Biologically Inspired CMOS Vision Sensors

Jan Van der Spiegel

jan@seas.upenn.edu

University of Pennsylvania

Department of Electrical and Systems Engineering

Philadelphia, PA 19104

*SSCS **Fort Collins Chapter, Colorado (Denver Section)***

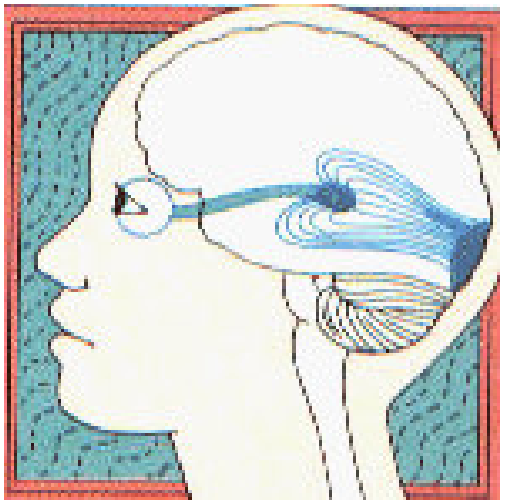
July 27, 2005

Overview

- **Why biologically inspired sensors?**
- **Biological visual system**
  - Computational strategies, Neural circuits
- Neuromorphic **Vision Sensors:**
  - Sensor for the detection of **image features (spatial)**
  - **Tracking sensor (spatio-temporal)**
- Conclusions

A picture is worth a thousand words

Biological Vision System

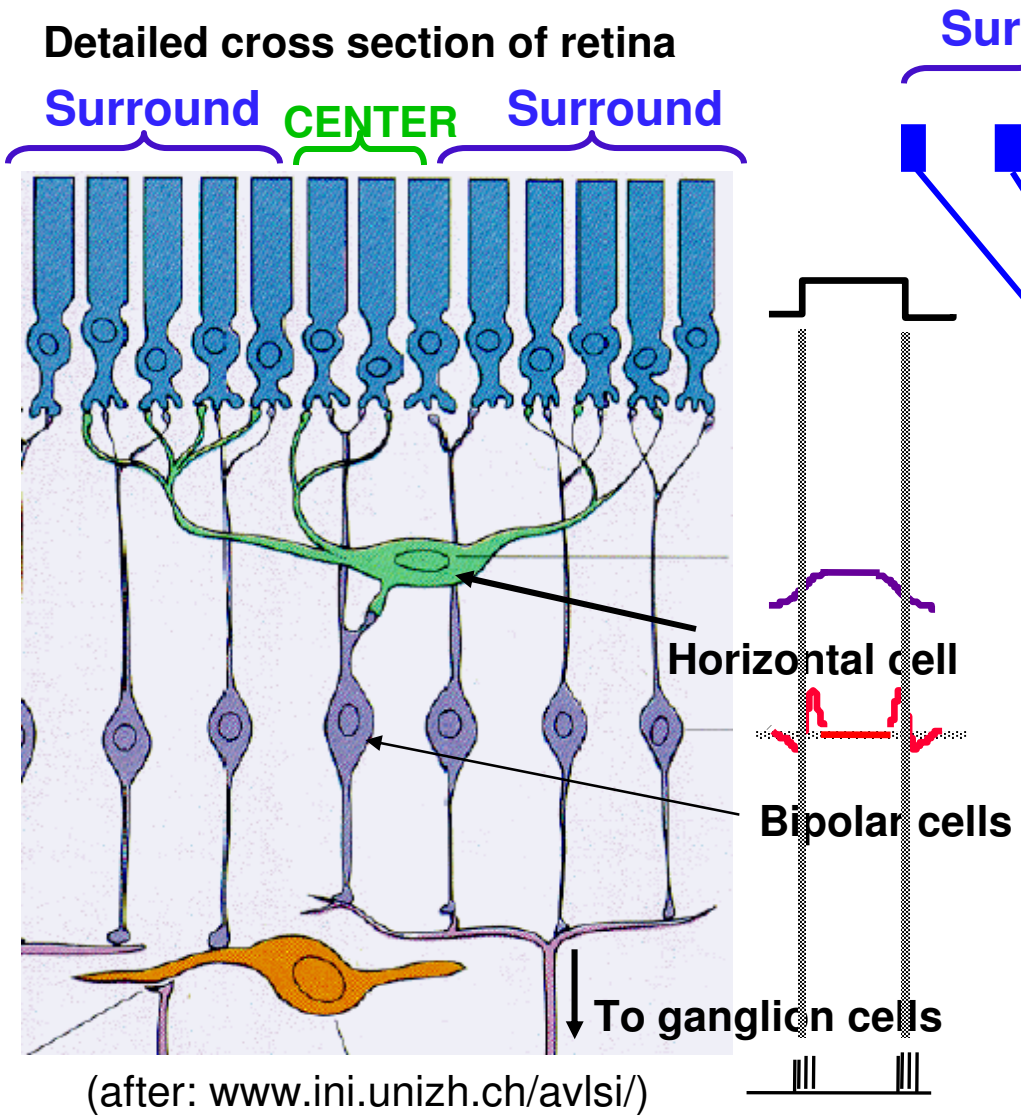


Retina

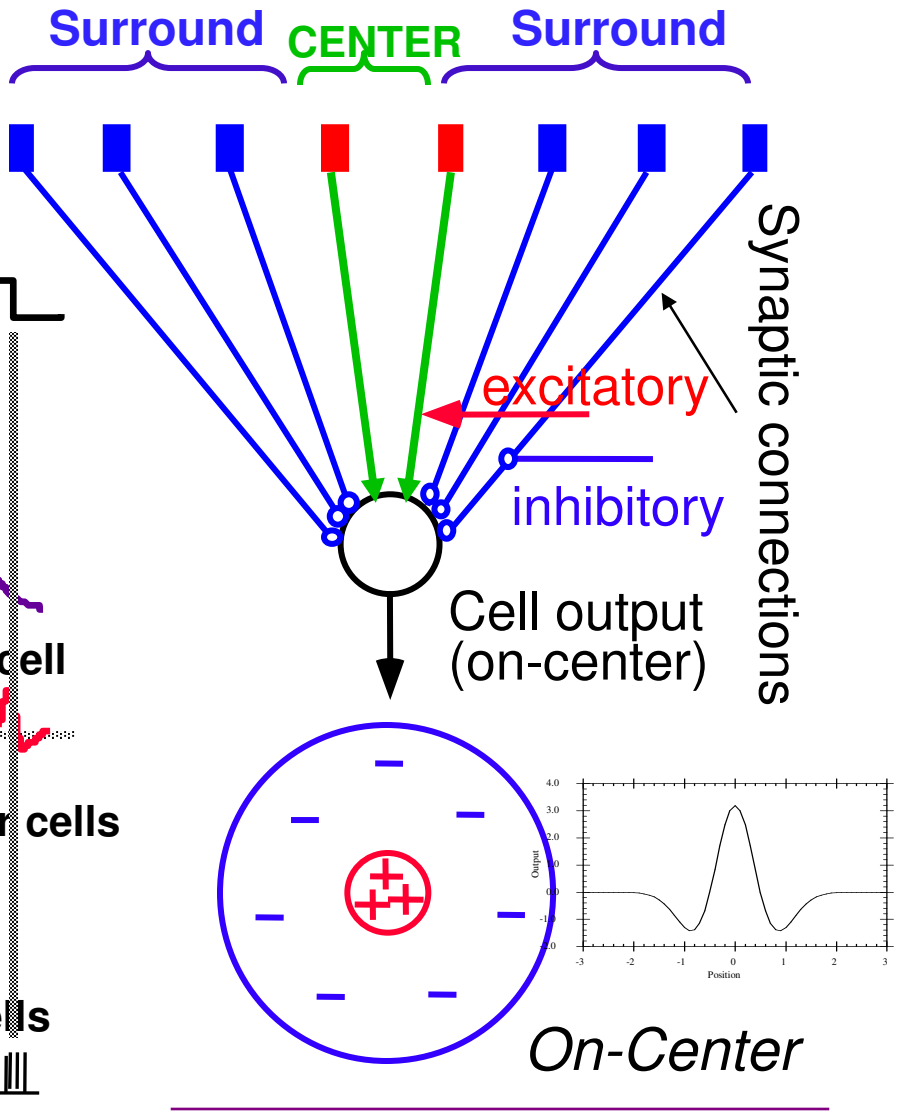
Receptive Fields

Visual Cortex

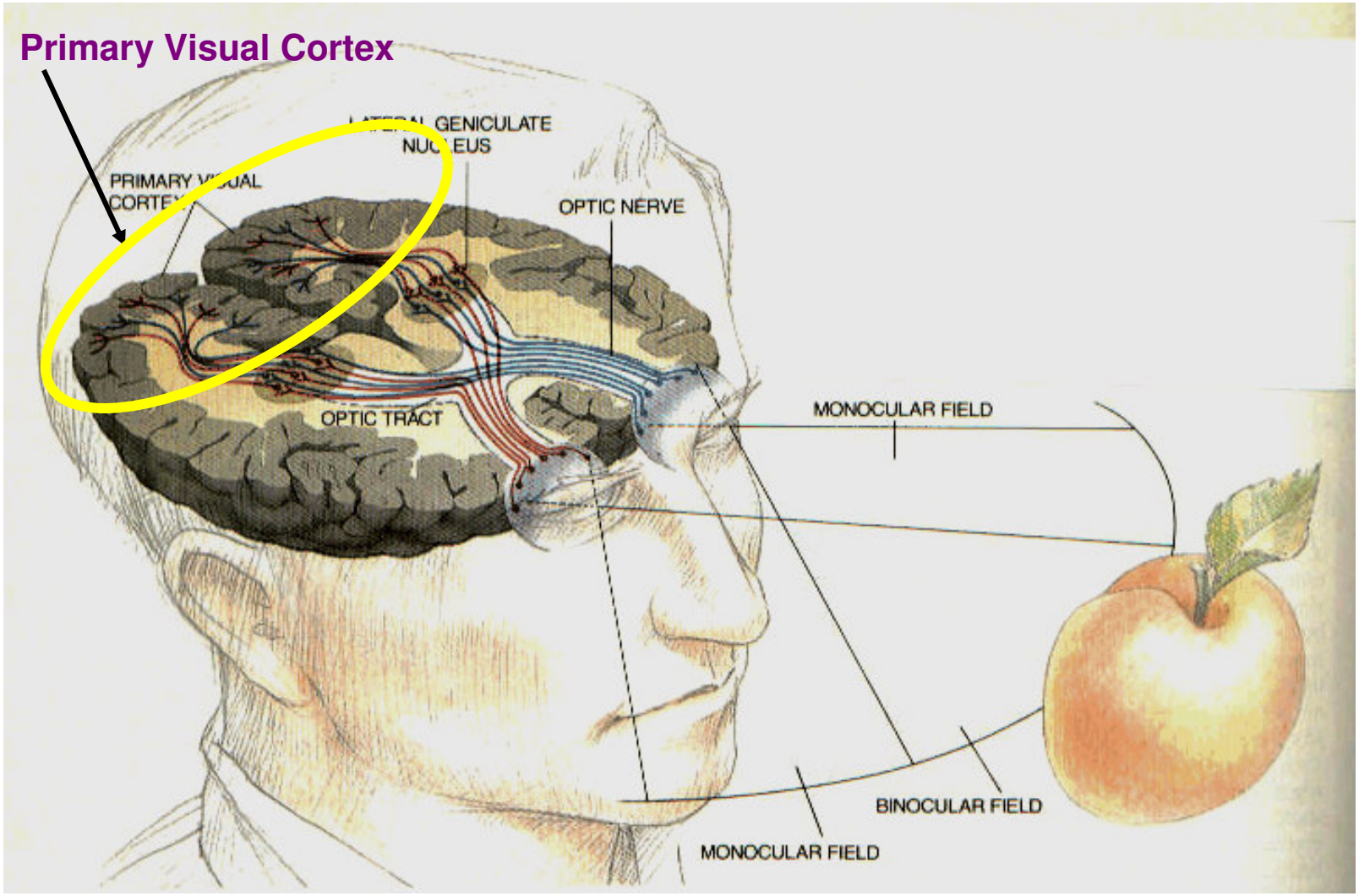
Retina - Center surround receptive field



(after: www.ini.unizh.ch/avlsi/)

