

# **FlowStyler: Artistic Video Stylization via Transformation Fields Transports**

Yuning Gong, Jiaming Chen, Xiaohua Ren, Yuanjun Liao, Yanci  
Zhang

Sichuan University

# Motivation & Challenge

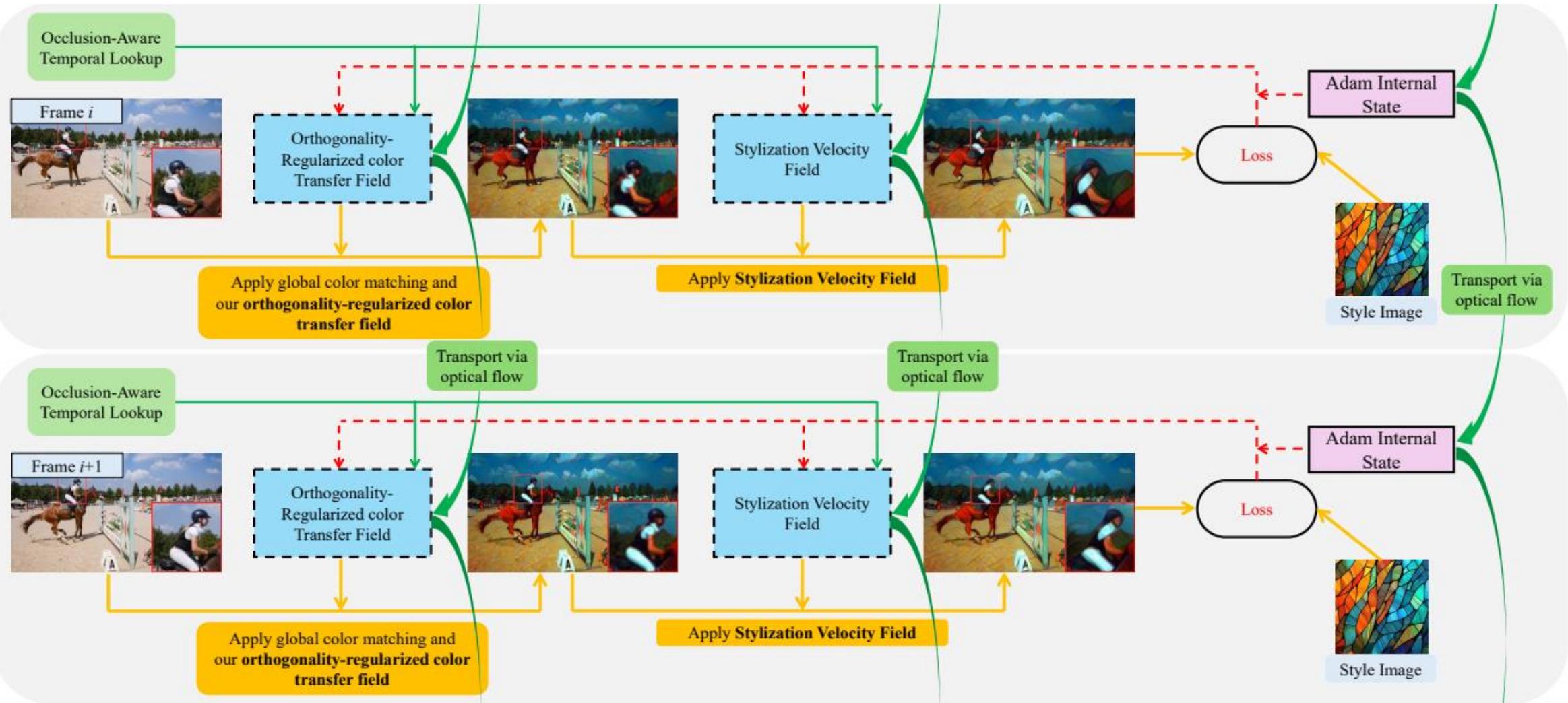
Method	Warping Error ( $10^{-3}$ ) ↓		DINO Score ↑	CLIP Score ↑	Peak GPU Memory (GB) ↓	
	1-frame	10-frame			50 frames	100frames
AnyV2V	25.0	30.7	0.61	0.58	40	62
TokenFlow	11.0	29.5	0.66	0.65	24	64
CAP-VSTNet	4.1	19.8	0.58	0.52	<b>0.44</b>	<b>0.44</b>
UniST	10.6	29.8	0.59	0.51	1.1	1.1
FlowStyler	<b>2.5</b>	<b>14.2</b>	<b>0.72</b>	<b>0.67</b>	3.6	3.6

- **Artistic Video Stylization:** Enables creators to turn ordinary videos into moving works of art.  
Challenge: Balancing *style fidelity* and smooth, *flicker-free* motion remains unsolved. This motivates new methods.
- **Current Limitations:** Existing methods struggle with **temporal artifacts**, require large **GPU memory**, or sacrifice visual quality for consistency. Addressing these issues is essential for practical applications.

# Prior Approaches

- **Non-Generative Methods:** Use *optical flow* or patch propagation but cause flickering under complex motion.
- **Generator-Based Methods:** Leverage *diffusion* models for consistency, but demand massive computational resources.
- **Unsolved Gaps:** Existing **field-based stylization** lacks temporal stability for videos. New strategies are needed to improve temporal coherence.

