High-speed Imaging through Turbulence with Event-based Light Fields

Yu-Hsiang Huang[1], Levi Burner[1], Sachin Shah[1], Ziyuan Qu[2], Adithya Pediredla[2], and Christopher A. Metzler[1]

[1] University of Maryland, College Park MD 20742, USA
[2] Dartmouth College, Hanover NH 03755, USA
https://justhowww.github.io/lf-ev-turb-project-page

Abstract. This work introduces and demonstrates the first system capable of imaging fast-moving extended non-rigid objects through strong atmospheric turbulence at high frame rate. Event cameras are a novel sensing architecture capable of estimating high-speed imagery at thousands of frames per second. However, on their own, event cameras are unable to disambiguate scene motion from turbulence. In this work, we overcome this limitation using event-based light field cameras: By simultaneously capturing multiple views of a scene, event-based light field cameras and machine learning-based reconstruction algorithms are able to disambiguate motion-induced dynamics, which produce events that are strongly correlated across views, from turbulence-induced dynamics, which produce events that are weakly correlated across view. Tabletop experiments demonstrate event-based light field can overcome strong turbulence while imaging high-speed objects traveling at up to 16,000 pixels per second.

Keywords: Event camera · Atmospheric turbulence · Light field · Video reconstruction

1 Introduction

Long-range imaging through the atmosphere is the primary goal of terrestrial astronomy, remote sensing, surveillance, and other applications. Atmospheric turbulence represents the central impediment to this goal. Regardless of sensor resolution, aperture size, platform stability, or cost, atmospheric turbulence significantly degrades image fidelity by introducing blur, geometric distortion, and wander.

Traditionally, turbulence has been mitigated with hardware-based solutions, such as adaptive optics (AO) [18], or software-based solutions, such as lucky imaging [12] or, more recently, deep learning [55–57]. However, almost all existing solutions share the same fundamental constraint: they rely upon conventional image sensors which integrate light over a fixed exposure window, irrecoverably

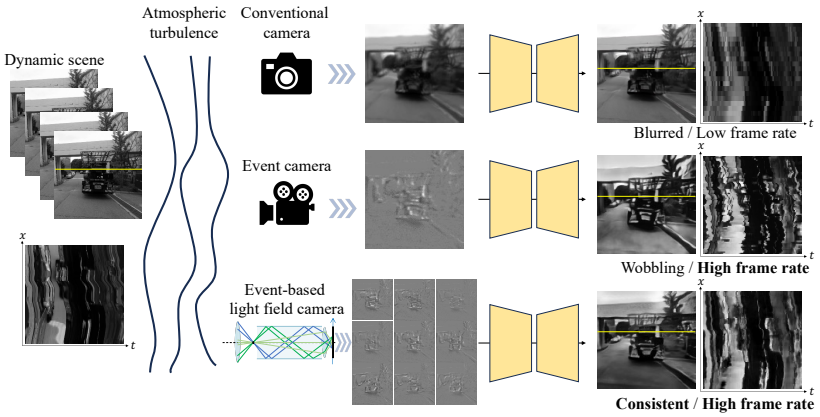

Fig. 1: Method overview. We introduce an event-based light field camera to image through turbulence at high speed. The event camera avoids motion blur by detecting brightness changes at microsecond time resolution. The light field offers cross-view constraints on the scene.

blending any scene dynamics or turbulence fluctuations that occur over a time scale shorter than the exposure time. In this work, we overcome this limitation with event cameras.

Event cameras mitigate motion blur by measuring brightness change asynchronously at microsecond timescales, making them a natural candidate for recovering high-speed dynamics. Several lines of work have used frames and events to image through turbulence [3,4,27,29]. However, they have relied upon restrictive rigid-body assumptions about scene motion [3, 4] or have restricted the reconstructions to the rate of the frame-based camera [27,29], thereby giving up event-camera's central advantage.

Event-only imaging of dynamic scenes through turbulence is nontrivial. If the scene were static and all events were due to turbulence, one may be able to exploit these events to help reconstruct the scene [19]. If the turbulence was static and the scene was dynamic, one could use existing event-to-video frameworks to reconstruct a (distorted) high-speed video [39, 40, 42]. However, when both the turbulence and the scene are dynamic, both will produce (largely indistinguishable) events, thereby invalidating the assumptions that existing methods rely upon.

In this work, we use an event-based light field camera (Event Fields) [38] to overcome this ambiguity. Light fields provide multiple sub-apertures simultaneously, each corrupted by a different turbulence realization, while scene dynamics remain largely consistent across all views. Thus, each light field view provides a constraint on the singular scene image, and inconsistent motion across views can be attributed to turbulence. We exploit this structure (implicitly) using the recurrent encoder-decoder architecture E2VID [39] with stacked multi-view event inputs and train it to reconstruct clean, high-speed video through turbulence.

In summary, our contributions are:

- The first event-only method for high-speed imaging through turbulence, breaking the frame-rate barrier of conventional cameras.
- Demonstration that light field optics provide sufficient information for removing turbulence while preserving scene motion.
- Real (in-lab) and simulated experiments demonstrating our method significantly outperforms single-view reconstruction.

2 Related works

2.1 Light Field Imaging

Light field imaging acquires multiple simultaneous views of a scene from distinct viewpoints, encoding both spatial and angular information of light [21,48]. Camera arrays achieve this by positioning multiple independent sensors at spatially separated viewpoints [51]. While conceptually straightforward, they require precise inter-sensor synchronization and their physical bulk limits deployment in constrained platforms. Single-shot alternatives using microlens arrays overcome the synchronization problem by imaging all sub-apertures onto a single sensor [1, 36]. However, the dense lenslet grid demands precise optical alignment and manufacturing tolerances. Kaleidoscope-based designs avoid both drawbacks [31, 38]. Planar mirrors placed at the intermediate image plane fold the optical path, routing light from distinct regions of the aperture to separate areas of the sensor and producing multiple virtual viewpoints without additional optics per channel. This construction requires no microlens fabrication and eliminates inter-camera synchronization entirely, accordingly we adopt it here.

2.2 Imaging Through Aberrations with Light Fields

Ng and Hanrahan first demonstrated that light fields could be used to correct for optical aberrations back in 2006 [35]. Light fields have since been used to image through turbulence on multiple occasions [30, 46, 47, 49]. Intuitively, by capturing multiple perspectives of the same scene at the exact same time, these cameras create valuable spatial redundancy that can be used to remove turbulence and other aberrations. These systems have relied on conventional image sensors. While a few recent works have combined light fields with event cameras to perform high-speed 3D imaging [16, 38], ours is the first work to use event-based light field cameras to image through turbulence.

