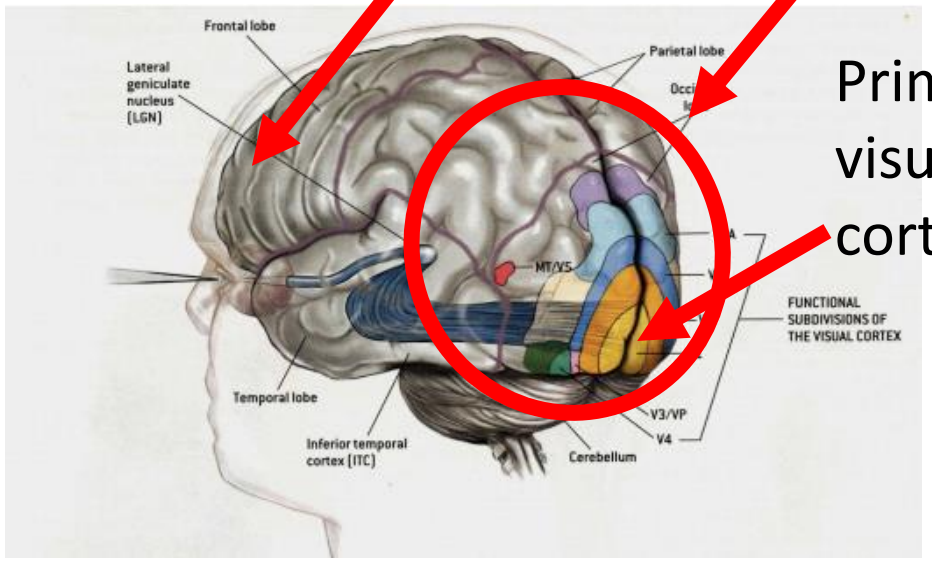
A new path to understanding vision

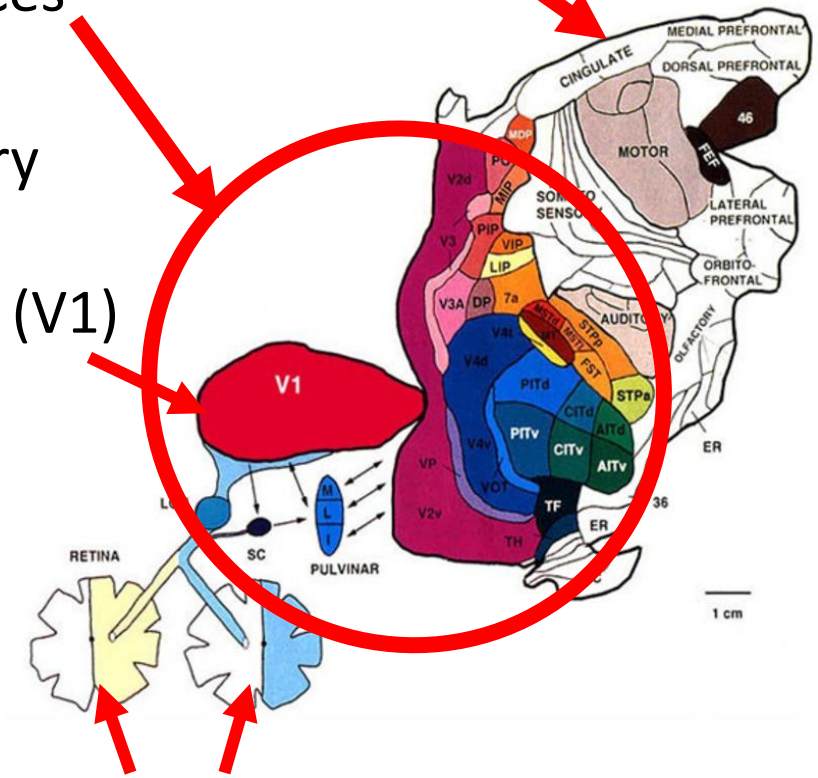
from the perspective of the primary visual cortex

Frontal brain areas

Visual cortices

Primary visual cortex (V1)





Retina

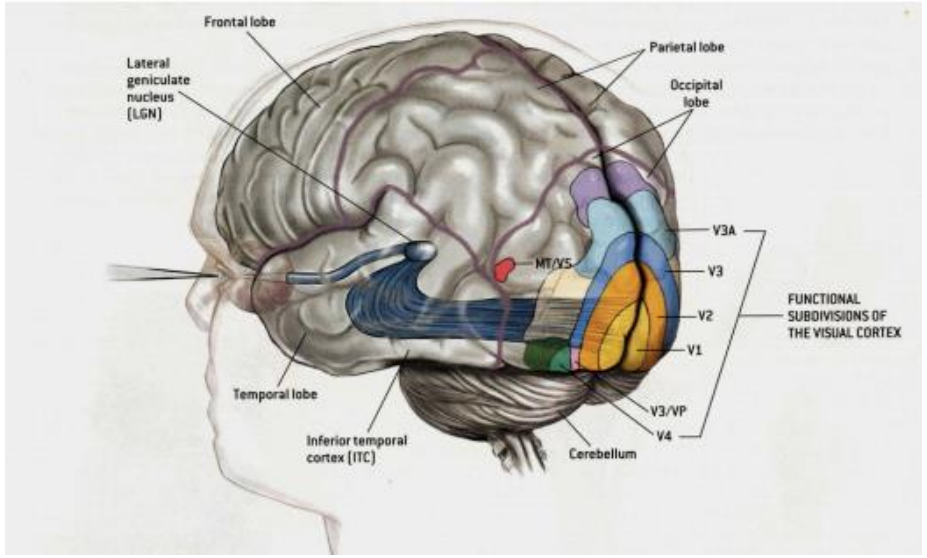
Li Zhaoping
July 16, 2018, presented at APCV 2018,
HangZhou, China

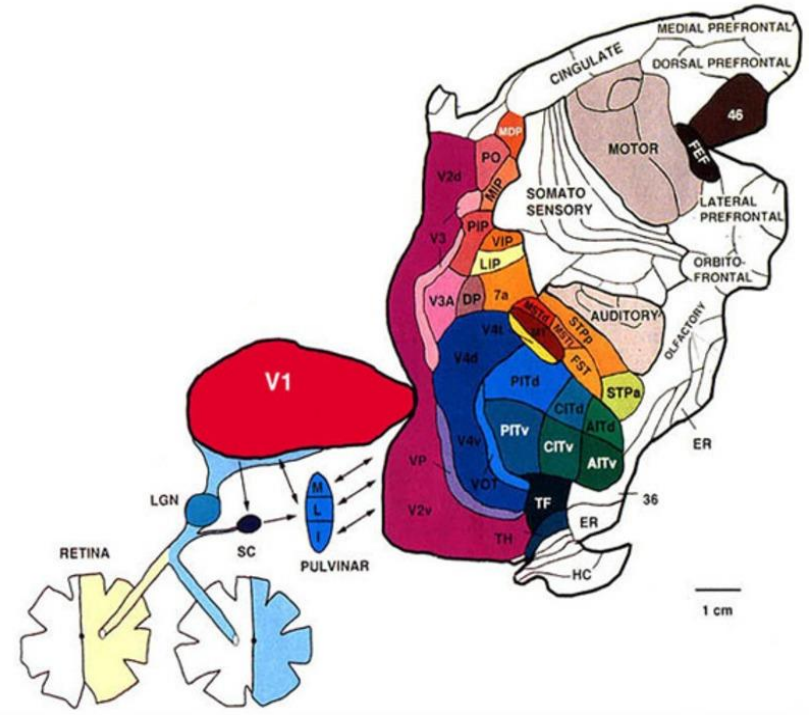
A new path to understanding vision

Traditional paths to understanding vision

(1) Low level vision, mid-level vision, high-level vision

(2) David Marr: primal sketch, 2.5 d sketch, 3-d model.





Li Zhaoping
July 16, 2018, presented at APCV 2018,
HangZhou, China

Talk outline

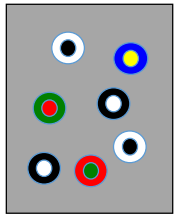
(1) The functional role of the primary visual cortex (V1)

(2) In light of V1's role →
a new plan to understanding vision

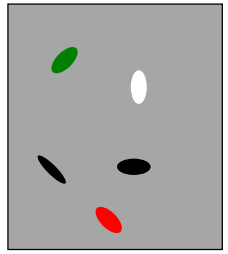
(3) A first example study in this new plan

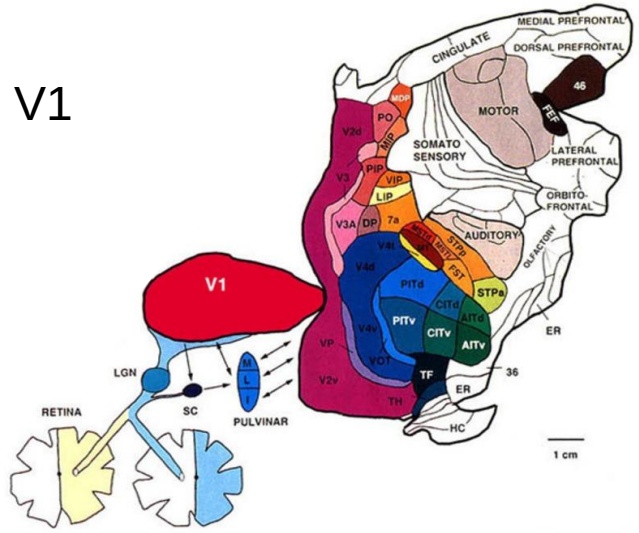
The primary visual cortex (V1)

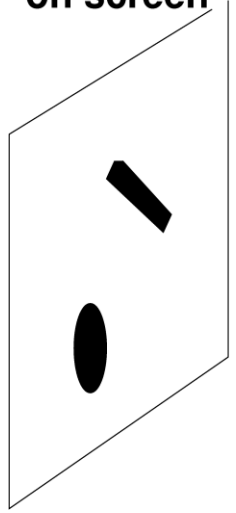
1953, Stephen Kuffler, retina,

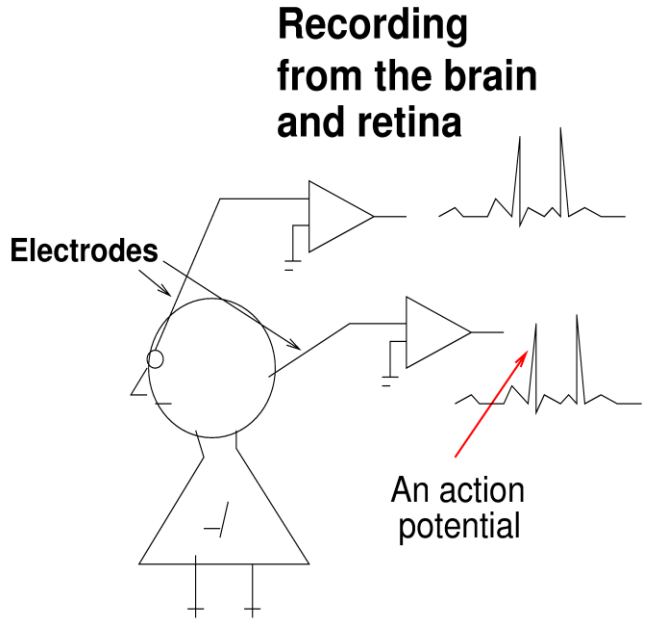


1959-- Hubel and Wiesel, V1



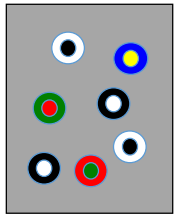


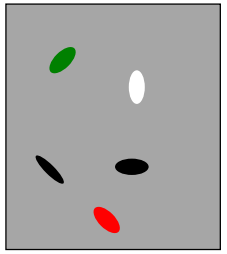


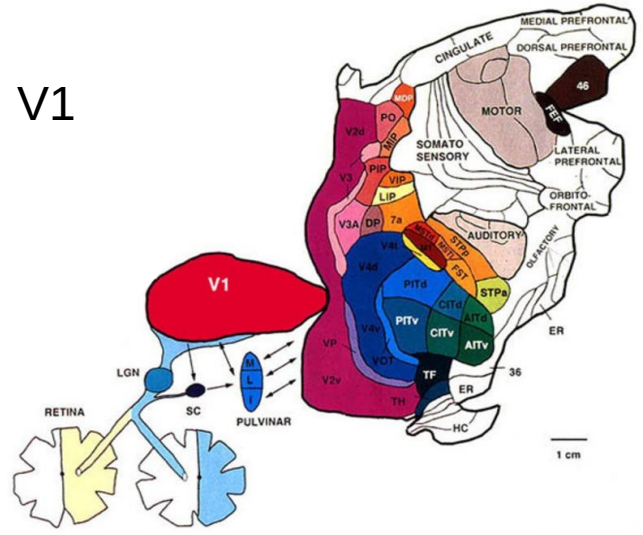


The primary visual cortex (V1)

1953, Stephen Kuffler, retina, 1959-- Hubel and Wiesel, V1







Then …

Experimentally: V1 and beyond

Theoretical/modelling, Reichardt, Marr, etc.

2005: How close are we to understand V1? Olshausen and Field 2005

Do we really know what the early visual system does?

Carandini, Demb, Mante, Tolhurst, Dan, Olshausen, Gallant, Rust, 2005

Standard models of V1 neural receptive field (combining filtering, rectification, squaring, normalization) **captures only 15-35% of the variances in V1 responses**.

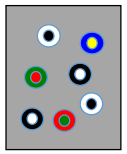
2012, David Hubel, in answer to "What Do You Feel Are the Next Big Questions in the Field?"

"**We have some idea … for the retina, the lateral geniculate body, and the primary visual cortex, but that's about it**."

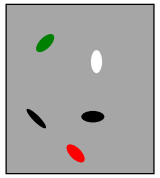
(Hubel & Wiesel 2012, Neuron)

The primary visual cortex (V1)

1953, Stephen Kuffler, retina,

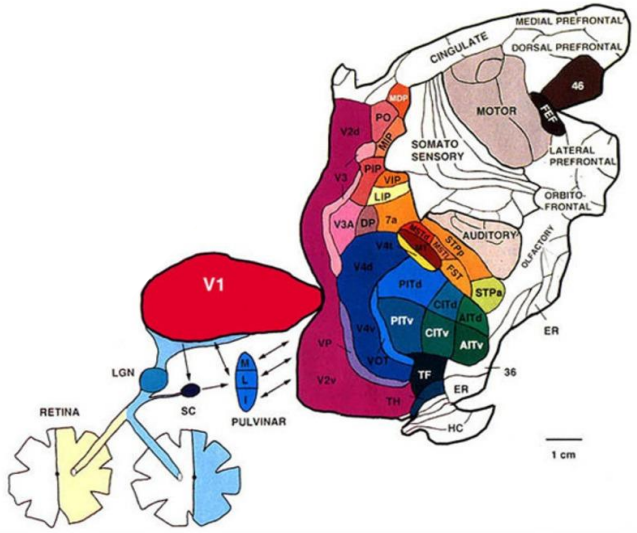


1959-- Hubel and Wiesel, V1



Then …

Experimentally: V1 and beyond
Theoretical/modelling, Reichardt, Marr, etc.



2005: How close are we to understand V1? Olshausen and Field 2005

Do we really know what the early visual system does?

Carandini, Demb, Mante, Tolhurst, Dan, Olshausen, Gallant, Rust, 2005

Standard models of V1 neural receptive field combining linear filtering, rectification and squaring, and response normalization **captures only 15-35% of the variances in V1 responses**.

2012, David Hubel, in answer to "What Do You Feel Are the Next Big Questions in the Field?"

"**We have some idea … for the retina, the lateral geniculate body, and the primary visual cortex, but that's about it**."

(Hubel & Wiesel 2012, Neuron)

Questions:

Is a lack of understanding of V1 hindering our progress beyond V1?

Physiologically

Functionally (in behaviour)?

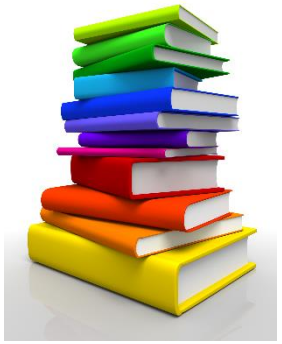
Information bottlenecks in the visual pathway:

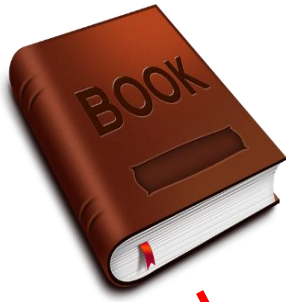
"To be or not to be,
This is the question .."

