

分类号: _____

密级: _____

UDC : _____

编号: _____

工学硕士学位论文

基于图像声纳的动目标检测技术研究

硕士研究生：董 静

指导教师：孙大军 教授

学科、专业：水声工程

论文主审人：黄益旺 教授

哈尔滨工程大学

2014 年 3 月

分类号：_____

密级：_____

UDC：_____

编号：_____

工学硕士学位论文

基于图像声纳的动目标检测技术研究

硕士研究生：董 静

指导教师：孙大军 教授

学位级别：工学硕士

学科、专业：水声工程

所在单位：水声工程学院

论文提交日期：2014 年 2 月 27 日

论文答辩日期：2014 年 3 月 15 日

学位授予单位：哈尔滨工程大学

Classified Index:

U.D.C:

A Dissertation for the Degree of M. Eng

Research on Moving Targets Detection Based on Image Sonar

Candidate: Dong Jing

Supervisor: Prof. Sun Dajun

Academic Degree Applied for: Master of Engineering

Specialty: Underwater Acoustic Engineering

Date of Submission: Feb. 27, 2014

Date of Oral Examination: Mar. 15, 2014

University: Harbin Engineering University

哈尔滨工程大学

学位论文原创性声明

本人郑重声明：本论文的所有工作，是在导师的指导下，由作者本人独立完成的。有关观点、方法、数据和文献的引用已在文中指出，并与参考文献相对应。除文中已注明引用的内容外，本论文不包含任何其他个人或集体已经公开发表的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

作者（签字）：

日期： 年 月 日

哈尔滨工程大学

学位论文授权使用声明

本人完全了解学校保护知识产权的有关规定，即研究生在校攻读学位期间论文工作的知识产权属于哈尔滨工程大学。哈尔滨工程大学有权保留并向国家有关部门或机构送交论文的复印件。本人允许哈尔滨工程大学将论文的部分或全部内容编入有关数据库进行检索，可采用影印、缩印或扫描等复制手段保存和汇编本学位论文，可以公布论文的全部内容。同时本人保证毕业后结合学位论文研究课题再撰写的论文一律注明作者第一署名单位为哈尔滨工程大学。涉密学位论文待解密后适用本声明。

本论文（☐在授予学位后即可 ☐在授予学位 12 个月后 ☐解密后）由哈尔滨工程大学送交有关部门进行保存、汇编等。

作者（签字）：

导师（签字）：

日期： 年 月 日

年 月 日

摘 要

伴随海洋资源的开发和水下探测技术的飞速发展,声纳探测设备凭借其得天独厚的优势成为水下探测的重要组成部分。利用声纳设备进行水下目标检测识别是目前海洋研究的重要方向,其中图像声纳利用声成像技术将水下运动目标信息转换为直观的序列图像呈现在人们面前,使之成为水下动目标检测领域的重要基础。由于水下声环境复杂,图像声纳采集到的水下声图像通常存在信噪比低、目标信息残缺等问题,因此对于声纳图像上的运动目标检测具有十分重要的意义与广阔的应用前景。

本文首先进行了声成像算法的分析与仿真,并针对声纳图像存在的问题研究了两种图像预处理的技术,对不同的图像预处理方法通过仿真进行性能验证。图像分割是图像目标检测的最关键技术之一,因此研究了基于单帧图像检测的两种常见图像分割算法(固定阈值分割法和天津阈值分割法),通过实验发现这两种算法对目标的检测性能较差。因此,研究了恒虚警(Constant False Alarm Rate, CFAR)检测法,经过对 CFAR 算法的理论、模型以及检测性能详细研究后,将 CFAR 算法应用于常规分割方法未能检测出目标的声纳图像中,并成功检测出这些图像中的目标。

图像声纳的动目标检测是基于声纳图像序列的,因此本文进行了多帧图像的动目标检测研究。研究了帧间差分法和背景估计法,并对其检测性能进行了仿真分析。其中背景估计法的重点就是对序列图像的背景建模,本文分析了单高斯背景建模法和混合高斯背景建模法,针对水下环境变化复杂的情况,通过混合高斯背景建模法建立背景的置信区间,实现静目标的剔除和背景噪声的抑制,最后与 CFAR 检测法相结合进行动目标的检测。仿真试验与水池数据处理结果表明,本文中基于图像声纳的水下动目标检测算法具有有效的检测能力。

关键词: 图像声纳; CFAR 检测技术; 背景估计; 动目标检测

ABSTRACT

Accompany the development of marine resources and underwater detection technology, sonar detection devices become an important part of underwater detection by virtue of its unique advantage. Using sonar devices to detect and recognize underwater targets is an important direction of recent ocean research, image sonar using acoustical imaging technology transforms underwater moving targets' information into intuitive sequence of images making it an important basis of underwater moving target detection field. Because of the complex underwater acoustic environments, underwater acoustic images collected by image sonar usually have the problem of low signal-to-noise ratio, targets information incomplete, so the moving object detection on sonar image plays a very important role and has broad application prospects.

Firstly, acoustical imaging algorithms are analyzed and simulated. Aiming at solving the problems of sonar images, two image processing techniques have been researched, and performance validation of different image pre-processing methods are discussed through simulation. Image segmentation is one of the most key techniques of image target detection, study of two common image segmentation(such as: fixed thresholding method and Otsu thresholding method)has been conducted based on single image detection. We find these two algorithms have poor detection performance through experiments. Therefore, CFAR (Constant False Alarm Rate, CFAR) detection method is studied. After a detailed research of CFAR's theories, models and detection performance, apply CFAR algorithm to the sonar images which conventional sonar image segmentation algorithms fail to detect in, and successfully detected these targets.

Moving target detection in sonar images is based on sonar image sequences, so a moving target detection research on multiple-frame images is carried out. Research of frame difference method and background estimation method has been conducted, and analyzes their detection performance by simulating. Among them, background modeling is the key of background estimation method. The Gaussian background modeling method and mixture Gaussian background modeling method are analyzed. Because underwater environment is complex, the confidence interval of background by mixture Gaussian background modeling is established, the culling of static goals and restraining of background noise is completed, and detect the targets end with CFAR detection algorithm. The processing result of simulation and pool experiments show that this underwater moving target detection algorithm based on sonar image has effective detection capability.

Key words: image sonar; CFAR detection technology; background estimation; moving target detection

目 录

第 1 章 绪论	1
1.1 引言	1
1.2 研究背景	1
1.3 动目标检测技术	3
1.4 本文的主要内容	4
第 2 章 声成像原理及图像预处理技术	6
2.1 引言	6
2.2 声成像技术	6
2.2.1 正交变换	6
2.2.2 宽带波束形成	8
2.2.3 契比雪夫加权	9
2.2.4 近场聚焦	9
2.3 声纳图像的预处理	12
2.3.1 图像的增强	12
2.3.2 图像的去噪	14
2.4 本章小结	18
第 3 章 基于图像声纳的二维检测技术	19
3.1 引言	19
3.2 图像分割	19
3.2.1 固定阈值分割	19
3.2.2 大津阈值分割	21
3.3 恒虚警(CFAR)检测法	23
3.3.1 瑞利分布模型	24
3.3.2 CFAR 概述	25
3.3.3 CFAR 检测模型	28
3.3.4 单帧声纳图像的 CFAR 检测	30
3.4 本章小结	33
第 4 章 基于图像声纳的动目标检测技术	35
4.1 引言	35
4.2 帧间差分法	35
4.3 背景估计法	37
4.4 统计背景建模法	38

4.4.1 单高斯背景模型	39
4.4.2 混合高斯背景模型	40
4.4.3 建立背景置信区间	45
4.5 本章小结	50
结 论	51
参考文献	52
攻读硕士学位期间发表的论文和取得的科研成果	55
致 谢	56

第1章 绪论

1.1 引言

蔚蓝的海洋占据了地球 71% 的面积，其中蕴含了丰富的各类矿产、水产资源等。随着对海洋资源的开发研究和海洋环境保护事业的快速发展，我们逐渐意识到海洋重要的战略地位及资源价值。中国约有 18000 公里的海岸线，且大多数经济发达的城市及重要的军事港口都分布在我国沿海地区，因此，对近岸和港口的防御工作已刻不容缓。以前，对海上的探测与防御工作具有较强的针对性，其主要目标是大型的舰船和潜艇等^[1]。而现在，水中蛙人、自主水下航行器(AUV)和小型的水下运载器等运动目标，常常利用薄弱的水下安保环节，实施对海港水域的破坏活动^[2]。因而，世界各国为了对抗动目标的侵入威胁，都在大力发展对动目标入侵的探测技术，对水下运动目标的检测与防御俨然成为海防的重要环节。

当今的水下探测技术主要有：红外探测技术、光探测技术、磁探测技术与声探测技术。而我们知道，在海洋中，水对电、光的等信号的吸收比率远远高于空气中的吸收比率，但声信号却能以比在陆地上快 4 倍多的速度在水下进行传输，可见水下环境是水声信号得天独厚的应用背景^[3]。声纳是指利用声波探测和定位水下目标的设备^[4]，声纳系统在水下导航定位与目标检测的研究中具有举足轻重的地位，其中图像声纳通过图像的形式反映目标的信息，成为声成像系统中的重要组成部分。

动目标检测技术是一门综合了计算机视觉、图像处理等多个学科，能够通过一系列的图像序列，从复杂背景中检测出运动目标的技术。因为声纳图像中信息含量较少，且存在严重的噪声污染，导致从声纳图像中实现水下运动目标的检测目的难度很大^[3]。因此，从技术研究及实际应用等方面来说在声纳图像中进行水下动目标检测都具有重要的意义。

1.2 研究背景

本文以图像声纳对动目标的检测进行港口预警为研究背景。过去，水声定位系统一般使用多个接收阵元，然后对接收到的数据用解定位方程的方法求得水下目标的方位信息。目前，这一技术随着科技的发展已经从单目标转向多目标。声成像系统近年来在水下目标的检测研究中发展迅速，不同于传统的水声定位系统，声成像系统不用安装同步时钟或应答器，操作便捷，缺点是作用距离较短。

目前，对水下声成像系统的研究已日渐成熟，尤其是一些起步较早的国家都已经成功研制出许多成熟的产品。已有的且在实际应用中表现较为出色的高分辨率前视声纳有：来自美国的 Blue View Technologies 公司和 Reson 公司，以及来自加拿大的

IMAGENEX 公司生产的产品；而侧扫声纳有：来自英国的 GeoAcoustic 公司、来自美国的 EdgeTech 公司和来自德国的 GEOMAR 公司生产的产品；三维成像声纳则主要是来自英国的 Coda Octopus 公司和 Tritech 公司、来自美国的 Klein Associates 公司的产品等^[5]。

Echoscope Mark II 是英国声纳设备公司—Coda Octopus 研制的一款三维回波图像声纳。该声纳内含 2034 路电子通道，相控二维基阵的大小为 48×48 ，可同时并行发出 16384 路波束，最远可获取 100 米左右的目标的声纳图像，同时具有高达 20 帧/秒的图像刷新率。该型号的图像声纳在水下建筑物、海港码头、船体外壳的成像中具有较广泛的应用，还能探测港口的水下蛙人^[5]。

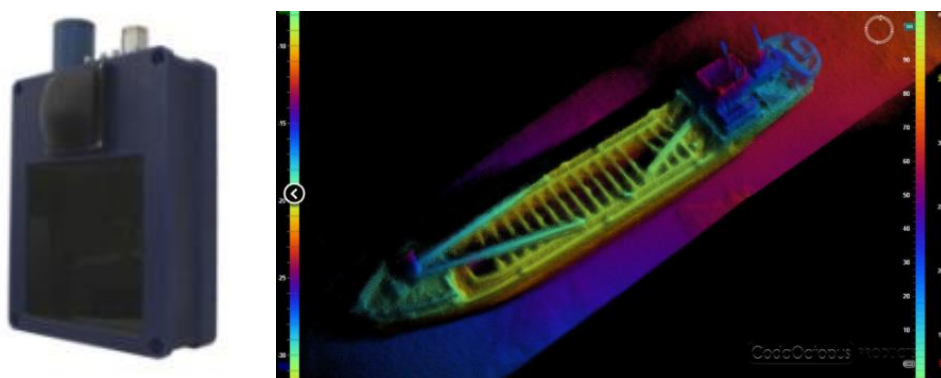


图 1.1 Echoscope Mark II 型声纳外观和成像效果图

因为水声信道是时变和空变的，因此图像声纳的中心频率一般都保持在几百千赫兹以上，这样才能满足对声图像分辨率的要求^[6]。而另一方面，海水中的介质对声能量以其中心频率增长的二次幂进行吸收，此外，在声传播过程中还伴随着体积扩散，导致在海水中声波的高频部分衰减严重^[7,8]。因此，由声成像系统得到的声纳图像质量存在分辨率低、图像中目标信息残缺等问题，造成目标检测的困难。

不同于光学图像，声纳图像具有以下的特点^[3]：

1. 图像的分辨率较低：由于声纳设备发出的探测声波的波长较长，同时，声纳头以较大的步进角度进行探测，造成相邻波束间的变化并不连续，而存在一定的间隔距离，这也就导致声纳图像的分辨率比常见的数字图像低。

2. 严重的噪声干扰：我们知道，海洋环境和水声信道十分复杂，声波在这种复杂的环境下传播会受到来自混响、衰减、散射等的干扰。此外，存在于海洋环境中的其他噪声（如：海洋生物发出的声音、水下火山爆发或者地震等）也干扰声纳图像，导致目标被这些干扰淹没，造成目标检测的困难。

3. 多途：声波在水中传播时，目标回波通过不同的途径抵达接收声纳，导致接收声纳设备得到目标的多个回波，且接收声纳将这些回波当作相距较近的多个目标来处理，使得在生成声纳图像的过程中在图像中出现重影现象，无法得到目标的真实位置。

4. 旁瓣干扰问题：在进行波束形成时，声纳系统不仅在当前较窄的扫描扇面区内接收到目标回波的极大值，而且也能在其旁边较宽的扫描区内接收到信号。当出现较强的目标回波时，在与目标回波等距离的附近波束上会出现一条回波亮线，这就是旁瓣干扰。

5. 存在多普勒效应：当声纳系统的载体或被探测的目标进行运动时，将导致声纳发射和接收到的声波出现偏移现象，进而导致不同时刻的声纳图像上同一个目标的轮廓形状不同，造成探测到的目标出现失真，甚至变形。

由于声纳图像的以上特殊性，直接在声纳图像中运用针对光学图像的各种去噪、分割处理算法效果大大降低，而且目前还没能形成较统一的、公认的声纳图像分割算法。近年来，研究较多的声纳图像分割算法包括：Bayesian 理论^[9]，马尔科夫 Gibbs 随机场^[9,10]，多尺度边缘检测，随机场(MRF)^[11]等。

1.3 动目标检测技术

对于动目标检测算法的研究，根据国内的文献总结可以将已有的动目标检测算法分为三类：光流法^[12]，相邻帧差分^[13,14]和背景差分法^[15,16]。背景差分法是一种典型的运动检测算法，其基本思想是对当前图像和背景图像进行差分运算，根据得到的差分图像分离出运动目标信息^[17]。该算法思想简单、容易实现，可以又快又好地从序列图像中得到动目标信息。但该算法的难点在于背景图像的选择，在实际应用中，由于光照变化、水波流动等都会造成背景图像的变化，也就无法得到完全静止的背景图像，这样就降低了背景差分法的准确性。对此，国内外的专家、学者为了改进该方法的检测效果进行了大量的研究，也提出了许多改进算法，如对背景模型进行重建、更新等。当场景中的背景变化缓慢时，混合高斯模型常被用于描述此时的背景，而如果场景中光线或者水流变化较快时，混合高斯模型的效果将大大降低^[18]。一种可以进行自适应更新的混合高斯模型被 Stauffer 与 Grimson 等提出，该模型的实时性较好^[16]。相邻帧差分法是利用两帧图像之间的差异来进行运动目标的分离，由于这种方法计算速度快，比较适宜硬件实现，因而也得到了很广泛的应用。Litpon 等人成功通过对实际视频中相邻的两帧图像做差分检测出运动目标^[13]。而光流法的优点是无需预先知道被检测场景内的任何信息，就可检测出出现在场景内的独立运动目标，缺点是算法所需运算时间较长，在实时性要求较高的场合不适用。为此，研究人员经过大量探索对各种光流法进行了改进，提出区域光流法^[19]、金字塔光流法^[20]等。

国外在对动目标检测的研究中，除了上述的三种方法外，还通过分析多路数据来进行动目标检测，也就是多传感器分析法。Varadarajan 2005 年在 Acoustics, Speech, and Signal Processsing, 2005 上发表的一篇名为 Target detection using dynamically reconfigurable sensor arrays 的文章与 Krishna 等在 Intelligent Robots and system, 2005 发

表名为 t-step ahead constrained optimal target detection algorithm for a multisensor surveillance system 一文中都介绍了多传感器在目标检测中的应用；Varadarajan 的论文中通过 Dynamically reconfigurable sensor arrays(DRSAs)进行动目标检测，文中直接利用噪声域得到 DRSAs，然后根据动态情况下选择出的最佳方向，指定某个传感器的 DRSAs 值最大，最后将其作为目标检测的依据；后一篇中通过利用多个传感器对某个矩形区域进行监控，此时运动物体将按泊松分布以匀速进入该监控区域^[21]。该文通过某种坐标机制对传感器的分布和方向进行控制，保持被监控区域的重复面积最小。

目标的阴影检测与抑制也是运动目标检测研究中不容忽视的问题之一。目标的阴影与背景不同，它随着目标的运动而运动，容易被误判为目标。所以，在目标检测中需要抑制阴影，为后续的跟踪、识别打好基础。McKelma 等通过将亮度信息与归一化颜色分量相结合的方法实现阴影的检测与抑制；Brisson 等利用混合高斯模型在 YUV (Y-亮度，U、V-色度)颜色空间上对阴影部分进行建模。Prati 等人则将几种阴影检测的抑制算法作了详细的分类比较。Leone 等通过对纹理的分析进行活动阴影的检测。而 Stauder 等则通过分析活动阴影的几何特征完成对视频序列图像中活动阴影的检测。Bai 等人依据背景在阴影的影响下在 HSV (H-色度，S-饱和度，V-亮度)颜色空间会变暗，并且根据背景的亮度在阴影覆盖前后的比值近似线性这一现象，实现阴影的抑制^[22]。不过，已有的这些阴影抑制方法都可能在消除阴影的同时将与阴影相似或者比阴影色度更低的目标区域也消除，从而使目标检测的结果中出现空洞。

除此之外，还有大量其他的运动目标检测方法^[23]。其中，较为典型的有块比较法、张量投票法、点状相关图分析法等，限于篇幅就不赘述了。虽然运动目标检测算法的种类很多，但是对声纳图像的动目标检测算法还没有成形，我国在声纳图像的动目标检测的研究还处于初级阶段，并且落后于发达国家，这就需要我们在今后的研究中不断地探索、发展完善。

1.4 本文的主要内容

本文主要围绕图像声纳对动目标检测进行研究。通过图像声纳成像得到水下运动目标的序列图像，进行动目标检测方法的研究。全文的研究内容分为四章：

第一章为绪论部分，主要对文章的研究背景进行介绍，特别对声纳图像的特点进行归纳，介绍了三种常规的动目标检测算法，还详述动目标检测算法的研究现状。最后，合理安排文章的研究内容和组织结构。

第二章对基于波束形成的声成像算法进行理论概括与仿真分析，由于声纳图像受到噪声的严重干扰，图像分辨率低，目标边缘残缺，为了抑制声纳图像中的噪声，突出目标边缘与细节，更好的检测目标，对声纳图像的预处理方法进行研究并进行仿真分析。

第三章首先对两种常规的图像分割算法进行了研究，即固定阈值分割算法、大津阈

值分割算法。由于声纳图像情况复杂，上述的算法无法较好的分割出目标，因此研究了恒虚警(CFAR)检测法。CFAR 检测法是一种自适应的门限检测算法，根据接收器接收到的回波数据进行干扰的统计特性估计，在此基础上算出检测门限，最后与得到的检测单元杂波样本进行比较，实现目标检测过程中的恒虚警性。利用三种检测法对水池数据进行处理，得到 CFAR 检测法在水池数据处理中的可行性。

第四章介绍了包括帧间差分法和背景估计差分法的动目标检测原理，对背景估计的动目标检测法中背景图像的获取进行详细研究，根据海洋混响包络的分布特性进行统计建模，分析单高斯模型与混合高斯模型的特点。最后通过混合高斯模型建立背景置信区间，与第三章中的 CFAR 检测法相结合进行动目标的检测，完成了仿真与水池数据处理。

最后进行论文总结，主要概括文章的研究成果，对进一步的研究工作进行展望。

第2章 声成像原理及图像预处理技术

2.1 引言

常用的主动声成像系统中，声纳基阵发射声波，遇到目标后声波发生反射，然后声纳成像系统接收到这些反射波，对接收数据按空间位置进行排列，得到二维数组^[24]。这个数组中的数据代表的就是声回波信号的幅度，回波幅度因为目标材质、距离等的不同产生变化，因此这个数组就是一幅能表达目标特性的图像^[24]。由此可见，声纳图像其实就是声纳系统接收到的目标声回波能量在平面中的分布，是与光学图像类似的二维空间信号。其中，波束形成技术是图像声纳成像算法的核心^[24]。本章首先进行波束形成的声成像算法理论概括与仿真，并对常用的声纳图像预处理方法进行比较分析。

2.2 声成像技术

水下声成像技术是进行水下探测的重要环节，与其他声成像原理相同，它也是根据物体对基阵发出的声波的反向散射进行成像的。常用的水下声成像方法有三种：波束形成法、声透镜法、声全息法，其中波束形成法已经十分成熟，且在图像声纳中已有较广泛的应用，故而本文主要研究波束形成的声成像算法^[25]。

