

## 基于 AIS 的轨迹插值方法

周 鹏<sup>1,2</sup>, 陈金海<sup>1,2</sup>, 林乾锋<sup>1,2</sup>, 梅 强<sup>1,2</sup>, 黄鹏飞<sup>1,2</sup>

(1. 集美大学航海学院, 福建 厦门 361021; 2. 船舶辅助导航技术国家地方联合工程研究中心, 福建 厦门 361021)

**[摘要]** 为了解决常用插值方法在 AIS (automatic identification system) 轨迹插值上存在精度差异, 在拉格朗日线性插值、三次样条插值、时间插值等几种常用插值方法的基础上, 提出一种新的插值方法。新方法通过连续两点间航向差值大小, 将轨迹分为直线和曲线两种, 并结合线性插值和航速航向, 以提高插值精度。最后以实际 AIS 轨迹数据进行插值实验。研究表明, 与其他常用插值方法相比较, 新方法在插值效果上更接近原有轨迹。

**[关键词]** 轨迹插值; AIS; 船舶轨迹; 稀疏轨迹

**[中图分类号]** U 675.7

## Research on Trajectory Interpolation Method Based on AIS

ZHOU Peng<sup>1,2</sup>, CHEN Jinhai<sup>1,2</sup>, LIN Qianfeng<sup>1,2</sup>, MEI Qiang<sup>1,2</sup>, HUANG Pengfei<sup>1,2</sup>

(1. Navigation College, Jimei University, Xiamen 361021;

2. National-local Joint Engineering Research Center for Marine Navigation Aids Services, Xiamen 361021, China)

**Abstract:** The commonly used interpolation method different levels of precision in the AIS (automatic identification system) trajectory interpolation. Several common interpolation methods are introduced, and a new interpolation method is proposed. The new method divides the trajectory into two types, straight line and curve, according to the difference of course differences between two consecutive points, and combines linear interpolation and speed course to improve the interpolation precision. Finally, the interpolation experiment is carried out with the actual AIS trajectory data, and the effects of other common interpolation methods are compared. The results show that the new method is closer to the original trajectory in terms of interpolation effect.

**Keywords:** trajectory interpolation; AIS; ship trajectory; sparse trajectory

## 0 引言

船舶自动识别系统 AIS (automatic identification system) 能够发射并接收船舶的动态信息和静态信息<sup>[1,4]</sup>。其中: 动态信息包括船舶发出信息的时刻, 该时刻船舶经度、纬度值, 船舶航向, 船舶向等<sup>[5]</sup>; 静态信息包括船舶长度、宽度, AIS 天线到船首距离和左右舷距离等。AIS 数据分析研究预处理<sup>[6-9]</sup>一般都包括对原始数据的插值<sup>[2]</sup>, 但常用的几种数值分析插值方法<sup>[10-11]</sup>并没有结合具体水域的交通状况, 忽视了由于插值方法的不同带来的误差<sup>[12]</sup>。为了进一步提高插值精度, 已有学者提出结合专业领域的插值方法<sup>[3]</sup>。王超等<sup>[13]</sup>提出了一种考虑船舶航速航向的 AIS 航迹插值方法, 刘立群

[收稿日期] 2018-03-28

[修回日期] 2018-09-05

[基金项目] 国家自然科学基金项目 (41501490); 福建省本科高校教育教改项目 (JZ160302); 中科院重点部署项目 (ZDRW-ZS-2016-6-3); 福建省教育厅 B 类项目 (JB13135)

[作者简介] 周鹏 (1996—), 男, 硕士生, 从事交通信息工程及控制研究。通信作者: 陈金海 (1980—), 男, 实验师, 硕士生导师, 从事海事大数据与 GIS 研究。E-mail: jhchen@jmu.edu.cn

<http://xuebaobangong.jmu.edu.cn/zkb>

等<sup>[14]</sup>提出三次样条插值结合船舶经纬度的方法; Wang 等<sup>[15]</sup>提出结合空间多个维度构建多维阵列来对轨迹进行插值。这些插值方法结合了 AIS 数据特有的属性,在一定程度上减小了插值方法本身造成的误差,但适用性有限。本文通过分析船舶轨迹研究中常用的插值方法,提出一种结合线性插值和航速航向<sup>[16]</sup>的新插值方法,并以实际 AIS 轨迹数据进行验证。

