Retina-like Visual Image Reconstruction via Spiking Neural Model

Lin Zhu [1,2], Siwei Dong [1], Jianing Li [1,2], Tiejun Huang [1,2], Yonghong Tian [1,2*]

[1] National Engineering for Video Technology, School of EE&CS, Peking University, Beijing, China [2] Pengcheng Laboratory, Shenzhen, China

CVPR SEATTLE WASHINGTON JUNE 16-18 2020

Motivation

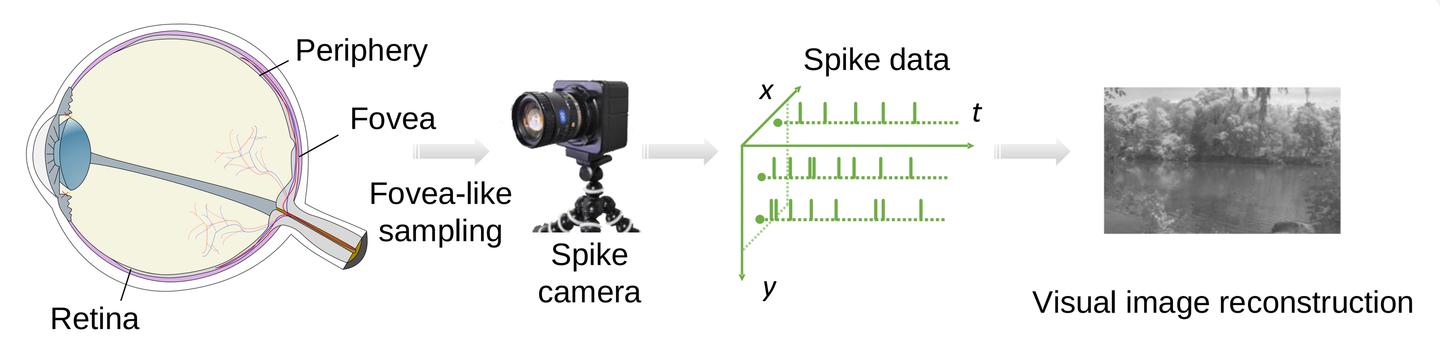


Fig. 1 The spike camera based on fovea-like sampling and visual image reconstruction

