Event-Based Sensor Fusion in Human-Machine Teaming

Benjamin Noack[1], Florian Röhrbein[2], and Gunther Notni[3]

[1] Autonomous Multisensor Systems Group, Institute for Intelligent Cooperating Systems, Otto von Guericke University Magdeburg

[2] Professorship of Neurorobotics, Faculty of Computer Science, Chemnitz University of Technology

[3] Group for Quality Assurance and Industrial Image Processing, Department of Mechanical Engineering, Technical University Ilmenau

ABSTRACT

Realizing intelligent production systems where machines and human workers can team up seamlessly demands a yet unreached level of situational awareness. The machines' leverage to reach such awareness is to amalgamate a wide variety of sensor modalities through multisensor data fusion. A particularly promising direction to establishing human-like collaborations can be seen in the use of neuro-inspired sensing and computing technologies due to their resemblance with human cognitive processing. This note discusses the concept of integrating neuromorphic sensing modalities into classical sensor fusion frameworks by exploiting event-based fusion and filtering methods that combine time-periodic process models with event-triggered sensor data. Event-based sensor fusion hence adopts the operating principles of event-based sensors and even exhibits the ability to extract information from absent data. Thereby, it can be an enabler to harness the full information potential of the intrinsic spiking nature of event-driven sensors.

Index Terms **–** Sensor Fusion, Neuromorphic Sensors, Event-Based Filtering

1. INTRODUCTION

Neuromorphic sensors are an emerging technology promising low latency, high performance, and better resource utilization as compared to traditional sensors. With dynamic vision sensors [1], the first neuromorphic cameras have been emerging as commercial products, which asynchronously output per-pixel brightness changes as lists of events. By mimicking the function of the human retina, these sensors have the potential to excel in highly dynamic contexts, making them a promising technology for a wide range of applications. The purpose of this note is to emphasize the importance of utilizing event-based sensors and data fusion in production systems that involve human-machine teaming. By leveraging the unique capabilities of event-based sensors and fusing their data with other sensor inputs, such systems can achieve comprehensive real-time situational awareness, faster response times, and improved decision-making [2]. These benefits are especially critical in complex, high-pressure environments where humans and machines must work together seamlessly to achieve common goals.
The fusion of neuromorphic and traditional sensor outputs requires algorithms that can incorporate synchronous and asynchronous sensor data simultaneously. Conventional

approaches to multisensor data fusion operate in discrete time at fixed rates and demand significant adaptation to process asynchronous sensor data effectively. Another direction can be deploying neuromorphic computing hardware for processing and fusing sensor data. However, such hardware is still under development and may face challenges in processing the large amounts of data generated by frame-based vision sensors. For this reason, the considered concepts focus on multisensor data processing implemented on clock-driven von Neumann computers by utilizing event-based filtering methods [3].
The following sections first discuss the benefits of event-based sensor modalities and sensor fusion architectures, in general, and then proceed to illustrate how these benefits are particularly effective when it comes to enhancing collaboration between humans and machines.

2. POTENTIALS OF EVENT-BASED SENSORS AND SENSOR FUSION

Neuromorphic sensing aims to depart from the conventional concept of periodic sampling and instead embraces an event-driven sensing scheme that closely mimics the sampling process in the nervous system [4]. In this approach, events are primarily generated when there is a change in the sensed stimulus. Fig. 1 compares time-periodic sampling with an event-based scheme that only triggers when the difference between the current input and the last event exceeds a threshold Δ.

