**Software Engineering and Applications** 软件工程与应用, 2019, 8(2), 58-64
Published Online April 2019 in Hans. http://www.hanspub.org/journal/sea
https://doi.org/10.12677/sea.2019.82007

Hans汉斯

# Realization and Application of Multipoint Geostatistics Algorithms by C#

**Fengyun Guo, Changyun Miao**

The 22nd Research Institute of CETC, Xinxiang Henan
Email: gfy432@163.com

Received: Mar. 24th, 2019; accepted: Apr. 8th, 2019; published: Apr. 15th, 2019

## Abstract

**In Snesim algorithm of multi-point geological statistics, the probability distribution can be determined by scanning the number of data events which are same as the distribution of conditional data in the training image, which can reflect the joint variability of multiple locations. In this paper, Snesim algorithm is implemented in C# language. Using the parallel operation ability of C# to build a search tree greatly improves the computational efficiency. The search algorithm of data events is studied in this paper, and the effect of changing training image to verify the treatment of simulation results is explored.**

## Keywords

**Training Image, Search Tree, Data Events, Parallel Computation**

# 多点地质统计学算法的C#实现及应用

**郭凤云，苗长运**

中国电子科技集团公司第22研究所，河南 新乡
Email: gfy432@163.com

收稿日期：2019年3月24日；录用日期：2019年4月8日；发布日期：2019年4月15日

## 摘 要

**在多点地质学统计Snesim算法中，通过在训练图像中扫描与待模拟点条件数据分布相同的数据事件个数来确定概率分布，可以反映出多个位置的联合变异性，本文用C#语言完整地实现了Snesim算法，利用C#的并行运算能力建立搜索树大大提高了计算效率，对搜索树中数据事件的查找算法进行研究，并探索了改变训练图像来验证对待模拟结果的影响。**

文章引用: 郭凤云, 苗长运. 多点地质统计学算法的 C#实现及应用[J]. 软件工程与应用, 2019, 8(2): 58-64.
DOI: 10.12677/sea.2019.82007

## 关键词

**训练图像，搜索树，数据事件，并行计算**

Copyright © 2019 by authors and Hans Publishers Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY).
http://creativecommons.org/licenses/by/4.0/
Open Access

## 1. 引言

在地质工程中，往往是研究测量地层、裂隙、断层等的空间展布特征及其物理力学参数。由于不可能对某一研究的相关地质变量进行连续的测量，因此往往取一些有代表性的点作为采样点，然后再运用各种不同的预测技术，来推测出整个研究区域内该地质变量的空间变化规律。

多点地质统计学是Journel等学者首先提出的，Strebelle等在此基础上，从训练图像着手提出了Snesim [1] (Single normal equation simulation)算法，使得多点地质统计学算法成为一种真正实用的随机建模算法。传统的两点地质学随机建模方法只能考虑空间两点之间的相关性，而多点地质统计学着重表达多点之间的相关性，克服了两点地质统计学的不足，是目前国际前沿研究方向。多点地质统计学是相对于传统的两点地质统计学而言的，可以反映空间多个位置点的几何形状与相互配位关系，在模拟复杂形状地质体分布中有较大优势。

本文从多点统计学的基础出发，对多点地质统计学的经典算法进行介绍，用C#进行程序的实现，进一步探索了多点统计学中的关键因素——训练图像选取对整个算法结果的影响并进行了实测数据的验证和探索。