## 1 插值原理

### 1.1 常用插值方法

在船舶轨迹分析中常用的插值方法可以分为两种:一种是以数值分析为基础,常用的有拉格朗日线性插值、三次样条插值等;另一种是以数理统计结合 AIS 固有属性的方法,主要有结合时间的插值方法和结合航速航向的插值方法。

#### 1) 拉格朗日插值方法

拉格朗日插值法能给出一个恰好穿过二维平面上若干个已知点的多项式函数。如对实际的船舶轨迹数据进行处理,有若干个已知点,拉格朗日插值法可以得到一个多项式,能够包含所有的已知点,最后再由多项式得到插值点的数据。这种方法思路简洁,在算法的实现上也简单,但是由于 AIS 数据的离散性、连续变化的无规律性,这种插值方法精度并不高;而且 AIS 数据量大,根据拉格朗日插值方法构造插值多项式时,插值点越多,插值多项式的次数也越高,当插值次数过高时就会出现龙格现象(即在一定范围内能较好逼近,但越接近插值端点,误差也就越大的现象),这样得到的插值结果误差会很大,还会产生虚假的波动,不具备保形效果。

#### 2) 三次样条插值方法

三次样条插值(Spline 插值)是通过求解三弯矩方程组得出曲线函数组的过程<sup>[17]</sup>,最终得到的是通过一系列形值点的一条光滑曲线。三次插值样条曲线在灵活性和计算速度之间进行了合理的折中。与更高次样条相比,三次插值样条只需较少的计算和存储,且较稳定。与二次插值样条相比,三次插值样条在模拟任意形状时显得更灵活。

#### 3) 结合时间的插值方法

将船舶运动分解为经度和纬度两种运动状态的变化,即将原有的经纬度变化视为(时间—经度)(时间—纬度)两个一维运动的组合,将两种方向上的运动都视为短时间的匀加速直线运动。假设待插值点为  $p_i(\text{Lon}_i, \text{Lat}_i, v_i, \theta_i, t_i)$ , 其中:  $\text{Lon}$  代表经度;  $\text{Lat}$  代表纬度;  $v$  代表船速;  $\theta$  代表航向;  $t$  代表发送时间。前后两点坐标为  $q_1(\text{Lon}_1, \text{Lat}_1, v_1, \theta_1, t_1)$  和  $q_2(\text{Lon}_2, \text{Lat}_2, v_2, \theta_2, t_2)$ , 由通过经纬度插值变为通过时间插值,通过插值时间序列  $T$  得到待插值时间  $t_i$  点处的坐标  $(\text{Lon}_i, \text{Lat}_i)$ , 再结合线性插值方法分别对两种运动状态插值,最后将插值得到的相同时间的经纬度组合,得到最终的插值结果。

### 1.2 改进插值方法原理

船舶轨迹可分为直线轨迹段和曲线轨迹段,若插值点附近航向差值的绝对值在  $5^\circ$  以内,则判定该插值点所处轨迹段为直线轨迹,否则为曲线轨迹<sup>[16]</sup>。

首先将 AIS 的大地坐标值(经纬度)经过高斯—克吕格投影转化为平面坐标值  $(x, y)$ <sup>[18]</sup>; 如果已知插值点前后点高斯坐标  $A(x_1, y_1)$ 、 $B(x_2, y_2)$ , 那么可以分别以  $A$ 、 $B$  两点为基准得到两组插值点的预测坐标

$$\begin{cases} x_{i1} = x_1 + v_1 \sin(\theta_1)(t_i - t_1), \\ y_{i1} = y_1 + v_1 \cos(\theta_1)(t_i - t_1). \end{cases} \quad (1)$$

$$\begin{cases} x_{i2} = x_2 + v_2 \sin(\theta_2)(t_i - t_2), \\ y_{i2} = y_2 + v_2 \cos(\theta_2)(t_i - t_2). \end{cases} \quad (2)$$

式中:  $\theta$  和  $v$  分别为插值基点处的航向和航速;  $t_i$  为插值点的时间,  $t_1$ 、 $t_2$  为插值基点处对应的时间。

对预测坐标  $(x_{i1}, y_{i1})$ 、 $(x_{i2}, y_{i2})$  进行加权平均,权重由插值点距离  $A$ 、 $B$  两点的时间差确定,距离

哪一点时间越近,相应点对应的权重也就越大。即

$$\begin{aligned} Q_1 &= 1 - (t_i - t_1)/(t_2 - t_1), \\ Q_2 &= 1 - (t_2 - t_i)/(t_2 - t_1)。 \end{aligned} \quad (3)$$

