

复杂流场特征区域模糊描述与提取方法*

徐华勋⁺, 李思昆, 马千里, 蔡 勋

(国防科学技术大学 计算机学院, 湖南 长沙 410073)

Fuzzy Description and Extracting Methods of Complex Feature Regions in Flow Fields

XU Hua-Xun⁺, LI Si-Kun, MA Qian-Li, CAI Xun

(College of Computer, National University of Defense Technology, Changsha 410073, China)

+ Corresponding author: E-mail: xxhxx@163.net

Xu HX, Li SK, Ma QL, Cai X. Fuzzy description and extracting methods of complex feature regions in flow fields. Journal of Software, 2011, 22(8): 1960-1972. <http://www.jos.org.cn/1000-9825/3851.htm>

Abstract: This paper presents a method to describe and extract the general feature structures based on the fuzzy theory. This paper first builds some fuzzy measure rules with three levels (the basic property Φ , the derived property γ , and the associated property Π). With the aid of measure rules, the feature vectors are established for fuzzy membership calculation and the feature region extraction. The analysis demonstrates that this method can obtain the optimal division of the flows on the principle of the minimum square sum. Further, the experiments show that the extent of typical flow structures extracted by this method are more effective than existing methods. In addition, the transfer function can be more flexible in design to avoid the cluttering and occlusion problems, which must be solved when visualizing 3D flows.

Key words: flow fields; feature visualization; feature structure; fuzzy theory

摘 要: 引入模糊理论对流场特征的描述,提出了一种流场特征区域描述与提取算法.通过采用模糊理论对流场特征区域进行描述,从基本属性 Φ 、衍生属性 γ 、关联属性 Π 这3个层次建立了测度规则,然后对各规则分析建立了相应的特征向量,并基于特征向量偏差给出了模糊测度隶属度的确定方法.理论分析表明,该描述与提取算法在最小平方和准则下是对流场区域的最优模糊划分,并具有良好的划分性质.实验结果表明,与传统的提取方法相比,该方法能够更准确地提取各种流场的典型特征区域,既可以有效地支持流场特征强度变化的定量分析,又易于设计转换函数,从而有效地解决3D流场可视化存在的遮挡和混淆问题.

关键词: 流场;特征可视化;特征结构;模糊理论

中图法分类号: TP391 **文献标识码:** A

流场具有复杂的特征结构.特征可视化方法对于流场性质分析具有重要意义.随着计算机性能的不不断提升,流场计算模型精度越来越高,数值计算结果规模越来越大.庞大的数据规模、复杂的特征结构,使得流场的高质量、实时交互式可视化面临严峻挑战.特征可视化方法通过对流场中复杂特征结构的描述,可以采用较少的几

* 基金项目: 国家自然科学基金(60873120, 60773020, 60970094); 国家重点基础研究发展计划(973)(G2009CB723803); 国家高技术研究发展计划(863)(2009AA01Z301)

收稿时间: 2009-10-20; 定稿时间: 2010-03-11

何绘制原语表示复杂的流场信息,在流场可视化方法中越来越得到研究人员的重视.

特征可视化方法大致可以分为两类:基于临界点理论的拓扑分析方法和特定特征结构提取方法.拓扑分析方法通过对不同类型临界点附近流线渐近行为的分析,得到流场的拓扑图.Helman 等人提出了矢量场拓扑结构分析方法^[1],相关研究很多.Scheuermann 对 2D 平面矢量场中的高阶临界点和一阶连续性进行了研究^[2,3].Scheuermann 等人引入了有界区域流场这一概念,研究了有界区域 2D 矢量场的拓扑特征,得到了边界鞍型点、边界源点以及边界汇点等类型^[4].Tricoche 则研究了基于拓扑分析流场的复杂流场拓扑简化和时变流场的拓扑跟踪方法^[5,6].Weinkauf 等人基于矢量场拓扑构架,采用符号标记法对 3D 流场的拓扑结构可视化方法进行了研究,对于研究 3D 流场拓扑性质起到了推动作用^[7,8].拓扑分析方法主要侧重于研究流场的流线渐近性等定性性质,但对流场特征的定量属性如强度、区域范围等缺乏有效的定义与描述,而这些属性与实际应用密切相关,如台风强度、范围与其危害性紧密联系.

特定特征的区域提取方法则主要针对流场的某种特定特征进行提取,目前主要集中于涡旋结构的提取^[9].在涡旋特征结构提取方法中,可提取涡旋特征区域的方法主要包括旋转参数(swirl parameter)方法^[10]、 λ_2 方法^[11]和流线方法^[12]等.由于临界点本身的特殊性,其与流场特征结构有着紧密联系,比如,Globus 将涡旋核结构看作由螺旋临界点出发的,沿第 3 个实特征向量的两条流线构成^[13].同时,临界点本身就可被看作流场区域中的一种特征.但目前,特征区域提取并没有考虑临界点理论,特征区域提取方法均从某特定特征性质出发,对于临界点,如源点、汇点和鞍型点等重要特征区域无法进行有效提取.流场中包含非常复杂的特征结构,尤其是 3D 流场;同时,流场所具有的矢量性也增加了特征结构统一描述的困难.模糊理论对于复杂现象和复杂结构具有较强的描述能力,但在科学计算可视化领域有关应用研究相对较少.

本文采用模糊理论对流场典型特征区域进行了统一描述,提出了一种基于模糊理论的流场特征区域提取方法,实现了流场特征的自动、准确提取.首先,采用模糊理论对流场特征区域进行了重新定义,并将流场性质描述变量分解为 3 个层次:基本属性 Φ (包括位置 X 和速度 V)、衍生属性 \mathcal{N} (从 Φ 进行推导得到的性质如梯度、涡度等)和关联属性 \mathcal{I} (流场内的其他物理属性变量);然后,从这 3 个层次建立了测度规则和相应的特征向量,并基于特征向量偏差给出了模糊测度隶属度的确定方法;最后,给出了形式化的流场特征模糊提取算法 FEFR(fuzzy extraction of feature regions).理论分析表明,FEFR 算法在最小平方和准则下是对流场区域的最优模糊划分,并具有良好的划分性质.实验结果表明,与传统的特征提取涡旋特征方法相比,FEFR 算法提取的涡旋特征区域更为准确.同时,对于传统方法不能提取的其他典型特征区域,如临界点特征,FEFR 算法也能进行有效提取.基于 FEFR 算法提取的流场特征区域,可以容易地设计合适的转换函数对特征区域和背景区域进行区分绘制,从而有效地解决传统流场可视化存在的遮挡和混淆问题.

1 流场特征拓扑分析

1.1 临界点分析

流场拓扑分析的基础是微分方程的一些基本概念和定理.流场可定义为向量场中所有积分曲线即流线的并集,临界点是向量场中流线唯一可能相交的点,在流场拓扑分析中占有重要地位.下面引入其定义^[14].

定义 1(临界点(critical point)). 向量场 V 的一个临界点 x_0 是向量的各个分量均为 0 的点: $\vec{V}(x_0) = 0$, 临界点又称为奇点(singular point)、不动点(fixed point)或平衡点(equilibrium point).

