MANCHESTER 1824

SpiNNaker-based Visual Systems

End-of-first-year report

Garibaldi Pineda García

Supervisor: Steve Furber
Co-supervisor: Dave Lester

Advanced Processing Technologies Group
School of Computer Science
University of Manchester
United Kingdom

SpiNNaker

University of Manchester
APT Group

Contents

Contents

Chapter 1

Introduction

A picture is worth a thousand words
Frederick R. Barnard - 1921

For most animals, an important part of perception is done through visual input.

This kind of input has given humans the possibility of culture through reading and writing, cognitive development.

We might take vision for granted, but we are reminded of its importance when we hear an unusual noise in a dark room.

An important aspect of vision is our ability to create mental maps of our current or, even past, locations.

The goal of the project is to create a 3D environment reconstruction.

There has been work on this field on “classic” computer vision, but for real-time they rely on high-performance power-hungry devices. Something that limits the actual utility of such systems for mobile applications. It would be a bit inconvenient to carry around a couple of car batteries in ones pocket.

We are able to do it with a highly-parallel 20-watts neural blob. There must be a more efficient way of doing this. A brief description of the brain and its function can be found in Chapter 2. We delve into the components of human vision in Chapter 3.

SpiNNaker provides a massively-parallel high-efficiency computing platform, inspired by the brain. It’s an excellent choice for neuroscience research, particularly to study spiking neural networks. Its software stack has many ready-to-use neural models and development of new models can be performed in a straight forward manner. Chapter 4 has a more detailed description of this and other neuromorphic hardware.

Input for spiking neural networks has to be in *spike trains*, which are a series of spikes emitted by a neuron in a given time slot. In order to use video sources they need conversion. Few solutions which, mostly, require the use of custom hardware which is expensive and has low availability. We propose a parallel software based encoding.

This years work consisted in creating an input system for our spiking neural networks; details of this can be found in Chapter 5.

While there are some examples of hardware based retinas, they are still expensive or they have limited availability. Implementing a retina model using consumer hardware is of great help for people that are unable to obtain a silicon retina.

Of special interest are mobile applications, if we can provide a low-power solution to a silicon retina emulator, we could enable millions of phones, tablets or computers to work as an input to neural computations (QUALCOMM CHIP, SPINNAKER) and keep the traditional camera functionality.

1.1 Objectives

3D environment reconstruction is a very active field of research. Advances in depth perception (KINECT) have made real-time simultaneous localization and mapping (SLAM) a possibility. One way of achieving is to use high-performance GPUs and solve the problem using raw power. Another is to use a mix of KINECT and RETINA, not fully neural??? We propose using an exclusively neural networks approach using SpiNNaker hardware.

1.2 Plan

Steps:

Image recognition

Time-based encoding, learning, classification, deep belief networks comparison, hierarchical structures

3D object recognition

Correlation in space and time, spiking neural networks should make an excellent match for this.

Depth perception

Binocular, depth-from-defocus, other sensors? Optic flow to infer motion?

Orientation and localization

Even more sensors? Make statistics/probabilistic models of past data?

Reconstruction

Get a top down approach? Interface 2 nets?

Chapter 2

A look into the brain

The brain is an exquisite piece of evidence of energy-efficient biological computation.

Neural systems in animals are different, from the simple ones found in insects to more complex ones in reptiles, birds and mammals. After millions of years of evolution, great apes, humans in particular, have one of the most intricate nervous systems. The human brain acts as regulator of this system and performs high level cognitive tasks.

It can perform the most diverse activities, from bird spotting to mathematics to art. All of this with about 20 watts of energy spread across many small computational units called *neurons*.

It consists of around 10^{12} individual neurons which are interconnected through about 10^{15} synapses.

Most of this neurons are arranged in thin sheet of about 1100 cm^2 area and a 2 to 4 mm. thickness.

So far, several functional units in the brain have been identified; but the exact mechanisms of how they perform is still unknown.

