谨以此论文献给所有关心支持我的人

------------------甄勇

基于渔业船舶轨迹数据的海洋捕捞规律挖掘研究与实现

学位论文完成日期：

指导教师签字：

答辩委员会成员签字：

**独 创 声 明**

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。据我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含未获得 （注：如没有其他需要特别声明的，本栏可空）或其他教育机构的学位或证书使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文作者签名： 签字日期： 年 月 日

---------------------------------------------------------------------

**学位论文版权使用授权书**

本学位论文作者完全了解国家有关保留、使用学位论文的法律、法规和学校有关规定，并同意以下事项：

1、学校有权保留并向国家有关部门或机构送交本学位论文的复印件和磁盘，允许论文被查阅和借阅；

2、学校可以将本学位论文的全部或部分内容编入学校学位论文数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编本学位论文；

3、学校可以基于教学及科研需要合理使用本学位论文。

需保密的学位论文在解密后适用本授权书。

学位论文作者签名： 导师签字：

签字日期： 年 月 日 签字日期： 年 月 日

基于渔业船舶轨迹数据的海洋捕捞规律挖掘研究与实现

摘 要

近年来，渔业船舶监控系统（VMS）在全球范围进行了广泛部署。最初服务于渔业航行安全，如今已成为观察渔业船舶运动、分析渔业资源变化、预测渔业经济产量的有力工具。它记录了渔业船舶轨迹数据，研究者们利用这些数据对捕捞努力量、渔业资源时空变化、渔民行为特点、海洋环境保护等问题进行了探索。

由于这些研究都侧重于渔业船舶轨迹中的捕捞行为，而VMS数据本身并不包含区分不同行为的信息，因此识别渔业船舶的捕捞行为是一个首要问题。为了解决这个问题，此前的研究或者利用渔业船舶速度的变化，或者利用航海日志等辅助信息。本文中提出了渔业船舶捕捞行为识别算法，仅用到VMS轨迹数据，不依靠其他辅助信息，利用形态学方法识别渔业船舶的捕捞作业轨迹。

本文使用的数据来自于浙江省温州市海洋渔业安全救助信息中心，收集了中国东海自2014年4月1日起至2016年6月30日止31条单拖渔业船舶共2,595,004条记录以及对应时段的渔业产量数据。通过以捕捞时间捕捞距离与产量对比，Pearson相关系数分别为89.80%和82.11%，验证了所提出的渔业船舶捕捞行为识别算法的准确性。

此外，以前的VMS数据分析研究忽略了渔业船舶在捕捞区域之间的航行轨迹。这种轨迹描述了捕捞区域之间的转换联系，表征着渔业船舶在海上的航道。借鉴于道路对城市规划的意义，渔业船舶航道有助于对捕捞区域进行切分。本文提出了渔业船舶航道识别算法，通过迭代的方式完成对捕捞区域的切分以及区域间航道的识别任务。进而，通过比较不同年份相同季度的航道变化，确认了捕捞区域的迁移趋势。

为了方便渔业资源管理人员对VMS轨迹数据的观察、分析和研究，本文设计并实现了渔业船舶轨迹数据分析系统。该系统分为数据管理层、业务层和展示层。数据管理层定义了文件结构并对VMS数据进行预处理；业务层的核心是渔业船舶捕捞行为识别算法和捕捞区域间航道发现算法；展示层通过业务逻辑的设计，能够方便地观察VMS数据的港口分布和特定时段内的轨迹及捕捞区域。

关键字: VMS；轨迹；渔业船舶；捕捞区域

Research and Reality on Fishing Vessel Trajectory Data

Abstract

In recent years, the Vessel Monitoring System (VMS) has been completed the worldwide deployment. It was used for the safety of fishery navigation originally. Nowadays, fishery workers use this powerful tool to observe fishing vessel activities, predict fishery economic and yield. VMS records the trajectory data which been used to explore the fishing effort, the spatio-temporal change of fishery resources, the characteristics of fishermen's behavior, and the protection of the marine environment.

Because these researches all focus on fishing activity in trajectory, and VMS data does not contain information to distinguish different activities, identifying the fishing behavior from VMS data is a primary problem. In order to solve this problem, previous approaches are basically according to change of vessel’s speed; or rely on the validated data from logbooks. In this paper, a fishing activity identification algorithm is proposed, which used an image processing method to complete this work. It uses only the VMS trajectory data without any auxiliary information. The data used in this paper come from Zhejiang Province Ocean and Fisheries Bureau. There are 2,595,004 records of 31 otter trawl vessels in the East China Sea from April 1, 2014 to June 30, 2016 and corresponding fishery yield data. By comparing the fishing time and distance with the yield, the Pearson correlation coefficient is 89.80% and 82.11% respectively, which verifies the accuracy of the algorithm.

Although there has been a lot of analysis on VMS data, the information of vessel sailing activities traveling from one fishing region to another are lost. These kinds of trajectory indicate the link relationship among fishing regions, which can be regard as ocean roads. Taking into account the significance of roads to urban planning, ocean roads can help to classify fishing regions. In order to make full use of this navigational trajectory, this paper proposes an algorithm of ocean roads recognition, which classify the fishing regions and identify the ocean roads by iteration. By comparing the roads change in the same quarter of different years, it is found that the fishing area is moving north-east.

In order to facilitate the fishery staff to observe, analyze and study the VMS trajectory data. In this paper, a fishing boat trajectory data analysis system is designed and realized. It is divided into the data management layer, the business layer and the display layer. The data management layer defines the file structure and VMS data preprocessing. The core of the business layer is fishing activity identification algorithm. The presentation layer can conveniently observe the port distribution and the fishing region a period of time from VMS data.

Keywords: VMS; Mathematical Morphology; trajectory; iteration

目 录

[1 引言 8](#_Toc508613065)

[1.1 研究背景与意义 8](#_Toc508613066)

[1.2 研究现状 9](#_Toc508613067)

[1.3 研究内容和贡献 10](#_Toc508613068)

[1.4 论文组织结构 11](#_Toc508613069)

[2 渔船捕捞行为识别算法 12](#_Toc508613070)

[2.1 算法结构 13](#_Toc508613071)

[2.2 VMS数据处理 13](#_Toc508613072)

[2.3 港口定位 14](#_Toc508613073)

[2.4 航次划分 17](#_Toc508613074)

[2.5 捕捞行为识别 17](#_Toc508613075)

[2.5.1 轨迹压缩 17](#_Toc508613076)

[2.5.2 数学形态学 19](#_Toc508613077)

[2.6 结果验证与分析 20](#_Toc508613078)

[2.6.1 结果验证 20](#_Toc508613079)

[2.6.2 结果分析 21](#_Toc508613080)

[2.7 本章小结 22](#_Toc508613081)

[3 渔船航道识别算法 24](#_Toc508613082)

[3.1 算法结构 25](#_Toc508613083)

[3.2 迭代计算的输入 25](#_Toc508613084)

[3.2.1 轨迹与航行线段 26](#_Toc508613085)

[3.2.2 捕捞热度图与分层 26](#_Toc508613086)

[3.3 投票 27](#_Toc508613087)

[3.4 迭代流程 28](#_Toc508613088)

[3.5 渔船航道 30](#_Toc508613089)

[3.6 本章小结 32](#_Toc508613090)

[4 软件实现 33](#_Toc508613091)

[4.1 系统架构 33](#_Toc508613092)

[4.2 数据管理层 34](#_Toc508613093)

[4.2.1 文件系统结构 34](#_Toc508613094)

[4.2.2 数据处理 35](#_Toc508613095)

[4.3 业务层 35](#_Toc508613096)

[4.3.1 港口定位模块 35](#_Toc508613097)

[4.3.2 航次划分模块 35](#_Toc508613098)

[4.3.3 捕捞区识别模块 36](#_Toc508613099)

[4.4 展示层 36](#_Toc508613100)

[4.4.1 导入数据 38](#_Toc508613101)

[4.4.2 保存 39](#_Toc508613102)

[4.4.3 退出 40](#_Toc508613103)

[4.4.4 时间选择 40](#_Toc508613104)

[4.4.5 渔船选择 41](#_Toc508613105)

[4.4.6 航次编号 42](#_Toc508613106)

[4.4.7 捕捞区统计热度图 43](#_Toc508613107)

[4.4.8 港口分布图 44](#_Toc508613108)

[4.4.9 示意图操作 44](#_Toc508613109)

[4.5 本章小结 45](#_Toc508613110)

[5 总结与展望 46](#_Toc508613111)

[5.1 总结 46](#_Toc508613112)

[5.2 展望 47](#_Toc508613113)

[6 参考文献 48](#_Toc508613114)

# 引言

本章首先介绍了船舶轨迹数据的研究背景；其次，概述了基于VMS数据的国内外研究现状,以及轨迹数据研究成果；之后，总结了本文的研究内容、贡献和难点；最后，对全文的组织架构进行了简要说明。

## 研究背景与意义

我国位于亚洲东部，太平洋西岸，大陆海岸线1.8万公里，海洋国土面积约300万平方公里，南北纬横跨近40度。这样的海域环境蕴含了丰富的渔业资源，不仅丰富了食物的来源，更增加了劳动收入[[[1]](#endnote-2)]。但是近年来，我国近海渔业资源日趋枯竭，原因一方面在于市场需求不断扩大，渔民在利益驱使下“竭泽而渔”，严重破坏了海洋生态平衡[[[2]](#endnote-3)]；另一方面，渔民的捕捞行为依赖于口耳相传的经验，容易造成局部区域的过度捕捞，缺乏系统科学的指导[[[3]](#endnote-4)]。为了规范渔业生产，实现海洋资源的可持续性利用，中国农业部出台了一系列举措。单以伏季休渔为例，在2013年延长各海域不同类型渔船禁渔期为2至3个月之后，2017年初，再次调整禁渔期，且在原有基础上又普遍延长一个月[[[4]](#endnote-5)]。同时，2017年初规定，在“十三五”期间各省控制渔船船数和功率数，淘汰老旧木制渔船和过度捕捞船型，如船底拖网、三角虎网等[[[5]](#endnote-6)]。

在海洋航运和生产安全的需求下，渔业船舶监控系统（Vessel Monitoring System，VMS）应运而生。如今已经发展为集全球卫星定位系统、地理信息系统、电子海图、计算机网络通讯和数据库技术于一体的综合应用系统。主要由船载终端与地面基站两部分构成[[[6]](#endnote-7)]。船载终端负责对船舶进行卫星定位，并通过传感器采集船舶运行状态信息。地面基站负责收集、存储数据，并向船舶反馈实时信息。

船舶监控系统起源于葡萄牙，最初是为了监控渔业活动并保障船舶航行安全[[[7]](#endnote-8),[[8]](#endnote-9)]。2000年，欧盟委员会通过了规定所有24米以上船舶安装VMS终端的条例。至2012年1月1日普及至所有12米以上船舶[[[9]](#endnote-10)]。我国也在近十年时间基本完成了终端推广工作，如浙江海域的覆盖率为95.5%。

虽然VMS在不同地区的部署略有不同，但采集的数据基本都包含了时间、经纬度、瞬时速度、船艏向等字段。这些数据为渔业管理提供了可靠的船舶运行状态信息，不仅用于揭示不同海域的捕捞强度[[[10]](#endnote-11)]，也有助于分析渔业资源的时空变化[[[11]](#endnote-12)]。综上所述，借助于VMS数据对渔业资源时空变化进行定量化分析，对于我国渔业管理有着重要的意义。本文将以VMS轨迹数据为研究对象，识别渔船的捕捞行为和渔船航道，挖掘海洋捕捞规律，并用软件对算法进行实现。

## 研究现状

VMS轨迹数据研究涉及多个领域。比如Coro等[[[12]](#endnote-13)]按月计算了加拿大近海捕捞努力量；Fonseca等[[[13]](#endnote-14)]人通过对拖网渔船的VMS数据和渔获量进行比较，分析了渔场的变迁情况；Watson和Gerritsen等[[[14]](#endnote-15)，[[15]](#endnote-16)]通过计算捕捞强度分析了渔业活动对海洋环境和海洋生物资源的影响。这些研究多是根据渔船的捕捞行为进行分析，但VMS数据本身并没有区分渔船行为的信息。

直接解决方法是借助于辅助数据，如航海日志记录了船舶在出海过程中航行、捕捞和停泊等主要情况。这种方法可以粗粒度的对渔船行为进行分类识别。但是我国缺少对航海日志的数字化管理手段，无法利用这些数据。因此这种方法并不适用。

另一种较直接的方法是利用VMS数据的字段作为特征进行分类。比如，传统研究多以速度作为判断标准，通过对样例数据中捕捞和航行两种行为的分析，确定速度区间作为渔船行为的判断标准[3，[[16]](#endnote-17)-，[[17]](#endnote-18)，[[18]](#endnote-19)，[[19]](#endnote-20)，[[20]](#endnote-21)]。这种方法处理速度较快，但是Bertrand等[[[21]](#endnote-22)]指出，这样处理容易使渔船捕捞行为的数据点数量被高估。Lee等[[[22]](#endnote-23)]通过统计不同研究中的速度阈值，发现在不同海域没有一个统一的速度阈值范围能够准确区分同一作业方式渔船的状态，且不同作业方式渔船的速度船艏向等字段差异明显。

区分渔船行为还可以通过对VMS轨迹数据进行模型训练与分类。Joo等[[[23]](#endnote-24)]在研究秘鲁鳀鱼时，发现利用速度判断捕捞行为的方法高估了捕捞数据，提出使用人工神经网络进行模型训练，提高了渔船行为的识别准确率。对于高时空分辨率的北斗卫星数据，Zhang等[[[24]](#endnote-25)]提出可以利用统计学模型判断捕捞状态。但是，这种方法需要大量标记了渔船行为的数据对模型进行训练，不仅标记数据难以获得，而且算法时间开销较大。

