# Detecting Salient Contours Using Orientation Energy Distribution

Hyeon-Cheol Lee and Yoonsuck Choe*

Department of Computer Science
Texas A&M University
College Station, TX 77843-3112
Email: hclee@cs.tamu.edu, choe@tamu.edu

***Abstract*— How does our visual system detect prominent contours? Our investigation begins with the observation that neurons in the visual cortex have receptive fields similar to oriented Gabor filters. Unlike plain gray-level intensity histograms which greatly vary across images, we found that Gabor orientation-response (or orientation-energy) histograms of natural images have a fairly uniform shape. Based on this observation, we derived a threshold criterion which only depends on the standard deviation of the orientation-energy distribution. Thus, the same principle could be uniformly applied to different natural images, either locally or globally. Comparison with thresholds chosen by humans showed that the criterion can accurately predict human performance. Further, the proposed criterion can be easily implemented in a neural network, which is currently under investigation.**

## I. Introduction

We are interested in how the human visual system detects salient contours. Our first observation is that detection of such contours must be based on early visual processing in the human visual system. At the retinal ganglion cells and the lateral geniculate nucleus (LGN), the receptive fields have a center-surround property [3, 7], and further downstream in the primary visual cortex (V1), the receptive fields exhibit an orientation-, phase-, and frequency-tuned properties [6, 9].

These early visual processes can be approximated by a sequence of convolutions with difference-of-Gaussian (DoG) and oriented Gabor filters (for instance, see [5]). The resulting response level (called *orientation energy*) can be interpreted as visual cortical activity in response to visual input. Because of the oriented nature of Gabor filters, the filtered response is usually high around locations in an image where there is a strong contour edge component.

Since we are interested in the *salience* of contours, one reasonable way of detecting it is to find a *threshold* of the filter responses. We found that orientation energy histograms of different natural images have a fairly uniform shape, unlike that of gray-level intensity histograms which can be greatly different as shown in Fig. 1. This discovery allowed us to derive a single threshold criteria that can be uniformly applied.

In the following sections, we will describe the computational procedure for calculating orientation energy of images, and what are the properties of orientation energy histograms. We will then show how to estimate the thresholds based on these properties, and present results on a selected set of images. The paper will conclude with a discussion on how easily our algorithm can be implemented in a neural network, and on future directions.

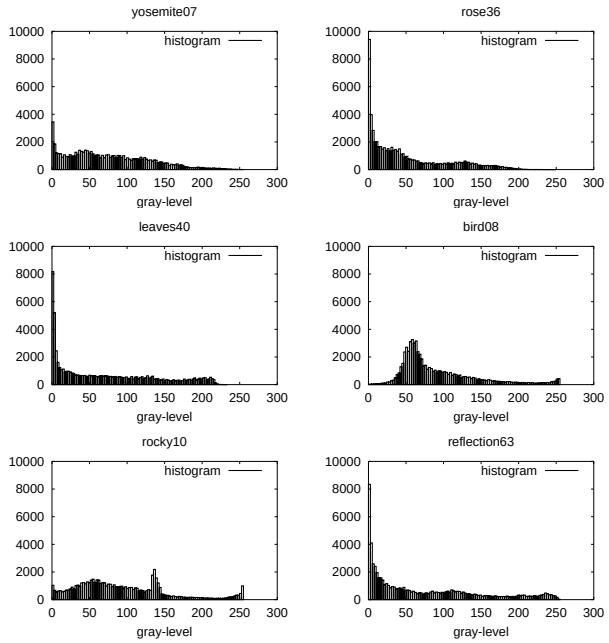



Fig. 1. **Gray-level intensity histogram of natural images.** Intensity histograms for 6 natural images are shown. The x-axis is the gray-level intensity (0 to 255), and the y-axis the frequency.

## II. Calculation of Orientation Energy

We follow a procedure similar to Geisler et al. [5] in calculating the orientation energy of natural images. First, we convolve the gray-level intensity matrix $I$ of a natural image with a difference-of-Gaussian (DoG) filter $D$ to obtain $I_D$ for initial edge detection:

$$I_D = I * D, \tag{1}$$

where $*$ is the convolution operator. The DoG filter $D$ is defined as a difference of two Gaussian functions $N_\sigma(x, y)$ with variance $\sigma^2$:

$$D(x, y) = N_{\sigma/2}(x, y) - N_\sigma(x, y), \tag{2}$$

$$N_\sigma(x, y) = \frac{1}{2\pi\sigma^2} \cdot exp(-\frac{x^2 + y^2}{2\sigma^2}), \tag{3}$$

where $(x, y)$ is the pixel location.

