J. Innovative Eng. Nat. Sci., vol. 5, no. 2, pp. 783-795, 2025.
http://doi.org/10.61112/jiens.1626247

Original Paper

Journal of Innovative Engineering and Natural Science
(Yenilikçi Mühendislik ve Doğa Bilimleri Dergisi)
https://dergipark.org.tr/en/pub/jiens

Image reconstruction of static scenes from a static event camera using an LCD projector

Beyza Eraslan[a,*], Gokhan Koray Gultekin[a]

[a]*Ankara Yildirim Beyazit University, Department of Electrical and Electronics Engineering, Ankara, 06010, Türkiye.*

ARTICLE INFO

Article history:
Received
Received in revised form
Accepted

Available online

Keywords:
Event camera
Image reconstruction
E2VID
FireNet
Static event data

ABSTRACT

Event cameras are promising sensors that show many advantages over frame-based cameras. Unlike conventional cameras, whose pixels share a common exposure time, event-based cameras represent a novel bio-inspired technology capable of capturing scenes with a high dynamic range and without motion blur. Due to their working principle, an event is generated when a pixel's brightness changes. Therefore, no event data is generated in a scenario where there is no relative motion between the event camera and the scene. However, in this study, we present a new method to enable event generation with a static event camera on a scene with static objects, aiming to eliminate the requirement of relative motion between event camera and the scene. By projecting custom designed grayscale pattern sequences onto static scenes, we successfully triggered a controlled event generation without requiring camera or object motion. Instead of a direct black-to-white transition, we used a sequence of contrast compatible grayscale projection pattern to regulate event rates and prevent bandwidth overload. To prevent event loss over time, we equalized all timestamps to the first timestamp. Since events in event cameras gradually lose their impact and reset over time, this adjustment prevents the decay of event information and ensures a continuous and stable event generation. Despite the absence of motion, we achieved reasonable results in image quality metrics such as MSE, LPIPS, SSIM. In this way, we aim to expand the usage areas of event cameras and make significant progress in data collection processes, especially for static camera and scenes.

I. INTRODUCTION

Event cameras are powerful new sensors that have microseconds level temporal resolution and high dynamic range without motion blur. Their strength is in asynchronously detecting brightness changes rather than directly capturing brightness images [1]. Event cameras have created a major change in the field of computer vision with their features, such as asynchronous detection of events, high dynamic range (HDR). These features allow event cameras to detect changes much faster and more accurately than traditional cameras and therefore are considered a revolutionary development in computer vision technologies [2-5]. When events are caused by the apparent motion of objects, event-based cameras sample visual information based on scene dynamics and are therefore more naturally suited than traditional cameras to capture motion, especially at high speeds where traditional cameras suffer from motion blur [6]. Over the last few years, their exceptional features have led to exciting vision applications with high speed and low latency. However, these sensors are still rare and expensive to obtain, which slows down the progress of the research community. There is a huge demand for inexpensive, high-quality real/synthetic, labeled events dataset for algorithm prototyping, deep learning, and algorithm benchmarking to address these issues [7]. In particular, low-latency perception, HDR scenes, and motion blur remain a profound problem for perception systems based on standard frame-based cameras. Instead of capturing intensity images synchronously, i.e., at regular time intervals, event cameras measure intensity changes asynchronously in the form

*Corresponding author. Tel: +90-312-906-1000; e-mail: beyza.eraslan@aybu.edu.tr

of an event stream. [6]. The events are produced based on the logarithmic pixel intensity function $L_{xy}(t)$, which represents the instantaneous light intensity (brightness) at a specific pixel (x, y) at a given time t. However, $L_{xy}(t)$ is not directly produced by the camera. Instead, the camera outputs a sequence of events in the format (x, y, t, p). Here, (x, y) denotes the image coordinates, t indicates the time of the event, and p represents the polarity ($p = \pm 1$), which indicates whether the intensity change at that pixel is a decrease or an increase [8].

An event ($e_k \doteq \boldsymbol{x}_k, t_k, p_k$) is triggered when the brightness increment ($\Delta L(\boldsymbol{x}_k, t_k)$) at a pixel ($\boldsymbol{x}_k \doteq (x_k, y_k)^{\mathrm{T}}$) reaches a predefined contrast threshold ($\pm C$) [9]. The brightness increment is given in the Equation 1.

$$\Delta L(\boldsymbol{x}_k, t_k) \doteq L(\boldsymbol{x}_k, t_k) - L(\boldsymbol{x}_k, t_k - \Delta t_k) \tag{1}$$

When the contrast threshold is reached, the situation becomes as in Equation 2.

$$\Delta L(\boldsymbol{x}_k, t_k) = p_k . C \tag{2}$$

Conventional cameras measure absolute brightness information, whereas event cameras provide information about brightness changes or their temporal derivatives [9]. Equation 3 indicates that events offer information about the temporal derivative of brightness changes at a pixel.

$$\frac{\partial L}{\partial t}\ (\boldsymbol{x}_k, t_k) \approx \frac{p_k C}{\Delta t_k}. \tag{3}$$

Due to their exceptional nature, event cameras promise to unlock robust and high-speed perception in situations currently out of reach of standard cameras [7]. However, a significant drawback of event cameras is that it is not possible to obtain information from a scene where the camera and the scene is static. The main aim of this study is to propose a technique to expand the area of use of event cameras particularly in scenarios where they cannot generate data (e.g. in a static scene (i.e. no motion) with a static event camera). Actually, the event cameras are not intended to be used in purely static scenarios. However, there are specific use cases where capturing or reconstructing a static view of the scene is necessary and useful, even in the absence of activity. For example, applications such as background modelling, camera calibration, visual localization, mapping, or object recognition require a conventional image or intensity frame. In dynamic scenarios, reconstruction can be performed from accumulated events. But when the camera and scene are static, no events are generated, which creates a challenge for reconstructing the background or ambient environment. In such cases, our proposed method can be utilized to trigger events that can be used for scene image reconstruction.

