

第6章 大规模多变量空间数据场可视化

数据拓扑 - 结构化网格

➤ 科学数据大致可分为

- 多维度(Multi-Dimensional, 包括时变数据)
- 多变量(Multi-Variate)
- 多模态(Multi-Modal)
- 多趟(Multi-Run)
- 多模型(Multi-Model)

- 多维度用于表达物理空间中独立变量的维数, 以及是否包含时间维度, 侧重于表达空间和时间概念
- 多变量则用于表达变量和属性的数目, 表示数据所包含信息和属性的多寡
- 多模态强调获取数据的方法不同, 以及各自对应的数据的组织结构和尺度的不同
- 多趟和多模型亦可表示数据所含信息, 但和多变量属于不同的概念, 例如单变量多值数据, 在不同的计算参数或不同的计算模型下获得的多个仿真结果, 每一个采样点含有属于同一数据属性的多个值, 其重点在于描述“值”的个数, 而不是数据属性和变量的个数

- 三维气象仿  间也是一个  数据集是时变数据, 则时

- 若源数据通过卫星探测、地面站点观测、数值模拟等多种方式获得, 则称为多模态数据

1 大规模空间标量场数据的实时可视化

大规模空间标量场数据的单机绘制

- 大规模标量场数据的单机绘制方法可以分为三类：
 - 硬件加速:基于图形硬件的快速绘制算法
 - 数据压缩(含多分辨率显示)
 - 外存计算

硬件加速

- 基于图形硬件的快速绘制算法
- Beyer等认为交互性可以通过使计算和可视化正比于屏幕上真实可见的数据量实现，即输出敏感(Output-Sensitive)的算法和系统设计，包括光线引导(Ray-Guided)算法、可视化驱动(Visualization-Driven)算法和显示感知(Display-Aware)算法
- Frey等人提出了一种时空误差控制的交互式渐进标量场数据可视化方法
- Hadwiger等人提出了基于GPU的SparseLeep[Hadwiger2018]，混合了物体空间和图像空间的技术，实现了基于八叉树的快速空区间跳跃，能够高效绘制大规模的稀疏标量场数据

数据压缩

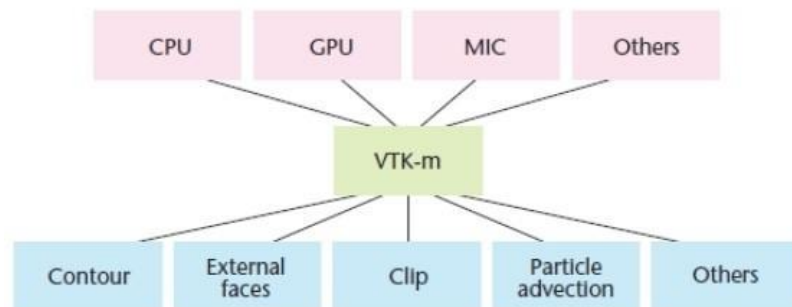
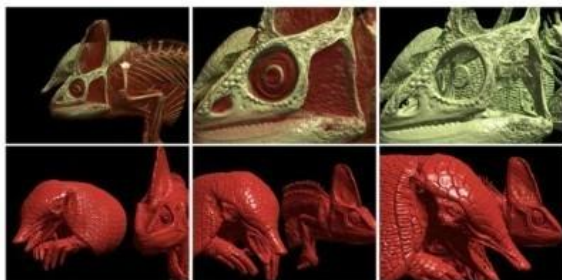
- 数据压缩是解决大规模空间数据处理的主要思路，Rodríguez等人系统地综述了基于GPU的压缩标量场数据可视化方法，数据压缩可以追求高精度，但不要求实时；数据解压则要求能够实时，实现大规模标量场数据的实时可视化 [Rodríguez2014]
- 体数据的多分辨率绘制方法将体数据投影到小波基函数张成的空间(即计算小波系数)，求得任意分辨率下的绘制结果。细节层次(Level-of-Detail)的方法 [Weiler2000]和基于视点的数据压缩优化 [Guthe2000]也可用于数据的自适应存取。将LBG(Linde Buzo Gray)算法应用于体数据压缩 [Schneider2003]，并分析原数据的多分辨相关性，进行层次化的矢量化，可大幅减小显存的占用
- Lindstrom等人 [Lindstrom2006]提出了一种对浮点格式数据进行快速压缩的算法，该方法能够与应用程序的I/O环节无缝连接且适用于可变精度的浮点或整型数据。压缩后的可视化质量可采用合适的图像质量度量标准 [Wang2007]
- Xu等人将压缩感知应用于标量场数据的原位数据缩减，可以在没有任何特征先验知识的情况下，从分解获得的部分傅里叶系数中准确重构出原始数据 [Xu2014]
- Soler等人提出了一种拓扑控制的数据压缩方法，该方法基于拓扑自适应的数值量化，能够保留重要特征，控制被去除的细小特征数量 [Soler2018]

数据的统计信息

- 数据的统计信息最初用来表示数据中的不确定性，近年来被用来表示大规模数据的紧凑代理，即每个数据块都使用其统计分布进行表示
- Johnson和Huang提出了一种通过模糊匹配用户指定分布和数据局部分布的特征查询方法 [Johnson2009]。Wei等人利用位图索引(Bitmap Indexing)的方法加速局部直方图的搜索 [Wei2015]
- Thompson等人提出了hixels数据表示方法，利用直方图存储一个数据块内的数据分布或一个采样点的一系列组集数据 [Thompson2011]
- Liu等人提出了利用高斯混合模型对数据块中的数据进行建模，并在GPU上基于高斯混合模型随机采样对数据进行绘制 [Liu2012]
- Dutta等人进一步改进了基于数据块的空间划分策略，采用简单线性迭代聚类方法划分空间，使块内的差异最小化，从而减少数据块的统计分布所需的存储空间 [Dutta2017a]
- Wang等人改进了已有的基于高斯混合模型的方法，增加了基于数值的空间概率分布，在获取给定空间位置的数值时，利用贝叶斯定律同时结合数值分布和空间分布，更准确地计算该位置上可能的数值概率分布 [Wang2017a]。与数据压缩相比，基于统计分布的数据表示更易于数据的分析、查询和可视化，也可以应用于原位可视化的数据缩减
- 对于无法存储于内存和显存的数据集，可采用外存计算(Out of Core)处理大规模数据 [Farias2001]。细节层次技术是适用于不规则体数据可视化的外存计算的方法 [Du2010]
- Gobbetti等人 [Gobbetti2008]提出了一种只需遍历数据一次的光线投射外存计算方法

大规模空间标量场数据的并行绘制

- 对于大规模数据的分布式绘制，如何提高算法的并发性、设计负载均衡都是研究人员首要关注的问题
- 大规模空间数据的并行绘制研究可以分为CPU与GPU集群并行计算两个阶段
 - 利用GPU构建分布式的计算与可视化平台是显卡工业发展的必然
 - 多核/众核架构是当前高性能计算的发展趋势，其并行性更强，但内存并没有按比例增长，使得算法需要更高的可扩展性



时变空间标量场数据加速绘制方法

- 时空连贯性是加速时变空间标量场数据可视化的基本思路
- 从时变空间标量场数据中提取时变等值面，是展示形状演化规律的基本方法。类似于单个时间步上的等值面提取加速，也可采用利用时空连贯性的加速方法，避免处理不必要的体素，提高时变等值面的提取效率
- 时变空间标量场具有连贯性强、规模大、多变量等特点，为直接体绘制的加速提供了不同的契机，如保特征的数据压缩、流水线绘制等。数据压缩一般分为分块、变换、量化和编码4个阶段，不同的数据划分方式将对应不同的压缩过程
 - 一类方法是将数据按照时间维度进行划分，先对单帧数据进行压缩，再考虑时间相关性
- 从时变标量场抽取的时变特征既可以通过融合多个时间步的可视化结果展示，也可以通过动画播放或并排放置多个时间步绘制结果以观察时变特征的差异和演化规律。对于相邻时间帧的特征信息，可采用表意性可视化方法叠加在当前时刻的可视化结果上

2 时变异构空间数据场的特征追踪与可视化

时变异构空间数据场的特征追踪与可视化

- 就大规模、时变、异构(高维)的特点而言，时变空间标量场数据可视化既需要高性能、自适应的显示手段，突出用户感兴趣的特征，又需要能够融合不同的绘制方法，从多个角度辅助用户进行分析。就可视化系统的设计而言，应当将数据、特征、感知与交互作为核心要素
- 时变异构空间数据场可视化的两大挑战是数据的不规则与多变量
 - 不规则表明了科学数据的复杂形态
 - 多变量则是物理空间的本质特征

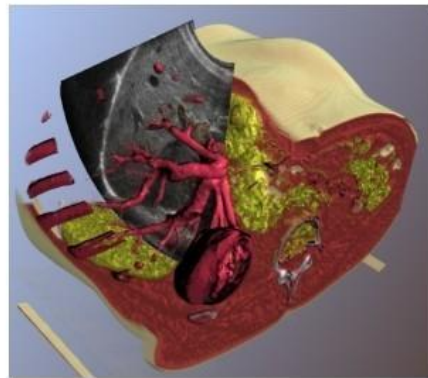
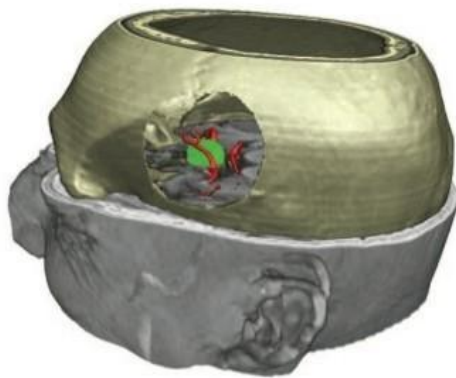
时变空间标量场数据的特征提取

➤ 特征

- 按照空间大小，时变数据的特征可以分为局部特征和全局特征；
- 按照时间变化规律，则可以分为常规(Regular)、周期(Periodic)和随机(Random)
 - 常规模式指特征在三维空间中稳定地移动或形变，其变化趋势既不剧烈，也不遵循周期性的路径
 - 周期模式指特征周期性地出现和消失，或沿着周期性路径进行移动
 - 随机模式指特征的变化规律较为随机，常见于湍流模拟
- 特征提取操作可在每个时间步单独提取。标量场数据分类是基本的可视化手段
- 特征提取的另一大类方法是将时变空间标量场数据看作一个四维数据，统一提取时变特征
- Fang等人 [Fang2007]将时变空间标量场等效成一个三维数组，每个元素都是一个时间活动曲线(Time-Activity Curve)，进而可采用三种相似性度量标准对这些曲线进行聚类 and 可视化，统一提取感兴趣的时变特征

