IBEC^3: IWE-Based Event Camera Calibration with Checkerboard

Taehun Ryu
Artificial Intelligence Graduate School, UNIST
taehunryu@unist.ac.kr

Abstract

Camera calibration is a prerequisite for accurate 3D vision tasks, involving the estimation of intrinsic, extrinsic, and distortion parameters. Traditional calibration assumes a static configuration between the camera and the calibration target at the time of image capture, requiring multiple views from diverse angles. However, event cameras necessitate motion to observe calibration patterns, making checkerboard edges difficult to detect. As a result, most existing methods rely on circle-grid patterns, which are known to introduce systematic bias under lens distortion. To address this limitation, we propose IBEC^3, a calibration method that leverages event warping to reproduce static views. By using a checkerboard, IBEC^3 avoids the inherent bias of circle patterns and remains compatible with classical calibration frameworks. To our knowledge, IBEC^3 is the first fully automatic event camera calibration method using raw events and a checkerboard, without blinking boards or neural network-based image reconstruction. The implementation of IBEC^3 is open-sourced at https://github.com/taehun-ryu/3D_vision_IBEC3

1. Introduction

Camera calibration plays a fundamental role in 3D computer vision applications, enabling accurate mapping between 3D world coordinates and 2D image projections. Classical calibration techniques estimate intrinsic, extrinsic, and distortion parameters by observing well-defined patterns, such as checkerboards or circle grids, from multiple static viewpoints.

Event cameras [5], with their asynchronous and high-temporal-resolution sensing, pose unique challenges to calibration. Unlike conventional frame-based cameras, event cameras require relative motion between the sensor and the calibration pattern to trigger events. As a result, traditional static-image-based calibration methods are not directly applicable.

In particular, checkerboard patterns, despite their widespread use in standard calibration, face difficulties in event-based settings. Due to the directional nature of event generation, camera motions restricted to single-axis translations (e.g., purely vertical or horizontal) often cause events to be triggered only along specific edge orientations [14]. This partial visibility hampers reliable checkerboard observation. To circumvent this, most existing event camera calibration methods [2, 7, 14] rely on circle-grid patterns, which generate omnidirectional edge features and are thus less sensitive to motion direction.

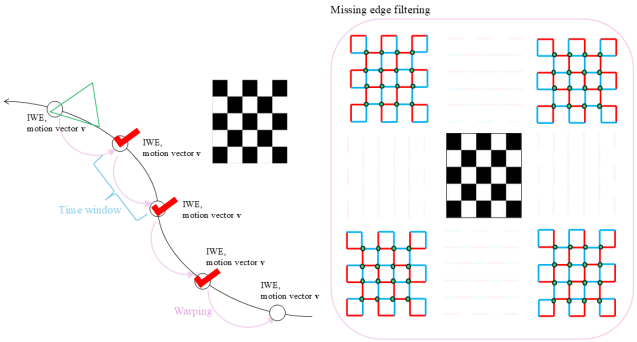

Figure 1. **Concept of our method.** We set a time window, estimate the motion vector, and obtain the IWE. Using motion vector, we can filter out unnecessary directions, such as including only vertical edges or only horizontal edges.

However, circle-grid patterns are known to introduce systematic biases under heavy lens distortion, especially in wide-angle or fisheye lenses [13, 17]. These biases persist even when averaged over multiple views and can degrade the accuracy of the calibration.

Several alternative approaches have been proposed to overcome the limitations of event-based calibration. Some use blinking light patterns to actively stimulate event generation [6, 10, 11], while others reconstruct intensity images from events using deep neural networks [8]. Despite their effectiveness, these methods require specialized hardware or heavy computation, which limits their practicality.

In this paper, we propose **IBEC^3**, a fully automatic calibration pipeline that uses raw event data and a checkerboard pattern without any active blinking or image reconstruc-

tion. Our method leverages event warping to generate Image of Warped Event (IWE), selects valid motion segments via gradient-based filtering, and refines checkerboard corner locations in a continuous space, as illustrated in Fig. 1. Finally, we estimate camera intrinsics using standard geometric optimization, maintaining compatibility with classical calibration frameworks.

The remainder of this paper is organized as follows: Section 2 reviews related work. Section 3 describes our proposed method in detail. Section 4 presents experimental results, and Section 5 concludes the paper.

