Subpixel-accurate dense 4D measurements of dynamic scenes with event cameras and structured illumination

Christoph Freitag* **, Stefan Heist*, Andreas Stark***, Aliya Turelli Khusainova*,
Peter Kühmstedt*, Gunther Notni* ****, Christian Franke*** *****

**Fraunhofer Institute for Applied Optics and Precision Engineering IOF, Jena, Germany*
***Institute of Applied Physics, Friedrich Schiller University Jena, Germany*
****Institute for Applied Optics and Biophysics, Friedrich Schiller University Jena, Germany*
*****Group for Quality Assurance and Industrial Image Processing, Technische Universität Ilmenau, Germany*
******Jena Center for Soft Matter, Friedrich Schiller University Jena, Germany*

mailto:christoph.freitag@iof.fraunhofer.de

The method presented leverages event cameras and structured light to reconstruct accurate 4D point clouds (3D+time) of static and dynamic scenes. Event streams are processed into 4D points with subpixel accuracy. The interplay between our illumination and event cameras potentially yields both, dense 4D points and information on the movement of objects in the scene.

1 Introduction

Event cameras [1] have introduced a paradigm shift away from measurements with fixed frame rates as performed by traditional cameras. Instead, pixels of event cameras asynchronously react to changes in illumination that cross certain thresholds, adapting to the dynamics of the observed scene. Only changes in brightness result in reported data in the form of events (neglecting noise events caused by electronic noise in this brief report). A stream of events (x, y, p, t) with pixel coordinates x and y, a polarity p (sign of change), and a timestamp t requires different algorithms than classical images for post-processing. This is even more the case for two streams of events from a stereo configuration of two event cameras for 4D (3D+time) measurements (see Fig. 1). We introduce challenges and opportunities with dynamic scenes for a specific 4D measurement and reconstruction pipeline.

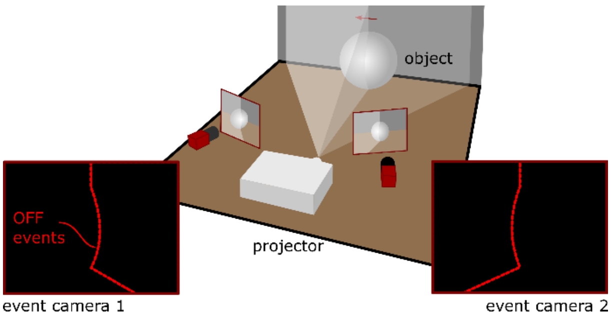

Fig. 1 *Setup of two event cameras (red), a projector (white), and a scene composed of a sphere and a background plane. The measurement volume is divided by a moving contrast edge (transition from a brighter to a darker section). The deformation of the contrast edge when projected on the event cameras' sensor planes is caused by variations in depth.*

2 Experimental setup and 4D reconstruction

The principle presented in Fig. 1 was realized in a lab setup. The projector for structured illumination is based on moving masks in front of an illumination source (see Fig. 2). Two event cameras observe the contrast edge's movement. Both event cameras are temporally synchronized, allowing to identify temporally correlating patterns in the two event streams they generate.

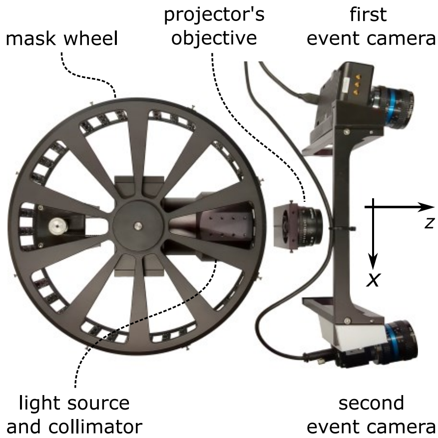

Fig. 2 *Top view of the event camera setup with part of the projector on the left and both event cameras on the right side.*

3D reconstruction based on event streams from static scenes has been done in a variety of ways [1]. Results of our previous approach can be seen in Fig. 3 with 3D reconstructions of sub-millimeter precision [2]. In contrast to the previous approach our new 4D reconstruction works directly on event streams instead of taking a detour over converting event streams to stacks of image frames. The algorithm processes small groups of events that potentially have seen the same point of the measured

DGaO Proceedings 2025 – http://www.dgao-proceedings.de – ISSN: 1614-8436 – urn:nbn:de:0287-2025-A034

submitted: 31.07.2025 published: 30.09.2025

scene (see Fig. 4). It finds pairs of each a reference pixel in the first event camera and a corresponding subpixel in the second camera which likely have seen the same point of the scene. They are identified by temporal proximity of the events and epipolar geometry. Intrinsic and extrinsic calibrations of the stereo set of cameras allows to triangulate those pairs of (sub)pixels (i.e. tracing and intersecting their corresponding pairs of rays) to obtain 3D points. These are augmented by the mean time of their generating events, which adds the fourth (temporal) dimension. The temporal resolution is adapted to the dynamics of measured events instead of fixed frame rates.

