

基于局部熵的高度场三角格网化研究

明德烈^{1,2}, 尤克非², 田金文², 柳健²

(1. 南京大学计算机软件新技术国家重点实验室, 江苏南京 210093;

2. 华中科技大学图像所, 图像信息处理与智能控制教育部开放实验室, 湖北武汉 430074)

摘 要: 不规则三角格网表示以其优越的性能在地形高度场数据的简化中得到了广泛的应用. 在不规则三角格网生成的过程中, 插入点的选取是关系到最终简化后的地形逼近质量好坏的关键. 由于我们通常对原始格网上的每一个点给出一个重要性度量, 并以该度量值作为插入点选取的依据, 因此重要性度量值计算方法的选取成为研究的重点. 本文在对传统重要性度量方法的性能比较和分析基础上, 提出了一种基于局部熵的重要性度量方法. 实验结果表明, 基于局部熵的插入点选取方法在各项性能上均优于局部误差估计的方法, 并且有效的克服了以往方法中的“短视性”问题.

关键词: 高度场; 地形简化; 不规则三角格网 (TIN); 重要性度量值; 局部熵

中图分类号: TP391.9 **文献标识码:** A **文章编号:** 0372-2112 (2002) 07-1009-04

Height Field TIN Representation Based on Local Entropy

MING De lie^{1,2}, YOU Ke fei², TIAN Jin wen², LIU Jian²

(1. State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, Jiangsu 210093, China;

2. Key Laboratory for Image Processing and Intelligent Control of the Education Department,
Institute for Pattern Recognition and Artificial Intelligence, HUST, Wuhan, Hubei 430074, China)

Abstract: Because of its predominant capability, TIN (triangulated irregular network) has been widely used in the simplification of terrains represented as height field. During the generation of TIN, selection of input points is the key to the approximation quality of simplified terrains. Usually, the input point is selected according to the importance measure on each point, so research on the calculation of importance measures has attained more and more attention. Based on analysis and comparison of traditional importance measures, we put forward a new importance measure based on local entropy. The results demonstrate that the local entropy criterion has a better performance than any traditional method. In addition, it can effectively conquer the “short sight” problem in previous methods.

Key words: height field; terrain simplification; triangulated irregular network (TIN); importance measures; local entropy

1 引言

在类似于三维场景重建、虚拟战场仿真以及地形导航这样的系统中, 我们经常需要对场景进行建模. 这里的建模概念有别于计算机图形学中的实体建模, 它主要是对场景中的地形表面进行建模. 通过构造对地表高程数据的逼近估计, 使得能够用最小的数据量, 尽可能准确地逼近真实地形的高程模型, 从而为下一步的三维场景的显示和交互奠定基础.

地形(高程)数据是在一个平面域上对高度的一系列采样所获得的一个高度场 (Height Field). 设 H 是在地形表面二维方向上的一个离散采样点的集合, 可以用一个离散函数 $H(x, y)$ 来表示, $Z = H(x, y)$ 表示的点 (x, y, z) 位于实际的地形表面上. 由于绝大多数图形硬件和 3D 引擎 (如 OpenGL) 用三角形作为描述物体的基元, 为了充分利用硬件的加速性能, 我们采用三角形网格元素来代表地形, 即通过三角划分的方法来重建地形表面. 设重建算子 T 为将一个定义在离散样点集上的函数 (如 $H(x, y)$), 映射到一个定义在连续域上所有点的函数. T 能做到这样是通过在采样点上建立三角划分, 并用经过三角划分后得到的重建表面给所有不在网格上的点赋值. 如果设 S 是输入点集合 H 的某个子集, 则 TS 便是重建的

表面, $(TS)(x, y)$ 就是经过重建后的表面在点 (x, y) 处的值. 这样的三角划分逼近, 常被称为不规则三角化网, 即 TIN. 不规则三角网是用来简化和模拟高度场的一种重要手段. 根据不同的具体要求, 可以利用 TIN 进行不同细节程度上的模拟. 也就是说, 可以在高细节程度 (高准确度) 和低数据量 (高速度) 之间进行取舍^[1], 以达到快速、逼真地动态演示大范围地形高度场数据的目的.

