Salience in Orientation-Filter Response Measured as Suspicious Coincidence in Natural Images

Subramonia Sarma

Amazon.com
1200 12th Avenue South, Suite 1200
Seattle, WA 98144-2734
sarma@amazon.com

Yoonsuck Choe

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

Abstract

Visual cortex neurons have receptive fields resembling oriented bandpass filters, and their response distributions on natural images are non-Gaussian. Inspired by this, we previously showed that comparing the response distribution to normal distribution with the same variance gives a good thresholding criterion for detecting salient levels of *edginess* in images. However, (1) the results were based on comparison with human data, thus, an objective, quantitative performance measure was not taken. Furthermore, (2) why a normal distribution would serve as a good baseline was not investigated in full. In this paper, we first conduct a quantitative analysis of the normal-distribution baseline, using artificial images that closely mimic the statistics of natural images. Since in these artificial images, we can control and obtain the exact saliency information, the performance of the thresholding algorithm can be measured objectively. We then interpret the issue of the normal distribution being an effective baseline for thresholding, under the general concept of *suspicious coincidence* proposed by Barlow. It turns out that salience defined our way can be understood as a deviation from the unsuspicious baseline. Our results show that the response distribution on white-noise images (where there is no structure, thus zero salience and nothing suspicious) has a near-Gaussian distribution. We then show that the response threshold directly calculated from the response distribution to white-noise images closely matches that of humans, providing further support for the analysis. In sum, our results and analysis show an intimate relationship among subjective perceptual measure of salience, objective measures of salience using normal distributions as a baseline, and the theory of suspicious coincidence.

Introduction

Edge detection algorithms have been widely used in the past with great success (Canny 1983). However, even after initial edge detection by these algorithms humans may have to determine which edge feature seem to be the most prominent. An alternative to standard edge-detection algorithms is the oriented Gabor filters (Daugman 1980). As it turns out, they are biologically grounded, i.e., the shape of the Gabor filters closely resemble experimentally measured receptive fields in the primary visual cortex (Jones & Palmer 1987).

An interesting property of such filters is that when applied to natural images, the response histogram shows a characteristic non-Gaussian shape with a sharp peak at zero. Thus, when compared to a normal distribution with the same variance, the response distribution has a heavy tail (Field 1987; Simoncelli & Adelson 1996). Such a response property has been found to be useful in tasks such as denoising (Simoncelli & Adelson 1996) and salient contour detection through thresholding. For instance, Lee & Choe (2003) showed that a simple thresholding criterion based on the comparison of the filter response distribution to a normal distribution of the same variance can accurately predict the salience level perceived by humans. Thus, these methods based on Gabor filters may help address the problem of "salience" in edge-detected images. Although the method by Lee & Choe (2003) was effective, (1) an objective quantitative measure of performance was not available, and (2) it was not shown why a normal distribution serves so well as a baseline in such a comparison.

In this paper, we first carry out a quantitative analysis of the normal-distribution baseline using artificially generated images that closely mimic the characteristics of natural images and which also yield to variations of the saliency information. We conduct different experiments to objectively measure the performance of fixed-percentile thresholding methods and others derived from the orientation energy distribution (OED) where the normal distribution is used as a baseline. Results show that the OED-derived thresholding methods overwhelmingly outperform the fixed-percentile thresholding methods.

Having objectively established the effectiveness of the normal distribution for thresholding, we then investigate why it would be so. We frame the problem in terms of the concept of *suspicious coincidence*, proposed by Barlow (1989): Two statistical events A and B are said to be a suspicious coincidence if they occur more often together than can be expected from their individual probabilities. In other words, the two events should be statistically non-independent in order for them to be deemed a suspicious coincidence: $P(A, B) > P(A)P(B)$. This approach is easily extended into the problem of image analysis, where suspiciousness can be determined by testing the inequality shown above where events correspond to pixels from different locations in the image (Barlow 1994).

Suspiciousness is directly linked to salience, i.e., more suspicious events may be seen as more salient to a perceptual system. According to the definition above, an image where each pixel is independent from each other would contain no suspicious coincidence between any pair of pixels. An example is a white-noise image, where we cannot see any noticeable salient feature. If we equate salience with suspiciousness, this also implies that algorithms such as Lee & Choe (2003) should fail to detect any salience in white-noise images. In other words, the distribution of orientation response to white-noise images should closely match a normal distribution.

