

## 一种面向空间多变量数据聚类的可视分析方法\*

吴斐然, 陈海东, 黄劲, 陈为

(CAD & CG 国家重点实验室(浙江大学), 浙江 杭州 310058)

通讯作者: 陈为, E-mail: chenwei@cad.zju.edu.cn, <http://www.cad.zju.edu.cn/home/chenwei>

**摘要:** 聚类是研究空间多变量数据的重要工具之一,但是自动聚类算法通常需要用户预设参数然后生成结果,缺乏一种有效的交互机制将用户介入到聚类的过程,使之动态改变参数并对结果进行调整和评估.为此提出一种面向空间多变量数据聚类的可视分析流程,首先运用自动聚类算法对原始三维空间进行聚类,针对三维空间不易交互的缺陷将数据点投影到二维平面进行交互选择和可视编码,设置多种视图使用户实时而全面地理解数据分布和模式,交互地修正聚类结果,并根据一些编码的统计信息来判断结果的合理性和正确性.整个流程是渐进式的,即用户通过迭代逐步细化结果,最终抽取兴趣域.案例分析表明,新的可视分析流程能够有效地提高空间自动聚类算法的精度,也极大地缩短了用户交互的时间.

**关键词:** 空间多变量数据场;可视化;交互;空间聚类

中文引用格式: 吴斐然,陈海东,黄劲,陈为.一种面向空间多变量数据聚类的可视分析方法.软件学报,2014,25(Suppl.(2)): 111-118. <http://www.jos.org.cn/1000-9825/14029.htm>

英文引用格式: Wu FR, Chen HD, Huang J, Chen W. Visual analysis approach for clustering multivariate spatial data. Ruan Jian Xue Bao/Journal of Software, 2014, 25(Suppl. (2)): 111-118 (in Chinese). <http://www.jos.org.cn/1000-9825/14029.htm>

### Visual Analysis Approach for Clustering Multivariate Spatial Data

WU Fei-Ran, CHEN Hai-Dong, HUANG Jin, CHEN Wei

(State Key Laboratory of CAD & CG (Zhejiang University), Hangzhou 310058, China)

Corresponding author: CHEN Wei, E-mail: chenwei@cad.zju.edu.cn, <http://www.cad.zju.edu.cn/home/chenwei>

**Abstract:** Clustering is one of important tools to study the multivariate spatial data. However, automatic clustering algorithms require the user to finely modulate parameters, imposing the need for an effective mechanism to manipulate the clustering process by dynamically changing the parameters and evaluating the results. This paper proposes a novel visual analysis approach for clustering multivariate spatial data. First, the underlying dataset is clustered in 3D using an automatic clustering algorithm. Second, the result is examined and refined on its 2D projection by leveraging a suite of visualization and analysis toolkits. The user is allowed to intuitively verify and adjust the clusters by referring to the visual encoding and visual patterns. The entire process is progressively performed in a raw-to-fine fashion. The case study on a high-dimensional symmetric tensor field verifies the effectiveness and robustness of the proposed approach.

**Key words:** multivariable; visualization; interaction; spatial clustering

空间多变量数据场可视化以及特征检测是科学可视化领域的核心问题之一<sup>[1]</sup>,在实际科研领域,如医学图像、生物信息学以及计算模拟等领域中有着重要的意义.针对多变量场的探索和分析,不同的应用场景各有其不同的需求.但在实际操作中,这些需求都可以抽象到寻找数据场的兴趣域(ROI)这一简单命题上,某一 ROI 其包含数据点的多变量数值表现是相似的,且有一定的空间连续性.针对这些特点,多变量数据场的 ROI 探寻可以借助一些已有的手段,其中聚类方法就是重要的工具之一.但是,自动聚类算法通常需要设定参数,同时其生成结果亦没有直观的可视化方案供用户评估分析和调整,这些都限制了自动聚类算法的效力.

\* 基金项目: 国家自然科学基金(81172124, 61232012); 国家高技术研究发展计划(863)(2012AA12090)

收稿时间: 2013-08-05; 定稿时间: 2014-03-13

本文提出一套构建在自动聚类算法之上的可视分析流程,使得用户得以参与自动聚类过程中的参数设定、聚类结果的评价以及调整,并利用 GPU 硬件平台使之适应大规模科学数据的需求.整个过程是渐进性的,用户可以通过逐步的参数选择、聚类标定等方式,由粗到精迭代式地把自身的判断融合进去冗存真的过程当中,这样,一方面通过去除冗余数据点提升了聚类的效率,另一方面构建出一种层次聚类的可视化通道,最终能够达到较好的聚类效果,抽离出 ROI 并加以分析.

