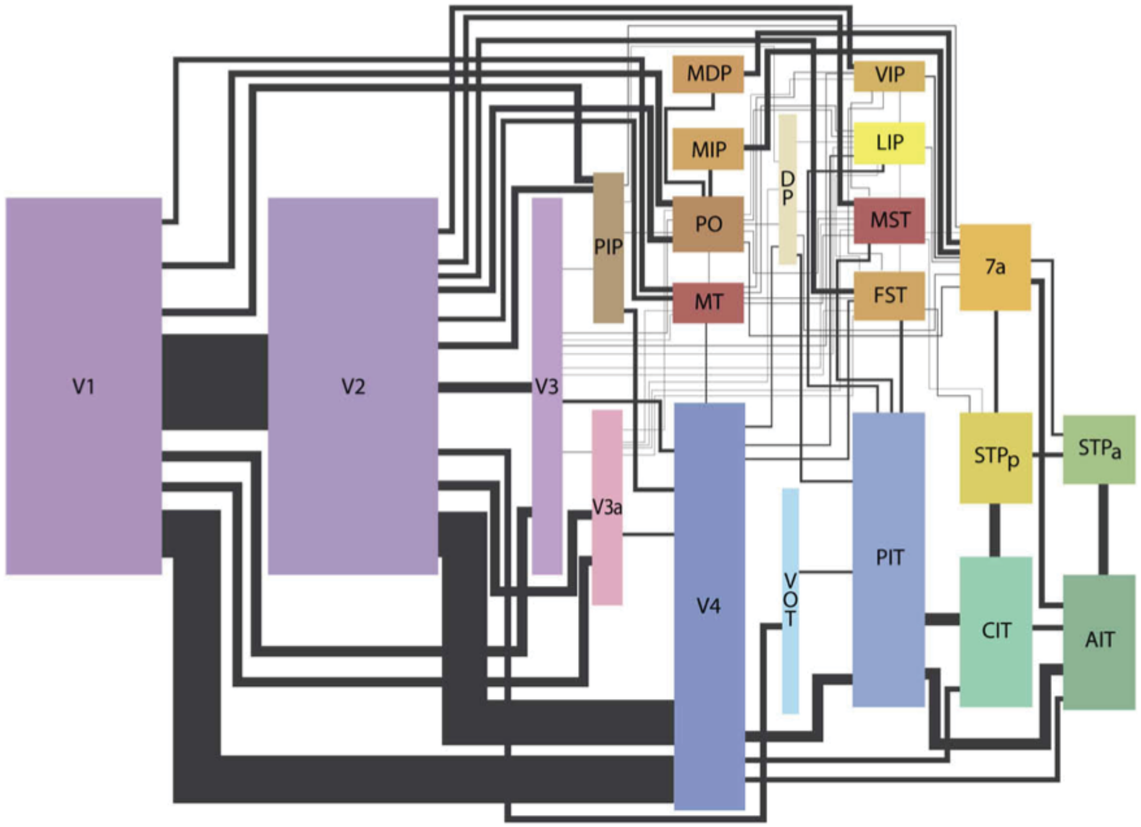
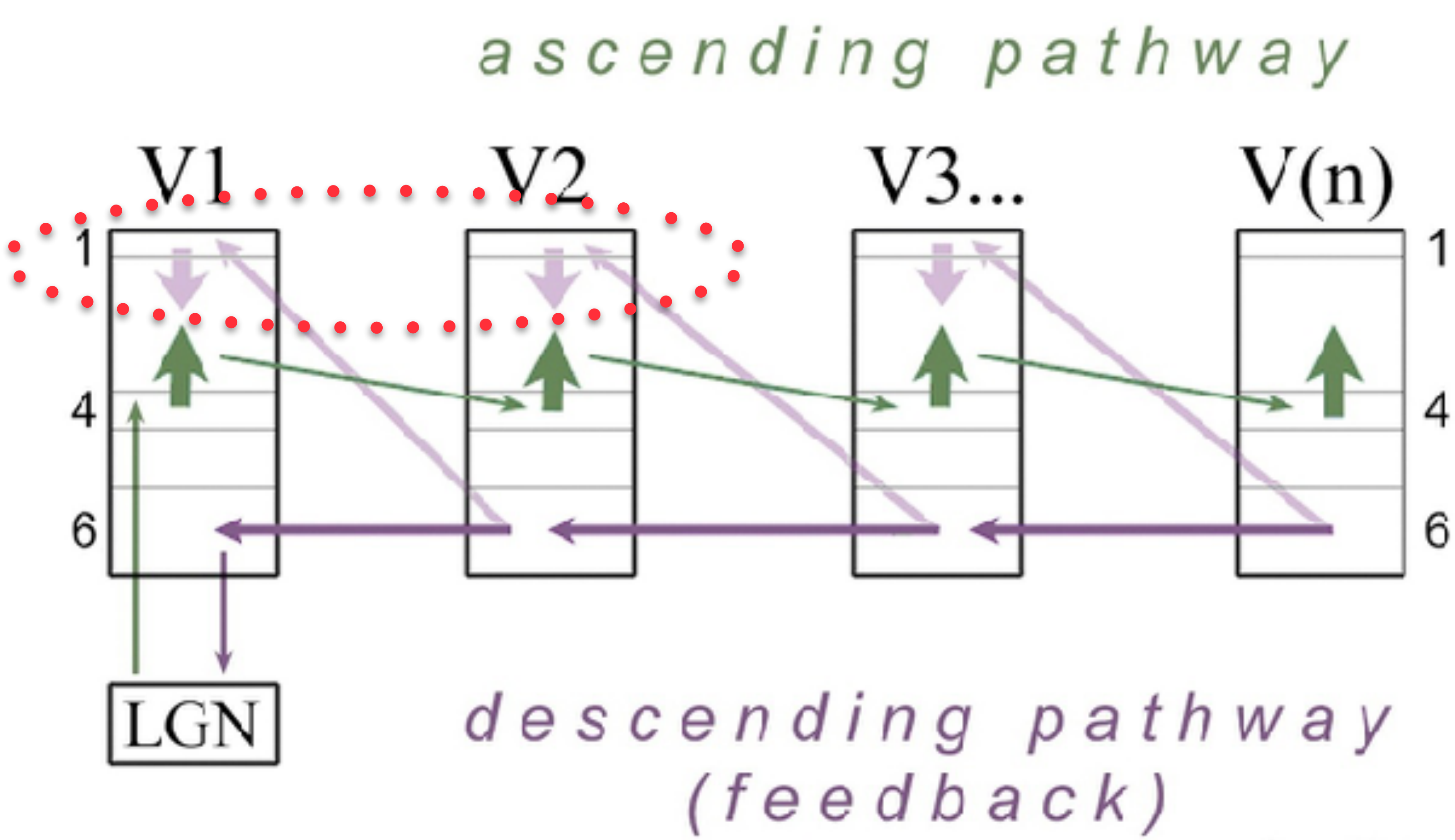
Bidirectional processing III:

feedforward & feedback networks for object perception

Focus on empirical studies in humans





Current Biology

connection to Bayes

$$p(S|I) = \frac{p(I|S)p(S)}{p(I)}$$

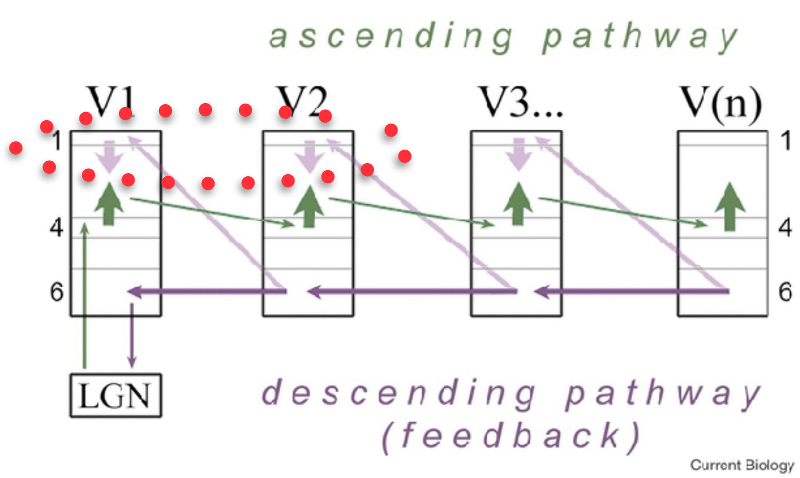
$$p(S|I) \propto p(I - f(S))p(S)$$

does the visual system use built-in knowledge of how images are naturally generated to predict the input I, based on candidate "explanations" f(S)?

If so, such a mechanism could be used to test and sort through competing explanations

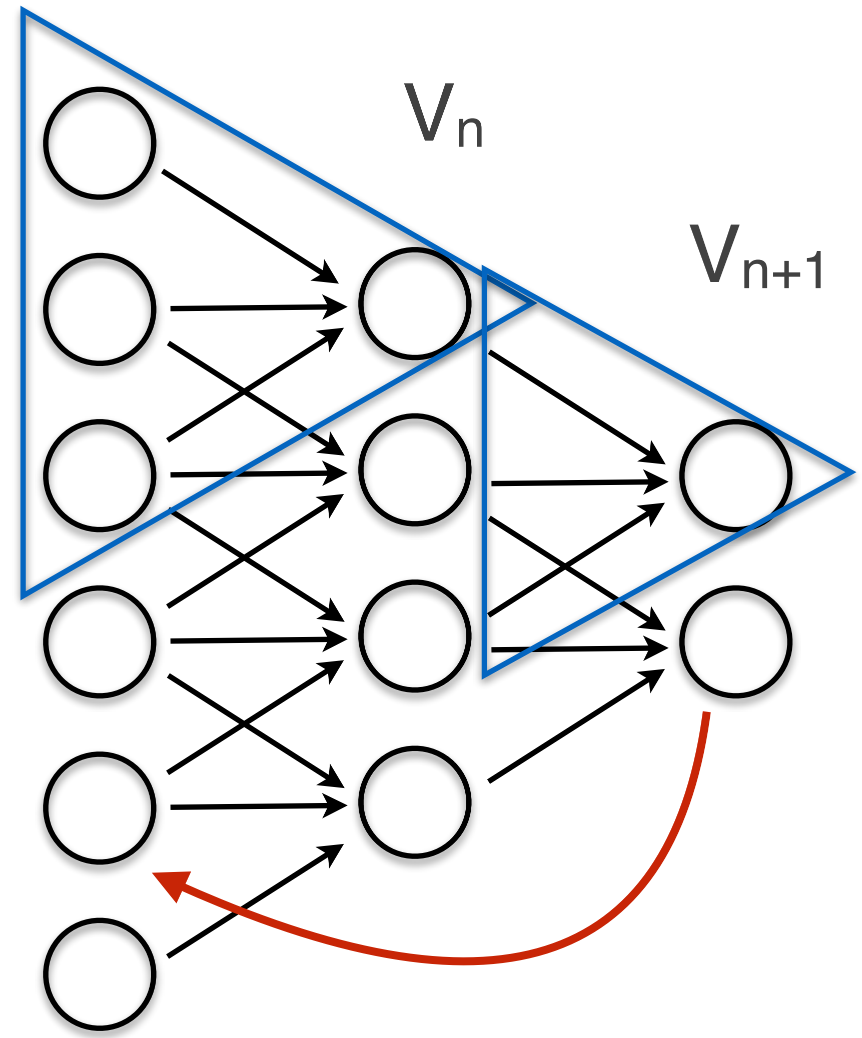
Evidence for the visual system to do anything like this?

How can one study feedback in humans? Psychophysics? Large-scale imaging?

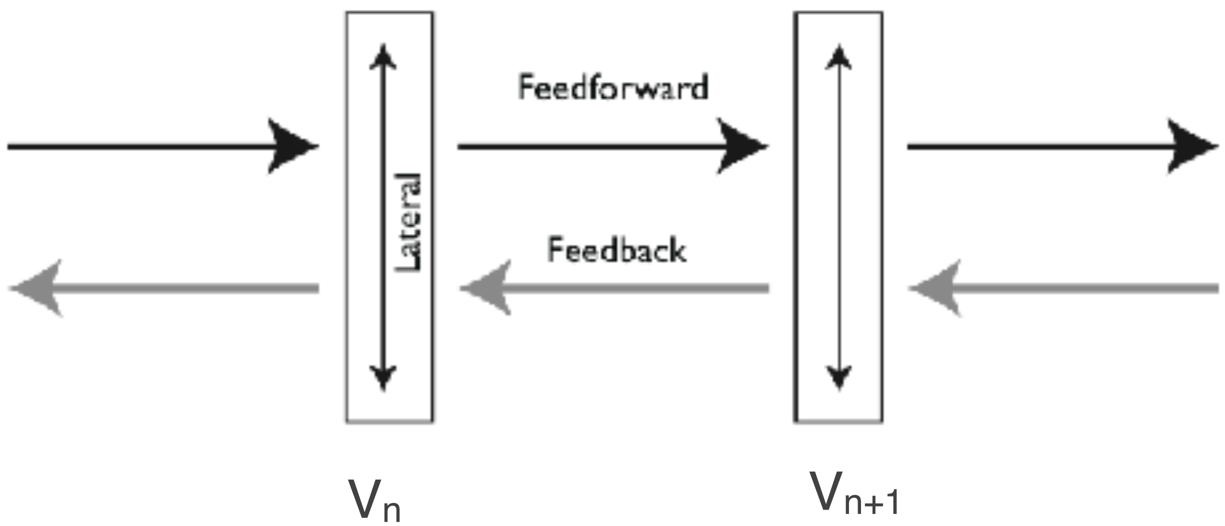


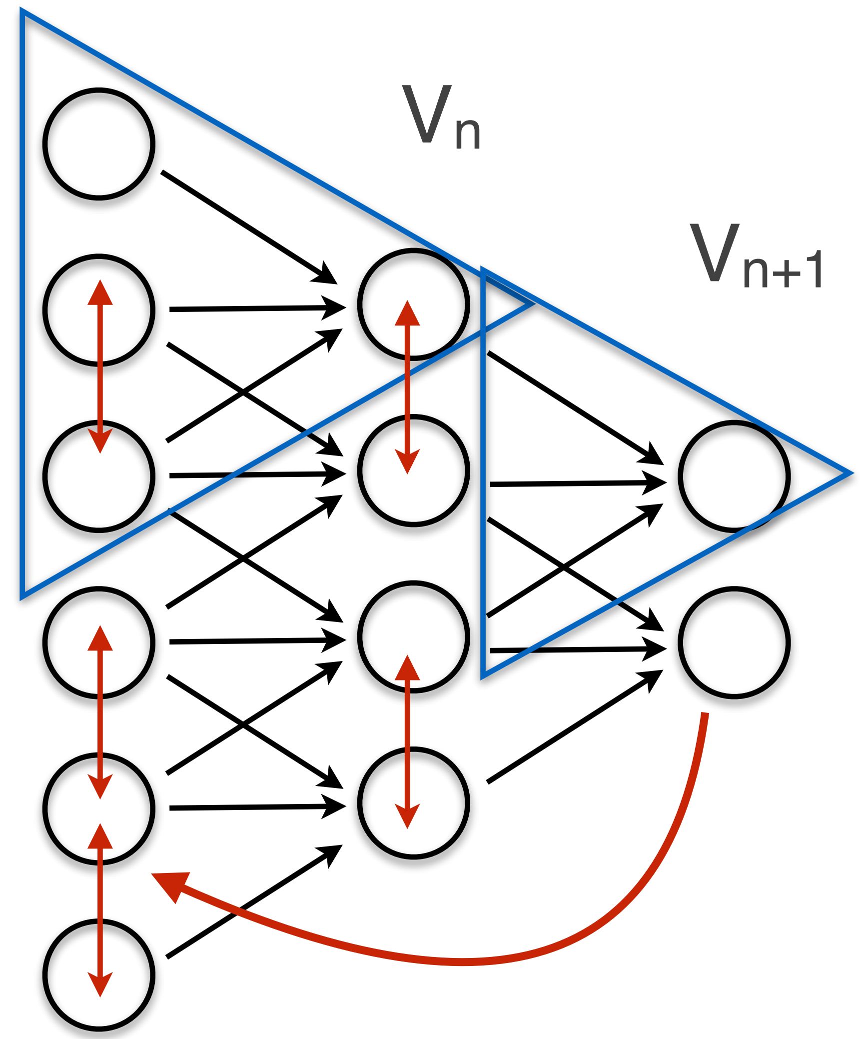
take advantage of the hierarchical structure of visual cortical areas

look for effects of spatial context on early, local processing

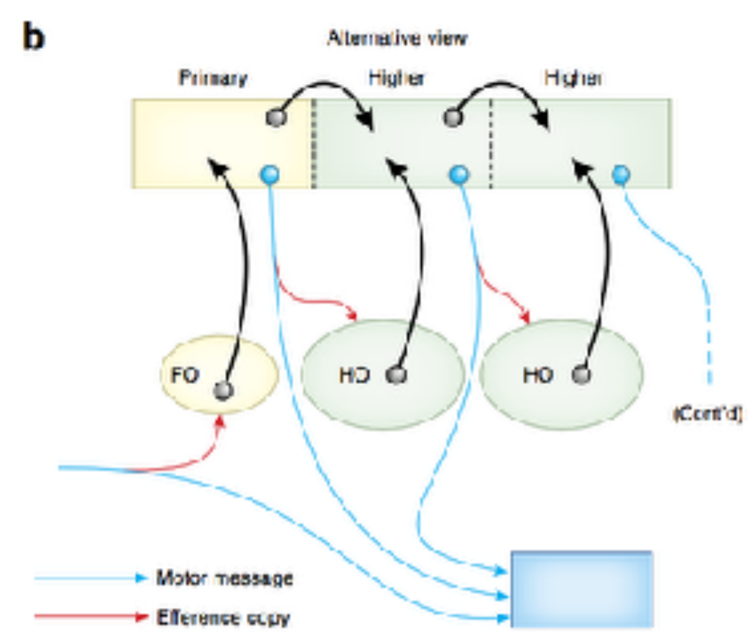


…some caveats





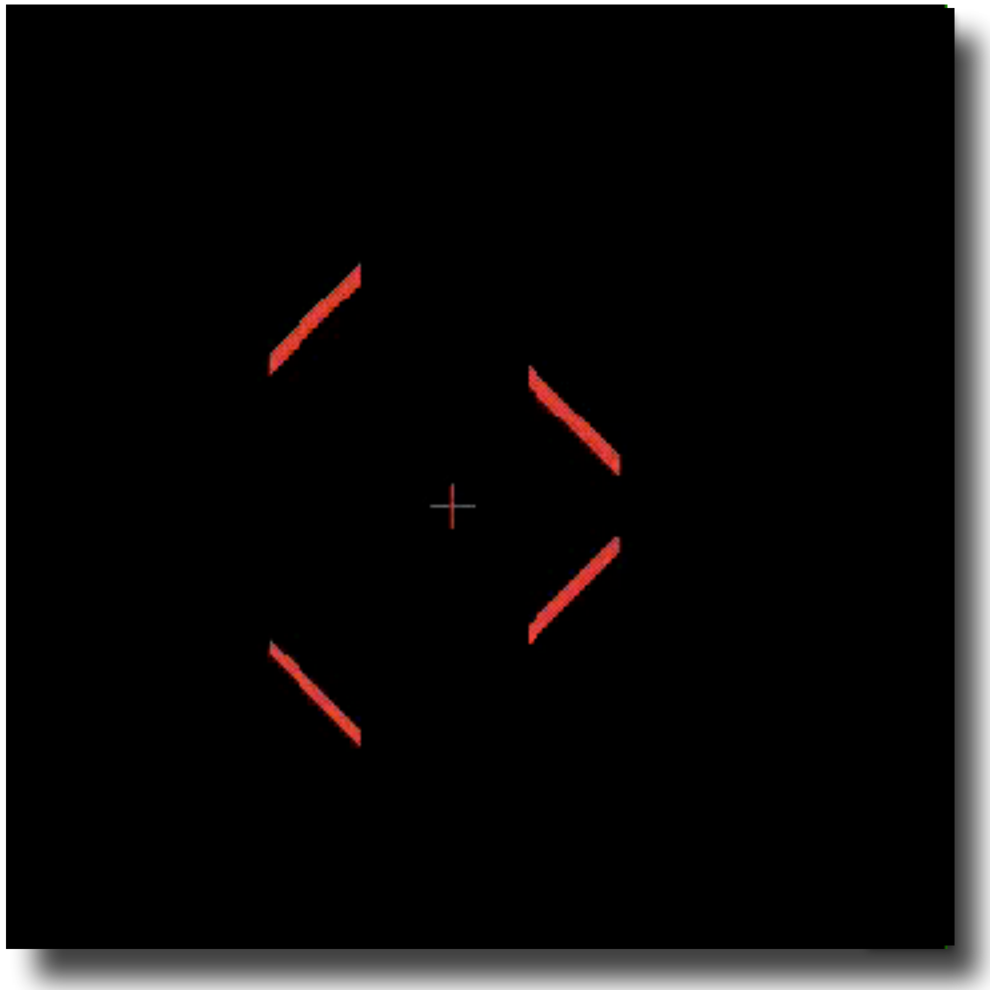
contextual information can be integrated feedforward, laterally within an area, and through feedback

