

# 三维地震数据场的并行体绘制算法<sup>\*</sup>)

杨磊<sup>1</sup> 李肯立<sup>1</sup> 齐雪生<sup>1</sup> 彭成斌<sup>2</sup>

(湖南大学计算机与通信学院 长沙 410082)<sup>1</sup> (壳牌石油公司 休斯顿 77478)<sup>2</sup>

**摘要** 光线投射算法是一种应用广泛的体绘制基本算法,其存在的主要问题是绘制速度较慢。为了提高光线投射算法的绘制速度,利用光线间和光线内的并行性,结合三维地震数据场的特点和文件存储特性,提出了一种基于工作站机群的三维地震数据场的并行体绘制算法。地震油气解释实际应用表明本算法能满足地震解释的要求。

**关键词** 并行体绘制,地震数据,光线投射

## Three-dimensional Seismic-data Parallel Volume Rendering Algorithm

YANG Lei<sup>1</sup> LI Ken-li<sup>1</sup> QI Xue-sheng<sup>1</sup> PENG Cheng-bin<sup>2</sup>

(School of Computer and Communication, Changsha 410082, China)<sup>1</sup> (Southwest Freeway Sugar Land, Texas 77478, USA)<sup>2</sup>

**Abstract** Ray-casting algorithm is an elementary algorithm which is used widely and based on volume rendering, but the key problem is the lower speed of rendering. In order to accelerate the rendering speed of Ray-casting algorithm, based on the parallelism of both between the ray and inner the ray, then using the characteristic of the three dimension of seismic-data field and the special method of file memory on the condition of cluster, we propose a highly efficiency parallel volume rendering algorithm for three dimension of seismic-data field. The proposed algorithm applies whole parallel from file I/O to rendering.

**Keywords** Parallel volume rendering, Seismic-data, Ray-casting

## 1 引言

大规模数据场并行可视化是科学计算可视化中极具挑战性的一个研究领域。体绘制技术在地震勘探数据分析和科学可视化领域有着极为广泛的应用,由于其巨大的计算开销,限制了其在单机上广泛的应用,因此许多研究人员致力并行体绘制算法的研究。科学计算可视化的核心是三维空间数据场的显示。体绘制技术是科学计算可视化的一个重要手段,它是分析和描述三维数据内部结构信息的有效方法,但由于体绘制技术的计算复杂、计算量大,即使对于小规模数据,也难以在普通 PC 上满足用户交互分析及实时绘制的要求。在对可视化的实时性要求越来越高的情况下,高度复杂的计算成为限制体绘制向更深领域发展的重大障碍。引入并行技术是提高体绘制速度的主要方法<sup>[1]</sup>。

地震数据的三维可视化是科学计算可视化在地震勘探领域的具体应用。把地震勘探中获取的大量数据转换为二维三维图像,以二维切片与三维数据联合显示等手段为地质解释人员提供直观的感觉,加深对地质体属性的认识和把握,提高解释工作的精度和效率。在三维地震解释过程中,不规则地质体识别、分割与显示是一个重要的技术环节,对于地下地质构造解释判定起到关键性的作用。将分割出来的不规则地质体融入到体绘制当中,使不规则地质体与其他地质构造之间的关系一目了然,便于对复杂地质体的解释。地震勘探数据的特点:数据量大,一般是数 T 的数据;有用的数据信息比较少,反映的是层次轮廓;经过流程处理之后,用来分析的数据一般都比较有规则。

大型的专业图形处理机价格昂贵,系统扩展性差,应用范

围限制于图形应用领域。PC 集群系统用便宜的 PC 机依托高速网络构建,可作为专业图形处理机的替代。它的优势表现为:高性价比、扩展性好、使用灵活、方便升级。针对三维地震数据场在可视化的应用要求,本文提出一种基于集群的并行光线投射体绘制算法。经实际地震解释实验验证,该算法可满足地震解释的要求。

