

Alternatives for the Implementation of the Crowley-Arbib Saccade Model in a Robotic System

Fortunato Flores Ando, Alfredo Weitzenfeld Ridel

Abstract—The robotic implementation of *Rana computatrix* uses a geometric algorithm to regain visual input that is “lost” as the result of its orienting behavior. In this paper the Crowley-Arbib saccade model is used in several simulations in order to give this problem a solution with a biological basis. Three approaches are proposed and evaluated.

Index Terms—Biomimetic, robot, saccade.

I. INTRODUCTION

SACCADES are eye movements used by primates to rapidly bring in to the fovea, the most sensitive area of the retina, points of the visual scene that are relevant for the individual. Saccadic movements are used in robots with the same purpose.

Here we present a partial implementation of a model of the oculomotor system involved in the generation of saccades in a virtual environment. It is used to keep track of visual input lost due to reorientation.

The behavior of frogs in certain specific conditions was simulated by a robot. However it kept losing track of relevant visual input that was needed to generate motor commands. This problem was solved using a geometric [1] compensating geometric algorithm but we propose a more biologically relevant solution using saccades.

This paper is divided as follows. In Section II the frog’s behavior model is presented along with the robot that embodied it. Section III is an explanation of the Crowley-Arbib saccade model. In Section IV three methods are proposed to use the saccade model within the robot along with results of simulations carried out in a virtual environment and Section V presents conclusions and future work.

II. RANA COMPUTATRIX, DETOUR BEHAVIOR AND ITS ROBOTIC IMPLEMENTATION

Rana computatrix is a biomimetic system that models the visually guided behavior seen in the common frog, *Rana pippiens* [2]. To do so, it includes depth perception, object-recognition, selection and other mechanisms abstracted as

schemas and implemented at a lower level as neural substrates. It has been simulated using the Neural Simulation Language (NSL) and a virtual world in which various experiments can be done. [3]

The simulated model specifically deals with the behavior of the common frog when confronted to a static barrier of vertical fence posts in its way to a prey. See Figure 1. In different experiments it has been seen that the real animal either shows a detour behavior, going around the barrier, or goes straight towards it, bumping several times and reorienting until it reaches the edge, where it is finally able to reach the prey. What it does, depends on how wide the barrier is and on the times it has gone through the experiment.

When the model is simulated and instantiated in a virtual frog that can move in a virtual world it consistently behaves as its natural counterpart. Visual information is processed by several schemas, creating “attractant” and “repellant” fields, i.e. excitatory or inhibitory spatial functions, that add up to create motor response. Visual input is the first step and the generation of motor action is the last one in a series of continuous cycles that integrated give rise to behavior [4].

Rana computatrix was not only simulated in a virtual world but it was also tested in the real world by means of a robot that performed the same experiments real frogs went through. Specific details of such implementation can be found in [1]. Its behavior matched expected results, however the real world posed more problems on the system than those encountered in the simulation trials.

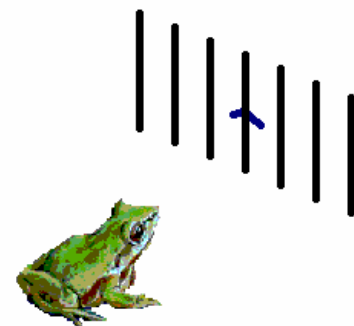


Fig. 1. Frog in prey and barrier task. The animal shows different kinds of behavior depending on how wide the barrier is and on previous experience.

One of such problems is pictured in figure 2. When the robotic frog made an orienting movement towards the edge of the barrier, its visual field would also change. As a result, the prey, which was previously visible, now could not be seen and visually guided behavior could not be calculated because there was no relevant input.

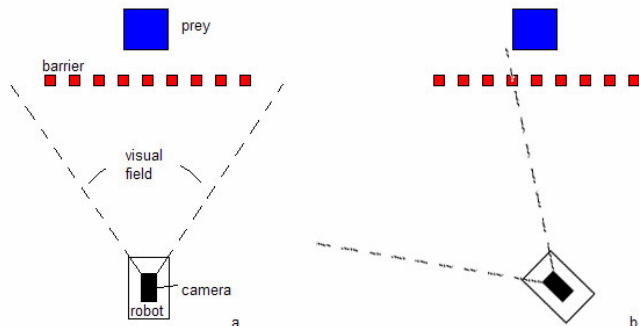


Fig. 2. Simplified top view representation of the robot experiment and the problem that arises as it moves. a) Initial setup of the experiment, the prey and the barrier are in the visual field of the robot so motor commands are calculated. b) After an orienting command has been done the visual field has changed and the prey cannot be seen, there is no relevant input for the frog.