This study presents a novel method that eliminates the requirement of a relative motion between the camera and the scene so that it enables event generation in such a case. Our method aims to allow effective data collection in

static environments with an event camera, even in the absence of motion in the scene and the camera. The main flow diagram of the study is given in Figure 1.

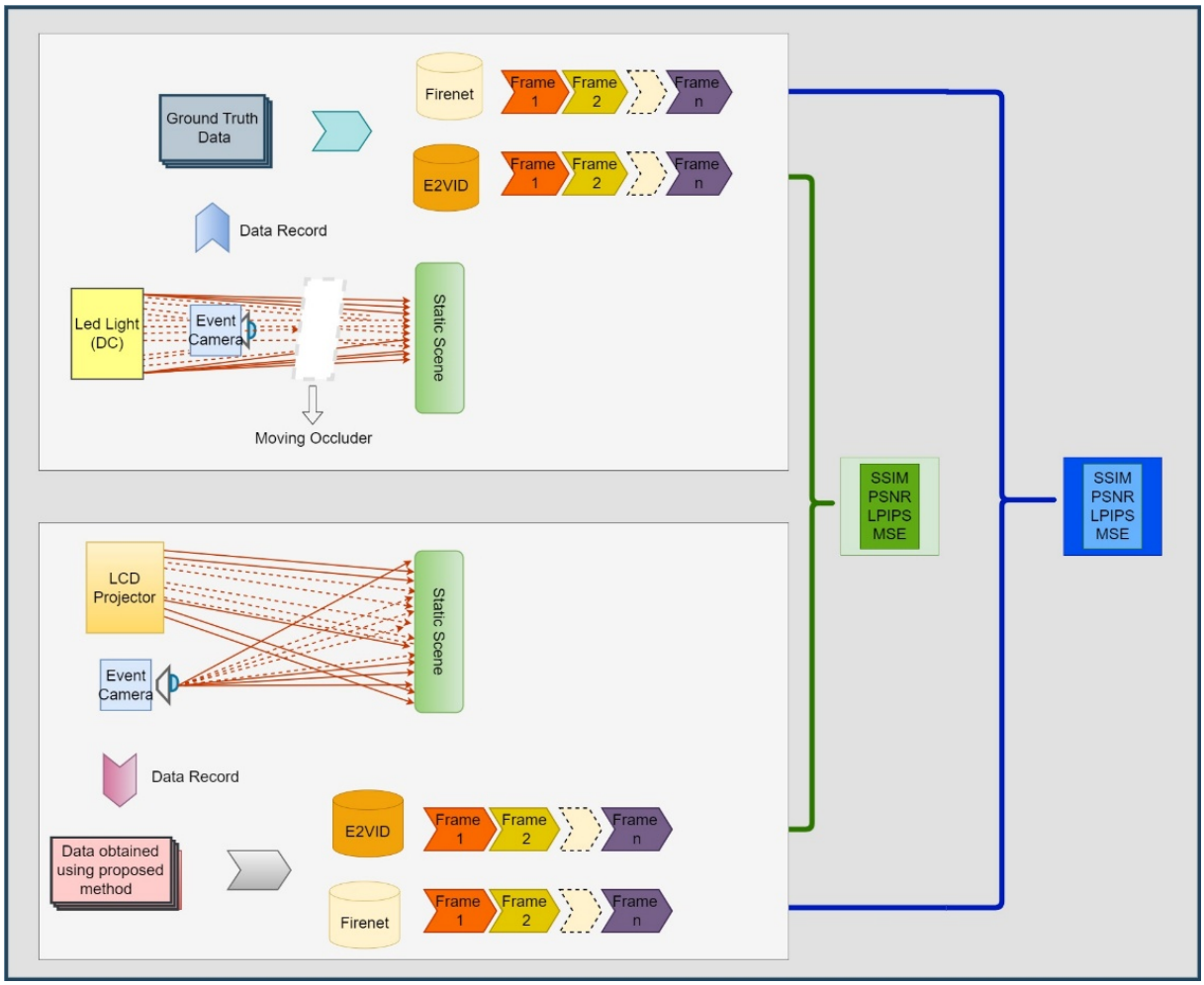

Figure 1. Main flow diagram of LCD projector -based event generation and model evaluation

II. METHOD

2.1. Image Reconstruction of a Scene from Event Data

Event cameras differ from traditional video cameras by producing data only when there is apparent motion or brightness change in the scene. Therefore, they do not directly output an image of the scene. However, there are methods in the literature such as E2VID and FireNet [1, 10] in order to reconstruct the image of the scene from the event data. E2VID transforms the event stream into a sequence of event tensors organized as a 3D spatiotemporal voxel grid, composed of fixed groups of events. This sequence is then processed by an iterative UNet model to reconstruct image frames. The E2VID architecture integrates elements of both UNet and LSTM, forming a fully convolutional and recurrent design. While UNet focuses on processing spatial details, LSTM captures and interprets temporal context [10-13].

FireNet, a fully convolutional network with recursive connections, is designed for convenience. Compared to E2VID, FireNet's compact and fast neural network architecture is a proof of this. Tailored for image reconstruction

from event data, it uses significantly fewer parameters, resulting in lower computational overhead. The use of recursive connections further enhances its efficiency, making FireNet a particularly advantageous choice for computational demands [14]. The features of the utilized E2VID and FireNet models are presented in the Table 1 [1].

