Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Neuromorphic High-Throughput Imaging via Transport-Induced Differential Sensing

Chutian Wang

ctwang@eee.hku.hk

The University of Hong Kong https://orcid.org/0000-0002-5941-8028

Shuo Zhu

The University of Hong Kong https://orcid.org/0000-0002-3000-1536

NI Chen

The University of Hong Kong https://orcid.org/0000-0002-0663-3867

Jingqian Wu

The University of Hong Kong

Shuo Li

The University of Hong Kong

Yuxing Li

The University of Hong Kong https://orcid.org/0009-0004-8128-1517

Yanmin Zhu

The University of Hong Kong

Rongzhou Chen

The University of Hong Kong

Yunping Zhang

Hong Kong Shue Yan University https://orcid.org/0000-0002-6929-7784

Edmund Lam

Department of Electrical and Electronic Engineering, Faculty of Engineering, The University of Hong Kong https://orcid.org/0000-0001-6268-950X

Article

Keywords:

Posted Date: June 17th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9045382/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

Neuromorphic High-Throughput Imaging via Transport-Induced Differential Sensing

Chutian Wang[1], Shuo Zhu[1], Ni Chen[1], Jingqian Wu[1], Shuo Li[1], Yuxing Li[1], Yanmin Zhu[1,2], Rongzhou Chen[1], Yunping Zhang[1,3], and Edmund Y. Lam[1,4,*]

[1]Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam, Hong Kong, China
[2]School of Biomedical Engineering, The University of Hong Kong, Pokfulam, Hong Kong, China
[3]Department of Applied Data Science, Hong Kong Shue Yan University, North Point, Hong Kong, China
[4]School of Innovation, The University of Hong Kong, Pokfulam, Hong Kong, China
[*]e-mail: elam@eee.hku.hk

ABSTRACT

High-throughput imaging is fundamentally constrained by the trade-off between acquisition speed and spatial resolution under finite spatio-temporal bandwidth. Here we show that this compromise can be restructured at the measurement stage: transport-induced intensity variations, when captured through differential sensing, encode spatial gradients into the temporal domain and thereby circumvent the gradient attenuation associated with exposure integration. By formulating this process as a physics-consistent differential operator under fixed sensing bandwidth, fine spatial structure can be recovered directly from motion without modifying the underlying optical hardware, enabling bandwidth-efficient spatial reconstruction. Building on this principle, we develop Neuromorphic Super-Resolution (NeuroSR), a framework that enables joint inference of super-resolved structure, transport dynamics, and task-dependent physical parameters. Through controlled transport experiments and demonstrations in holographic phase retrieval and biological microscopy, we show that transport-induced differential sensing provides a general mechanism for enhancing spatial observability in bandwidth-limited high-throughput imaging systems.

Introduction

High-throughput imaging systems aim to recover fine spatial structure under rapid dynamics, yet are fundamentally constrained by finite spatio-temporal bandwidth and power budgets, imposing system-level limits[1]. Under such constraints, the manner in which transient intensity variations are encoded at acquisition determines the observability of fine structure in subsequent reconstruction. Biological vision provides an instructive example of how temporally structured encoding can partially mitigate bandwidth limitations at the sensing stage[2,3]. In particular, even during nominal fixation, involuntary ocular motions and rapid saccades continuously modulate the retinal image[4,5], converting spatial structure into temporally structured signals that can be integrated by downstream circuitry[6]. This motion-induced encoding effectively decomposes visual information into complementary components: a photometric anchor that preserves stable intensity structure, and transient features that encode transport-induced photometric variations. These complementary signals are jointly integrated at subsequent cortical stages, where spatial context and motion cues are combined to infer fine structure beyond static sampling limits[7,8]. In this sense, biological vision illustrates how temporally structured sampling can partially offset bandwidth-induced loss of spatial information.

Multi-frame super-resolution (SR) in computational imaging infers fine spatial detail by aggregating motion-induced subpixel shifts of photometric gradients across multiple low-resolution frames[9–11]. Despite its broad adoption, SR remains fundamentally rooted in a synchronous sampling paradigm, in which spatial resolution is recovered through temporal redundancy under frame-based acquisition[12,13]. As acquisition speed and throughput increase, this redundancy becomes increasingly inefficient: finite exposure and photon budgets force a trade-off between noise and motion blur, while temporal integration progressively attenuates gradient contrast that constrains fine structure in inverse problems. Consequently, although nominal spatial sampling may be preserved, the effective gradient information available to computational reconstruction degrades rapidly, rendering many high-throughput computational imaging tasks increasingly ill-posed, including applications such as quantitative phase imaging[14,15]. These limitations are intrinsic to synchronous frame integration and persist across optical super-resolution strategies that rely on temporally aggregated snapshots to exceed sampling limits[16].

Asynchronous sensing therefore offers a natural alternative for high-throughput imaging by redistributing a finite sensing bandwidth toward temporally informative variations rather than uniformly integrating absolute intensity. Event cameras instantiate this paradigm by encoding temporal changes instead of absolute intensity. By reporting per-pixel brightness