2.3 Software-based Imaging through Turbulence

Classical turbulence mitigation algorithms often use spatial registration and temporal fusion [6, 32, 41] to recover imagery. Accordingly, these methods generally depend on static references or distortion-free patches and fail in dynamic scenes

or strong turbulence. Learning-based methods relax these requirements: self-supervised and test-time optimization use priors such as lucky images [12], blind degradation estimation [22,28], or pretrained mitigation models [5,11] and adapt without paired data, but are iterative and ill-suited to high-speed use; meanwhile physics-based simulators [33] enable supervised learning with fast inference via feed-forward sequence-to-sequence [57] or recurrent [13,55] networks. All of these approaches remain limited by conventional sensors' fixed exposure, which conflates scene dynamics and turbulence at high frame rates.

Event cameras, with microsecond-level temporal resolution, are a natural candidate for resolving this bottleneck. An early work, used an event camera to identify lucky regions in conventional frames [3, 4]. Two recent works explore this direction with machine learning: EvTurb [29] employs events to guide single-frame restoration, while EGTM [27] leverages event temporal resolution to compute adaptive fusion weights for conventional frames. In both, the final image formation depends on the integration time of the conventional sensor, leaving output constrained to standard frame rates.

A fundamental bottleneck thus persists across all existing methods: event cameras alone cannot distinguish scene-induced events from turbulence-induced events. We propose an event-only approach that leverages light field imaging to resolve this ambiguity and achieve high-speed imaging through turbulence.

2.4 Hardware-based Imaging through Turbulence

AO, wherein one measures and then physically compensates for optical aberrations using deformable mirror or spatial light modulator, is the primary hardware-based approach for imaging through turbulence [18]. The majority of AO systems rely upon wavefront sensors [8,17,44], which can be accelerated with event cameras [15,25,43,58]. However, with a few notable exceptions [10,17,23,37,50,52], existing AO systems rely upon guidestars—bright points in the scene that can be used for calibration—which greatly limits their flexibility. Ours is the only guidestar-free turbulence mitigation approach that can operate at kHz.

2.5 Events to Video

Reconstructing intensity frames from event streams has seen rapid evolution. Early methods [39, 40, 42] employ recurrent U-Net structures with ConvLSTM or ConvGRU modules to accumulate temporal state across asynchronous events, enabling continuous video reconstruction. HyperE2VID [9] extends this paradigm by introducing a hypernetwork that dynamically predicts convolutional kernel sizes, improving adaptability to varying event densities. To better capture long-range temporal dependencies, ETNet [45] adopts transformer architectures, achieving improved reconstruction quality at the cost of increased computational overhead. More recently, EventMamba [13] leverages state-space models to achieve efficient long-range modeling with linear complexity. A fundamental assumption underlies all these methods: events arise exclusively from scene dynamics or camera motion.

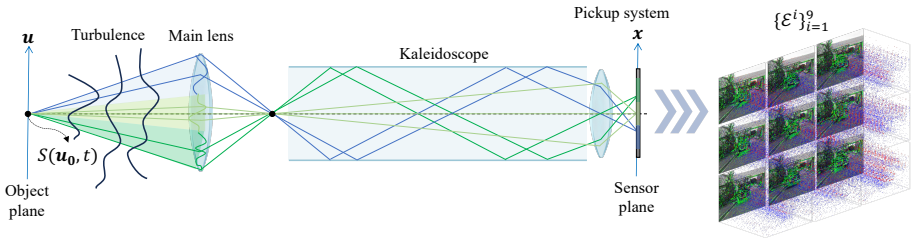

Fig. 2: Image model for event-based light field camera under turbulence. A light field camera captures N simultaneous sub-aperture views, each observing the same scene through an independent turbulence realization, which injects different turbulence-induced fluctuations across views while leaving the scene contribution identical.

Under atmospheric turbulence, events no longer arise from scene motion alone. With a single camera, both turbulence and scene dynamics introduce changes in intensity, and retraining these networks on turbulent data does not resolve this fundamental ambiguity. We overcome this limitation with a light field event camera that simultaneously captures multiple views of the scene. Events caused by scene dynamics are strongly correlated across views, while events caused by turbulence are weakly correlated, providing the cross-view structure a network can exploit to disentangle the two.

3 Methods

3.1 Modeling Turbulent Event Light Fields

Our imaging model is illustrated by Figure 2. A detailed description follows. The intensity at the sensor $I(\mathbf{x}, t)$ is given by the incoherent imaging integral under an anisoplanatic point spread function (PSF):

$$I(\mathbf{x}, t) = \int_{\Omega_s} |h_{\mathbf{u}}(\mathbf{x}, t)|^2 \, S(\mathbf{u}, t) \, d\mathbf{u}, \tag{1}$$

where $S(\mathbf{u}, t)$ is the scene radiance and $h_{\mathbf{u}}(\mathbf{x}, t)$ is the complex coherent PSF which incorporates the phase errors introduced by turbulence along the optical path [14]. The accumulated phase error leads to the geometric warp and spatially varying blur observed at the sensor, as dictated by $h_{\mathbf{u}}(\mathbf{x}, t)$ [7].

This forward model is bilinear in S and $|h|^2$: scene motion (changes in S) and turbulence (changes in h) are multiplicatively coupled, making their individual contributions ambiguous in the observed intensity I. Thus, recovering both from a single observation constitutes an ill-posed blind inverse problem.

Multiple views mitigate this ambiguity. If N observations share a scene S through separate turbulence realizations $\{h^{(i)}\}$, the system accumulates cross-view constraints on the common S:

$$I^{(i)}(\mathbf{x}, t) = \int_{\Omega_s} |h_{\mathbf{u}}^{(i)}(\mathbf{x}, t)|^2 \, S(\mathbf{u}, t) \, d\mathbf{u}, \quad i = 1, \ldots, N. \tag{2}$$

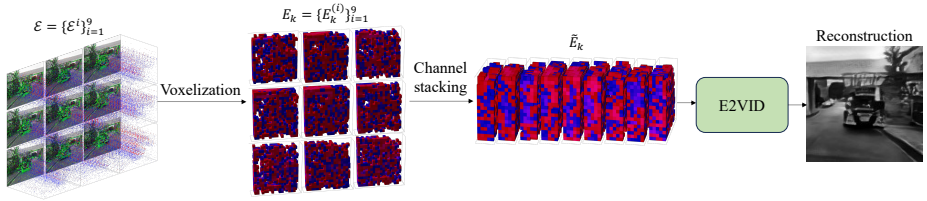

Fig. 3: Network architecture. Each sub-aperture event stream is converted to a voxel grid $E_k^{(i)} \in \mathbb{R}^{B\times H\times W}$; the N grids are stacked along the channel dimension to form the joint input $\tilde{E}_k \in \mathbb{R}^{NB\times H\times W}$. A recurrent convolutional encoder-decoder processes the stacked input window by window and outputs a reconstructed frame $\hat{V}_k$.