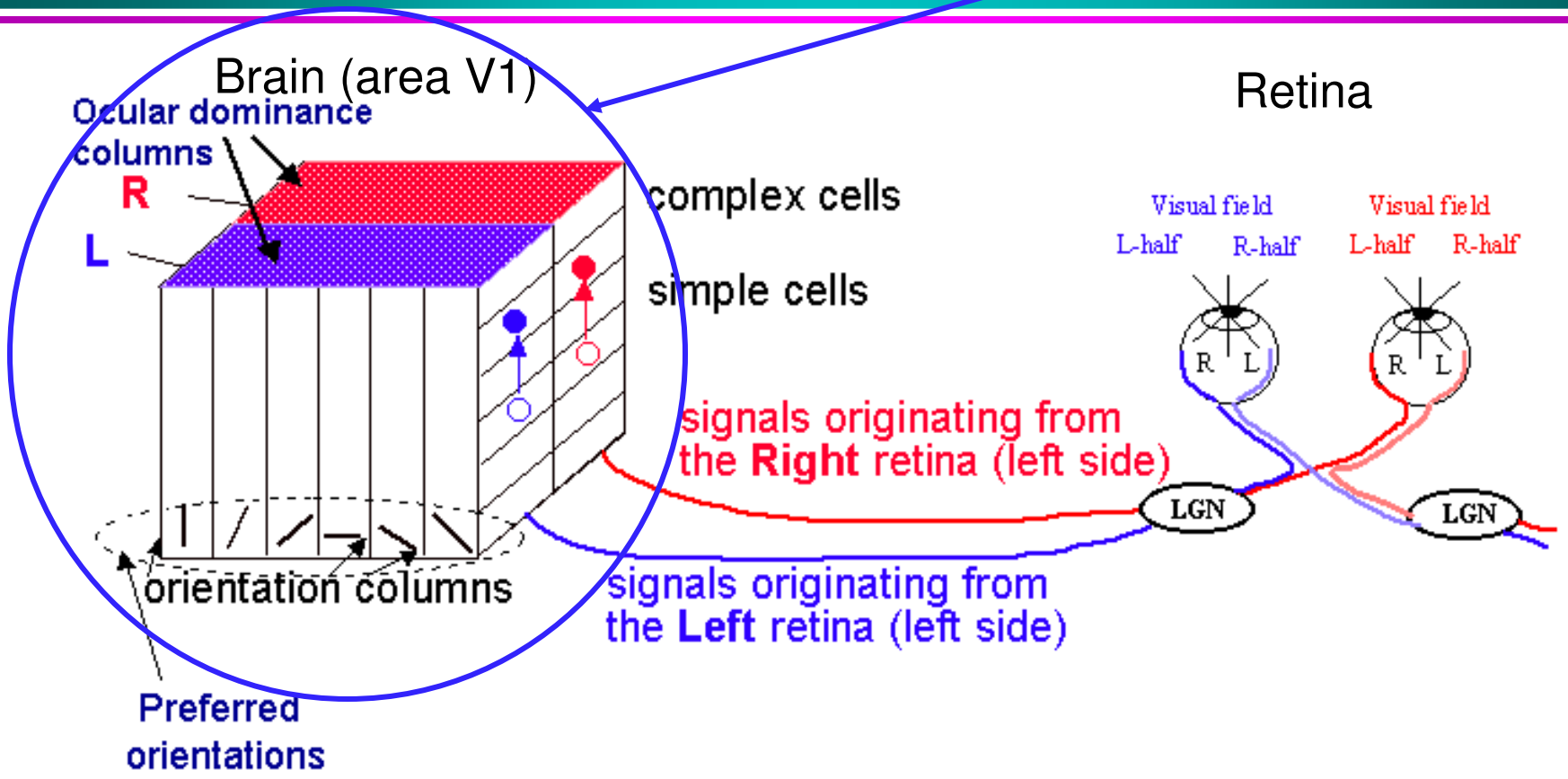


Processing in the Visual Cortex



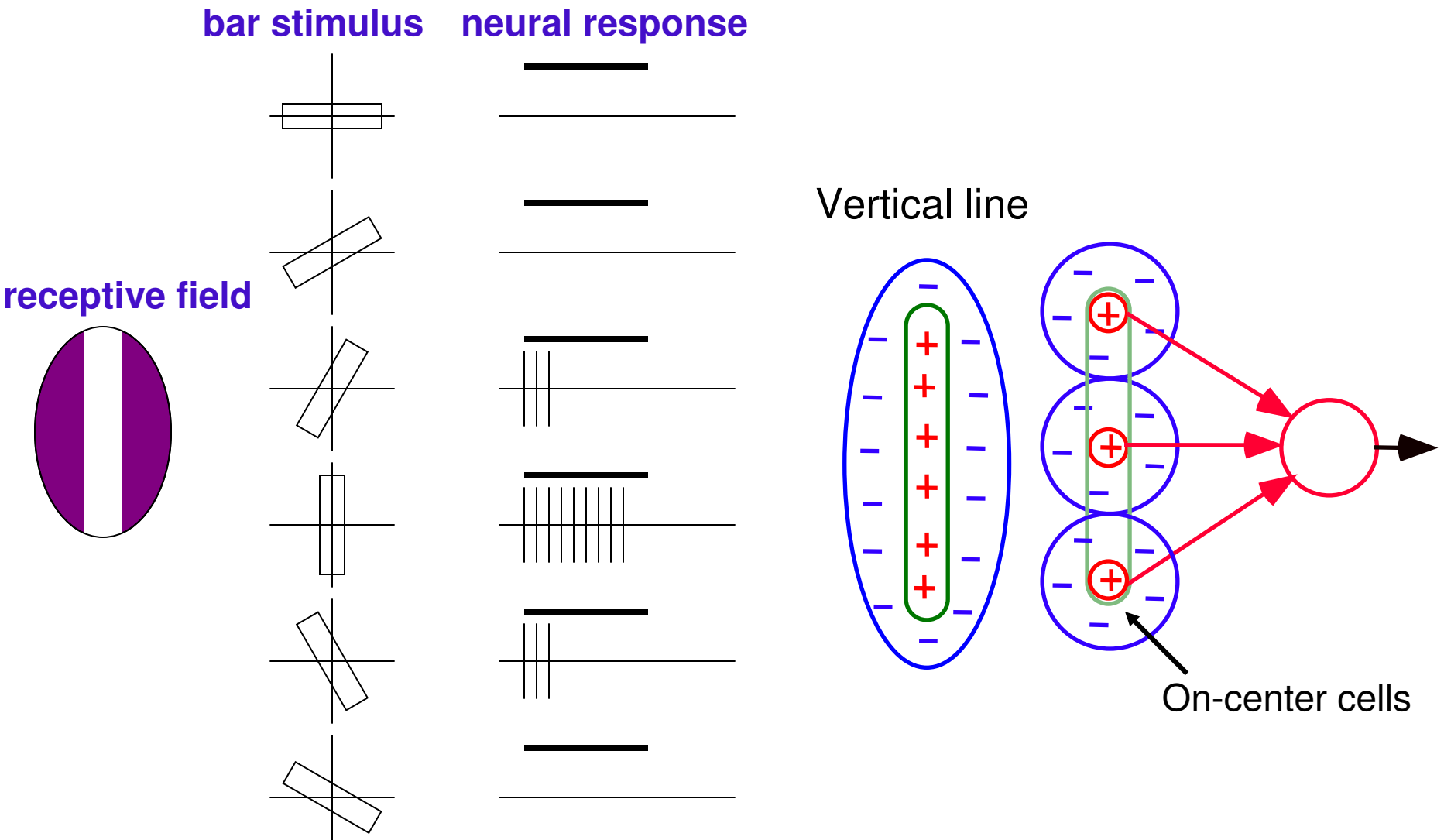
(Scientific American)

Processing in the Visual Cortex

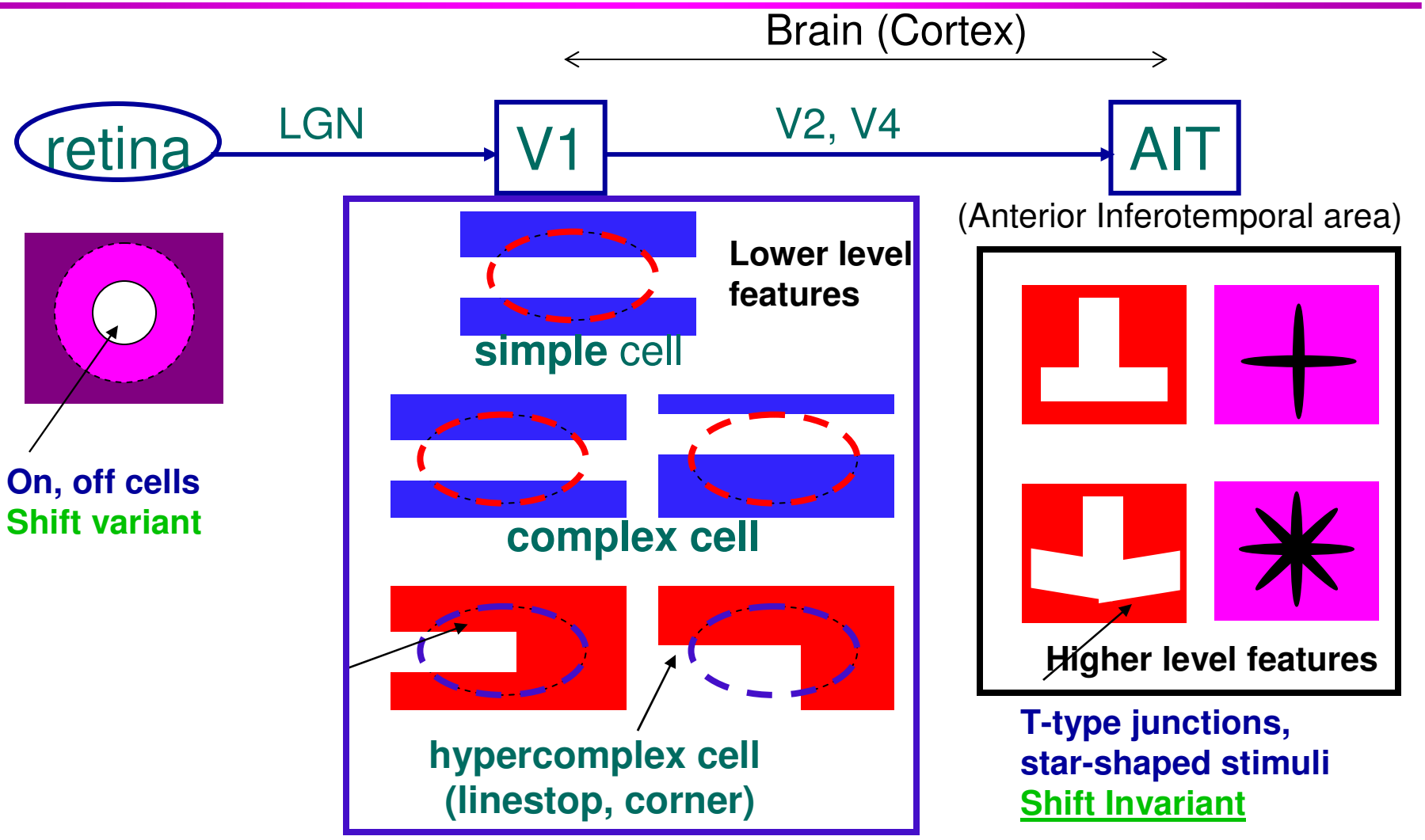


- Cortical area V1: hypercolumn structure
  - Orientation, edge, line stops, etc detectors

Orientation selectivity of a simple cell



Information flow from retina to brain



In summary: the biological system...

