Motivational principles for visual know-how development

Frederic Kaplan and Pierre-Yves Oudeyer
Sony Computer Science Laboratory, 6 rue Amyot, Paris, France
kaplan@csl.sony.fr, py@csl.sony.fr

Abstract

What dynamics can enable a robot to continuously develop new visual know-how? We present a first experimental investigation where an AIBO robot develops visual competences from scratch driven only by internal motivations. The motivational principles used by the robot are independent of any particular task. As a consequence, they can constitute the basis for a general approach to sensory-motor development.

1. Introduction

One of the challenges for research in epigenetic robotics is to find general principles to design robots capable to extend their sensory-motor competences during their lifetime. These robots usually start with crude capabilities for perception and action and try to bootstrap new know-how based on their experience. Several researchers have investigated how some particular competence can emerge using a bottom-up mechanism (e.g. (Andry et al., 2001, Metta and Fitzpatrick, 2002, Tani, 2002)). A possible approach consists in defining a reward function adapted to the behavior that the robot has to develop. Several state-of-the-art techniques in machine learning show how a robot can learn how to behave in order to maximize such a function (Kaelbling et al., 1996). But in most cases, this reward function is specific to the task the robot has to learn. It means that for each new behavior to be developed, the designer has to define a new reward function. In this paper we discuss the design of motivational principles that would be independent of a particular task and that could be used, as a consequence, for any sensory-motor development. Despite its relative simplicity, it can be argued that the architecture we present can overcome several limitations of current epigenetic artificial systems (as recently reviewed by (Zlatev, 2002)).

The paper focuses on a mechanism for bootstrapping a simple active vision system. In the first months of their life, babies develop sensory-motor competences almost from scratch to localize lights sources, pay attention to movement and track moving objects (Smith et al., 1998). The robotic model presented in this paper does not attempt to model precisely this developmental pathway but to illustrate how general motivational principles can drive the bootstrapping of such competences. The rest of the paper presents our developmental architecture and experimental results on its use for developing visual know-how.

2. An architecture for self-developing robots

2.1 Presentation of the problem

The AIBO ERS-210, Sony's four-legged robot, is equipped with a CCD camera and can turn its head in the pan and tilt directions (a third degree of liberty exists but is not exploited in this experiment). We have deliberately simplified the vision system to an extreme point. The robot extracts from each image it analyses the point of maximum intensity. The visual system perceives only the coordinates of this maximum (i_{dpan}, i_{dtilt}) expressed relative to the image center. The robot also perceives the position of its head in a pan-tilt coordinates system (h_{pan}, h_{tilt}). At each time step its perception can be summarized by a vector of dimension four.

$$S(t) = \begin{vmatrix} i_{dpan}(t) \\ i_{dtilt}(t) \\ h_{pan}(t) \\ h_{tilt}(t) \end{vmatrix} \tag{1}$$

The robot moves its head by sending motor commands (m_{dpan}, m_{dtilt}). So the sensory- motor vector $SM(t)$ at each time step is of dimension 6.

$$M(t) = \begin{vmatrix} m_{dpan}(t) \\ m_{dtilt}(t) \end{vmatrix} \tag{2}$$

$$SM(t) = \begin{vmatrix} m_{dpan}(t) \\ m_{dtilt}(t) \\ i_{dpan}(t) \\ i_{dtilt}(t) \\ h_{pan}(t) \\ h_{tilt}(t) \end{vmatrix} \tag{3}$$

Initially the robot does not know anything about its sensory-motor device. Can the robot develop a simple attention behavior in which it intentionally fixes its gaze on a certain number of things in its environment? To do this, it must discover the structure of several couplings in its sensory-motor device.

- How does a relative command (m_{dpan}, m_{dtilt}) affects the next position of (h_{pan}, h_{tilt}) of the head? This sensory-motor coupling is constrained by the head limit positions resulting of the structure of the robot's body.
- How does a relative command (m_{dpan}, m_{dtilt}) affects the movement of the visual field in particular the position of (i_{dpan}, i_{dtilt}).This sensory-motor coupling is again constrained by the robot's body and also by the structure of what happens in the environment.

