

分类号：_____

密级：_____

UDC：_____

编号：_____

工学硕士学位论文

水下声纳图像目标检测技术研究

硕士研究生：张元科

指导教师：叶秀芬 教授

学科、专业：模式识别与智能系统

论文主审人：王科俊 教授

哈尔滨工程大学

2013 年 12 月

分类号: _____

密级: _____

UDC : _____

编号: _____

工学硕士学位论文

水下声纳图像目标检测技术研究

硕士研究生：张元科

指导教师：叶秀芬 教授

学位级别：工学硕士

学科、专业：模式识别与智能系统

所在单位：自动化学院

论文提交日期：2013 年 12 月 16 日

论文答辩日期： 年 月 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 Target Detection Technique of Underwater Sonar Image

Candidate: Yuanke Zhang

Supervisor: Prof. Xiufen Ye

Academic Degree Applied for: Master of Engineering

Specialty: Pattern Recognition & Intelligent System

Date of Submission: Dec. 16, 2013

Date of Oral Examination:

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

摘 要

随着成像声纳系统的不断发展和完善,以及水下智能机器人的出现,开展基于水下声纳的自主目标检测与识别技术的研究,在军事和民用领域都具有重要的意义和价值。本论文在通过对不同的目标检测和分割算法进行大量实验的基础上,对水下目标检测识别系统中的图像数据预处理、声纳图像分割和声纳图像的特征提取技术进行了研究和实验分析。并在此基础上开发出了相应的实时目标检测方法。主要工作如下:

首先,通过在声纳图像的预处理中对声纳图像特征的分析,得知声纳图像普遍噪声污染严重。基于此,对水下声纳图像的空间域去噪和频率域去噪方法进行了分析与实验。着重将基于多分辨率的高斯金字塔滤波方法应用于声纳图像的预处理中,取得了一定的效果。

其次,针对声纳图像的特点,设计了一种新的基于 MRF 的非监督声纳图像分割方法。在声纳图像的海底混响区服从 Gamma 分布的情况下,通过一种快速有效的高斯金字塔模型对声纳图像进行预处理,使得处理后的声纳图像的海底混响区服从高斯分布。在此基础上构造了一个自动确定声纳图像分类个数的模型。并通过该模型结合一种局部能量极值化的方法对马尔科夫模型的初始化参数进行估计,从而形成一种完全自动的声纳图像分割模型。然后利用该模型对声纳图像数据进行了分割实验,并和其他典型的分割算法进行了比较,验证了该方法的有效性及其快速性。

再次,在分析声纳图像的纹理特征、形状特征和声纳图像特征提取技术的基础之上,着重对一种新的基于类 Haar 特征的声纳图像特征提取技术进行了实验分析与介绍。实验结果表明,采用该方法提取声纳图像特征的效果较为明显。

最后,基于以上技术的研究,针对 Klein5000 侧扫声纳和 BlueView 前视声纳,分别设计出了与之适应的实时目标检测方法。此外,对基于类 Haar 特征的目标检测方法也进行了相应地研究。针对 BlueView 前视声纳目标检测方法中同一检测算法不能适用于所有帧的目标,而使得某些帧的图像目标有漏检的情况,本文采用了基于卡尔曼预测相关联的目标检测技术,即将先前的目标检测结果和基于卡尔曼的目标预测结果进行综合,得到最终的检测结果。实验表明,该方法效果较为理想。

关键词: 声纳图像; 目标检测; MRF; 类 Haar 特征; 卡尔曼预测

ABSTRACT

Autonomous target detection and recognition based on underwater sonar shows its significance and value in both military and civilian fields with the continuous development of imaging sonar systems and the emergence of underwater intelligent robots. In this paper, a large number of experiments were carried out on the underwater target detection and recognition system with using different target detection and segmentation algorithms. Based on these, we studied the image data pre-processing image segmentation and sonar image feature extraction. And we designed corresponding real-time target detection systems. The main work is as follows.

Firstly, we learned that sonar image noise pollution is serious in the most sonar images through the analysis in the pre-processing of the sonar images. On this basis, both the spatial domain and frequency domain denoising methods of underwater sonar images are analyzed and experimented. And a Gaussian pyramid method based on multi-resolution sonar image filtering is applied to the pre-noising, and it achieves relatively good results.

Secondly, taking the sonar image characteristics into account, we optimize an unsupervised automatic sonar image segmentation method based on MRF (Markov Random Field). After analyzed a large number of sonar images, we find that the reverberation area of sonar images basically obey the rule of Gamma distribution. In this case, we use a fast and effective Gaussian Pyramid model for the sonar image preprocessing to make the underwater reverberation of these images obey Gaussian distribution. On this basis, we design a kind of model to classify the sonar images automatically. And we use a local energy maximum method to estimate the initialization parameters of MRF model. Then a fully automated sonar image segmentation model was formed by these two steps. Also, we design experiments to test this model. Compared with other typical sonar image segmentation algorithm, this method is effective and fast.

Thirdly, after having analyzed the texture features, shape features and feature extraction techniques of sonar image, a sonar image feature extraction method focusing on a new class-based Haar features are designed. The experimental analysis and results show that the effect of image feature extraction is improved obviously by this method.

Finally, based on the above researches, corresponding real-time target detection systems were designed for the Klein 5000 sidescan sonar and BlueView forward looking sonar. And a research on Haar features class-based object detection system is also carried out. For the

problem that the same detection algorithms cannot apply to targets in all frames on BlueView forward looking sonar target detection system, which causes missing target in some frames on particular situations. This paper uses a Kalman prediction associated with target detection technology. i.e., we combine the previously target detection results and predictions based on Kalman comprehensive goals in order to get the final test results. Experiments show that the method works better.

Key words: sonar image; target detection; MRF; similar Haar features; Kalman prediction;

目 录

第 1 章 绪论	1
1.1 研究的目的是和意义	1
1.2 国内外研究现状	2
1.2.1 水下目标检测与识别系统分析	2
1.2.2 声纳图像分割技术	2
1.2.3 水下特征提取技术	4
1.2.4 水下目标检测技术	4
1.3 论文的主要研究内容和结构安排	5
第 2 章 声纳图像的预处理	7
2.1 声纳图像数据分析	7
2.1.1 声纳的探测工作原理	7
2.1.2 声纳图像的特点	8
2.2 水下声纳图像去噪概述	9
2.3 声纳图像的空间域去噪	9
2.3.1 常见的空间域图像去噪方法	9
2.3.2 实验结果与分析	11
2.4 声纳图像的频率域去噪	13
2.4.1 小波变换理论	13
2.4.2 基于小波变换的声纳图像去噪	15
2.5 本章小结	17
第 3 章 水下声纳图像的分割技术	19
3.1 基于阈值的声纳图像分割	19
3.1.1 直方图方法	19
3.1.2 最大类间方差法	20
3.1.3 最大熵法	21
3.2 基于聚类的声纳图像分割	22
3.2.1 K 均值聚类	22
3.2.2 FCM 聚类	24
3.2.3 聚类实验与分析	25
3.3 基于 MRF 的非监督声纳图像分割方法	26
3.3.1 声纳图像的直方图分析	27

3.3.2 图像分割中的 Markov 随机场模型及算法	29
3.3.3 非监督的声纳图像分割模型	31
3.3.4 实验结果与分析	33
3.4 本章小结	36
第 4 章 声纳图像的特征提取与分析	37
4.1 常用的声纳图像特征	37
4.1.1 纹理特征	38
4.1.2 形状特征	43
4.2 基于类 Haar 特征的声纳图像特征提取	48
4.2.1 Haar 特征及类 Haar 特征	48
4.2.2 特征提取实验	49
4.3 声纳图像的特征提取方法分析	50
4.4 本章小结	51
第 5 章 水下声纳图像的目标检测技术	53
5.1 水下目标检测技术	53
5.2 基于聚类分析的侧扫声纳目标检测	54
5.2.1 Klein5000 侧扫声纳图像数据分析	54
5.2.2 侧扫声纳目标检测流程图	54
5.2.3 实验结果与分析	55
5.3 基于类 Haar 特征的目标检测技术	57
5.3.1 声纳图像目标分析	57
5.3.2 流程图与实验结果	58
5.4 卡尔曼预测相关联的目标检测技术	59
5.4.1 数据分析	59
5.4.2 卡尔曼预测原理	60
5.4.3 前视声纳目标检测流程图	61
5.4.4 实验结果及分析	63
5.5 本章小结	65
结 论	67
参考文献	69
攻读硕士学位期间发表的论文和取得的科研成果	77
致 谢	79

第 1 章 绪论

1.1 研究的目的和意义

进入 21 世纪以来,世界经济发展迅速,同时,陆地资源也被过度开发,人类生存空间也在不断缩小,以及世界上发达的国家对海洋资源争夺的日趋激烈化,使得海洋拥有的战略地位越来越重要^[1]。海洋是人类可持续发展极其宝贵的财富,主要原因是它蕴藏着极其丰富的矿产和生物资源,已探明的资源储量惊人,同时还有许多未被发现或尚未开发的潜在战略资源。在复杂的水下环境中,信息的获取主要依靠声学探测器材,这是因为相对于其它载体而言,声波是目前唯一能进行水下远程探测和通讯的手段,是海洋中能够实现信息传递的最有效的载体。声学探测器材主要是声纳,声纳是用于对水下目标进行探测、分类、定位和跟踪的主要工具;而且,声纳技术还广泛应用于鱼群探测、海洋石油勘探、船舶导航、水下作业、水文测量和海底地质地貌的勘测等^[2]。

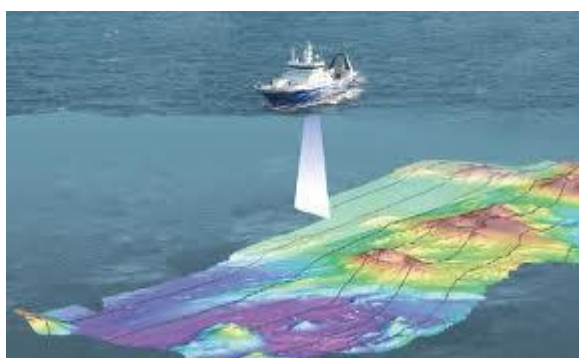


图 1.1 声纳应用于海底地貌绘制

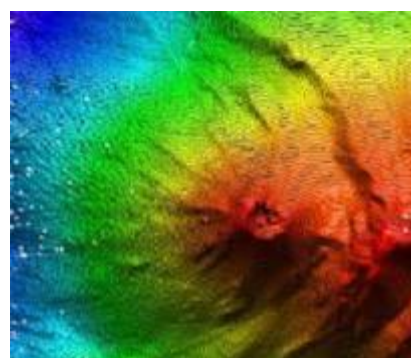


图 1.2 声纳应用于海底地质探测

随着信息化和智能化技术的不断发展,以及水下智能机器人的出现,水下声纳设备智能化的要求也越来越高。因此,开展基于水下声纳图像的目标自主检测与识别技术的研究,在军事和民用领域里都具有极其重要的意义和价值。

但在基于水下声纳图像的目标进行识别之前首先要进行目标的检测,只有更好地检测到目标,才能对目标进行更好更高精度地识别。水下声纳图像目标检测与识别存在的关键问题是图像分割,它不仅是图像处理中的关键步骤,也是机器视觉领域低层视觉中的经典难题。再加上水下复杂的环境及机器人本身设备环境的影响,声纳图像包含的各种污染噪声,使得水下声纳图像的分割和目标的检测识别变得更加困难。因此,如何快速准确地进行水下目标检测是实现实时快速水下目标识别技术的关键。

因此本文研究的目的是通过查阅国内外的相关文献资料和对不同的目标检测算法进行大量实验的基础上,分析已有的水下声纳图像目标检测算法中的优势与不足,结合声纳图像本身的特点和声纳目标检测识别系统对实时性和目标检测精确性的不同要求,对不同的水下声纳目标检测算法进行研究和改进。在提高分割算法的效率和准确

度的同时,对非监督的声纳图像分割方法进行了研究;并针对系统中对声纳目标检测的实时性、高检测率和低虚警率等的要求,对基于聚类分析的声纳目标检测和基于类 Haar 特征、以及卡尔曼预测相关联的声纳目标检测技术也进行了研究。此对水下目标实时自主识别技术的实现具有很重要的现实意义。

1.2 国内外研究现状

1.2.1 水下目标检测与识别系统分析

伴随着成像声纳系统的不断发展和完善,研究工作者们也将各种图像处理和模式识别技术应用到基于水下声纳图像的目标检测与识别系统中,在国内外相关的技术文献中,已经提出了许多种基于水下目标的分类识别系统和算法。传统的水下声纳图像目标检测与识别过程一般可分为图像预处理、图像分割、特征提取、目标认证和目标的分类识别等。系统框图如图 1.3 所示。

其中,目标的认证是指在特征提取之后,根据已建立的目标模型对目标候选区域进行的认证。目标的认证只需要验证候选区域是否属于目标,而不需区分是哪一个目标。而目标模型的建立则依赖于对目标的描述,目标模型参数通常由大量的目标样本训练得到。而目标的识别是指在目标认证之后,如果需要确定具体是哪一个目标,则要对目标进行进一步地分类。分类方法大多采用近邻法分类器、结构式分类器、树形分类器、或者基于学习的分类器来实现。

传统的水下声纳图像目标检测识别方法一般很少注重实际应用方面的研究,而一般的实际应用水下声纳图像目标检测识别方法要求整个过程是连续的,特别是目标检测阶段,实时性的要求更为重要。再者就是检测目标的检测率要高,以及虚警率要低等的要求。

结合以上问题的分析和实际声纳试验数据少,特征信息不足的情况考虑,本论文不准备对目标的识别和目标模型的建立进行研究,目标认证阶段将根据实际项目经验进行判定;而着重从实时性和高检测率的要求出发对目标检测技术中的图像分割和特征提取技术进行研究。

1.2.2 声纳图像分割技术

图像分割技术是图像处理中的一项极其关键的技术,在 20 世纪中期,研究学工作者们就对图像分割技术开始了研究,20 世纪 70 年代起开始引起广泛重视,到目前为止,已经提出了许多种有效的分割方法^[3]。根据不同的图像信息,图像分割方法通常被分为五大类:基于阈值的分割方法、基于区域的分割方法、基于边缘检测的分割方法、基于模型的分割方法和基于人工智能的分割方法等^[4]。

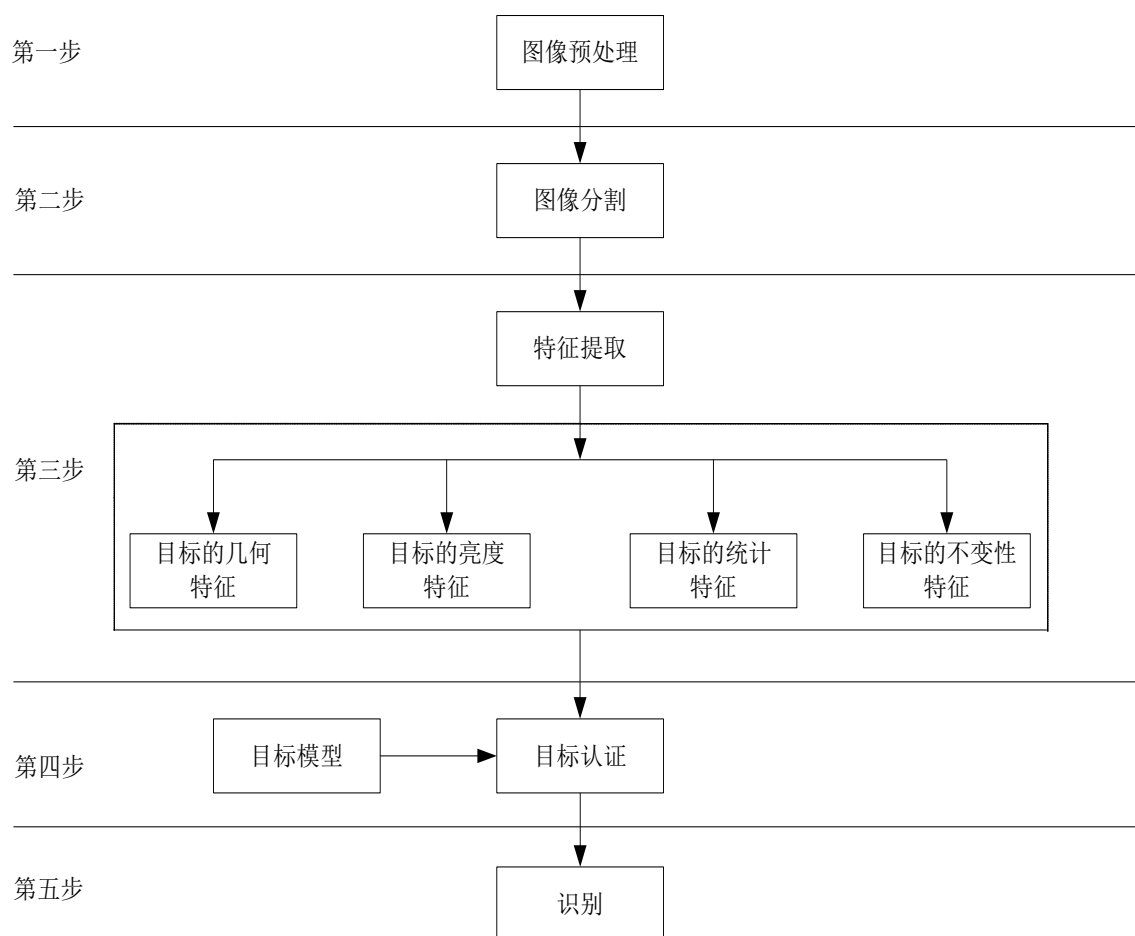


图 1.3 水下声纳图像目标检测与识别系统框图

由于声纳设备成像的非线性和水下声场环境的复杂性，所采集到的水下声纳图像具有对比度低和成像质量差等特点，其中基于阈值、区域和边缘检测的分割方法在声纳图像中的应用较少，而基于模型和人工智能的声纳图像分割方法在国内外都有较多的研究，下面就此进行介绍。

模糊聚类 and 神经网络是在图像分割领域应用广泛的人工智能技术，原理也就主要是利用人工智能方法得到图像分割的某一参数，然后基于得到的该参数采用图像分割的方法来分割图像。而基于人工智能的声纳图像分割方法主要是指基于聚类的分割方法。Stitt J P^[5-6]等利用声纳图像的灰度信息采用了一种改进的模糊 C 均值聚类方法，但分割效果不理想；郭海涛^[7]等又采用一种基于模糊聚类的声纳图像多区域分割方法，最终的分割效果相对较好。文献[8]针对侧扫声纳图像的特点，对 FCM 聚类算法进行了改进，采用了新的聚类准则和距离函数，该算法具有较好的抗噪性和准确性。文献[9]结合粒子群算法对 FCM 算法进行了改进，并用该方法实现了高分辨率声纳图像的分割。

近年来研究较热的基于模型的声纳分割方法主要有水平集方法和马尔科夫随机场（Markov Random Field, MRF）理论。Mumford-Shah 模型是一种基于区域的水平集模型方法，它将图像看作是若干个分片光滑的区域的组合^[10]。MS 模型是 Chan 等提出的

另一种简化的无边界主动轮廓模型^[11-13]。MS 模型认为图像被闭合边界划分为目标和背景两种分片常值区域，并用二相的水平集方法来进行数值来求解。而文献[14] [15]等对基于水平集的方法对声纳图像进行了分割研究，取得了一定的进展。

基于 MRF 的分割方法已被国内外的研究者在声纳图像上的应用进行了深入地研究，取得了重要的研究成果。如文献[16]为此作了开创性的工作，基于 Gibbs 分布和 MRF 的一致性建立了关于重建图像及其边缘的联合先验分布模型；文献[17]采用了基于分层的 MRF 声纳图像分割模型；文献[18]尽管使用 MRF 模型对声纳图像分割取得了较好的分割结果，但是其初始分割需要根据图像来人工选择窗口的大小，加上算法运算较为复杂，很难达到水下目标分割的自动和实时性要求；文献[19]中使用了快速的 K 均值聚类以及 FCM 算法对 MRF 初始化的参数进行估计，但是也要先人为地确定聚类的数目，而且算法速度较慢，很难用于实时性的声纳图像分割；文献[20]采用了一种非监督的声纳图像分割算法，也是在假设类数已知的情况下进行的。

从上面的分析可以看出，目前的分割算法大都是人工确定模型的参数或者是人为地确定聚类的类别数目，而没有一种完全自动的声纳图像分割方法。而实际的基于水下声纳的实时目标识别系统方法中的情况是复杂的，为了能够实现完全自动地水下目标识别，就不得不进行非监督的声纳图像分割算法的研究，同时要考虑系统方法的实时性。因此，本文设计了一种基于 MRF 的非监督声纳图像的自动分割方法，不仅能够自动地确定 MRF 模型的初始化参数，而且还能自动地确定声纳图像要分割的类别数。

1.2.3 水下特征提取技术

特征提取是为了提取出能反映物体本质的一些特征属性。不同的特征有不同的特性，如亮度、形状、灰度等，这些属于视觉上的特征，另外还有直方图特性、频谱特性等。在图像处理过程中，有时候需要针对具体问题来选取特征。如果特征选取的不适合，会影响整个系统的性能。

声纳图像目标的特征提取技术一般很少被单独研究，往往是结合目标的检测和识别技术进行研究和分析。本文也将结合具体的声纳实验数据进行特征的提取分析和研究，并着重把人脸识别中的 Haar 特征方法应用于声纳图像目标的特征提取中。

1.2.4 水下目标检测技术

目标检测技术是图像处理中的关键基础技术，利用运动目标检测技术可以对视频图像中感兴趣的目标进行实时地检测和识别等，是进行进一步视频图像处理的重要依据。目标的检测原理是将运动目标从视频序列图像背景中分离出来。

在光学成像的视频中，检测运动目标的方法已经非常成熟，常用的有点检测法、帧间差分法、光流法、背景减除法、基于分割的方法和监督学习的方法等。

水下目标的探测技术相对于声纳图像的分割技术，国内外相关的文献较少，文献[21]对基于多分辨率声纳的目标检测算法进行了研究；文献[22]对基于蛙人探测声纳的下目标进行了检测算法研究；文献[23]针对水下目标的几何特征，采用了一种基于弦中点的改进 Hough 变换椭圆检测方法，实现了水下多目标的特征提取，并对水下椭圆型目标进行了检测。文献[24]针对水声图像中人造物体的探测问题，给出了一种基于分形分析的方法，并对水下人造目标进行了检测实验和分析。

从以上的分析可以看出，现有的研究大都是对水池中的人造简单目标进行的，并没有针对复杂的、噪声干扰严重的湖试或海试水下目标进行研究与实验，也就是说，这样的研究没有结合具体的实际问题进行，往往是不可取的。再者由于水下目标的检测一般是针对实际运动的目标进行的，从而要求对目标的检测的准确度不仅要高，实时性也要有很高的要求。综合以上观点，本论文从快速性和准确性两个方面进行考虑和研究，设计了基于聚类分析的侧扫声纳目标检测方法、基于类 Haar 特征的方法以及卡尔曼预测相关联的目标检测方法，并通过实验验证了本文所提方法的有效性和快速性。

1.3 论文的主要研究内容和结构安排

随着计算机技术和微电子技术的发展，水下智能设备的要求也越来越高，其中尤为突出的是水下目标的自主检测识别技术，其研究意义的重要性和紧迫性不言而喻。本文通过对该领域国内外研究学者研究成果的分析，确定了本文的研究方向，结合实际的声纳实验数据，有针对性的研究水下声纳目标的自主检测技术，并针对不同型号的声纳开发出对应的实时水下目标检测方法。

