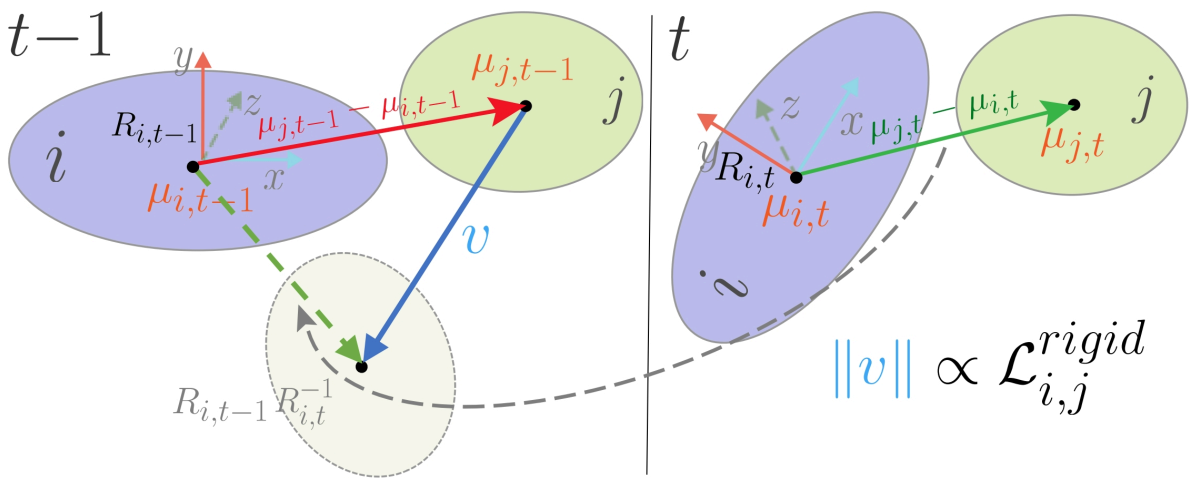
3D Gaussian Splatting 用于动态场景表示

(3DV 2024) Dynamic 3D gaussians: Tracking by persistent dynamic view synthesis



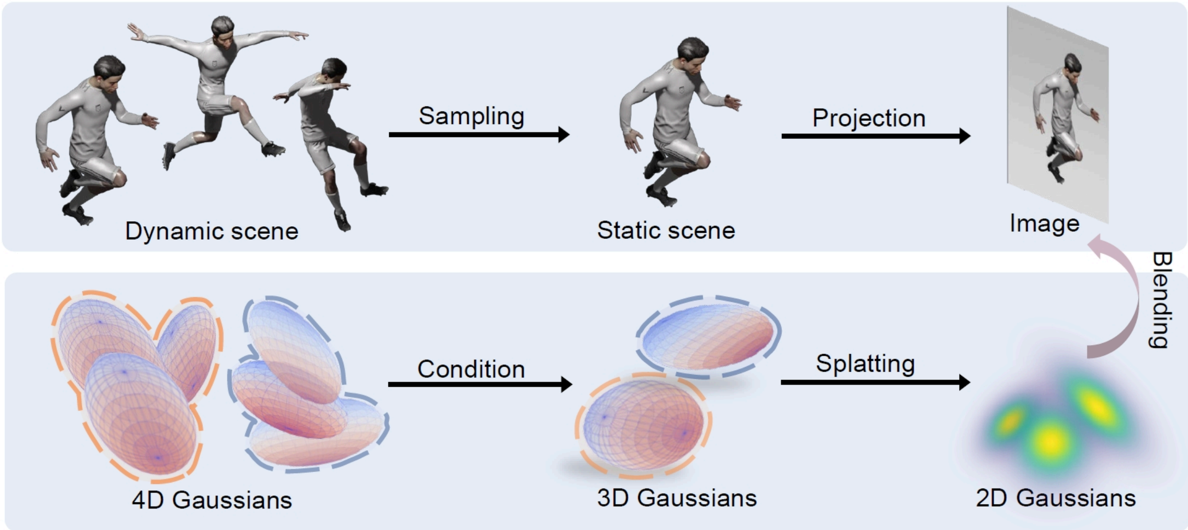
基于物理启发出了三个正则化项加在loss函数里面：

- local-rigidity loss 局部刚度：一个Gaussian绕着某个轴旋转，相近的Gaussian都要跟着绕这个轴转
  $\mathcal{L}_{i,j}^{\text{rigid}} = w_{i,j}\left\|(\mu_{j,t-1} - \mu_{i,t-1}) - R_{i,t-1}R_{i,t}^{-1}(\mu_{j,t} - \mu_{i,t})\right\|_2$
  $\mathcal{L}^{\text{rigid}} = \frac{1}{k|\mathcal{S}|}\sum_{i\in\mathcal{S}}\sum_{j\in\text{knn}_{i;k}}\mathcal{L}_{i,j}^{\text{rigid}}$
- local-rotation similarity 局部旋转相似性：相近的Gaussian相同旋转，作者加它的理由是实验发现效果更好
  $\mathcal{L}^{\text{rot}} = \frac{1}{k|\mathcal{S}|}\sum_{i\in\mathcal{S}}\sum_{j\in\text{knn}_{i;k}} w_{i,j}\left\|\hat{q}_{j,t}\hat{q}_{j,t-1}^{-1} - \hat{q}_{i,t}\hat{q}_{i,t-1}^{-1}\right\|_2$
- local-isometry loss 局部等距：相近的Gaussian相对位置不变，作者加它的理由是上面两个loss容易导致撕裂，加它防止撕裂
  $\mathcal{L}^{\text{iso}} = \frac{1}{k|\mathcal{S}|}\sum_{i\in\mathcal{S}}\sum_{j\in\text{knn}_{i;k}} w_{i,j}\left|\|\mu_{j,0} - \mu_{i,0}\|_2 - \|\mu_{j,t} - \mu_{i,t}\|_2\right|$

其中，这里的“相近的Gaussian”是用K-nearst给每个Gaussian找20个点，并将其距离作为算上述正则化项的权值：

$w_{i,j} = \exp\left(-\lambda_w\|\mu_{j,0} - \mu_{i,0}\|_2^2\right)$

(ICLR 2024) Real-time Photorealistic Dynamic Scene Representation and Rendering with 4D Gaussian Splatting



3D Gaussian形状参数里的位置、rotate和scale都变成4D，相当于3D Gaussian（椭圆）加一个时间维度变成了4D椭圆，渲染是**在时间轴上采样从而将这个4D椭圆投影到3D空间。**

文中介绍的3D Gaussian：

Differentiable rasterization via Gaussian splatting In rendering, given a pixel (u, v) in view $\mathcal{I}$ with extrinsic matrix E and intrinsic matrix K, its color $\mathcal{I}(u, v)$ can be computed by blending visible 3D Gaussians that have been sorted according to their depth, as described below:

$$\mathcal{I}(u,v) = \sum_{i=1}^{N} p_i(u,v;\mu_i^{2d},\Sigma_i^{2d})\alpha_i c_i(d_i) \prod_{j=1}^{i-1}(1 - p_i(u,v;\mu_i^{2d},\Sigma_i^{2d})\alpha_j), \tag{2}$$

