Biological Neural Processing as a Paradigm for Image Pattern Recognition

by Brad Morantz PhD

My Perspective

- I am a decision scientist
  - My goal is to decide what I am seeing
  - I do not care about programming language
    - The hammer in tool box paradigm
    - A story can be told in 200 languages
- Decision Science uses data and other skills to solve problems & make high quality decisions

Please Remember

- Don't only think of the primate visible light spectrum
  - Think about active or passive
- Image can be in many formats
- Image can be formed from many different kinds of sensors
- Logic and thought can affect an image
- A vision impaired (blind) person can form an image in their mind
    - Story of 3 blind men and an elephant
  - A thorough presentation would take 2 semesters, we have 90 minutes

Recognition vs Identification

- Recognition is a subset of identification
  - Recognition is the act or process of identification of a person or thing from previous encounters or knowledge; connecting previous stimuli with a new encounter
  - Identification is the act or process of naming or declaring someone or something
- The difference is that recognition is something where the stimuli have been previously encountered and this is not always true for identification.
- My work in cognitive identification

Question

- What is the major organ of vision?
- The Brain!

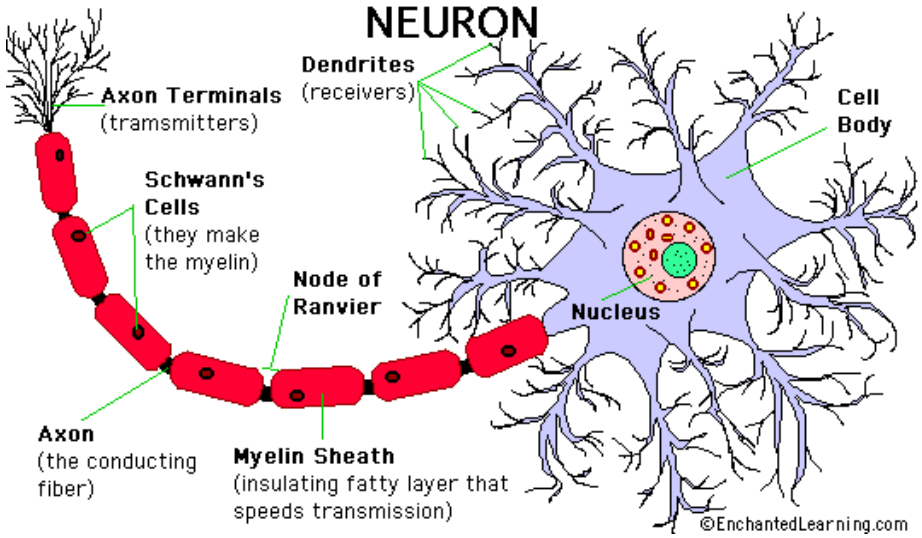
Overview

- What is a biological neural network?
- Why do we care?
- Visual Pattern Recognition in biological neural network
- Computer pattern recognition
- Comparison
- Steps forward

Biological Neural Network

- Brain of a living animal/CNS
- Massively parallel connection of neurons
  - Sea snail has 100
  - HomoSapien has from 4×10^{10} to 1×10^{11}
- 35+ neurotransmitter chemicals
  - Affect operation/performance
- Various glands and organs
- More is unknown than known
- “Clock Speed” is about 1 KHz

The Neuron



- Electrochemical
  - Signaling from diffusion of neurotransmitters
  - Excite or Inhibit
- Axon is output
- Dendrite is input
- Axon on dendrite creates synapse
  - Unidirectional junction
  - Electrochemical

