



FRAMEWORK PARA CONVERTER IMAGENS RGB PADRÃO EM DADOS NEUROMÓRFICOS

Resumo – Dispositivos neuromórficos são instrumentos que podem ser utilizados em diversas aplicações robóticas e computacionais a fim de solucionar e otimizar cada uma. Estas, podem ser câmeras baseadas em eventos, que com uma alta capacidade de processamento, podem detectar variações de luminosidade em uma cena, porém estas câmeras, para muitos grupos de pesquisa ainda é muito cara, tornando-a inacessível. Dessa forma, propomos um framework capaz de converter informações advindas de câmeras RGB padrão em dados neuromórficos. Com isso, os horizontes do campo da visão neuromórfica poderão ser ampliados, podendo os pesquisadores desenvolver novos algoritmos neuromórficos a partir de datasets da visão computacional, sendo esses convertidos através do framework criado. Nós obtivemos resultados similares, comparando informações convertidas pelo framework com uma câmera neuromórfica usual.

Palavras-Chave – Câmeras, Eventos, Framework, Neuromórfico, Software.

FRAMEWORK TO CONVERT STANDARD RGB IMAGES IN NEUROMORPHIC DATA

Abstract – Neuromorphic devices are instruments that can be used in several robotic and computational applications in order to solve and optimize each one. These can be event-based cameras, which with a high processing capacity, for many research groups, are still very expensive, making it inaccessible. Thus, we propose a framework capable of converting information from standard RGB cameras into neuromorphic data. With that, the horizons of the field of neuromorphic vision can be expended, and researchers can develop new neuromorphic algorithms from computer vision datasets, which are converted through the created framework. We obtained similar results, comparing information converted by the framework with a usual neuromorphic camera.

Keywords – Cameras, Events, Framework, Neuromorphich, Software.

I. INTRODUÇÃO

O processo de desenvolvimento de modelos computacionais baseados em sistemas neuromórficos ocorreu, principalmente, a partir das tecnologias bioinspiradas. Sendo essas, abstrações de atividades realizadas pela natureza para solução e otimização de problemas tecnológicos [1]. Dando início então à diversos algoritmos e tecnologias que encontramos hoje, como as redes neurais artificiais, as redes convolutivas, algoritmo genético e muitos outros [2]. Dessa forma, os dispositivos neuromórficos também fazem parte desse grupo, pois são inspirados em sistemas biológicos [3] [4].

Assim umas das grandes vantagens de trabalhar com sistemas neuromórficos em sua grande eficiência computacional e energética [5] [6] [7]. Sua concepção neuro-inspirada, tem por objetivo alcançar a sofisticação do processamento biológico, que possui um baixo consumo de energia e uma alta velocidade de processamento das informações [4] [8] [9]. Consequentemente, o foco em desenvolver sistemas neuromórficos têm se tornado estado da arte de grandes empresas de tecnologia, podendo então ser a próxima geração de inteligência artificial [6] [10] [11].

Um dos dispositivos neuromórficos que têm ganhado atenção são as câmeras baseadas em eventos [9] [11] [12] [13]. Esses dispositivos possuem um funcionamento semelhante ao da retina humana, e têm vantagens em relação as câmeras tradicionais [14]. Cada pixel possui um funcionamento independente e assíncrono, as informações são transmitidas com alta resolução temporal e baixa latência [15]. Essas características colaboram para uma grande vantagem das câmeras neuromórficas em relação as câmeras tradicionais, sobretudo em tarefas como detecção e rastreamento de movimento [16] [17].

Sendo esse um novo campo de estudo a ser explorado, torna-se cada vez mais necessário facilitar o acesso de pesquisadores a dados neuromórficos, para que os trabalhos realizados no campo de visão neuromórfica alcancem maturidade suficiente para que os novos algoritmos desenvolvidos possam ser comparados entre si.

