Neuromorphic Fringe Projection Profilometry

Ashish Rao Mangalore, Chandra Sekhar Seelamantula, *Senior Member, IEEE*,
and Chetan Singh Thakur, *Senior Member, IEEE*

***Abstract*—We address the problem of 3-D reconstruction using neuromorphic cameras (also known as *event-driven* cameras), which are a new class of vision-inspired imaging devices. Neuromorphic cameras are becoming increasingly popular for solving image processing and computer vision problems as they have significantly lower data rates than conventional frame-based cameras. We develop a neuromorphic-camera-based Fringe Projection Profilometry (FPP) system. We use the Dynamic Vision Sensor (DVS) in the DAVIS346 neuromorphic camera for acquiring measurements. Neuromorphic FPP is faster than a single-line-scanning method. Also, unlike frame-based FPP, the efficacy of the proposed method is not limited by the background while acquiring measurements. The working principle of the DVS also allows one to efficiently handle shadows thereby preventing ambiguities during 2-D phase unwrapping.**

***Index Terms*—Event cameras, Neuromorphic cameras, Depth map, 3-D reconstruction, Fringe projection profilometry, DAVIS, Dynamic Vision Sensor (DVS).**

I. Introduction

NEUROMORPHIC cameras [1]–[3] are a recent invention and are being explored for analyzing optical flow, feature detection, image reconstruction, segmentation, video synthesis, 3-D reconstruction, etc. [4]–[12]. Neuromorphic cameras are event-driven and have several advantages such as sparse data, low latency, low power consumption, asynchronous sensing, and a high dynamic range, which have made them attractive for solving computer vision problems. Neuromorphic cameras have been used in both mono and stereo configurations for 3-D reconstruction [9], [13]. In general, 3-D reconstruction using only event data is a hard problem. Hence, neuromorphic cameras have been used in conjunction with other sensors. Rebecq et al. [14] used an Inertial Measurement Unit (IMU) along with a Dynamic Vision Sensor (DVS) to perform sparse 3-D reconstruction. By back-projecting an event and voting for the events contained in every voxel, they could detect edges with a high accuracy. Kim et al. [15] used an IMU together with a neuromorphic sensor and performed dense real-time 3-D reconstruction, six Degrees-of-Freedom (DoF) motion estimation, and intensity image reconstruction. They made use of three interleaved Extended Kalman Filters (EKFs) to estimate the depth, 6-DoF motion and image intensity using only event data. They demonstrated how Simultaneous Localization and Mapping (SLAM) could be performed more efficiently using neuromorphic cameras than frame-based ones.

This work was supported by Pratiksha Trust.

A. R. Mangalore was with the Department of Electrical Engineering and the Department of Electronic Systems Engineering, Indian Institute of Science (IISc), Bengaluru. He is presently with the Department of Electrical and Computer Engineering, Technische Universität München. C. S. Seelamantula is with the Department of Electrical Engineering. C. S. Thakur is with the Department of Electronic Systems Engineering, IISc, Bengaluru - 560012, India. E-mail: ashish.rao.m@gmail.com,{css,csthakur}@iisc.ac.in.

Some approaches have made use of structured light projection and a neuromorphic camera. Brandli et al. [16] used a projector-camera pair to project lines over a terrain and use the deformation for mapping the terrain. Matsuda et al. [17] proposed Motion Contrast 3-D scanning (MC3D) using structured light. They converted the line-scanning method into a problem that could be solved using event data, thereby eliminating redundant sampling. Leroux et al. [18] also demonstrated the feasibility of using structured light patterns with event cameras.

Each pixel of a neuromorphic camera encodes instants when the intensity changes by a preset threshold and thereby generates an asynchronous stream of non-uniformly spaced events. Reconstruction from such time-encoded measurements has been addressed by Lazar and Tóth considering the bandlimited signal model [19], and by Gontier and Vetterli for signals belonging to shift-invariant spaces [20]. The reconstruction strategies employ alternating projections similar to the one proposed by Aldroubi and Gröchenig [21]. Recently, multi-channel time-encoding strategies [22] and extensions of the reconstruction framework to handle nonbandlimited signals [23], [24] have also been developed.

