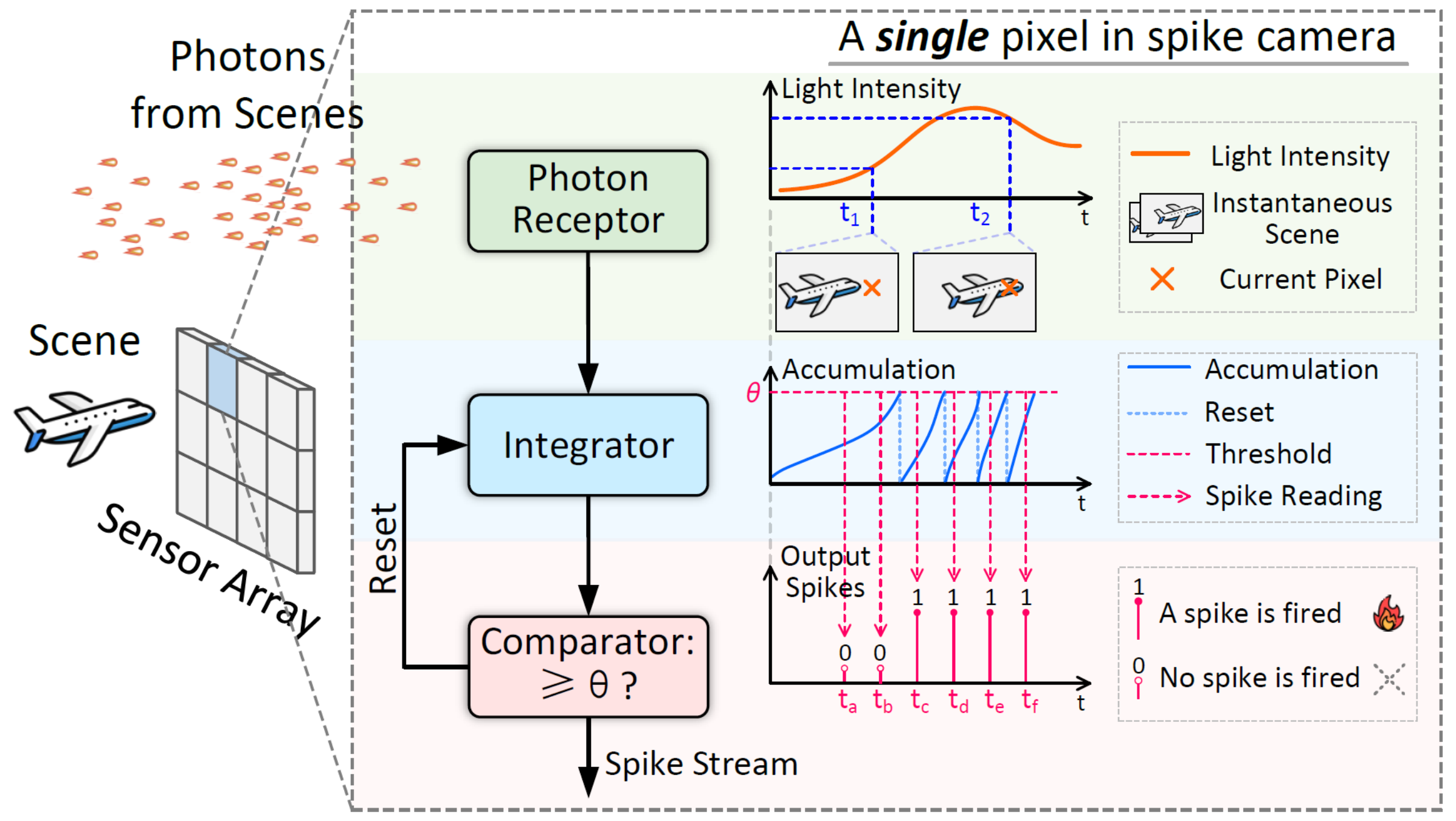
Boosting Spike Camera Image Reconstruction from a Perspective of Dealing with Spike Fluctuations

Rui Zhao[1,2] Ruiqin Xiong[1,2 *] Jing Zhao[1,2] Jian Zhang[3] Xiaopeng Fan[4] Zhaofei Yu[1] Tiejun Huang[1,2]
[1]SCS PKU [2]NKLMIP PKU [3]SECE PKU [4]SCST HIT
PEKING UNIVERSITY
Personal Homepage

1. Introduction

1.1 Spike Camera

Each pixel of the spike camera comprises three main components: a photon receptor, an integrator, and a comparator. The incoming photons are captured by the photon receptor and accumulated by the integrator. Whenever the number of accumulated photons reaches a predefined threshold θ, a spike is fired, and the integrator is reset.