*Corresponding author: Yoonsuck Choe.

*Hyeon-Cheol Lee and Yoonsuck Choe. Detecting salient contours using orientation energy distribution. In Proceedings of the International Joint Conference on Neural Networks, pages 206-211. IEEE, 2003.*

The edge-detected image $I_D$ was then convolved with even- and odd-phased oriented Gabor functions $G_{\theta,\phi,\sigma}(x,y)$ of orientation $\theta$, phase $\phi$, and width $\sigma$ to obtain the orientation energy matrix $E_\theta$ as follows:

$$E_\theta = (G_{\theta,0,\sigma} * I_D)^2 + \left(G_{\theta,\frac{\pi}{2},\sigma} * I_D\right)^2 \quad (4)$$

$$G_{\theta,\phi,\sigma}(x,y) = exp(-\frac{x'^2+y'^2}{2\sigma^2})cos(2\pi x'+\phi) \quad (5)$$

$$x' = xcos\theta + ysin\theta$$

$$y' = -xsin\theta + ycos\theta,$$

where $(x,y)$ is the pixel location.[1]

To obtain the combined orientation energy $E(x,y)$ for each pixel $(x,y)$, we calculated the vector sum of 6 different pairs of $(\theta, E_\theta(x,y))$ at an interval of $\pi/6$. The resulting vector $(\theta', E(x,y))$ determined the optimal orientation $\theta'$ and the orientation response $E(x,y)$ at pixel $(x,y)$. In the following, for simplicity, we will refer to the orientation energy at any point $(x,y)$ as simply $E$, instead of $E(x,y)$.

The properties of the $E$ responses to natural images will be analysed next.

## III. Orientation Energy Distribution

To better understand the properties of orientation energy ($E$) in natural images, we calculated the $E$ responses from 31 natural images obtained from the same source as in [5]. Since we are interested in salient contours, first we calculated the histogram of the orientation energy matrices.

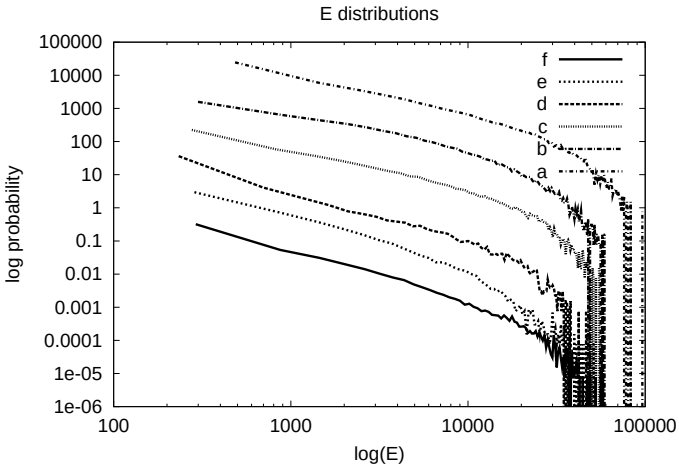



Fig. 2. **Orientation Energy ($E$) Distributions.** The $E$-distributions derived from the $E$-histograms of 6 natural images are shown in log-log plot (from $a$ to $f$). The same set of images from Fig. 1 was used. The curves are scaled by a factor of 10 successively to make the comparison easier. All curves are mostly straight with a similar slope, demonstrating a power law. Note that high energy area (toward the right) has a lot of noise due to the scarcity of samples.

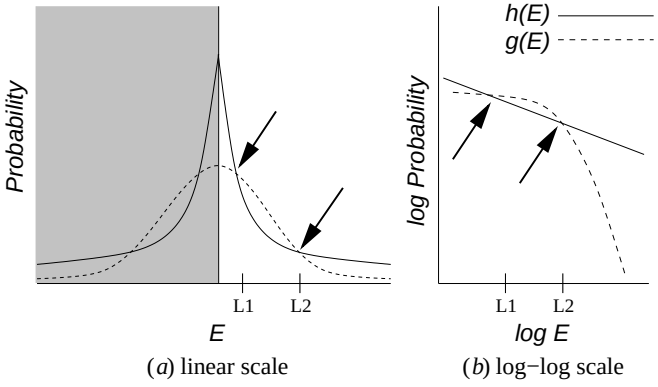
[1]There are other parameters associated with Gabor filters such as spatial frequency and aspect ratio, which is not made explicit in the above: All our Gabor filters were generated with spatial frequency 1 and aspect ratio 1. All of our convolution kernels were $7 \times 7$ in size, and the $\sigma$ parameter was adjusted accordingly.

