Review Article

A Survey on Event-driven 3D Reconstruction: Development under Different Categories

Chuanzhi Xu[1], Haoxian Zhou[1], Haodong Chen[1], Vera Chung[1], Vincent Qu[1]

1. School of Computer Science, University of Sydney, Australia

Event cameras have gained increasing attention for 3D reconstruction due to their high temporal resolution, low latency, and high dynamic range. They capture per-pixel brightness changes asynchronously, allowing accurate reconstruction under fast motion and challenging lighting conditions. In this survey, we provide a comprehensive review of event-driven 3D reconstruction methods, including stereo, monocular, and multimodal systems. We further categorize recent developments based on geometric, learning-based, and hybrid approaches. Emerging trends, such as neural radiance fields and 3D Gaussian splatting with event data, are also covered. The related works are structured chronologically to illustrate the innovations and progression within the field. To support future research, we also highlight key research gaps and future research directions in dataset, experiment, evaluation, event representation, etc.

Corresponding author: Chuanzhi Xu, chxu4146@uni.sydney.edu.au

I. Introduction

Event cameras, also known as neuromorphic cameras, silicon retina, or dynamic vision sensors, are bio-inspired sensors that respond asynchronously to changes in brightness. Unlike traditional RGB cameras with a fixed frame rate, each pixel in event cameras operates independently and asynchronously as a sub-sensor. When a brightness change surpasses a certain threshold, the pixel records the coordinates, timestamp, and polarity of the brightness change - this is known as event data.

The application of event cameras in 3D reconstruction has been mainly explored since the 2010s[1][2], whereas traditional camera-based 3D reconstruction has been studied since the 1990s. Referring to

Figure 1, 3D reconstruction can be done with stereo and monocular event cameras, and multimodal methods with event cameras. Recent methods also incorporate NeRF and 3D Gaussian Splatting to achieve high-fidelity reconstructions. The number of related studies is steadily increasing. However, no dedicated survey on event-driven 3D reconstruction can provide a summary and guidance for future research.

Author (Method)	Year	Priors	Event Rep.	Real-time	Output (Dense?)	Dataset
Zhu et al. (TSES)[3]	2018	Velocity	Event disparity volume	✓	3D depth map (✓)	MVSEC
Zhou et al.[4]	2018	Pose	Time surface	✗	3D depth map (✗)	MVSEC
Lin et al.[3]	2018	Pose	Time surface	✓	Point cloud (✗)	-
Steffen et al.[5]	2019	-	4D vector	✗	Voxel (✗)	-
Zhou et al. (ESVO) [6]	2021	Pose	Time surface	✓	Point cloud, depth map (✗)	RPG, ESIM
Nam et al. (SE-CFF) [7]	2022	Pose	Event Stack	✓	3D depth map (✓)	DSEC
Ghosh et al.[8]	2022	Pose	Disparity space image	✗	3D depth map (✗)	MVSEC, DSEC, ESIM

Table I. Methods with Stereo Event Cameras

In the field of event cameras, three key survey papers have been published. The first, by Gallego et al. (2020), included a section on 3D reconstruction[9], but parts of it are now outdated due to the rapid growth of deep learning. The survey by Chakravarthi et al. (2024) categorizes various tasks related to event cameras[10], but it lacks detailed discussions on specific domains, particularly 3D reconstruction. Similarly, the survey by Zheng et al. (2024) focuses exclusively on deep learning-based event camera tasks[11], but it only mentions 3D reconstruction in one subsection and does not provide a sufficient historical perspective on the evolution of 3D reconstruction. Additionally, no other survey has specifically mentioned event-driven 3D reconstruction. Therefore, in this survey, we aim to:

1. Provide a comprehensive review of the event-driven 3D reconstruction domain.
2. Establish a coherent categorization of the diverse event-based 3D reconstruction methods.
3. Present a timeline of technical advancements in event-driven 3D reconstruction.
4. Highlight existing research gaps and promising future directions for event-driven 3D reconstruction.

