

Is the world full of circles?

Carson C. Chow

Department of Mathematics and Center for the Neural Basis
of Cognition, University of Pittsburgh, Pittsburgh, PA, USA



Dezhe Z. Jin

Howard Hughes Medical Institute and Department of Brain and
Cognitive Sciences, Massachusetts Institute of Technology,
Cambridge, MA, USA



Alessandro Treves

Cognitive Neuroscience, Scuola Internazionale
Superiore di Studi Avanzati, Trieste, Italy



The statistical arrangement of oriented segments in natural scenes was recently proposed to be indicative of a cocircularity rule. In particular, the probability density function of the relative position of two oriented segments was found to be maximal along fixed angles on the plane, consistent with the two segments being tangent to two points of a circle. Does this observation point to a prevalence of circles in natural scenes? Here we demonstrate that similar statistics can be obtained even when circles are not very common in visual scenes. The reason is that circles or near circular objects can heavily skew the distribution in favor of the cocircularity rule.

Keywords: natural scenes, cocircularity, orientation selectivity, gestalt, object perception, closed contours, vision

1 Observations From Natural Scenes

An interesting recent work by [Sigman, Cecchi, Gilbert, & Magnasco \(2001\)](#) examines the relative placement of pairs of oriented segments in natural scenes, and relates the observed statistics to the properties of the visual system. The authors use a filtering procedure to extract, from a large database of natural scenes, the probability of finding a short segment of orientation ψ at a certain position relative to another short segment of orientation ϕ . They discuss both the scaling properties of natural scenes (essentially, how the probability above scales with the distance between the segments) and their spatial structure (how it varies along the circle, at a given distance).

Here we are concerned only with the spatial structure. Let us define as \mathcal{G} the angle at which the second segment may occur, at a fixed distance from a first segment placed at the origin. The observation made in the study by [Sigman et al. \(2001\)](#) is that the probability $p(\mathcal{G})$ is maximal, with very good precision, at $\mathcal{G}^* = (\psi - \phi)/2$. This is exactly what one would expect if the segments were mostly oriented so as to be tangent to circles of arbitrary size, because circles always contribute to $p(\mathcal{G})$ at \mathcal{G}^* . One possibility for generating this cocircularity rule is, therefore, that natural visual scenes typically include an abundance of circles. In the discussion of [Sigman et al. \(2001\)](#), the authors conclude that “[The line and the

circle] are, in the same order, the most significant structures in natural scenes.” However, they also note that “Cocircularity in natural scenes probably arises because of the continuity and smoothness of object boundaries; when averaged over objects of vastly different sizes present in any natural scene, the most probable arrangement for two edge segments is to lie on the smoothest curve joining them, a circular arc. These ideas, however, require an investigation that is beyond the scope of this paper.”

Our aim was to investigate the cocircularity hypothesis by constructing visual scenes consisting only of simple geometrical objects, of which circles are a subclass. We used ellipses and stadia (two semicircles connected by straight lines). We found that circles need not be a very significant structure to obtain the statistics observed in $p(\mathcal{G})$ ([Sigman et al., 2001](#)). The observed statistics arise, in our constructed scenes, from the fact that circles give a contribution to $p(\mathcal{G})$ concentrated exclusively at \mathcal{G}^* , while noncircular shapes contribute to the distribution in a range around \mathcal{G}^* . We note that although we have chosen to consider only closed smooth contours, our analysis easily generalizes to segments of these contours. The conclusions would be the same. The presence of a few circular arcs would also give a singular contribution to cocircularity. We propose that an effective over-weighting of circles together with the presence of smooth closed objects may be a possible explanation for the cocircular statistics observed by [Sigman et al. \(2001\)](#).