在轨迹数据挖掘方面，根据Zheng等[[[25]](#endnote-26)]总结的人体、车辆轨迹研究，轨迹处理包括噪声处理、停止点检测、轨迹压缩、轨迹分段等内容。以轨迹压缩为例，按照轨迹的离线和在线可以把处理方法分为两种类型。Douglas等[[[26]](#endnote-27)]基于垂直欧氏距离（perpendicular Euclidean distance）提出了Douglas-Peucker算法对离线轨迹数据进行压缩。为了解决在线轨迹的压缩问题，Vitter等[[[27]](#endnote-28)]人提出了reservoir sampling算法。之后，Maratnia等[[[28]](#endnote-29)]提出了改进的sliding window和Before Open Window算法。Potamias等[[[29]](#endnote-30)]提出了基于物体运动速度和方向的在线轨迹压缩方法。

## 研究内容和贡献

渔业船舶的捕捞行为与渔业压力、环境保护、经济产量等一系列因素相关，但是VMS数据本身不能区分渔船的停靠、航行和捕捞行为。因此，识别渔船的行为是进行VMS数据研究的首要问题。本文提出了渔船捕捞行为识别算法，通过港口定位、航次划分，并利用轨迹压缩和数学形态学识别了渔船的捕捞行为。之后，计算了捕捞时间和捕捞距离与渔业产量进行比较，验证算法的准确率。并对捕捞热度图进行统计，分析了渔业资源的整体变化。

渔业管理研究者侧重于捕捞行为分析，往往忽略了捕捞行为之间的航行轨迹。它象征着捕捞轨迹区域之间的航道，有助于挖掘捕捞区域间的时空变化规律。所以，本文提出了渔船航道识别算法，利用渔船在捕捞区域间的航行轨迹以及捕捞热度图，通过迭代和投票方法对捕捞区进行分类，并得到了渔船的航道信息。通过比较不同年份相同季度的航道变化，分析了渔业资源的时空变化，以及渔民的捕捞行为。

为了便于渔业工作人员对渔业轨迹进行分析，本文设计并实现了船舶轨迹数据分析系统。通过将导入VMS数据，可以通过可视化窗口观察港口分布、渔船轨迹、航次和对应捕捞区、捕捞区统计热度图。方便渔业从业人员操作使用。

本文的主要贡献如下：

（1）在缺少航海日志的条件下从VMS数据中识别了渔船的捕捞行为，并提出利用数学形态学这种图像学方法对捕捞轨迹进行识别。

（2）提出了渔船航道识别算法，在没有明确航道的条件下，根据渔船轨迹抽象出航道信息，为渔船行为和渔业资源时空变化规律的分析提供了新的手段。

## 论文组织结构

本文共分为六个章节，各章内容安排如下：

第一章 引言。介绍了船舶轨迹数据的研究背景；概述了基于VMS数据的国内外研究现状,以及轨迹数据研究成果；总结了本文的研究内容、贡献和难点。

第二章 渔船捕捞行为识别算法。按照数据处理、港口定位、航次划分、捕捞行为识别的顺序，介绍了算法的各个模块。最后对算法进行了评估和分析。

第三章 渔船航道识别算法。详细介绍了投票和迭代过程。并根据航道变化对渔民的捕捞行为和渔业资源的时空变化进行了分析。

第四章 船舶轨迹数据分析系统。按照系统设计的三个层次——数据管理层、业务层、展示层的顺序，介绍了船舶轨迹数据分析系统。着重介绍了其中对渔船行为识别系统的改进。

第五章 总结与展望。对全文内容进行了总结，分析了未来的研究方向。

# 渔船捕捞行为识别算法

船舶监控系统（Vessel Monitoring System，VMS）最早是船舶通过无线广播来发送自己的位置数据给附近海域的其他船舶，起到避免船舶碰撞的目的。VMS数据主要包括当前时间、船舶的当前位置、速度、航向等信息。在渔业船舶配备了卫星终端后，渔业管理部门可以接收相关海域的VMS数据。于是，通过VMS数据，渔业管理部门可以获得渔业船舶活动的完整轨迹。但是，VMS数据本身不能区分渔业船舶的行为是停靠、航行还是捕捞，而渔业资源管理研究通常更注重渔业船舶轨迹中的捕捞行为发生的部分。因此，根据渔业船舶的VMS轨迹数据，区分其中的捕捞、航行和停靠部分是基于渔业船舶VMS数据的渔业资源管理研究首先要解决的问题。

虽然渔业船舶的航海日志部分记录了船舶航行、捕捞和停泊的主要情况，但是目前仍没有航海日志规范记录标准和数字化管理手段，因此不借助或不完全借助航海日志，基于VMS数据的渔业船舶行为分类是目前VMS数据分析的关键问题。

针对基于VMS数据的渔船行为分类问题，研究者们已经开展了研究工作。相关研究可归纳为两类：基于统计的方法和基于模型的方法。前者是利用渔船航行过程中总结的统计经验来区分不同行为，比如渔船在进行捕捞时速度会下降，并且通过折返运动拖曳渔网进行捕捞作业。Witt等[3]以英国附近海域约400艘不同类型的渔船为统计对象，所使用的VMS数据包含自2000年1月1日至2004年12月31日共5,788,188条记录，确定了速度区间[3,10]km/h为捕捞行为行为区间。Fork等[[[30]](#endnote-31)]为不同类型的渔船确定不同的速度阈值来确定捕捞行为：拖网渔船< 8kn，刺网渔船< 5kn。这类方法的处理速度快，但分类准确率不高，且不具备海域及渔船类别的通用性。

基于模型的方法使用分类模型完成基于VMS轨迹的训练及分类，从而确定渔船的捕捞行为。Walker等[[[31]](#endnote-32)]在研究金枪鱼围网渔船的VMS数据时，将渔船行为分为寻找、捕捞、停泊和航行等四种状态，并通过船艏向和航速两个字段数据训练隐马尔可夫模型（Hidden Markov Model, HMM）进行船舶行为分类。在训练过程中，使用了96,325条渔业船舶的航海日志记录作为标定数据训练模型，并用3,510条数据进行验证。Joo等[[[32]](#endnote-33)]在研究秘鲁鳀鱼围网作业时，收集了2000-2007年共347,901条VMS数据及对应航海日志，使用神经网络对其中97,877条数据进行训练，其余航海日志记录进行验证，真阳性达到76%。使用基于模型的渔业船舶行为分类方法优点在于有较高的分类准确率，但是需要大量标定好的数据进行模型训练；而且时间开销较大。

由于本文研究对象仅为我国浙江海域的渔业管理部门记录的拖网渔船VMS数据，没有与之相配的航海日志，因此基于模型的分类方法无法适用。此外，基于统计的方法虽然运行速度快，但识别精度不高。于是，本文以拖网渔船的捕捞行为特点为切入点，探索使用图像处理方法识别渔业船舶轨迹中的捕捞行为。

## VMS数据

本文处理的数据来自于浙江省温州市海洋渔业安全救助信息中心。在浙江海域内，单船拖网是最常见的生产方式之一，数据量大，便于研究。因此，本文对该船型的VMS数据进行分析，收集了自2014年4月1日起至2016年6月30日止共31条渔船2,595,004条记录。单船最多153,690条记录，最少31,950条记录。

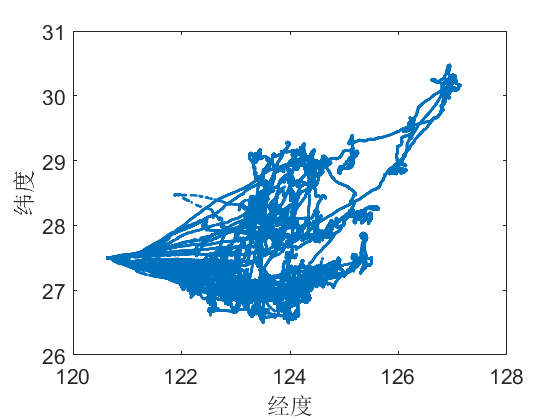
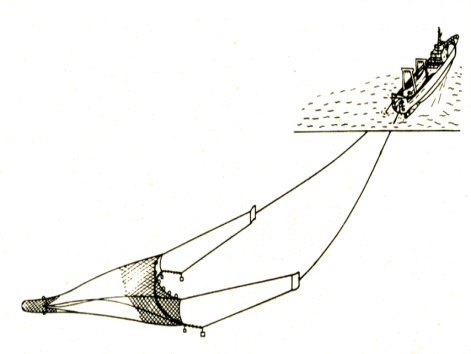


图2-1 单船拖网作业 图2-2 渔船轨迹图（终端ID：255368）

单拖渔船在进行捕捞作业时通过拖曳渔网在鱼群密集处往返低速航行（图2-1），反映密集在VMS轨迹数据上是轨迹连续且的区域。图2-2展示了设备终端ID为255368的VMS数据轨迹，包括了自2014年4月1日至2016年9月27日共134,622条记录。虽然可以找到个别的轨迹密集区域，但是由于叠加过多，难以分析。因此，需要对VMS数据进行分段，得到简明的轨迹。同时，可以通过迭代的方式，利用图像处理方法识别每一段VMS数据的捕捞行为。



图2-3 同一艘船的连续三个航次（终端ID：255368）

航次描述了渔船从离开港口出海捕捞直至返回港口的过程，常被用于渔船行为分析[3,24,[[33]](#endnote-34),[[34]](#endnote-35)]。因此，将VMS数据按照航次进行分段。图2-3展示了手动确认的三个航次轨迹，航次之间因港口而相互独立。但是VMS数据中并不包含港口信息，所以，在划分航次之前需要对港口的位置进行判断。

根据单拖渔船的作业特点，轨迹密集分布的区域被认为是捕捞轨迹区域，该区域内部的渔船活动即渔船捕捞行为。为了识别图2-3中红色标注的捕捞轨迹区域，本文先对航次进行轨迹压缩，降低了渔船航行轨迹对识别的干扰；然后使用数学形态学完成识别任务。航次轨迹进出捕捞轨迹区域的端点就是渔船捕捞行为的起点和终点，至此完成了渔船捕捞行为的识别。

## 算法结构

本文把研究问题拆分为四个子问题，如图2-4所示，根据这些问题，算法中依次设计对应的模块进行解决：

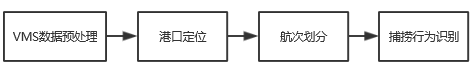


图2-4 渔船捕捞行为识别算法

（1）VMS数据预处理模块：由于卫星定位误差以及海上恶劣的通讯环境，VMS数据中会存在少量异常记录。本模块通过剔除异常数据，为后续研究还原渔船的真实轨迹。

（2）港口定位模块：港口定位是航次划分的前提。虽然可以借助于官方统计的港口列表进行判断，但是渔船在实际活动中经常停靠的一些小型码头或锚泊地难以统计。本模块根据VMS数据的特征得到了港口坐标，为划分航次提供了依据。

（3）划分航次模块：航次是研究渔船行为的基本单位。本模块利用港口定位结果按航次对VMS数据进行分段，将渔船捕捞行为识别问题的处理对象简化为一个航次的轨迹。

（4）捕捞行为识别模块：首先，对航次进行轨迹压缩，减少图像处理过程中渔船航行行为的干扰。然后，利用拖网渔船在捕捞作业时拖曳渔网往返运动这一特点，先将航次轨迹转化为图像，然后通过数学形态学识别图像上的捕捞轨迹区域。最终，将进出捕捞轨迹区域的位置即就是捕捞行为的起点和终点作为算法的输出。

## VMS数据预处理

VMS数据主要包括渔船ID、时间、经纬度、瞬时速度、船艏向等字段信息。对于原始数据，先按照时间顺序对数据进行排序。

轨迹中的异常数据源自于卫星定位误差以及海上恶劣的通信环境，我们借鉴了Yuan等在研究计程车轨迹中的方法[[35]](#endnote-36)，通过设定平均速度阈值的方式，剔除异常数据。具体而言，计算相邻采样点AB间的平均速度，剔除平均速度大于的数据。由于我国东海海域渔船航速小于30kn，设阈值。这里两点间取球面距离，计算方法如下。

其中，D表示采样点AB间的球面距离，R表示地球半径（取值6372公里），A点坐标，B点坐标。最终从原始数据中剔除异常数据5224条记录，占总数的0.20%

原始数据中存在数据缺失，本文没有通过插值对数据进行填补，以免对捕捞行为的识别造成干扰。

## 港口定位

渔船在进行捕捞作业时往返拖曳渔网，体现在轨迹图上就是局部连续折返、轨迹密集的区域。本文利用这一特征识别捕捞行为。图2-5展示了设备终端ID为255368的单拖渔船自2014年4月1日至2016年9月27日的轨迹图。海岸线附近由于多个航次叠加已经难以区分捕捞轨迹区域。由于轨迹相互叠加严重，因此需要通过港口位置对VMS数据进行航次划分。

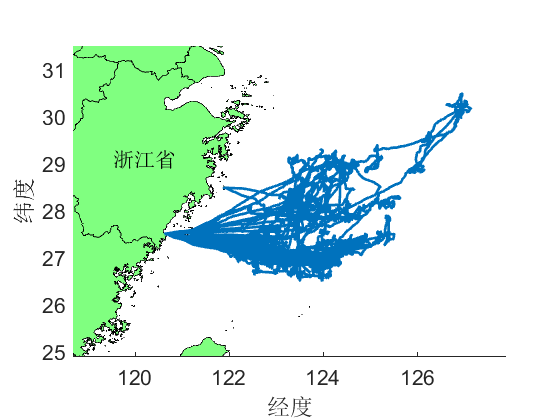


图2-5 渔船轨迹（终端ID：255368）

官方统计的港口信息可以作为参考，但不包含一些小渔港和锚泊地。为了在不依赖外部信息的条件实现港口定位，本节提出的“坐标驻留法”根据VMS经纬度信息得到港口坐标。

相邻两条记录的经纬度坐标值相同被称为“坐标驻留”。虽然有误差的因素存在，但仍然可以表明在这段时间内船舶近似静止。坐标驻留法就是利用这一点进行港口定位。虽然渔船出海夜间进行锚泊时也容易发生坐标驻留现象，但可以与港口进行区分。一方面，同样是锚泊，海上风浪对渔船的影响比港口大，坐标驻留现象在港口发生的概率远大于海上。另一方面，港口一般固定不变而海上锚泊位置随机性较大，在港口区域发生坐标驻留的频度远远高于海上任一区域。这种方法优点是判定条件简单，缺点同样明显：a.关闭VMS终端对判定结果有决定性影响，渔民为了省电甚至可能会在入港前提前关闭终端；b.难以识别访问次数少的港口,易与海上锚泊混淆。