- **Decomposes** a picture in many **features**: edges, orientation, line stops, junctions, onset (in time) etc.
- The **features are integrated at higher level** into a more conceptual representation.
- **Highly structured**, parallel and hierarchical.
- Distributed architecture leads to:
  - Data reduction; fast processing (parallelism), robustness

Neuromorphic Vision Sensors

(Sc. American, May 2005)

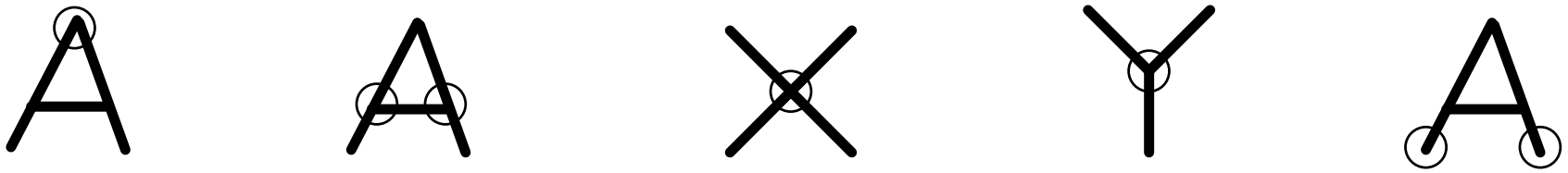
Confluence of electronics and biology

Example 1

Vision Sensor for the Detection of Image Features

Line orientations, line stops

Edges, corners, intersections:



(a) **Corner**

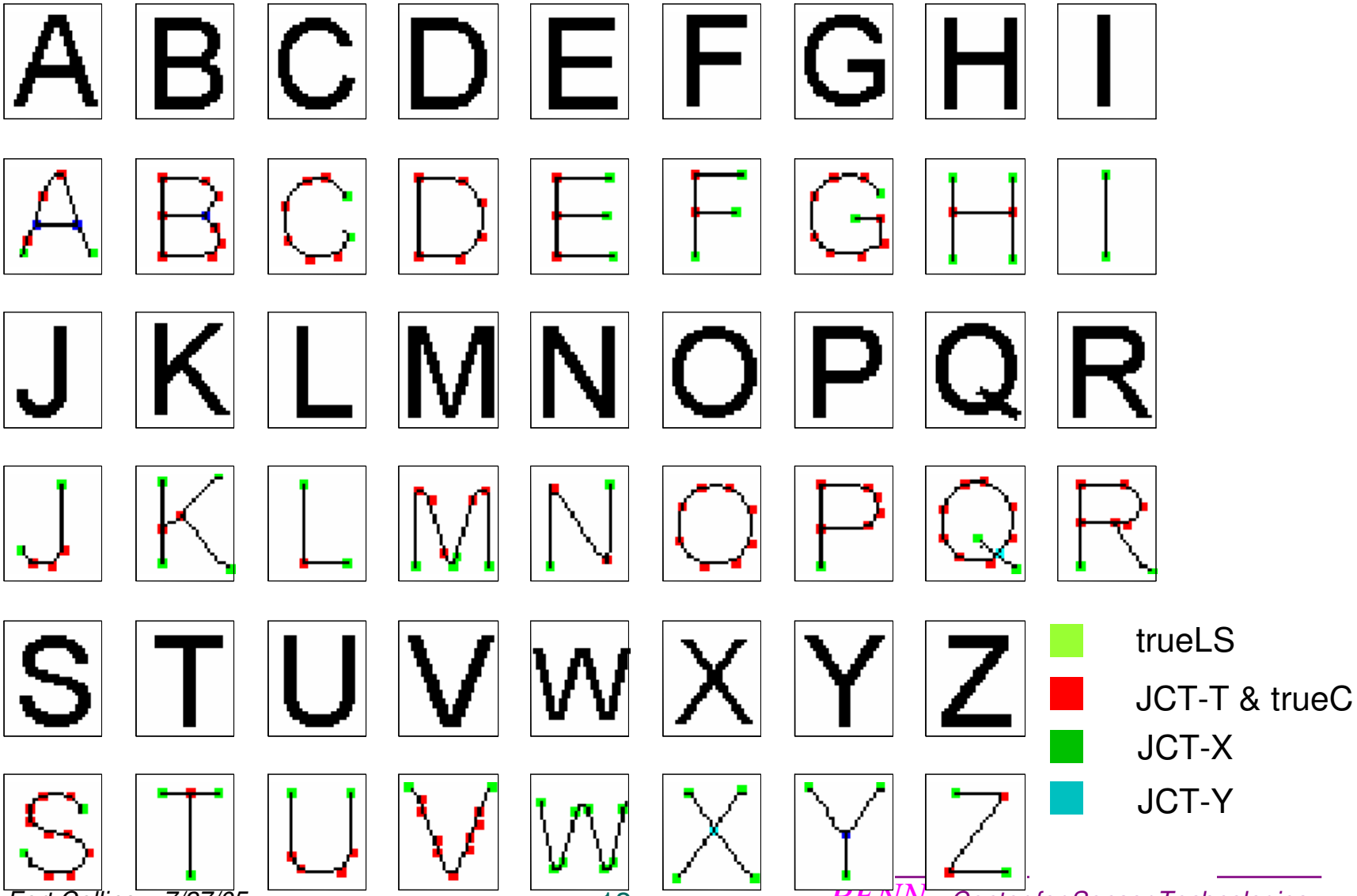
(b) **T-type**

(c) **X-type**

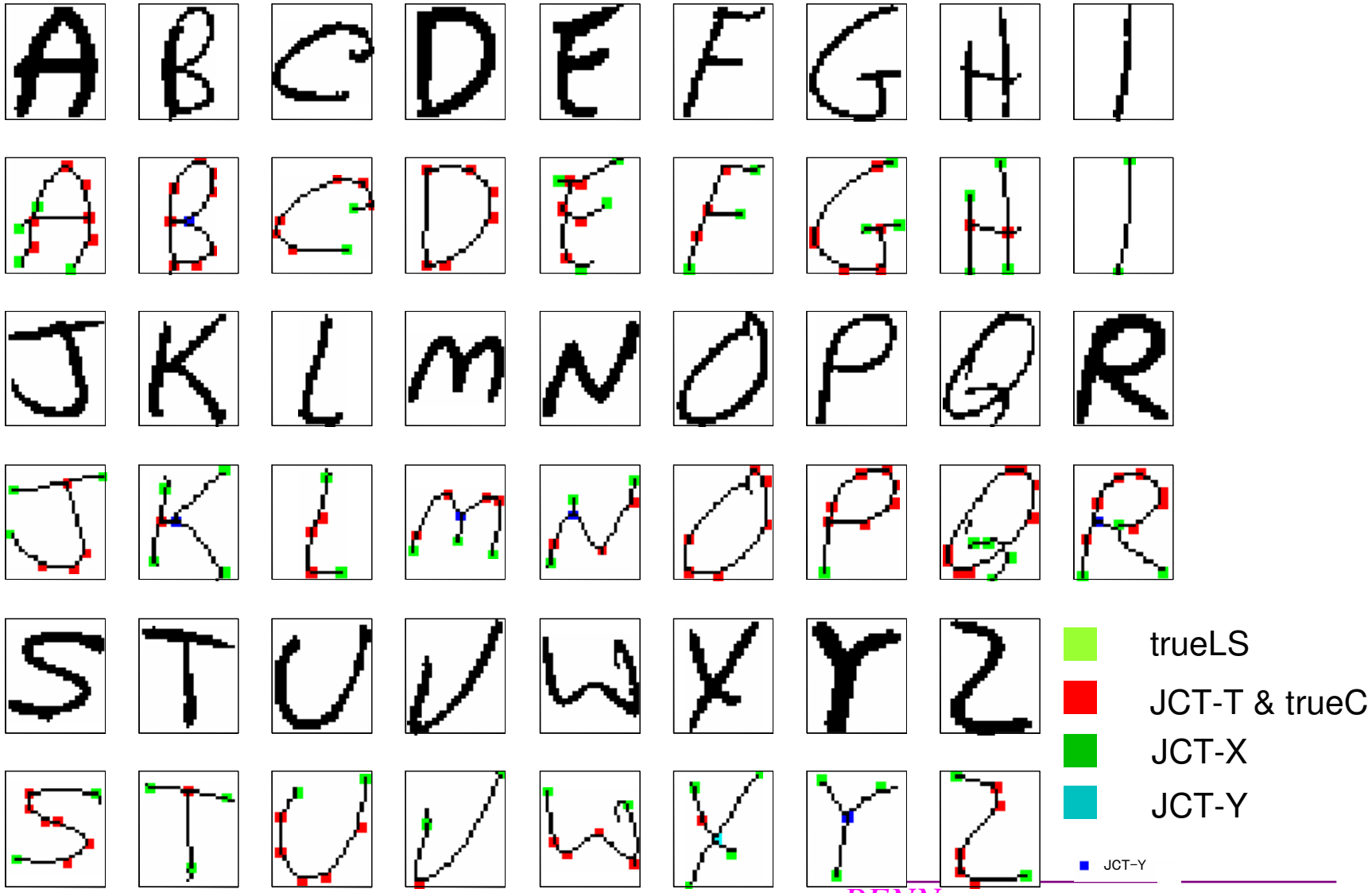
(d) **Y-type**

(e) **Line stops**

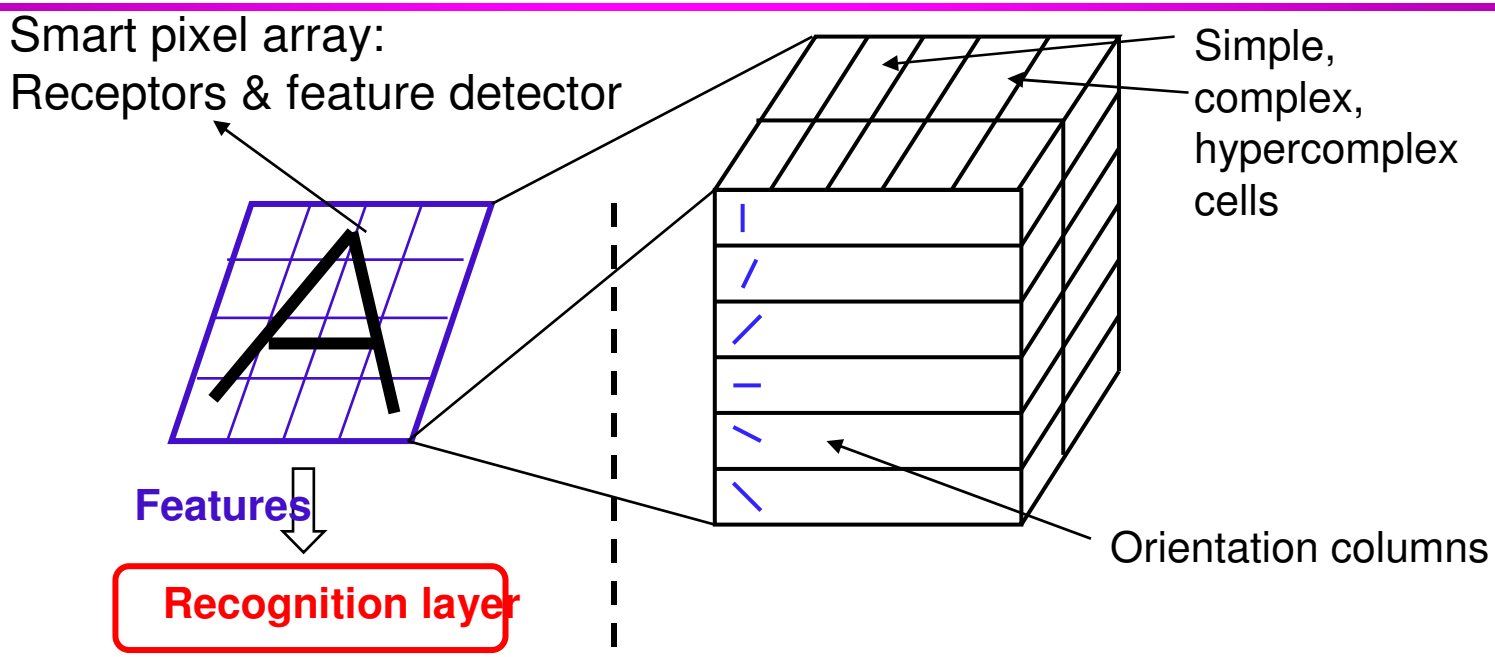
Detected features for printed characters



Detected features for handwritten characters



Retinal orientation sensor



Proposed approach:

- Template matching
- Programmable templates are executed in sequence

Trade-off between speed and area to reduce the number of processing elements.

