**谨以此论文献给所有关心支持我的人**

**------------------甄勇**

基于渔船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学校有关保留、使用学位论文的规定，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.学校可以将学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。同时授权清华大学“中国学术期刊(光盘版)电子杂志社”用于出版和编入CNKI《中国知识资源总库》，授权中国科学技术信息研究所将本学位论文收录到《中国学位论文全文数据库》。（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

基于渔船轨迹数据的海洋捕捞规律挖掘

研究与实现

摘 要

如何合理利用海洋渔业资源一直是困扰各国的难题。近年来，渔业政策制定不合理，渔船过度捕捞等因素，导致了渔业资源日益枯竭。因此，针对渔业资源以及渔船行为的分析在近些年受到越来越多海洋研究工作者的重视。在这个过程中，利用渔业船舶监控系统（VMS）采集的渔船轨迹数据进行的研究分析工作得到了普遍认同。尤其是当前该系统在全球范围部署日趋完善，它能够提供的服务也在不断地丰富，不仅包括基本的渔船定位、碰撞检测，还涉及到了渔业经济与渔业资源分布的量化分析预测，对渔业生产的可持续发展意义重大。

本文的主要研究工作如下：

针对国内缺少航海日志数据管理的情况，本文首先根据单拖渔船的作业特点设计了渔船捕捞行为识别算法，完成了港口定位与航次划分，并利用形态学方法找到捕捞区，进而完成对捕捞行为的识别。

在此基础上，利用渔船在捕捞区之间的航行轨迹信息，设计了针对捕捞区域的投票分类算法，完成了对捕捞区域统计热度图的划分。根据分类结果将各个捕捞区域的拓扑关系进行抽象，我们得到了捕捞区域间的路网信息，并以此量化分析了我国东海渔民的捕捞作业区域正逐步向东北转移，预示着东海近海渔业资源逐渐匮乏，同时渔民的捕捞成本在不断加大。

最后，将渔船行为识别系统整合为一体，设计并完成了船舶轨迹数据分析系统。通过数据管理层、业务层和展示层的设计，同时，出于对时间开销的考虑，对原算法进行了优化，方便对轨迹、航次、捕捞区进行可视化分析。

关键字:VMS；分类；渔业资源；行为分析

Mining the regular pattern of marine fishing based on the data of fishing vessel trajectory

Abstract（待修改）

As the worldwide deployment of the Vessel Monitoring System (VMS) becoming more and more perfect, the services it can provide are constantly enriched, including not only the basic functions such as fishing location and collision detection, also relates to the quantitative analysis and prediction of the fishery resources distribution and the fishery production, which is significant to sustainable development of fishery economy.

The previous research focused on identifying and analyzing the fishing area, aiming to reveal the spatial-temporal pattern of the fishing area. Based on this, we use the trajectory data between fishing areas to design a voting classification algorithm for fishing area. It achieves a reasonable division of fishing area. On the one hand, it can make more detailed quantitative analysis of the temporal-spatial pattern of fishery resources. On the other hand, the trajectory information can be used to analyze the fishing behaviors, which can be used as a reference for fishery management.

According to the operation characteristics of single trawler, the following works have been accomplished. First of all, by using the Fishing Activity Recognition System(FAR), achieving the port identification, track segmentation and fishing areas identification. Secondly, a voting classification algorithm is proposed exploiting the trajectory between fishing areas, to divide the fishing areas into fine-grained degree. And then, the fishery resources distribution and the fishing behaviors are analyzed after the algorithm. Finally, an application system is achieved to integrate the above algorithms and result. There is also an evaluation about it.

Keywords: VMS; Classification; Fishery Resources; Behavior Analysis

目 录

1. [引言 9](#_Toc508180145)
   1. [研究背景与意义 9](#_Toc508180146)
   2. [研究现状 10](#_Toc508180147)
   3. [研究内容和贡献 11](#_Toc508180148)
   4. [论文组织结构 12](#_Toc508180149)
2. [渔船捕捞行为识别算法 13](#_Toc508180150)
   1. [算法结构 14](#_Toc508180151)
   2. [VMS数据处理 14](#_Toc508180152)
   3. [港口定位 15](#_Toc508180153)
      1. [低速法 17](#_Toc508180154)
      2. [坐标驻留法 17](#_Toc508180155)
   4. [航次划分 18](#_Toc508180156)
   5. [捕捞行为识别 20](#_Toc508180157)
   6. [结果验证 24](#_Toc508180158)
   7. [本章小结 26](#_Toc508180159)
3. [渔船航道识别算法 27](#_Toc508180160)
   1. [研究的起点 27](#_Toc508180161)
   2. [迭代计算的输入 28](#_Toc508180162)
   3. [投票方案的设计 29](#_Toc508180163)
   4. [一次完整的迭代流程 31](#_Toc508180164)
   5. [渔船航道 33](#_Toc508180165)
   6. [本章小结 35](#_Toc508180166)
4. [软件实现 36](#_Toc508180167)
   1. [系统架构 36](#_Toc508180168)
   2. [数据管理层 37](#_Toc508180169)
      1. [文件系统结构 37](#_Toc508180170)
      2. [数据处理 38](#_Toc508180171)
   3. [业务层 38](#_Toc508180172)
      1. [港口定位模块 38](#_Toc508180173)
      2. [航次划分模块 38](#_Toc508180174)
      3. [捕捞区识别模块 39](#_Toc508180175)
   4. [展示层 39](#_Toc508180176)
      1. [导入数据 41](#_Toc508180177)
      2. [保存 42](#_Toc508180178)
      3. [退出 43](#_Toc508180179)
      4. [时间选择 43](#_Toc508180180)
      5. [渔船选择 44](#_Toc508180181)
      6. [航次编号 45](#_Toc508180182)
      7. [捕捞区统计热度图 46](#_Toc508180183)
      8. [港口分布图 47](#_Toc508180184)
      9. [示意图操作 47](#_Toc508180185)
   5. [本章小结 48](#_Toc508180186)
5. [总结与展望 49](#_Toc508180187)
   1. [总结 49](#_Toc508180188)
   2. [展望 50](#_Toc508180189)

# 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。但是近年来，我国近海渔业资源日趋枯竭，并且渔民收入下降难以维持收支平衡，这对矛盾越来越突出。主要原因是：一方面，在利益驱使下“竭泽而渔”，严重破坏了海洋生态平衡；另一方面，渔民的捕捞行为往往依靠口口相传的经验，容易造成局部区域的过度捕捞，缺乏系统科学的指导。为了规范渔业捕捞，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[[1]](#endnote-1)。同时，2017年初规定，在“十三五”期间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，并为每个省份制订了指标和进度计划[[2]](#endnote-2)。

在这样的形势下，船舶监控系统(Vessel Monitoring System，VMS)越来越受到重视。它起源于葡萄牙，最初是为了保障船舶航行安全[[3]](#endnote-3)。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[[4]](#endnote-4)。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，如警示与其他船只存在碰撞危险等。

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。相比于国外VMS数据采集间隔长达两小时[[5]](#endnote-5)，我国基于北斗卫星导航的VMS能够达到平均三分钟一条记录的采样频率，且传输可靠性高。这为科学研究提供了便利，不仅能够统计不同海域的捕捞强度，更能够揭示渔业资源的时空变化，分析渔民的捕捞策略。

通过定位系统对人或车辆进行跟踪，来分析人的轨迹和行为，进一步对地理区块之间的关系进行分析，一直是受到研究者们关注的热点问题。它可以揭示个体行为的潜在规律，以及地理区块关系的时空变化。但是，海洋上没有统一的道路约束，船舶航行相对自由，如何分析区域之间的时空变化有待进一步研究。

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。另一方面，基于渔业轨迹数据的分析研究已经具备一定基础，且可以借鉴人和车辆的轨迹研究成果，进行知识迁移。所以，渔业船舶轨迹数据具备进一步挖掘的潜力，本文将以此为基础，对海洋捕捞规律进行研究。

## 研究现状

VMS轨迹数据研究涉及多个领域。比如Coro等人按月计算了加拿大近海捕捞努力量[[6]](#endnote-6)；Fonseca等人通过比较拖网渔船的VMS数据和渔获量分析了渔场变迁[[7]](#endnote-7)；Watson和Gerritsen等人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[[8]](#endnote-8)，[[9]](#endnote-9)。这些研究的共同点是都涉及到了捕捞行为的识别。

渔船行为判断主要是利用VMS数据的不同字段作为特征进行分析。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，确定固定的速度区间作为渔船行为的判断标准3，[[10]](#endnote-10)，[[11]](#endnote-11)，[[12]](#endnote-12)，[[13]](#endnote-13)，[[14]](#endnote-14)。这种方法处理速度较快，但是Bertrand指出，这样处理容易使渔船捕捞行为的数据点数量被高估[[15]](#endnote-15)。Lee通过统计不同研究中的速度阈值，发现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，因此该方法的实用性较差[[16]](#endnote-16)。

另一种区分渔船行为的方法是对VMS轨迹数据进行模型训练与分类。Joo等人在研究秘鲁鳀鱼时，发现利用速度判断捕捞行为的方法高估了捕捞数据，通过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[[17]](#endnote-17)。对于高时空分辨率的北斗卫星数据，Zhang等人提出利用统计学模型分析航速和船艏向，进而判断捕捞状态的方法[[18]](#endnote-18)。但是，这种方法需要大量标记行为的数据进行模型训练，而训练过程一方面耗时较大。

