-1(c)为经高通滤波后的声音信号频谱。在实际应用中，可以根据具体需要选择滤波器。

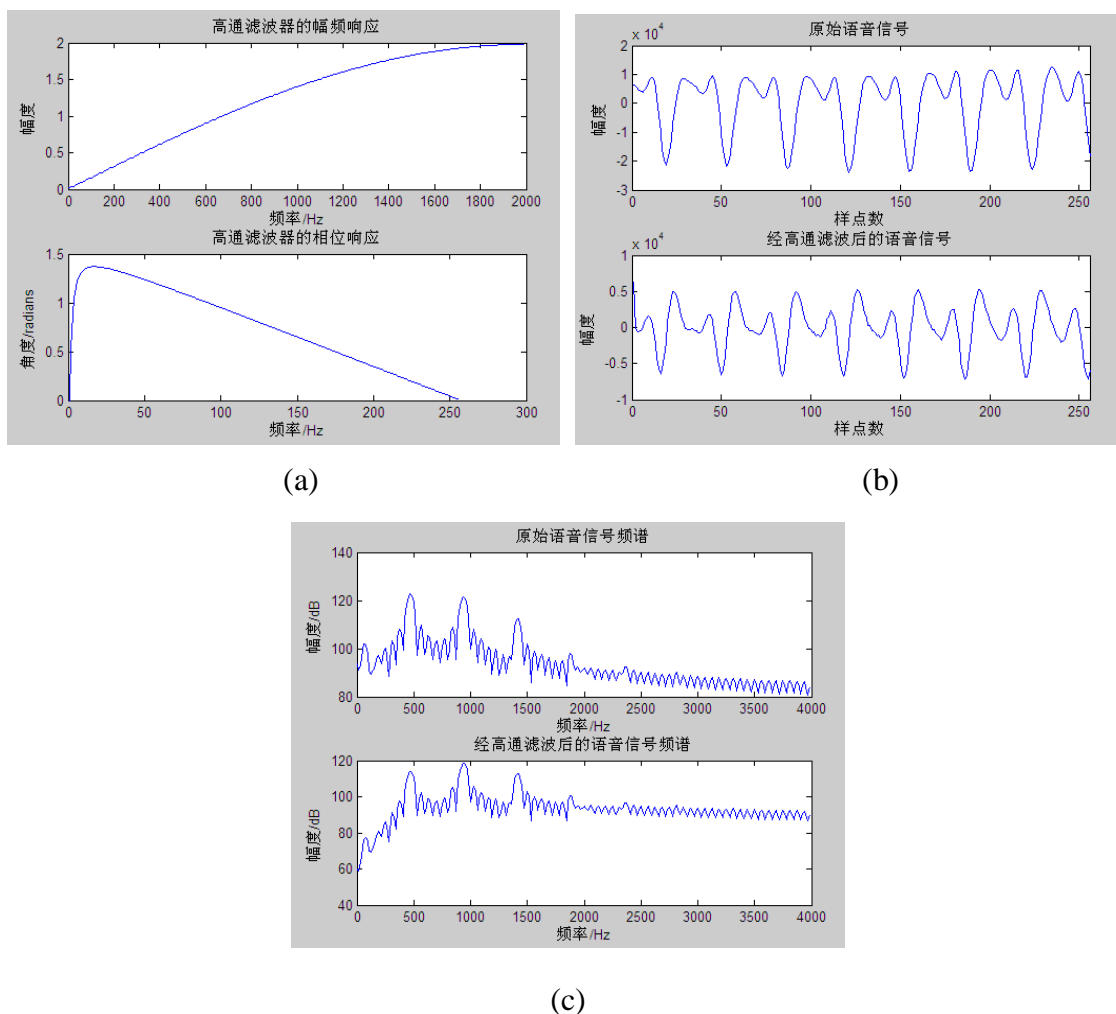


图 2-1 高通滤波器对声音信号进行滤波结果

(a)高通滤波器的幅频响应和相位响应；(b)经高通滤波后的声音信号时域图；
 (c)经高通滤波后的声音信号频谱

2.2.2 分帧加窗处理

声源信号为短时平稳声音信号，只在较短的时间间隔内，是特性不随时间变化的平稳随机信号，正是因为声音信号的这个特性，所以必须对声音信号作加窗分帧处理，使其满足要求。为了保证信号的连续性，通常情况下，都使用重叠分帧的方法，这样的好处是使帧与帧之间的过渡平滑，不会出现突变的情况。分帧处理采用的基本方法是连续交叠分段，交叠部分占 $1/3 \sim 1/2$ ，此时对声源信号进行分析能得到明显的效果，信号的细节特征也表现明显。如果重叠部分太多，会使总体特性被模糊，同时增加计算量。但是也不能重叠过小，否则会出现帧与帧之间细节突变明显，使得分析不具有连续性。在分帧加窗时，通常要考虑以下三个因素：交叠长度、窗函数的选择和帧长，因为它们将决定分帧处理的好坏。

假设原始信号用 $s(n)$ 表示，窗函数用 $w(n)$ 表示，分帧后的信号用 $s_w(n)$ 表示，则加窗后的信号加窗模型为：

$$s_w(n) = s(n) * w(n) \quad (2-1)$$

在通常的应用中，窗函数有矩形窗和汉明窗两种：

$$\text{矩形窗: } w(n) = \begin{cases} 1, & 0 \leq n \leq N-1 \\ 0, & \text{else} \end{cases}$$

$$\text{汉明窗: } w(n) = \begin{cases} 0.54 - 0.46 \cos[2\pi n / (N-1)], & 0 \leq n \leq N-1 \\ 0, & \text{else} \end{cases}$$

图2-2为矩形窗频谱分析图，图2-3为汉明窗的频谱分析图，图2-4是原始语音信号经过矩形窗和汉明窗处理的结果图（经归一化处理），其中帧长为256，帧移为128。

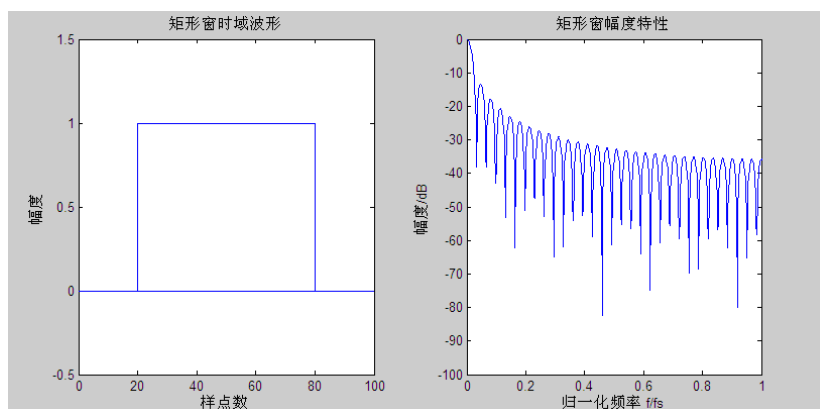


图 2-2 矩形窗频谱图

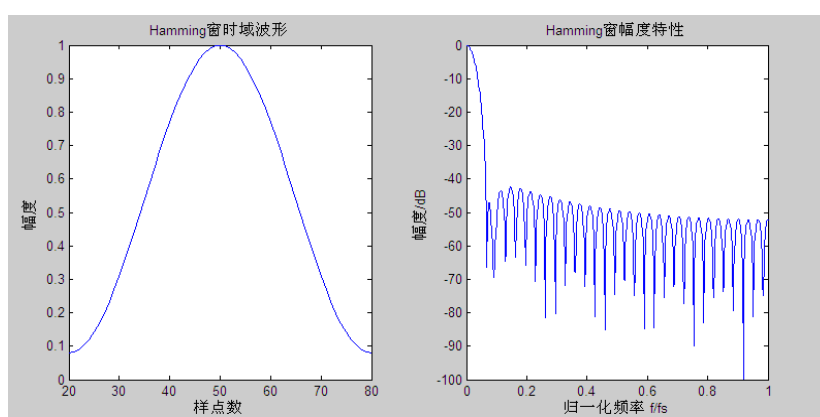


图 2-3 汉明窗频谱图

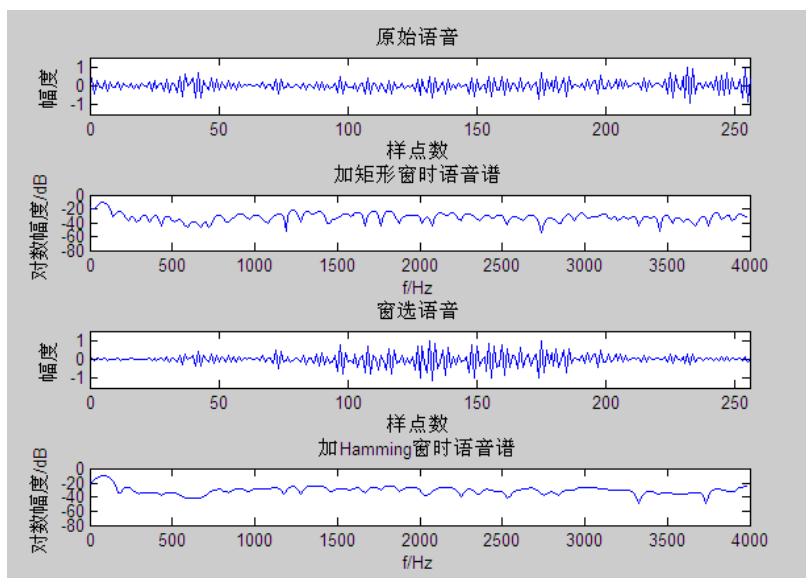


图 2-4 分帧加窗处理结果

比较图 2-2、图 2-3 和图 2-4 波形可知，矩形窗的频率分辨率较高，但却有较高的旁瓣；汉明窗主瓣较宽，但旁瓣衰减过快。再者，前者比后者频谱泄露严重，后者具有更平滑的低通特性。通常，根据实际应用要求，选择性能较优的窗函数。本文需要较高的频率分辨率，所以选用矩形窗作为分帧加窗函数。

2.2.3 端点检测

声音信号的检测对整个工作起至关重要的作用，如果能有效地从接收到的信号中提取出声音信号并正确确定其起始点和结束点将大大缩减计算时间。同时，又能避免误将噪声当作声音信号进行处理而出现错误结果，以及避免当前错误结果对后续信号帧处理工作造成影响。

如果接收的信号的信噪比较高，那么就可利用求短时能量法分辨出有用的声音部分，而剔除不需要的声音部分。因此，本文利用求短时能量法作为有用声音信号的检测法。

假设用 E_n 代表信号的短时能量，那么可得：

$$E_n = \sum_{m=-\infty}^{+\infty} [x(m)w(n-m)]^2 = \sum_{m=n-(N-1)}^n [x(m)w(n-m)]^2 \quad (2-2)$$

上式中， N 代表窗函数的长度，从 E_n 的表达式中可以看出，求短时能量法就是对 N 个样本点的能量进行平均求和。如图 2-5 为声音的短时能量，对 N 分别取 50、100、400、800 时的短时能量进行了比较，从图中可以看出当 N 的取值越大，短时能量的值就越大，越容易检测出信号，但是帧与帧之间的平滑性越差；当 N

的越小时，短时能量的值就越小，越不利于信号的检测，但是帧与帧之间平滑性越好。所以在实际的应用过程中，要根据实际需要选择 N 的取值，本文中取 $N=100$ 。

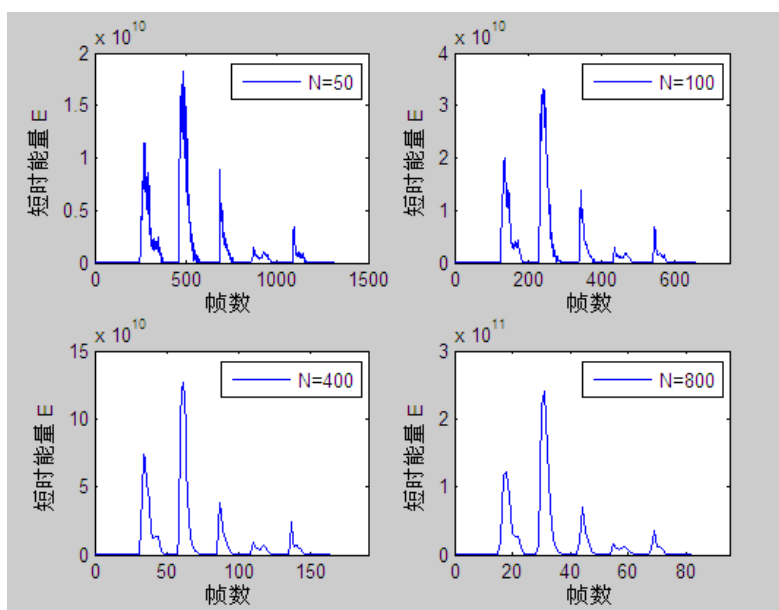


图 2-5 声音的短时能量

求短时能量检测法首先是求取当前时刻帧的能量值，然后与一个阈值进行对比，如果比设定阈值大则判定当前帧为信号帧，比设定的阈值小则认为当前帧不是信号帧。所以门限的选择是一个重要内容。例如，记录最小的信号短时能量，将其作为背景噪声能量，再将其加上 10dB 作为阈值，在更新信号最小短时能量的同时更新阈值；再如，记录最大的信号短时能量，将其乘以一个因子 α 作为阈值，通常取 $\alpha=0.1$ 。本文采用记录最大的信号短进能量乘以一个因子作为阈值。

2.3 麦克风阵列信号模型

采集声源的设备通常是麦克风阵列，麦克风阵列是把许多个麦克风在空间进行布阵，以利于对信号进行接收和处理。利用多个麦克风与只利用一个麦克风相比，有许多明显的好处。不仅能够接收声音信号作为后续处理的输入，而且能够起到抑制环境噪声的作用，而且多个麦克风在应用上，有更好的空间选择性。为了对麦克风接收的信号有一个具体的认识，将信号的传播过程用图 2-6 显示，从图中，可以看到麦克风接收到的信号包含的直达信号，反射信号还有噪声信号。

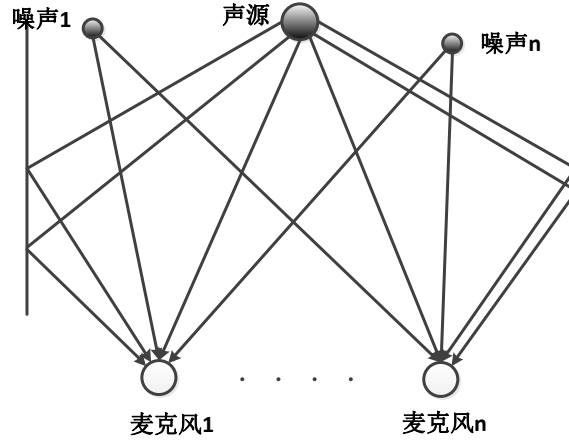


图 2-6 语音信号接模型

与电磁信号相似，声音信号的传播过程也相当复杂，从声音产生到到达麦克风，整个过程会受到许多因素的影响，这其中包括反射，多次反射，衰减等等，所以到达麦克风的信号将包含许多不同的成分，这会使得声音信号处理难度加大，同时声音信号也会越来越弱。因此，在真实情况下，采用理想环境模型不再合适。所以，必须在麦克风接收到的直达信号中，加入多径信号和噪声，模拟真实的声学环境。

为了模拟真实的声学环境，本文采用映像算法（IMAGE）来计算房间内的声学通道响应。声音信号在一个封闭的环境(例如室内)里传播时，有些是从声源直接传播到麦克风，有些却是经由墙壁、天花板、地面等障碍物反射一次或数次才能到达麦克风，这使得达到麦克风的语音信号的路径除了直接路径外还有其他的间接路径，因此导致了接收信号幅度衰减、能量变小等问题。在物理声学上，定义回声为反射波的延迟时间达到约 50ms 以上，其余的反射波所产生的物理效应为混响。从而反射声就可以看作是房间脉冲响应与声源信号的卷积结果。产生房间脉冲响应函数用 g_n 来表示，公式如下：

$$g_n = \sum_{p=0}^1 \sum_{r=-\infty}^{\infty} \beta_{x1}^{|n-q|} \beta_{x2}^{|n|} \beta_{y1}^{|l-i|} \beta_{y2}^{|l|} \beta_{z1}^{|j-k|} \beta_{z2}^{|j|} \times \frac{\delta(n - |R_p + R_r|/c)}{4\pi |R_p + R_r|} \quad (2-3)$$

其中， β_{x1} ， β_{x2} ， β_{y1} ， β_{y2} ， β_{z1} ， β_{z2} 分别是各面墙的反射系数， (x, y, z) 是声源位置的坐标， (x_1, y_1, z_1) 是麦克风位置的坐标， c 是声音的传播速度， L_x ， L_y ， L_z 分别是房间的长、宽、高，声源的镜像向量为 $R_r = (2nL_x, 2nL_y, 2nL_z)$ 和 $R_p = [(x - x_1 + 2qx_1), (y - y_1 + 2ir_1), (z - z_1 + 2kz_1)]$ 。仿真结果如 2-7 所示，从图中可以看出，真实环境下，传感器接收到的信号中除了一条直达信号外，还有多径产生的信号。

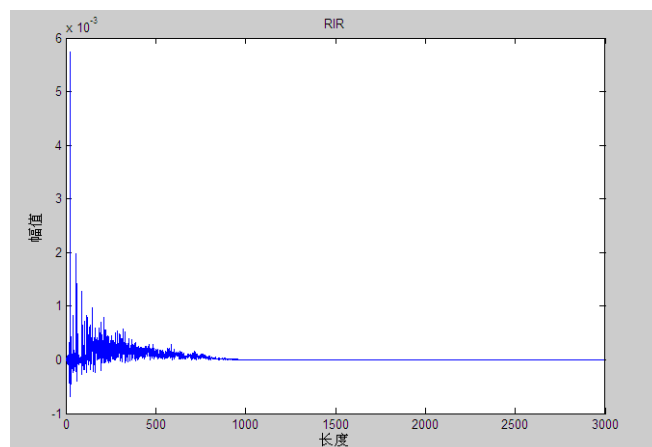


图 2-7 房间脉冲响应

2.4 本章小结

本章分析了声源信号的特性及传播模型，并仿真了真实环境下房间内的脉冲响应。然后研究了声音信号的预处理过程，主要包括：预滤波、加窗分帧、预加重、归一化和去噪。通过声源信号预处理过程，不仅可以减少后续工作的复杂性，还可以提高后续工作中求时延估计的精度。

第三章 时延估计算法研究

本章主要研究时延估计算法。时延估计算法有许多的优点，然而在实际运用过程中，由于噪声和混响的干扰，给时延估计带来了严格挑战。而产生的误差会直接影响到后续定位与跟踪的精度。因此，针对如何提高时延估计精度的问题，许多学者和研究人员提出多种适用性强的算法，包括互相关法（CC）、广义互相关法（GCC）、特征时延估计法、相位谱法和高阶统计量法等。然而，每种算法都各有利弊^[23]。在本文中主要采用经典的广义互相关（GCC）法及其改进算法。该类算法具有实现容易，计算量较小的特点。

3.1 时延估计技术

3.1.1 互相关时延估计法

互相关函数是经常使用的一种求时延估计法，利用互相关法求一对麦克风接收的声源信号时域波形的互相关函数，再从互相关函数找出最大点，该最大点对就的位置就是该麦克风对的时间延迟。通常情况下，目标发出的声音到达两个麦克风的时间是不同的，它们的差也就不为零。于是这个差值被称为目标到达麦克风对的时间差（TDOA），即时延，记为 τ 。首先介绍基本互相关法，通过互相关法可以比较时域波形的相关性，也就是相似程度。

