TNS Journal Club: "Scale Mixtures of Gaussians and the Statistics of Natural Images", Martin Wainwright and Eero Simoncelli. Advances in Neural Processing Systems, 2000

Rich Turner (turner@gatsby.ucl.ac.uk)

Gatsby Unit, 27/04/2005

General theme and area

- Last week we saw: A functional explanation of sensory non-linearities in terms of efficient coding of natural signals.

- The statistics of photographic images, when represented using multi-scale wavelet bases exhibit two forms of striking non-Gaussian behaviour:
  - marginals of the coefficients have heavy tails
  - the conditionals are bow-tie like

- Used a generalised form of feed-forward divisive normalisation to produce independent outputs from these coefficients.

- This week we develop these ideas further:

Goal and Outline of the paper

GOAL: A top-down statistical description of signal generation that captures key bottom-up statistical characteristics of filter responses to images

Going to motivate a forward-model for image generation in the following stages:

1. Recap: Show that responses of a linear model exhibit striking statistical dependencies.

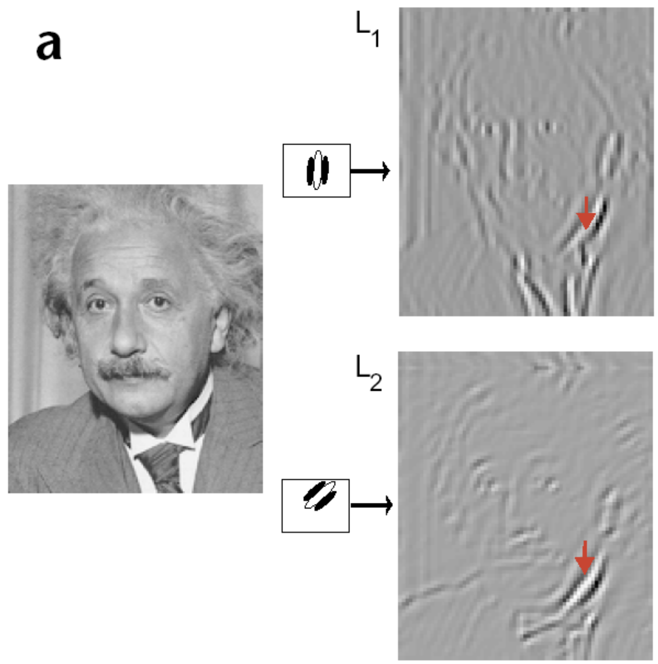
2. Show that Gaussian Scale Mixtures can characterise these pair-wise and marginal dependencies.

3. Derive a non-linear form of normalisation.

4. Introduce a parameterisation to capture the dependency between nearby filters that falls off with increased separation.

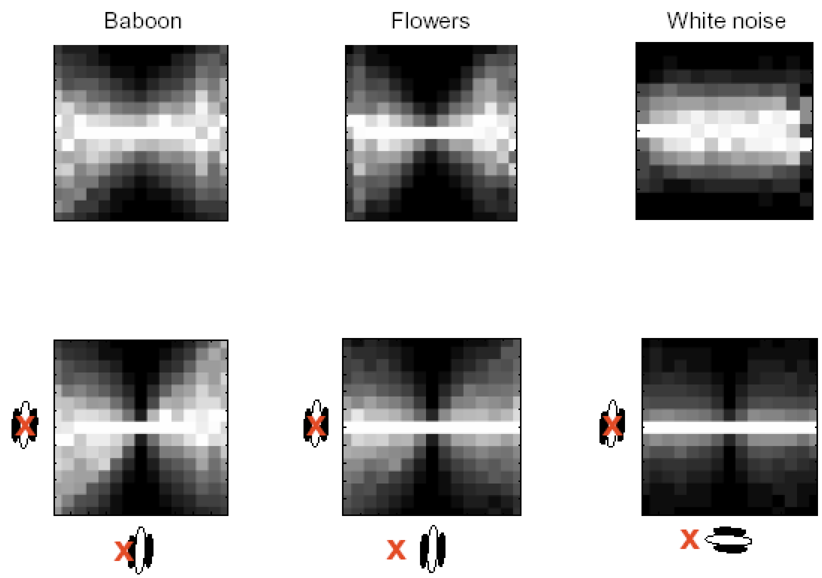
Recap: Statistics of linear filter responses, 1

stimulus $\longrightarrow$ 2 linear filters $\longrightarrow$ output
Filter pair: *Steerable pyramid shifted and rotated*



