Learning Hand-eye Coordination for a Humanoid Robot using SOMs

Ivana Kajić
Bernstein Center for Computational
Neuroscience
Berlin, Germany

Guido Schillaci, Saša Bodiroža,
Verena V. Hafner
Cognitive Robotics
Department of Computer Science
Humboldt-Universität zu Berlin
Berlin, Germany

ABSTRACT

Hand-eye coordination is an important motor skill acquired in infancy which precedes pointing behavior. Pointing facilitates social interactions by directing attention of engaged participants. It is thus essential for the natural flow of human-robot interaction. Here, we attempt to explain how pointing emerges from sensorimotor learning of hand-eye coordination in a humanoid robot. During a body babbling phase with a random walk strategy, a robot learned mappings of joints for different arm postures. Arm joint configurations were used to train biologically inspired models consisting of SOMs. We show that such a model implemented on a robotic platform accounts for pointing behavior while humans present objects out of reach of the robot's hand.

Keywords

cognitive robotics, HRI, sensorimotor coupling

1. INTRODUCTION

Detecting and manipulating the focus of attention of an interacting person is fundamental in human-robot interaction. In humans, the development of such social skills is preceded by the acquisition of early motor competences, such as hand-eye coordination [4].

In [2], learning hand-eye coordination and reaching capabilities through self-exploration led to the development of pointing behaviors in a humanoid robot. Learning consisted in the robot exploring its arm movements while collecting sensorimotor data (hand and arm joint positions), thus building up an internal model of its own body. A simple predictive algorithm provided the robot with a mechanism for producing motor commands to reach desired hand positions. Pointing behaviors emerged when target points were presented outside the reach of the robot, thus strengthening the hypothesis that pointing may arise from grasping.

Here, a similar experiment is presented, where a biologically inspired model consisting of self-organizing maps (SOMs)
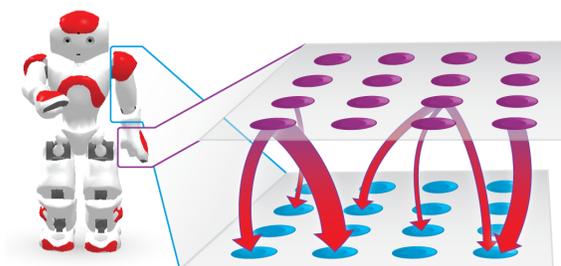
Figure 1: Scheme of the model architecture for learning hand-eye coordination with SOMs and connecting weights

has been used for modeling the hand-arm joints mapping. SOMs reproduce the phenomenon of self-organization of the developing brain in response to sensory input [1]. As in [2], the robot showed pointing behaviors when required to approach points outside its reach. The model architecture is inspired by the Epigenetic Robotics Architecture presented in [5], where a structured association of multiple SOMs has been adopted for mapping different sensorimotor modalities in a humanoid robot.

2. BIOLOGICALLY INSPIRED MODEL

We implemented a biologically inspired model consisting of two connected neural networks simulating two distinct sensorimotor brain areas. SOMs, or Kohonen networks [3], are a class of topological neural networks used to learn low-dimensional representations of data by neurons approximating multiple data points. We used 2D SOMs to represent left arm postures of a humanoid robot. Neurons in one map were used to represent the hand trajectory and neurons in the other the joint elbow and shoulder trajectory.[1] The schematic architecture of the model is shown in Figure 1.

In the random babbling experiment 74,143 input vectors were collected and used for training of SOMs.[2] Training

---

[1] Horizontal, vertical and depth dimensions were used to characterize the position of a hand. Elbow jaw, elbow roll, shoulder pitch and shoulder roll were used to characterize the rest of the arm trajectory.

[2] Random body babbling lasted approximately 40 minutes. Trajectories of the robot's left arm and the marker position on the hand were captured by a camera positioned in its head at a varying speed between 9 and 30 frames per second.
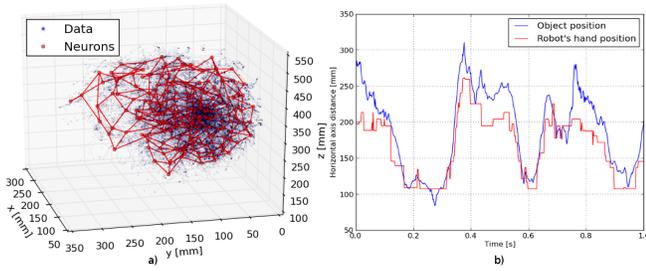
Figure 2: SOM with 225 neurons covering the left hand trajectory samples (a) and movements of the object and the robot's left hand along the horizontal axis (b)