假设麦克风收到的声音信号（经过离散采样）为：

$$x_1(n) = \alpha_1 s(n - \tau_1) + m_1(n) \quad (3-1)$$

$$x_2(n) = \alpha_2 s(n - \tau_2) + m_2(n) \quad (3-2)$$

其中 $x_1(n)$ 和 $x_2(n)$ 分别为麦克风1和麦克风2接收到的信号， τ_1 和 τ_2 是声源分别到达麦克风1和麦克风2的传播时间， $m_1(n)$ 和 $m_2(n)$ 是互不相关的高斯白噪声， $s(n)$ 为声源信号，同时， $s(n)$ 和 $m_1(n)$ 、 $m_2(n)$ 也是互不相关的， α_1 和 α_2 为声波的衰减系数（也就是说声音传播越久，能量就越越来越弱）。于是，定义 τ_{12} 为到达两个麦克风的时间差， $\tau_{12} = \tau_1 - \tau_2$ 。

信号 $x_1(n)$ 和 $x_2(n)$ 的互相关函数 $R_{12}(\tau)$ 可表示为：

$$R_{12}(\tau) = E[x_1(n)x_2(n - \tau)] \quad (3-3)$$

将式(3-1)、式(3-2)代入式(3-3)，可得：

$$R_{12}(\tau) = \alpha_1 \alpha_2 E[s(n - \tau_1)s(n - \tau_2 - \tau)] + \alpha_1 E[s(n - \tau_1)m_2(n - \tau)] + \alpha_2 E[s(n - \tau_2 - \tau)m_1(n)] + E[m_1(n)m_2(n)] \quad (3-4)$$

因 $s(n)$ 和 $m_1(n)$ 、 $m_2(n)$ 两两不相关，由式(3-4)，于是可得：

$$\begin{aligned} R_{12}(\tau) &= E[\alpha_1 \alpha_2 s(n - \tau_1)s(n - \tau_2 - \tau)] \\ &= \alpha_1 \alpha_2 R_s[\tau - (\tau_1 - \tau_2)] \end{aligned} \quad (3-5)$$

当 $\tau - (\tau_1 - \tau_2) = 0$ 时， $R_{12}(\cdot)$ 达到最大值，此时的 τ 就是时间延迟。因此就从 $x_1(n)$ 和 $x_2(n)$ 的互相关函数 $R_{12}(\tau)$ 中找出峰值点，该对应的 τ 值就是麦克风对的时间延迟 τ_{12} 。

从上面的公式中可知，利用传统的互相关法进行时间延迟估计运算量小，方法简单，然而该互相关法也有缺点^[24-25]：一是假设的噪声之间和信号与噪声之间相关性小，但是这在实际环境中是不存在的；二是由数学理论可知，在严格意义上的统计平均是不存在的，在实际应用到都是用时间平均来替代统计平均。除自身算法的限制外，各种环境噪声都会对相关函数有干扰，从而使 $R_{12}(\tau)$ 得到的峰值对应的不是真正的目标值。但是针对以上问题，为了提高时延估计的精度，许多科研工作人员经过长期的努力，提出了许多改进的算法，其中就有广义互相关时延估计法。

3.1.2 广义互相关时延估计法

经典的广义互相关算法是由 Knapp 和 Carter 提出。在没有外界干扰的理想情况下，相关函数的峰值只有一个，很容易区分出来。但在真实环境中，存在各种环境噪声。通常情况下，声音的传播过程相当复杂。所以求得的互相关函数的最大值不只一个，而且很难辨别哪个是目标所产生的。这造成了时延检测的困扰。为了保证估计的时间差精度足够高，应取尽量大的观测时间。卡特在 1946 年提出的经典的广义互相关函数法^[26]，能够有效地避开噪声和混响对时延的估计，从此以后，广义互相关法开始应用于各个领域。广义互相关算法（GCC）的主要思想就是将信号从时域转换到频域，计算信号的加权功率谱，使得峰值更加明显。互相关函数与互功率谱的关系如下：

$$\begin{aligned} R_{12}(\tau) &= \int_0^\pi X_1(\omega) X_2^*(\omega) e^{j\omega\tau} d\omega \\ &= \int_0^\pi G_{12}(\omega) e^{j\omega\tau} d\omega \end{aligned} \quad (3-6)$$

这其中 $X_1(\omega)$ 和 $X_2(\omega)$ 分别是 $x_1(n)$ 和 $x_2(n)$ 的傅立叶变换， $G_{12}(\omega)$ 是信号 $x_1(n)$ 和 $x_2(n)$ 的互功率谱。

接收到的结果中不只有需要的声源目标信号，还包括一些环境干扰，或者混响之类的。广义互相关法通过求出互功率谱并加权，然后白化信号、噪声，削弱噪声的干扰，放大声源目标信号强度大的频率成分，再反变换到时间域上，求得想要的广义互相关函数。因此，通常用如下式来表示广义互相关函数：

$$R_{12}^g(\tau) = \int_0^\pi \psi_{12}(\omega) G_{12}(\omega) e^{j\omega\tau} d\omega \quad (3-7)$$

其中 $\psi_{12}(\omega)$ 代表在频域内的加权函数， $R_{12}^g(\tau)$ 的峰值就是时间延迟 τ 。如图 3-1 的流程图显示了 GCC 的运算过程。

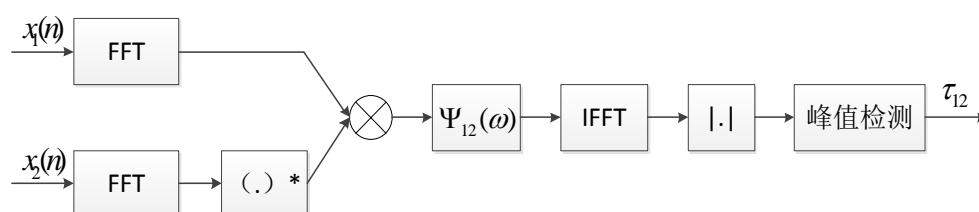


图 3-1 GCC 算法进行时间延迟估计流程图

表 3-1 广义互相关加权函数的优缺点

名 称	表达式	特性
CC	1	计算简单，但容易受到环境因素的影响。
ROTH	$\frac{1}{G_{12}(\omega)}$	可以对噪声较大的部分起来削弱作用，然而相关峰值不够尖锐。
SCOT	$\frac{1}{\sqrt{G_{11}(\omega)G_{22}(\omega)}}$	与 ROTH 方法相比，SCOT 同时对两个通道进行了分析。如果 $G_{11}(\omega) = G_{22}(\omega)$ ，SCOT 与 ROTH 功能相同，一样互相关峰不尖锐。但如果 $G_{11}(\omega) \neq G_{22}(\omega)$ 时，互相关函数并非标准的冲击响应，使得 TDE 的检测难度加大。
PHAT	$\frac{1}{ G_{12}(\omega) }$	PHAT 就是白化处理， $ G_{12}(\omega) $ 加权 $G_{12}(\omega)$ ，但可以看出，当 $ G_{12}(\omega) $ 的值比较小，整个式子会趋于无穷大，所以产生的误差就比较大；但是存在许多改进方法来对权函数进行纠正。
ML/HT	$\frac{1}{ G_{12}(\omega) } \cdot \frac{ \gamma_{12}(\omega) ^2}{[1 - \gamma_{12}(\omega) ^2]}$	$\gamma_{12}(\omega)$ 使得在 SNR 值偏高的部分权值较大，而在 SNR 值偏小的部分权值较小，所以对噪声有较强的抗干扰效果。理论上是一种最优的方法，但是它需要信号的先验知识，这也就限制了它的应用。

对于 $\psi_{12}(\omega)$ 的选择通常情况下，也没有统一的标准，所以选择上也是一个难点。如果 $\psi_{12}(\omega)$ 选择不当，得到的时延估计的误差就会比较大。所以在平时应用过程中，要综合分析各种 $\psi_{12}(\omega)$ 的**使用约束条件**，最终选择出性能最优的 $\psi_{12}(\omega)$ 。为此，将平时用到的一些 GCC 方法归纳如表 3-1 所示，下面主要对前四种方法进行分析。图 3-2 显示了 SNR=-10、0、10、20 时，表 3-1 中前四种互相关延时估计加权函数性能的仿真图，其中时间延迟为 0.018s。

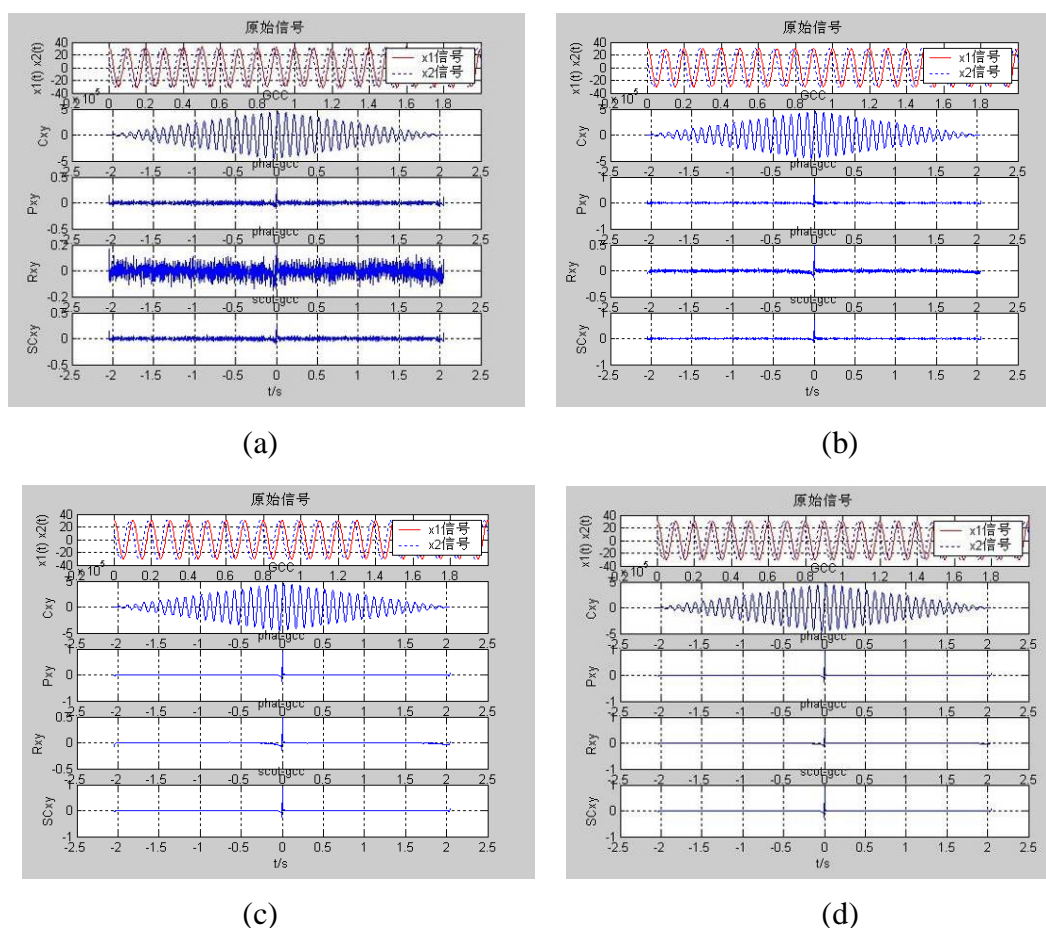


图 3-2 互相关延时估计加权函数性能分析

(a)SNR=-10dB; (b)SNR=0dB; (c)SNR=10dB; (d)SNR=20dB;

表 3-2 互相关延时估计加权函数运行时间比较

信噪比 SNR 互相关	-10dB	0dB	10dB	20dB
CC	0.030559s	0.033602s	0.031369s	0.029855s
PHAT-GCC	0.024260s	0.024468s	0.022588s	0.022433s
ROTH-GCC	0.024811s	0.021991s	0.020626s	0.023041s
SCOT-GCC	0.026194s	0.025038s	0.030971s	0.026718s

从图 3-2 和表 3-2 对比结果可得，从运算量上来分析，上述四种方法的运算量相差不大，基本一样；从峰值来分析，上述四种方法的峰值点都随 SNR 的降低而不明显，越不尖锐；从精度上来分析，上述四种方法的准确性都随 SNR 的降低而下降。当然它们在性质上也有属于自己的一些优于其它方法的特性：SCOT 对 SNR 的值特别敏感；GCC 和 RHOT 抗噪声效果比 SCOT 强，但是它的峰值在有干扰的情况下不是很明显。PHAT 方法，不仅在 SNR 高的时候估计精度高，而且在 SNR 低的时候也有较强的抗干扰能力。所以总的来说，PHAT 方法性能更优，可作为求时延的方法，但是也不是相当完善，因为 $\psi_{12}(\omega)$ 与 $G_{12}(\omega)$ 的倒数关系，PHAT 就是白化处理， $1/|G_{12}(\omega)|$ 加权 $G_{12}(\omega)$ ，但是可以看出，当 $|G_{12}(\omega)|$ 的值比较小，整个式子会趋于无穷大，所以产生的误差就比较大；但是存在许多改进方法来对权函数进行纠正。

综上所述，GCC 方法更倾向于比较强的声音信号的时延估计，同时，需要知道足够多的关于信号、噪声的先验信息，这样才能保证计算的精度足够高。所以它们都会受到信噪比 SNR 的干扰，从而使时延估计精度不高。尽管如此，GCC 计算量比其他方法小，而且实现起来容易的好处，使得它的应用相当广泛。本文将选用估计效果较好的 PHAT-GCC 算法，它能降解旁瓣并使互相关峰尖锐，抗干扰能力较强，估计精度较高，在实际系统应用中应用得最多的一种算法。

3.1.3 改进的 PHAT-GCC 定位算法

尽管 PHAT-GCC 能够获得精度很高的时延，有着较好的抗干扰能力。然而当所处的环境比较得复杂时，本算法求得的时延结果误差就比较大。于是又提出对 PHAT-GCC 方法作进一步的改进，从而使得估计值与真实值之间的误差减小，同时增加算法的抗噪性，使得该算法能应用于更广泛的场景和领域。

采用对加权函数 $\psi_{12}(\omega)$ 的倒数 $G_{12}(\omega)$ 加上一个 α ，于是可得新的加权函数，通常称之为指数相位互相关加权函数，于是新的加权函数可用如下形式表示：

$$\psi_{PHAT}(\omega) = \frac{1}{|G_{12}(\omega)|^\alpha}, 0 \leq \alpha \leq 1 \quad (3-8)$$

若 $\alpha=0$ 时， $\psi_{PHAT}(\omega)=1$ ，结果为标准广义互相关；若 $\alpha=1$ 时， $\psi_{PHAT}(\omega)=1/|G_{12}(\omega)|$ ，结果为基本互相关；若 $\alpha \in (0,1)$ 时，结果即为改进的 PHAT-GCC 算法。下面对三种 GCC 作一个比较，比较结果如表 3-3 所示：

表 3-3 三种 GCC 的比较

GCC 类别	α 的值	加权函数	性能比较
GCC	$\alpha = 0$	$\psi_{PHAT}(\omega) = 1$	考虑了全部谱信息，计算简单。
PHAT-GCC	$\alpha = 1$	$\psi_{PHAT}(\omega) = \frac{1}{ G_{12}(\omega) }$	没有考虑谱信息，结果是谱信息损失严重，计算量较大。
改进的 PHAT-GCC	$\alpha \in (0,1)$	$\psi_{PHAT}(\omega) = \frac{1}{ G_{12}(\omega) ^\alpha},$ $0 < \alpha < 1$	保留了一部分就谱信息，从而使得时延估计更加准确，计算量较大。

在实际应用过程中， α 的值也不是一成不变的，通常情况下， α 的值是随着环境中信噪比的变化而变化，即 α 随 SNR 而变化， α 的取值是通过实验经验总结出来的，它仅仅和空间环境的信噪比 SNR 有关。经过归纳，我们得出如下结论：当 SNR 小时， α 就小；当 SNR 大时， α 就大，总结 SNR 与 α 的对应关系如表 3-3 所示。由于本文所处的环境的 SNR $\in [10, 25]$ ，所以取 $\alpha = 0.75$ 。

表 3-4 信噪比 SNR 与参数 α 的对应关系

SNR / dB	α
SNR ≤ 10	$0.20 \leq \alpha \leq 0.55$ ，一般取 0.50
$10 < \text{SNR} \leq 25$	$0.55 < \alpha \leq 0.75$ ，一般取 0.75
$25 < \text{SNR}$	$0.75 < \alpha \leq 0.90$ ，一般取 0.85

除此之外，本文还对 PHAT-GCC 加权函数做了进一步的改进。如果声源信号的能量比较小时，式 (3-8) 的分母几乎为零，其结果是式 (3-8) 的权函数变成无穷大，为解决这个问题，于是采用相干函数来对分母进行加权，其定义如下：

$$\gamma_{12}^2(\omega) = \frac{|G_{12}(\omega)|^2}{G_1(\omega)G_2(\omega)} \quad (3-9)$$

很容易看出，当 $\gamma_{12}^2 = 0$ 时，表示两函数不相干；而当 $\gamma_{12}^2 = 1$ 时，表示两函数完全相干；相干函数的值在 0~1 之间时可能是由于噪声的干扰，而使得相干性降低。最终，得到更改后的加权函数为：

$$\psi_{PHAT}(\omega) = \frac{1}{|G_{12}(\omega)|^\alpha + |\gamma_{12}^2(\omega)|} \quad (3-10)$$

3.2 仿真

本节通过仿真比较 PHAT-GCC 和改进型 PHAT-GCC 这两种求时间延迟算法的性能。

3.2.1 仿真环境

设房间大小为 $3m \times 3m \times 2.5m$ ，根据 IMAGE 模型来产生房间脉冲响应。两个麦克风 M1 和 M2 分别位于 $(1.25m, 0.01m, 1.8m)$ 和 $(1.75m, 0.01m, 1.8m)$ 。选用目标声源为 1200ms 长、8KHz 采样的声音信号，目标位于 $(2.0m, 1.8m, 1.8m)$ 。在信噪比为 15dB, 5dB, 0dB, -5dB 和 -15dB 的情况下，对声音信号进行时延估计实验。

3.2.2 实验结果与分析

在通常情况下，要对一个系统进行评估的时候，都要引入一个评估标准，即估计结果与真实结果之间存在的误差。在本文中，引入时延估计精度作为评定算法优劣的评估标准，即真实时间差与估计时间差的误差大小评定。在大多数的应用场景中，用时延估计的标准差作为评定时延估计精度的标准。只要知道真实的值与目标的估计值，就可以计算出参量的均方差。标准差是用来反应实验结果相对真实时延值的偏离程度，偏离的越小，实验结果越准确；反之，实验结果误差越大。

