VECTOR SCAFFOLDING: Inter-Scale Orchestration for Differentiable Image Vectorization

Jaerin Lee, Kanggeon Lee, Kyoung Mu Lee
ECE & ASRI & IPAI, Seoul National University

ECCV 2026 | MALMÖ | SEPT 8–12

ComputerVisionLab
Seoul National University

TL;DR

- Problem = To Accelerate Vectorization
- Stabilization = Gradient Correction + Hierarchical Initialization
- Speed-Up = Stabilization + LR Boost + Rescheduling

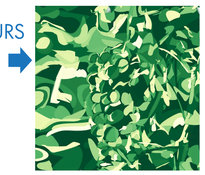
Raster → VECTOR

Motivation

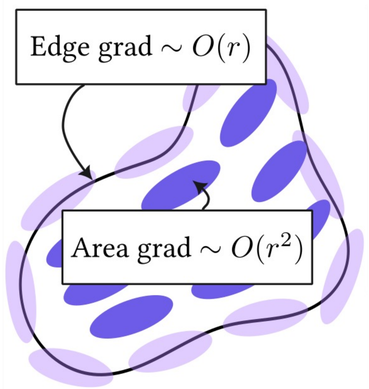
- Topology collapse (2D polygon soup)
- Real-images have scale hierarchies.
- Structure of the representation should match that of optimization alg.

Previous SotA (Bézier splatting)

OURS

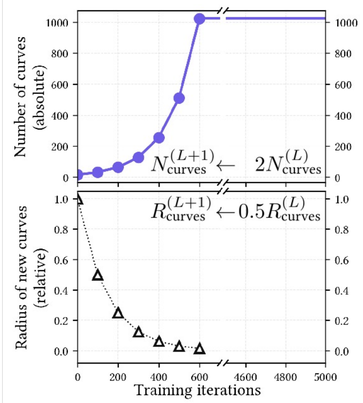


Interior Gradient Aggregation



stable training

Progressive Stratification

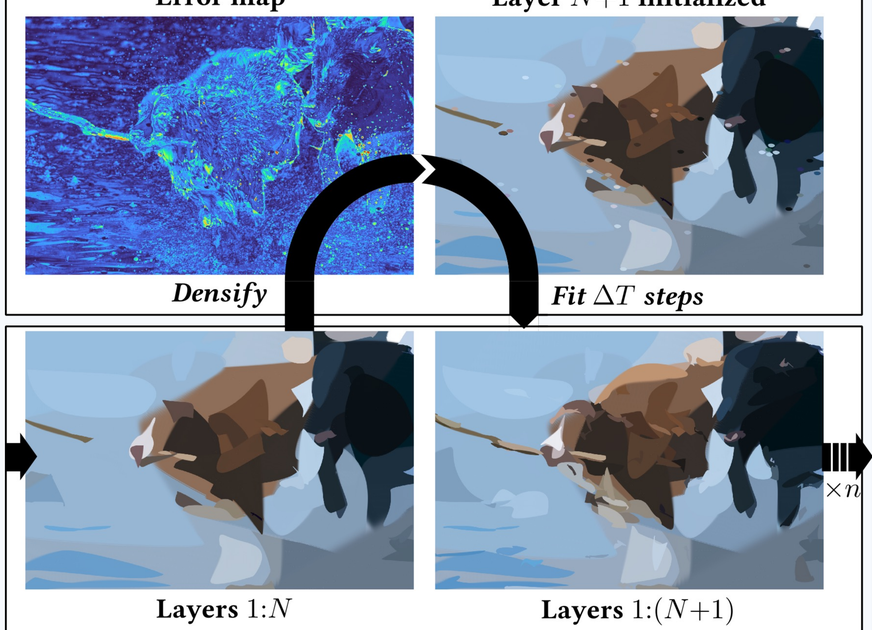


×50 learning rate

Rapid Inflation Scheduling

Progressive Stratification

Each successive layer doubles the number of curves while halving their size.