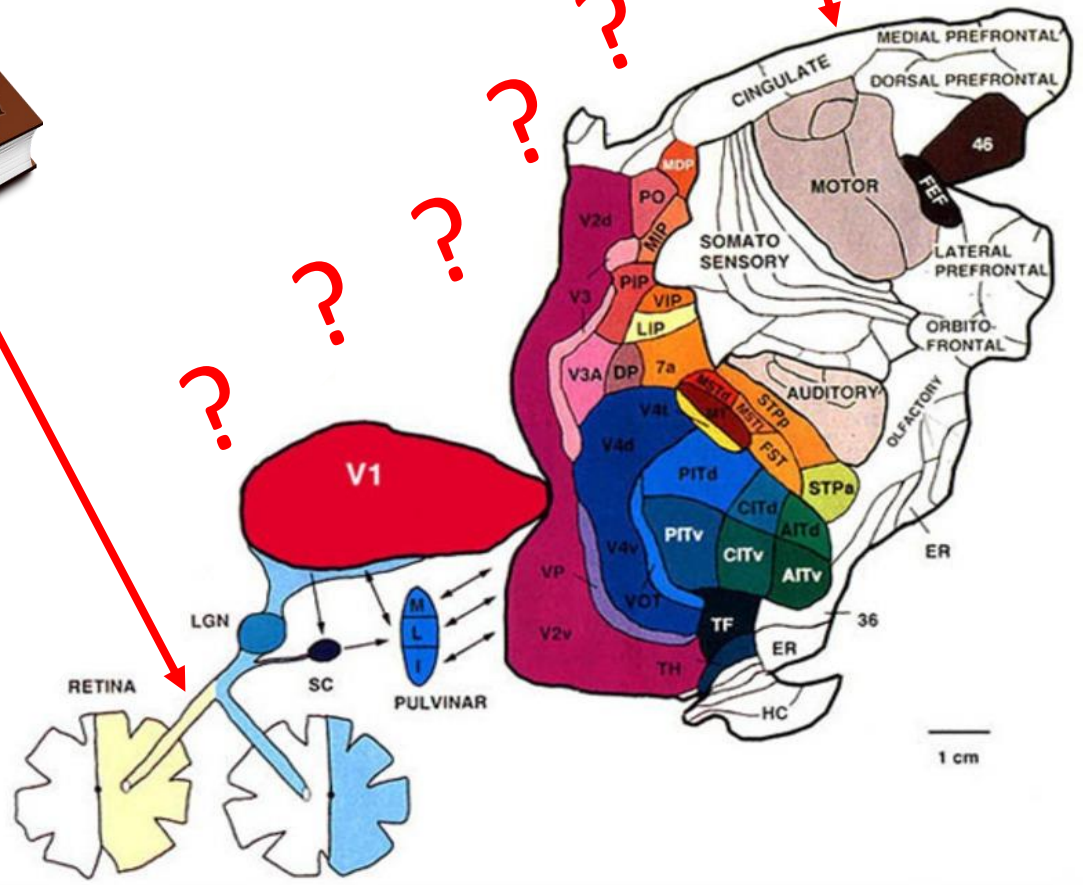
10^7 bits/second
~ 10^6 neurons ,
~10 spikes/neuron
~1 bit/spike

40 bits/second
(Sziklai, 1956)

10^9 bits/second (Kelly 1962)
~ 25 frames/second,
2000x2000 pixels,
1 byte/pixel







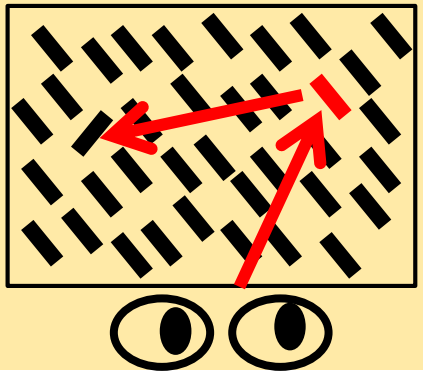
Information bottlenecks in the visual pathway:

We are nearly blind!

Vision ~ Looking (selecting) + Seeing

top-down vs. **bottom-up** selection

Task: find a uniquely oriented bar

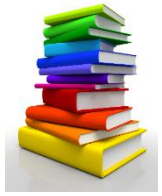


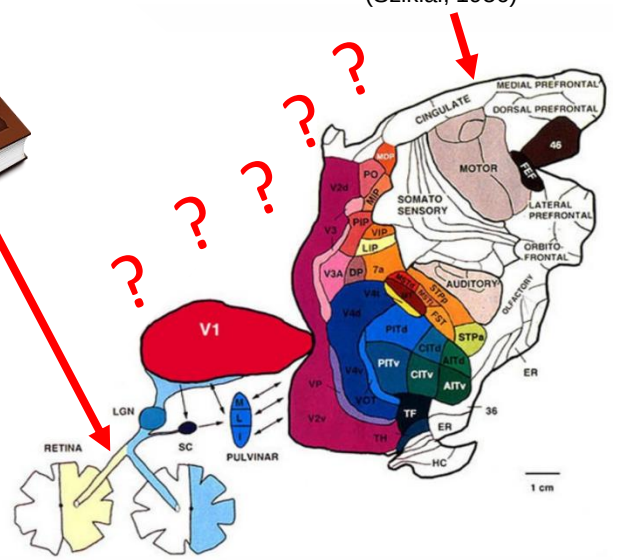
"To be or not to be,
This is the question .."

40 bits/second
(Sziklai, 1956)

10^7 bits/second
~ 10^6 neurons ,
~10 spikes/neuron
~1 bit/spike

10^9 bits/second
(Kelly 1962)
~ 25 frames/second,
2000x2000 pixels,
1 byte/pixel





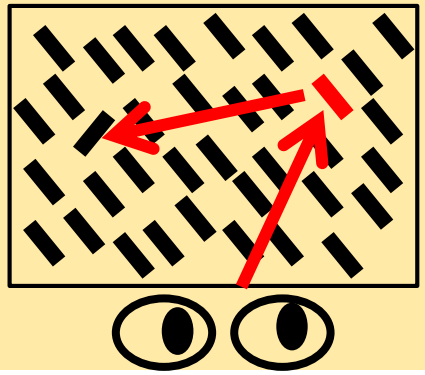
Information bottlenecks in the visual pathway:

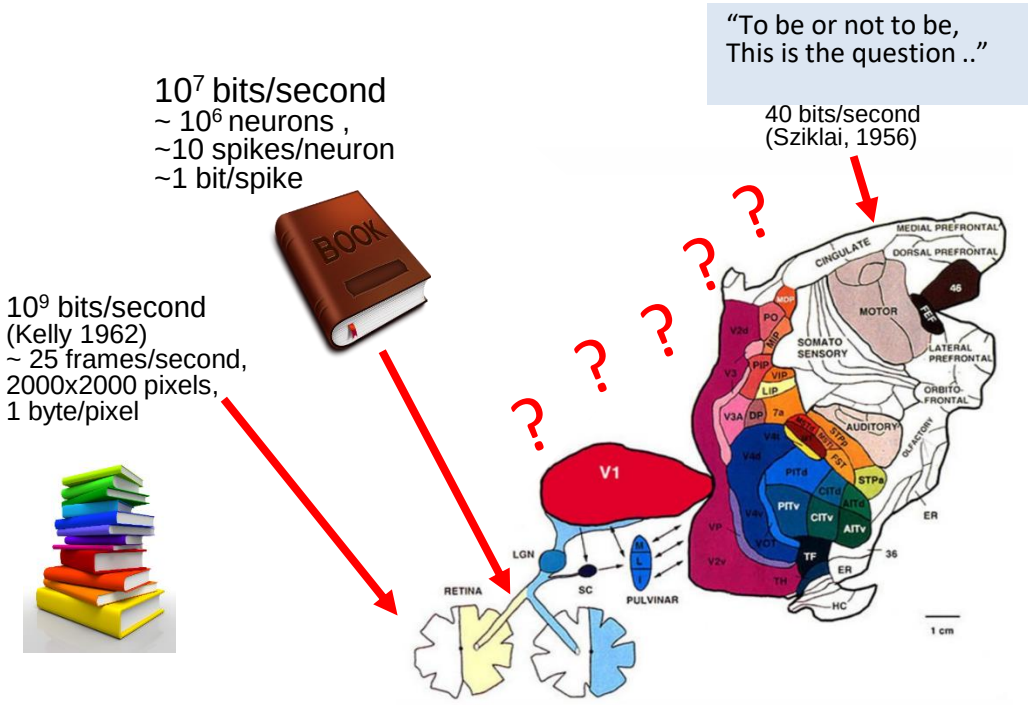
Questions:
which brain areas are doing the bottom-up selection?

Frontal? Parietal?

top-down vs. **bottom-up** selection

Task: find a uniquely oriented bar



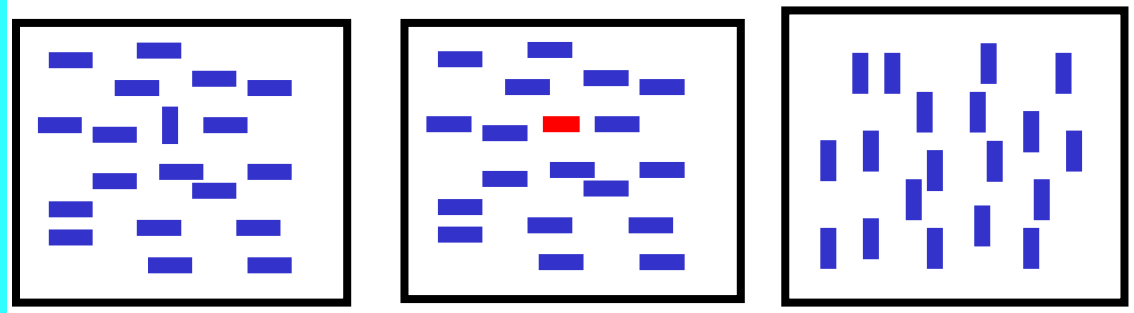


Information bottlenecks in the visual pathway:

Questions:
which brain areas are doing the bottom-up selection?

Frontal? Parietal?

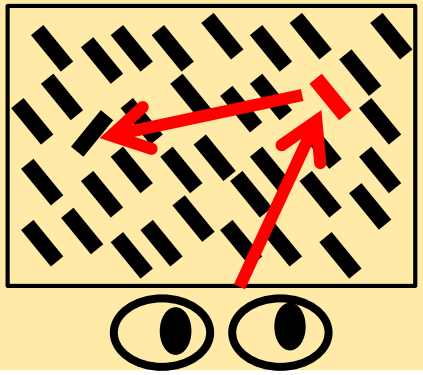
Saliency regardless of visual features



Koch & Ullman 1985, Itti & Koch 2001, etc

top-down vs. **bottom-up** selection

Task: find a uniquely oriented bar

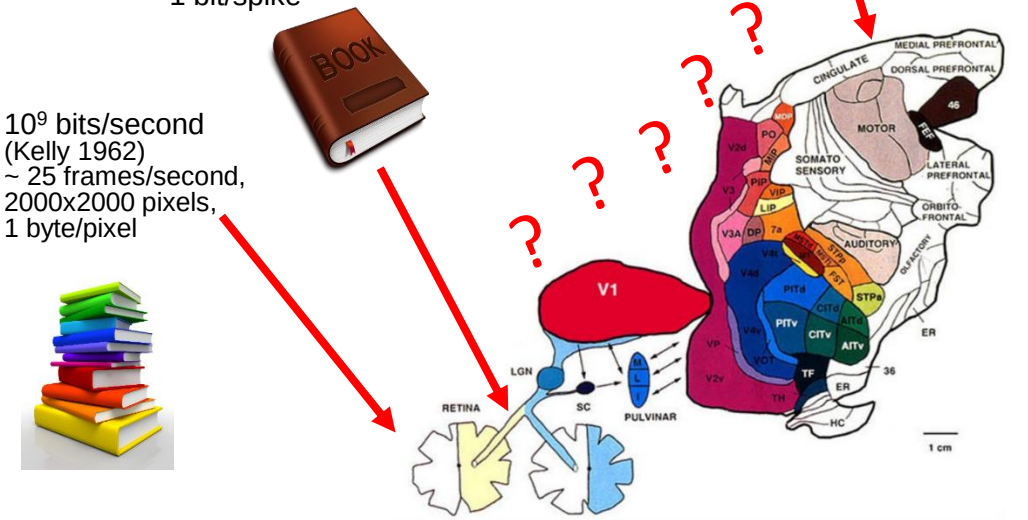


"To be or not to be,
This is the question .."

40 bits/second
(Sziklai, 1956)

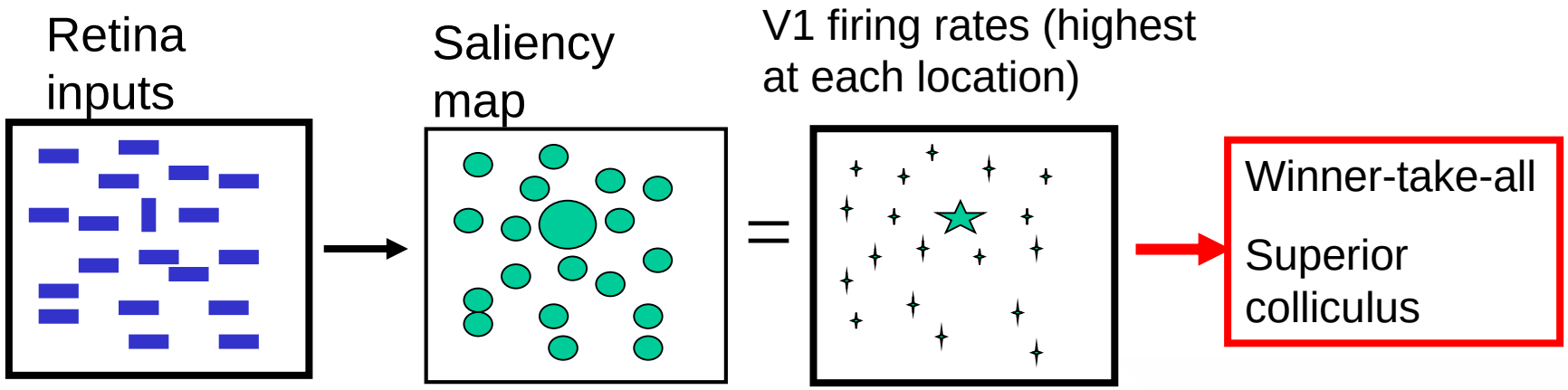
10^7 bits/second
~ 10^6 neurons ,
~10 spikes/neuron
~1 bit/spike

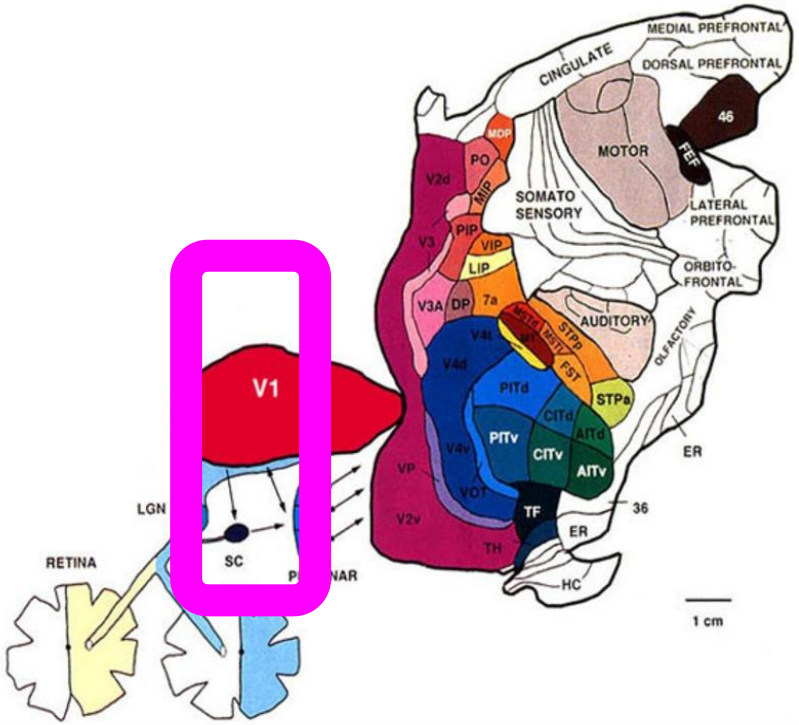
10^9 bits/second
(Kelly 1962)
~ 25 frames/second,
2000x2000 pixels,
1 byte/pixel



The V1 Saliency Hypothesis:

A bottom-up saliency map in the primary visual cortex (Li 1999, 2002)



