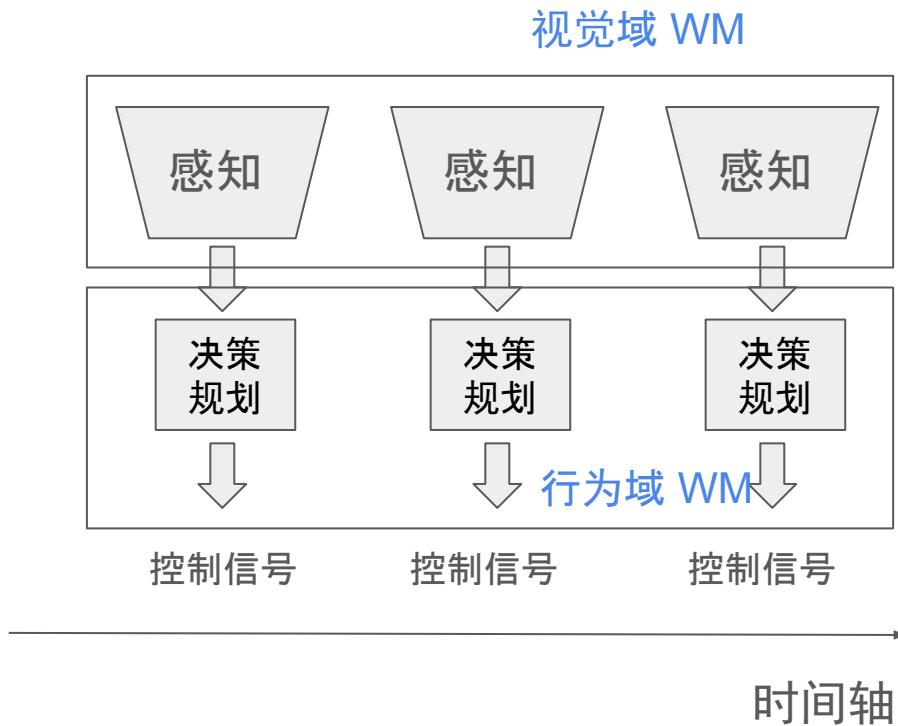
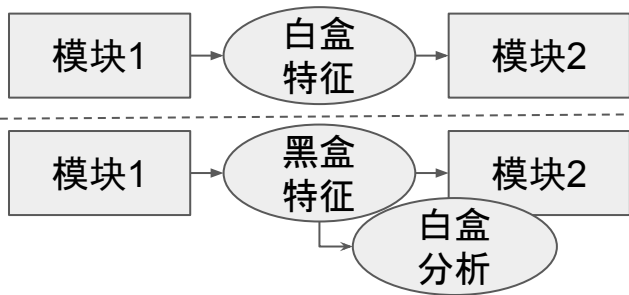


基于世界模型的生成式自动驾驶

世界模型在端到端中的应用

- 利用无标注数据
 - Next-token prediction 范式
 - 消化全量回流数据, 形成数据 飞轮
- 应用 RLHF (强化学习人类反馈)
 - 将人类反馈纳入数据迭代
 - 提供白盒分析面
- 减少工程环节, 提升组织对齐
 - 从模型整合到数据整合
 - 世界模型提供多视角重放能力: “共同看见”



世界模型利用全量回流数据

- 世界模型对场景、他车、自车统一时序建模
 - 场景的光影、天气、事件 变化
 - 更好处理路权博弈
 - 自车车况掌握
- 全量信息利用
 - 视觉信息不依赖标注(视觉域 WM)
 - 不依赖用户反馈(差多个数量级), 直接建模用户(行为域 WM)
 - 利用声音、车流状态等标注困难信息: 音视频联合建模
- 挑战
 - 跨车辆、跨车型的数据利用
 - 保核心信息的低码率高分辨率 tokenizer
 - 幻觉

工作积累(视觉域 WM)