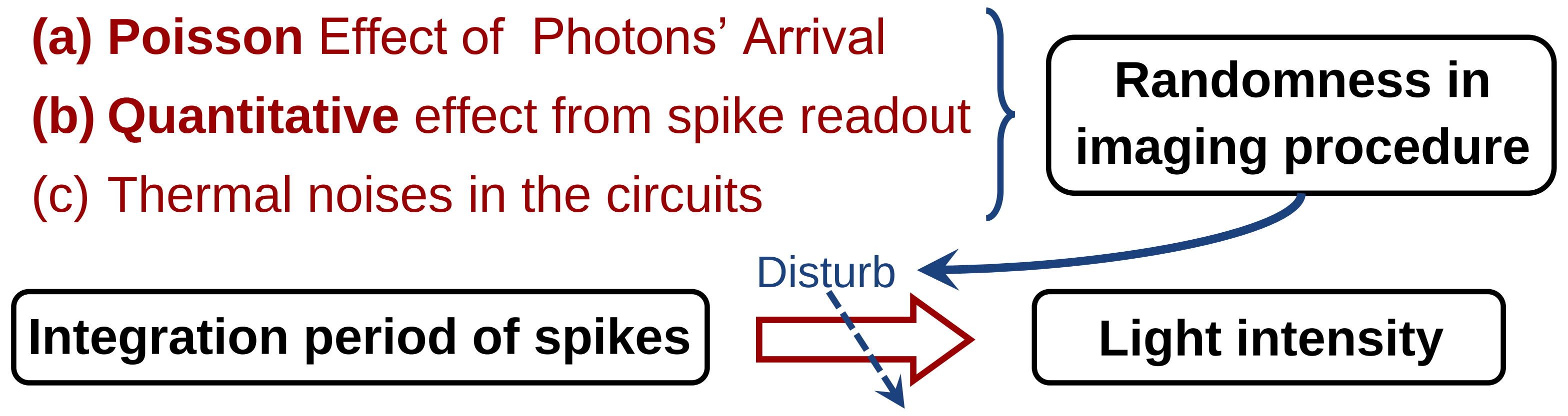
Key components of a pixel in spike camera

$$\mathbf{A}(\mathbf{x},t)=\int_0^t \alpha\mathcal{P}\left(\mathbf{L}(\mathbf{x},\tau)\right)\mathrm{d}\tau \mod \theta$$

α is the quantum conversion coefficient of photons. $\mathcal{P}$ means Poisson sampling. $\mathbf{L}(\mathbf{x},\tau)$ is the expected number of arrival photons at a pixel area per unit time.

1.2 Fluctuations in Spikes

Effects in the imaging of spike cameras.



- **(a) Poisson** Effect of Photons' Arrival
- **(b) Quantitative** effect from spike readout
- (c) Thermal noises in the circuits