Table 1. Features of E2VID and FireNet models

Feature	E2VID	FireNet
Downsampling	Yes	No
Memory (Mb)	43	0.16
Recurrent Units	LSTM	GRU
Max. Kernel Size	5x5	3x3
No. Parameters (k)	10700	38

2.2. Proposed Framework for Event Generation from a Static Scene and Camera

In this study, we propose a new framework for event generation from a static scene using an Liquid Crystal Display (LCD) projector to systematically control event production. The primary goal of this framework is to efficiently regulate the event generation rate in the proposed LCD projector based event data collection process. To collect the event data, a DVXplorer event camera (iniVation AG, Switzerland) was utilized. The camera features a spatial resolution of 640 × 480 pixels, a temporal resolution of 200 µs, and a high dynamic range up to 110 dB. It supports event output rates of up to 165 million events per second, with a typical latency of less than 1 ms. These specifications make it suitable for capturing subtle brightness changes in static scenes with high precision. The device also includes an integrated 6-axis IMU and supports synchronization through USB 3.0 and external trigger connectors. During event data collection, the built-in noise filtering feature provided by the DV (iniVation) software was enabled to reduce burst noise at the sensor level. After recording, the event data stored in .aedat4 files were exported into a text-based format containing timestamp, spatial coordinate, and polarity information using Python with the dv library. All data processing steps, including timestamp equalization and event stream reconstruction, were performed using independently developed Python scripts. A DC illumination source was used to illuminate a static scene, providing a consistent baseline illumination. To regulate the event rate and prevent excessive event generation, a sequence of pattern gradually transitioning from dark to light were projected onto the scene using an LCD projector. The Wanbo T2 Max projector, employed in this study, features a Full HD 1080p resolution, 450 ANSI lumens brightness, a 2000:1 contrast ratio, and automatic keystone correction. Its optical system utilizes a full glass coated lens with an LED light source, ensuring high image clarity and stable illumination during the pattern projection process. The projector's compact design and precise focusing capability provided effective and controlled event generation from the static scene. Two different pattern sequences were designed to regulate event generation in a controlled manner. In the first pattern sequence, the entire slide background was filled starting with pure black (RGB: 0,0,0) and incrementally transitioned towards white by increasing the RGB values by 25 units at each step, culminating in full white (RGB: 255,255,255) at the final frame. In the second sequence, a small gray rectangle was gradually expanded across a black background in successive slides, increasing the area of brightness while maintaining a slow and progressive transition. This approach was chosen to prevent a sudden surge in event generation, which could exceed the camera’s bandwidth and cause system bottlenecks. By implementing a controlled contrast transition, the pattern design was optimized to ensure gradual and continuous stimulation of events without exceeding the camera’s bandwidth, thereby achieving a balanced,

efficient, and stable event data collection process. To minimize unwanted noise and improve data quality, a noise filter [15] was applied within the processing software during event data collection before recording, which corresponds to the Background Activity Noise Filter module of the DV software. This filter evaluates the temporal and spatial coherence of events within a local pixel neighborhood. An event is retained if it is "supported" by at least one other event within its 8-pixel neighborhood and within a 2-millisecond time window. In our configuration, the support threshold was set between 1 and 8, polarity checking was disabled, and default settings were used without further manual adjustment. This filtering, applied in real time, helped removing isolated noise events while maintaining signal integrity for the reconstruction phase. It should be noted that different noise filtering methods and parameters may affect the results slightly. Dynamic parameter optimization for noise filtering in event-based image reconstruction can be performed by adjusting filter parameters (e.g. threshold values, time windows, etc.) iteratively to maximize reconstruction quality. However, since the primary focus of the work is on static scene event generation and reconstruction rather than noise filtering algorithm comparison, alternative denoising methods and the effect of noise filter parameter tuning were not investigated within the scope of this study. Nevertheless, future studies may explore the effects of applying advanced denoising algorithms such as BM3D filtering [16], non-local means denoising [17], and wavelet-based thresholding [18] on reconstruction quality. Moreover, inspired by recent works such as Noise2Image [19], where photon noise characteristics were leveraged for static scene recovery, more systematic analyses of noise filtering strategies could further enhance event-based imaging.

Furthermore, to reconstruct temporally distributed event occurrences as if they had occurred simultaneously, all timestamps were equalized to the first timestamp. Since event cameras gradually reset event data over time, this adjustment ensured continuous event retention and prevented information decay. Specifically, for each recorded event represented as (x, y, t, p), where x and y denote spatial coordinates, t denotes the timestamp, and p denotes the polarity, the timestamp t was replaced with t_0, the timestamp of the first recorded event. This adjustment effectively removed temporal variations across the event set, ensuring that the data appeared to originate from a single synchronized capture. Since event cameras gradually reset event information over time, timestamp equalization also prevented the decay of event relevance and ensured continuous event retention. This step was critical to maintaining spatial consistency and enhancing the quality of reconstructed images in the subsequent processing stages.

To evaluate the effectiveness of our method with a nonstatic case we collected data by triggering event generation by moving an occluder object (e.g. a white board) in front of the camera as illustrated in Figure 1. This manual movement allowed the camera to capture event data based on the apparent motion of the scene. The event data collected using this method serves as the ground truth for evaluating the effectiveness of our framework.

For a quantitative comparison, we used E2VID and FireNet to reconstruct images from both the collected ground truth dataset and the LCD projector-based event data. We then quantitatively evaluated the reconstruction quality by calculating SSIM, PSNR, LPIPS, and MSE metrics between the reconstructed frames, using the ground truth as a reference. The experimental setup is illustrated in Figure 2. The visual representations of the pattern sequences we used are provided in Figure 3.