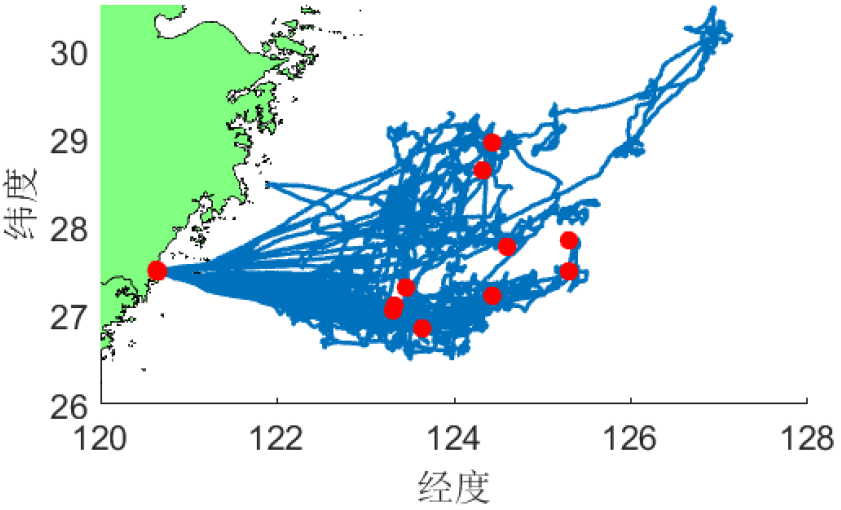


图2-6 “坐标驻留”现象分布（终端ID：255368）

图2-6是某单港口渔船VMS数据轨迹，共存在48次“坐标驻留”现象，位置如图上红点所示。经过判断，其中在港口区域发生38次，明显高于其他区域。

在找到坐标驻留点后，对图像划分网格并统计每个格子中坐标驻留点的数量，用阈值进行筛选，保留下的网格中心点作为港口坐标。在这里，网格大小和阈值选择都会影响到港口定位的结果，为了避免海上的坐标驻留点误判为港口，本文采用了较严格的约束条件。定义网格大小为0.1′经度×0.1′纬度，约0.01平方海里；当某个网格中统计的坐标驻留点数量大于总数的5%，将其中心点作为港口。

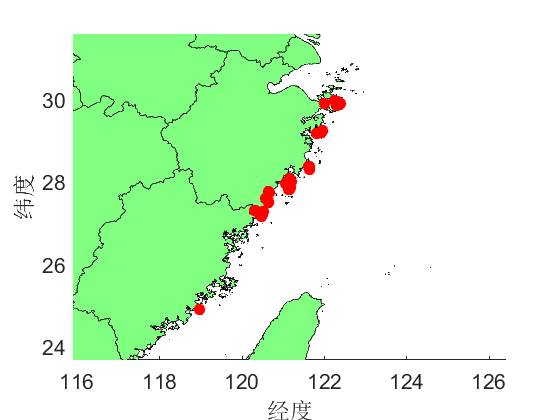


图2-7 港口识别结果

这样做会导致一些港口无法识别，于是将所有渔船的港口识别结果汇总为一个港口信息表，共识别86个不同坐标的港口（图2-7），作为下一节划分航次的依据。

需要指出的是，随着船舶监控系统的完善，渔民主动关机的现象必然会大大减少。而且随着VMS数据的进一步积累，“坐标驻留法”本身处理速度快的特点将进一步体现，并且准确率也会大幅提高。

## 航次划分

航次是渔船捕捞活动的基本单位，刻画了渔船从离开港口出海捕捞到返回港口的过程。本模块通过港口信息表，对VMS数据进行分段，使捕捞行为识别模块的处理对象从叠加的VMS数据简化为单一的航次数据。

具体步骤是：初始化一幅图像令全部像素值为0，每个像素表示1′经度×1′纬度的区域。将港口坐标映射到图像上令值为1。同理，将VMS轨迹数据映射到该图像上，标记所有离开港口像素和进入港口像素的数据。从离开港口到进入港口的轨迹称为一个航次。

需要注意的是，对于图2-3的例子（终端ID：255368），共划分了99个航次，远远超出了实际情况。这是因为渔船路过港口A到达港口B的轨迹，被视为先抵达港口A，后抵达港口B，从而将AB之间的轨迹作为一个航次。由于6小时内无法完成一个航次，已知平均数据采样间隔是3分钟，6小时平均产生120个数据点。经过统计，划分出的99个航次中有81个航次的数据量小于120个数据点。通过这一约束条件筛选得到18个航次作为航次划分模块的结果。

## 捕捞行为识别

拖网渔船在捕捞作业中通过拖曳渔网在渔区内做折返运动完成捕捞，我们利用这种特性，通过识别轨迹上的折返区域——本文称其为捕捞轨迹区域，完成对渔船捕捞行为的识别。捕捞行为识别模块首先对航次进行轨迹压缩，然后通过数学形态学识别轨迹图像上的捕捞轨迹区域，得到捕捞行为起始和结束点作为输出。

### 轨迹压缩

为了避免轨迹重叠严重对识别捕捞轨迹区域的影响，上文提出了航次划分模块对轨迹进行简化。从而可以很明显的指出一个航次中捕捞轨迹区域的位置。比如，图2-6上用红色圈注的区域。但是在实际处理中，蓝色圈注区域由于航线的交叉造成局部轨迹较为密集，容易被错误识别成捕捞轨迹区域。因此需要在保留捕捞轨迹区域特点的前提下，减少蓝色圈注区域的数据点数量。

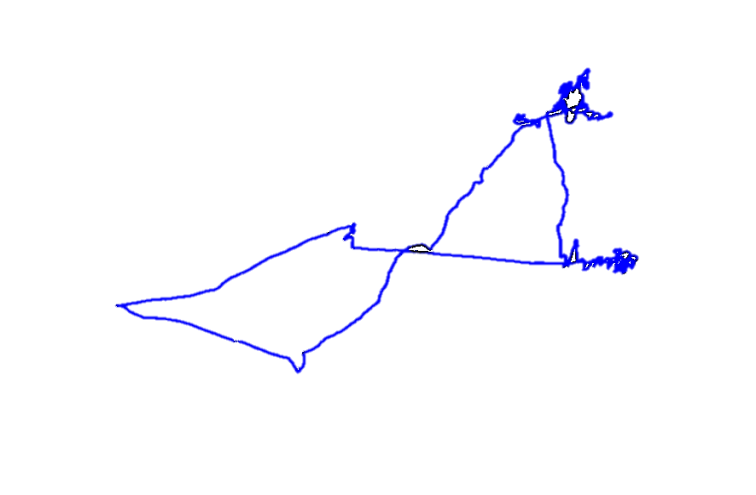
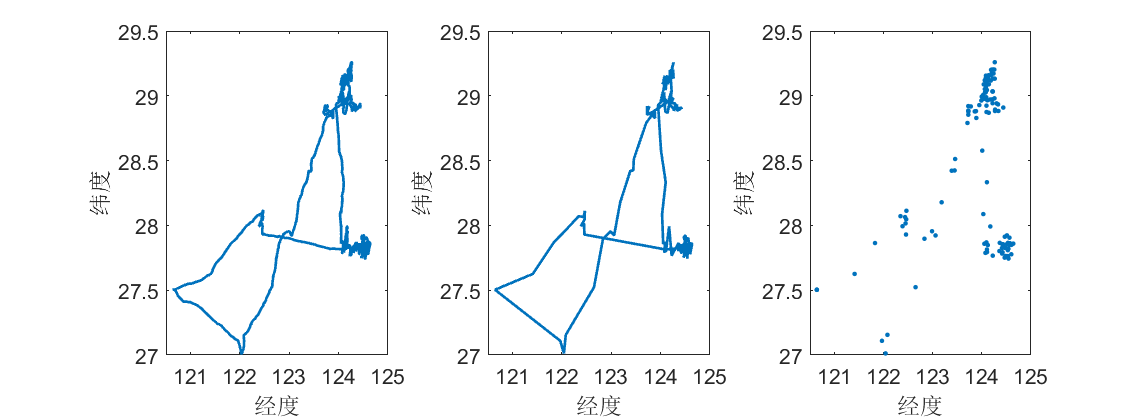


图2-8 根据轨迹密集区域判断捕捞轨迹区域

使用Douglas-Peucker算法对VMS数据进行轨迹压缩，它将垂直欧氏距离作为压缩误差，并与提前设好的阈值进行比较，作为轨迹压缩的终止条件。图2-8是对航次数据压缩前后的效果。图2-8（1）是压缩前的折线图，共7,183条记录；图2-8（2）是阈值取0.03得到的压缩后的折线图，保留了126条记录。

（1）压缩前折线图 （2）压缩后折线图 （3）压缩后点图

图2-9 利用Douglas-Peucker算法压缩轨迹数据的前后对照

由于Douglas-Peucker算法是利用轨迹的转折程度进行压缩。当渔船沿直线航行时，数据压缩率较大；反之，当渔船轨迹曲折时，数据压缩率较低。对应到实际轨迹上，航行行为数据大幅减少，而捕捞行为数据保留比例较大。图2-9（3）的点图与图2-9（2）折线图的数据相同，可以看出，图2-8蓝色圈注区域由于轨迹交叉造成局部轨迹复杂的问题已经基本解决。

### 数学形态学

数学形态学是一种图像处理方法，包含四种基本运算：膨胀、腐蚀、开启和闭合。它利用一种被称为结构元素的算子对图像进行代数运算，适用于各种图像形状和结构的分析和处理。本模块利用单拖渔船的捕捞作业特点，对航次轨迹进行数学形态学处理，得到对应的捕捞轨迹区域，并以进出该区域的位置作为模块的输出，之间的轨迹就是待识别的捕捞行为。

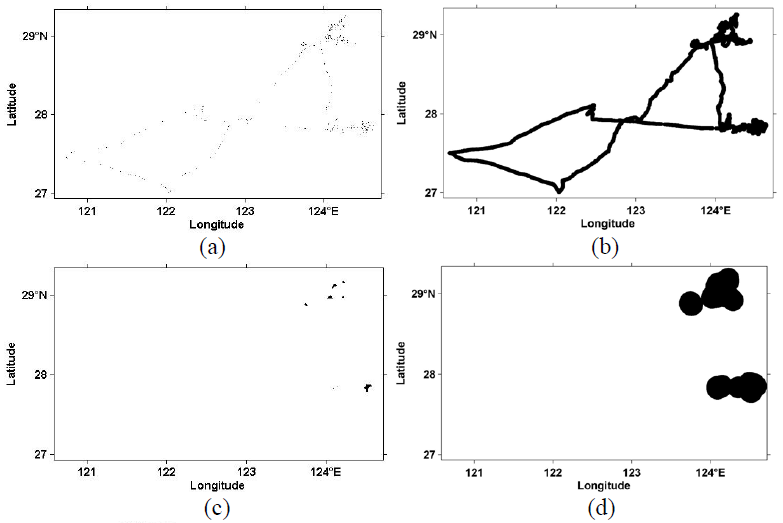


图2-10 利用形态学方法识别捕捞区的流程

具体实现是通过结构元素对图像进行“膨胀”与“腐蚀”操作。令结构元素为“圆盘”，半径为4。首先，将压缩后的轨迹点映射到图像上（图2-10（a））；通过腐蚀操作，将密集轨迹点连接成团（图2-10 (b)）；然后进行扩张操作，剔除航行轨迹，保留下的区域即为捕捞轨迹区域的“核心”（图2-10 (c)）；最后再进行腐蚀操作，将核心放大，即为识别的捕捞轨迹区域（图2-10 (d)）。

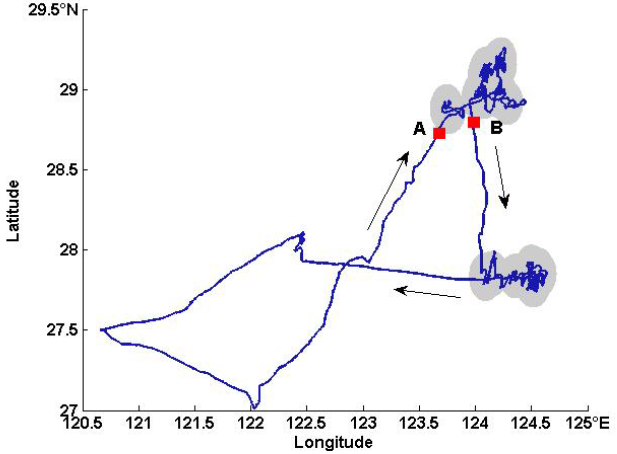


图2-11 完成捕捞区域识别的航次轨迹

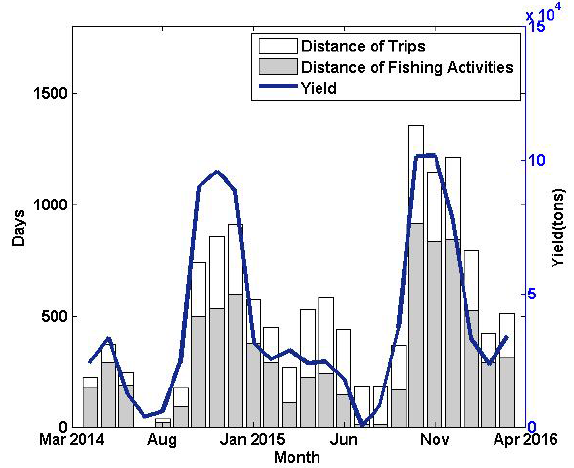
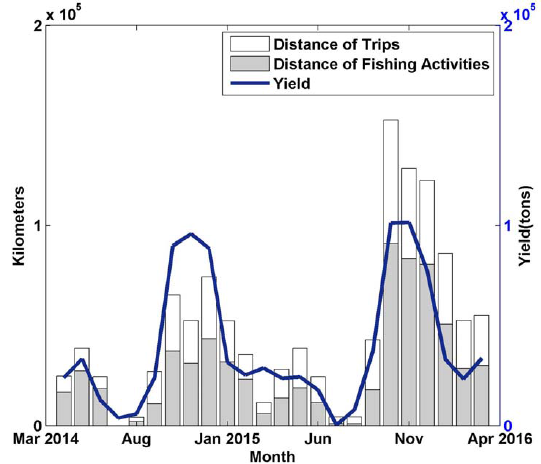
在成功识别捕捞轨迹区域后，就可以完成捕捞行为的识别：对于一个航次而言，我们把在捕捞轨迹区域内的轨迹视为渔船的捕捞行为，其余轨迹视为航行行为。图2-11中灰色区域是该航次通过数学形态学识别出的捕捞轨迹区域，AB两点是捕捞行为的起点和终点。最终将形如AB点对的序列作为算法的输出结果。

