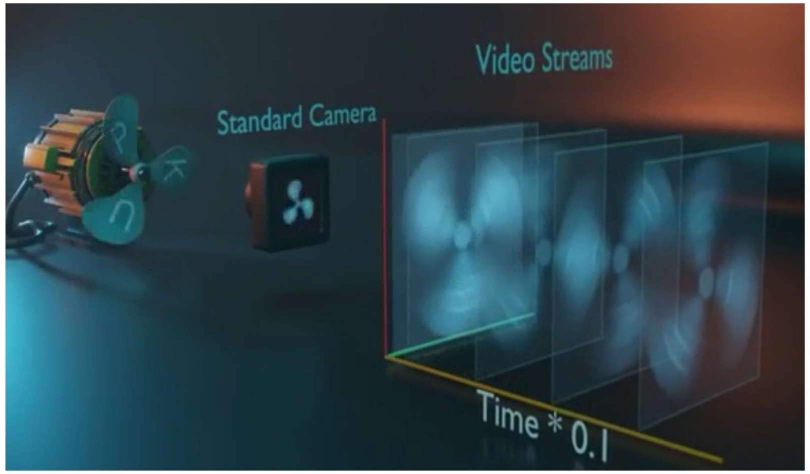
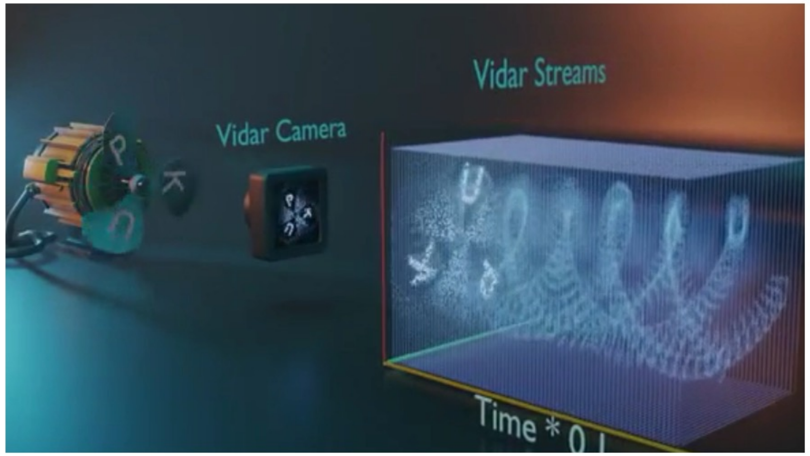
Spike Camera Autofocus via Frequency-Domain Spectral-Centroid Migration

Xijie Xiang, Lin Zhu, Yonghong Tian

Code and dataset: github.com/XIJIE-XIANG/CEN

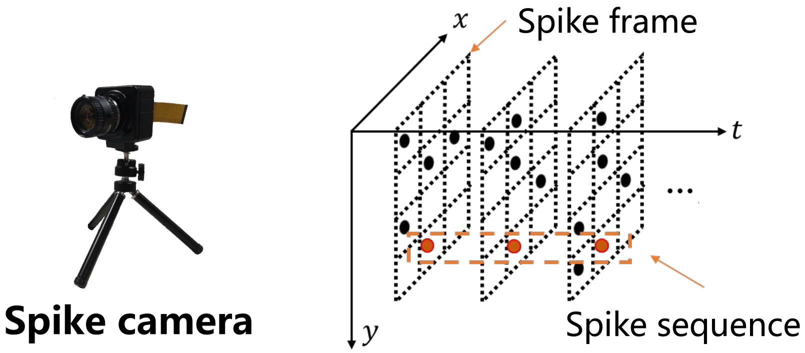
Spike camera





Principle

- Integrate-and-fire sensing
- A spike is emitted when: $\int I\,(t)\,dt \geq \theta$
- Spike frame: 0/1 matrix (40,000 fps)



Spike camera autofocus

Goal: Find the best-focused position during a focus sweep from spike-camera measurements.