We present results of comparing the oriented filter response histograms from white-noise images to their matching normal distributions that shows their close similarity and suggest that normal distributions can serve as a baseline for the detection of suspicious coincidence. The white-noise-derived response distributions are then used as a new baseline, and the results are shown to be consistent with results of psychophysical experiments showing human performance (Lee & Choe 2003).

Quantitative Analysis Using Synthetic Images

The synthetic image input for the quantitative analysis was created by generating random squares of various sizes where the gray-level was uniform within each square but different from adjacent squares. Uniform background noise was also embedded among these squares. A large number of such squares were generated and were subjected to a circular aperture to discount any artefactual orientation bias. Such synthetic images have the advantage that the error can be precisely measured, by comparison with the orientation energy matrix of the clean image without noise.

To find the orientation response (or energy) distribution, we followed the procedure described by Geisler *et al.* (2001). The method uses a sequence of convolutions: first the difference of Gaussian (DoG), and then the oriented Gabor filters to calculate the orientation filter response. The DoG filter uses the difference of two Gaussian functions whose widths differ by a factor of 0.5, as

$$D(x,y) = G_{(\sigma/2)^2}(x,y) - G_{\sigma^2}(x,y), \qquad (1)$$

where $G_{\sigma^2}(\cdot)$ is a Gaussian function with variance σ^2.

The gray level intensity matrix I of the input image is convolved with the DoG filter to obtain the resultant matrix I_d ($I_d = I * D$). We used a DoG filter of size 7×7 for all of our experiments. The filtered image is then convolved with oriented Gabor functions (Daugman 1980) of both even and odd phases with orientation θ, phase ϕ, and width σ to obtain the orientation energy matrix E_θ. The spatial frequency and aspect ratio parameters of the Gabor filters were set to 1 each, and the convolution kernels were sized 7×7 as usual. The orientation energy matrix for a single orientation θ is

$$
\begin{aligned}
E_\theta &= \left(\left(\exp^{-\frac{x'^2+y'^2}{2\sigma^2}} \cdot \cos(2\pi x') \right) * I_d \right)^2 \\
&+ \left(\left(\exp^{-\frac{x'^2+y'^2}{2\sigma^2}} \cdot \cos\left(2\pi x' + \frac{\phi}{2}\right) \right) * I_d \right)^2 \quad (2)
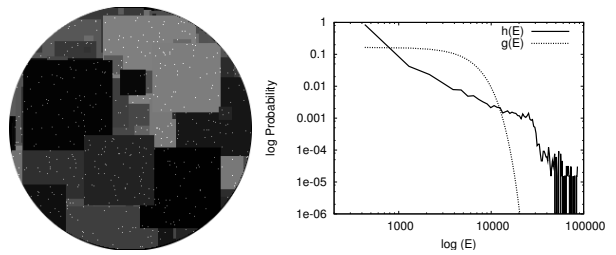\end{aligned}
$$

Figure 1: **Input Image and Response Distribution.** A synthetic image consisting of a set of overlapping squares with embedded noise and its corresponding orientation energy distribution ($h(E)$) is shown against the normal distribution of the same variance ($g(E)$). The synthetic image has a distribution similar to that of natural images (i.e., showing a power law: see Fig. 6).

where $x' = x\cos(\theta) + y\sin(\theta)$, $y' = -x\sin(\theta) + y\cos(\theta)$, and (x,y) is the pixel location as above.

For each location (x,y), we obtained the vector sum of six $(\theta, E_\theta(x,y))$ pairs in polar coordinates ($\theta = 0, \frac{\pi}{6}, \frac{2\pi}{6}, \frac{3\pi}{6}, \frac{4\pi}{6}, \frac{5\pi}{6}$) to find the combined orientation energy which gives the estimated orientation $\theta^*(x,y)$ and the associated orientation energy value $E^*(x,y)$ at that location. The orientation energy distribution is then estimated from the E^* responses using a histogram of bin size 100, followed by normalization, as h(E)$= \frac{f(E)}{\sum_{x\in B_h} f(x)}$ where $f(E)$ is the frequency of energy value E in the histogram, B_h is the set of histogram bin locations, and $h(E)$ is the resulting probability mass function which specifies the orientation energy distribution for the filtered image.

