

工学硕士学位论文

双站时差频差联合定位技术研究

刘宇

中国航天科工集团第二研究院

2018 年 3 月

国内图书分类号：

密级 公开

国际图书分类号：

工学硕士学位论文

双站时差频差联合定位技术研究

研 究 生：刘宇

导 师：王桂玲 研究员 冯胜 研究员

申请学位级别：工学硕士

学 科、专 业：信号与信息处理

所 在 单 位：二十三所

答 辩 日 期：2018 年 3 月

学位授予单位：中国航天科工集团第二研究院

Classified Index:

U.D.C:

Dissertation for Master Degree in Engineering

| | |
|-------------------------------------|-------------------------------------|
| Candidate: | Liu Yu |
| Supervisor: | Prof. Wang Guiling Prof. Feng Sheng |
| Academic Degree Applied for: | Master of Engineering |
| Specialty: | Signal and Information Processing |
| Date of Oral Examination: | March, 2018 |
| Degree-conferring Unit: | The Second Academy of CASIC |

本人声明

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究所得的成果。除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的科研成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确的方式表明。本人完全意识到本声明的法律责任由本人承担。

论文作者签名：_____

日 期：_____

关于学位论文使用授权的申明

本人完全了解中国航天第二研究院有关保留、使用学位论文的规定，同意中国航天第二研究院保留或向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅；本人授权中国航天第二研究院可以将本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、缩印或其他复制手段保存论文和汇编本学位论文。

（保密论文在解密后遵守此规定）

论文作者签名：_____

导师签名：_____

日 期：_____

摘要

ABSTRACT

目录

| | |
|----------------------------|-----|
| 本人声明..... | I |
| 摘要..... | III |
| ABSTRACT..... | IV |
| 第一章 绪论..... | 1 |
| 1.1 研究背景及意义 | 1 |
| 1.2 常见的无源定位体制 | 3 |
| 1.2.1 单站无源定位 | 3 |
| 1.2.2 双站无源定位 | 4 |
| 1.2.3 多站定位体制 | 6 |
| 1.3 时差频差定位技术研究现状 | 7 |
| 1.3.1 时差频差参数估计 | 7 |
| 1.3.2 时差频差定位算法及其误差分析 | 8 |
| 1.3.3 双站时差频差定位 | 9 |
| 1.4 论文主要内容和章节安排 | 9 |
| 第二章 双站测向时差频差定位..... | 11 |
| 2.1 双站测向时差定位 | 11 |
| 2.1.1 定位原理 | 11 |
| 2.1.2 定位误差分析 | 12 |
| 2.1.3 误差仿真分析 | 13 |
| 2.2 双站时差频差定位 | 16 |
| 2.2.1 定位原理 | 16 |
| 2.2.2 定位误差分析 | 18 |
| 2.2.3 误差仿真分析 | 19 |
| 2.3 双站定位解模糊 | 23 |
| 2.4 本章小结 | 25 |
| 第三章 基于深度神经网络的双站时差频差定位..... | 26 |
| 3.1 神经网络理论基础 | 27 |
| 3.1.1 神经网络结构 | 27 |
| 3.1.2 激活函数 | 28 |
| 3.2 深度神经网络模型 | 31 |
| 3.2.1 BP 神经网络算法原理 | 31 |
| 3.2.2 BP 神经网络的不足 | 34 |

| | |
|---------------------------|----|
| 3.2.3 深度神经网络模型 | 35 |
| 3.3 基于深度神经网络的定位仿真分析 | 41 |
| 3.3.1 模型训练 | 41 |
| 3.3.2 模型泛化 | 42 |
| 3.3.3 模型测试仿真 | 44 |
| 3.4 本章小结 | 45 |
| 第四章 时差频差参数估计 | 46 |
| 4.1 时差频差估计信号模型 | 46 |
| 4.2 时差频差估计理论下限 | 47 |
| 4.3 时差频差联合估计算法 | 49 |
| 4.3.1 时差频差粗估计 | 49 |
| 4.3.2 时差频差精估计 | 50 |
| 4.3.3 算法总结 | 52 |
| 4.4 时差频差分维估计 | 52 |
| 4.4.1 频差估计 | 53 |
| 4.4.2 时差估计 | 55 |
| 4.4.3 算法总结 | 59 |
| 4.5 算法仿真 | 60 |
| 4.6 本章小结 | 60 |
| 第五章 总结与展望 | 61 |
| 5.1 本文内容总结 | 61 |
| 5.2 展望 | 61 |
| 致谢 | 62 |
| 参考文献 | 64 |
| 攻读硕士学位期间发表的学术论文 | 71 |

第一章 绪论

1.1 研究背景及意义

无源定位系统自身不辐射电磁信号,通过被动接收辐射源发出的电磁信号对目标进行侦察定位^[1],具有作用距离远、隐蔽接收、不易被敌方发觉的优点。实现对辐射源快速高精度的定位^[1~2],对提高预警探测和战场态势感知具有重要意义,是近年来预警探测领域研究中的热点问题之一^[3]。无源定位系统通过被动接收目标辐射源信号实现目标探测、识别及定位,具有电磁隐蔽性,战场生存能力强^[4]。

根据接收辐射源信息的类型,将无源定位系统分为两大类:一类是基于外辐射源的无源定位系统,辐射源包括雷达、广播电视、卫星通信和移动通信信号等,其中,根据外辐射源的从属关系,可将基于外辐射源的无源定位系统分为合作式无源定位系统和非合作式无源定位系统;另一类是基于目标辐射源的无源定位系统,辐射源包括雷达、通信、干扰和热量等,其中根据定位对象的不同,可将基于目标辐射源的无源定位系统分为无源定位和无源导航。

基于外辐射源的无源定位系统是指通过接收除目标外的照射源辐射信号的直达波和经过目标反射后的回波来获取目标信息的系统。可能利用的外辐射源有雷达信号、广播信号(调频、调幅、数字音频广播)、电视信号(模拟电视、数字电视)、移动通信(GSM、CDMA)信号、卫星(GPS等)信号、手机基站信号等。

该无源定位系统具有下述优点:造价低廉、抗干扰能力强、隐蔽性好、不易被摧毁;各种广播电台等外辐射源分布广泛,探测系统的布站比较灵活,可以通过组网来提高协同作战能力;该系统的天线都设置在贴近地面的高处,这对探测低空飞行的飞机和巡航导弹十分有利;工作频率比较低,具有较好的反隐性。

合作式无源雷达就是传统双(多)基地雷达,其发射站和接收站在不同位置。早期的雷达体制都是双基地雷达。1922年,美国海军研究实验室(NRL)的 Taylor 博士和其助手 Young 采用类似于现在的双基地连续波雷达的设备,探测到正在河中航行的木质船。第一次世界大战中德国首先将 Klein-Heidelberg 应用于战争。

这套设备接收来自英国的海岸防空雷达 ChainHome 的直射信号和来自目标的反射信号，通过测量两信号的时间差和测量反射信号的到达角来定位目标。

但自 1936 年发明了天线收发开关和 1940 年出现大功率脉冲磁控管后，雷达的发展主要集中在单基地雷达。直至 20 世纪五六十年代，为了对付“四大威胁”的挑战，双（多）基地雷达再次受到广泛的重视和研究。如：英国 Plessey 公司研制的近程警戒双基地雷达系统、伦敦大学工学院的双基地雷达试验系统；美国的“圣殿”（Sanctuary）防空双基地雷达系统、战术双基地雷达验证（TBIRD）系统、双基地报警和指示（BAC）系统、多基地测量系统（MMS）；俄罗斯的 Barrier 雷达等。

非合作式无源定位系统采用机会发射源，其中包括电视、广播等民用辐射源。英、美、德等国的非合作式无源定位系统研究处于世界领先水平：英国主要研究基于电视信号的探测系统，美国主要研究基于调频广播信号的目标探测系统，德国主要研究基于手机信号的目标探测系统。主要包括的种类有：（1）基于广播信号的无源定位系统。（2）基于电视信号的无源定位系统。（3）基于卫星信号的无源定位系统。（4）基于手机信号的无源雷达。

根据观测平台的数量，无源定位系统可以分为单站、双站和多站无源定位。单站定位系统具有成本较低、系统复杂度低、机动灵活等特点^{[5][7]}，其缺陷在于要求对辐射源目标进行连续跟踪和较长时间的观测。双站和多站定位系统具有较高的定位精度和较快的定位速度，但是系统的同步性和各站之间的协调性要求比较高^[8-9]，对大量数据的传输和实时通信任务具有更高的要求。

按照量测的参数不同，定位系统又可以有不同的划分。常用的测量参数有：目标方位角（Bearing）、信号到达时间差（Time Difference of Arrival, TDOA）、到达频差（Frequency Difference of Arrival）、基站与目标之间的多普勒频移（Doppler shift）、信号的相位差变化率（Phase difference rate of change）等。量测参数之间可以进行不同的组合构成不同的定位跟踪体制。

本文重点研究双站无源定位跟踪技术，由于辐射源到两个基站的距离不同，信号传播到基站处会产生时间差；由于辐射源与基站间的相对运动会产生多普勒效应（对于窄带信号可以近似为多普勒频移），信号传播到基站处会产生频差。时差频差定位克服了单站无源定位的缺陷，与测角交叉相比具有较高的定位精度，

且最少只需要双站就可以完成，定位体制较好。通过联合估计时差和频差，在基站坐标和速度已知的情况下，利用时差和频差可以联立处定位方程进行解算，可以求出辐射源的二维坐标信息，实现对固定辐射源的定位，利用卡尔曼滤波方程，实现对运动辐射源的跟踪定位。

1.2 常见的无源定位体制

1.2.1 单站无源定位

在不考虑特殊情况时，运动的无源单站定位系统通常对静止目标和运动目标都有着一定的可观性^[10]。静止的无源单站定位系统对运动目标也具备可观性，而当单站定位系统和目标均处于静止状态或者均处于匀速直线运动状态时，系统不具备可观性。

单站无源定位系统常用的观测量是方位角，利用方位角及其变化确定目标与观测站之间的距离 r 。根据距离 r 的求解方法的不同，可以分为以下两种单站定位体制：单站切向运动定位和单站径向运动定位^[1]。

（1）单站切向运动定位

运用观测站和辐射源之间的相对切向运动。由运动学原理可以得到相对运动的切向速度 v_t 等于相对径向距离 r 与相对旋转角速率 $\dot{\beta}$ 的乘积。根据观测信号的方向 β 及其变化率 $\dot{\beta}$ 可以实现运动单站对静止目标的无源定位，如图 1-1 所示， v_x 和 v_y 分别表示观测站在二维平面坐标系中横轴与纵轴上的速度分量。利用角度变化率估计目标距离的公式为： $r = \frac{v_t}{\dot{\beta}} = (v_y \cdot \sin\beta - v_x \cdot \cos\beta) / \dot{\beta}$ ，在方位角和相对距离已知的情况下，坐标可以被解算出来。方位角变化率可以利用观测站携带的双干涉仪天线获取信号的相位差变化率 $\dot{\phi}$ 来估计，结合上述利用单站切向运动来定位的方法称为单站相位差变化率定位法^[11]。

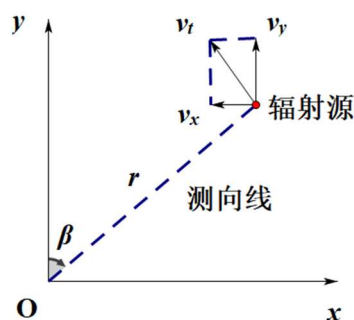


图 1-1 单站定位图

(2) 单站径向运动定位

运用观测站与辐射源之间的相对旋转运动，相对向心加速度 a_r 等于相对切向运动速度 v_t 的平方除以相对距离 r ，公式为 $a_r = \frac{v_t^2}{r} = (r\dot{\beta})^2/r$ ，得到径向距离 $r = a_r/\dot{\beta}^2$ ，再根据向心加速度 a_r 进行无源测距。若目标处于静止状态但观测站运动状态已知， a_r 的计算仅与辐射源方位角 β 及其变化率 $\dot{\beta}$ 有关，因此，根据方位角和方位角变化率可以计算出相对向心加速度和相对距离，在已知方位角的情况下即可对目标进行无源定位。

1.2.2 双站无源定位

观测平台数量的增多使得双站定位在参数选择上比单站定位更加丰富和灵活。双站定位系统观测站之间的基线距离越长定位精度就越高^[12]，然而在实际的应用中双站之间的间距通常会被限制在一定的距离之内，这样可以保障观测站之间的有效通信。根据不同的目的和实际情况，需要我们在两者之间进行权衡。常见的双站定位体制有：

(1) 双站测向交叉定位^[13]

考虑如图 1-2 所示的二维平面定位问题，观测站分别测量辐射源信号到各站的方位角。测向线的交点即为辐射源位置的估计值，但由于角度观测噪声的存在，通常会导致交点误差较大。

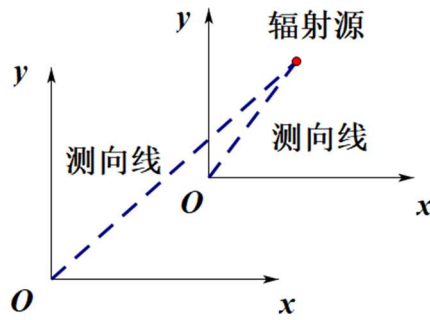


图 1-2 双站测向定位图

(2) 双站测向时差定位^[14]

双站定位系统从信号中提取出辐射源信号到观测站之间的到达时间差,并结合方位角可以实现无源定位。测向时差定位法将测向线与时差线的交点坐标作为辐射源位置的估计值,其定位原理图如图 1-3 所示。

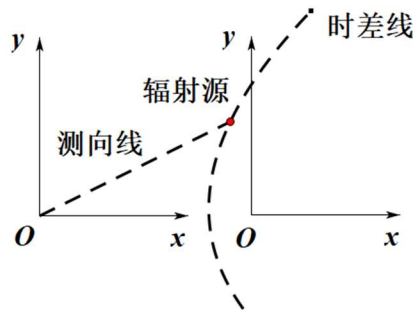


图 1-3 双站测向时差定位图

(3) 双站时差频差定位

频差是双站以及多站定位常用的观测量之一,指的是辐射源信号到各观测站的多普勒频移之差,频差是由目标与各观测站之间的相对运动引起的。若目标处于静止状态,那么只有运动的观测站才能从辐射源信号中提取出非零频差作为观测量。在目标运动速度已知的情况下,频差观测量仅是目标位置的函数,其观测方程与二维平面的频差线对应,如图 1-4 所示。而对于静止目标来说,时差频差定位通过将时差线和频差线相交,得到辐射源的位置估计。文献[15]中提到,频差的作用等价于测量辐射源到双站的距离,时差的作用等价于测量辐射源到双站的方位角。

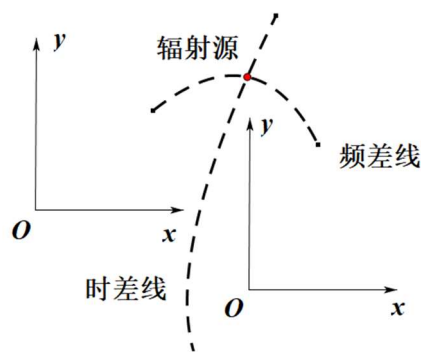


图 1-4 双站时差频差定位图

1.2.3 多站定位体制

以上的单站和双站定位体制为多站定位提供了理论和技术基础,前文介绍的所有定位体制均可作为多站定位的一部分。多站定位的优势在于:1.定位体制比较灵活,可以使用多种不同的测量参数,不同的观测量之间进行组合可以形成多种不同的定位方法;2.在相同观测噪声的情况下,定位误差随着观测站数量的增多而降低;3.相比于单站和双站定位,多站定位系统的可观测性大大提高[16]。

如图 1-5 所示,多站时差定位[17]是一种被广泛使用的多站定位体制。另外还有多站时差频差联合定位[18]、多站仅测角定位[19]等多种定位体制。

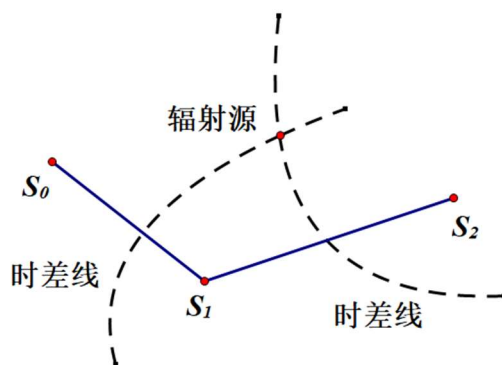


图 1-5 多站时差定位图

1.3 时差频差定位技术研究现状

针对辐射源信号,以时差频差测量为核心的定位体制因其定位精度比较高而获得很多关注。国内外也有很多很多文献和书籍探讨了时差频差定位的问题,主要包括以下两个方面:1.参数估计;2.定位解算。

1.3.1 时差频差参数估计

时差频差参数估计指的是利用无源定位系统接收到的辐射源信号,采用参数估计的方法,从中提取出辐射目标运动状态信息中的时差和频差的过程^[20]。

对于时差频差估计,最经典并且有效的方法仍然是互模糊函数。C.H.Knapp 和 G.C.Carter 在文献[21]中首先提出互相关函数估计时差的方法,并证明互相关函数是时差参数的最大似然估计。S.Stein 则在文献[22]中首先提出了用互模糊函数估计时差频差的方法,并证明互模糊函数法是时差频差参数估计的最大似然估计。对于窄带信号,文献[23~26]对互模糊函数进行了深入的研究。对于宽带信号,文献[27]提出了一种时差与时间伸缩因子联合估计的快速算法。文献[4]针对窄带脉冲串信号,运用了统计直方图法、时差序列法来改进互模糊函数法,实现了窄带脉冲串信号时差频差的快速估计;针对宽带脉冲串信号,运用频域分割,提出了分离估计法及频域累积互模糊函数方法,实现了时差与时间多普勒差联合估计。

在以上的基础上,发展出了很多延伸出来的方法,如文献[28]和文献[29]采用了三阶和四阶等高阶累积量的思想,在噪声分布为高斯分布且噪声之间具有相关性时可以有效消除噪声对定位参数估计的影响。文献[30]利用信号的循环平稳特性来进行参数估计,这可以适用于非平稳但呈现循环平稳特性的信号。文献[31]和文献[32]利用了分数阶统计量来描述定位参数的估计问题。文献[33]和文献[34]提出了自适应滤波的思想,文献[35]提出了一种基于交互小波变换的宽带模糊函数求解方法。然而在实际的应用中,噪声通常可以用零均值的加性高斯白噪声来描述,噪声之间往往是相互独立的,并且和辐射源无关。并且大多数的信号都是宽平稳的。所以,基于二阶统计量的参数估计方法仍然是时差频差估计的最有效的方法。

对于互模糊函数法,由于其需要在时间和频率两个维度上进行搜索,计算量比较大,因此研究的重点在于计算量的减少。文献[24]和[26]均提出了各自的补偿搜索算法,但计算量仍没有明显降低。文献[36]和文献[37]中将联合估计问题中的二维搜索彻底变成了两个独立的一维搜索,降低了运算量。文献[20]中对比分析了时差频差估计精度的理论下限,推导了实际被动定位系统中的时差频差参数估计理论下界,提出了时差频差参数估计的高效分维实现算法。

1.3.2 时差频差定位算法及其误差分析

在已知信号时差频差、接收站的位置和速度的情况下,对辐射源的定位就是一个求解非线性方程组的问题。时差频差定位计算方法主要有:搜索法、数值迭代和解析法等几种典型的方法。搜索法将辐射源位置状态在全空间内进行搜索,计算量较大;迭代法将观测方程在初始值附近进行一阶泰勒展开,得到线性化方程组,利用最小二乘估计得到对初始值的修正量,反复进行迭代,最终得到结果,其运算量较小,初始值对定位性能影响较大;解析法主要是将非线性方程进行数学变换,形成包括位置坐标的伪线性方程组,从而得到位置解算结果。

文献[38]和文献[39]通过引入额外的变量,将非线性方程组转化为线性方程组。当频差来源于接收机运动的前提下,文献[40]中推导了仅利用一组时差频差数据,对已知高度的静止辐射源的定位算法。文献[41]中推导了在超定情况下的定位算法,文献[42]中给出了利用时差频差数据对运动辐射源进行定位问题的代数解。文献[43]采用高维伸缩分析,提出了一种利用时差频差进行运动目标定位的新架构。

