Detecting Salient Contours Using Orientation Energy Distribution

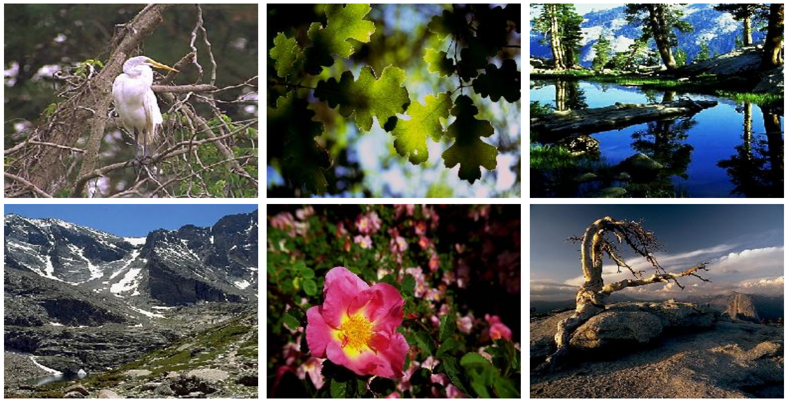
CPSC 636 Slide12, Spring 2015

Yoonsuck Choe

Co-work with S. Sarma and H.-C. Lee

Based on Lee and Choe (2003); Sarma (2003); Sarma and Choe (2006)

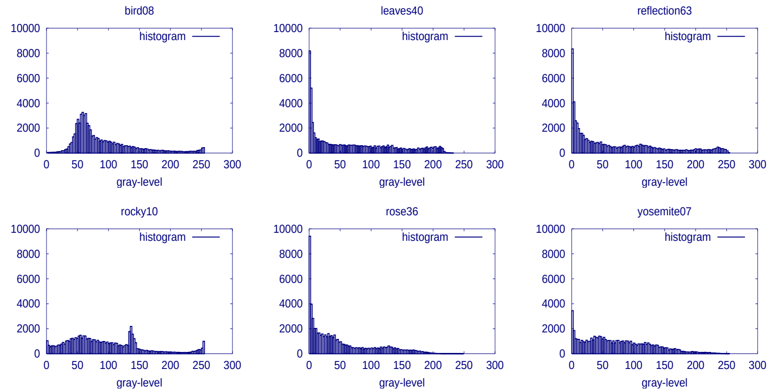
What Is Common in These Images?



What Is Common in These Images?

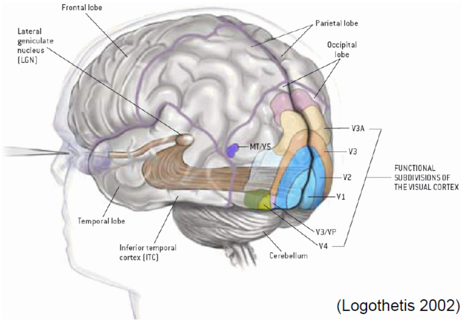
- In color, natural image, from the Kodak data set, ...
  - What about the brightness intensity histogram?

Brightness Intensity Histogram



- They are very different!
- What is similar then?

The Visual Cortical Response

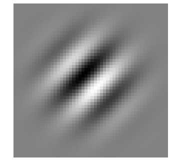


(Logothetis 2002)



- Retina: center-surround filter
- LGN (thalamus): center-surround filter
- Visual cortex: oriented Gabor filter

The Problem: How Does the Visual System Detect Salient Contours?



- Neurons in the visual cortex have Gabor-like receptive fields.
- Looking at the **response properties** of these neurons can help us answer the question.
- The simplest statistical property can be measured by looking at the **response histogram**.

