Comparing Learning Attention Control in Perceptual and Decision Space

Maryam S. Mirian[1], Majid Nili Ahmadabadi[1,2], Babak N. Araabi[1,2], and Ronald R. Siegwart[3]

[1] Control and Intelligent Processing Centre of Excellence
Dept. of Electrical and Computer Eng, University of Tehran
[2] School of Cognitive Science, IPM, Tehran, Iran
{mmirian,mnili,araabi}@ut.ac.ir
[3] ASL, ETHZ, Switzerland
rsiegwart@ethz.ch

Abstract. The first question answered in this paper is whether or not learning attention control in the decision space is feasible and how to develop an online as well as interactive learning approach for such control in this space, in case of feasibility. Here, decision space is formed by the decision vector of the agents each has allowed to dynamically observe just a subset of all available sensors. Attention control in this new space means active and dynamic selection of these decision agents to contribute in making final decision. The second debate is verifying the advantages of attention control in decision space over that in perceptual space. According to the tight coupling of attention control and motor action selection, in order to answer above mentioned questions, attention control and motor action selection are formulated in a unified optimization problem and reinforcement learning is utilized to solve it. In addition to the theoretic comparison of learning attention control in perceptual and decision space in terms of computational complexity, two proposed approaches are tested on a simple traffic sign recognition task.

Keywords: Attention Control, Learning, Multi-modal perceptual space, Decision fusion, Mixture of Experts, Soft Decision.

1 Introduction

Basically, attention control can be assumed as an active intelligent filter which trims down the dimension of the huge input sensory space and prevents reaching it entirely to the further processing units. In other words, it is a must for an agent to purposefully reduce the computational burden of sensory input processing before performing any cognitive task; such as object recognition or scene interpretation.

The great significance of attention control is in fact because of these requirements: reduction of probable confusion among multiple dimensions of the perceptual space, faster response and dealing with dynamicity of perceptual space. The mentioned dynamics is in sense of reliability and accuracy of multiple sensors or processing elements. These requirements in face of limited processing power necessitate a dynamic

L. Paletta and J.K. Tsotsos (Eds.): WAPCV 2008, LNAI 5395, pp. 242 – 256, 2009.

attention control strategy rather than designing just a simple sensor selection algorithm. The tight coupling of attention control and motor action selection in a sequential decision making is another concern which makes the problem even more challenging. There are not enough works done in the field of learning attention while learning the desired behavior. It means attention control strategy is task-dependent. As a result, we couple motor actions with those that are performed solely for change of attention focus. The later ones are called perceptual actions and include those are mental only – like giving more weights to color in comparison to shape for example- and the actions that involve control of physical sensors –such as saccadic movements. We call selection of pure motor actions decision making.

It is clear that information bottleneck gives meaning to attention control however; here we raise this question that what type of information should be attentively processed? In other words, we are interested to know if attention control is restricted to active sensor selection or there is another information space where attention control can be learnt more effectively or robustly. In this paper, we chose decision space – or more accurately the probability vector of selecting actions- as a candidate information space to apply attention control in it and compare the results with those we attain in the sensory space. To perform the mentioned comparison, we model attention control as an optimization problem and choose reinforcement learning for solving this problem. The reason behind such a choice is to provide the potential for interactively solving the problem when the agent is acting in its world. By doing so, the agent learns the attention control strategy in concert with learning its task in the framework of expected reward maximization.

In this paper, we first review the related works on learning attention control. After that, two proposed approaches are described in details. Then, we will express the test-bed and the results taken. Finally a comprehensive discussion, conclusions and future works are given.

2 Related Works

Surely, we implicitly know what we mean by attention. But, a psychological definition may be a good starting point: focusing mind in a clear manner on one of many subjects or objects that may simultaneously stimulates the mind [1]. Adopting engineering perspective, it can be considered as a filtering process which trims down the input sensory space to help us focusing on some thing which is more valuable to be processed, i.e., worth-focusing. Let's look at the attention problem from action perspective and this means using active perception instead of processing the entire sensory space. This is the viewpoint we have adopted and tried to realize it through learning. In this section, the review of related works is done with more focus on learning aspects of attention. Unfortunately, there are a few researches on learning and formation of attention control; rather they are mostly related to the attention modeling. [2] presents an RL[1] based approach in which visual, cognitive and motor processes are integrated to help an agent learn how to move its eyes in order to generate an efficient behavior of a human expert while reading. Using two spatial and temporal

[1] Reinforcement Learning.