2 Simple Models of Visual Scenes

We constructed model scenes consisting only of contours of simple geometrical objects, at random positions and orientations. We first note that the distribution $p(\mathcal{G})$ can be broken down into two parts. One part includes the contributions from pairs of line segments belonging to different objects. This part is a constant with respect to \mathcal{G} , because the relative positions and orientations of the different objects are random. The other part includes the contributions from the pairs belonging to the same object. This part, which we redefine as our new $p(\mathcal{G})$, is responsible for any nonuniform features in the distribution of \mathcal{G} values. Second, we note that objects whose contours include sharp angles are unlikely to favor any particular orientation. This is intuitive for very irregular polygons, but it also holds for more regular geometric shapes consisting of straight edges, such as rectangles. For example, with a square, the only possible (nonzero) orientation is given by two lines at $\pi/2$, and the angle \mathcal{G} between them can range from 0 to $\pi/2$. We therefore limit our objects to smooth closed contours. It is smoothness in the contours that should induce, as noted by Sigman et al. (2001), nonuniformity in $p(\mathcal{G})$. For simplicity, we consider just two smooth geometrical shapes: stadia and ellipses.

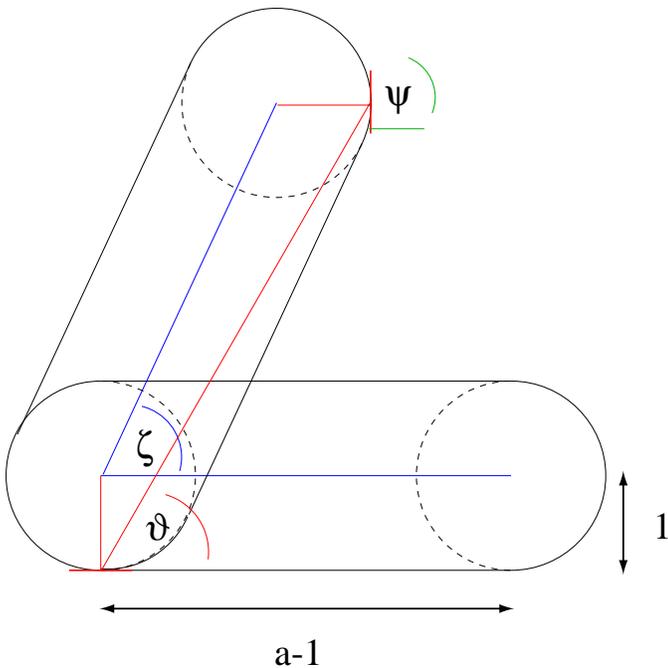


Figure 1. The geometry of the stadium.

A stadium is defined to be two semicircles of radius τ , connected by two straight lines of length $a-\tau$ (Figure 1).

An ellipse is specified by a semimajor axis a and a semiminor axis τ (Figure 2). Because we are not interested in scaling properties, we set $\tau = 1$ without loss of generality, and take the elongation of both shapes to be parameterized by a , with $a > 1$. (Stadia have total length $a + 1$ and ellipses have total length $2a$.) For $a = 1$, both shapes reduce to circles. Our elongation statistics are thus given by setting a model distribution $p(a)$. Any $p(a)$ that includes a finite density at $a = 1$ will include circles, although the number of exact circles in any large sample of shapes will be, strictly speaking, zero (because we are considering probability densities for the continuous parameter a). Without loss of generality, we take our shapes to intersect the origin tangent to the x -axis, so that $\varphi = 0$. We are thus left with an orientation angle ζ between the main axis of the shape and the x -axis. Circles are, of course, indifferent to the orientation angle.

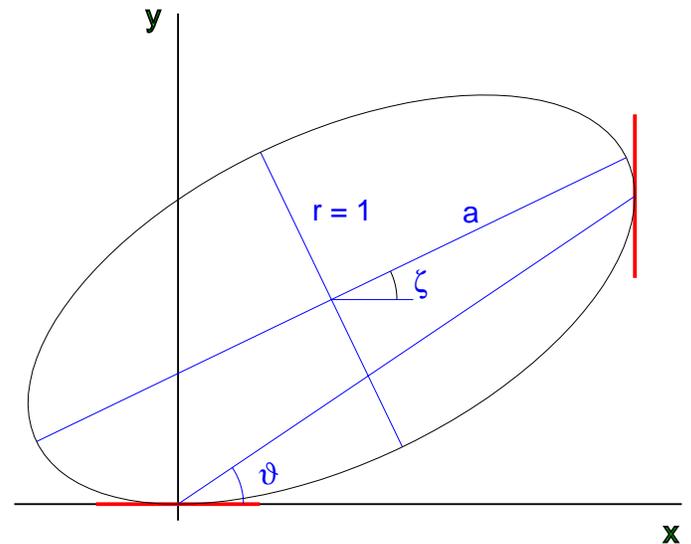


Figure 2. The geometry of the ellipse.