where c_i denotes the color of the i-th visible Gaussian from the viewing direction d_i, α_i represents its opacity, and $p_i(u, v)$ is the probability density of the i-th Gaussian at pixel (u, v).

拓展到4D Gaussian：

Problem formulation and 4D Gaussian splatting To extend the formulation of Kerbl et al. (2023) for modeling dynamic scenes, reformulation is necessary. In dynamic scenes, a pixel under view $\mathcal{I}$ can no longer be indexed solely by a pair of spatial coordinates (u, v) in the image plane; But an additional timestamp t comes into play and intervenes. Formally this is formulated by extending equation 2 as:

$$\mathcal{I}(u,v,t) = \sum_{i=1}^{N} p_i(u,v,t)\alpha_i c_i(d) \prod_{j=1}^{i-1}(1 - p_j(u,v,t)\alpha_j). \tag{5}$$

Note that $p_i(u, v, t)$ can be further factorized as a product of a conditional probability $p_i(u, v|t)$ and a marginal probability $p_i(t)$ at time t, yielding:

$$\mathcal{I}(u,v,t) = \sum_{i=1}^{N} p_i(t)p_i(u,v|t)\alpha_i c_i(d) \prod_{j=1}^{i-1}(1 - p_j(t)p_j(u,v|t)\alpha_j). \tag{6}$$

Let the underlying $p_i(x, y, z, t)$ be a 4D Gaussian. As the conditional distribution $p(x, y, z|t)$ is also a 3D Gaussian, we can similarly derive $p(u, v|t)$ as a planar Gaussian whose mean and covariance matrix are parameterized by equation 3 and equation 4, respectively.

缩放3D变4D在数学上是3x3对角矩阵变成4x4对角矩阵S；旋转3D变4D在数学上是两个啥矩阵相乘得到R；
从而4D Gaussian均值（中心点坐标）变成4维μ、协方差也变成4x4矩阵Σ：

Representation of 4D Gaussian To address the mentioned challenge, we suggest to treat time and space dimensions equally by formulating a coherent integrated 4D Gaussian model. Similar to Kerbl et al. (2023), we parameterize its covariance matrix Σ as the configuration of a 4D ellipsoid for easing model optimization:

$$\Sigma = RSS^T R^T, \tag{7}$$

where S is a scaling matrix and R is a 4D rotation matrix. Since S is diagonal, it can be completely inscribed by its diagonal elements as $\boxed{S = \mathrm{diag}(s_x, s_y, s_z, s_t)}$. On the other hand, a rotation in 4D Euclidean space can be decomposed into a pair of isotropic rotations, each of which can be represented by a quaternion.