估计偏差定义为：

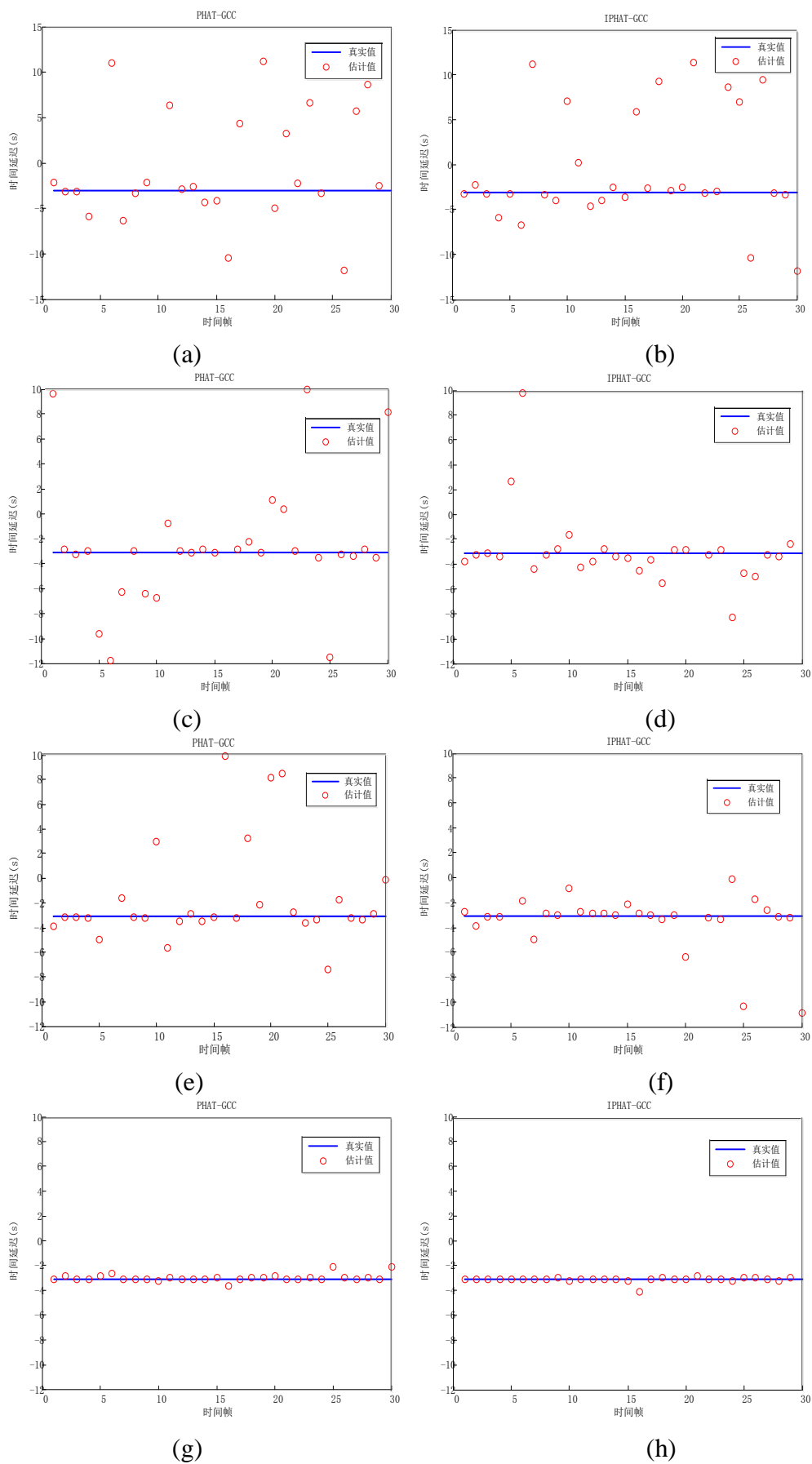
$$\hat{\sigma}_{\tau}(\omega) = \frac{1}{N} \sum_N \hat{\tau}(i) - \tau_0 \quad (3-11)$$

标准差的表达式：

$$Bias(\hat{\tau}) = \sqrt{\frac{1}{N} \sum_{n=0}^N (\hat{\tau}(i) - \tau_0)^2} \quad (3-12)$$

其中， $\hat{\tau}(i)$ 为第 i 个估计值， N 为总的估计的次数。

图 3-3 为 PHAT-GCC 和改进型 PHAT-GCC 求时延算法的估计值与真实值比较图，接收到的信号的信噪比 SNR 由 -15dB 到 15dB。



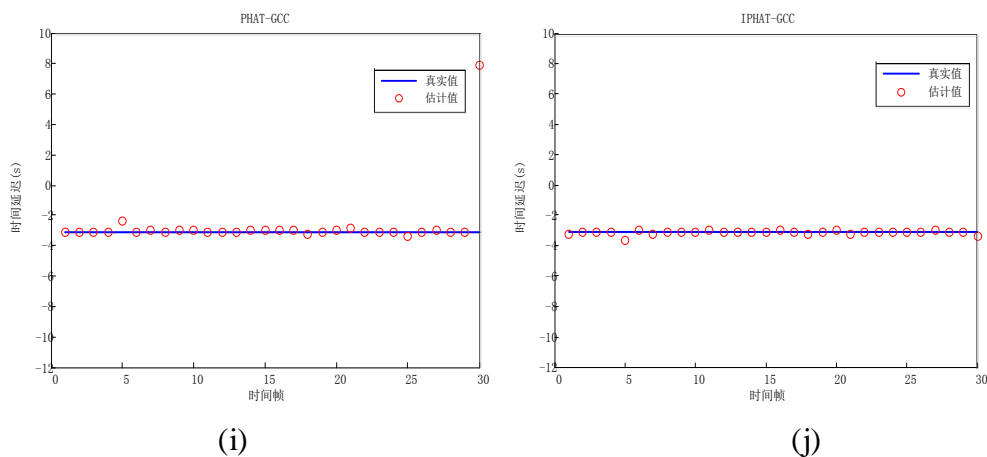


图 3-3 IPHAT-GCC 和 PHAT-GCC 结果比较

(a)(b)SNR=-15dB; (c)(d)SNR=-5dB; (e)(f)SNR=0dB; (g)(h)SNR=5dB; (i)(j)SNR=15dB;

比较图 3-3 左右两个图可知，改进的 PHAT-GCC 比 PHAT-GCC 计算效果好，当信噪比发生变化时改进的 PHAT-GCC 比 PHAT-GCC 性能更稳定。两种算法的 Bias 的统计结果如图 3-4 所示，图 3-5 是 SNR=10dB 时两种算法的 GCC 波形图。

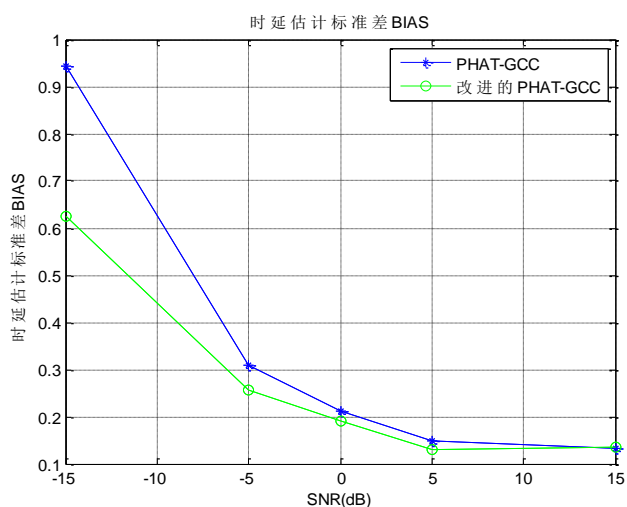


图 3-4 时延估计标准误差 (Bias)

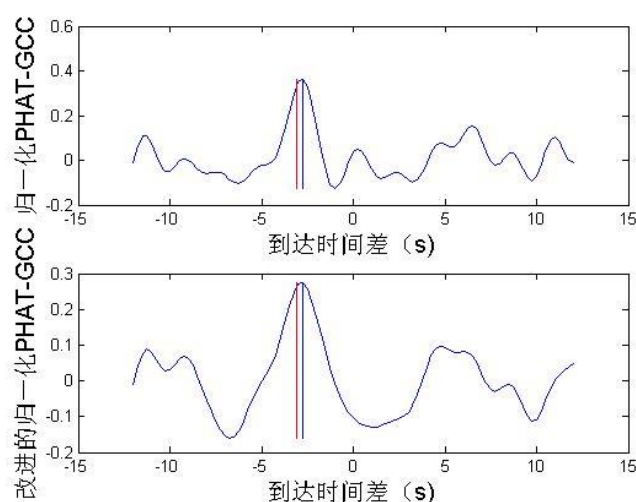


图 3-5 两种算法 GCC 波形图

从图 3-3、图 3-4 可得，信噪比好坏对算法的精度起着重要作用。高信噪比较高时，两种算法的性能都能达到满意的时延。但是当信噪比低时，两种算法的性能都下降明显。从图 3-5 可以看出，改进的 PHAT-GCC 方法能有效的抑制伪峰值，增强延迟处的峰值，而且改进的 PHAT-GCC 并没有大量增加算法的复杂度。本文选择改进的 PHAT-GCC 作为时延估计算法，时延估计好坏决定后续定位与跟踪性能的好坏，时延估计精度越高，定位跟踪性能越好；时延估计精度越低，定位跟踪性能越差。

3.3 本章小结

本章首先简单介绍了时延估计的物理意义以及几种传统的广义互相关时延估计算法，并对其进行了仿真，最终选择了 PHAT-GCC 作为本文的求时延估计的方法。并针对信噪比低的情况下 PHAT-GCC 时延估计算法估计时延不准确的问题，提出了改进的 PHAT-GCC 时延估计算法。通过大量仿真实验说明，改进的 PHAT-GCC 时延估计算法求得的互相关峰更加尖锐，降低了估计的误差。

第四章 多目标跟踪滤波算法研究与分析

在第三章，主要研究了时延估计算法，本章将对多目标跟踪算法进行研究。传统的多目标跟踪方法（如最近邻法（GNN）^[28]、概率数据关联法（PDA）^[29]以及多假设跟踪器法（MHT）^[30-31]等等）通过数据关联将多目标跟踪问题变成单目标跟踪问题，跟踪性能主要是取决于数据关联的效果，所以数据关联将对整个算法的性能起着决定性的作用，而且当目标数目多并存，再加上环境中的杂波和本身的存在虚警检测，使得关联本身处理起来就更加困难，运算量相当大，在一定的情况下呈指数增大等问题。于是，急需一种能够代替数据关联的方法，而概率假设密度滤波算法（PHD）应运而生，它有着严格数学理论推理，它适用于关联过程相对复杂的多目标跟踪问题（如群目标跟踪，密度问题跟踪等^[32-33]）。

PHD算法能够有效解决复杂环境下数据关联的问题，但算法本身拥有着复杂的积分运算，从而使其在解决实际问题时受到阻碍，这也就限制了PHD的应用和推广^[34]。为此，怎样解决实现PHD算法一直以来是很多研究人员都在从事的工作，例如，Vo等人证明在系统满足线性高斯模型时，PHD滤波递推过程存在闭合解。然而在实际许多应用模型中，系统是非线性非高斯的，PHD滤波递推过程不存在闭式解，于是Vo等人又提出PHD算法的粒子滤波实现^[35-36]，它采用蒙特卡罗的思想，利用大量的离散的带有权值的粒子集样本去逼近PHD函数，最终PHD函数用这些粒子近似表示，即SMC-PHD方法，它适用于非线性非高斯系统，有着广泛的适用性。正是通过许多学者不断的探索和研究，才使得PHD算法在各个领域的运用多了起来。

4.1 随机集的多目标跟踪理论

自从 Mahler 提出了基于随机集理论的概率假设密度滤波 (Probability Hypothesis Density Filter, PHDF)多目标跟踪算法^[37-39]后，该研究领域取得了一个重大的突破，不仅避免了数据的关联，还大大降低了计算量，它的核心思路是：将目标的状态以及传感器观测值用随机有限集表示，然后计算 PHD，从而把多目标跟踪问题转变成了单目标的跟踪问题。在真实环境情况下，元素及元素的个数提前都是不知道的或时变的，但是通常假设目标的个数也是有限的。于是，就可以用随机集的形式建立起状态模型以及观测模型。下面对随机集和相关理论进行研究学习。

4.1.1 随机集理论基础

4.1.1.1 随机集及其概率

所谓的随机集，与概率论中的随机变量相似，是一种推广，它所包含的元素是一个集合。但是它又与随机变量不同，它的元素个数和维数经常发变化，所以概率论中的定理或者结论在这里不再适用，在的实际应用过程中，目标的个数也即元素的个数是一定的且有限的，所以主要研究随机有限集。

首先，假设几个概念：设 E 为 R^n 上的有界闭子集， $B(E)$ 为 E 的全部子集的集合， Ω, A, P_r 为事件的概率空间。

于是可以得到随机有限集为映射 $E: \Omega \rightarrow B(E)$ 。

把概率论中随机变量中的一些重要结论在随机集的研究领域内进行扩展，可以得到随机集的重要结论。假设概率空间 Ω 上的概率测度用 P_r 表示，信任函数用 β_{Ξ} 表示：

$$\beta_{\Xi}(S) = P_r(\{\omega: \Xi(\omega) \subseteq S\}), \forall S \subseteq E \quad (4-1)$$

随机集的概率密度可以通过集导数求导得到， $F: B(E) \rightarrow [0, \infty)$ 是一个集类函数，于是得集导数 $\delta F / \delta x: B(E) \rightarrow [0, \infty)$ 如下示所示：

$$\frac{\delta F}{\delta x}(T) = \lim_{i \rightarrow \infty} \frac{F(T \cup \Delta_x^i) - F(T)}{\lambda(\Delta_x^i)} \quad (4-2)$$

式中， $x \in \Delta_x^i \subset B(E)$ ， $\lambda(\Delta_x^i)$ 是 Δ_x^i 的 Lebesgue 测度。

于是，在随机集 $X = \{x_1, \dots, x_M\}$ 上的集类函数就可以用递归的方法的推导如下：

$$\frac{\delta F}{\delta X}(T) = \frac{\delta^M F}{\delta x_1 \cdots \delta x_M}(T) \quad (4-3)$$

式中， M 代表求导重数，但是所得结果与求导次序没有直接关系。下面，将集导数满足的一些法则罗列如下：

- (1) 常数集导数为零：即在 $F = K$ (K 为一常数集合) 的条件下， $\frac{\delta F}{\delta X}(T) = 0$ 。
- (2) 数乘法则： $\frac{\delta(kF(T))}{\delta X} = k \frac{\delta F}{\delta X}(T)$ ， k 为一实数。
- (3) 加法法则： $\frac{\delta(F_1(T) + F_2(T))}{\delta X} = \frac{\delta F_1}{\delta X}(T) + \frac{\delta F_2}{\delta X}(T)$ 。
- (4) 乘积法则： $\frac{\delta(F_1(T)F_2(T))}{\delta X} = \sum_{W \subset X} \frac{\delta F_1}{\delta W}(T) \frac{\delta F_2}{\delta(X-W)}(T)$ 。

与之相反, 函数 $f = \frac{\delta F}{\delta X}(T)$ 的集积分可用表示为:

$$\int_S f(X) \delta X = \sum_{i=0}^{\infty} \frac{1}{i!} \int_S f(\{x_1, \dots, x_i\}) dx_1 \cdots dx_i \quad (4-4)$$

下面, 就将对集积分成立的法则罗列如下:

(1) 数乘法则: $\int_S (kf(X)) \delta X = k \int_S f(X) \delta X$, k 为一实数

(2) 加法法则: $\int_S (f_1(X) + f_2(X)) \delta X = \int_S f_1(X) \delta X + \int_S f_2(X) \delta X$

与函数微积分理论相似, 基于随机集的广义微积分的核心定理为:

$$f(X) = \frac{\delta F}{\delta X}(\emptyset) \Leftrightarrow \int_S f(X) \delta X \quad (4-5)$$

从该公式可以得出如下结论: 通过信任函数可以计算出真实的转移密度函数, 而利用集积分公式对多目标跟踪的信任函数作积分运算, 可以得到真实的似然函数。本节所述内容为下文讨论打下了良好的基础。

4.1.1.2 随机集概率假设密度

概率假设密度(PHD)与概率密度不是一个概率, 其表达的意义完全不同, 它概率论中学到的随机变量的期望非常的类似, 它表示一个集合的一阶矩。但是它又与随机变量中的数学期望是非常不同的, 与它的性质也完全不同: PHD 在 S 作积分结果就是该区域内目标数的期望值, 所以正是利用这一特性求得目标的个数, 其中 S 为侦察的状态空间。下面假设 X 为状态空间 S 上的集合, 则可以求得 X 在 S 中的元素个数, 记为 $N_X(S)$, 表达式如下:

$$N_X(S) = \sum_{x \in X} 1_x(S) \quad (4-6)$$

其中,

$$1_x(S) = \begin{cases} 1, & x \in S \\ 0, & x \notin S \end{cases} \quad (4-7)$$

目标个数的概率密度函数可以通过对式(4-6)求导得, 如下式:

$$f_X(y) = \sum_{x \in X} \delta_x(y) \quad (4-8)$$

对 $f_X(y)$ 在区域 S 上作积分运算, 即就为目标个数 $N_X(S)$ 。其中

$$\int_S \delta_x(y) = 1_x(y) \quad (4-9)$$

如果 v 代表已知的概率密度函数, 并且如果在区域 S 内求积分运算能得到目标个数的期望, 于是可得目标个数的期望值, 可以用下式表示:

$$\int_S v = E[N_X(S)] \quad (4-10)$$

那么，这个期望就是集合 X 有区域 S 上求得的，同时又满足下式：

$$E[N_X(S)] = E[\sum_{x \in X} 1_x(S)] \quad (4-11)$$

于是，对式(4-11)两边求导可得：

$$v = E[f_X(y)] = E[\sum_{x \in X} \delta_x(y)] = \sum_{x \in X} \int \delta_x(y) f(x) dx \quad (4-12)$$

4.1.2 随机集多目标跟踪模型

由于目标个数的不确定、目标出现时间的不确定性以及杂波的影响，使跟踪问题变得复杂。然而通过利用随机集可以使问题转化成简单的单目标跟踪问题，于是如何对多目标跟踪问题进行模型，将直接决定是否能有效对目标进行跟踪，同时也要保证跟踪模型要具有普遍适用性。为了阐述模型原理的方便，先以贝叶斯单目标为出发点，再推广到贝叶斯多目标跟踪模型。

4.1.2.1 Bayes 递归单目标跟踪模型

正如第 4.1 节介绍的一样，单目标的状态服从马尔可夫过程，定义 $f_{k|k-1}(\cdot|\cdot)$ 为概率转移密度， $g_k(\cdot|\cdot)$ 为观测空间 Z 的似然函数，于是可得 $k-1$ 时刻的状态 x_{k-1} 到 k 时刻状态 x_k 的概率转移密度为 $f_{k|k-1}(x_k | x_{k-1})$ ； k 时刻的观测 $z_k \in Z$ 的似然函数为 $g_k(z_k | x_k)$ 。

假设 $z_{1:k} = (z_1, \dots, z_k)$ 代表从起始时刻 1 到时刻 k 的所有量测集合，则当前时刻 k 的后验概率密度可表示为 $p_k(x_k | z_{1:k})$ 。 $p_k(\cdot | z_{1:k})$ 包含了时刻 k 所有状态信息。同时，使用最小均方误差（MMSE）和极大后验（MAP）准则来更新状态的值。给定初始的密度分布 p_0 ，根据 Bayes 理论，就可以计算每一时刻的后验概率密度分布。于是从式(4-13)和式(4-14)可以递归出 k 时刻的后验密度：

$$p_{k|k-1}(x_k | z_{1:k-1}) = \int f_{k|k-1}(x_k | \eta) p_{k-1}(\eta | z_{1:k-1}) d\eta \quad (4-13)$$

$$p_k(x_k | z_{1:k}) = \frac{g_k(z_k | x_k) p_{k|k-1}(x_k | z_{1:k-1})}{\int g_k(z_k | \lambda) p_{k|k-1}(\lambda | z_{1:k-1}) d\lambda} \quad (4-14)$$

4.1.2.2 Bayes 递推多目标跟踪模型

在多目标情况下，目标状态存在以下几种可能：第一种，当同一目标在当前时刻和后一时刻都出现的情况下，为目标存活；第二种，当目标在当时刻存活，在下一时刻不存在，称为目标消失；第三种，当目标在下一时刻出现，称为目标

