

High-throughput space-time Fourier ptychography for motile microorganisms: supplement

This document serves as a supplement to "High-throughput space-time Fourier ptychography for motile microorganisms". We first present the space-time Fourier ptychography (ST-FP) optimization framework along with the solver algorithm. We then present detailed characterization and benchmarking of our customized high-power LED illumination, demonstrating improvements in irradiance compared to commercial LED panels. Next, we analyze the limitations of conventional LED panels by visualizing their internal scanning mode and PWM-based brightness modulation. Further, we showcase experimental results using a fabricated grayscale phase target that more closely resembles real biological structures, and evaluate the performance of our system in live dynamic motion of microorganisms (brine shrimp) under different frame rate conditions. Finally, we compare the computational efficiency of our GPU-accelerated auto-differentiable solver with the original CPU-based implementation.

1. SPACE-TIME FOURIER PTYCHOGRAPHY FRAMEWORK

To jointly recover the time-varying object and its underlying motion, we formulate a unified optimization framework that incorporates the physics of the imaging system, motion-aware priors, and regularization constraints across both space and time. This model simultaneously estimates the complex object $\{\mathbf{o}_t\}_t$ and its corresponding motion field $\{\mathbf{v}_t\}_t$, while enforcing consistency with the measured data and promoting physically plausible dynamics. The overall objective function is given by:

$$\begin{aligned} \min_{\{\mathbf{o}_t, \mathbf{v}_t\}_t} \quad & \sum_{t=1}^T \|\mathcal{A}_t(\mathbf{o}_t) - \mathbf{b}_t\|_2^2 + \delta_1 \mathcal{R}_{\{\mathbf{o}_t\}_t}^{s, \mu_1} + \delta_2 \mathcal{R}_{\{\mathbf{o}_t\}_t}^{t, \mu_2} \\ & + \alpha \sum_{t=1}^{T-1} \|\mathbf{D}_t |\mathbf{o}_t| + \mathbf{D}_s |\mathbf{o}_t| \cdot \mathbf{v}_t\|_1 + \beta \mathcal{R}_{\{\mathbf{v}_t\}_t}^{s, \mu_3} \\ & + \gamma \left(\sum_{t=1}^{T-1} \|\mathcal{A}_{t+1}(\text{warp}(\mathbf{o}_t, \mathbf{v}_t)) - \mathbf{b}_{t+1}\|_2^2 + \sum_{t=2}^T \|\mathcal{A}_{t-1}(\text{warp}(\mathbf{o}_t, -\mathbf{v}_{t-1})) - \mathbf{b}_{t-1}\|_2^2 \right), \end{aligned} \quad (\text{S1})$$

where the first term is the data fidelity term that enforces consistency between the reconstructed object \mathbf{o}_t and the corresponding measurement \mathbf{b}_t at each time t , based on the forward model $\mathcal{A}_t(\cdot)$. The regularization terms $\mathcal{R}_{\{\mathbf{o}_t\}_t}^s$ imposes spatial smoothness constraint on the object sequence. $\mathcal{R}_{\{\mathbf{o}_t\}_t}^t$ applies a zero-sum Gaussian kernel along the time axis, enforcing local temporal coherence by suppressing abrupt changes across neighboring frames. This enables adaptive temporal consistency under different illumination conditions across frames [1, 2]. The second line is the optical flow coupling term [3], which encodes a brightness constancy constraint by linking the temporal gradient $\mathbf{D}_t |\mathbf{o}_t|$ to the spatial gradient $\mathbf{D}_s |\mathbf{o}_t|$ modulated by the estimated motion field \mathbf{v}_t . $\mathcal{R}_{\{\mathbf{v}_t\}_t}^s$ is the spatial smoothness regularizer on the flow field. The third line incorporates a bidirectional warp-and-project mechanism [4] to enforce temporal alignment of the reconstructed object across time. Specifically, the object estimate at time t is forward-warped to time $t+1$ using \mathbf{v}_t , then passed through the forward model \mathcal{A}_{t+1} , and compared to the measurement \mathbf{b}_{t+1} . A similar process is applied in the backward direction from t to $t-1$. The smoothness terms are implemented using Huber-norm-based penalties with μ_1, μ_2 , and μ_3 . The weights $\delta_1, \delta_2, \alpha, \beta$, and γ are scalar parameters used to balance the influence of each loss term in the overall objective.

