In W. Epstein & S. Rogers (eds.) *Handbook of perception and cognition, Vol 5; Perception of space and motion*. (pp. 69-117). San Diego, CA: Academic Press.

# Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth

James E. Cutting and Peter M. Vishton

The layout in most natural environments can be perceived through the use of nine or more sources of information. This number is greater than that available for the perception of any other property in any modality of perception. Oddly enough, how perceivers select and / or combine them has been relatively unstudied. This chapter focuses briefly on the issues inhibiting its study, and on what is known about integration, then in detail on an assessment of nine sources of information—occlusion, relative size, relative density, height in the visual field, aerial perspective, motion perspective, binocular disparities, convergence, and accommodation—and their relative utility at different distances. From a comparison of their ordinal depth-threshold functions, we postulate three different classes of distance around an observer--personal space, action space, and vista space. Within each space, we suggest a smaller number of sources act in consort, with different relative strengths, in offering the perceiver information about layout. We then apply this system to the study of representations of layout in art, to the development of the perception of layout by infants, and to an assessment of the scientific study of layout.

> In general, the visual world approaches the Euclidean ideal of veridical perception as the quantity and quality of perceptual information increases.
>
> Wagner (1985, p. 493)

How do we see and understand the layout of objects in environments around us? This question has fascinated artists and philosophers for centuries. Indeed, the long-term cultural success of painting, and more recently of photography and cinema, is largely contingent on their convincing portrayal of spatial relations in three dimensions. Of course, the information about spatial layout available in pictures and in the real world has also fascinated psychologists for 125 years. Indeed, the problems of perceiving and understanding space, depth, and layout were among those which helped forge our discipline.[1] Moreover, their sustaining importance is reflected in this volume.

Perhaps the most curious fact about psychological approaches to the study of layout is that its history is little more than a plenum of lists. Lists have been generated since Hering and Helmholtz (see, for example, Boring, 1942; Carr, 1935; Gibson, 1950; Graham, 1951; Woodworth & Schlosberg, 1954), and are found today in every textbook covering the field. These lists typically include a selection of the following: accommodation, aerial perspective, binocular disparity, convergence, height in the visual field, motion perspective, occlusion, relative size, relative density, and often many more. As an entrée into the field, such a list is prudent and warranted. Taxonomy is, after all, the beginning of science, of measurement, and is an important tool in pedagogy. What is most remarkable about such a list is that in no other perceptual domain can one find so many generally intersubstitutable information sources, all available for the perception of the same general property--the layout of objects in the environment around us.[2] But after 125 years more than list making is needed for our science.

Three fundamental questions about the perception of our surrounding world are, to our minds, too seldom asked. All are prior to the summary statement by Wagner (1985), given above. First, why does the human visual system make use of such a wide variety of sources of information—often called *cues*?[3]—in understanding and in deriving the structure in depth of a complex natural scene. Second, how do we come to perceive the three-dimensional layout of our environment with reasonable, even near metric, accuracy when taken singly none of the visual sources of information yields metric information throughout the range of distances we need? And third, given all these sources, does anything lay beyond list making which might allow us to begin to understand why so many sources are used and are necessary?

The purpose of this chapter is to attempt to answer all three questions. In Wagner's terms we want to be concrete about the quantity and quality of information about depth, and to provide the framework for a theory of environmental context and physiological state supporting the perception of layout. We will claim also that these three questions are linked in important ways, but before addressing them we must acknowledge that historically psychology has had serious qualms with the formulation of the second. Two such qualms have dominated discussion of the topic, those pertaining to illusions and to the non-Euclidean nature of perceived space.

## Illusions of Layout

First, it is not overly difficult to fool the eye even of a wizened observer about the layout of some small corner of the world. If one has sufficient skill, one can paint *trompe l'oeil* or carpenter nonrectilinearly joined surfaces, place individuals in confined viewpoints so they can peer at those surfaces or into those spaces, and then demonstrate that those observers do not correctly understand the layout of what they see. Indeed, the Pozzo ceiling (Pirenne, 1970) and the Ames room (Ames, 1955; Ittelson, 1951) are compelling and important examples of such large-scale visual illusions of layout. In particular, under certain controlled circumstances we perceive the Ames room as it is not, conventionalizing its layout to that of a rectilinear chamber. The typical theoretical statement about its perception is a variant of a traditional argument in philosophy, *the argument from illusion* (e.g. Hirst, 1967). That argument states that if the senses can be deceived, how can we be sure they do not deceive us all the time? The particular answer given for the Ames room illusion by transactionalists concerns the useful, but occasionally misleading, idea that knowledge and experience necessarily intervene in perception.

Whereas one would be foolish to deny the utility of knowledge and experience as it can shape perception, current perceptual theory suggests that that molding process is not omnipotent. Moreover, the physical world is hardly the plastic and ambiguous place that the transactionalist psychologists might have had us believe. For example, if the Ames room illusion to deceive the eye, the perceiver must be severely fettered, typically monocular and motionless. As demonstrated by Gehringer and Engel (1986), when the observer is allowed to use two eyes and to move even a modest amount the illusion almost completely disappears (see also Runeson, 1988).

Thus, one should conclude that when perception of an environment's layout is important to us--and we would claim it almost always is--our perception of that layout is typically adequate, even near-veridical, when exploration is allowed. Therefore, in the place of the argument from illusion, we offer a pragmatic retort, which one might call the *argument from evolution* (see also Cutting, 1993). This argument states that if the senses were deceived too frequently, we as a species would surely have, without extensive divine guidance, become extinct long ago. The pragmatics of potential failure over the long span of every evolutionary test of our senses renders null the possibility of their gross defectiveness. Thus, we suggest the measurement abilities of the senses are typically as good as they need to be under everyday requirements (see Cutting, Springer, Braren, & Johnson, 1992b), and for measuring layout we think they need to be quite good.

Gehringer and Engel's (1986) result presages a main point of this chapter: Through the combination of multiple sources of information--motion parallax, stereopsis, etc.--in a cluttered, well-lit environment we come to perceive, with reasonable accuracy, the layout of the world around us. Of course, the key phrase here is "with reasonable accuracy." As a working assumption we will claim that temporary, and relatively small, errors in judgment of distance—say, those of up to 15%—are not likely to have consequence in normal, day-to-day livelihood; relatively large errors, on the other hand,—say, those of an order of magnitude or even of simply 50%—might begin to tax our survival.

**Euclidean and non-Euclidean Layout**

The second qualm about the accuracy of perceived layout concerns the "shape" of space. Physical and perceptual layouts have not always been modeled by mathematicians and psychologists in an Euclidean manner (see, for example, Luneburg, 1947; Blank, 1978). Even in moderately complex environments straight lines can be perceived as curved (Battro, Netto, & Rozestraten, 1976; Helmholtz, 1866; Indow, 1982; Indow & Watanabe, 1984). Indeed, anyone with relatively high-powered glasses (1.5 or more diopters) can, upon taking them off and putting them back on, verify this fact by looking near and along the corners of a room. Such curvature, whether seen binocularly or monocularly, occurs in the periphery of the eye(s). Assessments of such curvature have often been placed in a Riemannian framework and used in accounting for the perception of both three-dimensional environments, and illusions in two-dimensional images (Watson, 1978). Such results have led philosophers to argue whether visual space is Euclidean or non-Euclidean (e.g., Daniels, 1974; Grünbaum, 1973; Putnam, 1963; Suppes, 1973).