A. Contributions of This Paper

We develop a Fringe Projection Profilometry (FPP) [25] system using a neuromorphic vision sensor for image capture. The term *fringe* in the standard optics literature essentially refers to a 2-D sinusoid. Conventional FPP systems employing frame-based cameras use a static fringe pattern, which is not suitable in the present setting since a neuromorphic vision sensor would not record any events for static stimuli. Hence, we use a moving fringe pattern that propagates laterally over the object being scanned. This is equivalent to performing several line-scans, all at once, with the advantage that the scan duration is shorter than a standard line-scanning method. A neuromorphic sensor is also effective at handling shadows because there aren't any temporal events corresponding to the shadows. One could therefore argue that shadows are regions in a scene that do not trigger any events. Thus, detecting shadows and eliminating their adverse influence on depth estimation is much easier in neuromorphic FPP than frame-based FPP systems.

This paper is organized as follows. We briefly review the functioning of neuromorphic cameras (Section II) before proceeding with the principle and description of the proposed neuromorphic FPP (Section III). In Section IV, we describe the experimental setup and demonstrate examples of depth scans obtained using the proposed method.

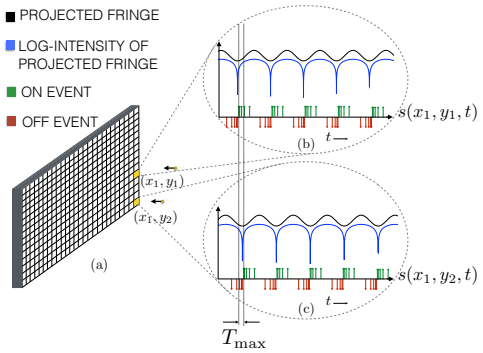

Fig. 1. (Color online) Encoding of ON and OFF events in a DAVIS346 image sensor in response to a moving fringe pattern. The simultaneous activity at two pixels is shown in (b) and (c). The sinusoid (black) corresponds to the fringe that the image sensor is exposed to and the log-intensity (blue) is the one sensed by the DVS.

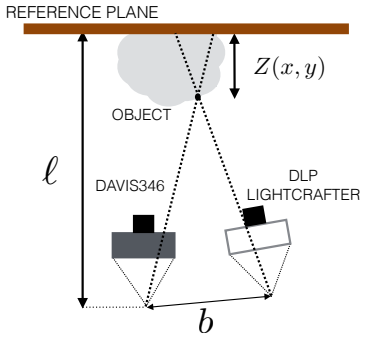

Fig. 2. The neuromorphic fringe projection profilometry setup.

II. The Principle Behind Neuromorphic Cameras

Neuromorphic cameras [12] are inspired by the magnocellular pathways in human vision [26], which specialize in detecting high-rate transient events. They are fast and asynchronous, i.e., they are triggered by events rather than an external clock. Mahowald [27] developed the *silicon retina* and invented the Address Event Representation (AER). Subsequently, Boahen [28] and Culurciello et al. [29] developed the address event arbitrated image sensor used in modern neuromorphic cameras. Lichtsteiner et al. [1], [30] developed a preliminary neuromorphic sensor called the Dynamic Vision Sensor (DVS), which is used by the ATIS [3] and DAVIS [2] cameras. The ATIS has two visual sensors: DVS, which captures events asynchronously and an APS (Active Pixel Sensor), which gives asynchronous gray-scale outputs at regions in the image sensor where the events occur. The two sensors in the ATIS operate in a coupled fashion. The DAVIS, has an IMU and two visual sensors — DVS and APS, which operate independently. While the DVS captures events, the APS captures frames at a uniform rate. In our experiments, we employed the DAVIS camera. A DVS generates an *event* whenever the following condition is satisfied:

$$\log\left(\frac{I_{t_i}(x,y)}{I_{t_{i-1}}(x,y)}\right) \underset{\text{OFF}}{\overset{\text{ON}}{\gtrless}} C, \tag{1}$$

where C is a threshold that is constant for the duration of operation, $I_{t_i}(x,y)$ is the intensity at pixel location (x,y) corresponding to an event occurring at time instant t_i. As indicated in (1), events are associated with a polarity: an "ON event" occurs when the difference in log-intensities is greater than the threshold and an "OFF event" occurs if it is below the threshold. When neither of the inequalities in (1) is satisfied, no event is generated. The DVS can also accommodate different thresholds for the ON and OFF events. Events can also occur simultaneously at multiple locations. The time stamp, spatial coordinates and polarities of the events are consolidated and represented in the AER format: $e_i^j = \left(t_i, x_i^j, y_i^j, p_i^j\right)$, where the time-stamp is t_i, the spatial coordinates corresponding to multiple locations (indexed by j) are (x_i^j, y_i^j), and p_i^j denotes the polarity of the j^{th} event associated with time-stamp t_i. The events coming out of a single pixel (x,y) over time constitute an *event-stream* $s(x,y,t)$, which at time instant t_i gives the polarity of the event p_i if an event exists at that instant, and zero otherwise. An event-stream is illustrated in Fig. 1.

III. Fringe Projection Profilometry

A. Phase Estimation

The projector-camera arrangement for fringe projection profilometry is shown in Figure 2. The projector emits a propagating fringe/sinusoidal pattern given by $f(x,y,t) = A\cos(\omega_c\, y + \omega_0\, t)$. The pattern is horizontally oriented and has a vertical spatial frequency of ω_c, and a temporal frequency of ω_0. The corresponding encoding of the ON and OFF events in a DAVIS346 image sensor is illustrated in Fig. 1. The event streams from all pixels are periodic because of the periodicity of the incident fringe pattern. Assuming noise-free operation, consider two event streams $s(x_{\text{ref}}, y_{\text{ref}}, t)$ and $s(x,y,t)$ from pixels at locations $(x_{\text{ref}}, y_{\text{ref}})$ and (x,y), respectively. The pixel $(x_{\text{ref}}, y_{\text{ref}})$ corresponds to a point in the background, which is taken as the reference. The event streams are periodic (with period $\frac{2\pi}{\omega_0}$) and time-shifted versions of one another. In order to determine the depth map, one would need the unwrapped phase, which has to be estimated from the event streams. To begin with, we compute the lag at which the cross-correlation between the event streams is maximum:

$$T_{\max}(x,y) = \arg\max_{\tau\in[0,\frac{2\pi}{\omega_0})} \langle s(x_{\text{ref}}, y_{\text{ref}}, t), s(x,y,t+\tau)\rangle_t, \tag{2}$$

where $\langle\cdot\rangle_t$ denotes temporal cross-correlation. Since the event streams are $\frac{2\pi}{\omega_0}$-periodic, the cross-correlation computation is limited to one period only. The above equation gives the time-lag relative to the fixed reference location $(x_{\text{ref}}, y_{\text{ref}})$. The corresponding phase is given by

$$\phi(x,y) = T_{\max}(x,y)\,\omega_0. \tag{3}$$

Since $T_{\max}(x,y) \in [0, \frac{2\pi}{\omega_0})$, the periodicity of the event streams gives only the wrapped phase, i.e., $\phi(x,y) \in [0, 2\pi]$. The steps are summarized in Algorithm 1 of the supplementary document. Matsuda et al. [17] showed that the pixel disparity

that one computes in a frame-based camera for measuring depth is encoded in the time difference between the events with and without the object in an event camera. In the proposed method, the depth is encoded in the phase of the periodic fringe pattern.

