利用三元组稀疏矩阵技术改进 HASM 算法
——以全球平均气温模拟为例

张 昊[1],岳天祥[2],李昌华[1],杜正平[2],范泽孟[2],孙晓芳[2],赵 娜[2]
(1. 西安建筑科技大学 信息与控制工程学院,西安 710055;
2. 中国科学院地理科学与资源研究所 资源与环境信息系统国家重点实验室,北京 100101)

摘要:采用预处理共轭梯度法(PCG)进行求解的高精度曲面建模(HASM)算法的计算过程需要大量矩阵计算,采用稀疏矩阵方式可以压缩存储空间。三元组稀疏矩阵存储是较为传统的稀疏矩阵存储结构,这种存储结构的稀疏矩阵技术已被广泛使用。根据 HASM-PCG 的特点,本文通过改进三元组稀疏矩阵的部分计算方式,调整 HASM-PCG 算法中的部分计算顺序,从而舍弃部分不需要存储的非零元素,提高了计算效率。根据全球 1998—2008 年的近 3000 个气象观测台站的气温观测数据,以及全球 DEM 数据,对 20 世纪末至 21 世纪初的 11 年来 5、6、7、8 月份的全球平均气温进行数字模拟。数值结果表明,采用改进的三元组稀疏矩阵技术有效地提高了 HASM 方法的模拟效率。

关键词:HASM;三元组;多元线性回归
DOI:10.3724/SP.J.1047.2012.00158

1 引言

高精度曲面建模(HASM)开始于 20 世纪 80 年代末,源于冰斗形态研究[1]。作为全新的建模理论,经过近 30 年坚持不懈的研究,HASM 算法基本已经发展成熟。与传统的曲面建模方法相比,如:空间自协方差最佳插值法(Kriging)、反距离加权插值法(IDW)、不规则三角网插值(TIN)、样条插值(Spline)等,HASM 算法极大地提高了插值精度[2-3]。高精度的插值结果在许多领域有着重要的意义[4-6]。随着理论研究的愈加深入成熟,我们需要及早地把理论和实践成果软件化、成品化,让更多的用户能广泛地使用 HASM 模型。

在 HASM 模型中,矩阵计算在其中占有重要的地位。在进行大规模的曲面建模时,计算过程中需要用到的很多矩阵都规模巨大,但其中非零元素的个数相对较少。大量的矩阵计算会严重降低计算效率,为了提高 HASM 模型的计算效率,采用稀疏矩阵计算技术变得顺理成章[7-8]。除了计算效率,HASM 算法运行过程中,还必须考虑其大量的内存消耗问题。在 C/C++方式中,程序员可以对内存进行完全的控制,并针对 HASM 算法中矩阵的特征设计稀疏矩阵算法和分配内存方式,从而增加程序的运行效率。

综上所述,为了更好地发挥 HASM 所具有的巨大价值,本文将三元组稀疏矩阵技术[9-10],应用到 HASM 算法,在保持计算精度的同时,提高了计算的效率。并且使用 C/C++语言编写,使系统基本能达到跨平台应用,增强其易用性和通用性。

2 稀疏矩阵建模的理论与方法

由微分几何可知,曲面可由高斯方程组决定,该高斯方程组为偏微分方程组。通过有限差分法,将该偏微分方程组离散为最小二乘问题,保证采样点处的模拟值与采样值相等。运用法方程组求解此最小二乘问题,由此导出了线性代数方程组。对于此线性代数方程组可通过预处理共轭梯度法求解[11]。

HASM 方法可以归纳为如下等式约束的最小二乘问题:

$$\begin{cases} \min \| HF^{n+1} - Q^n \|_2 \\ s.t. \quad L \times F^{n+1} = K \end{cases}$$

收稿日期:2011-12-01;**修回日期**:2012-03-20.
基金项目:国家"863"计划项目"面向新型硬件架构的复杂地理计算平台"(2011AA120305)。
作者简介:张昊(1987-),男,硕士研究生,主要研究方向为计算机数值模拟与空间数据处理。E-mail:zhanghao@lreis.ac.cn