In short, the robot must learn to perceive its environment by moving its head in the right manner. The developmental mechanism that we describe is only driven by a set of internal motivational variables. We claim that the dynamics resulting from these motivational variables are sufficient to lead the robot into a continuous increase of its sensory-motor mastery.

2.2 Overview of the architecture

The architecture of a self-developing device can be schematized by the interaction of three processes (Figure 1).

- The *Motivation process* is responsible for the evaluation of a given sensory-motor situation. A set of *motivational variables* $Motiv(t) = \{mot_i(t)\}$ is defined and associated with a set of *reward functions R*. A situation is desirable if it results in important rewards. An important feature of self-developing devices is the use of task-independent motivation variables. These variables typically result of internal computations based on the behavior of the two other processes (Prediction and Actuation). This process is used to evaluate anticipated situations and plays a role in the actuation process.
- The *Prediction process* tries to predict the evolution of the sensory-motor trajectories. It uses three prediction devices dedicated respectively to the prediction of $M(t)$, $S(t)$ and $Motiv(t)$. All the knowledge the device has about its environment, its "awareness", is resulting from these prediction devices.
- Eventually, the *Actuation process* decides based on the state of the two other modules which action should be performed in order to obtain rewards. This process goes through four phases : (a) Generation of possible motor commands, (b) Anticipation of the corresponding sensory-motor trajectories (using the Prediction process), (c) Evaluation of each simulated trajectories of the corresponding expected rewards (using the Motivation process) and eventually (d) Selection of the best motor commands.

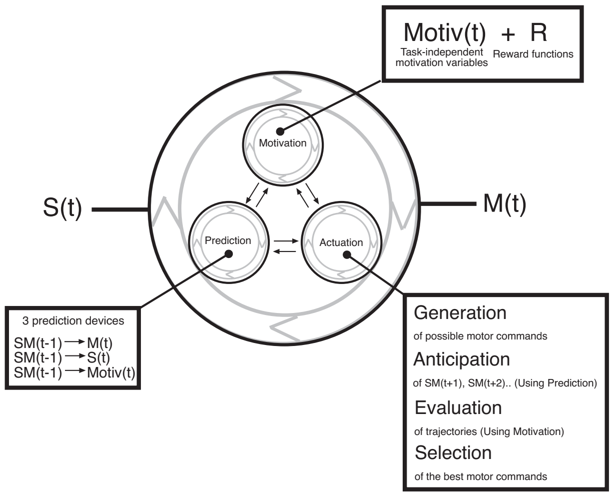

Figure 1: The architecture of a self-developing device

The three processes evolve based on the experiences of the agent. What the agent is aware of, what it is motivated for and the way it acts on its environment changes over time as the result of its developmental trajectory. The rest of the section goes into more details about each of these processes.

2.3 Motivation

The motivation process is based on a set of motivational variables mot_i. We have tried to design a set of motivations that are independent of the particular sensory-motor device that the system explores. Being rather abstract they can be used to drive the mastery of any sensory-motor device. In order to create the condition for an open-ended sensory-motor exploration, we have chosen variables which value depends on the developmental history of the robot. This means that the way to receive rewards for such motivations is constantly changing as the robot develops. Here are the three kind of variables used by the system described in this paper.

- Predictability: Can the robot predict the current sensory context $S(t)$ based on the previous sensory-motor context $SM(t-1)$? The robot is equipped with a prediction device that tries to learn sensory-motor trajectory. If $e(SM(t-1), S(t))$ is the current error for predicting $S(t)$, the predictability $P(t)$ can be defined as :

$$P(t) = 1 - e(SM(t-1), S(t)) \quad (4)$$

- Familiarity: Is the sensory-motor transition that leads from $SM(t-1)$ to $S(t)$ a common pathway ? The robot is equipped with a device evaluating the frequency of the sensory-motor transition for a recent period $t-T$. If $f_T(SM(t-1), S(t))$ is the current frequency of the transition that leads to S(t), the familiarity $F(t)$ can be defined as :

$$F(t) = f_T(SM(t-1), S(t)) \tag{5}$$

- Stability: Is the current sensory variable s_i of $S(t)$ far from its average value? The robot tracks the average value $< s_i >_T$ for the recent period $t-T$. So for each sensory variable s_i the stability $\sigma_i(t)$ can be defined as :

$$\sigma_i(t) = 1 - \sqrt{(s_i - < s_i >_T)^2} \tag{6}$$

Predictability and Familiarity share some similarities with internal variables experimented by other researchers like "novelty" (Huang and Weng, 2002) or "curiosity" (Kulakov and Stojanov, 2002). More generally the study of such kind of general basic motivation can be traced back to Piaget's research (Piaget, 1937).