B. Phase Inpainting in Shadows

Light emitted by the projector that is incident on the object casts shadows on the reference plane. Shadows do not activate pixels on the neuromorphic sensor. Hence, the phase estimated in (3) would be indeterminate in regions corresponding to the shadows, which causes errors in unwrapping the phase, and consequently errors in estimating the object profile. Detecting shadows is relatively harder with frame-based cameras since both shadows and dark regions of the fringe pattern have similar grayscale intensities. With a neuromorphic camera, the problem is much easier to deal with. Over the course of the experiment, all pixels except those that correspond to shadows record events. Hence, shadows can be detected by identifying regions of no activity. We *inpaint* the phase in the shadows by taking the corresponding values from a reference phase map obtained in a calibration phase, with the fringe projected only on the reference plane, without the object in place. This process completes the phase map computation. The procedure is summarized in Algorithm 2 of the supplementary document.

C. Phase Unwrapping and Depth Calculation

The phase $\phi(x, y)$ is unwrapped by using the algorithm proposed in [31], which follows a non-continuous path based on a reliability score. This procedure produces consistent results even in the presence of discontinuities or noise. Let the unwrapped phase maps corresponding to the object and the background (used as the reference) be denoted as $\phi_u(x, y)$ and $\bar{\phi}_u(x, y)$, respectively. The depth map $Z(x, y)$ relative to the reference plane (cf. Fig. 2) can be calculated using the *triangulation principle* [32] as follows:

$$Z(x,y) = \left(\frac{\phi_u(x,y) - \bar{\phi}_u(x,y)}{\omega_c \, b} \right) \ell, \tag{4}$$

where b is the separation between the centers of the camera and the projector and ℓ is the distance between the center of the camera and the reference plane.

IV. Experimental Results

A. Generation of a Moving Fringe Pattern

We used OpenCV to generate a static fringe pattern containing 10 cycles in one frame. The vertical frequency of the fringe is limited by the resolution of the camera. The DAVIS346 camera used in our experiments has a resolution of 346×260 pixels. We generated a video of a moving fringe pattern from the static one by rolling it over in a circular fashion with one-row-shift downward per frame. The corresponding vertical frequency is given as $\omega_0 = \dfrac{v_{\text{scan}} \, \omega_c}{h}$, where h is the number of rows in the pattern and v_{scan} is the scanning speed in rows per second. The video was generated at 20 frames/second, and one-row-shift downward, which corresponds to a scanning speed $v_{\text{scan}} = 20$ rows/sec.

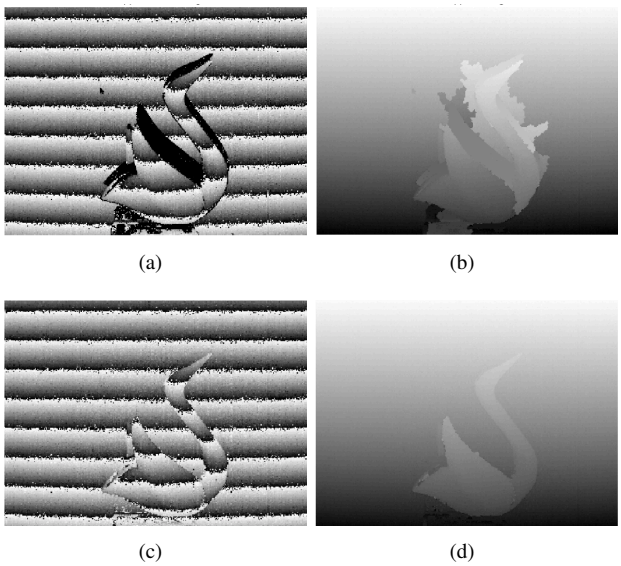

Fig. 3. Illustrating the effect of shadow-inpainting: (a) Wrapped phase without shadow inpainting; (b) the corresponding unwrapped phase; (c) Inpainted wrapped phase; and (d) the corresponding unwrapped phase.

