Noise & Synthesis Dataset Survey

Presenter: Hao Wang

Advisor: Prof. Chia-Wen Lin

Outline

- Only dataset
  - PolyU
  - LLRVD
- Dataset and Synthesis
  - LRD
  - LRID
- Discussion

Real-world Noisy Image Denoising: A New Benchmark

Jun Xu[1], Hui Li[1], Zhetong Liang[1], David Zhang[1,2], *Fellow, IEEE*, and Lei Zhang[1,*], *Fellow, IEEE*

[1]Department of Computing, The Hong Kong Polytechnic University, Hong Kong SAR, China

[2]School of Science and Engineering, The Chinese University of Hong Kong (Shenzhen), Shenzhen, China

arXiv 2018

Presenter: Hao Wang

Advisor: Prof. Chia-Wen Lin

Introduction

- Construct a large dataset of real-world noisy images with reasonably obtained corresponding “ground truth” images.


(a) 200,3.5,1/60


(b) 200,6.7,1/60


(c) 200,6.7,1/8


(d) 400,6.7,1/60


(e) 1600,6.7,1/60

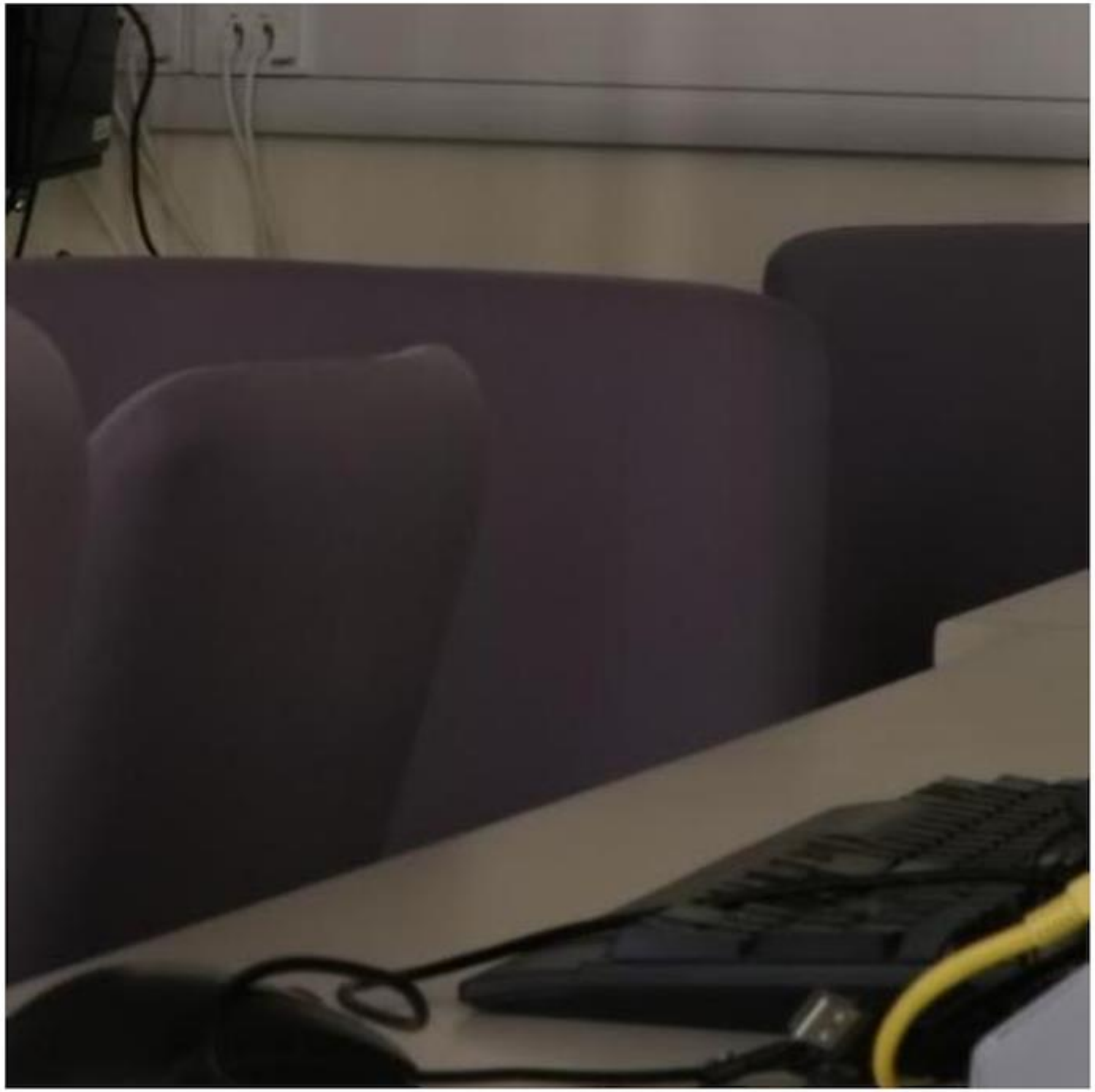

(f) 3200,6.7,1/60

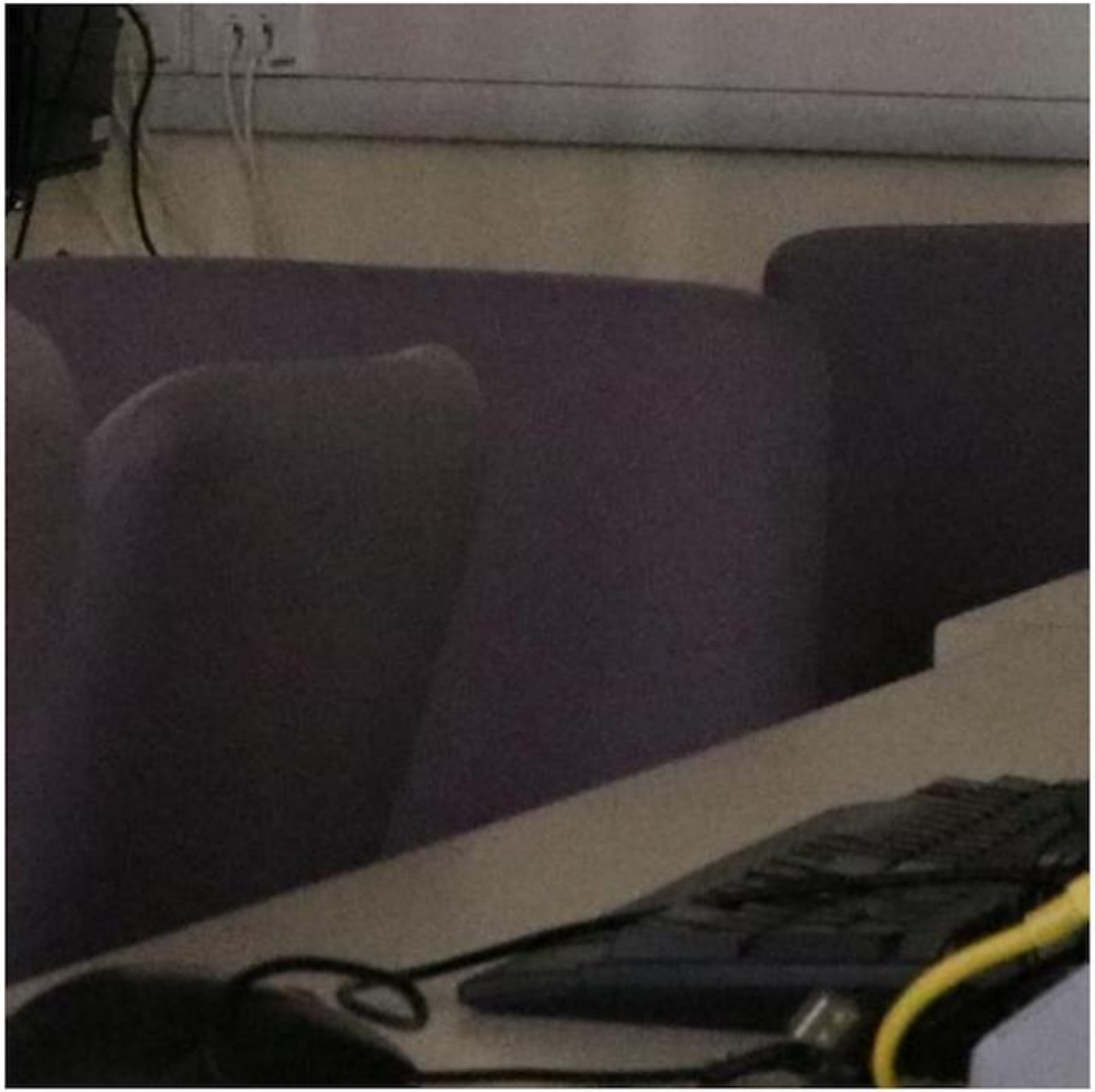

(g) 6400,6.7,1/60


(h) 6400,6.7,1/350


(i) 6400,16,1/60

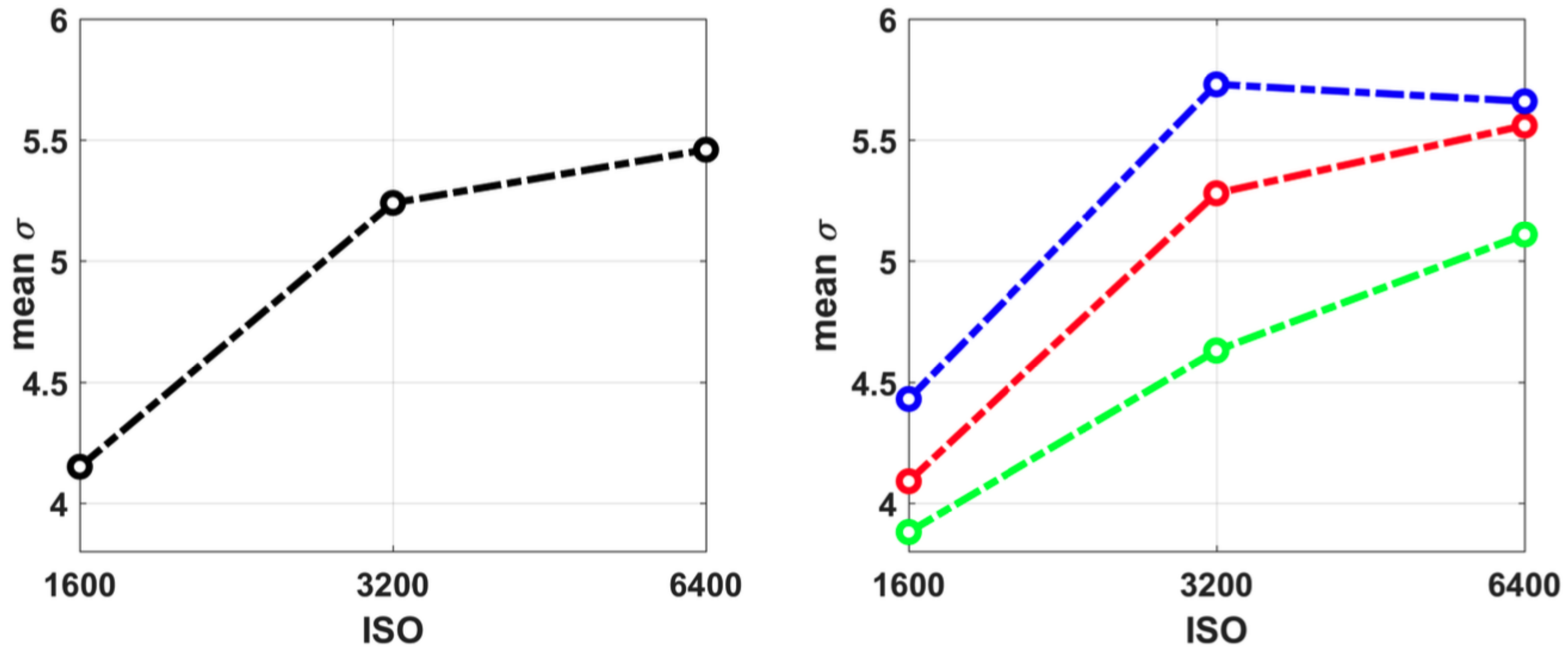
- good example