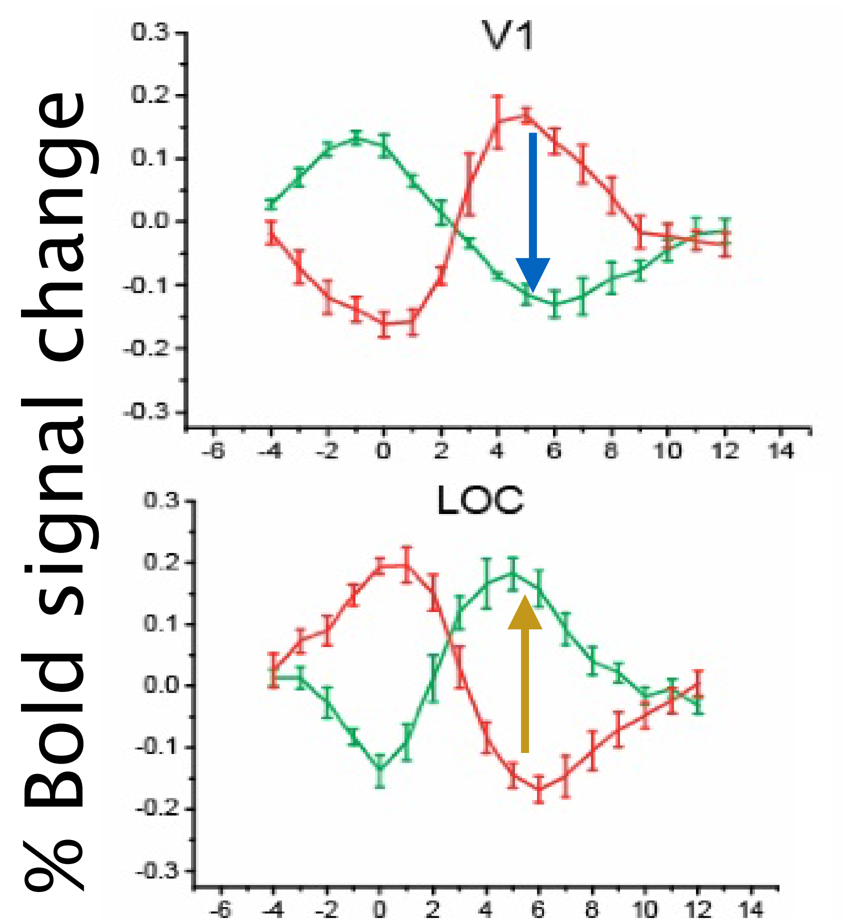


..and the elephant in the room

Sherman and Guillery

fMRI activity in V1



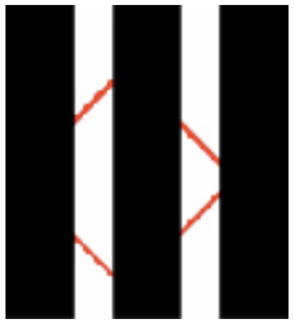


Diamond shape perceived

Line fragments perceived

V1 activity decreases when the diamond shape is perceived

one of the perceptual states - a "diamond" shape



LOC—a high-level object area— activity is increases when the diamond shape is perceived

Murray, Kersten, Olshausen, Schrater, & Woods (2002)

Fang, Boyaci, Kersten, Murray (2008)

But is the modulation of low-level activity localized to early feature detectors?

But is the suppression localized to cortical regions corresponding to the features and properties?

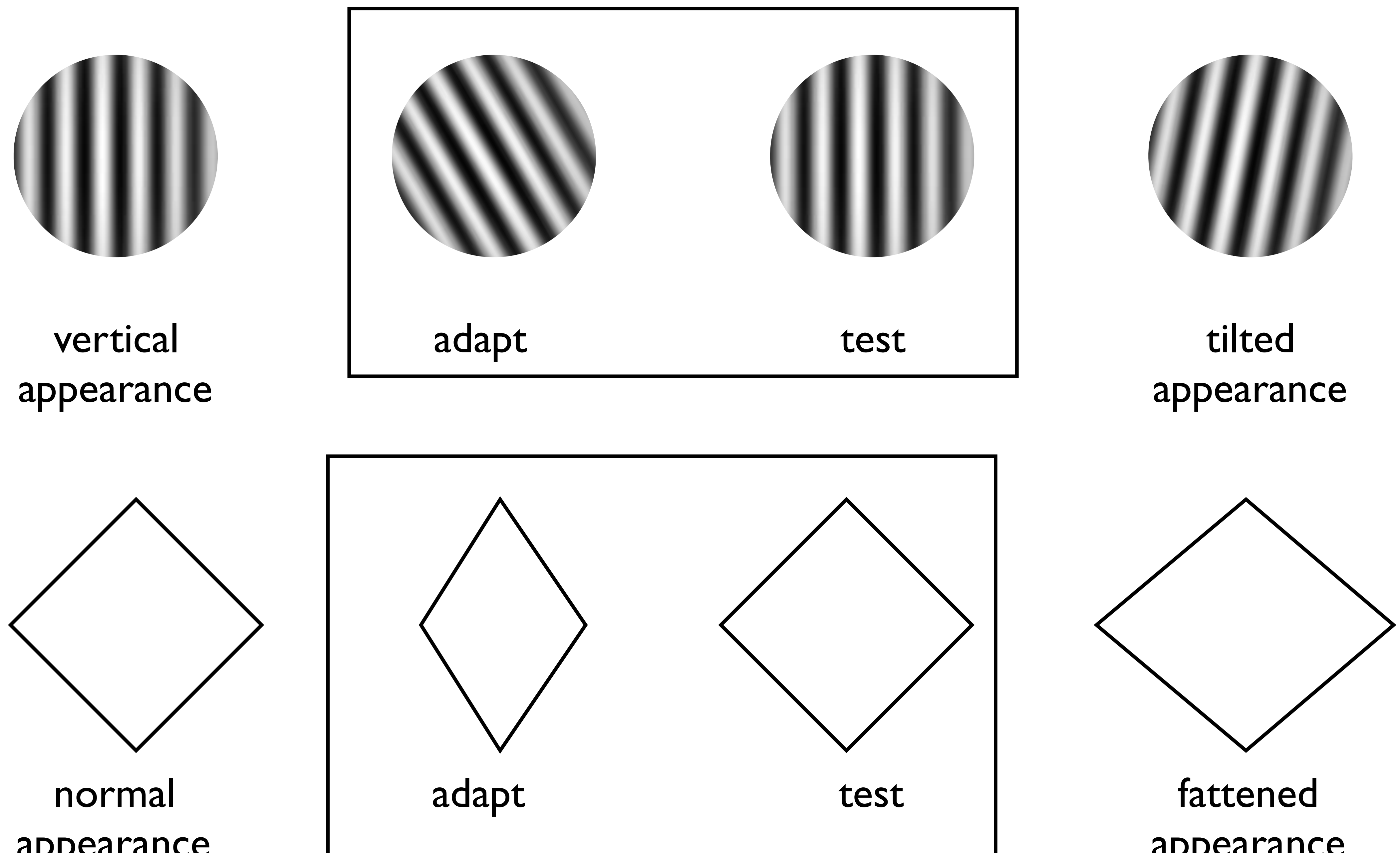
Take advantage of the high degree of orientation selectivity in early cortical areas, and selectivity to whole forms in higher cortical areas

Is the suppression localized to early feature detectors? A psychophysical test

use adaptation--psychophysicist's "electrode"

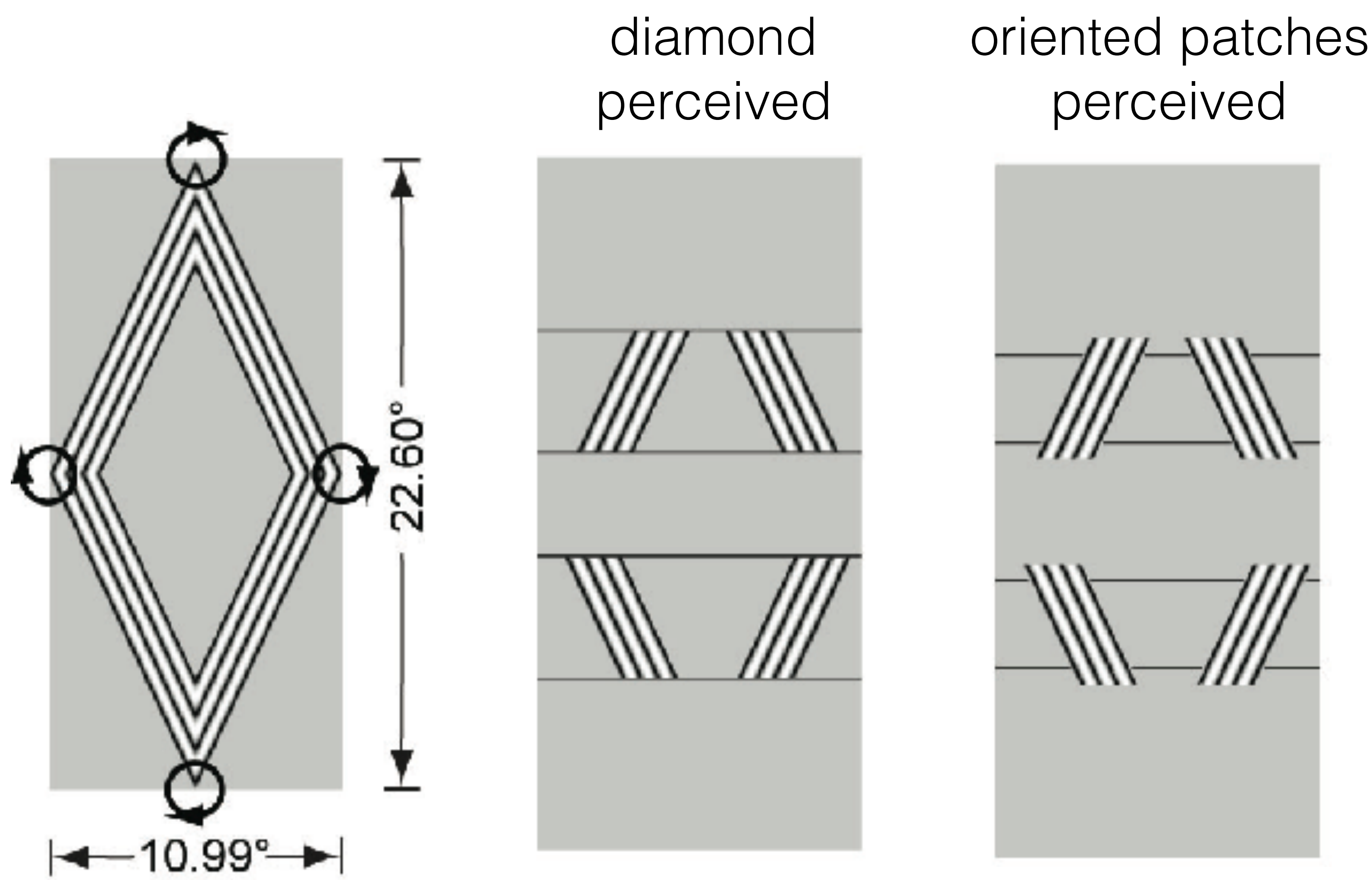
assumption: adapts neurons in early cortical areas, V1

assumption: adapts neurons in high-level cortical areas



He, D., Kersten, D., & Fang, F. (2012). Opposite modulation of high- and low-level visual aftereffects by perceptual grouping. Current Biology, 22(11), 1040–1045.

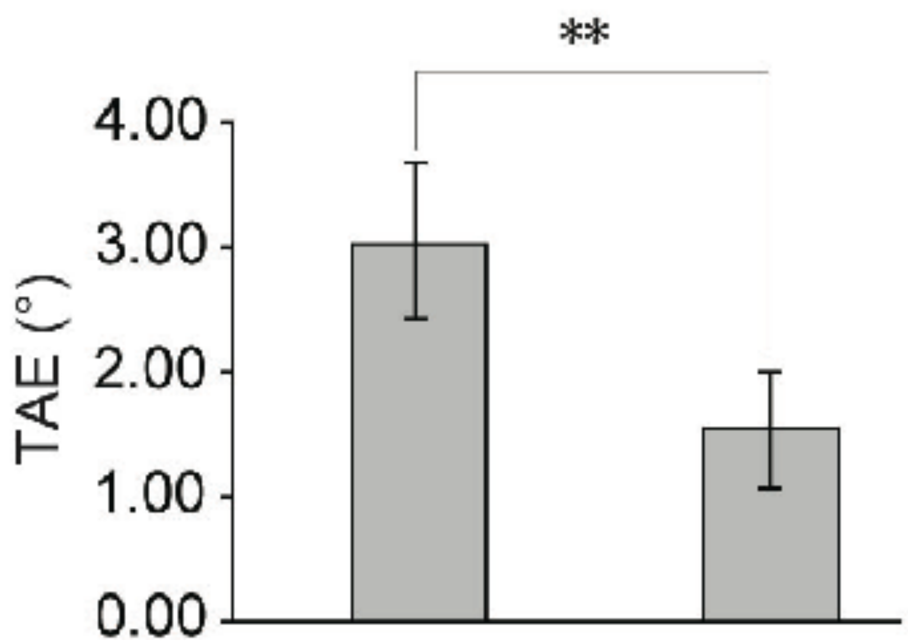
use occlusion cues to manipulate perception of diamond shape vs. four separate oriented grating patterns



He, D., Kersten, D., & Fang, F. (2012). Opposite modulation of high- and low-level visual aftereffects by perceptual grouping. Current Biology, 22(11), 1040–1045.

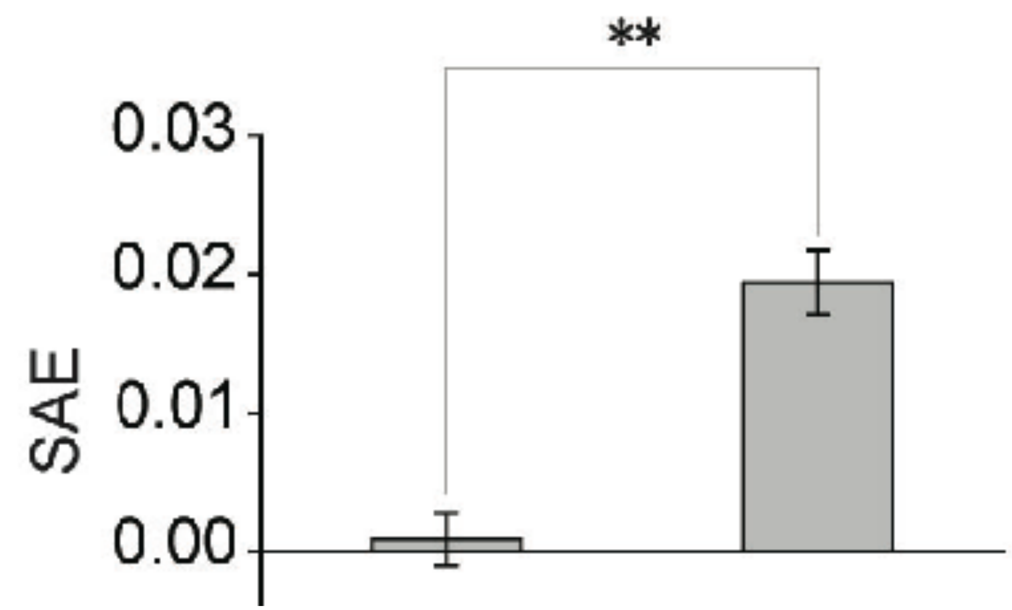
The results showed opposite modulation of high- and low-level visual aftereffects as a consequence of perceptual grouping

Tilt after-effect



Non-diamond Diamond

Shape after-effect



Non-diamond Diamond

Perceptual grouping ("diamond percept") reduces the strength of adaptation to local tilt, while amplifying the effect of adaptation to a whole shape, consistent with localized lower-level, feature-specific modulation, and with predictive coding—local, feature-specific suppression.

resolving ambiguity using high-level knowledge

Exploit the hierarchical organization of object knowledge, and use feedback to solve ambiguity through "explaining away"

"predictive coding" as top-down error detection

- suppress lower-level responses to features "explained" by a higher-level interpretation

and/or amplify those responses ("residuals") that are not explained

cf. Mumford, 1992; Rao & Ballard, 1999

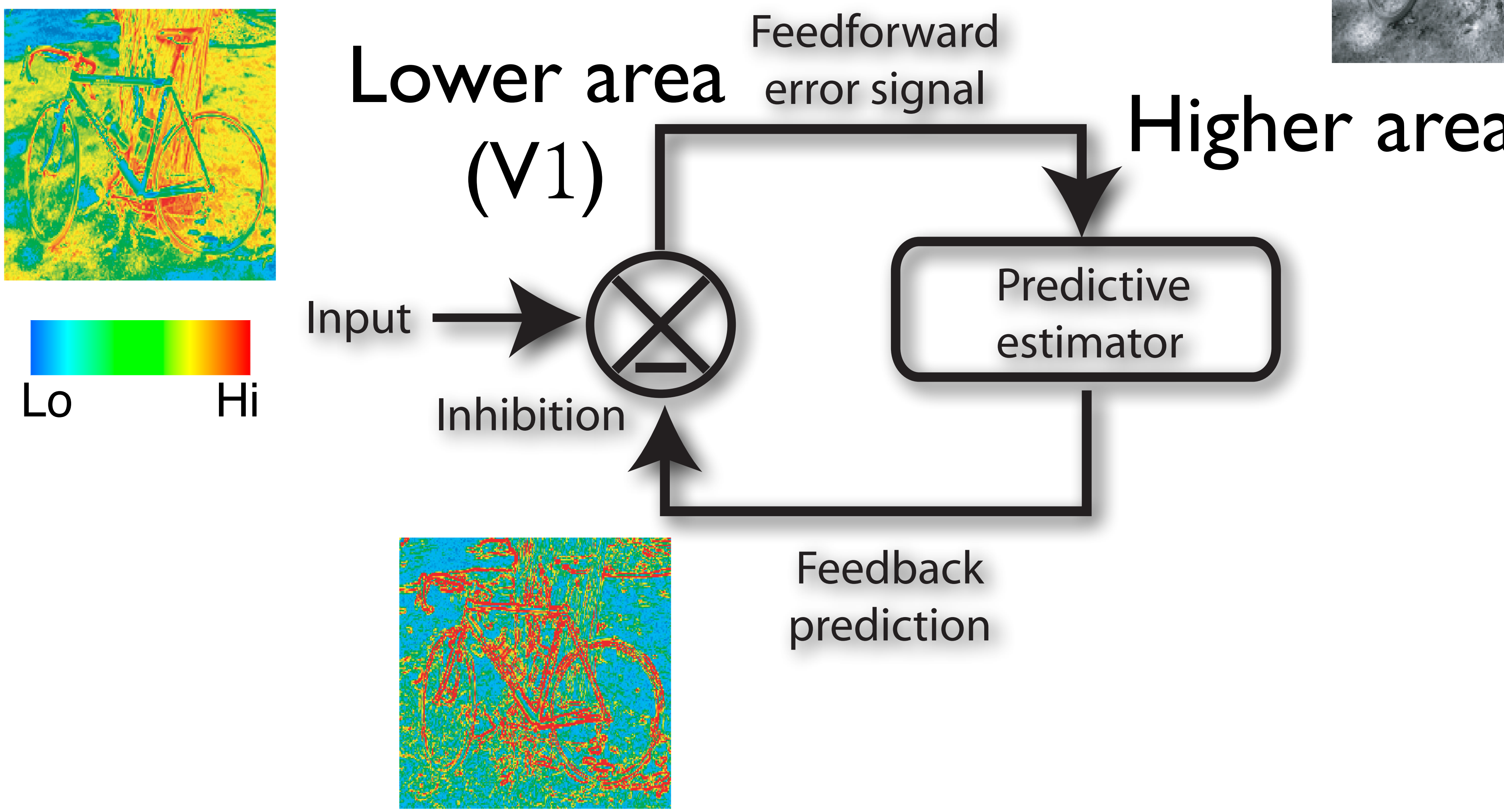
Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical Microcircuits for Predictive Coding. *Neuron*, *76*(4), 695–711.