在轨迹数据挖掘方面，根据Zheng等人总结的人体、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[[19]](#endnote-19)。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等人基于垂直欧氏距离（perpendicular Euclidean distance）提出了Douglas-Peucker算法对离线轨迹数据进行压缩[[20]](#endnote-20)。为了解决在线轨迹的压缩问题，Vitter等人提出了reservoir sampling算法[[21]](#endnote-21)。之后，Maratnia等人提出了改进的sliding window和Before Open Window算法[[22]](#endnote-22)。Potamias等人提出了基于物体运动速度和方向的在线轨迹压缩方法。[[23]](#endnote-23)

## 研究内容和贡献

为了能够在缺少标定数据的前提下识别捕捞行为，本文提出了渔船捕捞行为识别算法。在此基础上设计了渔船航道识别算法，分析渔船航道和捕捞资源的变化规律。最后，为了方便渔业从业人员使用，用软件实现了渔船捕捞行为识别算法。

渔船捕捞行为识别算法对VMS数据进行了分类，识别出捕捞行为对应的数据。在这个过程中完成了港口定位、航次划分，并利用数学形态学识别了捕捞轨迹。最后，我们计算了捕捞时间和捕捞距离与渔业产量进行比较，验证了算法的准确率。并对捕捞区域热度图进行统计，分析了渔业资源时空变化。

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区域热度图，通过迭代和一种投票方案对捕捞区进行分类，得到了渔船的航道信息。通过比较相同季度的航道变化，分析了渔业资源的时空变化，并对渔民的捕捞行为进行分析。

船舶轨迹数据分析系统是渔船捕捞行为识别算法的软件实现。将VMS数据导入软件，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）在缺少标定数据的条件下从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

由于VMS数据本身不包含区分船舶行为的信息，所以对于它的深入研究往往要先解决对不同渔船行为的分类问题。虽然航海日志记录了船舶航行和停泊过程中的主要情况，可以VMS数据进行粗粒度的分类，但我国渔业部门缺乏对它的数字化管理，无法作为参考。为了在不借助标定数据的条件下完成捕捞行为的识别，本章提出了渔船捕捞行为识别算法。算法的输入是两年的拖网渔船VMS数据。输出是一连串渔船捕捞行为的起点和终点。

对于从VMS数据中按渔船行为进行分类的问题，研究者们已经进行了大量的工作，可以把这些工作分为两类：基于经验的方法和基于分类模型的方法。前者是利用渔船实际航行过程中总结的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等人对英国附近海域约400艘不同类型的渔船进行了研究，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，将速度在[3,10]km/h区间内数据识别为捕捞行为3。Fork等人在研究捕捞区域变化时，对不同渔船类型采用不同的速度阈值区分捕捞行为：拖网渔船< 8kn，刺网渔船< 5kn。这种方法处理较快，但分类准确率不高，且不具备通用性。后者是利用模型对轨迹进行训练和分类。比如Walker等人在研究金枪鱼围网渔船的轨迹数据时使用状态空间模型（state-space model）将渔船行为分为寻找、捕捞、停泊、航行四种，并通过船艏向和航速数据训练了隐马尔可夫模型（Hidden Markov Model, HMM），完成了对不同行为的分类[[24]](#endnote-24)。这种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，且时间开销较大。区别于前人的工作，本文根据拖网渔船拖曳渔网折返运动的特点提出了一种利用图像处理方法识别捕捞行为的算法。

航海日志记录了包括船舶航行行为在内的多种信息，虽然粒度较粗，但仍然是研究VMS数据的重要参考。然而，渔业部门难以对航海日志进行收集整理。为了能够在没有标定数据的条件下实现渔船捕捞行为的识别，渔船捕捞行为识别算法首先对VMS数据进行分段，简化处理对象。然后，利用数学形态学识别捕捞轨迹区域，从而完成捕捞行为的识别。

## 算法结构

本文把研究问题拆分为四个子问题，如图2-1所示，根据这些问题，算法中依次设计对应的模块进行解决：

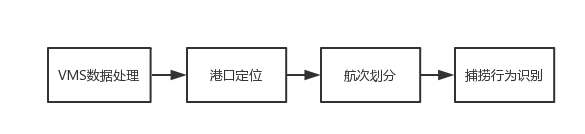


图2-1 渔船捕捞行为识别算法结构

（1）VMS数据处理模块：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——通过得到的港口位置信息，实现对VMS数据进行分段。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还经常停靠在一些小型码头或锚泊地，这些地点难以统计。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，刻画了渔船离开港口进行捕捞到返回港口这样一个完整的出海捕捞过程。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：利用拖网渔船在进行捕捞活动时折返拖曳渔网的特点，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，进出该区域的位置就是捕捞行为的起点和终点。

## VMS数据处理

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。该船型在捕捞作业时，渔船拖曳渔网在局部海域折返运动（图2-2），导致该区域作业渔船较少，避免了其他渔船的干扰，VMS数据相对独立；其作业方式灵活，适用于多种环境及各个深度的水层，因此VMS数据量大，便于研究。因此，本文对该船型的轨迹数据进行分析，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。

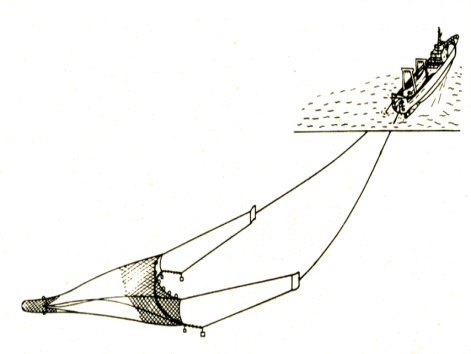


图2-2 单船拖网作业

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，经纬度信息按时间排序可以得到渔船轨迹。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[[25]](#endnote-25)，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2-3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。

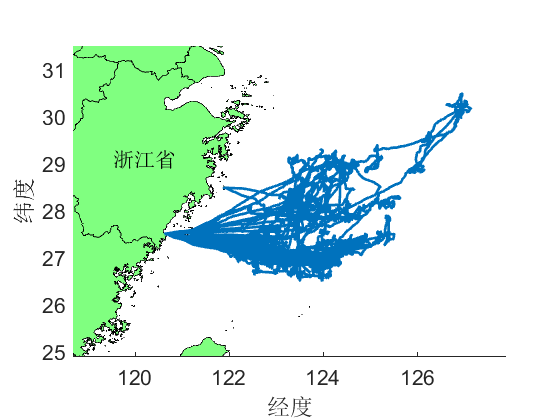


图2-3 渔船轨迹（终端ID：255368）

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。

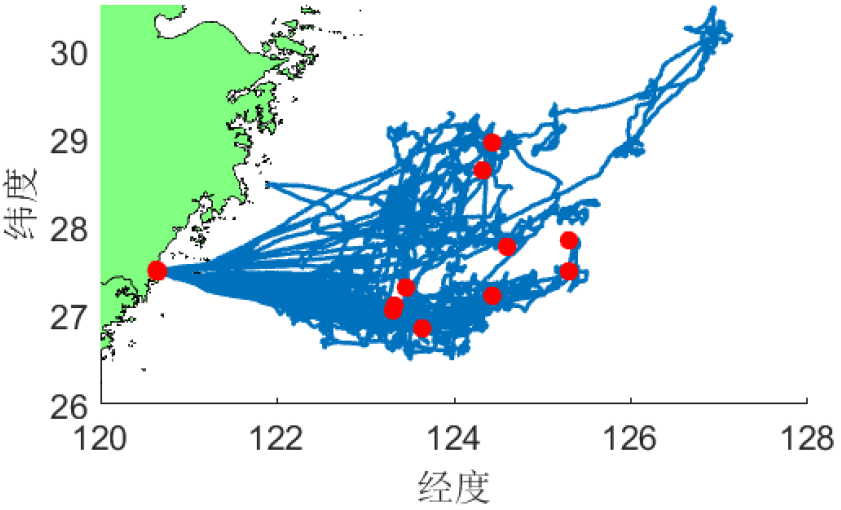


图2-4 “坐标驻留”现象分布（终端ID：255368）

图2-4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。

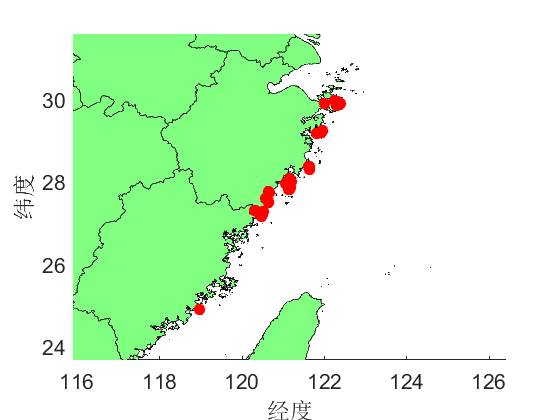


图2-5 港口识别结果

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2-5），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。