Portanto, com a finalidade de contribuir com o desenvolvimento de algoritmos neuromórficos, foi criado um framework que transforma as informações captadas por uma câmera RGB padrão em informações neuromórficas. O software foi construído tendo em vista os aspectos de

*lucas.vasconcelos@ufu.br

funcionamento de uma câmera neuromórfica, gerando informações similares as informações de uma câmera baseada em eventos nos pixels onde houve uma significativa variação de luminosidade [7] [9] [12] [14]. Logo, grupos de pesquisa na área podem se beneficiar e desenvolver novos softwares e algoritmos, utilizando as informações convertidas por esse framework, mesmo que não tenham acesso a uma câmera verdadeiramente neuromórfica.

II. MATERIAS E MÉTODOS

A. Dispositivos Neuromórficos

Em dispositivos neuromórficos, cada evento é gerado de maneira assíncrona, esse processo é feito através de um circuito comparador de luminosidade recebida nos fotodiodos. Assim, um evento é gerado quando há uma variação de luminosidade significativa em relação a luminosidade basal da cena [5] [12] [13].

O funcionamento baseado em eventos e a natureza assíncrona dos sensores de visão neuromórfico contribuem para sua alta capacidade de processamento e sua alta resolução temporal [12] [14] [18]. Ao gerar eventos apenas onde há variação de luminosidade, os sensores de visão neuromórfica eliminam informações redundantes na cena, enquanto os sensores de visão tradicionais capturam informações de todos os pixels, mesmo que não haja alteração no valor de luminosidade de determinada região.

Devido a sua natureza assíncrona, há uma ressignificação na concepção de frames nos dispositivos neuromórficos, uma vez que cada pixel dispara um evento a partir de um estímulo, ou seja, uma variação de luminosidade [12].

Os eventos que são capturados por esses sensores carregam informações primárias, relacionadas ao momento em que cada evento ocorreu, a posição e também sua intensidade, ou seja, sua variação [7] [12] [13] [14].

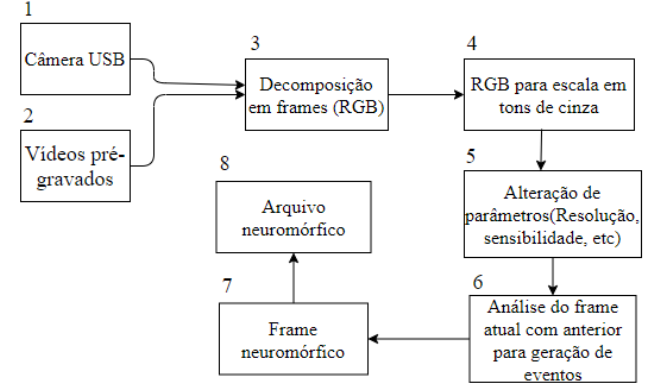
Para a realização deste trabalho, foi utilizado um dispositivo neuromórfico específico, a DVS128 (Dynamic Vision Sensor), que possui uma resolução de 128x128 pixels, sendo ela desenvolvida por outros grupos de pesquisadores [7] [9] [16].

B. Framework de Conversão

O processo de conversão de informações adquiridas de câmeras tradicionais em informações neuromórficas ocorreu através de um framework desenvolvido utilizando a linguagem Python. O software pode ser utilizado para converter informações em tempo real, usando qualquer tipo de câmera RGB comum, ou através de vídeos previamente gravados.

Assim, em nosso framework (Figura 1), a conversão de informações de sensores de visão tradicionais em dados neuromórficos é feita através da análise da variação de luminosidade em diferentes momentos, ou seja, realizando a comparação entre os frames dos dados tradicionais e convertendo-os em neuromórficos, com sua natureza binária baseada na concepção de eventos. Entretanto, a resolução temporal das informações convertidas pelo framework é diretamente ligada a taxa de transmissão de Frames Por Segundo (FPS) da câmera RGB utilizada para captação dos dados de entrada.

Figura1: Diagrama em blocos do framework desenvolvido.



