

三角形网格的分块 DCT 压缩

刘 波, 张鸿宾

(北京工业大学计算机学院, 北京 100022)

摘 要: 在现有的代表性三角形网格压缩方法中, 先采用一定的网格遍历方法来压缩连接信息, 同时用遍历路径上的相邻顶点来对每个顶点的几何坐标进行预测, 以压缩几何信息. 其主要缺点是只利用了遍历路径上的相邻顶点来进行预测, 并没有充分去掉顶点间的相关性. 其实在空间中一定局部范围内, 所有顶点的坐标间都存在着一定的相关性, 这些顶点虽然在空间上相邻, 但并不一定在遍历路径上相邻. 和图像压缩标准 JPEG 的思路类似, 本文提出一种新的基于分块 DCT 的网格几何信息压缩方法. 先将网格划分成很多基本同样大小的块, 利用每个块内的所有顶点按遍历次序排列成一维序列后, 坐标呈周期性分布的事实, 采用一维 DCT 变换来去除块内顶点间的相关性. 实验表明, 分块 DCT 方法取得了较好的几何信息压缩性能.

关键词: 三角形网格; 压缩编码; 几何信息; 分块; DCT

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2004) 02-0182-05

Block DCT Compression of Triangle Meshes

LIU Bo, ZHANG Hongbin

(Computer Institute, Beijing University of Technology, Beijing 100022, China)

Abstract: In some representative methods of triangle mesh compression, the connectivity is encoded using some kinds of traversal methods, the geometry is encoded by predicting the coordinates of each vertex with neighboring vertices along the traversal path. Their main drawback is the prediction only reduces the correlation between each vertex and its neighboring vertices in the traversal path. The coordinates of all vertices in a local spatial region have strong correlation, although these vertices are neighboring in space, they are usually not neighboring in the traversal path. Inspired by image compression standard JPEG, a novel block DCT geometry compression method is proposed. The input mesh is partitioned into many blocks. Given the fact that if all vertices in a block are ordered in traversal order, their coordinates distribute periodically, we use 2-dimensional DCT transform to reduce the correlation among all vertices in a block. The experimental results show the block DCT method has good geometry information compression performance.

Key words: triangle mesh; compression; geometry information; block partition; DCT

1 引言

三角形网格是计算机图形学中最常用的一种三维物体表示方式. 随着三角形网格的复杂性不断增加, 其数据量越来越大, 必须要进行有效的压缩, 以减少其对存储空间和网络带宽的要求. 与图像、视频压缩的研究相比, 三角形网格数据压缩^[1]的研究还是一个较新的课题. 一个三角形网格包含的信息可分为两部分: 连接信息和属性信息. 连接信息指定顶点间的连接关系; 属性信息指定每个顶点的各种属性, 主要为顶点的几何坐标, 称为几何信息, 以及颜色、法向、纹理坐标等用于绘制的信息.

当前三角形网格压缩的研究主要分为两大方向: 渐进 (progressive) 与非渐进 (non progressive) 压缩. 非渐进压缩, 或

称单位率 (single rate) 压缩, 是指将原始输入网格压缩为具有一个固定的目标位率的位流; 解码时, 解码器需要拥有整个压缩位流才能进行解码. 而渐进压缩是指将原始输入网格压缩为具有多个目标位率的位流; 解码时, 从某个目标位率重构出的网格为原始网格的一个近似. 本文探讨的是非渐进压缩方法中的几何信息压缩.

一个网格压缩方法需要探讨如何对连接信息与几何信息进行压缩. 目前各种非渐进压缩方法的基本思想都是先采用某种遍历方法对网格进行遍历, 并将遍历的结果表示成某种符号序列, 再对这个符号序列进行熵编码, 以压缩连接信息; 同时对每个顶点用遍历路径上的相邻顶点来进行预测, 以去掉相邻顶点的坐标间的相关性, 然后再对预测误差进行熵编码, 以压缩几何信息.

