Building and Interpreting Populations of Model Visual Cortical Neurons

Bert SHI

Dept. of ECE, HKUST
eebert@ee.ust.hk

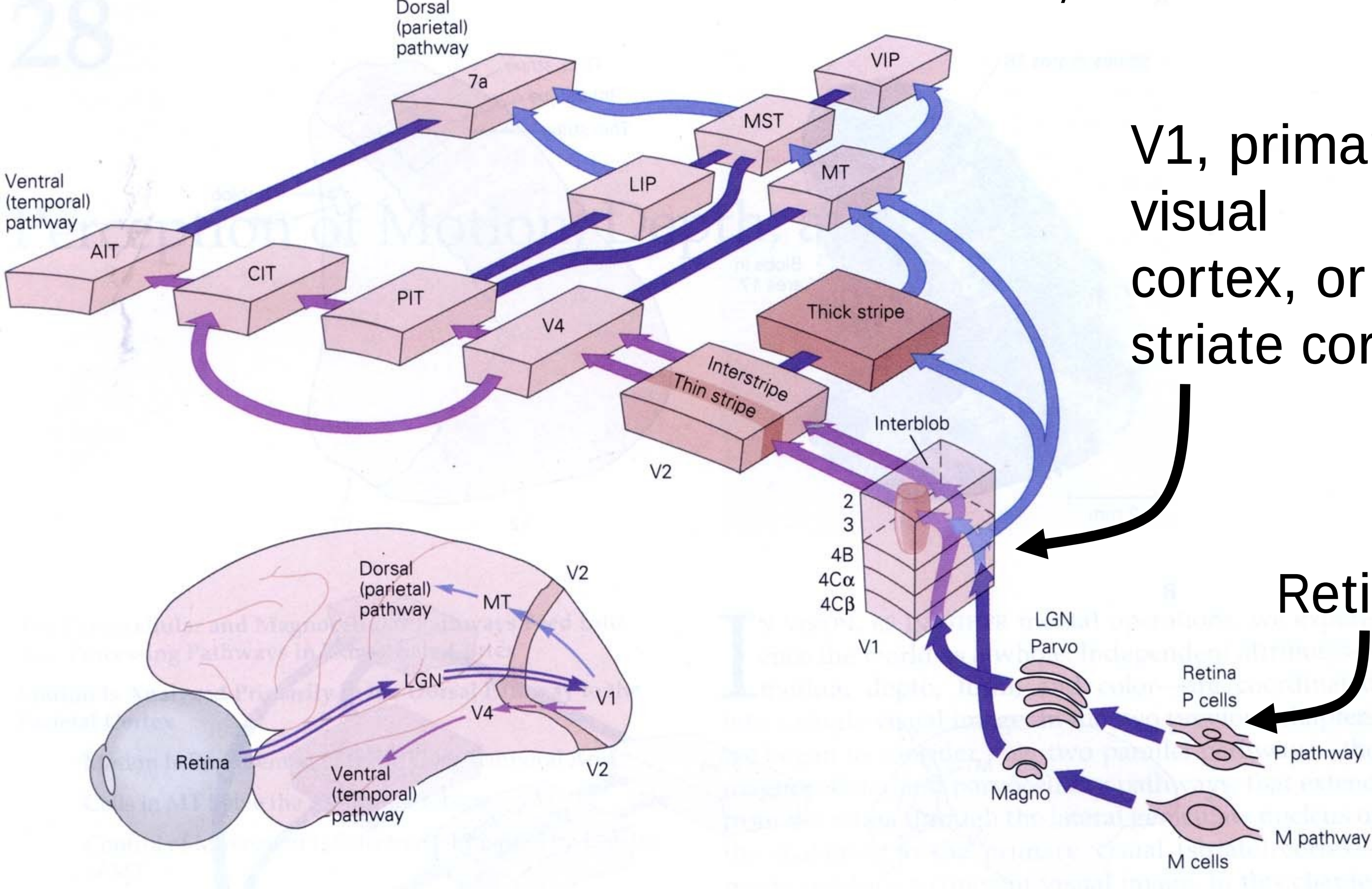
香港科技大学
HONG KONG UNIVERSITY OF SCIENCE & TECHNOLOGY

Overview

- Visual cortical processing
  - Multidimensional selectivity
- Orientation Selectivity
  - What is it?
  - How do we implement it?
    - AER Multi-chip Architecture
    - DSP/FPGA Architecture
- Joint Orientation/Disparity Selectivity
  - Disparity energy model: position versus phase shifts
  - Population responses versus tuning curves
  - Bigger is not necessarily better

Cortical Visual Processing

From E. R. Kandel et. al., *Principles of Neural Science*

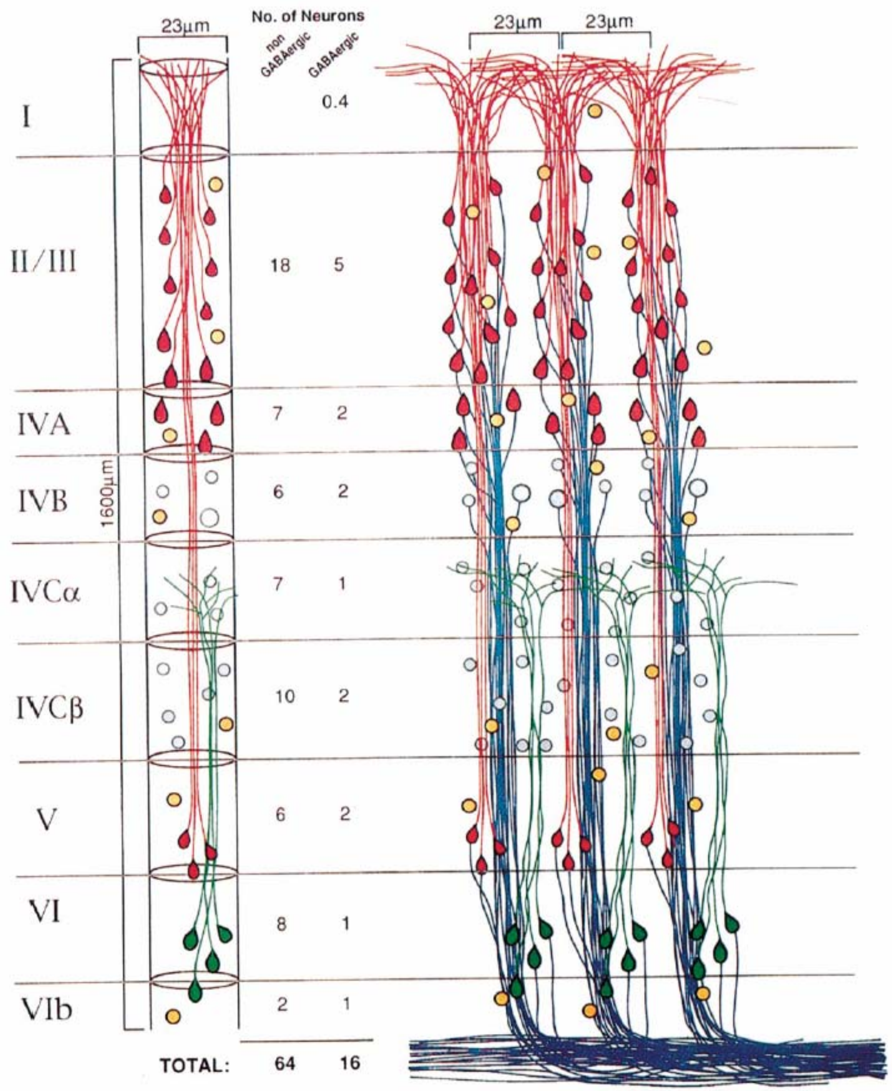


V1, primary visual cortex, or striate cortex

Retina

Level of Abstraction: Minicolumn

- "the effective unit of operation … is not the single neuron and its axon, but groups of cells with similar functional properties and anatomical connections."
- "The basic unit of the mature neocortex is the minicolumn, a narrow chain of neurons extending vertically across the cellular layers II–VI, perpendicular to the pial surface (Mountcastle, 1978). Each minicolumn in primates contains ~80–100 neurons, except for the striate cortex where the number is ~2.5 times larger."
  - V. M. Mountcastle, "The columnar organization of the neocortex", *Brain*, vol. 120, pp. 701–722, 1997.



Multidimensional selectivity

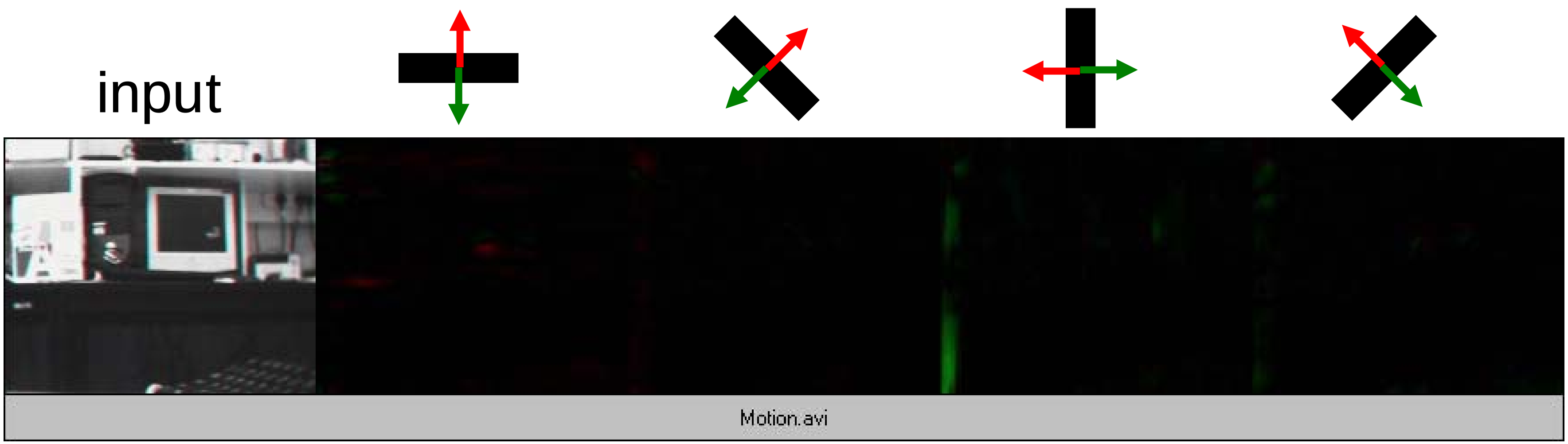
- Position
- Spatial frequency (size)
- Temporal frequency (change)
- Color

Retina/
LGN

- Orientation
- Binocular Disparity (depth)
- Direction/speed of motion
- Curvature

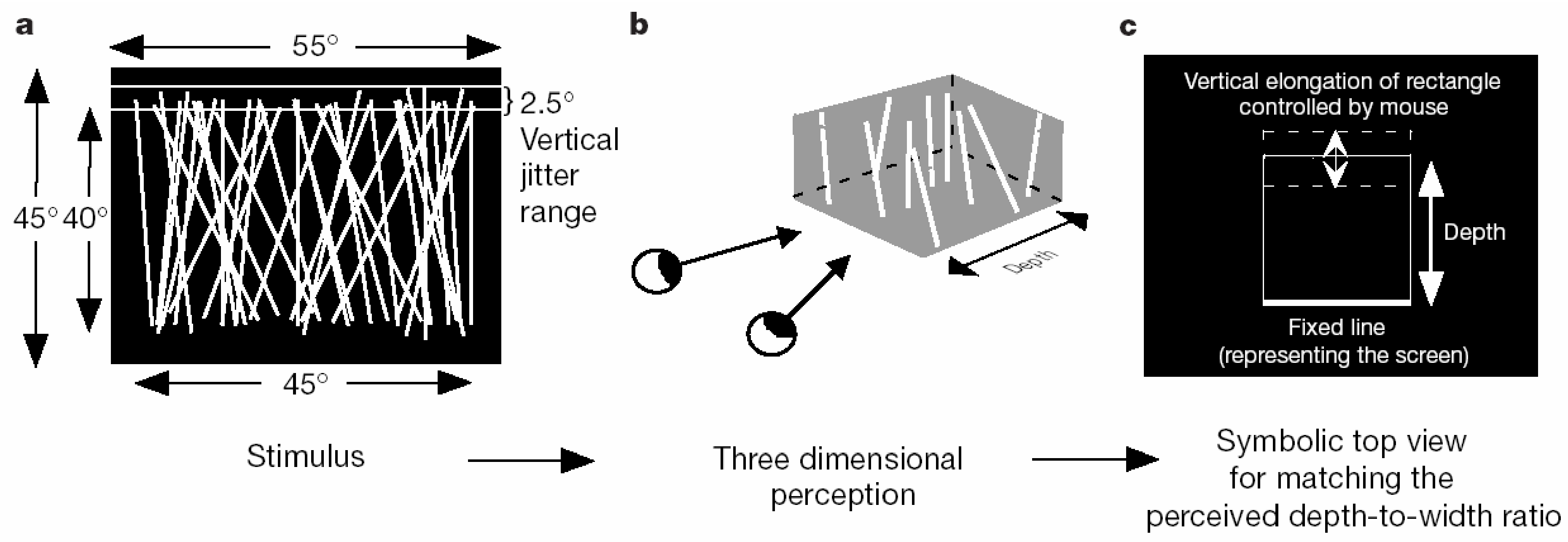
Primary
Visual
Cortex
(V1)

Examples

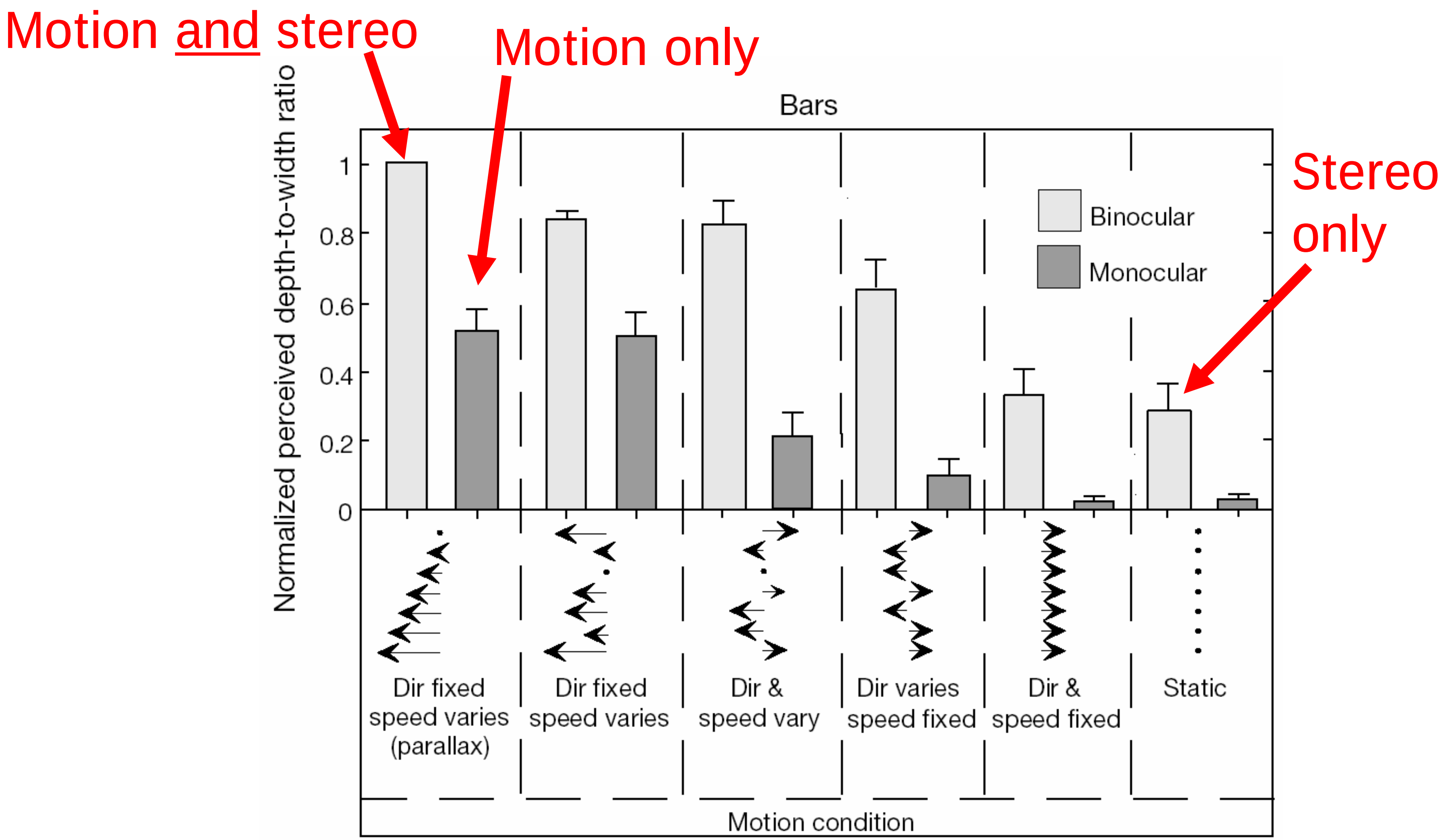


Why multidimensional selectivity?

- Measured in cortex
- Important perceptually
  - (van Ee and Anderson, *Nature*, 2001)



Improvement over isolated cues



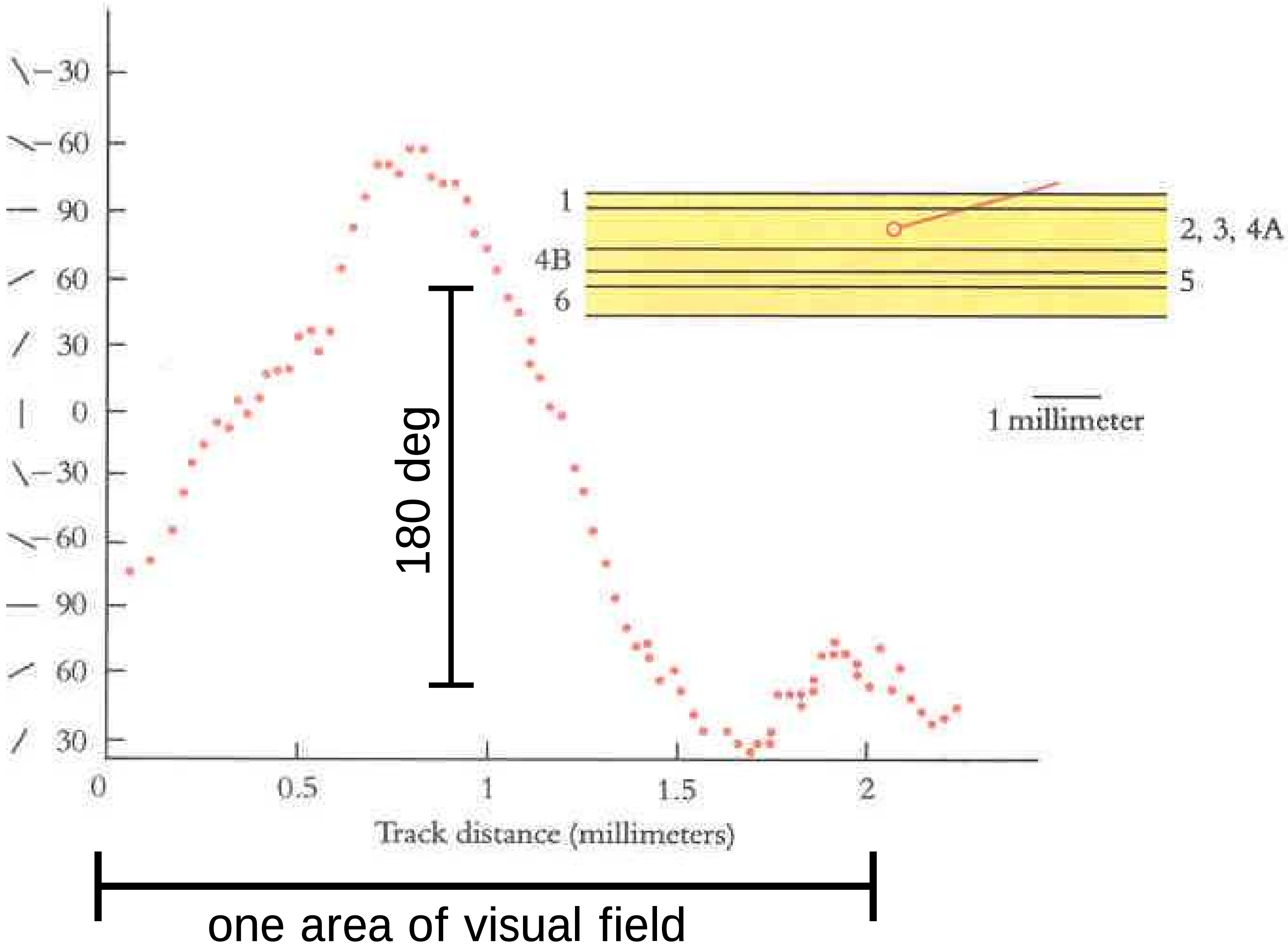
Conclusion

- V1 reformats the visual data so that it is easier to intepret
  - I/O ratio for retina: 100/1 (compression)
  - I/O ratio V1: input:output ratio ~ 1:50 (expansion!)
- A neuromorphic systems for visual perception should simultaneously integrate information from all cues (orientation, disparity, motion) at a very early stage.