…summary so far

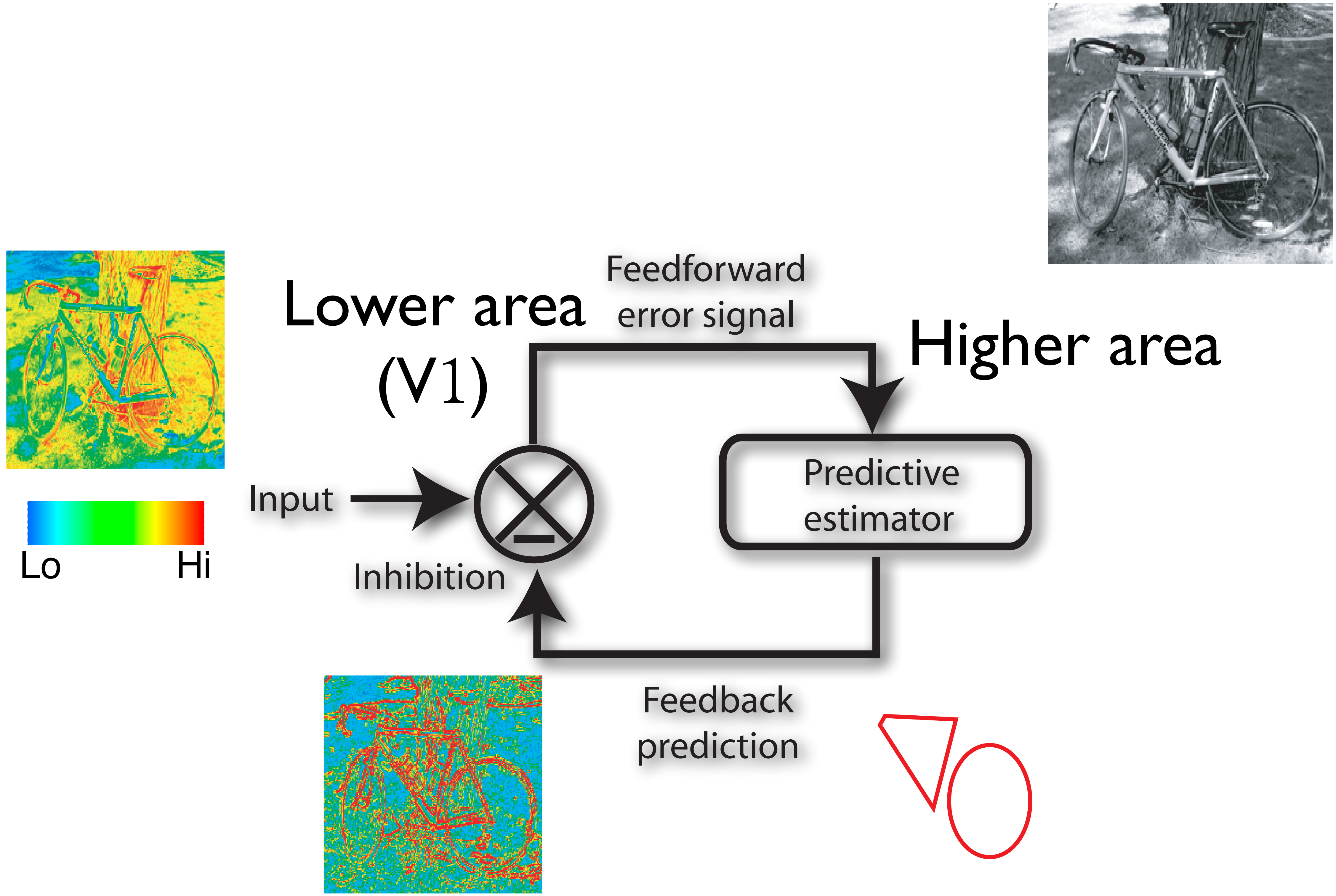
Evidence for suppression of local activity in V1 as a consequence of higher-level, global perceptual organization—i.e. suppression when all the local features have been "explained".

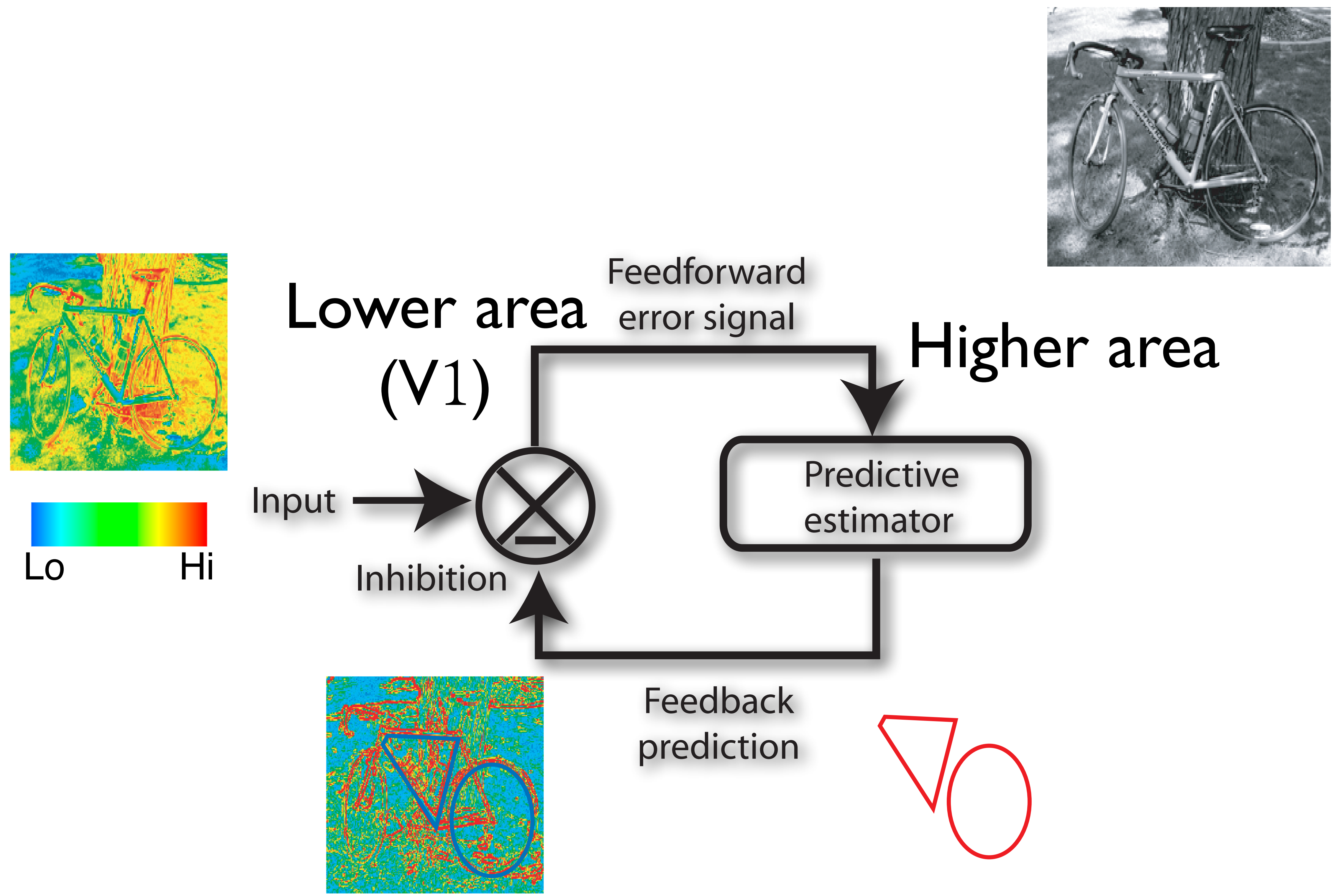
$$p(S|I) \propto p(I - f(S))p(S)$$

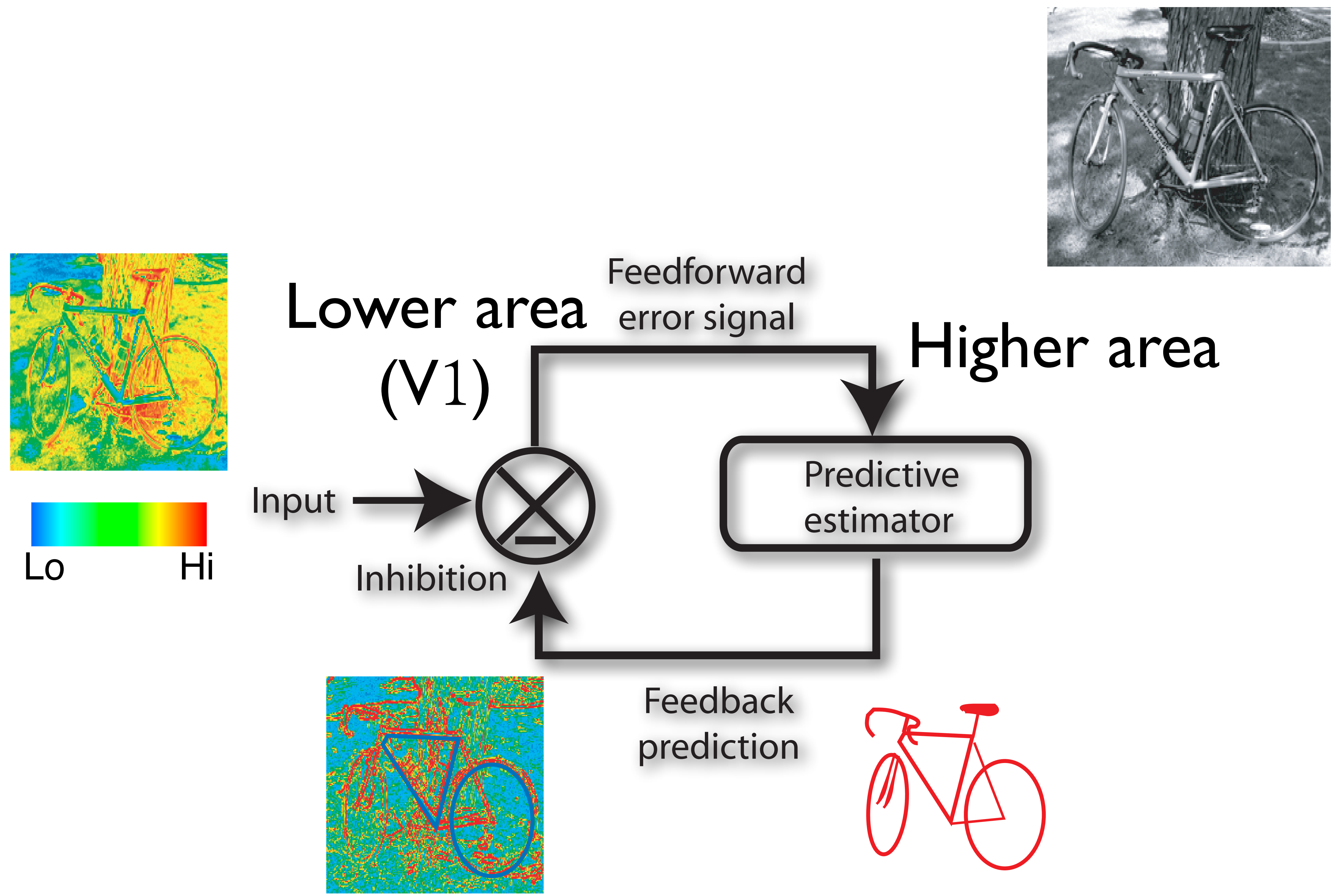
"predictive coding" through suppression of consistent features at lower levels

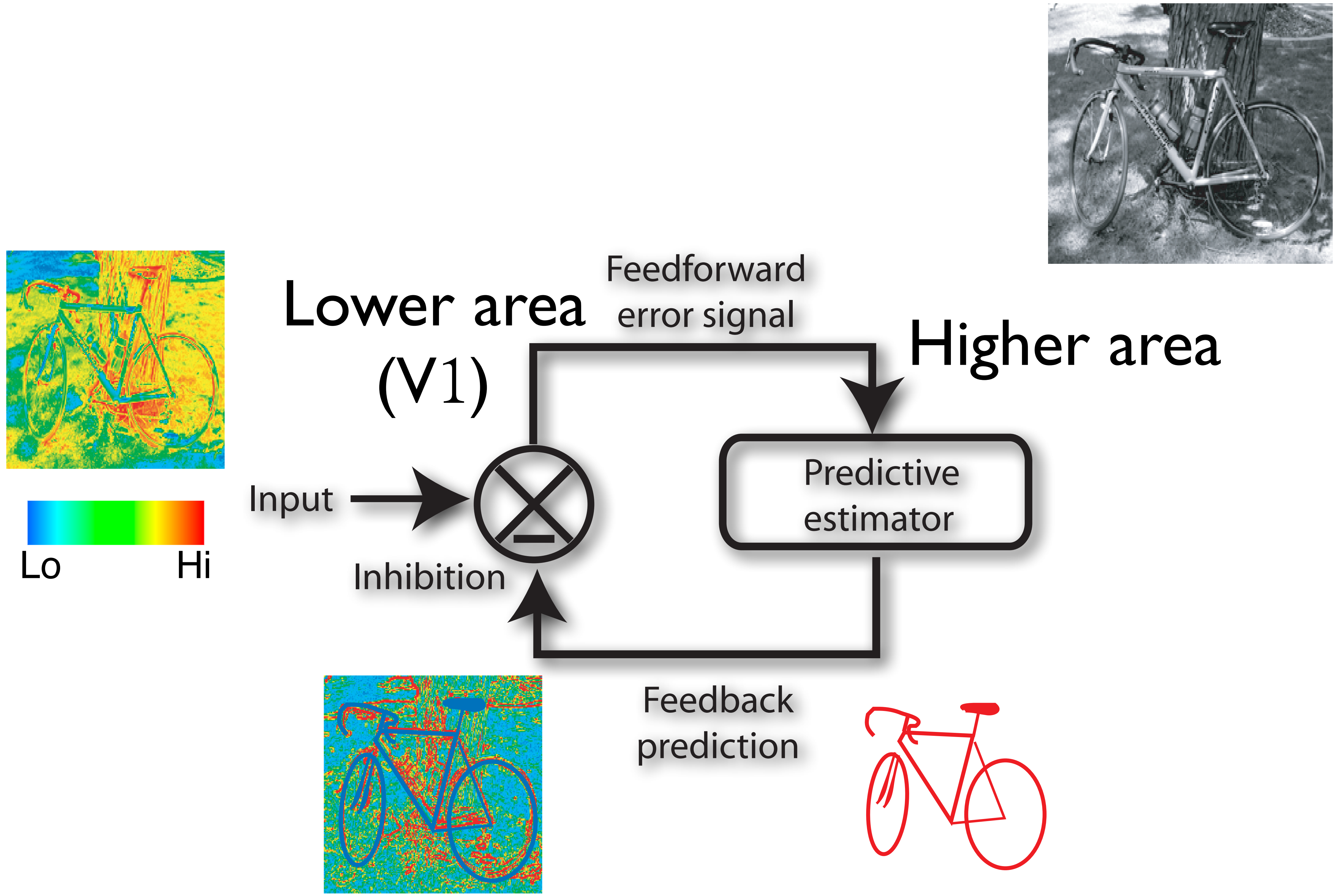


e.g. Rao, R. P., & Ballard, D. H. (1997). Dynamic model of visual recognition predicts neural response properties in the visual cortex. Neural Comput, 9(4), 721-763.

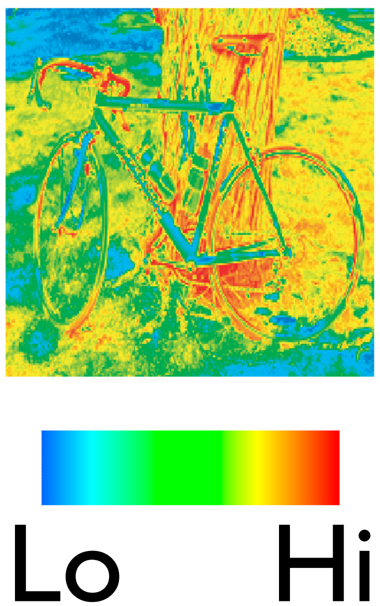

Lower area
(V1)
Feedforward
error signal
Higher area
Input
Inhibition
Predictive
estimator
Feedback
prediction
Lo
Hi

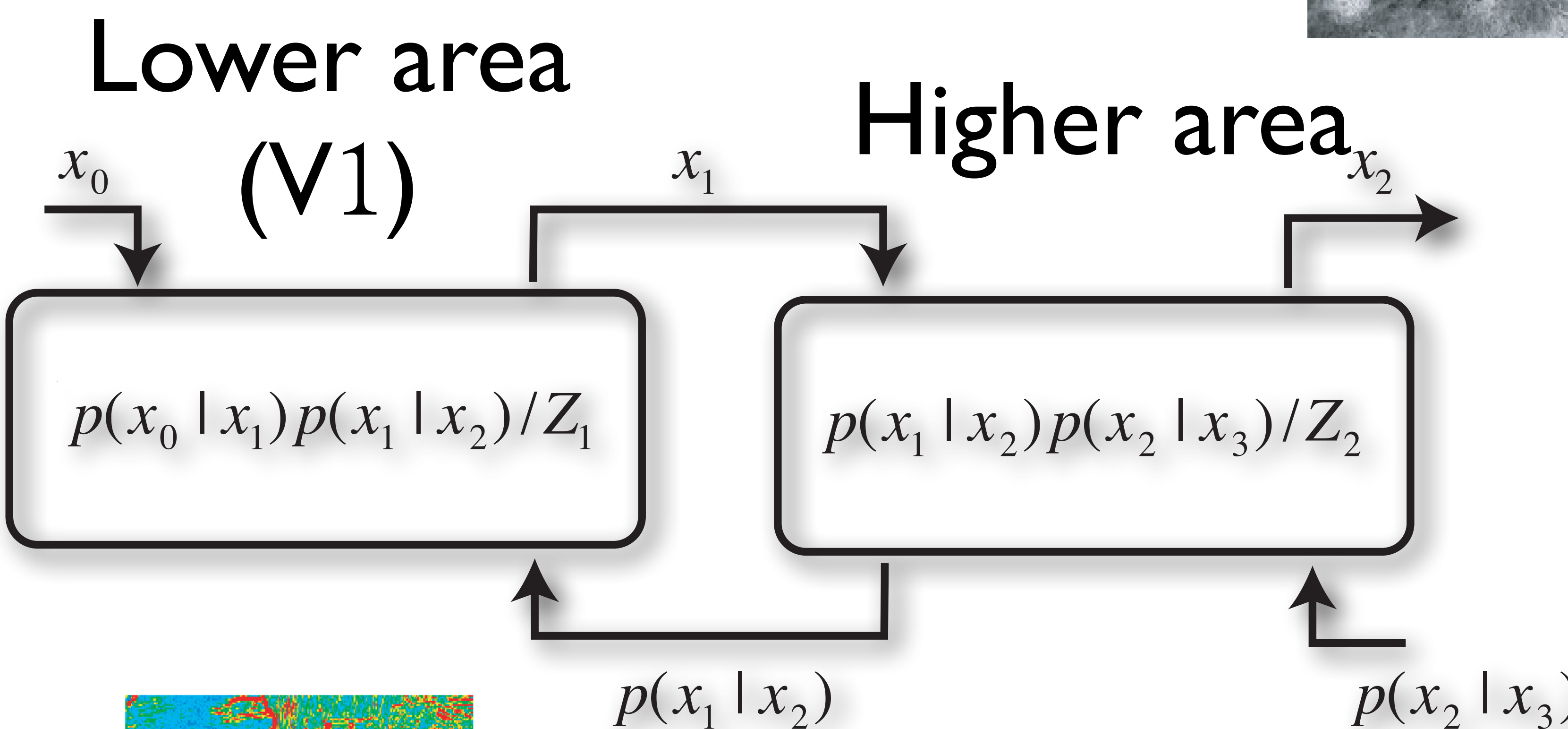

Lower area
(V1)
Feedforward
error signal
Higher area
Predictive
estimator
Input
Inhibition
Feedback
prediction
Lo
Hi

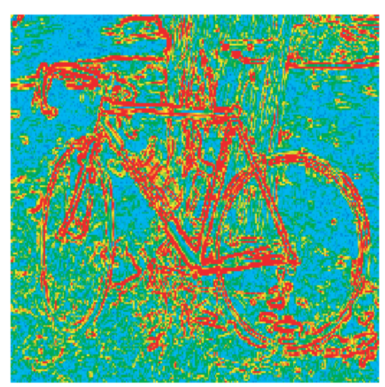

Lower area
(V1)
Feedforward
error signal
Higher area
Predictive
estimator
Input
Inhibition
Feedback
prediction
Lo
Hi