To solve the joint optimization problem in Eq. S1, we adopt an alternating strategy between the object sequence $\{\mathbf{o}_t\}_t$ and the motion field $\{\mathbf{v}_t\}_t$. See more details in our recent work [5]. In contrast to our original implementation, where the object subproblem was solved using ADMM-based approach with explicit splitting and proximal updates, the new solver leverages automatic differentiation through a fully differentiable forward model. Specifically, the object reconstruction

is now optimized using the Adam optimizer [6], with gradients computed and back propagated directly through the computational graph using PyTorch automatic differentiation engine [7]. This shift not only simplifies the implementation but also makes the solver more flexible and extensible, enabling integration with learned components or additional priors in the future.

2. LED OUTPUT POWER ENHANCEMENT

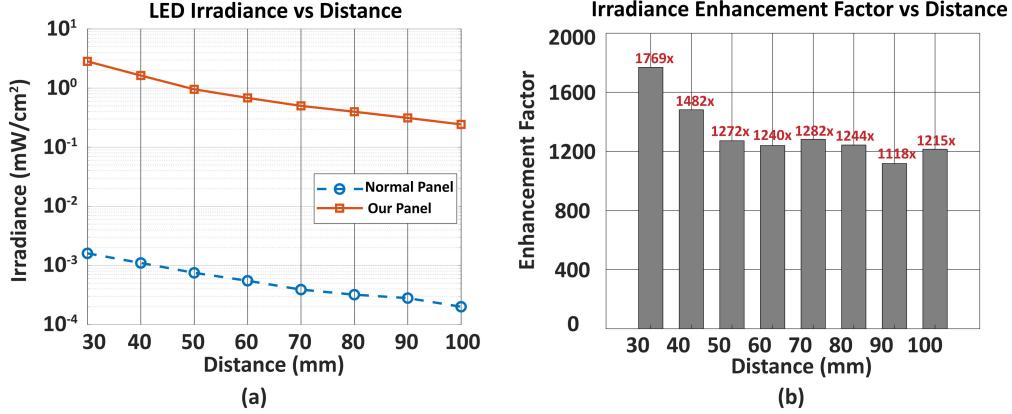


Fig. S1. LED output power comparison. (a) Irradiance of the LED emission measured (mW/cm²) at different distances for a commercial LED panel (Adafruit P4-2121-16S-HL1.0) and our custom-designed panel. (b) Corresponding irradiance enhancement factor, defined as the ratio of our panel's irradiance to that of the commercial panel at each distance.

We employed the Thorlabs PM100D digital optical power and energy meter console, paired with the S120C photodiode power sensor, to measure the irradiance of a single LED from each panel at various distances. Measurements were taken for our custom-designed LED board and a commercial normal panel (Adafruit P4-2121-16S-HL1.0) at distances ranging from 30 mm to 100 mm, as shown in Fig. S1(a). The measured power values were normalized by the sensor area to calculate irradiance (mW/cm²). The results in Fig. S1(b) demonstrate that our LED module consistently provides at least 1100 \times higher effective irradiance than the commercial board across all tested distances. This significant enhancement in optical power density is highly beneficial for high-speed FP imaging, particularly in dynamic scenes requiring fast capture rates, enabling shorter exposure times, improved signal-to-noise ratio (SNR), and robust synchronization performance with the camera.

3. ANALYSIS OF LED PANEL SCANNING AND PWM BRIGHTNESS CONTROL

To investigate the inherent limitations of conventional LED panels in high-speed imaging, we analyzed the panel's internal illumination behavior, which typically combines a column-based scanning mode with pulse-width modulation (PWM)-based brightness control. In such systems, LEDs are not illuminated simultaneously; instead, columns (or rows) are sequentially activated over time. Within each activated column, the brightness of individual LEDs is modulated by varying the duty cycle of their PWM signal. The fraction of time the LED remains on during its activation window. To visualize these effects, we designed an LED pattern where each column contains LEDs with systematically varied brightness levels, as shown in Fig. S2(a). The differences in brightness were programmed via PWM with distinct duty cycles. Fig. S2(b) illustrates two representative PWM signals: a lower duty cycle results in shorter LED on-time and reduced perceived brightness, whereas a higher duty cycle leads to longer on-time and greater brightness.