Overview

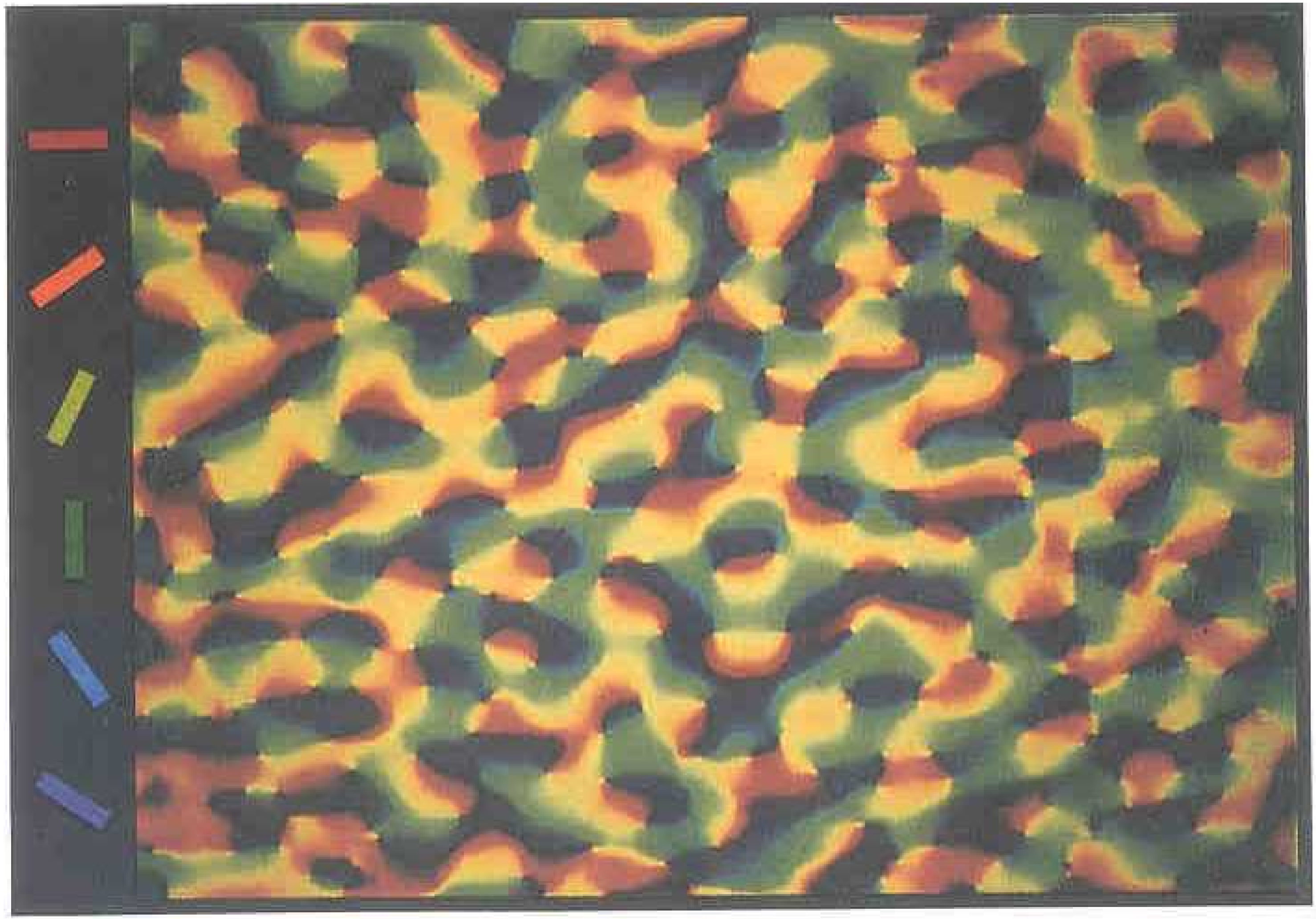
- Visual cortical processing
  - Multidimensional selectivity
- Orientation Selectivity
  - What is it?
  - How do we implement it?
    - AER Multi-chip Architecture
    - DSP/FPGA Architecture
- Joint Orientation/Disparity Selectivity
  - Disparity energy model: position versus phase shifts
  - Population responses versus tuning curves
  - Bigger is not necessarily better

Orientation Selectivity in V1



D. H. Hubel, *Eye, Brain and Vision*,
New York: Scientific American Library, 1995

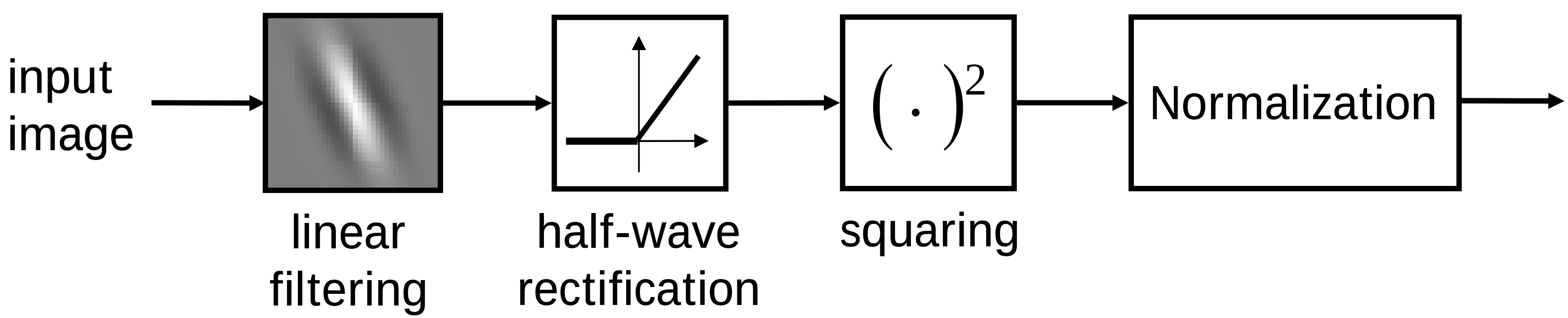
Orientation Map in V1



Model of Cortical Simple Cells

- Linear filtering to establish selectivity
- Half-wave rectifiction
- Squaring
- Normalization

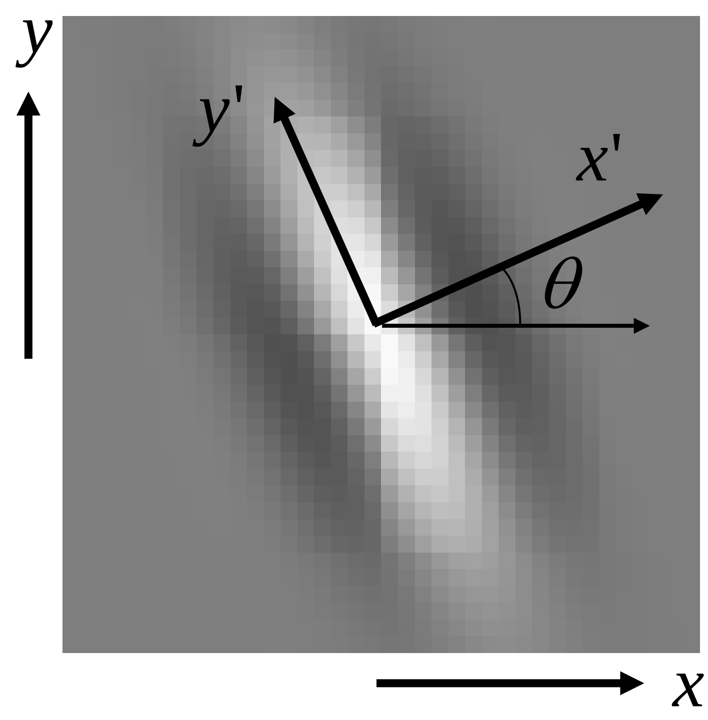
“half squaring”



Gabor Receptive Field (RF) Profiles

- Commonly used to model the 2D spatial receptive field profiles of orientation tuned cells.
- Shape determined by spatial frequency (Ω), width (σ_x, σ_y), orientation (θ) and phase (ϕ).
- Roughly speaking, describes the "best" stimulus for the neuron.

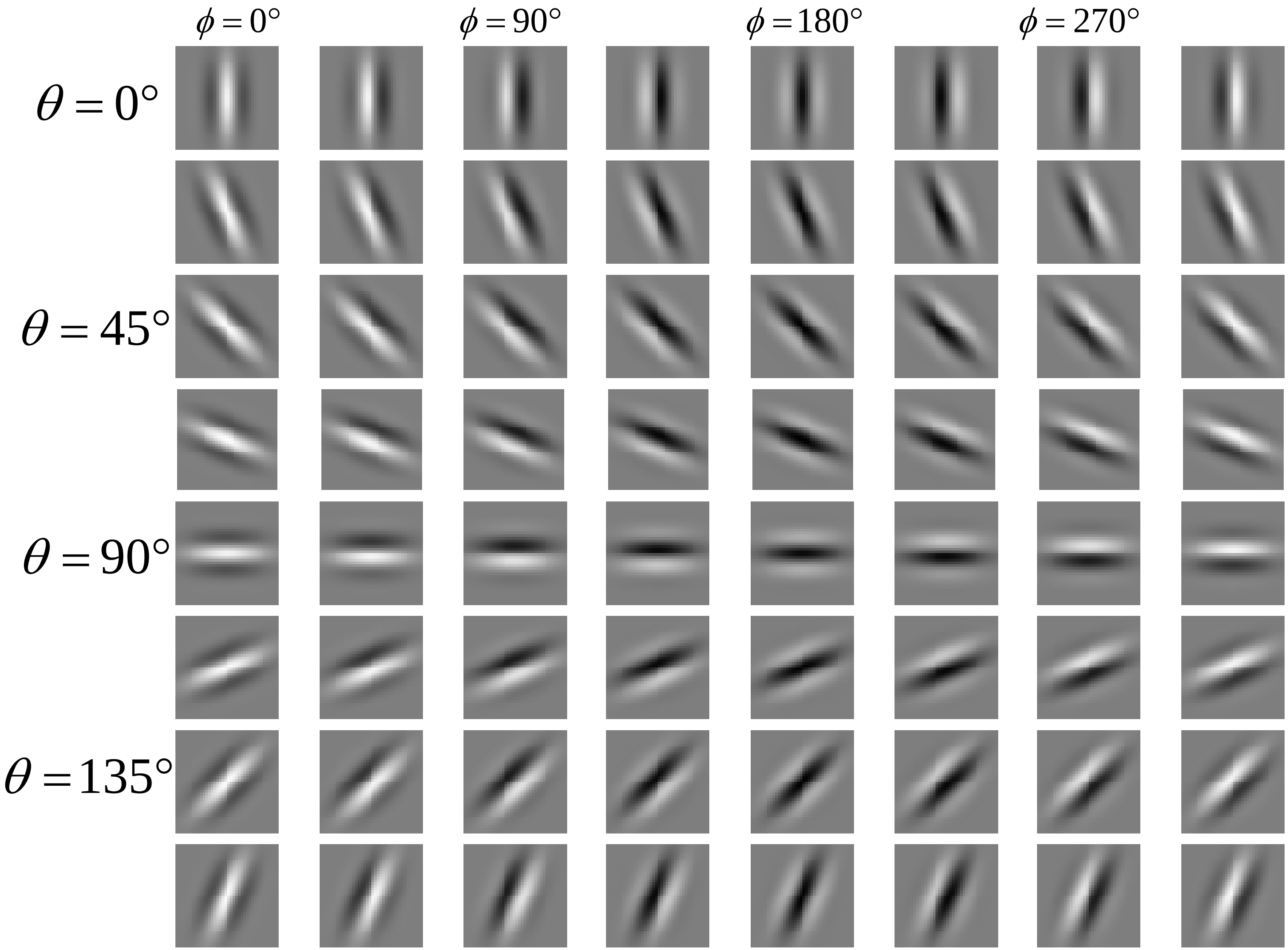
$$\phi = 0°, \quad \sigma_y = 2\sigma_x, \quad \sigma_x \Omega = 2$$



$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

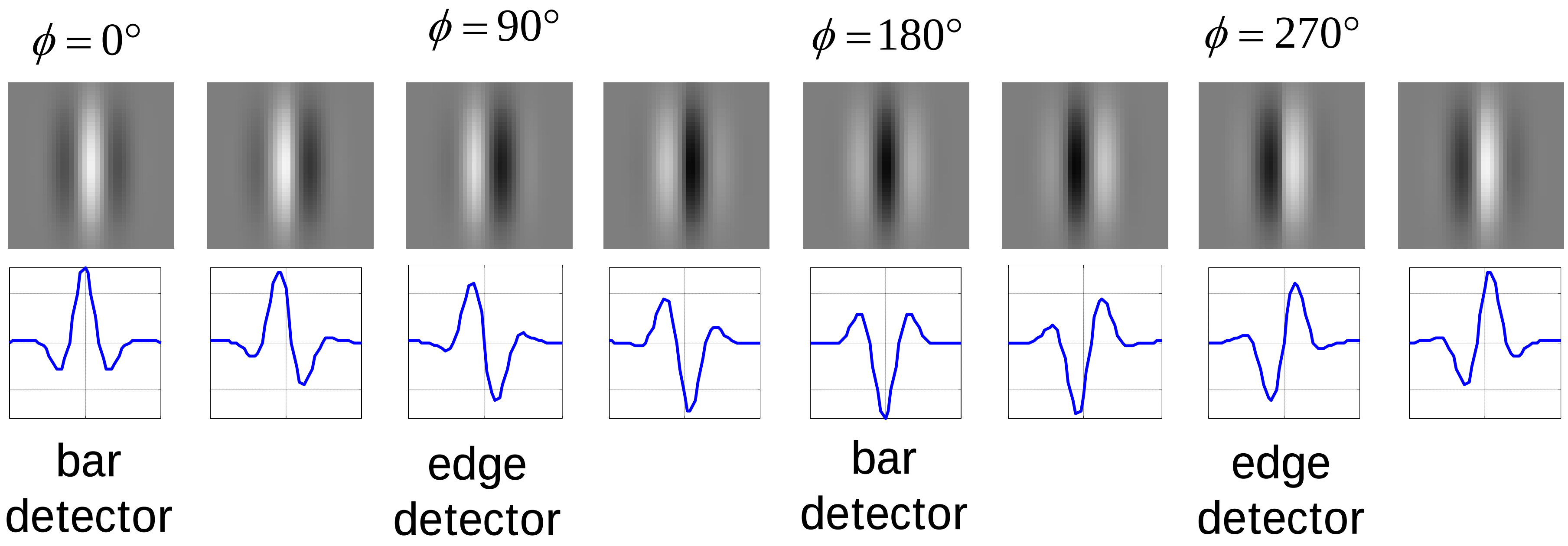
$$g(x,y) = \frac{1}{2\pi\sigma_x} \exp\left(-\frac{(x')^2}{2\sigma_x{}^2} \right) \cdot \frac{1}{2\pi\sigma_y} \exp\left(-\frac{(y')^2}{2\sigma_y{}^2} \right) \cdot \cos(\Omega x' + \phi)$$

Phase and Orientation Diversity



Phase Diversity

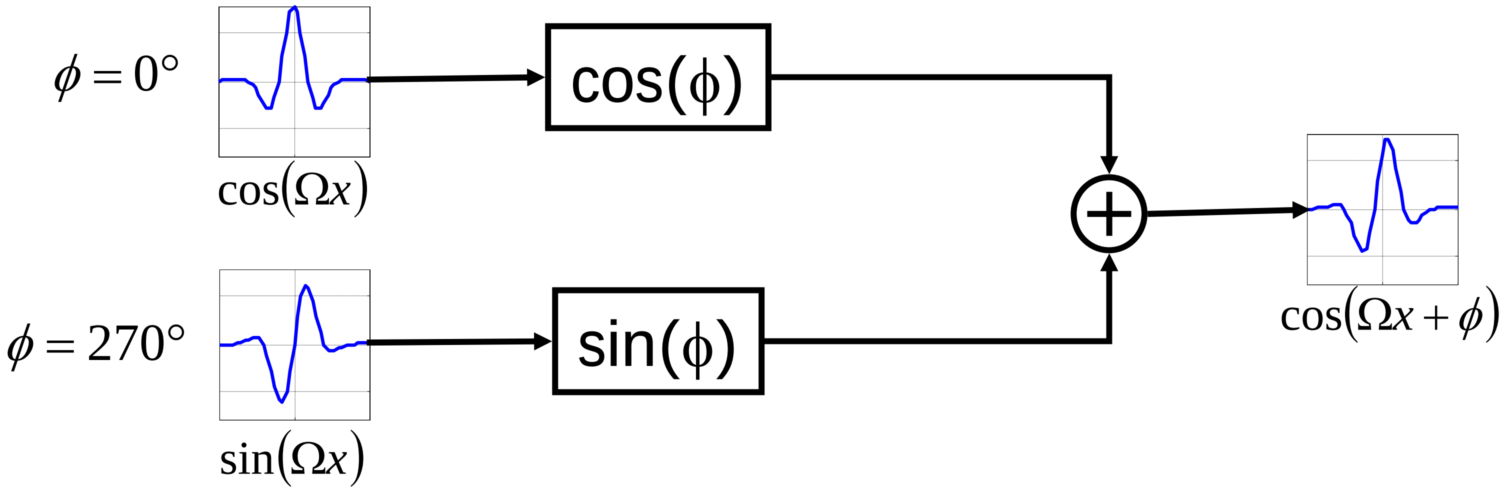
- The phase f controls the type of stimulus (edge or bar) that best excites the detector.
- Receptive fields with phases that differ 180° are tuned to stimuli with opposite polarity (e.g. dark bars versus light bars).



Two phases is enough

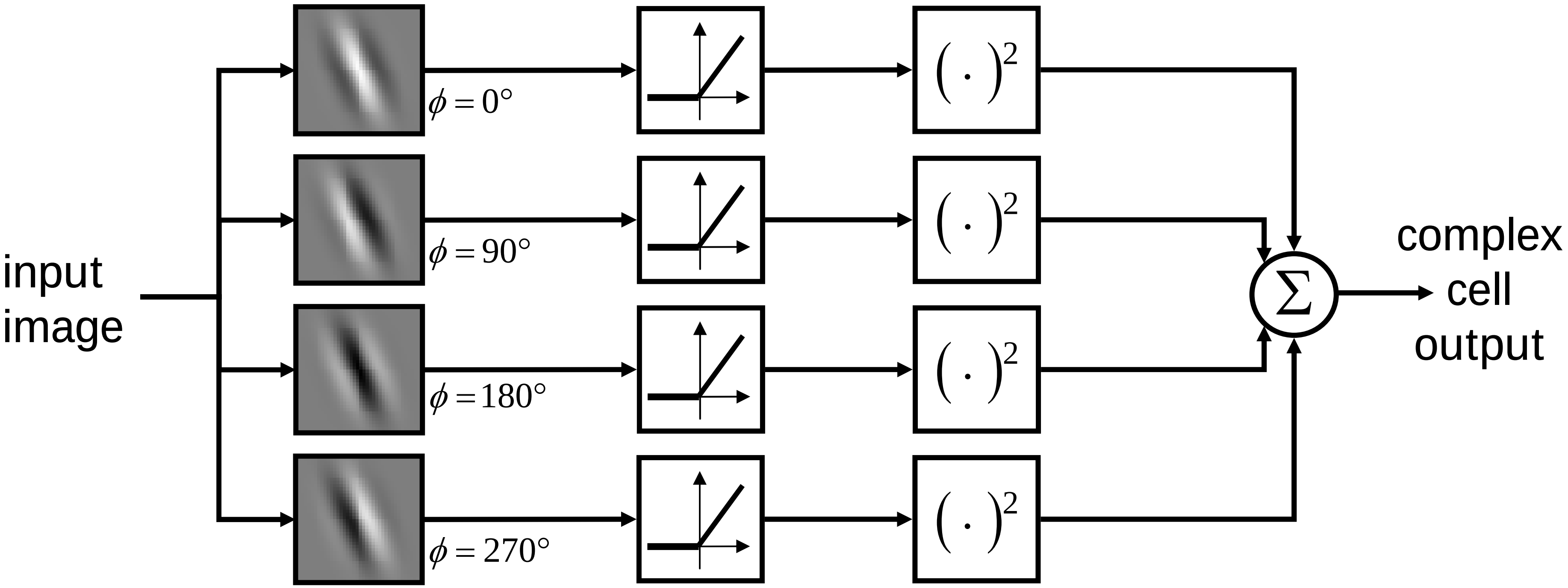
- Cortical neurons display the full range of phases, although there is some evidence that phases cluster around 0°, 90°, 180° and 270°.
- RFs with ϕ = 0° (180°) are referred to as *even*-symmetric.
- RFs with ϕ = 90° (270°) are referred to as *odd*-symmetric.
- Theoretically, we can compute the output of any phase RF given the just the outputs of 0° and 270°:

$$\cos(\Omega x + \phi) = \cos(\Omega x)\cos(\phi) + \sin(\Omega x)\sin(\phi)$$

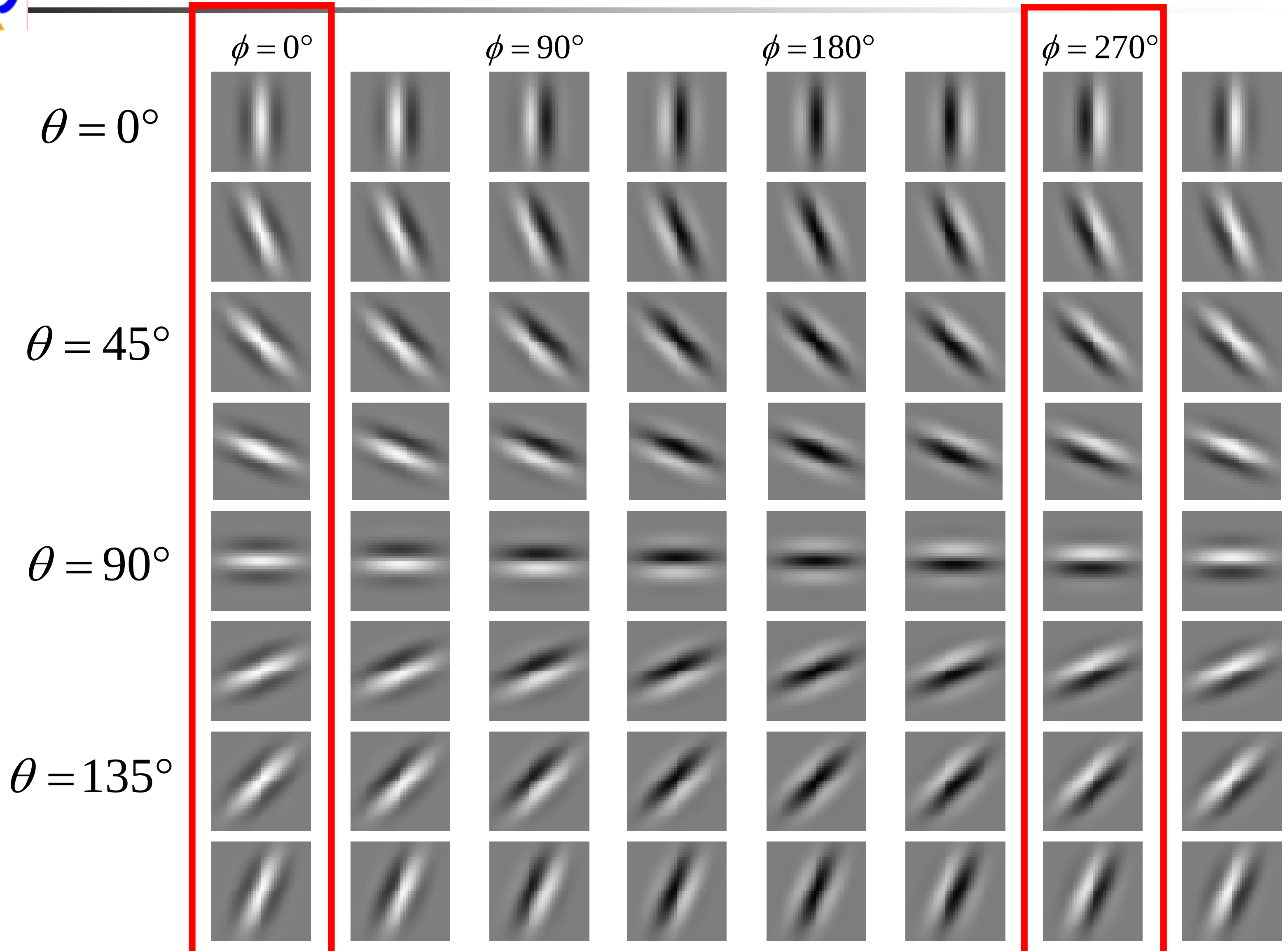


Orientation Energy

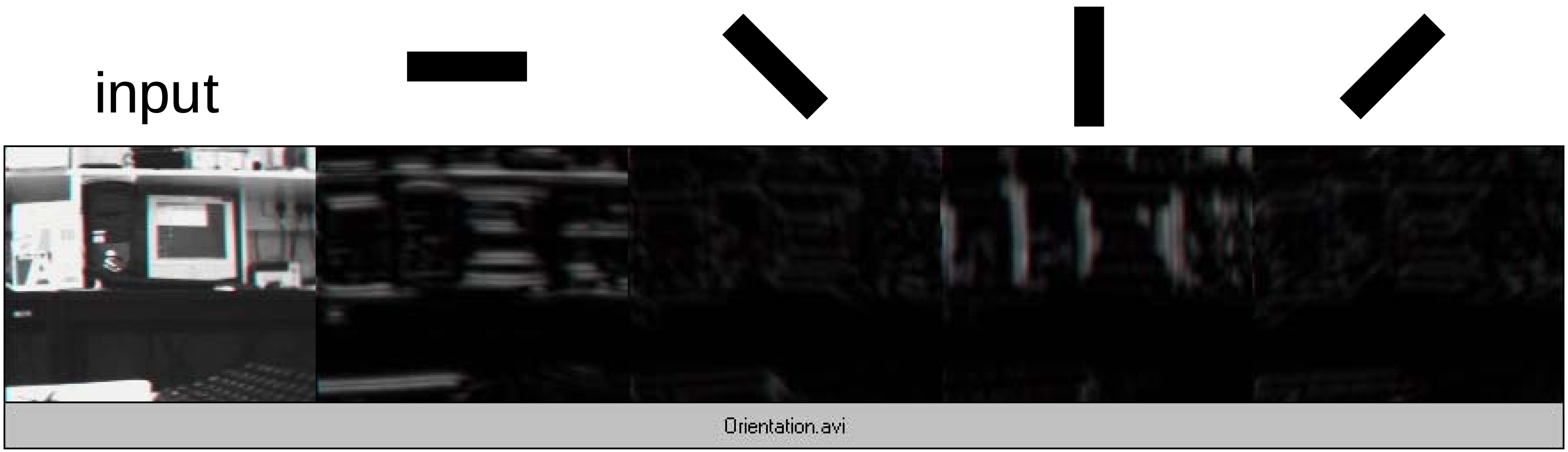
- V1 neurons are often differentiated as either
  - "simple": phase-dependent responses
  - "complex": phase-independent reponses
- Complex cells are often modelled as the sum of four simple cells with phase-quadrature RF profiles.