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

其中,

$$H=\begin{bmatrix}P_1\\P_2\end{bmatrix},Q^n=\begin{bmatrix}D^n\\E^n\end{bmatrix}$$

$$P_1=\begin{bmatrix}-2\mathrm{I}_{J\times J} & \mathrm{I}_{J\times J} & & & \\ \mathrm{I}_{J\times J} & -2\mathrm{I}_{J\times J} & \mathrm{I}_{J\times J} & & \\ & \ddots & \ddots & \ddots & \\ & & \mathrm{I}_{J\times J} & -2\mathrm{I}_{J\times J} & \mathrm{I}_{J\times J}\\ & & & \mathrm{I}_{J\times J} & -2\mathrm{I}_{J\times J}\end{bmatrix}_{(I*J)\times(I*J)}$$

$$\mathrm{I}_{J\times J}=\begin{bmatrix}1 & 0 & & & \\ 0 & 1 & 0 & & \\ & \ddots & \ddots & \ddots & \\ & & 0 & 1 & 0\\ & & & 0 & 1\end{bmatrix}_{J\times J}$$

$$P_2=\begin{bmatrix}\mathrm{B}_{J\times J} & & \\ & \ddots & \\ & & \mathrm{B}_{J\times J}\end{bmatrix}_{(I*J)\times(I*J)}$$

$$\mathrm{B}_{J\times J}=\begin{bmatrix}-2 & 1 & & & & & \\ 1 & -2 & 1 & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & 1 & -2 & 1 & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & 1 & -2 & 1\\ & & & & & 1 & -2\end{bmatrix}_{J\times J}$$

经过进一步转化,可以把 HASM 所要解决的问题集中于解方程组[12]:

$$Ax=b \tag{1}$$

其中:

$$A=\begin{bmatrix}H^T & \lambda L^T\end{bmatrix}\begin{bmatrix}H\\ \lambda L\end{bmatrix} \tag{2}$$

$$b=\begin{bmatrix}H^T & \lambda L^T\end{bmatrix}\begin{bmatrix}Q^n\\ \lambda K\end{bmatrix} \tag{3}$$

因为显然矩阵 A 是对称正定的,所以 A 是收敛的。

在实际应用中,所要处理的数据量往往是非常巨大的,之前的研究,如 HASM-PCG、HASM-AM、HASM-MGS 等策略用于提高速度,通过大量的数据验证,这些方法都取得了一定程度的效果。

由于采用 HASM-PCG(预处理共轭剃度法)技术求解代数方程组。迭代求解过程中主要的计算稀疏矩阵 A 与向量 x 的乘积,公式(1)这一过程也就是包含着大量 M^TMx。根据 HASM 模型中稀疏矩阵 M 的特殊结构,使用改进三元组的稀疏矩阵技术,对一些没有计算必要的非零元素进行省略,从而减小了计算过程中的存储量,增加了计算速度。

计算顺序包括:

(1)读入原始数据:首先,读入一个有初值的,包含整个模拟区域的栅格数据文件。然后,读入原始的通过文本文件存储的采样点数据,在读取过程中需要把采样点中坐标超出模拟区域的值剔除。

(2)确定采样点权重:λ 为采样点权重系数,随地形复杂度变化而变化,对于平坦区域该值较大,对于复杂多变区域该值较小。通过选择,所有采样点的权重一样统一为 λ,或者为每个采样点与周围点均值的差。

(3)生成采样向量与采样矩阵:确定方程(1)中的系数矩阵 A。

(4)有限差分离散高斯方程:确定方程(1)中的右端项 b。

(5)迭代求解代数方程组:求解方程(1)。

(6)输出插值结果:将最终的迭代求解结果输出为栅格数据的形式。

3　算法的实现过程

在生成采样向量和采样矩阵模块中存在着大量的稀疏矩阵运算。包括稀疏矩阵转置、矩阵与矩阵相乘、矩阵与向量相乘等。对于稀疏矩阵,这里采用了比较简单通用的三元组方式存储,数据结构定义如下:

```
typedef int Status;
typedef double ElemType;
typedef struct{
    int i,j;          //行列号
    ElemType e;       //元素值
}Triple;
struct TSMatrix{
    Triple *data;     //三元组表
    int mu,nu,tu;     //总计行列数和元素个数
};
```