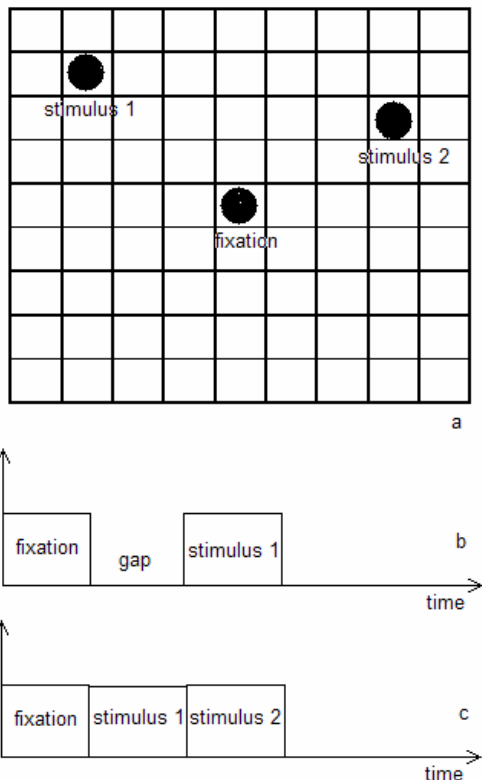


Fig. 3. a) Experimental setup for saccade paradigms. The individual's visual field is divided in an array of elements that may be active or inactive. b) In the gap saccade paradigm the fixation stimulus is turned off, then a temporal gap is introduced before a second stimulus in some other location is turned on. c) The double saccade paradigm consists of two stimuli in different locations presented in sequence after the fixation point is turned off.

To bring the prey back to the frog's visual field, a motor was included enabling the camera to reorient in order to compensate for then robot's movement. To control such subsystem, an algorithm, which made geometric calculations, was designed. In this way the problem was technically solved.

In this paper we investigate the use of alternative solutions, in particular, the use of saccades. This method is biologically inspired but it should be noted that we are not suggesting this mechanism actually occurs in the frog, in particular since it is based on a mammalian saccade model while toads already compensate with a large visual field.

III. CROWLEY-ARBIB SACCAD E MODEL

The primate's retina is a highly specialized system composed of two kinds of light receptors (cones and rods) and other kinds of neurons (amacrine cells, ganglion cells, horizontal cells and bipolar cells), performing visual processing. What we could call "high resolution" vision is achieved thanks to a small region of the retina called fovea, having a high density of cone receptors. However the portion of the visual field that this region subtends is very small and the eye has to be moved constantly around to "scan" the image of the outer world in search of relevant visual features. These "scans" or Saccades are very fast eye movements by which different points of the visual field are foveated.

The Crowley-Arbib [4],[5] saccade model, an extension of the Dominey-Arbib [6] model, tries to explain the generation of saccadic eye movements with a system that integrates the workings of cortical and subcortical structures in primates.

It is successful in explaining the gap saccade paradigm and the double saccade paradigm taking into.

Figure 3 shows the basic setup for such experiments. The subject's visual field is divided in a 2-dimensional array of locations which are either active or inactive, i.e. lighted or obscured. In the gap saccade task a fixation stimulus to the fovea is presented for some time. After it is turned off another location in the visual field is activated. The eye must saccade to this second point, i.e. must move its eyes so that the fovea directs to such location.

The double saccade task is an experiment in which after a fixation point is presented two targets appear in sequence. The individual must saccade to each one in the same order. The duration of both stimuli is smaller than the time needed to start the first saccade. In this paradigm several mechanisms are involved such as memory and remapping. Since the stimuli are not available during or after any of the saccades, there must be an internal spatial representation of their locations in the neural layers of the oculomotor system. When the first saccade is done this neural image of the second location suffers a change to compensate for the transformation of coordinates that happens when the eyes move, this is called a remapping. For a detailed explanation of the model refer to [5] and [6].