本文分几章对各个部分及所用的具体方法做了详细的论述。具体安排如下：

第1章为绪论，首先介绍了本文研究的目的和意义，明确了本研究的重要性和紧迫性，然后分别对声纳图像的分割技术、水下目标特征提取技术和水下目标检测技术的国内外研究进展进行了较为详细的阐述。最后对本文开发的检测方法进行了简单的介绍和对整体文章结构进行了梳理。

第2章为声纳图像的预处理，首先对声纳图像的数据进行分析，主要是对声纳的探测工作原理和声纳图像的特点进行了研究。然后对声纳图像的常用去噪方法进行研究实验，其中包括常见的空间域方法：中值滤波、均值滤波、维纳滤波和偏微分方程方法滤波，以及基于小波的频率域去噪滤波等，是后续目标分割的理论基础。

第3章为水下声纳图像的分割技术，主要研究了基于阈值的声纳图像分割算法、基于聚类的声纳图像分割算法和基于 MRF 的声纳图像分割算法。针对目前的声纳图像分割算法大都是人工确定模型的参数或者是人为地确定聚类的类别数目，而没有一种完全自动的声纳图像分割方法。本文对基于 MRF 的声纳图像分割方法进行了改进，设计了一种不仅能够自动地确定 MRF 模型的初始化参数，而且还能自动地确定声纳图像要分

割的类别数的模型，分割实验效果较好。

第4章为声纳图像的特征提取与分析，本章首先对常用的声纳图像特征提取技术进行了简单的介绍，然后对声纳图像的 **GMRF** 纹理、**Gabor** 纹理和 **GLCM** 纹理的纹理特征，以及面积、周长、长短轴、主轴方向、紧密度、实心度、偏心率等广泛应用的几何特征进行了实验与分析。最后，着重对基于类 **Haar** 特征的声纳图像特征提取技术进行了实验分析与介绍，而目前国内外还没有相关的研究及应用。

第5章为水下声纳图像的目标检测技术，主要是在对前面的目标检测技术进行大量研究与实验的基础上，基于 **Klein5000** 侧扫声纳，研究出相应的实时侧扫声纳目标检测方法；基于 **BlueView** 前视声纳，研究出了相应的实时前视声纳目标检测方法；对基于类 **Haar** 特征的实时目标检测方法也进行了简单的介绍。并分别对其进行了实验与分析。

最后对全文研究进行了总结和对下一步的研究及应用进行了展望。

第2章 声纳图像的预处理

由于水下环境和水声信道复杂多变，以及声波在传播过程中存在衰减、混响、散射、多径现象和旁瓣干扰等，生成的声纳图像干扰严重、分辨率低等^[25-26]。水下声纳的预处理主要包括水下目标对比度的增强、减少图像中噪声的干扰的图像平滑滤波、增强目标边缘轮廓和改善目标图像清晰度图像锐化^[27]。另外还包括声纳声纳图像数据的读取与图像的建立、声纳图像的数据分析和图像的插值等。总之，对声纳图像进行预处理的目的在于对声纳图像进行增强。

2.1 声纳图像数据分析

2.1.1 声纳的探测工作原理

声纳探测系统工作原理是要以一定的俯仰角和扇面向海底区域发射超声波和接收回波信号。声波在起伏不平的海底发生漫散射，其中一部分返回给声纳系统，先接收的信号来自离接收基阵近的区域，后接收的信号对应于较远的区域。声纳接收到回波信号后，经过相应的补偿，检波等处理，得到水底不同位置返还的声波幅度，最后显示成声图像。声纳系统的接收和发射换能器绕铅直轴旋转，波束在水中扫描，每次采集同一距离的某一扇面若干个波束的数据作为一行的数据，不同的列对应不同的方位^[28]。

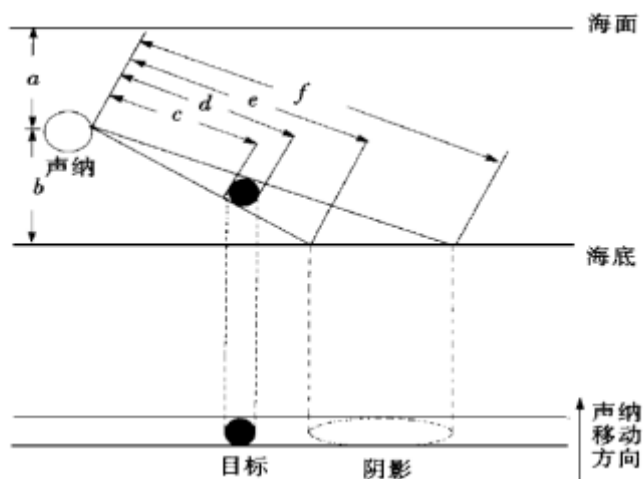


图 2.1 声纳回波记录示意图

图 2.1 显示的是声纳的回波记录几何示意图，它给出了声纳图像的基本组成和声纳成像的基本原理。其中 a 表示声纳设备在海平面以下的深度； b 表示声纳设备到海底以上的高度； c 表示目标的起始位置到声纳设备的距离； d 表示目标的终点位置到声纳设备的距离； e 表示阴影的起始位置到声纳设备的距离； f 表示阴影的终点位置到声纳设备的距离。声纳每一次扫描，都可获得不同的 c 、 d 、 e 、 f 值的记录，其中 a 和 b 的值可通过测深仪等传感器获得^[29-30]。

2.1.2 声纳图像的特点

声纳图像与光学图像在本质上都是能量平面或者空间分布图，但声纳图像与光学图像在成像机制上有着本质的区别。尽管人工智能学科已经发展比较成熟，但是很多可以用于光学图像处理的技术对声纳图像就不一定可行，特别是有一些前沿的图像处理技术没有在声纳图像处理中得到应用，因此是需要迫切地对声纳图像处理进行研究的。光学图像的成像原理是由感光元件通过接收目标物体发射或反射的光波而成像的。光波为感光元件和目标物体之间的单程直线传播，空气是其传播介质；而成像声纳则是通过发射换能器来发射声波，当遇到目标物体后，声波会发生反射，反射回波被接收换能器接收后，再利用目标物体材质、距离等不同所引起的回波信号幅度变换来生成声纳图像，是一种在换能器和目标之间双程传播的距离成像系统，其中，传播的介质对图像质量的影响相对较大，比如测量现场的水下环境、水温、水深、风速、噪声干扰、混响干扰、海洋内部的不均匀性、海底地形的不规则性等不确定因素都严重影响了声纳成像效果^[31-33]。

声纳图像的主要特征表现为：

- (1) 水下的噪声源丰富及环境比较复杂，噪声源有海洋环境噪声、舰船自噪声等，从而造成成像声纳图像普遍污染严重，相对较少的目标物体灰度级，和灰度级相对较为丰富背景噪声；
- (2) 由于成像声纳仪器中接收基阵可能存在着一一定的缺陷，例如仅在一个方向上形成分辨率及合成孔径声纳成像基阵载体运动不稳定等，致使得到的声纳图像分辨率低；
- (3) 由于声波的散射造成声纳图像中存在很多与周围像素灰度值有较大偏差的异常点；
- (4) 由于复杂的声波传播介质，使得接收到的声波信号被弱化，很多时候都不够完整，以至于声纳图像没有像光学图像那样拥有细致、精确、明显的边界特征，并且目标物体还经常出现边界残缺不全、不规则、不确定等特点；
- (5) 声纳图像清晰度一般较低，细节分量相对较少，图像主要以低频分量为主；
- (6) 有些声纳图像目标存在阴影，而有些声纳图像目标则没有阴影，同样使得目标的检测识别难度加大。

由于声纳图像的这些特性，使得对声纳图像的处理方法没有完全可靠的模型方法进行指导和完善，因此，这些因素在一定程度上阻碍了某些分析、处理和识别技术在声纳领域中的应用和发展。

2.2 水下声纳图像去噪概述

通过以上声纳图像特征的分析可知,相对于光学图像而言,声纳图像由于水下复杂的环境,使得成像声纳接受到的原始声纳图像数据质量急剧降低,噪声污染普遍严重。因此,对声纳图像的去噪技术进行研究对后序的声纳图像分割与识别都是至关重要的。

噪声一般都是随机产生的,因此都具有大小和分布的不规则性,有些噪声和图像信息互不相关、互相独立,而有些则是相关的。常见的图像噪声有乘性噪声和加性噪声,乘性噪声往往和图像交织在一起,而加性噪声往往又包括高斯噪声和椒盐噪声等典型噪声^[34]。因此,必须将声纳图像中的噪声滤除掉,得到更多更精细的原始声纳图像特征。

图像的有用信息一般情况下都集中在图像的低频部分,图像的高频部分主要是噪声,而往往图像的边缘部分和细节信息也集中在图像的高频部分,所以图像噪声的滤除要求所使用的滤波器不仅能够有效地去除噪声,而且还能尽量保留图像的原貌。因此根据图像的噪声不同,所使用的去噪方法往往也不同^[35]。

根据滤波所在空间的不同,图像去噪可分为基于空间域的方法和基于变换域的方法。前者是在原始图像上直接进行数据运算,对像素的灰度值进行处理。而后者是在变换域的图像上进行处理,然后进行反变换达到图像去噪的目的。常用的图像去噪方法有均值滤波、中值滤波、维纳滤波、小波变换、曲波以及基于偏微分方程的图像去噪等^[36-38]。

在上世纪50年代,傅里叶变换一直是变换域图像处理的基础,而小波变换解决了傅里叶变换对于非平稳信号处理中存在的缺陷。但是,小波变换在处理一维问题时所具有的优异特性并不能简单地推广到二维或更高维空间。为了能对高维奇异性的对象提供最优的表示方法,经过许多学者的努力,多尺度分析理论在图像去噪中也得以提出和发展。本文将在第三章对基于多尺度分析的理论进行简单介绍。

2.3 声纳图像的空间域去噪

2.3.1 常见的空间域图像去噪方法

1、中值滤波:

中值滤波方法是常用于数字图像预处理阶段的方法之一。中值滤波是一种减少边缘模糊的非线性平滑方法,它的基本思想就是利用该像素所在窗口的灰度中间值来替代该像素值,像素的原始值包含在中值计算结果中。

一般设有一个一维的序列 $f_1, f_2, f_3, \dots, f_n$, 取该窗口的长度(点数)为 m (m 为奇数), 对此一维序列进行中值滤波, 就是要从输入序列中相继抽出 m 个数, $f_{i-v}, \dots, f_{i-1}, \dots, f_1, \dots, f_{i+1}, \dots, f_{i+v}$, 其中 i 为窗口的中心位置, $v = (m-1)/2$ 。再将这 m 个点按其数值大小进行排序, 取中间的那个作为滤波器的输出, 用数学公式表示为:

$$\begin{cases} y_i = \text{med}\{f_{i-v}, \dots, f_{i-1}, f_i, f_{i+1}, \dots, f_{i+v}\} \\ i \in Z, v = (m-1)/2 \end{cases} \quad (2-1)$$

其中 Z 为一维序列点的个数。

对二维的图像序列进行中值滤波时，滤波窗口也是二维的，但这种二维窗口可以有各种不同的形状，如圆形、十字形、圆环形、线状、方形等。二维中值滤波可由下式给出：

$$y_{ij} = \text{med}\{F_{ij}\} \quad (2-2)$$

式(2-2)表示图像 F 中 (i, j) 点周围像素值的中间值。

2、均值滤波：

均值滤波方法是非常典型的线性滤波。其基本思想是用邻域均值来代替原来图像中的各个像素值，算法采用领域平均法来实现。其公式如下：

$$g(x, y) = \frac{1}{M} \sum_{(i,j) \in S} f(i, j) \quad (2-3)$$

式中 $x, y = 0, 1, 2, \dots, N-1$ ； S 是以 (x, y) 点为中心的邻域的集合， M 是 S 内坐标总数。

3、维纳滤波：

维纳滤波是一种自适应的滤波方法，它是在噪声和信号都是随机量，且其功率谱在互相交迭情况下提取信号的方法。它能根据图像的局部方差调整自身的输出，最终目标是使统计误差函数值 $\min MSE = E\{(f(x, y) - \hat{f}(x, y))^2\}$ 最小，其中， E 是期望值的操作符， f 是含噪声的图像， \hat{f} 是期望输出的图像。

4、偏微分方程方法：

数学分析的重要组成部分之一的偏微分方程，目前已被广泛应用到计算机视觉和模式识别等许多领域。在图像处理中应用偏微分方程，首要步骤是先建立图像的偏微分方程模型，假设图像是实数域上关于 x, y 的二维函数，由于图像是分片连续的，可以在除边缘外的连续区域内用函数来逼近原始图像^[39]。常见的偏微分方程去噪模型：热扩散方程、P-M扩散方程、TV扩散方程。下面简单的对偏微分方程的热扩散方程模型进行介绍：

热扩散方程是根据物理模型建立起来的，其物理模型为：一根局部被加热的铁管，随着导热过程它的热量会慢慢扩散，直到整个铁管的温度达到一致。将其应用在图像去噪中，就好比将噪声扩散使整个图像的灰度达到一致。

设原始图像为 $u(x, y, 0)$ ， $u(x, y, t)$ 为时间 t 时的扩散图像，则图像的热扩散偏微分方程为：

$$\frac{\partial u(x, y, t)}{\partial t} = \Delta u(x, y, t) \quad (2-4)$$

其中， $\Delta u(x, y, t)$ 为图像的拉普拉斯算子，其初始条件为 $u(x, y, 0)$ ，其解为：

$$u(x, y, t) = G_t * u(x, y, 0)$$

$$G_t(x, y) = \frac{1}{4\pi t} \exp\left(-\frac{x^2 + y^2}{4t}\right) \quad (2-5)$$

从式(2-5)可以看出，扩散后的图像即为原始图像与高斯滤波器的卷积，因此等价于低通滤波器，所以，热扩散方程是各向同性扩散，随着扩散的增强，图像的边缘等细节信息也被滤除了。

2.3.2 实验结果与分析

针对以上方法，实验选取了如图 2.2 的声纳图像进行滤波实验。下图分别给出了对图 2.2 的中值、均值等滤波效果图。

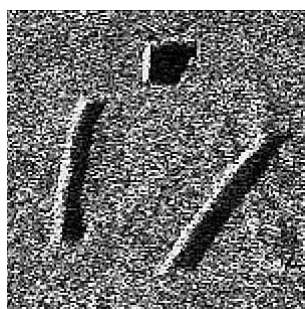


图 2.2 滤波实验用声纳图像

1、中值滤波效果：

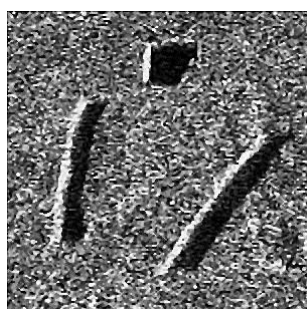


图 2.3 中值滤波效果图(3*3)

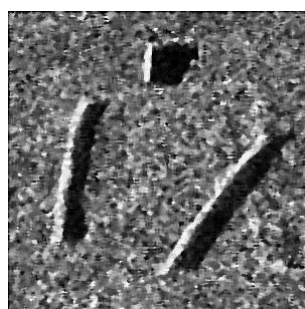


图 2.4 中值滤波效果图(5*5)

从总体上来说，中值滤波器对减少声纳图像中的随机干扰和脉冲干扰有很好的效果，并且算法简单，运行速度快。

但其缺点是去噪效果依赖于滤波窗口的大小，侵蚀孤立点的边缘，平滑噪声的同时会损失一些边缘细节。

2、均值滤波效果：

图像邻域平均法算法简单，计算速度快。但从图 2.5 和图 2.6 可以看出，它有在降低噪声的同时使图像产生模糊的缺点，特别在边缘和细节处，邻域越大，图像模糊越严重。



图 2.5 均值滤波效果图(3*3)



图 2.6 均值滤波效果图(5*5)

3、维纳滤波和偏微分方程滤波效果：

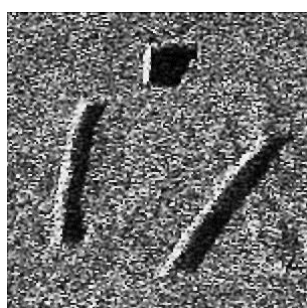


图 2.7 维纳滤波效果图

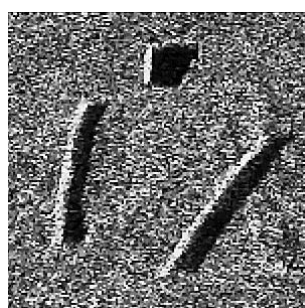


图 2.8 偏微分方程方法滤波

从上图 2.7 和图 2.8 可以看出，维纳滤波方法和基于偏微分方程的方法具有可以有效地去除光滑部分的噪声，将噪声和小尺度细节与纹理平滑掉的优点。但也有去噪后的图像有块儿现象，图像中的某些光滑区域呈现出块状，边缘和纹理也被模糊了的缺点。特别是基于偏微分方程的方法算法复杂，计算量大，不适用于实时性的声纳图像处理。

表 2.1 空间域去噪方法优缺点对比表给出了几种常见的空间域去噪方法所具有的优点和不足：

表 2.1 空间域去噪方法优缺点对比表

	优点	缺点
中值滤波	随机干扰和脉冲干扰效果好，算法简单	依赖于滤波窗口的大小、会损失图像的边缘细节
均值滤波	算法简单、计算速度快	图像容易模糊、滤波邻域越大，图像越模糊
维纳滤波	可以有效地去除光滑部分的噪声	光滑区域易呈现出块状，边缘和纹理也容易被模糊
偏微分方程方法滤波	具有将噪声和小尺度细节与纹理平滑掉的优点	边缘和纹理也容易被模糊；同时算法复杂、计算量大

2.4 声纳图像的频率域去噪

基于频率域的图像去噪方法最早可以追溯到傅里叶变换的使用,后来在此基础上提出了加窗傅里叶变换,同时也启发了小波多尺度分析思想的引入。小波理论是一门新兴的数学分支,传统的信号理论通常是建立在傅立叶分析的基础上的,与傅立叶变换相比,小波变换是空域和频率域的局部变换。小波变换继承了短时傅立叶变换局部化的思想,同时克服了窗口大小不随频率变化的缺点,多尺度特性为用户提供了更灵活的处理办法,可以选择任意的小波分解层数,用尽可能少的计算量得到满意的结果,目前已经成功应用于信号处理、图像处理和模式识别等领域^[40]。

2.4.1 小波变换理论

(一) 连续小波变换

1、一维连续小波变换

设函数 $f(x)$ 具有有限能量,即 $f(x) \in L^2(R)$, 则该函数的连续小波变换(Continuous Wavelet Transform)定义如下:

$$W_f(a, b) = \langle f, \psi_{a,b} \rangle = \int_{-\infty}^{\infty} f(x) \psi_{a,b}(x) dx \quad (2-6)$$

其中, $\psi_{a,b}(x)$ 是由母小波 $\psi(x)$ 平移、伸缩得到的小波函数, a, b, x 均为连续变量。对于任意的实数对 (a, b) , 小波母函数 $\{\psi_{a,b}(x)\}$ 能够通过平移和伸缩基本小波 $\psi(x)$ 来生成:

$$\psi_{a,b}(x) = \frac{1}{|a|} \psi\left(\frac{x-b}{a}\right) \quad (2-7)$$

其中 a 为尺度因子, b 为平移因子。通常情况下, 基本小波 $\psi(x)$ 以原点为中心, 因此 $\psi_{a,b}(x)$ 就以 $x=b$ 为中心。

连续小波逆向变换为:

$$f(x) = \frac{1}{C_\psi} \int_0^\infty \int_{-\infty}^\infty W_f(a, b) \psi_{a,b}(x) db \frac{da}{a^2} \quad (2-8)$$

2、二维连续小波变换

设 $f(x, y)$ 是一个二维函数, 则它的连续小波变换是:

$$W_f(a, b_x, b_y) = \int_{-\infty}^\infty \int_{-\infty}^\infty f(x, y) \psi_{a,b_x,b_y}(x, y) dx dy \quad (2-9)$$

其中 b_x 和 b_y 表示在两个维度上的平移。二维连续小波逆变换为:

$$f(x, y) = \frac{1}{C_\psi} \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty W_f(a, b_x, b_y) \psi_{a,b_x,b_y}(x, y) db_x db_y \frac{da}{a^3} \quad (2-10)$$

其中：

$$\psi_{a,b_x,b_y}(x,y) = \frac{1}{|a|} \psi\left(\frac{x-b_x}{a}, \frac{y-b_y}{a}\right) \quad (2-11)$$

而 $\psi(x,y)$ 是一个二维基本小波。

(二) 离散小波变换

1、一维离散小波变换

同傅里叶变换和沃尔什变换等正交变换类似，在实际应用尤其是在计算机上编程实现时，连续的小波变换必须离散化处理。

下面式子给出了对尺度参数和平移参数的离散化：

$$a = a_0^m, b = nb_0 a_0^m \quad (2-12)$$

其中， $a_0 > 1, b_0 \in R, m, n \in Z^2$ 。

连续小波和连续小波变换的离散化形式为：

$$\psi_{m,n}(x) = \frac{1}{\sqrt{a_0^m}} \psi\left(\frac{x - nb_0 a_0^m}{a_0^m}\right) = a_0^{-m/2} \psi(a_0^{-m} t - nb_0) \quad (2-13)$$

$$C_{m,n} = \langle f, \psi_{m,n} \rangle = a_0^{-m/2} \int_{-\infty}^{\infty} f(t) \psi(a_0^{-m} t - nb_0) dt \quad (2-14)$$

这里的 m 和 n 分别指频率范围指数和时间步长变换指数。

2、二维离散小波变换

图像处理中的小波变换，必然是二维的小波变换。二维基本小波为式(2-11)所示。

假定存在可分离的二维尺度函数 $\phi(x,y) = \phi(x)\phi(y)$ ， $\phi(x)$ 是一个一维尺度函数， $\phi(y)$ 为相应的小波函数。则可以得到三个二维基本小波变换函数：

$$\begin{aligned} \psi^1(x,y) &= \phi(x)\psi(y) \\ \psi^2(x,y) &= \psi(x)\phi(y) \\ \psi^3(x,y) &= \psi(x)\psi(y) \end{aligned} \quad (2-15)$$

由此建立了二维小波变换的基础。

一副 $N \times N$ 的图像 $f_1(x,y)$ ，其中下表指示尺度并且 N 是 2 的幂。对于 $j=0$ ，原图像的尺度为 $2^j = 2^0 = 1$ 。 j 值每一次增大都使尺度加倍，分辨率减半。每一次通过隔行隔列采样，图像被分解为 4 个大小为原来尺寸四分之一的子频带区域，这四个子频带区域的每一个都是由原图与一个小波基内积后，再经过 x 和 y 方向都进行 2 倍的间隔抽样而产生的。对于第一个层次 ($j=1$)，可写成：

$$\begin{aligned}
 f_2^0(m, n) &= \langle f_1(x, y), \phi(x-2m, y-2n) \rangle \\
 f_2^1(m, n) &= \langle f_1(x, y), \psi^1(x-2m, y-2n) \rangle \\
 f_2^2(m, n) &= \langle f_1(x, y), \psi^2(x-2m, y-2n) \rangle \\
 f_2^3(m, n) &= \langle f_1(x, y), \psi^3(x-2m, y-2n) \rangle
 \end{aligned} \tag{2-16}$$

而对于后继的层次 ($j > 1$)， $f_{2^j}^0(x, y)$ 都以完全相同的方式分解而构成 4 个在尺度 2^{j+1} 上更小的图像，如下图所示：