modeling parameters (fixation location of eyes as well as their fixation time) the optimal behavior is learned. In [3] a framework for attention control is presented which performs actively in high level cognitive tasks. It contains three phases: the first phase is learning attention control as in active perception. Then in the second phase it extracts those concepts learned previously and finally using mirror neurons it abstracts the learned knowledge to some higher level concepts. Continuing this work is one of our main motivations, but we are focused here on learning in the decision space rather than in perceptual space. In [4] attention control is applied in object recognition task but in a limited image database. The main idea is using information theoretic measures to find discriminative regions of the image in a general to specific manner. In [5], as a continuation of [4], a 3-step-architecture is presented which firstly extracts attention center according to information theoretic saliency measures. Then, by searching in pre-specified areas found from first step decides whether the object is available in the image and finally a shift for attention will be suggested. The final step is done using Q-Learning with the goal of finding the best perceptual action according to the search task. This research is related to our work because it also couples decision making and attention control and uses reinforcement based learning approach. In [6] two approaches for attention control are presented in a robotic platform with neck, eyes and arms. The first approach is a simple feed forward method uses back-propagation learning algorithm while the second uses reinforcement learning and a finite state machine for state space representation. Their results confirm that the second approach generates better performance in terms of finding previously observed objects even with fewer movements in head and neck and also in attention center shift. In [7] some approaches based on hidden states in reinforcement learning are proposed for active perception in human gesture recognition. This work proposes some solutions for perceptual aliasing. This problem is realized when there is a many to many correspondence among environment's state and agent's state. In such a situation, the agent's decision making has ambiguity and in order to reduce it, the agents decide to perform perceptual actions. This problem can be handled by merging similar (from utility perspective) states or splitting one state due to non-homogeneity in utility measure. The approaches for merging / splitting states presented in [7] are called Utile Distinction Memory and Perceptual Distinction Approach. Moreover, in order to handle the problem of requiring more than one shot observation, an approach called Nearest Sequence Matching is proposed which uses a chain of recent observations (state / action) to declare current state. The results show that by learning, they can find more informative set of features to attend for gesture recognition rather than just selecting them in a pre-specified manner. Unfortunately, it is mentioned that the computation load of these approaches are very high and can be problematic in real complex applications. In papers reviewed till now, the control policy was spatial. In [8] some biological evidences are presented which show that attention can also be directed to particular visual features, such as a color, orientation or a direction of motion. They showed effects of shifting attention between feature dimensions, rather than specific values of a given feature. In one condition the monkey was required to attend to the orientation of a stimulus in a distant location. In a second condition it was required to attend to the color of an un-oriented stimulus in the distant location. Finally, inspired from *Mirror Neuron* idea in [9], there is an indirect biological support for the action-based representation in the decision space as what we proposed in this paper. So, it can be assumed that for each stimulus in perceptual space, there is a corresponding action-based representation in the decision space and we have proposed two approaches for learning attention control in both spaces. Furthermore, according to the discussion presented by Rizzolatti in [10], there is a close relation between attention processes and motor planning processes. In fact, as claimed in their theory, there is a strict link between covert orienting of attention and programming explicit ocular movements.

3 Our Approach

Two approaches proposed here are based on these main concepts: Virtual Sensors and Decision Agents. Before going further into details, we define *virtual sensor* and *decision agent*. A *virtual sensor* is a processing element that gets the sensory information and extracts some high level features. A physical sensor can be regarded as a virtual sensor with the identity information processing function. According to this definition, attention control mechanism controls the physical sensors as well as the virtual ones, see Fig. 2.

A *decision agent* is a processing unit that resides inside the main agent and looks at the world through a set of virtual sensors. Its output is a probability vector. Element *i* of that vector is the suggested probability of selecting action *i* by that decision agent. Note that each action can be a pure motor action, a perceptual one or a combination of both. See Fig. 3.

As mentioned before, we have taken some primary steps to resolve the main problem of proposing a general framework for learning attention control in a dynamic and multi-modal perceptual space. Since, attention control and decision making are very closely correlated problems, we employ attention control alongside of decision making once in a high-dimensional perceptual space and once in a decision space. Therefore, in this paper, two models are proposed for a sequential, multi-step learning in each high dimensional space and the advantages and disadvantages are verified.

To summarize, in sensory space, based on the agent's current state, it learns which virtual sensors to look at in the next step in order to make the most beneficial decision, see Fig. 2. In that figure, the agent is at state *S* and has a set of action pairs each composed of a motor actions and a perceptual one; i.e. $A=\{(a_P, a_M)\}$ where *A* is the agent's action set. In other words, the agent takes a perceptual action (a_P) to select a virtual sensor and a motor action (a_M) to affect its environment.

Similarly, in decision space, the agent tries to find those decision agents –or local experts as such entities are named in multi-agent domain- to consult with to find the best decision, see Fig. 3. Again, in this scenario, the agent employs its perceptual action to select a decision agent. Note that any attentive selection –either selection of a virtual sensor or choosing a decision agent- involves processing the related sensory information.