From the histograms we obtained an orientation energy distribution by normalization[2]:

$$h(E) = \frac{f(E)}{\sum_{E \in B_E} f(E)}, \quad (6)$$

where $f(E)$ is the frequency at the histogram bin $E$, $B_E$ is the set of $E$ values of the histogram bins, and $h(E)$ is the resulting probability mass function.

Fig. 2 shows the orientation energy distribution for 6 sample images. We can see that the orientation energy distributions are similar to each other, and they share a unique feature – that of a power law (i.e., $p(x) = 1/x^a$, where $a$ is the fractal exponent). Thus when plotted in log-log scale, they show up as straight lines. One property of distributions following a power law is that extreme values are *not uncommon*, i.e. the distributions have a *heavy tail*. For example, when compared to a normal distribution with the same variance, power law distributions have higher probability near extreme values (see Fig. 3; especially, $E$ values greater than $L2$). Such a comparison can give us some clue on how to detect salient levels of orientation energy. The next section will further develop this idea.




(*a*) linear scale (*b*) log−log scale

Fig. 3. **Power Law distribution vs. Normal distribution.** A distribution following a power law (solid curve; $h(E)$) and a normal distribution (dashed curve; $g(E)$) having the same variance are shown. ($a$) On the left is the plot in linear scale, and ($b$) on the right the same plot in log-log scale (plotting for only positive $E$ values). The x-axis is the orientation energy $E$ and the y-axis is the probability. For the current work, we only consider positive $E$ values, and ignore the shaded part. There are two points where the two curves $h(E)$ and $g(E)$ intersect: $L1$ and $L2$. For $E$ values less than $L1$ or greater than $L2$, $h(E)$ is greater than $g(E)$.

## IV. Orientation Energy Distribution vs. Half-Normal Distribution

Comparing the $E$-distribution with a normal distribution can provide us with some insight into which $E$ values can be seen as unusually high (i.e. *salient*). Since $E$ values are only positive, we compared the $E$-distribution of natural images to *discretized half-normal* distributions (the unshaded part in Fig. 3). To make the two distributions have the same width, we calculated the *raw second moment* $\sigma^2$ of the $E$-distribution

[2]Note that the energy distribution is a *discrete* probability distribution.

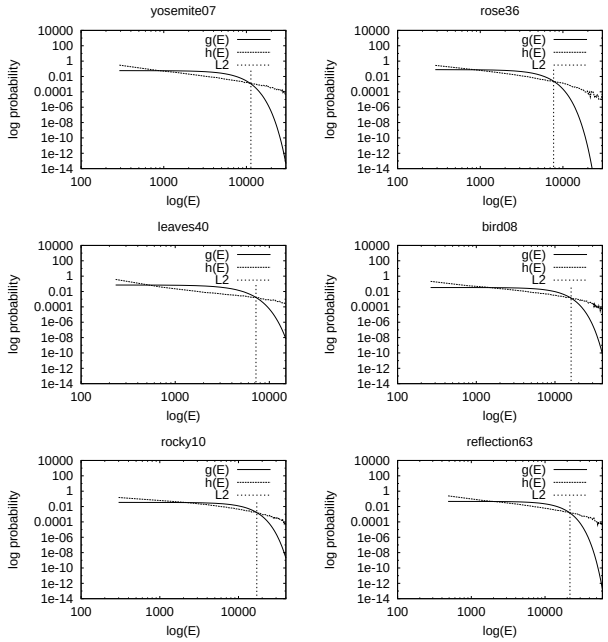



Fig. 4. **Comparison of Orientation Energy Distribution ($E$) and Normal Distribution with the Same Variance.** The $h(E)$ and $g(E)$ probability distributions (equations 6 and 8) of natural images are shown. The second intersection ($L2$) of the two curves are marked by a vertical line in each plot. We can see that beyond this point, $g(E)$ plummets, while $h(E)$ stays steady.

as follows:

$$\sigma^2 = \sum_{E \in B_E} E^2 h(E), \tag{7}$$

where $B_E$ is the set of $E$ values of the histogram bins, and $h(E)$ is the probability of the orientation energy level $E$ derived from the $E$-histogram.