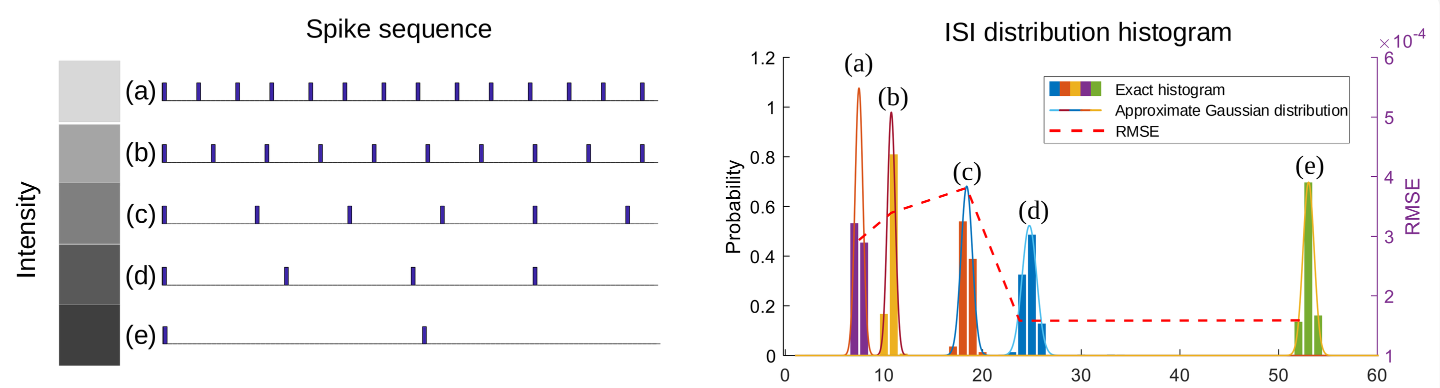


Fig. 2 The spike data distribution under different light intensities

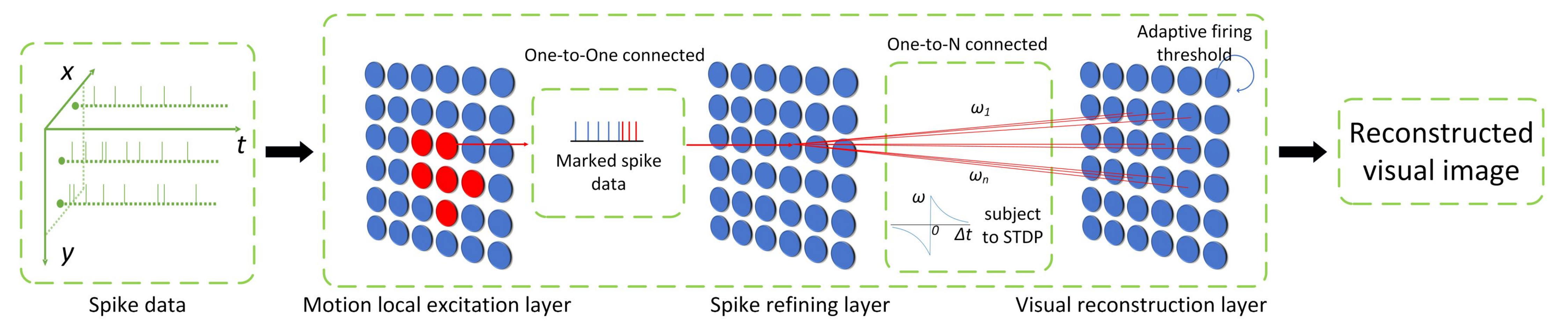
As a novel bio-inspired vision sensor, spike camera mimics the fovea to record the nature scenes by continuous-time spikes instead of frame-based manner. However, recon-structing visual images from the spikes remains to be a challenge. In this paper, we design a retina-like visual image reconstruction framework, which is flexible in reconstructing full texture of natural scenes from the totally new spike data.

The Main Contributions

- We propose a three layer spiking neural model which relies on a combination of biologically plausible mechanisms. Using time-continuous spike data, our method can reconstruct images at any sampling moment, and retain the details of high-speed motion and static background simultaneously.
- We propose a dynamic neuron extraction model to distinguish the neuron states (dynamic or static) in an incremental way, which is useful for reconstructing high quality high-speed motion scenes.
- We construct a new spike dataset for evaluating the reconstruction method, and make these available to the research community.