We can now consider $p(\mathcal{G})$ when presented with ellipses and stadia with a given statistical distribution $p(a)$, and with $p(\zeta) \equiv 1/(2\pi)$. Note that in Sigman et al. (2001), $p(\mathcal{G})$ was calculated by scanning each visual scene for edges and then computing the relative placement of oriented segments for the entire scene. Segments at orientation φ were compared to those at orientation ψ . The distribution $p(\mathcal{G})$ was then extracted by averaging over all orientations for a fixed $\varphi-\psi$, and over an ensemble of scenes. For computational ease, they considered 16 discrete angular values. In our model scenes, the equivalent calculation is to take each shape in the scene and translate it to the origin so that it is tangent to the x -axis, and then take all orientations ζ of the object. Thus \mathcal{G} is the angle to the location of a point on the shape that is tangent to a ψ -oriented segment. We also consider a continuum of angles.

2.1 Stadia

For two segments tangent to the end-circles of a stadium, elementary geometry shows that the relative angle \mathcal{G} for $\varphi = 0$ is given by $\mathcal{G} = f(\zeta, \psi; a)$, where

$$f(\zeta, \psi, a) = \tan^{-1} \left[\frac{(a-1)\sin \zeta + 1 - \cos \psi}{(a-1)\cos \zeta + \sin \psi} \right] \quad (1)$$

and ζ is the orientation of the stadium, with $\zeta = 0$ corresponding to the straight edge of the stadium being parallel to the x -axis. As seen in [Figure 1](#), [Equation 1](#) was derived by observing that the rotation of the stadium by ζ is equivalent to rotating a circle of unit radius about the point $(0,1)$, with center located a distance a away from $(0,1)$. By symmetry, we need only to focus on the range $0 \leq \zeta \leq \pi/2$ (see [Figure 1](#)). The stadium also constrains ψ to the range $\zeta \leq \psi \leq \pi + \zeta$. Note that $\tan^{-1}[(1 - \cos \psi)/(a - 1 + \sin \psi)] \leq \mathcal{G} \leq \tan^{-1}[(a - \cos \psi)/\sin \psi]$, for $0 \leq \zeta \leq \pi/2$.

When $\zeta = 0$, the straight edge of the stadium is tangent to the x -axis. Thus there is a range of positions the stadium can take and still maintain tangency. As seen in [Figure 1](#), the stadium can be slid along the axis, from one end of the straight edge at the origin, to the other end. In this case

$$\mathcal{G} = g(s, \psi, a) \equiv \tan^{-1} \left[\frac{1 - \cos \psi}{s + \sin \psi} \right] \quad (2)$$

where s , $0 \leq s \leq a - 1$, is the distance from the origin to the right end of the stadium straight edge. Applying the limits of s to [Equation 2](#) and applying the half angle formula for tangents shows that sliding along the x -axis will imply $\tan^{-1}[(1 - \cos \psi)/(a - 1 + \sin \psi)] \leq \mathcal{G} \leq \psi/2 \equiv \mathcal{G}^*$. For values greater than $\psi/2$, we must consider sliding along the y -axis.

Two further contributions should be considered. The first is for $\psi = \zeta$, that is, when the second segment is tangent to the straight edge of the stadium. This contribution, once integrated over ζ , is equivalent to the contribution above, integrated over s (by symmetry, exchanging end-circle and straight edge). A factor of 2 is, therefore, all that is needed. The second contribution is from segments that are both tangent to the same end-circle, that is for $\psi \leq \zeta$. This is obviously, and without further calculation, a pure cocircularity contribution, implying $\mathcal{G} = \psi/2$.

