Detecting Salient Contours Using Orientation Energy Distribution

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?

- 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

- 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?”

Observation

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

Part I: Thresholding Based on Response Distribution
A Typical Grayscale Image

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

A Typical Gabor Response (Orientation Energy)

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

Gabor Response Distribution

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

- Input and response distributions show quite different statistical properties.

**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 $L_2$?
- Clearly, there is a linear relationship between the two!

**Further Discoveries: $L_2$ and $\sigma$**

- Further, the raw standard deviation $\sigma$ of the orientation energy distribution is linearly related to $L_2$.
- Question: Is there an analytical solution to $1/x^{\alpha} = b \times exp(-x^2/c)$, where the constants $a$, $b$, and $c$ depend on $\sigma$?

**What to Make of the Power Law?**

- 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 $L_2$ where $h(E)$ becomes greater than $g(E)$ may be important, i.e., orientation energy is suspiciously high.
Using $\sigma$ to Estimate Optimal $E$ Threshold

- Relating $\sigma$ 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 $\sigma$ to estimate the appropriate $E$ threshold.

Application: Thresholding $E$

- Original, human-selected, 85-percentile, and $T_\sigma$.

Extraction of Salient Edges

- Using $T_\sigma$ as a threshold gives good results, comparable to humans’ preference.

Thresholding $E$: Limitations of Fixed Percentile

- Original, human-selected, 85-percentile, and $T_\sigma$. 

(a) Original Image  
(b) Thresholded Edges  
(c) Magnified (b)
Thresholding: Limitations of Global Thresholding

- Original, human-selected, 85-percentile, $T_\sigma$, and $T_\sigma$ local.

Summary of Thresholding Results

- Fixed percentile thresholding does not give consistent results.
- The $\sigma$-based $T_\sigma$ 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

- 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

<table>
<thead>
<tr>
<th></th>
<th>Global OED</th>
<th>Global 85%</th>
<th>Local OED</th>
<th>Local 85%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global OED</td>
<td>X</td>
<td>&lt; (p=0.018)</td>
<td>&lt; (p=0.35)</td>
<td>&lt; (p=0.021)</td>
</tr>
<tr>
<td>Global 85%</td>
<td>X</td>
<td></td>
<td>&gt; (p=0.019)</td>
<td>&gt; (p=0.014)</td>
</tr>
<tr>
<td>Local OED</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>&lt; (p=0.025)</td>
</tr>
<tr>
<td>Local 85%</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

### Experiment II: Vary Input Count
Experiment II: Vary Input Count

<table>
<thead>
<tr>
<th></th>
<th>Global OED</th>
<th>Global 85%</th>
<th>Local OED</th>
<th>Local 85%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global OED</td>
<td>X</td>
<td>&lt; (p=0.0107)</td>
<td>&gt; (p=0.258)</td>
<td>&lt; (p=0.014)</td>
</tr>
<tr>
<td>Global 85%</td>
<td>X</td>
<td>X</td>
<td>&lt; (p=0.011)</td>
<td>&gt; (p=0.006)</td>
</tr>
<tr>
<td>Local OED</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>&lt; (p=0.015)</td>
</tr>
<tr>
<td>Local 85%</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

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.

Relationship Between $T_\sigma$ 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 $B$:

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

33

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 $\sigma_h/\sigma_r$ where $\sigma_h$ and $\sigma_r$ are the STD in the natural image response and the white noise response.

• Recalculate the response distribution (if necessary).

35

Gabor Response to White Noise Images

• The orientation energy distribution is very close to a Gaussian, especially near the high $E$ values.

• Thus, the $T_\sigma$ thresholding will result in no salient contours in white noise images.

34

New Baseline for Salience vs. Humans

• 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. $\sigma$

![Graph showing the linear fit between New L2 and $\sigma$](image)

New $L_2$ vs. $\sigma$ ($r = 0.91$)

- The same linearity between $L_2$ and the $\sigma$ 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.

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