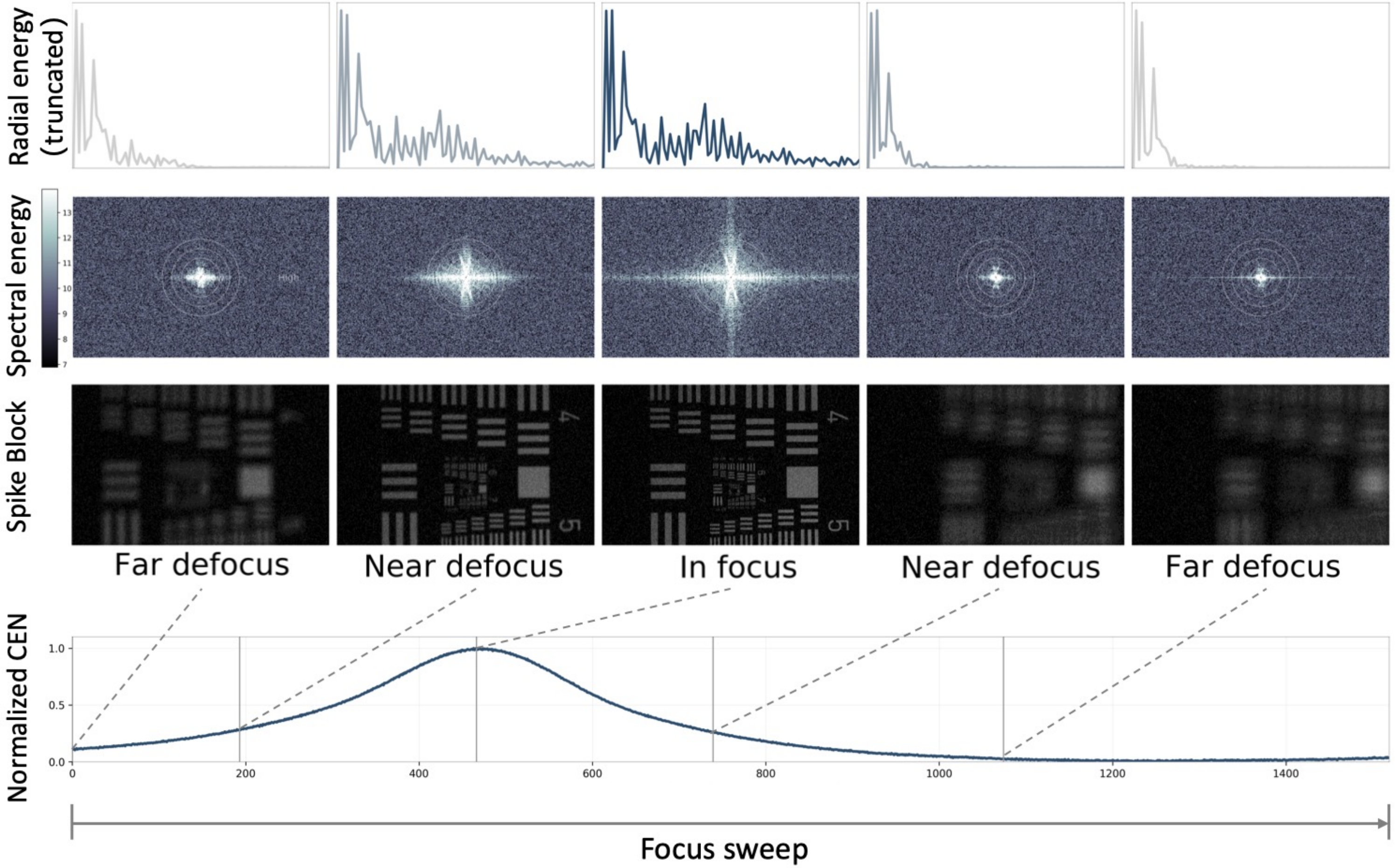
Intensity-Robust Autofocus for Spike Camera, CVPR 2024.

Challenge

- Spike streams are sparse, binary, and noisy.
- The special data format makes image-based focus measures difficult to apply directly.
- Reconstructing images before autofocus sacrifices the high-speed advantage and still falls into conventional image-based autofocus.

Key question: What focus cue is directly observable from binary spike streams?

Key Observation: Focus-Induced Spectral Migration



Observation

- **Focus**: **high frequencies**
  Defocus: **low frequencies**
- Focus changes induce **spectral-energy migration**
- Energy moves **to higher frequencies** when approaching focus
- Energy moves **back to lower frequencies** after passing focus
- The resulting **spectral centroid curve peaks near focus**

A stable focus cue is therefore available in the frequency domain.

X. Xiang, L. Zhu and Y. Tian, "[Spike Camera Autofocus via Frequency-Domain Spectral-Centroid Migration](#) ," ICML 2026.

CEN: Measuring Spectral Migration

How do we convert the frequency-domain observation into a focus score?

Pipeline

Spike sequence S_t → **Block accumulation A_b** → **FFT power spectrum P_b** → **Bounded centroid $CEN_r(b)$** → **Focus response peak**