本文第 1 节讨论相关工作.第 2 节介绍本文采用的方法和技术.第 3 节展示案例.第 4 节总结本文工作.

## 1 相关工作

### 1.1 变量三维空间数据场的数据分析与表达

在数据分析方面,传统的统计学方法依然有效.其将处理对象从原始数据空间变换到了统计空间(如直方图等)以帮助用户进行多变量数据分类,或是以统计分析从原始数据抽取特征,从而达到减少数据量同时保持关键信息、分析变量间关系的目的.如 Kniss 和 Hansen<sup>[2]</sup>设计的一个适用于多变量可视化的多维传输函数.之后,Kniss 等人<sup>[3]</sup>进一步提出了一种直接适用于复杂的多变量数据的高斯传输函数.

除了上述传统的方式以外,基于图像处理、物理学特征和拓扑逻辑分析的相关方法被广泛应用于流体可视化和向量场特征抽取.Sauber 等人<sup>[4]</sup>提出了相关场的概念,测量和计算多标量场数据各个属性之间的相关性,并设计了多场图(multifield-graphs)用于选择重要的感兴趣的相关场.

此外,数据降维,如自组织映射、多维标度分析、主元分析、等距特征映射和快速映射等在多变量可视化领域也有广泛应用.Guo 等人<sup>[5]</sup>还将基于多维标度的维度投影与平行坐标相结合,这样更有助于用户综合利用可视化手段,从而帮助分析和对比多维数据间的关联.

### 1.2 多变量三维空间数据场的可视化与交互

图标和纹理在多变量可视化技术中最为常见.其中图标法被广泛用于向量场和张量场可视化<sup>[6]</sup>.Ropinski<sup>[7]</sup>详细总结了图标在医学可视化中的应用,并提出应用图标编码多属性信息的 8 条指导方针.Kraus 和 Ertl<sup>[8]</sup>对使用组合图元进行多变量数据可视化进行了研究.在纹理方面,Ware 和 Knight<sup>[9]</sup>认为,人类感知的 3 个主要视觉通道是模式、大小和方向.结合颜色和纹理<sup>[10]</sup>进行多变量数据可视化也是常用的方法.Urness 等人<sup>[11]</sup>提出了色织(color weaving)的概念,克服了传统颜色混合的缺点,强化了颜色编码和纹理的混合应用.

在交互层面,特征选取和数据属性及关联分析是交互的两大根本目标.围绕这些目标出现的交互技术主要有 4 类,即画刷、高维传输函数、焦点+上下文以及多关联视图.在画刷技术方面,Doleisch<sup>[12]</sup>提出了光滑画刷的概念,以此实现数据子集的光滑选取.为了方便用户选取多个属性以便观察和分析数据属性之间的关联,Doleisch<sup>[13]</sup>又提出了一种特征定义语言,通过组合不同属性视图和逻辑操作的画刷以捕获一个特定的特征;在高维传输函数方面,Kniss 等人<sup>[14]</sup>设计了一个适用于多变量可视化的多维传输函数,又使用梯度等衍生信息构建多维直方图<sup>[15]</sup>.随后他们又进一步提出了一种直接适用于复杂的多变量数据的高斯传输函数<sup>[16]</sup>;焦点+上下文是多变量可视化领域常用的显示方法,不同焦点特征的定义和绘制手法可以获得不同的技术组合.Viola<sup>[17]</sup>通过视点的位置以及特征的描述定义数据的重要性;多关联视图是多变量可视化交互的基础而又重要的手段.Gresh 等人<sup>[18]</sup>第一次显式地集成三维物理空间视图和统计表达视图,并在统计表达视图中提供关联画刷.Piringer 等人<sup>[19]</sup>使用多关联视图将二维散点图和三维散点图连接起来,并使用颜色、点大小等深度隐喻帮助用户强化特征认知.