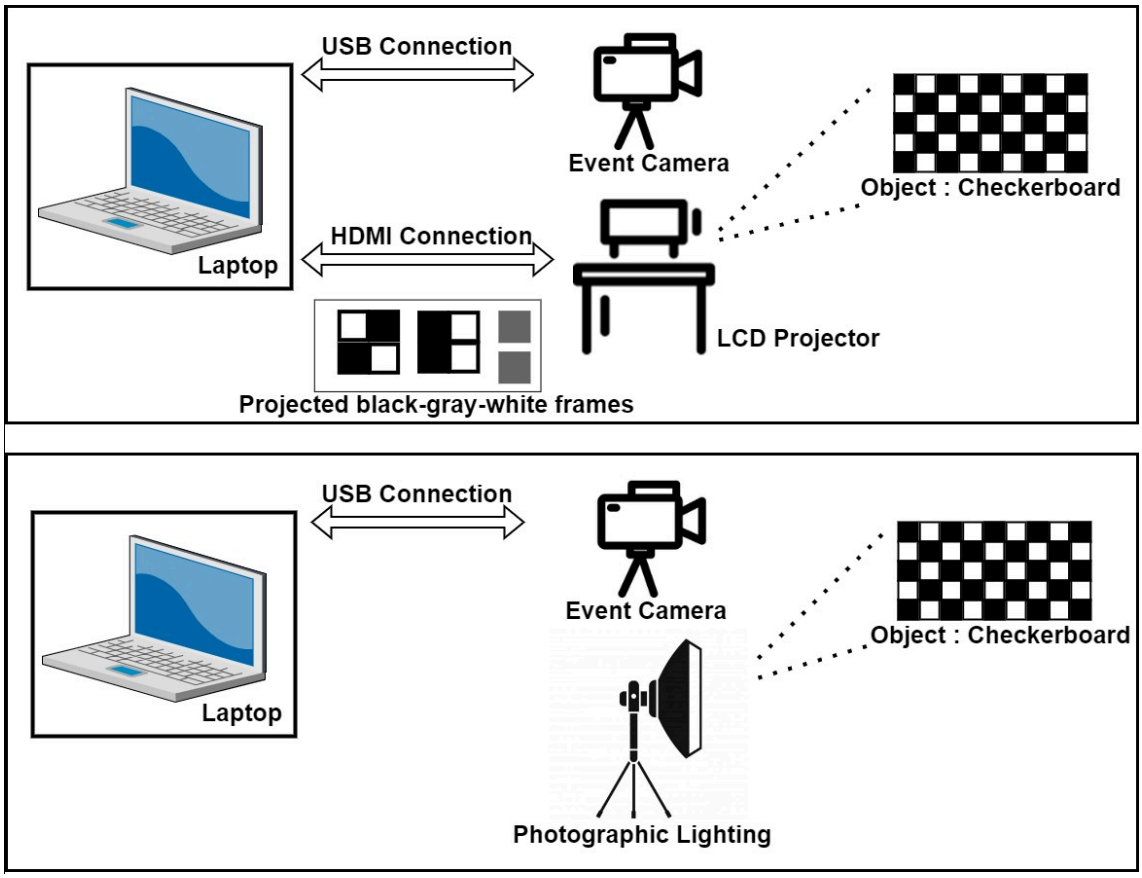

Figure 2. Experimental setup for using pattern sequences (upper) and ground truth (lower)

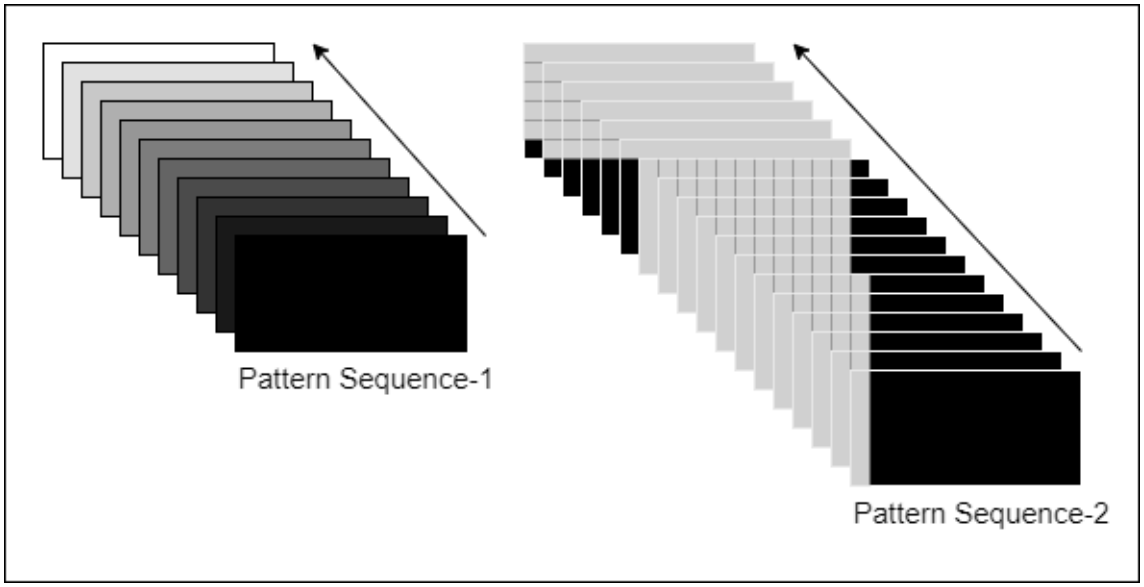

Figure 3. Images of the pattern sequences used

2.3. Performance Metrics

In our experimental evaluations, we employed widely used performance metrics from the literature to assess and compare the effectiveness of our approach. The following subsections provide details on these metrics and their relevance to our study.

