Relationship Between Suspicious Coincidence in Natural Images and Oriented Filter Response Distributions

Yoonsuck Choe and Subramonia Sarma
Department of Computer Science, Texas A&M University
College Station, TX 77843-3112
choe@tamu.edu,sps8556@cs.tamu.edu

Abstract

Visual cortical neurons have receptive fields resembling oriented band-pass filters, and their response distributions on natural images are non-Gaussian, with a sharp peak at zero and heavy tails. Inspired by this, we previously showed that comparing the response distribution to a normal distribution with the same variance gives us a good thresholding criterion for detecting salient levels of edginess in images. However, the question why the comparison to a normal distribution can be so effective was never answered. In this paper, we approach this issue under the general framework of *suspicious coincidence* proposed by Barlow. It turns out that white-noise images (where all pixels are independent, thus no suspiciousness) give rise to response distributions closely matching that of a normal distribution. Thus, salience defined our way can be understood as a deviation from the unsuspicious baseline. Further, we show that the response threshold directly calculated from the white-noise based distribution closely matches that of humans, providing further support for the analysis.

1 Introduction

Oriented Gabor filters have been used successfully to model visual cortical neuron response [1], and as it turns out, they are biologically grounded, i.e., the shape of the Gabor filters closely resemble experimentally measured receptive fields [2]. 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 [3, 4, 5] (see figure 1).

Such a response property has been found to be useful in tasks such as denoising [5] and salient contour detection through thresholding [6]. For instance, Lee and Choe [6] 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. Although the method was effective, it was not shown why a normal distribution serves so well as a baseline in such a comparison.

Yoonsuck Choe and Subramonia Sarma. Relationship between suspicious coincidence in natural images and oriented filter response distributions. Technical Report 2003-8-4, Department of Computer Science, Texas A&M University, 2003.

In this paper, we frame the problem in terms of the concept of *suspicious coincidence*, proposed by Barlow [7]: Two statistical events A and B are said to be suspicious if they occur more often together than can be expected from their individual probabilities. In other words, the joint probability for the two events should exceed the product of their individual probabilities in order for them to be deemed suspicious:

$$P(A, B) > P(A)P(B). \tag{1}$$

This approach is easily extended into the problem of image analysis, where each pixel is treated as a random variable. Suspiciousness in the image can then be determined by testing the inequality shown above where events correspond to pixels from different locations in the image [8].

Suspiciousness is directly linked to salience, i.e., more suspicious events may be seen as more salient to a perceptual system, and this is where we try to find the connection between oriented filter response distribution and the concept of suspiciousness. According to the definition above (equation 1), an image where each pixel is independent from each other (e.g., a white-noise image) would be seen as containing no suspicious coincidence between any pair of pixels. Thus a white-noise image can be said to have no suspicious feature in it, and this is generally what we humans perceive. If we equate salience with suspiciousness, then this also implies that algorithms such as [6] should fail to detect any salience in white-noise images.

In the following sections, we first describe the general procedure for calculating oriented filter response and its empirical distribution, and how to calculate the corresponding normal distribution with the same variance. Next, we measure the orientated filter response on white-noise images and compare its histogram to its matching normal distribution. The results show that the white-noise based distribution closely resembles a normal distribution of the same variance, suggesting that normal distributions can serve as a baseline for the detection of suspicious coincidence. The new distribution derived from white-noise images are then used as a new baseline, and the results are shown to be consistent with the findings in [6]. Finally, we discuss some related work and issues raised by our model, followed by the conclusion.

2 Calculation of Oriented Filter Response

To find the orientation response (or energy) distribution, we follow the procedure described in [9]. 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 two Gaussian functions whose widths differ by a factor of 0.5:

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

where $G_{\sigma^2}(\cdot)$ is a Gaussian function with variance σ^2, defined as follows:

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

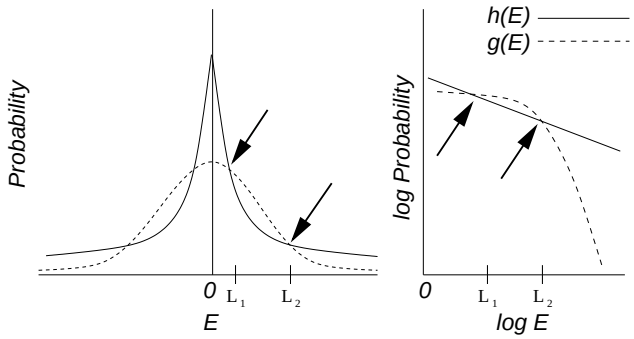
where (x, y) denotes the pixel location.