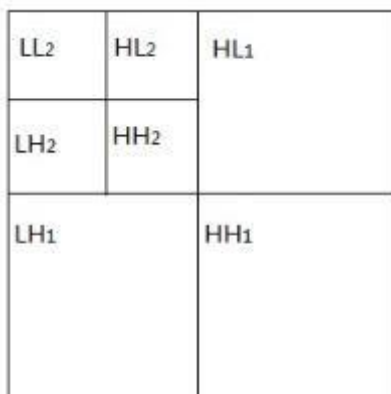


图 2.8 二维离散小波变换示意图

将内积写成卷积形式，可有：

$$\begin{aligned}
 f_{2^{j+1}}^0(m, n) &= \{[f_{2^j}^0(x, y) * \phi(-x, -y)](2m, 2n)\} \\
 f_{2^{j+1}}^1(m, n) &= \{[f_{2^j}^0(x, y) * \psi^1(-x, -y)](2m, 2n)\} \\
 f_{2^{j+1}}^2(m, n) &= \{[f_{2^j}^0(x, y) * \psi^2(-x, -y)](2m, 2n)\} \\
 f_{2^{j+1}}^3(m, n) &= \{[f_{2^j}^0(x, y) * \psi^3(-x, -y)](2m, 2n)\}
 \end{aligned} \tag{2-17}$$

并且在每一层次进行四个相同的间隔抽样滤波操作。

2.4.2 基于小波变换的声纳图像去噪

图像经过小波变换后，能量主要集中在最低分辨率子带图像上，而图像的加性随机噪声经过小波变换后，能量则分布在各个高分辨率子带图像上，因此，可以设定一个阈值，将绝对值小于阈值的小波系数当作噪声去除，达到去噪的目的。阈值去噪的关键是阈值的设定，因为过大的阈值会造成有用的高频信息(如边缘信息)丢失，使图像变得模糊；而过小的阈值又会保留过多的噪声，使去噪效果不明显。

一般的小波去噪过程如下图所示：

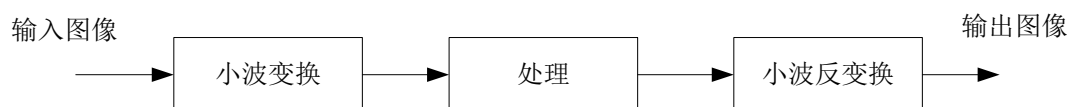


图 2.9 小波去噪流程图

通过利用小波变换处理分频带的图像信息是非常容易的，通过改变各个频带的小波

系数的幅值能够很好的完成滤波处理任务。每当经过一次小波分解后，如图 2.10 所示，图像被分解为 4 个四分之一的子频带区域，他们分别包含了相应频带的小波系数，*LL* 频



图 2.10 一次小波分解图

带是图像内容的缩略图，保持了源图像的内容信息，*HL* 频带含有水平方向的高频边缘信息，*LH* 频带含有竖直方向的高频边缘信息，*HH* 含有对角线方向的高频边缘信息，反映了水平和竖直方向上图像灰度的综合变化信息。

小波低通滤波的目的是保留低频带的小波系数，而衰减高频带的小波系数，从而可以去除图像中大量的边缘信息。

因此可以将 *HL*、*LH*、*HH* 三个频带中的一个或几个根据需要进行衰减，而主要保留 *LL* 频带的内容，再经过小波反变换即可得到平滑后的图像。

下面给出对图 2.2 滤波实验用声纳图像的小波去噪实验结果。

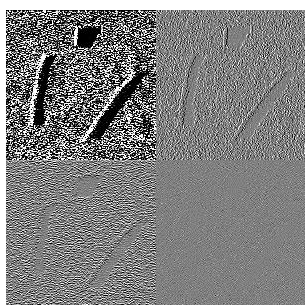


图 2.11 一层小波变换图

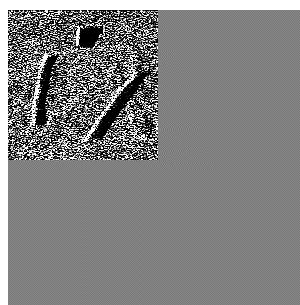


图 2.12 对高频区域进行衰减图

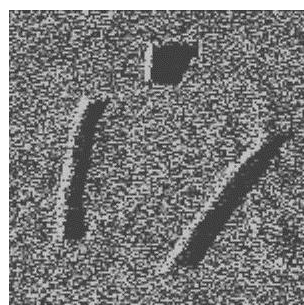


图 2.13 一层小波滤波效果图

从实验结果可以看出，由于水下混响在图像中表现出来的噪声是点奇异性，而小波在表示具有点奇异性的目标函数时是最优基，因此小波变化能够较好的对图像中的噪声起到抑制作用。但是利用小波分析处理后的图像边缘很容易变得模糊，也就是小波分析对声纳图像中边缘细节等重要信息的表达上有一定的局限性。

表 2.2 频率域去噪方法优缺点对比表给出了基于小波变换的声纳图像去噪方法所具有的优点和不足：

表 2.2 频率域去噪方法优缺点对比表

	优点	缺点
小波域声纳图像去噪	能够较好的对声纳图像中的噪声起到抑制作用	对声纳图像中边缘细节等重要信息的表达上有一定的局限性，图像容易模糊

2.5 本章小结

本章主要介绍了声纳的探测工作原理、水下声纳图像的主要特征和基于水下声纳图像的常用去噪方法。通过对水下声纳图像特征的分析可知，声纳图像由于水下复杂的环境，使得原始声纳图像数据质量急剧降低，噪声污染普遍严重。因此，不得不对水下声纳图像的滤波去噪方法进行研究。在声纳图像的去噪部分，首先对声纳图像的去噪技术进行了整体的介绍，然后对基于空间域的声纳图像去噪技术进行了研究与分析，主要包括：中值、均值、维纳和偏微分方程的方法。最后，对基于小波变换的频率域声纳图像去噪技术做了较为详细的分析与实验。但是基于上述的几种去噪方法，要么图像的边缘细节的表达上有一定的局限性，要么算法复杂、计算量大，不适用于实时性和精度要求都较高的声纳图像去噪中。本文将在第三章构造出基于多分辨率分析和高斯低通滤波相结合的声纳图像去噪方法，该算法简单，而且经过该方法处理后的声纳图像可以轻易地把图像目标和背景区域区分开来，实用性较好。

第3章 水下声纳图像的分割技术

水下声纳图像分割是水下目标自主识别的关键技术,是针对声纳图像进行分析和理解的基础,其分割的准确性和效率直接影响后期的目标特征提取和识别工作的有效性。虽然图像分割技术已经被人们广泛的利用,但是针对声纳图像分割技术的研究还相对缺乏,这主要是由于在声视觉系统应用的环境中,水声信道是时变和空变的,对在其中传播的声波信息产生各种复杂的作用。根据图像信息的不同,通常图像分割方法被分为五大类:基于阈值的分割方法、基于区域的分割方法、基于边缘检测的分割方法、基于模型的分割方法和基于人工智能的分割方法等。而由于水下声场环境的复杂性和声纳设备成像的非线性,所采集到的水下声纳图像具有成像质量差、对比度低等特点,本文主要介绍基于阈值、聚类、MRF 的声纳图像分割方法。

3.1 基于阈值的声纳图像分割

阈值分割是图像分割最常见也最简单的一种分割方法,阈值分割是基于对灰度图像的一种假设,目标或背景内相邻像素间的灰度值是相似的,但不同目标或背景的像素灰度上有差异,反应在图像的直方图上,不同目标和背景则对应不同的峰,选取阈值应位于两个峰之间的谷,从而将各个峰分开。阈值分割的阈值的确定方法常见的主要有直方图分割法、最大类间方差法和最大熵的方法等。

3.1.1 直方图方法

灰度直方图(histogram)是灰度级的函数,它表示图象中具有每种灰度级的像素的个数,反映图象中每种灰度出现的频率。灰度直方图的横坐标是灰度级,纵坐标是该灰度级出现的频率,是图象的最基本的统计特征。图 3.1 为阈值分割用的声纳图像。

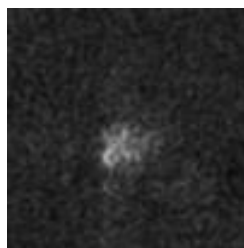


图 3.1 阈值分割用声纳图像

灰度级为 $[0, L-1]$ 范围的数字图像的直方图是离散函数 $h(r_k) = n_k$, 这里 r_k 是第 k 级灰度, n_k 是图像中的灰度级为 r_k 的像素个数。经常以图像中的像素点总数(用 n 表示)来除以它的每一个值得到归一化的直方图。因此,一个归一化的直方图由 $P(r_k) = n_k / n$ 给出, 这里 $k = 0, 1, \dots, L-1$ 。简单的说, $P(r_k)$ 给出了灰度级为 r_k 发生的概率估计值^[41]。

图 3.1 的灰度直方图如下图所示:

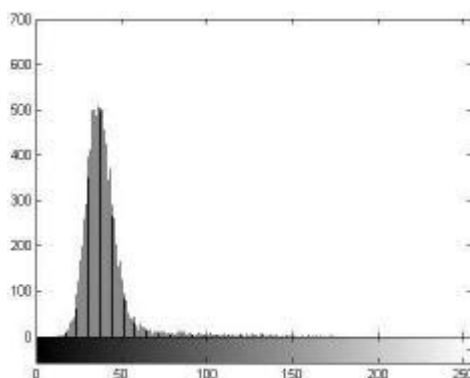


图 3.2 图 3.1 的灰度直方图

由上图可以看出，图 3.1 中的目标亮度特征不明显，背景区域较为丰富。通过人工选取峰值右边像素值 75 进行阈值化的处理，结果如图 3.3 所示：



图 3.3 直方图直接阈值图像

3.1.2 最大类间方差法

最大类间方差法是由日本学者大津(Nobuyuki Otsu)于 1979 年提出的，是一种自适应的阈值确定的方法，又叫大津法，简称 OTSU。它是按图像的灰度特性，将图像分成背景和目标两部分。背景和目标之间的类间方差越大，说明构成图像的两部分的差别越大，当部分目标错分为背景或部分背景错分为目标都会导致两部分差别变小。该方法的基本原理如下：

设图像有 L 个灰度等级，灰度级是 i 的像素个数为 n_i ，则总的像素个数是 $N = \sum_{i=0}^{L-1} n_i$ ，各灰度值出现的概率是 $p_i = n_i / N$ ，显然 $p_i > 0$ 且 $\sum_{i=0}^{L-1} p_i = 1$ 。设以灰度 t 为门限将图像分割成 2 个区域：灰度级为 $0 \sim t$ 的像素区域 A(背景)和灰度级为 $t+1 \sim L-1$ 的像素区域 B(目标)。A、B 出现的概率分别为：

$$P_A = \sum_{i=0}^t p_i, P_B = \sum_{i=t+1}^{L-1} p_i = 1 - P_A \quad (3-1)$$

A、B 两类的灰度均值分别为：

$$\omega_A = \left(\sum_{i=0}^t i p_i \right) / P_A, \omega_B = \left(\sum_{i=t+1}^{L-1} i p_i \right) / P_B \quad (3-2)$$

图像总的灰度均值为：

$$\omega_0 = P_A \omega_A + P_B \omega_B = \sum_0^{L-1} i P_i \quad (3-3)$$

由此可以得到 A、B 两区域的类间方差：

$$\sigma^2 = P_A (\omega_A - \omega_0)^2 + P_B (\omega_B - \omega_0)^2 \quad (3-4)$$

显然， P_A 、 P_B 、 ω_A 、 ω_B 、 ω_0 、 σ^2 都是关于灰度级的函数。为了得到最优分割阈值，Otsu 把两类的类间方差作为判别准则，认为使得 σ^2 值最大的 t^* 即为所求的最佳阈值：

$$t^* = \text{Arg Max}_{0 \leq t \leq L-1} [P_A (\omega_A - \omega_0)^2 + P_B (\omega_B - \omega_0)^2] \quad (3-5)$$

因此，使类间方差最大的分割意味着错分概率最小。

使用最大类间方差法分割图像的效果图如下图所示：



图 3.4 最大类间方差法分割图像 ($t^* = 68$)

3.1.3 最大熵法

图像熵是一种特征的统计形式，它反映了图像中平均信息量的多少。图像的一维熵表示图像中灰度分布的聚集特征所包含的信息量，令 p_i 表示图像中灰度值为 i 的像素所占的比例，则定义灰度图像的一元灰度熵为：

$$H = \sum_0^{255} -p_i \log p_i \quad (3-6)$$

其中 p_i 是某个灰度在该图像中出现的概率，可由灰度直方图获得。

图像的一维熵可以表示图像灰度分布的聚集特征，却不能反映图像灰度分布的空间特征，为了表征这种空间特征，可以在一维熵的基础上引入能够反映灰度分布空间特征的特征量来组成图像的二维熵。选择图像的邻域灰度均值作为灰度分布的空间特征量，与图像的像素灰度组成特征二元组，记为 (i, j) ，其中 i 表示像素的灰度值 ($0 \leq i \leq 255$)， j 表示邻域灰度均值 ($0 \leq j \leq 255$)：

$$p_{i,j} = f(i, j) / N^2 \quad (3-7)$$

上式能反应某像素位置上的灰度值与其周围像素灰度分布的综合特征，其中 $f(i, j)$ 为特征二元组 (i, j) 出现的频数， N 为图像的尺度，定义离散的图像二维熵为：

$$H = \sum_{i,j}^{255} -p_{ij} \log p_{ij} \quad (3-8)$$

构造的图像二维熵可以在图像所包含信息量的前提下，突出反映图像中像素位置的灰度信息和像素邻域内灰度分布的综合特征。

使用最大熵法分割图像的方步骤是先用阈值 t 将图像分割成两部分，分别计算两部分的一元灰度熵，再求和，得到熵 H 最大时的 t 值后，直接用 t 进行阈值分割即可。效果图如下图所示：



图 3.5 最大熵法分割图像效果图（ $t=68$ ）

3.2 基于聚类的声纳图像分割

聚类方法是将物理或抽象对象的集合组成为由类似的对象组成的多个类的过程成为聚类。聚类算法为图像分割中常用的算法之一，其本质为依据一种相似性测度，将大量数据中相似的数据聚集在一起。其中常用的有 K 均值算法、FCM 算法和谱聚类方法等。K-means 算法是硬聚类算法的典型代表，用欧式距离作为相似性测度，属于非监督学习方法^[42]；而 FCM 算法作为一种软聚类算法，是应用得最广泛的模糊聚类算法之一^[43]；然后就是近几年研究的热点之一谱聚类方法，它以谱图理论为基础，将聚类问题抽象成一个对无向图顶点的划分问题^[44]，其在水下声纳图像分割方面也值得较为深入地研究，但是由于其算法复杂，计算速度慢等特点，本文将不对其进行单独研究。

3.2.1 K 均值聚类

J.B.MacQueen 在 1967 年提出的 K-means 算法，是一种被广泛应用于科学研究和工业应用中的经典聚类算法^[45]。K-means 分割算法是一种基于距离聚类的分割方法，它是以欧式距离作为相似性测度，具体到图像，就是以每个像素点的灰度值的差，作为分类的标准。

K-means 算法的核心思想是把 n 个数据对象划分为 k 个聚类，使每个聚类中的数据点到该聚类中心的平方和最小，算法处理过程：

输入：聚类个数 k ，包含 n 个数据对象的数据集。

输出： k 个聚类。

- (1)从 n 个数据对象中任意选取 k 个对象作为初始的聚类中心。
- (2)分别计算每个对象到各个聚类中心的距离,把对象分配到距离最近的类中。
- (3)所有对象分配完成后,重新计算 k 个聚类的中心,将各个类的灰度求均值,得到的均值作为这一类新的聚类中心。
- (4)与前一次计算得到的 k 个聚类中心比较,如果聚类中心发生变化,转(2),否则转(5)。
- (5)输出聚类结果。

K 均值聚类的流程图如下图所示:

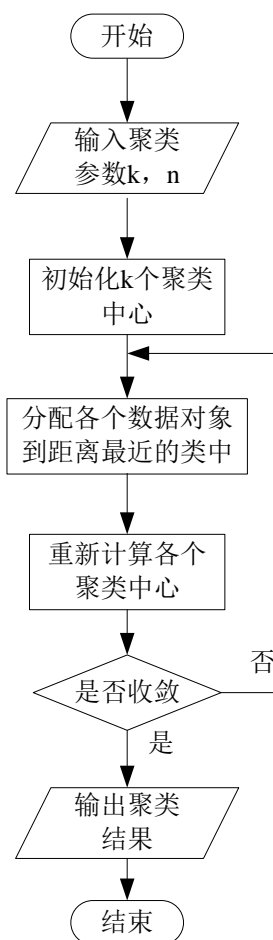


图 3.6 K 均值聚类流程图

首先从 n 个数据对象中任意选择 k 个对象作为初始聚类中心;而对于所剩下的其它对象,则根据他们与这些聚类中心的相似度(距离),分别将他们分配给与其最相似的(聚类中心所代表的)类中。然后再计算每个所新聚类的聚类中心(该聚类中所有对象的均值)。不断重复这一过程直到标准测度函数开始收敛为止。一般都采用均方差作为标准测度函数,具体定义如下:

$$E = \sum_{i=1}^k \sum_{p \in C_i} |p - m_i|^2 \quad (3-9)$$

其中 E 为数据库中所有对象的均方差之和； p 代表对象的空间中的一个点； m_i 为聚类 C_i 的均值。

3.2.2 FCM 聚类

模糊 C 均值聚类算法(FCM)，也就是模糊 ISODATA，是用隶属度确定每个数据点属于某个聚类的程度的一种聚类算法。1973 年，Bezdek 提出了该算法，作为早期 K-means 聚类方法的一种改进而被广泛试用。

给定数据集： $X = \{x_1, x_2, \dots, x_n\}$ ，其中每个样本包含 S 个属性。模糊聚类就是将 X 划分为 C 类 ($2 \leq C \leq n$)。 $V = \{v_1, v_2, \dots, v_c\}$ 是 C 个聚类中心，在模糊划分中，每一个样本不能严格的划分为某一个类，而是以一定的隶属度属于某一个类。

令 U_{ik} 表示第 k 个样本属于第 i 类的隶属度，其中 $0 \leq U_{ik} \leq 1$ ， $\sum_{i=1}^C U_{ik} = 1$ 。

模糊 C 均值聚类的准则如下：

首先定义目标函数：

$$J(U, V) = \sum_{k=1}^n \sum_{i=1}^C (U_{ik})^m (d_{ik})^2 \quad (3-10)$$

其中， $d_{ik} = \|x_k - v_i\|$ 为第 i 个聚类中心与第 k 个数据序列间的欧几里德距离。上式表示各类样本到聚类中心的加权距离平方和，权重是 x_k 对第 i 类的隶属度的 m 次方。且 $m \in [1, \infty)$ 是一个加权指数。

聚类准则为求 $J(U, V)$ 的极小值： $\min\{J(U, V)\}$ 。

算法步骤如下：

(1) 取定 C, m 和初始隶属度矩阵 U^0 ，迭代步数 $l=0$ ；

(2) 计算聚类中心 V ：

$$V_i^{(l)} = ((\sum_{k=1}^n (U_{ik}^{(l)})^m x_k) / (\sum_{k=1}^n (U_{ik}^{(l)})^m)) \quad (3-11)$$

$$i = \{1, 2, \dots, C\}$$

(3) 修正 U ：

$$U_{ik}^{(l+1)} = \frac{1}{\sum_{j=1}^C \left(\frac{d_{ik}}{d_{jk}} \right)^{\frac{2}{m-1}}}, \forall i, \forall k \quad (3-12)$$

(4) 对给定的 $\varepsilon > 0$ ，实际计算时应对取定的初始值进行迭代计算直至：

$$\max\{|U_{ik}^t - U_{ik}^{t-1}|\} < \varepsilon \quad (3-13)$$

则算法终止，否则 $l=l+1$ ，且迭代步数没有结束，转向(2)。

结果处理：

若 $U_{jk} = \max\{U_{ik}\} > 0.5$ ，则 $x_k \in j$ 。

3.2.3 聚类实验与分析

在对聚类算法介绍的基础上，本文对如图 3.7 的沉船的水下声纳目标图像进行聚类分割实验与效果分析：

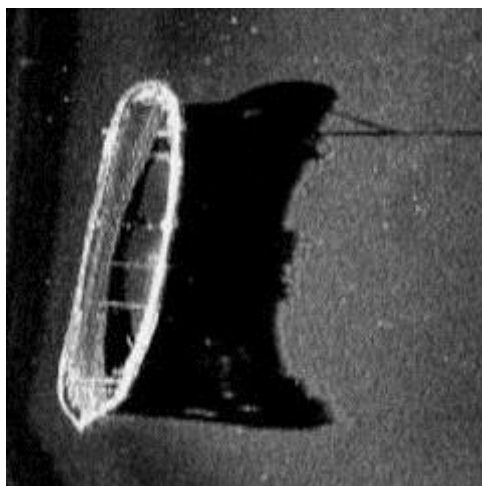


图 3.7 聚类实验用声纳图像

图 3.8 和图 3.9 分别为 K 均值聚类和 FCM 聚类后的效果图(聚为 3 类)。

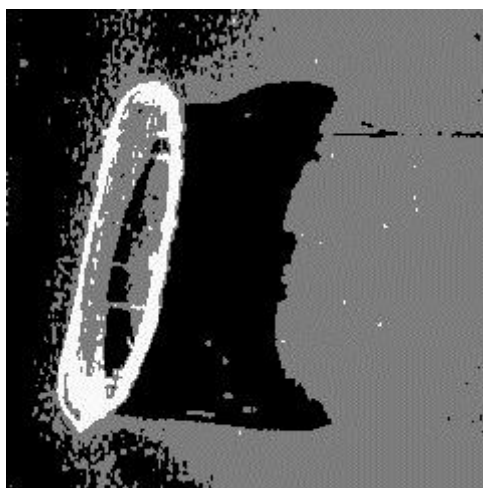


图 3.8 K 均值聚类图像

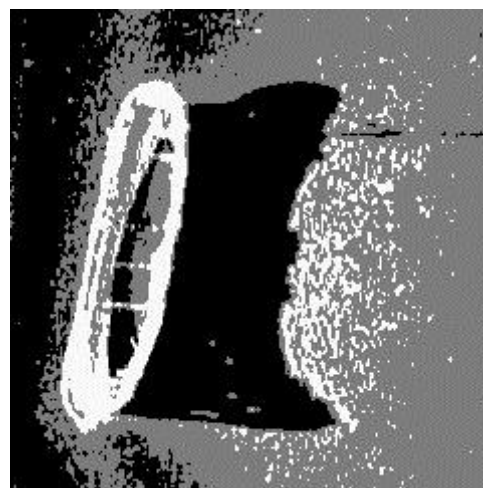


图 3.9 FCM 聚类图像

时间性能比较表如下：

表 3.1 两种聚类方法所用时间表

聚类方法	所用时间
K 均值	0.067s
FCM 聚类	2.85s

从上表中可以看出 K 均值聚类算法的速度远高于 FCM 聚类算法，实时性能较好。

3.3 基于 MRF 的非监督声纳图像分割方法

马尔可夫随机场 (Markov Random Field, MRF) 是声纳图像分割中的一个有效且常用的工具, 基于 MRF 的水下声纳图像检测方法能够准确地描述每个像素所属类别与周围像素类别之间的重要依赖关系, 是一种利用图像中像素间的空间相关性进行目标检测的方法^[46]。但现有的理论和算法都需人为地确定声纳图像要分类的类别个数, 还没有一种完全自动的声纳图像分割模型, 现有算法的复杂度较高, 耗费时间长, 不利于实时性的声纳图像处理与目标识别的要求。

针对声纳图像的特点, 本文设计了一种新的基于 MRF 的非监督声纳图像自动分割方法。在声纳图像的海底混响区服从 Gamma 分布的情况下, 通过一种快速有效的高斯金字塔模型对声纳图像进行预处理, 使得处理后的声纳图像的海底混响区服从高斯分布。并在此基础上构造了一个能够自动确定声纳图像分类个数的模型, 并通过该模型结合一种局部能量极化的方法对马尔科夫模型的初始化参数进行估计, 从而形成一种完全自动的声纳图像分割模型。最后, 利用该模型对声纳图像数据进行了分割实验, 并和其他典型的分割算法进行了比较, 验证了该方法的有效性及其快速性。

