Event-based perception learning for mobile robots

PhD description :

In this PhD project, we aim to explore the benefits of a new class of **biologically inspired vision sensors for mobile robotics applications**. These sensors mimic the operation of the primate retina and represent the visual scene via asynchronous event streams rapidly indicating *changes* in luminance at every pixel. These sensors have a number of decisive advantages compared to traditional video cameras. First, they have drastically improved speeds (equivalent to >10,000 frames per second), because pixels operate and communicate independently and asynchronously. This is important for any applications where short reaction times are critical. Second, they transmit information with massively reduced bandwidth, because they mostly register and communicate luminance *changes* in the image. Third, they can operate over a much greater dynamic range (>120 dB), which is important for open environments and especially outdoor applications. Fourth, they have substantially lower power consumption (<10 mW), which is of great importance to reduce the energetic requirements of mobile robots.
The advantages of these sensors could offer important benefits for mobile robotics applications. However, because of the fundamentally different data format of asynchronous streams of events instead of sequences of images, classic image and video processing techniques are either not applicable or do not support the unique advantages of these sensors. While first systems for, e.g., stereo vision or object tracking with these sensors have been demonstrated (e.g. Everding & Conradt, 2018), we are interested in ***active* perception** for mobile robots, where the robot moves through its environment and has movable cameras with which it samples the scene. Taking inspiration from biological vision and building on our extensive experience in developing vision systems that can autonomously self-calibrate their perception based on principles from information theory, we aim to address the following tasks:

T1: Build an active binocular event-based vision system for a mobile robot. The system will be built with a pair of commercially available event-based cameras (e.g. Prophesee Onboard, https://www.prophesee.ai). These will be mounted on a pan-tilt-vergence platform (3 degrees of freedom) or two separate pan-tilt units (4 degrees of freedom). The vision system will be mounted on a mobile robot from the lab.

T2: Learning fixational eye movements to maximize coding efficiency. An imaging sensor that is like the human retina and only responds to luminance changes is essentially “blind” unless there is either movement of the camera or in the scene. We will apply **reinforcement learning** to learn optimal strategies for camera movement in order to maximize the amount of information provided by the cameras while minimizing the associated data rate and energy costs. For example, we predict that the system will learn to make small fixational eye movements when the robot and the scene are stationary, allowing it to gain information about the scene through “jittering” its eyes as humans do.

T3: Event-based active stereo and motion vision on a mobile robot. In humans, perceiving motion and depth relies on pursuit and vergence eye movements, respectively. Over the last years, we have developed a theoretical framework called **Active Efficient Coding** (AEC) for how such sensor movements can be learned autonomously by robots (Zhao et al., 2012; Lonini et al., 2013; Teulière et al., 2015). Here we propose to build the **world’s first event-based binocular vision system that can self-calibrate its pursuit and vergence eye movements** using the AEC approach. For this, we will combine **unsupervised learning techniques** for learning compact representations of the binocular event streams with **reinforcement learning** for generating the motor commands for the cameras to maximize coding efficiency.

Context

This PhD is founded through the FrenchTech chair program. The candidate will prepare a PhD from the Université Clermont-Auvergne (Clermont-Ferrand, France). She/he will be joining the *Image, Perception Systems and Robotics* group of Institute Pascal which has long experience in computer vision and mobile robots.
This thesis is also the result of an existing collaboration between Institut Pascal and Prof. Jochen Triesch from the Frankfurt Institute for Advanced Studies (FIAS), who will be in Institut Pascal at least 3 months per year. The candidate will benefit from his strong expertise in computational neuroscience and will have the opportunity to have a research experience in FIAS.

Advisors:
- Prof. Jochen Triesch (Frankfurt Institute for Advanced Studies)
- Dr. Céline Teulière (Institut Pascal, UCA)
- Prof. Vincent Barra (LIMOS, UCA)

Research Group: Institut Pascal
University: Université Clermont Auvergne (UCA) – Clermont Ferrand - France
Contact: celine.teuliere@uca.fr

Position Requirements:

- MS Degree or equivalent in Computer Science, Computational Neuroscience, Robotics or other relevant subjects
- Experience with Machine learning (in particular reinforcement learning), Computational Neuroscience and/or Computer Vision are highly desirable.
- Strong programming skills (C/C++, Python)
- Solid analytical ability
- Taste for experimentation
- Good spoken and written English
- High motivation

References

Everding, L., & Conradt, J. (2018). Low-Latency Line Tracking Using Event-Based Dynamic Vision Sensors. *Frontiers in neurorobotics*, *12*, 4.

Lonini, L., Forestier, S., Teulière, C., Zhao, Y., Shi, B. E., & Triesch, J. (2013). Robust active binocular vision through intrinsically motivated learning. *Frontiers in neurorobotics*, *7*, 20.

Teulière, C., Forestier, S., Lonini, L., Zhang, C., Zhao, Y., Shi, B., & Triesch, J. (2015). Self-calibrating smooth pursuit through active efficient coding. *Robotics and Autonomous Systems*, *71*, 3-12.

Zhao, Y., Rothkopf, C. A., Triesch, J., & Shi, B. E. (2012, November). A unified model of the joint development of disparity selectivity and vergence control. In *ICDL, 2012* (pp. 1-6).