产生，第四种，当目标由当前时刻的目标分裂产生，称为目标衍生。假设 $x_{k,1}, \dots, x_{k,M(k)}$ 代表所有 k 时刻状态向量， $M(k)$ 代表 k 时刻的状态集合中状态向量的个数； $z_{k,1}, \dots, z_{k,N(k)}$ 表示 k 时刻的量测值， $N(k)$ 代表 k 时刻有量测个数。现把当前时刻 k 的所有目标状态向量以及量测值用集合 $X_k = \{x_{k,1}, \dots, x_{k,M(k)}\}$ 和 $Z_k = \{z_{k,1}, \dots, z_{k,N(k)}\}$ 分别来表示。 X 中的向量的个数通过离散的概率分布来表示出来；集合中元素的联合分布可以通过数个合适的概率密度来组合表示出来。

多目标状态集合 RFS 模型是时变的，它不仅包含了目标的运动、目标出现，还包含有目标衍生和目标消失的过程，整个模型将一个目标从出现到消失包含在内。 X_{k-1} 代表 $k-1$ 时刻的多目标状态， x_{k-1} 为集合 X_{k-1} 中的一个状态向量，也就是目标，即： $x_{k-1} \in X_{k-1}$ ，用 $S_{k|k-1}(x_{k-1})$ 表示 k 时刻对于目标状态 x_{k-1} 的生灭状态，当 $S_{k|k-1}(x_{k-1}) = \{x_k\}$ 表示目标存活，用 $e_{k|k-1}(x_{k-1})$ 表示存活的概率，那么 x_{k-1} 到 x_k 的转移概率密度可由式 $f_{k|k-1}(x_k | x_{k-1})$ 给出；当 $S_{k|k-1}(x_{k-1}) = \emptyset$ 代表目标消失，在这一时刻不再继续出现，用 $1 - e_{k|k-1}(x_{k-1})$ 代表目标消失的有可能性大小。通常情况下， k 时刻假设所有的目标想到之间不相关，各自服从各自的运动模型。于是当前时刻 k 的多目标状态集合的 RFS 模型如下式所示：

$$X_k = [\bigcup_{x \in X_{k-1}} S_{k|k-1}(x)] \cup [\bigcup_{x \in X_{k-1}} B_{k|k-1}(x)] \cup \Gamma_k \quad (4-15)$$

$B_{k|k-1}(x)$ 和 Γ_k 分别表示衍生的目标和以前没有现在出现目标的随机有限集。

通常情况下，对于 $x_k \in X_k$ ，集合 $\theta_k(x_k)$ 为其对应的有限集，当 $\theta_k(x_k) = \{z_k\}$ ，意味着目标出现差被检测到， $p_D(x_k)$ 代表检测概率，用 $g_k(z_k | x_k)$ 代表检测到目标的概率密度；当 $\theta_k(x_k) = \emptyset$ 则目标，用 $1 - p_D(x_k)$ 表示漏检概率。于是 k 时刻量测集合的 RFS 模型可由下式表示：

$$Z_k = C_k \cup [\bigcup_{x \in X_k} \theta_k(x)] \quad (4-16)$$

C_k 表示 k 时刻的杂波和虚警。

由此可以得到利用贝叶斯方法下的多个目标递归方程。用 $p_k(\cdot | Z_{1:k})$ 表示多目标后验密度，则有

$$p_{k|k-1}(X_k | Z_{1:k-1}) = \int f_{k|k-1}(X_k | X) p_{k-1}(X | Z_{1:k-1}) dX \quad (4-17)$$

$$p_k(X_k | Z_{1:k}) = \frac{g_k(Z_k | X_k) p_{k|k-1}(X_k | Z_{1:k-1})}{\int g_k(Z_k | X) p_{k|k-1}(X | Z_{1:k-1}) dX} \quad (4-18)$$

通过将式(4-13)、(4-14)与式(4-17)、(4-18)进行对比，单目标模型其实是多目标模型的一种特殊情而已。然而，式(4-17)、(4-18)中存在积分运算，而这在实际

实用中是不可能实现的，计算量会随着目标数目增加而成指数级增加，因此，必须找到一种可靠的方法来解决这个问题。于是在 2001 年 Mahler 教授提出了随机集的 PHD 滤波算法，下面将对 PHD 滤波算法进行重点研究。

4.1.2.3 PHD 滤波算法

PHD 滤波器不用像贝叶斯算法中要推导后验概率密度，它是用后验概率密度 PHD 作为后验概率密度，只对其状态的一阶矩进行递推，其原理更加简单，最后再对目标的确状态和目标个数进行估计，有较小的计算量。PHD 滤波流程如图 4-1 所示：

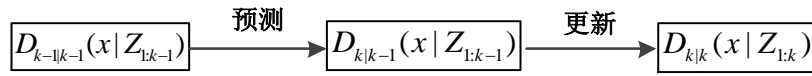


图 4-1 PHD 滤波算法

PHD 滤波算法必须满足以下条件：第一，杂波要服从泊松分布，与观测值之间相互独立；第二，提前预测的目标状态也满足泊松分布；第三，所在的目标想到之间相互独立，各自按自己的运动方式运动；第四，新生成的目标同样也要服从泊松分布。

结合以上假设，再通过一系列的推导和整理，可以得到 PHD 算法流程如下：

(1) 预测方程：如果假设 $f(\cdot|\cdot)$ 代表目标的转移概率密度； $b_{k|k-1}(\cdot|\xi)$ 表示目标在 $k-1$ 时刻衍生目标 $\beta_{k|k-1}(\{\xi\})$ 的 PHD； $\gamma_k(x)$ 表示 k 时刻新出现的目标 β_k 的 PHD； $e_{k|k-1}(\xi)$ 表示状态为 ξ 仍存在的概率。那么预测过程方程为：

$$D_{k|k-1}(x_k | Z_{1:k-1}) = \gamma_x(x) + \int (b_{k|k-1}(x | \xi) + e_{k|k-1}(\xi) f_{k|k-1}(x | \xi)) \cdots D_{k-1|k-1}(x_{k-1} | Z_{1:k-1}) dx_{k-1} \quad (4-19)$$

(2) 更新方程：如果， k 时刻的检测概率用 $P_{D,k}(x)$ 表示；一个目标的似然函数可以用 $g_k(\cdot|\cdot)$ 表示；杂波的平均数用 λ_k 表示；杂波的概率密度用 $c_k(z)$ 表示，则更新公式：

$$D_{k|k}(x | Z_{1:k}) = [1 - P_{D,k}(x) + \sum_{z \in Z_k} \frac{P_{D,k}(x) g_k(z | x)}{\lambda_k c_k(z) + \int P_{D,k}(x) g_k(z | x) D_{k|k-1}(x | Z_{1:k-1}) dx}] D_{k|k-1}(x | Z_{1:k-1}) \quad (4-20)$$

再对函数 $D_{k|k}$ 积分：

$$N_k = \int D_{k|k}(x | Z_{1:k}) dx \quad (4-21)$$

对 N_k 的值取整就可以得到 k 时刻的目标个数:

$$\hat{N}_k = [N_k]_{\text{int}} \quad (4-22)$$

同样, 对 PHD 进行递归时得不到解析解, 于是许多的研究者们又提出了将粒子滤波与 PHD 滤波相结合的 SMC-PHD 算法、以及将应用高斯混合的高斯 PHD 滤波算法, 用于解决“维数灾难”的问题。

4.2 PHD 的粒子滤波实现

针对 PHD 滤波算法中的积分的无法计算出闭式解, 根据粒子滤波的算法, 采用大量带权重的目标样本近似 PHD 函数, 这就是 SMC-PHD 滤波的思路^[40]。当粒子样本数量的足够多时, 粒子滤波接近于贝叶斯最优估计, 该方法不受系统模型的限制, 下面给出 SMC-PHD 实现的详细推导。

粒子 PHD 滤波算法包括: 初始化、预测、更新、重采样和状态估计, 下面将进行具体的研究:

第一: 初始化

在 $k=0$ 时刻, 目标的先验分布可以通过 L_0 个粒子 $\{w_0^{(i)}, x_0\}_{i=1}^{L_0}$ 来表征, 而 L_0 的值按目标数目的估计值成比例增加或者减少, 假设有 \hat{N}_0 个目标, 则 $w_0^{(i)} = \hat{N}_0 / L_0$ 。此时, 对 PHD 函数 D_0 采样可得样本 $x_0^{(i)}$, 其中 $i=1, \dots, L_0$ 。

函数 D_0 表示为:

$$D_0(x_k) \approx \sum_{i=1}^{L_0} w_0^{(i)} \delta_{x_k^{(i)}}(x_k) \quad (4-23)$$

第二: 预测

在 $k \geq 1$ 时, 给定重要性概率密度函数 q_k 和 p_k , 于是对这 q_k 和 p_k 进行采样可得采样样本。对从 $q_k(\cdot | x_{k-1}^{(i)}, Z_k), i=1, \dots, L_{k-1}$ 可得采样样本 $x_k^{(i)}$, 于是每个粒子的相应权值可计算如下:

$$\tilde{w}_{k|k-1}^{(i)} \approx \frac{\phi_{k|k-1}(\tilde{x}_k^{(i)}, x_{k-1}^{(i)})}{q_k(\tilde{x}_k^{(i)} | x_{k-1}^{(i)}, Z_k)} w_{k-1}^{(i)}, i=1, \dots, L_{k-1} \quad (4-24)$$

对于新出现目标, 就对 $q_k(\cdot | Z_k)$ 进行采样, 就可得到新的采样样本 $\tilde{x}_k^{(i)}, i=N_{k-1}+1, \dots, L_{k-1}+J_k$, 于是相应的权值计算式可表示为:

$$\tilde{w}_{k|k-1}^{(i)} \approx \frac{1}{J} \frac{\gamma_k(\tilde{x}_k^{(i)})}{p_k(\tilde{x}_k^{(i)} | Z_k)}, i=L_{k-1}+1, \dots, L_{k-1}+J_k \quad (4-25)$$

于是, 预测 $D_{k|k-1}$ 可表示如下:

$$D_{k|k-1}(x_k) = \begin{cases} \sum_{i=1}^{L_{k-1}} \tilde{w}_{k|k-1}^{(i)} \delta_{\tilde{x}_k^{(i)}}(x_k), i=1, \dots, L_{k-1} \\ \sum_{i=1}^{L_{k-1}+J_k} \tilde{w}_{k|k-1}^{(i)} \delta_{\tilde{x}_k^{(i)}}(x_k), i=1, \dots, L_{k-1} + J_k \end{cases} \quad (4-26)$$

第三：更新

利用传感器获得的观测向量，再结合相应的似然函数 $g(\cdot|\cdot)$ 就可以对权值进行更新。对于观测向量 $z \in Z_{1:k}$ ，其对应可得 $C_k(z) = \sum_{i=1}^{L_{k-1}+J_k} \psi_{k,z}(\tilde{x}_k^{(i)}) w_{k|k-1}^{(i)}$ ，于是得到更新的粒子权重为如下式所示：

$$\tilde{w}_k^{(i)} = \left[1 - P_{D,k} + \sum_{z \in Z_k} \frac{\psi_{k,z}(\tilde{x}_k^{(i)})}{\kappa_k(z) + C_k(z)} \right] w_{k|k-1}^{(i)} \quad (4-27)$$

于是可得：

$$D_k(x_k) \approx \sum_{i=1}^{L_{k-1}+J_k} \tilde{w}_k^{(i)} \delta_{\tilde{x}_k^{(i)}}(x_k) \quad (4-28)$$

第四：重采样

可通过下式可以估计出目标的个数：

$$\tilde{N}_k(x_k) = \sum_{i=1}^{L_{k-1}+J_k} \tilde{w}_k^{(i)} \quad (4-29)$$

然后归一化重采样粒子集 $\{\tilde{x}_k^{(i)}, \tilde{w}_k^{(i)} / \tilde{N}_k\}_{i=1}^{L_{k-1}+J_k}$ ，得到粒子集 $\{x_k^{(i)}, w_k^{(i)} / \tilde{N}_k\}_{i=1}^{L_k}$ ，再进行恢复权重并将权值乘以 \hat{N}_k ，最后得到重采样的粒子集 $\{x_k^{(i)}, w_k^{(i)}\}_{i=1}^{L_k}$ 。

第五：状态估计

PHD 函数确定如下：

$$D_k(x_k) \approx \sum_{i=1}^{L_k} \tilde{w}_k^{(i)} \delta_{\tilde{x}_k^{(i)}}(x_k) \quad (4-30)$$

将所在粒子的权值进行求和，就可以得到目标数目的估计值，结果如下：

$$\tilde{N}_k(x_k) = \sum_{i=1}^{L_k} \tilde{w}_k^{(i)} \quad (4-31)$$

4.3 PHD 的粒子滤波仿真

假设观测区域内有目标数为 3，目标的运动方程为：

$$x_k = F_{k-1}x_{k-1} + G_{k-1}w_{k-1}$$

量测方程为:

$$y_k = H_k x_k + v_k \quad (4-32)$$

三个目标的初始位置分别为: 目标1位置初始为(400;5;400;5), 在 $k=2$ 时刻出现, 在 $k=40$ 时刻消失; 目标2位置初始为(440;-5;440;5), 在 $k=2$ 时刻出现, 在 $k=30$ 时刻消失。目标3位置初始为(350;5;350;-5), 在 $k=11$ 时刻出现, 在 $k=50$ 时刻消失。目标检测概率 $P_d=0.95$, 目标存活概率 $P_s=0.95$, 杂波泊松分布参数 $\lambda=10$, 仿真所花费的时间 $T=50$, $Q=\text{diag}([0.05,0.3,0.02,0.3])$, 迭代时间步长 $\Delta T=1$, 最后对于矩阵 F_{k-1} 、 G_{k-1} 和 H_k 为:

$$F_{k-1} = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad G_{k-1} = \begin{bmatrix} T^2/2 & 0 \\ T & 0 \\ 0 & T^2/2 \\ 0 & T \end{bmatrix}, \quad H_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (4-33)$$

图 4-2 是每个目标的采样粒子数 $N_p=500$ 时, 新生目标的粒子数为 $N_b=100$, 目标状态的估计值与真实值(红色圆号代表目标真实值, 蓝色星号代表目标的估计值), 图 4-3 是经过 100 次蒙特卡罗运算后目标个数的真实值与估计值(红色代表目标真实值, 蓝色代表目标估计值), 其中 N_p 分别为 200, 400, 600, 1000。从图 4-3 可以知道, 随着每个目标的采样粒子数 N_p 的增加, 估计精度越好, 但是过大的 N_p 会使得计算的时间快速增长, 通常取 $N_p=500$, 表 4-1 显示了当 N_p 和 N_b 取不同值时的运行时间, 蒙特卡次数为 100。

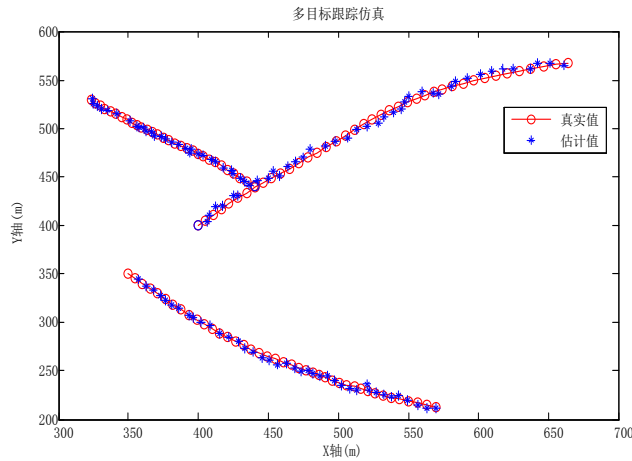


图 4-2 PHD 滤波算法

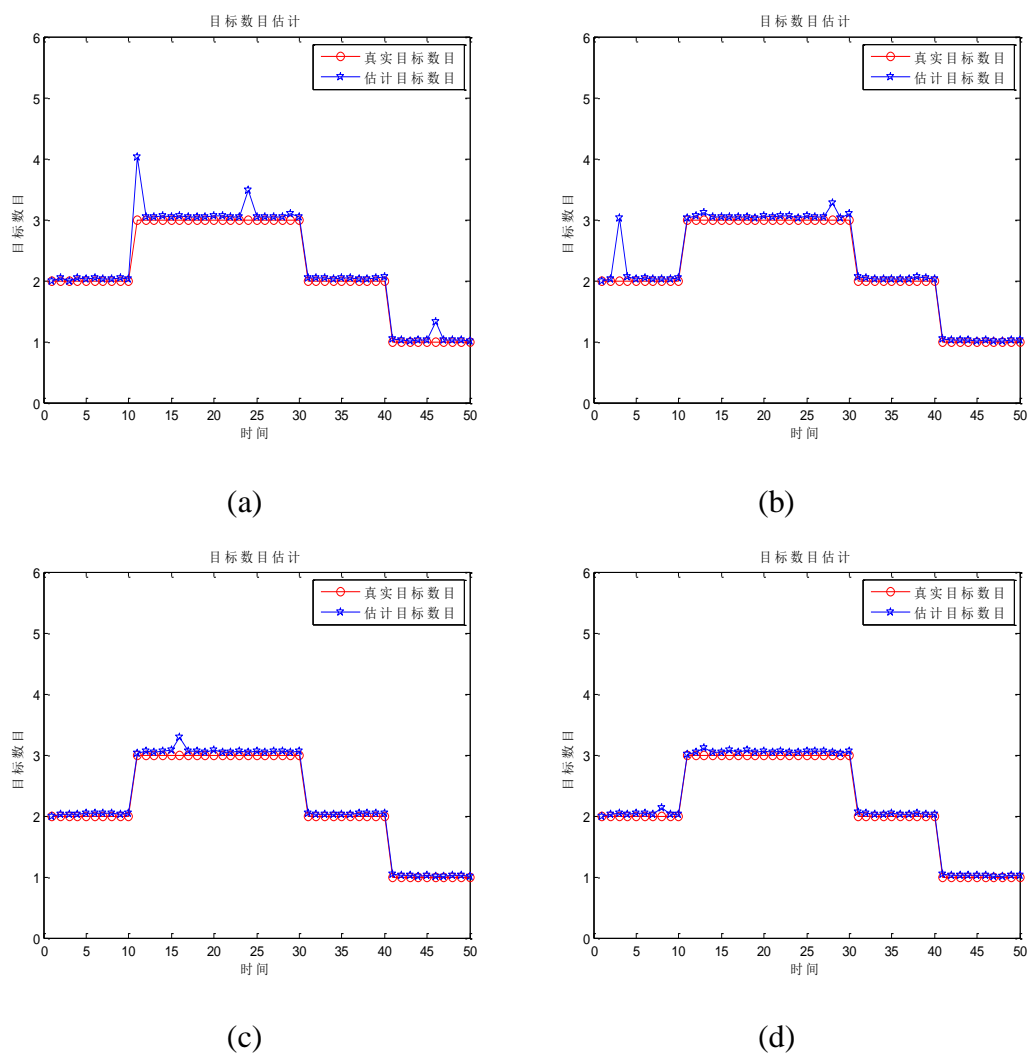
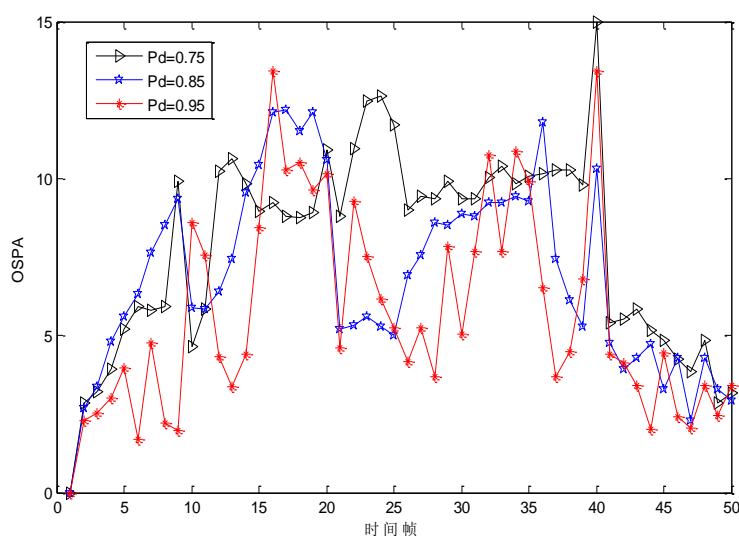