We constructed the a -dependent angular distribution $p(\mathcal{G}|a)$. The full distribution $p(\mathcal{G})$ is computed by integrating over a distribution of elongations a that is present in our mock visual scene. We consider the contribution to $p(\mathcal{G}|a)$ from rotating $p_r(\mathcal{G}|a)$ and sliding $p_s(\mathcal{G}|a)$ separately:

$$p_r(\mathcal{G}|a) = \frac{2}{\pi} \int_0^{\pi/2} \delta(\mathcal{G} - f(\zeta, \psi, a)) d\zeta \quad (3)$$

$$p_s(\mathcal{G}|a) = \frac{2}{(a-1)} \int_0^{a-1} \delta(\mathcal{G} - g(s, \psi, a)) ds \quad (4)$$

which reduce to

$$p_r(\mathcal{G}|a) = \frac{1}{\pi} \left| \frac{df}{d\zeta} \right|^{-1}_{f(\zeta, \psi, a) = \mathcal{G}} \quad (5)$$

$$p_s(\mathcal{G}|a) = \frac{2}{(a-1)} \left| \frac{dg}{ds} \right|^{-1}_{g(s, \psi, a) = \mathcal{G}} \quad (6)$$

[Equation 5](#) cannot be solved in closed form but we note that for parameter a large, $f(\zeta, \psi, a)$ is nearly linear in ζ . Thus $p_r(\mathcal{G}|a)$ is a function which is nonzero and nearly constant between $\tan^{-1}[(1 - \cos \psi)/(a - 1 + \sin \psi)]$ and $\tan^{-1}[(a - \cos \psi)/\sin \psi]$, with total area equal to one. When a is near 1, $p_r(\mathcal{G}|a)$ has a very large amplitude near $\mathcal{G} = \tan^{-1}[(1 - \cos \psi)/\sin \psi] = \psi/2$ (using the tangent half-angle formula). Thus, in the limit $a \rightarrow 1$, $p_r(\mathcal{G}|a) \rightarrow \delta(\mathcal{G} - \psi/2)$.

Evaluating the sliding contribution ([Equation 6](#)) yields

$$p_s(\mathcal{G}|a) = \frac{2 \sin^2(\mathcal{G})}{(a-1)(1 - \cos \psi)} \quad (7)$$

where $\tan^{-1}[(1 - \cos \psi)/(a - 1 + \sin \psi)] \leq \mathcal{G} \leq \psi/2$. $p_s(\mathcal{G}|a)$ is maximal at $\psi/2$. Sliding along the y -axis will give the contribution for $\mathcal{G} \geq \psi/2$, which decreases with \mathcal{G} . Again we find that this distribution is confined to a range around $\psi/2$ with a maximum at $\psi/2$. As $a \rightarrow 1$, $p_s(\mathcal{G}|a) \rightarrow 2\delta(\mathcal{G} - \psi/2)$.

Thus, for a visual scene consisting exclusively of stadia with a distribution of elongations a , we find that $p(\mathcal{G})$ will have a maximum near $\psi/2 \equiv \mathcal{G}^*$, indicating cocircularity, as observed in [Sigman et al. \(2001\)](#). The presence of a very small number of stadia with a near 1 (near circular) will contribute significantly to $p(\mathcal{G})$ at $\mathcal{G} = \mathcal{G}^*$. This will remain true when sampling only at discrete angles. We note that this result is perhaps not too surprising as the stadium is composed of two semicircles. In the next section, we consider ellipses where no arc segment is circular, and find similar results.

2.2 Ellipses

We now consider the distribution for a family of ellipses. For simplicity, we first consider the case $\psi = \pi/2$ (see [Figure 2](#)). We generalize to arbitrary ψ later. We first calculate $p(\mathcal{G}|a)$, which is the distribution of \mathcal{G} for ellipses with elongation a . We then calculate $p(\mathcal{G})$ for several distributions of elongations $p(a)$. By symmetry, we again only need to consider orientation angles $0 \leq \zeta \leq \pi/2$. The ellipse is described by the following equation:

$$\begin{aligned} (x \cos \zeta + y \sin \zeta - x_0)^2 / a^2 + \\ (x \sin \zeta - y \cos \zeta - y_0)^2 = 1 \end{aligned} \quad (8)$$