## 2. 多点地质统计学基础

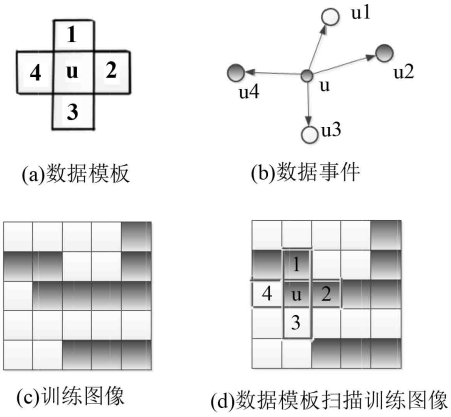
多点统计学着重表达多点之间的相关性，多点的集合用数据事件来表达。若在待模拟对象中存在某种属性 $S$，可取 $m$ 个状态，即为 $\{S_k, k=1,2,\cdots,m\}$，则一个以 $u$ 为中心，大小为 $n$ 的数据事件 $d_n$ 由两部分组成：由 $n$ 个向量 $\{h_\alpha, \alpha=1,2,\cdots n\}$ 确定的数据模板 $\tau_n$ 和 $n$ 个向量联合向量终点处数据值构成[2]。其中数据模板见图1(a)所示，数据事件见图1(b)所示，数据 $u_1$、$u_2$、$u_3$、$u_4$ 被称为待模拟数据 $u$ 的条件数据，简而言之数据事件就是数据模板中各条件数据及待模拟点 $u$ 分别取特定的属性状态。训练图像(见图1(c))在地质学中用来描述地层中各向异性，地质体的走向、分布等，包含了待模拟区域各种特征模式，只是一种概念上的特征模式集合，不需要有很高的精确度或者符合某种条件数据的分布。通过扫描训练图像，先验模型被明确而定量的引入到建模当中，先验模型包含了被研究的属性值中存在的结构特征，可以说训练图像中的概率信息决定了最终的模拟结果。

首先假设训练图像 $T$ 平稳，利用数据模板对训练图像进行扫描，当在训练图像中出现与数据模板相同的数据事件 $d_n$ 时，记为一次重复，重复次数 $C(d_n)$ 与“侵蚀的训练图像” $T_n$ (一个使得以 $u$ 为中心的数据模板中所有节点都在训练图像 $T$ 内的集合)的大小 $N_n$ 的比值为 $d_n$ 出现的概率[3]，表示为：

$$P\{d_n\} = P\{S(u_\alpha) = s_{k_\alpha}; \alpha = 1,2,\cdots,n\} \approx \frac{C(d_n)}{N_n} \tag{1}$$

对于任一取样点，在给定 $n$ 个条件数据的情况下，属性 $S(u)$ 取 $K$ 个状态中任一状态值的概率记为 $\text{Prob}\{S(u) = S_k | d_n\}$，其中 $d_n$ 为由 $n$ 个条件数据联合构成的数据事件。根据贝叶斯条件概率分布公式，

该概率可以表示为：




(a)数据模板　(b)数据事件　(c)训练图像　(d)数据模板扫描训练图像

**Figure 1.** Basic elements of multipoint geostatistics
**图 1.** 多点地质统计学基本要素

$$\text{Prob}\left\{S(u)=S_k \middle| d_n\right\}=\frac{\text{Prob}\left\{S(u)=S_k \text{ and } S(u_\alpha);\alpha=1,2,\cdots,n\right\}}{\text{Prob}\left\{S(u)=S_{k_\alpha};\alpha=1,2,\cdots,n\right\}} \tag{2}$$

式中：分母为条件数据事件，其出现的概率可以式(1)获取；分子为条件数据事件及待模拟点 $u$ 取 $S_k$ 的情况同时出现的概率，相当于在已有的 $C(d_n)$ 个重复中，$S(u)=S_k$ 的重复个数 $C_k(d_n)$ 与侵蚀的训练图像大小的比值。因此，局部条件概率分布函数可表示为：

$$P\left\{S(u)=S_k \middle| d_n\right\}\approx\frac{C_k(d_n)}{C(d_n)} \tag{3}$$

因此，通过扫描训练图像，可以获取待模拟点的条件概率分布函数，然后通过随机抽样，获得待模拟点的模拟值。

## 3. 多点地质统计学随机建模方法及 C#编程实现

### 3.1. Snesim 算法

基于搜索树方法，Strebelle and Journel 提出了多点地质统计随机模拟的 Snesim 算法(Strebelle and Journel, 2001; Strebelle, 2002)，整个多点地质统计学建模过程基本围绕 Snesim 算法进行的，本文主要是对 Snesim 算法的 C#语言实现进行展开的，具体实现流程如下：