## 2 空间多变量数据聚类优化可视分析

自动聚类算法的优势在于,其利用强大的计算能力为用户快速提供结果.随着数据量的提升,计算效率会极大地受到影响,对参数的掌控和效果的评估亦越来越困难.由于面对的数据不同,一般很难给出明确的优选方案,需要将用户自身的经验以及判断融入整个分析过程的方法,称为可视分析.我们设计了如图 1 所示的可视分

析环境,将数据预处理、空间聚类、低维投影等融入交互的可视界面。

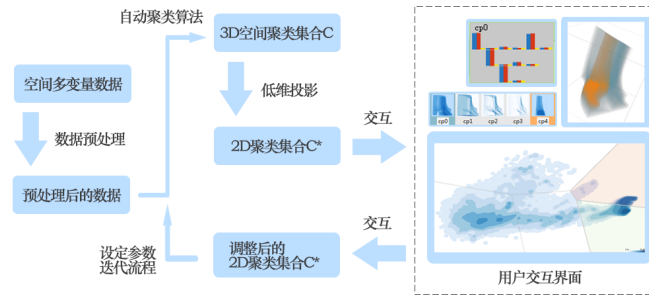


图1 空间多变量数据聚类优化的可视分析流程

## 2.1 数据预处理

由于空间多变量数据的形式各异,可以是多标量、多矢量、多张量乃至这几种的混合.需要通过数据预处理将这些形式统一到多标量,以方便接下来的步骤进行处理.处理不同形式的数据需要用到不同的方法,对矢量和张量数据可以抽取其方向性的表达(如矢量的方向坐标等),或者对其进行建模之后取其中的参数(如张量的矩阵表达,球谱分解重建等<sup>[20]</sup>).对于空间对称场<sup>[21]</sup>等张量场,数据场具有旋转无关性,需要通过特殊的预处理方式将其转化为多标量场,本文第3节将加以详述.

## 2.2 空间聚类

空间聚类算法将空间分成数块,提供了一个聚类的参考.后续步骤将围绕这个预置的聚类结果展开.在空间聚类分割的各种算法中,切割的粒度会影响整个分割的效果,导致过分割或欠分割.如果使切割粒度处于一个合理的水平,一方面切割块既能包含足够的信息又能保证不致出现不当切割,另一方面也能降低计算复杂度<sup>[22]</sup>.因此,通过参数控制分割的度,有助于算法融入整个迭代流程,由粗到精地进行尝试性分割,逐步分解出正确的结果.

设自动方法分割出的聚类集合为  $C = \{C_i, i = 1, 2, \dots, m\}$ , 每个聚类包含数量为  $n_i$  的数据点,下标对应其相应的聚类  $C_i$ .

## 2.3 低维投影

空间聚类得到的聚类结果表现为空间中分割出来的区块集合,而对于用户来说,在3D空间中进行点选以及区域的修正非常困难,在2D平面上对数据进行操作往往是有利于用户交互的更合理的选择.我们采用了低维投影的方式,将空间数据点间的相似程度表达为2D空间中的欧式距离.传统的全局投影技术,如MDS<sup>[23]</sup>等需要维持全部数据点的关系矩阵,无法应用于规模较大的数据场景,本文采用局部仿射多变量投影方法(LAMP)<sup>[24]</sup>通过并行加速其效率可达到供实时交互的程度.

运用局部投影方法需要设定控制点集  $P$ .根据原空间中的聚类集合  $C$  生成控制点,得到  $P = \{p_i(\bar{v}_i, \bar{x}_i), i = 1, 2, \dots, m\}$ , 其中的多变量参数由  $C_i$  所含数据点多变量均值  $\bar{v}_i$  和  $C_i$  的重心坐标  $\bar{x}_i$  确定.定义  $j, k$  两控制点间的不相似度为  $d_{p_j, p_k} = \left\| \bar{v}_j, w\bar{x}_j \right\| - \left\| \bar{v}_k, w\bar{x}_k \right\|$ ,  $j, k = 1, 2, \dots, m$ , 其中,  $w$  为控制重心位置差异在不相似度中所占比例的权重,根据构建出的控制点不相似度矩阵,可以将控制点利用MDS等方法从原空间先行投影到2D空间,剩余数据点再根据控制点在投影空间中的位置低维嵌入.

由于处理的数据量较大,数据点分布稠密,在视觉效果上会出现层叠粘连的情况,且大量的点聚集在狭小的屏幕空间内,用户视觉上对其位置精度的要求并不高.选择高斯核密度估计<sup>[25]</sup>,用热力图的方式表示出投影点分布的密度情况并进行颜色编码,取得了更多的点分布信息和更好的视觉效果(如图2所示).