其分割方法流程图见下图所示:

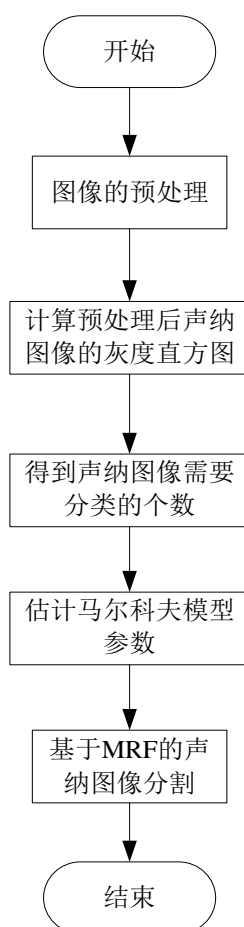


图 3.10 基于 MRF 的非监督声纳图像分割流程图

3.3.1 声纳图像的直方图分析

(一) 声纳图像的直方图处理

如下图 3.11 为原始人造声纳图像，图 3.12 为图 3.11 的直方图。

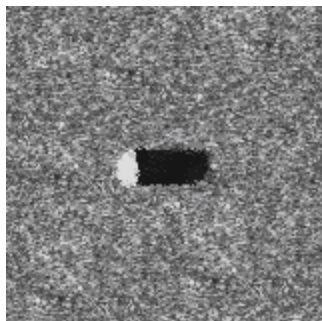


图 3.11 原始人造声纳图像

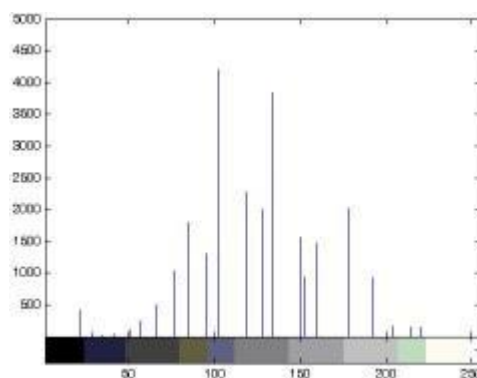


图 3.12 原始人造声纳图像的直方图

(二) 图像的高斯金字塔模型

图像处理的金子塔方法是将原始图像分解成不同空间分辨率的子图像，高分辨率(尺度大)子图像放在下层，低分辨率(尺度小)的图像放在上层，从而形成一个金字塔的形状^[46]，如下图 3.13 所示：

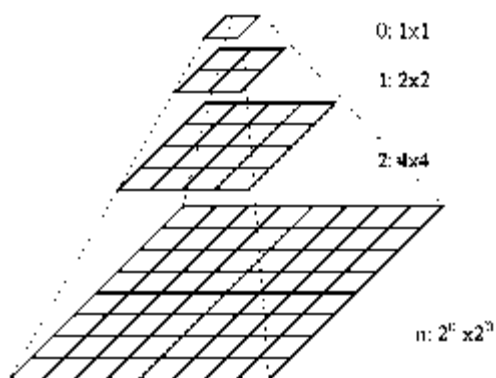


图 3.13 图像处理的金子塔模型

图像的高斯金字塔模型算法：

- (1) 先对图像 $f_l(i, j)$ 进行高斯卷积，下标 l 表示金字塔的层数，再对图像进行降采

样：

$$f_l'(i, j) = g_\sigma * f_l(i, j) \quad (i, j = 0, 1, 2, \dots, n) \quad (3-14)$$

$$f_{l-1}(i, j) = f_l'(2i, 2j) \quad (i, j = 0, 1, 2, \dots, \frac{n}{2}) \quad (3-15)$$

(2) 先对 (1) 处理后的图像进行上采样，再对图像进行高斯卷积：

$$f_l'(i, j) = f_{l-1}(\frac{i}{2}, \frac{j}{2}) \quad (i, j = 0, 1, 2, \dots, n) \quad (3-16)$$

$$g_l(i, j) = f_l'(i, j) * g_\sigma \quad (i, j = 0, 1, 2, \dots, n) \quad (3-17)$$

算法结束。

其中高斯卷积核为：

$$g_\sigma = \begin{bmatrix} \frac{1}{256} & \frac{1}{64} & \frac{3}{128} & \frac{1}{64} & \frac{1}{256} \\ \frac{1}{64} & \frac{1}{16} & \frac{3}{32} & \frac{1}{16} & \frac{1}{64} \\ \frac{3}{128} & \frac{3}{32} & \frac{9}{64} & \frac{3}{32} & \frac{3}{128} \\ \frac{1}{64} & \frac{1}{16} & \frac{3}{32} & \frac{1}{16} & \frac{1}{64} \\ \frac{1}{256} & \frac{1}{64} & \frac{3}{128} & \frac{1}{64} & \frac{1}{256} \end{bmatrix} \quad (3-18)$$

对原始的声纳图像进行高斯金字塔处理后的图像模型及其直方图如下图所示。

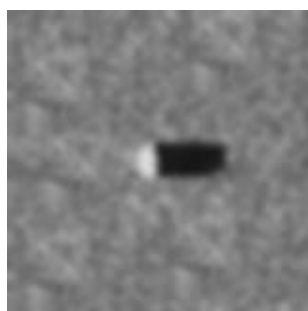


图 3.14 原始人造声纳图像的高斯金字塔处理结果

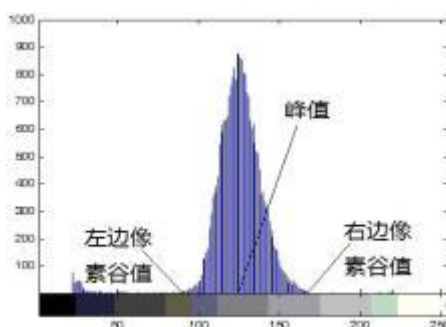


图 3.15 图 3.14 的灰度直方图

由图 3.15 可以看出, 经高斯金字塔处理后的声纳灰度直方图满足背景区域符合高斯分布的特点, 得知该滤波方法结合简单的阈值处理就能把目标、阴影以及背景分割出来, 而且算法简单, 运算速度快。表 3.2 给出了应用于本节设计的基于 MRF 的非监督声纳图像分割方法中使用高斯金字塔滤波方法的优点和不足:

表 3.2 声纳图像高斯金字塔滤波优缺点

	优点	缺点
高斯金字塔滤波方法	对噪声的平滑效果好, 算法简单, 结合简单的阈值处理就能把目标、阴影以及背景区域分割出来	适用于背景相对较为简单的声纳目标图像

本节所设计的基于 MRF 的非监督声纳图像分割方法也正是在此滤波方法的基础上进行改进得到的, 自动分割效果较为理想。

3.3.2 图像分割中的 Markov 随机场模型及算法

(一) 马尔科夫随机场模型

最大后验概率(MAP)是图像处理中最常用的最优化准则, 也是 MRF 建模中最常用的最优化准则。MRF 模型与 MAP 准则结合在一起就称 MAP-MRF 体系^[47]。

假定我们观测到的图像数据为 F , 图像上所有像素点的集合记为 S 。图像的分割问题是要求解的问题是满足最大后验概率准则, 对每个像素的分类标号(标号场), 记为 ω 。这样由 Bayes 后验概率准则:

$$P(\omega | F) = \frac{1}{P(F)} P(\omega) P(F | \omega) \quad (3-19)$$

其中 $P(F)$ 为观测数据的先验分布, 当数据给定后为常数, 所以不参与计算过程, 可以不予以考虑。

$P(\omega)$ 是标号场的先验的联合的 Gibbs 分布, 即满足马尔可夫性。假设 C 表示 S 所有的集簇, c 表示 C 中的元素, $U_2(\omega)$ 为能量函数, $V_c(\omega)$ 是与集簇相关的势函数, 那么:

$$P(\omega) \propto \exp(U_2(\omega)) = \exp\left(\sum_{c \in C} -V_c(\omega_c)\right) \quad (3-20)$$

$P(F | \omega)$ 是似然概率, 在很多情况下, 假定为各个位置的像素是独立同分布的, 即满足:

$$P(F | \omega) = \prod_{s \in S} P(F_s | \omega_s) \quad (3-21)$$

当我们假定每个 $P(F_s | \omega_s)$ 是高斯分布时, 每个类的类参数都是由两个参数唯一确

定该分布，即为 λ 和 σ 。将 Bayes 后验概率准则：

$$P(\omega|F) = \frac{1}{P(F)} P(\omega) P(F|\omega) \text{ 取对数，得到的目标函数为 } \ln P(\omega) + \ln P(F|\omega), \text{ 即我}$$

们要求的是使该表达式最大的时候 ω 的估计 $\hat{\omega}$ ，将似然函数和先验 Gibbs 分布的表达式带入，这时我们可以形成 MRF-MAP 下的目标函数的最优解问题：

$$\hat{\omega} = \arg \max_{\omega} \prod_{s \in S} P(F_s | \omega_s) \prod_{c \in C} \exp(-V_c(\omega_c)) \quad (3-22)$$

则我们根据声纳图像的性质，选取一阶的邻域系统，并且势函数为 ISING 模型的势函数，一阶邻域系统和势函数如图 3.16 所示：

势函数：

$$V(\omega_c) = V_{\{s,r\}}(\omega_c, \omega_r) = \begin{cases} -\beta & \text{if } \omega_c = \omega_r \\ +\beta & \text{if } \omega_c \neq \omega_r \end{cases} \quad (3-23)$$

其中 β 为耦合系数。

对每个像素点选取 $P(F_s | \omega_s)$ 服从高斯分布，那么取对数后其函数形式为：

$$\begin{aligned} \ln P(F|\omega) &= U_1(\omega, F) \\ &= \sum_{s \in S} \left(\ln(\sqrt{2\pi}\sigma_{\omega_s}) + \frac{(f_s - \mu_{\omega_s})^2}{2\sigma_{\omega_s}^2} \right) \end{aligned} \quad (3-24)$$

取 $\ln P(\omega) + \ln P(F|\omega)$ 为目标函数，得到的分割结果为：

$$\text{Arg}(\max_{\omega} (U_1(\omega, F) + U_2(\omega))) \quad (3-25)$$

其中 $U_2(\omega)$ 为能量函数，且

$$U_2(\omega) = \exp\left(\sum_{c \in C} -V_c(\omega_c)\right) \quad (3-26)$$

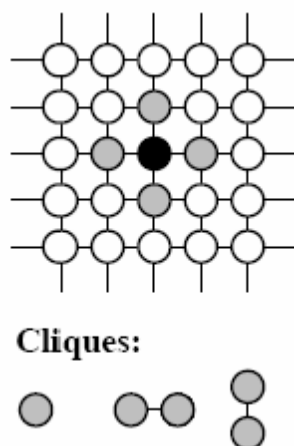


图 3.16 平面上的一阶邻域系统

(二) 条件迭代算法(ICE)

条件迭代算法是典型的确定松弛算法。其用于图像分割的算法流程如下：

(1) 通过训练样本得到所需似然函数 $P(F|\omega)$ 参数集合，即对应不同分类 λ 情况有

$$\mu_\lambda = \frac{1}{|S_\lambda|} \sum_{s \in S_\lambda} f_s \quad (3-27)$$

$$\sigma_\lambda = \sqrt{\frac{1}{|S_\lambda|} \sum_{s \in S_\lambda} (f_s - \mu_\lambda)^2} \quad (3-28)$$

初始化势函数中的耦合系数 β ，经验值是取区间 $[0.5, 1]$ 之间的数值。

(2) 依据似然概率即 $P(F|\omega)$ 最大化的准则选取初始的标记场 ω_0 ，即对每一个像素点 s 取 $\text{Arg}(\max_{\omega_s} U_1(\omega_s, F))$ ，遍历整个图像得到整个图像的初始分割 ω_0 ；

(3) 根据目标函数计算当前分割结果：取 k 为当前的迭代次数，每一个像素点 s 取

$$\text{Arg}(\max_{\omega_s} U_1(\omega_s, F) + U_2(\omega_s)) \quad (3-29)$$

遍历整个图像得到标记场 ω_k ；

(4) 判断收敛条件：以每次迭代过程中全局能量的变化量为收敛条件，计算当前全局能量的值：

$$E_k = \sum_{s \in S} U_1(\omega_s, F) + \frac{U_2(\omega_s)}{2}$$

(3-30)

如果 $\Delta \geq E_k - E_{k-1}$ 认为全局能量变化很小，标记场 ω_k 为最后的分割结果，式中 Δ 为常数，即预先设定的能量改变量的阈值，得到了 ICM 算法的分割结果；否则取 $k = k + 1$ ，转到步骤 (2)。

3.3.3 非监督的声纳图像分割模型

(一) 声纳图像自动分类的模型与算法

本文构造了一种能够自动确定声纳图像分类及分类个数的模型如下：

$$\begin{cases} \Gamma(y_i, n) = \hat{f}(p(r_k)) \\ i = 1, 2, 3 \\ k = 0, 1, \dots, L-1 \end{cases} \quad (3-31)$$

其中， $p(r_k)$ 为图像经过高斯金字塔处理后模型的归一化统计直方图。 y_1 为图像是否含有阴影区的判别函数， $y_1 \in \{0, 1\}$ ； y_2 为图像是否含有背景区的判别函数， $y_2 \in \{0, 1\}$ ； y_3 为图像是否含有目标区的判别函数， $y_3 \in \{0, 1\}$ 。 n 为图像分类的个数，此处声纳图像 $n \in \{1, 2, 3\}$ 。

如图 3.15 高斯金字塔处理后的图像直方图所示，定义像素峰值 r_{pv} 为图像统计直方图取得极大值时的灰度级；左边像素谷值 r_{vvl} 为图像统计直方图从像素峰值 r_{pv} 开始向左统计直方图变化较小时的灰度级；同理，右边像素谷值 r_{vvr} 为图像统计直方图从像素峰值 r_{pv} 开始向右统计直方图变化较小时的灰度级。

并定义判别函数：

$$q(x, \theta) = \begin{cases} 0, & x \leq \theta \\ 1, & x > \theta \end{cases} \quad (3-32)$$

自动分类模型算法的实现：

(1) 设定迭代条件 $iter = 0.01$ ，计数个数 $count = 0$ ，迭代步数 $l = 0$ ；

(2) 计算像素峰值：

$$r_{pv} = \arg \max_{r_k} p(r_k) \quad k = 0, 1, \dots, L-1 \quad (3-33)$$

并计算灰度级为 r_{pv} 时的概率： $p^l = p(r_{pv})$ ；

(3) 计算灰度区间 $[r_{pv} - l, r_{pv} + l]$ 的概率统计和：

$p^{l+1} = p([r_{pv} - l, r_{pv} + l])$ ；若 $|p^{l+1} - p^l| < iter$ 则 $count++$ ；若 $count$ 大于给定一整数 m ，则转 (4)，否则， $l = l + 1$ ，并转 (3)；

(4) 计算左右边像素谷值：

$$\begin{cases} r_{vvl} = r_{pv} - l \\ r_{vvr} = r_{pv} + l \end{cases}$$

(3-34)

若 $p([r_{vvl}, r_{vvr}]) > 1 - iter$ ，则令 $\begin{cases} y_1 = 0 \\ y_2 = 1 \\ y_3 = 0 \\ n = 1 \end{cases}$ ，算法终止；

(5) 计算图像的判别函数 y_i 及类别个数 n ：

$$\begin{cases} y_1 = q(\sum_{k=0}^{r_{vvl}} p(r_k), \frac{1 - p([r_{vvl}, r_{vvr}])}{10}) \\ y_2 = q(\sum_{k=r_{vvl}}^{r_{vvr}} p(r_k), 0.80) \\ y_3 = q(\sum_{k=r_{vvr}}^{L-1} p(r_k), \frac{1 - p([r_{vvl}, r_{vvr}])}{10}) \\ n = \sum_{i=1}^3 y_i \end{cases} \quad (3-35)$$

算法结束。

(二) 局部能量极值化的原理与算法

令 (x, y) 为某一图像中像素的坐标, 令 S_{xy} 表示一确定大小的邻域 (子图像), 其中心在 (x, y) 。则在 S_{xy} 中像素的平均值、能量和方差的计算公式如下[8]:

$$m_{S_{xy}} = \frac{1}{|S_{xy}|} \sum_{(s,t) \in S_{xy}} r_{s,t} \quad (3-36)$$

$$E_{S_{xy}} = \sum_{(s,t) \in S_{xy}} r_{s,t}^2 \quad (3-37)$$

$$\sigma_{S_{xy}}^2 = \frac{1}{|S_{xy}| * (|S_{xy}| - 1)} \left(\sum_{(s,t) \in S_{xy}} r_{s,t}^2 - m_{S_{xy}}^2 \right) \quad (3-38)$$

局部能量极值化也就是把图像分成很多大小相同的区域, 然后分别计算源图像中各个区域的均值、能量和方差。然后根据能量函数的大小, 以及图像分类的个数和判别函数来得到 MRF 分割模型的初始化参数。

其算法如下:

(1) 根据前面的算法得到的图像的高斯金字塔模型 $g_l(i, j)$ 、左右边像素谷值 r_{vvl} 和 r_{vvr} 、以及判别函数 y_i ;

(2) 分别以 r_{vvl} 和 r_{vvr} 为阈值得到金字塔模型的阈值化函数, 即阴影模型 g_{shadow} 、目标模型 g_{target} 和背景模型 g_{back} ;

(3) 根据上面得到的各个模型和判别函数 y_i 可以估计 MRF 模型初始化参数的所选取的区域的长 L 和宽 W ;

(4) 然后根区域能量公式计算各个区域的能量函数, 并将能量最小时的区域均值 μ_1 和均方差 σ_1 作为 MRF 模型的阴影区域参数; 将能量取得极大值时的区域均值 μ_2 和均方差 σ_2 作为 MRF 模型的目标区域参数; 而选取一块既非目标也非阴影的区域的均值 μ_3 和均方差 σ_3 作为 MRF 模型的背景区域参数;

即可得到 MRF 模型的初始化参数均值 μ_λ , 均方差 σ_λ , 其中 $\lambda \in \{1, 2, 3\}$ 为不同的分割区域。

最后, 用得到的初始化参数初始化 MRF 模型, 并利用 ICE 对声纳图像进行分割, 将分割后的模型分别与阴影模型 g_{shadow} 和目标模型 g_{target} 进行简单的与运算与高斯低通滤波等处理, 即可以得到最终的声纳图像分割结果。

3.3.4 实验结果与分析

通过对图 3.11 原始声纳图像和图 3.12 原始声纳图像的直方图以及大量实验数据的分析可以看出, 声纳图像含有比较复杂的海底混响噪声, 因此, 对声纳图像的分割是比较困难的。本文在对声纳图像进行大量实验的基础上, 得到了声纳图像经过适当的预处理之后也是有规律可循的, 在此基础上构造了能够自动确定分类个数以及自动确定

MRF 模型初始化参数的分割方法，不仅分割的效果较好，而且时间复杂度也较小，有利于实时性的声纳图像数据处理。

根据本文设计的算法，对图 3.11 既有目标又有阴影的声纳图像、图 3.17 较为复杂的声纳图像分别进行了实验，验证了本文算法的有效性。其中图 3.11 像素尺寸大小为 160×160 ，图 3.17 像素尺寸大小为 100×100 。



图 3.17 较为复杂的声纳图像

并和其他文献算法的分割结果进行了比较，比如文献[49]利用 FCM 快速聚类算法对 MRF 模型的参数进行初始化，但仍旧需要首先指定聚类的个数，文献[50]同样也需要指定分类的类别数，而且两者的算法的时间复杂度较大，不利于实时性的声纳图像数据处理。

图 3.18 为不同算法对图 3.11 原始声纳图像的分割效果。

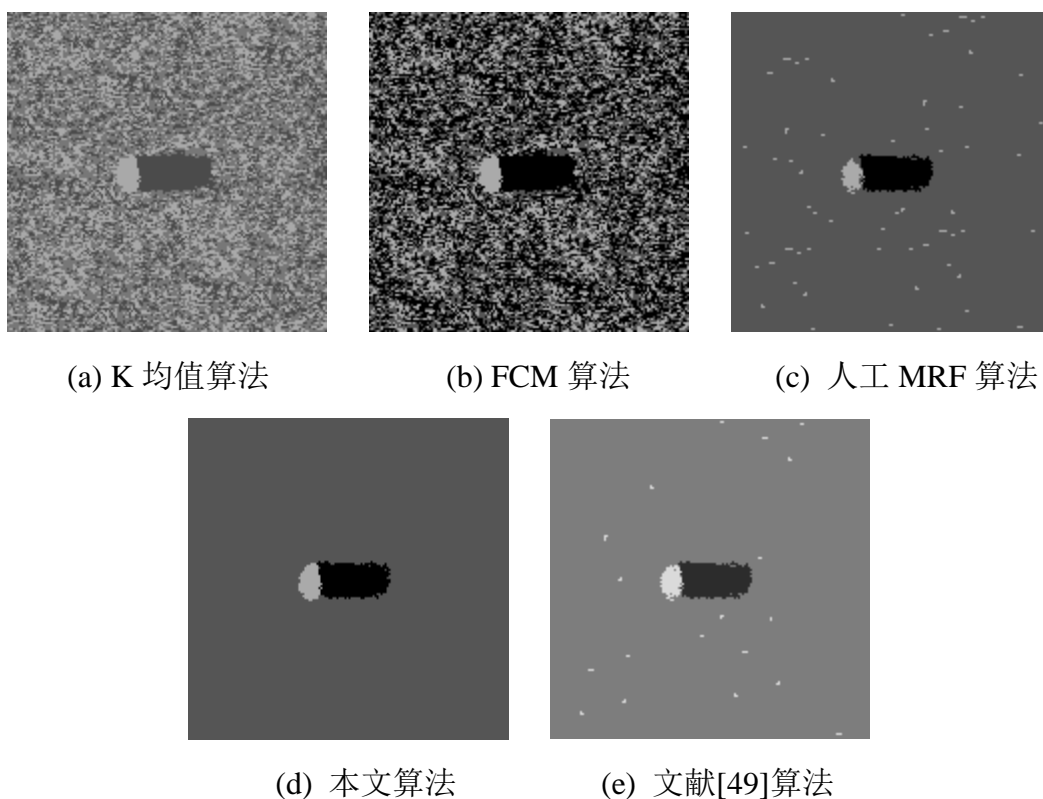


图 3.18 不同算法对图 2 的分割效果

图 3.19 为不同算法对图 3.17 较为复杂的声纳图像进行分割的效果。

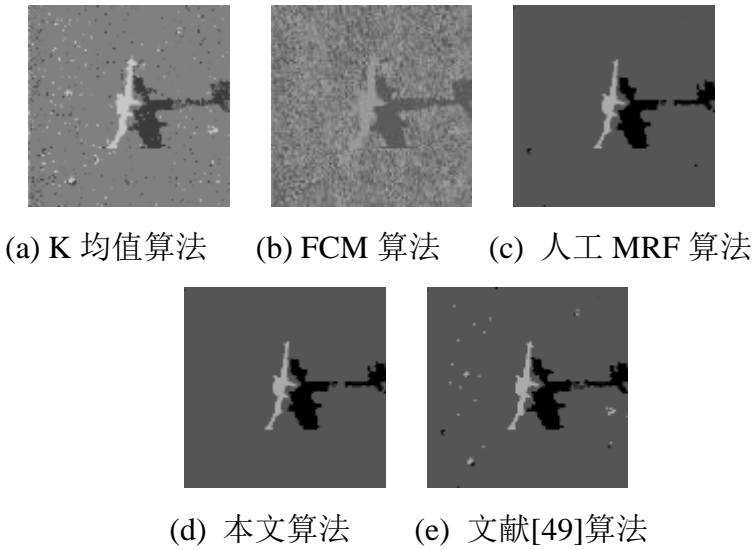


图 3.19 不同算法对图 3.17 的分割效果

图 3.20 为分别对声纳图像图 3.11 和图 3.17 进行分割所用时间的性能分析。由图可以看出，本文所采用的算法所用的时间只比人工确定 MRF 模型参数的分割算法多了 100ms 左右，所用的时间明显较少，算法复杂度较低。表 3.3 为模型方法性能比较表

