Initiative MSTD Postdoctoral fellowship

Maîtrise des Systèmes Technologiques Sûrs et Durables

« Control of Safe and Sustainable Technological Systems »

Event-based vision for autonomous driving safety regarding soft mobility users.

Postdoc advisors:

- *Advisors names, laboratories*: - Franck **Davoine**, Julien **Moreau**, Heudiasyc UMR 7253, **UTC**, firstname.lastname@hds.utc.fr ;
  - Catherine **Achard**, ISIR UMR 7222, **Sorbonne Université**, Paris catherine.achard (at) sorbonne-universite.fr

Heudiasyc lab. address: 57 avenue de Landshut, CS 60319, 60203 Compiègne Cedex, France.
ISIR lab. address: Campus Pierre et Marie Curie, 4 Place Jussieu, 75005 Paris, France.

Application deadline: November 15, 2020 (beginning of the postdoctoral fellowship: February 2021).

Send your resume and application letter, preferably in a single PDF document by email to **franck.davoine (at) hds.utc.fr**, with the subject: **[postdocMSTD]**

Location: Laboratoire Heudiasyc, UTC – Université de technologie de Compiègne, 57 avenue de Landshut, 60200 **Compiègne, France**. Robtic systems in interaction (SyRI) team. https://www.hds.utc.fr/

Project abstract:

Urban traffic lanes are increasingly shared between vehicles and soft mobility users (bicycles, kick scooters, etc.), whose erratic behavior might harm their safety. The project aims to develop vehicle perception systems, making it possible to reliably detect nearby vulnerable people. These can be difficult to see depending on the context (partial occlusions, rapid movements, unpredictable behavior, etc.). This requires research in suitable computer vision algorithms, multi-sensor information fusion, machine learning and artificial intelligence.

Scientific description:

Several recent studies insist on the importance of encouraging soft and less carbon-intensive mobility in urban areas, with a view to achieving the implicit objective of reducing greenhouse gas emissions by 5% per year.

The objective of the postdoctoral project is to demonstrate the benefit of using so-called “event” cameras [1] to improve the safety of vulnerable road users (that is to say people in soft mobility) vis-à-vis the road car. The reaction time of the system must be reduced to its minimum [2, 3], and its reliability must be maximum.
Unlike classic image sensors which capture snapshots of the scene to be rendered, taken at a precise moment and containing no dynamic information, the neuromorphic or bioinspired sensor (event camera) aims to capture visual movement. It emits a +/- 1 pulse for each independent pixel asynchronously as soon as the variation in brightness at the pixel exceeds a given threshold. By design, this sensor encodes visual dynamics in the form of high

resolution spatio-temporal patterns of events (contours of moving objects, trajectories, velocity, etc.) [1]. This technology makes it possible in particular to benefit from increased speed, a high dynamic range (perception of information under unfavorable light conditions, dazzling light, night, etc.), a reduction in calculation costs and energy consumption.

For this, our strategy is to contribute to test and compare methods using event cameras against methods using color cameras widely explored in the literature and for which the algorithms can be considered to have reached maturity, in particular deep learning-based detections which constitute the current state of the art [4, 5, 6].
In a complementary manner, using event cameras, we propose to contribute to the development of deep learning methods adapted to quickly and reliably detect the vulnerable near the vehicle. Grafted Networks type methods [7] could be tested to compensate for the lack of learning data in certain modalities such as event vision. Techniques such as embeddings [8] could be exploited to allow clustering of signals by networks to detect and track moving elements. In addition, events, by their impulsive nature, immediately lend themselves to the use of spiking neural networks, allowing the development of neural networks with reduced energy consumption [9] (the Achilles heel of classical neural networks).
The fusion of color and events is also a serious avenue that we will study to combine the advantages of the two modalities [10] (richness of color images making it possible to recognize behaviors [11], instantaneity and safety of operation of events in more various). We could also be interested in the comparison or even the fusion with a thermal camera, another high-performance sensor at night, for example. The Heudiasyc laboratory is equipped with color, thermal and event cameras (Prophesee, iniVation), works in this direction are already initiated [12].

The purpose of this project is to promote the most appropriate perception techniques to serve the safety of vulnerable road users in urban environments. We choose to focus our research on the sensor / algorithm set and not just on improving algorithms for a sensor chosen by default. The choice of the combination of event camera and pulse neural networks will optimize the safety of the system and also minimize its environmental costs. This is a promising area of research.