Fig. 4: Simulation pipeline. Turbulence is added to clean frames using a turbulence simulator based on the phase-to-space (P2S) transform [33]. V2E converts turbulent frames to events [20]. The process is repeated 9 times to simulate a 3×3 light field. The network is trained on simulated data and generalizes to tabletop experimental settings.

S can be assumed to be the same in each view, because at long range parallax is negligible. However, turbulence affects each view differently. Rays entering each sub-aperture traverse a spatially distinct atmospheric column, imparting a unique instantaneous phase error and associated PSF $h_{\mathbf{u}}^{(i)}(t)$.

We obtain N views with a light field event camera, with the goal of high speed imaging through turbulence. Following the standard event camera model [39], the intensity of each view is passed through a non-linear logarithmic threshold: an event $(\mathbf{x}, t, p) \in \mathcal{E}^{(i)}$ is emitted whenever the absolute log-intensity change $\Delta L^{(i)}(\mathbf{x}, t) = |\log I^{(i)}(\mathbf{x}, t) - \log I^{(i)}(\mathbf{x}, t-\Delta t)|$ exceeds a threshold C, with polarity $p_i = \mathrm{sgn}(\Delta L^{(i)})$.

3.2 Turbulence Mitigation Model

While the N different views provide critical cross-view constraints on the shared scene S, the system remains underdetermined without physical regularization on h. Rather than designing explicit priors for this non-convex joint estimation problem, we adopt a data-driven approach: a neural network is trained to directly approximate the inverse mapping from the N event streams to the clean video V. By training on a large dataset of simulated turbulence [33], the network implicitly learns the spatial statistics of anisoplanatic turbulence and the man-

ifold of natural scenes, leveraging cross-view consensus to resolve the bilinear ambiguity.

We implement this by adapting E2VID [39] to the multi-view setting with minimal modification, so that the performance gain can be attributed to the light field geometry rather than architectural choices. In its original single-view form, E2VID takes an event stream $\mathcal{E}$ as input and outputs a reconstructed video $\hat{V}$. E2VID represents input events as a spatio-temporal voxel grid with 5 temporal bins, whose values contain temporally weighted summed event polarity.

We extend E2VID to the multi-view setting by stacking each view's voxel representation along the temporal bin dimension as additional channels, forming a joint input illustrated in Figure 3. E2VID processes the stacked input with a recurrent convolutional encoder-decoder, whose recurrent state accumulates temporal evidence across windows, producing the reconstructed frames $\hat{V}$.

Training Details The network is trained to minimize the LPIPS loss [54] averaged over all reconstructed frames:

$$\mathcal{L} = \frac{1}{K} \sum_{k=1}^{K} \mathrm{LPIPS}(\hat{V}_k, V_k). \tag{3}$$

E2VID is traditionally trained with the calibrated perceptual loss LPIPS instead of mean squared error (MSE) because MSE losses often produce blurry images [39]. The network is trained from scratch for 200 epochs with a batch size of 8 using the Adam optimizer [24] with a constant learning rate of 10^{-4} on a single NVIDIA RTX A6000 GPU.

We follow the augmentation strategy of [39, 42]. Geometric augmentation applies random 112×112 crops and flips with probability 0.5. Noise augmentation adds $\mathcal{N}(0, 0.4)$ to $\tilde{E}_k$ and injects hot-pixel noise $\mathcal{N}(0, 0.1)$ at up to 0.01% of pixel locations. Pause augmentation zeros input events with probability 0.05 and maintains the paused state with probability 0.9, training the recurrent state to preserve the output without new events. Spectral normalization [53] is applied to the ConvLSTM layers to prevent artifacts from accumulating as the recurrent hidden state propagates across temporal windows.

4 Experiments

Simulated Dataset A synthetic light field event dataset is constructed by combining a physics-based turbulence simulator (P2S) [33] with an event simulator (V2E) [20], illustrated by Figure 4. Clean intensity frames are taken from the REDS dataset [34], whose 120 fps videos provide high-frame-rate intensity sequences for event synthesis. The central region is cropped and resized to 256×256, serving as ground truth. For each clip, $N = 9$ independent turbulence realizations are generated by running the turbulence simulator with the same parameters but different random seeds for the P2S, producing nine sequences of degraded frames. Each degraded sequence is then passed through the

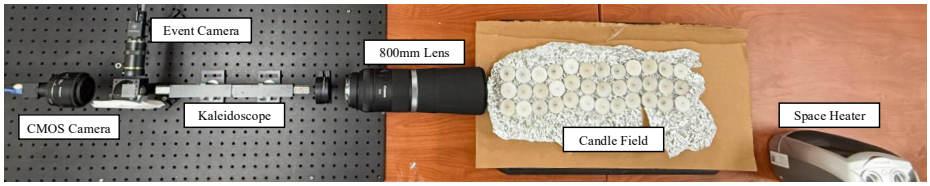

Fig. 5: Tabletop experimental setup. Real turbulence is generated from temperature fluctuations caused by candles and a space heater. We placed a beam splitter behind the kaleidoscope to simultaneously capture events and reference CMOS frames.

event simulator to produce the corresponding sub-aperture event stream $\mathcal{E}^{(i)}$. The contrast threshold C is uniformly sampled from $[0.1, 0.7]$ and the refractory period is set to the default 5 ms. The temporal window ΔT_k in the voxel grid is set to one frame interval.

To improve generalization across turbulence conditions, we sample turbulence parameters from the ATSyn dataset [55], which provides physically grounded turbulence parameters, and adapt P2S following the simulation procedure described in the same work. First, the pixel pitch p is computed from the focal length f, scene width, and propagation distance, and the geometric tilt is scaled by f/p. Second, the blur kernel is resized to match the scale specified by the ATSyn parameters. Finally, temporal correlation is introduced by relating the random seed $\boldsymbol{n}_t$ at time t to its predecessor via $\boldsymbol{n}_t = \alpha \boldsymbol{n}_{t-1} + \sqrt{1-\alpha^2}\,\boldsymbol{\epsilon}_t$, where α is the temporal correlation coefficient and $\boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$.