TIN 的生成主要分为两个方面: 特征点的提取和三角网的生成. 特征点的提取主要是依赖于地形中每个顶点的重要性度量值. 每个顶点按重要性度量值排序后, 顺序取前若干个作为特征插入点参与地形三角格网逼近. 在已有的特征点基础上生成三角化格网已经有比较成熟的算法, 如 Delaunay 三角划分法^[2]等, 这些方法在国内外的一些书籍和论文中已有论述.

我们的目标便是找到采样点集合 H 的某一个子集 S , 将 S 中的点作为插入点顺序插入到地形中, 并在插入的同时进行三角划分. 最后产生的三角化格网即为真实地形表面的一个近似逼近. 由此可见, 插入点集合 S 的选取是整个 TIN 生成算法的关键. 其中, 重要性度量值作为三角网生成过程中的插

入点的选取准则,其计算方法的选取直接影响到地形三角网逼近的质量.下面我们来简单比较和分析一下目前几种比较流行的重要性度量方法.

2 几种常用重要性度量方法的性能分析和比较

正如前面提到的,一个地形 TIN 简化算法好坏的关键取决于在某点上重要性度量值的计算方法.但是评价一个计算方法优劣的具体标准又是什么呢?显然,最终的判断要取决于它产生的结果的质量.此外,它必须计算起来简单快捷;适用于各种类型的高度场信息,不具有特殊性;并且从算法运行过程中性能优化角度考虑,它应该尽量只使用邻近局部范围内的“本地”信息^[3].计算每一点的重要性度量值有多种方法,我们将逐一地简要讨论一下局部误差估计,曲率估计,全局误差估计这三种最典型的重要性度量值的计算方法,并着重分析和比较一下它们的性能.

2.1 局部误差估计

它是实际地形表面和近似估计表面在二维平面上一点 (x, y) 处垂直方向上的误差.如果设 H 为真实地形表面, S 为插入点集合, T 为重建算子, TS 便是重建的表面,则点 (x, y) 处的重要性度量值可以表示为 $|H(x, y) - (TS)(x, y)|$.直觉上,我们希望通过消除局部误差产生高质量的地表平面近似.经过实验论证,基于局部误差的方法的确起到了这样的效果.同时它还具有以下优点:计算简单快捷,具有广泛适应性,计算过程中只需要利用周围的局部信息.

2.2 曲率估计

在日常生活中,尖峰、凹洞、山脊、山谷等具有高曲率的物体随处可见.这些现象提示我们可以用曲率作为原始地形网格上点的重要性的一种度量^[4].但是,它对高频率的变化却过于敏感.一系列快速剧烈的小尺度波动(如锯齿),由于尺度很小,从一定距离之外看时根本不明显,但用该方法计算重要性度量值时,每一个波动点都可能被看作是特征插入点而被确定下来.

2.3 全局误差估计

全局误差估计的方法实际上就是将所有点上局部误差的总和作为地形点重要性的一种度量.对于每一个点,我们计算将该点加入插入点集合 S 后所产生的地表逼近的全局误差 $\sum_{x,y} |H(x, y) - (TS)(x, y)|$.然后,选取产生最小全局误差的那一点作为新的逼近点插入到集合 S 中.与基于局部误差估计的方法相比,它直接选择使全局误差最小化的点,而不是简单地假设增加具有高局部误差值的插入点会降低实际的全局误差,因此在理论上这个“更智能化”的重要性度量方法可以产生比局部误差估计更高质量的结果.但在实际应用中发现,目前的实现方法都存在两个缺陷,具体表现在:

(a) 违背了计算过程局部化的原则.每选取一个插入点对地形进行一次新的三角划分之后,对于原始网格上的所有点,都要重新计算一遍将该点加入插入点集合 S 后所产生的地表逼近的全局误差,以此作为该点重要性的度量.这样一个过程的计算量是相当大的,从而限制了该方法的实时性应用.

(b) 算法本身固有的“短视性”^[5]——目光太短浅.基于

全局误差的插入点选取方法,对于每一个插入点的选取都一味追求全局的最优,但实际上经常需要在短期内做一些“坏”的预测来达到长期内的最优.例如,假设为了精确的表现某一地形特征需要插入两个候选点进行三角划分,往往第一个插入点的选取并不一定是当前全局范围内的最佳选择,但由其带来的逼近误差却能由于第二个插入点的选取而很好的弥补,并且达到全局内的最优.