(a) 未经核密度估计处理过的投影点 (b) 经过高斯核密度估计后的结果

图 2

## 2.4 交互式聚类调整

考虑到 2D 投影空间内点与点间的欧式距离体现了实际空间中点与点多变量数值和空间距离的差异,选用 Fortune 算法生成 2D Voronoi 图<sup>[26]</sup>将投影空间进行切割,用  $p_i$  的位置作为图的生成点,在 2D 平面上重新定义原空间的聚类  $C^* = \{C_i^*, i = 1, 2, \dots, m\}$ . 由于 Voronoi 图中每个区块  $C_i^*$  内的点到其生成点  $p_i$  的距离较其到其他生成点距离最小,而投影方式又决定了在原空间属于相同  $C_i$  的数据点在 2D 平面上的位置会趋近控制点  $p_i$ ,因此  $C_i^*$  和  $C_i$  之间存在对应和关联关系,而  $C^*$  和  $C$  之间的差异则会作为之后聚类调整的重要参考.

### 2.4.1 可视分析平台

如图 3 所示,其中,a 代表 3D 视图;b 代表 2D 视图;c 代表预览视图;d 代表变量数值列表.集成的可视分析平台综合不同的视觉通道编码方式,使用户可视地评价每个  $C_i^*$  的合理性,进而调整其边界,确定需要再挖掘或排除的区块集合.界面主要包括 4 种视图:3D 视图、2D 视图、预览切片视图和变量数值列表.每个视图的功能和含义如下.

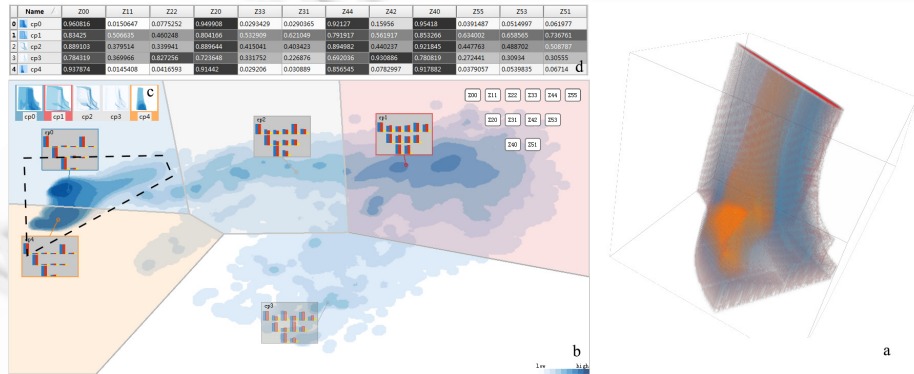


图 3 可视分析界面

3D 视图.对于 3D 空间的聚类调整,其最终调整的效果仍然需要在 3D 空间中加以展示.直接体绘制方法可不抽取中间结构,适合于交互地浏览空间数据场的内部结构.当用户选择  $C_i^*$  的子集时,子集所包含的点在原空间中位置所在体素被标定为代表  $C_i^*$  的特定数值,以区别选中的数个  $C_i^*$ .令这些值处在体素值域的均分点上,即属于  $C_i^*$  的数据点在原空间中的体素值为

$$s_i = i \frac{S}{m^+ + 1}, i = 1, 2, \dots, m^+ \quad (1)$$

其中, $S$  为体素值域, $m^+$ 为当前选中  $C^*$  的子集中成员数量.体绘制的颜色由一个一维传输函数控制,每个  $C_i^*$  都用其专属的颜色标识.3D 视图、2D 视图以及预览视图都采用这套颜色方案用以区分  $C_i^*$ .

2D 视图.这是用户进行交互活动的主要区域.2D 视图主要由核密度估计投影、2D 区块以及可视编码图标构成.投影可以表达出数据点的相似度,每个区块以一个多边形凸包表达,其边界可调,选中时会以颜色高亮.可视编码图标编码了每个区块的统计信息,如图 3 中的 b 所示,图标在投影空间内的默认位置为控制点所在坐标,上方注明其代表区块的名称.每个变量有其对应的图元,每个图元的位置可由用户自定义,用颜色区别了原空间

$C_i$  数据点多变量均值  $\bar{v}_i$ ,  $C_i^*$  包含数据点多变量的统计信息,如均值  $\bar{v}_i^*$  和标准差  $\sigma_i^*$ ,形状大小则编码了这几个量的数值,如图 4 所示.其中,a 给出图标上每个图元位置对应的属性名称,图元位置可由用户调整;b 给出可视编码图标、外框及区域内以用户设定的颜色编码;c 给出图标名称;d 给出图标对应  $C_i^*$  的控制点位置.e~g 描述了图元编码方案:以柱状图的方式,左侧色块 e 给出  $C_i$  在该变量上的均值;中间色块 f 给出  $C_i^*$  在该变量上的均值;g 给出  $C_i^*$  在该变量上的标准差.如此,用户可通过视觉感知直观地判断分割区域所处的状态,如当某个图元上左侧色块和中间色块几乎平齐且右侧色块较小时,区域包含点均值和控制点数值接近,且标准差很小,可以判定此图标所处的  $C_i^*$  和  $C_i$  较为一致.为适应不同数据的需要,图元的排布方案可视需求预设,有利于变量间对比和筛选分析.