式中:  $Q_1$ ,  $Q_2$  分别为  $A$ 、 $B$  两端的权重,则加权平均后的插值点坐标为

$$\begin{cases} x_i = Q_1 x_{i1} + Q_2 x_{i2}, \\ y_i = Q_1 y_{i1} + Q_2 y_{i2}。 \end{cases} \quad (4)$$

最后将相应平面坐标转换为经纬度坐标,即得到最终的插值结果。

## 2 整体实验设计

为了验证改进后的插值方法的插值效果,利用 MMSI 码为 412468324 的船舶在 20150104 这天一段时长为 20 min 的 AIS 轨迹数据,在 MATLAB 中进行仿真实验。对这段轨迹分别采用不同插值方法进行轨迹插值,通过比较与原始轨迹的贴近程度,先验证各种插值方法在轨迹较完整时的效果,再从完整轨迹中取 4 个抽样点(表 1 中加黑的数据点),模拟极端稀疏情况下的 AIS 数据,通过比较不同插值方法,对同一数据插值后与原轨迹的接近程度来进一步验证其方法的优劣程度。

船舶 AIS 原始数据如表 1 所示(表中加黑的为抽样点数据)、对应的坐标图如图 1 所示(图中绿色点为抽样点)

表 1 原始轨迹数据  
Tab. 1 Original trajectory data

Lat/(°)	Lon/(°)	$\theta/(^{\circ})$	$v/\text{kn}$	$t$
20.86040333	111.3120200	209.2	3.1	19:02:37
20.86003000	111.3117617	208.4	3.0	19:03:07
20.85964667	111.3115033	213.6	3.0	19:03:37
20.85928833	111.3112267	206.1	3.2	19:04:07
20.85891667	111.3109650	211.8	3.2	19:04:37
20.85853167	111.3106783	211.2	3.3	19:05:07
20.85781667	111.3101050	217.1	3.3	19:06:07
20.85706667	111.3095517	220.8	3.7	19:07:08
20.85652333	111.3087100	245.0	4.4	19:08:08
20.85636167	111.3081433	261.8	4.1	19:08:38
20.85625500	111.3075600	261.8	4.2	19:09:08
20.85606833	111.3058400	266.6	3.5	19:10:40
20.85600500	111.3047467	280.9	4.2	19:11:38
20.85605667	111.3041850	281.6	3.8	19:12:08
20.85651333	111.3031050	302.7	4.2	19:13:08
20.85684167	111.3026050	304.3	4.1	19:13:38
20.85762167	111.3017550	318.2	4.1	19:14:38
20.85804000	111.3013783	321.5	3.7	19:15:08
20.85898167	111.3007767	338.9	3.6	19:16:08
20.86051000	111.3003183	350.0	3.4	19:17:39
20.86098667	111.3002550	343.1	3.9	19:18:09
20.86147333	111.3002533	4.1	3.7	19:18:39
20.86287667	111.3001850	7.6	3.9	19:20:09
20.86377667	111.3003050	11.5	3.1	19:21:09
20.86461167	111.3004183	18.7	2.5	19:22:08

### 3 实验结果及分析

采用上述 4 种方法分别对原始轨迹点以及抽样点进行插值, 并输出插值图像, 直观对比插值轨迹与原轨迹的趋近程度, 分析改进后的插值方法在曲线轨迹时的效果。

在对原始轨迹的插值过程中, 4 种插值方法的效果如图 2 所示, 除拉格朗日线性插值轨迹较原始轨迹有较小幅度的偏差, 其余 3 种方法都很接近原始轨迹。主要原因在于原始轨迹点较密集, 这些方法的插值误差都很小。这也表明三次样条插值、时间维度插值、改进航速航向插值方法在原始轨迹密集的情况下都可以满足实验需求。