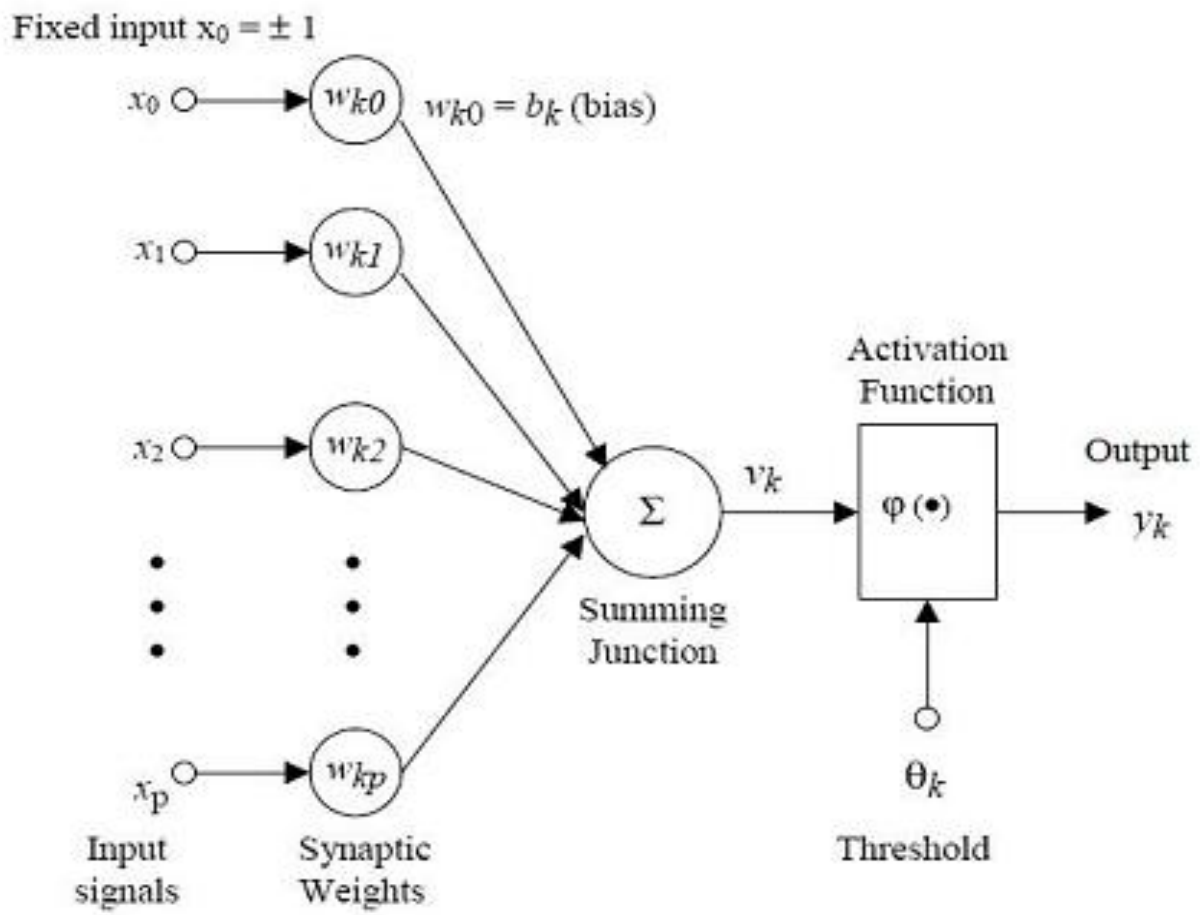
Neural Network

Non-linear
Parallel processing

If sum of weighted inputs exceeds the threshold, then the neuron fires via the activation function

This is just background information.

This presentation is NOT about neural networks

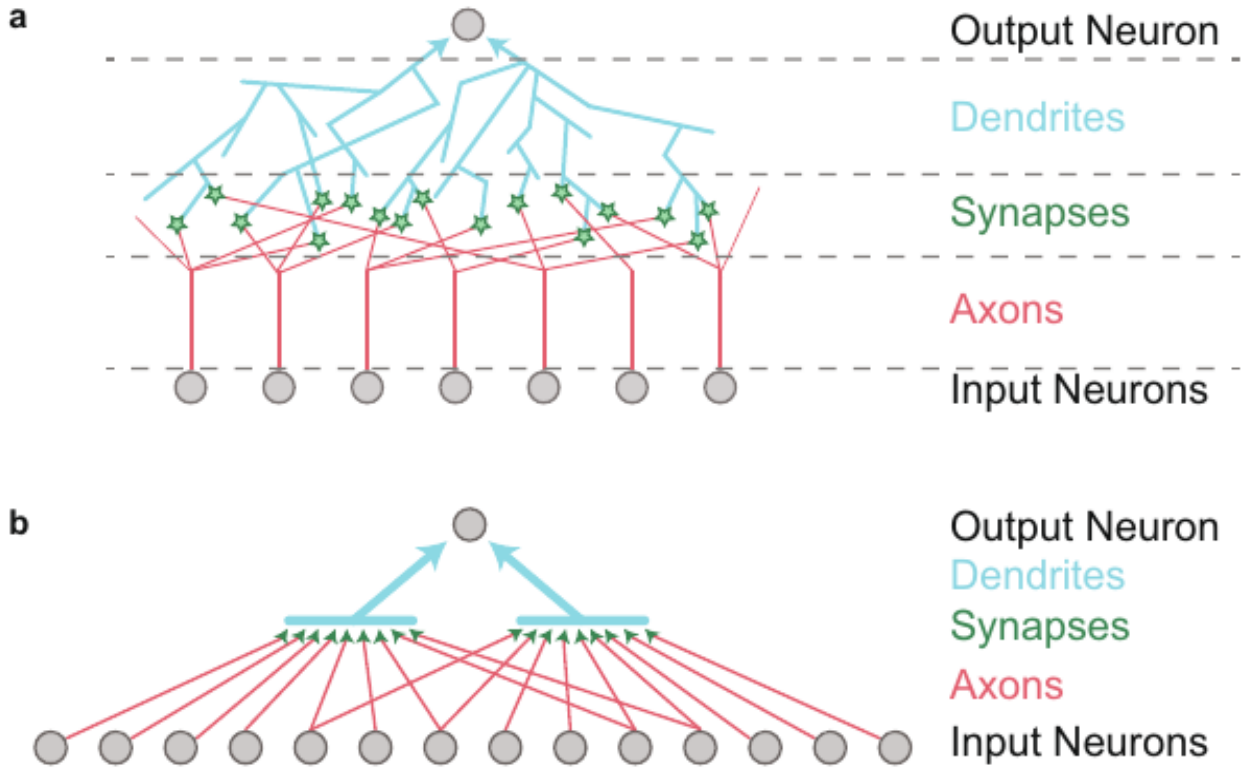


This is one layer of a neural network

Neural Models

- McCullough – Pitts
- Perceptron – Rosenblatt
- Hodgkins & Huxley
- Adaline – Widrow & Hoff
- Bar Ilan Model -Ido Kanter et al