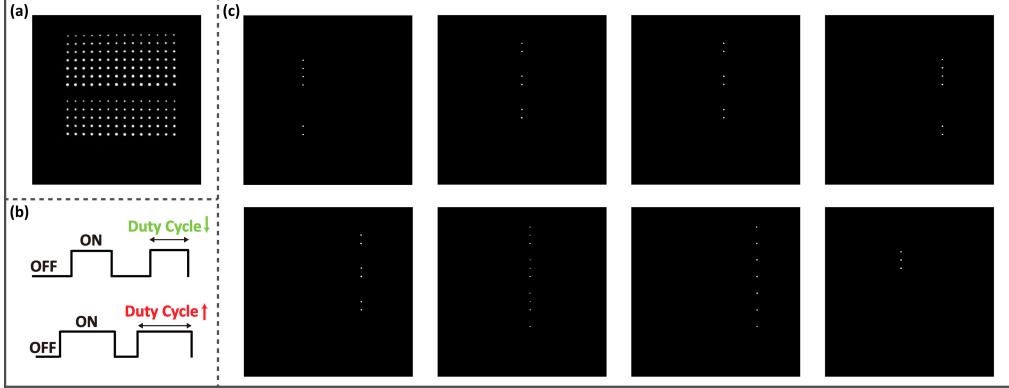


Fig. S2. Visualization of LED scanning mode and pulse-width modulation (PWM)-based brightness control. (a) The designed illumination pattern, where each column contains LEDs set to different brightness levels. (b) Illustration of PWM signals with varying duty cycles: a lower duty cycle (top) results in shorter LED on-time, while a higher duty cycle (bottom) increases the on-time and thus the perceived brightness. (c) Eight representative frames captured using a short exposure time ($10 \mu\text{s}$), demonstrating the scanning behavior of the LED panel. In each frame, only a single column is active due to the panel’s internal column-scanning scheme. Within each active column, LEDs with higher brightness settings are more frequently visible, consistent with their longer PWM duty cycles. This supports the presence of both column-wise scanning and PWM-based brightness modulation.

We captured a sequence of short-exposure images ($10 \mu\text{s}$) without synchronizing the camera to the LED panel. Due to the scanning nature of the panel, only a single column is typically active at any given time. As a result, each captured frame randomly samples the instantaneous state of one column. Figure S2(c) shows eight representative frames. Due to the lack of synchronization between the camera and LED panel, the captured images represent a random sampling of the illumination process. However, in each frame, some LEDs appear illuminated only within a single column, supporting the column-based scanning hypothesis. Not all LEDs in the active column are illuminated in each frame. This behavior suggests that individual LEDs have different duty cycles due to PWM control. LEDs with a higher brightness setting have a greater likelihood of appearing in the captured frames, indicating a longer on-time (higher duty cycle). This pattern confirms that brightness control is achieved via PWM rather than direct intensity modulation.

Based on the fact that scanning mode and PWM effects exist in normal LED panels, stable illumination with minimal artifacts or errors requires setting the LED exposure time to a relatively long duration. This constraint limits the system’s capturing speed. In contrast, we use a customized LED board with independent control of each LED element, effectively eliminating scanning and PWM-related issues. Our system incorporates an enhanced LED panel with significantly higher brightness than conventional off-the-shelf LED boards, enabling rapid and precise imaging of live biological specimens. This design removes the speed limitations imposed by conventional LED panels, shifting the primary performance bottleneck to the image sensor. To further improve efficiency, we utilize an array of LED-controlling shift-register chips, where each pin independently controls an individual LED. This architecture enables simultaneous activation of multiple LEDs based on preloaded patterns, eliminating delays and inefficiencies typically associated with standard LED switching. This design enables truly simultaneous, flicker-free activation of complex patterns at high speed, significantly improving the temporal fidelity of dynamic imaging.

4. EXPERIMENT RESULTS

A. Grayscale Phase Target Results

To extend our evaluation beyond binary phase targets (in the main text Section 3.2) and bridge the gap toward more realistic imaging scenarios, we further evaluated our high-speed ST-FP system using a fabricated grayscale phase target. We used direct-write grayscale lithography with a Heidelberg DWL 66+ mask writer to create high-precision 3D Ascaris eggs. The technique utilizes up to 1023 grayscale levels to precisely control laser exposure, allowing for the creation of

complex topographies in a single exposure step. AZ® 4562, a positive tone photoresist with an excellent dose response, spin-coated in soda-lime glass, soft-baked at 120 °C for 3 minutes, and stabilized for 3 hours. The pattern is then written, and the resist is developed using AZ® 726 MIF for 25 seconds to yield smooth 3D profiles without additional processing steps.