Core computation

1. Accumulate spike frames and remove DC bias

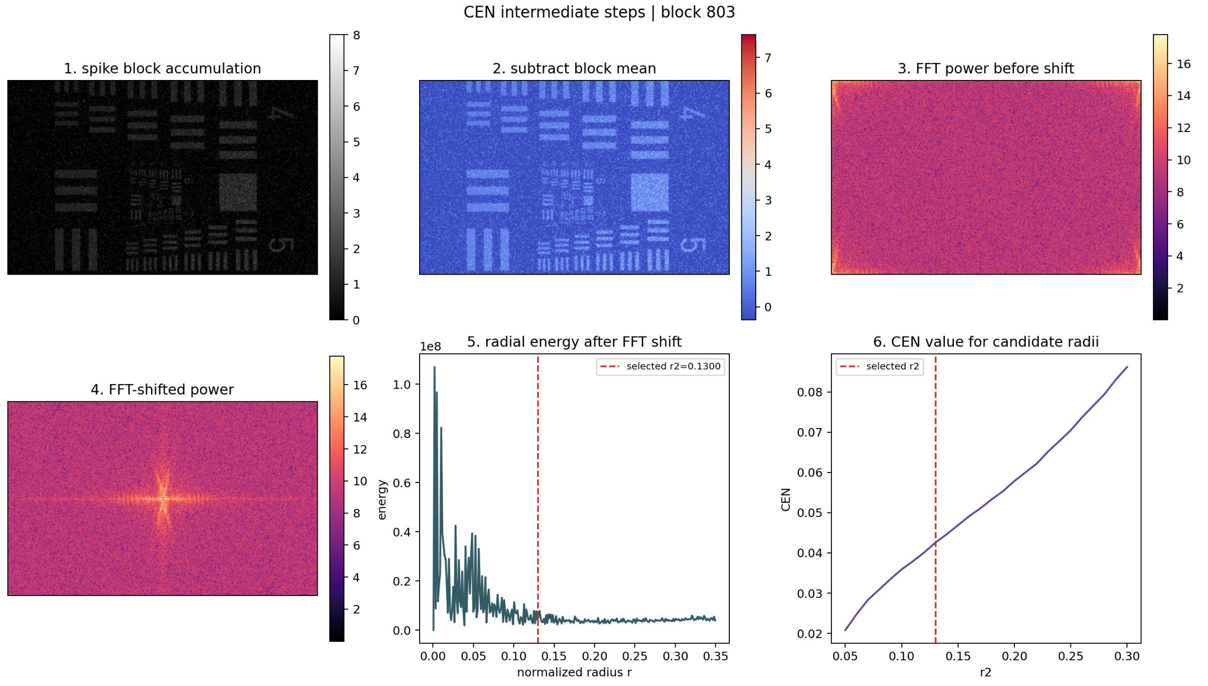
$$A_b = \Sigma_t S_t, \; \tilde{A}_b = A_b - \mu_b$$

2. Fourier transform, FFT shift and spectrum power

$$P_b(u,v) = \left|FFT_{shift}(F(\tilde{A}_b)(u,v)\right|^2$$

3. Compute bounded spectral centroid

$$CEN_r(b) = \sum_{\rho(u,v)\le r} \rho(u,v) P_b(u,v) \,/ \Big(\sum_{\rho(u,v)\le r} P_b(u,v) + \varepsilon \Big)$$



X. Xiang, L. Zhu and Y. Tian, "Spike Camera Autofocus via Frequency-Domain Spectral-Centroid Migration," ICML 2026.

Frequency-Bound Selection

Structure-consistent response selection

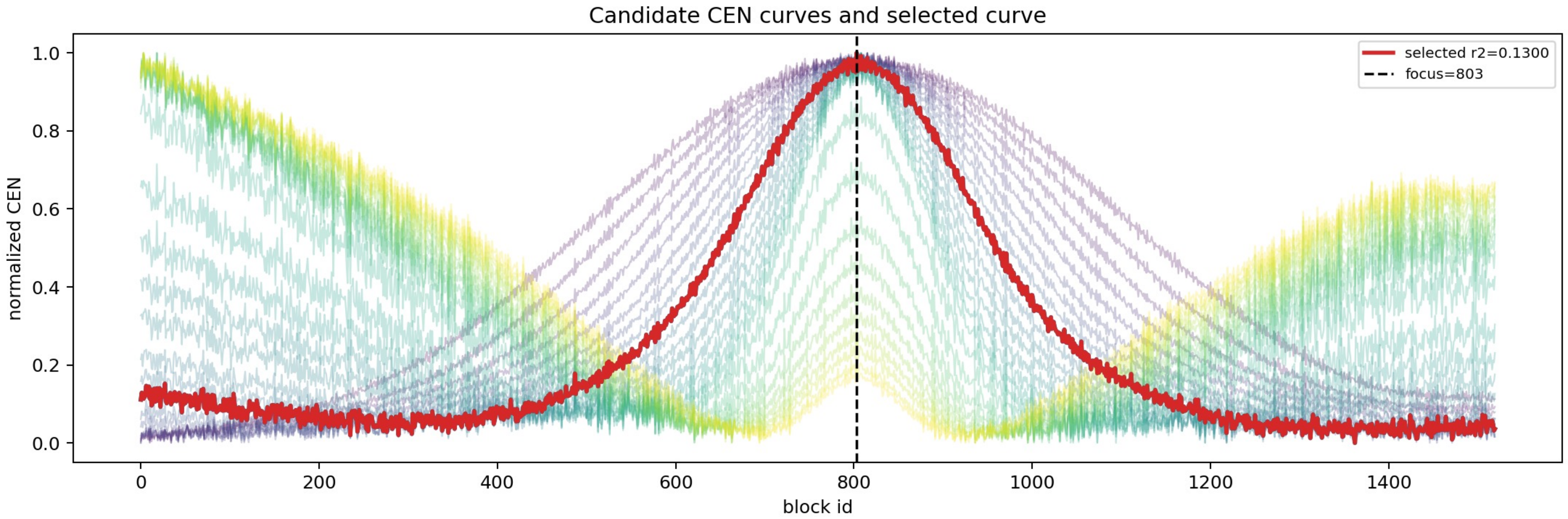
Different spectral bounds r generate different CEN curves.

$$y_r(b) = \frac{CEN_r(b) - \min_b CEN_r(b)}{\max_b CEN_r(b) - \min_b CEN_r(b) + \epsilon}$$

$$r^\star = \arg\max_{r \in R} S(y_r)$$

$$S(y_r) = \alpha_{\mathrm{prom}} S_{\mathrm{prom}} + \alpha_{\mathrm{curv}} S_{\mathrm{curv}} - \alpha_{\mathrm{width}} S_{\mathrm{width}} - \alpha_{\mathrm{plat}} S_{\mathrm{plat}} - \alpha_{\mathrm{edge}} S_{\mathrm{edge}}$$

Selection criteria: prominent, localized, sharp, non-plateau, interior peak.