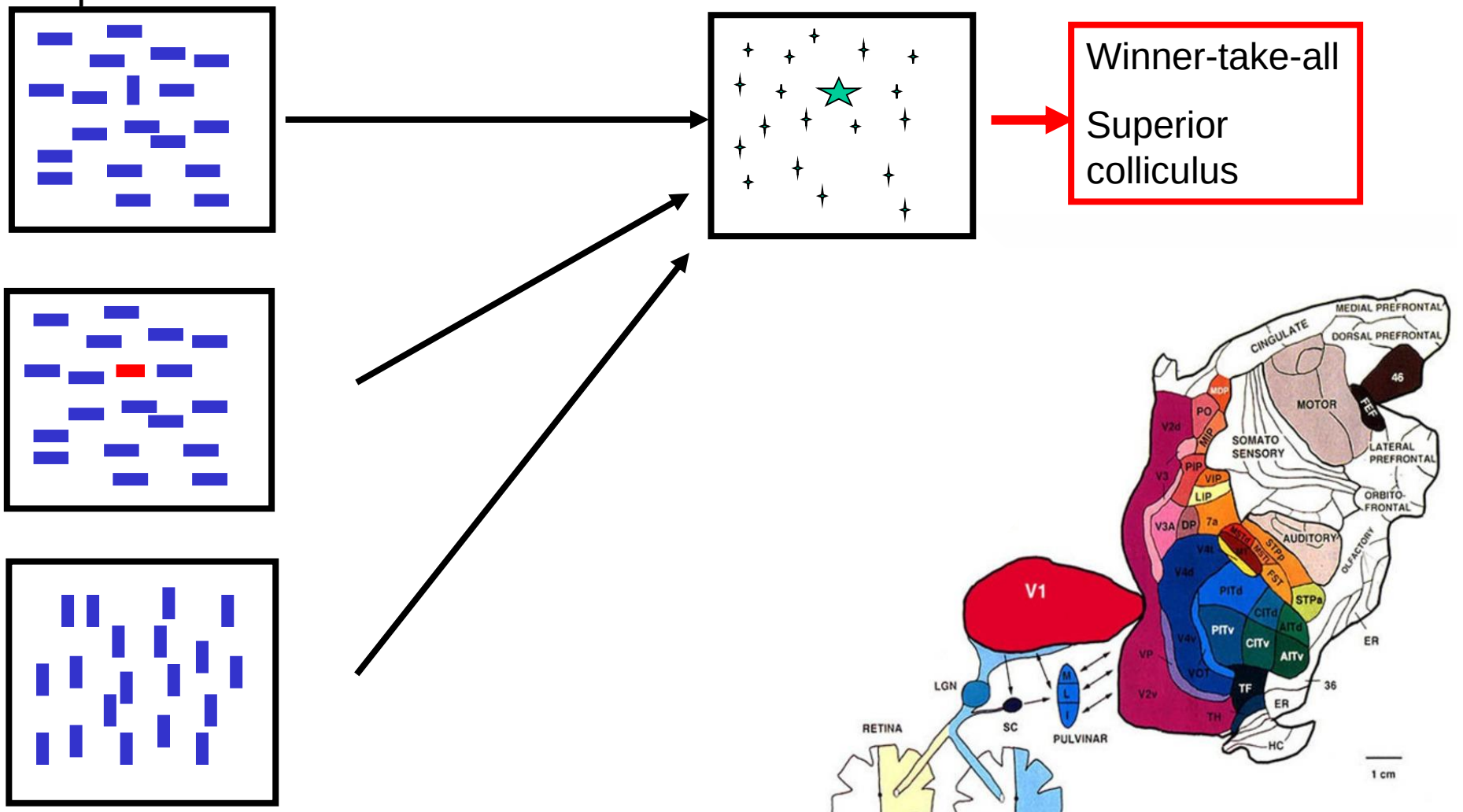


The V1 Saliency Hypothesis:

A bottom-up saliency map in the primary visual cortex (Li 1999, 2002)

Retina inputs

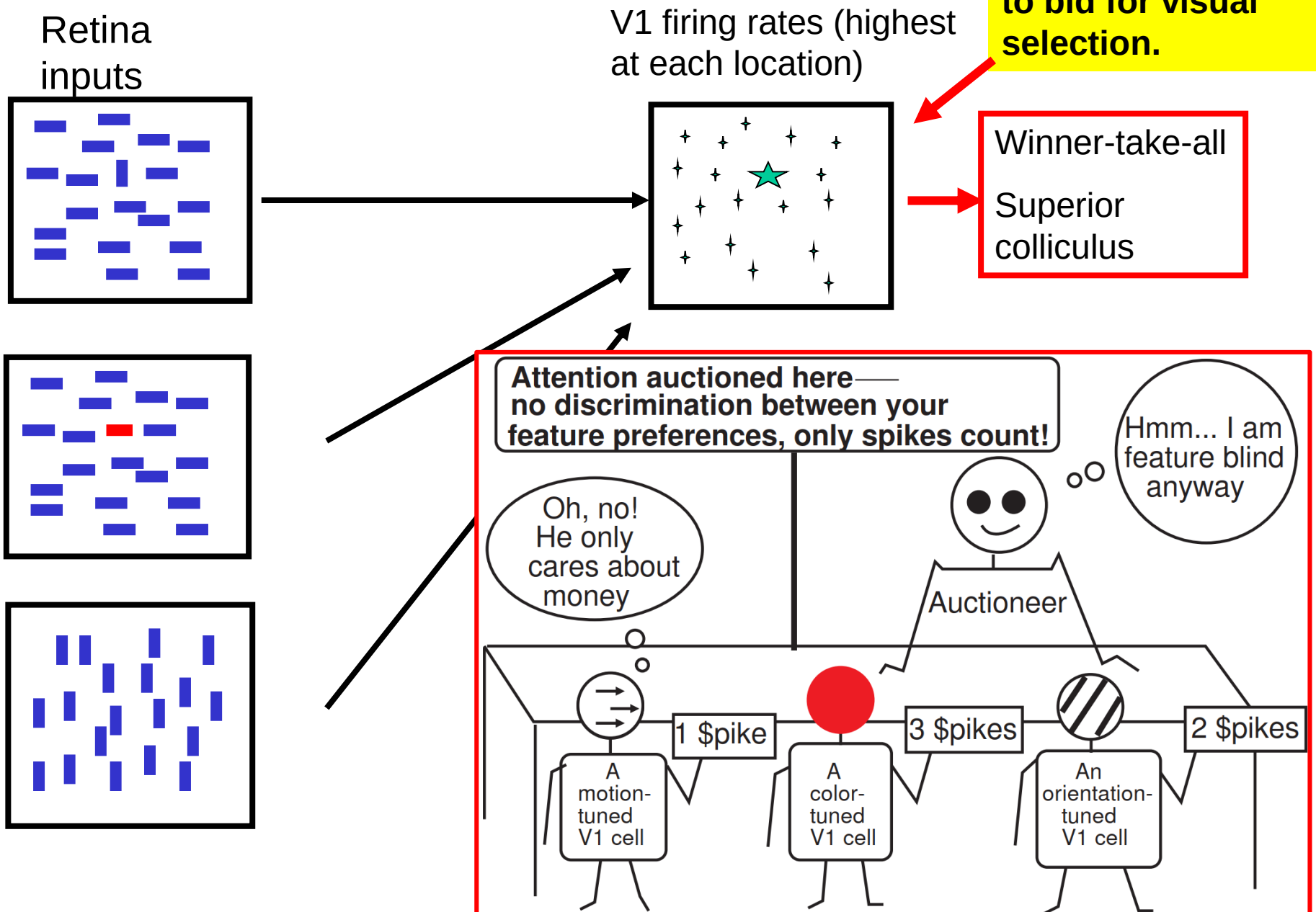
V1 firing rates (highest at each location)



The V1 Saliency Hypothesis:

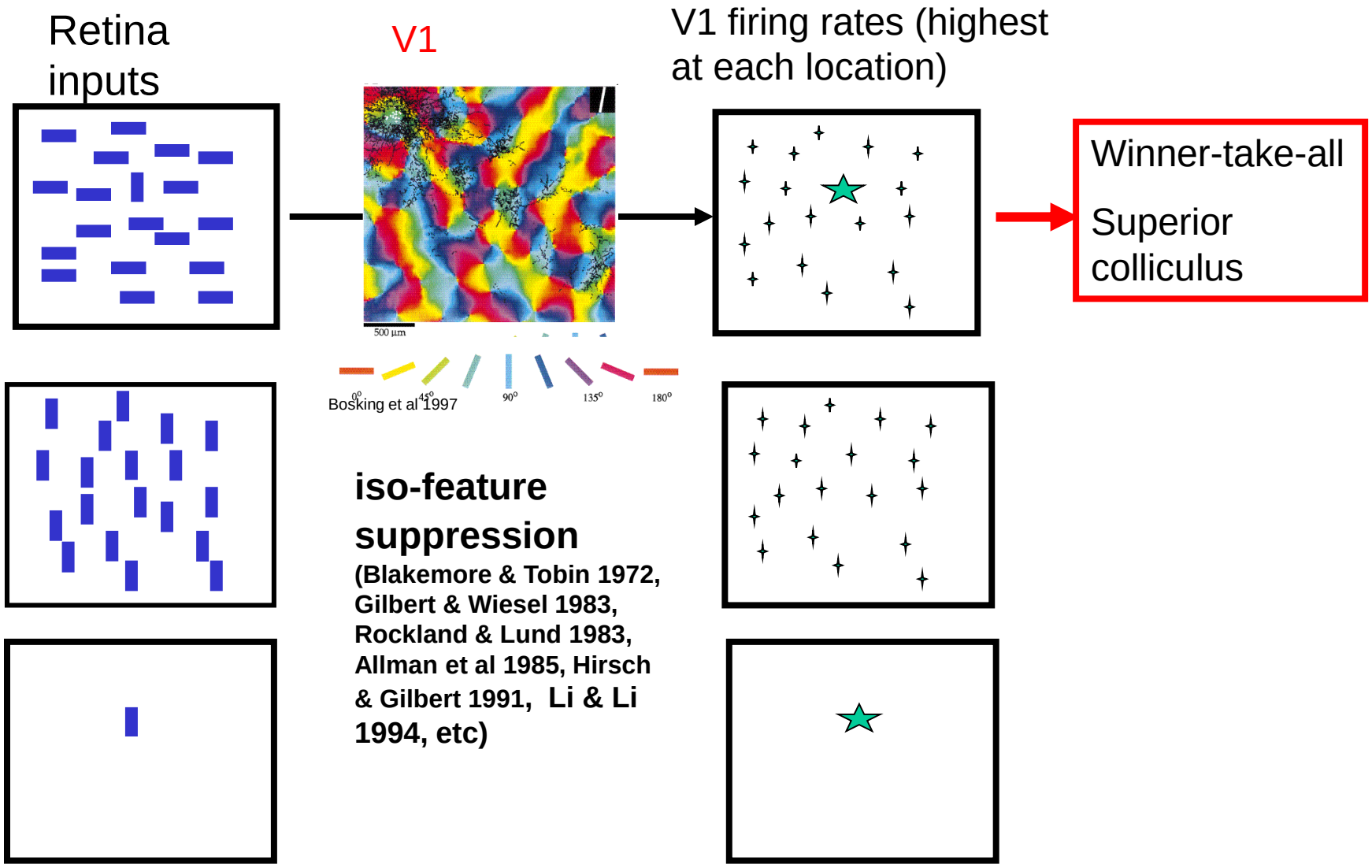
A bottom-up saliency map in the primary visual cortex (

Neural activities as universal currency to bid for visual selection.



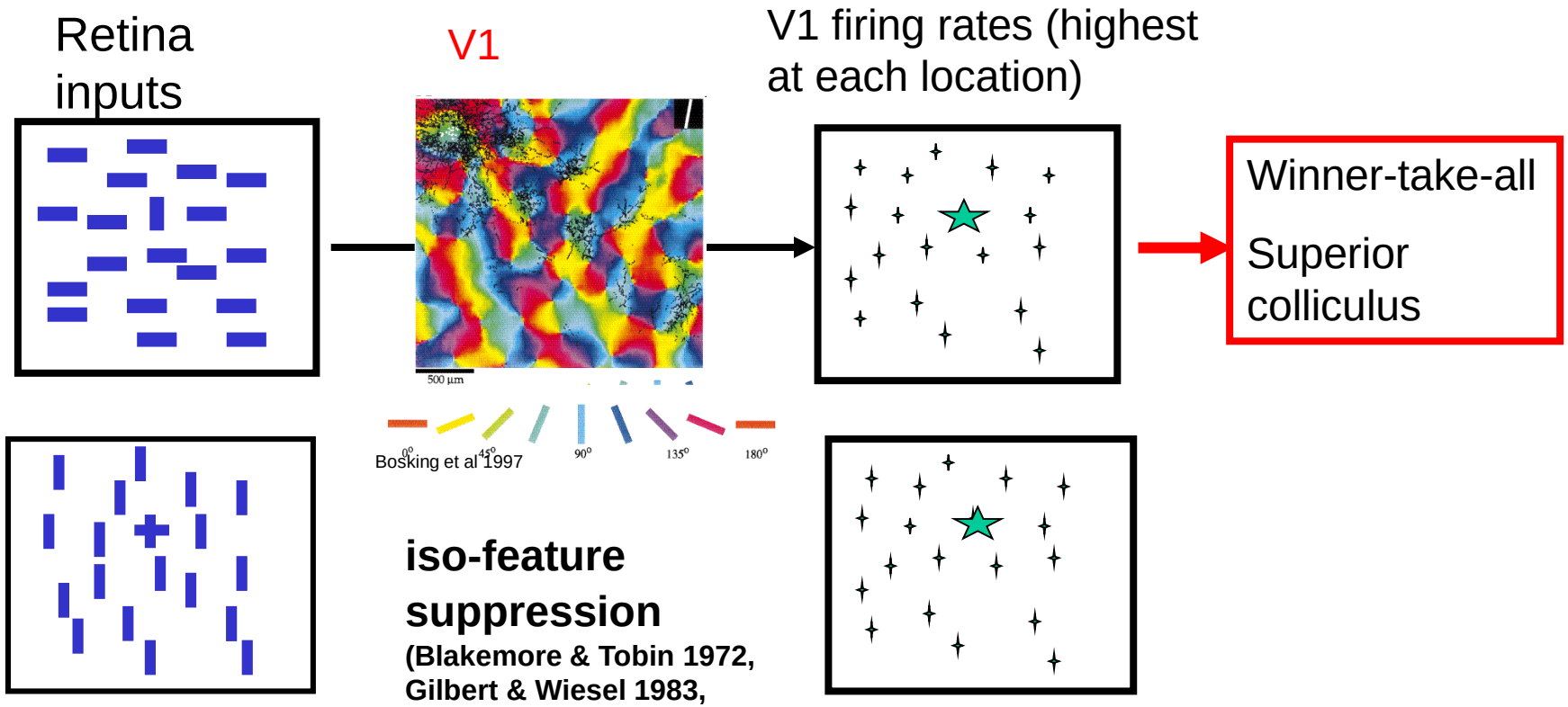
The V1 Saliency Hypothesis:

A bottom-up saliency map in the primary visual cortex (Li 1999, 2002)



The V1 Saliency Hypothesis:

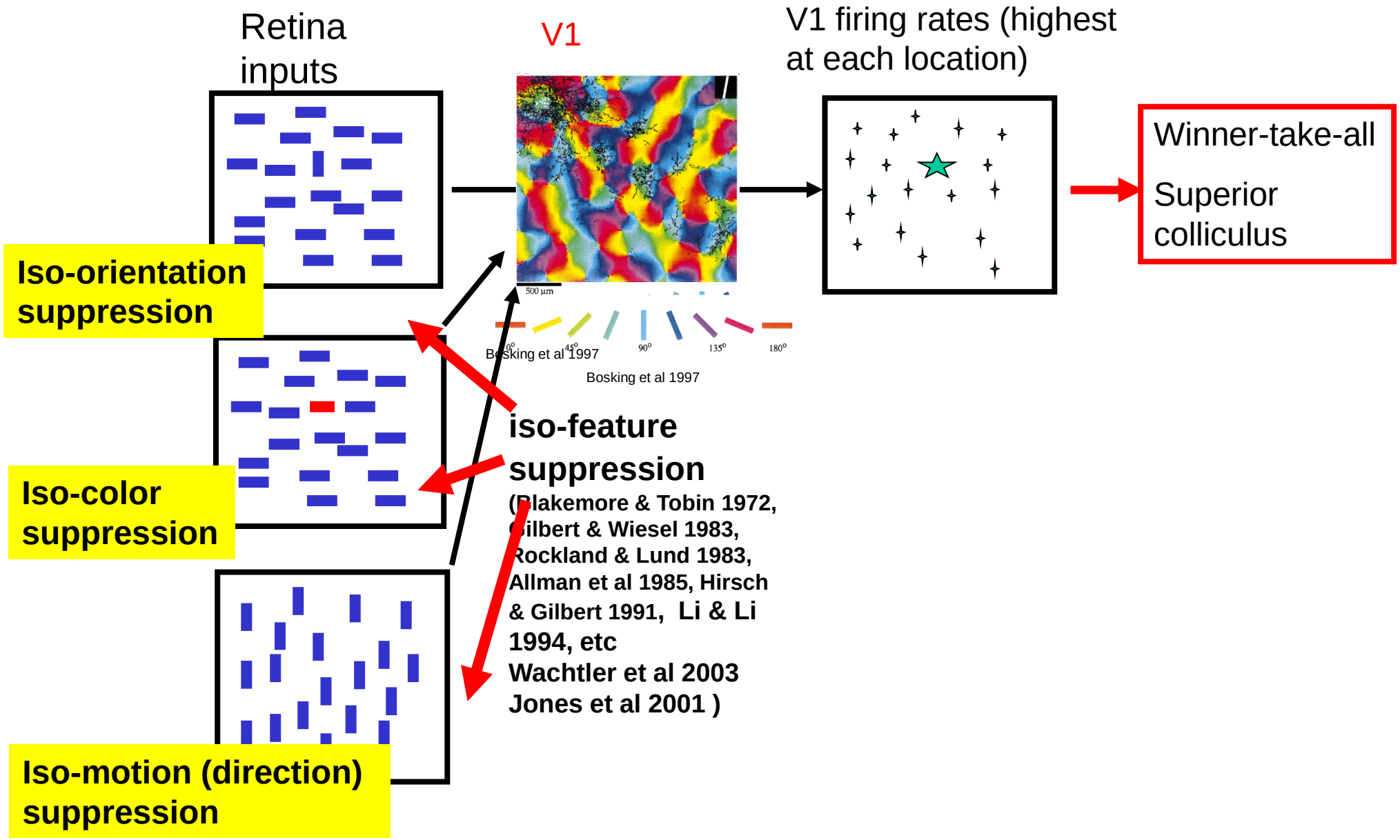
A bottom-up saliency map in the primary visual cortex (Li 1999, 2002)



iso-feature suppression
(Blakemore & Tobin 1972, Gilbert & Wiesel 1983, Rockland & Lund 1983, Allman et al 1985, Hirsch & Gilbert 1991, Li & Li 1994, etc)