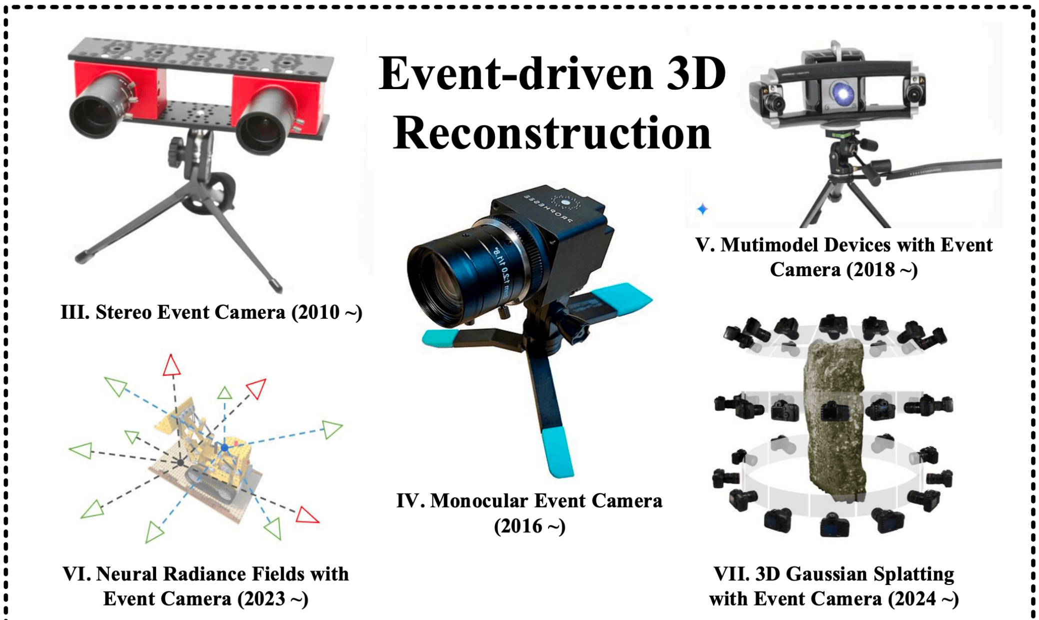

Figure 1. Different categories of event-driven 3D reconstruction.

II. Event Camera

When the event camera detects a brightness change at a pixel k, it generates event data containing event coordinates $\mathbf{x}_k = (x_k, y_k)$, timestamp t_k and the polarity p_k. The brightness $L(\mathbf{x}_k, t) = \log(I(\mathbf{x}_k, t))$ is set as the pixel's log intensity. The brightness change threshold C usually varies by 10-15%[12]. An event $e_k = (\mathbf{x}_k, t_k, p_k)$ is triggered when the brightness change ΔL at pixel k exceeds C, which can be expressed as:

$$|\Delta L(\mathbf{x}_k, t_k)| = |L(\mathbf{x}_k, t_k) - L(\mathbf{x}_k, t_{k-1})| \geq |p_k \cdot C| \tag{1}$$

where t_{k-1} represents the timestamp of the last event at the same pixel. The polarity value p_k is determined as follows:

$$p_k = \begin{cases} +1, & \text{if } \Delta L(\mathbf{x}_k, t_k) \geq C \\ -1, & \text{if } \Delta L(\mathbf{x}_k, t_k) \leq -C \\ \text{No event}, & \text{if -C } < \Delta L(\mathbf{x}_k, t_k) < \text{C} \end{cases} \tag{2}$$

When an event camera continuously captures events, it forms an event stream, which can be represented as a sequence of events ordered by timestamps:

$$\text{EventStream} = \{(t_k, x_k, y_k, p_k)\}_{k=1}^{N} \tag{3}$$

where N denotes the total number of recorded events.

Event cameras have some characteristics that distinguish them from traditional cameras. Event cameras offer microsecond-level temporal resolution, enabling the capture of rapid motion without the motion blur. With a dynamic range exceeding 120 dB, they effectively adapt to both extremely low and high illumination conditions. Unlike traditional cameras, event cameras respond solely to pixel intensity changes, leading to significantly reduced data bandwidth, lower power consumption, and minimal latency. These characteristics make event cameras highly applicable across a wide range of fields, such as object tracking[13], corner detection[14], object recognition[15], depth estimation[16], video generation[17], light field video enhancement[18], and 3D reconstruction, etc.

doi.org/10.32388/B4ZZQI

III. Methods with Stereo Event Cameras

Author (Method)	Year	Priors	Event Rep.	Output (Dense?)	Dataset
Kim et al.[19]	2016	Trajectory	Event gradient accumulation	3D depth map (✗)	-
Rebecq et al. (EVO)[20]	2016	Pose	Event frame	3D depth map (✗)	Multi-Keyframe
Rebecq et al. (EMVS)[12]	2018	Trajectory	Event-by-event	3D depth map (✗)	-
Guan et al. (EVI-SAM) [21]	2024	Trajectory	Time surface	3D mapping (✓)	DAVIS240c
Elms et al. (eSfO)[22]	2024	Trajectory	Time surface	Sparse point cloud (✗)	TOPSPIN

Table II. Monocular Event Camera: Real-time Geometry-based Methods

Author (Method)	Year	Priors	Pipeline	Event Rep.	Model	Output	Dataset
Baudron et al. (E3D)[23]	2020	Silhouette	✓	Event accumulate frames	E2S(CNN)	Mesh	ShapeNet
Xiao et al.[24]	2022	Pose	✓	Intensity image	E2VID(RNN-CNN)	Mesh	ESIM
Wang et al. (Evac3d)[24]	2023	Contour, Trajec.	✓	Voxel grid	Evac3d(CNN)	Mesh	MOEC-3D
Chen et al. (E2V) [25]	2023	-	✗	Event frame	ResNet-152	Voxel	SynthEVox3D
Xu et al.[26]	2025	-	✗	Sobel event frame	ResNet-152	Voxel	SynthEVox3D

Table III. Monocular Event Camera: Deep Learning-based Methods

doi.org/10.32388/B4ZZQI

Stereo event cameras typically refer to two or more rigidly mounted event cameras. Event-driven 3D-related tasks were initially pioneered using stereo event cameras in the 2010s[27][2]. However, many prior works, such as[28] and those in SLAM and visual odometry (VO), can only be regarded as prior estimation steps in the broader pipeline of 3D reconstruction. In this section, we focus only on methods that perform complete event-to-3D representations.