定位精度分析具有两个方面的意义:1.当观测量及参数的测量精度已经确定时,通过定位精度分析,判断定位精度能否达到预定的指标;2.从定位精度指标反推出达到某一定位精度所需的观测量及参数测量精度[20]。在无源定位中,通常用定位误差的几何稀释(GDOP: Geometrical Dilution of Precision)^[44]来描述定位误差。文献[45]给出了时差频差参数测量误差对双星几何定位误差的理论分析。文献[46]和文献[47]中对观测参数存在误差的情况下双星定位误差的GDOP分布进行了理论和蒙特卡洛方针分析,得到了定位误差的解析表达式。

1.3.3 双站时差频差定位

当今研究时差频差定位应用于星载平台的文献较多,而专门研究其在双站或者双机平台上应用的文献还相对较少。文献[48]给出了双机时差频差定位的模型,推导了定位误差的解析表达式,并比较了时差定位与测角定位的误差,显示了时差频差定位的优良特性。文献[49]中给出了典型场景下,不同时差频差及载机航向速度情况下的定位 GDOP 分布图,分析了各个因素对定位精度的影响。在此基础上,文献[50]建立了双机时差频差定位的系统模型,给出了状态方程和量测方程,推导了双机时差频差定位的扩展卡尔曼(EKF)滤波算法,通过仿真验证了算法的有效性。文献[51]使用一阶泰勒展开,运用迭代的方法进行时差频差定位,并且给出了定位算法的收敛性仿真和性能分析,文献[52]利用基于极坐标的时差频差方程,得到了无模糊测距解析解。

1.4 论文主要内容和章节安排

本文的主要内容与结构如下:

第一章概述了无源定位跟踪技术的研究背景,简要介绍了几种常用的无源定位跟踪体制,随后对时差频差估计和定位算法的研究现状进行了相关的介绍和分析,最后阐述了本文的章节安排。

第二章详细介绍了双站测向时差定位和双站时差频差定位的解算过程,推导了误差的 GDOP 分布,并对各自的误差分布进行了仿真,主要通过改变测量参数的精度来观察误差分布的变化情况,选取出对于两种定位体制较为重要的几组测量参数,最后介绍了几种常用的解模糊方法,本章主要是为了第三章和第四章内容的展开做好铺垫。

第三章将深度神经网络模型运用到了双站时差频差定位体制中,首先介绍了 BP 神经网络模型的运算流程,然后详细推导了反向传播算法,分析了 BP 神经网络算法的不足之处,并由此引出了深度神经网络模型,继而简要介绍了深度神经网络中常用的几种正则化和优化方法,最后对该算法进行了仿真建模,与传统的解方程定位方法进行了对比,验证了该模型算法的有效性。

第四章研究了无源定位系统中对辐射源信号的时差频差参数估计问题。首先建立了两路信号的数学模型，计算了时差频差估计的理论下限，然后介绍了两种参数估计的方法：一种主要采用二维联合搜索，然后利用二次曲面拟合的方法进行细化，得到参数的精估值；另一种是将时差和频差分维处理，首先估计出频差值，然后利用频差补偿，估计出时差值。最后对两种算法进行了仿真，验证了两种算法的有效性，并且通过对比得到了分维估计算法在运算量上的优越性。

第五章总结了全文，展望了下一步的研究方向。

第二章 双站测向时差频差定位

2.1 双站测向时差定位

2.1.1 定位原理

双站测向时差定位是利用辐射源目标到两个双站的方位角和信号到达两个基站的时差作为观测量进行定位求解。双站测向时差定位的示意图如图2-1所示。

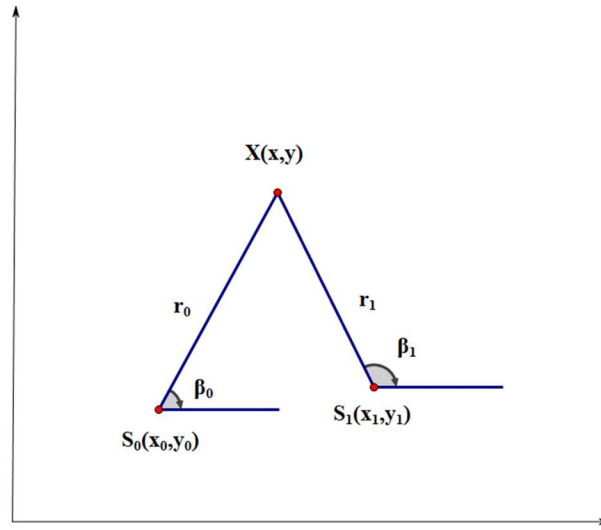


图 2-1 双站测向时差定位示意图

已知两个基站的坐标分别为 $S_0(x_0, y_0)$ 和 $S_1(x_1, y_1)$ ，信号传播速度为 c ，测量得到的目标信号到达两个基站的时间差为 Δt ，目标与两个基站的水平夹角分别为 β_0 和 β_1 ，通过这些已知数据求得目标的坐标 $X(x, y)$ 。

目标到两个基站的距离分别为：

$$r_1 = \sqrt{(x - x_1)^2 + (y - y_1)^2} \quad (2-1)$$

$$r_0 = \sqrt{(x - x_0)^2 + (y - y_0)^2} \quad (2-2)$$

得到关于时差的方程为：

$$c\Delta t = \Delta r = \sqrt{(x - x_1)^2 + (y - y_1)^2} - \sqrt{(x - x_0)^2 + (y - y_0)^2} \quad (2-3)$$

两个夹角的方程分别是：

$$\tan \beta_0 = \frac{y - y_0}{x - x_0} \quad (2-4)$$

$$\tan \beta_1 = \frac{y - y_1}{x - x_1} \quad (2-5)$$

对时差方程进行化简，并与公式(2-4)合并为方程组 $\mathbf{A}\mathbf{X} = \mathbf{F}$ 。

$$\text{其中 } \mathbf{A} = \begin{bmatrix} x_0 - x_1 & y_0 - y_1 \\ \tan \beta_0 & -1 \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} x \\ y \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} k + r_0 \Delta r \\ \tan \beta_0 x_0 - y_0 \end{bmatrix},$$

$$k = \frac{1}{2}[\Delta r^2 + (x_0^2 + y_0^2) - (x_1^2 + y_1^2)]。$$

可以求得目标的坐标为 $\mathbf{X} = \mathbf{A}^{-1}\mathbf{F}$ 。

2.1.2 定位误差分析

对公式(2-4)两边求微分并化简得到：

$$d\beta_0 = -\frac{\sin^2 \beta_0}{y - y_0} dx + \frac{\cos^2 \beta_0}{x - x_0} dy + \frac{\sin^2 \beta_0}{y - y_0} dx_0 - \frac{\cos^2 \beta_0}{x - x_0} dy_0 \quad (2-6)$$

距离差方程为 $\Delta r = \sqrt{(x - x_1)^2 + (y - y_1)^2} - \sqrt{(x - x_0)^2 + (y - y_0)^2}$ ，同样等式两边求微分并化简得到：

$$d\Delta r = \left(\frac{x - x_1}{r_1} - \frac{x - x_0}{r_0}\right)dx + \left(\frac{y - y_1}{r_1} + \frac{y - y_0}{r_0}\right)dy + \left(\frac{x - x_0}{r_0} dx_0 + \frac{y - y_0}{r_0} dy_0\right) - \left(\frac{x - x_1}{r_1} dx_1 + \frac{y - y_1}{r_1} dy_1\right) \quad (2-7)$$

转变为矩阵方程形式为

$$dV = \mathbf{C}d\mathbf{X} + \mathbf{U}dS_0 - \mathbf{W}dS_1 \quad (2-8)$$

$$\text{其中 } \mathbf{C} = \begin{bmatrix} -\frac{\sin^2 \beta_0}{y - y_0} & \frac{\cos^2 \beta_0}{x - x_0} \\ \frac{x - x_1}{r_1} - \frac{x - x_0}{r_0} & \frac{y - y_1}{r_1} + \frac{y - y_0}{r_0} \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} \frac{\sin^2 \beta_0}{y - y_0} & -\frac{\cos^2 \beta_0}{x - x_0} \\ \frac{x - x_0}{r_0} & \frac{y - y_0}{r_0} \end{bmatrix}, \quad \mathbf{W} =$$

$$\begin{bmatrix} 0 & 0 \\ \frac{x - x_1}{r_1} & \frac{y - y_1}{r_1} \end{bmatrix}, \quad dV = \begin{bmatrix} d\beta_0 \\ d\Delta r \end{bmatrix}, \quad d\mathbf{X} = \begin{bmatrix} dx \\ dy \end{bmatrix}, \quad dS_0 = \begin{bmatrix} dx_0 \\ dy_0 \end{bmatrix}, \quad dS_1 = \begin{bmatrix} dx_1 \\ dy_1 \end{bmatrix}。$$

进一步解得目标的误差矩阵为

$$P_{dX} = E(d\mathbf{X}d\mathbf{X}^T) = \mathbf{C}^{-1}[\mathbf{R}_v + \mathbf{U}\mathbf{R}_{S_0}\mathbf{U}^T + \mathbf{W}\mathbf{R}_{S_1}\mathbf{W}^T]\mathbf{C}^{-T} \quad (2-9)$$

其中 $\mathbf{R}_v = E(dVdV^T) = \begin{bmatrix} \sigma_\beta^2 & 0 \\ 0 & \sigma_{\Delta r}^2 \end{bmatrix}$ 为测量参数误差矩阵, $\mathbf{R}_v = E(dVdV^T) = \begin{bmatrix} \sigma_\beta^2 & 0 \\ 0 & \sigma_{\Delta r}^2 \end{bmatrix}$ 为测量参数误差矩阵, $\mathbf{R}_{S_0} = E(dS_0dS_0^T) = \begin{bmatrix} \sigma_S^2 & 0 \\ 0 & \sigma_S^2 \end{bmatrix}$ 为基站 S_0 的位置误差矩阵, $\mathbf{R}_{S_1} = E(dS_1dS_1^T) = \begin{bmatrix} \sigma_S^2 & 0 \\ 0 & \sigma_S^2 \end{bmatrix}$ 为基站 S_1 的位置误差矩阵。

σ_β 为角度均方根误差, $\sigma_{\Delta r} = c\sigma_{\Delta t}$ 为时差均方根误差, σ_S 为位置均方根误差。

几何精度因子为

$$GDOP = \sqrt{P_{dx}(1,1) + P_{dx}(2,2)} \quad (2-10)$$

2.1.3 误差仿真分析

本文使用 MATLAB 软件对上述定位算法的误差分析公式进行了数字仿真, 方针基本参数如下:

双站基线长度: 10km

时差测量误差: 20ns

方位角测量误差: 3mrad

基站位置测量误差: 5m

以上仿真参数中, 误差值均为高斯分布中的标准差。基站位置误差在每个坐标轴上相互独立, 上述值为每一维的标准差。仿真中将双站的位置坐标分别设为 $(-5, 0)$ km 与 $(5, 0)$ km。

得到上述场景下的辐射源定位 GDOP 值分布如图 2-2 所示。

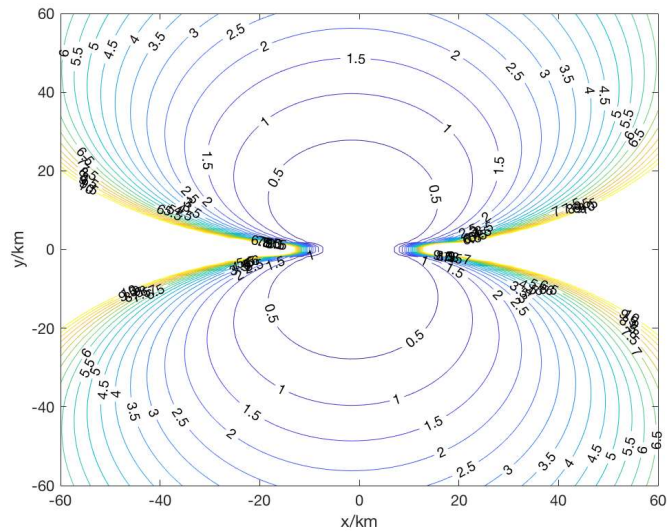
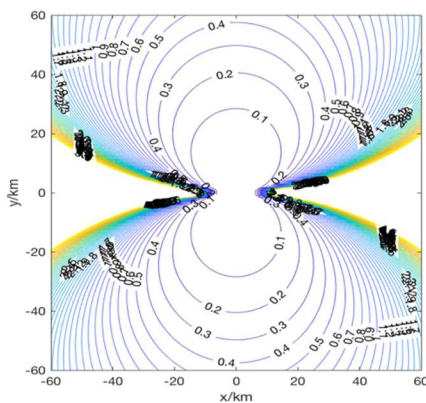
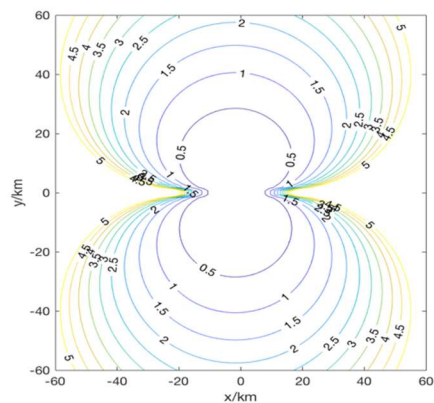


图 2-2 双站测向时差定位 GDOP 分布图

(1) 改变仿真参数，分别仅考虑时差测量误差、方位角测量误差与位置测量误差（仅考虑一个参数测量误差时，其余 2 个参数测量误差为零），得到这 3 种情况下的 GDOP 分布如图 2-3、2-4 所示。



(a) 仅考虑时差测量误差



(b) 仅考虑方向测量误差

图 2-3 仅考虑时差/方向测量误差时的 GDOP 分布图

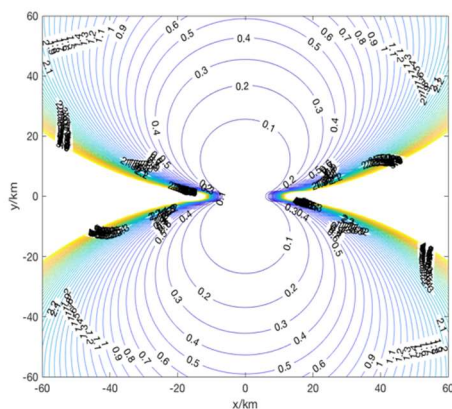


图 2-4 仅考虑基站位置测量误差时的 GDOP 分布图

从上述仿真结果可以看出，在本节开始时的仿真参数条件下，方位角的测量误差对定位结果的影响最大，其次是时差和基站位置的测量误差。

(2) 减小基站的基线距离，由原来的 10km 减小到 1km，得到仿真对比如图 2-5。

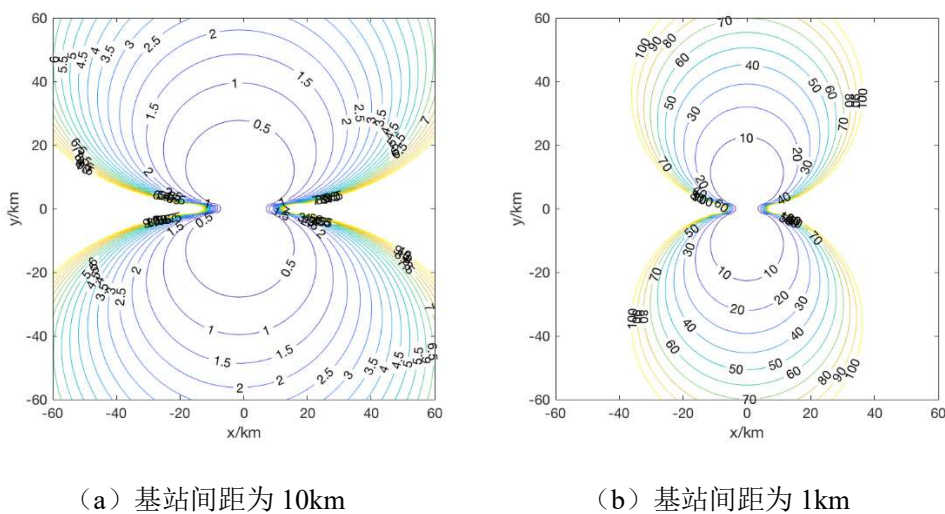


图 2-5 不同基站间距的 GDOP 分布对比图

从上面的仿真结果可以看出，基线距离越大，定位精度越高。

2.2 双站时差频差定位

2.2.1 定位原理

由于辐射源到两个基站的距离不同，信号传播到基站处会产生时间差；由于辐射源与基站间的相对运动会产生多普勒效应（对于窄带信号可以近似为多普勒频移），信号传播到基站处会产生频差。双站时差频差定位示意图如图 2-6 所示。

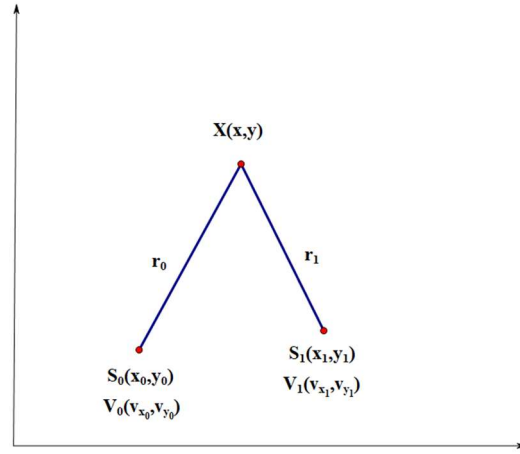


图 2-6 双站时差频差定位示意图

已知两个基站的坐标分别为 $S_0(x_0, y_0)$ 和 $S_1(x_1, y_1)$ ，两个基站的运动速度分别为 $V_0(v_{x_0}, v_{y_0})$ 和 $V_1(v_{x_1}, v_{y_1})$ ，信号传播速度为 c ，测量得到的目标信号到达两个基站的时间差为 Δt ，信号到达两个基站的频率变化差为 Δf_d ，通过这些已知数据求得目标的坐标 $X(x, y)$ 。

目标到两个基站的距离分别为：

$$r_1 = \sqrt{(x - x_1)^2 + (y - y_1)^2} \quad (2-11)$$

$$r_0 = \sqrt{(x - x_0)^2 + (y - y_0)^2} \quad (2-12)$$

得到关于时差的方程为：

$$c\Delta t = \Delta r = \sqrt{(x - x_1)^2 + (y - y_1)^2} - \sqrt{(x - x_0)^2 + (y - y_0)^2} \quad (2-13)$$

对上式两边求微分得到：

$$d\Delta r = dr_1 - dr_0 = \frac{-(x - x_1)dx_1 - (y - y_1)dy_1}{\sqrt{(x - x_1)^2 + (y - y_1)^2}} - \frac{-(x - x_0)dx_0 - (y - y_0)dy_0}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} \quad (2-14)$$

频率变化差相关方程为

$$d\Delta r = (f_{d1} - f_{d0})\lambda = \Delta f_d \frac{c}{f_0} \quad (2-15)$$

两式进行合并得到：

$$-\frac{c\Delta f_d}{f_0} = \frac{(x-x_1)v_{x1} + (y-y_1)v_{y1}}{\sqrt{(x-x_1)^2 + (y-y_1)^2}} - \frac{(x-x_0)v_{x0} + (y-y_0)v_{y0}}{\sqrt{(x-x_0)^2 + (y-y_0)^2}} \quad (2-16)$$

此处不妨设两个基站的运动方向和速度是一致的，即 $v_{x0} = v_{x1} = v_x$ ， $v_{y0} = v_{y1} = v_y$ ，并将时差和频率差方程合并为一个方程组，其矩阵形式为

$$\mathbf{A}\mathbf{X} = \mathbf{F} \quad (2-17)$$

其中 $\mathbf{A} = \begin{bmatrix} x_0 - x_1 & y_0 - y_1 \\ v_x & v_y \end{bmatrix}$ ， $\mathbf{X} = \begin{bmatrix} x \\ y \end{bmatrix}$ ， $\mathbf{F} = \begin{bmatrix} k_1 + \Delta r \cdot r_0 \\ k_2 r_0^2 + k_3 r_0 + k_4 \end{bmatrix}$ ， $k_1 = \frac{1}{2}[\Delta r^2 + (x_0^2 + y_0^2) - (x_1^2 + y_1^2)]$ ， $k_2 = \frac{c\Delta f_d}{f_0 \Delta r}$ ， $k_3 = \frac{(x_0 - x_1)v_x + (y_0 - y_1)v_y}{\Delta r} + \frac{c\Delta f_d}{f_0 \Delta r}$ ， $k_4 = x_0 v_x + y_0 v_y$ 。

