# Visualization of Flow Field based on Centroidal Voronoi Tessellation

Michael Shell, *Member, IEEE,* John Doe, *Fellow, OSA,* and Jane Doe, *Life Fellow, IEEE*

**Abstract**—重心沃罗努瓦分布（Centroidal Voronoi tessellation，CVT）是沃罗努瓦分布的一种，其要求每个沃罗努瓦区域（Voronoi region）的生成元恰好为该沃罗努瓦区域的重心。这样的沃罗努瓦分布相较一般的沃罗努瓦分布更能突出的表现流场的一些特征。而科学计算可视化中正要求对目标的特征的进行图形化。所以将重心沃罗努瓦分布与流场可视化结合起来，可以用少量的点和箭头表达大量的流场信息，生成人眼更易接受的流场图。

本文将介绍重心沃罗努瓦分布的基本信息与常用算法，并将其具体到科学计算可视化中，进而尝试连续多时间层的流场可视化。

A Centroidal Voronoi Tessellation (CVT) is a Voronoi Tessellation in which the generators are the centroids for each Voronoi region.It can convey us some characteristics while the common Voronoi Tessellation can't.Meanwhile,Visualization in Scientific Computing(ViSC) just requires us visualizing the representative characteristics of the field.In this case, it can be done with a few points or arrows to express a lot of information,by combining the CVT with the flow field visualization, and generating the image that can be more accessible to people.

In this paper,it is mentioned that what is ViSC and CVTs.Then focusing on how to generate a CVT and how to use CVT to generate a flow field visualization,in only one time point or more.

**Index Terms**—科学可视化（Scientific Visualization），科学计算可视化（Visualization in Scientific Computing），重心沃罗努瓦分布（Centroidal Voronoi tessellation），流场（Flow Field）

## 1 INTRODUTION

labelsec:introduction

科学计算可视化，是指将科学计算、实验获得的数据，转化为人类可以直接识别的的图像，图形的行为。因为人们对大量的数据很难直接理解与分析，同时人类获得的信息中，视觉信息占着绝大多数。所以我们通过可视化的方法，对数据进行一步转化，使的人们可以直观的看出某些数据的某些特点，为后续的精确分析、处理提供方向上的指引和规划。当然，因为人眼本身的的性质，我们无法通过图像获得精确的信息，所以可视化无法作为分析处理的一步，也始终没有产生对可视化效果优劣的明确评价标准；但也正因为如此，我们才需要不断改进可视化方法，使得其能够更加突出的表现目标的某些特征（突出到人眼能够区分识别）。这正是本文的目标，并且我们将这个指标作为定性的评价可视化优劣的标准。通过重心沃罗努瓦分布生成的图像，恰能明确反映目标流场的特征的，成为本文实验与改进目标。

本章将主要介绍可视化的发展历史与现状，从而更加明晰我们究竟为什么要作做可视化以及如何进行科学计算可视化可视化。

Visualization in Scientific Computing, witch refers to the action of converting from the scientific computing and experimental data to the images that people can understand.Since we human can't directly comprehend and analysis big amount of data, it is necessary to convert the data before.And because visual information accounted for the vast majority of all the information which human obtains, it will be more conveniently and successfully to convert the data to the image helping people comprehend the data briefly.So that scientists can get the general impression of the data which leads the direction for further research. However,due to the structure of human eye,it is hardly to distinguish the difference of images in detail.So that simply rendering every data to generate an image may cover the details which sometimes can be the information we focusing on.At this point,it is better to extract the characteristics of the data and then only visualizing the characteristic information.That is the reason why we steadily developing the algorithm of ViSC,both on the higher computing speed and better characteristic extraction.