Room-temperature nanoimprint lithography (NIL) is performed using the Obducat Eitre 3 system to replicate the patterned photoresist structures in a durable optical material. The grayscale AZ® 4562 masters serve as molds to transfer patterns to an OrmoComp, a UV-curable hybrid polymer with excellent optical and mechanical properties. A thin layer of OrmoComp resin is dispensed onto a soda-lime glass substrate. The mold is then pressed into it under uniform pressure, and UV light cures the material. This results in a high-fidelity replica of the original 3D microstructure, achieved without thermal processing, reducing stress and cost.

This process produces high-quality grayscale phase structures with smooth profiles ideal for optical testing. To simulate motion, the fabricated sample was mounted on a manual translation stage controlled via the Thorlabs MCM3001 Controller Box, connected to its knob interface. We manually translated the sample along a single lateral axis during acquisition. Note that some minor vertical deviations may be introduced due to the manual operation.

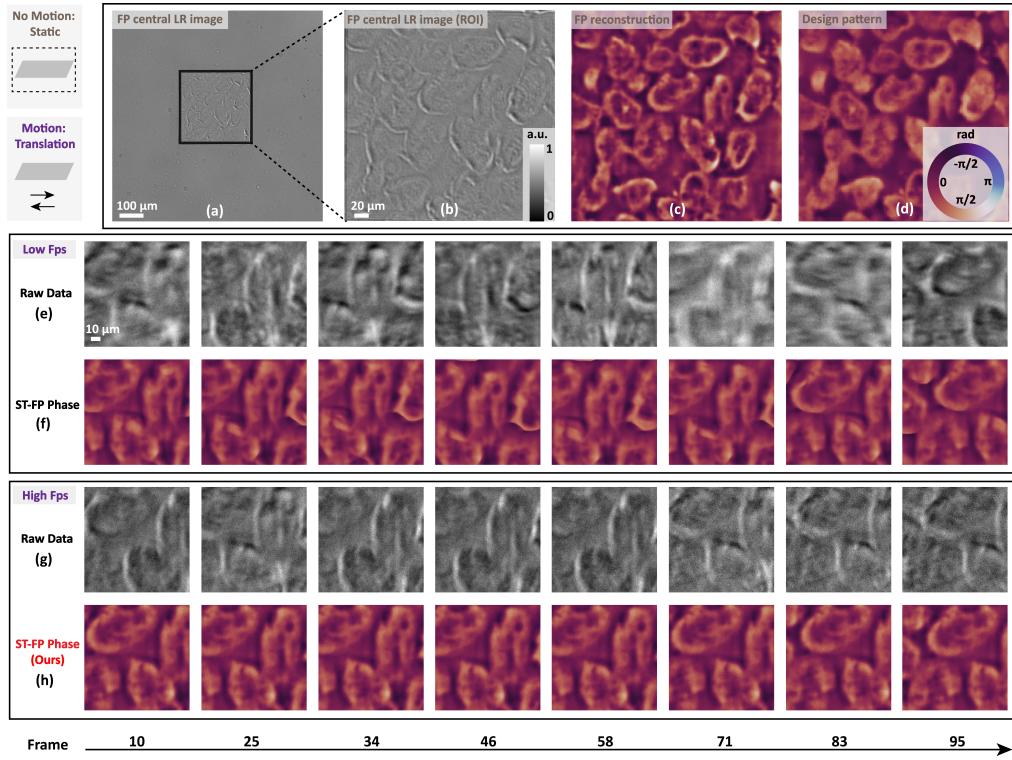


Fig. S3. Comparison of ST-FP under different frame rate conditions using a translating grayscale phase target (Ascaris eggs sample). (a) Raw intensity image from a single central LED under the static condition. (b) Zoomed-in region of interest (ROI) from (a). (c) Phase reconstruction from traditional FP. (d) Design pattern of the phase target for reference. (e) Raw intensity images acquired at low fps. (f) Phase reconstructions from ST-FP using the low-fps dataset in (d). (g) Raw intensity images acquired at high fps. (h) Phase reconstructions from ST-FP using the high-fps dataset in (f). The comparison highlights the advantage of high-speed acquisition in preserving fine phase structures during motion (see Visualization 3).

