

# Towards Real-World Neurorobotics: Integrated Neuromorphic Visual Attention <sup>★</sup>

Samantha V. Adams, Alexander D. Rast, Cameron Patterson, Francesco Galluppi,  
Kevin Brohan, José-Antonio Pérez-Carrasco, Thomas Wennekers, Steve Furber and  
Angelo Cangelosi

<sup>1</sup> Plymouth University, Plymouth, UK

<sup>2</sup> School of Computer Science, University of Manchester  
Manchester, UK M13 9PL

{rast}@cs.man.ac.uk

{steve.furber}@manchester.ac.uk

{samantha.adams,A.Cangelosi,thomas.wennekers}@plymouth.ac.uk

<http://www.cs.manchester.ac.uk/apt>

<http://www.plymouth.ac.uk>

**Abstract.** Neuromorphic hardware and cognitive robots seem like an obvious fit, yet progress to date has been frustrated by a lack of tangible progress in achieving useful real-world behaviour. System limitations: the simple and usually proprietary nature of neuromorphic and robotic platforms, have often been the fundamental barrier. Here we present an integration of a mature “neuromimetic” chip, SpiNNaker, with the humanoid iCub robot using a direct AER - address-event representation - interface that overcomes the need for complex proprietary protocols by sending information as UDP-encoded spikes over an Ethernet link. Using an existing neural model devised for visual object selection, we enable the robot to perform a real-world task: fixating attention upon a selected stimulus. Results demonstrate the effectiveness of interface and model in being able to control the robot towards stimulus-specific object selection. Using SpiNNaker as an embeddable neuromorphic device illustrates the importance of two design features in a prospective neurorobot: *universal configurability* that allows the chip to be conformed to the requirements of the robot rather than the other way round, and *standard interfaces* that eliminate difficult low-level issues of connectors, cabling, signal voltages, and protocols. While this study is only a building block towards that goal, the iCub-SpiNNaker system demonstrates a path towards meaningful behaviour in robots controlled by neural network chips.

**Keywords:** cognitive, robotics, attention, neuromorphic

---

<sup>★</sup> Alexander Rast and Steve Furber are with the School of Computer Science, The University of Manchester, Manchester, UK (email: rast@cs.man.ac.uk). Samantha Adams, Thomas Wennekers and Angelo Cangelosi are with Plymouth University, Plymouth, UK. Francesco Galluppi is with the Institut de la Vision, Paris, France. José-Antonio Pérez-Carrasco is with the Departamento de Teoría de la Señal y Comunicaciones, Universidad de Sevilla, Seville, Spain. Cameron Patterson is with Supplemango Networks Ltd., Aberdeen, UK.

## 1 Introduction: The Need for Practical Neuromorphic Robotics

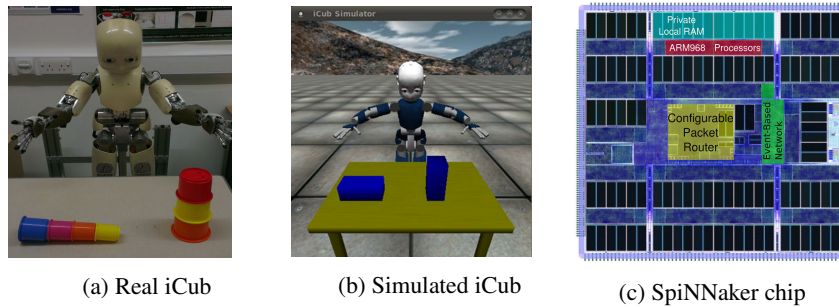


Fig. 1: The iCub Robot and SpiNNaker chip

Neural networks seem like an obvious fit for robots. Indeed, behavioural roboticists assert, with some justification, that *embodiment matters*: it is not enough, to achieve meaningful behaviour, to implement a neural network as an abstract “disembodied brain” operating on synthetic stimuli in an artificial environment [11]. Within neuro-robotics itself, advances in understanding of the neurobiology have many researchers suggesting that neural models more closely matching the biology [1] may more clearly reveal the computational principles necessary for cognitive robotics while illuminating human (and animal) brain function. Critically, some neurobiological experiments suggest that spike-based signalling is important at the cognitive/behavioural level as well as in learning [12]. If this is true perhaps spike-based neurobots can embody behavioural features that seem intractably hard without neural hardware [7]. However, until now, in practice most neurobotic systems, e.g. [3] have simulated the neural component on an external host PC due to internal hardware constraints or incompatible interfaces. Both the hardware [4] and the robotic systems [2] have now reached a point of maturity where *integrated* neurobots able to demonstrate effective behaviour in nontrivial real-world scenarios are within reach. Our aim is to create such a system in a way that both fulfils the long-held promise of practical neurorobotics and illustrates their potential to act as a tool for neurobiological experimentation. Here we introduce a system integrating 2 mature platforms over a direct interface: the humanoid iCub robot (Figs. 1a, 1b) and the embeddable neuromimetic SpiNNaker chip (Fig. 1c) to solve a behaviourally relevant task: goal-directed attentional selection.