A Equação 1, foi utilizada para converter as informações em dados neuromórficos, demonstrado pelo bloco 6 da Figura 1. Onde β representa o valor máximo de luminosidade possível, de acordo com a profundidade de bits. O fator α representa a sensibilidade de geração de um evento. A função $f(t)$ representa as informações convertidas, localizadas no bloco 7 da Figura 1, sendo todos eles armazenados e por conseguinte originando o arquivo neuromórfico. Já a função $s(t)$ corresponde aos frames de entrada, a serem analisados e comparados, em cada período de tempo.

$$f(t) = \begin{cases} \beta & , \text{ se } s(t) - s(t-1) \geq \alpha \times \beta \\ 0 & , \text{ se } s(t) - s(t-1) \leq -\alpha \times \beta \\ \frac{\beta}{2} & , \text{ de outra maneira} \end{cases} \quad (1)$$

O framework de conversão foi construído de uma maneira parametrizada, com a qual é possível salvar os dados convertidos e determinar alguns atributos das gravações, como: a resolução das gravações neuromórficas, a sensibilidade do sistema para estimular a geração de um evento, ou seja, a mínima variação de luminosidade necessária para gerar um evento e alterar digitalmente a profundidade da gravação através da sua quantidade de bits.

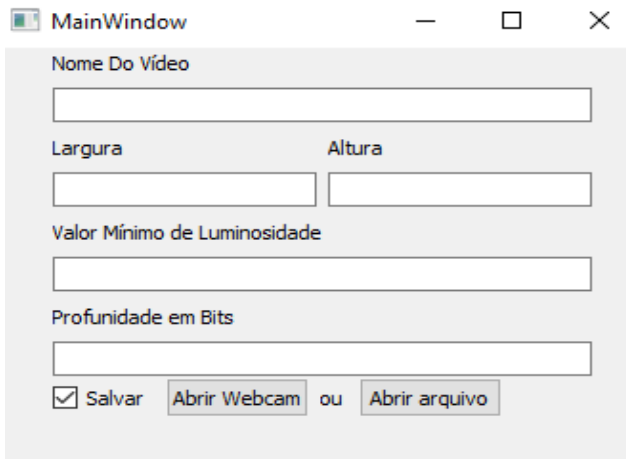
A escolha desses parâmetros é representada pelo bloco número 5 da Figura 1.

III. RESULTADOS

A Figura 2, representa a interface do framework construído, onde podem ser inseridos os parâmetros desejados, como a resolução espacial dos dados, a serem inseridos nos campos de Largura e Altura, o valor mínimo de luminosidade, representado por β na Equação 1, e a profundidade em bits desejada.

Além disso, é possível escolher um arquivo já existente, ou seja, vídeos pré-gravados, ou simplesmente utilizar uma câmera usb comum ou uma webcam. Além do mais, é permitido escolher, através da checkbox, salvar a conversão feita, seja ela através de arquivos já existentes ou não. Sendo admissível escolher o nome do arquivo a ser salvo, no primeiro campo da interface.

Figura 2: Interface gráfica do software de conversão desenvolvido.

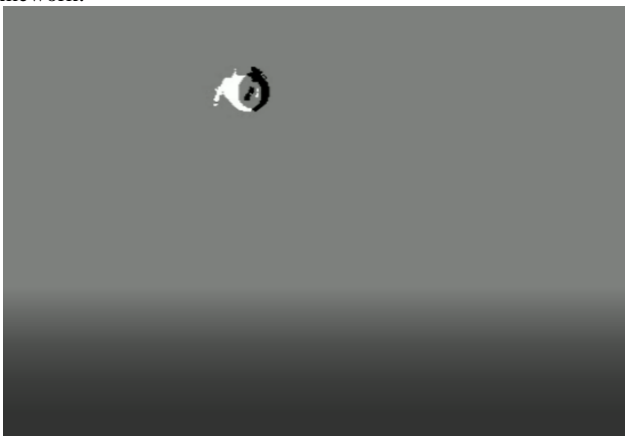


Com a finalidade de validar a eficiência da ferramenta desenvolvida, foram feitas duas gravações, em que utilizamos uma câmera neuromórfica, uma câmera RGB tradicional e o framework. Assim podemos perceber as semelhanças, principalmente na concepção de eventos de ambas, e suas diferenças, relacionadas principalmente às questões de resolução temporal e espacial.