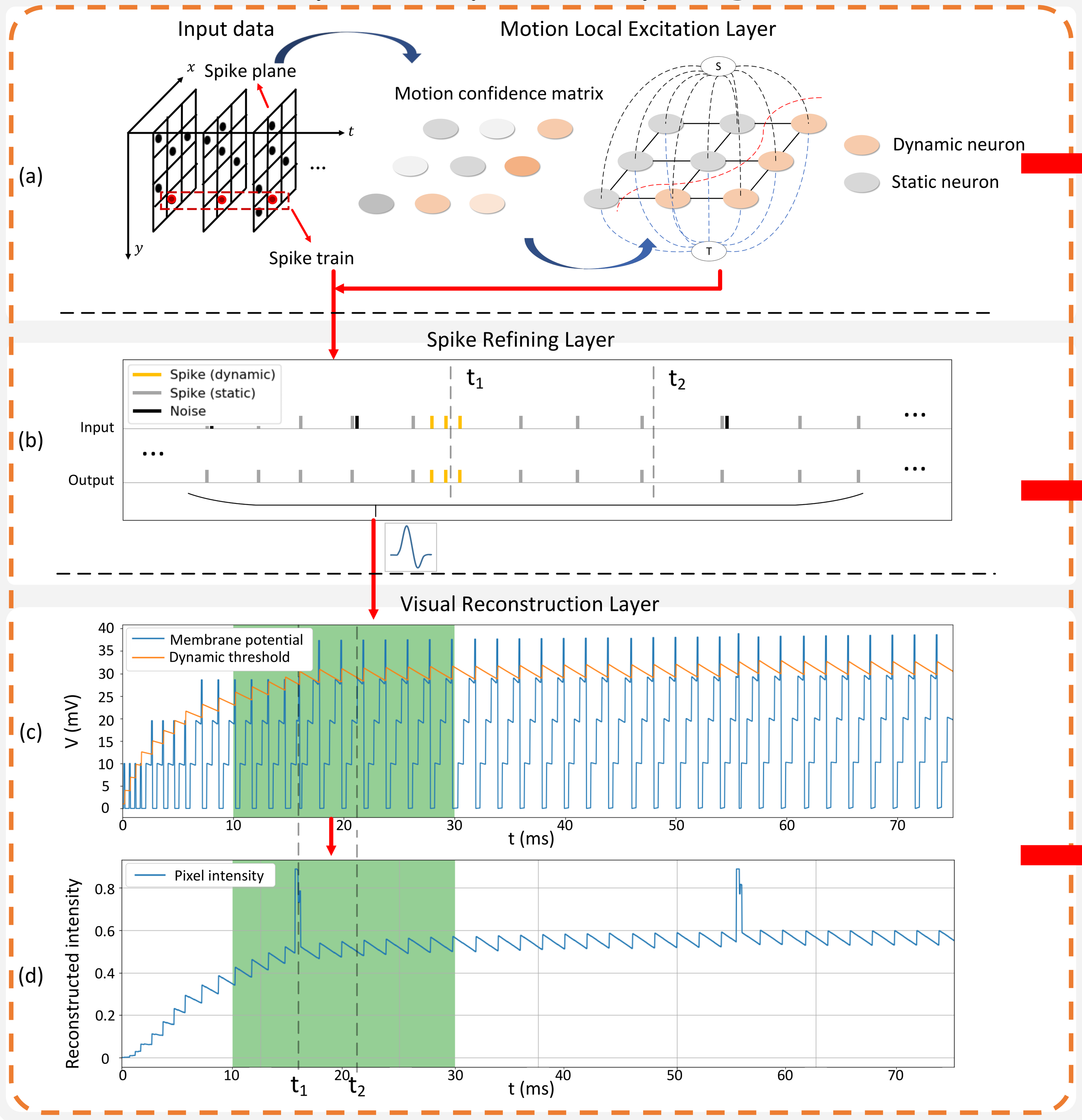
The Proposed Method: Visual Image Reconstruction via Spiking Neural Model

The overall architecture of spiking neural model



The spiking neural model consists of three layers: the motion local excitation layer, the motion local excitation layer and the visual reconstruction layer. The motion local excitation layer operates on the input spike data and outputs spike train with binary marks (dynamic or static). The spike refine layer models neuronal dynamics, the input spikes are filtered to keep the fast response to the motion while removing the noise. In visual reconstruction layer, the neuron and the synaptic connection make various adaptive adjustments. Finally, the visual image is reconstructed according to the state of the neuron.

The microscopic analysis of spiking neural model



- Motion confidence matrix

$$O = \begin{bmatrix} C_{11} & \cdots & C_{1n} \\ \vdots & \ddots & \vdots \\ C_{m1} & \cdots & C_{mn} \end{bmatrix} \in \mathbb{R}^{m\times n\times T}$$

- Dynamic neuron extraction

$$\min_{M_{ij}^t \in 0,1} \sum_{ij} (\alpha - \frac{1}{2}(O_{ij}^t)^2) M_{ij}^t + \beta \left\| E\mathrm{vec}(M^t) \right\|_1 + c$$

- Synapse Connection (STDP mechanism)

$$\Delta\omega = \sum_{t_{pre}} \sum_{t_{post}} W(t_{post} - t_{pre})$$

$$W(\Delta t) = \begin{cases} A_{pre} \exp\left(-\Delta t/\tau_{pre}\right) & \text{if } \Delta t > 0 \\ A_{post} \exp\left(\Delta t/\tau_{post}\right) & \text{if } \Delta t < 0 \end{cases}$$

- Dynamic firing threshold

$$\vartheta_{ij}(t) = \vartheta_0 + \int_0^{\infty} \theta_{ij}(s) S_{ij}^{'}(t-s) ds$$

where $\theta_{ij}(t) = \begin{cases} \eta_0 \exp(\frac{-(t-t_f)}{\tau}) & \text{if } M_{ij}^t = 0 \\ \eta_0 \exp(\frac{-(t-t_w)}{\tau \int_{t_w}^{t} S_{ij}(x)dx}) & \text{if } M_{ij}^t = 1 \end{cases}$

