

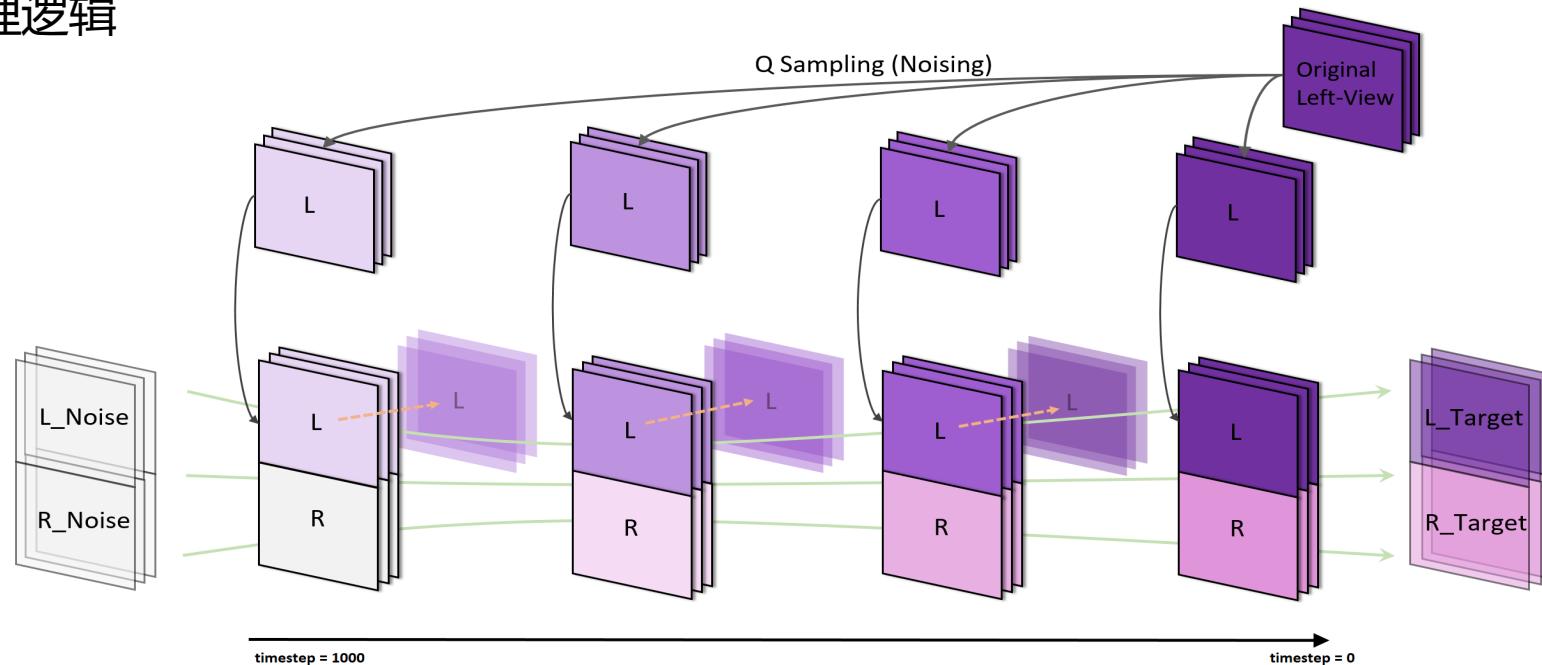
Stereo Prediction from DiT without Inductive Bias

一种单目生成双目视频的训推方法

Click Each One to Play



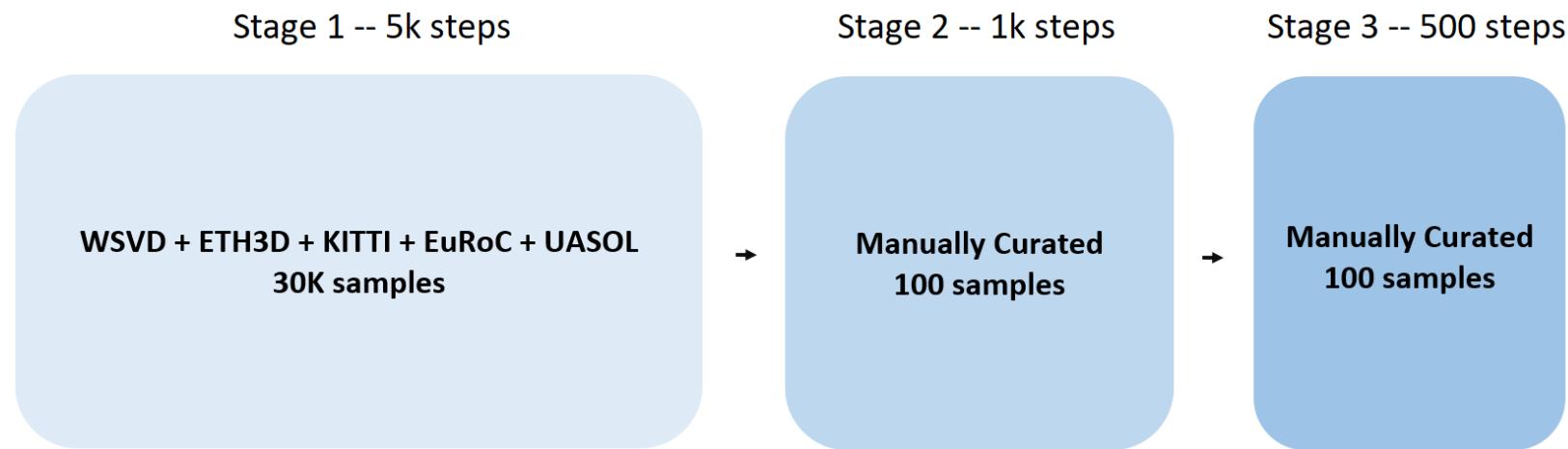
推理逻辑



推理过程：双目视频左右一起一步步去噪，每次迭代用原左眼加噪的特征来替换去噪过程中预测的左眼特征。

注：不同于SynCamMaster等方法的多视角“翻拍”，双目视频的左右视角对画面一致性要求更高。
现有的StereoCrafter等单目预测双目需要depth prediction + warp/disparity等规则类先验处理。

训练数据



三个训练阶段：1. 双目视频预训练，2. 训练左目预测右目的因果推理，3. 多阶段多任务左预测右，保持训推一致。

注：数据管道中低质量情形：

1. 左右视角颠倒，
2. 画面边界草木/中央人物的数据比例过多，
3. 左右时序未对齐，
4. 非全帧数立体(如幻灯片式移进移出)，
5. 填充边界的黑色部分过多，
6. 视差微弱，
7. 景深多样性缺乏，
8. 镜头动、物体不动(或物体动、镜头不动)比例过高，
9. 近景/远景、真实/合成数据量不平衡。

训练方式