In addition, it is worth mentioning that the selection strategy is sequential. It means that a selection is done after the selected entities are processed. It is also important to note that, similar to any motor action, each virtual sensor selection (and its processing) or expert consultation has a cost and the agent needs to minimize the total cost. The associated cost is related to the complexity of each virtual sensor or decision agent.

As Fig. 2 shows, learning attention control in the sensory space is straightforward. The agent tries to select (or in fact to attend to) those more relevant virtual sensors to the task at hand. It is done implicitly by learning a mapping between the agent's state and its optimum action in that state.

Learning attention control in the decision space looks more complex however; it is a new approach to the complicate problem of attention control. The approach benefits many interesting aspects of distributed and multi-agent systems as the agent's mind is composed of local decision makers each looking at a portion of sensory information. These local decision makers (when trained) form our local experts and the final decision of the agent can be shaped through a *mixture of experts* strategy.

Although the real world can be modeled by a MDP [11] from an absolute agent's point of view, our agent is a partial observer. So, we need to propose a POMDP approach. But to keep the problem manageable at this stage, we considered one coupled optimization problem in MDP framework. The Markov decision process provides the general framework to outline sequential attention for optimal decision making. A MDP is defined by a 4-tuple (*States, A, δ, R*) with state set *States*, action set *A*, probabilistic transition function δ and reward function *R.* In each transition, the agent receives reward from a critic according to $R : S \times A \rightarrow R$, Rt □ R. The agent must act to maximize the utility $Q(s, a)$. The decision process in sequential attention control is determined by the sequence of choices on perceptual actions - either in sensory or decision space- at specific states, see Fig. 1.

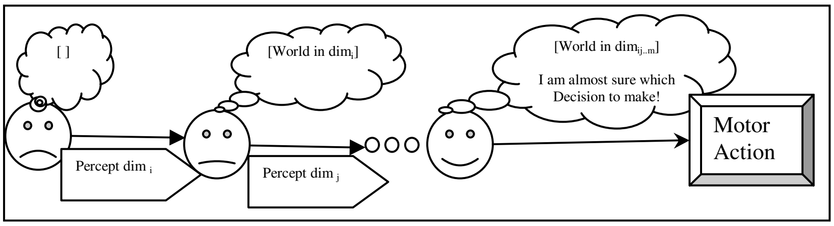


Fig. 1. A simple view of sequential perceptual state change

Fig. 1 simply shows the sequential change of agent's mental state due to performing multiple perceptual actions. At first, the agent's state is null, i.e. it knows nothing about the world's state. After a while, it decides to percept dim_i and its state changes accordingly. This continues until it can specifically decide which motor action is the most suitable to be performed. The following sections explain both learning models former in sensory space and latter in decision space.

3.1 Approach 1- Learning Attention control in Perceptual Space: Attentive Sensor Selection

In this approach, we want the agent to learn which features to attend in a state in order to gain maximum reward or in fact can perform the task as efficiently as possible from the critic's perspective, See Fig. 2.

Assume that the agent is allowed to use maximum *m* physical sensors to percept the environment and based on this information should perform one best *action* among

k available actions consisting of both perceptual and motor actions. Each physical sensor is equipped with a set of processing layers, let's say n. As mentioned before, we can assume each physical sensor plus its processing layer as a virtual sensor. The agent can either turn on all sensors at once which is very computationally expensive, time-consuming and maybe redundant or it can try to build up its percept based on a subset of its whole sensors; here, those it has found more rewarding. This can be thought as a very rough definition of agent's attention control problem. When a learning episode starts, the agent should decide whether to perform more perceptual actions to reduce ambiguity in its perception or just perform a motor action and terminate the episode. In this setting, action and state sets (A and S respectively) are defined as:

$$A = \{perceptual\_action, null\} \text{ x } \{motor\_action, null\} \tag{1}$$

$$S = \{s = (o_1, o_2, \ldots, o_m) : o_i = f_j(sensor_i)\} \quad i = 1,\ldots, m \qquad j = 1,\ldots, n \tag{2}$$

Where

$$f_j(sensor_i) \in \{v_1, v_2, \ldots, v_f, null\} \tag{3}$$

the output value of each sensor processing takes maximum f_i+1 values for sensor i including *null* when that sensor is not attended. For example, if we have a virtual sensor for temperature with three fuzzy labels, a two-valued-color and a two-valued-shape, S is:

S= {Hot, Cold, null} x {Red, Blue, null} x {Circle, Rectangle, null}. Note that a learning episode start from the null state and after a number of perceptions or after a time, when a motor action is performed, the current episode will end. Performing perceptual actions have different constant costs. This cost is a function of power consumption of the sensor and the associated processing time of its processing function. Also, when a correct motor action is performed a positive value is assigned to it. This is the common strategy of *Reward Function* of the MDP frameworks used in both approaches. Fig. 2 is a schematic view of the proposed decision making strategy coupled with attention problem (from sensor selection perspective) using RL as a learning

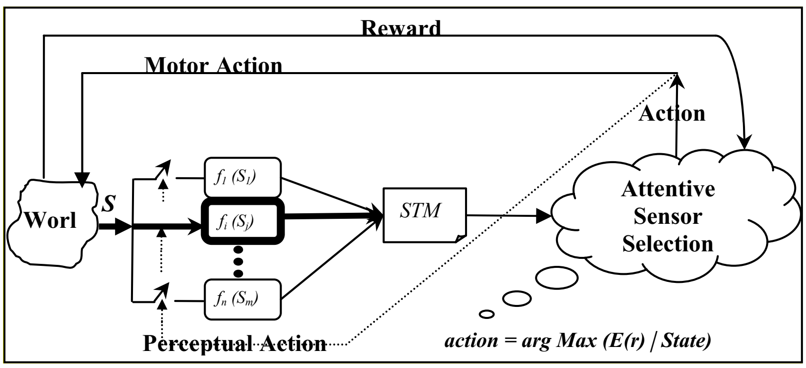


Fig. 2. Schematic view of attentive sensor selection method

method. In Fig. 2, f_1, f_2,.., f_n are processing functions like: dominant color finding, color segmentation, shape detection, straight line extraction, template matching on vision sensor and so on. STM is short term memory and here keeps the required present and past observations. The agent's state is in fact kept in STM.

3.2 Approach 2- Learning Attention Control in Decision Space: Attentive Decision Fusion

In this section, a general method for learning attention control is proposed in the decision space, see Fig. 3. Here, one simple implication from the decision space is proposed.

Again assume we have *m* sensors each observed by a tiny agent. These tiny agents are in fact our *local decision makers.* When they learned the decision making task individually in their own partial sensory space (and the learning is saturated), they start to propose their decisions (if the fuser asked them) and based on their non-greedy opinions, the agent should make the best decision which is actually performing one *action* among *k* available actions. The agent can either consider decisions made by every local expert, which is not a reasonable policy, or it can learn to build up its decision profile based on a subset of the whole decision set and on a need basis. After this introduction, let's define the decision space:

Decision sub-space is a sub-space formed by Boltzman probabilities of selecting each motor action$_j$ on the condition of $state_{S_i}$ *(as i-th sensor concerns) when the learning by agent$_i$ is finished.*

It means for each partial observation done by each tiny agent, there is one selection probability for a motor action. This definition named "decision template" is similarly introduced in [12]. Putting these templates together we will find a decision profile. It is noticeable that instead of using greedy decisions of each agent (their hard decisions) we used their soft decisions in order not to miss any probably helpful information. The mathematical definition of this subspace is expressed here:

$$D_{j|O_i} = P(action_j \mid state = O_{S_i}) = \frac{e^{Q(state_{S_i}, action_j)}}{\sum_k e^{Q(state_{S_i}, action_k)}} \tag{4}$$

in which $D_{j|O_i}$ is the agent$_i$'s decision to select action$_j$ on condition to the environment state O_i (which is the environment state from agent$_i$'s point of view) and $Q(state_{S_i}, action_j)$ is the Q-value of selecting $action_j$ in $state_{S_i}$. Therefore, by concatenating these conditional probabilities, we will find decision template of agent$_i$:

$$D_{O_i} = \left\lfloor D_{1|O_i} \mid D_{2|O_i} \mid \ldots D_{M|O_i} \right\rfloor \tag{5}$$

in which *M* is number of motor actions. The reason behind such conditional definition is that each decision is attached to a specific situation and the real environmental state is the link of the local or partial states (observed by each tiny agent). As in *Attentive*

Sensor Selection, when a learning episode starts, the agent should decide whether to perform more perceptual actions (consult more experts) to find a more descriptive state or just perform a motor action and terminate the episode. Note that a learning episode start from the null state and after a number of perceptions or after a time, when a motor action is performed, the current episode will end. Performing perceptual actions (consultation with experts) have different constant costs. Also, when a correct motor action is performed a positive value is assigned. In this setting, action and state sets (*A* and *S* respectively) are defined as:

$$A = \{perceptual\_action,\ null\} \text{ x } \{motor\_action,\ null\} \tag{6}$$

$$S = \{s = (D_{O1} \mid null), (D_{O2} \mid null), \ldots, (D_{Om} \mid null)\ \} \tag{7}$$

Fig. 3 shows the learning strategy for decision making coupled with learning attention control in the decision space.