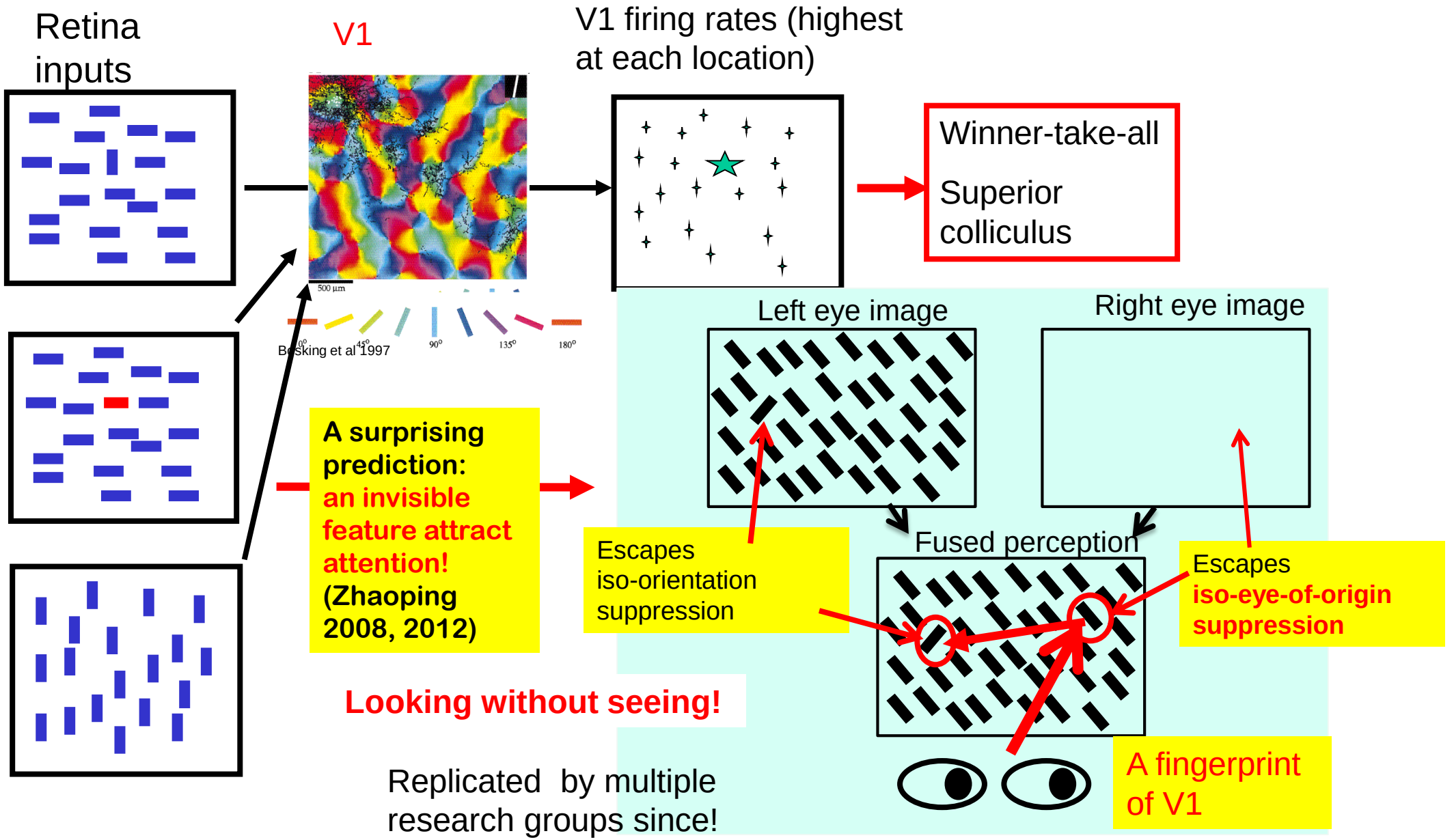
The V1 Saliency Hypothesis:

A bottom-up saliency map in the primary visual cortex (Li 1999, 2002)



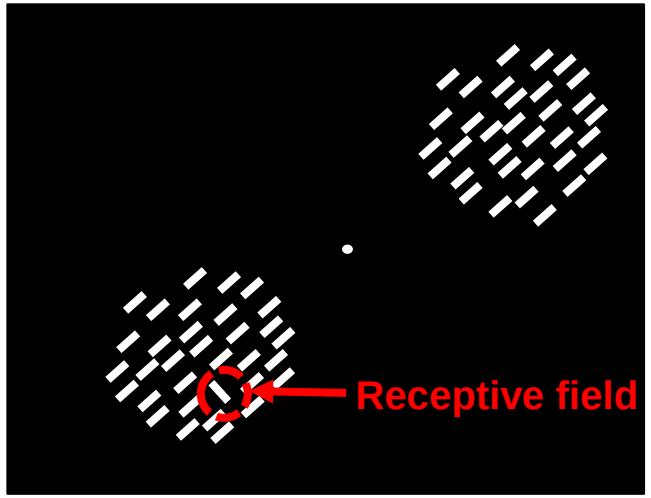
The V1 Saliency Hypothesis:

A bottom-up saliency map in the primary visual cortex (Li 1999, 2002)

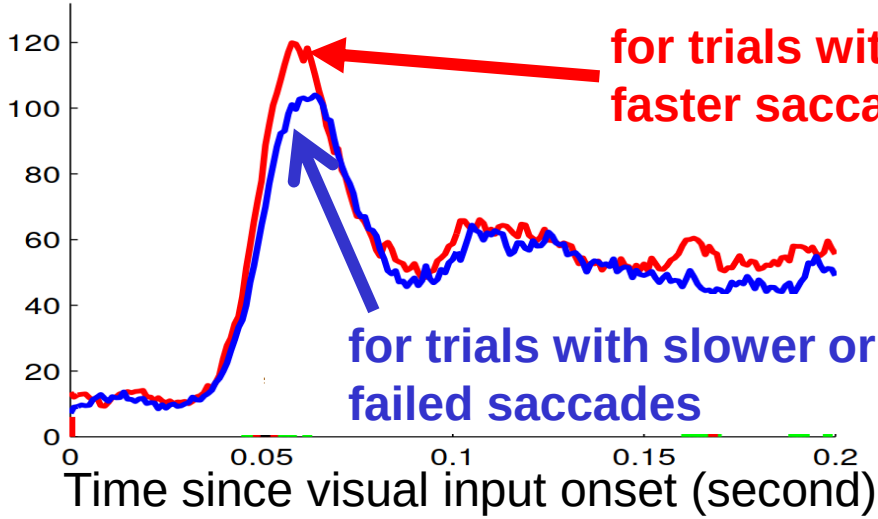


Testing the V1 theory on behaving monkeys --- Yan, Zhaoping, & Li, Submitted.

Saccade to an uniquely oriented bar ASAP

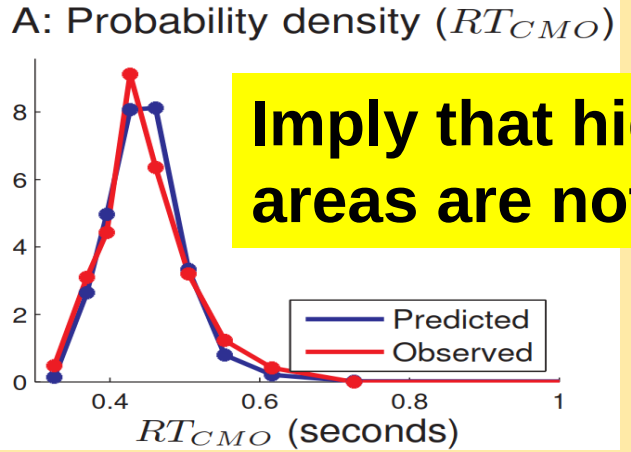


V1 neural responses to input stimulus (spikes/sec)

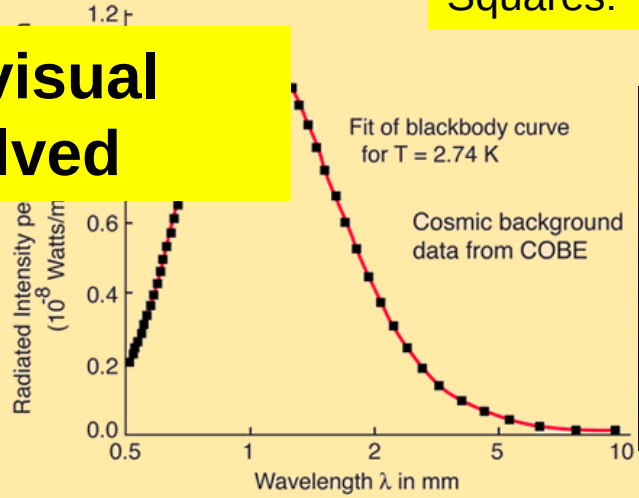


Quantitative, zero-parameter, predictions from theory

Solid curve ---Planck's law
Squares: --- data points



Imply that higher visual areas are not involved



Koene & Zhaoping 2007
Zhaoping & Zhe 2015,

Talk outline

(1) The functional role of the primary visual cortex (V1) ← **Attentional selection**

(2) In light of V1's role → a new plan to understanding vision

(3) A first example study in this new plan

Information bottlenecks in the visual pathway:

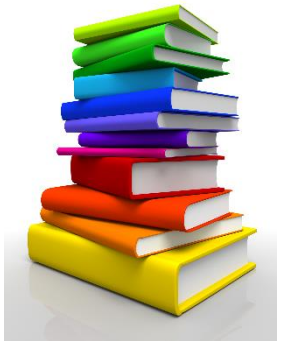
"To be or not to be,
This is the question .."

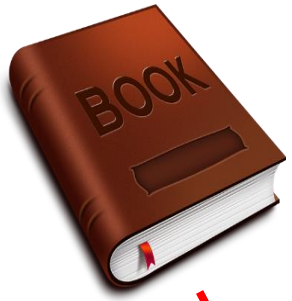
10^7 bits/second
~ 10^6 neurons ,
~10 spikes/neuron
~1 bit/spike

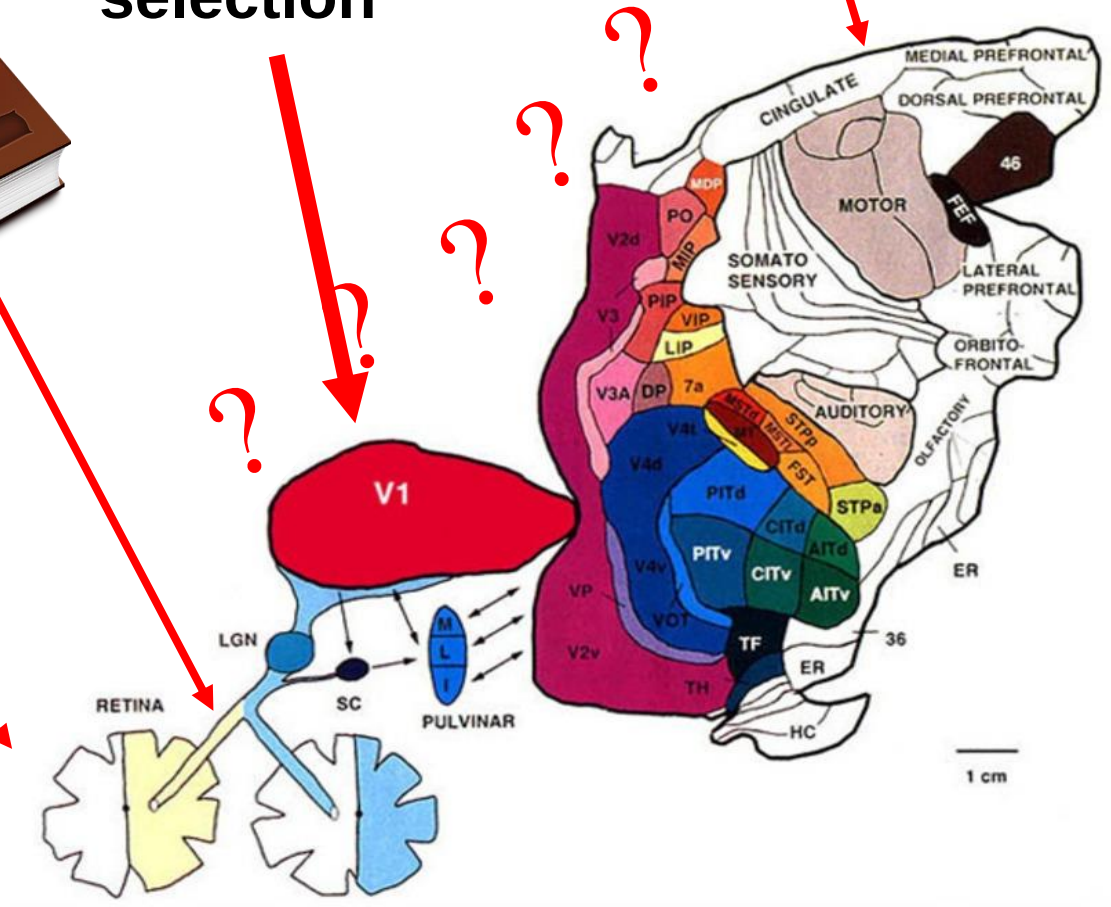
Bottom-up selection

40 bits/second
(Sziklai, 1956)

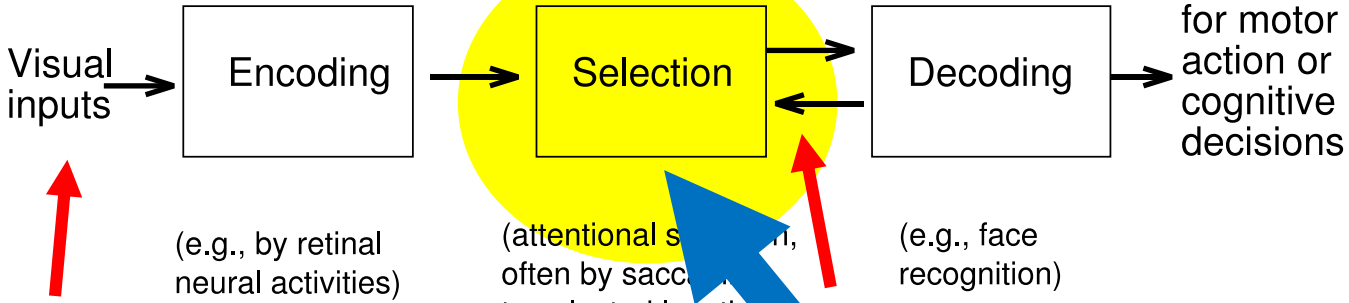
10^9 bits/second (Kelly 1962)
~ 25 frames/second,
2000x2000 pixels,
1 byte/pixel

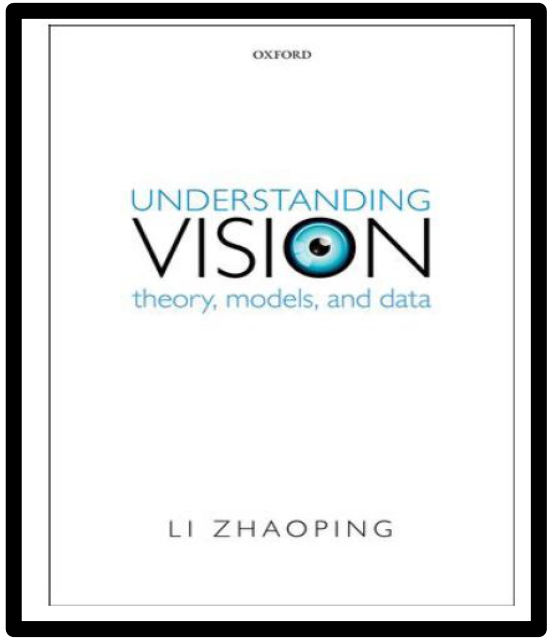






A new path to understanding vision





Traditional paths to understanding vision

(1) Low level vision, mid-level vision, high-level vision

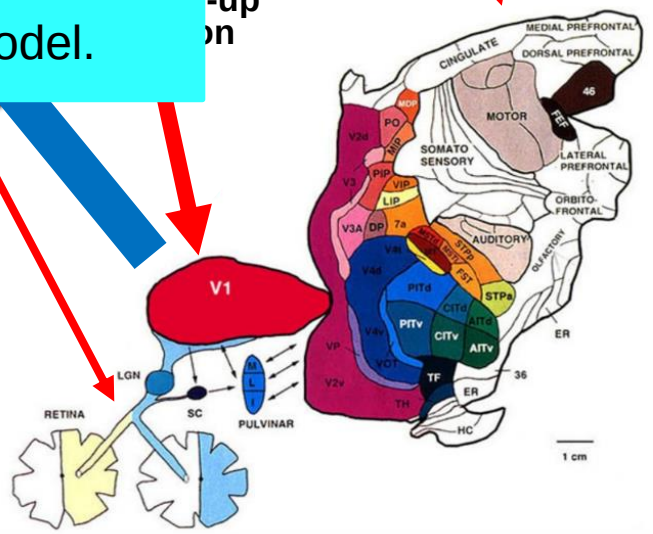
(2) David Marr: primal sketch, 2.5 d sketch, 3-d model.

