NeurIPS 2024

Quality-Improved and Property-Preserved Polarimetric Imaging via Complementarily Fusing

Chu Zhou[1] Yixing Liu[2,3] Chao Xu[4] Boxin Shi[2,3]

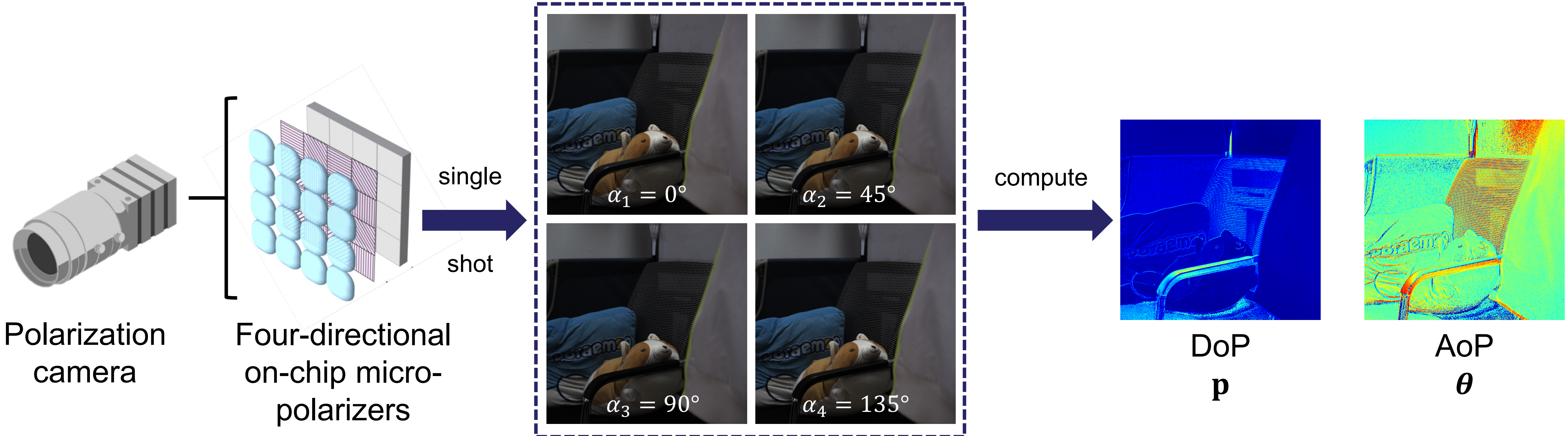
[1]National Institute of Informatics, Japan

[2]State Key Lab for Multimedia Info. Processing, School of CS, Peking University, China

[3]National Eng. Research Center of Visual Tech., School of CS, Peking University, China

[4]National Key Lab of General AI, School of IST, Peking University, China

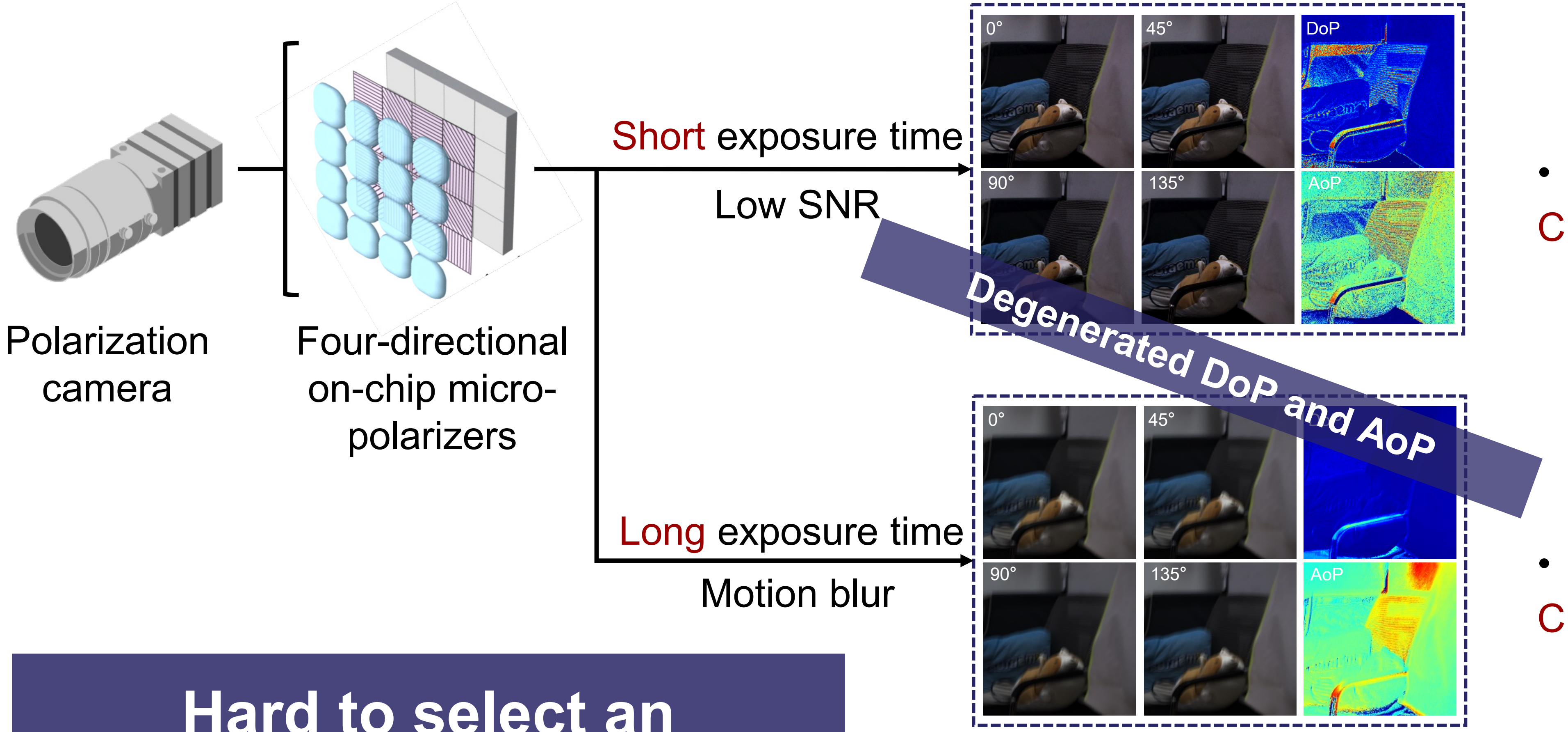
Polarimetric imaging



Polarized snapshot
(four polarized images $\mathbf{I}_{\alpha_{1,2,3,4}}$ with different polarizer angles $\alpha_{1,2,3,4}$)

- A polarization camera can capture a polarized snapshot in a single shot
  - The degree of polarization (DoP) and angle of polarization (AoP) can be directly computed from it
  - Useful for polarization-based vision applications

Polarimetric imaging: difficulties



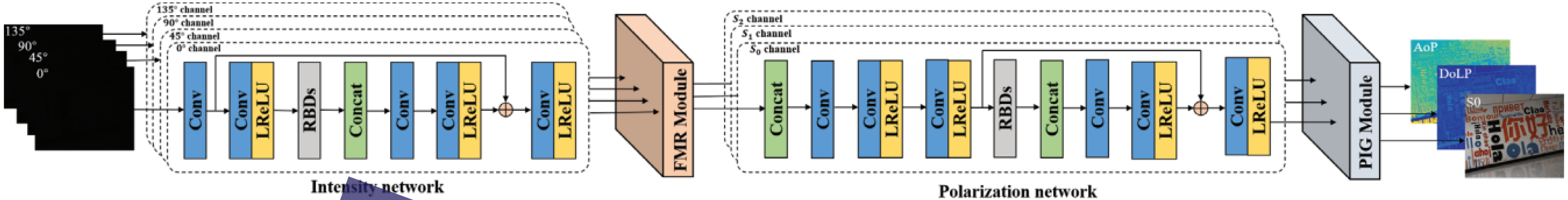
- $\mathbf{L}_{\alpha_{1,2,3,4}}$
  Clear but noisy

- $\mathbf{B}_{\alpha_{1,2,3,4}}$
  Clean but blurry

Hard to select an appropriate exposure time

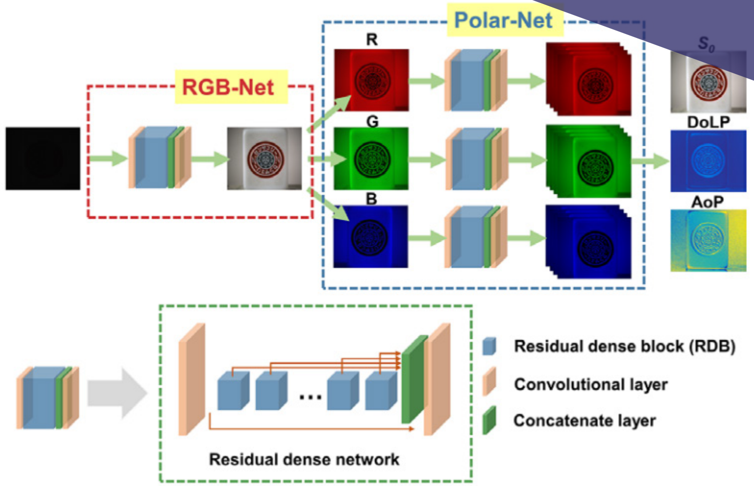
Handling noisy polarized images

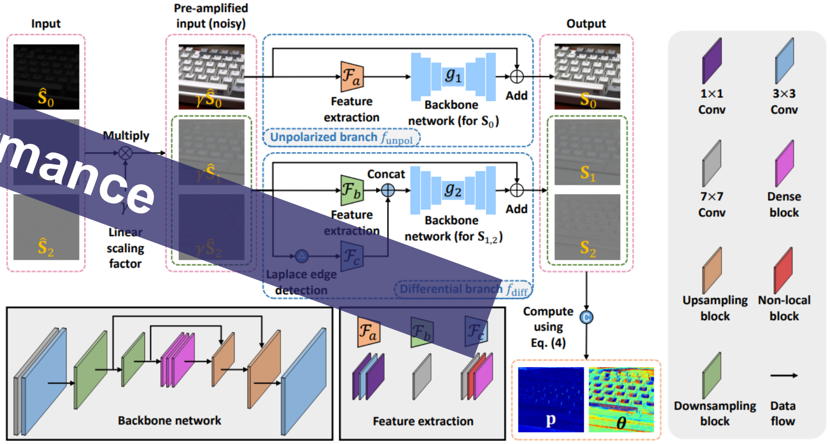
- Low-light enhancement methods for polarized images: handling $\mathbf{L}_{\alpha_{1,2,3,4}}$ only



- Denoise in both the intensity and Stokes domains (ColorPolarNet) [Xu *et al.*, TIM 22]