在进行稀疏矩阵转置的运算时,设计函数:Status TransposeSMatrix(TSMatrix M, TSMatrix &T,int n)。由于行号小的元素在三元组序列中的排序是靠前的,所以,我们按三元组中元素的列号

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

大小进行检索。算法从最小列号开始检索，直到最大列号[13]。每次首先检索到的元素，就一定会是行号也小的元素。

在迭代求解模块中，需要计算一个稀疏矩阵的转置 M^T 乘以它本身 M 再乘以向量 x(M^TMx)。如果使用一阶截断方式进行计算的话，M^TM 的最终的结果矩阵中除主对角线外的其他部分的值都是固定且事先知晓的，所以不需要存储，只有主对角线上面的有值元素需要存储下来。基于这一特点，这里设计了如下算法来简化计算：

首先，通过循环计数的方式统计出稀疏矩阵序列中存有多少行非零元素，以此来确定结果稀疏矩阵序列中存在多少个非零元素。根据非零元数量申请内存空间完毕后，第一步搜索存放稀疏矩阵的三元组序列，如果当前元素和下一个元素的行号一致，则平方这个元素值。再查看下一个元素跟第三个元素的行号是否一致，若依旧一致则平方后，将新得出的平方值累加到前面的平方值上。然后，再查看第四个元素，依此类推，直到行号不一致为止。最后，便是这个累加值乘积稀疏矩阵中，那一行上主对角线所对应的值。保存下来后，继续用上述方法计算下一个行号的元素，直至最终得到计算所需要的乘积稀疏矩阵。

在采用三阶或者五阶截断方式时，稀疏矩阵 M^TM 中需要存储的数据不再只集中在主对角线上，如果继续使用上述算法，将会大大提高算法复杂度，增加内存消耗和时间消耗。对于这种情况而言，我们可以通过改变计算顺序来达到相同的效果。

因为 $(M^TM)\cdot x=M^T\cdot(Mx)$，所以，先计算 Mx，再将其与 M^T 相乘，使原计算顺序变为 $[M^T(Mx)]$。通过进行两次稀疏矩阵与向量的乘积运算，来降低算法复杂度。

对于稀疏矩阵与向量的乘积运算，在生成采样向量和采样矩阵模块和迭代求解代数方程组模块中都有多次使用。

从第一个元素开始，逐个查看稀疏矩阵的三元组序列元素，记录行号。通过三元组中的列号来寻找所乘向量中的对应元素，并将二者相乘。然后查看三元组序列中的下一个元素，如果下一元素与当前元素的行号相同，根绝此元素的列号找到向量中的对应元素相乘，并把乘积累加到上一乘积后，再查看下一元素。直到当前元素与上一元素行号不同为止，此时的累加值便为结果向量中，此行号对应位置的值。以此类推，对当前元素进行上述操作，直到整个三元组序列遍历为止，我们就得到了新的结果向量。

4 模拟气温的数值实验分析

本文使用美国国家气候中心(National Climatic Data Center)的全球气候数据(Monthly Climatic Data for the World)中的气温数据作为基础数据，以美国 USGS 90m 的 DEM 数据(http://data.geocomm.com/dem/)作为模拟的辅助数据。主要步骤如下：

首先，运用 VC++编程，对全球气候数据中3013个台站的记录进行统计计算，求得 11 年(1998-2008年)中每个台站，在 5-8 月的平均气温数据。然后，以大洲为基本地域特征分区，根据经验选定台站记录中的经度坐标、纬度坐标、高程 DEM 作为因子，对这 3 个因子与全球地面温度进行回归分析，得到定量关系式[14-19]。最后，运用 HASM 方法进行插值运算。

需要说明的是，由于 HASM 是以平面直角坐标系为理论基础提出的，所以，在模拟开始前，需要一些特殊的准备工作来保证最终模拟结果的正确，包括：首先，对世界地图进行保直角的墨卡托投影，从而将经纬网的球面坐标系转换为直角坐标系。其次，在直角坐标系中分区，把具有相同或相似地域气候特征的台站观测点组织在一起。最后，再对台站数据中的经纬度球面坐标(Lat，Lon)，通过墨卡托投影转换为平面直角坐标(X，Y)。

