



点云编码综述

A Review of Point Cloud Compression

李厚强 / LI Houqiang¹, 李礼 / LI Li¹, 李竹 / LI Zhu²

(1. 中国科学技术大学, 中国 合肥 230027;

2. 美国密苏里大学堪萨斯分校, 美国 堪萨斯城 64110)

(1. University of Science and Technology of China, Hefei 230027, China;

2. University of Missouri-Kansas City, Kansas City 64110, USA)

摘要: 点云编码是支撑点云广泛应用的关键技术之一, 是近期技术研究和标准化领域的热点。对点云几何信息和属性信息编码技术演进进行了回顾, 并针对稠密点云和稀疏点云的几种典型编码方法的编码效率进行了比较。未来点云编码研究将集中于利用帧间预测去除动态点云的不同帧之间的相关性, 以及端到端点云编码、任务驱动的点云编码等方面。

关键词: 3D 点云编码; 几何信息编码; 属性信息编码

Abstract: 3D point cloud compression is one of the key technologies supporting the wide-spread use of point clouds. Recently, it is one of the focuses for both research and standardization groups. The latest advance of the compression technologies for both the 3D point cloud geometry and attribute information is reviewed. Compression efficiencies of several typical compression technologies for both the 3D dense and sparse point clouds are compared. In the future, more studies will focus on inter-frame prediction to exploit the correlations between different frames in 3D dynamic point clouds, end-to-end point cloud compression, and task-driven point cloud compression.

Keywords: 3D point cloud compression; geometry information coding; attribute information coding

DOI: 10.12142/ZTETJ.202101003

网络出版地址: <https://kns.cnki.net/kcms/detail/34.1228.TN.20210122.1752.002.html>

网络出版日期: 2021-01-25

收稿日期: 2020-12-08

点云是一系列高维空间点(例如三维空间点)的集合。每一个点包含几何信息(x, y, z)以及颜色和反射率等属性信息。根据点云中点的密度, 点云可以粗略地分为稠密点云和稀疏点云。稠密点云可以用来精细重建3D物体例如人物等, 可被广泛应用于虚拟现实和增强现实。稠密点云重建的

3D物体支持6自由度, 相比360°全景视频仅能支持3自由度, 可以给用户带来更好的视觉体验。稀疏点云可以高精度重建3D场景, 结合2D摄像头采集的高清图像视频, 可被用于自动驾驶和机器人视觉等应用中。由于点云数据量巨大, 点云编码成为了上述应用中不可或缺的一环。相比成熟的图像视频编码技术, 点云编码由于其独有的特点成为近期的研究热点。图像视频中的像素在2D空间中均匀分布, 而点云中的点在3D空间的分布是

稀疏且无规律的。点云的稀疏性是指3D空间仅有很小一部分3D位置被点占用。从压缩的角度来看, 相比于编码整个3D空间, 仅仅编码被占用的部分信息会更加高效。同时, 点云的无规律性使得点云的不同点之间的相关性难以被有效去除。点云编码可以根据其包含的信息分为两个部分: 几何信息编码指明空间中哪些位置存在3D点, 属性信息编码指明空间中3D点的颜色和反射率等属性信息。在大部分点云编码算法中, 都是先编码几何信

基金项目: 国家自然科学基金(62021001); 中国科学技术大学引进人才科研启动经费(KY2100000108)

息,然后基于重建的几何信息和原始点云对点云进行重着色,最后编码重着色之后的属性信息。

1 几何信息编码

几何信息编码主要分为3类:基于树结构的方法、基于表面近似的方法、基于映射的方法。下面我们将分别对这些方法进行详细介绍。

1.1 基于树结构的方法

基于树结构的方法是最直接的几何信息编码方法。其基本思想是对包含点云的最小立方体以树的形式进行迭代划分,如果划分完的子立方体包含点,则编码“1”,且会被进一步划分;不包含点,则编码“0”,且不会被进一步划分。在基于树结构的方法中,使用的树结构通常为八叉树和二叉树。

早在21世纪初,基于八叉树的方法就被用于编码点云^[1]。基于八叉树的方法首先迭代地把包含点云的最小立方体划分为8个子立方体,然后用一个字节编码8个子立方体是否包含点这一信息。由于父节点和子节点,以及相邻节点的字节之间存在很强的相关性,通常使用基于上下文的算术编码进一步去除该相关性。由于该方法简单有效,它在国际动态图像专家组征集的所有稀疏点云编码方法中取得了优胜,最终发展成为基于几何信息的点云编码标准之一^[2]。为了进一步提升编码效率,我们提出了使用该字节中1的个数和组合来代表该字节,1的个数和组合也可以使用父节点和邻近节点近似成的面来估计^[1]。八叉树的主要缺点是表征八叉树需要的比特数会随着树深度的增加而急剧增加,所以使用二叉树来编码几何信息的方法被提出^[3]。点云编码使用基于数据的二叉树可以一定程度上缓解因深度增加所需要的比特数,但是基于数据