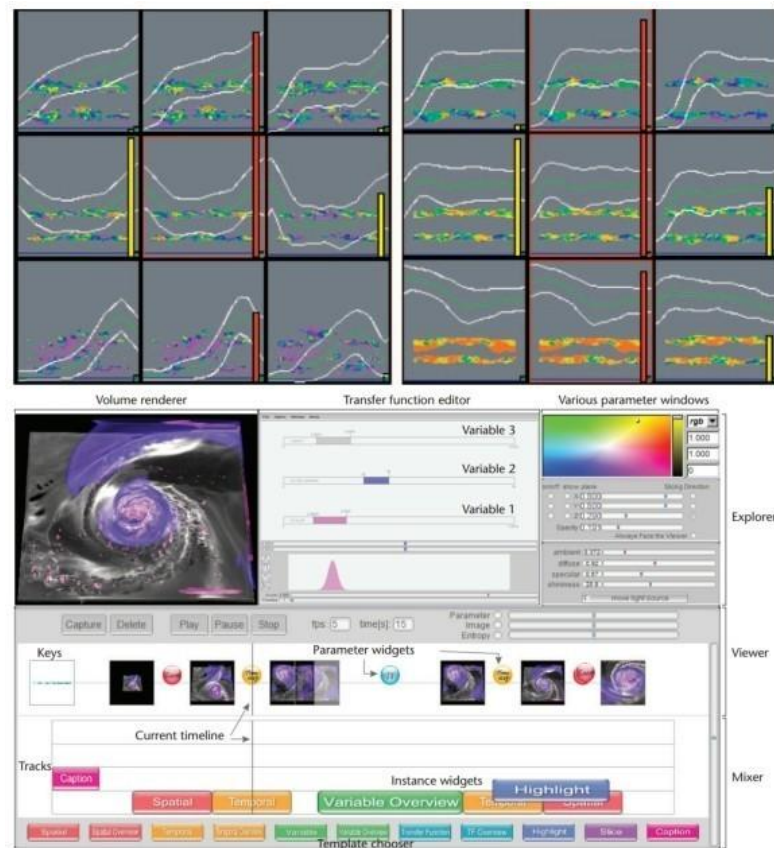
异构数据的特征融合

- 异构数据的特征融合涉及多维度的融合与多模态的融合两个方面。实现多维度融合可以采用信息可视化的相关方法
 - 如平行坐标、散点图、时间直方图、纹理、影线、图标(Glyph)等
 - 可以与人机交互、科学可视化紧密地结合
- 多模态融合的挑战在于数据的融合方式难以定义，空间分布也可能差异很大。因此，这类数据一般需先经过配准，并转化为相同格式后才能进行融合可视化



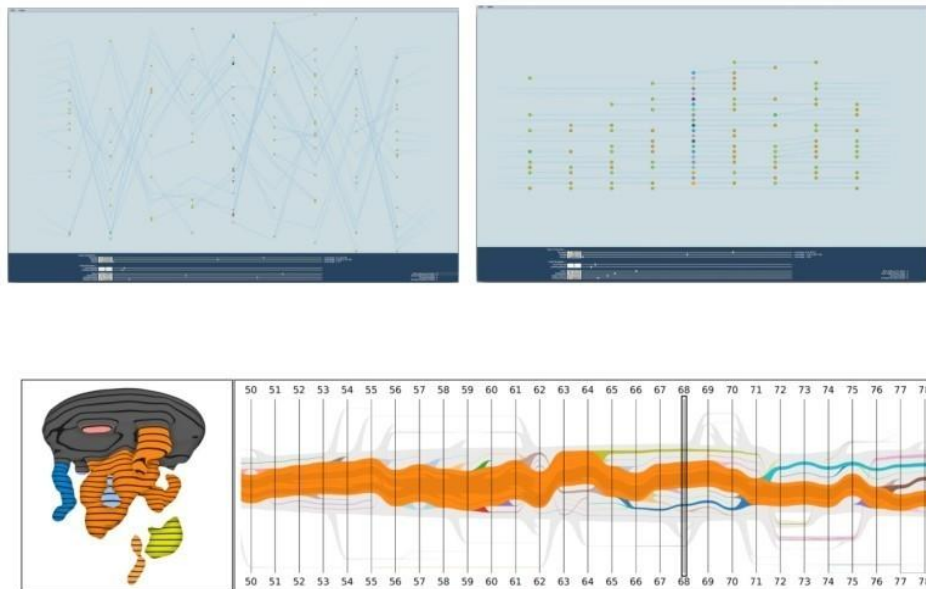
时变空间标量场数据的特征追踪

- 早期的特征跟踪方法 [Samtaney1994]采用计算机视觉、图像处理、计算几何等方法，直接保留和追踪流场特征
- 此后的方法基本围绕图示 (Illustration)[Joshi2005]、轮廓 (Contour)[Sohn2006]、视点 [Ji2006a]和纹理 [Caban2007]等方法展开



时变数据的特征追踪

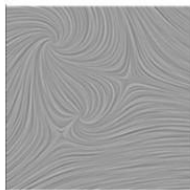
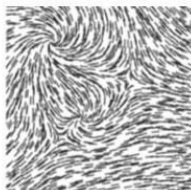
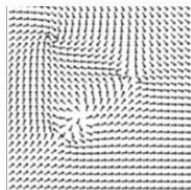
- 时变数据的特征追踪应结合多视图、用户交互以及可视分析方法，从时间、数据与表现三个角度出发确定具体的可视化原则



3 空间向量场数据可视化

空间向量场

- 向量场的每个采样点的数据是一个向量(一维数组), 向量场可视化的主要目标
 - 展示场的导向趋势信息
 - 表达场中的模式
 - 识别关键特征区域
- 流场可视化将物理过程产生的或与之相关的现象, 通过可视化以人眼能感知的图像形式显示, 使得对此过程的洞察更为清晰
- 流场可视化的核心目标, 是设计感知有效的流表示方式描绘其流动信息
 - 相关的重要问题, 包括效率(计算速度和存储成本)、数据尺寸和复杂需求、随时间变化的非定常流、复杂网格以及多种变量的可视化、流场特征提取和跟踪等



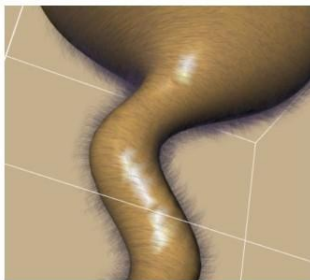
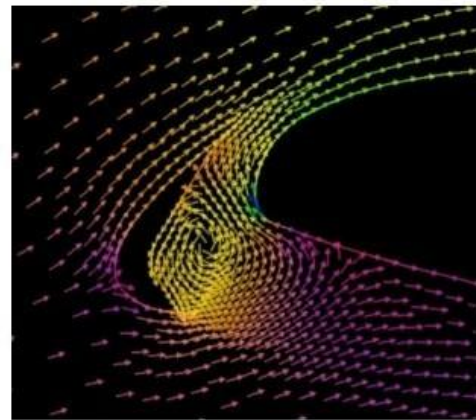
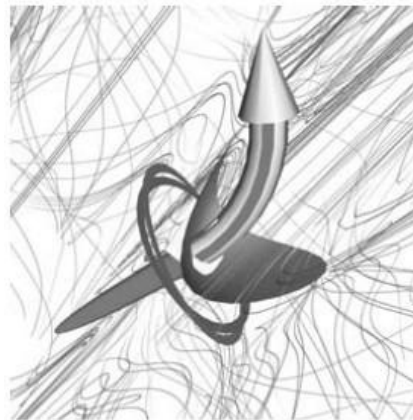
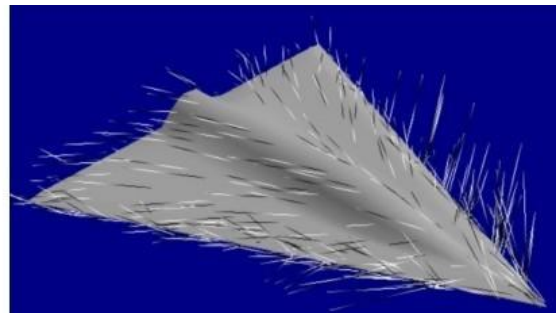
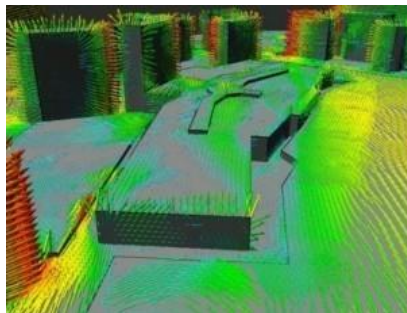
图标法

➤ 采用图标逐个表达向量

- 箭头的方向代表向量场的方向，长度表示大小。
图标的尺寸、颜色、形状等视觉通道可用来表示其他信息

➤ 图标法所采用的图标主要有

- 线条(Hedge Hogs)、箭头和方向标志符
- 除速度外，可以将向量场的雅可比矩阵分解，将其主成分以直观的方式映射到可视图标隐喻。典型的可编码数据包括：速度、加速度、曲率、局部旋度、切变速率和收敛度等



几何法

- 几何法，指采用不同类型的几何元素模拟向量场的特征
 - 如线(Curve)、面(Surface)和体(Volume)
- 不同类型的几何元素和方法适用于不同特征(稳定、时变)和维度(二维、三维)的向量场

基于曲线的可视化

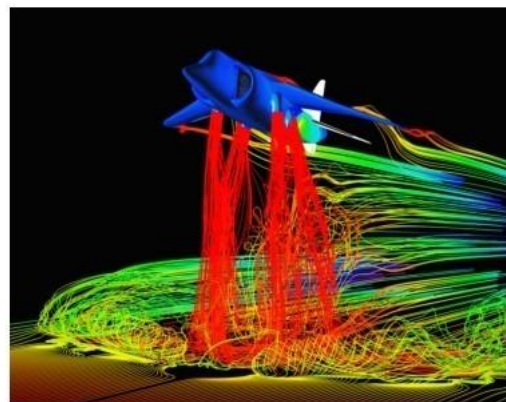
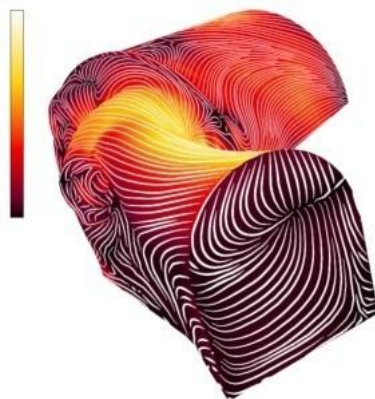
➤ 基于曲线的可视化方法包括两类

