Specularity Factorization for Low Light Enhancement

CVPR 2024

Saurabh Saini*
P J Narayanan

IIIT-Hyderabad, India

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

IIIT Hyderabad

CVPR
JUNE 17-21, 2024
SEATTLE, WA

**Now @Amazon (work done prior @IIIT-H)*

Overview

- Image Factorization:
  - Specularity based
  - Interpretable
  - Simple & low-cost
- Core Idea:
  - Recursively extract specularity
  - Identify illuminated regions
  - Enhance individually

Image Factorization

- WHAT :

  Disentangle the image into two or more meaningful factors .

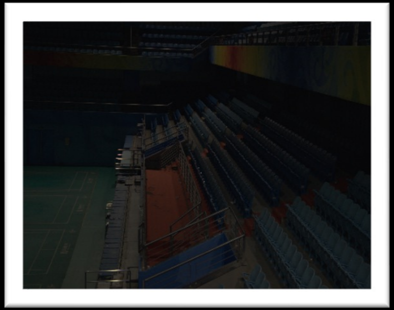
  Combined multiplicatively or additively.

- WHY :
  - Common and Crucial Step
  - Interpretable Editing

- WHEN :
  - Single Image
  - Multiple Images

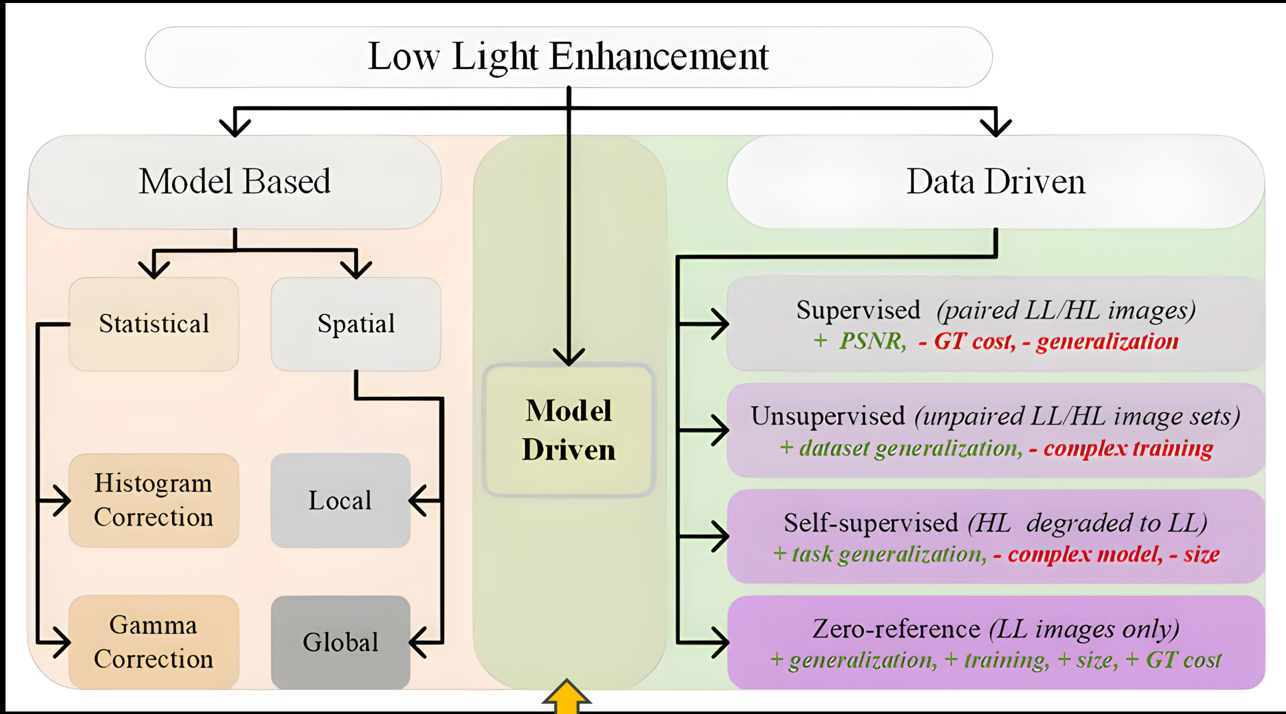
Criteria	Type	#Factors
Retinex	I = R.S	2
Details	I = High + Low pass	2
Spectral	I = F (phase*amp.)	2
Rank	I = L + e	2
Wavelets	I = HL + LL + LH + HH	2^n
Spatial	I = Pymd(I)	var.
Intensity	$I = I^1 + I^2 + .. + I^k$	var.
Specularity	$I = E^1 + E^2 + .. + E^k$	var.

Low Light Enhancement



Factorization → Fusion →

- Factorization + Fusion
- Model Driven
- Zero-reference
  (no paired/unpaired dataset)



Specularity Factorization

Dichromatic reflection model: X = A + E

Specular

Simple L1 minimization

$$\underset{\mathbf{E},\mathbf{A}}{\operatorname{argmin}} \|\mathbf{A}\|_* + \lambda\|\mathbf{E}\|_1 \quad \text{s.t.} \quad \mathbf{X} = \mathbf{A} + \mathbf{E},$$

Sparse

ADMM update (T iterations)

$$\mathbf{E}_{t+1} = \delta^1_{\alpha_t}\left(\mathbf{X} - \mathbf{A}_t - \mathbf{Y}^T_t/\mu_t\right)$$
$$\mathbf{A}_{t+1} = \delta^*_{\beta_t}\left(\mathbf{X} - \mathbf{E}_{t+1} - \mathbf{Y}^T_t/\mu_t\right)$$
$$\mathbf{Y}_{t+1} = \mathbf{Y}_t + \mu_t\left(\mathbf{A}_{t+1} + \mathbf{E}_{t+1} - \mathbf{X}\right)$$