With this $\sigma^2$, we calculated the continuous normal probability density function values $N(x; 0, \sigma^2)$ with mean 0 and variance $\sigma^2$ for all $E$ values, and then normalized them as follows:

$$g(E) = \frac{N(E; 0, \sigma^2)}{\sum_{E \in B_E} N(E; 0, \sigma^2)}, \text{ for } E \in B_E, \tag{8}$$

where $B_E$ is the set of $E$ values of the histogram bins, and $g(E)$ is the resulting discretized half-normal probability mass function of orientation energy level $E$. Fig. 3 illustrates the $E$-distribution $h(E)$ and the corresponding half-normal distribution $g(E)$.

Fig. 4 shows comparisons of $h(E)$ and $g(E)$ of 6 natural images in log-log scale. We can see that after the second intersection $L2$, $g(E)$ significantly drops in comparison to $h(E)$. To find out whether this point $L2$ has any significance in detecting salient contours, we compared the $L2$ values with orientation energy thresholds chosen by humans. The next section details the results of this comparison.

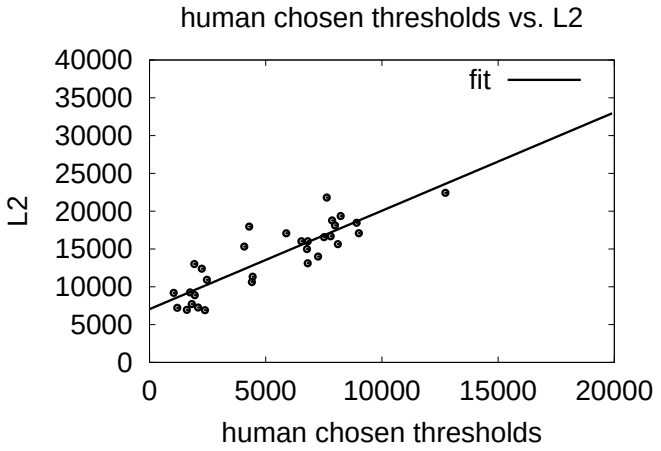



Fig. 5. **Orientation Energy Threshold Selected by Humans vs. $L2$.** The orientation energy thresholds selected manually are compared to the $L2$ values for each image. Each point in the plot corresponds to one of 31 natural images used in the calculation. The straight line shows a linear fit to the data.

## V. Orientation Energy Threshold Selected by Humans vs. $L2$

To test the significance of the $L2$ value where the orientation energy distribution $h(E)$ begins to diverge from the normal distribution $g(E)$, we compared the $L2$ values against orientation energy thresholds selected by humans. A total of 31 natural images were used to make the comparison. The $L2$ values were located computationally based on equations 6 and 8, and the orientation energy thresholds were selected manually by a single person (SB) in our research group. SB was shown thresholded $E$ visualized as shown in Fig. 8*b*. For each image from the set of 31, thresholded $E$ at 55% to 95% percentile of the $h(E)$ at an increment of 5% was shown to SB and he was asked to choose the best threshold according to the following criteria:

1) object contours should remain intact, and
2) noisy background edges must be removed as much as possible.

The results are shown in Fig. 6. The plot shows a clear linearity between $L2$ and the threshold selected by humans. Thus, the $L2$ value of orientation energy distributions can be used to predict the optimal threshold of $E$. However, computing $L2$ is a laborious process, because of the time and space spent in constructing $h(E)$ and $g(E)$, and also in locating $L2$. Interestingly, it turns out that the $L2$ value has a close linearity with the square root of the raw second moment ($= \sigma$)! Fig. 6 clearly demonstrates the strong linearity between the two values. As expected, the $\sigma$ of the orientation energy distribution $g(E)$ showed a clear linearity to the threshold selected by humans. Fig. 7 shows the results of this comparison.