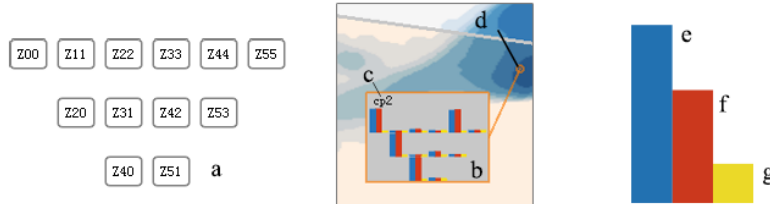


图 4 可视编码图标及图元编码的统计信息

预览视图.  $C_i^*$  包含的点会在 3D 空间呈现相应的结构,但在 3D 空间逐个检视这些结构的效率较为低下.若能够提供这些结构在一个视角的预览,则能大大缩小检视的范围.由此延  $z$  轴方向累计体素个数,用颜色编码,并标识区块名,获得了如图 3 中 c 所表示的预览图像集合.用户从而能够快速得知区块包含点的数量以及实际在 3D 空间的分布规律等信息.

变量数值列表.提供传统的列表查看方式,用颜色编码了数值的大小.

整个分析平台对同一  $C_i^*$  进行了不同的可视化处理,同样的  $C_i^*$  在各个视图中会用同样的颜色编码.如图 3 所示,标定的区域在 3D 视图、预览视图以及 2D 视图中均以同种颜色高亮以示关联,用户在某一视图上对  $C_i^*$  的操作可以反映在各个视图上.

平台运行硬件环境为: Intel Core i5-3470 3.20GHz 3.60GHz 双核 CPU, 8GB 内存, NVIDIA GeForce GTX 650 显卡.利用 GPU 对投影等过程进行了并行处理及优化,可以保证交互的实时性以及数据处理的效率.

#### 2.4.2 可视分析交互流程

整个分析交互流程的入口可以是预览视图,也可以是 2D 视图上的图标.通过这两者传达出的信息判断每一个  $C_i^*$  上的数值分布以及  $C_i^*$  和  $C_i$  的一致性.随后,用户选中  $C_i^*$  并调整传输函数以使其高亮,在 3D 空间中查看或者根据图标信息判断其中是否有 ROI.若其存在于一个  $C_i^*$  上,用户可根据需要调整  $C_i^*$  的边界将其细化并在 3D 视图中予以展现;若 ROI 分散在几个  $C_i^*$  上,用户也可以将它们同时选中调整.对于选中的  $C_i^*$ ,用户也可以对其包含数据点设定参数进行更细化的聚类,之后再先前所述步骤迭代式地发掘,直至找到需要的 ROI(如图 5 所示).详细的分析过程会在第 3 节中加以介绍.

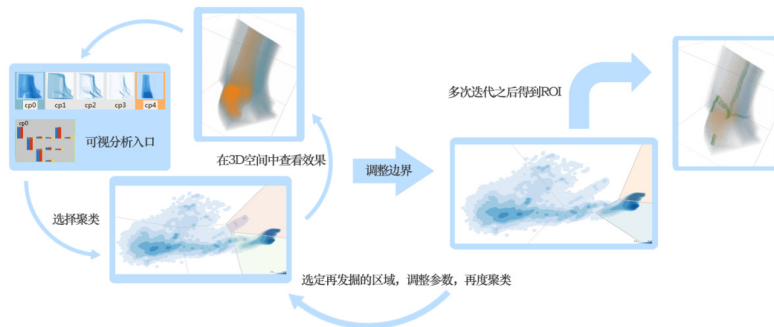


图 5 面向空间多变量数据聚类的可视分析交互流程图解

### 3 案例分析

空间对称场<sup>[21]</sup>是三维空间参数化、多面体网格生成、纹理合成等三维应用中生成的特征场.例如正六面体对称场,其在空间中的每一采样点都包含 6 条两两垂直且模为单位值的矢量.分析和理解这类空间对称场数据对于优化各类应用的设计、探索数据场的特征区域至关重要.已有的可视化方法只着眼于奇异线的表达,对其缺乏分析,且对于数据其他部分也没有挖掘.