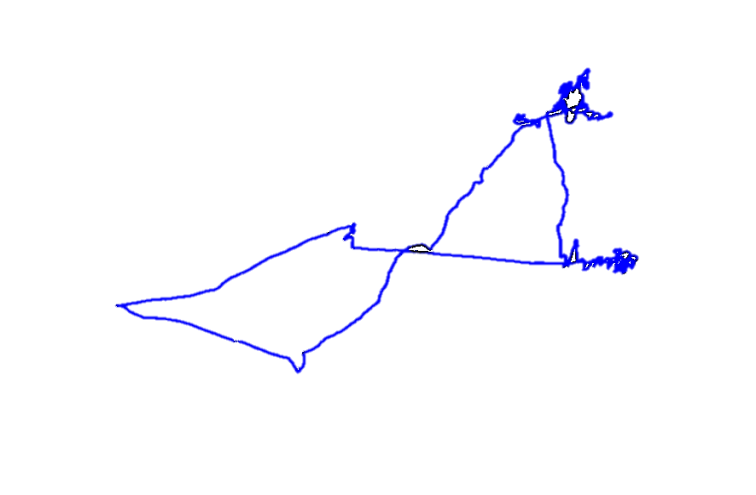
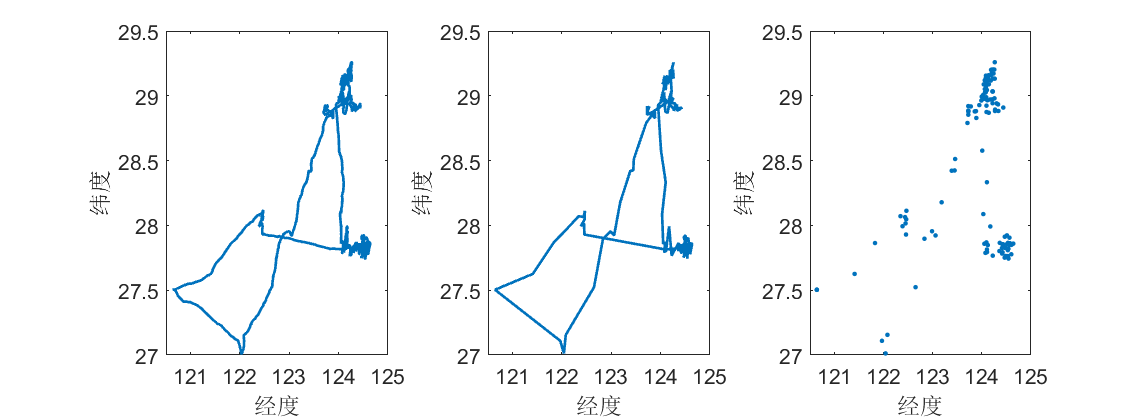


图2-6 根据轨迹密集区域判断捕捞轨迹区域

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2-7是对航次数据压缩前后的效果。图2-7（1）是压缩前的折线图，共7,183条记录；图2-7（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2-7 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2-7（3）的点图与图2-7（2）折线图的数据相同，可以看出，图2-6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。

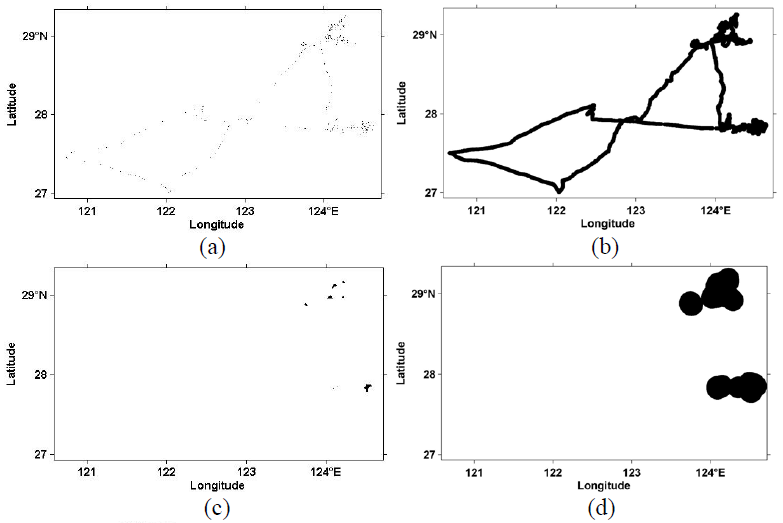


图2-8 利用形态学方法识别捕捞区的流程

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2-8（a））；通过腐蚀操作，将密集轨迹点连接成团（图2-8(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2-8(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2-8(d)）。

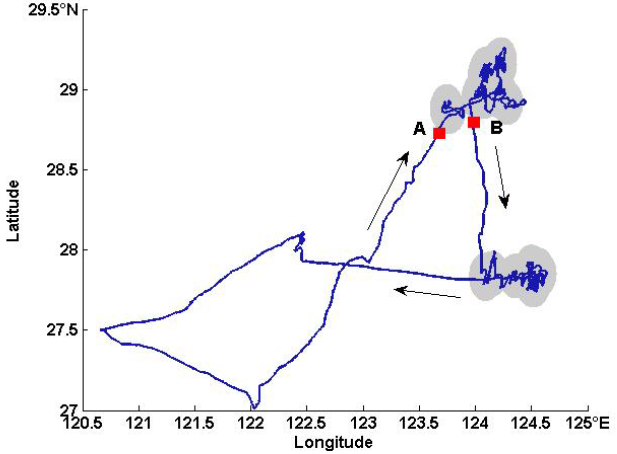


图2-9 完成捕捞区域识别的航次轨迹

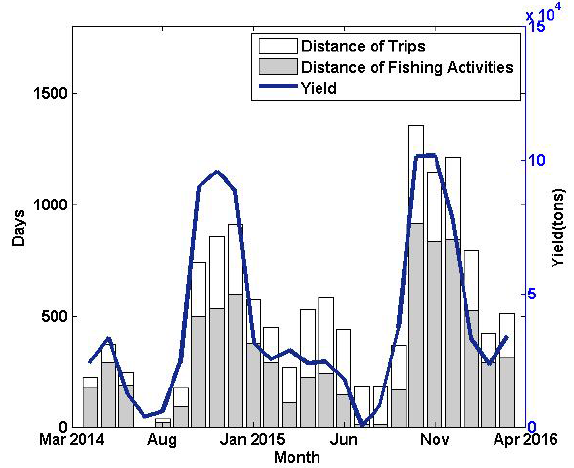
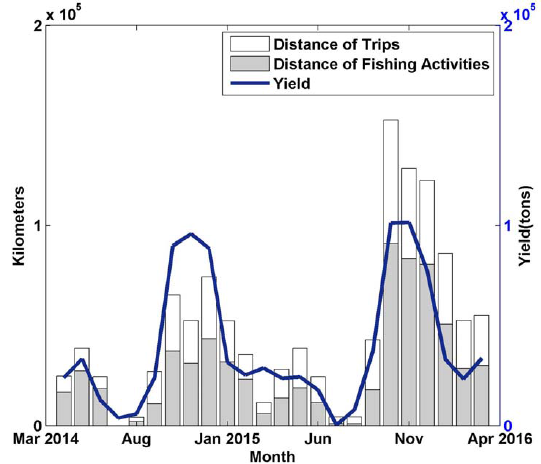
在成功识别捕捞轨迹区域后，就可以完成捕捞行为的识别：对于一个航次而言，我们把在捕捞轨迹区域内的轨迹视为渔船的捕捞行为，其余轨迹视为航行行为。图2-9中灰色区域是该航次通过数学形态学识别出的捕捞轨迹区域，AB两点是捕捞行为的起点和终点。最终将形如AB点对的序列作为算法的输出结果。

## 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进行分析。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2-10 捕捞时间/捕捞距离与渔业产量数据对比

图2-10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2-10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2-11 2014年第二季度到2016年第二季度的捕捞强度热度图

图2-11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2-11（a）（e）（i））可以明显看出这一趋势。与图2-10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

在上一章，通过从VMS数据中识别渔船的捕捞行为，发现了捕捞作业的东移趋势。但是，这种趋势如何变化？区域之间存在怎样的联系？这些问题难以回答。原因在于，刻画这种变化需要从更细粒度的角度分析。为了对捕捞热度图进行合理的分类，同时建立各分类区域之间的联系，本章提出了渔船航道识别算法。

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等人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，并对它们的范围变化进行了比较[[26]](#endnote-26)。Russo等人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[[27]](#endnote-27)。然而，这些研究并没有对区域进行分类，并且往往只关注捕捞行为，忽视了渔船在捕捞轨迹区域之间的航行轨迹信息。而这些轨迹信息描述了区域之间的联系，刻画了渔船的航道。同时，借鉴于陆地道路能够把城市划分成区块，海上航道同样有助于捕捞区域的划分。因此，确定算法思路：利用渔船在捕捞区之间的航行轨迹信息作为捕捞热度图的分类依据，并根据最后的分类结果抽象出渔船航道。