图 4-3 PHD 滤波算法目标数目估计

(a) $N_p = 200$; (b) $N_p = 400$; (c) $N_p = 600$; (d) $N_p = 1000$;

表 4-1 不同 N_p 和 N_b 值下运行时间比较

$N_p \backslash N_b$	50	100	150	200
200	21003.9321s	23040.6124s	23916.4413s	26019.9063s
400	39543.2289s	42951.6077s	43001.5675s	46542.7001s
600	57383.0971s	60345.7780s	62452.4207s	65024.2785s
1000	97641.8781s	105595.9596s	119249.7805s	136455.0457s

图 4-4 不同目标检测概率 P_d 下目标跟踪性能比较

从表 4-1 可以看出, N_p 和 N_b 取值不同时运行时间也不同, 而且当这两个指标都不断增大时, 运算时间随着 N_p 和 N_b 的增加而增加。图 4-4 为在不同目标检测概率 P_d 条件下目标跟踪性能比较, 由图 4-4 可以看出, 在目标存活概率 $P_s = 0.95$, 目标检测概率 P_d 不同时, 当 P_d 的值增大时, 目标的跟踪效果就提高。这是由于 P_d 值大时, 目标的漏检的概率就下降了, 从而提高了目标的跟踪性能。

4.4 本章小结

本章主要介绍了随机集及概率假设密度函数的基本知识, 在回顾 Bayes 递归单目标跟踪模型的基础上, 介绍了基于随机集的 Bayes 递归多目标跟踪模型, 以及 PHD 滤波器。并针对 PHD 滤波算法积分运算没有闭式解的问题, 研究了 SMC-PHD 滤波算法。通过仿真, 验证了 SMC-PHD 滤波算法对多目标定位与跟踪的有效性, 并分析了在不同目标检测概率下目标的跟踪效果。

第五章 基于声源信息的多目标定位与跟踪研究与实现

虽然基于随机集的 Bayes 递归多目标跟踪模型在雷达、图像、红外等研究领域已经有了很成熟的体系，但将随机集的 Bayes 递归多目标跟踪用于声源多目标跟踪领域还比较少。而且声源多目标跟踪问题面临的环境更加复杂性（如在房间内，存在电脑、空调等产生的噪声，同时还存在房间内桌椅板凳的反射）、目标本身的复杂多变（目标出现的时刻以及目标消失的时刻都具有不可遇见性），使得 Bayes 递归多声源目标模型更加复杂。为了体现由简单到复杂的原则，首先讨论单声源目标跟踪模型，然后讨论多声源目标跟踪模型。

5.1 单声源目标跟踪模型

对于单个声源目标，采用标准的状态空间模型^[41-42]，假设 α_k 代表第 k 个时刻声源的位置， $x_k = [\alpha_k^T, \phi_k^T]^T$ 代表状态变量，其中 ϕ_k 代表某些运动变量，比如加速度，速度等。那么目标的动态运动过程如下：

$$x_k = Ax_{k-1} + Bw_k \quad (5-1)$$

其中 A 和 B 是特定的矩阵， w_k 是时间不相关的高斯向量，其均值为零，方差为 I 。在声源定位与跟踪应用场景中，常采用 Langevin 方程来模拟目标声源的运动过程^[50]，其中， ϕ_k 是由 X 方向和 Y 方向的速度组成。于是，状态变量方程可表示如下：

$$\alpha_k = \alpha_{k-1} + T\phi_{k-1} \quad (5-2)$$

$$\phi_k = e^{-\eta T} \phi_{k-1} + \nu \sqrt{1 - e^{-2\eta T}} w_k \quad (5-3)$$

其中， η 和 ν 为模型参数。

下面，定义 z_k 代表麦克风对在时刻 k 测得的 TDOA 值，那么测得的 TDOA 值可表示为：

$$z_k = \tau(Cx_k) + v_k \quad (5-4)$$

其中 $Cx_k = \alpha_k$ ， $C = [I \ 0]$ ， v_k 是时间不相关的噪声。假设每对麦克风测得的噪声是不相关的，其服从均值为零，方差为 σ_v^2 的高斯分布。

定义 $f(x_k | x_{k-1})$ 代表状态转移密度， $g(z_k | x_k)$ 代表似然函数。通常 $f(x_k | x_{k-1}) = N(x_k; Ax_{k-1}, BB^T)$ ， $g(z_k | x_k) = N(z_k; \tau(Cx_k), \sigma_v^2)$ ， $N(\cdot; \mu, P)$ 代表均值为 μ ，方差为 P 的高斯密度函数。

为了估计声源的状态 x_k ，首先定义观测值集合 $z_{1:k} = \{z_i; i=1, \dots, k\}$ ，于是后验概率密度 $p(x_k | z_{1:k})$ 可通过贝叶斯递归实现如下：

$$p(x_k | z_{1:k-1}) = \int f(x_k | x_{k-1}) p(x_{k-1} | z_{1:k-1}) dx_{k-1} \quad (5-5)$$

$$p(x_k | z_{1:k}) = \frac{g(z_k | x_k) p(x_k | z_{1:k-1})}{\int g(z_k | x_k) p(x_k | z_{1:k-1}) dx_k} \quad (5-6)$$

5.2 多声源目标跟踪模型

本节将单声源目标跟踪模型进行扩展，延伸到多声源目标状态估计。通过应用第四章介绍的随机有限集的相关知识，建立起多声源目标的状态空间模型。

5.2.1 多声源目标状态空间模型

多声源目标场景满足下列特点：第一，每个声源目标服从 Langevin 运动方程；第二，整个房间墙面存在反射。在随机有限集的基础上，通过一个有限集定义目标的状态变量为：

$$\mathcal{X}_k = \{x_{1,k}, \dots, x_{N_k,k}\} \quad (5-7)$$

这里， N_k 是 k 时刻活动的声源目标个数，假设 $N_k \leq N_{\max}$ 。

对于一对麦克风对，用有限集定义 k 时刻 TDOA 观测值集如下：

$$\mathcal{Z}_k = \{z_{1,k}, \dots, z_{M_k,k}\} \quad (5-8)$$

这里， M_k 代表麦克风对在 k 时刻测得的 TDOA 个数。

对于声源多目标情况，还面临如下问题：

(1) 对于 \mathcal{X}_k 中任意一个 $x_{i,k}$ ，期望在 \mathcal{Z}_k 中有一个观测值由 $x_{i,k}$ 产生，也可能有多个观测 $x_{i,k}$ 产生，然而我们并不知道是哪一个。

(2) 对于 \mathcal{X}_k 中任意一个 $x_{i,k}$ ，有可能在 \mathcal{Z}_k 中没有一个观测值由 $x_{i,k}$ 产生，这可能是由于噪声的影响或者信号太弱。

(3) 在 \mathcal{Z}_k 中，有可能包含错误的 TDOA，他们可能是由于噪声产生或者由混响生成的。

考虑到以上问题，于是建立起了随机有限集模型，它能很好地处理以上的多目标多观测值问题。这里，多目标随机有限集 \mathcal{X}_k 被重新定义为：

$$\mathcal{X}_k = B_k(b_k) \cup \left\{ \bigcup_{i=1, \dots, |\mathcal{X}_{k-1}|} S_k(x_{i,k-1}, w_{i,k}) \right\} \quad (5-9)$$

其中， $B_k(b_k)$ 包含了所有在 k 时刻新出现声源目标的状态变量， $S_k(x_{i,k-1}, w_{i,k})$ 是

一个随机集包含先前声源目标的状态变量 $x_{i,k-1}$ ， $w_{i,k}$ 和 b_k 是随机变量。对于 $S_k(x_{i,k-1}, w_{i,k})$ ，于是作如下假设：

$$S(x_{i,k-1}, w_{i,k}) = \begin{cases} \phi, & H_{death} \\ \{Ax_{i,k-1} + Bw_{i,k}\}, \bar{H}_{death} \end{cases} \quad (5-10)$$

H_{death} 和 \bar{H}_{death} 分别代表死亡和没有死亡的假设。对于没有死亡的假设，状态变量过程和单声源目标处理过程一样， H_{death} 发生的概率为 P_{death} 。对于出生的情况，假设在一个时刻只有一个目标生成，当 $|\chi_{k-1}| = N_{max}$ 时， $B_k = \theta$ ；其他情况有：

$$B_k(b_k) = \begin{cases} \phi, & \bar{H}_{birth} \\ \{b_k\}, H_{birth} \end{cases} \quad (5-11)$$

H_{birth} 和 \bar{H}_{birth} 分别代表出生和没有出生的假设。 b_k 是一个初始化状态变量，服从高斯分布， H_{birth} 发生的概率为 P_{birth} 。

观测模型对于每个 Z_k ，被定义如下：

$$Z_k = \left\{ \bigcup_{i=1, \dots, |\chi_k|} T_k(x_{i,k}, v_{i,k}) \right\} \cup C_k \quad (5-12)$$

这里， C_k 是测得错误的 TDOA 观测值，其个数服从泊松分布。 T_k 的表达式如下：

$$T_k(x_{i,k}, v_{i,k}) = \begin{cases} \phi, & H_{miss} \\ \{\tau(Cx_{i,k}) + v_{i,k}\}, \bar{H}_{miss} \end{cases} \quad (5-13)$$

其中， H_{miss} 和 \bar{H}_{miss} 分别为未测到的假设和测到的假设。 H_{miss} 发生的概率为 P_{miss} 。

为了估计声源的状态 χ_k ，首先定义 $Z_{1:k} = \{Z_i; i=1, \dots, k\}$ ，于是后验概率密度 $p(\chi_k | Z_{1:k})$ 可通过贝叶斯递归实现如下：

$$p(\chi_k | Z_{1:k-1}) = \int f(\chi_k | \chi_{k-1}) p(\chi_{k-1} | Z_{1:k-1}) d\chi_{k-1} \quad (5-14)$$

$$p(\chi_k | Z_{1:k}) = \frac{g(Z_k | \chi_k) p(\chi_k | Z_{1:k-1})}{\int g(Z_k | \chi_k) p(\chi_k | Z_{1:k-1}) \mu d\chi_k} \quad (5-15)$$

求解以上两个方程不是一件容易的事，于是采用粒子滤波的方法来实现。

5.2.2 多声源目标跟踪实现算法研究

标准的 SMC-PHD 滤波方法虽然能估计出目标状态和目标个数，但并没有对目标轨迹进行关联，因此如何对目标进行关联，一直以来都是大家研究又一热点。目标已有的关联方法有：PHD-with-Association^[43]、MHT-with-PHD clutter 滤波器

^[44]。PHD-with-Association 方法先用 SMC-PHD 估计出的目标状态和个数结果，然后通过轨迹关联算法完成对目标的轨迹关联；而 MHT-with-PHD clutter 方法通过 SMC-PHD 滤除观测集合中的杂波，再采用多假设跟踪（MHT）对目标轨迹进行关联。但上述两种方法在实际应用满足不了对精度的要求。因此，本文采用了一种加粒子标签的多目标状态估计办法，其思路是：为每个粒子添加身份标签，然后根据标签的不同，对粒子进行分类，从而达到实现目标的分辨能力，最后实现目标轨迹关联。由于此方法中的目标状态估计和目标个数估计不受粒子标签的影响，因而性能更优。其算法如下：

对于 $k-1$ 时，粒子为 $\{w_{k-1}^{(i)}, x_{k-1}^{(i)}, j_{k-1}^{(i)}\}_{i=1}^{L_{k-1}}$ ，每个粒子对应的标签为 $\{j_{k-1}^{(i)}\}_{i=1}^{L_{k-1}}$ ，则下一时刻 k 的粒子及标签可以通过以下公式求得：

(1) 预测

$$\tilde{x}_k^{(i)} = \begin{cases} q_k(\cdot | x_{k-1}^{(i)}, Z_k), & i = 1, \dots, L_{k-1} \\ p_k(\cdot | Z_k), & i = L_{k-1} + 1, \dots, L_{k-1} + J_k \end{cases} \quad (5-16)$$

从而可得预测权重如下：

$$\tilde{w}_{k|k-1}^{(i)} = \begin{cases} \frac{\phi(\tilde{x}_k^{(i)}, x_{k-1}^{(i)}) w_{k-1}^{(i)}}{q(\tilde{x}_k^{(i)} | x_{k-1}^{(i)}, Z_k)}, & i = 1, \dots, L_{k-1} \\ \frac{\gamma(\tilde{x}_k^{(i)})}{J_k p_k(\tilde{x}_k^{(i)} | Z_k)}, & i = L_{k-1} + 1, \dots, L_{k-1} + J_k \end{cases} \quad (5-17)$$

采样后的粒子标签如下：

$$\tilde{j}_k^{(i)} = \begin{cases} j_{k-1}^{(i)}, & i = 1, \dots, L_{k-1} \\ 0, & i = L_{k-1} + 1, \dots, L_{k-1} + J_k \end{cases} \quad (5-18)$$

(2) 更新

对于每一个粒子更新权值：

$$\tilde{w}_k^{(i)} = [\nu(\tilde{x}_k^{(i)}) + \sum_{z \in Z_k} \frac{\psi_{k,z}(\tilde{x}_k^{(i)})}{\kappa_k(z) + C_k(z)}] \tilde{w}_{k|k-1}^{(i)} \quad (5-19)$$

其中， $z \in Z_k$ ， $C_k(z) = \sum_{j=1}^{L_{k-1}+J_k} u_k(\tilde{x}_k^{(j)}, z)$ ， $u_k(\tilde{x}_k^{(j)}, z) = \psi_{k,z}(\tilde{x}_k^{(j)}) \tilde{w}_{k|k-1}^{(j)}$ 。

(3) 估计目标状态和个数

(4) 重采样 $\{w_k^{(i)} / \tilde{N}_{k|k}, x_k^{(i)}, j_k^{(i)}\}_{i=1}^{L_{k-1}+J_k}$ ，得到粒子 $\{w_k^{(i)} / \tilde{N}_{k|k}, x_k^{(i)}, j_k^{(i)}\}_{i=1}^{L_k}$ ；每个子粒子的标签与父粒子的标签是一样的；再利用因子 $\tilde{N}_{k|k}$ 将粒子权重放大，得到粒子 $\{w_k^{(i)}, x_k^{(i)}, j_k^{(i)}\}_{i=1}^{L_k}$ 。

图 5-1 为在相同条件下标准的 SMC-PHD 与改进的 SMC-PHD 算法性能的比较,从图中可以看出,在多目标情况下,改进的 SMC-PHD 比标准的 SMC-PHD 跟踪性能好,其中蒙特卡次数为 100。

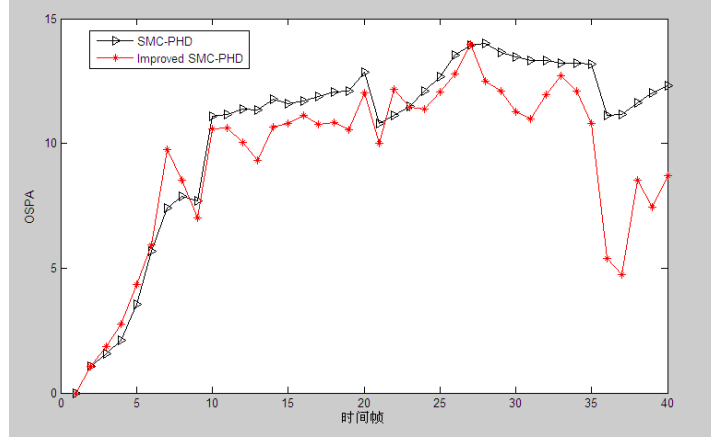


图 5-1 SMC-PHD 与改进的 SMC-PHD 性能比较

5.3 多传感器 PHD 滤波算法研究与仿真

通常情况下,都认为麦克风接收到的信号是来自于直达波,于是,在这种情况下,要想实现目标的定位与跟踪,必须用到两对或者两对以上的麦克风。在该模式下,定义 $z_k^{[q]}$ 代表第 q 对麦克风在时刻 k 测得的 TDOA 值,那么测得的所有 TDOA 值可表示为:

$$z_k^{[q]} = \tau_q(Cx_k) + v_k^{[q]}, \quad q = 1, \dots, Q \quad (5-20)$$

其中 Q 为麦克风的对数, $Cx_k = \alpha_k$, $C = [I \ 0]$, $v_k^{[q]}$ 是时间不相关的噪声,其服从均值为零,方差为 σ_v^2 的高斯分布。

那么目标的后验概率密度 $p(x_k | z_{1:k}^{[1:Q]})$, 可通过下面两个式实现:

$$p(x_k | z_{1:k-1}^{[1:Q]}) = \int f(x_k | x_{k-1}) p(x_{k-1} | z_{1:k-1}^{[1:Q]}) dx_{k-1} \quad (5-21)$$

$$p(x_k | z_{1:k}^{[1:Q]}) = \frac{\prod_{q=1}^Q g_q(z_k^{[q]} | x_k) p(x_k | z_{1:k-1}^{[1:Q]})}{f \prod_{q=1}^Q g_q(z_k^{[q]} | x_k) p(x_k | z_{1:k-1}^{[1:Q]}) dx_k} \quad (5-22)$$

多传感器 PHD 滤波算法(SMC-MS-PHD)的具体实现步骤与标准 PHD 相比,预测过程相同,更新过程不同。在这里假设粒子数为 L , 对于粒子 $i=1, \dots, L$, 可以从状态转移函数 $f(x_k^i | x_{k-1}^i)$ 得到下一时刻 k 的状态,然后再利用似然函数求得目标的权值更新过程如式(5-23),其它过程都与标准的 PHD 算法相同。

$$w_k^{(i)} = \prod_{q=1}^Q g_q(z_k^{[q]} | x_k^{(i)}) w_{k-1}^{(i)} \quad (5-23)$$

下面将利用上面的算法对房间内的两个目标进行定位和跟踪。

实验在房间内进行，房间的大小为 $3\text{m} \times 3\text{m} \times 2.5\text{m}$ ，整个房间的冲击响应用 IMAGE 的房间混响模型模拟，假设目标为两个在平面内做缓慢运动的声源（轨迹如图 5-1），分别为男声和女声：目标 1 位置初始为 $(1.0, 1.0, 1.8)$ ，在 $k=1$ 时刻出现，在 $k=30$ 时刻消失（如图 5-2 中实线）；目标 2 位置初始为 $(2.5, 0.5, 1.8)$ ，在 $k=21$ 时刻出现，在 $k=45$ 时刻消失（如图 5-2 中虚线）。4 对麦克风位置如图 5-1 所示，麦克风对 1 $(1.25, 0.01, 1.8)$ 、 $(1.75, 0.01, 1.8)$ ，麦克风对 2 $(0.01, 1.25, 1.8)$ 、 $(0.01, 1.75, 1.8)$ ，麦克风对 3 $(1.25, 2.99, 1.8)$ 、 $(1.75, 2.99, 1.8)$ ，麦克风对 4 $(2.99, 1.25, 1.8)$ 、 $(2.99, 1.75, 1.8)$ 。

