

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/269401785>

Learning Visual-Motor Cell Assemblies for the iCub Robot using a Neuroanatomically Grounded Neural Network

CONFERENCE PAPER · DECEMBER 2014

READS

111

5 AUTHORS, INCLUDING:



[Samantha Vanessa Adams](#)

University of Plymouth

12 PUBLICATIONS 9 CITATIONS

[SEE PROFILE](#)



[Angelo Cangelosi](#)

University of Plymouth

226 PUBLICATIONS 2,630 CITATIONS

[SEE PROFILE](#)



[Max Garagnani](#)

Freie Universität Berlin

39 PUBLICATIONS 365 CITATIONS

[SEE PROFILE](#)



[Friedemann Pulvermüller](#)

Freie Universität Berlin

245 PUBLICATIONS 13,104 CITATIONS

[SEE PROFILE](#)

Learning Visual-Motor Cell Assemblies for the iCub Robot using a Neuroanatomically Grounded Neural Network

S. V. Adams, T. Wennekers, A. Cangelosi
Centre for Robotics and Neural Systems
Plymouth University, Plymouth, UK
samantha.adams@plymouth.ac.uk

M. Garagnani, F. Pulvermüller
Brain Language laboratory
Freie Universität Berlin
Berlin, Germany

Abstract— In this work we describe how an existing neural model for learning Cell Assemblies (CAs) across multiple neuroanatomical brain areas has been integrated with a humanoid robot simulation to explore the learning of associations of visual and motor modalities. The results show that robust CAs are learned to enable pattern completion to select a correct motor response when only visual input is presented. We also show, with some parameter tuning and the pre-processing of more realistic patterns taken from images of real objects and robot poses the network can act as a controller for the robot in visuo-motor association tasks. This provides the basis for further neurobotic experiments on grounded language learning.

Keywords—Neurorobotics; Cell Assemblies; Visual-Motor Learning

I. INTRODUCTION

The wider context for the current work is the design of a biologically-inspired architecture for brain embodied language skills in robots. This is the goal of the interdisciplinary BABEL project, which aims at the synthesis of computational neuroscience modelling (informed by and informing actual fMRI experiments), of the SpiNNaker neuromorphic hardware [1] and the humanoid robot iCub [2] for language learning.¹ In particular an *embodied* approach to language acquisition is chosen where visual input and motor interaction with the world constitute the grounding for language. Research in behavioural and cognitive neuroscience demonstrates that language, action and perception are closely linked in the brain [3-5]. There is now substantial neuroscientific evidence that the use of language activates brain areas closely joining together motor, perceptual and speech-language mechanisms (see review in [6]). This is consistent with the embodied view of cognition in psycholinguistics and cognitive science where cognitive functions, such as language, are closely integrated with sensorimotor knowledge [7], and with the *situated* learning approach to studying language in context. Consequently, there have been advances in both

computational neuroscience in the design of neuroanatomically grounded models of word acquisition [8-10] and also in cognitive robotics by the use of such brain-inspired models in the integration of action and language learning in robots [11,12]. In [11], Morse et al. trained the humanoid robot iCub to learn the names of objects, replicating the same phenomena observed in child language experiments. The robot's neural control architecture learned to integrate visual, motor and linguistic representations through Hebbian-linked Kohonen maps. In [12], Caligiore et al. developed the TRoPICALS model to study how vision, action and language are integrated in the representation and activation of affordances. The TRoPICALS model architecture is based upon real brain areas with a path relating to object identity, name and current goal (PFC: prefrontal cortex, STC: superior temporal cortex, VOT: ventral occipito-temporal area) and a path relating to object shape, position and desired position of the arm and hand to form the correct grasp (PMC: premotor cortex, PC: parietal cortex, SSC: somatosensory cortex; VC: visual cortex). However, these robotic models are limited to the use of relatively small numbers of neural elements to model cortical processing.

As its core model for language learning, BABEL uses the existing work of [9] which is a multi-area neural network model whose architecture is neuroanatomically grounded in the left perisylvian language cortex and includes biologically inspired learning. In its original formulation, the model had six areas and was used to investigate the association of auditory and articulatory stimuli. This model is currently being extended to 12 areas to incorporate an extra 'semantic' set of 6 areas for visual and motor (action) input. The ultimate aim of BABEL is to produce a spiking neuron implementation of the 12 area architecture so that it can be deployed on SpiNNaker hardware and integrated with the iCub robot, delivering a better real time performance than a software neural network running on a host PC would be able to achieve. This integrated platform will be used in experiments for object naming and word associations with action composition. As a first step in the roadmap for robotics experiments we investigated if it was possible to integrate the existing standalone neural model with the iCub

This work was supported under EPSRC Grant EP/J004561/1 (BABEL)