One way to detect salient levels of orientation energy is by comparing the orientation energy distribution for the input image with a normal distribution of the same variance as proposed in Lee & Choe (2003), so that unusually high levels of orientation energy show up as salient. We calculate the raw second moment of the E distribution (i.e., the expected value of E^2) for the input image as $\sigma_h^2 = \sum_{x\in B_h} x^2 h(x)$. We use this calculated σ_h^2 to find the matching continuous normal probability density function $\mathcal{N}(x; 0, \sigma_h^2)$ with mean 0, variance σ_h^2 for all $E \in B_h$ and normalize it to find the discretized normal probability mass function $g(E)$ of the orientation energy level E:

$$g(E) = \frac{\mathcal{N}(E; 0, \sigma_h^2)}{\sum_{x\in B_h} \mathcal{N}(x; 0, \sigma_h^2)}. \qquad (3)$$

From the above, we derived the saliency threshold based on the L2 value (see Fig. 6(a)), following Lee & Choe (2003).

Note that E is always greater than zero (equation 2), i.e., $\forall E \in B_h, E \geq 0$, thus the above can be seen as a half-normal distribution. In the following section we will measure the orientation energy E of natural and artificial images, and their distributions $h(E)$ and the matching normal distributions $g(E)$.

The OED for the synthetic images were similar to that of natural images, and had a characteristic high peak and a heavy tail. Fig. 1 shows a sample synthetic image and its OED distribution plotted against the normal distribution

in log-log scale. We can see that the OED of the synthetic image closely mimics that of natural images. In fact Lee, Mumford, & Huang (2001) showed that similarly generated synthetic images have statistical properties very similar to natural images.

A quantitative analysis using synthetic images can then be focused on suitable thresholding of the images to remove the background noise and detect salient features given by the edges of the squares. A good thresholding method should perform two objectives: (1) detect as many salient edges as possible, and (2) remove as much background noise as possible. In the following sections, we describe two sets of experiments for the quantitative analysis of the thresholding methods where, (1) the number of overlapping squares was kept constant but the background noise was varied; (2) and the background noise was kept constant while the number of squares (the input count) was varied. For each of the types, we generated five different image configurations for a more thorough analysis.

Variation in noise and performance

For this experiment, we generated synthetic images by keeping the input count (the number of overlapping square elements) constant and varied the background noise. The embedded noise was uniformly distributed. Three density levels of noise were used corresponding to 10%, 5% and 1% noise. The combined resultant images were then subjected to a circular aperture, and the performance of the fixed and OED-derived adaptive thresholding methods were measured.

To better investigate the relative merits of the global and local thresholding methods as described in Lee & Choe (2003), both the variations were used for the fixed percentile and OED-derived methods on all the images. For each noise density level, 5 representative synthetic images were generated with different configurations of the square patterns in them, with the total number of squares fixed to 300. The OEDs for each image input were then calculated. A fixed threshold of 85-percentile was used for the fixed percentile thresholding. For the local thresholding, a sliding window of size 21×21 was used.

Lee & Choe (2003) described four different thresholding methods, which we used in this paper: (1) the global fixed percentile thresholding, (2) the local fixed percentile thresholding, (3) the global OED-derived adaptive thresholding, and (4) the local OED-derived adaptive thresholding.

Thresholding results for a sample image with 300 overlapping squares and with 95-percentile and above embedded noise level are shown in Fig. 2. We can see that the global OED-derived thresholding method provides the best performance for this kind of input.