Limited performance

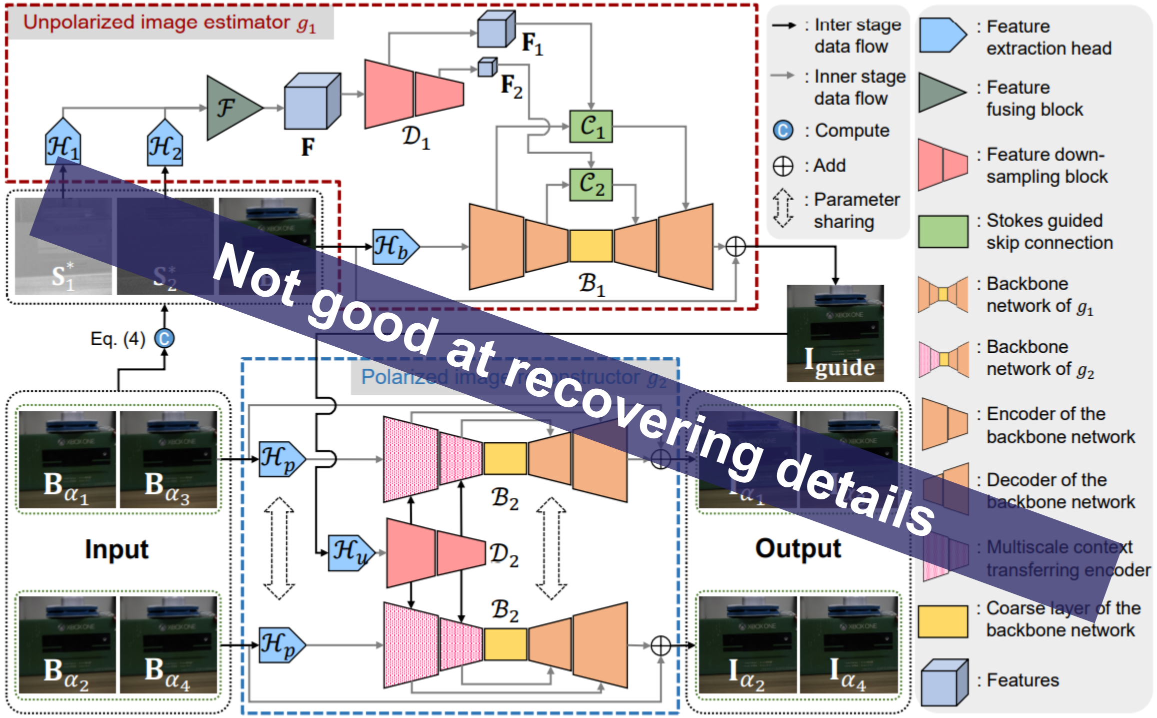




- Denoise in the intensity domain (IPLNet) [Hu *et al.*, OL 20]
- Denoise in the Stokes domain (PLIE) [Zhou *et al.*, AAAI 23]

Handling blurry polarized images

- Deblurring methods for polarized images : handling $\mathbf{B}_{\alpha_{1,2,3,4}}$ only



- Two-stage deblurring pipeline (PolDeblur) [Zhou *et al*., Arxiv 24]

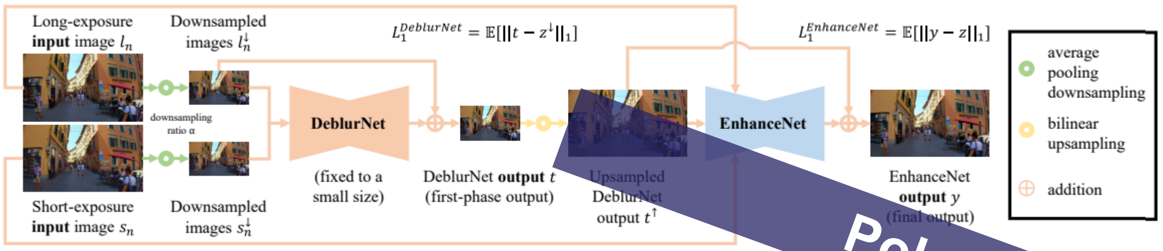
- Instead of enhancing $\mathbf{L}_{\alpha_{1,2,3,4}}$, another approach to obtain high-quality DoP and AoP is deblurring $\mathbf{B}_{\alpha_{1,2,3,4}}$
  - The problem is highly ill-posed

How to reduce the ill-posedness?

Fusing a pair of noisy and blurry images

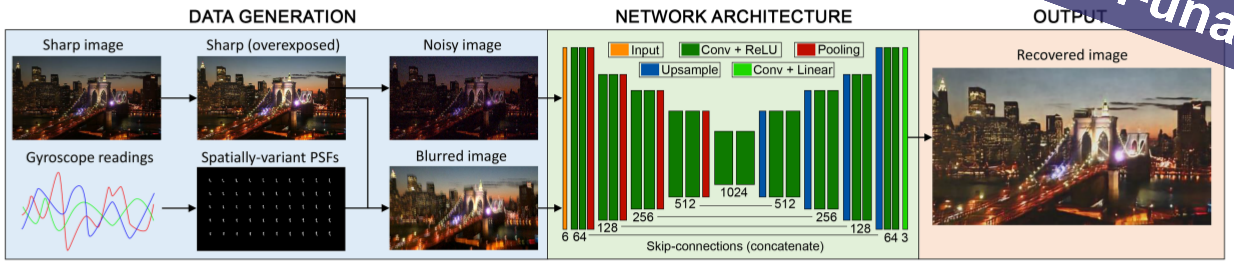
- A pair of noisy and blurry images would provide complementary knowledge

Clear but noisy image
Clean but blurry image
→ Fusing them could produce clean and clear results

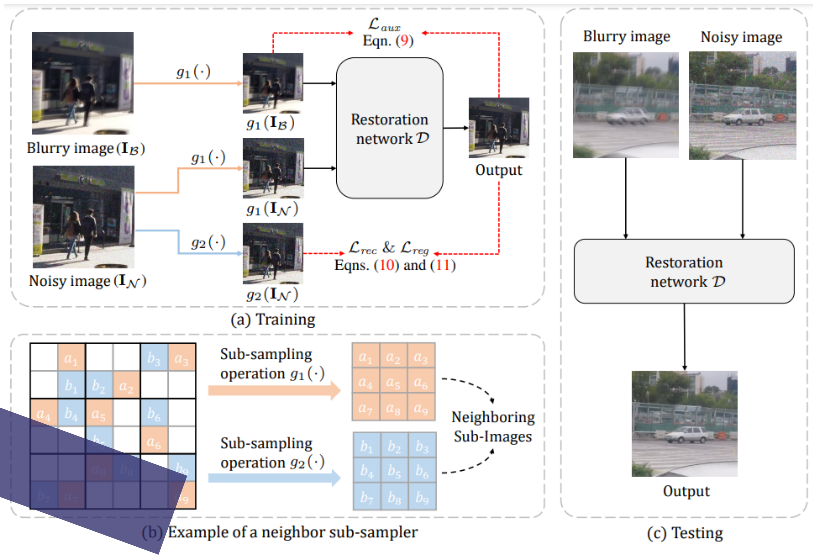


- Two-stage fusing pipeline (D2HNet) [Zhao et al., ECCV 22]

Polarization-unaware



- Single-stage fusing pipeline (LSD2) [Zhao *et al*., BMVC 20]



- Self-supervised fusing pipeline (SelfIR) [Zhang *et al*., NeurIPS 22]

A fusing framework for polarized images

- $\mathbf{L}_{\alpha_{1,2,3,4}}$ and $\mathbf{B}_{\alpha_{1,2,3,4}}$ would also provide complementary knowledge
  - Need a specially designed fusing framework for polarized images that can simultaneously
    - improve the image quality
    - preserve the polarization properties
  - Formulating the fusing framework as maximizing a posteriori estimation

$$\underset{\Psi}{\operatorname{argmax}} f\left(\mathbf{I}_{\alpha_{1,2,3,4}} \middle| \mathbf{L}_{\alpha_{1,2,3,4}}, \mathbf{B}_{\alpha_{1,2,3,4}}, \Psi\right)$$

- Implementing a fusing function f parameterized by Ψ

- Inputs:
  - noisy polarized images $\mathbf{L}_{\alpha_{1,2,3,4}}$
  - blurry polarized images $\mathbf{B}_{\alpha_{1,2,3,4}}$
- Output:
  - clean and clear polarized images $\mathbf{I}_{\alpha_{1,2,3,4}}$

How to implement the fusing function?

Stokes parameters

- Denoting the unpolarized images as $\mathbf{I}$
  - When placing a polarizer with polarizer angle α: $\mathbf{I}_\alpha = \frac{\mathbf{I}\left(1-\mathbf{p}\cos\left(2(\alpha-\boldsymbol{\theta})\right)\right)}{2}$ Malus' law
  - Reformulating the above equation into a polynomial form: $\mathbf{I}_\alpha = \frac{\mathbf{S}_0}{2} - \frac{\cos(2\alpha)\mathbf{S}_1}{2} - \frac{\sin(2\alpha)\mathbf{S}_2}{2}$

  where $\mathbf{S}_0 = \mathbf{I}$, $\mathbf{S}_1 = \mathbf{Ip}\cos(2\boldsymbol{\theta})$, and $\mathbf{S}_2 = \mathbf{Ip}\sin(2\boldsymbol{\theta})$ are called the Stokes parameters

$\mathbf{I}$

$\mathbf{I}_{\alpha_{1,2,3,4}}$

$\mathbf{S}_0$

$\mathbf{S}_1$

$\mathbf{S}_2$

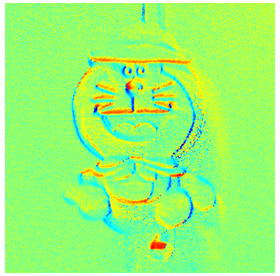
Stokes parameters

- $\mathbf{S}_{0,1,2}$ can be computed from $\mathbf{I}_{\alpha_{1,2,3,4}}$ directly
  - According to the physical meanings of the Stokes parameters:
    - $\mathbf{S}_0$ describes the total intensity of the light, which is polarization-unrelated
    - $\mathbf{S}_1$ describes the difference between the intensity of the vertical $(90°)$ and horizontal $(0°)$ polarized light
    - $\mathbf{S}_2$ describes the difference between the intensity of the $135°$ and $45°$ polarized light

$$\begin{cases} \mathbf{S}_0 = \mathbf{I}_{\alpha_1} + \mathbf{I}_{\alpha_3} = \mathbf{I}_{\alpha_2} + \mathbf{I}_{\alpha_4} \\ \mathbf{S}_1 = \mathbf{I}_{\alpha_3} - \mathbf{I}_{\alpha_1} \\ \mathbf{S}_2 = \mathbf{I}_{\alpha_4} - \mathbf{I}_{\alpha_2} \end{cases}$$

  - Once the Stokes parameters are available, the DoP $\mathbf{p}$ and AoP $\boldsymbol{\theta}$ could be easily acquired