而非空间的二叉树需要编码分割节点信息,尤其在树的深度较浅时会消耗大量比特。

1.2 基于表面近似的方法

因为完整点云很难被近似成一个参数化的表面,所以基于表面近似的方法通常与基于树的方法结合使用。首先使用八叉树或二叉树把点云分割成互不包含的小立方体,然后小立方体被近似成表面以进一步编码。表面近似的方法的本质是降维,编码一个小立方体相当于编码三维信息,而把小立方体近似成一个表面则仅需要编码二维信息。

在所有基于表面近似的方法中,最常用的表面是平面。我们首先对点云进行八叉树划分,划分到一定的深度后,再使用平面对立方体中的点进行近似,编码平面与立方体的交点来代表平面,最后对平面进行采样恢复最终的点。该方法在国际运动图像专家组征集的所有静态稠密点云编码方法中取得了优胜,最终发展成为基于几何信息的点云编码标准之一^[2]。除了使用八叉树作为树分割的方式,二叉树也可以作为一种树分割的方式使用。除了使用采样来恢复最终的点,也可以使用四叉树对近似形成的平面进行基于树结构的编码^[3]。基于平面的编码方法相比于基于树的编码方法,在低码率上可以带来明显的性能提升,但是由于平面近似始终存在误差,基于平面的编码方法无法实现无损编码。除此之外,为了进一步提升表面近似精度,二阶曲面也被用于表面近似^[4],但是二阶曲面相比平面需要传输更多的头信息,这会导致编码性能提升有限。

1.3 基于映射的方法

基于映射的方法最初是针对网格

(mesh)编码设计的。近些年来,基于映射的方法逐渐开始被用于点云编码。基于映射的方法的基本思想是把点云从3D空间映射到2D空间,然后使用成熟的2D图像视频编码方法进行编码。此方法的核心在于找到一种合适的映射,既能在投影的过程中减少点的损失,又能使投影之后的图像视频具有较高的时空相关性以更好地利用2D图像视频编码方法中高效的预测技术。

为了尽可能在投影过程中减少点的丢失,我们以一定顺序扫描点云八叉树,把3D点云转化为2D图像或视频^[5]。这种投影方式不会造成任何点的丢失,但形成的2D图像视频时空相关性弱,编码效率低。为了提高2D图像视频的时空相关性,我们提出把点云完整地投影到包围着该点云的圆柱体或立方体上^[6]。此方法的2D图像视频编码效率高,但会造成部分被遮挡的连续点丢失,从而导致3D点云质量较差。为了兼顾投影点的数量和2D图像视频编码效率,我们提出把具有相似法向量的点按片投影到包围该点云的立方体上,不同的点云会形成几十到数百个片^[2]。此基于片的投影不会导致被遮挡的连续大量点丢失,因为它们会形成一个新的片投影到2D空间。此外,基于片的投影方法把具有相似法向量的点投影成一个片,使得属于同一个片的点的深度方差较小,有利于提升编码效率。该方法在国际运动图像专家组征集的所有动态点云编码方法中取得了优胜,最终发展成为基于视频的点云编码标准^[2]。

2 属性信息编码

属性信息编码主要可以分为3类:基于变换的方法、基于预测的方法、基于映射的方法。下面我们将分别对这些方法进行详细介绍。

2.1 基于变换的方法

变换是编码中一种常用的去相关方法。基于变换的属性信息编码方法的基本思想是利用重建的几何信息来设计一个内容自适应的属性信息变换，以去除属性信息之间的相关性。去除相关性之后的属性信息经过量化和熵编码后形成属性信息码流。

