2023 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)
June 28-30, 2023. Seattle, Washington, USA

Understanding and Controlling the Sensitivity of Event Cameras in Responding to Static Objects

Qiyao Gao[1], Xiaoyang Sun[1], Zhitao Yu[1], Xu Chen[†,1]

***Abstract*— Event cameras are sensors that asynchronously measure the brightness change of each pixel with a microsecond-level time resolution. They have reduced motion blur and a dynamic response range of up to 140 dB, which allows them to handle extreme lighting conditions better than traditional frame cameras. Most of the research on event cameras has focused on dynamic vision, neglecting their potential applications in static scenes. This paper investigates how event cameras respond to stationary objects and provides a comprehensive theoretical analysis. We show that the event camera's output depends on the brightness of the objects and their circuit structure. We also derive a mathematical formula that relates the event camera's sensitivity to the number of absorbed photons. Furthermore, we propose a method to modulate the stationary output of event cameras.**

I. INTRODUCTION

Unlike traditional cameras that acquire full-field images at a specific frequency, in a dynamic vision sensor (DVS) such as an event camera, each pixel asynchronously and independently responds to changes in light intensity. When such light-intensity change exceeds a threshold, the DVS outputs ”events” containing coordinate, time, and polarity of the change. The amplification circuit deployed around each pixel enables microsecond-level time resolution and localized response, dramatically reducing data size and increasing sensing speed. However, vision algorithms designed for image sequences are not suitable for event cameras because they output in the form of address events (AEs). Developing algorithms compatible with the AE outputs or converting the output to image formats are the two mainstream methods for using event cameras. Previous work focused on developing algorithms to unleash their potential in dynamic scenes. For example, Davide Falanga et al. [1] equipped a drone with an event camera to avoid dynamic obstacles at a speed of 10m/s. The algorithm they designed compensates for the event camera's own motion (formally known as the ego-motion) to distinguish between static and dynamic objects. Olivier Bichler et al. [2] trained spiking neural networks through spike-timing-dependent plasticity to extract such temporally correlated features as car trajectories from the DVS.

Existing literature have explored the potential of event cameras in high-motion speed applications. However, we observe that event cameras also have a high dynamic response range of up to 140dB, which enables application potentials in static conditions that are difficult or infeasible for traditional frame cameras to handle. Capitalizing on this observation, this paper reveals a novel phenomenon related to how event cameras respond to static objects of different colors. Specifically, when there is a dark object in the field of view, the event camera will generate more ”events” in the corresponding pixel area, even without any noticeable light intensity changes (cf. Fig. 1). Our experimentation further indicates that: (a) the phenomenon can be noticeably observed only after at least a few seconds of output accumulation; (b) the event camera responds little to light-colored objects; and (c) the response is independent of the structure and material of the object.

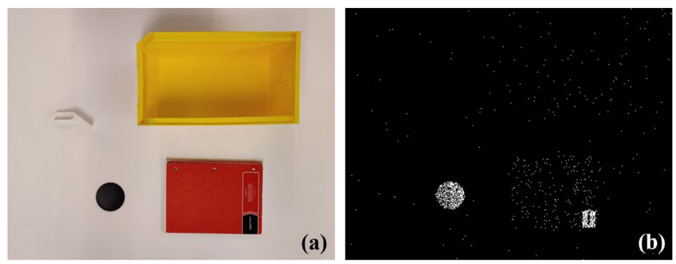


Fig. 1: Event-based image generated with 10 seconds of exposure in presence of objects with different colors, shapes, and materials

We propose that the event camera's sensitivity to static objects depends on their luminance. Fig. 2 shows how an event camera responds to a piece of paper with different grayscale values in a constant lighting environment. As shown, the event camera exhibits high sensitivity to low-grayscale objects, namely, it generates more events when responding to darker objects. The result breaks the traditional view that event cameras only respond to changes in light intensity, suggesting the possibility of its application in static scenes. The research on the sensitivity of event cameras in this paper provides understanding and control of the mechanism.