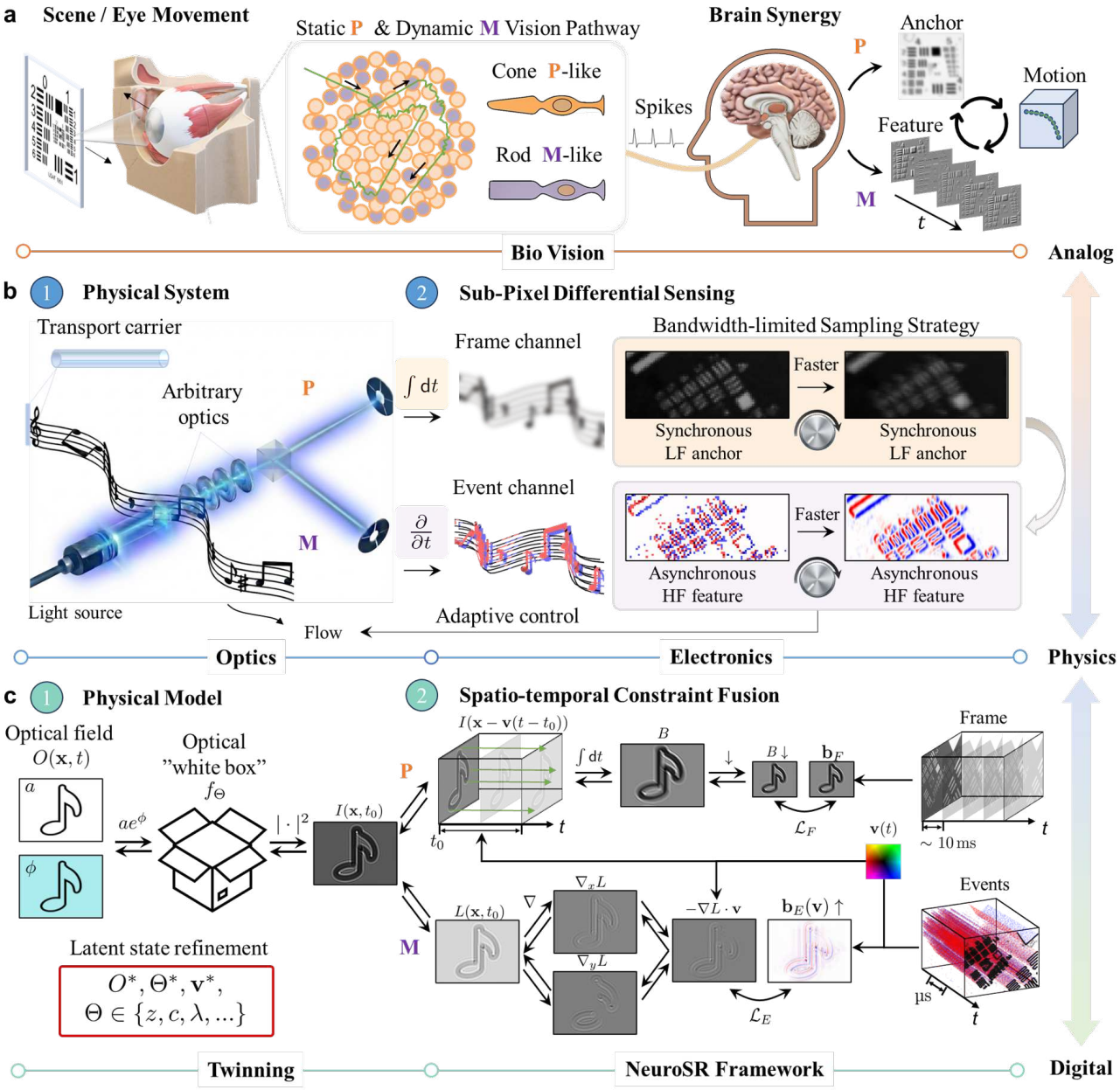


Fig. 1. Conceptual overview of NeuroSR. a, Biological motivation. Relative motion in primate vision transforms spatial structure into temporally modulated signals, which are encoded through complementary parvocellular (P) and magnocellular (M) pathways as a sustained photometric anchor and transient features, respectively. **b, Sub-pixel differential sensing.** Under a fixed spatio-temporal bandwidth budget typical of high-throughput imaging, synchronous frame-based sensing primarily constrains low-frequency (LF) intensity anchors through temporal integration, while remaining susceptible to noise and motion blur. In contrast, asynchronous event sensing selectively responds to transport-induced intensity variations, mapping continuous motion into subpixel-resolved high-frequency (HF) photometric gradient features. This measurement strategy redistributes the limited sampling bandwidth toward HF components where fine spatial detail is concentrated. **c, Physics-consistent reconstruction.** NeuroSR reconstructs high-resolution physical structures by jointly leveraging frame-based intensity constraints and spatio-temporal structures encoded in the event stream through a physically interpretable forward model. By parameterizing the imaging system as f_Θ, the framework supports plug-and-play integration across diverse computational imaging tasks. A representative application to complex-field recovery in inline digital holography is shown.

increments with microsecond latency and high dynamic range[17], the event channel suppresses redundant sampling in static

regions and concentrates measurements on transport-induced intensity variations[18,19]. From a measurement perspective, however, event sensing provides differential temporal increments and therefore imposes incomplete constraints on the underlying spatio-temporal signal when used in isolation—an incompleteness intrinsic to differential sensing that cannot be resolved by reconstruction algorithms alone. Prior efforts have addressed this limitation through hardware-level modulation strategies, such as mechanically induced microsaccades[20] or temporally sparse fluorophore blinking[21,22], as well as algorithmic fusion of event streams with low-frame-rate intensity images[23–25]. While effective in specific settings, these approaches are typically developed outside a unified high-throughput computational imaging framework: event measurements are often incorporated as auxiliary signals or intermediate representations rather than as explicit measurement operators coupled to the physical image formation process[26–29]. To date, principled information allocation and joint optimization of structure, motion, and system parameters under fixed spatio-temporal bandwidth constraints remain underexplored in high-throughput imaging.

This conceptual gap motivates a fundamental question: can neuromorphic sensing be rigorously incorporated into computational imaging in a physically interpretable and differentiable manner that extends, rather than replaces, existing inference frameworks? While synchronous frame-based solvers have been extensively developed across learning-based and model-based paradigms[30–35], a central challenge lies in embedding asynchronous event measurements as explicit constraints within intensity-domain forward models. Prior work has demonstrated that events can be triggered at microsecond latency even when scene translation remains within a fraction of a pixel[36–38]. Under continuous transport, such subpixel temporal responsiveness implies that differential intensity variations are registered before a full spatial sampling interval is traversed. When coherently aligned through motion compensation, events can constructively accumulate over extended temporal windows to reveal noise-mitigated subpixel structure[39,40]. This phenomenon reflects the fact that events encode photometric gradient information of the underlying scene in the presence of relative motion, such that joint inference of the transport dynamics and structure yields a motion-compensated representation at subpixel precision. Importantly, the apparent resolution enhancement arises from a redistribution of measurement density across motion-corrected coordinates[41], rather than from additional spatial sampling or heuristic interpolation. As illustrated in Figure 1(b), such gradient-domain measurements can be elicited through simple relative motion without specialized optical hardware, indicating that transport-induced differential sensing provides a physically grounded pathway toward enhanced spatial observability in high-throughput imaging systems.

In this work, we introduce Neuromorphic Super-Resolution (NeuroSR), a high-throughput computational imaging framework grounded in transport-induced differential sensing. Rather than treating event streams as an alternative sensing modality or post hoc representation, NeuroSR interprets events as measurements of transport-induced photometric gradients and embeds them explicitly as a differential measurement operator within physics-consistent forward models. Within this formulation, low-frame-rate intensity measurements provide a low-frequency photometric anchor, while asynchronously encoded gradient signals supply complementary high-frequency constraints that remain observable under bandwidth-limited, high-throughput acquisition. By coupling these heterogeneous measurements within a unified, fully differentiable inference framework, NeuroSR enables joint optimization of latent structure, motion, and task-dependent physical parameters, enabling high-resolution physical quantities to be recovered with subpixel precision. Taken together, NeuroSR establishes a principled framework for reallocating sensing bandwidth through motion-induced differential measurements, enabling enhanced spatial observability in bandwidth-limited high-throughput imaging systems.

Results

We consider a general high-throughput imaging regime representative of stream-based and flow-driven inspection systems, wherein neuromorphic imaging leverages asynchronous sampling to bypass the conventional trade-off between spatial resolution and acquisition speed. Within a spatiotemporal window $\Omega \times T$, the scene logarithmic intensity $L \triangleq \log I(\mathbf{x},t)$ evolves continuously and, over a sufficiently short temporal neighborhood around a reference time t_0, admits a local first-order approximation that serves as the effective observation model in the motion regime. Under the brightness constancy assumption, temporal intensity variations can be related to spatial motion, yielding

$$\begin{aligned} L(\mathbf{x},t) &= L(\mathbf{x},t_0) + \frac{\partial L}{\partial t}(\mathbf{x},t_0)\,(t-t_0) + \mathcal{O}\big((t-t_0)^2\big) \\ &\approx \underbrace{\log I(\mathbf{x},t_0)}_{\text{LF Anchor}} - \underbrace{\nabla L(\mathbf{x},t_0)\cdot\mathbf{v}(t_0)}_{\text{HF Feature}}\,(t-t_0), \end{aligned} \tag{1}$$

which decomposes the spatiotemporal signal into a low-frequency (LF) photometric anchor and a motion-induced variation component. Importantly, this relation makes explicit a time-to-space transfer in the motion regime: finer temporal sampling of the increment $(t-t_0)$ improves the estimation fidelity of the high-frequency (HF) gradient feature $-\nabla L\cdot\mathbf{v}$, enabling spatial detail to be recovered without requiring dense, high-rate intensity frames. Under a fixed spatiotemporal throughput budget, the measurement allocation can be characterized by a sampling-density model (Supplementary Note 1).

Based on this observation, the reconstruction problem is governed by the complementarity of frame-based and event-based measurements, which NeuroSR instantiates within a unified inference framework:

$$O^*, \mathbf{v}^*, \Theta^* = \arg\min_{O,\mathbf{v},\Theta} \underbrace{\mathcal{D}_F\left\{\mathbf{b}_F(\mathbf{x}), \int_T I(\mathbf{x},t)\,\mathrm{d}t \downarrow\right\}}_{\text{Frame Fidelity } \mathcal{L}_F} + \alpha \underbrace{\mathcal{D}_E\left\{\mathbf{b}_E(\mathbf{x};\mathbf{v}), -\nabla L(\mathbf{x})\cdot\mathbf{v}\right\}}_{\text{Event Fidelity } \mathcal{L}_E} + \sum_{n=1}^{N} \beta_n \mathcal{R}_n(O,\mathbf{v}),$$
$$\text{s.t.} \quad I(\mathbf{x},t) = f_\Theta(O), \quad O \in \Omega_O, \quad \mathbf{v} \in \Omega_\mathbf{v}, \quad \Theta \in \Omega_\Theta, \tag{2}$$

where the reconstruction is governed by the complementarity between the two fidelity terms: $\mathcal{L}_F$ anchors the global photometric baseline using temporally integrated intensity measurements, while $\mathcal{L}_E$ enforces asynchronously triggered, subpixel gradient constraints derived from the event stream, weighted by the coefficient α. The regularizers $\{\mathcal{R}_n\}$ and admissible sets $\Omega_{\{\cdot\}}$ encode task-dependent priors and feasibility constraints on $(O,\mathbf{v},\Theta)$. Crucially, these terms are coupled through a shared forward operator f_Θ, which parameterizes the image formation process mapping the latent object O to the encoded intensity, and are kept modular to support plug-and-play integration with existing computational imaging solvers. In the following sections, this framework is instantiated at progressively richer levels of forward modeling in order to isolate and demonstrate the role of neuromorphic sensing under different high-throughput regimes.