数据预处理.在预处理步骤中,用  $n \times n \times n$  的规则网格重采样输入数据.由于对称场的旋转无关性,无法用一般方法描述每个采样点上数据的特征.我们设计了一种特殊局部采样方法,如图 6 所示,其中,(a)表示  $m \times m \times m$  区域;(b)表示取某采样点和中心连线中角度最小的矢量,并取其在连线上的模作为该点的标量采样;(c)表示所有取到的矢量,以黑色箭头标识;(d)表示所有取到的矢量在中心连线上的模.以遍历网格中  $m \times m \times m, m \ll n$  的小区域.在该区域的采样点上,选取与区域中心连线夹角最小的矢量,并取其连线上的模,构建一个局部的标量场,对其进行 Zernike 分解,生成数个旋转无关的特征参数  $Z_n$ <sup>[27,28]</sup>.整个张量数据场转变成标量场,可根据参数和其对应算子的性质描述和解释当前区域对称场的性状.选用数据为空间六面体对称场,设  $n=160, m=4$ ,经过预处理步骤得到一个  $160 \times 160 \times 160$  的空间 12 标量数据场.

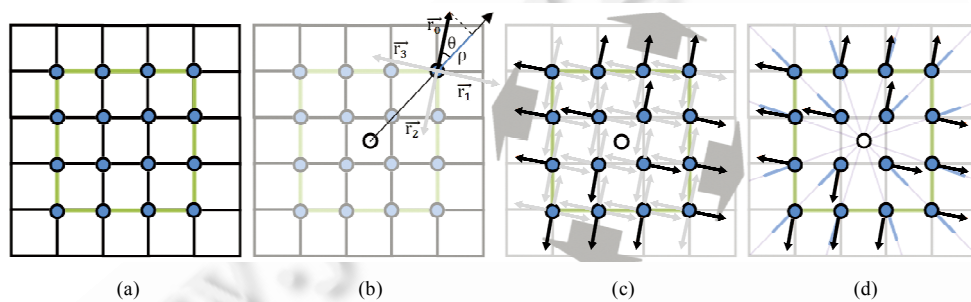


图 6 空间对称场预处理步骤图示

聚类优化分析.本案例中我们选用可调参数的统计区域融合体分割(SRM)<sup>[29]</sup>作为空间自动聚类方法.其基于区域生长,并运用统计原理进行区域融合,使得每个分割区域在空间上内部高度连贯,很好地保持了区域的空间属性.分割的粒度用参数  $Q$  来定义,它控制了分割结果的复杂度,其值越高,分割结果越细致.

首先设  $Q=16$ ,可得图 3 所示的投影分布,cp0,cp4 以及 cp1 内都有较多的点.通过观察预览视图,以及实际 3D 视图查证,cp0 和 cp4 包含区域处在整个模型的内部,而 cp1 则处在外部.观察图标上图元的状态,可见 cp0 和 cp4 的数值分布很相似,呈现一定的模式,且各参数数值分布体现出  $C_i^*$  和  $C_i$  较为一致;而 cp1 的数值分布与 cp0 和 cp4 差别较大,且各参数数值分布标准差更大,判断 cp4 和 cp0 这两个聚类中会有需要挖掘的 ROI,且其可能集中在投影点较密集的图 3 所示中黑色虚线框内.

调整 cp0 和 cp4 的边界以框选中黑色虚线框内投影点,设  $Q=32$ ,对其进行更细化的聚类操作.多次迭代将离群的数据点去除以后,得到图 7 所示的结果,观察得知,所有区域的变量数值分布情况较为接近,其中数值较大的变量是 Z00,Z20,Z40 和 Z44,对 3D Zernike 描述算子的性质分析可知,Z44 对正六面体对称场较为敏感,而对放射状的场较敏感的变量为 Z00,Z20 和 Z40,从实际情况来看,和上述的分析相一致.而继续对比图元的细节可以发现,尽管大体上 5 个聚类的变量数值分布相似,cp25 中  $l$  为奇数的变量  $Z_n$ ,如 Z33,Z53,较 cp24 和 cp26 要大,说明该部分的对称场体现出特殊的旋转倾向,可能有奇异线存在.