Such qualitative results are also backed by values for the quantitative measure of sum of squared error (SS Error). The SS Error for a thresholding result could be defined as the sum of the square of the difference in the orientation energy values of the thresholded image (containing background noise) from the un-thresholded image representing the ideal response with zero noise. The average SS Error values for a sample noise level are shown in the plot of

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

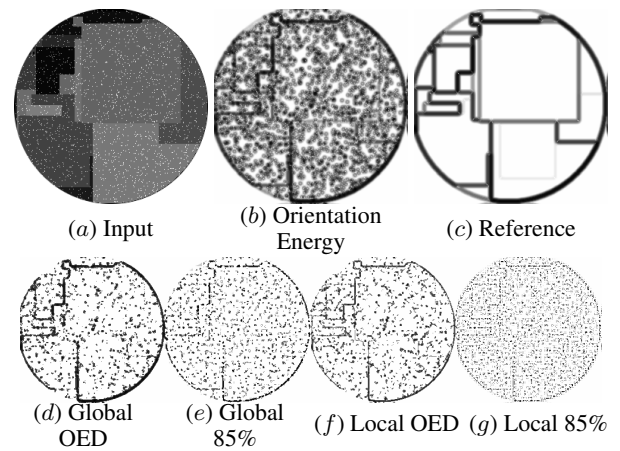(d) Global OED (e) Global 85% (f) Local OED (g) Local 85%

Figure 2: **Thresholding Results for Different Input Density.** (a) A synthetic image consisting of 300 overlapping squares with embedded noise. (b) The orientation energy matrix for (a). (c) The orientation energy matrix for the synthetic image without the noise that is used as the reference. Results of thresholding by the four different methods of (d) global OED, (e) global 85-percentile, (f) local OED, and (g) local 85-percentile respectively are shown from the left to the right. The global OED-derived adaptive thresholding method seems to offer the best performance for this input.

	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

Table 1: The p-values for the paired t-test of the difference in the mean sum of squared error values for the four thresholding methods, for a synthetic image with 300 squares and 5% noise. The '<' and '>' symbols indicate the relation between the mean SS Error values of the two thresholding methods.

Fig. 3. We can see that the sum of squared error is the lowest for the global OED derived method, followed closely by the local OED derived method. The SS Error values for the fixed 85-percentile methods were found to be significantly higher. The significance in this difference was evaluated using paired t-test. The results of the paired t-test are given in Table 1. The p-values of all the probable pairs of comparisons except the global OED vs. local OED comparison were found to be less than 0.025, which indicates that the mean SS Error values for the methods could indeed be significantly different. The p-value for the global OED vs. local OED comparison was found to be much higher, thus the difference was not significant.

The results for this experiment show that the OED-derived adaptive methods overwhelmingly outperformed the fixed percentile based methods in terms of detection efficiency of edges in the image input, and suppression of noise. However, the global OED-derived method offered a slight improvement in performance compared to the local method. This could probably be attributed to specific properties of the input images considered, such as the low density of the
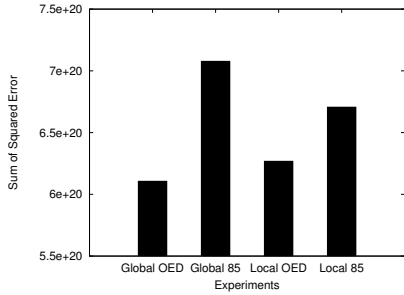
Figure 3: **Average Error in Thresholding (Experiment 1).** Bar plots show the average SS Error values for the thresholding results for a sample noise level for the synthetic images. The global and local OED-derived method have significantly smaller SS Error values than the fixed 85-percentile methods.

squares in the image.

The variation of noise seemed to have little effect on the relative performance of the thresholding methods. The global OED thresholding method offered the best performance closely followed by the local OED thresholding method. The global and local 85-percentile based thresholding methods performed differently for different inputs but were always weaker than the OED-derived methods.

Variation in input count and performance

The second experiment that was carried out was to keep the background noise constant while varying the number of input features. Three different image configurations were used for each, where the numbers of overlapping synthetic squares were 100, 300 and 500 respectively. The constant background noise level was kept as 1%. For each input configuration, 5 different image samples were generated as in the first experiment. A fixed threshold of 85-percentile was used for the fixed percentile thresholding methods, and a sliding window of size 21×21 was employed for the local thresholding method.

Thresholding results for a sample image with 500 overlapping squares and with 1% embedded noise level are shown in Fig. 4. The OED-derived thresholding methods again outperformed the fixed 85-percentile based methods for all the input configurations. Among the OED-derived methods, the global method offered better performance for smaller input count but fell behind the local method for larger input count. Thus the local thresholding result for the 500-square configuration was better than the corresponding global thresholding result. The average SS Error value was the lowest for the local OED-derived thresholding method, but quite high for the fixed-percentile methods. Fig. 5 shows this comparison. The paired t-test statistic for the average SS Errors for the different thresholding methods had p-values less than 0.015 for all the pairs of comparisons, except for the Global OED vs. Local OED comparison (Table 2).