$\mathbf{p}$



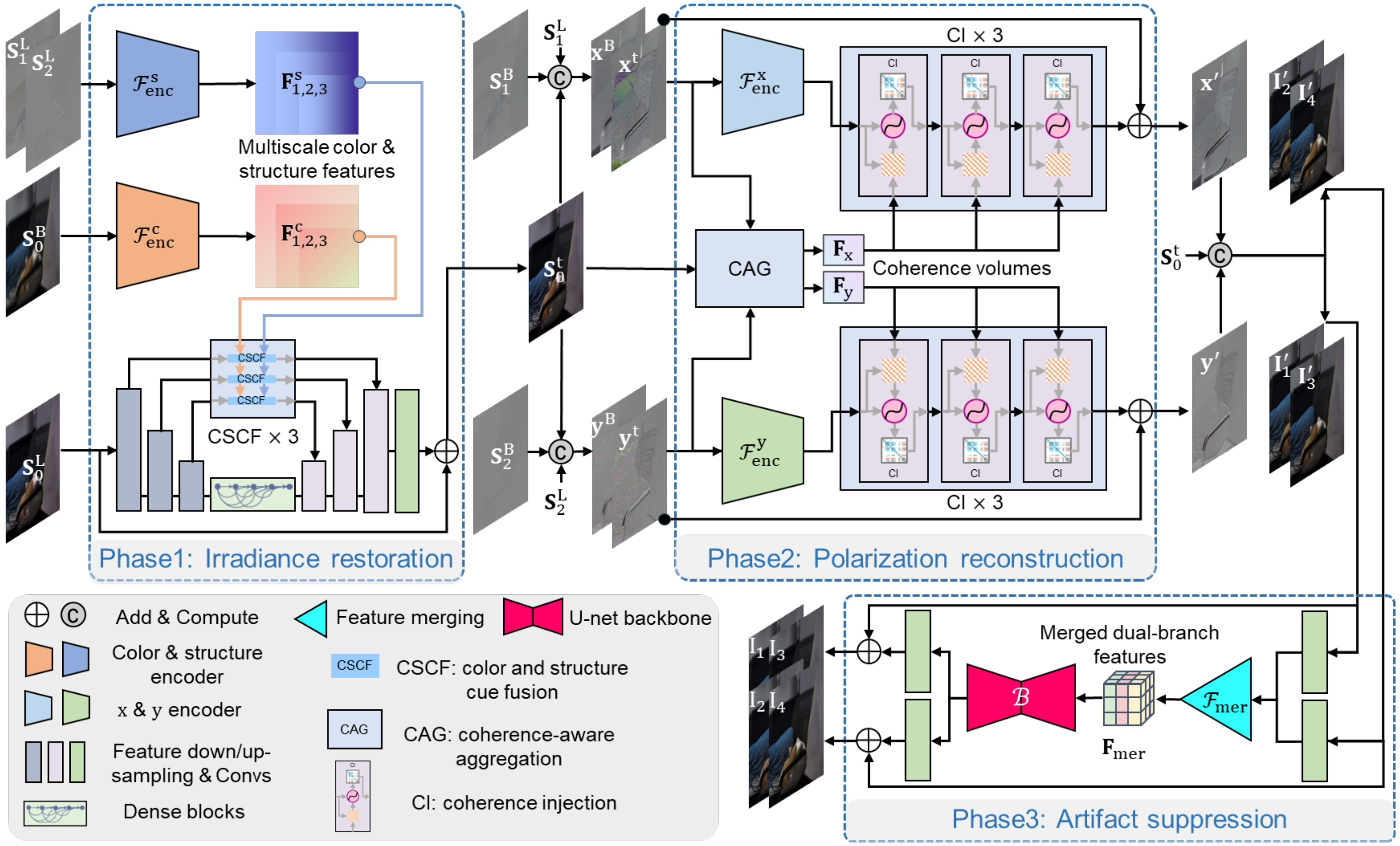
$\boldsymbol{\theta}$

$$\mathbf{p} = \frac{\sqrt{\mathbf{S}_1{}^2 + \mathbf{S}_2{}^2}}{\mathbf{S}_0}$$

$$\boldsymbol{\theta} = \frac{\tan^{-1}\left(\frac{\mathbf{S}_2}{\mathbf{S}_1}\right)}{2}$$

Framework design

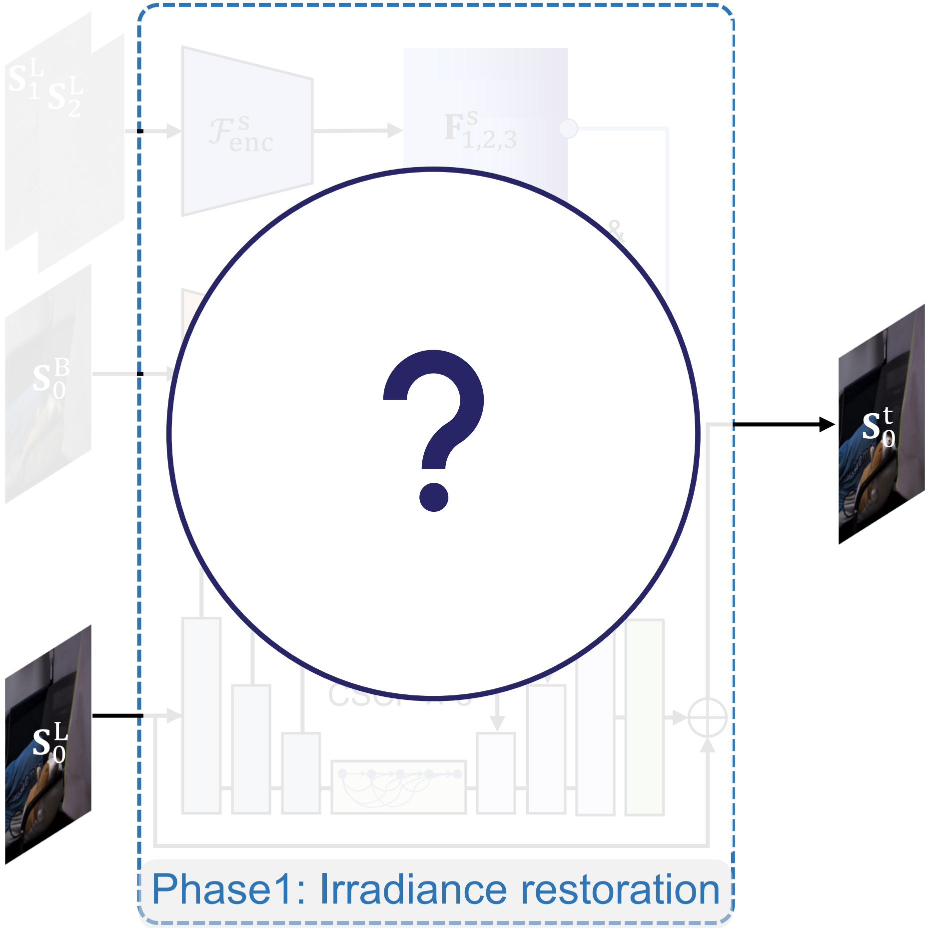
- A neural network-based three-phase fusing scheme



- Phase1: Irradiance restoration
  - Improve the image quality
- Phase2: Polarization reconstruction
  - Preserve the polarization properties
- Phase3: Artifact suppression
  - Refine the overall details

Phase1: Irradiance restoration

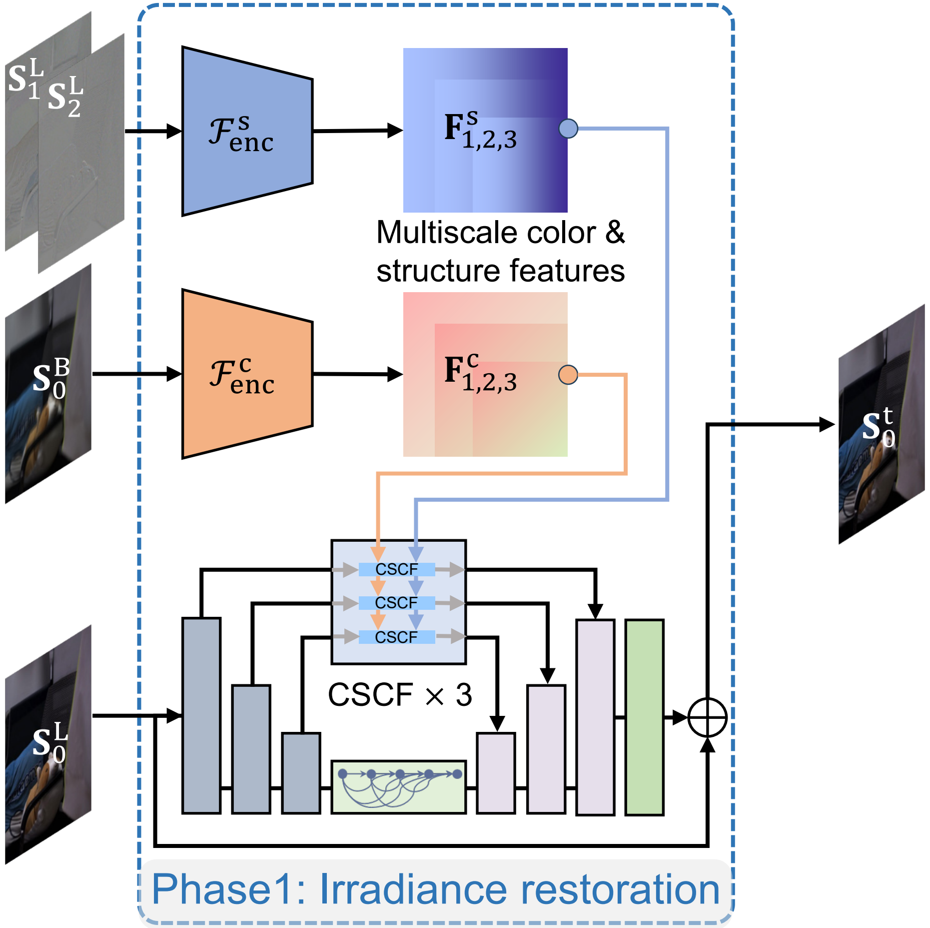
- Goal: restoring the polarization-unrelated high-level irradiance information
  - to obtain the coarse value of the total intensity of the light $\mathbf{S}_0^{\mathrm{t}}$ for providing further guidance



- Since $\mathbf{L}_{\alpha_{1,2,3,4}}$ would retain better contours than $\mathbf{B}_{\alpha_{1,2,3,4}}$
  - We choose to learn the residual between $\mathbf{S}_0^{\mathrm{L}}$ and $\mathbf{S}_0^{\mathrm{t}}$ instead of the residual between $\mathbf{S}_0^{\mathrm{B}}$ and $\mathbf{S}_0^{\mathrm{t}}$
- Difficulties: $\mathbf{S}_0^{\mathrm{L}}$ suffers from color bias and noise
  - Hard to extract features robustly
  → erroneous global tone and less salient local structure

Phase1: Irradiance restoration

- Goal: restoring the polarization-unrelated high-level irradiance information
  - to obtain the coarse value of the total intensity of the light $\mathbf{S}_0^{\mathrm{t}}$ for providing further guidance



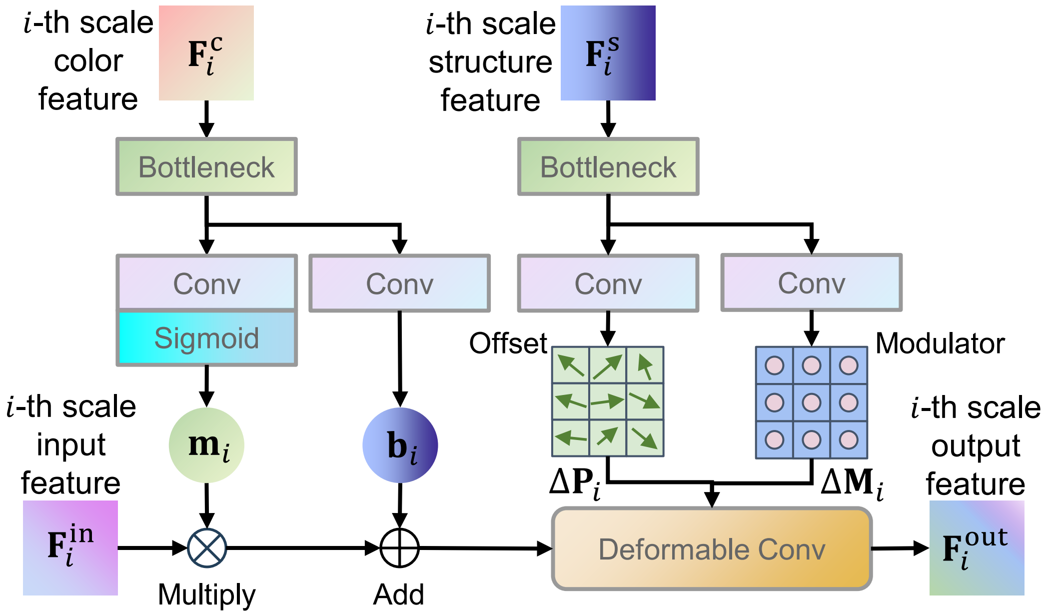
- We observe that $\mathbf{S}_0^{\mathrm{B}}$ and $\mathbf{S}_{1,2}^{\mathrm{L}}$ could provide some cues:
  - $\mathbf{S}_0^{\mathrm{B}}$ contains undamaged color information
    - due to the relatively high SNR of $\mathbf{B}_{\alpha_{1,2,3,4}}$