Questioning from a slightly different perspective, “**how can the particular response property of visual cortical neurons be utilized by later processing?**”

Part I: Thresholding Based on Response Distribution

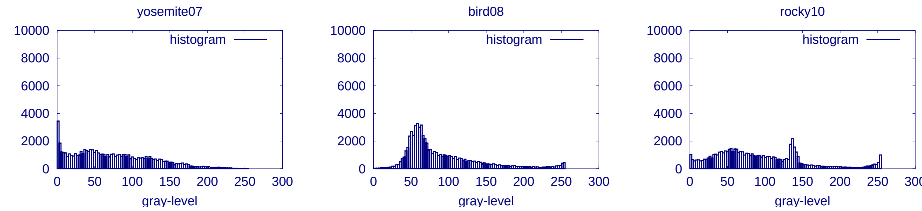
Observation

- Grayscale intensity distributions are quite **different** across different images.
- However, Gabor response distributions are quite **similar** across different images.

A Typical Grayscale Image

- Although not evident from the above, the intensity histogram can be widely different across different images.

Grayscale Intensity Distribution

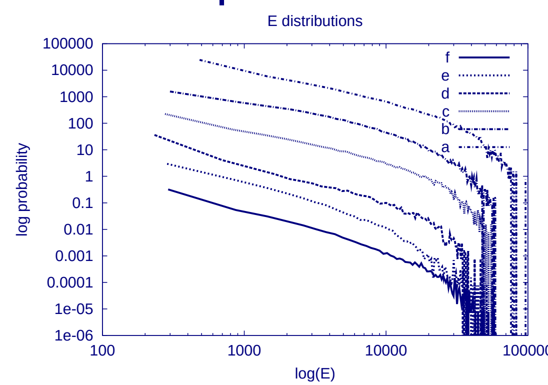


- Grayscale intensity histograms are drastically different across different images.
- Thus, a general algorithm for utilizing the intensity distribution cannot be easily derived.

A Typical Gabor Response (Orientation Energy)

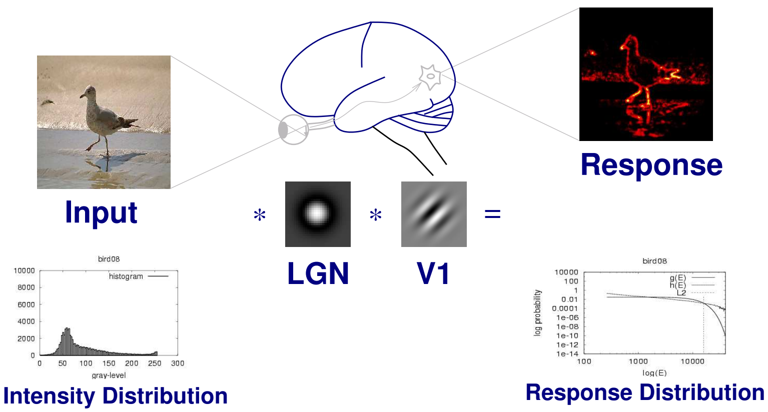
- High values near contours or edges.
- The energy distribution is strikingly uniform across images.

Gabor Response Distribution



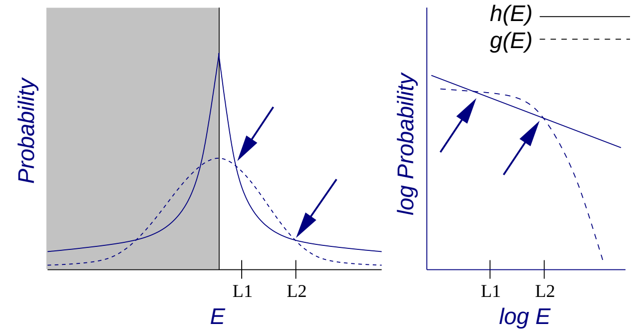
- The Gabor response (or **orientation energy**; E) distributions on the other hand are quite similar across differen images (shown in Log-Log plot).
- The distribution shows a power law property ($f(x) = 1/x^a$): sharp peak and heavy tail.

Summary So Far



- Input and response disrtibutions show quite different statistical properties.