"To be or not to be,
This is the question .."

40 bits/second
(Sziklai, 1956)

-up
on

(Kelly 1962)
~ 25 frames/second,
2000x2000 pixels,
1 byte/pixel



Talk outline

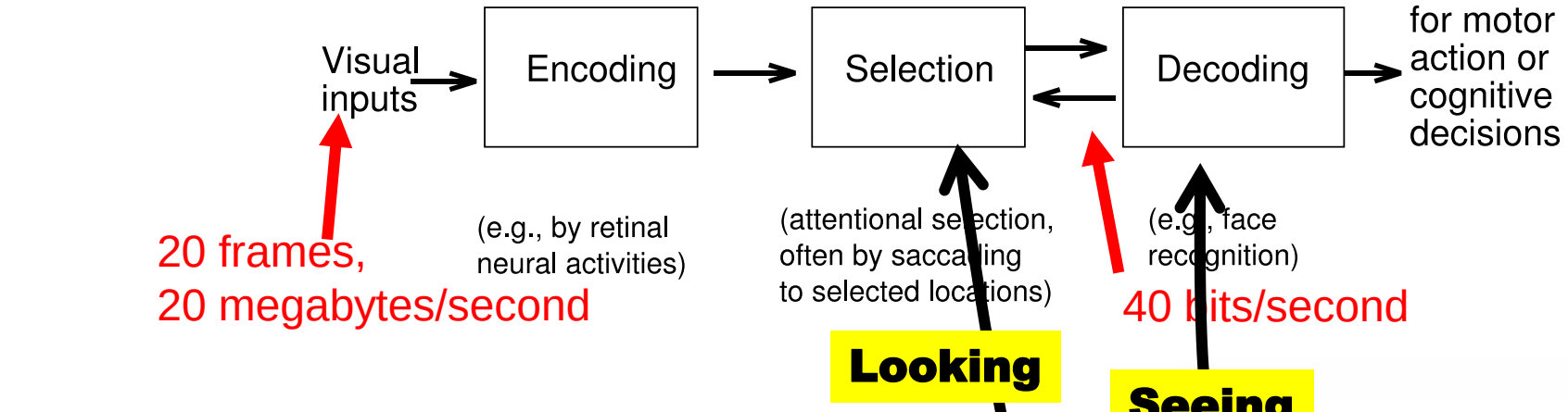
(1) The functional role of the primary visual cortex (V1) ← **Attentional selection**

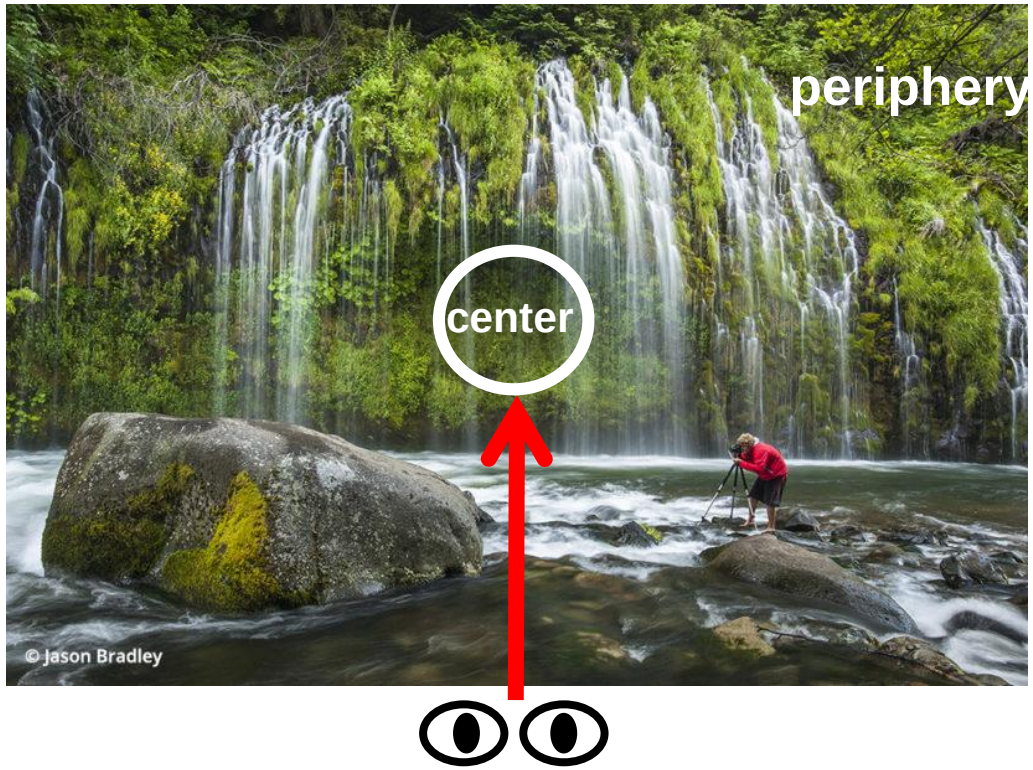
(2) In light of V1's role → a new plan to understanding vision

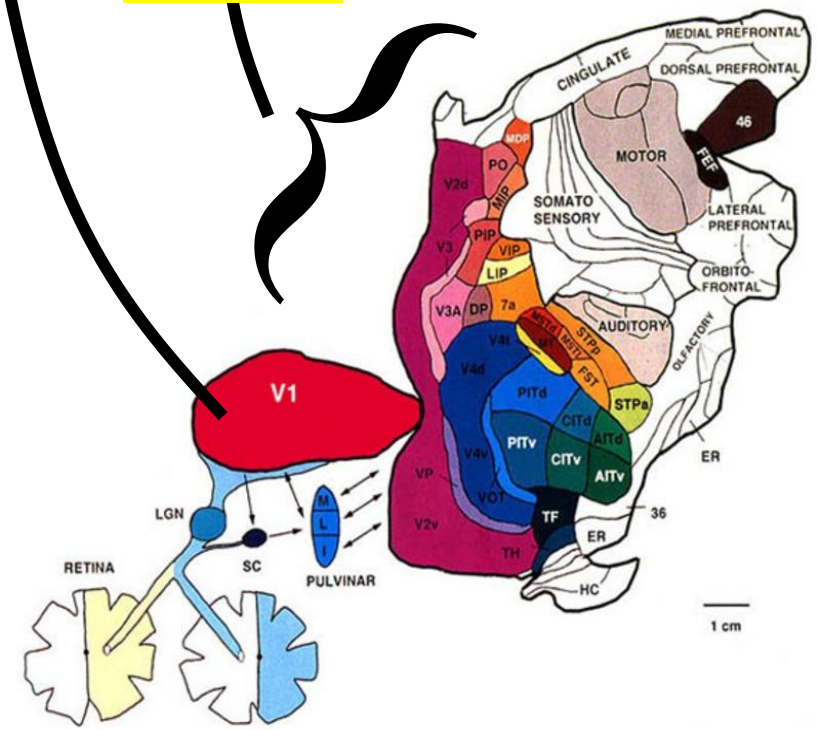
Encoding → Selection ←→ Decoding

(3) A first example study in this new plan

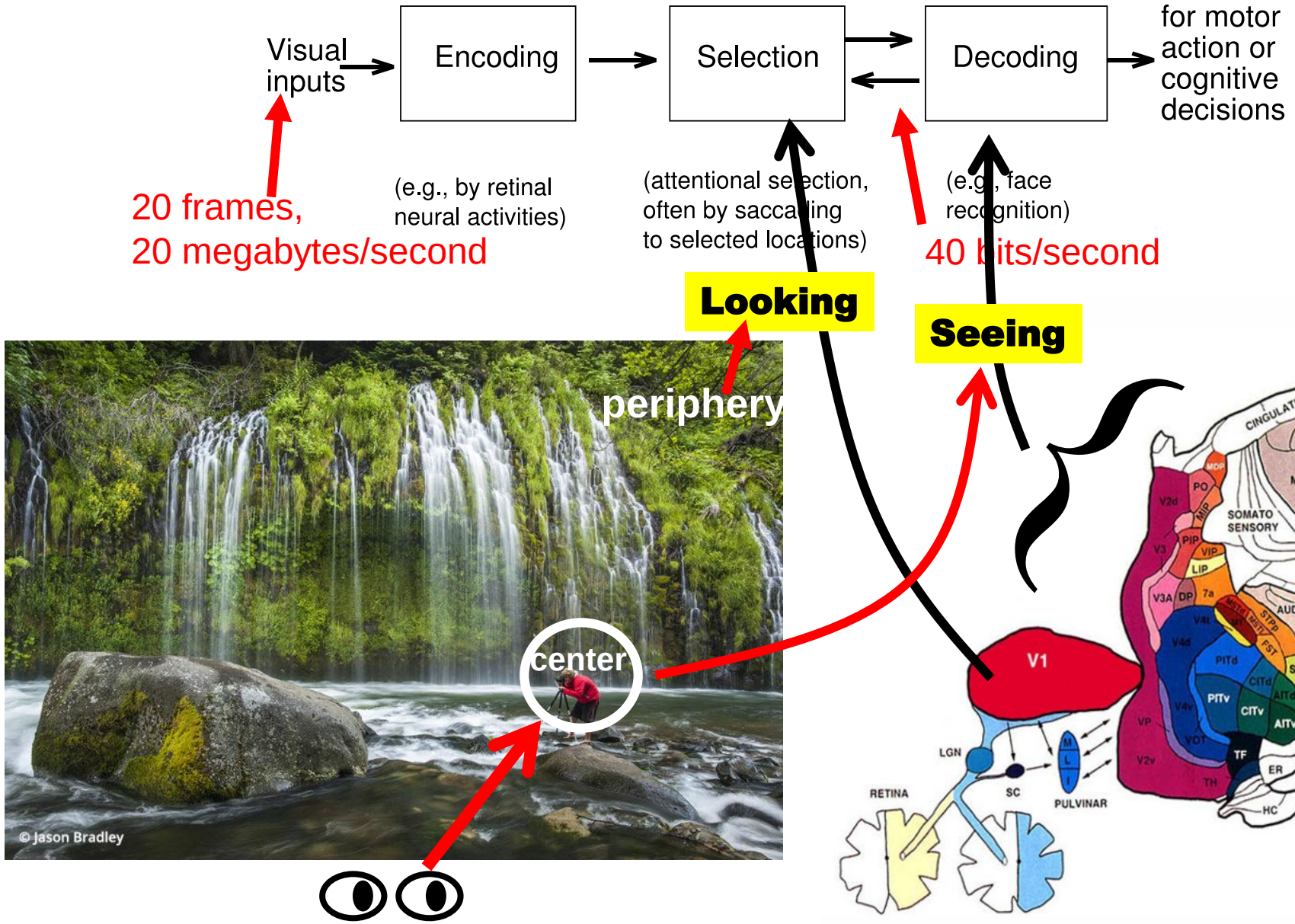
A new path to understanding vision







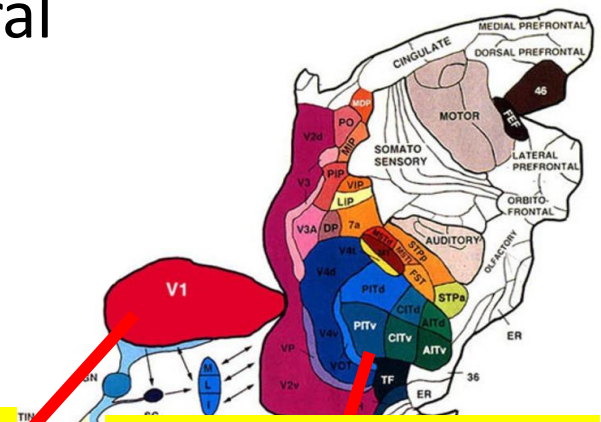
A new path to understanding vision



Looking and seeing --- peripheral and central

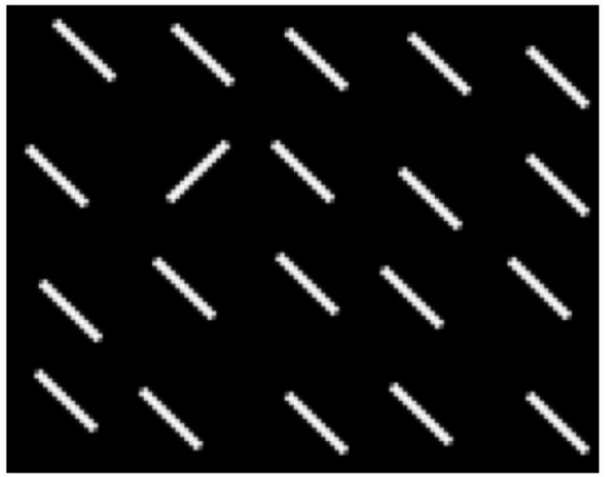
Two separate processes

Demo: (Zhaoping & Guyader 2007)

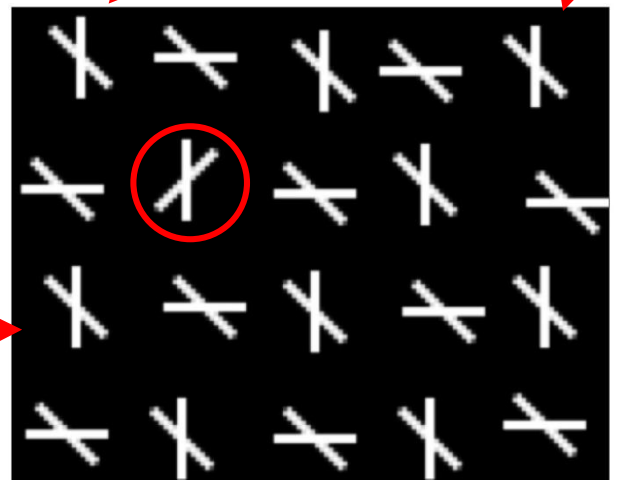


V1: tuned to primitive bars

IT & higher areas: 'X' shape recognition, rotationally invariant



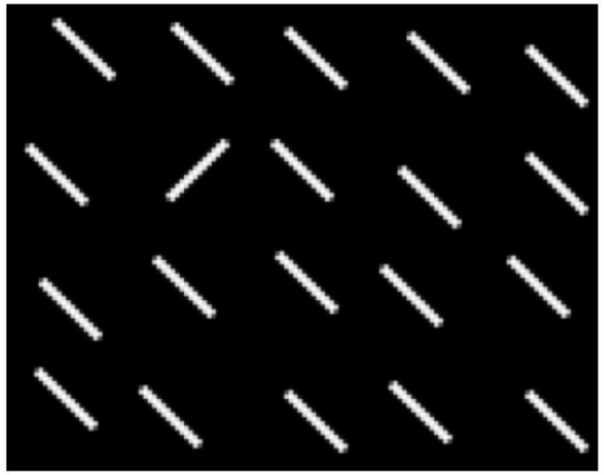
Add a horizontal bar or a vertical bar to each oblique bar



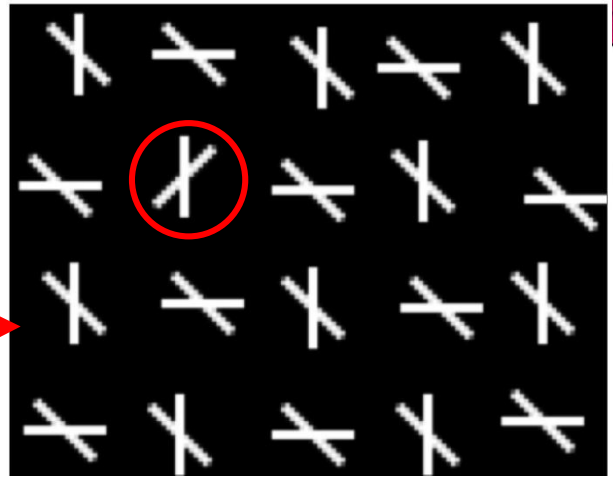
Looking and seeing --- peripheral and central