The network is trained exclusively on this simulated dataset but is evaluated on both simulated and experimental data. The dataset is split into 216 training, 27 validation, and 27 test clips, each elapsing 4 seconds. Using the turbulence severity criteria in [55], our simulated dataset consists of 60% strong, 30% medium, and 10% weak clips. A 768×768 version of the dataset was generated to compare single-view reconstructions estimated from the same number of pixels as the $3 \times 3 \times 256 \times 256$ light field. Complete characterization of the light-field/resolution tradeoff is provided in the supplemental material.

Tabletop Experiments As illustrated in Figure 5, the light field event camera comprises two optical stages in series. A Canon RF 800 mm f/11 main lens focuses the scene onto an intermediate image plane. To match the exit pupil of the main lens with the entrance pupil of the pickup system, a 100 mm pupil-matching relay lens is introduced. The pickup system forms an image from the intermediate plane onto the sensor plane. Through the reflections of a 305 mm-long kaleidoscope with a 12.6 mm aperture, similar to the setup in [38], a 3×3 array of sub-aperture views is formed. A beam splitter behind the kaleidoscope directs light simultaneously to a Prophesee EVK4 event sensor for event capture and a Blackfly S BFS-U3-13Y3C CMOS camera for reference imagery.

Turbulence is generated in the laboratory by placing a 3×12 grid of tea candles starting 61.0 mm from the main lens and a 1500 watt space heater 813 mm

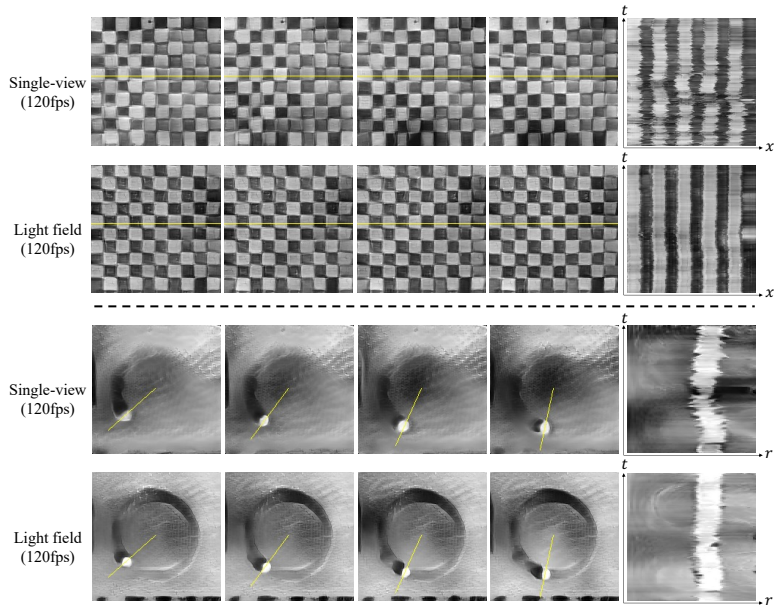

Fig. 6: Light fields improve temporal consistency on experimental data. In the laboratory, we imaged a static checkerboard and a rotating dot, both displayed on an LCD monitor, through turbulence. The left four columns show reconstructed frames at 120 fps. The right column visualizes x-t and r-t plots showing the time evolution the slice marked in yellow. The light field results in straighter lines, indicating less wander.

from the same lens, as illustrated in Figure 5. All experiments are conducted with targets at a distance of 9.14 m from the camera. This is not far enough to assume targets are at optical infinity, and the kaleidoscope alignment is not perfect, so each sub-aperture view is aligned to the center with a homography estimated from a calibration target presented simultaneously to each view.

5 Results

5.1 Event Light Field versus Single-View Reconstruction

To demonstrate light field optics capture sufficient information to remove turbulence while preserving scene motion, we compare the light field configuration against a single-view baseline using the same reconstruction architecture trained on turbulence-degraded data from the central sub-aperture only.

As illustrated in Figure 7, the model trained on light field views reconstructs edges more faithfully than the single-view baseline. This is apparent along structural lines: the single-view approach bends them into curved artifacts, whereas the light field model preserves sharper, straighter edges. This improvement is particularly evident in real-world experiments, where the single-view baseline

Table 1: Simulated quantitative results. On the REDS dataset, the light field improves frame quality in comparison to a single-view configuration.

Method	Resolution	PSNR ↑	SSIM ↑	LPIPS ↓	ITF ↑	E_{warp} ↓
Single-view	768×768	14.84	0.501	0.5185	22.71	0.0050
Single-view	256×256	14.84	0.461	0.4607	22.72	0.0047
Light field	$3 \times 3 \times 256 \times 256$	**15.40**	**0.505**	**0.4332**	**25.16**	**0.0022**

Table 2: Tabletop experiment quantitative results. Reconstructing REDS videos displayed through turbulence, light field outperforms single-view.

Method	PSNR ↑	SSIM ↑	LPIPS↓	ITF ↑	E_{warp} ↓
Single-view	12.37	0.351	0.543	20.87	0.0085
Light field	**12.99**	**0.367**	**0.531**	**23.30**	**0.0051**

suffers from blurred edges and dimming artifacts. The x-t plots further show that the light field reconstruction is more temporally consistent, with fewer frame-to-frame fluctuations caused by turbulence.

Table 1 quantitatively shows that the light field model outperforms the single-view baseline in both reconstruction quality and temporal stability. This holds even when the total pixel count is matched by using three times the spatial resolution in both the horizontal and vertical dimensions to match the aggregate pixel count of the 3×3 light field. We evaluate temporal stability using Inter-frame Transformation Fidelity (ITF) [2] and warping error (E_{warp}) [26]; complete implementation details are provided in the supplemental material.

Additionally, two real-world experiments are presented in Figure 6 to illustrate how the light field approach reconstructs consistent videos at a traditional frame rate of 120 fps. Higher frame rate results are shown later. First, a static checkerboard is reconstructed through a turbulent medium. In the x-t plot, one can see that the single-view baseline exhibits clear temporal jitter, whereas the light field method maintains highly stable edges over time. Next, a spinning dot is displayed on a flicker-free, LCD monitor. While the monitor displays frames discretely at 120 Hz, the motion-blur inherent to LCD monitors smooths the scene's motion. The r-t plot is constructed by creating a slice that follows the dot's spin. This plot demonstrates the light field configuration effectively reduces spatial wandering caused by turbulence. Quantitative results for the tabletop experiments are reported in Table 2. The ground truth is a separate capture acquired with the same experimental setup but without turbulence. Across metrics, the light field configuration consistently outperforms the single-view baselines.