1.如果 \mathbf{A} 可逆，通过 $\mathbf{X} = \mathbf{A}^{-1}\mathbf{F}$ 可以求得 x 和 y 关于的 r_0 表达式，再将结果代入方程 $r_0 = \sqrt{(x-x_0)^2 + (y-y_0)^2}$ 中，得到关于 r_0 的一元四次方程，可以解出目标的坐标，并带有模糊解。

2.如果 \mathbf{A} 不可逆

(1) $v_x \neq 0$ ， $v_y \neq 0$ ， $x_0 - x_1 \neq 0$ ， $y_0 - y_1 \neq 0$ ，即两个基站运动方向和连线方向一致，则可以设 $\frac{v_x}{x_0 - x_1} = \frac{v_y}{y_0 - y_1} = q$ ，将其带入公式(2-18)中，解出目标的坐标和模糊解。

$$\begin{bmatrix} x_0 - x_1 & y_0 - y_1 \\ v_x & v_y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} k_1 + \Delta r \cdot r_0 \\ k_2 r_0^2 + k_3 r_0 + k_4 \end{bmatrix} \quad (2-18)$$

(2) 特殊地， $v_x \neq 0$ ， $v_y = 0$ ， $x_0 - x_1 \neq 0$ ， $y_0 - y_1 = 0$ ，可以直接解出 $x = \frac{k_1 + \Delta r \cdot r_0}{x_0 - x_1}$ ， $y = y_0 \pm \sqrt{r_0^2 - (x - x_0)^2}$ ，再代入 $r_0 = \sqrt{(x - x_0)^2 + (y - y_0)^2}$ 中，解得目标坐标并带有模糊解。

(3) 或者 $v_x = 0$, $v_y \neq 0$, $x_0 - x_1 = 0$, $y_0 - y_1 \neq 0$, 同样可以直接解出 $y = \frac{k_1 + \Delta r \cdot r_0}{y_0 - y_1}$, $x = x_0 \pm \sqrt{r_0^2 - (y - y_0)^2}$, 再代入 $r_0 = \sqrt{(x - x_0)^2 + (y - y_0)^2}$ 中, 解得目标坐标并带有模糊解。

因此, 不管 \mathbf{A} 是否可逆, 解得的坐标都有模糊解, 解模糊的方法有很多, 比如可以增加方向测量等, 具体参见 2.4 节。

2.2.2 定位误差分析

将等式 $c\Delta t = \Delta r = \sqrt{(x - x_1)^2 + (y - y_1)^2} - \sqrt{(x - x_0)^2 + (y - y_0)^2}$ 两边进行微分, 化简得到

$$d\Delta r = \left(\frac{x - x_1}{r_1} - \frac{x - x_0}{r_0}\right)dx + \left(\frac{y - y_1}{r_1} - \frac{y - y_0}{r_0}\right)dy + \left(\frac{x - x_0}{r_0}dx_0 + \frac{y - y_0}{r_0}dy_0\right) - \left(\frac{x - x_1}{r_1}dx_1 + \frac{y - y_1}{r_1}dy_1\right) \quad (2-19)$$

记 $\Delta v_r = -\frac{c\Delta f_d}{f_0}$, 同样将频差等式两边进行微分, 得到

$$(a_1 - a_0)dx + (b_1 - b_0)dy = d\Delta v_r - (c_1dx_1 + d_1dy_1) + (c_0dx_0 + d_0dy_0) + (e_0 - e_1)dv_x + (f_0 - f_1)dv_y \quad (2-20)$$

其中 $a_i = \frac{v_x r_i^2 - (x - x_i)^2 v_x - (x - x_i)(y - y_i)v_y}{r_i^3}$, $b_i = \frac{v_y r_i^2 - (x - x_i)(y - y_i)v_x - (y - y_i)^2 v_y}{r_i^3}$, $c_i = \frac{(x - x_i)^2 v_x + (x - x_i)(y - y_i)v_y - v_x r_i^2}{r_i^3}$, $d_i = \frac{(y - y_i)^2 v_y + (x - x_i)(y - y_i)v_x - v_y r_i^2}{r_i^3}$, $e_i = \frac{x - x_i}{r_i}$, $f_i = \frac{y - y_i}{r_i}$, $i = 0, 1$ 。

将两个等式合并为矩阵方程形式, 得到

$$\mathbf{C}d\mathbf{X} = d\mathbf{Z} + \mathbf{U}d\mathbf{S}_1 + \mathbf{W}d\mathbf{S}_0 + \mathbf{V}d\mathbf{v} \quad (2-21)$$

其中 $\mathbf{C} = \begin{bmatrix} a_1 - a_0 & b_1 - b_0 \\ e_1 - e_0 & f_1 - f_0 \end{bmatrix}$, $\mathbf{U} = \begin{bmatrix} -c_1 & -d_1 \\ e_1 & f_1 \end{bmatrix}$, $\mathbf{W} = \begin{bmatrix} c_0 & d_0 \\ -e_0 & -f_0 \end{bmatrix}$, $\mathbf{V} = \begin{bmatrix} e_0 - e_1 & f_0 - f_1 \\ 0 & 0 \end{bmatrix}$, $d\mathbf{X} = \begin{bmatrix} dx \\ dy \end{bmatrix}$, $d\mathbf{Z} = \begin{bmatrix} d\Delta v_r \end{bmatrix}$, $d\mathbf{S}_1 = \begin{bmatrix} dx_1 \\ dy_1 \end{bmatrix}$, $d\mathbf{S}_0 = \begin{bmatrix} dx_0 \\ dy_0 \end{bmatrix}$, $d\mathbf{v} = \begin{bmatrix} dv_x \\ dv_y \end{bmatrix}$ 。

则定位误差矩阵为

$$\mathbf{P}_{d\mathbf{X}} = E(d\mathbf{X}d\mathbf{X}^T) = \mathbf{C}^{-1}(\mathbf{R}_Z + \mathbf{U}\mathbf{R}_{S_1}\mathbf{U}^T + \mathbf{W}\mathbf{R}_{S_0}\mathbf{W}^T + \mathbf{V}\mathbf{R}_v\mathbf{V}^T)\mathbf{C}^{-T} \quad (2-22)$$

其中时差和频率差误差矩阵为 $\mathbf{R}_z = \begin{bmatrix} \sigma_{\Delta v_r}^2 & 0 \\ 0 & \sigma_{\Delta r}^2 \end{bmatrix}$, 基站坐标误差矩阵为 $\mathbf{R}_{s_1} = \begin{bmatrix} \sigma_s^2 & 0 \\ 0 & \sigma_s^2 \end{bmatrix}$, $\mathbf{R}_{s_0} = \begin{bmatrix} \sigma_s^2 & 0 \\ 0 & \sigma_s^2 \end{bmatrix}$, 速度误差矩阵为 $\mathbf{R}_v = \begin{bmatrix} \sigma_{v_x}^2 & 0 \\ 0 & \sigma_{v_y}^2 \end{bmatrix}$ 。

几何精度因子为

$$GDOP = \sqrt{P_{dx}(1,1) + P_{dx}(2,2)} \quad (2-23)$$

2.2.3 误差仿真分析

本文应用 MATLAB 软件对上述定位算法的误差分析公示进行了数字仿真, 仿真基本参数如下:

信号载频: 2GHz

双站运动速率: 150m/s

基线长度: 10km

时差测量误差: 20ns

频差测量误差: 5Hz

基站位置测量误差: 5m

基站速度测量误差: 5m/s

以上仿真参数中, 误差均为高斯分布中的标准差。双站的位置、速度误差在坐标轴各个方向上相互独立, 上述值为每一维的误差。仿真中将双站的位置坐标分别设为 $(-5, 0)$ km 与 $(5, 0)$ km, 双站的运动速度均为 $(150, 0)$ m/s, 即均向 x 轴右侧运动。

得到上述场景下的辐射源定位 GDOP 值分布如图 2-7 所示。

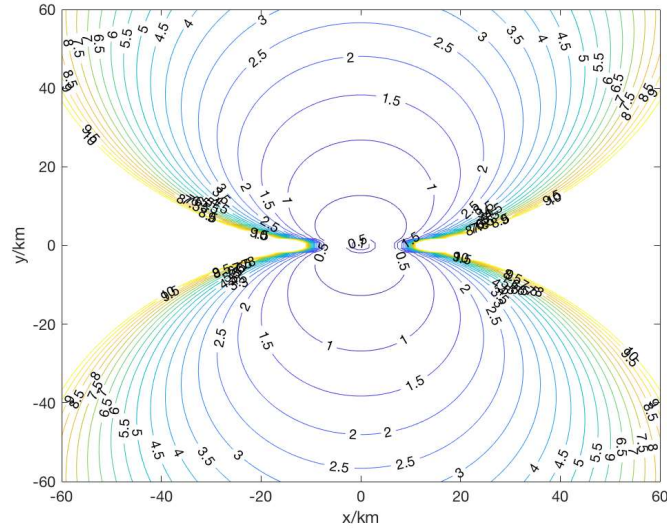
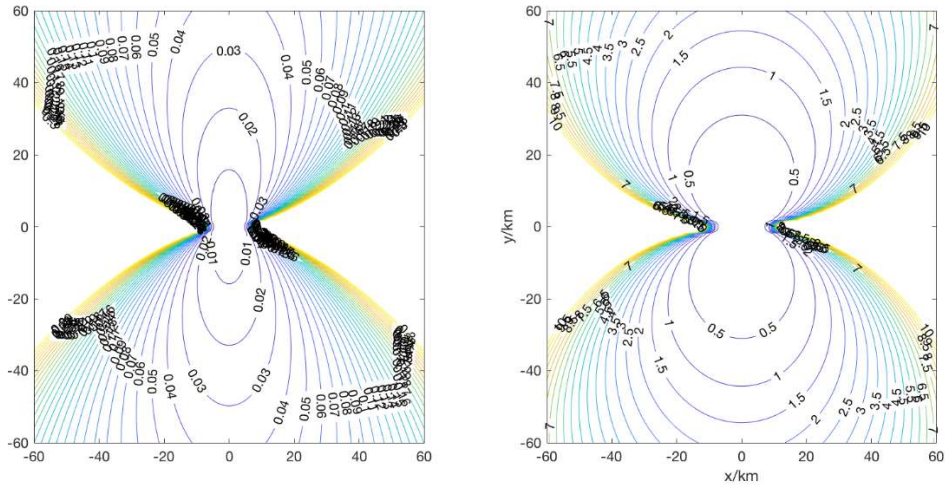


图 2-7 双站时差频差定位 GDOP 分布图

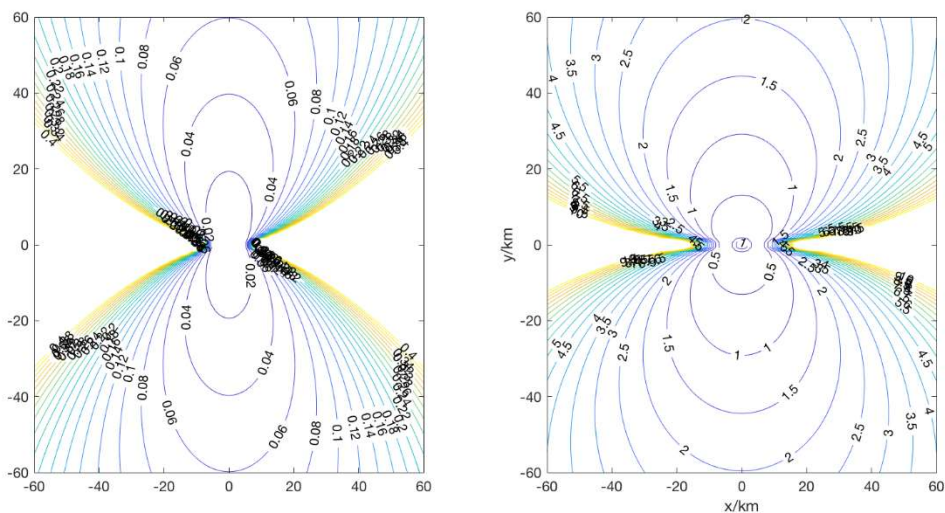
(1) 改变仿真参数，分别仅考虑时差测量误差、频差测量误差、位置测量误差和速度测量误差(仅考虑一个参数测量误差时，其余 3 个参数测量误差为零)，得到这 4 中情况下的 GDOP 分布如图 2-8、2-9 所示。



(a) 仅考虑时差测量误差

(b) 仅考虑频差测量误差

图 2-8 仅考虑时差/频差测量误差时的 GDOP 分布图



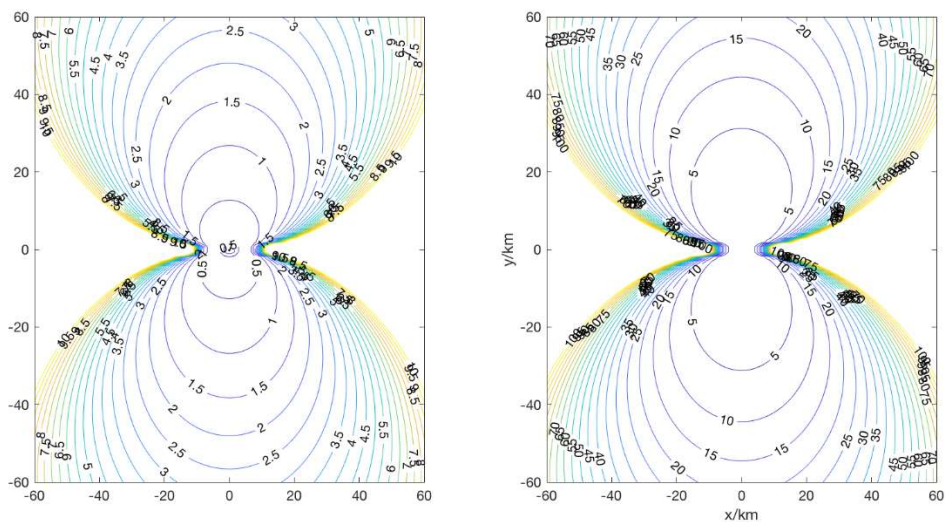
(a) 仅考虑位置测量误差

(b) 仅考虑速度测量误差

图 2-9 仅考虑位置/速度测量误差时的 GDOP 分布图

通过上述仿真结果可以看出，在本节开始时设定的仿真参数条件下，频率测量误差和速度测量误差对定位结果的影响较大，时差测量误差和位置测量误差的影响较小。

(2) 减小基站的基线距离，由原来的 10km 减小到 1km，得到仿真对比如图 2-10。



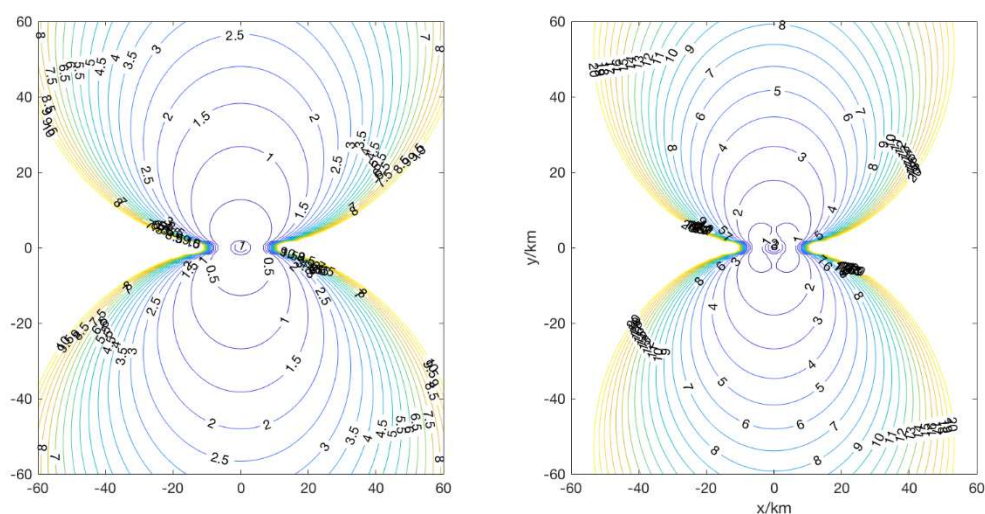
(a) 基站间距为 10km

(b) 基站间距为 1km

图 2-10 不同基站间距的 GDOP 分布对比图

通过上述仿真可以看出，基线距离越大，定位精度越高。

(3) 减小基站的运动速率，由原来的 150m/s 减小为 50m/s，得到仿真对比如图 2-11。



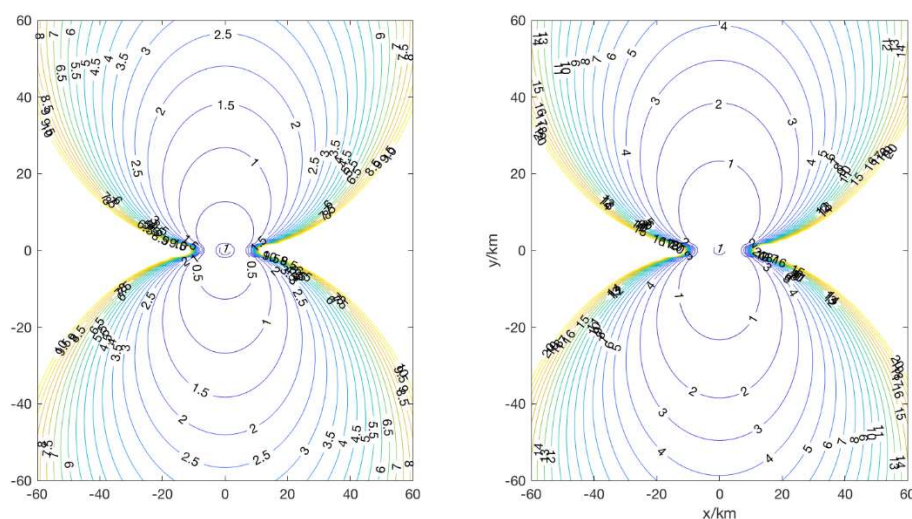
(a) 基站运动速率为 150m/s

(b) 基站运动速率为 50m/s

图 2-11 不同运动速率的 GDOP 分布对比图

从上面的仿真图像可以看出，基站的运动速率越大，定位精度越高。

(4) 减小信号载频，由原来的 2GHz 减小为 1GHz，得到仿真对比如图 2-12。



(a) 信号载频为 2GHz

(b) 信号载频为 1GHz

图 2-12 不同信号载频的 GDOP 分布对比图

从上面的仿真图像可以看出，信号载频越大，定位精度越高。

(5) 上面均为两个基站一前一后同向运动时的定位 GDOP 图，改变仿真场景设置，将双站的运动速度分别改为 $(0, 150)$ m/s 和 $(150, 150)$ m/s，得到定位 GDOP 分布图如图 2-13 所示。

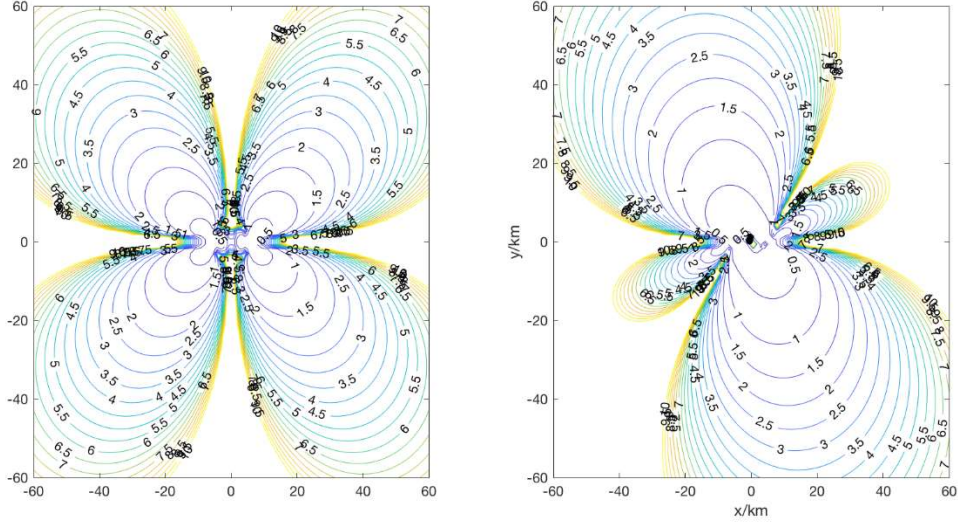


图 2-13 不同基站运动方向的 GDOP 分布对比图

从上面的仿真图像可以比较看出，在基站的运动方向和连线方向不同时，时差频差定位的 GDOP 分布会有很大差别。因此如何合理优化双站时差频差定位系统中双站的运动轨迹，使得对某一点或某一区域的定位精度更高也是一个需要解决的问题。

