

Garibaldi Pineda García

Supervisor: Prof. Steve Furber



The University of Manchester

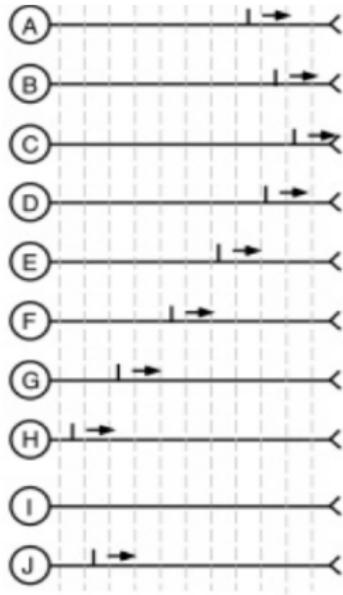
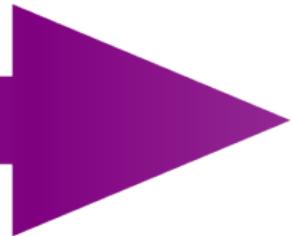
Computer Vision on the SpiNNaker Platform

A picture is worth a thousand words

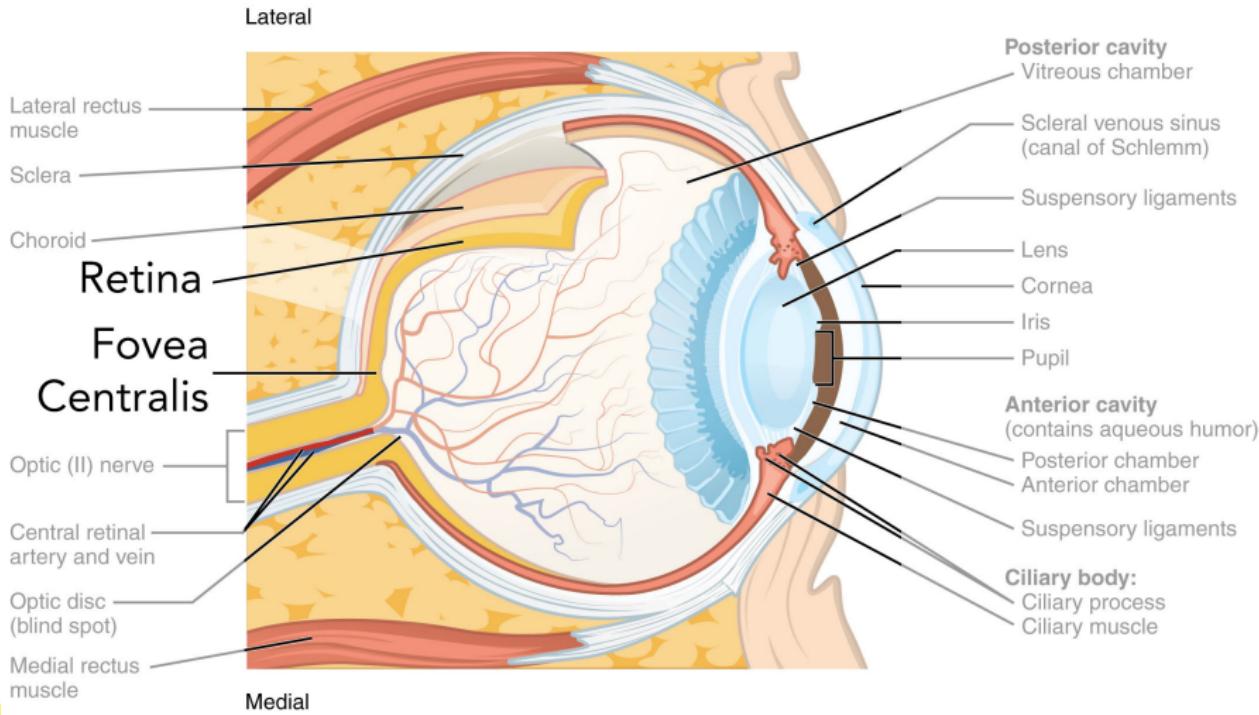


- Spiking Neural Network Architecture
- Massively parallel processing
- Inspired by the structure/function the human brain





