Estimating Spatial Position from Proprioceptive and Visual information: A Cognitive Robotics Model

Alessandro G. Di Nuovo[1] , Davide Marocco[2] , Santo Di Nuovo[1] , Angelo Cangelosi[2]
[1]University of Catania, Italy
[2]University of Plymouth, UK

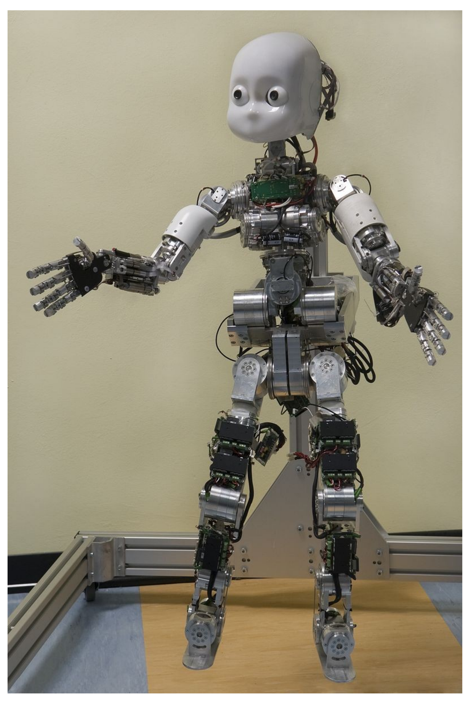
Abstract

In this work we presented the results of our preliminary study about mental imagery in cognitive robots.
In particular we focused on spatial position estimation from proprioceptive and visual information, demonstrating that cognitive robotics can be an interesting and complementary framework to study mental imagery by means of real tasks.

Introduction

Understanding the process behind the human ability of creating mental images of events and experiences is a crucial issue for psychologists (Kosslyn and Thompson, 2006). From a psychological perspective, mental imagery may be considered a multimodal biological simulation that activates the same, or very similar, sensorial and motor modalities that are activated when we interact with the environment in real time. For instance, research reports that the time needed for an imaginary walk is proportional to the time spent for the real action (Decety et al., 1989). Similarly, neuropsychological studies show that neural mechanisms underlying real-time visual perception and mental visualization are the same when a task is mentally re- called. Nevertheless, the neural mechanisms involved in the active elaboration of mental images might be different from those involved in passive elaborations. The enhancement of this active and creative imagery is the aim of most psychological and educational processes, although, more empirical effort is needed in order to understand the mechanisms and the role of active mental imagery in human cognition. In this work we present some preliminary results of on ongoing investigation about mental imagery using cognitive robotics. Here we focus on the capability to estimate, from proprioceptive and visual information, the position into a soccer field when the robot acquires the goal. The final objective of our work is to replicate with a cognitive robotics model the study presented in (Smith et al., 2001), where mental imagery is used during the training phase of athletes that are allowed to imaginary practice to score a goal. Results show that the imaginary practice enhances the train phase, obtaining better performance.