Many earlier approaches perform depth mapping reconstruction by computing disparity between events observed at the same timestamp across different viewpoints, followed by geometry-based multi-view stereo estimation to achieve real-time 3D depth reconstruction. In 2018, Zhu et al.[3] proposed a method that synchronizes events in time using known camera velocities, constructs a dense event disparity volume, and performs real-time sliding window matching, introducing a novel matching cost function combining ambiguity and similarity. Simultaneously, Lin et al.[29] reformulated event matching as a time-difference consistency problem across views.

However, disparity matching can be bypassed as well. In 2018, Zhou et al.[4] proposed a forward-projection based depth estimation method by directly optimizing a temporal consistency energy across stereo time surfaces, without requiring disparity computation. Later in 2021, Zhou et al.[6] extended this concept by integrating stereo time surfaces with a stereo visual odometry framework, optimizing a spatio-temporal consistency objective for real-time semi-dense reconstruction. In 2022, Ghosh et al. [8] proposed a robust stereo depth estimation framework that fuses multi-view event ray densities (via Disparity Space Image, DSI), achieving high-quality depth estimation without explicit disparity matching.

Additionally, one research explores unsupervised clustering for self-organizing structural modeling. In 2019, Steffen et al.[5] employed Self-Organizing Maps (SOMs) to embed events from multiple viewpoints into a high-dimensional space, enabling sparse voxelized 3D reconstruction under uncalibrated and unsupervised conditions.

More recent methods leverage neural networks to learn stereo disparity estimation and event representations. In 2022, Nam et al.[7] introduced a deep learning framework that combines multi-density event stacking with attention mechanisms using a UNet + ResNet encoder. By incorporating future-event prediction during training, their method achieves high-accuracy, real-time, and dense stereo event-based depth estimation.

 doi.org/10.32388/B4ZZQI

IV. Methods with Monocular Event Cameras

Monocular event cameras cannot directly obtain disparity information as in stereo methods, requiring additional prior estimation. These methods can be categorized as:

A. Geometry-based methods

Geometry-based methods typically achieve semi-dense real-time 3D reconstruction, relying on spatial scanning with a monocular event camera. One critical step is to estimate the physical prior.

In 2016, Kim et al.[19] proposed an approach utilizing three interleaved probabilistic filters to estimate camera trajectory, scene log-intensity gradient, and inverse depth. Rebecq et al.[20] introduced EVO, leveraging event projection, edge alignment, and DSI construction for 3D reconstruction. In the same year, Rebecq et al.[12], later in 2018, proposed EMVS, employing event space-sweep and ray density analysis to directly generate a semi-dense 3D depth map, without frame-level data association.

A recent study has improved geometry-based methods to achieve dense reconstruction. In 2024, Guan et al.[21] proposed EVI-SAM, a tightly coupled event-image-IMU SLAM framework. It achieves real-time dense 3D reconstruction on a standard CPU, integrating event-based 2D-2D alignment, image-guided depth interpolation, and TSDF fusion.

A unique innovation enables reconstruction when the event camera is stationary, while the object rotates. In 2024, Elms et al.[22] proposed eSfO, which performs 3D reconstruction through event corner tracking and factor graph optimization, but it only perform non-real-time sparse point cloud reconstruction.

B. Deep learning-based methods

Deep learning-based methods typically produce non-real-time dense reconstruction. However, traditional RGB image feature extraction techniques cannot be directly applied to event data, and raw events are also difficult to use as neural network inputs[30].

In 2020, Baudron et al.[23] proposed E3D, the first dense 3D shape reconstruction method based on monocular event cameras. It employs the E2S neural network to estimate silhouettes and leverages PyTorch3D for 3D mesh optimization, achieving high-quality multi-view 3D reconstruction trained on ShapeNet. In 2022, Xiao et al.[24] proposed a pipeline using the E2VID deep learning method[31] to process continuous event streams and generate normalized intensity image sequences. They then employed SfM

doi.org/10.32388/B4ZZQI

to estimate intrinsic and extrinsic parameters for sparse point clouds and used MVS for dense mesh reconstruction. In 2023, Wang et al.[32] proposed EvAC3D, which uses CNN to predict Apparent Contour Events (ACE), combined with Continuous Volume Carving and Global Mesh Optimization, to achieve dense 3D shape reconstruction with known camera trajectories.

Many studies have established an event-to-3D pipeline, a structured and modular event processing framework[23][24][30][20], including feature extraction, matching, and 3D computation. The extraction and estimation of priors are also essential. However, recent methods aim to eliminate the pipeline and priors. In 2023, Chen et al.[25] proposed E2V, which employs a modified ResNet-152 and a U-Net 3D decoder to directly predict dense 3D voxel grids from monocular event frames, achieving event-based 3D reconstruction without external priors. In 2025, Xu et al.[26] extended E2V by introducing a novel event representation, Sobel Event Frame, and an optimal binarization strategy for event-based 3D reconstruction. By enhancing E2V with ECA[33], their method significantly improved reconstruction quality.

