

# Visibility in three-dimensional cluttered scenes

Michael S. Langer\* and Fahim Mannan

School of Computer Science, 3480 University, Room 318, McGill University, Montreal H3A2A7, Canada

\*Corresponding author: michael.langer@mcgill.ca

Received January 3, 2012; revised May 11, 2012; accepted June 29, 2012;  
posted July 10, 2012 (Doc. ID 160878); published August 8, 2012

Three-dimensional (3D) cluttered scenes consist of a large number of small surfaces distributed randomly in a 3D view volume. The canonical example is the foliage of a tree or bush. 3D cluttered scenes are challenging for vision tasks such as object recognition and depth perception because most surfaces or objects are only partly visible. This paper examines the probabilities of surface visibility in 3D cluttered scenes. We model how the probabilities of visible gaps, depth discontinuities, and binocular and half-occluded points depend on scene parameters such as the size and density of the surfaces that make up the clutter, as well as on depth and inverse depth. Inverse depth is of particular interest since both binocular disparity and motion parallax depend directly on it. The probability models are verified using data from synthetic 3D cluttered scenes, which are generated using computer graphics. © 2012 Optical Society of America

OCIS codes: 150.6910, 330.1400.

## 1. INTRODUCTION

The term *visual clutter* usually refers to scenes that contain many objects. The word “clutter” typically has negative connotations, implying that there is an excessive number of objects that degrade visual performance in carrying out some task [1], for example, visual search for a target among a set of distractors [2]. In natural scenes, visual clutter can arise in many ways. A ground surface may be covered in a texture whose elements are objects of interest. Clutter also arises from linear perspective and occlusions, in that the image size of an object depends on the object’s three-dimensional (3D) size, its distance from the observer, and on haphazard occlusion relationships that are determined by the depth ordering relative to the viewer position. Note that objects have a wide range of 3D sizes and any given object may be viewed from many distances and positions. With few exceptions, such as the moon or the sun, there is no natural relationship between the 3D size of an object and the visual angle it subtends.

This paper addresses a particular kind of 3D clutter that arises from objects such as trees, bushes, and tall grass. Such scenes contain many surfaces that are distributed in a bounded 3D volume. Images of such scenes contain many depth discontinuities. Surfaces that are visible may occlude objects of interest and the surfaces may be the objects of interest themselves, for example, an apple or squirrel in a tree, or a golf ball in the tall grass. Objects of interest may also lie beyond the clutter, for example, a scene that is partly visible through a hedge.

Visibility in a cluttered scene depends probabilistically on depth. Surfaces that are deeper within the clutter are more likely to be occluded. But how does this falloff in visibility depend on the scene parameters? For example, suppose we have an object that is of some size. If the object were placed deeper within the clutter, then the object would project to fewer image pixels because of linear perspective. At the same time, the expected percentage of the pixels for which the object would be occluded would increase. How do these two effects

interact? A similar question arises when we consider depth discontinuities. As the object is placed deeper within a cluttered volume, it subtends a smaller visual angle and so the ratio of boundary to region pixels increases. This implies that the density of pixels lying on depth discontinuities increases at greater depths. Yet again, objects at greater depths are more likely to be partially occluded, which decreases the number of visible discontinuities at greater depths. How do these effects interact?

Similar questions about visibility and depth arise in binocular vision. Having a second eye allows a viewer to triangulate those surface points that are visible in both eyes. It also increases the chance that a surface will be visible to at least one eye [3]. If a surface is deeper in the scene, then it is more likely to be occluded in either eye. This implies that points that are binocularly visible (and hence can be triangulated) tend to be those that are closer to the viewer, and those points that are deeper in the volume are more likely to be at best monocularly visible. How should one pose these probabilistic statements, and how do these probabilities depend on the parameters of the 3D clutter, such as the size and density of objects? These are the questions that this paper addresses.

The paper is organized as follows. Section 2 discusses related work. Section 3 presents a model of the 3D cluttered scene model and reviews the visibility probabilities for a monocular observer in such scenes. The original contributions begin in Section 4, where we investigate the probabilities of visible 3D gaps in the scene. We show that the size distribution of visible gaps depends both on the depth and on the ratio of gap size to occluder size. These results can be stated equivalently in terms of the probabilities that an object of a given size is entirely visible. These results are fundamental for understanding the information that is available for solving the problem of detection and recognition of objects that are partially occluded in 3D clutter. In Section 5, we examine how the probabilities of depth discontinuities between neighboring pixels in an image depend on the depth and inverse depth and

on the cluttered scene parameters such as size and density of surfaces. We show that certain models that are commonly used in computer vision for 3D reconstruction of uncluttered scenes also capture reasonably well the probabilities of depth discontinuities in cluttered scenes. In Section 6, we examine binocular visibility. We show how the probability that a point is visible to both eyes decreases with depth and, as a consequence, how the binocular power [3] of the second eye increases with depth. Section 7 presents an example of a leafless tree. This example illustrates some of the differences that arise between the theoretical clutter model in which the clutter is uniform versus nonuniform clutter that sometimes arises in natural scenes. We conclude in Section 8.

## 2. RELATED WORK

The canonical example of a 3D cluttered scene is the foliage and branches of a set of plants. Researchers in the fields of agriculture, botany, and forestry have developed many models of the spatial distribution of plants and their branches and leaves, from single bushes and trees to entire forests [4]. These models often describe, for example, how light penetrates to different layers of a tree canopy or how the spacing of crops affects the distribution of illumination at different heights and at different times of the year. Such models are similar in flavor to the ones we develop in this paper, but the goals are slightly different. Those models address the visibility of the light source from the perspective of different points within the cluttered scene, whereas this paper addresses the visibility of the scene as viewed from the perspective of an observer. Later we will discuss certain aspects of these models [5–10]. For now, we turn our attention instead to a different branch of the literature, namely vision studies that are concerned with natural image statistics, in particular, the visibility and size distributions of the objects as seen by an observer.

An early seminal study by Ruderman and Bialek addressed forest scenes and the resulting scale invariant properties of image statistics [11]. This study was followed by others that used a “dead leaves” model [12,13] to show how image scaling effects [14] can arise from occlusions. The dead leaves model is a stochastic process whereby each image in the model is created by a sequence of opaque two-dimensional (2D) objects dropped into the image plane such that each object’s position is random and size is chosen according to some distribution. Ruderman [15] considered a power law distribution of object sizes and showed how two point correlations depend on the object size distributions [16]. Huang *et al.* [17] and Lee *et al.* [18] considered both intensity and range images, and used a scale invariant object size distribution to account for observed scaling. Although the dead leaves model uses orthographic projection, the range of object sizes allows images to contain many small objects and a smaller number of big objects, just as in many natural images.

More recent empirical studies include Yang and Purves [19], who used range images to estimate the image size distribution of visible objects, and Potetz and Lee [20], who examined joint statistics of luminance and range. The latter found that deeper objects within a volume tend to be darker, which may be related to the proximity-luminance bias in human vision. Natural scene range data have been used for generating models of image statistics such as the optical flow seen by a

moving observer [21,22] and the retinal disparities for a binocular observer [23,24].

The issue of binocular visibility has received relatively little attention in the natural image statistics literature even though partial occlusions are known to be important in human binocular vision. For example, in da Vinci stereopsis, surface points are visible to one eye but not the other. Such monocular points often can be perceived at their correct depth by applying visibility constraints [25]. Another important aspect of da Vinci stereopsis is that it increases the total number of surfaces that are visible to the observer. Changizi and Shimojo [3] refer to the boosted ability of the binocular observer to see beyond occluders as “X-ray vision.” Forte *et al.* showed that, to see behind occlusion using stereovision, the visual system relies on Gestalt grouping processes such as good continuation [26].

Partial occlusions in stereopsis have also been addressed in computer vision methods [27–31], for finding corresponding points when they exist. Other computer vision problems that consider partial occlusion include synthetic aperture, namely using large camera arrays to reconstruct the image of objects that are partly observable behind occluders such as the leaves of a bush [32]. A final example is the problem of how to place cameras to maximize scene visibility in the presence of random object occlusions [33].

## 3. MODEL OF 3D CLUTTER

Our model of visibility in 3D clutter is based on the theory of coverage processes [34–36].

### A. Theoretical Model: Frontoparallel Disks

We assume the cluttered scene consists of many surfaces that are randomly distributed over a 3D volume. To derive closed form expressions for visibility probabilities, we first assume surfaces are disks of radius  $R$  that are oriented so their normal is parallel to the observer’s optical axis. This assumption will be relaxed in Subsection 3.B and in our computer graphics experiments where the objects are squares of random 3D orientation. We also assume the spatial distribution of the disk centers is a Poisson process with density  $\eta$ , which is the average number of disk centers per unit volume.

We begin by reviewing a well-known result that the probability density of visible surfaces as a function of depth  $z$  in the clutter satisfies a Beer–Lambert law. Consider a line segment with one endpoint at the viewer and the other endpoint at distance  $z$ , as shown in Fig. 1. Let  $x$  be the pixel where this line segment intersects the image plane. For the 3D scene point to be visible, no disk center can fall in a cylinder  $C_x$  whose radius is  $R$  (the radius of the disks) and whose axis is the line

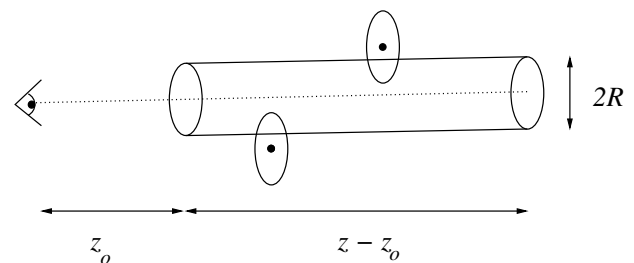


Fig. 1. A surface point  $(x, z)$  is visible if the cylinder centered at the line of sight to that point does not contain any disk centers.

segment of length  $z$ . The probability of this event is calculated as follows.

We assume the cluttered volume lies between depths  $z = z_0$  and  $z = z_1$ , which are taken as fixed. Then such a cylinder  $C_x$  has volume

$$V = \pi R^2(z - z_0). \quad (1)$$

According to the Poisson model, the probability that any 3D volume  $V$  contains zero disk centers is

$$P(\text{no disk centers in } V) = \exp\{-\eta V\}. \quad (2)$$

When  $\eta V$  is small, we have

$$\exp\{-\eta V\} \approx 1 - \eta V. \quad (3)$$

Thus, the probability that there is at least one disk center in the cylinder  $C_x$  is  $\eta V$ .