2. Related Work

2.1. Event camera calibration

Event camera calibration has gained increasing attention as event-based sensors become more prevalent in robotics and computer vision. Unlike traditional frame-based calibration, event camera calibration must account for the asynchronous and motion-dependent nature of event data, requiring specialized strategies for pattern observation and parameter estimation.

One of the most intuitive strategies for extracting structured patterns from event data is to adapt conventional calibration targets specifically for the event camera modality. This approach is exemplified in [10, 11], where the authors designed an active LED-based grid pattern that blinks at a fixed frequency. By leveraging the brightness sensitivity of event cameras, the blinking pattern consistently triggers meaningful events, enabling the construction of clean, low-noise event accumulations suitable for pattern extraction.

Most existing event camera calibration methods rely on modified calibration targets based on circle-grid patterns. For instance, E-Calib [14] detects events corresponding to circular edges and uses ST-DBSCAN [1] clustering to associate them spatially. It then applies an extended RANSAC with weighted least squares (eRWLs) to fit circular features with sub-pixel accuracy. Huang et al. [7] optimize event trajectories using B-spline representations and detect cylindrical structures along those trajectories. Among recent methods, e-Kalibr [2] shows the strongest performance on circle-grid patterns. It classifies events into polarity-based clusters—termed running and chase clusters—and matches them by minimizing the Frobenius norm between cluster pairings. The method fits a time-varying 3D ellipse in the spatiotemporal domain (x, y, t), then samples 2D ellipses at selected time points to extract circular features.

In the neural network category, E2Calib [8] reconstructs grayscale images from raw events using the E2VID [12] network. These images are then fed into conventional RGB calibration tools such as Kalibr. Excluding blinking-pattern approaches, E2Calib is the only prior method capable of checkerboard-based calibration.

In contrast, our proposed method, **IBEC**^{3}, is the first checkerboard-based event camera calibration algorithm that does not rely on any neural networks or active blinking board. While the idea of constructing an Image of Warped Events (IWE) may resemble grayscale reconstruction methods like E2VID, the underlying principles are fundamentally different. Unlike E2VID, which reconstructs intensity images using learned priors from event sequences, IWE represents a physically grounded accumulation of events warped along a motion hypothesis. While E2VID aims to restore appearance, IWE aims to reveal structure through contrast consistency.

2.2. Image of warping event, IWE

The Image of Warped Events (IWE) is a fundamental representation in event-based vision, formed by warping events according to a motion hypothesis. IWE-based contrast maximization has been extensively studied in the context of optical flow estimation, where it serves as the objective for aligning events along motion trajectories. This approach has also been adapted for other tasks such as feature tracking, depth estimation, and camera motion estimation.

The concept of the Image of Warped Events (IWE) was first formalized by Gallego et al. [4], who introduced the contrast maximization (CM) framework for motion estimation using event cameras. By warping events along a motion hypothesis and maximizing the resulting image variance, CM enables alignment of asynchronous events to reveal underlying scene structure. Since then, contrast maximization has become a central technique in event-based optical flow estimation, with the IWE serving as both a byproduct and a diagnostic tool for evaluating flow quality. Stoffregen and Kleeman [18] further extended the CM framework by exploring various objective functions and optimization strategies. Among these, we adopt the R1 loss in our proposed framework, as detailed in Section 3.

More recently, Shiba et al. [16] proposed a CM-based optical flow estimation model that incorporates time-aware flow and multi-reference focus loss function. Their approach demonstrated strong performance gains while highlighting the potential of CM-based formulations for self-supervised learning. In this work, we adapt the CM framework—originally developed for optical flow—to the task of camera calibration. Our proposed **IBEC**^{3} method leverages IWE-based contrast maximization to detect checkerboard features without requiring appearance reconstruction or supervised training.

3. Methodology

In this section, we first define the notation (Section 3.1). The overall pipeline of **IBEC**^{3} is shown in Fig. 2. It consists of three main steps, and each step is covered in one section. Section 3.2 describes how we warp events to obtain IWE.