The gray level intensity matrix I of the input image is convolved with the DoG filter to obtain the resultant matrix I_d:

$$I_d = I * D, \tag{4}$$

where $*$ is the convolution operator. We used a DoG filter of size 7×7 for all of our experiments.

The filtered image is then convolved with oriented Gabor functions $R_{\theta,\phi,\sigma}(x, y)$ [3] of both even and odd phases with orientation θ, phase ϕ, and width σ to obtain the orientation



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

Figure 1: **Power Law vs. Normal Distribution.** A distribution $h(E)$ following a power law (solid curve) and a normal distribution $g(E)$ with the same variance (dashed curve) are shown in (a) linear scale and (b) log-log scale. The x-axis is the orientation energy E and the y-axis the probability. The two points where the curves intersect are labeled L_1 and L_2. For the E values less than L_1 or greater than L_2, *h(E)* is greater than $g(E)$. (For the current work, we only consider positive E values.)

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 Gabor function is defined as

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

and using this, the orientation energy matrix for a single orientation θ is found as

$$E_\theta = (R_{\theta,0,\sigma} * I_d)^2 + (R_{\theta,\frac{\phi}{2},\sigma} * I_d)^2, \tag{6}$$

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:

$$h(E) = \frac{f(E)}{\sum_{x \in B_h} f(x)}, \tag{7}$$

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 gives 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 [6], 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

$$\sigma_h^2 = \sum_{x \in B_h} x^2 h(x) \tag{8}$$

We use this calculated σ_h^2 to find the matching continuous normal probability density function $N(x; 0, \sigma_h^2)$ 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:

$$g(E) = \frac{N(E; 0, \sigma_h^2)}{\sum_{x \in B_h} N(x; 0, \sigma_h^2)}. \quad (9)$$

Note that E is always greater than zero (equation 6), i.e., $\forall E \in B_h, E \geq 0$, thus the above is similar to a half-normal distribution discussed in [10].

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 $g(E)$ to investigate their relationship to the concept of suspicious coincidence.

3 Experiments and Results

For the experiment, we used 42 natural images (resized to a size of 256×256) from a set of freely available stock photos from Kodak (the same source as in [9]). For each image, the orientation energy E was calculated, and to avoid any artefactual border effect, only the circular central area of the energy matrix was used for the calculation of its distribution $h(E)$ and the matching normal distribution $g(E)$, following [9]. 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 (figure 1). For example, figure 2 shows a typical orientation energy distribution $h(E)$ compared to its matching normal distribution $g(E)$ calculated from a natural image (shown in figure 3a). The straight declining slope characteristic of a power law distribution is evident in $h(E)$. The two curves intersect at two points, near $E \sim 500$ (i.e., L_1; see figure 1 for notations) and $E \sim 7,000$ (L_2). Beyond L_2, $g(E)$ plummets, but $h(E)$ remains high relative to $g(E)$.

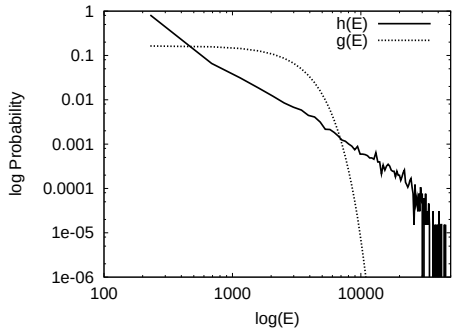


Figure 2: **Orientation Energy Distribution of a Natural Image vs. Its Matching Normal Distribution.** The orientation energy distribution $h(E)$ (solid line) of a natural image shown in figure 3 and the matching normal distribution $g(E)$ (dashed line) with the same variance are shown in log-log plot.

Lee and Choe empirically derived the effective threshold for the detection of salient contours, which was linear to the orientation energy corresponding to the second point of intersection (L_2) of the response distribution and its matching normal distribution [6]. For example, figure 3 shows the effect of thresholding on the image of a bird using the threshold derived from L_2. Only the salient edges in the image remain after the thresholding, and the resulting plot is very close to our perceived *edginess* in the image (figure 3c).