---> At its level, the project contributes to serving the objectives of the MSTD initiative which are safety and sustainability: (1) development of perception systems adapted to the particular case of driving the vehicle in an urban space shared with vulnerable people, (2) promotion of a desirable vision of carbon-free mobility, soft mobility, by making the vehicle safer for the vulnerable.

Scientific schedule:

- Bibliography of soft mobility users behaviors and interactions with intelligent vehicles, and of multi-sensor fusion studies applied to urban scene analysis, in order to extract a starting architecture with some possible improvement perspectives. Regular interactions with the postdoc and advisors of both involved laboratories.
- Preparation of use cases and data acquisitions on UTC's test track, making use of the instrumented vehicles of Heudiasyc laboratory [13]. Data sharing with ISIR team for the development of adapted vulnerable detection methods in hazardous situations (proximity to the vehicle, quick movements, views partially occluded by obstacles).
- Gradual comparison and improvement of proposed architectures.
- Publication in a journal or as a minimum in a top-ranking international conference; the publication will concern recorded datasets and proposed architecture.

References:

[1] Gallego, G., Delbruck, T., Orchard, G., Bartolozzi, C., Taba, B., Censi, A., ... & Scaramuzza, D. (2019). Event-based vision: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI), July 2020.
[2] Rasouli, Amir, and John K. Tsotsos. Autonomous vehicles that interact with pedestrians: A survey of theory and practice. IEEE Transactions on Intelligent Transportation Systems, 2019.
[3] Combs, T. S., Sandt, L. S., Clamann, M. P., McDonald, N. C. Automated Vehicles and Pedestrian Safety: Exploring the Promise and Limits of Pedestrian Detection. American journal of preventive medicine. 56(1), January 2019.
[4] Liu, L., Ouyang, W., Wang, X., Fieguth, P., Chen, J., Liu, X., & Pietikäinen, M. (2020). Deep learning for generic object detection: A survey. International journal of computer vision, 128(2), 261-318.
[5] Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. Preprint arXiv:2004.10934.
[6] Benzine, A. and Luvison, B. and Pham, Q.C. and Achard, C. Deep, robust and single shot 3D multi-person human pose estimation from monocular images. IEEE ICIP International Conference on Image Processing 2020.
[7] Hu, W., Delbruck, T., Liu, S.C. Learning to Exploit Multiple Vision Modalities by Using Grafted Networks, ECCV 2020.
[8] Istasse, M., Moreau J., De Vleeschouwer, C. Associative Embedding for *Game-Agnostic* Team Discrimination. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Jun 2019, Long Beach, United States.
[9] Tavanaei, A., Ghodrati, M., Kheradpisheh, S. R., Masquelier, T., & Maida, A. (2019). Deep learning in spiking neural networks. Neural Networks, 111, 47-63.
[10] Capellier, E., Davoine, F., Cherfaoui, V., Li, Y. Fusion of neural networks, for LIDAR-based evidential road mapping. Article soumis, en phase de révision mineure, 2020.
[11] Benzine, A., Chabot, F., Luvison, B., Pham, Q.C., Achard, C. PandaNet: Anchor-Based Single-Shot Multi-Person 3D Pose Estimation, IEEE CVPR 2020.
[12] Brebion, V., under the supervision of Davoine, F. and Moreau, J. Study and application of the event camera to the autonomous vehicle, Master thesis, UTC, September 2020.
[13] https://www.hds.utc.fr/recherche/plateformes-technologiques/vehicules-intelligents-autonomes.html

Requirements:

We are looking for a highly-motivated, independent, and skilled researcher with a Ph.D. degree and related experience in machine vision, deep learning, artificial intelligence, and/or computer science/engineering.
A background and track record of peer-reviewed conference/journal publications in machine vision/machine learning is compulsory (object recognition, deep learning, data fusion, etc.).
Experience with designing, managing, and executing experiments is a mandatory requirement. Relevant knowledge and experience in developing visual-based solutions with Python are mandatory. As the Research Fellow will be working in cooperation with universities and possibly with companies, we expect a pro-active and collaborative attitude and excellent communication skills both in presentations and writing.

Salary:

~ 2,600 – 2,700€ gross per month.