通过自动聚类、聚类调整、选定调整后的聚类并在其中继续细化聚类的迭代循环,完成对聚类结果的优化调整.结果显示,提取到的奇异线区域(图 7 中 cp25 所示标定区域)和用户标注相一致,而且通过编码信息能够解释这些区域的数据特性.

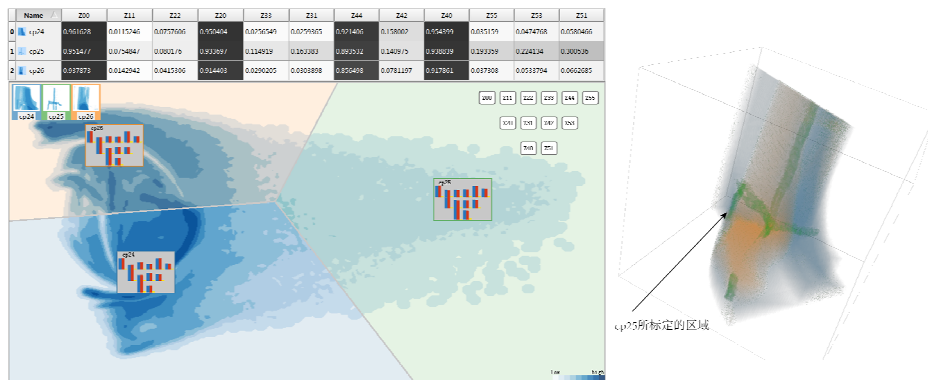


图7 多次迭代调整后的聚类结果

## 4 总结

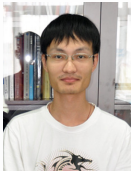
本文设计了一种面向空间多变量数据聚类的可视分析流程,主要思想是将用户本身的分析和决策能力有机地融入到自动算法之中,将参数等纯数字化的输入条件化作用户可以理解的可视化语言,以人之判断弥补自动算法的不足.在这个流程中,综合运用了降维、投影、可视编码、可视交互等多种手段,为用户提供了审视每一步操作结果的全方位视角,不仅可以迅速地对自动算法进行研判,也能灵活地根据自身所见进行适当调整.同时整个过程也遵循了层次式聚类的思想,逐步地将聚类细化,直至挖掘出用户需要的部分.

在将来的工作中,这个流程需要适应更大规模的数据.预期改进的流程不仅能对一种聚类算法的结果进行优化,也能采用多种聚类算法,综合比较其结果进行优化.

## References:

- [1] Johnson C. Top scientific visualization research problems. *IEEE Computer Graphics and Applications*, 2004,24(4):13–17.
- [2] Kniss J, Kindlmann G, Hansen C. Interactive volume rendering using multi-dimensional transfer functions and direct manipulation widgets. In: *Proc. of the IEEE Visualization*. 2001. 255–262.
- [3] Kniss J, Pemoze S, Ikits M, Lefohn A, Hansen C, Praun E. Gaussian transfer functions for multi-field volume visualization. In: *Proc. of the IEEE Visualization*. 2003. 497–504.
- [4] Sauber N, Theisel H, Seidel HP. Multifield-Graphs: An approach to visualizing correlations in multifield scalar data. *IEEE Trans. on Visualization and Computer Graphics*, 2006,12(5):917–924.
- [5] Guo H, Xiao H, Yuan X. Scalable multivariate volume visualization and analysis based on dimension projection and parallel coordinates. *IEEE Trans. on Visualization and Computer Graphics*, 2012,18(9):1397–1410.
- [6] Bueger R, Hauser H. Visualization of multivariate scientific data. *Computer Graphics Forum*, 2009,28(6):1670–1690.
- [7] Ropinski T, Preim B. Taxonomy and usage guidelines for glyph-based medical visualization. In: *Proc. of the Simulation and Visualization (SimVis)*. SCS Publishing House, 2008. 121–138.
- [8] Kraus M, Ertl T. Interactive data exploration with customized glyphs. In: *Proc. of the WSCG*. 2001.
- [9] Ware C, Knight W. Using visual texture for information display. *ACM Trans. on Graphics*, 1995,14(1):3–20.
- [10] Shenan HH, Interrante V. Compositing colour with texture for multi-variate visualization. In: *Proc. of the 3rd Int'l Conf. on Computer Graphics and Interactive Techniques in Australasia and South East Asia*. New York: ACM, 2005. 443–446.
- [11] Urness T, Interrante V, Longmire E, Marusic I, Ganapathis Ubramani B. Effectively visualizing multivalued flow data using color and texture. In: *Proc. of the IEEE Visualization*. Washington: IEEE Computer Society, 2003. 151–121.
- [12] Doleisch H, Hauser H. Smooth brushing for focus+context visualization of simulation data in 3D. *Journal of WSCG*, 2001,10(1): 147–154.
- [13] Doleisch H, Gasser M, Hauser H. Interactive feature specification for focus+context visualization of complex simulation data. In: *Proc. of the Symp. on Data Visualization*. Aire-la-Ville: Eurographics Association, 2003. 239–248.

