Local-Global Processing in Visual Crowding: a Bayesian Hierarchical Model

Shunan Zhang, Man Amanda Song, Angela J. Yu {s6zhang, , ajyu}@ucsd.edu University of California, San Diego, Department of Cognitive Science

Motivation

We study effect of perceptual grouping on visual crowding using hierarchical Bayesian model

'Local' vs 'Global' distinction in visual psychophysics and neuroscience

- Local processing: depends on spatially proximate visual elements
- Global processing: holistic, influenced by spatially distal elements

Recent experimental findings suggest more subtle and complex interactions between them [1]

Vernier Discrimination Task

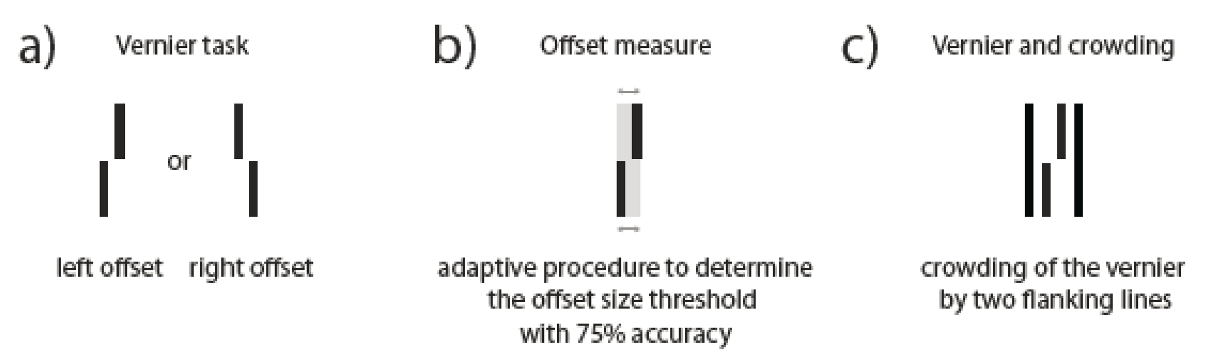

Figure 1: Vernier Discrimination Task

Generative Model

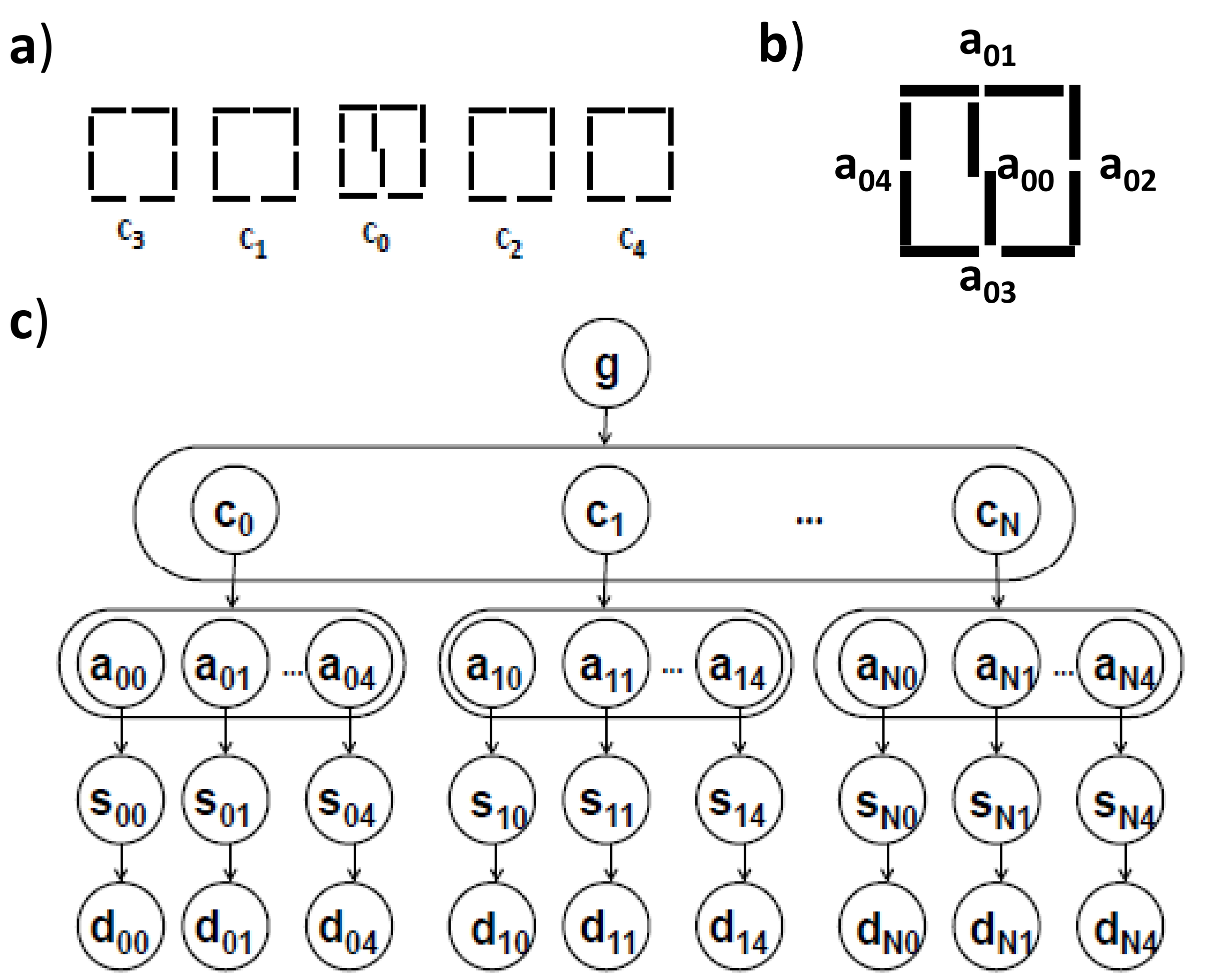

Figure 2: **a)** We model the visual scene as being made up as N decomposable stimuli; **b)** straight lines generates noisy sensory evidence for two well-aligned segments; **c)** generative model

Other Mechanistic Models

- Pooling [2]: neurons in higher visual areas with larger receptive fields pool information from lower-level neurons with smaller receptive fields
- Substitution [3]: features of the target and flankers are mislocalized or not "accessible" by attention

Both predict: crowding increase when additional flankers are added, since more irrelevant elements are pooled or confused

Generative Model Details

Group Level:

- g=1: same coherent object at all places
- g=0: unrelated objects/visual input

Object Level:

- c=0: no coherent visual object
- c=s: square (with or without vernier)
- c=w: window (vernier being part of object)

Line Level:

- a=0: no visual input
- a=1: well-aligned segments
- a=2: unrelated two segments

Conditional Probabilities:

- when g=1: c's same with large probability
- when g=0: c's are independent
- when c=0: a's are independent
- when c=s: large marginal probability for $a_1 = a_2 = a_3 = a_4 = 1$
- when c=w: large probability for $a_i = 1$, all i
- when a=1: s $\sim N(0, \sigma_{small})$
- when a=2: s $\sim N(0, \sigma_{large})$
- d $\sim N(s, \sigma)$

Recognition Model

$$\Pr(s_{00}, a_{00}, c_0, g|\mathbf{d}) \propto \Pr(g) \int \Pr(\mathbf{c}|g)...$$