Even though the thresholding criterion was effective as we have seen above, it does not tell us why the simple idea of comparing to a normal distribution has to be so effective. That is,

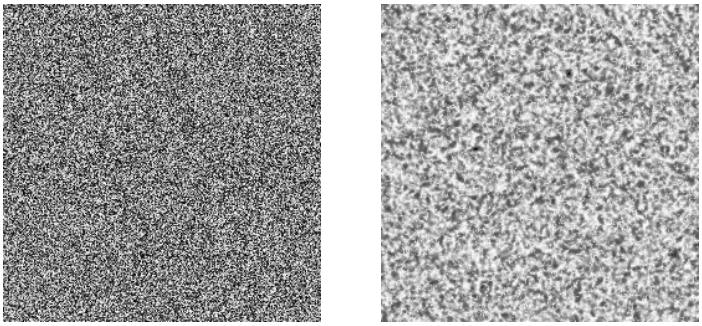
(a) Original Image (b) Orientation Energy (c) Thresholded Edges (d) Magnified (c)

Figure 3: **Orientation Energy Thresholding in a Natural Image.** An example of orientation energy thresholding is shown: (a) The original image; (b) its orientation energy in grayscale (black is high and white is low); (c) the pixels after thresholding shown as oriented edges representing the detected edge orientation at that location; and (d) the magnified view of the head of the bird in (c). The threshold result is close to the our perceived salience of edges in the original image (a).

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 [7]) 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 under our criterion. 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 the matching normal distribution. The white-noise image was a 256×256 intensity matrix of uniformly randomly distributed values between 0 and 255. The orientation energy distribution was then found using the procedure outlined in the previous section. The image and its orientation energy are shown in figure 4. 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 (figure 5). Results by Simoncelli and Adelson [5] 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 (i.e., denoising). 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 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 L_2 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 was the standard deviation before scaling, we multiply 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 L_2 values were then found computationally by comparing the two distributions. These values were compared to the orientation energy thresholds selected by humans. All of the thresholds for the 42 natural images were determined by a single person (SB) in our research group. For each image, the thresholded E at 55% to 95% percentile of $h(E)$ at an interval of 5% (simi-



(a) White-Noise Image (b) Orientation Energy

Figure 4: **Orientation Energy of a White-Noise Image.** (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.

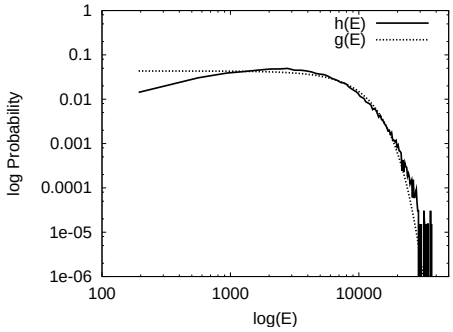


Figure 5: **Orientation Energy Distribution of a White-Noise Image vs. Its Matching Normal Distribution.** The log-log plot shows the orientation energy distribution $h(E)$ for the white-noise image in figure 4 against a normal distribution of the same variance $g(E)$. The two distributions show a close resemblance.

lar to figure 3c) was shown to SB and he determined the best threshold according to the following criteria [6]: (1) object contours should remain intact, and (2) noisy background edges must be removed, as much as possible. The results are shown in figure 6a. It is clear that the new white-noise based L_2 values also have a strong linear relationship with the human-selected thresholds, even more so than the old Gaussian-based L_2 values. (We are currently investigating the cause of this small difference.)

In sum, we have shown that our criterion for salience determined by comparing the orientation energy distribution to a normal distribution with the same variance can be understood as a deviation from the unsuspicious baseline where all pixels are independent (i.e., a white-noise image). Also, as expected, orientation energy thresholds directly based on white-noise derived distributions showed a close match to that of humans.

4 Discussion

The non-Gaussian nature of orientation energy (or wavelet response) histograms has been recognized and utilized for some time now, but in a different context compared to our approach: e.g., denoising and compression [11, 5]. Our approach is not unique however: For example, Barlow noted the usefulness of comparing peaked distributions with high kurtosis and distributions derived from an unsuspicious baseline [8]. On the other hand, to our knowledge, the current work is the first systematic study of the relationship between

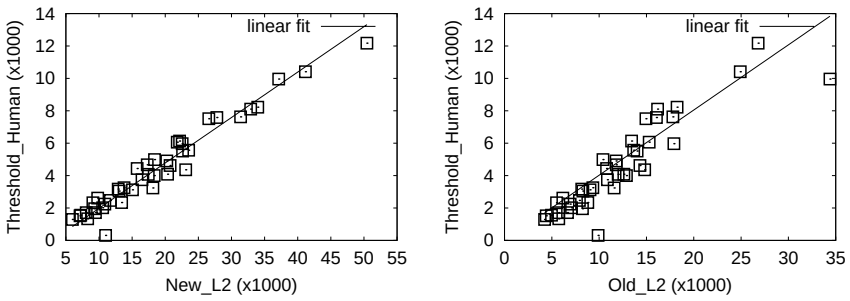


Figure 6: **Comparison of the White-Noise- vs. the Gaussian-Based L_2 Values.** (a) The new L_2 values derived from the white-noise based distribution against the human-chosen thresholds for a set of 42 natural images are shown (each square represents one image). The correlation coefficient was $r = 0.98$. (b) The same is shown for the case with the old L_2 values based on the normal-distribution baseline. The correlation coefficient was $r = 0.91$, lower than in (a). In both plots, the least-square fit is shown in the background.

