Visual Neuroscience (1991), **6**, 95–111. Printed in the USA.

The density recovery profile: A method for the analysis of points in the plane applicable to retinal studies

R. W. RODIECK

Department of Ophthalmology, The University of Washington, Seattle

(RECEIVED April 24, 1990; ACCEPTED September 17, 1990)

Abstract

The density recovery profile is a plot of the spatial density of a set of points as a function of the distance of each of those points from all the others. It is based upon a two-dimensional point autocorrelogram. If the points are randomly distributed, then the profile is flat, with a value equal to the mean spatial density. Thus, any deviation from this value indicates that the presence of the object represented by the point alters the probability of encountering nearby objects of the same set. Increased value near an object indicates clustering, decreased value near an object indicates anticlustering. The method appears to be unique in its ability to provide quantitative measures of the anticlustered state. Two examples are presented. The first is based upon a sample of the distribution of the somata of starburst amacrine cells in the macaque retina; the second is based upon the distribution of the terminal enlargements on the dendrites of a single macaque ganglion cell that projects to the superior colliculus. In both cases, the density recovery profile is initially lower than the mean density, and increases up to the plateau at the value of the mean density. Two useful measures can be derived from this profile: an intensive parameter termed the *effective radius*, which quantifies the extent of the region of decreased probability and is insensitive to random undersampling of the underlying distribution, and an extensive parameter termed the *packing factor*, which quantifies the degree of packing possible for a given effective radius, and is insensitive to scaling. An extension of this method, applicable to correlations between two superimposed distributions, and based upon a two-dimensional point cross-correlogram, is also described.

Keywords: Spatial distribution, Territorial domains, Nearest-neighbor analysis, Undersampling, Anticlustering

Introduction

Most of the cell types of the vertebrate retina are not randomly distributed but show some degree of spatial organization (e.g. Wässle & Riemann, 1978). Characteristically, near a given cell, there is a *decreased* probability of encountering another cell of the same type (anticlustering). However, many methods that investigate the distribution of neighboring points are designed primarily for characterizing the clustering of points (i.e. where there is an *increase* in the probability that a second point will lie near a given point; e.g. Ripley, 1981; Rose & Grimson, 1988).

One of these methods, introduced by Clark and Evans (1954) and based upon a histogram of nearest neighbors, has nevertheless been widely and effectively used to analyze the spatial distribution of various types of retinal cells (e.g. Wässle & Riemann, 1978; Vaney et al., 1981; Voigt, 1986). The observed distribution of nearest neighbors is compared to the theoretical curve for a random (Poisson) process, which leads to a test for significance. A weakness of this method is that only the nearest neighbor of a given cell is recorded, although there are related methods that make use of all neighbors. A second difficulty is that the resulting histogram provides only an indirect indication of the distribution of other cells in relation to a given cell.

Given data in which there is a decreased probability of encountering nearby cells, the possibility arises that there may be another approach in which this decreased probability might be more simply represented. If so, then it might be possible to find a useful measure for the effective "dead space" about a given cell. This paper describes one such approach, perhaps the simplest possible. It leads naturally to a measure of the dead space, termed the *effective radius*, which, somewhat remarkably, is insensitive to random undersampling. An additional measure, termed the *packing factor*, proves insensitive to scaling, and for some cell types at least, appears to change little over large variations in spatial density.

This method can demonstrate either the clustering or anticlustering of points in a plane, but does not address the issue of *why* the distribution of retinal cells or the distribution of the terminals of a single cell deviate from randomness (a forthcoming paper will present a specific hypothesis for the observed dis-

Reprint requests to: R. W. Rodieck, Department of Ophthalmology, RJ-10, The University of Washington, Seattle, WA 98195, USA.

tribution of the somata of starburst amacrines). Instead, this method is concerned with *how* to express such deviation, and how resulting measures of it can be interpreted.

Description of method

Figure 1A shows the distribution of a group of labeled cells in the ganglion cell layer that lay within a 1-mm-wide strip of a wholemounted macaque retina. These data came from a study of cells labeled by means of an antibody to choline-acetyltransferase (ChAT) (Rodieck & Marshak, 1989). A 1-mm^2 portion of this strip is shown in Fig. 1B. In the example to be presented, the aim is to characterize the relation between the presence of a point and the occurrence and distribution of other points that lie within 200 μm of it. The dotted rectangle is inset by 200 μm from the bounds so every point within it is surrounded by other points in all directions out to at least this distance. The inset region has an area of 0.36 mm^2, and contains 102 cells, so that the average density within it is 283 cells mm^{-2}. One of these cells is marked with a plus sign, and a circle having a radius of 200 μm has been drawn around it. Within this circle are 35 other cells, lying at varying distances and directions with respect to the reference cell.

The basic idea is to superimpose the distribution of other cells using every cell in the region as a reference cell. Imagine placing a reference point somewhere near the middle of a piece of translucent paper. Place the paper on the diagram, align the reference point with the plus sign, and mark the positions of the 35 other points in the diagram that lie within the 200-μm circle. The result is show in Fig. 1C.

In the upper right of the dotted rectangle is another labeled cell, again highlighted by means of a plus sign. Shift the translucent paper to align its reference point with the point that represents this cell, and again mark on this paper the positions of all of the other points that lie within 200 μm of the mark. The result is shown in Fig. 1D. Repeat this process for each of the remaining 100 cells within the dotted rectangle. The result, termed the *spatial autocorrelogram*, is shown enlarged in Fig. 2. Every circular symbol shows the average size of the somata of these cells at this retinal eccentricity; the superimposed reference cells are shown as a single black circle, and all those offset from each of the reference cells are shown in gray. By construction, this diagram shows the combined offsets from every cell in the dotted rectangle to every other cell that lies within 200 μm of it. There is a partial radial symmetry to this diagram, produced by those pairs of points for which each point lies within the dotted rectangle (e.g. if point *a* is to the left of point *b*, then *b* is to the right of point *a* by the same distance).

If the distribution were entirely random, then the occurrence of a cell should not alter the probability of occurrence of those away from it and thus the average density should be the same at all offsets. However, visual inspection suggests that there may be fewer cells near the reference cell than at larger offsets, and we now consider how such a deviation from randomness can be quantified.

Figure 3A is a replot of Fig. 2, with the somata now represented by points. There is also a series of circles, drawn about the reference point and separated by 10 μm, which defines a series of annuli. The expected number of points in each annulus is easily calculated. Let

A = area of the region containing the reference points (i.e. the dotted rectangle in Fig. 1B);
N = number of points in A;

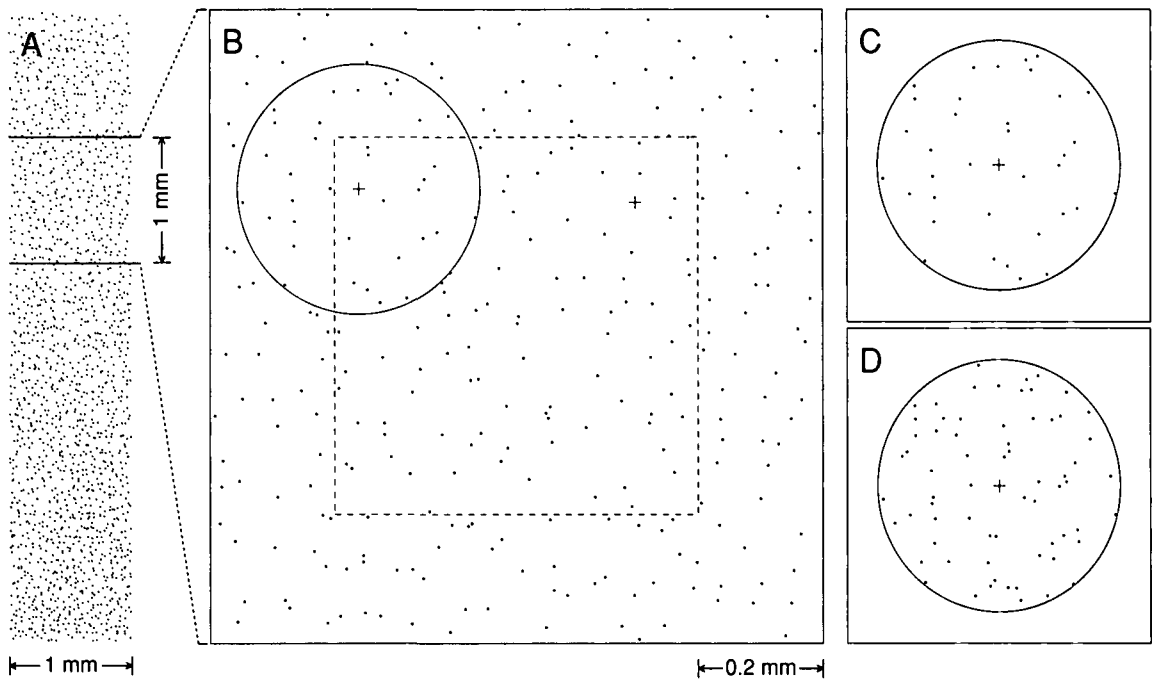


Fig. 1. Diagram to illustrate the construction of a spatial autocorrelogram. A: 1-mm-wide strip of ChAT-labeled cells from a macaque retina. Data from Rodieck and Marshak (1989). B: 1-mm^2 portion of this strip. Radius of circle = 200 μm. C: Offsets about one of the cells in the dotted rectangle, marked with a plus symbol in the upper left region of the dotted rectangle. D: Superimposed offsets about two of the cells in the dotted rectangle, each marked with a plus.

