

DiGS

论文：《DiGS : Divergence guided shape implicit neural representation for unoriented point clouds》

地址： <https://arxiv.org/abs/2106.10811>

年份： CVPR 2022

Introduction

任务： 表面重建

技术贡献：

- (1) 针对 SIREN 的网络结构设计了对应的几何初始化方法；
- (2) 引入关于 SDF 散度的 loss 用于提升重建质量。

Method

Geometric initialization for SIRENs

首先是 SIREN 的网络结构

$$\begin{aligned}\Phi(x; \theta) &= \mathbf{w}_n^T (\phi_{n-1} \circ \phi_{n-2} \circ \dots \circ \phi_0)(x) + b_n, \\ \phi_i(x_i) &= \sin(\mathbf{W}_i x_i + \mathbf{b}_i)\end{aligned}\quad (1)$$

where $\phi_i : \mathbb{R}^{M_i} \rightarrow \mathbb{R}^{N_i}$ is the i^{th} layer of the network, with input $x_i \in \mathbb{R}^{M_i}$ (so $x_0 = x$) and parameters $\theta = \{\mathbf{w}_n, b_n, \mathbf{W}_{n-1}, \mathbf{b}_{n-1}, \dots, \mathbf{W}_1, \mathbf{b}_1\}$ where $\mathbf{W}_i \in \mathbb{R}^{N_i \times M_i}$, $\mathbf{b}_i \in \mathbb{R}^{N_i}$, $\mathbf{w}_n \in \mathbb{R}^{M_n}$, $b_n \in \mathbb{R}$. Rather than approximating a norm using smooth SIRENs, we instead approximate the more tractable signed squared norm [43] and apply the following function to the output of the SIREN:

$$\nu(d) = \text{sign}(d) \sqrt{|d| + \varepsilon}. \quad (2)$$

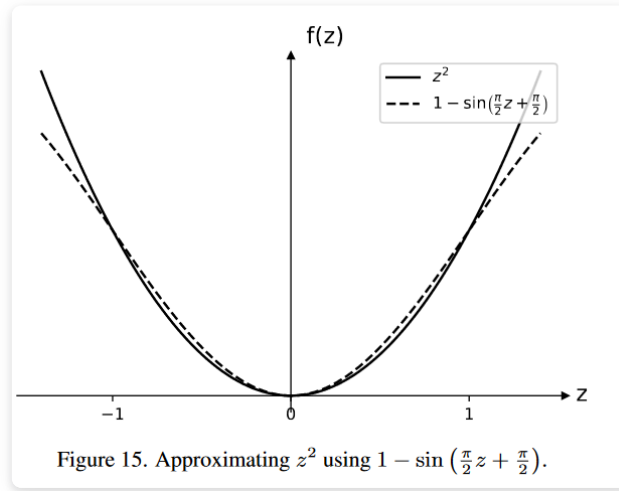
与 SAL 不同的地方就在于激活函数是 \sin 以及 ν 的选择。

由于激活函数的不同，所以需要重新设计初始化的方法。这里是先将点云数据归一化到单位球中，在后续设计几何初始化方法时只关注单位球内的区域。最后在减去一个 r 就可以得到一个半径为 r 的球。

Proposition 4.1. Let Φ be a single hidden layer SIREN ($n = 1$ in Equation 1) of dimension M_n and let x be a point within the unit ball. Set $\mathbf{W}_{n-1} = \frac{\pi}{2}I$, $\mathbf{b}_{n-1} = \frac{\pi}{2}\mathbf{1}$, $\mathbf{w}_n = -\mathbf{1}$ and $b_n = M_n$. Then, $\nu(\Phi(x)) \approx \|x\|_2$.

Proof of Proposition 4.1

定理 1 针对单层 SIREN 的情况，给定相关初始化参数，即可使输出的模长与输入的模长相同。这个定理证明比较简单，具体过程可以看论文附录。主要就是在单位球范围内，我可以用 $1 - \sin(\frac{\pi}{2}z + \frac{\pi}{2})$ 来拟合 z^2 。



有了单层 SIREN 的初始化方法，就要将其推广的多层的情况。论文的做法是将定理 1 给出的参数作为 SIREN 中最后一层参数的初始化方法，对于前面的 $n - 2$ 层，想要保持输出的模长的期望一直是 $\|x\|_2$ ，也就是有 $\mathbb{E}[\|\phi_i(x)\|_2] = \|x\|_2$ 这里一共分了 3 个小引理来证明。

Proof of Proposition 4.2

Lemma A.1.

Lemma A.1. Let X, Y be two random d -dimensional vectors with elements X_i and Y_i sampled i.i.d. from $\mathcal{U}(-\sqrt{\frac{3}{d}}, \sqrt{\frac{3}{d}})$. Then $\mathbb{E}[\|X\|] = \mathbb{E}[\|Y\|] = 1$ and $\mathbb{E}[\langle X, Y \rangle] = 0$.