For our problem, we have motivational vector of dimension 6.

$$Motiv(t) = \begin{vmatrix} P(t) \\ F(t) \\ \sigma_{idpan}(t) \\ \sigma_{idtilt}(t) \\ \sigma_{hpan}(t) \\ \sigma_{htilt}(t) \end{vmatrix} \tag{7}$$

Each motivational variable v is associated with a *reward function* $r(v,t)$. It takes the following general form:

$$r(v,t) = f_t(v(t), v(t-1), v(t-2), ..) \tag{8}$$

In the current implementation two kinds of functions are used.

- The robot is rewarded when it *maximizes* the value v of the *stability* motivations. This is similar with the way motivational variables are generally treated (e.g homeostatic models in (Breazeal, 2002)).

$$r_{max}(v,t) = v(t) \tag{9}$$

- But for *predictability* and *familiarity*, the robot tries to experience *increases* of the value of the variable instead of maximizing it. This means it does not look for predictable or familiar situations. It seeks "learning" experiences (predictability) and "discovery" situations (familiarity). As we will see, this small difference plays an important role for the dynamics of the system.

$$r_{inc}(v,t) = \begin{cases} (v(t) - v(t-1)) & : \quad v(t) > v(t-1) \\ 0 & : \quad v(t-1) \geq v(t) \end{cases} \tag{10}$$

A parameter α_i is associated to each motivational variable. It enables to specify the relative weight of each variable for determining the overall reward of vector $Motiv(t) = \{mot_i(t)\}$.

$$R_(Motiv(t)) = \sum_{mot_i} \alpha_i . r(mot_i, t) \tag{11}$$

2.4 Prediction

The awareness of the robot comes from its ability to predict sensory-motor trajectories. Recognizing a situation is recognizing a sensory-motor pathway. This standpoint follows the lines of current research that considers that perception emerges from motor actions (Gibson, 1986, Varela et al., 1991, OŔegan and Noe, 2001). This view, also known as active perception, is now shared by a growing number of robotic engineers (e.g. (Marocco and Floreano, 2002, Metta and Fitzpatrick, 2002)).

We can considerer that at a given time t, a robot experiences a particular sensory-motor context, that can be summarized in vector $SM(t)$ of dimension 6. The system uses three *prediction devices*: $\Pi_m, \Pi_s, \Pi_{motiv}$. The three devices take the current situation $SM(t)$ as an input and try to predict respectively the future motor situation $M(t+1)$, the future sensory situation $S(t+1)$ and the future state of the motivation vector $Motiv(t+1)$.

At each time step , the three devices learn the correct prediction by comparing the current situation with the previous one.

$$\Pi_m(SM(t-1)) \rightarrow M(t) \tag{12}$$

$$\Pi_s(SM(t-1)) \rightarrow S(t) \tag{13}$$

$$\Pi_{motiv}(SM(t-1)) \rightarrow Motiv(t) \tag{14}$$

The landscape of the motivation that Π_{mot} must learn is dependent on the performance of the two other devices. $P(t)$ is determined by the error rate of Π_s, and the other motivational variables change according to the action selection process which in turn results form the prediction of Π_m and Π_s (see below). As a consequence, Π_{mot} must adapt continuously during the bootstrapping process.

For this study we tried two kinds of implementation for the prediction devices:

- A recurrent Elman neural network with a hidden layer / context layer of 12 input nodes

(Elman, 1990). Because this network is recurrent, it predicts its output based on the value of the sensory-motor vectors several time steps before t.