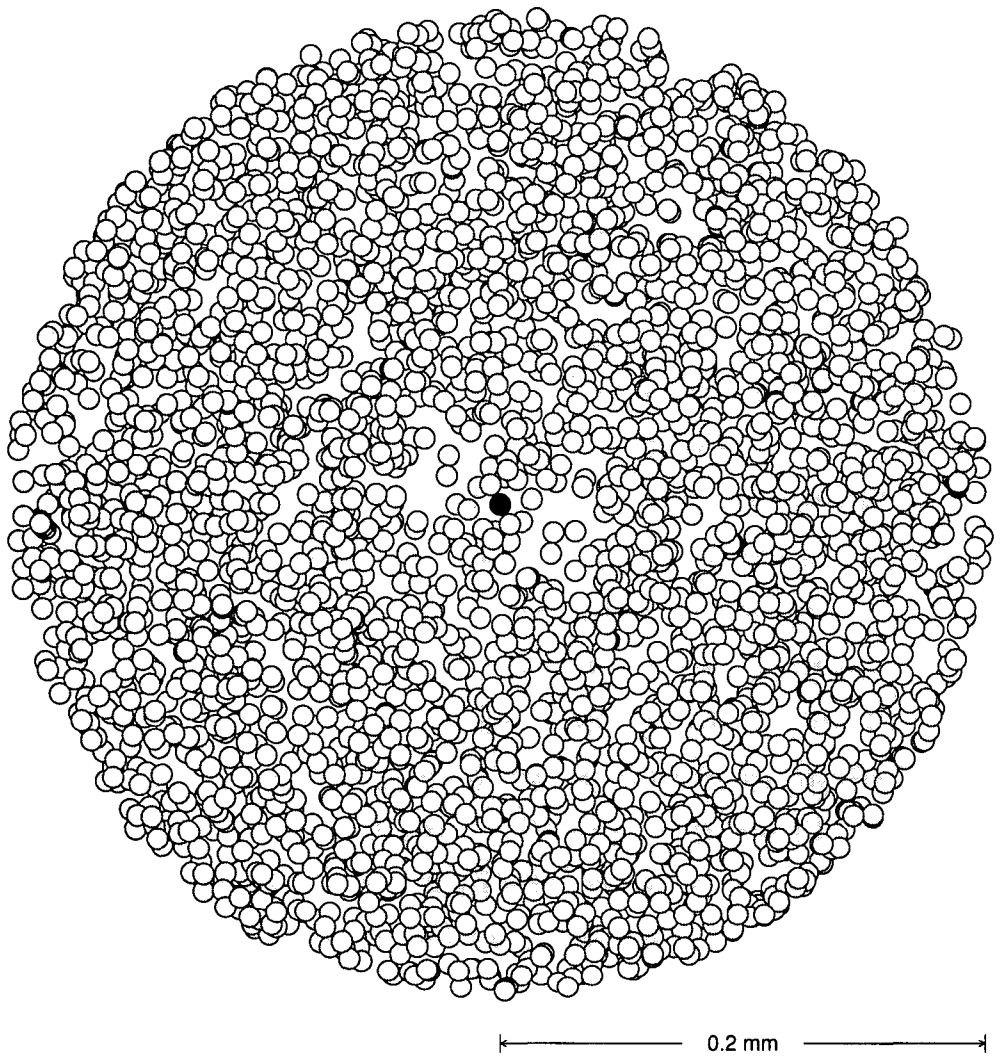


Fig. 2. Spatial autocorrelogram for the points lying within the dotted rectangle in Fig. 1B. The location of the reference point is indicated by a black circle whose size matches that of the somata of the labeled cells. The gray symbols show the offsets of all of the other cells from each reference cell that lies within a 200-μm radius of that reference cell.

D = density of points in the region, equal to N/A;
i = index to each annulus: 1,2,... (or the corresponding bin in the *density recovery profile*, as discussed below);
ΔA_i = area of annulus i;
Δr = width of each annulus (and bin width of the density recovery profile, described below);
n_i = measured number of points in annulus i;
λ_i = expected number of points in annulus i, assuming a random (Poisson) distribution; and
d_i = density measure for the density recovery profile, defined below.

The area of an annulus is easily shown to be

$$\Delta A_i = \pi \Delta r^2 (2i - 1). \tag{1}$$

The expected number of points in an annulus is the product of the expected number to be found in the annulus for each reference point times the number of reference points N. Assuming a random distribution of points in the region of interest, the expected number in the annulus for each reference point is simply the density of the points D times the area of the annulus ΔA_i:

$$\lambda_i = ND\Delta A_i. \tag{2}$$

Let

$$d_i = \frac{n_i}{N\Delta A_i}. \tag{3}$$

The expected value of this measure for a random distribution is then

$$d_{i,\text{random}} = \frac{\lambda_i}{N\Delta A_i} = D, \tag{4}$$

and thus independent of the bin number.

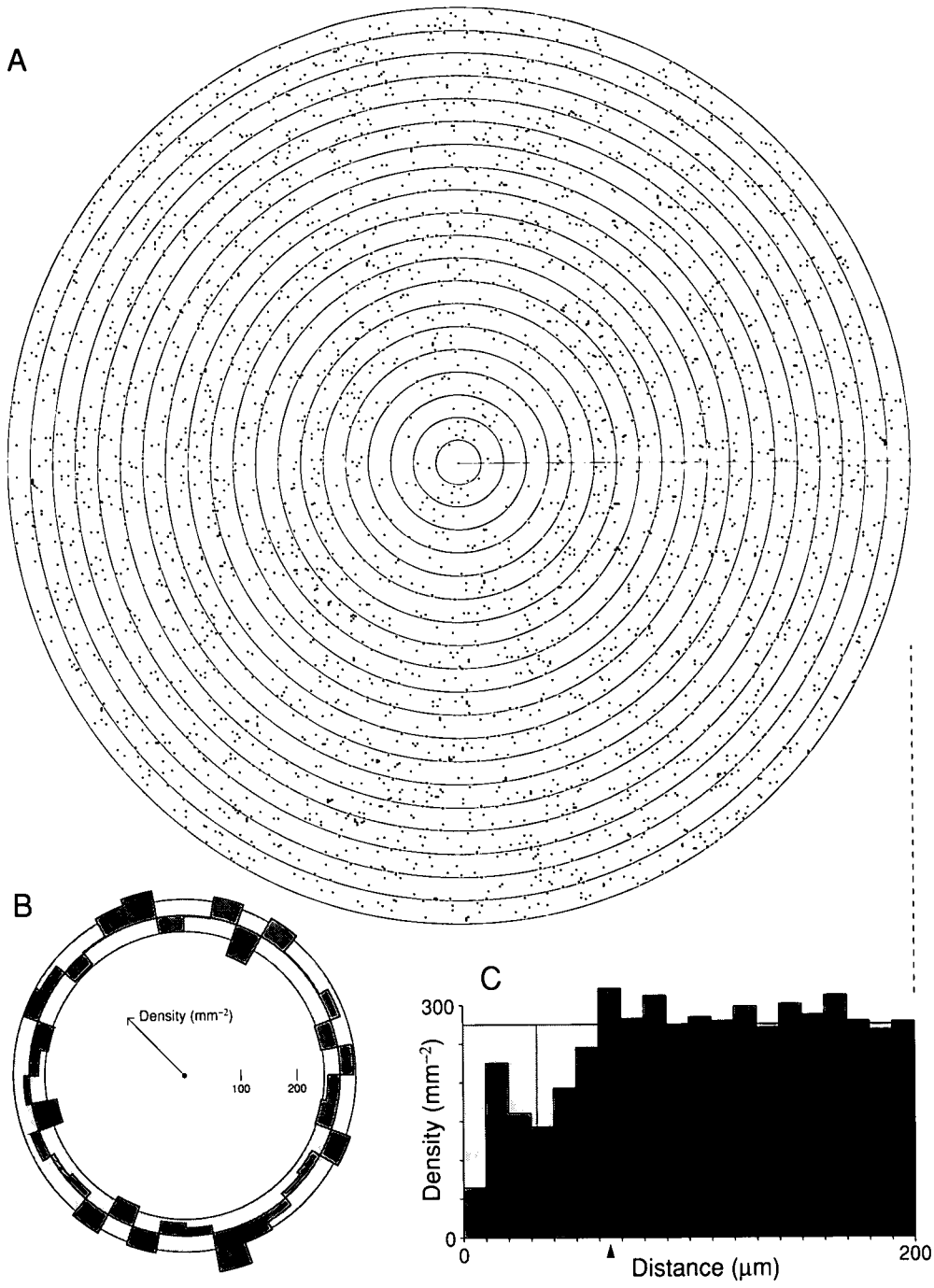


Fig. 3. A: Same as Fig. 2, with cells now represented by points, and a series of annuli of width 10 μm drawn about the reference point. B: Density of points in each 10-deg sector of the upper diagram, shown as a radial plot, whose axis (indicated by arrow to upper left) is spatial density. The tick marks to the right of the center are in increments of 100/mm^2. The mean density (283 cells/mm^2) appears as a circle superimposed upon the bounds defined by $\pm$ the standard deviation for Poisson statistics, shown as a light gray annulus. The region between the density of each sector and the mean density is shown as medium gray. The observations that the sector densities lie without apparent pattern on each side of the mean, and that only about a third of them extend beyond the standard deviation annulus are consistent with the hypothesis that there is no significant orientation of the distribution of points with respect to some axis in the plane. C: Density recovery profile, aligned to show the relation between each bin of this profile, and the corresponding annulus of the spatial autocorrelogram. The thin vertical line shows the effective radius, and the arrowhead shows the maximum radius; these terms are defined in the text.

Figure 3C plots the density measure d_i, for the successive annuli, and provides a compact means of showing the relation between the presence of a cell and the radial distribution of the cells that lie near it. This construction will be termed the *density recovery profile*. The term *recovery* is meant to express the return to the mean spatial density as distance increases from the reference point. The light gray zone shows the mean density. The annular densities, in darker gray, initially lie well below this

level, cross it at about 60–70 μm, and continue at an average level slightly above the mean level.

Object size

The somata considered here are shown in Fig. 2 as circles that approximate their shape and size, in order to make evident the fact that, when the somata lie in a single layer, there is a minimum distance between their centers. The somata of ChAT-labelled cells in this sample are approximately spherical and of uniform size, so this minimum distance is equal to their diameter (about 8 μm). Thus, a reduced probability over this distance may indicate nothing more than the fact that two cells cannot occupy the same space. The observed decrease in probability, however, extends well beyond this value, and thus cannot be so simply accounted for. A measure of the extent of this zone of decreased probability, termed the *effective radius*, is described below.

Orientation

The above construction is valid whether or not the distribution of points is oriented in the plane. However, certain inferences and measures that can be derived from the density recovery profile require that the distribution is radially symmetric (i.e. isotropic). Issues of orientation in the general sense of this term are somewhat involved, and will be taken up in the Discussion. However, the presence of global orientation, meaning the same orientation factors are applied to all points in the same way, is simple to test over the region of the data sample. Consider first a random (Poisson) distribution of points on a rubber sheet. Imagine this sheet stretched horizontally. It is simple to show that the density recovery profile will still be flat, and the autocorrelogram will still be radially symmetric, since the local spatial density, although reduced, remains everywhere constant. Thus, for purely random distributions, anisotropic stretching alters the spatial density of points, but does not disturb the radial symmetry. However, if the points are nonrandomly distributed so that in the autocorrelogram there is a "hill" (clustering) or "valley" (anticlustering) about the reference cell, then stretching will alter that shape, and radial symmetry will be lost. In effect, global orientation factors will make the density of cells offset from a reference cell dependent upon the *angle* of the offset.