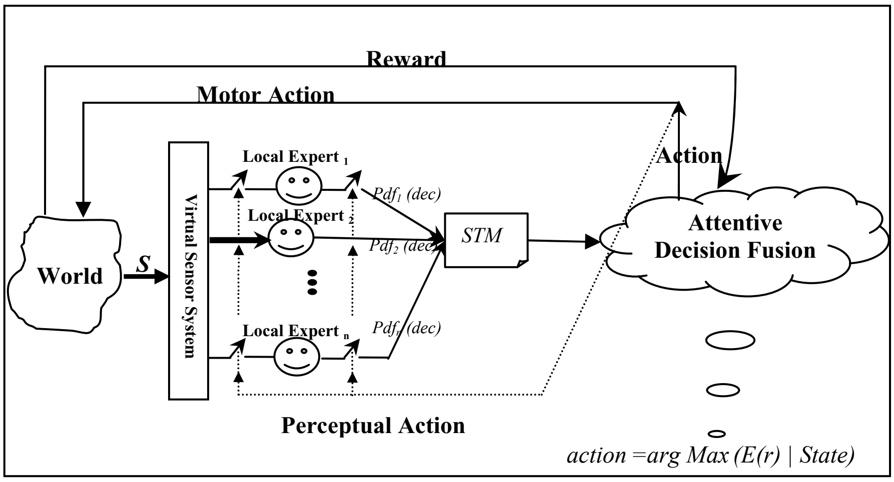


Fig. 3. The schematic view of Attentive Decision Fusion method

4 Testbed, Evaluation Measures and Simulation Results

In this section, first we introduce our testbed. Then the evaluation measures for comparing these two proposed learning strategies are defined. Finally the simulation results are given and analyzed.

4.1 Testbed

As a decision making problem, a simple cognitive task of Traffic Sign Classification is considered: "A*t the beginning of each episode, a single sign is shown to the agent. Using Attentive Sensor Selection or Attentive Decision Fusion it should decide which action to perform to minimize the total cost (of processing a feature or consulting a decision agent)".* There is a one to one correspondence between the signs and motor actions to perform. This is obviously a simple classification task which may be resolved with no attention control policy. But, there are some reasons for selecting such testbed to test our basic ideas:

- Without losing generality, any real cognitive application can be considered as a classification problem with a vast number of classes and different input data and it has the potential of extension to more complex tasks.
- This is a primary step of our ongoing research and we need to gradually test the ideas and make sure if they work. Therefore, the complexity of task should be kept small enough in order not to dominate the learning strategy.
- It is surely required in any real autonomous vehicle driving / assistant application which maybe a very good testbed for this research according to the great need to attention control in such applications.

There are three virtual sensors for the agent to percept the environment:

- Virtual Color Sensor to detect the dominant color of the sign
- Virtual Shape Sensor to detect the border shape of the sign
- Virtual Content Sensor to detect the text or symbol inside the sign.

We can consider three types of perceptual actions corresponding to attending these specific sensors (in Attentive Sensor Selection) or to consider the decision made by the agent observes these sensors (in Attentive Decision Fusion). The complexity of each processing function is implicitly considered in the cost of selecting that perceptual action. Fig. 4 shows the selected subset of traffic signs for classification.

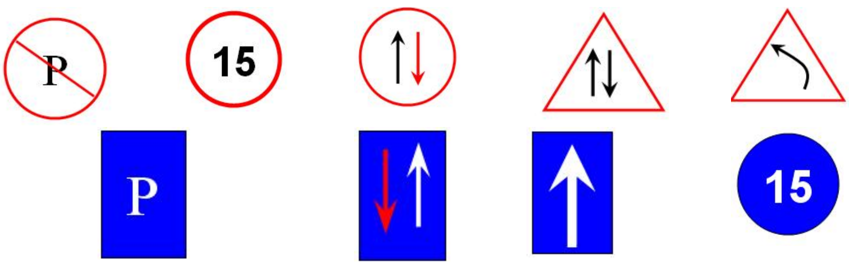


Fig. 4. Selected Traffic Signs for Recognition

According to the selected signs, we can define:

- C = Colors detected by Virtual Color Sensor = {Blue, Red}
- S = Shapes detected by Virtual Shape Sensor = {□, ❍, △}
- CN = Contents detected by Virtual Content Sensor {P, 15, ⮌, ↑, ↑↓}.