DreamLLM: 生成提升理解


		ID	$\mathcal{L}_{\text{align}}$	MM-Vet	VQAv2	COCO
0	Stable Diffusion	\times	-	-	-	12.43
1	Creation-only	\times	\times	-	-	8.50
2	Creation-only	\checkmark	\times	-	-	8.57
3	Comprehension-only	\times	\times	31.0	55.1	-
4	Comprehension-only	\checkmark	\times	34.4	54.3	-
5	Joint-learning	\checkmark	\times	35.9	56.6	8.46
6	Joint-learning	\checkmark	\checkmark	N/A	N/A	N/A

Vary: 多头输入专项提升解析力

复合层次化 token 设计

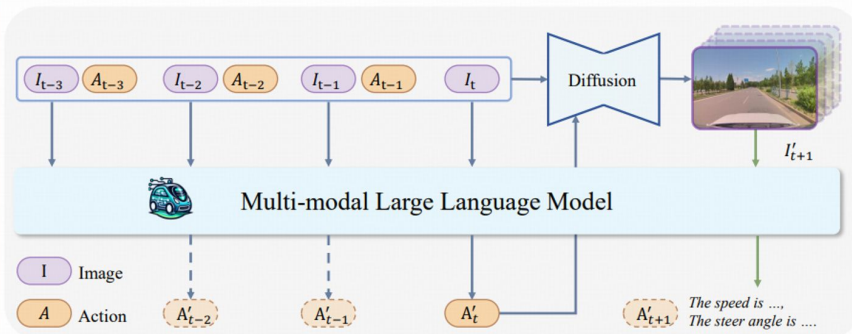
ADriver-I

The frame t is



The action t is

The speed is 3.36,
The steer angle is 3.56
The vehicle is ... a moderate angle, ... a steady speed...



Method	Input→Output	FID↓	FVD↓
DriveGAN	1F→12F	73.4	502.3
DriveDreamer	1F+12B+12M→12F	52.6	452.0
ADriver-I	4F → 4F	5.5	97.0

AR 方案

4F -> 4F

103

工作积累(行为域 WM)

- 行为模型: 直接在用户行为日志上预训练
- Reward Model: 建模人类偏好, 提供过程监督
 - PRM 驱动前进后退
- 增强推理: 从推理时 MCTS 到训练时 MCTS
- 反 Reward Model: 偏好驱动的行为模型
 - 行为模型基础上 fine-tune

用户模行为

个性化建模

MCTS

预训练

后训练

推理

Video PreTraining (VPT): Learning to Act by Watching Unlabeled Online Videos (2022)