- bad example

Noise level estimation



- The noise levels in R, G, B channels are estimated via some noise estimation methods
  - Single-image noise level estimation for blind denoising. (IEEE TIP,2013)
  - An efficient statistical method for image noise level estimation. (ICCV, 2015)

Key insight

- 可以採用這個方式，反正我們的 Low-Light 與 noise condition 都可以量化去學，不用亮度剛好一樣
  - 但亮度調到相同時，noise 似乎都不明顯
  - 但這樣才反而是真實表現

Outline

- Only dataset
  - PolyU
  - LLRVD
- Dataset and Synthesis
  - LRD
  - LRID
- Discussion

Low-Light Raw Video Denoising With a High-Quality Realistic Motion Dataset

Ying Fu, *Senior Member, IEEE*, Zichun Wang, Tao Zhang, and Jun Zhang

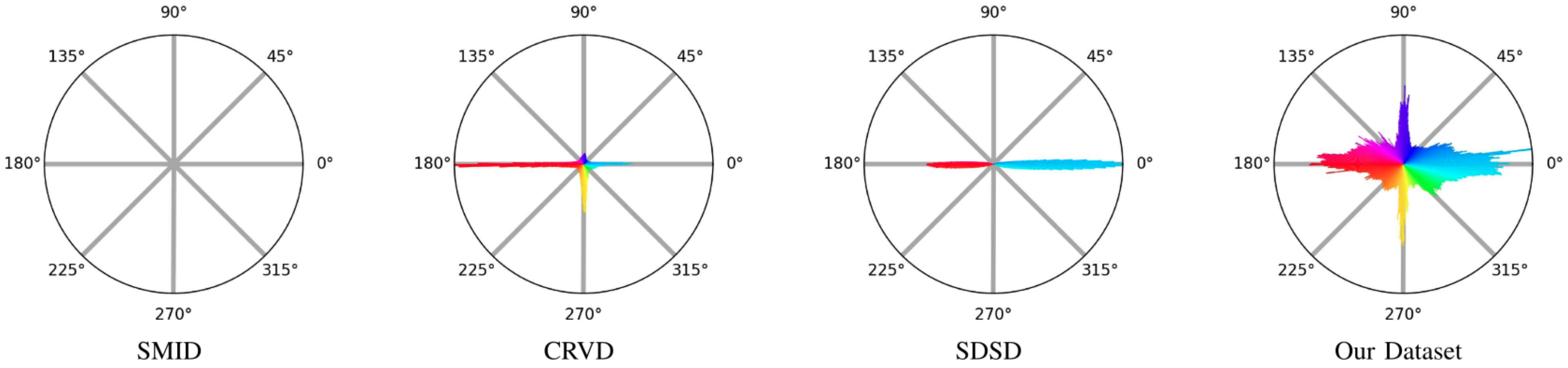
IEEE TOM 2023

Presenter: Hao Wang

Advisor: Prof. Chia-Wen Lin

Introduction

- collect a raw video denoising dataset in low-light with complex motion and high-quality ground truth



Data collection

- collect 70 high-quality 4 k videos from the internet, then play them on the DELL U2720QM monitor.
- Use a Sony Alpha 7R IV full-frame mirrorless camera.