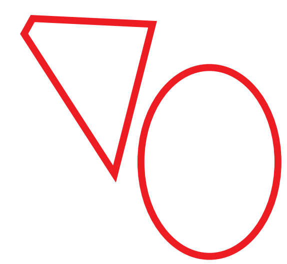

Lower area
(V1)
Feedforward
error signal
Higher area
Input
Predictive
estimator
Inhibition
Lo
Hi
Feedback
prediction

binding through enhancement of consistent features at lower levels

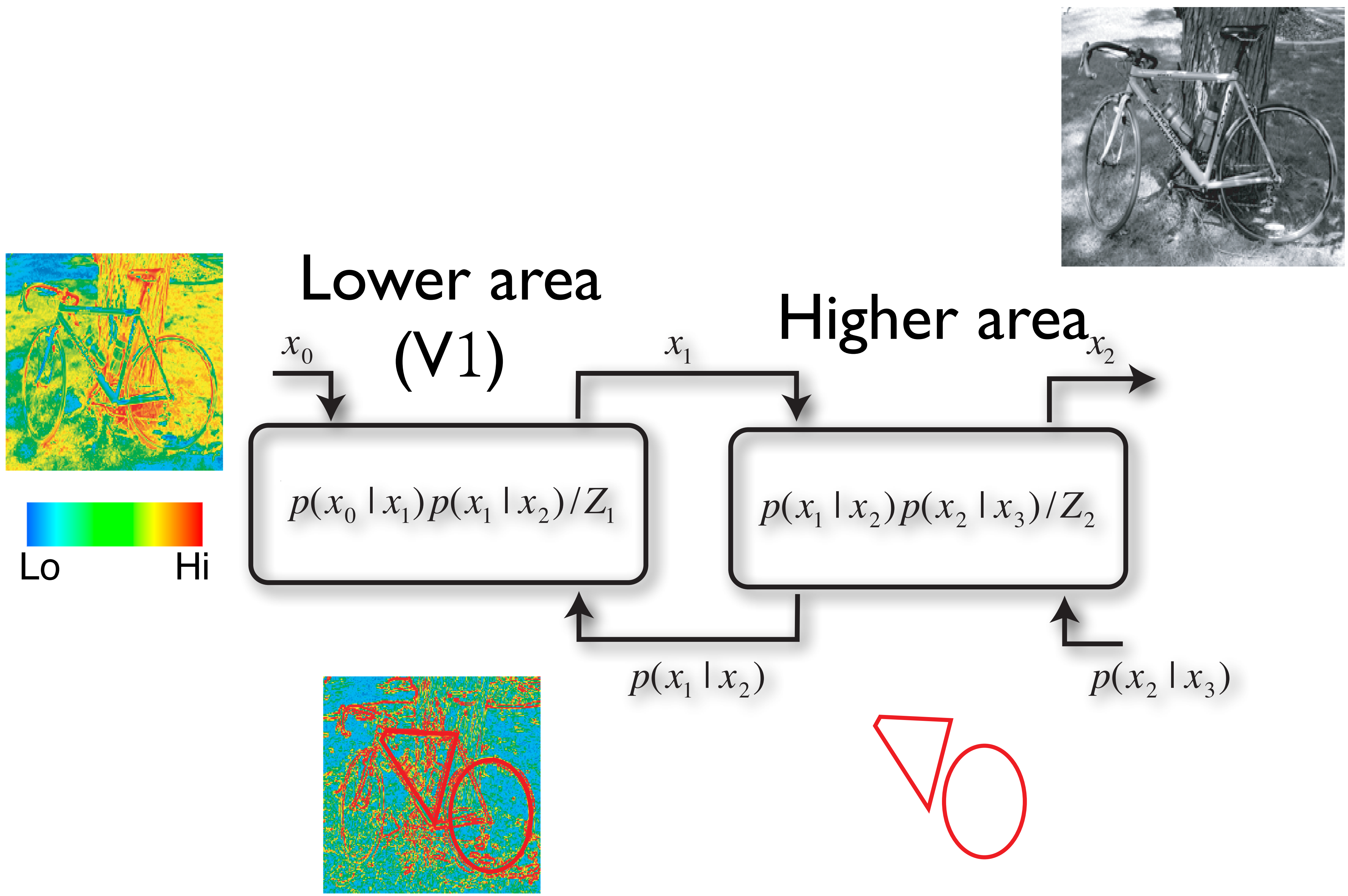


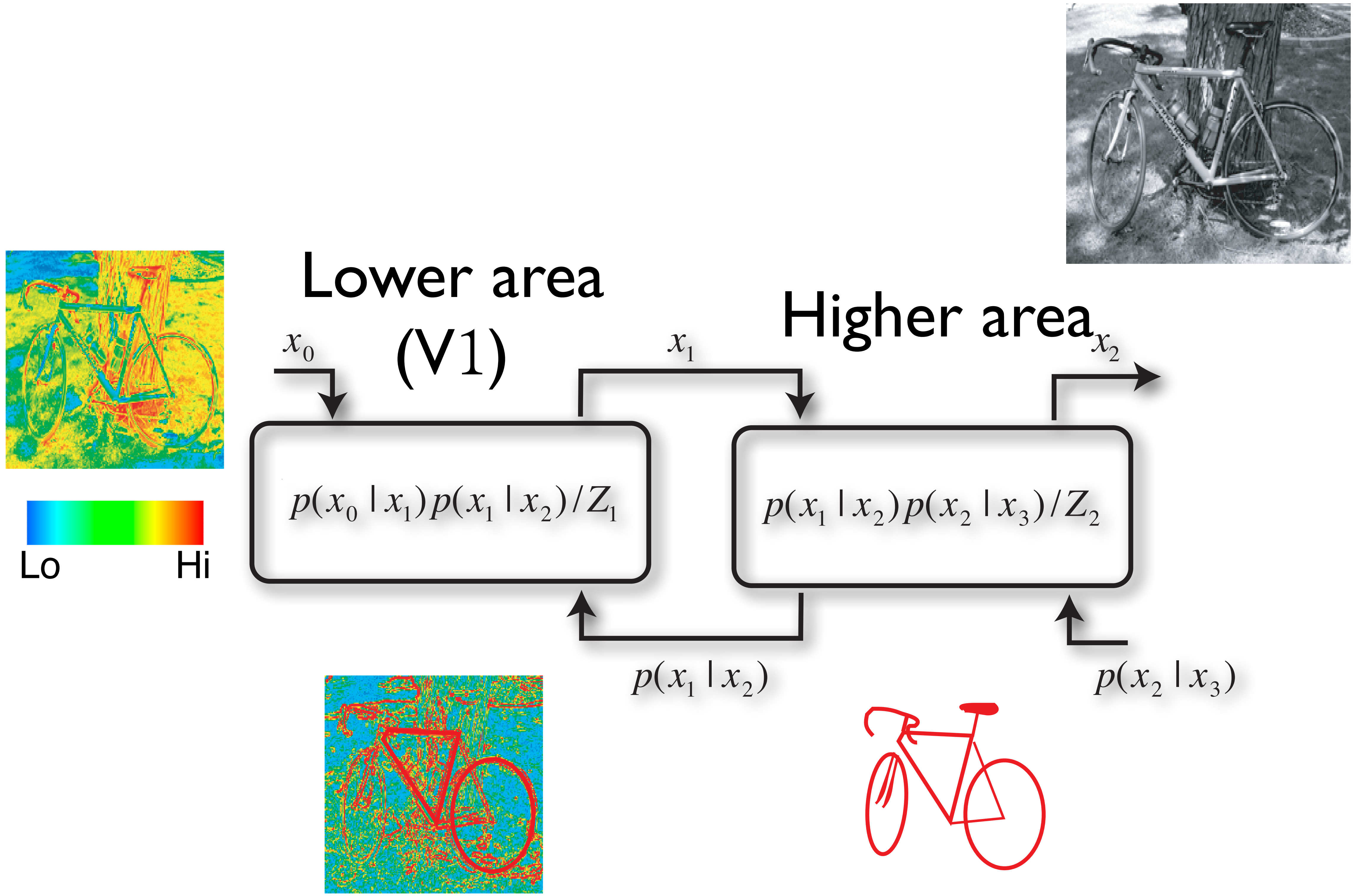




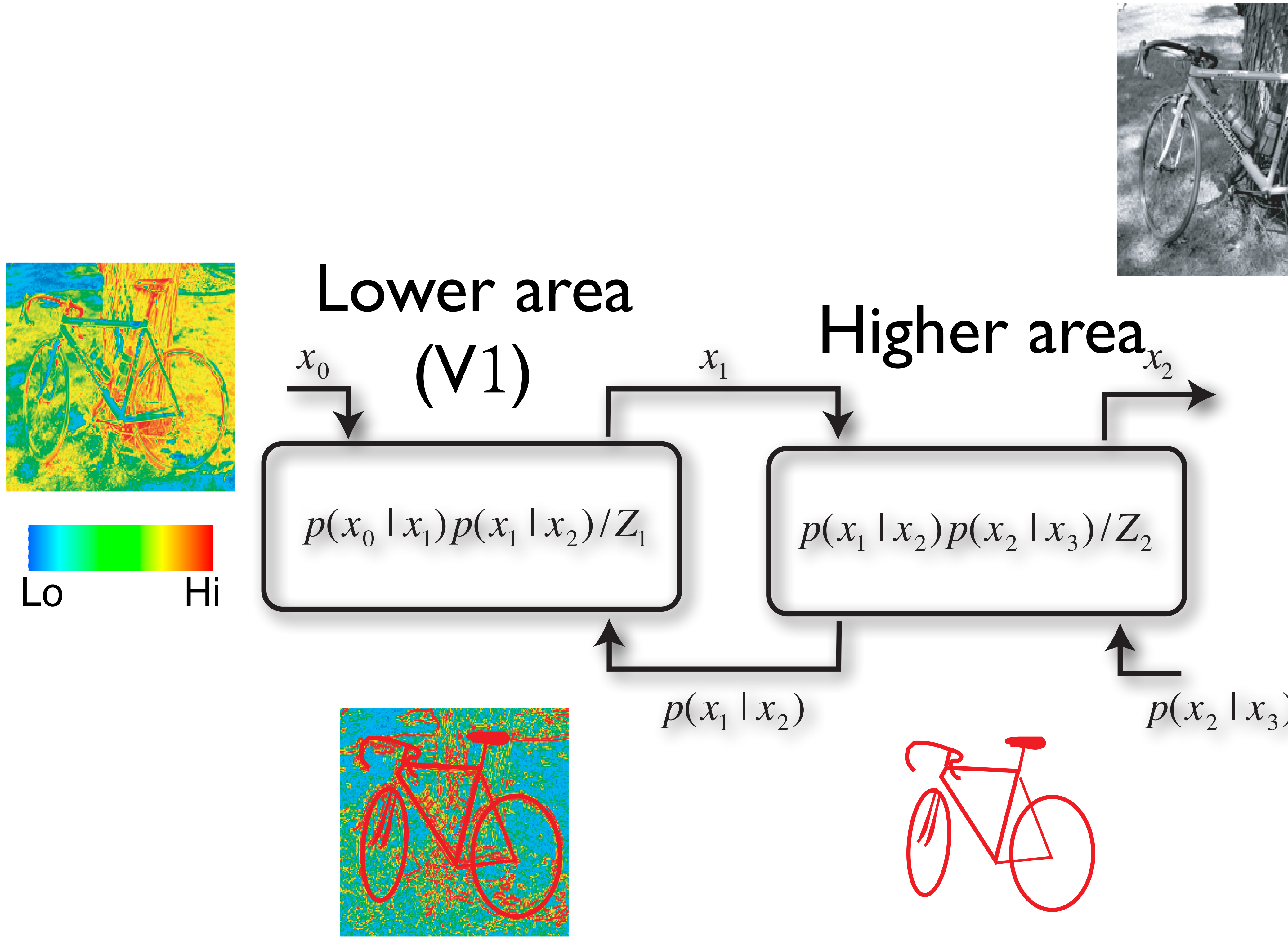


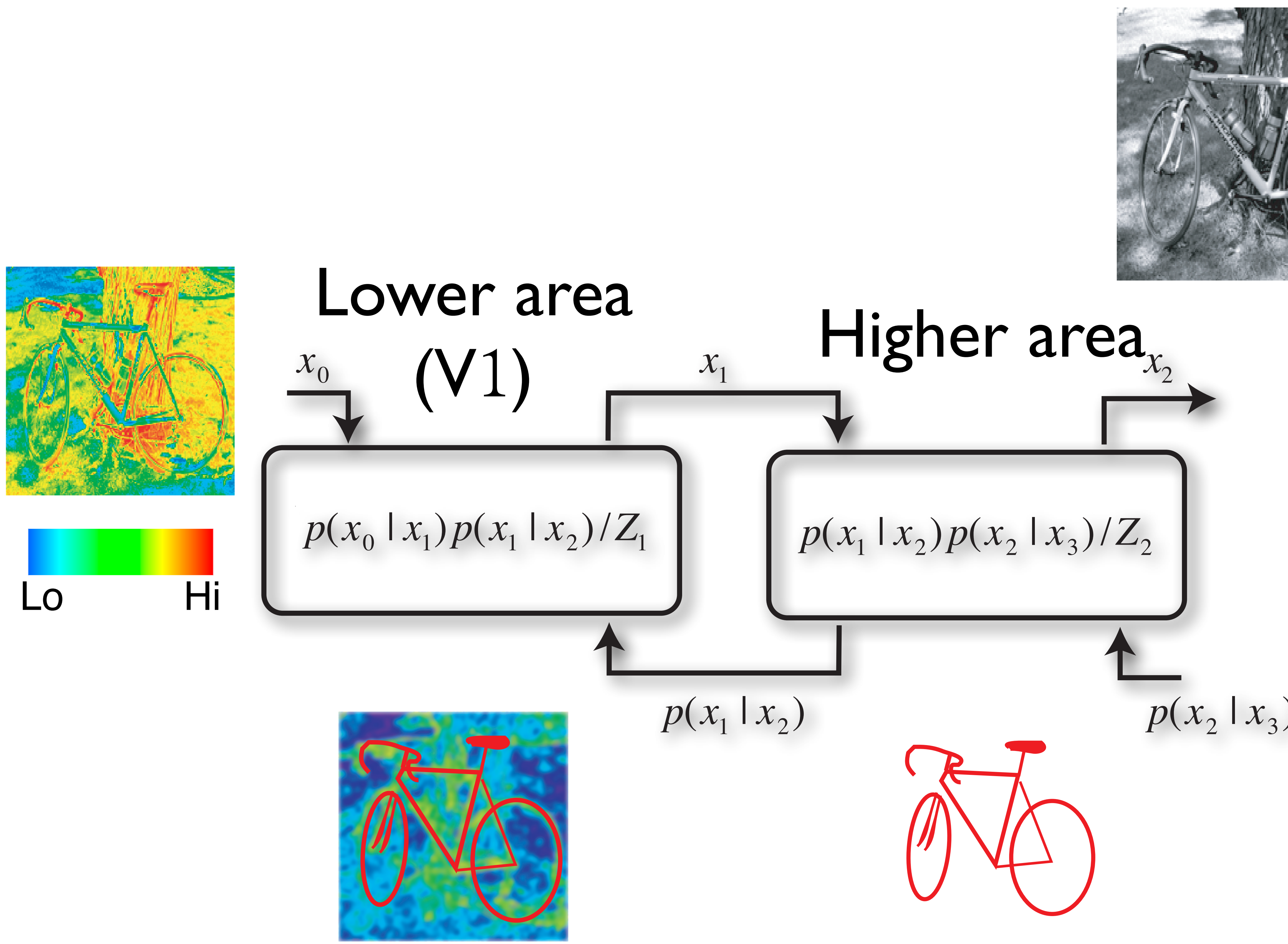
Lee & Mumford, 2003, JOSA


Lower area
(V1)
Higher area
x_0
x_1
x_2
$p(x_0 \mid x_1)p(x_1 \mid x_2)/Z_1$
$p(x_1 \mid x_2)p(x_2 \mid x_3)/Z_2$
$p(x_1 \mid x_2)$
$p(x_2 \mid x_3)$
Lo
Hi



Lower area
(V1)
Higher area
x_0
x_1
x_2
$p(x_0 \mid x_1)p(x_1 \mid x_2)/Z_1$
$p(x_1 \mid x_2)p(x_2 \mid x_3)/Z_2$
$p(x_1 \mid x_2)$
$p(x_2 \mid x_3)$
Lo
Hi


Lower area
(V1)
Higher area
x_0
x_1
x_2
$p(x_0 \mid x_1)p(x_1 \mid x_2)/Z_1$
$p(x_1 \mid x_2)p(x_2 \mid x_3)/Z_2$
$p(x_1 \mid x_2)$
$p(x_2 \mid x_3)$
Lo
Hi



Lower area (V1)
Higher area
x_0
x_1
x_2
$p(x_0 | x_1) p(x_1 | x_2) / Z_1$
$p(x_1 | x_2) p(x_2 | x_3) / Z_2$
$p(x_1 | x_2)$
$p(x_2 | x_3)$
Lo
Hi

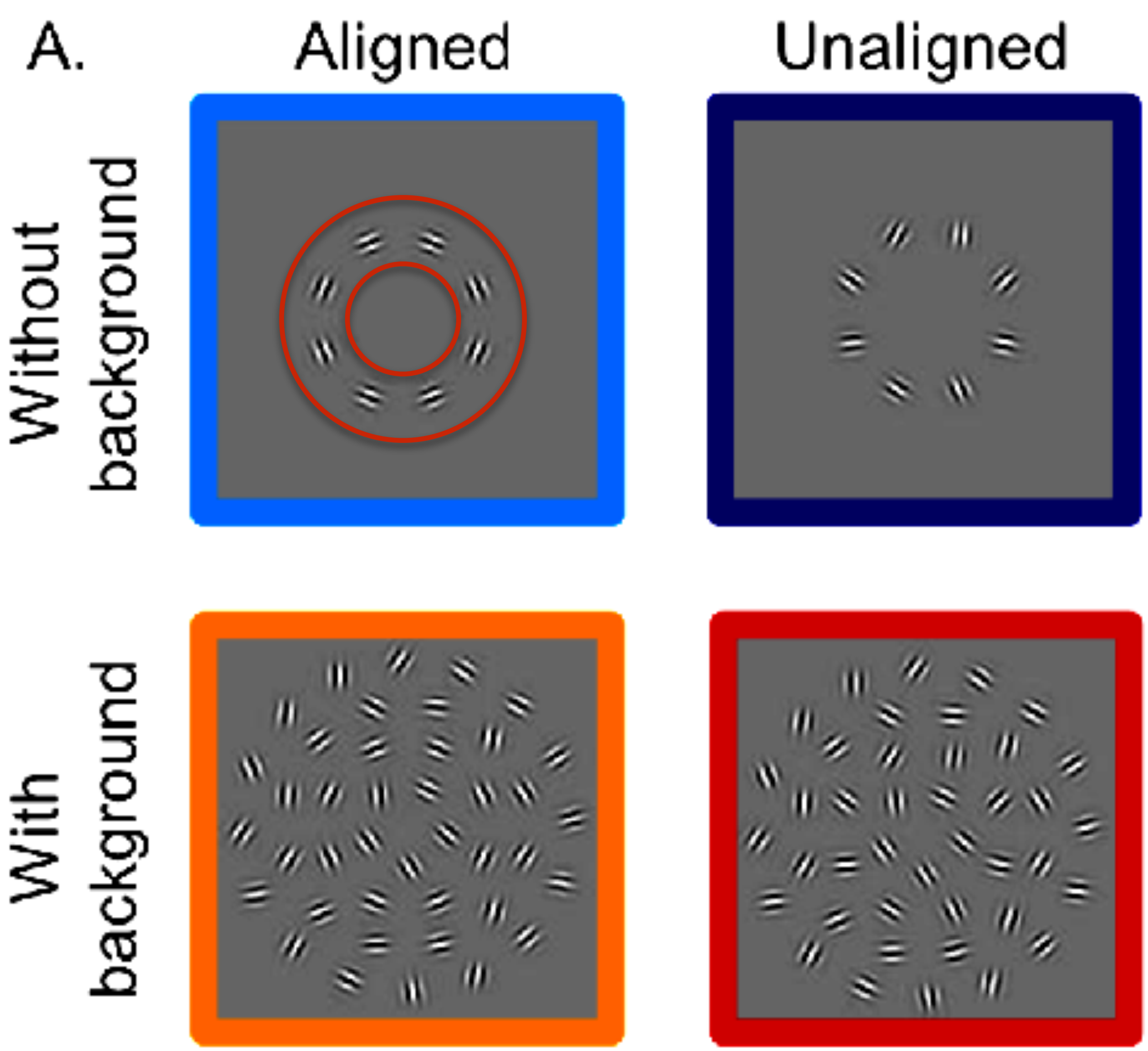
binding information across levels of abstraction