### 1.1 可视化的历史背景

可视化（Visualization），顾名思义，就是将非图形化的事物转换成图形化的事物，将非视觉信息转化为视觉信息，

- *M. Shell was with the Department of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, 30332. E-mail: see http://www.michaelshell.org/contact.html*
- *J. Doe and J. Doe are with Anonymous University.*

*Manuscript received April 19, 2005; revised August 26, 2015.*

亦可翻译为图形化，图示化。　　事实上，最早的可视化可以可以追寻到很古老。举个简单的例子，从抽象的数字“2”的到符号“$ii$”的转化，就是一种可视化。尽管其转化的符号也是一个抽象的符号，但是他完成了不可视到可视的转化，使得互相的信息的交流成为可能。 现代对可视化的分类，主要将可视化按照其处理的对象分成3类：科学计算可视化：将科学计算获得的数据藉由计算机图形学和图像处理技术的方法将其传化为图形或图像进行显示；信息可视化：将已知的信息，如每天股市的交易情况，转化为可以直观识别的图形；知识可视化：将人类的知识转化为图形从而便于交流，并帮助他人获得正确的知识。

作为学科术语,“可视化”一词正式出现在1987 年2月在美国国家科学基金会(NationalScienceFounda2tion,NSF) 召开的图形图像专题研讨会上。并在此会议上正式给出了科学计算可视化（Visualization in Scientific Computing，ViSC）的定义、涵盖的领域。从此，科学计算可视化成为了一门正式的专门学科。

但在计算机用于科学计算的早期,由于计算机硬件条件的限制,科学计算只能以批处理方式进行。由于没有交互功能,科学家不能对计算过程进行干预和引导,只有被动地等待计算结果的输出,且大量的输出结果也是用人工方式处理。最好的辅助方法是用绘图仪以二维的方式展示结果。这样以来,大量信息被处理方法丢失,科学家得不到计算结果的直观、形象的整体印象,处理结果数据的任务繁杂,花费时间往往是计算时间的十几倍到几十倍。随着计算机处理图形、图像功能的不断提高,科学计算可视化成为了可能。

## 1.2 科学计算可视化的现状

ViSC被定义为：用计算机图形学和图形处理技术,将科学计算过程中产生的数据及计算结果转换为图形或图像在屏幕上显示出来,并进行交互处理的理论、方法和技术。它涉及计算机图形学、图像处理、计算机视觉、计算机辅助设计及图形用户界面等多个研究领域。

科学计算可视化的目的是理解自然的本质。要达到这个目的,科学家把科学数据,包括测量获得的数值或是计算中涉及、产生的数字信息变为直观的、以图形图像形式表示的、随时间和空间变化的物理现象或物理量呈现在研究者面前,使他们能够观察、模拟和计算。这个过程可细化为以下四个步骤:a. 过滤,预处理原始数据、转换数据形式、滤掉噪声、抽取感兴趣的数据等;b. 映射,将过滤后的数据映射为几何元素,常见有点、线、面图元、三维体图元等;c.绘制,几何元素绘制,得到结果图像;d.反馈,显示图像,并分析得到的可视结果。ViSC技术的意义重大,它大大加快了数据的处理速度,使每日每时都在产生的庞大数据得到有效的利用;实现人与人、人与机器之间的图像通讯,增强了人们观察事物规律的能力;使科学家在得到计算结果的同时,知道在计算过程中发生了什么现象,并可改变参数,观察其影响,对计算过程实现引导和控制。

随着可视化技术应用领域不断扩大,人们对其需求越来越复杂,可视化概念扩展到工程数据和测量数据时,学术界把这种空间数据场的可视化称为体可视化(Volume visualization)技术。一般不将它看成是可视化的一个分支,而作 为科学计算可视化的一个方向处理。 体可视化技术研究如何表示、维护和绘制体数据集,从而观察数据内部结构和理解事物复杂特性。它是一种从体数据集中抽取内在的、本质的信息,并以图形图像方式交互地表现出来的技术。体数据集由三维笛卡尔网格体元(网格点上的数据值)构成,将网格体元存储在三维数组(体缓冲区)中。体数据集可来源于许多领域,这决定了体可视化的应用范围。如:工程建筑的空间场,气象卫星测量的空间场,航空航天实验、核爆炸模拟等对应的速度场和温度场,地震预报的力场,工业产品超声波探测的零件密度场,核磁共震产生的人体器官密度场等。