Algorithm 1: Dataset Capture Protocol.

Require: $t_b = 1\ s, g_b = ISO\ 100$;

Posit camera until moire pattern disappears. Distance between the camera and monitor should ensure one monitor pixel is smaller than a camera sensor pixel;

for Each Scene **do**

  Meter the scene to find the aperture size f that well exposes the video;

  Choose 3 out of 6 random low-light ratios r, $r \in [100, 320]$;

  for Each frame in the video **do**

    Take the reference frame at exposure setting (f, t_b, g_b);

    for Each low light ratio r **do**

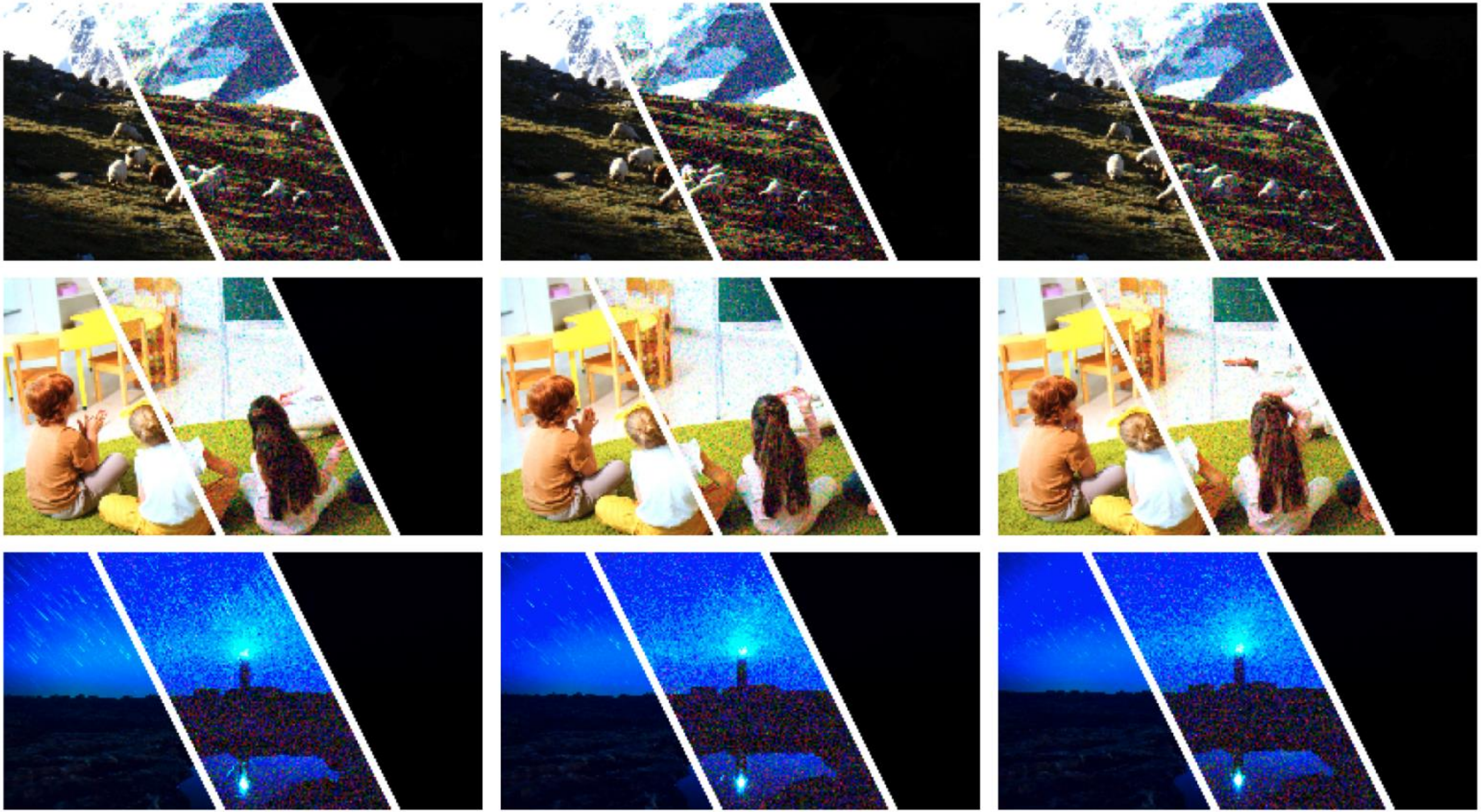
      Take the noisy frame at $(f, t_b/r, g_b)$;

    end

    Play the next frame of the video;

  end

end



Key insight

- noise 要如何量化？
  1. pixel wise 表示:
     - ➢ 用 LLRVD 的放大亮度再相減？
       - 但噪音並不是線性關係
       - Unreal
  2. 數值表示：
     - ➢ 用現成的 paper Single-image noise level estimation for blind denoising. (IEEE TIP,2013) 或 An efficient statistical method for image noise level estimation. (ICCV, 2015)
     - ➢ 直接用相機的 aperture, shutter speed 與 ISO 做為參數(explosure value)

Outline

- Only dataset
  - PolyU
  - LLRVD
- Dataset and Synthesis
  - LRD
  - LRID
- Discussion

Towards General Low-Light Raw Noise Synthesis and Modeling

Feng Zhang[1] Bin Xu[2] Zhiqiang Li[1,2] Xinran Liu[1] Qingbo Lu[2] Changxin Gao[1] Nong Sang[1†]
[1]National Key Laboratory of Multispectral Information Intelligent Processing Technology,
School of Artificial Intelligence and Automation, Huazhong University of Science and Technology
[2]DJI Technology Co., Ltd
{fengzhangaia, lxryx, cgao, nsang}@hust.edu.cn, {mila.xu, cristopher.li, qingbo.lu}@dji.com