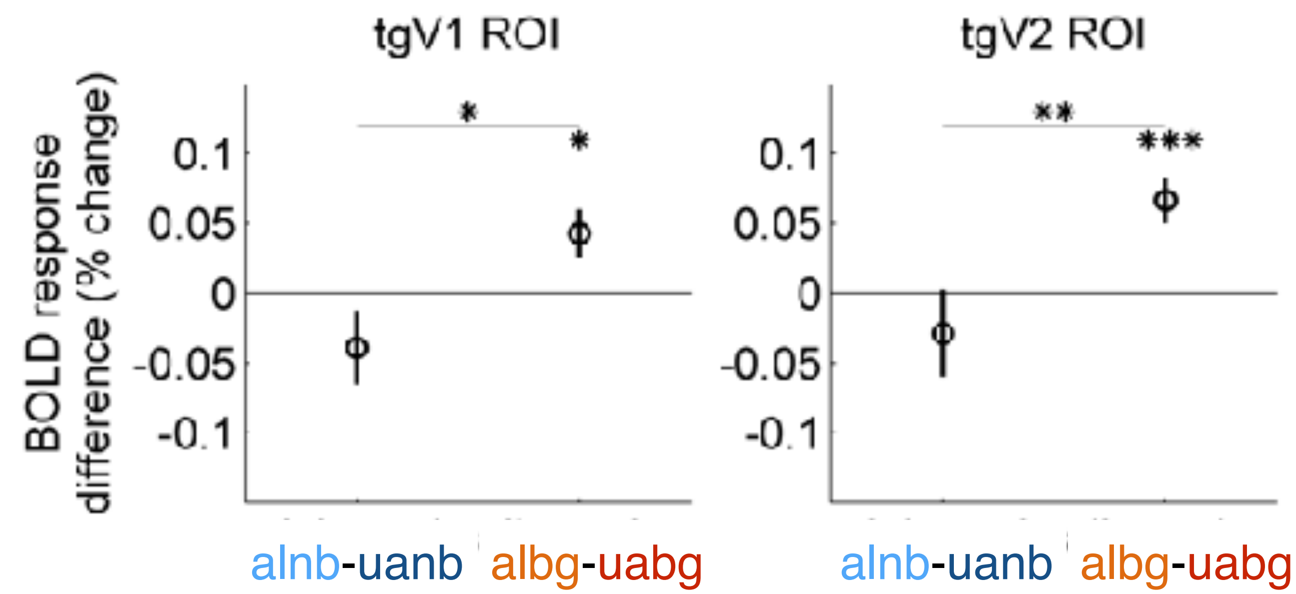
- amplify lower-level responses consistent with high-level a explanation
  - perhaps important given clutter
    - cf. Li, W., Piëch, V., & Gilbert, C. D. (2008). Learning to Link Visual Contours. Neuron, 57(3), 442–451.
    - Qiu, C., Burton, P. C., Kersten, D., & Olman, C. A. (2016). Responses in early visual areas to contour integration are context dependent. Journal of Vision, 16(8), 19–18.
  - and/or subsequent tasks that involve decisions across spatial scale within an object

cf. Grossberg & Mingolla, 1985; Ullman, S, 1995

localized enhancement of V1 & V2 voxel activity depends on the complexity of the perceptual selection/integration problem

~2mm fMRI in V1/V2





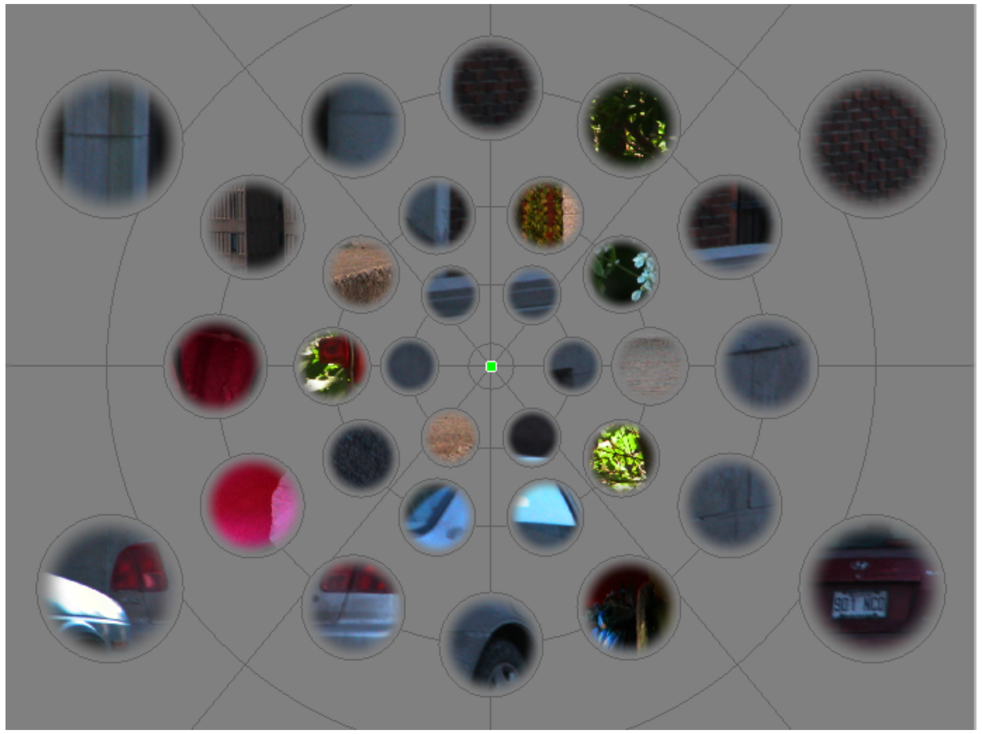
In background clutter, V1/V2 activity in target region increased for aligned vs. unaligned features

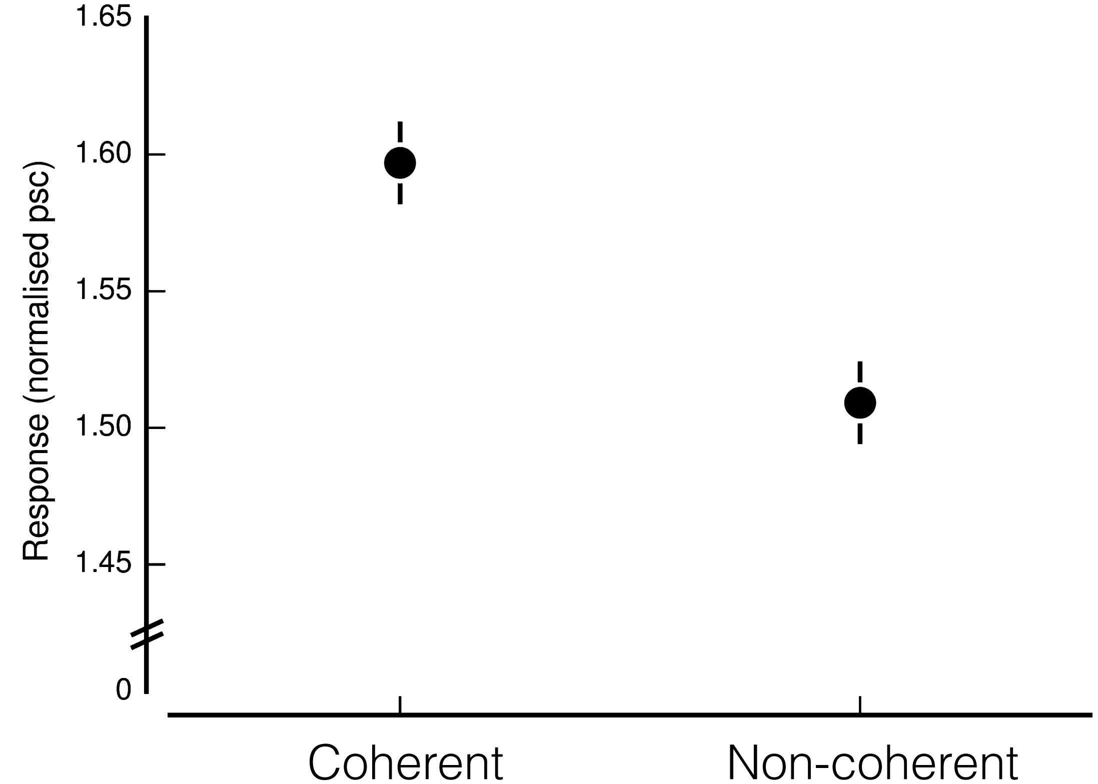
AND functional connectivity between V1-V2 also increased when perceiving aligned versus unaligned contours in background

Cheng Qiu, Philip Burton, Daniel Kersten, Cheryl Olman (2016) Responses in early visual areas to contour integration are context dependent. Journal of Vision

Li, W., Piech, V., & Gilbert, C. D. (2006). Contour saliency in primary visual cortex. Neuron, 50, 951– 962, doi:10.1016/j.neuron.2006.04.035.

Gilad, A., Meirovithz, E., & Slovin, H. (2013). Population responses to contour integration: Early encoding of discrete elements and late perceptual grouping. Neuron, 78, 389–402, doi:10.1016/j. neuron.2013.02.013.



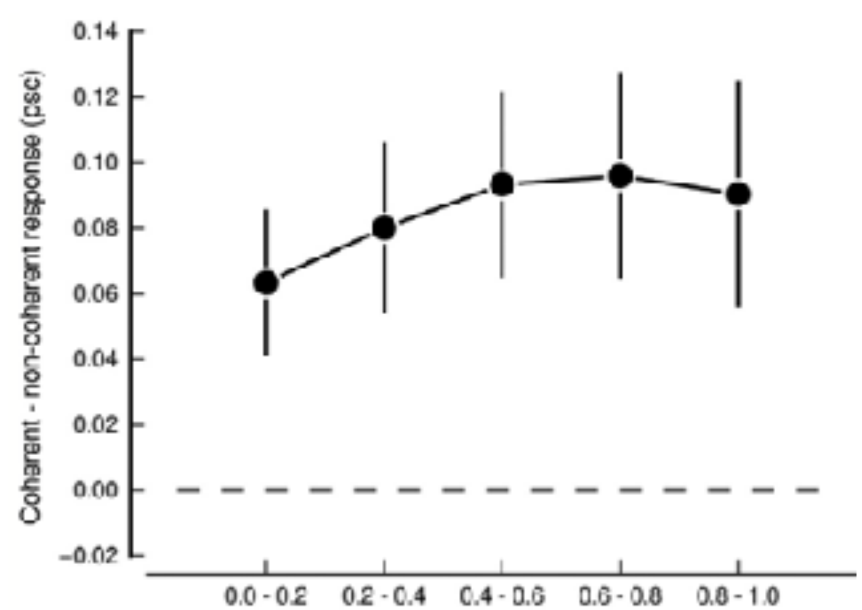


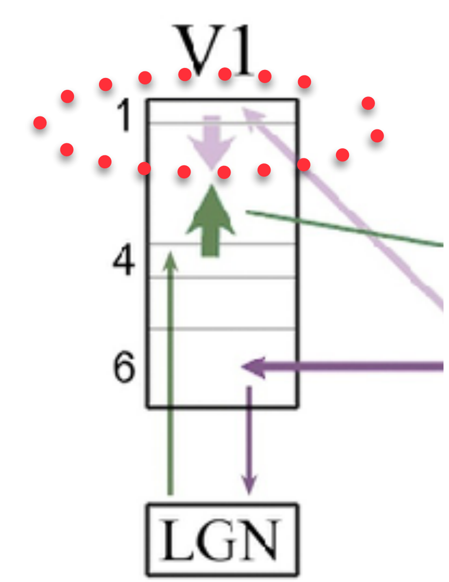
Mannion, D. J., Kersten, D. J., & Olman, C. A. (2015).

Larger fMRI responses to peripheral patches belonging to the perceived "coherent" image

Preference for coherent patches found in more superficial layers of V1

due to feedback and/or lateral connections?

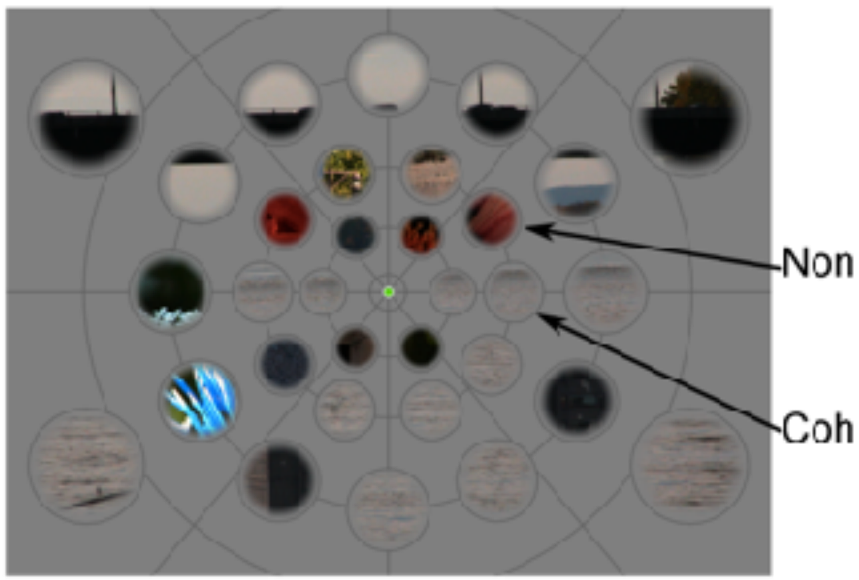


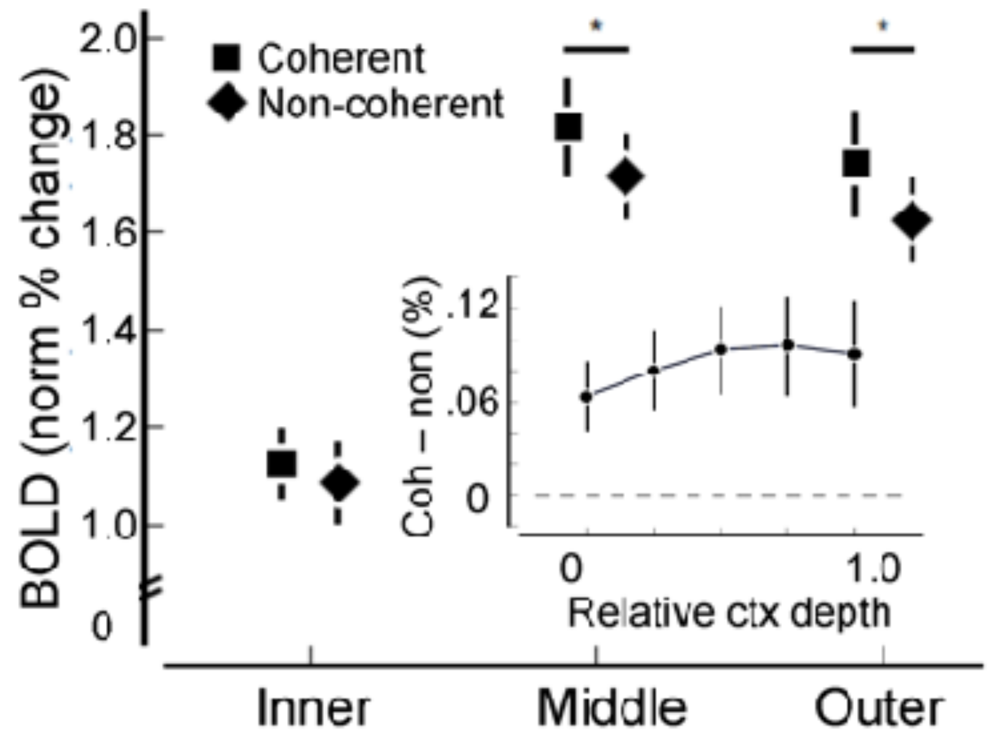


Consistent with: Muckli, L., De Martino, F., Vizioli, L., Petro, L. S., Smith, F. W., Ugurbil, K., Goebel, R. and Yacoub E. (2015). Contextual Feedback to Superficial Layers of V1.

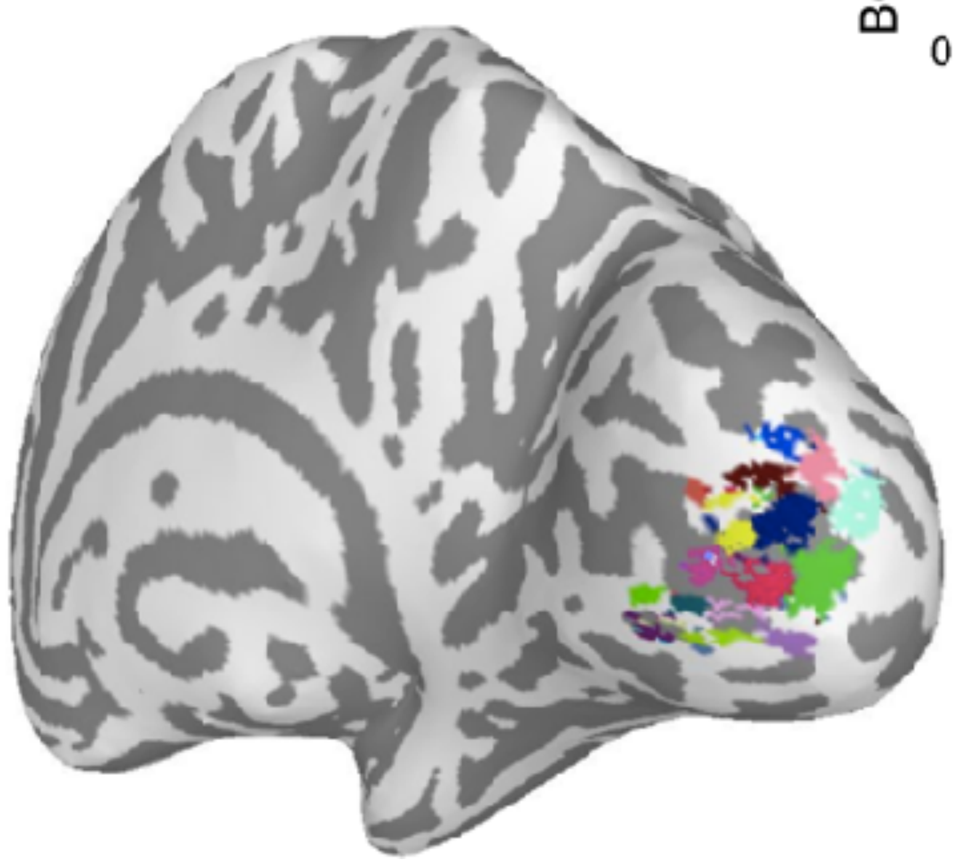
…but we haven't always found localized suppression when local patches "fit" the larger context

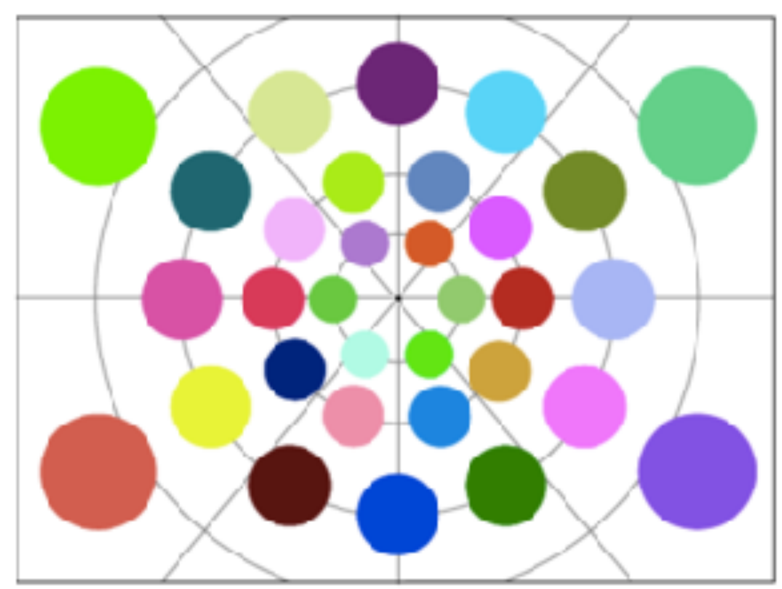
some patches are consistent with scene (Coh) and some not (Non)