The diagram in Fig. 3B shows the spatial density of offsets from a reference cell, plotted for each 10-deg sector of the data in the upper diagram (see legend for details). There is no systematic variation in density with angle, indicating that if global orientation over the sample region is present, then it is too small to have much influence upon the measures to be discussed below.

A second example

Figure 4A is a drawing of a ganglion cell, termed a maze cell; it is taken from a study on the projection of ganglion cell types to different regions of the macaque brain (Rodieck & Watanabe, 1988). Figure 4B (upper) is a photomicrograph that shows, at higher magnification, a portion of the dendritic tree of this cell. The dendrites of maze cells give off spines that terminate in small swellings, indicated in Fig. 4B (lower) by means of circles. Figure 4C shows the spatial distribution of these swellings. Figure 4D shows the density recovery profile for this distribution. Just as there can be a decreased probability of the occurrence of another cell near one of the same type, so too can there be a decreased probability of finding another swelling near one of the same cell. Although the scale is different, the method is the same. Figure 4D also illustrates the distribution of the nearest neighbor to each swelling in the sample.

Effective radius

Although the shape of the density recovery profile provides some insight into the manner in which the distribution deviates from randomness, there exists a useful scalar measure, based upon it, that quantifies one's intuitive notion of a "dead space" about each cell.

Consider again the density recovery profile for ChAT-labeled cells, shown in Fig. 3C. Seen in three dimensions, and relative to the mean density, there is a dip in average spatial density about the reference point that extends about 60 μm from the reference point. Although it is useful to think of this dip as a *volume* in this geometric space, the measure itself has *number* (#) as its dimension, since the third axis *density* has the dimensions of #/area. This "volume" is readily calculated from the data in Fig. 3. Let ΔV_i be the volume of the dip for annulus i and V_e the total volume of the dip, obtained by adding the volumes of the annuli until the density value for an annulus exceeds the mean density (i.e. until the calculated volume for the annulus becomes negative). Now the volume of the dip for a given annulus is simply the area of the annulus times its height:

$$\Delta V_i = \Delta A_i (D - d_i). \tag{5}$$

Substituting from eqns. (2) and (3),

$$\Delta V_i = \frac{\lambda_i - n_i}{N}, \tag{6}$$

so that

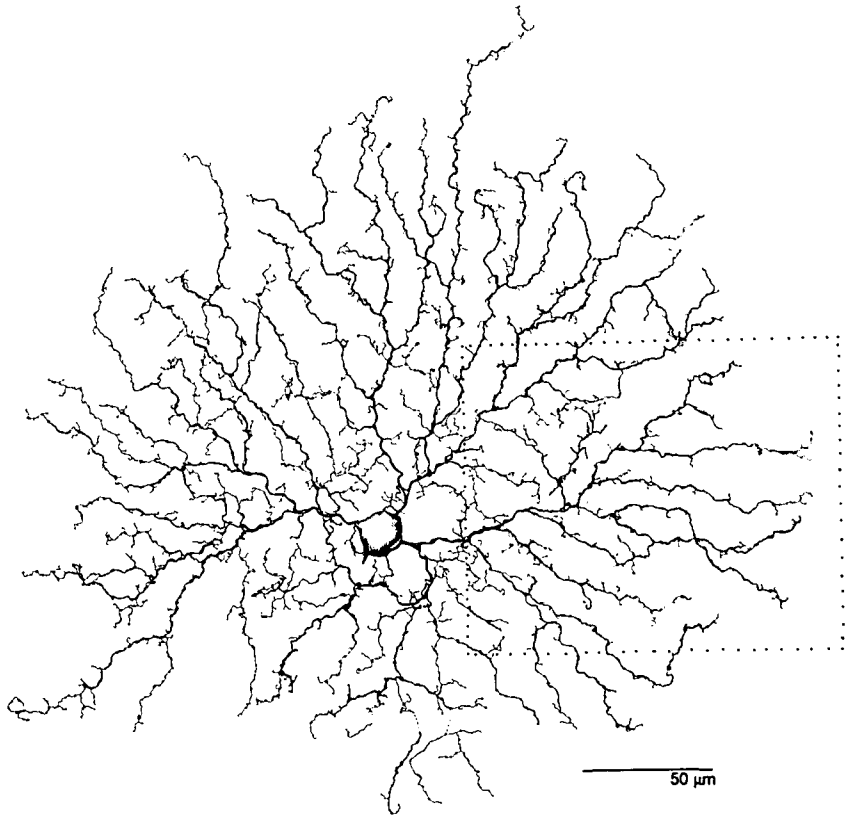
$$V_e = \frac{1}{N} \sum_{i=1}^{\text{until } n_i > \lambda_i} (\lambda_i - n_i). \tag{7}$$

In this case, $V_e = 0.95$ points/reference point. In itself, V_e is a useful measure of the magnitude of the average "dead space" about each point. However, just as dendritic-field diameter is an easier measure to visualize than dendritic-field area, a linear measure proves easier to visualize than this volumetric measure, which also depends as well on mean density. It is therefore convenient to reconfigure this volume into a cylinder having a height equal to the mean density and a radius calculated to give the cylinder the same volume as that of the dip. This calculated value will be termed the *effective radius*. The total decrease in cells within the dip is then made equivalent to one composed of a step change from zero to the mean density at the effective radius. Let r_e be the effective radius. Then, by definition,

$$V_e = \pi r_e^2 D, \tag{8}$$

so that

$$r_e = \sqrt{\frac{V_e}{\pi D}}. \tag{9}$$



A

Fig. 4. A: Drawing of a maze cell; data from Rodieck and Watanabe (1988). Dotted line indicates the area of the photomicrographs in part B. (*Figure continued on the facing page.*)

In this example, the effective radius is 32.7 μm, and is shown as a thin vertical line in the density profiles shown in Figs. 3C and 4. As discussed below, the reliability of this measure depends ultimately upon the expected number of counts per bin, where the distribution is random. Thus, a higher mean density, a larger sample area, and a wider bin width all improve the reliability of this measure, but do not directly influence its value.

Reliability

Too few points, or too narrow bins, can produce an unreliable estimate of effective radius—typically an underestimate, and thus a bias. The underlying reason for this potential bias is that the bins at the start of the profile contain fewer counts, and thus greater variability. This section analyzes this situation from a quantitative perspective, and describes a method for choosing the sample area and the bin width that will generate an estimate of the effective radius to a selected level of reliability and variance. For a random (Poisson) distribution, the variance σ^2 in the number of points in a given area is equal to the mean number λ:

$$\sigma^2 = \lambda.$$

Substituting from eqn. (2),

$$\sigma^2 = ND\Delta A.$$

The Poisson distribution approaches the binomial distribution as λ increases; assuming that to be the case,

$$\sigma = \sqrt{ND\Delta A}.$$

This is the standard deviation about the mean number of points in a bin. In order to plot it in the density recovery profile, we must multiply by the scale factor for each bin, as given in eqn. (3):

$$\sigma_i' = \frac{\sqrt{ND\Delta A_i}}{N\Delta A_i} = \frac{1}{\sqrt{A\Delta A_i}}. \tag{10}$$

Substituting for ΔA_i, from eqn. (1):

$$\sigma_i = \frac{1}{\sqrt{A\pi\Delta r^2(2i-1)}}.$$

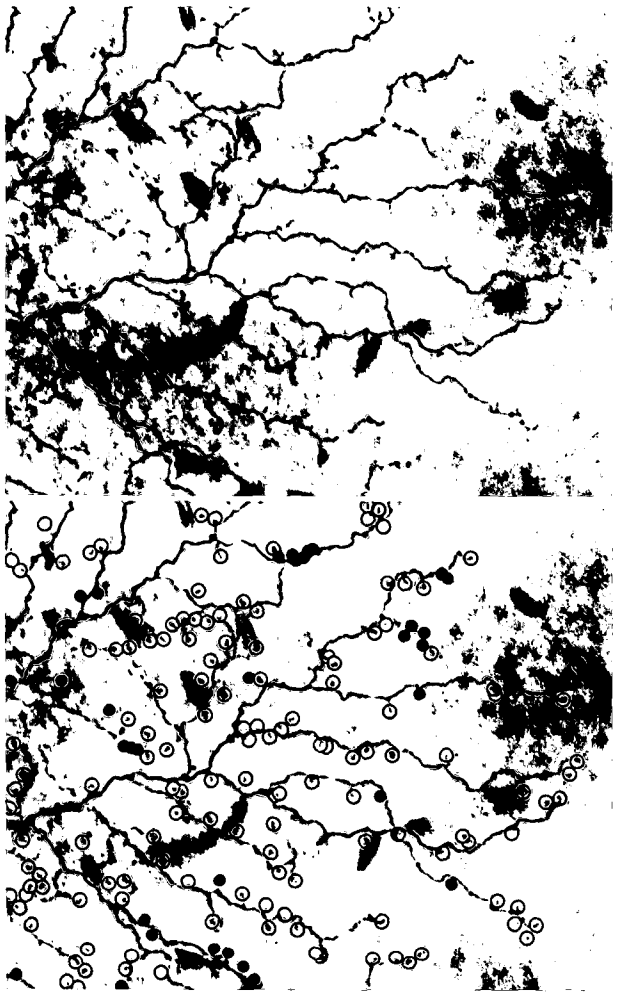
Let

$$D_c = \frac{1}{\sqrt{A\pi\Delta r^2}}, \tag{11}$$

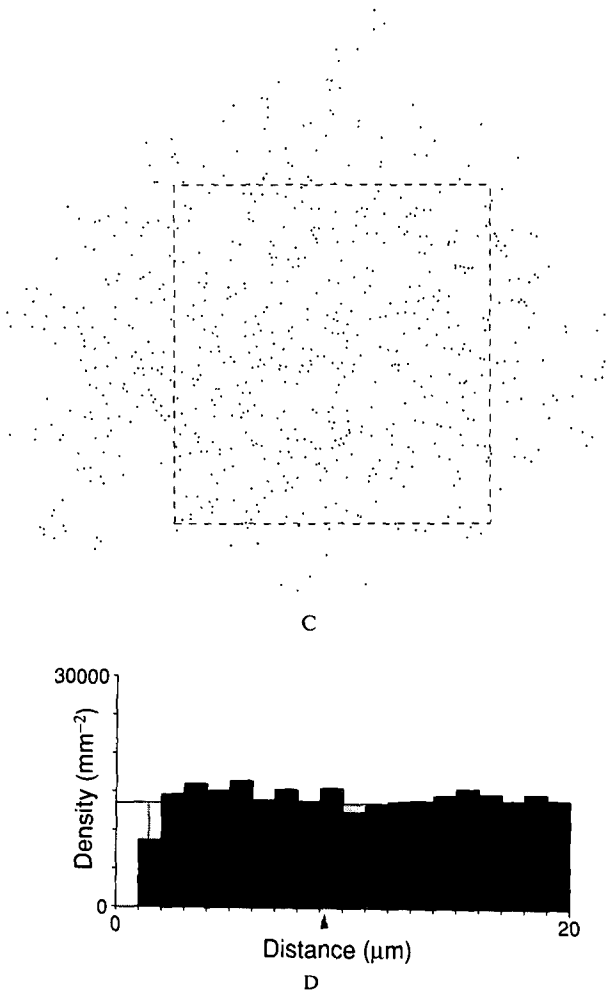
then

$$\sigma_i' = \frac{D_c}{\sqrt{2i-1}}. \tag{12}$$

The value D_c depends only on the sample area A and the bin width Δr. It has the dimensions of spatial density and is termed the *critical density*. For the sample area shown as the dotted rectangle in Fig. 1 (0.36 mm^2) and for a bin width of 10 μm (e.g. Fig. 3), D_c is equal to 94.0 mm^{-2}. Figure 5 shows the scaled standard deviation σ_i' plotted against bin number, about the