2.3 双站定位解模糊

对于 2.1 节中的双站方向时差定位不存在定位模糊问题，这里主要研究双站时差频差定位中的定位模糊问题。当基站和辐射源之间存在相对运动时，同时可以得到时差和频差定位方程：

$$c\Delta t = \Delta r = \sqrt{(x-x_1)^2 + (y-y_1)^2} - \sqrt{(x-x_0)^2 + (y-y_0)^2} \quad (2-24)$$

$$-\frac{c\Delta f_d}{f_0} = \frac{(x-x_1)v_{x1} + (y-y_1)v_{y1}}{\sqrt{(x-x_1)^2 + (y-y_1)^2}} - \frac{(x-x_0)v_{x0} + (y-y_0)v_{y0}}{\sqrt{(x-x_0)^2 + (y-y_0)^2}} \quad (2-25)$$

仿真场景与 2.2.3 中相同，双站时差频差二维定位示意图如下图所示。

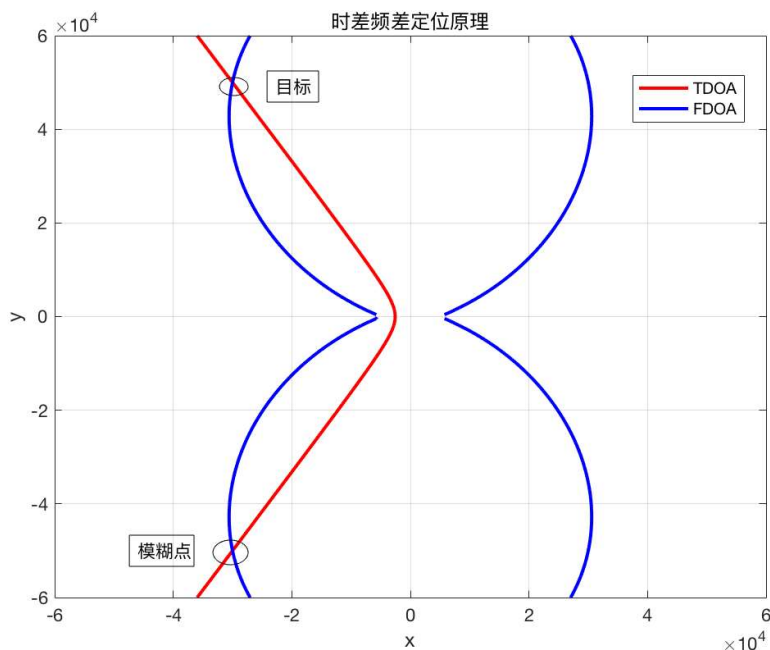


图 2-14 双站时差频差定位原理图

从图中可以看出，在 x 轴坐标对侧存在一个和目标相对应的定位模糊点。

对于消除模糊点的方法，主要有以下几种^[53]：

1.辅助测量方位角法。即增加一个测向信息，通过测向线与上图中的两个定位曲线进行相交，距离交点较近的那个即为辐射源目标。例如文献[54]中增加了测向信息，通过先验信息和工程实际消除了模糊解。

2.反代入法。将计算得到的目标和模糊点都反代入定位方程中，用误差最小的原则去除增根。

3.多次定位法。由于基站和目标之间存在相互运动，可以用多次定位的方法绘制出运动轨迹，根据运动轨迹去除增根。

4.根据探测区域进行解模糊。如文献[55]中根据模糊点与目标点的位置不同，分成了对称模糊和独立模糊，主要根据探测区域进行解模糊。如上图模糊点和目标位于基线两侧，如果探测区域仅仅是基线上方，则可以把下方的模糊点去除。

另外，文献[56]讨论了平面中三站时差定位系统的模糊分布取决于三站的布局，基线的长度与基线的夹角大小决定了该系统的模糊区。对于时差频差定位而言也有类似的结论。

2.4 本章小结

本章研究了双站测向时差和双站时差频差定位算法与误差分析问题。首先，推导了两种定位算法方程的求解过程，然后进行了误差分析，通过仿真 GDOP 分布图，得到对于定位而言较为重要的几个测量参数，为后续章节的展开奠定了基础，最后简要介绍了几种在工程上比较常用的解模糊方法。

第三章 基于深度神经网络的双站时差频差定位

在得到目标辐射源的时差频差定位数据之后,通过求解一组非线性定位方程即可得到辐射源的位置。但当获得了更多的测量信息时,方程定位解法没有办法利用这些多余的信息,比如在时差频差定位中,如果获得了目标相对于两个基站的方向信息,利用这些方向信息可以消除模糊点,但无法实现更精确的定位效果。另外,在实际的应用中,测量参数受噪声影响较大,导致定位不准,尤其当我们无法对噪声统计特性进行准确建模时,会导致很大的定位误差。

神经网络作为一种非线性动态数学模型,因为其广泛的适应能力和学习能力而被广泛应用与非线性系统数学建模、系统辨识和函数逼近等方面^[57]。神经网络模型通过对一组针对问题具有代表性的样本进行学习和训练,推断并建立满足这种输入和输出关系的网络模型。在本文中可以利用神经网络模型尤其时深度神经网络模型建立关于时差频差数据(输入)和目标位置坐标(输出)的较为精准的网络模型,通过所得数据对网络进行训练,得到基于深度神经网络模型的时差频差定位模型。加强多源信息的利用,提高定位精度和鲁棒性。

神经网络模型大多被用在室内或者无线定位领域,比如文献[58~61]运用神经网络将接收到的信号强度与接收距离之间的非线性关系进行拟合,然后利用接收距离和位置坐标的方程关系进行求解。另外,文献[62]则将测量得到的时差和方向信息输入到神经网络中进行修正,然后利用定位算法进行求解。这些方法大多被运用在室内的无线定位中,空间范围较小,所使用的神经网络的层数也较少。与采用全连接神经网络不同的是,文献[63~66]使用的是径向基函数(Radial Basis Function, RBF)神经网络,场景和测量参数也比较简单,然而在较为复杂的环境中径向基函数神经网络这种浅层结构的建模能力有限,其拟合效果就明显不如深层的神经网络[67~68]。同时,上述几乎所有的文献都没有使用相应的泛化方法,仅仅用训练时的样本点作为测试,得到的结果不具有说服力。

本章将使用深度神经网络,将样本点的多种测量信息输入到神经网络中,通过改变输入参数的个数神经网络的层数,并且运用正则化方法进行泛化,找到了

最佳的超参数组合，并且与单层的全连接神经网络和 RBF 神经网络的定位精度进行了对比，仿真实验证明该深度神经网络模型可以提高定位精度和鲁棒性。

3.1 神经网络理论基础

3.1.1 神经网络结构

神经元是神经网络的基本处理单元，它是利用数学方法对生物神经元的简化与模拟。神经元一般是一个多输入单输出的非线性器件，大量简单神经元相互联结构成了神经网络。

图 3-1 是一个含有 1 个输入层、1 个隐藏层和 1 个输出层的神经网络示意图。其中输入层的神经元有 3 个，依次将其编号为 1、2、3；隐藏层的神经元有 4 个，依次将其编号为 4、5、6、7；输出层的神经元有 2 个，依次将其编号为 8、9。这个神经网络为一个全连接的神经网络，每个节点都和上一层的所有节点有连接。比如隐藏层的节点 4 和输入层的三个节点 1、2、3 都有连接，其连接上的权重分别为 w_{41} ， w_{42} ， w_{43} 。节点 4 的输出值 a_4 的计算公式为：

$$a_4 = f(w_{41}x_1 + w_{42}x_2 + w_{43}x_3 + w_{4b}) \quad (3-1)$$

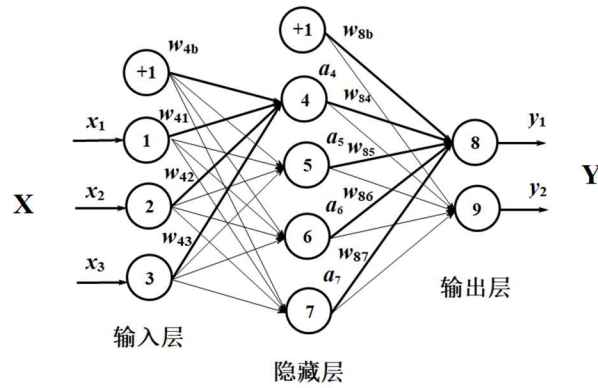


图 3-1 神经网络示意图

上式中的 w_{4b} 为节点 4 的偏置项， f 代表神经网络的激活函数，具体见 3.1.2 节。同样，通过计算出节点 5、6、7 的输出值 a_5 ， a_6 ， a_7 ，节点 8 的输出值 y_1 的计算公式为：

$$y_1 = f(w_{84}a_4 + w_{85}a_5 + w_{86}a_6 + w_{87}a_7 + w_{8b}) \quad (3-2)$$

3.1.2 激活函数

激活函数（Activation transfer function）是一个神经网络的核心。网络解决问题的能力与功效除了与网络结构有关，在很大程度上取决于网络采用的激活函数。神经网络的线性或非线性是由网络神经元中激活函数的线形或非线性决定的。

激活函数的作用主要有：1.控制输入对输出的激活作用；2.对输入、输出进行函数转换；3.将可能无限域的输入变换成指定的有限范围内的输出。

常用的激活函数类型包括：

1.Sigmoid 函数

Sigmoid 函数的公式形式为

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (3-3)$$

它的图形如图 3-2 所示

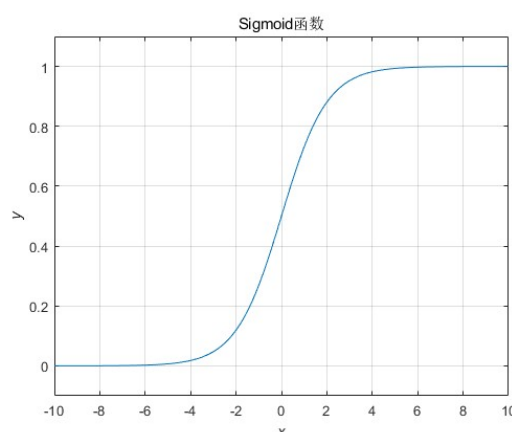


图 3-2 Sigmoid 函数示意图

Sigmoid 函数将输入为实值的数压缩到 0~1 的范围内，较大的负数被映射成 0，较大的正数被映射成 1，Sigmoid 函数在历史上流行过一段时间，因为它能够很好表达“激活”的意思，未激活为 0，完全的饱和激活就是 1。但是在近些年 Sigmoid 函数已经不怎么常用了，主要因为它有以下两个缺点：

（1）Sigmoid 函数容易饱和，并且当输入非常大或者非常小的时候，神经元的梯度接近 0，从图中可以看出梯度的变化趋势。这就使得在计算反向传播时得到接近 0 的梯度，导致权重几乎没有更新，我们就无法递归去学习数据。另外要注意参数初始值的设定，要避免初始值的设定落到饱和状态，如果初始值比较

大，大部分的神经元可能都会处于饱和状态而使得梯度无法更行，导致网络变得很难学习。

(2) Sigmoid 函数不是 0 均值的，这是我们不希望出现的状态，因为这会导致后面的神经元的输入是非 0 均值的信号，会对梯度产生影响：假设后层神经元的输入都为正数，那么对权重求局部梯度就都为正数，这样在反向传播的过程中权重要么都往正方向更新，要么都往负方向更新，导致存在这样一种类似的捆绑效果，使得收敛很慢。当然，如果是按照批量去训练，那么每个批量可能得到不同的符号，相加一下后此问题可以得到缓解。因此，非 0 均值的问题虽然会产生一定的影响，不过相比上面的第一个问题来说影响较小。

2.Tanh 函数

Tanh 函数和 Sigmoid 函数有着类似的效果，它的图像如图 3-3 所示

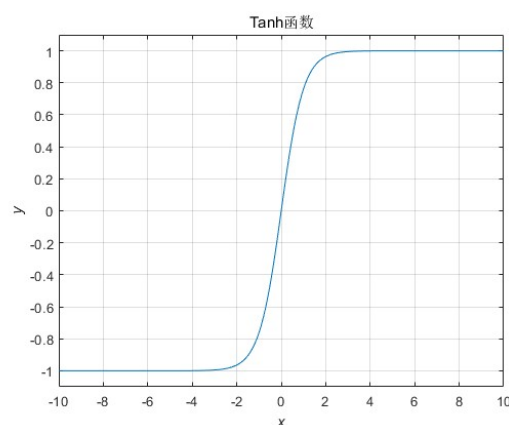


图 3-3 Tanh 函数示意图

不同的是它把输入的实值压缩到了-1~1 的范围，因此它基本上是 0 均值的，也就解决了上述 Sigmoid 函数中的第二个缺点，所以实际中会比 Sigmoid 函数更常用。但是它也存在梯度饱和的问题，从数学变换中可以看出，Tanh 是 Sigmoid 的变形：

$$\tanh(x) = 2\sigma(2x) - 1 \quad (3-4)$$

3.ReLU 函数

近年来 ReLU 函数变得越来越受欢迎，它的数学表达式为：

$$f(x) = \max(0, x) \quad (3-5)$$

其图像如图 3-4 所示

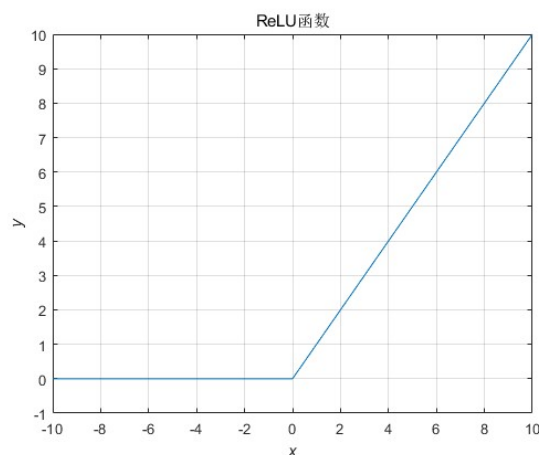


图 3-4 ReLU 函数示意图

从图中可以明显看出，当输入信号小于 0 时，输出为 0，当输入信号大于 0 时，输出等于输入。ReLU 函数的优点如下：

（1）文献中发现使用 ReLU 函数得到的批量梯度下降的收敛速度会比 Tanh 函数和 Sigmoid 函数快很多。

（2）相比于 Tanh 函数和 Sigmoid 函数而言，ReLU 函数只需要一个阈值就可以得到激活值，而不用去算一大堆复杂的运算。

当然 ReLU 函数也有它的缺点，就是训练的时候会导致梯度更新很容易“死亡”，比如一个非常大的梯度流过一个 ReLU 神经元，更新过参数之后，这个神经元再也不会对任何数据有激活现象了，那么这个神经元的梯度就永远都是 0。

在实际的操作中，如果设置的学习速率过大，那么很有可能神经网络中的 40% 的神经元都会处于上面所说的永远不会再更新的状态。

4. Leaky-ReLU、P-ReLU、R-ReLU 函数

（1）Leaky-ReLU 就是用来解决上面所说的 ReLU 函数的缺点的，与 ReLU 函数不同的是，Leaky-ReLU 的数学表达式如下：

$$f(x) = \begin{cases} \alpha x, & (x < 0) \\ x, & (x \geq 0) \end{cases} \quad (3-6)$$

这里的 α 是一个很小的常数，这样不仅修正了数据分布，又保留了一些负轴的值，使得负轴的信息不会被全部丢失。

但是，关于 Leaky-ReLU 的效果还没有清晰的定论，有人用了 Leaky-ReLU 做实验发现效果很好，而有些实验的效果则刚好相反。

(2) 对于 Leaky-ReLU 中的 α ，通常是通过先验知识人工进行赋值的，然而可以观察到，损失函数对 α 的导数我们是求得的，论文[69]指出，我们可以将它作为一个参数进行训练，而且效果更好，提高了 1.03%。

(3) R-ReLU 是 Leaky-ReLU 的一个随机版本，即 α 是随机指定的，核心思想为：在训练过程中， α 是一个服从高斯分布的随机值，然后再测试过程中进行修正。

以上只是介绍了主要的激活函数，实际上还有很多被提出来的改进的激活函数，它们大部分都是对以上几种函数的修正，因此不再表述。那么如何去选择激活函数就成为了一个问题，这个问题没有定论，只能靠一些实际的经验和具体的场景进行对比测试，从中选出一个表现最好的函数作为激活函数。

3.2 深度神经网络模型

3.2.1 BP 神经网络算法原理

BP 神经网络模型利用已经得到的数据和反向传播算法对神经网络中的权重进行学习，以一个三层的网络为例，并且假设它的激活函数为 Sigmoid 函数，我们首先直观地介绍反向传播算法和它的结论，然后再来介绍这个算法的推导过程，它的学习规则和过程如下。

假设每个训练样本为 (\vec{x}, \vec{t}) ，其中向量 \vec{x} 是训练样本的特征，而 \vec{t} 是样本的目标值。首先可以根据 3.1.1 节中介绍的算法，用样本的特征 \vec{x} 计算出神经网络中每个隐藏层节点的输出 a_i ，以及输出层每个节点的输出 y_i 。然后按照下面的计算方法计算出每个节点的误差项 δ_i ：

对于输出层节点 i ，

$$\delta_i = y_i(1 - y_i)(t_i - y_i) \quad (3-7)$$

其中， δ_i 是节点 i 的误差项， y_i 是节点 i 的输出值， t_i 是样本对应于节点 i 的目标值。举个例子，根据图 3-1，对于输出层节点 8 来说，它的输出值是 y_1 ，而样本的目标值是 t_1 ，带入上面的公式得到节点 8 的误差项 δ_8 应该是：

$$\delta_8 = y_1(1 - y_1)(t_1 - y_1) \quad (3-8)$$

对于隐藏层节点：

$$\delta_i = a_i(1-a_i) \sum_{k \in \text{outputs}} w_{ki} \delta_k \quad (3-9)$$

其中， a_i 是节点*i*的输出值， w_{ki} 是节点*i*到它的下一层节点*k*的连接权重， δ_k 是节点*i*的下一层节点*k*的误差项。例如，对于隐藏层节点 4 来说，计算方法如下：

$$\delta_4 = a_4(1-a_4)(w_{84}\delta_8 + w_{94}\delta_9) \quad (3-10)$$

最后，更新每个连接上的权值：

$$w_{ji} \leftarrow w_{ji} + \eta \delta_j x_{ji} \quad (3-11)$$

其中， w_{ji} 是节点*i*到节点*j*的权重， η 是一个被称为学习速率的常数， δ_j 是节点*j*的误差项， x_{ji} 是节点*i*传递给节点*j*的输入。例如，权重的更新方法如下：

$$w_{84} \leftarrow w_{84} + \eta \delta_8 a_4 \quad (3-12)$$

类似的，权重 w_{41} 的更新方法如下：

$$w_{41} \leftarrow w_{41} + \eta \delta_4 x_1 \quad (3-13)$$

偏置项的输入值永远为 1，例如节点 4 的偏置项 w_{4b} 应该按照下面的方法计算：

$$w_{4b} \leftarrow w_{4b} + \eta \delta_4 \quad (3-14)$$

上面介绍了神经网络每个节点误差项的计算和权重更新方法。显然，计算一个节点的误差项，需要先计算每个与其相连的下一层节点的误差项。这就要求误差项的计算顺序必须是从输出层开始，然后反向依次计算每个隐藏层的误差项，直到与输入层相连的那个隐藏层。这就是反向传播算法的名字的含义，当所有节点的误差项计算完毕后，我们可以根据式来更新所有的权重。

下面推导反向传播算法，反向传播算法其实就是链式求导法则的应用，按照机器学习的步骤，先确定神经网络的目标函数，然后用随机梯度下降优化算法去求目标函数最小值时的参数值。

取网络所有输出层节点的误差平方和作为目标函数：

$$E_d \equiv \frac{1}{2} \sum_{i \in \text{outputs}} (t_i - y_i)^2 \quad (3-15)$$

其中 E_d 表示样本 d 的误差。

然后使用随机梯度下降算法对目标函数进行优化：

$$w_{ji} \leftarrow w_{ji} - \eta \frac{\partial E_d}{\partial w_{ji}} \quad (3-16)$$

随机梯度下降算法也就是需要求出误差 E_d 对于每个权重 w_{ji} 的偏导数（也就是梯度），其中权重 w_{ji} 仅能通过节点 j 的输入值影响网络的其它部分，设 net_j 是节点 j 的加权输入，即

$$net_j = \overline{w_j} \cdot \overline{x_j} = \sum_i w_{ji} x_{ji} \quad (3-17)$$