## 2 相关工作的介绍

### 2.1 体绘制介绍

20 世纪 80 年代以来,体积数据的可视化技术得到了迅速发展,提出了大量的算法,主要有基于表面和基于体元的算法。体绘制是对体积数据的直接显示方法,其常用算法是光线投射算法。在光线投射成像算法中,光线间和光线内均存在并行性。光线间的并行性粒度大,一般适合采用机群或多处理器等解决方案。光线内的并行性粒度细(如每条光线上的重采样计算、法向量的密度梯度近似等),非常适合采用 SIMD 指令并行加速技术。分割技术则可以减少不必要的空采样,进一步提高成像速度。提高绘制速度是体绘制技术的难点和热点之一。剪切-弯曲算法已使并行体绘制成为一种基本算法,该算法可以利用多个 CPU 并行计算,因而大大提高了绘制速度。采用透视投影模式的光线投影法和以图像空间为序的体绘制法,运用简单的直线步进法和平行四边形步进法代替复杂的矩阵乘法,使重采样点从观察空间转换到物体空间的变换的时间减少来提高体绘制的速度<sup>[2]</sup>。

在大规模的体绘制研究过程中,一般的文章只是考虑到绘制过程中算法的改进,大部分忽略了文件 I/O 的操作,在现在的处理过程中, I/O 已成为了一个瓶颈。对于处理一个几

<sup>\*</sup>)国家自然科学基金(90715029);教育部重点项目(05128);国际合作项目“并行地震数据处理与解释一体化系统”资助。杨磊 讲师,博士生,研究领域为分布式系统及网络、并行可视化;李肯立 教授,博士后,研究领域为并行处理、科学可视化、DNA 计算机;齐雪生 硕士研究生,研究方向为并行可视化;彭成斌 博士。

GB或是更大的文件进行体绘制,这个时间是不可忽略的。改进 I/O 的处理速度,对体绘制有着不可忽视的作用,文献[10]提出了一种 NETCDF 格式的并行 I/O 操作,而现在 NETCDF 这种数据格式在海量数据的存储有着广泛的应用。本文提出了一种在机群环境下的基于三维地震数据的并行体绘制算法,该算法可以解决 I/O 瓶颈。

### 2.2 并行体绘制算法介绍

文献[5]提出并行体绘制算法中的任务划分方法一般有两类:基于体数据场的划分和基于图像空间的划分。前一种方法中,每个子进程存储一组二维体数据切片,该方法的优点是存储消耗小,但由于在射线投射法中,每一条射线的计算可能与所有的切片相关,因此每一条射线的计算可能需要几个子进程的存储数据,增加了通信的负担,且由于旋转时的视线方法不断变化,给编程和计算带来了困难,时间开销较大。在后一种划分方法中,每个子进程对应于一个图像子块,进程间具有很高的独立性,但由于视线方向在旋转时的变化,每个图像子块都与整个体数据相关,因此通常要在各个计算节点上保留一份原始体数据,存储消耗较大,增加了硬件负担。

并行体绘制算法按照绘制的流水线方式分为三种:全图像深度合成(Sort-First),前分布拼接合成(Sort-Middle)和中分布拼接合(Sort-Last)。全图像深度合成主要开销在图素点的合成和图像传输,主要瓶颈在图像的传输与深度合成速度限制了系统性能,它适合于大型机和图形处理硬件的实现。中分布拼接合主要开销是在预变换,主要瓶颈是在负载不平衡导致系统性能变差,它适合于工作站机群方式实现。中分布拼接合主要开销是在几何变换和绘制之间的数据传输,主要瓶颈是在负载不平衡以及处理器数的变化,它适合于图形处理硬件加速器<sup>[2]</sup>。

按照场景数据的控制方式可分为立即模式和保留模式<sup>[3]</sup>:立即模式是客户端捕获应用程序发出的绘制指令与数据指令,根据预先设定的任务划分打包发送到不同的绘制节点。绘制节点接收客户端发送的指令包,解码后执行。保留模式是客户端将绘制指令打包发送给所有的绘制节点,绘制节点接收指令包,解码执行。立即模式的特征模型数据集中存储在客户端,优势是使用灵活,兼容性强,弱点是每绘制一帧都要通过网络传输大量的几何指令,网络带宽成为最严重的瓶颈。保留模式特征模型数据分布存储在绘制服务器端;优点是减少网络负荷,利用帧间连贯性,充分利用系统资源;弱点是使用不灵活。