## 2 Test Materials and Methods

There are 3 main components to our test system: the iCub robot and its associated support systems, the SpiNNaker system with its AER interface, and the network model.

## 2.1 Neural Model

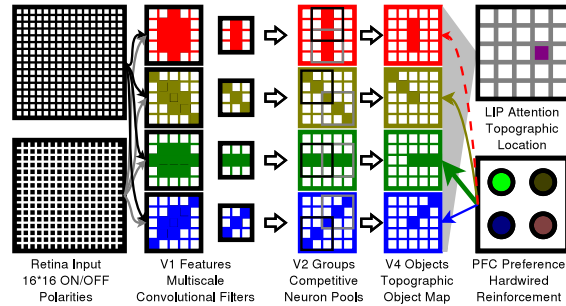


Fig. 2: The test model. The input retina layer is a real or simulated visual field taken either from the preprocessed robot imaging system or from a software image generator. Each of layers V1, V2, V4, and PFC are separated into 4 orientations per layer. Layer LIP merges orientations via a winner-take-all.

As a test network, we use the attentional model of F. Galluppi, K. Brohan, et al. in [5]. We chose this network because it exhibits realistic real-world behaviour in a non-trivial task while remaining simple, proven, and scalable. The network (Fig. 2) has 6 layers roughly corresponding to selected brain areas. We retained the original network parameters and sizes:  $16 \times 16$  neuron visual field,  $10 \times 10$  image maps in 4 orientations for both V1 and V2, and  $5 \times 5$  location fields for V4, PFC, and LIP. However we made the following modifications/additions to the network:

1. We sharpened the winner-take-all filtering in the LIP of the original model.
2. We extended the input stimulus to allow testing with either real or synthetic sources.
3. We inserted an automatic parameter scaling module in the PyNN script.
4. We added an option to enable STDP learning between the V2 and V4 layers.

## 2.2 SpiNNaker

The SpiNNaker chip (Fig. 1c) is a universal neural network platform designed for real-time simulation with an array of programmable cores operating in parallel over a configurable asynchronous multicast interconnect. While the typical neuromorphic device implements a fixed model, SpiNNaker can be easily programmed by users with a wide range of different models. Inter-processor communications occur exclusively through Address-Event Representation (AER) spikes: small packets carrying only the address of the neuron that spiked (and possibly a 32-bit data payload).

To interface to SpiNNaker, external devices (such as the robot) send spikes rather than pass data structures directly. This conveniently and data-efficiently abstracts away internal processing particulars on both sides of the interface. SpiNNaker supports the AEthernet general-purpose direct AER interface [10], transmitting up to 256 spikes per

frame as 32-bit words via UDP over an Ethernet connection. We enhanced the original AetheRnet interface with additional support for multiple (up to 6) input and output devices (spike sources and spike receivers) and an internal software router to allow multicast of output spikes to multiple devices simultaneously.

### 2.3 iCub

iCub is a flagship humanoid developmental robotics platform [8]. See Fig. 1. There has been a limited amount of research implementing spiking neural networks for control of the iCub, such as the work of Bouganis and Shanahan [2] and the iSpike library developed by Gamez, et al. [6]. These have relied on a host PC to run the neural network whereas our aim is direct execution on neuromorphic hardware.

We use YARP (Yet Another Robot Platform) for a communications protocol and also Aquila - an easy-to-use, high-performance, modular and scalable software architecture for cognitive robotics [9]. In particular we use the Tracker module for extraction of objects from the scene and basic image processing, transforming the 240x320 RGB raw image from a single iCub camera into a downsampled 16x16 image of black and white pixels which are then converted to spikes by mapping “ON” pixels to spike outputs. We also use the iCubMotor module which converts image coordinates into head motor movements to enable the iCub to look at a location corresponding to a point in a 2D image. We then configured the iCub as a virtual AetheRnet device with a bidirectional link which maps iCub camera input to the two input layers and receives spikes from the LIP output layer.