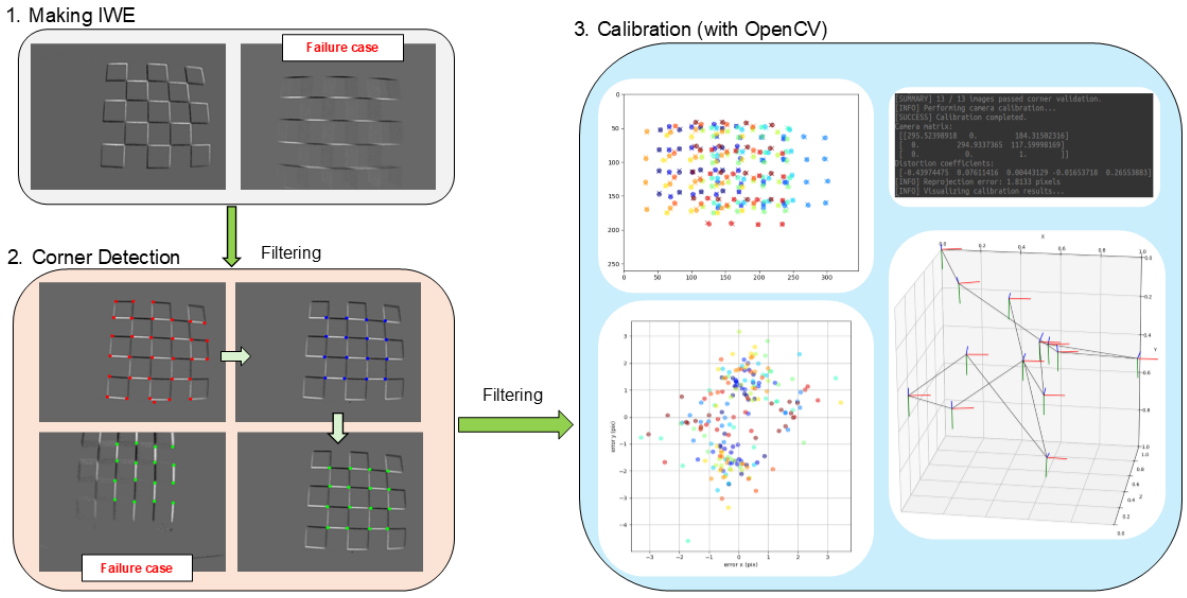

Figure 2. **High-level block diagram of IBEC^3.** The process consists of three main stages: (1) constructing and filtering IWEs from raw events, (2) detecting checkerboard corner candidates and snapping them to optimized grid points under geometric constraints, and (3) performing camera calibration using OpenCV.

Section 3.3 deals with the process of detecting checkerboard corners in the generated IWE. Section 3.4 deals with the final calibration work.

3.1. Notations

We denote an event as a tuple $e_n = (t_n, \mathbf{x}_n, p_n)$, where t_n is the timestamp, $\mathbf{x}_n = (x_n, y_n)^\top$ represents the pixel coordinates on the image plane, and $p_n \in \{-1, +1\}$ is the polarity indicating the sign of brightness change.

Let $\mathcal{E} = \{e_1, e_2, \ldots, e_N\}$ be a set of events captured within a predefined temporal window. A warping function $W(\cdot; \boldsymbol{\theta})$ maps each event position $\mathbf{x}_n$ at time t_n to a reference time t_{ref}, under a motion parameter $\boldsymbol{\theta}$:

$$\mathbf{x}'_n = W(\mathbf{x}_n, t_n; t_{ref}, \boldsymbol{\theta}) \tag{1}$$

We denote the resulting warped event image as an Image of Warped Events (IWE), denoted by $I(\mathbf{x}; \boldsymbol{\theta})$, constructed by accumulating warped events $\mathbf{x}'_n$ into a 2D histogram.

The camera intrinsic matrix is denoted by $\mathbf{K} \in \mathbb{R}^{3\times3}$ and has the following form:

$$\mathbf{K} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

where (f_x, f_y) are the focal lengths, and (c_x, c_y) are the coordinates of the principal point.

A 3D point $\mathbf{X} = (X, Y, Z)^\top$ is projected onto the image plane as:

$$\mathbf{x}_{\text{dist}} = \text{distort}\left(\pi(\mathbf{K}, \mathbf{R}, \mathbf{t}, \mathbf{X}); \mathbf{d}\right) \tag{3}$$

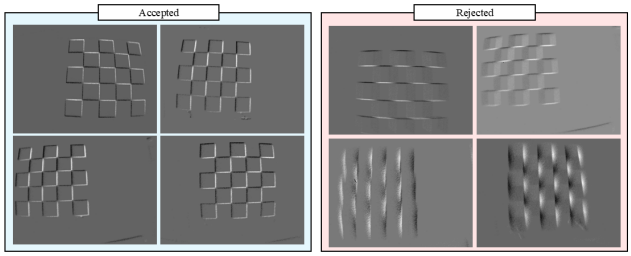

Figure 3. **IWE filtering results.** Left: accepted IWEs; Right: rejected IWEs. Filtering is based on the gradient energy metric.