## 结果验证与分析

以上四个模块共同构成渔船捕捞行为识别算法，为了验证算法的准确性，将统计的捕捞时间和捕捞距离与渔业产量进行了对比。对于不同季度的捕捞区热度图，分析捕捞区域的变化规律。

### 结果验证

由于渔业产量与捕捞时间和捕捞距离息息相关，可以通过计算两者的相关性验证捕捞行为的识别准确率。本文根据渔业数据的发布周期，按月统计了渔船的捕捞时间和捕捞距离，与浙江省温州市海洋渔业救助中心记录的渔业产量数据进行对比。

（1）捕捞时间与渔业产量 （2）捕捞距离与渔业产量

图2-12 捕捞时间/捕捞距离与渔业产量数据对比

图2-12中，蓝色折线为2014年4月至2016年3月浙江省渔业产量，灰色条形图为（1）捕捞时间（2）捕捞距离。从图中可以看出，统计结果与渔业产量数据呈现明显的正相关：禁渔期（7月-8月）捕捞累计时间短、产量低；冬捕期和春节旺季（10月-次年1月）捕捞累计时间长，产量高。捕捞作业总时间与渔业产量的Pearson相关系数为89.80%，捕捞作业总距离与渔业产量的Pearson相关系数为82.11%。由于总航行时间和总航行距离的计算本身就有偏差，最终得到的强相关性验证了捕捞作业行为被准确识别。

### 结果分析

从图2-12中，还可以看出，与2014年冬季捕捞数据（蓝色折线图的第一个峰值）相比，2015年冬捕数据（蓝色折线图的第二个峰值）并没有明显变化，说明两年产量相似，但捕捞时间和捕捞距离明显增大。这说明了捕捞难度加大，渔民需要投入更多的时间且捕捞成本提高。

统计一段时间内的捕捞轨迹区域，可以得到渔船在该时间段内的捕捞强度分布。



（a）2014第二季度 （b）2014第三季度 （c）2014第四季度



（d）2015第一季度 （e）2015第二季度 （f）2015第三季度



（g）2015第四季度 （h）2016第一季度 （i）2016第二季度

图2-13 2014年第二季度到2016年第二季度的捕捞强度热度图

图2-13是按季度统计的捕捞强度热度图，可以看出捕捞作业确实存在东移趋势。比如从2014年至2016年的第二季度捕捞区热度（图2-13（a）（e）（i））可以明显看出这一趋势。与图2-12的分析结果一致，说明渔民需要航行至更远的海域进行捕捞作业。这种行为一方面反映出近海捕捞资源不足，渔业资源枯竭；另一方面预示着捕捞成本的增加，可能会造成渔民生活水平下降、水产品价格上升、海洋环境进一步恶化。

## 本章小结

本章提出了渔船捕捞行为识别算法。按照系统实现的顺序，依次介绍了数据处理、港口定位、航次划分、捕捞区识别四个模块，并对识别结果进行了验证和分析。通过与渔业经济数据对比并计算Pearson相关系数可以看出，算法的输出结果准确的识别了渔船的捕捞行为。统计的捕捞强度强度图能够反映出捕捞作业东移的趋势，这种变化表明近海渔业资源枯竭和渔民捕捞成本提高。

# 渔船航道识别算法

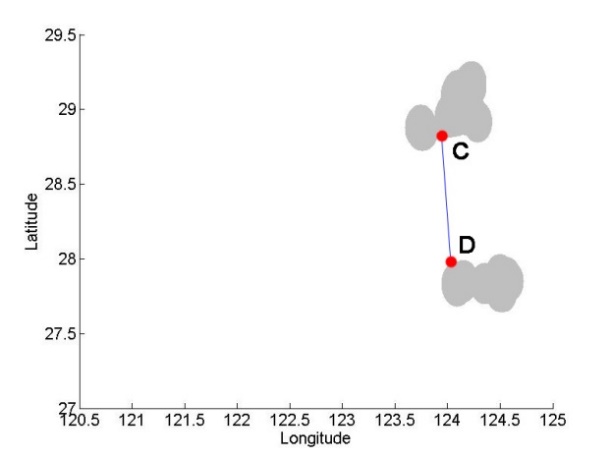
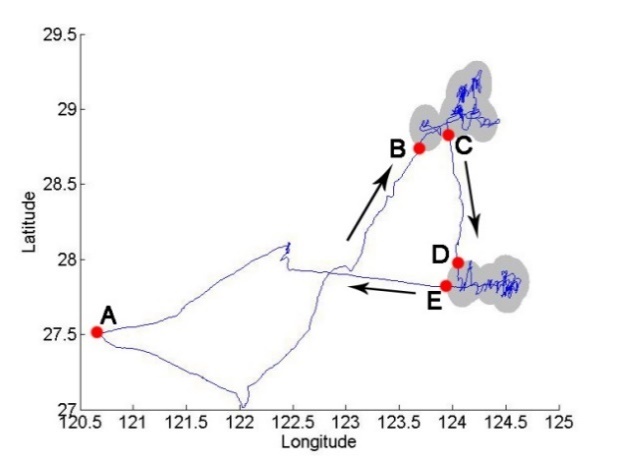
VMS数据研究多侧重于渔船的捕捞行为，根据捕捞活动分析捕捞努力量、生态环境变化等要素。然而，这忽略了渔船的航行行为，尤其是捕捞作业之间的航行轨迹。它描述了渔船的航道信息，为研究捕捞区域间的联系、渔业资源时空变化提供了新的途径。同时，这些航行轨迹有助于捕捞区域的划分，能够对渔业捕捞做出更清晰的指导。

之前的研究者们已经对捕捞作业区域的变化和联系进行了一些分析。Campbell等[[[36]](#endnote-37)]在研究英国西南海域海洋保护与可再生资源时，借助VMS数据对拖网、刺网、延绳钓等多种渔船类型2005至2008年的捕捞努力量分布进行了可视化分析，利用空间变化指数（index of difference in spatial pattern）对捕捞努力量的变化进行了比较。Russo等[[[37]](#endnote-38)]在利用意大利海域拖网渔船的VMS数据分析捕捞压力的经济学指标时，将海域按照3km×3km划分网格，按月统计了网格中的数据点数量，对每个网格按时间的变化序列计算了格里菲斯时空指数（Griffith’s spatio-temporal index），以此作为区域的时空变化规律。然而，这些研究或者对捕捞区域整体进行评估，或者将海域划分网格分别计算评价指标，缺乏灵活的捕捞区域分类方法。

由于海上不存在固定的航道，且缺乏直接验证的手段，在此前没有人对渔船航道进行研究。本文在渔船捕捞行为识别算法的基础上，提出了渔船航道识别算法。借助捕捞区域间的航行轨迹和捕捞热度图，探索了捕捞区域的合理分类方法，并确定渔船航道。

## 研究思路

在城市计算研究中，主干道路是划分城市区域的重要参考标准。在研究捕捞区域分类时借鉴这一思路，将捕捞行为之间的航行轨迹信息作为分类的依据。但是，由于捕捞行为识别的误差，以及渔船航行的灵活多变等原因，这些轨迹过于复杂，增加了分类难度。因此，需要对它们进行简化。



（1）渔船捕捞行为识别算法的输出 （2）航行线段

图3-1 对捕捞区之间的航行轨迹进行抽象

通过渔船捕捞行为识别算法，得到了捕捞的起始终止点，如图3-1（1）中的B点C点，以及D点E点。捕捞行为之间的轨迹是指C点到D点的轨迹。简化的方法是只保留它的端点，将轨迹CD抽象为图3-1（2）所示的线段CD，称该线段为航行线段。它表示线段两端点应处于不同的捕捞区域。

将按季度统计的捕捞热度图（图2-13）作为分类对象，与对应时间段内统计的航行线段叠加，如图3-2所示。

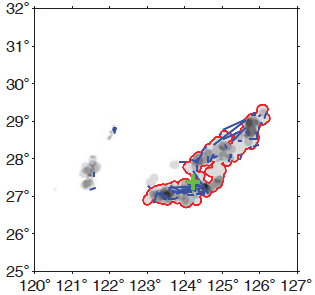
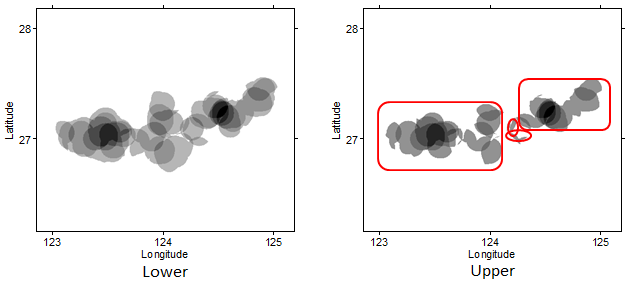


图3-2 2014年第二季度捕捞热度图与航行线段

灰度图是2014年第二季度的捕捞热度图，蓝色线段是该时间段内统计的航行线段，红色曲线包络的是一个相互连通的捕捞区域。由于热度图的数值本身就可以作为分类的依据，这里根据连通性，将捕捞热度图粗略地分为七个区域。可以看出，航行线段不仅仅分布在不同区域之间，还分布于区域内部。以红色曲线勾勒的区域为例，大量的航行线段两端点落于该区域内部。根据航行线段两端点应处于不同的捕捞区域这一条件，将航行线段作为捕捞区域分类的依据。



（1）下层： (2)上层：

图3-3 相邻图层某个出现“一对多”的区域

对于实际生产而言，如果两个捕捞区域相邻，可以把它们视为一个整体。所以，将捕捞区域划分成相互独立的个体对渔民捕捞作业更具有指导意义。为了使分类后的捕捞区彼此分离，本文利用热度图的权值将图3-2中的连通区域划分成多个区块。设置参数 ，比较相邻两图层——热度权值的区域和的区域。如图3-3所示，是2014年第二季度捕捞热度图的局部。称保留区域较大的为下层（图3-3（1）），较小的为上层（图3-3（2））。两图层之间存在三种区域变化方式：（Ⅰ）“零对一”：上层的空白位置对应到下层是一个独立的区块；（Ⅱ）“一对一”：上层中的区块对应到下层范围变大，但仍保持独立性；（Ⅲ）“一对多”：上层的两个或多个区块对应到下层是一个相互连通的区域。图3-3是情况（Ⅲ）的一个实例，上层中的四个区块在下层中合并为一个整体。将下层中（Ⅰ）（Ⅱ）两种情况视为独立的区域。

对于情况（Ⅲ），比较区块任意组合的方案，找出最符合航行线段含义一个作为该区域的分类方案。这个比较过程通过一种投票的方式实现，根据得票数确定对应区域的分类方案。最终，迭代的确定整个捕捞热度图的分类方案。分类方案各区域之间的航行线段就是渔船的航道。

## 算法结构

算法分为两部分，第一部分为渔船捕捞行为识别算法，不再赘述。取按季度统计的捕捞热度图以及渔船在捕捞行为之间的航行轨迹作为第二部分“迭代投票”的输入，如图3-4所示。

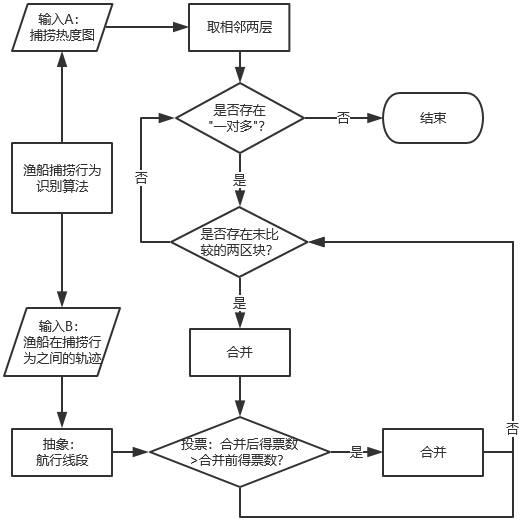
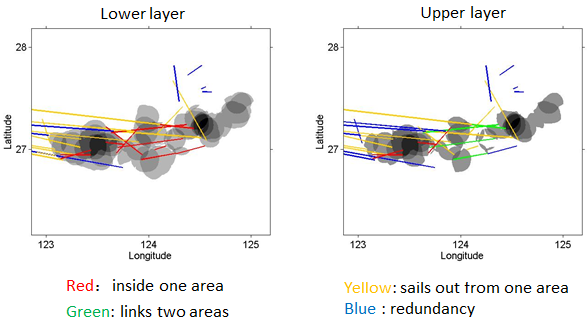