Mechanism demonstration under controlled transport

We first isolate the sensing mechanism of NeuroSR under controlled transport, where the relationship between temporal sampling density and spatial observability can be examined directly. A large-field-of-view (FOV) relay imaging system provides a convenient experimental platform for this regime, as it enables uniform transport across the sensor while maintaining fixed optical throughput[42,43] (see Supplementary Video 1). As illustrated in Figure 2(a), a motorized translational stage induces constant-speed motion across the field of view, yielding a known uniform transport field $\mathbf{v} = \text{const}$. Opaque workpieces are imaged under reflected illumination using a collimated LED source, and a $4f$ relay system with unit magnification provides aberration-free imaging over a $0.4\,\text{cm} \times 0.6\,\text{cm}$ field of view[44]. With illumination power and stage speed controlled, the acquisition can be treated as exposure-ideal, allowing the constraint in Equation (2) to reduce to a super-resolved intensity reconstruction problem from spatially downsampled observations.

To enable evaluation, a positive 1951 USAF resolution target is used as the test sample, which in principle contains spatial frequency components beyond the Nyquist rate of the sensor. As depicted in Figure 2(b), the etalon is also imaged under a standard wide-field (WF) microscope ($3\times$, NA = 0.2, transmission mode) to provide an experimental SR ground truth (GT) for reference, whose motion-aligned structural feature reference exhibits convincing consistency with the measured IWE$\uparrow$ from the event channel in the rightmost panel of Figure 2(c). NeuroSR is compared with representative model-based super-resolved intensity reconstruction methods, including single-frame SR (SFSR: Fourier upsampling and ADMM[45]), multi-frame SR (MFSR[46], using $n = 30$ frames), and the state-of-the-art event-only Ev-PnP[41]. As depicted in Figure 2(c), the raw frame (blue) cannot resolve element 4 in group 4 (G4-E4), being limited by the Nyquist rate in the frequency domain, as indicated by the dashed square and line in the 2D power spectrum and its radial section, *i.e.*, the modulation transfer function (MTF), respectively. The SFSR methods (orange, green) fail to extend the 2D power spectrum due to insufficient high-frequency fidelity, resulting in false frequency structures compared with methods that contain higher frequency fidelity, *i.e.*, MFSR (brown), NeuroSR (purple), and GT (pink). The event-only method, *i.e.*, Ev-PnP (red) fails to reconstruct global intensity, causing a severe mismatch in absolute values. Ours (purple) resolves element 3 in group 5 (G5-E3). The 2D power spectrum closely matches the GT and MFSR, demonstrating sampling enhancement beyond the Nyquist limit and approaching the diffraction limit, as marked by the MTF boundary in Figure 2(d).

Stabilizing snapshot holography with asynchronous differential sensing

We next examine whether transport-induced differential sensing can stabilize complex-field recovery in sweep-based lensless inline holography, a setting in which motion-induced super-resolution is commonly realized through pixel super-resolution holography[47–49]. Phase retrieval in inline holography is intrinsically ill-posed due to the compressive nature of the intensity-only measurement and the well-known twin-image ambiguity. Prior studies have shown that additional constraints, such as handcrafted support priors[50] or spatial regularization strategies[49], can partially mitigate these ambiguities and enable twin-image–artifact-free quantitative phase imaging. NeuroSR does not replace these priors; instead, it complements them by introducing motion-induced differential constraints derived from asynchronous events directly into the data fidelity term (see Supplementary Video 2).

We first examine the refocusing behavior of a standard USAF resolution target placed at an uncertain axial distance. As shown in Figure 3(c), NeuroSR consistently reconstructs sharper high-frequency bars by jointly adapting the propagation depth z and the sweeping trajectory $\mathbf{v}$, fully resolving Group 4 Element 6 and partially retrieving the first three elements of Group 5. At close distances, where the effective numerical aperture is higher, conventional methods such as backpropagation

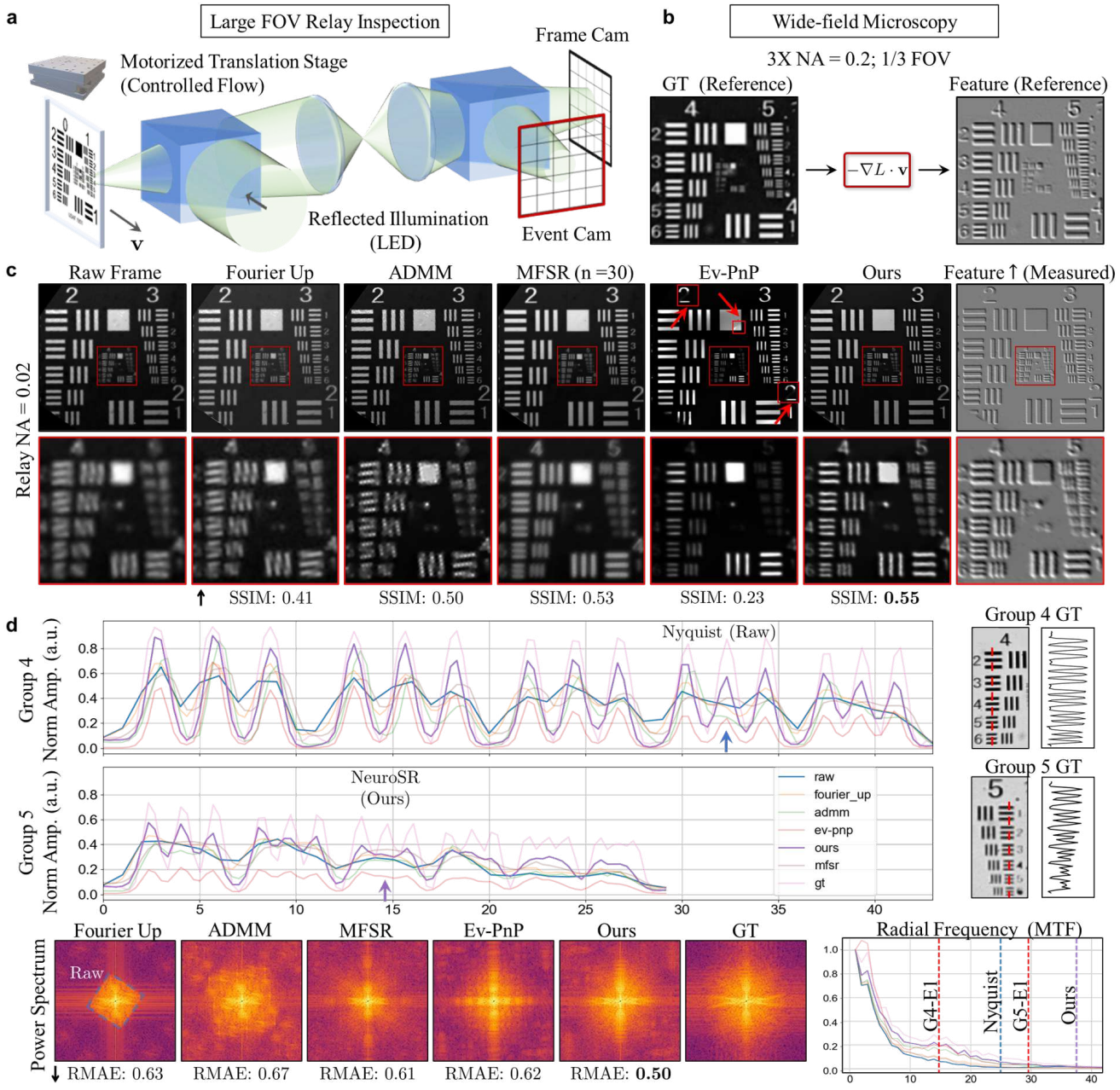


Fig. 2. NeuroSR enables high-fidelity reconstruction in large-FOV relay inspection. a, Optical setup. A reflective $4f$ relay system is used to image a resolution target with NA = 0.1 under uniform motion. **b, GT Acquisition.** A wide-field (WF) microscopy (NA = 0.2, 3×, transmission mode) is used to obtain the optical SR reference of the central $1/3$ FOV. **c, Reconstruction Comparison.** NeuroSR is compared to model-based benchmark methods fed with different inputs, including a single frame (Fourier up, ADMM), multiple frames (MFSR), and only events (Ev-PnP). Feature ↑ reveals sub-pixel-resolved gradient map consistent with the reference in **b**. Red boxes highlight missing edges and artifacts in event-only reconstructions. **d, MTF Analysis.** NeuroSR (purple) extends the resolution beyond the original Nyquist limit (blue), from element 4 in group 4 (G4-E4) to element 3 in group 5 (G5-E3), while preserving frequency structural similarity in the 2D power spectrum.