    → Extract color features $\mathbf{F}_{1,2,3}^{\mathrm{s}}$ from $\mathbf{S}_0^{\mathrm{B}}$
  - $\mathbf{S}_{1,2}^{\mathrm{L}}$ contain distinctive structure information
    - since both of them describe the difference between two polarized images

    → Extract structure features $\mathbf{F}_{1,2,3}^{\mathrm{c}}$ from $\mathbf{S}_{1,2}^{\mathrm{L}}$

CSCF: color and structure cue fusion

- How to mitigate erroneous global tone and less salient local structure?



CSCF (color and structure cue fusion) module

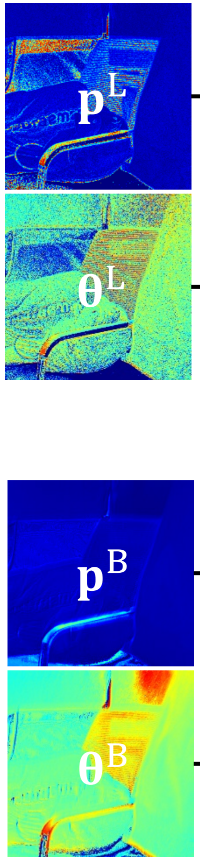
- In the feature space, we propose to
  - apply an affine transformation to $\mathbf{F}_i^{\mathrm{in}}$ to
    - adjust the color in the feature space

    → $\mathbf{F}_i^{\mathrm{t}} = \mathbf{m}_i \odot \mathbf{F}_i^{\mathrm{in}} + \mathbf{b}_i$
  - apply a deformable convolution layer to
    - align the gradients
    - overcome the possible shifts caused by the exposure interval

    → $\mathbf{F}_i^{\mathrm{out}} = \mathcal{D}(\mathbf{F}_i^{\mathrm{t}}, \Delta\mathbf{P}_i, \Delta\mathbf{M}_i)$

Phase2: Irradiance restoration

- Goal: establishing the physical correlation between the polarized images
  - by reconstructing the high-quality DoP and AoP



Shall we design Phase2 in this way?

High-quality DoP and AoP

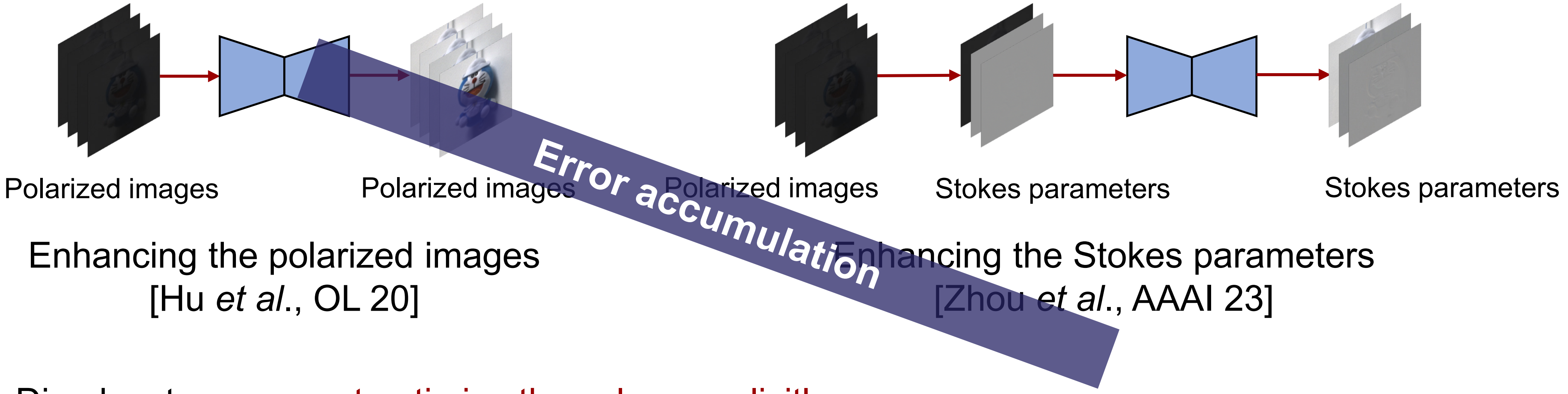
- No!
  - The degeneration patterns of the DoP and AoP could be complicated due to their non-linearity

$$\mathbf{p} = \frac{\sqrt{\mathbf{s}_1{}^2 + \mathbf{s}_2{}^2}}{\mathbf{s}_0}$$

$$\boldsymbol{\theta} = \frac{\tan^{-1}\left(\frac{\mathbf{S}_2}{\mathbf{S}_1}\right)}{2}$$

Repairing the degenerated values

- How to handle the non-linearity?
  - Previous solution: adopting an indirect approach
    - Repairing the degenerated values of the DoP and AoP in the image domain or Stokes domain
  - Let's take the low-light enhancement methods for polarized images as examples:



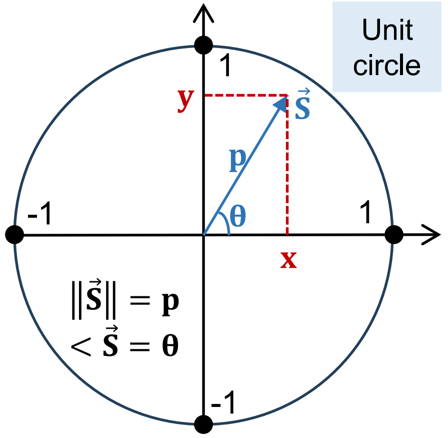
Enhancing the polarized images
[Hu *et al.*, OL 20]

Enhancing the Stokes parameters
[Zhou *et al.*, AAAI 23]

  - Disadvantage: cannot optimize the values explicitly

Cartesian coordinate representation

- Can we achieve this in a direct manner?



$$\mathbf{p} = \frac{\sqrt{\mathbf{S_1}^2 + \mathbf{S_2}^2}}{\mathbf{S_0}}$$

$$\boldsymbol{\theta} = \frac{\tan^{-1}\left(\frac{\mathbf{S_2}}{\mathbf{S_1}}\right)}{2}$$

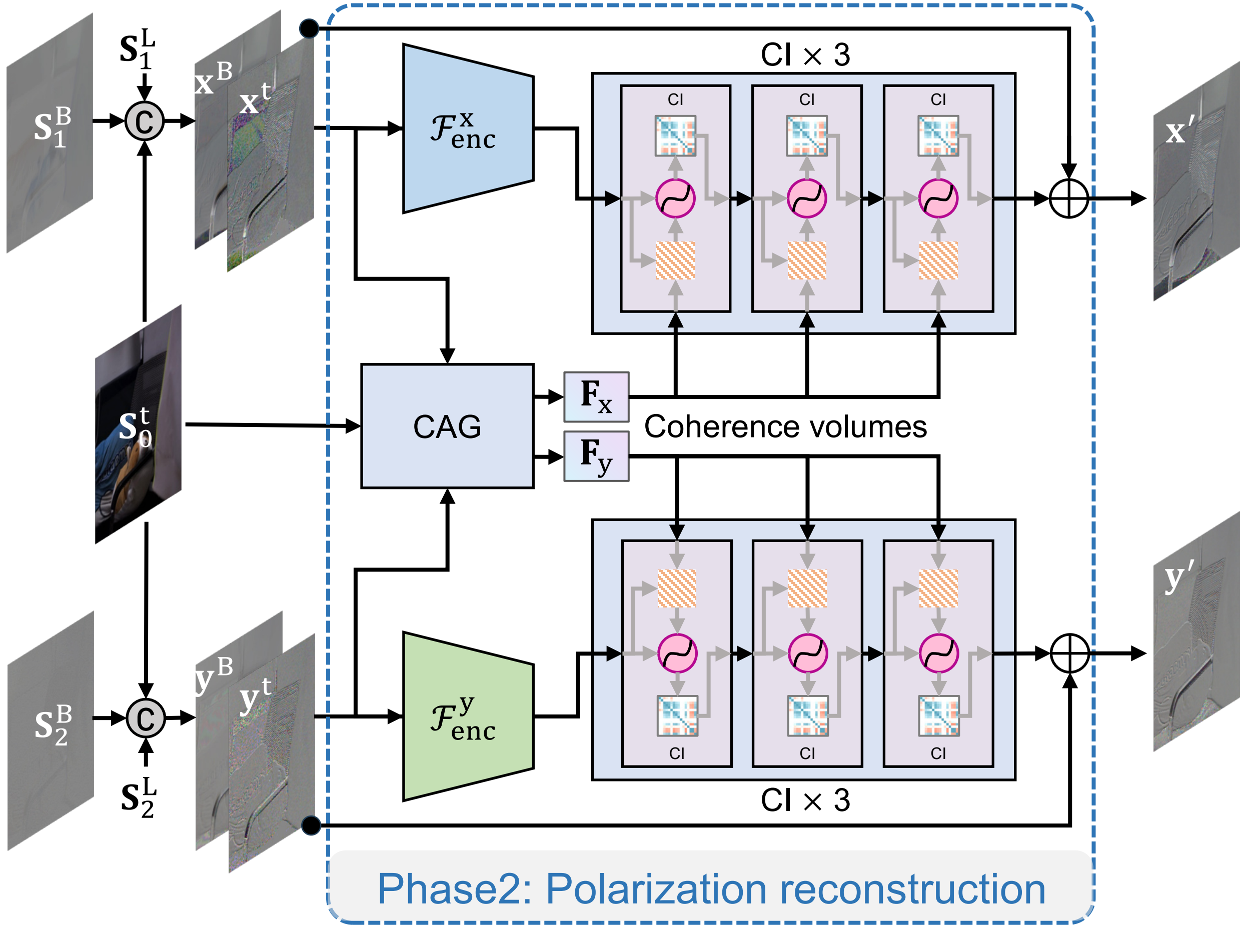
- Given a vector $\vec{\mathbf{S}}$ lying inside a unit circle
  - $(\mathbf{p}, \boldsymbol{\theta})$: the polar coordinate representation (PCR)
    - $\mathbf{p}$ is the magnitude
    - $\boldsymbol{\theta}$ is the angle
  - $(\mathbf{x}, \mathbf{y})$: the Cartesian coordinate representation (CCR)
    - $\mathbf{x}$ is the horizontal value
    - $\mathbf{y}$ is the vertical value