Perhaps the most damning empirical evidence against systematic accounts of non-Euclidean (curved) space is that, when individual perceivers' spatial judgments are modeled, they vary widely in curvature (Foley, 1964, 1966; but see also Wagner, 1985, for other problems). To accept curved perceptual space, then, is also to accept that each of us lives in a perceived world with different curvature. Embracing such a view leads one to the philosophical antinomy more typical in discussions of language—that of how private minds with quite different ideas can communicate about shared experiences.

We think the best solution to the problem of curved space is an elaboration of the pragmatic approach of Hopkins (1973): The degree of curvature of physical space is so small that, on a local level, it cannot be decided whether it and phenomenal space, which generally follows physical space, are Euclidean or only a good approximation thereof (see also Cutting, 1986). Again, notice the implication of "reasonable accuracy," or in this case reasonable Euclidean approximation. We strongly suspect that with modest exploration on the part of any observer, the curvature modeled to his or her perceptual judgments of layout will vary both with the environment modeled and with the particular individual perceiving it, but that the overall mean curvature of all terrestrial spaces we might find ourselves in is very nearly zero (see also Wagner, 1984). This would mean that perceptual spaces are sufficiently close to being Euclidean that little of practical import is gained by considering Riemannian curvatures, perhaps with the exception of understanding how a Riemannian space becomes a Euclidean space with additional sources of information.

## ON THE ACCURACY OF PERCEIVED SPACE

> An observer's ability to perceive distance varies with number of circumstances, most prominent of which is the degree to which the experimental situation makes available information for distance.
>
> Sedgwick (1986, p. 22-8)

As implied by Sedgwick (1986) it is not difficult in the laboratory to make the perception of layout difficult. Escaping artificial confinements and following general tenets of Gibson (1950, 1979), then, many researchers have conducted studies using various experimental techniques in large outdoor spaces. Most of these suggest that our ability to perceive layout and distances is quite good in situations with many sources of information. We appear to have rough metric knowledge in making distance judgments (E. Gibson & Bergman, 1954) which improve with feedback; relative distance judgments, on the other hand, do not appear to improve with feedback (Wohlwill, 1964).

### Distance Estimation by Stationary Observers

In psychophysical terms, the power function for distance judgments is well-described by an exponent with the value near 1.0 (see Baird & Biersdorf, 1967; Cook, 1978; DaSilva, 1985; Purdy & Gibson, 1955; Teghtsoonian & Teghtsoonian, 1969; see also Flückiger, 1991); such a result means that metric relations are fairly well perceived and preserved. Indeed, in keeping with the notion of "reasonable accuracy" Cook's (1978) data show stable exponents of individuals varying between about 0.78 and 1.22, with a mean close to 0.95; and with several different methods, Da Silva (1985) reported ranges of about 0.60 to 1.30 and means around 0.94.

An important fact about results from distance judging experiments is that mean egocentric depth (distance away from the observer) is systematically foreshortened when compared to frontal depth (distances extended laterally in front of the observer, orthogonal to a given line of sight).[4] Thus, even with an exponent of 0.95 objects at egocentric distances of 10, 100, and 1000 m would be seen to be at only 9, 79, and 710 m, respectively; whereas frontal distances would tend to show no such diminution. These egocentric foreshortenings represent 10, 21, and 29% errors in judgment, respectively, and might seem to impugn a claim of "reasonable accuracy" in perceiving layout, particularly in the great distance. However, there are two sets of facts available suggesting we generally have a more accurate perception of layout than even these experiments suggest.

### Distance Estimation by Moving Observers

In defense of the idea of reasonable accuracy the first set of facts comes from research on visually directed action. Thomson (1980), Laurent and Cavallo (1985), and Rieser, Ashmead, Talor, and Youngquist (1990), for example, showed that individuals are quite accurate in walking blindly to the distance of an object previously seen--indicating little, if any, foreshortening. Moreover, Loomis, DaSilva, Fujita, and Fukushima (1992) combined tasks in the same subjects, and found the typical foreshortening effect in the distance estimation task and general accuracy in a directed action task. Thus, with Loomis et al, we conclude that acting on static visual information yields accurate estimates of space, but numerical or algebraic estimates in the same space often do not.

A second important fact concerns motion information available to a moving observer. In standard psychophysical distance estimation tasks the exponents reported are for stationary observers. By allowing an observer to move more than minimally with eyes open, the egocentric origin of the foreshortened space (be it affine or vectorially defined, see Wagner, 1985) must also change. Such changes would make any simple foreshortening model of perceived distances unlikely, and would necessarily transform the space to a more Euclidean format.

### Perceived Interobject Distances Are Multiply Constrained

Another avenue of research supporting accurate perception of depth and layout concerns the following kind of experiment: Individuals are placed in a cluttered environment and then asked make judgments of distances among the various objects in that space (e.g. Kosslyn, Pick & Fariello, 1974; Toye, 1986). The half-matrix of interobject judgments is then entered into a multidimensional scaling program (e.g. Kruskal, 1964; Shepard, 1980), and the two-dimensional solution compared to the real layout in which observers made their judgments. Toye (1986), for example, had observers judge the distances between all 78 possible pairs of 13 posts in a courtyard set among four tall buildings. These judgments were treated nonmetrically (that is, as ranked information) and scaled in two dimensions. The original and the best-fitting derived solutions revealed detailed and accurate correspondence between the two. Since absolute distances were judged with reasonable accuracy, the overlap of the two spatial representations is correct in scale as well as in general configuration.

We take Toye's (1986) methods and results as an important analogy for everyday commerce in our environment. That is, judgments about the distances among objects in a cluttered environment can vary, even vary widely. However, when taken together they constrain each other in a manner well-captured by nonmetric multidimensional scaling (NMDS) procedures (see also Baird, Merrill, & Tannenbaum, 1979; Baird & Wagner, 1983). Any psychological process for understanding the layout of objects in a visual scene would do well to mimic such a multiple-constraints procedure. Incidental, anomalous over- and underestimates of distance would be then corrected through consideration of other distances between various objects under consideration.

## INFORMATION INTEGRATION: RULES, WEIGHTS, AND THE FEASIBILITY OF COMPLETE EXPERIMENTATION

The are two general empirical phenomena associated with experiments which have viewers estimate depth and which vary the number of sources of information available about layout. Künnapas (1968) found both. That is, adding information to a stimulus display generally increases the amount of depth seen, and adding information generally increases the consistency and accuracy with which judgments are made. The latter effect is surely the more important, but most investigations have focused on the former.