where

$$x_0 = \frac{a^2 \sin \zeta}{\sqrt{a^2 \sin^2 \zeta + \cos^2 \zeta}} \tag{9}$$

$$y_0 = -\frac{\cos \zeta}{\sqrt{a^2 \sin^2 \zeta + \cos^2 \zeta}} \tag{10}$$

At the location (x_i, y_i) (the point where the tangent to the ellipse is vertical), $dy/dx = \infty$. With this condition and Equation 8, we can solve for x_i and y_i . Because $\tan \vartheta = y_i/x_i$, we find

$$\vartheta = h(\zeta, a) = \tan^{-1} \left(\sqrt{\frac{a^2 \sin^2 \zeta + \cos^2 \zeta}{a^2 \cos^2 \zeta + \sin^2 \zeta}} \right) \tag{11}$$

and $\tan^{-1}(1/a) \leq \vartheta \leq \tan^{-1}(a)$.

Considering all orientations, $0 \leq \zeta \leq 2\pi$, we find

$$\begin{aligned} p(\vartheta|a) &= \frac{2}{\pi} \int_0^{\pi/2} \delta(\vartheta - h(\zeta, a)) d\zeta = \frac{2}{\pi} \left[\frac{dh}{d\zeta} \right]_{h(\zeta, a)=\vartheta}^{-1} \\ &= \frac{2(a^2 + 1)}{\pi \sqrt{(a^2 - \cot^2 \vartheta)(a^2 - \tan^2 \vartheta)}} \end{aligned} \tag{12}$$

and $p(\vartheta|a) = 0$ outside of the range $\tan^{-1}(1/a) \leq \vartheta \leq \tan^{-1}(a)$. With Equation 12, we obtain the distribution $p(\vartheta)$ via

$$p(\vartheta) = \int_{\max(\tan \vartheta, \cot \vartheta)}^{\infty} p(\vartheta|a) p(a) \tag{13}$$

From Equation 12, it is apparent that $p(\vartheta|a)$ reaches a minimum value of $(2/\pi)(a^2+1)/(a^2-1)$ at $\vartheta = \pi/4$. The height of this minimum however increases as $a \rightarrow 1$, and the width of the distribution decreases. As an example, we plot $p(\vartheta|a=1.2)$, $p(\vartheta|a=2)$ and $p(\vartheta|a=4)$ superimposed in Figure 3. For a given elongation a , $p(\vartheta|a)$ is restricted to a finite range around $\vartheta^* = \pi/4$ but with a *minimum* at $\pi/4$. Thus, for a visual scene of ellipses of a single elongation, cocircularity is not dominant.

However, as a decreases, the width will decrease and the height will increase. In the limit of $a \rightarrow 1$, $p(\vartheta|1) \rightarrow \delta(\vartheta - \pi/4)$ because its integral equals 1, while the interval over which it differs from zero vanishes, $\tan^{-1}(a) - \tan^{-1}(1/a) \rightarrow 0$. Hence, in a visual scene that includes ellipses with a distribution of elongations, if $p(a)$ is nonzero near $a = 1$, (i.e., the distribution includes near circular ellipses), then $p(\vartheta)$ can reach a maximum at $\vartheta = \pi/4$ (instead of a local minimum). The presence of a

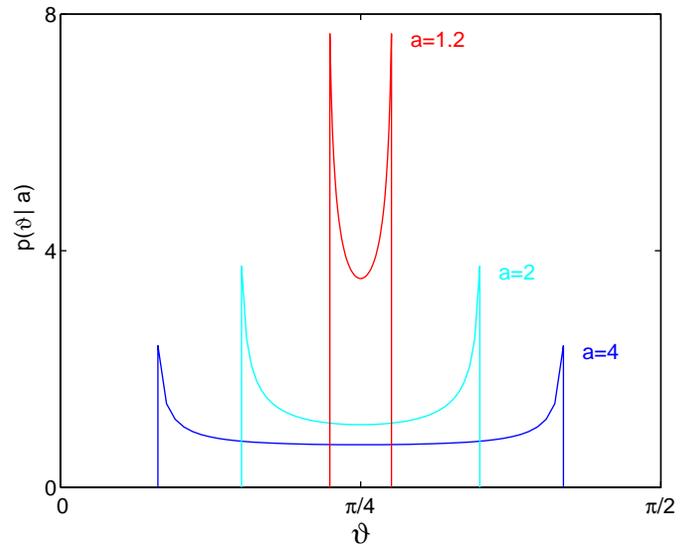


Figure 3. Distribution function $p(\vartheta|a)$ for ellipses with $a = 1.2, 2, 4$.

few near circular objects can skew the distribution toward cocircularity.