E_d 是 net_j 的函数，而 net_j 是 w_{ji} 的函数，根据链式求导法则可以得到：

$$\frac{\partial E_d}{\partial w_{ji}} = \frac{\partial E_d}{\partial net_j} \frac{\partial net_j}{\partial w_{ji}} = \frac{\partial E_d}{\partial net_j} \frac{\partial \sum_i w_{ji} x_{ji}}{\partial w_{ji}} = \frac{\partial E_d}{\partial net_j} x_{ji} \quad (3-18)$$

上式中 x_{ji} 是节点 i 传递给节点 j 的输入值，也就是节点 i 的输出值。

对于 $\frac{\partial E_d}{\partial net_j}$ 的推导，分为输出层和隐藏层两种情况：

（1）对于输出层来说， net_j 仅能通过节点 j 的输出值 y_j 来影响网络的其它部分，也就是说 E_d 是 y_j 的函数，而 y_j 是 net_j 的函数，其中 $y_j = \text{sigmoid}(net_j)$ 。所以可以再次使用链式求导法则：

$$\frac{\partial E_d}{\partial net_j} = \frac{\partial E_d}{\partial y_j} \frac{\partial y_j}{\partial net_j} \quad (3-19)$$

考虑上式第一项：

$$\frac{\partial E_d}{\partial y_j} = \frac{\partial}{\partial y_j} \frac{1}{2} \sum_{i \in \text{outputs}} (t_i - y_i)^2 = \frac{\partial}{\partial y_j} \frac{1}{2} (t_j - y_j)^2 = -(t_j - y_j) \quad (3-20)$$

考虑上式第二项：

$$\frac{\partial y_j}{\partial net_j} = \frac{\partial \text{sigmoid}(net_j)}{\partial net_j} = y_j(1 - y_j) \quad (3-21)$$

将第一项和第二项带入，得到：

$$\frac{\partial E_d}{\partial net_j} = -(t_j - y_j) y_j (1 - y_j) \quad (3-22)$$

这就得到了公式(3-7)。

将上述推导带入随机梯度下降公式中，可以得到：

$$\begin{aligned}
w_{ji} &\leftarrow w_{ji} - \eta \frac{\partial E_d}{\partial w_{ji}} \\
&= w_{ji} + \eta(t_j - y_j)y_j(1 - y_j) \\
&= w_{ji} + \eta\delta_j x_{ji}
\end{aligned} \tag{3-23}$$

这就得到了公式(3-11)。

(2) 对于隐藏层来说首先定义节点 j 的所有直接下游节点的组合 $\text{Downstream}(j)$ 。例如，对于节点 4 来说，它的直接下游节点是节点 8 和节点 9，可以看到 net_j 只能通过影响 $\text{Downstream}(j)$ 来影响 E_d 。设是 net_k 节点 j 的下游节点的输入，则 E_d 是 net_k 的函数，而 net_k 是 net_j 的函数。因为有多多个 net_k ，此处运用全导数公式，做出如下的公式推导：

$$\begin{aligned}
\frac{\partial E_d}{\partial \text{net}_j} &= \sum_{k \in \text{Downstream}(j)} \frac{\partial E_d}{\partial \text{net}_k} \frac{\partial \text{net}_k}{\partial \text{net}_j} \\
&= \sum_{k \in \text{Downstream}(j)} -\delta_k \frac{\partial \text{net}_k}{\partial \text{net}_j} \\
&= \sum_{k \in \text{Downstream}(j)} -\delta_k \frac{\partial \text{net}_k}{\partial a_j} \frac{\partial a_j}{\partial \text{net}_j} \\
&= \sum_{k \in \text{Downstream}(j)} -\delta_k w_{kj} \frac{\partial a_j}{\partial \text{net}_j} \\
&= \sum_{k \in \text{Downstream}(j)} -\delta_k w_{kj} a_j (1 - a_j) \\
&= -a_j (1 - a_j) \sum_{k \in \text{Downstream}(j)} \delta_k w_{kj}
\end{aligned} \tag{3-24}$$

因为 $\delta_j = -\frac{\partial E_d}{\partial \text{net}_j}$ ，带入上式得到：

$$\delta_j = a_j (1 - a_j) \sum_{k \in \text{Downstream}(j)} \delta_k w_{kj} \tag{3-25}$$

这就是公式(3-9)。

3.2.2 BP 神经网络的不足

BP 神经网络的限制和不足，主要表现在它训练过程的不确定性。具体表现为：

1. 训练时间过长

由于采用固定的学习速率，因此网络的收敛速度慢，需要较长的训练时间。对于一些复杂的问题或是大量的训练样本，BP 算法需要的训练时间可能会很长。

2.收敛到局部极小值

由于 BP 网络的训练采用梯度下降法，训练是从某一起点沿着误差函数的偏微分方向逐渐达到误差的极小值。对于复杂的网络，误差函数为多维空间中的一个凹凸不平的曲面，所以在网络训练过程中极有可能陷入某一个小的凹区，产生一个局部极小值^[57]。

3.网络隐含层的层数和单元数的选择缺乏理论上的指导

一般是根据经验选择或者多次实验进行确定。层数较浅的话网络表示能力就比较差，但如果层数较深的话，网络的学习时间和拟合能力都会有不确定性的变化。

4.网络的学习和记忆具有不确定性

若训练样本增加或者改变，训练好的网络就需要重新开始训练，对于以前训练得到的权值和阈值，网络是没有记忆的。

针对以上几点不足，3.3 节中提出了深度神经网络模型，利用深度学习中的各种优化算法和正则化的泛化方法，可以在一定程度上解决或者缓解上面的几点问题。

3.2.3 深度神经网络模型

1. 深度神经网络的模型特点

文献[70]表明，一个前馈神经网络如果具有线性输出层和至少一层具有“挤压”性质的激活函数（比如 sigmoid 函数）的隐藏层，只要给予足够数量的隐藏单元，它可以以任意精度来近似任何从一个有限维空间到另一个有限维空间的映射，前馈网络的导数也可以任意好地近似函数的导数^[71]。这个近似定理也被证明对于更广泛类别的激活函数也是适用的，其中就包括吸纳在常用的 ReLU 函数。然而不能保证训练算法能够学得这个函数，原因之一是用于训练的优化算法可能找不到用于期望函数的参数值，另外就是训练算法可能由于过拟合而选择了错误的参数。

近似定理表明，存在一个足够大的网络能够达到我们所希望的任意精度，但是并没有说明这个网络有多大。具有单层的前馈网络足以表示任何函数，但是网络层可能很大以至于不可实现，并且可能无法正确地学习和泛化。在很多情况下，

使用更深的模型能够减少表示期望函数所需要的单元的数量,并且可以减少泛化误差。因此,深度神经网络的研究重点就在于优化方法的选择和泛化误差的限制。

2. 深度神经网络的正则化

机器学习的一个核心问题是设计不仅在训练数据上表现好,而且能在新输入上泛化好的算法。在机器学习中,有许多被设计出来减少测试误差的方法,这些方法可能会以增大训练误差为代价,这些策略被统称为正则化。估计的正则化以偏差的增加换取方差的减少,一个有效的正则化方法能够显著减少方差而不过度增加偏差。在实践中,过于复杂的模型族不一定能够包括目标函数或真实数据的生成过程,甚至也并不包括近似过程。真实数据的生成过程有时候是未知的,它的生成过程有时候肯定会在模型之外。这就意味着我们需要找到一个合适复杂度的模型,这样一个最好的拟合模型应当是一个适当正则化的大型模型。

(1) 参数范数惩罚

正则化在深度学习出现之前就已经被使用了数十年,线性回归和逻辑回归都可以使用简单的正则化策略。许多正则化方法通过对目标函数 J 添加一个参数范数惩罚 $\Omega(\theta)$, 限制模型的学习能力, 将正则化后的目标函数记为 \tilde{J} :

$$\tilde{J}(\theta; \mathbf{X}, y) = J(\theta; \mathbf{X}, y) + \alpha \Omega(\theta) \quad (3-26)$$

其中 $\alpha \in [0, \infty)$ 是权衡范数惩罚项 Ω 和标准目标函数 $J(\mathbf{X}; \theta)$ 相对贡献的超参数, 将其设置为 0 表示没有正则化, 它的值越大对应正则化惩罚就越大。

L2 正则化和 L1 正则化是两种经常使用的参数范数惩罚方法, 其中 L2 正则化的正则项 $\Omega(\theta) = \frac{1}{2} \|\omega\|_2^2$, L1 的正则化项被定义为 $\Omega(\theta) = \|\omega\|_1$

(2) 提前终止

当训练有足够的表示能力甚至会过拟合的大型模型时, 经常会观察到训练误差会随着时间的推移逐渐降低但验证集的误差会再次上升, 图 是这种现象中的一个例子, 这种现象几乎是一定会出现的。这意味着我们如果返回使验证集误差最低的参数设置就可以获得更好的模型。在每次验证集误差有所改善后, 存储下模型的参数, 当训练算法终止时, 返回这些参数而不是最新的参数, 当验证集上的误差在事先指定的循环次数内没有进一步改善时, 算法就会终止。这种策略被称为提前终止, 它是深度学习中最常用的正则化形式, 并且十分简单有效。

这种方法可以被认为是一种高效的超参数选择算法，按照这种观点，训练步数仅是另一个超参数。从图中可以看到，这个超参数在验证集上具有 U 型性能曲线，很多控制模型容量的超参数在验证集上都是这样的 U 型性能曲线，在提前终止的情况下，我们通过拟合训练集的步数来控制模型的有效容量。大多数超参数的选择必须使用高代价的猜测和检查过程，需要在训练开始时猜测一个超参数，然后运行几个步骤检查它的训练效果。“训练时间”是唯一只要跑一次训练就能尝试很多值的超参数。通过提前终止自动选择超参数的唯一显著代价就是训练期间要定期评估验证集。提前终止的另一个额外代价是需要保持最佳的参数副本，但是这种代价一般是可以忽略的，因为可以将它存储在较慢较大的存储器上。

提前终止是一种非常不显眼的正则化形式，它几乎不需要改变基本训练过程、目标函数或一组允许的参数值，这就意味着不需要破坏学习动态就能很容易地使用提前终止。当然提前终止也可以与其他的正则化策略结合使用。

（3）Bagging 和其他集成方法

Bagging (bootstrap aggregating) 是通过结合几个模型降低泛化误差的技术，主要想法是分布训练几个不同的模型，然后让所有模型表决测试样例的输出，这是机器学习中常规策略的一个例子，被称为模型平均，采用这种策略的技术被称为集成方法。模型平均能够奏效的原因是不同的模型通常不会在测试集上产生完全相同的误差。

（4）Dropout

Dropout 是指在深度学习网络的训练过程中，对于神经网络单元，按照一定的概率将其暂时从网络中丢弃，对于随机梯度下降来说，由于是随即丢弃，故而对每一个小批量都在训练不同的网络。大规模的神经网络有两个缺点：1.费时；2.容易过拟合。为了解决过拟合问题，一般会采用集成方法，即训练多个模型做组合，此时费时就成为了一个大问题，不仅训练起来很费时，测试起来多个模型也很费时间。Dropout 的出现很好的解决了这个问题，每次做完 Dropout 相当于从原始的网络中找到一个更“瘦”的网络，如下图所示：

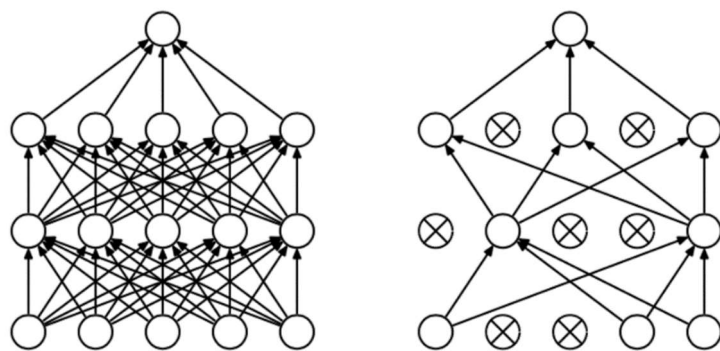


图 3-5 使用 Dropout 后网络对比示意图

Dropout 以概率 p 舍弃神经元并让其它神经元以概率 $q=1-p$ 保留，每个神经元被关闭的概率相同的。Dropout 提供了正则化一大类模型的方法，计算方便但功能强大，是大量深度神经网络的实用集成方法。

针对以上几种正则化方法进行分析。提前终止是通过观测训练集和测试集数据的学习曲线进行手动操作的方法，使用简单。文献[72]显示，Dropout 比范数正则化方法更有效，它的另一个优点是计算方便，训练过程中产生 n 个随机二进制数与状态相乘，每个样本每次更新只需 $O(n)$ 的计算复杂度， n 代表模型中的神经元总个数。另外文献[73]中将 Dropout 与其他的集成方法进行对比并得出结论：相比独立模型集成获得泛化误差，Dropout 会带来额外的改进，是对集成方法的一种近似，计算复杂度也较小。Dropout 的另一个显著优点是具有广泛的适用性，几乎在所有分布式表示且可以用随机梯度下降训练的模型以及各种场景中都表现很好。例如文献[74]将其运用在了贝叶斯神经网络中，文献[75]讨论了在循环神经网络中的使用情况，文献[76]研究了 Dropout 在长短期记忆网络中的应用。

无源定位通过对测量参数的运算得到目标的位置，其实质是一个在特定场景下将参数作为输入、将位置坐标作为输出的一个映射系统，又因为上文中提到的 Dropout 的各种优势，因此本文选择 Dropout 和提前终止作为正则化方法来提高深度神经网络的泛化性能。

3. 深度神经网络的学习优化算法

在一节中介绍了梯度下降，即沿着整个训练集的梯度方向下降。可以使用随机梯度下降法进行加速，沿着随机挑选的小批量数据的梯度下降方向，随机梯度下降及其变种是一般机器学习中应用最广泛的优化算法，特别是深度学习中。

随机梯度下降的整个过程如下所示：

Require: 学习率 ϵ

Require: 初始参数 θ

While 没有达到停止准则 do

从训练集中采包含 m 个样本 $\{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(m)}\}$ 的小批量, 其中 $\mathbf{x}^{(i)}$ 对应目标为 $\mathbf{y}^{(i)}$ 。

计算梯度估计: $\mathbf{g} \leftarrow \frac{1}{m} \nabla_{\theta} \sum_i L(f(\mathbf{x}^{(i)}; \theta), \mathbf{y}^{(i)})$

参数更新: $\theta \leftarrow \theta - \epsilon \mathbf{g}$

end while

虽然随机梯度下降是很受欢迎的优化算法, 但是它的学习过程有时候会比较慢, 因此文献[77]提出了一种动量方法用来加速学习, 特别是处理高曲率、小但一致的梯度或者是带噪声的梯度, 动量算法主要是积累了之前梯度指数级衰减的移动平均, 并且继续沿着这个方向移动。使用动量的随机梯度下降算法如下所示:

Require: 学习率 ϵ , 动量参数 α

Require: 初始参数 θ , 初始速度 \mathbf{v}

While 没有达到停止准则 do

从训练集中采包含 m 个样本 $\{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(m)}\}$ 的小批量, 其中 $\mathbf{x}^{(i)}$ 对应目标为 $\mathbf{y}^{(i)}$ 。

计算梯度估计: $\mathbf{g} \leftarrow \frac{1}{m} \nabla_{\theta} \sum_i L(f(\mathbf{x}^{(i)}; \theta), \mathbf{y}^{(i)})$

计算速度更新: $\mathbf{v} \leftarrow \alpha \mathbf{v} - \epsilon \mathbf{g}$

参数更新: $\theta \leftarrow \theta + \mathbf{v}$

end while

基于此, 文献[78]提出了动量算法的一个变种, 即 Nesterov 动量算法, 改进了额外误差收敛率。使用 Nesterov 动量的随机梯度下降算法步骤如下所示:

Require: 学习率 ϵ , 动量参数 α

Require: 初始参数 θ , 初始速度 \mathbf{v}

While 没有达到停止准则 do

从训练集中采包含 m 个样本 $\{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(m)}\}$ 的小批量，其中 $\mathbf{x}^{(i)}$ 对应目标为 $\mathbf{y}^{(i)}$ 。

参数临时更新： $\tilde{\theta} \leftarrow \theta + v$

计算梯度估计： $\mathbf{g} \leftarrow \frac{1}{m} \nabla_{\tilde{\theta}} \sum_i L(f(\mathbf{x}^{(i)}; \tilde{\theta}), \mathbf{y}^{(i)})$

计算速度更新： $v \leftarrow \alpha v - \epsilon \mathbf{g}$

参数更新： $\theta \leftarrow \theta + v$

end while

另外，模型中的学习率也是一个难以设置的超参数，近几年，大量学者提出了一些增量（或者基于小批量）的算法来自适应模型参数的学习率，主要有以下几种：

1. AdaGrad 算法

独立适应所有模型的学习率，缩放每个参数反比于其所有梯度历史平方总和的平方根。具有损失最大偏导的参数相应地有一个快速下降的学习率，而具有小偏导的参数在学习率上有相对较小的下降。效果就是在参数空间中更为平缓的倾斜方向会取得更大的进步。

2. RMSProp 算法

RMSProp 算法对 AdaGrad 加以修改使其在非凸设定下效果更好，改变梯度积累为指数加权的移动平均。经验上，RMSProp 已经被证明是一种有效并且实用的深度神经网络优化算法，目前这种算法是深度学习从业者经常采用的优化方法之一。

3. Adam 算法

Adam 算法是另一种学习率自适应的优化算法，它将动量直接并入了梯度一阶矩（指数加权）的估计，其次它修正了从原点初始化的一阶矩（动量项）和（非中心）二阶矩的估计。通常被认为对超参数的选择相当鲁棒，尽管学习率有时候需要从建议的默认修改。

然而，对于上述算法的选择这一点上，目前并没有达到共识，文献[79]展示了许多优化算法在大量学习任务上极具价值的比较，虽然结果表明具有自适应学

习率的算法族表现得相当鲁棒，不分高下，但没有哪个算法能够脱颖而出，因此算法的选择主要还是取决于具体的问题以及超参数的调节。

3.3 基于深度神经网络的定位仿真分析

3.3.1 模型训练

在-60km 到 60km 的方形区域内以均匀分布的概率随机选取 10000 个点的坐标值(x, y)，信号载频为 1GHz；双站的坐标分别为（-5km，0）、（5km，0），即基线距离为 10km；双站的速度都为（150m/s，0），即沿着 x 轴向右运动。然后算出时差值 Δt 、频差值 Δf_d 和辐射源相对于两个基站的方向值 β_0 、 β_1 作为测量参数。为了模拟噪声环境，分别对时差加上误差为 10ns、对频差加上误差为 1Hz、对方向加上误差为 3mrad 的高斯分布随机数。最后得到有噪声的 10000 个样本点以及其对应的测量参数，然后运用 式对输入的所有参数包括时差、频差和方向值进行归一化处理。

$$x^* = \frac{x - \mu}{\sigma} \quad (3-27)$$

式中 μ 为样本均值， σ 为样本数据的标准差。

然后将这 10000 个样本随机分为三组：训练样本组（8000 个），交叉验证样本组（1000 个），测试样本组（1000 个）。

损失函数设置为均方根误差（RMSE）：

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N [(x - x_{\text{output}})^2 + (y - y_{\text{output}})^2]} \quad (3-28)$$

式中 x 和 y 分别是样本中的真实坐标， x_{output} 和 y_{output} 分别是神经网络的输出坐标值。

然后将经过归一化处理的测量参数输入到神经网络中，参数个数依次为 2（ Δt 、 Δf_d ）、3（ Δt 、 Δf_d 、 β_0 ）、4（ Δt 、 Δf_d 、 β_0 、 β_1 ）。每个隐藏层的神经元个数设置为 100 个，得到训练结果如图 所示。