① 建立训练图像，选定训练图像后，按定义的坐标规则输入训练图像数据，本文定义的坐标原点为训练图像矩阵的左下角；

② 准备待模拟数据，将实测的井数据作为已知的条件数据标注在最近的网格结点上；

③ 用自定义的数据模板 $\tau_n$ 扫描训练图像，以构建搜索树。

构建如图 2 所示搜索树是一个很复杂的过程，需要遍历训练图像集合中的所有符合条件的某种数据事件，将事件重复次数存储成搜索树的数据结构形式。本文使用 Paraller. For()方法类循环多次执行一个任务，并行运行[4] [5]迭代，大大缩短了计算机的运算时间，以下是截取所用程序的部分代码：

```
int Ni = trainingImage.GetLength(1);
int Nj = trainingImage.GetLength(0);//读取训练图像的维数
```

```
Vector[] all = new Vector[Nj * Ni]; //存储训练图像中每个元素的坐标信息
var result = Parallel.For(0, Nj, J => Parallel.For(0, Ni, I => { all[J * Ni + I] = new
Vector { ReX = I, ReY = J }; }));
  //遍历训练图像的每个元素，将每个元素的坐标信息赋值给 all 数组
while (!result.IsCompleted) ; //判断循环是否结束
return createSearchTreeAccumulator( trainingImage,dataTemplate,Next(trainingImage,
 all, new Vector { ReX = 0, ReY = 0 }, 0), Next(trainingImage, all, new Vector { ReX
 = 0, ReY = 0 }, 1));//调用搜索树创建函数
```

Paraller. For()方法类似于 C#的 for 循环语句，多次执行一个任务。使用 Paraller. For()方法并行运行迭代，迭代的顺序没有定义。Paraller. For()方法的返回类型是 ParallelLoopResult 结构，它提供了循环是否结束的信息和最低迭代的索引。

④ 确定一个访问未取样结点的随机路径，从搜索树里提取多点概率分布，计算局部条件概率，建立未取样点局部条件概率分布；

⑤ 用蒙特卡洛方法[6]，根据随机数的值分布区间决定待模拟点的属性值 $u$。$u$ 的取值基于先验定义该区域的属性状态集合(本文定义两种状态 0 和 1)，根据随机数的概率分布决定 $u$ 取 0 或者 1，并将实现值加入到条件数据中；