B



C D

Fig. 4 continued. B: Duplicate photomicrographs of spines and swellings. In the lower one the swellings are highlighted with circles—open if the swelling is in focus and filled if it is out of the plane of focus of the micrograph. C: Spatial distribution of swellings, together with a dashed rectangle that indicates the data used for analysis. Note that this is a different rectangle than that shown in part A. D: Density recovery profile of swellings. Only the points that lay within the dashed rectangle of part C were included in the analysis, and a compensation factor was applied to allow all points to be included in the analysis, as discussed in the text. As in Fig. 3C, the light gray shows the mean density, and the medium gray shows the density recovery profile. The dark gray profile shows the fraction of the counts that were also the nearest neighbor to the reference cell. When the distances between points approach the resolution of the system (in this case the microscope and camera lucida), there exist two possible ways of generating an artifactual dip in the density recovery profile. The first is by failing to resolve two nearby objects; the second is by recording their positions with slightly larger separations, so that the points remain resolvable on the recording medium (e.g. pencil marks on a paper). In both cases, it is important to recognize that the number of counts in each bin of the density recovery profile depends upon the bin number, since annuli near the reference cell have smaller areas than those farther away. In effect, small changes in the counts in the first few bins have much larger effects than similar changes to the counts of subsequent bins. Thus, failing to resolve two nearby objects can have a profound effect on the shape of the density recovery profile. If positions are recorded with larger separations, then this will have no effect if the increase does not put the offset value into the next bin. But if it does, then the resulting decrease in the height of the previous bin will be larger than the increase made to the height of the next bin (this will change the shape of the density recovery profile, but will not alter a measure of the magnitude of the dip). In the case of the density recovery profile shown here, the swellings on the spines were dark, small ($\leq$0.4 μm), and there was a clear gap between all nearest neighbors. The same dip was seen for the three maze cells tested. Consequently, we are confident that the dip is real.

mean value for each bin, assuming a random distribution of points having a spatial density D of 283 cells mm^{-2} (i.e. the value for the points shown in Fig. 1A). The standard deviation decreases with bin number, as expected, starting from a value of D_c for the first bin.

Although the standard deviation bars are shown about the mean density D, their amplitude is independent of it, since this amplitude depends only upon the sample area and the bin width. In other words, if D were made smaller, the standard deviation bars would drop, but would not change in size (eventually the assumption of a binomial distribution would be violated). Clearly, if D is large compared to D_c, the sample es-

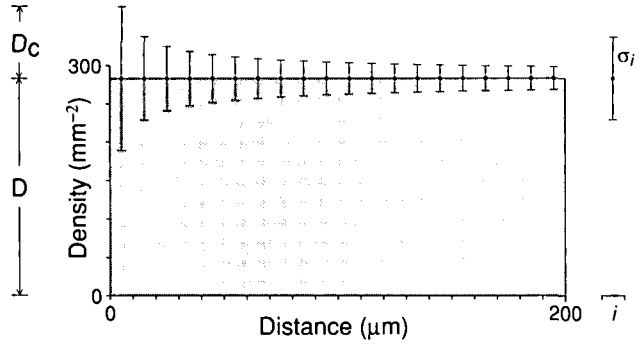


Fig. 5. Standard deviation bars above and below the mean spatial density (horizontal line) for the amplitudes of the bins in a density recovery profile, based upon a random distribution. Sample area equal to that of the dotted rectangle in Fig. 1A, a bin width of 10 μm, and a spatial density of 283 mm^{-2}. D: mean density, D_c: critical density [eqn. (11)]; and σ_i: scaled standard deviation for bin i.

timates for each bin will be good, and the estimate of the effective radius will be reliable. Thus, a practical measure of the robustness of this measure is simply

$$k = \frac{D}{D_c}. \tag{13}$$

In the case considered, it has a value of 283/94 or 3.0. The expected number of points in the first bin of the density recovery profile, assuming a random distribution, is the square of this value [substitutions from eqns. (1), (2), (11), and (13)]:

$$\lambda_1 = ND\Delta A_1 = ND\pi\Delta r^2 = D^2 A\pi\Delta r^2 = \frac{D^2}{D_c^2} = k^2. \tag{14}$$

It is convenient to consider the sample area consisting of a square of side s:

$$s = \sqrt{A}. \tag{15}$$

Substituting in the above equation, and rearranging terms yields

$$s\Delta r = \frac{k}{D\sqrt{\pi}}. \tag{16}$$

The observed spatial density D and a chosen reliability factor k define a minimum value for $s\Delta r$. One can thus calculate the minimal sample area, given the bin width, or the minimal bin width, given the sample area. For example, for a reliability factor of 4, a spatial density of 283 mm^{-2}, and an area of 1 mm^2, a bin width of not less than 8 μm should be used.

The section entitled *Statistical Testing* below describes how to determine the variation in effective radius as a function of the reliability factor.

Packing factor

Although the effective radius is a useful measure, its maximum possible value is constrained by the observed spatial density. This limit proves to depend upon how tightly a set of points with a given effective radius can be packed. Consideration of this issue leads to a dimensionless measure, termed the *packing factor*, that ranges from 0 for a random distribution to 1 for a packed hexagonal array of uniform disks (i.e. maximal packing of nonoverlapping disks). As discussed in the next section, the packing factor proves to be invariant to scaling.

A random distribution will show a density profile in which each bin will, on average, have a density measure equal to the average density; thus the effective radius will be zero if statistical fluctuations are ignored. The maximum radius r_m possible for a given density is found by determining the spacing of a maximally packed array having the same average density D. Maximum packing in two dimensions is achieved by hexagonal packing, where the area per point a is

$$a = \sqrt{\frac{3}{4}}\, r_m^2. \tag{17}$$

Since

$$D = \frac{1}{a}, \tag{18}$$

we have

$$r_m^2 = \frac{\sqrt{\frac{4}{3}}}{D}. \tag{19}$$

The volume of the maximum decrease V_m is a constant:

$$V_m = \pi r_m^2 D = \pi\sqrt{\frac{4}{3}} = 3.6 \text{ points/reference point}. \tag{20}$$

In this example, with a density of 283 cells/mm^2, the spacing between adjacent cells arranged in a hexagonal array is 63.9 μm. This point is shown as an arrow in the density recovery profile of Fig. 3C. The packing factor p is defined to be the square of the ratio of the effective radius to the maximum radius:

$$p = \left(\frac{r_e}{r_m}\right)^2. \tag{21}$$

Since the effective radius can range from 0 to r_m, the packing factor ranges from 0 to 1. In this example;

$$p = (32.7/63.9)^2 = 0.26.$$

The packing factor is also equal to the ratio of the volume of the observed dip V_e to the volume of the maximum dip V_m:

$$\frac{V_e}{V_m} = \frac{\pi r_e^2}{\pi r_m^2}\frac{D}{D} = \frac{r_e^2}{r_m^2} = p. \tag{22}$$

Thus, in this example, using the value of V_e calculated earlier;

$$p = 0.95/3.6 = 0.26.$$

Consider the packing of nonoverlapping disks of a given size. In this context, the packing factor also gives the fraction of the

maximum number of disks of that size that could be maximally (i.e. hexagonally) packed. Thus, in the above example, it would have been possible to have placed about four times more cells in the sample region, without their "dead spaces" overlapping, although they have to be arranged in an ordered hexagonal array to do so. If the packing factor is equal to 1, then the points are necessarily distributed in a hexagonal array.

It is easy to show that a square array has a packing factor of 0.87; thus, neglecting side effects, a box containing a single layer of 87 balls packed in a square array could hold 100 such balls if they were packed hexagonally. This example serves to illustrate that although a large packing factor implies some degree of regularity, the packing factor, as its name implies, is a measure of packing, not of regularity.

Scaling alters the effective radius, but does not alter the packing factor

Imagine placing a photographic negative of the distribution of points such as shown in Fig. 1B in a photographic enlarger, so that the linear scale is increased by a factor of 2. This will increase the area of any region by a factor of four. Since the number of points in the region remains unchanged the density of points will be reduced by four. Since the distance between every pair of points will be increased by a factor of 2, the effective radius must necessarily double. For the general case let

s = scaling factor;
D' = scaled density for some region;
r_e' = scaled effective radius; and
V_e' = scaled volume of the dip in density.

We have

$$r_e' = sr_e,$$

$$D' = \frac{D}{s^2},$$

so that

$$V_e' = \pi(r_e')^2 D' = \pi r_e^2 D = V_e. \tag{23}$$

Thus, the volume of the dip is not altered by scaling. Since V_m is a constant, it follows that V_e/V_m remains unaltered, and consequently, so does the packing factor, which, as shown above, is equal to this ratio. The invariance of the packing factor to scaling is consistent with the fact that it is a dimensionless measure.

Random undersampling reduces the packing factor, but does not alter the effective radius

Let us assume that the positions of only a fraction of the cells in some region were determined, using some method that randomized which of the cells would be included. Consider the annuli shown in Fig. 3A. On average, the number of cells in each annulus will be reduced by this fraction. As the amplitude of the bin in the density recovery profile is simply the number of points in the corresponding annulus divided by the area of the annulus, the amplitude of each bin will consequently be reduced by this fraction. The *shape* of the profile thus remains unchanged, only the vertical scale is affected by the undersampling. Let

γ = undersampling factor, as a fraction of full sampling;
D' = measured density for some undersampled region;
n_i' = measured number of points in annulus i for the undersampled region;
$\Delta V_i'$ = volume of the dip in annulus i for the undersampled region;
V_e' = measured volume of the dip in density for the undersampled region;
r_e' = calculated effective radius for the undersampled region; and
r_m' = calculated maximum radius for the undersampled region.

