

物探数据三维属性模型高效构建方法研究*

郑坤^{1,2}, 王 镭^{1,2}, 黄 飞^{1,2}, 李芬蕾³

(1. 中国地质大学 信息工程学院, 武汉 430074; 2. 教育部地理信息系统软件及应用工程中心, 武汉 430074; 3. 江苏省邮电规划设计院有限责任公司, 南京 210006)

摘 要: 针对传统三维地质属性建模方法进行了研究, 提出了一种物探数据三维属性模型的高效构建方法。该方法利用结构化数据快速搜索策略, 并采用了分块建模机制。在 PC 机上的实验表明, 该方法不仅克服了大规模数据引起的计算机系统硬件瓶颈, 而且在保证物探属性模型质量的前提下, 构建效率显著提高。

关键词: 三维属性建模; 物探数据模型; 样本点搜索; 分块

中图分类号: TP391.9 **文献标志码:** A **文章编号:** 1001-3695(2012)11-4367-03

doi: 10.3969/j.issn.1001-3695.2012.11.095

Efficient modeling method for 3D geophysical prospecting attribute model

ZHENG Kun^{1,2}, WANG Pei^{1,2}, HUANG Fei^{1,2}, LI Fen-lei³

(1. Faculty of Information Engineering, China University of Geosciences, Wuhan 430074, China; 2. China GIS Software Research & Application Engineering Center of the Ministry of Education, Wuhan 430074, China; 3. Jiangsu Posts & Telecommunications Planning & Designing Institute Co. Ltd., Nanjing 210006, China)

Abstract: Based on the study of the traditional 3D geological attribute modeling method, this paper proposed an efficient construction method for 3D geophysical prospecting attribute model. Through the fast search strategy of the structured data and block-division mechanism, it not only realized the efficient construction of attribute model, but also overcame the limit of system memory caused by massive data. The experiment shows that based on the block-division mechanism the high-speed modeling method has obvious effect in the construction of geophysical data attribute model.

Key words: 3D attribute modeling; geophysical data model; sample points search strategy; block-division

0 引言

地球物理勘探是地质调查的重要组成部分, 包括重力、磁法、电法、地震勘探等多种探测手段。与此相应获取了大量如密度、磁导率、电导率、弹性等真三维属性数据, 通过分析研究这些真三维数据, 可推断、解释地质构造和矿产分布情况。三维地质属性模拟技术是揭示地质体内部物化属性参量的一个重要手段, 能够反映属性参量的空间分布、变化规律以及参量之间的关联关系。

研究如何有效地利用三维地质属性模拟技术来真实反映异常信息, 验证地质解释成果, 是目前地球物理专业领域研究的重要方向之一^[1]。三维地质属性建模是三维地质属性模拟的关键技术, 是通过有限的采样资料, 采取合理的空间插值方法, 模拟真实的地质参数空间分布规律。在地学领域常用的插值算法主要包括几何方法、地质统计学法和随机模拟法^[2]。其中, 几何方法以距离反比加权法^[3]为代表, 计算比较简单, 易于实现, 但是生成网格曲面的速度较慢, 主要用于中、小规模散乱点的模拟^[4]; 以 Kriging 为代表的地质统计学法^[5,6]是以

地质统计学作为坚实的理论基础, 克服了模型构建中误差难以分析的问题, 能够对误差作出逐点的理论估计, 但是该方法计算量大, 常以耗费时间和计算资源为代价^[7]; 以序贯指示模拟^[8]等为代表的随机模拟方法定义的各种随机变量之间的空间相关, 可以根据相邻数据把高度不确定性的先验分布更新为低不确定性的后验分布, 但此方法建模困难、计算量大^[9]。

随着地球物理勘探技术的更新以及对三维地质属性模型本身精细程度要求的提高, 需要构建的属性模型规模正以惊人的速度增长, 已达到十亿字节 (GB) 甚至万亿字节 (TB), 使得传统的建模手段在面对大规模数据量时低效性问题凸显。因此, 寻求一种高效的适合于物探数据三维属性模型的高效构建方法, 已经成为三维地质属性模拟研究的一个重要内容。