Importance of Bounded parameter r

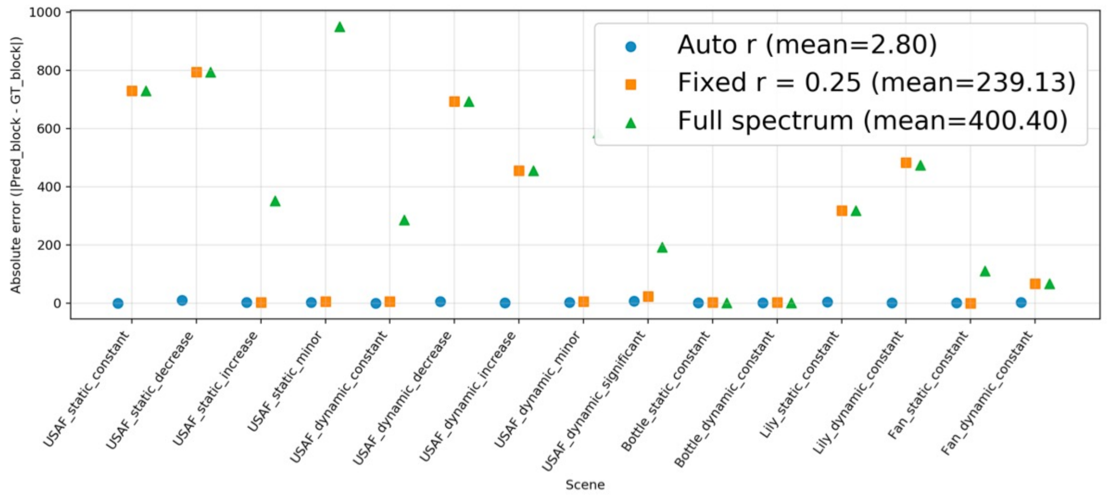


Figure 5. Ablation on the spectral radius r. We compare three settings: automatic radius selection (Auto-r), a fixed radius ($r = 0.25$), and using the full spectrum. Each marker corresponds to one scene, and the y-axis reports the absolute block error. The legend shows the mean error over all scenes.

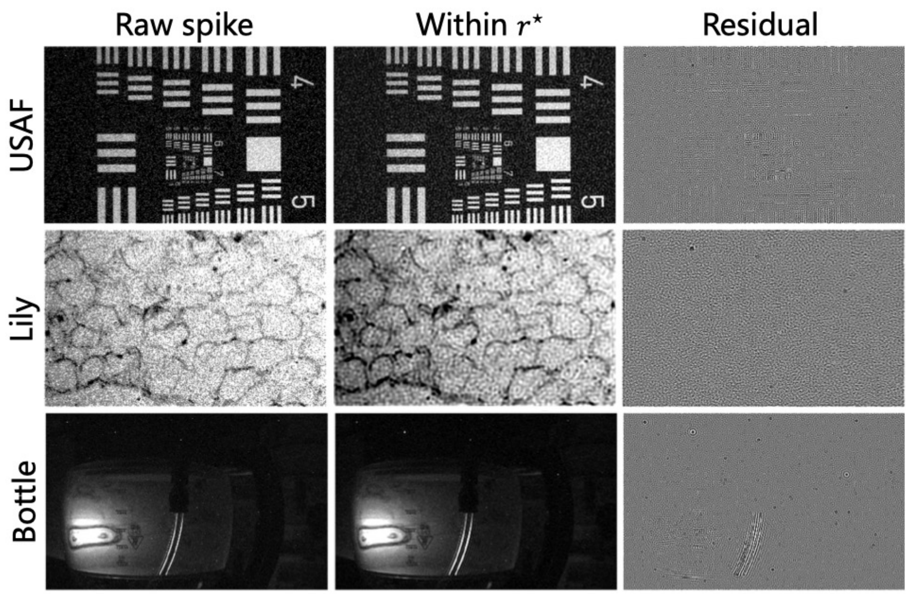


Figure 6. Effect of radius truncation in the frequency domain. Rows show representative scenes. From left to right: accumulated spike blocks, band-limited component within $r^\star$, and high-frequency residual outside $r^\star$ dominated by noise-like patterns.

X. Xiang, L. Zhu and Y. Tian, "Spike Camera Autofocus via Frequency-Domain Spectral-Centroid Migration ," ICML 2026.

Parameter ablation and sensitivity for focus curve selection

Table 7. Ablation of structural criteria used for focus curve selection.