- 面向稳定向量场，如流线(Stream Line)；面向不稳定/时变向量场，如迹线(Path Line)和脉线(Streak Line)
- 流线，描述向量场空间中任意一点处向量场的切线方向
 - 对于向量场空间中一个特定的位置(除临界点(Critical Point)外)，某一时刻有且仅有一条流线通过该点。该点处流线的切线方向即表示该点处向量场的方向
 - 生成流线时，首先，在向量场空间中播撒种子点，然后，从种子点发射粒子，对向量场进行采样。根据采样得到的向量，平移粒子，不断迭代得到一条完整的流线

$$\frac{dx_{stream}(s)}{ds} = u(x_{stream}(s), \tau) \quad x_{stream}(0) = x_0$$

s 为流线轨迹参数，

τ 为某个时间点的流场



基于曲线的可视化

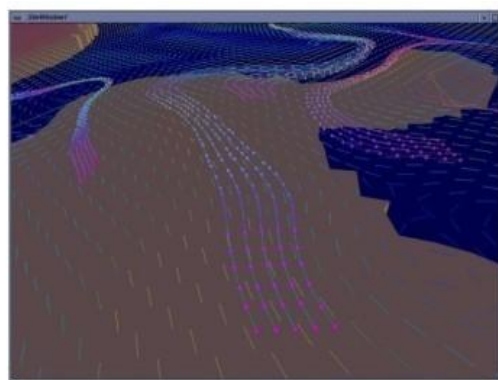
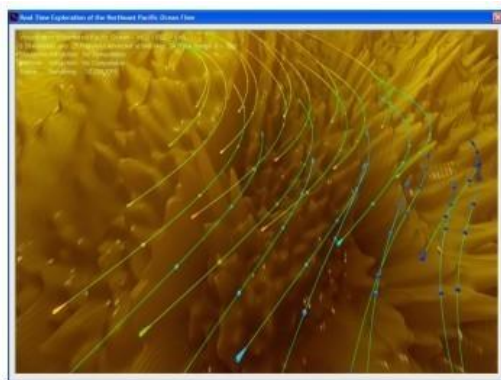
➤ 迹线，用于描述向量场中一个粒子在某个时间段的流动轨迹

➤ 适用于不稳定向量场，对于向量场空间中特定的位置 (除临界点外)，某一时刻有且仅有一条迹线通过该点

➤ 生成迹线时，同样首先要在向量场空间中撒入种子，然后从种子发射粒子，粒子仅在向量场的作用下运动，追踪其运动轨迹即获得一条迹线

$$\text{➤ } \frac{dx_{path}(t)}{dt} = u(x_{path}(t), t) \quad \text{or} \quad x_{path}(0) = x_0$$

➤ 流线是假想的曲线，是不存在的；而迹线是实际存在的曲线，随着粒子的运动产生



脉线与时线

脉线

- 适用于不稳定向量场，用于描述一个粒子集合从一个起始点不同时间发射，在之后的某一时刻获取该集合中粒子的位置并连接形成的轨迹
- 直观地说，脉线的形成过程可以看作在一个向量场的特定位置处持续释放不计质量的有色染料，染料在向量场的作用下运动而形成的几何曲线。由于不稳定向量场随时间变化，这些粒子都有各自的轨迹

$$\frac{dx_{path}(t)}{dt} = u(x_{path}(t), t) \quad , \quad x_{path}(\tau) = x_0$$

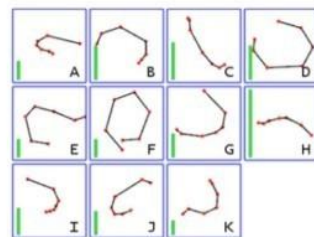
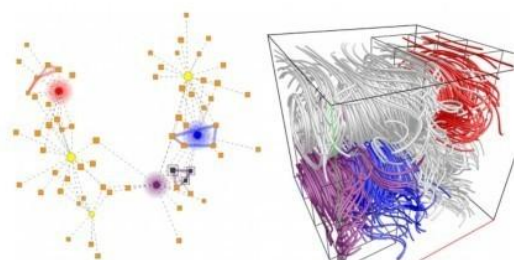
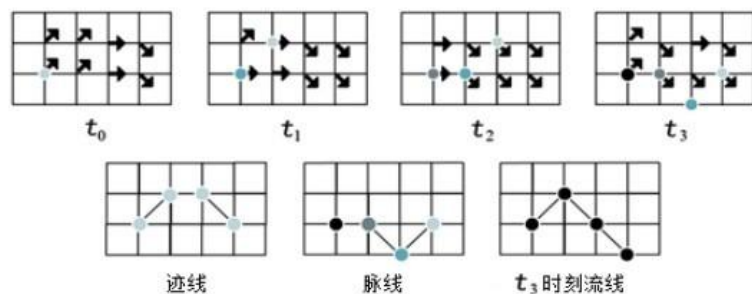
x_0 为释放粒子的位置, $0 \leq \tau \leq T$, T 是释放粒子的时间段

时线(Timeline)，脉线的一个扩展形式

- 指一个粒子集合从一条起始轨迹或一个起始区域上的不同位置发射，并以类似于脉线的方式获取其运动轨迹的方法
- 简而言之，时线的方法所产生的轨迹是脉线的一个集合，能够反映一个局部区域内的向量场的特征

向量场分析

- 稳定向量场：流线、迹线和脉线，相同
- 不稳定/时变向量场：迹线和脉线，不同
- 向量场分析，不仅仅需要上述曲线的可视化方法，还需要探索向量场中隐含的特征或模式
- Ma等人提出了基于图的流线和迹线可视分析框架 [Ma2014]，图的节点包括空间区域和流线集合两类，通过图的路径比较、区域比较、图的变换等操作，实现流线集合和空间区域的交互分析



基于积分曲线的方法

Constructing Streak Surfaces for 3D Unsteady Vector Fields

Tony McLoughlin, Robert S. Laramée and Eugene Zhang

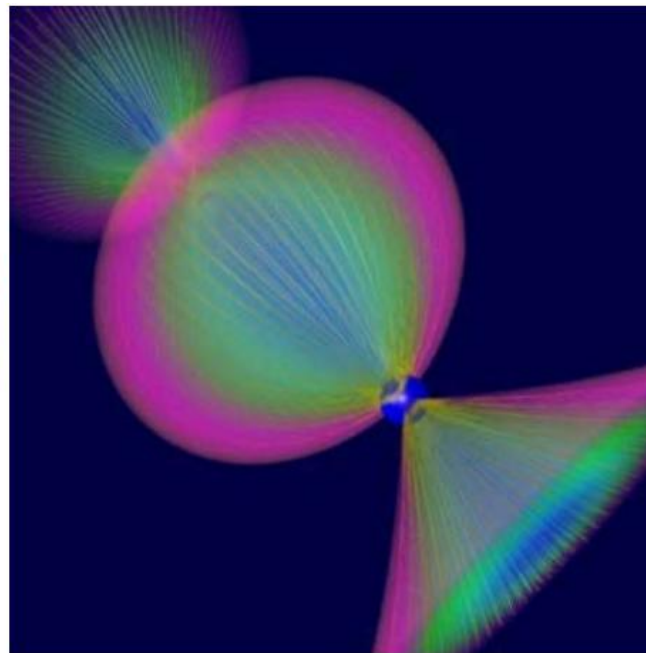
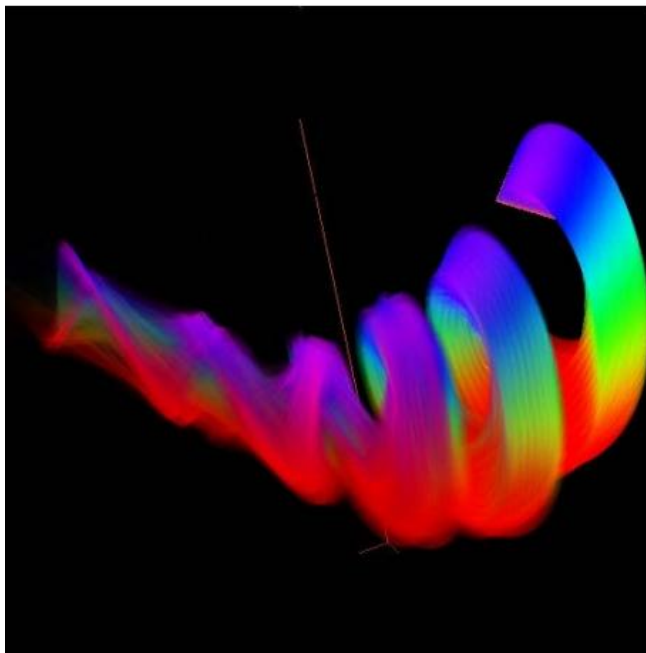


Supplementary video

Constructing streak surfaces for 3D unsteady vector fields, [SCCG '10](#)

基于积分曲线的方法

➤ 流体、流多边形



左：对龙卷风数据的流体可视化结果[Max1993]；右：对充电耦合(Charge Couple)数据集的隐式流体可视化结果[Xue2004]

基于积分曲线的方法

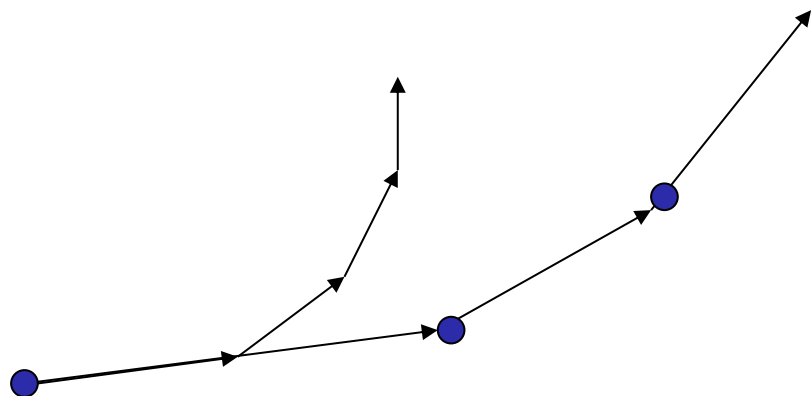
➤ 粒子



IEEE Visualization Contest

提升积分曲线的方法的质量

- 关键技术1：积分曲线的计算
- 欧拉方法：简单快速，但精度太低
- $x(t+dt) = x(t) + v(x(t)) * dt$
- 改进欧拉方法
- 龙格-库塔方法：二阶、四阶
- 误差控制、自适应步长选择