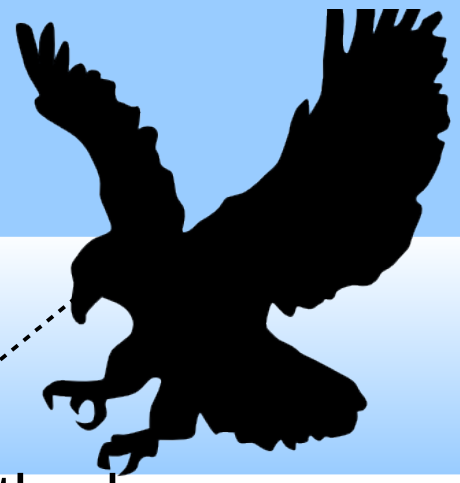
Bar Ilan Model - Kanter et al



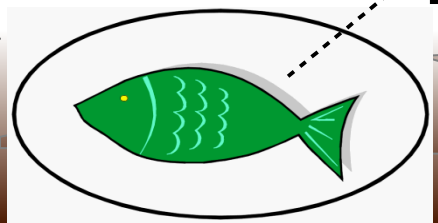
Hebbian

Bar Ilan Model

Performance



- Bat
  - Has brain the size of a plum
  - Can navigate through electric fan unscathed
  - Very powerful flight navigation system (echolocation)
- Frog
  - Very small brain
  - Can discriminate fast moving target insects
  - And catch them with its tongue
- Eagle can spot & identify target at long range
  - Can see fish from 500 ft up one mile away (20:4 Acuity)
  - Dive at 100 mph and catch it



Why Do We Care?

- Biological brain
  - Is smaller & less energy
  - Performs very fast
  - Works very well
- Can we learn from this to make our systems work better and faster?
- Can we use for recognizing target of interest?
- How does it adapt to changes over time?
- How does it adapt to varying aspect angles?
- Pattern generalizability

What If. . .

- We could emulate performance of bat's navigation
- Combine it with pattern recognition of frog and human and long range vision of eagle
- Put it all in a computer that ran 6 orders of magnitude faster than a biological processor
- Imagine . .
  - Navigation systems
  - Robotics systems
  - Computer vision applications

Image Pattern Recognition in a Biological Neural Network

The brain is the primary organ of vision

- The eyes are only the sensors

Vision is process of converting sensor information into knowledge of shape, identity, and configuration [F&F]

Very parallel process

A bat uses his hearing as sensors

- Active sensing (sends out a signal, then listens)

Pit viper uses infrared/heat sensing

Sensors can vary among systems

Pattern Recognition

How objects in the environment are identified is a basic description of pattern recognition.

In order to recognize a pattern a set of patterns must already exist in long term memory to which the current one can be compared

There is much variation in the object that is trying to be recognized from many things such as lighting, perspective, coloration, sound, background, timing, environment, and variability in the object or pattern itself

Perceptual generalization allows the recognition of such.

Pattern Generalizability

- Seeing the same thing as before but:
  - The lighting is different
  - Closer or further away from it
  - Turned a little bit, different angle
  - A person could be wearing different clothes
  - There are some minor changes
- Pattern generalizability allows one to recognize in spite of minor variations

Perception Model

Not the result of a set of patterns, but rather an interpretation of sensory data based on past experiences.

Theory is that the brain performs Bayesian style inference and integrates different sensory information to form a perception of the world

We need to extract & manipulate as we rarely see the exact same thing twice.

Creates need for pattern generalizability

Example

Ears: load roar

Eyes: Rectangle w/rectangles on side, Triangle at end

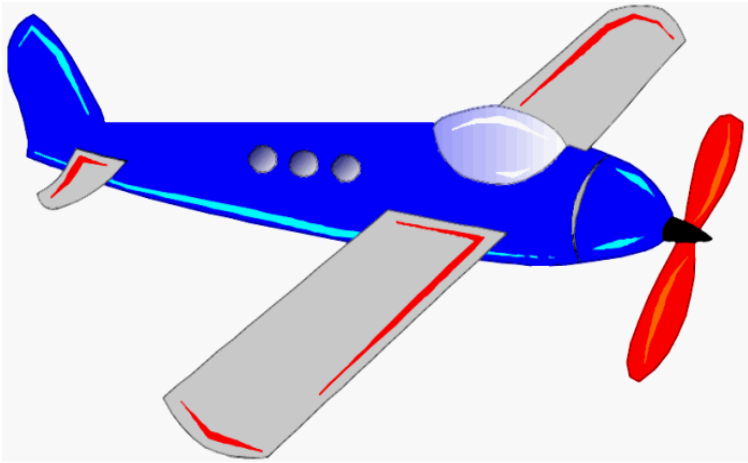
Eyes: it is in the air

Logic: it is flying

Reasoning

Memory

Answer: airplane



Frog Vision

- Four different neural structures
  1. Small dark moving approximately circular object
     a) Bug (food) detector
  2. Edge detector
  3. Moving contrast detector
  4. Dimming detector for when illumination dims
- Ganglion cell in eye is part of logic
  - Effectively distributed computation
- In lower organisms, vision is typically goal specific

Sensors

- A bat uses hearing as sensors
  - Active sensing (sends out signal, then listens)
- Pit viper uses infrared/heat senses (passive)
  - Can differentiate 200 millikelvin
- Sensors can vary among systems
- Eyes
- Ears
- Nose
- Feel/Touch
- Radar, sonar, AF, IR, EO, tHz, chemical (spectroscopy), etc, etc

Compund vs Camera Eye

- Compund eye
  - 1000's of facets per eye
    - Each has its own lens, set of nerves, and fiber
    - Each pointing slightly differently
  - Great for detecting motion
  - Think of large Phase Array Radar
- Camera eye
  - One lens per eye
    - High resolution
    - Color detection
  - Since we are discussing image recognition, we will only discuss this type