Variant	prom	width	plateau	edge_pos	Err ↓
prom only	✓				56.47
prom+width	✓	✓			33.40
Full - prom		✓	✓	✓	35.27
Full - edge_pos	✓	✓	✓		33.07
Full - plateau	✓	✓		✓	3.13
Full - width	✓		✓	✓	3.07
Full	✓	✓	✓	✓	2.80

Table 10. Parameter sensitivity analysis on real spike-camera sequences. Mean and median absolute error ranges are reported in block units.

Parameter	Values tested	Mean Err.	Median Err.
α_{prom}	0.55, 0.80, **1.10**, 1.40, 1.65	2.73–3.20	2.00–2.00
α_{curv}	0.10, 0.20, **0.35**, 0.50, 0.70	2.80–2.80	2.00–2.00
α_{width}	0.15, 0.30, **0.45**, 0.60, 0.75	2.73–3.13	2.00–2.00
α_{plat}	0.10, 0.20, **0.35**, 0.50, 0.70	2.73–2.93	2.00–2.00
α_{edge}	0.20, 0.40, **0.60**, 0.80, 1.00	2.80–33.07	2.00–2.00
τ	0.005, 0.008, **0.012**, 0.016, 0.020	2.80–4.13	2.00–3.00
γ	1.0, 1.5, **2.0**, 3.0, 4.0	2.80–2.80	2.00–2.00

X. Xiang, L. Zhu and Y. Tian, "Spike Camera Autofocus via Frequency-Domain Spectral-Centroid Migration ," ICML 2026.

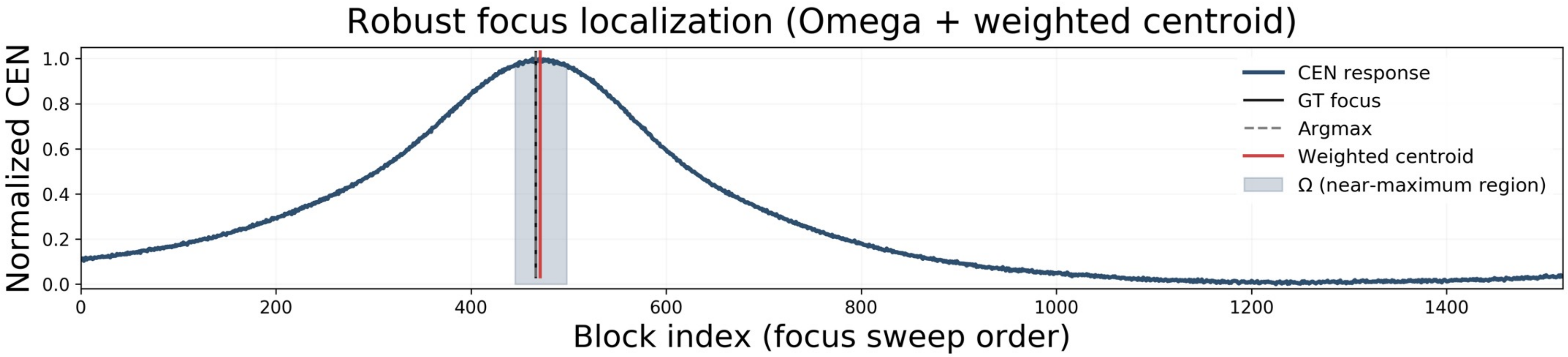
Robust Focus Localization

Instead of directly using a single argmax, we localize focus within the near-maximum region:

$$\Omega = \{b \mid y_{r^\star}(b) \geq (1-\tau)y_{r^\star}(p_0)\}, p_0 = \arg\max_b y_{r^\star}(b)$$

$$\hat{b} = \mathrm{round}\left(\frac{\sum_{b\in\Omega} b\ w_b}{\sum_{b\in\Omega} w_b + \epsilon}\right), w_b = \left(y_{r^\star}(b)\right)^\gamma$$

Goal: stabilize focus prediction under noisy or flat response peaks.



X. Xiang, L. Zhu and Y. Tian, "Spike Camera Autofocus via Frequency-Domain Spectral-Centroid Migration ," ICML 2026.

Synchronous dataset experiment

Table 1. Focus estimation error on synthetic focus sweeps under moderate-light and low-light conditions. Error is reported in centimeters; lower is better.

	Moderate-light				Low-light			
Method	simu01	simu02	simu03	**Mean**	simu01	simu02	simu03	**Mean**
GRAD	0.37	0.00	0.63	0.33	4.00	4.00	6.30	4.80
HF	0.37	0.37	0.00	0.24	**0.37**	0.37	5.00	1.90
LAP	0.37	2.20	0.00	0.86	4.00	4.00	5.60	4.60
MFDCT	0.37	0.37	0.00	0.24	4.00	4.00	5.60	4.60
SD	0.37	0.37	0.00	0.24	12.00	2.20	6.30	6.90
COUNT	4.00	4.00	0.63	2.90	12.00	6.60	2.50	7.20
CEN	**0.00**	**0.00**	**0.00**	**0.00**	2.90	**0.37**	**0.63**	**1.30**