and automatic differentiation[51] fail to resolve Group 4 Element 3 due to overexposure-induced texture loss. NeuroSR, in contrast, preserves usable hologram gradient information by leveraging the high dynamic range capability of the event sensor, thereby stabilizing the reconstruction under severe illumination saturation. These results confirm the robustness of the proposed framework across both axial defocus and challenging exposure regimes with 2× pixel super-resolution and significant noise mitigation. We then extend the evaluation to complex-valued biological samples. In Figure 3(d) (top), NeuroSR resolves

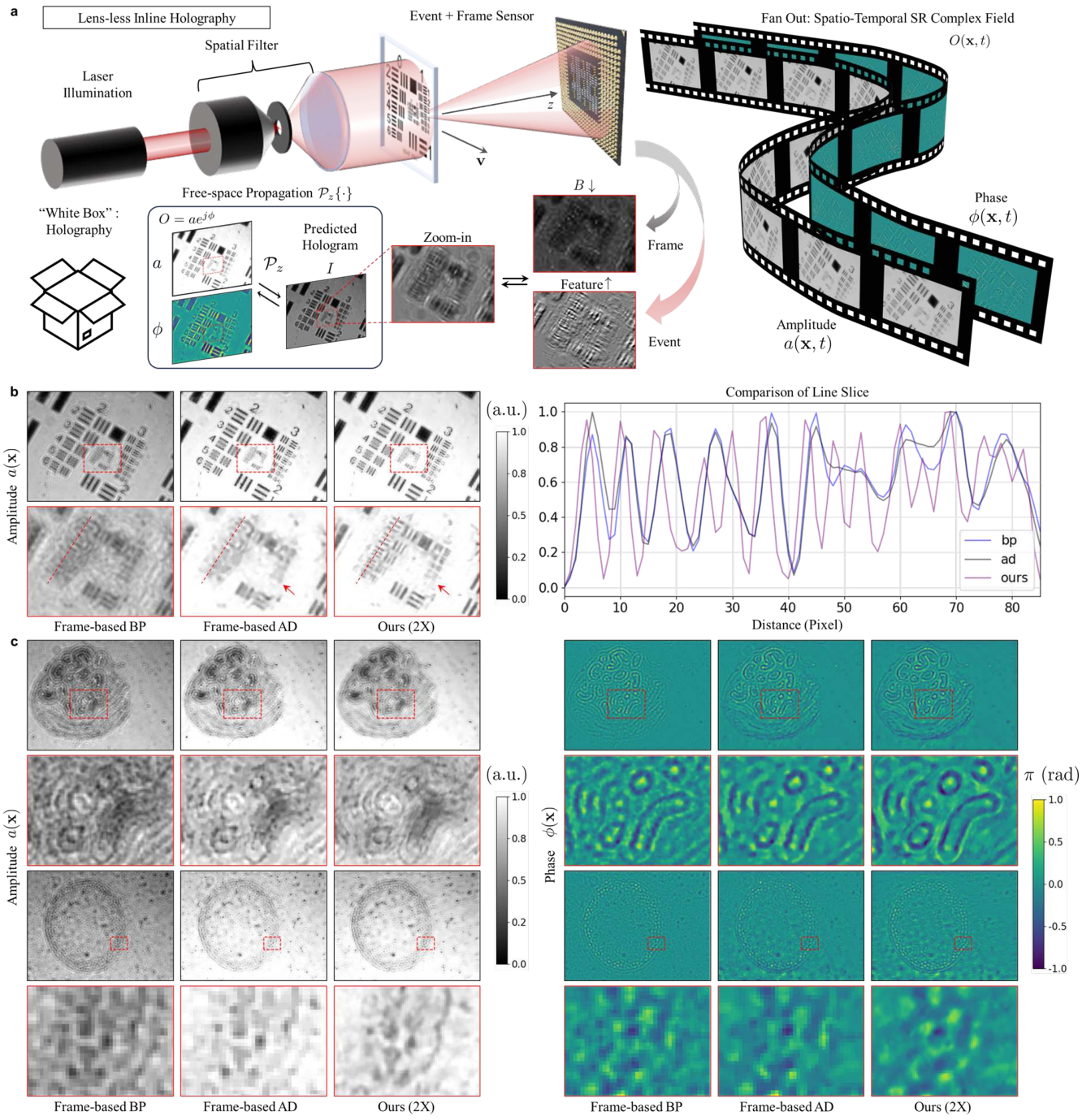


Fig. 3. Snapshot pixel-SR inline holography with a single frame. a, Experimental configuration with transmissive laser illumination and event-frame fusion sensor. **b,** Reconstruction pipeline based on motion-aligned structure alignment with hologram gradients. **c,** USAF target Refocusing. NeuroSR resolves finer structures across depths and retains contrast under overexposure. **d,** Complex field recovery. NeuroSR restores double-layer membranes in testis tissue and porous structures in monocotyledons, substantially outperforming the baseline backpropagation (BP) and the advanced twin-image-free method via automatic differentiation (AD).

double-layer phase structures in testis tissue that are otherwise suppressed by lack of sampling and noisy background in conventional reconstructions, enabled by the combined effect of physically informed regularization and improved gradient fidelity. In Figure 3(d) (bottom), the proposed method successfully recovers porous cellular structures in monocotyledon

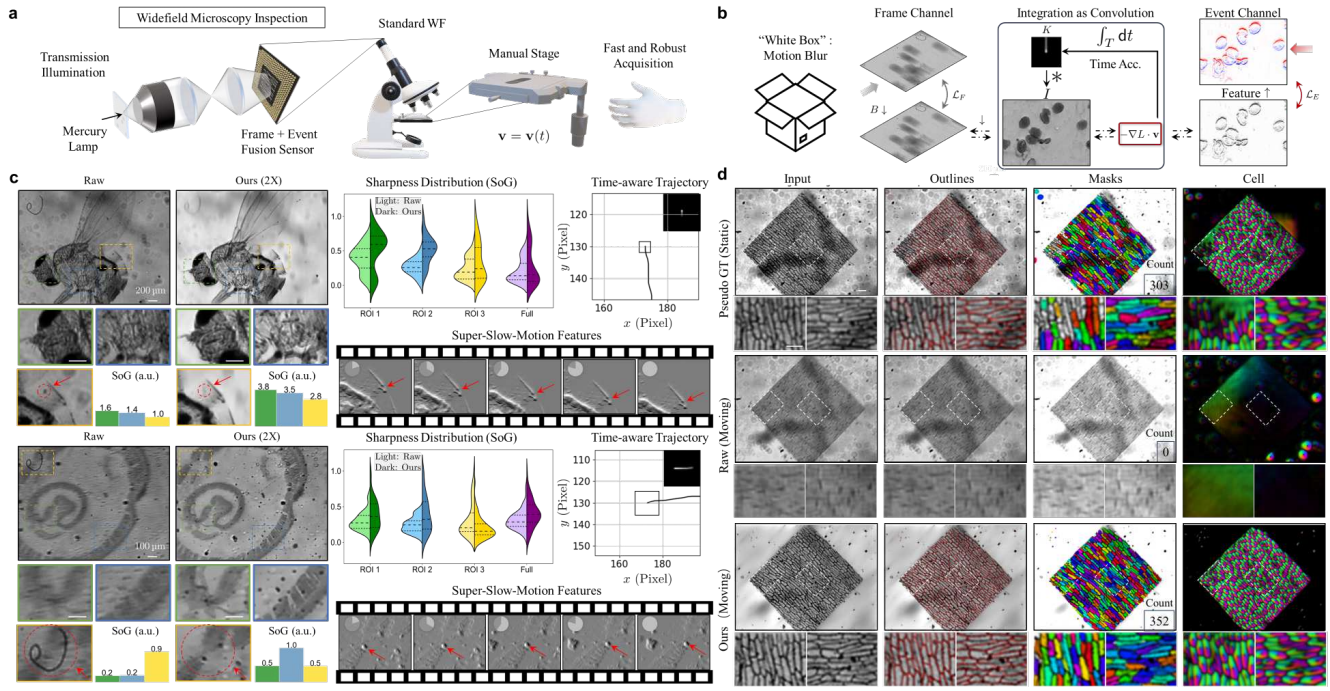


Fig. 4. High-throughput WF microscopy enhanced by motion-aware super-resolution. a, Simplified hardware setup using event-frame fusion and manual scanning. **b,** Reconstruction pipeline with asynchronously aligned HF gradients via motion-feature joint estimation. **c,** Representative reconstructions of biological samples. ROI 1 and 2 highlight sharpness enhancement, quantified by the sparsity of the gradient (SoG) and its distribution across the full FOV. ROI 3 demonstrates recovery of features occluded in the raw image, enabled by the temporally dense nature of IWEs, which is inherently resilient to static obstructions. **d,** Downstream segmentation on onion epidermal cells. NeuroSR restores detection precision and pose consistency from the destructive blurry case. The enhanced sub-pixel edges yield an evident 16.5 % increase in cell count compared to the static GT.

specimens and maintains high contrast despite static speckle background in the illumination path with 2× super-resolution. Together, these results demonstrate that transport-induced differential sensing provides additional constraints that stabilize joint amplitude and phase recovery from bandwidth-limited hologram measurements, enabling robust snapshot reconstruction under defocus, exposure saturation, and noise.