$$\int \Pr(\mathbf{a}_n|c_n) \prod \Pr(s_{ni}|a_{ni}) \Pr(d_{ni}|s_{ni})$$

Grouping Effects at Global/Object Level

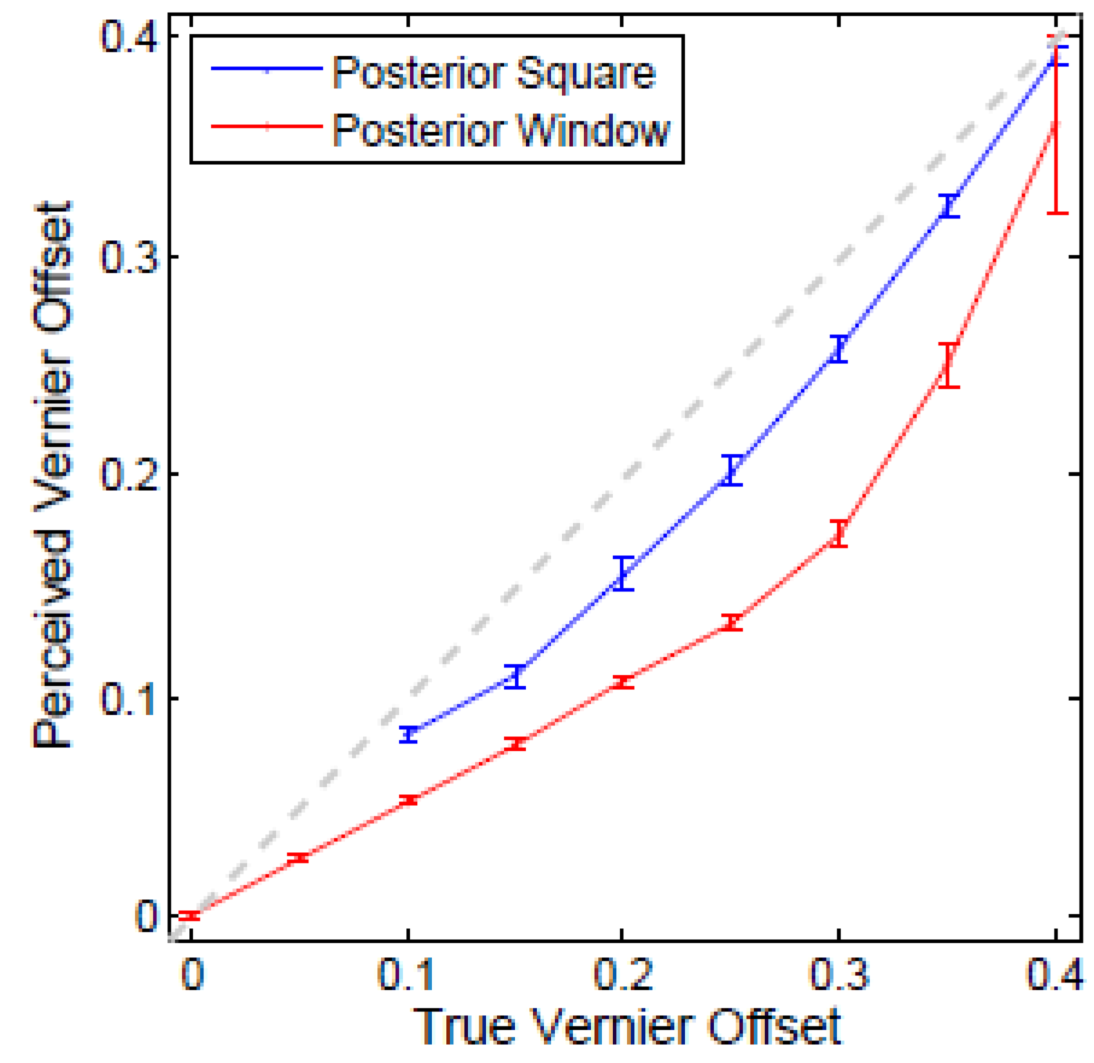

Figure 3: Bias of perceived offset distance is larger when posterior probability of "windowness" is greater than .5