为了充分利用已经编码的几何信息，我们提出使用图变换的方法对属性信息进行变换编码^[7]。首先根据点与点之间的距离构建图，然后对图拉普拉斯矩阵进行特征值分解，最后使用特征向量构建的变换对属性信息进行变换。除此之外，我们还提出使用高斯过程来近似点与点之间的关系，推导出高斯过程对应的K-L变换来编码属性信息^[8]。上述方法能达到较好的编码性能，但是需要进行复杂的特征值分解。这会导致很高的复杂度，不利于实际使用。为了更好地取得编码性能和复杂度之间平衡，我们提出使用基于区域的自适应分层变换对属性信息进行编码^[9]。基于区域的自适应分层变换本质上是加权Haar小波变换。根据八叉树的每一个子节点包含的点的数量，对属性信息进行加权小波变换，以利用几何信息。基于区域的自适应分层变换被基于几何信息的点云编码标准采纳，成为被推荐的静态点云属性编码方法^[2]。除了以上常规的基于变换的编码方法，基于几何信息的稀疏表达变换也被用于压缩属性信息^[10]，但是稀疏位置信息的编码限制了其效率。

2.2 基于预测的方法

除了变换以外，预测也是一种常用于编码的去相关方法。不同于变换对信号进行旋转使其更适合编码，预测本质上是已编码的信息作为条件，使用条件熵代替原信号的熵，从

而提升编码效率。和变换一样，预测之后的信号经过量化和熵编码后形成码流。

在图像视频编码中，基于邻近已重建图像块对当前图像块进行预测，在各代图像视频编码标准中一直沿用。在基于八叉树的几何信息编码中，点云被分割成多个等大的小立方体，基于邻近已经重建的小立方体的属性信息对当前小立方体进行预测，是2D预测编码到3D预测编码的一个简单扩展^[11]。但是一方面，点云的稀疏性导致邻近可用预测块较少，3D预测不如2D预测有效；另一方面，如果想要使3D预测和2D预测一样精细，在3D空间进行预测编码需要使用比2D空间多得多的预测方向。因此，针对3D点云进行类似图像视频的预测并不高效。3D点云属性信息预测通常使用分层预测^[2]。我们把点云属性信息分成不同的层进行逐层编码，并使用已经编码的层对待编码的层进行加权预测。在此种方法的发展过程中，涌现出了多种点云分层方式：基于点与点之间距离的分层方法，以及基于二叉树的分层方法等。

除此之外，我们还提出了基于提升的方式使用编码残差来进一步修正层间预测，以更好地提升性能^[2]。由于分层预测的方法在编码性能和复杂度之间取得了很好的均衡，该方法被基于几何信息的点云编码标准采纳，成为了推荐的属性压缩方法之一。

2.3 基于映射的方法

大部分属性信息编码方法利用几何信息去除属性信息之间的相关性，以提升编码效率；但是基于映射的属性编码方法则有所不同，它采用与基于映射的几何编码方法相同的投影方式，然后使用成熟的视频编码技术对重着色之后的属性视频进行编码。从基本的流程上来说，基于映射的属性编码方法和基于映射的几何编码方法并没有太大不同。在不使用几何信息的情况下，基于高效成熟的2D图像视频编码技术已经能够取得非常好的性能。基于已经编码的几何信息，我们对属性2D图像视频编码器进行运动矢量预测率，并对失真方面进行优化^[12]，这使得基于映射的属性视频编码方法取得了进一步的性能提升。基于映射的属性信息压缩方法和基于映射的几何信息压缩方法组成了基于视频的点云编码标准^[2]。

3 点云传统编码方法的比较

3.1 几何信息压缩性能

基于映射的方法和基于表面近似的方法都不适合稀疏点云，所以稀疏点云几何信息几乎只能使用基于树的方法进行压缩。不同于稀疏点云，稠密点云几何信息可以使用上述3种方法进行压缩。表1给出了相对于表面近似的方法，基于映射的方法、基于树结构的方法压缩稠密点云几何信息

▼表1 使用基于映射的方法和基于树结构的方法压缩稠密点云几何信息所得的率失真性能结果

测试序列	基于映射的方法		基于树的方法	
	D1/%	D2/%	D1/%	D2/%
Loot	-80.8	-72.0	117.4	67.9
Redandblack	-77.1	-60.6	103.0	117.9
Soldier	-74.7	-62.8	91.5	61.4
Queen	-87.3	-76.8	94.3	74.9
Longdress	-83.0	-75.0	51.2	54.5

D1: 点到点距离 D2: 点到平面的距离

的率失真性能结果。表1中, $D1$ 表示点到点的距离, $D2$ 表示点到平面的距离; 数字表示相同点云几何信息质量下的码率变化。从表1可以看出, 针对稠密点云, 基于映射的方法会比基于表面近似的方法带来显著的性能提升, 在相同的点到点和点到平面的距离下, 基于映射的方法分别会带来近80%和70%的码率节省。此外, 基于树的方法相比于基于表面近似的方法, 需要额外90%和70%的比特数。综上所述, 基于映射的方法可以带来最好的稠密点云几何信息压缩效果。