- A prototype-based prediction system that learns prototypic transitions and extrapolates the result for unknown regions. It takes the form of a set of vectors associating a static sensory-motor context $SM(t-1)$ with the predicted vector ($M(t)$,$S(t)$ or $Motiv(t)$). New prototypes are regularly learned in order to cover most of the sensory- motor space. The prediction is made by combining the results of the k closest prototypes. This prediction system is faster and more adaptive than the Elman network, but may be less efficient for complex sensory-motor trajectories.

The performances of the prediction devices are crucial for the system, but the architecture does not assume anything about the kind of devices that need to be used. As a consequence, any state-of-the-art techniques can be tried. For the problem we tried to tackle, the dynamics were roughly the same for the two kinds of prediction devices. The results presented in this paper are obtained with the prototype-based prediction system.

2.5 Actuation

The actuation process anticipates the possible evolutions of the sensory-motor trajectories and tries to choose the motor commands that should lead to the maximum reward. Several techniques taken from the reinforcement learning literature can be used to solve these kinds of problems(Kaelbling et al., 1996). In our system, the process can be split into four phases:

- Generation : The system constructs a set of possible motor commands $\{mi\}$. This phase can be trivial for simple cases but may require special attention when dealing with complex actuators.
- Anticipation : The system simulates the possible sensory-motor evolution $\{SM_{mi}\}$ over T time steps using the prediction devices in a recurrent manner. The system combines the result of both Π_m and Π_s to predict future sensory-motor situations and uses Π_{motiv} to predict the evolution of the motivation vector $Motiv(t)$.
- Evaluation : For each evolution $\{SM_{mi}\}$ an expected reward R_{mi} is computed as the sum of all the future expected rewards.

$$R_{mi}(t) = \sum_{j=t}^{t+T} R(Motiv(j)) \qquad (15)$$

- Selection : The motor command $\{mi\}$ corresponding to the highest R_{mi} is chosen.

3. Isolation of the dynamics in simulation

3.1 Simulated environment

The developmental dynamics of such an architecture can be rather complex. In order to better understand the role of each internal motivation we have conducted a series of experiments in a simple simulated environment. We simulate the presence of a light performing a sinusoidal movement in the environment.

$$light_{pan}(t) = K * sin(p(t)) \qquad (16)$$

$$light_{tilt}(t) = L * sin(p(t) + \beta) \qquad (17)$$

$$p(t+1) = p(t) + \delta \qquad (18)$$

The oscillations in the tilt domain have a smaller amplitude than in the pan domain ($L < K$).

The robot perceives the relative position of the light compared to its own position.

$$i_{dpan}(t) = light_{pan}(t) - h_{pan}(t) \qquad (19)$$

$$i_{dtilt}(t) = light_{tilt}(t) - h_{tilt}(t) \qquad (20)$$

At each time step it decides the most appropriate action $\{m_{dpan}, m_{dtilt}\}$ to perform. The effect of this action is simulated using the following simple rules :

$$g_{pan}(t+1) = m_{dpan}(t) + h_{pan}(t) \qquad (21)$$

$$g_{tilt}(t+1) = m_{dtilt}(t) + h_{tilt}(t) \qquad (22)$$

The constraints on the robot's body are simulated by imposing limits on the possible head positions: max_{pan}, min_{pan}, max_{tilt}, min_{tilt}.

$$h_{pan}(t+1) = \begin{cases} max_{pan} & : \quad g_{pan}(t+1) > max_{pan} \\ min_{pan} & : \quad g_{pan}(t+1) < min_{pan} \\ g_{pan}(t+1) & : \quad otherwise \end{cases} \qquad (23)$$

A similar equation is defined for $h_{tilt}(t+1)$.

3.2 Increase in predictability

For this experiment, we assume that the robot is only driven by its predictability motivation. It tries to experience increases in its predictability level $P(t)$ which means that it seeks for "learning" situations. As it learns, sensory-motor trajectories that used to give rewards tend to be less interesting. This dynamics push robot towards an open-ended dynamics of exploration.

Figure 2 shows the evolution of the average predictability level $P(t)$. It quickly reaches a high value. This shows that the robot has learned the overall effect of its movement on the light position and on the position of its own head. As the robot tries to experience *increases* in predictability and not simply to maximize it, small oscillations can be seen near the maximum value. They correspond to new sensory-motor trajectories that the robot explores.