Edgebreaker^[2,3] 和 TG^[4] 是目前非渐进压缩方法中压缩性能较好的代表性工作。以 Edgebreaker 方法为例, 它采用了三角形的宽度优先遍历, 每当从某条边进入到一个未被访问过的三角形时, 如果相对的那个顶点没有被访问过, 则输出符号 C (center), 标记该顶点与三角形为已访问, 并且下一个操作为访问右边的相邻三角形; 否则根据左、右两个相邻三角形被访问情况的四种可能组合, 分别输出 L (left), R (right), S (split), E (end) 等四个符号中的某一个, 标记该三角形为已访问, 并采取相应的下一个操作: 访问右边的相邻三角形, 访问左边的相邻三角形, 分叉或结束。这个过程不断进行, 直到所有的三角形都被访问过为止, 最后再对所得的符号序列进行熵编码。由于这五种符号的概率分布很不均匀, 其中 C 的概率约为 0.15, 所以 Edgebreaker 方法可获得很高的连接信息压缩比。其压缩位率的上限为 31.67b/v, 是第一个在理论上给出了压缩上限的方法。

TG 和 Edgebreaker 方法对几何信息的压缩方案是沿遍历路径对每个顶点的坐标进行平行四边形预测 (parallelogram prediction), 然后再对预测误差进行熵编码。如图 1 所示, 顶点 a, b, c 已经被遍历和编码, 在编码顶点 v 的几何坐标时, 用 p 点对其进行预测, 其中 p, a, b, c 构成平行四边形; 预测误差为 $E = r_v - (r_a + r_b - r_c)$ 。

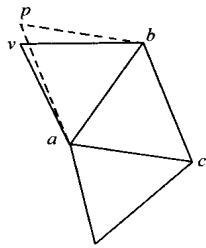


图 1 平行四边形预测

沿遍历路径对顶点进行预测的几何信息压缩方案的主要缺点是没有充分去掉顶点间的相关性, 压缩性能不太好。事实上, 在空间一定局部范围内的所有顶点间都存在着一定的相关性, 这些顶点虽然在空间上相邻, 但并不一定在遍历序列上相邻, 仅仅沿遍历序列进行预测是无法充分利用它们间的相关性的。

对图像而言, 一定区域范围内的所有像素的灰度 (颜色) 值之间存在着较强的相关性, 图像压缩标准 JPEG^[5] 采用分块 DCT 的方法来利用这种相关性。它先将图像分成很多的块, 然后利用二维 DCT 变换来去掉每个块内的所有像素灰度 (颜色) 值之间的相关性。由于图像的能量主要集中在低频部分, 大量高频 DCT 系数的值集中在零附近, 对 DCT 系数做量化和熵编码后, 可取得较好的压缩性能。

和图像压缩标准 JPEG 的思路类似, 我们试图将网格划分成很多块, 然后再设法去掉各块内所有顶点间的相关性。问题在于: (1) 网格存在不规则的连接关系, 需要进行连接关系的编码。将网格划分成块也比将图像进行分块要复杂的多。(2) 虽然网格表面是二维的, 但由于每个块内的连接关系不规则, 无法使用二维 DCT 变换来去掉块内顶点间的相关性。

本文提出一种新颖的基于分块 DCT 的网格几何信息压缩方法, 解决了上述问题。编码时, 为了利用空间一定范围内所有顶点间的相关性, 我们采用 Metis 工具^[7] 将网格划分成很多基本同样大小的块。由于 TG 和 Edgebreaker 方法所采用的遍历方法是螺旋式的宽度优先遍历, 每个块内的顶点按被遍历到的次序排列成一维序列后, 坐标值是呈周期性分布的, 对

这种周期性的坐标序列做一维 DCT 变换, 来去除块内顶点间的相关性。然后对所有块的 DCT 系数做量化和熵编码。解码时, 先解码连接信息, 这样在解码几何信息时, 整个网格的连接信息已知; 对这个解码后的连接关系进行网格划分, 经过熵解码、反量化以及 DCT 逆变换得到各块内顶点的几何坐标, 并将各块内的连接关系与该块内各顶点的几何坐标结合起来, 从而得到解码后的网格。

2 编码

在编码时, 先编码完连接信息后, 再编码几何信息。依次包括连接信息编码、网格划分、DCT 变换、DCT 系数的量化和熵编码等阶段。

2.1 连接信息压缩

本文的核心是探讨几何信息压缩方法。除了在对每一块内的顶点进行一维排序时受遍历方法影响外, 本文的分块 DCT 方法是独立于所采用的连接信息压缩方法的。可以采用 Edgebreaker 或 TG 方法来压缩连接信息, 具体可参看文献[2~4]。