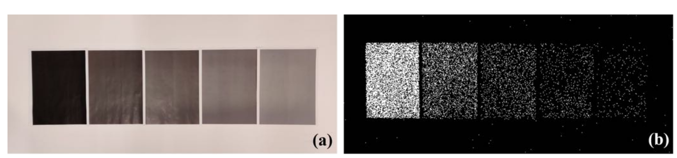


Fig. 2: Printing papers with zones of different grayscale values (left) and 10-second cumulated output of the event camera (Prophesee EVK3) (right). The results show that for paper with a lower grayscale value (darker), the event camera generates more events, and as the grayscale value increases (whiter), the number of events per pixel area decreases significantly

†: corresponding author.
[1]Department of Mechanical Engineering, University of Washington, Seattle, WA, USA {gao1997, dylansun, zhitaoyu, chx} @uw.edu

II. Theoretical analysis

To test the sensitivity of the event-based camera against static objects, we accumulated the output of the DVS over an interval and converted it to a frame image. We then used the number of events in the image to evaluate the sensitivity of the event camera to static objects. Under this imaging setting, we define the static sensitivity (S_s) of the event camera as the number of events generated per unit pixel area per unit time:

$$S_s = \frac{N_e}{A \times T}, \tag{1}$$

where N_e is the number of events, A is the pixel area, and T is the accumulation time.

The basic structure of the DVS is a logarithmic pixel circuit structure modeled on the three-layer Kufler retina and encodes pixel coordinates in the form of address events [3], [4], [5], [6], [7], [8]. The circuit of the DVS sensor consists

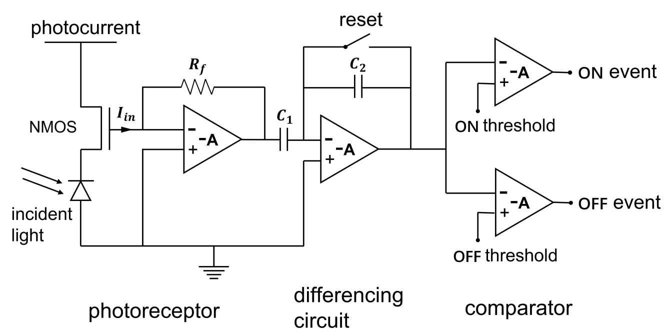


Fig. 3: Schematic of the DVS circuits (redrawn from [4])

of a fast logarithmic photoreceptor circuit, a differential circuit for amplifying changes with high precision, and an inexpensive two-transistor comparator, as shown in Fig. 3. The process of event generation can be divided into three steps. First, the photodiode in the photoreceptor circuit absorbs photons and the built-in saturated NMOS transistor generates a photocurrent, which is logarithmically converted to a voltage through a transimpedance configuration. Second, the differential circuit buffers the output of the photosensor through a source follower, reversely amplifies it through a capacitive feedback amplifier, and then outputs the result to a comparator. Third, the comparator compares the output to a global threshold offset from the reset voltage. If there is a change crossing the threshold, an event will be generated depending on the polarity. Unlike CCDs, each pixel of a CMOS sensor in the DVS has an independent gain circuit (through its logarithmic response) around it. The resulting DC mismatch will be removed by balancing the output of the differential circuit to a reset level.

Both the logarithmic conversion in the photoreceptor circuit and the DC elimination in the differential circuit contribute to the result that the pixel is sensitive to the temporal contrast (T_c) [4], which is defined as:

$$T_c = \frac{1}{I(t)} \times \frac{dI(t)}{dt} = \frac{d \ln I(t)}{dt}, \tag{2}$$

where I is the photocurrent. The temporal contrast decreases with increasing photocurrents. As the photocurrent is generated by the photodiode through the internal photoelectric effect, the sensitivity is therefore related to the number of absorbed photons.

A. Application of Photon Discreteness Theory

Event cameras can easily capture high-speed moving objects because they can respond asynchronously to the light-intensity changes of each pixel. The core component in this process is the photodiode. In a PIN photodiode, photons absorbed in the depletion or intrinsic region generate electron-hole pairs, most of which move to produce photocurrent. This phenomenon, known as the internal photoelectric effect and explained by Einstein's quantum theory of light, is widely used in the manufacture of optoelectronic devices. Due to the particle nature of light, the photon absorption is not continuous in time, and the number of photons absorbed at one moment may be significantly different from the previous moment. This fluctuation is pronounced when the basis is small, and leads to the shot noise, a type of noise prevalent in optical and electronic components.

Given that the photocurrent is proportional to light intensity, and that the light intensity is related to the number of photons per unit time, we hypothesize that the event camera is more sensitive to dark objects due to their lower reflectivity of photons, hence a small photocurrent received by the event camera. Such a small cardinality of photons leads to large photocurrent fluctuations. As the photocurrent gets compared to the global threshold, the shot noise at low photon cardinality will cause the photocurrent to cross the threshold frequently, triggering more events.