Soft thresholding

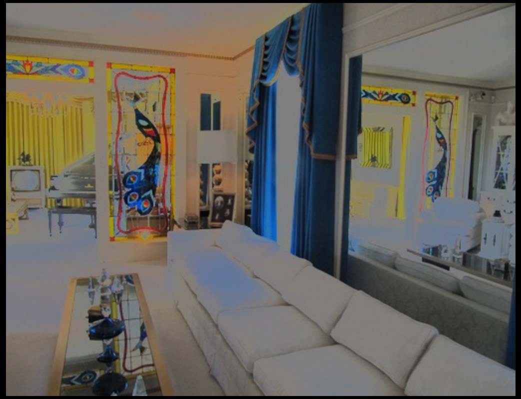
α,β functions of λ

μ step size

X

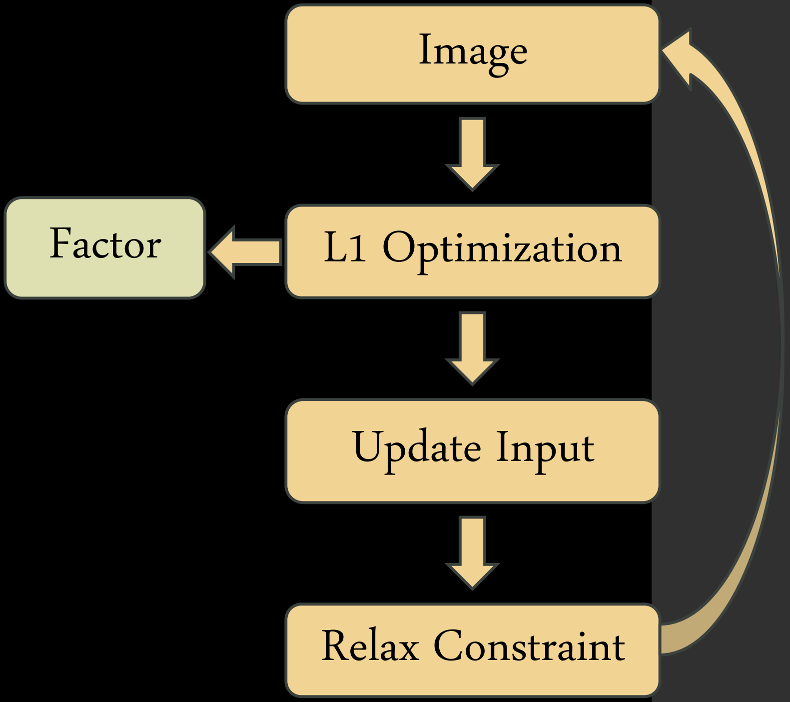
E

A



- Convex Optimization. Stephen Boyd and Lieven Vandenberghe. Cambridge University Press, 2004.
- Distributed optimization and statistical learning via the alternating direction method of multipliers. Boyd et al. Foundations and Trends in ML, 2011.
- Proximal algorithms. Neal Parikh and Stephen Boyd. Foundations and Trends in Optimization, 2014.

Recursive Specularity Factorization Model



Single L1 minimization

$$\operatorname*{argmin}_{\mathbf{E},\mathbf{A}} \|\mathbf{A}\|_* + \lambda\|\mathbf{E}\|_1 \quad \text{s.t.} \;\; \mathbf{X} = \mathbf{A} + \mathbf{E},$$

Update Input & relax sparsity constraint

$$\mathbf{X}^{k+1} = \mathbf{X}^k - \mathbf{E}^k$$

$$\lambda^{k+1} < \lambda^k$$

Recursive optimization

Repeat K times

$$\mathbf{I} = \mathbf{E}^1 + \mathbf{E}^2 + \ldots + \mathbf{E}^K = \sum_{k=1}^{K} \mathbf{E}^k$$

Core Idea:

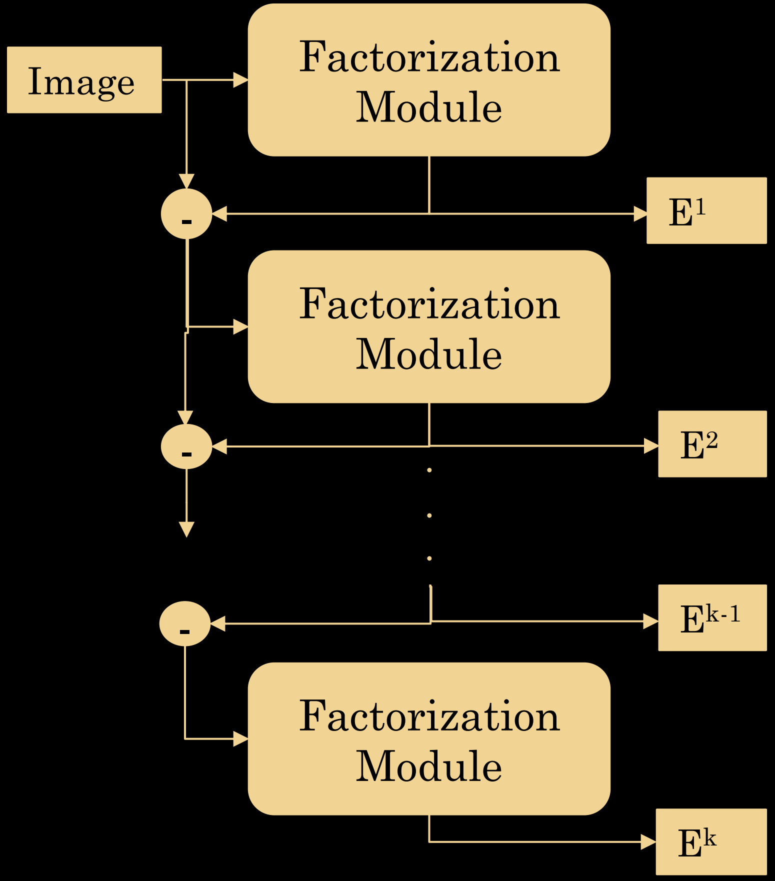
Decompose an image into K factors by recursively optimizing with gradually relaxed sparsity.

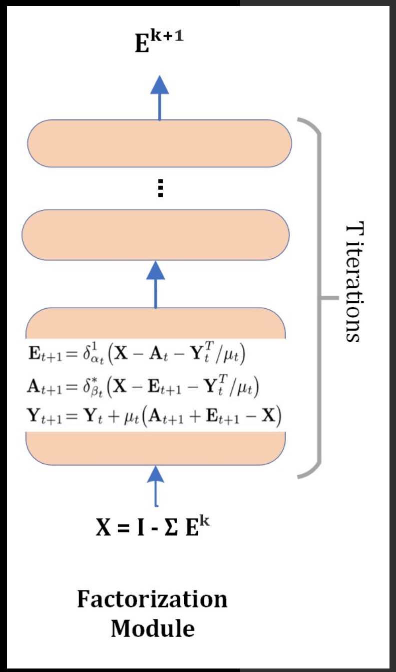
Unrolling (RSFNet)

- Recursive Specularity Factorization Network
- Learn 3 params per iteration
  - α β μ
- For K factors (T iterations each):
  - 3 K T parameters only !
- Factorization Loss