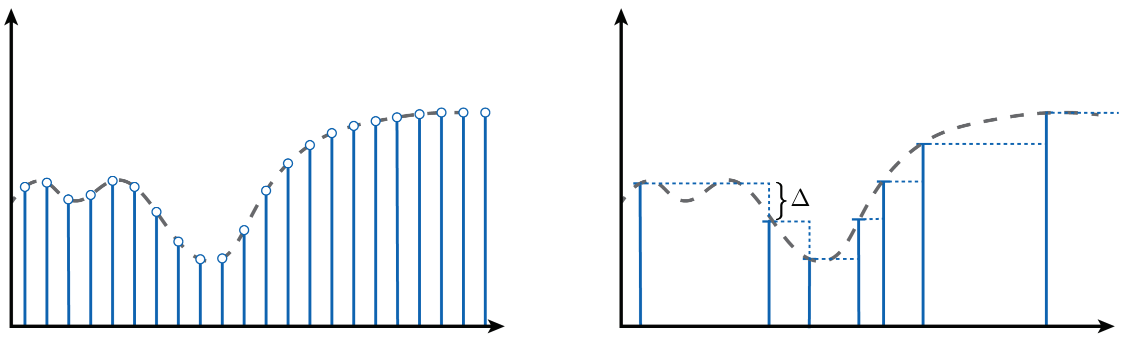

Fig. 1. Left: Time-periodic sampling. Right: Event-based sampling with a send-on-delta scheme.

A prominent example of a sensor system implementing this principle is event-based neuromorphic cameras, which bring numerous opportunities, especially in highly dynamic settings. However, to capitalize on their advantages, several challenges must be addressed, which is reflected by the increasing amount of research in recent years dedicated to exploring the potential of event-based sensing. This section summarizes the potential benefits of event-based sensor systems and data processing with a focus on human-machine environments.

2.1 Event-Based Sensing: What to Expect?

The advantages of, in particular, event-based vision can be better understood by comparing it to how the human eye and brain process visual information. Unlike a frame-based camera, our eyes process images as a continuous stream, allowing us to recognize objects based purely on motion. This was unveiled in Johannson's experiment [5], which demonstrates that humans or animals and their actions can be identified through a few moving dots. The same principle applies to event-based sensors. The frame-based nature of conventional cameras providing periodic temporal snapshots is replaced by aperiodic event streams enabling blur-free perception. In event cameras [6], brightness changes in the scene independently and asynchronously trigger events for each pixel, which can be captured with an extremely high temporal resolution. Event data can reach a temporal accuracy of microseconds and can be output with sub-millisecond latency [1]. Besides the robustness to motion blur, event-based

sensing hence allows fast response times making it well-suited for applications requiring rapid decision-making and feedback like robotics, autonomous driving, and human-machine collaboration. The pixel-wise responses also lead to a high dynamic range of above 120 dB making event cameras much less susceptible to challenging lighting conditions and variations. There are additional benefits to event-based sensing, such as the sparse data representation and reduced power consumption. Sparse data representation is made possible by only detecting and processing changes instead of complete frames. Therefore, utilizing event-based sensing leads to a natural reduction in redundancy. This, in turn, enhances the efficiency of storage, transmission, and processing while also facilitating faster data processing and reducing computational requirements. Reduced power consumption is also tied to avoiding the capturing and processing of entire frames. Hence, event-based sensing is well-suited for energy-constrained devices and applications. Additionally, event-based sensing promotes enhanced privacy since the transmission and processing of only relevant event information minimize the exposure of sensitive or unnecessary data, reducing potential privacy risks.
Although event-based cameras are the most prominent instance of neuromorphic sensors, there are other sensing modalities currently being developed. Examples are tactile sensors [7] that translate contact forces to spike trains and neuromorphic acoustic sensors [8] mimicking the functioning of the cochlea. Current results report similar advantages over traditional sensor systems, like low latency, high dynamic range, and energy efficiency. Event-based sensing hence constitutes a valuable addition to existing sensing modalities in human-machine environments.