Phase and Orientation Diversity



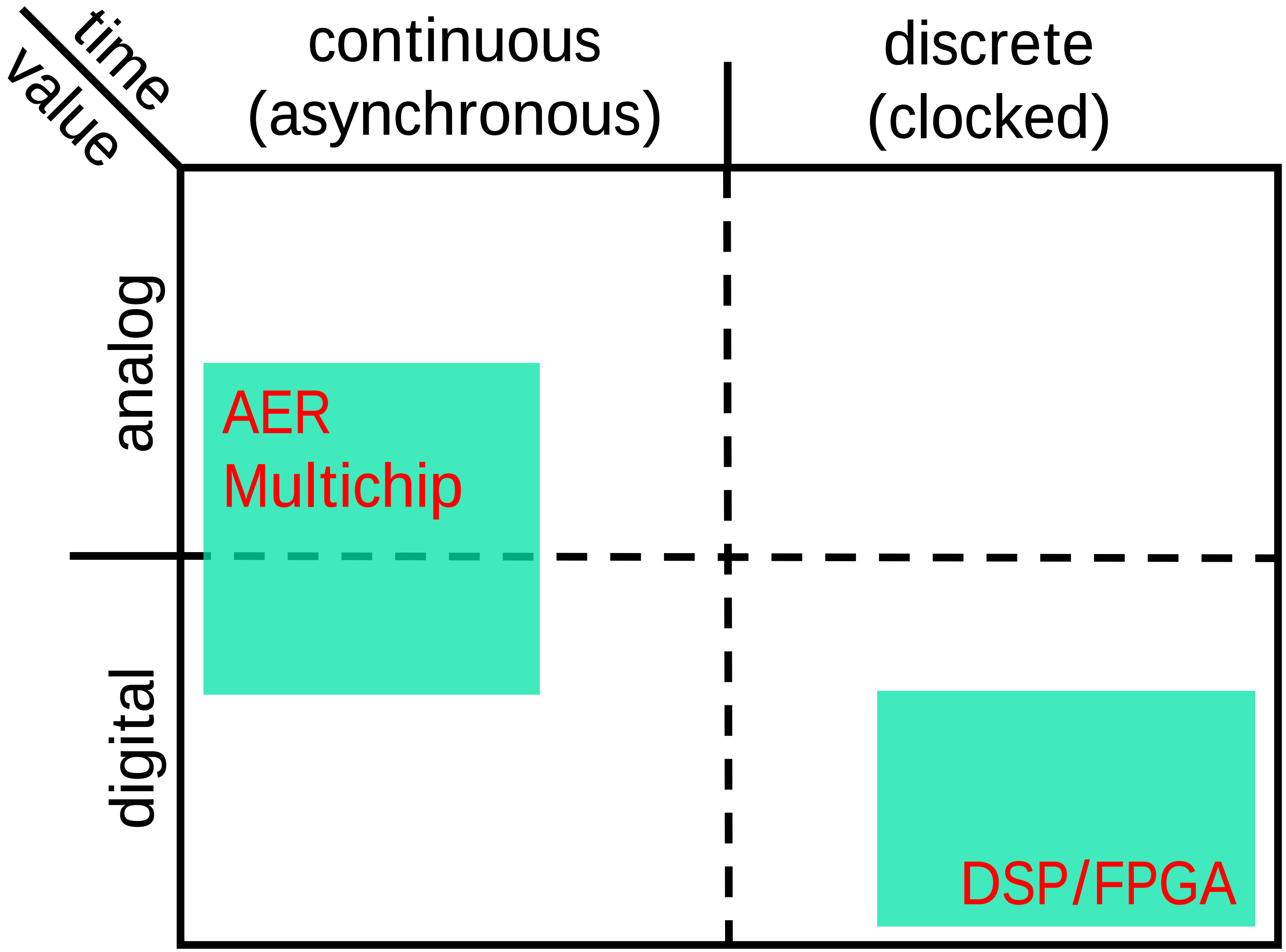
Video recap



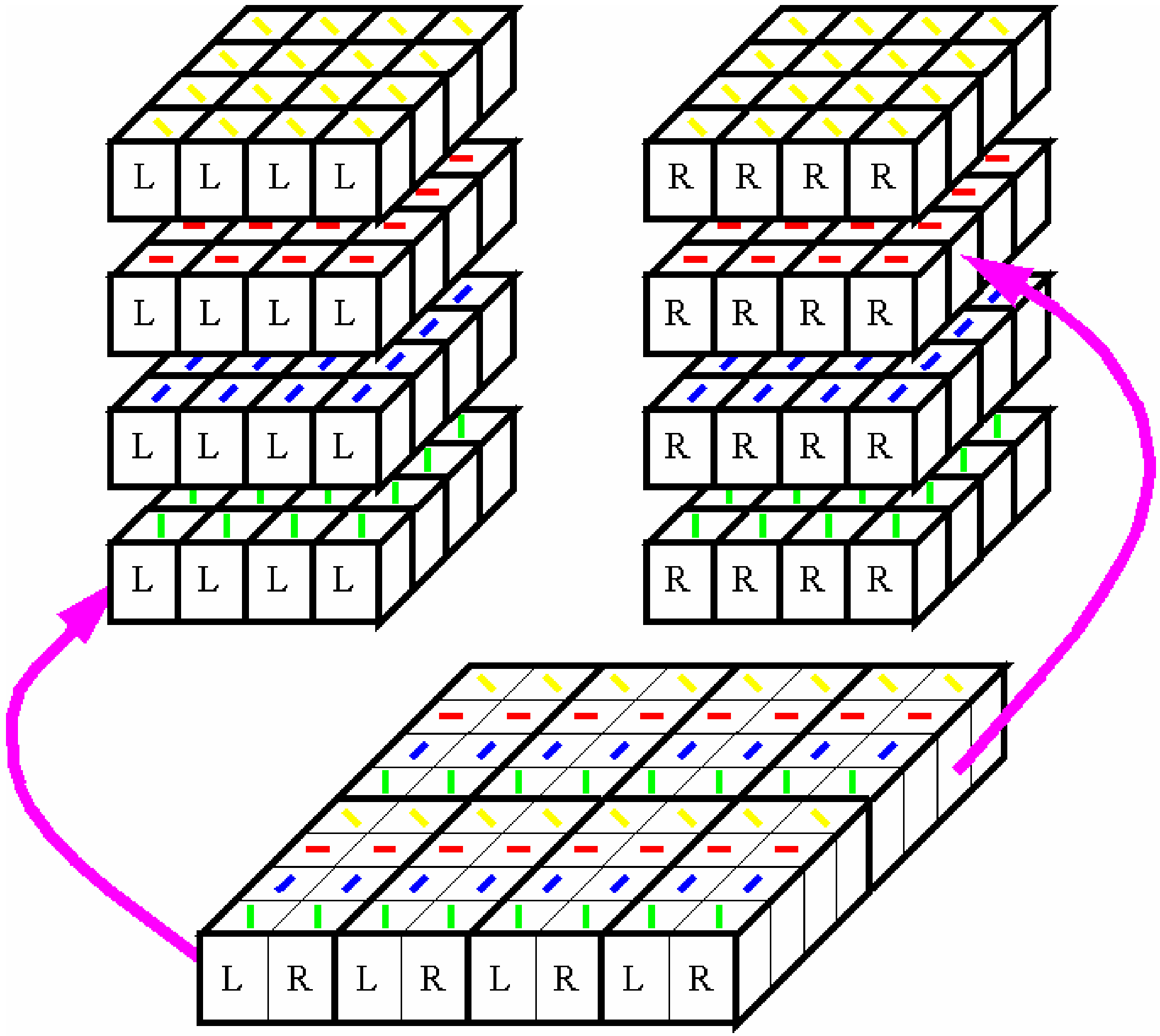
Overview

- Visual cortical processing
  - Multidimensional selectivity
- Orientation Selectivity
  - What is it?
  - How do we implement it?
    - AER Multi-chip Architecture
    - DSP/FPGA Architecture
- Joint Orientation/Disparity Selectivity
  - Disparity energy model: position versus phase shifts
  - Population responses versus tuning curves
  - Bigger is not necessarily better

Implementation: Design choices



Orientation Hypercolumns



Multi-chip Model

Ice Cube Model

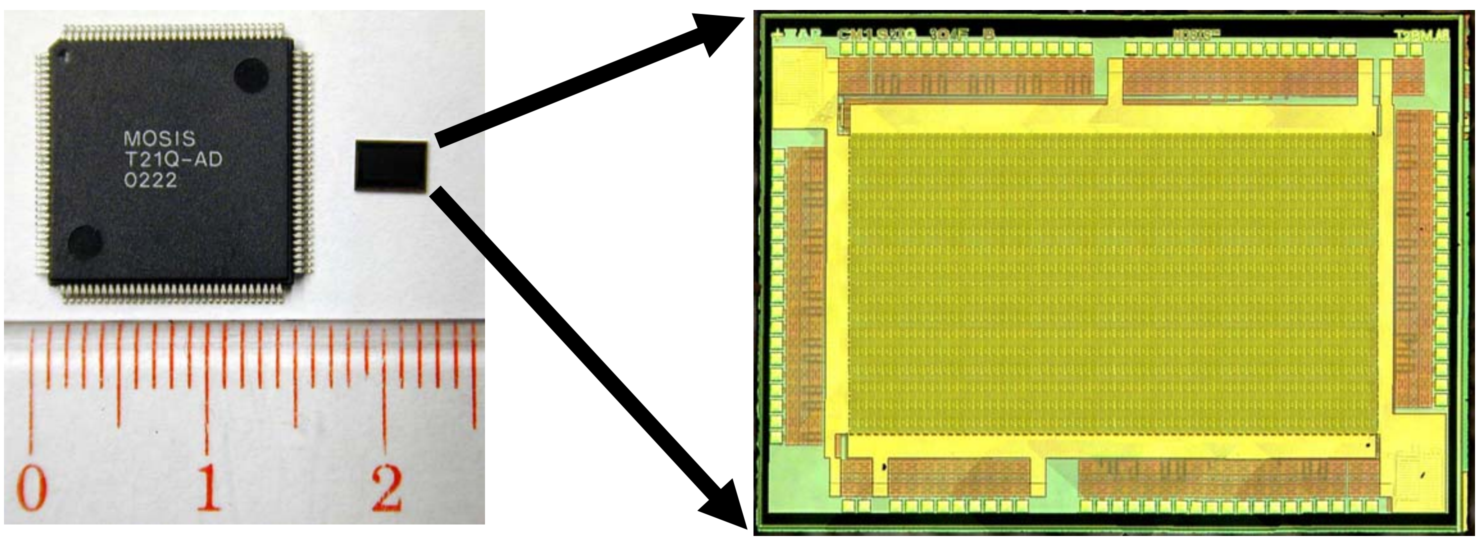
System Characteristics

- Retinotopic arrays of neurons with Gabor-type receptive fields
  - All neurons on one chip tuned to the same orientation and scale (electronically adjustable)
  - Phase quadrature receptive fields (EVEN and ODD)
  - Half wave rectification
- Continuous time operation
- Neurons on different chips communicate with spikes (AER)

Chip Data

- Neurons: 32x64x4 neurons (8000)
- Technology: TSMC 0.25um
- Die size: 3.84mm x 2.54mm (9.8mm^2)
- Power dissipation: 3mW

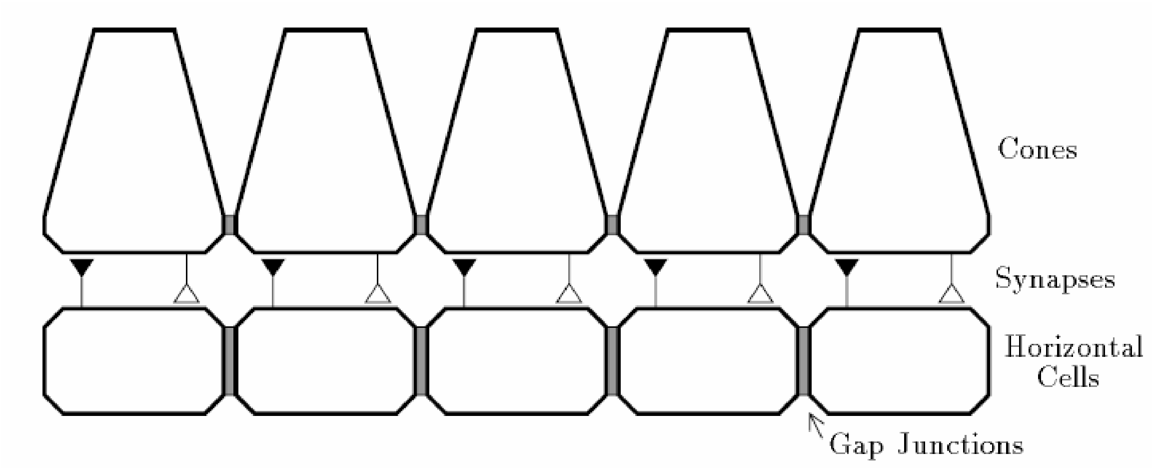
(cm)

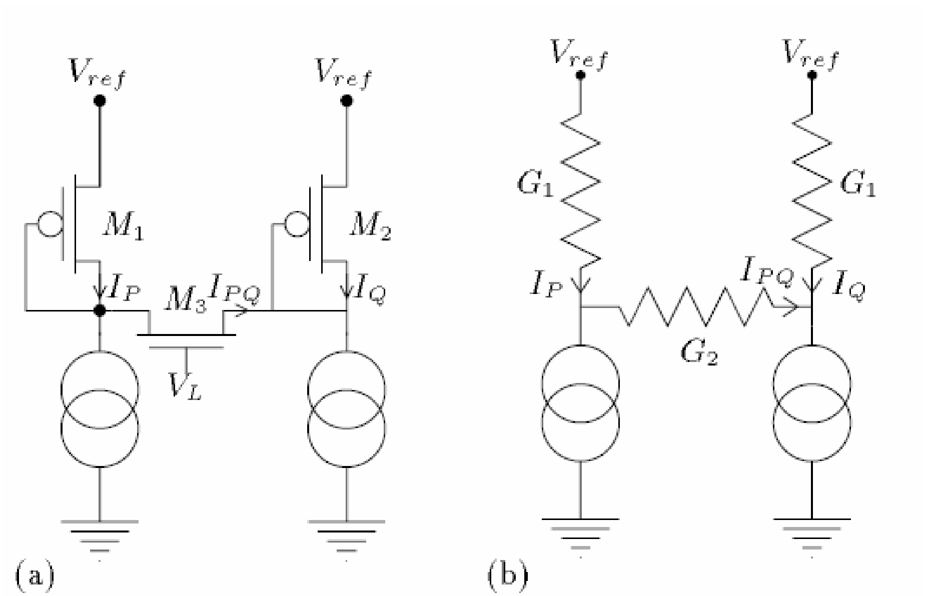


Architectural Advantages

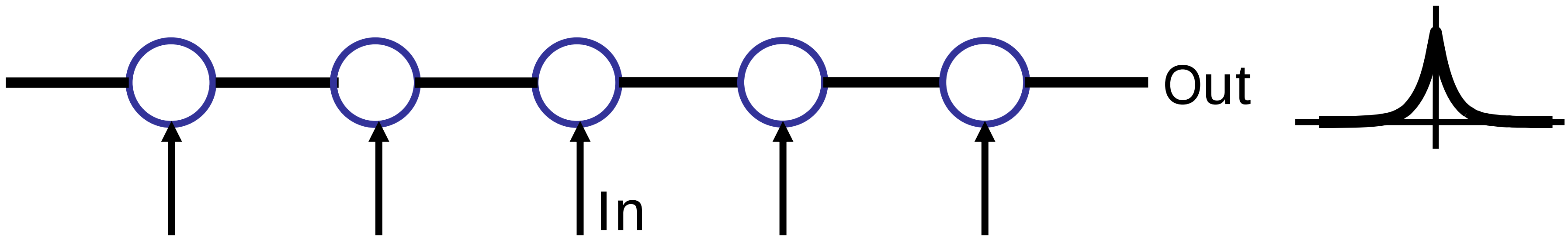
- Since neurons have identical response properties, retinotopic arrays are constructed by *tiling identical circuit blocks*.
- Neurons are only locally interconnected, which *simplifies wiring*.
- ON-OFF and spike based representation, *lowers power consumption and fixed pattern noise.*
- Continuous time operation enables *feedback interactions between maps.*