Now the undersampling affects both the measured density and the number of points collected in each annulus:

$$D' = \gamma D,$$

$$n_i' = \gamma n.$$

Therefore, from eqns. (3) and (5),

$$\Delta V_i' = \Delta A_i(D' - d_i') = \Delta A_i\left(D' - \frac{n_i'}{N\Delta A_i}\right)$$

$$= \gamma \Delta A_i\left(D - \frac{n_i}{N\Delta A_i}\right) = \gamma \Delta V_i.$$

Since the total volume is simply the sum of all of the annular volumes,

$$V_e' = \gamma V_e.$$

Hence, from eqn. (9):

$$r_e' = \sqrt{\frac{V_e'}{\pi D'}} = \sqrt{\frac{\gamma V_e}{\pi \gamma D}} = \sqrt{\frac{V_e}{\pi D}} = r_e. \tag{24}$$

The effective radius is thus insensitive to undersampling. The maximum radius, however, depends upon the mean density:

$$r_m' = \sqrt{\frac{\sqrt{\frac{4}{3}}}{D'}} = \sqrt{\frac{\sqrt{\frac{4}{3}}}{\gamma D}} = \frac{r_m}{\sqrt{\gamma}}. \tag{25}$$

For example, if we undersample by half, the mean density and the packing factor will be reduced by half, and the maximum radius will be increased by $\sqrt{2}$.

These considerations illustrate a fundamental difference between the effective radius and the maximum radius. The effective radius is an intensive (local) measure, characteristic of individual points, whereas the maximum radius, and consequently the packing factor, are extensive (global) properties of the distribution of points.

Cross-correlation

A natural extension of this approach is to compare how objects of one type are distributed with reference to objects of some other type. For example, how the somata of starburst amacrine

cells that lie in the inner nuclear layer are laterally disposed with respect to the somata of each of the starbursts of the ganglion cell layer. The resulting distribution, termed the *spatial cross-correlogram*, is similar to the spatial autocorrelogram (e.g. Fig. 2). Except for a rotation by 180 deg, it has exactly the same form whichever of the two groups is used as the reference points. If the two sets of points are statistically independent, then the points will be uniformly distributed, and the resulting profile will be flat, independent of the shapes of the density recovery profiles of each of the two groups.

The expected number of points in each annulus may be calculated as follows. Define

A = area of the region that contains the reference points;
N_1 = number of points in the region that belong to the first group;
N_2 = number of points in the region that belong to the second group;
D_1 = density of points in the region that belong to the first group = N_1/A;
D_2 = density of points in the region that belong to the second group = N_2/A;
i = index to each annulus: 1,2,... (or the corresponding bin in the *density recovery profile*);
ΔA_i = area of annulus i;
Δr = width of each annulus;
n_i = measured number of points in annulus i; and
λ_i = expected number of points in annulus i assuming a random distribution.

The expected number of points in the annulus is then the product of the expected number to be found in the annulus for each reference point times the number of reference points N_1. Assuming a random distribution of points in the region of interest, the expected number in the annulus is simply the density of the points D_2 times the area of the annulus ΔA_i:

$$\lambda_i = N_1 D_2 \Delta A = \frac{N_1 N_2 \Delta A_i}{A}. \tag{26}$$

Since, by substitution and rearrangement,

$$A = \sqrt{\frac{N_1 N_2}{D_1 D_2}}, \tag{27}$$

we have

$$\lambda_i = \sqrt{D_1 D_2}\sqrt{N_1 N_2}\Delta A_i. \tag{28}$$

Let

$$d_i = \frac{n_i}{\Delta A_i \sqrt{N_1 N_2}}, \tag{29}$$

then the expected value of this measure for a random distribution is

$$d_{i,\text{random}} = \frac{\lambda_i}{\Delta A_i \sqrt{N_1 N_2}} = \sqrt{D_1 D_2}. \tag{30}$$

This is the simplest scaling factor for the bins of the cross-correlogram that has the dimensions of density and goes to zero as either of the two density values goes to zero.

Compensation factor to maximize data inclusion for rectangular regions

In the density recovery profile for cell bodies, shown in Fig. 3C, the cells used as reference points were all restricted to the interior of the dotted rectangle shown in Fig. 1B. This region includes only 36% of the area of the 1-mm^2 sample area. In this special case, the bounds of the dotted rectangle could have been extended above and below the 1-mm^2 region, since additional sampled points lie in these directions. But if the bounds of the dotted rectangle were extended to the left or right, then a bias would arise, since large annular zones about these extended portions would include portions of the retina that were not sampled. Nevertheless, if the extent of the unsampled portion is known, then it can be used to calculate a compensation factor, assuming similar and uniform conditions in the unsampled regions.

Figure 6 illustrates how the compensation factor can be calculated for a given annular zone of radius r. We assume that there are no sampled points outside the rectangle drawn with the continuous line. For each of its sides a dotted line is drawn parallel to the side, and inset from it by r. These dotted lines divide the rectangle into nine regions. A reference point within the central region requires no compensation for this particular radius, for the annular zone lies wholly within the rectangle. There are four side regions, and a point lying in one of them is shown at the top (Fig. 6A). The shaded region of its annular zone lies outside the sample region, and the effective sample area of this annular zone is consequently reduced. Let h be the perpendicular distance from the point to the upper side of the sample region and f_s be the ratio of the actual sample area to the total sample area for a given annulus. Then

$$f_s = 1 - \frac{\cos^{-1}\left(\frac{h}{r}\right)}{\pi}.$$

given in Appendix 1.1. This factor varies from 0.5 for a point lying on the border ($h = 0$) to 1 for a point lying on the dotted horizontal line. The reduction as a function of vertical position is shown in the shaded region to the right (Fig. 6B). The ratio f_s of the *effective sample area* to the actual area for points lying in this side region, found by integrating h over the range from 0 to r, is (see Appendix 1.1):

$$\bar{f}_s = 1 - \frac{1}{\pi}. \tag{31}$$

For the four corner regions, two cases need to be considered. If a point lies within the shaded region at the top left (Fig. 6C), then there are two portions of the annular zone that lie outside the sample area. The extent of each portion can be calculated as described above; and for the purpose of calculation the point can be viewed as belonging to both the upper side region and the left side region. If a cell lies in the other portion of the corner (Fig. 6D), then we can again view it as belonging to both side regions that intersect at the corner, provided that the reduction in the sample area of the annulus that is shown in black is counted only once. These considerations lead to the following equation, whose derivation is outlined in Appendix 1.2:

$$f = 1 - \frac{2r}{\pi LW}(L + W) + \frac{r^2}{\pi LW}, \tag{32}$$

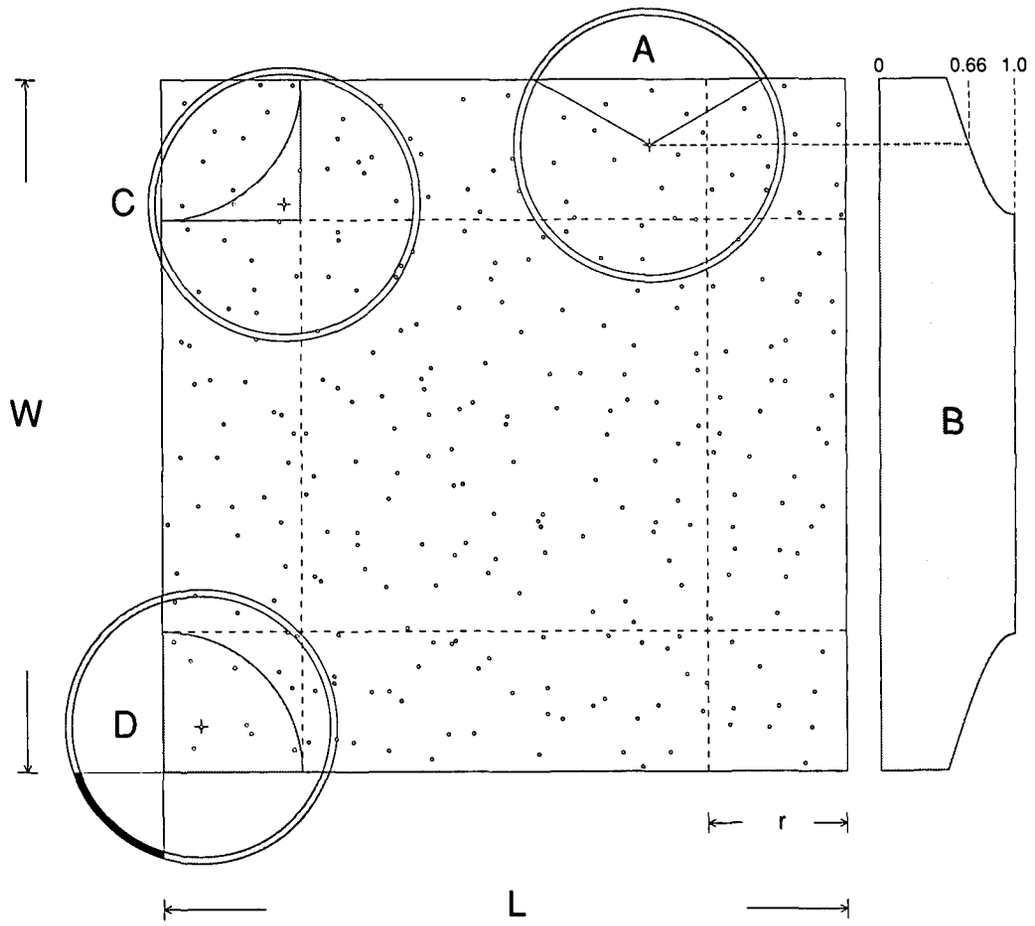


Fig. 6. Diagram to illustrate how data points lying outside the dotted rectangle can be included in the analysis, provided that a compensation factor is included in the calculation. See text for details.

where L and W are the length and width of the rectangular patch, as indicated in Fig. 6. This compensation factor varies with r, and thus must be applied bin by bin to a density recovery profile. Figure 7B shows this method applied to the cells contained in the full 1-mm^2 region of Fig. 1B, compared with the uncompensated density recovery profile for the 0.36-mm^2 sample region enclosed by the dotted rectangle, shown in Fig. 1B and reproduced in Fig. 7A.

Figure 7C shows the result of the nearest-neighbor analysis applied to the same data. In this type of analysis, the shape of the histogram of nearest neighbors is compared to that of the predicted distribution of randomly distributed points with the same spatial density (Rayleigh distribution), shown here as the smooth curve. Compared to the curve, the histogram tends to be shifted to the right, indicating a deviation from randomness in the anticlustering direction.