IV. THE SACCAD E MODEL IN VISUAL INPUT COMPENSATION

Saccades have been implemented in several robotic systems. In [8] saccades are used in an autonomous robot to redirect vision to potentially interesting targets. In [9] and [10] saccades are used along with other strategies in order to track selected moving targets. Here the saccadic system is proposed as a way of recovering visual information that has been lost due to the movement of the robot. Another distinction should

be made: the mentioned robots use saccades with biologically inspired algorithms, in this paper the saccadic system is a neural model where different neural layers are simulated as dynamical systems, with differential equations.

Three ways of actually embodying the Crowley-Arbib saccade model have been devised. They have been tested in the same virtual world that was used to experiment with the virtual Rana computatrix. No integration of models has yet been done; instead each implementation is being revised in order to determine if in a static context saccades can be done to locate the prey.

1) To divide each video frame sent by the camera in a 9x9 array so that the central element can be considered as the fovea. In this computational retina no difference in resolution among fovea and periphery is considered, but the spatial location of the former is used to determine whether the prey is in the center of the visual field or not. This approach would take advantage of the higher velocity of saccades compared with detour behavior computation in order to keep the “prey” focused at every time. See Figure 4. For this to work, calculation and realization of saccades would have to be done in parallel to that of the detour behavior so that the discrete steps commanded by the motor schemas of Rana computatrix are continuously compensated. However this is not something easily done and the robotic frog does not include that functionality built in.

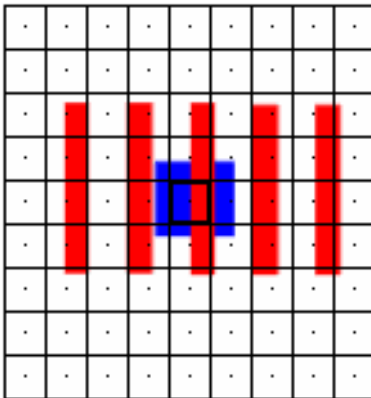


Fig. 4. The visual field is divided in a 9x9 array, with the fovea represented by the central element. Whenever the prey (dark square in background) is not foveated the system commands the camera to move in order to recenter it.

In simulations the information of the visual field is a 90x90 bitmap which is passed through a filter to determine which elements correspond to relevant information. If the element carries information of a prey it is set to an arbitrary activation level greater than 0. If the information corresponds to the barrier or something else then this element is set to 0. This filtered array is then reduced to a 9x9 array for further processing in the saccadic system. Each of the elements of this smaller array is calculated as the average of the values carried by 10x10 subarrays of the 90x90 array.

Figure 5a shows the initial setup, as the frog sees it, used to test the three approaches. It can be seen that the prey is at the center of the visual field. Figure 5b shows the 9x9 array used as visual input to the saccade model, it is seen that three

elements of the center column are active. Figure 5c shows the top view of the experiment, at the beginning. Figure 5d shows what is seen after a command issued by Rana computatrix is performed. A 20 degree rotation to the right has been done. The visual input to the saccade model that corresponds to this frame is seen in figure 5e. It can be seen, in figure 5f, how the visual field is modified. At this moment a saccade is started in order to bring the prey to the center of the visual field and while it is in progress no visual input is taken. In real (biological) systems such input exists but it is minimized thanks to the velocity of the movement [11]. Here we have simplified this process not letting any image to come in because with the robot such a rapid movement won't be able to be produced. Figures 5d through 5i show the result of the saccade, the prey is in the center of the visual field and the element corresponding to the fovea is active, preventing any saccades to occur.