For a surface point to be visible on a disk in the depth interval  $[z, z + \Delta z]$ , one also requires that there is a disk center in the thin cap of volume  $\pi R^2 \Delta z$ . When  $\Delta z$  is small, we can ignore the probability that there is more than one disk center in the thin cap. Thus, the probability that there is a disk center in the cap is  $\eta \pi R^2 \Delta z$ , and so the probability that the disk that is visible at  $x$  lies in the depth interval  $[z, z + \Delta z]$  is

$$P(\text{visible disk} \in [z, \Delta z]) = \eta \pi R^2 \exp\{-\eta \pi R^2(z - z_0)\} \Delta z \quad (4)$$

$$= \lambda \exp\{-\lambda(z - z_0)\} \Delta z. \quad (5)$$

Letting  $\Delta z \rightarrow 0$  gives the *probability density function* that the visible point at  $x$  is at depth  $z$ ,

$$p(z) = \lambda \exp\{-\lambda(z - z_0)\}. \quad (6)$$

This is a Beer–Lambert law, namely exponential decay in visibility. The decay constant  $\lambda$  is

$$\lambda = \eta \pi R^2. \quad (7)$$

$\lambda$  is the density of surfaces that intersect a line in the  $z$  direction and so  $\frac{1}{\lambda}$  is thus the average path length between surfaces.

We will also consider how surface visibility depends on *inverse depth*,

$$v = \frac{1}{z}, \quad (8)$$

which has units of  $\text{m}^{-1}$  or diopters. Inverse depth is a natural quantity to consider; for example, it determines motion parallax and binocular disparity, as we will discuss below. (Indeed optical blur also depends on inverse depth, but we will not discuss it in this paper since we are considering pinhole imaging only.) The probability density of visibility as a function of inverse depth,  $p(v)$ , can be obtained from Eq. (5) by a change of variables [37,38] from  $z$  to  $v$ , namely

$$dz = \frac{1}{v^2} (-dv), \quad (9)$$

where the minus sign indicates that a positive step in  $z$  gives a negative step in  $v$ . Letting  $v_0 = 1/z_0$ , we get

$$p(v) = \frac{\eta \pi R^2}{v^2} \exp\left\{-\eta \pi R^2 \left(\frac{1}{v} - \frac{1}{v_0}\right)\right\}. \quad (10)$$

Several interesting observations about  $p(v)$  can be made. First,  $p(v)$  is nonmonotonic. It has a single maximum, which occurs at  $v = \frac{1}{2}$ , or equivalently at  $z = \frac{2}{\lambda}$ , which is twice the average path length between surfaces. The depth  $z$  where this local maximum occurs depends on occlusions and perspective effects. Scene points that are further away are less likely to be visible because of occlusions, but tend to have similar  $v$  values because of the inverse relation between  $z$  and  $v$ . Note that the depth of the maximum of  $p(v)$  is independent of  $z_0$  and  $z_1$  and thus it does not necessarily fall within the range of depths  $[z_0, z_1]$ . Also, the depth of the maximum of  $p(v)$  shifts to smaller  $v$  when either the density  $\eta$  or size  $R$  of objects is decreased, since this simply increases the probability that a given distant point will be visible.

As mentioned above, Eq. (10) gives us expressions for the retinal speed [21,22] and stereo disparity [23,24] as a function of depth. Suppose an observer is translating laterally at a velocity  $T_x \text{ ms}^{-1}$  and uses a head and/or eye movement to track a point at depth  $z'$ . Then [39] the horizontal retinal velocity in radians per second of a point at depth  $z$  will be approximately  $T_x(\frac{1}{z} - \frac{1}{z'})$ . Similarly, if two eyes are converged at a distance  $z'$  and now  $T_x$  is the interocular distance (IOD), then the angular disparity of a point at depth  $z$  will be approximately  $T_x(\frac{1}{z} - \frac{1}{z'})$  radians. In both cases, the probability density  $p(v)$  is related to the probability density on image velocities or stereo disparities by a simple shift and scale.

## B. Empirical Model: Randomly Oriented Squares

We compare the theoretical probability densities to visibility data for scenes generated using the computer graphics package RADIANCE [40]. We used scenes that obeyed the assumptions of the model, namely frontoparallel disks, as well as scenes that consisted of square objects and random 3D orientation.

For both the disk and square scenes, we used two sizes of surfaces. One used a large number of small surfaces ( $R = .01$ ), and the other used a small number of large surfaces ( $R = 0.04$ ). The object centers were randomly placed in a depth interval  $[z_0, z_1] = [1, 4]$ . The decay constant was  $\lambda \approx 0.7$  [recall Eq. (6)]. Scenes were rendered under perspective projection with a field-of-view of  $12^\circ$ . The image size was  $1024 \times 1024$  pixels. For the disk scenes, the parameter  $R$  was the radius of each disk. For the square scenes, we chose the width  $w$  of each square so that the average projected area of an unoccluded frontoparallel square at a given depth would be equal to the area of an unoccluded frontoparallel disk at that depth; namely, we chose  $w$  such that  $\frac{w^2}{2} = \pi R^2$ . Figure 2 shows the two classes of scenes consisting of squares.

Figures 3 and 4 plot the means of normalized histograms for depth  $z$  and inverse depth  $v$ , for 100 instances of scenes with large (left) and small (right) disks and squares, respectively, along with error bars that show standard error of the means. Also shown are the theoretical  $p(z)$  and  $p(v)$  curves. The scenes with frontoparallel disks have lower variability than the scenes with randomly oriented squares, and the reason

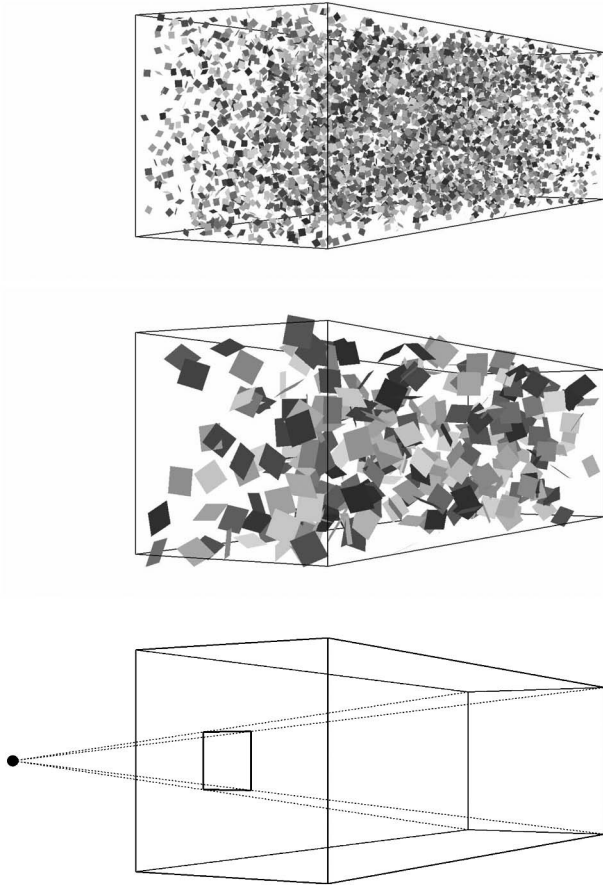


Fig. 2. (Top and middle) Examples of large and small square scenes. (Bottom) The black sphere shows the position of the eye. The dashed lines show the pyramid shaped view volume. The small thick black square on the front face of the rectangular volume shows the “window” on the  $z = z_0$  plane through which the viewer sees the scene. See top row of Figs. 3 and 4.

is that (apart from occlusion effects) the former have constant projected image area whereas the latter do not. In both cases, the means are a good fit to the theoretical curves.

For both the disk and square scenes, the histograms for the large occluder scenes had more variability than for small square scenes. This was expected because there is a smaller number of large occluders and when a large occluder does fall in a  $z$  or  $v$  bin, it contributes a greater number of pixels to that bin. Moreover, the large occluder scenes had larger standard errors at small  $z$  values. There are two reasons for this: first, each occluder subtends a larger visual angle, which leads to an “all or nothing” effect. Second, the  $p(z)$  values themselves are larger at smaller depths, and so the large mean naturally produces a larger variance. Note that for the  $p(v)$  data, these two effects just described go in opposite for larger values of  $v$  and so the sizes of the error bars are more uniform over  $v$ .

It is important to note that we are plotting standard errors of the mean, rather than standard deviations. For  $N = 100$  scenes, the standard deviation is 10 times as great as the standard errors. Histograms for the individual scenes indeed have large variability relative to the smooth theoretical curves  $p(z)$  and  $p(v)$ , especially the case for the large square scenes. This variability would make it difficult to estimate  $\lambda$  directly from the  $z$  or  $v$  histogram, which might be estimated from stereo or motion parallax information. We would expect that to

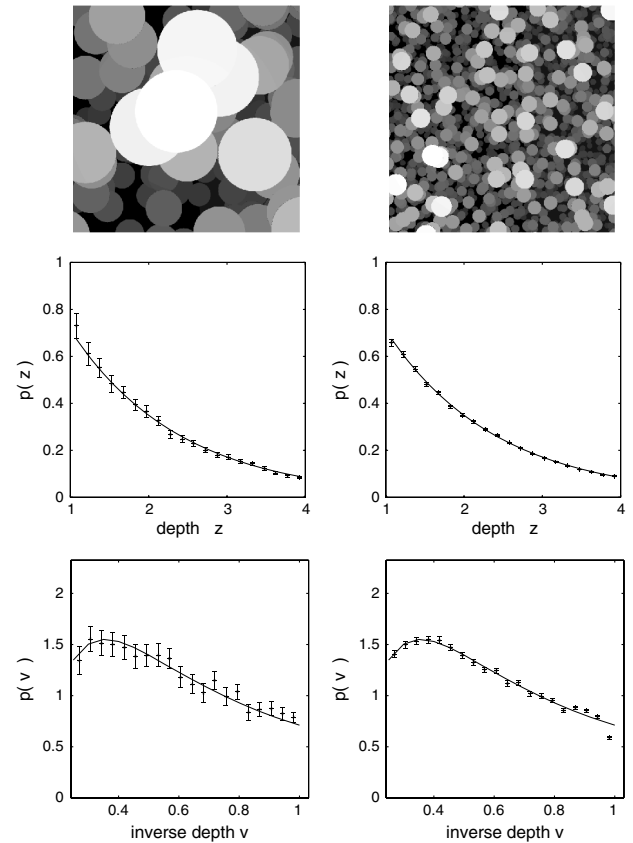


Fig. 3. Top row shows rendered images for scenes with (left) large and (right) small occluders. Here the image intensity is inversely related to depth to better illustrate the depth ordering of elements. Middle row shows  $p(z)$ . Bottom row shows  $p(v)$ . The depth and inverse depth intervals have been partitioned into  $n_{\text{bin}} = 20$  bins. Error bars show standard error of the mean normalized histograms from 100 scenes.

estimate  $\lambda$ , one would need to use more of the available information, such as the image size distribution of partially occluded objects, which might be estimated using a segmentation algorithm. The next section addresses a visibility property that is related to partial occlusion, namely the width of visible gaps in clutter as a function of depth.