根据临界点附近的流线渐近线,可将临界点周围流场分割为具有不同流场流线性质的扇区.这些扇区包括 3 类:抛物线型、双曲线型和椭圆型.在抛物线型扇区中,所有的流线或者起自该临界点,或者终止于该临界点.在双曲线型扇区中,除了分割线以外,所有流线都不经过该临界点;而在椭圆型扇区中,所有的流线均起始于该临界点并终止于该临界点.可根据临界点附近速度场的特性对临界点进行分类.也可依据临界点处的偏导数矩阵 $\nabla V = \partial V_i / \partial X_j$ 决定临界点的类型.偏导数矩阵特征值实部为正,表示矢量 V 从临界点发散;为负则表示矢量 V 向临界点靠拢.共轭复数表示 V 是螺旋入或出.根据 V 的特征值,2D 临界点可分为源点、汇点(包括排斥聚点和吸引

聚点)、马鞍点以及中心点这 4 类.

对于 3D 流场临界点 x_0 , 周围的扇区可以根据 x_0 附近的一个曲面 s 上的入流与出流行为分为如下几类: 入流曲面 C_{in} 、出流曲面 C_{out} 以及双曲型扇区和椭圆型扇区, 如图 1 所示. 入流曲面 C_{in} 和出流曲面 C_{out} 不能相交(属于 2D 情况下的抛物型扇区类), 中间为双曲型扇区和椭圆型扇区^[7].

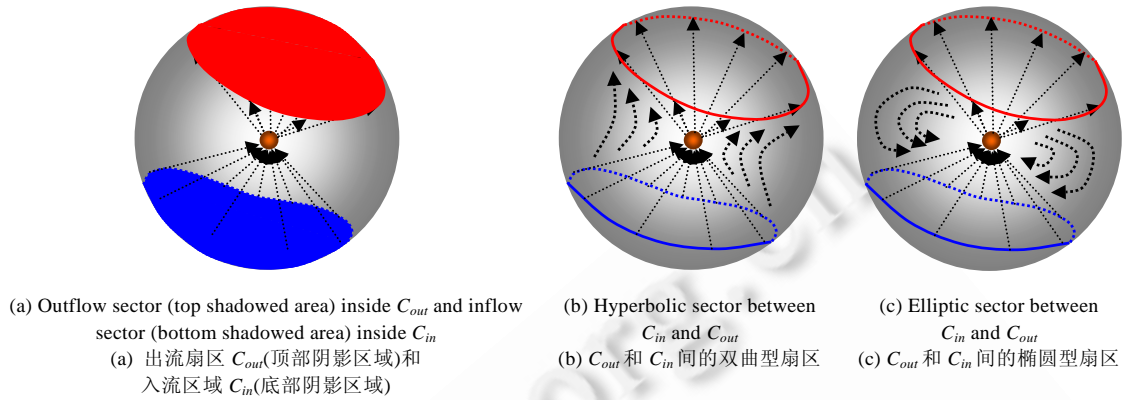


Fig.1 Sector classification of a 3D critical point

图 1 临界点附近的扇区分类

基于上述扇区分类, 可以区分源点、汇点和鞍型点. 源点和汇点有一个抛物型扇区, 覆盖了曲面 s , 分别如图 2(a)和图 2(b)所示, 源点附近曲面 s 上的流线均起始于 s , 而汇点附近曲面 s 上的流线则终止于 s . 鞍型点则由一个出流面和两个退化的入流面或一个入流面和两个退化的出流面组成, 分别如图 2(c)和图 2(d)所示. 图 2(c)中, 上下两端斜网格区域表示出流区, 而中间斜网格表示入流区; 图 2(d)的情况则相反. 因此, 一个鞍型点由两个双曲型扇区组成.

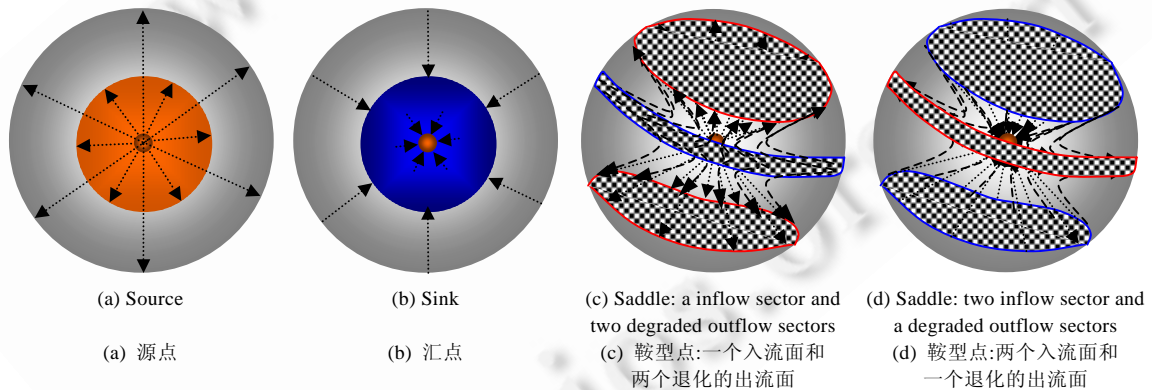


Fig.2 Classification of critical points

图 2 临界点分类

1.2 拓扑图分析

在传统的拓扑分析方法中, 临界点、分界线和闭合流线构成了描述流线渐近行为的拓扑图. 下面引入相关基本定义^[14]:

定义 2(轨线(orbit)). 过向量场 V 中一点 x 的轨线可定义为如下映射:

$$\alpha_x: R \supset I \rightarrow R^2,$$

其中, $\alpha_x(0)=x_0$, $\dot{\alpha}_x(t)=v(\alpha(t))$, $\forall t \in I$. 对应于定常流场和非定常流场, 轨线又可分别称作流线(stream line)和迹线(path line). 由于本文主要研究定常流场条件下的特征影响范围, 考虑上下一致性, 不对轨线和流线进行区分, 统一采用流线描述.

定义 3(分界线(separatrix)). 由马鞍点或边界马鞍点出发或终止的流线称为分界线.

流场是有方向的矢量场, 不同区域具有不同的流向属性, 表现为区域中每个点有唯一一条流线穿过, 不同流线有不同的渐近行为. 分界线的本质在于, 它将向量场分解为有相同流线渐近行为的若干子区域, 每个子区域内的流线以同一临界点或极限环为其 α -极限集或 ω -极限集.

定义 4(闭合流线(closed orbit)). 设 s 是向量场 V 中的一条积分曲线, 若存在一个周期 $t_0 \in \mathbb{R}, t_0 \neq 0$, 使得 $s(t+nt_0)=s(t)$ 对所有的 t 以及整数 n 成立, 则称 s 为闭合流线.

根据邻近流线渐近行为的不同, 闭合流线又可分为极限环和同心闭轨. 极限环根据其两侧流场的流线行为又可以分为双曲型和非双曲型两种. 在本文中, 主要侧重于临界点分析, 暂不考虑封闭流线的影

对于图 3 中所示流场, 其拓扑图如图 4 中的灰白色斜十字格纹理区域边界曲线所示, 临界点与分割线将平面流场分割为几个具有不同流动行为的区域: 左侧的灰黑色反斜线纹理区域为出流区; 右上侧的灰黑色斜线纹理区域为入流区; 中间的灰白色斜十字格纹理区域流场起自源点, 终于汇点; 右下角的灰黑色点状纹理区域自入流点流入, 从出流点流出.