十数年前,由于计算机的速度、显示器的分辨率等原因,使得体可视化在数据量即使较小的情况下也要花费很长的时间。解决这个问题来源于两个方面的努力:一是新型算法的研制成功,如表面拟合法(Surfacefitting)和体绘制(Volumerendering)。表面拟合法从数据场中提取等值面后对其进行展示, 其显示速度快。缺点是由于只显示指定等值面非常容易丢失数据体中的关键信息。体绘制算法无需几何元素(点、线、面等)作中介表示,而是直接将整个数据场投影到屏幕空间。由于每次显示均要涉及整个数据集,因此它的显示速度比表面拟合法要慢,但它的显示质量很高且不丢失任何信息。二是硬件技术和并行计算技术的发展。并行体可视化既提供快速绘制的可能,又能够处理大的分布式数据。一个体可视化系统一般具有下列功能:a.高质量的数据表示方法;b.处理大规模的数据集合,且分辨率较高;c. 快速的图形、图像绘制算法;d.支持交互式操作,使用户参与控制。

流场可视化是科学计算可视化的另一个分支方向,同时是流体力学的重要部分。正是流场可视化的研究使得计算流力学(CFD)快速发展。计算流体力学中的数据类型主要有向量场和张量场。流场可视化用箭头、流线和粒子跟踪技术研究二维流场。对三维流场,人们还没有找到一个普遍接受的、直观的、符合视觉习惯的可视化方法。主要通过以下几种方法进行研究:a.改造体可视化技术,将其拓展,使其可以用于流场可视化;b. 流面技术,用流面加梯度线索的方法表示流场;c.图标技术;d.基于特征的可视化,这是从较高层次进行的流场可视化。

## 1.3 本文的主要内容

本文主要研究基于重心沃罗努瓦分布生成的二维流场可视化，观察其对流场特征的表现效果，并试着对其细节进行改进。

本文的主要结构为：

第一章：可视化背景介绍。

第二章：重心沃罗努瓦分布算法介绍。

## 2 重心沃罗努瓦分布算法

### 2.1 重心沃罗努瓦分布

首先我们给出重心沃罗努瓦分布的严格定义：

对一个开集$\Omega \subseteq \mathbf{R}^n$, 记$d(x_1, x_2) \in \mathbf{R}$为$\mathbf{R}^n$中任意两个点$x_1, x_2$ 的距离，一系列点集$\{V_i\}_{i=1}^k$ 被称为一个沃罗努瓦分布，如果满足:①对$V_i$要求：$V_i \cap V_j = \{0\} \iff i \neq j, \bigcup_{i=1}^k \overline{V_i} = \overline{\Omega}$ 其中$\overline{V_i}, \overline{\Omega}$ 为$V_i, \Omega$的闭包；
②存在$mathbf R^n$中一族互不重合的点$\{v_i\}_i = 1^k$，满足[1]

$$V_i = \{x \in \Omega | d(v_i, x) < d(v_j, x), j = 1, \ldots, i-1, i+1, \ldots, k\}$$

即$\{V_i\}_{i=1}^k$是一个沃罗努瓦分布。 ③同时$\{v_i\}$也恰为$\{V_i\}$的重心，即$\forall i \quad x \in V_i$, 对重量泛函

$$E(x) = \int_{V_i} \rho(s) d^2(x, s) ds$$

每个$\{v_i\}$恰使得对应的$V_i$中重量泛函达到最小。其中$\rho(s)$ 为密度分布，对均匀物质可设置为1。 此时对应的$v_i$称为对应的沃罗努瓦分布的生成元（Voronoi Generators）。

从定义中我们可以看出，如果我们已知一个分布是重心沃罗努瓦分布，那么只需要获得$\{V_i\}$或$\{v_i\}$ 都可以推出另一方。然而定义中要求$\{V_i\}$和$\{v_i\}$ 互相约束。虽然每一组$\{v_i\}$只对应一个$\{V_i\}$，但是不同的$\{V_i\}$却可能对应相同的$\{v_i\}$；同时$\{v_i\}$生成的沃罗努瓦分布也并不一定以$\{v_i\}$ 为重心。所以重心沃罗努瓦分布并不能简单直接生成。同时需要指出的一点在于，上文定义的距离并不一定是欧氏距离，可以是广义距离；而密度函数也可以定义为许多东西，这在后面的分析中将会有所提及。下一节将介绍一种简单泛用生成重心沃罗努瓦分布的算法。

### 2.2 重心沃罗努瓦分布生成算法

当前比较广泛使用的一种简单生成重心沃罗努瓦分布的算法是罗伊德算法（Lloyd Algorithm）。对于给定的区域$\Omega$以及正整数$k$，我们可以生成$\Omega$ 中的$k$个区域的重心沃罗努瓦分布.