#### 4. GAP SIZES

The Beer–Lambert model of Eq. (6) described the probability density  $p(z)$  of visible surface points as a function of depth in the clutter. A more general question is whether a finite area of a surface is visible, or equivalently, whether an empty gap of some area is visible. Such questions are relevant to the problems of object detection and recognition since these problems become more difficult when the object is partially occluded.

In this section, we compute the probability that gaps of various sizes are visible. We use the same theoretical model of frontoparallel occluder disks of radius  $R$ , and again test the model with randomly oriented squares. Let  $P_{\text{gap}}(r, R, z)$  be the probability that a frontoparallel disk of radius  $r$  and at depth  $z$  would be entirely visible to the observer. This new disk could be a real surface that is an object of interest, or it could be a gap in empty space through which objects or occluders at greater depths might be seen.



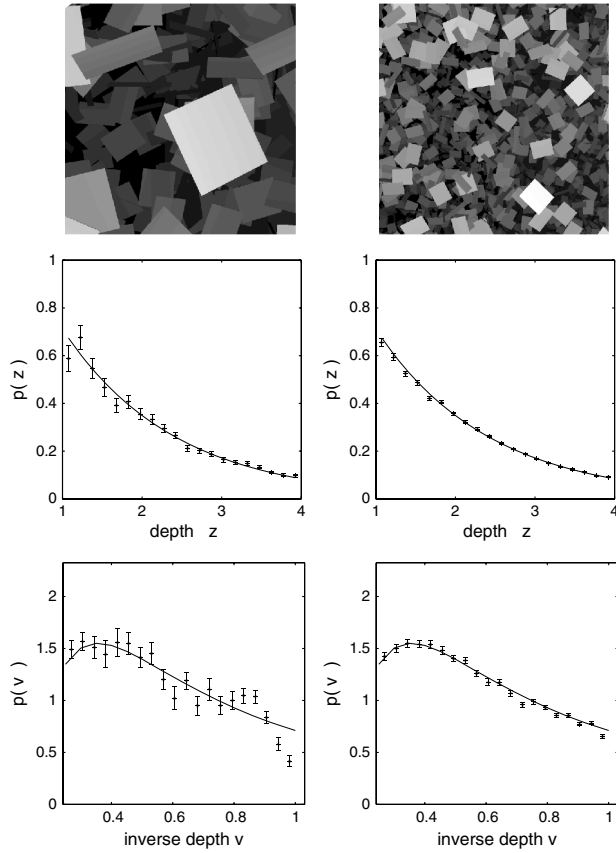


Fig. 4. Same as Fig. 3 except now the scene has randomly oriented squares.

To compute  $P_{\text{gap}}(r, R, z)$ , first note that a disk of radius  $r$  and at depth  $z$  defines the base of a cone of rays of length  $z$ , with apex at the observer. For the base of the cone to be entirely visible to the observer, no disk occluder center can lie within a distance  $R$  from the cone. We calculate the probability that the disk is entirely visible by considering a morphological dilation of the cone by a distance  $R$  and considering the volume  $V_{\text{gap}}$  of the portion of the dilated cone that lies in the clutter between depths  $z_0$  and  $z$ . A general cone of

length  $z$  and base radius  $r$  has volume  $\frac{1}{3}\pi r^2 z$ , and from this it can be shown that

$$V_{\text{gap}}(r, R, z) = \frac{\pi z}{3r} \left( (R+r)^3 - \left( R + \frac{z_0}{z} r \right)^3 \right). \quad (11)$$

From Eq. (2), we get that the probability that all points on the cone base are visible to the observer is

$$P_{\text{gap}}(r, R, z) = \exp\{-\eta V_{\text{gap}}(z, r, R)\}. \quad (12)$$

Substituting  $\eta = \frac{\lambda}{\pi R^2}$  from Eq. (7) gives

$$P_{\text{gap}}(r, R, z) = \exp\left\{-\frac{\lambda}{3} \left(\frac{R}{r}\right) \left( \left(1 + \frac{r}{R}\right)^3 - \left(1 + \frac{z_0}{z} \frac{r}{R}\right)^3 \right) z\right\}. \quad (13)$$

This is a generalization of the earlier analysis of the visibility probability at a single point, in that letting  $r \rightarrow 0$  gives

$$\lim_{r \rightarrow 0} P_{\text{gap}}(r, R, z) = \exp\{-\lambda(z - z_0)\}. \quad (14)$$

This is the same expression as in Eq. (6) except for a scale factor  $\lambda$ , which accounts for the requirement in Eq. (6) that a visible surface indeed exists in  $[z, z + dz]$ . Here we do not have that requirement since the “gap” can be in empty space or it can be a visible region of a surface.

Inspection of Eq. (13) reveals that  $P_{\text{gap}}(r, R, z)$  depends on the ratio  $r/R$ , but does not depend individually on  $r$  and  $R$ . Thus for a fixed  $\lambda$  and  $z$ , a scene with small occluders gives proportionally smaller gap sizes. Figure 5 shows  $P_{\text{gap}}(\frac{r}{R}, z)$  for four ratios  $\frac{r}{R} = 0, \frac{1}{2}, 1$ , and  $2$  and for the two values of  $R$  from earlier. The smooth curves show the theoretical model of Eq. (13). The plotted data are the percentages of points at depth  $z$  that are visible and whose gap size is at least  $r$ . (Gap sizes were computed using the depth maps for the randomly oriented square squares. Specifically, inward morphological dilation was applied to the set of pixels whose depth was less than  $z$ . The plots show the fraction of pixels that remain after the dilation.)

The probability of a visible gap of radius  $r$  falls off very quickly with  $r$ . For example, consider the curves for the case

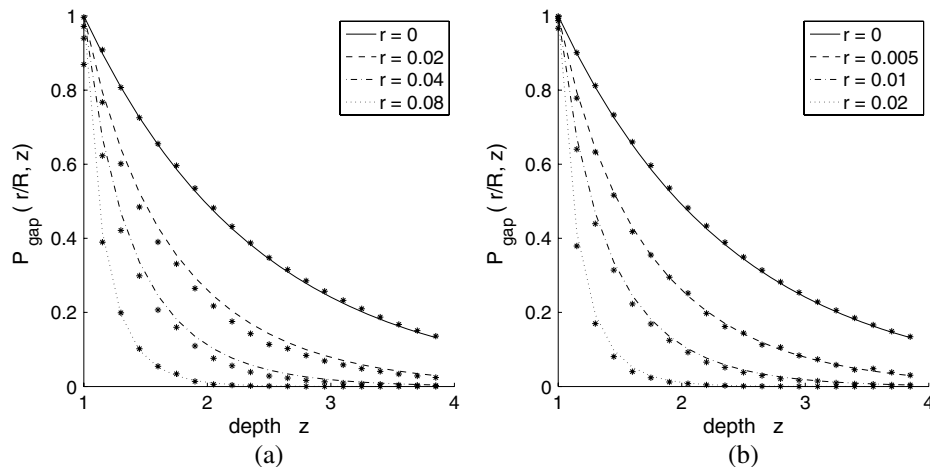


Fig. 5. Gap size probabilities  $P_{\text{gap}}(\frac{r}{R}, z)$  for the scenes with (a) large and (b) small randomly oriented squares. The four curves in each plot from top to bottom are for ratios  $\frac{r}{R} = 0, \frac{1}{2}, 1$ , and  $2$ .

$\frac{r}{R} = 2$  in which the gap sizes are twice as wide as the occluder objects. The gap probabilities are very small for even moderate depths, e.g.,  $z = 2$ . Moreover, even for gaps that are small relative to the objects in the clutter, e.g.,  $\frac{r}{R} = \frac{1}{2}$ , the gap probabilities are low—typically half as great as for the case  $r = 0$ .

To summarize, we have calculated the probability that there is a gap of some radius as a function of depth within the clutter. This probability falls off exponentially with depth  $z$ , and the falloff rate increases very quickly with the ratio of gap size to occluder size. For even moderate size gaps, e.g., half the width of the occluders, the falloff is very fast. Thus, not only is a given point on an object less likely to be visible for objects at greater depths within the clutter (Section 3), but the gaps through which regions of the object are visible tend to become smaller as well. The effect is more pronounced when the cluttered scene consists of a large number of small objects.

## 5. DEPTH DISCONTINUITIES

We next turn to a closely related question that concerns whether two nearby pixels “see” the same surface or whether the surfaces are separated by a depth discontinuity. We show how the probability of depth discontinuities depends on the parameters of the clutter and on depth. We follow the approach of Ruderman [15] and present the probabilities of depth discontinuities separately for the cases in which

neighboring image points are on the same surface versus different surfaces. Let  $p_{\text{same}}(\cdot, \cdot)$  and  $p_{\text{diff}}(\cdot, \cdot)$  be the joint probability densities, respectively, where the arguments can be either depths ( $z, z'$ ) or inverse depths ( $v, v'$ ). Expressions for these densities are derived in Appendix A for the theoretical model of frontoparallel disks of radius  $R$ .