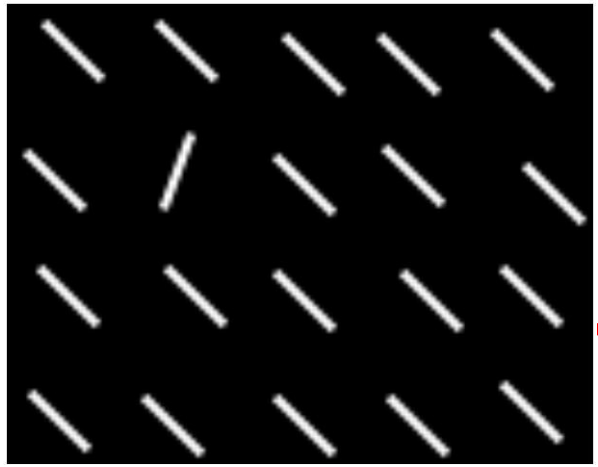
Two separate processes

Demo: (Zhaoping & Guyader 2007)

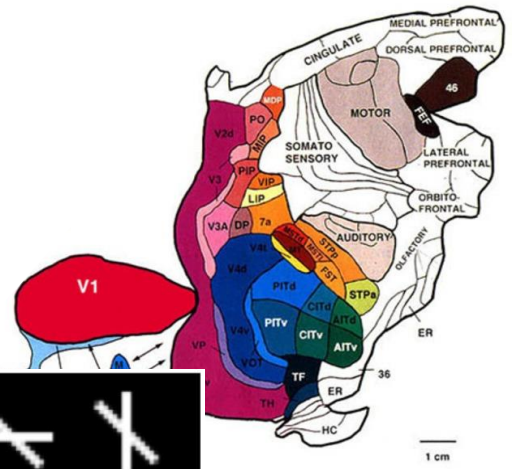


Add a horizontal bar or a vertical bar to each oblique bar

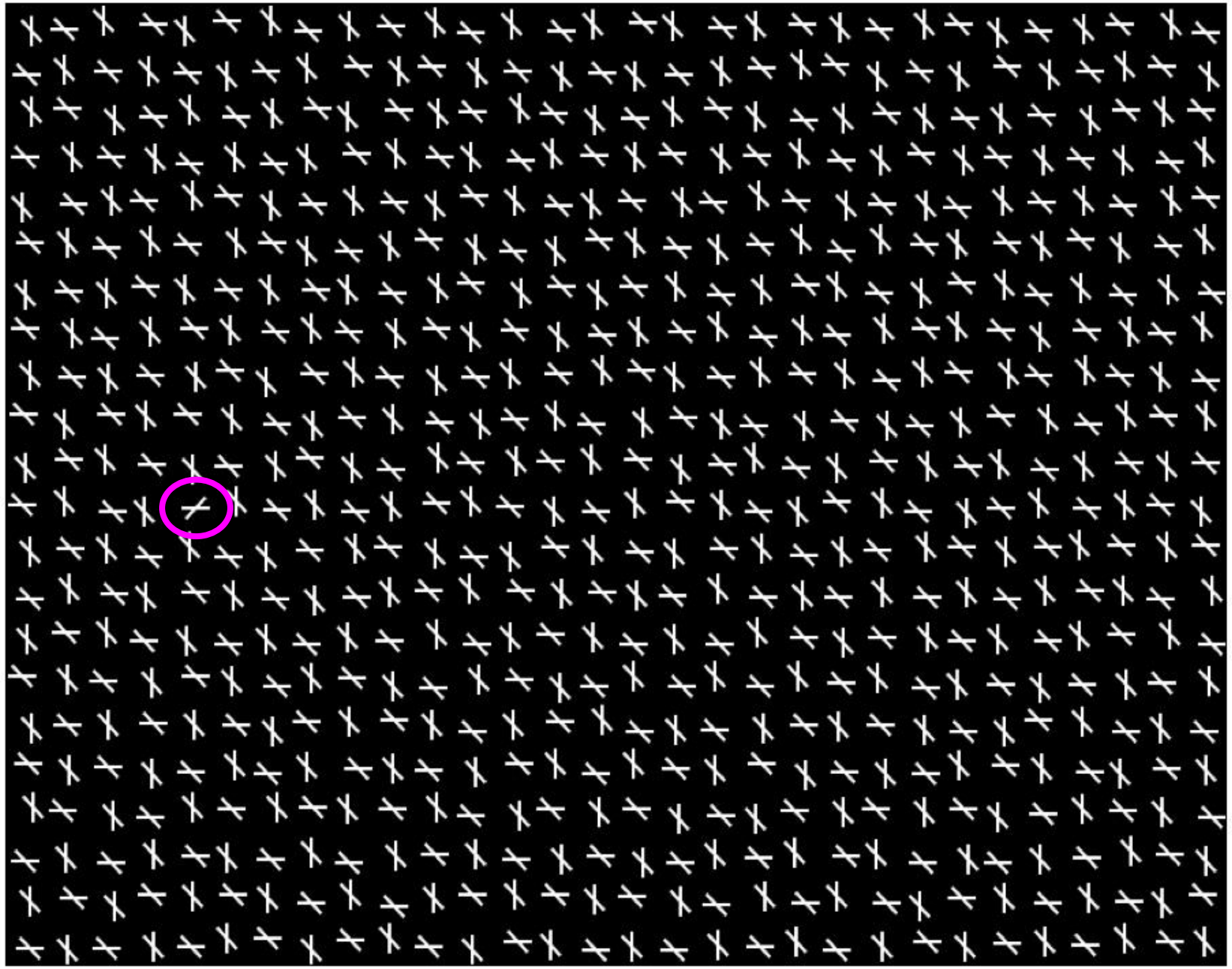




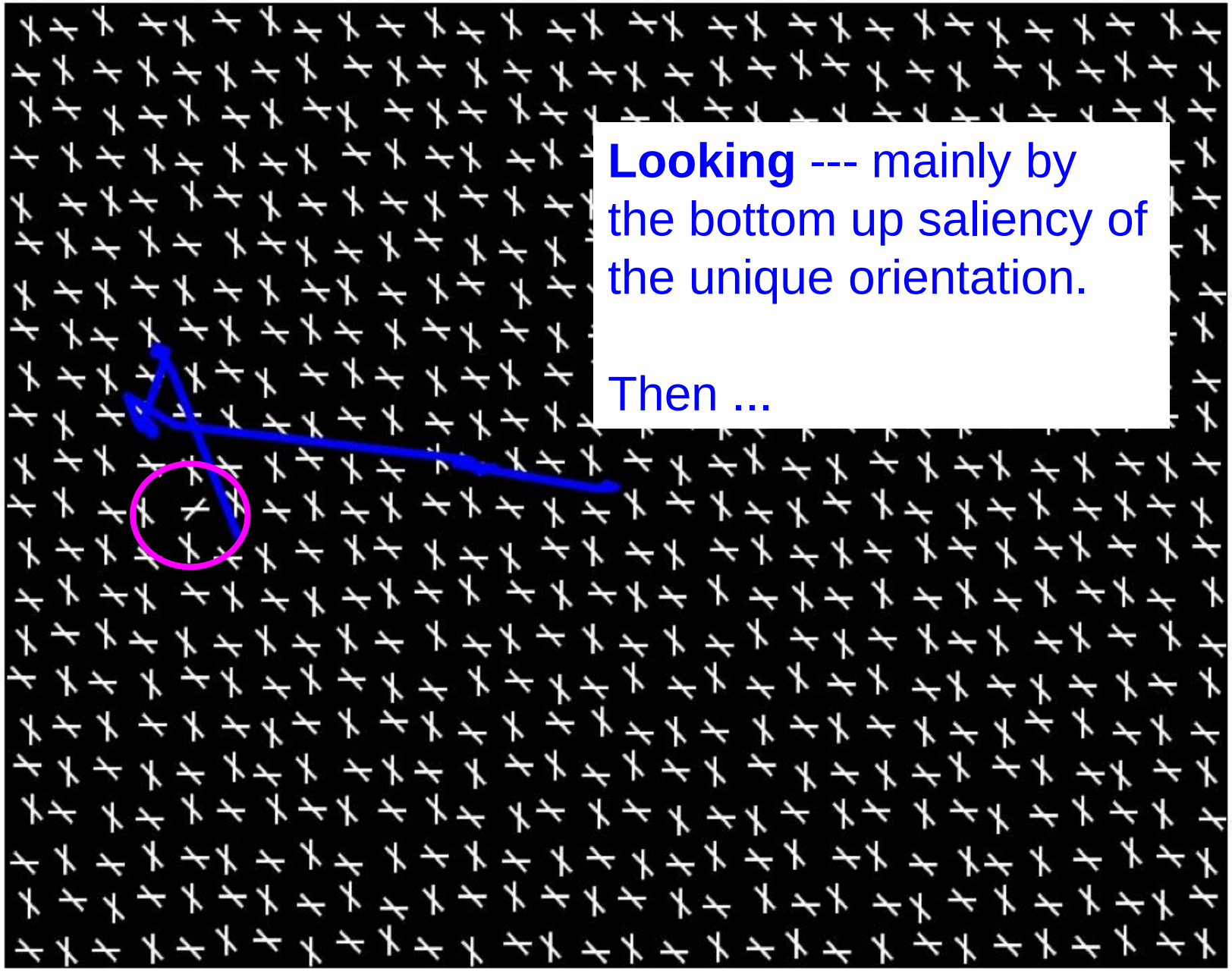
Add a horizontal bar or a vertical bar to each oblique bar



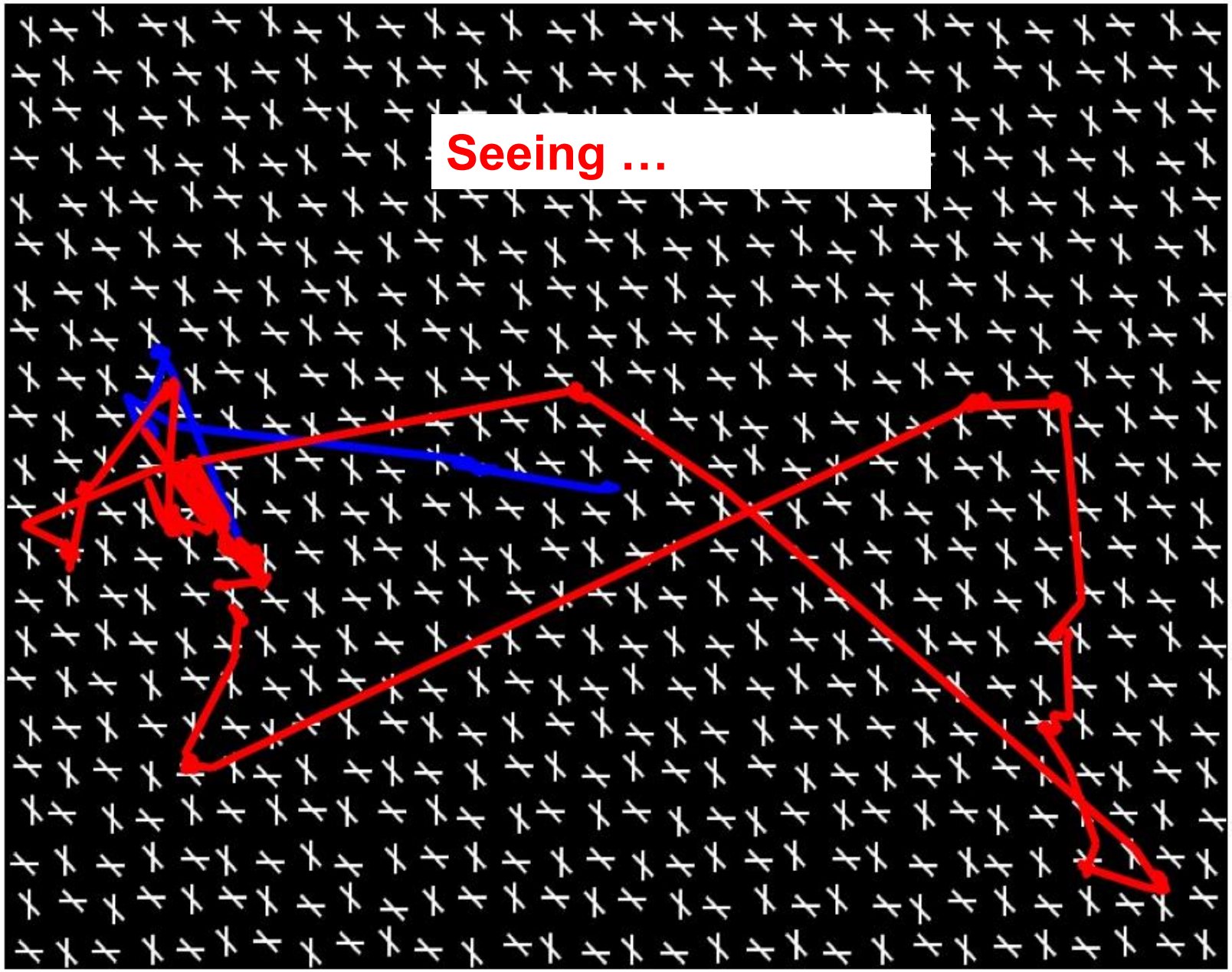
Display span 46x32 degrees in visual angle --- condition A



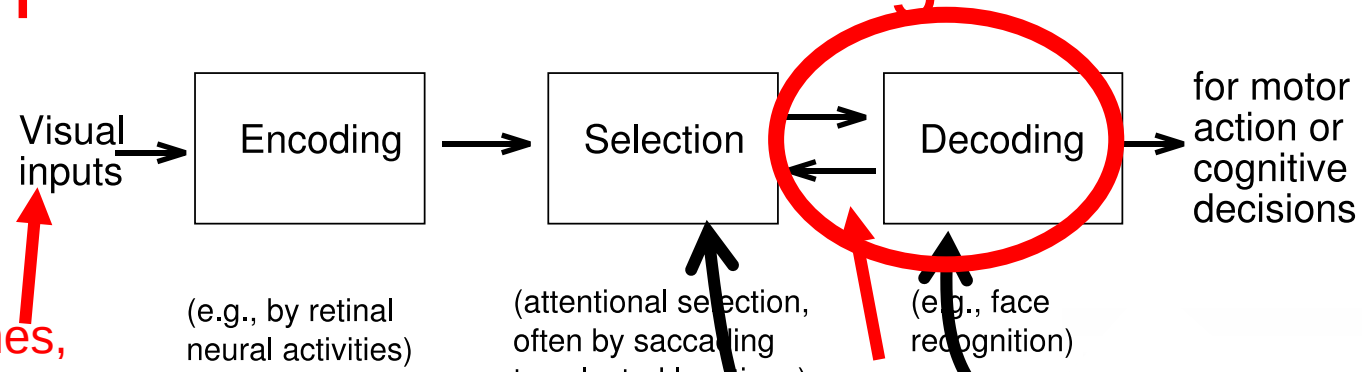
Gaze arrives at target after a few saccades

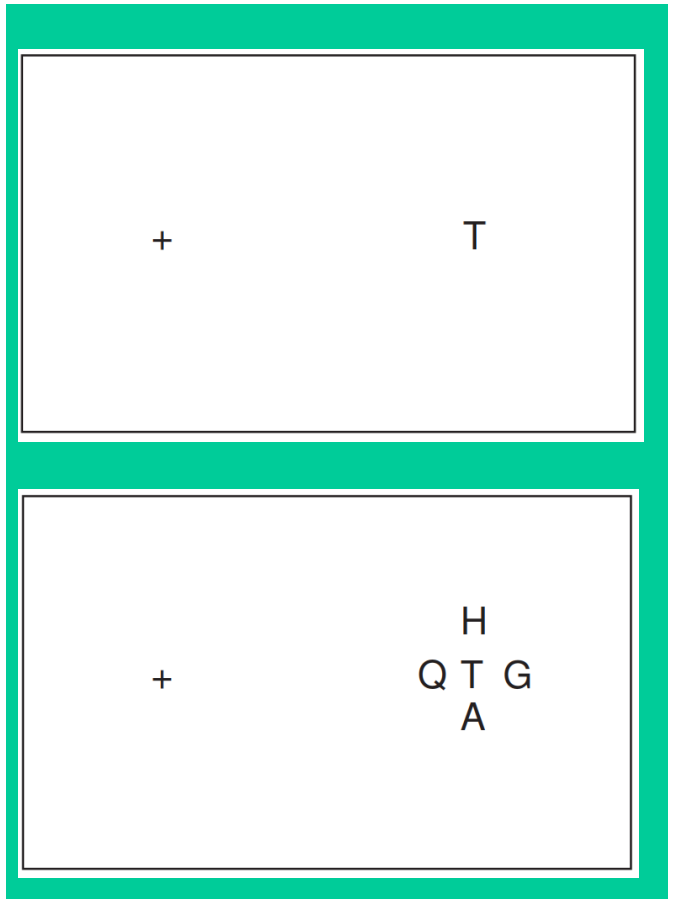


Gaze dawdled around the target, then abandoned and returned.