Figura 3: Imagem de uma esfera em movimento capturada por uma câmera RGB comum.



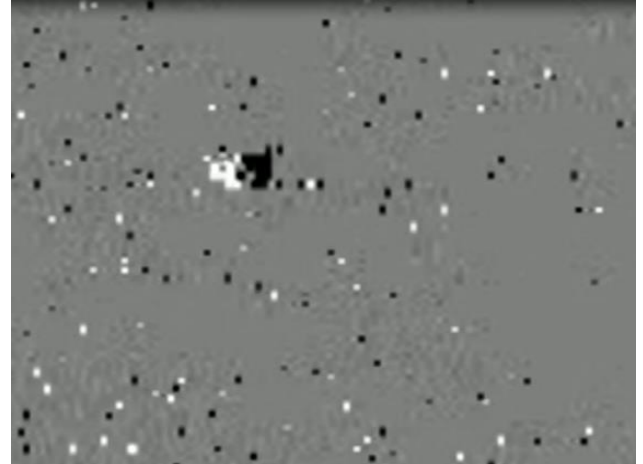
Figura 4: Imagem de uma esfera em movimento, convertida pelo framework.



As Figuras 3, 4 e 5 apresentam um experimento feito a partir da movimentação de uma esfera, para que, os dispositivos a serem comparados, consigam captar o movimento existente na cena. A Figura 3 representa a

obtenção de informações de uma câmera RGB comum. Já a Figura 4 representa a conversão destes dados pelo framework, enquanto a Figura 5 apresenta a mesma situação, porém obtida

Figura 5: Imagem de uma esfera em movimento, capturada por uma câmera neuromórfica.



por uma câmera neuromórfica. Dessa forma, é possível notar algumas diferenças, dentre elas o ganho de resolução espacial pelo framework, sendo ele superior em relação ao dispositivo neuromórfico, tendo escolhido para a conversão uma resolução de 480x480 pixels, e a DVS128, utilizada nesse trabalho, que possui uma resolução de 128x128 pixels.

Figura 6: Imagem de vários objetos e uma esfera em movimento adquirida por uma câmera RGB comum.

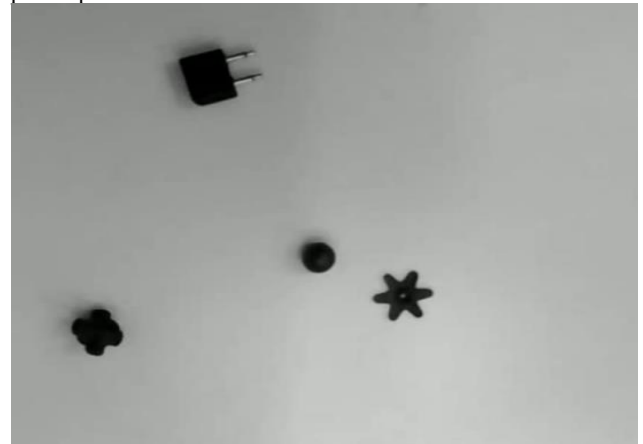
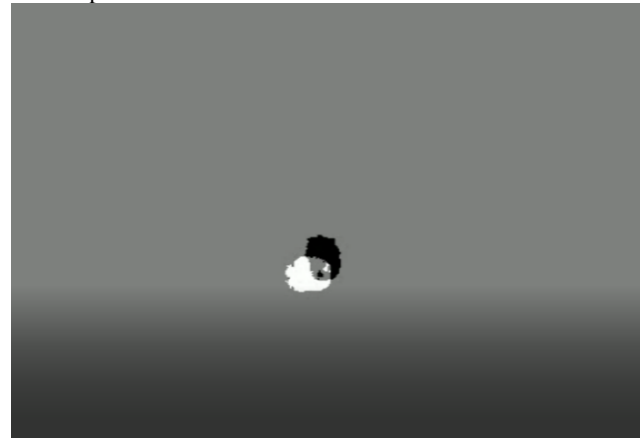