1 传统属性建模方法分析

距离反比加权法是目前众多传统方法中应用较为广泛的建模技术之一, 其算法简单, 易于实现, 适合将不规则离散点数据转换为规则的栅格数据, 具有重要的实用意义^[10]。该方法是对每一个网格点按照某种规则寻找与其距离最近的若干个

收稿日期: 2012-04-04; **修回日期:** 2012-05-14 **基金项目:** 国家教育部新世纪优秀人才支持计划资助项目 (NCET-07-0772); 武汉市学科带头人计划资助项目 (201271130443)

作者简介: 郑坤 (1977-), 男, 湖北襄樊人, 副教授, 博士, 主要研究方向为三维地学模拟; 王镭 (1987-), 男, 河北石家庄人, 硕士研究生, 主要研究方向为计算机三维建模与可视化 (myfrank0922@sina.com); 黄飞 (1984-), 男, 湖北武汉人, 硕士研究生, 主要研究方向为计算机三维数据存储; 李芬蕾 (1985-), 女, 浙江舟山人, 硕士, 主要研究方向为计算机三维可视化。

点,然后对这若干个点进行加权平均,得到网格点的值。若散布在空间上的数据点数量较少时,此方法效率较高;若散布在空间上的数据点数量很多时,其效率就急剧下降^[11]。因此,在实际应用中,传统距离反比加权法仅适合于数据量不大的空间数据的插值计算^[12]。

传统方法在面对大数据量时计算效率低的主要原因在于,对每一个待插点必须按照某种规则在全局范围内搜索所有数据点,以寻找与其距离最近的样本点。常用的搜索策略有固定搜索范围法和固定搜索数目法^[10]两种。

固定搜索范围是指首先确定一个固定大小的搜索体(如圆球体或椭球体)进行搜索,然后通过搜索体范围内的样本点计算待插点的值。该方法计算简捷,但是搜索体半径的确定是较难解决的问题:若半径过大,则搜索到的样本点数目过多,导致效率降低且不会提高模型精度;但若半径过小,则会导致样本点过少或无样本点。

固定搜索数目是指对每一个待插点采用半径大小可变的搜索体,以保证每个待插点搜索到的样本点数目相同。由于该方法对每个待插点都需要与所有样本点的距离进行比较,以确定最近的 n (搜索数目)个样本点来插值,其效率远不及固定搜索范围法。

由此可见,能否高效地搜索待插点周围的样本点是影响三维地质属性建模效率的重要因素,且数据量越庞大其影响程度越明显。当前物探采样点个数及其预构建的属性模型的网格密度通常在千万级以上,传统的建模方法在处理此类规模数据时突显出的低效性已经无法满足当前模型构建的要求。同时,有限的系统内存容量也成为成功处理大规模数据的瓶颈。因此,本文针对重力、磁法等物探采样数据的空间分布特点,提出了结构化数据快速搜索策略和分块建模机制,实现了物探数据三维属性模型的高效构建,并成功克服了因系统内存受限导致建模失败的问题。

2 属性模型高效构建方法

2.1 结构化数据快速搜索方法

本文通过对磁法、重力等物探属性采样资料的分析发现:其采样数据分布呈结构化特性,即在每一水平层上样本点分布密度大且规则,而垂直方向上稀疏且非等间距采样。每个样本点都具有其特定的行号、列号和层号。如图1所示,网格点代表原始离散样本点,空心点 A 、 B 、 C 、 D 代表待插点的四种位置属性。

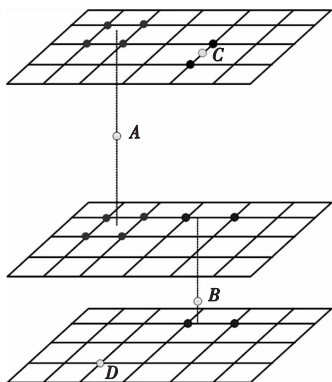


图1 物探数据空间分布及待插点位置属性

本文基于此规律提出了结构化数据的样本点快速搜索策略,即根据待插点在原始采样数据中的三维空间位置,高效确定周围样本点。其算法流程如图2所示。具体步骤如下:

a) 计算待插值点在待构建属性模型中的位置关系,即其三维索引与待构建属性模型格网点个数的比值,记为 (dX, dY, dZ) 。

b) 根据 (dX, dY, dZ) 反推出待插点在原始数据场中的位置关系 $d3CurPntPos$ 和三维空间坐标 $d3CurPntCoord$ 。

c) 以 $d3CurPntPos$ 各轴向值与该值向下取整值 ($l3Index$) 的差 ($d3Diff$) 来确定待插值点的位置属性。其中,待插点位置属性分为以下四种情况:

(a) $d3Diff$ 三轴向值均不为0,如图1中 A 点,则在正六面体内插值,即取出周围8个样本点(A 点周围8个实心点);

(b) $d3Diff$ 中两个轴向上值不为0,如图1中 B 点,则在平面内插值,即取出周围4个样本点(B 点周围4个实心点);

(c) $d3Diff$ 仅一个轴向上不为0,如图1中 C 点,则在直线上插值,即取出周围2个样本点(C 点周围2个实心点);

(d) $d3Diff$ 三轴向值均为0,如图1中 D 点,则直接取 $l3Index$ 所对应的样本点(D 点)。

d) 最终根据搜索得到的周围样本点估算待插点的属性值。

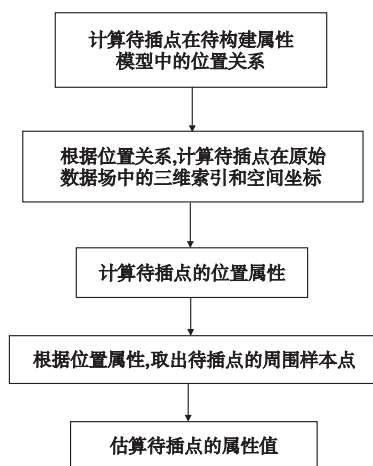


图2 结构化数据快速搜索方法流程

以 X 、 Y 、 Z 分别代表 $l3Index$ 三轴向上的值,则取出的周围样本点索引值如表1所示。

表1 待插点周围样本点三维索引描述

位置	待取点数	待取样本点三维索引描述
A	8	$(X, Y, Z), (X+1, Y, Z), (X, Y+1, Z),$ $(X, Y, Z+1), (X+1, Y+1, Z), (X+1, Y, Z+1),$ $(X, Y+1, Z+1), (X+1, Y+1, Z+1)$
		情况1 $(X, Y, Z), (X+1, Y, Z),$ $(X, Y+1, Z), (X+1, Y+1, Z)$
B	4	情况2 $(X, Y, Z), (X+1, Y, Z),$ $(X, Y, Z+1), (X+1, Y, Z+1)$
		情况3 $(X, Y, Z), (X, Y+1, Z),$ $(X, Y, Z+1), (X, Y+1, Z+1)$
		情况1 $(X, Y, Z), (X+1, Y, Z)$
C	2	情况2 $(X, Y, Z), (X, Y+1, Z)$
		情况3 $(X, Y, Z), (X, Y, Z+1)$
D	1	(X, Y, Z)

注:位置 A 对应于图1中的 A 点, B 、 C 、 D 同理。

2.2 分块建模机制

随着地球物理勘探技术的不断提高,需要处理的物探数据规模以爆炸式的速度增长。一个典型的物探数据属性模型通常包含数百层,每层具有数百万个格网点。若每个格网点属性值由一个 8 Byte 的数表示,则一次性构建数十亿字节 (GB) 规模的属性模型往往限于系统硬件瓶颈而导致建模失败^[13]。因此,本文采用分块建模机制,以适应当前大规模的物探三维属性模型的构建。分块建模是将原始采样数据均匀划分为较小的子数据块,对每一子块单独进行插值以构建相应的模型块,并加入合理的索引策略和内存调度机制,将各个子模型块组合成为完整的属性模型,实现了内存的合理调度和高效使用。算法流程如图 3 所示。

