

## I. 方法

本章给出所提 **SiT-PVG** 的完整方法。按顺序介绍：预备知识 (§3.1)，2D→4D 语义蒸馏 (§3.2)，双线索动态掩码 (§3.3)，语义驱动的时间约束 (§3.4)，时间一致性增强 (§3.5)，以及优化与实现细节 (§3.6)。

### A. 预备知识：3DGS 与 PVG

**3D Gaussian Splatting (3DGS)** 以一组各向异性的高斯基元集合显式建模场景。每个基元包含空间中心、各向异性形状与朝向、不透明度与外观参数；通过可微光栅化与按深度排序的透明度融合实现高效渲染与快速收敛。高斯基元的数学表达式为：

$$G_i(\mathbf{x}) = \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)^\top \boldsymbol{\Sigma}_i^{-1}(\mathbf{x} - \boldsymbol{\mu}_i)\right). \quad (1)$$

其中， $\boldsymbol{\mu}_i$  为三维位置； $\boldsymbol{\Sigma}_i$  为协方差矩阵，描述形状与朝向，通常由旋转矩阵  $\mathbf{R}$  与尺度矩阵  $\mathbf{S}$  参数化（如  $\boldsymbol{\Sigma} = \mathbf{R}\mathbf{S}\mathbf{S}^\top\mathbf{R}^\top$ ）。将三维高斯投影到像平面得到 2D 高斯，其投影协方差由

$$\boldsymbol{\Sigma}' = \mathbf{J}\mathbf{W}\boldsymbol{\Sigma}\mathbf{W}^\top\mathbf{J}^\top \quad (2)$$

给出，其中  $\mathbf{W}$  是世界到相机的外参变换 ( $\text{SE}(3)$ )， $\mathbf{J}$  为透视投影的雅可比近似。像素颜色采用按深度排序的  $\alpha$ -融合：

$$\mathbf{C} = \sum_{i=1}^N T_i \alpha_i \mathbf{c}_i, \quad T_i = \prod_{j<i} (1 - \alpha_j), \quad (3)$$

其中  $\alpha_i$  由点元不透明度与其投影协方差在该像素的覆盖贡献共同决定， $\mathbf{c}_i$  为外观（如球谐系数着色）。上述表示配合基于瓦片的可微光栅化，使 3DGS 在静态场景中实现实时渲染和快速收敛。

3DGS 在建模上默认静态：点元参数随时间不变，难以直接刻画道路场景中普遍存在的时变要素（车辆、行人等）。为此，**Periodic Vibration Gaussians (PVG)** 在 3DGS 的最小改动上引入时间参数化：令点元的空间位置与不透明度随时间围绕“寿命峰值” $\tau$  作可微振荡与衰减。具体地，对每个点元引入周期长度  $l$ 、速度方向/幅度  $v$ 、寿命尺度  $\beta$ ，定义

$$\tilde{\mu}(t) = \mu + \frac{l}{2\pi} \sin\left(2\pi \frac{t - \tau}{l}\right) v, \quad \tilde{o}(t) = o \cdot \exp\left(-\frac{1}{2}(t - \tau)^2 \beta^{-2}\right). \quad (4)$$

此时点元在时刻  $t$  的状态为  $H(t) = \{\tilde{\mu}(t), q, s, \tilde{o}(t), c\}$ ，整幅图像按

$$\hat{I}_t = \text{Render}(\{H_i(t)\}_{i=1}^N; E_t, I_t) \quad (5)$$

渲染。为衡量点元“静态度”，定义

$$\rho = \beta/l, \quad (6)$$

$\rho$  越大表示寿命相对周期更长、越趋近静止；当  $v = 0$  且  $\rho \rightarrow \infty$  时，PVG 退化回标准 3DGS。由此，静态/动态以统一参数化出现，仅通过  $\{v, \beta, l, \tau\}$  的取值加以区分。PVG 以最小改动继承了 3DGS 的高效与可扩展性，同时补足了动态建模与可编辑性，是本文面向道路环境的更合适表征选择。