    → $\mathbf{x} = \frac{\mathbf{S_1}}{\mathbf{S_0}}, \mathbf{y} = \frac{\mathbf{S_2}}{\mathbf{S_0}}$

- Advantages of reconstructing the DoP and AoP in CCR:
  - Not only reduce the non-linearity
  - But also optimize the values in a direct manner

Phase2: Irradiance restoration

- Goal: establishing the physical correlation between the polarized images
  - by reconstructing the high-quality DoP and AoP in a Cartesian coordinate representation $(\mathbf{x}', \mathbf{y}')$



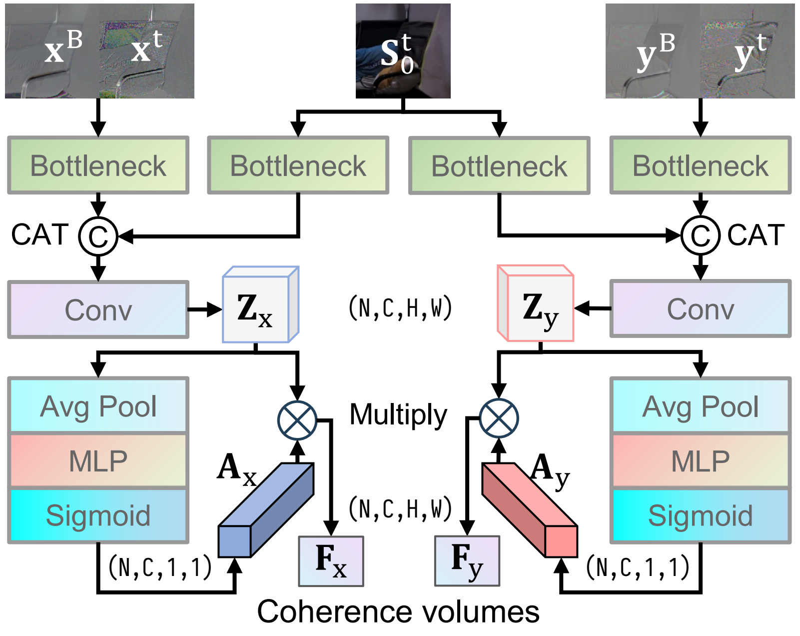
- Learn the residual between $(\mathbf{x}^{\mathrm{t}}, \mathbf{y}^{\mathrm{t}})$ and $(\mathbf{x}', \mathbf{y}')$ with the help of $(\mathbf{x}^{\mathrm{B}}, \mathbf{y}^{\mathrm{B}})$ and $\mathbf{S}_0^{\mathrm{t}}$

$$\mathbf{x}^{\mathrm{t}} = \frac{\mathbf{S}_1^{\mathrm{L}}}{\mathbf{S}_0^{\mathrm{t}}}, \mathbf{y}^{\mathrm{t}} = \frac{\mathbf{S}_2^{\mathrm{L}}}{\mathbf{S}_0^{\mathrm{t}}}$$

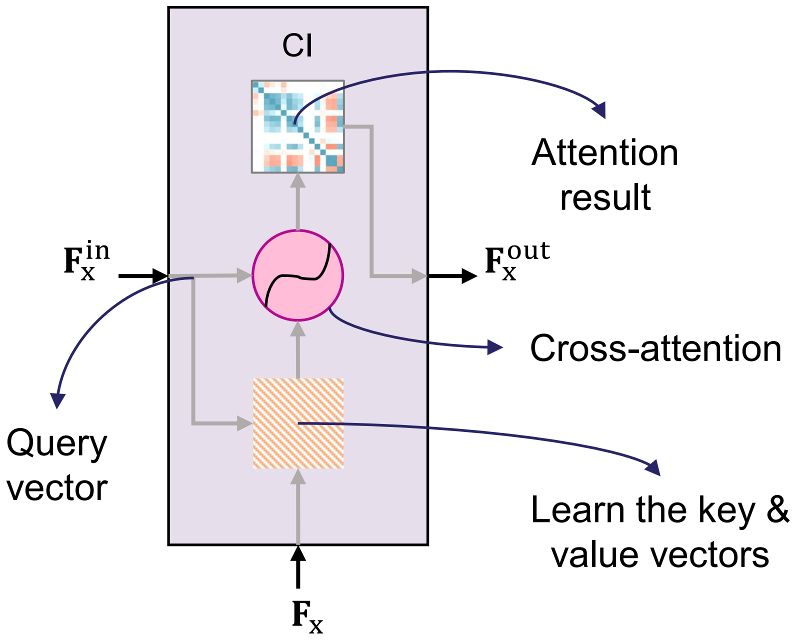
$$\mathbf{x}^{\mathrm{B}} = \frac{\mathbf{S}_1^{\mathrm{B}}}{\mathbf{S}_0^{\mathrm{B}}}, \mathbf{y}^{\mathrm{B}} = \frac{\mathbf{S}_2^{\mathrm{B}}}{\mathbf{S}_0^{\mathrm{B}}}$$

CAG & CI: Modules about coherence

- CAG: aggregate the coherence between the polarization properties and the irradiance information
- CI: inject the coherence into the reconstruction of $\mathbf{x}^{\mathrm{t}}$ and $\mathbf{y}^{\mathrm{t}}$



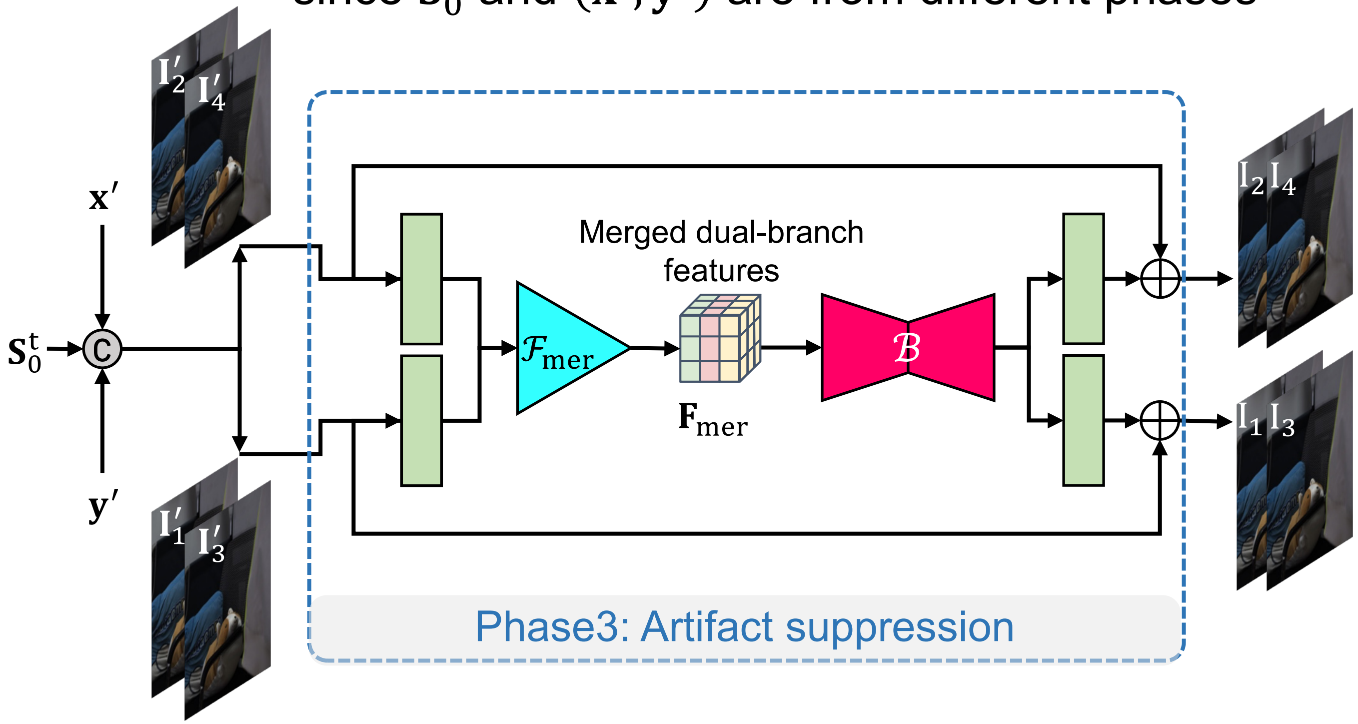
CAG (coherence-aware aggregation) module



CI (coherence injection) module

Phase3: Artifact suppression

- Goal: increasing the quality of details
  - by suppressing the artifacts in the image domain to obtain $\mathbf{I}_{\alpha_{1,2,3,4}}$
  - After Phase2, the quality of the coarse values of the polarized images $\mathbf{I}'_{\alpha_{1,2,3,4}}$ is still not satisfying since $\mathbf{S}_0^{\mathrm{t}}$ and $(\mathbf{x}', \mathbf{y}')$ are from different phases



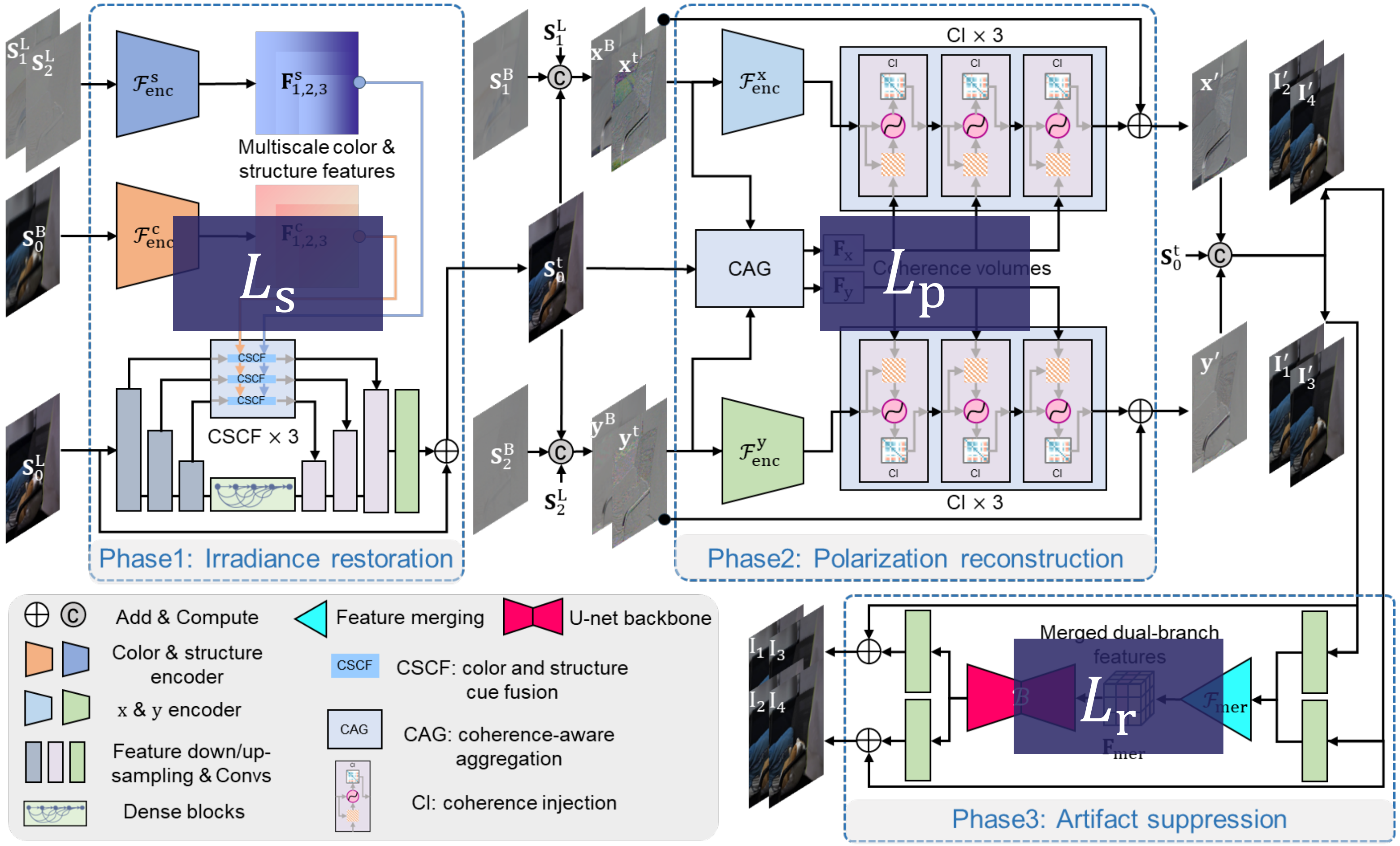
- Solution: add a refinement phase
  - Divide $\mathbf{I}'_{\alpha_{1,2,3,4}}$ into two groups ($\mathbf{I}'_{\alpha_{1,3}}$ & $\mathbf{I}'_{\alpha_{2,4}}$) since $\mathbf{S}_0 = \mathbf{I}_{\alpha_1} + \mathbf{I}_{\alpha_3} = \mathbf{I}_{\alpha_2} + \mathbf{I}_{\alpha_4}$