Randomness in imaging procedure

Disturb

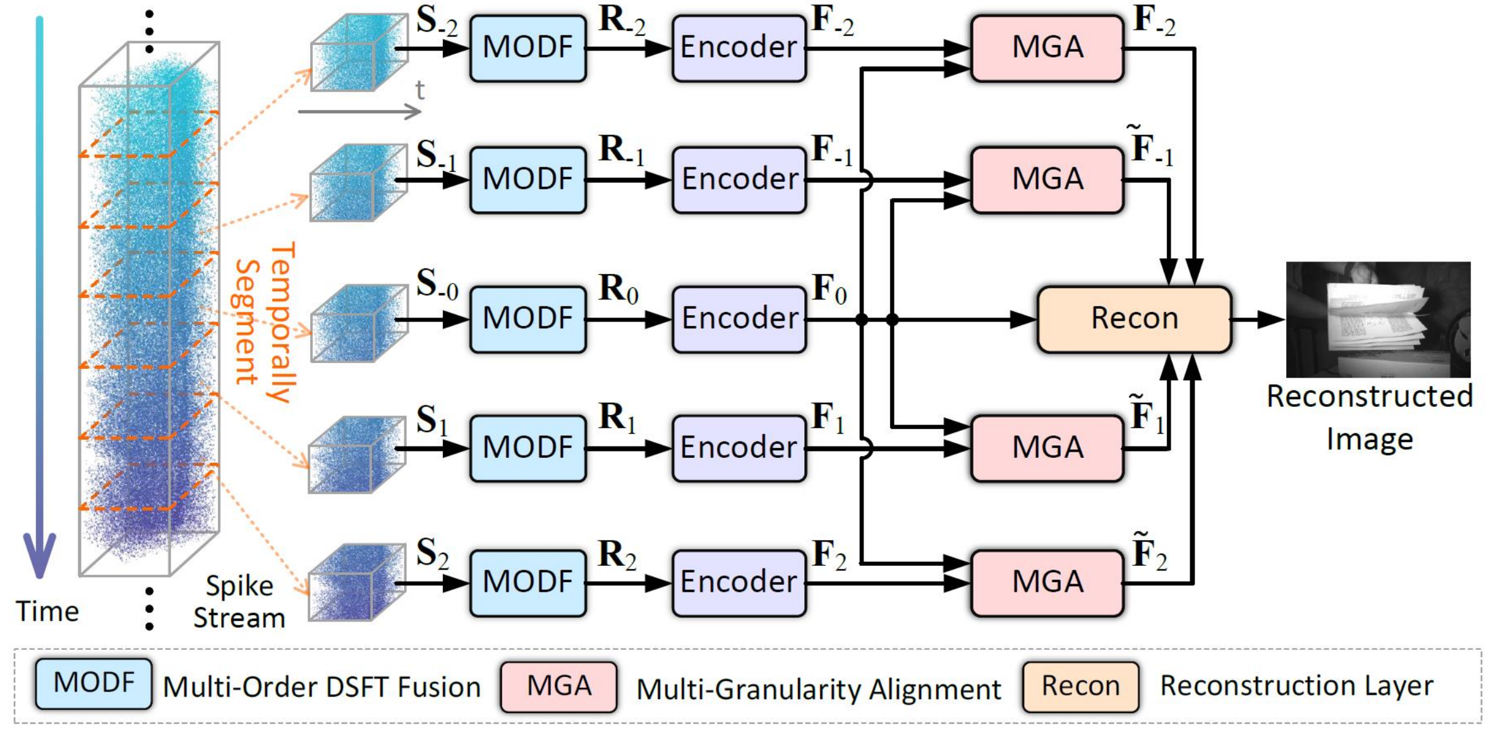
Integration period of spikes

Light intensity

Fluctuations: even when the light intensity is constant, the integration period of each spike changes over time.

2. Methods

2.1 Overall Architecture of the proposed method (BSF)



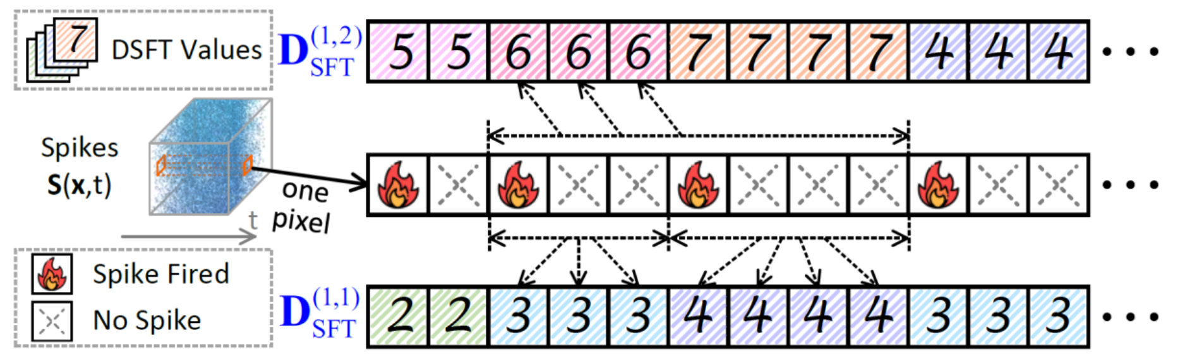
2. Methods

2.2 Multi-Order DSFT Fusion (MODF)

Proposition 1: Use DSFT as input.

DSFT: Differential of Spike Firing Time.

(1,1)-order DSFT, i.e., $\mathbf{D}_{SFT}^{(1,1)}$ ☞



Proposition 2: Processing DSFT in reciprocal domain.