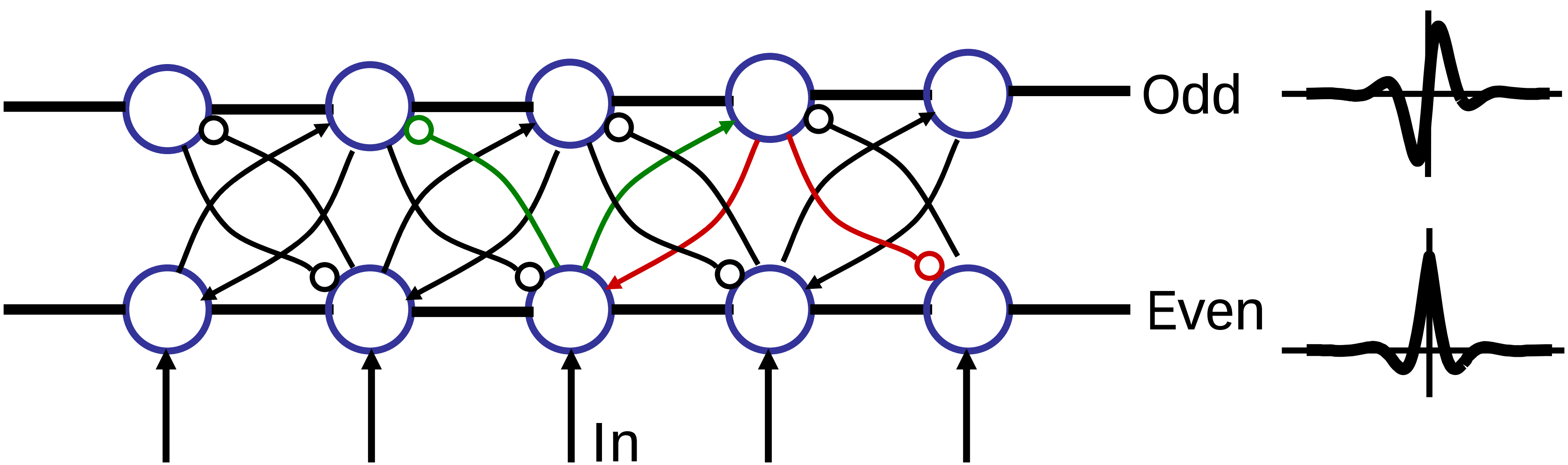
Neuromorphic analogue of gap junctions





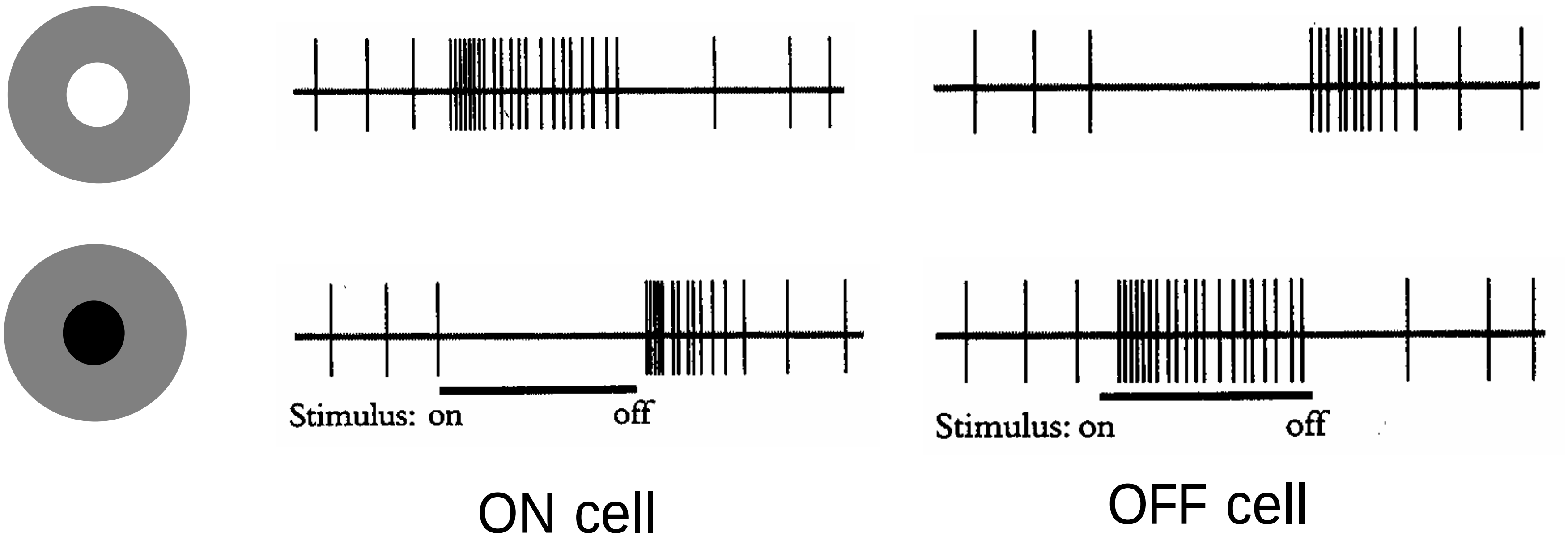
From one layer to two



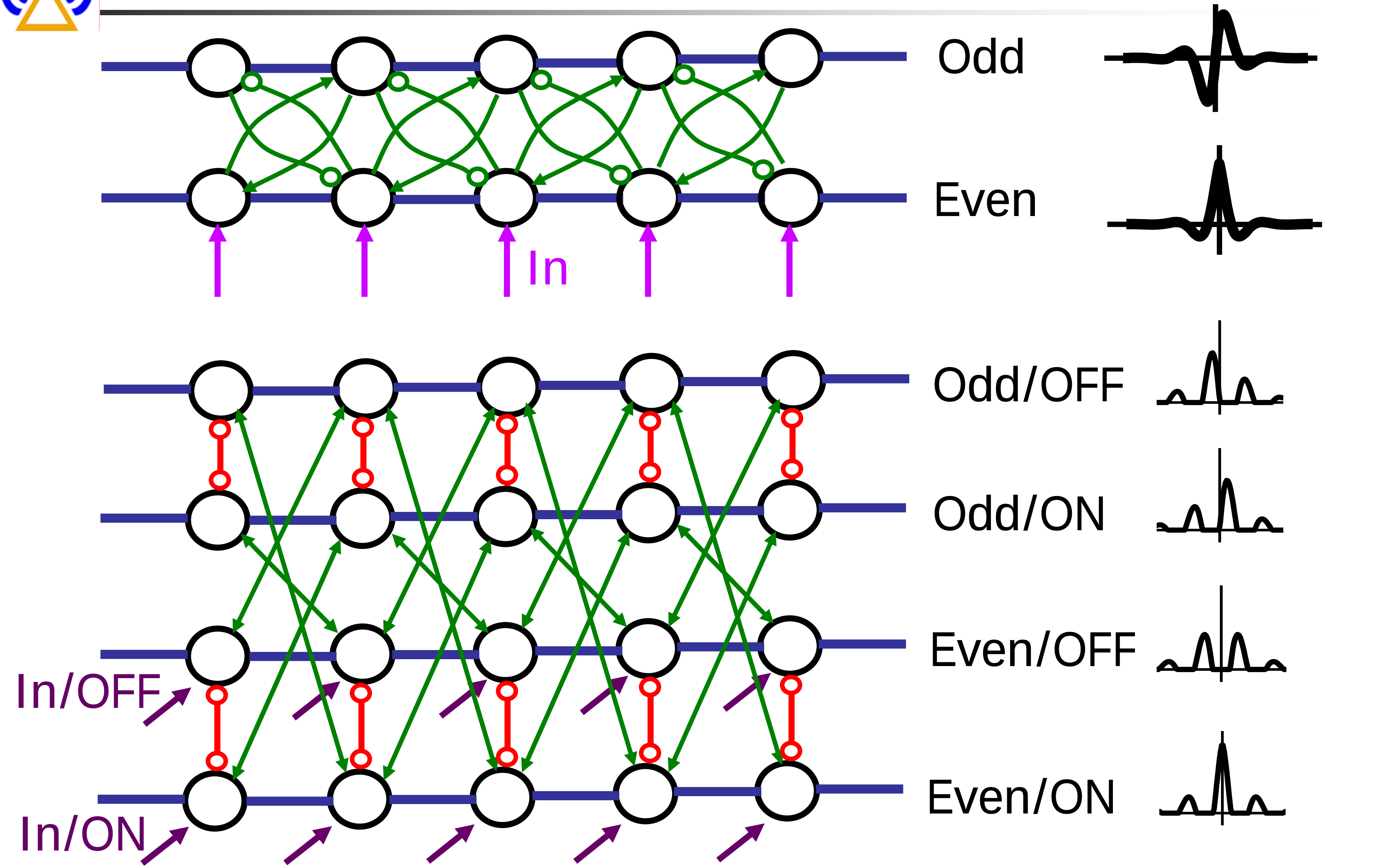


ON/OFF Signal Representation

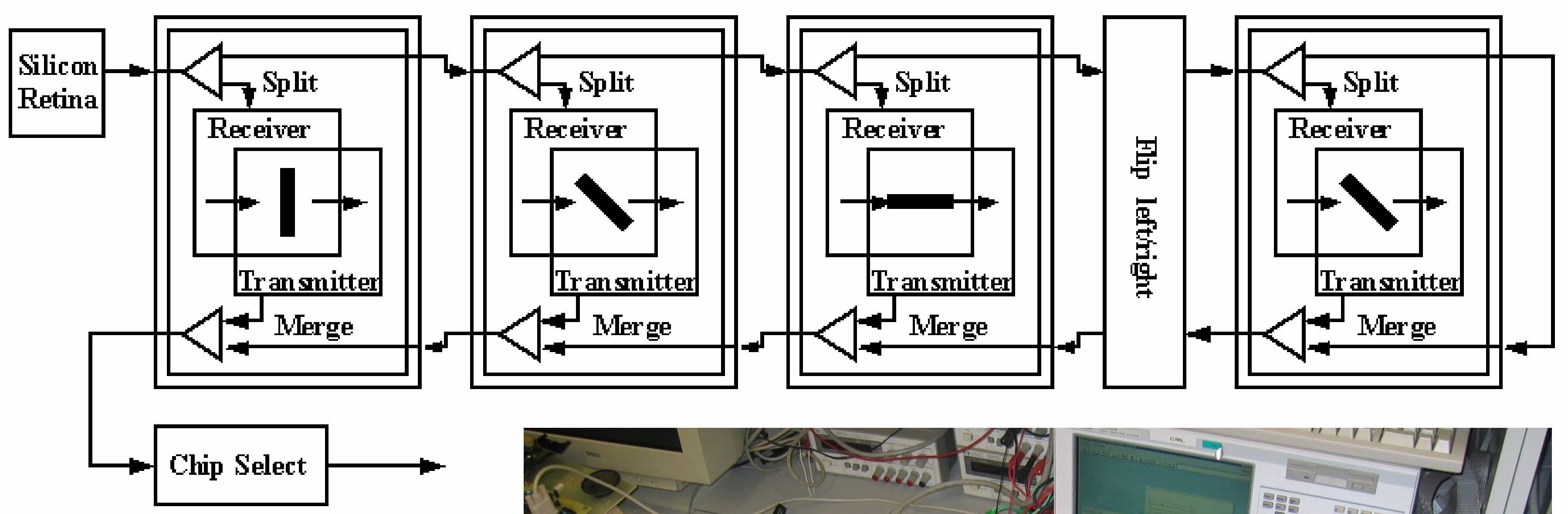
- Complementary channels encode positive and negative components of a signal
- Conserves metabolic resources by mapping background signals to near zero spike rates

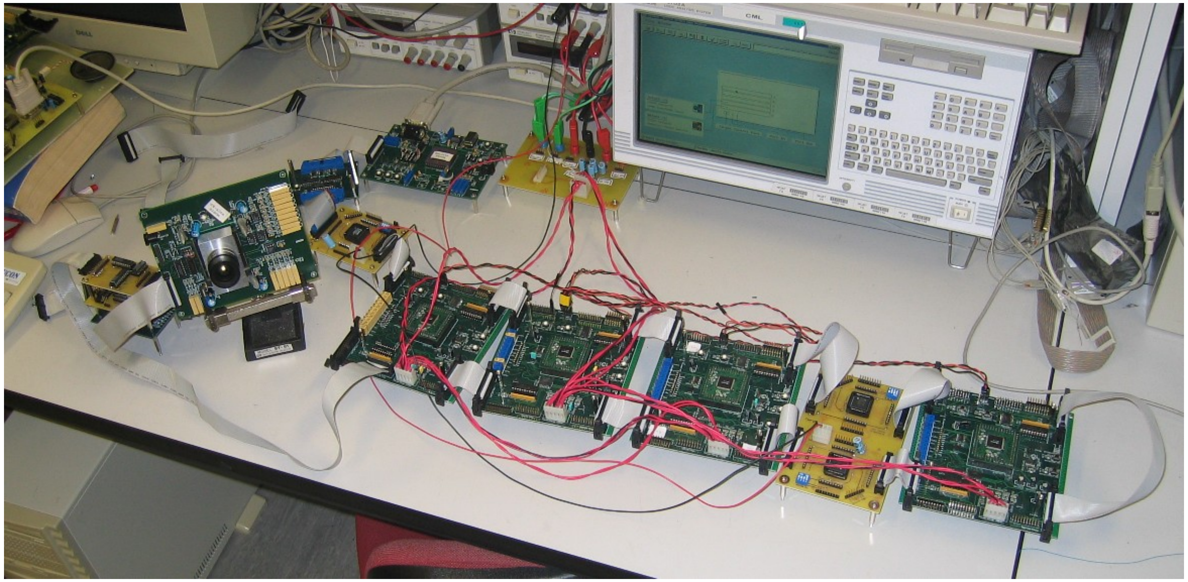


From two layers to four



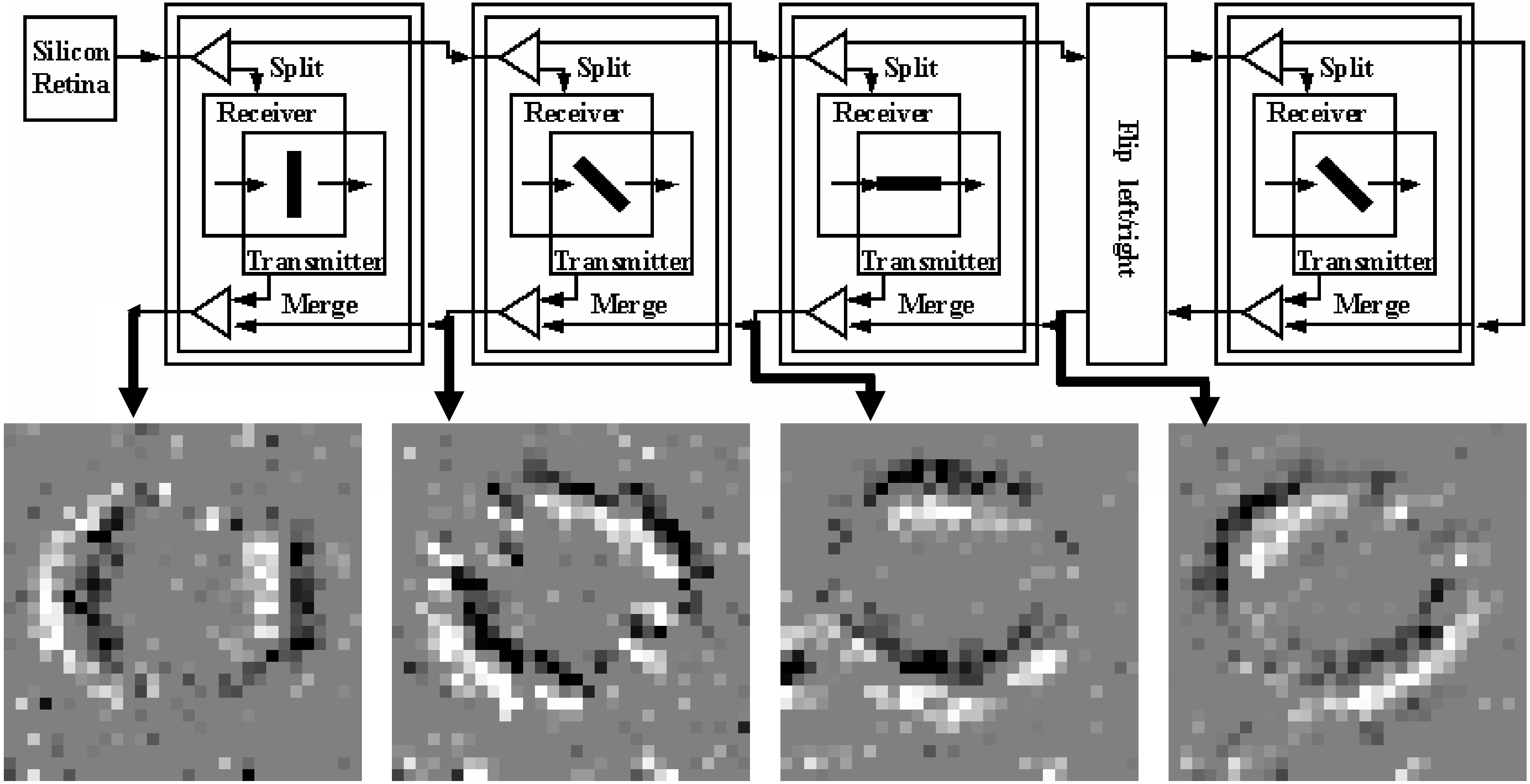
Feedforward System





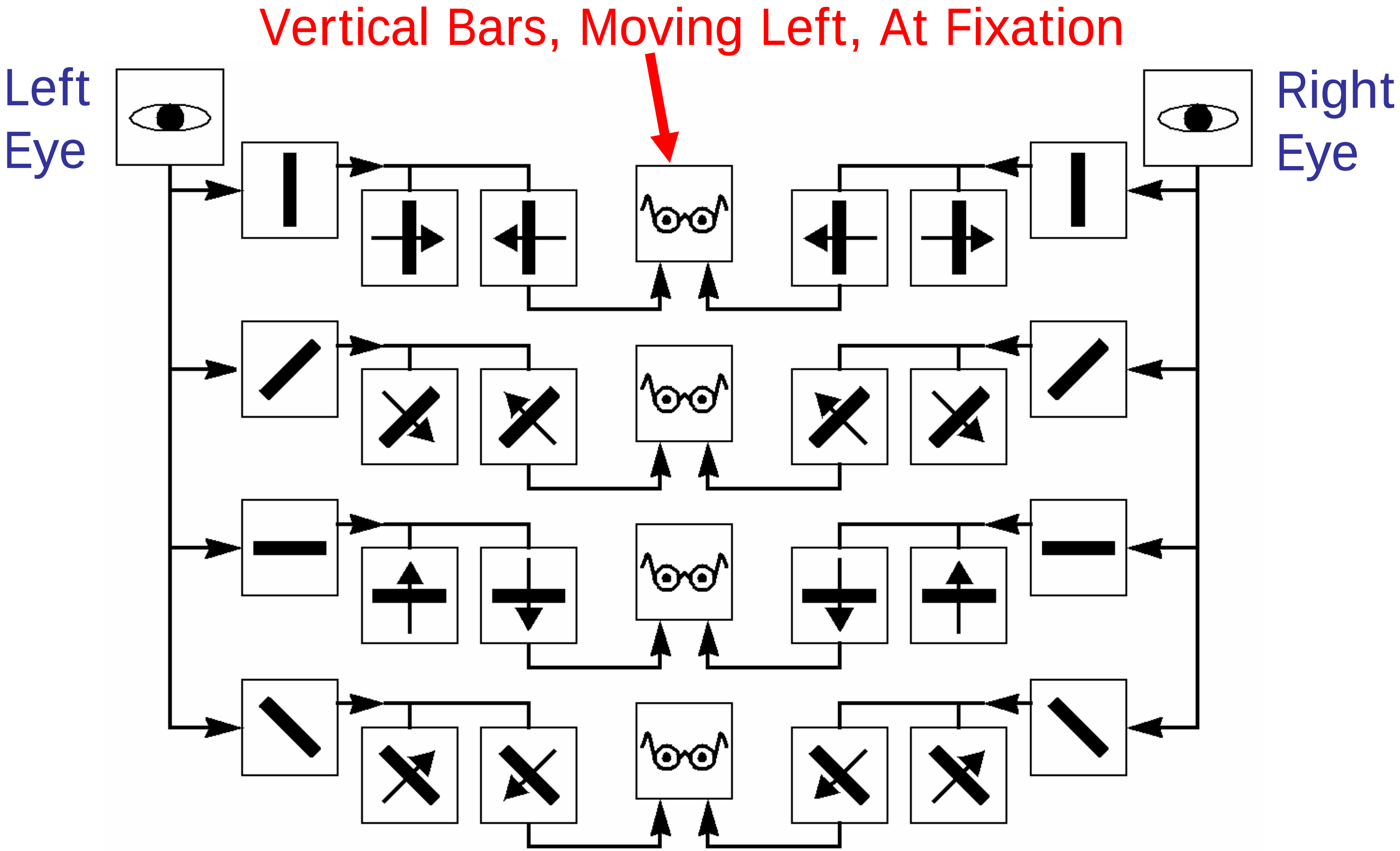
30,000 spiking neurons!