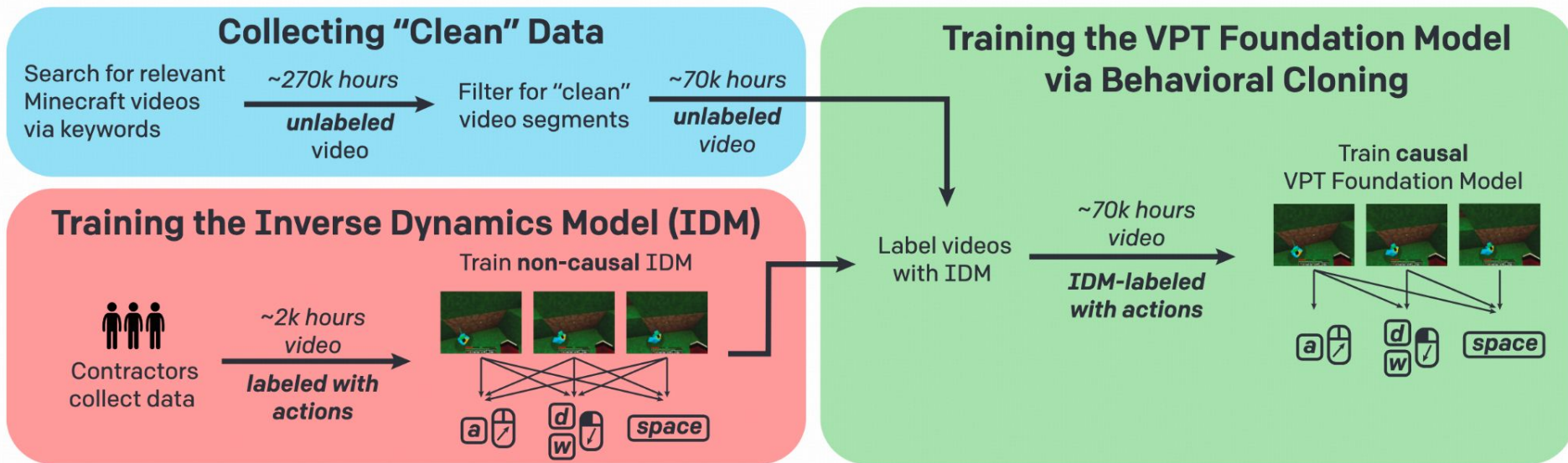


Figure 2: Video Pretraining (VPT) Method Overview.

IDM 对自动驾驶有没有用？

<https://zhuanlan.zhihu.com/p/688921480>

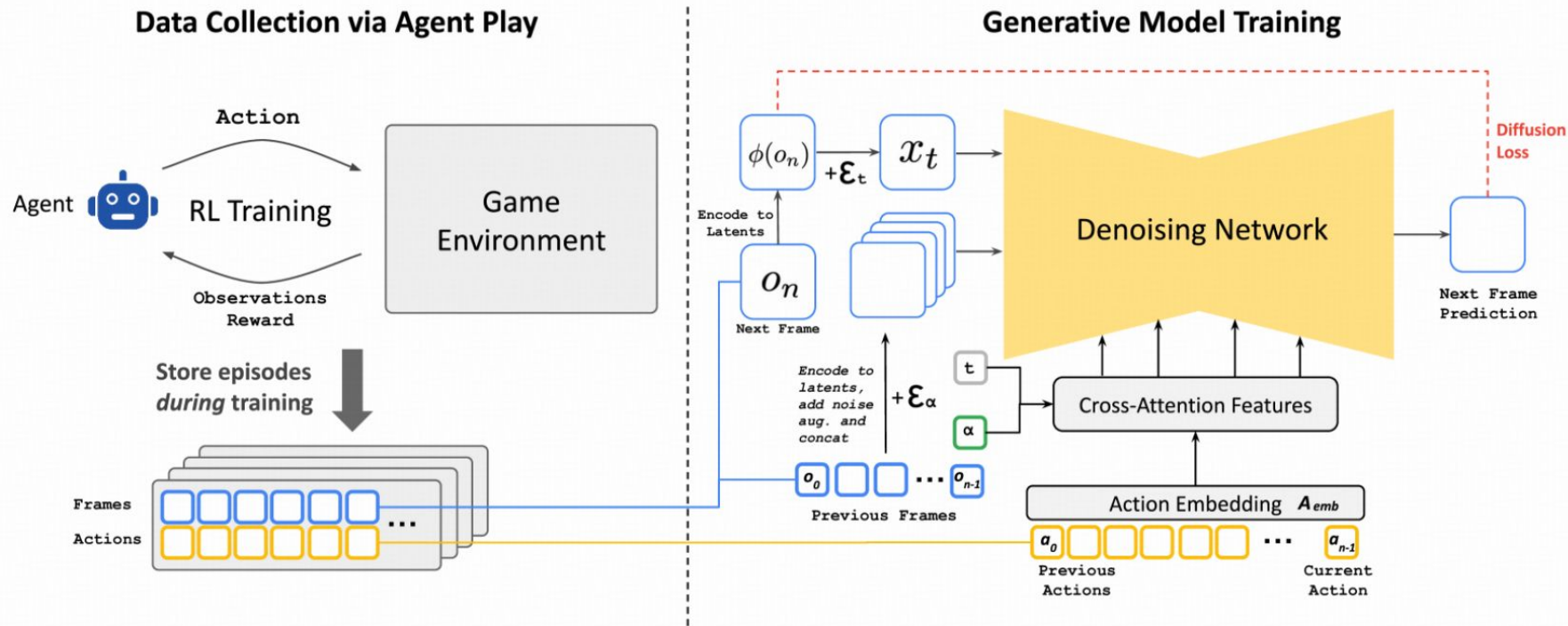


Figure 3: **GameNGen** method overview. v-prediction details are omitted for brevity.

RGB 世界

方块世界