Generalization to high-throughput biological microscopy

Finally, we demonstrate that NeuroSR generalizes to high-throughput wide-field microscopy, enabling motion- and occlusion-robust imaging by replacing high-frame-rate temporal redundancy with asynchronous differential constraints under a fixed optical spatial bandwidth product[52] (see Supplementary Video 3). This regime is central to high-content cell screening and pose inspection, where large-scale data acquisition must be performed rapidly without sacrificing spatial resolution[53,54]. In conventional wide-field microscopy, the coupling between objective resolution and field of view fundamentally constrains throughput, often requiring longer exposure times to mitigate photon starvation at the cost of motion blur[55]. Classical mitigation strategies, such as stop-and-stare or strobed imaging, alleviate blur but require precise stage control and hardware synchronization, and may introduce phototoxicity or photobleaching. By contrast, we show that combining a single low-frame-rate intensity image with microsecond-resolved event measurements enables at least a two-fold increase in effective acquisition throughput under continuous sweeping and moderate illumination while preserving super-resolved structural detail.

As illustrated in Figure 4(a), we employ a cost-effective wide-field microscopy configuration that replaces the conventional sensor with an event-frame fusion sensor and eliminates the need for precisely controlled motorized scanning by multiplexing a manual translation stage in a regular WF setup. During continuous sweeping, sub-pixel displacements are captured by asynchronous events, which encode high-temporal-resolution motion cues within millisecond-level exposure in a purely physics-based manner[56,57]. Because the sweep spans the entire field of view, the event stream provides sufficient motion information to transform an otherwise ill-posed blind deblurring problem into a well-conditioned reconstruction task, as schematically illustrated in Figure 4(b), with formal derivations deferred to the Methods. Leveraging this differential sensing

constraint, NeuroSR yields a two-to-five-fold sharpness enhancement and suppresses background clutter within event-dominant regions (blue and green boxes, ROI 1 and 2 in Figure 4(c)). Importantly, the temporally distributed accumulation of event renders the reconstruction robust to static occlusions, such as dust or contaminants on the optical path (yellow boxes, ROI 3): such artifacts are confined to isolated timestamps and therefore do not accumulate coherently in motion-aligned temporal gradients, as further evidenced by the temporally-resolved features in the occlusion area, shown in Figure 4(c). By exploiting dense temporal sampling over the exposure window, NeuroSR recovers structural details that are partially obscured in the raw snapshot, as demonstrated on Drosophila cross-section and earthworm transverse-section samples. This recovered structural fidelity transfers directly to downstream analysis: for cell pose estimation we employ Cellpose-SAM[58], which is selected for its strong segmentation performance[59–61]. Results of onion epidermal cells in Figure 4(d). show that the destructive blurry case (with detection rate 0) is resolved; comparison with a static GT frame in the zoomed-in ROI (white boxes) reveals finer pose textures and improved morphological consistency, yielding sharper segmentation and an increased cell count (352 NeuroSR vs. 303 static, a 16.5 % gain). These results confirm that temporally dense differential sensing improves structural observability in high-throughput microscopy and directly benefits downstream quantitative analysis.

Discussion

NeuroSR reformulates high-throughput computational imaging by interpreting events as direct physical measurements of photometric gradient information. In several computational imaging frameworks, including Fourier ptychography and holography, practical deviations from ideal photometry such as illumination non-uniformity and measurement noise have motivated performing inference in feature domains, for example through logarithmic gradient representations that stabilize alignment and strengthen phase-retrieval constraints[33,62]. However, in high-throughput imaging where relative motion is inherent or unavoidable, such feature-domain quantities are still numerically derived from temporally integrated intensity measurements and therefore degrade along the motion direction as blur and noise suppress high-frequency structures under finite spatio-temporal bandwidth. Event-based sensing introduces a complementary measurement mechanism by directly recording logarithmic intensity increments with microsecond latency, preserving motion-aligned differential information that is attenuated in frame-based acquisition. Under a fixed sensing bandwidth, synchronous integration necessarily concentrates measurements on low-frequency intensity components, resulting in an irreversible loss of motion-induced differential information. The complementary sensing characteristics of frame and event measurements are theoretically analyzed via Cramér–Rao lower bounds in Supplementary Note 4.

An important implication of this formulation is that asynchronous event measurements and synchronous frame measurements can be embedded within a unified physical twin that remains differentiable with respect to structure, motion, and system parameters[63,64]. By expressing event generation and optical image formation through a shared forward model, NeuroSR enables self-consistent refinement of latent variables within a single computational graph rather than relying on sequential estimation or external calibration. This perspective aligns with recent developments in differentiable optics and computational imaging, where physical image formation models are treated as optimizable components rather than fixed operators[34,65–67]. NeuroSR extends this paradigm by incorporating asynchronous differential measurements directly into the physics-consistent twin, enabling closed-loop optimization that improves both computational efficiency and sensing efficiency under bandwidth-limited acquisition.

More broadly, neuromorphic sensing has been explored across diverse domains that demand high-throughput perception, including autonomous driving, microscopy, wavefront sensing, and microfluidic analysis[19,22,68–71]. These studies have largely emphasized perception, control, or reactive estimation, treating event streams primarily as temporal cues for decision-making. By contrast, NeuroSR interprets events as direct physical measurements that can be embedded within model-based inference frameworks. This perspective suggests a broader paradigm in which neuromorphic measurements act not merely as an alternative sensing modality but as complementary physical constraints within computational imaging systems. Recent work has shown that neural signals originating from retinal circuits, when interpreted through models of fixational eye motion, can be decoded to reconstruct detailed visual scenes[11]. Embedding computational imaging principles into such neuromorphic measurement pipelines raises the possibility that physical inference tasks traditionally associated with optical systems—such as ghost imaging or non-line-of-sight imaging—may ultimately be realized through efficient neural or hybrid bio-artificial sensing architectures[72,73].

Method

Experimental configuration

To evaluate NeuroSR across representative high-throughput imaging scenarios, we implemented a unified dual-modality sensing pipeline that combines low-frame-rate intensity acquisition with asynchronous event sensing under varied motion controllability. Across all configurations, sensing is performed using a monolithic event-frame fusion sensor (DAVIS346,

iniVation) with a spatial resolution of 346×260 pixels and an 18.5 µm pixel pitch. This architecture provides simultaneous synchronous and asynchronous outputs: a frame-based active pixel sensor (APS) recording global intensity at up to 40 fps with an effective dynamic range of approximately 40 dB, and a neuromorphic event channel reporting per-pixel intensity changes with 1 µs temporal resolution and a dynamic range exceeding 120 dB. The shared focal plane ensures precise spatio-temporal co-registration without the parallax or synchronization errors inherent in multi-sensor configurations. In the industrial inspection configuration, a bright-field $4f$ relay system with unit magnification images opaque samples, including a 1951 USAF resolution etalon and semiconductor workpieces, under uniform broadband LED illumination. The specimen is mounted on a high-precision linear translation stage (WN262TA20, up to 5 mm s^{-1}, 2 µm repeatability) for continuous sweeping; in practice, residual mechanical vibration introduces a lateral positional uncertainty on the order of 40 µm. For biological microscopy, the sensing pipeline is integrated into a wide-field microscope (Eclipse Ni-U, Nikon) under transmissive halogen illumination, using a manual translation stage to intentionally introduce stochastic velocity fluctuations and trajectory non-uniformity. Similarly, the holography configuration adopts a lensless inline setup with a 632.8 nm HeNe laser (HNL100L, Thorlabs), where biological specimens such as testis tissue and monocotyledon sections are manually translated to induce transverse motion. In this unified acquisition strategy, the APS channel provides a global photometric reference, while the event channel captures motion-induced temporal variations, enabling consistent evaluation of NeuroSR across relay imaging, microscopy, and lensless holography.