¹ <http://www.tech.plym.ac.uk/SoCCE/CRNS/babel/>

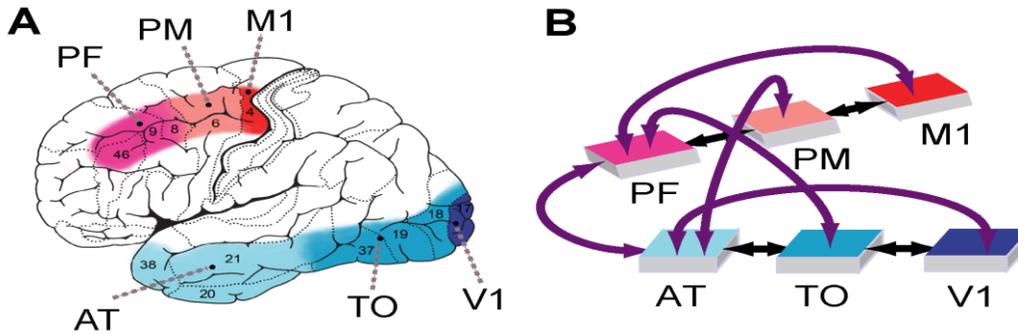


Fig. 1. Relevant brain areas and model architecture. (A) Sets of cortical areas modelled. Three areas relevant for the planning and execution of manual action (PF – prefrontal; PM – premotor; M1 – primary motor) in the frontal lobe, and three for processing information concerning visual object identity (AT – anterior-temporal; TO – temporo-occipital; V1 – primary visual) in the occipital and temporal lobes, are identified. Numbers indicate Brodmann Areas. (B) Overall between-area connectivity realised by the model. Colours indicate the mapping between model and corresponding cortical areas. A link (arrow) between two model areas indicates the presence of random and sparse – but not *all-to-all* – synaptic connections between cells of these areas. Links between model correlates of cortically adjacent areas are depicted in black; longer-distance cortico-cortical connections (i.e., between areas that are not cortically adjacent) in purple.

simulator to learn associations between the visual representation of an object and an appropriate grasping action. The structure of this paper is as follows: firstly we briefly describe the existing neural model and modifications that were required to allow learning of visual and motor associations with the iCub simulator. Next we introduce the iCub simulator and its communications protocol, YARP (Yet Another Robotic Platform) as well as describe the integration between the robot and the neural model. Lastly we describe experiments performed with this setup and present the results to show how well the system enabled iCub to learn the association of a visual object and the required grasping motion. We conclude with an evaluation of this prototype and future work required to enhance it in the wider context of the BABEL project aims.

II. LEARNING VISUAL-MOTOR ASSOCIATIONS

A. The Neural Model

The model architecture is shown in Fig. 1.B. In this section we present only a brief overview of salient features of the model—the mathematical formulation of the model, structural and connectivity details, and learning mechanisms, have been previously described elsewhere [9, 13-16]. We implemented a network consisting of six areas of artificial neurons (graded response cells) with reciprocal connections between and within areas. The model was constructed so as to mimic a range of properties of the brain, and especially the human cortex. We included the following features:

1. Area structure: six areas, modelling pre-specified sensorimotor and multimodal brain regions (see Fig. 1.A);
2. Between-area connectivity, constrained by specific neuroanatomical data and obeying general neuroanatomical principles in being sparse, random, and topographic;
3. Within-area connectivity, also sparse and random, and such that local connections are more likely than distant ones;
4. Local lateral inhibition and area-specific global regulation mechanisms;

5. Synaptic weight changes mediated by Hebb-type learning (simulating long-term potentiation and depression, LTP, LTD);
6. Neurophysiological dynamics of single cells including adaptation, spatial and temporal summation of inputs, nonlinear transformation of membrane potential into neuronal output (firing rate);
7. Constant presence of uniform uncorrelated white noise in all parts of the network.

In previous work [9,13-16] we have shown that networks that include this range of realistic neurobiological features tend to exhibit the formation of (input-specific) memory circuits, corresponding to what Hebb once postulated and labelled “Cell Assemblies” [17] or CAs. Below we identify the brain areas that the model simulated, and review the neuroanatomical evidence that we used as a basis to establish the presence, in the network, of direct links between pairs of such areas. The set of sensorimotor brain areas modelled include the left ventral visual “what”-system (blue-shaded areas in Fig. 1A) and the hand/arm motor system (red-shaded areas), involved in processing, respectively, visual object identity [18,19] and manual actions [20-23]. We take it that associative learning mechanisms at work in this system support the development of CAs binding together specific visual stimuli to hand/arm movements, as required, e.g., to pair up visual identity of objects with associated reaching and grasping actions [24-27], or, more generally, to acquire conditional visuomotor associations arbitrarily mapping visual features of objects to specific actions [28-31]. The neural model implemented (Fig. 1.B) reflects structure and connectivity of the above cortical areas. In particular, in the frontal lobe the hand/arm motor area M1 (in Brodmann Area, BA 4), adjacent PM (BA 6) and more rostral PF (BA 8/9/46) areas are reciprocally connected [20-22, 32-34], with cortico-cortical links documented between PF and M1 [35,36]. In addition, the primary visual, V1 (BA 17), temporo-occipital, TO (occipital lobe, inferior parts of BA 18/19, and posterior parts of the inferior and middle temporal gyri, BA 37) and anterior-temporal, AT (temporal pole, BA 38, and middle