    → both groups contain the full irradiance information

    → each group contains half of the polarization properties

Loss function

- The total loss function can be written as $L = L_s + L_p + L_r$



- L_s: irradiance loss
  - Phase1
- L_p: polarization loss
  - Phase2
- L_r: refinement loss
  - Phase3

Loss function

- Irradiance loss: $L_{\mathrm{s}} = \lambda_{\mathrm{s}}^{\mathrm{a}} L_1\left(\mathbf{S}_0^{\mathrm{t}}, \mathbf{S}_0^{\mathrm{gt}}\right) + \lambda_{\mathrm{s}}^{\mathrm{b}} L_{\mathrm{perc}}\left(\mathbf{S}_0^{\mathrm{t}}, \mathbf{S}_0^{\mathrm{gt}}\right)$
  - L_1: ℓ_1 loss
  - L_{perc}: perceptual loss
    - $L_{\mathrm{perc}}\left(\mathbf{S}_0^{\mathrm{t}}, \mathbf{S}_0^{\mathrm{gt}}\right) = L_2\left(\phi_h\left(\mathbf{S}_0^{\mathrm{t}}\right), \phi_h\left(\mathbf{S}_0^{\mathrm{gt}}\right)\right)$
      - L_2: ℓ_2 loss
      - ϕ_h: the feature map from h-th layer of VGG-19 network pretrained on ImageNet

- Polarization loss: $L_{\mathrm{p}} = \lambda_{\mathrm{p}}^{\mathrm{a}}\left(L_1(\mathbf{x}', \mathbf{x}^{\mathrm{gt}}) + L_1(\mathbf{y}', \mathbf{y}^{\mathrm{gt}})\right) + \lambda_{\mathrm{p}}^{\mathrm{b}}\left(L_{\mathrm{tv}}(\mathbf{x}') + L_{\mathrm{tv}}(\mathbf{y}')\right) + \lambda_{\mathrm{p}}^{\mathrm{c}} L_{\mathrm{pol}}^1(\mathbf{x}', \mathbf{x}^{\mathrm{gt}}, \mathbf{y}', \mathbf{y}^{\mathrm{gt}})$
  - L_{tv}: total variation loss
  - L_{pol}^1: a polarization-based regularization term to ensure the ratio between $\mathbf{x}'$ and $\mathbf{y}'$
    - $L_{\mathrm{pol}}^1 = L_2(\mathbf{x}' \odot \mathbf{y}^{\mathrm{gt}}, \mathbf{y}' \odot \mathbf{x}^{\mathrm{gt}})$

- Refinement loss: $L_{\mathrm{r}} = \lambda_{\mathrm{r}}^{\mathrm{a}} L_1\left(\mathbf{I}_{\alpha_{1,2,3,4}}, \mathbf{I}_{\alpha_{1,2,3,4}}^{\mathrm{gt}}\right) + \lambda_{\mathrm{r}}^{\mathrm{b}} L_{\mathrm{pol}}^2\left(\mathbf{I}_{\alpha_{1,2,3,4}}\right)$
  - L_{pol}^2: another polarization-based regularization term
    - $L_{\mathrm{pol}}^2\left(\mathbf{I}_{\alpha_{1,2,3,4}}\right) = L_2\left(\mathbf{I}_{\alpha_1} + \mathbf{I}_{\alpha_3}, \mathbf{I}_{\alpha_2} + \mathbf{I}_{\alpha_4}\right)$

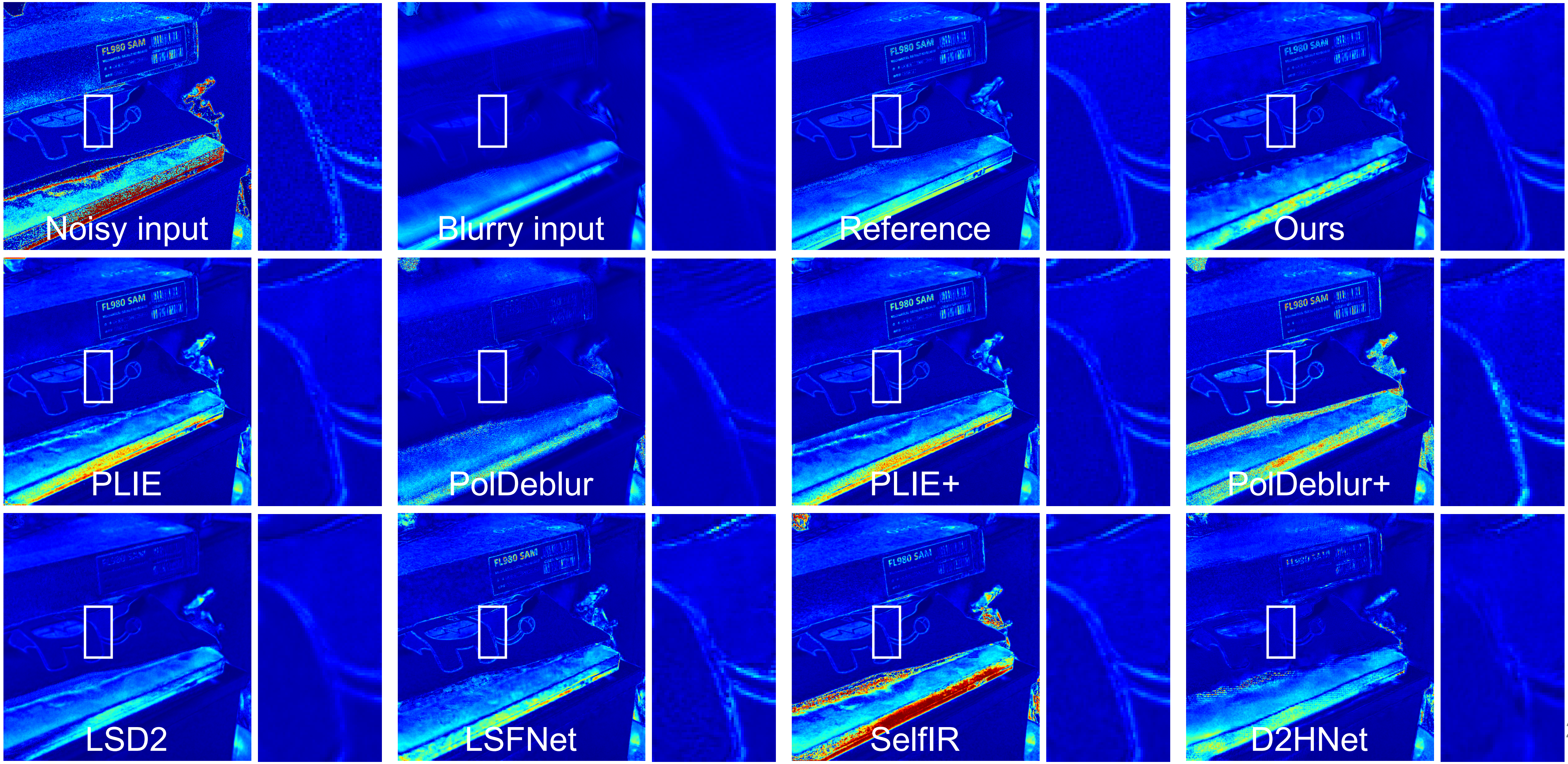
Quantitative evaluation on synthetic data

	PSNR-$\mathbf{p}$	SSIM-$\mathbf{p}$	PSNR-$\boldsymbol{\theta}$	SSIM-$\boldsymbol{\theta}$	PSNR-$\boldsymbol{S}_0$	SSIM-$\boldsymbol{S}_0$
Ours	**29.23**	**0.797**	**16.96**	**0.382**	**39.05**	**0.982**
PLIE	27.91	0.790	15.92	0.371	38.95	0.978
PLIE+	27.98	0.794	16.93	0.379	39.01	0.979
PolDeblur	24.52	0.676	15.73	0.280	26.12	0.794
PolDeblur+	25.31	0.758	16.75	0.374	39.04	0.981
LSD2	25.73	0.662	13.75	0.288	27.88	0.905
LSFNet	25.56	0.693	15.90	0.282	26.76	0.826
SelfIR	19.43	0.647	15.39	0.231	25.90	0.785
D2HNet	24.45	0.671	15.63	0.264	25.25	0.803

- The state-of-the-art polarized image low-light enhancement method and its improved version
  - **PLIE** & **PLIE+** [Zhou *et al*., AAAI 23]
- The state-of-the-art polarized image deblurring method and its improved version
  - **PolDeblur** & **PolDeblur+** [Zhou *et al*., Arxiv 24]
- Four learning-based image enhancement methods designed for conventional images that also fuse noisy and blurry pairs
  - LSD2 [Zhao *et al*., BMVC 20]
  - LSFNet [Chang *et al*., TMM 21]
  - SelfIR [Zhang *et al*., NeurIPS 22]
  - D2HNet [Zhao *et al*., ECCV 22]

