

李忠伟, 焦方涛, 李永, 等. 信息熵控制的海洋流线自适应步长算法研究[J]. 海洋学报, 2024, 46(12): 100–110, doi:10.12284/hyxb2024115
Li Zhongwei, Jiao Fangtao, Li Yong, et al. Research on adaptive step size algorithm of marine streamline controlled by information entropy[J].
Haiyang Xuebao, 2024, 46(12): 100–110, doi:10.12284/hyxb2024115

信息熵控制的海洋流线自适应步长算法研究

李忠伟¹, 焦方涛¹, 李永^{1*}, 曾伟¹, 杨俊钢², 崔伟²

(1. 中国石油大学(华东)海洋与空间信息学院, 山东 青岛 266580; 2. 自然资源部第一海洋研究所, 山东 青岛 266061)

摘要: 海洋流场的流线构造与放置对于认识和理解海洋流场具有重要意义。在流线绘制过程中, 积分步长的选择至关重要, 能够直接影响流线放置效果。固定步长算法由于无法适应曲率不断变化的情形往往不被采用, 以往自适应步长流线算法存在自由度低以及多尺度适用性差的问题。针对上述问题, 本文首次将信息熵引入到步长计算中, 提出了信息熵控制的海洋流线自适应步长算法。该算法首先通过计算流场信息熵得到熵场, 然后依据熵值大小将流场划分为高熵区域和低熵区域, 并为每个积分点赋予新的步长, 使得流场可以根据变化剧烈程度自适应调整步长大小, 即高熵区域(变化剧烈的区域)步长较小, 低熵区域(变化平缓的区域)步长较大。实验结果表明, 本文算法能够显著增加变化剧烈区域的积分点数目和流线条数, 更好地绘制特征处的流线细节, 同时在不会影响放置效果的前提下减少了非重要区域的积分点数目和流线条数以提高计算效率, 相比以往自适应步长算法显著提高了步长调节的自由度以及尺度适用性, 可以应用于不同尺度的海洋流场。

关键词: 海洋流场; 信息熵; 自适应步长; 流线放置; 数值积分

中图分类号: TP391

文献标志码: A

文章编号: 0253-4193(2024)12-0100-11

1 引言

随着海洋探测技术的不断发展, 人们对于海洋的认知需求和表现手段也在逐步提升, 流场可视化作为科学计算可视化中的一个重要组成部分, 能够将复杂的海洋数据转换成简单直观的图形, 以此来帮助人们更加直观地观察到海洋的运动规律、结构特征和拓扑信息。其中流场通常被定义为某个区域中离散网格上的矢量场, 其表达方法有很多, 主要可以分为直接法、纹理法、几何法、特征法和信息论法^[1]。其中几何法是指通过使用流线来表达流场信息, 能够清楚地展示流场的结构信息, 突出流场中的重要区域, 但该方法容易造成遮挡、视觉杂乱等问题, 严重影响人们对流场的理解和认识。针对该问题, 多种解决遮挡

和杂乱的方法被提出。表意性可视化方法着重于信息传递, 即以清晰和可理解的方式可视化数据, 增强可视化结果的表达能力, 目前在基于面的可视化、二维和三维数据可视化中被广泛应用^[2]。此外, 基于聚类算法的流线筛选方法也是解决流线遮挡和杂乱的一个重要手段^[3-4]。

流线放置的过程主要分为两个阶段: 种子点放置和流线积分。后者需要设置积分步长, 研究人员通常会在实际操作中将其设定为一个固定值, 但采用固定步长往往会导致生成的流线无法适应积分曲线多种曲率变化的情形, 使得积分曲线精度变差, 产生大量短流线并造成数据冗余, 扰乱整个流场的连续性。针对这些问题, 多种自适应步长算法被提出: 如通过计算流线上相邻 3 个样本点连接而成的 2 条流线之间

收稿日期: 2024-07-16; 修订日期: 2024-10-08。

基金项目: 国家自然科学基金重点基金(62231028)。

作者简介: 李忠伟(1978—), 男, 山西省晋城市人, 教授, 主要从事大数据处理与人工智能算法及其智慧应用方面的研究。E-mail: lizhongwei@upc.edu.cn

* 通信作者: 李永, 高级实验师, 主要从事中尺度涡方向、人工智能方向等的研究。E-mail: 20030019@upc.edu.cn

的夹角控制步长^[5-6]。该方法考虑了相邻向量点的流向,通过设定角度阈值来进行步长的自动调整,但由于仅仅只设置了两个阈值,导致步长大小只有3种情况,因此调整的自由度偏低。之后由于综合考虑了局部网格流速和流向,一种双自由度调整的自适应步长算法(AMFCA算法)被提出^[7],该算法相比以往算法提高了步长调节的自由度,能够使相邻两网格中同向矢量积分步长快速增加,反向矢量的积分步长快速减小,实现了全球流场主流线的生成及均匀化布置。李婷等^[8]对AMFCA算法进行了改进,通过重新限定积分步长以及参数的范围,解决了AMFCA算法中流线迭代死循环和追踪不连续等问题,使流线构造时具有更高的计算效率、更少的数据存储量和更合理的流线分布。该方法完善了AMFCA算法的诸多不足,但仍存在以下缺陷:(1)无论是AMFCA算法还是改进后的算法目前仅应用于全球流场,对于中低尺度下的流场细节绘制较为粗糙;(2)算法中参数的取值与数据关联性较强,不能直接应用于其他数据,缺少一定的泛用性。

如今,信息熵^[9]在海洋流场中得到了广泛应用,例如利用信息熵放置种子点^[10]、检测流场特征^[11]、确定噪声群的密度^[12]等,还未有学者将其引入到步长计算中。因此,为解决上述问题,本文首次将信息熵引入到步长计算中,首先通过计算流场信息熵得到熵场,然后综合考虑积分点熵值所占流场最大熵值的比例以及步长控制参数后为每一个积分点赋予新的步长,使熵值较高的区域(流场复杂区域)具有较小的步长,熵值较低的区域(流场平缓区域)具有较大的步长,相比以往自适应步长算法提高了步长调整的自由