2.2 Event-Based Multisensor Integration

Event-based sensing naturally promotes data processing on neuromorphic hardware and software that promise the same previously mentioned benefits, like low power consumption. Yet, they are highly specialized, and research in the field continues to expand their capabilities and applications. In this note, we focus on sensor integration using synchronous and asynchronous sensing and processing schemes executed on conventional von Neumann architectures. For camera data, this typically implies buffering events and processing them in batches, thus artificially creating frames from the events [9]. Still, applications can benefit from the sparse data representation, high dynamic range, and low latency. These advantages can also be leveraged to enhance the quality and efficiency of time-periodic sensor data processing.
Several studies showcase the potential benefits of utilizing both frame-based and event-based camera systems in a complementary manner. For example, the approach described in [10] employs two parallel feature extractor networks applied to frames and events, respectively. To implement sensor fusion, the results are concatenated and fed into a feature pyramid network. In doing so, the authors achieve robust object detection under adverse conditions like fog, snow, and varying illumination. Here, events are collected in temporal bins, which are translated into voxel representations, where the first two dimensions are the event pixel position and the third represents the time. Another example is efficient feature tracking, as proposed in [11]. Image frames from a standard camera are used to extract features, while the event streams allow tracking and associating the features with the next frame. Hence with the integration of event-based sensors, their advantages, like robustness and high temporal resolution, can be inherited, and the overall performance increases.

2.3 Event-Based Sensor Fusion Architectures

Synchronous sensor data processing, state estimation, and control can adopt event triggers to determine which data should be transmitted and processed. In event-triggered control [12], this implies executing the control task only when necessary. This approach can be more efficient than traditional time-periodic control as it reduces the number of task executions while maintaining the desired closed-loop performance. Similarly, event-based state estimation [3], [13], typically employs a trigger at the sensor to decide for transmission in order to find a trade-off between estimation quality and resource utilization in networked sensor and control systems. By implementing this estimator design, versatile sensor fusion frameworks can be realized that allow for seamless integration of both synchronous and asynchronous sensors and estimators. Periodic estimators can, for example, be extended to incorporate out-of-schedule measurements that are triggered by event-based sensors. One option is to activate the estimator on incoming events; another option is to map event-triggered measurements to the processing times of the estimator. The latter can use the underlying state-space model to synchronize the measurement information with the estimator.