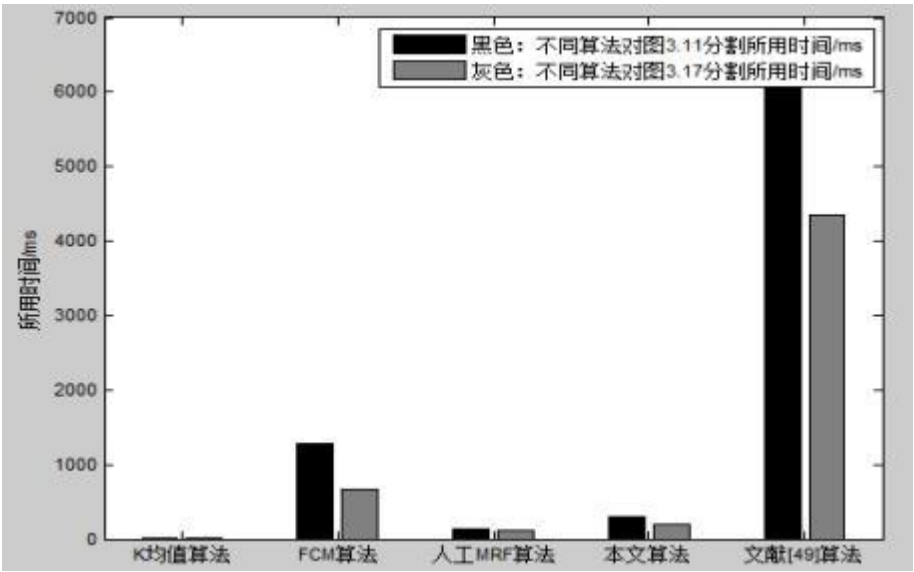


图 3.20 不同算法的时间性能分析

表 3.3 模型方法性能比较表

	K 均值	FCM 算法	人工 MRF	本文算法	文献[49]算法
算法复杂度是否简单	√	×	√	√	×
分类个数是否自动确定	×	×	×	√	×

模型参数是否自动确定	∕	∕	×	√	√
是否是非监督声纳图像分割	×	×	√	√	×

注：×代表“否”；√代表“是”；∕代表“无此项”。

3.4 本章小结

本章主要对基于阈值的声纳图像分割方法、基于聚类的声纳图像分割方法和基于 MRF 的声纳图像分割方法进行了研究与改进。其中，对基于阈值的声纳图像分割方法中的直方图、最大类间方差法和最大熵方法进行了简单的介绍与实验；对基于聚类分析中的 K 均值聚类和 FCM 聚类方法也进行了分析与比较，并得出 K 均值聚类算法的速度远高于 FCM 聚类算法，实时性能较好，具有较高的实用性。

而针对目前声纳图像分割方法中存在的缺点与不足，特别是已有的算法模型大都是事先确定声纳图像要分类的个数，基于 MRF 的声纳图像分割方法也大都是人工的确定 MRF 模型的初始化参数，同时，算法计算也都较为复杂。在本章的第三节中构造了一种能够自动确定声纳图像分类个数的模型，并在此基础上通过一种局部能量极化的方法来得到 MRF 模型的初始化参数，该法算法简单，实现容易，有利于实时性的声纳图像分割与目标识别。

第 4 章 声纳图像的特征提取与分析

特征是一个事物区别于其他事物的根本属性。如果要想从一堆物体中把我们需要的物体准确的寻找到，就要掌握它的特征^[51]。而特征提取是从图像中获取目标的数据信息并进行相关分析的前提条件和关键环节。虽然特征的具体定义往往由问题或者应用类型决定，但目前多数文献认为，特征是一幅图像中“有趣”的部分，能有效描述目标区域或场景，一个算法是否成功往往由它使用和定义的特征决定^[52-54]。声纳图像的成像原理与光学图像有很大的不同，所以在特征提取的时候，与光学中的特征提取技术也有所不同，在声纳图像处理过程中，需要针对声纳图像的具体特征来进行分析和研究，以便进行更为有效的特征分析与提取^[55]。

4.1 常用的声纳图像特征

图像的特征是图像的原始特性或属性，是人们对图像视觉感受的量化描述，从各个方面描述了图像的内在语义，可以作为图像的抽象表示，因此它是图像分析的重要依据。对于一个有效的图像特征来说，通常从以下四个方面对其性能的优劣进行评判^[56]。

(1)可区分性：对于不同类型的图像而言，它们的特征应该具有明显的差异，即应能较好地区分视觉上存在差异的图像。

(2)可靠性：对于同类的对象，特征值应该比较接近。

(3)独立性：所选择的特征之间彼此不相关。

(4)数量少：特征的数量少，这样能保证算法具有较为理想的时间复杂度。对于特征识别来说，算法的复杂度随着特征个数的增加而迅速的增加。

可以根据图像的特征属性对图像特征进行总结和分类，如下图所示：

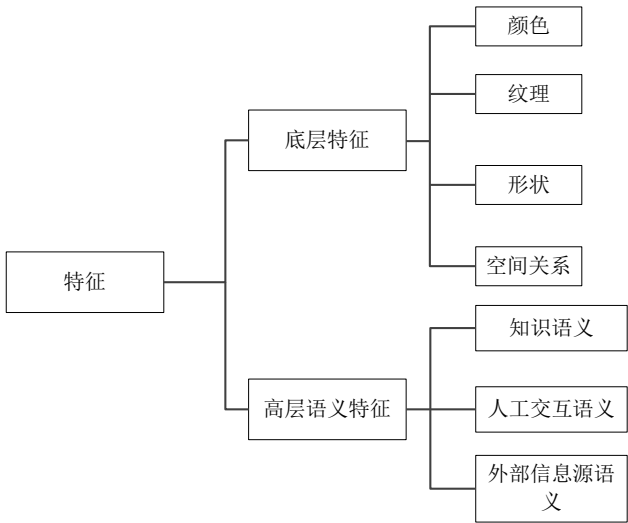


图 4.1 基于特征属性的特征分类

如上图所示的基于特征属性的特征分类图，常用的底层图像特征有颜色特征、纹理特征、形状特征、空间关系特征；和高层语义图像特征：知识语义、人工交互语义和外部信息源语义特征等。

其中，底层图像特征中的颜色特征是一种全局特征，描述了图像或图像区域所对应的景物的表面性质。一般颜色特征是基于像素点的特征，此时所有属于图像或图像区域的像素都有各自的贡献。由于颜色对图像或图像区域的方向、大小等变化不敏感，所以颜色特征不能很好地捕捉图像中对象的局部特征。而声纳图像主要是基于灰度特征的，所以颜色特征并不适用于声纳图像的特征提取与分析。

而图像的空间关系特征，是指图像中分割出来的多个目标之间的相互的空间位置或相对方向关系，以及图像特征中的高层语义特征，同样也不适用于声纳图像的特征提取与分析。所以本节只对声纳图像的纹理特征和形状特征进行分析与实验。

4.1.1 纹理特征

图像的纹理特征也是一种全局特征，它也描述了图像或图像区域所对应景物的表面性质。但由于纹理只是一种物体表面的特性，并不能完全反映出物体的本质属性，所以仅仅利用纹理特征是无法获得高层次图像内容的。作为一种统计特征，纹理特征常具有旋转不变性，并且对于噪声有较强的抵抗能力。但是，纹理特征也有其缺点，一个很明显的缺点是当图像的分辨率变化的时候，所计算出来的纹理可能会有较大偏差。

图像纹理特征在图像分割和识别中得到的广泛的应用，从结构分析法的角度来看，纹理特征分析就是要找出图像中的纹理基元以及基元间随机性，函数性或结构性的相互依赖关系。纹理作为图像的宏观特性，包含了图像的重要信息，而基本不受局部的噪声干扰影响。相对利用特征值来表述纹理区域的统计特性，利用纹理模型提取纹理图像能够更加全面的表述像素间相关性^[57-58]。其中用 GMRF 来反映纹理基元间的随机特性，可以表达像素和像素间的空间关系^[59]；利用 Gabor 小波纹理信息，可以在频域不同尺度、不同方向上提取相关的特征^[60]；而利用 GLCM 特征也能够得到图像不同方向上的纹理特征^[61]。

（一）GMRF 纹理

马尔可夫随机场方法是建立在 MRF 理论模型和 Bayes 估计的基础上，按统计决策理论和估计理论中的最优准则来确定问题的解。其突出的特点是通过适当定义的邻域系统引入结构信息，提供了一种一般而自然的用来表达空间上的相关随机变量之间的相互作用的模型，由此所生成的参数可以描述纹理在不同方向、不同形式的集聚特征，更符合人的感官认识。

当 MRF 邻域的激励噪声满足 Gauss 分布时，可得到一组由空域像素灰度表示的线性方程，称作高斯-马尔可夫（GMRF）随机场模型^[62]。

用 GMRF 随机场模型表达纹理，如果图像 X 的分辨率为 $w \times h$ ，其为灰度图像， N_s 是像素 x_s 的对称邻域。其中 $s = (i, j)$ ， $i \in (0, 1, \dots, w-1)$ ， $j \in (0, 1, \dots, h-1)$ 。

则 x_s 的 l 阶 GMRF 纹理，可以用下式表示。

$$x_s^l = \sum_{r \in N_s} \theta_r^l f_r + \frac{1}{\sqrt{2\pi\sigma_l^2}} \exp\left(\frac{-x_s^2}{2\sigma_l^2}\right) \quad (4-1)$$

其中参数 θ_r^l 是 l 阶邻域的 MRF 模型参数。如果像素 x_s 的对称邻域 N_s 包含 n 个元素，所有的模型参数 θ_r^l 组成 MRF 参数向量 $\theta_l = [\theta_l^1, \theta_l^2, \dots, \theta_l^n]$ 。其一阶邻域到四阶邻域的参数图如下图 4.2 所示：

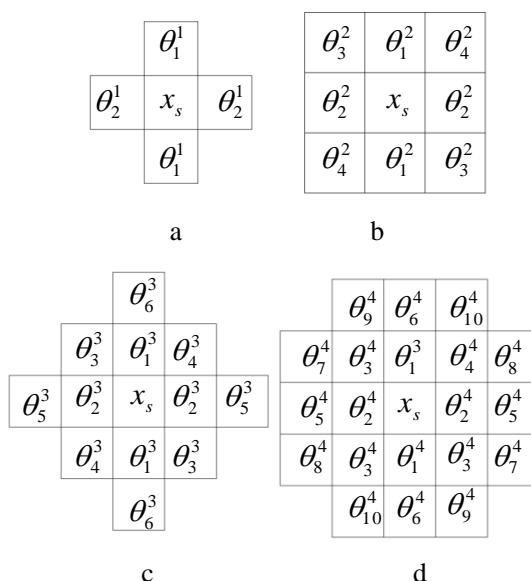


图 4.2 (a)-(d)一阶邻域到四阶邻域的参数图

再者，如果 x_{s+r} 和 x_{s-r} 为 x_s 的一对对称的邻域像素，则定义 $f_r = x_{s+r} + x_{s-r}$ ，其中所有 f_r 组成 $F_s^l = [f_1, f_2, \dots, f_n]$ ，公式(4-1)中的第二项为一个零均值的高斯噪声。

$$\hat{\theta}_l = \left[\sum_{s \in X} F_s^l F_s^{lT} \right]^{-1} \left[\sum_{s \in X} F_s^l x_s \right] \quad (4-2)$$

$$\hat{\sigma}_l = \frac{1}{M^2} \sum_{s \in X} (x_s - \hat{\theta}_l^T F_s^l)^2 \quad (4-3)$$

式(4-2)中的 $\hat{\theta}$ 是对 GMRF 随机场模型参数的渐进一致性估计，式(4-3)给出了参数估计的平方误差 $\hat{\sigma}$ 。当邻域系统的阶数较低时，用线性自回归的高斯-马尔可夫随机场模型描述复杂的图像特征有一定的局限性，随着阶数的增加， $\hat{\theta}$ 能描述更复杂的纹理。下面对如图 4.3 的声纳图像提取一到四阶 GMRF 纹理，得到其一到四阶的纹理图像和 GMRF 随机场模型参数 $\hat{\theta}$ 的估计值。

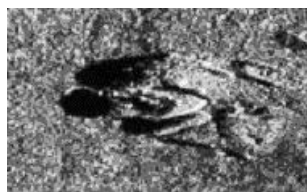


图 4.3 GMRF 纹理实验图



(a) 1 阶



(b) 2 阶



(c) 3 阶



(d) 4 阶

图 4.4 (a)-(d)图 4.3 的一阶邻域到四阶邻域的 GMRF 纹理图像

表 4.1 一阶邻域到四阶邻域的 GMRF 纹理参数表

l	参数					
	θ_l					
1	[0.2638 0.2415]					
2	[0.3982 0.3822 -0.1407 -0.1381]					
3	[0.4336 0.3536 -0.0912 -0.0871 -0.0758 -0.0330]					
4	[0.4058 0.3627 -0.0622 -0.0572 -0.0884 -0.0140 -0.0299 0.0090 0.0149 -0.0415]					

(二) Gabor 纹理

Gabor 纹理变换属于加窗傅立叶变换，Gabor 函数可以在频域不同尺度、不同方向上提取相关的特征，Gabor 变换是短时 Fourier 变换中当窗函数取为高斯函数时的一种特殊情况。另外，Gabor 函数与人眼的生物作用相仿，所以经常用作纹理识别上，并取得了较好的效果^[63]。

Gabor 在方向选择性和最佳的本地化空间和频域域中表现出应有的特点。一般来说，Gabor 小波(内核，过滤器)可以被定义如下^[64-66]：

$$g_{uv}(x, y) = \frac{k^2}{\sigma^2} \exp\left(-\frac{k^2(x^2 + y^2)}{2\sigma^2}\right) [\exp(ik \begin{pmatrix} x \\ y \end{pmatrix}) - \exp(-\frac{\sigma^2}{2})] \quad (4-4)$$

其中：

$$k = \begin{pmatrix} k_x \\ k_y \end{pmatrix} = \begin{pmatrix} k_v \cos \varphi_u \\ k_v \sin \varphi_u \end{pmatrix} \quad (4-5)$$

$$k_v = 2^{\frac{v+2}{2}} \pi \quad (4-6)$$

$$\varphi_u = u \frac{\pi}{K} \quad (4-7)$$

v 的取值决定了 Gabor 核函数的波长, u 的取值表示 Gabor 核函数的方向, K 表示总的方向数。参数 σ/k 决定了高斯窗口的大小, 这里取 $\sigma = \sqrt{2}\pi$ 。在对图 4.5 的声纳图像提取纹理图像时, 分别选取 $0, \pi/4, \pi/2, 3\pi/4$ 四个方向, 获得纹理图像如下图所示:



图 4.5 Gabor 纹理实验图

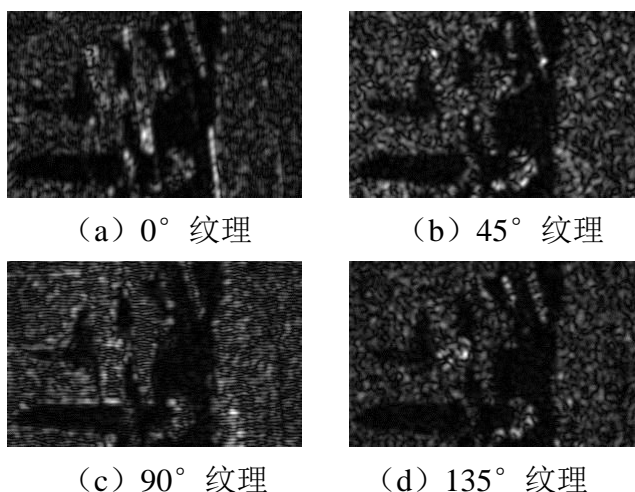


图 4.6 (a)-(d)图 4.5 的四个方向的 Gabor 纹理图像

(三) GLCM 纹理

由于纹理是由灰度分布在空间位置上反复出现而形成的, 因而在图像空间中相隔某距离的两像素之间会存在一定的灰度关系, 即图像中灰度的空间相关特性。灰度共生矩阵(GLCM)就是一种通过研究灰度的空间相关特性来描述纹理的常用方法^[67-70]。

取图像($N \times N$)中任意一点(x, y)及偏离它的另一点($x+a, y+b$), 设该点对的灰度值为($g1, g2$)。令点(x, y)在整个画面上移动, 则会得到各种($g1, g2$)值, 设灰度值的级数为 k , 则($g1, g2$)的组合共有 k 的平方种。对于整个画面, 统计出每一种($g1, g2$)值出现的次数, 然后排列成一个方阵, 再用($g1, g2$)出现的总次数将它们归一化为出现的概率 $P(g1, g2)$, 这样的方阵称为灰度共生矩阵。

距离差分值(a, b)取不同的数值组合, 可以得到不同情况下的联合概率矩阵。(a, b)

取值要根据纹理周期分布的特性来选择,对于较细的纹理,选取(1,0)、(1,1)、(2,0)等小的差分值。当 $a=1, b=0$ 时,像素对是水平的,即0度扫描;当 $a=0, b=1$ 时,像素对是垂直的,即90度扫描;当 $a=1, b=1$ 时,像素对是右对角线的,即45度扫描;当 $a=-1, b=1$ 时,像素对是左对角线,即135度扫描。这样,两个像素灰度级同时发生的概率,就将 (x, y) 的空间坐标转化为“灰度对” (g_1, g_2) 的描述,形成了灰度共生矩阵。

常见的特征有能量、对比度、相关性、熵和逆差矩等。其中能量表示为:

$$ASM = \sum_{i=1}^k \sum_{j=1}^k (G(i, j))^2 \quad (4-8)$$

图像能量是灰度共生矩阵元素值的平方和,所以也称能量,反映了图像的灰度分布均匀程度和纹理粗细度。

对比度表示为:

$$CON = \sum_{n=0}^{k-1} n^2 \{ \sum_{|i-j|=n} G(i, j) \} \quad (4-9)$$

对比度 CON 反映了图像的清晰度和纹理沟纹深浅的程度。纹理沟纹越深,其对比度越大,视觉效果越清晰;反之,对比度小,则沟纹浅,效果模糊。

相关度表示为:

$$IDM = \sum_{i=1}^k \sum_{j=1}^k \frac{G(i, j)}{1 + (i - j)^2} \quad (4-10)$$

相关度 IDM 的大小反映了图像中局部灰度相关性。当矩阵元素值均匀相等时,相关值就大;相反,如果矩阵像元值相差很大则相关值小。

熵表示为:

$$ENT = - \sum_{i=1}^k \sum_{j=1}^k G(i, j) \log G(i, j) \quad (4-11)$$

熵 ENT 表示了图像中纹理的非均匀程度或复杂程度。

自相关表示为:

$$COR = \sum_{i=1}^k \sum_{j=1}^k \frac{(ij)G(i, j) - u_i u_j}{s_i s_j} \quad (4-12)$$

其中:

$$\begin{aligned}
u_i &= \sum_{i=1}^k \sum_{j=1}^k i \cdot G(i, j) \\
u_j &= \sum_{i=1}^k \sum_{j=1}^k j \cdot G(i, j) \\
s_i^2 &= \sum_{i=1}^k \sum_{j=1}^k (i - u_i)^2 \cdot G(i, j) \\
s_j^2 &= \sum_{i=1}^k \sum_{j=1}^k (j - u_j)^2 \cdot G(i, j)
\end{aligned} \tag{4-13}$$

自相关 *COR* 反应了图像纹理的一致性。

下面对图 4.5 提取 GLCM 纹理特征如下表 4.2 所示。

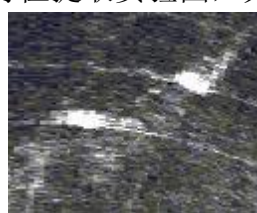
表 4.2 GLCM 纹理特征

方向	相关	能量	熵	对比度	逆差矩
0°	0.040710	0.190083	2.446943	3.875673	0.582619
45°	0.021508	0.193977	2.403110	4.001486	0.571866
90°	0.338188	0.206375	2.384692	2.097662	0.646219
135°	-0.045806	0.189026	2.421924	4.297231	0.562004
均值	0.088650	0.194865	2.414167	3.568013	0.590677
方差	0.021788	0.000048	0.000531	0.744061	0.001081

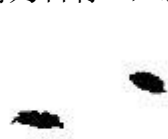
4.1.2 形状特征

在代表目标的众多参数中，目标的形状是描述目标的重要参数之一。目标的形状就是目标的外轮廓，或者可以被看作是一个代表目标范围的二值图像^[71]。面积、周长、长短轴、主轴方向、紧密度、实心度、偏心率等广泛应用的全局几何特征量，以及目标形状的矩不变量特征等^[72]。下面分别对其进行介绍。

图 4.7 为声纳形状特征提取实验图，其中右上角为目标 1，左下角为目标 2。



(a)原始声纳图像



(b)分割后图像

图 4.7 (a)-(b)形状特征提取用声纳图像

(一) 面积

也即区域面积，区域内的像素数。可以首先对图像进行标记，划分出不同的连通区域，在标记图上的每一个物体的像素值就是该物体的标号。求各种不同标号的总和，即求得不同区域的面积数。比如图 4.7 中目标 1 的像素值可以置为 1，目标 2 的像素值可以置为 2。设图像 $f(x, y)$ 的大小为 $M * N$ ，则目标物的面积为：

$$Area = \sum_{y=1}^N \sum_{x=1}^M f(x, y) \quad (4-14)$$

对属于目标区域的像素求和，即利用像素统计的方法，来计算目标的区域面积。提取目标面积如下表：

表 4.3 区域面积

	目标 1	目标 2
<i>Area</i>	361	432

（二）周长

对于内部无孔区域的边界，定义为其内部边界像素的集合。区域 R 的边界 B 是由 R 的所有边界点按照 4-连通方向或 8-连通方向连接组成的，区域 R 的每一个边界点 P 都应该满足两个条件，其一是 P 本身属于区域 R ；其二是 P 具有该区域 R 外的邻点。4-连通域边界和 8-连通域边界可以分别用以下公式来表示。

$$B_4 = \{(x, y) \in R | N_4(x, y) - R \neq 0\} \quad (4-15)$$

$$B_8 = \{(x, y) \in R | N_8(x, y) - R \neq 0\} \quad (4-16)$$

其中 x, y 表示像素点在图像中的坐标。判断边界点后，即可由此计算出目标边界的周长 B 。提取目标边界周长如下表：

表 4.4 区域周长

	目标 1	目标 2
<i>B</i>	83	109

（三）形状参数

区域的形状参数是根据区域的周长和面积计算得出的，定义如下：

$$F = \frac{\|B\|^2}{4\pi Area} \quad (4-17)$$

形状参数在一定程度上描述了区域的紧凑性，它没有量纲，对尺度变化不敏感。除掉由于离散区域旋转带来的误差，它对旋转也不敏感。由上式可见，形状参数 F 是用来描述区域形状接近圆形的程度的。当区域为圆形时 F 为 1，当区域为其它形状时 F 大于 1，当 F 与 1 差别越大时，其形状与圆的差别就越大。

提取目标形状参数如下表：

表 4.5 形状参数

	目标 1	目标 2
<i>F</i>	1.519	2.212

(四) 长轴和短轴

1、外接矩形：当图像的边界已知时，用外接矩形的尺寸来刻画基本形状是最简单的方法。求物体在坐标系方向上的外接矩形，只需要计算物体边界点的最大和最小坐标值，即可得到物体的水平和垂直跨度。但是对于任意朝向的物体，需要计算反映物体形状特征的主轴方向上的长度和与之垂直方向上的宽度。得到物体的最小外接矩形。