Total intensity ratio

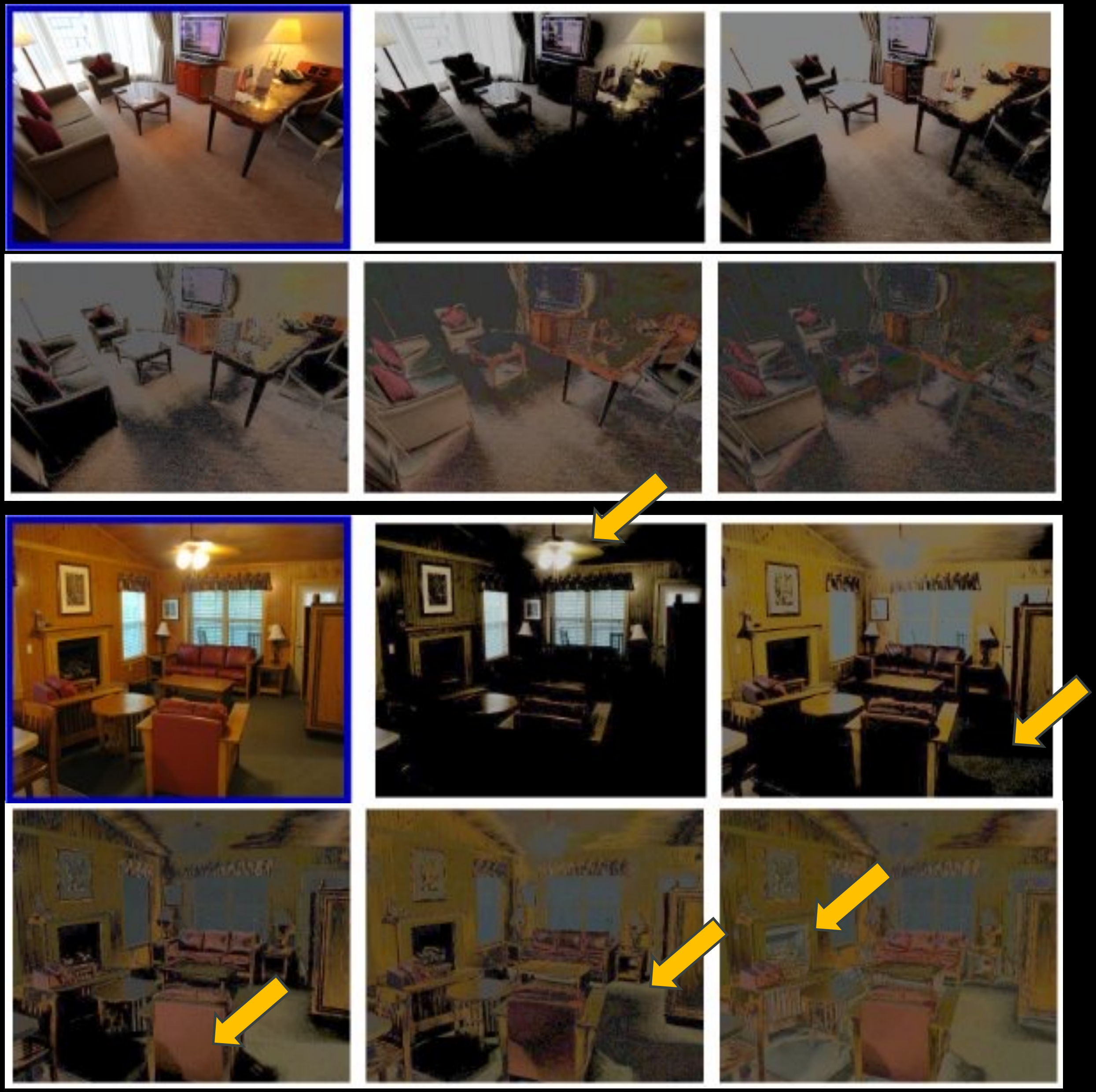
$$L_f^k = \left| \frac{E^k}{X^k} - \frac{k}{K} \right| \qquad L_f = \Sigma_{k=1}^{K} L_f^k$$



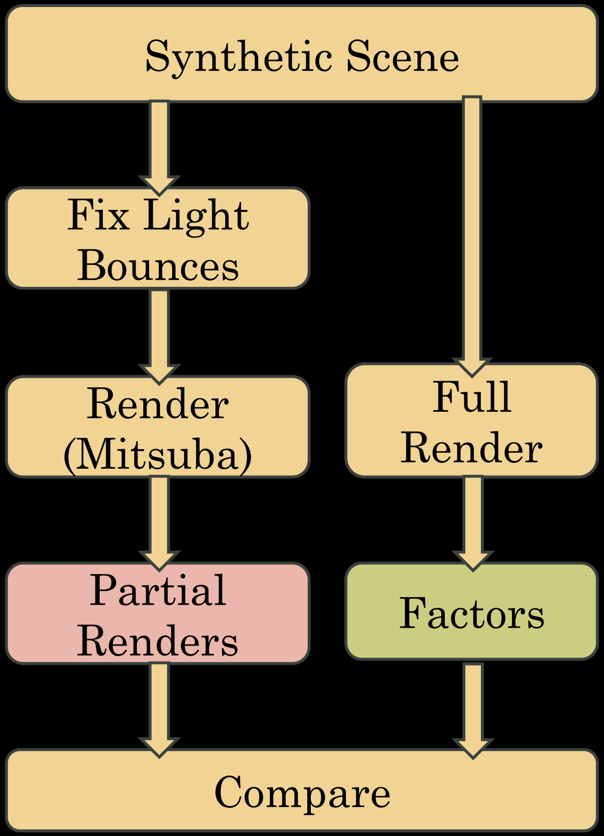


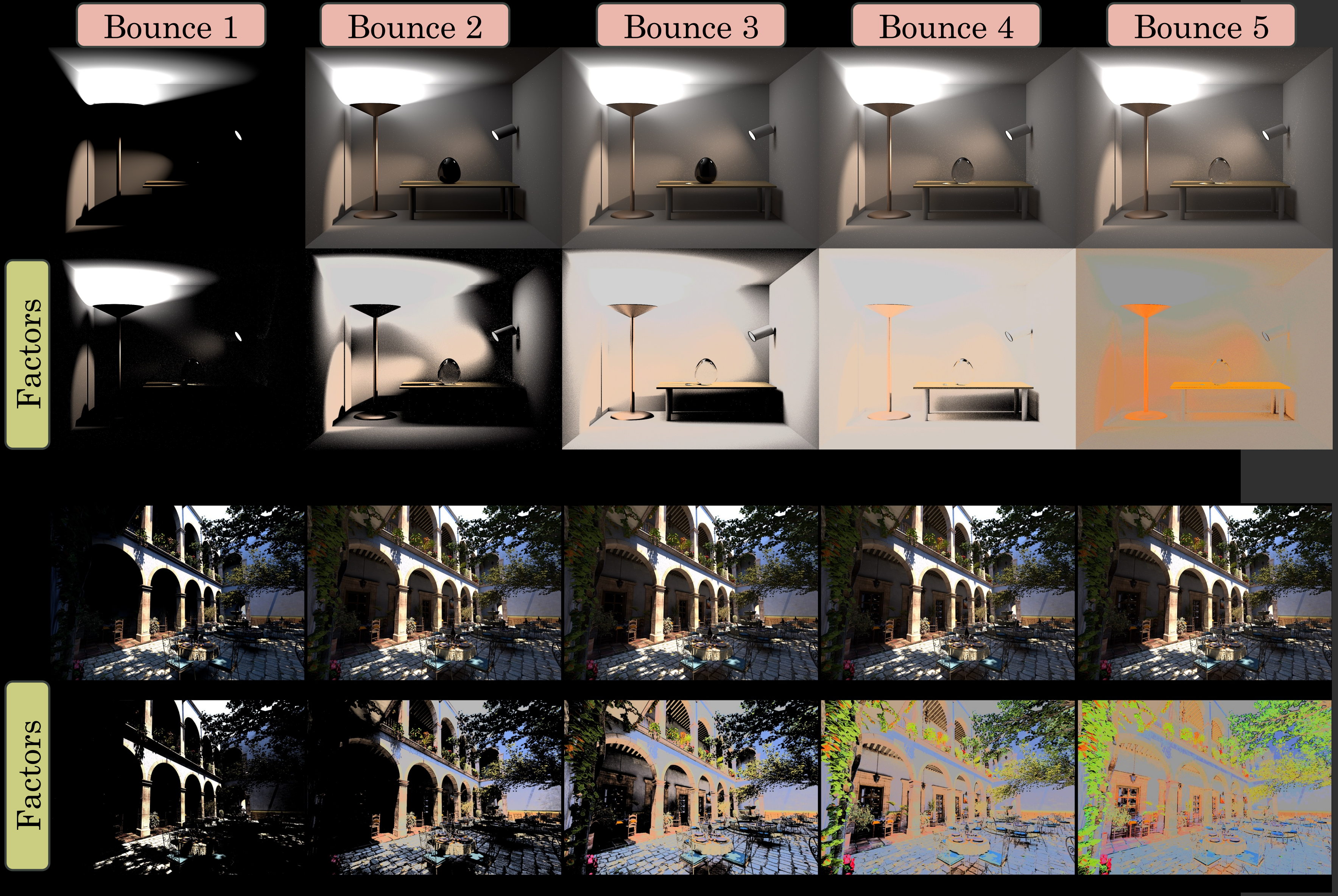
- Learned robust pca: A scalable deep unfolding approach for high-dimensional outlier detection. Cai et al. NeurIPS 2021.