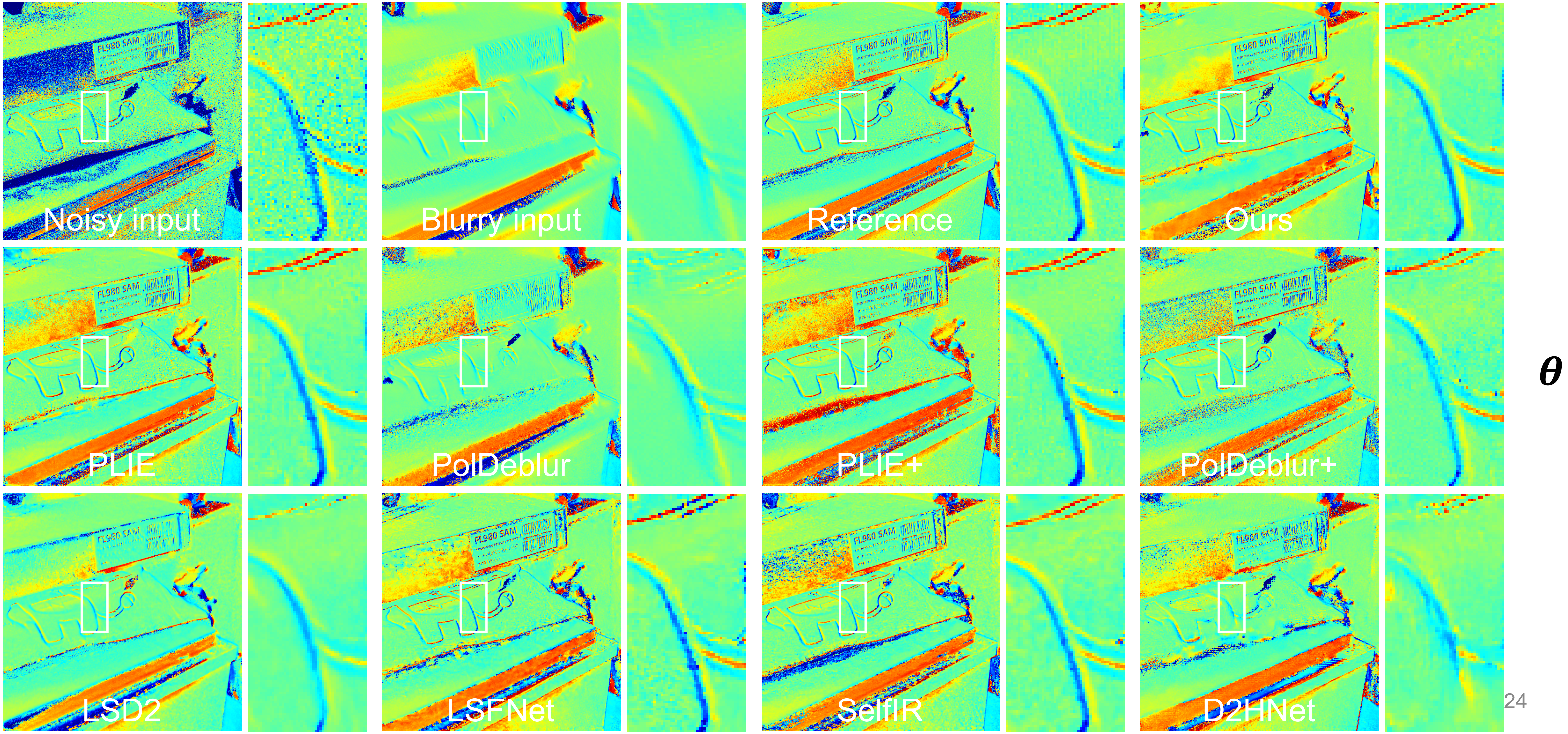
* The dataset is generated from the PLIE dataset [Zhou *et al*., AAAI 23]
* An "improved version" refers to making slight modifications to the original version in order to enable it to accept noisy and blurry input pairs.