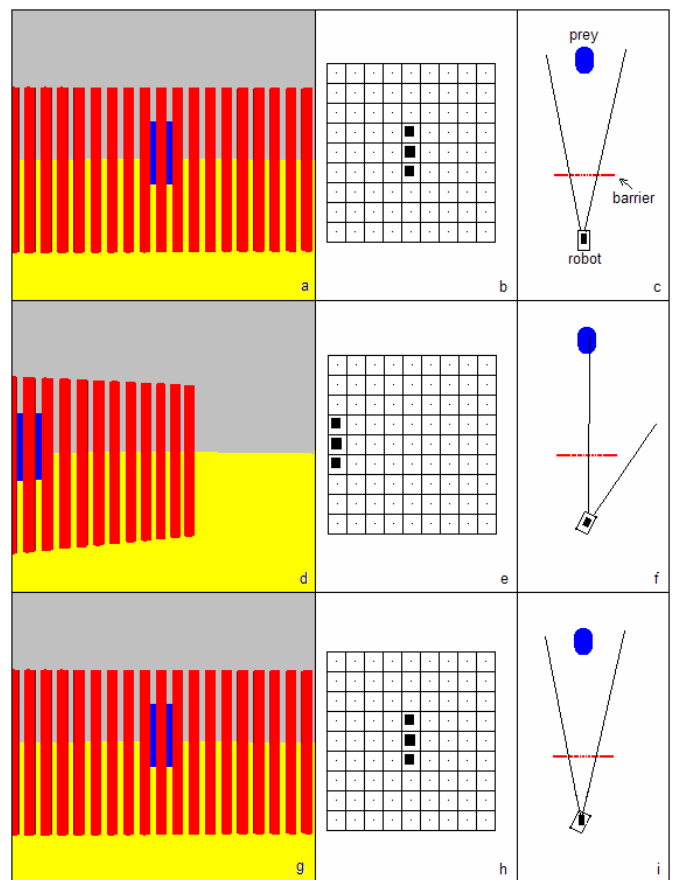


Fig. 5. a) Initial setup to test the proposals. b) Prey is foveated so no saccade is triggered. c) Top view of the initial setup. d) Image after a rotation has been done. e) Visual input carrying relevant information on the location of the prey in the visual scene f) Top view showing how the visual field has changed. g) The saccade has ended, the prey is centered in the visual field. h) Visual input is regained after the saccade; the fovea is active. i) The camera has rotated and the visual field is the same as it was at the beginning.

2) To consider the camera as the fovea and the rest of the retina to be coarsely and virtually calculated by means of a feedback signal coming from the motor schemas of Rana computatrix. See figure 6. In this way when the prey has disappeared from the actual camera image, the “virtual”

periphery of the retina would still carry information on where the prey is, or it is supposed to be, and a saccade could be made in order to bring it again in the visual field of the camera. The state of visual input out of the foveal region is calculated using a signal that codes the direction (right or left), of the rotation generated by detour behavior. Then the particular location of the LIP region is activated with this computed retinal image so that saccades are evocated.

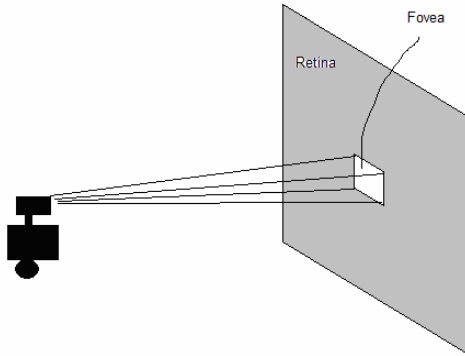


Fig. 6. What the camera sees is considered the fovea, the rest of the retina is calculated according to feedback coming from motor centers.

For this method to work the prey must be in the visual field at the beginning of the experiment. In this way by “remembering” the past location and knowing where the frog rotated to, the retina can be built. The “virtual” visual input is created in a very rough way. If the frog rotated left, the rightmost center location of the visual input is set to some arbitrary value. In this way a saccade will be made to that location. If the rotation was to the right then the leftmost center location is set and a saccade is elicited in that direction.

The activation of the fovea is determined by the presence of a prey stimulus in the visual field of the camera. The 90x90 bitmap is filtered to look up for relevant information for the model (if the prey is visible or not) as in the first case, and the average is calculated. The difference is that in this case this determines just the value of the central element of the input array.

Figure 7 shows different moments of an experiment done using this proposal. Figure 7a shows what the frog sees after a 25 degrees rotation to the right has been commanded. The initial setup is the same as that for experiment one. This rotation is a simulation of a command that the detour behavior model would issue to the robot. The prey is not visible any more but the “virtual” visual field shows activity in the leftmost center element, as can be seen in figure 7b. This elicits a saccade to that location. At the end of the saccade the prey is again in the visual field and it is almost centered.