One of the most consuming tasks is vision, about 30% of the cerebral cortex is used in visual perception.

2.1 Neurons and responses

The nervous system is composed of specialized cells called *neurons*. Their area of expertise is long-range communication. While most cells in the body can "talk" to their neighbours, neurons have structures that allow them to communicate for up to XX cm (in the human case).

Neurons are composed of a soma (body), this part has similar components to other cells in the body. One of the specialized communication structures is called the *axon*, through it the neuron outputs a signal. *Dendrites* are at the other end of the information exchange and receive messages that other neurons sent through their axon, thus they can be seen as inputs for the cell.

Theories suggest they perform some kind of calculation, most of the time modelled as a threshold activation function

Some neurons use analog/continuous signals to transmit information, though they are mostly act as an interface to the exterior world.

Latest evidence suggests that the complex cognitive functions are performed using spikes, on-off responses, as a means for communication.

The place where axons and dendrites meet, is called the SSS, synapses

When one neuron's output elicits another neuron to spike,

2.2 Different languages

Neurons communicate using different "languages", spike-codes

rate time rank-order

Input from sensors is most likely rate-based, though processing time and energy consumption in the brain suggests a different one is used for further processing

2.3 Artificial neural networks

First modelled as an on-off threshold gate, perceptron

Multi-layered networks and feedback, Hebbian learning

Spiking neural networks, third gen, include time as a factor, more accurate, more powerful mathematical properties,

learning studied using Hebbian back-prop, stdp, bcm

still work to be done on time-based learning

2.4 Conclusions

conclusions the brain

Chapter 3

Vision

Vision is one of the most important senses for animals; humans use it extensively for all kinds of tasks. Hunting, assessing danger, reading, driving, drawing, predicting rain from grey clouds, etc., these are all tasks that involve *seeing*.

There is a vast collection of knowledge about the components of vision, though a unified theory of vision (or the answer to *How do we see?*) has not yet been achieved.

Vision starts at the eye, which transforms electromagnetic radiation that assembles an image, into voltage pulses that our brain may interpret. This encoded images are sent to the posterior region of the brain through the optic nerves. The cortex then performs many computations that result in our ability to see.

3.1 The eye and the retina

Our everyday experience might lead us to believe that the eyes are sensory organs developed completely separate from the brain but, in fact, the retina is an extension of the brain that performs spatio-temporal compression of a continuous flow of "images" of the world.

The eye is composed of many parts that resemble a camera (LENS, CAMERA OSCURA, FILM)

After light has been transformed into an electrical representation, the retina takes over and computes a representation of it.

First layers (bipolar, horizontal cells) use analog signals, ganglion cells use spike trains.

Ganglion cells extend to the Lateral Geniculate Nucleus, where information is relayed and organized so that the cortex can interpret it.

Organization makes left visual field sent to right hemisphere, right field to left hemisphere.

3.2 The visual cortex

The portion of the cortex that is involved with visual processing has been estimated to about 30%.

It has been studied and areas have been labelled due to their function.

V1, V2, V...

3.3 Eyes for computers

Traditionally cameras have been used as they work similar to the first stage of the eye.

Recently dynamic vision sensor

They work in a way that more resembles the retina, event based

Still need development, resolution is low, they do not perform multiple-scale convolution, expensive

A mix of both can be a great tool for researchers and commercial applications see Chp 5

Chapter 4

Neuromorphic hardware

4.1 Classical computing

classical computing

4.2 Neuromorphic trends

neuromorphic hardware trends

4.3 SpiNNaker

spinnaker info

4.4 Event-based model

event-based programming/infrastructure

4.5 Conclusions

conclusions neuro hardware

Chapter 5

From images to spikes

5.1 Real-time encoding

real-time encoding

5.2 SpiNNaker implementation

spinnaker encoding

5.3 Dataset creation

dataset for article

5.4 Conclusions

conclusions rank-ordered images

Chapter 6

Conclusions

6.1 Conclusion

6.2 Further work

6.3 Plans for second and third year

Bibliography