Shaping of Robot Behaviors by Demonstrations

Louis Hugues
LIP6, Université Pierre et Marie Curie, Paris, France, Louis.Hugues@lip6.fr
Alexis Drogoul
LIP6, Université Pierre et Marie Curie, Paris, France, Alexis.Drogoul@lip6.fr

Abstract

This paper is concerned with the learning of robots behaviors in real environments. To face the constraints imposed by both physical and human spaces, it insists on the interest of a shaping process relying on learning by demonstrations. A mechanism for learning by demonstration is briefly described based on robot vision. The paper then discusses several general points related to learning by demonstrations, focusing particularly on practical issues.

1 Introduction

This paper is interested in the adaptation of robots to real environments by the way of learning by demonstrations.

Robots have sophisticated sensors and devices coming from their design process but no phylogenetic neither ontogenetic processes have occurred to really adapt them to our environments. Like any living being they have to deal with physical constraints, and moreover they have to be adapted to constraint of our human social space.

To fill this gap the methods like:

- *Learning by Demonstration* [Atkeson and Schall, 1997], [H.Friedrich and Dillman, 1995],
- *Learning By imitation* [Bakker and Kuniyoshi, 1996], [J.Demiris and Hayes, 1996], [G.Hayes and Demiris, 1994], [Gaussier et al., 1997]
- and *Supervised Learning* [Pomerlau, 1993]

take advantage of direct human demonstrations to obtain a better adaptation to physical and social spaces.

In the next chapter we present briefly our proposed technique for capture from demonstrations and reproduction of behaviors - a wider description can be found in [Hugues and Drogoul, 2001]. In a second part, starting from this technical proposal, we discuss more general points, always considering the practical issues.

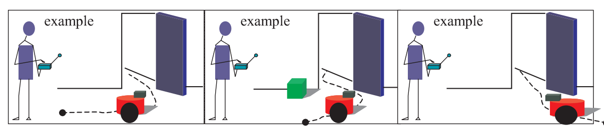

Figure 1: Three examples which may be used to shape the behavior ”Exit by the door”.

2 Behavior Capture

The proposed learning mechanism transforms a set of demonstrations (or examples) shown by a tutor into a synthetic behavior. This synthetic form will be used later by an autonomous robot so as to reproduce the behavior in situation. This *Behavior Capture* process uses as input the color video images seen by the robot and the value of its effectors (here wheels velocities).

To record an example the tutor controls the robot by a joystick (linked by radio communication) and produces a movie file containing video frames from robot camera and robot movements (effectors values). The movie files are first preprocessed and local properties $(prop_1, prop_2, ..., prop_p)$ are extracted from de perceptual field.

The capture mechanism relies on two data structures:

- The perception/actions relations are encoded into *a set of cells*.
- The context of the behavior is captured into *a context histogram*.

They are referred hereafter by $Cells$ and $Context$.

In the learning process, each example E_i is first recorded and captured separately. This produces a set of cells and an context histogram for each examples, denoted respectively by $Cells_i$ and $Context_i$. The final behavior is obtained by fusion of all $Cells_i$ and $Context_i$ into $Cells_{final}$ and $Context_{final}$ which can be used by the robot for real operation.

2.1 Cells Population

In the $Cells$ data structure the cells are organized along p dimensions, each dimension corresponding to a property. The properties used during experimentations where:

- x location in the image
- y location in the image
- color class property. Obtained by quantization in hue,saturation,value space.
- local density. For a given pixel, local density is the amount of neighboring pixels of the same color.This enables to distinguish several local configurations of a pixel neighborhood.