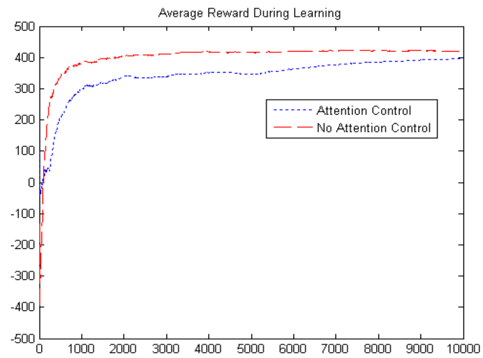
4.2 Evaluation Measures and Simulation Results

There are two sets of measures for evaluation of the proposed approaches. The first set which is tightly coupled to reward function design is accumulative reward and recognition rate. The second set contains secondary measures to evaluate our approaches: perceptual steps taken after learning and required number of episodes to complete the learning.

Approach 1- Learning Attention control in Perceptual Space: Attentive Sensor Selection

In order to show the effectiveness of the first approach, we compare it with the case where there is no attention control and the agent can utilize all its sensors at once. The results are shown in Table 1.

Table 1. Results of Simulating Approach 1 (Attentive Sensor Selection)

Measures	With Attention Control (Attentive Sensor Selection)	Without Attention Control
Recognition Rate after learning	100%	100%
perceptual steps taken	2.1	3
Average Reward gained during learning	 **Fig. 5.** The accumulative reward during learning in perceptual space	

The results justify that if we have enough time and processing power, there is no need to control the attention and the agent can learn the task even more quickly as its state space is three times smaller. However, when the attention control is necessary, *Attentive Sensor Selection* can gain perfect recognition rate while taking smaller number of perceptual steps; which means faster response and consuming less processing power.

In order to evaluate the amount of computational efficiency found by using the first approach, two other sets of results are also generated:

- Learning the task in uni-modular spaces
- Learning the task in bi-modular spaces (Color + Shape, Shape + Content and Color + Content): This is when the agent has pair of fixed sensors to percept the environment and selects its motor action accordingly.

Table 2 shows the recognition rate of the mentioned cases as well as the average reward of Attentive Sensor Selection vs. fixed bi-modular selection. The results clearly confirm that using *Attentive Sensor Selection* for attention control in the input sensory space can significantly enhance both the accumulative reward and also the recognition rate (a direct measure of success in decision-making). This is because, the agent autonomously and efficiently selects best pair of sensors to attend according to the state situated in, or maybe in some cases it pays to attend to all available sources to find the most rewarding decision.

Table 2. Results of Simulating Approach 1 (comparing with fixed selection in sub-modalities

Learning in Uni-modular Space		Learning in bi-modular Space	
Color	20%	Color + Shape	46%
Shape	30%	Shape + Content	80%
Content	50%	Color + Content	88%

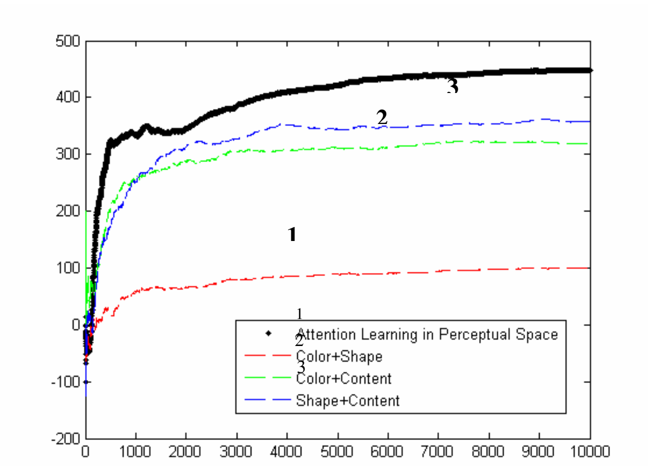


Fig. 6. The accumulative reward of Attentive Sensor Selection vs. fixed bi-modular sensor selection

Approach 2- Learning Attention Control in Decision Space: Attentive Decision Fusion

The effectiveness of the second approach (*Attentive Decision Fusion*) is shown in comparison with the first approach *(Attentive Sensor Selection)*. Learning in decision space starts with learnt pre-knowledge of each decision agent. It means, in the first step the decision agents learn the task in a parallel manner. Then, each proposes a decision vector to the main agent. The main agent uses the Max operator and selects the action with the highest probability value. All decision agents update their knowledge knowing the selected action and received reward. The agent starts learning attention control in decision space when the first step is finished. Note that, to have a fair comparison with attention control in sensor space, the learning cost of the first step is added to the cost of attention control in the decision space. The results show that the agent has learned attention control in decision space however; learning attention control in decision space is slower than learning it in the sensory space. Moreover, the number of perceptual steps taken in decision space is larger than that in the perceptual space. A detailed comparison is given in the next section.