度以及尺度的适用性。

2 流线生成方法

2.1 种子点放置

在流线放置过程中,种子点放置是整个过程的第一步。作为绘制流线的起始点,种子点的如何放置对于流线放置效果的好坏是一个重要的评判条件。一个好的种子点放置方法可以帮助促成覆盖整个流场并产生均匀而均匀的流线,同时还能减少后期对于绘制结果的计算成本^[13]。研究人员对于种子点放置问题已经做了大量的研究并形成了比较完备的理论体系。种子点的放置方法一般可以分为随机种子点放置方法、均匀种子点放置方法、基于特征的种子点放置方法、最远距离种子点放置方法、基于信息论的种子点放置方法等^[14-16],图1展示了其中3种种子点放置效果。上述种子点放置方法一般只针对结构化网格,对于非结构化网格,也有其独特的放置方法^[17]。不同的种子点放置方法各有特点,对于流场绘制时的侧重点也不同,随机种子点放置方法即利用一个简单的随机模型,在规定大小的场内放置种子点,该方法的优点是简单易实现,但会出现信息遗漏的问题。最远距离种子点放置方法是在生成一条流线的前提下在其周围和边界处放置种子点,该方法注重生成成长流线。基于特征的种子点放置方法和基于信息论的种子点放置方法则更加注重刻画流场特征,但忽略了平缓区域的绘制。相比上述放置方法,均匀种子点放置方法能够将生成的种子点均匀分布在流场中,生成的流线也能够较为均匀地覆盖整个流场,因此本文采用均匀种子点放置方法来验证算法的有效性。



图1 不同种子点放置效果

Fig. 1 Placement effects of different seed points

a. Places seed points based on features; b. randomly places seed points; c. evenly places seed points

a. Places seed points based on features; b. randomly places seed points; c. evenly places seed points

2.2 流线构造方法

流线是一条与流场处处相切的线,流线上的每个点都代表了流场在该处的切线方向。流线构造主要有两种方法:流函数生成法和数值积分法。在实际计

算中,数值积分方法构造的流线计算简单,更符合流体计算中的离散思想,成为了主流流线构造方法。目前数值积分法主要包括一阶欧拉法、二阶龙格库塔法和四阶龙格库塔法。欧拉法计算精度相比另外两

种方法略低但计算速度较快,尤其是在大范围流场计算时能够节省很多时间,因此本文采用欧拉法来进行流线积分,欧拉法的计算公式如下:

$$X(t + \Delta t) = X(t) + \Delta t \times V(X(t)), \quad (1)$$

式中, $X(t)$ 为流线的当前点, Δt 为积分步长, $X(t + \Delta t)$ 为下一点的位置, $V(X(t))$ 为当前点的矢量值。流线生成方式如图 2 所示,其中 μ_0 为种子点, μ_1 和 μ_2 依次为种子点经积分后计算得到的下一个点,将这些点连接起来即可生成流线。

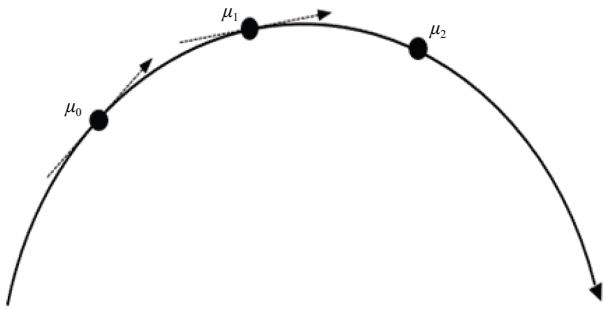


图 2 流线生成方式

Fig. 2 Streamline generation method

3 信息熵控制的海洋流线自适应步长算法

3.1 数据集介绍

实验数据为哥白尼海事局提供的全球海洋物理再分析数据,格式为 NETCDF,空间范围覆盖全球,时间跨度自 1993 年 1 月 1 日至 2024 年 6 月 25 日,该数据常用于海表面涡旋识别、流线放置等领域。本文选取的数据时间分辨率为逐日,空间分辨率包含 0.083° 和 0.04° 两种,包含了经纬度、温度、盐度、流场 U 分量、流场 V 分量等多维变量,实验选择印度洋海域作为实验区域。

3.2 熵场计算

1948 年 C. E. Shannon 等首次提出信息熵的概念。信息熵也称香农熵,用于度量信息的不确定性或复杂程度。当信息的不确定性越大或者复杂程度越高时信息熵越大;信息的不确定性越小或者复杂程度越低时信息熵越小。从数学角度可将其描述为离散型随机变量出现的概率值,即给定一个变量 X , X 可能的取值为 $\{x_1, x_2, \dots, x_n\}$, 若已知每个值出现的概率为 $p(x_i)$, 那么该变量的熵值可由以下公式计算:

$$H(X) = - \sum_{i=1}^n p_i \times \log_2 p_i. \quad (2)$$

由该式可以看出,当每个值出现的概率相等时,

即 X 的分布完全均匀时信息熵最大,此时信息量最大;反之,当变量 X 取某个值的概率为 1,其他值的概率为 0 时,即分布完全不均匀时信息熵的值为 0,此时信息量最小。

2010 年 Xu 等^[9] 将信息熵的概念引入到海洋流场中,并提出了一种基于信息熵的可视化框架,通过该框架可以计算流场的信息熵,具体方法如下:将流场按照矢量方向 ($0^\circ, 360^\circ$) 划分为 n 个大小相同的子区间,记为 $x_i, i = 1, 2, \dots, n$, 计算每个矢量的方向并根据结果将该矢量放入对应的子区间中。计算完成后统计每个子区间内的矢量个数 $C(x_i), i = 1, 2, \dots, n$, 根据公式 (3) 计算出每个子区间内矢量出现的概率 $p(x_i)$, 然后就可以根据公式 (2) 计算出信息熵。计算完成后该框架选取了熵值较高的区域放置种子点生成了一部分流线,随后根据信息量的测量结果继续迭代放置种子点。

$$p(x_i) = \frac{C(x_i)}{\sum_{i=1}^n C(x_i)}. \quad (3)$$

为了得到矢量场的信息分布情况就需要计算整个矢量场的熵, Xu 等^[9] 以某个矢量数据为例进行计算,若该矢量数据存在一个正方形邻域 L , 用该邻域的熵值近似等于矢量数据的熵值即可计算出整个矢量场的熵。具体计算过程如下:

(1) 计算矢量数据角度

根据矢量数据的 U 、 V 值可以通过公式 (4) 计算其角度,其中 $angle1$ 为第一象限计算公式, $angle2$ 为二三象限计算公式, $angle3$ 为第四象限计算公式^[18]。

$$\begin{cases} angle1 = \arctan\left(\frac{v}{u}\right) \times \frac{180}{3.14}, \\ angle2 = \arctan\left(\frac{v}{u}\right) \times \frac{180}{3.14} + 180, \\ angle3 = \arctan\left(\frac{v}{u}\right) \times \frac{180}{3.14} + 360. \end{cases} \quad (4)$$

(2) 建立概率统计模型

对某个矢量数据的邻域建立概率统计模型,通过该模型计算该矢量的信息熵值。首先将该邻域按照角度分为 N 份, Xu 等^[9] 经多次试验,发现二维矢量场当 N 取 60, 邻域大小 L 取 13 时能够取得较好的效果。因此在结构化网格数据中将矢量方向分为 60 份,并以某个矢量数据为中心点取其周围邻域内 13×13 的正方形作为局部邻域计算该点处的信息熵。

局部邻域内的每个点都可以根据公式 (5) 计算其所在的角度空间,计算完成后即可统计每个角度空间内的矢量数据个数。设某正方形邻域的中心点为 i ;

该邻域内矢量数不为0的角度空间个数为 $bin[i], i=1, 2, \dots, M; bin[i][j], j=1, 2, \dots, bin[i]$ 为该邻域内第 j 个角度空间中的矢量个数; $sumbin[i]$ 为该邻域内的矢量总数。由于所取邻域为 13×13 所以 $sumbin[i]$ 为169, 通过公式(6)即可计算出邻域内第 j 个区间的概率 p_j 。

$$I = \frac{floor(angle)}{6}, \quad (5)$$

$$p_j = \frac{bin[i][j]}{sumbin[i]}. \quad (6)$$

(3) 计算局部熵值生成熵场

对矢量场内所有存在邻域的网格点通过上述方法计算概率 p_j , 然后由公式(7)计算出邻域的熵值, 也就是该网格点的熵值, 对不存在邻域的网格点采用原始方法计算即可^[9]。

$$entropy[i] = \sum_{j=1}^{bin[i]} p_j \times \log_2 \left(\frac{1}{p_j} \right). \quad (7)$$

图3所用数据是空间分辨率为 0.083° 的原始数据, 通过提取数据中的经纬度以及 U, V 分量分别计算得到了多个典型时刻的二维流场图(图3a, 图3b, 图3c)和二维熵场图(图3d, 图3e, 图3f)。图中使用不同的颜色来表示不同的熵值大小, 黄色和橙色部分为熵值较高的区域, 蓝色和绿色部分为熵值较低的区域。通过对比流线图和熵场图可以发现流场变化剧烈的区域具有较高的熵值, 流场变化平缓的区域具有较低的熵值。

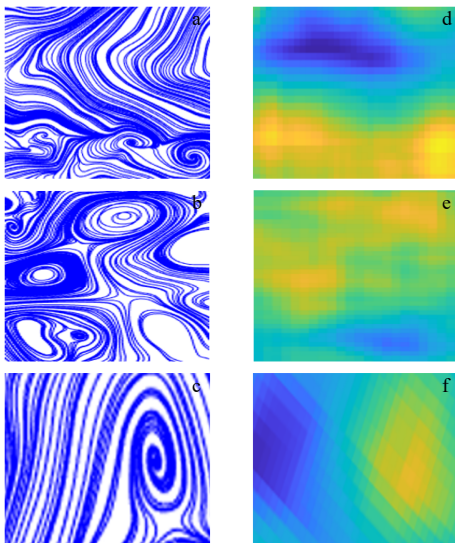


图3 二维流场和二维熵场
Fig. 3 Two-dimensional flow field and two-dimensional entropy field
a, b, c. 二维流场图; d, e, f. 二维熵场图
a, b, c. Two-dimensional flow field map; d, e, f. two-dimensional entropy field map

3.3 信息熵控制的海洋流线自适应步长算法

前文提到积分步长的选择至关重要, 能够直接影响流线的可视化效果。通常情况下较小的积分步长能够很好地刻画流场细节, 但计算时间较长, 较大的积分步长能够缩短计算时间, 但无法很好地刻画流场细节。图4所用数据是空间分辨率为 0.083° 的原始数据。从原始数据的部分区域中提取了经纬度以及 U, V 分量后分别使用不同固定步长算法绘制的流线放置效果图。其中图4a的放置步长为0.2, 图4b的放置步长为0.3。由于积分步长是一个粒子距离调整参数, 用于控制计算精度和流线长度, 取值会随着流场尺度大小进行变化, 因此没有具体单位。从图4可以看出不合适的步长大小会导致流场丢失很多细节, 尤其是在变化剧烈的区域, 如图中左上角的特征处。但对于变化平缓的区域来说, 步长的选择对流线放置效果影响较小, 如图中那些具有较低曲率的长流线。

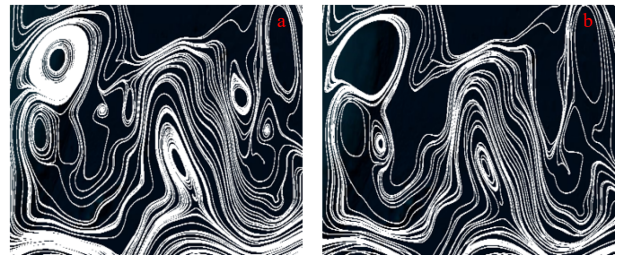


图4 不同固定步长下流线放置效果

Fig. 4 Streamline placement effects with different fixed step sizes

a. 步长为0.2的固定步长算法; b. 步长为0.3的固定步长算法
a. Fixed-step algorithm with a step size of 0.2; b. fixed-step algorithm with a step size of 0.3