分区过程中，考虑到格陵兰岛、阿拉伯半岛，以及澳大利亚的地域和气候的特殊性，选择对上述地域独立分区。考虑到没有明显的气候分界线，没有对亚欧大陆进行更加细致的分区。另一方面，由于墨卡托投影所使用的投影为 WGS1984 方式，这种投影方式使一些本来在地理环境中连续的区域人为的造成了割裂。为了保证多元线性回归的准确性，针对线性回归过程，对位于白令海峡以东小部分地区或者位于东太平洋诸岛的采样点的坐标(X，Y)行进了调整，使之能够完整准确地表达地理位置

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

关系。经过对比试验可以发现,经调整后的坐标参与线性回归的效果比没有调整的更好。

线性回归结束后分别得到不同年份各分区的常数项(k_0)、Y 系数(k_1)、X 系数(k_2)、DEM 系数(k_3)以及相关系数(R)和显著性检验参数。如表 1 至表 4 所示。

表 1　5 月份全球分区线性回归参数表

Tab. 1　Linear regression parameters for the subareas of the globe in May

分区	常数项(k_0)	Y 系数(k_1)	X 系数(k_2)	DEM 系数(k_3)	相关系数(R)	显著性检验参数
亚欧大陆	3.6086E+01	−3.2472E-06	−1.5366E-07	−3.4646E-03	9.2138E-01	0.0000E-00
格陵兰岛	1.7302E+01	−1.5078E-06	4.3219E-07	−8.8415E-03	9.4744E-01	2.9087E-05
北美洲	2.9541E+01	−3.9745E-06	−7.0748E-07	−3.0479E-03	9.5360E-01	5.4906E-231
南美洲	3.4927E+01	3.1668E-06	1.1955E-06	−3.0110E-03	9.5245E-01	1.9264E-187
太平洋诸岛	2.6966E+01	1.9498E-06	−1.3356E-08	−5.3272E-03	8.0516E-01	4.0885E-50
南极洲	3.3482E+01	4.2495E-06	−1.8700E-07	−9.7286E-03	9.7044E-01	4.1398E-24
澳大利亚	3.8131E+01	4.1503E-06	−4.1647E-07	−5.6350E-03	9.7231E-01	5.7156E-63
非洲	2.7103E+01	1.1279E-07	−8.9569E-09	−5.3372E-03	4.1829E-01	3.7838E-16
阿拉伯半岛	2.7503E+01	−3.0489E-06	2.4292E-06	−2.7188E-03	7.8801E-01	9.6259E-15

表 2　6 月份全球分区线性回归参数表

Tab. 2　Linear regression parameters for the subareas of the globe in June

分区	常数项(k_0)	Y 系数(k_1)	X 系数(k_2)	DEM 系数(k_3)	相关系数(R)	显著性检验参数
亚欧大陆	3.4474E+01	−2.4271E-06	−2.8072E-08	−2.6142E-03	8.5612E-01	0.0000E-00
格陵兰岛	1.0646E+01	−5.5343E-07	1.0589E-07	−7.4438E-03	8.2487E-01	7.7279E-03
北美洲	3.1070E+01	−2.9817E-06	−4.4926E-07	−2.8742E-03	8.9782E-01	4.7633E-158
南美洲	3.5989E+01	3.4352E-06	1.4182E-06	−3.1956E-03	9.5639E-01	1.6124E-192
太平洋诸岛	2.6693E+01	2.3552E-06	−2.7203E-08	−5.5736E-03	8.6211E-01	1.3155E-62
南极洲	3.1966E+01	4.3086E-06	−1.0457E-07	−9.4280E-03	9.7348E-01	1.1924E-25
澳大利亚	3.1882E+01	4.0021E-06	−2.1315E-07	−7.5102E-03	9.5028E-01	1.6339E-50
非洲	2.5580E+01	1.2221E-06	1.8340E-07	−4.5836E-03	6.2014E-01	6.8235E-42
阿拉伯半岛	2.5929E+01	−1.6220E-06	2.3117E-06	−1.5949E-03	6.2685E-01	3.3147E-08