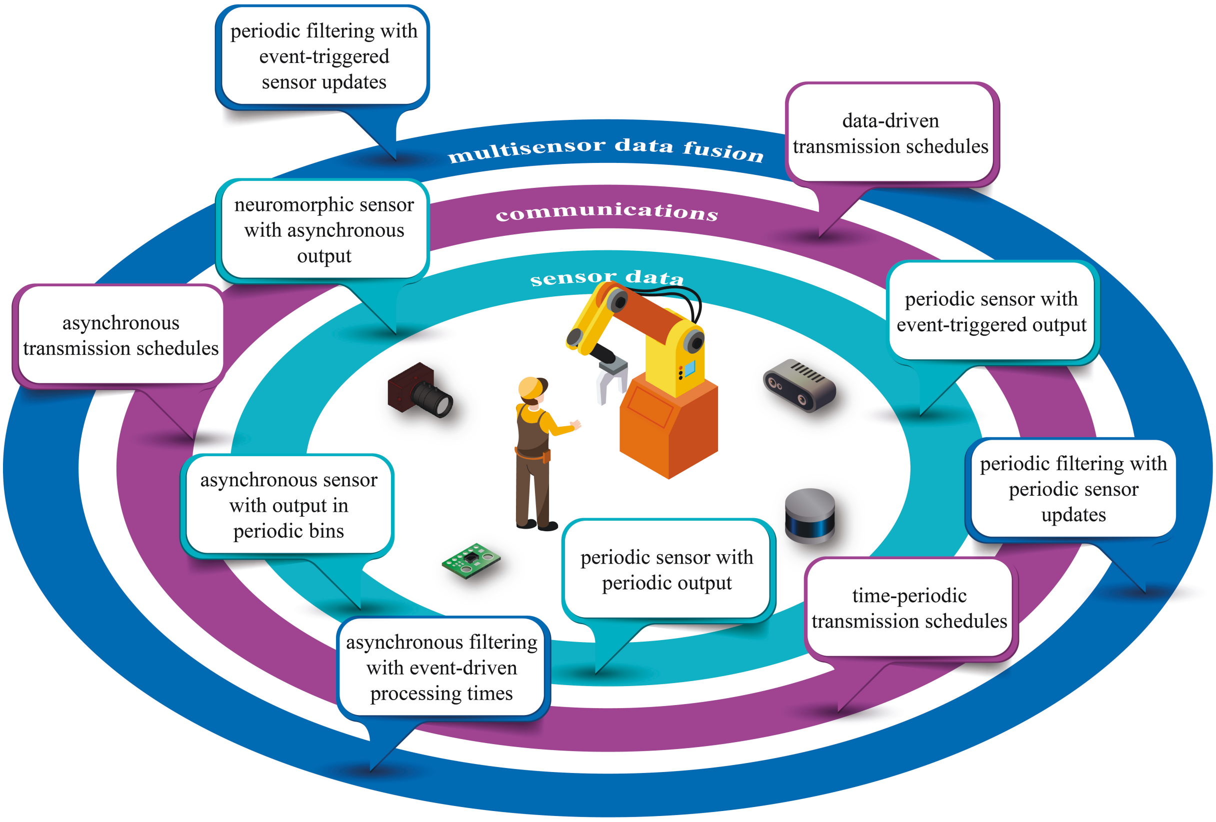

Fig. 2. Existing fusion architectures can integrate event-based technologies and methods implementing arbitrary combinations of event-based and period sensor outputs, transmission schedules, and data fusion methods.

The deployment of event-based sensors, transmissions, and fusion algorithms should not be viewed as a replacement of existing systems and architectures. Rather, event-based technologies can integrate into the existing fusion architectures to enhance their functionality and effectiveness. As illustrated in Fig. 2, different combinations of event-based and periodic systems and methods can be implemented. For instance, a sensor that is time-periodic can be equipped with an event trigger to determine when to output readings, resembling a neuromorphic sensor. Similarly, such a trigger could be implemented in the communications system to address communication costs and constraints [14]. As stated before, a time-perioding fusion algorithm can be additionally triggered by an asynchronous sensor output [13] or could use filters tailored to out-of-sequence measurements [15]. In summary, frameworks that incorporate both periodic and event-triggered approaches can be very beneficial, in particular, for collaborative human-machine environments.

3. SEAMLESS HUMANE-MASCHINE TEAMING WITH EVENT-BASED FUSION

The numerous benefits outlined in the previous section make event-based sensors and data processing a valuable addition to human-machine production systems. Low latency and robustness to motion blur contribute significantly to ensuring the safety of human workers by allowing machines to react quickly and smoothly. By promoting resource efficiency, event-based units can be deployed in distributed environments equipped with wireless communications and battery-driven nodes. We scrutinize these potentials from the perspective of human-machine cooperation in the following.

3.1 Event-Based Object Tracking

Vision-based tracking systems serve as indispensable components of a human-machine collaborative production environment as they enable accurate monitoring and tracking of human actions and machine movements [16], facilitating efficient and safe collaboration between humans and machines. As stated in the introduction, conventional multi-camera tracking environments can greatly benefit from incorporating complementary neuromorphic vision systems. In order to process their data in a synchronous fashion, the fixed-rate tracking algorithms need to map the event data to the processing time steps, as already indicated in Section 2.2. To address this issue, the approach from [3] and [17] can be adopted, which employs motion models of the tracked targets to map the pixel-based events to the considered time step. If the event-to-target association is unknown, a track-before-detect approach can be followed, which involves using multiple motion models for the mapping and subsequently determining the most likely association. Multiple hypothesis tracking for event data has, for instance, been proposed in [18]. Similarly, [19] employs a probabilistic data association approach that combines probabilities that events are generated by certain points with optical flow constraints. To further improve tracking performance, the tracker can also take into account absent events. Non-moving objects relate to the absence of pixel events but may still be detected in frame-based sensors. The integration of this implicit motion information improves tracking performance. More generally, the object's motion determines where events are spiked; hence, motion information can be inferred from absent events as well. Such concepts are known as implicit information in event-based filtering [17]. As discussed in the subsequent subsection, the tracking performance of event-based systems can be leveraged in particular in human-machine systems.