We acquired two datasets under different acquisition regimes: a low-fps mode (20 Hz, 20 ms exposure) and a high-fps mode (260 Hz, 1 ms exposure). As shown in Fig. S3, the central low-resolution image captured under static conditions using a single central LED is displayed in Fig. S3(a), with a zoomed-in region of interest in Fig. S3(b). Figure S3(c) shows the high-resolution reference phase reconstruction by traditional FP under static conditions. Figure S3(d) displays the design pattern used to fabricate the grayscale phase target based on Ascaris egg data. Figures

S3 (e)-(f) show the raw intensity frames and corresponding ST-FP phase reconstructions under the low-fps condition, respectively. Motion blur is clearly visible in the raw data, leading to degraded reconstructions where fine structures are lost. In contrast, the high-fps condition shown in Figs. S3 (g)-(h) preserves sharp intensity frames, allowing our ST-FP algorithm to recover the object with high fidelity and structural clarity.

B. Live Microorganism Results

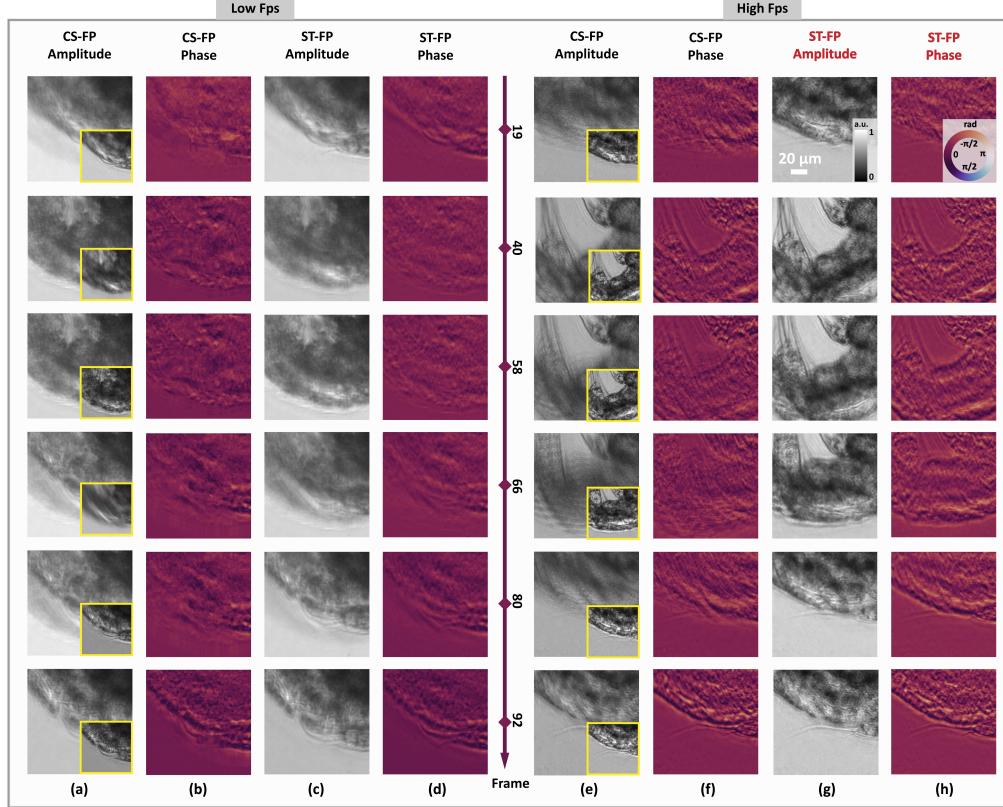


Fig. S4. Comparison of CS-FP and ST-FP reconstructions for a dynamically swimming brine shrimp under low and high frame rate conditions. (a)-(b) Amplitude and phase reconstructions using conventional CS-FP at low-fps. (c)-(d) Corresponding amplitude and phase reconstructions using ST-FP at low-fps. (e)-(f) Amplitude and phase results of CS-FP at high-fps. (g)-(h) Corresponding results of ST-FP at high-fps. The insets in (a) and (e) show the raw images. The frame indices illustrate the temporal evolution of the sample across selected frames. High-speed ST-FP demonstrates improved reconstruction fidelity and reduced motion artifacts, particularly in the dynamic regions (see Visualization 6).