V. Multimodal Methods with Event Cameras

Author (Method)	Year	Device	Priors	Event Rep.	Real-time	Output (Dense?)
Leroux et al.[34]	2018	Structured light	Pose	Time surface	✓	Point cloud (✓)
Huang et al.[35]	2021	Structured light	Pose	Event-by-event	✗	Point cloud (✓)
Zuo et al. (Devo)[36]	2022	D-RGB camera	Trajectory	Time surface	✓	Point cloud (✗)
Xiao et al.[37]	2023	Structured light	Pose	Event frame	✓	Point cloud (✓)
Fu et al.[38]	2023	Structured light	Pose	Time surface	✓	Point cloud (✓)
Li et al.[39]	2024	Structured light	Pose	Event-by-event	✓	Point cloud (✓)

Table IV. Multimodal Methods with event cameras

Some recent multimodal approaches combine event cameras with structured light or depth-RGB sensors

doi.org/10.32388/B4ZZQI

to improve 3D reconstruction. These systems enhance robustness through asynchronous processing, joint calibration, and information fusion, enabling real-time 3D point cloud generation under high-speed motion or low-light conditions. As most of them rely on self-collected datasets, they are not included in Table IV.

Structured light, an active 3D sensing technique, projects coded patterns onto surfaces and reconstructs depth via triangulation. Leroux et al.[34] used frequency-coded structured light with event cameras to recover depth. Huang et al.[35] combined structured light projection with digital image correlation (DIC) for high-speed scanning. Xiao et al.[37] employed alternating binary speckle patterns and DIC-based stereo matching for fast and accurate reconstruction. Fu et al.[38] introduced spatio-temporal coding (STC) with an enhanced matching scheme (STEM) for improved stereo robustness. Li et al.[39] proposed eFPSL, using time-frequency analysis to extract high-SNR fringe maps from events and an event-count-based shadow mask to reduce errors.

Some methods fuse event data with depth-RGB (D-RGB) sensors for improved scene understanding. Zuo et al.[36] proposed DEVO, combining time surface maps from events and depth supervision from a calibrated sensor. Their system performs semi-dense 3D-2D edge alignment to estimate poses and incrementally build point clouds under fast motion and poor lighting.

 doi.org/10.32388/B4ZZQI

VI. Neural Radiance Fields with Event Cameras

Author	Model	Yr-Mo	Inputs	Event Rep.	Colorful Recon.	Dataset
Klenk et al.[40]	E-NeRF	2023-01	Event stream	EAF	✗	ESIM
Hwang et al.[41]	Ev-NeRF	2023-03	Event stream	EAF	✗	IJRR, HQF
Rudnev et al.[42]	EventNeRF	2023-03	RGB bayer event stream	EAF	✓	NeRF dataset
Qi et al.[43]	E^2NeRF	2023-10	Blurry RGB, Event stream	EAF	✓	NeRF dataset
Bhattacharya et al.[44]	EvDNeRF	2023-12	Event stream	EAF	✗	Real-Fork
Cannici et al.[45]	Ev-DeblurNeRF	2024-06	Blurry RGB, Event stream	EAF	✓	Ev-DeblurBlender
Wang et al.[30]	NeRF(Enhanced)	2024-05	Event stream	EAF	✗	PAEv3D
Feng et al.[46]	AE-NeRF	2025-01	Event stream	Event-by-event	✗	TUM-VIE

doi.org/10.32388/B4ZZQI

Author	Model	Yr-Mo	Inputs	Event Rep.	Colorful Recon.	Dataset
Weng et al.[47]	EaDeblur-GS	2024-09	Blurry RGB, Event stream	EGA	✓	E^2NeRF
Wu et al.[48]	Ev-GS	2024-09	Event stream	EGA	✗	NeRF dataset
Deguchi et al.[49]	E2GS	2024-10	Blurry RGB, Event stream	EAF	✓	NeRF dataset
Xiong et al.[50]	Event3DGS	2024-10	Blurry RGB, Event stream	EGA	✓	EventNeRF
Han et al.[51]	Event-3DGS	2024-10	Blurry RGB, Event stream	EAF	✓	DeepVoxels
Huang et al.[52]	IncEventGS	2024-10	Event stream	EAF	✗	Replica dataset
Yu et al.[53]	EvaGaussians	2024-12	Blurry RGB, Event stream	EGA	✓	EvaGaussians
Yura et al.[54]	EventSplat	2024-12	RGB bayer event stream	EAF	✓	NeRF dataset

Table V. Neural Radiance Fields Methods with Event Cameras (EAF: Event accumulate frame)

Author	Model	Yr-Mo	Inputs	Event Rep.	Colorful Recon.	Dataset
Weng et al.[47]	EaDeblur-GS	2024-09	Blurry RGB, Event stream	EGA	✓	E^2NeRF
Wu et al.[48]	Ev-GS	2024-09	Event stream	EGA	✗	NeRF dataset
Deguchi et al.[49]	E2GS	2024-10	Blurry RGB, Event stream	EAF	✓	NeRF dataset
Xiong et al.[50]	Event3DGS	2024-10	Blurry RGB, Event stream	EGA	✓	EventNeRF
Han et al.[51]	Event-3DGS	2024-10	Blurry RGB, Event stream	EAF	✓	DeepVoxels
Huang et al.[52]	IncEventGS	2024-10	Event stream	EAF	✗	Replica dataset
Yu et al.[53]	EvaGaussians	2024-12	Blurry RGB, Event stream	EGA	✓	EvaGaussians
Yura et al.[54]	EventSplat	2024-12	RGB bayer event stream	EAF	✓	NeRF dataset