⑥ 重复步骤④、⑤，如果需要多个实现，只需改变随机访问路径，重复步骤④、⑤、⑥即可；

⑦ 改变随机路径，产生另一个随机模拟实现。

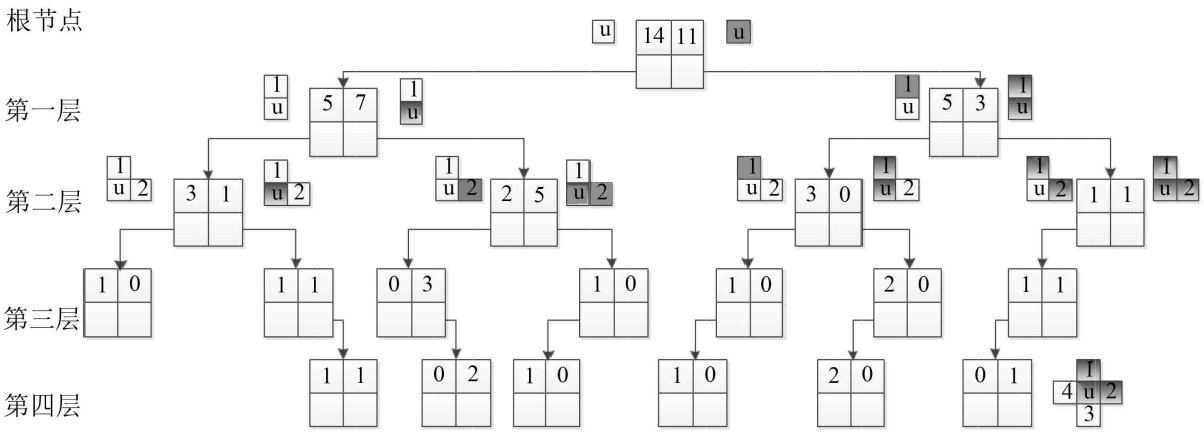



**Figure 2.** Search tree structure
**图 2.** 搜索树结构

## 3.2. 搜索树中数据事件查找及条件概率提取算法

搜索树数据结构构造成功后，另一关键算法便是如何高效准确查找给定数据事件在搜索树种存储的节点位置，以便提取数据事件的重复次数，从而得到待模拟点 $u$ 的条件概率分布函数，本文结合对搜索树数据结构的研究，提出了以下数据事件重复数查找算法：

① 假如待模拟点 $u$ 的条件数据 $u_1$、$u_2$、$u_3$、$u_4$ 均不存在，则返回搜索树根节点；

② 假如待模拟点 $u$ 的条件数据 $u_1$、$u_2$、$u_3$、$u_4$ 均存在，还应判断条件数据的状态值是否可以在搜索树中找到相应的节点，若条件数据带 $u_1$、$u_2$、$u_3$、$u_4$ 的数据事件在搜索树节点中不存在，则需依次减少条件数据的个数，直至找到去除所有的条件数据返回搜索树的根节点；

③ 假如待模拟点 $u$ 的条件数据 $u_1$、$u_2$、$u_3$、$u_4$ 不完全存在，则不存在的条件数据分两种状态(本文暂定两种状态 0 和 1)与存在的条件数据状态组合成完全的条件数据，然后分两个分支判断节点是否存在，若存在节点，返回该节点中，若不存在则依次去掉较远的条件数据再去查找节点；两个分支取较少层数的长度(节点短的分支长度)，然后取相同层数的节点中的 0 和 1 状态的事件重复数对应相加，最终得到待模拟点为不同岩性状态的条件概率分布。

## 4. 数据仿真及实际地层岩性模拟应用

首先，先给定一图 1(c)中取值的训练图像及二维实验待模拟数据矩阵进行测试，测试过程如下：

① 读入训练图像数据(见图 3)，并定义一定的坐标系统对其进行存储。

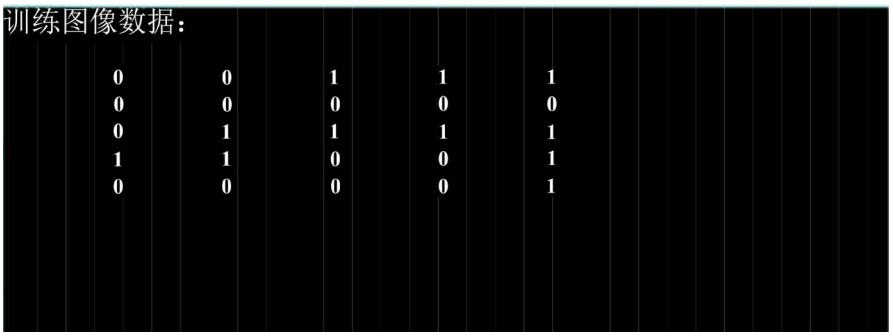



**Figure 3.** Training image data
**图 3.** 训练图像数据

② 将待模拟数据矩阵按与训练图像同样的坐标系统输入，并对待模拟矩阵进行一系列预处理，如矩阵边界的处理、待模拟点数据及其坐标的提取。

③ 确定数据模板的样式，本文采用的是图 1(a)所示模板，数据模板确定后，便可构造搜索树，然后对训练图像进行扫描、数据事件的查找及待模拟点条件概率分布函数的生成等。

