A Biologically motivated software retina for robotic sensors for ARM-based mobile platform technology+

J. Paul Siebert[1], Adam Schmidt[2], Gerardo Aragon-Camarasa[1], Nick Hockings[3], Xiaomeng Wang[1] & W. Paul Cockshott[1].

[1]School of Computing Science, Computer Vision for Autonomous systems Group, University of Glasgow, Scotland, UK

[2]CETAPS laboratory, Faculty of Sports Sciences, Rouen University, Rouen, France.

[3]Robotics Laboratory, School of Computer Science, University of Bath, Bath, UK.

Contact: {paul.siebert@glasgow.ac.uk; schmidt.adam@gmail.com}

A key issue in designing robotics systems is the cost of an integrated camera sensor that meets the bandwidth/processing requirement for many advanced robotics applications, especially lightweight robotics applications, such as visual surveillance or SLAM in autonomous aerial vehicles. 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 from quadrotor drone suppliers to children's toys, such as the Meccanoid robot [5], employ a smartphone to provide a vision system/control system [7,8].

Accordingly, many research groups are attempting to optimise image analysis, computer vision and machine learning libraries for the smartphone platform. However current approaches to robot vision remain highly demanding for mobile processors such as the ARM, and while a number of algorithms have been developed, these are very stripped down, i.e. highly compromised in function or performance. For example, the semi-dense visual odometry implementation of [1] operates on images of only 320x240pixels.

In our research we have been developing biologically motivated *foveated* vision algorithms based on a model of the mammalian retina [2], potentially 100 times more efficient than their conventional counterparts. Accordingly, vision systems based on the foveated architectures found in mammals have also the potential to reduce bandwidth and processing requirements by about x100 - it has been estimated that our brains would weigh ~60Kg if we were to process all our visual input at uniform high resolution. We have reported a foveated visual architecture [2,3,4] that implements a functional model of the retina-visual cortex to produce feature vectors that can be matched/classified using conventional methods, or indeed could be adapted to employ Deep Convolutional Neural Nets for the classification/interpretation stage. Given the above processing/bandwidth limitations, a viable way forward would be to perform off-line learning and implement the forward recognition path on the mobile platform, returning simple object labels, or sparse hierarchical feature symbols, and gaze control commands to the host robot vision system and controller.

We are now at the early stages of investigating how best to port our foveated architecture onto an ARM-based smartphone platform. To achieve the required levels of performance we propose to port and optimise our retina model to the mobile ARM processor architecture in conjunction with their integrated GPUs. We will then be in the position to provide a foveated smart vision system on a smartphone with the advantage of processing speed gains and bandwidth optimisations. Our approach will be to develop efficient parallelising compilers and perhaps propose new processor architectural features to support this approach to computer vision, e.g. efficient processing of hexagonally sampled foveated images.

Our current goal is to have a foveated system running in real-time on at least a 1080p input video stream to serve as a front-end robot sensor for tasks such as general purpose object recognition and reliable dense SLAM using a commercial off-the-shelf smartphone. Initially this system would communicate a symbol stream to conventional hardware performing back-end visual classification/interpretation, although simple object detection and recognition tasks should be possible on-board the device. We propose that, as in Nature, foveated vision is the key to achieving the necessary data reduction to be able to implement complete visual recognition and learning processes on the smartphone itself.

1. **Semi-Dense Visual Odometry for AR on a Smartphone** (T. Schöps, J. Engel, D. Cremers), In *International Symposium on Mixed and Augmented Reality*, 2014.
2. **A computational model of space-variant vision based on a self-organised artificial retina tessellation**, S.L. Balasuriya, *PhD Thesis*, School of Computing Science, University of Glasgow, Glasgow Scotland, UK, 2006
3. **Hierarchical Feature Extraction using a Self-organised Retinal Receptive Field Sampling Tessellation** L.S.Balasuriya & J.P.Siebert, *Neural Information Processing - Letters & Reviews*, 2006.
4. **Simulating the Effects of Laser Dazzle on Human Vision** C.A. Williamson, E. Strachan, & J.P. Siebert, *International Laser Safety Conference 2013*, Orlando, Florida, USA, March 18-21, 2013.
5. **Meccanoid G15KS Personal Robot**, http://www.meccano.com/product/p10945/meccanoid-g15ks-personal-robot
6. **SmartCopter - Enabling Autonomous Flight in Indoor Environments with a Smartphone as On-Board Processing Unit** A. Mossel, M. Leichtfried, C. Kaltenriener, H. Kaufmann, *International Journal of Pervasive Computing and Communications (invited)*, 10 (2014), 1; 92 - 114.
7. **Phonedrone Ethos**, Xcraft Enterprises, http://xcraft.io/phone-drone/
8. **A Smartphone Is the Brain for This Autonomous Quadcopter** E. Ackerman, *IEEE Spectrum*, 31 January, 2015. http://spectrum.ieee.org/automaton/robotics/drones/a-smartphone-is-the-brain-for-this-autonomous-quadcopter
   Video material: https://www.youtube.com/watch?v=mtTiPnNUKj8