Odd Response to Ring

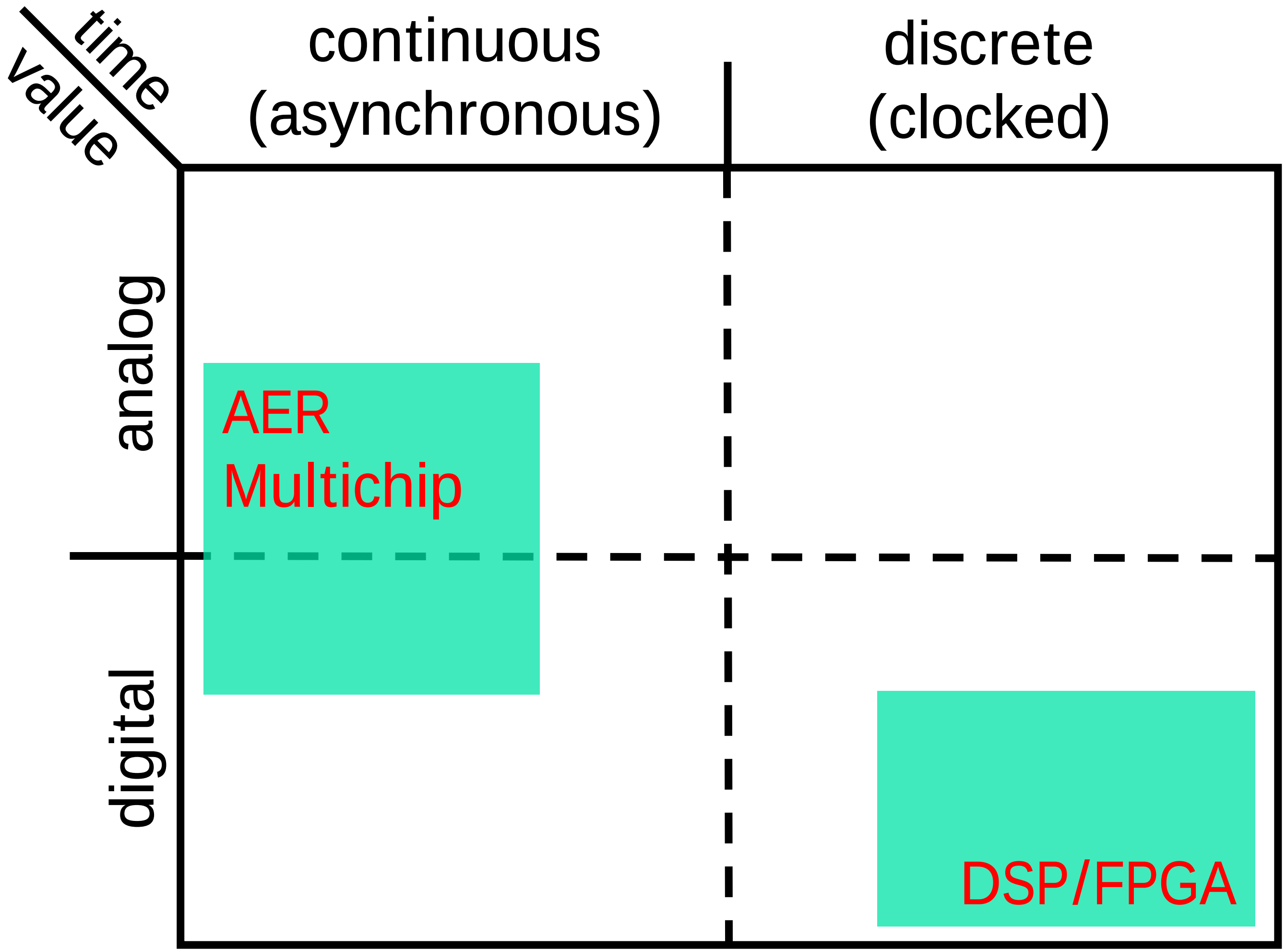


white = ON spikes black = OFF spikes

The eventual system



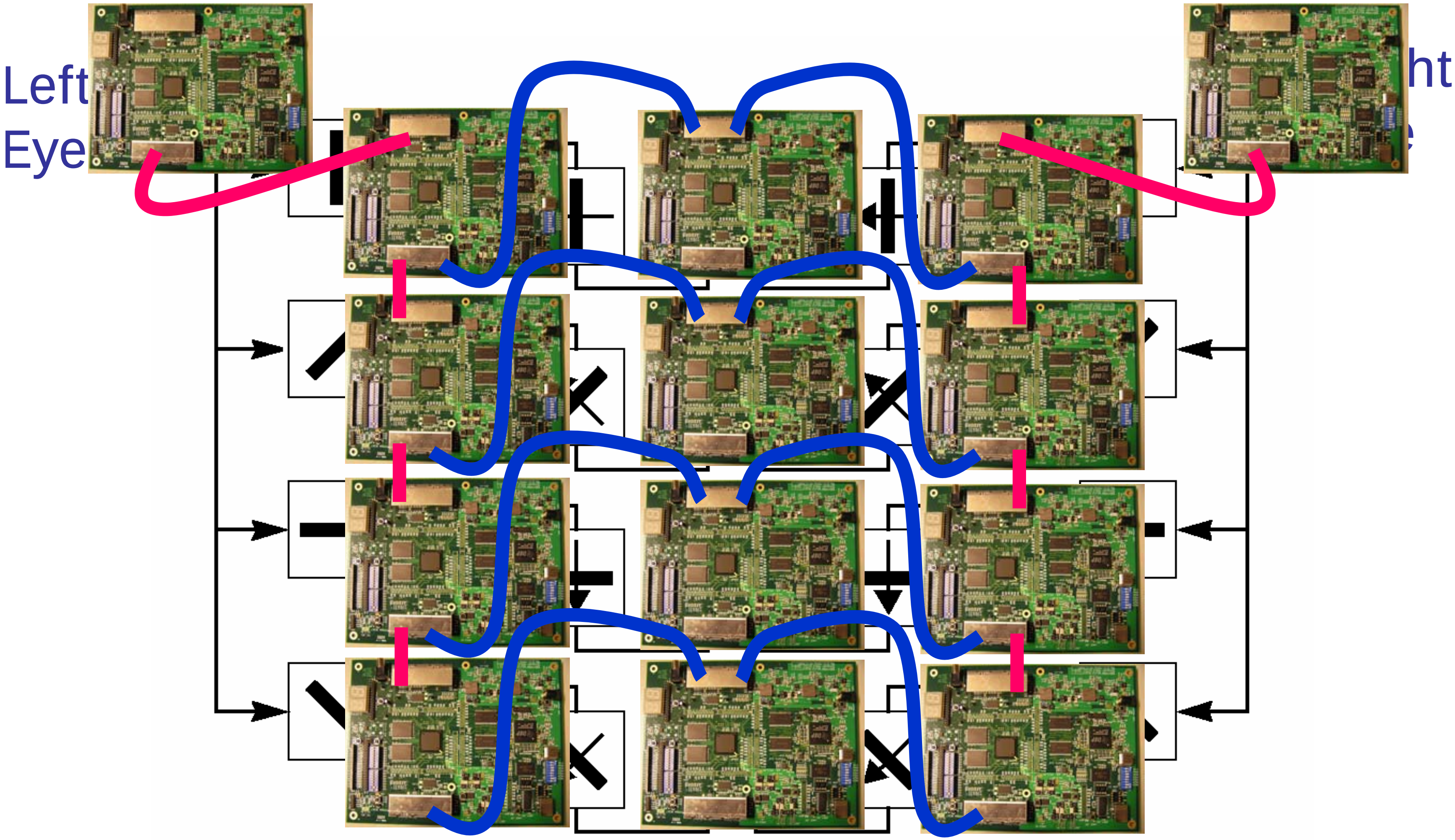
Implementation: Design choices



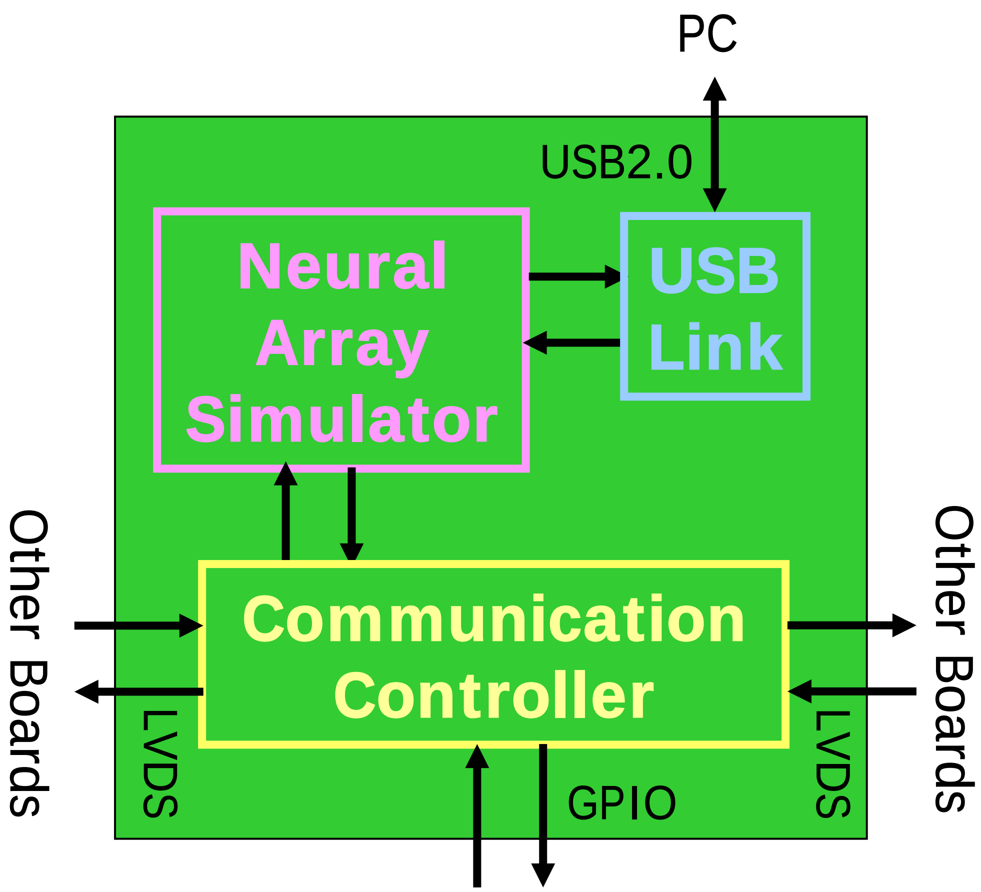
Goal

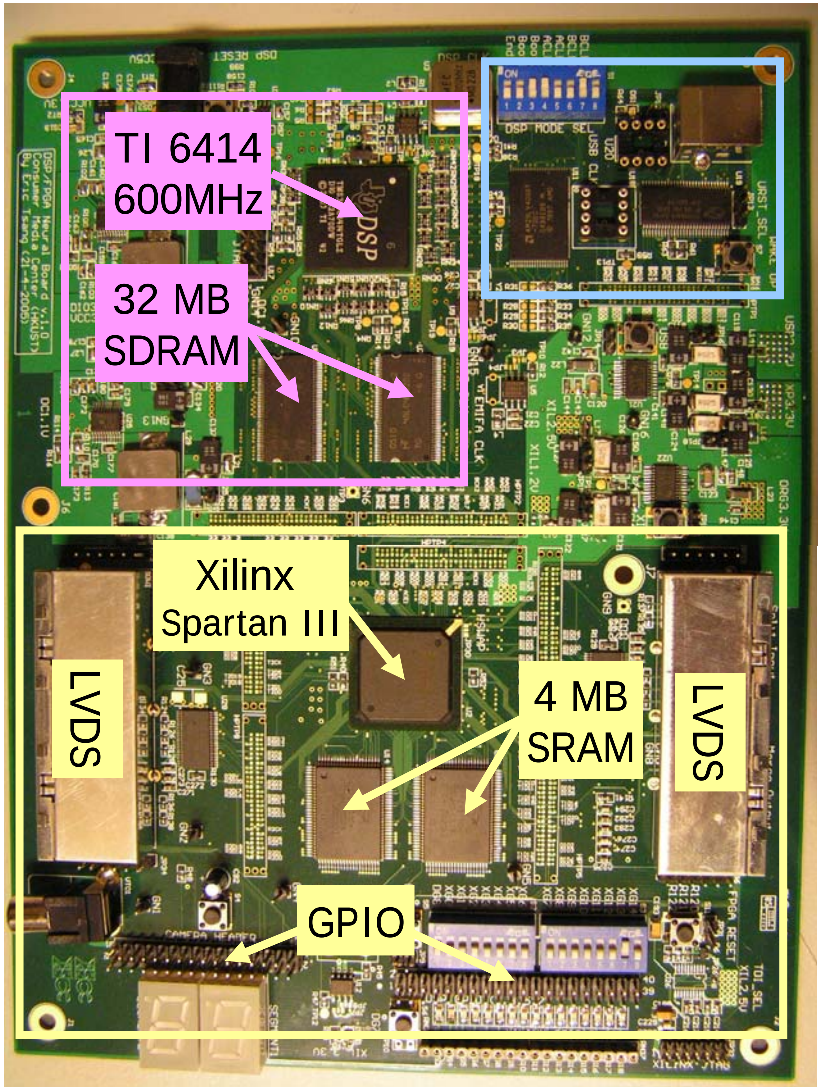
- **Rapidly reconfigurable** system for computing and combining outputs of many cortically inspired maps.
- Based on an expandable system architecture **that can be translated eventually to multi-chip AER neuromorphic systems**.
- Operates fast enough to support behavioral interaction with the environment.

System Architecture

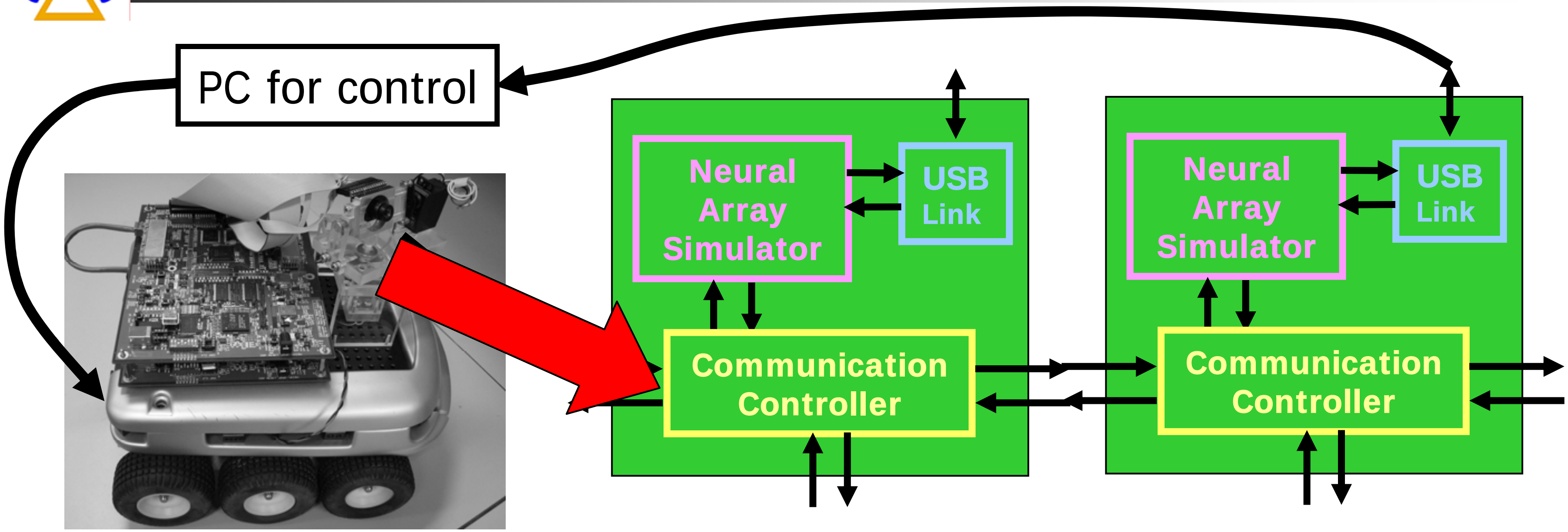


Board Architecture



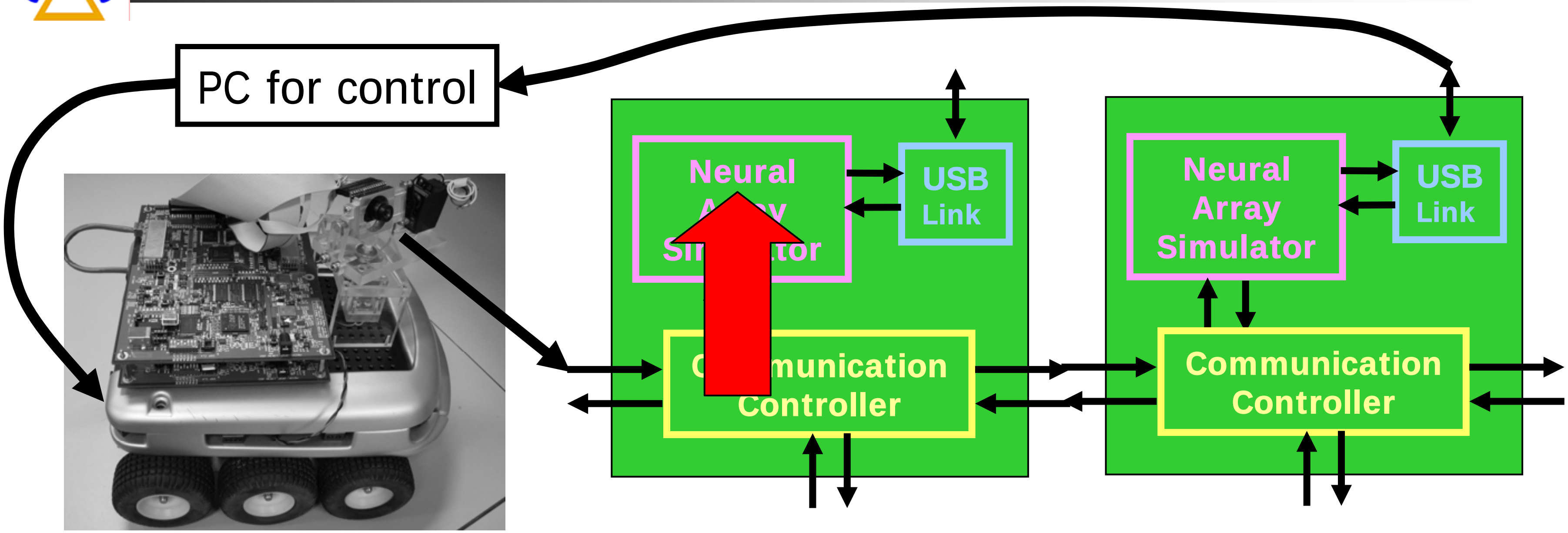


Telluride Project (Vision Systems)



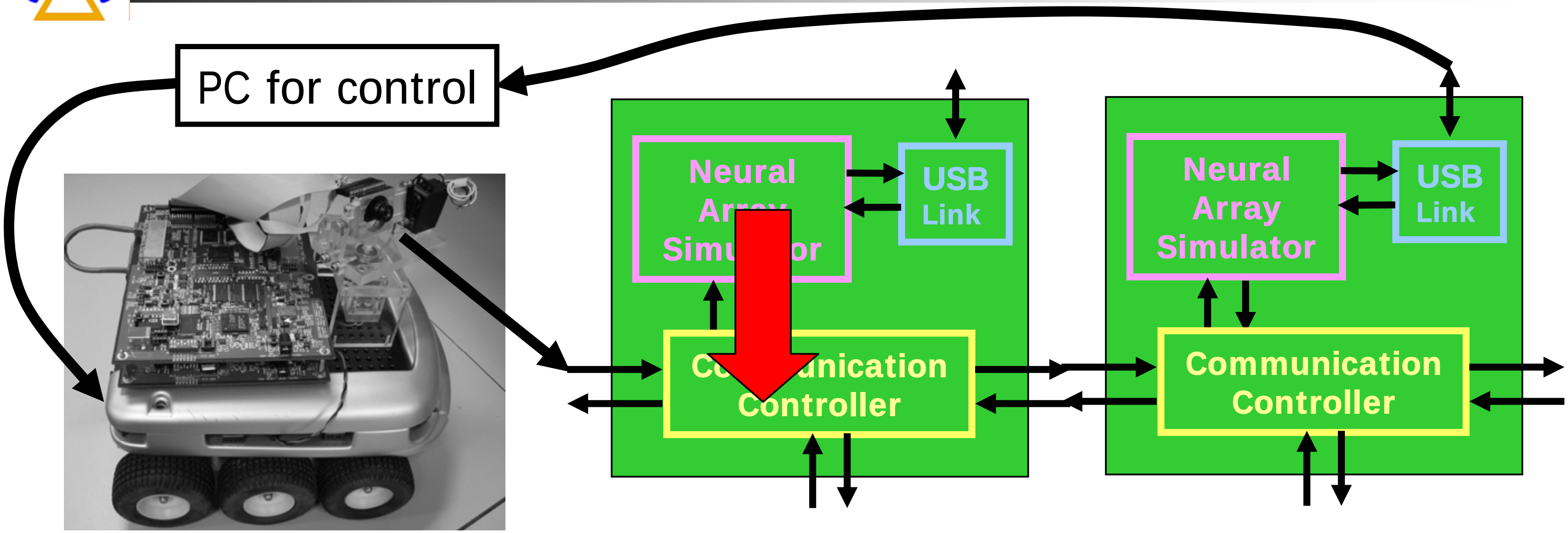
- First Board
  - Acquires image from camera at ~30fps
  - Computes Feature Maps
  - Converts to Spikes
  - Transmits to next board
- Second Board
  - Receives Spikes
  - Computes Additional Feature Maps
  - Transmits data to PC which integrates sensor information to generate behaviour

Telluride Project (Vision Systems)



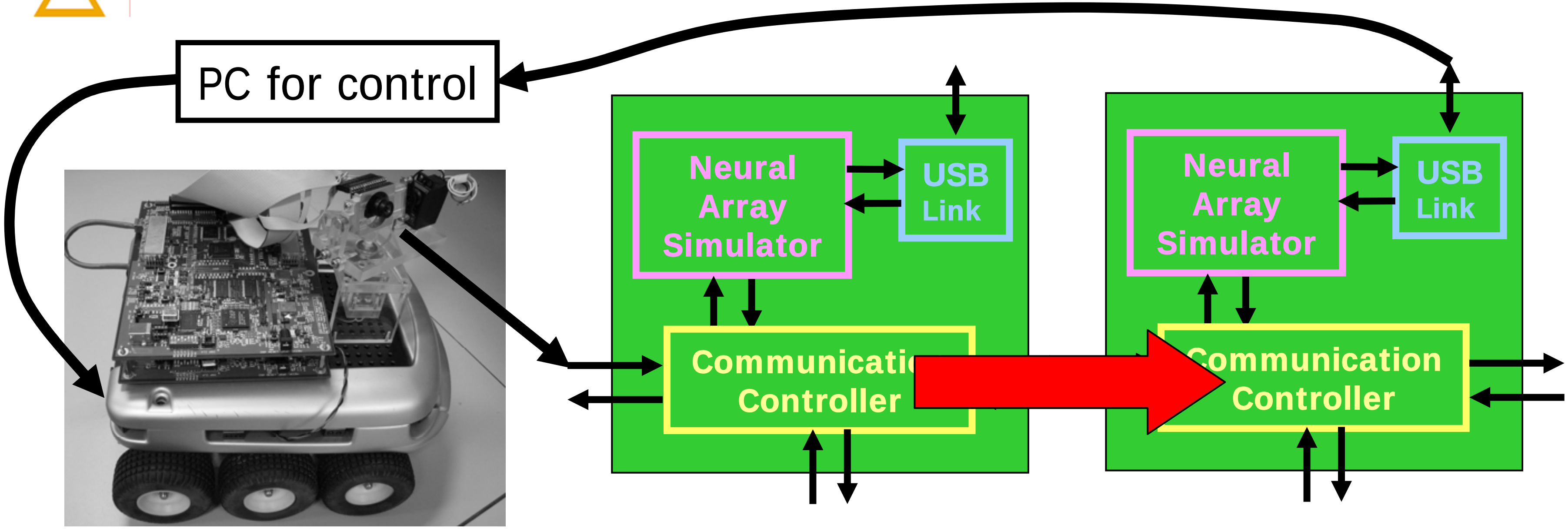
- First Board
  - Acquires image from camera at ~30fps
  - Computes Feature Maps
  - Converts to Spikes
  - Transmits to next board
- Second Board
  - Receives Spikes
  - Computes Additional Feature Maps
  - Transmits data to PC which integrates sensor information to generate behaviour

Telluride Project (Vision Systems)



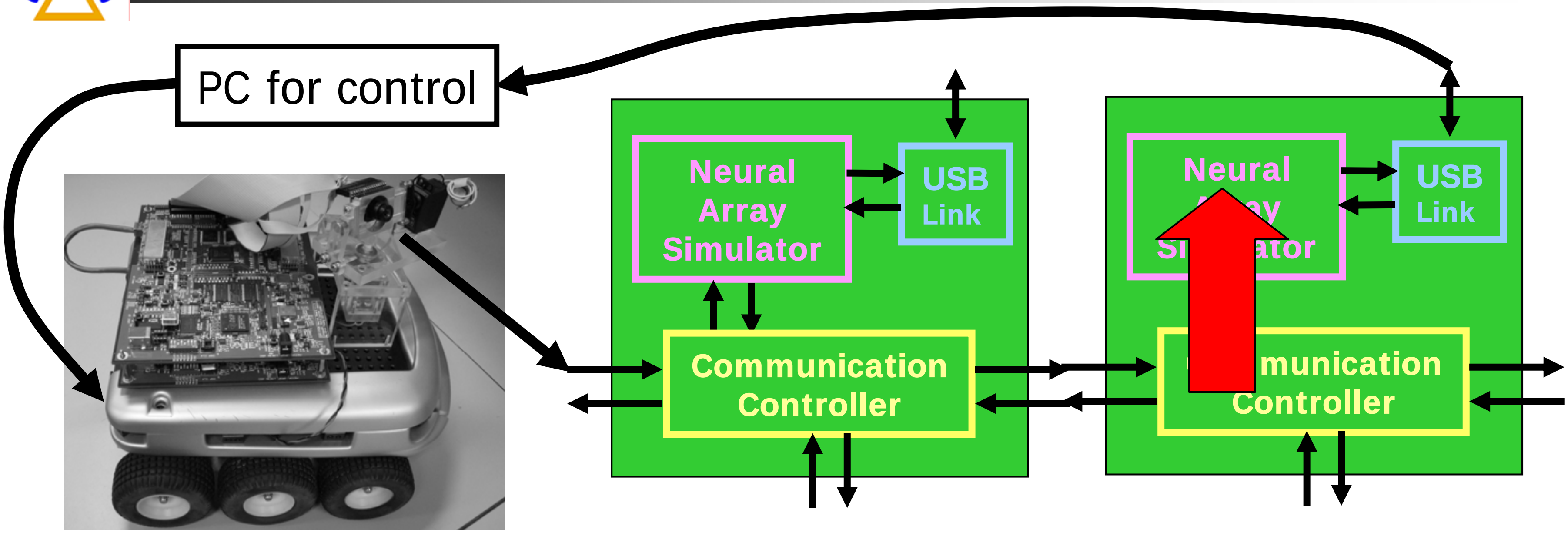
- First Board
  - Acquires image from camera at ~30fps
  - Computes Feature Maps
  - Converts to Spikes
  - Transmits to next board
- Second Board
  - Receives Spikes
  - Computes Additional Feature Maps
  - Transmits data to PC which integrates sensor information to generate behaviour

Telluride Project (Vision Systems)



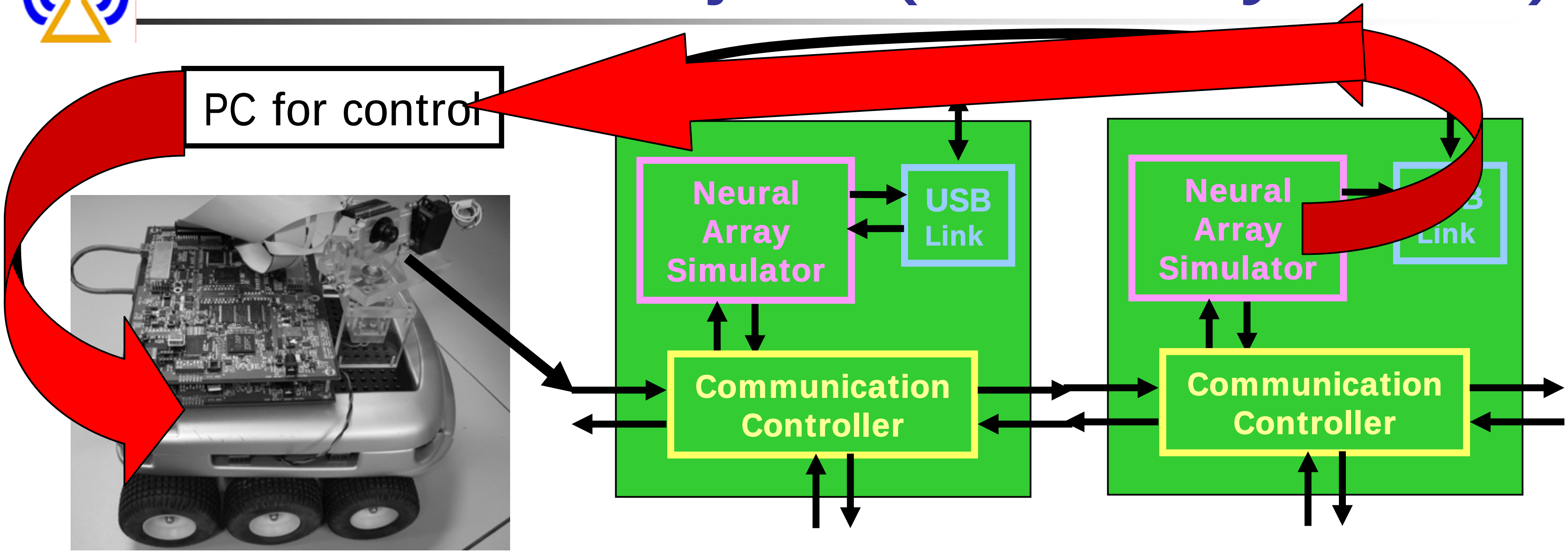
- First Board
  - Acquires image from camera at ~30fps
  - Computes Feature Maps
  - Converts to Spikes
  - Transmits to next board
- Second Board
  - Receives Spikes
  - Computes Additional Feature Maps
  - Transmits data to PC, which integrates sensor information to generate behaviour

Telluride Project (Vision Systems)