④ 按蒙特卡洛方法生成随机数决定待模拟点的岩性状态，替代原始待模拟数据矩阵中的未知状态 −1。

由图 4 中可以看出，在多点地质统计学的储层建模时，本文实现的算法完成了对未知待模拟点的未知岩性的估计。

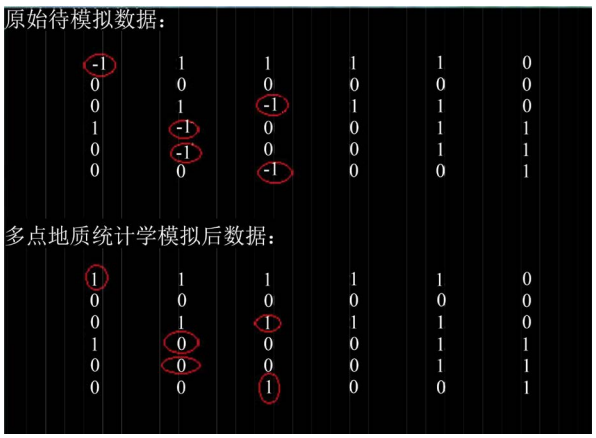



**Figure 4.** The comparison of the raw data with the data after multi-point statistical simulation
**图 4.** 原始待模拟数据矩阵及多点统计学模拟后数据对比

为了验证本文 C#语言实现的多点地质统计学算法的实际应用效果，本文取得了某石油局某区块的原始井况数据，并针对储层的岩性进行了模拟，仿真模拟如下：

取 10 * 10 的训练图像 T1 对实际井况数据的某特定目的层进行多点地质统计学计算模拟，结果如图 5。

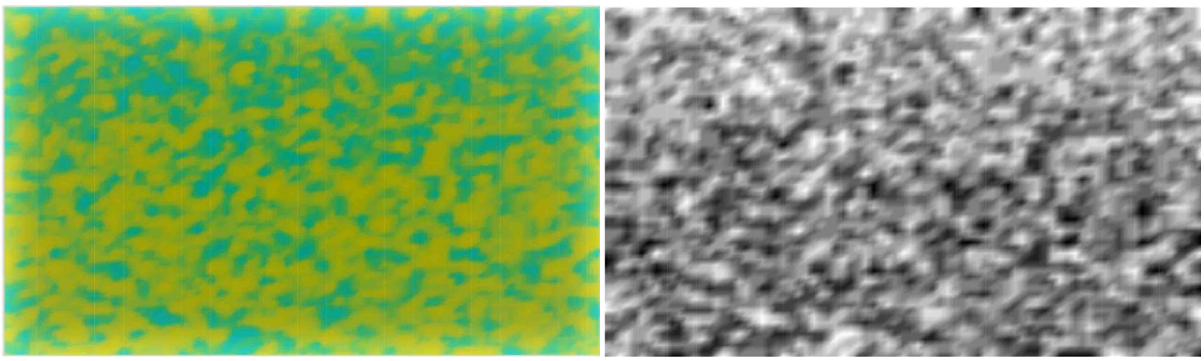


**Figure 5.** The simulated result graph using training image T1
**图 5.** 运用训练图像 T1 模拟结果图

当改变训练图像的岩性状态，及把训练图像 T1 的岩性状态(暂定 0 代表泥岩，1 代表沙岩)进行互换得到训练图像 T2，然后对同样的待模拟数据进行同样的多点地质统计学计算模拟结果如图 6。