3.2 Human-Machine Interaction

The techniques of event-based object tracking described in the preceding subsection are predestined to human 3D pose estimation. For instance, [20] implement a high-frequency motion capture system using event cameras. Their approach includes the generation of event frames by aggregating events over intervals. [21] propose an event-only approach to estimate the skeletal pose. Unlike high-speed cameras, these approaches bypass the need for extensive computational power and storage capacity. Further applications involve high-speed gaze tracking, utilizing event sensors to capture subtle and rapid eye movements. [22] achieve sampling rates of $10\ kHz$ such that microsaccades and sudden movements can be detected. When humans work alongside machines, this can help the machines to evaluate the attention and reactions of their human counterparts more effectively. Event sensors can also be leveraged for drowsiness detection, which is demonstrated in [23] for driver assistance systems. Accordingly, production systems can greatly benefit from using sensor systems that can detect fatigue and diminishing levels of attention to enhance the safety of human workers. In general, event-based sensors have the potential to excel in human activity recognition tasks and may surpass frame-based systems, as demonstrated in [24].

Employing event-based fusion is not only a tool to incorporate neuromorphic sensors but also to improve resource utilization in conventional sensor systems. For example, body-worn inertial measurement units may produce data at high rates. A neuro-inspired pre-processing may aid in reducing data by translating the measurements into events, e.g., by using thresholds. Event-based filters can still retrieve information from absent events. The machine can respond based on the frequency of events or use it to evaluate the speed at which a human partner reacts, among other possibilities. In conclusion, event-based approaches show great potential for enabling comprehensive real-time situational awareness and offering new ways to process and interpret data.

3.3 Event-Based Multisensor Data Fusion for Human-Machine Teaming

Within networked estimation architectures, the use of event-based sensing and communication has the potential to enhance resource allocation and optimize communication resource utilization. As indicated in Section 2.3 and illustrated in Fig. 3, each stage of a control loop in a human-machine system can have an event-triggering mechanism, where communication and processing systems can utilize the principles of event-based sensing. For instance, body-worn, battery-driven sensor nodes can benefit from event-triggered transmission schedules for saving energy. Simple decision rules are sent-on-delta rules to implement event-triggered transmissions, which compare the difference between the current and last transmitted measurement against a user-defined threshold [25]. The sensor node hence assesses the innovation with respect to the previous transmission to decide on sending its current measurements or estimates. Intelligent sensors can keep track of the system's state to predict future sensor readings and can thereby further decrease transmission rates [26]. For example, kinematic models can be used to predict movements and only deviations from the predicted trajectory trigger new events. Instead of hard thresholds, [27] proposes to use stochastic trigger rules that ease the exploitation of implicit information at the receiver. Implicit information is the concept of extracting information from the trigger rule in the absence of transmissions: At the receiver, the absence of a transmission still carries useful information about the sensor data, e.g., that the innovation falls below the trigger threshold. For example, a body-worn sensor network could trigger transmission only when a certain stress level is indicated. The remote fusion system can infer from the absence of transmissions that the stress level is below the triggering threshold. In summary, event-based mechanisms offer new ways of processing and interpreting data and designing fusion architectures in human-machine teams.

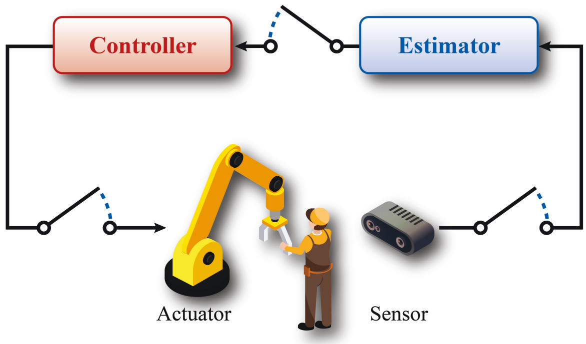

Fig. 3. A human-machine control loop can implement event triggers at different components.