提升积分曲线的方法的质量

➤ 关键技术1：积分曲线的计算

➤ RK-4

➤ Christian Teitzel , Roberto Grosso , Thomas Ertl, Efficient and Reliable Integration Methods for Particle Tracing in Unsteady Flows on Discrete Meshes, Visualization in Scientific Computing '97

$$\Delta y = \frac{1}{6} [k_1 + 2k_2 + 2k_3 + k_4]$$

$$k_1 = \Delta h [f(x, y)]$$

$$k_2 = \Delta h \left[f \left(x + \frac{1}{2} \Delta h, y + \frac{1}{2} k_1 \right) \right]$$

$$k_3 = \Delta h \left[f \left(x + \frac{1}{2} \Delta h, y + \frac{1}{2} k_2 \right) \right]$$

$$k_4 = \Delta h [f(x + \Delta h, y + k_3)]$$

		RK2	RK3	RK4	RK2(1)	RK3(2)	RK4(3)
vortex	integr. steps	3694	1849	1233	691	239	166
	v-evaluations	7387	5545	4929	2582	1957	2211
air flow	integr. steps	5000	4147	3290	146	53	46
	v-evaluations	9999	12439	13157	636	225	235
laminar flow	integr. steps	291	146	97	94	34	29
	v-evaluations	581	436	387	346	180	189

提升积分曲线的方法的质量

- 关键技术2：种子点的选取、流线放置
- (1) 过多  视觉混乱
- (2) 过少  表达不完整

研究方向：

- (1) 种子点的放置
- (2) 流线的简化

种子点的放置策略

好的种子点策略的一般特征：

覆盖性 (Coverage)

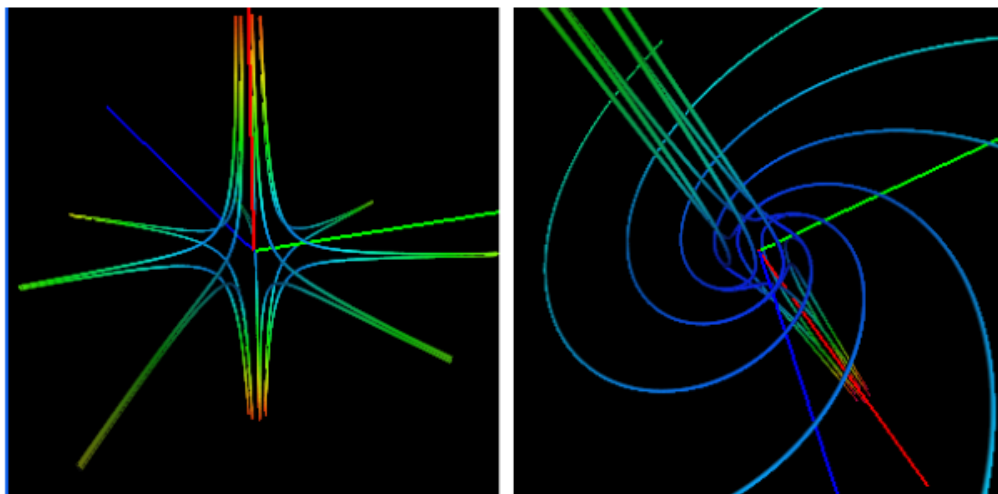
均匀性 (Uniformity) 连续性 (Continuity)

种子点的放置策略

► 图像指导

均匀放置

拓扑指导

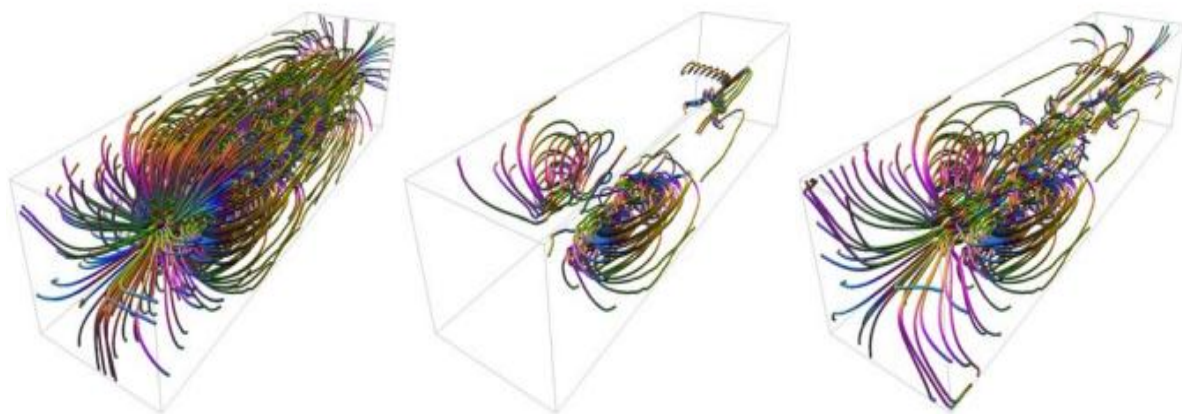


【 Ye. et al. 05 】

流线的简化

► 核心思想：使用指定的度量标准进行流线删除或聚类。

相似度



信息论

【 Marchesin. et al. 10 】

长度熵
$$E_L = -\frac{1}{\log_2(m+1)} \sum_{j=0}^m \frac{D_j}{L_S} \log_2 \frac{D_j}{L_S}$$

角度熵
$$E_A = -\frac{1}{\log_2(m)} \sum_{j=0}^{m-1} \frac{A_j}{L_A} \log_2 \frac{A_j}{L_A}$$

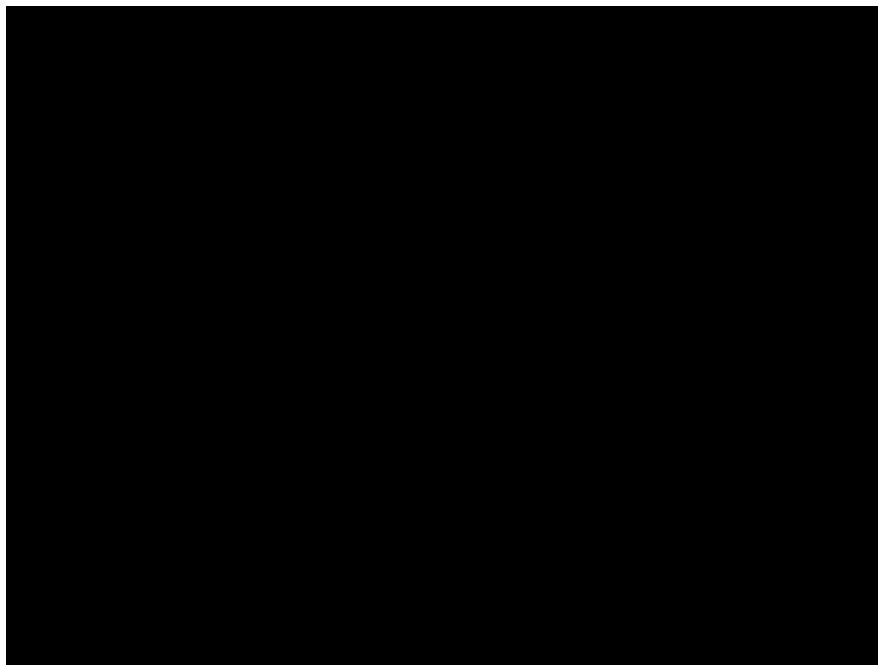
基于积分曲线的方法

Evenly-Spaced Streamlines for Surfaces: An Image-Based Approach

Ben Spencer, Robert S. Laramée, Guoning Chen
and Eugene Zhang

B.Spencer, et al. Evenly-Spaced streamlines for surfaces: An image-based Approach. Computer Graphics Forum 2009.

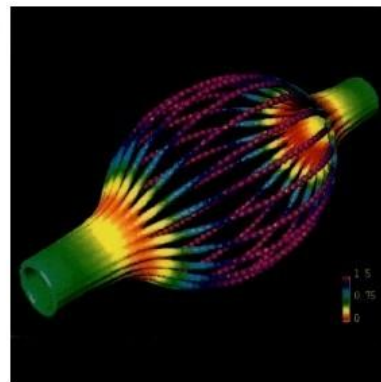
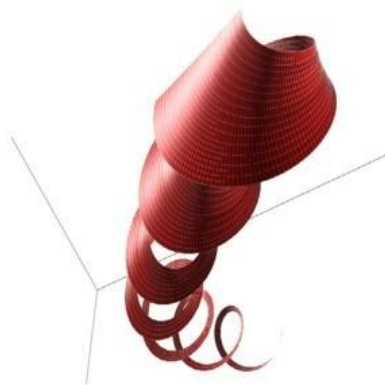
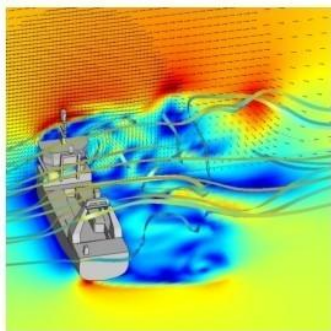
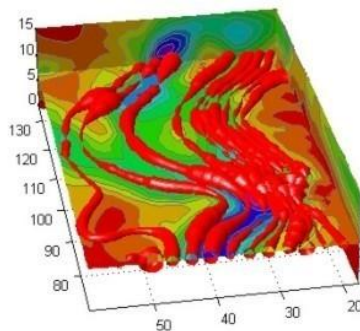
基于积分曲线的方法



M Edmunds, et al. Automatic Stream Surface Seeding: A Feature Centered Approach.
Computer Graphics Forum 2012

