Getting to Know the World: Active Acquisition of Categories

Gert Kootstra
Artificial Intelligence Institute
University of Groningen
The Netherlands
gert@ai.rug.nl
http://www.ai.rug.nl/~gert

The Problem

Figure 1: Objects look very different from different viewpoints.

One of the main problems in computer vision is the **object constancy problem**: Objects appear very different depending on the viewpoint (fig. 1). A **passive** system that has learned the object from one viewpoint, will not be able to classify it from another viewpoint.

We propose to tackle this problem by using **active perception** to explore the object similar to how humans learn and categorize objects (fig. 2).

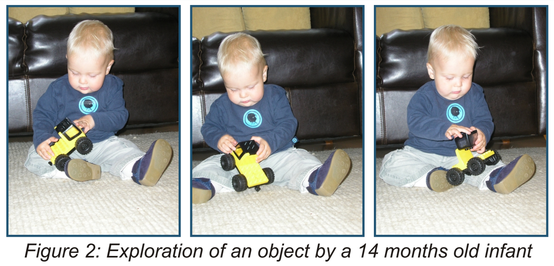

Figure 2: Exploration of an object by a 14 months old infant

Why Active Vision?

- Simplify foreground-background segmentation
- Gathering more information about the object from different viewpoints (fig. 3)
- Autonomous viewpoint selection to solve ambiguity (*e.g.,* Nolfi 1996)
- Some perspectives are easier to learn and recognize than others

Research Questions

- Which viewpoint invariant features can be used?
- How should the 3D objects be represented? Which features distinguish between objects?
- How can a robot learn to actively change its viewpoint in order to overcome ambiguity?

Method

We will develop methods for a **mobile robot** equipped with a camera and a robotic arm for object manipulation.

Why a robot?

- Possibility to explore objects.
- Dealing with the real world requires robust methods.
- Long term goal: Application in personal robotics.

Techniques

We will **not** use **complete 3D models** and computational exhaustive calculations. We believe that **exploration simplifies** the object recognition task. Therefore, we will use techniques like:

- Scale and rotation invariant keypoints (Lowe 2004)
- Viewpoint selection (Deinzer *et al.* 2003)
- Discrimination games (Steels 1997)

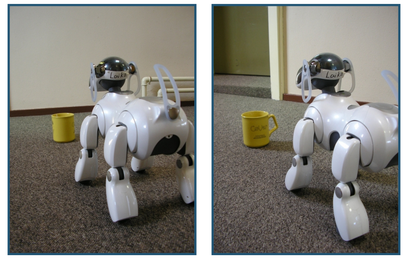

Figure 3: Exploration of the object

Towards a complete system

1. Extracting features from a sequence of images of an object being explored.
2. Making a representation of the object based on the features + a comparison function.
3. Close the loop between object recognition and exploration: calculate actions to disambiguate objects.
4. Add a learning component: Create representations from examples and refine existing ones based on perceived ambiguities + learn the manipulations to distinguish objects

Related Research

- Learning about objects through manipulation, and active segmentation at CSAIL MIT (*e.g.,* Fitzpatrick 2003)
- Developmental approach to autonomous exploration of the world in the Babybot project at LIRA-Lab, Genua (*e.g.,* Natale *et al.* 2005).
- Robot learning of categories (Steels & Kaplan 2000)

References

Deinzer, F., Denzler, J. and Niemann, H. (2003) Viewpoint Selection Planning Optimal Sequences of Views for Object Recognition. In: *Proc. Of Computer Analysis of Images and Patterns*, Groningen, The Netherlands, August 2003.

Fitzpatrick, P (2003) From First Contact to Close Encounters: A developmentally deep perceptual system for a humanoid robot. PhD thesis at MIT, 2003.

Lowe, D. G. (2004) Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision,* **60**, 2 (2004), pp. 91110.

Natale, L., Metta, G and Sandini, G. (2005) A developmental approach to grasping. In: *Developmental Robotics AAAI Spring Symposium,* 2005. Stanford, CA, March 2005

Nolfi S. (1996). Adaptation as a more powerful tool than decomposition and integration. In: T.Fogarty and G.Venturini (Eds.), *Proceedings of the Workshop on Evolutionary Computing and Machine Learning, 13th International Conference on Machine Learning*, University of Bari, Italy

Steels, L. (1997) Constructing and Sharing Perceptual Distinctions. In: van Someren, M. and Widmer, G. (eds) *Proceedings of the European Conference on Machine Learning, ECML'97*, Lecture Notes in Computer Science, vol. 1224, pp. 413, Berlin: Springer-Verlag.

Steels, L. and Kaplan, F. (2000) AIBO's first words, the social learning of language and meaning. *Evolution of Communication 4(1)*