Measuring Human Probabilistic Segmentation Maps

Jonathan Vacher (jonathan.vacher@einsteinmed.org), Pascal Mamassian, Ruben Coen-Cagli

Einstein Albert Einstein College of Medicine · Montefiore · CNRS · ENS · PSL

Introduction

Visual segmentation is a core function of biological vision:

- involves Gestalt principles, *e.g.* grouping by similarity, proximity and good continuation [1]
- visual cortical neurons are sensitive to those cues [2]

Feedforward models: comparing the local summary statistics of low-level visual features [3, 4].

Alternative view: perceptual segmentation is probabilistic.
To test that:

- we propose a new protocol to measure segmentation maps and variability
- we measured segmentation maps of composite artificial images
- we compared how these two models predict human responses

Probabilistic Segmentation Maps: Measure and Reconstruction

A new task to measure segmentation maps:

- Ask the participant to decompose the image in K segments
- Show the image for 3 s
- Run a sequence of M trials: does the pair belong to the same segment ?

For any response model $p_{\mathbf{i},\mathbf{j}}(\Theta)$, the MLE estimate is

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmin}} \sum_{(\mathbf{i},\mathbf{j})\in\mathcal{P}} \|k_{\mathbf{i},\mathbf{j}} - p_{\mathbf{i},\mathbf{j}}(\Theta)\|^2 + \text{reg.} \quad (1)$$

When $\Theta = ((p_\mathbf{i})_\mathbf{i})$, the probability of response verifies

$$p_{\mathbf{i},\mathbf{j}}(\Theta) = \langle p_\mathbf{i}, p_\mathbf{j} \rangle.$$

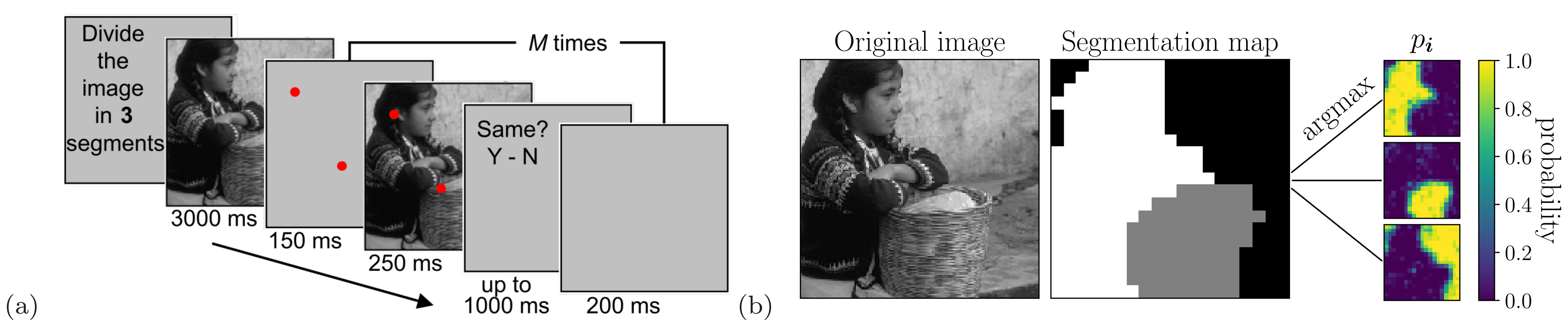
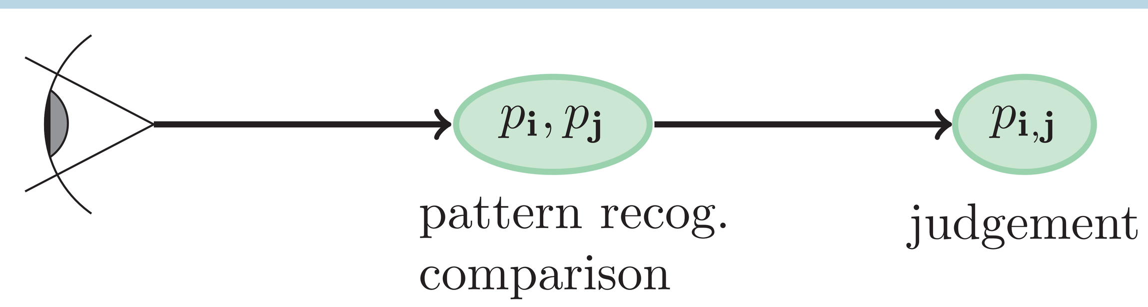


Figure 1: (a) Experiment trial layout. (b) Segmentation map of a natural image and probability of assignment to each segment, obtained with our protocol.