2.2 网格划分

与文献[6]一样, 我们采用了 Metis 软件包^[7] 来进行网格划分。Metis 采用了多层划分 (multilevel partitioning) 方法, 具有快速、划分均匀等优点。只需输入网格的连接关系和需要划分的块数, Metis 就可以划分好网格, 而无需网格的几何信息。

如果将网格划分为较少的块数, 则每个块中的顶点数目比较多, 能够去掉更多的相关性, 但后续的 DCT 变换将更耗时; 且随着块的增大, 靠近块边界的顶点与中心区域顶点的相关性越来越小, 去相关效果逐渐趋于饱和, 计算复杂度却迅速增加。如果将网格划分为较多的块数, 则每个块中的顶点数目比较少, 计算速度快, 但去掉的相关性会少一些, 且块效应 (block artifact) 严重一些。在实验中, 我们将所有网格都划分成 50 块。但也可以根据每个网格自身的大小来决定划分的块数。图 2 显示了对 horse 和 bunny 网格用 Metis 进行划分以后的效果。

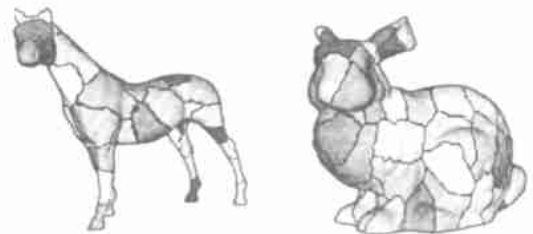


图 2 网格划分

用 Metis 将网格划分后, 就可以确定每一块中包含了哪些顶点, 然后本文再用 DCT 变换来去掉每一块内所有顶点间的相关性。

2.3 用 DCT 变换去相关

由于每个块内的连接关系不规则, 无法使用二维 DCT 变换来去掉块内顶点间的相关性。我们注意到, 由于 TG 和 Edgebreaker 方法所采用的遍历方法是螺旋式的宽度优先遍历, 在整个遍历过程中, 遍历路径将会多次进入和离开每一块。因

此, 如果将每个块内的顶点按被遍历到的先后次序排列成一维序列, 它们的坐标分布将有一定的周期性. 图 3 显示了在将 cow 网格(见图 5)划分成 50 块时, 将位于左侧躯干的第 13 块内的所有顶点如上所述排列成一维序列后, 序列上顶点的 x ,

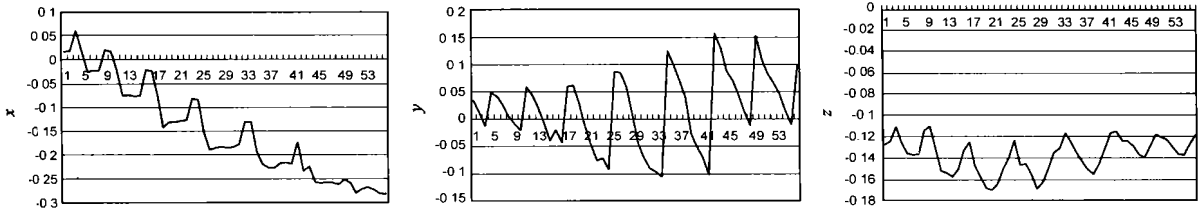


图 3 块内顶点坐标的周期性分布

由于剧烈变化的细节并不多, 用 DCT 变换作频谱分解后, 除了少数低频系数外, 大量的高频系数的值应该都很小, 从而有利于后续的熵编码. 图 4 显示了对 cow 网格各块的顶点序列的 x 坐标做 DCT 变换后, DCT 系数的分布. 在图中, 按从低频到高频的顺序, 将所有块的 DCT 系数连在一起. 从图 4 可以明显看出, 确实大量的高频系数都分布在零附近.