文献[6]提出了在三维地震解释过程中,不规则地质体的识别、分割与显示是一个很重要的技术环节,对于地下地质构造的解释判定起到关键性的参考作用。对于分割出来的不规则地质体,将其融入到体绘制当中,使得不规则地质体与其他地质构造之间的关系一目了然,便于对复杂地质体的解释。文献[8]提出了一种在集群环境下海量数据的交互式并行直接体绘制方法。该方法采用7个绘制节点和一个显示节点,提供了每秒9帧实时的绘制方法。文献[7]提出了一种在集群环境下基于纹理的大数据集的直接体绘制算法。这些数据根据空间位置分割成不同的数据块,根据生成中间图像通过 Sort-Last 方法合成最终的图像。文献[8]提出了一种基于集群的分布式共享图像并行光线投射法。

## 3 一种并行光线投射体绘制算法

### 3.1 光线投射模型

定义 1(物质的不透明度  $O$ ) 光线从外部进入物质时,在单位长度内衰减的比例。

定义 2(物质的微分不透明度  $\Omega$ ) 从外部进入的光穿过物质时衰减的比率,即,在一定无穷小的距离  $da$  内,只有  $(1-\Omega(\delta))d\delta$  部分的光线穿过。若不透明度  $\Omega$  为常数,则光线穿过厚度为 1 的物质时,物质的不透明度为:

$$\Omega = \log\left(\frac{1}{1-O}\right) \quad (1)$$

定义 3(物质的发光强度  $I$ ) 物质在单位长度内所发出的光。

定义 4(物质的微分发光  $E$ ) 物质发光的比率,即在一个无穷小的距离内所发出的光。

$$\frac{dT}{dz} = -\Omega(z)T(z) \quad (2)$$

$$\frac{dI_c}{dz} = -\Omega(z)I_c(z) + E_c(z) \quad (3)$$

$T(0)=0$  且  $I_c(0)=0$ , 设  $d$  为顶点间的距离,颜色  $c$  的累加光强为  $I_c(d)$ ,不透明度为

$$O_{cum} = (1-T(d)), \text{根据微分方程可解得:} \\ C_c = C_a + C_b(1-\alpha_a) \quad (4)$$

$$\alpha_c = \alpha_a + \alpha_b(1-\alpha_a) \quad (5)$$

$C_c$  和  $\alpha_c$  分别为合成的颜色和透明度。

### 3.2 光线投射算法的缺点

光线投射法最主要的问题就是不能按照物理存储的顺序来存取体数据,因为视线方向是任意的,导致视线与体素可以以任意方向相交。任何与视点相关参数的改变,都会引起光线投射算法的完全重绘制。这样就不得不花大量的时间去计算采样点的位置,及其属性值。光线投射法存在的另一个问题是对内存的需求量大。由于算法对数据存取的任意性,使具有空间数据结构的体数据阵列无用武之地。因此在集群计算机上采用平行光线投射算法来提高速度。

### 3.3 三维地震数据场的并行体绘制框架。

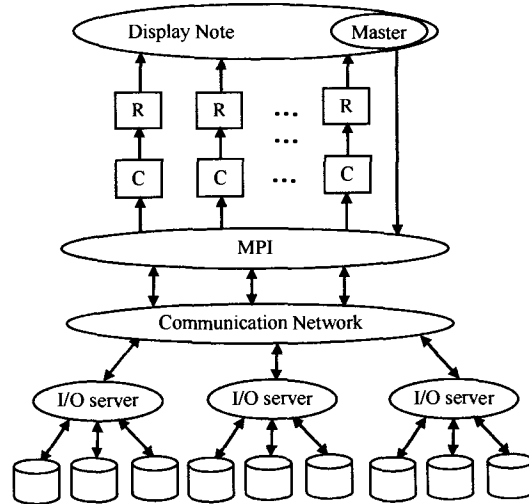


图1 并行框架

在图1中描述了并行框架。底层是文件服务器,用来存储文件,各个文件服务器通过局域网连接。MPI层包括并行 NETCDF 库负责并行文件读操作。每个子进程读取文件后进行计算,把结果作颜色映射,处理后的数据传给主进程,最后通过主进程把图像显示出来。本文采用专用的 SEG Y 和 NETCDF 文件格式文件服务器,SEG Y 文件的特点是接道的