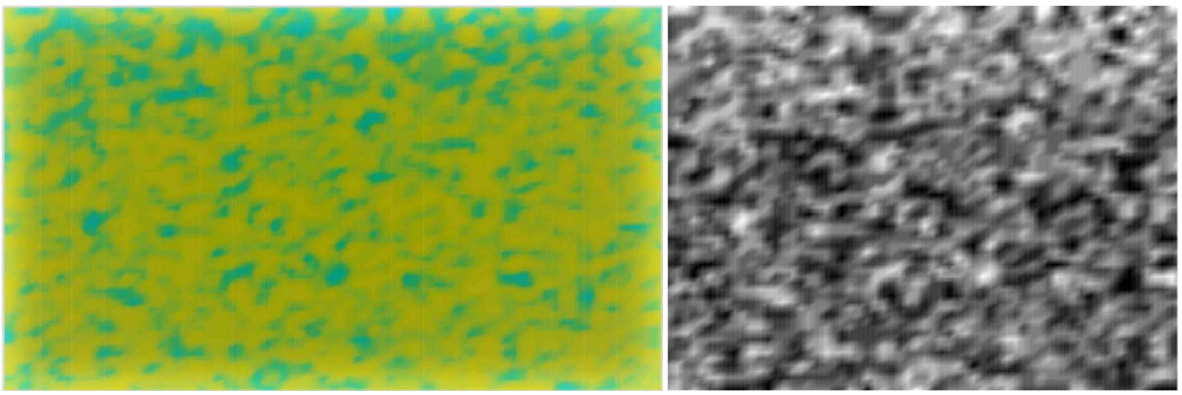


**Figure 6.** The simulated result graph using training image T2
**图 6.** 运用训练图像 T2 模拟结果图

将训练图像的泥沙特性反转，模拟结果能很好的体现出训练图像的影响，由此得出，本文实现的多点统计学算法，改变训练图像能对模拟结果造成影响。

## 5. 结论

本文实现了多点地质统计学的 Snesim 算法，采用了 C#语言中的并行计算及递归循环等关键技术来实现了树型数据类型存储和查找，大大地提高了代码的效率，节省了计算时间，并进行了数据的仿真及实际地层数据的模拟，通过改变训练图像，测试选取不同的训练图像对最终模拟结果的影响。经过验证说明本文实现的 Snesim 算法在多点地质统计学应用中具有可行性，本文只是探索了训练图像的选取对结果的影响，实际的地层建模中，训练图像的选取参考因素更复杂[7]，训练图像的获取是多点地质统计学中很关键的因素，将成为未来多点地质统计学研究的重要方向。

## 参考文献

[1] 李少华, 张昌民, 尹艳树. 储层建模算法剖析[M]. 北京: 石油工业出版社, 2012: 139-142.

[2] Strebelle, S. and Journel, A. (2001) Reservoir Modeling Using Multiple-Point Statistics. *SPE Annual Technical Conference and Exhibition*, New Orleans, Louisiana, 30 September-3 October 2001, SPE 71324. https://doi.org/10.2118/71324-MS

[3] 王芳芳, 陈小宏, 李景叶. 多点地质统计学整合地震数据的研究算法[C]//石油地球物理勘探物探技术研讨会专刊, 2012: 219-222.

[4] 高璐, 马玉志. 浅谈 Microsoft Visual Studio 2010 新特性[J]. 黑龙江科技信息, 2010(32): 96.

[5] 邢陈思. 探索 Visual Studio 2010 对并行计算的支持[J]. 金融科技时代, 2012, 20(1): 56-58.

[6] 张挺. 基于多点地质统计的多孔介质重构方法及实现[D]: [博士学位论文]. 合肥: 中国科学技术大学, 2009: 36-37.

[7] 熊哲, 严锡, 金杨波. 油藏地质建模中多点训练图像应用研究[J]. 江汉石油职工大学学报, 2012, 25(4): 1-2.

Hans汉斯

**知网检索的两种方式：**

1. 打开知网页面 http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD
   下拉列表框选择：[ISSN]，输入期刊 ISSN：2325-2286，即可查询
2. 打开知网首页 http://cnki.net/
   左侧“国际文献总库”进入，输入文章标题，即可查询

投稿请点击：http://www.hanspub.org/Submission.aspx
期刊邮箱：sea@hanspub.org