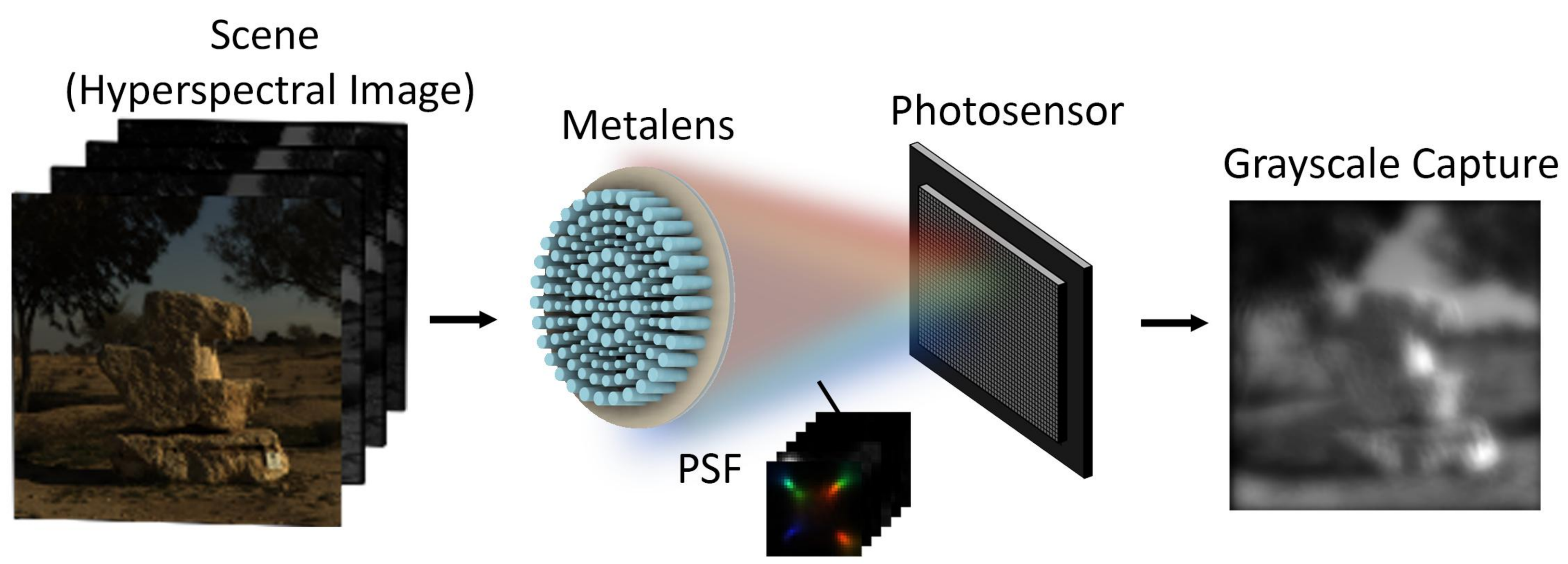
Filterless Snapshot Hyperspectral Imaging Using Guided Patch Diffusion

ECCV 2026 | MALMÖ | SEPT 8–12

Dean Hazineh, Luca Sacchi, Davide Cassara, Federico Capasso, Todd Zickler
Harvard University