2、最小外接矩形的求法：

(1)首先定义图像的矩函数：将一幅图像看成一个二维密度函数分布 $f(x, y)$ ，这里，函数值 $f(x, y)$ 表示点 (x, y) 处图像像素的亮度值。用 ξ 表示 x - y 平面上图像的区域，也就是函数 $f(x, y)$ 的定义域范围。

设一幅图像的亮度值为 $f(x, y)$ ，它的 $(p+q)$ 阶矩函数的一般定义如下：

$$\Phi_{pq} = \iint_{\xi} \Psi_{pq}(x, y) f(x, y) dx dy \quad (4-18)$$

$$p, q = 0, 1, 2 \dots$$

这里， $\Psi_{pq}(x, y)$ 是一个 ξ 内关于 (x, y) 的连续函数，它被称为矩的权核或基本集。下表 pq 通常表示函数 $\Psi_{pq}(x, y)$ 中定义的坐标 x, y 的次数。

(2)几何矩的定义：定义： $p+q$ 阶二维几何矩用 m_{pq} 表示，其表达式为：

$$m_{pq} = \iint_{\xi} x^p y^q f(x, y) dx dy \quad (4-19)$$

$$p, q = 0, 1, 2 \dots$$

这里， ξ 是图像亮度函数 $f(x, y)$ 定义的像素空间区域。

(3)几何矩的快速算法：一个 $N \times N$ 图像的几何矩，可用如下的公式求解，其表达式为：

$$m_{pq} = \sum_{x=1}^N \sum_{y=1}^N x^p y^q f(x, y) \quad (4-20)$$

对于一副二值图像，图像目标区域上的亮度函数取值“1”，背景取值“0”。这种情况下，式(4-20)可简化成：

$$m_{pq} = \sum \sum x^p y^q \quad (4-21)$$

这里，求和中的 x, y 是 $f(x, y)=1$ 的像素坐标。

除此之外，还有利用目标区域的边界点求取几何矩的快速算法，比如 σ 法、矩形积分法和围线积分法等。由于需要先求目标的边界，算法实现并不很高效，所以此处不做介绍。

(4)所用到的几何矩函数：

零阶几何矩：零阶几何矩 m_{00} 代表一幅图像的总亮度。对于剪影图像， m_{00} 是目标

区域的几何面积。

一阶几何矩：一阶几何矩 m_{10} ， m_{01} 是图像关于 x 轴和 y 轴的亮度矩。其亮度“矩心” (x_0, y_0) 为：

$$x_0 = \frac{m_{10}}{m_{00}} \quad y_0 = \frac{m_{01}}{m_{00}} \quad (4-22)$$

对于一副剪影图像，点 (x_0, y_0) 给出了图像区域的几何中心，又称为“形心”或“重心”。

中心矩：一副图像相对于亮度矩心所计算出的几何矩称为中心矩。它表示为：

$$\mu_{pq} = \iint_{\xi} (x - x_0)^p (y - y_0)^q f(x, y) dx dy \quad (4-23)$$

$$p, q = 0, 1, 2, \dots$$

方位角 θ ：定义：本原坐标轴之一对于 x 轴的方位角 θ 由下式给出：

$$\theta = \frac{1}{2} \tan^{-1} \left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}} \right) \quad (4-24)$$

其中本原坐标轴的定义：将穿过图像矩心的两条正交线定义为图像参考坐标系时，惯量分量为零时的参考坐标轴称为本原坐标轴。

(5)由图像几何矩函数的定义，以及相关几何矩的公式，可以提取出目标的一些特征^[73-75]：

由公式(4-22)可以求得目标的形心坐标，由公式(4-24)可以求得本原坐标轴之一对于 x 轴的方位角 θ 。然后通过坐标旋转很容易的求出其最小外接矩形。

坐标旋转公式如下：

$$\begin{cases} x_2 = x_0 + \cos\theta \cdot (x_1 - x_0) + \sin\theta \cdot (y_1 - y_0) \\ y_2 = y_0 - \sin\theta \cdot (x_1 - x_0) + \cos\theta \cdot (y_1 - y_0) \end{cases} \quad (4-25)$$

其中， (x_0, y_0) 为目标区域的中心，也称为“矩心”。 (x_1, y_1) 为原图像目标区域的像素坐标。 (x_2, y_2) 为旋转后的目标区域像素坐标。

然后可以在旋转后的目标图像上画出其最小外接矩形，再利用下面的反坐标旋转公式变换回原坐标。反坐标变换公式如下：

$$\begin{cases} x_1 = x_0 + \cos(-\theta) \cdot (x_2 - x_0) + \sin(-\theta) \cdot (y_2 - y_0) \\ y_1 = y_0 - \sin(-\theta) \cdot (x_2 - x_0) + \cos(-\theta) \cdot (y_2 - y_0) \end{cases} \quad (4-26)$$

最后，由最小外接矩形，可以直接得到目标的长轴和短轴等特征信息。

其最小外接矩形及参数信息如下：



图 4.8 目标最小外接矩形图

表 4.6 最小外接矩形参数

	目标 1	目标 2
长轴	33	43
短轴	15	14

（五）偏心率

偏心率也可以叫做伸长度，它描述的是区域图像的复杂度，可以通过下式进行计算：

$$r = \frac{L_{MER}}{W_{MER}} \tag{4-27}$$

其中， W_{MER} 为最小外接矩形的宽度， L_{MER} 为最小外接矩形的长度。一般定义最小外接矩形的短边为宽度，长边为长度，因此偏心率的取值大于 1。当 r 越接近 1，其形状与圆越接近。

偏心率 r 如下表：

表 4.7 偏心率

	目标 1	目标 2
r	2.239	3.128

（六）不变矩

矩的概念最早是在 1961 年由 Hu 提出来的。他将代数的不变量引入了矩的不变量，并对几何矩进行线性组合，得出了一组对图像平移、缩放、尺度变换不变的矩。在运用矩来处理图像的时候，把图像看做是一个二维密度分布函数，这样，用矩来处理图像就有了坚实的数学理论基础^[76-79]。

其中式(4-22)和(4-23)已经给出了矩函数及中心距的简单定义，中心矩相对于前面提到的简单矩，它有自己的优势，就是具有平移不变性，但是不具备比例不变性，我们在做特征提取希望提取的那些对平移、比例、旋转、镜像等都具有不变性的特征，因此我们提出了归一化的中心矩：

$$\eta_{pq} = \frac{\mu_{pq}}{m_{00}^{\frac{p+q}{2}+1}} \tag{4-28}$$

M . K .Hu 在 1962 年，基于以上归一化中心矩的线性组合的基础，提出了 Hu 不变矩，其计算公式如下：

$$h_1 = \eta_{20} + \eta_{02} \tag{4-29}$$

$$h_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (4-30)$$

$$h_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (4-31)$$

$$h_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (4-32)$$

$$h_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} - \eta_{03})^2 \right] \\ + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \left[3(\eta_{21} + \eta_{03})^2 - (\eta_{21} + \eta_{03})^2 \right] \quad (4-33)$$

$$h_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - \eta_{21} + \eta_{03}^2] \\ + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \quad (4-34)$$

$$h_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12}) \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 \right] \\ + (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (4-35)$$

4.2 基于类 Haar 特征的声纳图像特征提取

4.2.1 Haar 特征及类 Haar 特征

Haar 特征，也即 Haar-like 特征，是计算机视觉领域一种常用的特征描述算子。它最早是由 Papageorigiou 等人用于人脸描述。目前常用的 Haar-like 特征可以分为三类：线性特征、边缘特征、点特征（中心特征）、对角线特征^[80-84]。如下图所示：

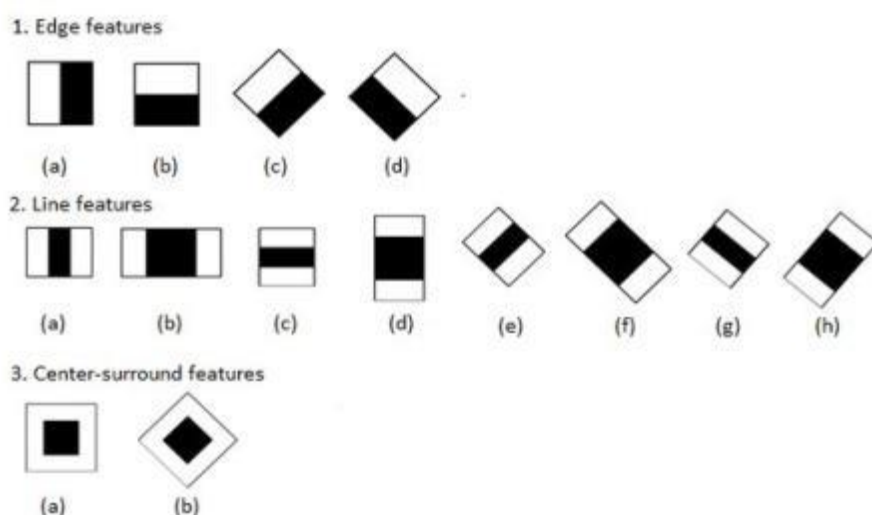


图 4.9 Haar-Like 特征

每一种特征的计算都是由黑色填充区域的像素值之和与白色填充区域的像素值之和的差值。而计算出来的这个差值就是所谓的 Haar-like 特征的特征值。

而本文所述的类 Haar 特征是在分析声纳图像本身特点（目标区域具有中心聚集特

性)的基础上构造出来的,它是基于 Haar 特征的中心特征(a),但是其计算方法不同。如图 4.10 所示:

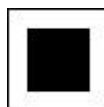


图 4.10 类 Haar 特征

其计算方法为:黑色填充区域的含有目标点(255)的像素个数和与白色填充区域的含有目标点(255)的像素个数和的差值。具体见下面对声纳图像的分类 Haar 特征提取实验。

4.2.2 特征提取实验

本实验选取经过滤波、分割等处理后的 8 幅声纳图像,其分割后的目标的面积、偏心率等都已经满足要求,下面图为分割后的结果图像及原始声纳灰度图像:



图 4.11 分割处理后的声纳图像

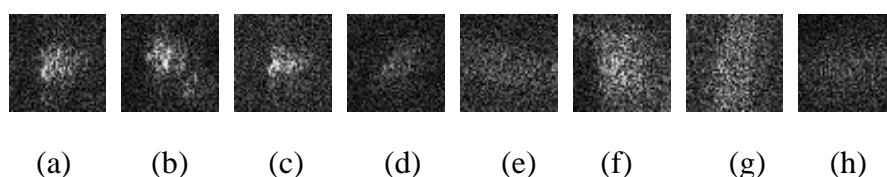


图 4.12 图 4.11 对应的原始声纳图像

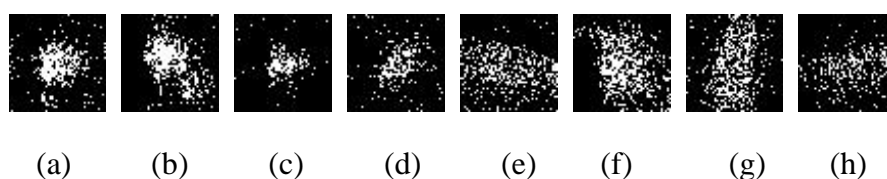


图 4.13 适当阈值分割后的声纳图像

其中,前面四个(a-d)为真实目标声纳图像,后面四个(e-h)为虚假目标声纳图像。下面对图 4.13 进行分类 Haar 特征提取实验与分析。

根据声纳图像目标的大小,选取图 4.10 的类 Haar 特征的外部边界大小为 50×50 ,内部黑色区域边界大小为 40×40 。

特征提取结果如表 4.8 所示。

由上表 4.8 可以看出,直接根据黑色区域或白色区域的特征不能够把目标与非目标区分开来,或者根据差值与黑色区域像素个数的比率也不能很好的把目标与非目标区分开来,而根据黑色区域与白色区域的像素个数的差值则可以很容易的把目标与非目标区

分开。

表 4.8 类 Haar 特征提取表

	a	b	c	d	e	f	g	h
黑色区域	116	102	80	93	234	216	219	178
白色区域	33	20	11	17	93	57	67	37
差值	83	82	69	76	141	159	152	141
差值与黑色区域比	0.715	0.804	0.862	0.817	0.602	0.736	0.694	0.792

特别是，根据图 4.14 对类 Haar 特征的分析与描述，很容易的区分出(a-d)为真实目标，而(e-h)为虚假目标。

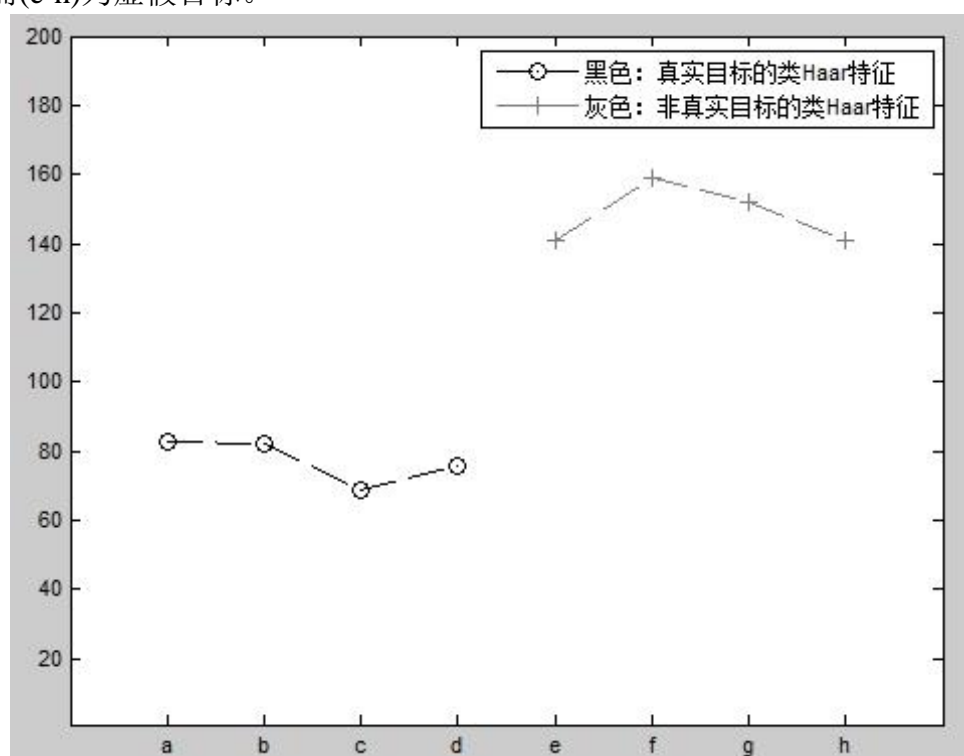


图 4.14 类 Haar 特征分析图

4.3 声纳图像的特征提取方法分析

通过对以上的不同种类的声纳图像特征提取方法进行分析和实验，现将上述不同种类的特征提取方法，在声纳图像的特征提取中的以下三个方面进行总结，分别是针对声纳图像提取的目标特征是否明显、算法的运算复杂度，以及对不同型号的声纳特征提取的适用情况进行列表说明，见表 4.9 所示。

从表中可以看出，声纳图像的特征提取中的几何特征明显，算法复杂度不高，而且对于大多数的声纳图像适用；类 Haar 特征则只适用于目标区域具有中心聚集特性的声纳图像；而纹理特征和颜色特征由于提取的特征不明显，基本上不适用于声纳图像目标的特征提取中。

表 4.9 不同种的特征提取方法对比表

	针对声纳图像提取的目标特征是否明显	算法复杂度（高、中、低）	对不同型号声纳特征提取的适用情况
颜色特征	不明显	低	不适用
纹理特征	不明显	高	基本适用
几何特征	明显	中	大都适用
类 Haar 特征	明显	低	适用于目标区域具有中心聚集特性的声纳图像

4.4 本章小结

本章首先对常用的声纳图像特征进行了总体的介绍与分析,并着重对基于纹理的声纳图像特征和基于形状的声纳目标特征进行了研究与提取实验,取得了较好的效果。其中纹理特征主要包括:GMRF 纹理、Gabor 纹理和 GLCM 纹理特征;形状特征提取了目标的面积、周长、长短轴、主轴方向、紧密度、实心度、偏心率等是广泛应用的全局几何特征量以及不变矩等。

然后在分析声纳图像本身特点的基础上构造出了用于声纳图像特征提取的类 Haar 特征,其原理与 Haar 特征的计算有所不同。通过实验验证了本文构造的特征提取方法的有效性和易用性,有利于整个目标检测系统方法的实时性及功能的完善性要求。

第5章 水下声纳图像的目标检测技术

目标检测技术是图像处理中的关键基础技术,利用运动目标检测技术可以对视频图像中感兴趣的目标进行实时地检测和识别等,是进一步视频图像处理的重要依据^[85-86]。目标检测是将运动目标从视频序列图像背景中分离出来的方法。而且目标的检测是目标跟踪中的一个重要的环节,目标的跟踪和识别的前提就是要把目标从背景中分离出来,之后再对目标进行跟踪和进一步的识别。特别是水下背景相对比较复杂,噪声污染严重等,在目标的检测方面更加的困难,所以,选取较为合适有效的目标检测方法就显得尤为重要^[87-88]。

5.1 水下目标检测技术

在光学成像的视频中,检测运动目标的方法已经非常成熟,常用的有点检测法、帧间差分法、光流法、背景减除法、基于分割的方法和监督学习的方法等^[89-92]。

由于水下环境的复杂性和声纳设备本身固有的特点,所形成的声纳图像噪声污染非常严重,同一背景下的两帧图像也有很大的差别,所以基于帧间差分法、光流法和背景减除法的目标检测技术在水下声纳图像的目标检测中都不适用。而基于监督学习的方法需要大量的声纳图像数据,目前也没有足够的实验条件。常用的水下目标检测技术主要是基于分割的方法,这也是本文的重点。而基于角点检测的方法能取得一定的效果,但是效果还不是很理想。下面对其进行简单的介绍。

基于点检测的方法,即角点检测算法主要有 Moravec 算法, SUSAN 算法, Harris 角点检测算法等,文献中对几种角点检测算法进行了比较和评价,认为 Harris 的定位性能和鲁棒性较好^[93-94]。

Harris 的自相应矩阵 M 为:

$$M = \begin{bmatrix} A & C \\ C & B \end{bmatrix} = e^{-\frac{x^2+y^2}{2\sigma^2}} \otimes \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \rightarrow R^{-1} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} R \quad (5-1)$$

I 为像素点的灰度, x , y 分别为水平方向和垂直方向上的分量, R 为响应系数,对实对称矩阵对角化处理后,得到特征值 λ_1 和 λ_2 , 然后分析边缘区域、角点区域以及平坦区域。

角点的响应因子测量:

$$R = \det(M) - k \bullet \text{trace}^2(M) > T \quad (5-2)$$

其中 trace 为矩阵的迹, k 取 0.04 ~ 0.06, 判断点 (x, y) 是否为所求特征点, 是依据其角点值是不是邻域最大值, 同时它的值还要超过设置的阈值 T , 若满足则确定该点为

所要求的角点。

Harris 算法只使用了灰度的一阶差分以及滤波，操作简单，计算量小。在特征点信息比较丰富的区域，*Harris* 算子能提取出大量有用的特征点。由于 *Harris* 算法在计算过程中只用到一阶偏导数，所以即使图像存在灰度变化、噪声、旋转和透视变换，也可以比较稳定的提取出角点。下面为角点检测的简单实验。

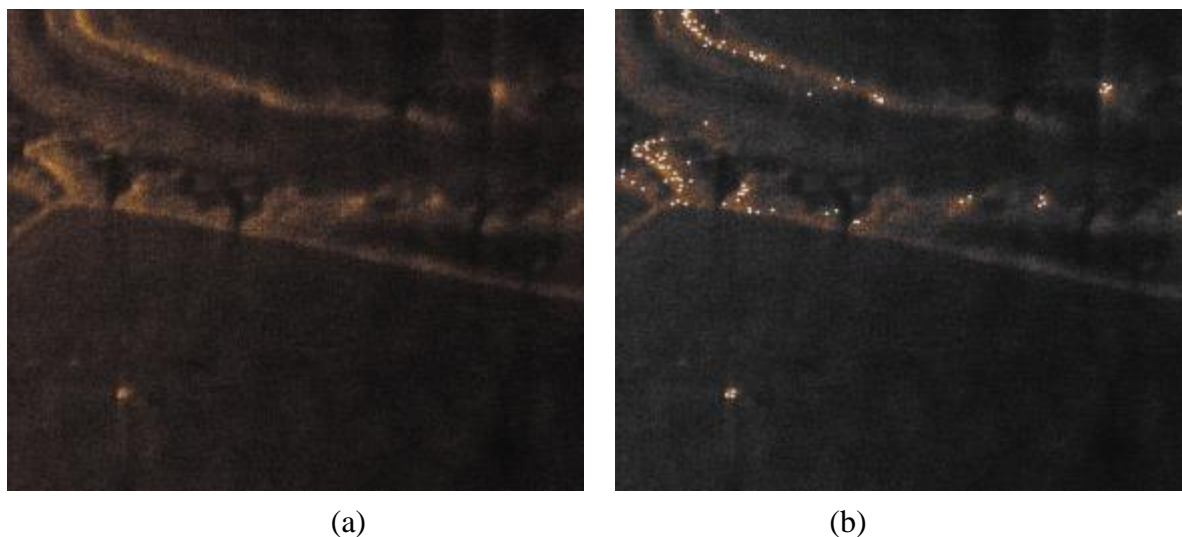


图 5.1 角点检测实验效果图

5.2 基于聚类分析的侧扫声纳目标检测

5.2.1 Klein5000 侧扫声纳图像数据分析

基于聚类分析的侧扫声纳目标检测技术是在分析 Klein5000 侧扫声纳图像数据的基础上设计的，如下图 5.2 Klein5000 侧扫声纳图像。



图 5.2 Klein5000 侧扫声纳图像

其数据分为左右两部分，其中右侧有一目标，其特点是：目标有阴影，且阴影在目标的右边；相同的，如果目标在左边，则其阴影也应该在目标的左边；目标的亮度信息不是很明显；而阴影区域相对明显，且有一定的宽度等。可以根据这些特点对目标进行分割和检测。

5.2.2 侧扫声纳目标检测流程图

如图 5.3 基于聚类分析的侧扫声纳目标检测流程图，程序首先开始读取声纳图像数据（包括图像数据的解码和插值等预处理），读取 x 行的声纳数据后开启一个线程进行

处理，同时主线程继续实时读取声纳数据， x 行的声纳图像如上图 5.2 所示；然后对得到的该声纳图像进行聚类，得到聚类后的目标图像和阴影图像；然后对聚类后的结果图像进行高斯金字塔滤波处理，使目标和阴影区域连续；再就是提取目标和阴影的特征；最后根据目标和阴影的这些特点以及目标和阴影的位置关系得到检测的结果；判断是结束还是继续读取声纳图像数据和继续处理。

