Parametric Shadow Control for Portrait Generation in Text-to-Image Diffusion Models

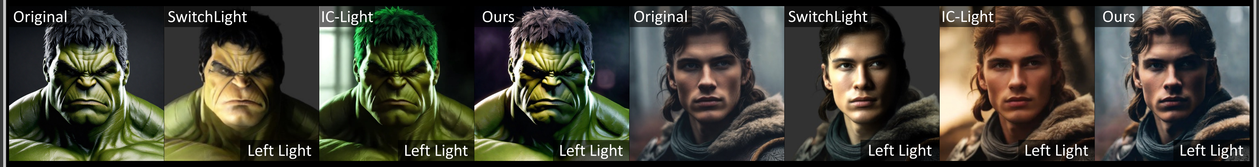
Haoming Cai, Tsung-Wei Huang, Shiv Gehlot, Brandon Y. Feng, Sachin Shah, Guan-Ming Su, Christopher Metzler
University of Maryland, Dolby Labs, MIT

Homepage | Contact

Haoming is looking for 2026 research internships!

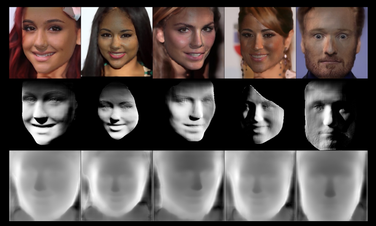
Goal – Enable fine-grained shadow control for Text-2-Image generation

Background – Existing editing methods are data-hungry and generalize poorly to generated images



Baselines suffer from identity shift and inaccurate relighting direction on generated images

Key Ideas – Leverage the rich lighting priors embedded in the pre-trained diffusion models



No light-stage data required
Only synthetic data (RGB, Shadow, Depth)

	SwitchLight	IC-Light	Ours
Type	editing	editing	Cond Gen
Philosophy	neural rendering	scaling	Attribute Mining
No Real-world OLAT	—	—	✓
Data Quantities	30k	10000k+	5.7k
GPU Hours and Type	224 A6000 Hrs	1120 H100 Hrs	16 A6000 Hrs

99.95% less data (vs. IC-Light) and **14× fewer GPU hours** (vs. SwitchLight)

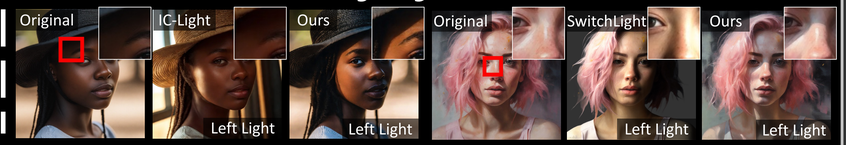
Application & Capability

Progressive Shadow Strength Control

Shadow Shape Control

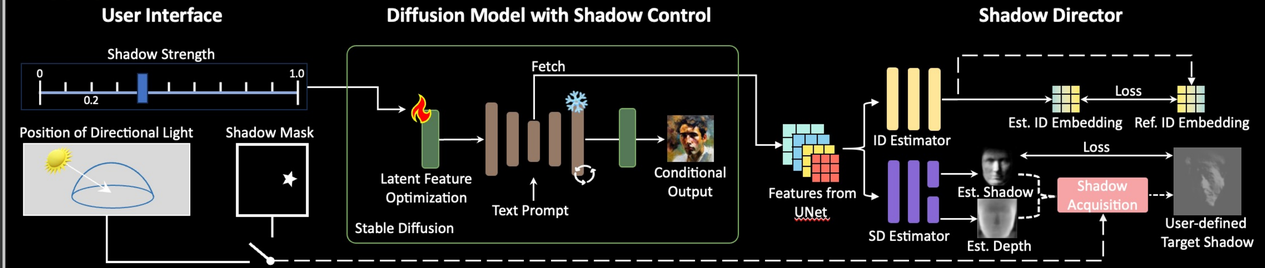
Relighting

Object Relighting w/o Finetune



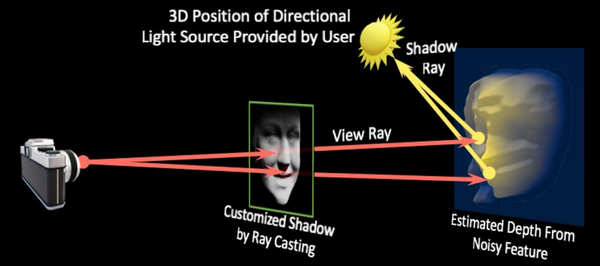
Obstacle-aware

Method

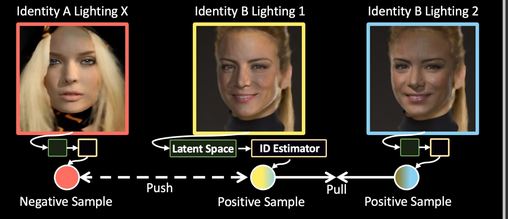


Training: Train a module to estimate shadow, depth, and identity from noisy latent features
Inference: Iteratively optimize latent features until the desired shadow emerges
User Inputs : Text Prompt & Shadow Strength & Directional Light Location (X,Y,Z) or Shadow Binary Mask

Method Details



Given the user-specified light location (XYZ), the **desired shadow map** is computed via **ray casting** on the estimated depth.



Identity Constraint Loss