Table VI. 3D Gaussian Methods with Event Cameras (EAF: Event accumulate frame, EGA: Event gradient accumulation)

Neural Radiance Fields (NeRF) is a neural network-based method for representing 3D scenes, proposed by Mildenhall et al. in 2020[55]. It learns a continuous and differentiable 3D radiance field from multi-view 2D images and synthesizes novel views. Since 2023, NeRF-based methods have been adapted for event-based 3D reconstruction. These approaches typically use event accumulation frames and assume known or estimated camera trajectories. By modeling brightness changes and event-triggering probability, they enable dense 3D reconstruction and novel view synthesis.

doi.org/10.32388/B4ZZQI

Methods without RGB input focus solely on events, capturing grayscale scene structure under challenging conditions. Klenk et al.[40] proposed E-NeRF, using an event-triggered brightness model and no-event loss for high-fidelity reconstruction. Hwang et al.[41] introduced Ev-NeRF, aggregating volumetric event data via multi-view consistency. Bhattacharya et al.[44] proposed EvDNeRF, a dynamic NeRF model supporting rigid and non-rigid motion. Wang et al.[30] introduced Physical Priors Augmented EventNeRF with motion priors and geometric consistency loss. Feng et al.[46] proposed AE-NeRF, combining pose correction with a two-stage event-based NeRF for reconstruction from noisy, asynchronous data.

Methods incorporating RGB images combine events with visual texture for color reconstruction. Rudnev et al.[42] proposed EventNeRF with self-supervised training and event-based rendering. Qi et al. [43] introduced E^{2}NeRF, jointly using blurry images and events with blur and event rendering losses. Cannici et al.[45] proposed Ev-DeblurNeRF, using an event double integral (EDI), learnable event response function (eCRF), and explicit feature volumes to reconstruct under extreme motion blur.

VII. 3D Gaussian Splatting with Event Cameras

3D Gaussian Splatting, proposed by Kerbl et al.[56], represents a volumetric primitive in 3D space with attributes such as position, shape, orientation, and color. In computer graphics, it serves as an explicit 3D representation that enables efficient differentiable rendering.

Recently, combining event cameras with 3D Gaussian Splatting has become a new trend in implicit 3D reconstruction. These methods typically use event accumulation frames or gradient maps as input and apply Gaussian modeling with differentiable rendering to achieve smooth and scalable 3D reconstruction and novel view synthesis (NVS). They usually assume known or estimated camera trajectories, aiming for dense reconstruction with real-time rendering performance.

First, methods without RGB input rely solely on event streams, focusing on brightness change, edges, and temporal cues. They are well-suited for low-light or image-infeasible environments. Wu et al. [48] introduced Ev-GS, the first Gaussian splatting framework driven by event data, using a logarithmic brightness accumulation model and a lightweight renderer. Huang et al.[52] proposed IncEventGS, which follows a SLAM-style “tracking and mapping” design and jointly optimizes motion and scene structure using continuous-time trajectory modeling.

 doi.org/10.32388/B4ZZQI

Second, methods that combine events with RGB images fuse texture and edge information for higher-quality reconstructions. Deguchi et al.[49] proposed E2GS, using the Event Double Integral (EDI) and event-based loss for blur removal. Yu et al.[53] introduced EvaGaussians with learnable pose offsets and event supervision. Weng et al.[47] developed EaDeblur-GS, featuring adaptive pose correction and combined loss for fast and accurate results. Xiong et al.[50] presented Event3DGS, using DSSIM loss, sparsity-aware sampling, and progressive training. Yura et al.[54] proposed EventSplat, initializing reconstruction via event-to-video SfM and cubic spline interpolation. Han et al.[51] introduced Event-3DGS, which integrates photovoltage estimation, contrast-based rendering, and event-driven loss for robust dense reconstruction in challenging conditions.

VIII. Event-driven 3D Reconstruction: Research Gaps & Future Direction

Despite recent advances, event-driven 3D reconstruction still faces key challenges across simulation, evaluation, modeling, and deployment.

First, datasets and benchmarks can be explored. The main modeling platforms do not support event-based 3D modeling. While simulating tools like ESIM[57] and Video-to-Event[58] exist, datasets for event-based 3D reconstruction remain limited, and the methods proposed so far rely on inconsistent and non-standardized data. The reconstruction outputs vary greatly across methods due to differences in datasets, result types, and metrics, hindering fair comparison. A unified dataset and standardized evaluation metric for accuracy and speed are needed to benchmark progress.

Second, event representation can be explored. Event representation significantly affects feature extraction, yet its impact on 3D reconstruction quality and efficiency remains underexplored. Although many representations have been proposed[59][60][26], no systematic comparison exists within a unified pipeline.

Third, experiments under extreme scenarios and with special objects remain insufficient. Although event cameras are known to perform well under extreme conditions (e.g., brightness, speed, low light), few studies benchmark them against traditional cameras in such scenarios[26]. Reconstruction of non-Lambertian objects like mirrors and glass also remains underdeveloped, with only limited attempts such as EventPS[61].

doi.org/10.32388/B4ZZQI

From a system perspective, sparse and uneven event data limits reconstruction in low-texture or static regions; polarity cancellation further weakens structure encoding[7][8][21][25][50]. Many methods rely on precise synchronization and accurate extrinsics, posing challenges for practical deployment[8][21][22][42][43][53][47]. Most approaches assume static scenes and struggle with dynamic or deformable objects[41][42][40][49][47].