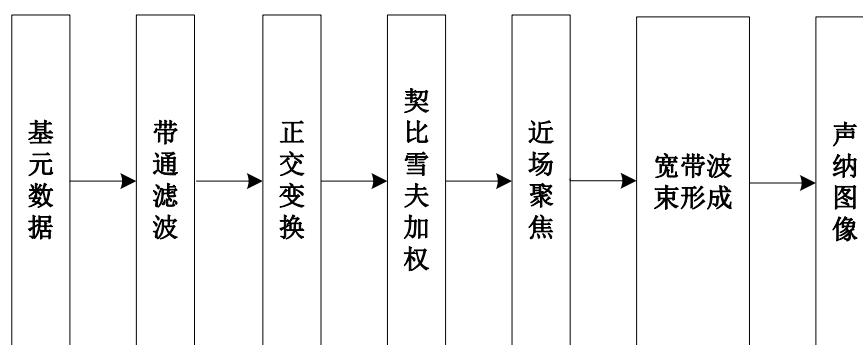


图 2.1 声成像算法流程图

一般情况下，基元数据依次经过带通滤波、正交变换、加权聚焦和波束形成等之后就可成像。而其中波束形成又是成像的核心，对一定几何形状（如直线、圆柱、弧形等）排列的多元基阵，波束形成就是对其各个阵元输出进行（如加权、求和、时延等）处理从而形成空间指向性的方法^[4]。

2.2.1 正交变换

正交变换能除去载波信号调制后的信号，将信号转换到零中频（基带），实信号频

谱的是具有共轭对称性的双边谱，可用 $x(t) = x^*(-t)$ 表示，也就是说，频谱上的正频率分量和负频率分量相位相反，且是关于频谱对称的^[26]。若对一个目标回波信号（实信号）直接进行FFT波束形成处理，那么波束形成之后图像中不仅有一个真正的目标，同时也会存在一个关于 0° 角对称的虚目标，造成目标识别的困难，此时，应该采用其单边谱信号进行波束形成。在时域上，复信号就是单边谱信号，所以在进行波束形成之前应把复信号从实信号中分解出来，这时的复信号中携带所有的信息（幅度和相位）^[27]。

任意的一个实信号 $x(t)$ 都可分解出一个与之对应单边谱信号 $\tilde{x}(t)$ ：

$$\tilde{x}(t) = x(t) + j \cdot \hat{x}(t) \quad (2-1)$$

其中 $\hat{x}(t)$ 是 $x(t)$ 的 Hilbert 变换。

实际应用中所接收到的目标回波，其中心频率远远大于频带宽度，可被近似的看作高频的窄带信号。假设一个高频窄带信号 $x(t)$ 可以表示为：

$$x(t) = a(t) \cos[\omega_0 t + \varphi(t)] \quad (2-2)$$

其中 $\varphi(t)$ 是 $x(t)$ 的相位函数， $a(t)$ 为 $x(t)$ 的包络函数。对于载波 $\cos \omega_0 t$ 来说， $a(t)$ 和 $\varphi(t)$ 都是包含了目标信息且变化缓慢的时间函数^[27]。

由公式(2-1)和(2-2)公式求出 $\tilde{x}(t)$ ：

$$\tilde{x}(t) = a(t)e^{j[\omega_0 t + \varphi(t)]} = \tilde{m}(t)e^{j\omega_0 t} \quad (2-3)$$

其中 $\tilde{m}(t) = a(t)e^{j\varphi(t)} = a(t)\cos\varphi(t) + j \cdot a(t)\sin\varphi(t) = m_c(t) + j \cdot m_s(t)$ ，称为信号 $x(t)$ 的复包络或复基带信号。由基带波束形成理论可知，利用正交变换能够得到复基带信号，具体过程如图2.2所示。

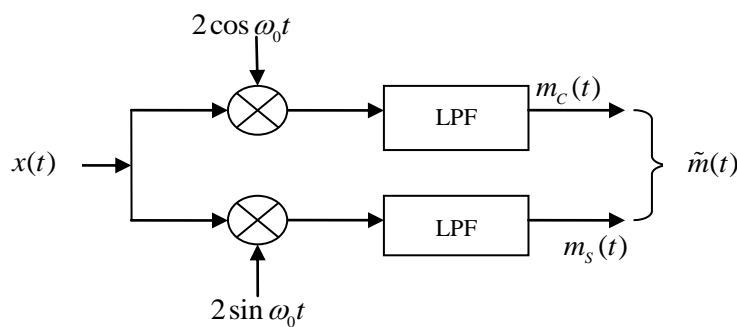


图 2.2 正交变换获取复信号

换言之，用频率为 ω_0 的复信号 $\tilde{x}(t)$ 对 $e^{-j\omega_0 t}$ 进行调制时，可获取复包络 $\tilde{m}(t)$ 。也就是复信号 $\tilde{x}(t)$ 线性搬移后得到包含目标信息的 $\tilde{m}(t)$ 的频谱，因此仅需对 $\tilde{m}(t)$ 进行波束形成处理。

2.2.2 宽带波束形成

本文进行的宽带波束是通过一个窄带带通滤波组将宽带信号分成若干窄带信号,对窄带信号进行FFT波束形成,此时得到每一个窄带信号的空间谱,再将每个窄带信号的空间谱能量相加成宽带信号的空间谱,则得到的输出就是声纳图像,如图2.3。

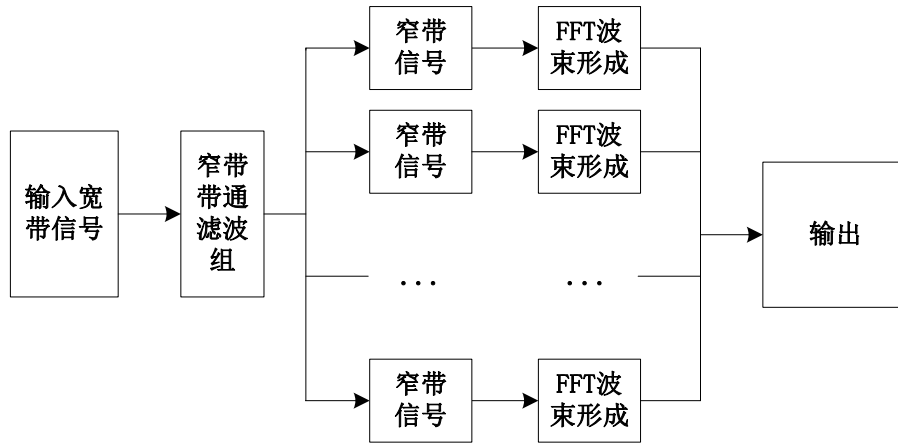


图 2.3 宽带波束形成流程图

下面对FFT波束形成技术进行介绍,波束形成器可以抑制其他方位的信号而只对空间上某个方位的信号做出响应。相移波束形成和时延波束形成是目前最为常见的两种波束形成方法,相移波束形成是通过在阵元间插入不同的相移量对主极大值的方位进行控制,而时延波束形成是通过在阵元间插入不同的时延量对主极大值的方位进行控制^[28]。但是在进行相移和时延波束形成时所需的计算量巨大,不利于图像声纳对实时性的要求。为此,下面对FFT快速波束形成方法进行介绍。

一个等间隔的 N 元直线阵,阵元间隔为 d ,那么该阵的第 m 个阵元的复数输出信号用下式表示:

$$X_m(t) = w_m e^{j\omega t} \cdot e^{jk\varphi} (-\pi \leq \varphi \leq \pi), (m=0,1,\dots,N-1) \quad (2-4)$$

式(2-4)中的 $\varphi = (2\pi d/\lambda) \sin\theta$ 表示相邻阵元所接收到的信号之间的相位差,其中信号波长是 λ ,角频率是 ω , θ 表示入射声线和基阵的法线之间的夹角, w_k 代表各个阵元的幅度加权系数^[28]。如果将相移 β_r 插入相邻两个阵元间,那么各阵元此时的输出为:

$$X_m(t) = w_m e^{j\omega t} \cdot e^{jm(\varphi - \beta_r)} \quad (2-5)$$

不记时间因子,基阵的输出为:

$$A_r = \sum_{m=0}^{N-1} X_m e^{-jm\beta_r} \quad (2-6)$$

β_r 作为一个补偿相位,可反映出波束被控制的方向,当 $\beta_r = 2\pi r/N$ 时,线阵输出

可表示为:

$$A_r = \sum_{m=0}^{N-1} X_m e^{-jmr \frac{2\pi}{N}} \quad (2-7)$$

由公式(2-7)可以看出,此时的基阵的输出实际上为离散傅里叶变换的形式。此时可以认为,对一个等间隔直线阵各波束输出值的计算就等于对 $X_m(t)$ 的离散傅里叶变换的计算^[29]。

当相移 β_r 被插入各个阵元之间时,主极大值被控制到:

$$\sin \theta_r = \frac{\beta_r \lambda}{2\pi d} = \frac{2\pi}{N} \cdot r \cdot \frac{\lambda}{2\pi d} = \frac{\lambda}{Nd} r \quad (2-8)$$

方向上。当接收到的信号在 $\theta = \theta_r = \arcsin(\lambda / (Nd)r)$ 方向上时, $|A_r| = 1$ 得到最大值,此时接收到的信号在各个阵元上实现同相相加。在同一时刻, X_m 是基阵中不同阵元在空域上对信号的等间隔采样,而对其多 FFT 达到了提取信号空间谱线的目的,此空间谱线是信号在空间中的能量分布^[30]。

2.2.3 契比雪夫加权

在图像声纳的声成像过程中,波束形成后旁瓣对声纳图像的影响也需要考虑。如果旁瓣干扰对主瓣过大,此时声纳系统就无法获得清晰的图像。加权可以通过改变阵元灵敏度或放大的前置放大器来实现。声纳系统的基阵的各个基元都具有不同的发射或接收灵敏度,基阵的指向性基于指向性乘积定理获得。为了提高基阵的指向性,通常可通过对阵元幅度进行加权使旁瓣对主瓣的影响降低。较常被使用到的加权标准是:满足给定主瓣宽度条件下,旁瓣最低;满足给定旁瓣高度的条件下,主瓣最窄;给定阵元个数的条件下,得到符合要求的主旁瓣高度比等^[27]。契比雪夫加权是在等间隔线阵加权中最常用的幅度加权方法,基于主旁瓣高度比,契比雪夫加权可以获得等旁瓣级;基于旁瓣级,契比雪夫加权可以获得最窄的主瓣;基于主瓣宽度,契比雪夫加权可以获取最低旁瓣级^[4]。但是,对幅度进行契比雪夫加权使旁瓣级降低需要增加主瓣宽度,所以不能一味坚持主旁瓣比,应该同时考虑主瓣宽度和主旁瓣比。

2.2.4 近场聚焦

根据第二节中内容可知,本文采用的是 FFT 快速波束形成方法,这种方法适用于声纳系统在远场条件下,且需假设信号是一个平面波。而在实际中声纳系统往往在近场工作,假定基阵的长度为 L ,目标距离为 r ,信号波长为 λ ,近场目标距离 r 满足下式:

$$r \leq 2L^2 / \lambda \quad (2-9)$$

此时,回波信号不是平面波而是球面波,波阵面依球面波规律衰减,若再用 FFT 波束形成会造成波束性能恶化,而无法形成准确的图像。因此,根据声波在近场中依球

面波规律传播的特点，阵元间相差相移的大小需按照球面波规律重新进行计算，如果在波束形成前补偿此相位差，便能获得效果较好的波束形成结果^[27]。

这里对典型的聚焦方法进行介绍：一种叫做正前方单波束聚焦，另一种叫做所有波束都聚焦。

1. 正前方单波束聚焦

如图 2.4，假设一个线阵，每一个线阵基元的位置用 $\overline{X_i}$ 表示，阵元之间的间距为 d ， \overline{F} 表示沿着声轴方向与阵中心的距离为 R 的焦点位置，参考阵元为 $\overline{X_k}$ ，各基元在 \overline{F} 处同相叠加^[31]。

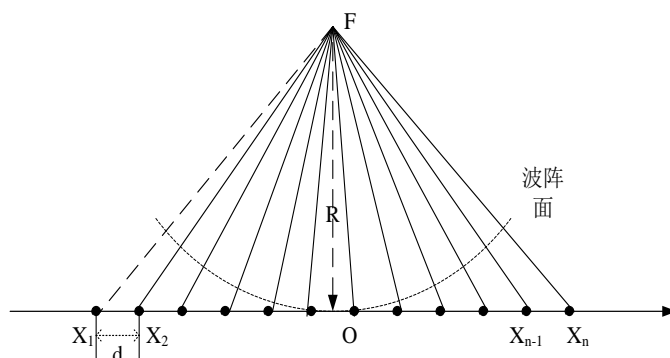


图 2.4 正前方单波束聚焦波束形成几何关系图

对于一个指定参考阵元，设 $\phi_1, \phi_2, \dots, \phi_N$ 为参考阵元与各阵元的接收信号之间的相位差，那么正前方单波束聚焦方法就是用 $\phi_1, \phi_2, \dots, \phi_N$ 对空间中其他方向的波束进行相位补偿。其中， $\phi_1, \phi_2, \dots, \phi_N$ 称作聚焦校正因子。对于平面阵来说，此时的相位补偿因子 $\phi_{n,m}$ 为任意阵元与参考位置（沿着声轴方向与阵中心距离为 R ）的声程差。

虽然对空间各个方位都进行相位补偿，但只有沿着声轴方向的波束才能实现同相相加。若波束偏离声轴的方位，那么偏离的方位越远产生的误差越大，此时通过正前方单波束聚焦法聚焦之后的波束变宽，同时主旁瓣比变小。由此可知，正前方单波束聚焦在基阵扫描角度小的情况下有较好的适用性。

2. 所有波束都聚焦

在波束形成之前，计算扫描空间内各个方位波束的聚焦校正因子，并用它们分别补偿其所对应的波束，这就是所有波束都聚焦。同样以线阵举例，如图 2.5：

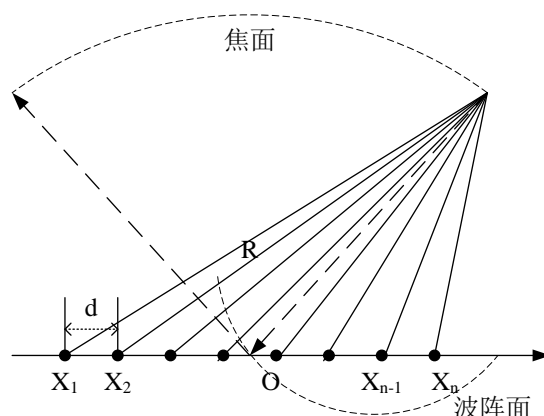


图 2.5 所有波束都聚焦波束形成几何关系图

从上图中可以看出，假设一个线阵， \bar{F} 是声源所在的位置。对于指定的参考阵元，设 $\phi_1, \phi_2, \dots, \phi_N$ 为该参考阵元与基阵中各阵元接收信号之间的相位差。那么计算扫描空间范围内所有方位的波束聚焦校正因子 $\phi_1, \phi_2, \dots, \phi_N$ ，并用其进行相位补偿，使各阵元的输出沿声源方位实现同相相加。一个阵元数为 N 的等间隔线阵，在其扫描空间上存在 N 个独立的波束，对扫描范围空间内 N 个方位都需要进行相位补偿。因此，对于这个阵元数为 N 的等间隔线阵需要对 $N \times M$ 个方位做补偿，此时其焦面在一个的球心与基阵中心重合，半径为 R 的切面上。但是所有波束都聚焦法运算量庞大，无法满足成像系统的实时性要求，实际工程中一般不采用该方法。

对于聚焦波束形成方法需要根据具体的波束形成条件进行选择，在成像区域较小的情况下，前者正前方单个波束聚焦比较适合采用，其只需要对空间内各个方位的波束补偿一次聚焦校正因子，就能进行 FFT 快速波束形成。在成像区域范围较大的情况下，后者所有波束都聚焦比较适用。本文中使用的就是正前方单个波束聚焦方法。

本文的成像算法是基于宽带信号的波束形成的，仿真条件为：信号频率 25kHz，脉宽为 2ms，带宽 10kHz，阵元个数为 16，间距为 7cm，采样频率 200kHz，且目标位于 10° ，9m 处，声成像的仿真结果如下：

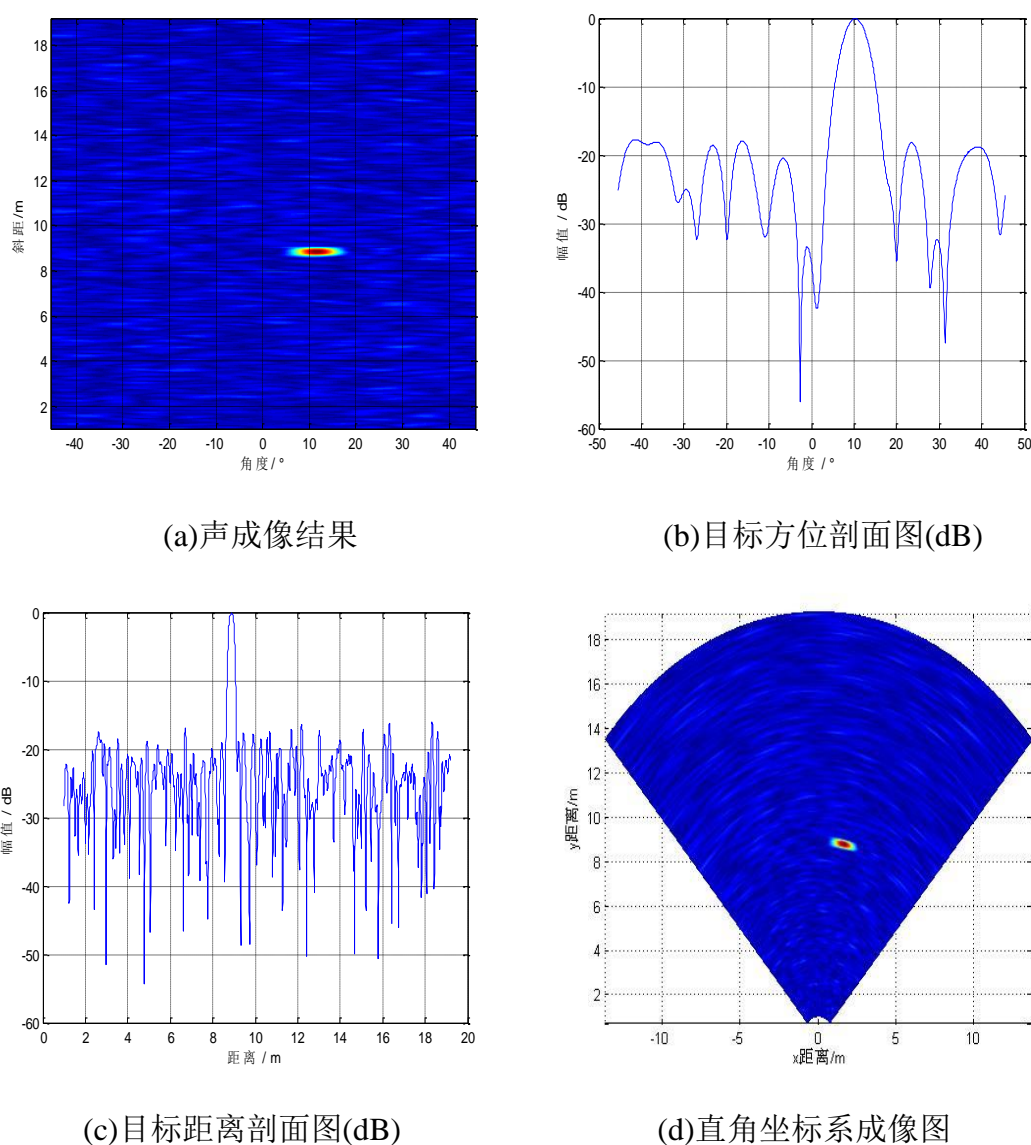


图 2.6 声成像仿真图

2.3 声纳图像的预处理

一般情况下，声纳图像由三部分组成：海底混响区、阴影区和目标区^[32]。海底混响噪声严重，具有多途、起伏、传播损失和多普勒效应等问题，它们对声纳图像产生严重的噪声干扰，使得图像分辨率低，目标边缘残缺，对目标检测带来极大的阻碍。为了抑制声纳图像中的噪声，突出目标边缘与细节，我们需要对声纳图像进行增强去噪的预处理，主要采取的方法有：直方图均衡化、灰度变换、滤波等^[33]。

2.3.1 图像的增强

图像增强指运用一系列技术实现图像视觉观测效果的提高，或者是把图像转成更适于人眼观察的形式^[33]。图像质量一般由使用者进行主观评判，并无统一的客观标准。图

像的增强方法通常具有很强的针对性,需要反复进行多次实验才能得到一种方法,在没有图像的先验知识的情况下,很难预测具体方法的使用效果。较常采用的方法是根据图像的具体情况将几种增强技术组合使用^[33]。

声纳图像中,像素点的像素值表示对应空间内回波的强弱。在处理声纳图像的时候,我们采用灰度级变换的方法进行增强操作。灰度级变换是指通过统一的方法对整个图像或者图像的某一区域的灰度进行调整,实现图像的对比度的增强,提高图像质量。因此,灰度变换又被称为对比度增强或者对比度的拉伸^[34]。