由图3可以看出流场中某点熵值大小与该点处流变化的剧烈程度有关。结合图4我们就可以得到以下结论: 某点的熵值越高, 可以认为该点处的流场越复杂, 那么在流线积分过程中应给予较小的步长, 以便更好地绘制流场细节; 某点的熵值越低, 可以认为该点处的流场越平缓, 在流线积分过程中应给予较大的步长, 以便提高计算效率。依据该结论, 算法可以在流线积分过程中实时为每个积分点设置新的步长, 克服了以往自适应步长算法存在的诸多问题, 算法具体模型如公式(8)所示:

$$h_i = d_{max} - \mu \times \left(\frac{e_i - e_{min}}{e_{max} - e_{min}} \right), \quad (8)$$

式中, h_i 代表某点的步长, d_{max} 为设定的步长最大值, μ 为步长控制参数, 取值范围为 $0 < \mu < d_{max}$, 用户可以根据具体情况调整该参数的取值, e_i 为当前点的熵值, e_{min} 为流场最小熵值, e_{max} 为流场最大熵值。本算

法综合考虑流场的步长取值范围和流场信息熵两个要素共同控制步长大小,使每个积分点拥有独特的步长,大大提高了步长调节的自由度。此外,由于流场信息熵的计算与数据本身的分辨率无关,因此通过使用流场信息熵作为控制步长大小的因素可以有效缓解以往算法多尺度适用性差的问题。

4 实验结果分析

本节的对比实验区域如图 5 所示,将它们分别定义为流场 A 和流场 B,两组数据的空间分辨率分别为 0.083° 和 0.04° 。本文用于展示流线放置效果的地球模型为 Cesium,该引擎允许用户自行导入数据,能够在浏览器或移动设备上展示 3D 地球、地图和地理



图 5 实验区域

Fig. 5 Experimental area

数据,因此本文依托该引擎实现对流线放置效果的可视化。

由于不同流场需要设置的积分步长大小各不相同,因此本文采用了最常见的 3 种固定步长值进行对比。图 6 分别对应用于流场 A 中步长大小分别为 0.1、0.2、0.3 的 3 种固定步长算法、考虑速度方向变化的自适应步长算法、改进后的 AMFCA 算法以及本文算法的流线放置效果进行可视化。将红色虚线方框区域定义为区域一,黄色虚线方框区域定义为区域二,区域一中的绿色实线方框内的 3 个特征依次定义为特征 1、特征 2 和特征 3,种子点数目与放置方法均相同。其中图 6a、图 6b 和图 6c 分别是 3 种固定步长算法绘制的流场,图 6d 是考虑速度方向变化的自适应步长算法绘制的流场,图 6e 是改进后的 AMFCA 算法绘制的流场,图 6f 为本文算法绘制的流场。从可视化效果方面来看,以往算法对区域一的绘制效果较差,尤其是在变化剧烈的区域均存在绘制不完全的情况。具体看来图 6c、图 6d 对特征 1 的绘制效果较差,其中图 6c 更是缺少了该特征。图 6a、图 6b 以及图 6d 对特征 2 的绘制效果较差,存在大量空白区域。除图 6d 和图 6f 外,其余算法绘制的特征 3 也存在大量空白。由于区域二为平缓区域,流线曲率变化较小,因此步长大小对该区域流线放置效果影响很小。综上,本文算法不仅能够有效解决区域一特征处绘制不完全的

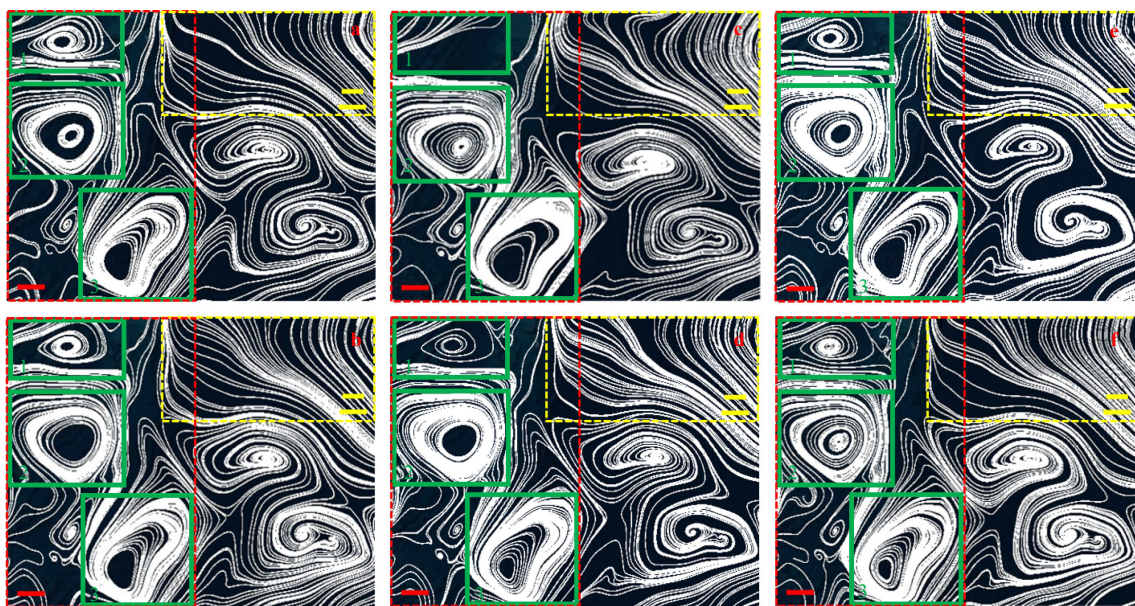


图 6 流场 A 流线放置效果对比

Fig. 6 Comparison of flow field A streamline placement results

- a. 步长为 0.1 的固定步长算法; b. 步长为 0.2 的固定步长算法; c. 步长为 0.3 的固定步长算法; d. 考虑速度方向变化的自适应步长算法; e. 改进后的 AMFCA 算法; f. 本文算法
- a. Fixed-step algorithm with a step size of 0.1; b. fixed-step algorithm with a step size of 0.2; c. fixed-step algorithm with a step size of 0.3; d. adaptive step size algorithm that takes into account changes in velocity direction; e. improved AMFCA algorithm; f. this paper's algorithm