Convert images/video from a regular camera to spike trains



- **Foveal pit** model with *rank-ordered* output¹
- Center-surround behaviour



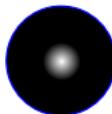
Midget
Off-centre



Midget
On-centre

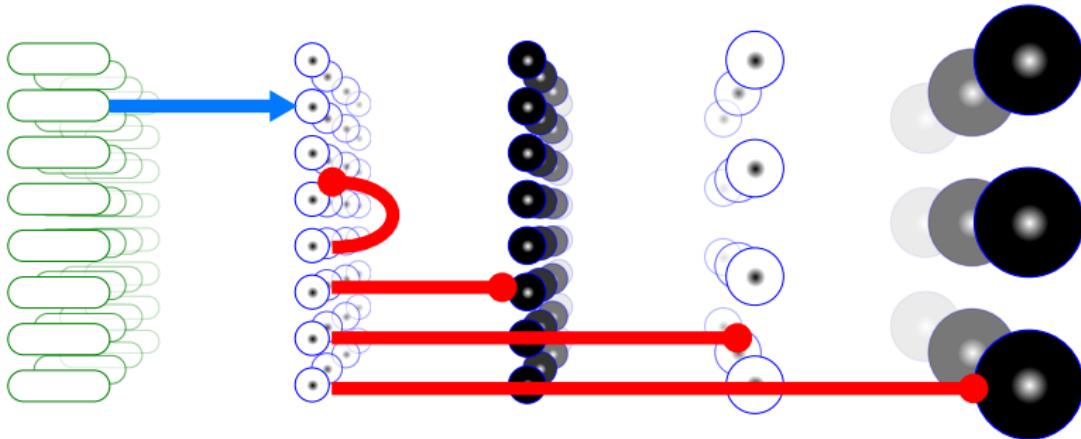


Parasol
Off-centre

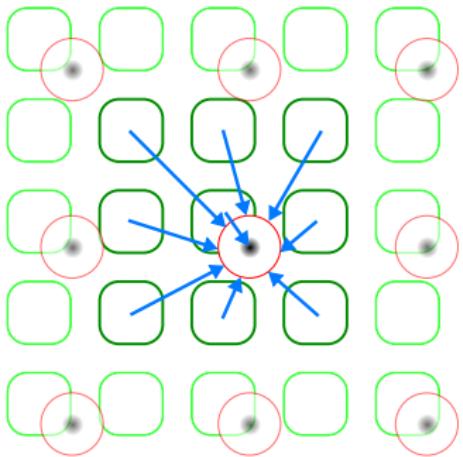


Parasol
On-centre

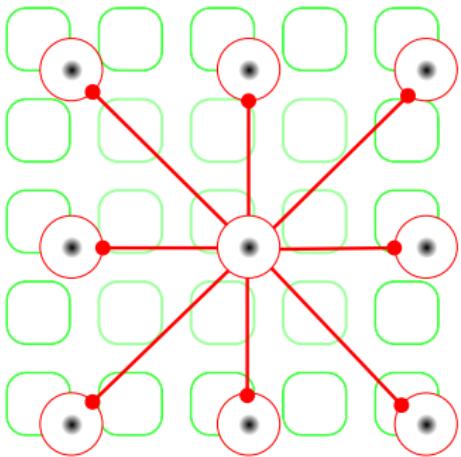
Foveal pit connections



- Excitatory connections from photoreceptors to ganglion cells.
- Inhibitory connections to and from ganglion cells.