Factors Visualization



Hypothesis Validation (synthetic scenes)



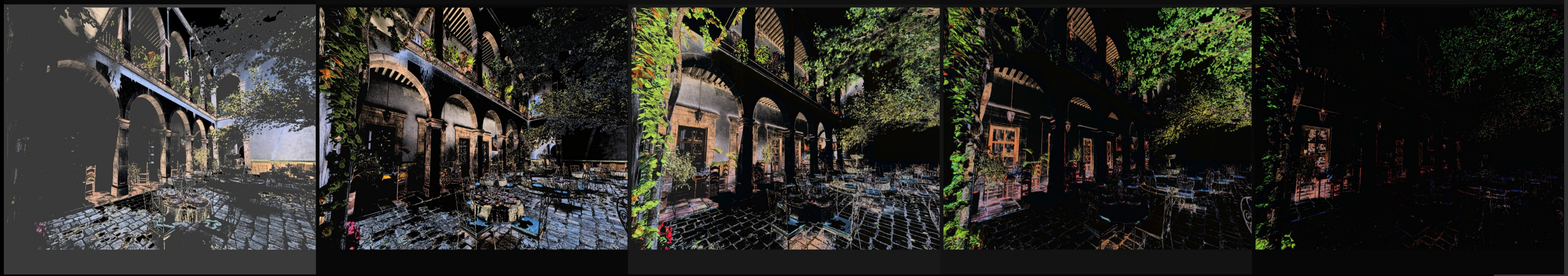


Hypothesis Validation (synthetic scenes)

Image

Factors

Differences



Hypothesis Validation (real scenes)

Shadow Dataset

↓ ↓

Shadow / bright regions | Factors

↓ ↓

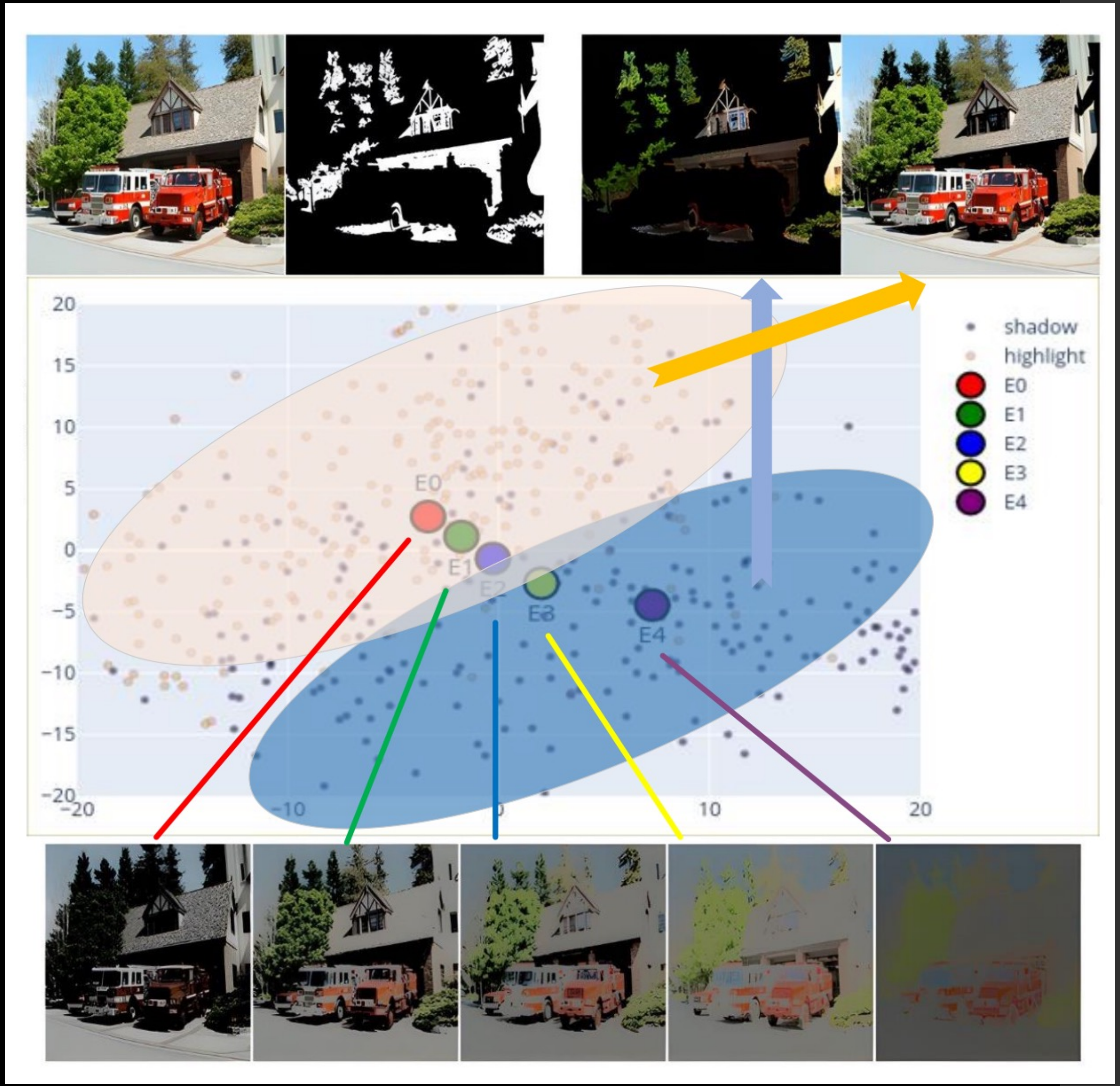
Image features (DINOv2)