图3-4 渔船航道识别算法

算法的第二部分是通过迭代的方式运行的。（1）从捕捞热度图中提取出相邻的两个图层，遍历所有“一对多”的情况。（2）通过投票比较区块合并前后的得票数，选取最高票数的方案作为捕捞热度图的分类方案。（3）将航行线段中连接分类方案两区域的子集抽象为渔船航道。

## 投票

上文提到航行线段可以作为区域划分的依据。为了比较不同区块组合的合理性，设计了一种投票机制，将航行线段作为选票，计算各种组合方案的得票数，从而确定最终的分类方案。



（1）下层投票 （b）上层投票

图3-5 四种选票。“红色”反对票，“绿色”支持票，

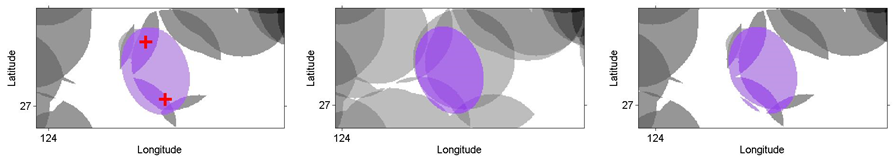
“黄色”部分支持票，“蓝色”弃权票

根据航行线段两端落点的不同，共有四种不同的选票，如图3-5所示：两端在不同区块内的用绿色表示，称为支持票，是对两个区块划分方法的肯定，得票数记作+2；两端在同一个区块内的选票用红色表示，称为反对票，得票数记作-2；只有一端在区块内的用黄色表示，是对其中一个区块的肯定，称为部分支持票，记作+1；两端均不在区块内的用蓝色表示，称为弃权票，记作0。所以，对于一个分类方案进行投票，就是根据落点判断每一个选票的类型，将对应数字相加求和。图3-5是对图3-3的案例进行投票，图3-5（1）得票数为-11，图3-5（2）得票数为11。所以根据航行线段，图3-5（2）的分类方案要比图3-5（1）更佳。

## 迭代流程

整个算法的第二部分是通过迭代的方式运行的，遍历所有“一对多”的情况，并比较不同组合的得票数，逐一确定对应区域的最佳分类方案。

如果对全部组合进行投票，组合数将随着区域数量上升呈指数型增长，大大增加了计算量。实际上，当三个区域排成一条直线时，可以只对相邻两区域进行分析。根据这一条件对组合数进行剪枝：取各个捕捞区的几何中心点两两连线，按线段长度排序；根据线段长度从小到大的顺序比较对应两区域合并前后的得票数；若合并后得票数较高，则将合并后区域视为一个整体，不再分开讨论其他分组情况。对于n个捕捞区域，这种处理方案最少只需要进行（n-1）次比较，即n个捕捞区域合并为一个整体；最坏情况只进行(n-1)n/2次比较，对应着n个捕捞区域全部独立的情况。在算法实现过程中，利用并查集算法简化判断归属的时间开销，且并没有出现时间开销过大的情况。这是因为当n比较大时（最多为14个），其中包含了大量的形如图3-3中间部分的小捕捞区，这些区域往往合并前后得票数相同。为了简化分类结果，将捕捞区尽量合并为一个整体，对于得票数相同的情况，判断为合并而非独立。



1. （2） （3）

图3-6 两捕捞区的合并策略

合并的具体方法如图3-6所示。在对两捕捞区进行合并时，以两个捕捞区的几何中心点作为椭圆焦点并连线，延长线较短一侧作为长轴，作椭圆连接这两个捕捞区如图3-6（1）。将椭圆与下一层捕捞区取交集如图3-6（2），计算结果即为两捕捞区域的合并方式如图3-6（3）。这种合并方法虽然存在与其他捕捞区域产生交集的可能，但胜在计算简便，能够满足合并的需求。

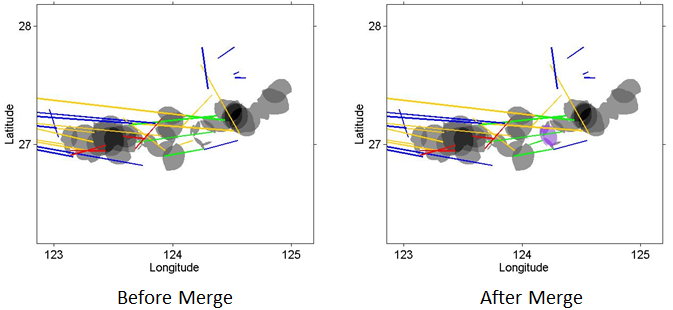


图3-7 比较合并前后的得票数

对于每次合并的尝试，比较合并前后的得票数大小。如图3-7所示，是对图3-6合并前后进行了投票，图3-7（1）的得票数为11，图3-7（2）的得票数为12，所以合并两个区块。接下来，继续按照线段长度从小到大的顺序进行迭代，依次比较捕捞区合并前后的得票情况，重复上述步骤。确定最终的分类方案。

## 渔船航道

通过上面介绍的算法，实现了对捕捞热度图的细粒度划分。在此基础上，利用航行线段统计渔船在各个区域之间的转移关系，可以得到渔船航道。

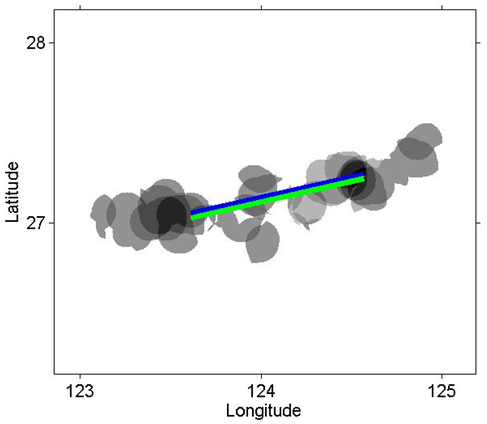
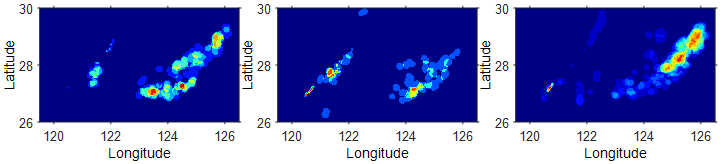


图3-8 渔船航道

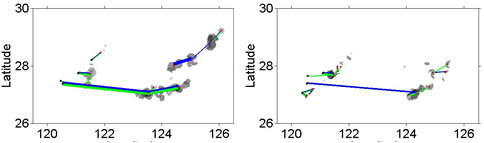
如图3-8所示，灰色部分是通过渔船航道识别算法确定的分类方案：从最初的四个区域变为两个区域。统计“支持票”，作图上蓝绿色线段，并称其为渔船航道。线段端点定位在对应区域的几何中心点上。蓝色表示自西向东，绿色表示自东向西，线段宽度表示船次。图3-8两个区域之间自西向东航行有三个船次，自东向西有三个船次。



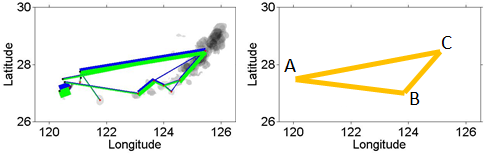
（1）2014 （2）2015 （3）2016

图3-9 连续三年第二季度捕捞区热度图

如图3-9是2014年至2016年第二季度的捕捞热度图，在上一章中分析了捕捞区域向东北移动的趋势。通过对渔船航道的统计，能够建立与陆上交通相似的渔船航道网络。以此为基础，可以进一步分析渔业资源的变化以及渔民的捕捞行为。



（1）2014年第二季度 （2）2015年第二季度



（3）2016年第二季度 （4）渔船航道形状

图3-10 连续三年渔船航道变化

图3-10展示了2014-2016年第二季度的渔船航道信息。可以将其抽象为一个三角形的航道网络（图3-10（4））：A点表示港口，BC两点表示两个捕捞热点区域。在2014年第二季度，大部分渔船出海沿着AB对应航道进行捕捞作业；2015年第二季度船次整体明显减少；2016年第二季度，只有少量船只沿AB进行捕捞作业，大部分的船只都选择直接航行到C区域进行捕捞。相比于B区域，选择更远的C区域进行捕捞，这充分说明B处渔业资源减少。同时，2014年便开始在C点捕捞的渔民冒险和开拓意识较强，而2016年仍在B点捕捞的渔民相对而言比较保守。

## 本章小结

通过对渔船在捕捞行为之间的航行轨迹的观察，发现它有助于捕捞区域的分类以及渔船航道的确定。据此，本章提出了渔船航道识别算法。它包含两部分，第一部分是第二章介绍的渔船捕捞行为识别算法，得到捕捞行为之间的航行轨迹和捕捞区热度图这两组参数。第二部分是通过迭代方式对捕捞热度图的分类方案进行投票，确定最终了分类结果，并根据航行线段统计渔船航道。根据2014-2016年第二季度渔船航道的流量变化，可以指出捕捞区域的变化，并且区分不同渔民的捕捞策略。

# 软件实现

前两章介绍了渔船捕捞行为识别算法和渔船航道识别算法。前者能够仅依靠经纬度信息识别渔船的捕捞行为，后者对捕捞区进行细粒度分类并得到了渔船航道信息，为分析渔业资源的变化规律和渔民行为提供了新的依据。为了方便渔业从业人员对VMS数据进行分析，本文设计并实现了船舶轨迹数据分析系统。通过导入VMS数据，可以观察渔船的港口识别、划分航次、捕捞区域等结果。

## 系统架构

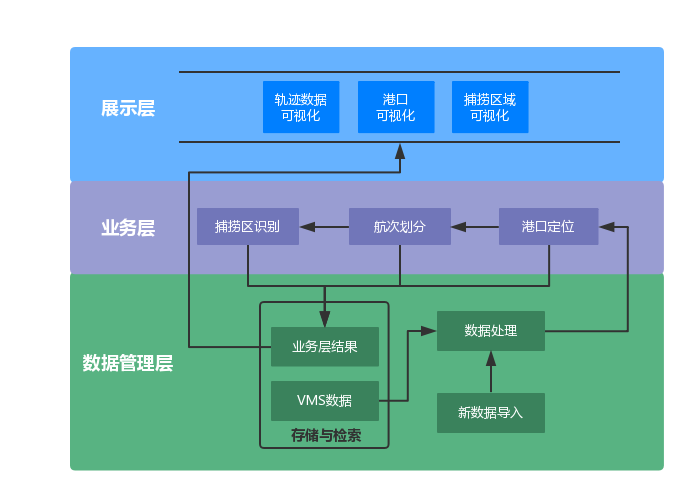


图4-1 船舶轨迹数据分析系统的架构设计

如图4-1所示，将船舶轨迹数据分析系统分为三个层次：数据管理层、业务层和展示层。数据管理层设计了整个系统的文件存储结构，完成了对原始数据的预处理操作。业务层实现了渔船捕捞行为识别算法的主要功能，包括港口定位、航次划分和捕捞区识别三个模块。展示层主要从轨迹数据可视化、港口可视化、捕捞区域可视化三个角度展示软件运算结果。

## 数据管理层

数据管理层作为整个系统的底层，实现了对数据的基本操作：定义文件系统结构，并完成了对VMS数据的预处理。

### 文件系统结构

船舶轨迹数据分析系统是一个可执行文件，双击启动，存放位置没有特殊要求。其余文件结构如图4-2所示。系统根目录“\”下存放有三个项目：

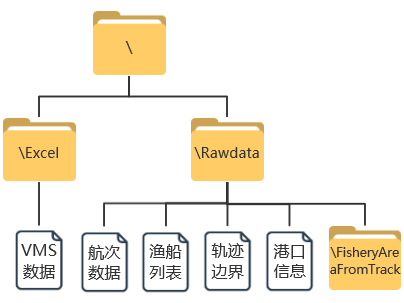


图4-2 文件系统结构示意图

“\”为系统根目录，用于存放数据，包括原数数据和系统运行产生的中间结果。Excel文件夹：存放Excel格式的VMS数据文件，作为整个系统的输入，必须与软件同处于根目录下。Rawdata文件夹：存放业务层执行过程的中间结果，作为展示层的输入。在业务层执行时如果不存在该文件夹则自动生成；如果已经存在，且包含文件，则替换原文件。子文件航次数据存储了VMS数据去除异常值后的航次信息，渔船列表存储了渔船的名称及对应编号，轨迹边界用于调整可视化窗口大小，港口信息用于展示港口识别结果，子文件夹FisheryAreaFromTrack存储了所有航次对应的捕捞区识别结果。

### 数据预处理

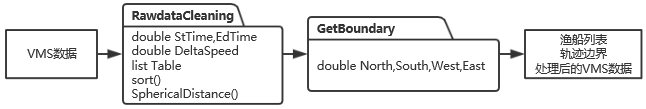


图4-3 数据处理流程

数据管理层针对VMS数据进行一系列基本处理，如图4-3所示。输入是VMS数据，输出是记录渔船名称和编号的渔船列表，轨迹边界，以及处理后的VMS数据。通过调用RawdataCleaning()中的sort()函数对VMS数据按时间字段排序，SphericalDistance()函数用于计算球面距离，删除VMS数据中的速度异常值。同时，记录轨迹的起始时间StTime和终止时间EdTime，以及渔船信息列表Table。调用GetBoundary()，记录轨迹边界以备后续处理。

## 业务层

业务层是整个软件的核心，主要包含港口定位、航次划分和捕捞区识别三个模块。详细类图如图4-4，本节将详细介绍各个函数之间的业务逻辑关系。

### 港口定位模块

在渔船捕捞行为识别算法中，我们仅根据经纬度数据来进行港口定位，如图4-4所示。

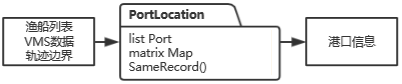


图4-4 港口定位流程

输入是渔船列表、处理后的VMS数据和轨迹边界，输出是对港口的识别结果。通过调用PortLocation()函数，利用SameRecord()实现“坐标驻留”的识别方法，将港口信息存储在列表Port中，并利用矩阵Map去重。