Table 3. Results of comparing two approaches

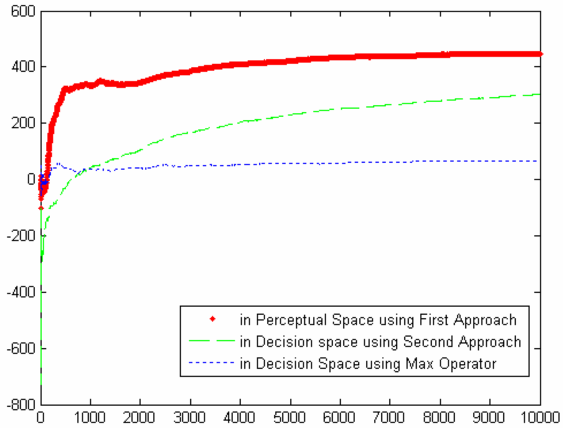
<table>
<tr><td></td><td>Attentive Sensor Selection</td><td>Attentive Decision Fusion</td></tr>
<tr><td>Recognition Rate(Test)</td><td>100%</td><td>100%</td></tr>
<tr><td>Perceptual Steps (Test)</td><td>2.1</td><td>2.8</td></tr>
<tr><td>Required episodes[2]</td><td>1000</td><td>1900</td></tr>
<tr><td>Average Reward (Learning)</td><td colspan="2">

Fig. 7. The accumulative reward during learning to compare methods</td></tr>
</table>

5 Discussions

The results show the feasibility of attention control in decision space. There are some general advantages for learning in this new space. The major ones are listed below:

- The local knowledge gained by different experts is utilized in a distributed manner by decision agents to make a unified and more confident decision. This is in fact the main justification behind any fusion algorithm.
- Decision agents share the decision space. So, their decisions can be verified to anticipate which decision agents are redundant, which decisions are more informative

[2] The number of episodes required to reach a perfect recognition rate.

and even which ones contain complementary information. It is obvious that there is no such information straightforward available in perceptual space. This information can be utilized to further reduce the learning time in decision space.

- By attention control in decision space, we can take advantage of diverse available types of learning methods for decision agents. In fact, each decision agent can use the most suitable learning method regardless of what methods the other ones employ. This benefit is gained because all agents share the decision space. Possibility of using different learning methods across decision agents enables the designer to use dissimilar types of information -such as training data, expert knowledge, etc- and sensors for training different decision agents.
- Another issue to discuss is the fact that, transferring the attention control learning from perceptual space to decision space results in learning decision fusion. Decision fusion has some major advantages (like reliability, robustness and survivability) not only because of fusion [13] but also due to its boosting characteristics. Schapire in [14] describes: “Boosting is a general method for improving the accuracy of any given learning algorithm. It refers to a general and provably effective method of producing a very accurate prediction rule by combining rough and moderately inaccurate rules of thumb.” The reason behind the claim that our proposed structure for attention control in decision space implements boosting is that “while the performance of each local expert (decision agent) is less than or equal to chance, by using learning attention control we can improve the performance considerably.” Despite the motioned general benefits, the proposed representation of the decision space seems not to be theoretically compact. This problem can be quantified through a simple order computation for the two approaches which comes in Table 4:

Table 4. Comparing Order of State-Action for both approaches

<table>
<tr><td></td><td colspan="3">In Decision Space</td><td colspan="3">In Feature Space</td></tr>
<tr><td>Parameters</td><td colspan="6">M: Number of Motor Actions
m: number of sensors
f: discretization level in sensory space
c: discretization level in decision space
n: number of decision agents
k: number of sensors observed by each decision agent</td></tr>
<tr><td>Theoretical Order of States-Action</td><td colspan="3">$M .(n f^k + c^{n(M-1)})$</td><td colspan="3">$M. f^m$</td></tr>
<tr><td rowspan="2">Example (Theoretical Number of State-Action)</td><td>$M = 9$</td><td>$m = n = 3$</td><td colspan="2">$f = 4$</td><td>$c = 10$</td><td>$k = 1$</td></tr>
<tr><td colspan="3">10^{24}</td><td colspan="3">576</td></tr>
<tr><td rowspan="2">Practical Number of State-Action</td><td colspan="3">$M. (nf^k + C)$
C = number of sparse points in decision space</td><td colspan="3">$M. f^m$</td></tr>
<tr><td colspan="3">1008 with $C \leq 100$</td><td colspan="3">576</td></tr>
</table>

Above computation theoretically shows that the number of states in decision space is very large and expresses state explosion. While, as tested in practice, the number of exiting states in decision space is very much fewer than $c^{n(M-1)}$. It means that the agent does not even go into most of the theoretically mentioned states. In other words, the space is considerably sparse. So, there is no need to reserve any space for non-existing states and be aware of their values; which results in reasonable learning speed. We are not sure if the mentioned sparseness is hold such strongly in all practical cases. Therefore, it is one of our main concerns to find a more compact representation for the decision space to preferably speed up the learning and become robust to missing information and noise. One solution is not quantizing the decision space and using continuous space RL methods [15].