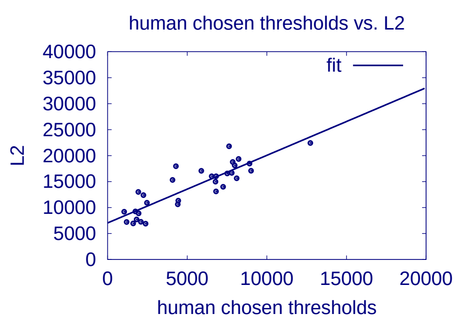
What to Make of the Power Law?



(a) linear scale (b) log−log scale

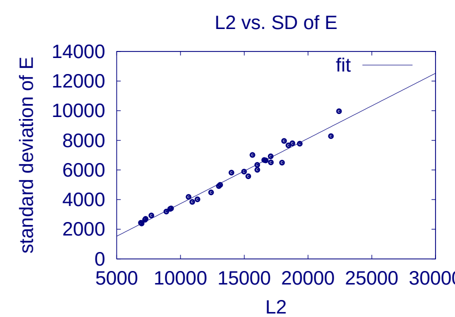
- Comparing the power law distribution with a normal distribution with the **same variance** can provide us with some information.
- **Assumption**: normal distribution can be a suitable standard.
- The point $L2$ where $h(E)$ becomes greater than $g(E)$ may be important, i.e., orientation energy is **suspiciously high**.

Exploiting the Power Law in E



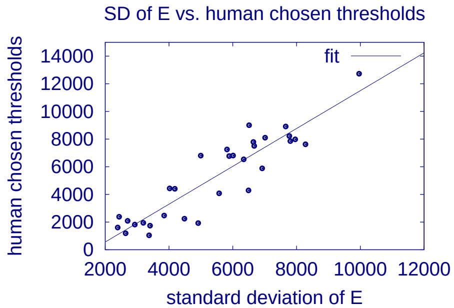
- High orientation energy E indicate a strong edge component in images.
- Can there be a relationship between the threshold of E above which humans see it as **salient** and the point $L2$?
- Clearly, there is a **linear relationship** between the two!

Further Discoveries: $L2$ and σ



- Further, the raw standard deviation σ of the orientation energy distribution is **linearly related** to $L2$.
- Question: Is there an analytical solution to $1/x^a = b \times exp(-x^2/c)$, where the constants a, b, and c depend on σ?

Using σ to Estimate Optimal E Threshold

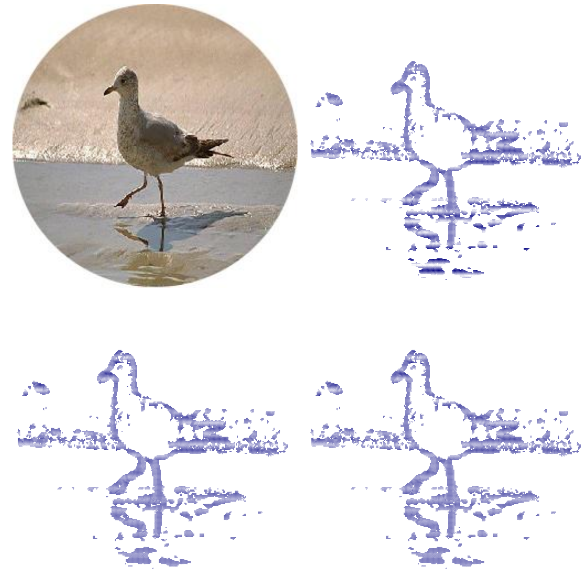


- Relating σ back to the human-chosen E threshold gives again a **linear relation**:

$$T_\sigma = 1.37\sigma - 2176.59.$$

- Thus, instead of calculating the histogram, etc., we can simply calculate the raw standard deviation σ to estimate the appropriate E threshold.