### 航次划分模块

根据港口定位结果可以进行航次划分，如图4-5所示

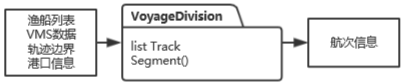


图4-5 航次划分流程

输入包括渔船列表、处理后的VMS数据、轨迹边界和港口信息，输出包含航次信息的VMS数据。通过调用VoyageDivision()的Segment()函数，利用港口信息对VMS数据进行分段，并将每个航次的起止位置存储在列表Track中。

### 捕捞区识别模块

对于一个航次，根据渔船的捕捞作业特点利用轨迹压缩和数学形态学识别捕捞轨迹区域，如图4-6所示。



图4-6 捕捞区识别流程

输入包括渔船列表、处理后的VMS数据、航次数据和轨迹边界。输出是每个航次对应的捕捞轨迹区域识别结果。函数DouglasPeucker()的参数epsilon=0.03，通过PerpendicularDistance()计算垂直欧氏距离评估轨迹压缩率，实现对航次轨迹的压缩。数学形态学函数Morphology()利用了imdilate()和imerode()两个算子完成捕捞轨迹区域的识别。

## 展示层

展示层主要包含三个部分：轨迹数据可视化模块、港口可视化模块和捕捞区域可视化模块。他们将业务层的输出结果通过可视化的方式展示在界面上。系统的初始化界面如图4-7所示：

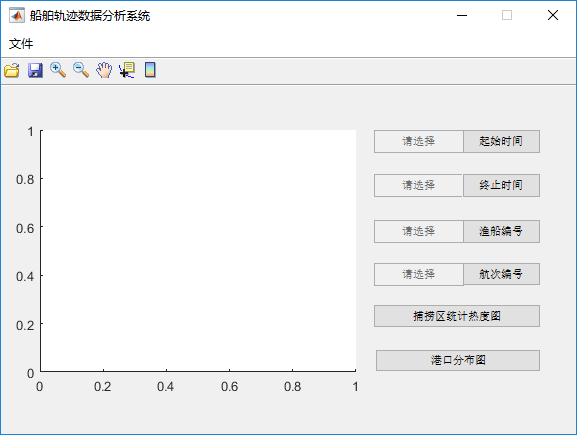


图4-7 船舶轨迹数据分析系统的主界面

主界面包括菜单栏、工具栏、窗口四部分。菜单栏的文件菜单主要功能包括：导入数据、保存图片、退出。工具栏的工具按钮包括导入数据、保存、放大、缩小、平移、数据游标、插入颜色栏。窗口分为左右两个部分，左侧是数据可视化窗口，用来展示轨迹图和捕捞热度图。右侧是操作面板，通过选择不同的选项，可以在可视化窗口上观察到不同的结果。

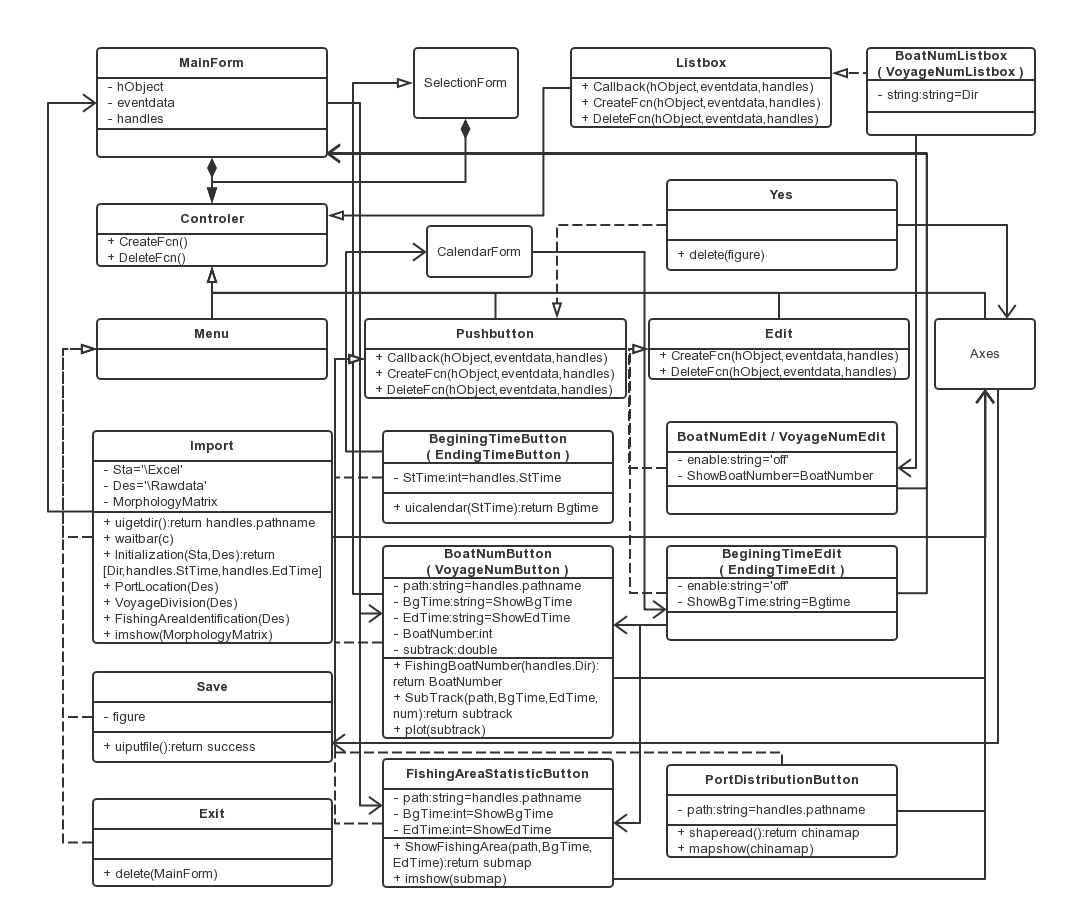
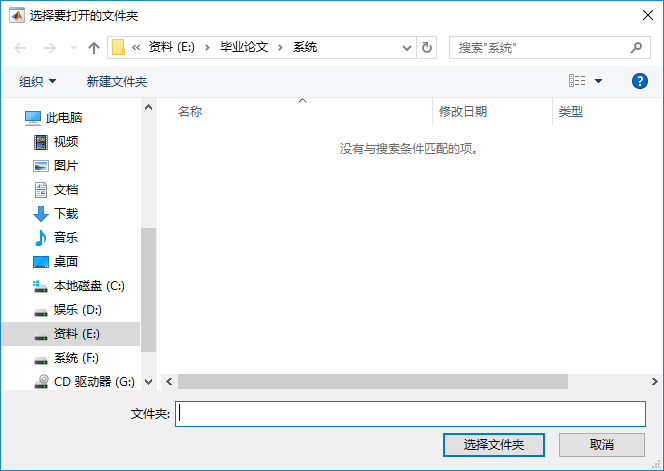
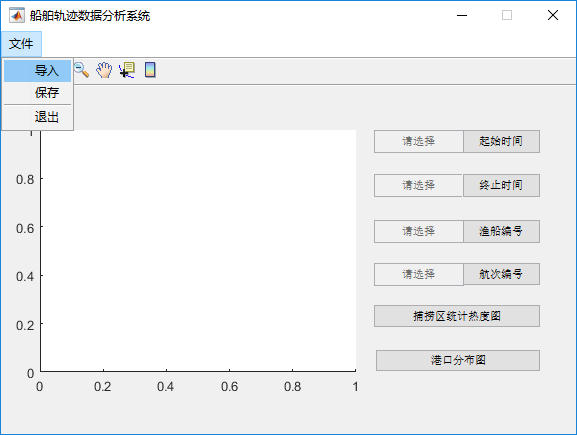


图4-8 展示层类图

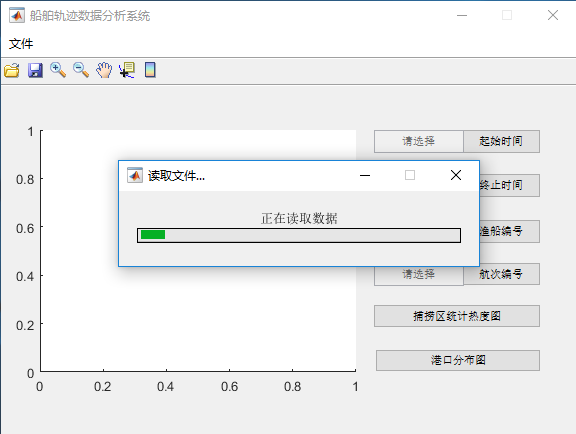
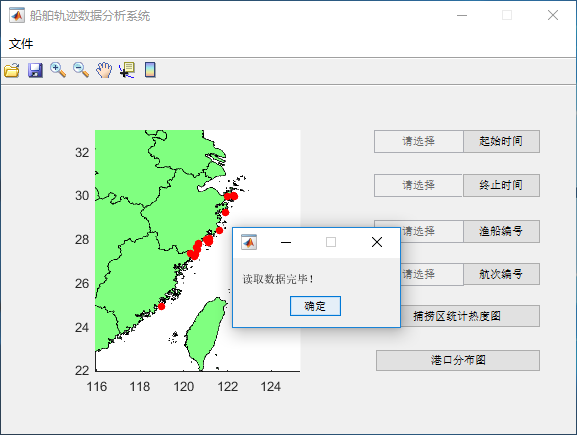
展示层的类图设计如图4-8所示。主界面（MainForm）包含了多个类型的控件（Controler）：菜单（Manu）、按钮（Pushbutton）、编辑框（Edit）、图片框（Axes）等。所有控件都包含有基本的触发事件：Callback()，CreateFcn()和DeleteFcn()。控件之间通过一个核心变量handles进行参数传递。业务层的主要功能通过导入菜单项（Import）执行。展示功能通过按钮（Pushbutton）的一系列实例完成。

### 导入数据

双击可执行文件启动程序后，第一步就是导入数据。在这个过程，将完成数据处理以及业务层各个模块的任务，具体工作如图4-9：



（1）文件菜单下的导入按钮 （2）路径选择对话框

（3）开始导入 （4）导入完成

图4-9 导入数据的操作流程

依次点击文件菜单、导入按钮（图4-9（1），工具栏的第一个按钮“导入”功能相同），弹出路径选择对话框，找到VMS数据所在文件夹，点击选择文件夹按钮进行数据导入（图4-9（2））。数据导入过程以进度条的形式展示，平均耗时约194s（图4-9（3））。最终导入完成后，窗口左侧展示了港口定位的识别结果，并弹出对话框提示读取数据完毕，点击确定进行下一步操作。

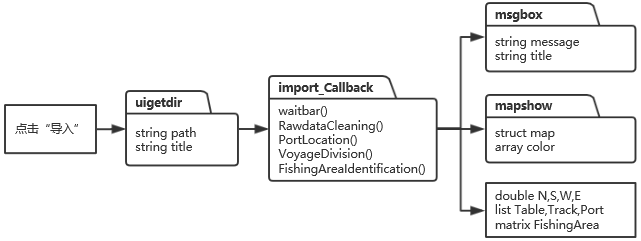


图4-10 导入数据的业务逻辑

业务逻辑如图4-10所示：

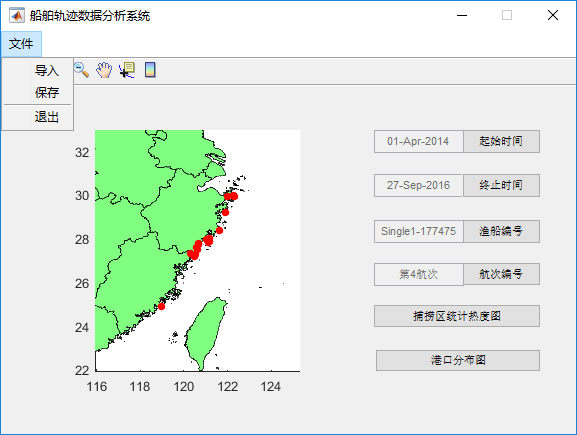
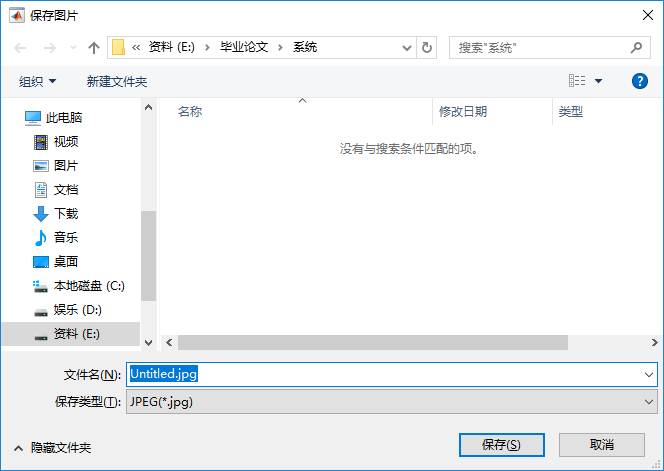
（1）点击“导入”按钮，调用uigetdir()函数，弹出路径选择对话框，选择VMS数据的路径并赋值给参数path。

（2）调用import\_Callback()函数，依次完成数据处理RawdataCleaning()、港口定位PortLocation()、航次划分VoyageDivision()和捕捞轨迹区域识别FishingAreaIdentification()。调用waitbar()函数显示进度信息，最终完成数据导入任务。中间产生的结果保存在路径path的\Rawdata目录下，包括渔船列表Table、轨迹边界（东西南北）、港口信息Port、航次数据Track和航次对应的捕捞轨迹区域FishingArea。