parts of the middle and inferior temporal gyri, BA 20/21) areas (Fig. 1, blue-shaded areas) are reciprocally connected [37,38]. Direct (jumping) links from V1 to anterior temporal (AT) regions via the inferior longitudinal fascicle have also been documented [39, 40]. Neuroanatomical [41, 42], inactivation [43-45] and lesion [29,46,47] studies in the monkey indicate the presence of direct connections also between anterior-temporal and prefrontal cortices (AT and PF in Fig. 1.B). Evidence suggests the presence of direct links (via the external capsule) also between temporo-occipital (TO) and prefrontal areas (PF) [48-50] and between anterior/middle temporal (AT) and premotor (PM) areas [32]. This between-area connectivity structure is accurately replicated in the model (see Fig. 1.B).² In addition, local, within-area connections realised in the model reflect connectivity features that are shared by sensory and motor systems of the mammalian brain [51,52].

B. Input Patterns

We used three different types of input pattern depending upon the experiment. These are described in the subsections below.

1) Neural Model Representation (NMR)

In the standalone model of [9] input patterns are generated within the model code itself as 25x25 pixel pseudo-random patterns with minimal overlap to guarantee best training of CAs (henceforth referred to as Neural Model Representations or NMRs). We used this type of pattern in Experiment 1 for verification of the Cell Assembly learning (see section IV A). See Fig. 2 for an example of an NMR encoding for one pattern pair.

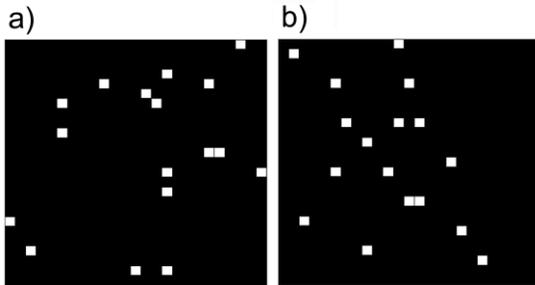


Fig. 2. Example Neural Model pattern pair a) visual, b) motor

2) ICub Simulator Representation (ISR)

We also required encodings to represent images of objects and grasp poses within the simulator. We used a repertoire of 6 objects with shapes loosely matching those of everyday real objects we plan to use in the final experiments with real iCub (Ball, Can, Cup, Pen, Peg, Phone) and that require different grasping actions (‘power’ or ‘precision’). Fig. 5 shows an example of iCub looking at one of the objects generated in the simulated world. The visual encoding was generated from

² Note that some of the links in the model (purple arrows in Fig. 1.B) represent long-distance cortico-cortical connections: in particular, those between areas AT and PF, which are “adjacent” in the model but not so in neuroanatomical terms.

iCub’s view of an object by performing the following image-processing steps using OpenCV: convert from RGB to greyscale, Canny edge detection, thresholding to produce a binary image and resizing to 25x25. Grasping poses were generated by taking the joint angles from the iCub head, right arm and torso required to make a particular grasping action, totalling 18 Degrees Of Freedom (DOF). These were then encoded as a 25x25 pixel pattern by representing the DOF by rows 0-17 (leaving 7 rows unused but with the potential to add more DOF later) and the value of the joint angle as the x pixel using the formula in (1).

$$x = \text{floor} \left[\frac{(\text{angle} - \text{min angle})}{(\text{max angle} - \text{min angle})} * \text{max } x \right] \quad (1)$$

where: x is the calculated x pixel value for the DOF, angle is the joint angle for the DOF, min angle and max angle are the allowed joint ranges for the DOF, max x is the maximum x pixel used (as the joint ranges vary for each DOF we do not always use all 25 x pixels in order to make the calculation more convenient). The ‘floor’ operation rounds down to the nearest whole pixel value. These patterns are henceforth referred to as iCub Simulator Representations or ISRs. For the simulator experiments the ISR patterns were used as an intermediate step in identifying the training information to be sent to the neural model: the actual neural model training was done with patterns in NMR form as described in Section II B 1). The reason for this was that the neural model had previously only been trained with the pseudo-random patterns and more realistic data presented problems in that a small overlap between patterns and a similar number of ‘on’ pixels per pattern could not be guaranteed. In the iCub controller code a simple translation between representations is done so that recognition of the object seen and the correct grasp pose to execute are done using the ISR but the data sent to the neural model are in NMR form. See Fig. 3 for an example of the ISR encoding for the ‘Ball’ object and its corresponding grasp pattern.