**算法1**：$\mathbf{R}^n$中开集的重心沃罗努瓦分布生成算法
任取$\Omega$中$k$个互不相同的点，记为$\{v_i\}_{i=1}^k$
①对$\Omega$由$\{v_i\}_{i=1}^k$构造其沃罗努瓦分布，记为$\{V_i\}_{i=1}^k$。
② 对$\{V_i\}_{i=1}^k$最小化重量泛函，得到每一个的重心，记为$\{\overline{v}_i\}_{i=1}^k$。
③计算变化距离$\delta = \sum_{i=1}^k d(v_i, \overline{v}_i)$，比较$\delta$与预设阈值$\epsilon$。

如果$\delta \leq \epsilon$，则终止，获得的$\{\overline{v}_i\}_{i=1}^k$ 即为某一沃罗努瓦分布的生成元；

如果$\delta > \epsilon$，则令$v_i = \overline{v}_i, i = 1, \ldots, k$，返回①循环。

其中③的终止条件可以改成其他适合的条件，目标是使得每次循环的变动得以控制。

1. 内容

### 2.3 流场重心沃罗努瓦分布生成算法

相比于一个单纯的重心沃罗努瓦分布，例如上节所讨论的沃罗努瓦分布，相当于对某个只有一个标量函数$\rho$的区域进行划分，获得其沃罗努瓦分布。对一个流场来说，其上至少有一个向量值函数。当然，我们可以用对标量场相同的操作来进行沃罗努瓦划分；但这意味着我们将无法通过重心的分布来获得流场区别于普通标量场的信息。一个很直接的想法就是我们通过变更距离函数和密度函数，使得在沃罗努瓦划分的过程，考虑流场的向量信息。
下面给出一个求解流场重心沃罗努瓦分布的算法范例，这是

DU, Q., AND WANG, X. 2004. Tessellation and clustering by mixture models and their parallel implementations. In Proceedings of the 2004 SIAM data mining conference, Orlando, FL, SIAM.

中提出的一种流场重心沃罗努瓦分布生成算法。 对$\Omega \subset \mathbf{R}^n, \forall x \in \Omega$，记该流场为$y = F(x) \in \mathbf{R}^m$。 定义距离$\forall x_i \in \mathbf{R}^n, y_i = F(x_i), i = 1, 2$给定正常数$w$

$$d(x_1, x_2) = \sqrt{|y_1|^2 - |y_1| y_1 * y_2 + w|y_1|^2 |x_1 - x_2|^2}$$

于是我们得到了沃罗努瓦划分：

$$V_i = \{x \in \Omega | d(x_i, x) < dx_j, x, \forall j \neq i\}$$

以及重量泛函：

$$E(x_i) = \int_{V_i} |y_i|^2 - |y_i| y_i * y_s + w|y_i|^2 |x_i - x_s|^2 dx_s$$

基于最小化重量泛函，我们得到了如下的算法。
**算法2**：流场重心沃罗努瓦分布生成算法

任取$\Omega$中$k$ 个不重合点，记为$\{x_i\}_{i=1}^k$
①由这$k$个点生成$\Omega$的一个沃罗努瓦分布$\{V_i\}_{i=1}^k$

$$V_i = \{x \in \Omega | d(x, x_i) < d(x, x_j), j \neq i\}$$

②由每个沃罗努瓦区域生成其重心

$$\overline{x}_i = \frac{\int_{V_i} |F(x)|^2 x dx}{\int_{V_i} |F(x)|^2 dx}, \quad \overline{y}_i = \frac{\int_{V_i} |F(x)| F(x) dx}{|\int_{V_i} |F(x)| F(x) dx|}, i = 1, \ldots, k$$

③比较$\{x_i, y_i\}, \{\overline{x}_i, \overline{y}_i\}_{i=1}^k$
如果其相等或满足某些容许条件（例如$\sum_{i=1}^k d(x_i, \overline{x}_i) < \epsilon$），终止运算，$\{\overline{x}_i, \overline{y}_i\}_{i=1}^k$ 即为所求的沃罗努瓦生成元；
如果不满足容许条件，则将$\{x_i, y_i\}$赋值为$\{\overline{x}_i, \overline{y}_i\} i = 1^k$，然后回到①中循环。