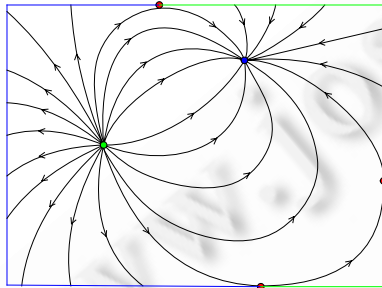


Fig.3 Classification of critical points
图 3 边界点分类

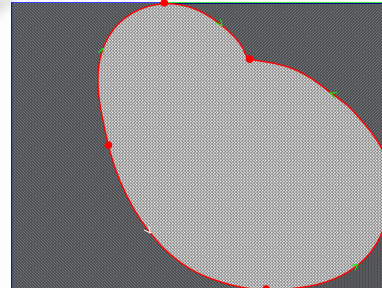


Fig.4 Topology graph of limitary flow fields
图 4 有界区域流场拓扑图

定义 5(拓扑图(topology graph)). 平面向量场中所有的临界点、分割线以及闭合轨线组成该平面向量场的拓扑图.

传统的拓扑分析方法侧重于流线渐近行为的研究, 拓扑图将流场分割为有不同流线渐近行为的区域. 但对于流场中的典型特征区域, 如涡旋、临界点等, 则缺乏有效的描述方法, 从而无法对流场特征结构、性质进行更好的研究. 同时, 特征结构区域的难以确定, 导致 3D 流场可视化中无法突出重点, 遮挡现象严重影响了可视化效果. 考虑到临界点在流场拓扑分析中的重要性, 本文基于临界点分类, 将流场特征分为源、汇和鞍型这 3 种类型. 下一节将在此基础上采用模糊理论对特征区域的界定进行研究.

2 特征区域模糊描述与测度

流场特征结构比较复杂, 特征区域精确界定困难, 缺乏有效的统一描述. 模糊理论对于复杂现象与结构具有较强的描述能力, 本节首先基于模糊理论对流场相关概念进行定义, 并建立特征区域的提取规则, 给出实现特征区域的模糊界定方法.

2.1 流场特征区域模糊描述

定义 6(流场). 流场为特定区域 X 上具有基本属性 Φ 、衍生属性 Υ 、关联属性 Π 的三元组 $U(\Phi, \Upsilon, \Pi)$. 基本属性 Φ 包括位置 X 和速度 V ; 衍生属性 $\Upsilon = \Upsilon(X, V)$, 是指从基本属性 X, V 进行推导得到的描述流场性质的属性, 如梯

度、涡度等;关联属性 Π 则是流场区域内的其他物理属性,如温度 T 、压强 P 、密度 ρ 等.

基本属性 Φ 描述了流场的基本性质,即空间性和方向性.衍生属性 $\gamma = \gamma(X, V)$ 与流场的典型特征性质描述密切相关,比如,涡旋与涡度密切相关,涡度越大的地方,涡旋特征越明显.关联属性表面看来似乎与流场性质描述无关,但实际上两者之间有着紧密的联系,如,气象风场中的风向实际由气压高低决定.基于上述定义,从模糊理论出发,可以给出流场特征的定义:

定义 7(流场特征). 对于流场 $U(\Phi, \gamma, \Pi)$, 具有性质 i 的流场特征是 U 上的一个模糊集. 它可以表示为一组有序数对:

$$F^i = \{x, \mu_F^i(U(x))\}, x \in X, \\ \mu_F^i(U(x)) = \mu_F^i(\Phi(x), \gamma(x), \Pi(x)) \in [0, 1].$$

上式中,隶属函数 μ_F^i 定义了从 U 到区间 $[0, 1]$ 的一个映射, $\mu_F^i(x)$ 称为 $x \in X$ 对 F^i 的隶属度.

基于上述定义,我们可以得到流场特征区域的概念:

定义 8(特征区域). 设 F^i 是第 i 类流场特征, 则 $\mu_F^i(x) > 0.5$ 的所有 $x \in X$ 组成的集合称为特征 F^i 的特征区域, 记为 D_{Fi} , 即 $D_{Fi} = \{x | \mu_F^i(U(x)) > 0.5\}$.

此外,可以方便地引入如下概念:

定义 9(特征核). 设 F^i 是具有特征属性 i 的流场特征, 则 $\mu_F^i(x) = 1$ 的所有 $x \in X$ 组成的集合称为特征 F^i 的核, 记为 D_{Ki} , 即 $D_{Ki} = \{x | \mu_F^i(U(x)) = 1\}$.

定义 10(边界区域). 设 F^i 是第 i 类流场特征, 称 $D_{FBi} = \{x | x \in X, 0 < \mu_F^i(U(x)) \leq 0.5\}$ 为特征 F^i 的边界区域, 记为 D_{FBi} ; 同时, 将 $B^i = \{x | x \in X, \mu_F^i(U(x)) = 0.5\}$ 构成的集合称为第 i 类流场特征的特征边界.

定义 11(背景区域). 对流场区域 X , 称 $D_B = \{x_k | x_k \in X, \forall i, 0 \leq i \leq c, \mu_{ik} < 0.5\}$ 为流场的背景区域, 记作 D_B . 其中, c 表示临界点个数, μ_{ik} 表示 x_k 对第 i 个临界点的隶属度.

2.2 特征区域模糊测度规则

在定义完流场特征和特征区域的概念后,需要在模糊理论意义上对流场中任意点进行特征性质测度,即确定 $\mu_F^i(U(x))$. 下面将从基本属性 Φ 、衍生属性 γ 和关联属性 Π 这3个层面探讨并建立特征区域测度规则.

2.2.1 基本属性规则

对于基本属性 Φ ,主要考虑空间变量 X 和速度变量 V .由第1节易知,临界点在流场分析中具有重要作用,由临界点定义 $\vec{v}(x_0) = 0$ 易知其仅与 X, V 相关,因此可建立如下基本属性规则(简称 Φ 规则):

Φ 规则:

$$\forall x \in C^i \mapsto \mu_F^i(U(x)) = 1 \quad (1)$$

其中, $C^i \subseteq X$ 表示具有特征属性 i 的临界点集合.

Φ 规则主要对临界点集合进行隶属度测度,将不同性质的临界点作为相应特征区域的特征核.由于临界点是典型流场特征的中心,故肯定属于相应特征区域.

2.2.2 衍生属性规则

对于衍生属性,首先引入如下变量定义:

定义 12(流线距离). 对于向量场中任意两点 A, B , 两者之间的流线距离定义如下:

$$L = \begin{cases} \int_A^B f(x) ds, & \text{若流线 } f(x) \text{ 过 } A, B \text{ 两点} \\ \infty, & \text{否则} \end{cases} \quad (2)$$

其中, $f(x)$ 表示流线曲线函数, s 为 $f(x)$ 的弧长参数. 流线距离描述了流场中任意两点沿流线方向的距离远近.