Biological approach for feature detection:

- Multiple receptive fields
- Massively parallel execution
- Requires many processing element (cells)

Implementation considerations

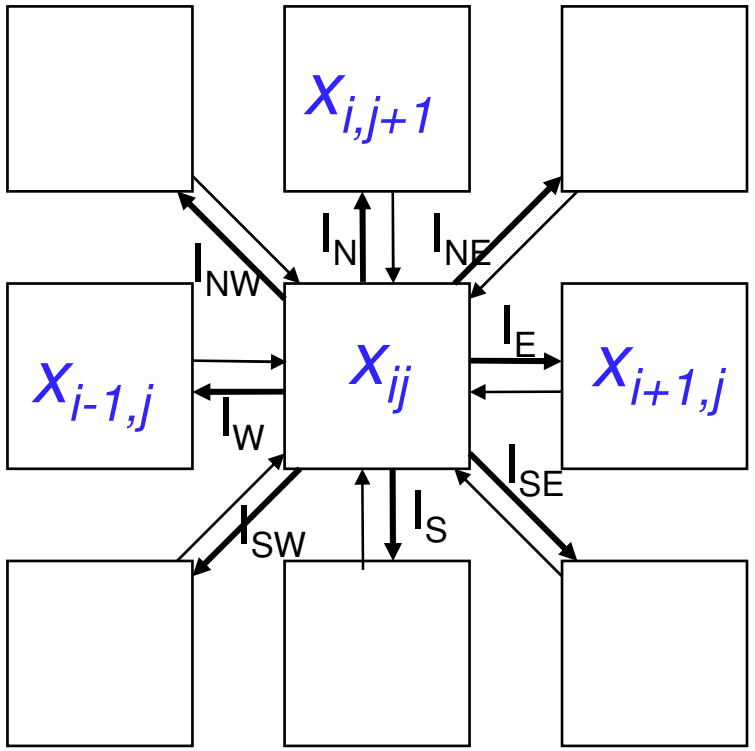
Template matching (or convolution)

$$x_{ij}(n+1) = f\left(\sum_{l=-1}^{1}\sum_{m=-1}^{1} x_{i+l\,j+m}(n)\, r_{lm}, I\right)$$

$$= f\left(\sum_{l=-1}^{1}\sum_{m=-1}^{1} x_{i-l\,j-m}(n)\, k_{lm}, I\right)$$

$f(x;I) = 1$ if $x \geq I$
$= 0$ of $x < I$

I: Threshold



Template:

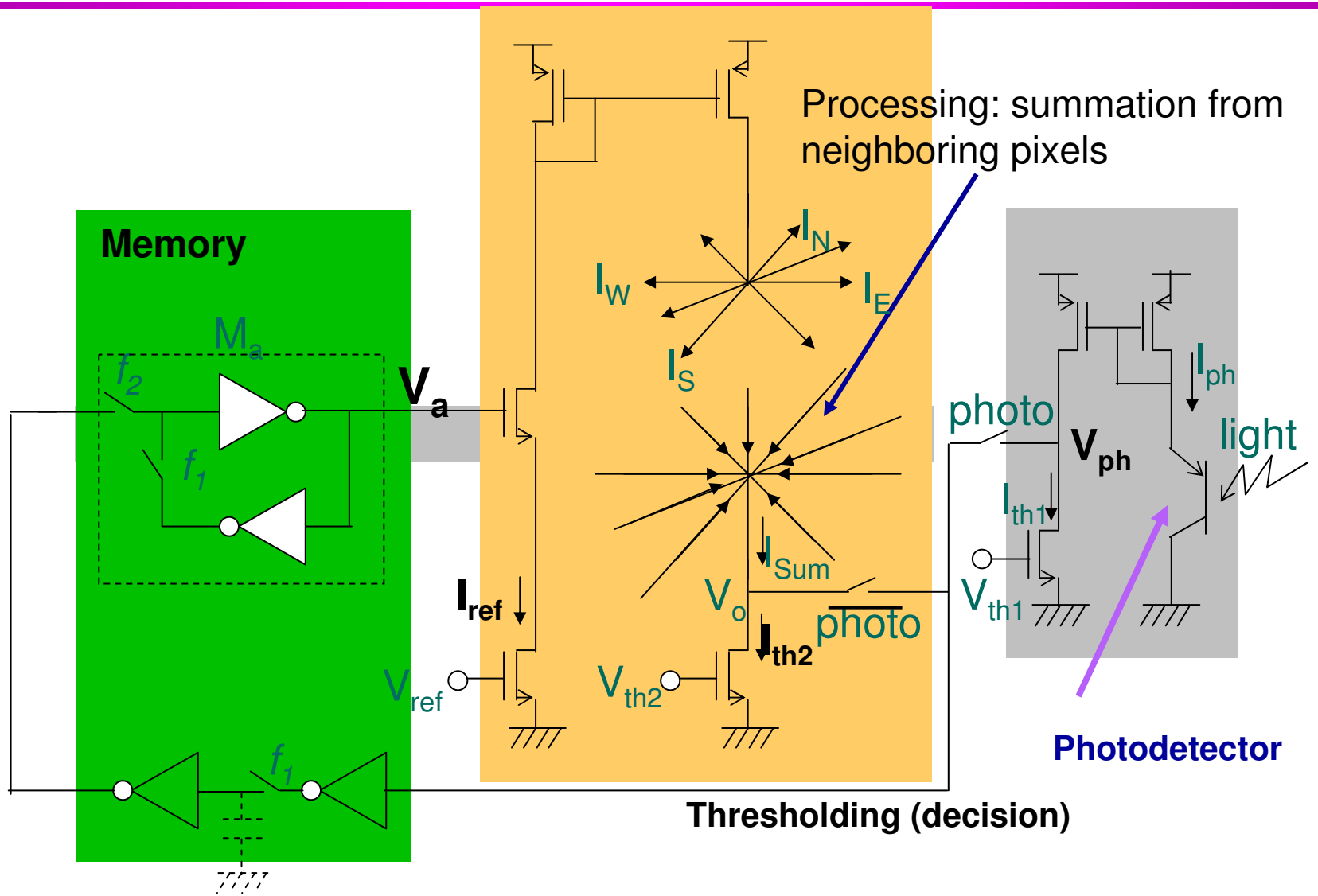
$$K = \begin{pmatrix} r_{11} & r_{10} & r_{1-1} \\ r_{01} & r_{00} & r_{0-1} \\ r_{-11} & r_{-10} & r_{-1-1} \end{pmatrix}$$

$$r_{ij} = 1, -1, 0$$

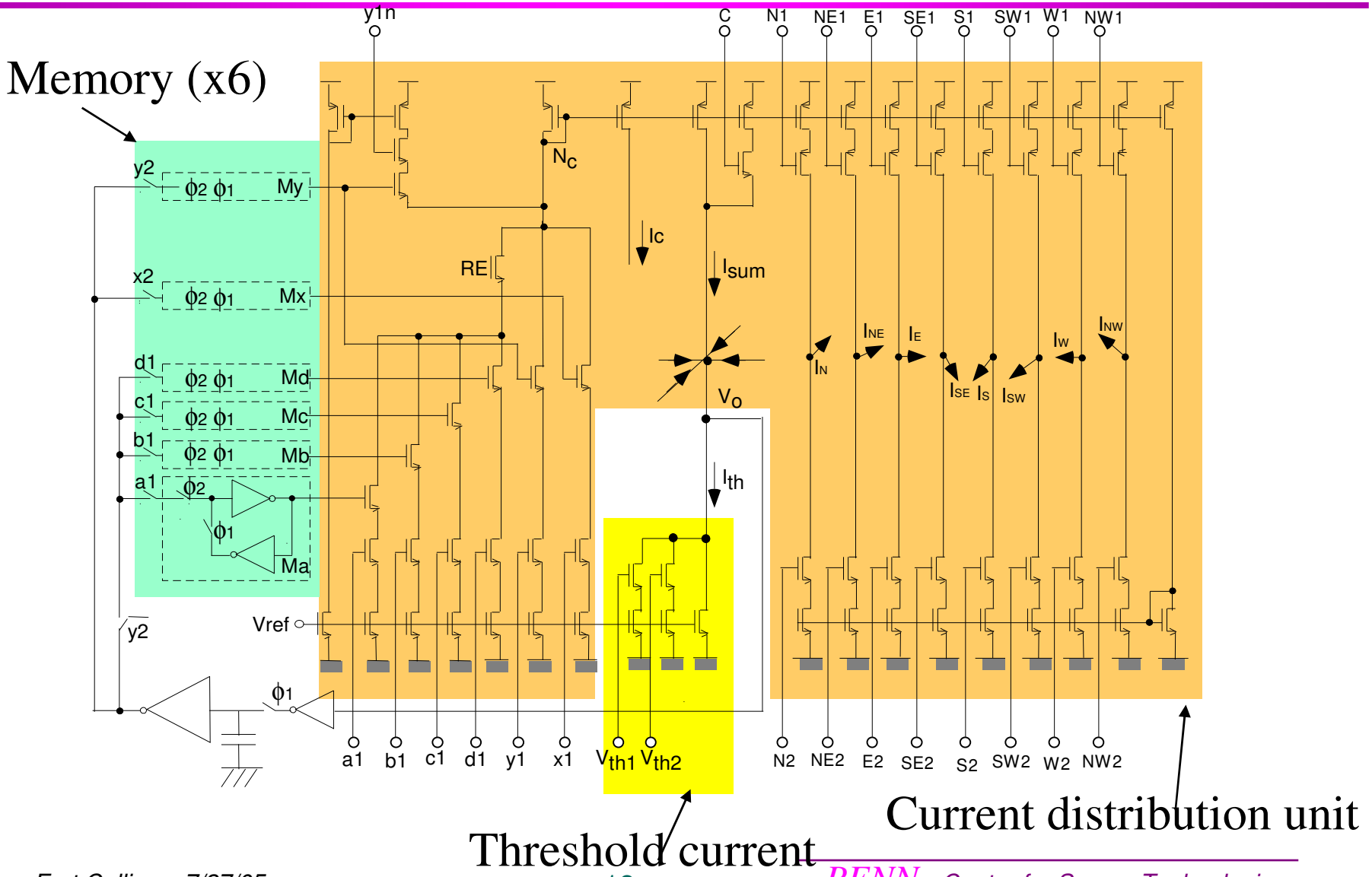
can be naturally mapped onto hardware by current distribution

Template is made programmable to detect a variety of features and perform a set of operation