这里假设的是均匀场，即$\rho(x) \equiv 1$，如果对于非均匀场$\rho(x)$，可将②中的公式替换为： ②'

$$\overline{x}_i = \frac{\int_{V_i} \rho(x) |F(x)|^2 x dx}{\int_{V_i} \rho(x) |F(x)|^2 dx}, \quad \overline{y}_i = \frac{\int_{V_i} \rho(x) |F(x)| F(x) dx}{|\int_{V_i} \rho(x) |F(x)| F(x) dx|}$$
$$i = 1, \ldots, k$$

### 2.4 离散情况

由于实际的数据场，通常是离散的点数据，而非一个统一的函数场。这时只需将上节的算法中积分变成求和便可以直接平推到离散情况。但同时我们可以得到一个更直接的算法。

**算法3**：对离散数据场$\{c_i\}_{i=1}^n$，其上向量值函数$F(x)$，任取$k$个互异点$\{x_i\}_{i=1}^k$。

①以$\{x_i\}$为生成元生成一个沃罗努瓦分布。，将所有数据点$c_i$划分为$k$组，每组记为$V_k$。

②对每一组$c_i$求其新重心：

$$\overline{x}_i = \frac{\sum_{V_k} |F(x)|^2 x dx}{\sum_{V_k} |F(x)|^2 dx} \quad , \quad \overline{y}_i = \frac{\sum_{V_i} |F(x)| F(x) dx}{|\sum_{V_i} |F(x)| F(x) dx|}$$

③对比新重心与原重心，如果相差足够小就终止，否则将$\overline{x}_i$赋值给$x_i$，返回①。

### 2.5 关于CVT生成算法的说明

之前两节所介绍的CVT生成算法，其原型是劳埃德算法（Lloyd Algorithm）。劳埃德算法解决K——平均问题时优势是收敛快，劣势是不能保证全局最优。然而由于CVT 并不需要要求其重量泛函的和是全局最优（最小）的，正好克服了劳埃德算法的这一劣势。

当然，尽管说劳埃德算法的收敛相对快，但也只是相对而言。对于大规模的数据场，其计算时间依旧可能十分长。现在对于时间代价的减小方法主要是拟牛顿法

Fast Methods for Computing Centroidal Voronoi Tessellations James C. Hateley · HuayiWei · Long Chen

其具体思路是：????????????????????????

### 2.6 多个相近时间层的CVT方法

有时我们需要处理多个连续时间层的流场可视化，从而展示一定空间内流场的变化情况。而对一个固定空间中多个相近时间层求其CVT 时，除了每一个时间层分别求其CVT以外，我们希望能够有一些更快的，相临时间层之间有更多联系的方法。

考虑流场的变化对CVT重心位置的影响。事实上在迭代过程中计算新的CVT重心时，可以看成对某个区域中的所有点进行对其向量模长的加权平均。也就是说，只要模长不变，或者模长平方的比例保持不变，对同一个点集来说，则流向不管如何变化，CVT 重心依然不变；同时，当向量模长整体变化不大时，新的重心也不会有较远的偏移。 基于以上的考虑，在处理连续多个时间层的CVT时，可以将上一个时间层的CVT 重心作为新的时间层的重心初始点$\{x_i\}_{i=1}^k$。 这样可以在一定程度上节约时间成本。事实上在具体的实验中可以节约$\frac{5}{6}$的时间，甚至更多。

需要注意到前几节最终得到的$\{x_i, y_i\}$ 是在$x_i$ 处的单位向量。也就是说，重心沃罗努瓦分布最终获得的仅仅是一个代表点的分布。距离实际的可视化还差一些工作。下一章将说明具体的可视化手段。

## 3 可视化方法

上一章介绍了重心沃罗努瓦分布的定义以及获得方法。通过Lloyd 算法，我们可以得到误差任意小的一个CVT。这样我们就有了一组在CVT的每个重心的单位长度的向量，于是我们就可以进行一些具体的可视化步骤了。

然而不同的可视化，因其目标差异，所以关注点亦有不同。这就造成了可视化并无法评判对错的现象。广泛的来说，只要能展现出所想要表达，凸出的的特征，我们就可以说这个可视化是成功的。下面我将就长度，颜色，图形表示（箭头，椭圆盘，三角）给出一些说明与例子。