↓

Dimensionality Reduction (PCA)

↓

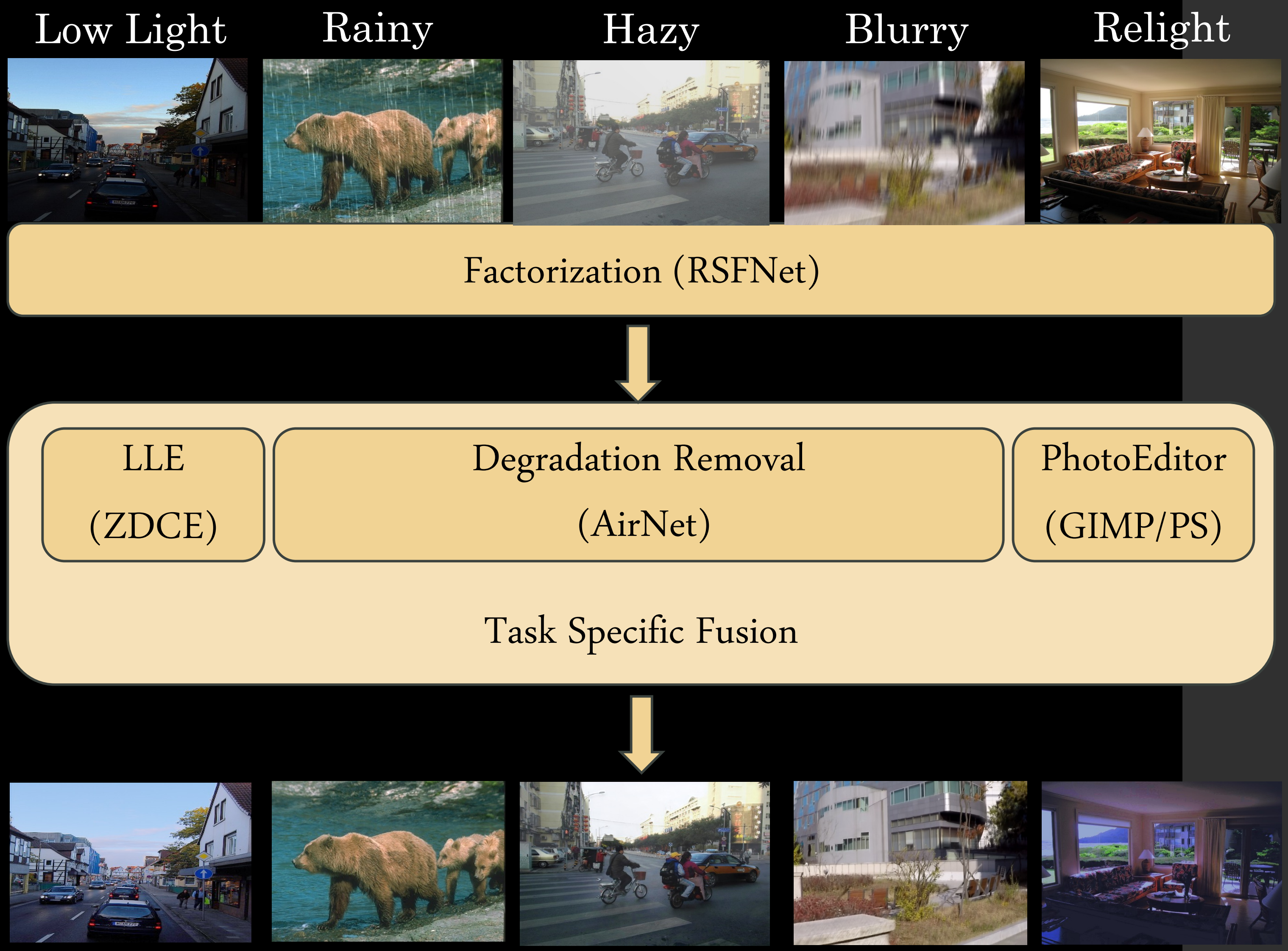
Plot



- Revisiting shadow detection: A new benchmark dataset for complex world. Hu, Xiawei et al. *IEEE TIP 2021*.
- Emerging properties in self-supervised vision transformers. Caron, Mathilde et al. *ICCV 2021*.

Fusion

- Off-the-shelf <u>task based fusion</u>
  - Zero-reference LLE network (ZDCE - UNet)
  - Supervised multi-task enhancement (AirNet)
  - Direct fusion (w/o learning)
    - Weighted Average (by intensity)
    - Manual blending
    - Exposure Fusion



- **ZDCE:** Zero-reference deep curve estimation for low-light image enhancement. CVPR, 2020. *Guo et al. CVPR 2020.*
- **AirNet:** All-In-One Image Restoration for Unknown Corruption. *Li et al. CVPR 2022.*
- **Exposure Fusion:** A simple and practical alternative to high dynamic range photography. *Mertens et al. Computer Graphics Forum 2009.*

Experimentation

- Datasets: (13)
  - Full-reference (4)
    - Lol-v1
    - Lol-v2 (real)
    - Lol-v2 (synthetic)
    - VE-Lol
  - No-reference (5)
    - DICM
    - MEF
    - LIME
    - VV
    - NPE
  - Application Datasets (4)
    - Rain100L
    - GoPro
    - RESIDE
    - IIW

- Metrics: (7)
  - Full-reference
    - PSNRc (rgb)
    - SSIMc (rgb)
    - PSNRy (Y of YCbCr)
    - SSIMy (Y of YCbCr)
    - LPIPS
  - No-reference
    - NIQE
    - LOE

- Comparisons: (20+)
  - LLE
    - Traditional (3)
    - Zero-reference (7)
    - Unsupervised (5)
    - Supervised (4)
  - Applications (3)
    - Deraining
    - Dehazing
    - Deblurring

LLE Results

Input GT Reconstruction Input GT Reconstruction

- Lolv1 : Deep retinex decomposition for low-light enhancement. Wei et al. *BMVC 2018*.
- Lolv2 : Sparse gradient regularized deep retinex network for robust low-light image enhancement. Yang et al. *IEEE TIP 2021*.