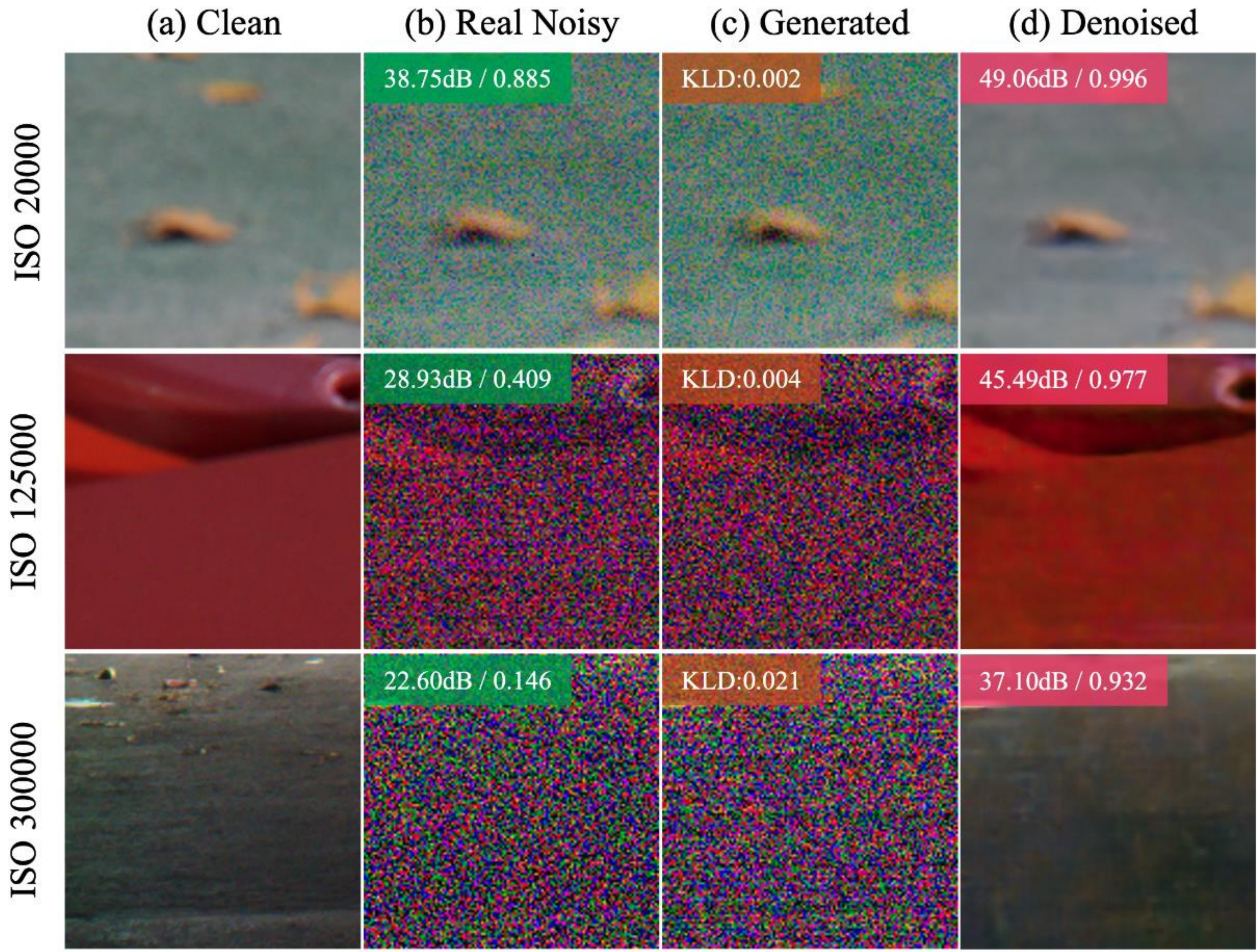
ICCV 2023

Presenter: Hao Wang

Advisor: Prof. Chia-Wen Lin

Introduction

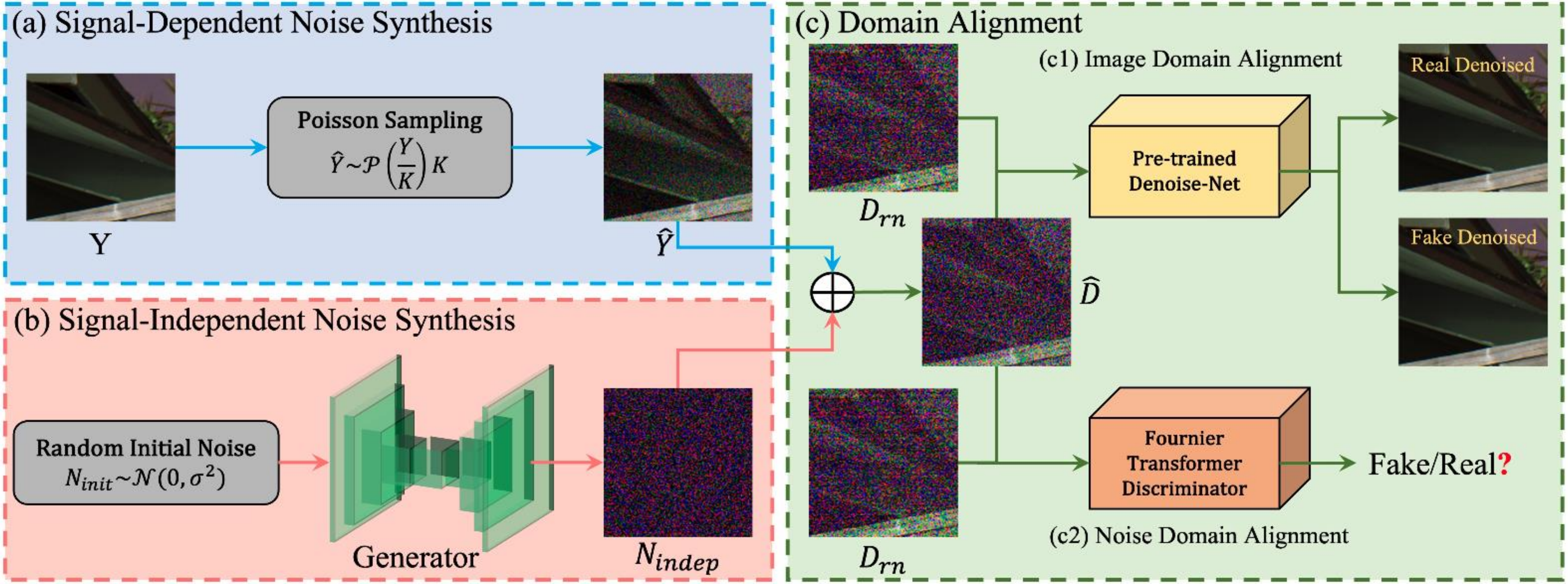
- Propose a general **noise model** to **imitate accurate low-light raw noise** on different sensors.
- Establish **Fourier transformer discriminator (FTD),** which **encourages the generator** to favor solutions.
- Collect a **new large-scale dataset**.



Framework

$$\mathcal{L}_1 = \| P(\hat{D}) - P(D_{rn}) \|_1,$$

$$\mathcal{L}_{per} = \| \phi(P(\hat{D}) - \phi(P(D_{rn}) \|_2^2,$$



$$\mathcal{L}_{adv} = \mathop{\mathbb{E}}_{\hat{D} \sim \mathbb{P}_g} [D_F(\hat{D})] - \mathop{\mathbb{E}}_{D_{rn} \sim \mathbb{P}_r} [D_F(D_{rn})] + \lambda \mathop{\mathbb{E}}_{\tilde{x} \sim \mathbb{P}_{\tilde{x}}} [(\|\nabla_{\tilde{x}} D_F(\tilde{x})\|_2 - 1)^2],$$

ISO levels,
exposure times,
in-camera noise profiles

Data collection

- Clean
  - long-exposure image at **ISO 100** to get a noise-free reference image

- Degraded
  - Total image:1800 pairs (**100 images $\times$ 6 ISO levels $\times$ 3 exposure value**)
  - different ISO levels and Evs
  - 6 different **ISO levels** ranging from **200 to 6400**