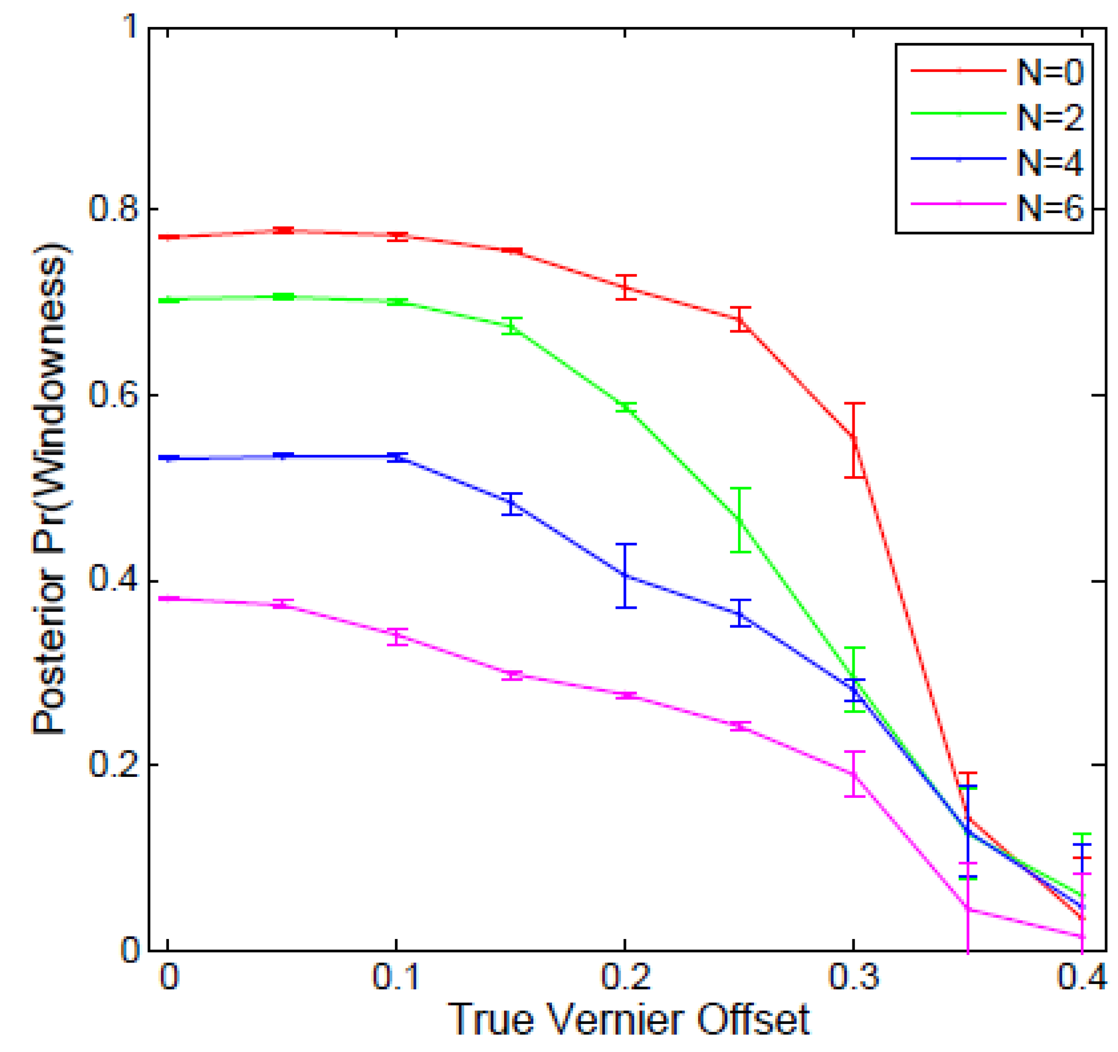

Figure 4: The posterior of "windowness" decreases as the number of additional flankers increases