To illustrate this point, we construct scenes with ellipses at random positions and orientations, with elongation a sampled from three different distribution functions. The first is the Γ distribution

$$p(a) = \frac{1}{\sigma^2} (a-1) e^{-(a-1)/\sigma} \tag{14}$$

where $\sigma > 0$ is a parameter. This distribution peaks at $a = 1 + \sigma$. The second is the triangular distribution

$$p(a) = \begin{cases} \frac{2(a-1)}{(a_0-1)(a_{\max}-1)} & \text{if } 1 < a < a_0 \\ \frac{2(a_{\max}-a)}{(a_{\max}-a_0)(a_{\max}-1)} & \text{if } a_0 < a < a_{\max} \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

Here a_0 and a_{\max} are the parameters. The third is the uniform distribution

$$p(a) = \begin{cases} 1/a_{\max} & \text{if } 1 < a < a_{\max} \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

In Figure 4, we show an example of the scene when a is sampled from the Γ distribution with $\sigma = 0.5$. As can be seen, nearly circular ellipses are rare.

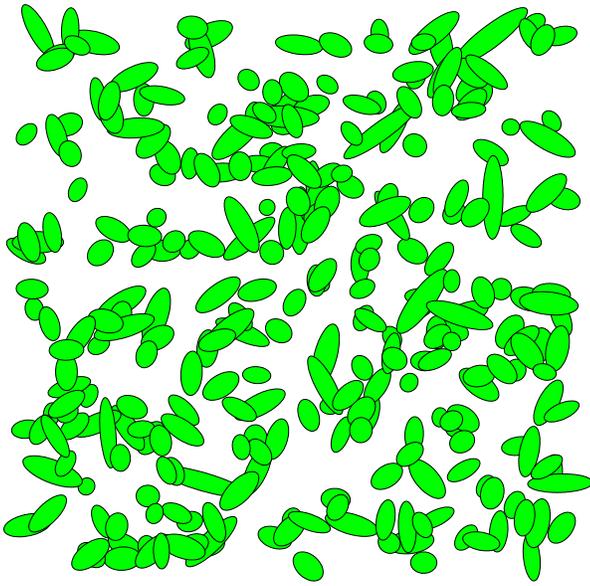


Figure 4. Scene generated with random ellipses. The ellipses have the short axis b randomly chosen from 1 to 1.5, and elongation a sampled from the Γ distribution defined in Equation 14 with $\sigma = 0.5$. The centers and orientations are randomly distributed.

Using Equations 12 and 14-16, we can numerically integrate Equation 13 to get $p(\mathcal{G})$. In Figure 5, we plot $p(\mathcal{G})$ for three cases of $p(a)$: a Γ distribution with $\sigma = 0.5$; a triangular distribution with $a_0 = 1.2$ and $a_{\max} = 4$; and a uniform distribution with $a_{\max} = 4$. From the curves, we find that $p(\mathcal{G})$ always peaks at $\mathcal{G} = \pi/4$. The flat distribution has a significant peak at $\mathcal{G} = \pi/4$.