由变换函数的不同形式,可以把灰度变换划分为非线性变换与线性变换,而常用的非线性变换有两种:对数函数变换与指数函数变换。当选择对数函数时,图像的高灰度区域可通过灰度变换实现压缩,低灰度区域实现扩展;当选择指数函数时,图像的低灰度区域可通过灰度变换进行压缩,高灰度区域实现扩展;当选择线性函数时,可以将图像的各像素值进行同等比例扩展^[33]。

另一种图像增强方法就是直方图均衡化,所谓的直方图就是统计一幅图像中所有像素的灰度分布,然后将灰度值用横坐标表示,与每一个灰度值对应的分布概率用纵坐标表示^[33]。图像的直方图能反映出一幅图像的整体性质,比如整个图像的明暗、图像细节的清晰度等等。当图像在整个灰度级上都有分布,且分布较为均匀时,该图像就是一幅明暗细节较好的图像。

然而,当我们所得到的图像并不都是明暗效果良好的。图像直方图的纵坐标表示的是对应像素灰度在整个图像中的比例,该峰值的大小表示该灰度值上像素个数的多少。当分布偏右时,表示图像整体较亮,有可能过度曝光;当分布偏左时,表示图像缺少明亮细节,可能是曝光不足带来的。此时,为了提高画面质量,我们可以进行直方图均衡化处理^[34]。直方图均衡化就是对一幅画质不太良好的图像,缩减图像中分布较少的灰度级,而展宽分布较多的灰度级,最终实现图像清晰度改善的目的。

仿真条件为:信号频率 25kHz,脉宽为 2ms,带宽 10kHz,阵元个数为 16,间距为 7cm,采样频率 200kHz,加入 5dB 的随机噪声,且目标位于 10°, 9m 处,对于原始图像图 2.7 中(a),用对比度拉伸的方法进行增强处理,结果如图 2.7 中(b):

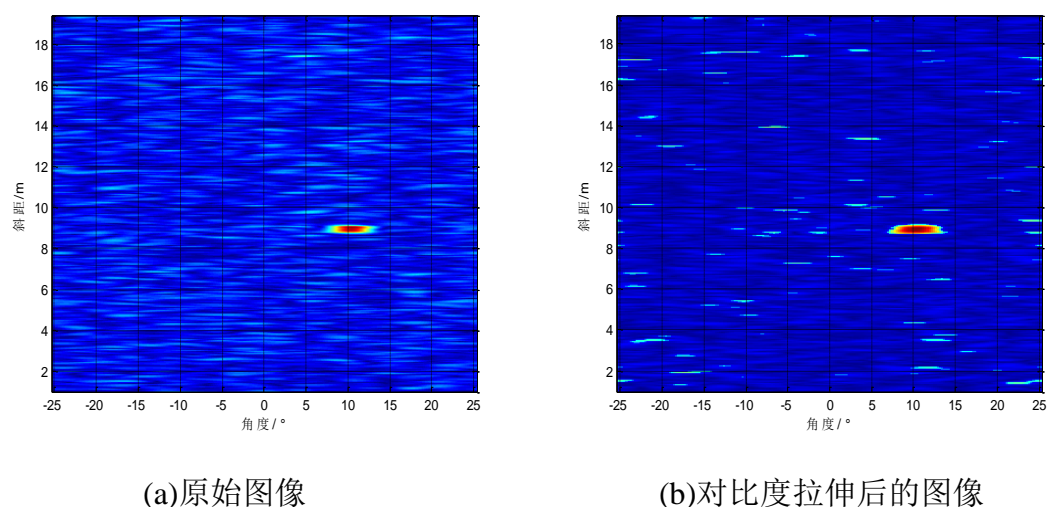


图 2.7 图像增强结果

观察上述处理结果，从图 2.7(b)中可以看出经过对比度拉伸后，改善了图像的画面效果，使图像更加清晰，并且图像中我们感兴趣的部分得到了加强，可以更加清晰的获取目标信息。图 2.8 表示的是原始图像与均衡化后图像的直方图，原始图像的直方图中峰值分布偏左，通过图 2.8 中(b)可以明显的看出此时的直方图均匀分布在灰度级范围中。

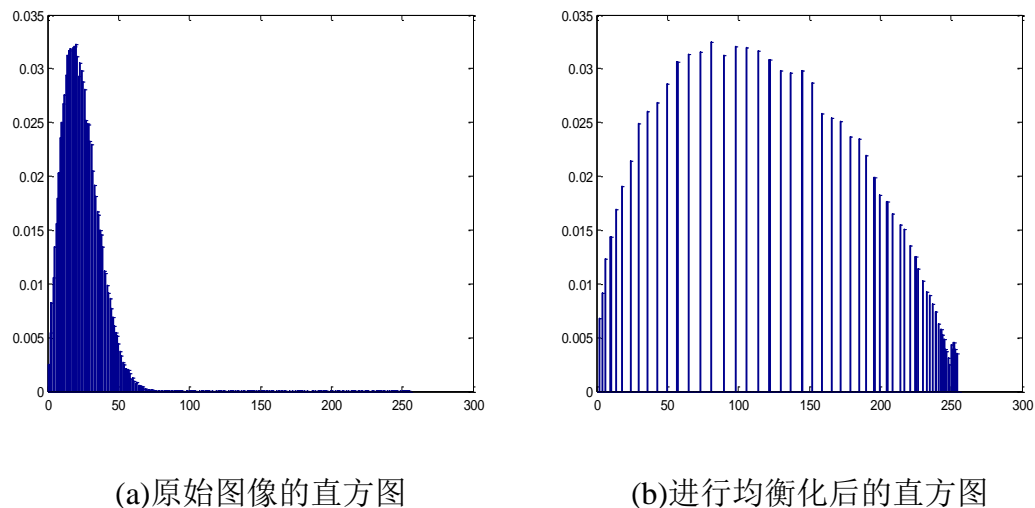


图 2.8 直方图均衡前后对比

2.3.2 图像的去噪

图像去噪是指去除在实际成像过程中由声纳设备和环境造成的失真现象，只提取图像中的有用信息。一般情况下，图像的去噪可以分为：全局去噪，也就是对含有噪声图像的整体或其中较大的块进行去噪，但是全局处理用时较长，不能应用于实时性要求高的处理中；局部去噪，即当需要对某一像素进行去噪时，只对其局部较小邻域内的像素

点进行处理，相比全局去噪，具有计算效率高的优势。

常用的图像去噪方法有：低通滤波、均值滤波、中值滤波。对几种滤波方法进行比较发现：这些去噪方法都能满足对图像的后期处理要求，在实际处理过程中，我们可根据任务要求的不同，来进行去噪方法的选择。当环境中孤立点噪声较大时可优先考虑中值滤波法，而当对实时性要求较高的时候可考虑低通滤波法。

在一幅图像中，噪声通常都是较为尖锐的高频部分，相对而言，背景区域和其中灰度变化较慢的像素区域是图像中的低频部分。由此，我们可以通过利用一个合理的低通滤波器过滤掉其中的高频噪声部分，保存低频部分，但是低通滤波在去噪的过程中会造成较明显的细节模糊。通常我们在时域中将低通卷积的模板与图像进行卷积便可得到去噪后的图像^[35]，而在频域中，可通过下图得到：

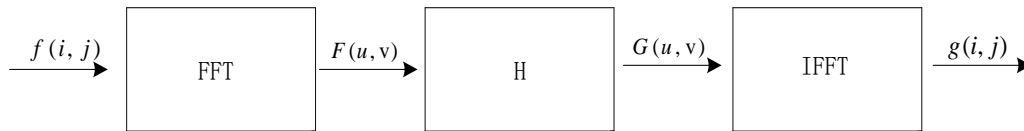


图 2.9 频域低通滤波原理

图 2.9 中 $f(i, j)$ 表示的是带噪声的图像， $F(u, v)$ 是其傅里叶变换，通过传递函数 H 后， $F(u, v)$ 中的高频部分被衰减，得到平滑后图像的傅里叶变换 $G(u, v)$ 。理想低通滤波器的传递函数需要符合：

$$H(u, v) = \begin{cases} 1 & D(u, v) \leq D_0 \\ 0 & D(u, v) > D_0 \end{cases} \quad (2-10)$$

的条件。上式(2-10)中， D_0 为一个非负数， $D(u, v)$ 表示点 (u, v) 到频域上平面原点的距离，可以表示为：

$$D(u, v) = \sqrt{u^2 + v^2} \quad (2-11)$$

均值滤波的原理是将图像中某一点的像素值用该点的某一个邻域的平均值进行替代。对于一幅给定的图像 $f(i, j)$ ， (x, y) 代表图像上的每一个像素点，其像素值用滤波窗口 G 中 N 个像素值的平均值进行替换，如下式：

$$\bar{f}(x, y) = \frac{1}{N} \sum_{(i, j) \in G} f(i, j) \quad (2-12)$$

中值滤波的原理是将图像中某一点的像素值用该点的某一个邻域的中值替代^[32]。中值就是指：一组给定的数值 $x_1, x_2, x_3, \dots, x_n$ ，将这 n 个数按照数值的大小排序如下：

$$x_{j1} \leq x_{j2} \leq x_{j3} \leq \dots \leq x_{jn} \quad (2-13)$$

$$y = Med\{x_1, x_2, x_3, \dots, x_n\} = \begin{cases} x_{j\frac{n+1}{2}} & n \text{ 为奇数} \\ \frac{1}{2}(x_{j(\frac{n}{2})} + x_{j(\frac{n}{2}+1)}) & n \text{ 为偶数} \end{cases} \quad (2-14)$$

上式中的 y 就是数组 $x_1, x_2, x_3, \dots, x_n$ 的中值。我们将中心像素点的具有特定形状和尺寸的邻域称为窗口。这就是一维中值滤波器的理论，如此类推至二维信号处理中则产生了二维中值滤波器，此时滤波器的窗口也拓展至二维。对于一幅给定的图像 $f(i, j)$ ， (x, y) 代表每个像素点，其像素值用滤波窗口 G 中 N 个像素值的中值进行替换，如下式：

$$\bar{f}(x, y) = med_{(i, j) \in G} \{f(i, j)\} \quad (2-15)$$

对于孤立的噪声点，中值滤波器在保证较完整的图像像素细节的同时能最好的消除孤立噪声点。

滤波器窗口的形状和尺寸对中值滤波器和均值滤波器的滤波效果影响较大，常用与实际中的二维滤波器的窗口形状有圆形、十字形、方形等。如图 2.10 所示：

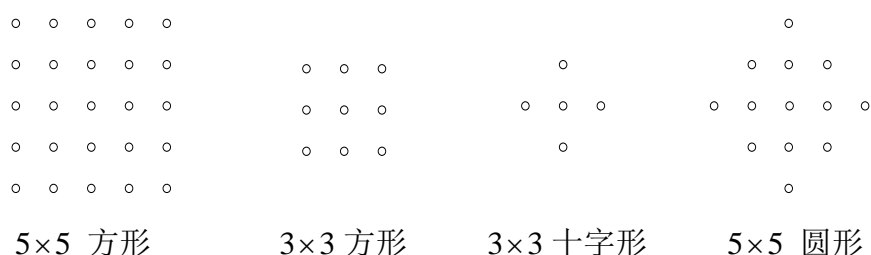


图 2.10 几种滤波窗

图像的噪声抑制效果可通过图像质量的客观评价指标来进行评定，也就是通过计算处理后图像与原始图像的偏离误差实现客观评价。常用的方法有均方误差(MSE)与峰值信噪比(PSNR)，经过处理后质量较好的图像一般 PSNR 值也较高，但是也可能与人眼看到的图像质量不一样，所以 MSE 与 PSNR 仅仅为图像质量的参考值^[33]。

均方误差(MSE)可表示为：

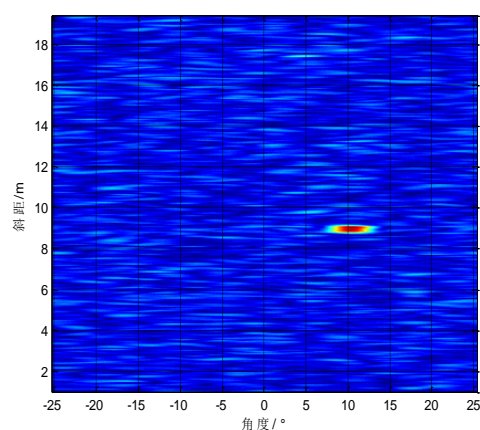
$$MSE = \frac{\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (f(i, j) - g(i, j))^2}{m \times n} \quad (2-16)$$

上式中， $f(i, j)$ 表示原始图像， m 与 n 代表图像的宽和高， $g(i, j)$ 则为处理后的图像。

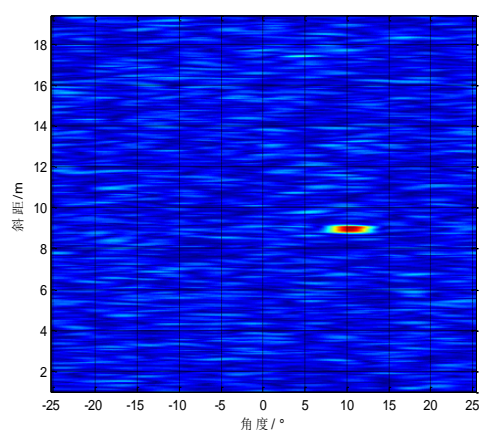
峰值信噪比(PSNR)与均方误差(MSE)在本质上是相同的，可表示为：

$$PSNR = 10 \times \lg \frac{255 \times 255}{MSE} \quad (2-17)$$

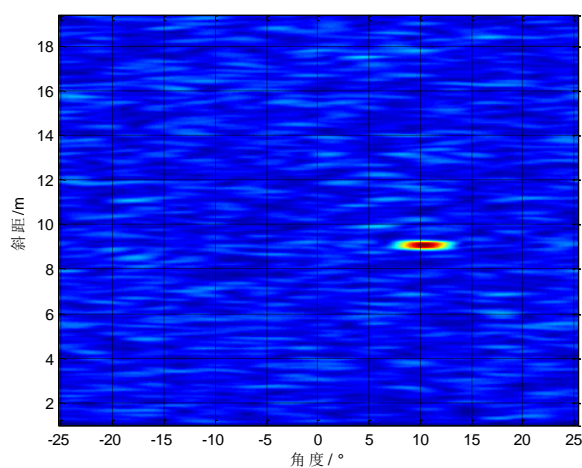
对上一节中仿真出的原始图像进行去噪处理，得到以下结果：



(a)中值滤波后的图像



(b)均值滤波后的图像



(c)低通滤波后的图像

图 2.11 图像去噪结果对比

表 2.1 三种去噪方法的峰值信噪比(PSNR)值

三种去噪方法	峰值信噪比(PSNR)
中值滤波	50.2348
均值滤波	47.2137
低通滤波	38.8826

从图 2.11 的去噪对比可以看出,对于人眼观察,低通滤波的方法对本图的去噪效果较好,通过观察原始图像与经过均值滤波和中值滤波的图像可以知道后两种滤波器也实现了滤波的功能,虽然出现了与人眼观察图像质量不同的效果,但从表 2.1 中三种去噪方法的 PSNR 值可以知道中值滤波的值最高,而低通滤波的像素模糊度最大。

2.4 本章小结

本章首先对声成像原理进行了研究，然后通过仿真实验对声成像算法进行分析与验证，为后续的声纳图像预处理提供了基础。接着又探讨了声纳图像的预处理方法，包括增强和去噪两个部分，分别对不同的增强和去噪方法进行了对比，最后在理论的基础上进行了仿真实验，通过实验得到了图像增强和去噪的处理结果。这些预处理方法可以为后续的目标检测提供良好的基础。

第3章 基于图像声纳的二维检测技术

3.1 引言

经过预处理的图像有较明显的目标信息，声纳图像目标检测的重点就是进行图像的分割运算，得到图像中的目标区域。由于海洋混响不同于其他噪声，所以对声纳图像分割难度也很大。本章首先对基于单帧图像的分割理论进行了研究，包括固定阈值分割法和天津阈值(Ostu)分割法，然后详述了 CFAR 的概念，对 CFAR 检测法在单帧声纳图像目标检测中的可行性进行了详尽的分析。

3.2 图像分割

声纳图像的目标分割算法是声纳图像分析的关键，在对图像的研究中，我们并非对图像中的所有部分都有兴趣。这时，我们就需要将感兴趣的部分（目标或前景）从图像中提取出来，将图像分割为目标区域与背景区域。在不同的领域，图像分割也有其他的名称，如：阈值化技术、目标检测技术、目标轮廓技术等，声纳图像在进行目标检测时所要克服的最大难题就是声纳图像像素的模糊与噪声干扰^[36]。

目前图像分割方法多种多样，究其本质就是对图像中不同特性的像素进行分类。应用较为广泛的图像分割方法有：根据图像的灰度值直接进行分割的方法，即阈值法和根据图像灰度分布的不同而进行分割的区域分割法。阈值法就是通过将某一个或者几个固定的灰度值设为阈值，将其他像素值与其进行判比，把图像分为不同的部分，在同一个部分的像素值被看为同一个物体。阈值法的分割效果与阈值的选取有极大的关联，因此，阈值法的关键技术就是有效阈值的选取。如果阈值选的太高，就会造成目标的丢失；而阈值选的太低，那么虚假目标的泛滥。下面主要比较了固定阈值分割算法和天津阈值分割算法。

3.2.1 固定阈值分割

固定阈值分割算法的思想是：依据某一准则在图像 $f(i, j)$ 中找到某一个灰度值作为分割的阈值 T ，使得分割之后的图像为：

$$g(i, j) = \begin{cases} b_0 & f(i, j) < T \\ b_1 & f(i, j) \geq T \end{cases} \quad (3-1)$$

一般情况下，我们将与阈值比较后的像素值中大于或等于阈值 T 的赋成 1，即 $b_1 = 1$ ，而小于阈值 T 的赋成 0，即 $b_0 = 0$ 。固定阈值分割法是一种全局单阈值分割方法，图像分割效果完全取决于阈值额选取，常见的固定阈值分割法有：直方图双峰法和二分法。直方图双峰法需要事先得到图像的灰度直方图，如果直方图呈双峰分布，那么就选取双

峰之间的峰谷像素值作为图像的固定分割阈值。实际上，不同的图像可能都对对应着同一幅灰度直方图，所以该方法对图像的先验信息有较大的依赖，当图像中只有单峰、双峰差别较大或者两峰之间的谷在直方图上分布较宽时，该方法无法实现目标分割^[37]。二分法进行固定阈值的分割计算的思想就是：固定阈值 T 选取图像中最大像素值与最小像素值的均值，即：

$$T = \frac{1}{2} [\max(\max(f(i, j))) + \min(\min(f(i, j)))] \quad (3-2)$$

对水池数据进行分析，水池数据的基本参数：线阵为 16 元线阵，阵元间距 7cm，LFM 信号 20-30kHz，脉宽 2ms，带宽 10kHz，采样频率 200kHz，目标由远及近以 0.1m/s 进行匀速运动，具体位置如下：

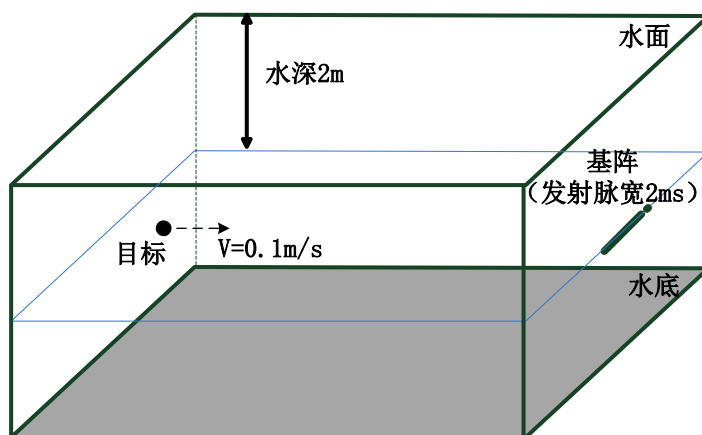
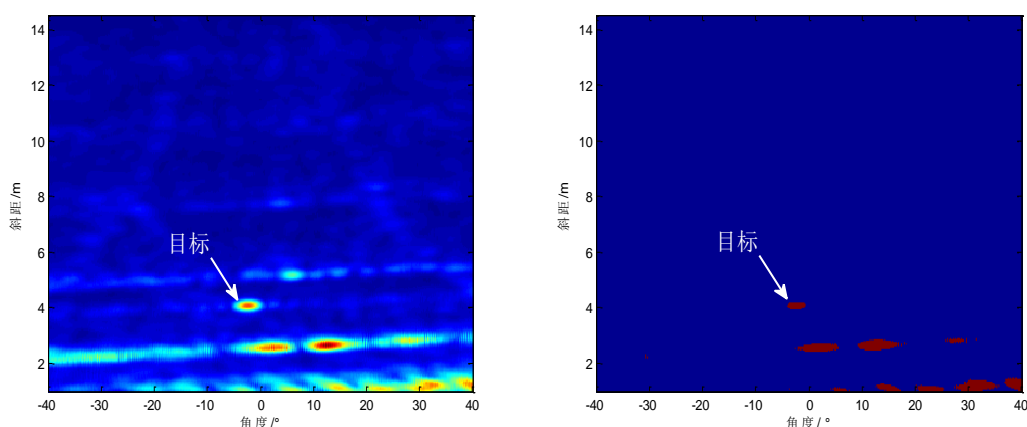