Specifically, given $q_l = (a, b, c, d)$ and $q_r = (p, q, r, s)$ denoting the left and right isotropic rotations respectively, R can be constructed by:

$$\boxed{R = L(q_l)R(q_r) = \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix} \begin{pmatrix} p & -q & -r & -s \\ q & p & s & -r \\ r & -s & p & q \\ s & r & -q & p \end{pmatrix}}. \tag{8}$$

The mean of a 4D Gaussian can be represented by four scalars as $\boxed{\mu = (\mu_x, \mu_y, \mu_z, \mu_t)}$. Thus far we arrive at a complete representation of the general 4D Gaussian.

最后，每个时刻的3D Gaussian是从这个4D Gaussian中采样而来（4D椭圆于t时刻在3D空间中的一个投影）：

Subsequently, the conditional 3D Gaussian can be derived from the properties of the multivariate Gaussian with:

$$\begin{aligned} \mu_{xyz|t} &= \mu_{1:3} + \Sigma_{1:3,4}\Sigma_{4,4}^{-1}(t - \mu_t), \\ \Sigma_{xyz|t} &= \Sigma_{1:3,1:3} - \Sigma_{1:3,4}\Sigma_{4,4}^{-1}\Sigma_{4,1:3} \end{aligned} \tag{9}$$

Since $p_i(x, y, z|t)$ is a 3D Gaussian, $p_i(u, v|t)$ in equation 6 can be derived in the same way as in equation 3 and equation 4. Moreover, the marginal $p_i(t)$ is also a Gaussian in one-dimension:

$$p(t) = \mathcal{N}(t; \mu_4, \Sigma_{4,4}) \tag{10}$$

So far we have a comprehensive implementation of equation 6. Subsequently, we can adapt the highly efficient tile-based rasterizer proposed in Kerbl et al. (2023) to approximate this process, through considering the marginal distribution $p_i(t)$ when accumulating colors and opacities.

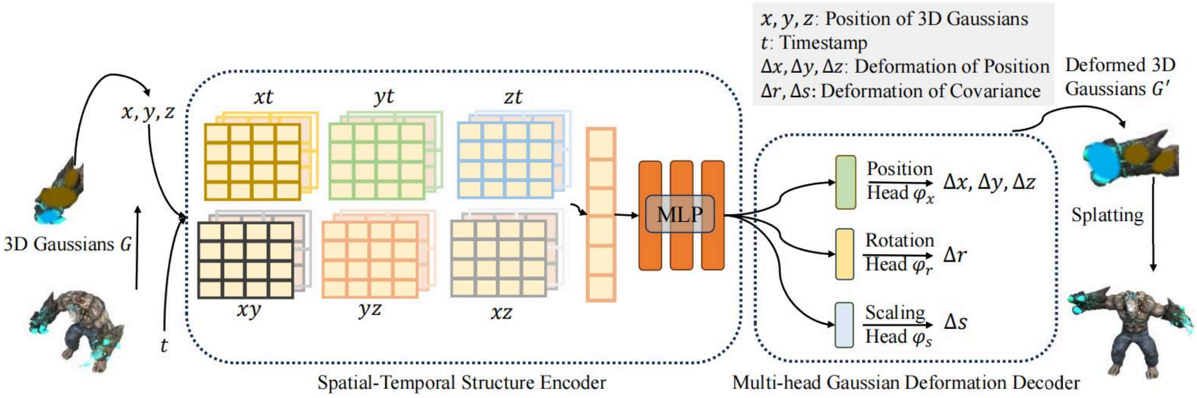
球谐系数加上一个维度用傅里叶级数组成的函数表示：

Inspired by studies on head-related transfer function, we propose to represent $c_i(d, t)$ as the combination of a series of 4D spherindrical harmonics (4DSH) which are constructed by merging SH with different 1D-basis functions. For computational convenience, we use the Fourier series as the adopted 1D-basis functions. Consequently, 4DSH can be expressed as:

$$Z_{nl}^m(t, \theta, \phi) = \cos\left(\frac{2\pi n}{T} t\right) Y_l^m(\theta, \phi), \tag{11}$$

where Y_l^m is the 3D spherical harmonics. The index $l \geq 0$ denotes its degree, and m is the order satisfying $-l \leq m \leq l$. The index n is the order of the Fourier series. The 4D spherindrical harmonics form an orthonormal basis in the spherindrical coordinate system.

4D Gaussian Splatting for Real-Time Dynamic Scene Rendering



高斯点云只有一个，用Triplane存储运动信息。

任意高斯点位置xyz和时间t输入Triplane得到位移/旋转/缩放的变化情况，从而对高斯点进行变换。

（只有形状方面的变换，没有颜色和球谐系数的变化）