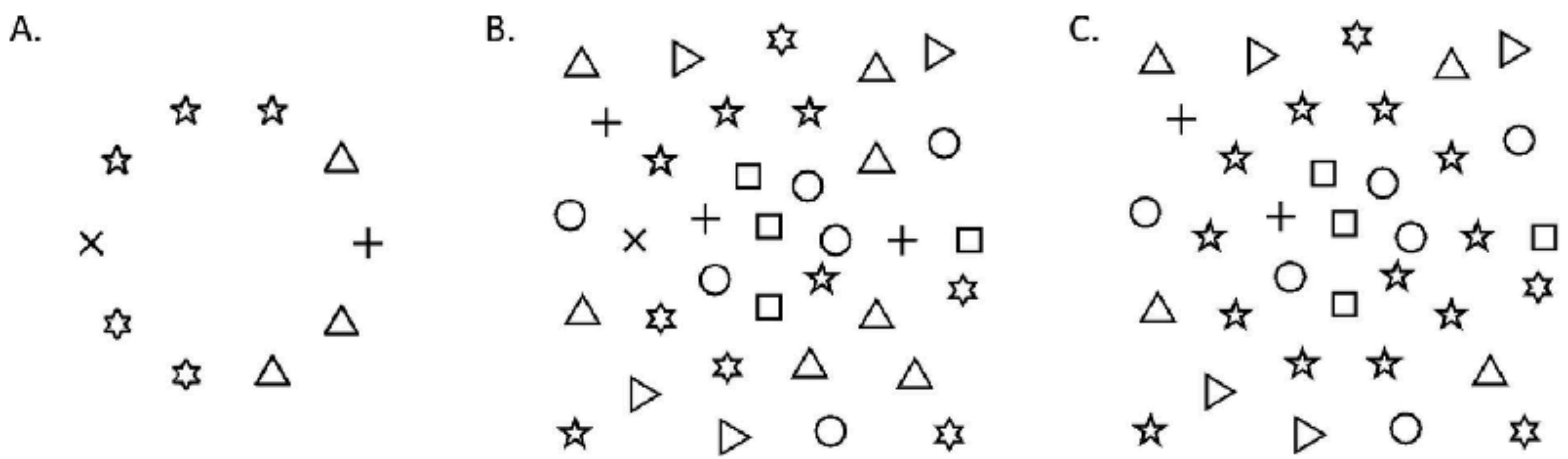
~1mm fMRI in V1





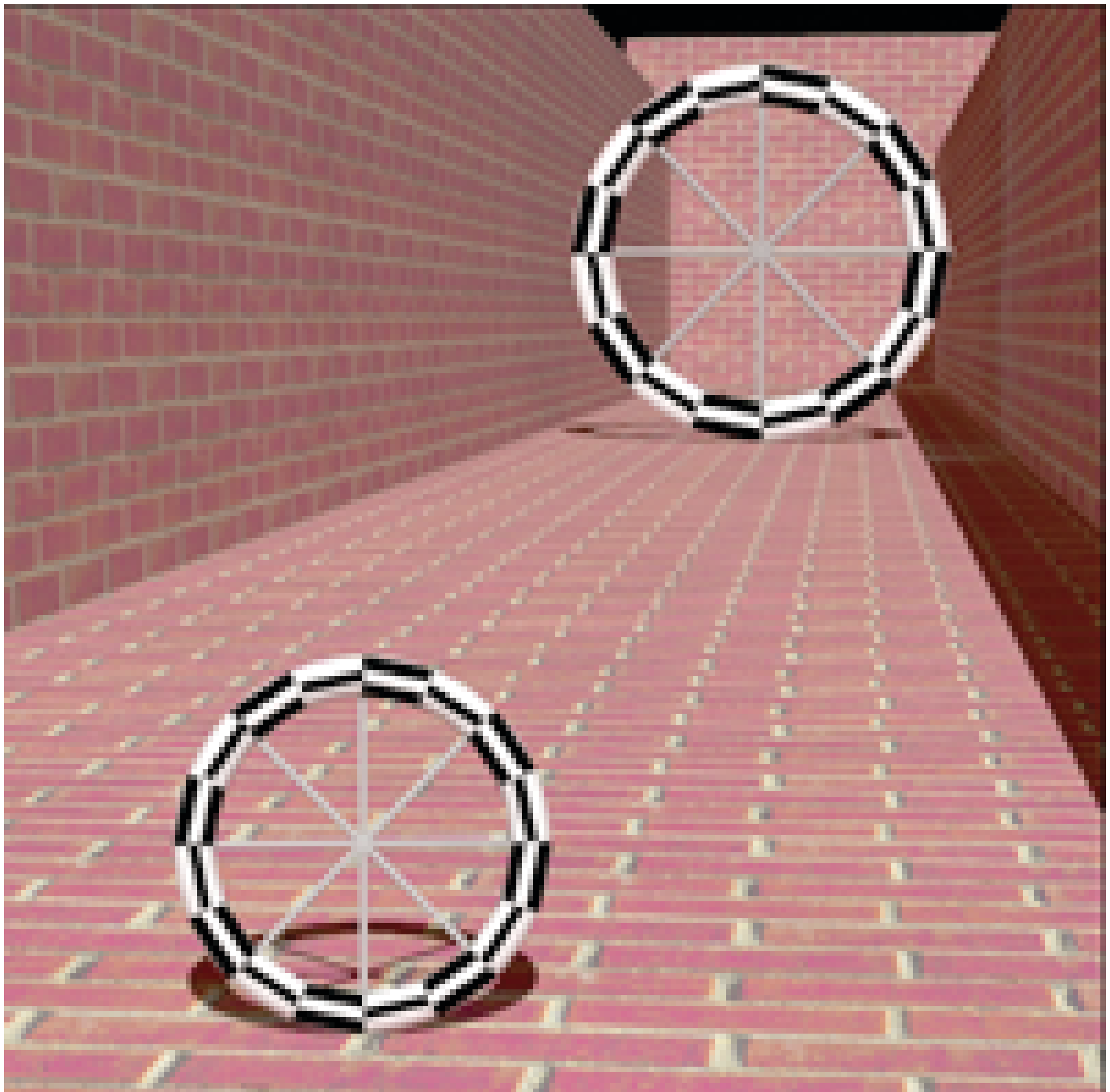
Mannion, Kersten & Olman

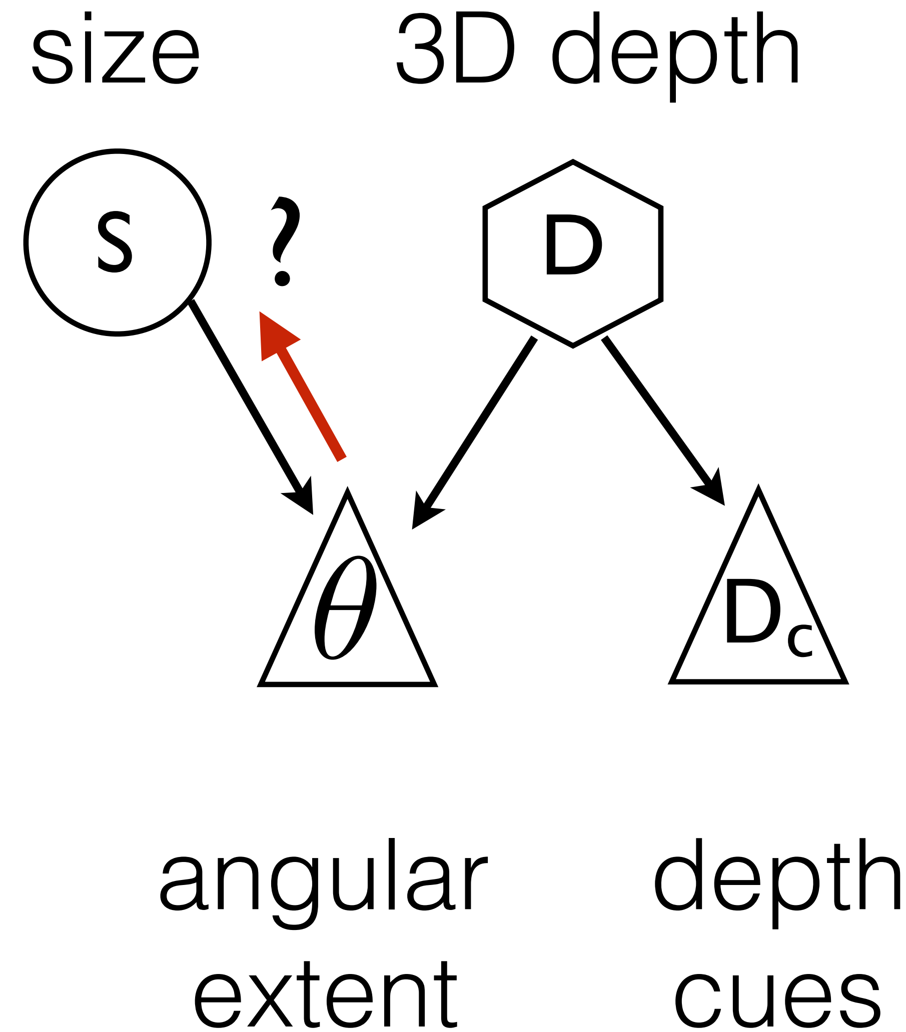
suppression vs. enhancement mechanisms: a flexible feedback/lateral strategy?



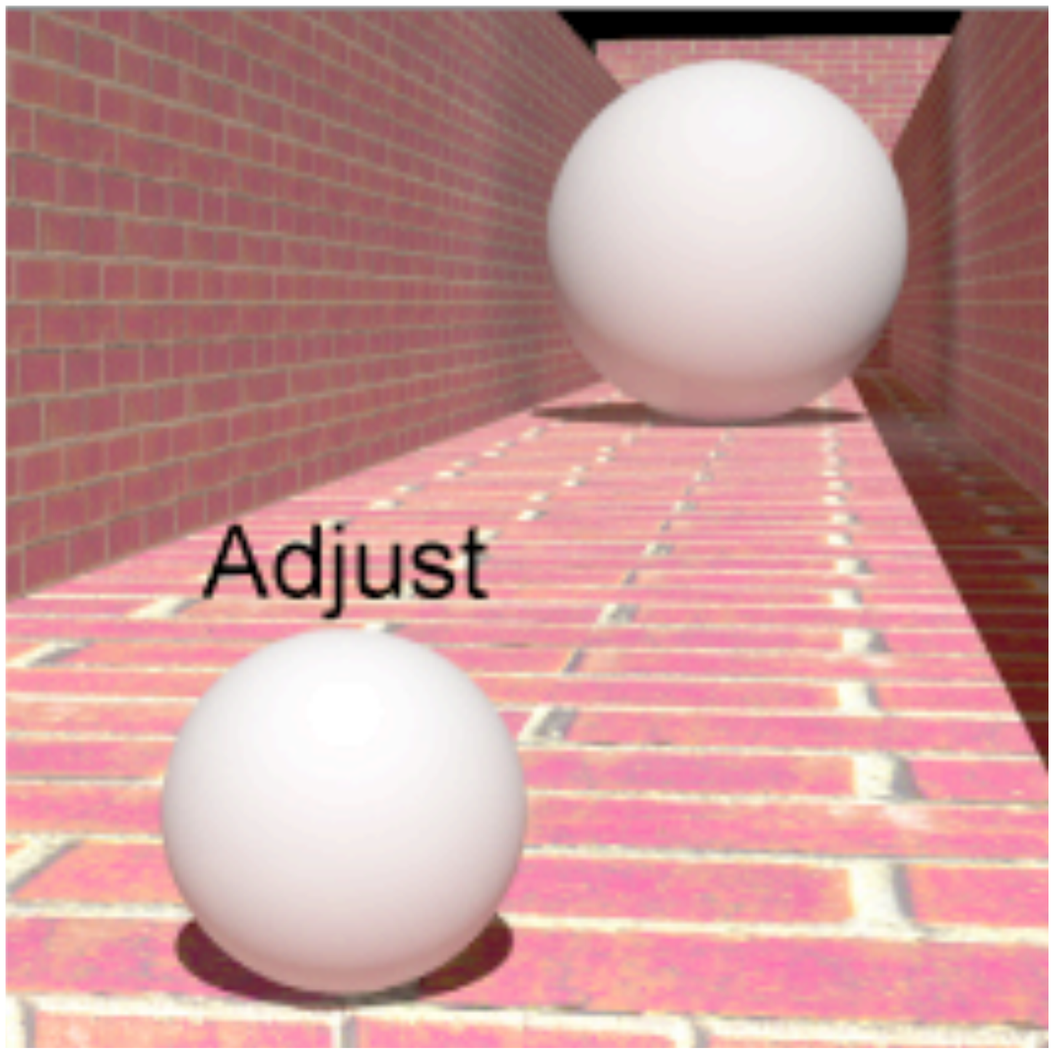
Qiu, C., Burton, P. C., Kersten, D., & Olman, C. A. (2016). Responses in early visual areas to contour integration are context dependent. *Journal of Vision*, *16*(8), 19–18.

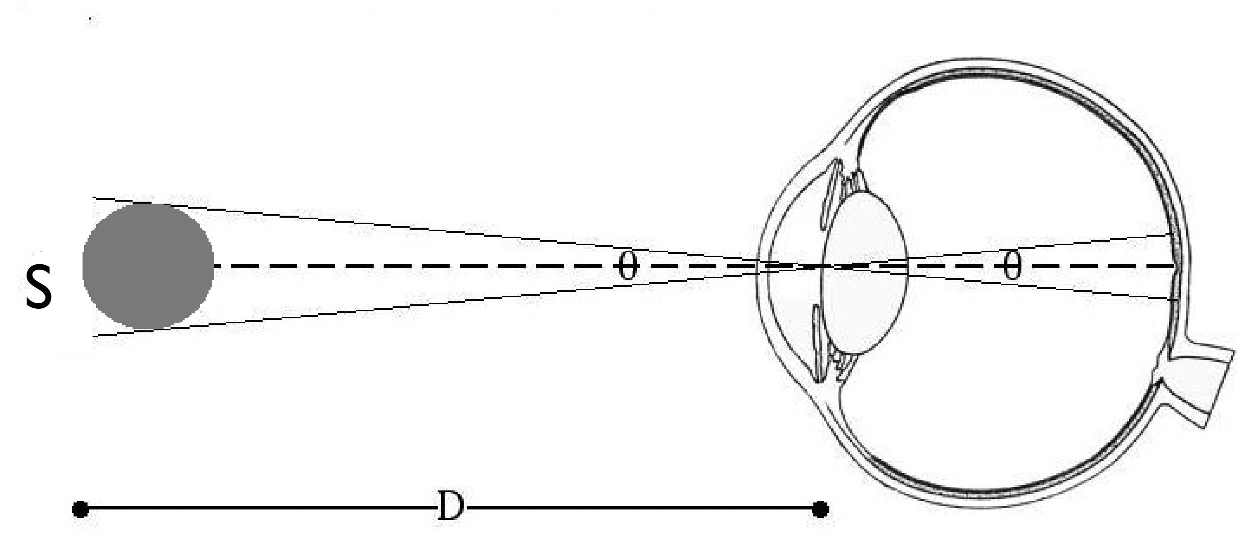
inferring the size of an object





perceptual estimation of the size of an object





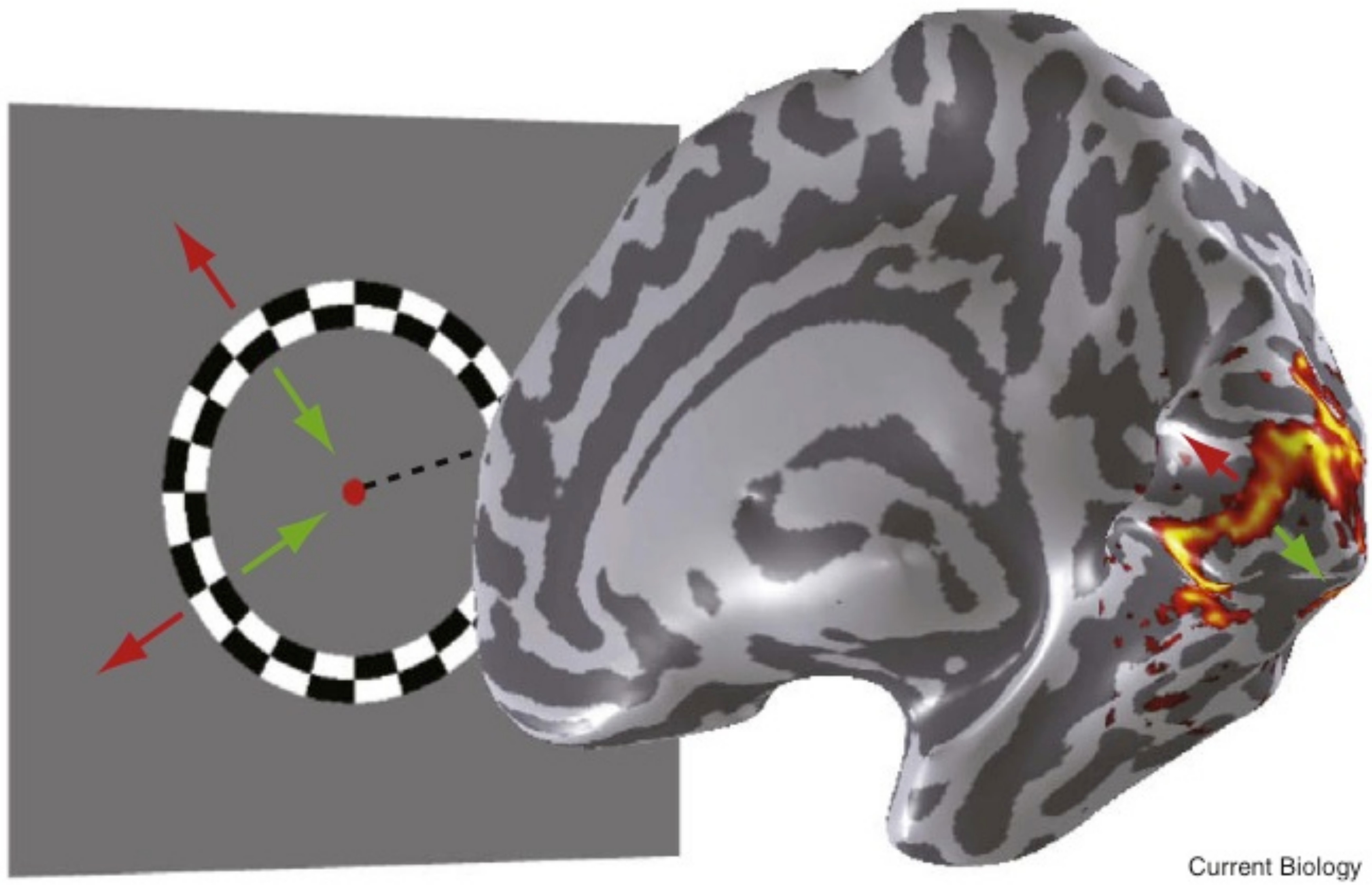
$$\theta \approx S/D$$

Perceptual effect: ~17%

http://vision.psych.umn.edu/users/boyaci/Vision/SizeAppletLarge.html

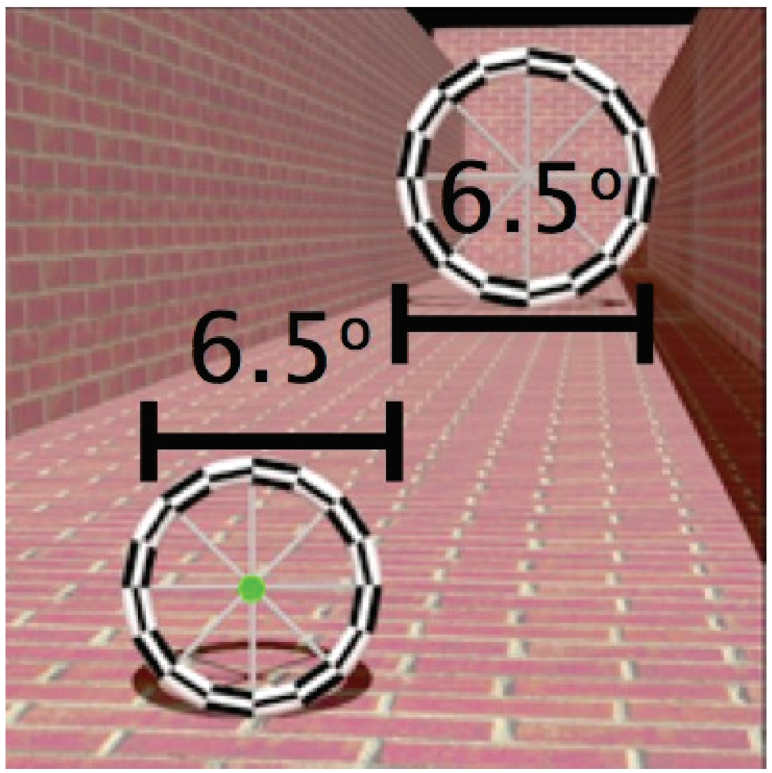
does 3D context modulate the size of the "neural image" in human V1?