表 3　7 月份全球分区线性回归参数表

Tab. 3　Linear regression parameters for the subareas of the globe in July

分区	常数项(k_0)	Y 系数(k_1)	X 系数(k_2)	DEM 系数(k_3)	相关系数(R)	显著性检验参数
亚欧大陆	3.4341E+01	−2.0853E-06	9.0122E-09	−2.7247E-03	8.3958E-01	0.0000E-00
格陵兰岛	1.0763E+01	−3.6848E-07	−4.0188E-08	−1.0136E-02	7.1399E-01	5.8849E-02
北美洲	3.2020E+01	−2.3794E-06	−2.5749E-07	−2.3898E-03	8.4514E-01	1.5505E-119
南美洲	3.5301E+01	3.4978E-06	1.3179E-06	−3.0378E-03	9.4740E-01	1.1731E-180
太平洋诸岛	2.6450E+01	2.4848E-06	−3.1609E-08	−5.1736E-03	8.8462E-01	2.7650E-70
南极洲	3.1667E+01	4.3624E-06	−8.3909E-08	−9.4271E-03	9.7352E-01	1.1534E-25
澳大利亚	3.2358E+01	3.9859E-06	−2.8892E-07	−6.8191E-03	9.5868E-01	5.6007E-55
非洲	2.4580E+01	1.8048E-06	2.6973E-07	−3.7391E-03	7.5748E-01	5.6289E-73
阿拉伯半岛	2.7406E+01	−7.5506E-07	1.7077E-06	−2.0225E-03	4.9509E-01	1.0812E-04

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

表 4 8 月份全球分区线性回归参数表

Tab. 4 Linear regression parameters for the subareas of the globe in August

分区	常数项(k_0)	Y 系数(k_1)	X 系数(k_2)	DEM 系数(k_3)	相关系数(R)	显著性检验参数
亚欧大陆	3.5851E+01	−2.4263E-06	6.9937E-09	−3.4251E-03	8.8996E-01	0.0000E-00
格陵兰岛	1.4767E+01	−7.2333E-07	1.3190E-07	−5.9956E-03	9.0640E-01	4.5180E-04
北美洲	3.3076E+01	−2.5954E-06	−2.3416E-07	−3.0046E-03	8.8679E-01	4.7744E-148
南美洲	3.6930E+01	3.3417E-06	1.4654E-06	−2.8855E-03	9.2883E-01	2.1355E-157
太平洋诸岛	2.6524E+01	2.3657E-06	−3.4225E-08	−5.6437E-03	8.7138E-01	3.9152E-67
南极洲	3.5989E+01	4.8796E-06	−9.9839E-08	−8.8736E-03	9.8372E-01	4.2194E-29
澳大利亚	3.3773E+01	4.1981E-06	−2.6884E-07	−5.3447E-03	9.6540E-01	8.5934E-59
非洲	2.4383E+01	1.6482E-06	3.2642E-07	−3.3801E-03	7.6043E-01	3.9003E-73
阿拉伯半岛	2.8329E+01	−4.6205E-07	1.3455E-06	−1.7118E-03	3.8489E-01	7.5800E-03

经过相关性分析，可以看到在全球范围内气温与高程(DEM)，转换后经度(X)、纬度(Y)具有明显的相关性。显著性检验参数远远小于显著性水平 0.05，可以认定符合回归要求。公式表示为：

$$Tem(x,y,dem) = k_0 + k_1 * Y(i) + k_2 * X(i) + k_3 * Dem(i);$$

确定回归成功后，使用气象台站记录的气温数值 $Tem(n)$ 与通过线性回归取得的趋势面的气温值 $Tem(x,y,dem)$ 作差，得到采样点的残差数据 $I(n)$。

$$I(n) = Tem(n) - Tem(x,y,dem);$$

然后根据采样点残差数据进行 HASM 插值，取得插值曲面。对残差面进行返回归，加上前一步中的趋势面，得到全球气温分布。因为全球 DEM 数据缺少南极部分，以及海洋气候与陆地气候决定因素的极大差异，所以，本文的模拟区域选取了包括除海洋与南极洲的所有地区。但采样点数据中包含有气象台站的 DEM，所以在趋势面模拟的过程中地域分布连续，南极洲的减少未对最后结果产生影响。