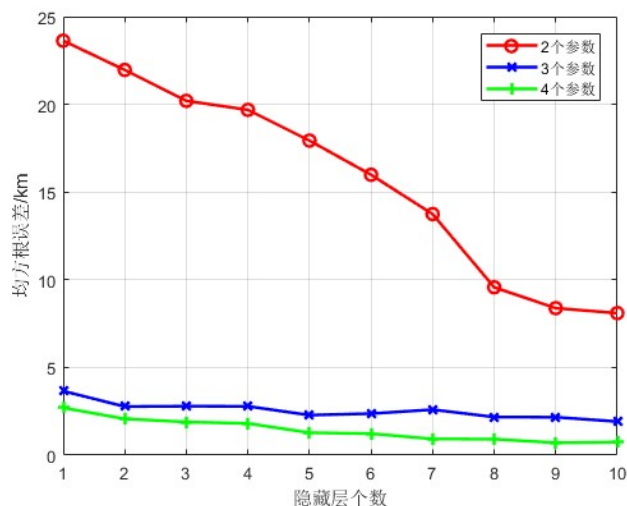


图 3-6 训练组的 RMSE 值随隐藏层个数增加的变化曲线

从图 中可以看出：（1）随着隐藏层个数的增多，每条曲线的整体趋势是先下降后趋于变化缓慢，说明当隐藏层的个数过多时，神经网络已经学习不到样本的特征。（2）随着输入参数的增多，RMSE 值在逐渐减小，当输入参数由 2 个增加到 3 个时下降尤其明显，证明了神经网络对多源信息的综合利用。

3.3.2 模型泛化

在训练的过程中，可以记录下训练组和交叉验证组数据的 RMSE 值随着迭代次数的增加而变化的曲线图，以输入参数为 4 个、隐藏层个数为 7 个的情况为例，如图 所示。

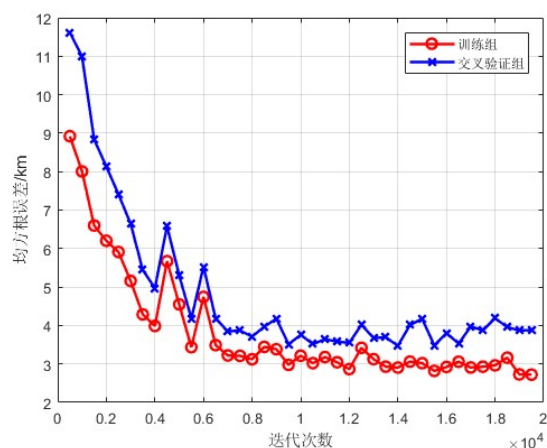


图 3-7 训练组和交叉验证组的 RMSE 值随迭代次数增加的变化曲线

从图 中可以看出，随着迭代次数的增加，训练组和交叉验证组的 RMSE 值都在一直降低，但是训练组的 RMSE 值一直都比交叉验证组的 RMSE 值小，说明训练出来的模型在训练组上表现很好，但对于未训练的数据而言表现较差，因此需要运用正则化方法进行泛化。

根据图 中训练组中输入参数为 4 个的情况，当隐藏层个数在 1~6 时，RMSE 的值都在减小，而在 7~10 个隐藏层时，RMSE 的值变化缓慢，因此将 7~10 作为隐藏层个数的备选值。Dropout 是一个 0~1 的概率数字，过大或过小都会反而令测试集的误差升高，尝试将 0~0.5 作为 Dropout 这一维度的备选值。另一方面，将迭代次数作为第三个备选值，即在每一次迭代时如果交叉验证组的 RMSE 值减小，则保存下当前模型的参数值，并更新当前交叉验证组的 RMSE 最小值，从图 中可以看出当迭代次数达到 20000 次时模型的表现基本稳定，因此将迭代次数的最大值定为 20000 次。得到交叉验证组的 RMSE 值与隐藏层个数和 Dropout 值的关系表。

表 3-1 交叉验证组 RMSE 值随隐藏层个数及 Dropout 值不同的变化情况/km

| Dropout 值 | 7 层 | 8 层 | 9 层 | 10 层 |
|-----------|------|------|------|------|
| 0 | 3.28 | 3.27 | 3.91 | 2.69 |
| 0.01 | 3.05 | 3.23 | 3.28 | 3.37 |
| 0.02 | 2.68 | 3.73 | 3.01 | 3.26 |
| 0.03 | 3.51 | 3.97 | 4.56 | 4.15 |
| 0.04 | 3.89 | 4.06 | 4.37 | 4.21 |
| 0.05 | 3.54 | 4.29 | 5.15 | 4.38 |
| 0.1 | 3.38 | 3.46 | 4.12 | 4.81 |
| 0.15 | 3.11 | 3.45 | 3.76 | 4.92 |
| 0.2 | 3.36 | 3.55 | 3.74 | 4.80 |
| 0.3 | 3.74 | 3.70 | 4.18 | 5.12 |
| 0.4 | 5.39 | 4.85 | 4.88 | 5.21 |
| 0.5 | 5.73 | 6.01 | 7.69 | 7.67 |

从表 3-1 中可以看出，当 Dropout 取较为合适的值时，交叉验证组的 RMSE 值是要比不采用 Dropout 方法时要更小，针对本课题，当隐藏层个数为 7，Dropout 值为 0.02 时交叉验证组的 RMSE 值最小，为 2.68km，选取当前保存下的模型参数值作为深度神经网络的最终参数。

3.3.3 模型测试仿真

将测试组中的 1000 个样本输入到 3.3.2 节中得到的最终模型中，计算出来得到的测试误差为 2.92km，对比交叉验证组的误差（2.68km）稍有增加但相差不多，说明得到的深度神经网络模型具有较强的泛化能力。另一方面，将测试组中的 1000 个样本分别用 2.2 节中的解方程的方法去求解，然后计算最终获得的 RMSE 值为 33.02km，经过对比可以发现深度神经网络模型与方程解算方法相比在 -60km 到 60km 范围内的平均定位精度要明显优于方程解算方法。

深入研究方程解算方法为何会有这么大的误差，通过选取几个样本点可以发现，由于时差和频差测量误差的存在，导致在解方程时求得的 r_0 ($r_0 = \sqrt{(x - x_0)^2 + (y - y_0)^2}$) 是小于 0 的，此时定位系统可以退化为测向时差定位系统；如果计算出的 r_0 仍然是小于 0 的，定位系统只能退化为双站测向定位，用两根方向线来定位，导致误差很大。选取几个样本点作为参考，如表 3-2 中的数据所示。

表 3-2 两种定位方法解出的坐标对比/km

| 准确位置坐标 | DNN 解算坐标 | 方程解算坐标 |
|------------------|------------------|------------------|
| (6.62, 23.85) | (6.60, 23.84) | (6.47, 23.64) |
| (32.92, 0.38) | (26.91, -0.62) | (5.20, 0.03) |
| (-33.88, -31.48) | (-33.72, -30.77) | (-33.83, -31.41) |
| (28.46, 50.07) | (28.90, 50.83) | (27.45, 48.35) |
| (-34.96, 46.81) | (-35.14, 46.01) | (-35.46, 47.46) |

从表 3-2 中可以看出，在大部分情况下，两种定位方法都会得到较为准确的结果，但是在一些特殊位置比如 x 轴线附近的(32.92km, 0.38km)位置处，DNN 方法就要明显优于方程解算方法，从图 2-7 中也可以看出，在 x 轴线附近方程解算方法的定位误差很大。

3.4 本章小结

本章提出了深度神经网络模型,详细介绍了 BP 神经网络的计算和训练方法,然后讲述深度学习中常用的正则化和优化方法。在模型建立的整个过程中,将样本点的多种测量信息输入到深度神经网络中,通过改变输入参数的个数和神经网络的层数,并且运用正则化方法进行泛化,找到了最佳的超参数组合,并且与方程解算方法的定位精度进行了对比,仿真实验证明该深度神经网络模型可以提高定位精度和鲁棒性。

第四章 时差频差参数估计

对于时差频差定位系统来说,怎样从两路不同位置处的接收站接收到的信号中快速准确地估算出时差频差是需要解决的首要问题。1.3.1 节中对时差频差参数估计问题国内外的研究现状做了总结性的介绍。到目前为止,基于互模糊函数的时差频差估计方法仍然是最有效并且在工程上广泛应用的方法。本章将具体分析文献[26]中的细节步骤,简化算法流程,针对具体的定位场景提出更符合实际的时差频差估计算法。

4.1 时差频差估计信号模型

多站被动定位系统由空间上分布在不同位置处的接收站组成,每个接收站在实现同步后对感兴趣的信号进行接收,从多个接收信号中提取定位参数,实现对目标的定位。具体到本文中的问题,文献[20]从运动学的原理的角度分析了多普勒效应对接收信号的影响,并且证明了对于窄带信号来说,多普勒效应都可以近似看作多普勒频移,即接收信号的频率在原来载波频率的基础上发生了微小的变动,该频率移动包含了相对运动信息。因此,双站接收到的辐射源信号可以表示为:

$$y_i(t) = r_i s(t - \tau_i) e^{j(\omega_i t + \varphi_i)} + n_i(t), i = 1, 2 \quad (4-1)$$

其中 r_i 表示接收信号的幅度衰减, τ_i 表示接收信号的时延, ω_i 表示接收信号的频率偏移量, φ_i 表示接收信号的载波相位, $n_i(t)$ 表示加性噪声。两路信号的时差定义为 $\tau = \tau_2 - \tau_1$, 频差定义为 $\omega = \omega_2 - \omega_1$ 。本文将接收信号的载波相位设定为在 $[-\pi, \pi]$ 上均匀分布的随机变量,将噪声建模为与辐射源信号无关的零均值加性白噪声。将中频信号离散化后, 信号模型变为

$$y_i(n) = r_i s(n - D_i) e^{j(\omega_i n + \varphi_i)} + n_i(n), i = 1, 2 \quad (4-2)$$

其中 $n = 0, 1, 2, \dots, N - 1$, $N = f_s T$ 表示信号的采样总点数, f_s 表示满足奈奎斯特带通采样定理的采样频率, T 为信号持续时间, D_i 表示接收信号时延的离散值。

4.2 时差频差估计理论下限

在参数估计的问题当中，对无偏估计量的方差确定一个下限是很重要的。在参数估计理论中，Cramer-Rao 下限有着广泛的应用，参数估计的均方误差常常和 Cramer-Rao 下限进行比较，用来评价该参数估计方法的性能。对于矢量参数而言，CRLB 定理为：

假定 PDF $p(\mathbf{x};\theta)$ 满足正则条件，即对于所有的 θ 都有：

$$E\left[\frac{\partial \ln p(\mathbf{x};\theta)}{\partial \theta}\right] = 0 \quad (4-3)$$

其中数学期望是对 $p(\mathbf{x};\theta)$ 求出的。那么对于任何无偏估计量 $\hat{\theta}$ 的协方差矩阵满足：

$$\mathbf{C}_{\hat{\theta}} - \mathbf{I}^{-1}(\theta) \geq 0 \quad (4-4)$$

其中大于等于 0 表示半正定。Fisher 信息矩阵 $\mathbf{I}(\theta)$ 由下式给出，即

$$[\mathbf{I}(\theta)]_{ij} = -E\left[\frac{\partial^2 \ln p(\mathbf{x};\theta)}{\partial \theta_i \partial \theta_j}\right] \quad (4-5)$$

其中导数是在 θ 的真值上计算的，数学期望是对 $p(\mathbf{x};\theta)$ 求出的。而且对于某个 p 维函数 \mathbf{g} 和某个 $p \times p$ 的矩阵 \mathbf{I} ，当且仅当

$$\frac{\partial \ln p(\mathbf{x};\theta)}{\partial \theta} = \mathbf{I}(\theta)(\mathbf{g}(\mathbf{x}) - \theta) \quad (4-6)$$

可以求得达到下限 $\mathbf{C}_{\hat{\theta}} = \mathbf{I}^{-1}(\theta)$ 的无偏估计量。这个估计量是 $\hat{\theta} = \mathbf{g}(\mathbf{x})$ ，它是 MVU 估计量，它的协方差矩阵是 $\mathbf{I}^{-1}(\theta)$ 。

文献[7]中指出，时差频差参数估计的 Cramer-Rao 下限为：

$$\sigma_{TDOA} = \frac{1}{\beta \sqrt{B_n T \gamma}} \quad (4-7)$$

$$\sigma_{FDOA} = \frac{1}{T_e \sqrt{B_n T \gamma}} \quad (4-8)$$

其中 B_n 为接收机的噪声带宽， T 为积累时间长度， γ 为信噪比， β 和 T_e 的定义如下：

$$\beta = 2\pi \sqrt{\frac{\int_{-\infty}^{+\infty} f^2 W_s(f) df}{\int W_s(f) df}} \quad (4-9)$$

$$T_e = 2\pi \sqrt{\frac{\int_{-\infty}^{+\infty} t^2 |u(t)|^2 dt}{\int |u(t)|^2 dt}} \quad (4-10)$$

其中 $u(t)$ 为时域信号， $W_s(f)$ 为信号的功率谱密度。当信号载频为矩形谱时，有

$$\beta = \frac{\pi}{\sqrt{3}} B \approx 1.8B \quad (4-11)$$

$$T_e = \frac{\pi}{\sqrt{3}} T \approx 1.8T \quad (4-12)$$

其中 B 为信号带宽，此时时差频差参数估计的 Cramer-Rao 下限为：

$$\sigma_{TDOA} = \frac{0.55}{B\sqrt{B_n T \gamma}} \quad (4-13)$$

$$\sigma_{FDOA} = \frac{0.55}{T\sqrt{B_n T \gamma}} \quad (4-14)$$

由上式可以看出，信号的持续时间、接收机噪声带宽以及接收机的信噪比决定了频差的估计精度，而时差的估计精度除了和上述的参数有关以外，还受到信号噪声带宽的影响。例如，某一窄带信号的信噪比为 0dB，接收机噪声带宽为 200kHz，那么当信号持续时间分别取 10ms、20ms、50ms 和 100ms 的时候，信号带宽分别取 10kHz、20kHz 和 50kHz 的时候，时差和频差的理论估计下限如下表所示。

表 4-1 时差理论下限

| | 10kHz | 20kHz | 50kHz |
|-------|---------|---------|---------|
| 10ms | 1.230μs | 0.615μs | 0.246μs |
| 20ms | 0.870μs | 0.435μs | 0.174μs |
| 50ms | 0.550μs | 0.275μs | 0.110μs |
| 100ms | 0.389μs | 0.194μs | 0.078μs |

表 4-2 频差理论下限

| | 10kHz | 20kHz | 50kHz |
|------|---------|---------|---------|
| 10ms | 1.230Hz | 1.230Hz | 1.230Hz |
| 20ms | 0.435Hz | 0.435Hz | 0.435Hz |

| | | | |
|-------|---------|---------|---------|
| 50ms | 0.110Hz | 0.110Hz | 0.110Hz |
| 100ms | 0.039Hz | 0.039Hz | 0.039Hz |

文献[20]从载波相位已知、载波相位未知且相等、载波相位未知且不相等三个角度分别讨论了时差频差估计精度的理论下限，再通过 Cramer-Rao 下限指出了现有时差频差联合估计的精度分析成立的潜在前提条件为：接收信号的载波相位未知且不相等。这个前提条件在绝大多数实际情况下都可以得到满足。

此外文献[20]还讨论了时差频差参数估计的物理分辨率和计算分辨率两个概念。所谓“物理分辨率”是由 式和 式确定的，与信号积累时间、信号带宽、噪声带宽和信噪比有关，它取决于信号模型本身所携带的信息。当将信号模型离散化之后，其参数估计的分辨率就会受到采样频率的影响，成为“计算分辨率”，其取决于参数估计的具体算法。该文献中还指出，只要满足奈奎斯特采样定理，离散信号就携带了原始模拟信号的全部信息，片面地通过提高采样率来提高“计算分辨率”是没有实际意义的，并不能提高参数估计的精度，应避免盲目提高采样率而得不偿失。

4.3 时差频差联合估计算法

4.3.1 时差频差粗估计

接收信号 $y_1(t)$ 和 $y_2(t)$ 的互模糊函数定义为：

$$CAF(\tau, f) = \int_{-\infty}^{+\infty} y_1^*(t) y_2(t - \tau) e^{-j2\pi ft} dt \quad (4-15)$$

现对接收信号以采样频率 f_s （采样周期 $T_s = \frac{1}{f_s}$ ）进行离散化，令 $t = nT_s$ ，则上式的离散形式为：

$$CAF(d, k) = \sum_{n=0}^{N-1} y_1^*(n) y_2(n - d) e^{-j2\pi \frac{kn}{N}} \quad (4-16)$$

式中变量之间的关系为 $N = \frac{T}{T_s}$ ， $f = \frac{kf_s}{N}$ ， $\tau = d\Delta\tau$ 。通过搜索 $|CAF(d, k)|$ 的最大值可以估计时差和频差，设 $|CAF(d, k)|$ 的极值点为 (d_0, k_0) ，则时差和频差的估计值分别为 $t_d = d_0\Delta\tau = d_0T_s$ 和 $f_d = k\Delta f = \frac{kf_s}{N}$ （ $\Delta\tau$ 和 Δf 可视作估计精度）。

式(4-16)可以看成：

$$CAF(d, k) = \sum_{n=0}^{N-1} [y_1^*(n) y_2(n-d)] e^{-j2\pi \frac{kn}{N}} = DFT[y_1^*(n) y_2(n-d)] = DFT[r(n, d)] \quad (4-17)$$

其中 $r(n, d) = y_1^*(n) y_2(n-d)$ ， $CAF(d, k)$ 的结果实际上是对相关序列 $r(n, d)$ 进行 DFT 变换，在具体实现时运用 FFT 可以提高运算速度。

基于公式(4-17)，文献[4]使用了基于滤波抽取的快速计算方法。采样频率 f_s 通常远大于相关序列 $r(n, d)$ 对应的信号带宽 B ，然而根据带通采样定理可知，为了保证信号不失真，要求采样频率大于 $2B$ ，并不要求采样频率远大于带宽 B 。鉴于此，可以采用滤波抽取技术降低采样频率，减小计算互模糊函数所需的运算量。

将相关序列 $r(n, d)$ 包含的 N 个数据分成 M 段，每段 L 个数据，按照式(4-18)获得经滤波抽取后的相关序列

$$R(m, d) = \sum_{p=0}^{L-1} r(mL + p, d) h(L-1-p) \quad (4-18)$$

式(4-18)可以理解为，将相关序列 $r(n, d)$ 通过冲激响应序列为 $h(n)$ 、长度为 L 的低通滤波器后，以 L 为间隔进行抽取，从而获得序列 $R(m, d)$ ，其中 L 为抽取比。

经过上述滤波抽取后，互模糊函数为

$$CAF(d, k) = \sum_{m=0}^{M-1} R(m, d) e^{-j2\pi \frac{km}{M}} \quad (4-19)$$

峰值处所对应的时差为 $\tau = d\Delta\tau$ ，频差为 $f = k\Delta f$ ，频差的分辨率为 $\Delta f = \frac{f_s}{ML} = f_s/N$ 。

4.3.2 时差频差精估计

由于收到采样周期 T_s 的限制，最小的时差搜索间隔只能等于 T_s ，无法实现分数倍采样周期的时差搜索，从而限制了时差频差估计精度。针对该问题，为了进一步提高估计精度，可以采用两种方法：一是信号插值法；利用 sinc 插值技术，基于离散采样点可以重构信号 $s(t)$ ，然后进行分数倍采样周期的移动，接着对时间移动后的连续信号进行重采样，可获得时间移动为分数倍采样周期的离散信号，从而实现分数倍采样周期的时差搜索。二是二次曲面拟合法；在已获得的时差频差估计点附近，选取若干点，进行二次曲面拟合，从而进一步提高时差频差的估

计精度。具体来说就是由于时差频差精度的存在，有可能出现如图 4-1 所示的情况：

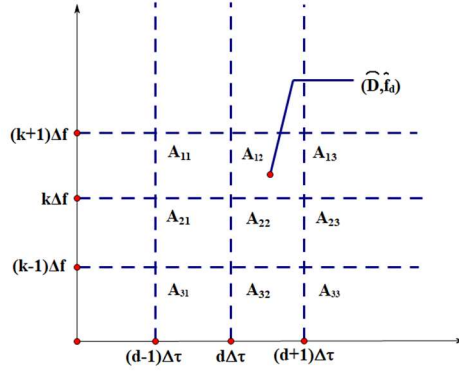


图 4-1 二次曲面拟合示意图

即实际的时差和频差为 (\hat{D}, \hat{f}_d) ，但是解得的时差频差估计值为 $(d\Delta\tau, k\Delta f)$ ，针对上述两种方法，文献[4]对其进行对比分析，选用了小运算量的二次曲面拟合。二次曲面拟合的思想为：可以把 $|CAF(\tau, f)|$ 视作关于 τ 和 f 的二次函数，可写成：

$$|CAF(\tau, f)| = c_1\tau^2 + c_2\tau f + c_3f^2 + c_4\tau + c_5f + c_6 \quad (4-20)$$

然后对其求极值点，分别让其偏导数为零：

$$\frac{\partial |CAF(\tau, f)|}{\partial \tau} = 2c_1\tau + c_2f + c_4 = 0 \quad (4-21)$$

$$\frac{\partial |CAF(\tau, f)|}{\partial f} = c_2\tau + 2c_3f + c_5 = 0 \quad (4-22)$$