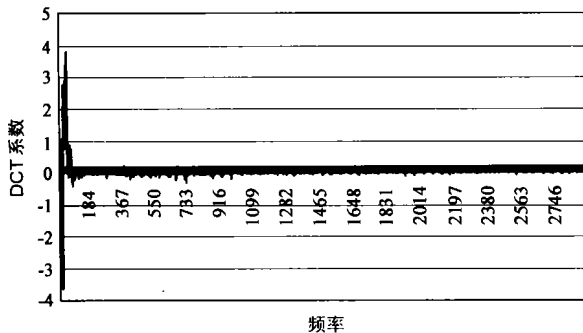


图 4 DCT 系数的大小

2.1.4 量化和算术编码

做 DCT 变换后, 得到的 DCT 系数为浮点数, 需进行量化后, 才能进行熵编码. 具体采用的量化方案为死区量化 (dead zone quantization), 量化位数通常取 8 至 16 位.

对 DCT 系数进行量化后, 大量的高频系数都在零附近分布, 概率分布很不均匀. 对量化后的 DCT 系数做熵编码可以有效利用这种概率分布的不均匀性, 进一步减少数据量. 常用的熵编码方法有 Huffman 编码和算术编码, 由于算术编码^[8]比 Huffman 编码的压缩性能略好, 且无需事先知道输入符号的概率分布, 因此本文采用了算术编码. 具体实现时, 考虑到 DCT 系数在零左右呈一定的对称分布, 本文将量化后 DCT 系数的符号和大小 (magnitude) 分开, 分别进行算术编码, 以取得更好的压缩性能.

3 解码

解码基本上是以上描述的编码过程的逆过程. 先将整个网格的连接信息解码出来, 再了解码几何信息. 连接关系解码出来后, 对这个解码后的连接关系用 Metis 进行网格划分. 划分确定了每一块包含了哪些顶点以及块内顶点的排列顺序, 其结果应该与编码端划分的结果一致. 在解码几何信息时, 先

y, z 坐标值的分布, 其中横轴为顶点号. 从图中可以看到明显的周期性, 且剧烈变化的细节并不多. 对这种周期性的坐标分布, 本文采用一维 DCT 变换来进行频谱分解, 以去除块内顶点坐标间的相关性.

进行算术解码和反量化, 得到 DCT 系数序列; 由于此时网格的划分结果已知, 我们将属于各块的 DCT 系数从序列中取出来, 再进行 DCT 逆变换, 得到每一块中各个顶点的几何坐标. 最后将解码后的连接关系与几何坐标结合起来, 就得到了完整的解码后的网格.

DCT 系数的量化引入了一定的失真, 实验发现, 在低位率压缩时, 失真较为明显. 为提高解码网格的质量, 本文在低位率解码时, 对解码网格进行平滑后处理. 具体采用 Taubin 的 K-L 平滑算法^[9]. 而高位率解码时则不需要做平滑后处理.

4 实验结果

在实验中, 我们实现了基于分块 DCT 的网格几何信息压缩, 采用的连接关系压缩方法是 Edgebreaker. 为了对比, 我们也实现了基于 Edgebreaker 的平行四边形预测几何压缩. 图 5 显示了实验中所用的一些网格数据. 表 1 列出了对各个网格分别用平行四边形预测和分块 DCT 方法压缩后的几何信息位率, 单位为 bits/vertex. 其中两种方法都采用了 11 位量化; 在分块 DCT 方法中, 所有网格都被划分为 50 块; 压缩倍数是相对于 ASCII 格式的网格文件大小而言的.

从表 1 可见, 在相同的量化位数时, 分块 DCT 方法的几何信息位率比平行四边形预测方法明显要小. 同时可以看出, 随着网格分辨率的增加, Edgebreaker 和分块 DCT 方法的几何信息位率越来越低, 且分块 DCT 方法的几何信息位率比 Edgebreaker 方法下降得更快. 这是因为网格分辨率越高, 顶点间的距离就越近, 从而相关性越强, 能去掉的相关性也就越多.

减少量化位数时, 可以获得更高的压缩比, 但失真也会相应随着增加. 实验发现, 在低位数量化时, 分块 DCT 方法的失真较为明显. 图 6(a) 显示了在 9 位量化时, 解码出的 cow 网格. 与原始网格相比, 人眼能略微看到一些失真, 主要是在细节较多的区域, 如牛角、牛嘴等. 本文将在第 5 节中讨论失真的原因. 如前所述, 我们对低位率解码后的网格进行平滑后处理来提高解码质量. 图 6(b) 显示了用 K-L 算法对图 6(a) 的解码网格进行平滑后处理的效果, 它与原始网格已经没有明显差别.