6 Conclusions and Future Works

The proposed approaches are our primary steps taken to bold the main requirements of a general framework for learning attention control in a multi-modal as well as dynamic perceptual space during learning to perform a complex decision making task, such as autonomous driving which surely contains many different distracters. It is expected that if there were many distracters, the attention control algorithm would try to remove those irrelevant dimensions thus accelerate learning process considerably. The main outcome of the paper is to show that learning attention control is feasible in decision space and the results are comparable with those attained in the perceptual space. Learning attention control in decision space benefits some interesting advantages over learning attention control in perceptual space. The major ones are sharing the common space (decision space) among tiny decision agents, utilizing not necessary similar learning algorithms for decision agents and finally making a more confident decision. There are many extensions planned for the proposed approach and the most important one is finding a more compact and yet meaningful decision space to learn attention in it with preferably higher advantages such as faster learning speed, lower cost and maybe more robustness. Another extension is learning to expand the perceptual space in a gradual manner.

Acknowledgments. This research was supported by University of Tehran and has been realized in close collaboration with the BACS project supported by EC-contract number FP6-IST-02´140, Action line: Cognitive Systems.

References

1. James, W.: The principles of psychology. Holt, New York (1890)
2. Reichle, E., Laurent, A.: Using Reinforcement Learning to Understand the Emergence of “Intelligent” Eye-Movement Behavior During Reading. Psychological Review Copyright 2006 by the American Psychological Association 113(2), 390–408 (2006)
3. Shariatpanahi, H.F., Ahmadabadi, M.N.: Biologically Inspired Framework for Learning and Abstract Representation of Attention Control. In: Proceedings of International Workshop on Attention in Cognitive Systems, at IJCAI 2007, Hyderabad, India, pp. 63–80 (2007)

4. Fritz, G., Seifert, C., Paletta, L., Bischof, H.: Attentive object detection using an information theoretic saliency measure. In: Paletta, L., Tsotsos, J.K., Rome, E., Humphreys, G.W. (eds.) WAPCV 2004. LNCS, vol. 3368, pp. 29–41. Springer, Heidelberg (2005)
5. Paletta, L., Fritz, G., Seifert, C.: Cascaded Sequential Attention for Object Recognition with Informative Local Descriptors and Q-learning of Grouping Strategies. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (2005)
6. Gonc, L., Giraldi, G., Oliveira, A., Grupen, R.: Learning Policies for Attentional Control. In: IEEE International Symposium on Computational Intelligence in Robotics and Automation, pp. 294–299 (1999)
7. Darrell, T.: Reinforcement Learning of Active Recognition Behaviors. In: Portions of this paper previously appeared in Advances in Neural Information Processing Systems (NIPS 1995), vol. 8. MIT Press, Cambridge (1995); Vidyasagar M. (ed.) Intelligent Robotic Systems, pp. 73-80. Tata Press (1998)
8. Maunsell, J., Treue, S.: Feature-based attention in visual cortex. TRENDS in Neurosciences, TINS special issue: The Neural Substrates of Cognition 29(6) (2006)
9. Rizzolatti, G., Fogassi, L., Gallese, V.: Neurophysiological mechanisms underlying the understanding and imitation of action. Review Neuroscience (2001)
10. Rizzolatti, G., Riggio, L., Dascola, I., Umilta, C.: Reorienting Attention across the Horizontal and Vertical Meridians: Evidence in Favor of a pre-motor theory of attention. NeuroPhysiologia 25(1A), 31–40 (1987)
11. Putterman, M.L.: Markov Decision Processes. John Wiely and Sons, New York (1994)
12. Kuncheva, L., Bezdek, J., Duin, R.: Decision templates for multiple classifier fusion. The Journal of the Pattern Recognition Society 34(2), 299–314 (2001)
13. Zhu, Y.: Multi-sensor Decision and Estimation Fusion. The International Series on Asian Studies in Computer and Information Science 14 (2002) ISBN: 978-1-4020-7258-1
14. Schapire, R.: The Boosting Approach to Machine Learning: An Overview, AT&T Labs, Research Shannon Laboratory, Nonlinear Estimation and Classification. Springer, Heidelberg (2003)
15. Sutton, R.S., Barto, A.G.: Reinforcement Learning. MIT Press, Cambridge (1998)