$$\mathrm{EV} = \log_2 \frac{N^2}{t}$$

Data collection

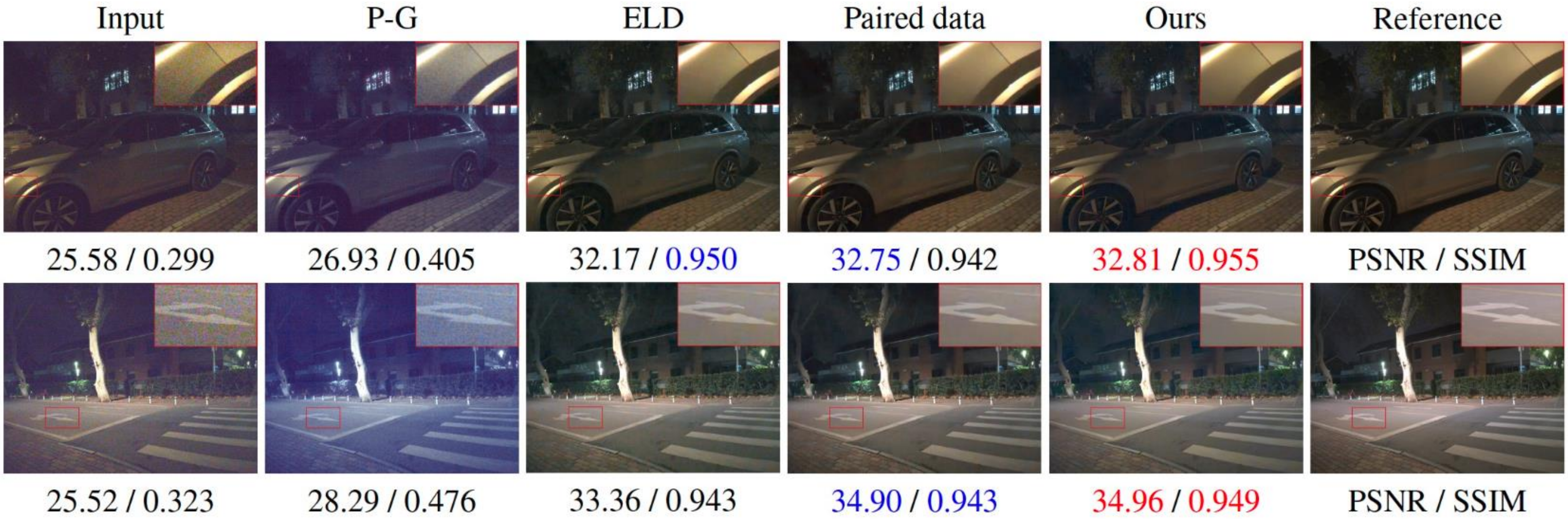
Figure 4. Example images of the LRD dataset. First column: long exposure reference (ground truth) images. Second column: low-light images with -1EV. Third column: low-light images with -2EV. Fourth column: low-light images with -3EV.

Experiment

Dataset	Ratio	Physics-based		Real-noise-based	DNN-based	
		Poisson-Gaussian	ELD	Paired data	Noise Flow	Ours
		PSNR / SSIM	PSNR / SSIM	PSNR / SSIM	PSNR / SSIM	PSNR / SSIM
SID	$\times100$	37.51 / 0.856	41.21 / 0.952	41.39 / 0.954	36.75 / 0.787	41.95 / 0.956
	$\times250$	31.67 / 0.765	38.54 / 0.929	38.90 / 0.937	33.98 / 0.739	39.25 / 0.931
	$\times300$	28.53 / 0.667	35.35 / 0.908	36.55 / 0.922	31.82 / 0.713	36.03 / 0.909
ELD	$\times100$	39.46 / 0.785	45.06 / 0.975	43.80 / 0.963	38.68 / 0.793	44.95 / 0.979
	$\times200$	33.81 / 0.615	43.21 / 0.954	41.54 / 0.918	36.30 / 0.713	43.32 / 0.966
LRD	-1EV	33.77 / 0.895	38.31 / 0.968	38.80 / 0.970	35.19 / 0.874	38.89 / 0.971
	-2EV	32.99 / 0.856	37.35 / 0.959	37.88 / 0.961	34.55 / 0.842	37.95 / 0.962
	-3EV	31.44 / 0.770	36.49 / 0.950	36.92 / 0.951	33.72 / 0.826	37.01 / 0.953

- Denoising models are optimized using the generated training pairs from the trained generator
- Even partially **outperforms** the denoiser trained with **real paired data**
  - the real image pairs still suffer from **luminance misalignment** and **pixel misalignment**

Experiment



Input
P-G
ELD
Paired data
Ours
Reference
25.58 / 0.299
26.93 / 0.405
32.17 / 0.950
32.75 / 0.942
32.81 / 0.955
PSNR / SSIM
25.52 / 0.323
28.29 / 0.476
33.36 / 0.943
34.90 / 0.943
34.96 / 0.949
PSNR / SSIM

Key insight

- 噪音採集方式跟 PolyU 差不多
  - 調整 ISO 與 Exposure Value, EV
  - Loss 的設計用 adversarial loss，也不是直接使用 L1 Loss
- 有將 signal dependent 與 independent noise 分開處理

Outline

- Only dataset
  - PolyU
  - LLRVD
- Dataset and Synthesis
  - LRD
  - LRID
- Discussion

Learnability Enhancement for Low-Light Raw Image Denoising: A Data Perspective

Hansen Feng, Lizhi Wang, *Member, IEEE*, Yuzhi Wang, Haoqiang Fan,
and Hua Huang, *Senior Member, IEEE*