Simulation Results

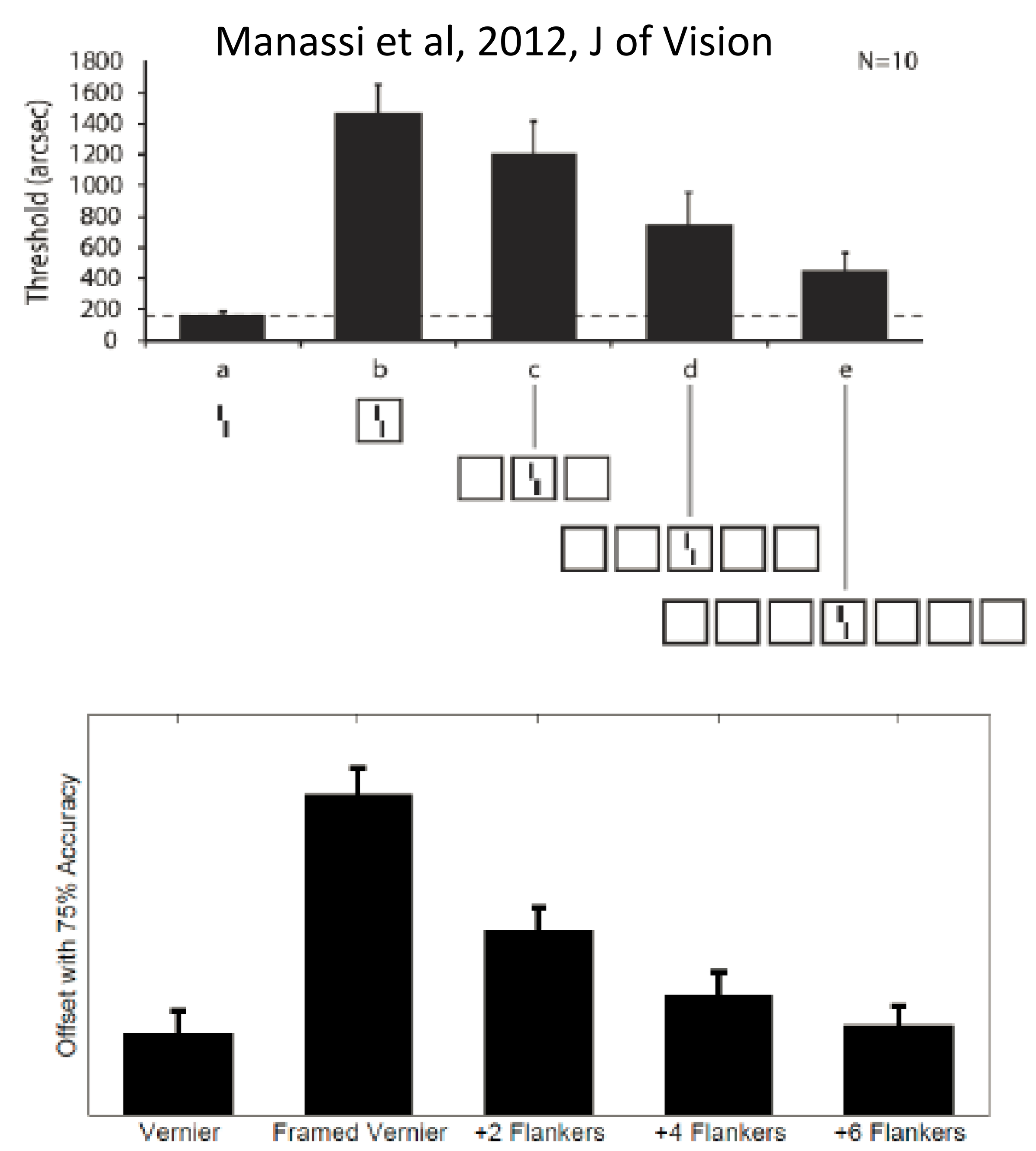

Figure 5: Top: see [1]; Bottom: model simulated thresholds for 75% accuracy: adding one embedding square deteriorates performance, while adding additional flankers improves performance

Summary

- Our Bayesian model makes simultaneous inferences about (relatively) global grouping membership and local visual features
- New approach for explaining uncrowding; bidirectional information flow
- related to other mechanistic models that involve the feed-forward and feedback loops

Future Work

- Principled decision component
- Temporal aspect
- Explain broader crowding vs uncrowding
- Account for phenomena involving complex interactions among stimuli
- Adding front end

References

[1] Manassi, Sayim, Herzog (2013) When crowding of crowding leads to uncrowding. *J of vis*.

[2] Parkes et al (2001) Compulsory ave. of crowded orientation signals in human vis. *Nat neurosci*.

[3] Huckauf, Heller (2002) What various kinds of errors tell us about lateral masking effects. *Vis Cogn*.