声音在空气中的传播速度为 $c = 344\text{m/s}$ ，设房间的四个顶点坐标分别为 $O(0, 0, 1.8)$ 、 $A(0, 3, 1.8)$ 、 $B(3, 3, 1.8)$ 、 $C(3, 0, 1.8)$ 。一般房间有六个墙壁， β_{x1} 、 β_{x2} 、 β_{y1} 、 β_{y2} 、 β_{z1} 、 β_{z2} 分别是各面墙的反射系数，根据实验经验，通常设置对应的值为 $(0.85, 0.85, 0.25, 0.25, 0.25, 0.25)$ ， (x, y, z) 是声源位置的坐标， $L_x = 3$ 、 $L_y = 3$ 、 $L_z = 2.5$ 分别是房间的长、宽、高。选用的帧长度是 1024，采样频率为 8KHz，信号长度是 7.68s，信噪比为 15dB，目标个数为 1， $\beta = 10\text{s}^{-1}$ 和 $v = 1\text{ms}^{-1}$ 为模型参数， $P_{\text{birth}} = 0.05$ ， $P_{\text{death}} = 0.01$ ， $P_{\text{miss}} = 0.25$ ，每个目标的采样粒子数 $N_p = 500$ ，新生目标的粒子数为 $N_b = 100$ 。

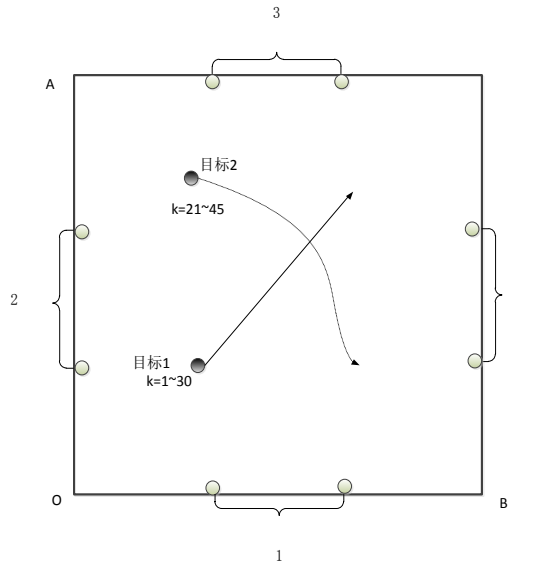


图 5-2 两个目标运动轨迹图

图 5-3 显示了利用四对麦克风实现对目标的定位与跟踪结果图。图 5-3 (a)显示了目标横坐标的真实值与估计值的对比；图 5-3 (b)显示了目标纵坐标的真实值与估计值的对比。图 5-4 显示了蒙特卡罗次数为 100 次时目标个数的真实值与估计值之间的对比。从图 5-3 和图 5-4，可以看出多传感器 PHD 滤波算法(SMC-MS-PHD)能够有效地实现对目标的定位与跟踪，但是当两个目标同时存在时，目标的跟踪性能下降，这是因为两个声源信号存在相互干扰，存在一定的相关性。

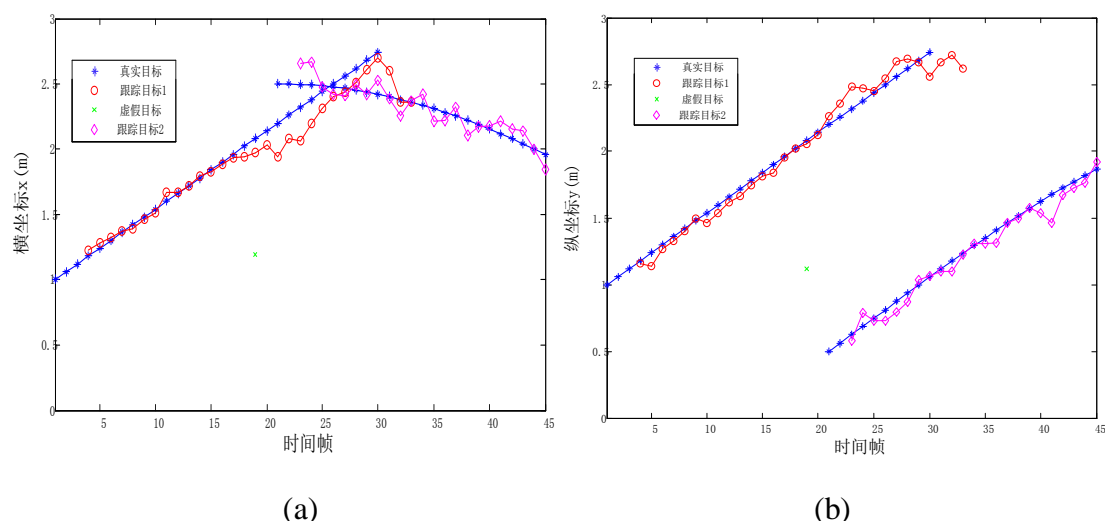


图 5-3 利用四对麦克风实现对目标的定位与跟踪结果

(a)目标横坐标的真实值与估计值对比；(b)目标纵坐标的真实值与估计值对比

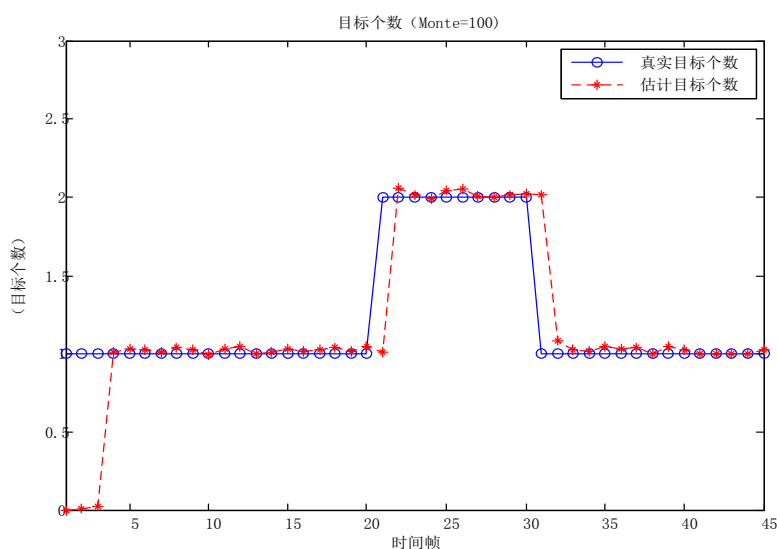


图 5-4 目标个数的真实值与估计值

从图 5-4，可以看出，在 $k = 20$ 和 $k = 30$ 目标个数发生变化时，真实目标个数与估计目标个数相差较大，这是由于目标个数发生变化时，采样粒子不能及时地

跟踪到新出现的目标。

5.4 多检测 PHD 滤波算法研究

大部分传统的目标跟踪算法都假设在一帧内最多产生一个检测值。然而，在许多实际的目标跟踪应用中，由于多径传播、高分辨率的接收器或者其它因素，一个目标可能在一帧内产生多个检测值。如果从一个目标接收到的多个检测值能够得到有效的利用，它定能对提高多目标跟踪的效果有很大帮助。然而，这种算法面临的挑战是目标个数的不确定以及观测数据与目标之间的关联关系不确定增加了目标跟踪算法的复杂性。为了解决这些问题，将用到随机有限集模型和随机有限集统计学。与标准的 PHD 算法相比，没有额外的假设被定义，一种多检测 PHD (MD-PHD) 更新公式用于公式推导。更进一步，本文采用粒子滤波的方式实现 MD-PHD 算法，即 SMC-MD-PHD 滤波器。

5.4.1 多检测目标跟踪背景

在这一部分，首先介绍多检测目标跟踪的相关背景，正如上面提到的一样，在许多的应用场景都可以应用多检测目标跟踪的算法。下面来讨论室内麦克风定位与跟踪的模型，如图 5-5 所示：

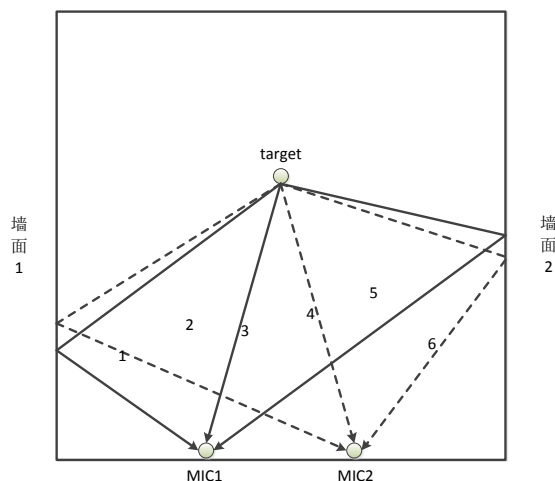


图 5-5 室内声音多径传播模型

声源目标发出的声音由于室内墙面的反射，到达两麦克风 MIC1 和 MIC2 的声音信号不仅仅只有直射波的信号，对于麦克风 MIC1 接收到的信号就包括经过路径 1、路径 3、路径 5 到达的信号（如图中实线所示）；对于麦克风 MIC2 接收到的信号就包括经过路径 2、路径 4、路径 6 到达的信号（如图中虚线所示）。如图 5-6，

在测得的 GCC 结果中不仅包涵有直射波与直射波之间的时间差，还包涵有反射波与直射波之间的时间差，如果能有效利用这些时间差，在一对麦克风的情况下，将能够对声源目标进行定位与跟踪。如何有效地利用这些多径信息将是本文研究的重点。

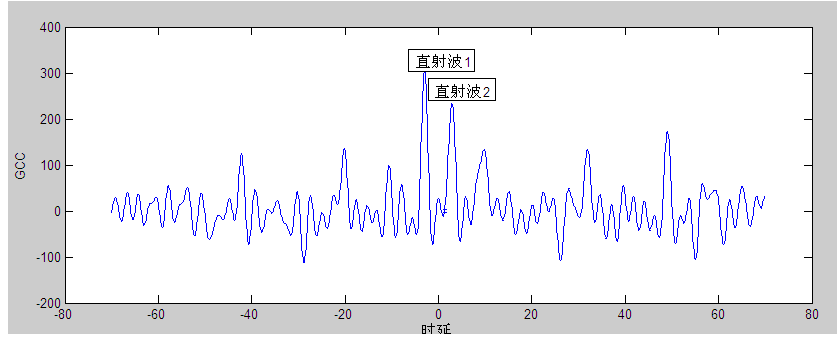


图 5-6 GCC 结果

对于多个观测值的系统下，假设有 L 种可能的检测模型，也就是传播路径，并且每一种的检测方程已知。 z_k^j 代表第 j 种检测模型的观测值，那么有如下定义：

$$z_k^j = \begin{cases} h_k^1(x_i) + w_k^1 & \text{if } l = 1 \\ h_k^2(x_i) + w_k^2 & \text{if } l = 2 \\ \vdots & \\ h_k^L(x_i) + w_k^L & \text{if } l = L \\ \kappa & \text{if clutter} \end{cases} \quad (5-24)$$

其中， $h_k^l(\cdot), l=1,2,\dots,L$ 代表第 l 种模型下的检测方程， w_k^l 代表相应的检测噪声，这里假设检测噪声是一个零均值的高斯噪声，与目标状态不相关。对于跟踪来说，目标 x_k^i 、观测值 z_k^j 和模型 l 之间的对应关系未知。假设虚假观测数据 κ 平均分布在整个观测平面上，并且每一帧的虚假个数服从泊松分布，其均值为 λ 。

然而，需注意的地方是，在许多实际的应用环境中，多个观测值可能由未知的模式产生，也就是说， $h_k^l(\cdot)$ 的准确表达式可能不知道，或者说检测模型是随着时间变化的。但是，这不是本文要研究的问题，为了简化，假设检测模型个数 L 是常数， $h_k^l(\cdot)$ 的准确表达式也知道，可能目标的检测模型有多种，可以根据先验信息，选择其中最重要的 L 种作为检测模型^[45-46]。

本文的主要目的是从观测集 Z_k 中，估计出目标的状态集 X_k ，于是将用到贝叶斯框架下的多目标跟踪理论，如第四章所述。

5.4.2 多检测 PHD (MD-PHD) 滤波器

因为本文提出的 MD-PHD 滤波器与标准的 PHD 滤波器相比, 预测过程相同, 所以主要研究更新过程的算法。因此在这里, 只介绍 MD-PHD 滤波器的更新过程的算法。

5.4.2.1 多检测模型

为了推导 MD-PHD, 做出如下假设:

假设 1: 假设这里有 L 种检测模型, 每种模型有一个已知的似然函数:

$$f_{k+1}^l(z|x) = f_{w_{k+1}^l}^l(z - h_{k+1}^l(x)), \quad l = 1, \dots, L \quad (5-25)$$

其中, $f_{w_{k+1}^l}^l(\cdot)$ 是检测噪声, w_{k+1}^l 的概率密度分布函数。

假设 2: 一帧内, 一个目标通过一个检测模型最多只能产生一个观测值。

假设 3: 每一个检测模型的检测概率 $p_D^l(x)$ 可以不同, 并且每个状态的检测概率相互独立。

基于以上假设, 通过应用参考文献[47]中 FISST 公式可以得到一个目标的似然函数。对于单目标状态 x 和相应的观测集 W , 如果 $W = \theta$, 也就是没有检测到目标的情况, 其似然函数仅仅是所有模型的未检测概率的乘积:

$$f_{k+1}(W|x) = f_{k+1}(\theta|x) = \prod_{l=1}^L (1 - p_D^l(x)) \quad (5-26)$$

对于 $W = \{z_1, \dots, z_m\}$, 其中 $|W| = m \leq L$, 可得到其似然函数如下:

$$f_{k+1}(W|x) = \sum_{\theta} \left(\prod_{\theta(l)=0} (1 - p_D^l(x)) \right) \left(\prod_{\theta(l)>0} p_D^l(x) h_{k+1}^l(z_{\theta(l)}|x) \right) \quad (5-27)$$

在这里, $h_{k+1}^l(z|x)$ 为第 l 种模型下的检测似然函数, 式 (5-27) 所作的求和包括了所有可能性的假设 $\theta: \{1, \dots, L\} \rightarrow \{0, 1, \dots, m\}$, 如参考文献[47]所示。下面将举例说明, 假设 $L=3$ 和 $W = \{z_1, z_2\}$, 其相应的似然函数如下:

$$\begin{aligned} f_{k+1}(W|x) = & (1 - p_D^1(x)) \cdot p_D^2(x) \cdot p_D^3(x) \cdot f_{k+1}^2(z_1|x) \cdot f_{k+1}^3(z_2|x) \\ & + (1 - p_D^1(x)) \cdot p_D^3(x) \cdot p_D^2(x) \cdot f_{k+1}^3(z_1|x) \cdot f_{k+1}^2(z_2|x) \\ & + (1 - p_D^2(x)) \cdot p_D^3(x) \cdot p_D^1(x) \cdot f_{k+1}^3(z_1|x) \cdot f_{k+1}^1(z_2|x) \\ & + (1 - p_D^2(x)) \cdot p_D^1(x) \cdot p_D^3(x) \cdot f_{k+1}^1(z_1|x) \cdot f_{k+1}^3(z_2|x) \\ & + (1 - p_D^3(x)) \cdot p_D^2(x) \cdot p_D^1(x) \cdot f_{k+1}^2(z_1|x) \cdot f_{k+1}^1(z_2|x) \\ & + (1 - p_D^3(x)) \cdot p_D^1(x) \cdot p_D^2(x) \cdot f_{k+1}^1(z_1|x) \cdot f_{k+1}^2(z_2|x) \end{aligned} \quad (5-28)$$

5.4.2.1 MD-PHD 滤波

现在只讨论 MD-PHD 滤波更新过程的算法,它是利用麦克风接收到的检测值,通过似然函数来重新计算权值,再对分布进行更新^[48]。现在必须考虑在当前时刻的量测集 Z_{k+1} 中可能存在多个值来自于同一个目标。因此,在对式(5-28)作求和运算时,必须考虑 Z_{k+1} 的所有种可能的划分。于是,定义了一种 L_{\max} 划分:它规定对于所有的 $W \in P$,其中, W 中元素的个数不超过 L , P 包含了全部的 L_{\max} 划分^[49-50]。下面举例说明:假设 $Z_{k+1} = \{z_1, z_2, z_3\}$, $L=2$, $|Z_{k+1}|=3$, 其划分结果如下:

$$P_1 = \{\{z_1\}, \{z_2\}, \{z_3\}\}, \text{ where } W_1 = \{z_1\}, W_2 = \{z_2\}, W_3 = \{z_3\}$$

$$P_2 = \{\{z_1, z_2\}, \{z_3\}\}, \text{ where } W_1 = \{z_1, z_2\}, W_2 = \{z_3\}$$

$$P_3 = \{\{z_1, z_3\}, \{z_2\}\}, \text{ where } W_1 = \{z_1, z_3\}, W_2 = \{z_2\}$$

$$P_4 = \{\{z_2, z_3\}, \{z_1\}\}, \text{ where } W_1 = \{z_2, z_3\}, W_2 = \{z_1\}$$

基于标准的 PHD 更新公式,于是得出 MD-PHD 的更新公式如下:

$$L_{Z_{k+1}}(x) = (1 - \tilde{p}_D(x)) + \sum_{P \subset Z_{k+1}} \omega_P \sum_{W \in P} \frac{L_W(x)}{\kappa_W + \tau_W} \quad (5-29)$$

其中,

$$L_W(x) = f_{k+1}(W | x) \quad (5-30)$$

$$1 - \tilde{p}_D(x) = L_\theta(x) = f_{k+1}(\theta | x) \quad (5-31)$$

$$\kappa_W = \frac{\delta \log \kappa_{k+1}}{\delta W} [0] \quad (5-32)$$

$\kappa_{k+1}[g]$ 是杂波的生成概率,这里,假设杂波的随机集服从 Poisson 分布,与目标产生的观测值之间相互独立。最终,可以得到如下结果:

$$\kappa_W = \begin{cases} -\lambda_{k+1} & \text{if } W = \theta \\ \kappa_{k+1}(z) & \text{if } W = \{z\} \\ 0 & \text{if } |W| > 1 \end{cases} \quad (5-33)$$

在本文中, $W = \theta$ 不考虑。在公式(5-29)只是一个简单的符号公式,它不代表任何特殊的检测概率 \tilde{p}_D 或者目标似然函数 $L_W(x)$ 。这使得该公式可以运用于许多的改进的 PHD 算法中。

将式(5-26)、(5-27)代入式(5-29)中,可以得到如下的 MD-PHD 的似然函数:

$$\begin{aligned}
 L_{Z_{k+1}}(x) &= (1 - \tilde{p}_D(x)) + \sum_{P \in Z_{k+1}} \omega_P \sum_{W \in P} \frac{L_W(x)}{\kappa_W + \tau_W} \\
 &= \left(\prod_{l=1}^L (1 - p_D^l(x)) \right) + \sum_{P \in Z_{k+1}} w_P \sum_{W \in P} \\
 &\quad \frac{\sum_{\theta} \left(\prod_{\theta(l)=0} (1 - p_D^l(x)) \right) \left(\prod_{\theta(l)>0} p_D^l(x) h_{k+1}^l(z_{\theta(l)} | x) \right)}{\kappa_W + \tau_W}
 \end{aligned} \tag{5-34}$$

在这里必须考虑所有的 L_{\max} 划分。但是，如果 $L=1$ ，式(5-27)的似然函数将退化成单目标单检测公式。这是因为在量测集中，将得到 $|Z_{k+1}|=1$ ，因此式(5-34)将变成标准的 PHD 中的似然函数。