3.2 属性信息压缩性能

针对稠密点云属性信息, 基于变换、预测和映射的方法均可以使用。但是基于映射的属性信息压缩方法通常和基于映射的几何信息压缩方法结合起来使用, 而基于变换和预测的方法通常和基于树和表面的几何信息压缩方法结合起来使用。不同的几何信息压缩方法会带来不同的重着色之后的点云, 所以很难单独对基于映射的方法和基于预测和变换的方法进行直接对比。表2给出了使用基于变换的方法和基于预测的方法压缩稠密点云属性信息的率失真性能比较。从表2可以看出, 相比于基于预测的方法, 基于变换的方法对于亮度分量会带来大约3.6%的码率增加, 对于色度分量的性能损失则更加明显。因此, 针对稠密点云, 基于预测的方法可以带来比基于变换的方法更好的属性信息压缩性能。

针对稀疏点云属性信息, 基于映射的方法难以使用, 所以我们主要对基于变换的方法和基于预测的方法进行了对比, 率失真性能如表3所示。从表3可以看出, 针对稀疏点云属性信息, 相比基于预测的方法, 基于变换的方法能带来大约3%的码率节省。

综上所述, 基于变换的方法是目前业界效果比较好的稀疏点云属性信息压缩方法。

4 点云编码最新进展和发展方向

近几年来, 点云几何和属性信息编码技术的发展取得了长足进步, 但是和传统的图像视频编码标准所能达到的编码效率相比, 仍有较大的距离。如何进一步提升编码性能是点云编码未来发展的目标之一。

帧间预测是传统视频编码中提升压缩效率最显著的部分, 但是对于动态点云而言, 帧间预测效率目前还远远不够。对于稠密动态点云帧间预测, 基于片的映射方法取得了目前最优的性能, 但基于片的映射方法仍存在两个问题: 首先, 点云按片映射到2D视频的过程复杂度很高, 不同于视频编码存在成熟的市场优化方案, 此映射过程目前还不适合实时应用; 另外, 按片映射过程不可避免地破坏了视频的时空相关性。尽管一些人们尝试在视频编码过程中通过寻找3D空间对应块来解决此问题^[12], 但如何从更源头产生时空更连续的视频仍然是稠密动

态点云编码中一个非常关键的问题。对于稀疏动态点云帧间预测, 需要直接在3D空间进行运动估计和运动补偿。但由于相邻点云帧点数不完全相同, 且不同点之间不存在和视频像素一样的——对应关系, 所以3D运动估计和运动补偿是业界一个非常困难的问题, 目前还没有一个成熟的解决方案。

基于深度学习的端到端图像视频编码近期取得了长足的进步, 几乎达到或超越了传统图像视频编码的性能, 这就促进了以端到端的方式对点云进行压缩编码的方法的使用。端到端点云几何属性信息压缩是目前的研究热点^[13]。几何信息编码使用3D普通或稀疏卷积神经网络来编码每个空间位置是否存在3D点这一信息; 属性信息编码使用神经网络结合坐标信息编码对应的颜色和反射率等。目前端到端点云编码仅在稠密静态点云方面取得了较好的效果, 而针对稀疏点云和动态点云, 目前都没有较好的解决方案。另外, 稀疏点云主要针对机器视觉, 易于被端端点云压缩利用, 也是未来非常值得尝试的方向。

▼表2 使用基于变换的方法压缩稠密点云属性信息所得的率失真性能结果

测试序列	Y/%	U/%	V/%
Loot	-4.2	21.3	19.2
Redandblack	7.7	12.2	42.3
Soldier	2.7	21.5	21.3
Queen	5.0	22.1	2.9
Longdress	6.6	22.4	53.4

U、V: 统称为色度分量 Y: 亮度分量

▼表3 使用基于变换的方法压缩稀疏点云属性信息所得的率失真性能结果

测试序列	反射率/%
Ford_01	-2.5
Ford_02	-1.5
Ford_03	-1.5
Qnxadas-junction-approach	-5.4
Qnxadas-junction-exit	-4.6
Qnxadas-motorway-join	-5.9
Qnxadas-navigating-bends	-1.3

5 结束语

点云几何和属性信息编码是支撑点云广泛应用的关键技术之一。点云几何和属性信息编码近年来取得了长足的进步,但在帧间预测、编码应用等方面仍有许多悬而未决的问题。未来人们需要进一步研究帧间预测、基于深度学习的端到端点云编码等技术,以更高层的应用为目标设计更高效的点云几何和属性信息压缩技术。