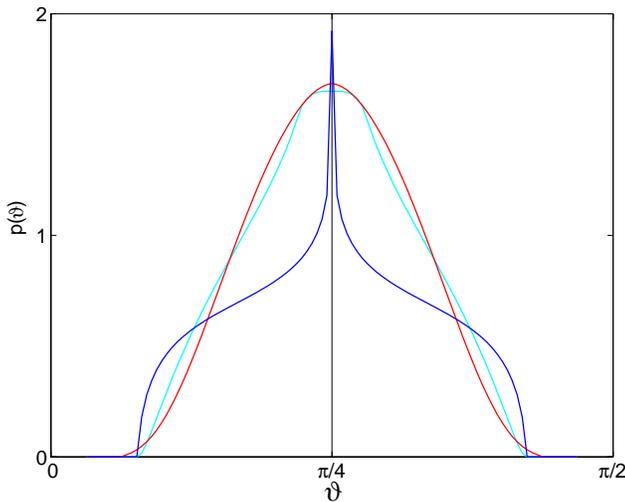


Figure 5. Distribution function $p(\mathcal{G})$ for three different $p(a)$ for the ellipses with $\psi = \pi/2$. For the red line, a is sampled from the Γ function defined in Equation 14 with $\sigma = 0.5$. For the cyan line, a is sampled from the triangular function defined in Equation 15 with $a_0 = 1.2$ and $a_{\max} = 4$. For the blue line, a is sampled from the uniform distribution defined in Equation 16 with $a_{\max} = 4$. The black line indicates the position of $\mathcal{G} = \pi/4$.

Results similar to those for $\psi = \pi/2$ can be obtained for other values of ψ . In the general case, however, it is tedious to derive an analytical expression for $p(\mathcal{G})$. We demonstrate the results, therefore, with computer simulations. Through a derivation similar to the one for the case of $\psi = \pi/2$ we find

$$\theta = \tan^{-1} \left[\frac{A^2 B + A a^2 \tan \zeta (\tan \psi - \tan \zeta) - A \tan \psi \tan \zeta - A}{B(a^2 - 1) \tan \zeta + A a^2 (\tan \psi - \tan \zeta) + A \tan \zeta (1 + \tan \zeta \tan \psi)} \right] \quad (17)$$

where

$$A \equiv \sqrt{a^2 \tan^2 \zeta + 1},$$

$$B \equiv \sqrt{a^2 (\tan \psi - \tan \zeta)^2 + \tan \zeta \tan \psi + 1}$$

We obtain $p(\mathcal{G})$ by sampling ζ uniformly from $(0, \pi)$ and by sampling a from the three distributions used for the case of $\psi = \pi/2$. For each sampled pair of ζ and a , we calculate \mathcal{G} using Equation 17. By constructing the histogram of the resulting values of \mathcal{G} , we get $p(\mathcal{G})$. The results for three values of ψ are shown in Figure 6 ($\psi = \pi/8$), Figure 7 ($\psi = \pi/4$), and Figure 8 ($\psi = 3\pi/8$). The number of sampling points is 500,000. As can be seen in these figures, the distributions are always peaked at $\mathcal{G}^* = \psi/2$, in exact agreement with the cocircular statistics observed by Sigman et al. (2001).

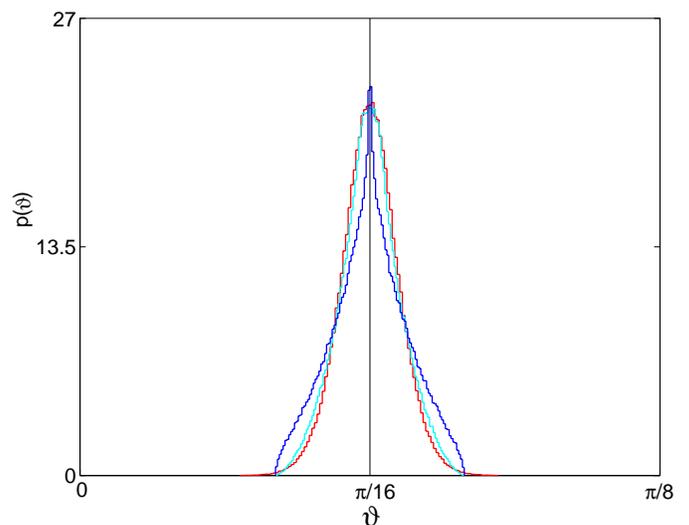


Figure 6. The same as in Figure 5, except that $\psi = \pi/8$. The black line indicates the position of $\mathcal{G} = \pi/16$.