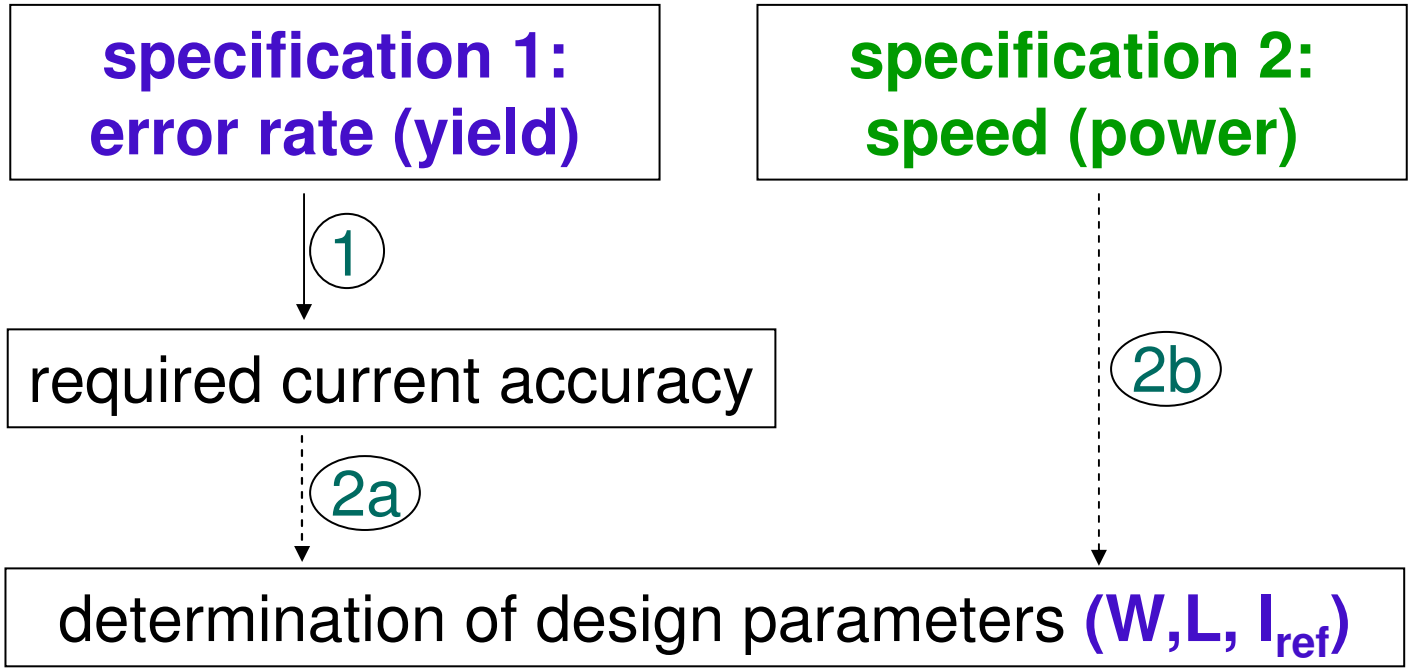
Conceptual pixel architecture



Schematic of the processing circuit



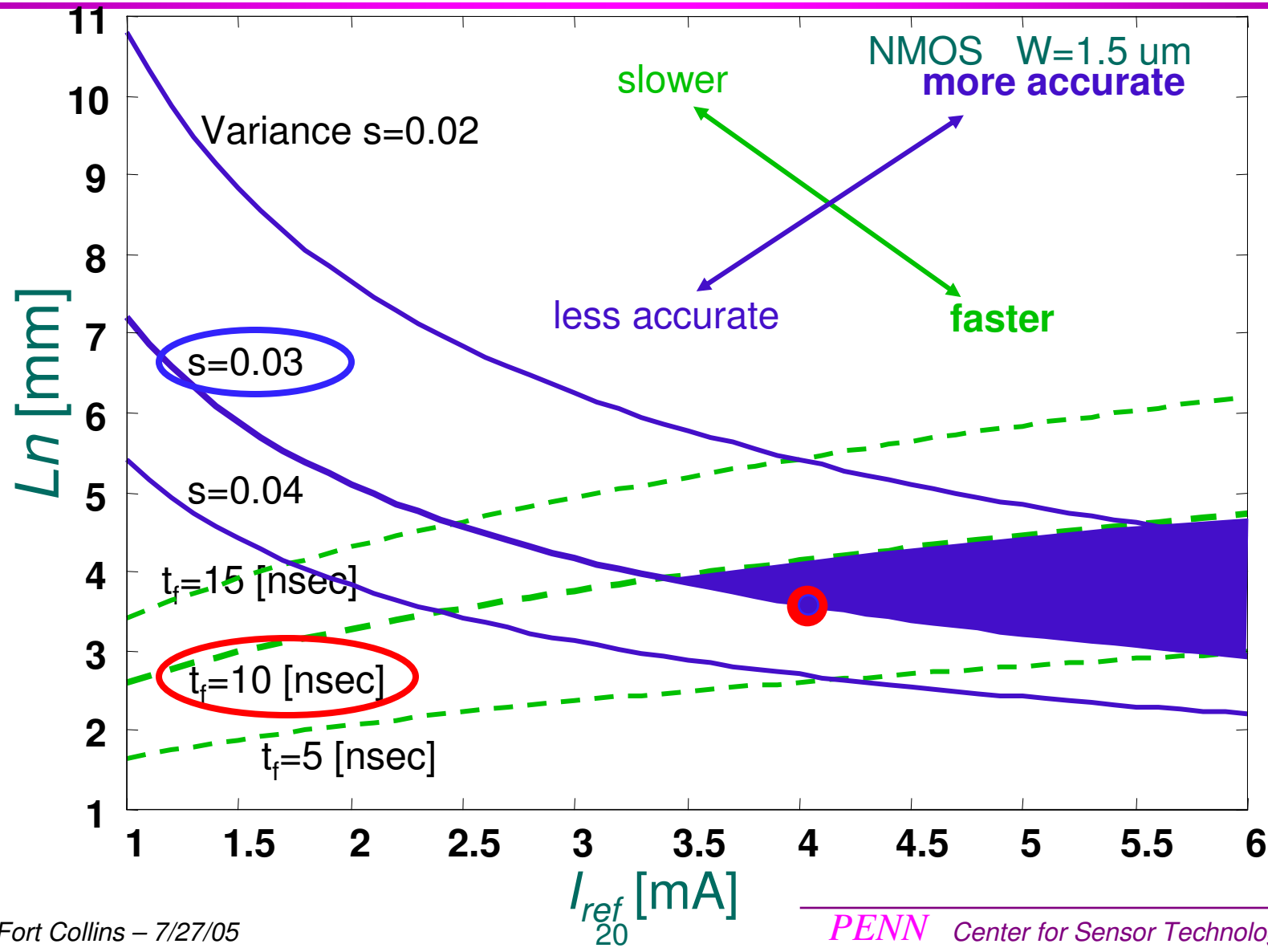
Design procedure based on transistor mismatch analysis



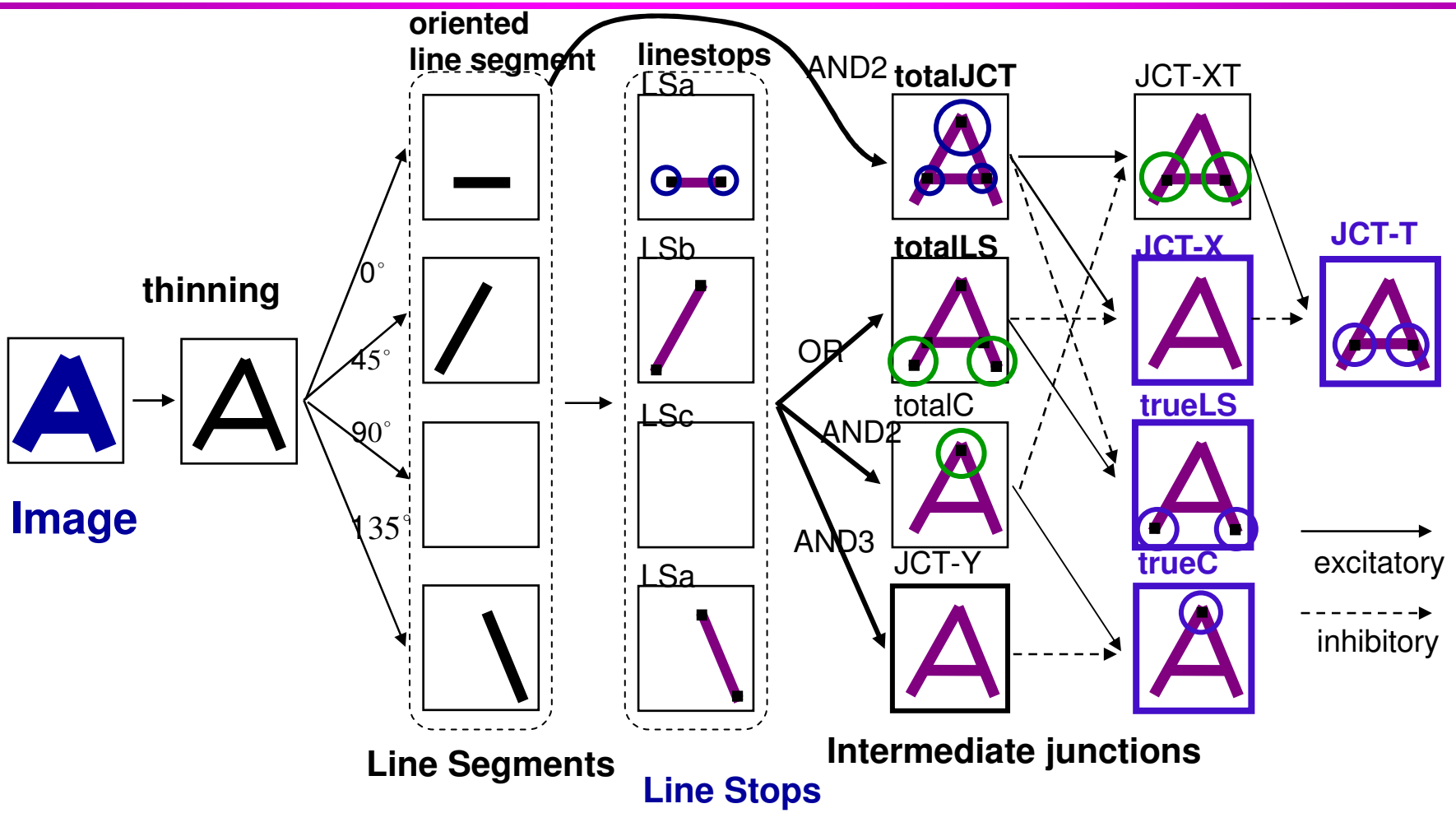
Current variation is function of:

- design parameters (W,L,I_{ref})
- mirroring operation
- summation and subtraction operations

Speed and Accuracy as a function of the reference current

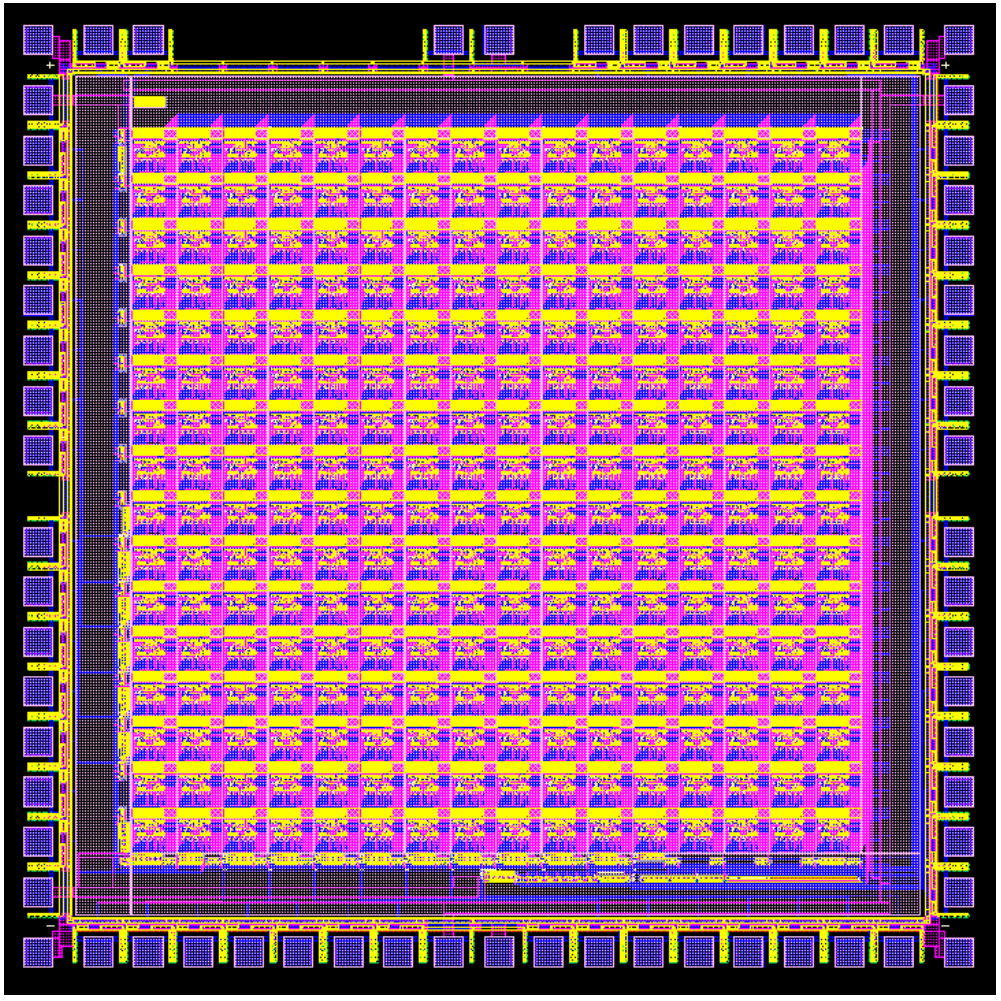


Processing Flow on the Sensor



Additional operations: Line completion ; elimination of isolated points; line elongation