B. Proposed Static Sensitivity Analysis Formula

To quantify the event camera's sensitivity, we need to establish a mathematical relationship between the number of events and the number of absorbed photons. Patrick Lichtsteiner et al. [4] gave the formula of the event-generation rate:

$$f(t) = \text{Event Rate}(t) \approx \frac{T_c}{\theta} = \frac{1}{\theta}\frac{d \ln I}{dt}, \tag{3}$$

where θ is the temporal contrast threshold and the photocurrent intensity comes from:

$$I = R \times P, \tag{4}$$

where R is the detector responsivity (built-in parameter), and P is the optical power, that is, the total energy of photons absorbed per unit of time:

$$P = \frac{N \times h_v}{\Delta T}, \tag{5}$$

where N is the number of photons absorbed in a time window of ΔT seconds, and h_v is the energy of a single photon. Recalling Equation (1), the static sensitivity becomes:

$$S_s = \frac{\int_{t_1}^{t_2} f(t)dt}{A \times (t_2 - t_1)} = \frac{\frac{1}{\theta}\ln(\frac{N_{t_2}}{N_{t_1}})}{A \times (t_2 - t_1)}, \tag{6}$$

where N_{t_1} and N_{t_2} are the numbers of photons absorbed at time t_1 and t_2. Notice that the number of photons is not available in practice. However, in a static scene with the same light source, the incident light is a beam carrying the same density of photons. Hence, the main factor affecting the response of the event camera is the reflectivity of the object. Dark objects reflect less photons due to their smaller reflectivity, leading to a more sensitive response of the event camera.

III. Experiment and Analysis

We designed a series of experiments to verify the proposed theory, including an event-counting experiment to verify the formula of static sensitivity and an experiment to show the influence of reflectivity on sensitivity. Next, we will describe in detail each experiment's conception, implementation, and results.

A. Event Counting Experiment

Our hypothesis was that the fewer the number of photons reflected by an object, the more events will be generated in the corresponding pixel area. However, the reflectivity of physical objects is difficult to model accurately due to the many associated factors such as color, material, surface smoothness, etc. From the field of computer vision and human perception, grayscale value has been applied to measure reflectance and color [9][10]. We repurpose the concept here to evaluate the number of reflected photons.

The specific steps of the experiment are as follows. First, we print dark regions with different grayscale values in the order of the arithmetic sequence (cf. each gray box in Fig. 2). Papers of the same material and shape are used in four repeated trials to eliminate other interference factors. We use an event camera to observe these papers, accumulate the output for 0.3s, and form the obtained events into an image, as shown in Fig. 4. Then, we select regions at the same location in each image, count the number of events contained within, and calculate the event density (the number of events per unit time per unit pixel area). Finally, we draw the relationship diagram between event density and grayscale value, and compare it with the theoretical sensitivity curve, as shown in Fig. 5.

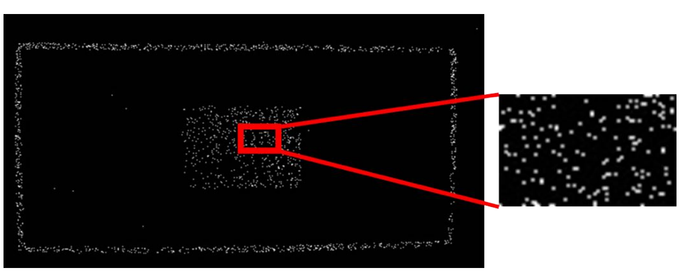
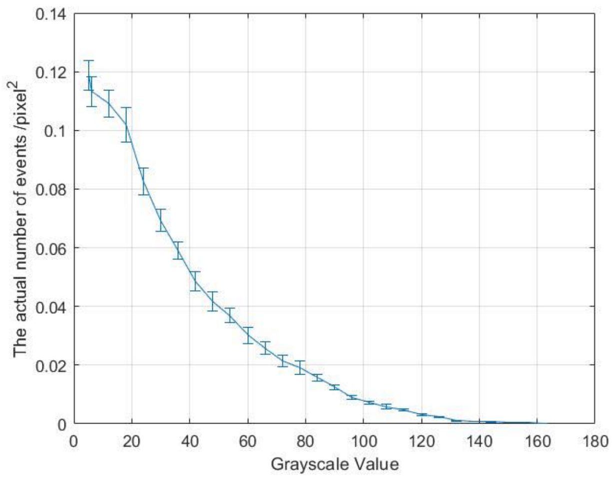
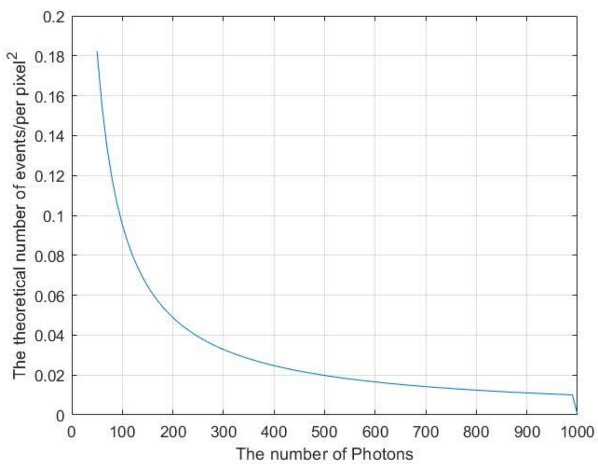