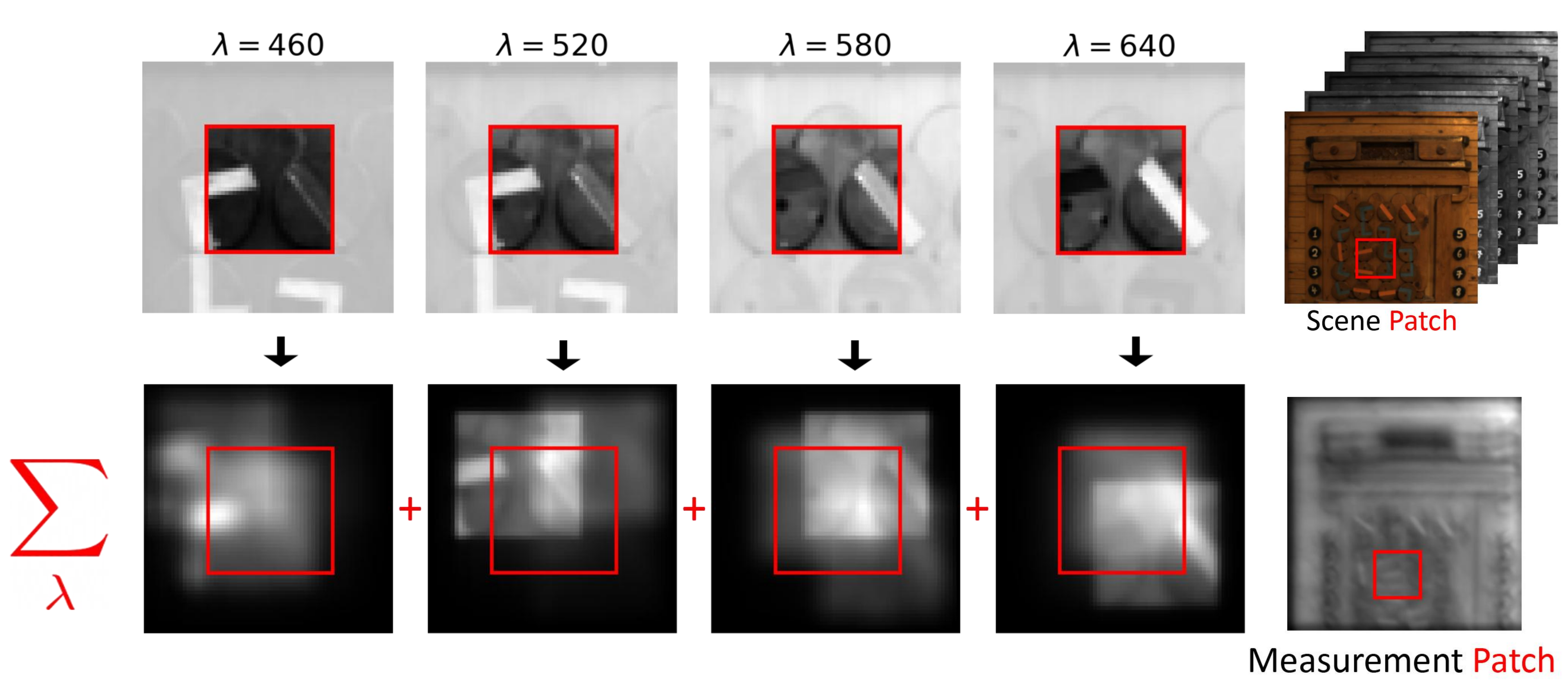
Overview:

We reconstruct a 31-channel Hyperspectral Image (HSI, H x W x 31) using only the chromatic aberration in a snapshot grayscale measurement (H x W x 1). We remove extra optical components, RGB filter arrays, and temporal scanning.



