

Motivation

- Single-view (2D) Reflection separation is highly under-determined.
- Traditional approaches work by simply capturing paired flash and no-flash images: their difference will be a reflection-free image.
- Such pairs must be strictly aligned. Even tiny vibrations like pressing the shutter button produce artifacts, even when using tripod.
- Our approach performs robust 2D reflection separation *without* paired flash/no-flash measurements.



Paired capture with tripod. Wireless shutter control.



Paired capture with tripod. Shutter control pressed by finger.

Comparing Different Paradigms

Comparing to software-only methods and flash-based ones, we present a best-of-both-worlds approach that gradually removes reflections against misaligned flash/no-flash pairs in the latent space with diffusion.



Flash-Split: 2D Reflection Removal with Flash Cues and Latent Diffusion Separation

Tianfu Wang*, Mingyang Xie*, Haoming Cai, Sachin Shah, Christopher A. Metzler

*Equal Contribution.

Our Contributions



- Stage 1: the core idea here is to perform flash-informed reflection separation iteratively in a low-dimensional latent space.
- The dual-branch latent diffusion U-Net with cross-attention takes in a latent pair of flash/no-flash images and gradually separates the composite scene into reflection and transmission latent images.





Stage 2: to avoid hallucinations in the Stage 1 latent diffusion process.

reconstructed image, we use the input image as a guide to condition the decoding process of the separated latents predicted from our





Cross-Latent Decoder Beats Vanilla Decoder