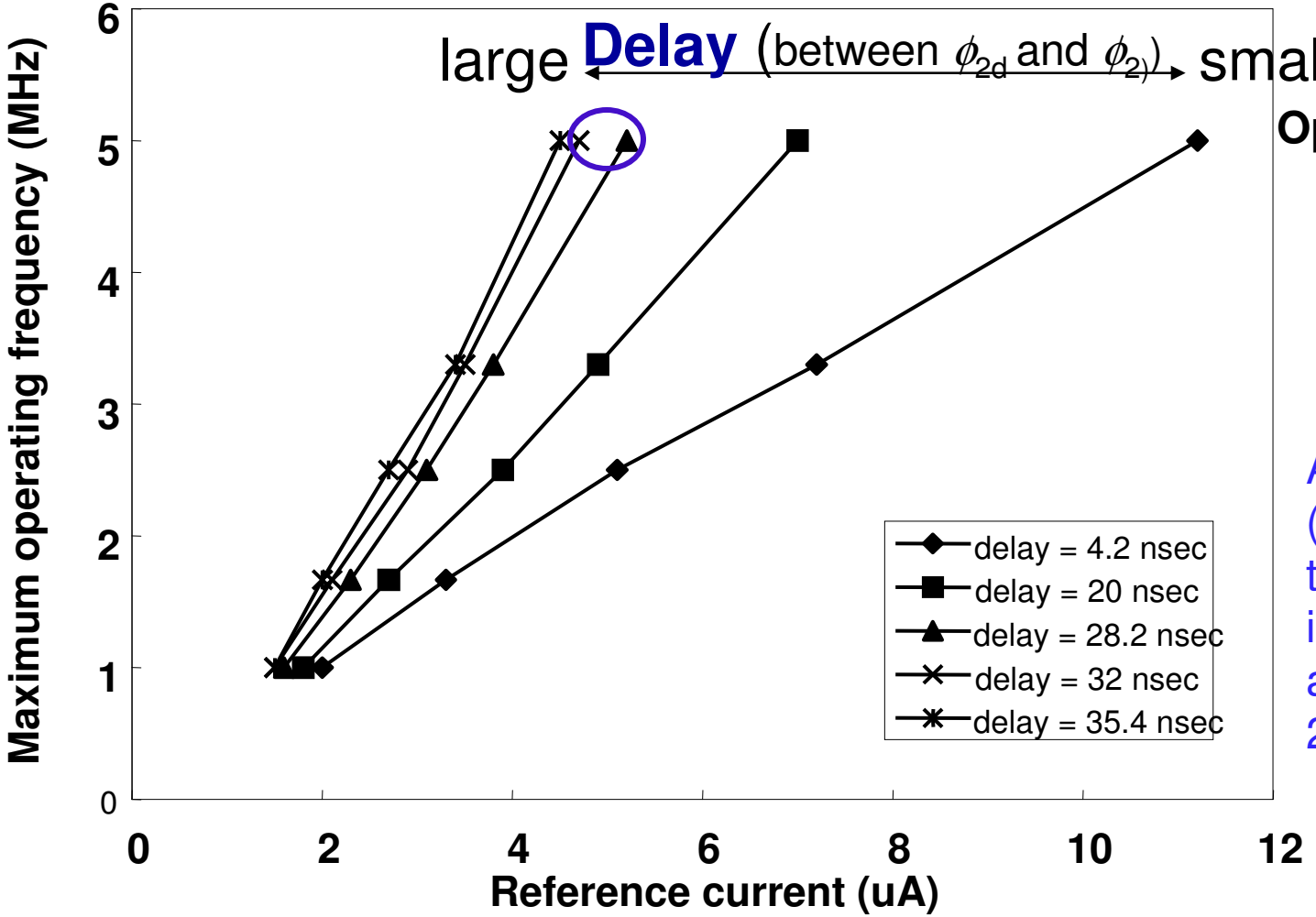
Prototype Test Chip



- technology: HP CMOS 0.5µm (3 metal 1 poly)
- chip area: 3.2mm x 3.2mm
- pixel number: 16x16
- pixel area :154.5 µm x 153.3 µm
- number of transistors: 147tr/pixel
- fill factor: 12.5 %
- Each pixel is programmable: 27 types of operations (30-bit word) involving up to 270 individual steps

Results

Maximum operating frequency as a function of the reference current



Operating conditions

V_{DD} = 4 [V] and

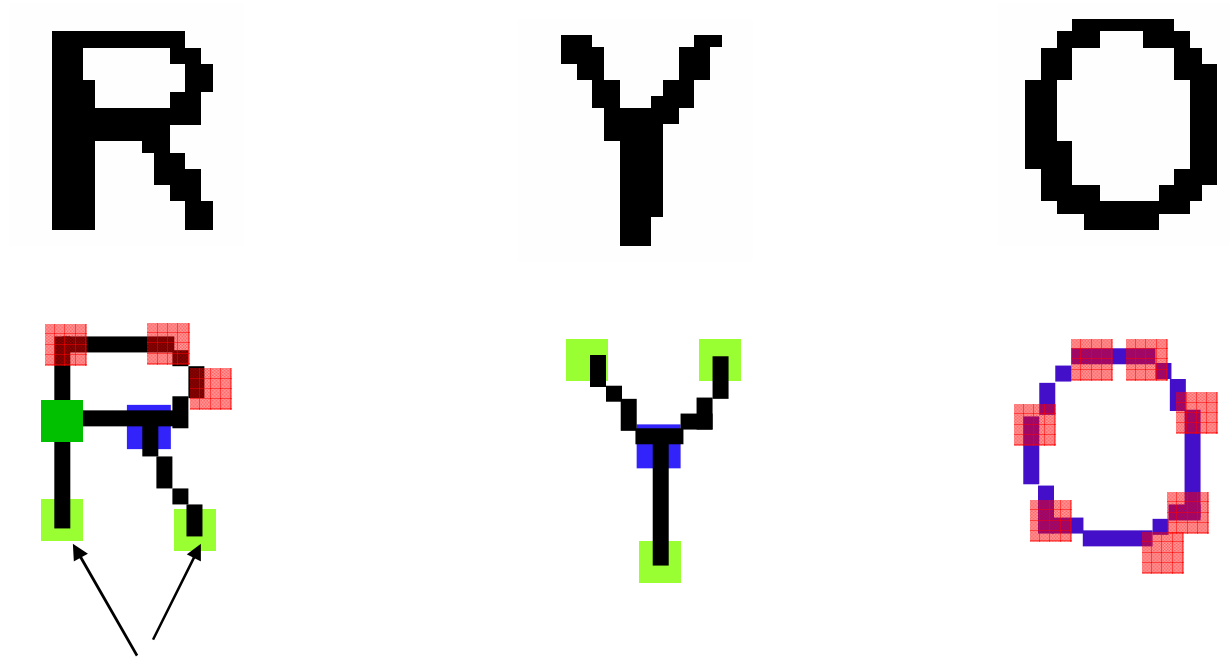
I_{ref} = 4.5 [μA]

100 mW

5 MHz

An operation (template matching, thresholding, storing in memory) over full array takes only 200ns.

Sensor responses to letter images



True LineStops

Corner

T-JCT

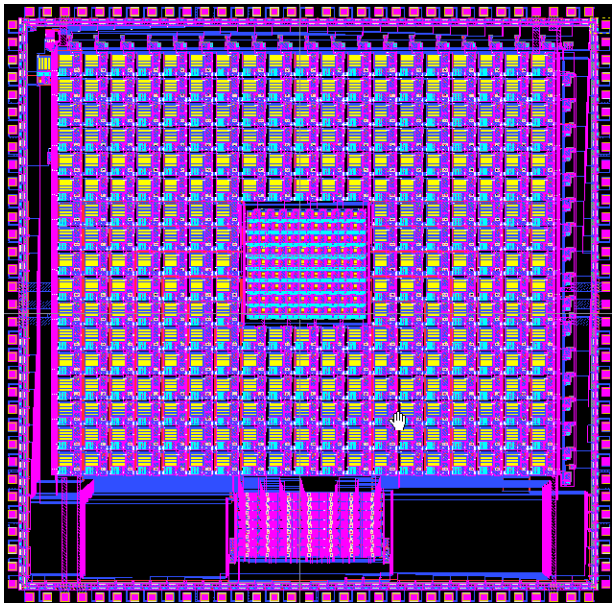
JCT-Y

JCT-X

Processing time @ 5MHz:
54.2 µsec (271 steps)

Example 2

Silicon Retina for 2-D Tracking

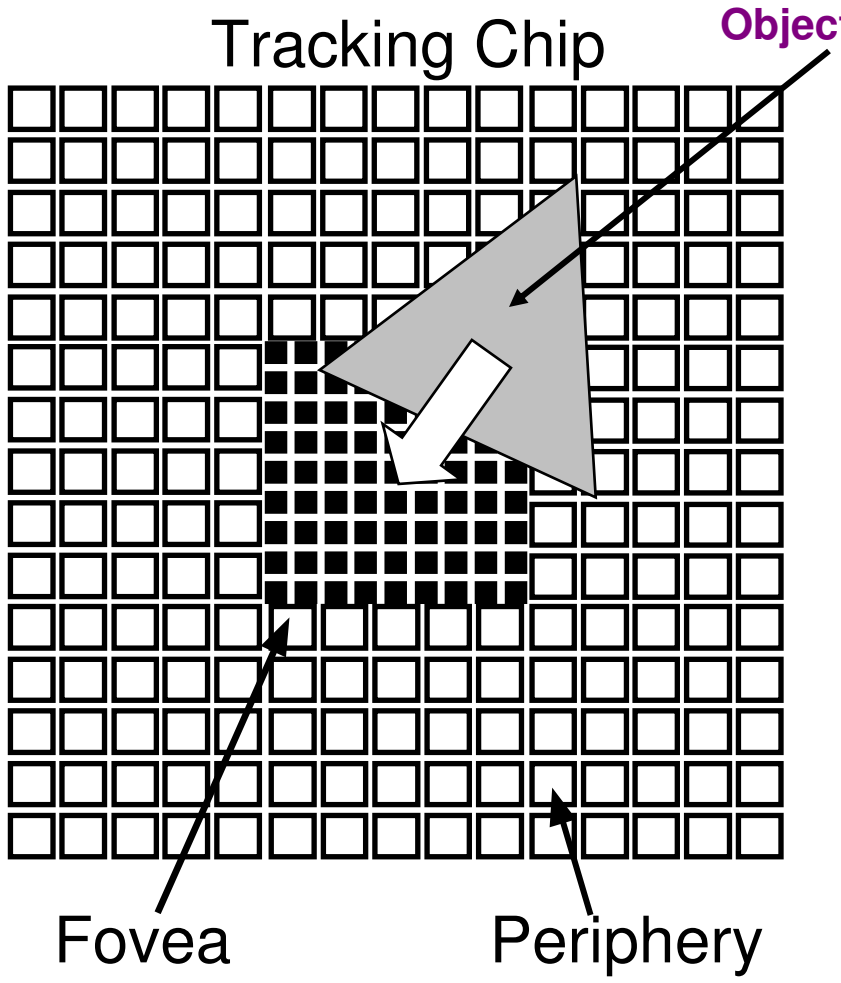


(Ref: R. Etienne, J. Van der Spiegel, P. Mueller, M. Zhang, IEEE CAS II, June 2000)

Proposed approach

- **Loosely modeled after the primate oculomotor system:**
  - Retinal photoreceptor organization
  - Retinal photosensing and early processing
  - Visual cortex for smooth pursuit
- Superior colliculus for saccadic generation
- **Capture the *functions*** found in biology and use the most efficient way to implement it using hardware (vs. wet ware)