顺序连续存储,每道数据的大小都是由 240 个字节作为道头和后面的道数据存储在一起。NETCDF 数据格式是把所有的道头信息放在文件的前面,后面存储的是每道的数据,通过道头索引能够快速读到某道数据,在它的文件头里存储每道数据的相关信息。因此这两种文件格式都能很方便地并行读取数据。在文献[10]里详细介绍了 NETCDF 格式的并行读写方式,而且提供了相应的并行读写库。大多数 SEGY 数据格式的文件通过地震勘探各种流程处理后写成 NETCDF 文件格式,因此对文件进行体绘制时,只对 NETCDF 这种文件格式进行处理。由于这种数据有很好的并行性,因此它在并行体绘制时有着很高的加速比

程序通过一个 Master 节点启动。具体运行如图 2;程序初始化,分一个 Master 主进程和多个 Slave 子进程,由主进程初始化参数,参数数据信息包括每个子进程从那道数据读起,每次读取多少道。主进程初始化参数之后就处于等待状态,每个子进程获取主进程的参数,读取文件放在自己的内存中,当多个节点同时运行时,每个节点的读取的文件不会很大,而文件是通过高速的以太网传输,因此读取文件的速度不会受到影响。每个子进程根据读取的文件进行快速光线投射,虽然快速光线投射算法会对图像的质量有一定的损失,但对于地震勘探数据来说,它反映的是地层断层与断层之间的关系,对图像的精度要求不是很高,处理的难度主要是数据量大,因此以图像的质量换取处理时间是可以接受的。

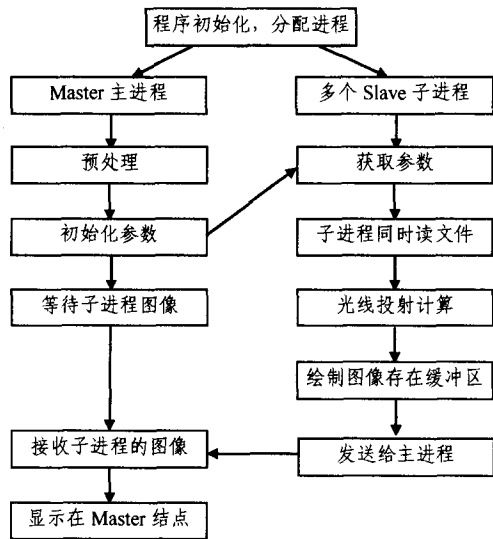


图 2 程序流程图

### 3.4 并行算法

算法 1 并行读文件算法(Parallel Read File Algorithm){

- Step 1 主进程读取文件存放地址,文件大小和文件头信息。
- Step 2 主进程获取子进程个数并根据子进程个数计算分发信息。
- Step 3 主进程分发任务给各个子进程。
- Step 4 各子进程接收主进程分发的任务开始读文件。

算法 2 Parallel Render Algorithm{

基本思想:读取数据后进行预处理与分类,同时给数据点赋颜色值和数据点赋不透明度值。通过平行发射线,重采样和光照效应计算采样点光亮度值和采样点不透明度值。合成

图像,最后显示出来。

每个子进程采用如下步骤:

- Step 1 坐标和投影变换;
  - Step 2 插值。算法中还是采用光线投射法中应用最多三线性插值法;
  - Step 3 分类与着色;
  - Step 4 明暗计算;
  - Step 5 发送数据到主进程。
- 主进程执行如下步骤:
- Step 1 主进程监听并等待子进程发送的数据;
  - Step 2 接收到子进程发送过来的,图像融合;
  - Step 3 接收完毕转到 Step 4,否则转到 Step 1;
  - Step 4 显示图像;

## 4 实验结果及其分析

### 4.1 实验环境及实验结果

本实验是在 HP P4(3G)处理器,512 内存,OpenPBS (Portable Batch System)的并行管理器和高速以太网搭建的集群系统下进行的。操作系统为 ReadHet9.0,采用 C++ 语言及 OpenGL 硬件加速库和 parallel-NETCDF 库,Qt 可视化开发工具包,数据采用 NETCDF 格式。本文对一个 512×512×512 和一个 1024×1024×1024 地震数据文件加以实现,通过实验可以得到预期的结果。通过实验得到如下图片。