图 7 显示了 Edgebreaker 方法和分块 DCT 方法的率失真曲线, 其中失真用 Metro^[10] 工具测量的, 且在图中被扩大了一万倍. 尽管在相同的量化位数时, 分块 DCT 方法的几何信

息位率要低于 Edgebreaker 方法,但从图 7 可以看出,分块 DCT 方法在低位率时的率²失真性能比 Edgebreaker 方法要差一些.对 Venus 网格,在同样的位率时,分块 DCT 方法的失真比 Edgebreaker 约大 215×10^{-4} 左右,这是一个很小的值,人眼看不出明显差别来.从图 7 还可以看出,网格的分辨率越高,两种方法的失真的差别就越小.在较高位率时,两种方法的率²失真曲线趋于一致,失真变得很小.

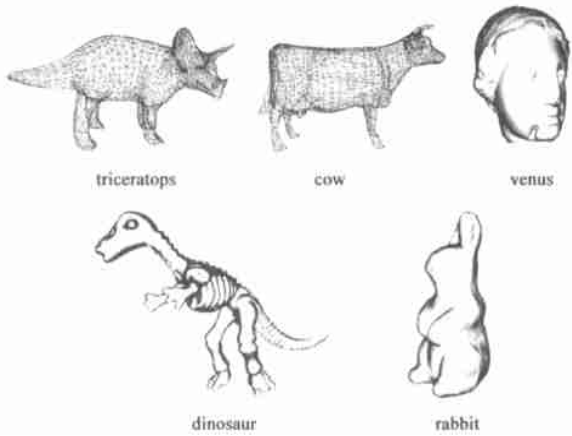


图 5 实验中所用的一些网格模型

表 1 平行四边形预测和分块 DCT 方法的比较(11 位量化)

网格	顶点数	平行四边形预测 的几何信息位率	分块 DCT 方法的 几何信息位率	分块 DCT 方法 的总压缩倍数
cow	2904	20.4	13.6	31
triceratops	2832	19.5	12.9	31
venus	33587	15.2	6.7	57
rabbit	16760	17.3	7.7	51
dinosaur	14050	17.2	6.6	55

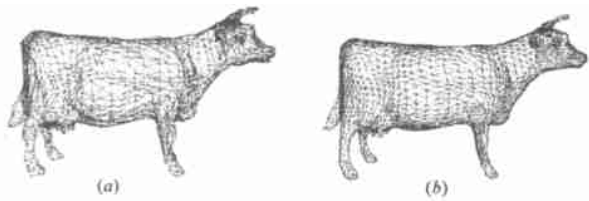


图 6 解码和平滑后的 cow 网格

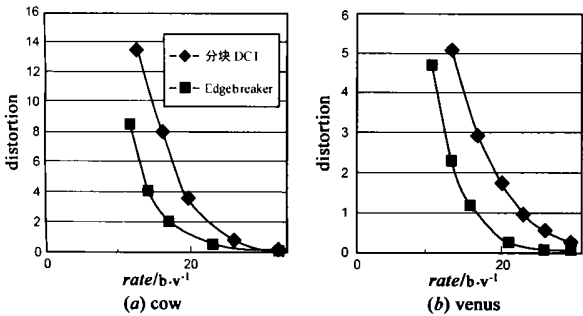


图 7 Edgebreaker 与分块 DCT 方法的率²失真曲线的比较

5 讨论

先分析一下在分块 DCT 方法中引起失真的原因.对 DCT 系数做量化和反量化后,得到的值与真实值有一定的误差,做 DCT 逆变换后,顶点的坐标将沿三个坐标轴方向相对于真实位置有一定的移动.而人眼所感知到的三维物体表面细节其实是顶点的坐标对局部邻域的法向偏离,如果顶点坐标沿邻域的切向有少量移动,人眼并不能明显感知出物体表面细节有何变化.但在分块 DCT 方法中,顶点坐标的误差是沿坐标轴方向,而不是沿邻域的切向,所以失真相对要明显一些.