Theorem: When the photons' arrival is constant

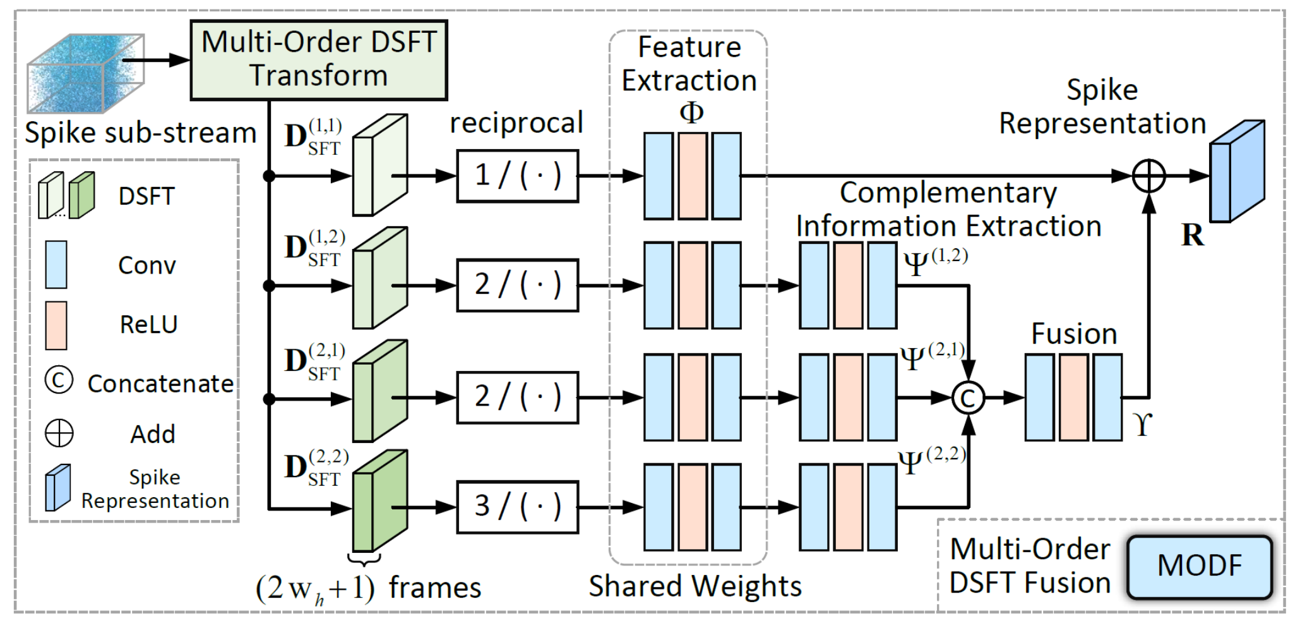
$$\begin{cases}\Pr\left\{\mathbf{D}_{SFT}^{(1,1)}=\lfloor\theta/L\rfloor\right\}=p_1=\left(\lceil\theta/L\rceil-\theta/L\right)\cdot\dfrac{\lfloor\theta/L\rfloor}{\theta/L}\\ \Pr\left\{\mathbf{D}_{SFT}^{(1,1)}=\lceil\theta/L\rceil\right\}=p_2=\left(\theta/L-\lfloor\theta/L\rfloor\right)\cdot\dfrac{\lceil\theta/L\rceil}{\theta/L}\end{cases} \Longrightarrow \mathbb{E}\left(\frac{1}{\mathbf{D}_{\mathrm{SFT}}^{(1,1)}}\right)=\frac{1}{\lfloor\theta/L\rfloor}\cdot p_1+\frac{1}{\lceil\theta/L\rceil}\cdot p_2=\frac{L}{\theta}$$

The reciprocal of DSFT corresponds to **unbiased** estimation of light intensity

Proposition 3: Fusing DSFT with multiple orders.

