



浙江大學
ZHEJIANG UNIVERSITY

数据可视化技术期末作业

Instant Neural Graphics Primitives with a
Multiresolution Hash Encoding 读书报告

姓名：_____李亚伦_____

学号：_____22351118_____

学院：_____软件学院_____

2023 年 1 月 14 日

摘要

全连接神经网络参数化的神经图形原语训练和评估成本可能很高。论文通过一种多功能的新输入编码来降低训练成本，该编码允许使用较小的网络而不牺牲质量，从而显著减少浮点运算和内存访问操作的数量：一个小型神经网络通过一个多分辨率哈希表增强，该表包含可通过随机梯度下降优化的可训练特征向量。多分辨率结构允许网络消除哈希冲突，使得架构简单，易于在现代 GPU 上并行化。论文利用这种并行性，使用全融合 CUDA 内核实现整个系统，重点是最小化浪费的带宽和计算操作。作者实现了几个数量级的组合加速，使得在数秒内训练高质量神经图形基元成为可能，并且能够在 1920×1080 的分辨率下以数十毫秒的速度进行渲染。

关键词：神经图形原语，多分辨率哈希，GPU 并行化

1 引言

计算机图形基元本质上是由参数化外观的数学函数表示的。数学表示的质量和性能特征对于视觉真实性至关重要：作者希望表示法在捕捉高频、局部细节的同时保持快速和紧凑。先前方法的编码，将神经网络输入映射到更高维度的空间，这对于从紧凑模型中提取高近似质量至关重要。这样的数据结构依赖于启发式和结构修改（如剪枝、分割或合并），这可能会使训练过程复杂化，限制方法用于特定任务，或限制在 GPU 上的性能，其中控制流程和指针追踪是昂贵的。

作者通过多分辨率哈希编码解决了这些问题，该编码适应性强且高效，不依赖于任务。作者通过多分辨率哈希表层次结构来实现完成任务独立适应性和效率：

适应性：作者将一系列网格映射到相应的固定大小特征向量数组。在粗分辨率下，网格点与数组条目之间存在 1:1 映射。在细分辨率下，数组被视为哈希表，并使用空间哈希函数进行索引，其中多个网格点别名为每个数组条目。这样的哈希冲突导致碰撞训练梯度平均，意味着最大的梯度——那些与损失函数最相关的——将占主导地位。因此，哈希表自动优先处理具有最重要细节的稀疏区域。与以往的工作不同，训练期间无需对数据结构进行任何结构更新。

效率：作者的哈希表查找是 $O(1)$ 的，不需要控制流。这非常适合现代 GPU，避免了树遍历中固有的执行分歧和串行指针追踪。所有分辨率的哈希表可以并行查询。作者在四个代表性任务中验证了多分辨率哈希编码的有效性：（1）千兆像素图像：MLP 学习从 2D 坐标到高分辨率图像的 RGB 颜色的映射。（2）神经符号距离函数（SDF）：MLP 学习从 3D 坐标到表面距离的映射。（3）神经辐射缓存（NRC）：MLP 学习蒙特卡洛路径。

2. 方法

2.1 频率编码

频率编码使用三角函数将低频信号转换为高频信号，这是通过线性变换来完成的。参数编码除了权重和偏差之外，还增加了新的数据结构，如网格或树，允许根据这些结构内的输入向量进行值的插值。

这种方法以更高的内存使用量换取了减少的计算成本。在完全连接的 MLP 网络中，每个通过网络反向传播的梯度都必须更新每个权重。然而，对于可训练的输入编码参数（“特征向量”），只有一小部分受到影响。例如，使用三线性插值计算网格中一个点的值只涉及周围的八个点。

尽管参数更多，但在训练期间所需的 FLOPs 和内存访问次数并没有显著增加。这种方法可以通过减小 MLP 的大小来提高训练速度，同时保持质量。

2.2. 参数编码

参数编码是在权值和偏置之外增加一种辅助数据类型（如网格或树）的结果。在这些结构（网格或树）上，可以根据输入向量使用插值法获得数值。这种安排以较小的计算成本换取了较大的内存占用：对于通过网络向后传播的每个梯度，全连接 MLP 网络中的每个权重都必须更新，而对于可训练的输入编码参数（“特征向量”），只有极少数参数会受到影响。例如，网格类型中一个点的位置值是用三线插值法计算的，只需对周围的八个点进行计算。这样，尽管参数编码中的参数总数远高于固定输入编码，但训练过程中执行更新所需的 FLOP 和内存访问次数并没有显著增加。通过减小 MLP 的大小，可以在不降低质量的情况下提高训练速度