Application: Thresholding E



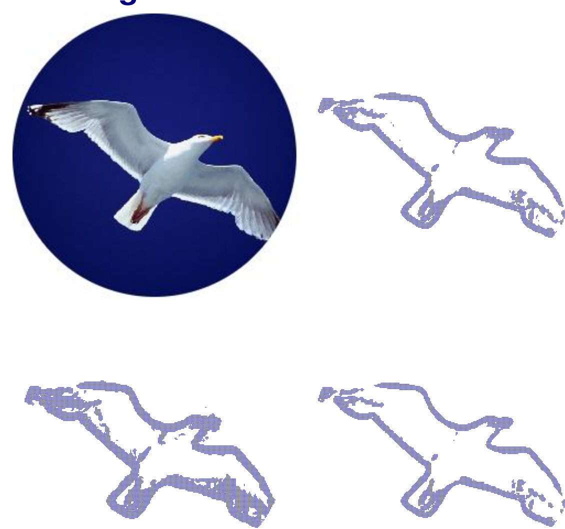
- Original, human-selected, 85-percentile, and T_σ.

Extraction of Salient Edges

(a) Original Image (b) Thresholded Edges (c) Magnified (b)

- Using T_σ as a threshold gives good results, comparable to humans' preference.

Thresholding E: Limitations of Fixed Percentile



- Original, human-selected, 85-percentile, and T_σ.

Thresholding E: Limitations of Global Thresholding

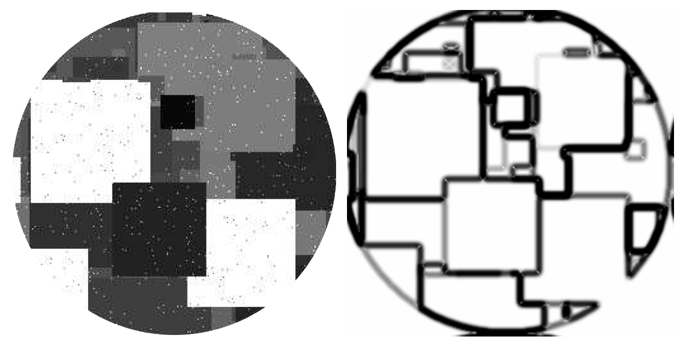
- Original, human-selected, 85-percentile, T_σ, and T_σ local.

Summary of Thresholding Results

- Fixed percentile thresholding does not give consistent results.
- The σ-based T_σ threshold works well.
- However, globally applying the same threshold has limitations.
- This problem can be overcome by applying the **same principle** derived here to calculate the **local thresholds**.
- The proposed method is an **efficient** way of detecting salient contours.

Part II: Quantitative Comparison

Approach

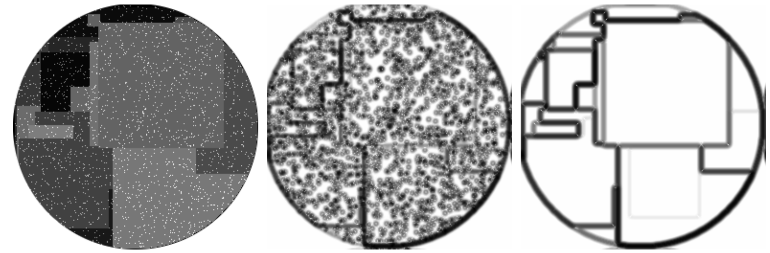


- Generate synthetic image (left) with known salient edges (right) and compare the thresholded version to this ground truth.
- Add noise and vary number of objects to make it interesting.

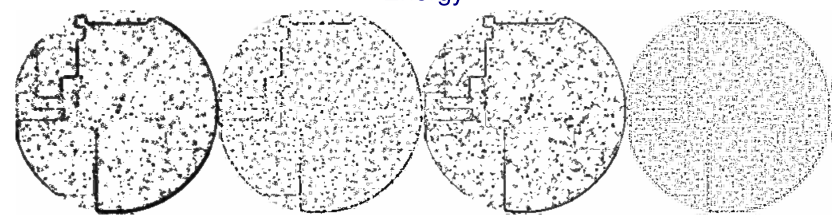
Compared Thresholding Methods

- Global OED: threshold derived from OED from the entire response matrix, using the normal distribution baseline.
- Local OED: threshold derived from OED from the local surrounding area of a pixel in the response matrix. using the normal distribution baseline.
- Global 85%: threshold derived from OED from the entire response matrix, using the 85-percentile point as the threshold.
- Local 85%: threshold derived from OED from the local surrounding area of a pixel in the response matrix, using the 85-percentile point as the threshold.