参考文献

- [1] SCHNABEL R, KLEIN R. Octree-based point-cloud compression [C]//IEEE VGTC conference on Point-Based Graphics. Goslar, Germany: IEEE VGTC, 2006, (6): 111-120
- [2] SCHWARZ S, PREDA M, BARONCINI V, et al. Emerging MPEG standards for point cloud compression [J]. IEEE journal on emerging and selected topics in circuits and systems, 2019, 9(1): 133-148. DOI:10.1109/jet-cas.2018.2885981
- [3] KATHARIYA B, LI L, LI Z, et al. Scalable point cloud geometry coding with binary tree embedded quadtree [C]//2018 IEEE International Conference on Multimedia and Expo (ICME). San Diego, CA, USA: IEEE, 2018: 1-6. DOI:10.1109/icme.2018.8486481
- [4] XU Y Z, ZHU W J, XU Y L, et al. Dynamic point cloud geometry compression via patch-wise polynomial fitting [C]//ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Brighton, United Kingdom: IEEE, 2019: 2287-2291. DOI:10.1109/icassp.2019.8682413
- [5] BUDAGAVI M, FARAMARZI E, HO T, et al. Samsung response to CFP for point cloud compression (Category 2) [R]. Macau, China, 2017
- [6] HE L Y, ZHU W J, XU Y L. Best-effort projection based attribute compression for 3D point cloud [C]//2017 23rd Asia-Pacific Conference on Communications (APCC). Perth, Australia: IEEE, 2017: 1-6. DOI:10.23919/apcc.2017.8304078
- [7] ZHANG C, FLORENCIO D, LOOP C. Point cloud attribute compression with graph transform [C]//2014 IEEE International Conference on Image Processing (ICIP). Paris, France: IEEE, 2014: 2066-2070. DOI:10.1109/icip.2014.7025414
- [8] DE QUEIROZ R L, CHOU P A. Transform coding for point clouds using a Gaussian process model [J]. IEEE transactions on image processing, 2017, 26(7): 3507-3517. DOI:10.1109/tip.2017.2699922
- [9] DE QUEIROZ R L, CHOU P A. Compression of 3D point clouds using a region-adaptive hierarchical transform [J]. IEEE transactions on image processing, 2016, 25(8): 3947-3956. DOI:10.1109/tip.2016.2575005
- [10] GU S, HOU J H, ZENG H Q, et al. 3D point cloud attribute compression using geometry-guided sparse representation [J]. IEEE transactions on image processing, 2020, 29: 796-808. DOI:10.1109/tip.2019.2936738
- [11] COHEN R A, TIAN D, VETRO A. Point cloud attribute compression using 3-D intra prediction and shape-adaptive transforms [C]//2016 Data Compression Conference (DCC). Snowbird, UT, USA: IEEE, 2016: 141-150. DOI:10.1109/dcc.2016.67
- [12] LI L, LI Z, ZAKHARCHENKO V, et al. Advanced 3D motion prediction for video-based dynamic point cloud compression [J]. IEEE transactions on image processing, 2020, 29: 289-302. DOI:10.1109/tip.2019.2931621
- [13] QUACH M, VALENZISE G, DUFAUX F. Learning convolutional transforms for lossy point cloud geometry compression [C]//2019 IEEE International Conference on Image Processing (ICIP). Taipei, Taiwan, China: IEEE, 2019: 4320-4324. DOI:10.1109/icip.2019.8803413

作者简介



李厚强, 中国科学技术大学教授; 主要研究领域为视频编码与通信、图像处理与计算机视觉、多媒体信息检索等; 主持国家自然科学基金重点项目、“973”项目、“863”项目等国家级科研项目 10 余项; 获 2019 年国家技术发明二等奖(排名第 2)、2015 年国家自然科学二等奖(排名第 2)、2012 年安徽省科学技术一等奖(排名第 1); 发表论文 200 余篇, 获授权发明专利 60 余项, 被视频编码国际标准采纳提案 45 项。



李礼, 中国科学技术大学特任研究员; 主要研究领域为图像视频编码、3D 点云编码等; 获 2019 年国家技术发明二等奖(排名第 5); 发表论文 50 余篇, 获授权发明专利 9 项, 被视频编码国际标准采纳提案 8 项。



李竹, 美国密苏里大学堪萨斯分校副教授; 主要研究领域为图像视频编码、图像视频处理、图像视频通信等; 获国际会议 ICIP 2006 最佳论文奖; 发表论文 100 余篇, 获授权美国发明专利 40 余项。