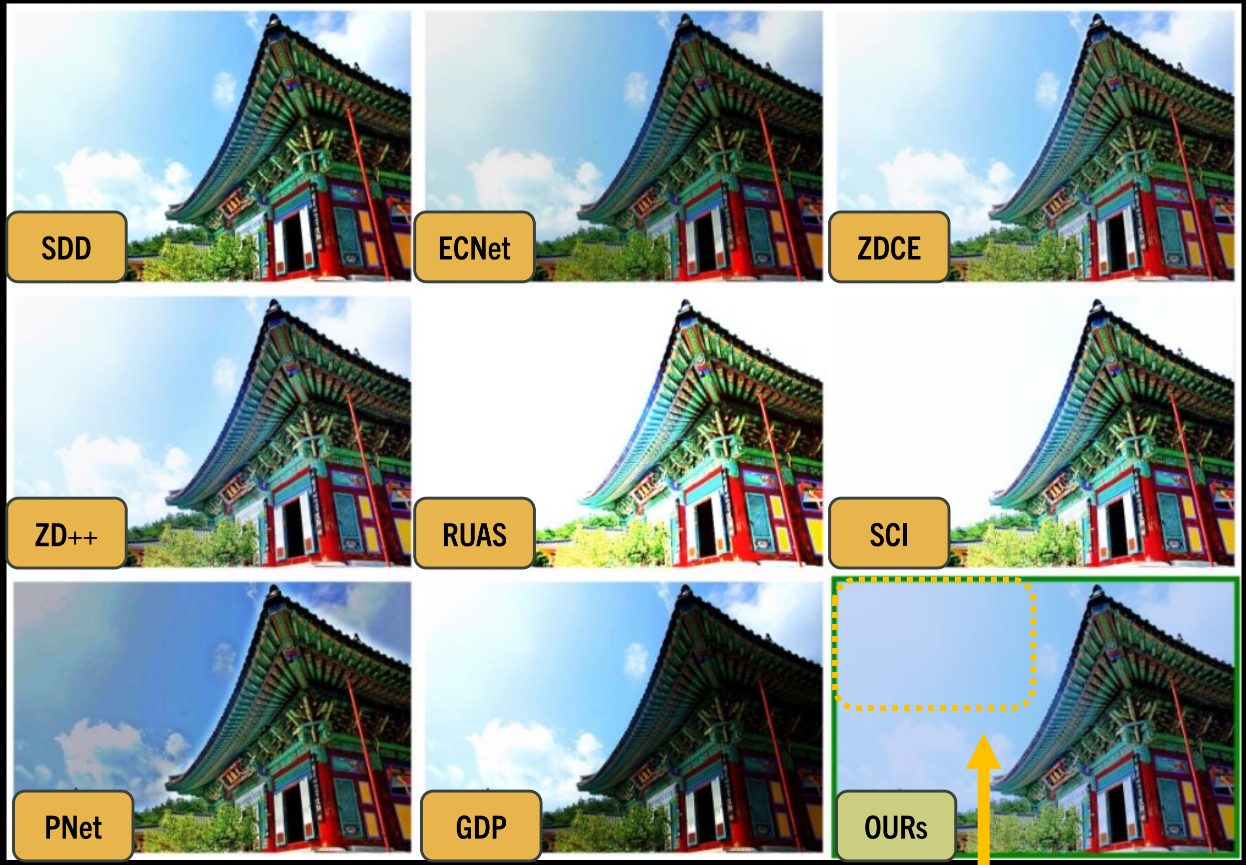
Quantitative Comparison

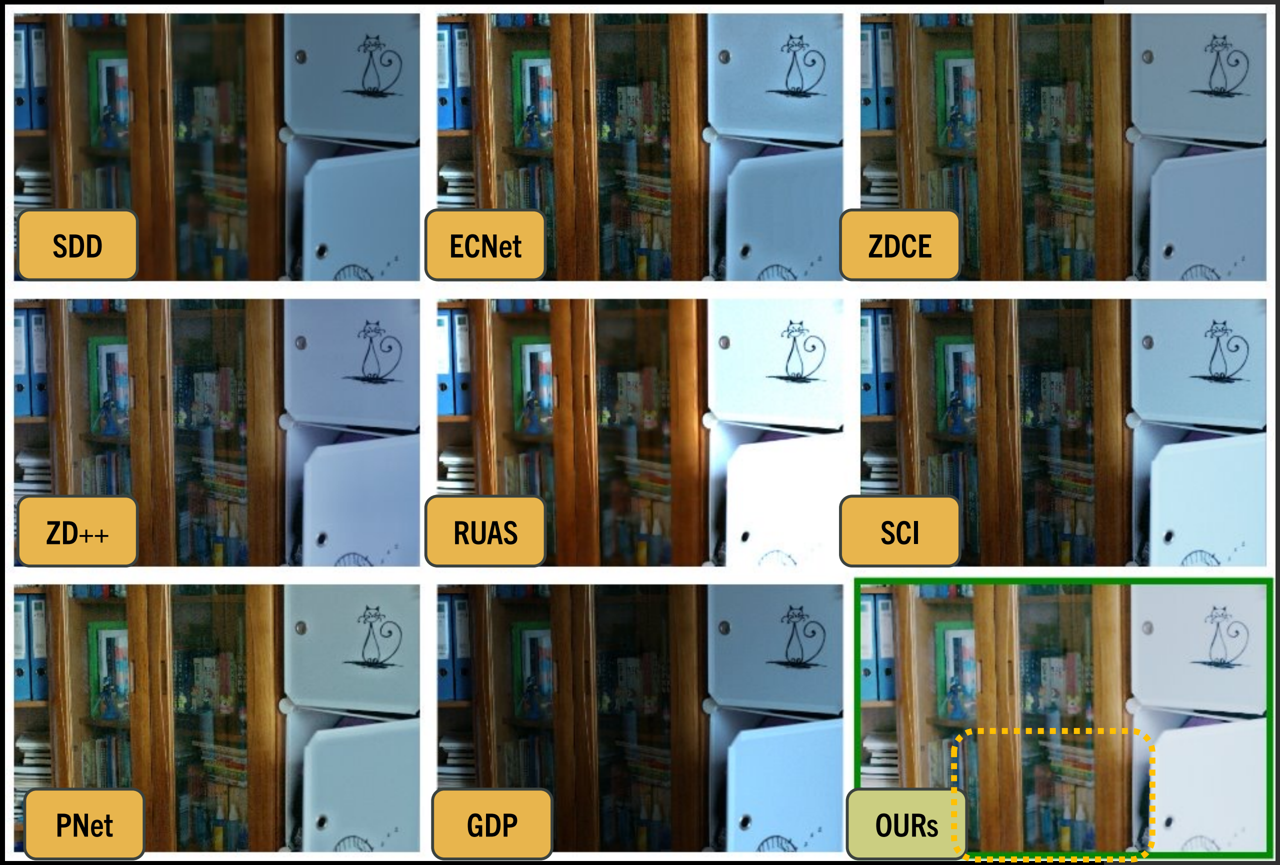
Mean scores on LoLv1, LoLv2-real, LoLv2-syn and VE-LoL datasets