Karni 和 Gotsman^[6]提出了基于谱方法的网格几何信息压缩.他们把由连接关系唯一确定的 Laplace 矩阵的本征矢量作为一组正交基,对网格顶点的坐标做谱展开,然后对谱系数进行量化和熵编码.从理论上说,谱分解的本质更类似于图像中的二维 DCT 变换,其率²失真性能比本文的分块 DCT 方法也要好一些.但谱方法的问题在于求 Laplace 矩阵的本征矢量时,计算复杂度太高,难以做到实时解码.为了减少解码复杂度,文献[6]先将网格分割成一些小的子网格,然后对每一子网格单独做谱分解.即使这样,对复杂网格,谱方法还是很难做到实时.相比之下,分块 DCT 方法可以采用快速 DCT 逆变换而做到实时解码.

最近,Gu 等人提出了几何图像(geometry image)^[11]方法.他们先把一个网格切开,然后将其参数化到一个矩形区域上,从而得到所谓的几何图像.在将网格转换成几何图像后,可以利用图像压缩方法来进行压缩.几何图像方法与分块 DCT 方法的主要区别在于:几何图像方法尽管能用于网格压缩,但它主要是一种重新网格化方法,目的是把非正规(irregular)的连接关系转换为完全正规的,所以在解码后连接关系是有损的,这对某些应用并不合适;而分块 DCT 方法则是一种连接关系无损压缩方法.

6 结论

本文提出一种新颖的基于分块 DCT 的网格几何信息压缩方法.为了利用空间一定局部范围内所有顶点间的相关性,我们将网格划分成很多基本同样大小的块.由于连接关系压缩所采用的遍历方法一般是螺旋式的宽度优先遍历,每个块内的顶点按遍历次序排列成一维序列后,坐标分布呈一定的周期性,我们对这种周期性的坐标序列做一维 DCT 变换,来去除块内顶点间的相关性.实验表明,分块 DCT 方法取得了较好的几何信息压缩性能.

参考文献:

[1] Taubin G, Rossignac J Eds. 3D geometry compression [A]. SIG2 GRAPH. 99 Course Notes 21 [C]. New York, USA: ACM Press, ACM 1999.
[2] Rossignac J. Edgebreaker: Connectivity compression for triangle meshes [J]. IEEE Transactions on Visualization and Computer Graphics, 1999,

- 5(1): 47261.
- [3] Rossignac J, Safonova A, Szymczak A. 3D compression made simple: Edgebreaker on a come2table[A]. International Conference on Shape Modeling and Applications[C]. Genoa, Italy: IEEE, 2001. 278- 283.
- [4] Tuma C, Gotsman C. Triangle mesh compression[A]. Graphics Inter2 face. 98[DB/OL]. <http://www.graphicsinterface.org/proceedings/1998>, 1998. 26- 34.
- [5] Pennebaker W B, Mitchell J L. JPEG Still Image Data Compression Standard[S]. New York: Van Nostrand, 1993.
- [6] Karni Z, Gotsman C. Spectral coding of mesh geometry[A]. Computer Graphics(SIGGRAPH. 00 Proceedings) [C]. New York, USA: ACM, Press 2000. 279- 286.
- [7] <http://www.users.cs.umn.edu/~karypis/memis/memis.html> [DB/ DL].
- [8] Witten I H, Neal R M, Cleary J G. Arithmetic coding for data compression[J]. Communications of the ACM, 1987, 30(6): 520- 540.
- [9] Taubin G. A signal processing approach to fair surface design[A]. Computer Graphics (SIGGRAPH. 95 Proceedings) [C]. New York, USA: ACM Press, 1995. 351- 358.
- [10] Cignoni P, et al. Metro: measuring error on simplified surfaces[J]. Computer Graphics Forum, 1998, 17(2): 167- 174.
- [11] Gu Xianfeng, Gortler S, Hoppe H. Geometry images[A]. Computer Graphics (SIGGRAPH. 02 Proceedings) [C]. New York, USA: ACM Press, 2002. 355- 361.

作者简介:



刘 波 男, 1975 年 10 月生于湖南双牌, 北京工业大学计算机学院博士研究生, 主要研究方向为计算机图形学, 多媒体数据压缩。



张鸿宾 男, 1944 年 7 月生于北京, 教授, 博士生导师, 1968 年清华大学自动控制系毕业, 1981 年清华大学模式识别与智能控制专业研究生毕业, 1986~ 1989 年日本京都大学、1993~ 1994 年美国 RPI 高级访问学者。目前主要从事模式识别、图像与机器视觉、人工神经网络以及数字水印等方面的研究工作。