

DISS. ETH NO. 30131

# A SCIENTIFIC EVENT CAMERA: THEORY, DESIGN, AND MEASUREMENTS

A thesis submitted to attain the degree of  
DOCTOR OF SCIENCES  
(Dr. sc. ETH Zurich)

presented by  
RUI GRAÇA  
M.Sc., University of Porto  
born on 14.02.1992

accepted on the recommendation of

Prof. Dr. Tobi Delbruck  
Prof. Dr. Shih-Chii Liu  
Prof. Dr. Teresa Serrano-Gotarredona  
Prof. Dr. Benjamin Grewe

2024

# Abstract

This thesis explores the physical limits of neuromorphic event cameras and proposes a novel highly-sensitive event camera targeting scientific applications.

Certain scientific applications, such as fluorescent imaging of neural activity (especially voltage imaging), as well as tracking of particles/objects moving at high speed, require vision sensors operating near the limits of physics.

These applications consist in the detection of low contrast changes in light intensity during time intervals of only a few milliseconds or less, in observations that can last several minutes.

Currently, these applications rely on scientific image sensors capable of capturing thousands of frames per second. While the paradigm of frames is highly prevalent in computer vision, it has significant downsides. Namely, for the applications described, the acquisition of thousands of frames per second during several minutes leads to a highly redundant output, resulting in an extremely inefficient data utilization.

Vision in animals is the product of millions of years of evolution through natural selection, resulting in visual systems orders of magnitude more efficient than frame-based cameras. Neuromorphic silicon retinas such as the Dynamic Vision Sensor (DVS) event camera draw inspiration from biological visual systems to build more efficient vision sensors.

Characteristics of the DVS such as its high-speed performance with low latency and high data efficiency, as well as its high dynamic range, make it an emerging technology with growing popularity over the last years. These characteristics make the DVS a promising candidate for the scientific applications mentioned above. However, DVS implementations proposed before this work did not demonstrate sufficient sensitivity in the light-constrained settings required by these applications.

**The main purpose of the work presented in this thesis is the development of a novel DVS event camera with improved sensitivity under dim light.** To achieve this goal, this thesis investigates the physical limits of the DVS technology, demonstrating that the DVS is limited to

a minimum of 2x shot noise, and providing the conditions for the camera to operate near this limit. It also shows that spatial and temporal integration of light are fundamental to improve sensitivity in the dark - a result known from other visual systems, but never fully exploited in the DVS. This new knowledge, resulting from extensive measurements of DVS cameras and supported by theoretical analysis, resulted in a more realistic model of the DVS pixel, capable of reproducing measured phenomena and aligning with theory.

The results obtained are useful for DVS users, by providing optimal biasing strategies, for algorithm developers, by providing novel interpretations and insight about DVS data encoding, and for DVS designers, by defining the limits of the technology and optimization goals.

Finally, supported by an improved understanding of the DVS pixel and its limits, this thesis proposes SciDVS: A scientific event camera capable of responding to edges of 1.7% contrast under dim light settings at 0.7 lx on-chip illuminance. SciDVS features an array of  $126 \times 112$  pixels with a pitch of  $30 \mu\text{m}$ , implemented on a 180 nm CMOS Image Sensor process. The SciDVS pixel introduces novelty such as an auto-centering high dynamic range pre-amplifier, improved bandwidth control achieving cutoff frequencies down to 3.5 Hz, and pixel binning.

# Zusammenfassung

Diese Dissertation erforscht die physikalischen Grenzen von neuromorphen Event-Kameras und schlägt eine neuartige, hochsensible Event-Kamera für wissenschaftliche Anwendungen vor.

Bestimmte wissenschaftliche Anwendungen, wie die fluoreszierende Bildgebung neuronaler Aktivität (insbesondere Voltage Imaging), sowie das Verfolgen von Partikeln/Objekten, die sich mit hoher Geschwindigkeit bewegen, erfordern Bildsensoren, die nahe an den Grenzen der Physik arbeiten.