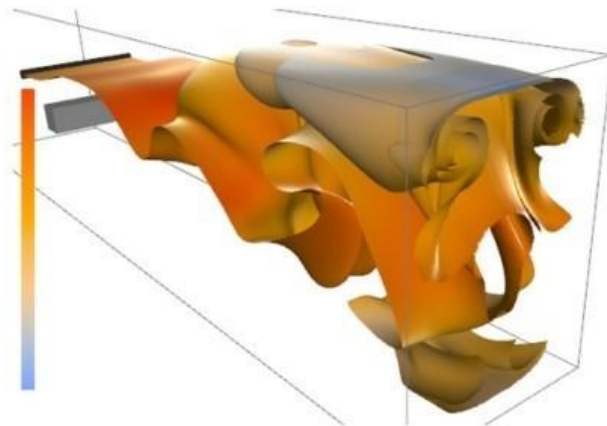
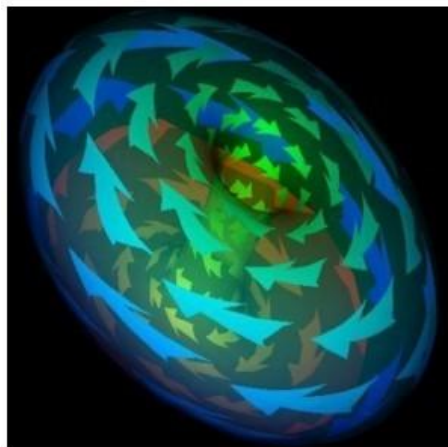
基于曲面的可视化

- 提供了更好的用户体验和感知，并能显著降低视觉混淆，更为真实地揭示向量场的结构
 - 面向稳定向量场，包括流面(Stream Surface)、流球(Stream Ball)和流形箭头(Stream-Arrow)
 - 面向不稳定/时变向量场，例如脉面(Streak Surface)等
- 流面的可视化方法，以流线为基础，在三维空间的特定位置(如曲线、曲面和四面体网格等)播撒种子点并生成流线簇，以此揭示向量场特征
- 流球的可视化方法，借用元球(Metaball)的概念，将连续的流线离散化为流线上的一组点，用一个流球表示每个点



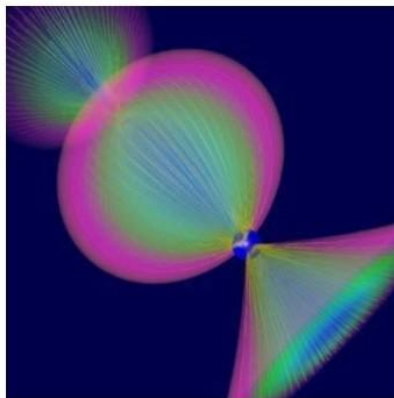
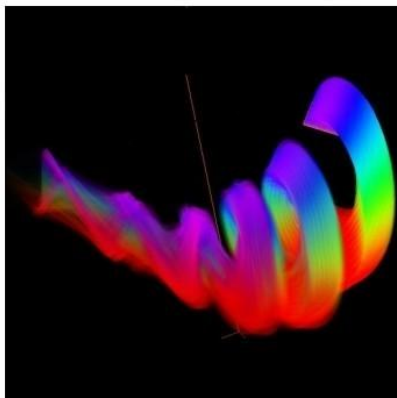
扩展

- 流形箭头 [Löffelmann1997]将箭头的纹理图案镶嵌于原来的流面上，不同曲率的流面应用不同分辨率的纹理图案，以最大限度地消除图案的扭曲
- 脉面的可视化，以脉线为基础，适用于不稳定时变向量场
 - 由于向量场的高度复杂性(四维数据)和人眼对时变数据的不易感知性，脉面方法面临的技术挑战是降低计算复杂度、保持数据精确性和降低视觉混淆



基于体的可视化

- 基于曲线或者曲面的可视化方法通常难以展示三维向量场完整的整体性特征，如流的聚合和分散、漩涡、剪切和断裂等拓扑信息。基于体的可视化方法能有效地弥补这一缺陷，帮助用户探索向量场的拓扑结构
 - 主要包括流体(Flow Volume)等
- 流体可视化方法的主要思路：将追踪粒子撒入流场，捕获粒子的运动轨迹，形成三维空间内的一个流体，继而将该流体通过特定算法转化为一个四面体的集合，并利用体绘制技术进行可视化



纹理法

- 由于设计合理的种子点策略难度较大，往往要求用户对数据具有先验知识，因此传统的基于几何的方法在处理复杂向量场时经常产生不理想的可视效果
 - 例如，曲线和曲面的聚集与覆盖导致视觉混淆和结果的不完整，最终造成关键特征的丢失或掩盖
- 以纹理图像的形式显示向量场的全貌，能够有效地弥补图标法和几何法的缺陷，揭示向量场的关键特征和细节信息。根据处理对象的不同，纹理法主要包括三大类
 - 点噪声(Spot Noise)
 - 线积分卷积(Line Integral Convolution, LIC)
 - 纹理平流(Texture Advection)

点噪声

- 点噪声方法 [Wijk1991]，以单点(通常表示为除一有限区域外处处为0的函数)作为生成纹理的基本单元，将随机位置(均匀分布)、随机强度的点混合形成噪声纹理，并沿向量方向对噪声纹理进行滤波，纹理图像中条纹的方向反映了向量场的方向
 - 随机排列一些圆点，按照局部流场方向对圆点变形，将变形后的圆点用滤波器扩散到纹理中
- 点噪声方法的扩展方法适用于复杂多样的应用情形，包括加强的点噪声方法(Enhanced Spot Noise)[Leeuw1995]、并行点噪声方法(Parallel Spot Noise)[Leeuw1997a]和面向时变向量场的点噪声方法(Unsteady Spot Noise)[Leeuw1997b]等



线积分卷积

➤ 线积分卷积

- 与点噪声方法有一定相似性
- 基本思路，以随机生成的白噪声作为输入纹理，根据向量场数据对噪声纹理进行低通滤波，这样生成的纹理既保持了原有的模式，又能体现出向量场的方向

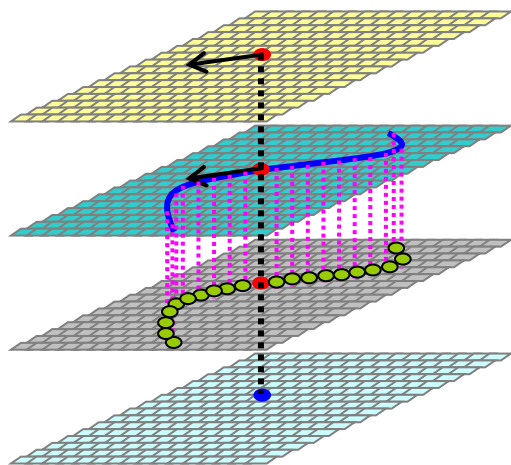


基于纹理的方法

➤ 线积分卷积 (line integral convolution, LIC)

➤ [Cabral and Leedom, Siggraph'93]

➤ 将矢量场与白噪声进行卷积



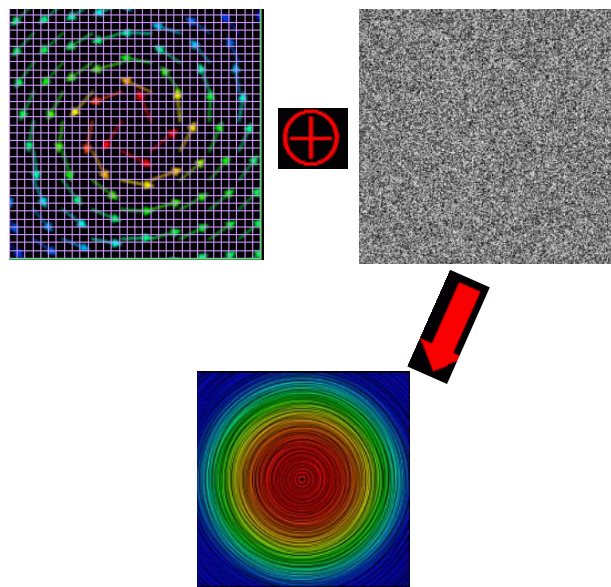
原始2D 矢量场

沿该点做双向积分

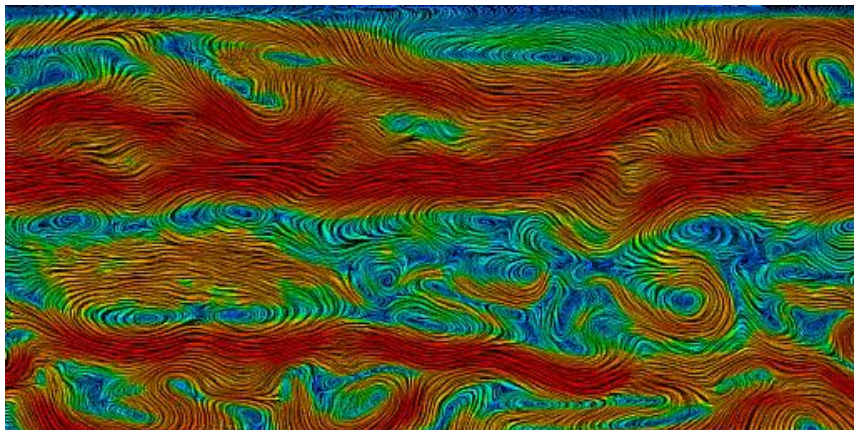
在白噪声纹理中查找积分路径上对应的各点

对各点进行卷积后，得到结果图像上该点的像素值

【 <http://zhanpingliu.org/Research/FlowVis/LIC/LIC.htm> 】

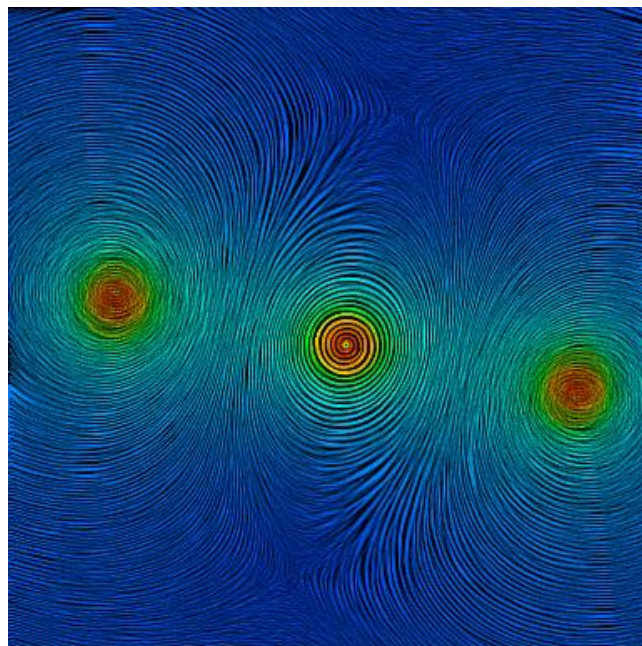


线卷积积分(LIC)