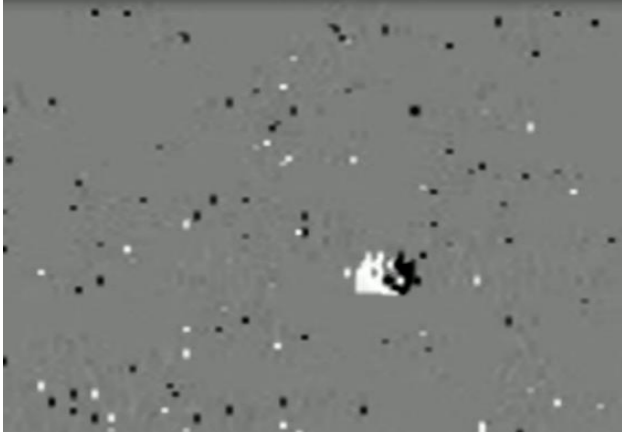
Figura 7: Imagem de vários objetos e uma esfera em movimento, convertida pelo framework.



As Figuras 6, 7, 8 apresentam outro experimento realizado, porém a cena possui vários objetos estáticos, e uma esfera em

movimento. É possível perceber que, nem o framework, nem a câmera neuromórfica detectam os objetos imóveis na cena, como esperado, uma vez que ambas as ferramentas apenas assinalam eventos, ou seja, onde houve alguma alteração de luminosidade.

Figura 8: Imagem de vários objetos e uma esfera em movimento capturada por uma câmera neuromórfica.



Além do mais, é notável que a informação adquirida por uma câmera neuromórfica, Figuras 5 e 8 são muito mais ruidosas, assinalando eventos onde não há movimento relativo entre o sensor e a cena. Podendo esse fator ser prejudicial para certas aplicações desejadas, enquanto que o framework detecta de uma maneira muito mais livre de ruídos, o movimento da esfera, ficando até mesmo mais evidente o sentido do movimento em que o objeto se desloca.

IV. DISCUSSÃO

Os sensores neuromórficos usuais recriam a noção de frame integrando os eventos em um curto período de tempo, de acordo suas posições no espaço e seus respectivos valores de intensidade [7] [12]. Enquanto isso, o framework construído e utilizado para fazer as conversões de dados, possui características diretamente relacionadas com a câmera utilizada para a captura das informações de entrada. Porém, outros aspectos tornam-se mais vantajosos no software, como a resolução espacial, que, dependendo da câmera utilizada para captura de dados, pode ser muito superior, se comparada com câmeras neuromórficas já existentes [7] [9] [11]. Dependendo das aplicações e tarefas desejadas a se realizar, o ganho em resolução espacial pode ser muito benéfico, e a perda em resolução temporal pode não ser prejudicial para o trabalho almejado.

Dessa forma, existe uma diferença entre os dados convertidos pelo framework e aqueles que são capturados por uma câmera neuromórfica, que se dão principalmente pelo processo de funcionamento de cada um. Sendo que, o framework, possui um processo de geração de eventos a partir da análise da variação de luminosidade de acordo a intensidade dos pixels de cada frame. Já dispositivos neuromórficos implementados em hardware, como em FPGAs por exemplo, possuem diversos tipos de geração de eventos, a partir da análise da cena, sendo uma possibilidade, a utilização de circuitos integra-dispara no equipamento, reconstruindo a cena a partir da integração dos eventos no tempo, recriando assim a noção de frame existente em câmeras tradicionais [5] [12] [13] [14].

