Active Vision for Robotic Applications

S. Rougeaux
Department of Systems Engineering,
Research School of Information Sciences and Engineering
Australian National University
e-mail: sebastien.rougeaux@anu.edu.au

Keywords: **robotics, active and real-time vision**

Extended abstract

The importance of eye movement to biological visual systems is obvious. In contrast, controlled camera movement have played a small role in computer vision research, but are becoming increasingly recognized as important capabilities in robotic visual perception [1, 2, 8, 6, 7].

For sure, active vision signals a decisive shift from *data driven* to *task driven* applications. Instead of building a whole set of representations from passive images before making decisions, an active sensor, with a particular goal in mind, can select useful information and simply discard the irrelevant parts of the scene, changing the very nature of *both* the goals and means of computer vision:

> "Participation changes what the vision system is capable of perceiving and changes the nature of the requirements on what needs to be perceived." Brooks, 1992 [3].

This paradigm is clearly demonstrated, for example, in the work on gaze fixation for the Rochester Robot [4]. In their experiment, they show how dynamic vergence[1] control can maintain the observed object within a narrow range of disparity, enabling segmentation with a simple disparity filtering algorithm. As a result, tracking can be achieved without requiring any model of the target shape, color or texture.

[1] Vergence is the symmetric rotations of the eye about vertical axes that adapts to depth changes of the target during fixation.

Movements of the camera can also bring substantial advantages over passive sensing devices. To name just a few examples, active vision can:

- provide the robot with a **full panoramic range of view**. Contrary to fixed sensors where the field of view is restricted by the specifications of the lens, a moving camera can greatly extend the sensing area of the robot and thus improve its global awareness of the environment.
- enable **foveal vision**. A passive camera cannot obtain a high resolution image of a large scene without having to sustain an enormous amount of information. An agile system may, on the opposite, exhibit only a small area of high resolution, yet be able to scrutinize any details of a scene by shifting attention to the regions of interest.
- **facilitate stereo fusion**. With vergence control, it is possible to observe any points of interest in the scene with a limited range of disparity. For the field of depth reconstruction, this ability can drastically reduce the search area for matching points in the pair of images and thus improve the speed of correlation or feature correspondence techniques.
- provide **moving images of static scenes**. It is fairly obvious that a robot perceive more when it is moving in a static environment than when it is stationary. In this eventuality, the sequences of acquired images usually present

temporal and spatial consistencies that can be used to derive useful representations of the observed field. Controlled camera motions, for example, have been successfully applied in the field of structure from motion [5].

Technological opportunities have also contributed significantly to the recent advances in active vision. With computer hardware becoming faster and cheaper, scientists are now more than ever capable of experimenting the link between perception and action.

The paper will present several designs of active robotics heads and their application for robotic applications. In the first section, we will introduce a 2 DOFs prototype mounted on mobile robot capable of following moving objects in real-time and in cluttered environments. Using disparity information, the system is capable of locating and actively following other robots. On the basis of the visual information alone, the robot is then able to cooperate with other robots, removing obstacles on their path when needed.

We will then present a more complex robotics head called ESCHeR that exhibits mechanical and optical properties similar to the human visual system. Using visual cues such as motion and disparity, ESCHeR can detect, acquire and follow arbitrary targets as they move in cluttered surroundings, without any *apriori* knowledge about the object's shape, color, texture or trajectory. The tracking holds for the observation of human gestures such as motions of head and limbs as well. The tracking behavior draws its robustness from the space-variant properties of ESCHeR's lenses, as well as from the interaction between image processing and dynamic control of the eye movements.

We will also show how the basic visual skills exhibited by ESCHeR can support the emergence of more complex behaviors. As an example, we present a hand-eye coordination experiment that requires no visual model of the end-effector, only a minimum knowledge about the head and arm kinematics and no calibration procedure at all. The coordination is established by learning the motor-motor mapping between head and arm joint spaces while ESCHeR is tracking the end-effector of a robotic manipulator.

To conclude, we will present the latest active heads designed at the RSISE, nicknamed **HyDrA** [9] (**Hy**brid **Dr**ive **A**ctive Vision Head) and **CeDAR** [10] (**C**abl**e** **D**rive **A**ctive Vision **R**obot). HydrA is based on the combination of cable-drive transmissions for the neck rotations and direct-drive motors for the control of eye vergence. The head has been used to validate a novel and compact algorithm capable of controlling the motion of both saccades and smooth pursuit with optimal accelerations and velocities. CeDAR has been fully designed on cable-drive transmissions to reduce the load and inertia on its fast rotating axes while completely eliminating backlash effects. CeDAR exhibits rotational eye accelerations and velocities superior to the performance of the human visual system, while achieving position accuracy and repeatability of less than 0.01 degree.

References

[1] J. Aloimonos, I. Weiss, and A. Bandopadhay. Active vision. *International journal on Computer Vision*, 1:333–356, 1988.

[2] D.H. Ballard. Animate vision. *Artificial Intelligence*, 48:57–86, 1991.

[3] Rodney Brooks. Foreword. In Andrew Blake and Alan Yuille, editors, *Active Vision*, pages 155–172. MIT Press, Cambridge MA, 1992.

[4] D. Coombs and C. Brown. Real-time smooth pursuit tracking for a moving binocular robot. In *in Proc., Int. Conf. on Computer Vision and Pattern Recognition*, pages 23–28, 1992.

[5] O. Faugeras, F. Lustman, and G. Toscani. Motion and structure from motion from point and line matches. In *in Proc., 1st Int. Conf. on Computer Vision*, pages 25–34, 1987.

[6] Ian D. Reid and David W. Murray. Active tracking of foveated feature clusters using affine structure. *International Journal of Computer Vision*, 18(1):41–60, April 1996.

[7] Sébastien Rougeaux and Yasuo Kuniyoshi. Robust tracking by a humanoid vision system. In *in Proc., First Int. Workshop Humanoid and Human Friendly Robotics, Tsukuba, Japan*, 1998.

[8] P. M. Sharkey, D. W. Murray, S. Vandevelde, I. D. Reid, and P. F. McLauchlan. A modular head/eye platform for real-time reactive vision. *Mechatronics*, 3(4):517–535, 1993.

[9] O. Sutherland, S. Rougeaux, S. Abdallah, and A. Zelinsky. Tracking with hybrid-drive active vision. In *In Proc., Australian Conference on Robotics and Automation, Melbourne*, pages 155–160, 2000.

[10] H. Truong, S. Abdallah, S. Rougeaux, and A. Zelinsky. A novel mechanism for stereo active vision. In *In Proc., Australian Conference on Robotics and Automation, Melbourne*, pages 161–166, 2000.