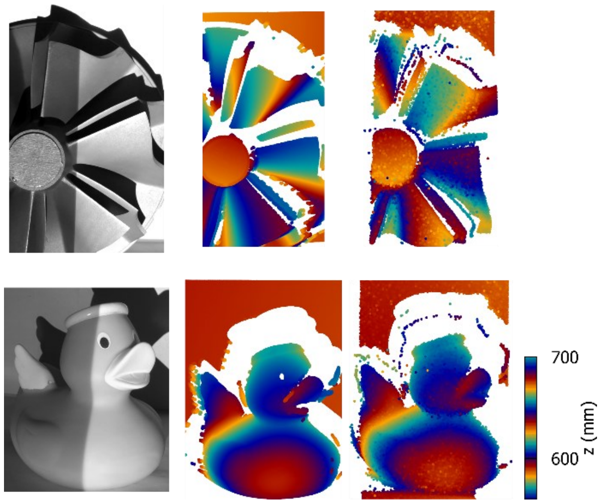

Fig. 3 Two static scenes, a metal turbine and a rubber duck, as measured by a conventional camera (left column), a conventional optical 3D sensor (central column), and our event-camera-based setup (right column). Color in the central and right columns represents depth.

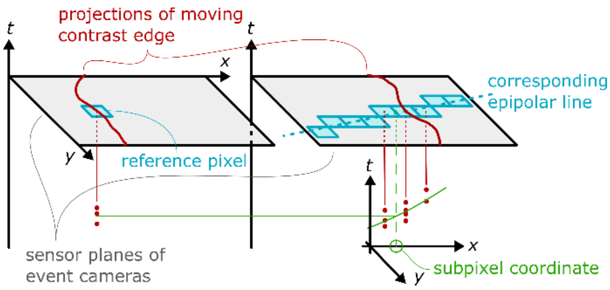

Fig. 4 *The principle of event-based stereo matching. The contrast edge (red) passes over the sensor planes (gray) of both event cameras with their pixels (blue). Red dots represent events lined up on the temporal axis. A local neighbourhood of pixels in the right camera has seen the contrast edge in temporal proximity to the reference pixel of the left camera (green horizontal line). Subpixel interpolation (model fit to event dots in right camera) intersected with the reference time yields a subpixel coordinate that temporally corresponds to the reference pixel and is the basis for triangulation.*

3 Challenges and opportunities with dynamic scenes

When scanning a static scene, only the projector's moving contrast edge stimulates events in the event cameras. In contrast, when there is movement, such as a bright sphere moving in front of a dark background as in Fig. 5, additional events are produced. On one hand, temporal matching becomes ambiguous as there are multiple stimuli to match at a time as visualized in Fig. 5. We attempt to resolve this by filtering based on a calibrated model of the projector. On the other hand, the additional events stimulated by an object's movement can be used for inferring parameters of that movement. This information is contained in the raw event streams. In contrast, raw data of traditional stereo 3D measurement setups with frame-based cameras does not contain information on the movement of objects, which instead must be inferred in post-processing.

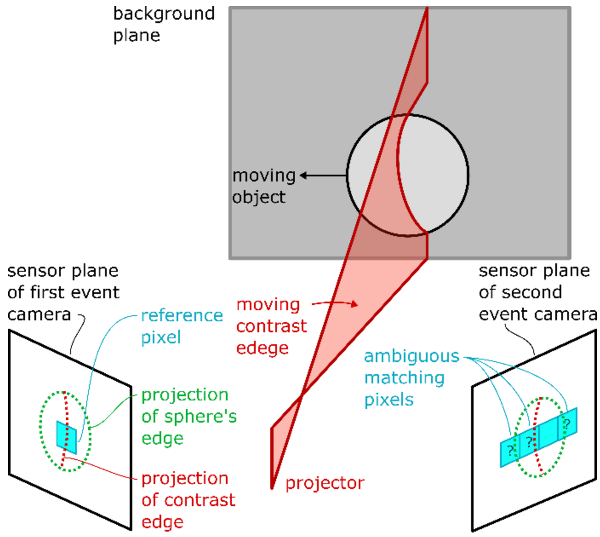

Fig. 5 *Principle of event cameras observing a dynamic scene (moving sphere) scanned by the projector's contrast edge (red). Due to the movement of the object itself, more changes in brightness (red and green on sensor planes) are measured, making the problem of temporal matching of events ambiguous (unclear which pixels of the right camera have seen the same change as the reference pixel of the left camera).*

4 Conclusion and outlook

Based on the presented setup and algorithms we intend to report experimental results of 4D measurements of dynamic scenes in the near future. This is to be extended by a systematic comparison of frame- and event-camera-based 4D sensors.

References

[1] G. Gallego, T. Delbrück, G. Orchard, C. Bartolozzi, T. Brian, A. Censi, S. Leutenegger, A. J. Davison, J. Conradt, K. Daniilidis und D. Scaramuzza, „Event-Based Vision: A Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Bd. 44, Nr. 1, pp. 154-180, 2022.

[2] C. Freitag, S. Heist, A. Stark, P. Kühmstedt, G. Notni and C. Franke, "Photometrically Optimized Event-based Stereo 3D Measurements," *Optics Express*, pp. 2924-2939, 2025.