University of Glasgow

Jan Siebert
11/22/2016

University Partner:
University of Glasgow

Academic Supervisor - Name and Email Address:
J. Paul Siebert, paul.siebert@glasgow.ac.uk

Suggested Project Title:
A Smartphone-based Universal Vision Sensor Combining a Software Retina and Deep Learning

Suggested Project Summary:
In this project we propose to develop a low-cost and highly integrated camera sensor for egocentric and robotic vision applications. Our goal is to address a key issue in designing robotics systems, namely the cost of an integrated camera sensor & vision system that meets the bandwidth, processing and analysis requirements for many advanced robotics applications. We are especially targeting lightweight robotic vision required for as visual surveillance and SLAM in autonomous aerial vehicles. We also propose to address or wearable intelligent monitoring devices that serve as assistants, and/or are intended to interpret human actions from an egocentric perspective, e.g. for robot training or activity monitoring, e.g. safety systems for public vehicle operators. Therefore, this is a heartland visual data processing project, with implications for autonomous systems requiring powerful, yet low-cost, vision systems.

There is currently much work going on to adapt smartphones to provide complete robot vision systems, as the smartphone is so exquisitely integrated by having camera(s), inertial sensing, sound I/O and excellent wireless connectivity. Mass market production makes this a very low-cost platform and manufacturers for quadrotor drone suppliers and small low-cost autonomous robots are now employing a smartphone to provide a vision system/control system. Accordingly, there is strong commercial potential in developing “universal” vision systems to operate on self-contained platforms as smartphones, or low-power high-performance embedded modules such as NVIDIA’s Jetson processor.

However, a current issue in adopting the smartphone as a self-contained vision system is the inability of the low-power phone processors to analyse the large-scale camera data provided by the phone itself. A comparatively new approach to data analysis termed, Deep Learning (DL), has the potential to accomplish all the visual data analysis required for advanced robotics and egocentric vision applications. DL systems can process only very small images patches of ~100x100 pixels at a time and are therefore slow to compute over large scale images as captured using a smartphone.

The underlying novel scientific basis of our approach is to apply a biologically-motivated "software retina" transformation we have developed to reduce the dimensionality of the input image space by a factor of ~x100 prior to training and processing by means of DL nets. In this approach the retina also transforms images spatially and maps image rotations and scale changes into spatial shifts. This simplifying transformation of the input image reduces the DL net's requirement to learn image scale and rotation transformations (which contribute to the large computational demands for training and opaqueness of the learned structures the net produces).

By reducing the input image size accordingly, and therefore learning and processing requirements, we aim to develop compact and lightweight egocentric and robot vision sensor using a smartphone as the target platform.

The proposed development has profound implications for efficient visual data processing using DL and has the potential to provide Scotland core expertise in this field to support Scottish commercial activity in the development of embedded real-time vision systems for mass-market applications such as driverless cars and home robotic devices.

Collaboration Sought for the Project:
An ideal collaborator would be developing either autonomous applications, such as driverless vehicles, robot systems, surveillance and security systems, inspection systems (or similar), or wearable camera systems for monitoring the activity of an individual from an egocentric perspective. Egocentric applications include intelligent assistants instructing equipment or plant operators, or warning them of hazard conditions. Applications also include monitoring systems, perhaps checking the awareness of the wearer, or training data capture for autonomous robot operation.

The collaborator would already have, or be prepared to establish, a development pipeline whereby the proposed research could be evaluated quickly during its development within a commercially relevant testbed scenario to facilitate its speedy product integration and commercial impact. This collaboration could satisfy both the needs of a systems integrator/developer requiring low-cost but advanced vision capability and inform the research team with valuable insights into the operation of the developed system in commercially relevant scenarios.

Published or Private?:
Yes