

基于轨迹图像特征匹配的渔船轨迹相似度计算和轨迹分类^①



徐文进¹, 解 钦¹, 黄海广²

¹(青岛科技大学 信息科学技术学院, 青岛 266061)

²(温州大学 计算机与人工智能学院, 温州 325035)

通讯作者: 解 钦, E-mail: xieqin1013@163.com

摘 要: AIS (Automatic Identification System) 是一种船舶的自动识别系统, 可以提供船舶的时间戳、经纬度、航向角度、速度等数据信息. 本文针对船舶航行轨迹多维度的特点以及对船舶轨迹预测的精确度和实时性的需求, 提出了一种基于图像检测和匹配的计算轨迹相似度的方法. 该方法首先将所有渔船轨迹数据进行可视化, 再通过 ORB (Oriented FAST and Rotated BRIEF) 算法和 BF (Brute-Force) 匹配来计算轨迹图片相似度用于划分渔船轨迹类型. 实验结果显示, 通过该计算相似度的方法具有精度高、易实现的特点, 与传统计算方法相比, 其在处理轨迹数据的效率和速度更具有优越性.

关键词: AIS 数据; ORB; BF; 轨迹图片相似度

引用格式: 徐文进, 解钦, 黄海广. 基于轨迹图像特征匹配的渔船轨迹相似度计算和轨迹分类. 计算机系统应用, 2021, 30(8): 232-236. <http://www.c-s-a.org.cn/1003-3254/8054.html>

Similarity Calculation and Trajectory Classification of Fishing Boats Based on Trajectory Image Feature Matching

XU Wen-Jin¹, XIE Qin¹, HUANG Hai-Guang²

¹(College of Information Science and Technology, Qingdao University of Science and Technology, Qingdao 266061, China)

²(College of Computer Science and Artificial Intelligence, Wenzhou University, Wenzhou 325035, China)

Abstract: The Automatic Identification System (AIS) can provide the time stamp, latitude and longitude, heading angle, speed, and other data information of ships. In light of multi-dimensional ship trajectories and the demand for accuracy and timeliness in trajectory prediction, this study proposes a calculation method for trajectory similarity based on image detection and matching. To be specific, the trajectory data of all fishing boats are first visualized, and then the Oriented FAST and Rotated BRIEF (ORB) algorithm and Brute-Force (BF) matching are used to calculate the similarity between trajectory pictures for classifying fishing boat trajectories. The experimental results show that the proposed method has high accuracy and can be easily implemented and is superior to the traditional methods in the processing efficiency and speed of trajectory data.

Key words: AIS data; ORB; BF; similarity between trajectory pictures

据农业部 2019 年统计, 我国现拥有渔船 73.12 万艘, 从事渔业的人口为 1828.02 万人. 全社会渔业经济

总产值 26406.50 亿元, 其中海洋捕捞产值为 2116.02 亿元^[1]. 在海洋捕捞给人们带来丰厚经济收入的同时, 也

① 基金项目: 浙江省基础公益研究计划 (LGN20F020001)

Foundation item: Basic Research Plan for Public Welfare of Zhejiang Province (LGN20F020001)

收稿时间: 2020-11-20; 修改时间: 2020-12-21; 采用时间: 2021-01-08; csa 在线出版时间: 2021-07-31

使得我国的渔业发展遇到了许许多多的问题。随着出海捕捞成本变高,近海捕捞作业能力增强,导致我国渔业资源衰退较为严重和分布范围变得不均衡。由于渔民作业方式不规范及相关部门监管能力不足,海洋生态环境遭到严重的破坏。另外,作为海洋经济价值的重要组成部分,越来越多的沿海国家开始大力发展海洋渔业的科技信息化和现代化。因此,为了能够解决渔业发展面临的问题和应对其他国家的竞争,我国也开始推动渔业信息化的建设。其中,实现现代化至关重要的过程就是实现数据的采集和合理分析应用。AIS系统作为渔业监管中不可或缺的手段,它被广泛应用于海上行驶、信息通信和生产作业等方面。另外,该系统能够实时采集海上作业渔船的位置、速度和转向角等数据,并通过卫星与地面基站进行接收和通信,从而实现作业渔船的海上管理和监控。但是,AIS数据只包含了渔船自身的信息和实时航行数据,并不包含渔船的捕捞方式、作业状态和捕鱼热点等深层次信息,这就要求进一步的通过分析航行状态数据来获取。传统的方法对AIS数据的研究和挖掘不够深入,但随着人工智能的快速发展,将机器学习应用到渔船轨迹的分析中,对实现渔船行为的实时监控、作业渔场的发现、捕捞强度的挖掘和掌握渔业资源变动等问题具有重要的意义^[2]。