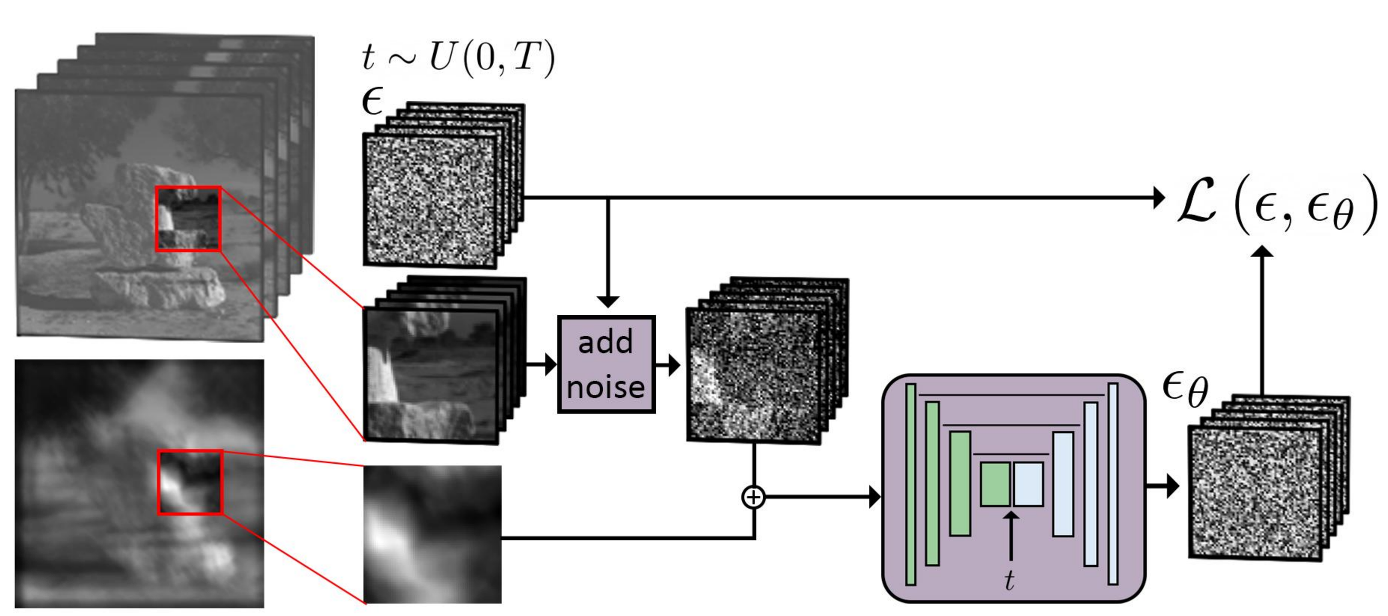
Challenges:

1. **Data Scarcity:** Natural-scene hyperspectral image datasets are too small to train large reconstruction models (<1K images).
2. **Patch-Based Processing Fails for Large Spatial Blur:** Training models to process small measurement patches instead of the full image is more ambiguous. The optical cue is spread over an area larger than the patch.



Approach (1/2): Diffusion on HSI Patches

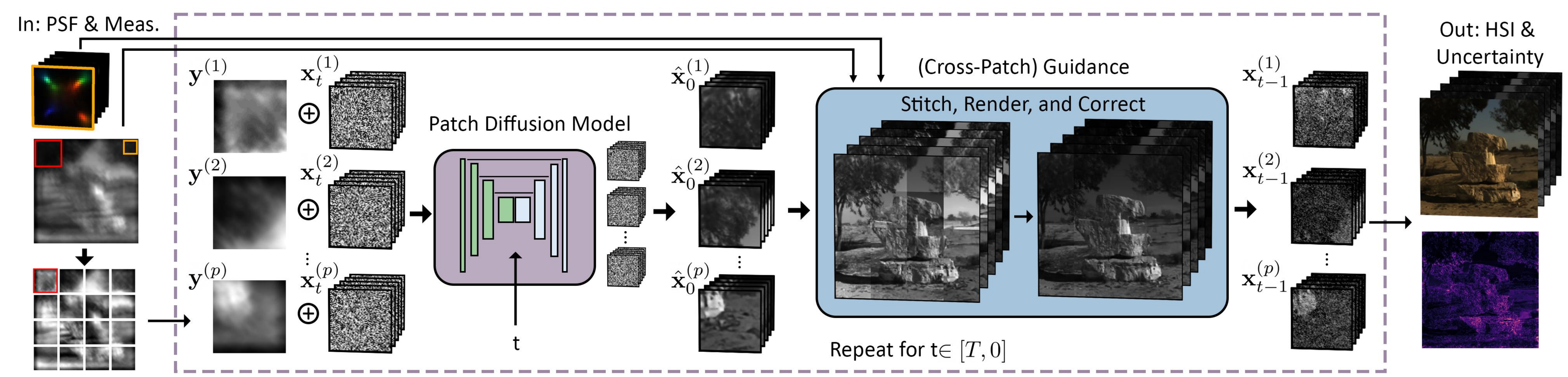
We simulate the chromatic aberration of our diffractive lens and render the measurement for a small dataset of hyperspectral images. We train a denoising diffusion model to generate hyperspectral patches from measurement patches.