图 3.1 水池实验图示

利用二分法进行分割，由于声纳图像进行过归一化处理，此时阈值为像素级的一半，结果如下图：



(a)近处目标原始图像

(b)固定阈值法分割后图像

图 3.2 固定阈值分割法对近处目标(SNR=13-15dB)的分割结果

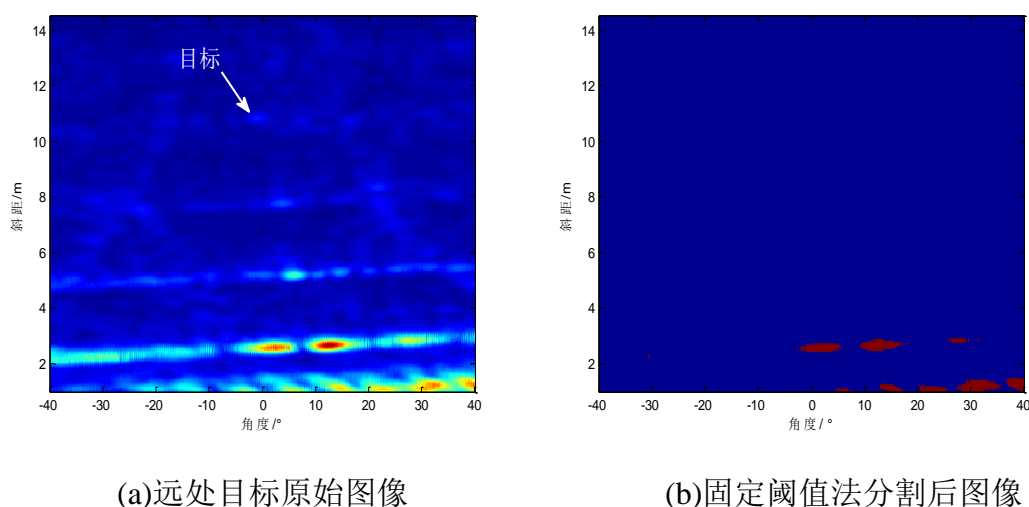


图 3.3 固定阈值分割法对远处目标(SNR=3-5dB)的分割结果

观察图 3.2 和图 3.3 可见，利用二分法分别对水池中目标位于近处和远处的两幅图像进行分割。目标位于近处时，其中心区域像素值几乎接近最大像素级，明显高于用二分法确定的阈值，所以能被成功的检测出来，如图 3.2 中(b)。而当目标位于远处时，目标信噪比较低，几乎淹没在背景像素区域中，远远低于二分法所确定的阈值，未能被检测出，如图 3.3。

3.2.2 大津阈值分割

大津阈值分割法又叫类间类内最大方差比法（Ostu 法）的思想是：图像的目标与背景具有一定的可分性灰度分布，即图像的像素值能被区分为目标与背景。从统计的意义上来说，方差是一种统计量，它用于表征数据分布的不平衡。类间方差大就说明不同的类具有较大的差异，不同的部分能被区分开来；类内方差小，即同一类间有较好的一致性。因此，选择类间类内方差比最大化就能保证所选取的分割阈值能在最大限度保护同类目标的情况下最好的分割不同的类。其具体的步骤是：

假设原始图像有 L 个灰度级，通过阈值 T 将图像分成灰度为 $1 \sim T$ 和灰度为 $(T+1) \sim L$ 的两个类，分别命名为 Δ_0 和 Δ_1 。假定 Δ_0 中的像素个数为 N_0 ， Δ_1 中的像素个数为 N_1 ，像素的总数为 $N = N_0 + N_1$ ，两类的出现概率和均值表示为：

$$p_0 = \frac{N_0}{N} = p(T) \quad (3-3)$$

$$p_1 = \frac{N_1}{N} = 1 - p(T) \quad (3-4)$$

$$\mu_0 = \frac{1}{N_0} \sum_{f(i,j) \in \Delta_0} f(i,j), \mu_1 = \frac{1}{N_1} \sum_{f(i,j) \in \Delta_1} f(i,j) \quad (3-5)$$

可由上式得到这两类的方差，表示为：

$$\sigma_0^2 = \frac{1}{N_0} \sum_{f(i,j) \in \Delta_0} (f(i,j) - \mu_0)^2 \quad (3-6)$$

$$\sigma_1^2 = \frac{1}{N_1} \sum_{f(i,j) \in \Delta_1} (f(i,j) - \mu_1)^2 \quad (3-7)$$

整个图像的均值表示为：

$$\mu = \frac{1}{N} \sum_{f(i,j) \in \Delta_0 \cup \Delta_1} f(i,j) \quad (3-8)$$

先假定两个类间方差为 σ_{out} ，类内方差为 σ_{in} ，利用上式表示其为：

$$\sigma_{out}^2 = p_0(\mu_0 - \mu)^2 + p_1(\mu_1 - \mu)^2 \quad (3-9)$$

$$\sigma_{in}^2 = p_0\sigma_0^2 + p_1\sigma_1^2 \quad (3-10)$$

那么依据类间类内最大方差比准则，此时的最佳阈值可通过下式得到：

$$\eta(T) = \max\left(\frac{\sigma_{out}^2}{\sigma_{in}^2}\right) \quad (3-11)$$

用大津阈值分割法对与上一节图 3.2(a)和图 3.3(a)中的原始图像进行处理分析，结果如下：

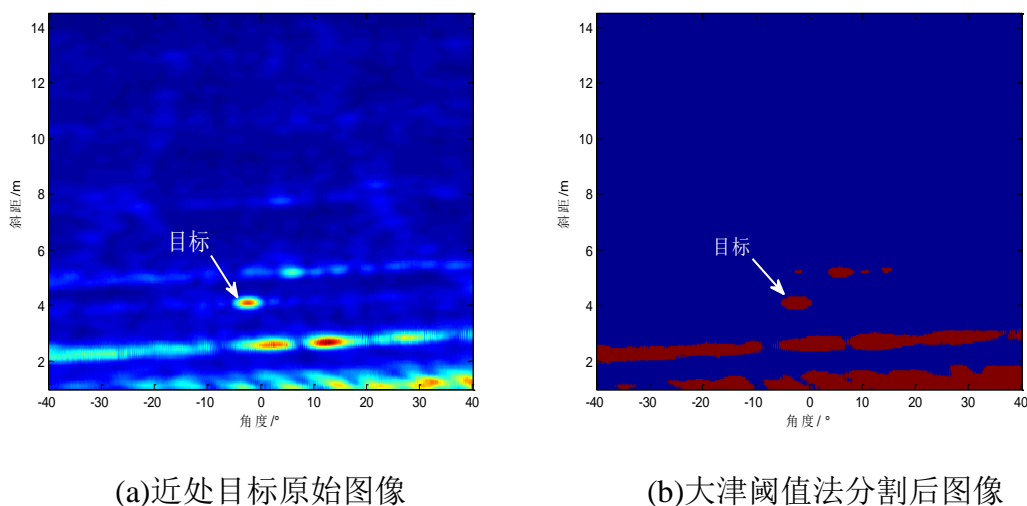


图 3.4 大津阈值分割法对近处目标(SNR=13-15dB)的分割结果

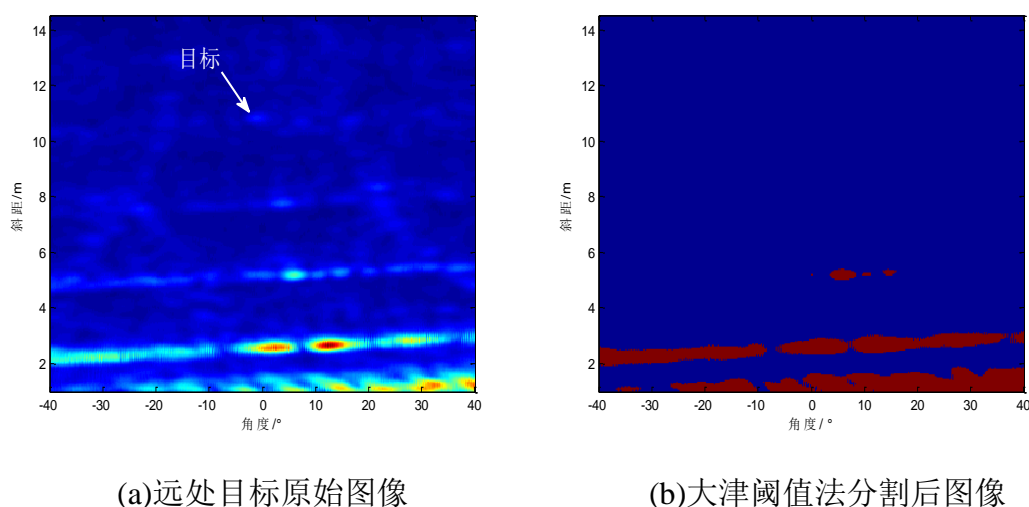


图 3.5 大津阈值分割法对远处目标(SNR=3-5dB)的分割结果

观察发现大津阈值分割法能将图 3.4 中近处目标分割出来，而对图 3.5 中远处目标不具有分割能力。根据大津阈值分割法的分割思想可以知道，当目标位于近处时，其信噪比较高，目标的像素值与背景像素值的类间方差较大，因此能被检测出来。而目标位于远处时，目标信噪比低，其像素值与背景像素值相差较小，目标与背景类间方差也很小，无法通过最大类间类内方差法确定的阈值，也就无法被分割出来。

由于传统的图像分割方法对于水下声纳图像的分割能力有限，仅能分割出高信噪比的图像目标。但我们又知道声纳图像具有分辨率低、干扰严重的现象，显然上述的分割方法在声纳图像的目标分割应用上存在不足。因此，寻找对声纳图像分割效果有保障的方法是我们的研究方向。

3.3 恒虚警(CFAR)检测法

当声波在水中传播的时候，海洋的不均匀结构、海底海面的反射、海洋中的生物及非生物（如气泡、海底泥沙等）都会导致声波的散射，所有的散射波信号在接收水听器叠加便形成了混响。根据海洋中散射体的不同，它们对声波信号的散射也不同，可以将它们形成的混响分为三种：

1. 海面混响：由分布在海面上散射面积中的散射体（如由海面波浪形成的气泡层）产生的混响；
2. 体积混响：由海水自身或者分布于海洋中的具有一定体积的散射体（如海洋生物、海水中的冷热水团及海水中的沙砾等）产生的混响；
3. 海底混响：由分布在海底散射面积上的散射体（如起伏不平的海底地形等）产生的混响。

由于目标回波与混响有较强的相关性，在主动声纳的目标检测中，作为随机信号的海洋混响是其主要干扰，当出现目标的时候，接收器能接收到含有目标信号和混响的回

波；而没有目标的时候，接收器只能接收到混响的回波^[38]。在目标出现的时候，如果信混比较大，那么能更好的判断回波中是否存在目标；若信混比较小的话，对目标存在性的判断就相对困难一些。因此对于混响的研究在声纳出现以后未曾停止过，主动声纳系统最重要的工作之一就是检测目标是否存在。

海水中混响的平均强度并不是平稳的随机过程，它会随着时间不断衰减，此时用固定阈值进行目标检测效果较差。根据混响统计理论可知，混响是由海洋中点散射体对声波信号的反向散射迭加形成的，当一个单元格内的散射体数目较多时，参考中心极限定理，此时匹配滤波器输出的同相和正交分量服从的是高斯分布，且其包络服从的是瑞利分布^[38]。然而，现代的高分辨率声纳系统在对信混比、方位分辨率和距离分辨率进行改善的同时，也使得某一单元格内有效散射体的个数减少，中心极限定理不在适用^[38]。此时，匹配滤波器的输出包络也无法满足瑞利分布，我们称此时的包络分布为非瑞利分布或者是广义瑞利分布。广义瑞利分布的概率密度函数有较严重的拖尾现象，此时再用以前的检测器进行检测可能会造成目标检测过程中虚警概率的增大，对目标检测带来困难。因此，为了达到更好更有效地检测目标，我们需要研究在恒定虚警概率情况下的目标检测。

目前，在假设匹配滤波器输出的包络各个分量为统计独立同分布的前提下，对混响描述常用的分布模型包括：瑞利分布、瑞利混合分布、K 分布、Weibull 分布、Log-normal 分布等^[39]。本文中的实验数据是在水池中获得的，由于水池单帧图像的背景分布为瑞利分布，且该分布模型方便、易于处理，因此下面主要介绍基于瑞利分布的检测。

3.3.1 瑞利分布模型

当一个单元格内的散射体数目过多时，那么参考中心极限定理，此时匹配滤波器输出的同相和正交分量服从的是高斯分布，且其包络服从的是瑞利分布^[38,40]。

设 X 为随机变量， x 为任意实数，瑞利分布随机变量 X 的概率密度函数(PDF, probability density function)如下式：

$$f_x(x) = \begin{cases} \frac{x}{\lambda} e^{-x^2/2\lambda} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (3-12)$$

分布函数(CDF, cumulative distribution function)如下式：

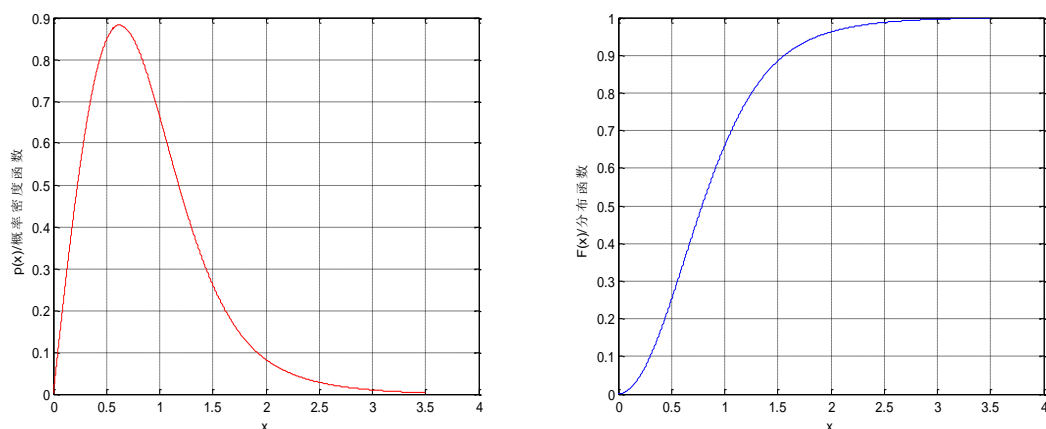
$$F_x(x) = 1 - e^{-\frac{x^2}{2\lambda}} \quad x \geq 0 \quad (3-13)$$

随机变量 X 的数学期望与方差分别为：

$$E(X) = \sqrt{\pi\lambda} / 2 \quad (3-14)$$

$$\text{Var}(X) = \lambda \quad (3-15)$$

瑞利分布的概率密度函数与分布函数由图 3.6 所示：



(a)瑞利分布的概率密度函数图

(b)瑞利分布的分布函数图

图 3.6 瑞利分布

可以通过矩估计法和极大似然估计法进行瑞利分布的参数估计。假设一组样本 X_1, X_2, \dots, X_N ，若它的 k （ k 是正整数）阶原点矩存在，那么根据大数定律有，对一个任意小的 $\varepsilon > 0$ ，当该样本容量 $N \rightarrow \infty$ 时，有下式：

$$\lim_{N \rightarrow \infty} p\left\{\left|\frac{1}{N} \sum_{i=1}^N X_i^k - E(X^k)\right| < \varepsilon\right\} = 1 \quad (3-16)$$

也就是说样本的 k 阶原点矩依据概率收敛于总体的 k 阶原点矩。这样我们可以根据矩估计的定义，算出 λ 的矩估计，如下式：

$$\hat{\lambda} = \frac{1}{N} \sum_{i=1}^N X_i^k \quad (3-17)$$

极大似然法的估计法主要是通过对 $L(x_1, x_2, \dots, x_N; \lambda) = \prod_{i=1}^N f_x(x_i, \lambda)$ 的自然对数求偏

导进行解算，即令 $\frac{\partial \ln L}{\partial \lambda} = 0$ ，得到的 $\hat{\lambda}$ 与上式一样。

3.3.2 CFAR 概述

利用 CFAR 进行处理是现代自动检测雷达系统中不可或缺的组成部分。确定检测准则是 CFAR 检测的基本问题，该方法在事先给出的系统参数和最大检测距离的情况下，利用某种最优判决准则进行目标的自动检测，并且能在保持虚警概率恒定的情况下，使目标检测概率最大。在实际的雷达或者是声纳系统中，无法事先精确地确定噪声与干扰

的统计特性, 因此很难事先得到固定的检测门限。CFAR 处理法是一种自适应的门限检测算法, 根据接收器接收到的回波数据进行干扰的统计特性估计, 在此基础上算出检测门限, 最后与得到的检测单元杂波样本进行比较, 实现目标检测过程中的恒虚警性。雷达系统中的杂波是指雷达波入射到目标表面后产生的反向散射, 具有较强的动态特性。利用声呐系统中的混响与雷达系统中的杂波统计特点类似的特征, 可以发现, CFAR 法也适用于水下目标的检测。

Nitzberg 把雷达的自动检测方法依据检测阈值的获取分为: (1)固定的阈值; (2)由外部干扰的平均幅度获取的阈值; (3)根据外部干扰统计分布的部分先验知识获取的阈值; (4)当外部干扰统计分布的先验信息未知时, 通过自由分布的统计假设检验获取的阈值。除了上述的第一种方法其它三种就是自适应阈值检测^[41,42]。

目前仍未形成统一的 CFAR 检测算法分类标准, 通常可根据以下的几种标准来进行划分: 按杂波分布模型划分为: 瑞利(Rayleigh)分布、K 分布、韦布尔(Weibull)分布、对数正态(Log-normal)分布模型下的 CFAR 检测算法^[43]; 通过数据形式划分为: 标量与矢量的 CFAR 检测算法^[44]; 依数据的处理域划分为: 时域与频域的 CFAR 检测算法; 按信号的相关程度划分为: 相关信号、部分相关信号与不相关信号的 CFAR 检测算法; 根据数据的处理方式划分为: 参量与非参量的 CFAR 检测算法^[39,41]。除了上述的 CFAR 检测算法外, 还有单参数与多参数分布式、单传感器与多传感器分布式的 CFAR 检测算法。

1. 二元假设检测

如果接收到的回波信号包络服从独立同分布的瑞利分布, 那么它的同相分量 I 与正交分量 Q 通过平方检波之后, 可以得到一个背景功率服从参数为 λ 的指数分布的杂波序列。如此, 对在滑动参考窗中的某个检测单元 X , 可根据单元 X 分布的二元假设检验进行是否含有目标的判决。对于假设 H_0 (无目标), X 中只有功率为 λ 的背景杂波; 而对于假设 H_1 (有目标), X 中不仅包含背景杂波, 还含有信噪比为 S 的目标回波。表达式如下:

$$H_0 : X \sim \exp(\lambda) \Rightarrow \text{无目标} \quad (3-18)$$

$$H_1 : X \sim \exp(\lambda(1+S)) \Rightarrow \text{有目标} \quad (3-19)$$

检测概率 P_d 是指存在目标并检测出目标的概率, 虚警概率 P_{fa} 是指不存在目标时检测出目标的概率。可由式(3-20)和式(3-21)得到检测概率 P_d 和虚警概率 P_{fa} 的表达式:

$$P_d = P[X > X_0 | H_1] = \int_{X_0}^{\infty} p(x | H_1) dx \quad (3-20)$$

$$P_{fa} = P[X > X_0 | H_0] = \int_{X_0}^{\infty} p(x | H_0) dx \quad (3-21)$$

上式中 X_0 表示检测门限, $p(x | H_1)$ 和 $p(x | H_0)$ 分别表示检测单元 X 在 H_1 和 H_0 的

假设下的概率密度函数，如下式：

$$p(x|H_1) = \frac{1}{\lambda(1+S)} \exp\left(\frac{-x}{\lambda(1+S)}\right) \quad (3-22)$$

$$p(x|H_0) = \frac{1}{\lambda} \exp\left(\frac{-x}{\lambda}\right) \quad (3-23)$$

为了达到在上述的假设下自动检测的目的，需要设计出相应的检测准则，实现一定意义上的最佳检测。

2. 奈曼皮尔逊最佳检测

在检验器的设计过程中，通过检测概率 P_d 和虚警概率 P_{fa} 对 CFAR 的性能进行评价，我们希望能够达到 P_d 尽可能大，而 P_{fa} 尽可能小，显然，这二者是互相矛盾的。奈曼皮尔逊准则在保持虚警概率 P_{fa} 在预定的虚警概率水平上，使检测概率 P_d 最大。当确定杂波和目标回波信号的分布后，可以通过极大似然检验法来表示奈曼皮尔逊最佳检测， $l(x)$ 代表对数似然比，则有：

$$l(x) = \ln\left(\frac{p(x|H_1)}{p(x|H_0)}\right) = \frac{x}{\lambda} \left(\frac{S}{1+S}\right) - \ln(1+S) \quad (3-24)$$

那么，判决准则可表示如下：

$$\begin{cases} l(x) < l_0 \Rightarrow H_0 : \text{无目标} \\ l(x) \geq l_0 \Rightarrow H_1 : \text{有目标} \end{cases} \quad (3-25)$$

若假设：

$$\mu_0 = \lambda \left(\frac{1+S}{S}\right) (l_0 + \ln(1+S)) \quad (3-26)$$

那么，判决准则(3-25)可写为：

$$\begin{cases} x < \mu_0 \Rightarrow H_0 : \text{无目标} \\ x \geq \mu_0 \Rightarrow H_1 : \text{有目标} \end{cases} \quad (3-27)$$

通过上面的式(3-27)可以得到奈曼皮尔逊准则所对应的检测概率 P_d 与虚警概率 P_{fa} ，表达式如下：

$$P_d = P[X > \mu_0 | H_1] = \exp\left(-\frac{\mu_0}{\lambda(1+S)}\right) \quad (3-28)$$

$$P_{fa} = P[X > \mu_0 | H_0] = \exp\left(-\frac{\mu_0}{\lambda}\right) \quad (3-29)$$