The capture of perception/action relations is obtain by projecting all video frames successively onto the cells structure. A cell is *activated* if all its properties $(prop_1, prop_2, ., prop_p.)$ are detected. The cell stores a statistical representation of the actions. This representation is the mean of the effectors (here wheels velocities) computed over periods of activation of the cell (Eq. 1).

$$action_{cell_i} = E\left[\overrightarrow{effectors}/activated(cell)\right] \quad (1)$$

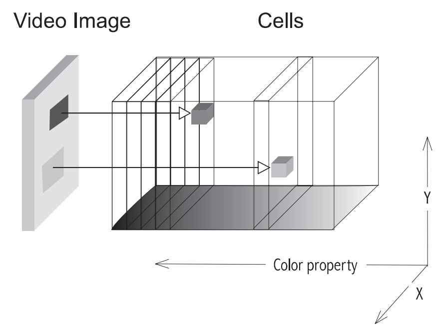

Figure 2: Cells data structure used to capture a behavior. The fourth dimension (neighborhood density) is not shown.

2.2 Context Histogram

The properties informations observed along an example, and more precisely the color information, are used to build $Context_i$ the context histogram of example E_i. $Context_i[color]$ represents the importance of *color* for the whole example. Every color is not taken into account and a measurement of color dispersion is used so as to favor colors that belong to compact and big objects. $Context$ gives the importance of each color for the behavior. The fusion of several $Context$ accentuates again those colors.

2.3 Fusion of examples

Each example is captured separately in $Context_i$ and $Cells_i$. All example are finally fused into $Context_{final}$ and $Cells_{final}$ by simple arithmetic operations. $Context_{final}$ is obtained by multiplying separate histograms to emphasize common features. $Cells_{final}$ is obtained by taking for each cell the mean of all corresponding cells.

2.4 Reproducing Behavior Autonomously

To reproduce the behavior in real time the robot camera image is projected again onto the $Cells_{final}$ structure. This determines a set of active cells from which current effectors values can be deduced. A majority scheme is used to determine the effectors values from this population of cells.

In the majority scheme, cells importance is ponderated by their corresponding $Context_{final}[color]$ so as to consider especially important features.

2.5 Experiment I : learning "approach object and stop"

The complete learning mechanism has been implemented on a robot pioneer 2DX running Linux and equipped with a color monoscopic camera.

In a first experiment we want the robot to approach a green box and then stop at approximately one meter of the box, Three examples have been recorded to learn the behavior. Like the one in Fig. 3,each one corresponds to a different pose. Once the behavior is active, the robot can be settled at various places, it then reproduces the behavior correctly if the box is sufficiently visible. The behavior works also if the green box is settled before another background.

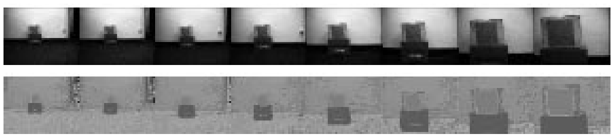

Figure 3: A first example (resumed) to learn "approach box and stop" behavior. Figure shows the original video and below the corresponding encoded video with color properties.

The Fig. 4 shows the response of $Cells_{final}$ to a video image . Intensity of pixels in response image corresponds to intensity of forward velocity stored in the cells. In response Fig. 4 a lot of pixels "tells" to go forward. A lot of indicates zero velocity but are not taken into account due to context histogram ponderation. In Fig. 5 much less forward intensity is visible and robot is near to stop.

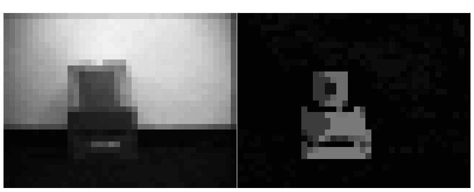

Figure 4: Forward velocity response (right) of the $Cells$ structure to the video image (left). Response is strong , the robot goes forward.

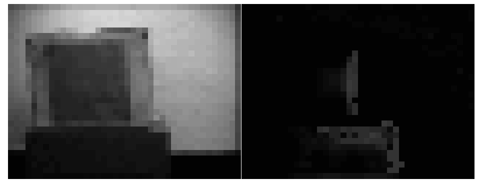

Figure 5: Forward velocity response (right) of the $Cells$ structure to a video image (left). Response is low , robot is near to stop.