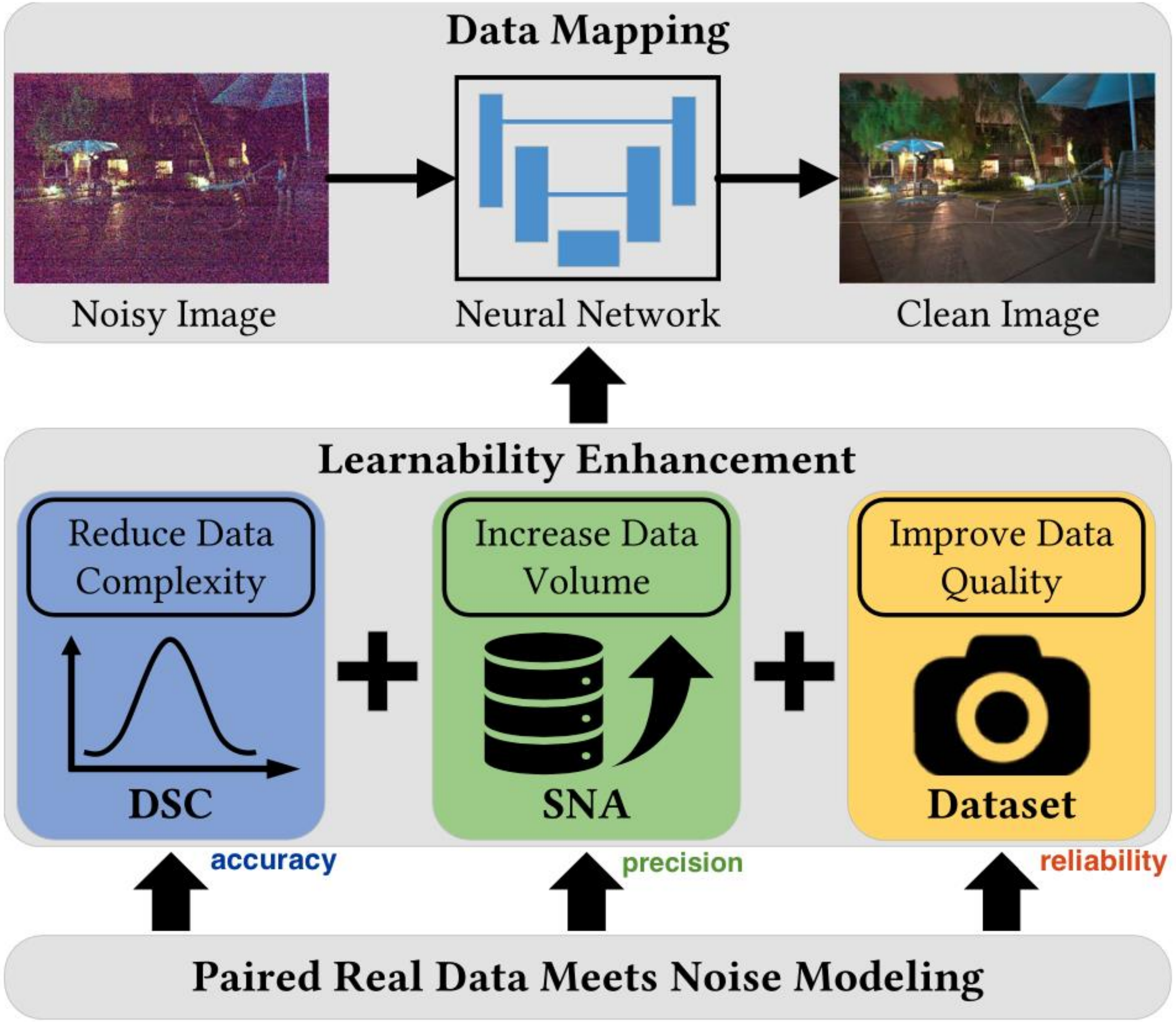
IEEE TPAMI 2024

Presenter: Hao Wang

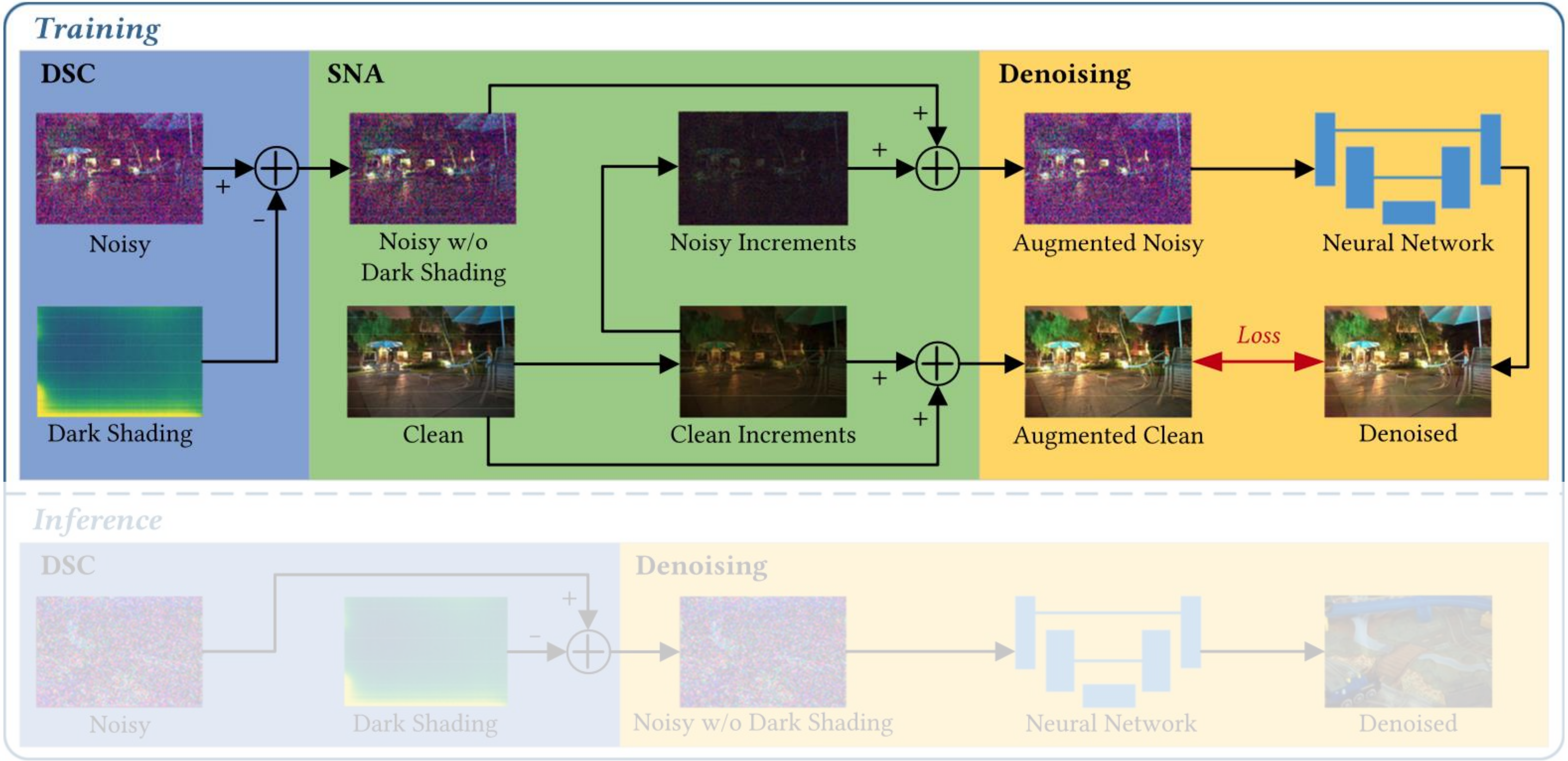
Advisor: Prof. Chia-Wen Lin

Introduction

- Introduce a learnability **enhancement strategy** for low-light raw image denoising by **reforming paired real data**.

