Perceptual Abstraction for Robotic Cognitive Development

DRIANCOURT remi
Intelligent Robot Lab., University of Tsukuba, JAPAN
driancourt@roboken.esys.tsukuba.ac.jp

Abstract

We are concerned with the design of a developmental robot that learns from scratch simple models about itself and its surroundings. A particular attention is given to perceptual abstraction from high-dimensional sensors.

1. Introduction

The hand-programming of autonomous robot is a tiresome process. In unstructured environments, it may be impossible for a designer to select relevant features in regard to the robot's capabilities. Of particular interest would be an "intelligent" system able to autonomously acquire more adapted representations and controls...a little bit like biological organisms naturally do during their cognitive development. The design of such an entity is a fundamental question in Artificial Intelligence, and it brings in its tray many debates in philosophy around themes like "Symbol grounding" or "consciousness" (Ziemke, 2001).

Our aim is to design a robotic system that emulates such a cognitive development. As such, our research belongs to the recent field of Developmental or Epigenetic Robotics [(Zhang and Weng, 2002), (Zlatev and Balkenius, 2001)]. Though sharing the same inspirations and vocabulary of "schemas", "stages" or "biological motivation", our research may be characterized by its object-oriented approach and the attention given to high-dimension sensors.

2. Problem setting

Though the goals of this research are general, we consider a simplified "ecological niche", where the robot is the only agent of change. All actions induce sensory feedback and are reversible. The agent recognizes innately an obstacle as a state where actions do not bring internal feedback. Practically, we use a wheeled mobile robot with high-dimension redundant sensory inputs: vision and laser range sensor. Though we intend to use soon real robots, our first experiments were accomplished using a simulator (written in Java).

The robot has three internal drives: curiosity (exploration), avoid pain (obstacles) and seek pleasure (recompense). A recompense is given by a human trainer when the robot is close to a visible goal landmark. The robot should therefore be motivated to navigate towards these landmarks while avoiding obstacles. This supposes the development of mental structures for the characterization and recognition of obstacles and goals concepts and controls to act on them and naviguate.

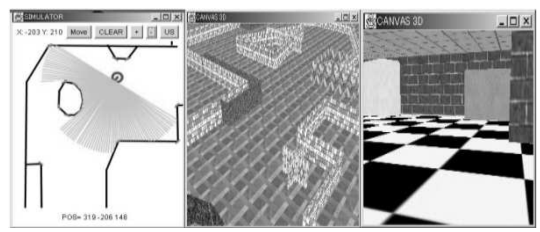

Figure 1: Pictures from our simulator environment

3. Development Plan

Very classically, we took inspiration from constructivist theories to propose stages of development during which the robot progressively constructs hierarchies of perceptual concepts and refines the motor schemas that manipulates them.

1. Calibration and self-organization of its sensory space so that change can be defined as motion.
2. Coarse perceptual abstraction and categorization. The robot "play" with its basic reflex actions to discover their general effect.
3. From reliably observed state/action relations, characterization of obstacle and goal states. Reflex actions with a one-time step visibility for avoidance.
4. Assimilation of reactive control policies to prolong certain interactions. The robot stays away from obstacles and is attracted by goals.
5. Proto-intentionality: the robot learns control laws for manipulating visible objects from one configuration to another.
6. Intentionality and object permanency. The robot can build a world model and plan its actions.

4. Approach

To represent abstraction hierarchies we use an object-oriented approach using "concepts" as basic bricks and cluster analysis as a unifying principle to abstraction from low-level sensory processing to higher level cognition. We define a perceptual state as a union of concepts, each defined by a list of attributes (e.g. raw perceptual data is $U_n O_n(v, p)$ with v measurement and p sensor reference). The goal of abstraction is to acquire an adequate low-dimension representation of the stimulus: compress the numerous raw data in a few high-level concepts.

We distinguish two types of bottom-up abstraction. 1. The redescription of actual perceptual data in a more suitable "summary" representation, and 2. the categorization of objects traceable through time for later recognition. A top-down process characterizes obstacle and goal states from defined concepts. Concepts can be dynamically linked by "is-a" (belongs to a category), "part-of" (belongs to a larger group of parts) and "related-to" (possibility of passage from one concept to another by change or action) links. Each cluster abstract underlying data by a statistical summary.