为了获取渔船之间的作业关系、发现捕鱼热点以及精准预测渔船未来的轨迹,划分渔船轨迹类型成为了至关重要的一步。然而,划分渔船轨迹类型的主要方式就是计算轨迹相似度,文献[3]中谢彬等基于欧式距离对移动对象进行轨迹相似度的计算,该算法适用于获取的轨迹中轨迹点的个数要跟随时间序列一一对应。文献[4]采用弗雷歇算法对船舶轨迹相似度进行建模,然后使用DBSCAN对轨迹进行聚类,最后得到了经典轨迹模型。Zhen等在文献[5]里整合k-medoids聚类和朴素贝叶斯分类器,实现了船舶作业行为的分类和异常行为的检测。文献[6]使用原始的动态规划算法和DBSCAN聚类实现了渔船轨迹的快速分类识别。涂刚凯等^[7]使用传统的LCSS算法从海量人员时空轨迹数据中挖掘人员之间的关系,但是最长公共子序列算法出现了时间阈值选取的敏感性问题,并且该算法允许跳过某些轨迹点所以是一种非常粗糙的轨迹相似度度量方法。Lee等^[8-11]使用了多次的轨迹聚类算法,对轨迹段的形状、特征角点、线段角度等进行线性化处理来

描述相似度,并对渔船的轨迹实现了线性化的分段和聚类。然而,轨迹的切片标准难以确定,很难将轨迹段划分为同类型的片段,这也导致轨迹的分段聚类方法效果不稳定以及计算复杂等问题^[12]。以上方法虽然都取得了不错的效果,但是均未充分考虑海上捕捞航行轨迹的特点,因而存在种种不足。

本文在对渔船AIS数据进行可视化之后,发现某些渔船的轨迹是很相似的,这是因为捕鱼区的位置在较长时间内是不会发生变化以及渔船之间存在协同作业的行为。针对这种特点,论文提出了一种基于轨迹图像的相似度计算方法,这种方法使用了图像特征点检测、特征点描述与特征点匹配,然后设定阈值来进行轨迹的分类。通过这种方法使得计算轨迹相似度的方法转换为计算图像相似度的方法,从而使计算效率和速度提高,并且结果可视化。经过划分渔船轨迹之后,可以更好的对轨迹数据进行挖掘,能够有效地发现渔船之间的协作关系、训练分类渔船预测模型。

1 相关概念

ORB将FAST算法和BRIEF算法进行了结合,并在两种算法的基础上添加了图像旋转与尺度不变性以及特征提取速度快的优点^[13]。首先,它通过FAST来获取图像的特征点,然后使用BRIEF对特征点进行描述。

1.1 FAST特征点

FAST (Features from Accelerated Segment Test) 是一种快速的特征检测算法。首先,它通过计算随机选取的像素点与其邻域的像素差值,若存在连续数个像素点的灰度值大于预先设定的阈值,即认为该像素点为特征点,但是该特征点不具有旋转不变性^[14]。针对这一个问题,ORB算法使用灰度质心法为每个检测到的特征点添加一个主方向。特征点的灰度矩阵定义为:

$$m_{pq} = \sum_{xy} x^p y^q I(x, y) \quad (1)$$

其中, $I(x, y)$ 为点 (x, y) 处的灰度值,那么图像的质心 C 为:

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad (2)$$

那么质心与特征点之间的夹角定义为该检测点的方向:

$$\theta = \arctan(m_{01}, m_{10}) \quad (3)$$

另外,将 x 和 y 的取值范围设定在半径为 r 的圆形区域内,即 $x, y \in [-r, r]$,这能够提升算法的旋转不变性。

1.2 BRIEF 特征描述符

BRIEF (Binary Robust Independent Elementary Features) 用来对特征点进行描述. 首先, 它在特征点附近随机选取若干像素点, 计算这些点对的二值响应值, 然后由这些响应值构成特征点的描述符.