最后对全球气温分布图进行反墨卡托投影，得到常用的球面坐标下的最终结果。如图 1-4 所示。

全球气温分布的一般规律为：因太阳辐射的分布是从低纬向高纬递减；故地形因素的影响同纬度高原、山地的气温比平原、低地的气温低。

可明显看到，随着时间推移，5 月到 8 月全球的高气温带向北明显推移，北半球进入盛夏，南半球温度逐渐变低，并且格陵兰岛、西伯利亚等北极地区持续低温，与自然规律一致。因为通过残差进行模拟得出残差面后再添加趋势面影响，所以，保证了在曲面整体分布较为光滑平顺的前提下，充分考虑局部地区地形高程和经纬度的影响，细节特征也得以体现。

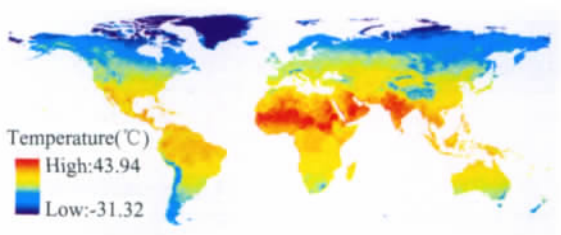

图 1 5 月全球平均气温模拟

Fig. 1 Global mean temperature simulation in May

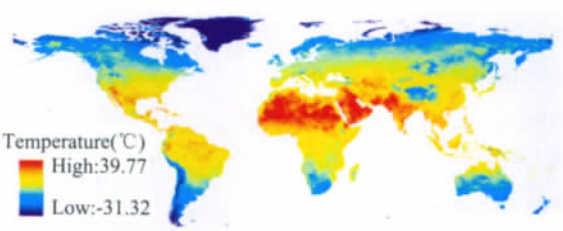

图 2 6 月全球平均气温模拟

Fig. 2 Global mean temperature simulation in June

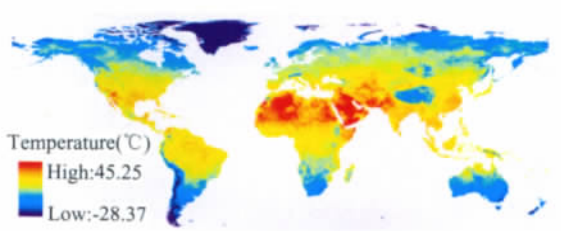

图 3 7 月全球平均气温模拟

Fig. 3 Global mean temperature simulation in July

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

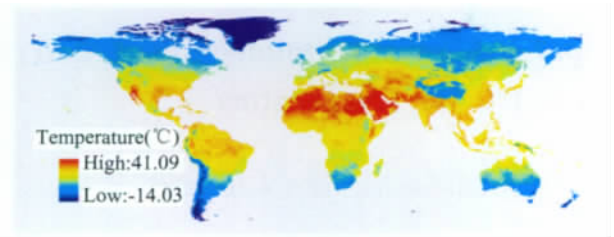

图 4　8 月全球平均气温模拟

Fig. 4　Global mean temperature simulation in August

5　结论与展望

模拟选择全球 25km 分辨率,矩阵规模 1603×1080,内迭代次数 10 000 次,外迭代 5 次。因为模拟过程中没有在数学层面改变计算原理,所以,最终结果的精度并没有提高,保持了与原始版本程序结果的一致。但是实验证明,由于采用三元组稀疏矩阵存储技术,计算过程中所要存储的数据量减少,从而使整个模拟的区域更大、分辨率更高。并且因为计算过程中计算方式有所简化,模拟同样规模的时间也有所降低。

尽管经过优化,计算过程在迭代求解部分的时间耗费是最大的,少则几千次多则几十万次的高密度简单数字计算迭代过程如果能够采用 GPU 并行的方式计算,那么可以预见的时间消耗将会大大的减少,这也是下一步的工作方向。可以预见,在提高计算速度之后,一个高精度、高速度的建模方法必将得到更为广泛的认同和应用。

参考文献:

[1] 岳天祥,艾南山. 冰斗形态的数学模型[J]. 冰川冻土,1990,12 (3): 227-234.

[2] 岳天祥,杜正平. 高精度曲面建模与经典模的误差比较分析[J]. 自然科学进展,2007,16(8):986-991.

[3] Yue T X,Du Z P,Song D J,*et al*. A new method of surface modeling and its application to DEM construction [J]. Geomorphology,2007,91(1-2): 161-172.