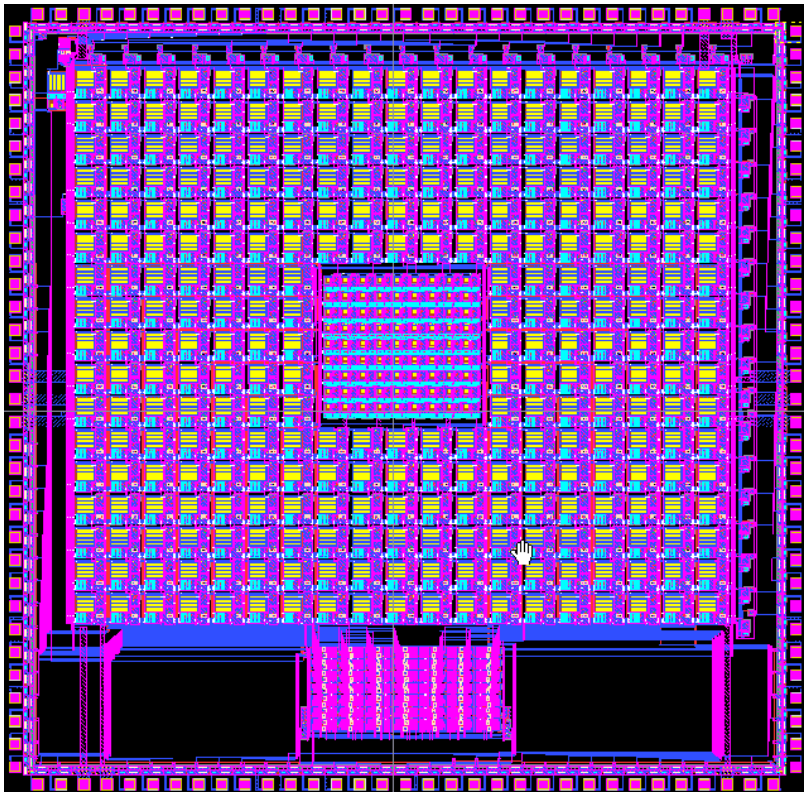
Tracking Chip Architecture



- **Keep object centered on fovea**
- **Fovea**: smooth pursuit
- **Periphery**:
  - localization
  - saccadic generation
- **Interaction** fovea-periphery
- Select target based on **motion**

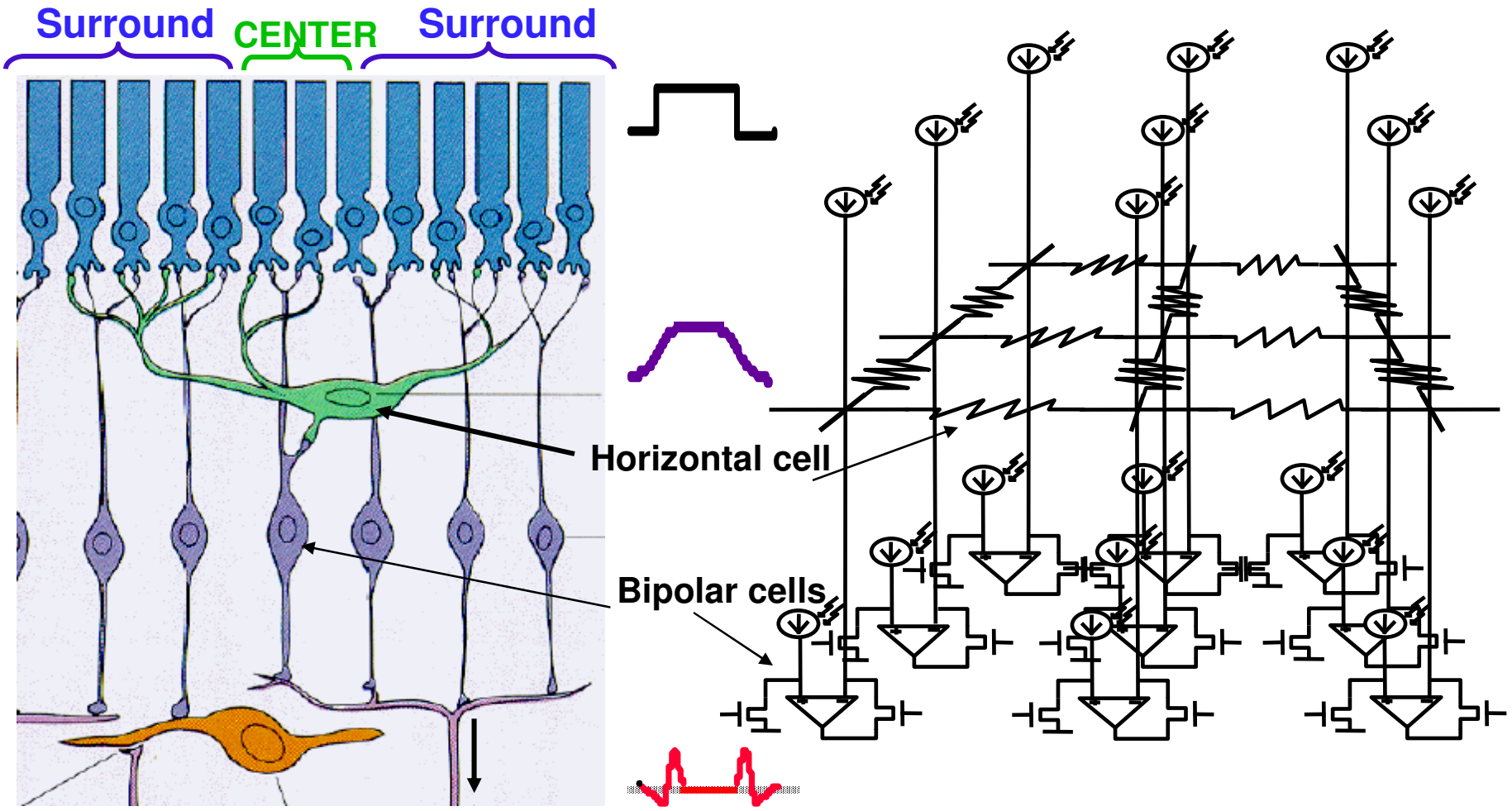
(After R. Etienne, J. Van der Spiegel, Mueller, M. Zhang, ISSCC 1997 and IEEE CAS II, June 2000)

Tracking Chip



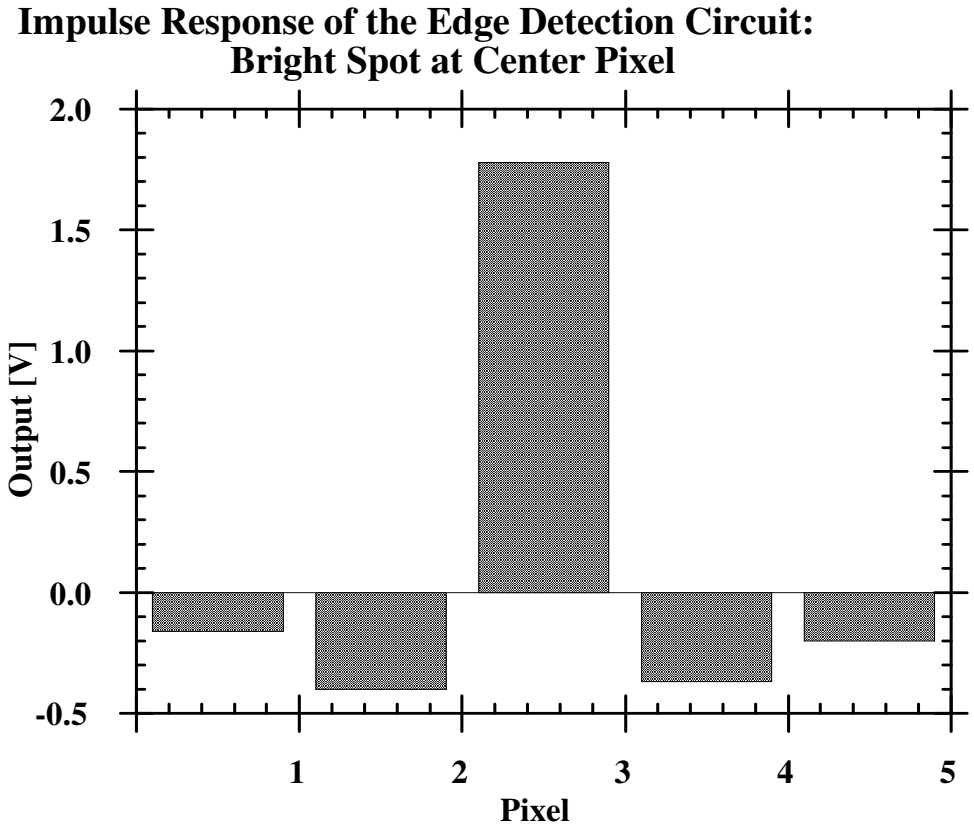
- Fovea: 9x9 cells
- Periphery: 19x17
- 2µm CMOS
- 6.4x6.8 mm^2

Retina with edge detection

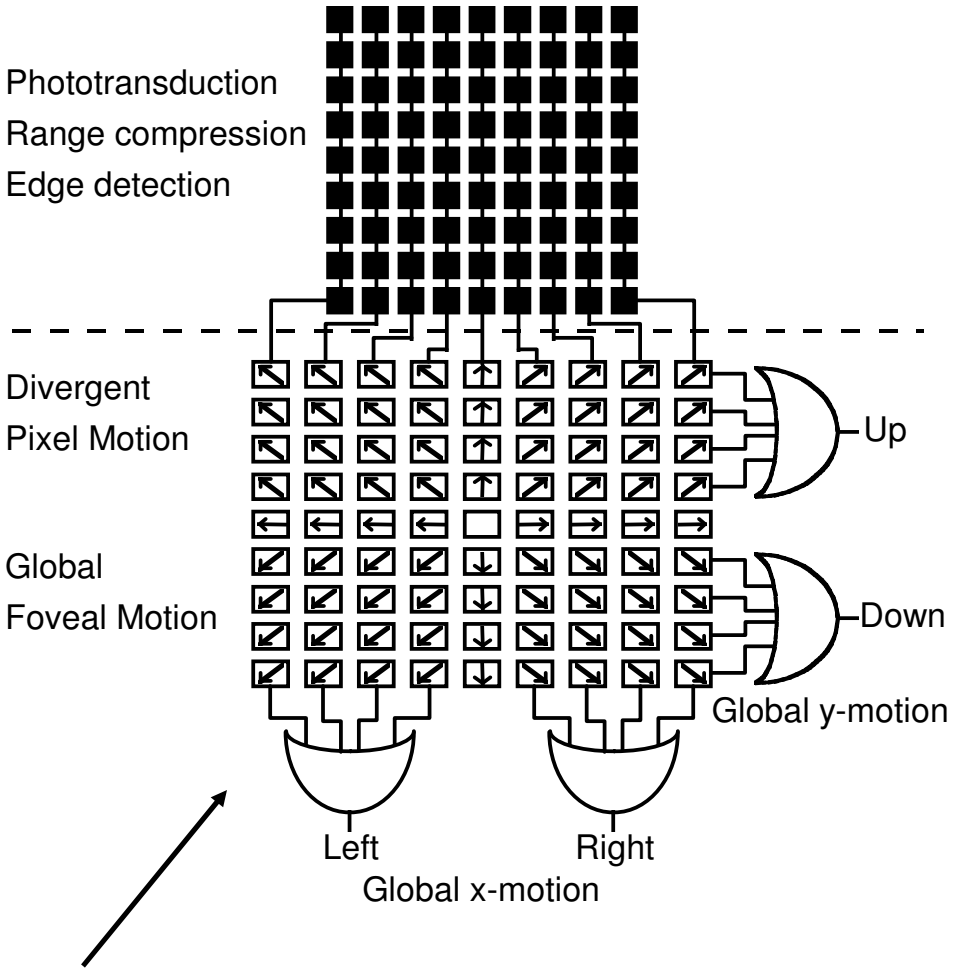


(after: www.ini.unizh.ch/avlsi/)