要充分体现全局误差概念的算法不仅复杂,且运算量也相当大,目前的实现方法难于将其用到地形的实时简化中.因此从实用意义上说,全局误差估计的整体性能目前尚不如局部误差估计方法,尚有待进一步的算法改进.

经过很多研究人员在实际工作中的应用发现,按照目前的算法实现条件,以上三种重要性度量方法以第一种——基于局部误差的重要性度量方法的综合性能为最优,它不仅是最容易实现的,比其它方法产生的逼近结果也更精确,而且比除了曲率估计之外的任何方法都快.但由于其一味强调局部范围内的最优,即最大程度的减小局部范围内的逼近误差,结果往往造成在某一地形局部插入点过于集中,使得其它地形区域由于没有足够的插入点,而不能得到很好的近似,因此同样存在和全局误差估计类似的“短视性”问题.

3 基于局部熵的重要性度量方法

通过前两节的论述,可以发现重要性的度量方法直接影响到地形 TIN 简化的质量.通过对几种常用重要性度量方法性能的比较和分析,我们发现其中普遍存在的“短视性”问题——在选取插入点的过程中,一味追求某一方面的最优,而忽视算法的最终整体效果.为了在一定程度上减小“短视性”问题的影响,我们尝试提出一种基于局部熵的重要性度量方法,下面是该计算方法的基本思想.

设 $Z = H(x, y)$ 为地形表面二维方向上的点 (x, y) 处的高程,点 (x, y, z) 位于实际的地形表面上.显然 $H(x, y) \geq 0$. 对于一个范围为 $M \times N$ 大小的地形区域,定义 H_f 为该地形区域的熵,即

$$H_f = - \sum_{i=1}^M \sum_{j=1}^N p_{ij} \lg p_{ij} \quad (1)$$

式中 $p_{ij} = [H(i, j) - k] / \sum_{x=1}^M \sum_{y=1}^N [H(x, y) - k]$ 为相对高程分布,其中 $k = \min\{H(x, y); x \in [1, M], y \in [1, N]\} - 1$, k 减 1 是为了防止分母为零时式(1)发生除零错误(局部高程数据完全一致——绝对平坦的情况).如果 $M \times N$ 是地形的一个局部区域,则称 H_f 为地形的局部熵.

局部熵反映了地形高程的起伏程度.在局部熵小的地方,地形起伏比较大,地形高程数据的离散性较大,局部熵大的地方,地形相对平坦.所以根据地形的局部熵可以把地形中起伏较为剧烈的目标区域分割出来,加以特殊处理.因为局部熵是由窗口内多点高程数据共同的贡献,对于单点噪声不敏感,所以,局部熵本身具有一定的滤波效果.由式(1)定义的熵涉及到对数运算,计算量大,运行速度慢.由定义可知, $p_{ij} \ll 1$, 因此可通过泰勒级数展开舍去高次项得到近似计算公式

$$H_f \approx - \sum_{i=1}^M \sum_{j=1}^N p_{ij} (p_{ij} - 1) = 1 - \sum_{(i,j) \in (M,N)} p_{ij}^2 \quad (2)$$

基于局部熵的重要性度量方法与基于局部误差的方法一样计算简单快捷, 适用于各种类型的高度场, 而且不难从计算方法看出, 它的计算过程高度局部化。更为重要的是, 基于局部熵的重要性度量方法从算法本身特性出发, 没有象基于全局误差和局部误差的方法那样一味追求某一方面的最优, 而是从全局考虑, 根据地形高程数据的起伏程度合理地分配插入点。因此, 它能够在一定程度上克服短视性的影响。

但是, 单纯基于局部熵的重要性度量方法本身仍然存在一定的不足。在计算过程中, 由于局部熵值较小的点都集中在地形起伏比较剧烈的地方, 这势必使得地形起伏比较剧烈地方的插入点过于密集, 从而使得地形起伏相对平坦的位置由于没有足够的插入点而得不到很好的近似。为了避免这种情况的发生, 我们在原算法的基础上做了进一步改进, 将它与曲率估计的方法结合起来, 通过引入一个随局部曲率变化的自适应距离因子 r , 来对插入点分布的疏密程度进行调节。