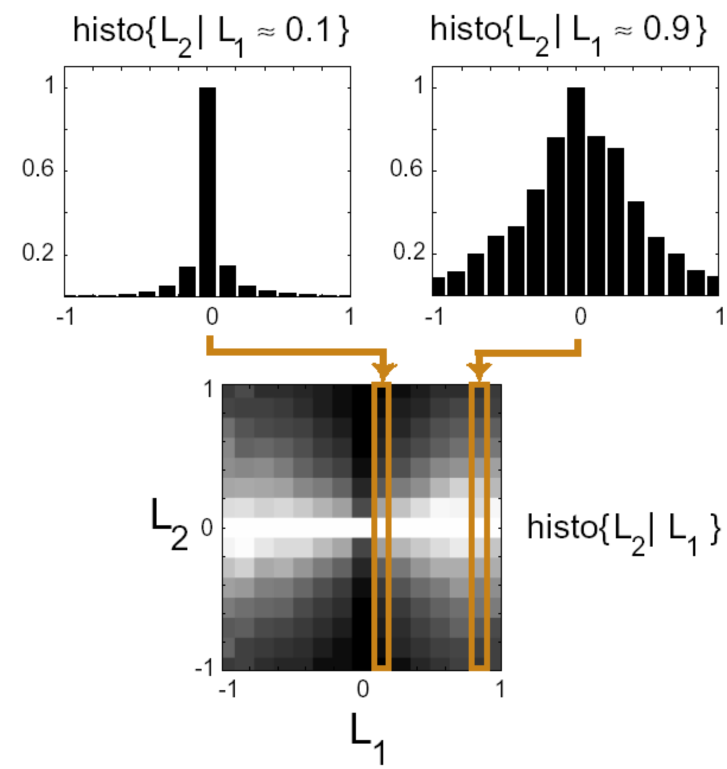
Recap: Statistics of linear filter responses, 2

Conditional histograms. Vertical cross-sections are not identical. **Top**: Previous filter pairs with different images. The dependency is strong for natural stimuli but weak for white noise. **Bottom**: Fixed stimulus, different filters. The strength of the dependence depends on the filter pair.



Recap: Statistics of linear filter responses, 3

Vertical slices of the conditional histogram. Differing widths indicate the dependency. The distributions appear to be heavy tailed (non-Gaussian). **Marginals are heavy tailed too.**



Gaussian Scale Mixtures (GSMs)

- $Y = zU$
  - $z \geq 0$ a scalar random variable
  - $U \sim G(0, Q)$
  - z and U are independent
- density of these *semi parametric* models can be expressed as an integral:

$$P(Y) = \int P(Y/z)P(z)dz = \int \frac{1}{[2\pi]^{\frac{N}{2}}|z^2Q|^{1/2}} exp\left(-\frac{Y^TQ^{-1}Y}{2z^2}\right)\psi(z)dz \qquad (1)$$

- One example is the MOG model [$\psi(z)$ is discrete]
- Investigate $\psi(z) = Gamma$, look at the marginals, joints and conditionals