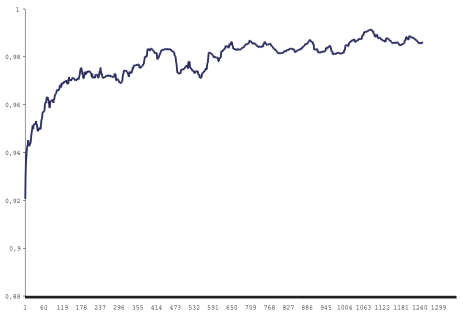

Figure 2: Evolution of the average of the predictability level $P(t)$

Figure 3 shows the evolution of the pan position of the head during 1000 time steps. The corresponding evolution of $light_{pan}$ is also indicated. A very similar curve can be plotted for the tilt dimension. The movement is rather complex as the robot gets away from predictable sensory-motor trajectories and tries to explore new ones. The evolution of the average h_pan position shows that the system progressively explores the amplitude of the possible pan positions by oscillating around the zero position.

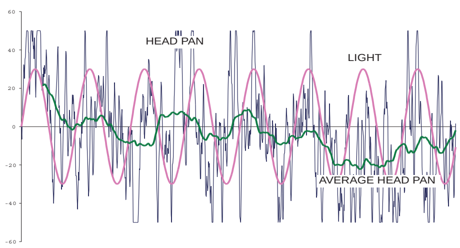

Figure 3: Evolution of the h_{pan} position (and its average) following the increase predictability rule. The evolution of $light_{pan}$ is also indicated

3.3 Increase in familiarity

For this experiment, the robot is driven only by its familiarity motivation. It tries to experience increases in its familiarity level $F(t)$. In a similar way than for predictability, unfamiliar situations tend to become familiar after a while and, as a consequence, less rewarding. This dynamics drives the robots into a continuous exploration behavior.

Figure 4 shows the evolution of the average familiarity level $F(t)$. The robot manages progressively to reach a very high level of familiarity. Similarly to the evolution of the previous experiment, we see oscillations due to the pressure of experiencing *increases* in familiarity. Each reduction of the familiarity level corresponds to the exploration of new parts of the sensory- motor space.

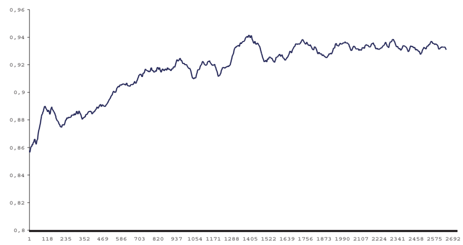

Figure 4: Evolution of the average of the familiarity level $F(t)$

Figure 5 shows the evolution of the pan position of the head during 1000 time steps. The movement looks a bit like the one obtained in the previous experiment but some differences can be noticed. The average position curve shows the robot first explored position corresponding mostly to high pan values then switched progressively to low pan values. This switch, that seems to occur independently of the oscillation of the light, did not appear as clearly as in the experiment on predictability. The familiarity motivation pushes the robot to explore trajectories in the sensory- motor space independently of how well it masters them. At the end of the experiment, the system has covered the entire set of possible pan positions. The familiarity and predictability motivations can be seen as two complementary ways to explore a sensory-motor device.

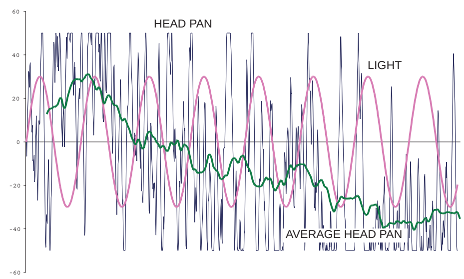

Figure 5: Evolution of the h_{pan} position (and its average) following the increase familiarity rule. The evolution of $light_{pan}$ is also indicated

3.4 Maximization of sensory stability

The last four motivational variables concern the stability of each component of the sensory vector $S(t)$. They are all associated with the *maximize* reward function r_{max}.