基于上述定义,可以建立如下衍生属性规则(简称 γ 规则):

γ 1 规则:

$$\forall x, y \in U, c \in C^i, L(x, c) < L(y, c) \mapsto \mu_F^i(U(x)) > \mu_F^i(U(y)) \quad (3)$$

在 γ_1 规则中, $L(A,B)$ 表示 A,B 两点间的流线距离, C^i 表示具有性质 i 的临界点集合.下文相同.

对于流场而言,属性传播具有方向性,比如,风场中热量将沿速度场方向传播.而流线能够反映流场的方向性. γ_1 规则反映出流场特征区域的这种方向性,也就是说,从流线方向距离具有特征属性 i 的临界点越近的位置,其特征属性 i 的隶属度越大.

对于源点和汇点特征而言, γ_1 规则可以有效地反映其特征区域范围,如图4所示.对于各向同性的流场,距离源点和汇点流线距离相等的点会构成一个等值面,从而有效反映了该类特征的区域范围.但对于鞍型点而言,其由两个双曲面组成,若采用流线距离则只能得到入流区和出流区的特征范围.对于椭圆型扇区和双曲型扇区部分,则没有特征所属区域.同时,流场中存在封闭流线等结构,流线距离计算可能无法测度该结构区域特征隶属度.因此,引入如下规则作为补充,对这些区域进行模糊测度:

γ_2 规则:

$$\forall x, y \in X, c \in C^i, D(x, c) < D(y, c) \mapsto \mu_F^i(U(x)) > \mu_F^i(U(y)) \quad (4)$$

在 γ_2 规则中, $D(A,B)$ 表示 A,B 两点之间的欧拉距离.需要注意的是, γ_2 规则仅适用于距离各临界点欧拉距离小于某特定阈值 d_i 的椭圆型扇区和双曲型扇区范围.

此外,针对流场中最重要的涡旋特征,引入传统的易于GPU加速的 λ_2 方法,建立如下规则:

γ_3 规则:

$$\forall x, y \in X, c \in C^i, \lambda_2(x) < \lambda_2(y) \mapsto \mu_F^i(U(x)) > \mu_F^i(U(y)) \quad (5)$$

2.2.3 关联属性规则

关联属性与应用数据场密切相关.本文实验的应用数据是气象风场数据,考虑涡旋与气压的关联性,建立如下规则:

Π_1 规则:

$$\forall x, y \in X, P(x) < P(y) \mapsto \mu_F^i(U(x)) > \mu_F^i(U(y)) \quad (6)$$

其中,性质 i 是指气旋涡旋性. Π_1 规则从气压值上对流场区域上的气旋涡旋性进行了测度,气压越低的地方隶属于气旋涡旋的隶属度越大.

2.3 特征区域隶属度的计算

上节中的规则只是对流场特征区域的定性描述,并不能确定 $\mu_F^i(U(x))$,本节将介绍 $\mu_F^i(U(x))$ 的计算方法.为方便叙述,首先引入如下概念:

定义 13(规则作用区域). 规则 R 的适用范围称为规则作用区域,记作 D_R ,其中, $R \in \{\Phi, \gamma_1, \gamma_2, \gamma_3, \Pi_1\}$.

定义 14(确定区域、模糊区域). 流场区域 X 内隶属度属于 $\{0,1\}$ 集合的范围称为确定区域,记作 $D_H = \{x | x \in X \wedge \mu(x) \in \{0,1\}\}$. $D_F = X - D_H$ 称为模糊区域.

隶属度的确定主要包括3步:(1) 确定区域隶属度的计算;(2) 测度规则特征向量的构造;(3) 模糊区域隶属度的计算.

(1) 确定区域隶属度的计算

根据 Φ 规则易知,对 $\forall x_k \in D_{\Phi} = C$,有 $\mu_{ik} = \mu_i(x_k) = \{1 | k=i, 0 | k \neq i\}$.对于 $\forall x_k \in X - \left(D_{\Phi} + \sum_{i=1}^2 D_{\gamma_i} \right)$,则将其归属到背景区域,即有 $\mu_{ik} = 0, 1 \leq i \leq c$.

(2) 对各规则给出特征向量

作如下假设:假设共有 c 个临界点,流场论域 X ,当前样本为 x_k .

① 对 γ_1 规则进行分析. γ_1 规则的本质是对流场中 x_k 点与各临界点的流线共性进行分析,因此可建立如下的 c 维特征向量:

$$\begin{cases} V_{r,1}(\mathbf{x}_k) = (r_{k_1}, r_{k_2}, \dots, r_{k_c}) \\ r_{k_j} (1 \leq j \leq c) = \begin{cases} 0, & \text{iff } L(\mathbf{x}_k, C_j) = \infty \\ 1, & \text{else} \end{cases} \end{cases} \quad (7)$$

其中, C_j 表示第 j 个临界点. 比如, 对 $\mathbf{x}_k \in X$, 若 \mathbf{x}_k 位于起始于第 j 个临界点终止于第 l 个临界点的流线上, 则根据公式(7)有 $V_{r,1}(\mathbf{x}_k) = (0, 0, \dots, r_{k_j} = 1, 0, \dots, r_{k_l} = 1, \dots, 0)$.

② 对 r_2 规则进行分析. r_2 规则的本质是对流场内位置 \mathbf{x}_k 与各临界点欧拉距离进行分析, 因此可采用位置坐标作为其特征向量, 即

$$V_{r,2}(\mathbf{x}_k) = (x_k, y_k, z_k) \quad (8)$$

③ 对 r_3 规则和 l_1 规则进行分析. 两者的本质分别是对流场内位置 \mathbf{x}_k 与各临界点的 λ_2 值和气压 P 进行分析, 因此可分别采用位置 \mathbf{x}_k 处的 λ_2 值和气压作为其特征向量, 即

$$\begin{cases} V_{r,3}(\mathbf{x}_k) = \lambda_2(\mathbf{x}_k) \\ V_{l,1}(\mathbf{x}_k) = p(\mathbf{x}_k) \end{cases} \quad (9)$$

因此, 根据公式(7)~公式(9)可最终建立 \mathbf{x}_k 的 $c+5$ 维特征测度向量(如式(10)所示), 并可根据原始流场数据计算特征向量数据.

$$V(\mathbf{x}_k) = (V_{r,1}(\mathbf{x}_k), V_{r,2}(\mathbf{x}_k), V_{r,3}(\mathbf{x}_k), V_{l,1}(\mathbf{x}_k)) \quad (10)$$

由于各特征区域的核心是临界点, 因此将临界点处的特征向量作为标准特征向量. 对于背景区域, 由于不存在临界点, 因此以背景区域内样本的均值作为其标准特征向量如下, 若记 $D_B^* = X - (D_\phi + \sum_{i=1}^2 D_{r,i})$, 则有:

$$V'(D_B) = \frac{1}{n} \sum_{\mathbf{x}_k \in D_B^*} (V_{r,1}(\mathbf{x}_k), V_{r,2}(\mathbf{x}_k), V_{r,3}(\mathbf{x}_k), V_{l,1}(\mathbf{x}_k)) \quad (11)$$

其中, n 为 D_B^* 内的样本个数.

