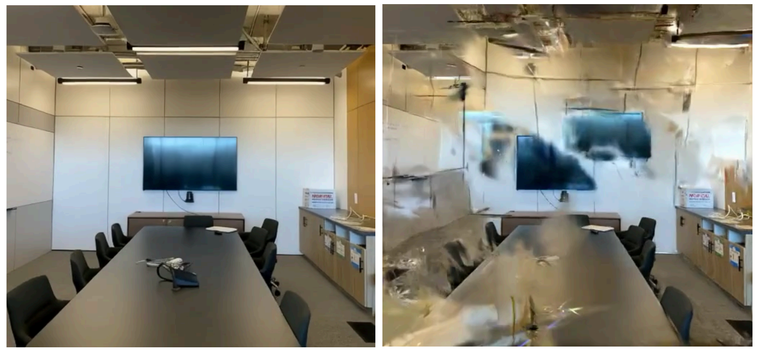
Novel View Synthesis from A Few Glimpses via Test-Time Natural Video Completion

Yan Xu Yixing Wang Stella X. Yu

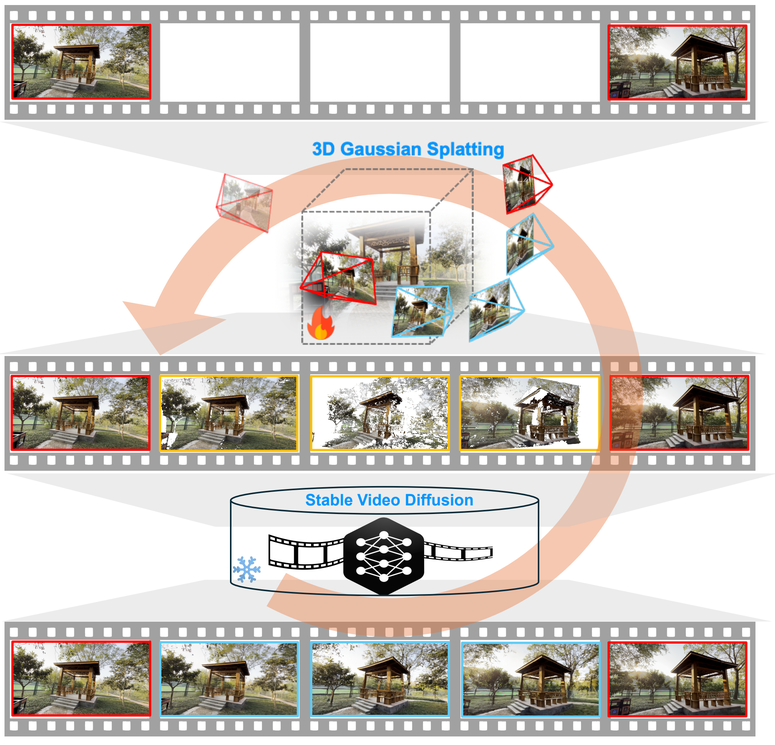
UC Berkeley · COMPUTER SCIENCE & ENGINEERING, UNIVERSITY OF MICHIGAN

NEURAL INFORMATION PROCESSING SYSTEMS

Q: Sparse Input Views Challenge Novel View Synthesis



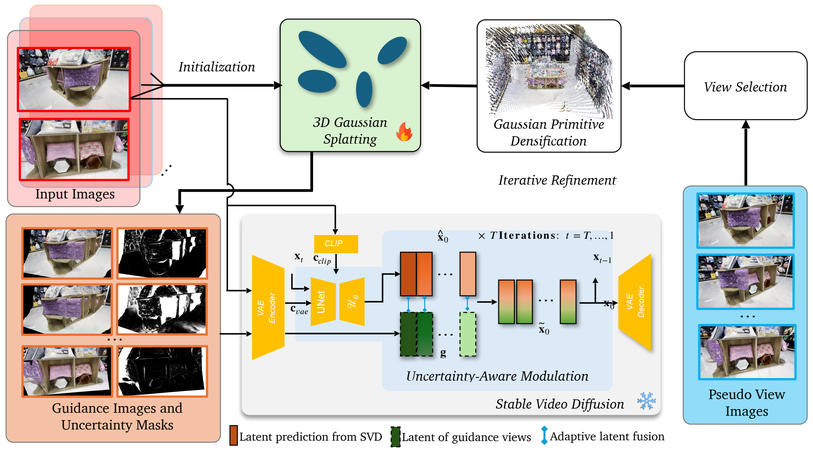
A: Geometric Consistency+Spatiotemporal Visual Priors



First Test-Time Sparse Novel View Synthesis Method

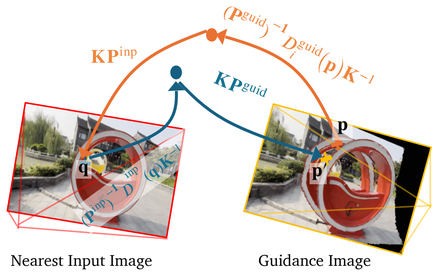
1. Integrate 3D-GS with stable video diffusion with uncertainty-aware modulation for controllable pseudo-view generation
2. Gaussian primitive densification to enhance scene completeness
3. SOTA: 2.5+ dB PSNR gain on DL3DV, strong on LLFF, DTU