问题,并且在区域二设置较大步长的情况下,与其他算法相比也没有显著区别。

流线是由一系列积分点组成的,因此流场在存在特征的情况下除了可视化流线放置效果外还可以通过统计特征区域内的积分点数目和流线条数辅助评价算法的可行性。本文算法实际上也是在确保计算效率大致相同的情况下通过控制每条流线上的积分点数目并间接影响某些区域内的流线条数来放置流线。由于高熵区域存在更多的流场特征,因此我们希望该区域有更多的流线数目和积分点数目来描绘这些特征。低熵区域存在较少流场特征,设置较大的步长也不会导致出现信息遗漏,因此在不影响绘制效果的前提下积分点数目应越少越好,流线也应尽量避免重叠和杂乱情况,以便减小后期数据的存储量。表1、表2分别对流场A区域一、区域二中积分点数目和流线条数进行了统计。

图7为流场A的流线放置效果图、温度图和海表面高度图,其中红色部分代表值较高的区域,蓝色部分代表值较低的区域。由于流场特征处存在海表温度、海表面高度异常的特性,因此综合考虑温度图

表1 流场A区域一中不同算法对比结果
Table 1 Comparison results of different algorithms in region 1 of flow field A

图像	积分点数目/个	流线条数/条
图6a	91 707	105
图6b	89 604	127
图6c	83 025	115
图6d	79 807	118
图6e	78 832	113
图6f	97 789	130

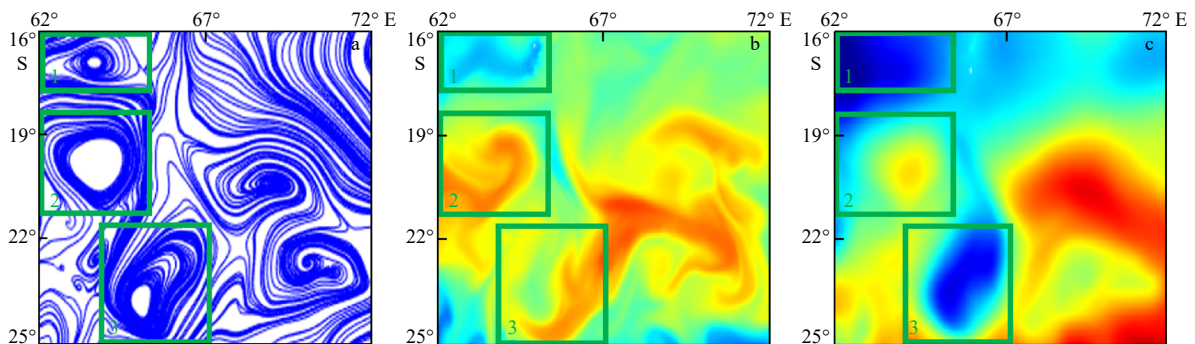


图7 流场A流线、温度和海表面高度

Fig. 7 Streamline, temperature and sea surface height of flow field A

a. 流线图; b. 温度图; c. 海表面高度图

a. Streamline map; b. temperature map; c. sea surface height map

表2 流场A区域二中不同算法对比结果
Table 2 Comparison results of different algorithms in region 2 of flow field A

图像	积分点数目/个	流线条数/条
图6a	4 931	67
图6b	4 827	75
图6c	4 649	65
图6d	4 761	83
图6e	7 677	64
图6f	4 181	64

和海表面高度图后可以直观地看出绿色方框内存在异常值,即证明这3处特征存在。

图8分别对流场B中不同算法的流线放置效果进行可视化,可以看到本文算法对区域一即高熵区域的绘制效果仍优于其他算法,能够绘制出该区域所有的特征,区域二的绘制效果与其他算法相比也没有显著区别。表3、表4分别对流场B中两个区域的积分点数目和流线条数进行了统计。图9为流场B的流线放置效果图、温度图和海表面高度图。

为证明上述实验结果代表客观规律,我们从欧洲中期天气预报中心(European Centre for Medium-Range Weather Forecasts)获得了同时期的卫星数据,如图10所示绘制了流场A和流场B的流线图并与本文实验结果进行对比。通过对比总体海流趋势以及绿色方框内的特征可以看到两种数据得到的流线图基本相同。

为进一步证明本文算法相比于固定步长算法和两种以往自适应步长算法具有优势,分别从计算结果运行时间、本文算法步长的空间分布以及流态更复杂区域的放置效果对比3个方面进行验证。表5和表6分别展示了上述两个实验中不同算法结果的运

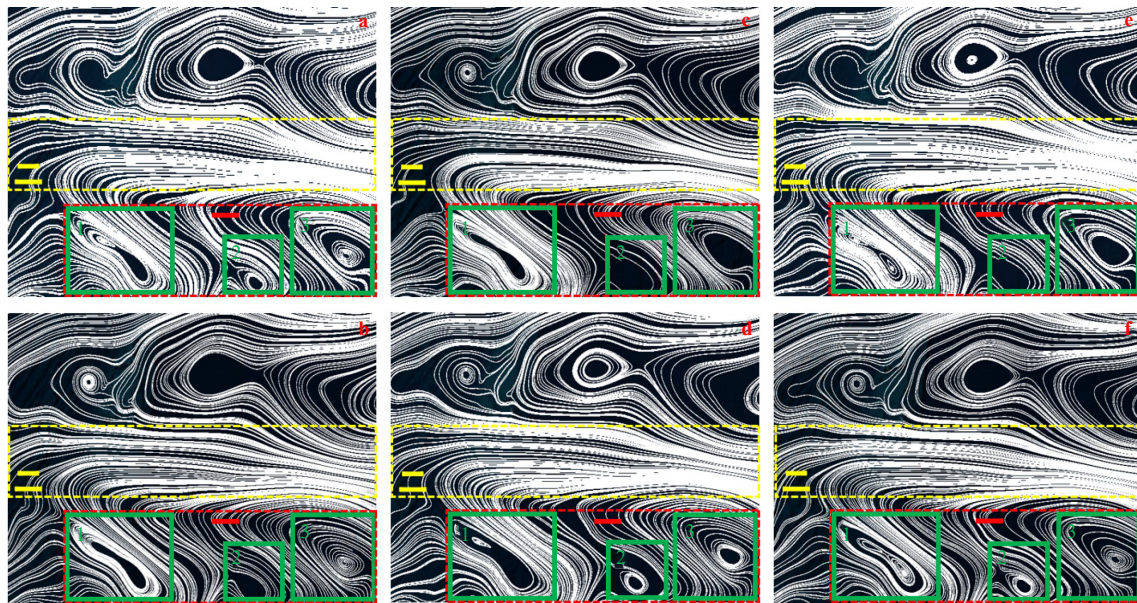


图 8 流场 B 流线放置效果对比

Fig. 8 Comparison of flow field B streamline placement results

a. 步长为 0.1 的固定步长算法; b. 步长为 0.2 的固定步长算法; c. 步长为 0.3 的固定步长算法; d. 考虑速度方向变化的自适应步长算法; e. 改进后的 AMFCA 算法; f. 本文算法