5.2 Event Light Fields Overcome Motion Blur

We test the ability of event light fields to mitigate motion blur by simulating motion blur and comparing to MambaTM [56], the state-of-the-art frame-based

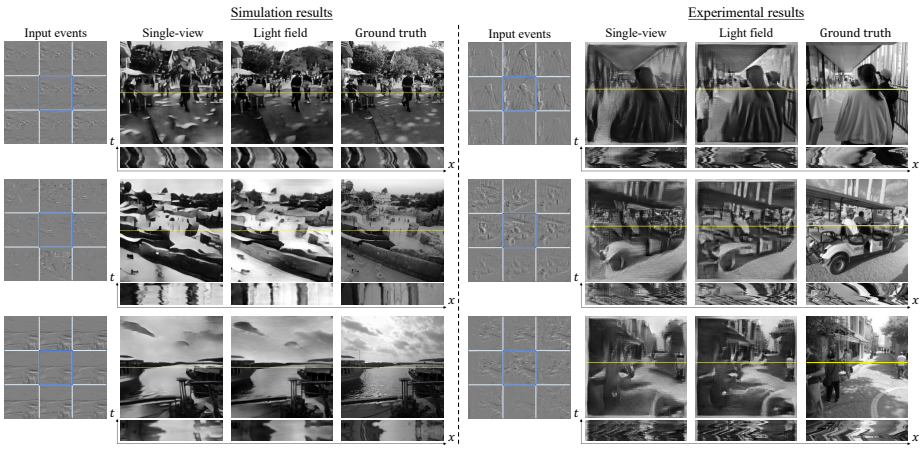

Fig. 7: Reconstruction with light field input outperforms single-view on simulated and experimental data. Each group shows the input events (blue box indicates the central sub-aperture used for the single-view baseline), single-view reconstruction, the light field reconstruction, and the ground truth. Simulated single-view results are at 768×768 resolution, which has the same number of input pixels as the 3×3 light field whose views are simulated at 256×256. Experimental single-view and light field results are at 176×176, with the target videos displayed on an LCD monitor. The light field model recovers sharper edges with fewer artifacts than the single-view baselines at either resolution. The x-t plots below each group show the time evolution of the slice marked in yellow. Their relative smoothness shows that the light field reconstruction maintains a consistent appearance over time, while the single-view output exhibits frame-to-frame instability caused by turbulence. Videos in supplemental material more clearly illustrate turbulence reduction than images.

turbulence mitigation method. High speed imagery is simulated by averaging each group of four consecutive frames.As illustrated in Figure 8, MambaTM reconstructions still contain blur and distortions. Our method, in contrast, ingests event streams, which are free from the camera integration window, simulated from the same high-speed frames, and reconstructs video directly. The most relevant prior methods for single-view turbulence mitigation using events, EvTurb [29] and EGTM [27] have not publicly released code, precluding direct comparison. These methods use both frames and events.

5.3 Recovering High-Speed Video Through Turbulence

We performed multiple in-lab, tabletop experiments to demonstrate that event-based light field imaging enables high-speed long-range imaging through turbulence. Unless noted otherwise, we reconstruct video at 600 fps. As shown in Figure 9, we captured a spinning stripe and a bouncing ball in lab. Since the spinning stripe rotates uniformly, the θ-t plot should show straight sloped lines. The single-view configuration produces a wavy trajectory, revealing resid-

Fig. 8: Simulated high-speed scenes show that event light fields recover sharper images than frame-based methods. Motion-blur cannot be compensated for by frame-based methods, which occurs when scene dynamics or turbulence are fast. Event-based light fields operate directly on the event stream and reconstruct the underlying scene without blur caused by integration. Motion-blur is simulated by aggregating 4 conventional frames into a single blurred frame.

ual turbulence corruption; the light field reconstruction recovers a straighter slope, consistent with the expectation. For the bouncing ball, the single-view x-t plot shows geometric distortion and loss of surface texture, while the light field reconstruction preserves the ball's texture along the full cross-sectional trajectory.

Additionally, as shown in Figure 10, we captured a water balloon bursting. The light field configuration captures the complex water dynamics without confusing them with turbulence. Finally, as shown in Figure 11, we captured a Nerf dart traveling at 16,000 pixels per second. We successfully estimate the video at 12,000 fps, with the light field reconstructing a noticeably clearer and more accurate straight line trajectory when compared to a reference captured with a high-speed camera (Photron UX100) recording at 4,000 fps with additional back-lighting.

5.4 Comparison with High-Speed Camera

To our knowledge, ours is the first event-only turbulence mitigation method for dynamic scenes, leaving no directly comparable high-speed baseline. The closest event-based methods, EvTurb [29] and EGTM [27], require both events and

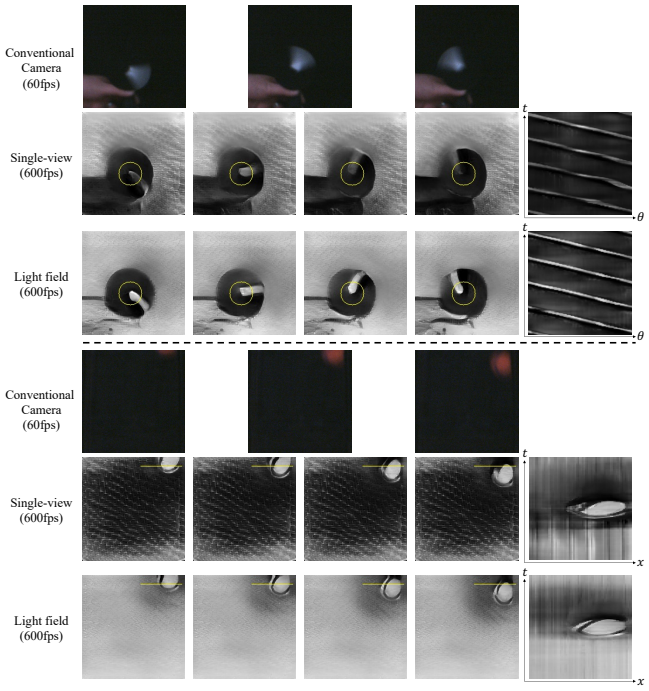

Fig. 9: Event-based light field design corrects turbulence in experimental high-speed regimes. In lab, we imaged a spinning reflective stripe (top) and a bouncing ball (bottom). A 600 fps video is reconstructed, with samples shown four frames apart. Our method successfully reconstructs frame detail while correcting for wobble, as shown in the last column. The θ-t and x-t plots show a 200 ms interval. Videos in supplemental material more clearly illustrate turbulence reduction than images.