为了更进一步分析这 4 种插值方法的适用范围, 在原始轨迹中提取 4 个抽样点, 只以这 4 个点为已知, 模拟稀疏情况下的船舶轨迹, 输出结果如图 3。

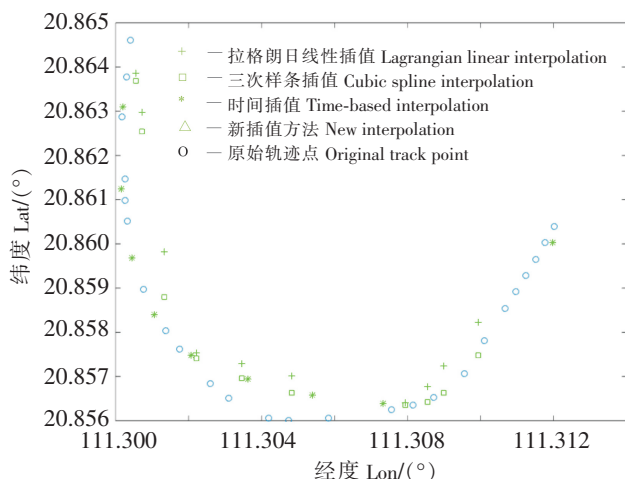


图 2 原始轨迹插值效果

Fig.2 Original track interpolation effects

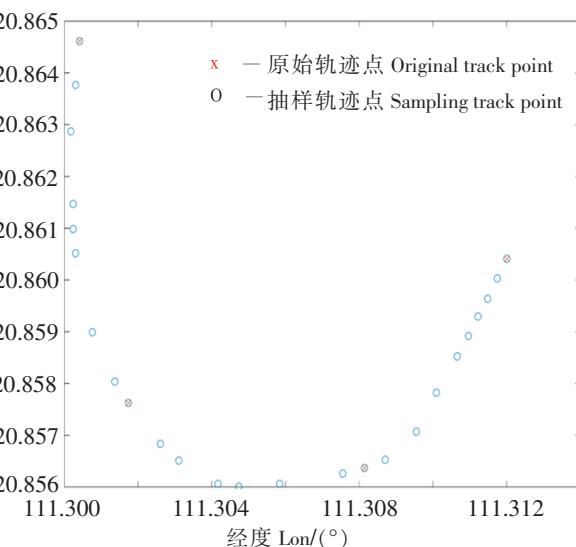


图 1 原始轨迹点及抽样点

Fig.1 Original track points and sampling points

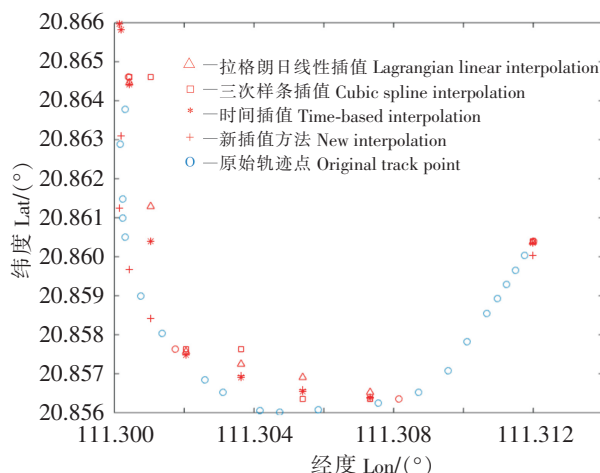


图 3 抽样点插值效果

Fig.4 Sampling point interpolation effect

从图 3 可以看出, 三次样条插值由于初始已知条件太少, 插值得到的轨迹点发生跳变, 轨迹的拟合度最差; 拉格朗日线性插值就是将已知点相连作为拟合轨迹, 插值点也在轨迹上对应取值, 误差太大, 且在插值过程中经常出现轨迹跨越陆地的情况, 也不可取; 基于时间的插值方法虽然在经度  $111.302^{\circ} \sim 111.308^{\circ}$  区域内插值效果与改进后的插值方法精度相差不大, 但在第一个抽样点到第二个抽样点范围内误差很大, 并且由于原始数据时间间隔不统一, 插值还出现了插在抽样点范围以外的情况; 改进后的新插值方法在整体上拟合原轨迹效果最好, 插值精度误差也最小, 证明改进后的插值方法在弯曲稀疏轨迹段也能取得较为理想的效果。

通过对比这些插值方法, 可以发现单纯以数值分析为原理的插值方法在 AIS 轨迹插值上效果并不理想, 主要原因在于远海 AIS 数据稀疏, 在插值时提供不了数值分析插值方法所需的大量临近已知点, 并且 AIS 中的船舶轨迹数据点并不同于数学上的离散函数点, 很难发现其数学意义上具有的函数规律, 因此当 AIS 数据大量缺失时, 轨迹插值的最好方法还是结合航速、航向、时间这些数据, 将轨迹分类处理, 直线轨迹采用数值分析方法插值, 保质并且高效, 曲线轨迹采用改进后的航向航速插值