Figure 3: Human-robot interaction

comprised of 20,000 iterations and in each iteration a random input vector was chosen – 3D hand coordinates for the first SOM, and 4D joint coordinates for the second SOM. Neurons with the smallest Euclidean distances to the input vector in each SOM were pulled together using a Gaussian neighborhood function centered around the winning neuron ($\sigma = 0.7$) at a learning rate $\eta = 0.9$. After the first half of iterations, parameters η and σ were annealed exponentially. To contrast the precision of arm movements, we trained two instances of a model, one consisting of two SOMs with 25 neurons each (5x5 model) and the other instance with SOMs with 225 neurons (15x15 model) mimicking the increase in the number of neurons for the more advanced stage in the skill development. The 15x15 network covering the left hand trajectory obtained in the babbling experiment is shown in Figure 2(a). The weights linking neurons of the two SOMs were increased according to the Hebbian learning rule: if a postsynaptic neuron is activated after the presynaptic neuron, the connection between these two neurons is strengthened. Both networks were presented with vectors from the babbling knowledge base, and the connection weight between winning neurons in each network was increased proportionally to their activations and the scaling factor $\eta = 0.01$. We adapted minimalistic SOM implementation [6] for our purposes.

3. RESULTS

We ran the experiment with a 15x15 model configuration by moving an object tagged with a marker in front of the robot for around 2.5 minutes. Movements were random and covered the space within and beyond the reach of the robot's hand. For positions out of its reach, the robot pointed to the object as presented in the experimental setting in Figure 3. Every 100 ms we collected sensory data which were the position of the object held by a human and captured by the robot's camera, and arm postures determined by both SOMs for 5x5 and 15x15 model configurations. The position along the horizontal axis of a robot is constrasted with the position of a marker in Figure 2(b).

As a simple quantification measure of pointing we introduced pointing precision error, which is defined as the sum over all dimensions of Euclidean distances between the position of a marker and the hand position.

A paired-samples t-test was conducted to compare precision errors for the 5x5 network and the 15x15 network. There was a significant improvement in the performances
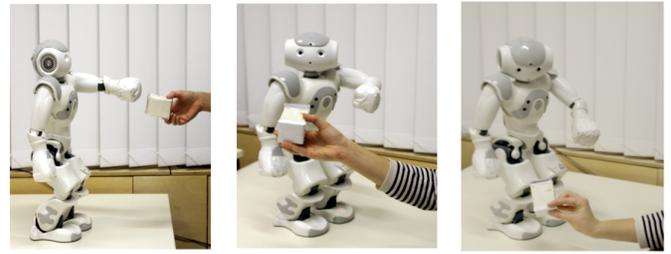
from the 5x5 network ($Mean\ error = 99.93\ mm$) , $Std.Dev. = 32.10\ mm$) to the 15x15 network ($Mean\ error = 80.32\ mm$, $Std.Dev. = 33.41\ mm$) conditions; t(1400)=76.47, $p < 0.05$.

4. FUTURE WORK

The preliminary results have shown that SOMs are well suited for sensorimotor coupling on a humanoid robotic platform. In particular, they explain how hand-eye coordination during random body babbling contributes to the development of pointing skills. Future work includes further analysis of the influence of the number of neurons and training parameters of the network on pointing precision. In addition, one might be able to attribute the number of neurons in the network to different degrees of pointing skill development in infants and children. Using extended SOM models such as growing neural gas, one could observe progressive development of the skill. Understanding cognitive development of this important social skill would improve human-robot interaction by making it more adaptive and natural.

5. ACKNOWLEDGMENTS

6. REFERENCES

[1] J. L. Elman, E. A. Bates, M. H. Johnson, A. Karmiloff-Smith, D. Parisi, and K. Plunkett. *Rethinking Innateness: A Connectionist Perspective on Development.* MIT Press, 1996.

[2] V. V. Hafner and G. Schillaci. From field of view to field of reach - could pointing emerge from the development of grasping? *Frontiers in Computational Neuroscience*, 2011.

[3] T. Kohonen. Self-organized formation of topologically correct feature maps. *Biological cybernetics*, 43(1):59–69, 1982.

[4] M. Lungarella and G. Metta. Beyond gazing, pointing, and reaching: A survey of developmental robotics. In *Epigenetic Robotics*, pages 81–89, 2003.

[5] A. Morse, J. de Greeff, T. Belpaeme, and A. Cangelosi. Epigenetic robotics architecture (era). *Autonomous Mental Development, IEEE Transactions on*, 2(4):325–339, 2010.

[6] G. Vettigli. MiniSom: minimalistic and numpy based implementation of the self organizing maps. https://github.com/JustGlowing/minisom, 2013.