To further evaluate the robustness of high-speed ST-FP across different specimens and dynamic behaviors, we conducted an additional experiment using a freely swimming brine shrimp sample. As in the main text Section 3.3, we acquired datasets under two acquisition regimes: a low-frame-rate condition (20 Hz, 20 ms exposure) and a high-frame-rate condition (260 Hz, 0.8 ms exposure). While the exact motion of the organism could not be repeated, we ensured qualitatively similar swimming patterns within a short interval, allowing fair comparison between methods.

Figure S4 presents the amplitude and phase reconstructions from conventional CS-FP and our proposed ST-FP method under both low-fps and high-fps conditions. Figures S4 (a)-(b) show CS-FP results at low frame rate, where strong motion blur severely degrades both amplitude and phase reconstructions, especially in dynamic regions. Figures S4 (c)-(d) show corresponding ST-FP results at low-fps, where incorporating motion estimation improves continuity and sharpness of the reconstructed features, though the long exposure time still imposes limitations. Under high-fps acquisition, CS-FP reconstructions benefit from reduced motion blur, yet still suffer from

assumptions of static behavior over multiple illumination patterns, leading to jitter and phase inconsistencies in fast-moving areas, as shown in Figs. S4(e)-(f). In contrast, ST-FP results in Figs. S4(g)-(h) demonstrate significant improvements in spatial resolution and temporal coherence. Structural edges are sharper and phase features are more continuous, even in regions undergoing rapid deformation.

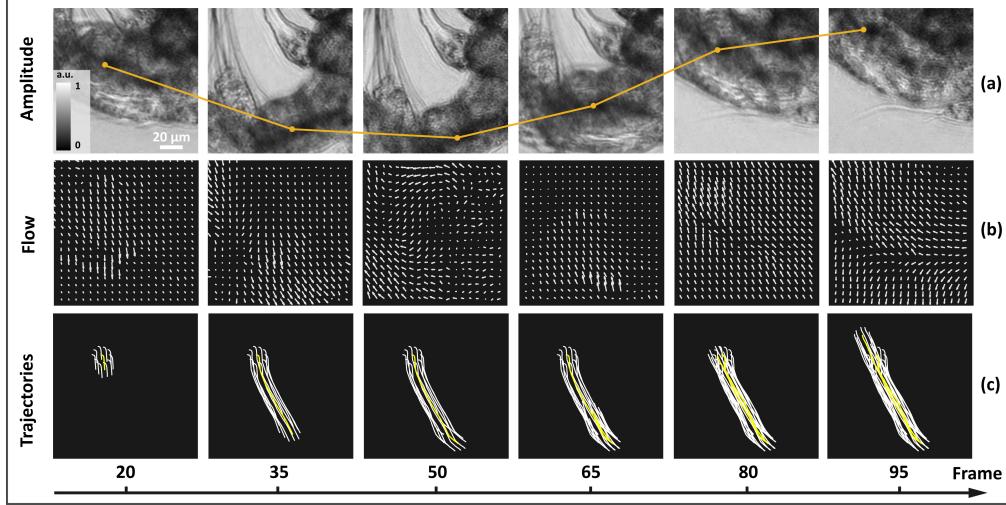


Fig. S5. Time-resolved motion analysis of a brine shrimp using ST-FP (see Visualization 7). (a) Representative amplitude reconstructions from high-fps ST-FP frames, showing continuous structural changes. Yellow points indicate the center of the tracked elliptical seed point cloud. (b) Corresponding 2D optical flow fields estimated between successive frames. (c) Tracked trajectories of an elliptical seed point cloud (10 points) over time, highlighting the motion evolution of the selected region. The yellow line marks the center trajectory in each frame.

Figure S5 presents dynamic motion analysis. Figure S5(a) displays amplitude reconstructions that resolve structural transitions of the specimen over time. The associated flow maps in Fig. S5(b) capture complex, time-varying motion. In Fig. S5(c), we visualize the temporal trajectories of two manually selected point clusters, revealing smooth and coordinated locomotion patterns.