Human Vision

- Saccadic vision
  - Peak angular speed of 1000 degrees per second
  - Can maintain for 20 to 200 milliseconds
  - Increases visual resolution and reduces blurring
  - Microsaccades
    - 20 Arcseconds excursion
    - At 60 Hz (because frequency of artificial light?)
    - Refreshes and keeps image
- Saccadic movement
  - Like SuperResolution
  - Like Synthetic Aperture Radar (SAR) and Synthetic Aperture Sonar (SAS)

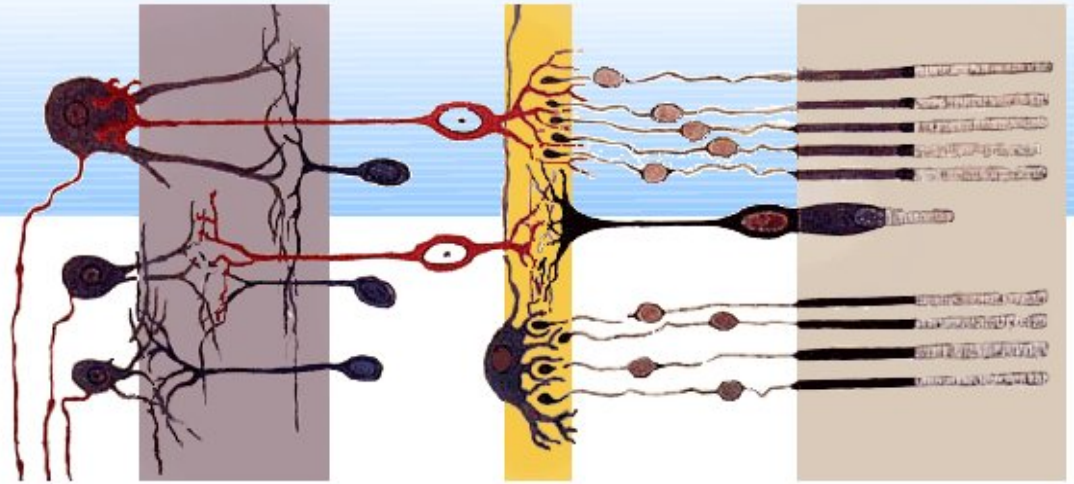
Human Eye

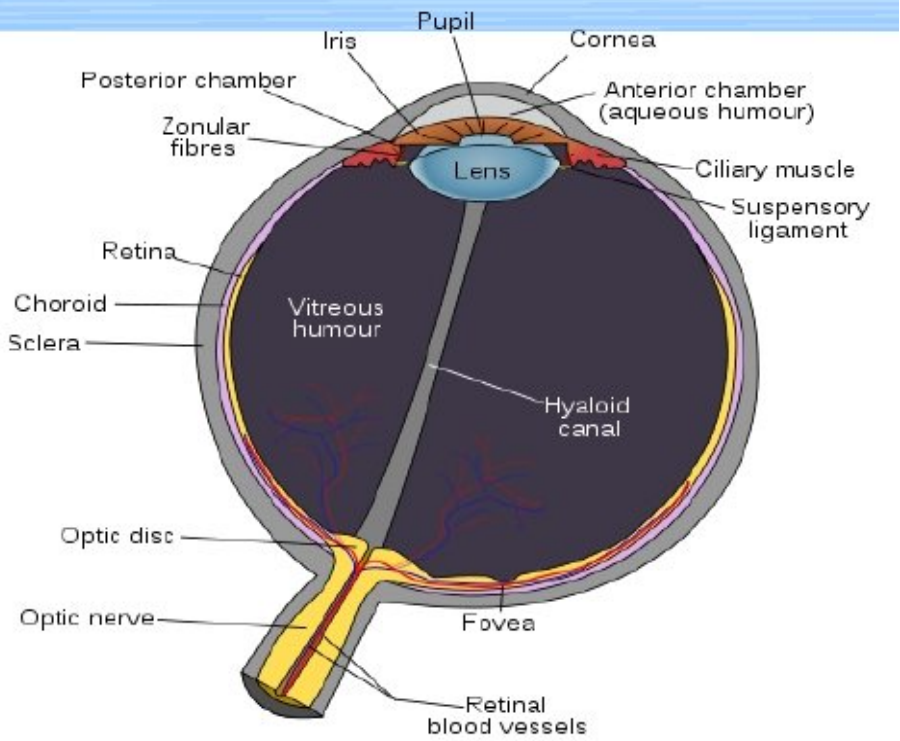
- 6 million cones on fovea centralis in center of retina
  - For color vision: 3 types, Red, Blue, Green
  - Opsins (Pigmentation) to vary color sensitivity
  - Sharpest focus
  - Peak sensitivity at 550 nm
- 120 million rods in 2 rings around fovea
  - Black & white vision
  - 100 times more sensitive photon counters
  - Peak sensitivity at 500 nm
- Ganglion in retina at back of eye for preprocessing
- One million fibers in octave nerve

Human Retina

- Five types of neurons on retina
  - Provides sensor information to optic nerve
    - Receptor cells (rods & cones)
    - Bipolar neurons (connect receptors to ganglion)
    - Ganglion cells (type of neuron for preprocessing)
  - Communicate between retinal cells in INL
    - Horizontal cells
      - Help integrate & regulate from photoreceptors
    - Amacrine cells
      - Inhibitory
      - Interact with retinal ganglion & bipolar cells
  - Preprocessing done on the retina

The Human Eye





Axial organization of the retina (Modified from Santiago Ramón y Cajal, 1911) Retina's simplified axial organization. The retina is a stack of several neuronal layers. Light is concentrated from the eye and passes across these layers (from left to right) to hit the photoreceptors (right layer). This elicits chemical transformation mediating a propagation of signal to the bipolar and horizontal cells (middle yellow layer). The signal is then propagated to the amacrine and ganglion cells. These neurons ultimately may produce action potentials on their axons. This spatiotemporal pattern of spikes determines
the raw input from the eyes to the brain.