## 3 Results

We ran 4 experiments. Experiment I was a simulation using synthetic visual inputs to develop an initial weight tuning for the network. Experiments II, III and IV tested the ability of the network (with optimised weights, STDP off), and the iCub robot, to locate and attend to the location of a preferred object when it was the only object present and when two objects (one preferred and one aversive) were present in the scene.

### 3.1 Experiment I

To explore the weight parameters of the network, we first ran a set of preliminary tests using the synthetic stimulation model noted in 2.1. The stimulus (3a) was a pair of black horizontal and vertical bars with slow upwards and rightwards motion respectively. We successively tuned the spike outputs for each of layers V1, V2, V4, and LIP to respond just below its continuous spiking threshold - at the point of maximum sensitivity (Fig. 3). As can be seen in Fig. 3f, when correctly tuned the output LIP layer was able, with synthetic inputs, to narrow its output (area of visual fixation) to a single neuron.

### 3.2 Experiment II

In this experiment we tested the ability of the robot to attend to single stimuli (either a horizontal or vertical object). Figure 4 shows the raw and preprocessed input. Figure 5

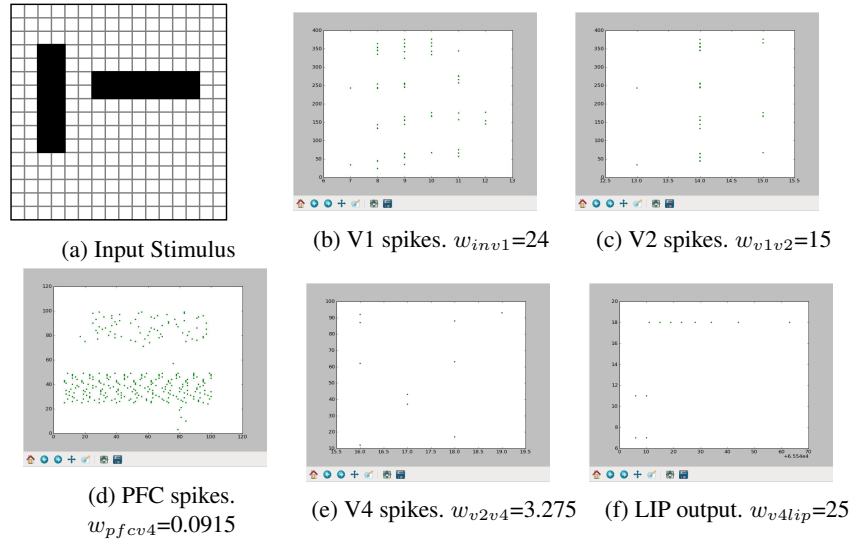


Fig. 3: Synthetic input and spike raster plots for each layer. The x-axis is time. Neuron numbers are depicted on the y-axis in ascending order, by orientation first, then by row, then by column. For each layer other than LIP, each orientation is  $\frac{1}{4}$  of the population. The LIP layer has only one combined orientation. Neuron 18 thus corresponds to the neuron at row 4, column 4 in the output field.

shows the LIP maps (fixation frequency by location) produced. The maps show spike count over a 1 second run (normalised so that values lie between 0.0 and 1.0) with lighter coloured areas indicating higher saliency. In both cases the network is able to determine whether the object to left or to the right is the preferred object.

### 3.3 Experiment III

In this experiment we tested the ability of the robot to attend to one object when both were present in the scene. Figures 4g and 4h show the input stimulus in this case. Figure 5c shows the LIP saliency map (attentional preference) produced when the horizontal object is preferred and the vertical object is aversive and Fig. 5d the LIP map when the vertical object is preferred and the horizontal is aversive. In both scenarios, although there is activity produced for both objects, the network is able to determine which one is the preferred object and thus the more active location. Table 1 compares the estimated mean firing rates in both scenarios for the V2 and V4 layers and shows the effect of the biasing of V4 by the PFC. In the V2 layer there is very little discrimination between the preferred and aversive stimuli but in V4 the activity of the preferred stimulus is amplified and that of the aversive stimulus is suppressed.