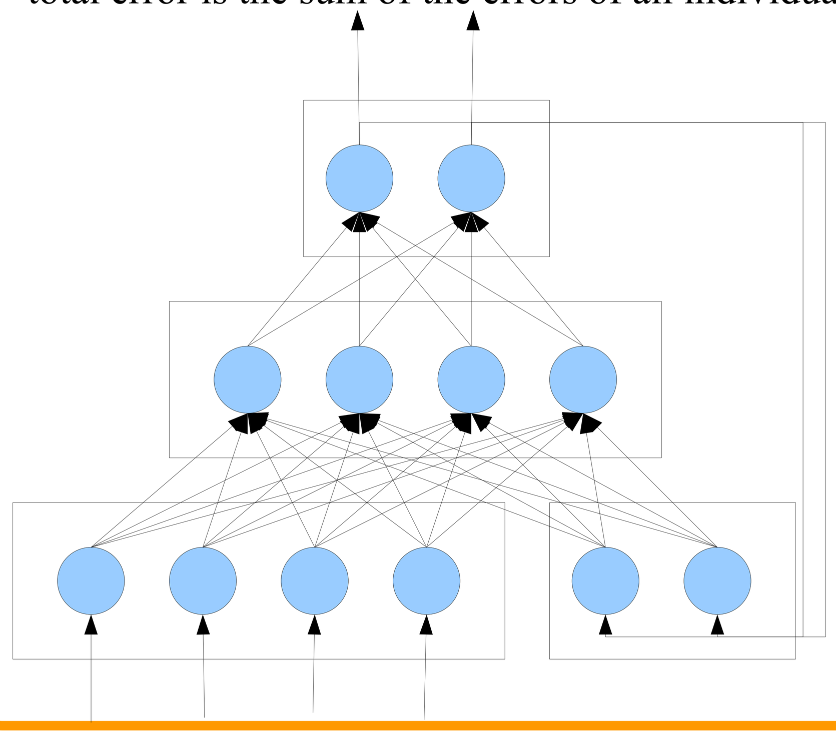
iCub



The iCub is an open source cognitive humanoid robot platform. It has 53 motors that move the head, arms & hands, waist, and legs. It can see and hear, it has the sense of proprioception (body configuration) and movement (using accelerometers and gyroscopes).

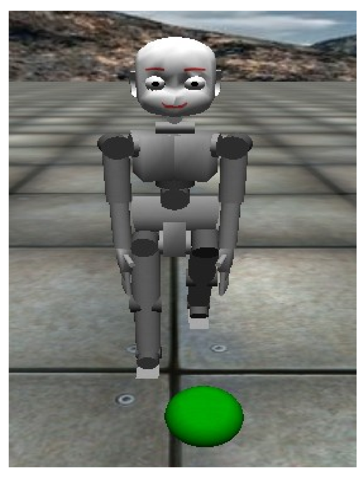
Recurrent neural networks

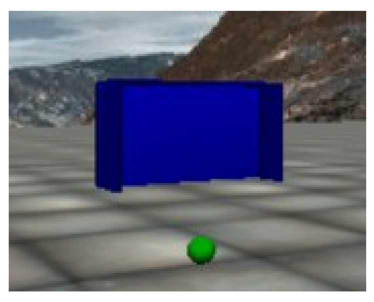
Mimicking the human brain, which is a network of neurons with feedback connections, a recurrent neural network (RNN) is a class of neural network where connections between units form a directed cycle. This creates an internal state of the network which allows it to exhibit dynamic temporal behavior. Unlike feedforward neural networks, RNN can use their internal memory to process arbitrary sequences of inputs. For supervised learning in discrete time settings, training sequences of real-valued input vectors become sequences of activations of the input nodes, one input vector at a time. At any given time step, each non-input unit computes its current activation as a nonlinear function of the weighted sum of the activations of all units from which it receives connections. There may be teacher-given target activations for some of the output units at certain time steps. For each sequence, its error is the sum of the deviations of all target signals from the corresponding activations computed by the network. For a training set of numerous sequences, the total error is the sum of the errors of all individual sequences.



Materials and Methods

The robotic model used for the experiments is a simulation of the iCub humanoid robot (Tikanoff, 2008) controlled by a recurrent artificial neural network. For the present work have been used the vision system and only a small subset of the available degrees of freedom.
The environment is a square portion of a soccer field, whose length and width are both 15 meters. At one end is placed a goal 1.94m wide. The robot can be positioned anywhere in this square.
The recurrent neural network has 37 input, 16 hidden and 6 output units. The input variables are the following: polar coordinates of the robot with respect to the goal (represented by the radius and the polar angle). The current angle of, respectively, the neck (left-right movement), the torso (left-right movement) and the body rotation joint (a joint attached to the body that allows the robot to rotate on its vertical axis). Visual information are provided through a vector of 32 bits, which encodes a simplified visual image obtained by the left eye and allows the robot to locate the goal in its visual field. The 6 outputs are, respectively, the polar coordinates of the robot respect to the goal, the desired angle for the neck, the torso and the body rotation joint, and an additional binary output that makes the robot kick the ball when its value is 1. In the learning phase this output was set to 1 only when the motion ends and the robot must kick the ball to the goal. All input and output variables are normalized in *[0,1]*.
Using the iCub simulator the robot was positioned into the environment in 8 different positions and then rotated to acquire the goal. During this movement, proprioceptive and visual information were sampled in order to build 20 input-output sequences corresponding to the 8 positions. After this movement, the robot was allowed to kick the ball. The kicking movement was pre-programmed. With the 8 series recorded, the neural network was trained to predict the next sensory state (excluding the visual input) by means of a backpropagation through time algorithm (Rumelhart and McClelland, 1986) for 50.000 epochs, after which the MSE error in estimating the 6 output variables was 0.0056.