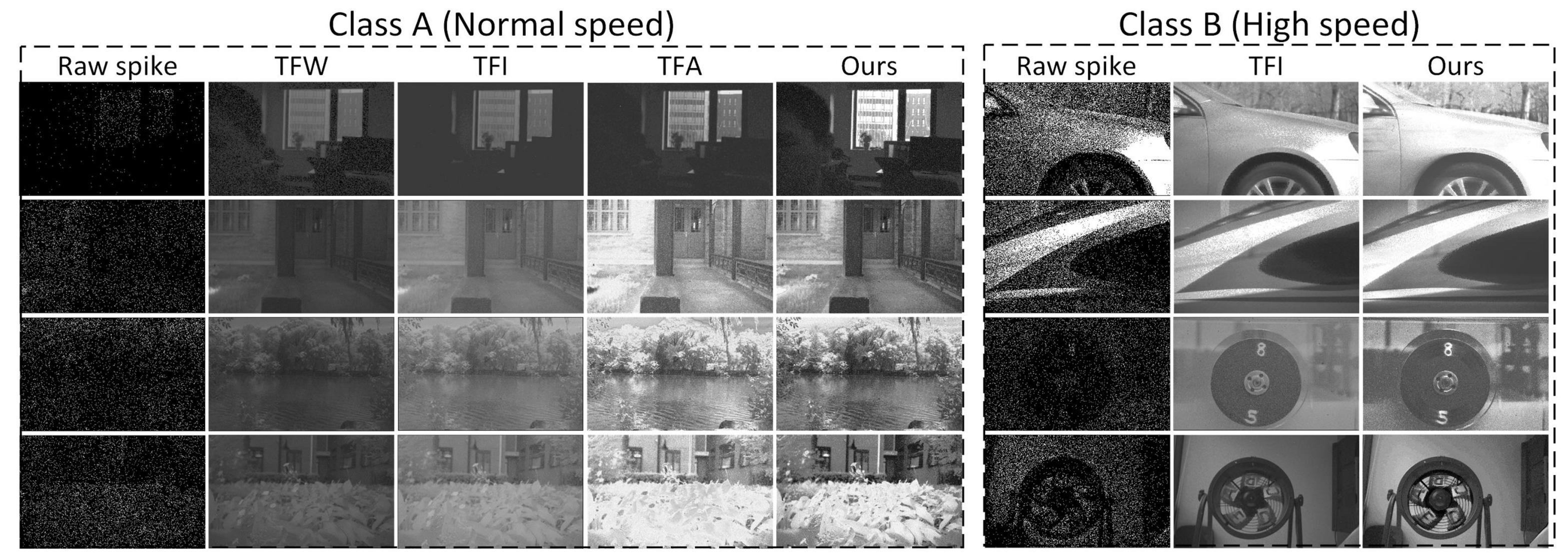
- Visual image reconstruction

$$G_{i,j,t} = \begin{cases} \vartheta_{ij}(t) & \text{if } M_{ij}^t = 0 \\ \vartheta_{ij}(t^-) * \kappa(t)^{\gamma} & \text{if } M_{ij}^t = 1 \end{cases}$$