### Nearly Linear Systems, Combination Rules, Cooperativity, and Data Fusion

An early indication that more depth is often seen with more sources of information stems from research by Jameson and Hurvich (1959). Kaufman (1974) called this the *superposition principle* from linear systems theory. More recently, Bruno and Cutting (1988) investigated the perception of exocentric depth from four sources of information: occlusion, relative size, height in the visual field, and motion perspective. By manipulating the image of three vertically oriented and parallel planes, Bruno and Cutting (1988) orthogonally varied the presence and absence of these four sources of information. Using a variety of techniques (direct scaling, preference scaling, and dissimilarity scaling) they found evidence of nearly linear additive combination of the four sources of information. That is, although different viewers weighted the four different sources differently, each source was generally intersubstitutable and with each addition generally yielding the perception of more exocentric depth.

After reanalyzing the direct scaling data of Bruno and Cutting (1988), Massaro (1988) questioned their conclusion of additivity. Comparing the fits of an additive and multiplicative model (the fuzzy-logical model of perception, or FLMP, see Massaro, 1987, 1989; Massaro & Cohen, 1993; Massaro & Friedman, 1990). Massaro (1988) found that among the ten individual viewers, the data of five were better fit by the additive model and those of five others by FLMP. Moreover, Dosher, Sperling, and Wurst (1986) found similar multiplicative results in the integration of information from stereopsis and size. Thus, Massaro claimed Bruno and Cutting's (1988) results were indeterminate with respect to additivity. Cutting, Bruno, Brady, and Moore (1992a) then ran additional studies similar to those of Bruno and Cutting (1988). They found that among a total of 44 viewers, the data of 23 were better fit by an additive model and the data of 21 by FLMP. Clearly, it is still indeterminate which model fits the data better, but it is equally clear that the data are almost linear in the middle range of the scale. The near linearity of these data suggest that, in an approach to the study of information combination, the notion of "nearly decomposable systems" (Simon, 1969)—what Marr (1981) and Fodor (1983) later called modules--is not far wrong.

Others have investigated the integration of various sources of information about layout (e.g. Berbaum, Tharp, & Mroczek, 1983; Braunstein & Stern, 1980; Nawrot & Blake, 1991; Rogers & Collett, 1988; Stevens & Brookes, 1988; Terzopoulos, 1986; van der Meer, 1979; Wanger, Ferwerda, & Greenberg, 1992). From a theoretical perspective these, and the previously outlined research, can be discussed in terms of two concepts—*cooperativity* (e.g. Kersten, Bülthoff, Schwartz, & Kurtz, 1992) and *data fusion* (e.g. Luo & Kay, 1989, 1992). Unfortunately, across the larger literature cooperativity is a term which seems to mean little more than the possibility of finding all possible combinations of additivity and interactions as dependent on context. Nonetheless, a particularly useful cooperative approach to integration was suggested by Maloney and Landy (1989; Landy, Maloney, & Young, 1991). They proposed that each source of information may leave one or more parameters about depth unknown, and that various parameters from various sources can disambiguate one another. For example, within their system, they suggest that absolute metric qualities of stereopsis can scale the unachored metrics of relative size, and size can offer depth order to the rotations of kinetic depth information. Robust estimators are also used to reduce the weight of any source of information generally out of line with the others.

Data fusion, on the other hand, is a term from robotics. Luo and Kay (1992), for example, discussed the integration of information in terms of four levels--integrating signals, pixels, features, and symbols. The first two forms of integration generally deal with relatively low-level electronics, but the latter two can be used to distinguish the approaches of Cutting et al (1992) and Massaro (1987; Massaro & Cohen, 1993). The approach of Cutting et al is, as extended in this chapter, one of integrating features, looking for geometrical correspondence among sources of information; the approach of Massaro, on the other hand, is one of integrating symbols and increasing the truth value of what is perceived. Thus, in this context, the computational goal of feature fusion is to build a map; the computational goal of symbol fusion is to measure the surety of a percept and the grounds on which that surety is based.

## Not Rules but Weights

Consider a modeling context. Whereas an understanding of the rules of information integration in visual perception is important, it is surely less interesting than an understanding of the weights. That is, what we and everyone else really want to ask is: Which sources of information are most important? When are they important? and Why? How much does relative size contribute to impressions of layout, say, as compared to stereopsis? How important is accommodation compared to aerial perspective? And so forth. These are the questions about weights for the various sources of information. In this chapter we will provide a framework within which to suggest answers, but in the past, this type of exercise has proven extremely difficult, for at least two reasons.

First, through a process called ecological sampling Brunswik (1956) and Brunswik and Kamiya (1953) tried to assess the "cue validity" of various sources of information—essentially the probability any proximal source is lawfully connected to distal affairs—then imputing weights to these probabilities. However, as principled as this approach is, it is logistically unfeasible. It has never been able to deliver a proper ecological survey of sources of information about depth, and in principle it probably cannot (Hochberg, 1966).

Second, the study of information weights is context and adaptation-level dependent. For example, Wallach and Karsh (1963a, 1963b; Wallach, Moore, & Davidson, 1963), have shown that stereopsis is extremely malleable as a source of information about depth. One day of monocular viewing even by an otherwise stereoscopically normal individual will, immediately after eye patch removal, render temporarily useless the stereoscopic information in judgments about depth. Moreover, through the use of random-dot stereograms (which remove all other depth information) Julesz (1971) reported that stereoblindness and stereoweakness not uncommon in the normal population. For our purposes, however, we will suggest that the variability found in stereopsis is unusual for an information source about layout; most other sources seem not show large individual differences nor adaptational differences (but see Wallach & Frey, 1972a, 1972b).

**Can One Empirically Study the Perception of Layout Given All Its Sources of Information?**

The sheer number of information sources about layout renders implausible any blindly systematic and thorough experimentation. Consider the list of sources given above—accommodation, aerial perspective, binocular disparity, convergence, height in visual field, motion perspective, occlusion, relative size, and relative density—plus others which have been suggested in the literature, such as linear perspective (e.g. Carr, 1935), light and shading (e.g. Gibson, 1948; Graham, 1951), texture gradients (Gibson, 1948, 1950), kinetic depth (e.g. Maloney & Landy, 1989), kinetic occlusion and disocclusion (Kaplan, 1969; Yonas et al, 1987), and gravity (Watson, Banks, von Hofsten, & Royden, 1992). Granting each source be singular--and the texture gradients clearly are not (see Stevens, 1981; Cutting & Millard, 1984)—and granting further this list be complete, there are fifteen different sources to be considered and integrated by the visual system. Given such a lengthy list it is a wonder how the visual system can function. Moreover, it is a wonder as to what researchers should do.