In terms of efficiency, NeRF-based methods are computationally expensive and slow to train and render[41][43][53]. Even fast Gaussian Splatting pipelines have yet to reach real-time performance in event-based settings[50][62]. Under blur, reflections, or low light, outputs remain noisy or sparse[32][38][43][47][21][36].

Future work should focus on dynamic scene modeling[22][41][42][49][53], asynchronous multimodal fusion (events + images + IMU)[8][21][36], efficient and lightweight networks[7][25][49][50], and self-/weak supervision for generalization[41][43]. Enhancing robustness to extreme conditions and improving pose estimation, energy efficiency, and privacy support are also promising directions[38][47][7].

IX. Conclusion

This survey provides a structured review of event-driven 3D reconstruction, covering stereo, monocular, and multimodal methods. We categorized approaches by different devices and highlighted emerging trends of NeRF and 3D Gaussian Splatting. In addition to summarizing key technical advances, we also presented a chronological timeline to reflect the evolution of the field. However, significant challenges remain in standardized datasets, evaluation metrics, event representation, and handling dynamic or extreme scenes, etc. We also suggested several directions for improvement. We hope this work serves as both a comprehensive reference and a guide for future research in event-driven 3D reconstruction.

References

1. ^ *Carneiro J, Ieng S-H, Posch C, Benosman R (2013). "Event-based 3D reconstruction from neuromorphic retinas". Neural Networks.* ***45****: 27–38.*
2. a, b *Piatkowska E, Belbachir A, Gelautz M (2013). "Asynchronous stereo vision for event-driven dynamic stereo sensor using an adaptive cooperative approach." In: ICCV Workshops. pp. 45--50.*

 doi.org/10.32388/B4ZZQI

3. a, b, c Zhu AZ, Chen Y, Daniilidis K (2018). "Realtime time synchronized event-based stereo." In: ECCV. pp. 433–447.
4. a, b Zhou Y, Gallego G, Rebecq H, Kneip L, Li H, Scaramuzza D. "Semi-dense 3D reconstruction with a stereo event camera." In: ECCV; 2018. p. 235-251.
5. a, b Steffen L, Ulbrich S, Roennau A, Dillmann R. "Multi-view 3d reconstruction with self-organizing maps on event-based data." In: 2019 ICAR. IEEE; 2019. p. 501–508.
6. a, b Zhou Y, Gallego G, Shen S (2021). "Event-based stereo visual odometry". IEEE Transactions on Robotics. **37** (5): 1433–1450.
7. a, b, c, d, e Nam Y, Mostafavi M, Yoon KJ, Choi J (2022). "Stereo depth from events cameras: Concentrate and focus on the future." In: CVPR. pp. 6114–6123.
8. a, b, c, d, e Ghosh S, Gallego G (2022). "Multi-Event-Camera Depth Estimation and Outlier Rejection by Refocused Events Fusion". Advanced Intelligent Systems. **4** (12): 2200221.
9. ^ Gallego G, Delbrück T, Orchard G, Bartolozzi C, Taba B, Censi A, Leutenegger S, Davison AJ, Conradt J, Daniilidis K, et al. (2020). "Event-based vision: A survey." IEEE Transactions on Pattern Analysis and Machine Intelligence. **44** (1): 154–180.
10. ^ Chakravarthi B, Verma AA, Daniilidis K, Fermüller C, Yang Y (2024). "Recent event camera innovations: A survey". arXiv:2408.13627.
11. ^ Zheng X, Liu Y, Lu Y, Hua T, Pan T, Zhang W, Tao D, Wang L (2023). "Deep learning for event-based vision: A comprehensive survey and benchmarks". arXiv:2302.08890.
12. a, b, c Rebecq H, Gallego G, Mueggler E, Scaramuzza D (2018). "EMVS: Event-based multi-view stereo⊠3D reconstruction with an event camera in real-time." International Journal of Computer Vision. **126** (12): 1394⊠14.
13. ^ Zuo Y, Xu W, Wang X, Wang Y, Kneip L (2024). "Cross-modal Semidense 6-DOF tracking of an event camera in challenging conditions." IEEE Transactions on Robotics. **40**: 1600–1616.
14. ^ Zhang N, Han S, Chen X, Chen H, Tan L, Chung YY (2024). "Event vision-based corner detection with count-normalized multi-layer perceptron and throughput indicator." Computers and Electrical Engineering. **118**: 109432.
15. ^ Millerdurai C, Luvizon D, Rudnev V, Jonas A, Wang J, Theobalt C, Golyanik V. "3d pose estimation of two interacting hands from a monocular event camera." In: 2024 3DV. IEEE; 2024. p. 291-301.
16. ^ Hidalgo-Carri⊠3 J, Gehrig D, Scaramuzza D. "Learning monocular dense depth from events." In: 2020 3DV. IEEE; 2020. p. 534-542.