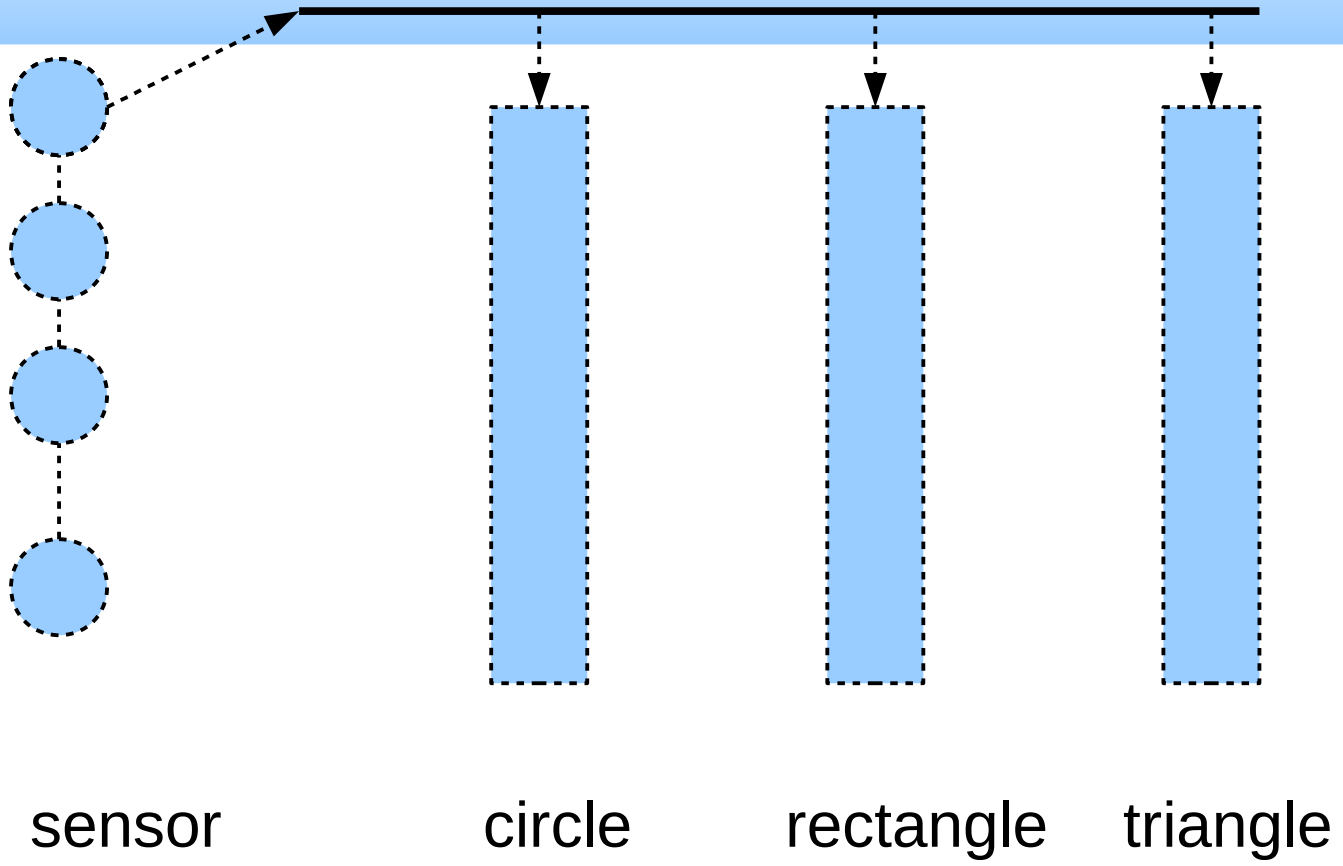
Color Understanding

- Most primates have 3 cones, R, B, G
- A few have 4 cones
  - Tetrachromacy
  - Can differentiate more colors
- Some birds, fish, reptiles, others
  - Tetrachromatics
  - UVB sensitive 4^{th} cone
  - Expands range down to 290 nm
- IR sensors on some snakes
  - Expands range up to 5000 nm

Processing

- Cat has columns of cells in cerebral cortex that are responsible for processing various shapes
  - Each one for a specific shape
- This columnar organization is prevalent in neocortex of most primates, but not in other brain areas
- A set of close columns receives the same signal but each one performs different operations.

Schematic



Human vision Processing

- Left side of each eye connects to left visual cortex
- Right side of each eye connects to right visual cortex
  - Receives signals from eyes
  - Interprets shapes, colors, movement
- Visual association cortex analyzes patterns in visual information and compares with things seen previously to form images
  - Compares to what we have seen in past and what we expect to see

Primate Vision Features

- Animals with image forming visual systems
- The brain controls the motion and saccades as needed
- Image on fovea centralis
  - Highest resolution part of retina
  - Scans continuously to build high resolution image
  - Rods & cones only detect changes in intensity
- The brain controls the sensor movement
- We look at a moving object, yet see no blur

Recognition by Components

- Proposed 1987, 1990 by Biederman
  - Recognize by breaking into components
  - Look in memory- which object matches this combination
- Human system has small number of geons
  - Simple 3-D geometric forms
  - Note where they are connected
    - These 'intersections' are very important
  - Use edges to recognize geons