O software foi construído de maneira parametrizada, onde algumas características das gravações podem ser alteradas durante o processo de conversão, como a sensibilidade da câmera para estimular a geração de um evento, podendo até diminuir a geração de ruídos nos dados, uma vez que uma câmera neuromórfica gravando uma cena em que não há movimento relativo entre ela e o cenário não gera eventos, ou apenas registros ruidosos, enquanto que no framework é possível alterar este parâmetro [7].

Além da possibilidade de alteração dessa propriedade, também é possível alterar a resolução das informações a serem convertidas, o que pode ser benéfico em algumas aplicações onde a resolução da câmera é uma característica importante. Esse processo de alteração da resolução é feito modificando as dimensões de cada frame de entrada para as dimensões escolhidas, utilizando então uma quantidade de pixels diferente da disponível pela câmera.

É possível também alterar digitalmente a profundidade de cada frame. Essa é uma característica que também está ligada à quantidade de bits disponíveis na câmera. Porém, o framework desenvolvido permite que se realize a alteração digital desses valores, atribuindo uma maior faixa de variação de luminosidade para cada pixel.

Portanto, com as diferenças citadas em relação as informações convertidas pelo software e aquelas geradas por uma câmera neuromórfica convencional, consideramos que os dados gerados pelo framework são pseudo-neuromórficos. Dessa maneira, o framework desenvolvido pode ser aproveitado para realização de trabalhos futuros, como a implementação de algoritmos de classificação, segmentação, rastreamento e processos de aprendizado de máquina [6] [10] [15] [16] [17] [18].

V. CONCLUSÃO

Nesse trabalho, foi apresentado um framework capaz de converter informações advindas de câmeras RGB padrão em informações pseudo-neuromórficas. A ferramenta se mostrou muito eficiente para essa tarefa, mesmo tendo limitações devido à perda em resolução temporal e outras diferenças devido a concepção de funcionamento dos sensores de visão neuromórficos, quando comparados com os tradicionais. O software desenvolvido pode ser muito explorado para ampliar a quantidade de dados neuromórficos existentes hoje, convertendo datasets clássicos do campo da visão computacional para que eles também sejam usados no campo neuromórfico, podendo ser utilizado para validar algoritmos desenvolvidos a partir dos datasets já existentes. Logo, através da democratização dessas informações, dados e ferramentas, o campo da visão neuromórfica torna-se cada vez mais amplo e sólido na comunidade científica.

Portanto, mesmo com as limitações existentes neste framework, principalmente em abordagens que necessitem de extrema resolução temporal de dispositivos neuromórficos comuns, ele se torna uma excelente proposta baixo custo para prover dados neuromórficos para aqueles que desejam ingressar neste novo campo de estudo, mas não possuem acesso à uma câmera neuromórfica. Assim, muitas aplicações podem vir a se beneficiar desta ferramenta de conversão de dados.