[4] 范泽孟,岳天祥,宋印军. 基于 YUE-HASM 方法的气温与降水时空变化趋势[J] . 地理研究,2009,28(3): 643-652.

[5] Yue T X,Fan Z M and Liu J Y. Scenarios of land cover in China[J]. Global and Planetary Change,2007,55: 317-342.

[6] Yue T X,Fan Z M. and Liu J Y. Changes of major terrestrial ecosystems in China since 1960[J]. Global and Planetary Change,2005,48: 287-302.

[7] 岳天祥,杜正平,刘纪远. 高精度曲面建模与误差分析[J]. 自然科学进展,2004,14(2): 83-89.

[8] 岳天祥,杜正平,宋敦江. 高精度曲面建模:HASM4. 中国图像图形学报,2007,12(2): 343-348.

[9] 殷人昆. 数据结构(用面向对象方法与 C++语言描述)[M]. 北京:清华大学出版社,2007,43-181.

[10] 严蔚敏,吴伟民. 数据结构[M]. 北京:清华大学出版社,2001,1.

[11] 陈传法. 高精度曲面建模(HASM)自适应算法研究[D]. 中国科学院地理科学与资源研究所,2010.

[12] 岳天祥,杜正平. 高精度曲面建模:新一代 GIS 与 CAD 的核心模块[J]. 自然科学进展,2005,15 (4) : 423-432.

[13] 徐光联. 三元组稀疏矩阵快速转置 C 语言算法[J]. 决策管理,2008,5:92-93.

[14] Yue T X. Surface modeling: High accuracy and high speed methods [M]. CRC Press Taylor & Francis Group,2011,474-477.

[15] 范泽孟,岳天祥,陈传法,孙晓芳. 中国气温与江水的时空变化趋势分析[J]. 地球信息科学学报,2011,13(4):526-533.

[16] 朱永禔. CO_2 等温室气体对全球气候影响的数值模拟研究[J]. 上海环境科学,1990(9):33-36.

[17] 丁裕国. 全球气候变化综述(II):最新模拟结果及其评价[J]. 气象教育与科技,2001,1:6-9.

[18] 蔡福,于贵瑞,祝青林,何洪林,刘新安,李正泉,郭学兵. 气象要素空间化方法精度的比较研究——以平均气温为例[J]. 资源科学,2005,5(27):173-179.

[19] 宋敦江,岳天祥,杜正平,散乱数据插值的 HASM 方法[J]. 地球信息科学,2007,9(3): 45-51.

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

The Application of Modified Three-Tuple Sparse Matrix Technology in HASM: A Case Study in the Simulation of Global Temperature

ZHANG Hao[1], YUE Tianxiang[2], LI Changhua[1], DU Zhengping[2], FAN Zemeng[2], SUN Xiaofang[2] and ZHAO Na[2]

(1. *College of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an* 710055, *China*; 2. *State Key Laboratory of Resources and Environment Information System, Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing* 100101, *China*)

Abstract: In terms of the fundamental theorem of surfaces, High Accuracy Surface Modeling (HASM) method, which is based on the differential geometry theory, transforms the surface modeling to a linear system by using the principle of least square method and describes the properties of the surface. Comparing with the classical surface methods in GIS application, such as inverse distance weight (IDW), triangulated irregular network (TIN), KRIGING and SPLINE, HASM has a much higher accuracy and the error problem of the surface modeling is solved in theory by HASM. However, the computational cost of HASM is high. The cost of HASM algorithm is expensive when it applies preconditioned conjugate gradient method (PCG). Using compressed storage formats can effectively save memory. Among these formats, the three-tuple storage format of the sparse matrix is widely used in the past. In this paper, based on the characteristic of HASM-PCG, we modify a large part of the sparse matrix-vector multiplication by using three-tuple technology. The computational efficiency is enhanced though adjusting the calculation order of HASM-PCG and abandoning some nonzero elements that are not need to store. Long term (1998-2008) averaged monthly data were obtained from site measuring temperature (3000 sites). Combining with the digital elevation model of the globe, the data are used to simulate the global mean temperature in May, June, July, and August of the eleven years from 1998 to 2008. Numerical results show that with this three-tuple technique, the computing efficiency of HASM is great improved.

Key words: HASM; three-tuple technique; multiple linear regression

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net