改进后的基于局部熵的重要性度量方法, 在选好插入点后并不是直接将它插入到集合 S 中进行三角划分, 而是首先对它进行合理性验证。在以当前所选插入点为中心, 半径为 r 的范围内进行搜索, 检查是否已经有集合 S 中的插入点存在, 如果没有则认为是合理点, 继续下一步的操作, 否则认为该点的插入不合理而将其忽略掉。 r 在程序的执行过程中并不是一成不变的, 它会随着插入点局部范围内曲率的变化而变化, 在曲率变化趋于剧烈的位置, r 值逐渐减小, 在曲率变化趋于平缓的位置, r 值逐渐增大, 即 r 值与曲率之间是一种成反比的关系, 其中的比例系数是根据具体地形高程数据而给定的一个经验参数。下面给出经过改进后的基于局部熵的地形简化算法流程。

基于局部熵的地形简化算法流程

ENTROPY_INSERT()

begin

用高度场的四个对角顶点进行第一次三角划分,
构造两个初始化三角形;

对所有的输入点计算局部熵作为该点的重要性度量值;

计算所有输入点上的局部曲率;

/* 在本系统用, 我们将由 Sobel 算子计算得到的

梯度值作为该点局部曲率的一种近似 */

输入点按照重要性度量值从小到大排序;

/* 顺序插入前 m 个输入点进行三角划分 */

count = 4;

$i = 1$;

while(count < m)

{

if(第 i 个顶点被标记已经使用)

{

$i++$;

continue;

}

/* 在将第 i 个顶点加入插入点集合 S 之前,

先对其进行合理性检验 */

根据第 i 个顶点上的局部曲率估算距离因子

$r(r \propto \frac{1}{\text{当前点上的局部曲率}})$;

/* 搜索插入点集合 S , 查询其中是否有顶点落在

以第 i 个顶点为中心, 半径为 r 的范围之内 */

valid = TRUE;

for $k = 1$ to count do

{

if(第 $S[k]$ 个顶点在以第 i 个顶点为中心,
半径为 r 的范围之内)

{

valid = FALSE;

break;

}

}

if(valid == FALSE)

{

/* 将经过验证后不合理的插入点忽略掉 */

$i++$;

continue;

}

将第 i 个顶点标记为已使用;

将第 i 个顶点加入到插入点集合 S 中,

并进行 Delaunay 三角划分;

count++;

}

end

值得指出的是, 改进后的算法虽然加入了插入点合理性的判断, 但是并没有增加算法的平均时间复杂度。

4 实验结果与分析

为了测试我们算法的实际性能, 我们分别用基于局部误差和基于局部熵的方法对同一块地形进行简化处理, 以作对比性实验。以下是相关实验分析的结果。

图 1(a) 是由美国国家地质勘测局(USGS)提供的美国大峡谷高程图, 原图大小为 4096×2048 , 我们从上面裁减了一块大小为 150×150 的区域(图 1(b) 所示), 分别用基于局部误差和基于局部熵的方法进行了三角划分, 作对比性实验。图 1(c)、1(e)、1(g) 是用基于局部误差的方法分别选 50、70 和 100 个插入点进行三角划分的结果, 图 1(d)、1(f)、1(h) 是用基于局部熵的方法分别选 50、70 和 100 个插入点进行三角划分的结果。

图中颜色较深的黑色条纹部位为天然大峡谷裂缝, 它是该地形高程图中的主要地形特征。在地形简化中, 主要地形特征应该用较密的三角格网表示, 而其它区域则可以用较稀疏的三角格网进行大数据量地形压缩, 以实现复杂地形的快速

