Integration of haptic senses and predetermined motion for autonomous object segmentation.

Daniel Beale
University of Bath
Department of Computer Science.

Dr. Peter Hall
(Supervisor)
University of Bath
Department of Computer Science.

Dr. Pejman Iravani
(Supervisor)
University of Bath
Department of Mechanical Engineering.

Introduction

The overall goal for my project is to build a robot which can build an understanding of the world and objects around it by imitating natural learning, and interacting physically with the objects of interest.

We present here, a method for autonomous object segmentation, problems with the method, and a potential solution.

Learning like a child

It is relatively new in the robotics literature to view and implement machine learning in a developmental way. The psychological foundation as well as potential solutions have been reviewed recently [2003, Lungarella et al.]

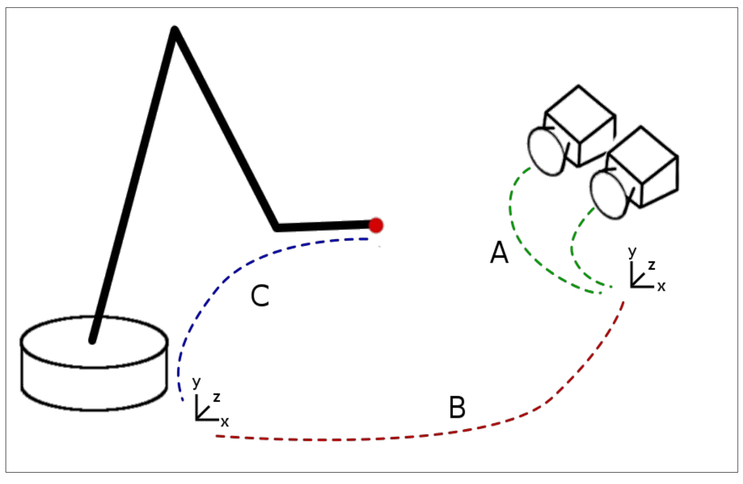
The developmental learning paradigm requires us to use only a very low set of assumptions which ideally reflect mechanisms inherently found in nature, and to acquire knowledge progressively.

This approach is useful for a variety of reasons, but most importantly because it makes the potential applications very robust to noise and uncertainty, allowing the robot to adapt to its environment.

Hand-eye Coordination

To integrate the vision system with a robotic manipulator, a map from the cameras to the end effector must be learned.

- The vision literature [2000, Hartley and Zisserman] gives a way of getting from the camera system into 3d, by learning matrix models for each camera and triangulating [1997, Hartley and Sturm].
- The kinematics of the manipulator is solved analytically [2007, Spong], assuming knowledge of the length of each link. This gives a mapping from three dimensions in to the joint space.
- Finally a linear map between the arm system frame and vision frame is formulated as a least squares problem, and solved using the standard techniques.

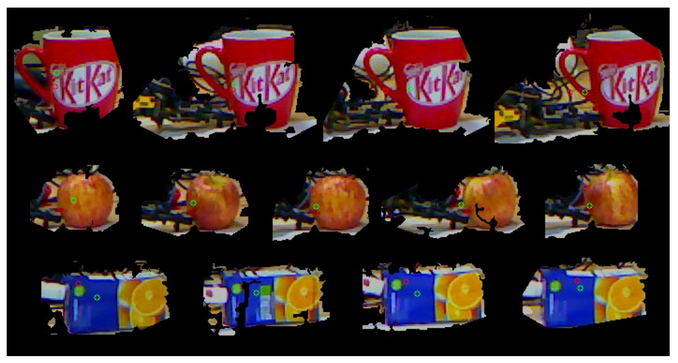


This map is invertible. The robot can place its hand anywhere in the scene it can find a point match, and also locate its hand in each camera.

Segmentation based on predetermined motion

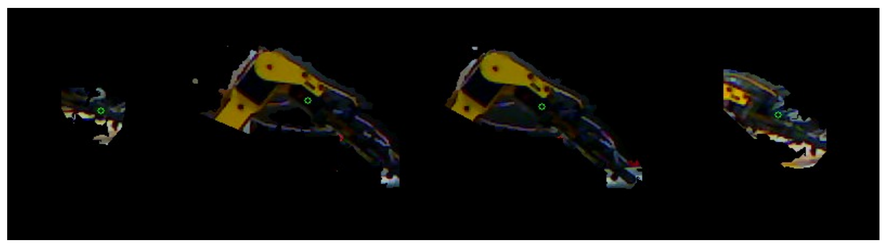
Now that the robot has an understanding of where its arm is located it can then predict the optic flow in each camera. Anything that coheres with this movement can be used to give a segmentation [Fitzpatrick and Metta].

The segmentation uses subtraction and a graph cut algorithm to clean up the final result.



The key to success in this kind of segmentation is determining exactly when the touch has occurred and is particularly difficult since optical flow is naturally very volatile and prone to noise.

This noise and also the complex nature of movements of the arm can cause drastic expansions of optic flow and trigger a premature visual touch, leading to incorrect segmentations.



Adding a touch sensor

Adding a haptic sensor could help with knowledge of exactly when the touch occurred. Using a this sensor alone is not enough however, since any small discrepancy in the timing could lead to premature touch detection, it is also difficult to detect objects which are very light.

Initially finding a maximum of the posterior probability at each time interval seems ideal.

$$p(touch|vision, haptics)$$

In fact it is not this simple. The camera system is slow due to the large amount of information being processed, and as a result the two signals consistently appear at different points in time.

To ensure the classification of a touch is robust enough, the time delay is integrated in statistically.

Current and future work

Currently we are working on a solution to the above problem to efficiently aid and improve the quality of touch recognition.

We also plan to go on to implement and improve the quality of arm removal from the active segmentations, another difficult and visible problem.

References

[2007, Mark W. Spong, *Robot Dynamics and Control*, Wiley, John & Sons, Incorporated

[2003, Lungrella et al.] Max Lungarella, Giorgio Metta, *Beyond Gazing, Pointing, and Reaching: A Survey of Developmental Robotics*, ERIPROB, 2003.

[2003, Fitzpatrick and Metta] Paul Fitzpatrick and Giorgio Metta *Grounding vision through experimental manipulation*. Philosophical transactions of the Royal Society: Mathematical, Physical, and Engineering Sciences.

[2000, Hartley and Zisserman] Richard Hartley and Andrew Zisserman *Multiple View Geometry - Second Edition* Cambridge University Press, 2000.

[1997, Hartley and Sturm] R.Hartley and P. Sturm. *Triangulation.* Computer Vision and Image Understanding, 68(2):146-157, 1997.