2.3.1. Structural similarity index measure (SSIM)

The Structural Similarity Index (SSIM) is a widely recognized metric for assessing the similarity between two images. It is regarded as closely aligned with the human visual system's (HVS) perception of quality. Unlike conventional methods based on error accumulation, SSIM evaluates image distortion with a comprehensive approach, considering three key components: correlation loss, luminance variation, and contrast variation. The SSIM value is always between 0 and 1. 0 means no similarity between the images, and 1 means the images are completely identical [20, 21]. Formula of SSIM is given in Equation 4 [21].

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x{}^2 + \mu_y{}^2 + c_1)\,(\sigma_x{}^2 + \sigma_y{}^2 + c_2)} \tag{4}$$

2.3.2. Mean squared error (MSE)

MSE is the most common estimator of image quality measurement metrics. MSE is a nonnegative value, where values closer to zero represents a better quality. A higher MSE means more error or difference. It is widely used to provide dispersion between true and predicted values, especially in fields such as image processing, machine operations, and signal processing [22, 23]. The formula for MSE is given in Equation 5.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{5}$$

2.3.3. Peak signal to noise ratio (PSNR)

Peak Signal-to-Noise Ratio (PSNR) is a widely used metric to assess the quality of digital images. It's important to note that a higher PSNR value indicates higher image quality. PSNR is usually calculated as the logarithmic term of the decibel scale due to signals having a wide dynamic range. This dynamic range varies between the largest and smallest possible values and can vary depending on their quality. PSNR is primarily calculated via Mean Square Error (MSE). The PSNR value approaches infinity as the MSE approaches zero, indicating that a higher PSNR value provides higher image quality. PSNR refers to situations where the numerical differences between images are large [20, 22]. Formula of PSNR is given in Equation 6.

$$PSNR = 20.log_{10}\left(\frac{MAX}{\sqrt{MSE}}\right) \tag{6}$$

2.3.4. Learned perceptual image patch similarity (LPIPS)

The Learned Perceptual Image Patch Similarity (LPIPS) metric is a perceptual similarity measure that utilizes deep features extracted from trained neural networks to assess visual similarity between images. This metric evaluates

relative image similarity. Unlike traditional metrics such as Mean Square Error (MSE) or Peak Signal-to-Noise Ratio (PSNR), which generally do not align with human perception, LPIPS has demonstrated a remarkable ability to align closely with human perception [24].

III. RESULTS AND DISCUSSIONS

This study investigates an approach to expand the data collection capacity of event cameras. The unique approach of creating an event in a static scene with no motion using a fixed event camera sets it apart. The use of pattern sequences projected through an LCD projector is a technique that provides controlled illumination and prevents instantaneous generation of high number of events that may possibly exceed the bandwidth. This approach aims to optimize the data collection process with event cameras. The quality scores calculated as a result of applying the E2VID and FireNet algorithms to the data obtained through the DV software are given in Table 2 and Table 3.

For E2VID, the images of Pattern Sequences 1 and Pattern Sequences 2 corresponding to the none and step decay functions are presented in Figure 4. For FireNet, the images of Pattern Sequences 1 and Pattern Sequences 2 corresponding to the exponential decay function are provided in Figure 5. The ground truth images with equalized timestamps are shown in Figure 6.

To better clarify the transition functions mentioned above, the decay mechanisms used in this study are defined as follows [25]:

- None Decay: No decay is applied to pixel intensity over time. After an event updates a pixel, its intensity remains unchanged unless another event occurs.
- Step Decay: After frame generation, all pixel potentials are immediately reset to a predefined neutral value, effectively clearing the accumulated contributions at once. This behavior is mathematically formulated in Equation (7) :

$$I(t) = \begin{cases} I_{0,} & t \leq t_0 \} \\ I_{neutral,} & t > t_0 \end{cases} \tag{7}$$

where $t_{0,}$ is the frame generation time and $I_{neutral,}$ is the neutral potential value.

- Exponential Decay: Pixel intensity decreases gradually over time following an exponential law, which is defined in Equation (8):

$$I(t) = I_0 \times exp\left(-\frac{t}{\tau}\right) \tag{8}$$

where I_0 is the initial intensity, t is the elapsed time, and τ is the decay time constant that governs the speed of attenuation. These decay strategies were specifically selected to manage the event accumulation dynamics and to prevent potential issues such as bandwidth saturation during the controlled event generation process.

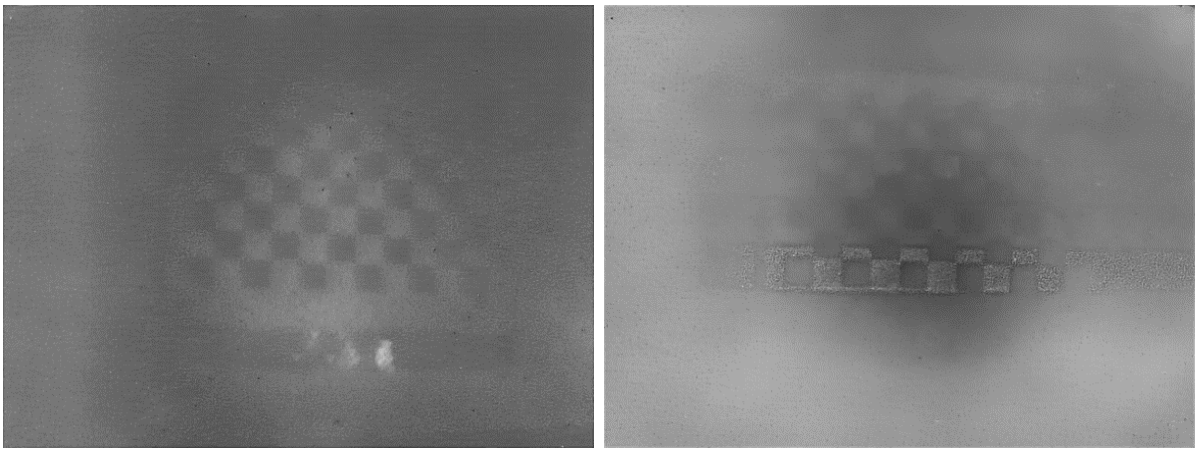

Figure 4. Frame of pattern sequences 1 (left) for none decay function and pattern sequences 2 (right) for the step decay function (E2VID)