17. ^Qu Q, Shen Y, Chen X, Chung YY, Liu T (2024). "E2HQV: High-quality video generation from event camera via theory-inspired model-aided deep learning." In: AAAI. **38** (5): 4632–4640.

18. ^Lu Z, Chen X, Chung YY, Cai W, Shen Y (2023). "EV-LFV: Synthesizing Light Field Event Streams from an Event Camera and Multiple RGB Cameras". IEEE TVCG. **29** (11): 4546–4555. doi:10.1109/TVCG.2023.3320271.

19. a, bKim H, Leutenegger S, Davison AJ. "Real-time 3D reconstruction and 6-DoF tracking with an event camera." In: ECCV. Springer; 2016. p. 349-364.

20. a, b, cRebecq H, Horstschaefer T, Gallego G, Scaramuzza D (2016). "Evo: A geometric approach to event-based 6-dof parallel tracking and mapping in real time." IEEE Robotics and Automation Letters. **2** (2): 593–600.

21. a, b, c, d, e, fGuan W, Chen P, Zhao H, Wang Y, Lu P (2024). "EVI-SAM: Robust, Real-Time, Tightly-Coupled Event--Visual--Inertial State Estimation and 3D Dense Mapping". Advanced Intelligent Systems. **6** (12): 2400243.

22. a, b, c, dElms E, Latif Y, Park TH, Chin T-J. "Event-based structure-from-orbit." In: CVPR; 2024. p. 19541-19550.

23. a, b, cBaudron A, Wang ZW, Cossairt O, Katsaggelos AK (2020). "E3d: event-based 3d shape reconstruction". arXiv:2012.05214.

24. a, b, c, dXiao K, Wang G, Chen Y, Nan J, Xie Y. "Event-based dense reconstruction pipeline." In: 2022 ICRAS. IEEE; 2022. p. 172-177.

25. a, b, c, dChen H, Chung V, Tan L, Chen X. "Dense voxel 3d reconstruction using a monocular event camera." In: 2023 ICVR. IEEE; 2023. p. 30-35.

26. a, b, c, dXu C, Chen L, Qu V, Chen H, Chung V (2025). "Towards End-to-End Neuromorphic Voxel-based 3D Object Reconstruction Without Physical Priors". arXiv:2501.00741.

27. ^Kogler J, Sulzbachner C, Humenberger M, Eibensteiner F (2011). "Chapter Address-Event Based Stereo Vision with Bio-Inspired Silicon Retina Imagers." IntechOpen. 2011.

28. ^Lin X, Qiu C, Shen S, Zang Y, Liu W, Bian X, Mfcller M, Wang C, et al. (2023). "E2pnet: event to point cloud registration with spatio-temporal representation learning." Advances in Neural Information Processing Systems. **36**: 18076–18089.

29. ^Ieng S-H, Carneiro J, Osswald M, Benosman R (2018). "Neuromorphic event-based generalized time-based stereovision." Frontiers in Neuroscience. **12**: 442.

30. a, b, c, dWang J, He J, Zhang Z, Xu R. "Physical priors augmented event-based 3d reconstruction." In: 2024 ICRA. IEEE; 2024. p. 16810-16817.

31. ^Rebecq H, Ranftl R, Koltun V, Scaramuzza D (2019). "High speed and high dynamic range video with an event camera". IEEE Transactions on Pattern Analysis and Machine Intelligence. **43** (6): 1964–1980.

32. [a, b] *Wang Z, Chaney K, Daniilidis K. "Evac3d: From event-based apparent contours to 3d models via continuous visual hulls." In: ECCV. Springer; 2022. p. 284-299.*
33. [^] *Xu C (2024). "The Application of Federated Learning in Face Recognition: A Systematic Investigation of The Existing Frameworks." Applied and Computational Engineering. **52**: 21–30.*
34. [a, b] *Leroux T, Ieng S-H, Benosman R (2018). "Event-based structured light for depth reconstruction using frequency tagged light patterns". arXiv:1811.10771. Available from: https://arxiv.org/abs/1811.10771.*
35. [a, b] *Huang X, Zhang Y, Xiong Z (2021). "High-speed structured light based 3D scanning using an event camera". Optics Express. **29** (22): 35864–35876.*
36. [a, b, c, d] *Zuo YF, Yang J, Chen J, Wang X, Wang Y, Kneip L. "Devo: Depth-event camera visual odometry in challenging conditions." In: 2022 ICRA. IEEE; 2022. p. 2179–2185.*
37. [a, b] *Xiao C, Chen X, Xi J, Li Z, He B. "Speckle-Projection-Based High-Speed 3D Reconstruction Using Event Cameras." In: 2023 CISP-BMEI. IEEE; 2023. p. 1-6.*
38. [a, b, c, d] *Fu J, Zhang Y, Li Y, Li J, Xiong Z (2023). "Fast 3d reconstruction via event-based structured light with spatio-temporal coding." Optics Express. **31** (26): 44588–44602.*
39. [a, b] *Li Y, Jiang H, Xu C, Liu L (2024). "Event-driven fringe projection structured light 3-D reconstruction based on time--frequency analysis". IEEE Sensors Journal. **24** (4): 5097–5106.*
40. [a, b, c] *Klenk S, Koestler L, Scaramuzza D, Cremers D (2023). "E-nerf: Neural radiance fields from a moving event camera". IEEE Robotics and Automation Letters. **8** (3): 1587–1594.*
41. [a, b, c, d, e, f] *Hwang I, Kim J, Kim YM (2023). "Ev-nerf: Event based neural radiance field." In: WACV. pp. 837–847.*
42. [a, b, c, d, e] *Rudnev V, Elgharib M, Theobalt C, Golyanik V. "Eventnerf: Neural radiance fields from a single colour event camera." In: CVPR; 2023. p. 4992–5002.*
43. [a, b, c, d, e, f] *Qi Y, Zhu L, Zhang Y, Li J. "E2nerf: Event enhanced neural radiance fields from blurry images." In: ICCV; 2023. p. 13254-13264.*
44. [a, b] *Bhattacharya A, Madaan R, Cladera F, Vemprala S, Bonatti R, Daniilidis K, Kapoor A, Kumar V, Matni N, Gupta JK. "Evdnerf: Reconstructing event data with dynamic neural radiance fields." In: WACV; 2024. p. 5846–5855.*
45. [a, b] *Cannici M, Scaramuzza D. "Mitigating motion blur in neural radiance fields with events and frames." In: CVPR, 2024. p. 9286–9296.*
46. [a, b] *Feng C, Yu W, Cheng X, Tang Z, Zhang J, Yuan L, Tian Y (2025). "AE-NeRF: Augmenting Event-Based Neural Radiance Fields for Non-ideal Conditions and Larger Scene". arXiv:2501.02807.*