2.6 Experiment II : learning "exit by the door"

In a second example we want the robot to learn to exit by a blue door. In this case four examples have been recorded. To test behavior reproduction the robot is settled at various poses near the door and if the door is visible, the robot is able to orientate, adjusts its direction and succeeds in 60 percents of trials. However the robot sometimes jams itself in the door embrasure, this shows that and avoidance reflex (ie: based on sonars) should be associated to obtain smoother navigation.
The figure 7 and 9 show the reactions cells population in two different situations. Each active cell is represented by the vector that it proposes.

3 Discussion

The capture mechanism described above permits to integrate several examples demonstrated from different poses. In the experiments , partial overlapping of current perceptions with previously seen examples is sufficient to generate appropriates movements. The final context histogram captures in a rough form some contextual features of the environment. Fusion highlights only the few colors supporting the behavior.

The capture mechanism has certainly to be improved and extended. However from this practical departure it is possible to envisage several directions.

3.1 Pedagogy

In a learning by demonstration scheme the human tutor has to interacts with the robot so as to improve robot's capabilities. This dialectical process suggests the use of some "pedagogical" tricks that will help to support and

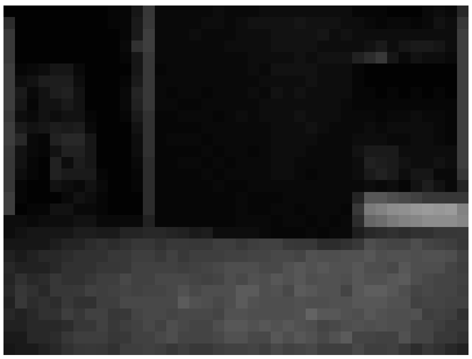

Figure 6: Robot Vision at frame ta

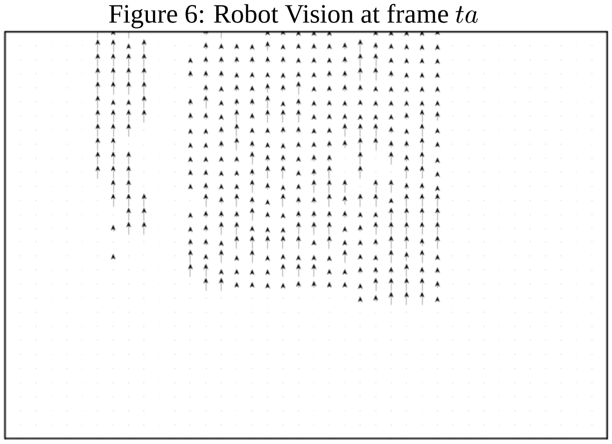

Figure 7: Reactions proposed by active cells at frame ta. Robot is far from the door and cells propose to go straight on.

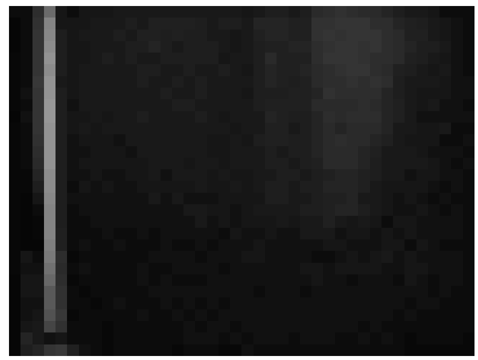

Figure 8: Robot Vision at frame tb

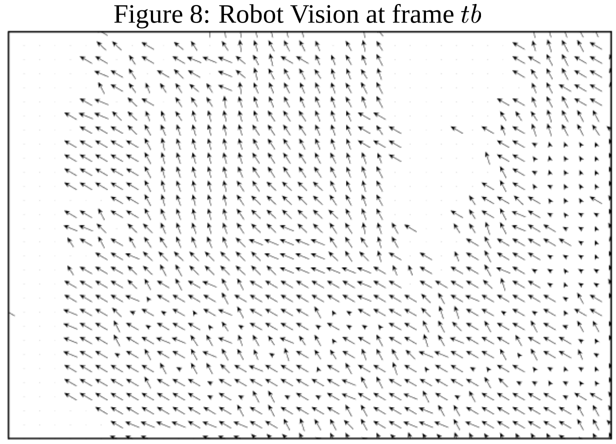

Figure 9: Reactions proposed by active cells at frame tb. Robot is just near the door and now cells propose to turn.

structure the learning process.