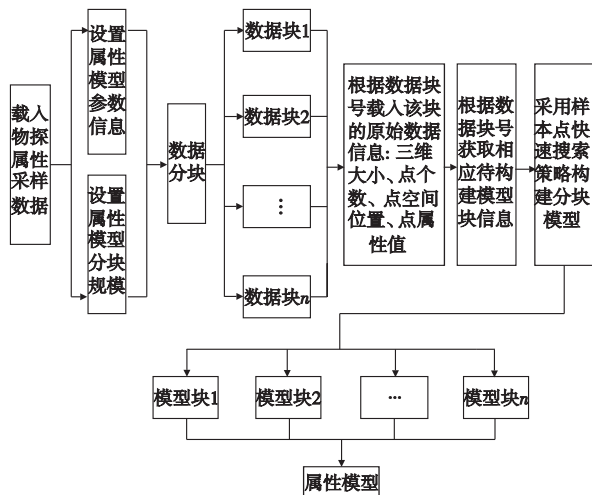


图3 物探三维属性模型分块建模流程

在分块建模机制中,只申请一个子块的内存,建模时依次加载各个子块,使得在整个建模过程中系统内存只驻留一个子块,从而克服了物探三维属性模型构建过程中内存受限的问题,避免了 PC 硬件的瓶颈效用。

数据块大小的设置需要根据具体的机器配置而定:若分块大小设定过大,可能仍超出内存可承受范围;若分块大小设定过小,则将耗费更多的时间用于数据调度。而且必须使当前块的起始位置应与前一块的终止位置严格重合,以确保各分块模型在拼接时边界的连续性。

3 应用实例

本文采用两个河北某地区磁法 $-3000 \sim 0$ m 数据作为测试数据,其 X 、 Y 、 Z 轴向的原始样本点规模分别为 $23 \times 51 \times 21$ 和 $361 \times 297 \times 21$; X 、 Y 轴向等间距采样,采样间距分别为 1000 m 和 100 m; Z 轴向非等间距采样, $-1000 \sim 0$ m 采样间距为 100 m, $-3000 \sim -1000$ m 采样间距为 200 m。在主频 2.70 GHz、内存 2.0 GB 的 Dual-Core E5400 计算机上进行了实验。选取规模为 $23 \times 51 \times 21$ 的样本数据,通过与 GMS 5.0、3D Surfer 两款三维地质建模常用软件中的距离反比加权法进行对比,验证本文提出的结构化数据快速搜索方法的高效性。实验结果如表 2 所示。

从表 2 中的数据可以得出:结构化数据快速搜索建模方法的效率明显高于 GMS 5.0 和 3D Surfer 的距离反比加权法,且当

构建的属性格网点密度越大,效率提高越明显。由此可见,本文算法在对物探属性参量的数据点空间分布情况作深入分析的基础上,提出的结构化数据快速搜索策略是简单高效的;其次,选取规模为 $361 \times 297 \times 21$ 的样本数据,构建一系列大规模的属性模型,验证了分块机制的有效性。实验结果如表 3 所示。

表 2 不同软件建模方法运行时间

模型三维格 网点数	GMS 距离 反比加权/s	3D Surfer 全局距离 反比加权/s	结构化数据 快速搜索/s
$64 \times 64 \times 64$	95.9	99.7	0.86
$128 \times 64 \times 64$	185.4	252.8	2.15
$128 \times 128 \times 64$	367.4	508.6	4.22
$128 \times 128 \times 128$	724.9	1 013.9	8.50
$150 \times 180 \times 128$	1 203.9	1 676.9	14.47
$200 \times 200 \times 128$	1 788.9	2 480.2	20.84

表 3 不同数据量属性模型分块构建运行时间

三维格网点数	模型规模/GB	分块数量/块	建模时间/s
$512 \times 512 \times 128$	0.25	16	135.7
$512 \times 512 \times 256$	0.5	16	271.7
$1\,024 \times 1\,024 \times 128$	1	256	460.0
$1\,024 \times 1\,024 \times 256$	2	256	915.9
$2\,048 \times 2\,048 \times 128$	4	1 024	2 144.3
$2\,048 \times 2\,048 \times 256$	8	1 024	4 288.5

本文采用基于硬件加速的体绘制方法^[14,15],对由规模为 $361 \times 297 \times 21$ 的样本数据构建的格网密度为 $2048 \times 2048 \times 256$ 的属性模型进行了多级分块可视化。图 4 为该模型某级 LOD 层的可视化及其分析效果。可见基于分块机制的结构化数据快速搜索建模策略构建的物探三维属性模型的正确性。