由式(3-28)和(3-29)可以看出，当信噪比 $S = 0$ 时，检测概率和虚警概率满足 $P_d = P_{fa}$ 。

当已知杂波背景功率 λ 时,固定阈值检测法容易实现,是此时的最佳的检测法,它的检测阈值可通过虚警概率 P_{fa} 和杂波背景功率 λ 确定,表达式如下:

$$\mu_0 = -\lambda \ln P_{fa} \quad (3-30)$$

奈曼皮尔逊准则具有固定的阈值和恒定的虚警概率,而实际的检测过程中,随着检测环境的变化,固定阈值无法满足检测需求,只能作为 CFAR 检测过程中的参考,因此要寻找更好的检测方法。

3.3.3 CFAR 检测模型

CFAR 检测方法最重要的是保持虚警概率恒定,为了达到此目的,需要实时估计接收器接收到的信号的概率分布情况,在此基础上形成自适应的判决门限。

一般情况下,对雷达或是声纳信号的检测都是将其中某个坐标单元规定为当前检测单元,并对该单元进行处理。根据接收到的回波信号的相关特性,检测单元的背景估计可通过估计相邻位置的杂波得到。也就是说,在当前检测单元相邻位置设置一个滑窗(即参考窗),令参考窗中的杂波样本为背景估计的参考单元。图 3.7 就是典型的一维和二维 CFAR 参考滑窗图。

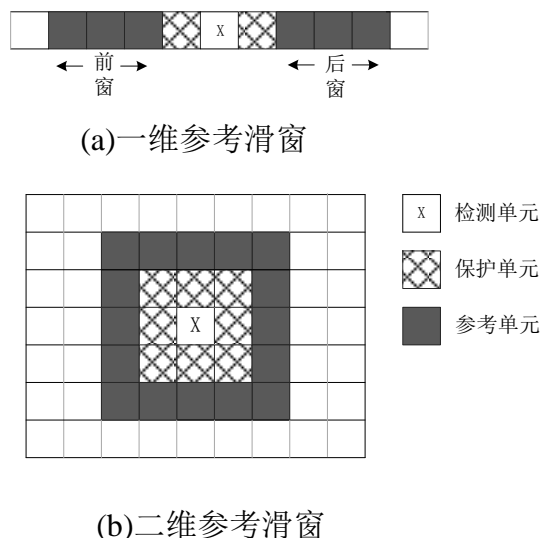


图 3.7 CFAR 目标检测参考滑窗模型

在瑞利包络杂波环境下主要采用均值类 CFAR 检测器,其共同特点是都利用参考窗的均值进行局部背景杂波的估计,均值类检测器的检测方法包括单元平均(CA)法、最小(SO: small of)法、最大(GO: great of)法、和加权平均(WCA: weighted cell-averaging)法。

下面的图 3.8 为均值类 CFAR 检测器方框图,其中的正交分量 Q 与同相分量 I 通过平方律检波之后形成杂波序列,依次进入一个长度为 $2N+1$ 的移位寄存器中。分别利用寄存器中的 N 个后窗和 N 个前窗杂波样本组成 CFAR 检测器的参考窗,令参考窗内的中心样本 $D=d_0$ 为当前检测单元。那么就可以通过参考窗内其他 $2N$ 个杂波样本实现参

考窗中的背景杂波功率估算,在此基础上确定检测门限来判决当前检测单元内目标的有无,若设 $D_0 = TZ$ 为检测门限的话,则自适应的判决准则可由下式表示:

$$\begin{cases} d_0 < TZ \Rightarrow H_0 : \text{无目标} \\ d_0 \geq TZ \Rightarrow H_1 : \text{有目标} \end{cases} \quad (3-31)$$

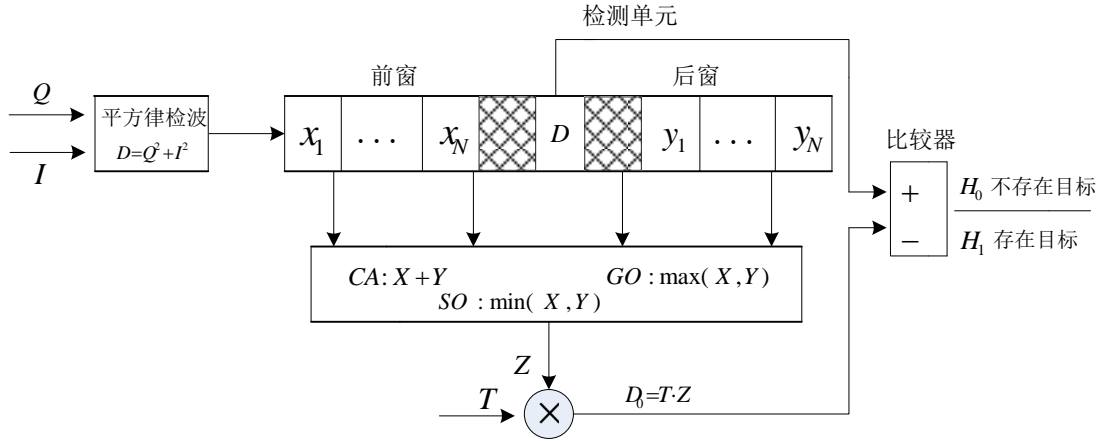


图 3.8 均值类 CFAR 检测器方框图

式(3-31)中的 Z 表示背景杂波功率估计, T 表示标称系数,可由参考窗的长度 N 和虚警概率 P_{fa} 得到。服从高斯分布的回波(包络服从瑞利分布)经过平方率检波之后,各参考单元的采样服从指数分布, Z 的概率密度函数为 $f_z(Z)$,则虚警概率的计算式如下:

$$P_{fa} = E_Z [P[D > TZ | H_0]] = \int_0^\infty f_z(Z) \int_{TZ}^\infty \frac{1}{\lambda} \exp\left(-\frac{x}{\lambda}\right) dx dz = M_z(u) \Big|_{u=T/\lambda} \quad (3-32)$$

$M_z(u)$ 为 Z 的矩母函数(moment generating function, MGF),多个随机变量和的矩母函数为各个随机变量矩母函数的乘积。在 $u = T / (\lambda(1+S))$ 的时候,上式可转化为均匀杂波背景情况下的检测概率表达式:

$$P_d = M_z(u) \Big|_{u=\frac{T}{\lambda(1+S)}} \quad (3-33)$$

1. CA—CFAR

对于 CA—CFAR 检测器,是将参考窗内的 $2N$ 个杂波样本的平均功率当作背景杂波功率的估计值 Z 。为方便计算,可把因子 $1/2N$ 归入标称系数中。对于参考窗的前窗和后窗中的杂波样本 $x_i, i=1, \dots, N$ 和 $y_i, i=1, \dots, N$, CA-CFAR 的背景杂波功率的估计值 Z 可如下表示:

$$Z = \sum_{i=1}^N x_i + \sum_{i=1}^N y_i \quad (3-34)$$

在此引入 Γ 简化推导,那么其概率密度函数(PDF)的表达式为:

$$f(x) = \beta^{-\alpha} x^{\alpha-1} e^{-x/\beta} / \Gamma(\alpha), x \geq 0, \beta \geq 0, \alpha \geq 0 \quad (3-35)$$

记服从 Γ 分布的随机变量为 $X \sim G(\alpha, \beta)$ ，此时 X 的矩母函数 $M_X(u)$ 为：

$$M_X(u) = (1 + \beta u)^{-\alpha} \quad (3-36)$$

由于IID假设可知道， $x_i \sim G(1, \lambda)$ 和 $y_i \sim G(1, \lambda)$ 。根据上面提及到的多个随机变量和的矩母函数为各个随机变量矩母函数的乘积的性质，可以通过式(3-36)得到 $Z \sim G(2N, \lambda)$ 。再将式(3-36)带进式(3-33)中，就可以得到CA-CFAR检测器的检测概率

$$P_d = \left(1 + \frac{T}{1+S}\right)^{-2N} \quad (3-37)$$

当式(3-37)中 $S=0$ 的时候，可以得到标称系数 T 和虚警概率之间的关系，如下式：

$$P_{fa} = P_d|_{S=0} = (1+T)^{-2N} \quad (3-38)$$

从式(3-37)和式(3-38)的表达式可以看出，CA-CFAR的检测概率 P_d 和虚警概率 P_{fa} 不依赖背景杂波功率 λ ，以此表明CA-CFAR确实为恒虚警的检测。

2. SO—CFAR 和 GO—CFAR

在运用CA—CFAR对杂波边缘进行检测时，虚警概率会上升。为了改善检测器在杂波边缘的性能，对检测器进行改进，得到SO—CFAR和GO—CFAR检测器。

而当参考窗中发生空间中多个相邻的目标相互干扰，从而导致CA—CFAR检测器性能下降的情况时，可选取SO—CFAR进行检测。此时选择参考窗前窗和后窗的局部最小估计作最后的背景杂波功率估计，背景杂波功率的估计值 Z 可如下表示：

$$Z = \min \left(\sum_{i=1}^N x_i, \sum_{i=1}^N y_i \right) \quad (3-39)$$

对于改善杂波边缘检测性能的需求，主要采用GO—CFAR进行检测，该方法主要是选择参考窗前窗和后窗的局部最大估计作最后的背景杂波功率估计，背景杂波功率的估计值 Z 可如下表示：

$$Z = \max \left(\sum_{i=1}^N x_i, \sum_{i=1}^N y_i \right) \quad (3-40)$$

3.3.4 单帧声纳图像的CFAR检测

上面详细研究了CFAR检测的理论，在图像域中对背景服从瑞利分布的目标运用单元平均的CFAR法进行检测，可以在恒虚警概率的条件下确定出每一个检测单元的阈值。当混响背景的包络服从瑞利分布时，可根据(3-38)式计算出这里虚警概率 P_{fa} 和单元检测概率之间的关系如式(3-41)，对于任意一个检测单元 i ，其像素值 x_i 服从一个瑞利分布，此时该检测单元的阈值可通过下式得到：

$$T_i = \sqrt{c \cdot \hat{\lambda}_i} \quad (3-41)$$

上式中:

$$\hat{\lambda}_i = \frac{1}{2M} \sum_{i \in G_i} (x_i)^2, c = 2M \left(\left(\frac{1}{P_{fa}} \right)^{1/M} - 1 \right) \quad (3-42)$$

其中 $\hat{\lambda}_i$ 是单元 i 背景功率的 MLE (极大似然估计), T_i 是单元 i 的检测阈值, G_i 是围绕在检测单元 i 的划窗, M 为该划窗包含的单元总数。

选择围绕在检测单元周围的背景值来进行局部背景估计, 此外还要依靠保护单元将用于局部背景估计的单元与检测单元隔开, 避免目标信号对背景估计的影响。具体情况如图 3.9 所示:

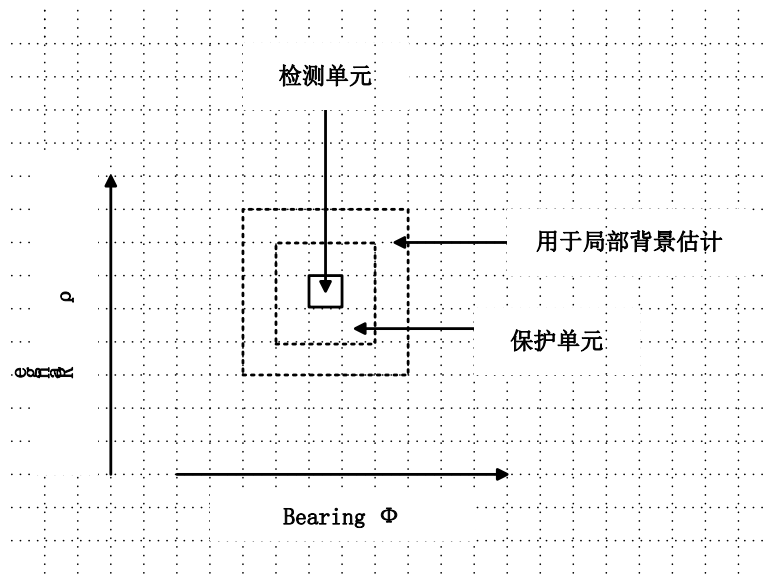


图 3.9 基于瑞利分布背景的 CA 检测模型

对于 CFAR 检测法, 当背景服从瑞利分布时, 目标的检测概率满足下式:

$$P_d = \left(1 + (P_{fa}^{-1/M} - 1) \frac{\lambda}{\lambda + d} \right)^{-M} \quad (3-43)$$

其中 d 为目标的强度, 可以推出:

$$\frac{\lambda}{\lambda + d} = \frac{1}{1 + SNR} \quad (3-44)$$

$$P_d = \left(1 + (P_{fa}^{-1/M} - 1) \frac{1}{1 + SNR} \right)^{-M} \quad (3-45)$$

图 3.10 根据上面的计算式给出了 CFAR 法虚警概率和检测概率的关系:

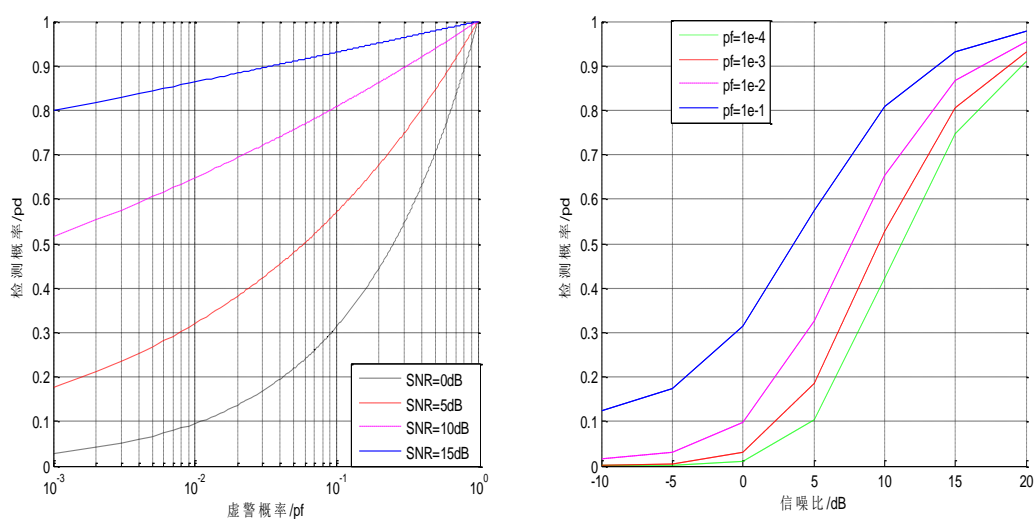


图 3.10 CFAR 法的 ROC 曲线

根据前面章节介绍的图像分割检测算法，首先对水池背景数据进行建模，结果如图 3.11:

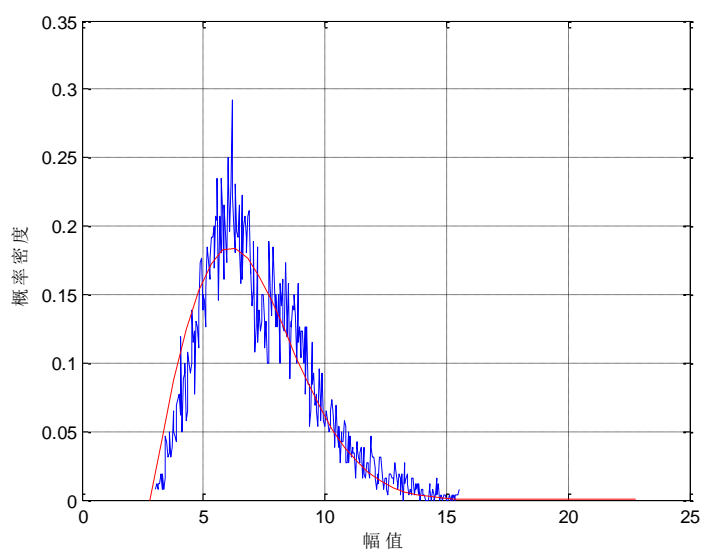


图 3.11 背景建模

根据上图可知，水池背景服从瑞利分布，因此本文所用的基于背景服从瑞利分布 CFAR 检测法可以进行水池数据的处理，其结果如下：

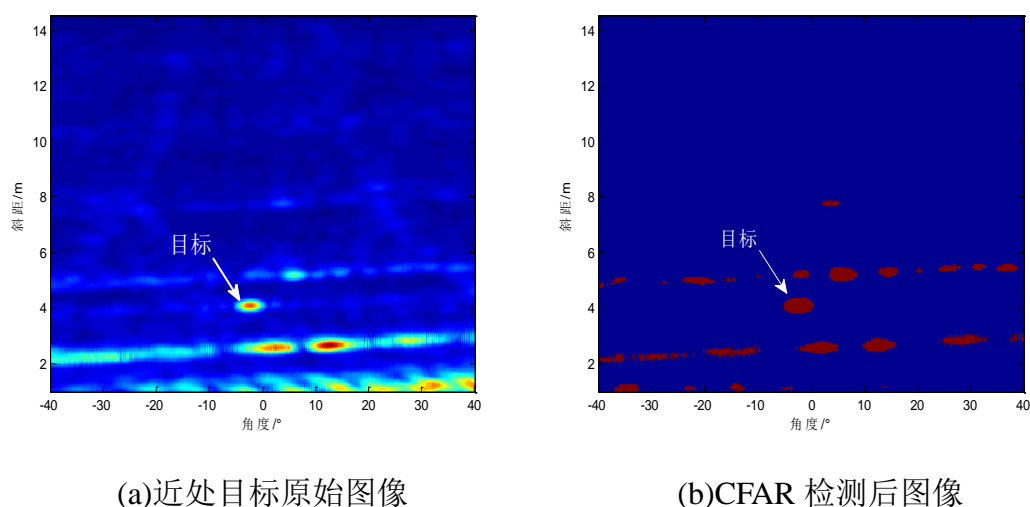


图 3.12 CFAR 法对近处目标(SNR=13-15dB)的检测结果

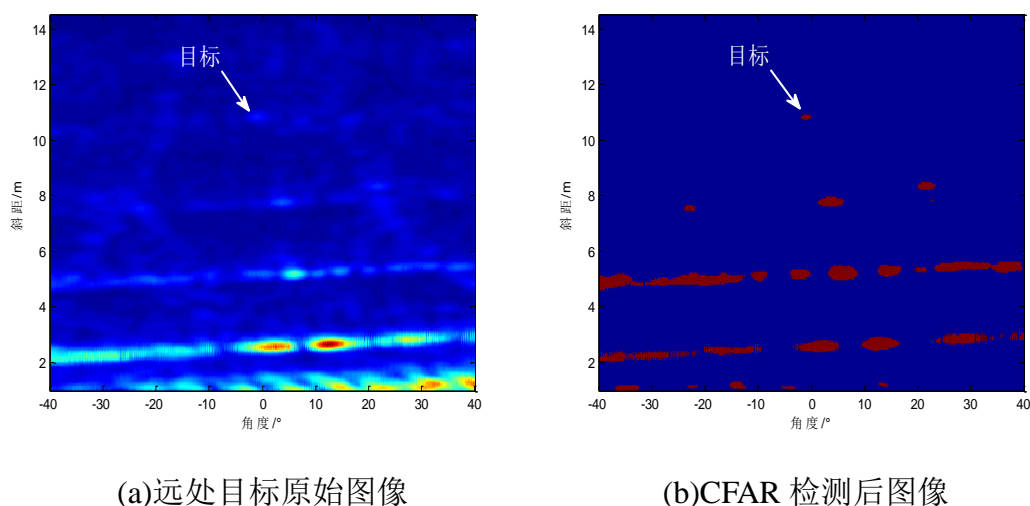


图 3.13 CFAR 法对远处目标(SNR=3-5dB)的检测结果

观察图 3.12 和图 3.13 可以发现，CFAR 检测法不仅能检测出近处目标，还能在目标位于远处、信噪比较低的情况下，将目标检测出来。检测效果优于固定阈值检测法和 Ostu 阈值检测法，这是由于 CFAR 可通过局部背景估计确定出的自适应门限，这样一来弱信噪比的信号也不容易漏检。通过检测结果图的比较，我们发现了 CFAR 检测法在声纳图像检测中的可行性，但是也发现 CFAR 检测法无法实现动静目标的分离，对于 0-5m 处的水面水底混响，CFAR 法也将其检测出来，使结果中包含较多的目标（包括静目标），因此下一章中需要抑制背景带来的静态干扰，提高 CFAR 检测法的性能。

3.4 本章小结

本章首先研究了两种图像分割算法（固定阈值分割法和天津阈值分割法），然后对 CFAR 检测法的原理和方法进行了详细的研究。由于本文 CFAR 检测是在背景服从瑞利

分布模型的基础上展开的,因此本章中对瑞利分布模型与其参数估计方法做了深入的探讨,同时还对二元假设检验和奈曼皮尔逊最佳检测准则进行了介绍。在 CFAR 检测的理论基础上,对瑞利背景下的声纳图像进行 CFAR 检测,并将其检测结果与该图像的固定阈值分割结果和大津阈值分割结果进行比对,确定了 CFAR 法在基于声纳图像的目标检测中的适用性。最后通过仿真得到 ROC 曲线。但 CFAR 检测后图像中的目标点过多,无法分辨出静态的干扰,还有待改善。

第4章 基于图像声纳的动目标检测技术

4.1 引言

计算机视觉处理中信息提取重要的一步就是目标的检测，目标的检测效果直接影响后续的目标跟踪结果，在提取目标特征后，利用适当的算法，将运动目标从图像序列中分离出来就是运动目标的检测^[45]。