- *Representation of robot's inner functioning*: The human tutor should have a clear idea of what the robot is really able to do. The tutor should be able to *imaginate* how a robot can pass from one step to another, what features of the environment the robot will use to support and construct the behaviors. Thus the tutor should be provided with some idealized representation of robot's innermost functioning. It is important to notice that this representation can be entirely fake, its unique purpose is to help the tutor to establish some guide marks.
  With the presented technique the tutor might think for instance that compact colored objects are important, that things present in each examples are important also, and so on ...

- *Converging views* : The robot and human tutor live in totally different perceptual spaces. It is clear that from a perceptual and ability point of view a 50cm round box evolving at 200 mm/s and a human tutor don't share much in common. This has of course impact on the significance of demonstrations. However the experimental setup designed to conduct demonstrations can be tailored so as to bring the tutor nearer from the robot. This can be done by reducing the tutor's vision and controls via appropriate apparatus (computer displays, control device with feed back,etc ...).

- *Facets of user's intentions:*

  The objective of Learning by demonstrations methods is somehow to transmit the intention of the user into the robot's behavior. From every day experience we know that intention is a polymorphic and versatile object which is not so easy to tackle.

  In a simplified view the user's intention can be decomposed and cut into elementary facets. Facets are just minimal scenarii, which, put together, form the real behavior (ie: $facet_1$: approaching the door like that i would do that, $facet_2$: approaching it like this i would do this, $facet_3$) etc..)

  In the capture mechanism a facet is a single demonstration. The final purpose of the learning mechanism is to aggregates the facets and provide correct behavior for all the intermediate situations. Defining and showing the facets constitutes a practical pedagogy.

3.2 Context

The *context* where a behavior occurs is of great importance. In our thinking a behavior is not a succession of actions that can be performed anywhere.

First, a synthetic behavior has to be *independent* of irrelevant features of the learning context. For instance it is clear for a human observer that the box which appears in Fig. 1 in only one example should not be taken into account into the final behavior.

Beside this, the context independence has a counterpart which is context *detection*. The actions only make sense in precise contexts and thus the ability to recognize valid contexts is very important to trigger the appropriated behaviors.

The literature often refers to similar (or inverse) notion of perceptual aliasing in the framework of Markov's decisions processes. However for the learning methods that involve humans we prefer the notion of context. This notion clearly points out the deep differentiation work that a robot should ideally do to adapt itself to its environment. A robot has not to perform well in the most complicated case. It has to perform well in most of the cases that occurs really.

Robots do not have model of the world at their disposal and therefore to determine what a context is, is extremely difficult. In capture mechanisms we use a rough approaches suggesting that simple statistics computed over the perceptions provide valuable indicators for context identification. In [Hugues, 2000] we proposed how perceptions could be differentiated and classified using Kohonen's Features Map (SOMF). This was done for communication purpose but can also be used in the context identification problem.

3.3 Elementary affordances

The theory of Affordances proposed by J.J.Gibson [Gibson, 1986] suggests that a behavior is a complementary relation between an animal and its environment. The environment provides a support for what the animal can afford. For real living beings this relation is the product of a complex phylogenetic and eventually ontogenetic process. This approach can be (and has been) transposed somehow in robotics and robot behaviors can be thought of as elementary affordances. In an affordance point a view the robot is no more trying to pick-up/recognize perceptual features in the environment so as to conform to its running behavior. Inversely it is permanently keeped aware by the environment itself of what is possible to do.

In the Capture mechanism this point of view is used to generate actions from the flow of perceptions. For instance, in the presence of the blue door the robot relates directly the perceptions to possible actions. Robots is somehow ”impregnated” with passing of the blue door.

On the tutor side, the affordances point of view permits to conceived behaviors in an homogenous framework and eventually combinate them more easily than purposive behaviors.