# FlowStyler Method



# FlowStyler Method

- **Dual Transformation Fields**: Separates geometric (velocity) and color (orthogonality-regularized) stylization for precise control.
  - Velocity field: Geometric stylization via Lagrangian advection.
  - Color field: Orthogonality-regularized transfer prevents over-stylization and drift.
- **Temporal Consistency**: Field and optimizer state propagation (momentum-preserving) plus occlusion-aware temporal lookup eliminate artifacts and vibration.

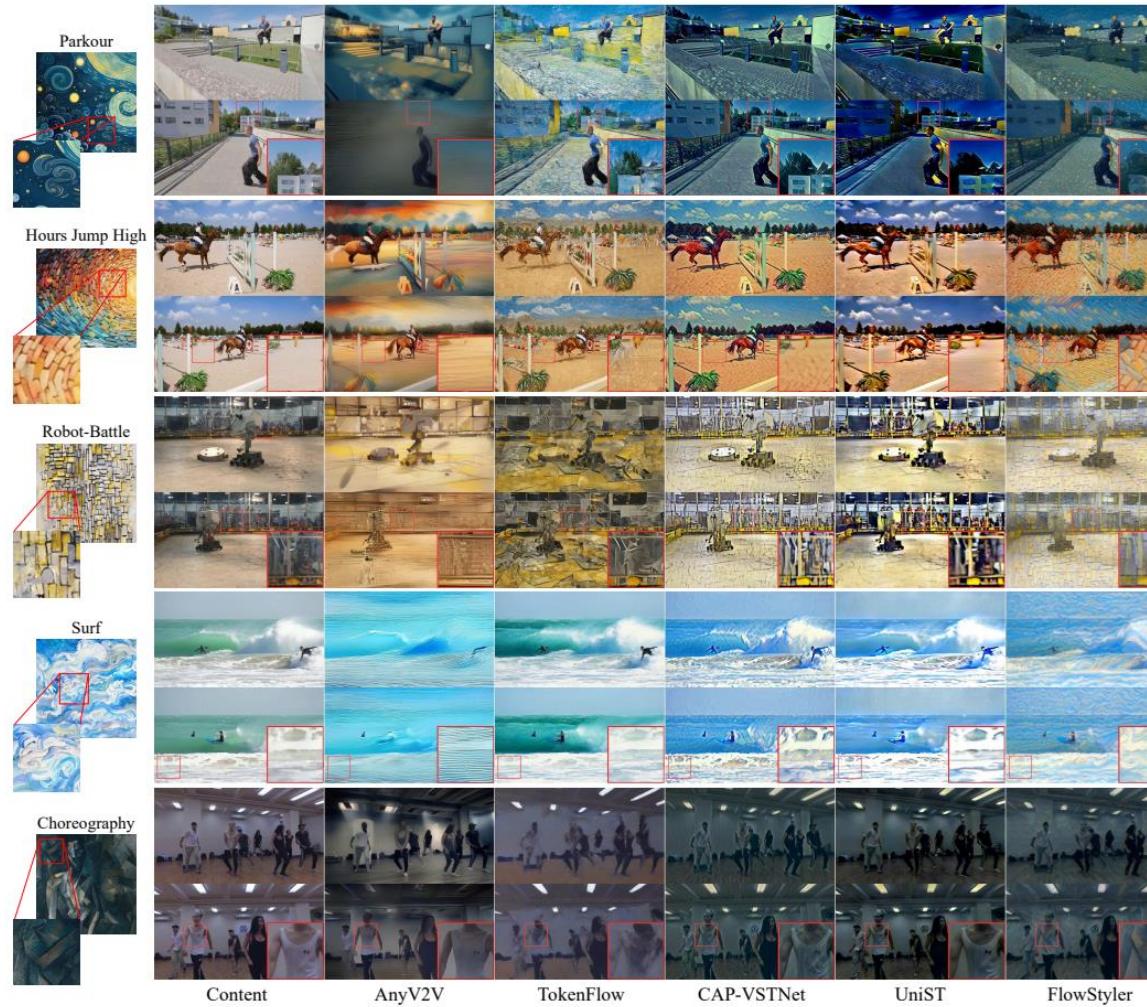
# Technical Highlights

- 1) **Orthogonality loss** ensures stable color mapping.
- 2) **Momentum-preserving optimization** reduces vibration artifacts.
- 3) **Occlusion-aware lookup** fills missing field data for *seamless* stylization.

# Key Results

- **Superior Consistency:** FlowStyler achieves up to  $4\times$  lower *warping* errors than generator approaches.
- **Style Fidelity:** Delivers 22% higher style match than non-generative baselines (DINO/CLIP metrics).
- **Efficiency:** Runs on 3.6GB GPU memory for 100 frames, far below generator methods (>60GB).

# Visual Comparison



# Visual Comparison

- **Qualitative Evaluation:** FlowStyler preserves fine artistic details and strict temporal alignment compared to AnyV2V, TokenFlow, CAP-VSTNet, and *UniST*.
- **Ablation Insights:** Each core component (color field, advection, momentum, occlusion) is critical for state-of-the-art results.