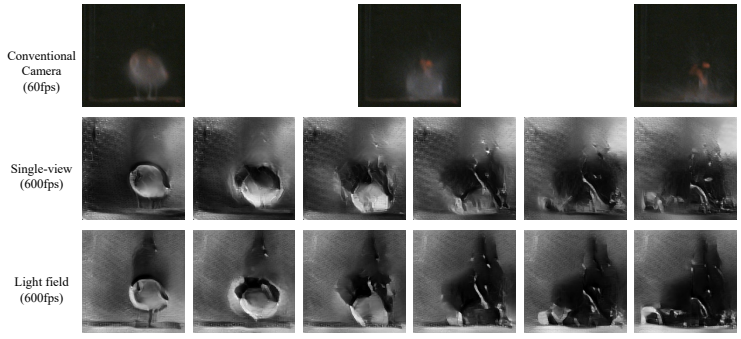

Fig. 10: High-speed video reconstruction in experimental turbulent scenes with complex dynamics. In lab, our method recovers a water balloon bursting on top of pins at 600 fps. The displayed images are spaced 5 frames apart. Please see the videos included in the supplemental material for a clearer comparison.

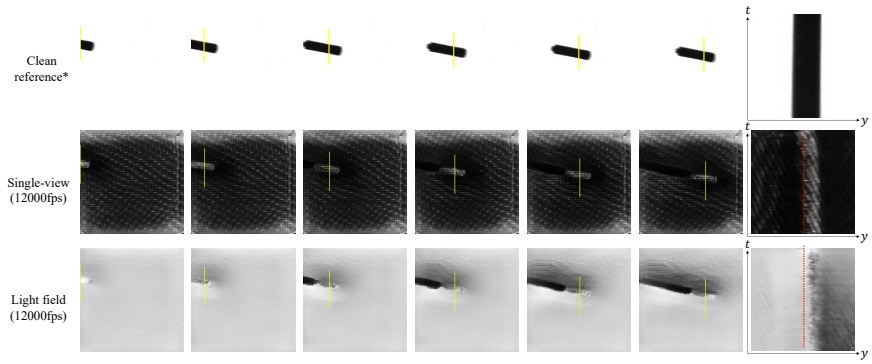

Fig. 11: Experimental reconstruction of a Nerf dart at 12,000 fps. The first row shows a turbulence-free HSC capture, where a Nerf dart follows a straight trajectory. The second two rows show reconstructions of a dart taking a similar, but not identical, trajectory imaged through turbulence with the event camera. Compared with the single-view baseline, the light-field model more accurately reconstructs the straight path of a passing dart traveling at 16,000 pixels per second. The y-t plot shows 8.3 ms of data, with images spaced 16 frames apart.

frames, target static scenes, and have no released code. We therefore compare against MambaTM paired with a Photron UX100 high-speed camera (HSC) operating at 500–4000 fps. This camera is an order of magnitude more expensive ($100k vs. $5k), far heavier (1.5 kg vs. 40 g camera body), has a limited capture duration (2 s maximum vs. unlimited), and requires substantially more light (targets had to be lit with 2–45× more light than previous experiments). As illustrated in Fig. 12, this powerful baseline still exhibits residual blur and distortion compared to our method. MambaTM's batching also causes glitches, visible in the θ-t plots.

6 Limitations

Our system inherits the limited spatial resolution of event cameras (∼1 MP maximum) and further trades spatial resolution for angular diversity. A multi-camera light field array avoids this tradeoff, but would increase cost and require synchronization across sensors. The current design assumes that the scene is approximately at optical infinity, with limited parallax between sub-aperture views. This is reasonable for long-range atmospheric imaging settings, but does not hold for close-range applications such as biomedical imaging through optical aberrations. Generalization beyond laboratory scenes remains an open challenge, particularly for real-world high-dynamic-range scenes where event simulation can be less accurate. The current reconstruction algorithm has a latency of 3.4 ms at 176×176 resolution. This supports real-time operation at the evaluated resolution, but higher-resolution deployment requires further optimization.

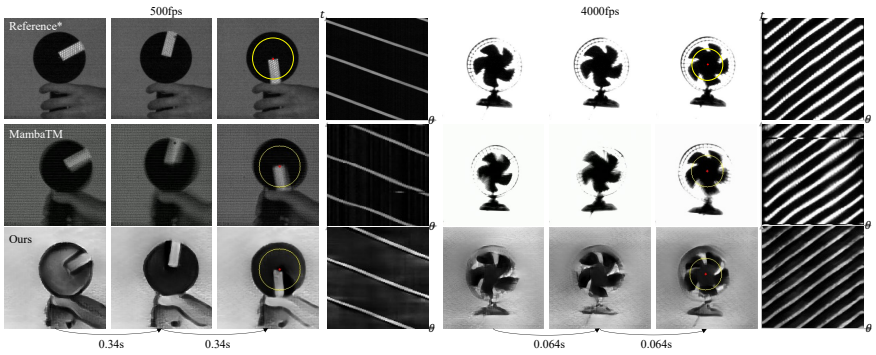

Fig. 12: Comparison with high-speed camera. With a $100k, high-speed camera MambaTM retains residual distortion, while ours, at ∼20× lower cost, stays sharp. Note that the reference was captured separately from the turbulent HSC and event data, and therefore temporally post-aligned.

7 Conclusion

We introduced and demonstrated an approach to recovering high-speed video through atmospheric turbulence using a light field event camera, which does not rely upon and is not constrained by a conventional frame-based sensor. The event camera offers high temporal resolution to resolve fast dynamics. However, on its own, it cannot discern scene dynamics from turbulence. We showed that a light field design resolves this issue. Each sub-aperture sees the same scene through a different path in the atmosphere. Then, the scene signal is consistent across views while the turbulence signature varies. Experiments with neural network based event to video reconstruction confirm that event fields mitigate turbulence better than a single-view, event fields overcome the motion-blur inherent to frame-based turbulence mitigation, simulated event fields can be used to train models that work on real-data, and event fields can reconstruct complex scene dynamics without confusing them with turbulence.