For the theory plots in Fig. 6, define the joint probabilities

$$P(z, z') = \begin{cases} p_{\text{diff}}(z, z')(\Delta z)^2, & \text{if } z \neq z' \\ p_{\text{same}}(z, z')(\Delta z), & \text{if } z = z' \end{cases} \quad (15)$$

from Eqs. (A1) and (A3). The factor  $(\Delta z)^2$  arises in the  $p_{\text{diff}}(z, z')$  case because we have two different surfaces that both must be present in  $[z, z + \Delta z]$ , whereas the factor  $(\Delta z)$  arises in the  $p_{\text{same}}(z, z')$  case because we have just one surface that must be present in this depth interval. Similarly, we define joint probabilities

$$P(v, v') = \begin{cases} p_{\text{diff}}(v, v')(\Delta v)^2, & \text{if } v \neq v' \\ p_{\text{same}}(v, v')(\Delta v), & \text{if } v = v' \end{cases} \quad (16)$$

from Eqs. (A2) and (A4).

Figure 6 plots the negative log of these joint probabilities. We plot  $-\log$  probabilities in order to tone map the enormous range of  $P(\cdot, \cdot)$  values. The off-diagonal elements in the plot are due to depth discontinuities, since the theoretical model

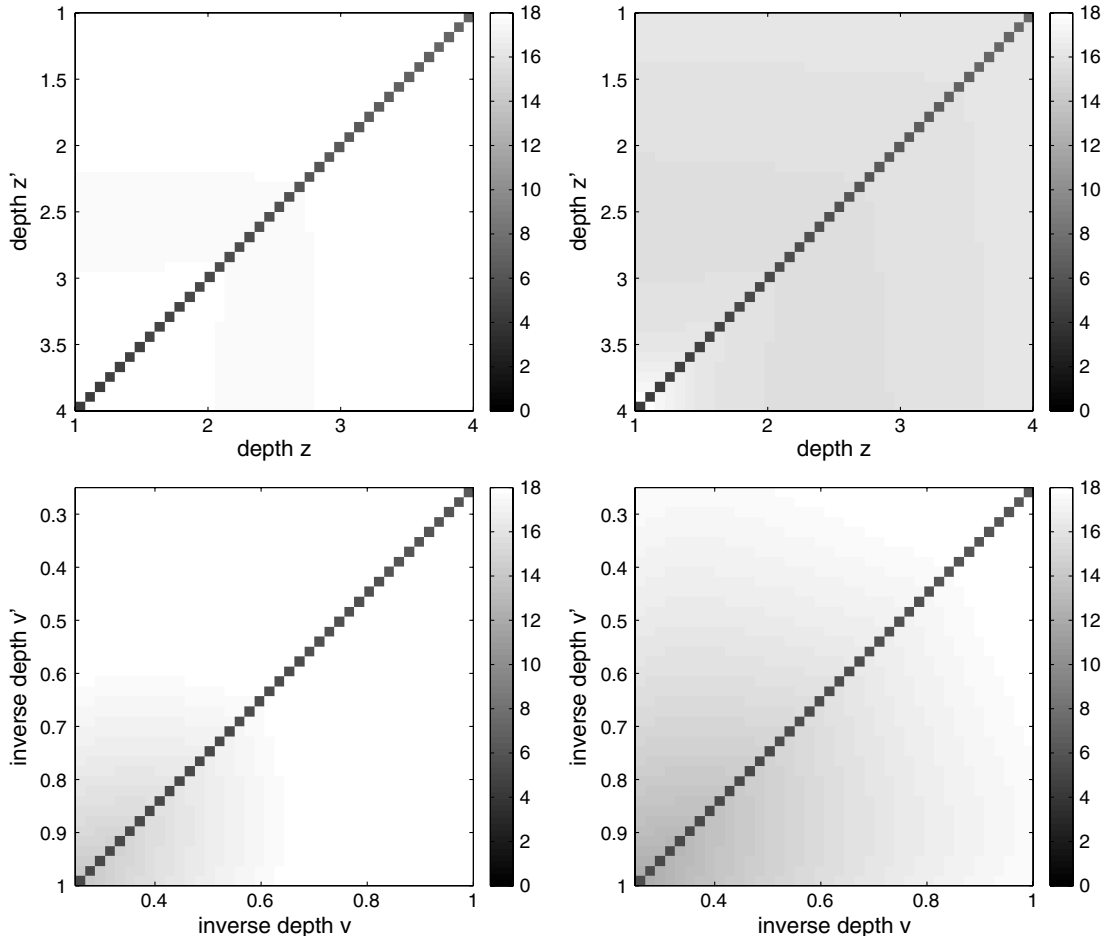


Fig. 6. Theory plots for visibility at neighboring pixels, as derived in Appendix A. Top row are  $-\log P(z, z')$  values. Bottom row are  $-\log P(v, v')$  values. Left and right columns correspond to the large and small disk scenes, respectively.

assumes frontoparallel surfaces. As expected, for each off-diagonal element, the probability of a depth discontinuity is greater in the small  $R$  scenes than in the large  $R$  scenes. The probabilities for the rendered squares surfaces are shown in Fig. 7. The first off-diagonals are elevated relative to the other off-diagonals. The reason is that the square surface has random 3D orientations and so two neighboring pixels can see the same surface and yet fall into a different depth or inverse depth bins by straddling a bin boundary. Such an event leads to an extra  $p_{\text{same}}(z, z')$  or  $p_{\text{same}}(v, v')$  contribution to the off-diagonals that is not captured by the theoretical model. The probability of straddling a boundary in this way is proportional to the bin width ( $\Delta z$  or  $\Delta v$ ). Note that, while it is possible for the depths to straddle two or more depth bins, this event can only occur in the rare circumstance that the surface slant is very high. This is why the joint probabilities in the experiment plots of Fig. 7 are noticeably elevated only in the first off-diagonal.

For the reader familiar with computer vision methods of stereo reconstruction, we note that these log probabilities are related to surface smoothness penalties. In the theory underlying these methods, the diagonally dominant joint probabilities are often interpreted as smoothness priors in a Markov

random field [41]. A typical form of the penalty function on inverse depths (disparities) is

$$E_{x,x'}(v, v') = \min(|v - v'|^k, \tau), \quad (17)$$

where  $v, v'$  are integer multiples of  $\Delta v$ , the exponent  $k$  is typically either 1 or 2, and a small threshold  $\tau$  is used for robustness to depth discontinuities. A simple example is the Potts model, where  $\tau = \Delta v$ . The elevated probability of the first off-diagonal, which is due to slanted surfaces, could be captured by using  $\tau = 2\Delta v$ , for example.

This discussion of the stereo problems brings us to our next set of questions, which concern the visibility of a single point as seen from two points of view. The visibility model is related to what was presented above, except, rather than considering two rays that diverge out from a monocular observer, we consider two rays (one from each eye) that converge on a scene point.

## 6. BINOCULAR VISIBILITY

When a binocular observer views a scene, each surface point can be visible to both eyes, to one eye only (a half-occluded

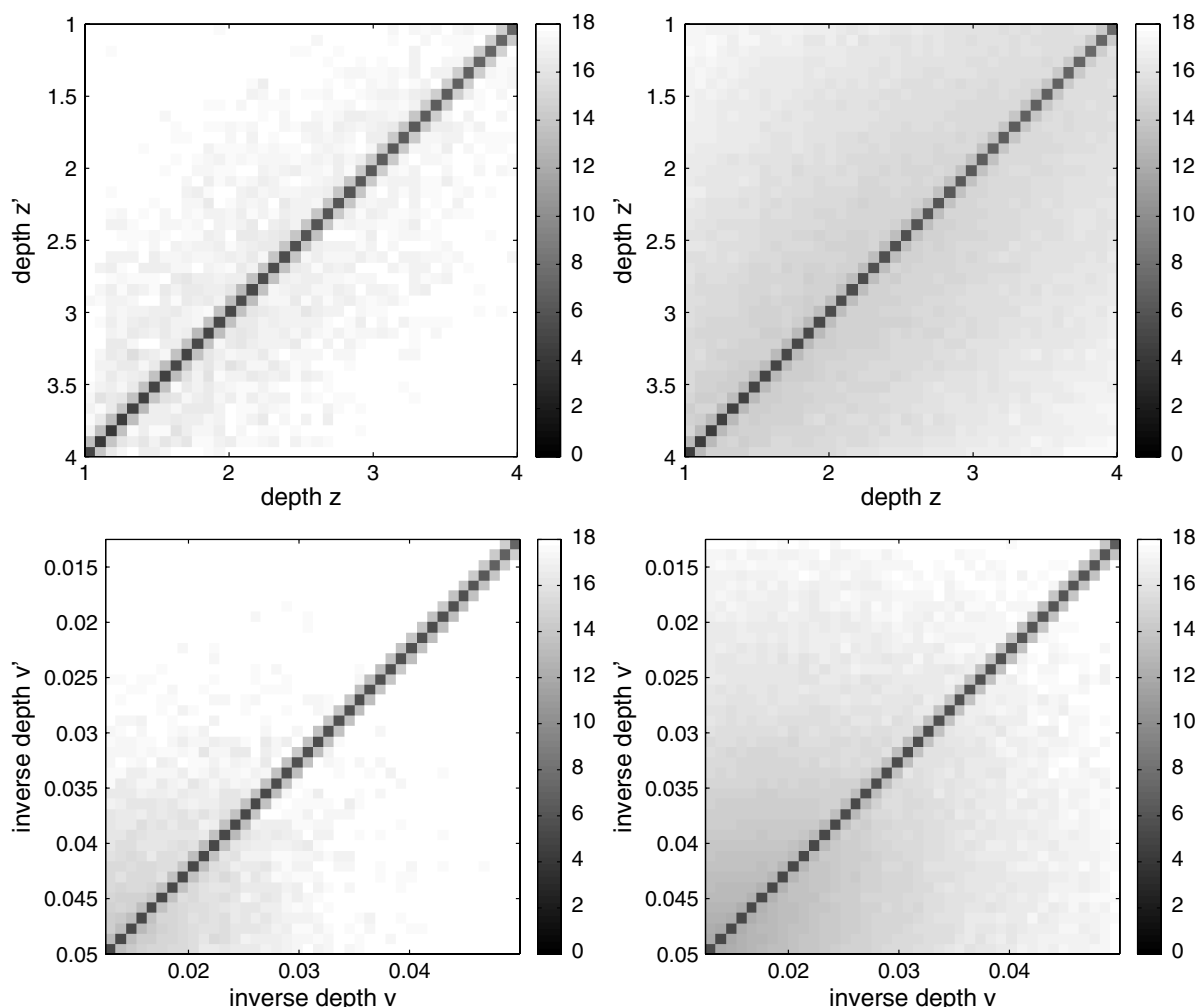


Fig. 7. Probabilities estimated from joint histograms over 100 scenes. (Top row)  $-\log P(z, z')$ . (Bottom row)  $-\log P(v, v')$ . Left and right columns are for large and small square scenes, respectively. The data show means over 100 scenes. The log probabilities within each scene are considerably more variable than illustrated here.

point), or to neither eye. If the surface point is visible to both eyes and the correspondence between left and right eyes can be found, the observer in principle can infer the depth from the disparity. When a surface point is visible to one eye only, depth is more difficult to infer as the visual system has to make use of 3D geometrical constraints [25] and/or monocular cues.