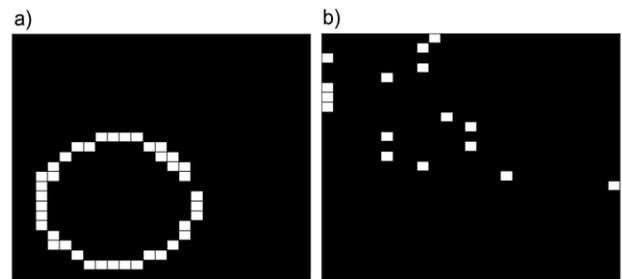


Fig. 3. Simulator representation for the Ball object a) visual, b) motor

3) Real data Representation

Given the aforementioned issues with using real data, a separate strand of our work has been to investigate a parameter tuning of the model and/or data pre-processing that allows training with real world data so that we can use the real iCub and real visual input directly with the model rather than using intermediate steps as described in the previous section.

We report our preliminary results for this in Experiment 3 (Section IV C). For this we created a visual encoding from six real images selected from the Amsterdam Library of Object Images (ALOI)³ (Ball, Can, Cup, Peg, Pen, Phone) [53]. OpenCV was used to perform basic image processing (convert from RGB to greyscale, Canny edge detection, thresholding to produce a binary image and resizing to 25x25). To achieve successful training with the neural model also required that some noise be added to the resulting pixel patterns (some pixels added and some removed) to break up the outlines. See Fig. 4 for an example of the processing stages from image to final pattern for the Cup object. We used the same encoding for the motor part of the pattern as described in Section II B 2) without any further processing.



Fig. 4. Image processing of a real image of a cup to training pattern

III. THE ICUB SIMULATOR AND YARP

The iCub simulator is an accurate physical simulation of the real iCub allowing prototyping of applications that can easily be transferred directly to the real robot with minimal changes. See Fig 5. As well as basic motor control and visual processing, for more complex tasks iCub is integrated with external libraries such as the OpenCV image processing framework. For communications, iCub uses YARP, a generic and flexible protocol which we have used to connect with the neural model (a standalone C program).



Fig. 5. The simulated iCub and world

Fig. 6 gives an overview of the integrated system. The communication is handled through a main iCub/YARP C++ program which retrieves visual and motor information from the simulator via YARP ports (oval shapes in Fig. 6.). This program also manages the iCub ‘world’ by adding and removing objects and issuing movement commands. The main

program also creates and manages the YARP port that the external model connects to in order to receive training data. At the model end this port is connected to as if it were a regular socket but with some special handshaking commands required to establish the protocol for exchanging information with a YARP port.

IV. EXPERIMENTS AND RESULTS

A. Experiment 1 - Verification of Cell Assembly learning

We firstly wanted to verify that the neural model could learn associations between a candidate set of 6 visual and motor inputs (in NMR form) well enough for use in a practical robotics scenario. Firstly, that the network could encode 6 robust Cell Assemblies concurrently and secondly that the pattern completion ability was good enough so that when only the visual part of the pattern was presented, the network would produce an unambiguous motor response. To achieve robust learning of 6 CAs required some parameter retuning from the original study of [9] which was optimised for 12 patterns. We investigated various parameters (e.g. Noise, Global inhibition, Learning Rate) and found that the Noise parameter had the greatest effect on learning. We increased noise from 145 to 200 in our networks which gave a robust and consistent performance. We also found that fewer training presentations were required: minimum 150-200 presentations of each pattern to see learning of CAs as opposed to a minimum of 500 in [9]. Training in the neural model proceeded as described in [9]. In summary: A pattern pair was selected randomly from the set of 6 (with a mechanism to ensure that roughly equal numbers of each have been presented by the end of the run). The visual part was applied to area 1 (V1 in Fig. 1.B) while the motor part was presented simultaneously to layer 6 (M1 in Fig 1.B). This input was presented for a fixed number of model timesteps followed by a fixed number of ‘noise’ presentations. Then the process was repeated for another pattern until the desired number of presentations of each pattern had been achieved. For Experiment 1 we performed 7 independent runs to be sure that the performance was robust to different network initialisations, and trained each network with 150 presentations of each pattern pair (900 presentations in total per run). For a first analysis we presented only the visual half of the input and observed the network response. See Fig. 7 for an example of the network activation when the visual component of a pattern was presented and the target motor component for comparison. In this figure the left hand side shows the visual input (first box) and activation in area 1 (V1) through to area 6 (M1). The right hand side of the figure shows the visual pattern and the target motor pattern. By visual inspection, for all of the 6 patterns a cell assembly was present (a clear subset of neurons in each area were activated all the way through to area 6). For a quantitative analysis we calculated the percentage pixel match between the activated motor response and all possible target patterns – here we were trying to assess if the pattern completion was good enough to distinguish between the possible responses.