Experiments

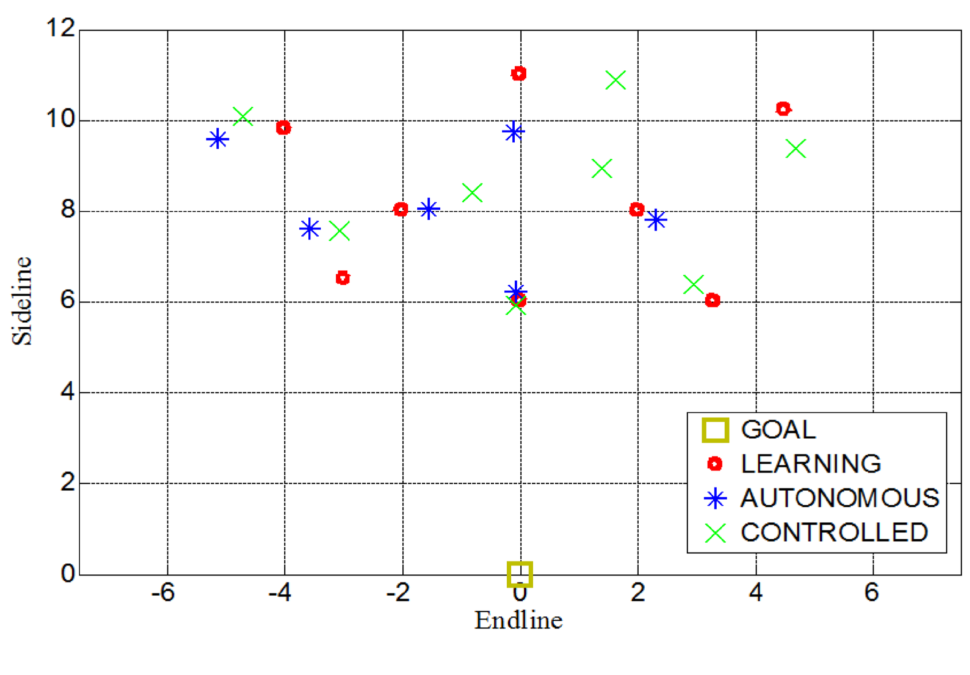


Figure 1 : Results on Learning positions

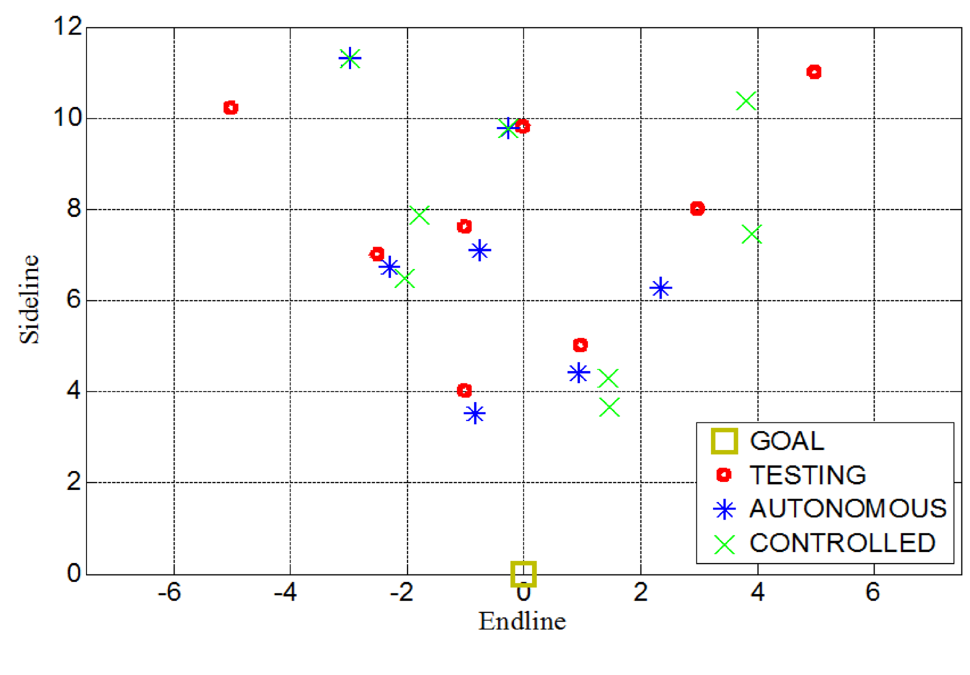


Figure 2 : Results on Testing positions

Two case studies were carried out to assess the generalization performance of the neural network and to evaluate the use of visual and proprioceptive information for the estimation of the robot position with respect to the goal. In a first study the robot was controlled by the same algorithm used for collect the train series (Controlled condition), meanwhile the neural network was used to estimate the position coordinates only. In the second study the robot was fully moved by means of the neural network (Autonomous condition), which controlled the neck angle, the torso angle and the body rotation angle, as well as the output commanding the kick.
Figure 1 and 2 show the environment with the 8 real and imagined positions respectively for train set and test set. As the figures show, overall the robot is able to estimate its position in the environment to a good extent in both conditions. Surprisingly, the Autonomous condition performs better than the Controlled. In the Controlled condition the average error in estimating its position is the 14.125%, whilst in the Autonomous condition the error is 9.245%. On the other hand, the Controlled condition shows better performance in terms of scores, since the positioning movement is governed by the hand-written algorithm. In this condition the robot scores 100% of times. In Autonomous condition the robot misses the 50% of the scores, but it is worth to mention that errors were mostly made in the test set (5 out of 8) and when the goal was very far and even a little error in the position leads the ball far from the goal.

Future Work