Sub-pixel transport-induced differential sensing and hybrid observation model

We model the event channel as a motion-aligned feature constraint by starting from the asynchronous triggering principle of an event camera[18]. Within a spatiotemporal window $\Omega \times T$ with $T = [t_0, t_0 + \Delta t]$, the sensor outputs an event set $\mathcal{E} = \{e_k\}_{k=1}^{N_e}$, where each $e_k = (\mathbf{x}_k, p_k, t_k)$ is generated at pixel $\mathbf{x}_k = (x_k, y_k)$ when the brightness increment since the last event at the same pixel reaches a fixed temporal contrast threshold c, *i.e.*,

$$\Delta L(\mathbf{x}_k, t_k) \triangleq L(\mathbf{x}_k, t_k + \Delta t_k) - L(\mathbf{x}_k, t_k) = p_k c, \qquad p_k \in \{-1, +1\}, \tag{3}$$

with $L \triangleq \log I$ and Δt_k denoting the inter-event interval at the same pixel location $\mathbf{x}_k$. See Supplementary Note 2. Under the brightness constancy assumption used in Equation (1), temporal variation is induced by relative motion and satisfies the first-order relation $\frac{\partial L}{\partial t}(\mathbf{x}, t) \approx -\nabla L(\mathbf{x}, t) \cdot \mathbf{v}(t)$[74,75], so that over a short interval a finite-difference approximation yields $\Delta L(\mathbf{x}; \Delta t) \approx -(\nabla L \cdot \mathbf{v})\Delta t$; this is the key mechanism by which events encode the motion-projected spatial gradient feature $\nabla L \cdot \mathbf{v}$ with subpixel sensitivity when $\|\mathbf{v}\Delta t_k\| < 1$ pixel. Practically, directly accumulating raw events over a fixed temporal window provides only a coarse estimate of this feature because it exhibits an intrinsic trade-off between low SNR and motion blur: short windows are dominated by stochastic triggering noise, whereas longer windows mix contributions along the motion trajectory and induce effective blur. Since the photometric distribution and optical flow are intrinsically coupled, accurate motion estimation is essential for reliable gradient feature extraction. We therefore adopt the contrast maximization framework[39,57]. Projecting and aggregating the transformed event cluster onto the image plane produces an image of warped events (IWE)[57], *i.e.*,

$$\frac{\Delta \hat{L}(\mathbf{x}; \Delta t)}{c} = \sum_{k=1}^{N_e} p_k \, \delta\big(\mathbf{x} - \mathbf{x}'_k(\mathbf{v})\big) \;\approx\; \sum_{k=1}^{N_e} p_k \, \kappa_\Sigma\big(\mathbf{x} - \mathbf{x}'_k(\mathbf{v})\big) \;=\; \mathrm{IWE}(\mathbf{x}; \mathbf{v}), \tag{4}$$

where $\delta(\cdot)$ is the Dirac delta and $\kappa_\Sigma(\cdot)$ denotes its differentiable Gaussian approximation with covariance Σ. In practice, the IWE can be rendered using standard isotropic kernels (e.g., Dirac, bilinear, or Gaussian) or anisotropic ones (See Supplementary Note 6), which yield subpixel-precise and differentiable event measurements. This motion compensation step makes the subpixel displacement explicit: the warped coordinates $\mathbf{x}'_k$ are continuous-valued, so the aggregated IWE represents a subpixel-resolved, motion-aligned gradient feature map rather than a blurred event frame. Following the linear inverse formulation of event-based reconstruction with optical flow[41], we define the event-channel measurement $\mathbf{b}_E(\mathbf{x}; \mathbf{v})$ as a standard IWE observation scaled by the contrast threshold c and model it as a feature-domain constraint in the logarithmic intensity space,

$$\mathbf{b}_E(\mathbf{x}; \mathbf{v}) \triangleq \frac{c}{\Delta t} \mathrm{IWE}(\mathbf{x}; \mathbf{v}) = -\nabla L(\mathbf{x}) \cdot \mathbf{v} + \mathbf{n}_E, \tag{5}$$

where $\mathbf{n}_E$ captures the effective event noise after temporal aggregation. See Supplementary Note 3.

To explicitly align the frame-based measurement with the continuous motion dynamics encoded by events, we model frame formation as temporal integration of relative motion over the exposure interval $T = [0, \Delta t]$. In the general case, such integration induces a spatially varying aggregation of intensity that cannot be expressed as a global convolution; however, under the sweep-based inspection and small-field regimes considered in this work, wherein relative motion is approximately uniform over the support of local structures, the frame observation admits a compact convolutional form. Let $I(\mathbf{x})$ denote the latent sharp intensity at the reference timestamp $t_0 = 0$. Under this locally uniform motion assumption, the instantaneous intensity

can be written as a motion-shifted version of the reference image, and the frame measurement corresponds to the temporal integration of this motion-induced intensity evolution over the exposure window,

$$\begin{aligned} B(\mathbf{x}) &= \frac{1}{\Delta t}\int_T I(\mathbf{x},t)\,\mathrm{d}t \approx \frac{1}{\Delta t}\int_T I\big(\mathbf{x}-\mathbf{v}(t)\big)\,\mathrm{d}t \\ &= I(\mathbf{x}) * \left(\frac{1}{\Delta t}\int_T \delta\big(\mathbf{x}-\mathbf{v}(t)\big)\,\mathrm{d}t\right), \end{aligned} \tag{6}$$

where $*$ is the convolution operation, and $K = \frac{1}{\Delta t}\int_T \delta\big(\mathbf{x}-\mathbf{v}(t)\big)\,\mathrm{d}t$ is denoted as a kernel and represents the projection of the relative motion trajectory onto the sensor plane and serves as an effective motion-integration operator[36,76]. This formulation compactly captures how photometric information is accumulated over time along the motion path, irrespective of scene support. Together with the inherent spatial sampling imposed by the sensor grid, the frame observation is modeled as

$$\mathbf{b}_F(\mathbf{x}) \triangleq B(\mathbf{x})\downarrow + \mathbf{n}_F, \tag{7}$$

where $\downarrow$ denotes sampling to the sensor lattice and $\mathbf{n}_F$ represents measurement noise.

Throughout this work, event measurements are written with explicit dependence on the motion field, while frame measurements depend on motion only through the forward model, ensuring a physically consistent formulation without ambiguity during inference. Importantly, the same motion trajectory $\mathbf{v}(t)$ governs both the frame integration here and the event warping in Equation (4), thereby establishing a consistent hybrid observation model for joint inference. The schematic of the full workflow, including sub-pixel event warping and IWE rendering, is summarized in Supplementary Figure S1.

Differentiable parameterization of motion and optical models

From the perspective of manifold hypotheses, motion cues often reside on a low-dimensional latent space rather than occupying the full spatiotemporal degrees of freedom[77,78]. Under the sweep-based acquisition condition considered in this work, we reparameterize the timestamp-wise optical flow $\mathbf{v}(t_k)$ as a temporally piecewise-linear function, yielding a compact representation that is well conditioned for subsequent inference. Specifically, the temporal interval $[t_0, t_0+\Delta t]$ is uniformly divided into M segments, and each temporal node t_m is associated with a two-dimensional flow vector $\mathbf{v}_m \in \mathbb{R}^2$. These nodes form the motion parameterization $\mathbf{v} = [\mathbf{v}_1,\ldots,\mathbf{v}_M] \in \mathbb{R}^{2\times M}$. For each event timestamp $t_k \in [t_m, t_{m+1}]$, the instantaneous flow is obtained by linear interpolation between adjacent nodes,

$$\mathbf{v}(t_k) = \alpha_k\,\mathbf{v}_m + (1-\alpha_k)\,\mathbf{v}_{m+1}, \tag{8}$$

where the interpolation weight $\alpha_k = \frac{t_{m+1}-t_k}{t_{m+1}-t_m} \in [0,1]$ reflects the temporal proximity to the neighboring nodes.

Beyond motion, NeuroSR adopts a differentiable imaging formulation to reparameterize the physical image formation process itself. To accommodate different computational imaging modalities within a unified framework, the forward operator is expressed in a generic form $f_\Theta(\cdot)$, which maps a physically parameterized object representation to measurable intensities while explicitly exposing task-dependent system parameters.