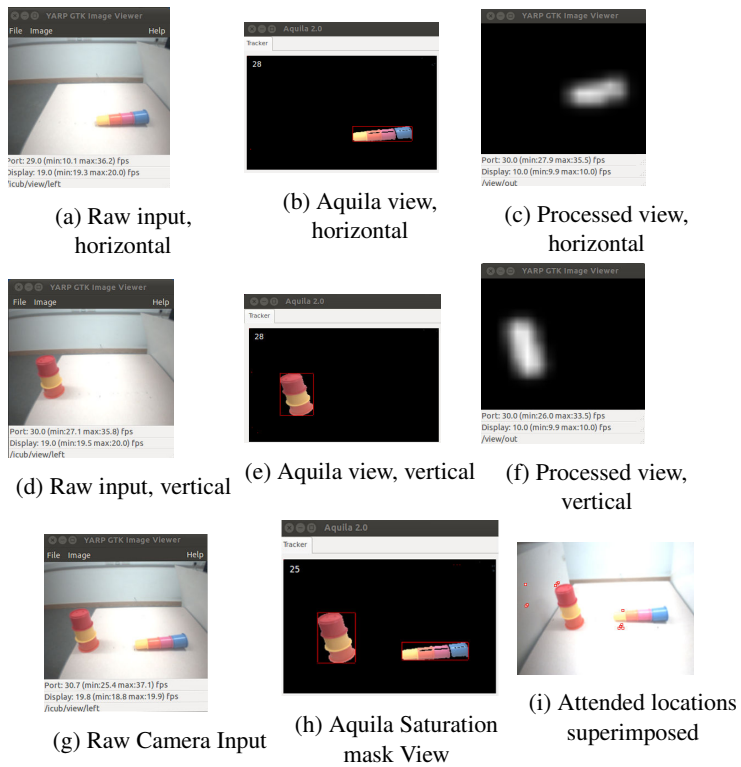


Fig. 4: iCub camera view and Aquila saturation masks for objects taken separately and together. Downsampled black and white images for the horizontal and vertical objects (top). Superimposed attended location map (bottom right).

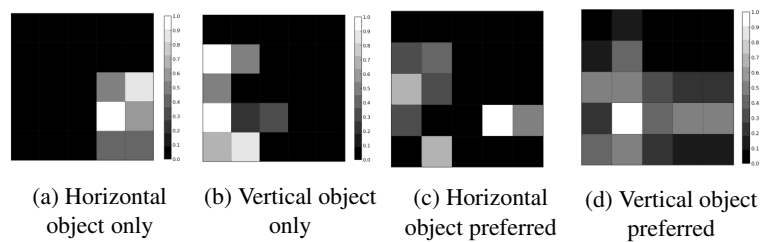


Fig. 5: LIP saliency maps for Experiment II (single stimuli) and Experiment III (dual stimuli)

Table 1: Estimated mean firing rates (Hz) in V2 and V4 for dual stimuli

Preferred	Aversive	V2 (unbiased)		V4 (biased)	
		P	A	P	A
Vertical	Horizontal	8.04	7.63	55.48	4.92
Horizontal	Vertical	8.34	7.83	23.0	4.48

### 3.4 Experiment IV

In this experiment we ran 20 repeats of Experiment III, 10 of each with the preferred orientation set as horizontal and vertical respectively. Figure 4i shows the raw camera view with a selection of the locations that were attended to superimposed as red squares. Of the 20 runs 14 resulted in the correct object being selected. Nearly equal activation in both areas of the LIP map caused the wrong object to be selected in 6 cases.

## 4 Discussion: The Test of Real-World Applications

Our system significantly expands the model complexity that can be embedded in a neuro-robot. Some issues remain to be addressed. *Scalability*: The small network we used allows only a broad distinction between 2 objects; we are scaling it to sizes with finer discrimination. *Interfaces*: Translating from the frame-based camera was rather artificial. Experiments with a neuromorphic version of the iCub show that native AER is clearly the preferred interface. *Learning*: Disabling STDP during tests clarified interface aspects at the expense of cognitive aspects. SpiNNaker-only trials of the model with STDP on have run successfully and will be tested on the iCub. *Context*: Using vision alone tends to lead to a fixed context but future plans include expanded top-down processing in the PFC and the addition of bio-inspired auditory processing to assist with learning the preferred stimulus. When these systems are in place we hope to be in a position to run a genuinely practical cognitive test. We propose several key criteria about what kind of application would demonstrate achievement of practical neurorobotics:

**Adaptive Behaviour:** The application should require the robot to synthesise behaviours that have not been specified imperatively.