In general we researchers have studied selected combinations of sources almost in effort to avoid thinking about a full-fledged, frontal attack on the general problem of layout and its perception. We have explored the effectiveness of pairs, triples, or even quadruples of the various sources of information in a given context. Examples of this strategy are rife (Bülthoff & Mallot, 1988; Bruno & Cutting, 1988; Dees, 1966; Dosher, et al, 1986; Landy, et al, 1991; Nakayama, Shimojo & Silverman, 1989; Ono, Rogers, Ohmi, Ono, 1985; Terzopoulos, 1986; Uomori & Nishida, 1994; Wanger, et al, 1992); and such research is important. However, given these fifteen sources of information there would be 105 possible pairs of information sources to study, 455 possible triples, 1365 possible quadruples, not to mention higher-order combinations. These are surely more than enough to keep visual scientists busy well past the millennium; but they are also sufficiently plentiful one wonders how overall progress is to be made. Such combinatorics suggest that researchers must set aside global experimentation as being simply unfeasible. As an example, if one uses only two levels (presence or absence) of each source listed above, he or she would need $2^{15}$ (or more than 32,000) different stimuli for a complete orthogonal design; and with three levels per source (necessary for thorough assessment of additivity; Anderson, 1981, 1982), there would be $3^{15}$, or more than 14,000,000 different stimuli. This explosion negates thorough experimentation, and even most theoretically selective experimentation. The major impetus of this chapter, then, is to explore logically the separateness of these sources, and their efficacy at different distances, in an attempt to prune the apparent richness of information about layout at all distances to a more manageable arrangement within particularly domains. This will also provide a set of concrete predictions.

## NINE SOURCES OF INFORMATION ABOUT LAYOUT: MEASUREMENT, ASSUMPTIONS, AND RELATIVE EFFICACY

In this section we will compare and contrast nine sources of information about depth, and then eliminate six others from consideration as either being not independent of those previously analyzed or not demonstrably useful for understanding general layout. Each source of information here will be discussed in three ways. First, each source provides information inherently measured along a particular scale type. Second, each is based on a different set of assumptions about how light structures objects in the world, and these will play a role in our discussion. Third and most important, many sources vary in their effectiveness at different distances, but some do not.

We will use the weakest common scale for each source of information--the ordinal scale--and plot the threshold for judging two objects at different distances, using previous data where available and logical considerations elsewhere. Reduction to a common scale is an example of scale convergence (Birnbaum, 1983), a powerful tool for any discussion of perception, of psychology, and of science in general. We will next compute distance thresholds by analogy to the computation of contrast sensitivity in the spatial-frequency domain. That is, in considering the distances of two objects, D1 and D2, we will plot the ratio of the just-determinable difference in distance between them over their mean distance, 2[D1-D2]/[D1+D2], as a function of their mean distance from the observer, [D1+D2]/2. In this manner our metric compensates for the often-noted decrease in accuracy with distance, such as that noted by Gogel (1993, p.148):

> Distance or depth errors are apt to occur in distance portions of the visual field because cues of depth are attentuated or are below threshold and therefore are unable to support the perception of depth between distant objects at different positions.

We will also assume also that distance judgments are made on the two objects separated by less than, say, about 5° of visual angle measured horizontally. This latter assumption is necessary in order that the environment is sufficiently cluttered with objects and surfaces to let all sources of information operate as optimally as conditions might allow. Having converted ordinal distance judgments to the same axes, we can then compare the efficacy of the various sources of information. Nagata (1991) was the first to present such plots and comparisons, but we will offer much in contrast to what he proposed. As a working assumption, we will regard depth thresholds of 10%, or 0.1 on the ordinate of Figure 1, as the useful limit in contributing to the perception of layout.

---Insert Figure 1 about here---

Before presenting these thresholds, however, we must discuss two important caveats: First, we assume the observer can look around, registering differences in each source of information on the fovea and neighboring regions of the retina. Thus, thresholds will typically be discussed in terms of foveal resolution even though the perception of layout, by definition, must extend well beyond the fovea at any instant. Clearly, then, we also assume visual stability and the integration of information across saccades (see, for example, Bridgeman, Van der Heijden, & Velichkovsky, 1994). Second, we assume that each source of information pertains to a set of objects at an appropriate retinal size. Thus, although an observer will surely not be able to discern from the distance of a kilometer the relative distance of two postage stamps, we claim this is a problem of object resolution, not a problem reflecting source threshold. The relative distance of two buildings at the same distance is in many circumstances quite easily seen. Thus, in the context of our discussion we will always assume that the observer can easily resolve what he or she is looking at.

We will begin our discussion with five more or less traditional pictorial sources of information (historically, often called "secondary cues," e.g. Boring, 1942), then move to the discussion of motion, and then finally to the ocular and physiologically based sources of information (often called "primary cues").

## 1. Occlusion

> Although the principle [of occlusion] is too obvious ordinarily to receive special mention as an artistic technic, it eventually got into the lists of secondary criteria for the perception of distance, as for example in Helmholtz's in 1866.
>
> Boring (1942, p. 264)

***Measurement***. Occlusion occurs when one object hides, or partially hides, another from view. It is ordinal information; it offers information about depth order, but not about the amount of depth. At an occluding edge one knows nothing other than that there is discontinuity in egocentric depth, and that along each surface on either side of the edge the depths are likely to change more or less smoothly and continuously. Thus, a given pair of objects--occluding and occluded--might be at 5 and 10 cm from the observer; at 50 and 100 m; or equally at 5 cm and 1000 m. Initially, ordinal information may not seem impressive, as suggested in the quote by Boring above. Indeed, some researchers have not considered occlusion as information about depth at all (e.g. Maloney & Landy, 1989). But consider two facts; first, an NMDS procedure based only on ordinal information can yield near-metric information in a scaling solution (Shepard, 1980), and second for occlusion to yield ordinal information one need make only four assumptions: the linearity of light rays, the general opacity of objects, the Gestalt quality of good continuation of boundary contours, and luminance contrast.

***Assumptions***. The first assumption—the rectilinearity of rays—is a general principle of light. Its failures are neither common nor generally consequential; they occur only with changes in densities of transmitting media, whether graded (such as those yielding mirages, Fata Morgana, and the like; see, for example, Forel, 1892; Minnaert, 1993) or abrupt (such as those at lens surfaces or at reflective surfaces). This assumption is further based on considering the eye (or any dioptric light-registration device) as a pinhole camera, an assumption known as the *Gaussian approximation* (Pirenne, 1970). But the rectilinearity of rays has been an axiom of optics and visual perception since Euclid (see Burton, 1945); it is an assumption which must be made for all visual sources of information about layout; and thus it need not be considered again.

The second assumption—the general opacity of objects—holds in most circumstances. Light does not usually pass through most of the things which furniture our world. This assumption is pertinent, strong, and useful. Moreover, it can be violated to quite some degree in cases of translucency and transparency (see the discussion of aerial perspective below). That is, even when light does pass through objects it is often changed in either luminance, chromaticity, or both such that depth order can be maintained (Gerbino, Stultiens, & Troost, 1990; Metelli, 1974).