X, Y 都是 d 维向量，其中每个元素都是从 $\mathcal{U}(-\sqrt{\frac{3}{d}}, \sqrt{\frac{3}{d}})$ 中采样得到。

我们很容易知道 $\mathbb{E}[X_i] = \mathbb{E}[Y_i] = 0$, $\text{var}(X_i) = \text{var}(Y_i) = \frac{1}{d}$, $\mathbb{E}[X_i^2] = \mathbb{E}[Y_i^2] = \frac{1}{d}$ 然后用了弱大数定律的切比雪夫不等式：对于任何随机变量 X 和实数 $b > 0$ ，都有

$$P(|X - \mathbb{E}(X)| \geq b) \leq \frac{\text{Var}(X)}{b^2}$$

这里我们令上式中的 X 为 $\frac{X_1^2 + \dots + X_d^2}{d}$ ，因此上式中的 $\mathbb{E}(X) = \mathbb{E}(X_i)$, $\text{Var}(X) = \frac{\text{Var}(X_i^2)}{d}$ ，带入到公式中就和论文原文一致了，最后的结果就是 d 越大则 $\|X\|_2^2$ 越接近于 1。因此就会有 $\mathbb{E}[\|X\|_2] = \mathbb{E}[\|Y\|_2] = 1$

Lemma A.2.

引理 2 很简单，就是把引理 1 的结论带入，相当于经过网络的线性层后 x 模长不变。

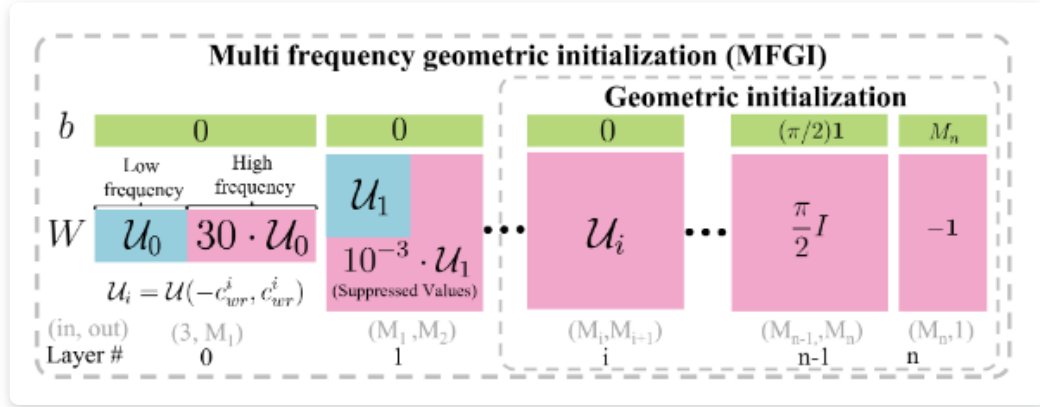
Lemma A.3.

引理 3 就是处理激活函数 \sin ，具体就是让矩阵的行数 $\gg 10$ ，而模长是固定为 1 的，相当于降低了各个元素的值大小，使得能够应用近似 $\sin(x) \approx x$

$$\begin{aligned}
\|x_i\|_2 &\approx \|\mathbf{W}_i x_i\|_2 & (*) \\
&\approx \|\sin(\mathbf{W}_i x_i + \mathbf{b}_i)\|_2 & (\dagger) \\
&= \|x_{i+1}\|_2 & (31)
\end{aligned}$$

最后就是将几个引理拼起来，经过 SIREN 每一层后模长不变。

Multi frequency geometric initialization

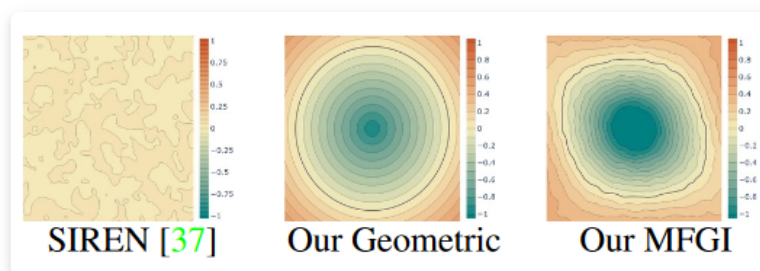


文章参照 SIREN 提出了一种引入高频信息的几何初始化方法，具体如上图所示 (这个图里的 W 好像是安装右乘输入向量的方式排列的，下面描述的时候我是按照左乘来描述的，会有行和列转置的差别)。
对于 $\mathbf{W}_0 \in \mathbb{R}^{N_0 \times M_0}$ 的前 k_r ，我们保持初始化的原样，对于其余的 $N_0 - k_r$ 行，我们乘上一个 scale $n_p = 30$ ，这样对输入向量做变换后，输出向量的前 k_r 个元素会在 \sin 的第一个周期内，其余元素会在第 n_p 个周期内，这样引入高频信息。