(3) 模糊区域隶属度的计算

首先需要对原始特征测度数据进行标准化处理. 除了 $V_{r,1}(\mathbf{x}_k)$ 测度向量以外, $V_{r,2}(\mathbf{x}_k)$, $V_{r,3}(\mathbf{x}_k)$ 和 $V_{l,1}(\mathbf{x}_k)$ 的各维数据取值范围变化大, 因此首先分别对其进行 $[0, 1]$ 区间标准化处理. 以 $V_{r,2}(\mathbf{x}_k)$ 的 y_k 为例加以说明, 标准化处理如公式(12)所示, 并记 $V(\mathbf{x}_k)$ 处理后得到 $V'(\mathbf{x}_k)$.

$$y'_k = \frac{y_k - \text{Min}}{\text{Max} - \text{Min}} \quad (12)$$

对原始数据进行标准化处理后, 隶属度的计算区分如下两种情况: 对于 $\forall k$, 定义集合 $I_k = \{i | 1 \leq i \leq c, d_{ik} = 0\}$ 和 $\bar{I}_k = \{i | 1 \leq i \leq c\} - I_k$, 其中, $d_{ik} = \|V(\mathbf{x}_k) - V(c_i)\| = \left(\sum_{j=0}^{c+5} (V'_j(\mathbf{x}_k) - V'_j(c_i))^2 \right)^{1/2}$ 表示当前位置 \mathbf{x}_k 与标准样本 c_i 间的偏差. 可根据如下公式计算其隶属度:

$$\mu_{ik} = \begin{cases} 1 / \left(\sum_{j=1}^{c+1} d_{ik} / d_{jk} \right), & \text{iff } I_k = \emptyset \\ 1 - L(\mathbf{x}_k, c_i) / L_{len}(\mathbf{x}_k), & \text{iff } I_k \neq \emptyset \text{ and } i \in I_k \text{ and } i \leq c \\ 0, & \text{iff } I_k \neq \emptyset \text{ and } i \in \bar{I}_k \text{ and } i \leq c \\ 1 - \sum_{i=1}^c \mu_{ik}, & \text{iff } I_k \neq \emptyset \text{ and } i = c + 1 \end{cases} \quad (13)$$

其中, $L(\mathbf{x}_k, c_i)$ 表示 \mathbf{x}_k 与第 i 个临界点 c_i 两点间的流线距离, $L_{len}(\mathbf{x}_k)$ 为过 \mathbf{x}_k 的流线在流场区域 X 内的长度.

2.4 特征区域模糊提取算法

在给出隶属度确定方法后, 本节介绍具体的特征提取算法. 算法中的符号含义同上, $[\cdot]$ 内为可选项.

算法 1. 特征区域模糊提取算法 FEFR.

输入: $X, V, C, d_v, [P]$.

输出: $D_{F1}, D_{F2}, \dots, D_{Fc}$.

1. $I_k \leftarrow \emptyset, D_{Fi} \leftarrow \emptyset$.
2. **for** (C 内的每个体元 x_k)
3. $\{D_{Fi} \leftarrow x_k, \mu_{ik} = 1\}$
4. **for** (X 内的每个体元 x_k)
5. {对过 X 内每个体元 x_k 的流线进行跟踪, 得到 $L(C_i, x_k), 0 \leq i \leq c$.
6. 根据跟踪结果 $L(C_i, x_k)$ 建立 $V_{r1}(x_k)$.
7. 计算 x_k 与各临界点 C_i 的欧拉距离 $d_{ik} = \{d_{ik} | d_{ik} < d_v, 0 | d_{ik} \geq d_v\}$, 建立 $V_{r2}(x_k)$.
8. 计算 x_k 处的 λ_2 值, 建立 $V_{r3}(x_k)$.
9. [根据 x_k 处的 P 值, 建立 $V_{r1}(x_k)$.]
10. **for** (X 内的每个体元 x_k)
11. **for** ($i=1; i \leq c+1; i++$)
12. $\{d_{ik} \leftarrow \|V(x_k) - V(c_i)\|\}$.
13. **if** ($d_{ik} = 0$) $\{I_k \leftarrow I_k \cup i\}$
14. **if** ($I_k = \emptyset$)
15. **for** ($i=1; i \leq c+1; i++$)
16. $\{ \mu_{ik} = 1 / \left(\sum_{j=1}^{c+1} d_{ik} / d_{jk} \right) \}$
17. **else**
18. **for** ($i=1; i \leq c+1; i++$)
19. **if** ($i \in I_k$ and $i \leq c$)
20. $\{ \mu_{ik} \leftarrow 1 - L(x_k, c_i) / L_{len}(x_k) \}$
21. **else if** ($i \in \bar{I}_k$ and $i \leq c$)
22. $\{ \mu_{ik} = 0 \}$
23. **else**
24. $\{ \mu_{ik} \leftarrow 1 - \sum_{i=1}^c \mu_{ik} \}$
25. **for** ($i=1; i \leq c; i++$)
26. **if** ($\mu_{ik} > 0.5$) $\{D_{Fi} \leftarrow x_k\}$

算法中, 第 4 行~第 9 行与下面的循环分开, 主要是考虑这部分可以作为预处理过程, 以提高算法的性能.

3 特征区域模糊描述性质

本节对本文方法的性质进行分析, 首先引入如下定义:

定义 15(模糊描述图(fuzzy description graph)). 若流场区域 X 采用 FEFR 算法后得到各特征区域集合为 $D_F = \{D_{F1}, D_{F2}, \dots, D_{Fc}\}$, 则特征区域的边界构成该平面流场 X 的模糊描述描述图, 简称 FDG.

传统的拓扑图 TG 可将流场特征区域划分成多个不相交的子集 X_1, X_2, \dots, X_n , 且满足 $X_1 \cup X_2 \cup \dots \cup X_n = X$, $X_i \cap X_j = \emptyset, 1 \leq i \neq j \leq n$, 此类划分属于硬划分. 若将样本 $x_k (1 \leq k \leq n)$ 对子集 X_i 的隶属关系表示为 μ_{ik} , 则硬划分中 $\mu_{ik} \in \{0, 1\}$. 若将 μ_{ik} 从 $\{0, 1\}$ 二值扩展到 $[0, 1]$ 区间, 则可到样本集的模糊划分:

定义 16(模糊划分(fuzzy partition))^[15]. 样本集 X 被划分为 c 个模糊子集 $\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_n$, 样本 $x_k (1 \leq k \leq n)$ 对子集 \tilde{X}_i 的隶属关系表示为 μ_{ik} , 若划分后 μ_{ik} 满足如下 3 个条件:

$$E_h = \left\{ \textcircled{1} : \mu_{ik} \in [0,1]; \textcircled{2} : \sum_{i=1}^c \mu_{ik} = 1, \forall k; \textcircled{3} : 0 < \sum_{k=1}^n \mu_{ik} < n, \forall i \right\},$$

则该划分称为 X 的一个模糊划分.

定理 1. FDG 将流场区域 X 分割成的集合 $D_F = \{D_{F1}, D_{F2}, \dots, D_{Fc}, D_B\}$ 构成 X 上一个模糊划分.

证明:首先证明定义 16 中条件①成立.对于确定区域,根据其隶属度计算方法易知 $\mu_{ik} \in \{0,1\}$.而对于模糊区域,根据公式(13)易知, $\mu_{ik} \in [0,1]$.因此,对 $\forall \mathbf{x}_k \in X$ 有 $\mu_{ik} \in [0,1]$,即条件(1)成立.