3.4.1 Head stability

First we will consider the case where the stability concerns the head position. It corresponds to the variables $\sigma_{hpan}(t)$ and $\sigma_{htilt}(t)$. It means that the robot seeks sensory-motor trajectories in which its head position remains stable in time. Figure 6 shows the evolution of average stability for an experiment when the robot uses this reward system. In this context the task is rather easy: the robot simply has to discover that it has to stop moving its head in order to obtain important rewards. Stability is reached rapidly for both the pan and tilt direction.

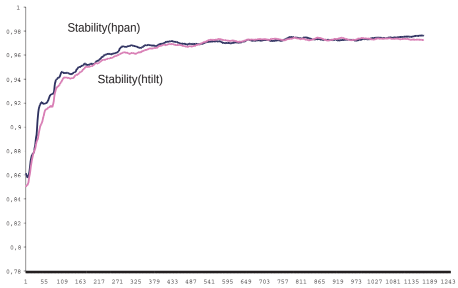

Figure 6: Evolution of the average of the stability level for the head position $\sigma_{hpan}(t)$ and $\sigma_{htilt}(t)$

The evolution of figure 7 shows that the head position stabilizes around its initial position after a short period of oscillation.

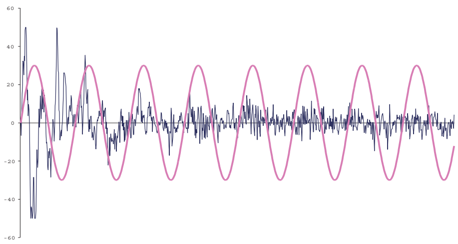

Figure 7: Evolution of the h_{pan} position following the maximization stability rule for the head position. The evolution of $light_{pan}$ is also indicated

3.4.2 Light stability

We now consider the case where stability concerns the relative position of the perceived light. The task is in this case a bit more complex as the light is not directly controlled by the robot. The robot has to discover that it can act upon it by moving its head in the appropriate directions. Figure 8 shows the evolution of the average stability for an experiment with this reward system. The robot manages to control the stability of the light in the tilt domain faster than in the pan domain probably because the movement has a smaller amplitude in the tilt domain ($L < K$).

Figure 9 shows the evolution of the head position during the same experiment. After a short time for tuning, the robot develops a tracking behavior and follows the light quite precisely. As the robot seeks

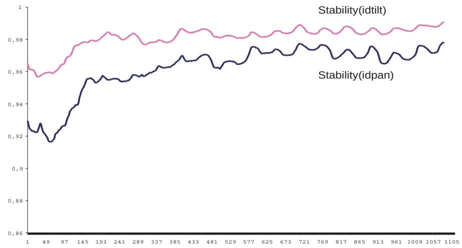

Figure 8: Average evolution of the stability level for the light relative position $\sigma_{idpan}(t)$ and $\sigma_{idtilt}(t)$

for sensory stability, each movement of the light can be seen as a perturbation that it learns to compensate. The development of this visual know-how results directly from the effect of the environment on the sensory-motor device.

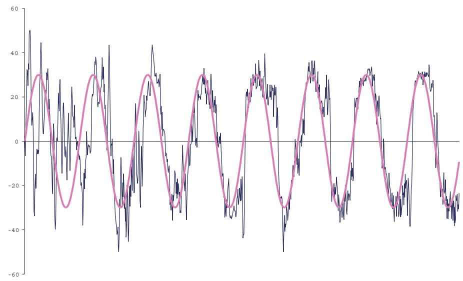

Figure 9: Evolution of the h_{pan} position following the maximization stability rule for the light relative position. The evolution of $light_{pan}$ is also indicated

With this series of experiments, we have a clearer idea of the effect of each reward system on the bootstrapping process. The two first motivations, increase in predictability and familiarity, push the robot to explore its sensory-motor device. The last four ones, maximization of sensory stability, lead the robot, on the one hand, to stop moving its head, and on the other hand, to develop a tracking behavior.

4. Experiment on the robot

This last experiment is conducted on one AIBO ERS-210. The software components are written in C++ using the publicly available OPEN-R SDK. The software runs on board, and the data for the experiment are directly written on the MemoryStick for later analysis. In this experiment we are using a small number of the degrees of freedom possessed by the robot. Nevertheless, the fact that the architecture can be used on a real robot shows that it is sufficiently light to perform on-line learning in real-time on a modest computer and that it is sufficiently robust to cope with noise on both sensory data and motor commands.