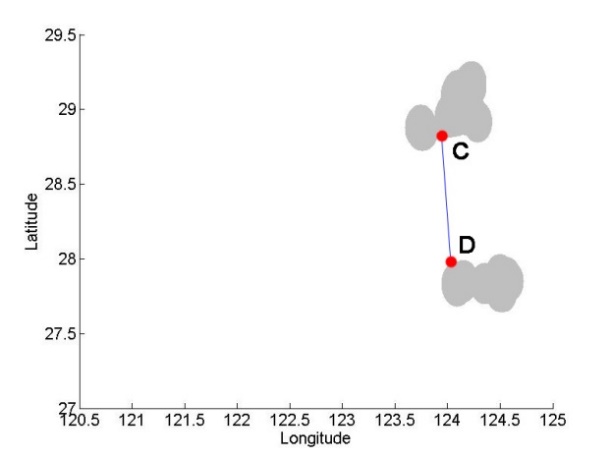
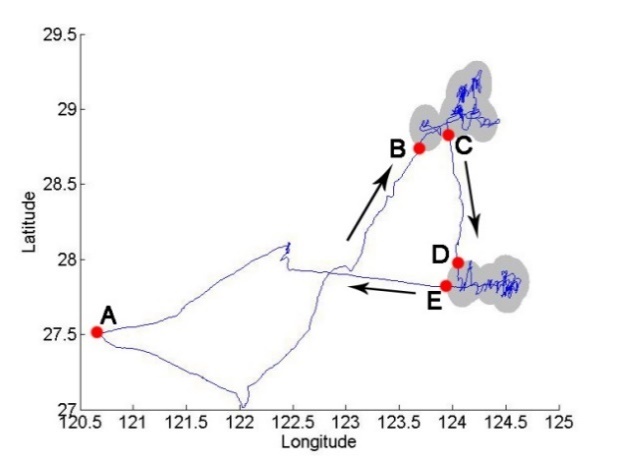
研究的难点在于，海上本身是没有航道和区块的。对于统计得到的捕捞热度图缺乏验证手段。而捕捞轨迹区域识别的准确性又会影响到捕捞区域之间的轨迹，进而影响分类结果。两者相互依存又相互矛盾。

于是，我们提出了渔船航道识别算法来同时完成捕捞热度图的分类和渔船航道的识别任务。这个算法包含两个部分：捕捞区粗粒度识别和迭代计算。前者利用上一章提到的渔船捕捞行为识别算法，得到捕捞轨迹区域以及捕捞行为的起点终点，这部分不再赘述。后者将捕捞热度图和渔船在捕捞轨迹区域之间的轨迹信息作为输入，利用一种投票方法迭代地尝试各种分类方案，最终确定捕捞热度图的分类结果，并抽象出渔船航道。下面介绍这种迭代方法的细节。

## 迭代计算的输入

上一节中，我们提到了迭代输入由两部分组成：捕捞热度图和渔船在捕捞区之间的轨迹信息。本节中分别对两者进行介绍。

捕捞区热度图产生于渔船捕捞行为识别算法。通过数学形态学方法，从每个航次中提取对应的捕捞区，并按照季度进行统计，得到如图3-10的捕捞区热度图。



1. （b）

图3-1 对捕捞区之间的航行轨迹进行抽象

渔船在捕捞区之间的轨迹信息并不是直接使用真实的轨迹数据，而是将其抽象为一条线段。如图3-1（a）展示了某艘船一个航次的轨迹，灰色区域为识别出的捕捞区域，各个节点的访问顺序为ABCDEA，可以看出，CD间的轨迹就是渔船在两个捕捞区之间的航行轨迹。但是这段数据的数据量和变化多样，不利于接下来的处理。所以将其抽象为图3-1（b）所示的“航行线段”——只保留了离开捕捞区的C点和进入下一个捕捞区的D点。这样处理虽然忽略了大量的信息，但是能够得到统一的航行信息：线段两端应该落在不同的捕捞区内部。

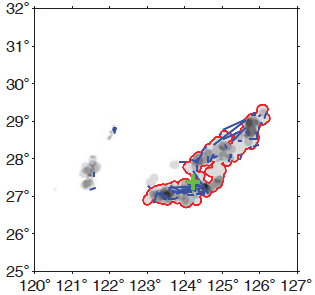
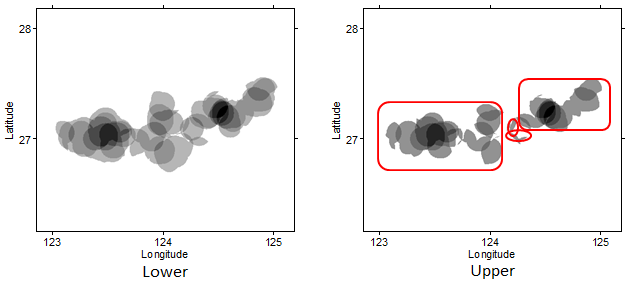


图3-2 2014年第二季度捕捞区统计热度图与航行线段

将迭代计算的两个输入叠加在一起进行分析，如图3-2所示。灰色区域为统计2014年第二季度得到的捕捞区热度图，红色曲线包络的是一个相互连通的捕捞区域，蓝色线段是对应时间段内统计的所有航行线段。首先，根据连通性，我们把捕捞区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，还分布于它的内部。以红色曲线勾勒的区域为例，作为一个面积较大的捕捞区域，内部包含了大量的航行线段。回顾我们对航行线段的定义可以提出问题：是否应该把一个连通的捕捞区划分为多个子区域？我们给出的答案是应该进行分类。对于较短的航行线段也许可以认为是渔船捕捞行为识别的误差导致的，但是长线段不能忽视。接下来的投票方案将介绍如何实现分类方法。

## 投票方案的设计

在上一节的最后，我们确定了要对捕捞区进行分类。但是如何分类呢？首先排出了形如K-means的分类方式——在一个捕捞区域内通过边界线的形式将其分割为多个小区域。因为这样做缺乏实际的指导作用。对于渔民而言，如果两个捕捞区之间的区分只是一条线，那么分类前与分类后就没有区别。所以，分类的目标是要将捕捞区划分为多个互不连通的小区域。

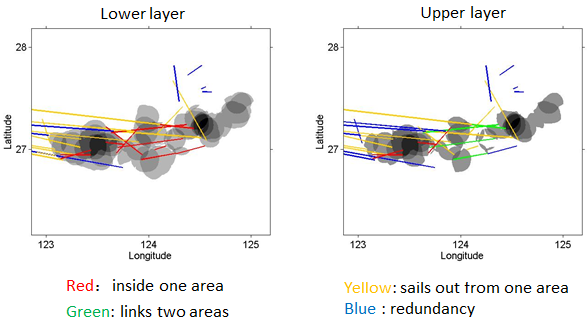


1. (b)

图3-3 相邻图层某个出现合并的区域

为了使分类后的捕捞区彼此分离，我们决定从捕捞区热度图的层次入手，保留热度数值大的区域，忽略热度值小的区域。设阈值θ=0，当热度值大于θ，可以保留图上全部的捕捞区信息；随着θ增大，保留的信息减少，且逐渐以高热度值为核心相互独立。通过分析对比，当θ取值为热度最大值的25%时，分类效果较好。如图3-3所示，是2014年第二季度的捕捞区的局部。将阈值θ=热度最大值\*25%作为热度图的切入点，观察到的相邻两个图层（a）下一层（b）上一层。在相邻图层间捕捞区存在三种变化方式：（Ⅰ）新增，上一层的空白位置对应到下一层是一个独立的捕捞区域；（Ⅱ）扩张，上一层中的某个捕捞区对应到下一层范围变大，但仍保持独立性；（Ⅲ）合并，上一层的两个或多个捕捞区域在下一层构成了一个连通的捕捞区。和前文一样，按照捕捞区的存在与否，把（Ⅰ）（Ⅱ）两种情况视为独立的捕捞区，对（Ⅲ）中捕捞区合并的情况单独分析。图3-3（a）中的捕捞区就符合方式（Ⅲ），是由图3-3（b）中的四个捕捞区合并组成的一个连通的整体。

利用这些“上一层”的小区域，我们来讨论如何进行分类。为此设计了一个投票方法：对于两种不同的分类方案，我们进行投票，得票数多的就是较优的方案。下面介绍具体的实现方法。



1. （b）

图3-4 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

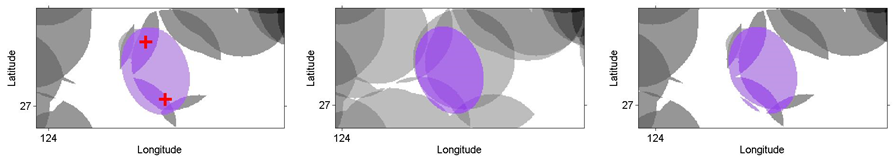
我们把航行线段作为“选票”。根据线段两端落点的不同，共有四种不同的选票，如图3-4所示：两端在同一个捕捞区内的选票用红色表示，称为反对票，得票数记作-2；两端在不同捕捞区内的用绿色表示，称为支持票，得票数记作+2；只有一端在捕捞区内的用黄色表示，是对其中一个捕捞区分类的支持，称为部分支持票，记作+1；两端均不在捕捞区内的用蓝色表示，称为弃权票，记作0。所以，对于一个分类方案进行投票，就是根据落点判断每一个选票的类型，将对应数字相加求和。图4-4是对图4-3的案例进行投票，图3-4（a）得票数为-11，图3-4（b）得票数为11。所以根据航行线段，图3-4（b）的分类方案要比图3-4（a）更佳。