In optical field recovery problems, the object is naturally represented as a complex field $O(\mathbf{x}) = a(\mathbf{x})\exp\big(j\phi(\mathbf{x})\big)$, where $a(\mathbf{x})$ and $\phi(\mathbf{x})$ denote the amplitude and phase, respectively. The physical forward operator then corresponds to wave propagation followed by intensity detection. Inline digital holography serves as a canonical example in this work. Under a scalar diffraction model, the complex field propagated from the object plane to the sensor plane at axial distance z is governed by free-space propagation. As such, the forward model that maps the object to the hologram intensity $I(\mathbf{x})$ at the sensor plane can be compactly written as

$$I(\mathbf{x}) = f_\Theta(O) = \big|\mathcal{P}_z\{O\}\big|^2, \tag{9}$$

where $\Theta = \{z\}$ in this holography instantiation (with other geometric and sampling parameters calibrated), and $\mathcal{P}_z$ denotes the free-space propagation operator implemented via Fourier optics[79] and is expressed as

$$\mathcal{P}_z\{O\}(\mathbf{x}) = (O * h_z)(\mathbf{x}) = \mathscr{F}^{-1}\{\mathscr{F}\{O(\mathbf{x})\}H_z(\mathbf{f})\}, \tag{10}$$

where $\mathscr{F}$ denotes the Fourier transform, $\mathbf{f} = (f_x, f_y)$ is the spatial frequency coordinate, h_z is the free-space propagation kernel, and $H_z = \mathscr{F}\{h_z\}$ is its transfer function. Under the Rayleigh-Sommerfeld diffraction theory, the propagation transfer function takes the form

$$H_z(\mathbf{f}) = \begin{cases} \exp\left(-j\dfrac{2\pi z}{\lambda}\sqrt{1-\lambda^2|\mathbf{f}|^2}\right), & |\mathbf{f}| < \dfrac{1}{\lambda}, \\ 0, & \text{otherwise}, \end{cases} \tag{11}$$

where λ denotes the illumination wavelength.

To enable joint optimization under the unified NeuroSR objective in Equation (2), both the event rendering process and the optical forward model are implemented as differentiable operators with respect to their parameters. Because each warped event location depends explicitly on the reparameterized motion field, differentiable rendering of the image of warped events is required to propagate gradients from the reconstruction objective to the motion parameters $\mathbf{v}$. Likewise, expressing optical image formation as $f_\Theta(\cdot)$ makes physically meaningful system parameters, *e.g.*, the propagation distance z in Equation (9), explicit and differentiable, allowing parameter refinement within the same computational graph[34]. Modern automatic differentiation frameworks support these requirements and enable complex-valued optimization via Wirtinger calculus[80].

References

1. Yang, Z. *et al.* A vision chip with complementary pathways for open-world sensing. *Nature* **629**, 1027–1033 (2024).
2. Nature Editorial. Big data needs a hardware revolution. *Nature* **554**, 145–146 (2018).
3. Roy, K., Jaiswal, A. & Panda, P. Towards spike-based machine intelligence with neuromorphic computing. *Nature* **575**, 607–617 (2019).
4. Martinez-Conde, S., Macknik, S. L. & Hubel, D. H. The role of fixational eye movements in visual perception. *Nat. reviews neuroscience* **5**, 229–240 (2004).
5. Rucci, M., Iovin, R., Poletti, M. & Santini, F. Miniature eye movements enhance fine spatial detail. *Nature* **447**, 852–855 (2007).
6. Intoy, J. & Rucci, M. Finely tuned eye movements enhance visual acuity. *Nat. Commun.* **11**, 795 (2020).
7. Barlow, H. Reconstructing the visual image in space and time. *Nature* **279**, 189–190 (1979).
8. Poggio, T., Fahle, M. & Edelman, S. Fast perceptual learning in visual hyperacuity. *Science* **256**, 1018–1021 (1992).
9. Jackson, R. R. & Harland, D. P. One small leap for the jumping spider but a giant step for vision science. *J. Exp. Biol.* **212**, 2129–2132 (2009).
10. Patrick, J. A., Roach, N. W. & McGraw, P. V. Motion-based super-resolution in the peripheral visual field. *J. Vis.* **17**, 15–15 (2017).
11. Wu, E. G. *et al.* Fixational eye movements enhance the precision of visual information transmitted by the primate retina. *Nat. Commun.* **15**, 7964 (2024).
12. Chan, K. C., Zhou, S., Xu, X. & Loy, C. C. Basicvsr++: Improving video super-resolution with enhanced propagation and alignment. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 5972–5981 (2022).
13. Bhat, G., Danelljan, M., Van Gool, L. & Timofte, R. Deep burst super-resolution. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 9209–9218 (2021).
14. Kellman, M., Chen, M., Phillips, Z. F., Lustig, M. & Waller, L. Motion-resolved quantitative phase imaging. *Biomed. optics express* **9**, 5456–5466 (2018).
15. Gao, Y. & Cao, L. Motion-resolved, reference-free holographic imaging via spatiotemporally regularized inversion. *Optica* **11**, 32–41 (2024).
16. Schermelleh, L. *et al.* Super-resolution microscopy demystified. *Nat. cell biology* **21**, 72–84 (2019).
17. Brandli, C., Berner, R., Yang, M., Liu, S.-C. & Delbruck, T. A 240 $\times$ 180 130 db 3 μs latency global shutter spatiotemporal vision sensor. *IEEE J. Solid-State Circuits* **49**, 2333–2341 (2014).
18. Gallego, G. *et al.* Event-based vision: A survey. *IEEE transactions on pattern analysis machine intelligence* **44**, 154–180 (2020).
19. Gehrig, D. & Scaramuzza, D. Low-latency automotive vision with event cameras. *Nature* **629**, 1034–1040 (2024).
20. He, B. *et al.* Microsaccade-inspired event camera for robotics. *Sci. Robotics* **9**, eadj8124 (2024).
21. Mangalwedhekar, R. *et al.* Achieving nanoscale precision using neuromorphic localization microscopy. *Nat. Nanotechnol.* **18**, 380–389 (2023).
22. Cabriel, C., Monfort, T., Specht, C. G. & Izeddin, I. Event-based vision sensor for fast and dense single-molecule localization microscopy. *Nat. Photonics* **17**, 1105–1113 (2023).
23. Jing, Y., Yang, Y., Wang, X., Song, M. & Tao, D. Turning frequency to resolution: Video super-resolution via event cameras. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 7772–7781 (2021).

24. Kai, D., Zhang, Y. & Sun, X. Video super-resolution via event-driven temporal alignment. In *2023 IEEE International Conference on Image Processing (ICIP)*, 2950–2954 (2023).

25. Kai, D., Lu, J., Zhang, Y. & Sun, X. Evtexture: event-driven texture enhancement for video super-resolution. In *Forty-first International Conference on Machine Learning* (2024).

26. Guo, R. *et al.* EventLFM: event camera integrated fourier light field microscopy for ultrafast 3d imaging. *Light. Sci. & Appl.* **13**, 144 (2024).

27. Zhu, S., Ge, Z., Wang, C., Han, J. & Lam, E. Y. Efficient non-line-of-sight tracking with computational neuromorphic imaging. *Opt. Lett.* **49**, 3584–3587 (2024).

28. Zhu, S. *et al.* Neuromorphic encryption: combining speckle correlography and event data for enhanced security. *Adv. Photonics Nexus* **3**, 056002–056002 (2024).

29. Ge, Z., Wang, C., Huang, J. & Lam, E. Y. Event-driven neuromorphic holography for dynamic particle imaging. *Opt. Lett.* **50**, 1496–1499 (2025).

30. Barbastathis, G., Ozcan, A. & Situ, G. On the use of deep learning for computational imaging. *Optica* **6**, 921–943 (2019).

31. Wang, F. *et al.* Phase imaging with an untrained neural network. *Light. Sci. & Appl.* **9**, 77 (2020).

32. Wang, K. *et al.* On the use of deep learning for phase recovery. *Light. Sci. & Appl.* **13**, 4 (2024).

33. Zhang, S., Pan, A., Sun, H., Tan, Y. & Cao, L. High-Fidelity Computational Microscopy via Feature-Domain Phase Retrieval. *Adv. Sci.* 2413975 (2025).

34. Chen, N., Brady, D. J. & Lam, E. Y. Differentiable imaging: Progress, Challenges, and Outlook. *Adv. Devices & Instrumentation* (2025).

35. Chen, N. *et al.* Uncertainty-aware Fourier ptychography. *Light. Sci. & Appl.* **14**, 236 (2025).

36. Zhang, P., Zhu, S., Wang, C., Zhao, Y. & Lam, E. Y. Neuromorphic imaging with super-resolution. *IEEE Transactions on Circuits Syst. for Video Technol.* (2024).

37. Li, S. *et al.* Event stream super-resolution via spatiotemporal constraint learning. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 4480–4489 (2021).

38. Wang, L., Kim, T.-K. & Yoon, K.-J. Eventsr: From asynchronous events to image reconstruction, restoration, and super-resolution via end-to-end adversarial learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 8315–8325 (2020).