a. Fixed-step algorithm with a step size of 0.1; b. fixed-step algorithm with a step size of 0.2; c. fixed-step algorithm with a step size of 0.3; d. adaptive step size algorithm that takes into account changes in velocity direction; e. improved AMFCA algorithm; f. this paper's algorithm

表 3 流场 B 区域一中不同算法对比结果

Table 3 Comparison results of different algorithms in region 1 of flow field B

图像	积分点数目/个	流线条数/条
图8a	28 835	92
图8b	23 727	97
图8c	20 048	99
图8d	20 627	101
图8e	24 352	100
图8f	29 817	108

表 4 流场 B 区域二中不同算法对比结果

Table 4 Comparison results of different algorithms in region 2 of flow field B

图像	积分点数目/个	流线条数/条
图8a	20 348	153
图8b	20 033	146
图8c	18 252	142
图8d	25 575	141
图8e	27 521	143
图8f	15 938	137

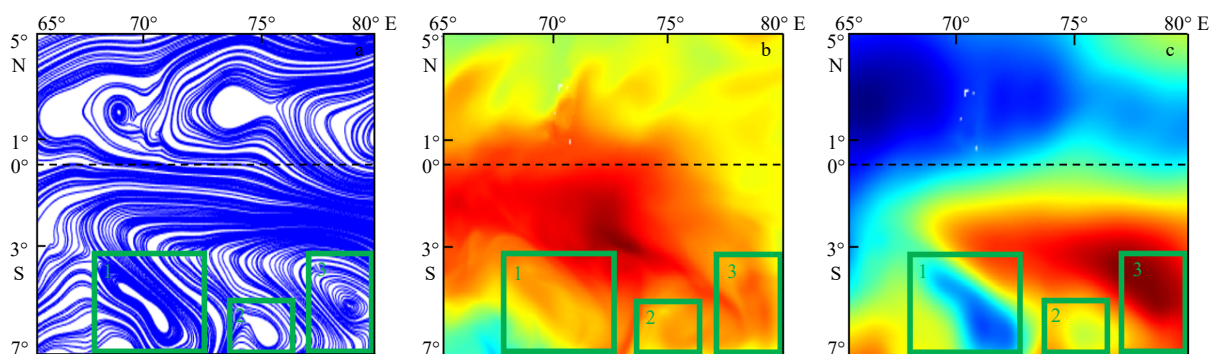


图 9 流场 B 流线、温度和海表面高度

Fig. 9 Streamline, temperature and sea surface height of flow field B

a. 流线图; b. 温度图; c. 海表面高度图

a. Streamline map; b. temperature map; c. sea surface height map

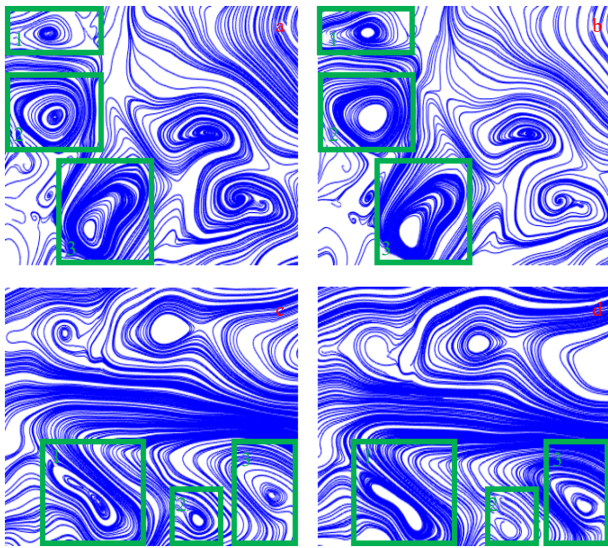


图 10 流线图相似性对比

Fig. 10 Similarity comparison of streamline map

- a. 本文实验结果(流场 A); b. 卫星数据流线图(流场 A); c. 本文实验结果(流场 B); d. 卫星数据流线图(流场 B)
- a. Experimental results in this paper (flow field A); b. streamline map of satellite data (flow field A); c. experimental results in this paper (flow field B); d. streamline map of satellite data (flow field B)

行时间。通过比较 3 种固定步长算法的运行时间可以看出运行时间与步长大小成反比,这是因为步长越小,积分点的数目越多,计算结果的数据量越大,所以运行时间更长。由于本算法在实验中设置的最大步长为 0.3,因此运行时间会略高于步长为 0.3 的固定步长算法。从后 3 列的统计结果来看本文算法在时间成本上要低于以往两种自适应步长算法。

具体结果运行时间是由流场中特征的数目和尺度决定的。

表 5 流场 A 不同算法结果运行时间
Table 5 Running time of different algorithm results of flow field A

图像	图6a	图6b	图6c	图6d	图6e	图6f
运行时间/s	1.733	1.625	1.528	1.714	1.729	1.642

表 6 流场 B 不同算法结果运行时间
Table 6 Running time of different algorithm results of flow field B

图像	图8a	图8b	图8c	图8d	图8e	图8f
运行时间/s	0.865	0.812	0.766	0.839	0.858	0.786

实验过程中通过选择部分积分点并记录它们的步长即可得到如图 11 所示的本文算法的步长空间分布图。图 11a 是流场 A 的部分区域(特征 1 和特征 2 处),从上往下各红色积分点的积分步长依次为 0.045、0.218、0.278、0.283、0.164、0.056;图 11b 是流场 B 的部分区域(特征 2 和特征 3 处),从左往右各红色积分点的积分步长依次为 0.061、0.177、0.282、0.274、0.185、0.047,结果均保留 3 位有效数字。通过该图可以直观地看出不同积分点的步长大小是不同的,变化剧烈的区域处积分点步长较小,变化平缓的区域处积分点步长较大。