A new path to understanding vision





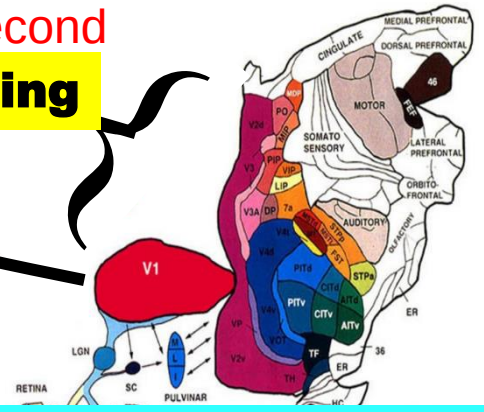
Looking

Seeing

Peripheral vision

Must qualitatively differ in Seein

Central visior



Computational algorithms?

Focus on feedforward-feedback / V1

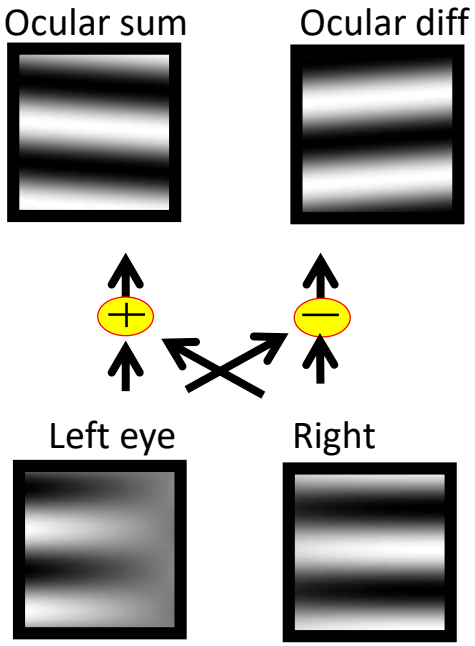
A demo of crowding in the periphery

Summation percept dominant

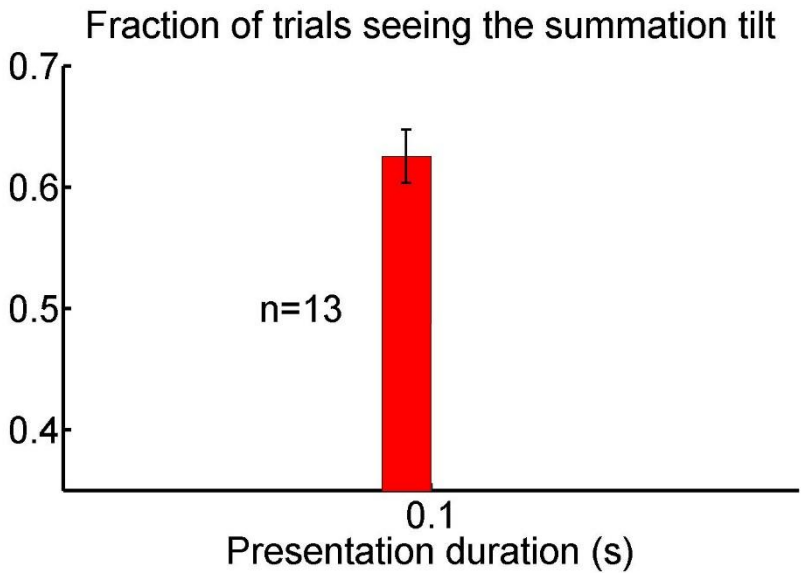
Perception = ?

In V1, signals are efficiently encoded by these two de-correlated channels
(Li & Atick 1994)

$C_+ = C_-$



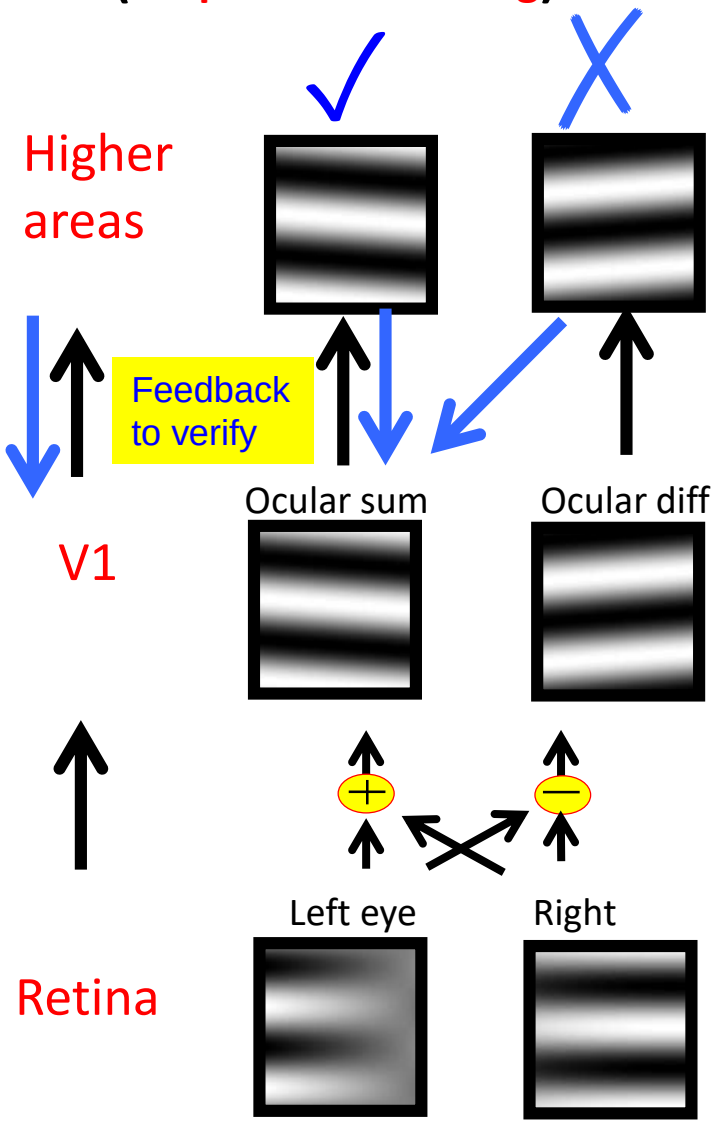
Subject task: report the perceived tilt.



Zhaoping 2017

Why does perception prefer ocular summation? (Zhaoping 2017)

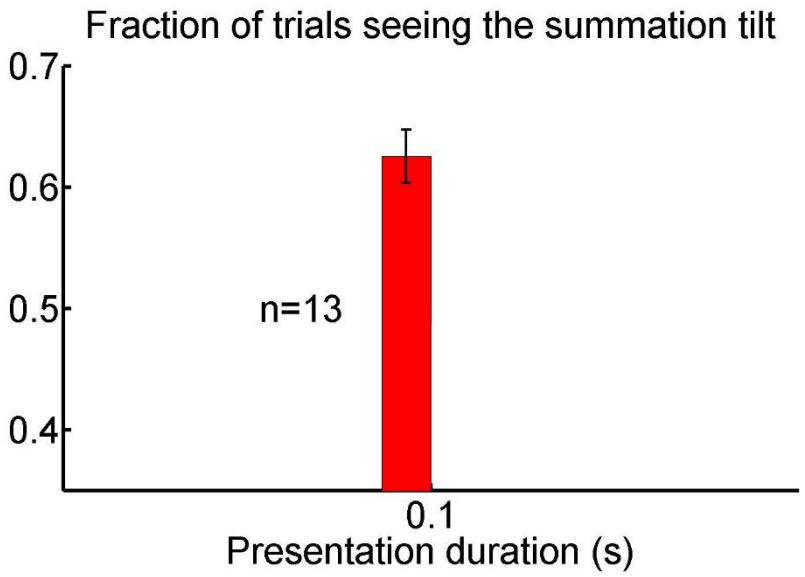
for analysis-by-synthesis
(c.f. predictive coding)



Feedforward, feedback, verify, and re-weight (FFVW)

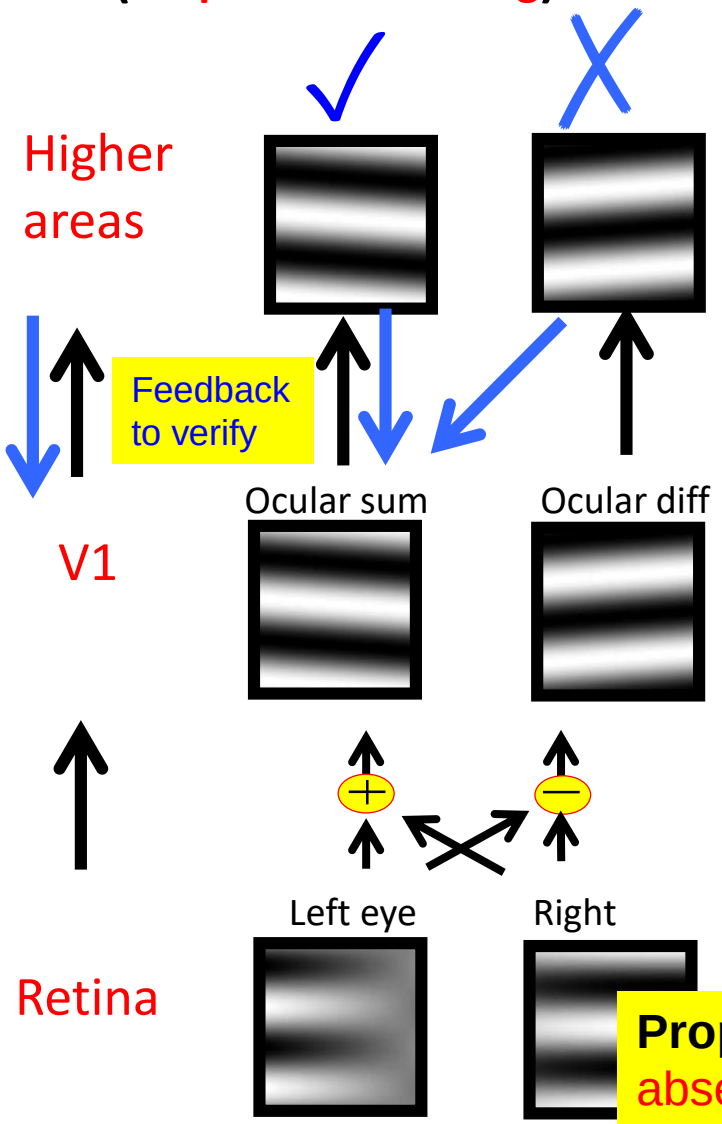
If I perceive it, it is likely (prior) shown to both my left and my right eyes,
so it should resemble the input in the sum channel!!!

The Bayesian(?) monster



Why does perception prefer ocular summation? (Zhaoping 2017)

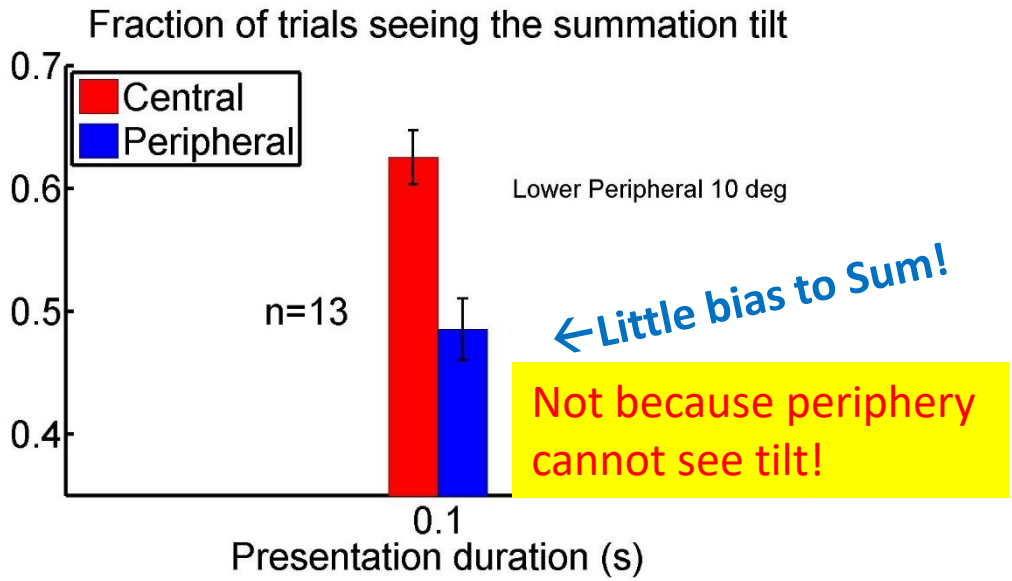
for analysis-by-synthesis
(c.f. predictive coding)



Feedforward, feedback, verify, and re-weight (FFVW)

If I perceive it, it is likely (prior) shown to both my left and my right eyes,
so it should resemble the input in the sum channel!!!

The Bayesian(?) monster



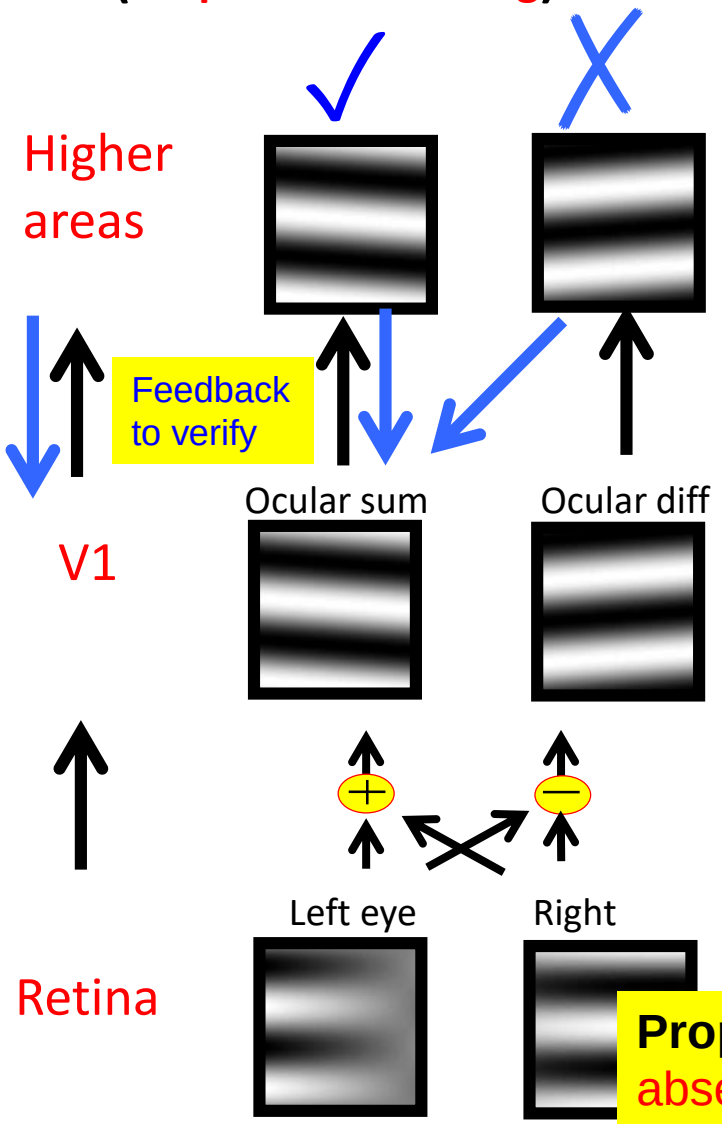
←Little bias to Sum!

Not because periphery cannot see tilt!

Proposal: Top-down feedback to V1 is weaker or absent in peripheral vision for analysis-by-synthesis!

Why does perception prefer ocular summation? (Zhaoping 2017)

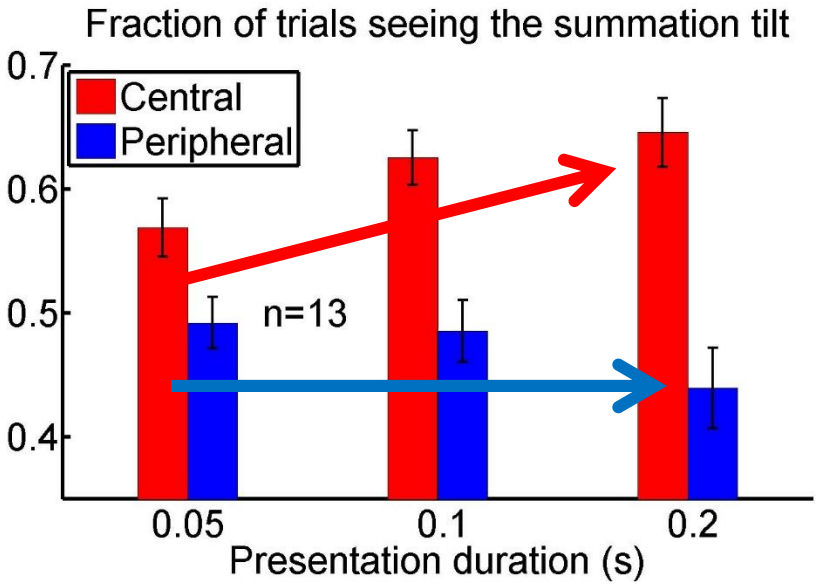
for analysis-by-synthesis
(c.f. predictive coding)



Feedforward, feedback, verify, and re-weight (FFVW)

If I perceive it, it is likely (prior) shown to both my left and my right eyes,
so it should resemble the input in the sum channel!!!