3) To consider the camera as one spot in the retina and use several image frames, each one taken from a different angle, to create the whole image of the retina. See figure 8. This approach means that before each saccade, the system would need to make a “scan” of the visual field by rotating the camera and integrate that information in order to create visual input to the saccade model.

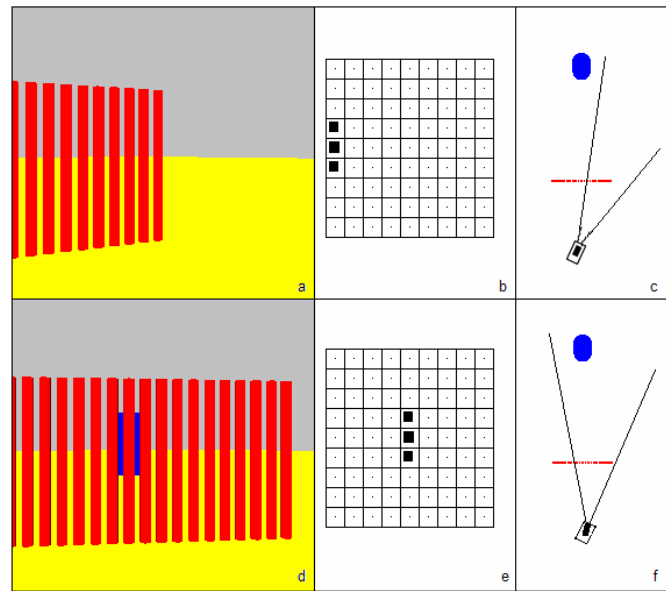


Fig. 7. a) Image seen by the frog after the rotation commanded by Rana computatrix has been done. The prey is out of the foveal region. b) Visual input carrying relevant information on the location of the prey in the visual scene. c) Top view showing rotation of the robot. d) The saccade has ended, the prey is centered in the visual field. e) Visual input is regained after the saccade; the fovea is active. f) Top view showing the reorientation of the camera and hence of the visual field.

Nine different positions of the camera have been used in simulations, each varying 25 degrees. The images obtained with each position are also filtered in search of prey stimuli. Each filtered image is used to calculate the numerical values of each element of each of the columns of the visual input array. This is done by dividing each of the 9 bitmaps in 9 subarrays of 10x90 elements each and then averaging the elements contained in each one to obtain a single value that corresponds to a location in the visual input.

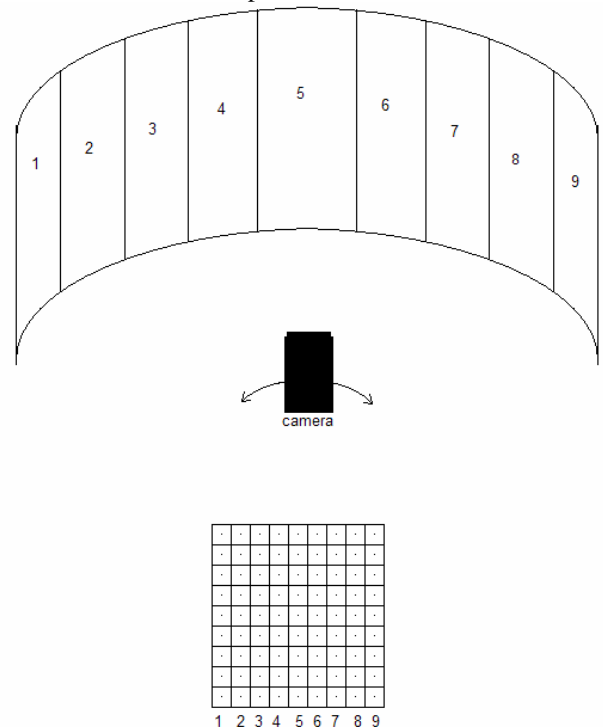


Fig. 8. The camera rotates to get 9 different images. Each image is then used to create the visual input for the saccade model.