Fig. 4: We accumulated the event camera's response to papers with different grayscales, output the results in image format, and then selected the area at the same position to count the number of events within. The operation is repeated four times for each paper and the average value is taken as the number of events corresponding to the grayscale value



(a)



(b)

Fig. 5: a) Diagram of event density versus grayscale value. b) Relation between the measured events and the number of received photons by the event camera at the start of capturing an integrated timelapse image for the event camera. More specifically, we assume that the light source is a flow of photons with uniform density. Namely, the environmental lighting is stationary and the number of photons absorbed by the camera at each moment is consistent. We test the static sensitivity at uniformly sampled integration time instances where t_2 equals t_1 plus a fixed sampling time. Let ΔN = N_{t_2} - N_{t_1} be the difference between the number of photons absorbed at time t_1 and t_2. This number does not change the fundamental shape of the log function in Eq. (6), and in the demonstration, we fix it to ten. The abscissa represents N_{t_1} in Eq. (6).

We compute the event density by counting the number of events in a fixed region, which equals the static sensitivity. From Fig. 5, the results validate that the grayscale value and the number of photons are indeed highly dependent, and the number of photons reflected by an object increases as its surface grayscale value increases.

B. Reflectivity Experiment

The previous experiment indicates that the static output of event cameras is influenced by the object's surface grayscale value. To practically apply this theory, we suggest utilizing object surface reflectivity as a means of regulating the static output of event cameras. To test the feasibility of this technique, we conducted a series of controlled experiments (as shown in Fig. 6). The first group involved a black acrylic plate with a rough surface, resulting in a uniform event domain in Fig. 6 (d). We then applied scotch tape to the surface of the plate to alter its surface reflectance without affecting its grayscale value, resulting in missing events in the tape-covered area, as shown in Fig. 6 (e). This effect was further amplified in the third set of experiments where we covered the plate's surface with a transparent plastic bag, as shown in Fig. 6 (f).

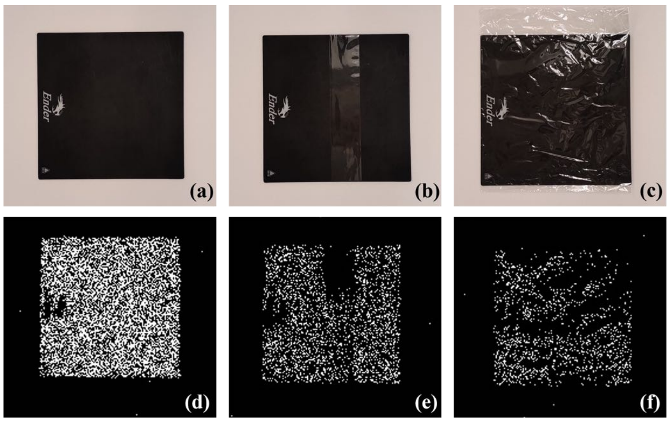


Fig. 6: Upper row: a) standard blackboard b) blackboard with scotch tape in the middle, and c) blackboard covered by a transparent bag. Bottom row: corresponding response results (cumulative 10s output)

The experimental results show that the number of events generated in the corresponding pixel area can be effectively controlled by changing the object's reflectivity. As a result, we can control the response of event cameras when observing static objects. For example, we can eliminate interference from the environment by using a highly reflective whiteboard (as shown in Fig. 7).