（3）完成数据导入后，调用msgbox()函数提示导入数据完成。同时，调用mapshow()函数在窗口左侧结合地图展示港口定位结果。

### 保存

窗口左侧展示的示意图可以保存到指定目录下，具体操作如图4-11：

（1）文件菜单下的保存按钮 （2）保存图片对话框

图4-11 保存的操作流程

点击文件菜单下的保存按钮（图4-11（1），工具栏的第二个按钮“保存”功能相同），弹出保存图片对话框，确定保存路径、修改文件名称并选择文件类型后，点击“保存”按钮完成保存（图4-11（2））。

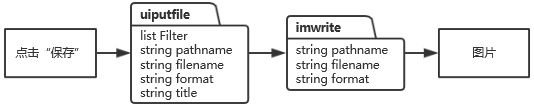


图4-12 保存的业务流程

业务逻辑如图4-12所示：

（1）点击“保存”按钮，调用uiputfile()函数弹出保存图片对话框，列表Filter限制保存类型，确认后得到保存路径pathname、文件名filename和文件类型format等信息。

（2）调用imwrite()函数，按照保存路径pathname、文件名filename和文件类型format对主界面窗口左侧示意图进行保存。

### 退出

点击文件菜单下的退出按钮，关闭主界面窗口。



图4-13 关闭的业务流程

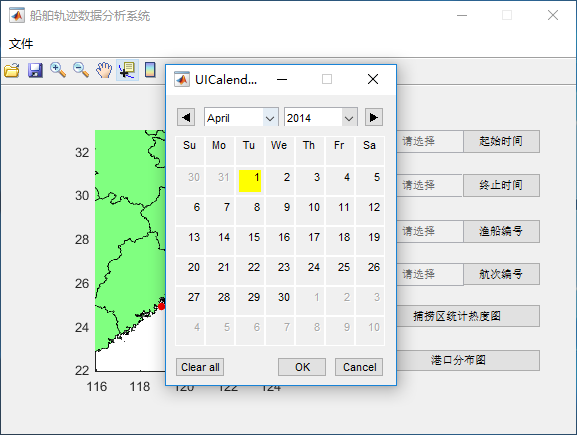
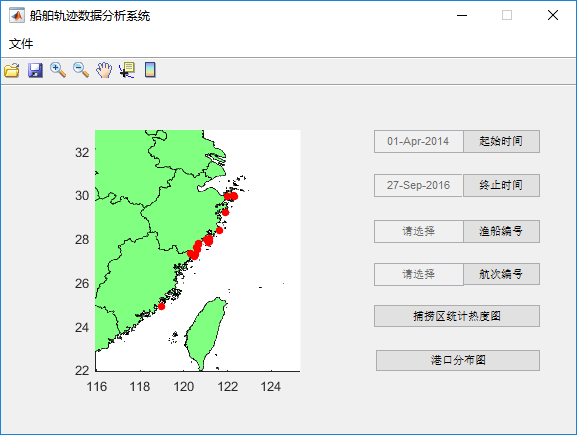
业务流程如图4-13所示：

（1）点击退出按钮，调用clear()函数，清除系统在运行过程中产生的临时变量。

（2）调用close()函数关闭主界面并终止进程，完成系统退出。

### 时间选择

时间选择是观察渔船轨迹数据和捕捞区统计热度图的前提。通过选择起始和终止时间，可以观察渔船在这段时间内的轨迹，或者是在这段时间内所有捕捞区识别结果叠加的热度图。具体操作如图4-14：

（1）在UICalendar界面选择时间 （2）时间选择完成效果

图4-14 时间选择的操作流程

点击起始时间按钮（或终止时间按钮），弹出时间选择窗口，时间精确到天。默认时间为VMS数据最早（或最晚）的记录时间，用黄色标记。图4-14（1）所示了默认时间是2014年4月1日。起始时间和终止时间选择完成后，会在按钮左侧的文本框中显示（图4-14（2））。



图4-15 时间选择的业务逻辑

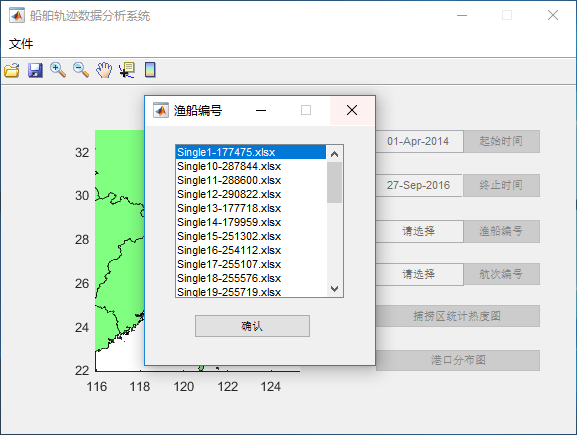
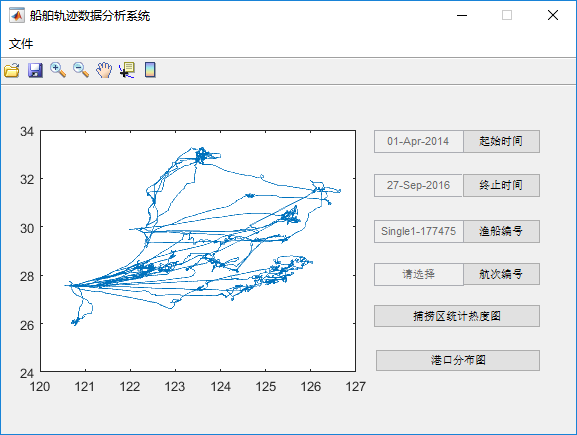
业务逻辑如图4-15所示：

（1）点击“起始时间”或“终止时间”按钮，调用uicalendar()函数，弹出时间选择对话框，初始默认值为StTime或EdTime，选择年月日并确认。若未选择时间，则返回为空值。函数waitfor()使得在退出时间选择窗口后执行后续操作。

（2）返回结果若为日期，显示在按钮左侧的文本框中；若为空值，在文本框中显示“请选择”。

### 渔船选择

通过渔船编号的选择，我们可以观察渔船的VMS数据轨迹，如图4-16所示

（1）选择渔船编号 （2）渔船的VMS数据轨迹

图4-16 渔船选择的操作流程

点击“渔船编号”按钮，弹出渔船编号窗口，从列表中选择要查看轨迹的渔船（图4-16（1））。点击确认后，主界面左侧示意图将显示渔船在起止时间段内的VMS数据轨迹（图4-16（2））。若未选择起止时间，即起始时间和终止时间都显示“请选择”。可以直接观察全部的轨迹数据。

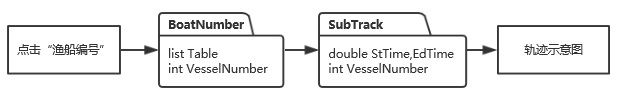


图4-17 渔船选择的业务逻辑

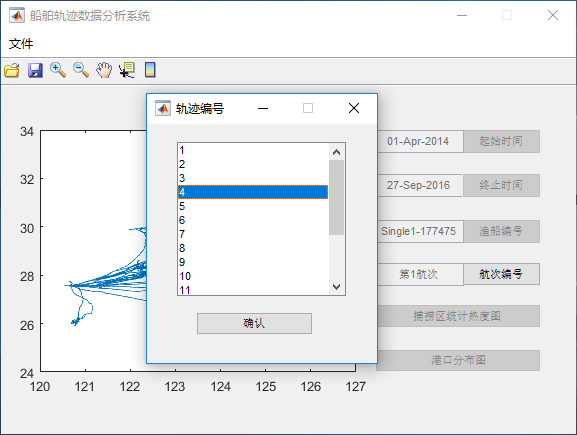
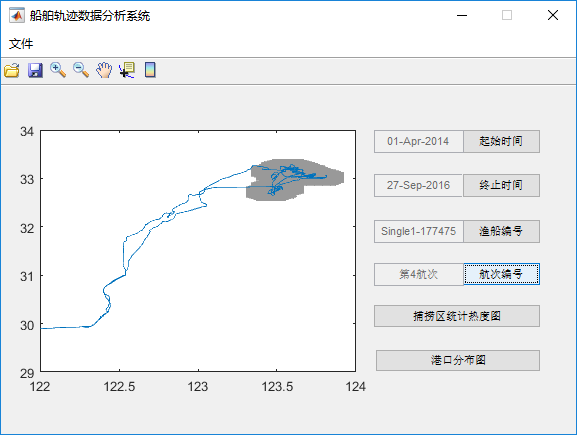
业务逻辑如图4-17所示：

（1）点击“渔船编号”按钮，调用BoatNumber()函数弹出渔船编号窗口，从渔船列表Track中选择其一后点击确认，窗口关闭并返回所选渔船编号VesselNumber。

（2）调用SubTrack()函数，按照起止时间StTime和EdTime截取对应时间段的VMS数据轨迹，并显示在主界面窗口左侧的示意图上。

### 航次编号

航次编号选择的前提是完成渔船编号的选择。通过航次选择，可以观察该渔船对应航次的轨迹以及识别出的捕捞轨迹区域，如图4-18所示

（1）选择航次编号 （2）航次轨迹及对应捕捞区

图4-18 航次选择的操作流程

点击“航次编号”按钮，弹出航次编号窗口，从列表中选择要查看的航次编号（图4-18（1））。点击确认后，主界面左侧示意图将显示该船所选航次轨迹及识别出的捕捞区（图4-18（2））。

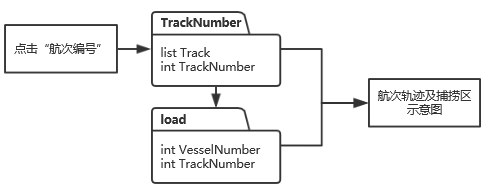


图4-19 航次选择的业务逻辑

业务逻辑如图4-19所示：

（1）点击“航次编号”按钮，调用TrackNumber()函数弹出航次编号窗口，从航次列表Track中选择其一后点击确认，窗口关闭并返回航次编号TrackNumber。

（2）根据渔船编号VesselNumber和航次编号TrackNumber，调用load()函数导入航次数据及对应的捕捞轨迹区域识别结果，显示在主界面窗口左侧的示意图上。

### 捕捞区统计热度图

根据所选择的起止时间，我们可以观察到不同的时间段捕捞区的统计热度图，如图4-20所示：

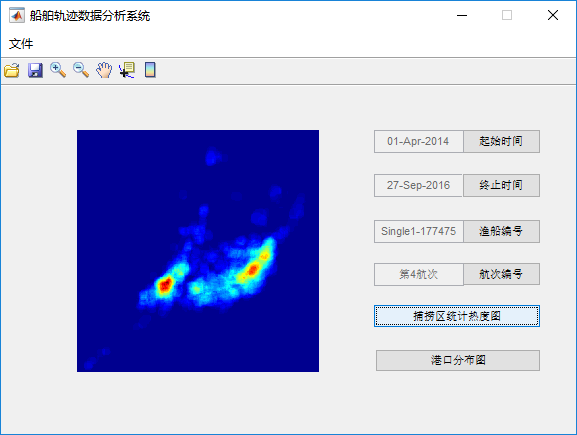
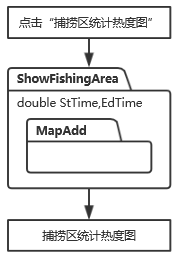
 

图4-20 捕捞区统计热度示意图 图4-21 捕捞区统计的业务逻辑

点击“捕捞区统计热度图”按钮，主界面窗口左侧显示对应时间段的捕捞区统计热度图（图4-20）。

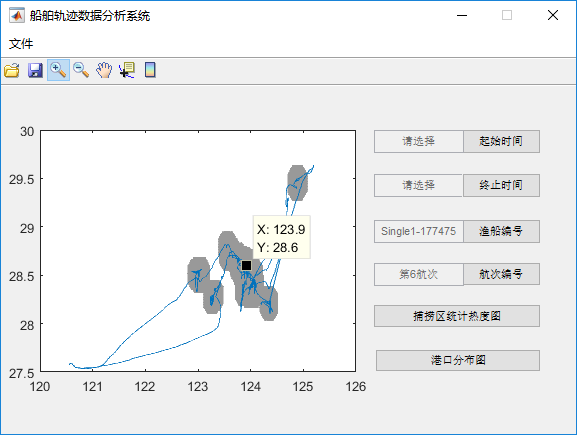
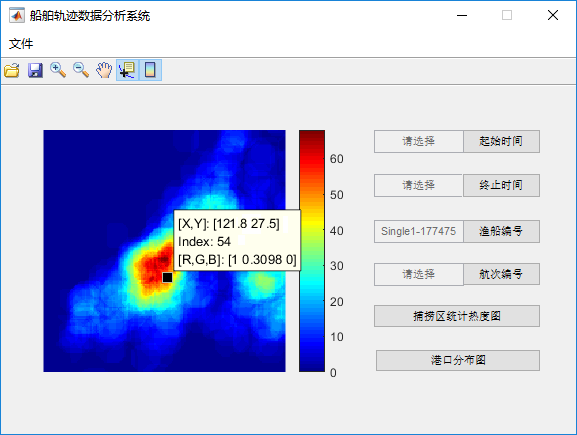
业务逻辑如图4-21所示：点击“捕捞区统计热度图”按钮后，调用ShowFishingArea()函数，函数内通过递归调用MapAdd()函数对每艘渔船在起止时间StTime和EdTime之间的捕捞轨迹区域识别结果进行叠加。显示在主界面窗口左侧的示意图上。若起止时间未选择，即起始时间和终止时间都显示“请选择”，则显示全部的捕捞区叠加结果。

### 港口分布图

点击“港口示意图”按钮，左侧示意图显示与数据导入完成后的示意图一致。详情参考“导入数据”。

### 示意图操作

工具栏自第三个按钮起是一系列对窗口左侧示意图进行处理的功能按钮，包括放大、缩小、平移、数据游标、插入颜色栏等。如图所示

（1）观察航次及对应捕捞区 （2）观察捕捞区统计热度图

图4-22 示意图放大、缩小、平移、添加数据游标、插入颜色栏

点击“放大”、“缩小”、“平移”按钮，可以对示意图简单观察。点击“数据游标”按钮后，点击图像，可以看到对应数据。对于轨迹数据，数据游标可以查看轨迹点的坐标（图4-22（1））。对于捕捞区热度图，数据游标不仅可以得到对应点坐标，还可以查看热度数值大小和RGB颜色（图4-22（2））。颜色栏能够将热度图上的颜色与热度数值对应。