This section examines the visibility probabilities of binocular points and half-occluded points and how these probabilities depend on depth. We begin by deriving expressions for the probability that a point at depth  $z$  is binocularly visible. This point can either be a surface point or it can be a point in free space. The reason for allowing the latter condition is similar to the analysis of gap probabilities; namely, we are also interested in the visibility of isolated objects that may have a different size than the occluders. We are also interested in the visibilities of scene points within the cluttered volume and beyond.

Consider two lines that connect a point at depth  $z$  to the two eyes. Let  $C_l$  and  $C_r$  be the two cylinders whose axes are the two lines, and which are truncated at the ends of the cluttered volume. The point at depth  $z$  is binocularly visible if neither cylinder contains an occluder center, that is, if  $C_r \cup C_l$  does not contain an occluder center. Note that the set union  $C_r \cup C_l$  is equal to the union of  $C_r$  and the set difference  $C_l \setminus C_r$ , and so their volumes are related by

$$V(C_r \cup C_l) = V(C_r) + V(C_l \setminus C_r). \quad (18)$$

An expression of  $V(C_r)$  is obtained from Eq. (1). To obtain an expression for  $V(C_l \setminus C_r)$ , we consider two different cases in which the occluders are either very large or small relative to the IOD  $T_x$ . We refer to these as the *large leaf* and *small leaf* cases. The distinction between the two cases is illustrated in Fig. 8.

An expression for  $V(C_l \setminus C_r)$  for the large leaf case is obtained by assuming the point at  $z$  is visible from all eye positions between the left and right eye. This gives the model

$$V_{\text{large}}(C_l \setminus C_r) = \frac{RT_x(z - z_0)^2}{z}. \quad (19)$$

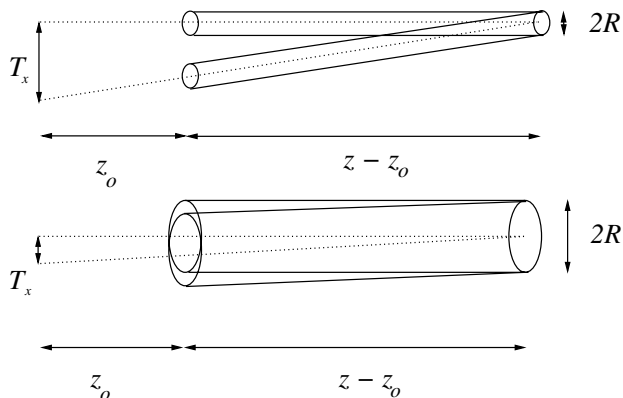


Fig. 8. A point at depth  $z$  is visible to both eyes if there is no surface center laying in the cylinders whose axes join the point to the two eyes. (Top) When surfaces are small relative to IOD, the two cylinders overlap only a little. (Bottom) When surfaces are large relative to IOD, the cylinders overlap almost entirely.

The large leaf model becomes exact as  $T_x/R \rightarrow 0$ . An expression for the small leaf case is obtained by approximating the two cylinders as disjoint, which gives the model

$$V_{\text{small}}(C_l \setminus C_r) = V(C_l) = \pi R^2(z - z_0). \quad (20)$$

The small leaf model becomes as  $R/T_x \rightarrow 0$ . For finite values of  $T_x$  and  $R$ , one can show that both the small and large leaf models underestimate the actual values. The differences between the models and data will be discussed below.

To compare the visibility in the two eyes, we consider the conditional probability that a point at depth  $z$  is visible to the left eye, given that it is visible to the right eye,

$$\begin{aligned} P(\text{left}|\text{right}; z) &= \frac{P(\text{left and right}; z)}{P(\text{right}; z)} = \frac{e^{-\eta V(C_l \cup C_r)}}{e^{-\eta V(C_r)}} \\ &= e^{-\eta V(C_l \setminus C_r)}. \end{aligned} \quad (21)$$

Figure 9 compares these conditional probabilities for two theoretical large and small leaf models from Eqs. (19) and (20) to data from simulations with the randomly oriented squares. We now include in the plots depth beyond the clutter. In particular, when we rendered the scenes, we included a slanted plane ( $75^\circ$  slant), which was placed beyond the clutter. (This plane was not shown in Fig. 2.) The plane was chosen to intersect the edge of the view volume at  $z_1 = 4$ . This plane allowed us to empirically measure the visibility of points beyond depth.

The IOD was again  $T_x = 0.05$ , which is approximately the width of the squares in the large leaf condition. We computed which points were visible to both eyes by rendering the scenes from the viewpoint of the right eye, and placing the light source at the position of the left eye. The depth buffer gave the  $z$  value at each pixel. If a scene point at depth  $z$  had non-zero intensity, it was visible to both eyes, whereas if the intensity was zero, then the surface was in shadow and so it was visible to the right eye but not the left eye. Conditional probabilities were estimated by taking the fraction of pixels at each depth bin  $z$  that were binocularly visible.

We expect the large leaf model to be a better fit than the small leaf model when the scene consists of larger leaves, and this is what we found. For the  $R = .04$  scenes, the large leaf model was a better fit, and for the  $R = .01$  scene, the small leaf model was a better fit. Note how the conditional probability falloff rate varies between the two scene types. Although the two types of scene have the same  $p(z)$  behavior (recall Fig. 4), the conditional probabilities decrease faster with depth in the small  $R$  case. For example, at  $z = 4$ , the conditional probability for the large square scenes is about 0.5, whereas it is about 0.2 for the small square scenes.

Cluttered scenes have fewer binocular points, which implies there are fewer points whose disparity can be measured directly by finding correspondences. While this obviously is a disadvantage for a vision system that is trying to infer the 3D geometry of the scene, there are advantages too; namely there are more scene points that are visible overall [3]. For many tasks such as recognition, it may be sufficient to see an object with just one of the two eyes. This raises the question of how much the overall visibility of the scene is increased by having the second eye.

The conditional probability model derived above can be used to express the probability that a binocular observer sees



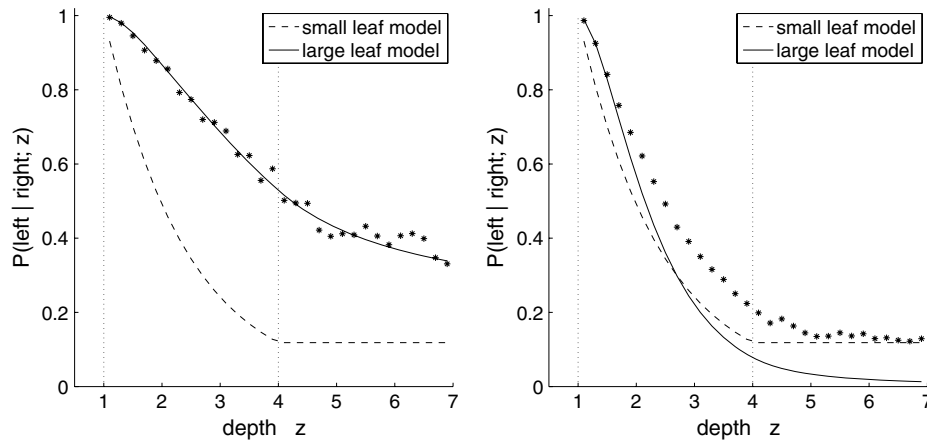


Fig. 9. Conditional probabilities that a surface at depth  $z$  is visible to left eye, given it is visible to right eye, for (left) large and (right) small square scenes. The vertical dotted lines mark the near and far boundaries of the clutter. The  $z$  values beyond the right dotted line are in the empty space behind the clutter.

a point with *at least* one of the two eyes. The probability model is similar to that of Changizi and Shimojo except that they modeled the scene clutter as a “wall of leaves” that is a partially occupied occluder plane. We are generalizing this model by letting the scene be a 3D volume that allows us to examine the increased visibility both within the cluttered volume as well as beyond it. The *binocular power* [3] of the viewer is the ratio of the probability that a point at depth  $z$  is visible to a binocular observer in either the left or right eye to the probability that the point is visible to the right eye:

$$power(z) \equiv \frac{P(\text{left and right}; z)}{P(\text{right}; z)}. \quad (22)$$

Using a simple Venn diagram argument, we have the relationship

$$P(\text{left or right}; z) = P(\text{left}; z) + P(\text{right}; z) - P(\text{left and right}; z). \quad (23)$$

Now substitute Eq. (23) into Eq. (22). Since

$$P(\text{left}; z) = P(\text{right}; z), \quad (24)$$

we get

$$power(z) = 2 - P(\text{left|right}; z), \quad (25)$$

which gives the power in terms of the conditional probability that was shown Fig. 9. For both the small and large  $R$  scenes,  $power(z)$  would rise from its theoretical minimum of 1 at the front of the cluttered volume to its theoretical maximum near 2 well beyond the cluttered volume. Note that power rises more quickly with depth in the small leaf scenes. Thus, even though the two scenes have the same monocular densities  $p(z)$ , the percentage of points that are visible to at least one of the two eyes is greater for small occluder scenes.

We also observe from Eq. (23) that

$$P(\text{right}; z) = \frac{P(\text{left or right}; z) + P(\text{left and right}; z)}{2}. \quad (26)$$

For a fixed  $P(\text{right}; z)$ , there is a direct tradeoff between the number of points that are visible to either eye and the number of points that are visible to both eyes. This tradeoff becomes

more significant as one goes further in depth. For example, consider the empirical data for the large  $R$  case in Fig. 9. Beyond the clutter, the probability that a point is visible to both eyes is less than 50% of the probability that it is visible to the right eye. It follows that the probability that the point is visible to either the left or the right eye is more than 50% greater than the probability that it is visible to the right eye, that is,  $power > 1.5$ . The effect is even more pronounced in the small  $R$  case.