We can now derive a linear equation of the preferred orientation energy threshold $T_\sigma$ as a function of $\sigma$:

$$T_\sigma = 1.37\sigma - 2176.59, \tag{9}$$

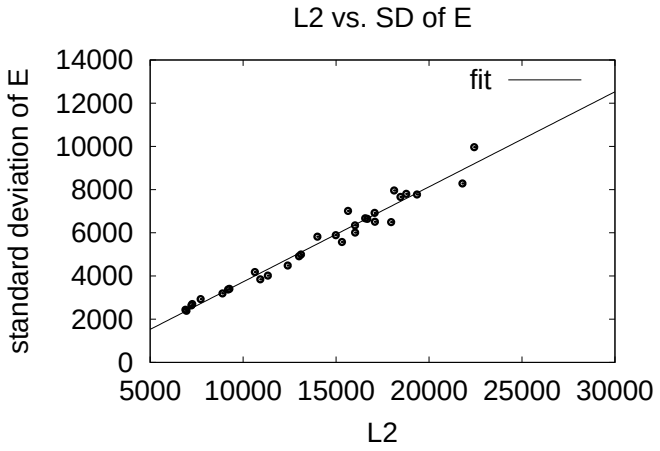



Fig. 6. $L2$ **vs.** $\sigma$ **of the Orientation Energy Distribution.** The $L2$ values derived from the comparison of $h(E)$ and $g(E)$ are compared with the square root of the raw second moment ($= \sigma$; or SD above) of $g(E)$. Each point represents one of 31 images. The straight line shows a linear fit to the data.

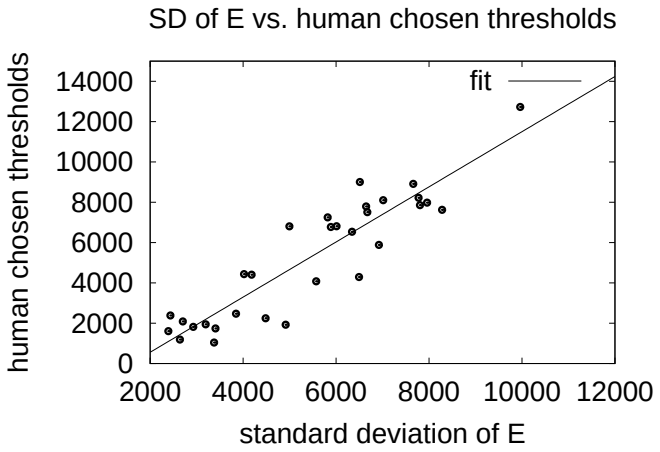



Fig. 7. **Orientation Energy Threshold Selected by Humans vs.** $\sigma$**.** The orientation energy thresholds manually selected are compared to square root of the raw second moment ($= \sigma$; or SD above) of $g(E)$. Each point represents one of 31 images. The straight line shows a linear fit to the data (see equation 9).

which is the linear fit shown in Fig. 7. This equation gives us a compact and easy way of determining orientation energy thresholds for detecting salient contours.

In summary, we found that comparing the orientation energy distribution $h(E)$ with a normal distribution $g(E)$ with the same variance can give us a good threshold criterion, and the value can be easily estimated by just calculating the raw second moment.

## VI. Thresholding Results

In this section, we present selected examples of applying the linear equation we derived above (equation 9). As a comparison, we took the 85-percentile of the orientation energy distribution $h(E)$, and contrasted the results with the