where $\pi(\cdot)$ denotes the ideal pinhole projection with intrinsics $\mathbf{K}$ and extrinsics $(\mathbf{R}, \mathbf{t}) \in \text{SE}(3)$. The function distort$(\cdot)$ models radial and tangential distortion using the distortion parameters $\mathbf{d} = (k_1, k_2, p_1, p_2, k_3)$, following the OpenCV distortion model.

3.2. Making IWE

We divide the input event stream into static time windows of fixed duration and independently process each segment. Within each window, we construct an Image of Warped Events (IWE) by temporally aligning asynchronous events onto a reference timestamp using a linear motion model. Each event is defined as $e_n = (t_n, \mathbf{x}_n, p_n)$, and is warped to:

$$\mathbf{x}'_n = \mathbf{x}_n + (t_{\text{ref}} - t_n) \cdot \mathbf{v} \tag{4}$$

where $\mathbf{v}$ is the motion vector on the image plane.

The warped event locations $\mathbf{x}'_n$ are accumulated into a

2D histogram, with polarity p_n applied as the weighting factor. Using bilinear interpolation, we define the IWE as:

$$I(\mathbf{x};\boldsymbol{\theta}) = \sum_{n=1}^{N_e} p_n \cdot \delta(\mathbf{x} - \mathbf{x}'_n) \tag{5}$$

where $\delta(\cdot)$ denotes bilinear distribution around $\mathbf{x}'_n$ and $\boldsymbol{\theta} = \{\mathbf{v}(\mathbf{x})\}_{\mathbf{x}\in\mathcal{E}}$.

To optimize the IWE construction, we adopt the R1 contrast-based objective function from [18]:

$$\mathrm{R1}(I) = \sum_{i,j} I(i,j)^2 + \sum_{i,j} \epsilon^{I(i,j)} \tag{6}$$

where ϵ is a small positive constant used for contrast regularization and (i, j) is pixel location.

To evaluate IWE quality during filtering, we compute the Gradient Energy (GE) metric. Let $G(x, y)$ be the IWE convolved with a Gaussian kernel $\mathcal{G}_\sigma$:

$$G(x,y) = \mathcal{G}_\sigma * I(x,y) \tag{7}$$

The final GE score is defined as:

$$\text{GradientEnergy}(I) = \sum_{x,y} \sqrt{G_x(x,y)^2 + G_y(x,y)^2} \tag{8}$$

where $G_x = \frac{\partial G}{\partial x}$, $\quad G_y = \frac{\partial G}{\partial y}$.

This metric captures the spatial contrast of the warped image and serves as our filtering criterion. Specifically, we discard IWEs whose gradient energy falls below a predefined threshold τ_{GE}, as they tend to result from degenerate motion directions or weak structure. Additionally, motion vectors with magnitude $r < 50$ or directions within $\pm 10^\circ$ of pure horizontal or vertical axes are excluded.

Figure 3 illustrates the effectiveness of this filtering process. Accepted IWEs (left) exhibit well-aligned and high-contrast checkerboard edges, while rejected IWEs (right) suffer from directional blur or poor event alignment. This pre-filtering step is critical for robust corner detection in the following stage.

3.3. Corner detection

We extract checkerboard corners from each accepted IWE using a hybrid approach that combines sparse feature detection and grid-constrained optimization.

We first apply the Shi–Tomasi corner detector [15] to the normalized IWE to extract a set of candidate points $\{\mathbf{c}_i\}$. We use OpenCV's `goodFeaturesToTrack` with a relaxed threshold to allow for noisy but dense detections. The maximum number of corners is set to $2 \times (W \times H)$, where W and H denote the checkerboard dimensions.