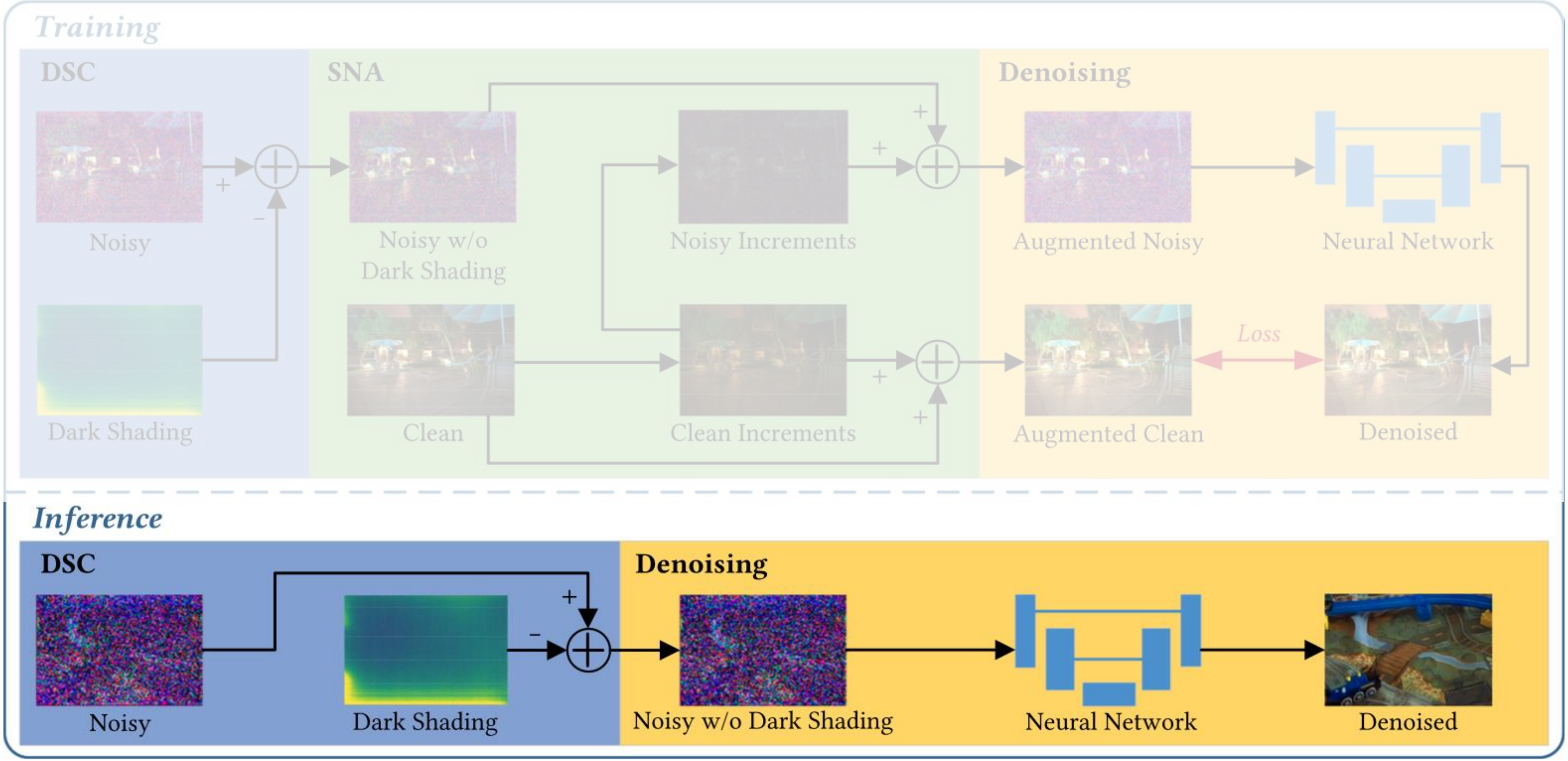


Framework

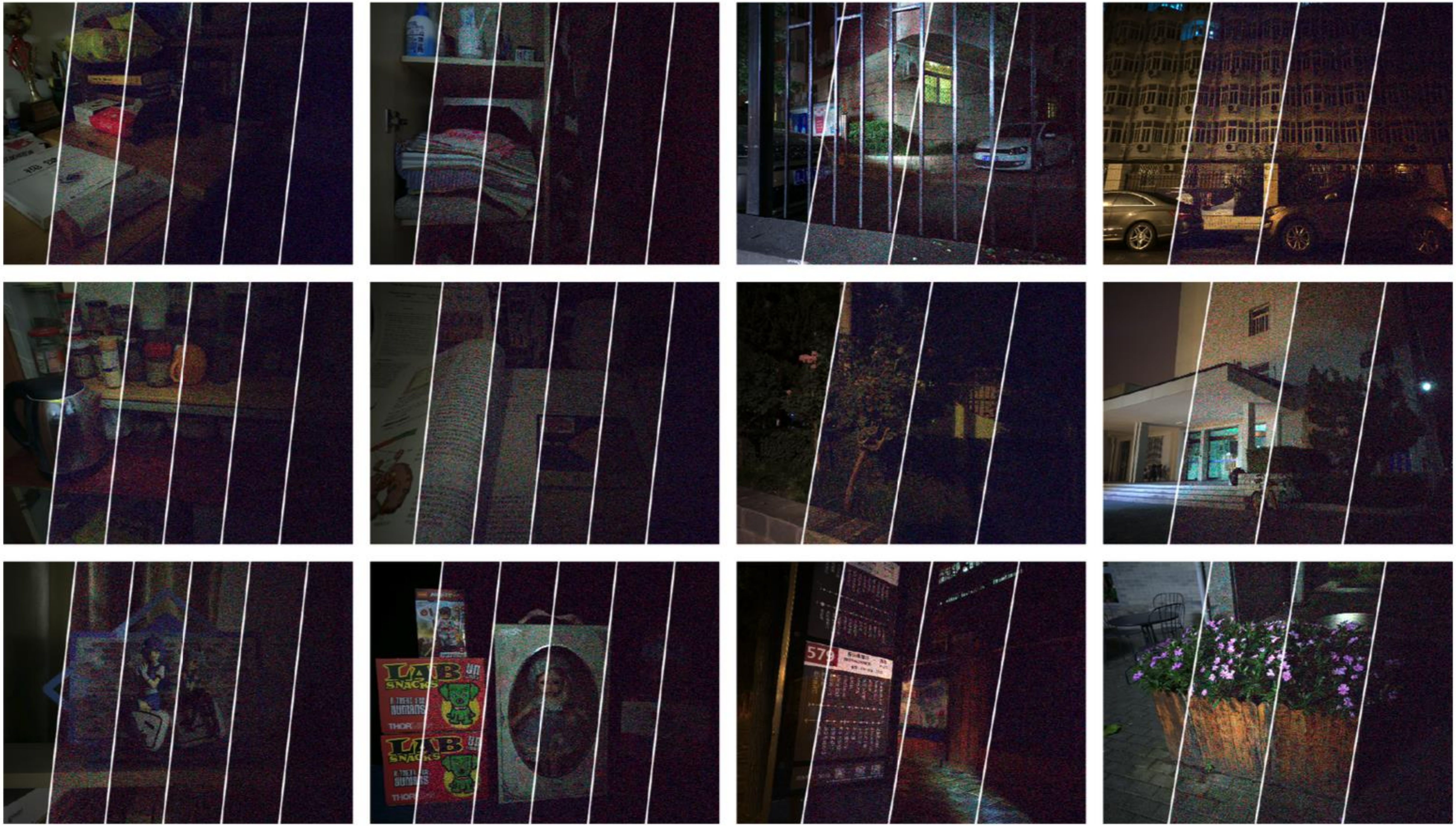


Training
DSC
SNA
Denoising
Noisy
Dark Shading
Noisy w/o
Dark Shading
Clean
Noisy Increments
Clean Increments
Augmented Noisy
Augmented Clean
Neural Network
Denoised
Loss
Inference
DSC
Denoising
Noisy
Dark Shading
Noisy w/o Dark Shading
Neural Network
Denoised

Framework



Training
DSC
SNA
Denoising
Noisy
Dark Shading
Noisy w/o Dark Shading
Clean
Noisy Increments
Clean Increments
Augmented Noisy
Augmented Clean
Neural Network
Denoised
Loss
Inference
DSC
Denoising
Noisy
Dark Shading
Noisy w/o Dark Shading
Neural Network
Denoised


LAB SNACKS
THOR
LAB SNACKS
THOR
579

Experiment result

Dataset	Ratio	Input PSNR / SSIM	P-G [38] PSNR / SSIM	ELD [21] PSNR / SSIM	SFRN [46] PSNR / SSIM	Paired PSNR / SSIM	Ours PSNR / SSIM
ELD	×100	30.85 / 0.5045	42.05 / 0.8721	45.45 / 0.9754	46.38 / 0.9793	44.47 / 0.9676	46.99 / 0.9840
	×200	25.92 / 0.2607	38.18 / 0.7827	43.43 / 0.9544	44.38 / 0.9651	41.97 / 0.9282	44.85 / 0.9686
	Average	28.38 / 0.3826	40.12 / 0.8274	44.44 / 0.9649	45.38 / 0.9722	43.22 / 0.9479	45.92 / 0.9763
SID	×100	29.10 / 0.5266	39.44 / 0.8995	41.95 / 0.9530	42.81 / 0.9568	42.06 / 0.9548	43.47 / 0.9606
	×250	23.95 / 0.3595	34.32 / 0.7681	39.44 / 0.9307	40.18 / 0.9343	39.60 / 0.9380	41.04 / 0.9471
	×300	22.00 / 0.2752	30.66 / 0.6569	36.36 / 0.9114	37.09 / 0.9175	36.85 / 0.9227	37.87 / 0.9344
	Average	24.81 / 0.3793	34.52 / 0.7666	39.05 / 0.9303	39.82 / 0.9349	39.32 / 0.9374	40.59 / 0.9465
LRID-Indoor	×64	32.81 / 0.6728	46.14 / 0.9872	48.19 / 0.9898	47.94 / 0.9899	48.77 / 0.9906	49.24 / 0.9916
	×128	29.10 / 0.4621	44.98 / 0.9809	46.55 / 0.9836	46.52 / 0.9854	47.00 / 0.9860	47.47 / 0.9868
	×256	25.07 / 0.2380	43.31 / 0.9682	44.39 / 0.9730	44.74 / 0.9789	44.74 / 0.9786	45.36 / 0.9804
	×512	20.53 / 0.0872	40.80 / 0.9429	41.56 / 0.9452	42.46 / 0.9652	42.40 / 0.9647	43.09 / 0.9671
	×1024	15.43 / 0.0241	37.74 / 0.8905	37.50 / 0.8915	40.10 / 0.9453	40.07 / 0.9437	40.20 / 0.9453
	Average	24.59 / 0.2968	42.59 / 0.9539	43.64 / 0.9566	44.35 / 0.9729	44.60 / 0.9727	45.07 / 0.9743
LRID-Outdoor	×64	33.25 / 0.7255	42.16 / 0.9796	45.00 / 0.9841	45.05 / 0.9850	45.84 / 0.9876	46.27 / 0.9884
	×128	29.49 / 0.5100	41.48 / 0.9709	43.48 / 0.9734	43.67 / 0.9766	44.50 / 0.9821	44.86 / 0.9834
	×256	25.26 / 0.2557	40.36 / 0.9525	41.31 / 0.9450	41.89 / 0.9591	42.66 / 0.9709	42.99 / 0.9703
	Average	29.33 / 0.4971	41.33 / 0.9677	43.26 / 0.9675	43.54 / 0.9736	44.33 / 0.9802	44.71 / 0.9807

The red color indicates the best results and the blue color indicates the second-best results.