## 本章小结

为了方便渔业从业人员从VMS数据中对港口、轨迹、捕捞区域进行可视化分析，本章设计并实现了船舶轨迹数据分析系统。它的核心是渔船捕捞行为识别算法。系统共分为三个层次：数据管理层、业务层、展示层。数据管理层主要负责文件系统的定义以及数据预处理；业务层实现了核心算法；展示层通过起止时间、渔船编号、航次编号、捕捞区统计热度图、港口分布图按钮，借助可视化工具观察港口、轨迹及捕捞区。

# 总结与展望

本章总结了全文的研究工作，并分析了基于VMS数据的轨迹挖掘研究今后的发展趋势。

## 总结

基于船舶监控系统（VMS）的轨迹数据研究自部署之日起一直受到国际海洋研究者的关注，虽然我国在这个领域起步较晚，但近年来，随着系统终端推广工作日益完善，已经初步完成了数据积累工作。本文在此基础上开展VMS渔船轨迹数据分析与挖掘工作。提出了渔船捕捞行为识别算法和渔船航道识别算法，分析了捕捞区域的时空变化以及渔民的捕捞行为。设计并实现了渔船轨迹数据分析系统，方便了渔业从业人员对VMS轨迹数据的观察、分析和研究。

本文的主要工作如下：

（1）对基于VMS渔船轨迹数据的研究背景、发展现状进行综述，指出捕捞行为的识别有待于进一步研究。

（2）提出了渔船捕捞行为识别算法。首先，在不依靠航海日志等辅助数据的条件下，利用轨迹数据“坐标驻留”的特点分析判断港口位置，并在此基础上将VMS数据按航次分段研究。然后，利用轨迹压缩和数学形态学得到渔船捕捞轨迹区域，最终完成捕捞行为的识别。最后，通过与浙江省温州市海洋渔业救助中心记录的渔业产量数据进行对照，计算Pearson相关系数验证了捕捞行为识别的准确率。

（3）提出了渔船航道识别算法。它由渔船捕捞行为识别算法和迭代投票分类两部分组成。利用前者输出的捕捞行为之间的航行轨迹，对季度捕捞热度图进行分类，并抽象出渔船航道。在此基础上，分析了浙江东部海域渔业资源随时间向东北转移的时空变化规律，以及渔民的捕捞行为特点。

（4）设计并实现了船舶轨迹数据分析系统。系统分为数据管理层、业务层、展示层三个层次。数据管理层定义了文件存储结构并对VMS数据预处理；业务层的核心是渔船捕捞行为识别算法；展示层通过界面和业务逻辑的设计，方便渔业从业人员观察不同时间段的港口、轨迹和捕捞区域识别结果。

## 展望

VMS数据分析仍有较大研究空间。相较于陆上的轨迹分析研究，VMS轨迹分析关注度低且手段较为单一，难以同时将天气、洋流、船舶轨迹、生物习性等多方面知识结合。致使大部分的研究都只能针对局部海域进行分析，难以适用于不同海域不同船舶类型。

目前为止，对于捕捞区定位的研究已经有很多成果，对渔业资源的分布变化情况已经有了一个大概的认识，可以开展下一步更深层次的数据挖掘工作。本文提出的渔船航道识别算法就是一次尝试，在此基础上可以对渔业资源的变化进行更细致的分析，此外还可以对渔民的捕捞作业进行行为分析，对渔业经济进行预测。

参考文献

1. [] 陈曦浩, 毛雄斌, 雷科,等. 基于渔捞日志的舟山渔场单拖网渔获量时空特征分析[J]. 浙江海洋学院学报(自然科学版), 2013, 32(3):202-207. [↑](#endnote-ref-2)
2. [] 王洋. 舟山渔场外侧海域单拖渔船轨迹模拟及其秋冬季渔获量估算与时空特征分析[D]. 浙江海洋学院, 2014. [↑](#endnote-ref-3)
3. [] 郭建兴. GPS渔船监控系统解决方案探讨[J]. 中国渔业经济, 2006(2):51-53. [↑](#endnote-ref-4)
4. [] 中华人民共和国农业部渔业局.农业部办公厅关于做好2017年海洋伏季休渔工作的通知[Z].2017-2-20 [↑](#endnote-ref-5)
5. [] 中华人民共和国农业部渔业局.农业部关于进一步加强国内渔船管控 实施海洋渔业资源总量管理的通知[Z].2017-1-16 [↑](#endnote-ref-6)
6. [] 陈曦浩, 毛雄斌, 雷科,等. 基于渔捞日志的舟山渔场单拖网渔获量时空特征分析[J]. 浙江海洋学院学报(自然科学版), 2013, 32(3):202-207. [↑](#endnote-ref-7)
7. [] Witt M J, Godley B J. A Step Towards Seascape Scale Conservation: Using Vessel Monitoring Systems (VMS) to Map Fishing Activity[J]. Plos One, 2007, 2(10):e1111. [↑](#endnote-ref-8)
8. [] Davies A J, Roberts J M, Hall-Spencer J. Preserving deep-sea natural heritage: Emerging issues in offshore conservation and management[J]. Biological Conservation, 2007, 138(3–4):299-312. [↑](#endnote-ref-9)
9. [] EC. Commission Implementing Regulation (EU) no.404/2011 of 8th April 2011. Laying down detailed rules for the implementation of Council Regulation EC No 1224/2009 establishing a community control system for ensuring compliance with the rules of the Common Fisheries Policy. Official Journal of the European Union; 2011.L112/1. [↑](#endnote-ref-10)
10. [] Gerritsen H, Lordan C. Integrating vessel monitoring systems (VMS) data with daily catch data from logbooks to explore the spatial distribution of catch and effort at high resolution[J]. Crop Science, 2011, 30(3):493-500. [↑](#endnote-ref-11)
11. [] Walker E, Rivoirard J, Gaspar P, et al. From forager tracks to prey distributions: an application to tuna vessel monitoring systems (VMS)[J]. Ecological Applications, 2015, 25(3):826-833. [↑](#endnote-ref-12)
12. [] Coro G, Fortunati L, Pagano P. Deriving fishing monthly effort and caught species from vessel trajectories[C]// Oceans. IEEE, 2013:1 - 5. [↑](#endnote-ref-13)
13. [] Fonseca T, Campos A, Afonso-Dias M, et al. Trawling for cephalopods off the Portuguese coast—Fleet dynamics and landings composition[J]. Fisheries Research, 2008, 92(2):180-188. [↑](#endnote-ref-14)
14. [] Watson R A, Cheung W W L, Anticamara J A, et al. Global marine yield halved as fishing intensity redoubles[J]. Fish & Fisheries, 2013, 14(4):493-503. [↑](#endnote-ref-15)
15. [] Gerritsen H D, Minto C, Lordan C. How much of the seabed is impacted by mobile fishing gear? Absolute estimates from Vessel Monitoring System (VMS) point data[J]. Ices Journal of Marine Science, 2013, 70(3):523-531. [↑](#endnote-ref-16)
16. [] Jonsen I, Myers R, James M. Identifying leatherback turtle foraging behaviour from satellite telemetry using a switching state-space model[J]. Marine Ecology Progress, 2007, 337(12):255-264. [↑](#endnote-ref-17)
17. [] Dinmore T A, Duplisea D E, Rackham B D, et al. Impact of a large-scale area closure on patterns of fishing disturbance and the consequences for benthic communities[J]. Ices Journal of Marine Science, 2003, 60(2):371-380. [↑](#endnote-ref-18)
18. [] Deng R, Dichmont C, Milton D, et al. Can vessel monitoring system data also be used to study trawling intensity and population depletion? The example of Australia's northern prawn fishery[J]. Canadian Journal of Fisheries & Aquatic Sciences, 2005, 62(2):611-622. [↑](#endnote-ref-19)
19. [] Murawski S A, Wigley S E, Fogarty M J, et al. Effort distribution and catch patterns adjacent to temperate MPAs[J]. Ices Journal of Marine Science, 2005, 62(6):1150-1167. [↑](#endnote-ref-20)
20. [] Walter J F, Hoenig J M, Gedamke T. Correcting for effective area fished in fishery-dependent depletion estimates of abundance and capture efficiency[J]. Ices Journal of Marine Science, 2007, 64(9):1760-1771. [↑](#endnote-ref-21)
21. [] Bertrand S, Burgos J M, Gerlotto F, et al. Lévy trajectories of Peruvian purse-seiners as an indicator of the spatial distribution of anchovy ( Engraulis ringens )[J]. Ices Journal of Marine Science, 2005, 62(3):477-482. [↑](#endnote-ref-22)
22. [] Lee J, South A B, Jennings S. Developing reliable, repeatable, and accessible methods to provide high-resolution estimates of fishing-effort distributions from vessel monitoring system (VMS) data.[J]. Ices Journal of Marine Science, 2010, 67(6):1260-1271. [↑](#endnote-ref-23)
23. [] Joo R, Bertrand S, Chaigneau A, et al. Optimization of an artificial neural network for identifying fishing set positions from VMS data: An example from the Peruvian anchovy purse seine fishery[J]. Ecological Modelling, 2011, 222(4):1048-1059. [↑](#endnote-ref-24)
24. [] 张胜茂, 杨胜龙, 戴阳,等. 北斗船位数据提取拖网捕捞努力量算法研究[J]. 水产学报, 2014, 38(8):1190-1199. [↑](#endnote-ref-25)
25. [] Zheng Y. Trajectory Data Mining: An Overview[M]. ACM, 2015. [↑](#endnote-ref-26)
26. [] Douglas D H, Peucker T K. Algorithms for the Reduction of the Number of Points Required to Represent a Digitized Line or Its Caricature[J]. Cartographica the International Journal for Geographic Information & Geovisualization, 1973, 10(2):112-122. [↑](#endnote-ref-27)
27. [] Vitter J S. Random sampling with a reservoir[J]. Acm Transactions on Mathematical Software, 1985, 11(1):37-57. [↑](#endnote-ref-28)
28. [] Meratnia N, Rolf A. Spatiotemporal compression techniques for moving point objects[C]//International Conference on Extending Database Technology. Springer, Berlin, Heidelberg, 2004: 765-782. [↑](#endnote-ref-29)
29. [] Potamias M, Patroumpas K, Sellis T. Sampling trajectory streams with spatiotemporal criteria[C]//Scientific and Statistical Database Management, 2006. 18th International Conference on. IEEE, 2006: 275-284. [↑](#endnote-ref-30)
30. [] Fock H O. Fisheries in the context of marine spatial planning: Defining principal areas for fisheries in the German EEZ[J]. Marine Policy, 2008, 32(4):728-739. [↑](#endnote-ref-31)
31. [] Walker E, Bez N. A pioneer validation of a state-space model of vessel trajectories (VMS) with observers’ data[J]. Ecological Modelling, 2010, 221(17):2008-2017. [↑](#endnote-ref-32)
32. [] Joo R, Bertrand S, Chaigneau A, et al. Optimization of an artificial neural network for identifying fishing set positions from VMS data: an example from the Peruvian anchovy purse seine fishery[J]. Ecological Modelling, 2011, 222(4): 1048-1059. [↑](#endnote-ref-33)
33. [] Bueno-Pardo J, Ramalho S P, García-Alegre A, et al. Deep-sea crustacean trawling fisheries in Portugal: quantification of effort and assessment of landings per unit effort using a Vessel Monitoring System (VMS)[J]. Scientific Reports, 2017, 7:40795. [↑](#endnote-ref-34)
34. [] Joo R, Salcedo O, Gutierrez M, et al. Defining fishing spatial strategies from VMS data: Insights from the world's largest monospecific fishery[J]. Fisheries Research, 2015, 164:223-230. [↑](#endnote-ref-35)
35. [] Yuan J, Zheng Y, Zhang C, et al. T-drive: driving directions based on taxi trajectories[C]// Sigspatial International Conference on Advances in Geographic Information Systems. ACM, 2010:99-108. [↑](#endnote-ref-36)
36. [] Campbell M S, Stehfest K M, Votier S C, et al. Mapping fisheries for marine spatial planning: Gear-specific vessel monitoring system (VMS), marine conservation and offshore renewable energy[J]. Marine Policy, 2014, 45(1):293-300. [↑](#endnote-ref-37)
37. [] Russo T, Parisi A, Cataudella S. Spatial indicators of fishing pressure: Preliminary analyses and possible developments[J]. Ecological Indicators, 2013, 26(1):141-153.

    致谢

    由衷感谢我的导师，洪锋教授。在研究中给予我大量的指导和帮助，对出现的问题及时指正，一起分析课题的问题。他严谨的工作态度和丰富的理论知识令我感动由衷的钦佩。

    感谢宗源师姐和黄海光师兄，在本文的研究中向我提出了大量的意见，并帮助我寻找问题，指出不足。

    感谢实验室的朋友们，在交流讨论中给予我灵感，在我灰心丧气时鼓励我前行。

    个人简历

    1992年9月30日出生于山东省青岛市。

    2012年9月考入浙江理工大学启新学院计算机科学与技术专业，2015年6月本科毕业并获得工学学士学位。

    2015年9月考入中国海洋大学信息科学与工程学院计算机技术专业攻读硕士学位至今。

    发表的学术论文

    1. Hong F, Zhen Y, Zong Y, et al. OceanRoads: Fishing area separation and relation discovery through VMS trace analysis[C]// Oceans. IEEE, 2017:1-6.
    2. Zong Y, Huang H, Hong F, et al. Recognizing fishing activities via VMS trace analysis based on Mathematical Morphology[C]// Techno-Ocean. IEEE, 2017.

    [↑](#endnote-ref-38)