Intuition for the properties of these models

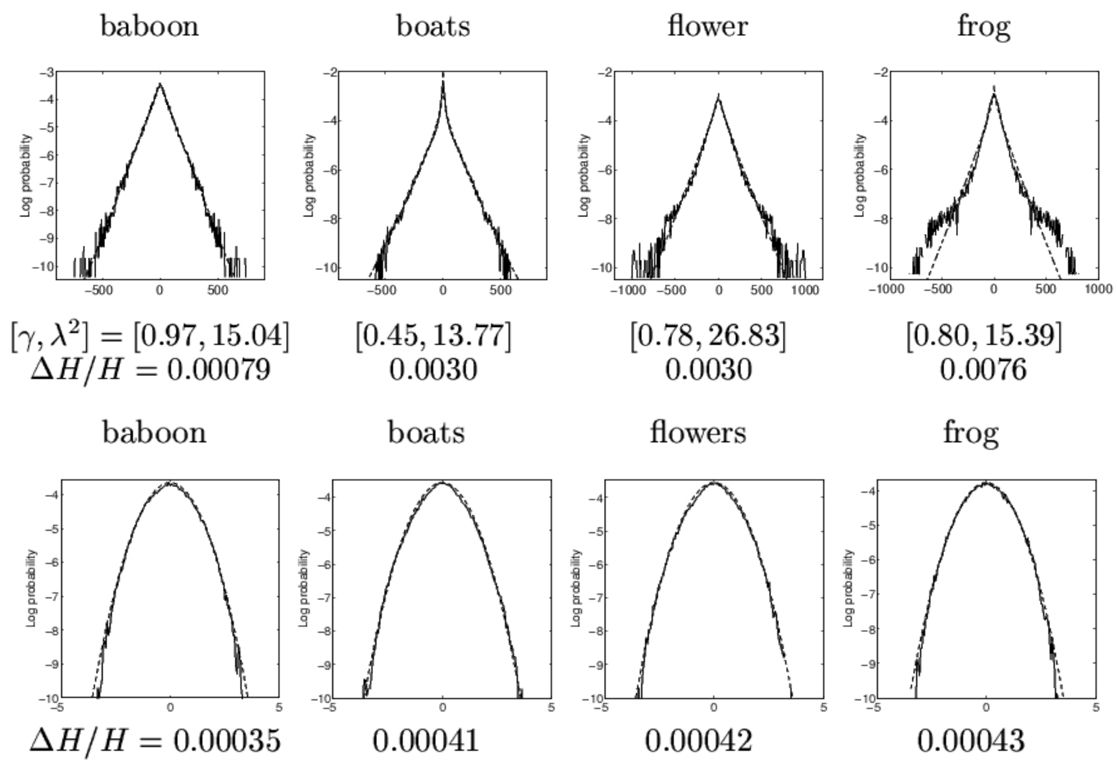
- Viewing the model as a 'weighted' sum of Gaussians we can see how it produces heavy-tailed marginal and variance-scaling joint densities.

Testing the model: Normalisation of components 1

- $Y = zU$ and therefore the prediction is: $Y/z \sim G(0, Q)$
- assume that 11 filters (different scales, orientations, positions) in a neighbourhood share a mixer
- estimate z by maximising the likelihood $\hat{z} = argmax_z \log P(Y|z)$ where $P(Y|Z) = \prod_i Norm_{y_i}(0, Qz^2)$.
- thus $\hat{z} = \sqrt{Y^t Q^{-1} Y/N}$
- identical to the divisive normalisation we saw last week up to an additive constant, if the filters are decorrelated: $\hat{z} = \sqrt{\sum_j w_j y_j^2 + \sigma_0^2}$.

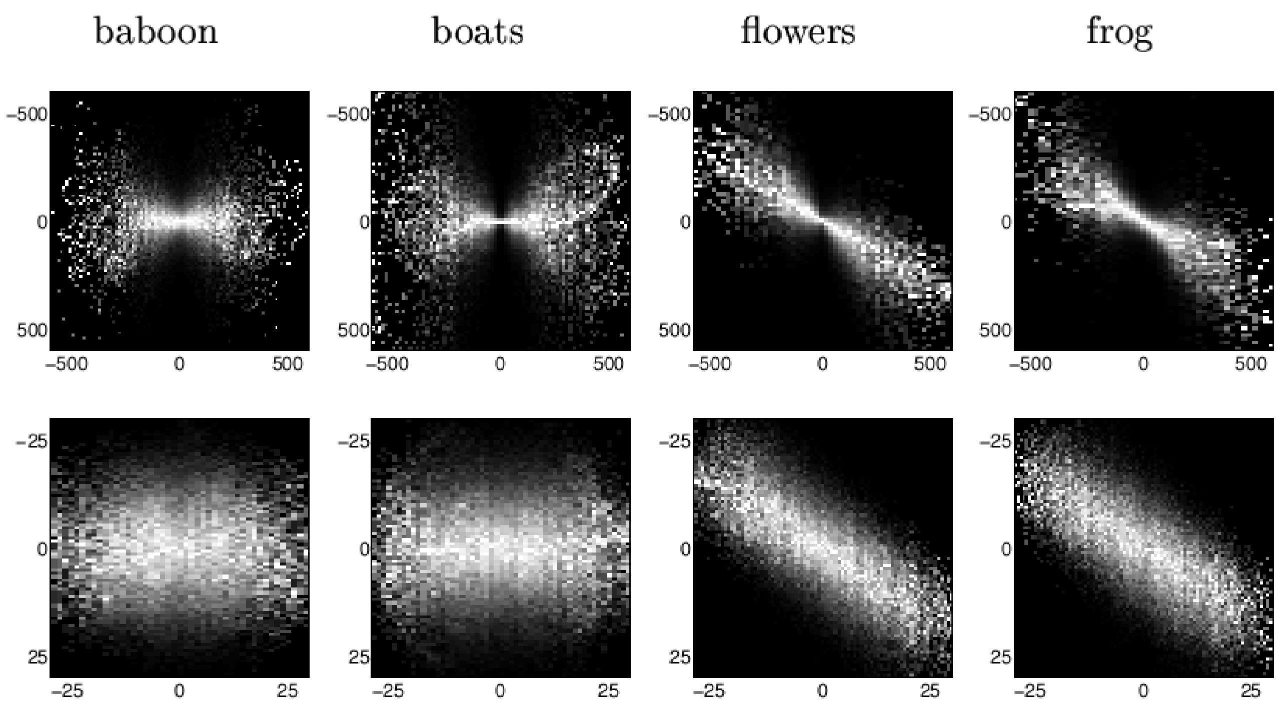
Normalisation of components 2 : Marginal statistics