4. CONCLUSIONS

Event-based sensors are an emerging technology that shows great potential for making human-machine collaboration safer and more seamless. By integrating these technologies and processing principles into existing production environments, one can reap the benefits of both conventional and aperiodic, event-based systems. The entire processing chain can adopt an event-triggering mechanism to decide when to communicate or process information, thereby reducing the load on the network and processing infrastructure. Architectures that integrate event-based sensors, communications, filtering, fusion, and control enable a tight and effective coupling of human and machine work.

REFERENCES

[1] P. Lichtsteiner, C. Posch and T. Delbruck, "A 128 x 128 120dB 15µ Latency Asynchronous Temporal Contrast Vision Sensor," *IEEE Journal of Solid-State Circuits,* vol. 43, p. 566–576, 2008.

[2] G. Chen, H. Cao, J. Conradt, H. Tang, F. Röhrbein and A. Knoll, "Event-Based Neuromorphic Vision for Autonomous Driving: A Paradigm Shift for Bio-Inspired Visual Sensing and Perception," *IEEE Signal Processing Magazine,* vol. 37, p. 34–49, July 2020.

[3] J. Sijs and M. Lazar, "Event Based State Estimation With Time Synchronous Updates," *IEEE Transactions on Automatic Control,* vol. 57, p. 2650–2655, October 2012.

[4] C. Bartolozzi, A. Glover and E. Donati, "Neuromorphic Sensing, Perception and Control for Robotics," in *Handbook of Neuroengineering*, Springer Singapore, 2021, p. 1–31.

[5] G. Johansson, "Visual Perception Of Biological Motion And A Model For Its Analysis," *Perception & Psychophysics,* vol. 14, p. 201–211, June 1973.

[6] G. Gallego, T. Delbruck, G. Orchard, C. Bartolozzi, B. Taba, A. Censi, S. Leutenegger, A. J. Davison, J. Conradt, K. Daniilidis and D. Scaramuzza, "Event-Based Vision: A Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 44, p. 154–180, January 2022.

[7] T. Birkoben, H. Winterfeld, S. Fichtner, A. Petraru and H. Kohlstedt, "A Spiking And Adapting Tactile Sensor For Neuromorphic Applications," *Scientific Reports,* vol. 10, October 2020.

[8] C. Lenk, P. Hövel, K. Ved, S. Durstewitz, T. Meurer, T. Fritsch, A. Männchen, J. Küller, D. Beer, T. Ivanov and M. Ziegler, "Neuromorphic Acoustic Sensing Using an Adaptive Microelectromechanical Cochlea with Integrated Feedback," *Nature Electronics,* vol. 6, p. 370–380, May 2023.

[9] G. Lenz, Neuromorphic Algorithms and Hardware for Event-based Processing, Ph.D. Thesis, Sorbonne Université, 2021.

[10] A. Tomy, A. Paigwar, K. S. Mann, A. Renzaglia and C. Laugier, "Fusing Event-Based and RGB Camera for Robust Object Detection in Adverse Conditions," in *Proceedings of the 2022 International Conference on Robotics and Automation (ICRA)*, Philadelphia, 2022.

[11] Y. Dong and T. Zhang, "Standard and Event Cameras Fusion for Feature Tracking," in *2021 International Conference on Machine Vision and Applications*, 2021.

[12] W. P. M. H. Heemels, K. H. Johansson and P. Tabuada, "Event-Triggered and Self-Triggered Control," in *Encyclopedia of Systems and Control*, Springer London, 2013, p. 1–10.

[13] J. Sijs, B. Noack, M. Lazar and U. D. Hanebeck, "Time-Periodic State Estimation with Event-Based Measurement Updates," in *Event-Based Control and Signal Processing*, M. Miskowicz, Ed., CRC Press, 2018.

[14] F. Rosenthal, B. Noack and U. D. Hanebeck, "Scheduling of Measurement Transmission in Networked Control Systems Subject to Communication Constraints," in *Proceedings of the 2018 American Control Conference (ACC 2018)*, Milwaukee, 2018.