V1 has a retinotopic map, so for an actual increase in ring size in the image, we expect:

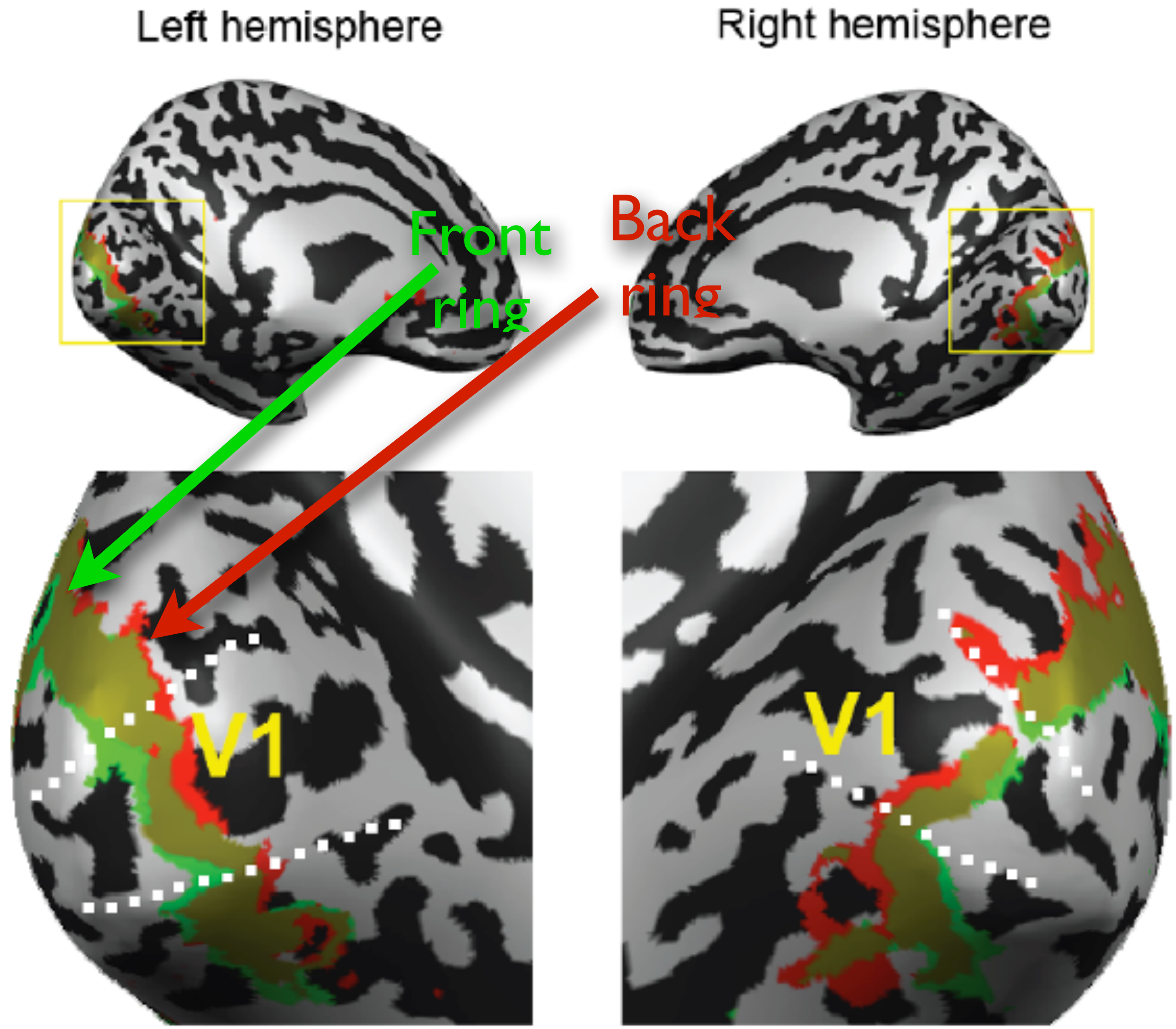


Huk, A. C. (2008) Visual Neuroscience: Retinotopy meets Percept-otopy, Current Biology, 18, 21, R1005-1007.

what was found for an illusory increase in ring size



Fang, Boyaci, Kersten, & Murray, S. O. (2008). Attention-dependent representation of a size illusion in human V1. Current Biology



attend-to-ring condition

Ni, A. M., Murray, S. O., & Horwitz, G. D. (2014). Object-Centered Shifts of Receptive Field Positions in Monkey Primary Visual Cortex. *Curbio*, 1–6

some proposed functions of feedback between visual cortical areas

- resolving local ambiguity using high-level knowledge
- binding information across levels of abstraction in the visual hierarchy
- accessing lower-level “expertise” as the task requires it

accessing lower-level “expertise”

hierarchically organized expertise

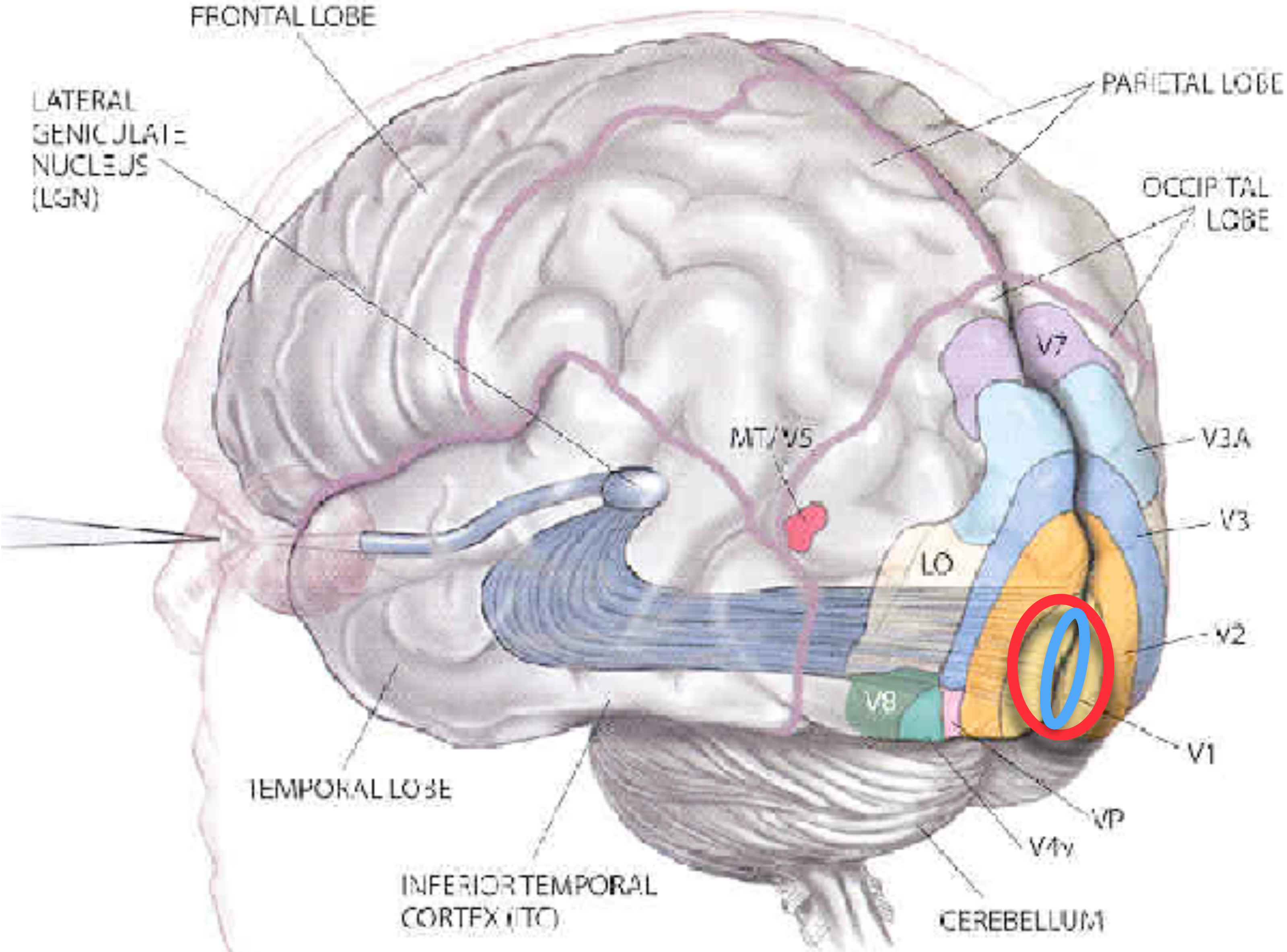
- Lee, T. S., Mumford, D., Romero, R., & Lamme, V. A. (1998); “Spatial buffer hypothesis”
- Hochstein, S., & Ahissar, M. (2002); “Reverse hierarchy theory”

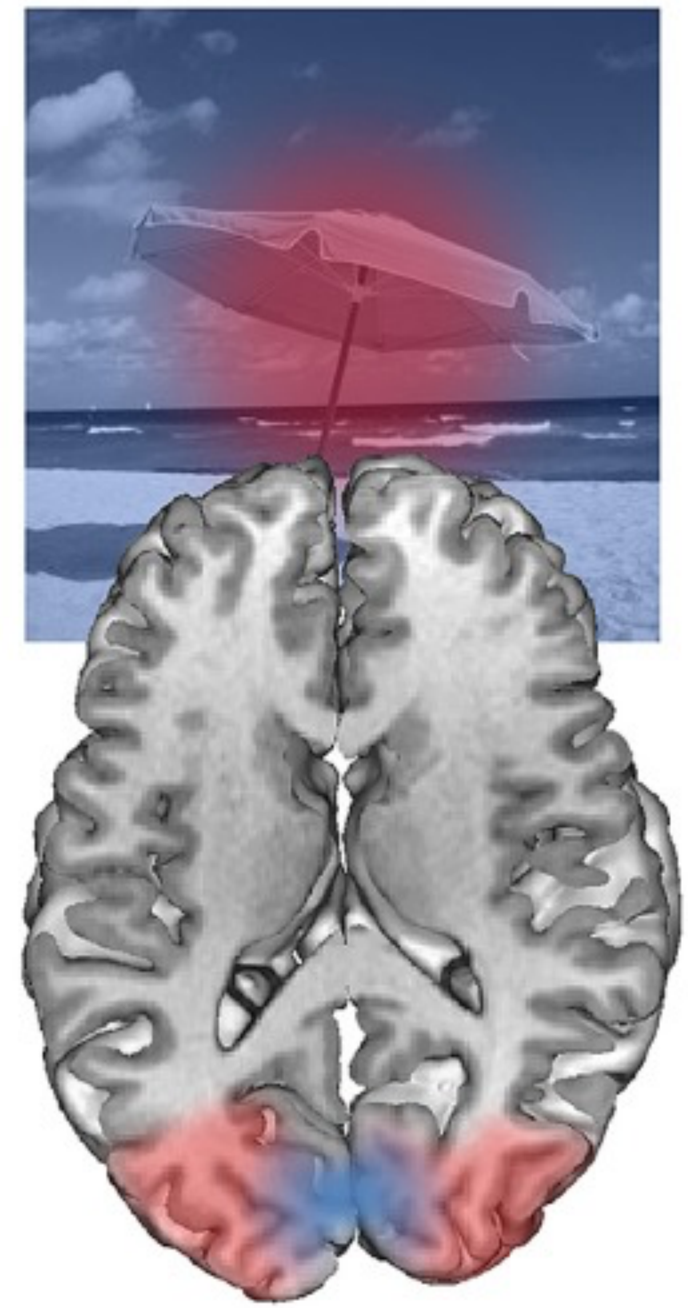
“executive metaphor” — emphasizes flexible top-down computations are foveal cortical neurons “consulted” for the analysis of detail in the absence of direct stimulation?

evidence from psychophysics

Fan, X., Wang, L., Shao, H., Kersten, D., & He, S. (2016). Temporally flexible feedback signal to foveal cortex for peripheral object recognition. PNAS.

some background

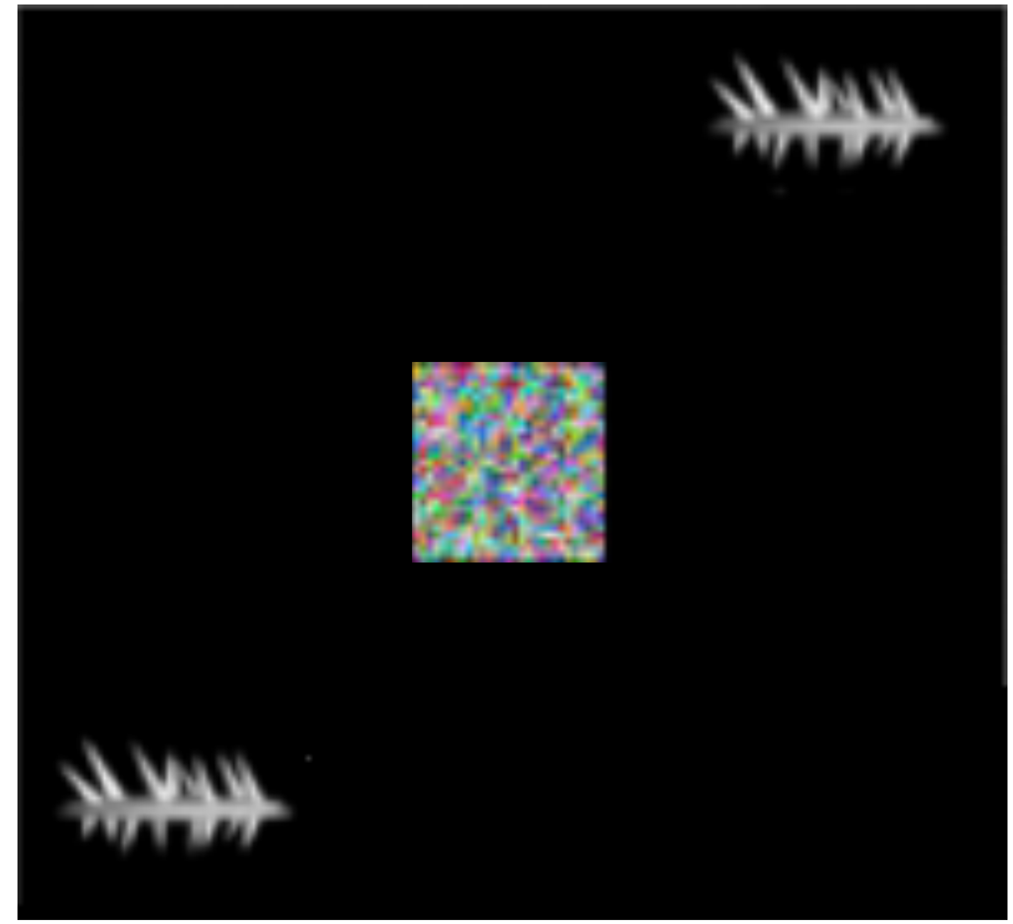




retinotopic property of early visual areas

some background

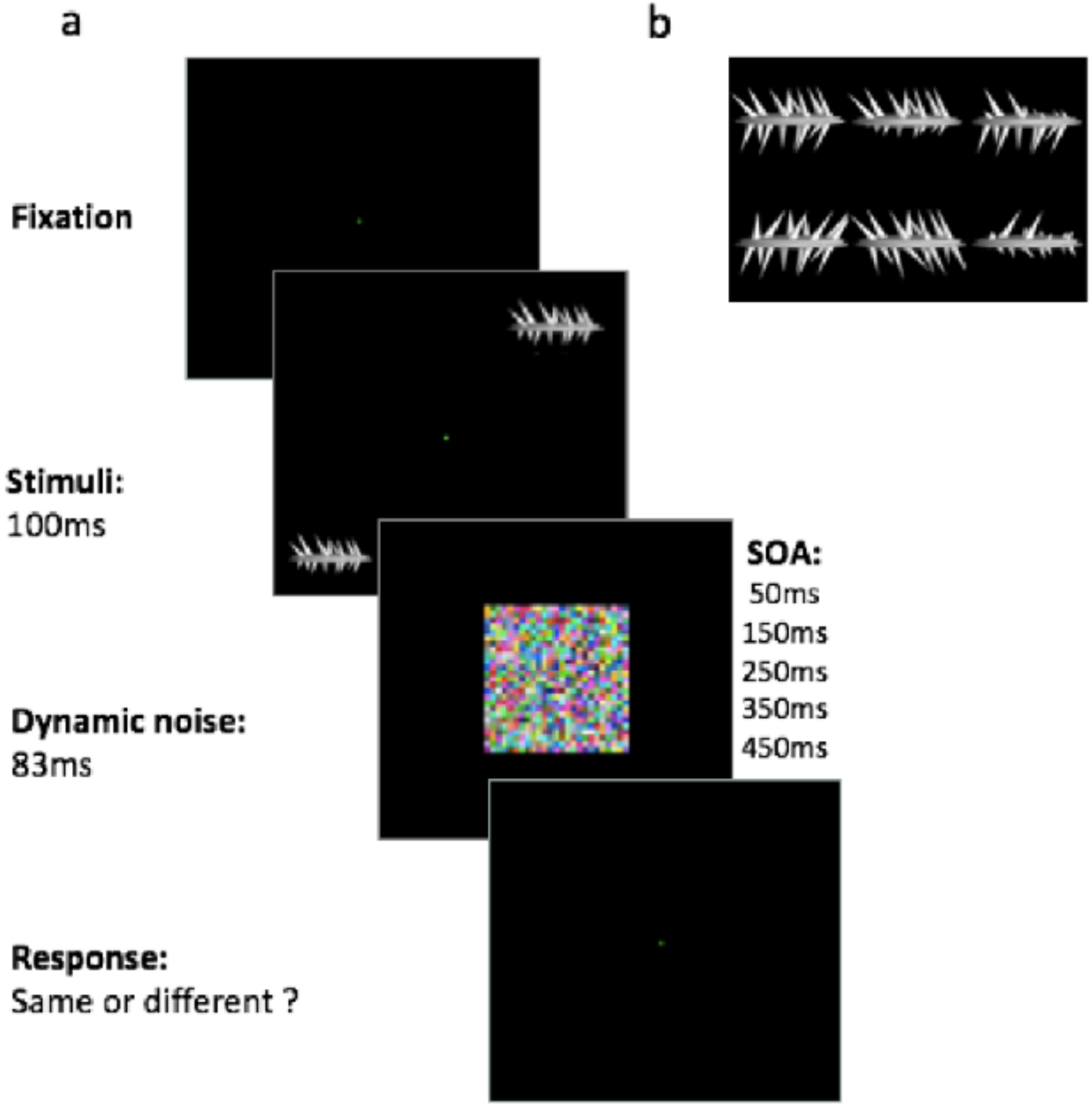
- Voxels in non-stimulated foveal V1 contain information about object category when observers make within-category discriminations
  - Williams, M. A., Baker, Op de Beeck, H. P., Shim, W. M., Dang, S., Triantafyllou, C., & Kanwisher, N. (2008)
- Transcranial magnetic stimulation (TMS) to foveal cortical areas most effective disrupting performance 350-400 msec after stimulus onset.
  - Chambers, C. D., Allen, C. P. G., Maizey, L. & Williams, M. (2013)
- Visual noise presented to fovea has a similar disruptive effect on task performance.
  - Wheldon et al. (2016); Yu Q & Shim WM (2016);