Models

(i) a non-parametric model $[\Theta = ((p_\mathbf{i})_\mathbf{i}, \alpha)]$

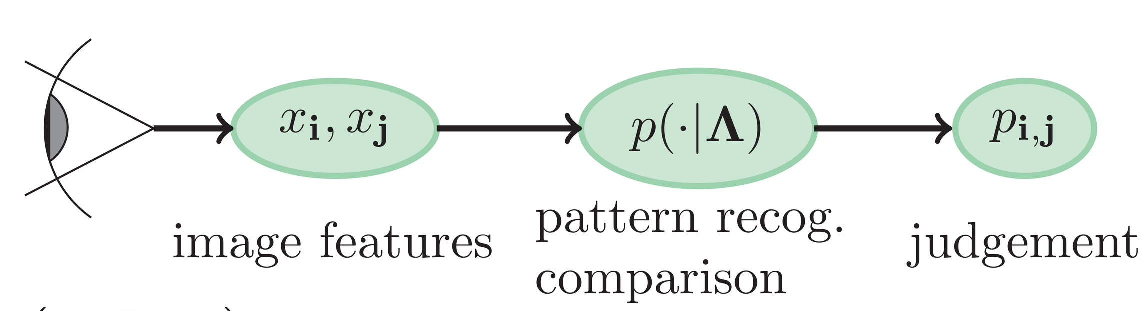
$$p_{\mathbf{i},\mathbf{j}}(\Theta) = \alpha + (1-2\alpha)\langle p_\mathbf{i}, p_\mathbf{j}\rangle \quad \text{(NP)}$$



- assumes the existence of underlying probability maps
- $p_{\mathbf{i}_k}$: probability that pixel $\mathbf{i}$ belong to segment k
- α: lapse rate

(ii) a generative model $[\Theta = (\Lambda, \alpha)]$

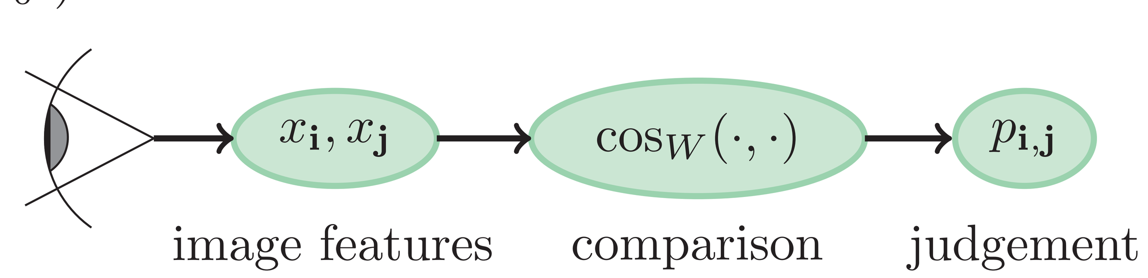
$$p_{\mathbf{i},\mathbf{j}}(\Theta) = \alpha + (1-2\alpha)\langle p(x_\mathbf{i}|\Lambda), p(x_\mathbf{j}|\Lambda)\rangle \quad \text{(GM)}$$



- assumes the probability maps are obtained via probabilistic inference
- $p_k(x|\Lambda) \propto \exp\left(-\frac{x^T \Lambda_k x}{2}\right)$ w/ $\Lambda_k = (\Sigma_k + \sigma_0 \mathrm{I})^{-1}$
- Σ_k, σ_0: feature covariance and internal noise

(iii) a discrimination model $[\Theta = (W, \mu, \sigma, \alpha)]$

$$p_{\mathbf{i},\mathbf{j}}(\Theta) = \alpha + (1-2\alpha)S_{\sigma,\mu}(\cos_W(x_\mathbf{i}, x_\mathbf{j})) \quad \text{(FD)}$$



- assumes that local features are directly compared
- $S_{\sigma,\mu}(u) = \left(1 + \exp\left(-\frac{1}{\sigma}\left(\log\left(\frac{u}{1-u}\right) - \mu\right)\right)\right)^{-1}$
- μ, σ: subjective eq. and inverse sensitivity

Results

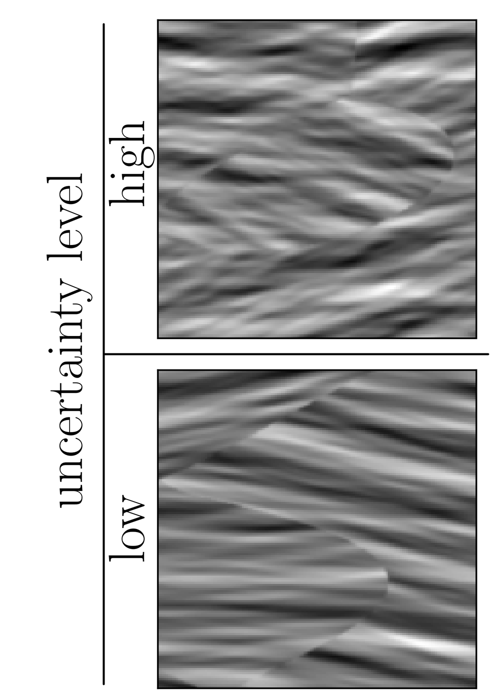


Figure 2: High and low uncertainty stimuli.