Expected variance in r_e for a random distribution

Even a random distribution of points can generate r_e values greater than 0, since by chance, the first and one or more successive bins in the density recovery profile may have values less than the mean density. In order to investigate this effect, we return to the measure of the volume of each annulus used to calculate the effective radius:

$$\Delta V_i = \begin{cases} \dfrac{\lambda_i - n_i}{N} & \lambda_i > n_i\,, \\ 0 & \lambda_i \le n_i\,. \end{cases} \tag{33}$$

This variable includes the parameter N, the total number of reference points. In order to express the variation in a more general manner, we introduce a new variable ΔQ_i defined as follows:

$$\Delta Q_i = \frac{N\Delta V_i}{\lambda_1} = \frac{\lambda_i - n_i}{\lambda_1}, \tag{34}$$

which specifies the value of this measure for a given bin i and a given count in the bin n_i. Assume for the moment that we have a specific count for each bin $(n_1, n_2, n_3, \ldots)$. Then the summed value Q for all bins is, from eqns. (7) and (34),

$$Q = \sum_{i=1}^{\text{until } n_i > \lambda_i} \frac{(\lambda_i - n_i)}{\lambda_1} = \frac{N}{\lambda_1}\left(\frac{1}{N}\sum_{i=1}^{\text{until } n_i > \lambda_i} (\lambda_i - n_i)\right)$$

$$= \frac{NV_e}{\lambda_1} = \frac{N\pi r_e^2 D}{N\pi \Delta r^2 D} = \left(\frac{r_e}{\Delta r}\right)^2 \tag{35}$$

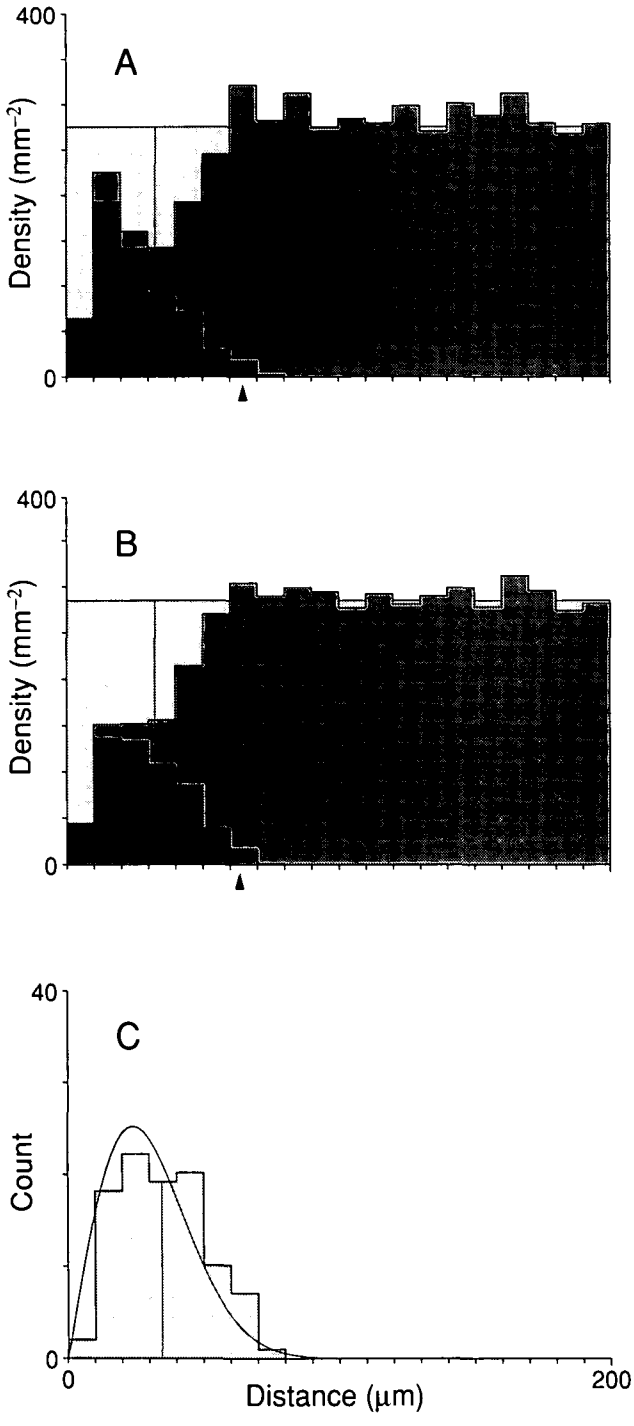


Fig. 7. A: Same density recovery profile as shown in Fig. 3C, with nearest neighbors shown in darkest gray. The data points are those contained within the dotted rectangle of Fig. 1B. Effective radius = 32.7 μm. Arrowhead here and below indicates the maximum radius. B: Same data as for diagram A, except that all data points in the 1 mm^2 are used, together with a compensation factor discussed in the text. Again, the contributions of nearest neighbors are shown in darkest gray. Effective radius = 32.4 μm. C: Nearest-neighbor analysis for the same data used in diagram A. The gray histogram shows the distribution of nearest neighbors; the curve shows the Rayleigh distribution for an equal number of points distributed randomly with the same density. The shape of this histogram differs from that of the nearest-neighbor histogram shown in darkest gray in diagram A, because a scaling factor has been applied to each bin of the latter histogram, as described in the text. The mean value (34.6 μm) is shown as a thin vertical line.

or

$$\frac{r_e}{\Delta r} = \sqrt{Q}. \tag{36}$$

This is the result for a given set of bin values $(n_1, n_2, n_3, \ldots)$. The only independent factor on the right-hand side of this equation is λ_1 [eqn. (34)]. Since, by eqn. (14) λ_1 is the square root of the reliability factor k, we can also use k as the independent factor, and this is the underlying reason for devising the parameter Q. The probability density function of $r_e/\Delta r$ can now be determined by calculating the probabilities of all combinatorial possibilities of the set of bin values $(n_1, n_2, n_3, \ldots)$. The number of counts in a given bin n_i is a random variable governed by the Poisson distribution:

$$P(n_i = X) = \frac{\lambda_i^x}{X!} e^{-\lambda_i}. \tag{37}$$

It is not actually necessary to calculate separately the probabilities of each combinatorial set $(n_1, n_2, n_3, \ldots)$, since a probability tree terminating in convolution integrals can be used to combine the summed probability distributions from different bins. The algorithm devised for this calculation is straightforward, but detailed, and will not be described.

Significance testing

The aim behind these calculations is to determine the distribution of values of $r_e/\Delta r$ for a random distribution having a given reliability factor k, so that one can test whether or not a measured value for r_e is significantly different from that produced by a random distribution of points. It is therefore convenient to plot the parameter

$$1 - P[k, (r_e/\Delta r)], \tag{38}$$

where $P[k, (r_e/\Delta r)]$ is the probability distribution function, found by summing the probability density functions. Figure 8 shows this parameter plotted against $r_e/\Delta r$ for different values of the reliability factor k calculated by the method outlined above. For example, the distribution of cells shown in Figs. 1-3 had a measured value for r of 32.7 μm, using a 10-μm bin width. Their ratio (3.27) is shown as a vertical line along the abscissa in Fig. 8. The reliability factor is 3.0, as described earlier. The probability that the observed value for the effective radius was drawn from a random distribution is the intersection of this vertical line with the curve labeled $k = 3$. In this case, the intersection point lies well below the graph, indicating a high significance level for the hypothesis that the distribution of points was not randomly distributed.

Distribution of r_e for a nonrandom distribution

The calculated value for the effective radius of the density recovery profile in Fig. 3C was 32.7 μm, based upon 102 reference points, contained in an area of 0.36 mm^2, a 10-μm bin width, and a calculated reliability factor, k of 3.0, as discussed above. We know qualitatively that the expected variation in this measure should vary inversely with the reliability factor k, but this information is insufficient to put standard deviation bars on

each side of the effective radius. In order to do so, we need to model the observed density recovery profile, and use this model to determine the expected distribution of $r_e/\Delta r$ for a given k value, using the same general approach as outlined above for the case of the random distribution.

The model chosen for the dip in the density recovery profile is shown in Fig. 9A. It is an inverted cone, whose apex lies at the origin, with a radius of R. The volume of this cone V_c is

$$V_c = \frac{\pi R^2 D}{3}, \tag{39}$$

so that the expected value for the effective radius, from eqn. (24), is

$$r_e = \sqrt{\frac{V_c}{\pi D}} = \frac{R}{\sqrt{3}}. \tag{40}$$

Figure 9B shows the normalized probability density function of $r_e/\Delta r$ for a given value of the reliability factor k. This function was calculated in the same way as for those of Fig. 8 (probability tree leading to convolution integrals), except that all volumes were restricted to the range 0 to R. The reason for this restriction is that actual density recovery profiles rise above the mean value for a while because of second-order effects, as noted earlier, and this feature is not included in the model. In fact, because of these second-order effects, which are a direct consequence of the dip itself, the distribution of counts in one bin is not entirely statistically independent of that in an adjacent bin. The assumptions used to perform the calculation are thus not exactly correct, although the error appears to be small.

The shaded tails of Fig. 9B each compose 15.87% of the total area. This fraction is the area of a tail of a normal distribution from one standard deviation, and the distance to a tail is thus equivalent to that of an ordinary standard deviation bar. Figure 9C shows the size of the standard deviation bars as a function of the reliability factor.

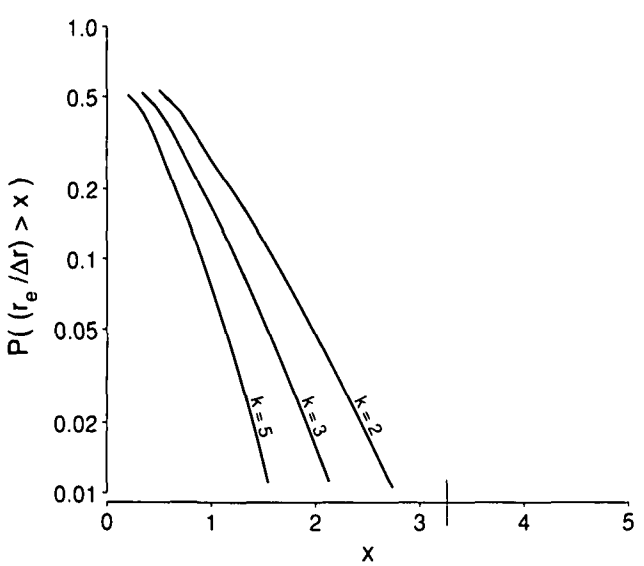


Fig. 8. Probability that a measured value of $r_e/\Delta r$ from a random distribution is greater than a given value, as a function of the reliability factor k.