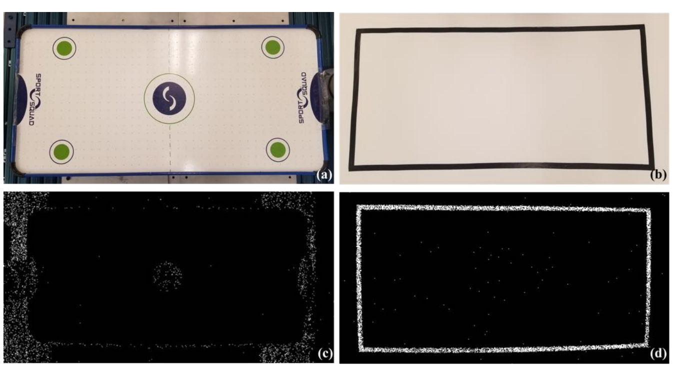


Fig. 7: Use a whiteboard with high reflectivity to eliminate ambient noise. Upper row: a) before placing the whiteboard b) after placing the whiteboard. Bottom row: corresponding response results (cumulative 5s output)

IV. Conclusion

This paper studied the response mechanism of event cameras to static objects and proposed a Photon Discreteness Theory to explain the phenomenon that event cameras are more sensitive to objects with low brightness. We defined a static sensitivity of the event camera and provided a computation method and experimentation verification of the sensitivity. Furthermore, we propose a technique based on the reflectivity of the object surface to stably adjust the static output of the event camera.

V. Limitations and Future Work

A. Photon Number versus Grayscale Value

In Section II and Section III, we established the relationship between the event camera's static sensitivity and the number of photons based on the internal photoelectric effect of the photodiode. In the experimental section, we repurposed grayscale values from computer vision to indirectly represent the number of photons. Although the black-body radiation theory proves that darker objects reflect fewer photons, future work remains to quantify the relationship between the grayscale value and the photon number.

B. Verifying the Presence of Shot Noise

We mentioned in Section II that shot noise dominates under low photocurrent conditions. It is difficult to separate shot noise from other noises. Work is underway to measure the magnitude of the shot noise. For example, we can count the probability distribution of the number of events generated in the case of low cardinality and observe whether it satisfies the Poisson Distribution. Another approach to obtain the shot noise voltage is by measuring the photocurrent flowing through the photodiode using carefully selected resistors.

References

[1] D. Falanga, K. Kleber, and D. Scaramuzza, "Dynamic obstacle avoidance for quadrotors with event cameras," *Science Robotics*, vol. 5, no. 40, p. eaaz9712, 2020.

[2] O. Bichler, D. Querlioz, S. J. Thorpe, J.-P. Bourgoin, and C. Gamrat, "Extraction of temporally correlated features from dynamic vision sensors with spike-timing-dependent plasticity," *Neural networks*, vol. 32, pp. 339–348, 2012.

[3] M. Mahowald, *An analog VLSI system for stereoscopic vision*. Springer Science & Business Media, 1994, vol. 265.

[4] P. Lichtsteiner, C. Posch, and T. Delbruck, "A 128×128 120 db $15\mu s$ latency asynchronous temporal contrast vision sensor," *IEEE journal of solid-state circuits*, vol. 43, no. 2, pp. 566–576, 2008.

[5] C. Posch, D. Matolin, and R. Wohlgenannt, "A qvga 143 db dynamic range frame-free pwm image sensor with lossless pixel-level video compression and time-domain cds," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 1, pp. 259–275, 2010.

[6] ——, "A qvga 143db dynamic range asynchronous address-event pwm dynamic image sensor with lossless pixel-level video compression," in *2010 IEEE International Solid-State Circuits Conference-(ISSCC)*. IEEE, 2010, pp. 400–401.

[7] C. Brandli, R. Berner, M. Yang, S.-C. Liu, and T. Delbruck, "A $240\times$ 180 130 db 3 μs latency global shutter spatiotemporal vision sensor," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 10, pp. 2333–2341, 2014.

[8] R. Berner, C. Brandli, M. Yang, S.-C. Liu, and T. Delbruck, "A $240\times$ 180 10mw 12us latency sparse-output vision sensor for mobile applications," in *2013 Symposium on VLSI Circuits*. IEEE, 2013, pp. C186–C187.

[9] M. Stevens, C. A. Párraga, I. C. Cuthill, J. C. Partridge, and T. S. Troscianko, "Using digital photography to study animal coloration," *Biological Journal of the Linnean society*, vol. 90, no. 2, pp. 211–237, 2007.

[10] P. Kevan, N. Grainger, G. Mulligan, and A. R. Robertson, "A grayscale for measuring reflectance and color in the insect and human visual spectra," *Ecology*, vol. 54, no. 4, pp. 924–926, 1973.