Next, we initialize a coarse grid of shape (H, W) by distributing points uniformly in the spatial range of the detected candidates. This grid is treated as a continuous variable and optimized using a differentiable loss that enforces geometric constraints:

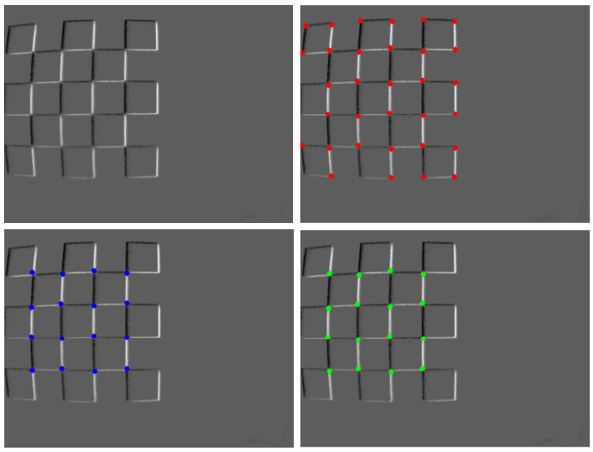

Figure 4. **Corner detection pipeline.** Top-left: input IWE; top-right: initial corner candidates detected by GFTT; bottom-left: optimized continuous grid points; bottom-right: final snapped corners used for calibration.

- **Spacing ratio loss**: encourages uniform distance between adjacent corners.
- **Orthogonality loss**: enforces right angles between row and column vectors.
- **Candidate proximity loss**: minimizes distance to the nearest candidate point.

We optimize the grid via gradient descent with the Adam optimizer.

The resulting grid is then snapped to the nearest unused corner candidates using a KD-tree structure, ensuring one-to-one matching. The resulting grid is then snapped to the nearest unused corner candidates using a KD-tree structure, ensuring one-to-one matching. We validate the final snapped grid by checking three structural criteria:

- The standard deviation of row-wise and column-wise spacing must be below a fraction τ_s of their respective mean values (default $\tau_s = 0.1$).
- The average cosine of angles between row and column vectors (measuring deviation from orthogonality) must be smaller than a threshold τ_o (default $\tau_o = 0.2$).

Only grids satisfying both spacing and orthogonality thresholds are accepted as valid detections. The overall pipeline of our corner detection method is illustrated in Fig. 4, including the initial GFTT detection, grid optimization, and final snapping.

3.4. Calibration

Given the snapped grid of checkerboard corners obtained from each accepted IWE, we perform standard pinhole camera calibration using OpenCV's `calibrateCamera` function.

Let $\{\mathbf{c}_k^{(i)}\}$ denote the set of 2D corner detections in the i-th image, ordered in row-major layout from the snapped

Table 1. Estimated intrinsic parameters and distortion coefficients.

Method	fx	fy	cx	cy	k1	k2	k3	p1	p2
E2Calib [8] (event-to-video)	280.57	280.71	172.30	132.54	-0.378	0.158	—	-0.000	-0.002
E-Calib [14] (circle-grid)	284.12	283.78	166.98	128.03	-0.383	0.194	0.002	-0.001	-0.036
Ours (A)	270.40	266.59	144.52	115.53	-0.704	2.617	0.007	0.015	-4.551
Ours (H)	**295.52**	**294.93**	184.31	117.59	-0.439	0.076	0.004	-0.016	0.265
GT (frame-based)	**291.95**	**292.73**	**163.41**	**130.17**	**-0.427**	**0.355**	**0.000**	**0.003**	**-0.233**

ξ: reprojection error; A: auto filtering in corner detection; H: human filtering in corner detection.

grid. Let $\{\mathbf{X}_k\}$ be the corresponding 3D coordinates of the checkerboard corners, generated by a known square size and fixed (W, H) layout. We jointly estimate the camera intrinsics $\mathbf{K}$, distortion parameters $\mathbf{d}$, and per-image extrinsic parameters $\{\mathbf{R}^{(i)}, \mathbf{t}^{(i)}\}$ by minimizing the reprojection error:

$$\min_{\mathbf{K},\mathbf{d},\{\mathbf{R}^{(i)},\mathbf{t}^{(i)}\}} \sum_i \sum_k \left\| \mathbf{c}_k^{(i)} - \pi(\mathbf{K}, \mathbf{d}, \mathbf{R}^{(i)}, \mathbf{t}^{(i)}, \mathbf{X}_k) \right\|^2 \tag{9}$$

All parameters are initialized internally by OpenCV and optimized via Levenberg–Marquardt.

4. Experiments

4.1. Setup

To evaluate the performance of IBEC^3, we conduct experiments using real-world data captured by a DAVIS346 camera. The DAVIS346 is an event-based camera with a resolution of 346×260, providing both asynchronous events and grayscale images.