在该算法中使用长度为 n 的二值码串来描述特征点, 其中, 这个二值码串是由特征点周围的 n 个像素点对生成的, 将这 $2n$ 个点 $(x_i, y_i), i = 1, 2, \dots, 2n$ 组成一个矩阵 S :

$$S = \begin{pmatrix} x_1 & x_2 & \cdots & x_{2n} \\ y_1 & y_2 & \cdots & y_{2n} \end{pmatrix} \quad (4)$$

然后, 使用邻域方向 θ 和对应的旋转矩阵 R_θ , 构建 S 的一个矫正版本 S_θ , 那么:

$$S_\theta = R_\theta S \quad (5)$$

$$R_\theta = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \quad (6)$$

其中, θ 是式 (3) 中特征点求得的主方向.

2 本文算法

计算轨迹相似度的过程中主要是轨迹特征点检测、特征点描述, 最后是特征匹配. 图像特征点检测的关键在于角点检测, 角点的特性是向任何方向移动像素的变化都很大. 本文中使用了 ORB 特征检测和 BF 匹配计算轨迹图片的相似度, 其中 BF 匹配是一种暴力匹配的方式, 它将计算样本图片中的每一个特征点与对照图片中的所有特征点之间汉明距离, 然后选择最近邻和次近邻的两个特征点, 设定最近邻距离除以次近邻距离的比率小于某个阈值来进行特征点筛选.

假设样本图片的特征点个数为 num_x , 样本某一特征点与最近邻特征点 M 、次近邻特征点 N 的距离分别为 dis_M 、 dis_N , 则阈值 η 为:

$$\eta > \frac{dis_M}{dis_N} \quad (7)$$

满足阈值 η 的特征点个数为 num_y , 则相似度 ζ 为:

$$\zeta = \frac{num_y}{num_x} \quad (8)$$

为了计算轨迹图像相似度, 设计了算法 1.

算法 1. 渔船捕捞轨迹图像相似度算法

- 1) 将渔船 AIS 数据转换为可视化图像, 图像只保留轨迹, 并设定轨迹图像大小;
- 2) 从文件夹中读取轨迹图像, 并将轨迹图像转换为灰度图;

- 3) 使用 ORB 特征检测器对轨迹图像进行特征检测, 并对特征点使用 BRIEF 算法进行描述;
- 4) 使用 BF 暴力匹配的方式, 将样本图片中的每一个特征点与对照图片中的所有特征点计算汉明距离;
- 5) 选择最近邻 M 和次近邻 N 两个特征点计算阈值 η . 然后计算满足阈值的特征点的个数占总特征点的比例 num_y/num_x , 此值即为相似度;
- 6) 将相似度满足阈值的轨迹划分为同一类别.

3 实验

3.1 实验数据

渔船 AIS 数据来源于浙江省海洋与渔业局, 其中渔船均在浙江登记, 数据格式如表 1 所示. 由于设备或者其它异常原因会出现一些错误, 要对异常、错误或者不符合范围的数据进行分析和过滤. 首先设定了渔船的海上航行范围: 经度范围 $[105^\circ, 130^\circ]$, 纬度范围 $[10^\circ, 40^\circ]$. 其次设定渔船的速度区间, 因为渔船的行为分为: 抛锚、捕捞、航行, 所以每种行为对应不同的速度区间. 根据统计可得, 渔船抛锚时的速度接近于 0 节, 捕捞时的速度为 0~5.5 节, 正常航行的速度为 5.5~12 节, 所以选择 0~12 节作为速度区间.

表 1 渔船 AIS 数据格式表

字段名称	字段类型	数据样例	描述
渔船ID	INT	167918	船舶的ID, 具有唯一性
时间戳	INT	1459440109	unix时间戳, 单位为秒
纬度	INT	18232946	北纬为正数, 南纬为负数, 单位: 1/600000度
经度	INT	75255639	东经为正数, 西经为负数, 单位: 1/600000度
转向角度	INT	2640	方向, 单位: 0.1度, 正北方向为0度
速度	INT	19	速度, 单位0.1节(Knot)

如图 1, 编号为 29902 的渔船在 2016 年 4 月~2016 年 6 月 3 个月, 渔船的真实轨迹点数据共计 18049 条.