5. COMPUTATION TIME

To address the computational burden of long-sequence reconstructions in ST-FP, we replace the original CPU-based ADMM solver [5] with a GPU-accelerated, auto-differentiable (AD) optimization framework. Specifically, we reformulate the object subproblem using PyTorch’s automatic differentiation engine [7], enabling gradient computation through the forward model and loss function. The gradients are then backpropagated to optimize the object sequence using the Adam optimizer [6]. This transition not only simplifies the optimization pipeline but also enables efficient utilization of modern GPU hardware.

In addition, we implement a recurrent processing scheme to handle long temporal sequences, which reduces peak memory consumption and computational complexity. As shown in Fig. S6(a), the GPU implementation leads to a substantial reduction in computation time across all tested configurations. For instance, in a large-scale reconstruction of size $120 \times 1200 \times 1200$, the computation time is reduced from 49,482 seconds (13.7 hours) on an AMD EPYC 7763 64-Core CPU to 2,131 seconds (0.59 hours) on an NVIDIA A100 GPU—yielding a ~ 23 -fold speedup. Fig. S6(b) further illustrates the absolute time savings achieved by the GPU implementation over CPU for different sequence sizes, highlighting the scalability of our framework with respect to both spatial and temporal dimensions.

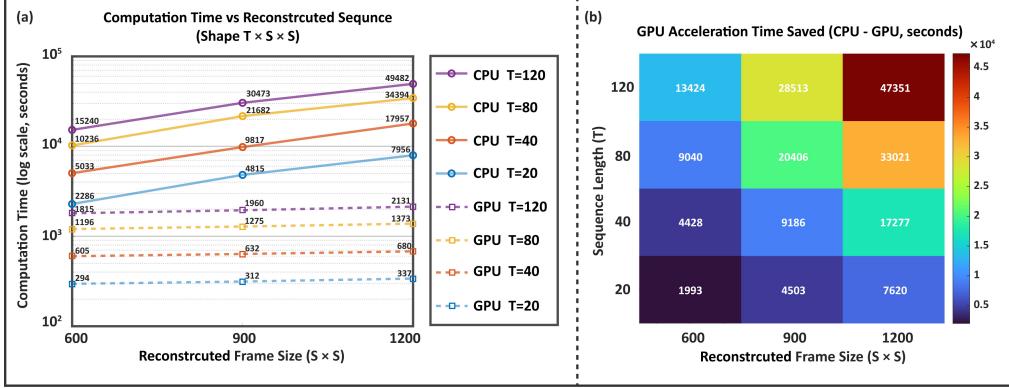


Fig. S6. Comparison of ST-FP computational performance between the original CPU-based ADMM implementation [5] and our GPU-accelerated auto-differentiable (AD) framework. (a) Computation time (log scale) for sequences of size $T \times S \times S$ reconstructed on CPU (solid lines) and GPU (dashed lines). GPU acceleration significantly reduces computation time, particularly for large spatial and temporal dimensions. (b) Heatmap showing time saved by GPU over CPU (CPU – GPU, in seconds), highlighting the scaling advantage of GPU computation with increasing sequence length and frame size. All numeric values shown are rounded to the nearest integer. The results are computed from the data in Section 3.1 in the main text.

REFERENCES

1. P. Getreuer, “A survey of gaussian convolution algorithms,” *Image Process. Line* **3**, 286–310 (2013).
2. S. Pollnow, N. Pilia, G. Schwaderlapp, *et al.*, “An adaptive spatio-temporal gaussian filter for processing cardiac optical mapping data,” *Comput. Biol. Med.* **102**, 267–277 (2018).
3. M. Werlberger, W. Trobin, T. Pock, *et al.*, “Anisotropic Huber-l1 optical flow,” in *Proc. Brit. Mach. Vis. Conf.*, (2009), pp. 1–11.
4. G. Zang, R. Idoughi, R. Tao, *et al.*, “Warp-and-project tomography for rapidly deforming objects,” *ACM Trans. Graph.* **38** (2019).
5. M. Sun, K. Wang, Y. N. Mishra, *et al.*, “Space-time Fourier ptychography for in vivo quantitative phase imaging,” *Optica* **11**, 1250–1260 (2024).
6. D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” arXiv preprint arXiv:1412.6980 (2014).
7. A. Paszke, S. Gross, S. Chintala, *et al.*, “Automatic differentiation in pytorch,” in *31st Conference on Neural Information Processing Systems (NIPS 2017)*, (2017).