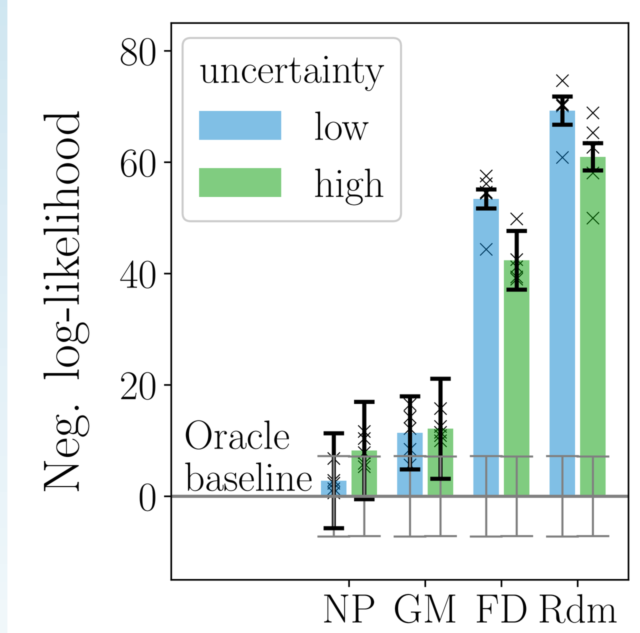


Figure 3: Fit quality (cross-val. negative log-lkl, lower is better). Rdm: chance level.

- manipulating the orientation and spatial frequency distributions of the textured segments changes the segmentation uncertainty – Figure 2
- the probabilistic inference model (GM) explains the data better than the feature discrimination model (FD) – Figure 3
- variability of human segmentation correlates with image uncertainty – Figure 4
- GM captures the variability that is intrinsic to image uncertainty, differences with NP account for other factors such as measurement noise and inter-participants variability – Figure 4
- variability is concentrated around contours, this effect is stronger for low uncertainty stimuli (blue) where contours are more spatially localized – Figure 5

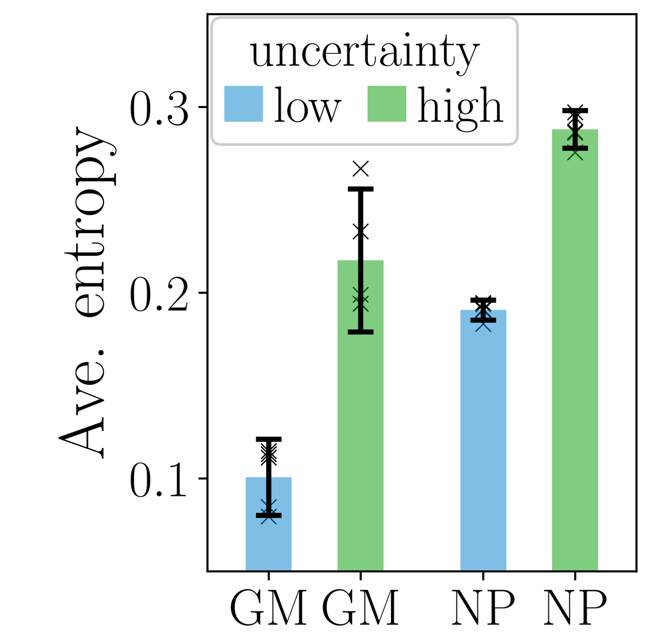


Figure 4: Average entropy.