³ <http://aloi.science.uva.nl/>

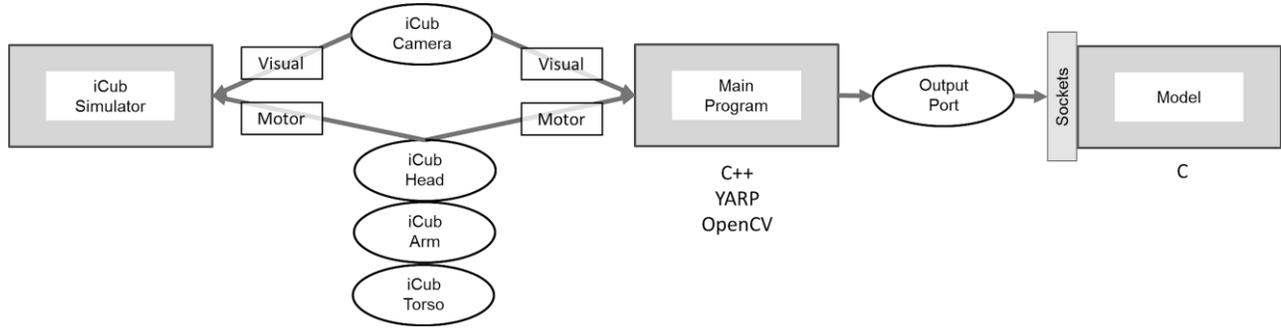


Fig. 6. Integration of the iCub Simulator and the Neural Model – oval shapes are YARP ports, larger rectangles are software components

Table I shows a matrix of all possible matches with values averaged over all runs (to nearest whole %). Although in many cases there is some activation of pixels from multiple target patterns, in all cases the pattern match with the intended target is the highest (diagonal in Table I). Therefore we concluded that the network could learn Cell Assemblies for 6 patterns well enough for use with the iCub simulator.

TABLE I. EXPERIMENT I PIXEL MATCHES

Pattern	1	2	3	4	5	6
1	66	1	1	3	1	0
2	3	71	1	3	0	1
3	2	0	84	0	1	2
4	0	1	1	54	0	0
5	1	2	2	2	76	2
6	0	2	0	3	3	71

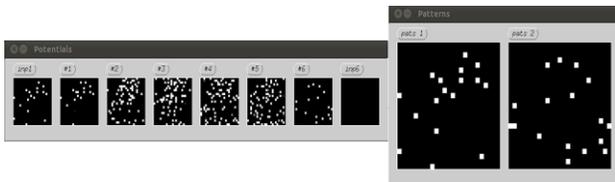


Fig. 7. Example of pattern completion from Experiment 1

B. Experiment 2 - Learning with the iCub simulator

This experiment consisted of two parts: firstly training online with the iCub simulator generating the inputs to the network in real time and secondly testing where only the visual input was applied to see if the network could generate the correct motor response to pass back to the simulator.

The training procedure:

1. Set iCub in 'ready' pose (See Fig. 5)
2. Select a random number between 1-6

3. Create the corresponding object in the world
4. iCub looks at object
5. Image of object is processed to work out what iCub saw (ISR)
6. iCub grasps object using appropriate pose (ISR)
7. Appropriate NMR data sent to neural model
8. Network is trained
9. Set iCub back to ready pose
10. Repeat 2-9 as required

The testing procedure:

1. Set iCub in 'ready' pose (as for training sequence)
2. Select a random number between 1-6
3. Create the corresponding object in the world
4. iCub looks at object
5. Image of object is processed to work out what iCub saw (ISR)
6. NMR data for visual input is sent to neural model (motor is blank)
7. Neural model produces a motor response
8. Completed pattern pair sent back to iCub
9. iCub translates NMR to ISR representation
10. iCub grasps object
11. Set iCub back to ready pose
12. Repeat 2-11 as required

Fig 8. Shows example screenshots of the training and testing scenarios. We trained a network for 900 presentations (150 for each of the 6 objects) and then in testing mode presented 50 randomly selected objects. We collected data on the actual object presented, whether it was recognised correctly as one of the candidate objects and whether the motor response back from the neural model was correct. We found that the object was recognised correctly in 100% of the cases. The motor response generated by the model was correct for 94 % of the cases. Because the NMR representation is translated to ISR at the iCub end a valid grasping pose is always executed even if it is the incorrect pose for the object.

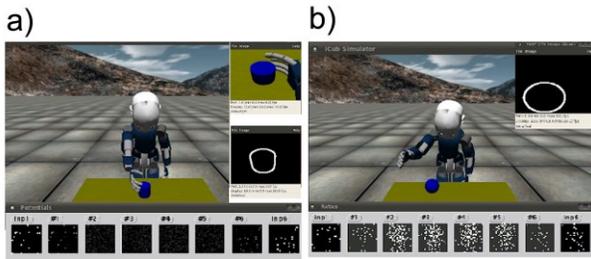


Fig. 8. Screenshots from a) the training and b) the testing scenarios in Experiment 2.