The third assumption, sometimes called Helmholtz's rule (Hochberg, 1971, p. 498), is that the contour of an object in front does not typically change its direction where it intersects the object seen as being behind. The extreme unlikelihood of alignment of the eye, the location of sudden change in a contour of an occluding object, and the point of occlusion is, in modern parlance, a *nonaccidental* property (Witkin & Tenenbaum, 1983).[5]

Finally, for occlusion information to be useful, the luminance contrast at the occluding edge must be above threshold, and probably considerably above. Contrast is importance because the world does not unusually present itself in line-drawing form, the structure implied by any straightforward application of Helmholtz's rule.

***Effective range***. The range over which occlusion works is extremely impressive. Quite simply, occlusion can be trusted at all distances in which visual perception holds. As shown in Figure 1, the effectiveness of occlusion does not attenuate with distance, and indeed the threshold for ordinal depth judgments generally exceeds all other sources of information. We suggest that with conventional natural objects and artifacts occlusion can, throughout the visible range, provide constant depth thresholds about 0.1%. This is the width of a sheet of paper seen at 1.0 m, or the width of a car against a building seen at 2 km. The separate functions for paper, for plywood, for people, for cars, and for houses are shown in the first panel of Figure 2. In a deep sense, the efficacy of occlusion is limited only by the physical properties of objects in the world; If we could make large opaque objects vanishingly thin, occlusion thresholds would increase with that ability. The caveat, of course, is that an occluding edge does not guarantee a difference of 0.1%; the difference in depth might equally be 1000%.

--Insert Figure 2 about here--

**2. & 3. Relative Size and Relative Density**

> There is no object so large ... that at a great distance from the eye it does not appear smaller than a smaller object near. Among objects of equal size, that which is most remote from the eye will look the smallest.
>
> Leonardo (Taylor, 1960, p. 32)

***Measurement and assumptions***. Relative size is the measure of the projected *retinal* size of objects or textures which are physically similar in size but at different distances. Relative density concerns the projected *retinal* density of a cluster of objects or textures, whose placement is stochastically regular, as they recede into the distance. Unlike occlusion, relative size and relative density can, in principle, both yield more than ordinality; they can yield scaled information. Making the similar-physical-size assumption for similarly shape objects, the ratio of their retinal sizes or the square-root of their ratio of their densities will determine the inverse ratio of their distances from the observer. But there are three assumptions.

First, for relative size there must be more than one object and they cannot be too large and too near. For example, if the center of one planar surface of an object is orthogonal to the line of sight and subtends a visual angle of 45°, an identical object half its distance would subtend 79°, not 90°.[6] But when objects are smaller than about 20°, which will include virtually everything seen, this submultiplicativity is minimal. For example, a rectilinear object which subtends 10° will, at half the distance, subtend 19.85°, a near-perfect doubling of retinal size.

Second, without assuming knowledge of the particular objects themselves (which transforms this source of information into "familiar size" or "assumed size"; see, for example, Epstein, 1963, 1965), differences in retinal size of 3:4:5 and differences in retinal density of 9:16:25 occur equally for three objects (or sets of objects or textures) at 1, 1.33, and 1.67 cm; and at 50, 66.5, and 83.3 m. Thus, differences are metrically scaled, but without absolute anchor. If the objects are known to the observer, then relative size becomes familiar size, and absolute information is available. Aside from knowledge the only other difference between the two is that with familiar size there need be only one object present; with relative size, more than one must be visible, and with relative density there must be many. In general, we will not assume knowledge of objects here.

Third, the assumption of "similarity" in physical size must not be taken too strictly. For example, mature trees of the same kind, age, and shape will be roughly, but not exactly, the same size (e.g. Bingham, 1993). A 10% range in variability in size in strategically placed objects could cause a similar variability (error) in the assessment in distance. This can reduce the power of the scaling factor, but even in its reduction the information scale has stronger assumptions than that of mere ordinality; it might be called *fuzzily scaled* information.[7]

***Ordinal conversion and effective ranges***. Although relative size can offer scaled information, for our purposes we must reduce it to ordinal judgments. Teichner, Kobrick, and Wehrkamp (1955) measured the just-noticeable-distance thresholds for large objects mounted on jeeps stationed in a desert and similar terrains.[8] These data are replotted and shown in Figure 1 with the same value as given by Nagata (1991). Notice that, like occlusion, relative size can generally be trusted throughout the visible range of distances, from say 0.5 m to 5000 m (the horizon for an average adult standing on an earthen surface approximating a flat plane or at the edge of a large calm lake), and beyond. However, relative size provides a depth threshold of about 3%, a bit more than an order of magnitude worse than occlusion. Moreover, if natural objects are allowed to vary in size by 10%, the effectiveness of size is further diminished, as shown in the top panel of Figure 2.

On logical grounds alone Nagata (1991) suggested that one's sensitivity to density ought to be the square-root-of-two times greater than relative size. However, the data of Cutting and Millard (1984) have shown that density is psychologically less than half as effective as the size gradient in revealing exocentric depth (see also Stevens, 1981). Thus we have plotted it as weaker than relative size in Figure 1, at just about our 10% threshold, equal to our margin for considering a source of information effective for determining layout. This means that a viewer ought to be able to just discriminate the difference between two patches of random elements, one containing 95 and the other 105 dots. For further discussions of density, see Allik, Helsper, and Vos (1991), Barlow (1978), and Durgin (1994). Since we have already assumed stochastic regularity in textures, any effects of variability in element spacing would seem likely to be related to the parameters modeling the stochastic regularity. However, since we know of no data relevant to this concern, no panel in Figure 2 is devoted to factors of variation which would generate a function different than that shown in Figure 1.

## 4. Height in Visual Field

> In the case of flat surfaces lying below the level of the eye, the more remote parts appear higher.
>
> Euclid (Burton, 1945, p. 359)

***Measurement and assumptions.*** This fourth source of information is in the projected relations of the bases of objects in a three-dimensional environment to the viewer. Such information, moving from the bottom of the visual field (or image) to the top, yields good ordinal information about distance from the point of observation (see, for example, Dunn, Gray, & Thompson, 1965). Height in the visual field also has the potential of yielding absolute distances. The assumptions which must be made for such information, however, may in many situations be unacceptably strong. They are four and in decreasing acceptability, if not plausibility, they are: (a) opacity of the ground plane; (c) gravity, and that each object has its base on the surface of support (it is not suspended, or floating); (c) that the observer's eye is at a known distance (say, about 1.6 m for any individual about 5'9" in height) above the surface of support; and (d) that the surface of support is generally planar and orthogonal to gravity. If all assumptions are valid then 10° of visual angle below the horizon, the width of a fist held vertically at arm's length, is 9 m away assuming an eye height of 1.6 m for an observer standing on level ground; 2° below the horizon, the width of the thumb at arm's length, is about 50 m away.[9] However, the plethora of assumptions needed to validate height in the visual field may occasionally impugn its effectiveness, as suggested in our later section on source conflict. We feel that assumptions (a) and (b) are the only ones generally valid.