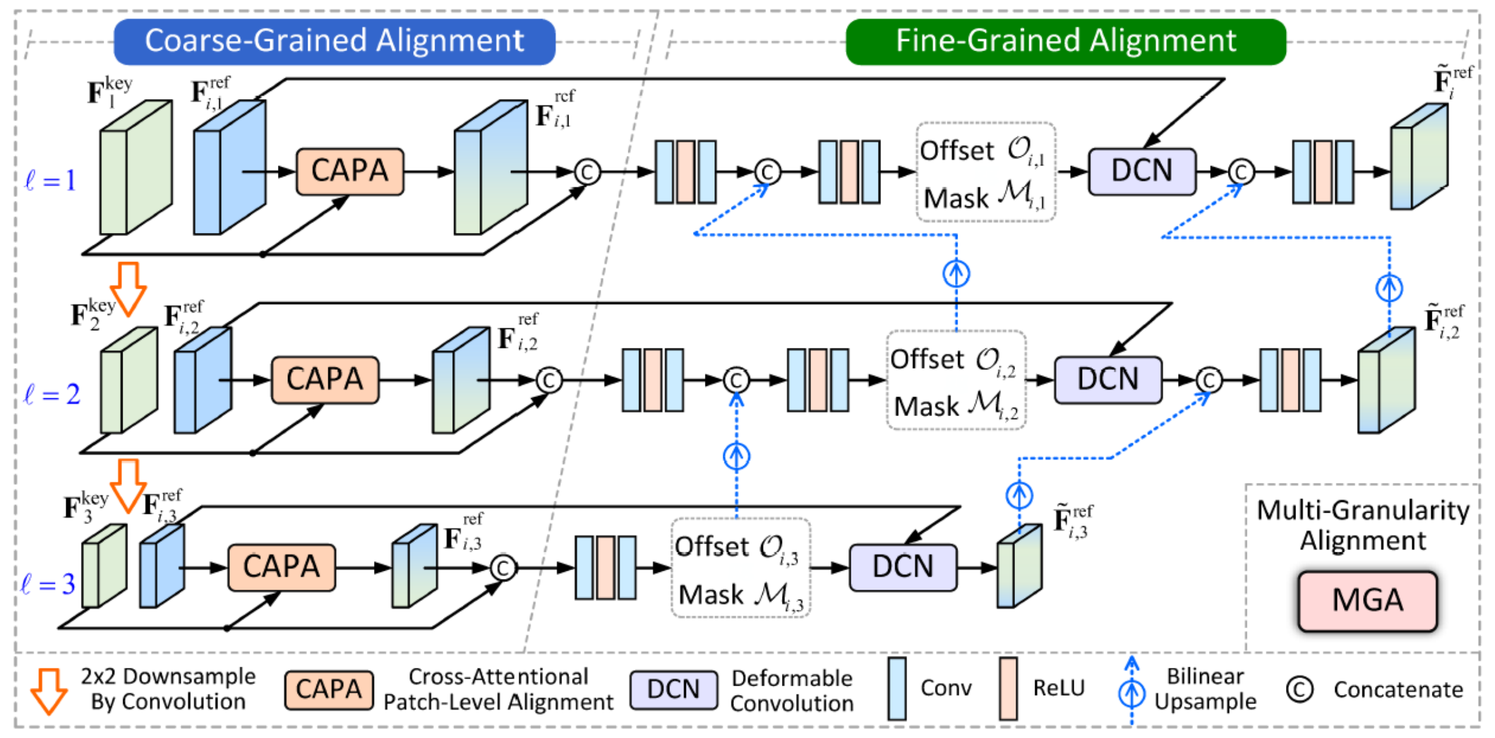
For reducing Poisson effects and light changes: purchase a more stable DSFT.

$$\mathbf{D}_{\mathrm{SFT}}^{(n_1,n_2)}(\mathbf{x},t)=\mathbf{T}_{\mathrm{next}}^{(n_2)}(\mathbf{x},t)-\mathbf{T}_{\mathrm{prev}}^{(n_1)}(\mathbf{x},t)$$
$$=\min\left\{\tau\;\middle|\;\sum\nolimits_{k=t+1}^{\tau}\mathbf{S}(\mathbf{x},k)=n_2,\ \tau>t\right\}$$
$$-\max\left\{\tau\;\middle|\;\sum\nolimits_{k=\tau}^{t}\mathbf{S}(\mathbf{x},k)=n_1,\ \tau\le t\right\},$$