39. Gallego, G., Rebecq, H. & Scaramuzza, D. A unifying contrast maximization framework for event cameras, with applications to motion, depth, and optical flow estimation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 3867–3876 (2018).

40. Gallego, G., Gehrig, M. & Scaramuzza, D. Focus is all you need: Loss functions for event-based vision. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 12280–12289 (2019).

41. Zhang, Z., Yezzi, A. J. & Gallego, G. Formulating event-based image reconstruction as a linear inverse problem with deep regularization using optical flow. *IEEE Transactions on Pattern Analysis Mach. Intell.* **45**, 8372–8389 (2022).

42. Orji, N. G. *et al.* Metrology for the next generation of semiconductor devices. *Nat. electronics* **1**, 532–547 (2018).

43. Zhu, S., Yin, Q., Wang, C., Huang, J. & Lam, E. Y. Ultrafast dynamic defect inspection with computational neuromorphic imaging. *Adv. Sci.* **12**, e10338 (2025).

44. Shimizu, Y. *et al.* An insight into optical metrology in manufacturing. *Meas. Sci. Technol.* **32**, 042003 (2021).

45. Aly, H. A. & Dubois, E. Image up-sampling using total-variation regularization with a new observation model. *IEEE Transactions on Image Process.* **14**, 1647–1659 (2005).

46. Yue, L. *et al.* Image super-resolution: The techniques, applications, and future. *Signal processing* **128**, 389–408 (2016).

47. Bishara, W., Su, T.-W., Coskun, A. F. & Ozcan, A. Lensfree on-chip microscopy over a wide field-of-view using pixel super-resolution. *Opt. express* **18**, 11181–11191 (2010).

48. Luo, W., Zhang, Y., Feizi, A., Göröcs, Z. & Ozcan, A. Pixel super-resolution using wavelength scanning. *Light. Sci. & Appl.* **5**, e16060–e16060 (2016).

49. Chen, N. & Lam, E. Y. Differentiable pixel-super-resolution lensless imaging. *Opt. Lett.* **50**, 1180–1183 (2025).

50. Zhang, W. *et al.* Twin-image-free holography: a compressive sensing approach. *Phys. review letters* **121**, 093902 (2018).

51. Chen, N., Wang, C. & Heidrich, W. ∂ H: Differentiable Holography. *Laser & Photonics Rev.* **17**, 2200828 (2023).

52. Lohmann, A. W., Dorsch, R. G., Mendlovic, D., Zalevsky, Z. & Ferreira, C. Space–bandwidth product of optical signals and systems. *J. Opt. Soc. Am. A* **13**, 470–473 (1996).

53. Zanella, F., Lorens, J. B. & Link, W. High content screening: seeing is believing. *Trends biotechnology* **28**, 237–245 (2010).

54. Boutros, M., Heigwer, F. & Laufer, C. Microscopy-based high-content screening. *Cell* **163**, 1314–1325 (2015).

55. Phillips, Z. F., Dean, S., Recht, B. & Waller, L. High-throughput fluorescence microscopy using multi-frame motion deblurring. *Biomed. optics express* **11**, 281–300 (2019).

56. Shiba, S., Aoki, Y. & Gallego, G. Secrets of event-based optical flow. In *European Conference on Computer Vision*, 628–645 (2022).

57. Shiba, S., Klose, Y., Aoki, Y. & Gallego, G. Secrets of event-based optical flow, depth and ego-motion estimation by contrast maximization. *IEEE Transactions on Pattern Analysis Mach. Intell.* **46**, 7742–7759 (2024).

58. Pachitariu, M., Rariden, M. & Stringer, C. Cellpose-sam: superhuman generalization for cellular segmentation. *bioRxiv* 2025–04 (2025).

59. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: a generalist algorithm for cellular segmentation. *Nat. Methods* **18**, 100–106 (2021).

60. Pachitariu, M. & Stringer, C. Cellpose 2.0: how to train your own model. *Nat. Methods* **19**, 1634–1641 (2022).

61. Stringer, C. & Pachitariu, M. Cellpose3: one-click image restoration for improved cellular segmentation. *Nat. Methods* 1–8 (2025).

62. Zhang, S. *et al.* FPM-WSI: Fourier ptychographic whole slide imaging via feature-domain backdiffraction. *Optica* **11**, 634–646 (2024).

63. Wang, C., Chen, N. & Heidrich, W. do: A differentiable engine for deep lens design of computational imaging systems. *IEEE Transactions on Comput. Imaging* **8**, 905–916 (2022).

64. Chen, N., Cao, L., Poon, T.-C., Lee, B. & Lam, E. Y. Differentiable imaging: A new tool for computational optical imaging. *Adv. Phys. Res.* **2**, 2200118 (2023).

65. Yang, X., Fu, Q. & Heidrich, W. Curriculum learning for ab initio deep learned refractive optics. *Nat. Commun.* **15**, 6572 (2024).

66. Wang, Z., Peng, Y., Fang, L. & Gao, L. Computational optical imaging: on the convergence of physical and digital layers. *Optica* **12**, 113–130 (2025).

67. Bian, L. *et al.* Physical twinning for joint encoding-decoding optimization in computational optics: a review. *Light. Sci. & Appl.* **14**, 162 (2025).

68. Wang, C. *et al.* Angle-based neuromorphic wave normal sensing. *Laser & Photonics Rev.* **19**, 2400647 (2025).

69. Shiba, S., Hamann, F., Aoki, Y. & Gallego, G. Event-based background-oriented schlieren. *IEEE transactions on pattern analysis machine intelligence* **46**, 2011–2026 (2023).

70. Lin, S. *et al.* Embodied neuromorphic synergy for lighting-robust machine vision to see in extreme bright. *Nat. Commun.* **15**, 10781 (2024).

71. He, W. *et al.* Neuromorphic-enabled video-activated cell sorting. *Nat. Commun.* **15**, 10792 (2024).

72. Wang, G. & Faccio, D. Computational ghost imaging with the human brain. *Intell. Comput.* **2**, 0014 (2023).

73. Abdulkader, S. N., Atia, A. & Mostafa, M.-S. M. Brain computer interfacing: Applications and challenges. *Egypt. Informatics J.* **16**, 213–230 (2015).

74. Gallego, G., Forster, C., Mueggler, E. & Scaramuzza, D. Event-based camera pose tracking using a generative event model. *arXiv preprint arXiv:1510.01972* (2015).

75. Gehrig, D., Rebecq, H., Gallego, G. & Scaramuzza, D. Eklt: Asynchronous photometric feature tracking using events and frames. *Int. J. Comput. Vis.* **128**, 601–618 (2020).

76. Zhou, C. *et al.* Deblurring low-light images with events. *Int. J. Comput. Vis.* **131**, 1284–1298 (2023).

77. Langdon, C., Genkin, M. & Engel, T. A. A unifying perspective on neural manifolds and circuits for cognition. *Nat. Rev. Neurosci.* **24**, 363–377 (2023).

78. Seung, H. S. & Lee, D. D. The manifold ways of perception. *Science* **290**, 2268–2269 (2000).

79. Goodman, J. W. *Introduction to Fourier optics* (Roberts and Company publishers, 2005), 4 edn.

80. Candes, E. J., Li, X. & Soltanolkotabi, M. Phase retrieval via wirtinger flow: Theory and algorithms. *IEEE Transactions on Inf. Theory* **61**, 1985–2007 (2015).

Data availability

All key data and methods needed to evaluate the conclusions of this work are presented in the main text and Supplementary Information. Additional data can be requested from the corresponding author.

Code availability

The code used to process event and frame data and to reproduce the results in this study will be made publicly available upon publication.

Acknowledgements

This work is supported by the Research Grants Council of Hong Kong (GRF 17201620, T45-701/22-R) and HKU Seed Fund for Basic Research (Seed 109001083, Seed 109002091). We thank Yuzhe Zhang for his insightful discussion.

Author contributions

C.W. conceived the idea, conducted the experimental demonstrations, coordinated the writing process, and wrote the main parts of the manuscript. S.Z., S.L., J.W., Y.Z., and Y.X. provided comprehensive comparisons and extensive supplementary results. N.C. contributed to discussions on physical twinning and provided code support. E.L. supervised the project, provided conceptual guidance and critical feedback on the methodology and manuscript, and acquired funding. All authors reviewed and approved the manuscript.

Correspondence and requests for materials should be addressed to Edmund Y. Lam.

Competing financial interests

C.W., Z.S., J.W., and E.L. are co-inventors of a pending patent application on the presented method. The remaining authors declare no competing interests.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryVideo3.mp4
- SupplementaryVideo2.mp4
- SupplementaryVideo1.mp4
- suppnote.pdf