X. Xiang, L. Zhu and Y. Tian, "Spike Camera Autofocus via Frequency-Domain Spectral-Centroid Migration ," ICML 2026.

Real dataset experiment

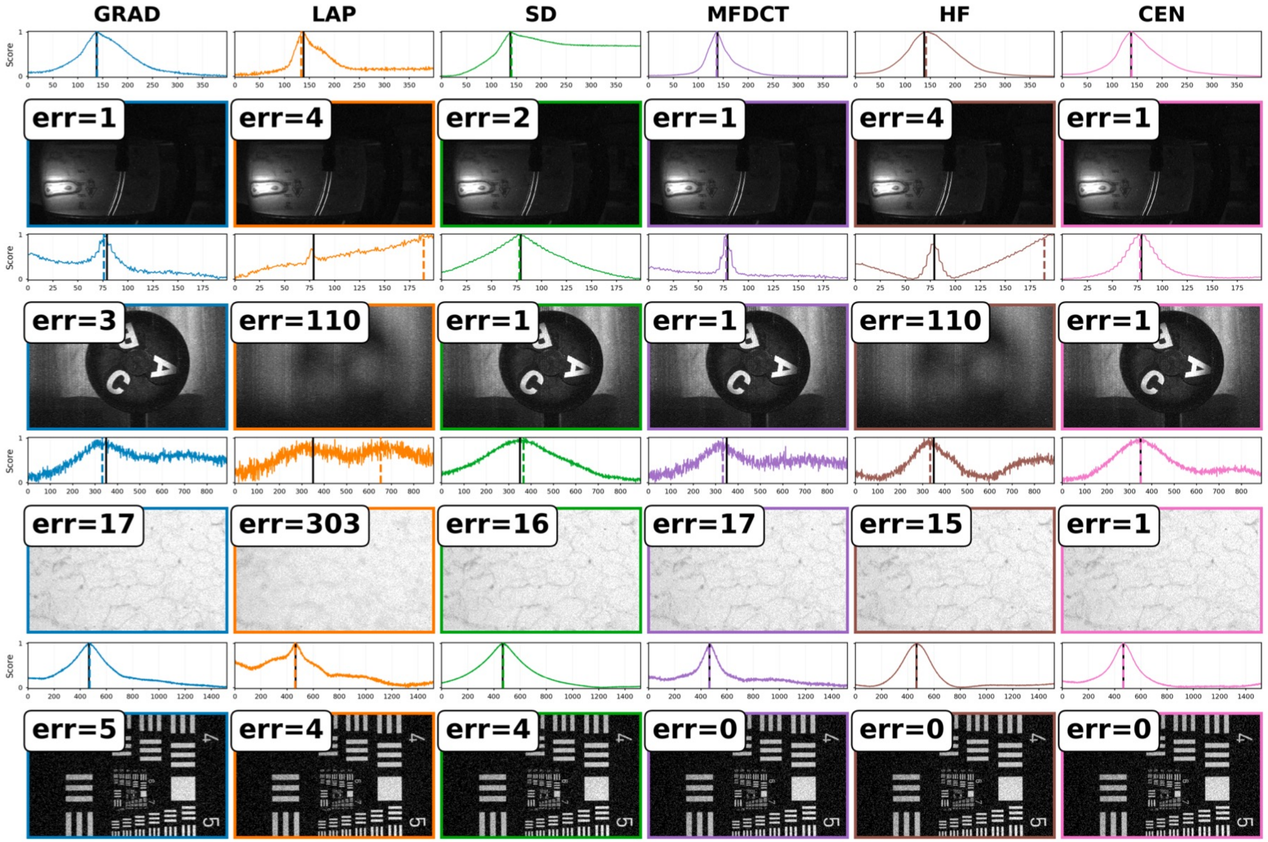


Figure 2. Qualitative comparison of autofocus responses and predicted focus blocks on representative spike-camera scenes. Each column corresponds to one method. For each scene (Bottle, Fan, Lily, USAF), the top subplot shows the normalized response curve over blocks with the ground-truth focus block indicated by a black vertical line and the method prediction indicated by a colored dashed line. The bottom subplot visualizes the accumulated block image at the predicted focus, and the block-level absolute error $|\hat{b} - b^*|$ is annotated in the corner. Overall, CEN produces a sharper and better-aligned extremum and yields more accurate focus blocks across diverse scenes.

X. Xiang, L. Zhu and Y. Tian, "Spike Camera Autofocus via Frequency-Domain Spectral-Centroid Migration ," ICML 2026.

Real dataset experiment

Table 2. Autofocus accuracy across representative spike-camera scenes. Mean AbsErr (RelErr, %) is reported; lower is better. Best and second-best results are shown in bold and underlined.