## 一次完整的迭代流程

上一节介绍了如何根据投票结果获得更优的分类方案。在这一节中，将介绍如何利用迭代的方式筛选出最终的分类方案。

通过投票我们可以得出结论：相比于图3-4（a）中组合成一个整体，图3-4（b）中各个小区域保持独立会更合理一些。那么会不会存在更优的分类方案呢？我们尝试对多种组合进行投票。

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。实际上当三个区域排成一条直线时，基本上可以排除两端合并为一个整体而中间区域保持独立这种情况。根据这个剪枝条件，我们设计了一个由近及远的、不可逆的组合方案。具体步骤：取各个捕捞区的几何中心点两两连线，按线段长度排序；根据排序顺次进行合并、投票；若相比于合并前，合并后得票数较高，则将两个区域视为一个整体，不再分开讨论其他分组情况；当尝试合并的两个几何中心点对应的捕捞区归属于同一个整体，则不再进行尝试。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。我们倾向于将捕捞区尽量合并为一个整体，于是，对于得票数相同的情况，判断为合并而非独立。



1. （2） （3）

图3-5 两捕捞区的合并策略

合并的具体方法如图3-5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3-5（1）。将椭圆与下一层捕捞区取交集如图3-5（2），计算结果即为两捕捞区域的合并方式如图3-5（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。

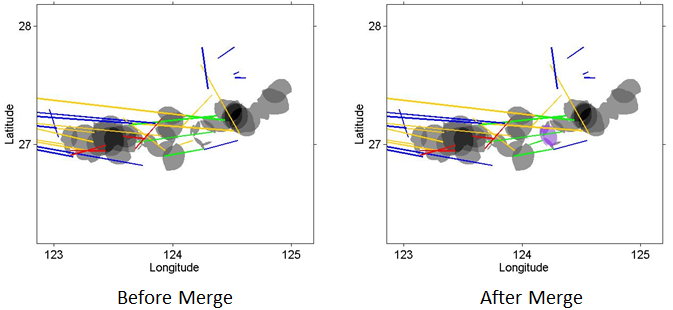


图3-6 比较合并前后的得票数

对于每次合并的尝试，比较合并前后的得票数大小。如图3-6所示，是对图3-5合并前后进行了投票，图3-6（1）的得票数为11，图3-6（2）的得票数为12，所以确定合并两个捕捞区域。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，实现了对捕捞区的细粒度划分。在此基础上，通过统计各个捕捞区的出入度和转移关系，我们得到了对渔船在捕捞区之间航行轨迹的抽象结果——渔船航道。

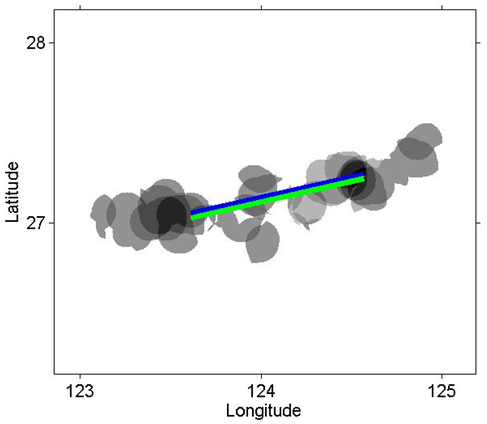
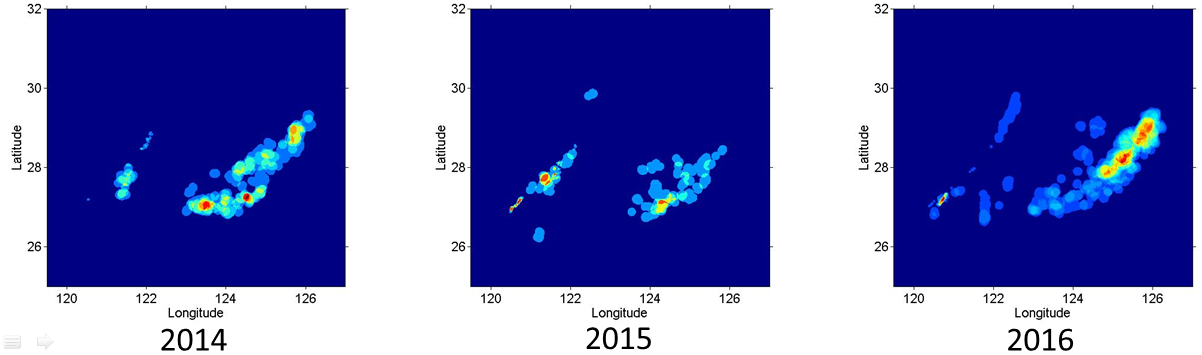


图3-7 渔船航道

如图3-7所示，灰色部分是图3-6中样例通过迭代，最终确定的分类方案，从最初的四个区域变为两个区域。统计渔船在先后在两个区域进行捕捞作业的轨迹，即选票中的“支持票”，作图上蓝绿色线段，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，线段粗细表示船次，也可以称为流量。图3-7上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。

我们称这样的连线为渔船航道。通过构建渔船航道，就可以构筑与陆地上的路网相似的道路网络，并以此为基础分析渔业资源的分布变化情况。



1. （2） （3）

图3-8 连续三年第二季度捕捞区热度图

比如图3-8是2014年至2016年第二季度的捕捞热度图，通过观察可以发现捕捞区有向东北移动的趋势，还可以通过捕捞区的几何中心位置、轮廓等特征描述这种变化。但是这些都是笼统的分析，无法进行局部捕捞区域的刻画。

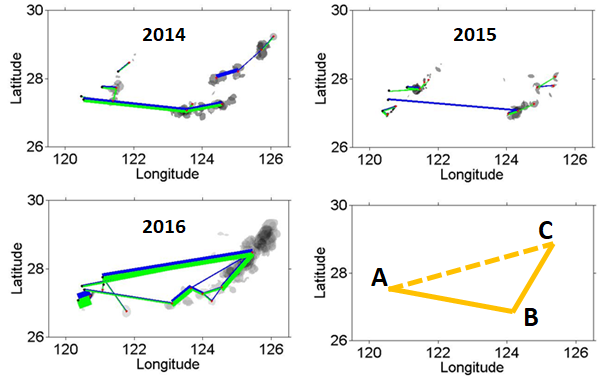


图3-9 连续三年海上路径变化

通过分类算法对捕捞区进行划分并构筑路网，有助于对渔业资源及渔民行为进行分析。如图3-9所示，对于连续三年第二季度的海上路径，我们把它抽象为一个三角形进行分析：A点表示港口，BC两点表示两个捕捞区域。在2014年第二季度，大部分的渔船出海沿着AB边航行，2015年对应季度船次整体明显减少，而到了2016年，只有少量船只沿AB航线捕鱼，大部分的船次都选择直接沿AC边捕捞。相比于B点，选择更远的C点进行捕捞，这充分说明B点渔业资源减少。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

为了将捕捞区热度图进行细粒度划分，同时找出渔船航道，本章提出了渔船航道识别算法。它包含两部分，一个是第二章介绍的渔船捕捞行为识别算法，得到两组参数：捕捞区之间的航行轨迹和捕捞区热度图。另一个是通过迭代的方式对两两分类方案进行投票，保留得票数高的方案，得到最终的分类结果。并根据出入度统计出渔船航道。根据渔船航道的流量变化，我们发现了渔民捕捞策略的变化。

# 软件实现

前面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和捕捞区，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。

为了能够实际应用，我们设计并实现了船舶轨迹数据分析系统，将渔船捕捞行为识别算法封装为一个可视化系统，方便观察和使用。本章将从系统架构入手，介绍系统的设计流程及功能实现。

## 系统架构

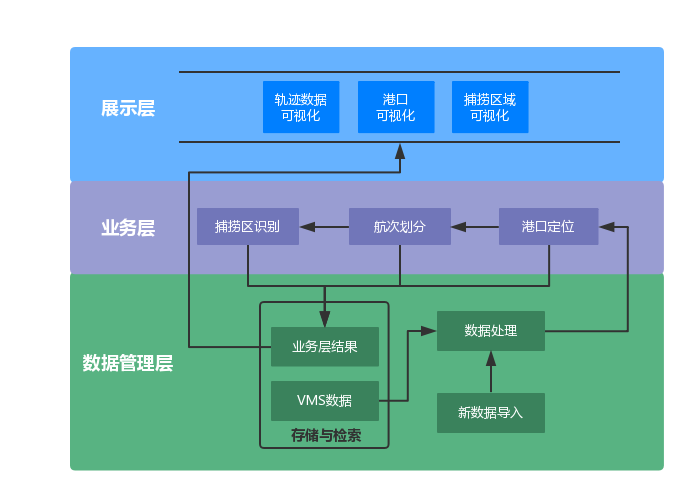


图4-1 船舶轨迹数据分析系统的架构设计

如图4-1所示，我们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，具体完成了两部分内容：定义了文件系统结构，完成了对VMS数据的基本处理。