Input
P-G [38]
ELD [21]
SFRN [46]
Paired
Ours
Ground Truth
LRID-Indoor
25.82 / 0.2597
44.37 / 0.9721
45.27 / 0.9765
45.75 / 0.9810
45.49 / 0.9806
46.27 / 0.9837
PSNR / SSIM
21.34 / 0.0811
41.24 / 0.9060
42.26 / 0.9336
43.35 / 0.9639
43.13 / 0.9619
43.70 / 0.9678
PSNR / SSIM
LRID-Outdoor
26.25 / 0.2687
41.72 / 0.9662
43.53 / 0.9722
44.06 / 0.9811
44.52 / 0.9815
44.94 / 0.9824
PSNR / SSIM
25.55 / 0.2505
42.76 / 0.9468
41.19 / 0.9155
42.47 / 0.9367
44.73 / 0.9749
45.24 / 0.9783
PSNR / SSIM

Data collection

- Clean
  - **ISO-100**
  - Long exposure time
  - fusion of multiple raw images for various misalignments flexibly
- Degraded
  - **ISO-6400**
  - Short exposure time
- **Exposure time ratios** of long- and short-exposure images are 64, 128, 256, 512, and 1024

Key insight

- 噪音採集方式跟 PolyU 差不多
  - 調整 ISO 作為 noise 的主要來源
  - 再提供多種的 exposure time ration
- 一樣有將 signal dependent 與 independent noise 分開處理
  - 我們的模型需要也分開處理嗎？
- Synthesis 的兩篇都沒有特別提到亮度需要保持一致，可能在生成 noise 時，一併處理？

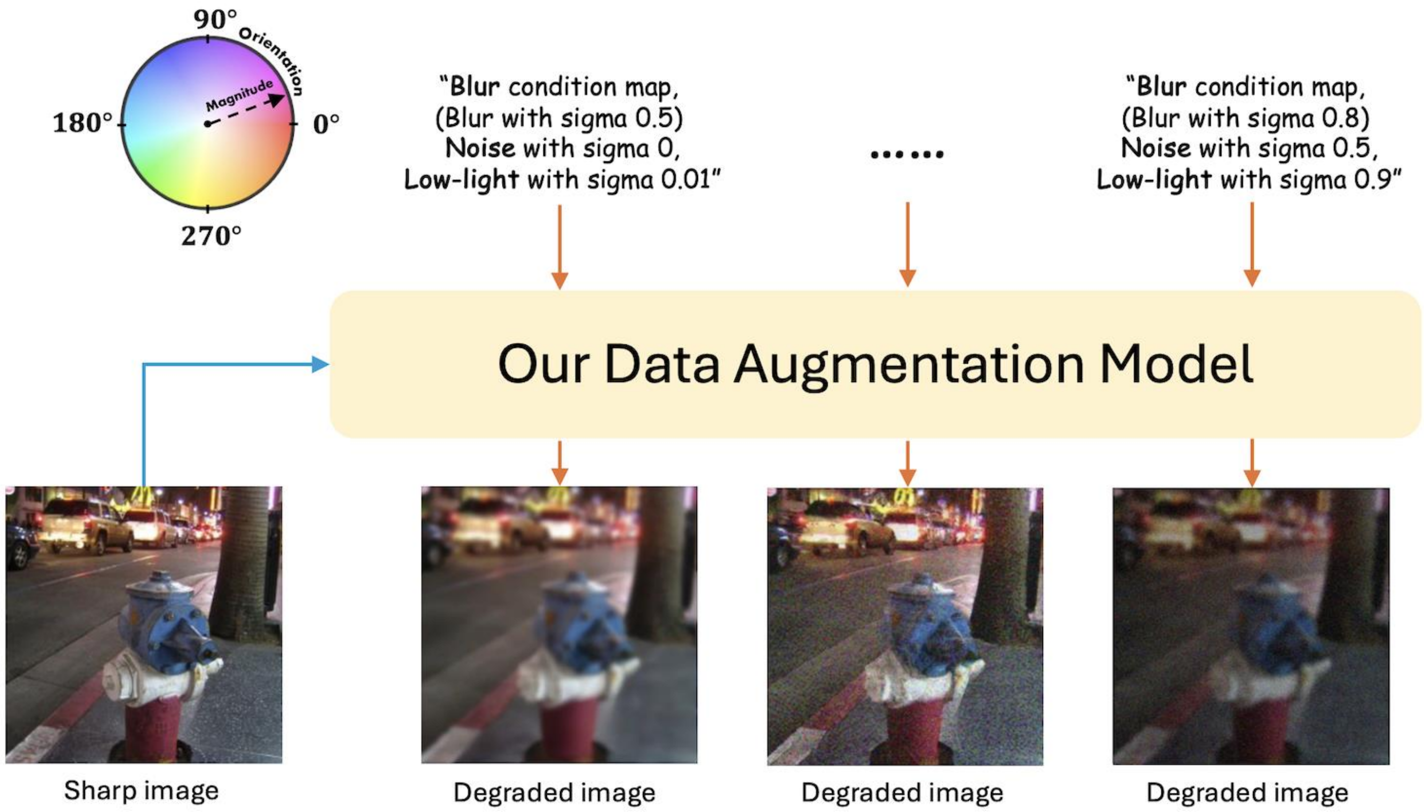
Outline

- Only dataset
  - PolyU
  - LLRVD
- Dataset and Synthesis
  - LRD
  - LRID
- Discussion

Conclusion

- noise 如何得到
    1. 結論：**光暗 ISO 很大**
    2. 收集 Noise 時，**通常** exposure time 要 short，long exposure time 給 Clean image
    3. **但我們需要 blur，所以長必須給 Degraded**，我們只能**進而調 ISO 與 aperture光圈**，有的還將多張圖像都平均以減少 misalignment，但我們是雙相機系統，可以不用平均，以免失真或是拉長曝光時間造成blur

Noise control



- noise 要如何量化？
  1. pixel wise 表示:
     - ➢ 用 LLRVD 的放大亮度再相減？
       - 但噪音並不是線性關係
       - Unreal
  2. 數值表示：
     - ➢ 用現成的 paper Single-image noise level estimation for blind denoising. (IEEE TIP,2013) 或 An efficient statistical method for image noise level estimation. (ICCV, 2015)
     - ➢ 直接用相機的 aperture, shutter speed 與 ISO 做為參數

Blur map

- 隨便一張新的 unseen sharp image 並沒有 optical flow， 要如何用 pixed-wise motion field map 來表達 blur
  - 用 segmentation 來協助？
  - 用 depth estimation 來協助？
  - 單純用數字來表示， 不用 map