Proving the normalised marginals are Gaussian. Top: The model can capture the marginal image statistics (fit via KL). Bottom: proving the normalised marginals are Gaussian



Normalisation of components 3: Joint Statistics

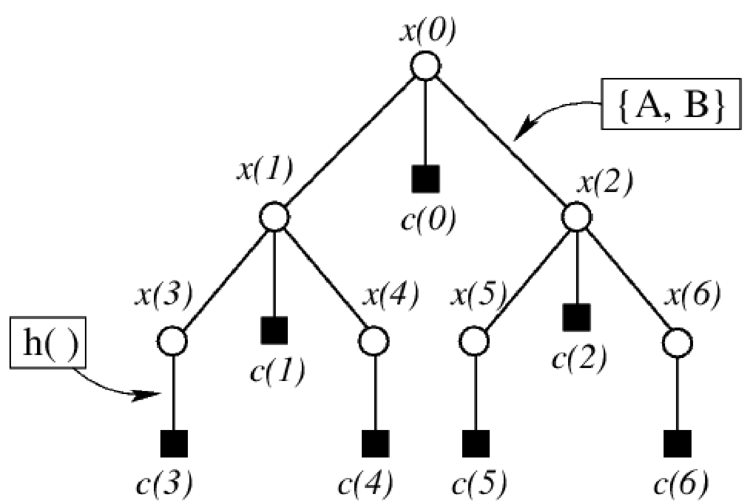
Proving the normalised joint distributions are Gaussian. Top: conditionals. Bottom: normalised joints.



$\Delta H/H = 0.0643$ 0.0743 0.0572 0.0836

Joint distributions

- Nearby filter coefficients are highly dependent, but this dependency decreases as (say) the spatial separation increases.
- Idea: specify smoothness of the mixer variables via a tree (equivalent to a GP)
- This then specifies the full forward model for image generation.



Multiscale Autoregressive Process (MAR)

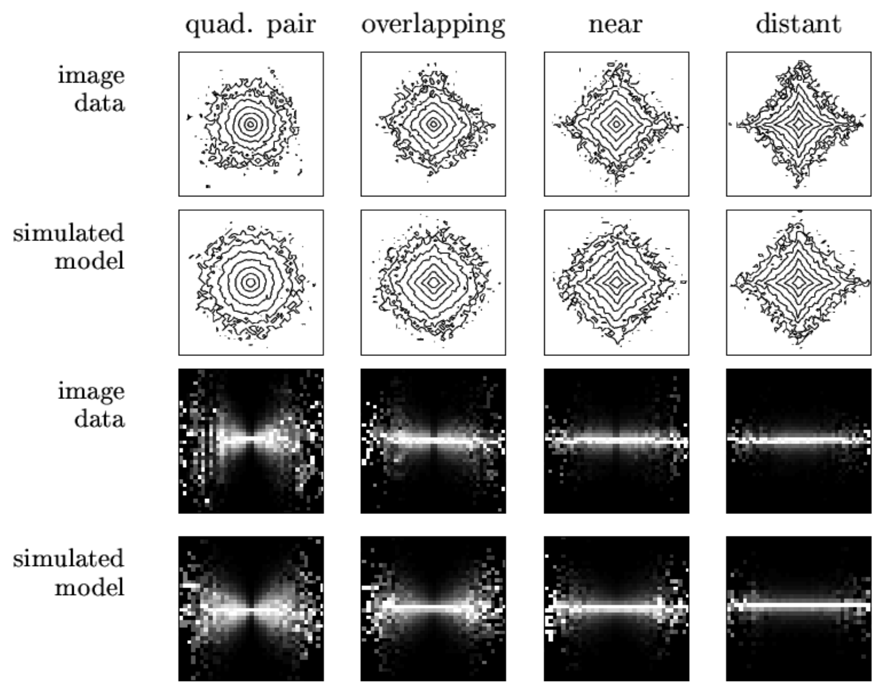
$$
\begin{aligned}
y(s) &= ||x(s)||u(s) && (2)\\
x(s) &\sim Norm && (3)\\
x(s) &= \mu x(p(s)) + \sqrt{1-\mu^2}w(s) && (4)\\
y(s) &= ||\mu^{d(s,s\wedge t)}x(s\wedge t) + v_1||u(s) && (5)\\
y(t) &= ||\mu^{d(t,s\wedge t)}x(s\wedge t) + v_1||u(t) && (6)
\end{aligned}
$$