REFERÊNCIAS

- [1] Calvo, R. (2012). Sistemas bio-inspirados para coordenação de múltiplos robôs móveis. Doctoral Thesis, Instituto de Ciências Matemáticas e de Computação, University of São Paulo, São Carlos. doi:10.11606/T.55.2012.tde-04092012-084821.
- [2] VARGAS, Patricia Amancio. Sistemas computacionais bio-inspirados: síntese e aplicação em inteligência computacional e homeostase artificial. 2005. 192 p. Tese (doutorado) - Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e Computação, Campinas, SP. Disponível em: <<http://www.repositorio.unicamp.br/handle/REPOSIP/260780>>. Acesso em: 6 ago. 2018.
- [3] Tan Cheston, Lalle Stephane, Orchard Garrick. Benchmarking neuromorphic vision: lessons learnt from computer vision *Frontiers in neuroscience*. 2015;9:374.
- [4] Osborn, L. E., Dragomir, A., Bethauser, J. L., Hunt, C. L., Nguyen, H. H., Kaliki, R. R., & Thakor, N. V. (2018). Prosthesis with neuromorphic multilayered e-skin perceives touch and pain. *Science Robotics*, 3(19), [eaat3818]. <https://doi.org/10.1126/scirobotics.aat3818>.
- [5] T. Delbrück, B. Linares-Barranco, E. Culurciello and C. Posch, "Activity-driven, event-based vision sensors," *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*, Paris, 2010, pp. 2426-2429, doi: 10.1109/ISCAS.2010.5537149.
- [6] Diamond Alan, Nowotny Thomas, Schmuker Michael. Comparing neuromorphic solutions in action: implementing a bio-inspired solution to a benchmark classification task on three parallel-computing platforms *Frontiers in neuroscience*. 2016;9:491.
- [7] G. Gallego, T. Delbruck, G. Orchard, C. Bartolozzi, B. Tabá, A. Censi, S. Leutenegger, A. Davison, J. Conradt, K. Daniilidis, D. Scaramuzza: Event-based Vision: A Survey, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, July 2020.
- [8] Aversa A, Tedesco M, Bonifazi P, Difato F, Massobrio P, Levi T, Chiappalone M. A Neuromorphic Prosthesis to Restore Communication in Neuronal Networks. *iScience*. 2019 Sep 27;19:402-414. doi: 10.1016/j.isci.2019.07.046. Epub 2019 Aug 1. PMID: 31421595; PMCID: PMC6706626.
- [9] Delbruck Tobi. Frame-free dynamic digital vision in *Proceedings of Intl. Symp. on Secure-Life Electronics, Advanced Electronics for Quality Life and Society*:21–26 CiteSeer 2008.
- [10] J. Kaiser et al., "Towards a framework for end-to-end control of a simulated vehicle with spiking neural networks," 2016 IEEE International Conference on Simulation, Modeling, and Programming for Autonomous Robots (SIMPAN), San Francisco, CA, 2016, pp. 127-134, doi: 10.1109/SIMPAN.2016.7862386.
- [11] Rebecq, H., Gehrig, D. & Scaramuzza, D. (2018). ESIM: an Open Event Camera Simulator. *Proceedings of The 2nd Conference on Robot Learning*, in PMLR 87:969–982.
- [12] Culurciello, E., Etienne-Cummings, R., and Boahen, K.A. (2003). A biomorphic digital image sensor. *IEEE Journal of Solid-State Circuits* 38, 281–294.
- [13] Ieng, S. H., Carneiro, J., Osswald, M., & Benosman, R. (2018). Neuromorphic Event-Based Generalized Time-Based Stereovision. *Frontiers in neuroscience*, 12, 442.
- [14] Brandli Christian, Berner Raphael, Yang Minhao, Liu Shih-Chii, Delbruck Tobi. A 240× 180 130 db 3 μs latency global shutter spatiotemporal vision sensor *IEEE Journal of Solid-State Circuits*. 2014;49:2333–2341.
- [15] Delbruck Tobi, Lichtsteiner Patrick. Fast sensory motor control based on event-based hybrid neuromorphic-procedural system in 2007 IEEE international symposium on circuits and systems:845–848 IEEE 2007.
- [16] Gallego Guillermo, Lund Jon EA, Mueggler Elias, Rebecq Henri, Delbruck Tobi, Scaramuzza Davide. Event-based, 6-DOF camera tracking from photometric depth maps *IEEE transactions on pattern analysis and machine intelligence*. 2017;40:2402–2412.
- [17] Ni Zhenjiang, Ieng Sio-Hoi, Posch Christoph, Regnier Stéphane, Benosman Ryad. Visual tracking using neuromorphic asynchronous event-based cameras *Neural computation*. 2015;27:925–953.
- [18] Paredes-Valles F, Scheper KYW, de Croon GCHE. Unsupervised Learning of a Hierarchical Spiking Neural Network for Optical Flow Estimation: From Events to Global Motion Perception. *IEEE Trans Pattern Anal Mach Intell*. 2020 Aug;42(8) 2051-2064. doi:10.1109/tpami.2019.2903179. PMID: 30843817.