5.5 实验仿真及结果分析

本节将按照上节的算法进行实验和仿真，本文实验，使用 matlabR2008b 平台进行仿真，基于随机有限集下的多声源目标跟踪包括如下几个步骤：

第一步：声源信号的生成，整个房间的冲击响应用 IMAGE 的房间混响模型模拟。

第二步：声源信号预处理，主要包括：预滤波、加窗分帧、归一化和去噪等。

第三步：求取到达时间差，本文利用改进的 PHAT-GCC 求得到达时间差。

第四步：跟踪，采用粒子 MD-PHD 滤波器，求得目标的状态和个数。

在本实验中，假设声源经过墙面反射的信号与声源相对于墙面的对称点发出的信号等价，同时，假设目标的检测模型只有5种，且目标与麦克风的高度相同，假设目标位置 (x, y, z_0) ，麦克风1坐标为 (x_1, y_1, z_0) 、麦克风2坐标 (x_2, y_2, z_0) ， L_0 代表房间的长度，那么第 j 种检测模型检测值 z_k^j 代表的有如下定义：

$$z_k^j = \begin{cases} (\sqrt{(x-x_2)^2 - y^2} - \sqrt{(x-x_1)^2 - y^2}) + w_k^1 & \text{if } l=1 \\ (\sqrt{(-x-x_2)^2 - y^2} - \sqrt{(x-x_1)^2 - y^2}) + w_k^2 & \text{if } l=2 \\ (\sqrt{(2L_0-x-x_2)^2 - y^2} - \sqrt{(x-x_1)^2 - y^2}) + w_k^3 & \text{if } l=3 \\ (\sqrt{(x-x_2)^2 - y^2} - \sqrt{(-x-x_1)^2 - y^2}) + w_k^4 & \text{if } l=4 \\ (\sqrt{(x-x_2)^2 - y^2} - \sqrt{(2L_0-x-x_1)^2 - y^2}) + w_k^5 & \text{if } l=5 \\ \kappa & \text{if clutter} \end{cases} \tag{5-35}$$

5.5.1 单声源目标定位与跟踪实现

本实验选择房间作为仿真环境，房间的大小为 $3\text{m} \times 3\text{m} \times 2.5\text{m}$ ，采用 1 对麦克

风，麦克风的坐标为：麦克风 1 (1.5,0.01,1.8)、麦克风 2 (1.0,0.01,1.8)，两个麦克风这间的距离是 0.5m,如图 5-7 所示：

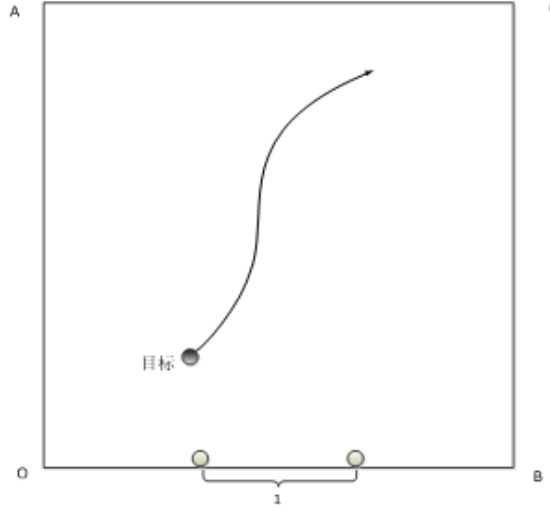


图 5-7 仿真模型

所用的单声源为女声，整个房间的冲击响应用 IMAGE 的房间混响模型模拟，声音在空气中的传播速度为 $c = 344 \text{ m/s}$ ，设房间的四个顶点坐标分别为 $O(0,0,1.8)$ 、 $A(0,3,1.8)$ 、 $B(3,0,1.8)$ 、 $C(3,3,1.8)$ 。一般房间有六个墙壁， β_{x1} 、 β_{x2} 、 β_{y1} 、 β_{y2} 、 β_{z1} 、 β_{z2} 分别是各面墙的反射系数，根据实验经验，设置对应的值为 (0.85,0.85,0.25,0.25,0.25,0.25)， (x,y,z) 是声源位置的坐标， $L_x = 3$ 、 $L_y = 3$ 、 $L_z = 2.5$ 分别是房间的长、宽、高。选用的帧长度是 1024，采样频率为 8KHz，信号长度是 7.68s，信噪比为 10dB，目标个数为 1， $\eta = 10 \text{ s}^{-1}$ 和 $\nu = 1 \text{ ms}^{-1}$ 为模型参数， $P_{\text{birth}} = 0.05$ ， $P_{\text{death}} = 0.01$ ， $P_{\text{miss}} = 0.25$ ，每个目标的采样粒子数 $N_p = 500$ ，新生目标的粒子数为 $N_b = 100$ 。

在这里，假设目标的初始位置在 (1.0,1.0,1.8)，在平面内做随机自由运动， $L = 5$ ， $|Z_{k+1}| = 3$ ，同时假设各种检测模型的检测概率相互独立， $P_{34} = 0.9$ ， $P_{14} = P_{54} = P_{32} = P_{36} = 0.4$ 。图5-8为测得的单目标GCC结果图（其中实线为直达波的时延，空心圆为大于门限的峰值点，实心圆为最大峰值点），其中门限值为最大值的0.25倍，图5-9为SMC-MD-PHD单目标跟踪结果，图5-9(a)和图5-9 (b)分别显示 SMC-MD-PHD算法的跟踪情况与真实情况的横坐标和纵坐标对比。从图5-9可以看出，基于一对麦克风情况下，采用SMC-MD-PHD算法，能够很好地对单目标实施跟踪。

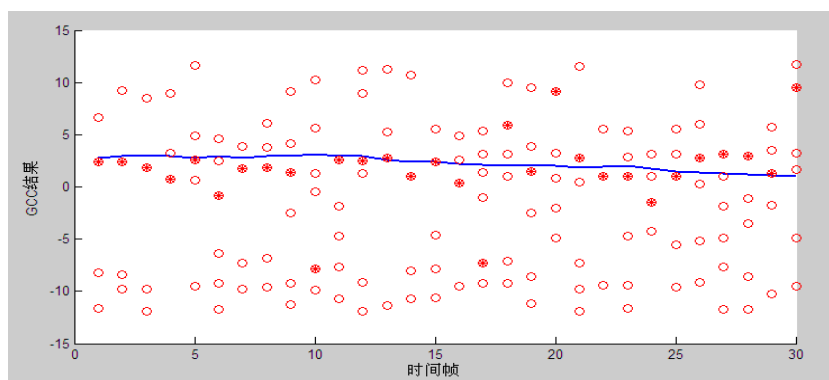


图 5-8 单目标 GCC 结果（实线为真实直达波的时延，空心圆为大于门限的峰值点，实心圆为最大峰值点）

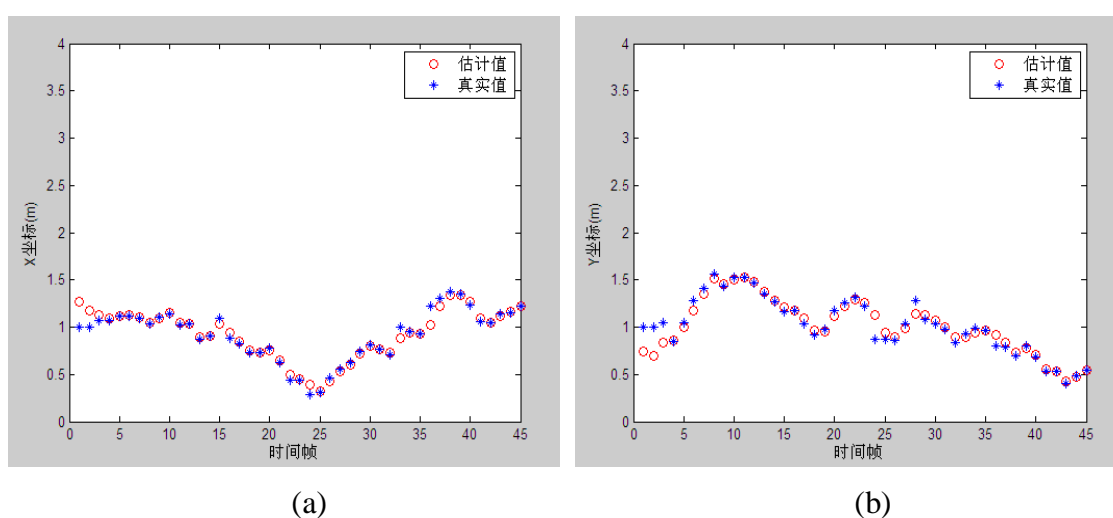


图 5-9 SMC-MD-PHD 单目标跟踪结果

(a)目标横坐标的估计值与真实值；(b)目标纵坐标的估计值与真实值

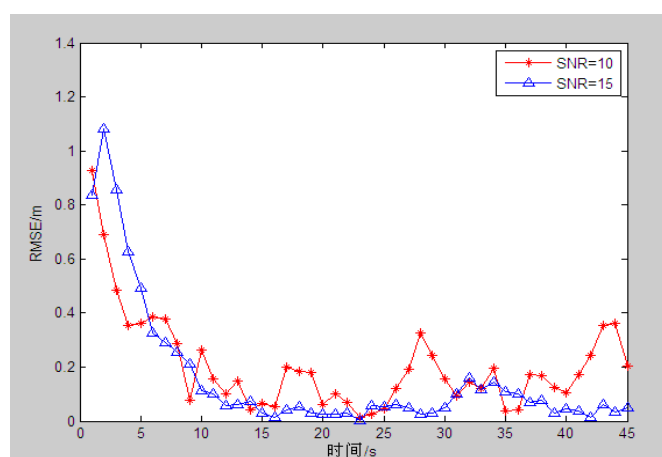


图 5-10 RMSE 结果

图5-10显示了在单目标状态下，目标的RMSE误差，从图中可以看出，当SNR

值较高时，目标的RMSE结果偏小；当SNR值较低时，目标的RMSE值偏大。这是因为当SNR较低时，时延估计的精度较低，当SNR较高时，时延估计的精度较高。

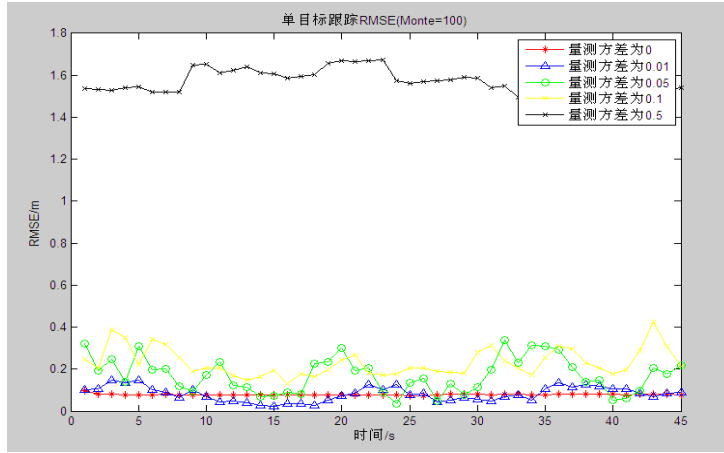


图 5-11 量测方差对跟踪性能的影响

虽然 SMC-MD-PHD 算法对单目标能够进行有效的跟踪，但是从图 5-11 可以看出，量测结果的好坏对跟踪性能的有明显的影晌，可以看出量测方差越小，估计精度越高；方差越大，估计精度越低。当方差足够小时，RMSE 的值就很小，也就对目标的跟踪性能非常好。

5.5.2 多声源目标定位与跟踪实现

为了简化问题的复杂性，假设两个声源之间相关性较小，所用的声源为多频声音信号，整个房间的冲击响应用IMAGE的房间混响模型模拟，假设目标为两个静止目标：目标1位置为(1.0,1.0,1.8)，在 $k=1$ 时刻出现，在 $k=25$ 时刻消失；目标2位置为(2.0,2.0,1.8)，在 $k=10$ 时刻出现，在 $k=30$ 时刻消失。两个麦克风位置：麦克风1(1.25,0.01,1.8)，麦克风2(1.75,0.01,1.8)。这里 $L=5$ ， $|Z_{k+1}|=6$ ，其他参数设置与单声源目标相同。

图 5-12 为测得的多目标 GCC 结果图（其中实线为直达波的时延，空心圆为大于门限的峰值点，实心圆为最大峰值点），其中门限值为最大值的 0.25 倍，这些测得的结果就用于后面多检测滤波器中。图 5-13 为 SMC-MD-PHD 算法在不同时刻粒子分布情况。

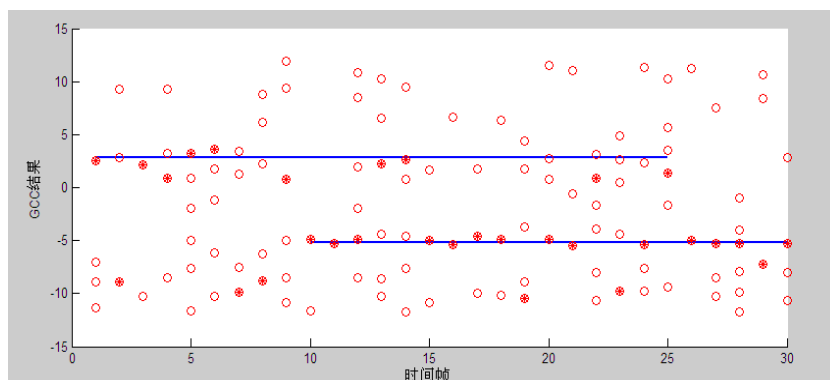


图 5-12 多目标 GCC 结果（实线为真实直达波的时延，空心圆为大于门限的峰值点，实心圆为最大峰值点）

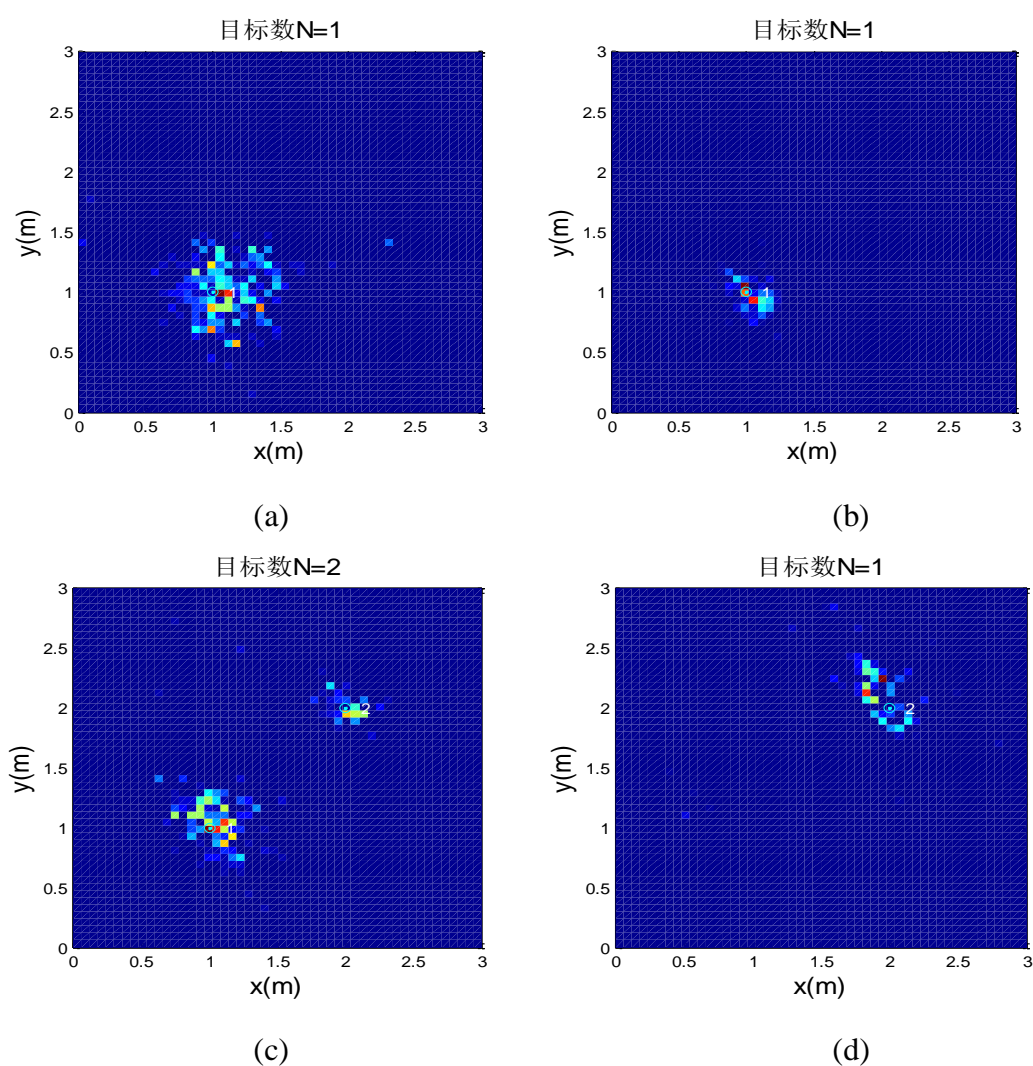


图 5-13 不同时刻粒子分布情况

(a) $k=3$ ，目标数 $N=1$ ；(b) $k=9$ ，目标数 $N=1$ ；(c) $k=15$ ，目标数 $N=2$ ；(d) $k=21$ ，目标数 $N=1$ ；