B. Experimental Setup

The experimental setup (cf. Fig. 2) consists of a DAVIS346 camera and a Texas Instruments Digital Light Projector (DLP) LightCrafter 4500 projector module arranged in a stereo setup separated by a baseline distance b. Light projected on the DLP's micromirror array gets reflected into a lens to generate an image on the screen. The reference plane is at a distance ℓ from the camera. To begin with, the moving fringe pattern was projected onto the reference plane without the object. This would allow us to compute the reference phase that is needed for phase estimation and shadow inpainting. The unwrapped reference phase is denoted by $\bar{\phi}_u(x, y)$. We then placed an object in front of the reference plane and projected the moving fringe pattern by operating the DLP LightCrafter module in the video mode. The DAVIS346 camera can output three types of simultaneous measurements: (i) asynchronous event data; (ii) grayscale or color images; and (iii) IMU data at predefined rates. We recorded only the ON events for reconstructing the phase although one could also use the OFF events for the purpose.

C. Data Acquisition from DAVIS346

There exist several software frameworks to acquire data from the DAVIS camera, some well-known ones being `jAER` [33] and `libcaer` [34] libraries, which are publicly available. We used the `rpg_dvs_ros` package [1], [2], [35], which is based on `libcaer`. The driver of `rpg_dvs_ros` is equipped with a graphical user interface to enable setting appropriate biases for the experiment. Since `rpg_dvs_ros` is integrated into the Robot Operating System (ROS), packages for camera calibration are readily available. Both the DVS and the APS use the same lens system and image sensor, which allows one to use the camera calibration packages [36] designed for frame-based cameras to calibrate the DVS as well.