3.2 捕鱼热点实验

针对前文所述, 本文使用 2015 年、2016 年的渔船数据做了热力图对比, 如图 2 所示. 通过实验发现, 沿海区域以及远海区域的捕鱼热点变化较小, 主要是海洋中部区域有些许变化, 捕鱼热点整体呈稳定趋势.

3.3 实验结果

经过本文中提出的轨迹图像相似度计算方法, 可以快速、有效地进行计算和分类, 文中选取部分划分类型的渔船轨迹进行比对. 由图 3、图 4、图 5 所示, 渔船轨迹划分类型准确率高, 各类型的主要轨迹相同. 通过这种方式, 也可以清晰获取渔船之间的协作关系.

以及捕鱼热点的发现. 下文中选取两组渔船的轨迹图像, 通过匹配特征来验证算法.

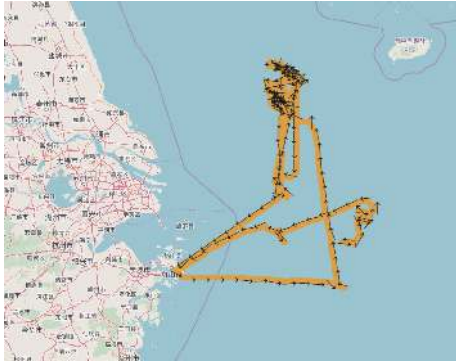
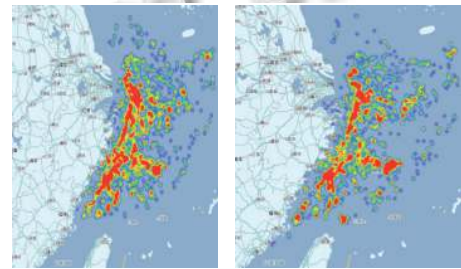


图1 编号为29902的渔船轨迹

如图6所示, 选择两组轨迹图像其部分特征点进行特征匹配, 并且特征点之间匹配度高. 其中, 存在某些特征点匹配发生偏移的现象, 主要原因在于轨迹图像过于繁杂, 导致部分特征模糊.

本文选择传统的轨迹相似度算法 DTW (Dynamic

Time Warping) 和 LCSS (Longest Common SubSequence) 作为对照. 通过表2可知, 在浙江渔船的数据集上, 本文算法在时间和准确率上都有较好的效果, 其中时间一栏代表一条渔船轨迹与其他渔船轨迹匹配一轮所花费的时间. 传统方法的不足之处在于, 一方面两者均是动态算法, 复杂的计算带来了较大的时间消耗. 另一方面, LCSS 只关注于局部匹配, 而 DTW 对噪音和离群点的抑制效果较差. 对于本文算法来说, 本质是将轨迹点之间的计算转换为图像点之间的计算, 这种方法提高了计算效率和准确率.