5. Experimental results

Up to stage 2, our first experiments aim at the abstraction of perceptual information from laser sensor and vision, without considering the grounding of concepts in action.

Having in mind the Gestalt grouping criteria (Elder and Goldberg, 2002) of proximity, continuation and common fate, we first enrich raw data with new attributes. We determine proximity from the idea that "close" sensors should produce similar measures. The method was inspired from (Pierce, 1997) and makes use of Multidimensional Scaling. For continuation, a codebook of "local shape" feature vectors was learned in an unsupervised manner from stimulus statistics using Independant Component Analysis (Hyvarinen and Oja, 1999). Common fate -or motion- of raw data objects is determined by a global optical flow method based on Dynamic Programming (Bellman, 1997). This method can produce quickly approximate dense motion fields.

Perceptual abstraction is done with a fast agglomerative clustering method. The decision to group together objects in a more abstract cluster is classically done by comparing a similarity measure with a threshold. The originality of our method is that this distance is learned from the environment. The distance between two attributes A and B is the probability $p_{A/B}$ for a concept of attribute A to have a neighbor of attribute B. A table of distance can be easily learned online and accessed during clustering. Clusters of raw data are characterized by a statistical summary: mean values and distribution histograms (correlation will be considered in a future work). Distance between clusters is a cross product on the histograms of the previous distances: $\sum_i \sum_j h(i) * h(j) * p_{i/j}$ with h(x) size of bin in histogram. In a few iterations meaningfull data groups can be extracted (fig.2).

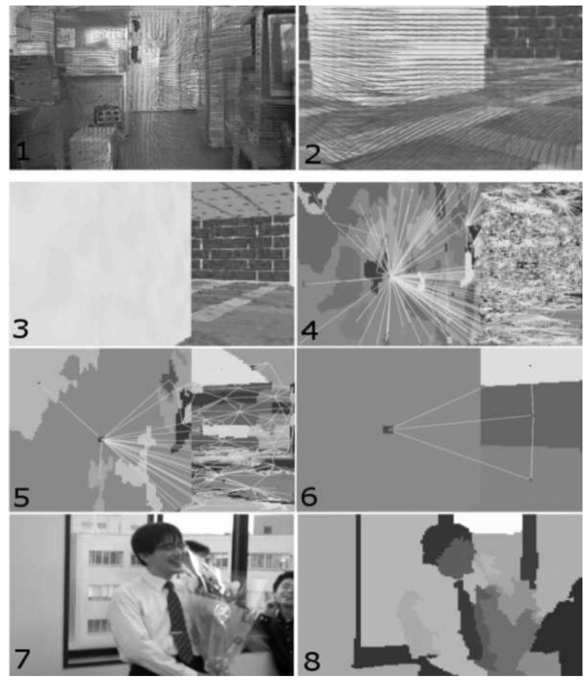

Figure 2: 1,2: Motion flow with real and simulator images. 3: raw image level. 4: 1st abstraction level. 5: 2nd level. 6: 4th level. Graphs show neighborhood relations. 7,8: real image and segmentation results

References

Bellman (1997). Dynamic programming. *Princeton University Press, Princeton, New Jersey.*

Elder and Goldberg (2002). Ecological statistics of gestalt laws for the perceptual organization of contours. *Journal of Vision, 2(4), 324-353.*

Hyvarinen, C. and Oja (1999). A fast algorithm for estimating overcomplete ica bases for image windows. *Proc. Int. Joint Conf on Neural Net.*

Pierce, K. (1997). Map learning with uninterpreted sensors and effectors. *A.I. Journal 92.*

Zhang and Weng (2002). Developmental robots: a new paradigm. *Proc. EPIROB 2002.*

Ziemke (2001). The construction of 'reality' in the robot. *Foundations of Science, 6(1), 163-233.*

Zlatev and Balkenius (2001). Why "epigenetic robotics"? *Proc. EPIROB 2001.*