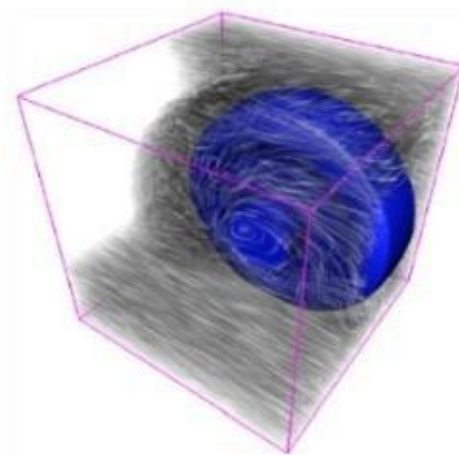
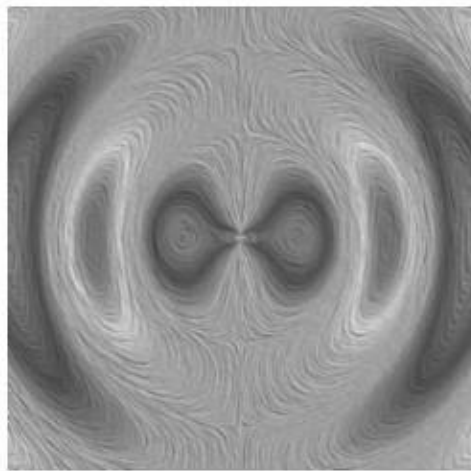
American Wind Field

<http://hint.fm/wind/gallery/mar-14.js.html>



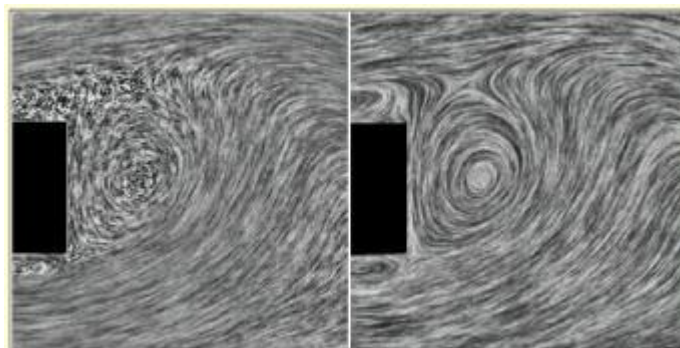
线积分卷积

- 与点噪声方法相比，线积分卷积方法成像质量较高，纹理细节清晰，能够精确地刻画向量场的特征，特别是鞍点、螺旋点和交点等拓扑特征，以及具有高曲率的局部区域
- 扩展到时变和三维向量场
 - 快速线积分卷积方法(Fast LIC)[Stalling1995]；面向时变向量场的线积分卷积方法(Unsteady LIC)[Forssell1995]；三维线积分卷积方法(3D LIC)[Rezk1999]；动态线积分卷积方法(Dynamic LIC)[Sundquist2003]；显示频率控制线积分卷积方法 [Matvienko20015]等



基于纹理的方法

➤ 点噪声方法与线积分卷积方法的比较



基于纹理的方法

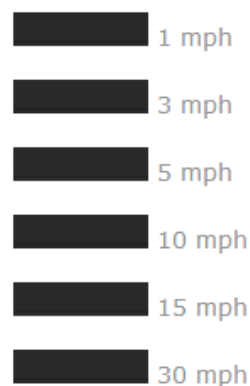
March 14, 2012

2:00 pm EDT

(time of forecast download)

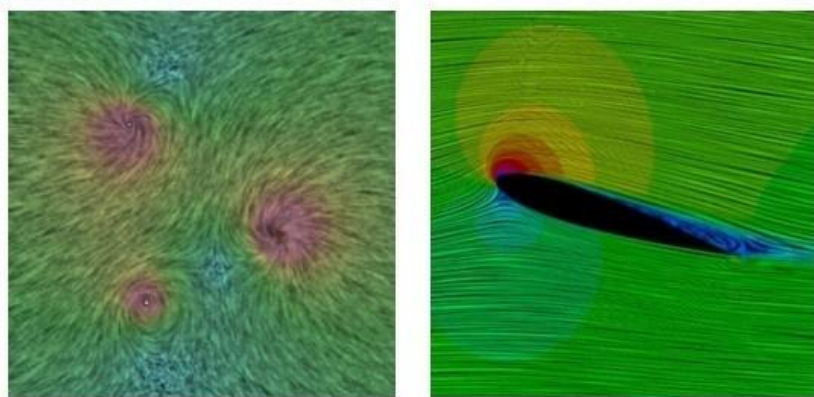
top speed: **34.0 mph**

average: **8.5 mph**



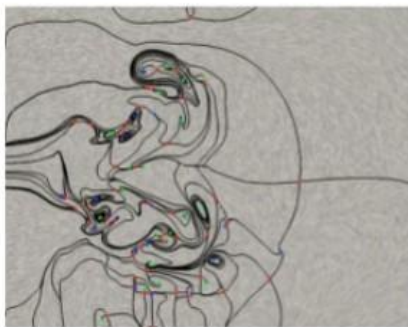
纹理平流

- 纹理平流法，时变流场可视化的标准方法之一。根据向量场方向移动一个纹元(纹理单元, Texel)或者一组纹元，以达到刻画向量场特征的目的。代表性
 - IBFV(Image Based Flow Visualization)[Wijk2002]
 - 将之前的若干帧图像和一系列经过滤波的噪声背景图像作为输入，经过卷积生成下一帧图像
 - UFLIC(Unsteady Flow LIC)
 - 从质点平流的角度重新阐释了LIC思想。其思路是：初始白噪声纹理中的每个像素点沿向量场方向运动，在经过的像素位置以灰度值的方式留下“印记”。经过多次平流后，对每个像素点记录的“印记”进行卷积，获得最终的向量场纹理图像



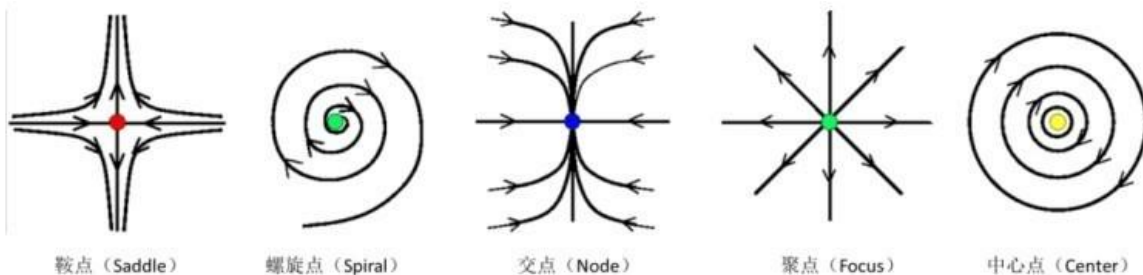
拓扑法

- 拓扑方法，主要基于临界点理论：任意向量场的拓扑结构由临界点和链接临界点的曲线或曲面组成
 - 临界点，指向量场中各个分量均为零的点。该方法是一种对向量场抽象描述的方法，让用户抓住主要信息，忽略其他次要信息，并且能够在此基础上对向量场进行区域分割
 - 由于具备丰富的数学理论基础，该方法适用于任意维度、离散或者连续的向量场
- 传统的向量场拓扑可视化方法主要由两步组成
 - 临界点位置的计算与分类；链接临界点的积分曲线或曲面，即向量场区域边界的计算。



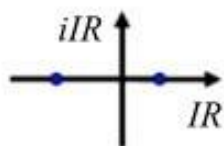
临界点位置的计算与分类

- 临界点为向量场中各个分量均为零的点。可以通过求解向量场方程或通过二分法求得临界点的位置
- 向量场的临界点特性
 - 所有的流线(stream line)汇聚于这些临界点，因此这些临界点被认为是流线的“交点”
 - 在这些临界点处速度光滑的变化为零
 - 临界点的类型决定了向量场的局部格局
- 临界点可以分为5类
 - 鞍点(Saddle)、螺旋点(Spiral)、交点(Node)、聚点(Focus)和中心点(Center, 同心圆临界点)

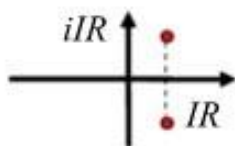


临界点

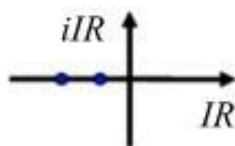
- 对于每一个临界点，可求其向量场的偏导，计算其雅可比矩阵 J ，进而计算 J 的特征值及特征向量，并根据特征值在复平面上的分布对临界点进行分类
- 越稳定的临界点通常意味着更重要的现象，Wang等人量化分析了二维临界点的稳定性，通过计算用来消除局部邻域内的临界点的最小扰动量，来度量该临界点的稳定性，可用于分析稳定和不稳定的特征



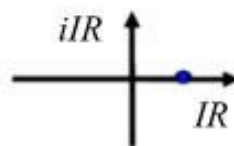
鞍点 (Saddle)



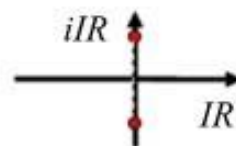
螺旋点 (Spiral)



交点 (Node)



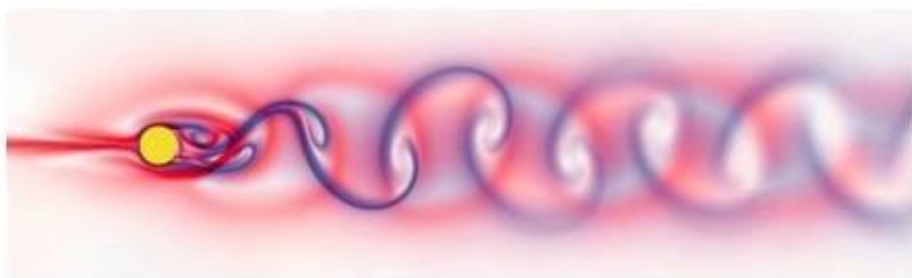
聚点 (Focus)