其次证明条件②成立.若 $I_k = \emptyset$,则根据公式(13)不难得到 $\sum_{i=1}^{c+1} \mu_{ik} = \sum_{i=1}^{c+1} \left(1 / \left(\sum_{j=1}^{c+1} d_{ik} / d_{jk} \right) \right) = 1$; 若 $I_k \neq \emptyset$,则根据公式(13)有 $\sum_{i=1}^{c+1} \mu_{ik} = \sum_{i \in I_k} (1 - L(\mathbf{x}_k, c_i) / L_{len}(\mathbf{x}_k)) + \left(1 - \sum_{i=1}^c \mu_{ik} \right) = 1$.故对 $\forall \mathbf{x}_k \in X$,有 $\sum_{i=1}^{c+1} \mu_{ik} = 1$,即条件(2)成立.

最后,根据条件①和条件②,条件③显然成立.故定理 2 成立.证毕. □

通常情况下,模糊划分可得到多种不同的划分结果.比如,对第 2.3 节中的隶属度计算也可以采用其他方法进行定义,则可得另外特征区域集合.因此,需要确立合理的划分准则来衡量模糊划分结果的优劣.本文采用常见的最小平方误差和准则,若用 $U = [\mu_{ik}]_{c \times n}$ 表示模糊划分矩阵, P 表示对 U 的模糊划分,则该方法定义的目标函数如下^[15]:

$$J_2(U, P) = \sum_{i=1}^c \sigma_i^2 = \sum_{i=1}^c \left(\sum_{k=1}^n (\mu_{ik})^2 (d_{ik})^2 \right) \tag{14}$$

其中, σ_i^2 表示模糊划分中第 i 类的模糊偏差,它反映了该类中的样本分布情况.最小平方和准则为,当 $J_2(U, P)$ 取得最小值时,该模糊划分最合理.

定理 2. FDG 对流场区域 X 的模糊划分是最小平方和准则下 X 的最优模糊划分.

证明:由于 $U = [\mu_{ik}]_{c \times n}$ 中各列独立,根据公式(14),问题可描述为

$$\begin{cases} \min\{J_2(U, P)\} = \min \left\{ \sum_{i=1}^c \left(\sum_{k=1}^n (\mu_{ik})^2 (d_{ik})^2 \right) \right\} = \sum_{k=1}^n \min \left\{ \sum_{i=1}^c (\mu_{ik})^2 (d_{ik})^2 \right\} \\ \sum_{i=1}^{c+1} \mu_{ik} = 1 \end{cases} \tag{15}$$

上述问题可采用拉格朗日乘数法进行求解,根据公式(15),可令

$$F = \sum_{i=1}^c (\mu_{ik})^2 (d_{ik})^2 - \lambda \left(\sum_{i=1}^c \mu_{ik} - 1 \right),$$

然后,根据最优化的一阶必要条件,类似于文献[15]不难得到使得 $J_2(U, P)$ 为最小的 μ_{ik} 值为

$$\begin{cases} \mu_{ik} = 1 / \sum_{j=1}^{c+1} (d_{ik} / d_{jk}), & \text{iff } I_k = \emptyset \\ \mu_{ik} = 0, & \text{iff } I_k \neq \emptyset \text{ and } i \in \bar{I}_k \\ \sum_{i \in I_k} \mu_{ik} = 1, & \text{otherwise} \end{cases} \tag{16}$$

显然,公式(13)满足公式(16)中的各约束条件,故定理 3 成立,证明完毕. □

定理 3. 若 FDG 将流场区域 X 分割成集合 $D_F = \{D_{F1}, D_{F2}, \dots, D_{Fc}, D_B\}$, 则 $\sum_{i=1}^{c+1} (D_{Fi} + D_{FBi}) = X$, 且 $D_{Fi} \cap D_{Fj} = \emptyset$,

$\forall i \neq j$.

证明:根据定义,显然 $\sum_{i=1}^{c+1} (D_{Fi} + D_{FBi}) \subseteq X$. 而由定理 1 可知, D_F 是 X 上的模糊划分,则 $\forall \mathbf{x}_k \in X$ 有 $\sum_{i=1}^{c+1} \mu_{ik} = 1$, 故必 $\exists i, 1 \leq i \leq c+1$, 使得 $\mu_{ik} > 0$, 即 $\mathbf{x}_k \in (D_{Fi} + D_{FBi})$. 故 $\sum_{i=1}^{c+1} (D_{Fi} + D_{FBi}) = X$ 成立.

然后,假设 $\exists i \neq j, D_{Fi} \cap D_{Fj} \neq \emptyset$,不妨设 $x_k \in (D_{Fi} \cap D_{Fj})$,则根据特征区域定义有 $\mu_{ik} > 0.5$ 和 $\mu_{jk} > 0.5$.

从而 $\sum_{i=1}^{c+1} \mu_{ik} \geq (\mu_{ik} + \mu_{jk}) > 1$,与 $\sum_{i=1}^{c+1} \mu_{ik} = 1$ 矛盾.因此, $D_{Fi} \cap D_{Fj} = \emptyset, \forall i \neq j$ 成立. \square

定理3表明了算法1对流场区域划分的完整性和无二义性.完整性保证了流场区域内任意一点或者属于某特征区域或边界区域,或者属于背景区域;无二义性保证了流场区域内任意一点不会同时属于两个不同的特征区域.

4 实验结果与分析

为了验证本文方法的有效性,本文对多个数据集进行了验证,包括洛伦茨吸引子数据集 $v_2=(10(y-x), 28x-y-xz, xy-(8/3)z)$ 、海棠台风数据以及一个飞行器实验数据.

洛伦茨吸引子数据集是目前国际上公认的具有代表意义的复杂流场模型^[16].模型的所有非平衡解最终都趋向于洛伦茨吸引子,即环绕模型中两个具有负特征值和正实部的复共轭特征值的临界点形成的两瓣缠绕区域.缠绕区域具有旋转性质,本文采用传统的涡旋特征区域提取 λ_2 方法与本文方法进行了比较.图5(a)~图5(c)中采用 λ_2 方法对洛伦茨吸引子数据集中的两瓣缠绕特征区域进行了提取,并对 λ_2 进行处理得到标准化值 $N(\lambda_2)$.图5(a)为 $N(\lambda_2) > 0$ 的特征区域,图5(b)和图5(c)分别为 $N(\lambda_2) > 0.6$ 和 $N(\lambda_2) > 0.9$ 的提取区域.由于吸引子的两瓣缠绕区域与具有旋转性质的区域并不完全等价,因此, λ_2 方法对吸引子的提取结果中包含了缠绕特征之外的旋转区域,引起遮挡与混淆,可视化效果不理想.图5(d)~图5(f)中为基于本文方法提取的两瓣缠绕特征区域,并采用模糊理论中的语气算子^[15] $\mu_{H_2A}(x) = [\mu_A(x)]^\alpha$ 进行了处理.图5(d)为未进行语气算子运算即 $\alpha=1$ 时提取到的特征区域,图5(e)为语气算子参数 $\alpha=0.5$ 时对提取的特征区域进行散漫化处理得到的特征区域,图5(f)为语气算子参数 $\alpha=2.0$ 时对提取的特征区域进行集中化处理得到的特征区域.从图中可以看出,相对于 λ_2 方法,本文算法提取的洛伦茨吸引子特征区域更准确,能够很好地剔除数据集中不关注的背景区域,突出吸引子的两瓣缠绕特征区域.