Structure of the multi-order DSFT fusion module based on the 3 propositions ☞

2.3 Multi-Granularity Alignment (MGA)



☜ Structure of the MGA module

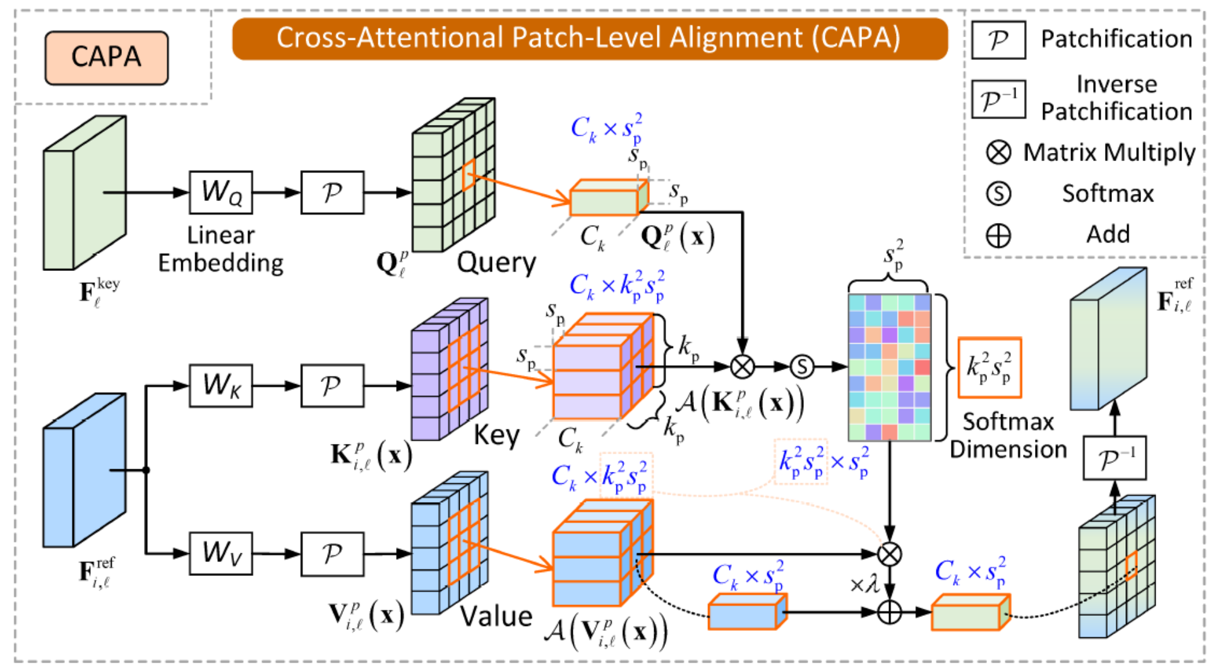
- Pyramidal structure
- At each pyramidal level

  First coarse-grained alignment

  Then fine-grained alignment

(A) Coarse Grained Alignment: (CAPA) Cross-Attentional Patch-level Attention

$$\mathbf{Q}_\ell^{\mathrm{p}}=\mathcal{Z}[\mathbf{Q}_\ell]=\mathcal{Z}[W_Q\mathbf{F}_\ell^{\mathrm{key}}]$$
$$\mathbf{K}_{i,\ell}^{\mathrm{p}}=\mathcal{Z}[\mathbf{K}_{i,\ell}]=\mathcal{Z}[W_K\mathbf{F}_{i,\ell}^{\mathrm{ref}}]$$
$$\mathbf{V}_{i,\ell}^{\mathrm{p}}=\mathcal{Z}[\mathbf{V}_{i,\ell}]=\mathcal{Z}[W_V\mathbf{F}_{i,\ell}^{\mathrm{ref}}]$$
$$\widehat{\mathbf{V}}_{i,\ell}^{\mathrm{p}}(\mathbf{x})=\mathcal{A}(\mathbf{V}_{i,\ell}^{\mathrm{p}}(\mathbf{x}))\sigma\left(\frac{(\mathbf{Q}_\ell^{\mathrm{p}})^\top(\mathbf{x})\,\mathcal{A}(\mathbf{K}_{i,\ell}^{\mathrm{p}}(\mathbf{x}))}{\sqrt{C_k}}\right)$$
$$\mathcal{A}(\mathbf{V}_i^{\mathrm{p}}(\mathbf{x}))=\{\mathbf{V}_i^{\mathrm{p}}(\mathbf{x}+\boldsymbol{\delta})\}_{\boldsymbol{\delta}\in\mathcal{N}(\mathbf{x};k_{\mathrm{p}})}$$



(B) Fine Grained Alignment: Deformable Convolution

$$\widetilde{\mathbf{F}}_i^{\mathrm{ref}}(\mathbf{x})=\sum_{\boldsymbol{\delta}\in\mathcal{N}(\mathbf{x};k_{\mathrm{d}})}K(\boldsymbol{\delta})\widehat{\mathbf{F}}_i^{\mathrm{ref}}\big(\mathbf{x}+\boldsymbol{\delta}+\mathcal{O}_i(\mathbf{x},\boldsymbol{\delta})\big)\mathcal{M}_i(\mathbf{x},\boldsymbol{\delta})$$