中心点 (Center)

向量场区域边界的计算

- 为了构造向量场的拓扑结构，需要用积分曲线或曲面连接临界点，并以当前临界点为起始点，沿着临界点处雅可比矩阵特征向量的方向，进行积分曲线的计算
- 向量场的区域边界，也被称为拉格朗日相干结构(Lagrangian Coherent Structures,LCS)[Garth2007]。近年来，研究者通过引入FTLE(Finite-Time Lyapunov Exponents)的概念来计算LCS。FTLE测量相邻的虚拟粒子在向量场中运动轨迹的分离程度，通过对这一测量的分析，进而得到向量场的区域边界，并对该向量场进行可视化



4. 空间张量场数据可视化

张量(Tensor)

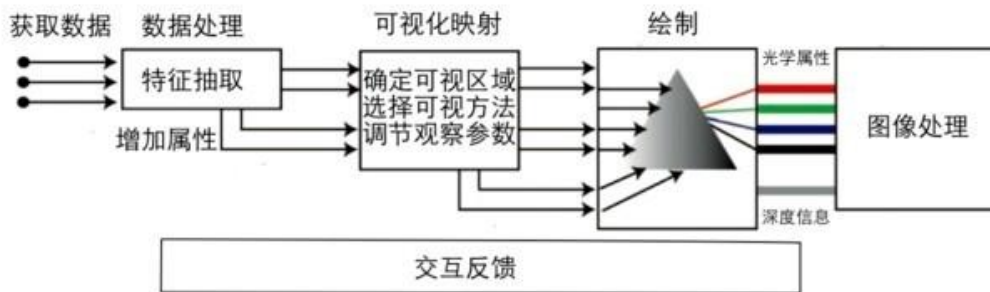
➤ 张量的数学定义是：

- 由若干坐标系改变时满足一定坐标转化关系的有序数组成的集合。张量是矢量的推广：标量可看作0阶张量，向量可看作1阶张量。如果在全部空间或部分空间里的每一点都有一个张量，则该空间确定了一个张量场。每个采样点的数据是一个张量的数据场，称为张量场(Tensor Field)

5 多变量空间数据场可视化

多变量空间数据

- 多变量空间数据在每个节点上包含多个不同类型的数据属性
 - 多个标量场数据、多个标量场混合多个向量场甚至张量场数据。由于数据本身多模态、多维度、多变量的复杂性，多变量可视化比普通的空间数据场可视化更加具有挑战性
- 多变量空间数据可视化的目的抽取和表达数据场中多个属性及其相互关系
- 其难点和挑战在于：
 - 克服多变量、复合类型、内在结构复杂且互相交织的数据特性，设计有效的视觉编码辅助用户同步地分析提取和表达这些信息，观察和研究数据属性及其相互之间的关系，发现未知的新特征和新现象



多变量空间数据场的特征表达与关联分析

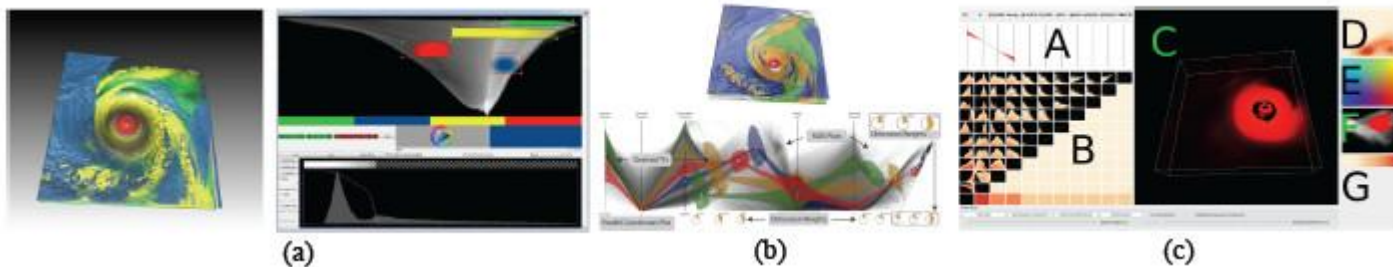
- ▶ 多变量数据场可视化的基本手段是采用特征表达和关联分析方法
 - ▶ 抽取多变量数据场的特征，计算特征的数值、几何、拓扑等信息
 - ▶ 并结合统计分析、信息论、数据挖掘等方法，实现变量、体素、数值、特征及其相互之间的关联分析，探索数据隐含的相似性和差异性。

特征表达与抽取

- 多变量空间数据场的特征表达与抽取的主要任务是构建和处理多变量数据场的特征(几何、拓扑、网格结构等), 提取可视化以及可视表达的对象, 即数据属性和衍生特征、相互之间的关系以及变化趋势等
- 特征指数数据场中包含重要信息的相关结构、现象和区域, 如体数据的物质边界(标量场特征)、临界点与分离线、面(向量场特征)和退化点(张量场特征)等
- 多变量数据场的特征相对抽象和复杂, 特征抽取通过构建几何的、拓扑的、图像的各种表达方式计算特征, 其方法大致可以分为基于多变量属性的交互特征提取、基于拓扑结构的特征抽取和基于数据挖掘的特征抽取
- 在实际应用中, 特征常依赖于特定的应用领域以及用户的需求, 没有普适的、统一的定义和描述, 甚至是未知的
- 基于多变量的交互特征抽取通常使用传输函数和统计分析, 将多变量的原始数据空间变换到统计空间(直方图、散点图等), 用统计学方法表达原始数据或抽取的特征, 实现减少数据量, 同时保留关键信息, 用户可以在统计空间交互选择感兴趣的特征。结合单变量数据分类的方法

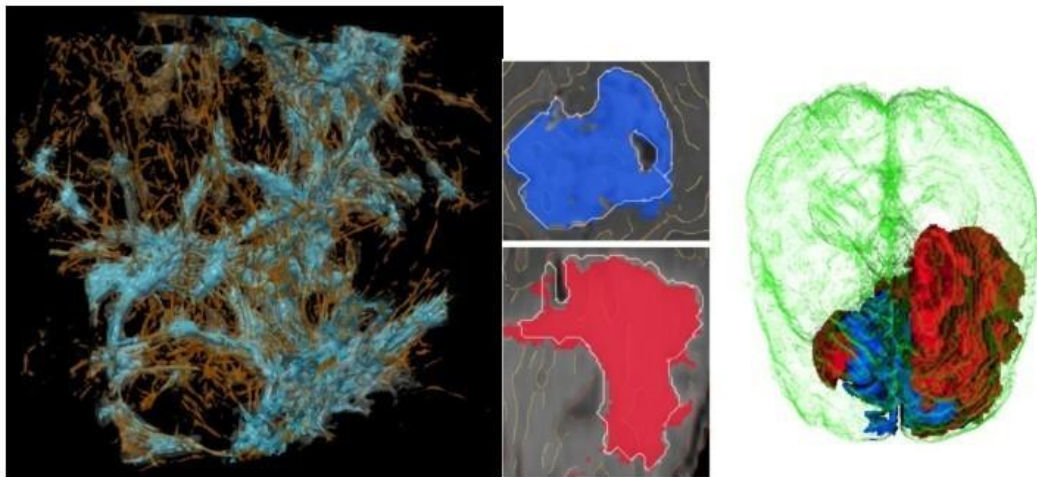
基于多变量的交互特征抽取

- 基于多变量的交互特征抽取通常使用传输函数和统计分析



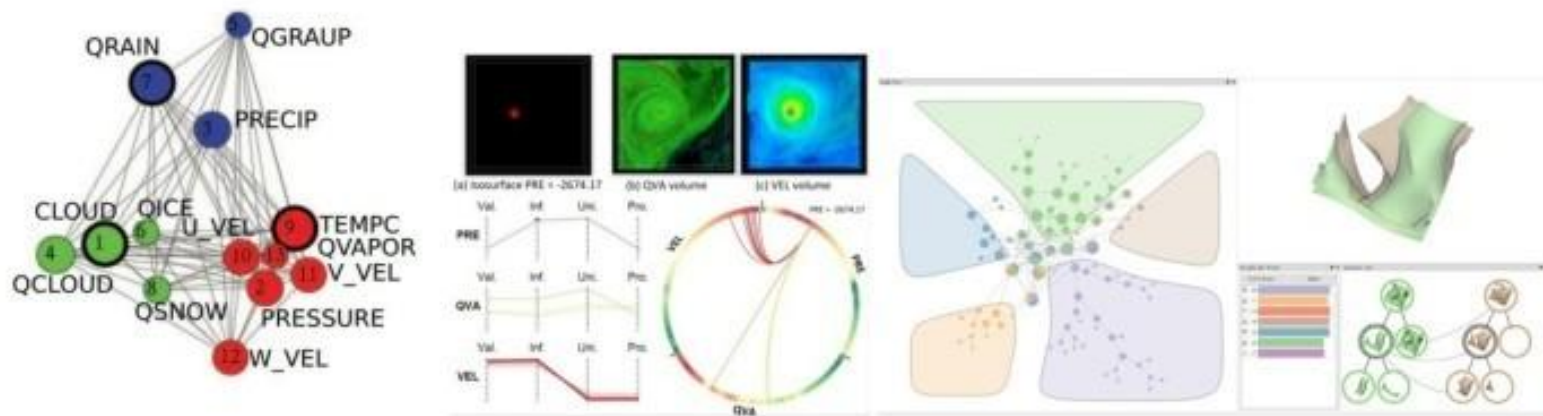
基于拓扑结构的特征抽取

- 基于拓扑结构的特征抽取类似于单变量数据的等值面抽取，轮廓树是一种抽象的拓扑表示方法，能够捕捉标量场中水平集的变化，也用于识别重要的等值面。对多变量数据来说，可以通过不同变量中的等值面构建变量间的关系