### 文件系统结构

文件结构如图4-2所示。系统根目录“\”下存放有三个项目：

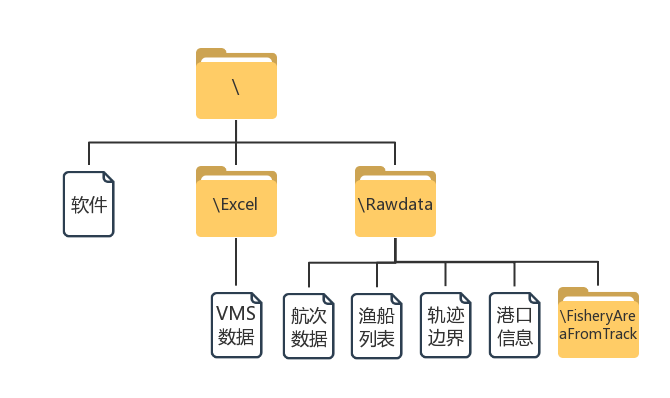


图4-2 文件系统结构示意图

软件：一个可执行文件，双击启动船舶轨迹数据分析系统。

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

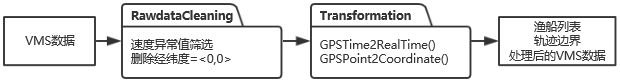


图4-3 数据处理流程

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()对VMS数据中的速度异常值和经纬度异常值进行删除；调用Transformation()，利用GPSTime2RealTime()和GPSPoint2Coordinate()将转换字段格式，方便后续处理。

## 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。

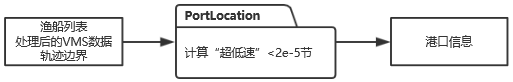


图4-4 港口定位流程

输入是渔船列表、处理后的VMS数据和轨迹边界，输出是港口信息。通过调用PortLocation()函数，利用“超低速”<2e-5节这一条件对VMS数据进行筛选，并利用轨迹边界划分网格去重，最终汇总得到港口信息表。

### 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、港口信息，输出包含航次信息的VMS数据。通过调用VoyageDivision()函数，利用港口信息对VMS数据进行分段，将每艘船的轨迹划分为多个航次。

### 捕捞区识别模块

对于一个航次，我们根据渔船的捕捞作业特点利用数学形态学识别捕捞区，如图4-6所示。

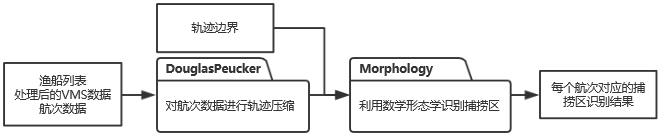


图4-6 捕捞区识别流程

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区识别结果。通过调用DouglasPeucker()函数，对航次轨迹进行压缩，并将压缩后的数据作为函数Morphology()的输入，利用数学形态学识别航次轨迹对应的捕捞区。

## 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：

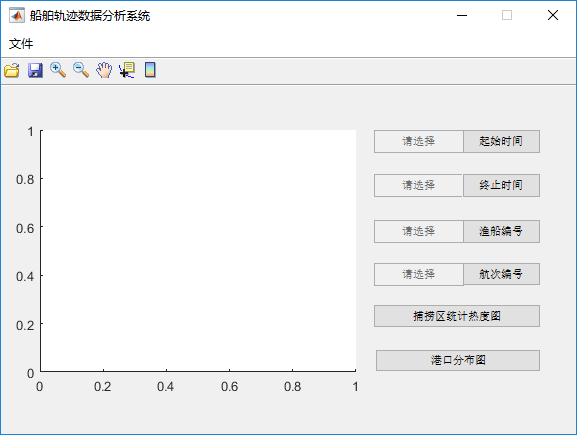


图4-7 船舶轨迹数据分析系统的主界面

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。

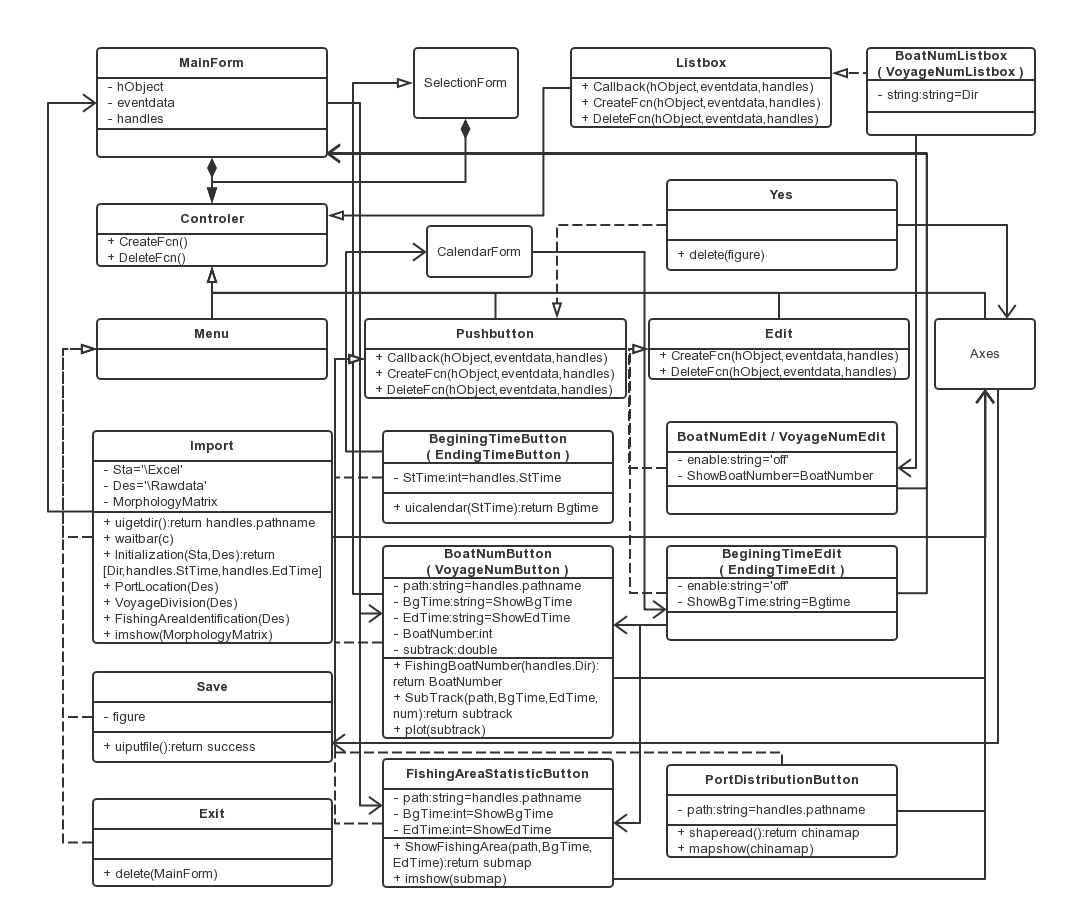
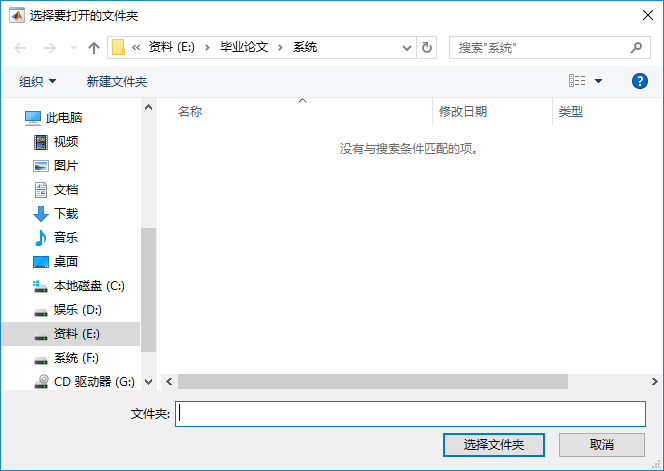
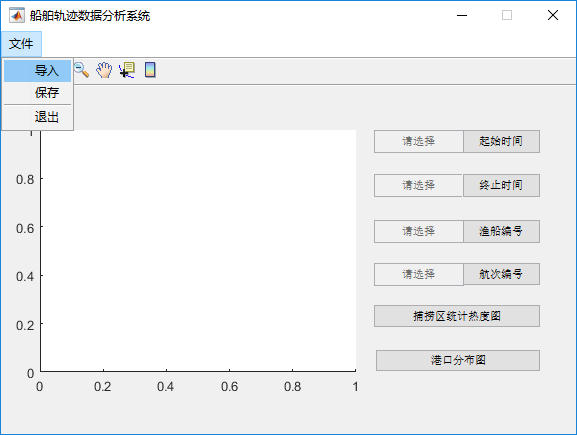