3.4 Incremental learning

Psychology suggests that parts of past experiments are transformed and reused along the childhood. This *re-use* process can be somehow mimicked on robots side. Finding ways to reuse elementary behaviors in more complex situations simplifies the learning process and can generate complex behaviors made of robust parts.

This behavioral elements learned with the Capture mechanism could be reused in a more complex behavior.

- In a first phase, we can ”teach” a set of elementary affordances to a robot so as to form a first level that grounds the robot into its environment and its acting capabilities.

- In a second phase we can show to the robot more complex demonstrations involving the reuse of elementary affordances. The learning mechanism in this phase relies on the robot’s capacity to recognize effectively previous elements. A complex behaviors at this level can be think of as a *network of elementary affordances* where robot passes from affordances to affordances.

The capture mechanism can be extended in this direction. The candidates behaviors can be compared to parts of a complex demonstration by using effectors values proposed by $Cells$ structure and indications given by $Contexts$ matching. At this level, each demonstration is a possible path in the network of affordances. The objective of the learning process is to collect several paths and construct the network. This has great advantages over direct learning from complex demonstrations:

- the search-space is reduced to a few elements. Those elements point to what is really possible to do in the environment (from physical and social point of view).

- the tutor can build a pedagogy starting from those basic elements.

4 Conclusion - Future work

In this paper we have presented briefly our ongoing work dedicated to robots behavior learning by the way of vision-based demonstrations. We have then discussed some more general points related to this approach:

- A pedagogical point of view is quickly necessary when using learning by demonstrations.

- A behavior is largely determined by the contexts where it can be reproduced.

- Elementary affordances provides a way to conceive grounding of actions in physical and social space.

- Incremental learning should offer a simple way to increase the complexity of learned behaviors.

In future work we plan to investigate the following directions:

- Evaluation and extension of the Capture mechanism so as to deal with sequential and memory aspects of behaviors.

- Improvement and extension of the Context structure with multi-modal informations.

- Learning of complex and composed behaviors by incremental learning.

References

[Atkeson and Schall, 1997] Atkeson, G. and Schall, S. (1997). Robot learning from demonstration. In Fisher, J., editor, *Proceedings of the 1997 international Conference on Machine Learning (ICML'97)*.

[Bakker and Kuniyoshi, 1996] Bakker, P. and Kuniyoshi, Y. (1996). Robot see, robot do : An overview of robot imitation. In *AISB'96 Workshop on learning in Robots and Animals*.

[Gaussier et al., 1997] Gaussier, P., Moga, S., Banquet, J., and Quoy, M. (1997). From perception-action loops to imitation processes: A bottom-up approcha of learning by imitaion. *Applied Artificial Intelligence*, 1(7).

[G.Hayes and Demiris, 1994] G.Hayes and Demiris, J. (1994). A robot controller using learning by imitation. In *Proceedings of the 2nd International Symposium on Intelligent Robotic System*.

[Gibson, 1986] Gibson, J. (1986). *The Ecological Approach to Visual Perception*. Lawrence Erlbaum Associates.

[H.Friedrich and Dillman, 1995] H.Friedrich and Dillman, R. (1995). Robot programming based on single demonstration and use intention. In *Third European Workshop on Learning Robots - ECML'95*.

[Hugues, 2000] Hugues, L. (2000). Collective grounded representations for robots. In *5th International Symposium on Distributed Autonomous Robotic Systems (DARS 2000)*.

[Hugues and Drogoul, 2001] Hugues, L. and Drogoul, A. (2001). Robot behavior learning by vision-based demonstrations. In *4th European workshop on advanced mobile robots (EUROBOT 01)*.

[J.Demiris and Hayes, 1996] J.Demiris and Hayes, G. (1996). Imitative learning mechanisms in robots and humans. In Klingspor, V., editor, *Proceedings . of the 5th European Workshop on Learning Robot*.

[Pomerlau, 1993] Pomerlau, D. (1993). Knowledge-based training of artificial neural networks for autonomous robot driving. In J.Connell and Mahadevan, S., editors, *Robot Learning*.