Our Method: 3D-GS + Test-Time SVD Video Completion

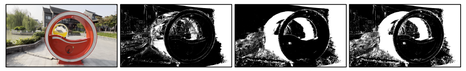


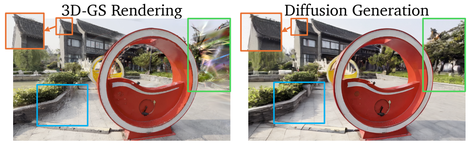
1. Objective function: $\tilde{\mathbf{x}}_0[i] = \arg\min_{\mathbf{x}} \|\mathbf{x} - \hat{\mathbf{x}}_0[i]\|_2^2 + \gamma_{t,i}\|\mathbf{x} - \mathbf{g}[i]\|_2^2$

2. Guidance generation and uncertainty

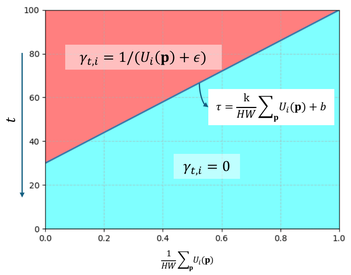


$$U_i(\mathbf{p}) = 1 - \exp\left(-\frac{1}{s_1}\|\mathbf{p} - \mathbf{p}'\|_2^2 - \frac{1}{s_2}\|I_i^{gs}(\mathbf{p}) - I^{inp}(\mathbf{q})\|_2^2\right)$$





3. Uncertainty-aware diffusion



- More consistency, less uncertainty
- Diffusion follows the guide when confident, the generation when not

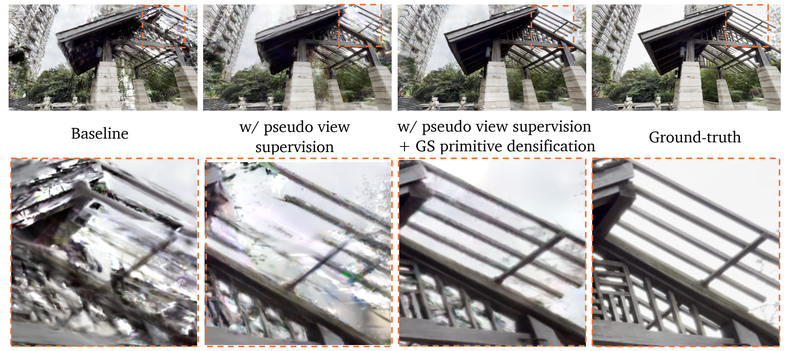
4. 3D-GS Optimization

Input views:
L_s = L1+D-SSIM+Pearson Corr

Generated views:
L_g = LPIPS+D-SSIM+Pearson Corr

	LLFF (3 Views)			*DTU (3 Views)*			*DL3DV (3 Views)*			*DL3DV (6 Views)*			*DL3DV (9 Views)*		
	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Mip-NeRF	16.11	0.401	0.460	8.68	0.571	0.353	10.92	0.191	0.618	11.56	0.199	0.608	12.42	0.218	0.600
3D-GS	17.43	0.522	0.321	10.99	0.585	0.313	10.97	0.248	0.567	12.34	0.332	0.598	12.99	0.403	0.546
DietNeRF	14.94	0.370	0.496	11.85	0.633	0.314	–	–	–	–	–	–	–	–	–
RegNeRF	19.08	0.587	0.336	18.89	0.745	0.190	11.46	0.214	0.600	12.69	0.236	0.579	12.33	0.219	0.598
FreeNeRF	19.63	0.612	0.308	19.92	0.787	0.182	10.91	0.211	0.595	12.13	0.230	0.576	12.85	0.241	0.573
SparseNeRF	19.86	0.624	0.328	19.55	0.769	0.201	–	–	–	–	–	–	–	–	–
SparseGS	–	–	–	18.89	0.834	0.178	–	–	–	–	–	–	–	–	–
FSGS	20.31	0.652	0.288	–	–	–	12.22	0.296	0.535	13.73	0.429	0.540	15.52	0.468	0.416
DNGaussian	19.12	0.591	0.294	18.91	0.790	0.176	11.10	0.273	0.579	12.67	0.329	0.547	13.44	0.365	0.539
IPSM	20.44	0.702	0.207	–	–	–	11.70	0.279	0.534	12.82	0.332	0.521	13.41	0.361	0.529
Ours	20.61	0.705	0.201	20.51	0.840	0.137	14.62	0.471	0.491	17.35	0.566	0.396	19.19	0.616	0.335

SVD Prior + GS Densification Improve Photorealism



Ours Has Better Geometric and Photometric Realism