从图 5-13 可以看出，在不同时刻，采用 SMC-MD-PHD 算法得到的粒子分布

能够很好地实现对目标的跟踪。

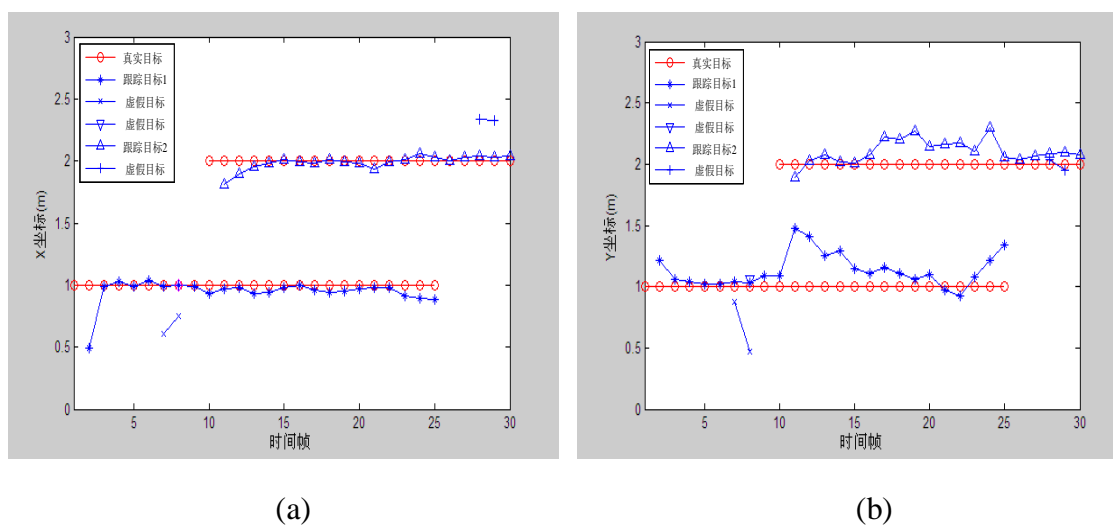


图 5-14 SMC-MD-PHD 多目标跟踪结果

(a)目标横坐标的真实值与估计值；(b)目标纵坐标的真实值与估计值

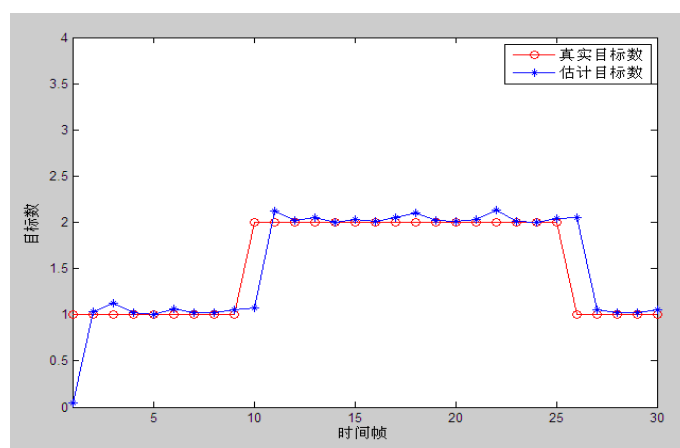


图 5-15 目标个数的真实值与估计值

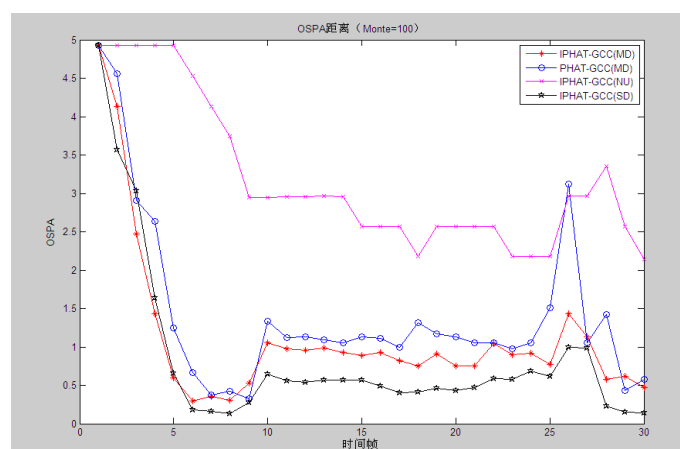


图 5-16 OSPA 距离

图 5-14 为 SMC-MD-PHD 多目标跟踪结果图 5-14(a)显示的是目标真实值与估计值的横坐标, 图 5-14(b)显示的是目标真实值与估计值的纵坐标, 图 5-15 显示的是蒙特卡罗次数为 100 时目标个数的真实值与估计值, 从图 5-14、图 5-15 可以看出, SMC-MD-PHD 算法能够对两个目标进行有效的跟踪。

图 5-16 显示的是蒙特卡罗次数为 100 时的 OPSA 距离, 可以看出, 多传感器 PHD 算法 (IPHAT-MS-GCC) 的跟踪效果最好, 而改进的 PHAT-MD-GCC (IPHAT-MD-GCC) 比 PHAT-MD-GCC 跟踪效果好。同时, 还可以看出未使用多径信息的粒子滤波概率假设密度算法 (SMC-NU-PHD) 不能对目标进行有效的跟踪, 其中 $p=2, c=25$ 。下面, 分析了不同 SNR 值对跟踪性能的影响, 如图 5-17 所示。

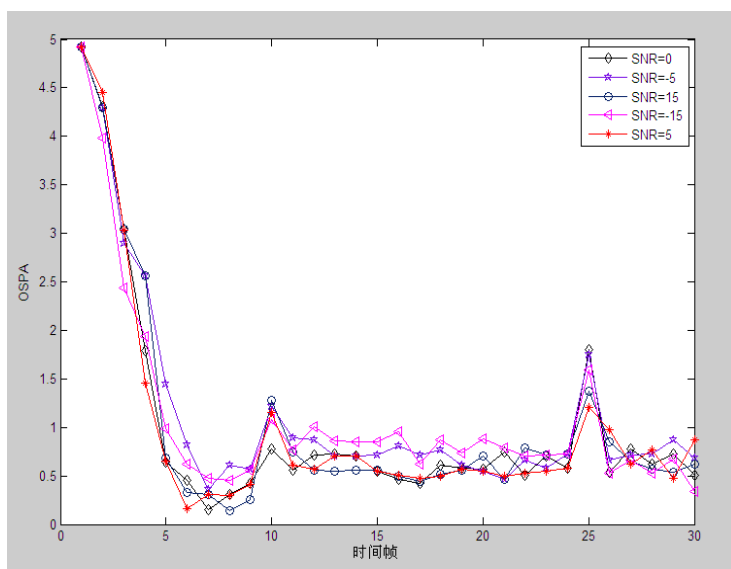


图 5-17 不同 SNR 条件下 OSPA 距离比较

图 5-17 为 SNR 分别为-15, -5, 0, 5, 15 时的 OSPA 距离, 从图中可以看出, 当 SNR 较高时, PHAT-GCC 时延估计结果误差较小, 目标的跟踪性能较好; 当 SNR 较低时, PHAT-GCC 时延估计结果误差较大, 目标的跟踪性能较差。当在 $k=10$ 和 $k=25$ 时会出现波峰, 这是因为目标状态和目标数在该时刻发生了变化。

5.6 本章小结

本章简要概述了单声源目标和多声源目标的状态空间模型, 研究了多传感器下多目标定位与跟踪算法并通过仿真验证了其有效性。针对在实际的应用场景中存在多径信息, 本文研究了一种利用多径信号的多检测概率假设密度多目标跟踪算法。通过该算法实现了利用一对麦克风对多目标的定位与跟踪。在实验仿真中, 本章主要从两个方面对多检测概率假设密度多目标跟踪算法进行仿真, 首先是针

对单目标的情况，其次是针对多目标的情况。然后通过仿真比较了 IPHAT-MS-GCC、IPHAT-MD-GCC、PHAT-MD-GCC、SMC-NU-PHD 四种算法的性能，通过比较验证了 IPHAT-MD-GCC 的有效性。最后通过仿真分析了 SNR 对 IPHAT-MD-GCC 的跟踪性能的影响。

第六章 总结与展望

6.1 总结

利用声源信息的多目标定位与跟踪技术目前还不成熟，但是却有着广泛的应用前景，如声源目标控制、智能机器、反声源武器等领域。本文正是从应用出发，针对实际声源的传播特点对声源多目标进行了定位与跟踪，并仿真实验环境进行了验证。

在只利用直达信号的情况下，一对麦克风是定位不出目标位置的。但在实际测得的 GCC 结果中，不仅包涵有直射波与直射波之间的时间差，还包涵有反射波与直射波之间的时间差。这些时间差是由于信号在传播过程中遇到障碍物的结果。于是认为，如果能有效利用这些时间差，在一对麦克风的情况下，将能够对声源目标进行定位与跟踪。

因此，以这一发现为出发点，本文主要展开了以下内容研究：

(1) 提出了利用声源信息的多目标定位与跟踪的思路，针对声音信号的特点进行了分析，并针对声音信号的特点进行了一系列预处理，包括预滤波、加窗分帧、预加重、归一化和去噪等，并进行了仿真分析。同时建立了声源信号的传播模型，为后续信号的处理打下基础。

(2) 研究分析了传统的广义互相关算法，比较了他们的优缺点。最终选择了性能较优的 PHAT-GCC 时延迟估计算法。但是传统的 PHAT-GCC 方法已经不能满足实际的应用需求，于是提出了一种改进的 PHAT-GCC 求时延方法，通过仿真验证了其在抗噪声等方面的优越性。

(3) 介绍了随机集的相关理论，通过利用随机集可以使复杂的多目标的问题转化成简单的单目标跟踪问题。然后，对多目标跟踪问题进行建模，并针对 PHD 算法积分运算困难的问题，研究了其粒子滤波实现。最后，针对声源信号的特点，建立了基于随机有限集下的多声源跟踪模型。

(4) 基于一对麦克风的情况下，利用房间内的多径信息和多个检测值，建立了多检测模型，研究了 MD-PHD 滤波算法（与标准 PHD 滤波器相比，预测过程相同，更新过程不同），并详细介绍了其推导过程。最后，分别在单声源和多声源情况下，进行了实验仿真，实验结果验证了该算法的可行性。

6.2 工作展望

在本文中，利用声音信号传输过程中的多径信息对目标进行定位，但是从实际应用来讲，还有许多需要继续研究的地方：

（1）首先是信号的获取。只有获得精确的信号才能对目标进行精确的定位。然而通常情况下，声音信号所处的环境相当复杂，不仅包含目标的声音信号，还包含自然界的各种强干扰信号。所以如何有效地对声音信号进行识别和提取将对定位与跟踪起到决定性的作用，这也将是下一步研究的重点。

（2）深入理解粒子滤波的算法原理，在重要性函数的构建及重采样算法的改进方面进行进一步的研究，提高粒子滤波 PHD 算法的跟踪性能。

（3）在本文中，假设目标所有可能的检测方程式和检测方程数已知且固定不变。但在实际实用，检测方程式和检测方程数都是未知的，检测方程式和检测方程数有可能是时变的，还有可能检测方程不能一一列举。因此，如何准确合理地确定检测方程式将是进一步研究的重点。

（4）虽然本文提出的 MD-PHD 算法在跟踪一些环境比较恶劣情况下的目标有较好的效果，但是其计算量比较大。因此，下一步的工作将是如何减少计算量。同时，考虑将 MD-PHD 算法用到更多的应用场景中去。

致 谢

三年的时间非常短暂，又该是说再见的时候。三年的研究生生活带给了我许多人生有益的东西，我想这些东西将伴随我以后的工作生涯。此时此刻，我内心充满了对我身边的人的感激之情。

首先，我要感谢我的指导老师唐续，通过三年时光，我加深了对唐老师的了解。唐老师是一位工作认真，治学态度非常严谨的人。我还记得刚到校的时候，唐老师就对我说，做研究要静得下心来，只有心静下来了，什么困难都能解决。唐老师还对我说，你做得好，我就表扬你，你如果做得不好，我就会批评你，所以希望你要努力做好每一件事。正是记住了这两句话，我一直以来，都对自己要求严格，早起晚归，尽善尽美地做好自己的每一项工作任务。但唐老师在生活上却是非常的体贴，他经常关心我们的吃穿，让我们从内心里面感觉很温暖。

其次，我要感谢我们教研室的老师，在这三年里，我每天都能看到你们，是你们的一举一动教会了我生命的意义、学习的意义。学无止境，这是你们教给我们最好的课程。

再次，我要感谢教研室的同学们，与你们一起生活、一起娱乐，让我在三年的人生旅程里多了一份精彩，感谢你们给我的鼓励，感谢你们给我的帮助，我相信我们之间的友谊将万世永存。

我还要感谢我的室友张博军、吴俊和高现文，没有你们给我生活上的帮助，我不可能有这么好的成绩。

最后，我还要特别感谢我的家人，在三年时光里，你们无时无刻不挂念着我，在生活和精神上，都给了我无比多的支持。你们是最亲的人，你们将成为我继续努力的动力。

参考文献

- [1] MARTI A, COBOS M, LOPEZ JJ. Real time speaker localization and detection system for camera steering in multi-participant videoconferencing environments[C]//Proc of IEEE International Conference on Acoustics, Speech, and Signal Processing. [S.l.]: IEEE Press, 2011: 2592-2595.
- [2] Perrodin F, Nikolic J, Busset J, et al. Design and calibration of large microphone arrays for robotic applications[C]// IEEE/RSJ International Conference on Intelligent Robots & Systems. IEEE, 2012:4596-4601.
- [3] CANCLINI A, ANTONACCI F, SARTI A, et al. Acoustic source localization with distributed asynchronous microphone networks[J]. IEEE Trans on Audio, Speech and Language Processing, 2013, 21(2) : 439-443.
- [4] 杨俊梅, 方厚林, 成龙, 等. 反狙击系统传感器阵列的设计[C]// 2014.中国西部声学学术交流会议论文集. 2014.
- [5] 郭俊成. 基于传声器阵列的声源定位技术研究[D].南京航空航天大学, 2007.
- [6] 赵玉洁. 声测定位技术的现状与发展趋势[J]. 声学与电子工程. 1993.
- [7] “巴格达之战”考验美军巷战武器装备. 兵器快讯. 2003.04.08.
- [8] Brandstein M S, Silverman H F. A Practical Methodology for Speech Source Localization With Microphone Arrays[J]. Computer Speech & Language, 1997, 11(2):91-126.
- [9] 蔡卫平, 殷侠, 徐健. 一种基于粒子滤波的混合声源跟踪算法[J]. 计算机工程与应用, 2012, 48(15):113-118.
- [10] 蔡卫平, 徐健, 周琳. 一种基于粒子滤波的多声源跟踪算法[J]. 计算机应用研究, 2014, 31(8):2326-2330.
- [11] 黄海军. 基于传声器阵列的声源定位系统的初步研究[D]. 东华大学, 2013.
- [12] Dibiase J H. A high-accuracy, low-latency technique for talker localization in reverberant environments using microphone Arrays[J]. European Journal of Biochemistry, 2010, 216(1):281-91.
- [13] Brandstein M S, Silverman H F. A Practical Methodology for Speech Source Localization With Microphone Arrays[J]. Computer Speech & Language, 1997, 11(2):91-126.
- [14] Schmidt R O. Multiple emitter location and signal parameter estimation. IEEE Transactions on Antennas and Propagation, 1986, 34(3):276-280.
- [15] Cadzow J A, Kim Y S, Shiue D C. General direction of arrival estimation: a signal subspace

- approach. IEEE Transactions on Aerospace and Electronic Systems, 1989, 25(1):31-47.
- [16] Pillai S U, Kwon B H. Forward/backward spatial smoothing techniques for coherent signal identification. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1989,37(1):8-15.
- [17] Yang Y, Wan C, Sun C, et al. DOA estimation for coherent sources in beamspace using spatial smoothing[C]// Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint Conference of the Fourth International Conference on. 2004:1028-1032 vol.2.
- [18] Kirlin R L, Moore D F, Kubichek R F. Improvement of delay measurements from sonar arrays via sequential state estimation[J]. IEEE Transactions on Acoustic, Speech, Signal Processing,1981,7(29):514-519.
- [19] Reed F A, Feintuch P L, Bershad N J. Time delay estimation using the LMS adaptive filter-static behavior[J]. IEEE International Conference on Acoustic, Speech, Signal Processing.1981,vol.29:561-571.
- [20] Youn D H, Ahmed N, Carter G C. On using the LMS algorithm for time delay estimation. IEEE Transactions on Acoustics, Speech and Signal Processing, 2006, 30(5):798-801.
- [21] Hahn W, Tretter S A. Optimum processing for delay-vector estimation in passive signal arrays[J]. IEEE Transactions on Information Theory, 1973,19:608-614.
- [22] Rudzyn B, Kadous W, Sammut C. Real time robot audition system incorporating both 3D sound source location and voice characterization[C]. IEEE International Conference on robotics and automation,Roma,Italy,2007,4733-4738.
- [23] 崔玮玮, 曹志刚, 魏建强. 声源定位中的时延估计技术[J]. 数据采集与处理, 2007, 22(1):90-99.
- [24] 林志斌, 徐柏龄. 基于传声器阵列的声源定位,扬声器与传声器,2004,5,18~22.
- [25] 李海城. 基于传声器阵列的自动声源定位方法,辽宁师范大学学报,2006,9(2):168~171.
- [26] Knapp C H, Carter G C. The generalized correlation method for estimation of time delay.IEEE Transactions on Acoustics, Speech and Signal Processing, 1976, 24(4):320-327.
- [27] 张雪英. 语音信号处理及 MATLAB 仿真[M]. 北京: 电子工业出版社, 2010.
- [28] Bar-Shalom Y. Tracking and Data Association. San Diego:Academic Press, 1988.
- [29] Bar-Shalom Y, Kirubarajan T, Lin X. Probabilistic data association techniques for target tracking with applications to sonar, radar and EO sensors. IEEE Aerospace and Electronic Systems Magazine, 2005, 20(8): 37 56.
- [30] Musicki D, Evans R. Joint integrated probabilistic data association: JIPDA. IEEE Transactions

- on Aerospace and Electronic Systems, 2004, 40(3): 1093-1099.
- [31] Blackman S S. Multiple hypothesis tracking for multiple target tracking. IEEE Aerospace and Electronic Systems Magazine, 2004, 19(1): 5-18.
 - [32] 张鹤冰. 概率假设密度滤波算法及其在多目标跟踪中的应用[D]. 哈尔滨工程大学, 2012.
 - [33] Mahler R P. Statistical Multisource-Multitarget Information Fusion. Norwood: Artech House, 2007.
 - [34] 杨峰, 王永齐, 梁彦,等. 基于概率假设密度滤波方法的多目标跟踪技术综述[J]. 自动化学报, 2013, 39(11):1944-1956.
 - [35] Vo B N, Singh S, Boucet A. Sequential Monte Carlo methods for multi-target filtering with random finite sets. IEEE Transactions on Aerospace and Electronic Systems. 2005,41(4): 1224-1245.
 - [36] Vo B N, Ma W K. The Gaussian mixture probability hypothesis density filter. IEEE Transactions on Signal Processing, 2006, 54(11): 4091-4104.
 - [37] Mahler R P S. Multitarget Bayes Filtering via First-Order Multitarget Moments[J]. IEEE Transactions on Aerospace & Electronic Systems, 2003, 39(4):1152-1178.
 - [38] Goodman I, Mahler R, Nguyen H. Mathematics of data fusion[M]. Boston: Kluwer Academic Publishers, 1997.
 - [39] Mahler R. Random set theory for target tracking and identification[M]. Boca Raton FL: CRC Press, 2002.
 - [40] 田淑荣, 王国宏, 何友. 多目标跟踪的概率假设密度粒子滤波.海军航空工程学院学报. 2007, 22(4): 417-420.
 - [41] Vermaak J, Blake A. Nonlinear filtering for speaker tracking in noisy and reverberant environments[C]// IEEE International Conference on Acoustics. IEEE, 2001:3021-3024.
 - [42] Ward D B, Lehmann E A, Williamson R C. Particle filtering algorithms for tracking an acoustic source in a reverberant environment.[J]. IEEE Transactions on Speech & Audio Processing, 2003, 11(6):826-836.
 - [43] Clark D E, Bell J. Data Association for the PHD Filter[C]// Intelligent Sensors, Sensor Networks and Information Processing Conference, 2005. Proceedings of the 2005 International Conference on. IEEE, 2006:217-222.
 - [44] Panta K, Vo B N, Singh S. Improved probability hypothesis density (PHD) filter for multi-target tracking[C]. Proc. IEEE 3rd International Conference on Intelligent Sensing and Information Processing, Bangalore, India, December 14-17, 2005.
 - [45] Habtemariam B, Tharmarasa R, Thayaparan T, et al. A Multiple-Detection Joint Probabilistic

- Data Association Filter[J]. IEEE Journal of Selected Topics in Signal Processing, 2013, 7(3):461-471.
- [46] T. Sathyan, C. Tat-Jun, S. Arulampalam, and D. Suter, "A multiplehypothesis tracker for multitarget tracking with multiple simultaneousmeasurements," IEEE J. Sel. Topics Signal Process., vol. 7, no. 3, pp.448-460, Jun. 2013.
- [47] R. Mahler, Statistical Multisource-Multitarget Information Fusion.Norwood, MA, USA: Artech House, 2007.
- [48] R. Mahler, ""Statistics 101" for multisensor, multitarget data fusion,"IEEE Trans. Aerosp. Electron. Syst., vol. 19, pp. 53-64, Jan. 2004.
- [49] M. N. van Lieshout, Markov Point Processes and their Applications.London, U.K.: Imperial College Press, 2000.
- [50] I. Goodman, R. Mahler, and H. Nguyen, Mathematics of Data Fusion. Norwell, MA: Kluwer Academic, 1997.