2015年 2016年

图2 2015年、2016年捕鱼热力图

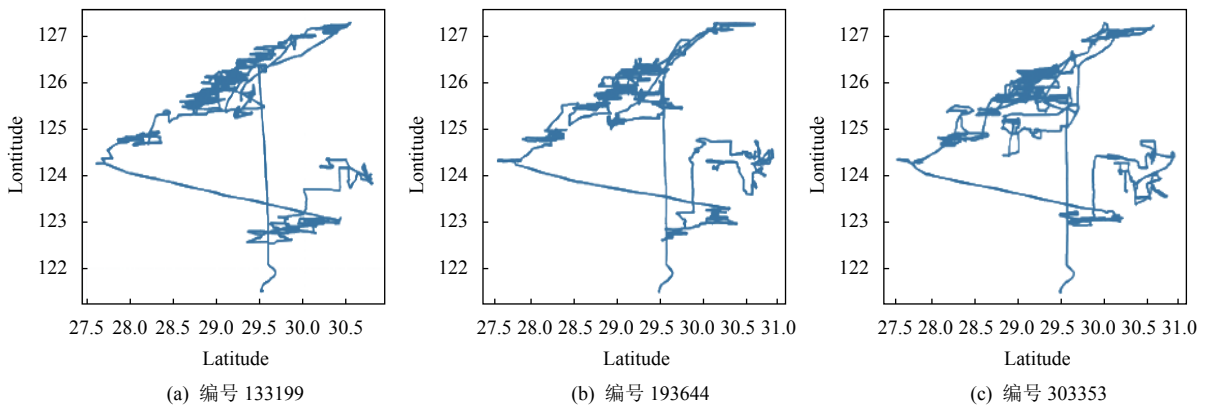


图3 A组部分渔船2016年4月至5月轨迹

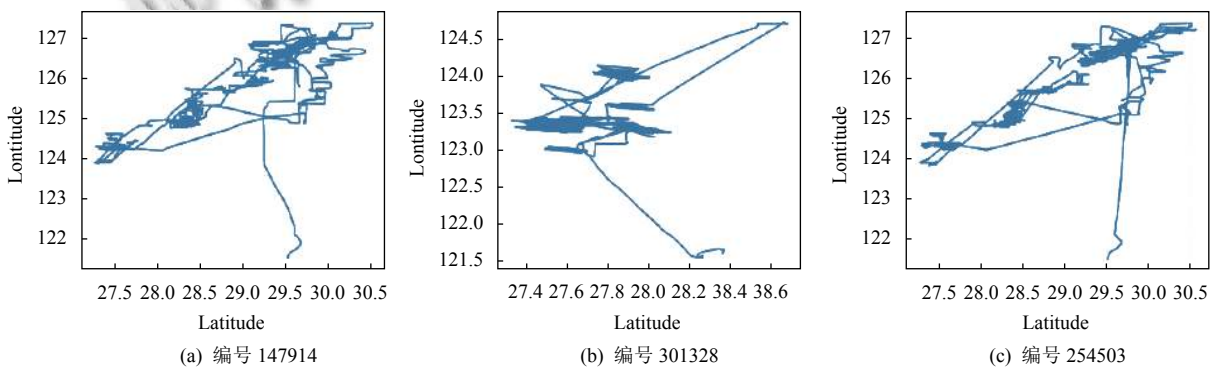


图4 B组部分渔船2016年4月至5月轨迹

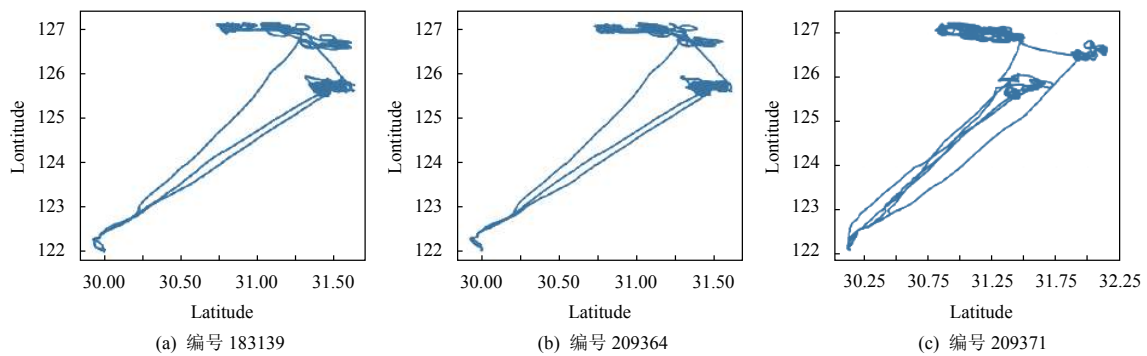


图5 C组部分渔船2016年4月至5月轨迹

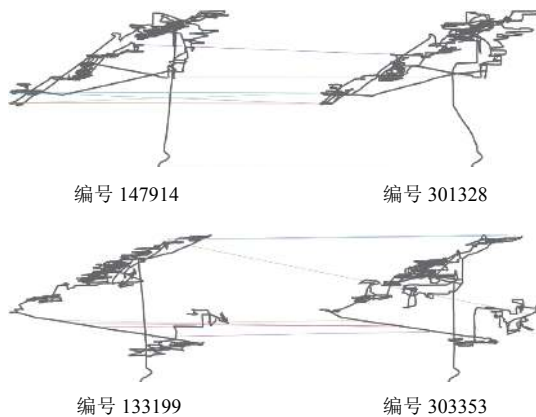


图6 两对渔船轨迹特征点匹配

表2 算法效果对比

算法	时间(s)	准确率(%)
DTW	107	86
LCSS	87	71
本文算法	29	91

4 结论与展望

通过实验验证了本文所提出的图像轨迹相似度算法可以有效并快速地对轨迹进行计算和分类,也验证了捕鱼热点区域随时间变化较小.文中提出的ORB算法和暴力匹配的方法可以对图像进行特征点检测、描述以及匹配,但是这种方法需要多次实验尝试不同的阈值.另外,由实验结果可发现,处理之后的轨迹图像存在冗余的现象.所以接下来的研究工作,一方面要进一步处理渔船AIS数据,压缩轨迹数据;另一方面需要尝试改进算法,将人为设定的阈值改为算法设定.

参考文献

- 1 农村农业部渔业渔政管理局. 2019年全国渔业经济统计公报. 中国渔业报, 2020-06-22(002).
- 2 Takahashi Y, Komeyama K. Simulation of the capture process in set net fishing using a fish-schooling behavior

- model. Fisheries Science, 2020, 86(6): 971-983. [doi: 10.1007/s12562-020-01462-w]
- 3 谢彬, 张琨, 张云纯, 等. 基于轨迹相似度的移动目标轨迹预测算法. 计算机工程, 2018, 44(9): 177-183.
- 4 江玉玲, 熊振南, 唐基宏. 基于轨迹段 DBSCAN 的船舶轨迹聚类算法. 中国航海, 2019, 42(3): 1-5.
- 5 Zhen R, Jin YX, Hu QY, et al. Maritime anomaly detection within coastal waters based on vessel trajectory clustering and Naïve Bayes Classifier. The Journal of Navigation, 2017, 70(3): 648-670. [doi: 10.1017/S0373463316000850]