- First Board
  - Acquires image from camera at ~30fps
  - Computes Feature Maps
  - Converts to Spikes
  - Transmits to next board
- Second Board
  - Receives Spikes
  - Computes Additional Feature Maps
  - Transmits data to PC which integrates sensor information to generate behaviour

Telluride Project (Vision Systems)



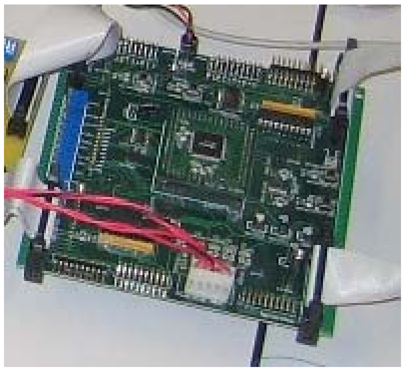
- First Board
  - Acquires image from camera at ~30fps
  - Computes Feature Maps
  - Converts to Spikes
  - Transmits to next board
- Second Board
  - Receives Spikes
  - Computes Additional Feature Maps
  - Transmits data to PC which integrates sensor information to generate behaviour

Speed

- Feature map computation (352 x 288 pixel images)
  - Edge map: 1.2ms
  - Gabor-like orientation map: 2.3ms
- Inter-board map transmission
  - 0.4ms (assuming an average of 5% spikes/frame)

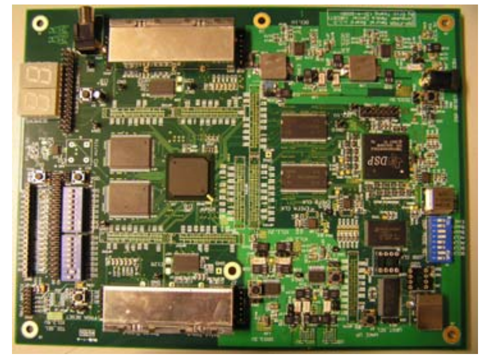
Power dissipation

AER Based



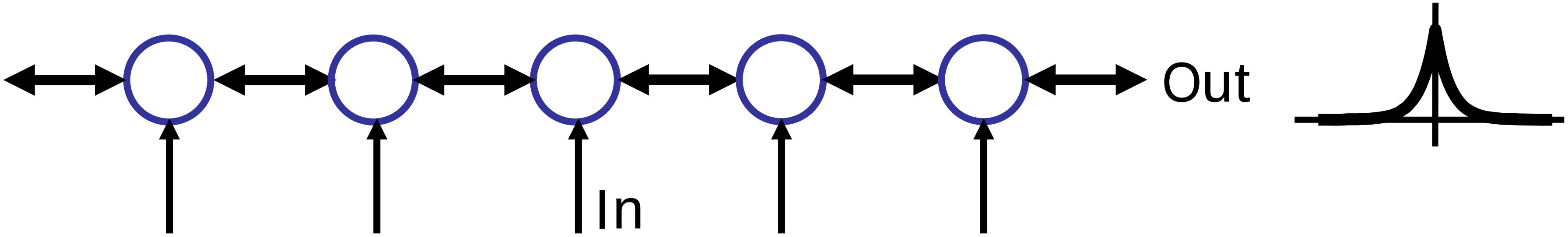
- 45mW
  - Chip: 3mW (1% analog, 99% digital)
  - Support: 42 mW
- 8192 neurons, 1ms settling time
- 5.4nJ/(neuron*map)

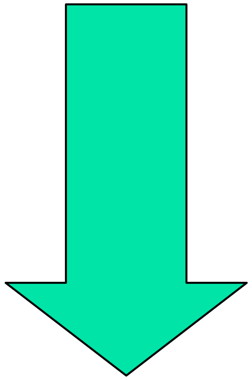
DSP / FPGA Based



- 3.5W
- 405,504 neurons, 2.3 ms map computation time
- 19.9nJ/(neuron*map)

Digital analogue of the gap junction

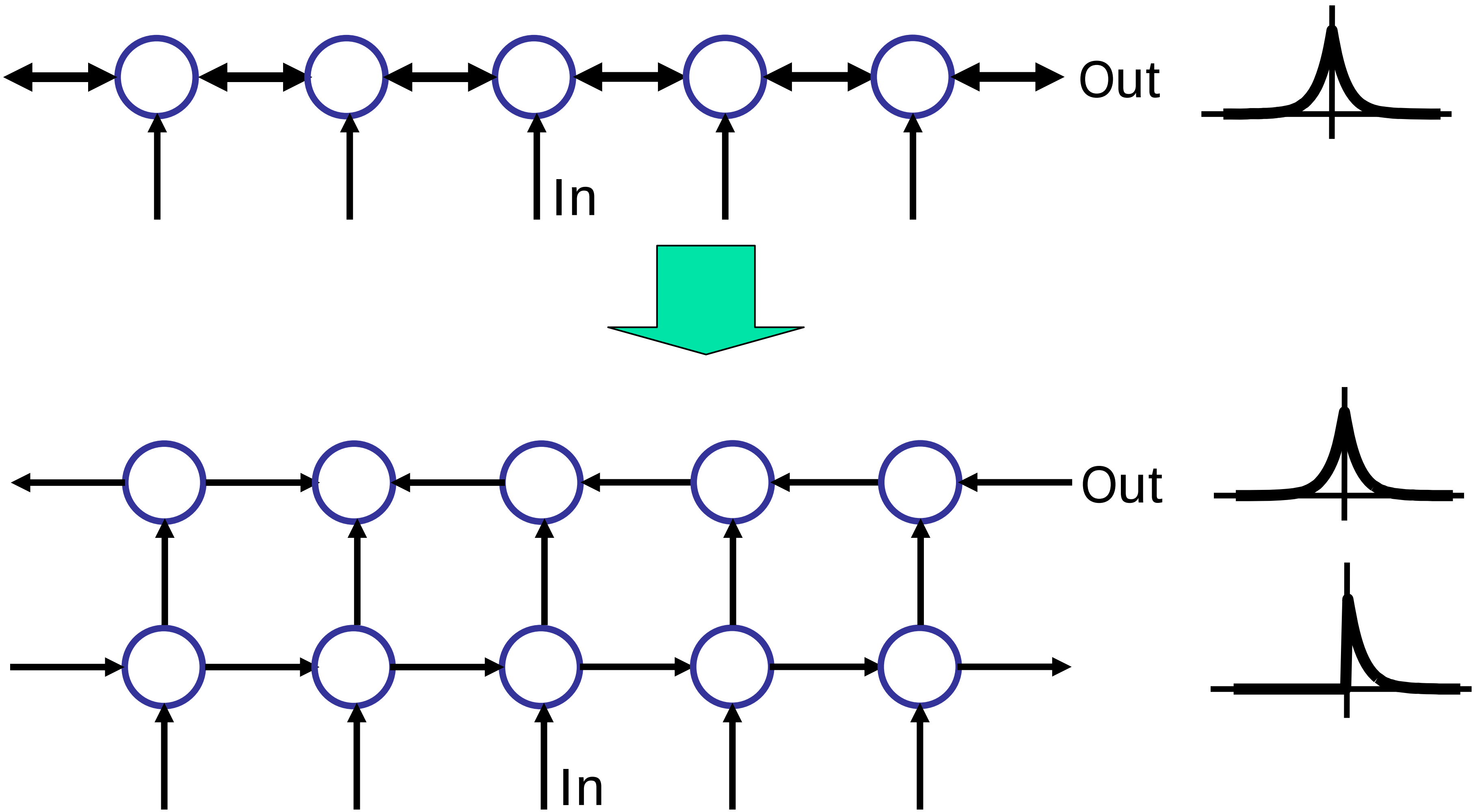




convolution?

integrate the differential equations?

Forward Backward Filtering



Overview

- Visual cortical processing
  - Multidimensional selectivity
- Orientation Selectivity
  - What is it?
  - How do we implement it?
    - AER Multi-chip Architecture
    - DSP/FPGA Architecture
- Joint Orientation/Disparity Selectivity
  - Disparity energy model: position versus phase shifts
  - Population responses versus tuning curves
  - Bigger is not necessarily better

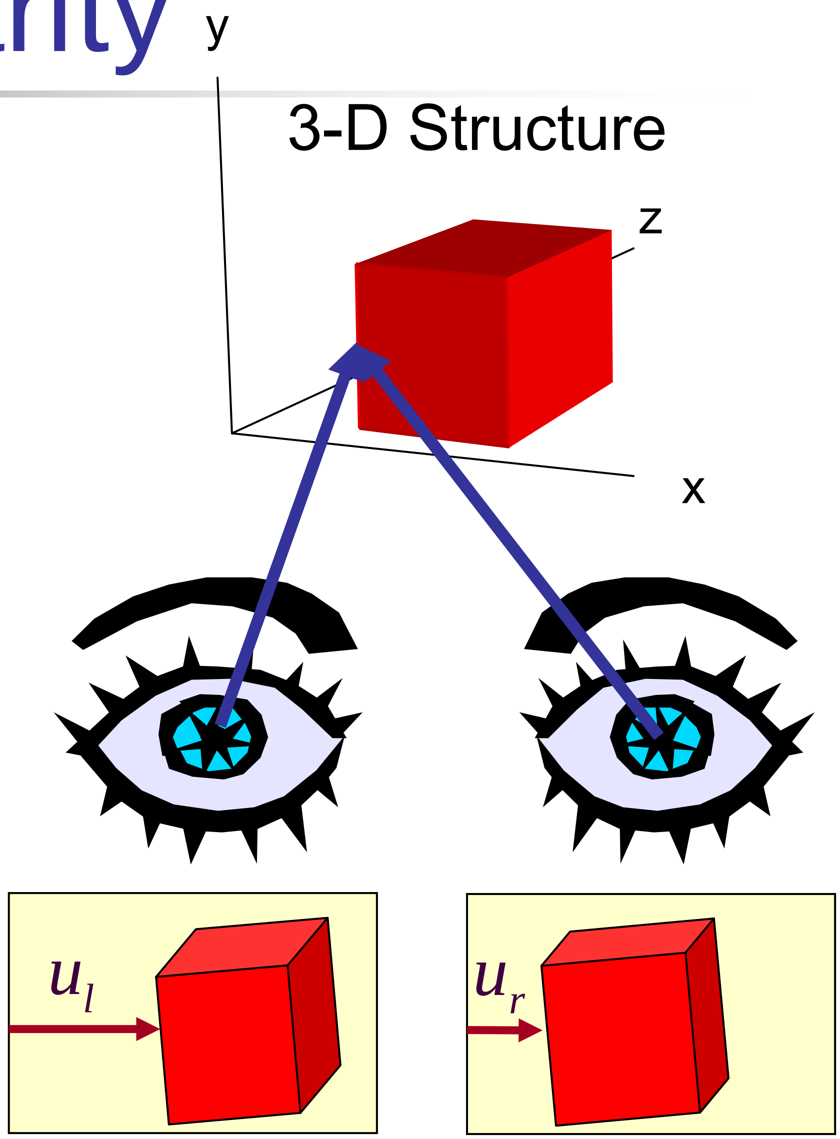
Binocular disparity

- Image disparity is the relative displacement between image points in the left and right eyes corresponding to the same environmental point.

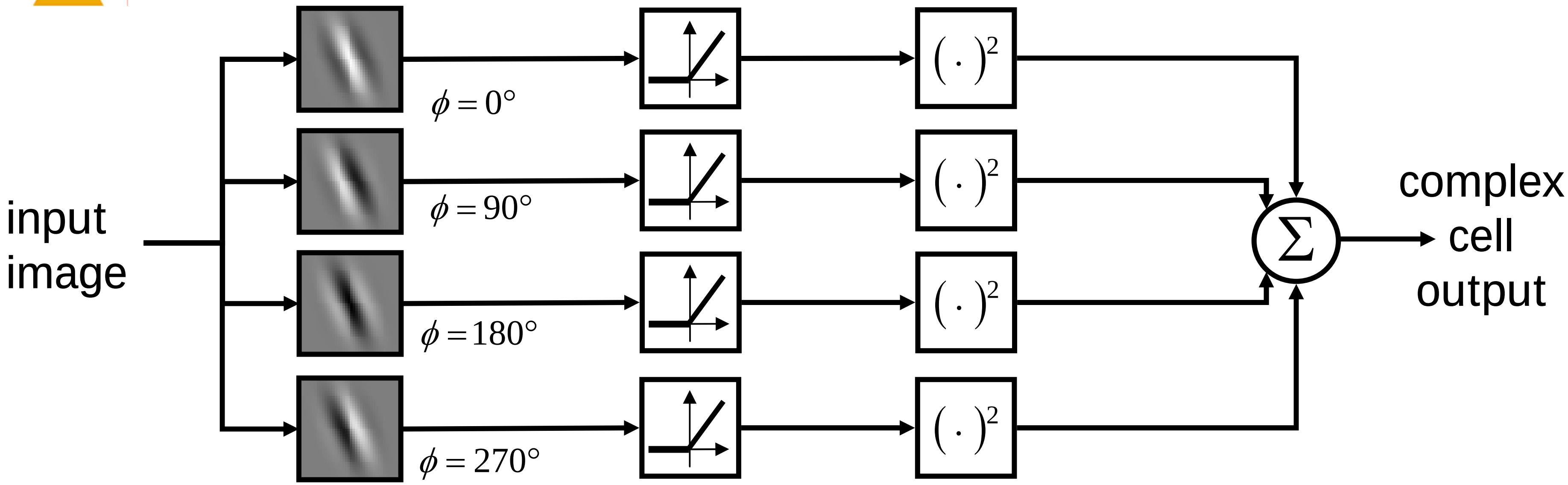
$$d = u_l - u_r$$

- Depth can be inferred from disparity and eye position.

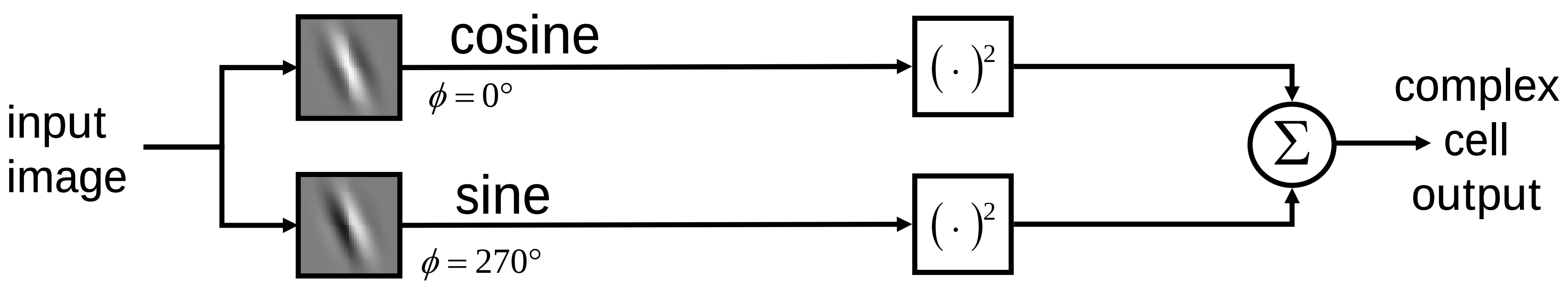
$$\text{Depth} \propto \frac{1}{d}$$



Review: Orientation Energy

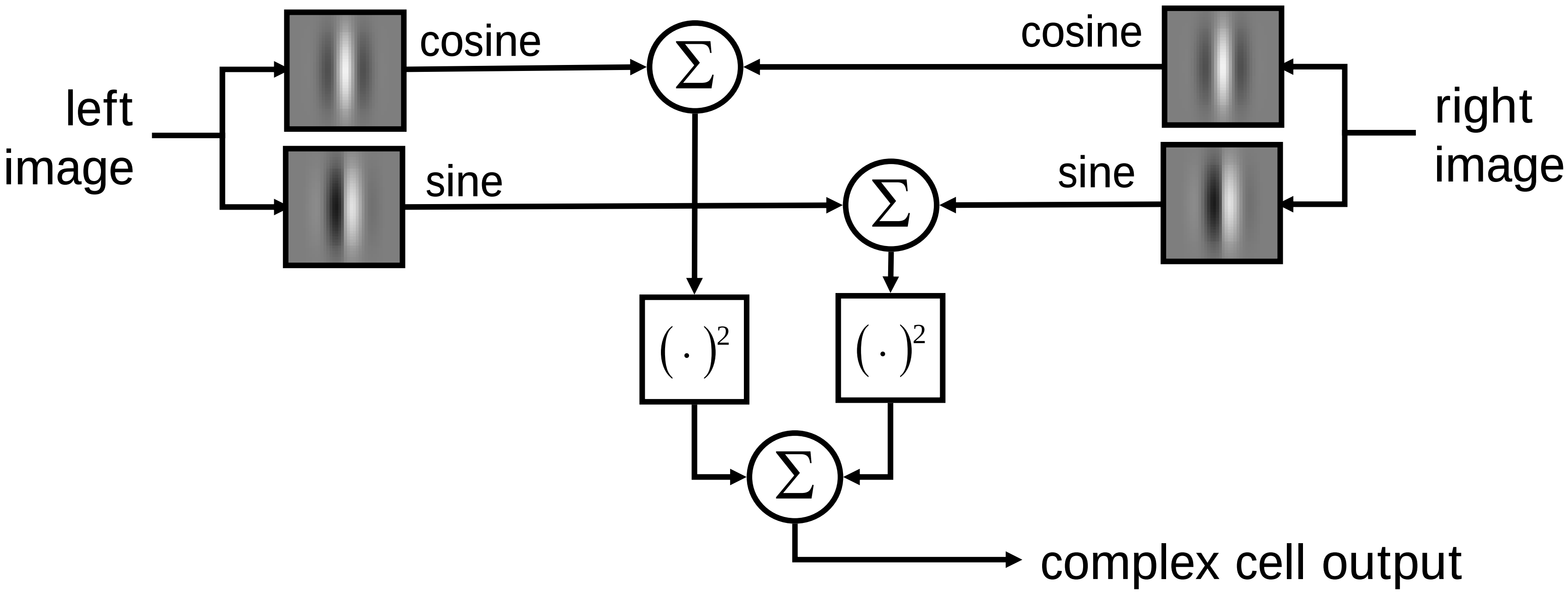


is equivalent to:



Binocular Energy Model

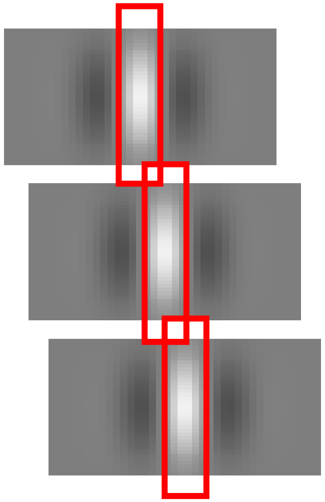
- Zero disparity tuned complex cell:



- I. Ohzawa, G. C. DeAngelis, and R. D. Freeman, "Stereoscopic depth discrimination in the visual cortex: Neurons ideally suited as disparity detectors," Science, vol. 249, pp. 1037-1041, 1990.

Position vs. phase shifts

- Position shifts:

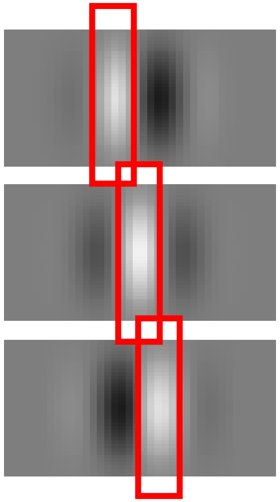


center = -1

center = 0

center = 1

- Phase shifts:



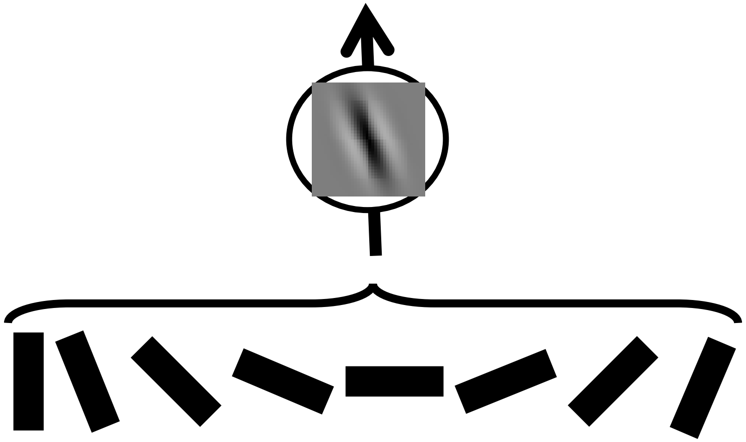
$\phi = 90°$

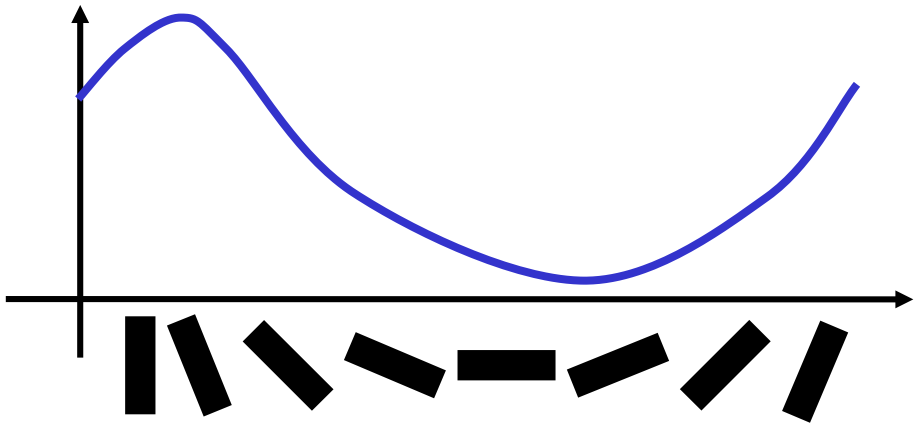
$\phi = 0°$

$\phi = -90°$

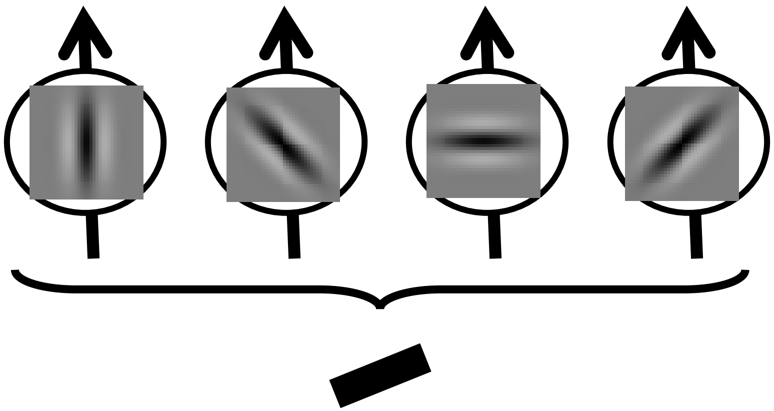
Tuning Curves vs. Population Responses

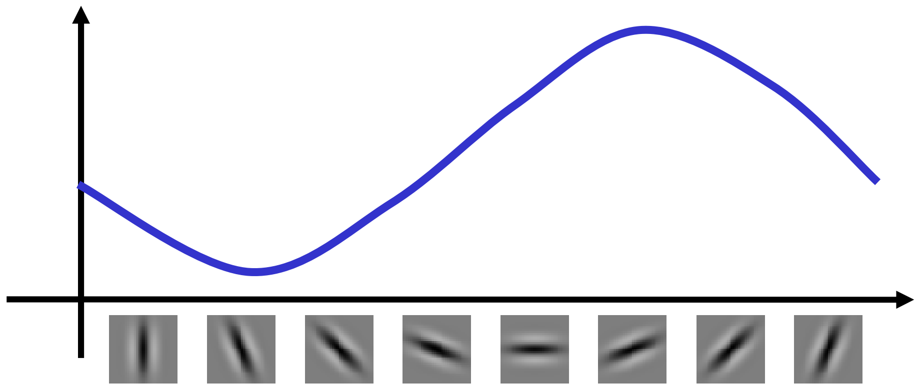
- Tuning curve: one neuron, many inputs



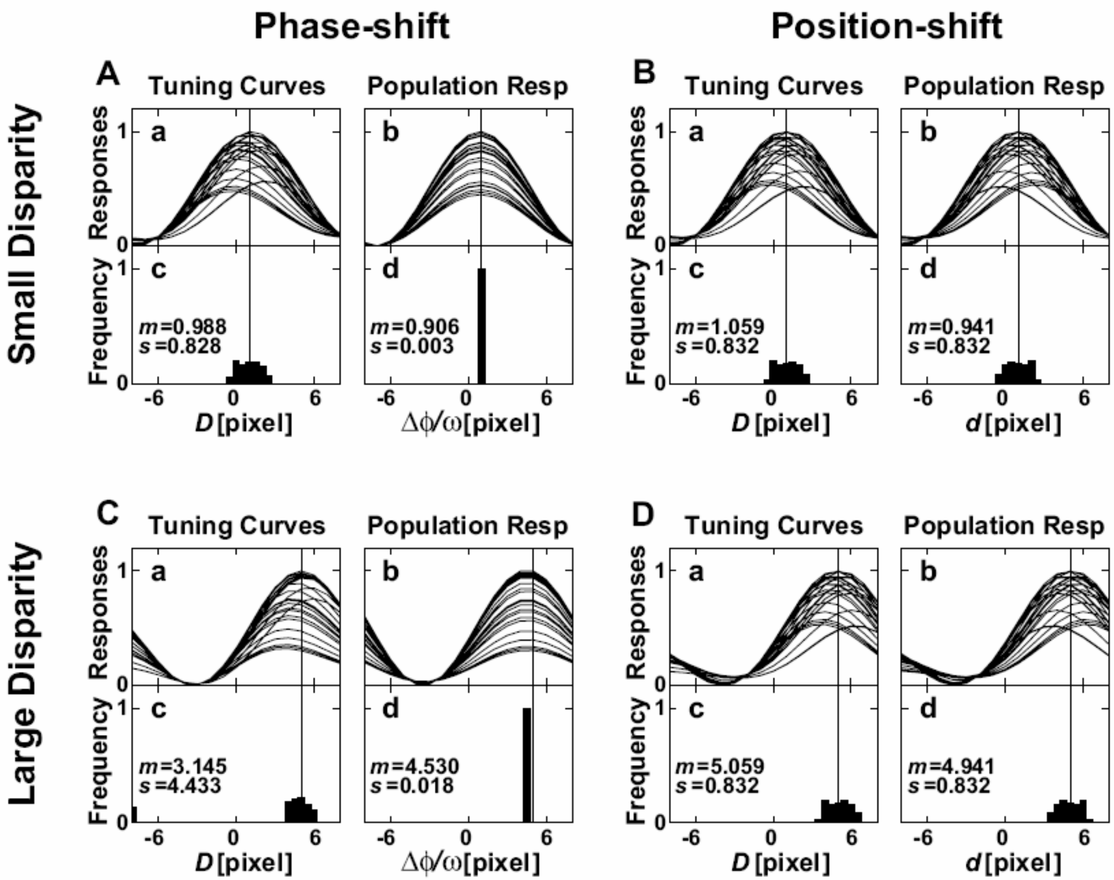


- Population responses: one input, many neurons





Phase-tuned populations are more reliable



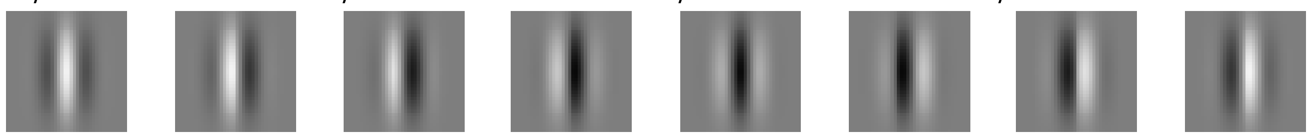
- Y. Chen and N. Qian, "A course-to-fine disparity energy model with both phase-shift and position-shift receptive field mechanisms," *Neural Computation*, vol. 16, pp. 1545-1578, 2004.

Limited disparity ranges in populations

- The disadvantage of phase-tuned neurons is that their preferred disparity range is limited by the periodicity of the cosine in the Gabor.

$\phi = 0^\circ$ $\phi = 90^\circ$ $\phi = 180^\circ$ $\phi = 270^\circ$

$\theta = 0^\circ$



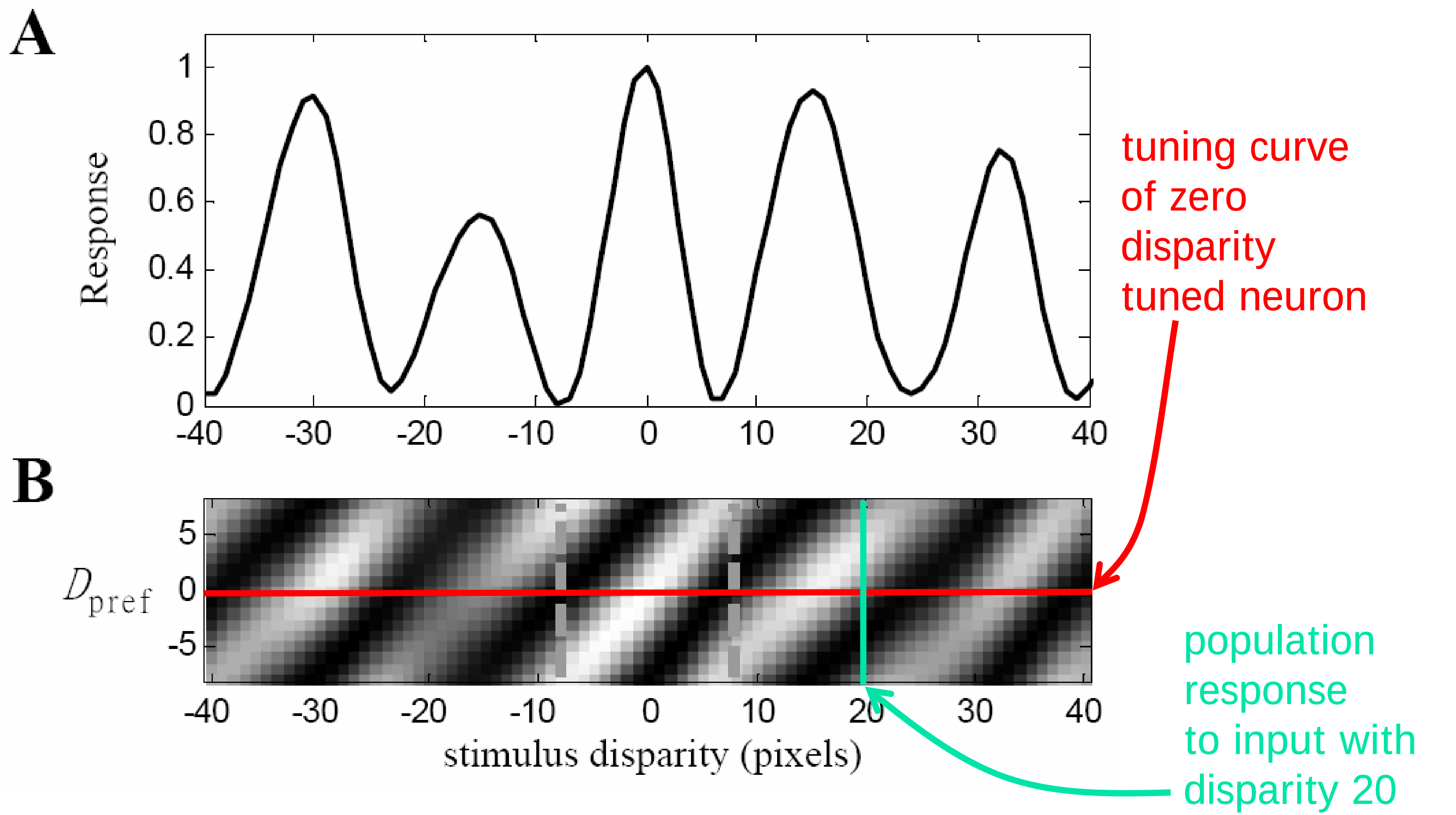
- The measured disparities of V1 neurons ranges over a few degrees, but actual scene disparities range over tens of degrees.
- Psychophysical results indicate that we fuse binocular stimuli only over a few degrees (Panum's fusional area).

Detecting out of range disparities

- The mismatch between the small range of preferred disparities in cortex and the large range of scene disparities suggests that there should be a mechanism for detecting whether a population response is due to an input disparity “in the range” of preferred disparites or “out of the range”

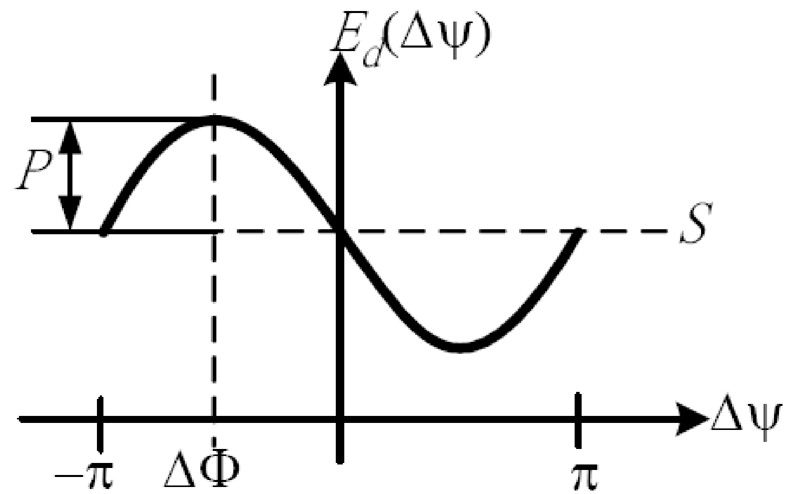
- Unfortunately, this problem is complicated due to the fact that the tuning curves of disparity tuned neurons are not unimodal.

False peaks in phase-tuned populations



Phase tuned population responses

- The population response of a phase-tuned disparity neuron population depends on three parameters:
  - S = average activation
  - P = difference between peak and average activation
  - DF = peak location

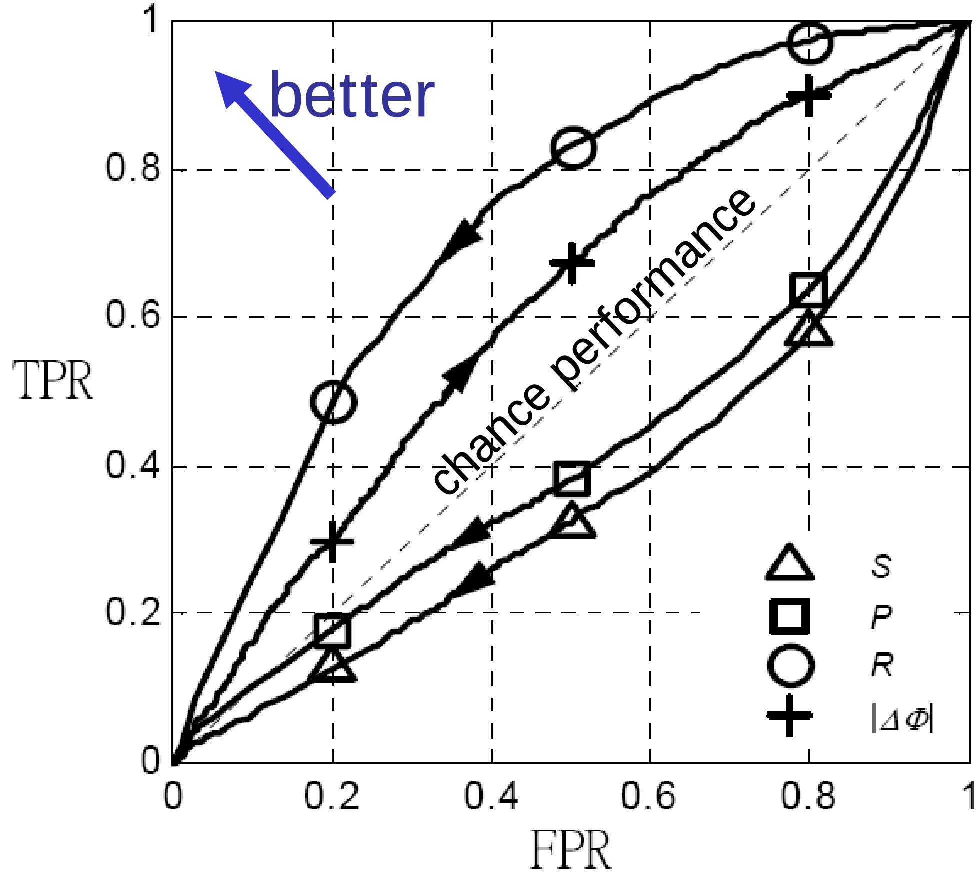


- Hypothesis: Since neurons respond strongly to their preferred inputs, if the average (S) activation or the difference between the peak and average activation (P) is large, the response is more likely to be reliable (i.e. “in the range”)

Bigger is not necessarily better

- TPR = true positive rate
  - The fraction of “in the range” inputs correctly classified
- FPR = false positive rate
  - The fraction of “out of range” inputs incorrectly classifed as “in the range.”
- Thresholding S and P performs worse than chance, indicating that ***smaller*** responses are actually more indicative of “in the range” disparities.

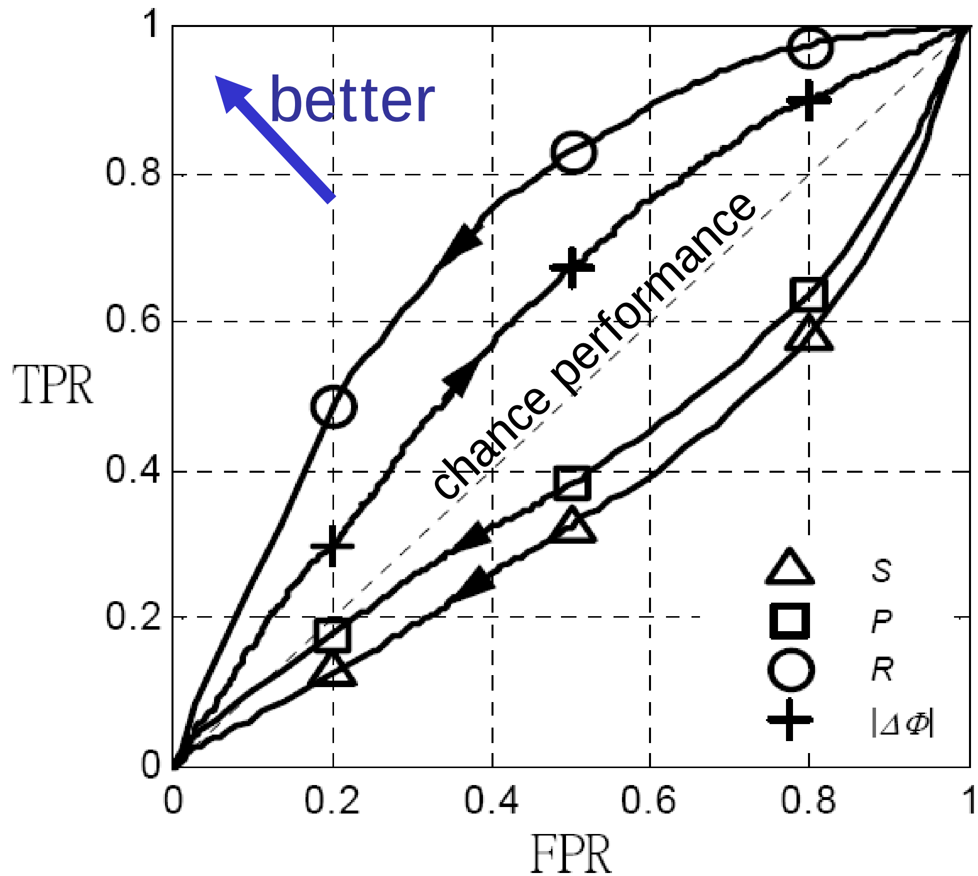
ROCs of thresholding classifier



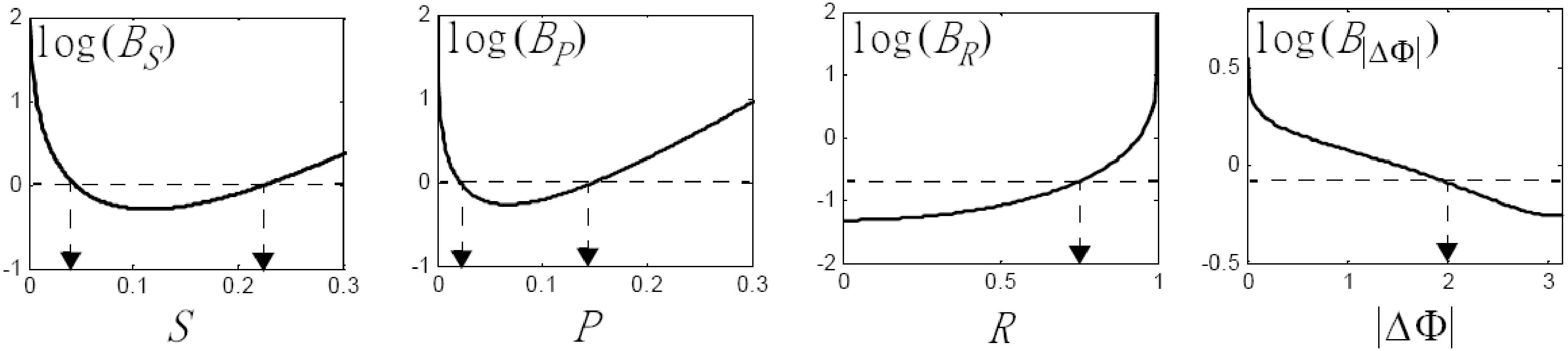
Normalization enables robust validation

- R = P/S performs reliably
  - The difference between the peak and average activation normalized by the average activation.
- Normalization is commonly used to account for nonlinear properties observed in cortical neurons.
- These results suggest a functional role for normalization.

ROCs of thresholding classifier



Statistical Analysis



- B_F = Bayes factor for feature F: a measure of the evidence that the input disparity is “in the range” given the observed F.
  - $\log(B_F) > 0$: positive evidence for “in the range”
  - $\log(B_F) > 0$: positive evidence for “in the range”

Disparity Video

input | d = 0 | d = 0 | d = 1 | d = 2

raw disparity energy

original image gated by R>T

Summary

- Visual cortical processing
  - Multidimensional selectivity
- Orientation Selectivity
  - What is it?
  - How do we implement it?
    - AER Multi-chip Architecture
    - DSP/FPGA Architecture
- Joint Orientation/Disparity Selectivity
  - Disparity energy model: position versus phase shifts
  - Population responses versus tuning curves
  - Bigger is not necessarily better

End of Presentation

Thanks for listening!

email: eebert@ee.ust.hk

香港科技大学
HONG KONG
UNIVERSITY OF
SCIENCE &
TECHNOLOGY

Acknowledgements

- Thomas CHOI (HKUST)
- Eric TSANG (HKUST)
- Stanley LAM (HKUST)
- Yicong MENG (HKUST)
- Kwabena BOAHEN (Stanford)
- Paul MEROLLA (U. Penn)
- John ARTHUR (U. Penn)

Architectural Advantages

- Since neurons have identical response properties, retinotopic arrays are constructed by *tiling identical circuit blocks*.
- Neurons are only locally interconnected, which *simplifies wiring*.
- ON-OFF and spike based representation, *lowers power consumption and fixed pattern noise.*
- Continuous time operation enables *feedback interactions between maps.*

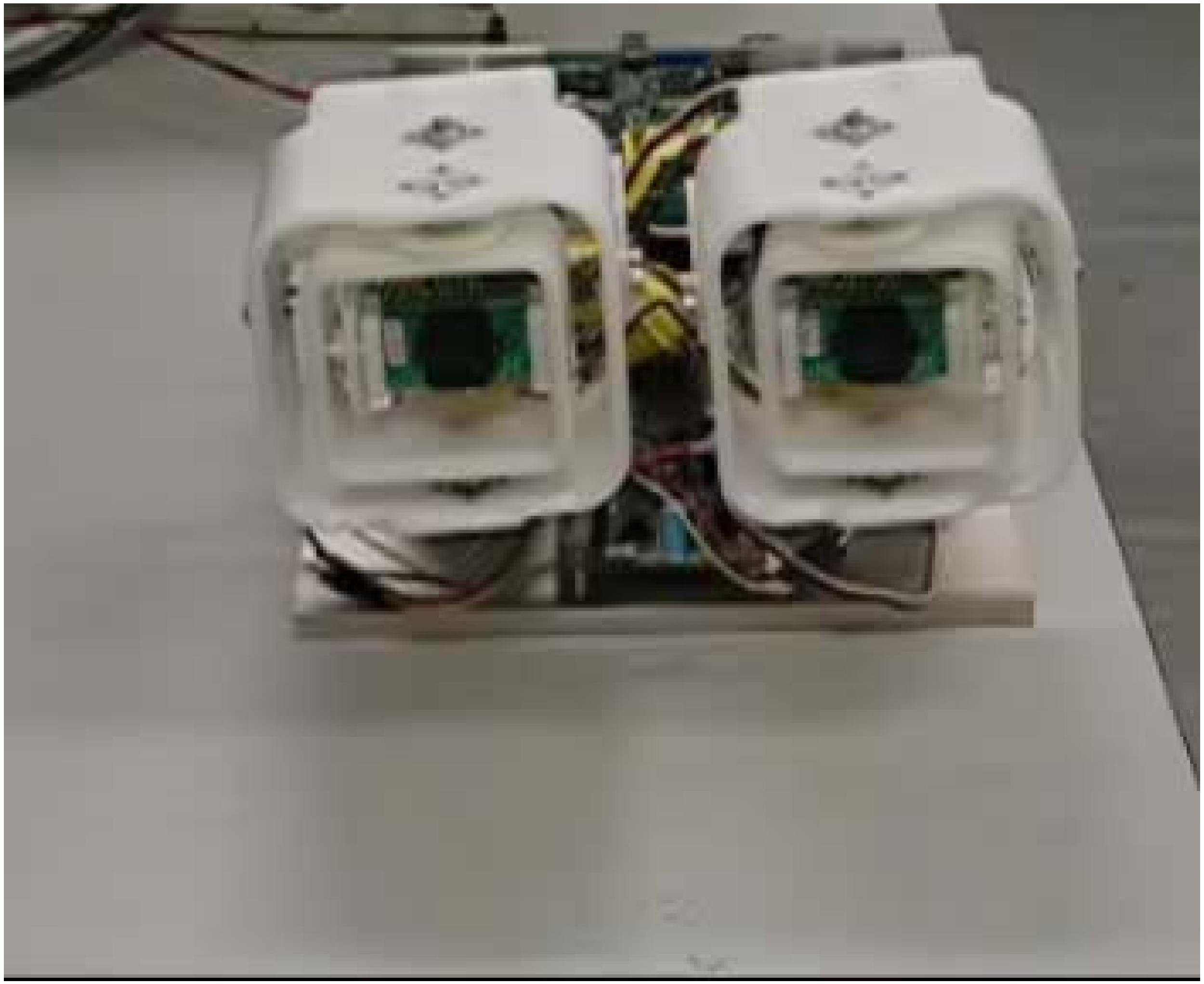
Conclusion

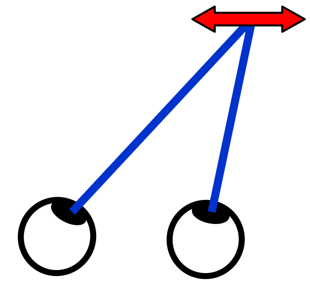
- We have developed digital hardware for real-time simulations of feature maps inspired by the visual cortex.
- This system is a rapidly reconfigurable test bed for investigating active visual perception based upon the outputs of model visual cortical columns.
- Algorithms and architectures developed on this hardware will guide the development of mixed signal neuromorphic chips, and multi-chip networks.