声纳图像由于成像的环境、机理及信息传输媒介的特殊性，具有存在声影区、噪声严重、对比度差的缺陷，这也就造成了声纳图像中的目标检测难度较大^[3]。所以，如何克服声纳图像的劣势，提高声纳图像中目标检测的准确度就显得十分重要，水下动目标的检测成为声纳探测系统研究中的重点。

目前，动目标的分析方法最常用的有：基于帧间差分的算法、基于背景估计的算法、基于光流场的算法。在检测运动目标的过程中，将场景中的背景分为静态背景和动态背景^[46]。当检测过程中只有目标在运动，而探测设备保持静止，那么这种目标检测的背景是静态的；相反，探测设备随着目标运动也发生移动（平移、旋转），那么此时目标检测的背景就是动态的。上一章中介绍的对单帧图像使用 CFAR 检测法虽然能检测出目标，但是静态干扰较为严重，为此本章中利用多帧图像建立混合高斯背景模型，在此基础上进行 CFAR 检测，改善其检测效果。

4.2 帧间差分法

差分法即两帧图像进行减运算，获得含有运动目标的图像。差分法要求两帧图像的背景是静态的，且目标在两帧图像的时间内保持相对匀速的运动。基于帧间的差分法分为连续帧的差分法和多帧差分法等。连续帧的差分法是利用连续相邻的两帧图像之间的差运算得到运动目标的信息，它可以被看作是所有其他帧间差分法的基础。相邻帧间的强相关性是连续帧的差分法成功检测目标的重要保证，该算法流程如下图：

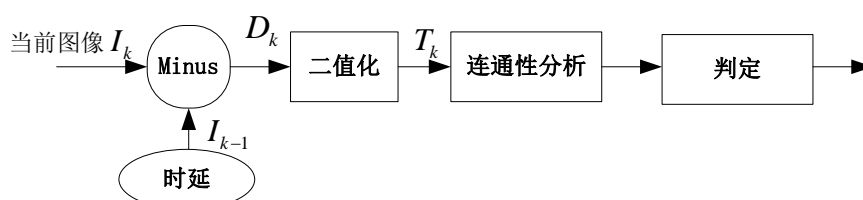


图 4.1 基于连续帧的差分法流程

连续帧的差分法可以定义为：

$$D_k(x, y) = |I_k(x, y) - I_{k-1}(x, y)| \quad (4-1)$$

图中 $I_k(x, y)$ 表示第 k 帧图像的 (x, y) 点的像素值，第 k 帧图像表示当前帧图像。

$I_{k-1}(x, y)$ 表示第 $k-1$ 帧图像的 (x, y) 点的像素值，第 $k-1$ 帧图像表示参考帧的图像。通过连续帧的差分法获得差分图像之后，将每一点像素与预先选定的阈值进行判定，大于阈值的像素值判定其属于目标区域，小于阈值的像素值则归于背景区域。其数学表达式定义为：

$$\begin{cases} T_k(x, y) = 1, & D_k(x, y) \geq T \\ T_k(x, y) = 0, & D_k(x, y) < T \end{cases} \quad (4-2)$$

连续帧的差分法具有运算简单、检测速度快的优点，能够提取出我们感兴趣的运动目标。但是该算法还是存在一些缺陷，比如：当目标运动速度较小时，连续帧的差分可能导致两帧图像目标的重叠区域被减掉，从而形成空洞，这种情况即使连通性分析也无法解决，那么就可能会丢失目标；而当目标运动速度较大时，连续帧的差分会造成差图像中检测出的目标面积远大与其真实面积，甚至出现两个目标。

为了弥补连续帧的差分法的上述不足，出现了不同的多帧差分法，如隔帧差分法和对称帧的差分法。隔帧差分法就是用当前第 k 帧图像和第 $k-i(i > 1)$ 帧图像做差分运算，这种方法可以解决目标运动速度较小时差分结果产生的“空洞”问题；对称帧的差分法是将连续的三帧图像进行相邻帧的差分运算，再对得到的两幅差分图像做与运算，最后得到当前帧的目标区域^[47]。但对称帧的差分法无法用自适应的方法确定差分阈值。

通过仿真利用帧间差分法对动目标进行检测，结果如下图：

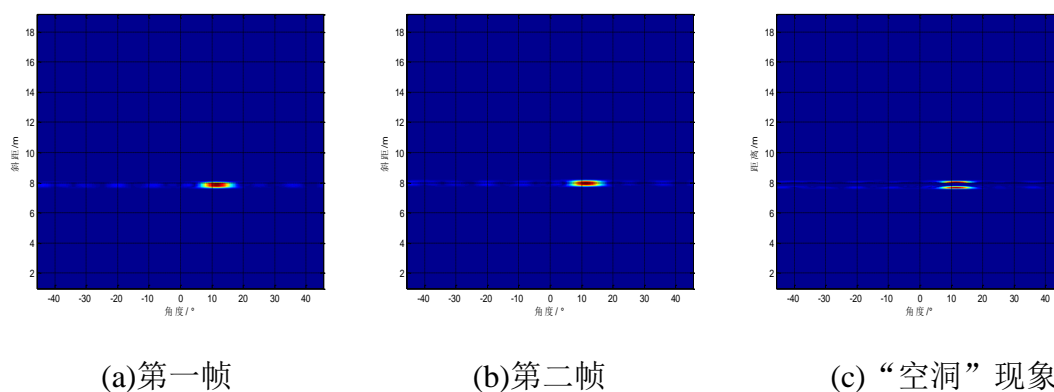


图 4.2 帧间差分法的“空洞”现象

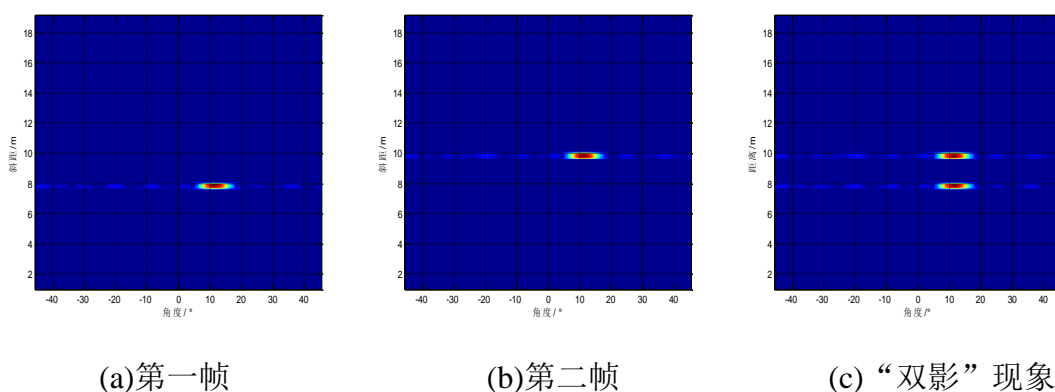


图 4.3 帧间差分法的“双影”现象

根据仿真图 4.2 可以看出，当目标的运动速度较慢时，前后相邻帧图像中目标区域有较严重的重叠，利用帧间差分法进行检测时，这些重叠区域无法被检测出来，导致检测到的结果出现“空洞”现象；而图 4.3 中目标的运动速度又过快，前后帧图像中的目标间距较远，再利用帧间差分法进行检测会出现“双影”现象。鉴于帧间差分法存在这些缺陷，本文不再采用其进行水池数据分析。

4.3 背景估计法

基于背景估计的算法是常见的一种动目标检测算法，它是用当前帧图像减去背景图像，得到信息量相对较小的差分图像。然后将差分后图像的所有像素点与预先取定的阈值做遍历运算，判定其中大于阈值的像素点属于目标区域，而小于阈值的像素点属于背景区域，这样实现运动目标的检测。基于背景估计的算法容易实现，且提取出的目标信息较为完整，具有广泛的应用空间。基于背景估计的算法的检测流程如下图：

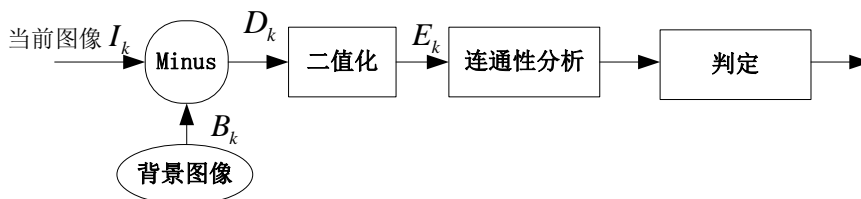


图 4.4 基于背景估计的目标检测流程

$B_k(x, y)$ 表示第 k 帧背景图像的 (x, y) 点的像素值，基于背景估计的算法首先将当前帧 I_k 和背景图像 B_k 相减，得到其差分图像 D_k ，将差分图像用预先选好的阈值进行二值化判决，最后从二值化的图像中获得目标信息。

$$D_k(x, y) = |I_k(x, y) - B_k(x, y)| \quad (4-3)$$

再对其进行二值化处理，得到二值化图像 $E_k(x, y)$ ：

$$E_k(x, y) = \begin{cases} 0 & |D_k(x, y)| \leq T \\ 1 & |D_k(x, y)| > T \end{cases} \quad (4-4)$$

基于背景估计的算法可被看作特殊的帧间差分法，它只是将相邻帧图像用算法估计出来的背景图像代替进行图像的差分运算，不过它却能较好的解决帧间差分法对目标区域较大时，目标内出现空洞的问题。基于背景估计的动目标检测算法需要保证背景是静态的，对背景图像的要求较高，声纳图像中目标的运动或者是水波的动荡都会造成背景的变化。因此我们需要根据不同的实际情况通过不同的方法建立背景模型，常用的有统计背景建模法、统计平均法和内核密度估计法。下面主要介绍的是第二种统计平均法，而统计背景建模法将在第五节进行详细的研究。

统计平均法主要用于背景较为简单的情况，首先假一组序列背景图像 $I_k(x, y)$ $k=1, 2, 3 \dots N$ ，则背景图像表示为：

$$B = I_k(x, y) \quad (4-5)$$

为了使背景变化对目标检测的影响降低，得到更适用的背景图像，进行下式计算：

$$B_k = \frac{1}{N} (I_k + I_{k-1} + I_{k-2} + \dots + I_{k-N+1}) \quad (4-6)$$

上述方法用一段序列内的图像平均灰度值构成背景图像，在适当选取 N 值后，我们能得到较真实的背景估计图像，但该方法获得的背景图像对于目标运动速度较快、目标在图像中出现时间较短的情况检测效果较好；而若目标的在图像内运动速度慢、出现时间长，那这个方法就会出现误差。

4.4 统计背景建模法

从第三节中我们知道，基于背景估计的目标检测算法是常用的动目标检测算法，它在动目标检测领域具有重要的地位，但是如果想要提高该算法性能，更好的检测出动目标，我们首先需要做好的就是背景模型的建立及其更新工作。

在理想的情况下，我们希望得到的背景图像就是场景中无运动目标时获得的图像。不过，在陆上，光照的时间变化和空气流动带来的扰动（树叶晃动等）；在水下，水流的波动、鱼群的游动等都会造成背景的变化，因此理想状态下毫无干扰的纯背景图像在实际生活中是不可能得到的。其次，对于长时间的目标监视，背景图像也会随之变化，因此背景图像还需要定期进行更新，用以保证得较好的检测效果^[48]。

传统的背景估计法是直接将探测设备观测到的没有动目标的图像作为背景图像，它只适用于外界条件稳定性较好的场景。现在的科学研究早已开始转向自适应的背景估计方法，比如在第三节中介绍的统计平均法的背景估计方法等。但是统计平均法还未能建立起背景完整的统计模型，所以其背景建模及目标检测效果也不理想。下面介绍的高斯

背景建模法就能较好的解决场景变化引起的背景变化，具有较好的目标检测能力。

背景建模法的思想是：建立背景图像中所有像素点的分布模型，同时对建立出的背景模型进行更新，当运动目标进入场景后，其像素分布不能满足背景图像的像素分布模型，根据此原则我们就可以将目标检测出来。一般，我们将背景模型中每个像素点的像素值分布较为集中的情况称为单模态，而背景模型中每个像素点的像素值较为分散的情况则被称为多模态。通常用高斯分布来描述在场景中背景的像素分布的概率密度函数，接下来就分别从单模态和多模态的情况介绍高斯背景建模法。

4.4.1 单高斯背景模型

顾名思义，单高斯背景模型就是只用一个高斯分布来描述场景中背景图像像素分布。此时，背景的像素值集中于一个峰值附近，是典型的单模态背景模型，如图 4.5。在 t 时刻，该高斯分布的定义为：

$$\eta(x_t | \mu_t, \Sigma_t) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma_t|^{\frac{1}{2}}} \cdot \exp\left(-\frac{(x_t - \mu_t)^T}{2\Sigma_t}(x_t - \mu_t)\right) \quad (4-7)$$

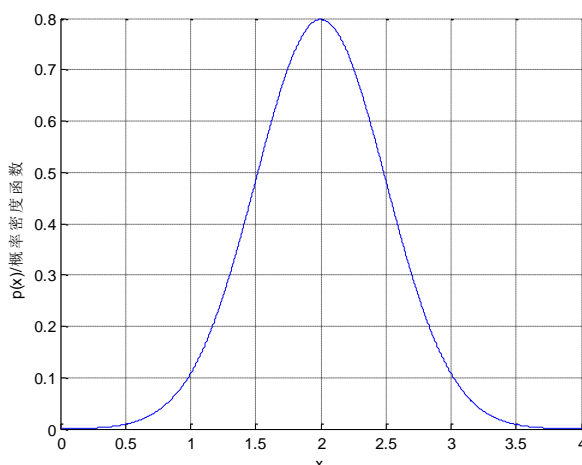


图 4.5 单高斯分布曲线

其中的 n 表示高斯分布中 x_t 的维数， Σ_t 表示一个 $n*n$ 维的协方差矩阵， μ_t 表示高斯分布的均值， x_t 表示图像的像素值。背景图像建成后，对当前帧图像中每一点的像素值与单高斯背景模型进行遍历匹配。若满足给定的匹配原则，则认为该点是背景点，反之则是目标点。匹配原则一般如下所示：

$$x_t \in [\mu_t - k\sigma_t, \mu_t + k\sigma_t] \quad (4-8)$$

上式的 k 是常数，可根据对要求的背景判别严苛程度确定。一般情况下，当 $k=1$ 时，将有 70% 的样本将落入高斯分布的置信区间内；选取 $k=2, k=3$ 的时候，落入样本置信区间的样本百分比分别约有 95% 以及 99%。通常我们取 $k=3$ 进行判断，也就是 3sigma

法。

经过一段时间得到的多帧的序列图像一般会发生变化,这种变化我们的人眼可能观察不到,但是随着阳光照射的不同、水波的动荡产生的细微变化在计算机进行处理时都很容易被发现。因此,为了保证良好的检测效果,我们还需要对背景图像进行更新,具体到单高斯模型背景的话就是对其中的函数参数进行更新。下面将参数的更新速度用 α 表示,对式(4-7)的均值 μ_t 和协方差 Σ_t 进行更新,其更新公式如下:

$$\mu_{t+1} = (1-\alpha)\mu_t + \alpha x_t \quad (4-9)$$

$$\Sigma_{t+1} = (1-\alpha)\Sigma_t + \alpha(x_t - \mu_t)^T(x_t - \mu_t) \quad (4-10)$$

上式中, μ_t 表示当前背景像素均值,同时也是高斯分布的均值, μ_{t+1} 表示更新后的背景像素值, Σ_{t+1} 表示更新后的协方差矩阵。不同更新速度 α 值对目标检测结果有不同的影响,当 α 的值太小,这时背景模型的更新比不上实际场景中背景的变化,造成背景模型的“滞后”,用这种情况下得到的模型进行目标检测效果必然不理想;而当 α 的值太大,此时背景的更新速度过快,可能导致场景中运动速度较慢的目标被误判为背景,再进行目标检测的话,检测可能会产生空洞或者拖尾问题,更有甚者造成漏报。

通常,单高斯背景模型对背景无变化或者变化缓慢微小的简单场景有相对稳定的适应力。而当实际场景中出现随风摆动的树枝、飘动的旗帜、监视器的屏幕和水面荡漾的波纹等情况时,背景的起伏变化很大,或造成背景像素值的分布出现多峰分布的情况。此时背景的像素值分布不再是单模态的,用单高斯背景建模的方法已经无法满足目标检测的需要,而是要根据背景像素值的多模态特点选择混合高斯背景建模法进行检测。

4.4.2 混合高斯背景模型

建立背景模型是为了能够有效地将场景中我们感兴趣的动目标提取出来,并将那些不感兴趣的动目标和静态背景归为场景的背景部分,因此,所建立背景模型的好坏直接影响动目标检测效果的优劣。我们已经知道,对于多变的复杂场景建模不能再用基于单模态的背景建模法。混合高斯模型能有效地描述复杂场景中的背景变化,它是利用多个单高斯模型来描述复杂的多模态背景,如图 4.6,若其由 M 个单高斯分布混合而成,其定义为:

$$\eta(x_t | \mu_t, \Sigma_t) = \sum_{k=1}^M \pi_{k,t} \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma_{k,t}|^{\frac{1}{2}}} \cdot \exp\left(-\frac{(x_t - \mu_{k,t})^T}{2\Sigma_{k,t}}(x_t - \mu_{k,t})\right) \quad (4-11)$$

上式中 $\pi_{k,t}$ 表示在 t 时刻观测到的像素值由第 k 个高斯分布产生的概率,且

$$\sum_{k=1}^M \pi_{k,t} = 1。$$

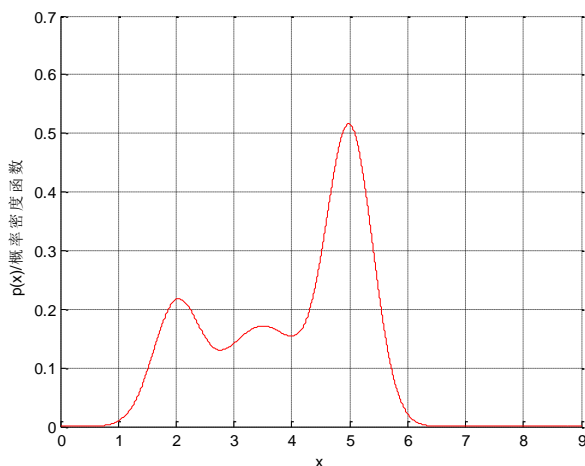


图 4.6 混合高斯分布曲线

混合高斯建模法的基本思想是：图像中各个像素点的值都可用 M 个高斯分布模型的混合进行建模，再通过对各个高斯模型录入图像数据的历史进行归一化的加权，加权主要是为了表示单个模型构成背景像素的可能性。通常， M 的取值为范围为 3-5，具体取决于硬件环境和对速度的要求， M 取值越大，模型能够更准确地表达图像，但所需的计算时间也就越长； M 取值越小，模型就越为模糊，不过计算速度变快。

Stauffer 等人利用回归滤波原理提出一种在线最大期望算法，并利用该算法进行模型参数的更新，各高斯分布分别对应不同的权值和优先级，用一个小常量 α 建立背景模型并进行更新，这种方法虽然保障了背景模型的稳定性，但是也降低了高斯分布的收敛速度。Cheng 通过与 EM 算法结合，在最大后验概率准则下推导出了高斯分布模型的参数迭代公式，这一算法分割效果良好、有较强的适应性，但是选取最优参数还有一定的困难。KaewTraKulPong 等人通过选取较大的学习率因子提高了背景模型的学习效率，缺点是学习率因子缺乏自适应性。之后，Lee 又在学习率因子上做出了改进，通过与时间建立的联系，实现了学习算法的自适应性，同时也提高了高斯分布的收敛速度。

下面将对混合高斯分布的参数估计与模型建立进行具体介绍。

1. EM 算法原理

EM (Expectation Maximization, 即期望最大化) 算法来源于极大似然估计原理，是 Dempster、Rubin 和 Laird 提出的一种能在测量数据不完整的情况下进行迭代的算法^[49]。EM 算法有两个核心：第一就是求解期望值，是指在已估计出当前时刻的参数的条件下，根据极大似然估计法求出完整数据的似然函数的对数期望值；第二就是实现最大化，是指利用已求得的使期望函数最大的参数值进行新参数值的计算。EM 算法通过上

述两步的连续迭代循环使得到的参数与预计的参数误差最小。

测量数据不完整主要由两种情况产生：第一种是在数据的采集过程中，由于探测设备受限或是测量错误等原因造成的数据丢失情况；第二种则是在数据处理的过程中，由于直接对参数的似然函数进行处理难以实现，而引入一些新的分布特征参数来简化似然函数，当定义加了新分布特征参数的数据为完整数据时，原始的数据也就相对变得不完整。一般情况下，EM 算法的收敛性比较可靠，因此该算法有较广泛的应用，本文中混合高斯背景建模法中的参数估计就是利用 EM 算法实现的。

我们将新加入的分布特征参数称为“隐含数据”定义为 Y ，而直接的到的原始数据为 X ，那么完整数据可定义为 $Z = \{X, Y\}$ 。对于完整的样本集合，它的样本都可表示为 $Z = \{Z_1, Z_2, \dots, Z_N\}$ ，并且当中的任何一个样本 Z_i 都可以用 $Z_i = \{X_i, Y_i\}$ 来表示，这也就表明每一个样本数据都是原始数据和隐含数据组成的。

