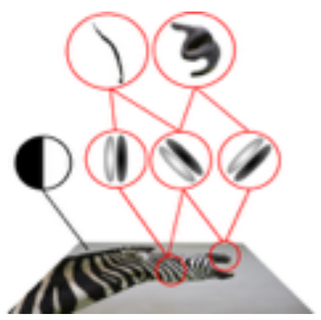
Sampling in a hierarchical model of images reproduces top-down effects in visual perception

Mihály Bányai, Gergő Orbán

Computational Systems Neuroscience Lab, Wigner Research Centre for Physics, Budapest, Hungary

Introduction

- Higher-order statistics
- Gestalt effect
- CSM model



$$p(v \mid g) = \mathcal{N}(v; 0, \sum_{j=1}^{K} g_j C_j)$$

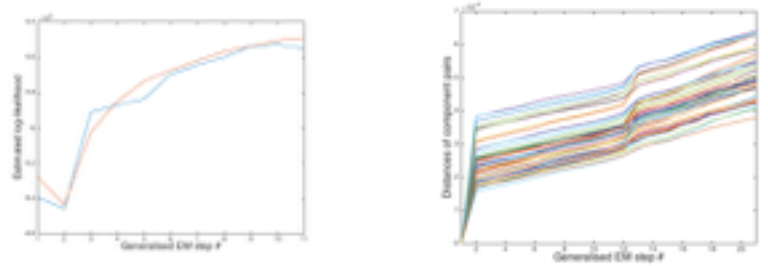
$$p(x \mid v, z) = \mathcal{N}(x; zAv, \sigma_x I)$$

Learning the components

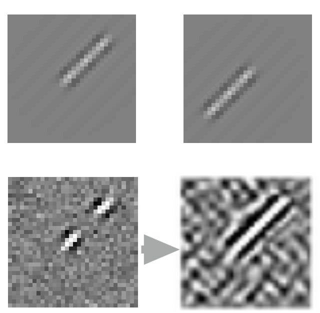
$$C_v = \sum_{k=1}^{K} g_k U_k^T U_k$$

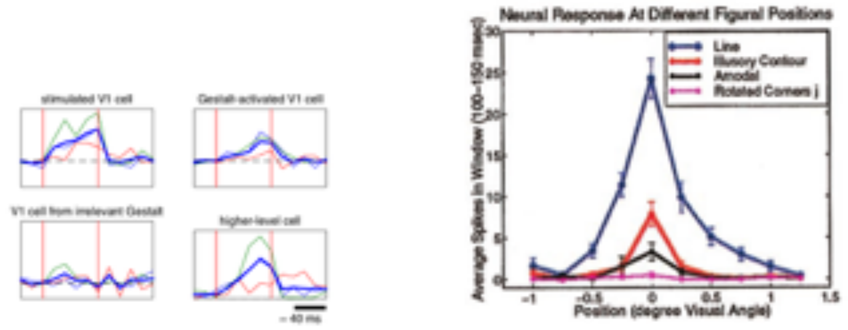
$$\frac{\partial \mathcal{L}}{\partial \left[U_k\right]_{i,j}} = \sum_{l=1}^{NL} \mathrm{Tr}\left[\frac{\partial \log p(x^l, v^l, g^l \mid U_{1 \ldots K})}{\partial C_v^l} \frac{\partial C_v^l}{\partial \left[U_k\right]_{i,j}}\right]$$

$$[U_k]_{i,j}^{new} = [U_k]_{i,j}^{old} + \epsilon \frac{\partial \mathcal{L}}{\partial \left[U_k\right]_{i,j}}$$

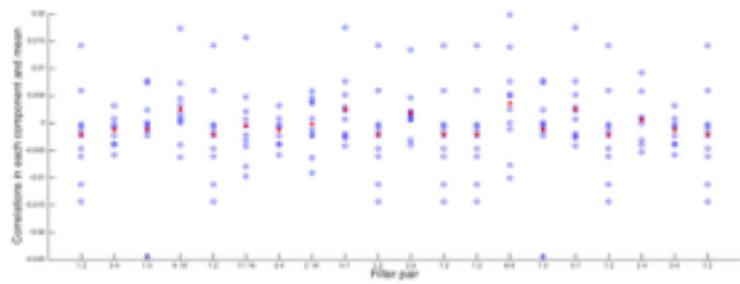


Predicted response to illusory contours





Relation to GSM

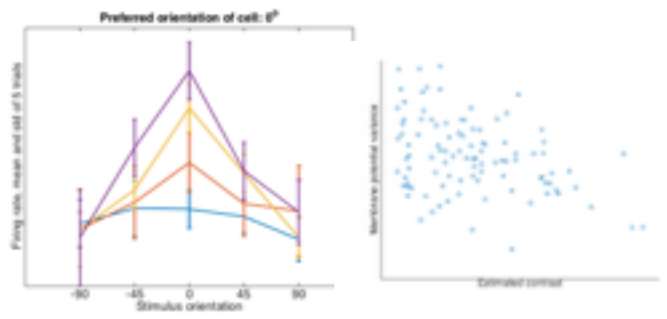


Sampling the posterior

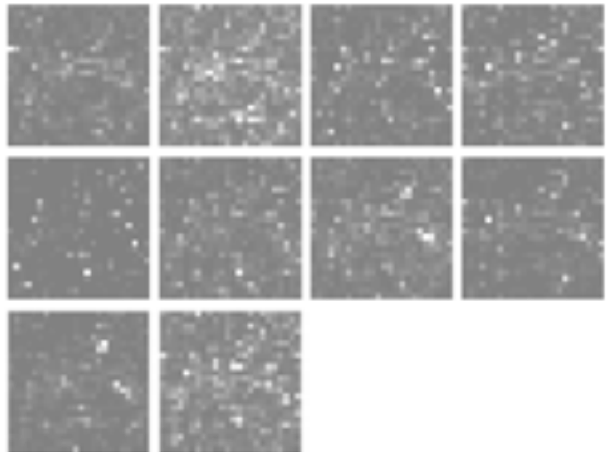
$$p(v \mid x, g, z) = \mathcal{N}\left(v; \frac{z}{\sigma_x} C_{v|x,g,z} A^T x, C_{v|x,g,z}\right), \quad C_{v|x,g,z} = \left(\frac{z^2}{\sigma_x} A^T A + \left(\sum_{j=1}^{K} g_j C_j\right)^{-1}\right)^{-1}$$

$$\log p(g \mid x, v, z) \sim -\frac{1}{2}\left[\log\left(\det\left(\sum_{k=1}^{K} g_j C_j\right)\right) + v^T \left(\sum_{k=1}^{K} g_j C_j\right)^{-1} v\right] + \log p(g)$$

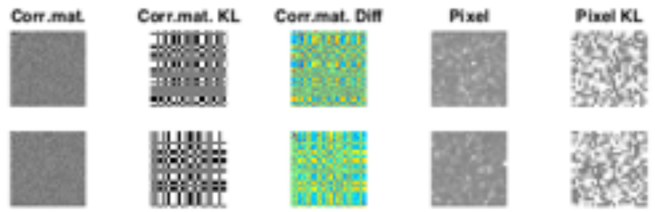
$$\log p(z \mid x, v, g) \sim -\frac{1}{2}\left[D_x \log(\sigma_x) + \frac{1}{\sigma_x}(x - zAv)^T (x - zAv)\right] + \log p(z)$$



Correlations implied by natural statistics



Relation to component models



Conclusions

- Higher-order statistics
- Gestalt effect