At each time step, the robot computes the point of maximum light intensity in its visual field. The relative position of this point provides the two inputs $i_{dpan}(t)$ and $i_{dtilt}(t)$. The robot measures its own head position $h_{pan}(t)$ and $h_{tilt}(t)$. Contrary to the simulation, this measure is not completely accurate. In the same way, due to different mechanical constraints, the relative movement resulting from the action $m_{dpan}(t)$ and $m_{dtilt}(t)$ can be rather noisy.

The reward system used can potentially include the six motivational variables previously studied. As we mentioned, the relative weight of each variable of the computation of the overall reward is determined by the set of parameters α_i. For this experiment, we set these weights so that the robot developed the know-how for paying attention to the different light patches present in its environment. This means it should develop a tracking behavior but also an exploratory skill for not being stuck in front of a given light. As head stability is to some extent counterproductive for such a goal, we decide that $\sigma_{hpan}(t)$ and $\sigma_{htilt}(t)$ should not play a role for this experiment. As a consequence, all the reward functions were associated with the same weight $\alpha_i = k$, except the two controlling the head stability that received the value $\alpha_i = 0$[1].

The experiments lasted 10 minutes. The robot was placed in front of an uncontrolled office setting. Figure 10 shows the evolution of the six motivational variables. As expected the four variables associated with the weight k obtained high values. The relative position of light reached rapidly a plateau, but predictability and familiarity kept increasing. The motivational variables for head stability oscillate at a lower level.

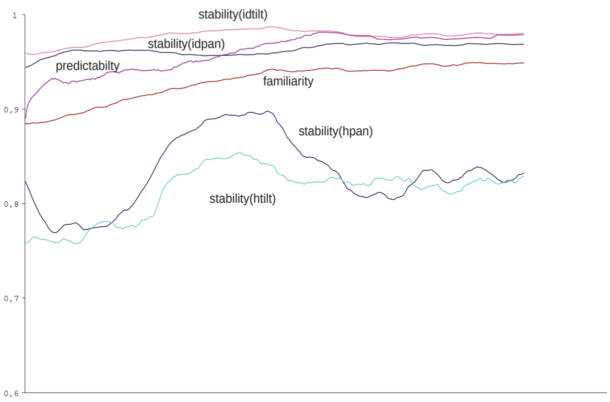

Figure 10: Evolution of six motivational variable for a 10 min experiment on the AIBO ERS-210

Figure 11 shows the evolution of head pan position during the experiment as well as the position of the perceived light. The robot seems to track the light, but motivated for exploration, its position oscillates around a local light maximum permitting the robot to find another one.

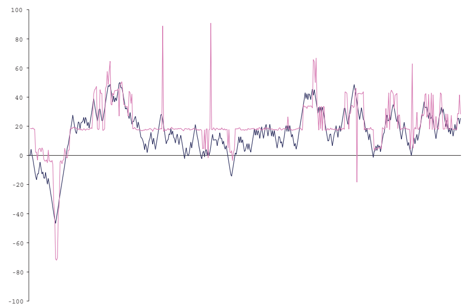

Figure 11: Evolution of the head pan position and of the perceived light position

This behavior can be seen more clearly on figure 12 which magnifies a detail in figure 11. The pan position increases to approach a local maximum, then oscillates around it for a while. At some point a larger oscillation makes it discover a higher local maximum. The robot switches back and forth several times between the two maxima and finally continues its exploration towards higher pan values. This kind of behavior is a typical result of the search of increase in predictability and familiarity. The robot uses familiar and predictable contexts as bases for progressively continuing its exploration.