Result of edge detector

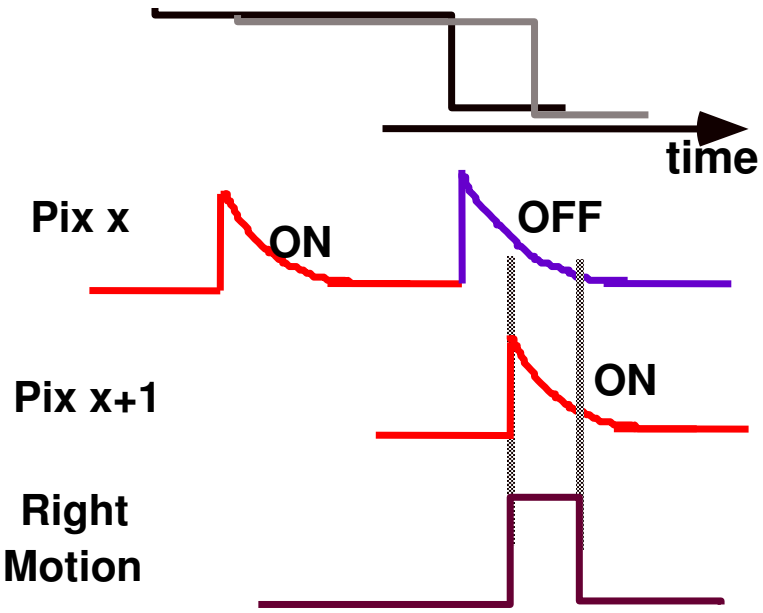


Motion detection in the fovea: smooth pursuit

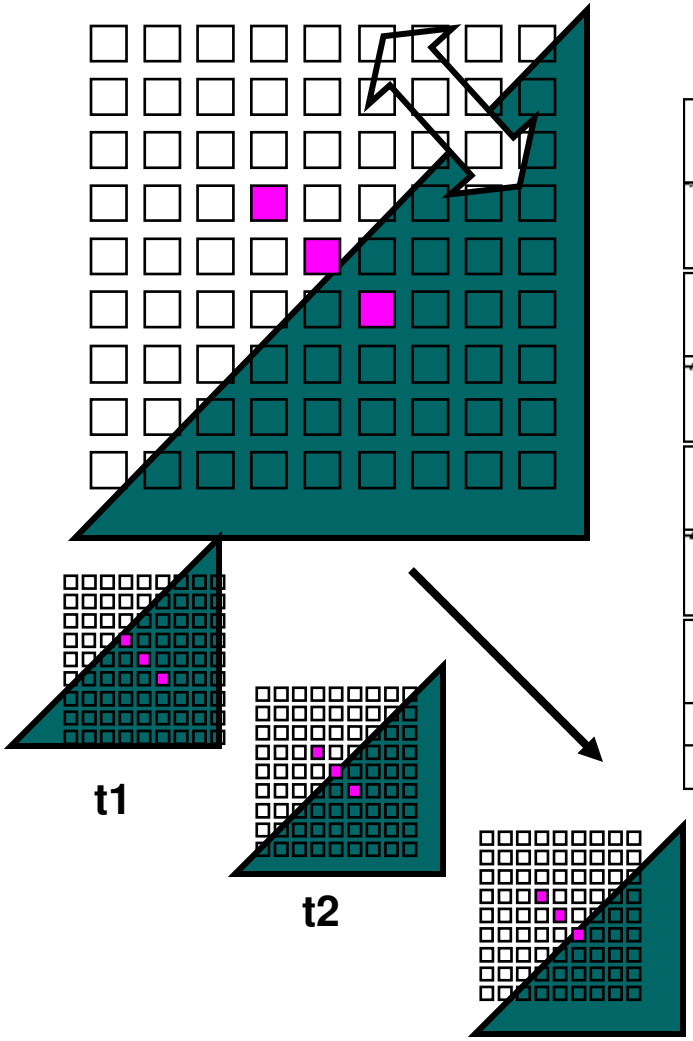


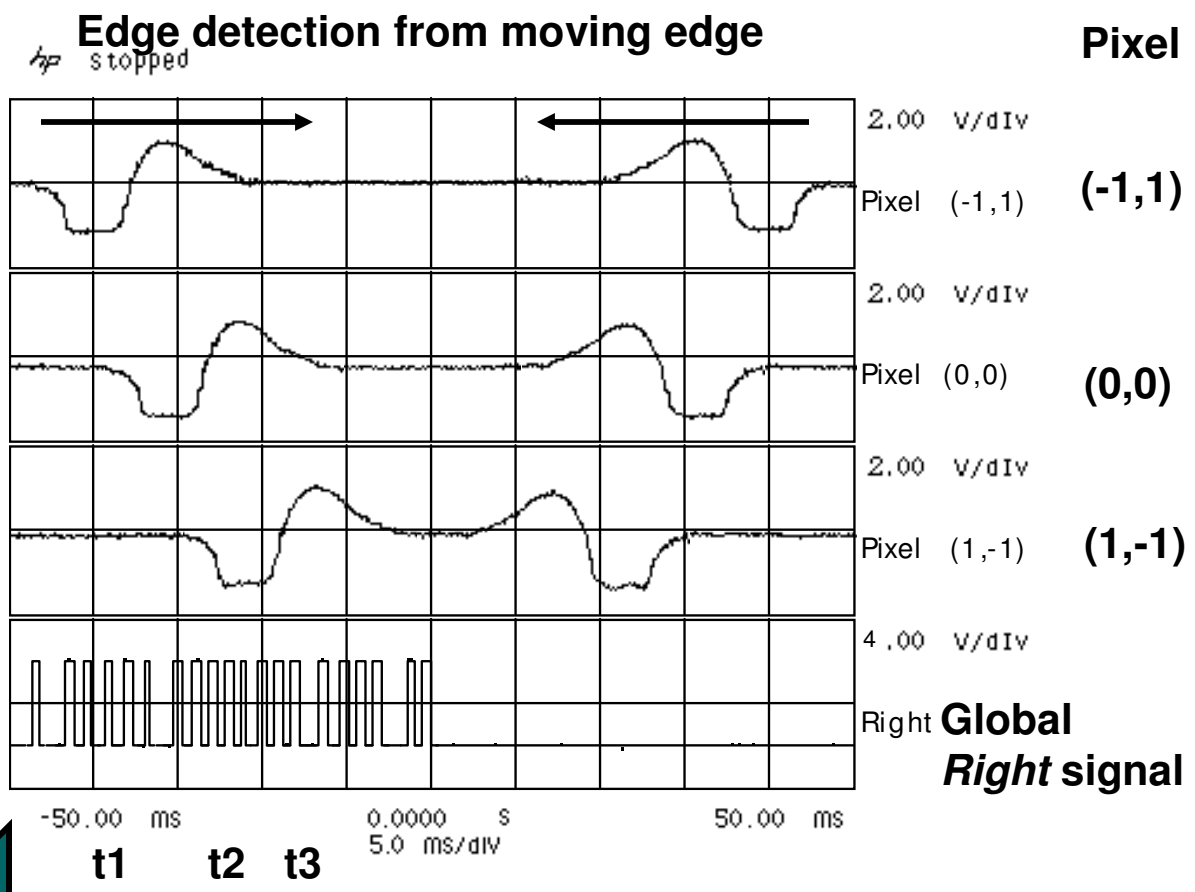
Performed outside array

- **Correlation** based
- X-motion (conceptual):

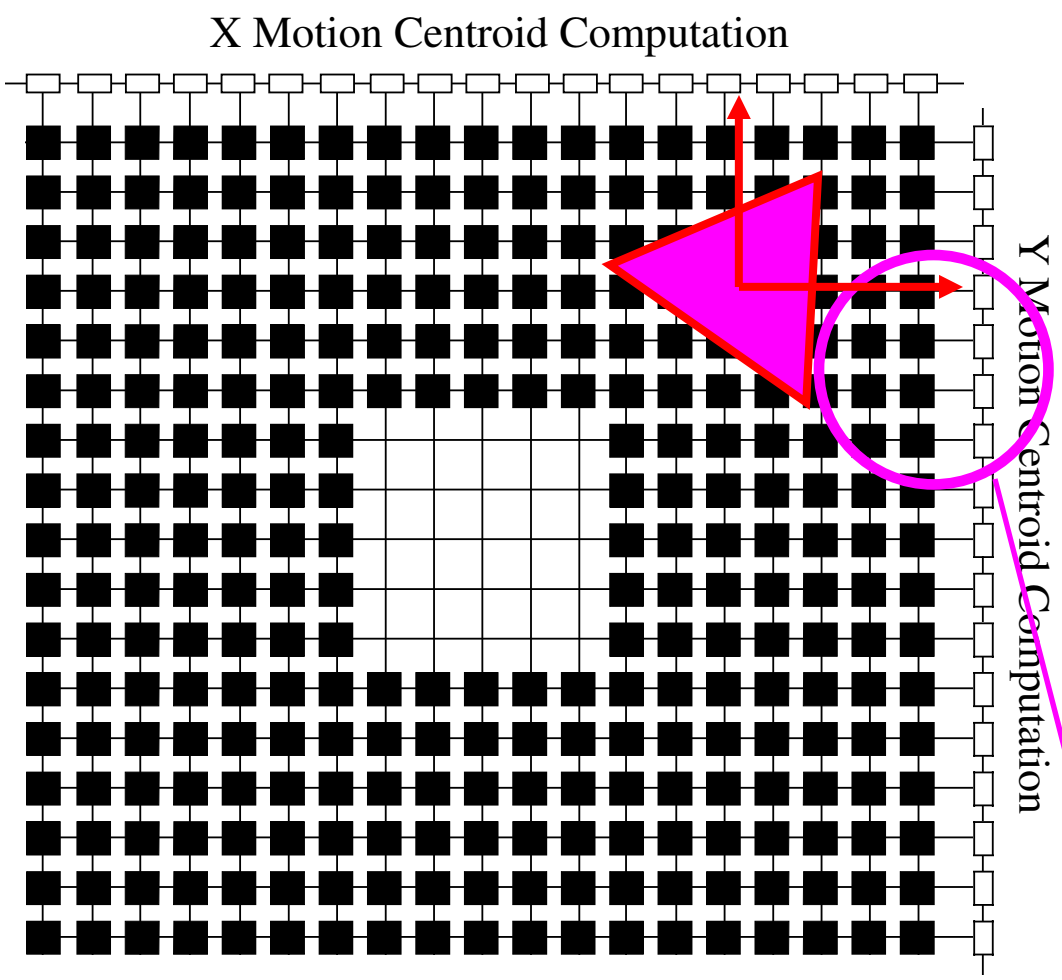


Measurement of the Foveal Motion detection



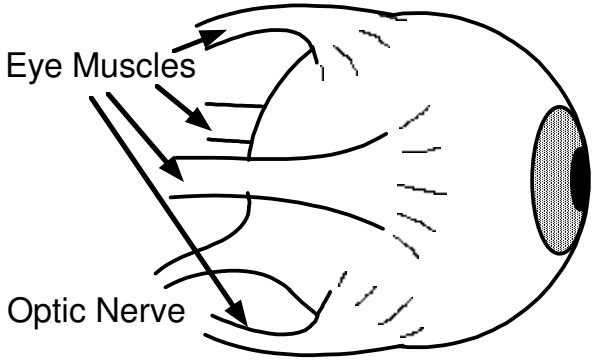


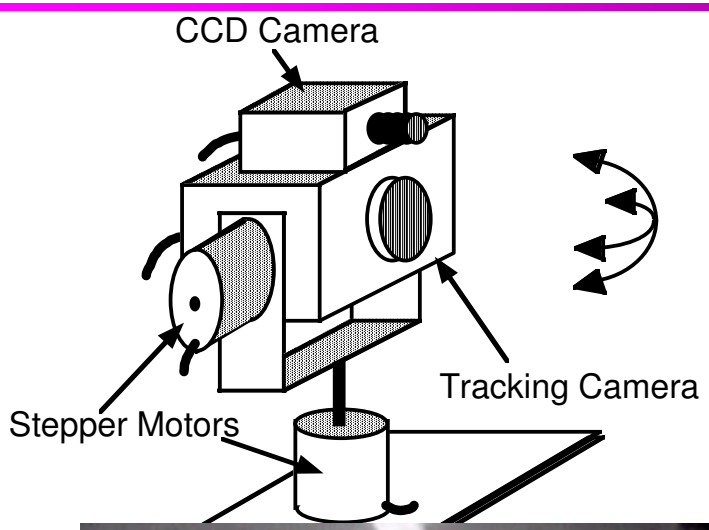
Periphery: Target Acquisition



- Lower resolution
- Edge-detection
- ON-set detection (temporal differentation but No motion detection)
- Localization of centroid in relatation to the spatio-temporal boundaries
- Row & column labeling: X and Y

Tracking Experiments





Conclusions

Summary

- **Biological systems provide a viable paradigm for building vision sensors:**
  - Compact, low power, robust under different conditions
  - Massively parallel pixel-level processing
- **Two Vision sensors:**
  - Higher level features (X-, Y, and T-type) : Incorporates processing functions found in area V1
  - Target tracking: space variant (foveal and periphery) for smooth pursuit and saccadic motion generation
  - Implements the **functions** and **algorithms** of biology.
  - Optimized for **information extraction**, not image rendering
  - Limitations: limited resolution, large pixel size and small fill factor.

Thank you