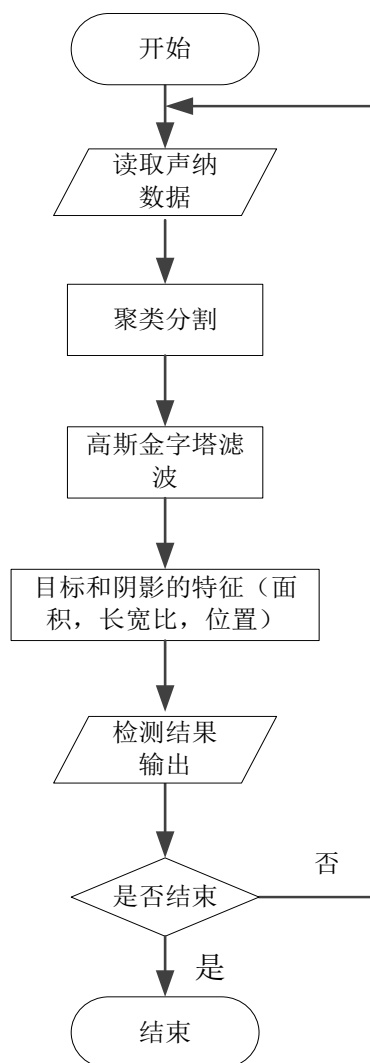
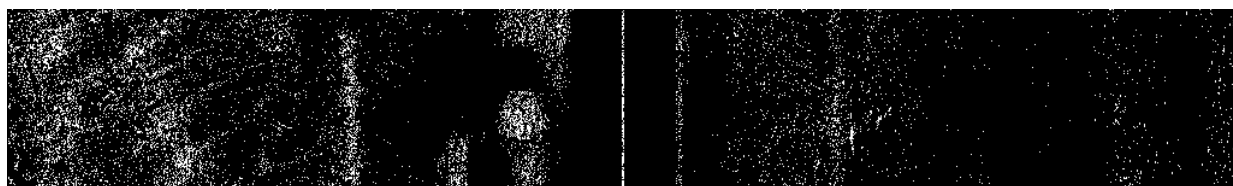


图 5.3 基于聚类分析的侧扫声纳目标检测流程图

5.2.3 实验结果与分析

下面就上面的分析结果对图 5.2 进行试验的结果如下：



(a) 聚类阈值化后的目标图像



(b) 聚类阈值化后的阴影图像(反色)

图 5.4 聚类阈值化处理后的结果图像

其中，聚类阈值化后的目标结果的阈值选择分为两种，因为可能一个阈值会使得目标结果不存在，而阴影图像也取值为反色后的图像，便于后续处理。

下面为进一步滤波阈值后的结果：



(a) 图 5.4 进一步滤波阈值后的目标图像



(b) 图 5.4 进一步滤波阈值后的阴影图像

图 5.5 图 5.4 进一步滤波阈值后的结果图像

对图 5.5 的阴影图像进行处理如下：

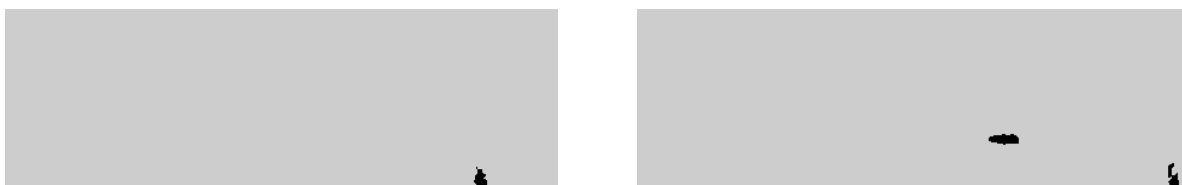


(a) 左半部分

(b) 右半部分

图 5.6 对图 5.5 的阴影图像去除较大较小区域后的结果

图 5.6 为对阴影部分，提取目标的面积特征之后，并去掉其面积较大和较小者后的结果图像，其中的右半部分为左右对换后的图像，只是为了进行统一处理，这样的话阴影都会在目标的左侧。进一步的处理如下：



(a) 左半部分

(b) 右半部分

图 5.7 对图 5.6 去除面积与长宽比的比率较小者

图 5.7 为对图 5.6 去除其面积与其长宽积的比率较小者后的图像。

最后根据其右边是否有目标，以及目标的特点，得到最终的检测结果如下图所示：



图 5.8 基于聚类分析的侧扫声纳目标检测结果

可以看出，该方法检测效果较好，速度和精度都能满足实时性目标检测系统的要求。

5.3 基于类 Haar 特征的目标检测技术

5.3.1 声纳图像目标分析

如下图所示的基于类 Haar 特征的目标检测技术所用声纳图像，如图 5.9 所示。

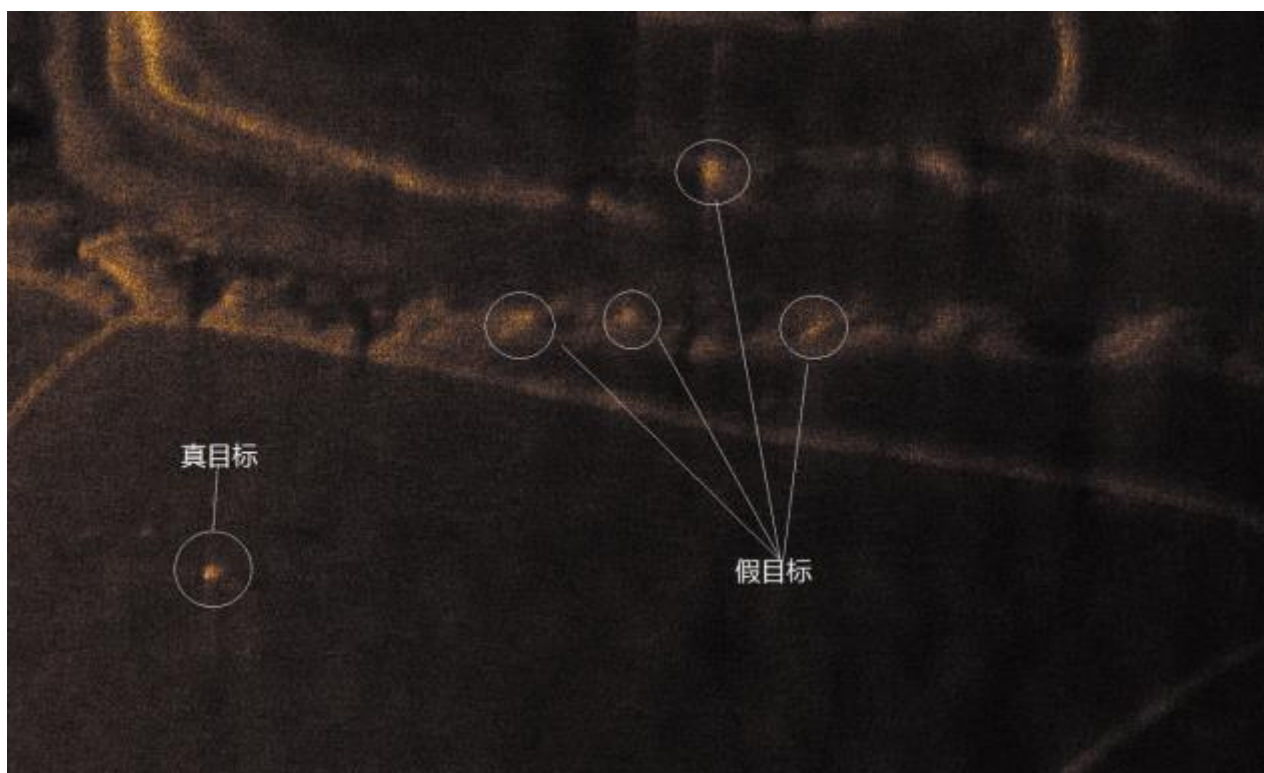


图 5.9 基于类 Haar 特征的目标检测技术用声纳图像

从图中可以看出，该侧扫声纳图像目标有着自己的特点，和前面的侧扫声纳图像有很大的不同：首先是目标没有阴影；其次是目标亮度特征不是很明显；再就是假目标的特征和真实目标比较接近，不易区分。如图所示的红色标记的目标为真实目标，而蓝色标记的区域则不是目标，从图像中可以看出，真实目标与其周围像素区别较大，而假目标周围有较多噪声干扰。

5.3.2 流程图与实验结果

图 5.10 为基于类 Haar 特征的目标检测系统流程图。

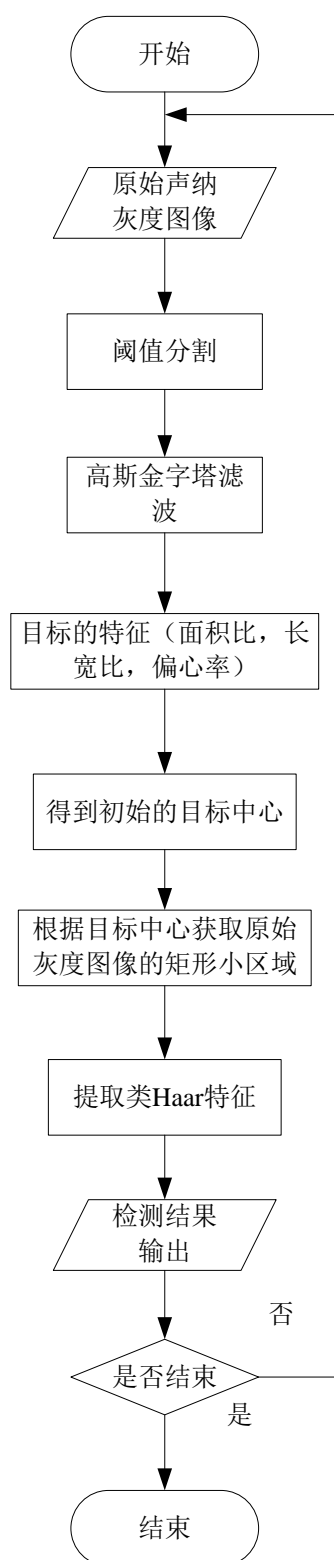


图 5.10 基于类 Haar 特征的目标检测系统流程图

其目标检测过程如下：首先获得原始声纳灰度图像，系统中选取 X 行进行一次处

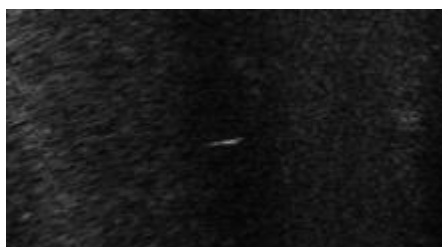
理，每次选取重叠 Y 行；然后经过简单的阈值分割，得到阈值后的结果；再就是滤波和提取目标的特征；得到初始检测结果，此时的检测结果有较多的虚假目标，如图 4.11 所示，需要进一步的去掉假目标；然后对当前检测的结果，提取中心，并以该中心取原始灰度图像的中心小区域如图 4.12 所示；然后对得到的目标小区域取阈值后提取类 Haar 特征，可以直接得到检测的目标结果，其检测方法简单，效果相对较好。

5.4 卡尔曼预测相关联的目标检测技术

5.4.1 数据分析

本节针对 BlueView 前视声纳实验数据进行实验分析，本次的目标检测实验是目标固定不动，而载着前视声纳的拖船对着目标向前移动。由于前视声纳的成像机理，可知前后帧的图像目标是有关系的，也就是可以认为当前帧的目标位置只和上一帧有关系，满足马尔科夫性，这也是卡尔曼预测模型的一个特点。再就是同一检测算法不能适用于所有帧的目标，使得某些帧的图像目标漏检，为了克服目标检测系统中的此种不足，本文采用了基于卡尔曼预测的目标检测技术，原理是：在先前的目标检测的基础上，如果确定是目标，则用卡尔曼预测目标下一帧可能的位置，最后将先前的目标检测结果和基于卡尔曼的目标预测结果进行综合，得到最终的检测结果。

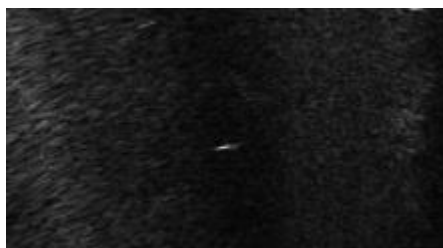
下图为 BlueView 前视声纳目标效果图(连续三帧)。从图像中可以看出，目标的亮度信息比较明显，其形状特征在多帧图像中变化较大，不适用于目标的特征提取，但是其面积等特征在前后帧的变化相对较小。



(a)



(b)



(c)

图 5.11 BlueView 前视声纳目标效果图

5.4.2 卡尔曼预测原理

(一) 马尔科夫模型

首先，卡尔曼预测模型具有马尔可夫性质，马尔可夫模型是一种简化的、高效的随机模型。马尔可夫模型假定在已知系统现在所处的状态的情况下，系统将来的演变与过去的状态无关的一种模型^[95-97]。

一维马尔科夫随机过程：设有随机过程 $\{X_n, n \in T\}$ ，若对于任意整数 $n \in T$ 和任意的 $i_0, i_1, \dots, i_n \in I$ ，条件概率满足 $P\{X_{n+1} = i_{n+1} | X_n = i_n, \dots, X_0 = i_0\} = P\{X_{n+1} = i_{n+1} | X_n = i_n\}$ 就称 $\{X_n, n \in T\}$ 为马尔可夫(MARKOV)过程，该随机过程的统计特性是完全由条件概率所决定的^[98]。

也即是假如已经知道当前目标的位置，则可由马尔科夫性，直接根据当前位置预测下一帧目标的位置，而不需要知道以前目标的位置状态。

(二) 卡尔曼预测理论

卡尔曼预测算法是由卡尔曼等人在上世纪 60 年代提出的一种算法，主要是利用系统的过程模型，来预测系统的下一状态^[99-100]。假设现在的系统状态是 k ，根据系统的模型，可以基于系统的上一状态而预测出现在状态：

$$X(k|k-1) = A X(k-1|k-1) \quad (5-3)$$

式(5-3)中， $X(k|k-1)$ 是利用上一状态预测的结果， $X(k-1|k-1)$ 是上一状态最优的结果， $U(k)$ 为现在状态的控制量，如果没有控制量，它可以为 0 本文中该项为 0。

令现在的状态为：

$$X = \begin{bmatrix} x \\ y \\ dx \\ dy \end{bmatrix} \quad (5-4)$$

而预测下一时刻的状态为：

$$X^- = \begin{bmatrix} x^- \\ y^- \\ dx^- \\ dy^- \end{bmatrix} \quad (5-5)$$

以及状态变量 A 为:

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5-6)$$

则目标的预测结果为:

$$\begin{cases} x^- = x^+ + d \\ y^- = y^+ + d \end{cases} \quad (5-7)$$

其中, dx^- 和 dy^- 需要根据下一次的检测结果和预测结果的综合来更新。

5.4.3 前视声纳目标检测流程图

如图 5.12 所示 BlueView 前视声纳目标检测系统流程图。介绍如下:

首先是 BlueView 前视声纳的数据读取与解码的过程, 这个过程需要用到生产厂家的数据读取的 API 库。公司是以动态链接库的形式提供的, 无法得到声纳内部的数据格式定义, 进而可以得到原始的灰度图像, 作为下一步继续处理的基础。

为了实现目标的预测相结合的检测, 本文首先对目标进行分割检测, 然后对检测到的目标进行统计计数, 如果该目标出现的次数大于某个值, 则认为该目标是真目标, 同时把它加入到一个新的目标链中。其中判断某一个目标是否为同一个目标是根据目标之间的距离判断的。

在对原始灰度图像进行处理的同时, 判断当前的目标链是否为空, 如果不为空, 说明目标链中已经有真实目标的位置等信息, 可以对其进行预测相关的目标检测。最后将检测的结果和正常检测的结果进行综合考虑, 得到最终的检测结果。

最后对目标的状态进行更新和目标的信息显示等处理。然后循环该过程。

需要注意的是, 如果当前已经存在某个目标链, 但是正常检测的结果却很多次都没有检测到目标, 则认为该目标链中的目标为假目标, 需要对目标链进行清空操作。而如果正常检测和预测都能检测到结果, 则以正常检测的结果来更新目标的最新位置状态等信息。如果, 只是某一帧正常检测没有检测到目标, 可以直接用预测结果来更新当前的状态, 但也要注意的, 并不是直接用预测的结果来更新当前的状态, 而是对预测区域进行进一步的分割检测结果, 只是其检测限制条件没有正常检测的要求高, 一般情况下都能检测的目标。

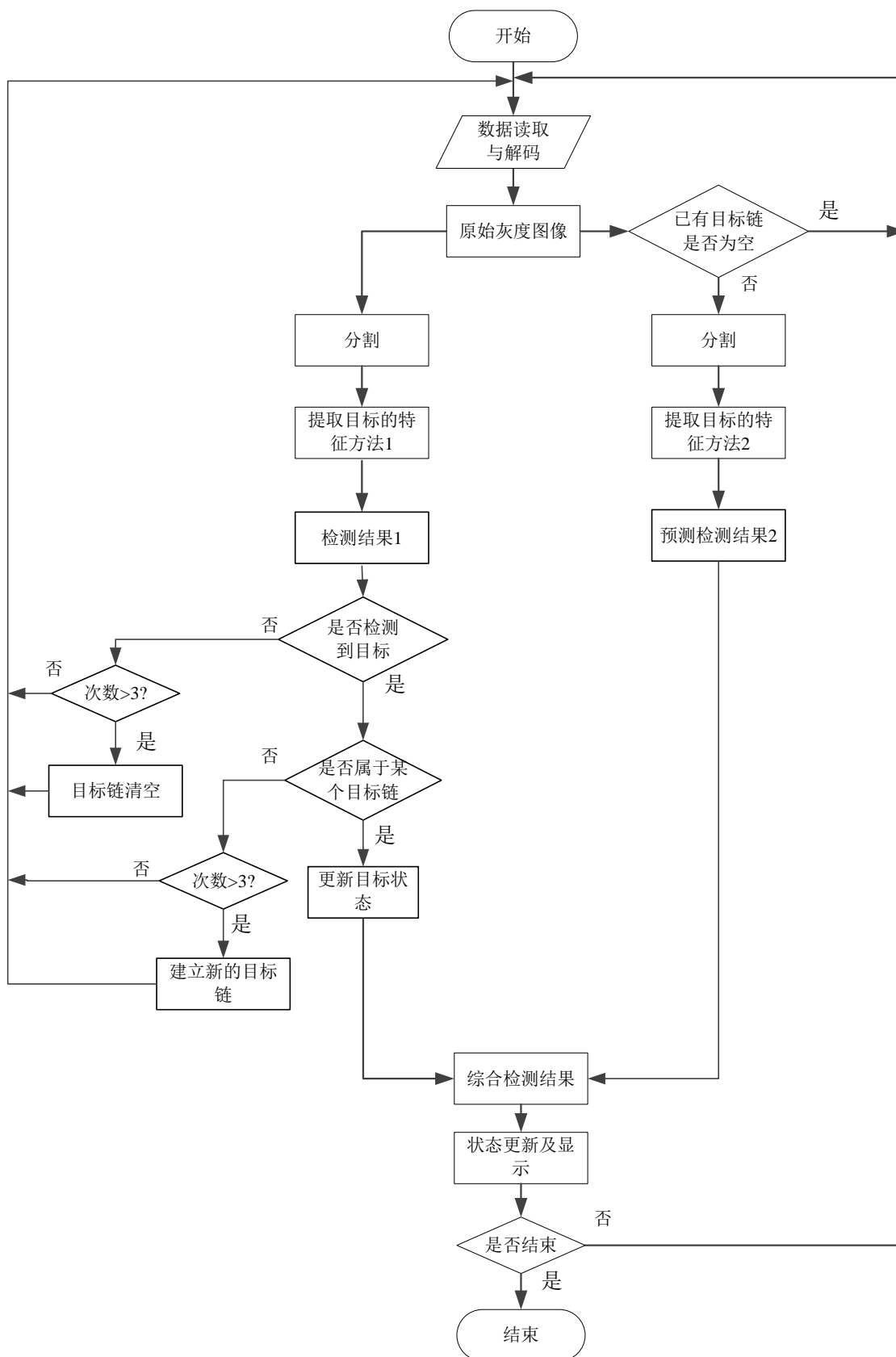
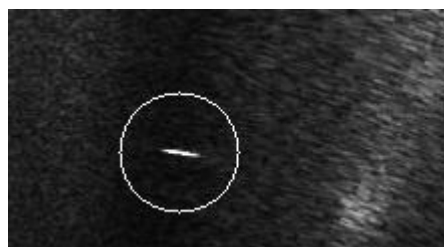


图 5.12 BlueView 前视声纳目标检测系统流程图

5.4.4 实验结果及分析

本文选取连续几帧的图像进行目标检测的实验与分析。

首先给出没有加入预测的目标检测结果：



(a)



(b)



(c)



(d)

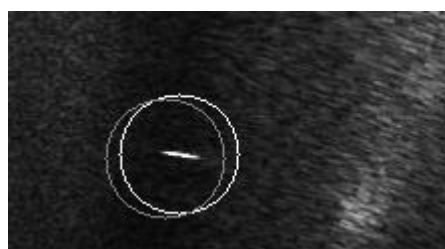


(e)

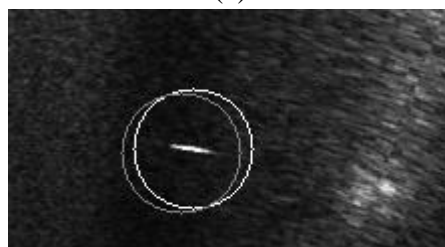
图 5.13 未加入预测的目标检测效果

由图 5.13 可以看出，未加入预测的目标检测结果显示，目标可能出现有漏检的情况。

下面给出加入预测之后的检测效果图。



(a)



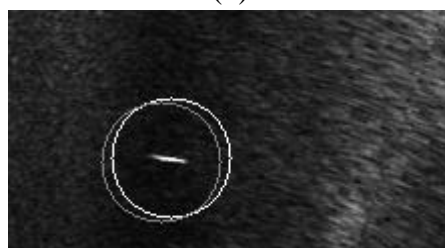
(b)



(c)



(d)



(e)

图 5.14 加入卡尔曼预测后的目标检测效果

从图 5.14 中可以看出，加入卡尔曼预测后的目标检测效果较好，目标能够进行连续的检测，其中灰色的圈表示预测目标下一帧可能的位置，白色圈为当前综合处理后的目标检测结果。

实验结果较好。

下面给出其他的几个目标检测效果。

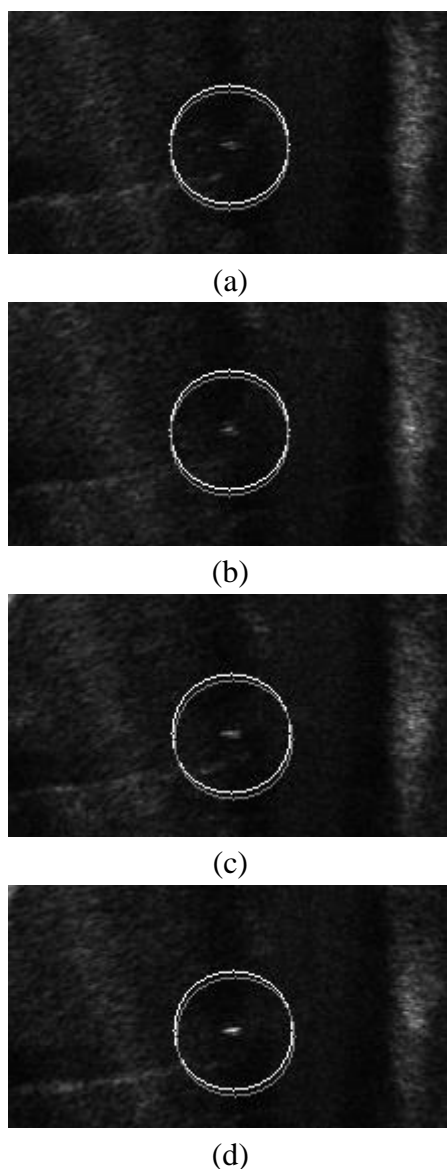


图 5.15 卡尔曼预测相关联的其它目标检测效果

5.5 本章小结

本章首先介绍了主要的光学目标检测方法，主要包括点检测法、帧间差分法、光流法、背景减除法、基于分割的方法和监督学习的方法等。分析声纳图像的特点，指出基于分割的方法在水下声纳目标检测中应用较广。在此基础上，实现了基于 Klein5000 侧扫声纳的目标检测方法、基于类 Haar 特征的目标检测方法和基于卡尔曼预测相关联的前视声纳目标检测方法，并分别对其进行了实时仿真实验与分析。