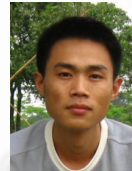
- [14] Kniss J, Kindlmann G, Hansen C. Interactive volume rendering using multi-dimensional transfer functions and direct manipulation widgets. In: Proc. of the IEEE Visualization. Washington: IEEE Computer Society, 2001. 255–262.
- [15] Kniss J, Hansen C, Grenier M, Robinson T. Volume rendering multivariate data to visualize meteorological simulations: A case study. In: Proc. of the 4th Joint IEEE TCVG—EUROGRAPHICS Symp. on Visualization. Aire-la-Ville: Eurographics Association, 2002. 189–194.
- [16] Kniss J, Premoze S, Ikits M, Lefohna, Hansen C, Praun E. Gaussian transfer functions for multi-field volume visualization. In: Proc. of the IEEE Visualization. Washington: IEEE Computer Society, 2003. 65–73.
- [17] Viola I, Kanitsar A, Groller ME. Importance-Driven feature enhancement in volume visualization. IEEE Trans. on Visualization and Computer Graphics, 2005,11(4):408–418.
- [18] Gresh D, Rogowitz B, Winslow R, Scollan D, Yung C. WEAVE: A system for visually linking 3D and statistical visualizations applied to cardiac simulation and measurement data. In: Proc. of the IEEE Visualization. Los Alamitos: IEEE Computer Society, 2000. 489–492.
- [19] Piringer H, Kosara R, Hauser H. Interactive focus+context visualization with linked 2D/3D scatterplots. In: Proc. of the 2nd Int'l Conf. on Coordinated & Multiple Views in Exploratory Visualization. Washington: IEEE Computer Society, 2004. 49–60.
- [20] Hess CP, Mukherjee P, Han ET, *et al.* Q-Ball reconstruction of multimodal fiber orientations using the spherical harmonic basis. Magnetic Resonance in Medicine, 2006,56(1):104–117.
- [21] Huang J, Tong Y, Wei H, *et al.* Boundary aligned smooth 3D cross-frame field. ACM Trans. on Graphics (TOG), 2011,30(6):143.
- [22] Zitnick C L, Kang SB. Stereo for image-based rendering using image over-segmentation. Int'l Journal of Computer Vision, 2007, 75(1):49–65.
- [23] Kruskal JB. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika, 1964,29(1):1–27.
- [24] Joia P, Paulovich FV, Coimbra D, *et al.* Local affine multidimensional projection. IEEE Trans. on Visualization and Computer Graphics, 2011,17(12):2563–2571.
- [25] Silverman BW. Density Estimation for Statistics and Data Analysis. CRC Press, 1986.
- [26] Fortune S. A sweepline algorithm for voronoi diagrams. Algorithmica, 1987,2:153–174.
- [27] Khotanadz A, Hong YH. Invariant image recognition by Zernike moments. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1990,12(5):489–497.
- [28] Novotni M, Klein R. 3D Zernike descriptors for content based shape retrieval. In: Proc. of the 8th ACM Symp. on Solid Modeling and Applications. ACM, 2003. 216–225.
- [29] Nock R, Nielsen F. Statistical region merging. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2004,26(11):1452–1458.



吴斐然(1986—),男,浙江杭州人,博士,主要研究领域为信息可视化,可视分析。  
E-mail: tomwfr007@gmail.com



陈海东(1987—),男,博士,主要研究领域为科学可视化,可视分析。  
E-mail: chenhd925@gmail.com



黄劲(1978—),男,博士,副教授,CCF 会员,主要研究领域为计算机图形学,几何处理,物理仿真。  
E-mail: hj@cad.zju.edu.cn



陈为(1976—),男,博士,教授,CCF 高级会员,主要研究领域为科学可视化,信息可视化,可视分析。  
E-mail: chenwei@cad.zju.edu.cn