Experiment I: Vary Noise Level

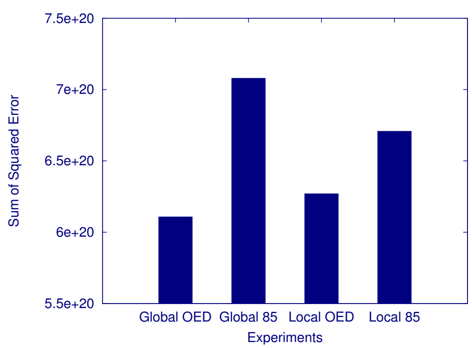


(a) Input (b) Orientation Energy (c) Reference



(d) Global OED (e) Global 85% (f) Local OED (g) Local 85%

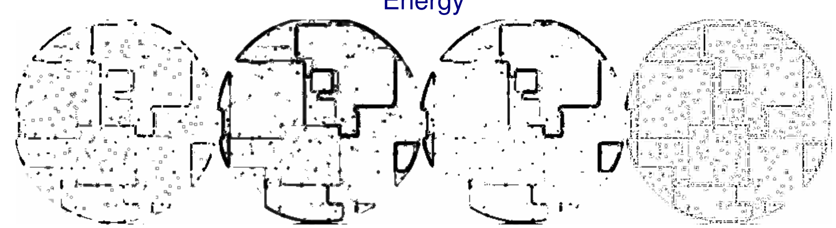
Experiment I: Vary Noise Level



	Global OED	Global 85%	Local OED	Local 85%
Global OED	X	$<$ (p=0.018)	$<$ (p=0.35)	$<$ (p=0.021)
Global 85%	X	X	$>$ (p=0.019)	$>$ (p=0.014)
Local OED	X	X	X	$<$ (p=0.025)
Local 85%	X	X	X	X

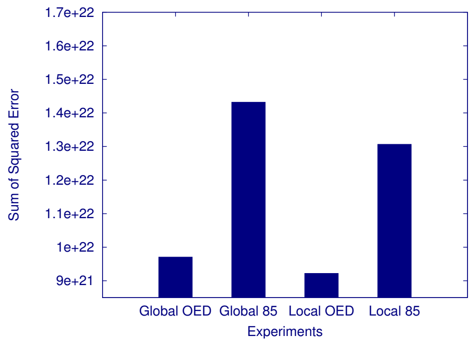
Experiment II: Vary Input Count

(a) Input (b) Orientation Energy (c) Reference



(d) Global OED (e) Global 85% (f) Local OED (g) Local 85%

Experiment II: Vary Input Count



	Global OED	Global 85%	Local OED	Local 85%
Global OED	X	$<$ (p=0.0107)	$>$ (p=0.258)	$<$ (p=0.014)
Global 85%	X	X	$>$ (p=0.011)	$>$ (p=0.006)
Local OED	X	X	X	$<$ (p=0.015)
Local 85%	X	X	X	X

Summary

- Thresholding based on orientation energy distribution (OED) did consistently better than fixed-percentile methods in noisy, synthetic images.
- Differences between glocal and local OED thresholding were not significant.

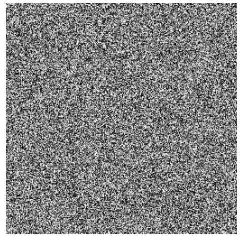
Part III: Analysis

Relationship Between T_σ Thresholding and Suspicious Coincidence

- What is the relationship between **salience** defined as **super-Gaussian** and the conventional definition of **suspiciousness** (Barlow 1994, 1989)?