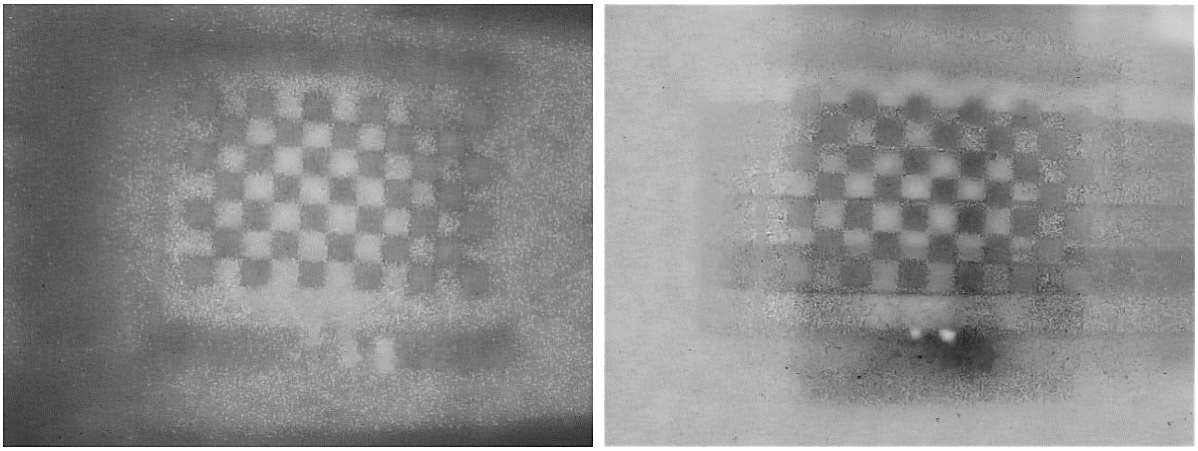

Figure 5. Frame of pattern sequences 1 (left) and pattern sequences 2 (right) for the exponential decay function (FireNet)

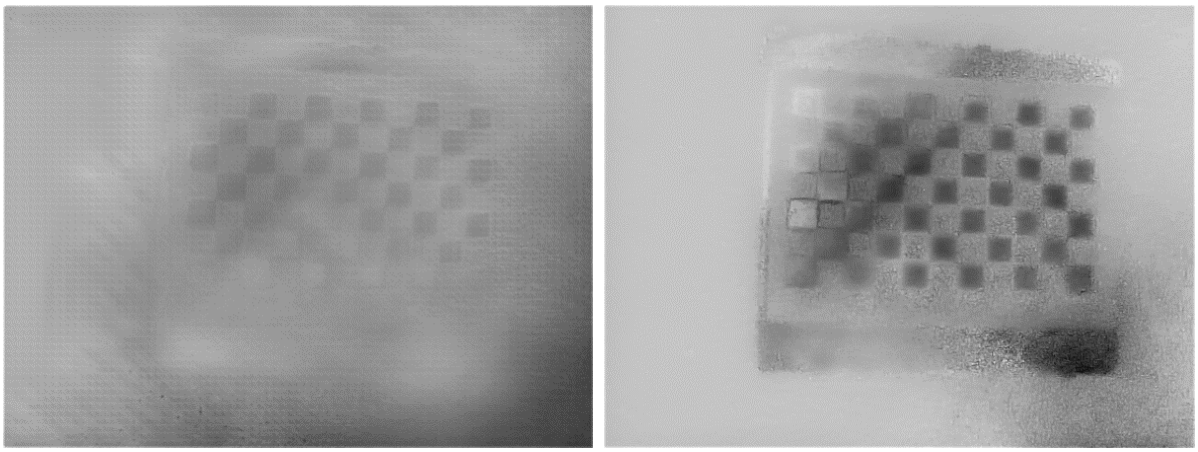

Figure 6. Frames of ground truth for the exponential decay function using E2VID (left) and FireNet (right)

Table 2. Result of E2VID on event camera dataset

Data	Decay Function	MSE	PSNR *(dB)*	SSIM	LPIPS
Pattern Sequences 1	None	0.0254	15.96	0.7860	0.3064
Pattern Sequences 2	None	0.0414	13.82	0.7743	0.3167
Pattern Sequences 1	Exponential	0.0154	18.11	0.8357	0.3444
Pattern Sequences 2	Exponential	0.0064	21.95	0.8439	0.2805
Pattern Sequences 1	Step	0.0091	20.42	0.8474	0.2723
Pattern Sequences 2	Step	0.0062	22.07	0.8495	0.2608

Table 3. Result of FireNet on event camera dataset

Data	Decay Function	MSE	PSNR *(dB)*	SSIM	LPIPS
Pattern Sequences 1	None	0.0289	15.38	0.4451	0.5944
Pattern Sequences 2	None	0.0232	16.35	0.5649	0.5538
Pattern Sequences 1	Exponential	0.0341	14.67	0.4683	0.6196
Pattern Sequences 2	Exponential	0.0151	18.21	0.5688	0.5493
Pattern Sequences 1	Step	0.0289	15.38	0.4179	0.6210
Pattern Sequences 2	Step	0.0195	17.10	0.5643	0.5779

The results demonstrate that using the proposed projection-based event generation method, reconstructed images can be obtained when employing existing E2VID and FireNet methods. Consistent with findings in the literature, E2VID also tends to produce better image reconstruction results in our experiments. While the tables indicate that reasonable image reconstructions can be achieved, challenges remain, particularly in terms of noise and distortion, as reflected in the PSNR values (~15 dB to 22 dB). These findings highlight both the strengths and limitations of the proposed method, pointing to potential areas for future improvements.