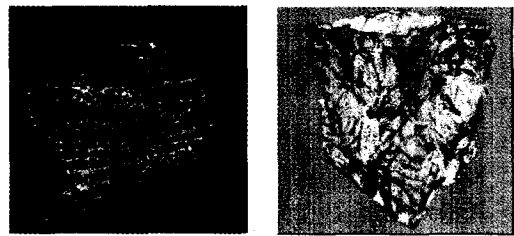


图 3 实验结果

### 4.2 加速比分析

本实验在工作站集群下得到的加速比如图 4。由于每个结点之间依赖弱、通信开销较少,因此算法的加速比能达到理想值。文件的大小对本算法影响不是很大,因此它能适合地震勘探数据的分析。

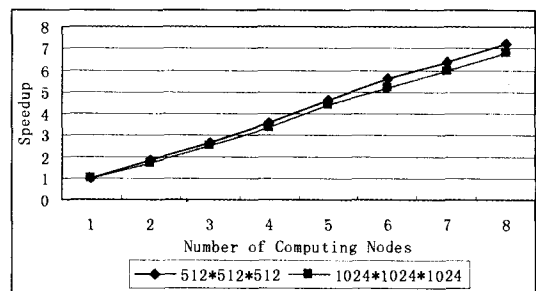


图 4 加速比分析

结束语 本文研究实现了一种基于地震勘探数据场的并行体绘制算法,该算法上能很好地解决地震勘探数据的体可视化问题,并能用集群系统解决文件 I/O 给并行带来的瓶颈,因此能有效、快速地实现三维地震数据场的体绘制。本算

(下转第 264 页)

- (6)特征区域内像素点的密度要高;  
 (7)特征模板在同类图像中所占的区域面积要大;  
 (8)模板的特征要稳定、抗干扰能力强,能容忍一定的偏差。

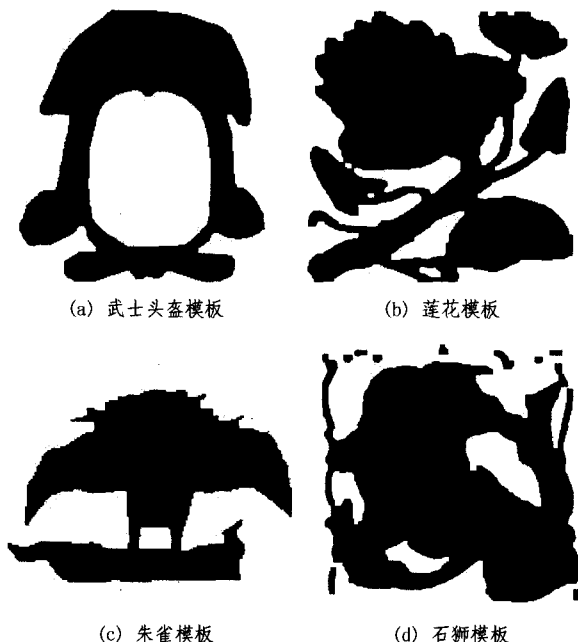


图8 川南石刻图像模板选择效果图

#### 4.3 川南石刻图像特征匹配方法

川南墓葬石刻图像的特征强烈依赖于图像本身的形状信息,几乎不涉及色彩和纹理方面的内容,因此,形状是特征提取研究的重点和关键。形状在计算机模式识别中,通常是对目标范围的二值表示,可以看作是目标的轮廓,也可以看作是目标的区域,具有突出、稳定、对光照与尺度旋转不敏感等性质。目前用于图像识别的形状描述方法主要是基于边缘的和基于区域的形状匹配方法。基于区域的形状匹配方法是通过图像分割技术在图像中提取出感兴趣的“区域”,然后根据特征“区域”内像素点的分布信息来实现图像之间的匹配对应关系,最后利用相似性测度来确定识别的结果,它适合于区域能被较为准确地分割出来的图像,是目前最具有鲁棒性的和有效的图像匹配方法。我们在川南石刻图像识别的前期预处理阶段,由于对图像的边缘轮廓进行了较好的分割,为特征提取与识别研究提供了良好的前提,因此,适合采用基于区域的形状匹配方法。该方法根据图像的区域特征建立两幅图像之间的匹配对应关系,再利用区域内部的像素点集来实现图像之