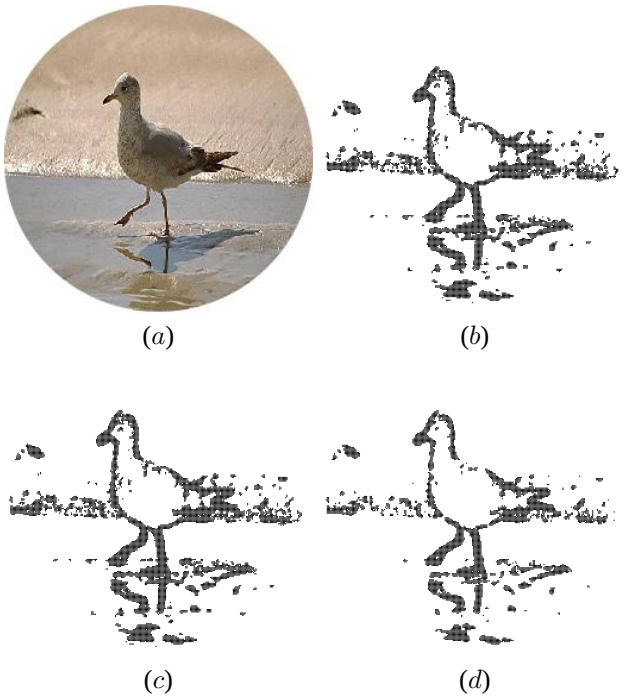



Fig. 8. **Global Thresholding Results (I).** Global thresholding results on the original image shown in ($a$) are presented. From ($b$) to ($d$) shows the thresholded orientation energy matrix where live pixels are replaced by oriented lines indicating the local orientation ($\theta'$ in Section II). ($b$) Threshold selected by a human. ($c$) Threshold at the 85-percentile of $h(E)$. ($d$) Threshold derived from the linear equation of $\sigma$ ($T_\sigma$, equation 9).

$T_\sigma$ thresholds. First, we will consider the case where a single threshold was found for the entire orientation energy matrix.

### A. Global Thresholding

Applying a single threshold value to the entire orientation energy matrix (i.e., *global thresholding*) can often be an effective strategy in detecting salient contours. For example, Fig. 8 shows global thresholding results. Three threshold values are compared in the figure: (1) threshold selected by a human, (2) 85-percentile of $h(E)$, and (3) $T_\sigma$ derived in equation 9. For this example, all three threshold values resulted in a similar outcome. However, such a case only arises when the orientation energy distribution is not too broad or not too narrow.

In cases where there are huge open space in the background, the fixed 85-percentile approach can give suboptimal results because of the concentration of energy around $E = 0$. This effectively pulls the 85-percentile point to a lower value than the desired threshold. Fig. 9 demonstrates this phenomenon. Due to the *under-thresholding*, we can see that in Fig. 9$c$, a lot of edges around the wing tips are not thresholded, leaving a mess. On the other hand, for our method ($T_\sigma$), because we take into consideration the spread of the distribution, we are able to effectively deal with such a condition.

### B. Local Thresholding

Even though global thresholding can be effective in many cases, it may still be inadequate in cases where there is a large

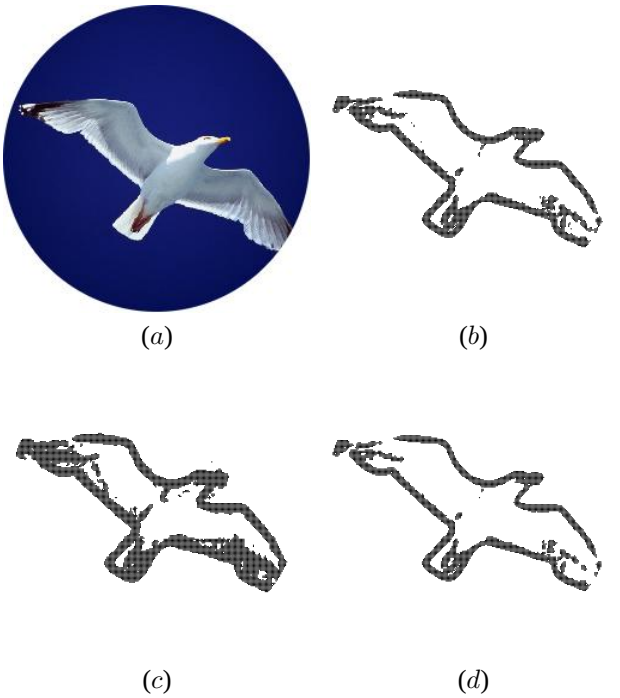


Fig. 9. **Global Thresholding Results (II).** Global thresholding results on an original image shown in ($a$) are presented. The details are the same as Fig. 8. ($b$) Threshold selected by a human. ($c$) Threshold at the 85-percentile of $h(E)$. ($d$) Threshold derived from the linear equation of $\sigma$ ($T_\sigma$, equation 9).

variation of $\sigma$ across different local patches in an orientation energy matrix. Fig. 10$a$ shows such an example. In this image, we can see that there are thin thread-like features and bold edges like the blade of the leaves. All global thresholding methods we have used so far fails in detecting both fine and bold contour features (Fig. 10$b$–$d$).