where $\kappa(t) = t^- / (t_{isi} \int_0^{t^-} S_{ij}(x)dx)$

Experimental Results

The qualitative results

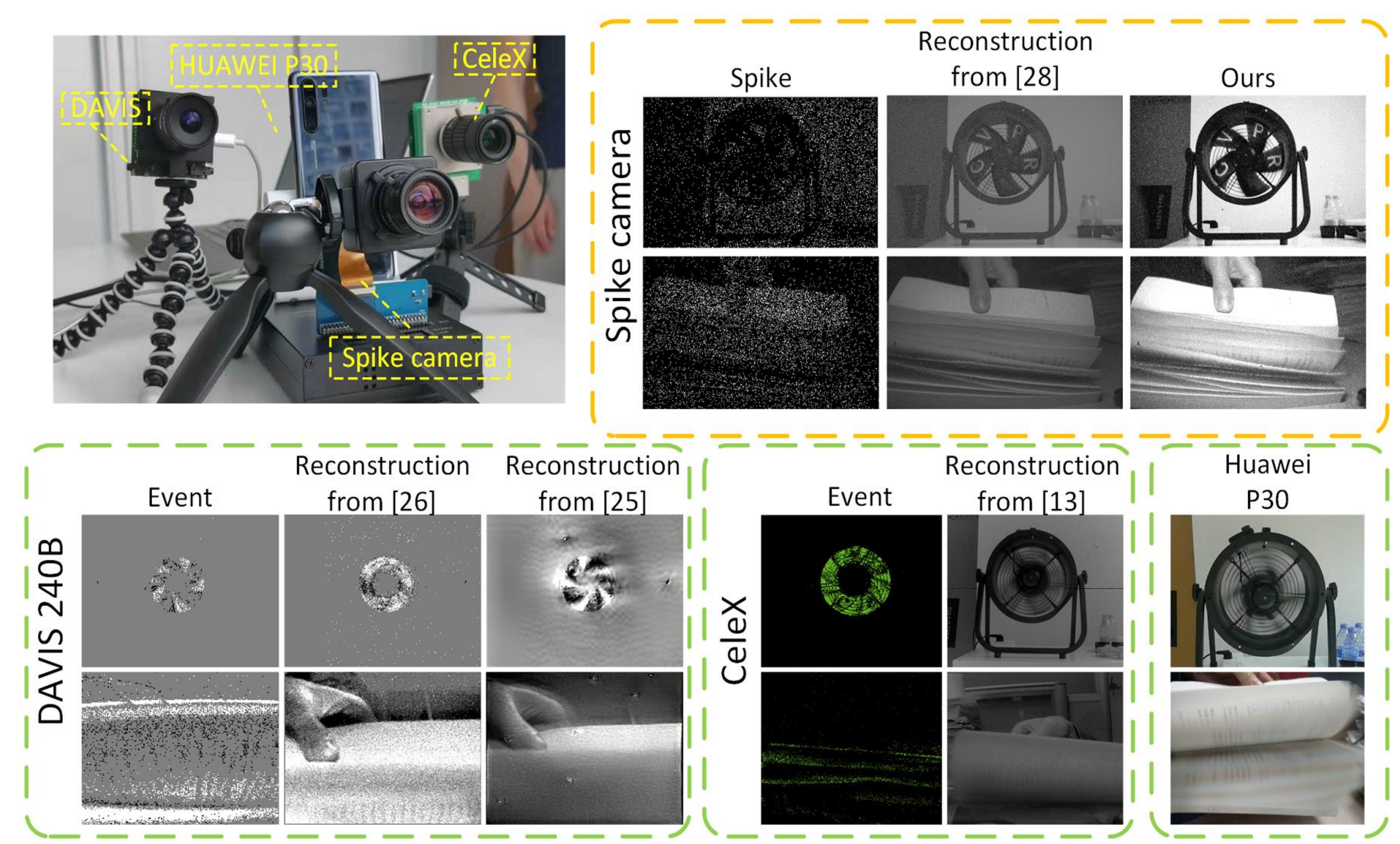


The quantitative results

Table 1. **The quantitative metrics on Class A and B.** (a higher value means better image quality)

Metric	Method	Class A (Normal speed)				Class B (High speed)				
		Office	Gallery	Lake	Flower	Car	Train	Ro1	Ro2	Mean
2-D entropy	TFW	9.12	8.68	9.45	9.22	-	-	-	-	9.12
	TFA	7.38	**12.41**	12.69	11.88	-	-	-	-	11.09
	TFI	10.01	9.85	10.51	10.01	10.71	**11.01**	10.23	9.81	10.27
	Ours	**10.38**	12.38	**12.83**	**12.63**	**10.76**	10.85	**11.72**	**11.69**	**11.66**
OG-IQA [20]	TFW	0.5729	0.4445	0.7575	**0.8959**	-	-	-	-	0.6677
	TFA	0.2737	0.6707	0.8369	0.8660	-	-	-	-	0.6618
	TFI	0.5738	0.4134	0.5110	0.8523	0.8637	0.7829	0.5602	0.5305	0.6359
	Ours	**0.5921**	**0.6879**	**0.8379**	0.8727	**0.8720**	**0.7897**	**0.6009**	**0.8105**	**0.7579**

Comparison with other vision sensors.



- Dataset is available: https://www.pkuml.org/resources/pku-spike-recon-dataset.html

Acknowledgments. This work is partially supported by grants from the National Natural Science Foundation of China under contract No. 61825101 and No. U1611461.