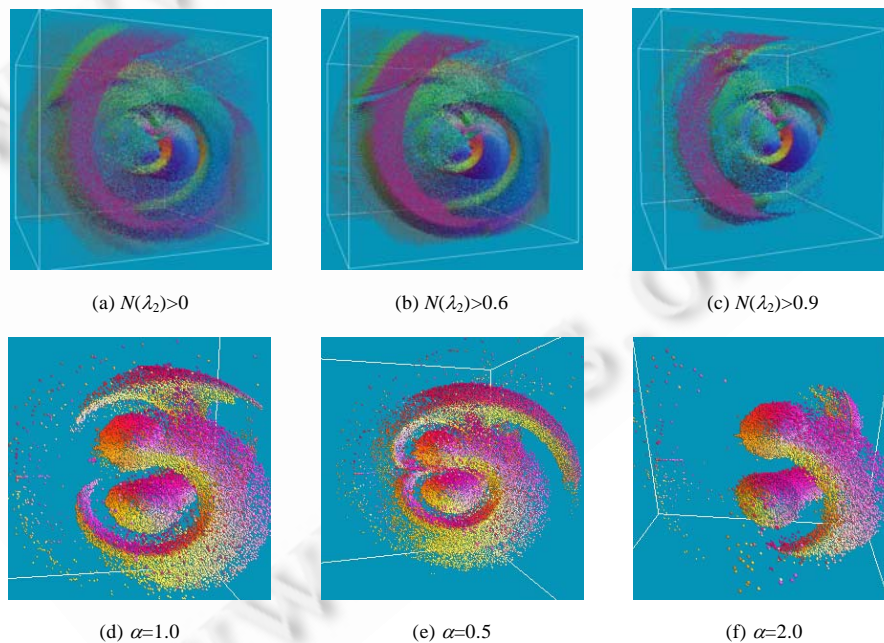


Fig.5 Extracting flow features of Lorenz attractor dataset by different fuzzy tone operator

图5 采用模糊语气算子对洛伦茨吸引子数据进行特征提取

台风数据具有复杂特征结构的 3D 流场.本文实验采用我国数值天气预报系统产生的海棠台风计算数据.图 6(a)为采用传统粒子系统方法对海棠台风数据进行可视化绘制的效果图,其中,颜色代表速度,速度值越大,颜色越浅.可以看到,流场中粒子相互遮挡,流场内的特征结构淹没在杂乱运动粒子中.图 6(b)和图 6(c)为分别采用速度过滤方法和本文方法对台风区域进行提取后的可视化效果图.其中,白色粒子区域为提取的台风区域,灰黑色粒子区域为背景流场区域.同时,对背景区域进行了提高透明度处理.从图 6(b)中可以看到,速度过滤方法虽然能够部分地提取台风区域,但对于流场上侧和左侧风速也比较大的地方,该方法并不能有效过滤,存在错误提取的特征区域.而图 6(c)中,基于本文方法提取台风特征区域可以比较准确而清晰地反映出台风的结构和运动情况,并通过透明度处理,较好地解决了遮挡影响,从而使用户能够更好地分析台风的性质.

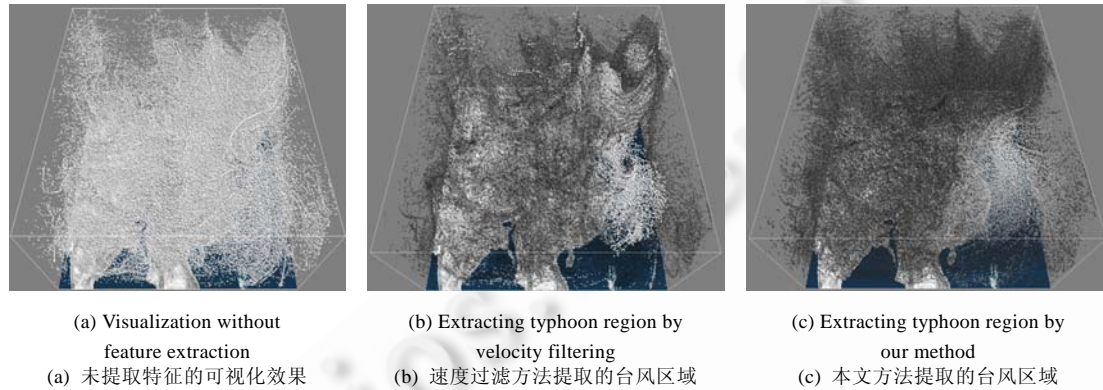


Fig.6 Comparison of traditional method and our method for visualizing Haitang typhoon data

图 6 传统方法与本文方法对海棠台风数据可视化比较

本文方法不仅对 3D 流场有效,而且对 2D 流场同样适用.图 7(a)~图 7(c)为对海棠数据场中间某层中的反气旋、鞍型点以及气旋涡旋区域进行提取后得到的可视化效果图.图 7(a)中,白色粒子区域为提取的气旋特征区域,灰黑色粒子区域为背景区域.图 7(b)中,白色粒子区域为反气旋特征区域.图 7(c)中,白色粒子区域为提取的鞍型点特征区域.从图中可以清晰地看到,不同特征区域的影响范围可以帮助用户更好地分析关注特征的运动和变化.

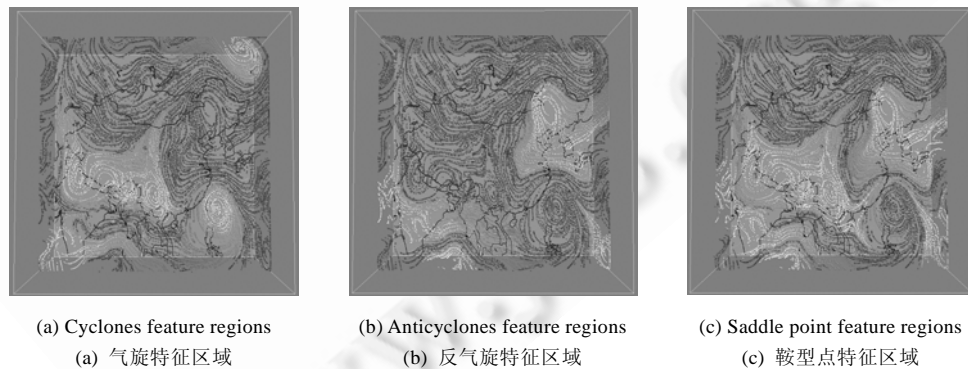


Fig.7 Extracting cyclones, anticyclones and saddles feature regions for one middle layer typhoon data

图 7 对台风中间某层数据的气旋、反气旋和鞍型特征区域进行提取

飞行器数据内包含了多个涡结构,如图 8(a)所示为传统 λ_2 方法对涡结构区域提取的结果,比较准确地得到了飞行器后侧的涡结构.图 8(b)为本文方法提取得到的涡区域结果,图 8(c)为速度过滤方法得到的结果.三者对