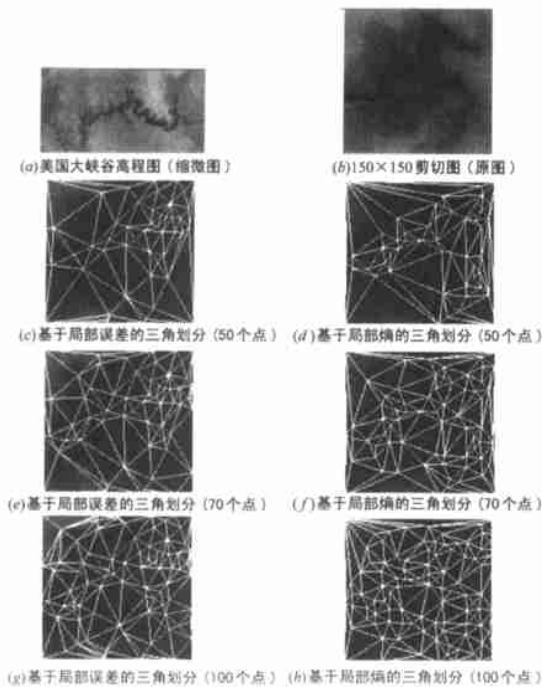


图1 实验结果

显示和浏览^[6]. 通过对比性实验, 不难看出基于局部熵的地形简化算法所生成的三角格网在布局上比基于局部误差的简化算法更加合理, 这一点同样可以从表1中所列出的基于两种重要性度量方法的地形逼近误差对比看出来. 这里的逼近误差是通过式(3)计算得到的.

$$\text{逼近误差} = \frac{\sum_{x=1}^M \sum_{y=1}^M |H(x, y) - TS(x, y)|}{\sum_{x=1}^M \sum_{y=1}^M H(x, y)} \times 100 \quad (3)$$

其中 H 为原始地形平面, TS 为重建后的地形逼近平面.

表1 地形三角格网逼近误差

插入点个数	50	70	100
重要性度量方法			
局部误差	27.41333%	17.95137%	12.46047%
局部熵	11.95816%	9.65427%	6.25565%

设输入点集合 H 中有 n 个点, 我们要取其中 m 个点到插入到集合 S 中进行三角划分. 对于基于局部误差的简化算法, 平均时间复杂度为 $O((m+n)\log m)$. 对于基于局部熵的简化算法, 平均时间复杂度为 $O(m+n)$. 由于后者在时间复杂度上大幅度降低, 使得有可能将其应用在地形的实时简化当中.

5 结束语

本文提出了一种新的基于局部熵的地形简化算法. 该方法不仅计算简单快捷, 适用于各种类型的高度场, 计算过程高度局部化, 而且从算法本身特性出发, 有效的克服了各种传统重要性度量方法所具有的短视性. 该算法可以以最小的插入点, 尽可能准确地逼近真实地形的高程模型, 而且算法执行时间复杂度较局部误差法有很大的降低, 使得可以将其用于地

形的实时简化. 最终的实验结果证实了本算法的优越之处.

参考文献:

- [1] James H Clark. Hierarchical geometric models for visible surface algorithms [J]. Communications of the ACM, 1976, 19(10): 547-554.
- [2] L Paul Chew. Constrained Delaunay Triangulations [J]. Algorithmica, 1989, 11(4): 97-108.
- [3] Michael Garland, Paul S Heckbert. Fast polygonal approximation of terrains and height fields [R]. Pittsburgh: Carnegie Mellon University, 1995.
- [4] Pankaj K Agarwal, Subhash Suri. An efficient algorithm for terrain simplification [A]. In Proceedings of the Eighth Annual ACM-SIAM Symposium on Discrete Algorithms [C]. New Orleans: ACM, 1997. 139-147.
- [5] P Cignoni, C Montani. A comparison of mesh simplification algorithms [J]. Computers & Graphics, 1998, 22(1): 37-54.
- [6] Hugues Hoppe. Progressive meshes [A]. SIGGRAPH'96 Proceedings [C]. New Orleans: ACM, 1996. 99-108.

作者简介:



明德烈 男, 1974 年 2 月生于湖北武汉, 1999 年于武汉理工大学计算机科学系获计算机及应用专业硕士学位, 现为华中科技大学图像识别与人工智能研究所模式识别与智能系统专业博士研究生, 目前研究的主要内容包括图像处理、模式识别、虚拟现实及增强现实等.



尤克非 男, 1977 年 2 月生于河北石家庄, 1999 年于华中科技大学电子与信息工程系获信息工程专业学士学位, 现为华中科技大学电子与信息工程系硕士研究生, 目前研究的主要内容包括三维仿真、三维数据压缩、虚拟现实等.



田金文 男, 1960 年 3 月生于河北石家庄, 华中科技大学教授, 1998 年于华中科技大学图像识别与人工智能研究所获模式识别与智能系统专业博士学位, 目前研究的主要内容包括小波理论, 图像压缩, 计算机视觉等.



柳健 男, 1939 年 4 月生于湖南长沙, 华中科技大学教授、博士生导师, 1961 年毕业于华中科技大学无线电技术专业, 目前研究的主要内容包括遥感图像处理, 计算机视觉等.