图像数据通常被看作是离散的，对于一个完整样本数据 Z ，样本测量值是 $\{z_1, z_2, \dots, z_N\}$ ，那么发生事件 $\{Z_1 = z_1, Z_2 = z_2, \dots, Z_N = z_N\}$ 的概率计算如下式：

$$\prod_{i=1}^N p(Z_i = z_i) = \prod_{i=1}^N p(z_i; x_1, x_2, \dots, y_1, y_2, \dots) \quad (4-12)$$

式(4-12)表示的是分布中的未知参数 $\{x_1, x_2, \dots, y_1, y_2, \dots\}$ 的似然函数，可以写成 $L\{z_1, z_2, \dots, z_N; x_1, x_2, \dots, y_1, y_2, \dots\}$ ，可以通过下式求得极大似然估计值：

$$L\{z_1, z_2, \dots, z_N; \hat{x}_1, \hat{x}_2, \dots, \hat{y}_1, \hat{y}_2, \dots\} = \max_{x_1, x_2, \dots, y_1, y_2, \dots} L\{z_1, z_2, \dots, z_N; x_1, x_2, \dots, y_1, y_2, \dots\} \quad (4-13)$$

这里的 $\{\hat{x}_1, \hat{x}_2, \dots, \hat{y}_1, \hat{y}_2, \dots\}$ 就是通过极大似然函数求得的极大似然估计值。对只含原始数据 X 的似然函数运用式(4-12)和式(4-13)进行如下计算：

$$L(\Phi | X) = p(X | \Phi) = \prod_{i=1}^N p(x_i | \Phi) \quad (4-14)$$

上式中的 Φ 表示需要估计的分布参数，因为数据并不完整，所以上式的优化处理比较困难，将“隐含数据” Y 补上之后，似然函数表示为：

$$L(\Phi | Z) = L(\Phi | X, Y) = p(X, Y | \Phi) \quad (4-15)$$

如果将 Φ 和 X 看作常量，则只有“隐含数据” Y 是自变量，此时该似然函数依然是随机变量，其最大化还是不易求得，需要寻找确定的函数来转换。我们发现它的期望满足确定性的条件，优化相对容易，其定义如下：

$$Q(\Phi; \Phi') = E[\ln p(X, Y | \Phi) | X, \Phi'] \quad (4-16)$$

式(4-16)中 $Q(\Phi; \Phi')$ 是一个关于 Φ 的函数，其中的 Φ' 是参数 Φ 现有的估计， Φ' 与 X 取常量。(4-16)式中右边可以求出缺失信息的对数似然函数期望值，上式假设 Φ' 表示整

个分布的真实参数，可以通过原始观测数据 X 和估计参数 Φ' 得到“隐含数据” Y 的边缘概率密度函数 $f(Y/X, \Phi')$ 。

式(4-16)是 EM 算法中最为核心的公式， Φ' 为对当前分布的最优估计参数，直接用于计算期望，在当前估计的基础上，再作进一步改良优化得到参数 Φ ，根据极大似然估计法的思想可知，求解极大似然函数就能得到事件发生的最大概率。因此，式(4-16)可以增大似然函数 $Q(\Phi; \Phi')$ 的值，似然函数的值越大得到的结果越接近真实值。不断地迭代求解期望值和实现最大化，就能不断的更新参数 Φ' ，提高 $Q(\Phi; \Phi')$ 的值，当其满足事先假定一个收敛判据 Th 时，停止迭代，使得似然函数值逼近一个局部的最大值，这时得到的 $\hat{\Phi}$ 就是最终选取的参数估计值。EM 迭代算法的步骤如下：

- (1) 参数的初始化： Φ^0 和 Th ；
- (2) 求解期望值：计算 $Q(\Phi; \Phi')$ 的值；
- (3) 实现最大化： $\Phi^{t+1} \leftarrow \arg \max Q(\Phi; \Phi')$ ；
- (4) 根据收敛判据，进行迭代计算： $Q(\Phi^{t+1}; \Phi^t) - Q(\Phi^t; \Phi^{t-1}) \leq Th$ ；
- (5) 返回 $\hat{\Phi} \leftarrow \Phi^{t+1}$ ，得到最大估计值；

通过上面的计算，我们就能确定最终的参数值，而参数值直接决定混合高斯背景模型的建立，因此下面就将介绍混合高斯背景模型的参数估计。

2. 混合高斯背景模型的参数估计

混合高斯背景模型是典型的多模态背景模型，它主要是用 M 个独立同分布的高斯模型来描述某一个像素点的像素值变化。在进行运动目标检测时，因为运动目标出现时间短，且只占场景中很小的区域，而其他大部分的区域都是背景区域。因此，动目标前景只需要用少数的状态来描述，而其他多数状态用于表示背景。在检测中，我们通过探测到的原始数据估计出当前时刻哪一状态出现的概率最大，然后根据该状态所属的类别，来判决它属于动目标前景还是背景。

假设场景中的某个像素点的像素值由 M 个独立同分布的高斯函数组成，且单个点像素值的时间序列分布是离散的，将其设为一个随机过程 $\{x_1, x_2, \dots, x_t\} = \{I(x_0, y_0, i) : 1 \leq i \leq t\}$ ， $I(x_0, y_0, i)$ 表示的是 (x_0, y_0) 点在第 i 帧图像中的像素值，其 t 时刻的采样值 x_t 的概率密度可由下式表示：

$$f(X | \Phi) = \sum_{k=1}^M \pi_k f(X | k, \theta_k) \quad (4-17)$$

$$f(X | k, \theta_k) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma_k|^{\frac{1}{2}}} \cdot \exp\left(-\frac{(X - \mu_k)^T}{2\Sigma_k(X - \mu_k)}\right) \quad (4-18)$$

式(4-17)和(4-18)中的 $f(X | k, \theta_k)$ 表示在 t 时刻该像素的观测值 X 是由第 k 个高斯

分布产生，且该高斯分布对应的参数为 θ_k 时，随机变量 X 的概率密度函数。式中

$\pi_k = p(k)$, $k = 1, 2, \dots, M$ ，并且满足 $\sum_{k=1}^M \pi_k = 1$ ，表示在 t 时刻观测到的像素值由第 k 个

高斯分布产生的概率。式(4-17)混合高斯分布参数集是 $\Phi = (\pi_1, \pi_2, \dots, \pi_k, \theta_1, \theta_2, \dots, \theta_k)$,

$\theta_k = (\mu_k, \Sigma_k)$ 表示第 k 个高斯分布的均值与协方差。为了便于计算，根据像素值的时间序列也是独立同分布的，可将协方差矩阵简化为对角阵。

采用上面介绍的 EM 算法，利用随机变量 X 的观测值可以估计出混合高斯分布的参数 Φ 。假设在进行参数估计时，将产生第 k 个高斯分布的过程与产生像素值的过程都看作平稳的随机过程，它们所有的统计特性保持稳定。同时，我们还将随机变量 X 的观测数量固定为 N 帧，以此保证混合高斯分布的参数 Φ 在这 N 帧观测数据中也不产生变化，这样我们就可以根据最大化似然函数 $L(\Phi | X)$ 来估计参数的值。对于事件 $\{X_1 = x_1, X_2 = x_2, \dots, X_N = x_N\}$ ，当前采集到的 N 个观测值被随机变量 X 取得的概率为：

$$L(\Phi | X) = p(X | \Phi) = \prod_{i=1}^N p(x_i | \Phi) \quad (4-19)$$

$$\hat{\Phi} = \arg \max_{\Phi} L(\Phi | X) \quad (4-20)$$

在求解式(4-20)时，根据对数的单调递增性特点，我们知道 $\ln L$ 和 L 可以在同一点取得最大值，显然利用对数似然函数 $\ln(L(\Phi | X))$ 进行计算更易实现，并且我们可以直接将对数似然函数中求得的 Φ 参数最大值作为参数的估计值 $\hat{\Phi}$ 。常用的导数取零方法只对简单的分布有用，且用导数取零方法求解混合高斯分布的参数最大值时，得到的解不是唯一的，也就无法确定最佳估计值。采用 EM 算法求解参数 Φ ，首先需要的计算出“隐含数据”的分布式，并通过它的分布式辨别出随机变量 X 的当前观测值 x_t ($t = 1, 2, \dots, N$) 是哪一个高斯分布产生的，通过贝叶斯法则可以得到 Y 的分布式，如下所示：

$$p(k | x_t, \Phi) = \frac{\pi_k f(x_t | k, \theta_k)}{f(x_t | \Phi)} \quad (4-21)$$

$$f(x_t | \Phi) = \sum_{l=1}^M \pi_l f(x_t | l, \theta_l) \quad (4-22)$$

式(4-21)中，取 k 值的最大后验估计 \hat{k} 可以使后验概率值 $p(k | x_t, \Phi)$ 取到最大值，此时像素观测值 x_t 是由第 k 个高斯分布产生的，同样也采用 EM 算法来求解 \hat{k} ，因为

$\sum_{l=1}^M \pi_l f(x_t | l, \theta_l)$ 和 k 的值无关, 可以得到下式:

$$\hat{k} = \arg \max_k p(k | x_t, \Phi) = \arg \max_k \pi_k f(x_t | k, \theta_k) \quad (4-23)$$

混合高斯分布的参数可以用以下的几个迭代方程式进行解算:

$$\pi_k^{t+1} = \frac{\sum_{t=1}^N p(k | x_t, \Phi^t)}{N} \quad (4-24)$$

$$\mu_k^{t+1} = \frac{\sum_{t=1}^N x_t \cdot p(k | x_t, \Phi^t)}{\sum_{t=1}^N p(k | x_t, \Phi^t)} \quad (4-25)$$

$$\Sigma_k^{t+1} = \frac{\sum_{t=1}^N p(k | x_t, \Phi^t) (x_t - \mu_k^{t+1})(x_t - \mu_k^{t+1})^T}{\sum_{t=1}^N p(k | x_t, \Phi^t)} \quad (4-26)$$

上面的四个推导过程需要事先假定观测像素值 x_t 的整个产生过程是一个平稳随机过程, 该过程内所有的统计特征保持恒定, 并且采集到的观测值的数量为固定值 N 。这样就完成了混合高斯背景的参数估计, 在此基础上建立高斯背景模型便可以进行目标的检测^[50]。

4.4.3 建立背景置信区间

利用混合高斯背景模型进行动目标检测时, 首先对其参数进行估计。通过观测采得一段时间内的序列图像, 对单一的像素点进行分析, 则其样本集是该点像素值在时间轴上的集合, 再根据得到的样本集合计算出高斯函数的初始均值 μ_0 与初始方差 σ_0^2 , 其计算式如下:

$$\mu_0 = \frac{1}{N} \sum_{t=0}^{N-1} I_t \quad (4-27)$$

$$\sigma_0^2 = \frac{1}{N} \sum_{t=0}^{N-1} (I_t - \mu_0)^2 \quad (4-28)$$

在进行式(4-27)和(4-28)的计算时, 需要较大的内存来满足多帧图像的存储需求。但是, 我们若对混合高斯中参数初始化的速度要求不高, 根据像素点的像素值取值范围为 $[0, 255]$, 将 M 个高斯分布的方差 σ_0^2 直接初始化为较大的值, 那么第 i 个高斯分布的权

重与均值的初始化如下所示：

$$\pi_i = \frac{1}{M} \quad (4-29)$$

$$\mu_i = 255 \cdot \frac{i}{M}, i = 1, 2, \dots, M \quad (4-30)$$

通常情况下，第一帧图像中出现背景的几率不太大，所以为了加快背景模型的建立速度，我们直接选取第一帧图像的像素值进行混合高斯模型其中一个高斯函数的均值的初始化，并假设该高斯函数占有较大的权重值。与此同时，假设其他高斯函数的均值初始值都为零，权重取较小的等值，而所有的高斯函数方差都取较大等值。

设置好各个高斯分布的参数后，还需要对参数进行更新。混合高斯背景模型不仅需要对各高斯分布的均值和方差进行更新，还需要更新各高斯分布对应的权值。设置好各个高斯分布的参数后，我们就可以利用建成的混合高斯背景模型对当前的像素点进行匹配，我们按照权重值依次将其和各个高斯分布进行匹配。匹配之后，若新的像素值属于 M 个高斯分布中任何一个，那就需要将新的像素值加入该高斯分布中，并利用新的像素值对该高斯分布的参数进行更新，参数的更新如下式：

$$\pi_{i,t} = (1 - \alpha)\pi_{i,t-1} + \alpha \quad (4-31)$$

$$\mu_{i,t} = (1 - \rho)\mu_{i,t-1} + \rho I_t \quad (4-32)$$

$$\sigma_{i,t}^2 = (1 - \rho)\sigma_{i,t-1}^2 + \rho(I_t - \mu_{i,t})^2 \quad (4-33)$$

$$\rho = \frac{\alpha}{\pi_{i,t}} \quad (4-34)$$

上式中的 ρ 表示参数的学习率，而 α 如在单高斯背景模型那节中介绍的一样，表示背景的更新速度，且 $0 \leq \alpha \leq 1$ 。当 α 的值太小，这时背景模型的更新比不上实际场景中背景的变化，造成背景模型的“滞后”，用这种情况下得到的模型进行目标检测效果必然不理想；而当 α 的值太大，此时背景的更新速度过快，可能导致场景中运动速度较慢的目标被误判为背景，再进行目标检测的话，检测可能会产生空洞或者拖尾问题，更有甚者造成漏报^[49]。

为了验证混合高斯建模法的有效性，对算法进行仿真，参数分别为 $u_1 = 1.5 = 1$ ， $u_2 = 3.5$ ， $u_3 = 5$ ， $\sigma_1 = 0.6$ ， $\sigma_2 = 0.7$ ， $\sigma_3 = 0.5$ ， $\pi_1 = 0.2$ ， $\pi_2 = 0.3$ ， $\pi_3 = 0.5$ ，其流程图如下：

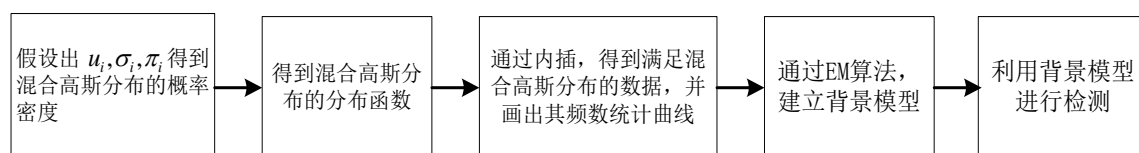


图 4.7 混合高斯背景建模仿真流程图

根据仿真流程图进行混合高斯背景建模，结果如下：

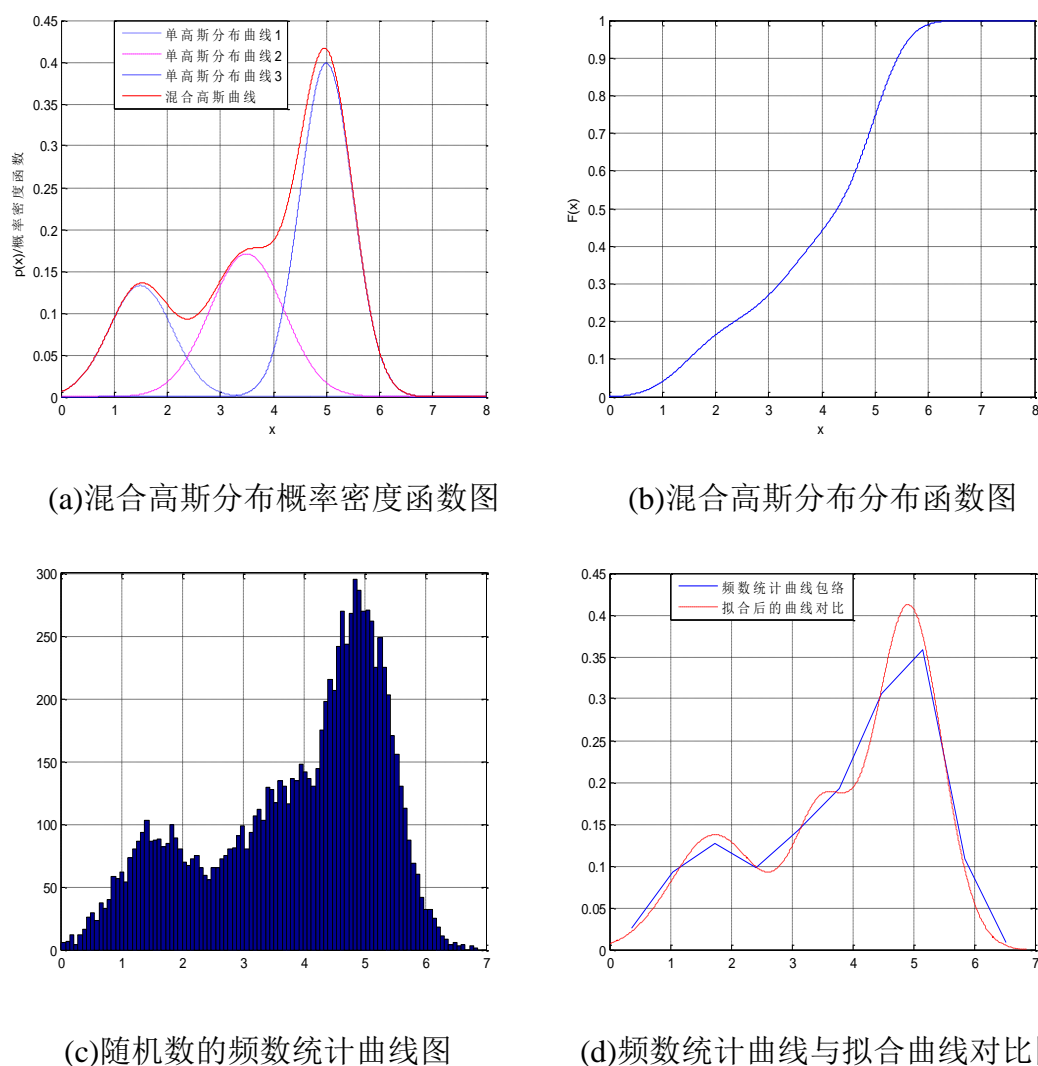


图 4.8 混合高斯背景建模法仿真结果

观察上面的仿真结果可以看出，经过拟合后的曲线与随机数的频数统计曲线十分接近，说明拟合效果良好，本文的算法能够较好的建立混合高斯模型。

在处理水池数据时，通过采集一系列背景图像，如图 4.9 所示，将每一个像素点的变化都描绘出来，根据 EM 迭代法直接利用这些序列背景图像各点的像素值建立其背景混合高斯分布的权值、均值和方差，再利用得到的参数建立各个背景高斯分布的置信区间，最后将当前帧图像中各点的像素值与置信区间进行匹配。其匹配原则是：当前像素值 x_t 与混合高斯背景模型中各个高斯模型的均值差值的绝对值要小于对应的高

斯分布标准差的 3 倍（如在单高斯背景模型一节中所介绍）。也就是说，当前点的像素值落入置信区间中的话，该像素点就满足背景的高斯分布模型，可以被判定为背景区域，而当前点的像素值未在置信区间中就表示该点位于目标区域中。

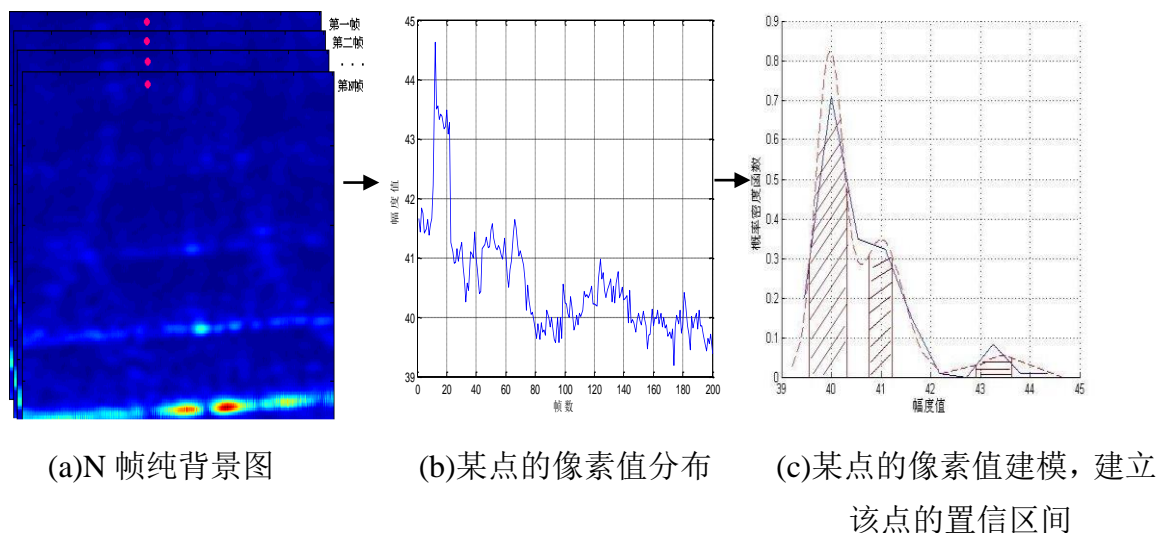


图 4.9 混合高斯背景建模法置信区间建立

图 4.10 给出了基于混合高斯背景建模法的动目标检测流程图：

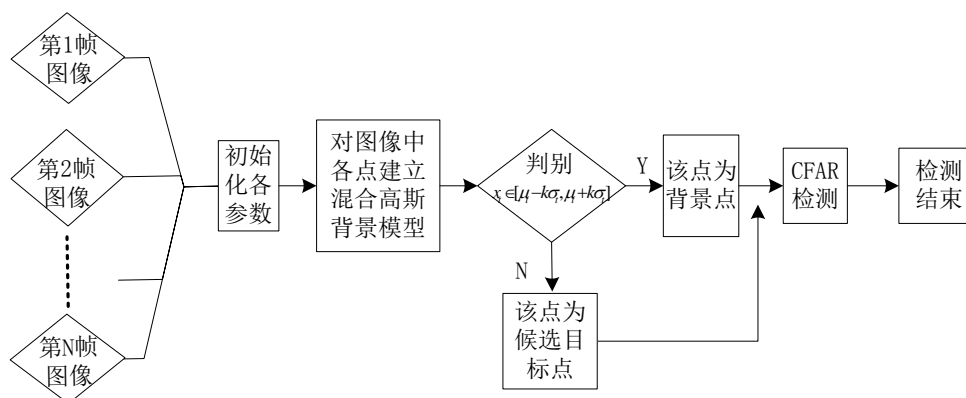
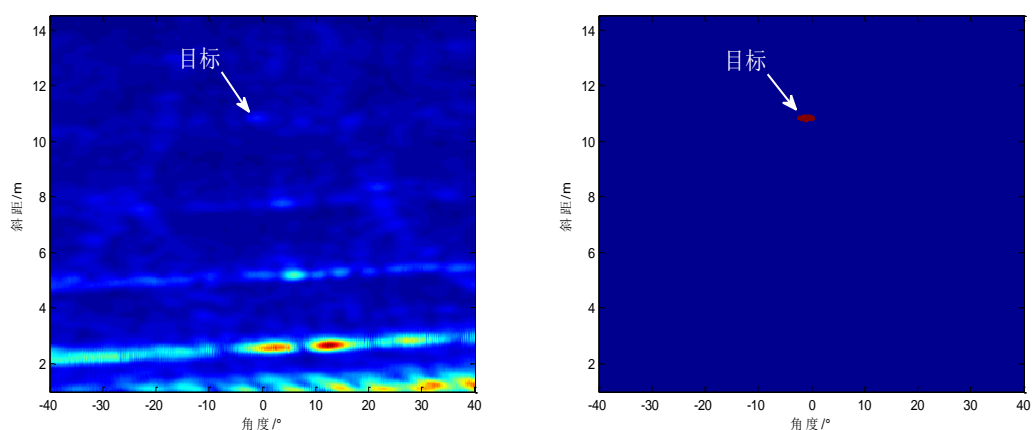