Scene	GRAD	LAP	SD	COUNT	MFDCT	HF	CEN
Bottle	**1.00 (0.3)**	<u>2.00 (0.5)</u>	**1.00 (0.2)**	129.50 (32.5)	**1.00 (0.3)**	3.50 (0.9)	**1.00 (0.3)**
Lily	20.50 (2.3)	298.00 (33.6)	12.00 (1.4)	427.50 (48.0)	18.00 (2.0)	<u>10.50 (1.2)</u>	**2.50 (0.3)**
Fan	4.50 (2.2)	88.50 (44.5)	<u>2.50 (1.2)</u>	79.50 (40.0)	3.50 (1.8)	88.50 (44.5)	**2.00 (1.0)**
USAF	368.67 (24.3)	386.11 (25.4)	7.67 (0.5)	729.56 (48.0)	500.89 (33.0)	**3.33 (0.2)**	<u>3.44 (0.2)</u>
MEAN	224.67 (15.2)	283.47 (25.7)	<u>6.67 (0.7)</u>	522.60 (44.9)	303.53 (20.3)	15.67 (6.3)	**2.80 (0.3)**

Table 3. Additional comparison with reconstruction-based image autofocus baselines. Mean AbsErr (RelErr, %) is reported; lower is better. The left block uses direct count-normalized reconstruction, while the right block uses TFI reconstruction. Best and second-best results in each row are shown in bold and underlined, respectively.

	Count-normalized reconstruction				TFI reconstruction				
Scene	GRAD	LAP	RDF	DRDF	GRAD	LAP	RDF	DRDF	CEN
Bottle	**0.50 (0.3)**	1.50 (1.0)	2.00 (1.4)	2.00 (1.4)	<u>1.00 (0.7)</u>	<u>1.00 (0.7)</u>	4.50 (3.0)	3.50 (2.3)	<u>1.00 (0.7)</u>
Lily	13.50 (3.8)	330.50 (92.7)	14.50 (4.1)	17.00 (4.7)	34.50 (9.6)	340.00 (95.0)	150.00 (42.8)	<u>11.50 (3.3)</u>	**2.50 (0.7)**
Fan	<u>3.50 (3.1)</u>	88.50 (89.5)	4.50 (4.4)	52.00 (43.3)	89.50 (87.0)	89.50 (87.0)	91.00 (92.0)	91.00 (92.0)	**2.00 (1.9)**
USAF	<u>43.11 (6.7)</u>	550.11 (91.6)	308.78 (50.2)	447.11 (70.2)	95.22 (14.1)	138.78 (19.6)	85.56 (12.7)	85.56 (12.7)	**3.44 (0.5)**
MEAN	<u>28.20 (5.0)</u>	386.13 (79.4)	188.07 (31.4)	277.73 (48.7)	73.80 (21.4)	140.67 (36.1)	84.07 (26.0)	65.47 (20.7)	**2.80 (0.8)**

X. Xiang, L. Zhu and Y. Tian, "Spike Camera Autofocus via Frequency-Domain Spectral-Centroid Migration ," ICML 2026.

Robustness under Motion and Illumination Variations

Table 5. Autofocus accuracy (AbsErr (RelErr) ↓) under static and dynamic scenes.

Motion Type	GRAD	LAP	SD	COUNT	MFDCT	HF	CEN
Static	391.00 (23.26)	183.71 (30.43)	3.00 (0.39)	541.71 (46.34)	356.86 (32.30)	33.86 (15.46)	**3.14 (0.33)**
Dynamic	201.63 (13.70)	245.13 (21.61)	9.50 (0.94)	505.88 (43.72)	256.88 (16.25)	12.25 (4.62)	**2.50 (0.36)**

Table 6. Autofocus accuracy (AbsErr (RelErr) ↓) under different illumination variation patterns during focus sweeping.

Illumination	GRAD	LAP	SD	COUNT	MFDCT	HF	CEN
Constant	98.38 (7.55)	188.38 (29.32)	4.38 (0.79)	327.00 (41.23)	96.88 (7.03)	25.75 (11.66)	**1.50 (0.40)**
Decrease	242.00 (15.90)	743.00 (48.90)	**6.50 (0.40)**	624.00 (41.10)	620.00 (40.80)	7.00 (0.45)	7.50 (0.50)
Increase	21.50 (1.40)	403.00 (26.55)	**1.50 (0.10)**	881.00 (58.00)	239.00 (15.75)	3.00 (0.20)	**1.50 (0.10)**
Minor	753.50 (49.60)	147.50 (9.70)	4.00 (0.25)	754.50 (49.70)	758.50 (49.95)	**0.50 (0.05)**	2.50 (0.20)
Significant	549.00 (36.10)	154.00 (10.10)	37.00 (2.40)	696.00 (45.80)	543.00 (35.70)	8.00 (0.50)	**7.00 (0.50)**