Relationship to Suspicious Coincidence

In this section we compared the orientation energy E of natural and artificial images, and their distributions $h(E)$ and

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

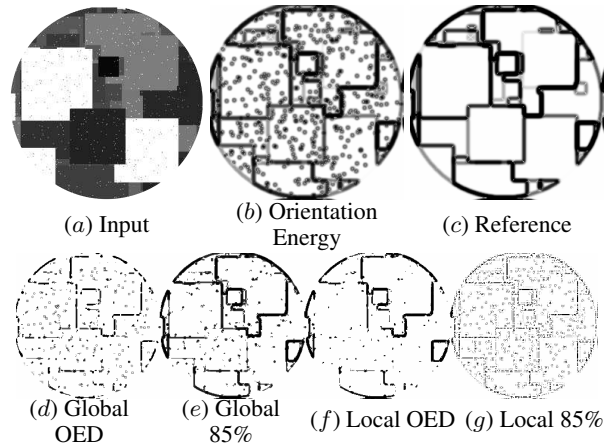(d) Global OED   (e) Global 85%   (f) Local OED   (g) Local 85%

Figure 4: **Thresholding Results with Varying Noise Level.** (a) A synthetic image consisting of 500 overlapping squares with embedded noise. (b-g) The results reflecting those in Fig. 2 are shown. The local OED-derived adaptive thresholding method seems to offer the best performance for this kind of input.
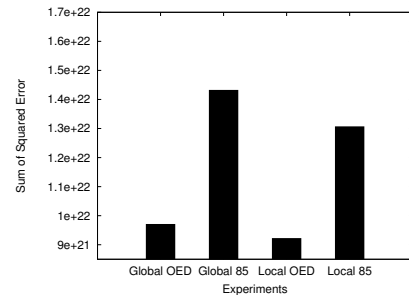

Figure 5: **Average Error in Thresholding (Experiment 2).** The thresholding results for a sample noise level for the synthetic images is shown. The same superiority of OED-derived methods is shown as in Fig. 3.

the matching normal distributions $g(E)$ to investigate their relationship to the concept of suspicious coincidence. For the experiment, we used the 42 natural images and the human performance data from Lee & Choe (2003). The orientation energy distribution of the images follows approximately a power law (i.e., $p(x) = 1/x^a$ where a is the fractal exponent), and as such, it has a heavy-tail where extreme values have higher probability of occurrence compared to a normal distribution of the same variance For example, Fig. 6(a) shows an illustration of power law distribution compared to a normal distribution having the same variance, and Fig. 6(b) shows the orientation energy distribution calculated from a natural image $h(E)$ compared to its matching normal distribution $g(E)$. The straight declining slope characteristic of a power law distribution is evident in $h(E)$. The two curves intersect at two points, near $E \sim 500$ (L1) and $E \sim 7,000$ (L2). Beyond L2, $g(E)$ plummets, but $h(E)$ remains high relative to $g(E)$.

Lee & Choe (2003) empirically derived the effective threshold for the detection of salient contours, which was linear to the orientation energy corresponding to the second

	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

Table 2: The p-values for the paired t-test of the difference in the mean sum of squared error values for the four thresholding methods, for a synthetic image with 500 squares and 1% noise. The '<' and '>' symbols indicate the relation between the mean SS Error values of the two thresholding methods.
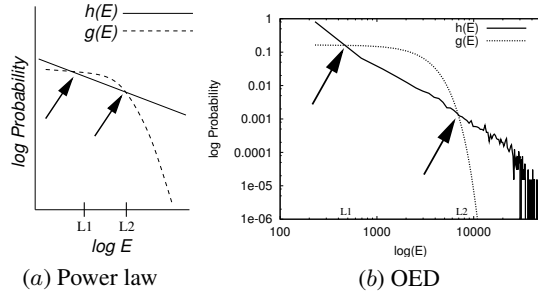

(a) Power law   (b) OED

Figure 6: **Orientation Energy Distribution of a Natural Image vs. its Matching Normal Distribution.**