$$P(A, B) > P(A)P(B).$$

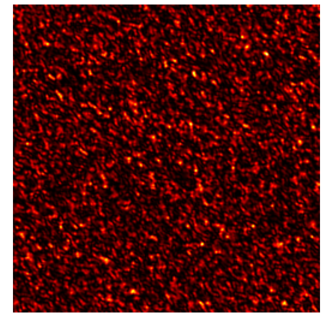
White-Noise Analysis

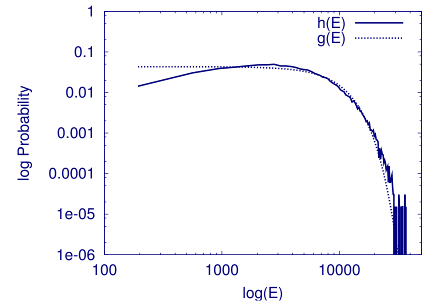


- If the Gaussian baseline assumption was correct, the E response distribution to white noise images should not be perceived as salient compared to a Gaussian with the same variance.
- In white-noise images, each pixel is independent, so, given pixel A and pixel B:

$$P(A, B) = P(A)P(B).$$

Gabor Response to White Noise Images



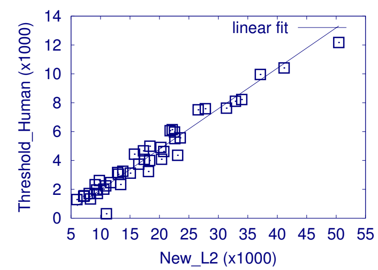


- The orientation energy distribution is very close to a Gaussian, especially near the high E values.
- Thus, the T_σ thresholding will result in no salient contours in white noise images.

Use of White Noise Response as a Baseline

- Can we use the white-noise response as a baseline for thresholding E?: Yes!
- Generate white noise response, and scale it by σ_h/σ_r where σ_h and σ_r are the STD in the natural image response and the white noise response.
- Recalculate the response distribution (if necessary).

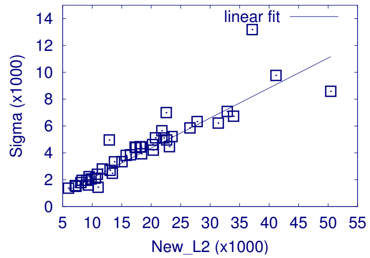
New Baseline for Salience vs. Humans



New L_2 vs. Human Chosen Threshold ($r = 0.98$)*

- Strong linearity is found between the new L_2 and the human selected threshold.
- * This is much tighter than the Gaussian baseline ($r = 0.91$)!

New Baseline for Salience vs. σ



New L_2 vs. σ ($r = 0.91$)

- The same linearity between L_2 and the σ is maintained.

Related Work

- Malik et al. (Malik et al. 1999) used peak values of orientation energy to define boundaries of regions of coherent brightness and texture.
- The non-Gaussian nature of orientation energy (or wavelet response) histograms has also been recognized and utilized for some time now, especially in denoising and compression (Simoncelli and Adelson 1996).
- Other kinds of histograms, e.g., spectral histogram by Liu and Wang (2002), or spatial frequency distributions (Field 1987), may be amenable to a similar analysis.

Discussion

- The local (or even global) threshold calculation can be easily implemented in a neural network.

$$\sigma^2 = \sum_{i,j} w_{ij} g(V_{ij}),$$

  where w_{ij} are connection weights serving as normalization constants, $g(x) = x^2$, and V_{ij} is the V1 response at location i, j.
- The resulting value can be passed through another activation function $f(x) = \sqrt{x}$.
- These are all plausible functions that can be implemented in a biological neural network.

Yet Another Power Law!

- Power law seems to be ubiquitous in nature and in human-made artifacts:
  - 957,000 documents returned by Google Scholar!
  - Power law phenomena range from www topology, financial market fluctuation, to word frequency and much more (see e.g., Clauset et al. 2009).
- However, it is not often asked:
  - What use is it?
  - What fundamental mechanisms underlie such phenomena?