其中，基于聚类分析的目标检测方法和基于卡尔曼预测相关联的目标检测方法都能够进行实时性的目标检测，实验效果较好。而基于类 Haar 特征的目标检测方法，由于该声纳的成像效果不好，而且声纳本身系统还在实验阶段，该方法有待进一步的研究与实验。

结 论

本文通过查阅国内外的相关文献资料和对不同的目标检测和分割算法进行大量实验的基础上,对水下目标检测识别系统中的图像数据预处理、声纳图像分割和声纳图像的特征提取技术进行了研究和实验分析,并在此基础上研究出了相应的实时目标检测识别方法。总的来说,本文主要包括如下几个方面的研究工作:

1、对不同型号的声纳数据进行了分析和去噪实验。在对声纳图像进行直方图分析的基础上,着重对基于高斯金字塔的多分辨率声纳图像滤波技术进行了研究与应用。由于其在声纳图像的滤波中使用范围广,而且效果较好,速度快等特点,本文将该滤波技术应用于声纳图像目标检测系统的预处理部分当中。

2、针对声纳图像的特点,主要对几种简单的基于阈值的声纳图像分割方法和基于聚类的声纳图像分割方法进行了研究与实验分析。并在此基础上对基于 **MRF** 的声纳图像分割方法也进行了研究和改进,设计了一种新的基于 **MRF** 的非监督声纳图像分割方法,该法通过一种快速有效的高斯金字塔模型对声纳图像进行预处理,使得处理后的声纳图像的海底混响区服从高斯分布。在此基础上构造了一个能够自动确定声纳图像分类个数的模型,并通过该模型结合一种局部能量极化的方法对马尔科夫模型的初始化参数进行估计,从而形成一种完全自动的声纳图像分割模型。

3、在分析声纳图像的特征和声纳图像特征提取技术的基础之上,着重对一种新的基于类 **Haar** 特征的声纳图像特征提取技术进行了实验分析与介绍,实验结果表明,该方法提取的特征效果较为明显。

4、针对 **Klein5000** 侧扫声纳,研究出了基于聚类分析的侧扫声纳实时目标检测方法。主要功能包括:声纳数据的实时读取与解码、数据瀑布显示、检测的实时检测与圈取,以及目标信息(GPS 信息等)的实时显示等。

5、针对 **BlueView** 前视声纳,研究出了基于卡尔曼预测相关联的前视声纳实时目标检测方法。主要功能包括:声纳数据的实时读取、量程的选择、通信、图像的显示、检测到目标的信息显示(主要包括目标的方位角和距离)、检测到目标的实时圈取等。该系统已经进行了湖试实验,其目标检测效果较为理想。

同时,由于作者精力和能力的限制,本课题的一些方面还有待进一步的改进,下面对其做进一步的研究展望:

- 1、实时目标检测系统方法整体功能的进一步完善。
- 2、目标检测识别系统中的目标模型的建立,此部分的工作需要大量的目标训练样本。
- 3、通过大量的实验与分析,开展对目标识别工作的研究,意义重大。

参考文献

- [1] 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, 2009: 1-3 页
- [2] 田坦, 刘国枝, 孙大军. 声纳技术. 哈尔滨工程大学出版社, 2000: 1-2 页
- [3] 章毓晋. 图像工程(中册)图像分析(第 2 版). 清华大学出版, 2005: 73-75 页
- [4] 黄长专, 王彪, 杨忠. 图像分割方法研究[J]. 计算机技术与发展, 2009, 19(6): 76-79 页
- [5] Stitt J P, Tutwiler R L, Lewis A S. Synthetic Aperture Sonar Image Segmentation using the Fuzzy C-Means Clustering Algorithm [J]. Autonomous Control and Intelligent Systems Division the Pennsylvania State Applied Research Laboratory. USA, 2001.
- [6] Stitt J P, Tutwiler R L, Lewis A S. Fuzzy C-Means image segmentation of side scan sonar images[C]//Proc. IASTED. 2001: 27-32P
- [7] 郭海涛, 田坦, 张春田. 基于模糊聚类的声纳图像多区域分割. 海洋技术, 2004, 23(3): 39-45 页
- [8] 王雷, 叶秀芬, 王天. 模糊聚类的侧扫声纳图像分割算法[J]. 华中科技大学学报, 2012, 40(9): 25-29 页
- [9] 伍海. 基于粒子群 FCM 的声纳图像分割[J]. 计算机工程与应用, 2010, 46(98): 238-241 页
- [10] 高梅. 基于主动轮廓模型的图像分割算法. 漳州师范学院学报, 2007, 4: 41-46 页
- [11] Chan T F, Vese L A. Active contours without edges [J]. Image Processing, IEEE Transactions on, 2001, 10(2): 266-277P
- [12] T. Chan and L. Vese. Active contour without edges for vector-valued image [J]. Journal of Visual Communication and Image Representation, 2001, 10(2): 266-277P
- [13] Chan T. F., Sandberg B.Y., VeseL. An Active contour without edges for vector-valued images [J]. J. Visual Communication Image Representation, 2000, 11(2), 130-141P
- [14] 刘光宇, 卞红雨等. 基于声纳图像的水平集分割算法研究[J]. 传感器与微系统, 2012, 31(1) : 29-31 页
- [15] Xiu F Y, Zhe H Z, Peter X. L, et al. Sonar image segmentation based on GMRF and level-set models[J]. Ocean Engineering. 2010, 37(10):891-901P
- [16] SGeman and D.Geman. Stochastic relaxation, Gibbs distribution and Bayesian restoration of image. IEEE Trans. Patten Anal. Machine Intel, 1984. 6: 721-741P
- [17] Max Mignotte, Christophe Collet, Patrick Perez, et al. Sonar image segmentation using

- an unsupervised hierarchical MRF model [J]. IEEE Transactions on Image Processing, 2000, 9(7):1216-1231P
- [18] 阳凡林, 独知行, 李家彪, 等. 基于 MRF 场的侧扫声呐分割方法[J]. 海洋学报, 2006, 28(4): 43-48 页
- [19] 叶秀芬, 王兴梅, 张哲会, 等. 新的声纳图像自动分割方法[J]. 计算机工程与应用, 2009, 45(2): 168-172 页
- [20] 汪西莉, 刘芳, 焦李成. 基于不完全分层 MRF 的非监督图象分割[J]. 电子学报, 2004, 32(7): 1087-1089 页
- [21] Klausner N, Azimi-Sadjadi M R, Tucker J D. Multi-sonar target detection using multi-channel coherence analysis[C]//OCEANS 2010. IEEE, 2010: 1-7P
- [22] 毛盾, 刘忠, 程远国. 基于蛙人探测声纳序列图像的水下小目标检测算法[J]. 传感技术学报, 2011, 07: 1027-1032 页
- [23] 吴丽媛, 徐国华, 余琨. 基于前视声纳的成像与多目标特征提取[J]. 计算机工程与应用. 2013, 49(2): 222-225 页
- [24] 田杰, 张春华. 基于分形的水声图像目标探测[J]. 中国图象图形学报, 2005, 10(4): 479-483 页
- [25] Karabchevsky S, Braginsky B, Guterman H. AUV real-time acoustic vertical plane obstacle detection and avoidance[C]//Autonomous Underwater Vehicles (AUV), 2012 IEEE/OES. IEEE, 2012: 1-6P
- [26] 许文海, 续元君, 董丽丽, 等. 基于水平集和支持向量机的图像声呐目标识别[J]. 仪器仪表学报, 2012, 33(1): 49-55 页
- [27] 洪义平. 鲁棒的目标检测与识别方法研究[D]. 中国科学院研究生院博士学位论文, 2005: 5-8 页
- [28] 汤春瑞. 水下目标声图像多分辨率分析及识别研究[D]. 哈尔滨工程大学, 2009: 45-46 页
- [29] Abraham D A, Lyons A P. Reverberation envelope statistics and their dependence on sonar bandwidth and scattering patch size[J]. Oceanic Engineering, IEEE Journal of, 2004, 29(1): 126-137P
- [30] 李轩, 徐宏宇, 刘玉英, 等. 合成孔径声纳成像聚焦问题的研究[J]. 微处理机, 2004, 25(3): 41-42 页
- [31] 石红. 声纳图像处理关键技术研究[D]. 哈尔滨工程大学, 2011: 2-3 页
- [32] 沈郑燕. 声纳图像去噪与分割技术研究[D]. 哈尔滨工程大学, 2010: 1-3 页
- [33] Padmavathi G, Subashini P, Kumar M M, et al. Comparison of filters used for underwater image pre-processing[J]. IJCSNS, 2010, 10(1): 58 P

- [34] Buades A, Coll B, Morel J M. A review of image denoising algorithms, with a new one[J]. Multiscale Modeling & Simulation, 2005, 4(2): 490-530 P
- [35] 万小红. 常用图像去噪方法探析与比较[J]. 运城学院学报, 2011.5: 019 页
- [36] Isar A, Firoiu I, Naforita C, et al. SONAR Images Denoising[J]. Sonar Systems, 2011:139-170P
- [37] Buades A, Coll B, Morel J M. Image denoising methods. A new nonlocal principle[J]. SIAM review, 2010, 52(1): 113-147 P
- [38] 李洪均. 基于多尺度几何分析与偏微分方程的图像去噪研究与应用[D]. 南京航空航天大学, 2012: 2-3 页
- [39] 王伟, 王宏志, 宋宇. 基于偏微分方程的图像去噪算法[J]. 长春工业大学学报: 自然科学版, 2011, 32(6): 554-557 页
- [40] 韩涛. 水下图像预处理技术研究[D]. 中国石油大学, 2010: 20-25 页
- [41] 冈萨雷斯. 数字图像处理 (第二版). 电子工业出版社, 2005. 9: 70-83 页
- [42] Zhao Z L, Liu B, Li W. Image Clustering Based on Extreme K-means Algorithm[J]. IEIT Journal of Adaptive & Dynamic Computing, 2012, 2012(1): 12-16 P
- [43] Khan K, Sahai A, Campus A. A fuzzy c-means bi-sonar-based metaheuristic optimization algorithm[J]. IJIMAI, 2012, 1(7): 26-32 P
- [44] Von Luxburg U. A tutorial on spectral clustering[J]. Statistics and computing, 2007, 17(4): 395-416 P
- [45] 冯超. k—means 聚类算法的研究[D]. 大连理工大学, 2007: 15-17 页
- [46] Li S Z. Markov random field modeling in image analysis[M]. Springer, 2009. 1-100 P
- [47] 赵春江. C# 数字图像处理算法典型实例[M]. 人民邮电出版社, 2009: 213-219 页
- [48] 张鹏, 张桂林. Markov 随机场在图像处理中应用的研究[D]. 武汉: 华中科技大学, 2005: 18-22 页
- [49] 叶秀芬, 王兴梅, 张哲会, 等. 新的声纳图像自动分割方法[J]. 计算机工程与应用, 2009, 45(2): 168-172 页
- [50] 汪西莉, 刘芳, 焦李成. 基于不完全分层 MRF 的非监督图像分割[J]. 电子学报, 2004.7, 32(7): 1087-1089 页
- [51] Yang X M, Wu W, Qing L B, et al. Image feature extraction and matching technology[J]. Optics and Precision Engineering, 2009, 9: 033 P
- [52] Chitaliya N G, Trivedi A I. Feature Extraction using Wavelet-PCA and Neural network for application of Object Classification & Face Recognition[C]//Computer Engineering and Applications (ICCEA), 2010 Second International Conference on. IEEE, 2010, 1: 510-514 P

- [53] Wang Y W, Zhang S Q, Lin B, et al. Feature Point Extraction Method of X-Ray Image Based on Scale Invariant[J]. Applied Mechanics and Materials, 2013, 274: 667-670 P
- [54] 曹健, 李海生, 蔡强, 等. 图像目标的特征提取技术研究[J]. 计算机仿真, 2013, 30(1): 409-414 页
- [55] 郭丽华, 刘超华, 丁士圻. 水下目标特征提取方法比较研究[J]. 吉林大学学报: 信息科学版, 2008, 26(4): 359-363 页
- [56] 孙晶, 褚金奎. 图像局部不变特征提取技术研究及其应用[D]. 大连理工大学, 2009: 15-25 页
- [57] Shi S, Xu D. A Robust Approach of Sonar Image Feature Detection and Matching[C]//Computer Science and Information Engineering, 2009 WRI World Congress on. IEEE, 2009, 6: 523-527 P
- [58] Liu Z, Wang H, Peng S. Texture segmentation using directional empirical mode decomposition[C]//Image Processing, 2004. ICIP'04. 2004 International Conference on. IEEE, 2004, 1: 279-282 P
- [59] Huazhong J, Minyi K, Xiwei Y, et al. Texture Analysis Using GMRF Model for Image Segmentation on Spectral Clustering[C]//Information Technology and Computer Science (ITCS), 2010 Second International Conference on. IEEE, 2010: 64-67 P
- [60] Mahajan P M, Kolhe S R, Patil P M. Segmentation of Texture Images Using Magnitude Response of Gabor Filters[J]. Computer Vision And Information Technology: Advances And Applications, 2010: 121 P
- [61] Rampun A, Strange H, Zwiggelaar R. Texture segmentation using different orientations of GLCM features[C]//Proceedings of the 6th International Conference on Computer Vision/Computer Graphics Collaboration Techniques and Applications. ACM, 2013: 17 P
- [62] 王业琴, 王辉. GMRF 随机场在纹理特征描述与识别中的应用[J]. 计算机工程与应用, 2011, 47(25): 202-204 页
- [63] Wang X, Huang C, Ni G, et al. Face Recognition Based on Face Gabor Image and SVM[C]//Image and Signal Processing, 2009. CISP'09. 2nd International Congress on. IEEE, 2009: 1-4 P
- [64] Chen L, Hagenah J, Mertins A. Texture analysis using gabor filter based on transcranial sonography image[M]//Bildverarbeitung für die Medizin 2011. Springer Berlin Heidelberg, 2011: 249-253 P
- [65] 杨宏雨, 余磊, 王森. 基于 Gabor 纹理特征的人脸识别方法倡[J]. 计算机应用研究, 2011, 28(10): 3974-3976 页

- [66] 付仲良, 张文元, 孟庆祥. 灰度和纹理特征组合的 SAR 影像 SVM 分类[J]. 应用科学学报, 2012, 30(5): 498-504 页
- [67] 韩晶, 邓喀中, 李北城. 基于灰度共生矩阵纹理特征的 SAR 影像变化检测方法研究[J]. 大地测量与地球动力学, 2012, 32(4): 94-98 页
- [68] 陈美龙, 戴声奎. 基于 GLCM 算法的图像纹理特征分析[J]. 通信技术, 2012: 2-034 页
- [69] Rampun A, Strange H, Zwiggelaar R. Texture segmentation using different orientations of GLCM features[C]//Proceedings of the 6th International Conference on Computer Vision/Computer Graphics Collaboration Techniques and Applications. ACM, 2013: 17 P
- [70] Cavalin P R, Kapp M N, Martins J, et al. A multiple feature vector framework for forest species recognition[C]//Proceedings of the 28th Annual ACM Symposium on Applied Computing. ACM, 2013: 16-20 P
- [71] Choi J, Ahn S, Chung W K. Robust sonar feature detection for the SLAM of mobile robot[C]//Intelligent Robots and Systems, 2005.(IROS 2005). 2005 IEEE/RSJ International Conference on. IEEE, 2005: 3415-3420 P
- [72] 高苗. 水声图像特征提取技术研究[D]. 哈尔滨工程大学, 2009: 10-33 页
- [73] 卢蓉, 范勇, 陈念年, 等. 一种提取目标图像最小外接矩形的快速算法[J]. Computer Engineering, 2010, 36(21): 178-180 页
- [74] 陈优广. 边界跟踪, 区域填充及链码的应用研究[D]. 上海: 华东师范大学, 2006: 10-50 页
- [75] 陈优广, 陈芳. 一种计算图像几何矩的快速算法[J]. 计算机应用与软件, 2007, 24(11): 172-173 页
- [76] Huang Z, Leng J. Analysis of Hu's moment invariants on image scaling and rotation[C]//Computer Engineering and Technology (ICCET), 2010 2nd International Conference on. IEEE, 2010, 7: V7-476-V7-480 P
- [77] Žunić J, Hirota K, Rosin P L. A Hu moment invariant as a shape circularity measure[J]. Pattern Recognition, 2010, 43(1): 47-57 P
- [78] 张伟, 何金国. Hu 不变矩的构造与推广[J]. 计算机应用, 2010, 30(9): 2449-2452 页
- [79] 周萧, 王岩. 基于 Hu 不变矩的叶片图像识别技术的研究[J]. 伺服控制, 2011 (6): 70-72 页
- [80] Cho J, Mirzaei S, Oberg J, et al. Fpga-based face detection system using Haar classifiers[C]//Proceedings of the ACM/SIGDA international symposium on Field

- programmable gate arrays. ACM, 2009: 103-112 P
- [81] 文学志, 方巍, 郑钰辉. 一种基于类 Haar 特征和改进 AdaBoost 分类器的车辆识别算法[J]. 电子学报, 2011, 39(5): 1121-1126 页
- [82] Erdem C E, Ulukaya S, Karaali A, et al. Combining Haar Feature and skin color based classifiers for face detection[C]//Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on. IEEE, 2011: 1497-1500 P
- [83] 甘玲, 朱江, 苗东. 扩展 Haar 特征检测人眼的方法[J]. 电子科技大学学报, 2010, 39(2): 247-250 页
- [84] Yang M, Lv F, Xu W, et al. Human action detection by boosting efficient motion features[C]//Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on. IEEE, 2009: 522-529 P
- [85] 景阳, 黄柔, 周婧琳. 运动目标检测方法概述[J]. 计算机光盘软件与应用, 2012 (23): 36-37 页
- [86] 刘明玺, 孟放. 运动目标检测与跟踪算法的研究进展[J]. 软件, 2010, 31(12): 85-88 页
- [87] Wei S, Leung H. A Markov Random Field Approach for Sidescan Sonar Change Detection[J]. 2012. 659-669 P
- [88] Quintal R T, Kiernan J E, Shannon J, et al. Automatic contact detection in side-scan sonar data[C]//Technologies for Homeland Security (HST), 2010 IEEE International Conference on. IEEE, 2010: 270-275 P
- [89] Yilmaz A, Javed O, Shah M. Object tracking: A survey [J]. Acm Computing Surveys (CSUR), 2006, 38(4): 7-15P
- [90] 万纓, 韩毅, 卢汉清. 运动目标检测算法的探讨[J]. 计算机仿真, 2006, 23(10): 221-226 页
- [91] 蔡荣太, 吴元昊, 王明佳, 等. 视频目标跟踪算法综述[J]. 电视技术, 2010 (012): 135-138 页
- [92] Leibe B, Schindler K, Cornelis N, et al. Coupled object detection and tracking from static cameras and moving vehicles[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2008, 30(10): 1683-1698 P
- [93] 龚平, 刘相滨, 周鹏. 一种改进的 Harris 角点检测算法[J]. 计算机工程与应用, 2010, 46(011): 173-175 页
- [94] Rosten E, Porter R, Drummond T. Faster and better: A machine learning approach to corner detection[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2010, 32(1): 105-119 P

- [95] 尹清波, 张汝波, 李雪耀, 等. 基于线性预测与马尔可夫模型的入侵检测技术研究[J]. 计算机学报, 2005, 28(5): 900-907 页
- [96] 闫新娟, 谭敏生, 严亚周, 等. 基于隐马尔科夫模型和神经网络的入侵检测研究[J]. 计算机应用与软件, 2012, 29(2): 294-297 页
- [97] Xi L, Fondufe-Mittendorf Y, Xia L, et al. Predicting nucleosome positioning using a duration Hidden Markov Model[J]. BMC bioinformatics, 2010, 11(1): 346 P
- [98] Bauer H. Harmonic spaces and associated Markov processes[M]//Potential Theory. Springer Berlin Heidelberg, 2011: 23-67 P
- [99] 王书玲. 基于卡尔曼滤波的跟踪技术的研究与应用[J]. 河北工业大学硕士学位论文, 2010: 14-16 页
- [100] Dept. of Electr, Nat.ChiaoTung, Hsinchu.Hybrid Unified Kalman Tracking Algorithms for Heterogeneous Wireless Location Systems.Vehicular Technology[J], IEEE Transactions on visualization and computer graphics. 2010, 12(2): 219-230P

攻读硕士学位期间发表的论文和取得的科研成果

- [1] 叶秀芬, 张元科. 一种基于 MRF 模型的非监督声纳图像分割方法. 中国知识产权局, 专利(已受理). 申请号: 201310148765.5. 申请日期: 2013.04.25.
- [2] 叶秀芬, 张元科. 基于 MRF 的非监督声纳图像分割方法. 哈尔滨工程大学学报 (已投稿).
- [3] Xiufen Ye, Jian'guo Zhang, Zhenhua Wang, Yuanke Zhang. Research on Cutting Algorithm Based on Local Mesh Projection in Virtual Surgery Simulation. *Mechatronics and Automation (ICMA)*, 2013 IEEE International Conference on , 2013: 279-284P
- [4] 叶秀芬, 李鹏, 徐世洋, 张元科. 一种基于 GPS 定位辅助的多波束侧扫声纳图像配准方法. 中国知识产权局, 专利(已受理). 申请号: 201310303759.2. 申请日期: 2013.07.19.
- [5] 叶秀芬, 张建国, 王振华, 张元科, 郭亚博. 一种软组织形变的局部质点弹簧模型的实时动态生成方法. 中国知识产权局, 专利(已受理). 申请号: 201310697661.x. 申请日期: 2013.12.18.
- [6] 叶秀芬, 张元科. 基于 MRF 的非监督声纳图像分割方法. 哈尔滨工程大学自动化学院 2013 年学术年会.

致 谢

两年半的硕士学习生涯即将结束，而硕士学位论文是我两年半学习成果的一个总结，在论文完成的过程中，除了我自己的潜心学习和研究之外，也凝聚了很多其他人的心血。在此，我要对帮助我完成科研和论文，以及一直以来陪我学习和生活的所有人表示感谢。

首先，我要对我的导师叶老师表示最由衷的感谢，叶老师学识渊博，为人和蔼，无论在学习上还是生活上都给了我很多的帮助。尽管平时她对我们要求较为严格，但正是由于这种严格的管理制度，才让我们能够投入更多的时间和精力用于学习和科研，才能够有今天的成果。叶老师平时工作较为繁忙，很少有休息时间，希望老师以后在工作的同时，能多注意身体。在此，我再次感谢叶老师在我攻读硕士学位期间对我无私的关心和帮助。

其次，要感谢实验室的两位博士成员，李朋师兄和张建国博士。每次有论文、专利和专业知识不懂时，我都会向李朋师兄请教，而他都能非常详尽地给我讲明白，不仅如此，师兄还为人和善，对人诚恳。和张建国博士则是经常一起吃饭、一起上课的挚友，建国兄乐于助人，每次有什么需要帮助的时候首先都会想到建国兄，怀念我们一起讨论学习的日子。在此也希望他和他的女朋友能一直走下去！

再次，要特别感谢实验室项目组成员和王璘小师弟，以及实验室的其他成员，我们一起学习和生活，度过了人生中最美好的两年，有你们真好。

此外，我还要感谢我的父母，他们一直是我努力和学习的动力，感谢他们为我所付出的一辈子的辛劳，感谢我的朋友和亲人，感谢所有支持过我、帮助过我、批评过我、鼓励过我和理解过我的人们。

最后，向百忙之中抽出时间评审论文和参加答辩的各位专家、教授致以衷心的感谢！