Acknowledgements

We gratefully acknowledge support from the Maryland Robotics Center. L.B. was supported by the TRX Systems Postdoctoral Fellowship. Z.Q. and A.P. were supported in part by the National Science Foundation under Grant No. 2403122. S.S. was supported by an NDSEG Fellowship through the Air Force Office of Scientific Research under award number FA955025CB010 in the amount of $153,984. Y.H., L.B., and C.A.M. were supported in part by ARO award no. W911NF2420113, ONR award no. N000142312752, and NSF CAREER award no. 2339616. This work is also supported in part by the Defense Advanced Research Projects Agency (DARPA) under the CIDAR program.

References

1. Adelson, E., Wang, J.: Single lens stereo with a plenoptic camera. IEEE Transactions on Pattern Analysis and Machine Intelligence **14**(2), 99–106 (1992)
2. Aguilar, W.G., Angulo, C.: Real-time video stabilization without phantom movements for micro aerial vehicles. EURASIP Journal on Image and Video Processing **2014**(1), 46 (2014)
3. Boehrer, N., Nieuwenhuizen, R., Dijk, J.: Using event cameras for imaging through atmospheric turbulence. In: COAT-2019-workshop (Communications and Observations through Atmospheric Turbulence: characterization and mitigation) (2019)
4. Boehrer, N., Nieuwenhuizen, R.P., Dijk, J.: Turbulence mitigation in imagery including moving objects from a static event camera. Optical Engineering **60**(5), 053101–053101 (2021)
5. Cai, H., Chen, J., Feng, B.Y., Jiang, W., Xie, M., Zhang, K., Fermuller, C., Aloimonos, Y., Veeraraghavan, A., Metzler, C.A.: Temporally consistent atmospheric turbulence mitigation with neural representations. Advances in Neural Information Processing Systems **37**, 44554–44574 (2024)
6. Caliskan, T., Arica, N.: Atmospheric turbulence mitigation using optical flow. In: 2014 22nd International Conference on Pattern Recognition. pp. 883–888. Ieee (2014)
7. Chan, S.H., Chimitt, N.: Computational imaging through atmospheric turbulence. Foundations and Trends in Computer Graphics and Vision **15**(4), 253–508 (2023)
8. Chimitt, N., Almuallem, A., Guo, Q., Chan, S.H.: Wavefront estimation from a single measurement: Uniqueness and algorithms. IEEE Transactions on Computational Imaging **11**, 1600–1613 (2025)
9. Ercan, B., Eker, O., Saglam, C., Erdem, A., Erdem, E.: HyperE2VID: Improving event-based video reconstruction via hypernetworks. IEEE Transactions on Image Processing **33**, 1826–1837 (2024)
10. Feng, B.Y., Guo, H., Xie, M., Boominathan, V., Sharma, M.K., Veeraraghavan, A., Metzler, C.A.: NeuWS: Neural wavefront shaping for guidestar-free imaging through static and dynamic scattering media. Science Advances **9**(26), eadg4671 (2023)
11. Feng, B.Y., Xie, M., Metzler, C.A.: Turbugan: An adversarial learning approach to spatially-varying multiframe blind deconvolution with applications to imaging through turbulence. IEEE Journal on Selected Areas in Information Theory **3**(3), 543–556 (2023)
12. Fried, D.L.: Probability of getting a lucky short-exposure image through turbulence. Journal of the Optical Society of America **68**(12), 1651–1658 (1978)
13. Ge, C., Fu, X., He, P., Wang, K., Cao, C., Zha, Z.J.: EventMamba: Enhancing spatio-temporal locality with state space models for event-based video reconstruction. Proceedings of the AAAI Conference on Artificial Intelligence **39**(3), 3104–3112 (2025)
14. Goodman, J.: Introduction to Fourier Optics. Electrical Engineering Series, McGraw-Hill (1996)
15. Grose, M., Schmidt, J.D., Hirakawa, K.: Convolutional neural network for improved event-based shack-hartmann wavefront reconstruction. Applied Optics **63**(16), E35–E47 (2024)
16. Guo, R., Yang, Q., Chang, A.S., Hu, G., Greene, J., Gabel, C.V., You, S., Tian, L.: EventLFM: event camera integrated fourier light field microscopy for ultrafast 3d imaging. Light: Science & Applications **13**(1), 144 (2024)

17. Guo, Y., Hao, Y., Wan, S., Zhang, H., Zhu, L., Zhang, Y., Wu, J., Dai, Q., Fang, L.: Direct observation of atmospheric turbulence with a video-rate wide-field wavefront sensor. Nature Photonics **18**(9), 935–943 (2024)
18. Hampson, K.M., Turcotte, R., Miller, D.T., Kurokawa, K., Males, J.R., Ji, N., Booth, M.J.: Adaptive optics for high-resolution imaging. Nature Reviews Methods Primers **1**(1), 68 (2021)
19. He, B., Wang, Z., Zhou, Y., Chen, J., Singh, C.D., Li, H., Gao, Y., Shen, S., Wang, K., Cao, Y., et al.: Microsaccade-inspired event camera for robotics. Science Robotics **9**(90), eadj8124 (2024)
20. Hu, Y., Liu, S.C., Delbruck, T.: v2e: From video frames to realistic DVS events. In: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE (2021)
21. Ihrke, I., Restrepo, J., Mignard-Debise, L.: Principles of light field imaging: Briefly revisiting 25 years of research. IEEE Signal Processing Magazine **33**(5), 59–69 (2016)
22. Jiang, W., Boominathan, V., Veeraraghavan, A.: NeRT: Implicit neural representations for unsupervised atmospheric turbulence mitigation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 4236–4243 (2023)
23. Jiang, W., Guo, H., Metzler, C.A., Veeraraghavan, A.: Guidestar-free adaptive optics with asymmetric apertures. ACM Transactions on Graphics (2026)
24. Kingma, D., Ba, J.: Adam: A method for stochastic optimization. In: International Conference on Learning Representations (ICLR). San Diego, CA, USA (2015)
25. Kong, F., Lambert, A., Joubert, D., Cohen, G.: Shack-hartmann wavefront sensing using spatial-temporal data from an event-based image sensor. Optics Express **28**(24), 36159–36175 (2020)
26. Lai, W.S., Huang, J.B., Wang, O., Shechtman, E., Yumer, E., Yang, M.H.: Learning blind video temporal consistency. In: European Conference on Computer Vision (2018)
27. Li, H., Fan, R., Guan, J., Hao, W., Rui, L., Wu, T., Wang, Y., Gu, L.: EGTM: Event-guided efficient turbulence mitigation. arXiv preprint arXiv:2509.03808 (2025)
28. Li, N., Thapa, S., Whyte, C., Reed, A.W., Jayasuriya, S., Ye, J.: Unsupervised non-rigid image distortion removal via grid deformation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). pp. 2522–2532 (October 2021)
29. Liu, Y., Teng, M., Xia, Y., Duan, P., Shi, B.: EvTurb: Event camera guided turbulence removal. IEEE Transactions on Computational Imaging (2026)
30. Loktev, M., Soloviev, O., Savenko, S., Vdovin, G.: Speckle imaging through turbulent atmosphere based on adaptable pupil segmentation. Optics letters **36**(14), 2656–2658 (2011)
31. Manakov, A., Restrepo, J., Klehm, O., Hegedus, R., Eisemann, E., Seidel, H.P., Ihrke, I.: A reconfigurable camera add-on for high dynamic range, multispectral, polarization, and light-field imaging. ACM Transactions on Graphics **32**(4), 47–1 (2013)
32. Mao, Z., Chimitt, N., Chan, S.H.: Image reconstruction of static and dynamic scenes through anisoplanatic turbulence. IEEE Transactions on Computational Imaging **6**, 1415–1428 (2020)
33. Mao, Z., Chimitt, N., Chan, S.H.: Accelerating atmospheric turbulence simulation via learned phase-to-space transform. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) (October 2021)