C. Experiment 3 – Training with Real Data

As mentioned previously, the real dataset inevitably has more overlap between patterns and (for the visual inputs) different numbers of ‘on’ pixels per pattern. We found that the best regime to learn the CAs stably and avoid merging of the responses was with further increased noise (220) and increased global inhibition (100 instead of the original value of 85). We also found that more training presentations were required (approximately 400 per pattern; 2400 in total). Similarly to Experiment 1 we performed 7 independent runs and trained each network with 400 presentations of each pattern pair. Again patterns were selected randomly, but balanced to ensure roughly equal presentations of each pattern. On presentation of only the visual input to a trained network, we calculated the percentage pixel match between the activated motor response and all possible target patterns. Table II shows a matrix of all possible matches with values averaged over all runs (to nearest whole %). In most cases the diagonal values are the highest with the exception of the match of the Pen object which is also activating the Peg object quite strongly.

TABLE II. EXPERIMENT 3 PIXEL MATCHES

Pattern	Ball	Can	Cup	Peg	Pen	Phone
Ball	38	6	0	0	2	8
Can	6	33	2	2	3	4
Cup	2	4	38	18	11	4
Peg	1	1	24	63	38	7
Pen	1	1	28	51	51	8
Phone	5	0	0	0	0	34

Although they are generally good enough to distinguish between the different CAs, the pixel match percentages are lower than those for Experiment 1 reflecting the difficulties with using more realistic data.

V. DISCUSSION AND CONCLUSIONS

We demonstrated how it was possible to ‘plug and play’ an existing neural model with a robotic simulator. However, there was a trade-off between the convenience of re-using an existing model and the compromises needed to get it working with ‘real’ data. Although minimal changes were needed to

the neural model code (mainly in the communications and I/O) there were limitations in the kind of patterns that could be successfully learned as the model was not originally designed for use in this type of scenario. We have made progress in showing that, with suitable processing and parameter tuning it would be possible for the model to learn more realistic data but more work needs to be done in fully integrating this with real iCub: in particular, a method for online presentation and processing of the images of real objects. We also found (at least with the current setup) that motor patterns are never 100% completed (best individual case was 94 %). In the current simulator setup the motor response from the network is translated back via its ISR form thereby ensuring that iCub always receives a valid motor command, but for future implementations it would be desirable to use the output directly from the network without recourse to intermediate processing. The issue of appropriate pre-processing and encoding of visual, motor and auditory data is relevant to the next steps for BABEL as it is likely that any spiking version of the model will also be sensitive to overlaps and varying information content in the input patterns. The principle of including noise and employing a global inhibition mechanism may also work well for this type of network.

An important next step for our prototype is to extend to the upcoming 12-layer architecture incorporating auditory and motor articulatory input so that experiments with object naming can be done.

As well as the presented results, our takeaway message is that there are advantages to be had in this type of plug and play integration of neural models and robots (simulated or otherwise) rather than creating whole systems from scratch. In technical terms integration need not be difficult but the biggest issue is the compromises required depending upon assumptions made in the respective components. Even if such systems are not likely to provide robust permanent solutions they are ideal as intermediate steps to explore candidate neural models.

REFERENCES

- [1] X. Jin, M. Lujan, L.A. Plana, S. Davies, S. Temple and S. Furber, “Modeling spiking neural networks on SpiNNaker”, *Computing in science & engineering*, vol. 12, 2010.
- [2] G. Metta, G. Sandini, D. Vernon, L. Natale and F. Nori. “The iCub humanoid robot: an open platform for research in embodied cognition”, in *Proc. IEEE Workshop on Performance Metrics for Intelligent Systems (PerMIS08)*, 2008.
- [3] F. Pulvermüller, “Words in the brain’s language”. *Behavioral and Brain Sciences*, vol. 22, pp.253-336, 1999.
- [4] F. Pulvermüller and L. Fadiga, “Active perception: Sensorimotor circuits as a cortical basis for language”, *Nature Reviews Neuroscience*, vol. 11(5), pp. 351-360, 2010.
- [5] M.A. Arbib, “Mirror system activity for action and language is embedded in the integration of dorsal and ventral pathways”, *Brain and Language*, vol. 112(1), pp. 12-24, 2010.
- [6] L.W. Barsalou, “Grounded cognition”, *Annual Review of Psychology*, vol. 59, pp. 617-645, 2008.
- [7] F. Pulvermüller, “Brain mechanisms linking language and action”, *Nature Reviews Neuroscience*, vol. 6(7), pp. 576-582, 2005.