The magnitude of the standard deviation from $r_e/\Delta r$ in Fig. 9C is shown by two curves that plot the equations $\sigma_{lower} = 0.6k^{-0.859}$ and $\sigma_{upper} = 0.74k^{-0.957}$, whose coefficients were adjusted to fit the integer values of k for 2-5. The equations were empirically fit, and thus without theoretical significance; however, their exponents are close to -1, indicating that the standard deviation bars decrease in an approximately reciprocal manner with an increase in the reliability factor. The actual curves are fractal staircases, although the irregularity is relatively small.

Implementation

The density recovery profiles shown here were calculated and printed using a Macintosh application written by me and named MacDRP, which is available upon request.

Discussion

In this paper, I have presented a method that allows one to examine the relations between objects in the plane and surrounding objects. These relations are expressed in terms of a concatenated profile of the average spatial density of the surround objects, and two measures derived from it: the *effective radius* of an average object and the *packing factor* of the objects that surround it. A method of determining the reliability of these measures is presented. Several other aspects of the distribution of objects are also important.

Orientation

The term orientation, when applied to the distribution of points in the plane, implies some dependence upon one or more directions within the plane. We need to consider the possibility of orientation, because the interpretation applied to measures extracted from the density recovery profile depend upon it. The term *global orientation* was introduced earlier, in the section titled Orientation, and used to refer to orientation factors that are presumed to be constant over the sample area. For the spatial distributions that formed the examples used here, there was no detectable global orientation over the sample region. However, if significant anisotrophic growth of the retina occurs after the distribution of points has been established then, like stretching, it should produce an orientation detectable via a radial plot of density *vs.* angle of offset, such as shown in Fig. 3B.

The size of the domain being considered is critical to the notion of orientation. For example, the spatial distribution of cones in the retinas of many species, including primates, show a local hexagonal packing even in the periphery, where they are well-separated from one another. The axes of this orientation are, however, not constant throughout the retina. Consequently, this type of orientation may not be readily detectable over sample areas that are large compared to the rate of change of the axes of orientation. Likewise, orientation detectable within any sample region may vary with retinal position, depending upon factors such as anisotropic growth. For additional discussion of orientation factors in spatial autocorrelograms, see Cliff and Ord (1973); more sensitive tests for detecting it are described by Zar (1974). However, both effective radius and packing factor appear to be reasonably robust as far as a limited amount of orientation is concerned, so that orientation

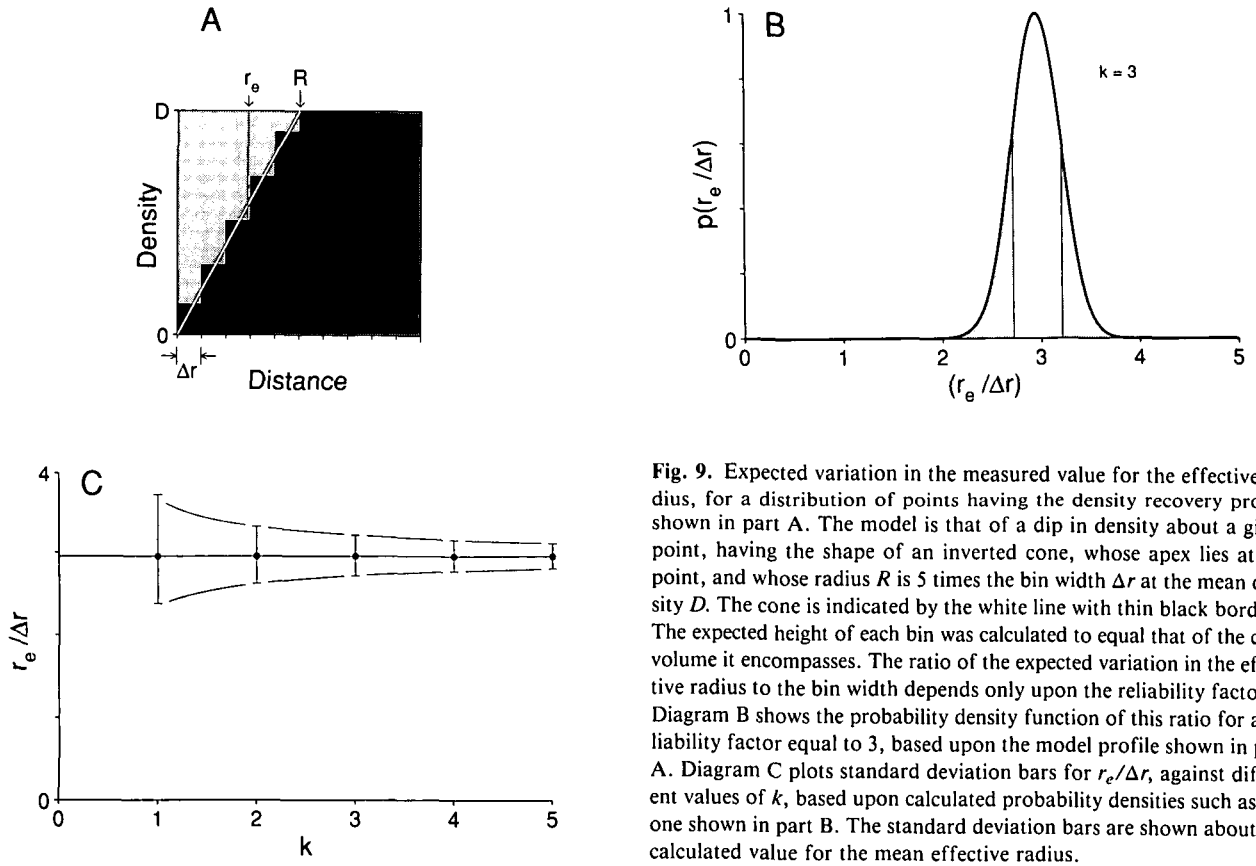


Fig. 9. Expected variation in the measured value for the effective radius, for a distribution of points having the density recovery profile shown in part A. The model is that of a dip in density about a given point, having the shape of an inverted cone, whose apex lies at the point, and whose radius R is 5 times the bin width Δr at the mean density D. The cone is indicated by the white line with thin black borders. The expected height of each bin was calculated to equal that of the cone volume it encompasses. The ratio of the expected variation in the effective radius to the bin width depends only upon the reliability factor k. Diagram B shows the probability density function of this ratio for a reliability factor equal to 3, based upon the model profile shown in part A. Diagram C plots standard deviation bars for $r_e/\Delta r$, against different values of k, based upon calculated probability densities such as the one shown in part B. The standard deviation bars are shown about the calculated value for the mean effective radius.

constitutes an additional measure rather than a necessary invalidation of those presented here.

Autocorrelograms and spatial frequencies

The autocorrelogram is a standard method in signal analysis, which readily lends itself to the distribution of points in any dimensional space. Rodieck (1967) applied it to the maintained discharge of retinal ganglion cells in order to determine how the occurrence of an action potential influenced the occurrence of subsequent action potentials. The Laplace transform of the autocorrelogram gives the *power density spectrum*. In the examples presented here, the dip in spatial density about a given cell in the spatial domain would be reflected in the frequency domain as a peak in the power density spectrum somewhere near the reciprocal of the width of the dip. For the starburst amacrines, with an effective radius 32.7 μm, the peak would lie somewhere around 30 cycles/mm; for the swellings of maze cells, with an effective radius of 1.4 μm, the peak would lie around 700 cycles/mm, with a magnitude, width, and position that depends upon the packing factor.

Related methods

Plant ecologists have developed a number of measures of distributions in the plane, which they have used to determine the degree to which trees or bushes of the same species cluster (e.g. Ripley, 1981). Some are devised for and suited to measurements in the field, where it is often easier to measure the distances between objects, or upon the number of objects contained within some area, than to determine absolute positions. For the latter case, the *L-function* of Ripley (1981) may be viewed as a transformation (scaled integral) of the density recovery profile described here. Diggle (1986) extended a related measure and applied it to ChAT-labeled cells in the rabbit retina. Shapiro et al. (1985) devised an elastic ball model for the distribution of blue cones in the primate retina, and used these measures to fit their data to the model.

These methods may be somewhat stronger than the approach presented here when it comes to testing for statistical significance in critical situations, since they are based upon cumulative distributions rather than scaled histograms, and thus avoid the complications associated with bin widths. The books by Ripley (1981), Diggle (1983), and Cliff and Ord (1973) should be consulted should this prove to be an issue. However, when the deviation from randomness is obvious (e.g. Fig. 3C), and highly significant (e.g. Fig. 8, where the probability of the null hypothesis is so small that it lies off the diagram), then the method and measures described here provide an alternative that presents the data in a readily comprehensible form.

Perhaps the most widely used method for the spatial analysis of retinal distributions is based upon the distance of each point to its nearest neighbor, introduced as a quantitative method by Clark and Evans (1954) and first used for retinal analysis by Wässle and Riemann (1978). The basic method is to compare the observed histogram of nearest neighbors with the distribution derived for a random (Poisson) arrangement of points, and known as the *Rayleigh distribution*. Figure 7 relates

this nearest-neighbor measure to the density profile shown in Fig. 3C. Figure 7A shows the nearest-neighbor contribution to the density recovery profile, and Fig. 7C shows the bin count of nearest neighbors compared to that of a Rayleigh distribution of points having the same spatial density.

It is obvious from Fig. 7 that the density recovery profile provides a substantially different measure than the more traditional nearest-neighbor/Rayleigh-distribution analysis, and possesses three advantages when compared with it. First, estimates are based upon all of the other points about a given point, rather than just one, thereby increasing reliability. Second, any deviation from randomness is readily seen in the deviation of the amplitude of the density recovery profile from the mean density. But the real value of the density recovery profile lies in its ability to provide a local measure of the effective "dead space" around an average object (i.e. the effective radius), and a measure characteristic of the spatial distribution itself (i.e. the packing factor).

Acknowledgments

This research was supported in part by NIH Grants EY02923, EY06098, and EY01730, by the E.K. Bishop Foundation, the Chatlos Foundation, Research to Prevent Blindness Inc., and by NIH Grant RR00166 to the Regional Primate Research Center at the University of Washington. The data used as examples in this paper are drawn from collaborative studies with Masami Watanabe and David Marshak. I would like to thank Clyde Oyster, I. Uramoto, and the two reviewers, for a number of useful comments and suggestions; I would also like to thank Toni J. Haun and Frank J. Hekel for their many contributions to this study.