Since E2Calib [8] is the only existing method that supports checkerboard-based calibration, we compare our method with other publicly available baselines that use circle-grid targets [2, 14]. Among them, we attempted to evaluate e-Kalibr [2], the most recent and accurate circle-grid-based calibration method. However, e-Kalibr consistently failed to track the circle-grid pattern in our dataset and was therefore excluded from comparison.

We use the frame-based calibration result obtained from grayscale images as ground truth (GT), as DAVIS346 natively provides synchronized intensity images. We follow the original E2Calib setup, where grayscale images are reconstructed using E2VID [12] and passed to a conventional calibration backend such as Kalibr [3] (or OpenCV). In our experiment, we adopt Kalibr for consistency.

4.2. Quantitative result

Table 1 reports the estimated intrinsic parameters and distortion coefficients for each calibration method. We evaluate two variants of our pipeline: **Ours (A)** uses automatic filtering of checkerboard corners based on structural validation, while **Ours (H)** involves human supervision to manually accept or reject each detection result.

Values highlighted in red indicate significant deviations from the ground truth (GT), particularly in the principal point estimates (c_x, c_y). Bolded values indicate the closest match to the GT for each parameter, most notably observed in the focal lengths of **Ours (H)**.

In terms of focal length (f_x, f_y), both versions of our method produced values comparable to GT, with **Ours (H)** achieving the closest match. This suggests that IBEC^3 successfully recovers the global scale of the camera model.

However, the principal point estimates show noticeable errors in both **Ours (A)** and **Ours (H)**. This is attributed to the spatial distribution of the accepted IWEs: since both automatic and manual filtering tend to favor samples with high contrast and visible grid structure, the selected IWEs are often biased toward central regions of the image plane. As a result, the calibration lacks peripheral image coverage, leading to biased estimates of the image center. We will revisit this limitation in Section 5.2.

The distortion coefficients $(k_1$–$p_2)$ are generally consistent in sign and magnitude across all methods. Although **Ours (A)** shows slightly larger deviations, the results remain within a reasonable range.

Overall, IBEC^3 achieves competitive intrinsic calibration performance compared to prior art, without requiring LED boards or neural image reconstruction. These results also highlight the importance of spatially diverse event samples for accurate principal point estimation.

4.3. Qualitative result

To further assess the quality of the estimated parameters, we visualize the calibration results from **Ours (H)** in Figure 5.

The middle column illustrates the reprojection of 3D checkerboard points onto the image plane using the optimized intrinsic and extrinsic parameters. The green dots correspond to the detected 2D corner positions from the IWE, while the red dots indicate the reprojected locations of

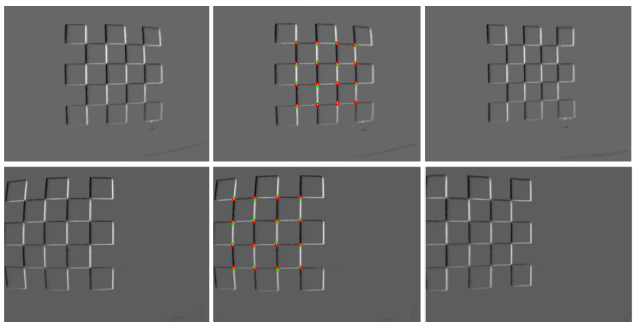

Figure 5. **Calibration visualization.** Left: original IWE; Middle: reprojection of estimated 3D points using optimized parameters (green: observed corners, red: reprojected points); Right: undistorted IWE using the estimated distortion coefficients.

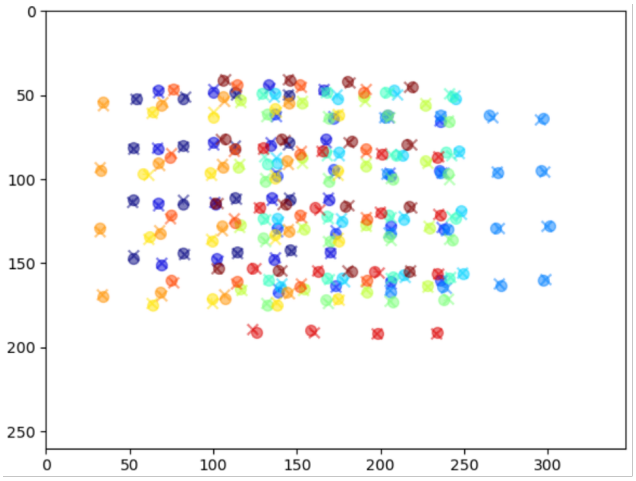

Figure 6. **Corner distribution across valid IWEs.** Each color indicates a different IWE; ∘ marks observed corners, × marks reprojected points. Most corners lie near the image center, causing spatial bias and limiting accurate principal point estimation.