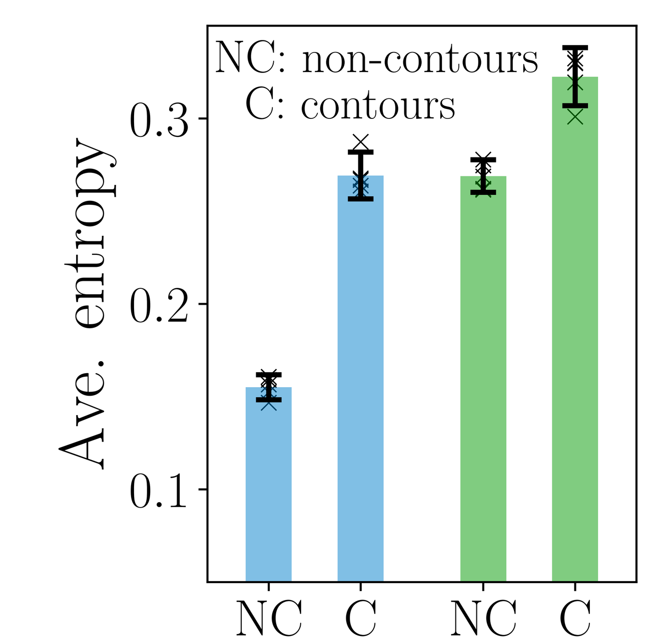


Figure 5: Average entropy for contour and non-contour areas obtained with NP model. Error bars: 99.7% conf. interval.

Optimal observer

- we compared the fitted covariances $\hat{\Sigma}_k$ (*i.e.* the internal representation of the average participant) to the ground truth covariances of the stimuli
- participant covariances were narrower for low-uncertainty than for high-uncertainty stimuli, and qualitatively followed the ground truth despite being broader overall

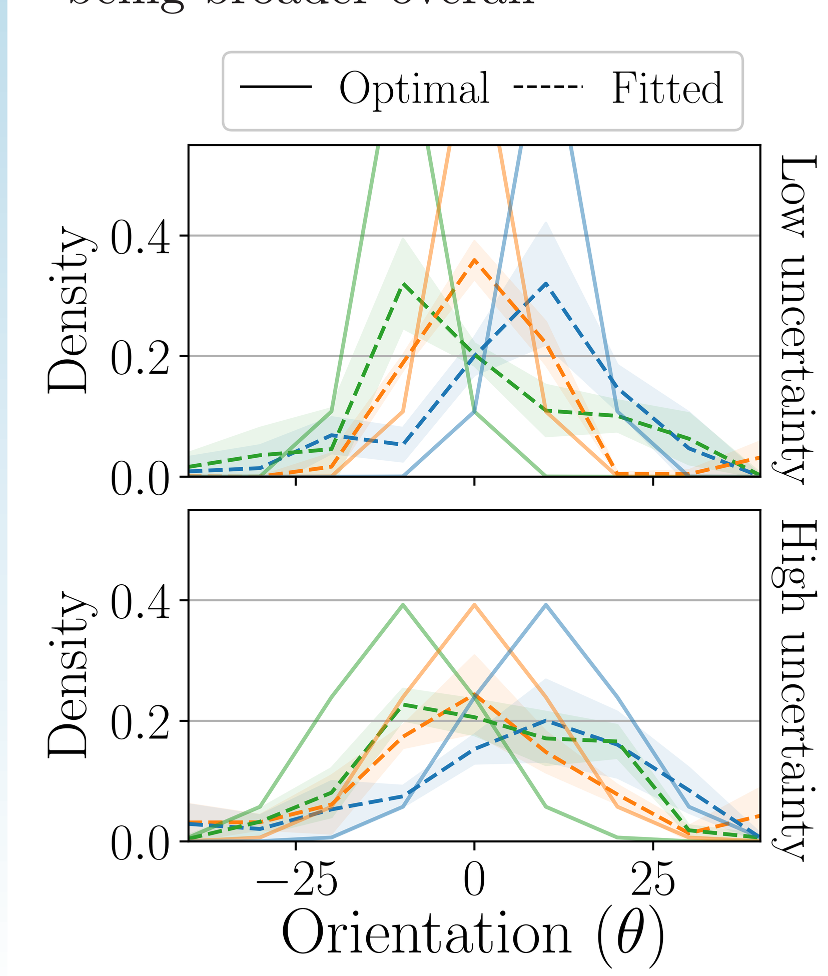


Figure 6: Ground truth covariances from the stimuli (optimal) and the fitted covariances of GM.

Summary

- human variability correlates with image uncertainty
- variability is localized around contours
- strong evidence that human segmentation is probabilistic
- a new protocol that will allow studying natural image segmentation

Funding

This collaborative work is jointly funded by NIH (NIH/CRCNS EY031166) in USA and ANR in France.

References

[1] J. Wagemans et al. In: *Psychological bulletin* 138.6 (2012), p. 1172.

[2] P. R. Roelfsema. In: *Annu. Rev. Neurosci.* 29 (2006), pp. 203–227.

[3] M. S. Landy and J. R. Bergen. In: *Vision research* 31.4 (1991), pp. 679–691.

[4] J. Freeman and E. P. Simoncelli. In: *Nature neuroscience* 14.9 (2011), p. 1195.