	Traditional LLE			Zero-Reference LLE							
	LIME	DUAL	SDD	ECNet	ZDCE	ZD++	RUAS	SCI	PNet	GDP	RSFNet
#Param (*10^3)	-	-	-	16.5k	79.42	10.56	3.43	**0.26**	15.25	552k	2.11
$PSNR_y$ ↑	18.50	17.83	17.50	18.45	19.26	18.73	17.09	18.07	**19.65**	15.88	**21.16**
$SSIM_y$ ↑	0.737	0.728	**0.781**	0.677	0.777	0.674	0.743	0.745	0.743	0.634	**0.854**
$PSNR_c$ ↑	16.53	15.88	15.77	16.25	17.19	16.76	15.12	16.20	**17.35**	14.15	**18.45**
$SSIM_c$ ↑	0.596	0.583	**0.679**	0.538	0.634	0.548	0.532	0.587	0.605	0.504	**0.758**
NIQE ↓	7.855	7.478	**4.077**	7.543	4.270	7.468	5.841	7.626	7.791	6.726	**3.763**
LPIPS ↓	0.291	0.297	**0.266**	0.329	**0.273**	0.296	0.346	0.295	0.302	0.379	0.276

- **LIME**: Low- light image enhancement via illumination map estimation. Guo et al. *IEEE TIP 2016.*
- **DUAL**: Dual illumination estimation for robust exposure correction. Zhang et al. *CFG 2019.*
- **SDD**: Low-light image enhancement with semi- decoupled decomposition. Hao et al. *IEEE TMM 2020.*
- **ECNet**: Zero-shot restoration of back-lit images using deep internal learning. Zhang et al. ACM MM 2019.
- **ZDCE/ZD++**: Zero-reference deep curve estimation for low-light image enhancement. CVPR 2020.
- **RUAS**: Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement. Risheng et al. CVPR 2021.
- **SCI**: Toward fast, flexible, and robust lowlight image enhancement. Ma et al, CVPR 2022.
- **PNet**: Progressive self-enhancement network for unsupervised extreme-light image enhancement. Nguyen et al. WACV 2023.
- **GDP**: Generative diffusion prior for unified image restoration and enhancement. Fei et al, CVPR 2023

Qualitative Comparisons





- Preserves Natural-ness
- Over/under exposed regions
- No artifacts

Generalization (0-Reference)

- 5 no-reference datasets (DICM, LIME, MEF, NPE, VV)
- 2 no-reference metrics (NIQE/LOE)

Generalization (5 no-reference benchmarks)

NIQE↓ & *LOE*↓	ECNet [89]	ZDCE [24]	ZD++ [42]	RUAS [66]	PNet [57]	SCI [51]	**RSFNet** (Ours)
DICM [37]	3.37—*676.7*	3.10—*340.8*	**2.94**—*511.9*	4.89—*1421*	3.00—*590.3*	3.61—*321.9*	3.23—***303.1***
LIME [26]	**3.75**—*685.1*	3.79—*135.0*	3.89—*332.2*	4.26—*719.9*	3.84—*223.2*	4.14—*75.5*	3.80—***68.3***
MEF [50]	3.30—*863.3*	3.31—*164.3*	3.18—*458.5*	4.08—*784.2*	3.25—*363.0*	3.43—***95.0***	**3.00**—*100.7*
NPE [76]	**3.24**—*936.1*	3.52—*312.9*	3.27—*532.2*	5.75—*1399*	3.29—*601.1*	3.89—*239.8*	3.31—***221.5***
VV [73]	2.15—*292.4*	2.75—*145.4*	2.53—*222.9*	3.82—*583.7*	2.56—*260.2*	2.30—***109.0***	**1.96**—***109.0***
Mean	3.16—*690.7*	3.29—*219.7*	3.16—*411.5*	4.56—*981.7*	3.19—*407.5*	3.47—*168.2*	3.06—*160.5*

- **NIQE:** Making a "Completely Blind" Image Quality Analyzer, Mittal et al. IEEE Signal Processing Letters *2013*.
- **LOE:** Naturalness preserved enhancement algorithm for non-uniform illumination images. Wang et al. *IEEE TIP 2013*.

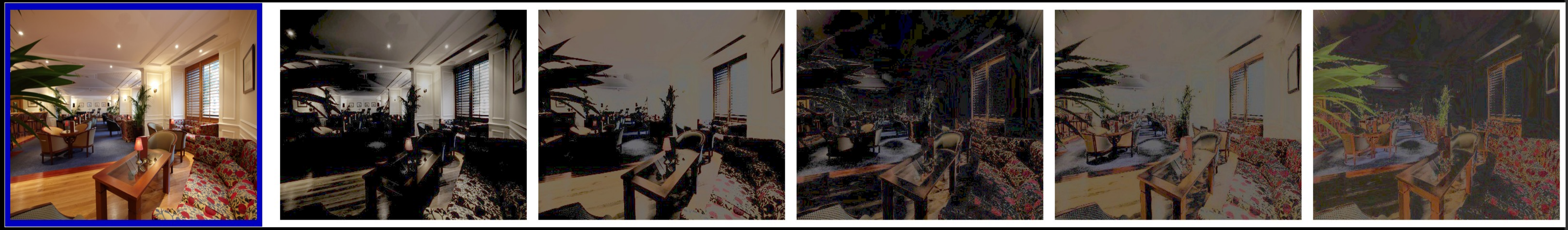
Generalization (Supervised/Unsupervised LLE)

- 5 No-reference Datasets (DICM, LIME, MEF, NPE, VV)
- 2 Supervised LLE (SNR, RFormer)
- 2 Unsupervised LLE (HEP, NeRCo)

	Supervised LLE		Unsupervised LLE		0-ref
NIQE ↓	SNR	RFormer	HEP	NeRCo	RSFNet
DICM	3.62	**3.08**	4.06	3.55	3.23
LIME	3.75	3.91	3.98	**3.42**	3.80
MEF	3.92	3.14	3.65	3.15	**3.00**
NPE	3.54	3.63	**2.99**	3.24	3.31
VV	2.89	2.18	3.60	3.17	**1.96**
Mean	3.54	3.19	3.66	3.31	**3.06**

- **SNR:** Snr-aware low-light image enhancement. *Xu et al. CVPR 2022.*
- **RFormer:** Retinexformer: One-stage retinex-based transformer for low-light image enhancement. *Cai et al. ICCV 2023.*
- **HEP:** Unsupervised low-light image enhancement via histogram equalization. *Zhang et al. arXiv:2112.01766 2021.*
- **NeRCo:** Implicit neural representation for cooperative low-light image enhancement. *Yang et al. ICCV 2023*

Prior Induction. (beyond LLE)



Image+Factors → **Existing Enhancement Nets** → **Improvement**

Concatenate input | Edit input channels | Finetune

- As Structural Priors

- Other Enhancement Extensions
  1. DeHazing
  2. DeRaining
  3. DeBlurring

- Modify Existing Enhancement Models
  1. **AirNet**: All-In-One Image Restoration for Unknown Corruption. CVPR 2022.
  2. **DePIO**: A general de-coupled learning framework for parameterized image operators. ECCV 2018.
  3. **TWeather**: Transformer-based restoration of images degraded by adverse weather conditions. CVPR 2022.
  4. **TAPE**: Task-agnostic prior embedding for image restoration. ECCV 2022.