则可以得到极值点处的时差和频差值：

$$\tau = \frac{2c_3c_4 - c_2c_5}{c_2^2 - 4c_1c_3} \quad (4-23)$$

$$f = \frac{2c_1c_5 - c_2c_4}{c_2^2 - 4c_1c_3} \quad (4-24)$$

将其变为离散形式可以写成一组矩阵方程：

$$\mathbf{MC} = \mathbf{A} \quad (4-25)$$

$$\text{式中 } \mathbf{M} = \begin{bmatrix} 1 & -1 & 1 & -1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & -1 & -1 & 1 \\ 0 & 0 & 1 & 0 & -1 & 1 \\ 1 & -1 & 1 & 1 & -1 & 1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} c_1(\Delta\tau)^2 \\ c_2\Delta\tau\Delta f \\ c_3(\Delta f)^2 \\ c_4\Delta\tau \\ c_5\Delta f \\ c_6 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} A_{11} \\ A_{12} \\ A_{13} \\ A_{21} \\ A_{22} \\ A_{23} \\ A_{31} \\ A_{32} \\ A_{33} \end{bmatrix}。$$

通过公式 $\mathbf{C} = \mathbf{M}^+ \mathbf{A}$ 可以解得 $c_1, c_2, c_3, c_4, c_5, c_6$ 的值，拟合后的估计值变为：

$$\hat{D} = d\Delta\tau + \frac{2c_3c_4 - c_2c_5}{c_2^2 - 4c_1c_3} \quad (4-26)$$

$$\hat{f}_d = k\Delta f + \frac{2c_1c_5 - c_2c_4}{c_2^2 - 4c_1c_3} \quad (4-27)$$

4.3.3 算法总结

4.4 时差频差分维估计

在时差频差联合估计中，可以直接通过互模糊函数搜索峰值的位置来估计时差和频差，但这种方法需要在时域和频域二维空间上进行计算和搜索，计算量很大。根据互模糊函数的性质，在最大峰处，当某一维参数已经被准确补偿的情况下，信号沿频差轴的分布完全取决于信号的频域相关函数，而沿时差轴的分布完全取决于信号的互相关函数，这就给降低运算量提供了思路，文献[80]将联合估计问题中的二维搜索问题彻底变成了两个独立的一维搜索问题，而文献[20]则根据接收信号自相关函数与时延无关的特性，实现了彻底的分维处理，避免了互模糊函数的时域和频域二维搜索的难题。在此基础上，文献[81]充分挖掘降低运算量的空间，在频差和时差估计的每一步均采取先粗估后精估的方法。

4.4.1 频差估计

实现对互模糊函数的分维计算，需要首先估计出时差或频差两者之一，补偿后再估计另外一个变量。文献[80]提出的先估计频差的思路是基于两路信号的相关函数之比与时差无关的特性。下式表示信号的相关函数：

$$\begin{aligned} r_i(\tau) &= E[y_i^*(t)y_i(t+\tau)] = r_i^2 E[s_i^*(t)s_i(t+\tau)]e^{j\omega_i\tau} + N_i\delta(\tau) \\ &= r_i^2 r_s(\tau)e^{j\omega_i\tau} + N_i\delta(\tau), i=1,2 \end{aligned} \quad (4-28)$$

其中 $r_s(\tau)$ 为信号 $s(t)$ 的自相关函数， N_i 为噪声功率。则两路信号自相关函数的比值如下：

$$\begin{aligned} \xi(\tau) &= \frac{r_2(\tau)}{r_1(\tau)} = \frac{r_2^2 r_s(\tau)e^{j\omega_2\tau} + N_2\delta(\tau)}{r_1^2 r_s(\tau)e^{j\omega_1\tau} + N_1\delta(\tau)} \\ &= \frac{r_2^2}{r_1^2} e^{j\omega_d\tau} + \left(\frac{r_2^2 r_s(0) + N_2}{r_1^2 r_s(0) + N_1} - \frac{r_2^2}{r_1^2} \right) \delta(\tau) \\ &= r^2 e^{j\omega_d\tau} + \eta\delta(\tau) \end{aligned} \quad (4-29)$$

其中 $r^2 = \frac{r_2^2}{r_1^2}$ ， $\eta = \frac{r_2^2 r_s(0) + N_2}{r_1^2 r_s(0) + N_1} - \frac{r_2^2}{r_1^2}$ ，均为常数，由此可知两路信号的自相关函数的比值 $\xi(\tau)$ 除一个冲击点是单谐波函数外，谐波分量就是两路信号的频差，且与时差无关，实现了对信号时差频差的分离。

对式(4-2)的离散信号模型定义自相关函数如下：

$$r_i(k) = \frac{1}{N} \sum_{n=0}^{N-1} y_i^*(n)y_i(n+k), i=1,2 \quad (4-30)$$

其中 $k = -N+1, \dots, -1, 0, 1, \dots, N-1$ ， $r_i(k)$ 的长度为 $2N-1$ 。两路信号的自相关函数之比为：

$$\xi(n) = \frac{r_2(n)}{r_1(n)} \quad (4-31)$$

用式(4-30)计算两路信号的自相关函数，根据式(4-31)求得 $\xi(n)$ ，做FFT取其峰值处的频率就可以得到频差 ω_d 。

计算信号自相关函数的传统方法为FFT。为了避免混叠，用FFT计算相关函数时需要补零至 N_f ($N_f \geq 2N-1$ 且 N_f 为2的整数次幂)点，对式(4-2)的信号做FFT如下：

$$Y_i(k) = FFT(y_i(n)) = \sum_{n=0}^{N_f-1} y_i(n) e^{-j2\pi nk/N_f}, i=1,2 \quad (4-32)$$

其中 $k = 0, 1, \dots, N_f - 1$ 。每一路信号的功率谱为：

$$R_i(k) = Y_i^*(k) Y_i(k), i=1,2 \quad (4-33)$$

对其做 IFFT 便可以得到信号的自相关函数为：

$$r_i(n) = IFFT(R_i(k)) = \frac{1}{N_f} \sum_{k=0}^{N_f-1} R_i(k) e^{j2\pi nk/N_f}, i=1,2 \quad (4-34)$$

其中 $n = 0, 1, \dots, N_f - 1$ ，自相关函数序列长度为 N_f 。

在实际中信号的频差一般比较小，远远达不到信号的采样率 f_s 的量级，因此序列 $\xi(n)$ 的谐波频率很低，不需要信号的采样率为 f_s 。如果根据先验知识，可以将采样频率降至 f'_s ，那么降采样的倍数为 $M = f_s/f'_s$ ，将自相关函数进行 M 倍抽取，得到长度为 L 的自相关函数序列 $r_{iL}(n)$ ，两者相除得到长为 L 的 $\xi_L(n)$ 序列：

$$\xi_L(n) = \frac{r_{2L}(n)}{r_{1L}(n)} \quad (4-35)$$

其中 $L = N_f/M$ 。

对于抽取后的自相关函数 $r_{iL}(k)$ 的计算，可以采用先计算自相关函数 $r_i(k)$ ，再进行抽取的方法，也可以采用文献[20]中提出的用多相滤波的方法降采样率来实现。但文献[81]中对两者的运算量进行了对比，多相滤波并不能降低运算量，并且会使得流程较为复杂，因此直接采用了 FFT 计算自相关函数，再进行抽取的方法。

得到式(4-35)所示的抽取后自相关函数之比 $\xi_L(n)$ 序列之后，对其做 FFT 得到 $\xi_L(n)$ 序列的频谱如下：

$$\psi(k) = FFT(\xi_L(n)) = \sum_{n=0}^{L-1} \xi_L(n) e^{-j2\pi nk/L} \quad (4-36)$$

搜索其峰值便可以得到两路信号的频差，即：

$$\hat{\omega}_{d0} = \arg \max_k (\psi(k)) \cdot 2\pi f'_s / L \quad (4-37)$$

上述算法得到的频差的分辨率只有 f'_s/L ，距离理论下界差得很远，还需要进一步估计。由数字信号处理的基本原理可以知道，信号的频谱分辨率取决于信号

的时长。对信号的时域补零等效于对频谱的插值，相当于提高了频谱的分辨率。设要求的频谱分辨率为 Δf ，则补零后的序列长为 $N' = f_s' / \Delta f$ ， $\xi_L(n)$ 序列补零后变为：

$$\xi'(n) = \begin{cases} \xi_L(n), n \in [0, L/2 - 1] \cup [N' - L/2, N' - 1] \\ 0, \text{others} \end{cases} \quad (4-38)$$

对 $\xi'(n)$ 做 FFT 便可以得到细化全部频谱。为了减小运算量，文献[20]中在对 $\xi_L(n)$ 进行补零后，用 CZT 得到粗估结果 $\hat{\omega}_{d0}$ 附近的局部细化频谱，而避免计算整个的频谱。这样做确实减小了频差精估阶段的运算量，但文献[81]通过对典型信号进行仿真，发现受到频差估计理论下限的限制，频差分辨率 Δf 的要求往往不大，直接对 $\xi'(n)$ 做 FFT 的运算量与前面式(4-32)一式(4-34)计算两路信号自相关函数的运算量相比要小得多。所以可以直接运用 FFT 计算 $\xi'(n)$ 的频谱：

$$\psi'(k) = FFT(\xi'(n)) = \sum_{n=0}^{N'-1} \xi'(n) e^{-j2\pi nk / N'} \quad (4-39)$$

搜索其峰值便可以得到频差的高精度估计，即：

$$\hat{\omega}_d = \arg \max_k (\psi'(k)) \cdot 2\pi f_s' / N' \quad (4-40)$$

综上所述，频差估计的算法具体步骤如下：

- (1) 根据式(4-32)计算两路信号 $y_i(n)$ 的频谱 $Y_i(k)$ ；
- (2) 根据式(4-33)计算两路信号的功率谱 $R_i(k)$ ；
- (3) 根据式(4-34)计算两路信号的自相关函数 $r_i(n)$ ；
- (4) 对 $r_i(n)$ 进行 M 倍抽取得到 $r_{iL}(n)$ ；
- (5) 根据式(4-35)计算自相关函数之比 $\xi_L(n)$ ；
- (6) 根据式(4-38)对 $\xi_L(n)$ 进行补零得到 $\xi'(n)$ ；
- (7) 根据式(4-39)计算 $\xi'(n)$ 的频谱 $\psi'(k)$ ；
- (8) 根据式(4-40)搜索 $\psi'(k)$ 的峰值，得到信号的频差 $\hat{\omega}_d$ 。

4.4.2 时差估计

对于模拟信号而言，文献[82]已经证明互相关函数是时差参数的最大似然估计。文献[20]对时差估计问题在对中频离散信号进行数字下变频处理的基础上，提出了先粗估后精估的思想，即先对离散信号直接处理获得计算分辨率较低的时

差粗值，然后在时差粗值附近对时域峰值进行细化处理，最终获得精确的时差估计。这一过程并不需要较高的采样频率来保证精确的时差估计，既可以降低接收信号离散数据的速率，又能获得最佳的时差估计精度。时差粗估算法的主要思想是互相关函数无失真分段表示，时差精估算法的主要思想是插值和拟合处理。论文设计了外场实验系统进行了试验验证，实验中时差估计的统计标准差很逼近 Cramer-Rao 下限的理论标准差，验证了算法的正确性和有效性。

上述算法在不提高信号采样率的前提下，成功获得了最佳的时差估计精度，且运算量较小，是一种工程中切实可行的方法。文献[81]对该算法进行了深入分析，进一步减小了运算量，文中指出在时差粗估和精估环节中均需要计算信号的互相关函数，因此需要进行多段 FFT 和 IFFT。在时差精估环节中，要对信号做 CZT 以得到其互功率谱，再做 ICZT 以得到互相关函数，其计算中也要用到 FFT 和 IFFT。时差粗估计和精估计环节都要涉及到信号的时域与频域之间的相互转化，由冗余计算。因此文献[81]中将时差粗估环节中的互相关函数的计算过程与精估环节进行了融合，减少了多余的 FFT 和 IFFT 过程，降低了运算量。

本节只考虑信号模型已经补偿了频差信息，两路信号中只含时差信息，那么式(4-1)中的信号模型就变成了：

$$y_i(t) = r_i s(t - \tau_i) e^{j\varphi_i} + n_i(t), i = 1, 2 \quad (4-41)$$

其中 r_i 表示接收信号的幅度衰减， τ_i 表示接收信号的时延， φ_i 表示接收信号的载波相位， $n_i(t)$ 表示加性噪声。同理式(4-2)中离散信号模型就变成了：

$$y_i(n) = r_i s(n - D_i) e^{j\varphi_i} + n_i(n), i = 1, 2 \quad (4-42)$$

其中 $n = 0, 1, \dots, N - 1$ ， $N = f_s T$ 表示信号的采样点数， f_s 表示满足奈奎斯特带通定理的采样频率， T 为信号持续时间， D_i 表示接收信号时延的离散值。其互相关函数的计算公式为：

$$r(k) = \frac{1}{N} \sum_{n=0}^{N-1} y_1^*(n) y_2(n+k) \quad (4-43)$$

其中 $k = -N + 1, \dots, -1, 0, 1, \dots, N - 1$ ， $r(k)$ 的长度为 $2N - 1$ 。

直接计算式(4-43)的运算量较大，文献[20]中的解决方法是在时差粗估环节采用互相关函数的分段无失真表示来降低运算量，在时差精估计环节采用 CZT 将其变换到频域处理。由于计算分段互相关函数中的每一段时仍然需要做 FFT

将其变换到频域，所以可以和精估计环节的 CZT 融合处理。当信号是基带信号时，不需要专门做带限频谱截取，CZT 等同于 FFT。为了避免混叠，用 FFT 计算相关时需要补零至 $N_f (N_f \geq 2N - 1 \text{ 且 } N_f \text{ 为 } 2 \text{ 的整数次幂})$ 点。将两路信号做 N_f 点 FFT 如下：

$$Y_i(k) = FFT(y_i(n)) = \sum_{n=0}^{N-1} y_i(n) e^{-j2\pi nk/N_f}, i=1,2 \quad (4-44)$$

其中 $k = 0, 1, \dots, N_f - 1$ 。由两路信号的频谱可以得到其互功率谱如下：

$$R(k) = Y_1^*(k) Y_2(k) \quad (4-45)$$

将 $R(k)$ 做 IFFT 便可以得到两路信号的互相关函数 $r(n)$ 如下：

$$r(n) = IFFT(R(k)) = \frac{1}{N_f} \sum_{k=0}^{N_f-1} R(k) e^{j2\pi nk/N_f} \quad (4-46)$$

其中 $n = 0, 1, \dots, N_f - 1$ ，搜索其峰值就可以得到两路信号的时差，即：

$$\hat{\tau}_0 = \arg \max_n (r(n)) / f_s \quad (4-47)$$

文献[20]中在时差粗估计环节用互相关函数的分段无失真表示来减小互相关函数的计算量。若单看时差粗估计环节，这种方法的运算量比式(4-44)一式(4-46)直接计算确实小了很多。但在时差精估计环节，若信号为基带信号，CZT 退化为 FFT，式(4-44)和式(4-46)的计算还是避免不了。综合考虑时差粗估和精估环节，分段无失真表示并没有减小运算量，所以可以对式(4-44)一式(4-46)直接进行计算，得到的互功率谱谱可以直接在时差精估计环节使用，避免了重复计算。

上述算法得到的互相关函数的分辨率只有 $1/f_s$ ，距离理论下界还差得很远，还需要进行进一步的估计。为了获得小于采样周期的时差分辨率，通常的思路是对互相关函数的峰值附近进行插值和拟合。文献[20]中采用的在互功率谱中间补零以提高采样率的方法是切实可行的。若要求的互相关函数的分辨率为 $1/f'_s$ ，则频谱点数变为 $N'_f = N_f \cdot f'_s / f_s$ ，补零后的互功率谱变为：

$$R'(k) = \begin{cases} R(k), k \in [0, N_f/2 - 1] \cup [N'_f - N_f/2, N'_f - 1] \\ 0, \text{others} \end{cases} \quad (4-48)$$

得到补零后的互功率谱之后，对于做 IFFT 便可以得到互相关函数的全部精细时域波形。文献[20]中指出，由于通过式(4-47)已经得到了时差的初步估计值，并不需要利用式(4-48)计算互相关函数的全部精细时域波形，而只需要计算 $\hat{\tau}_0$ 附近的点即可，这就可以利用 Chirp-Z 变换的思路，利用 ICZT 计算 $\hat{\tau}_0$ 附近的点即可，避免计算其他不必要的时间段上的精细时域波形。这样做确实减小了时差精估环节的运算量，但需要先进行时差粗估以确定 ICZT 的时域起始时刻。因此文献[81]中将时差粗估和精估环节相融合，不用进行时差粗估直接得到精确时差值。要实现这一点，就必须对补零后的互功率谱做 IFFT，其运算量又很大。为了实现在时差估计精度不降低的前提下减小运算量，可以借鉴上一节中频差估计利用先验频差信息对信号进行降采样的思路，采用利用先验时差信息先在频域采样降时长，后补零做 IFFT 的方法来计算互相关函数的精细时域波形。

与在频差估计可以根据先验信息降采样类似，若根据先验信息可以将相关函数的时长降至 T' ，那么可以对互功率谱进行 M 倍抽取($M = N_f / (T' \cdot f_s)$)，得到长度为 L 的互功率谱序列 $R_L(k)$ 。与式(4-38)类似，若要求的互相关函数的分辨率为 $1/f'_s$ ，那么频谱点数变为 $N'_f = f'_s \cdot T'$ ，补零后的互功率谱变为：

$$R'(k) = \begin{cases} R_L(k), k \in [0, L/2 - 1] \cup [N'_f - L/2, N'_f - 1] \\ 0, \text{others} \end{cases} \quad (4-49)$$

对上式做 IFFT，得到精细的互相关函数如下：

$$r'(n) = \text{IFFT}(R'(k)) = \frac{1}{N'_f} \sum_{k=0}^{N'_f-1} R'(k) e^{j2\pi nk/N'_f} \quad (4-50)$$

搜索其峰值就可以得到两路信号的时差，即：

$$\hat{\tau} = \arg \max_n (r'(n)) / f'_s \quad (4-51)$$

总结上述的算法具体步骤如下：

- (1) 根据式(4-44)由两路信号 $y_i(n)$ 做 FFT 计算其频谱 $Y_i(k)$ ；
- (2) 根据式(4-45)计算互功率谱 $R(k)$ ；
- (3) 对互功率谱 $R(k)$ 进行 M 倍抽取得到 $R_L(k)$ ；
- (4) 根据式(4-49)得到补零后的互相关函数的频谱 $R'(k)$ ；
- (5) 根据式(4-50)由 $R'(k)$ 做 IFFT 计算互相关函数的精细时域波形 $r'(n)$ ；

(6) 根据式(4-51)搜索 $r'(n)$ 的峰值，得到两路信号的时差 $\hat{\tau}$ 。

4.4.3 算法总结

上面两节分别推导了两路输入信号的频差和时差参数的高精度测量算法，在实际的处理过程中，首先用 4.4.1 节的算法估计出两路信号的频差，在补偿了频差之后，再用 4.4.2 节中的算法估计信号的时差。若估计出的频差为 $\hat{\omega}_d$ ，则补偿后的信号 $y'_2(n)$ 为：

$$y'_2(n) = y_2(n) \cdot e^{-j\hat{\omega}_d n / f_s} \quad (4-52)$$

用 $y'_2(n)$ 代替式(4-44)中的 $y_2(n)$ ，就可以用补偿后的信号进行时差的估计。综上所述，整个流程包括频差估计、频差补偿、时差估计三个主要步骤，总的流程如图 4-3 所示：