3. Experiments

3.1 Quantitative Results on REDS-SCIR Dataset

(1) Reference-based metrics

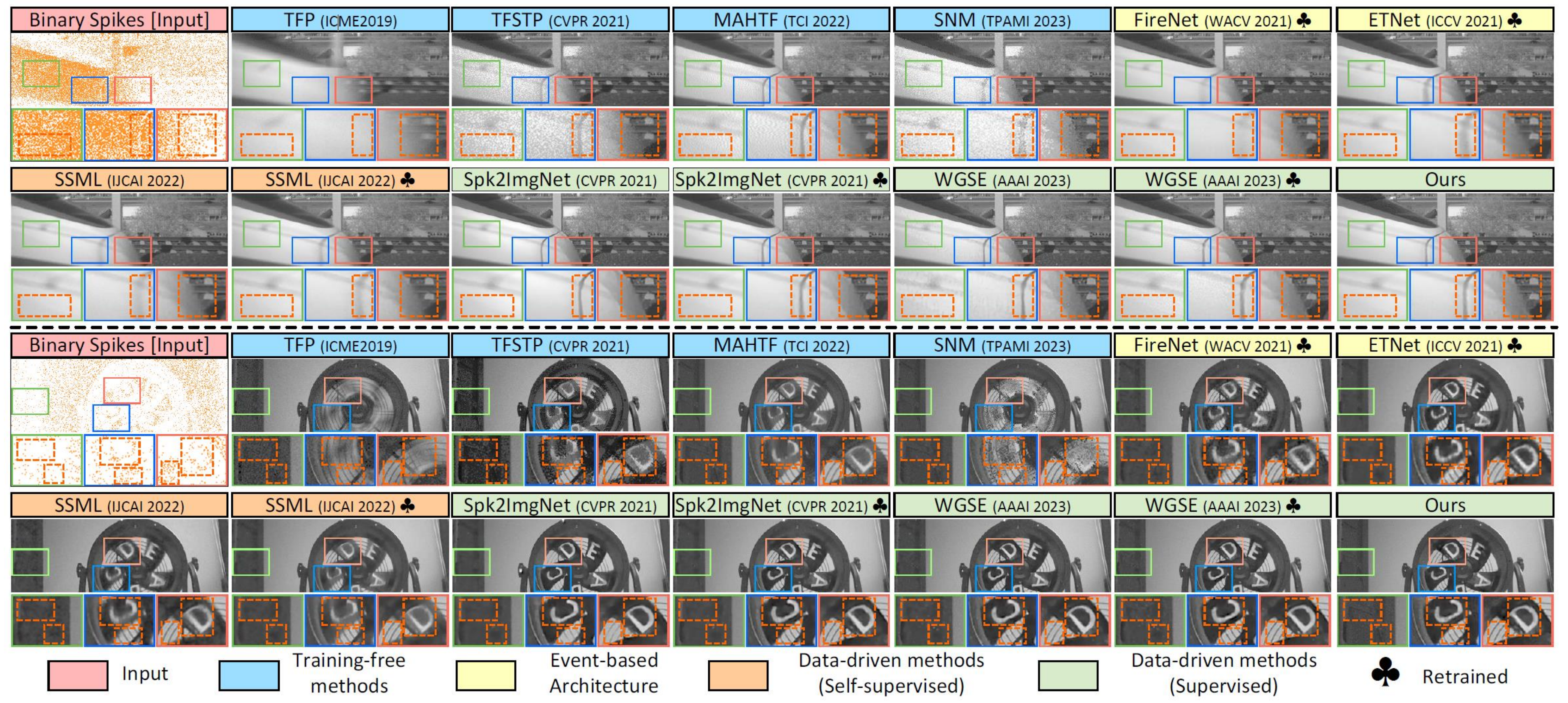
Part	Method	$\eta=1.00$			$\eta=0.75$			$\eta=0.50$			Params (M)
		PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	
(A)	TFP [73]	27.27	0.711	0.265	26.73	0.669	0.300	25.62	0.581	0.370	—
	TFI [73]	23.55	0.634	0.329	24.77	0.673	0.293	26.77	0.713	0.249	—
	TFSTP [69]	20.35	0.678	0.270	19.62	0.685	0.252	21.10	0.707	0.247	—
	MAHTF [65]	29.57	0.879	0.112	30.07	0.884	0.113	29.65	0.869	0.136	—
(B)	FireNet [39] ♣	34.38	0.922	0.077	33.87	0.911	0.084	32.62	0.884	0.105	0.038
	ETNet [46] ♣	33.24	0.918	0.082	32.85	0.909	0.089	31.96	0.889	0.109	22.179
(C)	SSML [5]	32.60	0.920	0.088	32.09	0.907	0.097	31.00	0.879	0.122	2.385
	SSML [5] ♣	33.94	0.923	0.075	33.27	0.909	0.088	32.01	0.883	0.116	2.385
(D)	Spk2ImgNet [64]	35.21	0.953	0.036	34.70	0.945	0.044	33.75	0.926	0.064	3.904
	Spk2ImgNet [64] ♣	39.16	0.966	0.024	38.27	0.958	0.032	36.59	0.940	0.051	3.904
	WGSE [58]	35.21	0.950	0.039	34.98	0.947	0.042	34.11	0.931	0.057	3.806
	WGSE [58] ♣	38.97	0.964	0.027	38.23	0.957	0.034	36.75	0.940	0.049	3.806
	BSF (Ours)	**39.76**	**0.970**	**0.021**	**39.09**	**0.964**	**0.027**	**37.76**	**0.951**	**0.040**	2.477