- 6 Zhao LB, Shi GY. A trajectory clustering method based on Douglas-Peucker compression and density for marine traffic pattern recognition. Ocean Engineering, 2019, 172: 456-467. [doi: 10.1016/j.oceaneng.2018.12.019]
- 7 涂刚凯, 耿鑫. 时空轨迹相似度计算方法研究与实现. 计算机与数字工程, 2020, 48(5): 1114-1120. [doi: 10.3969/j.issn.1672-9722.2020.05.023]
- 8 Lee JG, Han JW, Whang KY. Trajectory clustering: A partition-and-group framework. Proceedings of 2007 ACM SIGMOD International Conference on Management of Data. Beijing, China. 2007. 593-604.
- 9 肖潇, 邵哲平, 潘家财, 等. 基于 AIS 信息的船舶轨迹聚类模型及应用. 中国航海, 2015, 38(2): 82-86. [doi: 10.3969/j.issn.1000-4653.2015.02.020]
- 10 魏照坤, 周康, 魏明, 等. 基于 AIS 数据的船舶运动模式识别与应用. 上海海事大学学报, 2016, 37(2): 17-22, 71.
- 11 彭祥文, 高曙, 初秀民, 等. 基于 Spark 的船舶航行轨迹聚类方法. 中国航海, 2017, 40(3): 49-53, 68. [doi: 10.3969/j.issn.1000-4653.2017.03.011]
- 12 Zheng Y. Trajectory data mining: An overview. ACM Transactions on Intelligent Systems and Technology, 2015, 6(3): 29.
- 13 Rublee E, Rabaud V, Konolige K, et al. ORB: An efficient alternative to SIFT or SURF. Proceedings of 2011 International Conference on Computer Vision. Barcelona, Spain. 2011. 2564-2571.
- 14 Calonder M, Lepetit V, Strecha C, et al. BRIEF: Binary robust independent elementary features. Proceedings of the 11th European Conference on Computer Vision. Heraklion, Crete, Greece. 2010. 778-792.