47. [a, b, c, d, e, f, g] *Weng Y, Shen Z, Chen R, Wang Q, Wang J (2024). "Eadeblur-gs: Event assisted 3d deblur reconstruction with gaussian splatting". arXiv:2407.13520.*

48. [a, b, c] *Wu J, Zhu S, Wang C, Lam EY. "Ev-GS: Event-based gaussian splatting for efficient and accurate radiance field rendering." In: 2024 MLSP. IEEE; 2024. p. 1-6.*

49. [a, b, c, d, e, f] *Deguchi H, Masuda M, Nakabayashi T, Saito H. "E2gs: Event enhanced gaussian splatting." In: 2024 ICIP. IEEE; 2024. p. 1676–1682.*

50. [a, b, c, d, e, f] *Xiong T, Wu J, He B, Fermuller C, Aloimonos Y, Huang H, Metzler CA (2024). "Event3dgs: Event-based 3d gaussian splatting for high-speed robot egomotion". arXiv:2406.02972.*

51. [a, b, c] *Han H, Li J, Wei H, Ji X (2024). "Event-3DGS: Event-based 3D Reconstruction Using 3D Gaussian Splatting". Advances in Neural Information Processing Systems.* ***37****: 128139–128159.*

52. [a, b, c] *Huang J, Dong C, Liu P (2024). "IncEventGS: Pose-Free Gaussian Splatting from a Single Event Camera". arXiv:2410.08107.*

53. [a, b, c, d, e, f] *Yu W, Feng C, Tang J, Yang J, Tang Z, Jia X, Yang Y, Yuan L, Tian Y (2024). "Evagaussians: Event stream assisted gaussian splatting from blurry images". arXiv:2405.20224.*

54. [a, b, c] *Yura T, Mirzaei A, Gilitschenski I (2024). "EventSplat: 3D Gaussian Splatting from Moving Event Cameras for Real-time Rendering". arXiv:2412.07293.*

55. ^ *Mildenhall B, Srinivasan PP, Tancik M, Barron JT, Ramamoorthi R, Ng R (2021). "Nerf: Representing scenes as neural radiance fields for view synthesis." Communications of the ACM.* ***65*** *(1): 99–106.*

56. ^ *Kerbl B, Kopanas G, Leimkühlerr T, Drettakis G (2023). "3D Gaussian Splatting for Real-Time Radiance Field Rendering". ACM Trans. Graph..* ***42*** *(4): 139–1.*

57. ^ *Rebecq H, Gehrig D, Scaramuzza D. "Esim: An open event camera simulator." In: Conference on Robot Learning. PMLR; 2018. p. 969-982.*

58. ^ *Gehrig D, Gehrig M, Hidalgo-Carri⊠ J, Scaramuzza D. "Video to events: Recycling video datasets for event cameras." In: CVPR, 2020. p. 3586--3595.*

59. ^ *Qu Q, Chen X, Chung YY, Shen Y (2024). "EvRepSL: Event-Stream Representation via Self-Supervised Learning for Event-Based Vision". IEEE Transactions on Image Processing. IEEE.*

60. ^ *Yu Z, Qu Q, Zhang Q, Zhang N, Chen X (2025). "LLM-EvRep: Learning an LLM-Compatible Event Representation Using a Self-Supervised Framework". arXiv preprint arXiv:2502.14273.*

61. ^ *Yu B, Ren J, Han J, Wang F, Liang J, Shi B. "EventPS: Real-time photometric stereo using an event camera." In: CVPR, 2024. p. 9602–9611.*

62. ^Chen H, Chen R, Qu Q, Wang Z, Liu T, Chen X, Chung YY (2024). *"Beyond Gaussians: Fast and High-Fidelity 3D Splatting with Linear Kernels". arXiv preprint arXiv:2411.12440.*

Declarations

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.