方法,其精度得到了明显提高。

## 4 结论

通过总结现有 AIS 数据的插值方法,提出一种用于海上交通研究分析的新插值方法。在前 3 种插值方法的基础上进行改进,将轨迹曲线分类处理,按照插值前后点间航向的变化,将轨迹分为直线和曲线两种轨迹段,直线轨迹段各插值方法精度区别不大,因此采用分段线性插值方法,插值精度和效率都能得到保证;曲线轨迹段插值较为困难,因此结合 AIS 航速航向以及时间这些属性,分别以两端点坐标和航速航向为基准得到两组预测差值坐标,再加入时间作为权重,实验结果表明插值坐标精度有了明显提高,并且在插值点前后时间间隔很长的情况下效果较其他方法更好,尤其在曲线轨迹的插值上也明显优于其他方法。因此此种新插值方法在船舶数据分析上具有一定的实际意义。

## [ 参考文献 ]

- [1] 陈金海,陆锋,李明晓. 海上主航迹带边界统计推断与海西航路警戒区布局优化分析 [J]. 地球信息科学学报, 2015, 17(10): 1196-1206.
- [2] 陈金海,陆锋,彭国均. 海洋运输船舶轨迹分析研究进展 [J]. 中国航海, 2012, 35(3): 53-57.
- [3] CHEN J, LU F, PENG G. A quantitative approach for delineating principal fairways of ship passages through a strait [J]. Ocean Engineering, 2015, 103: 188-197.
- [4] 甄荣,邵哲平,潘家财,等. 基于 AIS 信息的航道内船舶速度分布统计分析 [J]. 集美大学学报(自然科学版), 2014, 19(4): 274-278.
- [5] ZHAO L, SHI G, YANG J. Ship trajectories pre-processing based on AIS data [J]. Journal of Navigation, 2018, 71(5): 1210-1230.
- [6] 潘家财,邵哲平,姜青山. 数据挖掘在海上交通特征分析中的应用研究 [J]. 中国航海, 2010, 33(2): 60-62.
- [7] 高强,张凤荔,王瑞锦,等. 轨迹大数据:数据处理关键技术研究综述 [J]. 软件学报, 2017, 28(4): 959-992.
- [8] YANG C, HU Q, TU X, et al. An integrated vessel tracking system by using AIS, inmarsat and china beidou navigation satellite system [J]. Transnav the International Journal on Marine Navigation & Safety of Sea Transportation, 2012, 6(2): 175-178.
- [9] ZHANG L, MENG Q, XIAO Z, et al. A novel ship trajectory reconstruction approach using AIS data [J]. Ocean Engineering, 2018, 159: 165-174.
- [10] 陈志军,吴超仲,吕能超,等. 基于改进三次 Hermite 插值的车辆时空轨迹重构研究 [J]. 交通信息与安全, 2013, 31(6): 43-46.
- [11] SANG L Z, WALL A, MAO Z, et al. A novel method for restoring the trajectory of the inland waterway ship by using AIS data [J]. Ocean Engineering, 2015, 110(A): 183-194.
- [12] 张海燕,吴方. 几种插值算法的比较研究 [J]. 计算机技术与发展, 2012, 22(2): 234-237.
- [13] 王超,纪永刚,黎明,等. 一种考虑船舶航速航向的 AIS 航迹插值方法 [J]. 舰船科学技术, 2015, 37(4): 60-64.
- [14] 刘立群,吴超仲,褚瑞峰. 基于 Vondrak 滤波和三次样条插值的船舶轨迹修复研究 [J]. 交通信息与安全, 2015, 33(4): 100-105.
- [15] WANG J, ZHU C, ZHOU Y, et al. Vessel spatio-temporal knowledge discovery with AIS trajectories using co-clustering [J]. Journal of Navigation, 2017, 70(6): 1383-1400.
- [16] 戴鹏睿. AIS 轨迹动态插值在实时视景显示中的应用 [J]. 电子设计工程, 2016, 24(14): 172-175.
- [17] 田利芹,贾杰. 三次样条插值在船舶轨迹修复中的研究 [J]. 舰船科学技术, 2016, 38(3A): 79-81.
- [18] 宗文鹏,李广云,王力,等. 直接地理参考中轨迹插值算法对精度影响的评估 [J]. 测绘通报, 2015(11): 35-38.

(责任编辑 陈 敏 英文审校 周云龙)