[15] Y. Bar-Shalom, H. Chen and M. Mallick, "One-Step Solution for the Multistep Out-of-Sequence-Measurement Problem in Tracking," *IEEE Transactions on Aerospace and Electronic Systems,* vol. 40, p. 27–37, January 2004.

[16] Y. Zhang, S. Müller, B. Stephan, H.-M. Gross and G. Notni, "Point Cloud Hand–Object Segmentation Using Multimodal Imaging with Thermal and Color Data for Safe Robotic Object Handover," *Sensors,* vol. 21, p. 5676, August 2021.

[17] E. J. Schmitt, B. Noack, W. Krippner and U. D. Hanebeck, "Gaussianity-Preserving Event-Based State Estimation With an FIR-Based Stochastic Trigger," *IEEE Control Systems Letters,* vol. 3, p. 769–774, July 2019.

[18] I. Alzugaray and M. Chli, "Asynchronous Multi-Hypothesis Tracking of Features with Event Cameras," in *2019 International Conference on 3D Vision (3DV)*, 2019.

[19] A. Z. Zhu, N. Atanasov and K. Daniilidis, "Event-Based Feature Tracking with Probabilistic Data Association," in *2017 IEEE International Conference on Robotics and Automation (ICRA 2017)*, 2017.

[20] S. Zou, C. Guo, X. Zuo, S. Wang, P. Wang, X. Hu, S. Chen, M. Gong and L. Cheng, "EventHPE: Event-Based 3D Human Pose and Shape Estimation," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021.

[21] G. Scarpellini, P. Morerio and A. D. Bue, "Lifting Monocular Events to 3D Human Poses," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2021.

[22] A. N. Angelopoulos, J. N. P. Martel, A. P. Kohli, J. Conradt and G. Wetzstein, "Event-Based Near-Eye Gaze Tracking Beyond 10,000 Hz," *IEEE Transactions on Visualization and Computer Graphics,* vol. 27, p. 2577–2586, May 2021.

[23] G. Chen, L. Hong, J. Dong, P. Liu, J. Conradt and A. Knoll, "EDDD: Event-Based Drowsiness Driving Detection Through Facial Motion Analysis With Neuromorphic Vision Sensor," *IEEE Sensors Journal,* vol. 20, p. 6170–6181, June 2020.

[24] F. J. Moreno-Rodríguez, V. J. Traver, F. Barranco, M. Dimiccoli and F. Pla, "Visual Event-Based Egocentric Human Action Recognition," in *Pattern Recognition and Image Analysis*, Springer International Publishing, 2022, p. 402–414.

[25] J. Sijs, B. Noack and U. D. Hanebeck, "Event-Based State Estimation with Negative Information," in *Proceedings of the 16th International Conference on Information Fusion (Fusion 2013)*, Istanbul, 2013.

[26] B. Noack, C. Öhl and U. D. Hanebeck, "Event-Based Kalman Filtering Exploiting Correlated Trigger Information," in *Proceedings of the 25th International Conference on Information Fusion (Fusion 2022)*, Linköping, Sweden, 2022.

[27] D. Han, Y. Mo, J. Wu, S. Weerakkody, B. Sinopoli and L. Shi, "Stochastic Event-Triggered Sensor Schedule for Remote State Estimation," *IEEE Transactions on Automatic Control,* vol. 60, p. 2661–2675, October 2015.

CONTACTS

Prof. Dr.-Ing. Benjamin Noack	email:	benjamin.noack@ieee.org
	ORCID:	https://orcid.org/0000-0001-8996-5738
Prof. Dr. rer. nat. Florian Röhrbein	email:	florian.roehrbein@informatik.tu-chemnitz.de
	ORCID:	https://orcid.org/0000-0002-4709-2673
Prof. Dr. rer. nat. Gunther Notni	email:	gunther.notni@tu-ilmenau.de
	ORCID:	https://orcid.org/0000-0001-7532-1560