Qualitative evaluation on synthetic data

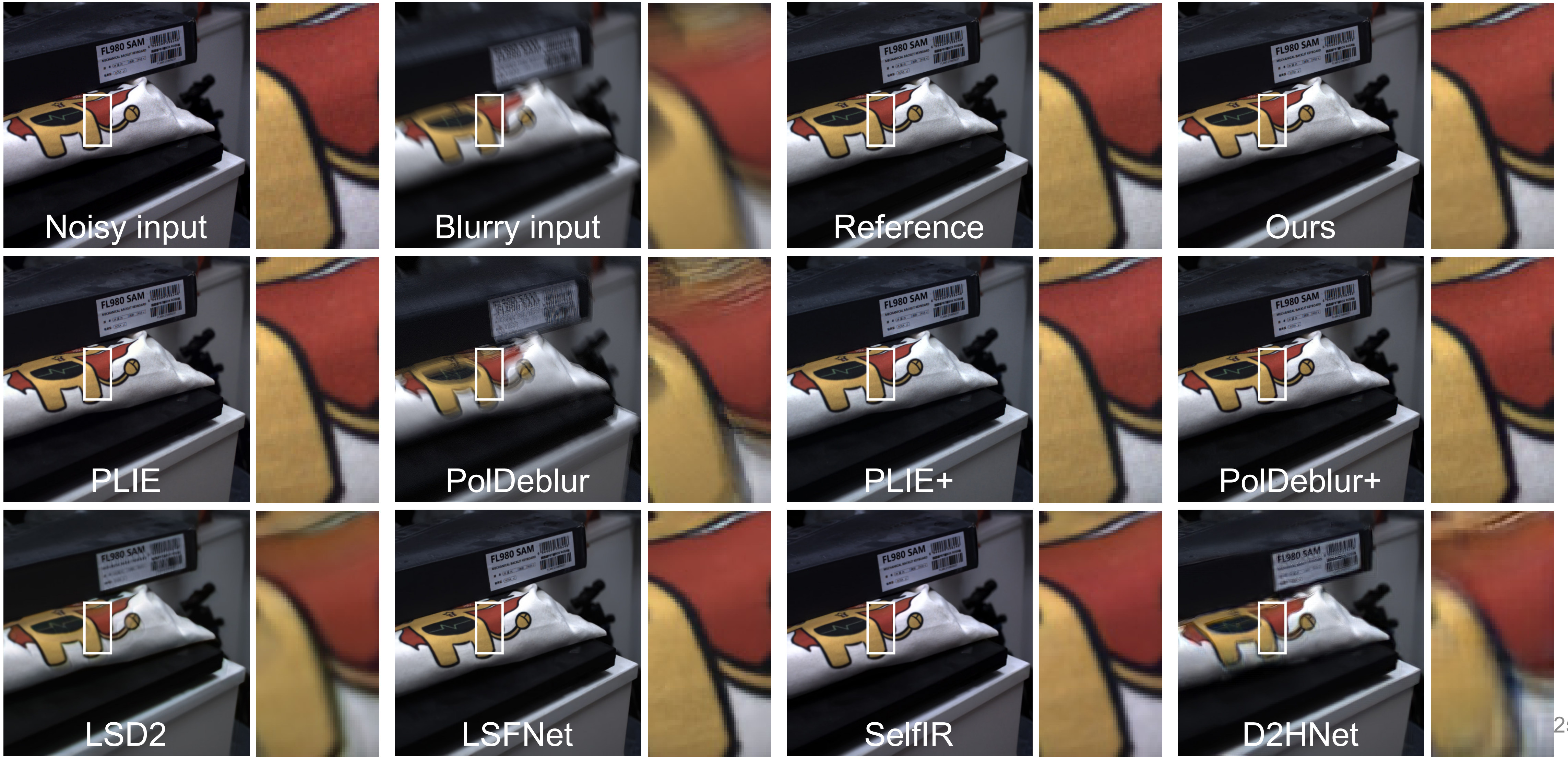


p

Qualitative evaluation on synthetic data

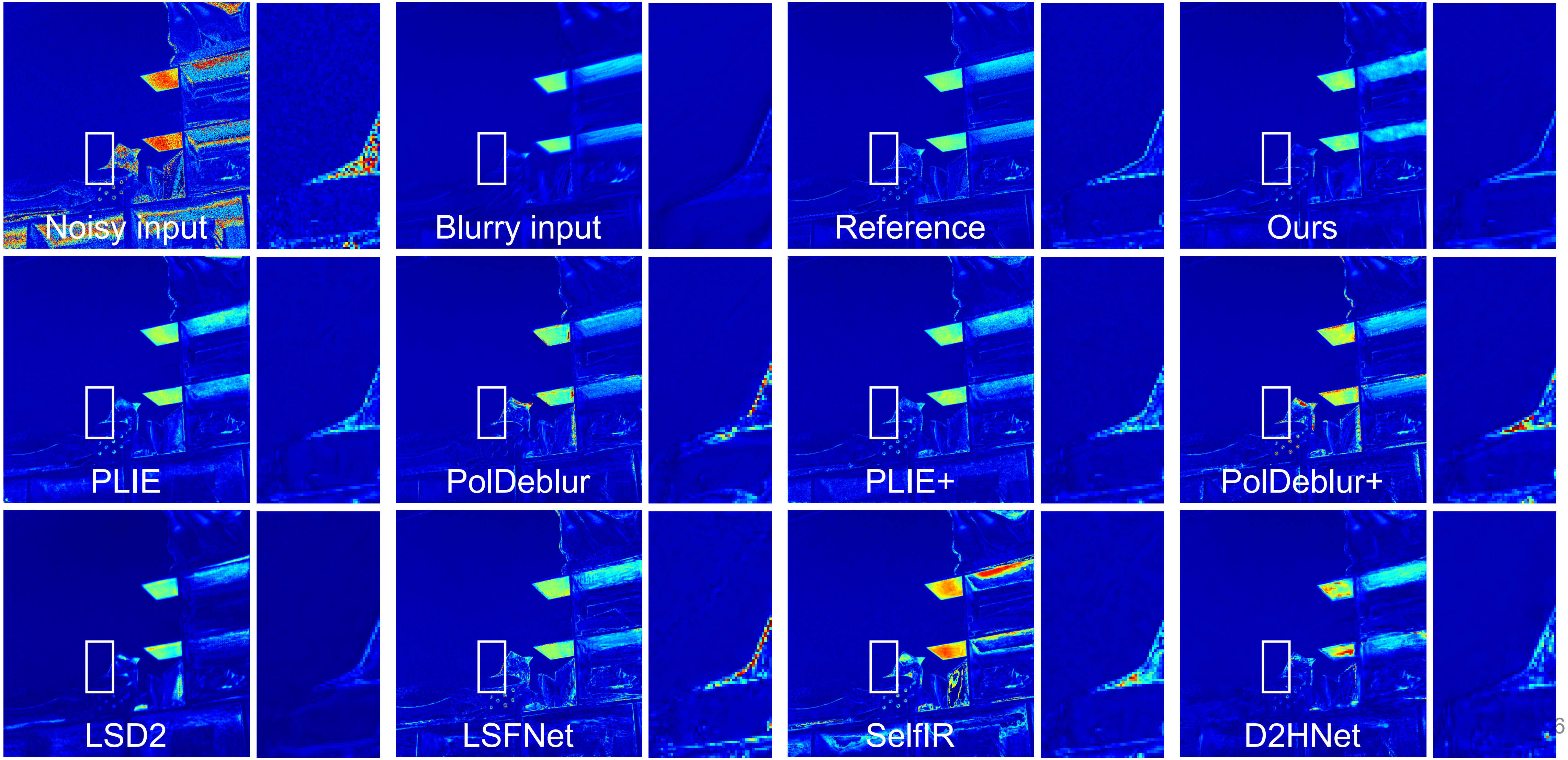


Qualitative evaluation on synthetic data



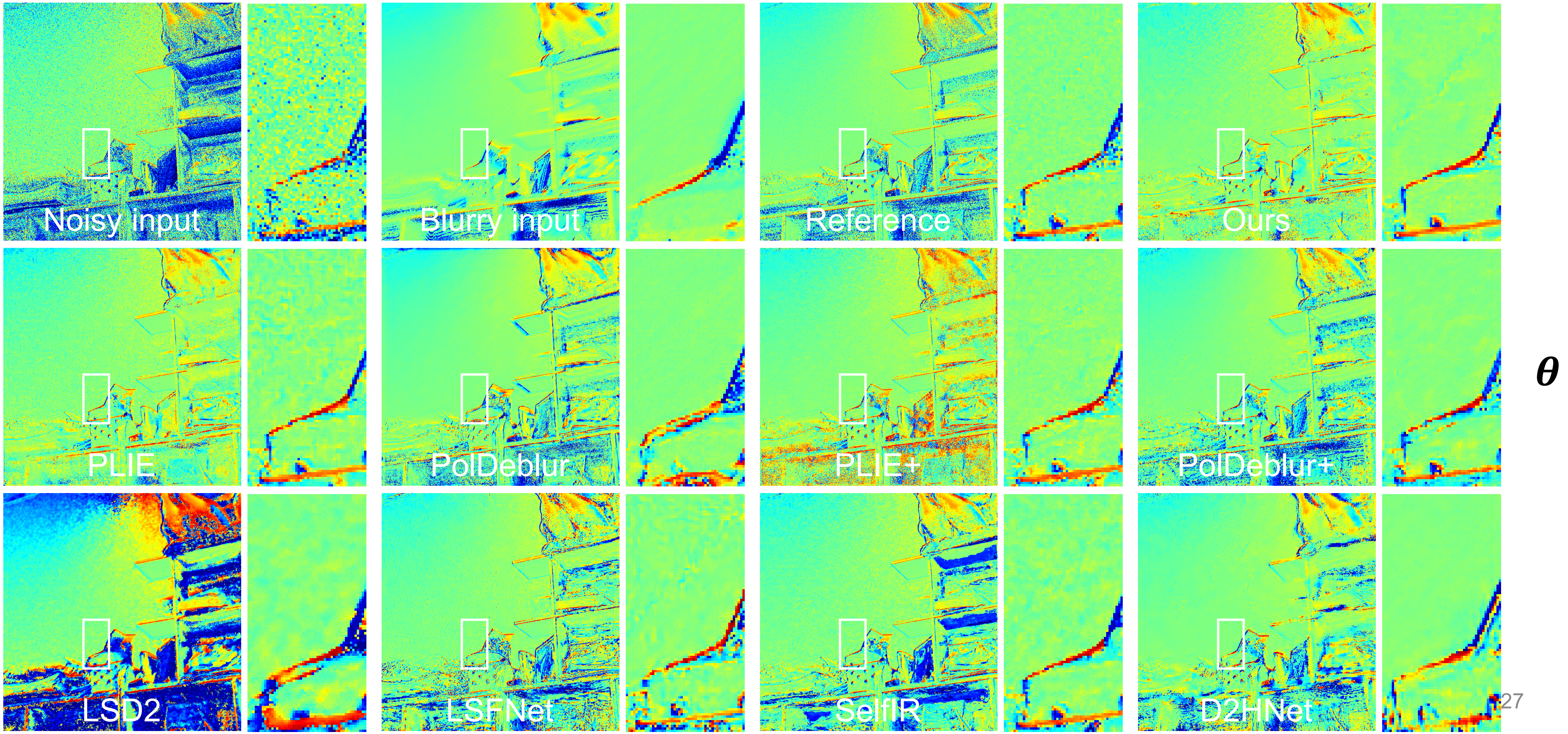
$\boldsymbol{s}_0$

Qualitative evaluation on real data

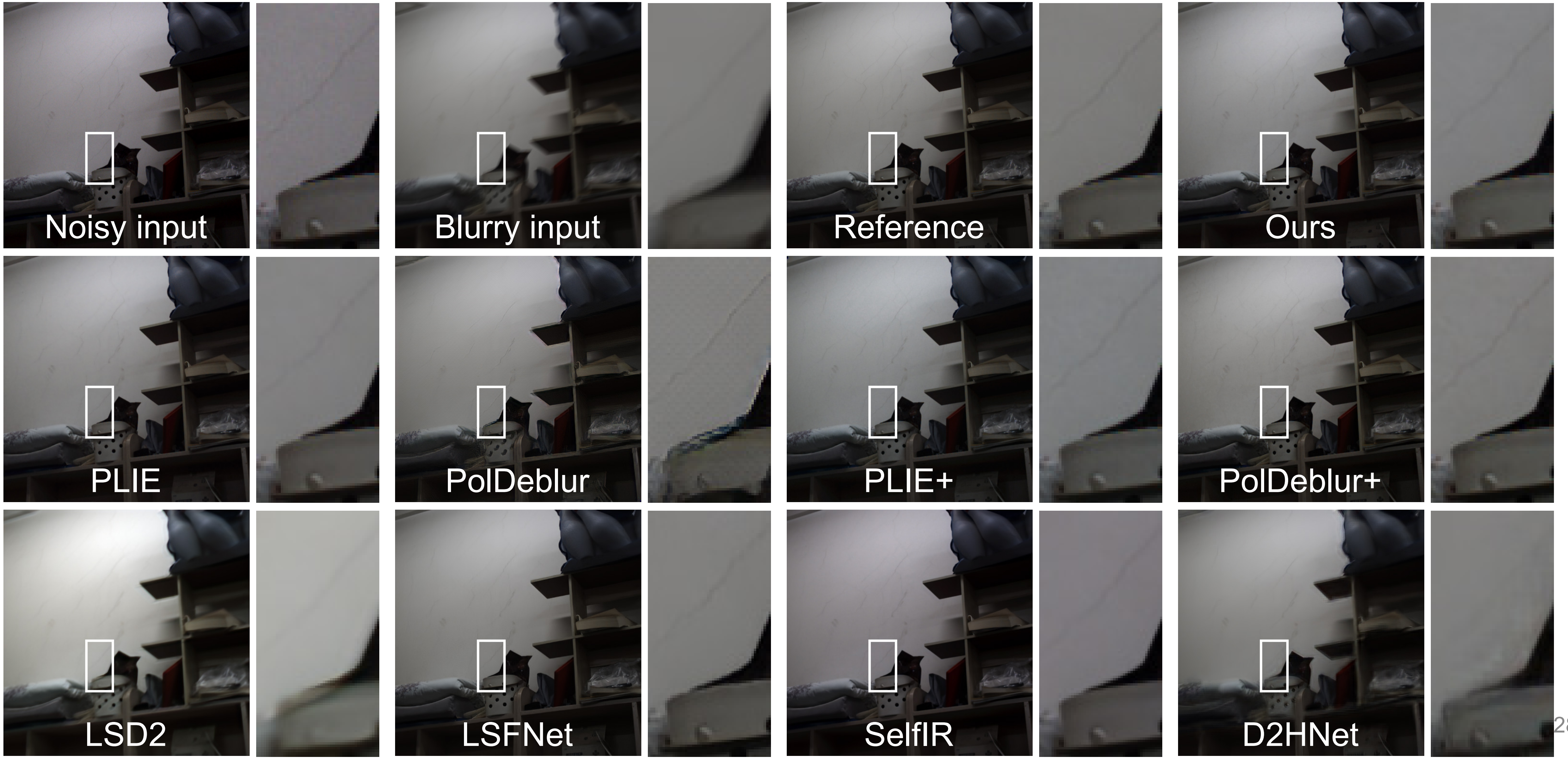


p

Qualitative evaluation on real data



Qualitative evaluation on real data



$\boldsymbol{s}_0$

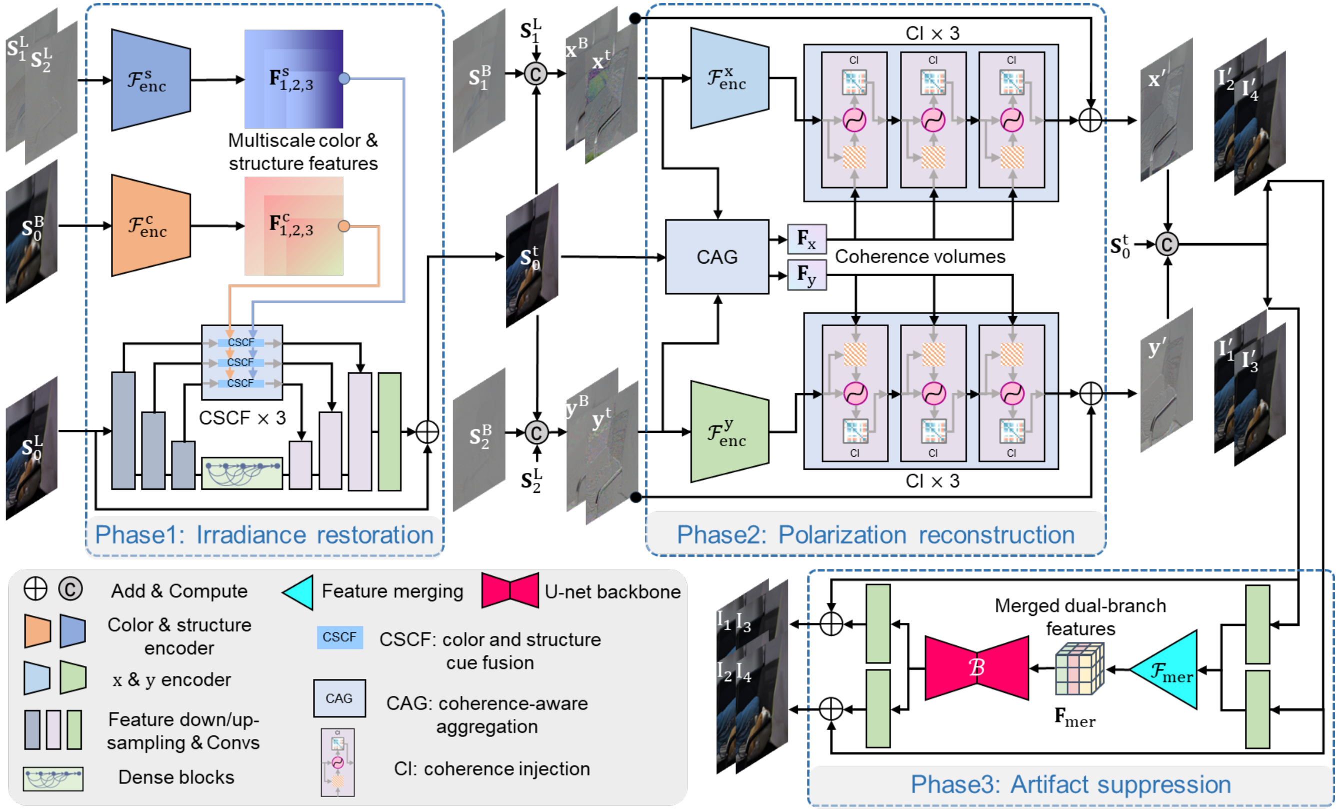
Application: reflection removal

Reflection-contaminated input

Noisy | Blurry | Fused by our framework

Reflection-removed output

Conclusion



- A quality-improved and property-preserved polarimetric imaging framework
  - by complementarily fusing a degraded pair of noisy and blurry polarized snapshots
- A neural network-based three-phase fusing scheme
  - fully utilizing the complementary knowledge from the noisy and blurry pairs in a polarization-aware manner
- Specially-designed modules tailored to each phase
  - effectively exploring the usage of different physical quantities to improve the overall performance

My GitHub: https://github.com/fourson
My homepage: https://fourson.github.io/

Thank you!

Q & A

Chu Zhou (周 矗，シュウ チク)
National Institute of Informatics, Japan
zhou_chu@hotmail.com