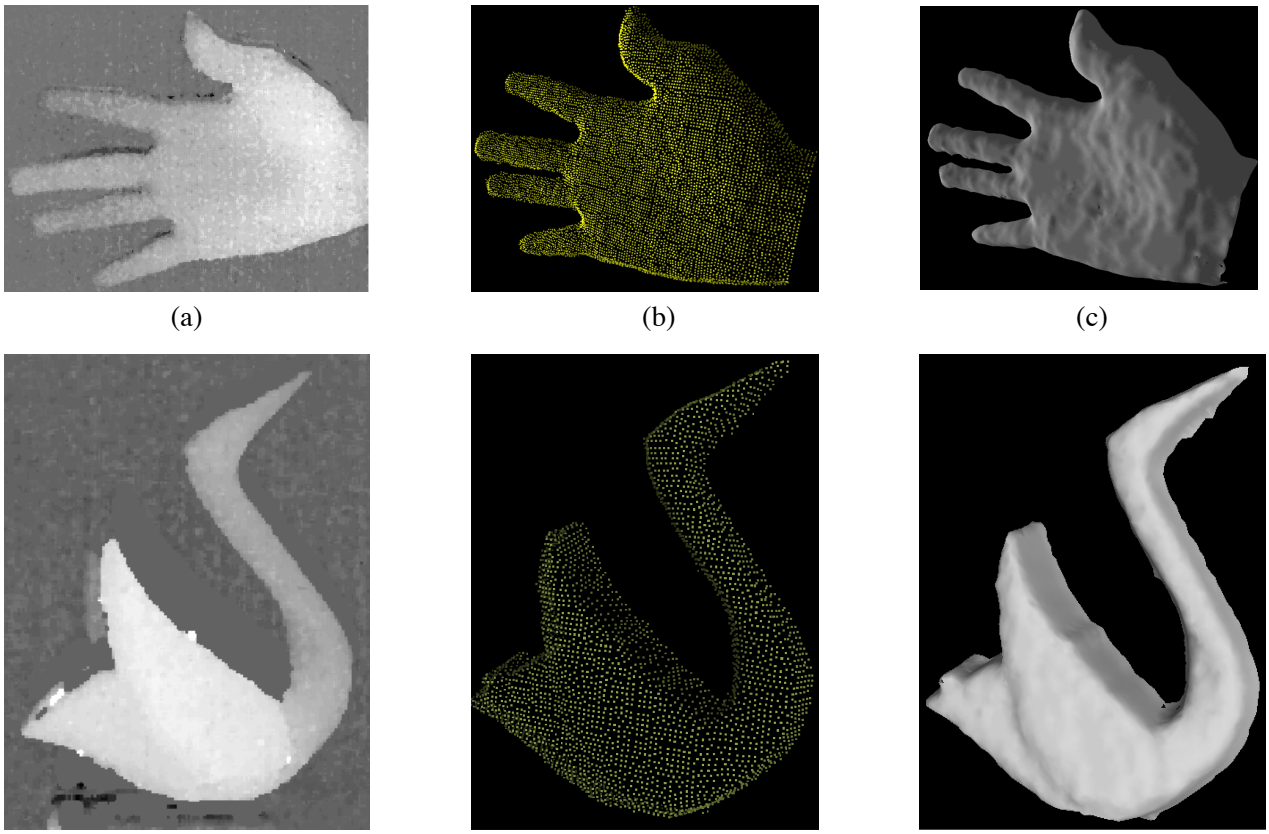

Fig. 4. Examples of scans obtained using the proposed FPP system: (a) and (d) show the depth maps obtained from the triangulation process outlined in Section III-C; (b) and (e) show the corresponding 3-D point clouds; and (c) and (f) show the corresponding meshes.

D. Results

We used the figurine of a *swan* as the object. The wrapped phase-map obtained and its unwrapped counterpart are shown in Figs. 3(a) and 3(b), respectively. Since there is a shadow, the phase-unwrapping algorithm gives an erroneous estimate. The shadow inpainted following the approach described in Sec. III-B is shown in Fig. 3(c). The resulting unwrapped phase shown in Fig. 3(d) is more accurate. In practice, the event data has noise, but we found that the phase-unwrapping algorithm was quite robust. The depth-maps were obtained following the triangulation procedure explained in Sec. III-C. Similarly, a depth-map was obtained for the *right hand* of the first author of this paper. The depth-maps were post-filtered using a 5×5 median filter to suppress outliers. The final depth-maps are shown in Figs. 4(a) and 4(d). The background was removed by a thresholding operation based on the histogram of the estimated depth values. The resulting point clouds are shown in Figs. 4(b) and 4(e). The point cloud obtained was Poisson-disk sampled [37], and the ball-pivoting surface reconstruction algorithm [38] in MeshLab [39] was used to generate the 3-D meshes shown in Figs. 4(c) and 4(f). This experiment demonstrates that using only the event data suffices to perform reliable 3-D scanning. Our Python code to obtain the point cloud from the event streams is available at [40].

V. Conclusions

We have developed a new technique to perform fringe projection profilometry using a neuromorphic sensor. Our approach relies only on the event stream output by the sensor to reconstruct the surface profile of an object. The proposed technique is faster than a line-scanning method since it is equivalent to performing several simultaneous line-scans. The factor of improvement in the time taken to scan the object would only be limited by the spatial frequency of the pattern. However, there is a limit on the number of events that the DVS can output per second, which is determined by the hardware. Operating beyond the limit would lead to some events being dropped. Working with event-stream data also has the added advantage that it is robust to handling shadows, which prevents ambiguities during phase unwrapping. We have demonstrated a proof of concept by scanning objects and reconstructing the corresponding surface profiles. The results show that only the event data suffices to perform reliable 3-D scanning. The resolution of the 3-D reconstruction is limited by the resolution of the neuromorphic sensor array, which is currently much lower than that of frame-based cameras. However, with rapid advances taking place in neuromorphic sensor development, the resolution is poised to increase in the years to come, which holds great promise for superior-quality reconstruction.

REFERENCES

[1] P. Lichtsteiner, C. Posch, and T. Delbruck, "A 128×128 120 dB 15μs latency asynchronous temporal contrast vision sensor," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 2, pp. 566–576, 2008.