***Effective range***. Unlike occlusion, size, and density the effectiveness of height in the visual field attenuates with distance. Since the bases of seen objects must be touching the ground plane for it to be useful, since we are considering an upright and moving observer, and since the observer's eye is already at a height of 1.6 m, no base closer to the eye than 1.6 meters will generally be available.

We will assume that height-in-visual-field differences of 5 min of arc between two nearly adjacent objects is just detectable. This is an order of magnitude above standard resolution acuity (which is about 0.5 min) but its allows for quite a few degrees of separation in the visual field for the two objects under consideration. Moreover, a different value would simply shift the function up or down a bit. Under these assumptions and considerations, as shown in Figure 1, the utility of height in the visual field is truncated short of about 2 m; at 2 m it is nearly as effective as occlusion (provided one is looking at one's feet); and beyond 2 m its effectiveness diminishes curvilinearly. At about 1000 m its threshold value has fallen to a value of 10%, our benchmark for assumed utility. However, when assumption (d) above is violated, ordinal information can even improve beyond what is shown in Figure 1 for objects whose bases are still visible. This due to the rotation of the particular region of the ground plane towards the eye of the observer; and sample functions are shown in the third panel of Figure 2.

### 5. Aerial perspective

> There is another kind of perspective which I call aerial perspective because by the atmosphere we are able to distinguish the variations of distance ... [I]n an atmosphere of equal density the remotest objects ... appear blue and almost of the same hue as the atmosphere itself ...
>
> Leonardo (Taylor, 1960, p. 37,38)

Aerial perspective is determined by the relative amount of moisture and / or pollutants in the atmosphere through which one looks at a scene. When air contains a high degree of either, objects in the distance become bluer and / or decreased in contrast with respect to objects in the foreground. Such effects have been known in art since before Leonardo and are called effects of *participating media* in computer graphics.

***Measurement and assumptions***. In principle aerial perspective ought to allow interval information about depth. Assuming, as Leonardo did, that the participating medium is uniform, then the increasing blueness or lightness of objects in the distance will be linear with their increasing distance from the observers. In practice, however, it would seem that aerial perspective allows only ordinal comparisons; we know of no data on the topic, although contemporary computer-graphics technology would easily allow preliminary study. Nonetheless, real-world experimental control of participating media would be extremely difficult to obtain.

***Effective range***. As shown in Figure 1 and unlike all other sources of information, the effectiveness of aerial perspective increases with the logarithm of distance. This is due to the fact that, assuming the medium is uniform across linear distance, logarithmic increases in distance will encompass logarithmically increasing amounts of the medium. The limitation, however, is that with great distances objects become indistinct, and the source of information becomes rapidly ineffective. The function shown in Figure 1 is Nagata's (1991), but as shown in fourth panel of Figure 2 the effective range can vary greatly, depending on whether there is fog, haze, or clear air.

The concept of aerial perspective is also easily modified to consider underwater environments, condensing the range still further; and is also justifiably applied to the perception of transparency (Gerbino, et al, 1990; Metelli, 1974). That is, the study of the depth order of two or more colored but transparent sheets is no different than aerial perspective, except that the sheets become the medium and that rather than being a graded continuous function, the transparency generates a discrete-valued step function. Transparency is the only source of information about layout that does not assume the opacity of objects.

With these five our assessment of the pictorial sources of information is complete. Other sources of information might be added to this list, but we will defer discussion of them until our full list is complete. We now move to the discussion of motion, then to accommodation, convergence, and binocular disparities.

### 6. Motion Perspective

> In walking along, the objects that are at rest by the wayside stay behind us; that is, they appear to glide past us in our field of view in the opposite direction to that in which we are advancing. More distant objects do the same, only more slowly, while very remote bodies like the stars maintain their permanent positions in the field of view ...
>
> Helmholtz (1866, p. 295)

Motion is omnipresent for a mobile observer; it is even claimed to be the foundation of human vision (Lee, 1980). To be sure, when stabilized images are projected onto the retina, removing all motion, the world disappears; and without motion it is very difficult to get infants to respond to any aspects of layout (Yonas & Granrud, 1985). Nonetheless, we think the role of motion in our perception of layout may have been overplayed; indeed, in landscape paintings there is no motion and yet these have pleased people in many cultures for many centuries. Moreover, empirically Schwartz and Sperling (1983) have shown that relative size can dominate motion information in judgments of depth, and Vishton, Nijhawan, and Cutting (1994) have shown that size and other sources of information can dramatically influence observers judgments of their heading in sequences simulating observer motion through cluttered environments. In this light, then, we would expect motion information to be relatively important for the perception of layout, but not all encompassing.

The relative movement of the projections several objects stationary caused by observer movement is called motion parallax; the motions of a whole field of such objects is called motion perspective (Gibson, 1950, 1966). Ferris (1972) and Johansson (1973) demonstrated that, through motion perspective, individuals are quite good at judging distances up to about 5 m (but see Gogel & Tietz, 1973; Gogel, 1993). Introspection suggests our accuracy would be high at considerably greater distances as well.

***Measurement and assumptions***. Motion perspective assumes little other than that one is moving through a rigid environment. Some depictions of that environment also assume planarity of the surface of support (Gibson, 1966, 1979; Gibson, Olum, & Rosenblatt, 1955; see also Bardy, Baumberger, Flückiger, & Laurent, 1992; Flückiger & Baumberger, 1988), but this second assumption is typically made only for purposes of pedagogical or methodological clarity. In principle, the complete array of motions specifies to a scaling factor the layout of the environment and the instantaneous position of the moving observer (e.g. Gibson, 1966, 1979; Koenderink, 1986; Lee, 1980; Prazdny, 1983). That is, with no other information available, one can mathematically discern from the flow both the relative position of objects and one's own relative velocity (in eye heights per second).

What is relative in this context concerns clutter and reveals two interrelated sources of motion information—edge rate and global flow rate (see, for example, Larish & Flach, 1990). Assume that one is moving at a constant height and velocity over a flat ground plane with textures at regular separation, and that one is in a vehicle with a windscreen and maintaining constant orientation with respect to the plane. Edge rate is the number of countable objects, or edges, per unit time that pass any point on the windscreen as one moves forward, and is related to relative density as discussed above. Given a uniformly textured ground plane edge rate will be constant everywhere there is texture.[10] Global flow rate is the pattern of relative motions everywhere around the moving observer; rapid beneath one's feet, decreasing with distance, nearly zero at the horizon all around, and generally following a sine function as one moves from 0° (directly ahead) to 180° (directly behind) and then back to 360°. This hemispherical pattern is multiplied as a function of velocity, or as the reciprocal of altitude.

More concretely, edge rate is dependent on velocity and texture density, but not on altitude; flow rate is dependent on velocity and altitude, but not texture density. Larish and Flach (1990) demonstrated that impressions of velocity depend much more on edge rate than flow rate, accounting for the efficacy in Denton's (1980) study of car drivers slowing down at tollbooths and roundabouts on motor ways through the use of decreasingly spaced markers. It may be that impressions of layout and depth are similarly governed more by edge rates than flow rates, but the relation between the functions for density and motion perspective in Figure 1 suggest it may not. In any event, since regular or even stochastically regular textures are often not available to a moving observer, we will concentrate on flow rates at 2 m/s, about the speed of a pedestrian.