间的变换映射。该方法操作的关键在于特征提取的有效性、模板选择的精确性以及图像匹配算法与阈值的设定是否恰当等,这些都是特征提取研究必须要考虑的问题。

## 5 实验结果分析

本研究的实验环境是:操作系统是 Window2000 的 Server;具有 2GB 的内存;仿真软件是 MATLAB;测试数据集使用自建的川南地区南宋墓葬石刻图像识别库,该识别库有彩色石刻图像 300 幅,采用本文提出的特征提取方法,系统识别率达到 92.3%(如表 1 所示)。实验结果表明我们提出的特征提取方法,比基于灰度的识别方法,正确率提高了数倍,比目前人工识别的方法,准确率又有了较大的改善,达到了较好的识别效果。

表1 川南石刻图像识别方法对比表

方法排序	采用的不同识别方法	识别率
1	基于灰度的模板匹配方法	20~30%
2	人工识别石刻图像的方法	65.8%
3	基于特征提取的模板匹配方法	92.3%

结束语 ①针对川南地区石刻图像鲜明的地域特征,本文设计了川南石刻图像识别的系统结构与特征提取的方法,并对特征提取与识别中的关键环节进行了重点研究;②本文结合艺术领域研究的前期成果,分析了川南地区石刻图像的唯一特征,并通过深入的比较剖析形成了这些唯一特征的主要构成元素;③本文根据川南石刻图像的特征情况,建立了石刻图像特征模板选择的基本原则;④本文提出的特征提取方法简单、有效,大大降低了计算的复杂度,在识别中取得了较好的效果;⑤本研究是计算机特征提取研究在艺术领域应用的初步尝试,具有开拓性的意义。如何改进现有特征提取算法使识别精度提高,是进一步的研究方向。

## 参考文献

- [1] 四川省文物考古研究所编著. 泸县宋墓. 北京:文物出版社,2004
- [2] 王家文,曹宇. MATLAB6.5 图形图像处理. 国防工业出版社,2004
- [3] 张洪刚,陈光,郭军. 图像处理与识别. 北京邮电大学出版社,2006
- [4] 郭茂来,著. 视觉艺术概论. 人民美术出版社,2000
- [5] 美·欧文·洛克. 知觉之谜. 武夷山,译. 科学技术文献出版社,1989
- [6] 鲁道夫·阿恩海姆. 艺术与视知觉. 加州大学出版社,1974

(上接第 260 页)

法只在理想的情况下(每个结点读取相同大小数据,不考虑各进节点之间等待的条件)下实现。下一步考虑在复杂情况对并行算法进行研究。

## 参考文献

- [1] 李晓梅,等. 并行与分布式可视化技术及应用[M]. 北京:国防工业出版社,2001
- [2] 金哲凡,林海,石教英. 数据分布型 sort-first 并行图形绘制系统的研究与实现. 计算机研究与发展,2004,41(2):376-382
- [3] 袁非牛,廖光焯,范维澄,等. 基于光线投射算法的混合场景可视化. 中国图象图形学报,2005,10(7):850-855
- [4] 刘杰,李政,康克军,等. 高效光线投射体绘制算法研究. 计算机

辅助设计与图形学学报,2002,14(7):629-631

- [5] 马仁安,张二华,杨静宇,等. 不规则地质体的分割与体绘制方法研究. 计算机研究与发展,2005,42(5):883-887
- [6] 马仁安,张二华,杨静宇. 三维地震数据场的快速体绘制方法. 计算机辅助设计与图形学学报,2005,17(5):1008-1012
- [7] Liang K. Interactive parallel visualization of large particle datasets. Parallel Computing,2005
- [8] Strengert M. Large volume visualization of compressed time-dependent datasets on GPU clusters. Parallel Computing,2005
- [9] David E. DeMarle Memory sharing for interactive ray tracing on clusters. Parallel Computing, 2005
- [10] Jianwei, Li Wei - keng, Alok L. Choudhary Parallel netCDF: A High-Performance Scientific I/O Interface. ACM SC'03, 2003(11):15-21