比可以发现, λ_2 方法提取的涡结构相对比较准确,但是无法描述所有穿过涡结构的气流(粒子),速度过滤方法则在一定程度上得到了过涡结构的粒子,但存在间断现象,如顶端和尾部的粒子不连续,相对而言,本文方法对过涡结构的气流粒子描述较好,从图 8(b)中可以比较清晰地分析过涡结构的气流在流场内的行为。

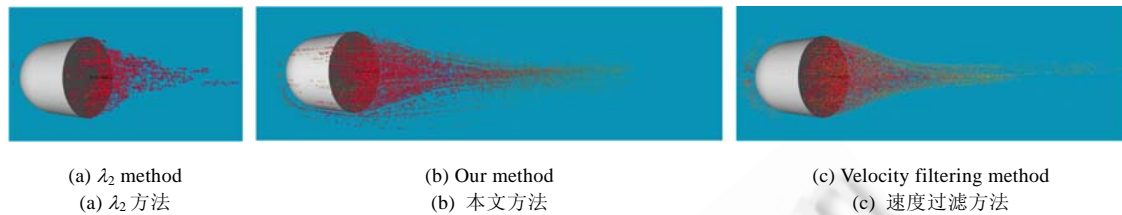


Fig.8 Comparison results for extracting the feature regions of vortex in a spacecraft data by different methods

图 8 采用不同方法对飞行器数据内涡特征区域提取对比效果图

5 结束语

传统的拓扑分析方法侧重于对流线渐近性质的描述,难以合理界定流场中特征结构的区域.本文基于模糊理论对流场特征区域进行了描述,从基本属性、衍生属性和关联属性这 3 个层次给出了相关模糊测度规则,并设计了一种在最小平方和准则下的流场的最优模糊划分算法.理论分析和实验均表明,本文的方法具有良好的性质,对流场中的典型特征结构区域能够合理地进行描述,并能有效缓解 3D 流场可视化面临的遮挡问题,使用户能够更好地分析流场的性质.

目前,本文的方法仍然局限于结构化网格数据.我们下一步将在非结构化网格数据上对本文方法进行测试,并对更为复杂的流场特征,如湍流进行验证.此外,FEFR 算法计算复杂度与临界点个数紧密相关.当临界点个数较大时,如何选取合适的规则减少临界点个数以及加速特征模糊区域提取,也需要进一步研究.

References:

- [1] Helman JL, Hesselink L. Visualizing vector field topology in fluid flows. *IEEE Computer Graphics and Applications*, 1991,11(3): 36–46. [doi: 10.1109/38.79452]
- [2] Scheuermann G, Hagen H, Krüger H, Menzel M, Rockwood A. Visualization of higher order singularities in vector fields. In: *Proc. of the IEEE Visualization 1997*. Los Alamitos: IEEE Computer Society Press, 1997. 67–74. [doi: 10.1109/VISUAL.1997.663858]
- [3] Scheuermann G, Tricoche X, Hagen H. C^1 -Interpolation for vector field topology visualization. In: *Proc. of the IEEE Visualization 1999*. Los Alamitos: IEEE Computer Society Press, 1999. 271–278. [doi: 10.1109/VISUAL.1999.809897]
- [4] Scheuermann G, Hamann B, Joy KI, Kollmann W. Visualizing local vector field topology. *Journal of Electronic Imaging*, 2000, 9(4):356–367. [doi: 10.1117/1.1289350]
- [5] Tricoche X, Scheuermann G, Magen H. Continuous topology simplification of planar vector fields. In: *Proc. of the IEEE Visualization 2001*. Los Alamitos: IEEE Computer Society Press, 2001. 159–166. [doi: 10.1109/VISUAL.2001.964507]
- [6] Tricoche X, Wischgoll T, Scheuermann G, Hagen H. Topology tracking for the visualization of time-dependent two-dimensional flows. *Computers & Graphics*, 2002,26(2):249–257. [doi: 10.1016/S0097-8493(02)00056-0]
- [7] Weinkauff T, Theisel H, Hegel HC, Seidel HP. Topological construction and visualization of higher order 3D vector fields. *Computer Graphics Forum*, 2004,23(3):469–478. [doi: 10.1111/j.1467-8659.2004.00778.x]
- [8] Weinkauff T, Theisel H, Shi K, Hege HC, Seidel HP. Extracting higher order critical points and topological simplification of 3D vector fields. In: *Proc. of the IEEE Visualization 2005*. Los Alamitos: IEEE Computer Society Press, 2005. 559–566. [doi: 10.1109/VISUAL.2005.1532842]
- [9] Kaufman A, Mueller K. *The Visualization Handbook*. Oxford: Elsevier Inc., Academic Press, 2005. 295–309.

- [10] Berdahl CH, Thompson DS. Education of swirling structure using the velocity gradient tensor. *AIAA Journal*, 1993,31(1): 97–103. [doi: <http://dx.doi.org/10.2514/3.11324>]
- [11] Jeong J, Hussain F. On the identification of a vortex. *Journal of Fluid Mechanics*, 1995,285:69–94. [doi: 10.1017/S0022112095000462]
- [12] Sadarjoen IA, Post FH, Ma B, Banks DC, Pagendarm HG. Selective visualization of vortices in hydrodynamic flows. In: Ebert DS, Rushmeier H, Hagen H, eds. *Proc. of the IEEE Visualization 1998*. Alamitos: IEEE Computer Society Press, 1998. 419–422. [doi: 10.1109/VISUAL.1998.745333]
- [13] Globus A, Levit C, Lasinski T. A tool for visualizing the topology of three-dimensional vector fields. In: Nielson GM, Rosenblum LJ, eds. *Proc. of the IEEE Visualization 1991*. Alamitos: IEEE Computer Society Press, 1991. 33–39.
- [14] Tricoche X. *Vector and tensor field topology simplification tracking and visualization* [Ph.D. Thesis]. Rheinland-Pfalz: University of Kaiserslautern, 2002.
- [15] Gao HB. *Fuzzy Cluster Analysis and its Application*. Xi'an: Xidian University Press, 2004. 37–48 (in Chinese).
- [16] Morris WL, Stephen S, Robert LD, Wrote; Gan SB, Trans. *Differential Equations, Dynamical Systems, and an Introduction to Chaos*. 2nd ed., Beijing: Posts & Telecom Press, 2008. 244–262 (in Chinese).

附中文参考文献:

- [15] 高洪波. 模糊聚类分析及其应用. 西安: 西安电子科技大学出版社, 2004. 37–48.
- [16] Morris WL, Stephen S, Robert LD, 著; 甘少波, 译. 微分方程、动力系统与混沌导论. 第2版, 北京: 人民邮电出版社, 2008. 244–262.



徐华勋(1977—),男,河北平乡人,博士生,主要研究领域为科学计算可视化.



马千里(1983—),女,博士生,主要研究领域为科学计算可视化,虚拟现实技术.



李思昆(1941—),男,教授,博士生导师,主要研究领域为虚拟现实与可视化技术,嵌入式系统,SoC设计方法学.



蔡勋(1971—),男,博士,副研究员,主要研究领域为科学计算可视化技术.