$$
d(s, s\wedge t) \propto log_2(\Delta(s,t)) + 1 \qquad (7)
$$

Dependencies are put in by hand (cf. Schwartz, et al)

Generation from the model

Judiciously set the parameters, draw from the forward model, and compare to real images. a) a quadrature pair b-d) Δ=[4,8,128] (would have been nice to sample from the forward model, cf Simoncelli, 2005)



Summary

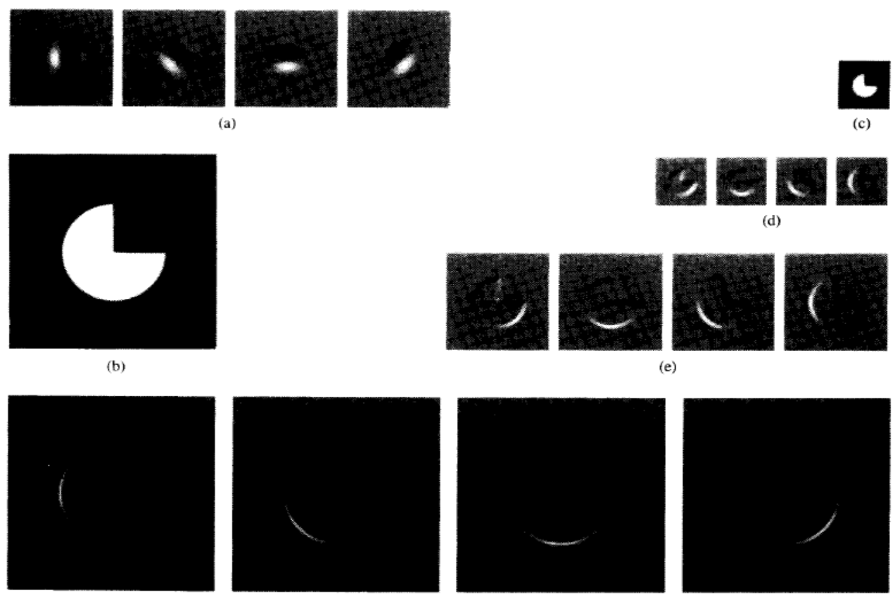
1. Shown **Gaussian Scale Mixtures** can characterise the pair-wise and marginal dependencies of wavelet filter coefficients.

2. Derived a non-linear form of normalisation.

3. Introduced a parameterisation (MAP) to capture the dependency between nearby filters that falls off with increased separation.

More on the steerable pyramid

Rotation and translation invariant. Steerable in orientation and multi-scale. Non-orthogonal, invertable and over-complete. a) filters at one scale. b) test image c) residual low pass d-f) coefficients at 3 scales. Residual high pass not shown.



Other useful references:

"Radom cascades on wavelet trees and their use in analysing and modeling natural images", Wainwright and Simoncelli, Journal of Applied Computational and Harmonic Analysis, 2001

"Statistical Modeling of Photographic Images", Eero Simoncelli, to appear in Handbook of video and image processing, 2005

"Shiftable Multiscale Transforms", Simoncelli, Freeman, Adelson, Heeger, Ahrens, IEEE transactions on information theory vol 38 no 2 March 1992

"Assignment of Multiplicative Mixtures in Natural Images", Schwartz, Sejnowski, Dayan. NIPS 2004?