### 3.1 长度

首先我们要对那些单位长度的向量进行长度修正。这里的修正可以按照最关心的量进行修正，如直接取区域中最大（小）的向量模长，或者梯度等等。当然，从人们的一般直观来说，第一想到的都会是区域的向量大小的某种表示，所以我们通常首选区域向量大小的平均：

$$L_i = \frac{1}{|V_i|} \int_{V_i} |F(x)| dx$$

或

$$L_i = \frac{1}{V_i} (\int_{V_i} |F(x)|^2 dx)^{1/2}$$

对非均匀场：

$$L_i = \frac{(\int_{V_i} \rho(x) |F(x)|^2 dx)^{1/2}}{\int_{V_i} \rho(x) dx}$$

这样将CVT输出的$\overline{y}_i$修正为$\overline{y}_i * L_i$，即在$\overline{x}_i$ 处绘制$\overline{y}_i * L_i$的向量箭头即可。

### 3.2 颜色

最通常的可视化都会用颜色来表达大小，例如天气预报中的温度图，亮色表征高温，冷色表征低温，这是在标量场中非常常用的。而到了向量场，由于有了位于CVT 重心的有长度向量，颜色可以用来表达一些其他的东西，例如某个区域数据点的疏密、该区域重量泛函的相对大小、或是该区域向量的差异。以最后一个为例，我们可以将某一个沃罗努瓦区域的向量差异大小定义为：

$$Var(V_i) = \frac{\int_{V_i} |y|^2 - |y| y \times y_i dx}{|V_i|}$$

然后对所有区域的$Var(V_i)$进行归一化处理，再对应到灰度值或者其他的颜色数值即可。这样可以在展示向量场模长的同时，再多表现一个特征量。

### 3.3 图形表示

至于对在CVT重心处向量的表示图形的选择，比较常用的是箭头和椭圆。相比较来说，各有优劣；箭头对方向的表现更加明确，但相对对颜色的展示较差；椭圆对颜色的表述很好，但让人很难确认其起始点（CVT重心）。

与此同时，因为点的疏密程度不同，如果CVT 区域较多且分布不均匀时，很可能出现绘制表现的交叉（箭头交叉和椭圆重叠）。因此可以考虑在局部使用短流线段来保证不交叉。具体如下。

**算法4**：流线生成算法；设定流线每步步长$L$与终止条件（如步数）

①取定初始点$x_0$，对应$y_0 = F(x_0)$

②计算$x_1 = x_0 + L * y_0$，对应$y_1 = F(x_1)$

③$y = \frac{y_0+y_1}{2}$延$y$的方向走$L$到达$x_2$，即$x_2 = x_0 + L * y$

③如果$x_2$满足终止条件，则终止，否则令$x_0 = x_2$，回到①

这样我们将初始点选取为CVT的重心，非数据点的向量通过插值拟合获得，就可以得到任意接近真实流线的在CVT重心的短流线。这样的短流线在表征区域整体方向的同时，还可以一定程度上表征较近区域的流向，例如涡流在短流线中就会出现明显的弯曲，但是相同的场用相同的CVT数进行可视化而不通过流线时，却无法明确的体现涡流。

通过上述的手段，我们已经可以将数据转化成我们需要的图片了。但是需要注意的一点在于：上述的很多方法，包括CVT过程中距离的选取，可视化时颜色，形状的决定等，都有很大的选择空间。这些需要根据具体情况来决定。下一章将展示一些具体的实例。

## 4 应用与数值实验

本章将主要展示运用CVT方法进行的流场可视化实例。

#### 4.0.1 *Subsubsection Heading Here*

Subsubsection text here.

## 5 Conclusion

The conclusion goes here.

## Appendix A
## Proof of the First Zonklar Equation

Appendix one text goes here.

## Appendix B

Appendix two text goes here.

## Acknowledgments

The authors would like to thank...

## References

[1] H. Kopka and P. W. Daly, *A Guide to LATEX*, 3rd ed. Harlow, England: Addison-Wesley, 1999.

PLACE PHOTO HERE

**Michael Shell** Biography text here.

**John Doe** Biography text here.

**Jane Doe** Biography text here.