图4-8 展示层类图

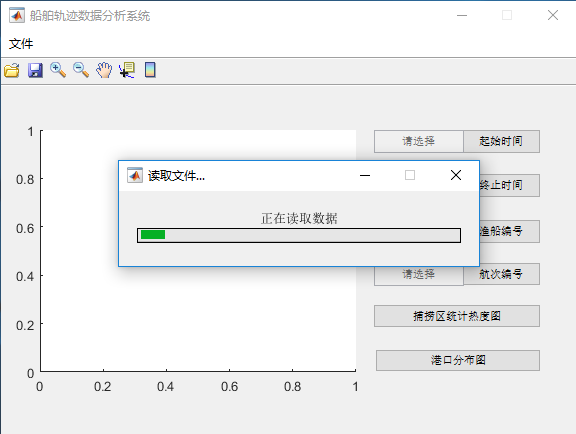
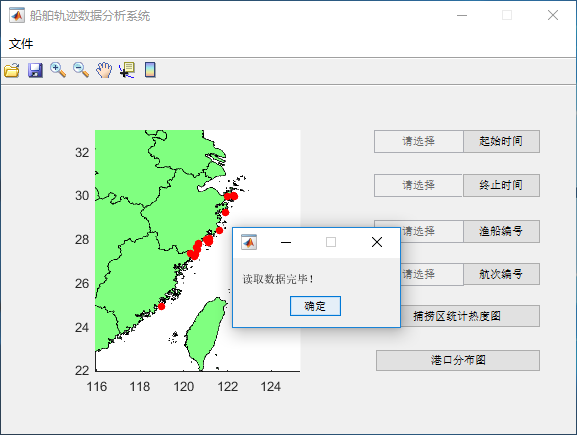
展示层的类图设计如图4-8所示。主界面（MainForm）包含了多个类型的控件（Controler）:菜单（Manu）、按钮（Pushbutton）、编辑框（Edit）、图片框（Axes）等。所有空间都包含有基本的触发事件：Callback(),CreateFcn()和DeleteFcn()。控件之间通过一个核心变量handles进行参数传递。业务层的主要功能通过导入菜单项（Import）执行。展示功能通过按钮（Pushbutton）的实例BeginingTimeButton/ EndingTimeButton、BoatNumButton/VoyageNumButton、Yes和FishingAreaStatisticButton完成。下面将对具体功能进行详细介绍。

### 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。

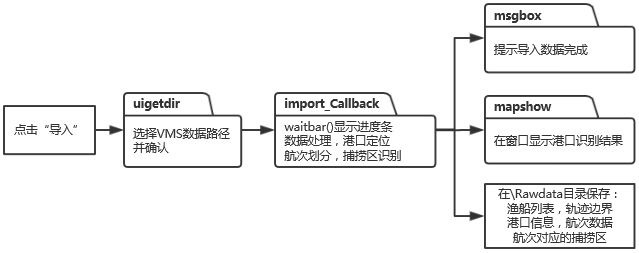


图4-10 导入数据的业务逻辑

业务逻辑如图4-10所示：

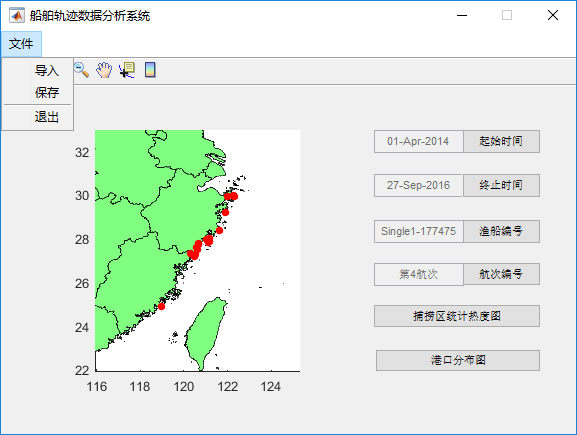
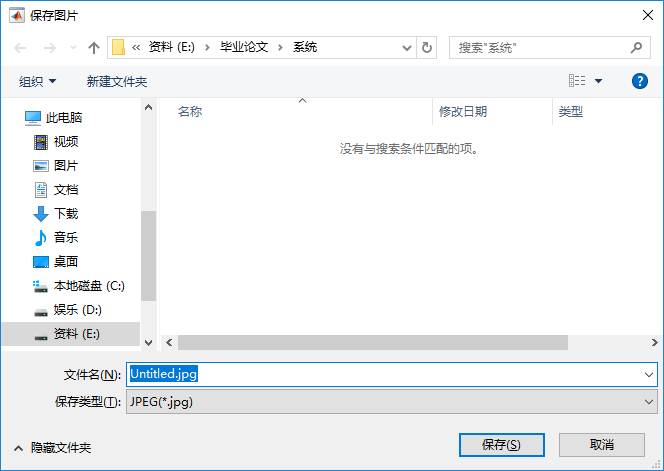
（1）点击“导入”按钮，调用uigetdir()函数，进行VMS数据的路径选择。

（2）确认后获取路径信息path，并调用import\_Callback()函数，依次处理完成数据处理、港口定位、航次划分和捕捞区识别操作。调用waitbar()函数显示进度信息，最终完成数据导入。中间产生的结果保存在路径path的\Rawdata目录下，包括渔船列表、轨迹边界、港口信息、航次数据和航次对应的捕捞区。

（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））。

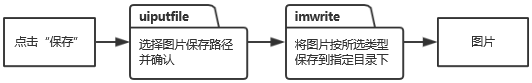


图4-12 保存的业务流程

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，对图片进行保存，确认后得到保存路径、文件名、文件类型等信息。

（2）调用imwrite()函数，按照保存路径、文件名、文件类型等信息将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。

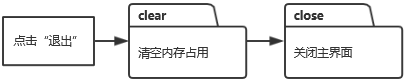


图4-13 关闭的业务流程

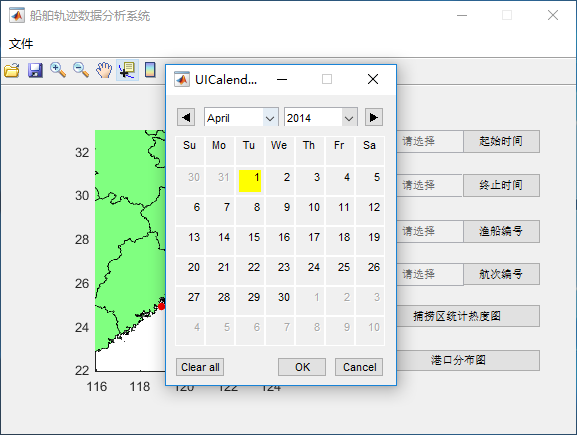
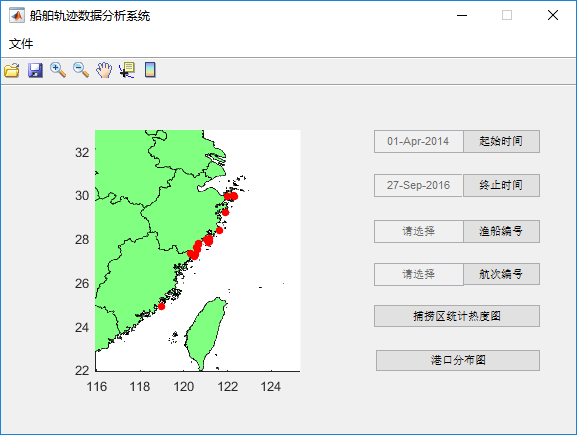
业务流程如图4-13所示：

（1）点击退出按钮，调用clear()函数，清除系统在运行过程中占用的内存资源。

（2）调用close()函数关闭主界面，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

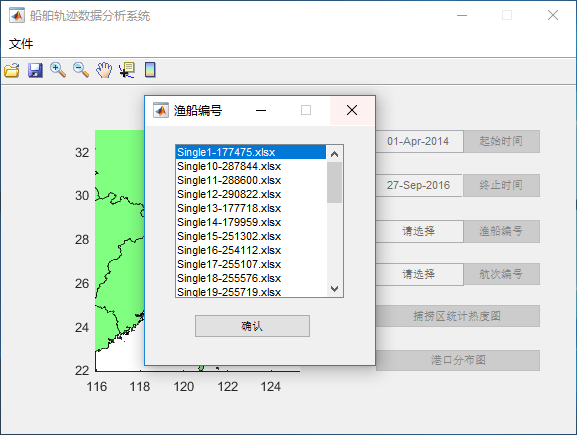
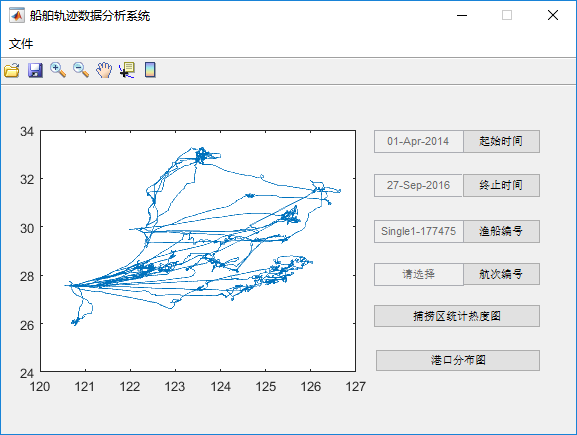
业务逻辑如图4-15所示：