## 7. EXAMPLE OF NONUNIFORM CLUTTER: A TREE

The uniform clutter model that we have used has the advantage that one can derive closed form expressions for how the visibility probabilities depend on scene parameters. However, this model does not capture some of the interesting properties of clutter in real scenes. For example, the spatial distribution of branches and leaves is nonuniform since it is determined by branching [5,6], which leads to clumping [7–10]. Another example is that clutter is often not restricted to thick slab volumes such as we have assumed.

To examine these nonuniformities, we repeated the experiments for several scenes using leafless tree models that were rendered using Blender [42]. We used ngPlant [43] to remove the foliage, leaving only the trunk and branches. Each tree was placed above a ground plane. The camera was tilted down so that the horizon appeared at  $25^\circ$  above the center of the image. A typical example is shown in Fig. 10. The tree trunk was centered at a distance of  $z = 2$  and the bulk of the branches ranged from  $z = 1$  to 3. The cropped image field-of-view was  $35^\circ$  in both horizontal and vertical direction, sampled by  $512 \times 512$  pixels. IOD was 0.05. Sixty views of the tree were rendered by rotating the tree by  $360^\circ$  about the vertical axis at  $6^\circ$  intervals. The plots showed data pooled over all views of the tree.

The uniformity assumptions do not hold exactly for tree scenes. A tree has roughly cylindrical symmetry, and the main trunk and branches of the tree are thicker near the trunk that lies near the image center. Moreover, the left and right peripheral regions of the image cover only the flanking regions of the cylinder and hence have a smaller range of depths. Also, the branch widths become smaller and more dense in the periphery. The  $p(z)$  and  $p(v)$  plots in Fig. 10 show that the

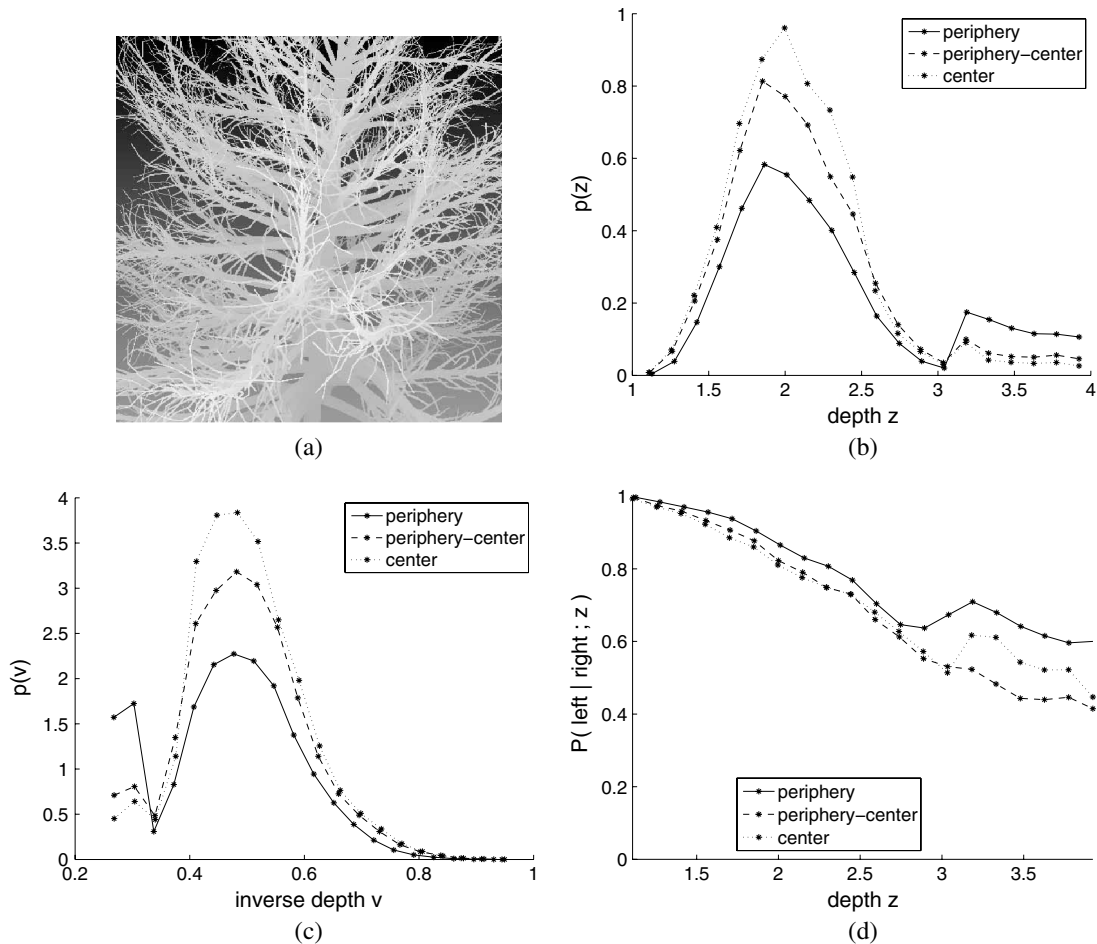


Fig. 10. (a) Range map where brighter means closer. (b)–(d) are  $p(z)$ ,  $p(v)$ , and  $P(\text{left}|\text{right}; z)$ . Each plot shows three curves corresponding to three different image regions. The image is partitioned into six vertical rectangles, each one-sixth of the image width. The central two rectangles define *center*, the two regions next to the center define *periphery-center*, and the outer two regions define *periphery*.

clutter is less dense as one moves from the image center to periphery.

Figure 11 shows the  $-\log$  of the joint probabilities  $P(z, z')$  and  $P(v, v')$ . Each plot consists of two regions. In one region, both  $z$  and  $z'$  correspond to points on the tree. In the other region, at least one of  $z$  or  $z'$  is on the slanted ground plane behind the tree. The closest point on the slanted ground plane is at about  $z = 3$ , which accounts for the visible discontinuity in the probabilities in the upper row, at the depth where the background plane suddenly comes into view. There was no such discontinuity in the earlier plots in Fig. 7 because those plots only included depth values in the clutter, not in the background.

Probabilities  $P(z, z')$  peak near the depth of the center of the tree,  $z \approx 2$ , i.e.,  $v = 0.5$ . As in Fig. 7, the probabilities are greatest on the diagonals. The probabilities are slightly elevated on the first off-diagonal, which is due to binning effects discussed earlier. This effect does not occur on the ground plane, however, since we plot horizontal neighbors only. When we include statistics for vertical neighbors, we see probability mass in the first off-diagonals where the  $z$  values of neighboring pixel straddle a  $z$  bin boundary because of the vertical slant of the plane (verified, but not shown here).

Finally, probabilities are nonuniform in the off-diagonal regions. Within each plot, probabilities are greater near the

depth of the tree trunk as expected from  $p(z)$  curves in Fig. 10. Between plots, probabilities of discontinuities are slightly lower in the periphery than in the center (compare left and right columns of Fig. 11), as expected since this just follows the  $p(z)$  and  $p(v)$  curves.

## 8. DISCUSSION AND CONCLUSION

Visibility is a fundamental issue in 3D cluttered scenes. Clutter complicates the task of inferring depth and detecting and recognizing objects, since objects can become partly occluded and more difficult to detect and to segment from their surroundings. Clutter also makes binocular stereopsis more difficult. 3D clutter tends to reduce the number of binocular points. This decreases the number of points with well-defined disparity and also increases the complexity of solving for correspondences between images.

Our main goal in this paper was to model visibility probabilities in 3D cluttered scenes and to understand how visibility depends on depth within or beyond the clutter, and on properties of the clutter, namely the size and density of occluders. We first examined the distribution of gap sizes, where “gap” refers to the space between the occluders that make up the clutter. The distribution of gaps at a depth  $z$  determines the sizes of the fragments of an object that would be visible if