Pattern Detection in High Level Primates

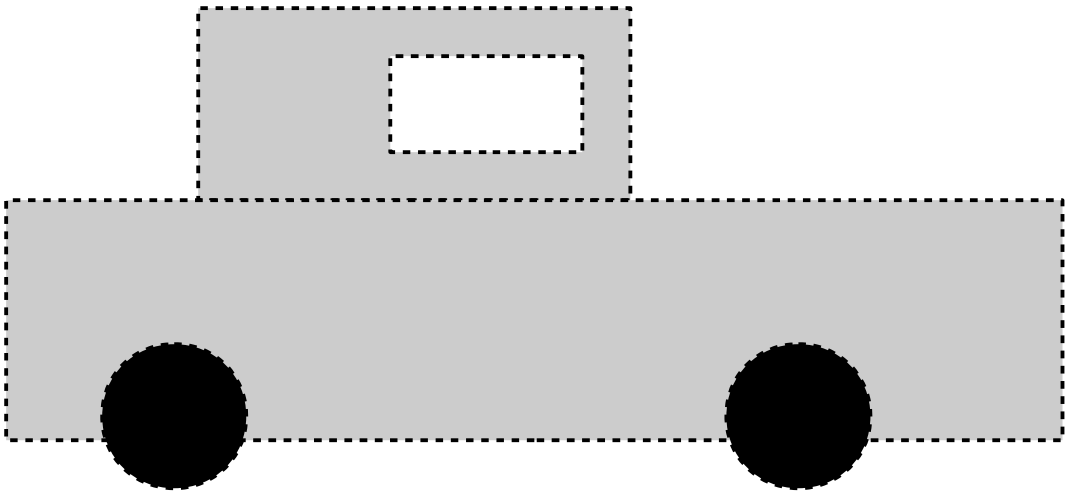
Shape: The visual cortex detects generic image features: line, circle, square. The shape is first detected from the edges.

Structural: The relationship between the recognized shapes

Matching: Now with reduced sample space, match image to history/library

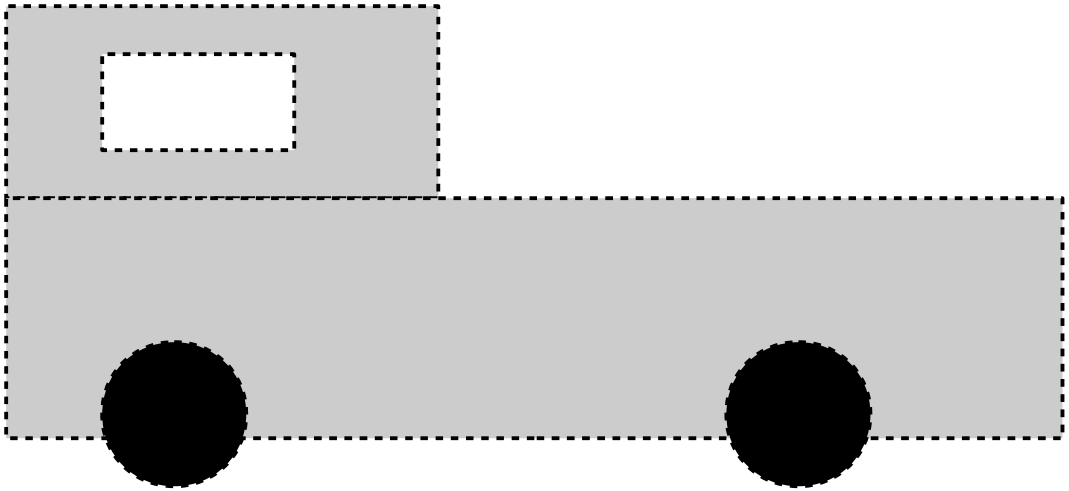
Additional information, such as time, location, situation, etc adds into the logic and reduces the search space

Example



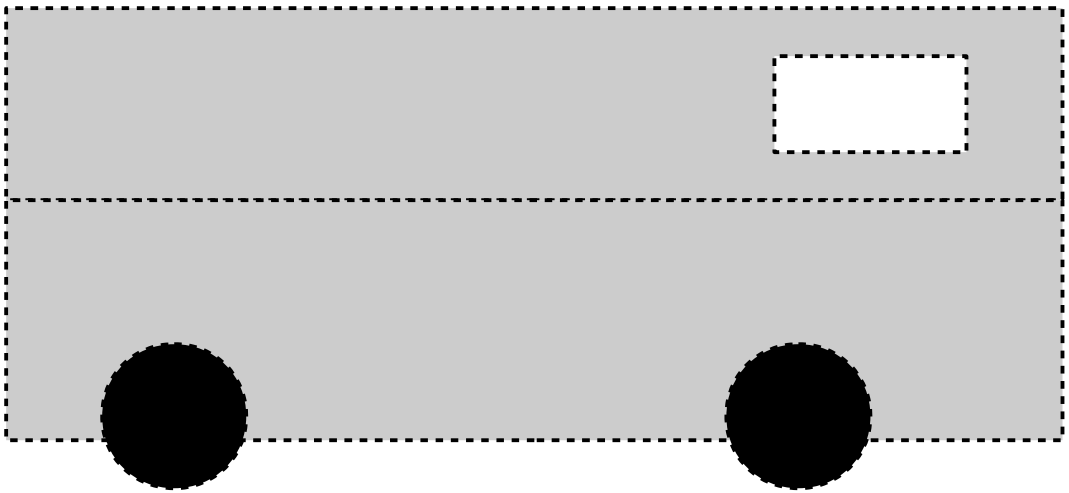
Three rectangles and two circles
1956 Lincoln ? 1964 Mustang?