the corresponding 3D points. The close agreement between these two sets of points across all valid IWEs demonstrates that the estimated camera model provides a geometrically consistent explanation of the observations. In particular, the residuals between reprojected and observed points are small and uniformly distributed, which suggests that both the intrinsics and the estimated board poses are well-optimized.

The right column shows the undistorted IWEs generated using the estimated distortion coefficients. Compared to the original warped images (left column), the undistorted images exhibit straight and regularly spaced checkerboard edges, especially along lines that were previously curved due to radial or tangential distortion. This visual improvement indicates that the distortion parameters successfully compensate for lens nonlinearity, further supporting the validity of the calibration outcome.

Taken together, these qualitative results confirm that IBEC^3 can produce accurate and interpretable calibration parameters even under the minimalistic setup of event-only input and static checkerboard motion.

5. Conclusion

5.1. Summary

We presented IBEC^3, a calibration method for event cameras using static checkerboards and contrast-based event warping. Unlike existing approaches that require blinking light patterns or deep neural networks, our method leverages standard printed patterns and low-level motion compensation to construct Image of Warped Events (IWE). We proposed a corner detection pipeline based on grid-constrained optimization and geometric validation, enabling the use of conventional calibration frameworks such as OpenCV.

Through experiments on real data from a DAVIS346 camera, we showed that IBEC^3 achieves comparable intrinsic accuracy to state-of-the-art baselines. Qualitative analysis confirmed that our reprojection and distortion correction are geometrically consistent and interpretable. Notably, our approach remains fully self-contained and does not rely on any grayscale imagery, learned features, or active stimuli.

By reintroducing the use of checkerboards into event-based calibration pipelines, IBEC^3 offers a simple yet effective solution that bridges the gap between event-only input and traditional camera calibration techniques.

5.2. Limitations

Despite the promising results of IBEC^3, several limitations remain that must be addressed in future work.

First, the proposed pipeline assumes a linear motion model to construct the Image of Warped Events (IWE). While effective for smooth trajectories, this approximation breaks down under fast or non-linear camera motion, leading to blurred IWEs with degraded contrast. As a result, both IWE quality and corner detection stability suffer in such cases.

Second, the checkerboard must remain clearly visible throughout the motion. When the pattern is significantly tilted or viewed at an oblique angle, edges fail to trigger consistent events, causing checkerboard corners to disappear from the IWE. This makes both detection and optimization unreliable in high-angle views.

Third, a large portion of the acquired data is discarded due to the strict IWE filtering pipeline. Events with insufficient motion magnitude or near-horizontal/vertical directions are excluded, and low-contrast IWEs are removed based on the GE metric. As shown in Figure 6, this results in most valid IWEs being concentrated near the image center. Consequently, spatial coverage is biased, and intrinsic parameter estimation—especially for the principal point

(c_x, c_y)—becomes unstable.

These limitations reveal clear boundaries of applicability for IBEC^3 and point to several research directions that can improve robustness, scalability, and generalization.

5.3. Future works

There are several promising directions for improving and extending IBEC^3. First, the current method assumes linear image-plane motion during event warping, which may fail under rapid or nonlinear motion. Adopting a time-varying motion model or allowing spatially-varying motion—e.g., applying different velocity vectors to local image patches like Shiba et al. [16]—could improve IWE quality and enhance the robustness of corner detection.

Second, the filtering pipeline tends to favor IWEs with strong gradients, which are often spatially concentrated near the center of the image. To mitigate this bias, future work could incorporate spatial coverage-aware filtering strategies or multi-view selection mechanisms to ensure more uniform image-plane coverage.

Third, the corner detection process can be further improved in accuracy and stability. Incorporating temporal consistency across frames, or introducing learned refinement modules such as graph neural networks, may help resolve ambiguities and reduce reliance on individual frame quality.