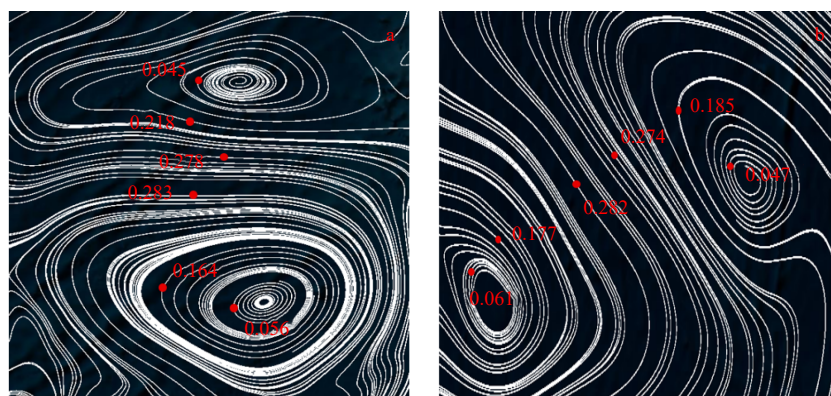


图 11 步长空间分布

Fig. 11 Step size spatial distribution

- a. 流场 A 部分区域步长空间分布; b. 流场 B 部分区域步长空间分布
- a. Step size spatial distribution of flow field A; b. step size spatial distribution of flow field B

图 12 是本文算法与以往两种自适应步长算法在流态更复杂区域处的放置效果对比图。图 12a 是本文算法绘制的流场,图 12b 是考虑速度方向变化的自

适应步长算法绘制的流场,图 12c 是改进后的 AM-FCA 算法绘制的流场,其中绿色方框处代表其他算法没有捕捉到的特征。由此可见,本文算法相比其他两

种自适应步长算法在复杂流场中具有更强的特征提 取能力。

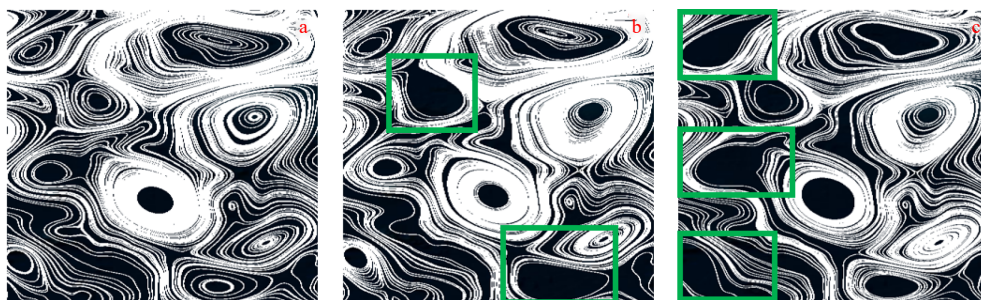


图 12 3种自适应步长算法对比结果

Fig. 12 Comparison of three adaptive step size algorithms

a. 本文算法; b. 考虑速度方向变化的自适应步长算法; c. 改进后的 AMFCA 算法

a. This paper's algorithm; b. adaptive step size algorithm that takes into account changes in velocity direction; c. improved AMFCA algorithm

5 结论

本文针对固定步长算法和以往自适应步长算法存在的自由度低以及多尺度适应性差的问题,提出了一种新的自适应步长流线放置算法。该算法以流场信息熵理论为基础,首先计算二维流场数据的信息熵得到二维熵场,然后依据熵值大小为流线上的每个积分点都赋予了新的步长。最后通过对比不同算法与本文算法的放置效果可知本文算法在保证计算效率的同时能够更好地绘制流场中变化剧烈区域的特征,显著增加了高熵区域(变化剧烈的区域)的积分点数目,有效解决了以往算法出现的特征缺失

和绘制不完全的问题。对于低熵区域(变化平缓的区域),减少积分点数目后并不会影响该区域的绘制效果。综上,本文算法能够实现不同尺度下流线的合理放置。

目前,该算法仅应用于结构化网格数据以及二维流场。对于没有规则拓扑结构的非结构化网格数据,在进行熵场计算时因不具有正方形邻域,所以需要选择合适的邻域专用于此类数据熵场的计算。对于三维流场,杂乱与遮挡是目前最关键的问题,其信息熵的计算相较于二维流场也更加复杂,后期可以通过将该算法引入三维流场解决上述问题,增加算法的适用性。

参考文献:

- [1] 王盛波, 潘志庚. 二维流场可视化方法对比分析及综述 [J]. 系统仿真学报, 2014, 26(9): 1875-1881, 1888.
Wang Shengbo, Pan Zhigeng. Comparison, analysis and review of 2D flow visualization[J]. Journal of System Simulation, 2014, 26(9): 1875-1881, 1888.
- [2] 邵绪强, 程雅, 金佚钟. 表意性方法在三维流线可视化中的应用综述 [J]. 图学学报, 2022, 43(5): 753-764.
Shao Xuqiang, Cheng Ya, Jin Yizhong. A review of the application of illustrative methods in 3D streamline visualization[J]. Journal of Graphics, 2022, 43(5): 753-764.
- [3] 黄智灏, 傅广涛, 曹凌婧, 等. 基于多视图聚类算法的三维流场关键点附近的流线筛选 [J]. 计算机辅助设计与图形学学报, 2022, 34(12): 1930-1942.
Huang Zhibin, Fu Guangtao, Cao Lingjing, et al. Streamline selection around critical points of 3D flow fields by the multi-view clustering[J]. Journal of Computer-Aided Design & Computer Graphics, 2022, 34(12): 1930-1942.
- [4] Liu Fan, Zhou Wensheng, LIU Bingxuan, et al. Flow field description and simplification based on principal component analysis down-scaling and clustering algorithms[J]. *Frontiers in Earth Science*, 2022, 9: 804617.
- [5] 杨光, 成诗明. 基于四面体的三维流线构造 [J]. 北京航空航天大学学报, 2008, 34(9): 1061-1064.
Yang Guang, Cheng Shiming. Tetrahedron-based constructing 3D_streamlines in visualization[J]. Journal of Beijing University of Aeronautics and Astronautics, 2008, 34(9): 1061-1064.
- [6] 鲁大营, 朱登明, 王兆其. 三维流场的流线提取算法 [J]. 计算机辅助设计与图形学学报, 2013, 25(5): 666-673.
Lu Daying, Zhu Dengming, Wang Zhaoqi. Streamline selection algorithm for three-dimensional flow fields[J]. Journal of Computer-Aided Design & Computer Graphics, 2013, 25(5): 666-673.
- [7] 季民, 陈丽, 靳奉祥, 等. 自适应步长的海洋流线构造算法 [J]. 武汉大学学报·信息科学版, 2014, 39(9): 1052-1056.
Ji Min, Chen Li, Jin Fengxiang, et al. Adaptive-step based marine fluid flow streamline constructing algorithm[J]. Geomatics and Information Science of Wuhan University, 2014, 39(9): 1052-1056.
- [8] 李婷, 季民, 靳奉祥, 等. 海洋流线积分自适应步长计算模型研究 [J]. 海洋学报, 2018, 40(3): 95-101.