Excitatory
Convolution
Parallel

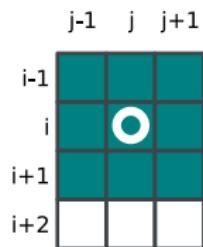


Inhibitory
Adjust weights
Serial

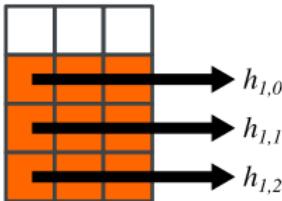
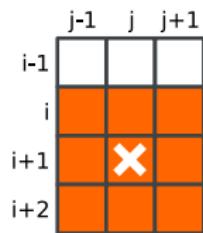
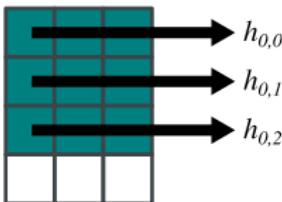
$$\begin{matrix} \text{Surround} \\ \begin{matrix} \textcolor{blue}{\square} \\ \textcolor{red}{\square} \\ \textcolor{blue}{\square} \end{matrix} \end{matrix} * \begin{matrix} \textcolor{blue}{\square} & \textcolor{red}{\square} & \textcolor{blue}{\square} \end{matrix} + \begin{matrix} \textcolor{red}{\square} \\ \textcolor{blue}{\square} \\ \textcolor{red}{\square} \end{matrix} * \begin{matrix} \textcolor{blue}{\square} & \textcolor{red}{\square} & \textcolor{blue}{\square} \end{matrix} = \begin{matrix} \textcolor{red}{\square} & \textcolor{yellow}{\square} & \textcolor{blue}{\square} & \textcolor{yellow}{\square} & \textcolor{red}{\square} \\ \textcolor{yellow}{\square} & \textcolor{blue}{\square} & \textcolor{blue}{\square} & \textcolor{blue}{\square} & \textcolor{yellow}{\square} \\ \textcolor{blue}{\square} & \textcolor{blue}{\square} & \textcolor{red}{\square} & \textcolor{blue}{\square} & \textcolor{blue}{\square} \\ \textcolor{red}{\square} & \textcolor{yellow}{\square} & \textcolor{blue}{\square} & \textcolor{yellow}{\square} & \textcolor{red}{\square} \end{matrix}$$

Separated convolution, perform four 1D convolutions instead of one 2D. Better for kernels bigger than 3×3 .

FoCal - convolution



$$h_{i,j} = k_0 \cdot I_{i,j-1} + k_1 \cdot h_{i,j} + k_2 \cdot h_{i,j+1}$$



$$h_{i,j} = k_0 \cdot I_{i,j-1} + k_1 \cdot h_{i,j} + k_2 \cdot h_{i,j+1}$$

$$v_o = q_0 \cdot h_{0,0} + q_1 \cdot h_{0,1} + q_2 \cdot h_{0,2}$$

$$v_x = q_0 \cdot h_{I,0} + q_1 \cdot h_{I,1} + q_2 \cdot h_{I,2}$$



$$h_{I,0} = h_{0,1}$$

$$h_{I,1} = h_{0,2}$$

Shared between
pixels \textcircled{o} and \texttimes



$$v_x = q_0 \cdot h_{0,1} + q_1 \cdot h_{0,2} + q_2 \cdot h_{I,2}$$

Tiled convolution, based on kernel
separability property, reuse calculations.

- Convolution of 4 cell types using OpenCL @ 12 fps
- Issues: Memory transfers for huge kernels

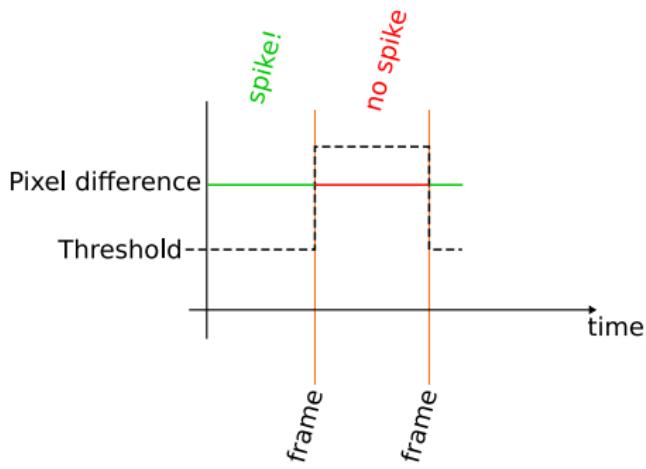


- Weight correction is slow,
SpiNNaker and/or custom hardware
- Issues: massive connectivity
(all-to-all), serial process



Second retinal model

- Emulate silicon retina
- Sense changes in illumination
- Per-pixel adaptive threshold



- Camera bound dynamic range, spacial and temporal resolution
- Cheaper and easier to use!



- Real-time convolution is possible
- SpiNNaker based FoCal
- DVS emulator, working @ 25 fps on OpenCL ... but needs more testing



- Keep improving video sources
- Learning with time-based spikes
- Spatio-temporal patterns
- Polychronous networks



Thank You!

Contact me:

pinedagg@cs.man.ac.uk

