



home
papers
study
life
About me

Flow-DPO

- Improving Video Generation with Human Feedback
- <https://arxiv.org/abs/2501.13918>
- NeurIPS 2025

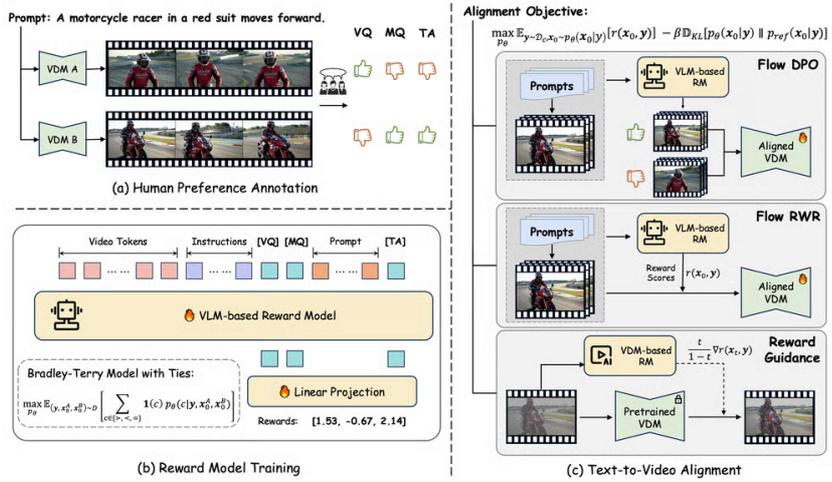


Figure 1: **Overview of Video Alignment Framework.** (a) **Human Preference Annotation.** We construct 182k prompt-video triplets, each annotated on Visual Quality (VQ), Motion Quality (MQ), and Text Alignment (TA). (b) **Reward Modeling.** A VLM-based reward model is trained under the Bradley-Terry-Model-with-Ties formulation. (c) **Video Alignment.** From a unified RL perspective, we introduce three alignment algorithms for flow-based video generation: Flow-DPO, Flow-RWR, and Reward Guidance (Flow-NRG), and provide a systematic comparison.

RLHF
VLM
Flow-DPO,
FlowRWR
Flow-NRG

Flow-DPO

Older

2026-3-5

WildActor

leicheng © 2022-2025

Archive RSS feed GitHub Email QR Code

Made with Montaigne and by anton