关联分析

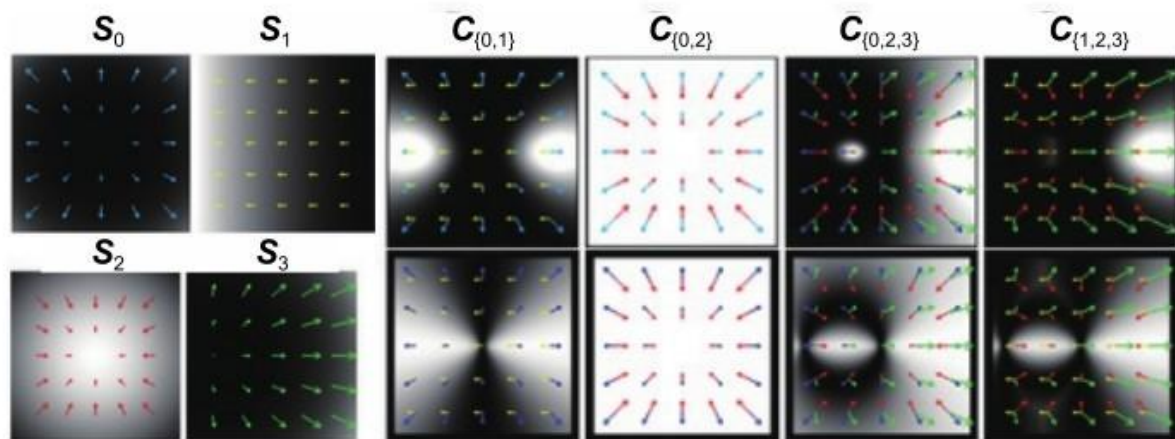
- 多变量数据的交互关系错综复杂，根据分析角度，可以将多变量数据关系划分为五种不同的表现形式：变量间的相关性、体素间的相关性、数值间的相关性、特征间的相关性和混合相关性
- 变量间的相关性
 - 变量对之间的交互关系。相关系数是一种标准的、常见的统计学度量方式，通过比较两个标量函数与其各自均值的偏差是否一致来判断它们是否线性相关。相关系数作为一种全局度量方式，也通常被用来计算两个变量间的局部相关性



多变量数据关系-体素间的相关性

➤ 体素间的相关性

- 衡量数值及其梯度等导出属性之间的相似性。
- Sauber等人 [Sauber2006]提出了相关场的概念，使用梯度相似性度量(Gradient Similarity Measure,GSIM)和局部相关性因子(Local Correlation Coefficient,LCC)测量和计算多标量场数据各属性之间的相关性



多变量数据关系

➤ 数值间的相关性

➤ 分析哪些变量的数值对与特征之间的关系。Haidacher等人对两个变量的所有等值面分布建立距离场，通过度量不同变量的等值面的距离场之间的互信息，计算两个变量的等值之间的相似性，生成两个变量所有标量值之间的相似性矩阵，用来选择和分析不同变量的等值面及相关性

➤ 特征间的相关性是分析不同变量的特征之间的相似性和差异性

➤ 混合相关性

➤ 分析变量、体素、标量和特征之间的关联性。变量间的相关性可以通过累加所有体素间的相关性获得，如体素的梯度相似性

多变量空间数据场的可视化与交互

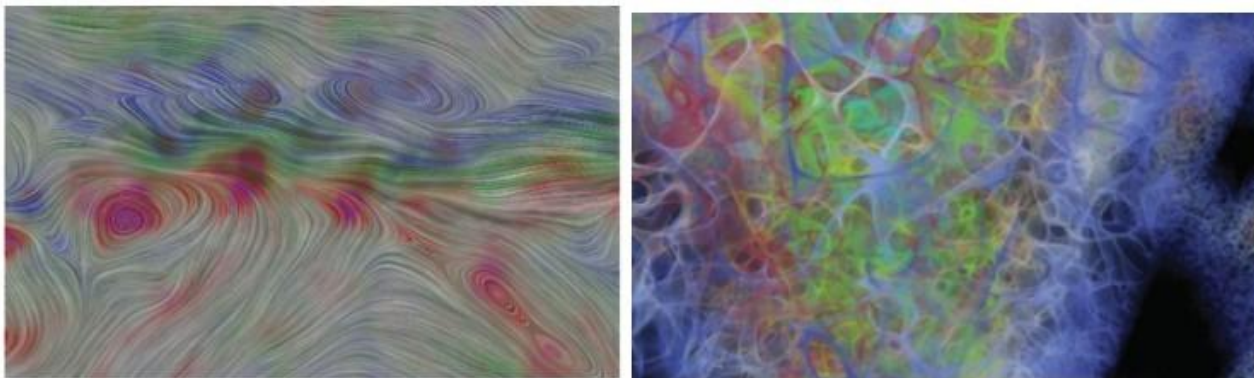
- 多变量空间数据场的可视化主要指基于图形学技术，对多变量数据的绘制着色和最终成像过程
- 交互则指便于用户对数据进行观察和分析的一系列操作，贯穿整个可视化进程。下面将分别进行讨论

视觉通道与融合

- 多变量空间数据场可视化的手段分为两类
 - 用不同的视觉通道映射和编码各个属性及其相互之间的关系和关联
 - 分别可视化各个数据场的数据属性，再进行融合
- 两者都可以归结为将数据属性和特征映射到不同的视觉通道(图标信息、纹理信息、颜色和透明度、轮廓、形状等)，再以合理的方式进行结合，如数据处理阶段的数据融合、绘制阶段的混合绘制、图像处理阶段的图像融合等
- 图标信息和纹理信息作为强大的、可携带和编码丰富信息的表达载体，在多变量可视化技术中最为常见
 - 图标法可清晰地编码多个变量属性，尤其是可表达方向信息，因此图标法被广泛用于向量场和张量场可视化
 - 另一种途径是使用纹理。人类可以感知的与纹理有关的三个视觉通道是模式、大小和方向。研究表明，人类视觉系统对于纹理的上述三个视觉元素的感知和颜色一样，属于一种底层的视觉感受，可以被迅速而精确地感知到

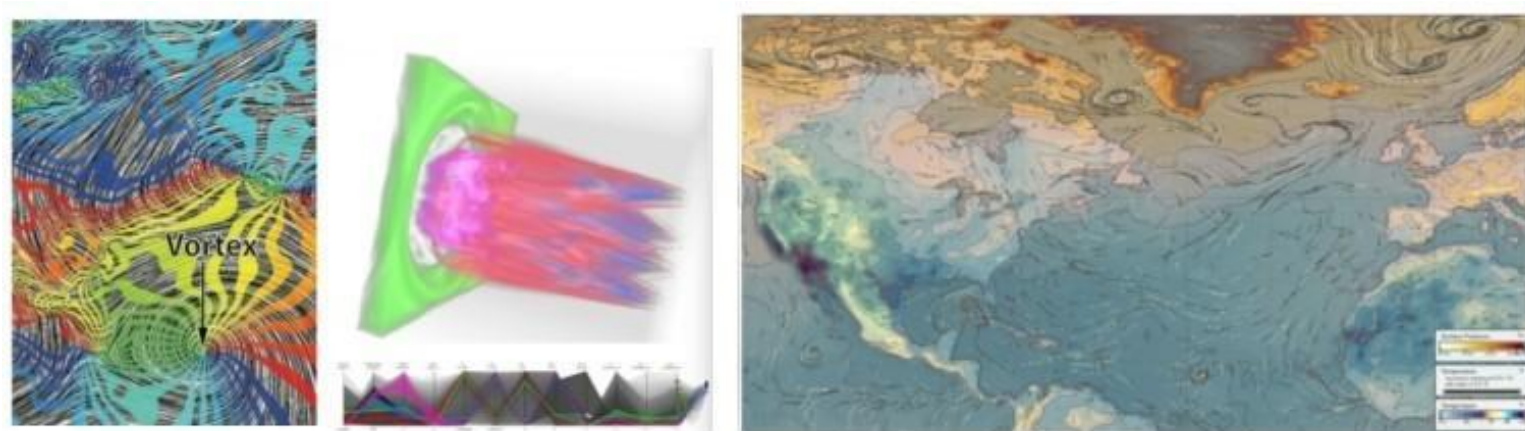
视觉通道与融合

- 使用色织技术混合纹理和颜色编码
- 基于噪声纹理的融合绘制



数据融合和混合绘制

- 数据融合和混合绘制是多变量可视化的另一个重要研究方向
- 基于数据融合和混合绘制的多变量可视化方法直接处理数据属性，可以有机结合统计分析以及特征抽取等数据处理手段进行可视分析，因而研究手段更加丰富



交互探索

- 多变量可视化领域的交互基本上已经形成了以多关联视图为基础、二维视图用于显示和分析属性关系、三维视图用于显示物理数据空间、使用高维传输函数和画刷及其衍生技术为特征选取技术、以焦点+上下文或其变种为数据和特征显示方法的基本框架

时 间	交互目标	
	特征选取	数据属性及关联分析
90-00	/	多关联视图 + 切片 [Wijk1993]
00-05	多维传输函数 [Kniss2001b][Kniss2003]、画刷选取 [Hauser2002]、 焦点 + 上下文 [Doleisch2002][Doleisch2003][Doleisch2005]、人工智 能、机器学习 + 分类 [Tzeng2005]	多关联视图 + 统计分析 [Gresh2000]、多关联视图 [Roberts2000]、 多关联视图 + 散点图 [Pringer2004]
06-12	焦点 + 上下文 [Novotny2006]、机器学习 + 特征抽取、 追踪 [Ma2007]、机器学习 + 画刷选取 [Fuchs2009]，层次 传输函数 [Zhou 2012]	多关联视图 + 多场图 [Saubert2006]、降维 + 平行坐标 [Zhao 2006]、多关联视图 + 平行坐标 + 维度投影 [Gao2012]、多 关联视图 + 算子 [Woodring2006]、多关联视图 + 机器学习 [Blaas2007]、多关联视图 + 平行坐标 [Keefe2009]
13-18	套索 [Zhou2014]、多关联视图 + 聚类 [Wu2015]、视觉特征 + 画刷选取 [Wang2016]、画刷 [Schroeder2016]、散点图矩阵 [Lu2017]	变量相似图 + 散点图 [Biswas2013]、联合轮廓图 [Carr2014]、平 行坐标 + 弦图 [Liu2016]、特征网 [Wang2018]

交互探索

- 多关联视图是多变量可视化交互的基础和重要的手段，在分析变量之间的相互关系以及建立数据空间和各种抽象空间的联系方面，有着无可替代的作用
- 近年来多变量数据越来越复杂，数据分析越来越抽象，研究重心逐渐偏向于对数据属性之间的关系和变化趋势的分析
- 从方法来看，无论是特征选取还是数据属性和关联分析，基于统计分析和机器学习的交互方法在近年的研究工作中所占的比重越来越大并且成为主流