- [8] P.F. Dominey and T. Inui, "Cortico-striatal function in sentence comprehension: insights from neurophysiology and modeling", *Cortex*, vol. 45(8), pp. 1012-8, 2009.
- [9] M. Garagnani, T. Wennekers, and F. Pulvermüller, "A neuroanatomically-grounded Hebbian learning model of attention-language interactions in the human brain" *European Journal of Neuroscience*, vol. 27, pp. 492-513, 2008.
- [10] F.T. Husain, M.A. Tagamets, S.J. Fromm, A. Braun and B. Horwitz "Relating neuronal dynamics for auditory object processing to neuroimaging activity: a computational modeling and an fMRI study", *Neuroimage*, vol. 21, pp. 1701-1720, 2004.
- [11] A. Morse, T. Belpaeme, A. Cangelosi and L. Smith, L., "Thinking with your body: modelling spatial biases in categorization using a real humanoid robot" in *The Proceedings of the Annual meeting of the Cognitive Science Society (CogSci 2010)*, 2010.
- [12] D. Caligiore, A. Borghi, D. Parisi, R. Ellis, A. Cangelosi and G. Baldassarre, "How affordances associated with a distractor object affect compatibility effects: A study with the computational model TRoPICALS" *Psychological Research*, vol. 77, pp. 7-19, 2012.
- [13] M. Garagnani, T. Wennekers and F. Pulvermüller, "Recruitment and consolidation of cell assemblies for words by way of Hebbian learning and competition in a multi-layer neural network", *Cognitive Computation*, vol. 1(2), pp 160-176, 2009.
- [14] M. Garagnani and F. Pulvermüller, "From sounds to words: a neurocomputational model of adaptation, inhibition and memory processes in auditory change detection", *Neuroimage*, vol. 54(1), pp 170-181, 2011.
- [15] M. Garagnani and F. Pulvermüller, "Neuronal correlates of decisions to speak and act: spontaneous emergence and dynamic topographies in a computational model of frontal and temporal areas", *Brain and Language*, vol. 127(1), pp. 75-85, 2013.
- [16] F. Pulvermüller and M. Garagnani, "From sensorimotor learning to memory cells in prefrontal and anterior-temporal cortex: A neurocomputational study of disembodiment", *Cortex*, vol. 57, pp. 1-21, 2014.
- [17] D.O. Hebb, *The organization of behavior*. New York: John Wiley, 1949.
- [18] L.G. Ungerleider and M. Mishkin, "Two cortical visual systems", in *Analysis of Visual Behaviour*. D.J. Ingle, M.A. Goodale, and R.I.W. Manfield, Eds. Cambridge (MA): MIT Press, 1982, pp. 549-586.
- [19] L.G. Ungerleider and J.V. Haxby, "'What' and 'where' in the human brain". *Curr Opin Neurobiol*, vol. 4(2), pp. 157-65, 1994.
- [20] R.P. Dum and P.L. Strick, "Frontal lobe inputs to the digit representations of the motor areas on the lateral surface of the hemisphere", *J Neurosci*, vol. 25(6), pp.1375-86, 2005.
- [21] M.T. Lu, J.B. Preston, and P.L. Strick, "Interconnections between the prefrontal cortex and the premotor areas in the frontal lobe". *J Comp Neurol*, vol. 341(3), pp. 375-92, 1994.
- [22] R.P. Dum and P.L. Strick, "Motor areas in the frontal lobe of the primate", *Physiol Behav*, vol. 77(4-5), pp. 677-82, 2002.
- [23] M.-P. Deiber et al., "Cortical areas and the selection of movement: a study with PET". *Experimental Brain Research*, vol. 84, pp. 393-402., 1991.
- [24] S.T. Grafton, L. Fadiga, M.A. Arbib and G. Rizzolatti, "Premotor cortex activation during observation and naming of familiar tools", *Neuroimage*, vol. 6(4), pp. 231-6, 1997.
- [25] M.A. Arbib, "The nervous system, motor control", in *Handbook of Physiology*, V.B. Brooks, Ed. 1981, Baltimore:Williams and Wilkins, 1981, pp. 1449-1480.
- [26] J. Grezes and J. Decety, "Does visual perception of object afford action? Evidence from a neuroimaging study". *Neuropsychologia*, vol. 40(2), pp. 212-22, 2002.
- [27] V. Gallese, L. Fadiga, L. Fogassi and G. Rizzolatti, "Action recognition in the premotor cortex", *Brain*, vol. 119(Pt 2), pp. 593-609, 1996.
- [28] I. Toni, M.F. Rushworth, and R.E. Passingham, "Neural correlates of visuomotor associations. Spatial rules compared with arbitrary rules" *Exp Brain Res*, vol. 141(3), pp. 359-69, 2001.
- [29] M.J. Eacott and D. Gaffan, "Inferotemporal-frontal disconnection: the uncinate fascicle and visual associative learning in monkeys". *Eur J Neurosci*, vol. 4(12), pp. 1320-1332, 1992.