***Ordinal conversion and effective range***. Again, and with others who have assessed the efficacy of motion perspective, such as Nagata (1991), we will assume that the moving observer is a pedestrian and free to move his or her eyes about (although all considerations are monocular). Thus, the thresholds plotted in

Figure 1 are foveal for a roving eye (one undergoing pursuit fixation during locomotion as well as saccadic movement, see Cutting et al, 1992b), where motion detection is best, and our situation then equally assumes an individual has sampled many locations in the visual field over a period of time.

Unlike the monotonic function shown by Nagata (1991), we have indicated that motion perspective acuity declines below about 2 m due to the difficulty in tracking and even seeing differences in rapid movement. Graham, Baker, Hecht, and Lloyd (1948), Zegers (1948), and Nagata (1991) measured the difference thresholds for motion detection. If the experimental values are used as input measures for the pedestrian motion perspective at relatively short distances (<10 m), and if the absolute motion threshold (1 min of arc/sec) is used for distances beyond (see also Graham, 1951),[11] the function shown in Figure 1 is generated. However, that function also assumes that one is looking at 90° to the direction of movement, a somewhat unusual situation for a pedestrian. Thus, a family of threshold functions is shown in the fifth panel of Figure 2 for an observer looking at 1, 2, and 10° from his or her direction of movement. Since observers do not usually change direction to assess layout, the function in Figure 1 probably overestimates the importance and power of motion perspective for a pedestrian; but, of course, when riding in a car or fast train (TGV in France), the function would amplify and slide to the right, as shown in the sixth panel of Figure 2.

### 7. & 8. Convergence and Accommodation

> We alter the disposition of our eyes, by lessening or widening the distance between the *pupils*. This disposition or turn of the eyes is attended with a sensation, which seems to me to be that which in this case, brings the *idea* of greater or lesser distance into the mind.
>
> Berkeley (1709, sections 16-20)

Convergence is measured by the angle between the optical axes of the two eyes. When this angle is large, the two eyes are canted inward as if to focus on a spot very near the nose; when this angle approaches 0° the two eyes are parallel aligned as if to focus on the horizon or beyond. Accommodation is the change in the shape of the lens of the eye, allowing it to focus on objects near or far while still keeping the retinal image sharp. The effective degree to which accommodation can change in a given eye is defined by the distance between the near point (the point measured along the optical axis—from the fovea, through the middle of the pupil, outward—to the nearest orthogonal plane at which something can be placed and remain in focus due to the thickening of the lens) to the far point (that point along the same axis farthest away at which an object can remain in focus due to the thinning of the lens). These points vary considerably across individuals and, with age, within individuals. At any instant, the effectiveness of accommodation may not be very great, since blur occurs equally for objects nearer and farther than the mean focal depth, providing measures which fail ordinality. However, across fixations at different depths it has the potential of offering such information.

Convergence and accommodation are reflexively yoked (Fincham & Walton, 1957; Kersten & Legge, 1983; Morgan, 1868) and thus we have considered them together.[12] Both can be measured in diopters, or the reciprocal of distance in meters; thus the changes in distance from 0.33 to 0.5 m, from 0.5 to 1.0 m, and from 1 m to infinite distance are each 1 diopter.

***Measurement and assumptions.*** The assumptions of these two sources of information are few: For convergence to be effective one must have an object to fixate, knowledge of (or familiarity with) the distance between one's eyes—for an adult usually about 6.4 cm. In addition, for both convergence and accommodation to be effective the target must have reasonable luminance contrast and a relatively complex spatial frequency distribution (Fisher & Ciuffreda, 1988). Thus, in principle, these sources could be extremely effective in measuring distance, yielding metric information within the near distance.

***Ordinal conversion and effective range***. As suggested by the metric of diopters, neither convergence nor accommodation are effective at great distance. Indeed, evidence suggests that the effectiveness of convergence alone as a source of information for depth is confined to a range up to only about 2 m or so (see Gogel, 1961; von Hofsten, 1976; Lie, 1965; Richards & Miller, 1969), and accommodation a bit less so (Fisher & Ciuffreda, 1988), although both can interact with other sources (Wallach & Norris, 1963). Even when combined their useful range still seems to be less than 3 m (Leibowitz, Shina, & Hennessy, 1973). The function shown in Figure 1 stems from the experiments and logical considerations of Nagata (1991), and shows the utility of both sources declining linearly with distance. The functions shown in seventh panel of Figure 2 remind us that since the ability to accommodate generally declines with age, its efficacy as a source of information about layout must also change.

## 9. Binocular Disparity, Stereopsis, and Diplopia

> ... the mind perceives an object of three dimensions by means of two dissimilar pictures projected by it on the two retinae ...
>
> Wheatstone (1838, p. 373)

Binocular disparities are the differences in relative position of the projections of the same object on the retinas of the two eyes. When looking at an object at a given distance there are other locations in the environment which, if objects are located there, fall on a zero-disparity surface called the horopter, generally circular in depth according to Thales theorem (see Arditi, 1986) and linear along declining planes when measured vertically (Nakayama, 1977). When these disparities are sufficiently small they yield stereopsis—or the impression of solid space. When greater they yield diplopia—or double vision—also a good source of information about relative depth (Ogle, 1952; Duwaer & van den Brink, 1981). No other source of information about layout has been more studied than binocular disparity and we cannot do justice here to its vast literature. For a lucid discussion see Kaufman (1974) and for comprehensive overviews, see Arditi (1986) and Gulick and Lawson (1976).

***Measurement and assumptions***. The measurement scale implied by retinal disparities and stereopsis has often been taken to be the strongest among all sources of information about layout. Indeed, many have claimed that disparities offer an absolute metric for distance (Dees, 1966; Maloney & Landy, 1989). The assumptions underlying stereopsis are few and seldom violated—one must have, first, some knowledge of or familiarity with the distance between the eyes,[13] the state of vergence (which beyond 2 m is always less than 1°), and then one must find the locations of appropriately corresponding points with generally lateral retinal displacements in the two eyes. These corresponding points need not be identified with parts of objects, indeed random dots will do (Julesz, 1971), and the relative luminance and relative size of those dots is not vitally important. The paucity of assumptions associated with the utility binocular disparities may reinforce their effectiveness in perception.

***Ordinal conversion and effective range***. Ordinal thresholds for depth are can be computed from existing discussions and data (Arditi, 1986; Berry, 1948; Bülthoff, Fahle, & Wegmann, 1991; Ebenholtz & Walchli, 1965; Foley, 1991; Nagata, 1991; Ogle, 1952, 1958), and are shown in Figure 1. Critically, these assume that vergence can vary across fixations at different distances, and that both stereopsis and diplopia are useful for depth judgments. Clearly, disparities are most effective in the near field (regardless of fixation depth), and their relative effectiveness attenuates linearly with distance. Notice also the similarity in threshold acuity for binocular disparity and motion perspective when both are at their best for pedestrians, a result which fits snugly with the comparative results of Dees (1966), who assessed the accuracy of perceived depth using these two source of information independently and in combination. Since stereoweakness and stereoblindness are not uncommon, the functions in the eighth panel of Figure 2 are designed to capture this fact.