To better contextualize our findings within the existing literature, a comparison with the recently proposed Noise2Image method [19] is provided. Noise2Image aims to reconstruct static scenes by leveraging spontaneous photon noise events captured by event cameras, without the use of active illumination. In contrast, our approach stimulates event generation through systematically projected contrast patterns, offering controlled event production rather than relying on noise characteristics. Although the underlying methodologies differ, the PSNR (~17–25 dB) and SSIM (~0.5–0.74) values reported in Noise2Image are comparable to our results (~15–22 dB PSNR and ~0.44–0.84 SSIM). This comparison suggests that the proposed controlled-illumination framework achieves reconstruction quality in line with recent developments, while providing an alternative event generation strategy for static scene recovery. In addition to PSNR and SSIM, the MSE and LPIPS metrics further characterize the reconstruction quality. Moderate MSE values observed (~0.0062 to ~0.0414) indicate that while pixel-wise intensity errors are present, they remain within acceptable ranges for event-based reconstruction, where asynchronous data collection inherently introduces noise. Similarly, LPIPS scores (~0.26 to ~0.62) suggest perceptual discrepancies between reconstructed and ground-truth images, particularly regarding fine texture details and local contrast. It should also be noted that capturing event data from static scenes poses inherent challenges due to the absence of natural brightness changes. Although controlled contrast patterns are projected onto the scene, surface reflections and scattering can cause localized, unintended brightness variations. These phenomena may generate noise-like events even under carefully designed illumination conditions. As a result, moderate reconstruction errors reflected in the MSE and LPIPS values are expected outcomes of static scene recovery using event cameras. These observations are consistent with the qualitative inspection of reconstructed images shown in Figures 4, 5, and 6.

The proposed framework offers potential contributions to both academic research and industrial applications. Academically, it enables controlled event data collection from static scenes, providing a valuable experimental platform for studying and benchmarking event-based image reconstruction algorithms. This can facilitate the development of more robust event processing models and datasets. Industrially, the approach could be applied to enhance monitoring systems operating under low-light or static conditions, such as long-term surveillance cameras, automotive perception systems with static field-of-view sensors, or optical quality control in manufacturing lines. By regulating event generation in static settings, the proposed method helps ensure reliable

data acquisition without overloading sensor bandwidth, paving the way for practical implementations of event cameras in broader application domains However, the proposed method is primarily designed for static scenes and may encounter limitations under dynamic conditions. When objects move during pattern projection, the superposition of motion-induced and pattern-induced events can introduce artifacts, impairing the reconstruction quality. Similarly, highly reflective surfaces may cause optical scattering, while extremely low-contrast regions may generate insufficient or excessive events, both of which affect the reliability of the reconstructed images. These factors highlight that although the method offers a controlled framework for static scene recovery, its applicability may be constrained in more complex or dynamic environments.

IV. CONCLUSIONS

In this study, a framework method is proposed in order to enable event data gathering in scenarios with a static event camera in a static scene. An analysis was conducted using event data obtained from an event camera, leveraging custom designed grayscale pattern sequences projected onto a static scene. The generated event data were processed using the E2VID and FireNet methods to reconstruct the image of the scene and the performance of our proposed method is evaluated. It is observed that the event data collected with our method yields a reasonable reconstruction of the scene both with E2VID and FireNet. Although the PSNR values obtained in this study were quite low, and the reconstructed images did not reach a high quality visual representation of the scene, this study demonstrates the possibility to obtain measurements with a static event camera in a static scene. Results indicate that further improvements in reconstruction algorithms are necessary to enhance event-based imaging.

Additionally, the careful selection of contrast-compatible grayscale tones instead of abrupt black-to-white transitions allowed for better event rate control, preventing excessive event generation and potential system bottlenecks. Furthermore, the implementation of timestamp equalization helped prevent event loss over time. Despite its limitations, the findings suggest that event cameras have the potential to expand their application areas including static scenes with a static event camera.

Future research can focus on improving image quality by exploring alternative reconstruction techniques, advanced denoising methods, and dynamic parameter optimization strategies. These efforts aim to enhance the usability of event cameras in static environments more effectively. Finally, based on the results of these experiments, it can be concluded that an LCD projector-based approach, when combined with timestamp equalization and optimized contrast transitions, may serve as an effective method for generating events from static objects in applications utilizing static event cameras.

While the proposed method aims to expand the application areas of event cameras to static scenes, it should be noted that direct comparison with traditional dynamic scene event camera applications is not methodologically appropriate due to the fundamentally different nature of event generation. In this study, performance comparisons were made with the recent Noise2Image approach [19], which also targets static scene recovery, albeit with different underlying methodologies. Both works share the common objective of reconstructing intensity information from static scenes where conventional event triggering mechanisms are absent. Given the novelty of static scene applications for event cameras, the available literature for direct comparison is currently limited. Future research could focus on developing broader benchmarks and systematic evaluations for static event-based imaging.

ACKNOWLEDGMENT

This study was supported by the Scientific and Technological Research Council of Türkiye (TÜBİTAK) under project number 123E693.

ETHICAL STATEMENT

All event data used in this study were collected by the authors using their own experimental setup involving a static scene and an LCD projector. No human subjects, personal data, or third-party datasets were involved. Therefore, ethical approval was not required for this study.