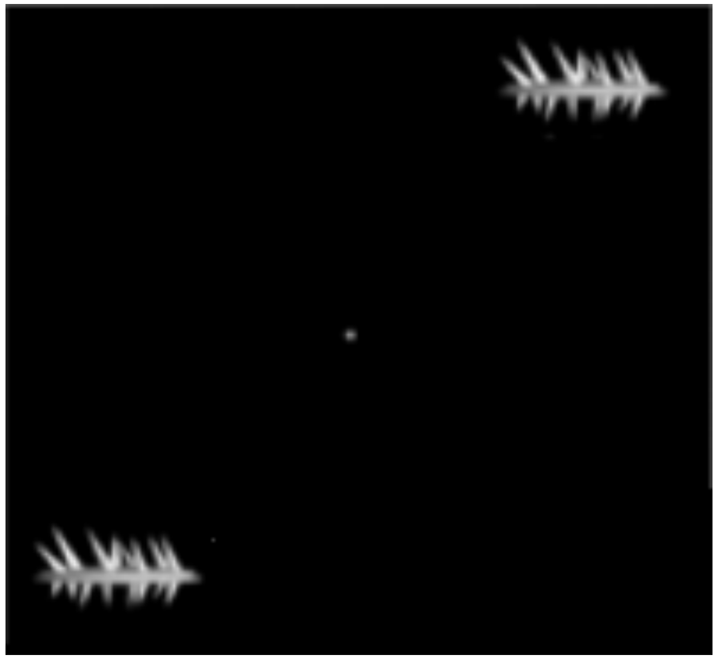


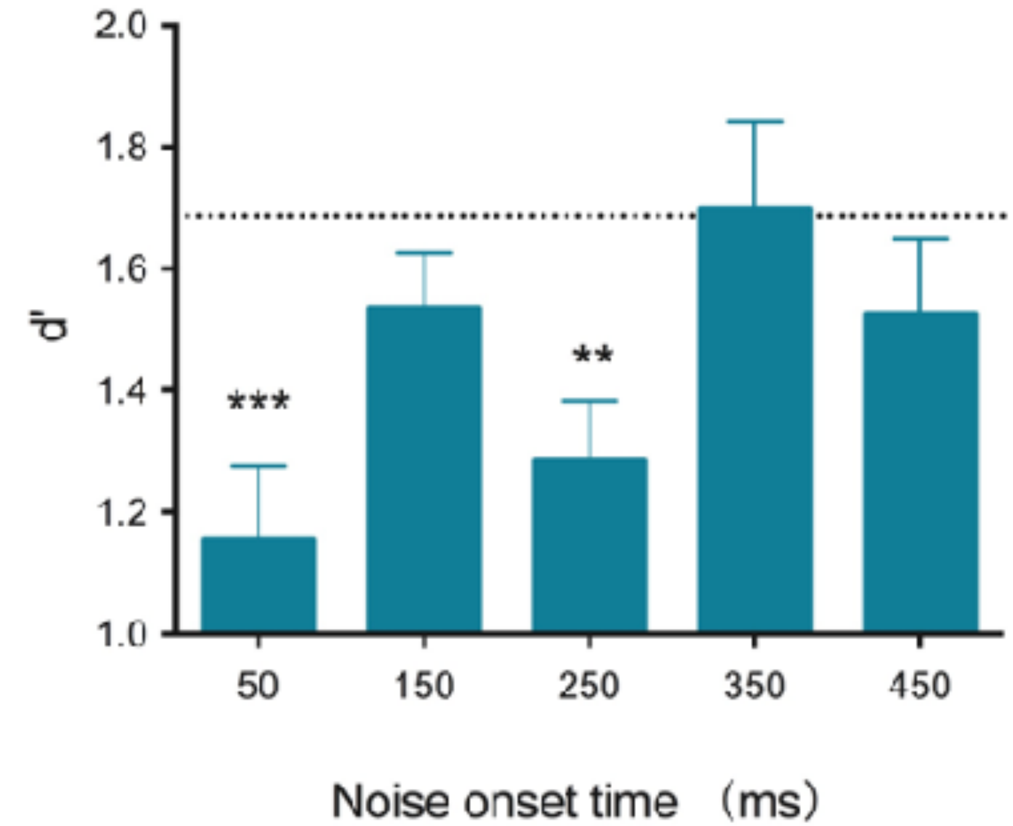
Is foveal processing only engaged for tasks requiring fine spatial detail?

Is deployment automatic or only when the task requires it?

Fan, X., Wang, L., Shao, H., Kersten, D., & He, S. (2016). Temporally flexible feedback signal to foveal cortex for peripheral object recognition. PNAS.


a
b
Fixation
Stimuli:
100ms
Dynamic noise:
83ms
SOA:
50ms
150ms
250ms
350ms
450ms
Response:
Same or different ?

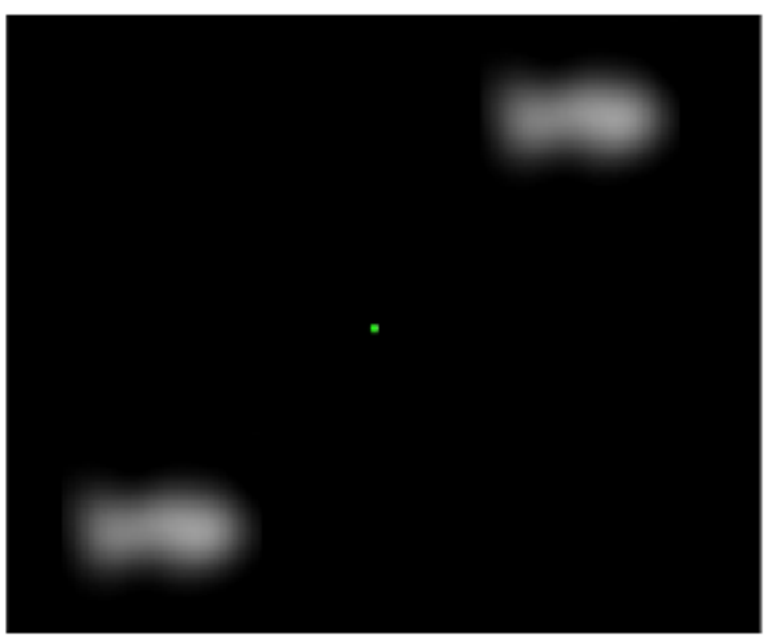




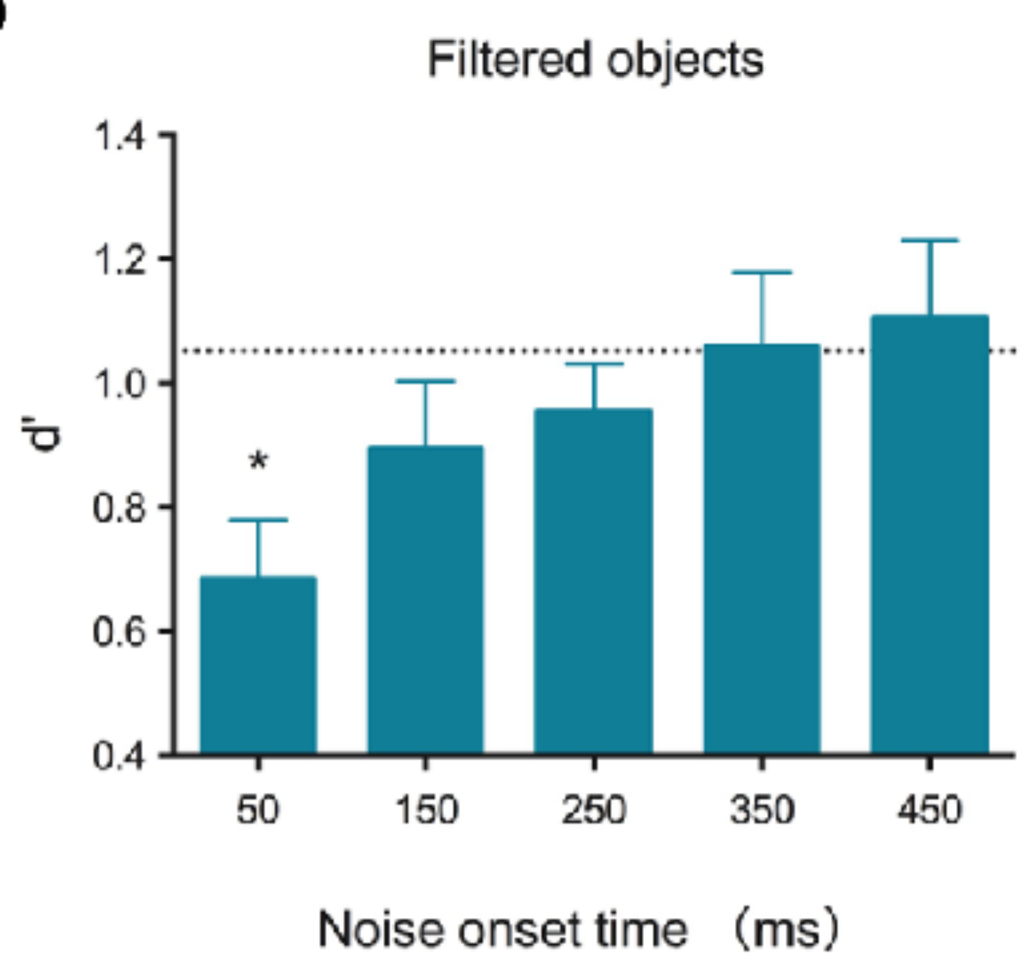
The temporal window when foveal noise disrupts the peripheral object discrimination occurs around 250 msec.

a

Filtered objects



b

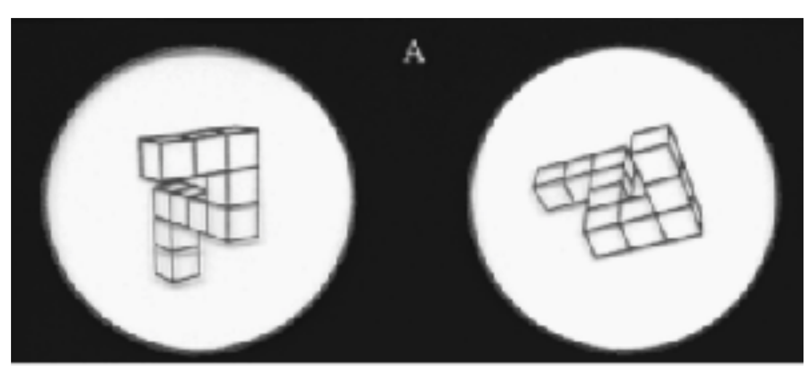


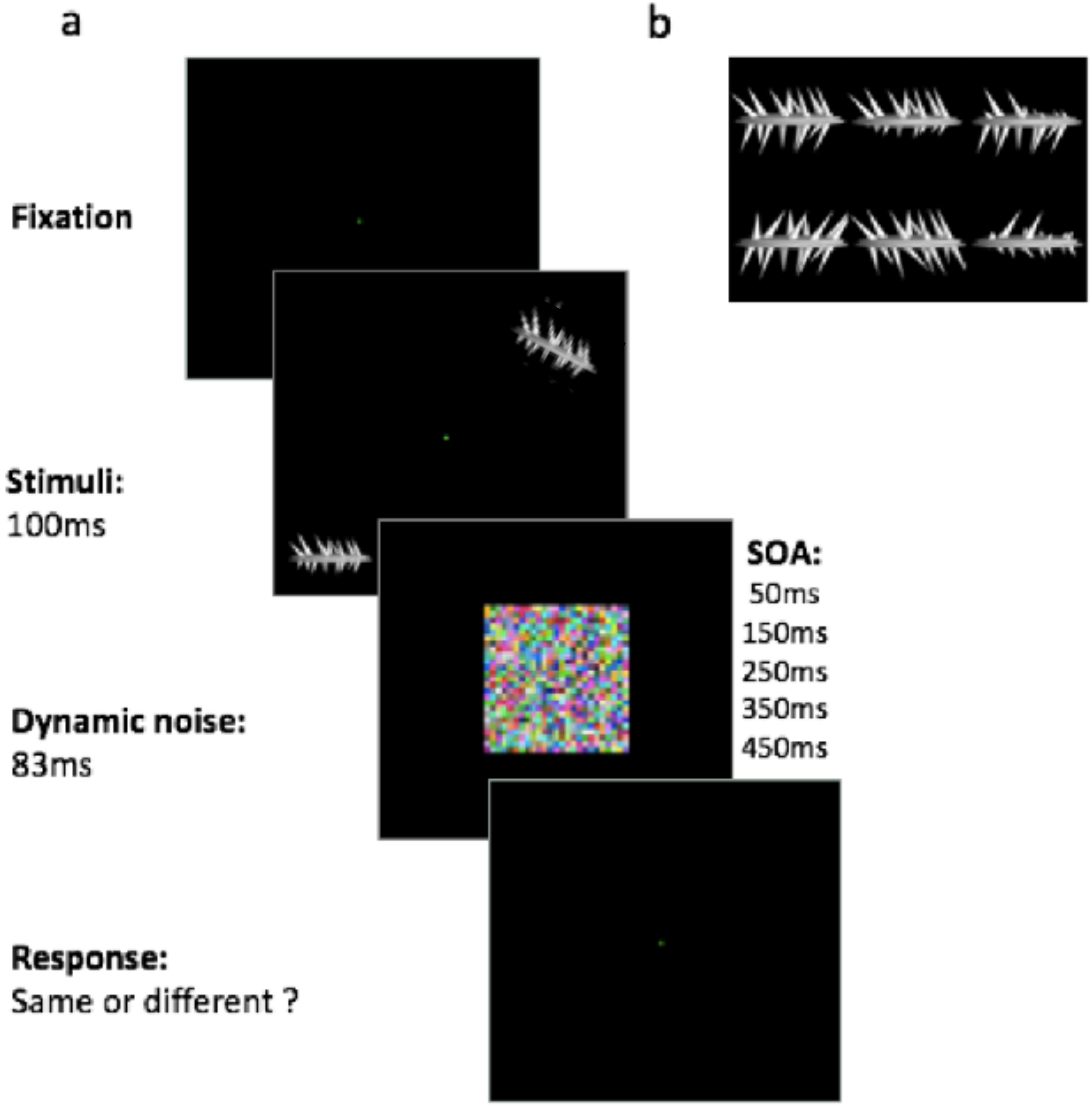
but no corresponding drop for low-pass filtered images of the objects

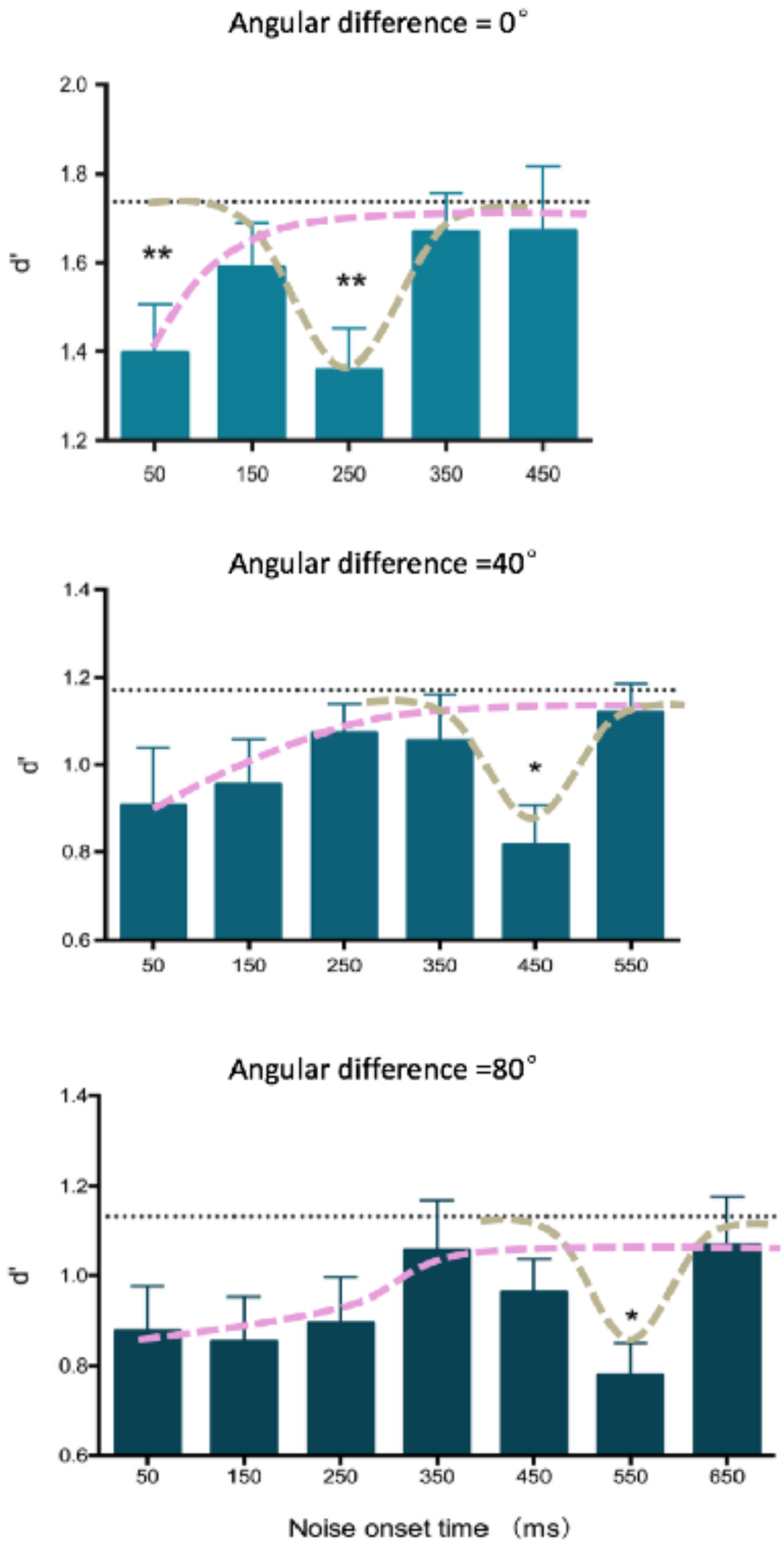
Is deployment automatic or only when
the task requires it?

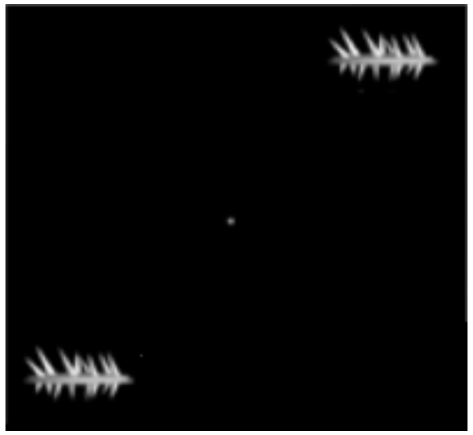
do the same experiment, but now incorporate
mental rotation

Shepard & Metzler, 1971; Cooper & Shepard, 1973; …

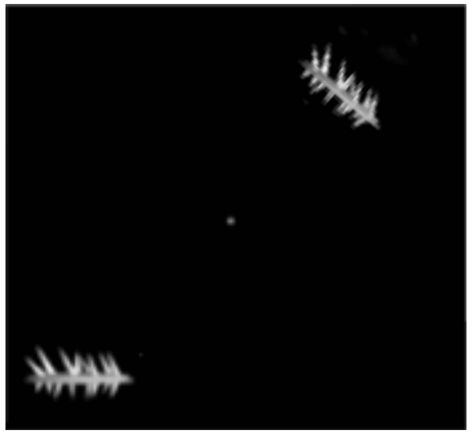


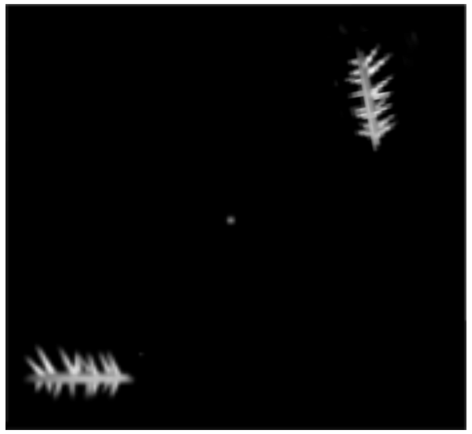

a
b
Fixation
Stimuli:
100ms
SOA:
50ms
150ms
250ms
350ms
450ms
Dynamic noise:
83ms
Response:
Same or different ?





The temporal window shifts the time that foveal noise disrupts the peripheral object discrimination when mental rotation was required as part of the peripheral object discrimination task.





Results are consistent with the idea that the foveal retinotopic cortex is not automatically engaged at a fixed time following peripheral stimulation, rather it occurs at a stage when higher level cortical areas are ready for and can use foveal cortical computations.

Further experiments show

- narrow time window
- closely coupled to saccade preparation
- fMRI: both category and image property information (patch orientation) could be recovered from patterns of activity in foveal voxels, not directly stimulated

Computational functions of feedback: evidence in early human visual cortex?

neuroimaging and psychophysics consistent with

- predictive coding
  - reduction of local ambiguity and signaling "unexplained" features
- binding
  - depending on segmentation complexity and/or access to low-level features
- psychophysical timing experiments requiring fine-grain discrimination of peripherally viewed objects consistent with feedback as accessing lower-level "expertise"

Thanks to

Sheng He, Cheryl Olman, Cheng Qiu, U Minn

X Fan, L Wang, Y Han, Chinese Academy of Sciences

Fang Fang, Dongjun He, Peking U

Damien Mannion, UNSW

Supported by NIH R01 EY015261 and ONR N000141210883., the WCU program of the Ministry of Education, Science and Technology through the National Research Foundation of Korea (R31-10008).