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

[3] 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, 2011.

[4] T. Stoffregen and L. Kleeman, "Event cameras, contrast maximization and reward functions: An analysis," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[5] J. Manderscheid, A. Sironi, N. Bourdis, D. Migliore, and V. Lepetit, "Speed invariant time surface for learning to detect corner points with event-based cameras," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[6] H. Rebecq, R. Ranftl, V. Koltun, and D. Scaramuzza, "Events-to-video: Bringing modern computer vision to event cameras," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[7] I. Alonso and A. C. Murillo, "EV-SegNet: Semantic segmentation for event-based cameras," in *Proc. IEEE Intl. Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2019.

[8] S. Pini, G. Borghi, R. Vezzani, and R. Cucchiara, "Video synthesis from intensity and event frames," in *Proc. Image Analysis and Processing (ICIAP)*, 2019, pp. 313–323.

[9] A. Z. Zhu, L. Yuan, K. Chaney, and K. Daniilidis, "Unsupervised event-based learning of optical flow, depth, and egomotion," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[10] B. R. Pradhan, Y. Bethi, S. Narayanan, A. Chakraborty, and C. S. Thakur, "N-HAR: A neuromorphic event-based human activity recognition system using memory surfaces," in *Proc. IEEE Intl. Symp. on Circuits and Systems (ISCAS)*, 2019, pp. 1–5.

[11] A. Lakshmi, A. Chakraborty, and C. S. Thakur, "Neuromorphic Vision: Sensors to Event-based Algorithms," *Wiley-Wires Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 9, no. 4, pp. e1310, 2019.

[12] G. Gallego, T. Delbruck, G. Orchard, C. Bartolozzi, B. Taba, A. Censi, S. Leutenegger, A. Davison, J. Conradt, K. Daniilidis, and D. Scaramuzza, "Event-based vision: A survey," Apr. 2019, arXiv:1904.08405 [Online]. Available: https://arxiv.org/abs/1904.08405.

[13] Y. Zhou, G. Gallego, H. Rebecq, L. Kneip, H. Li, and D. Scaramuzza, "Semi-dense 3D reconstruction with a stereo event camera," in *Proc. European Conf. on Computer Vision (ECCV)*, 2018, pp. 235–251.

[14] H. Rebecq, G. Gallego, E. Mueggler, and D. Scaramuzza, "EMVS: Event-based multi-view stereo — 3D reconstruction with an event camera in real-time," *Intl. Journal of Computer Vision*, vol. 126, no. 12, pp. 1394–1414, 2018.

[15] H. Kim, S. Leutenegger, and A. J. Davison, "Real-time 3D reconstruction and 6-DoF tracking with an event camera," in *Proc. European Conf. on Computer Vision (ECCV)*, pp. 349–364.

[16] C. Brandli, T. Mantel, M. Hutter, M. Höpflinger, R. Berner, R. Siegwart, and T. Delbruck, "Adaptive pulsed laser line extraction for terrain reconstruction using a dynamic vision sensor," *Frontiers in Neuroscience*, vol. 7, pp. 275, 2014.

[17] N. Matsuda, O. Cossairt, and M. Gupta, "MC3D: Motion Contrast 3D Scanning," in *Proc. IEEE Intl. Conf. on Computational Photography (ICCP)*, 2015, pp. 1–10.

[18] T. Leroux, S. H. Ieng, and R. Benosman, "Event-based structured light for depth reconstruction using frequency tagged light patterns," Nov. 2018, arXiv:1811.10771 [Online]. Available: https://arxiv.org/abs/1811.10771.

[19] A. A. Lazar and L. T. Tóth, "Perfect recovery and sensitivity analysis of time encoded bandlimited signals," *IEEE Transactions on Circuits and Systems*, vol. 51, no. 10, pp. 2060–2073, 2004.

[20] D. Gontier and M. Vetterli, "Sampling based on timing: Time encoding machines on shift-invariant subspaces," *Appl. Comput. Harmon. Anal.*, vol. 36, no. 1, pp. 63–78, 2014.