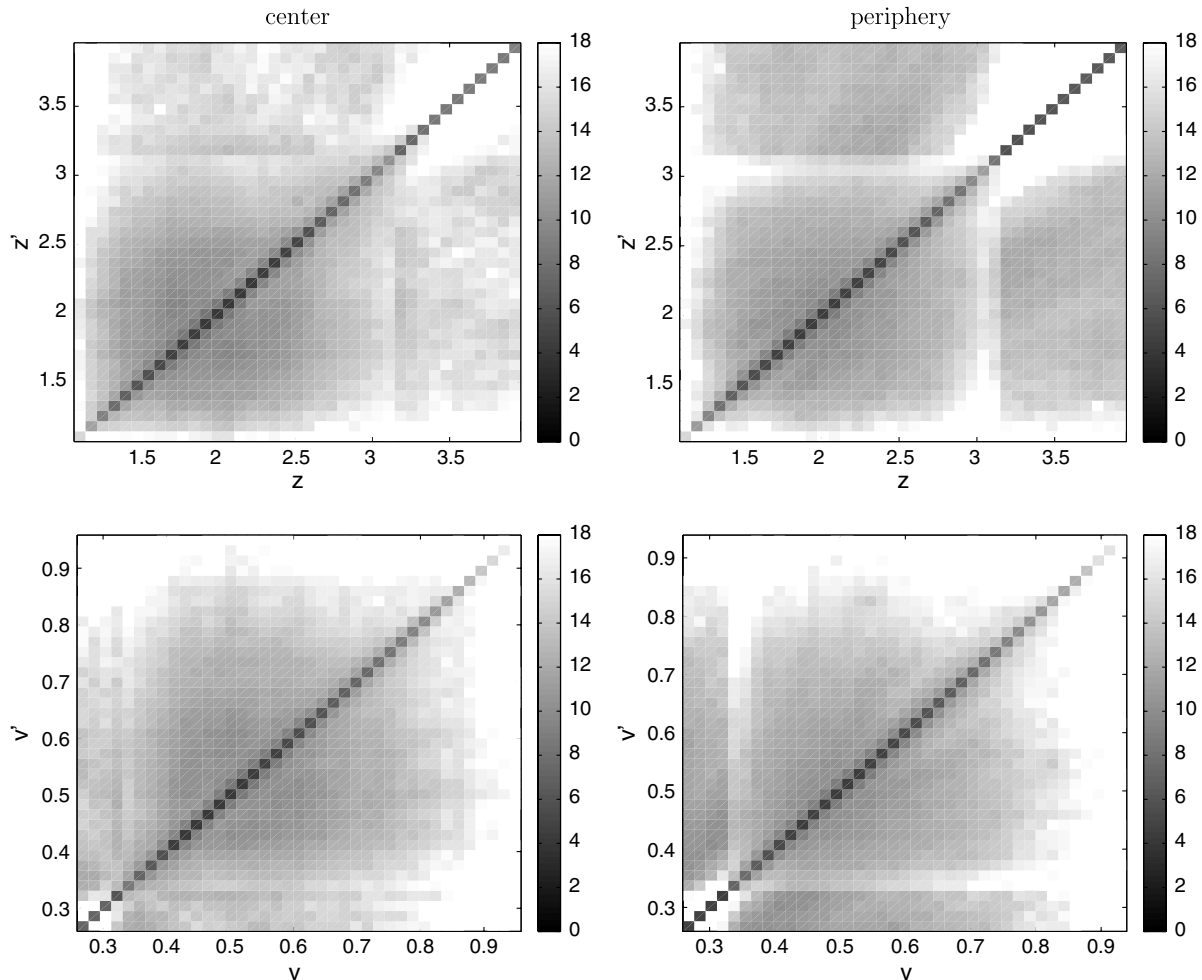


Fig. 11. Joint probabilities for a tree scene. (Top) Plots of  $-\log P(z, z')$ . (Bottom) Plots of  $-\log P(v, v')$ . (Left) Central third of the image. (Right) Left and right peripheral regions of the image. White regions correspond to points with 0 probability, i.e., infinite  $-\log$  probability.

the object were at depth  $z$ . How the visual system completes objects that are partly occluded has been investigated by many [44], but most studies have considered simple scenes only. We showed how, as an object descends into a cluttered volume, the distribution of gaps in the clutter through which the object is visible becomes dominated by small gaps.

The second visibility problem addressed the depths of surfaces seen at neighboring pixels in an image, and whether surfaces lie on the same or different surface. We examined the probability of a change in discrete depth and inverse depth that arises when neighboring pixels straddle a depth bin boundary on a surface that is slanted in depth. We argued that the results are consistent with stereo models used in computer vision that penalize disparity changes between neighboring pixels.

The third visibility problem addressed binocular vision. We compared the probability that a scene point is binocularly visible to the probability that it is visible by at least one of the two eyes. We showed that there is a tradeoff between the visibility to both eyes and the visibility to either eye. This tradeoff is relevant for tasks of depth estimation from stereo (which benefits from binocular visibility) and the task of detection and recognition of objects (which presumably benefits from having more surface points visible to either eye).

We also carried out experiments using a synthetic model of a leafless tree where the clutter was nonuniform over space and so the theoretical model of uniform clutter does not apply exactly. For example, because the tree has cylindrical symmetry, the range of depths of the occluders is different in the flanking regions of the tree than near the center of the tree.

The above example is just one of the ways in which the model that we have developed could be generalized to better fit real scenes. As another example, Hansard [45] recently proposed a model in which the distance  $z_0$  to depth plane where the clutter begins was a random variable rather than constant. This generalization gave him a good fit to the range data from the Brown dataset [17]. He also rewrote the binocular visibility model such that the probabilities depended on visual direction rather than depth. He showed that binocular visibility probabilities are quite different in directions far from the  $z$  axis, i.e., in the periphery. In our model, visual direction did not play much of a role since we considered only small field-of-view images and thus we only considered directions near the  $z$  axis.

We next turn to a different question of how results of our study and other related studies could be used. One example is in understanding human perception of 3D clutter. The approach could be analogous to the classical shape from texture problem. That problem addresses how humans perceive the

depth gradient of a smooth surface and how well observers use available cues such as size, density, and foreshortening [46]. For 3D cluttered scenes, these cues are present as well, but now the cues indicate a range of depths in a volume. It would be interesting to examine how well human perception of cluttered volumes makes use of each of these cues. The models we have derived could potentially be used to define ideal observers against which human performance could be compared. Similarly, shape-from-X cues such as motion parallax and stereo, which have been well studied for uncluttered scenes, could also be studied in the context of cluttered scenes [47].

Another avenue for research is to ask how (or if) these probability models may be used to improve the performance of state-of-the-art stereo algorithms in computer vision. One approach might be to try to estimate the probabilities for a given stereo pair by applying a state-of-the-art algorithm as a first pass and then fitting a simple model of the probabilities from the computed disparity map. The model could then be used to choose new penalties, both for visibility (half occlusions) and for smoothness (disparity discontinuities). Such a two-step approach has been fruitful before [48], and could be interesting for stereo in cluttered scenes as well.

The models might also be useful in visualization applications where one wishes to examine a 3D dataset by rendering it [49]. Examples include medical images such as MRI or angiograms of blood vessel branching patterns, or scientific visualization of flow lines in a fluid. Such volume visualization methods often use stereo or motion cues to improve the perception of surfaces. The visibility models developed in this paper might be useful for automatically choosing the parameters of such visualizations such that one maximizes the visibility of surfaces.

To conclude, this paper has provided an analysis of surface visibility in 3D cluttered volumes. We developed quantitative models of how visibility depends on depth and on the properties of clutter. We examined the distributions of gaps through which a surface can be seen through clutter, the distributions of depth discontinuities in single images, and the distribution of binocular and half-occluded points. These results could help future researchers gain a more fundamental understanding of the constraints, challenges, and limitations of vision in 3D cluttered scenes.

## APPENDIX A: DERIVATION OF MODELS FROM SECTION 5

Here we derive expressions for the joint probabilities of the depths and inverse depths at neighboring pixels. We assume

the scene consists of frontoparallel disks of radius  $R$  as in Section 3. Take a pixel  $\mathbf{x}$  and suppose there is a disk that is visible at  $\mathbf{x}$  and this disk has depth  $z$ . Let  $\mathbf{x}'$  be a neighboring pixel and let its depth be  $z'$ . Let  $\mathbf{x}$  and  $\mathbf{x}'$  lie in the projection plane  $z = 1$ , and so the 3D position seen at  $\mathbf{x}$  is  $z\mathbf{x}$  and the 3D position of the point seen at  $\mathbf{x}'$  is  $z'\mathbf{x}'$ .

For the points seen at  $\mathbf{x}$  and  $\mathbf{x}'$ , there are three possibilities that are illustrated in Fig. 12 (left).

1.  $\mathbf{x}'$  lies on a closer disk and so  $z' < z$  (see  $\mathbf{x}_1, \mathbf{x}'_1$ ),
2.  $\mathbf{x}'$  lies on a further disk or beyond the clutter, and so  $z' > z$  (see  $\mathbf{x}_2, \mathbf{x}'_2$ ),
3.  $\mathbf{x}$  and  $\mathbf{x}'$  see the same disk and so  $z = z'$  (see  $\mathbf{x}_3, \mathbf{x}'_3$ ).

Let  $C_{\mathbf{x}'}$  and  $C_{\mathbf{x}}$  be the cylinders that are centered on the rays through  $z\mathbf{x}$  and  $z'\mathbf{x}'$ .

Case 1:  $z < z'$

For  $\mathbf{x}'$  to see a closer disk, two events must occur. First, the pixel  $\mathbf{x}$  must see a disk in depth interval  $[z, z + dz]$ . The probability density of this event is obtained from Eq. (6). Second, the pixel  $\mathbf{x}'$  must see a different disk whose depth  $z'$  is less than  $z$ . Thus, the set difference of the cylinders  $C_{\mathbf{x}'} \setminus C_{\mathbf{x}}$ , which has volume  $R(z'^2 - z_0^2)|\mathbf{x} - \mathbf{x}'|$ , cannot contain a disk center, and the cap at depth  $z'$ , which has volume  $2R(z' - z_0)|\mathbf{x} - \mathbf{x}'|dz'$ , must contain a disk center, namely the disk seen by  $\mathbf{x}'$ . The probabilities of these two events are obtained similarly to Eq. (6). The joint probability below is obtained by multiplying together the probabilities of the two events.

$$p_{\text{diff}}(z, z') = 2\eta^2 \pi R^3 (z' - z_0) |\mathbf{x} - \mathbf{x}'| \exp\{-\eta(\pi R^2(z - z_0) + R(z'^2 - z_0^2)|\mathbf{x} - \mathbf{x}'|)\}. \quad (\text{A1})$$

We obtain  $p_{\text{diff}}(v, v')$  by substituting as in Eq. (10) for the case  $v < v'$ :

$$p_{\text{diff}}(v, v') = \frac{2\eta^2 \pi R^3 |\mathbf{x} - \mathbf{x}'|}{v^2 v'^3} \exp\left\{-\eta\left(\pi R^2\left(\frac{1}{v} - \frac{1}{v_0}\right) + R\left(\frac{1}{v'^2} - \frac{1}{v_0^2}\right)|\mathbf{x} - \mathbf{x}'|\right)\right\}. \quad (\text{A2})$$

Case 2:  $z' < z$

In this case, the joint probabilities are obtained by swapping  $z, z'$  in Eq. (A1) and  $v, v'$  in Eq. (A2).

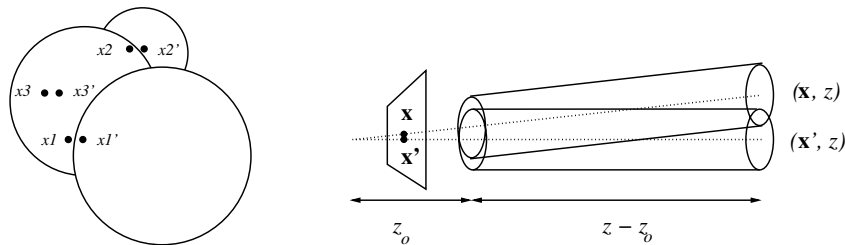


Fig. 12. (Left) For any point that lies on a disk at depth  $z$  and is visible at pixel  $\mathbf{x}$ , a neighboring pixel  $\mathbf{x}'$  may or may not see that same disk. Three possible cases are shown. See text. (Right) Given that the surface at pixel  $\mathbf{x}$  has depth  $z$ , the probability that the neighboring scene point  $(\mathbf{x}', z)$  is occluded is determined by the volume of  $C_{\mathbf{x}'} \setminus C_{\mathbf{x}}$  according to the Poisson model.