(2) Non-reference-based metrics

Part	Method	BRISQUE ↓	PIQE ↓	HOSA ↓
(A)	TFP [73]	37.502	45.956	35.436
	TFI [73]	37.708	45.148	30.892
	TFSTP [69]	37.585	38.714	29.173
	SNM [75]	32.089	41.927	29.334
	MAHTF [65]	30.910	30.068	26.757
(B)	FireNet [39] ♣	25.545	25.076	35.305
	ETNet [46] ♣	33.403	46.682	36.482

Part	Method	BRISQUE ↓	PIQE ↓	HOSA ↓
(C)	SSML [5]	29.240	25.491	35.399
	SSML [5] ♣	32.234	26.981	35.554
(D)	Spk2ImgNet [64]	29.351	26.745	25.761
	Spk2ImgNet [64] ♣	29.180	39.593	31.287
	WGSE [58]	24.637	27.831	25.657
	WGSE [58] ♣	23.429	30.673	27.434
	BSF (Ours)	**18.529**	**23.477**	**25.523**

3.2 Visualization Results on Real-Captured Data



3.3 Ablation Studies

Ablations on Proposed Modules

Case	MODF	MGA			PSNR ↑		
		DCN	CAPA	Pym	$\eta=1.00$	$\eta=0.75$	$\eta=0.50$
(1)	Spike				38.44	37.69	36.29
(2)	D (1,1)				38.78	38.10	36.80
(3)	✓				38.99	38.32	37.00
(4)	✓	✓		1	38.91	38.26	36.96
(5)	✓	✓	✓	1	39.06	38.39	37.06
(6)	✓	✓		2	39.02	38.39	37.11
(7)	✓	✓	✓	2	39.39	38.73	37.41
(8)	✓	✓		3	39.38	38.77	37.53
(9)	✓	✓	✓	3	**39.76**	**39.09**	**37.76**

Ablations on the number of input frames

N_{IF}	$\eta=1.00$			$\eta=0.75$			$\eta=0.50$		
	P ↑	S ↑	L ↓	P ↑	S ↑	L ↓	P ↑	S ↑	L ↓
21	38.14	0.959	0.032	37.44	0.952	0.039	36.08	0.935	0.055
41	39.35	0.969	**0.021**	38.67	0.963	**0.027**	37.34	0.949	0.041
61	**39.76**	**0.970**	**0.021**	**39.09**	**0.964**	**0.027**	**37.76**	**0.951**	**0.040**
81	39.70	0.969	0.022	39.03	0.963	0.028	37.69	0.950	0.042

Ablations on Hyper-parameters of the CAPA module

MGA			$\underline{s_{\mathrm{p}}}$	k_{p}	PSNR ↑		
Pym	DCN	CAPA			$\eta=1.00$	$\eta=0.75$	$\eta=0.50$
3	✓	✗	✗	✗	39.380	38.768	37.533
3	✓	✓	$\underline{3}$	3	39.759	39.088	37.764
3	✓	✓	5	3	39.704	39.041	37.714
3	✓	✓	7	3	39.693	39.030	37.714
3	✓	✓	9	3	39.700	39.037	37.721

MGA			s_{p}	$\underline{k_{\mathrm{p}}}$	PSNR ↑		
Pym	DCN	CAPA			$\eta=1.00$	$\eta=0.75$	$\eta=0.50$
3	✓	✗	✗	✗	39.380	38.768	37.533
3	✓	✓	3	$\underline{3}$	39.759	39.088	37.764
3	✓	✓	3	5	39.712	39.047	37.726
3	✓	✓	3	7	39.731	39.069	37.749
3	✓	✓	3	9	39.694	39.020	37.685