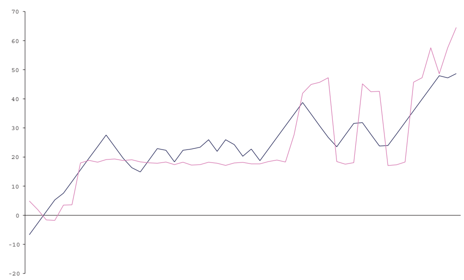

Figure 12: Magnification of a detail in figure 11

Figure 13 shows the overall pan tilt trajectory for the duration of the experiment. It appears that the robot has concentrated its exploration on the right part of the scene. It seemed to have highly explored one particular area and progressively search for other maxima in its immediate neighborhood. It result from this exploration a kind of "map" of the position of the local light maxima as shown on figure 13. This representation does not exist as such for the robot but is the result of the know-how it has developed with its sensory-motor device. The robot is not capable to perceive all these light positions at the same time, but it is confident that they are there because of its sensory-motor visual know-how. This kind of visual awareness can be seen as a technical illustration of what O'Regan and Noe call "the world as an outside memory" and the "impression of seeing

[1] It is possible to design another system that would control these weights automatically according to some predefined criteria. It all depends on the kind of general development strategies one wishes to observe

everything" (OŔegan and Noe, 2001).

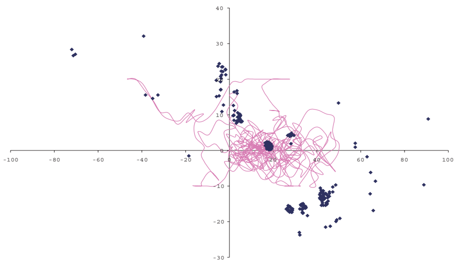

Figure 13: Pan tilt trajectory during the experiment and local light maxima identified.

5. Conclusion

We have illustrated how a robot can develop visual know-how driven by task-independent internal motivations. The experimental setup was deliberately simple in order to illustrate the basic dynamics of such a device. In further work, we will investigate how far the same set of motivational principles can account for an efficient exploration of other sensory-motor devices. We hope that these investigations will help us to define the characteristics of a general architecture that could account for open-ended sensory-motor development.

References

Andry, P., Gaussier, P., Moga, S., Banquet, J., and Nadel, J. (2001). Learning and communication in imitation: an autonomous robot perspective. *IEEE Transaction on Systems, Man and Cybernetics, Part A : Systems and Humans*, 31(5):431–444.

Breazeal, C. (2002). *Designing sociable robots.* Bradford book - MIT Press.

Elman, J. (1990). Finding structure in time. *Cognitive Science*, 14:179–211.

Gibson, J. (1986). *The ecological approach to visual perception.* Lawrence Erlbaum Associates.

Huang, X. and Weng, J. (2002). Novelty and reinforcement learning in the value system of developmental robots. In *Proceedings of the 2nd international workshop on Epigenetic Robotics - Lund University Cognitive Studies 94*, pages 47–55.

Kaelbling, L., Littman, M., and Moore, A. (1996). Reinforcement learning: A survey. *Journal of artificial intelligence research*, 4.

Kulakov, A. and Stojanov, G. (2002). Structures, inner values, hierarchies and stages : Essentials for developmental robot architectures. In *Proceedings of the 2nd International worksop on Epigenetic Robotics - Lund University Cognitive Studies 94*, pages 63–69.

Marocco, D. and Floreano, D. (2002). Active vision and feature selection in evolutionary behavioral systems. In et al, H., (Ed.), *From Animals to Animats 7*, Cambridge, MA. MIT Press.

Metta, G. and Fitzpatrick, P. (2002). Better vision through manipulation. In Prince, C., Demiris, Y., Marom, Y., Kozima, H., and Balkenius, C., (Eds.), *Proceedings of the 2nd international workshop on Epigenetics Robotics - Lund University Cognitive Studies 94*, pages 97–104.

OŔegan, J. and Noe, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioural and Brain Sciences*, 24(5).

Piaget, J. (1937). *La construction du reel chez lénfant.* Delachaux et Nieslte, Neuchatel et Paris.

Smith, P., Cowie, H., and Blades, M. (1998). *Understanding childrenś development.* Blackwell.

Tani, J. (2002). Articulations of sensory-motor experiences by forwarding forward model. In *From animals to animats 7*, Cambridge, MA. MIT Press.

Varela, F., Thompson, E., and Rosch, E. (1991). *The embodied mind : Cognitive science and human experience.* MIT Press, Cambridge, MA.

Zlatev, J. (2002). A hierarchy of meaning systems based on value. In *Proceedings of the 1st international workshop on Epigenetic Robotics - Lund University Cognitive Studies 85.*