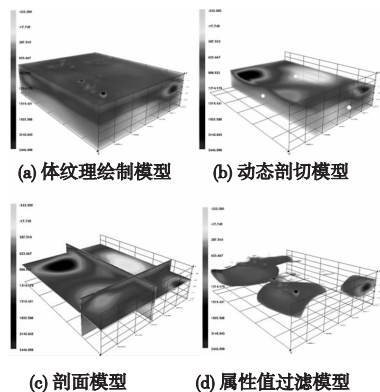


图4 LOD层可视化分析效果

4 结束语

a) 鉴于传统的距离反比加权法中样本点搜索效率低下的问题,结合物探数据属性采样资料的分布特征,提出一种结构化数据快速搜索策略,极大地提高了物探数据三维属性模型的构建效率。但需指出的是,该样本点快速搜索策略只适用于具有此类空间分布特征的样本资料。

b) 分块建模技术较好地解决了大规模数据模型无法构建的问题。但需要注意分块大小的合理确定以及利用边界数据冗余保证模型整体连续性。

c) 基于分块机制并采用结构化数据快速搜索策略的属性模型构建方法,在物探数据属性模型构建中取得了较好的成果,为三维地质属性模拟的研究提供了新思路。 (下转第 4393 页)

$O(dn^2)$ 。

4 结束语

步态识别技术是近十几年来新兴起来的生物特征识别技术,由于步态具有无须配合、可远距离获取、难以伪装等优势而受到广大生物特征研究者的青睐,并被许多政府部门、单位机构寄予厚望。基于步态图像的身份识别近年来成为计算机视觉领域中较新的研究方向。本文提出了一种监督最大差伸展的步态识别算法。实验结果表明该算法是有效可行的。进一步的工作将集中于更大规模数据库上的算法性能测试,确定所提取的特征对于视角变化的敏感度以及寻求更有效的步态图像特征提取算法和更有效的分类器。

参考文献:

- [1] 叶波,文玉梅. 基于步态的人身份识别技术综述[J]. 计算机应用, 2005, 25(11): 2577-2580.
- [2] 贾晓辉,徐森,王科俊. 行人步态的特征表达及识别综述[J]. 模式识别与人工智能, 2012, 25(1): 71-81.
- [3] LIU Ling-feng, JIA Wei, ZHU Yi-hai. Gait recognition using Hough transform and principal component analysis[C]//Lecture Notes in Computer Science, vol. 5754. [S. l.]: Springer-Verlag, 2009: 363-370.
- [4] LIU Ling-feng, JIA Wei, ZHU Yi-hai. Survey of gait recognition [C]//Lecture Notes in Artificial Intelligence, vol. 5755. [S. l.]: Springer-Verlag, 2009: 652-659.
- [5] XUE Zhao-jun, MING Dong, SONG Wei, et al. Infrared gait recognition based on wavelet transform and support vector machine[J]. Pattern Recognition, 2010, 43(8): 2904-2910.
- [6] 王斐,闻时光,张育中,等. 基于时-频分析的步态模式自动分类[J]. 北京科技大学学报, 2012, 34(1): 32-36.
- [7] 贾晓辉,徐森,王科俊. 基于 Trace 变换的步态识别算法研究[J]. 吉林大学学报:工学版, 2012, 42(1): 156-160.
- [8] 贾晓辉,安实,王健,等. 基于线性插值的张量步态识别算法[J]. 计算机应用研究, 2012, 29(1): 355-358.
- [9] 贾晓辉,王科俊,马慧. 视频下的正面人体身份自动识别[J]. 智能系统学报, 2012, 7(1): 71-76.
- [10] TAO Da-cheng, LI Xue-long, WU Xin-dong, et al. General tensor discriminant analysis and Gabor features for gait recognition[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2007, 29(10): 1700-1715.
- [11] TENENBAUM J B, De SILVA V, LANGFORD J C. A global geometric framework for nonlinear dimensionality reduction[J]. Science, 2010, 290(5500): 2319-2323.
- [12] HE Xiao-fei, YAN Shui-cheng, HU Yu-xiao, et al. Face recognition using Laplacianfaces[J]. IEEE Trans on Pattern Analysis Machine Intelligence, 2005, 27(3): 328-340.
- [13] YAN Yan, ZHANG Yu-jin. Discriminant projection embedding for face and palmprint recognition[J]. Neurocomputing, 2008, 71(16-18): 3534-3543.
- [14] BEN Xian-ye, AN Shi, MENG Wei-xiao, et al. Subpattern complete two dimensional locality preserving principal component analysis and its application to gait recognition[C]//Proc of the 6th International ICST Conference on Communications and Networking, 2011: 747-752.
- [15] WEINBERGER K Q, SAUL L K. An introduction to nonlinear dimensionality reduction by maximum variance unfolding[C]//Proc of the 21st National Conference on Artificial Intelligence and the 8th Innovation Application of Artificial Intelligence Conference. Boston: AAAI Press, 2006: 1683-1686.
- [16] BEN Xian-ye, MENG Wei-xiao, YAN Rui. Dual-ellipse fitting approach for robust gait periodicity detection[J]. Neurocomputing, 2012, 79(1): 173-178.
- [17] ETEMAD K, CHELLAPA R. Discriminant analysis for recognition of human face images[J]. Optics of Am A, 1997, 14(8): 1724-1733.
- [18] ZHAO Hai-tao, SUN Shao-yuan, JING Zhong-liang, et al. Local structure based supervised feature extraction[J]. Pattern Recognition, 2006, 39(8): 1546-1550.
- [19] YANG Jian, ZHANG D, YANG Jing-yu, et al. Globally maximizing, locally minimizing: unsupervised discriminant projection with applications to face and palm biometrics[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2007, 29(4): 650-664.