Conclusion

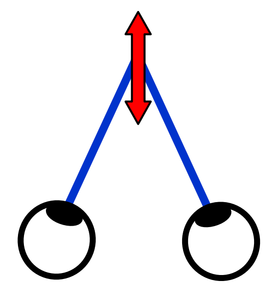
- We have developed silicon networks of neurons selective for position, spatial frequency, orientation and (binocular disparity <u>or</u> direction/speed of motion)
- Neurons support both feedforward and feedback interactions even though they may reside on different chips.
- The largest of these systems contains over 30,000 recurrently connected continuous-time neurons.
- Next step: Retinotopic arrays of neurons simultaneously selective for disparity <u>and</u> motion.

Active Binocular Tracking





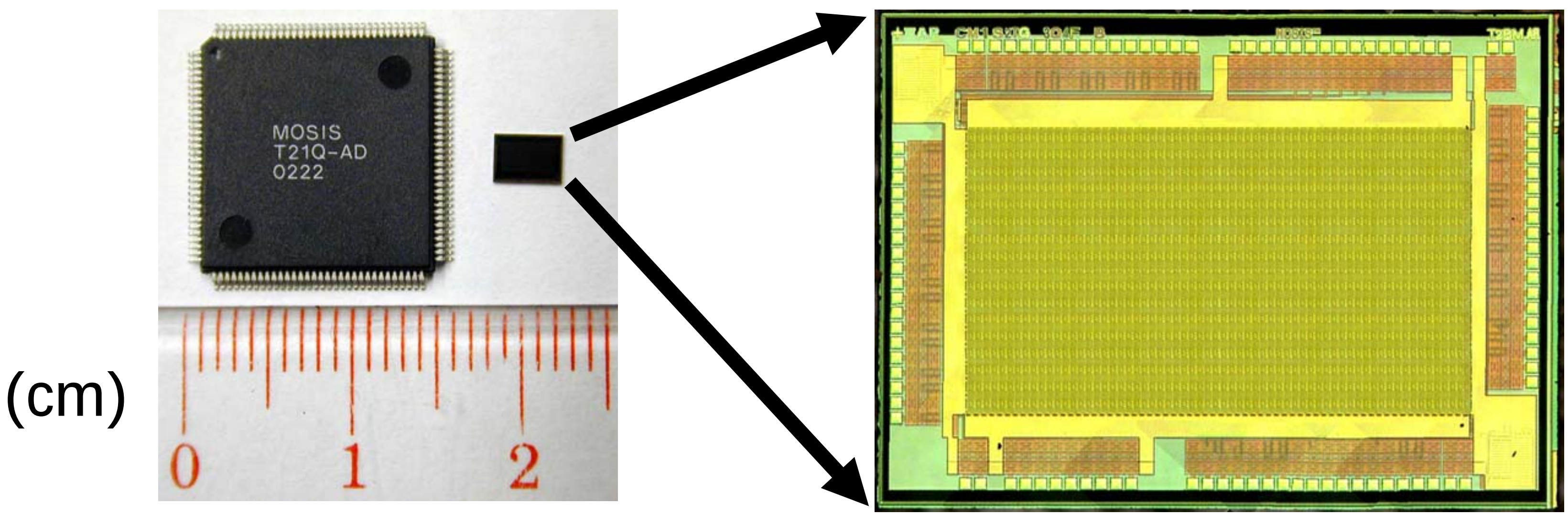
Version
controlled by centroid of zero disparity cells



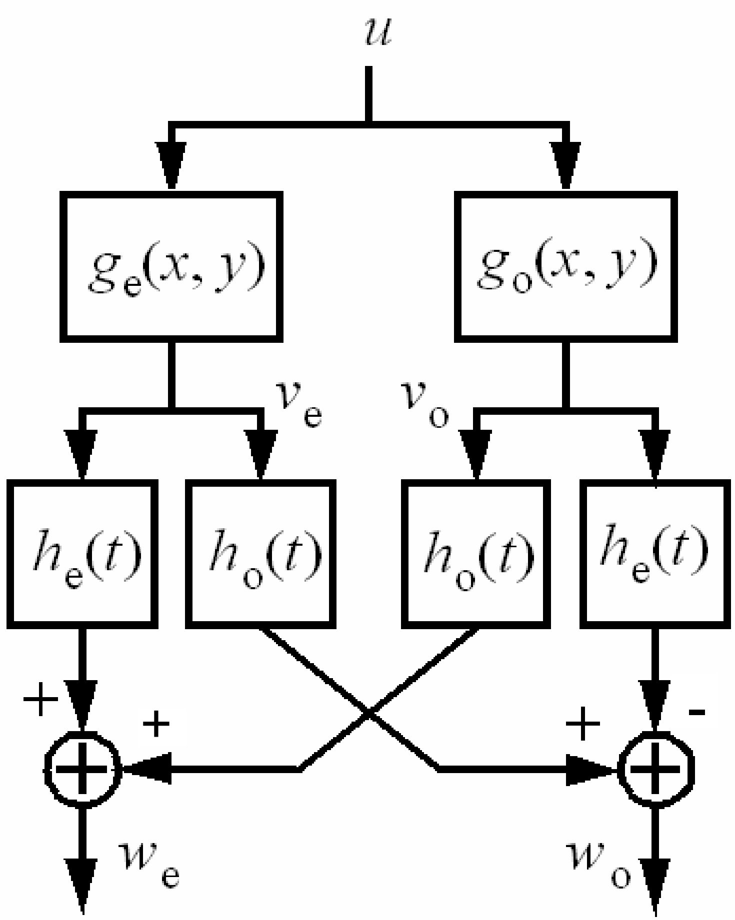
Vergence
controlled by difference between near and far disparity cells

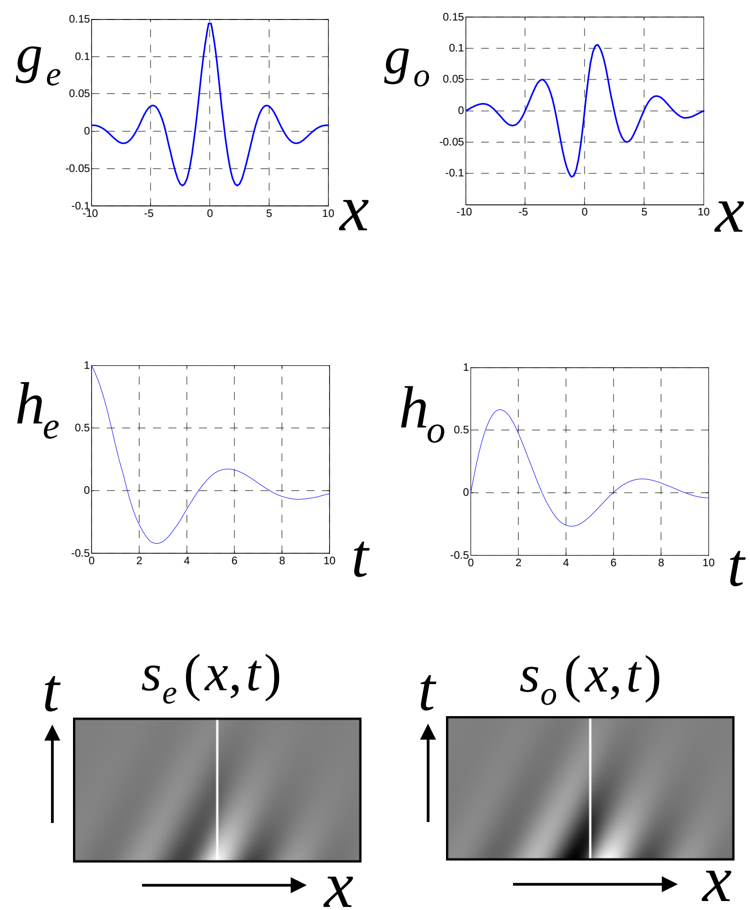
Chip Data

- Neurons: 32x64x4 neurons (8000)
- Technology: TSMC 0.25um
- Die size: 3.84mm x 2.54mm (9.8mm^2)
- Power dissipation: 3mW



Motion Energy Model



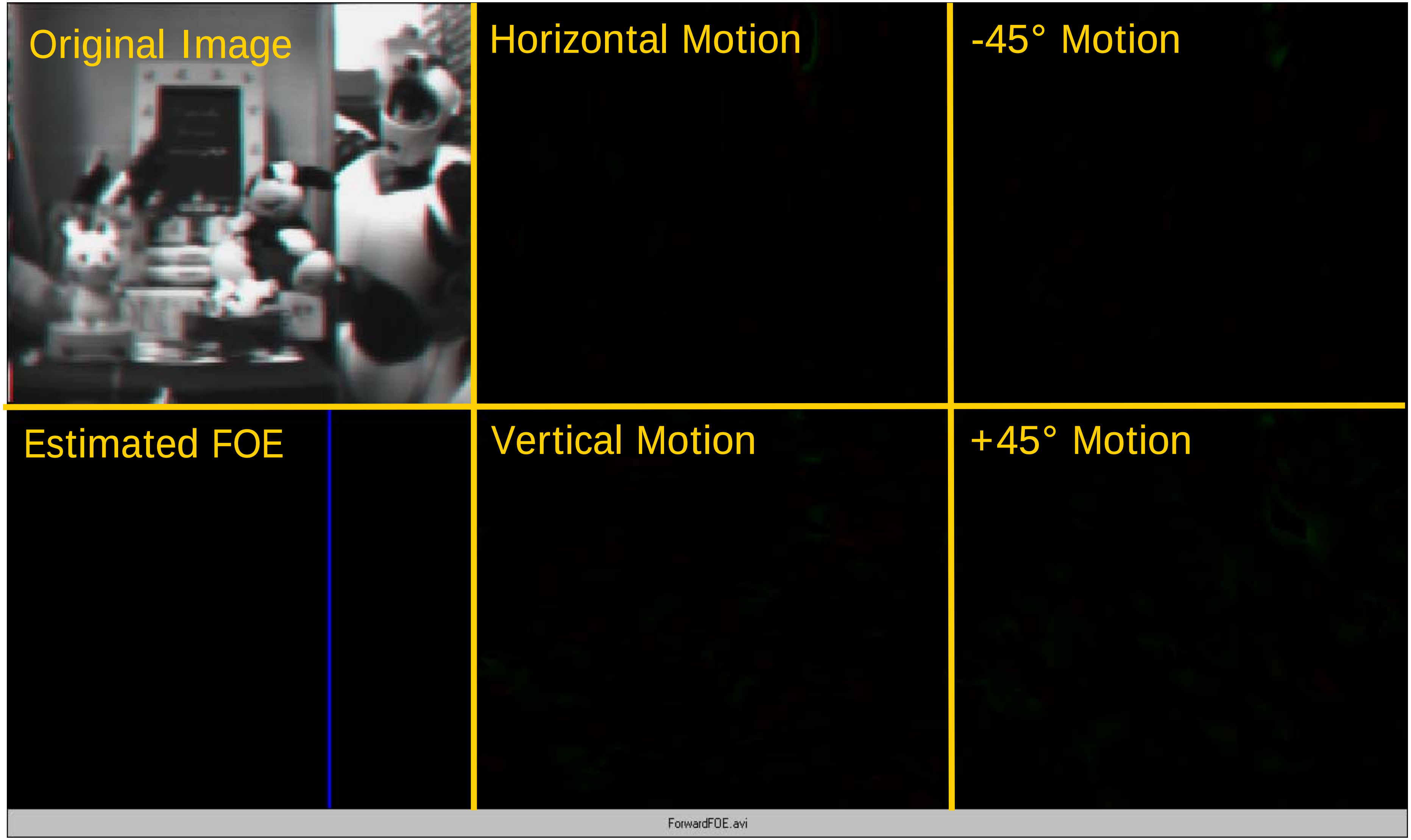


Spatial Filtering

Temporal Filtering

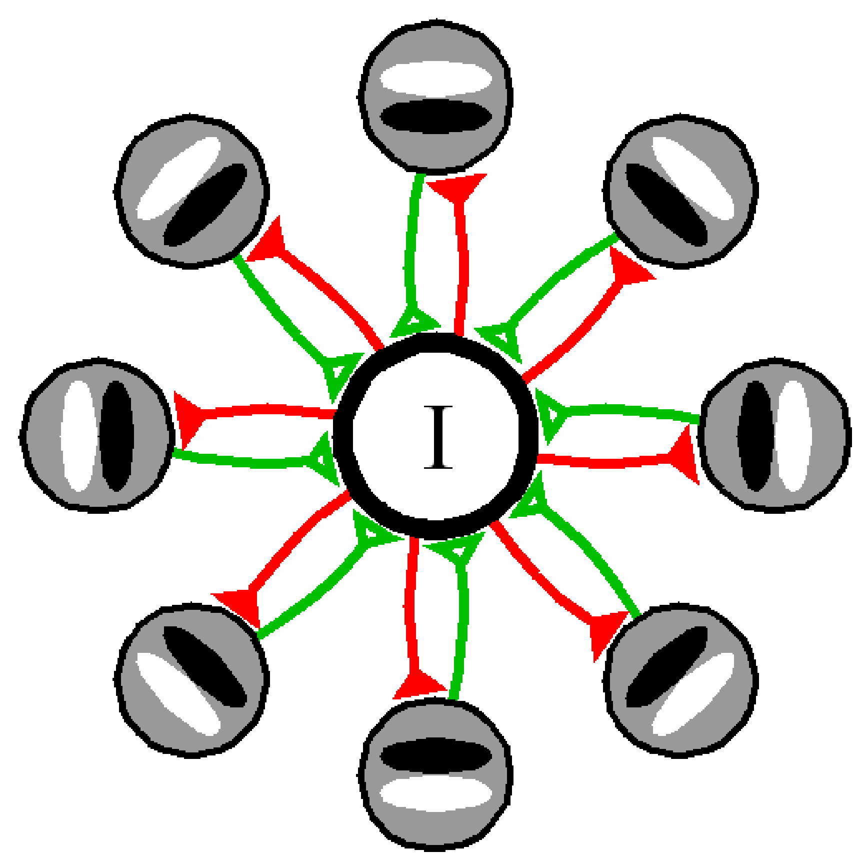
Spatio-Temporal Filtering

Estimation of Focus of Expansion

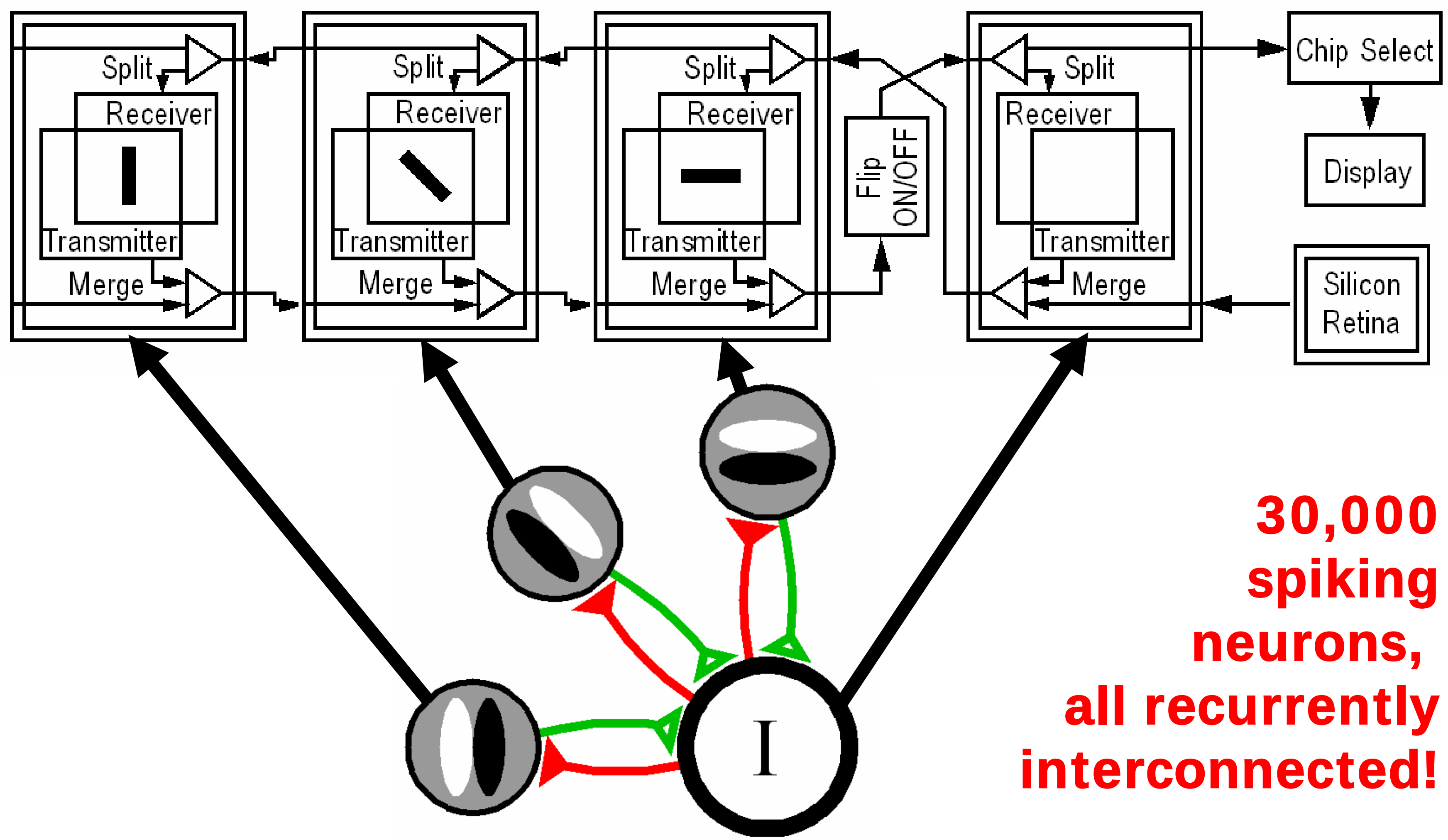


Feedback System

- A global inhibitory neuron pools input across different orientations from neurons tuned to the same retinal location with the same polarity (ON/OFF) and symmetry (EVEN/ODD).
- The inhibitory neuron sends inhibition to the neurons that excite it.



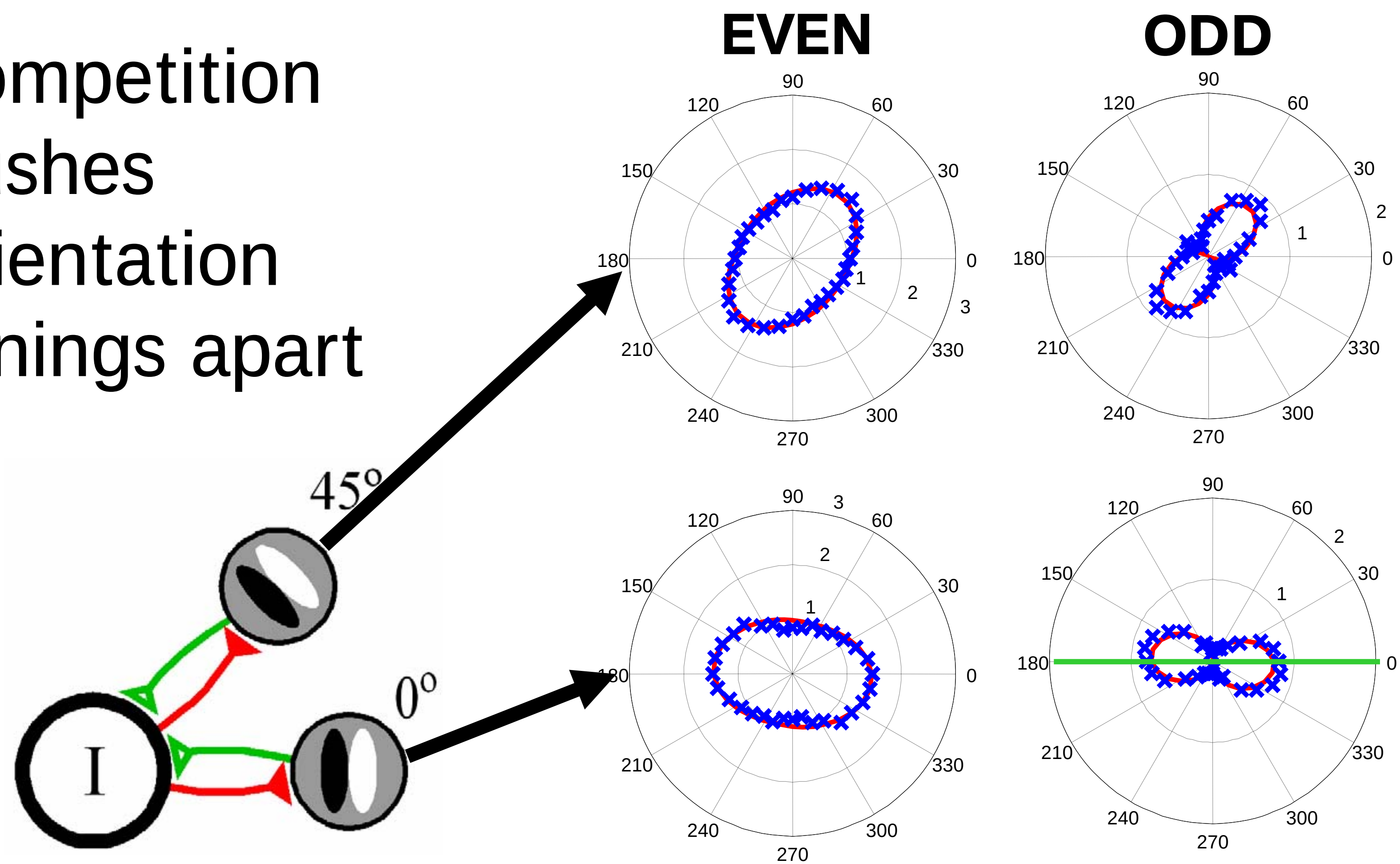
Feedforward System



30,000 spiking neurons, all recurrently interconnected!

Shifts in orientation tuning

- Competition pushes orientation tunings apart



— = **prediction**   x = measured

Model vs. Implementation

Model

- Graded interactions between neurons
- Linear model
- Perfect matching

Implementation

- Spike interactions between neurons
- Nonlinear model due to ON/OFF rectification
- Lots of mismatch

Applications

- Binocular tracking
- FOE estimation