[21] A. Aldroubi and K. Gröchenig, "Nonuniform sampling and reconstruction in shift-invariant spaces," *SIAM Review*, vol. 43, no. 4, pp. 585–620, 2001.

[22] K. Adam, A. Scholefield, and M. Vetterli, "Sampling and reconstruction of bandlimited signals with multi-channel time encoding," *arXiv preprint arXiv:1907.05673*, 2019.

[23] R. Alexandru and P. L. Dragotti, "Time-based sampling and reconstruction of non-bandlimited signals," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2019, pp. 7948–7952.

[24] R. Alexandru and P. L. Dragotti, "Time encoding and perfect recovery of non-bandlimited signals with an integrate-and-fire system," in *Proc. Int. Conf. Sampling Theory and Applications (SampTA)*, 2019.

[25] V. Srinivasan, H. C. Liu, and M. Halioua, "Automated phase-measuring profilometry of 3D diffuse objects," *Appl. Opt.*, vol. 23, no. 18, pp. 3105–3108, 1984.

[26] A. H. C. van der Heijden, *Selective Attention in Vision*, Routledge, 1992.

[27] M. Mahowald, *An Analog VLSI System for Stereoscopic Vision*, Boston, MA: Springer US, 1994, pp. 4–65.

[28] K. Boahen, "A throughput-on-demand address-event transmitter for neuromorphic chips," in *Proc. 20th Anniversary Conf. on Advanced Research in VLSI*, 1999, pp. 72–86.

[29] E. Culurciello, R. Etienne-Cummings, and K. Boahen, "Arbitrated address event representation digital image sensor," in *Proc. IEEE Intl. Solid-State Circuits Conf.*, 2001, pp. 92–93.

[30] P. Lichtsteiner and T. Delbruck, "A 64×64 AER logarithmic temporal derivative silicon retina," in *Research in Microelectronics and Electronics*, 2005, pp. 202–205.

[31] M. A. Herráez, D. R. Burton, M. J. Lalor, and M. A. Gdeisat, "Fast two-dimensional phase-unwrapping algorithm based on sorting by reliability following a noncontinuous path," *Appl. Opt.*, vol. 41, no. 35, pp. 7437–7444, 2002.

[32] C. Zhang, P. S. Huang, and F. Chiang, "Microscopic phase-shifting profilometry based on digital micromirror device technology," *Appl. Opt.*, vol. 41, no. 28, pp. 5896–5904, 2002.

[33] "The jAER library, https://github.com/SensorsINI/jaer," Last Accessed: June 25, 2020.

[34] "The libcaer library, https://gitlab.com/inivation/dv/libcaer," Last Accessed: June 25, 2020.

[35] E. Mueggler, B. Huber, and D. Scaramuzza, "Event-based, 6-DoF pose tracking for high-speed maneuvers," in *Proc. IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems*, 2014, pp. 2761–2768.

[36] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.

[37] M. Corsini, P. Cignoni, and R. Scopigno, "Efficient and flexible sampling with blue noise properties of triangular meshes," *IEEE Trans. Visualization and Computer Graphics*, vol. 18, no. 6, pp. 914–924, 2012.

[38] F. Bernardini, J. Mittleman, H. Rushmeier, C. Silva, and G. Taubin, "The ball-pivoting algorithm for surface reconstruction," *IEEE Trans. Visualization and Computer Graphics*, vol. 5, pp. 349–359, 1999.

[39] P. Cignoni, M. Callieri, M. Corsini, M. Dellepiane, F. Ganovelli, and G. Ranzuglia, "MeshLab: an open-source mesh processing tool," in *Proc. Eurographics Italian Chapter Conf.*, 2008, pp. 129–136.

[40] "Python code for Neuromorphic Fringe Projection Profilometry, https://github.com/ashishrao7/NFPP," Last Accessed: June 25, 2020.