An advantage of our $\sigma$-based threshold $T_\sigma$ is that the value does not always need to be collected from the entire orientation energy matrix. Using the same principle, we can estimate locally optimal thresholds by using a sliding window. Fig. 10$e$ shows the result of *local thresholding* with a sliding window of size $21 \times 21$. The threshold at the center of the window was determined by the $\sigma$ value calculated from the $21 \times 21$ window. We can see that local thresholding method preserves both fine and bold contour features while suppressing background noise.[3]

## VII. Discussion

The proposed thresholding criteria only depends on the calculation of the raw second moment $\sigma^2$ of the orientation energy matrix. The $\sigma^2$ value can be directly calculated from an orientation matrix using a simple neural network. Suppose $V$ is the visual cortical (V1) response matrix, approximated by $E$ in this paper. Then an appropriate threshold on V1 response

[3]There is one issue with local thresholding. A careful look at Fig. 7 reveals that $\sigma$ values less than 2000 can result in a threshold value less than 0. This problem only occurs when the matrix is small, and for this special condition, we imposed a minimum threshold of 700.

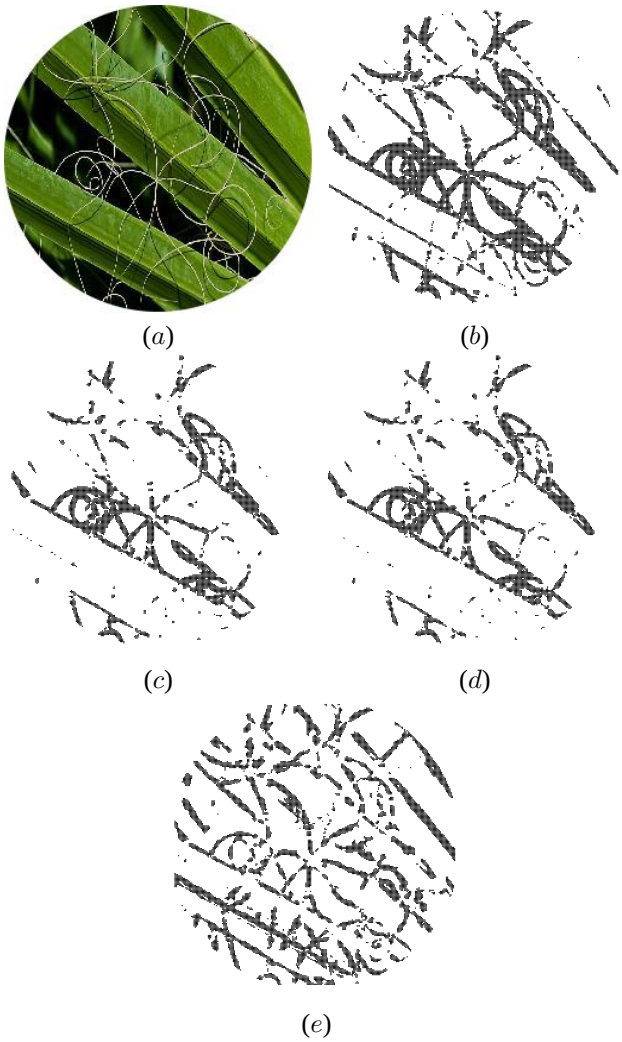


Fig. 10. **Global Thresholding vs. Local Thresholding.** Thresholding results on an original image shown in ($a$) are presented. The details are the same as Fig. 8. ($b$) Threshold selected by a human (globally applied). ($c$) Threshold at the 85-percentile of $h(E)$ (globally applied). ($d$) Threshold derived from the linear equation of $\sigma$ ($T_\sigma$, equation 9; globally applied). ($e$) Local thresholding based on $\sigma$ (sliding window size of $21 \times 21$).

can be easily calculated as follows:

$$\sigma^2 = \sum_{i,j} w_{ij} g(V_{ij}), \tag{10}$$

where $i, j$ are indices, $w_{ij}$ 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 an activation function of the form $f(x) = \sqrt{x}$ to finally obtain $\sigma$. Thus, a weighted sum of V1 response passed through a quadratic nonlinearity can then be passed through a square root activation function to easily obtain the desired value. It may be interesting to see if such a mechanism actually exists in the visual cortex or higher visual areas. One way of indirectly testing this speculation is to ask human subjects to do local thresholding similar to what our algorithm did in Section VI-B (see Fig. 10$e$). Instead of asking the subject to choose a single global threshold, we can show the subject all

local image patches from a single image and ask to choose a threshold for each patch. We expect the results to be similar to our results reported above.