Approach (2/2): Synchronize Patch Predictions Using Inference-time Guidance

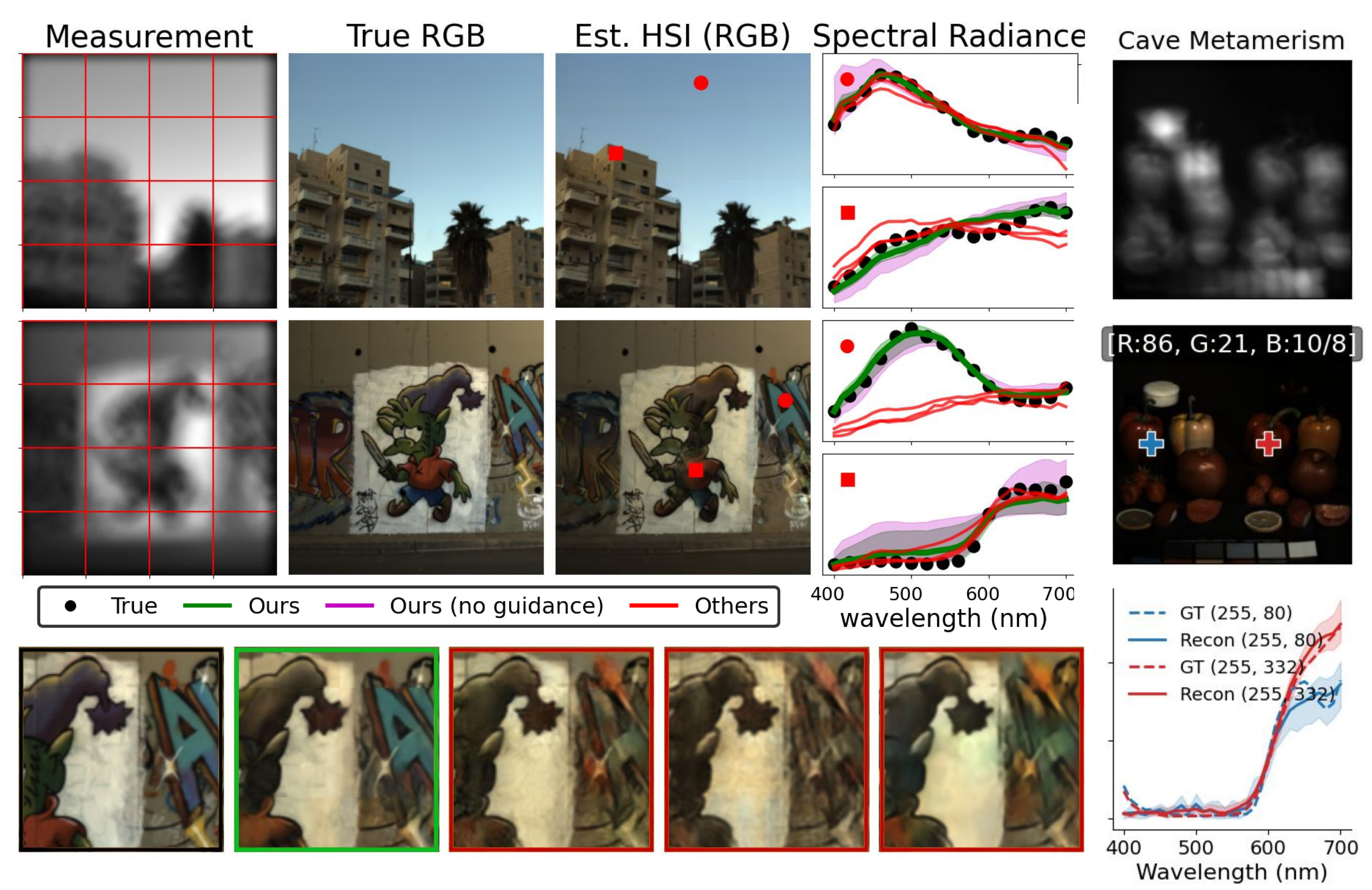
Hyperspectral patch predictions from the trained diffusion model (purple box) are made independently so neighboring patches don't agree. To fix this, we:

1. **Stitch** together the intermediate denoised patch predictions $\hat{x}_0$ into one full hyperspectral image,
2. **Render** the corresponding grayscale measurement using the known optical PSF,
3. **Correct** the denoising trajectory using the gradient between the rendered and true measurements, nudging all patches toward agreement.

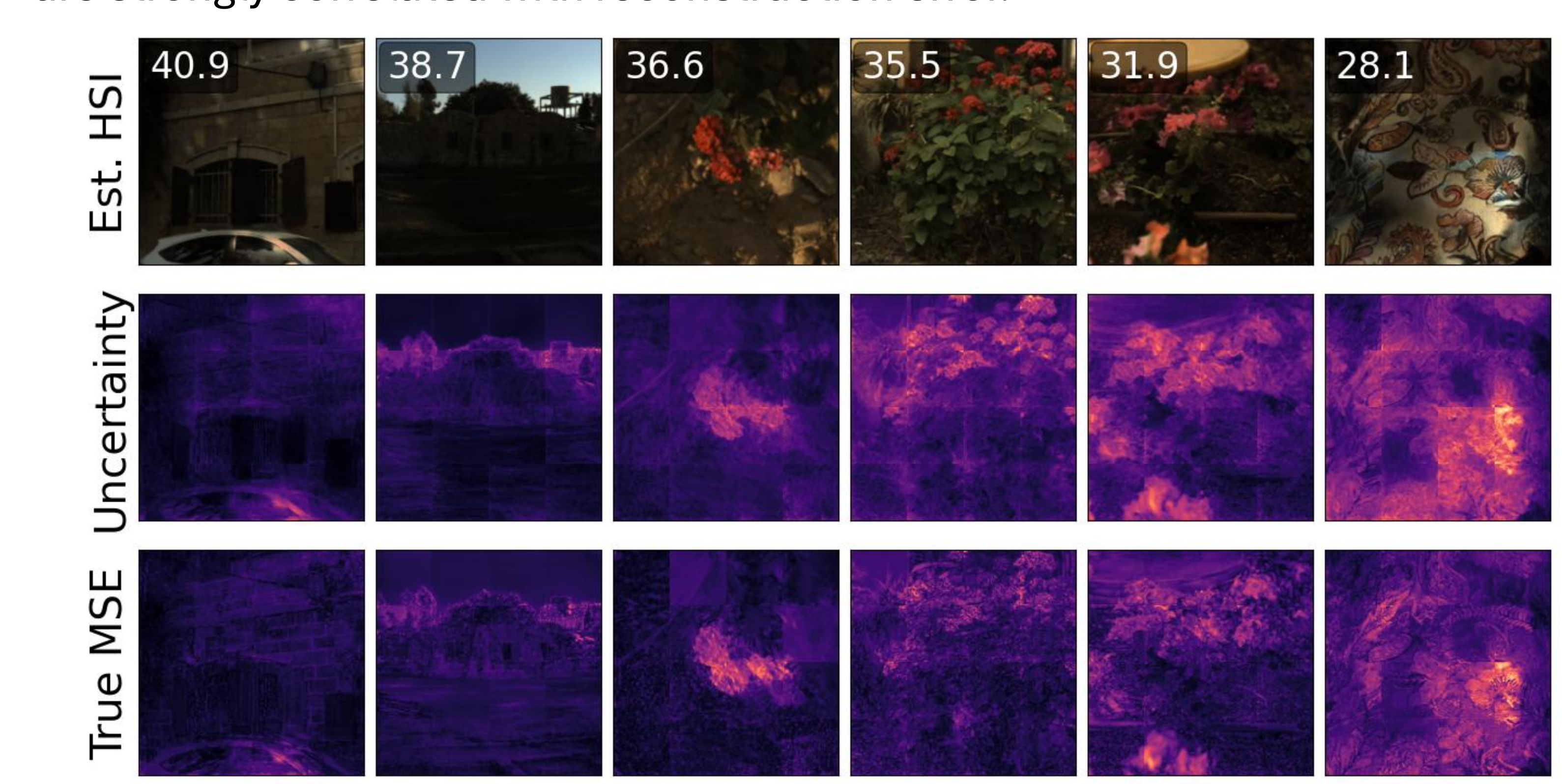


Results:

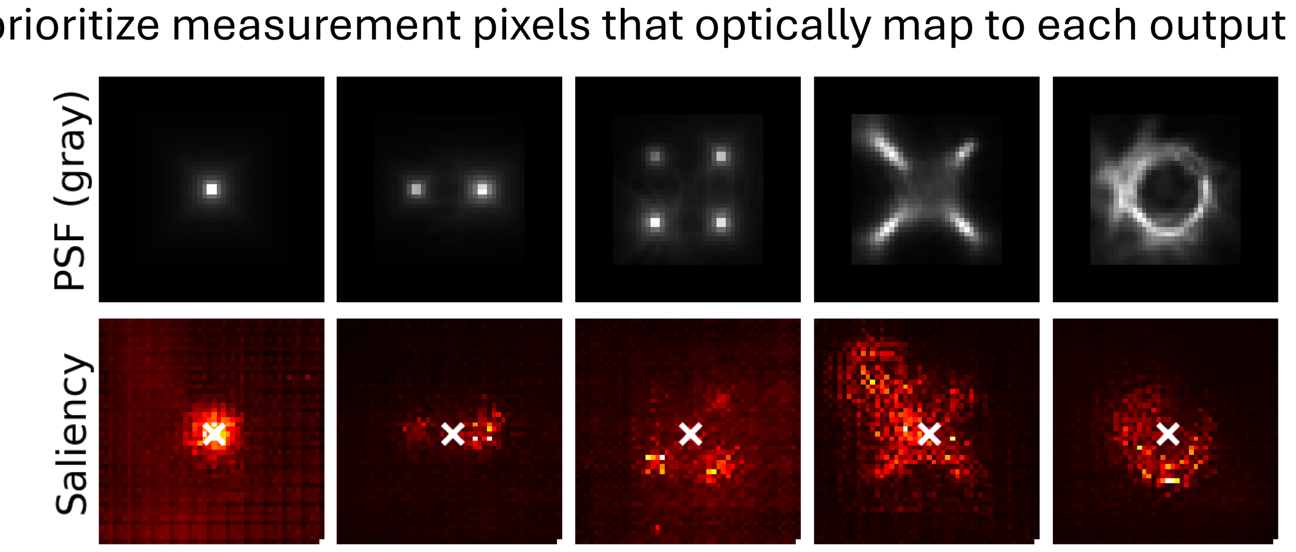
State-of-the-Art Reconstructions: The method generalizes across measurements rendered from other HSI datasets and metameric scenes.



Spatial-Spectral Uncertainty Estimates: Computing the variance over repeated reconstructions using different noise seeds produces uncertainty estimates that are strongly correlated with reconstruction error.



Model Interpretability: We train separate models for lenses with different PSFs, then compute perturbation-based saliency maps. Each model learned to prioritize measurement pixels that optically map to each output location.



Experimental Validation: We build a prototype camera and reconstruct real tabletop scenes. Guidance significantly helps to overcome calibration errors.