Image DeHazing

Factors

Results

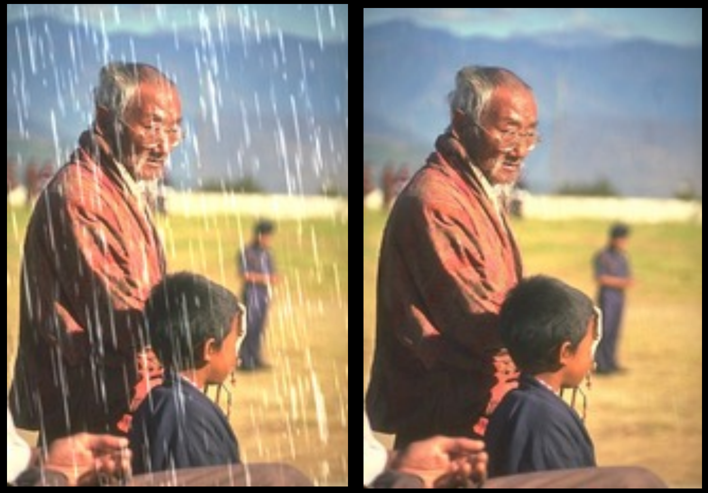
DeHazing (RESIDE-SOTS)		
Method	PSNR	SSIM
DePIO	20.54	0.826
TWeather	21.32	0.885
Tape	22.16	0.861
AirNet	23.18	0.900
AirNet + (Ours)	24.96	0.929
% Improvement	**+7.68**	**+3.22**

RESIDE Dataset: Benchmarking single-image dehazing and beyond. IEEE TIP 2019.

Image DeRaining

Factors

Results



DeRaining (Rain100L)		
Method	PSNR	SSIM
DePIO	21.96	0.762
TWeather	29.43	0.905
Tape	29.67	0.904
AirNet	34.90	0.966
AirNet + (Ours)	36.19	0.972
% Improvement	**+ 3.70**	**+ 0.60**

Rain100L Dataset: Deep joint rain detection and removal from a single image. CVPR 2017.

Image DeBlurring

Factors

Results

DeBlurring (GoPro dataset)		
Method	PSNR	SSIM
DePIO	19.86	0.672
TWeather	25.12	0.757
Tape	24.47	0.763
AirNet	26.42	0.801
AirNet + (Ours)	27.29	0.827
% Improvement	**+3.29**	**+3.25**

GoPro Dataset: Deep multi-scale convolutional neural network for dynamic scene deblurring. CVPR 2017.

User Controlled Relighting

Extract Factors

↓

Import Layers (GIMP, Photoshop)

↓

Global Operators (no local selection)

↓

Controllable Editing

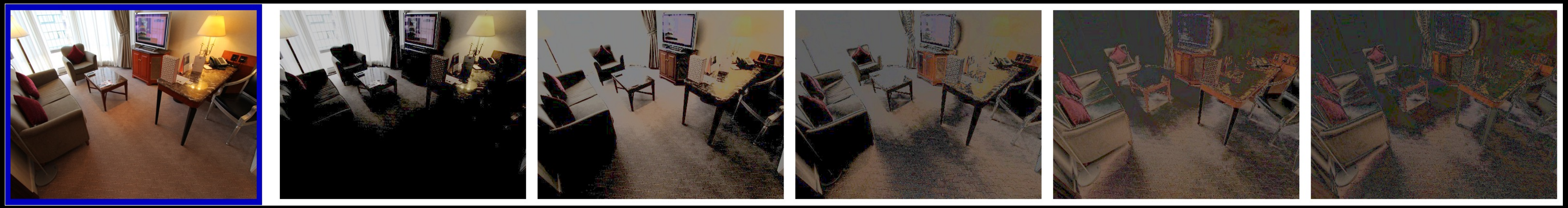
Original

Edited

Low Ambient

Indoor Light Color

Outdoor Light Intensity



Summary

- Novel image factorization criterion.
- New recursive specularity extraction model.
- Model-driven unrolled framework.

- Zero-reference LLE solution
- SOTA generalization performance.

- Induction as a structural prior
- User-controlled relighting.

Future Extensions

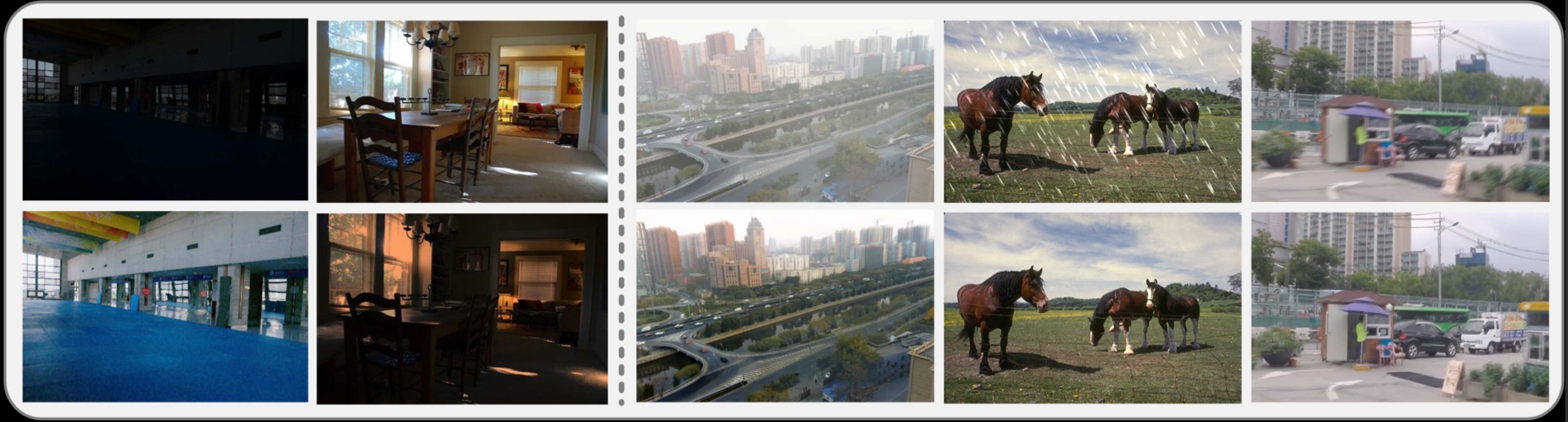
- Data simulation (~50k images)

- Image Adaptive Factorization ?
- Spatially varying parameterization ?

- Other Signals (hyperspectral, lidar) ?
- Applications :
  - Denoise
  - Reflection removal
  - Shadow Removal
  - Object compositing
  - Fg/Bg matting
  - Retexturing ….

Thank You

Poster 226
(Arch-4E)



https://sophont01.github.io/data/projects/RSFNet/