Another future direction is to consider other types of response histograms. For example, spatial frequency power spectrum also shows a power law [4], so a similar approach can be employed in the spatial frequency domain. Also, various kinds of histograms called *spectral histograms* by Liu [8] can be considered. By analysing the response distribution in these different spectral histograms, we can gain further insight into how simple mechanisms can be used in the detection of salient features in the visual scene.

The way we define salient levels of $E$ by comparing the response distribution to a normal distribution of the same variance may be related to the concept of detecting *suspicious coincidence* [1, 2]. Barlow proposed that the perceptual system needs to effectively detect suspiciously coinciding events. For two events $A$ and $B$, if the joint probability $P(A, B)$ is greater than the product of the marginals $P(A)P(B)$, then the events $A$ and $B$ co-occurring are seen as suspicious. Of course this criterion is based on the comparison to the baseline case when $A$ and $B$ are independent, i.e., $P(A, B) = P(A)P(B)$. In our case, the baseline was a normal distribution, and it is tempting to speculate about a potential link between the two baselines. Now let us think of $A$ and $B$ as pixel intensities at two locations in the visual field. Then we can talk about the pair of values $(A, B)$ being suspicious or not. We can then look at the orientation energy value in the corresponding area, and see if salience by our measure matches suspiciousness by Barlow's measure. One thing we can try is to see what is the orientation energy distribution for white-noise images. In this case, $A$ and $B$ will surely be independent, and if our measure is to come close to Barlow's measure the orientation energy distribution should be similar to a normal distribution, i.e., we should not be able to detect any salient contour with our approach. We predict that (1) either this will be the case, or (2) at least the tail of the $E$-distribution of white-noise images (which contains the $L2$ point) will be close to a normal distribution.[4] If this turns out to be true, we should consider using simple thresholding of neural responses to detect suspicious coincidence instead of explicitly calculating the joint probabilities, which may be a formidable task.

## VIII. Conclusion

In this paper, we have shown that natural images have similar orientation energy distributions across different images. The distributions follow a power law, and comparing them to normal distributions with the same variance gave us a good thresholding criteria. This threshold value can be efficiently estimated by the square root of raw second moment of the orientation energy, and this same principle can be applied either globally or locally. Because of the simplicity of the calculation, the measure can be easily implemented in a neural network. It may be worthwhile checking if a similar mechanism underlies the human ability to detect salient contours.

[4]In fact, preliminary results suggest that the latter is indeed the case.

## Acknowledgments

This research was supported in part by Texas A&M University, and by the Texas Higher Education Coordinating Board ARP/ATP program under grant #000512-0217-2001 (to YC). We would like to thank J. C. Liu for the lively discussions, and S. K. Bhamidipati and S. P. Sarma for assistance with basic image processing routines and for the help with manual selection of thresholds. Also, the comments from the anonymous reviewers helped us clarify several points, for which we are grateful.

## References

[1] 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.

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

[3] Casagrande, V. A., and Norton, T. T. (1989). Lateral geniculate nucleus: A review of its physiology and function. In Leventhal, A. G., editor, *The Neural Basis of Visual Function*, vol. 4 of *Vision and Visual Dysfunction*, 41–84. Boca Raton, Florida: CRC Press.

[4] 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.

[5] Geisler, W. S., Perry, J. S., Super, B. J., and Gallogly, D. P. (2001). Edge Co-occurrence in natural images predicts contour grouping performance. *Vision Research*. 711–724.

[6] Jones, J. P., and Palmer, L. A. (1987). An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex. *Neurophysiol.*, 58(6):1233–1258.

[7] Kaplan, E. (1989). The receptive field structure of retinal ganglion cells in cat and monkey. In Leventhal, A. G., editor, *The Neural Basis of Visual Function*, vol. 4 of *Vision and Visual Dysfunction*, 10–40. Boca Raton, Florida: CRC Press.

[8] Liu, X. (2000). *Computational Investigation of Feature Extraction and Image Organization*. PhD thesis, Ohio State University.

[9] Marcelja, S. (1980). Mathematical description of the response of simple cortical cells. *Journal of Optical Society of America A*, 70(11):1297–1300.