2.3. 稀疏参数编码

密集型 Grid 中的可训练参数比神经网络中的权重使用更多的内存

密集型网格有两种类型的浪费：在空间中的空白部分，也会进行分配，也会进行特征计算，但是它们是无用的。参数的数量是 N 的三次方，但是有用的物体表面的数据是 N 的 2 次方，这个 N 可以认为是分辨率。但是密集型网格会容易学习的过度平滑

2.4 多分辨哈希编码

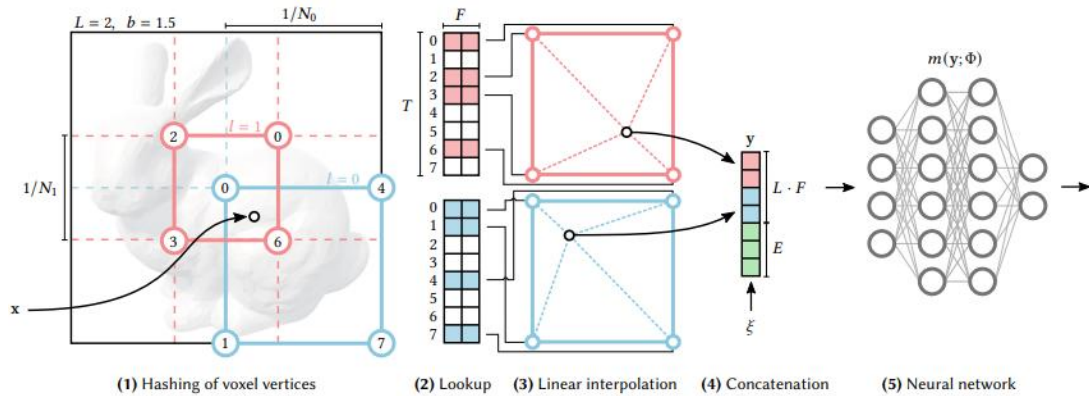


图 1 多分辨哈希编码过程

这种编码是作者在文中提出的一种方法，这种编码能在广泛的应用中提高逼近质量和训练速度，同时又不会产生明显的性能开销。作者的神经网络不仅有可训练的权重参数 Φ ，还有可训练的编码参数 θ 。这些参数被排列成多个层级，每个层级最多包含 T 个特征向量，其维度为 F 。每个层（图中红色和蓝色显示的是其中的两层）都是独立的，在概念上将特征

向量存储在一个网格的顶点上。每层所代表的分辨率，是最粗糙的分辨率和最精细的分辨率之间的插值：

$$N_l := \lfloor N_{min} \cdot b^l \rfloor$$
$$b := \exp\left(\frac{\ln N_{max} - \ln N_{min}}{L - 1}\right)$$

每层的分辨率以 N_{min} 为系数，乘上 b 的 l 层指数。

在作者的所有应用中都是一个有利的帕累托最优值，因此作者在所有其他结果中都使用了这些值，并推荐使用这些值。因此，作者在所有其他结果中都使用了这些值，并推荐将其作为默认值。隐式哈希碰撞解决这种编码能在存在哈希碰撞的情况下忠实地重建场景。这种编码能在哈希碰撞的情况下忠实地重建场景。其成功的关键在于不同的解码级别有不同的优势，可以相互补充互补。较粗的分辨率以及整个编码是注入式，也就是说，它们根本不会发生碰撞。但是，它们只能表示场景的低分辨率版本，因为它们提供的特征是线性插值的。它们提供的特征是从间距很大的点网格中线性插值出来的。点网格线性插值的特征。

相反，精细等级由于网格分辨率较高，可以捕捉到较小的特征但却会出现许多碰撞，也就是说，不同的点会散列到同一个表项中。附近的输入具有相等整数坐标不视为碰撞；当不同的整数坐标散列到同一索引时，就会发生碰撞。相同的索引时，就会发生碰撞。但是，这种碰撞是伪随机散布在空间中的在空间中，而且不太可能在给定的一对点的每个层级同时发生。同时发生。

当训练样本以这种方式发生碰撞时，它们的梯度就会平均。考虑到这些样本对最终重建的重要性样本的重要性很少相同。例如，辐射场可见表面上的一个点会对重建图像有很大的贡献（具有高可见度和高密度）。图像的贡献很大（具有高可见度和高密度，二者都会成倍地影响梯度的大小影响梯度的大小），导致其表项发生较大变化。表项的变化，而空地上的一个点如果恰好指向同一表项而空白处的一个点如果恰好指向同一个条目，其权重就会小得多。因此，更重要样本的梯度将主导碰撞平均值，而别名表项自然就会被优化，从而反映出碰撞的需求。的方式进行优化，以反映权重较高点的需求。哈希编码的多分辨率覆盖了从粗分辨率从保证无碰撞的粗分辨率 N_{min} 到任务所需的最小分辨率 N_{max} 。因此，它因此，它保证了所有可以进行有意义学习的尺度都包括在内，而不考虑稀疏性。