Diese Anwendungen bestehen in der Erkennung von Kontraständerungen geringer Intensität im Licht während Zeitintervallen von nur wenigen Millisekunden oder weniger, bei Beobachtungen, die mehrere Minuten dauern können.

Derzeit verlassen sich diese Anwendungen auf wissenschaftliche Bildsensoren, die in der Lage sind, tausende von Bildern pro Sekunde aufzunehmen. Während das Paradigma der frame-basierten Bildfolgen in der Computer Vision sehr verbreitet ist, hat es bedeutende Nachteile. Nämlich führt die Aufnahme von tausenden Bildern pro Sekunde über mehrere Minuten zu einem hochgradig redundanten Ausgang, was zu einer extrem ineffizienten Datennutzung führt.

Das Sehen bei Tieren ist das Ergebnis von Millionen Jahren Evolution durch natürliche Selektion, was zu Sehsystemen führte, die um Größenordnungen effizienter sind als frame-basierte Kameras. Neuromorphe Silizium-Netzhäute, wie die Dynamic Vision Sensor (DVS) Event-Kamera, lassen sich von biologischen Sehsystemen inspirieren, um effizientere Bildsensoren zu entwickeln.

Charakteristika der DVS, wie ihre Hochgeschwindigkeitsleistung mit niedriger Latenz und hoher Dateneffizienz, sowie ihr hoher Dynamikbereich, machen sie zu einer aufkommenden Technologie, die in den letzten Jahren zunehmend beliebter wird. Diese Charakteristika machen die DVS zu einem vielversprechenden Kandidaten für die oben genannten wissenschaftlichen Anwendungen. Jedoch haben vor dieser Arbeit vorgeschlagene DVS-Implementierungen nicht die erforderliche Sensitivität in den lichtbeschränkten Einstellungen gezeigt,

die von diesen Anwendungen gefordert werden.

Das Hauptziel der in dieser Dissertation vorgestellten Arbeit ist die Entwicklung einer neuartigen DVS-Event-Kamera mit verbesserter Sensitivität unter schwachem Licht. Um dieses Ziel zu erreichen, untersucht diese Dissertation die physikalischen Grenzen der DVS-Technologie und demonstriert, dass die DVS auf ein Minimum von  $2x$  Shot Noise begrenzt ist, und liefert die Bedingungen dafür, dass die Kamera nahe dieses Limits operieren kann. Es zeigt auch, dass räumliche und zeitliche Integration von Licht grundlegend sind, um die Sensitivität im Dunkeln zu verbessern - ein Ergebnis, das von anderen Sehsystemen bekannt ist, aber nie vollständig in der DVS ausgenutzt wurde. Dieses neue Wissen, das aus umfangreichen Messungen von DVS-Kameras resultiert und durch theoretische Analyse unterstützt wird, führte zu einem realistischeren Modell des DVS-Pixels, das gemessene Phänomene reproduzieren und sich mit der Theorie in Einklang bringen kann.

Die erzielten Ergebnisse sind nützlich für DVS-Benutzer, indem sie optimale Voreinstellungsstrategien bieten, für Algorithmientwickler, indem sie neue Interpretationen und Einblicke in die DVS-Datenkodierung bieten, und für DVS-Designer, indem sie die Grenzen der Technologie und Optimierungsziele definieren.

Schließlich schlägt diese Dissertation, unterstützt durch ein verbessertes Verständnis des DVS-Pixels und seiner Grenzen, SciDVS vor: Eine wissenschaftliche Event-Kamera, die auf Kanten von  $1.7\%$  Kontrast unter schwachen Lichtbedingungen bei einer  $0.7\text{ lx}$  Beleuchtung auf dem Chip reagieren kann. SciDVS verfügt über ein Array von  $126 \times 112$  Pixeln mit einem Pitch von  $30\text{ }\mu\text{m}$ , implementiert in einem  $180\text{ nm}$  CMOS Image Sensor Prozess.