**Intrinsic Unpredictability:** The environment should generate unpredictable events whose impact on behaviour is more than perturbative.

**Multidimensional Cognition:** Tasks should involve the integration of multiple sensory and response modes.

**Real-World Applicability:** Behaviour should be meaningful in a real human context.

## 5 Conclusions

We have integrated neuromorphic computing and humanoid robotics to perform a cognitive task that, while equally feasible with traditional robotics techniques, can here be

achieved with a model scalable to a variety of contexts. In possibly nonstationary situations, where decisions may need to be made faster than their consequences can be predicted, classical offline machine learning and Bayesian approaches may fall short. A cognitive system using configurable neural hardware with standard interfaces, running biologically derived models, offers perhaps the critical margin of flexibility. Our experiments point the way to one of the goals of cognitive robotics: self-directed robots able to respond adaptively and appropriately rather than imperatively to the combination of unexpected events and indeterminate consequences characteristic of the real world.

## Acknowledgements

This work was supported under EPSRC Grant EP/J004561/1 (BABEL). The SpiNNaker project is supported by EPSRC Grant EP/G015740/1, and by industry partners ARM, Silistix, and Thales.

## References

1. Arbib, M., Metta, G., van der Smagt, P.: Springer Handbook of Robotics, chap. Neuro-robotics: from vision to action, pp. 1453–1480. Springer-Verlag (2008)
2. Bouganis, A., Shanahan, M.: Training a spiking neural network to control a 4-DOF robotic arm based on spike timing-dependent plasticity. In: Proc. 2010 Int'l Joint Conf. Neural Networks (IJCNN2010). pp. 4104–4111 (2010)
3. Chersi, F.: Learning Through Imitation: a Biological Approach to Robotics. *IEEE Trans. Autonomous Mental Development* 4(3), 204–214 (Sep 2012)
4. Furber, S.B., Lester, D.R., Plana, L.A., Garside, J.D., Painkras, E., Temple, S., Brown, A.D.: Overview of the SpiNNaker system architecture. *IEEE Trans. Computers* PP(99) (2012)
5. Galluppi, F., Brohan, K., Davidson, S., Serrano-Gotarredona, T., Carrasco, J.P., Linares-Barranco, B., Furber, S.: A Real-Time, Event-Driven Neuromorphic System for Goal-Directed Attentional Selection. In: Proc. 19th Int'l. Conf. on Neural Information Processing (ICONIP 2012). pp. 226–233 (2012)
6. Gamez, D., Fidjeland, A., Lazdins, E.: iSpike: a spiking neural interface for the iCub robot. *Bioinspiration and Biomimetics* 7(2) (2012)
7. Kuniyoshi, Y., Berthouze, L.: Neural learning of embodied interaction dynamics. *Neural Networks* 11(7-8), 1259–1276 (Oct-Nov 1998)
8. Metta, G., Sandini, G., Vernon, D., Natale, L., Nori, F.: The iCub humanoid robot: an open platform for research in embodied cognition. In: Proc. IEEE Workshop on Performance Metrics for Intelligent Systems (PerMIS08) (2008)
9. Peniak, M., Morse, A., Cangelosi, A.: Aquila 2.0 software architecture for cognitive robotics. In: Proc. IEEE Third Joint International Conference on Development and Learning and Epigenetic Robotics (ICDL) (2013)
10. Rast, A., Partzsch, J., Mayr, C., Schemmel, J., Hartmann, S., Plana, L.A., Temple, S., Lester, D.R., Schüffny, R., Furber, S.: A Location-Independent Direct Link Neuromorphic Interface. In: Proc. 2013 Int'l Joint Conf. Neural Networks (IJCNN2013). pp. 1967–1974 (2013)
11. Scheier, C., Pfeifer, R., Kuniyoshi, Y.: Embedded neural networks: exploring constraints. *Neural Networks* 11(7-8), 1551–1569 (Oct-Nov 1998)
12. Shmiel, T., Drori, R., Shmiel, O., Ben-Shaul, Y., Nadasdy, Z., Shemesh, M., Teicher, M., Abeles, M.: Temporally precise cortical firing patterns are associated with distinct action segments. *The Journal of Neurophysiology* 96, 2645–2652 (2006)