3. 实验



图 2 多分辨率哈希编码逼近分辨率为 $20\,000 \times 23\,466$ 的 RGB 图像

作者使用比 ACORN 小得多的全连接层。这使得作者的方法相比 ACORN，有 $10\times-100\times$ 速度提升的主要原因。作者还认为多分辨率哈希编码的最大作用在于多分辨率哈希编码的最大价值在于它的简单性。ACORN 依靠训练过程，对场景进行自适应细分，而作者提出的多分辨率哈希编码进行自适应细分。

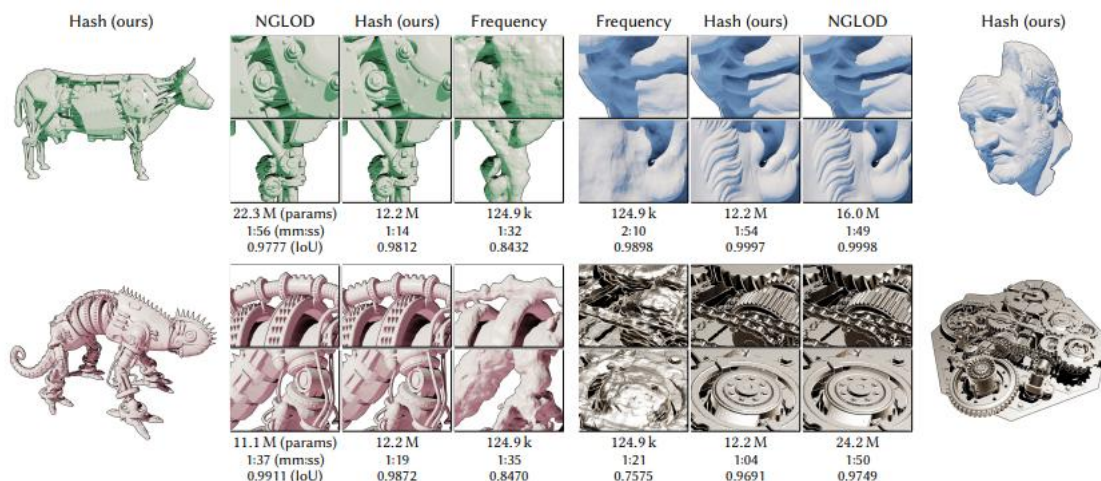


图 3 多分辨率哈希编码在 3 维重建中的效果

作者使用阴影模型将 SDF 可视化。由此产生的颜色对表面法线的微小变化都很敏感，这比其他预测颜色的图形原型更能突出预测中的微小波动。这种敏感性揭示了作者的散列编码在尺度上不希望出现的微观结构。

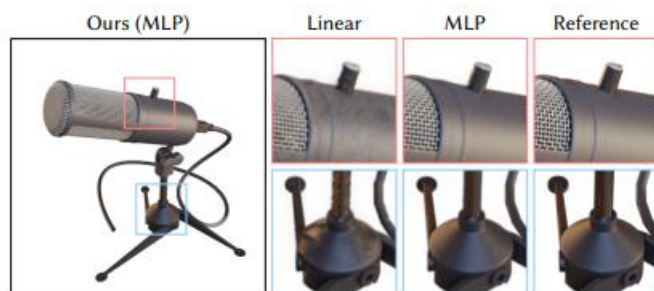


图 4 与其他方法在 NeRF 任务上的对比

作者在模型中使用了 MLP 进行特征处理，MLP 可以解析镜面细节，并减少散列碰撞造成

的背景噪声。由于 MLP 体积小、执行效率高，作者以 15% 的计算成本的增加换取了质量的明显改善。

4. 结论

许多图形问题都依赖于特定任务的数据结构，以利用手头问题的稀疏性或平滑性。而作者提出的多分辨率哈希编码提供了一种实用的基于深度学习的自动关注相关细节的替代方案，不依赖与特定任务。具有低成本的优势，即使在在线训练和推理等时间受限的情况下也能使用。在神经网络输入编码方面，它可以直接替换，例如，它可将 NeRF 的速度提高几个数量级，并与并发的非神经网络编码性能相匹配。与非神经三维重建技术的性能相当。