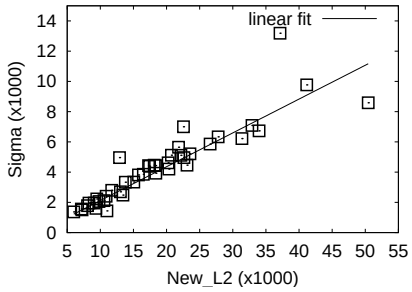


Figure 7: **White-Noise Based L_2 vs. σ_h.** The new L_2 values derived from the white-noise based distribution against the raw standard deviation σ_h of the distribution $h(E)$ is shown. (Each square represents one image.) We can see a clear linear relationship ($r = 0.91$).

perceived salience in humans and the orientation energy distribution under the framework of suspicious coincidence.

Another interesting observation made by Lee and Choe is that the L_2 points and the raw standard deviation σ_h of the orientation energy distribution of natural images have a close linear relationship [6]. As expected from the experiments in this paper, we get the same kind of linear relationship with the new white-noise based L_2 (see figure 7). As pointed out by Lee and Choe in [6], this result suggests that a very simple neural mechanism can extract the appropriate orientation energy threshold. A quadratic activation function on V1 response, followed by a weighted sum, with a final activation function in the form of a square root can quickly produce σ_h without even constructing a histogram, nor the matching normal distribution. It is quite striking that a measure as simple as σ can be used in such a useful way.

Response histograms in general are also widely used. For example, Liu and Wang used what they call the spectral histogram (which is a combination of many different kinds of filter response histograms) to segment and synthesize texture images [12]. It would be interesting to find out whether other response histograms can be analysed and used in a similar manner as described in this paper for salience detection in different image feature spaces.

5 Conclusion

In this paper, we have shown that the good performance shown in orientation energy thresholding based on the comparison of the orientation energy distribution and its matching normal distribution 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.

6 Acknowledgments

We would like to thank J. C. Liu for helpful discussions, and S. Kumar Bhamidipati for assistance with threshold selection. This research was supported in part by the Texas Higher Education Coordinating Board ARP/ATP program grant 000512-0217-2001.

References

[1] J. G. Daugman, "Two-dimensional spectral analysis of cortical receptive field profiles," *Vision Research*, vol. 20, pp. 847–856, 1980.

[2] J. P. Jones and L. A. Palmer, "An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex," *Journal of Neurophysiology*, vol. 58, no. 6, pp. 1233–1258, 1987.

[3] J. G. Daugman, "Entropy reduction and decorrelation in visual coding by oriented neural receptive fiends," *IEEE Transactions on Biomedical Engineering*, vol. 36, pp. 107–114, 1989.

[4] D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," *Journal of the Optical Society of America A*, vol. 4, pp. 2379–2394, 1987.

[5] E. P. Simoncelli and E. H. Adelson, "Noise removal via bayesian wavelet coring," in *Proceedings of IEEE International Conference on Image Processing*, vol. I, 1996, pp. 379–382.

[6] H.-C. Lee and Y. Choe, "Detecting salient contours using orientation energy distribution," in *Proceedings of the International Joint Conference on Neural Networks*. IEEE, 2003, in press.

[7] H. B. Barlow, "Unsupervised learning," *Neural Computation*, vol. 1, pp. 295–311, 1989.

[8] H. Barlow, "What is the computational goal of the neocortex?" in *Large Scale Neuronal Theories of the Brain*, C. Koch and J. L. Davis, Eds. Cambridge, MA: MIT Press, 1994, pp. 1–22.

[9] W. S. Geisler, J. S. Perry, B. J. Super, and D. P. Gallogly, "Edge Co-occurrence in natural images predicts contour grouping performance," *Vision Research*, vol. 41, pp. 711–724, 2001.

[10] C. Y. Chou and H. R. Liu, "Properties of the half-normal distribution and its application to quality control," *Journal of Industrial Technology*, vol. 14, pp. 4–7, 1998.

[11] M. H. Hansen and B. Yu, "Wavelet thresholding via MDL: Simultaneous denoising and compression," *IEEE Tranasctions on Information Theory*, vol. 45, pp. 1778–1788, 2000.

[12] X. Liu and D. Wang, "A spectral histogram model for texton modeling and texture discrimination," *Vision Research*, vol. 42, pp. 2617–2634, 2002.