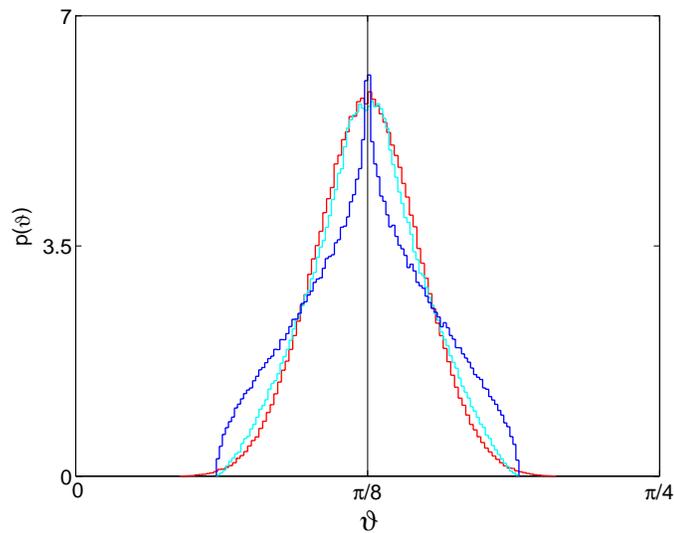


Figure 7. The same as in Figure 5, except that $\psi = \pi/4$. The black line indicates the position of $\vartheta = \pi/8$.

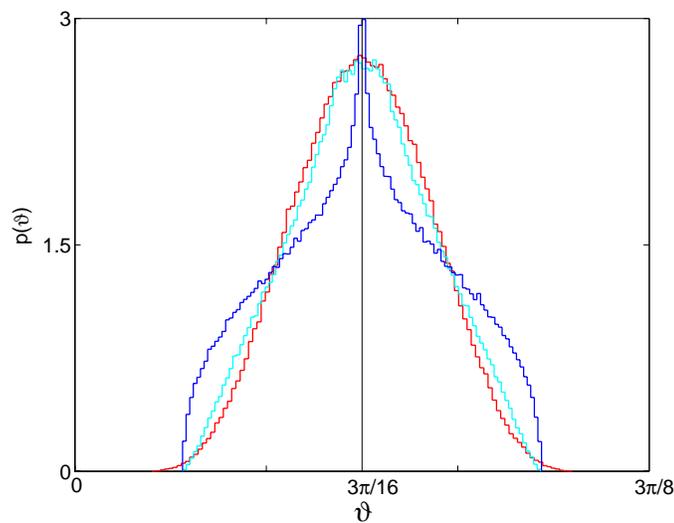


Figure 8. The same as in Figure 5, except that $\psi = 3\pi/8$. The black line indicates the position of $\vartheta = 3\pi/16$.

3 Implications

The authors of Sigman et al. (2001) discuss their findings in relation to properties of the visual system. They also relate them to psychological notions, such as the Gestalt idea of good continuation (Kofka, 1935; Wertheimer, 1938). These implications are not affected by our remark. We were intrigued by the thought that circles were a significant structure in natural scenes, independently of how our brain might strive to simplify them. We find that closed smooth objects will contribute to $p(\mathcal{G})$ in a restricted range around the value for pure circles, and that even a small number of nearly circular objects can give statistics indicating cocircularity. We propose that the results of Sigman et al. (2001) may simply indicate that there are many closed smooth contours in natural visual scenes.

Acknowledgments

We thank the gracious hospitality of the Institute for Theoretical Physics, where a seminar by C. Gilbert inspired this work. We also wish to thank D. Kleinfeld for his encouragement and his stimulating questions. This work is partially supported by the National Institutes of Health (C.C.), the A.P. Sloan Foundation (C.C.), and the Howard Hughes Medical Institute (D.Z.J.).

Commercial relationships: none.

References

- Kofka, K. (1935). *Principles of gestalt psychology*. New York: Harcourt & Brace.
- Sigman, M., Cecchi, G. A., Gilbert, C. D., & Magnasco, M. O. (2001). On a common circle: Natural scenes and Gestalt rules. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 1935-1940. [PubMed]
- Wertheimer, M. (1938). *Laws of organization in perceptual forms*. London: Harcourt, Brace & Jovanovitch.