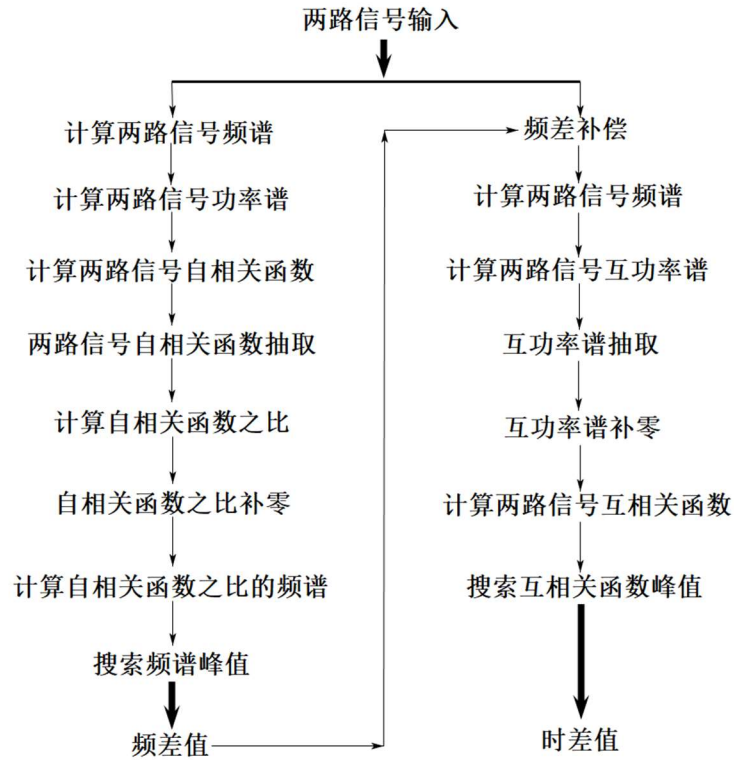


图 4-2 时差频差分维估计流程图

要注意的是在时差和频差的估计过程中，需要一些先验信息来减小运算量：

- (1) 两路信号可能产生的最大时差，由双站的间距和最大侦查范围等确定；
- (2) 两路信号可能产生的最大频差，由双站的速度和最大侦查范围等确定；
- (3) 频差估计的精度，由信噪比、积累时间和噪声带宽等确定；

(4) 时差估计的精度，由信噪比、积累时间、信号带宽和噪声带宽等确定。

4.5 算法仿真

为了验证上述时差频差估计算法的有效性，本文利用 MATLAB 软件进行数字仿真，信号选取窄带线性调频信号，仿真基本参数如下：

信号载频：1GHz

采样频率：1MHz

信号带宽：200kHz

信号时长：1ms

(1) 未加噪声

表 4-3 时差频差联合估计结果

| 真实时差 | 真实频差 | 一步估计 时差 | 一步估计 频差 | 两步估计 时差 | 两步估计 频差 |
|---------|------------|------------|------------|------------|------------|
| 19936ns | 64.2528Hz | 20000ns | 63.8962Hz | 20391ns | 63.8462Hz |
| 17204ns | 128.2487Hz | 17000ns | 127.7924Hz | 17165ns | 127.9565Hz |
| 23524ns | 44.4524Hz | 24000ns | 44.8227Hz | 23131ns | 44.4008Hz |

缺加噪声的结果，分维估计算法以及两者的对比

4.6 本章小结

第五章 总结与展望

5.1 本文内容总结

5.2 展望

本文的工作解决了时差频差定位体制应用到双站平台上的一些重要的技术问题，给出的算法明确清晰，实用性较强，且均通过了数字仿真得到了验证。时差频差技术作为各种无源定位体制中优势明显且得到深入研究的一种，其在各种观测平台上会得到越来越多的应用，这其中会不断出现新的问题需要解决。作者认为在以后的工作研究中需要深入展开的包括以下几点：

（1）对于时差频差参数估计问题，文中提出了两种估计算法，一种是二维联合搜索，另一种是分维搜索，相对于互模糊函数法的二维联合搜索来说，分维算法的适应性还有待检验。时差频差分维估计算法的思路和提出都较晚，其适用的信号形式和环境还没有得到充分验证，文中也只是对线性调频信号在加性高斯白噪声环境下进行了仿真分析。当信号为非窄带信号或者脉冲串信号时，分维算法在使用上与窄带信号时的区别需要进一步研究。

（2）对于时差频差定位问题，本文提出的深度神经网络模型算法实质上是一种机器学习算法，需要在特定的环境下依据大量的样本学习才能获得较为理想的定位结果，怎样将其运用在未知的场景或者通过少量样本就能够快速定位仍然是两个比较困难的问题。另外，时差频差与其他参数的联合定位算法还有待总结，当时差（或者频差）测量精度很差时，仅利用频差（或者时差）与角度以及其他参数进行联合定位的算法以及相关的误差分析还需要进一步研究。

（3）对于时差频差跟踪问题，本文并没有在跟踪问题上进行研究，通过观察定位公式可以看出，时差频差定位公式属于高度的非线性方程，如何建立以及求解跟踪方程也是一个值得深入探讨的问题。

致谢

在本文完成之际，谨向给予我帮助、指导、关心和支持的老师、领导、同学和亲友致以诚挚的谢意。

首先要衷心感谢导师王桂玲和冯胜老师。王老师为人正直、待人诚恳，在两年多的研究生学习期间，一直潜移默化地影响着作者，当作者在课题研究过程中遇到方向性问题时，总能得到王老师耐心的、具有启发性的指导，每次跟王老师谈话总有一种如沐春风的感觉，不仅如此，王老师还不断鼓励作者探索更具挑战性和关键性的问题，在生活中，王老师一直给予作者诸多关心和帮助。特别感谢冯胜老师，在学习上，冯老师循循善诱，治学有方，是一位非常有责任感的老师，他丰富的研究经验、工程阅历以及渊博的知识都极大地拓宽了我的视野，让我从中受益良多，感谢冯老师对我在学术上的一些浅薄的思考和见解给予了充分的鼓舞和肯定，感谢他对我论文的开题、写作都给予了非常宝贵的意见和建议；在科学研究上，他与时俱进，时刻站在科研的最前线，冯老师严谨的科学态度和敏锐的思维方式时刻感染着身边的每一个人；在生活中，他平易近人，言辞幽默风趣，给予了作者无微不至的关怀，是学生的好老师，更是学生的好朋友。值此论文完稿之际，再次向王老师和冯老师表示崇高的敬意和衷心的感谢。

感谢安洁和刘纪斌老师，安老师对待学生始终非常亲切，对于我们研究生阶段的课程学习及论文的进度，安老师十分关心并悉心督促，让我们的学习和论文都能够按部就班的进行。在生活 and 工作中遇到困难时安老师帮我分析困难，使我能够克服苦难，安老师的贴心关怀使我能够更加专注地投入学习和工作中。感谢刘老师在学习和生活上对我们的帮助，刘老师平易近人，并且十分关心研究生的论文进度情况，每次遇到工作上的困难，刘老师总会为我们妥善解决。

感谢二十三所给我们提供了良好的学习和工作环境；感谢十二室的领导和同事对我学习和工作上的支持和关心；感谢郭廷伟主任、刘晓宏副主任、沈小马、朱应申、任满堂、董华智、曲行根、丁晓宁等对我提出的疑问的耐心解答和指导。

感谢研究生院的赵昕虹老师、王淑敏老师、纪军丽老师、宋洁老师，感谢老师们在我研究生阶段的入学、选课等日常教学管理工作和在研究生院的生活中付出的辛勤汗水。

感谢王柳柳学姐、菅毛学长、林家豪学长、潘超群学长等在学习中对我的帮助，感谢岳天翔、宋鑫华、于沐尧等同学与我一起研究解决学术问题，在与他们的讨论中，我获益匪浅。

感谢父母对我生活、学习和工作上的关心和支持，在我遇到困难和逆境彷徨时，是他们给了我战胜一切的勇气和决心。

最后，向百忙之中评审论文以及参加答辩的各位专家学者致以诚挚的谢意。

参考文献

- [1] 孙仲康,郭福成.单站无源定位跟踪技术[M].国防工业出版社,北京.2008
- [2] 胡来招.无源定位[M].国防工业出版社,北京.2004
- [3] Richard A.Poisel 著.屈晓旭, 罗勇等译.电子战目标定位方法[M].电子工业出版社, 北京.2008
- [4] 贾兴江.运动多站无源定位关键技术研究[D].长沙: 国防科学技术大学研究生院, 2011
- [5] 许耀伟.一种快速高精度无源定位方法的研究[D].长沙: 国防科学技术大学研究生院, 1998
- [6] 郭福成.基于运动学原理的单站无源定位技术研究[D].长沙: 国防科学技术大学研究生院, 2002
- [7] 冯道旺.利用径向加速度的单站无源定位技术研究[D].长沙: 国防科学技术大学研究生院, 2003
- [8] 张敏, 冯道旺, 郭福成.基于多普勒变化率的单星无源定位[J].航天电子对抗, 2009, 25 (5) :11~13
- [9] 陆安南, 孔宪正.单星测频无源定位法[J].通信学报, 2004, 25 (9) :160~168
- [10] Bar-Shalom Y, X R Li, Kirubarajan T. Estimation with applications to tracking and navigation: Theory algorithms and software[M]. New York: John Wiley & Sons, 2004. 1-2, 165-236
- [11] Tianyuan X, Shunlan L. Single observer passive location using phase rate of change with the extended Kalman particle filter[C]. ISECS International Colloquium on Computing, Communication, Control and Management. CCCM 2009. Sanya: IEEE, 2009. 65-68
- [12] 李静, 张立东.短基线时差定位技术研究[J].舰船电子对抗,2014,06:22-25
- [13] 汪珺.侧向交叉定位技术[J].电子科技.2011,07:129-132
- [14] 崔焱, 王玉林.双站方位时差定位系统精度分析[J]. 无线电通信技术,2011,03:44-46

- [15]Musicki D, Koch W. Geolocation using TDOA and FDOA measurements[C].
2008 IEEE 11th International Conference on Information Fusion. Cologne:
IEEE,2008,1-8
- [16]李曦.双站无源目标跟踪算法研究[D].无锡: 江南大学, 2016
- [17]王永诚, 张令坤.多站时差定位技术研究[J].现代雷达, 2003, 02:1-4
- [18]郭福成, 李腾.基于时差和频差的固定多站定位方法及分析[J].系统工程与电子技术, 2011,09:1954-1958
- [19]X Lu, K C Ho. Taylor-series Technique for Source Localization using AoAs in
the Presence of Sensor Location Errors[C]. Fourth IEEE Workshop on Sensor
Array and Multichannel Processing. Waltham. WA: IEEE,2006.190-194
- [20]魏和文.被动定位参数估计与多维标度[D].成都: 西南电子通信技术研究所,
2009
- [21]C.H.Knapp, G.C.Carter. The generalized correlation method for estimation of
time delay[J]. IEEE Trans.Acoust., Speech, Signal Processing, vol.ASSP-24,
Aug.1976:320-327
- [22]S.Stein. Algorithm for ambiguity function processing[J]. IEEE Trans. Acoust.,
Speech, Signal Processing, vol.ASSP-29, no.3, June 1981:588-599
- [23]Stein S. Differential Delay/Doppler ML Estimation with Unknown Signal[J].
IEEE Transactions on Signal Processing, 1993, 41(8):2717~2719
- [24]Yatrakis C L. Computing the Cross-Ambiguity Function-A Review[D]. USA: the
Graduate School of Binghamton University, 2005
- [25]Johnson J J. Implementing the Cross Ambiguity Function and Generating
Geometry-Specific Signals[D]. USA: Research Office Naval Postgraduate School,
2001
- [26]Tao R, Zhang W Q, Chen E Q. Two-stage Method for Joint Time Delay and
Doppler Shift Estimation[J]. IET Radar Sonar Navig., 2008,2(1): 71~77
- [27]Chan Y T, Ho K C. Joint Time-Scale and TDOA Estimation: Analysis and Fast
Approximation[J]. IEEE Transactions on Signal Processing, 2005, 53(8):
2625~2634

- [28] C.L.Nikias, R.Pan. Time delay estimation in unknown Gaussian spatially correlated noise[J]. IEEE Trans. Acoust. Speech Signal Process., vol 36, no.11, Nov. 1988:1706-1714
- [29] A.V.Dandawate, G.B.Giannakis. Differential delay-doppler estimation using second and high-order ambiguity functions[J]. IEE Proc.-Radar Sonar Navig., vol.140, no.6, Dec. 1993:410-418
- [30] W.A.Gardenr, A.Napolitano, L.Paura. Cyclostationarity: Half a century of research[J]. Signal processing, vol.86, no.4, April 2006: 639-697
- [31] P.G.Georgiou, P.Tsakalides, C.Kyriakakis. Alpha-stable modeling of noise and robust time-delay estimation in the presence of impulsive noise[J]. IEEE Trans. Multimedia, Sept. 1999:291-301
- [32] G.A.Tsihrintzis, U.Tureli, C.L.Nikias. Fractional lower-order statistics-based ambiguity functions for differential delay Doppler estimation[J]. IEE Pro.-Radar Sonar Navig., vol.143, no.6, Dec.1996: 358-365
- [33] H.C.So, P.C.Ching. Comparative study of five LMS-based adaptive time delay estimators[J]. IEE Pro.-Radar Sonar Navig., vol.148, no.1, Feb.2001:9-15
- [34] S.R.Dooley, A.K.Nandi. Adaptive time delay and frequency estimation for digital signal synchronization in CDMA systems[C]. Thirsty-Second Asilomar Conf. Signals, Syst., Comput., vol.2, 1998: 1838-1842
- [35] Weiss L.G. Wavelets and wideband correlation processing[J]. IEEE Signal Processing, June 1994:13-32
- [36] A.R.Naghsh-Nilchi. Joint estimation of differential time delays and frequency offsets[D]. M.S.Thesis, University of Utah, Salt Lake City, Utah, USA, Dec. 1990
- [37] A.R.Naghsh-Nilchi, V.J.Mathews. An efficient algorithm for joint estimation of differential time delays and frequency offsets[C]. ICASSP-92, vol. 5, March 1992: 309-312
- [38] J.O.Smith, J.S.Abel. Closed-form least-squares source location estimation form range-difference measurements[J]. IEEE Trans. Acoust., Speech, Signal Process., vol.ASSP-35, Dec. 1987: 1661-1669

- [39]B.Friedlander. A passive localization algorithm and its accuracy analysis[J].
IEEE J.Ocean.Eng., vol. OE-12, Jan. 1987:234-245
- [40]R.O.Schmidt. An Algorithm for Two-Receiver TDOA/FDOA Emitter
Location[J]. ESL, Inc., Tech. Memo. TM-1229, May 1980
- [41]K.C.Ho, Y.T.Chan. Geolocation of a known altitude object form TDOA and
FDOA measurements[J]. IEEE Trans. Aerosp. Elect. Syst., vol.33, July
1997:770-782
- [42]K.C.Ho, Wenwei Xu. An accurate algebraic solution for moving source location
using TDOA and FDOA measurements[J]. IEEE Trans. Signal Proc., vol. 52,
no.9, Sept. 2004:2453-2463
- [43]He-Wen Wei, Rong Peng, Qun Wan, Zhang-Xin Chen, Shang-Fu Ye.
Multidimensional scaling analysis for passive moving target localization with
TDOA and FDOA measurements[J]. IEEE Trans. Signal Proc., vol.58, no.3,
March 2010:1677-1688
- [44]周一宇, 安玮, 郭福成, 柳征, 姜文利.电子对抗原理[M]. 北京: 电子工业
出版社, 2009
- [45]Tim Pattison, S.I.Chou. Sensitivity analysis of Dual-Satellite Geolocation[J].
IEEE Transactions on Aerospace and Electronic Systems, vol.36, No.1, January,
2000:56-71
- [46]郭福成, 樊昀.双星 TDOA/FDOA 无源定位方法分析[J].航天电子对抗, 2006
(6):20-23
- [47]郭福成, 樊昀.双星时差频差联合定位方法及其误差分析[J]. 宇航学报, 2008
(7):1381-1386
- [48]陆安南. 双机 TDOA/DD 无源定位方法[J]. 电子科技大学学报, 2006, 35
(1):17-20
- [49]冯天军, 贾良仁. 双机 TDOA-FDOA 定位精度理论分析[J]. 科技信息 (学术
版), 2008, 9:37-40
- [50]冯天军. 双机时差-频差组合定位研究[D]. 长沙: 国防科学技术大学研究生院,
2008

- [51]杨洁,刘聪锋,田中成,等.迭代时差频差联合定位算法及其性能分析[J].西安电子科技大学学报(自然科学版),2015,42(4):140-146
- [52]郁涛.基于时差频差测量的双机无源定位的线性解析方法[J].中国电子科学研究院学报,2016,11(1):29-31
- [53]曾辉,曾芳玲.空间三站时差定位的模糊及无解问题[J].太赫兹科学与电子信息学报,2010,8(2):139-142
- [54]洪伟,陈婷,雷连发.平面时差定位模糊性及精度研究[J].火控雷达技术,2015,44(4):31-35.
- [55]袁罡,陈鲸.三站时差定位模糊问题解决方法[J].中国电子科学研究院学报,2014,9(1):89-92.
- [56]韩丽君.平面三站时差定位中的模糊研究[J].计算机与数字工程,2012,40(5):10-11.
- [57]曲婧.基于神经网络的无源时差定位算法研究[D].中北大学,2010.
- [58]董哲,吴瑶,孙德辉.室内定位技术的多源数据融合算法研究[J].计算机工程与设计,2014,35(5):1526-1530.
- [59]刘晓晨,张静.基于改进 BP 神经网络的室内无线定位方法[J].计算机应用与软件,2016,33(6):114-117.
- [60]石晓伟,张会清,邓贵华.基于 BP 神经网络的距离损耗模型室内定位算法研究[J].计算机测量与控制,2012,20(7):1944-1947.
- [61]吴超,张磊,张琨.基于 BP 神经网络的 RFID 室内定位算法研究[J].计算机仿真,2015,32(7):323-326.
- [62]沈冬冬,李晓伟,宋旭文,等.基于多层神经网络的超宽带室内精确定位算法[J].电子科技,2014,27(5):161-163.
- [63]龚阳,崔琛,余剑,等.基于 RBF 神经网络的室内定位算法研究[J].电子测量技术,2016,39(10):57-60.
- [64]宋慧敏,杨社堂,赵栋栋.基于人工神经网络的无线传感器定位算法[J].计算机测量与控制,2014,22(2):473-475.
- [65]曲婧.基于神经网络的高速飞行体定位技术研究[J].电子测试,2010(2):27-30.

- [66]张宝军,毛永毅.基于 RBF 神经网络的泰勒级数展开定位算法[J].计算机工程与应用,2009,45(15):245-248.
- [67]覃玉清.基于深度学习的 WIFI 定位算法[D].南京大学,2014.
- [68]刘侃,张伟,张伟东,等.一种基于深度神经网络的无线定位方法[J].计算机工程,2016,42(7):82-85.
- [69]He K, Zhang X, Ren S, et al. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification[J]. 2015:1026-1034.
- [70]Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators[J]. Neural Networks, 1989, 2(5):359-366.
- [71]Hornik K, Stinchcombe M, White H. Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks[J]. Neural Networks, 1990, 3(5):551-560.
- [72]SRIVASTAVA N, HINTON G, KRIZHEVSKY A, et al. Dropout: a simple way to prevent neural networks from overfitting[J]. Journal of Machine Learning Research, 2014, 15(1):1929-1958.
- [73]WARDEFARLEY D, GOODFELLOW I J, COURVILLE A, et al. An empirical analysis of dropout in piecewise linear networks[J]. Computer Science, 2013.
- [74]LI Y, GAL Y. Dropout Inference in Bayesian Neural Networks with Alpha-divergences[J]. 2017.
- [75]KRUEGER D, MAHARAJ T, KRAMÁR J, et al. Zoneout: Regularizing RNNs by Randomly Preserving Hidden Activations[J]. 2016.
- [76]CHENG G, PEDDINTI V, POVEY D, et al. An Exploration of Dropout with LSTMs[C]// INTERSPEECH. 2017:1586-1590.
- [77]Polyak B T. Some methods of speeding up the convergence of iteration methods [J]. Ussr Computational Mathematics & Mathematical Physics, 1964, 4(5):1-17.
- [78]Sutskever I, Martens J, Dahl G, et al. On the importance of initialization and momentum in deep learning[C]// International Conference on International Conference on Machine Learning. JMLR.org, 2013:III-1139.

- [79]Schaul T, Antonoglou I, Silver D. Unit Tests for Stochastic Optimization[J].
Nihon Naika Gakkai Zasshi the Journal of the Japanese Society of Internal
Medicine, 2013, 102(6):1474-83.
- [80]Nilchi A R N, Mathews V J. An efficient algorithm for joint estimation of
differential time delays and frequency offsets[C]// IEEE International Conference
on Acoustics, Speech, and Signal Processing. IEEE, 1992:309-312 vol.5.
- [81]张文. 无人机时差频差定位关键技术研究[D]. 国防科学技术大学, 2011.
- [82]Knapp C, Carter G. The generalized correlation method for estimation of time
delay[J]. IEEE Transactions on Acoustics Speech & Signal Processing, 2003,
24(4):320-327.

攻读硕士学位期间发表的学术论文