34. Nah, S., Baik, S., Hong, S., Moon, G., Son, S., Timofte, R., Lee, K.M.: NTIRE 2019 challenge on video deblurring and super-resolution: Dataset and study. In: CVPR Workshops (June 2019)
35. Ng, R., Hanrahan, P.M.: Digital correction of lens aberrations in light field photography. In: International Optical Design Conference. p. WB2. Optica Publishing Group (2006)
36. Ng, R., Levoy, M., Brédif, M., Duval, G., Horowitz, M., Hanrahan, P.: Light field photography with a hand-held plenoptic camera. Ph.D. thesis, Stanford university (2005)
37. Pellizzari, C., Strong, D., Hardy, T., Metzler, C., Spencer, M.: Speckle-robust digital adaptive optics for coherent reflective imaging (2026)
38. Qu, Z., Zou, Z., Boominathan, V., Chakravarthula, P., Pediredla, A.: Event fields: Capturing light fields at high speed, resolution, and dynamic range. In: Proceedings of the Computer Vision and Pattern Recognition Conference. pp. 26910–26920 (2025)
39. Rebecq, H., Ranftl, R., Koltun, V., Scaramuzza, D.: High speed and high dynamic range video with an event camera. IEEE Transactions on Pattern Analysis and Machine Intelligence **43**(6), 1964–1980 (2019)
40. Scheerlinck, C., Rebecq, H., Gehrig, D., Barnes, N., Mahony, R., Scaramuzza, D.: Fast image reconstruction with an event camera. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. pp. 156–163 (2020)
41. Shimizu, M., Yoshimura, S., Tanaka, M., Okutomi, M.: Super-resolution from image sequence under influence of hot-air optical turbulence. In: 2008 IEEE Conference on Computer Vision and Pattern Recognition. pp. 1–8 (2008)
42. Stoffregen, T., Scheerlinck, C., Scaramuzza, D., Drummond, T., Barnes, N., Kleeman, L., Mahony, R.: Reducing the sim-to-real gap for event cameras. In: European Conference on Computer Vision. pp. 534–549. Springer (2020)
43. Wang, C., Zhu, S., Zhang, P., Huang, J., Wang, K., Lam, E.Y.: Neuromorphic shack-hartmann wave normal sensing. arXiv preprint arXiv:2404.15619 (2024)
44. Watnik, A.T., Gardner, D.F.: Wavefront sensing in deep turbulence. Optics and Photonics News **29**(10), 38–45 (2018)
45. Weng, W., Zhang, Y., Xiong, Z.: Event-based video reconstruction using transformer. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 2563–2572 (2021)
46. Wu, C., Ko, J., Davis, C.C.: Imaging through strong turbulence with a light field approach. Opt. Express **24**(11), 11975–11986 (May 2016)
47. Wu, C., Paulson, D.A., Rzasa, J.R., Davis, C.C.: Comparison between the plenoptic sensor and the light field camera in restoring images through turbulence. OSA Continuum **2**(9), 2511–2525 (2019)
48. Wu, G., Masia, B., Jarabo, A., Zhang, Y., Wang, L., Dai, Q., Chai, T., Liu, Y.: Light field image processing: An overview. IEEE Journal of Selected Topics in Signal Processing **11**(7), 926–954 (2017)
49. Wu, J., Guo, Y., Deng, C., Zhang, A., Qiao, H., Lu, Z., Xie, J., Fang, L., Dai, Q.: An integrated imaging sensor for aberration-corrected 3d photography. Nature **612**(7938), 62–71 (2022)
50. Xie, M., Guo, H., Feng, B.Y., Jin, L., Veeraraghavan, A., Metzler, C.A.: WaveMo: learning wavefront modulations to see through scattering. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 25276–25285 (2024)

51. Yang, J.C., Everett, M., Buehler, C., McMillan, L.: A real-time distributed light field camera. In: Proceedings of the 13th Eurographics Workshop on Rendering. p. 77–86. EGRW '02, Eurographics Association, Goslar, DEU (2002)
52. Yeminy, T., Katz, O.: Guidestar-free image-guided wavefront shaping. Science advances **7**(21), eabf5364 (2021)
53. Yoshida, Y., Miyato, T.: Spectral norm regularization for improving the generalizability of deep learning. arXiv preprint arXiv:1705.10941 (2017)
54. Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O.: The unreasonable effectiveness of deep features as a perceptual metric. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2018)
55. Zhang, X., Chimitt, N., Chi, Y., Mao, Z., Chan, S.H.: Spatio-temporal turbulence mitigation: A translational perspective. In: 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 2889–2899. IEEE Computer Society (2024)
56. Zhang, X., Chimitt, N., Wang, X., Yuan, Y., Chan, S.H.: Learning phase distortion with selective state space models for video turbulence mitigation. In: Proceedings of the Computer Vision and Pattern Recognition Conference. pp. 2127–2138 (2025)
57. Zhang, X., Mao, Z., Chimitt, N., Chan, S.H.: Imaging through the atmosphere using turbulence mitigation transformer. IEEE Transactions on Computational Imaging **10**, 115–128 (2024)
58. Ziemann, M.R., Rathbun, I.A., Metzler, C.A.: A learning-based approach to event-based shack-hartmann wavefront sensing. In: Unconventional Imaging, Sensing, and Adaptive Optics 2024. vol. 13149, pp. 199–207. SPIE (2024)