Finally, although IBEC^3 is demonstrated with checkerboard patterns, the framework could be extended to support alternative fiducials such as AprilTags or other high-contrast markers that produce reliable edge structures in event streams. Recent work has shown that structured patterns like QR codes can even be reconstructed from event data alone [9], suggesting that using only event data with fiducial patterns. This indicates that accurate calibration using only event data and simple fiducial markers—without relying on grayscale frames—could be a viable and effective approach in future event-based vision systems.

References

[1] Derya Birant and Alp Kut. ST-DBSCAN: An algorithm for clustering spatial–temporal data. *Data & Knowledge Engineering*, 2007. 2

[2] Shuolong Chen, Xingxing Li, Liu Yuan, and Ziao Liu. e-Kalibr: Dynamic intrinsic calibration for event cameras from first principles of events. *IEEE Robotics and Automation Letters*, 2025. 1, 2, 5

[3] Paul Furgale, Joern Rehder, and Roland Siegwart. Unified temporal and spatial calibration for multi-sensor systems. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2013. 5

[4] Guillermo Gallego, Henri Rebecq, and Davide Scaramuzza. A unifying contrast maximization framework for event cameras, with applications to motion, depth, and optical flow estimation. In *IEEE Conf. Comput. Vis. Pattern Recog.*, 2018. 2

[5] Guillermo Gallego, Tobi Delbruck, Garrick Orchard, Chiara Bartolozzi, Brian Taba, Andrea Censi, Stefan Leutenegger, Andrew Davison, Joerg Conradt, Kostas Daniilidis, and Davide Scaramuzza. Event-based vision: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2022. 1

[6] Thomas Gossard, Andreas Ziegler, Levin Kolmar, Jonas Tebbe, and Andreas Zell. ewand: A calibration framework for wide baseline frame-based and event-based camera systems. In *IEEE International Conference on Robotics and Automation*, 2024. 1

[7] Kun Huang, Yifu Wang, and Laurent Kneip. Dynamic event camera calibration. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2021. 1, 2

[8] Manasi Muglikar, Mathias Gehrig, Daniel Gehrig, and Davide Scaramuzza. How to calibrate your event camera. In *IEEE Conf. Comput. Vis. Pattern Recog. Worksh.*, 2021. 1, 2, 5

[9] Jun Nagata, Yusuke Sekikawa, Kosuke Hara, Teppei Suzuki, and Aoki Yoshimitsu. Qr-code reconstruction from event data via optimization in code subspace. In *IEEE Winter Conference on Applications of Computer Vision*, 2020. 7

[10] Garrick Orchard. Dvscalibration. https://github.com/gorchard/DVScalibration.git, 2025. accessed: January 3, 2025. 1, 2

[11] R. (Robotics and P. Group). rpg_dvs_ros. https://github.com/uzh-rpg/rpg_dvs_ros.git, 2025. accessed: January 3, 2025. 1, 2

[12] Henri Rebecq, René Ranftl, Vladlen Koltun, and Davide Scaramuzza. High speed and high dynamic range video with an event camera. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2019. 2, 5

[13] Victoria Rudakova and Pascal Monasse. Camera matrix calibration using circular control points and separate correction of the geometric distortion field. In *Proc. Canadian Conference on Computer and Robot Vision*, 2014. 1

[14] Mohammed Salah, Abdulla Ayyad, Muhammad Humais, Daniel Gehrig, Abdelqader Abusafieh, Lakmal Seneviratne, Davide Scaramuzza, and Yahya Zweiri. E-Calib: A fast, robust, and accurate calibration toolbox for event cameras. *IEEE Trans. Image Process.*, 2024. 1, 2, 5

[15] Jianbo Shi and Tomasi. Good features to track. In *IEEE Conf. Comput. Vis. Pattern Recog. Worksh.*, 1994. 4

[16] Shintaro Shiba, Yannick Klose, Yoshimitsu Aoki, and Guillermo Gallego. Secrets of event-based optical flow, depth, and ego-motion by contrast maximization. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2024. 2, 7

[17] Chaehyeon Song, Jaeho Shin, Myung-Hwan Jeon, Jongwoo Lim, and Ayoung Kim. Unbiased estimator for distorted conics in camera calibration. In *IEEE Conf. Comput. Vis. Pattern Recog.*, 2024. 1

[18] Stoffregen, Timo, Kleeman, and Lindsay. Event cameras, contrast maximization and reward functions: An analysis. In *IEEE Conf. Comput. Vis. Pattern Recog.*, 2019. 2, 4