Case 3:  $z = z'$

For  $\mathbf{x}'$  to see the same disk as  $\mathbf{x}$ , three conditions must be met. First, as in case 1, the disk seen at  $\mathbf{x}$  lies in the interval  $[z, z + dz]$ . The probability density of this event is obtained from Eq. (6). Second,  $\mathbf{x}$  and  $\mathbf{x}'$  must lie within the image projection of a single disk. Recall that  $\mathbf{x}$  and  $\mathbf{x}'$  are separated horizontally. When  $|\mathbf{x} - \mathbf{x}'|$  is small, the probability that  $\mathbf{x}$  and  $\mathbf{x}'$  lie on the same disk is

$$\frac{\pi R^2 - 2R(\mathbf{x} - \mathbf{x}')z}{\pi R^2},$$

where the numerator is approximately the area of a crescent of the disk containing  $\mathbf{x}$ , such that  $\mathbf{x}'$  does not lie in this disk. Third, case 1 cannot hold, namely that  $\mathbf{x}'$  is not occluded by a nearer disk. The probability of this event, given the first condition, is

$$\exp\{-\eta(\pi R^2(z - z_0) + R(z^2 - z_0^2)|\mathbf{x} - \mathbf{x}'|)\}.$$

Multiplying the probability densities of these three events gives

$$p_{\text{same}}(z, z) = \eta(\pi R^2 - 2R|\mathbf{x} - \mathbf{x}'|z) \exp\{-\eta(\pi R^2(z - z_0) + R(z^2 - z_0^2)|\mathbf{x} - \mathbf{x}'|)\}. \quad (\text{A3})$$

To obtain the joint density on inverse depth, we substitute  $1/v$  for  $z$ . A factor  $1/v^2$  must also be included because of the change in variables, giving

$$p_{\text{same}}(v, v) = \frac{\eta}{v^2} \left( \pi R^2 - \frac{2R}{v} |\mathbf{x} - \mathbf{x}'| \right) \times \exp \left\{ -\eta \left( \pi R^2 \left( \frac{1}{v} - \frac{1}{v_0} \right) + R \left( \frac{1}{v^2} - \frac{1}{v_0^2} \right) |\mathbf{x} - \mathbf{x}'| \right) \right\}. \quad (\text{A4})$$

Note that substituting  $\mathbf{x} = \mathbf{x}'$  gives the marginal distributions  $p(z)$  and  $p(v)$  from Eqs. (6) and (10).

## ACKNOWLEDGMENTS

This research was supported by grants to M. Langer from the Natural Sciences and Engineering Research Council of Canada (NSERC) and Fonds de recherche du Québec—Nature et technologies (FQRNT).

## REFERENCES

- R. Rosenholtz, Y. Li, and L. Nakano, "Measuring visual clutter," *J. Vision* **7**, 1–22 (2007).
- J. M. Wolfe, "Guided search 2.0: a revised model of visual search," *Psychon. Bull. Rev.* **1**, 202–238 (1994).
- M. A. Changizi and S. Shimojo, "X-ray vision and the evolution of forward-facing eyes," *J. Theor. Biol.* **254**, 756–767 (2008).
- T. E. Avery and H. E. Burkhart, *Forest Measurements*, 5th ed. (McGraw-Hill, 2002).
- P. Prusinkiewicz and A. Lindenmayer, *The Algorithmic Beauty of Plants* (Springer-Verlag, 1990).
- P. Prusinkiewicz, "Modeling of spatial structure and development of plants: a review," *Scientia Horticulturae* **74**, 113–149 (1998).
- T. Nilson, "A theoretical analysis of the frequency of gaps in plant stands," *Agric. Meteor.* **8**, 25–38 (1971).
- P. R. van Gardingen, G. E. Jackson, S. Hernandez-Daumas, G. Russell, and L. Sharp, "Leaf area index estimates obtained for clumped canopies using hemispherical photography," *Agric. Forest Meteor.* **94**, 243–257 (1999).
- J. M. Chen and J. Cihlar, "Plant canopy gap-size analysis theory for improving optical measurements of leaf-area index," *Appl. Opt.* **34**, 6211–6222 (1995).
- H. Sinoquet, G. Sonohat, J. Phattaralerphong, and C. Godin, "Foliage randomness and light interception in 3-d digitized trees: an analysis from multiscale discretization of the canopy," *Plant Cell Environ.* **28**, 1158–1170 (2005).
- D. L. Ruderman and W. Bialek, "Statistics of natural images: scaling in the woods," *Phys. Rev. Lett.* **73**, 814–817 (1994).
- G. Matheron, *Random Sets and Integral Geometry* (Wiley, 1975).
- J. P. Serra, *Image Analysis and Mathematical Morphology* (Academic, 1982).
- D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," *J. Opt. Soc. Am. A* **4**, 2379–2394 (1987).
- D. L. Ruderman, "Origins of scaling in natural images," *Vis. Res.* **37**, 3385–3398 (1997).
- R. M. Balboa, C. W. Tyler, and N. M. Grzywacz, "Occlusions contribute to scaling in natural images," *Vis. Res.* **41**, 955–964 (2001).
- J. Huang, A. Lee, and D. Mumford, "Statistics of range images," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2000), pp. 324–331.
- A. Lee, D. Mumford, and J. Huang, "Occlusion models for natural images: a statistical study of a scale-invariant dead leaves model," *Int. J. Comput. Vis.* **41**, 35–59 (2001).
- Z. Yang and D. Purves, "Image/source statistics of surfaces in natural scenes," *Network Comput. Neural Syst.* **14**, 371–390 (2003).
- B. Potetz and T. S. Lee, "Statistical correlations between two-dimensional images and three-dimensional structures in natural scenes," *J. Opt. Soc. Am. A* **20**, 1292–1303 (2003).
- D. Calow and M. Lappe, "Local statistics of retinal optic flow for self-motion through natural sceneries," *Network Comput. Neural Syst.* **18**, 343–374 (2007).
- S. Roth and M. J. Black, "On the spatial statistics of optical flow," *Int. J. Comput. Vis.* **74**, 33–50 (2007).
- P. B. Hibbard, "A statistical model of binocular disparity," *Vis. Cogn.* **15**, 149–165 (2007).
- Y. Liu, A. Bovik, and L. Cormack, "Disparity statistics in natural scenes," *J. Vision* **8**(2):17, 1–22 (2008).
- J. Harris and L. Wilcox, "The role of monocularly visible regions in the perception of three-dimensional scenes," *Vis. Res.* **49**, 2666–2685 (2009).
- J. Forte, J. W. Peirce, and P. Lennie, "Binocular integration of partially occluded surfaces," *Vis. Res.* **42**, 1225–1235 (2002).
- P. Belhumeur, "A Bayesian approach to binocular stereopsis," *Int. J. Comput. Vis.* **19**, 237–260 (1996).
- S. Kang, R. Szeliski, and J. Chai, "Handling occlusions in dense multi-view stereo," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2001), pp. 1:103–110.
- G. Egnal and R. Wildes, "Detecting binocular half-occlusions: Empirical comparisons of five approaches," *IEEE Trans. Pattern Anal. Mach. Intell.* **24**, 1127–1133 (2002).
- V. Kolmogorov and R. Zabih, "Computing visual correspondence with occlusions via graph cuts," in *Proceedings of 8th IEEE International Conference on Computer Vision* (IEEE, 2001), pp. II:508–515.
- Y. Wei and L. Quan, "Asymmetrical occlusion handling using graph cut for multi-view stereo," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 2 (IEEE, 2005), pp. 902–909.
- V. Vaish, M. Levoy, R. Szeliski, C. L. Zitnick, and S. B. Kang, "Reconstructing occluded surfaces using synthetic apertures: stereo, focus and robust measures," in *IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2006), pp. 2331–2338.
- A. Mittal and L. Davis, "A general method for sensor planning in multi-sensor systems: extension to random occlusion," *Int. J. Comput. Vis.* **76**, 31–52 (2007).

34. P. Hall, *Introduction to the Theory of Coverage Processes* (Wiley, 1988).
35. S. Zacks, *Stochastic Visibility in Random Fields, Lecture Notes in Statistics 95* (Springer-Verlag, 1994).
36. B. Nadler, G. Fibich, S. Lev-Yehudi, and D. Cohen-Or, "A qualitative and quantitative visibility analysis in urban scenes," *Comput. Graphics* **23**, 655–666 (1999).
37. J. H. van Hateren, "Theoretical predictions of spatiotemporal receptive fields of fly LMCs, and experimental validation," *J. Comp. Physiol. A* **171**, 157–170 (1992).
38. M. Langer, "Surface visibility probabilities in 3d cluttered scenes," in *Proceedings of 10th European Conference on Computer Vision* (Springer-Verlag, 2008), pp. I:401–412.
39. H. Longuet-Higgins and K. Prazdny, "The interpretation of a moving retinal image," *Proc. R. Soc. B* **208**, 385–397 (1980).
40. G. W. Larson and R. Shakespeare, *Rendering with Radiance: The Art and Science of Lighting Visualization* (Morgan Kaufmann, 1998).
41. R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother, "A comparative study of energy minimization methods for Markov random fields with smoothness-based priors," *IEEE Trans. Pattern Anal. Machine Intell.* **30**, 1068–1080 (2008).
42. "The blender greenhouse," <http://yorik.uncreated.net/greenhouse.html>.
43. "Ngplant—open source plant modeling suite," <http://ngplant.sourceforge.net/>.
44. G. Kanizsa, *Organization in Vision: Essays on Gestalt Perception* (Praeger, 1979).
45. M. Hansard, "Binocular projection of a random scene," in *Proceedings of British Machine Vision Conference* (Springer-Verlag, 2012), to be published.
46. D. C. Knill, "Surface orientation from texture: Ideal observers, generic observers and the information content of texture cues," *Vis. Res.* **38**, 1655–1682 (1998).
47. R. van Ee and B. Anderson, "Motion direction, speed, and orientation in binocular matching," *Nature* **410**, 690–694 (2001).
48. L. Zhang and S. M. Seitz, "Estimating optimal parameters for MRF stereo from a single image pair," *IEEE Trans. Pattern Anal. Mach. Intell.* **29**, 331–342 (2007).
49. K. Engel, M. Hadwiger, J. M. Kniss, and D. W. C. Rezk-Salama, *Real-Time Volume Graphics* (AK Peters, 2006).