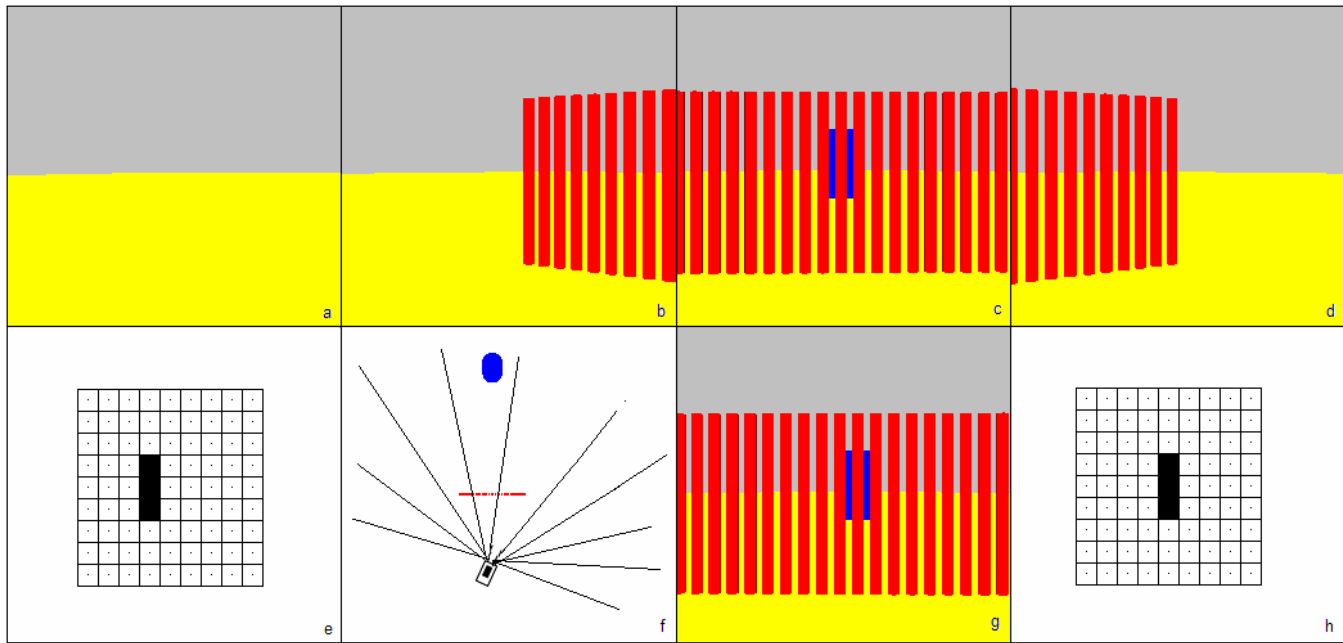


Fig. 9. a) - c) Images obtained after 75°, 50° and 25° rotations to the left. d) Image obtained with no rotation. e) Calculated input array after all rotations. f) Top view that shows how the rotations are done in order to do the scan. g) The result of the saccade. h) The fovea is active, indicating that the prey is visible.

Figures 9a through 9d show 4 of the 9 images taken to calculate the whole visual input. Figure 9e shows such input array after it has been calculated. The activation shown in three of the elements will elicit a saccade to the left. It can be seen in figures 9g and 9h that at the end of the saccade the prey is visible and the foveal location of the input array is active.

V. CONCLUSIONS

Three schemes were proposed to use the Crowley-Arbib Saccade model as an error-correction mechanism of the visual input in the robotic implementation of the visually guided detour behavior of *Rana computatrix*.

Simulations were done in order to test the performance of each proposal. The first alternative worked well but it has two drawbacks. The first is related to the video processing that has to be done to create the visual input. The second and most important is that it does not solve the problem completely because it achieves saccades when the prey is not centered but is still in the visual field. It was proposed that by exploiting the relative higher processing speed that takes place in the oculomotor system in contrast with that of *Rana computatrix*, this method could be used. However it may be difficult to coordinate both systems to achieve smooth processing. As both models have not been integrated, this alternative cannot be totally disregarded.