Appendices

Appendix 1.1

Derivation of eqn. (31)

With reference to Fig. 10A,

$$\cos\theta = \frac{h}{r}.$$

For the annulus:

$$\text{Total Sample Area} = 2\pi r\Delta r,$$

$$\text{Effective Sample Area} = (2\pi - 2\theta)r\Delta r,$$

$$f_s = \frac{\text{Effective Sample Area}}{\text{Total Sample Area}},$$

$$= \frac{2\pi - 2\theta}{2\pi} = 1 - \frac{\theta}{\pi},$$

$$= 1 - \frac{\cos^{-1}\left(\frac{h}{r}\right)}{\pi}.$$

The average value is found by integrating x over the range from zero to r:

$$\bar{f}_s = \frac{1}{r}\int_0^r f(x)\,dx = 1 - \frac{1}{\pi}.$$

Appendix 1.2

Derivation of eqn. (32)

Figure 10B shows a sample area divided into nine regions in the same manner as for Fig. 6. The symbol in each region corresponds to its area:

$$\boldsymbol{a} = (L - 2r)(W - 2r),$$

$$\boldsymbol{b} = (W - 2r)r,$$

$$\boldsymbol{c} = (L - 2r)r,$$

$$\boldsymbol{d} = r^2.$$

Area $\boldsymbol{a}$ requires no compensation factor. Areas $\boldsymbol{b}$ and $\boldsymbol{c}$ require the compensation factor for sides described above. The corner

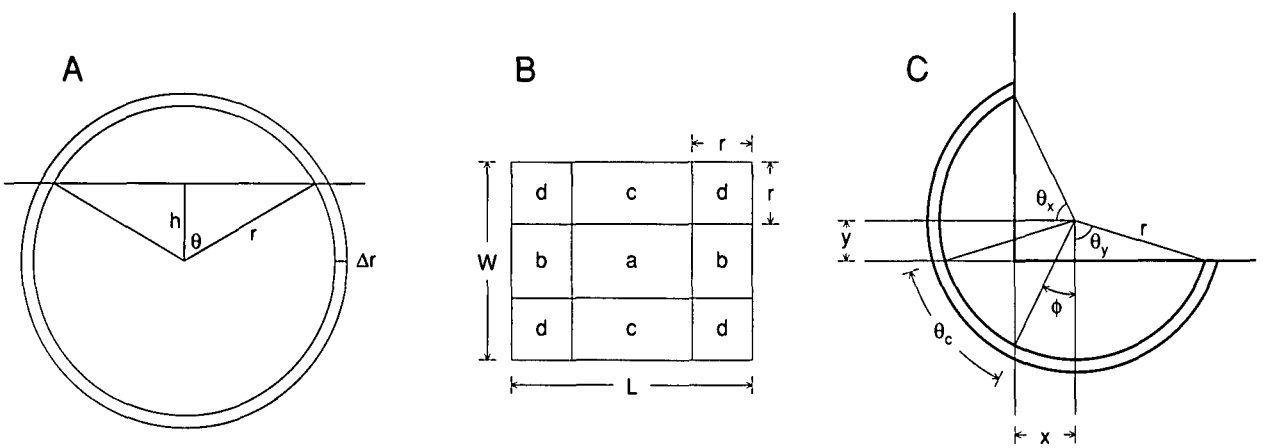


Fig. 10. A: Parameters for the calculation of the compensation factor for the sides of the sample area in Appendix 1.1. B: Parameters for the calculation of the compensation factor for the corners of the sample area in Appendix 1.2. C: Parameters for the calculation of compensation factor f_c in Appendix 1.2.

areas ($\boldsymbol{d}$) also require a compensation factor f_c, as discussed below. Let

$$\boldsymbol{b}' = \bar{f}_s \boldsymbol{b},$$

$$\boldsymbol{c}' = \bar{f}_s \boldsymbol{c},$$

$$\boldsymbol{d}' = \bar{f}_c \boldsymbol{d}.$$

For the annulus,

$$\text{Effective Sample Area} = \boldsymbol{a} + 2\boldsymbol{b}' + 2\boldsymbol{c}' + 4\boldsymbol{d}',$$

$$\text{Total Sample Area} = \boldsymbol{a} + 2\boldsymbol{b} + 2\boldsymbol{c} + 4\boldsymbol{d} = LW,$$

$$f = \frac{\text{Effective Sample Area}}{\text{Total Sample Area}},$$

$$= \frac{\boldsymbol{a} + 2f_s(\boldsymbol{b} + \boldsymbol{c}) + 4f_c\boldsymbol{d}}{LW}.$$

The only factor that remains to be evaluated in the above equation is f_c. The essential problem is to determine what fraction of the annulus in Fig. 6D lies outside the sample area. Figure 10C shows a number of parameters that allow this fraction f_c to be calculated. For an annulus of radius r, and width Δr,

$$\text{Total Sample Area} = 2\pi r \Delta r,$$

$$\text{Effective Sample Area} = (2\pi - 2\theta_s) r \Delta r,$$

$$f_c = \frac{\text{Effective Sample Area}}{\text{Total Sample Area}},$$

$$f_c = \frac{2\pi - \theta_s}{2\pi} = 1 - \frac{\theta_s}{2\pi}.$$

Within this region, any annulus of radius r touches or overlaps both the left and the bottom borders. The angular portion lying outside the border, θ_s, may be viewed as consisting of two portions: one due to the sum of the two side contributions, taken separately; the other being a negative factor, equal to the angular overlap between the two side contributions θ_c:

$$\theta_s = 2\theta_x + 2\theta_y - \theta_c.$$

Substituting into the above equation for f_c,

$$f_c = 1 - \frac{2\theta_x + 2\theta_y - \theta_c}{2\pi},$$

$$= 1 - \frac{\theta_x}{\pi} - \frac{\theta_y}{\pi} + \frac{\theta_c}{2\pi} = f_{c1} + f_{c2},$$

where

$$f_{c1} = 1 - \frac{\theta_x}{\pi} - \frac{\theta_y}{\pi} = 1 - \frac{\cos^{-1}\left(\frac{x}{r}\right)}{\pi} - \frac{\cos^{-1}\left(\frac{y}{r}\right)}{\pi},$$

$$f_{c2} = \frac{\theta_c}{2\pi}.$$

This separation is necessary because the limits of integration differ for the two terms. The mean value for f_{c1} is readily evaluated:

$$\bar{f}_{c1} = \frac{1}{r^2} \int_0^r \int_0^r f_{c1}(x,y)\, dy\, dx = 1 - \frac{2}{\pi}.$$

In evaluating θ_c, we need to consider two situations shown in Fig. 6C and 6D. For all points shown in the shaded area in Fig. 6C, there is no overlap between the upper and left portions of the annulus that lie outside the sample region. Hence, θ_c is equal to 0 everywhere in the shaded area of Fig. 6C. For the rest of area $\boldsymbol{d}$, shown shaded in Fig. 6D,

$$\theta_c = \theta_x - \phi = \theta_x - \left(\frac{\pi}{2} - \theta_y\right) = \theta_x + \theta_y - \frac{\pi}{2}.$$

Hence,

$$f_{c2} = \frac{\theta_x + \theta_y - \frac{\pi}{2}}{2\pi} = \frac{\cos^{-1}\left(\frac{x}{r}\right) + \cos^{-1}\left(\frac{y}{r}\right)}{2\pi} - \frac{1}{4}.$$

The mean value for f_{c2} leads to a difficult integration, because the radical in the upper limit of the inner integral creates complications for the outer integral; but it evaluates to a simple value:

$$\bar{f}_{c2} = \frac{1}{r_2} \int_0^r \int_0^{\sqrt{r^2 - x^2}} f_{c2}(x,y)\, dy\, dx = \frac{1}{4\pi}.$$

Hence,

$$\bar{f}_c = \bar{f}_{c1} + \bar{f}_{c2} = 1 - \frac{2}{\pi} + \frac{1}{4\pi}.$$

The compensation factor is thus

$$f = \frac{\text{Mean Effective Sample Area}}{\text{Total Sample Area}},$$

$$f = \frac{\boldsymbol{a} + 2\bar{f}_s(\boldsymbol{b} + \boldsymbol{c}) + 4\bar{f}_c\boldsymbol{d}}{LW}.$$

After substitutions and rearranging terms,

$$f = 1 - \frac{2r}{\pi LW}(L + W) + \frac{r^2}{\pi LW}.$$

This equation is valid for any combination of the parameters r, L, and W that leave area $\boldsymbol{a} \geq 0$. The smallest value of this factor $(1 - 7/(4\pi) = 0.44)$ occurs when $L = W = 2r$.

References

Clark, P.J. & Evans, F.C. (1954). Distance to nearest neighbour as a measure of spatial relationships in populations. *Ecology* **35**, 445-453.

Cliff, A.D. & Ord, J.K. (1973). *Spatial Autocorrelation*. London: Pion Limited.

DIGGLE, P.J. (1983). *Statistical Analysis of Spatial Point Patterns*. London: Academic Press.

DIGGLE, P.J. (1986). Displaced amacrine cells in the retina of a rabbit: analysis of a bivariate spatial point pattern. *Journal of Neuroscience Methods* **18**, 115–125.

RIPLEY, B.D. (1981). *Spatial Statistics*. New York: John Wiley.

RODIECK, R.W. (1967). Maintained activity of cat retinal ganglion cells. *Journal of Neurophysiology* **30**, 1043–1071.

RODIECK, R.W. & MARSHAK, D.W. (1989). Spatial distribution of choline acetyl-transferase (ChAT) labeled cells in the macaque retina. *Society of Neuroscience Abstracts* **15**, 1207.

RODIECK, R.W. & WATANABE, M. (1988). Morphology of ganglion cell types that project to the parvocellular laminae of the lateral geniculate nucleus, pretectum, and superior colliculus of primates. *Society of Neuroscience Abstracts* **14**, 1120.

ROSE, R.D. & GRIMSON, R.C. (1988). N-Dimensional clusters of neuronal somata: a statistical analysis. *Society of Neuroscience Abstracts* **14**, 550.

SHAPIRO, M.B., SCHEIN, S.J. & DEMONASTERIO, F.M. (1985). Regularity and structure of the spatial pattern of blue cones of macaque retina. *Journal of the American Statistical Association* **80**, 803–812.

VANEY, D.I., PEICHL, L. & BOYCOTT, B.B. (1981). Matching populations of amacrine cells in the inner nuclear and ganglion cell layers of the rabbit retina. *Journal of Comparative Neurology* **199**, 373–391.

VOIGT, T. (1986). Cholinergic amacrine cells in the rat retina. *Journal of Comparative Neurology* **248**, 19–35.

WÄSSLE, H. & RIEMANN, H.J. (1978). The mosaic of nerve cells in the mammalian retina. *Proceedings of the Royal Society B (Biology)* **200**, 441–461.

ZAR, J.H. (1974). *Biostatistical Analysis*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.