REFERENCES

1. Scheerlinck C, Rebecq H, Gehrig D, Barnes N, Mahony RE, Scaramuzza D (2020) Fast image reconstruction with an event camera. IEEE Winter Conf Appl Comput Vision (WACV) 2020:156–163. https://doi.org/10.1109/WACV45572.2020.9093366
2. Alonso I, Murillo AC (2019) EV-SegNet: Semantic segmentation for event-based cameras. IEEE/CVF Conf Comput Vis Pattern Recognit Work (CVPRW) 2019:1624–1633. https://doi.org/10.1109/CVPRW.2019.00205
3. Salah M, Ayyad A, Humais M, Gehrig D et al (2024) E-Calib: A fast, robust, and accurate calibration toolbox for event cameras. IEEE Trans Image Process 33:3977–3990. https://doi.org/10.1109/TIP.2024.3410673
4. Sun R, Shi D, Zhang Y, Li R, Li R (2021) Data-driven technology in event-based vision. Complexity 2021:1–19. https://doi.org/10.1155/2021/6689337
5. Zhu AZ, Thakur D, Ozaslan T, Pfrommer B, Kumar V, Daniilidis K (2018) The multivehicle stereo event camera dataset: An event camera dataset for 3D perception. IEEE Robot Autom Lett 3(3):2032–2039. https://doi.org/10.1109/LRA.2018.2800793
6. Stoffregen T, Gallego G, Drummond T, Kleeman L, Scaramuzza D (2019) Event-based motion segmentation by motion compensation. IEEE/CVF Int Conf Comput Vis (ICCV) 2019:7243–7252. https://doi.org/10.1109/ICCV.2019.00734
7. Rebecq H, Gehrig D, Scaramuzza D (2018) ESIM: An open event camera simulator. Conf Robot Learning (CoRL) 87:969–982
8. Pan L, Hartley R, Scheerlinck C, Liu M, Yu X, Dai Y (2022) High frame rate video reconstruction based on an event camera. IEEE Trans Pattern Anal Mach Intell 44(5):2519–2533. https://doi.org/10.1109/TPAMI.2020.3036667
9. Gallego G, Delbruck T, Orchard G, Bartolozzi C et al (2022) Event-based vision: A survey. IEEE Trans Pattern Anal Mach Intell 44(1):154–180. https://doi.org/10.1109/TPAMI.2020.3008413
10. Scheerlinck C, Rebecq H, Stoffregen T, Barnes N et al (2019) CED: Color event camera dataset. Conf Comput Vis Pattern Recognit Work (CVPRW) 2019:1684–1693. https://doi.org/10.1109/CVPRW.2019.00215
11. Adra M, Dugelay J-L (2024) TIME-E2V: Overcoming limitations of E2VID. IEEE Int Conf Adv Video Signal Based Surveillance (AVSS) 2024:1–7. https://doi.org/10.1109/AVSS61716.2024.10672575
12. Ronneberger O, Fischer P, Brox T (2015) U-Net: Convolutional networks for biomedical image segmentation. CoRR abs/1505.04597. http://arxiv.org/abs/1505.04597
13. Zhu L, Wang X, Chang Y, Li J, Huang T, Tian Y (2022) Event-based video reconstruction via potential-assisted spiking neural network. Proc IEEE/CVF Conf Comput Vis Pattern Recognit (CVPR) 2022:3594–3604.
14. Mei H, Wang Z, Yang X, Wei X, Delbruck T (2023) Deep polarization reconstruction with PDAVIS events. Proc IEEE/CVF Conf Comput Vis Pattern Recognit (CVPR) 2023:22149–22158.
15. Inivation (2025) Enhanced noise filtering from Inivation's DV platform. https://inivation.com/enhanced-noise-filtering-from-inivations-dv-platform/. Accessed 30 January 2025
16. Pakdelazar O, Rezai Rad G (2011) Improvement of BM3D algorithm and employment to satellite and CFA images denoising. Int J Inf Sci Tech 1(3):23–33. https://doi.org/10.5121/ijist.2011.1303
17. Buades A, Coll B, Morel J (2011) Non-local means denoising pixelwise implementation. Image Process Line 1:208–212.

18. Shalini K, Prasad LVN (2014) Wavelet based soft thresholding approach for color image denoising. Int J Adv Eng Technol 7(4):1233–1237.
19. Cao R, Galor D, Kohli A, Yates JL, Waller L (2024) Noise2Image: Noise-enabled static scene recovery for event cameras. Optica. https://doi.org/10.1364/OPTICA.538916
20. Horé A, Ziou D (2010) Image quality metrics: PSNR vs. SSIM. 20th Int Conf Pattern Recognit (ICPR) 2010:2366–2369. https://doi.org/10.1109/ICPR.2010.579
21. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: From error visibility to structural similarity. IEEE Trans Image Process 13(4):600–612. https://doi.org/10.1109/TIP.2003.819861
22. Sara U, Akter M and Uddin M (2019) Image Quality Assessment through FSIM, SSIM, MSE and PSNR—A comparative study. J Comput Commun 7:8-18. https://doi.org/10.4236/jcc.2019.73002
23. Wang Z, Bovik AC (2009) Mean squared error: Love it or leave it? A new look at signal fidelity measures. IEEE Signal Process Mag 26(1):98–117. https://doi.org/10.1109/MSP.2008.930649
24. Zhang R, Isola P, Efros AA, Shechtman E, Wang O (2018) The unreasonable effectiveness of deep features as a perceptual metric. IEEE/CVF Conf Comput Vis Pattern Recognit (CVPR) 2018:586–595. https://doi.org/10.1109/CVPR.2018.00068
25. Event Accumulation (2025) https://dv-processing.inivation.com/rel_1_7/accumulators.html. Accessed 29 April 2025.