- Li Ting, Ji Min, Jin Fengxiang, et al. Research on adaptive-step calculation model of marine fluid flow numerical integration[J]. Haiyang Xuebao, 2018, 40(3): 95–101.
- [9] Xu Lijie, Lee T Y, Shen Hanwei. An information-theoretic framework for flow visualization[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2010, 16(6): 1216–1224.
- [10] 黄冬梅, 杜艳玲, 张律文. 基于信息熵种子点选取的流线可视化 [J]. 计算机工程与科学, 2018, 40(3): 411–417.
Huang Dongmei, Du Yanling, Zhang Lvwen. Two information entropy-based seeding methods for 3D flow visualization[J]. *Computer Engineering and Science*, 2018, 40(3): 411–417.
- [11] 刘晓帆. 基于信息熵与 Clifford 代数的流场特征检测 [D]. 北京: 北京理工大学, 2016.
Liu Xiaofan. Flow feature detection based on entropy and Clifford algebra[D]. Beijing: Beijing Institute of Technology, 2016.
- [12] 李梦依, 方霞, 郑红波, 等. 基于信息熵的流场定向线积分卷积算法 [J]. 计算机应用, 2023, 43(4): 1233–1239.
Li Mengyi, Fang Xia, Zheng Hongbo, et al. Oriented line integral convolution algorithm for flow field based on information entropy[J]. *Journal of Computer Applications*, 2023, 43(4): 1233–1239.
- [13] 张倩, 肖丽. 基于流线的流场可视化绘制方法综述 [J]. *计算机科学*, 2021, 48(12): 1–7.
Zhang Qian, Xiao Li. Review of visualization drawing methods of flow field based on streamlines[J]. *Computer Science*, 2021, 48(12): 1–7.
- [14] Jobard B, Lefer W. Creating evenly-spaced streamlines of arbitrary density[C]//Proceedings of the Visualization in Scientific Computing '97. Vienna: Springer, 1997: 43–55.
- [15] Mebarki A, Alliez P, Devillers O. Farthest point seeding for efficient placement of streamlines[C]//VIS 05. IEEE Visualization, 2005. Minneapolis: IEEE, 2005: 479–486.
- [16] Verma V, Kao D, Pang A. A flow-guided streamline seeding strategy[C]//Proceedings Visualization 2000. VIS 2000 (Cat. No. 00CH37145). Salt Lake City: IEEE, 2000: 163–170.
- [17] 李忠伟, 徐斌, 李永, 等. 基于非结构化三角网格的海洋流场可视化 [J]. 图学学报, 2022, 43(3): 486–495.
Li Zhongwei, Xu Bin, Li Yong, et al. Visualization of ocean flow field based on unstructured triangular mesh[J]. *Journal of Graphics*, 2022, 43(3): 486–495.
- [18] 牛婵. 基于信息熵的流场特征提取与可视化研究 [D]. 秦皇岛: 燕山大学, 2018.
Niu Chan. Research on feature extraction of flow field based on information entropy and visualization[D]. Qinhuangdao: Yanshan University, 2018.
- [19] Du Xiaofu, Liu Huilin, Tseng H W. Adaptive method to locate seed points based on information entropy and quadtree[J]. *Sensors and Materials*, 2021, 33(2): 789–804.

Research on adaptive step size algorithm of marine streamline controlled by information entropy

Li Zhongwei¹, Jiao Fangtao¹, Li Yong¹, Zeng Wei¹, Yang Jungang², Cui Wei²

(1. College of Oceanography and Space Informatics, China University of Petroleum, Qingdao 266580, China; 2. First Institute of Oceanography, Ministry of Natural Resources, Qingdao 266061, China)

Abstract: The streamline construction and placement of the marine flow field is of great significance for recognizing and understanding the marine flow field. In the process of streamline drawing, the selection of integration step is very important, which can directly affect the effect of streamline placement. The fixed step size algorithm is often not used because it cannot adapt to the changing curvature. The previous adaptive step size streamline algorithm has the problems of low degree of freedom and poor multi-scale applicability. In view of the above problems, this paper introduces information entropy into the step size calculation for the first time, and proposes an adaptive step size algorithm of marine streamline controlled by information entropy. Firstly, the entropy field is obtained by calculating the information entropy of the flow field, and then the flow field is divided into high entropy region and low entropy region according to the entropy value, and each integration point is given a new step size, so that the flow field can adaptively adjust the step size according to the intensity of change, that is, the step size of the high entropy region (the region with sharp change) is smaller, and the step size of the low entropy region (the re-

gion with gentle change) is larger. The experimental results show that the proposed algorithm can significantly increase the number of integration points and streamlines in the rapidly changing region, better draw the details of the streamline at the feature, and reduce the number of integration points and streamlines in the unimportant region without affecting the placement effect to improve the computational efficiency. Compared with the previous adaptive step size algorithm, the proposed algorithm significantly improves the degree of freedom of step size adjustment and the scale applicability, and can be applied to different scales of marine flow field.

Key words: ocean flow field; information entropy; adaptive step size; streamline placement; numerical integration