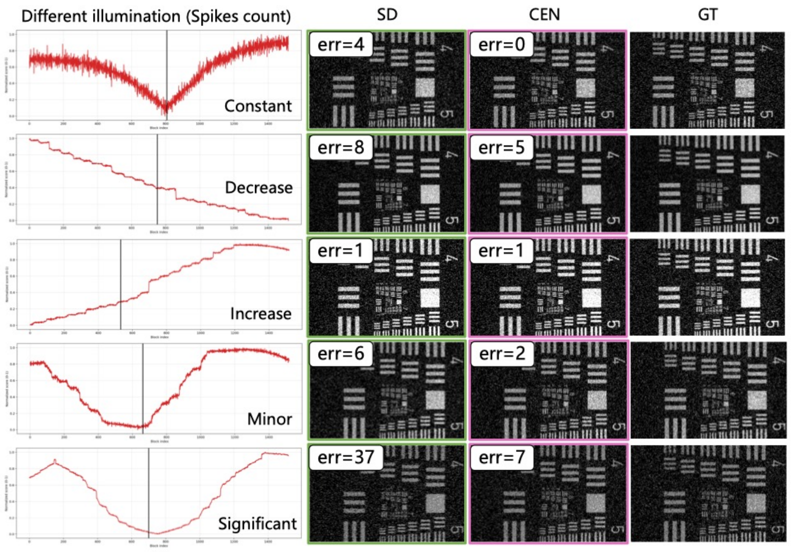


Figure 4. Autofocus comparison under different illumination changes. Each row shows one illumination pattern (constant, decrease, increase, minor, significant). From left to right: spike count curve, SD result, CEN result, and GT. Absolute block error is shown in the upper-left of each estimated focus image.

X. Xiang, L. Zhu and Y. Tian, "Spike Camera Autofocus via Frequency-Domain Spectral-Centroid Migration ," ICML 2026.

Limitations and failure cases

1. CEN works when focus variation produces a stable spectral migration pattern. It may fail when the cue is corrupted or missing.

- **Severe noise:** response tail may be dominated by noise
- **Insufficient spikes:** near-zero illumination makes the focus cue unobservable

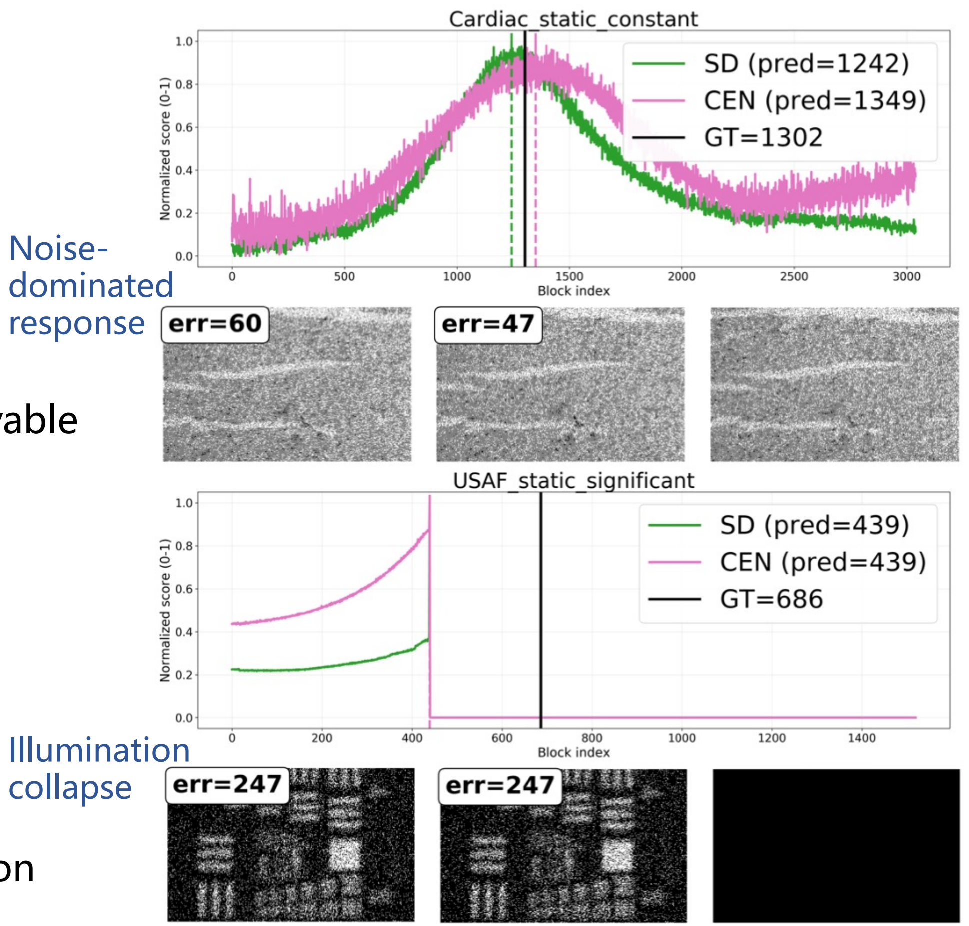


Figure 8. Representative failure cases. Top: an extremely noisy scene where noise dominates the response tail after the GT focus position. Bottom: the USAF static significant scene, where illumination drops to near zero and no spikes are generated at the GT focus. For each case, response curves of SD and CEN are shown together with the GT, and the accumulated spike blocks at the predicted focus are visualized with absolute block error indicated.

2. Edge deployment and computational cost

- CEN requires repeated block-wise 2D FFTs， complexity: $O(N \log N)$
- Latency and energy cost under strict edge constraints still need verification

X. Xiang, L. Zhu and Y. Tian, "Spike Camera Autofocus via Frequency-Domain Spectral-Centroid Migration ," ICML 2026.

Thank you !