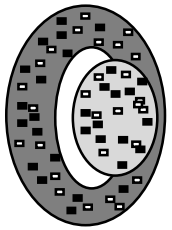
The Bayesian(?) monster



Proposal: Top-down feedback to V1 is weaker or absent in peripheral vision for analysis-by-synthesis!

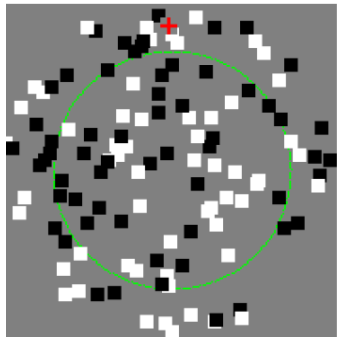
Testing it in depth perception

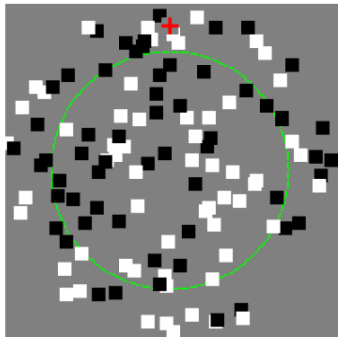
Zhaoping & Ackermann, 2018



A central disk (non-zero disparity)
and a surrounding ring (zero disparity)

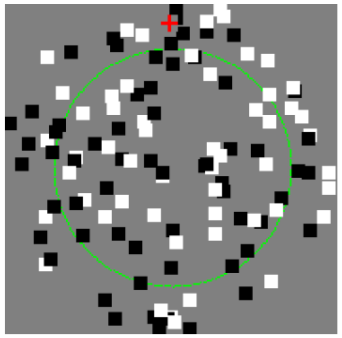
Dots for the central disk correlated

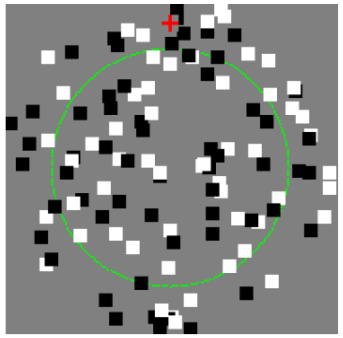




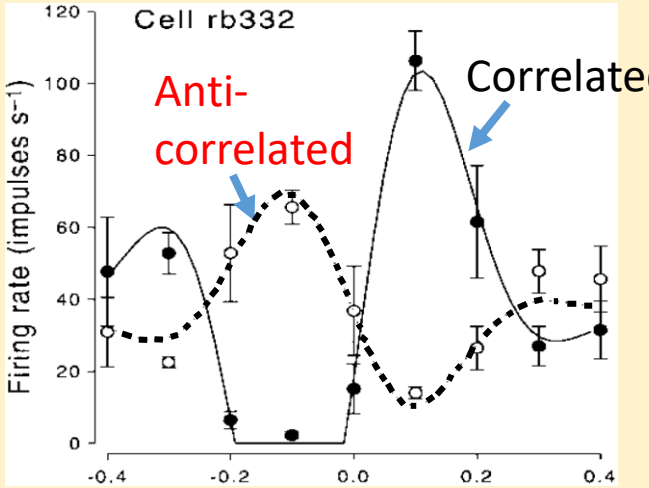
(Green circle not part of the stimuli)

Dots for the central disk Anti-correlated





A V1 neuron's response to disparity in random dot stereograms



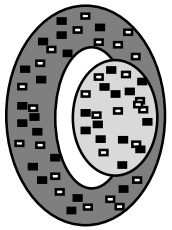
(Cumming and Parker 1997)

Doi et al 2011

Proposal: Top-down feedback to V1 is weaker or absent in peripheral vision for analysis-by-synthesis!

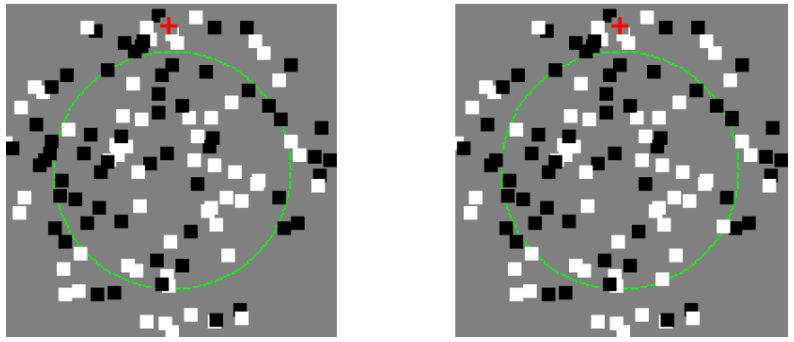
Proposal: Top-down feedback to V1 is weaker or absent in peripheral vision for analysis-by-synthesis!

Testing it in depth perception
Zhaoping & Ackermann, 2018

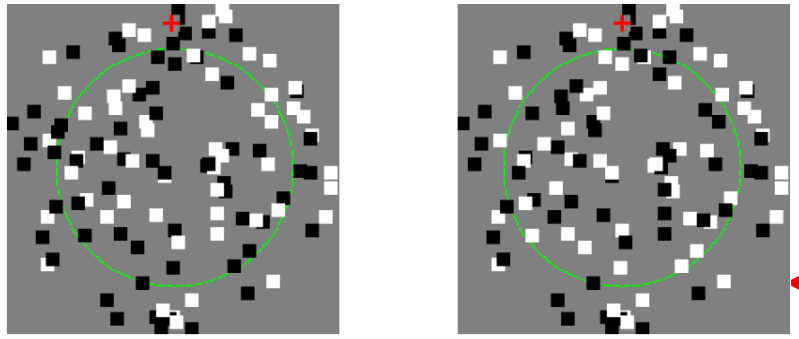


A central disk (non-zero disparity)
and a surrounding ring (zero disparity)

Dots for the central disk correlated



Dots for the central disk Anti-correlated



No reversed depth percept

V1's report vetoed

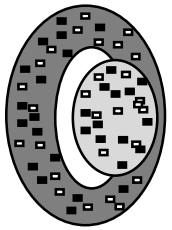
Expected disparity and binocular correlation not found in V1

Top-down feedback to verify V1's report

V1 feeds forward reverse depth to higher brain areas!

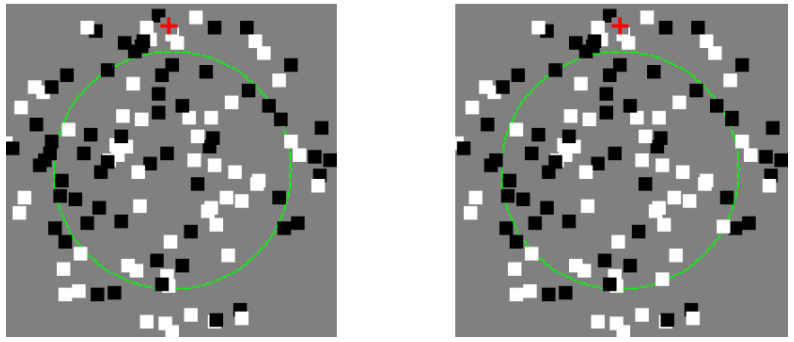
Proposal: Top-down feedback to V1 is weaker or absent in peripheral vision for analysis-by-synthesis!

Testing it in depth perception
Zhaoping & Ackermann, 2018



A central disk (non-zero disparity)
and a surrounding ring (zero disparity)

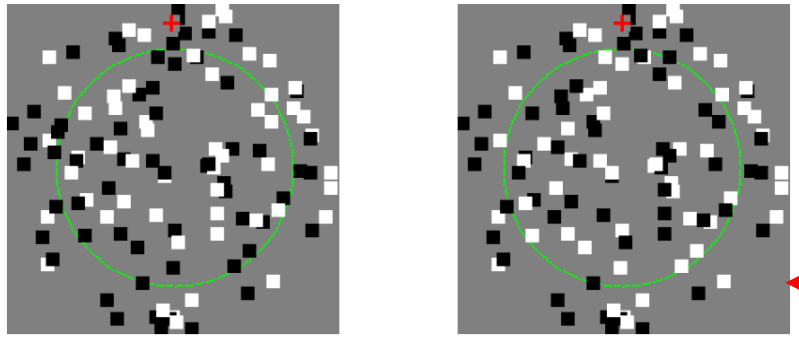
Dots for the central disk correlated



~~No reversed depth percept~~

If peripheral vision has no feedback

Dots for the central disk Anti-correlated

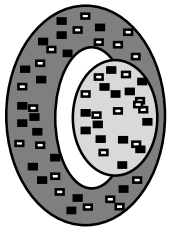


V1 feeds forward reverse depth to higher brain areas!

Proposal: Top-down feedback to V1 is weaker or absent in peripheral vision for analysis-by-synthesis!

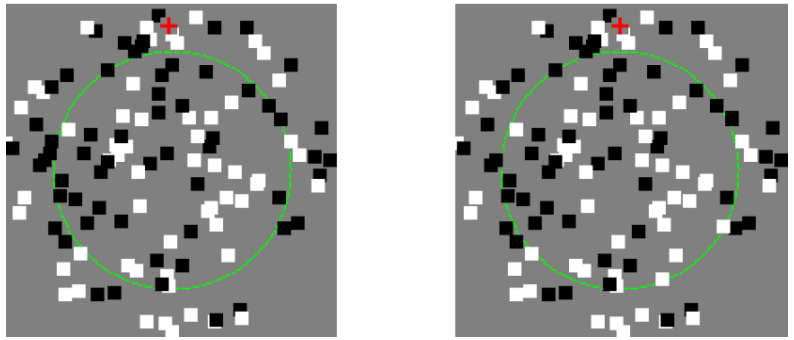
Testing it in depth perception

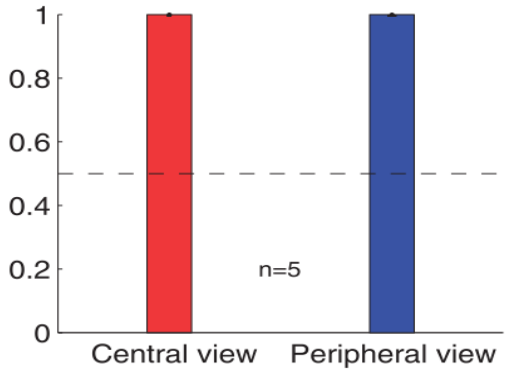
Zhaoping & Ackermann, 2018



A central disk (non-zero disparity) and a surrounding ring (zero disparity)

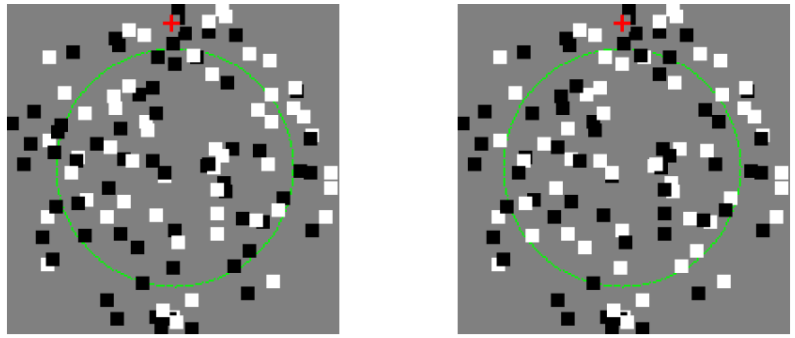
Dots for the central disk correlated

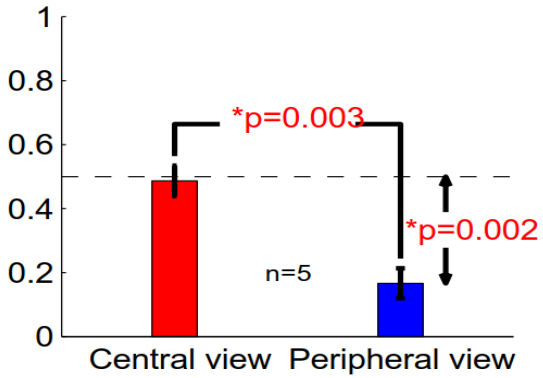




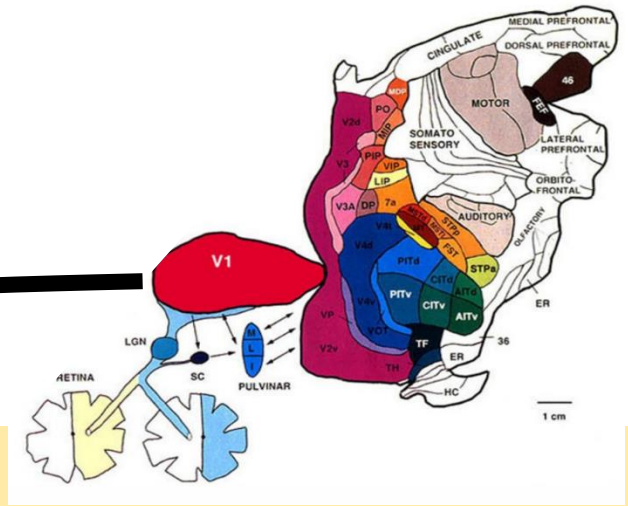
Accuracy reporting disparity-defined depth

Dots for the central disk Anti-correlated

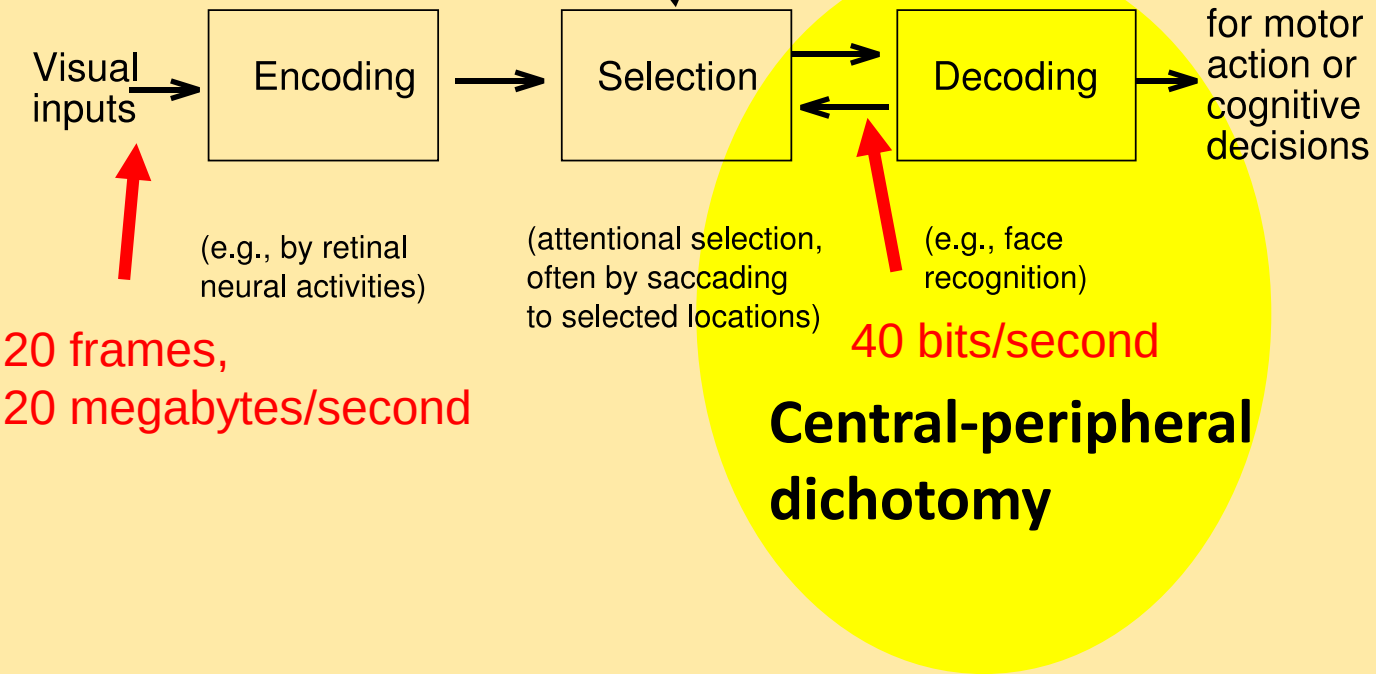




Thanks to
Funding source: Gatsby Foundation, EPSRC, BBSRC.
Collaborators: Wu Li, Yin Yan, Ansgar Koene, Li Zhe, Joealle Ackermann, etc



A new path to understanding vision



Falsifiable

Opening the window to our brain