图 4.10 基于混合高斯背景建模法的目标检测流程图

根据本节中对混合高斯背景建模法的仿真分析可以看出，混合高斯背景建模法有较好的背景建模能力，在此，首先利用混合高斯背景建模法建立背景的置信区间，对待测帧的像素点进行判断，将不在置信区间的像素点归类为候选目标点。最后再用第三章中介绍的 CFAR 检测法对通过置信区间后的图像进行检测，得到运动目标的信息。

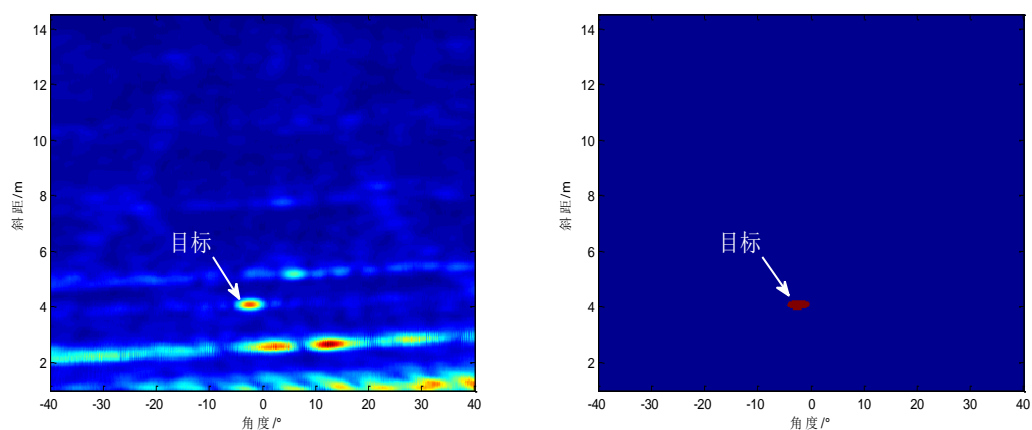
水池数据的基本参数如 3.2 节中介绍，第一组数据是单个目标由远及近进行匀速运动，检测结果如下：



(a)目标在远处的原始图像

(b)目标在远处的检测结果

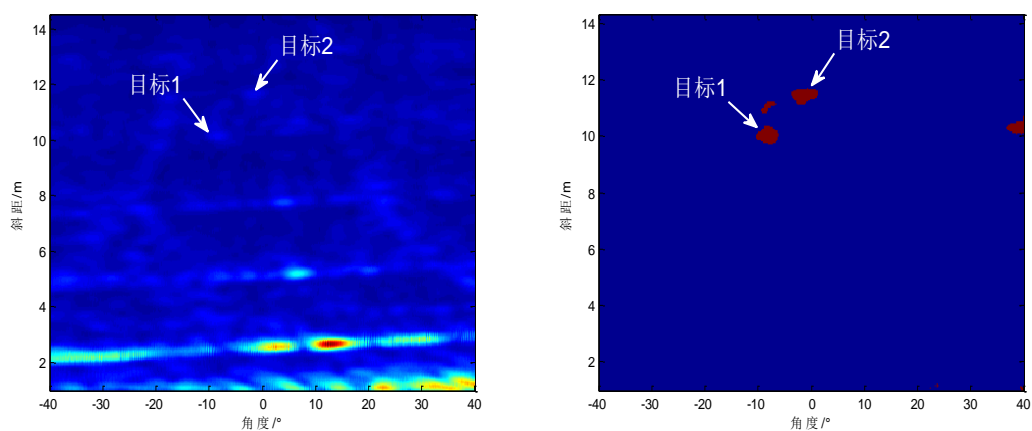
图 4.11 单目标在远处经过混合高斯背景置信区间后检测结果图



(a)目标在近处的原始图像

(b)目标在近处的检测结果

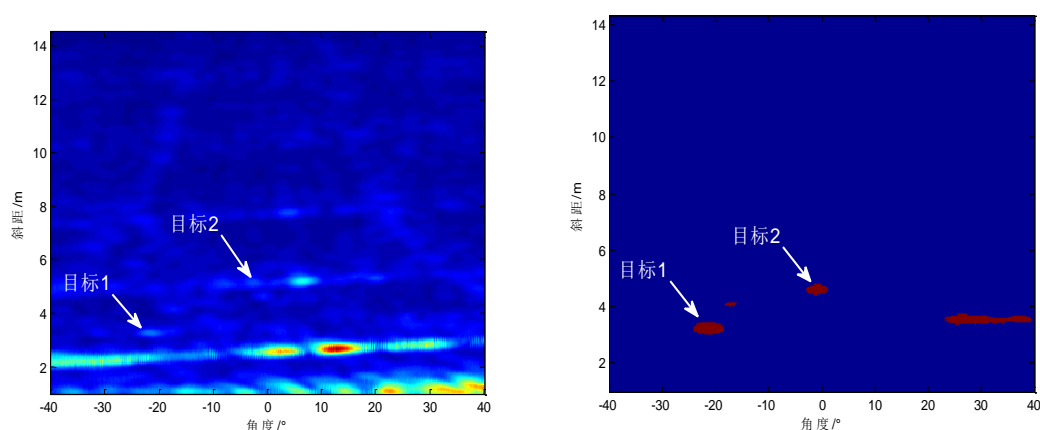
图 4.12 单目标在近处经过混合高斯背景置信区间后检测结果图



(a)目标在远处的原始图像

(b)目标在远处的检测结果

图 4.13 双目标在远处经过混合高斯背景置信区间后检测结果图



(a)目标在近处的原始图像

(b)目标在近处的检测结果

图 4.14 双目标在近处经过混合高斯背景置信区间后检测结果图

首先观察图 4.11 中(b)图对远处单目标的检测结果,与第二章中未经混合高斯背景建立的置信区间判决而直接进行的 CFAR 检测结果相比,由水底水面混响干扰在图像中产生的静态噪声得到了抑制,远处目标被清晰的检测出来,说明本文所采用的基于混合高斯背景建模法的背景置信区间能够较好的剔除静态目标和抑制背景干扰。

然后对双目标在远处的图 4.13 和在近处的图 4.14 可见,经过混合高斯背景置信区间判决后进行 CFAR 检测,能将成功地检测出两个弱信号的乒乓球目标,说明本文的算法有较好的检测效果。其中图 4.14 右侧出现的静态干扰带为目标位于近处时的旁瓣产生的静态背景干扰。

4.5 本章小结

本章对常见的动目标检测方法进行了详细介绍,并进行了仿真分析。然后在给出了本文所采用的混合高斯背景建模法之前,对单高斯模型的概念及优缺点进行详细的讲解。然后,根据混合高斯背景建模法得到背景的置信区间,结合第二章中的 CFAR 检测法进行了水池数据实验处理,并给出了结果。观察所得到的实验处理结果,可以说明这两种算法的结合的具较好的动目标检测能力。

结 论

本文围绕声纳图像中的动目标检测进行研究,由于海洋中的声背景复杂、存在的噪声污染严重,因此得到的声纳图像噪声严重、边缘恶化、对比度差,使得从声纳图像中有效的检测到运动目标十分困难。本文在声成像技术的基础上对声纳图像的预处理和分割进行了研究,并探讨了基于恒虚警(CFAR)的目标检测算法。通过混合高斯背景建模法建立背景的置信区间与 CFAR 检测法相结合,实现声纳图像的动目标检测。论文主要进行的工作及相关结论有:

1. 对基于波束形成的声成像算法进行理论概括与仿真分析,针对声纳图像受到噪声的严重干扰,图像分辨率低,目标边缘残缺的问题。对声纳图像的预处理方法包括图像增强(对比度展宽和直方图均衡化)与平滑去噪(低通滤波、中值滤波和均值滤波)进行了仿真分析。

2. 在对两种常见的声纳图像分割法,固定阈值分割法和天津阈值分割法的研究基础上,对 CFAR 法的理论进行详细的探讨,并将其应用于图像域的目标检测中,即通过对当前检测单元建立一个二维滑窗,利用滑窗内背景像素值计算出在恒虚警概率的条件下该检测单元的阈值,实现目标的检测。然后对在水池中获取的同一幅声纳图像用上述三种方法进行目标检测,最终发现 CFAR 法能够较好的检测出信噪比较低的目标,证实了图像域 CFAR 法的可行性。

3. 针对声纳图像中动目标的检测,分析了常用的帧间差分法,并通过仿真实验对该方法的检测性能进行了讨论,发现其存在“空洞”和“双影”的缺陷。对基于背景估计的动目标检测算法进行研究,在单高斯背景模型的基础上讨论了混合高斯背景建模法。文中利用混合高斯背景建模法建立背景的置信区间,然后用 CFAR 法对通过置信区间判决后的图像进行检测。水池数据的处理结果表明,这种方法有较好检测性能。

由于水下环境的特殊性,声纳图像中的动目标检测的实现相当复杂,本文所做的工作还不够,本文目前还在水池实验阶段,因此在分析恒虚警检测时只对混响服从瑞利分布的情况进行了研究。事实上,在高分辨率声纳时代,非瑞利分布对海洋中的混响统计特性描述应用更广泛,例如其中的 K 分布模型。因此,为了本文算法今后在海洋图像声纳中的应用,还需要展开对非瑞利分布的 CFAR 检测法的研究。此外本文中的通过混合高斯背景建模法建立的背景置信区间未进行更新,这些都是未来研究工作的改进方向。

参考文献

- [1] 李轲, 刘忠, 毛盾. 基于形态特征的水下小目标识别方法. 舰船科学技术. 2012, 34(1): 91-94 页
- [2] 尹凯华. 蛙人探测声纳在水下安保的技术实现. 上海船舶运输科学研究所学报. 2011, 34(1): 71-75 页
- [3] 董佳佳. 基于声纳图像水下运动目标识别与跟踪技术研究. 中国海洋大学硕士学位论文. 2011
- [4] 田坦, 刘国枝, 孙大军. 声呐技术. 哈尔滨: 哈尔滨工程大学出版社, 2006: 5-6 页
- [5] 桑恩方, 沈郑燕, 高云超. 水下机器人关键技术研究. 机器人技术与应用. 2008(4): 12-15 页
- [6] 王兴梅. 水下声纳图像的 MRF 目标检测与水平集的轮廓提取方法研究. 哈尔滨工程大学博士学位论文. 2010
- [7] 景志宏, 林钧清. 水下目标识别技术的研究. 舰船科学技术. 1999(4): 38-44 页
- [8] 张铁栋. 前视声呐的后置图像处理算法研究. 哈尔滨工程大学硕士学位论文. 2004
- [9] S. Geman and D. Geman, Stochastic Relaxation. Gibbs Distribution and Bayesian of Images. IEEE TPAMI, 1984, 6(11): 721-741P
- [10] Besag, J. On the Statistical Analysis of Dirty Picture. Journal of the Royal Statistical Society, 1986, 48(3): 259-302P
- [11] V. Murino. Reconstruction and segmentation of underwater acoustic image combining confidence information in MRF models. Pattern Recognition, 2001, 34(5): 981-997P
- [12] BRON J, FLEET D, BEAUCHEMIN S. Performance of optical flow techniques, International Journal of Computer Vision, 1994, 12(1): 42-47P
- [13] LIPTON A, FUJIYOSHI H, PATIL R. Moving target classification and tracking from real-time video. Proc of wacv'98, 1998: 8-14P
- [14] 孙志文. 复杂背景中移动目标的快速识别与跟踪, 微电子学与计算. 2004, 11: 48-55 页
- [15] ROBER T, T COLLINS. A system for video Surveillance and Monitoring, Pittsburg: Carnegie Mellon University, 2000: 53:62P
- [16] STAUFFER C, GRIMSON W. Adaptive background mixture models for real-time tracking, Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1999, 2: 248-252 P
- [17] Jain R, Nagel IH H. On the analysis of accumulative difference pictures from image sequences of real world scenes. Pattern Analysis and Machine Intelligence, IEEE

- Transactions on,1979(2):206-214P
- [18] Mittal A, Paragios N. Motion-based background subtraction using adaptive kernel density estimation//Computer Vision and Pattern Recognition,2004.CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on. IEEE, 2004,2: II-302-II-309 Vol. 2
 - [19] 李超, 熊璋, 赫阳等. 基于帧间差的区域光流分析及其应用. 计算机工程与应用, 2006, 41(31): 195-197 页
 - [20] 杨靖宇, 张永生, 于美娇等. 基于小波变换多分辨率分析特性的遥感影像动态重构的金字塔模型. 测绘科学, 2007, 32(5): 50-51 页
 - [21] 段昶. 活动序列图像中的运动目标检测. 电子科技大学硕士学位论文. 2006
 - [22] 王典. 基于混合高斯的背景建模与阴影抑制算法研究. 西安: 西北工业大学博士学位论文. 2006
 - [23] Fan J, Yau D K Y, Elmagarmid A K, et al. Automatic image segmentation by integrating color-edge extraction and seeded region growing. Image Processing, IEEE Transactions on,2001,10(10):1454-1466P
 - [24] 沈郑燕. 声纳图像去噪与分割技术研究. 哈尔滨工程大学博士学位论文. 2010
 - [25] 丁迎迎. 海底物体回波模拟与图像生成技术研究. 西安: 西北工业大学硕士学位论文. 2006
 - [26] 杨星, 杨万麟. 基于 FPGA 的数字下变频的研究与实现. 现代电子技术. 2009, 32 (9): 73-75 页
 - [27] 巫琴. 成像声纳数据传输和显控软件的设计与实现. 哈尔滨工程大学硕士学位论文. 2011
 - [28] 杨蕾. 时延, 相移波束形成技术研究. 哈尔滨工程大学硕士学位论文. 2004
 - [29] 杨威, 赵极远, 刘鑫. 基于 FPGA 的成像声纳 FFT 波束形成器设计. 电子技术应用. 2011, 10: 015 页
 - [30] 曹莉笠. 高分辨率成像声纳显控软件的设计与实现. 哈尔滨工程大学硕士学位论文. 2009
 - [31] 唐利娜. 水下三维声成像及声纳图像处理技术研究. 哈尔滨工程大学硕士学位论文. 2009
 - [32] 阳凡林, 独知行, 李家彪等. 基于 MRF 场的侧扫声呐图像分割方法. 海洋学报. 2006, 28(4): 43-48 页
 - [33] 朱虹. 数字图像技术与应用. 2001, 5(1): 131-135 页
 - [34] 方培生, 陆岚, 李广等. Visual C++在数字图像增强中的应用. 传感技术学报. 2005, 17(4): 600-602 页
 - [35] Chan T F, Osher S, Shen J. The digital TV filter and nonlinear denoising. Image

- Processing, IEEE Transactions on, 2001,10(2):231-241P
- [36] Ho K H L, Ohnishi N. FEDGE—fuzzy edge detection by fuzzy categorization and classification of edges//Fuzzy Logic in Artificial Intelligence Towards Intelligent Systems. Springer Berlin Heidelberg, 1997:182-196P
- [37] Wu S, Amin A. Automatic thresholding of gray-level using multistage approach//Document Analysis and Recognition, 2003. Proceedings. Seventh International Conference on. IEEE, 2003:493-497P
- [38] 孙文俊. 主动声纳抗混响模型与抗混响信号处理. 西北工业大学硕士学位论文. 2006
- [39] 张维. 复杂杂波背景恒虚警检测技术研究. 南京: 南京航空航天大学博士学位论文. 2009
- [40] 刘竹青. 非高斯混响背景下的信号检测. 哈尔滨工程大学硕士学位论文. 2008
- [41] Anitori L, Otten M, van Rossum W, et al. Compressive CFAR radar detection//Radar Conference (RADAR), 2012 IEEE. IEEE,2012:0320-0325P
- [42] 何友, 彭应宇. 雷达自动检测和 CFAR 处理方法综述. 系统工程与电子技术, 2001, 23(1): 9-14 页
- [43] 石炜. 海杂波的建模及 CFAR 目标检测研究. 哈尔滨工程大学硕士学位论文. 2009
- [44] 高宝山. 海杂波背景下的 CFAR 目标检测. 大连海事大学硕士学位论文. 2011
- [45] 覃剑. 视频序列中的运动目标检测与跟踪研究. 重庆: 重庆大学博士学位论文. 2008
- [46] 安博. 动态背景下运动目标检测的研究. 天津理工大学硕士学位论文. 2009
- [47] Emanuele Trucco, Tiziano Tommasini, Vito Roberto. Near-Recursive Optical Flow from Weighted Image Differences Systems, Man and Cybernetics, Part B,IEEE Transactions . 2005:77-85P
- [48] 普婕. 复杂动态场景背景建模与目标检测技术研究. 电子科技大学硕士学位论文. 2013
- [49] 邱道尹, 郑万晨, 顾波等. 基于背景模型的视频运动目标分割研究. 黑龙江科技信息, 2009(16): 30-30 页
- [50] Verbeek J J, Vlassis N, Krose B. Efficient Greedy Learning of Gaussian Mixture Models.Neural computation, 2003,15(2):469-485P
- [51] Gandhi P P, Kassam S A. Analysis of CFAR processors in homogeneous background, Aerospace and Electronic Systems, IEEE Transactions on, 1988,(24):427-445P
- [52] 何友, 关键, 孟祥伟等. 雷达目标检测与恒虚警处理. 北京: 清华大学出版社, 2011: 36-41 页
- [53] Daum F. Multitarget-Multisensor Tracking: Principles and Techniques. Aerospace and Electronic Systems Magazine, IEEE, 1996,11(2):41P

攻读硕士学位期间发表的论文和取得的科研成果

致 谢

光阴如梭，白驹过隙间，在北国哈尔滨近三年的硕士研究生学习和生活即将结束，我即将踏入人生中新的征程。但是，回望这三年美好的硕士求学生活，我的心中久久不能平静，这是我人生中一段既辛苦而又充实快乐的时光，在这些日子里，老师、同学、家人以及朋友们都给予了我无私的帮助和关怀，给予了我无比的信任和支持，使得我在三年里收获了太多太多，在此，我衷心的感谢他们。

首先，我要感谢我的导师孙大军教授。从论文的选题到修改，以致最终的完成，孙大军教授都对我的论文进行无私的指导和帮助。孙老师具有严谨的治学态度、深厚扎实的理论知识和极度负责任的敬业精神，这些都对我产生了极大的影响，这里我由衷的感谢孙老师。

其次，我还要感谢特别水声学院的滕婷婷老师，在她的悉心指导和帮助下，我的专业水平不断提高，理论知识不断丰富，在我的论文撰写过程中，滕老师提出许多宝贵意见。无论在学习还是生活遇到难题的时候，滕老师总是耐心地对我进行指导和帮助。滕老师对我来说亦师亦友，而对她的感激之情我将永存心中。

感谢实验过程中给予我帮助的黄聪博士，黄聪博士不但专业知识扎实，而且乐于助人。同时还要感谢 1001 实验室的所有同学，这是一个温暖的大集体，在 1001 实验室的这些日子我们一起欢笑一起成长，每一个人散发了自己的光和热。其次，还要感谢答辩组的老师们，他们认真审查论文，提出意见，共同的探讨给我的学习更添了一份动力。感谢我的室友，她们是我本科时候同伴，我们已经共同走过了 7 年的时光，这些感情是我大学生活中最宝贵的财富。

最后，我要向我的父母、弟弟以及陈波同学表示感谢，他们总是默默的站在我的身后给我最强有力的支持，为我解决了生活中的其他难题，让我一心专注于论文的撰写，他们是我强大的精神支撑，我会在以后的生活和工作中用我的认真和努力来回报他们。