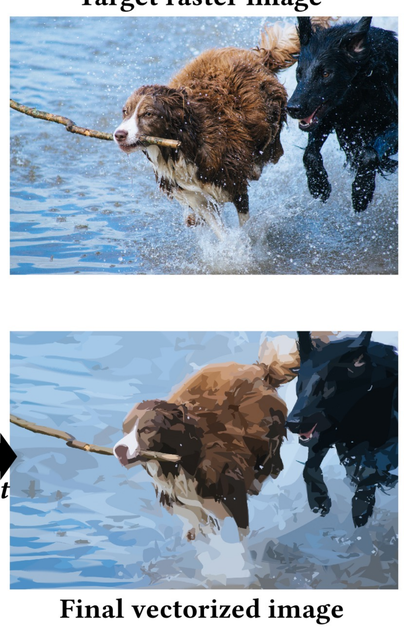


Rapid Inflation Scheduling

Densification happens in burst at the beginning of the training.

Internal Gradient Aggregation

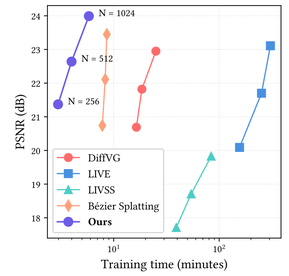
Use gradients from edges *and* interiors.



Vector Scaffolding

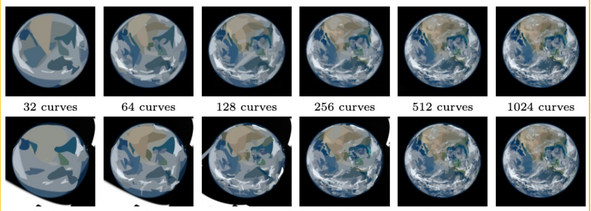
Results

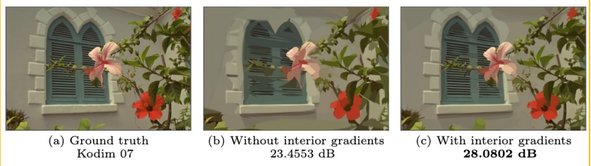
- **+1.4 dB** PSNR
- **x2.5** Speed-Up Wall Clock Time
- **Editability**
- **Level-Of-Detail** Control



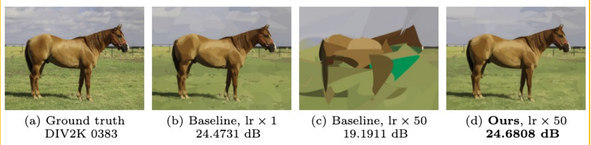
	256 Curves			512 Curves			1024 Curves		
	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
DiffVG [13]	20.69	0.578	0.548	21.82	0.601	0.531	22.95	0.631	0.509
LIVE [17]	20.09	0.576	0.543	21.70	0.611	0.521	23.11	0.648	0.495
LIVSS [23]	17.71	**0.586**	**0.542**	18.71	**0.630**	0.530	19.83	**0.678**	0.517
Bézier Splatting [14]	20.74	0.580	0.546	22.11	0.607	0.528	23.45	0.639	0.507
Ours	**21.37**	0.564	0.548	**22.64**	0.598	**0.503**	**23.99**	0.640	**0.450**

	256 Curves			512 Curves			1024 Curves		
	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
DiffVG [13]	24.11	0.622	0.513	25.34	0.666	0.475	26.66	0.719	0.420
Bézier Splatting [14]	24.19	0.621	0.519	25.61	0.664	0.485	26.91	0.708	0.448
Ours	**25.18**	**0.631**	**0.427**	**26.68**	**0.687**	**0.353**	**28.30**	**0.749**	**0.275**





(a) Ground truth Kodim 07 (b) Without interior gradients 23.4553 dB (c) With interior gradients **28.0802 dB**



(a) Ground truth DIV2K 0383 (b) Baseline, lr × 1 24.4731 dB (c) Baseline, lr × 50 19.1911 dB (d) **Ours**, lr × 50 **24.6808 dB**

Acknowledgement

This work was supported in part by the IITPgrants [No. RS-2021-II211343, Artificial Intelligence Graduate School Program (Seoul National University), No. RS-2025-02303870, No.2022-0-00156] funded by the Korea government (MSIT).