- [30] I.M. White and S.P. Wise, "Rule-dependent neuronal activity in the prefrontal cortex", *Exp Brain Res*, vol. 126(3), pp. 315-35, 1999.
- [31] R. Muhammad, J.D. Wallis, and E.K. Miller, "A comparison of abstract rules in the prefrontal cortex, premotor cortex, inferior temporal cortex, and striatum", *J Cogn Neurosci*, vol. 18(6), pp. 974-89, 2006.
- [32] G. Rizzolatti and G. Luppino, "The cortical motor system", *Neuron*, vol. 31(6), pp. 889-901, 2001.
- [33] D.N. Pandya and E.H. Yeterian, "Architecture and connections of cortical association areas", in *Cerebral cortex. Vol. 4. Association and auditory cortices*, A. Peters and E.G. Jones, Eds, London: Plenum Press, 1985, pp. 3-61.
- [34] T.K. Arikuni, T., K. Watanabe, and K. Kubota, "Connections of area 8 with area 6 in the brain of the macaque monkey", *J Comp Neurol*, vol. 277(1), pp. 21-40, 1988.
- [35] M. Guye, et al., "Combined functional MRI and tractography to demonstrate the connectivity of the human primary motor cortex in vivo", *Neuroimage*, vol. 19(4), pp. 1349-60, 2003.
- [36] M.P. Young, J.W. Scannell, and G. Burns, *The analysis of cortical connectivity*. Heidelberg: Springer, 1995.
- [37] H. Nakamura, R. Gatass, R. Desimone and L.G. Ungerleider, "The modular organization of projections from areas V1 and V2 to areas V4 and TEO in macaques". *J Neurosci*, vol. 13(9), pp. 3681-91, 1993.
- [38] C. Distler, D. Boussaoud, R. Desimone and L.G. Ungerleider, "Cortical connections of inferior temporal area TEO in macaque monkeys", *J Comp Neurol*, vol. 334(1), pp. 125-50, 1993.
- [39] M. Catani, D.K. Jones, R. Donato and D.H. Ffytche, "Occipito-temporal connections in the human brain", *Brain*, vol. 126(Pt 9), pp. 2093-107, 2003.
- [40] S. Wakana, H. Jiang, L.M. Nagae-Poetscher, P.C.M. van Zijl and S. Mori, "Fiber tract-based atlas of human white matter anatomy", *Radiology*, vol. 230(1), pp. 77-87, 2004.
- [41] L.G. Ungerleider, D. Gaffan, and V.S. Pelak, "Projections from inferior temporal cortex to prefrontal cortex via the uncinate fascicle in rhesus monkeys", *Exp Brain Res*, vol. 76(3), pp. 473-84, 1989.
- [42] M.J. Webster, J. Bachevalier, and L.G. Ungerleider, "Connections of inferior temporal areas TEO and TE with parietal and frontal cortex in macaque monkeys", *Cereb Cortex*, vol. 4(5), pp. 470-83, 1994.
- [43] M.V. Chafee and P.S. Goldman-Rakic, "Inactivation of parietal and prefrontal cortex reveals interdependence of neural activity during memory-guided saccades", *J Neurophysiol*, vol. 83(3), pp. 1550-66, 2000.
- [44] J.M. Fuster, R.H. Bauer, and J.P. Jervey, "Functional interactions between inferotemporal and prefrontal cortex in a cognitive task", *Brain Research*, vol. 330, pp. 299-307, 1985.
- [45] R.H. Bauer and J.M. Fuster, "Delayed-matching and delayed-response deficit from cooling dorsolateral prefrontal cortex in monkeys", *J Comp Physiol Psychol*, vol. 90(3), pp. 293-302, 1976.
- [46] A. Parker and D. Gaffan, "Interaction of frontal and perirhinal cortices in visual object recognition memory in monkeys", *Eur J Neurosci*, vol. 10(10), pp. 3044-57, 1998.
- [47] A. Parker, A. and D. Gaffan, "Memory after frontal/temporal disconnection in monkeys: conditional and non-conditional tasks, unilateral and bilateral frontal lesions", *Neuropsychologia*, vol. 36(3), pp. 259-71, 1998.
- [48] N. Makris and D.N. Pandya, "The extreme capsule in humans and rethinking of the language circuitry", *Brain Struct Funct*, vol. 213(3), pp. 343-58, 2009.
- [49] B. Seltzer and D.N. Pandya, "Frontal lobe connections of the superior temporal sulcus in the rhesus monkey" *J Comp Neurol*, vol. 281(1), pp. 97-113, 1989.
- [50] D.N. Pandya and C.L. Barnes, "Architecture and connections of the Frontal Lobe", in *The Frontal Lobes Revisited*, E. Perecman, Ed, New York: The IRBN Press, 1987, pp. 41-72.

- [51] V. Braitenberg and A. Schüz, *Cortex: statistics and geometry of neuronal connectivity*. Berlin: Springer, 1998.
- [52] R.J. Douglas and K.A. Martin, "Neuronal circuits of the neocortex", *Annu Rev Neurosci*, vol. 27, pp. 419-51, 2004.
- [53] J. M. Geusebroek, G. J. Burghouts, and A. W. M. Smeulders, "The Amsterdam library of object images" *Int. J. Comput. Vision*, vol. 61, pp. 103-112, 2005.