Move The Top Box



If going to the right, then Kenworth Tractor
If going to the left, then a car hauler

Stretch the Top Box



Now this looks like a bus or motor-home

Combination of Information

- Combined information
  - All of our senses
  - Memories
  - Learned patterns
  - Other information
- All combined together
  - Multiprocessing
  - Reasoning
  - Create decision of what we think that it is
  - In background, not conscious of reasoning

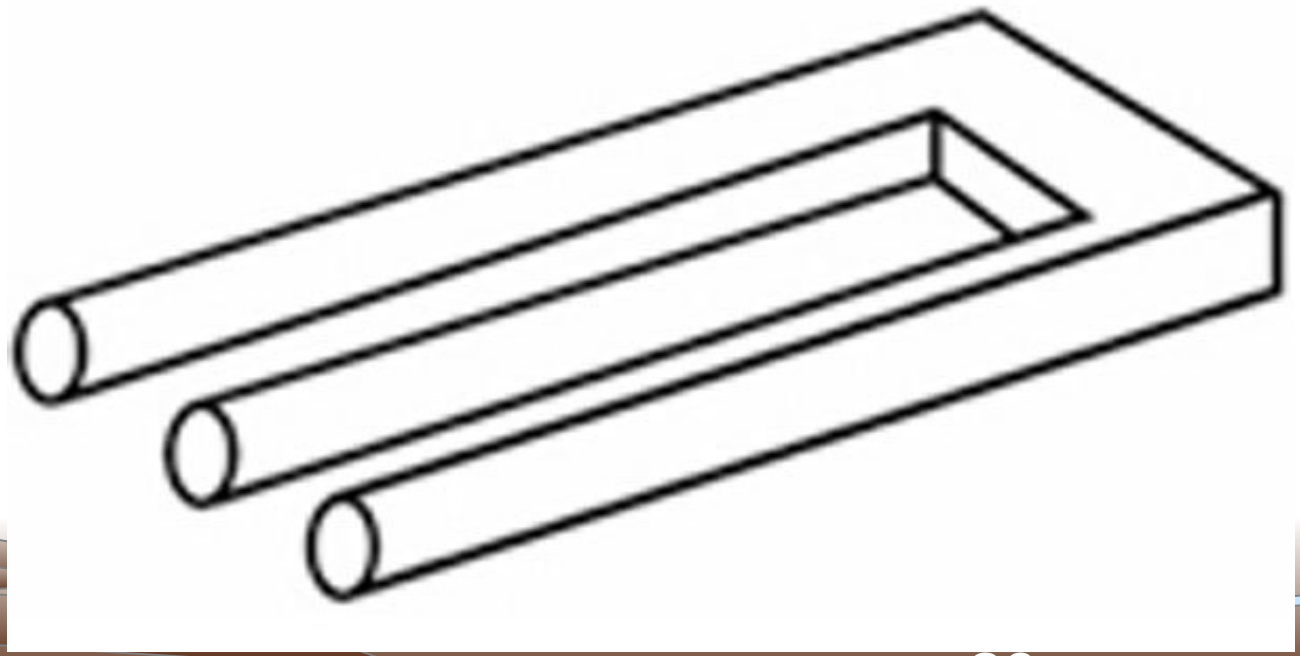
Missing Part of Combination

- Ever go someplace and see a familiar face?
- You know that you know this person
  - Recognize the face
  - Do not know who it is
- Your recognition is a combination of the person AND other things
- You think and think and might never know who it is
  - This has happened to me

Optical Illusions

- Our vision system is tricked (fooled)
  - Visual percept differs from reality
- Three Classes (Richard Gregory)
  - Physical
  - Physiological
  - Cognitive

This is one of the most famous optical illusion pictures of an impossible object. It has two rectangular prongs at one end that morph into three cylindrical prongs at the other.



Model Based Decision Support System

- Have model (s) of what you think it is
- Have image under analysis
- Have fitness function
  - Use all sensor data
- Put all images/models into fitness function one at a time, record scores
- Now have probability of what is the image

Other Logic & Math Methods

- Dempster-Shafer Evidence Theory
  - Reasoning with Uncertainty
  - Combine evidence from different sources to get a degree of belief
  - Plausibility
- Bayesian probability
  - Additional information
  - Posterior probability
  - Naive Bayes

Human Vision Performance

- Robust
  - Recognize noisy images, sloppy writing, colors, textures, occluded objects
- Predators have stereoscopic (depth perception) while prey have panoramic (larger field of view)

Typical Computer Vision Functions

- Edge detection
- Segmentation
- Feature extraction
  - Texture
  - Shading
  - Stereoscopic (create 3D)
  - Motion
  - Recognition

Typical Computer Pattern Recognition

- Template matching
  - Large catalog of templates
  - Exhaustive match of templates
  - Computationally intensive & time consuming
- Cross correlation
- Feature matching
- Hough Transforms

Previous Work

- Photoperceptron (1961)
  - Responded to optical patterns
  - Contained:
    - Sensory – 400 photocells (retina function)
    - Association – 512 neuron like units (-1, 0, +1)
    - Response – manually adjusted
- Neocognitron (1980)
  - Could account for invariance in position
- SEEMORE (1997)
  - Analyzed texture, color, & contours
  - Worked with changes in position, scale, & rotation
- Google Cloud Vision API

Comparison

Biological Neural	Brad's Computer
Speed 1 Khz	Speed 4 GHz
10^6 to 10^{11} processors	Eight core = 8 CPU

How Can We Implement

- Current computers are over 10^6 faster
  - Can we emulate multiprocessing?
- Multi-core processors
  - 16 cores (hypertreaded) on the market now
  - More coming
  - 80 Core units announced
  - GPU with 2560 (RISC) cores (GeForce 1080)
  - Xeon Phi and Nvidia Tesla
  - Neuromorphic chips
- Can we implement MIMD?