Mathematical/Statistical Implications

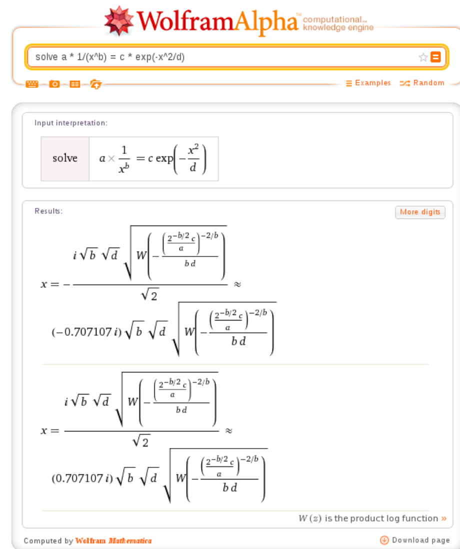
Is there an analytical solution to $a\frac{1}{x^b} = c \times \exp(-\frac{x^2}{d})$?

- This leads to another obscure yet surprisingly ubiquitous function called the Lambert W function $W(x)$ which is defined as the inverse of the following function:

$$x = W\exp(W)$$

- The Lambert W function is popping up everywhere: delay differential equations (with applications in population dynamics, economics, control theory), projectile trajectory calculation, voltage/current/resistance in a diode, etc. (see Hayes 2005 for a review)–A *déjà vu?*
- **Speculation**: Power law, Gaussian, and Lambert W function are deeply related.

Power Law, Gaussian, and Lambert W function



Basically, $x = \pm ip\sqrt{W(-q)}$

- How I found out: Wolfram Alpha (Mathematica, prior to that).

Summary

- Gaussian baseline was found to have a close relationship to the idea of suspicious coincidence by Barlow (1994)

Lesson Learned

- Studying statistical properties of raw natural signal distributions can be useful in determining why the visual system is structured in the current form (e.g., PCA, ICA, etc. predicts the **receptive field** shape).
- However, **what's more interesting** is that the response properties of cortical neurons can have certain **invariant properties** and this can be **exploited**.
- So, we need to **go beyond** finding out what receptive fields look like and why, and start to explore how cortical neuron response can be utilized by the rest of the brain.

Conclusion

- Cortical response distribution has a unique invariant property (the power-law).
- Such properties can be exploited in tasks such as salient contour detection.
- Gaussian distribution forms a good baseline for determining the threshold.
- The above may be related to the idea of suspicious coincidence.
- Power law, Gaussian baseline, and Lambert W function intricately interrelated.
- **Lesson:** Power law is there for a reason, and it can greatly simplify things downstream.

References

Barlow, H. (1994). What is the computational goal of the neocortex? In Koch, C., and Davis, J. L., editors, *Large Scale Neuronal Theories of the Brain*, 1–22. Cambridge, MA: MIT Press.

Barlow, H. B. (1989). Unsupervised learning. *Neural Computation*, 1:295–311.

Clauset, A., Shalizi, C. R., and Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Review*, 51:661–703.

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A*, 4:2379–2394.

Hayes, B. (2005). Why W?. *American Scientist*, 93:104–109.

Lee, H.-C., and Choe, Y. (2003). Detecting salient contours using orientation energy distribution. In *Proceedings of the International Joint Conference on Neural Networks*, 206–211. IEEE.

Liu, X., and Wang, D. (2002). A spectral histogram model for texton modeling and texture discrimination. *Vision Research*, 42:2617–2634.

Malik, J., Belongie, S., Shi, J., and Leung, T. K. (1999). Textons, contours and regions: Cue integration in image segmentation. In *ICCV(2)*, 918–925.

Sarma, S., and Choe, Y. (2006). Salience in orientation-filter response measured as suspicious coincidence in natural images. In Gil, Y., and Mooney, R., editors, *Proceedings of the 21st National Conference on Artificial Intelligence(AAAI 2006)*, 193–198.

Sarma, S. P. (2003). *Relationship between suspicious coincidence in natural images and contour-salience in oriented filter responses.* Master's thesis, Department of Computer Science, Texas A&M University.

Simoncelli, E. P., and Adelson, E. H. (1996). Noise removal via bayesian wavelet coring. In *Proceedings of IEEE International Conference on Image Processing*, vol. I, 379–382.