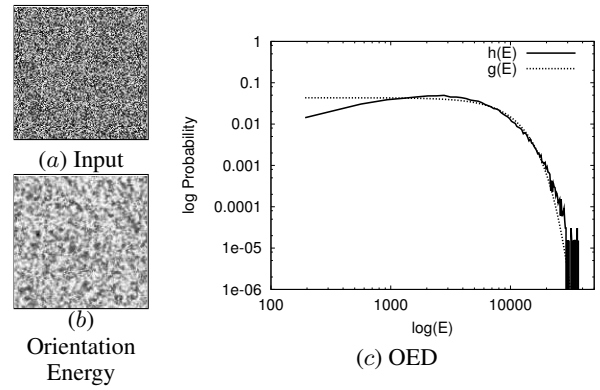(a) Input

(b) Orientation Energy

(c) OED

Figure 7: **Orientation Energy Distribution of a White-Noise Image vs. Its Matching Normal Distribution.** (a) A white-noise image shown in gray-scale. (b) The orientation energy E of the image in (a). There is no clear structure visible. (c) The log-log plot shows the orientation energy distribution $h(E)$ for the white-noise image against a normal distribution of the same variance $g(E)$. The two distributions show a close resemblance.

point of intersection (L2) of the response distribution and its matching normal distribution. However, it was not clear why the simple idea of comparing to a normal distribution has to be so effective. That is, why does a Gaussian distribution form a reasonable baseline for comparison? We observe that the detected salience in our method may correspond to a suspicious event in the image, i.e., a suspiciously high degree of edginess. If suspiciousness (as defined by Barlow (1989)) is indeed related to salience defined in our way, we can expect that a random image with completely independent statistical features (e.g., a uniformly randomly distributed white-noise image) may not show suspicious (i.e., salient) levels of orientation energy. (Note that we are not arguing that non-edge features in natural images have a white-noise statistic. Rather, our argument is that images containing white-noise statistic will not show any salience.) For this to happen in our method, the orientation energy distribution of white-noise images should not have a heavy tail, and in a more strict sense, it should coincide with its matching normal distribution. That is, it should be near-Gaussian.

To test if this is the case, we calculated the orientation energy distribution from a white-noise image and compared it with its matching normal distribution. The white-noise image was a 256×256 intensity matrix of uniformly randomly distributed values between 0 and 255 (Fig. 7(a)). The orientation energy distribution was then found using the procedure outlined in the previous section. We then compared the orientation energy distribution to the matching normal distribution of the same variance to see if there is any similarity between the two. It turns out that the two distributions closely overlap as expected (Fig. 7). Simoncelli & Adelson (1996) also point to a similar result, where they showed that wavelet response histograms from white-noise images

are near-Gaussian. However, they applied that finding in a different context than ours. These results suggests that normal distributions correspond to a baseline where all pixel values are independent (and thus no suspicious coincidence), and any deviation from this baseline can be seen as suspicious, or salient. Thus, salience as defined in our work can be understood as a deviation from the unsuspicious baseline.

From this result, we expect that the white-noise input based orientation energy distribution can also be used directly in finding the appropriate threshold. To test this, we conducted another experiment in which we generated new L2 values by comparing the orientation energy distribution with the white-noise based distribution. Since the standard deviation of a random variable scaled by the factor of c is $c \times \sigma$ where σ is the standard deviation before scaling, we multiplied the orientation energy matrix of the white-noise image with a constant σ_h/σ_r, where σ_h and σ_r are the standard deviations from a natural image and the white-noise image, respectively. Then the resulting orientation energy matrix has the same variance as the reference distribution calculated from a given natural image. The new L2 values were then found computationally by comparing the two distributions. These values were compared to the orientation energy thresholds selected by humans. The results are shown in Fig. 8(a). It is clear that the new white-noise based L2 values also have a strong linear relationship with the human-selected thresholds, even more so than the old Gaussian-based L2 values (correlation of $r = 0.98$ for the new L2, and $r = 0.91$ for the old one).