Please Remember

- The image is created in the brain, not the sensor
- Preprocessing in the sensors
- Do not limit concepts to visible light spectrum
- More than one sensor can contribute information
- Memory contributes to the image
- Think about how one can read a novel and create a picture in their mind

Models

MIMD processing is necessary

Multiple processes must be occurring simultaneously

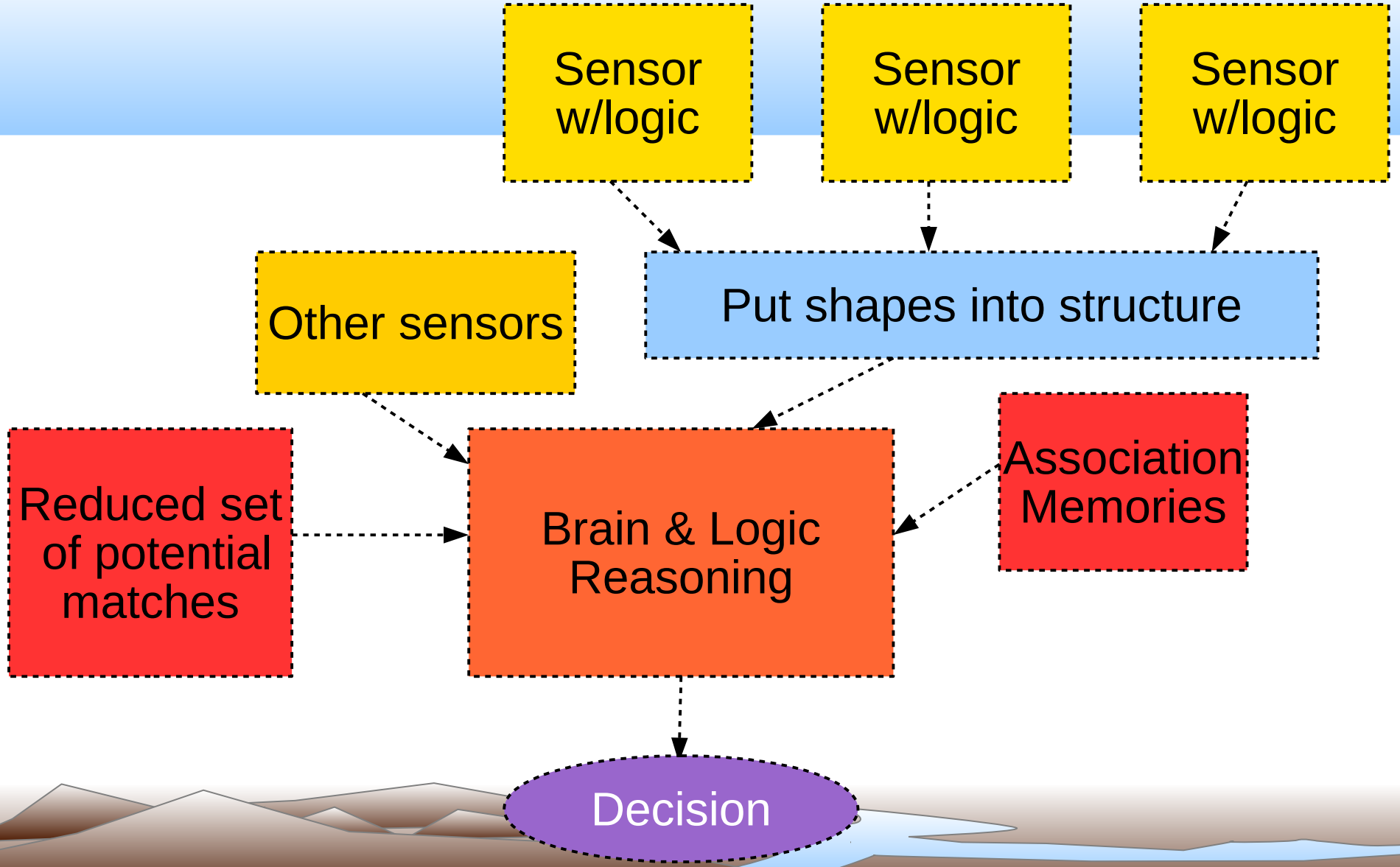
Use speed on computer to emulate more parallel processing

Use multiprocessor systems

Distributed processing

- Some processing at the sensor

Block Diagram



References

Intelligence, The Eye, the Brain, and the Computer, Fischler & Firschein, Addison-Wesley, 1987

www.machine-cognition.com

Cognition, Ashcraft, Prentice-Hall, 2002

Artificial Intelligence: a Modern Approach, *2nd Edition,* Russell & Norvig, Prentice-Hall, 2003

Computer Vision and Fuzzy-Neural Systems, Kulkarni, Prentice-Hall, 2001

Pattern Recognition and Image Analysis, Gose, Johnsonbaugh, & Jost, Prentice-Hall, 1997

www.scholarpedia.org

www.ieee.org

Numerous other articles and books

Contact Info

Dr Brad Morantz

www.machine-cognition.com

bradscientist@machine-cognition.com

Questions?