为了保证几何初始化仍然有效，就需要通过下一个 \mathbf{W}_1 ，来对这些值进行抑制，我们想要抑制向量中后面 $N_0 - k_r$ 个元素，因此可以对 \mathbf{W}_1 中的后面 $N_0 - k_r$ 列进行缩放 $s = 10^{-3}$ ，但这样与论文还是不同，论文中还对 $N_0 - k_r$ 行进行了缩放，可以看 <https://zhuanlan.zhihu.com/p/650679177> 中有作者对这样做的解释：

“To see that the L_shapes suppression does this, the first k_r elements of x_2 come from the first k_r elements of x_1 that only hit one period without any suppression, as well as elements of x_1 that hit up to n_p periods with suppression. Thus the first k_r elements of x_2 can be considered to mostly have only hit one period. On the other hand, the remaining $N_1 - k_r$ elements come from elements of x_1 that hit one or up to n_p periods with equal amounts of suppression, so we can consider it to have hit up to n_p periods but in general be suppressed.”

大致的意思是，如果采用第一种做法，得到的结果 x_2 会是 x_1 中前 k_r 个元素以及受抑制的 $N_0 - k_r$ 个元素，这样所有的元素都会是位于第一个周期的，因为前 k_r 个元素占主导；采用第二种方法的话， x_2 前 k_r 个元素不变，其余元素是来自 x_1 中受抑制的所有元素，相当于是相对地提高了在 n_p 周期的元素的地位，使得结果仍然能够保留一定的高频信息。



可以看到, “Our MFGI initialization is a noisy version of our geometric initialization, where the amount of noise can be tuned by np, kr and s.”

DiGS Loss

先列举一下训练 SDF 可能会用到的 loss:

(A) manifold constraint: 位于表面的点必须在 SDF 的 zero level set 上:

$$L_A = \int_{\Omega_0} \|\Phi(x; \theta)\|_2 dx$$

(B) normal constraint: 位于表面的点的梯度需要与法向一致 (如果有):

$$L_B = \int_{\Omega_0} (1 - \langle \nabla_x \Phi(x; \theta), n_{GT}(x) \rangle) dx$$

(C) (C1) non-manifold constraint: 不在表面上的点需要与 ground truth 的 SDF 值一致; (C2) Eikonal constraint: 所有点需要有单位梯度:

$$L_{C2} = \int_{\Omega} |||\nabla_x \Phi(x; \theta)||_2 - 1| dx$$

(D) non-manifold penalisation constraint: 不在表面上的点需要有非零的 SDF 值:

$$L_D = \int_{\Omega/\Omega_0} \exp(-\alpha|\Phi(x; \theta)|) dx$$

以上这些 loss 中, (A) 和 (C) 是必要的, (B) 和 (D) 是用来提升效果的。

如果有连续的约束, 只用 (A) 和 (C2) 是能得到唯一解的, 但实际输入是离散的, 所以会导致有无穷多解, 因此需要其他 loss 来得到更好的解。

Second order unsupervised constraints

因此文章就希望在 Ω/Ω_0 上找二阶信息来约束 SDF。给定梯度场 $\nabla \Phi$, 可以计算旋度 $\nabla \times \nabla \Phi$ 和散度 $\nabla \cdot \nabla \Phi$ 。旋度会在每一处都等于零 (<https://zhuanlan.zhihu.com/p/336928427>), 因此不能带来帮助。文章观察到了在绝大多数区域中散度的值都会很小, 因此就引入了以下约束:

$$L_{div} = \int_{\Omega/\Omega_0} |\Delta \Phi(x; \theta)| dx = \int_{\Omega/\Omega_0} |\nabla_x \cdot \nabla_x \Phi(x; \theta)| dx$$

需要注意的是这项 loss 要求神经网络存在二阶导如 SIREN, 对于使用 ReLU 的网络则无法使用。最后文章提了一下 Dirichlet Energy

$$E[\Phi] = \frac{1}{2} \int_{\Omega} \|\nabla \Phi(x)\|_2^2 dx$$

Eikonal term 相当于是去最小化 Dirichlet Energy 了, 因为“if the gradient is constrained to have unit norm on the sampled points, then the Dirichlet energy, as far as can be determined from at those sampled points, is already determined”, 然后加入散度 loss 能够更好地约束, 因为引入了二阶的信息。

Experiments

Failure case 也是不能重建细的物体，以及会有过光滑的现象，一部分原因是这些区域点太少了，没有充足的信息，另一部分原因就是引入的散度 loss，<https://zhuanlan.zhihu.com/p/649921965> 中有提到原因：“一个好的SDF其实只需要其法线方向上的二阶导数为0，如果在切线方向上的二阶导数为0的话，得到的SDF轮廓会非常平滑，不利于学习到一些细节。 L_{div} 同时惩罚法线和切线方向上的二阶导数，所以 DiGS 将很难学习到一些几何细节”