Discussion

The usefulness of identifying significant values of orientation energy has been studied previously, in applications such as denoising, texture perception, and image representation. For example, 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 orienta-
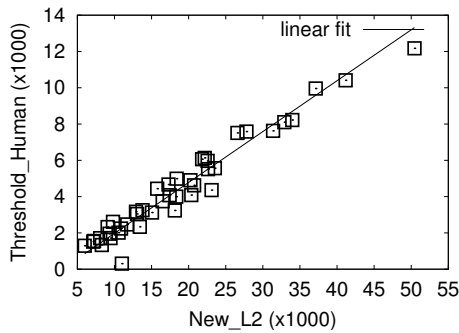
Figure 8: **Comparison of the White-Noise and Gaussian L2 Values with Human-Chosen Thresholds.** The new L2 values derived from the white-noise based distribution against the human-chosen thresholds for a set of 42 natural images are shown (each square represents one image). The correlation coefficient was $r = 0.98$ (higher than $r = 0.91$ in Lee & Choe (2003)). The least-square fit is shown in the background.

tion energy (or wavelet response) histograms has also been recognized and utilized for some time now, especially in denoising and compression (Simoncelli & Adelson 1996). Our approach was inspired by Barlow (1994) that comparing peaked distributions with high kurtosis and distributions derived from an unsuspicious baseline could be useful. The current work, to our knowledge, is the first systematic study of the relationship between measures of salience (perceived and objective measures) derived from a normal-distribution baseline and the concept of suspicious coincidence.

Response histograms in general are also widely used in computational vision. For example, Liu & Wang (2002) used what they call the spectral histogram to segment and synthesize texture images. It would be interesting to find out whether other response histograms can be analyzed and used in a similar manner as described in this paper for salience detection in different image feature spaces.

Finally, there are alternative theoretical frameworks of salience related or complementary to our work. Itti & Baldi (2006) defined saliency based on KL divergence between the prior and the posterior probability after an observation has been made. However, their model requires extensive computation, while ours only require the local variance in orientation energy, thus is more efficient. There are other theoretical definitions of salience based on information theory. Bruce & Tsotsos (2006) defined saliency based on information maximization, but they did not give a threshold criterion for saliency, unlike our approach. Jägersand (1995), on the other hand, used differential information gain across scale space to define salience, but again, he did not address the issue of how to set the saliency threshold.

Conclusion

In this paper, we have shown that salience measures derived from comparisons of the orientation energy distribution with a normal distribution have an objective basis. Further, we have demonstrated that the good performance seen in OED-based thresholding can be analyzed and understood under the general framework of suspicious coincidence. We directly used a scaled energy distribution from white-noise images to further demonstrate this point through a comparison to human performance. The results suggest that a similar approach can be applied to other sensory tasks where a similar response distribution is found.

References

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

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

Bruce, N., and Tsotsos, J. 2006. Saliency based on information maximization. In Weiss, Y.; Schölkopf, B.; and Platt, J., eds., *Advances in Neural Information Processing Systems 18*. Cambridge, MA: MIT Press. In press.

Canny, J. 1983. A variational approach to edge detection. In *Proc. of the 3rd National Conference on Artificial Intelligence*, 54–58. Menlo Park, CA: AAAI.

Daugman, J. G. 1980. Two-dimensional spectral analysis of cortical receptive field profiles. *Vis. Res.* 20:847–856.

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.

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* 41:711–724.

Itti, L., and Baldi, P. 2006. Bayesian surprise attracts human attention. In Weiss, Y.; Schölkopf, B.; and Platt, J., eds., *Advances in Neural Information Processing Systems 18*. Cambridge, MA: MIT Press. In press.

Jägersand, M. 1995. Saliency maps and attention selection in scale and spatial coordinates: An information theoretic approach. In *Proc. of ICCV*, 195–202.

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. *J. Neurophy.* 58:1233–1258.

Lee, H.-C., and Choe, Y. 2003. Detecting salient contours using orientation energy distribution. In *Proc. Int. Joint Conf. on Neural Networks*, 206–211. IEEE.

Lee, A. B.; Mumford, D.; and Huang, J. 2001. Occlusion models for natural images: a statistical study of a scale-invariant dead leaves model. *International Journal of Computer Vision* 41:35–59.

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.

Simoncelli, E. P., and Adelson, E. H. 1996. Noise removal via bayesian wavelet coring. In *Proc. IEEE Int. Conf. on Image Proc*, Vol. I, 379–382.