（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，选择年月日并确认。若未选择时间，则返回为空值。

（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。

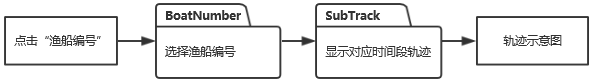


图4-17 渔船选择的业务逻辑

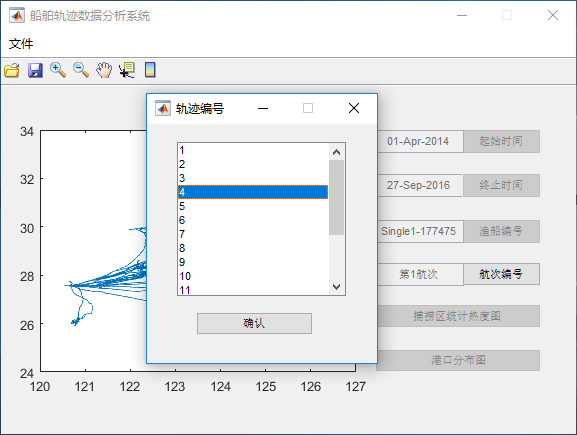
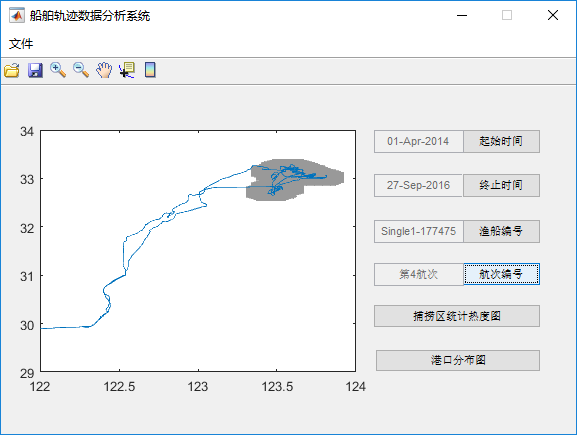
业务逻辑如图4-17所示：

（1）点击“渔船编号”按钮，调用BoatNumber()函数，弹出渔船编号窗口，选择一艘渔船点击确认。

（2）调用SubTrack()函数，按照起止时间选择对应时间段的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，我们可以观察该渔船对应航次的轨迹以及识别出的捕捞区，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））。

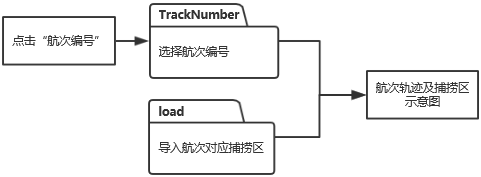


图4-19 航次选择的业务逻辑

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，选择一个航次点击确认。

（2）根据渔船编号及航次编号，调用load()函数导入对应的捕捞区识别结果，并与航次轨迹叠加，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：

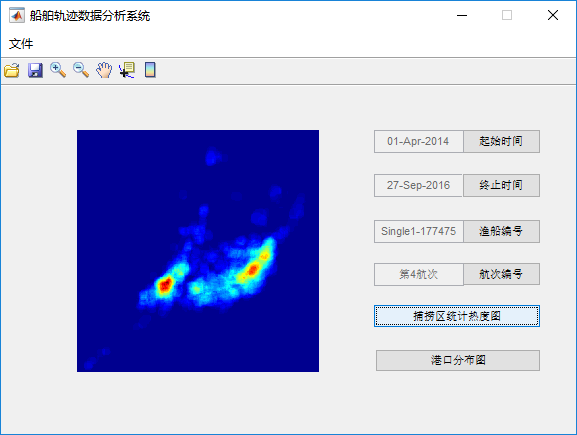
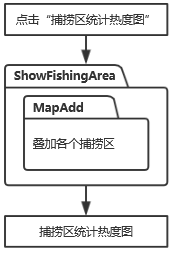
 

图4-20 捕捞区统计热度示意图 图4-21 捕捞区统计的业务逻辑

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）。

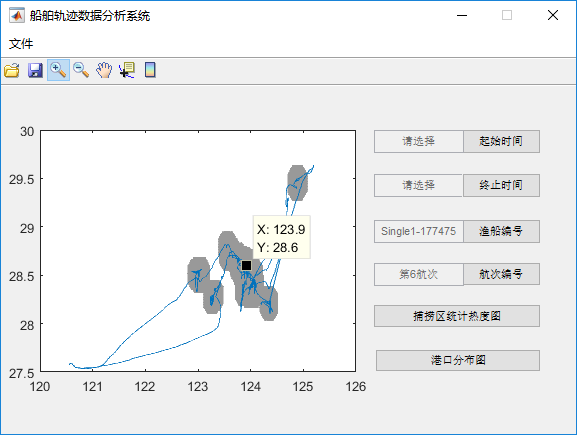
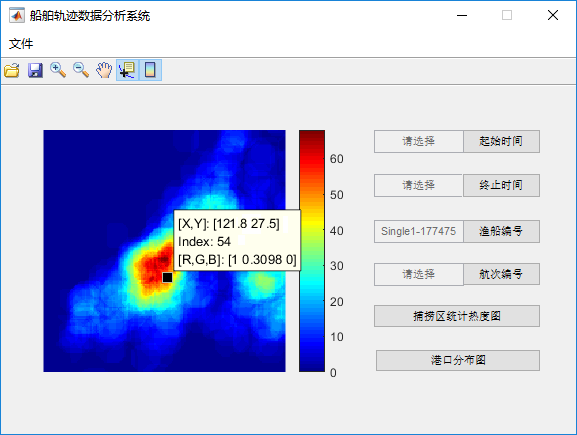
业务逻辑如图4-21所示：点击“捕捞区统计热度图”按钮，调用ShowFishingArea()函数，函数内通过递归调用MapAdd()函数将每艘渔船在起止时间段内的捕捞区叠加。显示在主界面窗口左侧的示意图上。若起止时间未选择，即起始时间和终止时间都显示“请选择”，则显示全部的捕捞区叠加结果。

### 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

本章通过设计船舶轨迹数据分析系统，将渔船捕捞行为识别算法可视化，方便对港口、轨迹、捕捞区进行可视化分析。系统共分为三个层次：数据管理层、业务层、展示层。其中，数据管理层主要负责文件系统的定义以及数据处理；业务层实现了算法；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、港口分布图等按钮，通过工具按钮观察港口、轨迹及捕捞区，同时可以对图像进行保存。

# 总结与展望

本章将对论文内容进行回顾，总结研究中遇到的问题，并对未来基于VMS数据的分析挖掘工作进行展望。

## 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，再加上北斗卫星导航系统提供了更高的采样频率，已经初步完成了数据积累工作。本文正是在此基础上开展VMS渔船轨迹数据分析与挖掘工作。

本文的主要研究工作如下：

（1）对当前基于VMS渔船轨迹数据的研究背景、发展现状进行综述，分析了当前研究的主要挑战与问题。同时，将行人、车辆轨迹分析相关研究作为参照对象，借鉴并融入到渔船轨迹分析上。

（2）提出了渔船行为识别系统（FAR）。首先，它针对国内渔民缺少记录航海日志习惯的情况，通过轨迹数据点的分布密度及<瞬时速度,船艏向>等于零等特征对港口进行定位，进而对VMS轨迹数据划分航次。然后，利用形态学方法对渔船单个航次轨迹的捕捞行为进行识别，并统计了捕捞区的范围，分析了捕捞区的变化。

（3）提出了投票分类算法（VCA）。它将渔船在捕捞区之间的航行轨迹作为分类判断的依据，对FAR中识别出的捕捞区进行进一步的划分，从而描述更细粒度的捕捞区变化情况。同时将捕捞区之间的航行轨迹抽象为两者间的通路，可以根据出度入度分析海上路径变化，从而定量描述渔业资源的分布变化情况。

（4）设计并实现了船舶轨迹数据分析系统。通过MATLAB的GUI编程，它将FAR整合为一个简洁的整体，便于观察渔船轨迹、航次及捕捞区识别结果、捕捞区统计热度图，对渔业数据进行分析。整个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进行预处理；业务层实现了FAR的算法，并根据时效性进行了优化；展示层通过界面窗口的设计，可以通过简单的按钮点选来控制时间段、渔船、航次，从而观察轨迹数据和捕捞区识别结果。

## 展望

渔业资源分析仍有较大研究空间。相较于陆上的轨迹分析研究，海上研究关注度低且手段较为单一，主要是通过VMS数据进行分析，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单一物种进行分析。

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。本文提出的投票分类算法就是一次尝试，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。另一方面，由于海上通信质量差，VMS数据存在明显的数据缺失情况，对于此类数据，随着渔船与捕捞区关系的揭示，可以尝试进行预测，而不是简单的差值操作。

1. 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-1)
2. 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-2)
3. 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-3)
4. 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-4)
5. Mills C M, Townsend S E, Jennings S, et al. Estimating high resolution trawl fishing effort from satellite-based vessel monitoring system data[J]. Ices Journal of Marine Science, 2007, 64(2):248-255. [↑](#endnote-ref-5)
6. Coro G, Fortunati L, Pagano P. Deriving fishing monthly effort and caught species from vessel trajectories[C]// Oceans. IEEE, 2013:1 - 5. [↑](#endnote-ref-6)
7. 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-7)
8. 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-8)
9. 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-9)
10. 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-10)
11. 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-11)
12. 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-12)
13. 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-13)
14. 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-14)
15. 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-15)
16. 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-16)
17. 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-17)
18. 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-18)
19. Zheng Y. Trajectory Data Mining: An Overview[M]. ACM, 2015. [↑](#endnote-ref-19)
20. 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-20)
21. Vitter J S. Random sampling with a reservoir[J]. Acm Transactions on Mathematical Software, 1985, 11(1):37-57. [↑](#endnote-ref-21)
22. 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-22)
23. 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-23)
24. 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-24)
25. 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-25)
26. 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-26)
27. Russo T, Parisi A, Cataudella S. Spatial indicators of fishing pressure: Preliminary analyses and possible developments[J]. Ecological Indicators, 2013, 26(1):141-153. [↑](#endnote-ref-27)