So far the second proposal has proven to be the most useful. The camera is seen as the fovea and the rest of the retinal

image is virtually generated with a simple algorithm. This approach may be useful in the development of new models regarding the orienting behavior shown by real frogs and in models trying to explain remapping of saccadic targets when not just the eyes but the whole body is moving. It may also be helpful in determining how such remappings are done when visual information is missing or is not the only one that has to be taken account of. The simulation results showed that the “virtual” visual input that was calculated had some error. It estimated the direction correctly but the size incorrectly, i.e. the prey was closer to the fovea than it was computed. To overcome this problem a similar approach to that used in the saccadic model remapping mechanism may be searched for: by previous leaning the robot could learn how to correctly compensate for its own movements. The remapping mechanism as it is, is insufficient for this purpose but an effort to find such a mechanism would be interesting from the neural modeling point of view. More research has to be done in order to find such a process.

The third alternative proved to work well but it requires more video processing than the other methods. The constant movement of the camera may be something difficult to coordinate and to artificial to work.

The results of the simulations show that the use of the saccadic model may be useful. However a lot of improvements and simplifications must be done so that the computational power needed to perform the task is not prohibitive.

Camera movements and video acquisition in the real world may not be as smooth and noiseless as in the virtual environment. The output signal of the saccade model depends on actual dynamical properties of the eye so in the virtual model the way the calculation is done is not of great importance, however in the real robot this calculations might prove crucial in the workings of the physical system. Further

evaluation is required in order to use the model in the real robot.

Along with the difficulties of adapting a saccade model that was created to explain very specific experimental results obtained through the course of very specific experimental protocols, new challenges arise as it is necessary to integrate it with Rana Computatrix, not just glue it together as this would have no biological relevance [1][12]. The effort of integrating both models may prove to be the most interesting and fructiferous task because the models are based in very different neural systems. These differences may suggest what and how strategies to solve specific problems evolved as the neural systems developed higher capacities.

Embodying the Crowley-Arbib saccade model in a robot is still a job in process.

REFERENCES

- [1] A. Weitzenfeld, "A Multilevel Approach to Biomimetic Robots: From Schemas to Neural Networks", Technical Report, ITAM, 2004.
- [2] M. Arbib, *The metaphorical brain 2*. New York: John Wiley and Sons, 1989.
- [3] *The Neural Simulation Language: A system for Brain Modeling*, Aut. A. Weitzenfeld, M. Arbib and A. Alexander, MIT Press, 2002.
- [4] F.J. Corbacho and A. Weitzenfeld, "Learning to Detour", in *The Neural Simulation Language: A system for Brain Modeling*, Aut. A. Weitzenfeld, M. Arbib and A. Alexander, MIT Press, 2002, pp. 319–342.
- [5] M. G. Crowley, *Modelling Saccadic Motor Control: Normal function, Sensory Remapping and Basal Ganglia Disjunction*. Ph.D. dissertation, Fac. Grad. School, University of Southern California, California, 1997.
- [6] M. Crowley, E. Oztop and S. Marmol, "Crowley-Arbib Saccade model", in *The Neural Simulation Language: A system for Brain Modeling*, Aut. A. Weitzenfeld, M. Arbib and A. Alexander, MIT Press, 2002, pp. 283–304.
- [7] P.F. Dominey and M.Arbib, "A cortico-subcortical model for the generation of spatially accurate sequential saccades", *Cerebral Cortex* vol. 2, pp 153-175, 1992.
- [8] M.R. Blackburn and H.G.Nguyen, "Autonomous Visual control of a Mobile Robot", Presented at the 1994 Arpa image understanding workshop, 1994, Monterey, CA.
- [9] Ch. Balkenius and L. Kopp, "Visual Tracking and Target Selection for Mobile Robots" in *Proceedings of EUROBOT '96*.
- [10] H. Kim, B. Lau and J. Triesch, "Adaptive Object Tracking with an Antropomorphic Robot Head" in *Proceedings of SAB'04*, CA, USA, July 2004.
- [11] D.L. Sparks, "The Brainstem Control of Saccadic Eye Movements", *Nature*, vol. 3, pp. 952-964, 2002.
- [12] B. Webb, "Can Robots make good models of biological behaviour", *Behavioral and Brain Sciences* 24(6), 2001.