(上接第 4369 页)

参考文献:

- [1] 荆永滨. 矿床三维地质混合建模与属性插值技术的研究及应用[D]. 长沙:中南大学, 2010.
- [2] 梁双华,汪云甲,杨敏. ESDA_GIS 支持下复杂岩溶地面塌陷区特征提取[J]. 煤田地质与勘探, 2011, 39(3): 52-57.
- [3] LU G Y, WONG D W. An adaptive inverse-distance weighting spatial interpolation technique[J]. Computer & Geosciences, 2008, 34(9): 1044-1055.
- [4] 何涛,张武,李玲,等. 基于距离反比加权的地质插值方法优化研究[J]. 石油天然气学报, 2010, 32(2): 213-216.
- [5] KRIGE D G, MAGRI E J. Studies of the effects of outliers and data transformation on variogram estimates for a base-metal and a gold orebody[J]. Journal of the International Association for Mathematical Geology, 1982, 14(6): 557-564.
- [6] 徐爱萍,胡力,舒红. 空间克里金插值的时空扩展与实现[J]. 计算机应用, 2011, 31(1): 273-276, 272.
- [7] 刘海英,刘修国,李超岭. 基于地质统计学法的三维储量估算系统研究与应用[J]. 吉林大学学报:地球科学版, 2009, 39(3): 541-546.
- [8] 胡望水,熊金红,舒志恒,等. 沉积微相随机模拟方法对比研究及应用[J]. 工程地球物理学报, 2012, 9(1): 108-113.
- [9] 王英伟,张建民,王满,等. 基于序贯指示模拟方法的火山岩储层岩性及孔隙度模拟[J]. 吉林大学学报:地球科学版, 2010, 40(2): 455-460.
- [10] 王小利. 基于 IDL 的矿体三维可视化开发与研究[D]. 昆明:昆明理工大学, 2008.
- [11] 陈学工,赵灿. 基于 Java3D 的三维栅格生成与可视化[J]. 计算机应用研究, 2009, 26(7): 2759-2761.
- [12] 陈学工,张坤. 带断层线的等值线生成方法[J]. 计算机应用研究, 2010, 27(8): 3144-3146, 3164.
- [13] 王威,周杰,王水林,等. 基于径向基函数的三维地层分块建模方法研究[J]. 岩土力学, 2012, 33(3): 939-944.
- [14] KRONANDER J, JONSSON D, LOW J, et al. Efficient visibility encoding for dynamic illumination in direct volume rendering[J]. IEEE Trans on Visualization and Computer Graphics, 2011, 18(3): 447-462.
- [15] DINESHA V, ADABALA N. Uncertainty visualization using HDR volume rendering[J]. Visual Computer, 2012, 28(3): 265-278.