The next step of this research will be to incorporate self-generated "mental" position in the training set and compare the performance with and without the help of the imagery training, aiming to replicate and to corroborate experimentally some of the data already known about the role of mental imagery in humans and sport practice, in particular. Moreover, we believe this work could be an additional step toward the incorporation of embodied cognition principles to the current research in Cognitive Science. The tight relation between mental simulations and real actions is a perfect example of embodied cognition, where the mind represents, or simulates, the body in action (Iachini, 2002). For this reason, the integration of traditional psychological studies with cognitive robotics models capable of embodied mental simulations can led to mutually fruitful insights and results, both for psychology and robotics.

References

Decety, J., M., J., and C., P. (1989). The timing of mentally represented actions. Behavioral & Brain Research, 34:35–42.

Iachini, T. (2002). Scienze della mente, chapter Spazio, movimento e immagini mentali, pages 176–184. Mulino, Bologna.

Kosslyn, S. M. and Thompson, W. L. (2006). The Case for Mental Imagery. Number 39 in Oxford Psychological Series. Oxford University Press, New York.

Rumelhart, D. and McClelland, J. (1986). Parallel distributed processing: Explorations in the microstructure of cognition. M.I.T. Press, Cambridge.

Rumelhart, D. and McClelland, J. (1986). Parallel distributed processing: Explorations in the microstructure of cognition. M.I.T. Press, Cambridge.
Smith, D., Holmes, P., Whitemore, L., Collins, D., and Devonport, T. (2001). The effect of theoretically-based imagery scripts on field hockey performance. Journal of Sport Behavior, 24:408–419.

Tikhanoff, V., Cangelosi, A., Fitzpatrick, P., Metta, G., Natale, L., and Nori, F. (2008). An open source simulator for cognitive robotics research: the prototype of the icub humanoid robot simulator. In PerMIS '08: Proceedings of the 8th Workshop on Performance Metrics for Intelligent Systems, pages 57–61, New York, NY, USA. ACM.