**Sources Not Considered**

In our list of nine sources of information about layout, we have left out several commonly accepted candidates. Their omission has been purposeful, but needs explanation; so let's consider each in turn.

***Texture gradients.*** Gibson (1950) was first to note the importance of texture gradients for the perception of layout, and we agree they are powerful sources of information (e.g. Cutting & Millard, 1984). However, there are at least three gradients of texture in surfaces receding in depth—the gradients of size, density, and compression. The size gradient (the change in the largest extent of a texture element, measured parallel to the axis of tilt) is effectively no different than relative size working continuously on many elements across a surface.[14] It has thus already been considered and plotted in Figure 1. Likewise, the effectiveness of relative density has also been given in Figure 1. The compression gradient is isomorphic with the slant gradient considered by Nagata (1991), and his analysis, based on the results of Freeman (1966) suggested that it is quite powerful (indeed about three times as effective as relative size) and unchanging with distance. However, we find Nagata's analysis unconvincing; the data of Cutting and Millard (1984) suggest that compression is extremely ineffective in its psychological effect of revealing depth—much less so than even density—and it is thus not included in our list. Instead, compression (or slant) is good information about object shape, particularly its curvature near the object's self-occluding contour (Marr, 1981).

***Linear perspective***. It will surely strike some as surprising that we have not placed linear perspective in our list. Our reason is not because its effects are not powerful; they are (e.g. Kubovy, 1986). We omit discussion here because linear perspective is a systematic combination (even an asystematic combination, see Elkins, 1992) of several other sources of information and a choice of display elements. For example, it combines all three texture gradients—size, density, and compression—with occlusion and then is separated from natural perspective (see Leonardo, in Taylor, 1960) by the copious use of parallel lines. Receding parallel lines are an extremely effective way of reducing noise in the measurement of object sizes, densities, and compressions; but all the relevant elements of linear perspective have already been considered.

***Brightness, light, and shading***. Some conflation of these terms has appeared on many lists of sources of information about layout (e.g. Boring, 1942; Nagata, 1991), but we find their inclusion anomalous (as did Gibson, 1950). To be sure, it has been demonstrated that gradually increasing the luminance of an object causes it to appear to move forward. However, this appears to be something of a "theatrical" effect, one which assumes either that the foreground is better lit than the background or that the light source of a scene is necessarily near and at roughly the same distance from the object as one's eye. Since the position of the sun negates both of these assumptions we think something else is at work. In fact, we believe that luminance in such cases may acting as a surrogate for relative size, and indeed has been used experimentally in just this manner (Dosher et al, 1986). More ecologically, the play of light on a form yields shadows, which have also been construed as a source of information about depth (e.g. Bülthoff & Mallot, 1988; Yonas & Granrud, 1985). However, with Cavanagh & Leclerc (1990) we think is it better to consider shadows as information about object shape, not depth per se. Moreover, shadows can be considered an application of the phenomenon of transparency, which in turn was considered a variety of aerial perspective.

***Kinetic depth***. Quite a number of researchers have discussed kinetic depth (Wallach & O'Connell, 1953)—structural information revealed through motion (e.g. Ullman, 1979, 1983) as an important source of information about depth (e.g. Maloney & Landy, 1989). However, we suggest that this information, like shading, concerns object shape and contributes little to the study of layout. Moreover, in rotating wireframe forms it often fails the criterion of ordinality—figures reverse in depth, even when perspective information is added (Braunstein, 1962, 1976)—and a rotating, generally solid object in silhouette often does not even allow perception of rigidity (Wallach & O'Connell, 1953).

***Kinetic occlusion and disocclusion.*** Kaplan (1969) and Yonas et al (1987), among others, have discussed the property of kinetic occlusion—sometimes called accretion and deletion of texture—as a source of information about depth. However, we regard kinetic occlusion as covered in our list above in two ways. First, kinetic occlusion is simply occlusion as revealed through motion; it differs only in that it need assume no luminance contrast at the occluding edge. Under conditions of low visibility, such as approaching a descending set of textured steps at night, kinetic occlusion may be the only source of information available to the pedestrian. Nonetheless, aside from small differences in the assumptions on which it is based, it does not have separate qualities or properties from occlusion. Second, kinetic occlusion can also be considered a two-valued function of motion perspective, yielding ordinal information, but again already considered in our list.

***Gravity***. It is interesting that researchers continue to discover more sources of information about layout and depth; and we expect this trend to continue at a slow pace. Recently, for example, Watson, et al (1992) postulated that gravity can serve as a source of information. That is, since the acceleration patterns of dropped or thrown objects on the retina will vary with distance, these patterns are available as information. Clever as this idea is, gravity is not likely to be very useful in determining layout because one seldom sees objects thrown or in free fall when viewing one's surround. In addition, "gravity" as such appears to be a source of information about an isolated object, not about layout.

### Overview

The functions in Figure 1 are ordinal depth-threshold measures for various sources of information about layout as a function of mean distance from the observer. If we assume that the relations among these threshold functions can generalize to suprathreshold situations, then we may have a powerful tool. But these functions are somewhat idealized and may not apply to all situations; those shown in Figure 2 show the variability that can be obtained according to different environments and states of the observer. Nonetheless, we assume that in every environment for every observer one can, in principle, plot a similar array of functions, and that that array can be used to understand how we perceive layout as we do.

## A SITUATION-SENSITIVE DOMINANCE HIERARCHY OF INFORMATION SOURCES

> Objects and events in a natural environment can be multiply specified by many different sources of information, each of which is detected by a specialized processing module with its own individual limitations. In any given situation, we would expect to obtain erroneous outputs from some of these modules because of inappropriate viewing conditions, but it would be most unlikely for two of more of them to fail in exactly the same way.
>
> Todd (1985, p. 708)

Why are there so many sources of information about layout? Part of the answer must be that environments are typically rich, but they can also be extremely varied, with certain kinds of information present in some situations but not in others. Another part of the answer must be that perceiving layout is extremely important to us, as it must have been to our forebears and is to other organisms, so that systems with multiple mechanisms for the gathering of information about layout have an advantage over those with single mechanisms. Now, the analysis of the relationships in Figure 1 suggests many things that allow us to go beyond such vague notions, and beyond Todd's instructive beginning.

Why might a source of information fail? Figure 1 suggests that with varying distance some sources fail to deliver an adequate quality of information. The panels of Figure 2 suggest that with varying environments (and with varying states of the organism) this adequacy will also vary. The differential adequacies as a function of distance may account for some traditional findings in the field.

### Conflicting Information

In Figure 1 we find rationale for certain well-known results in experimental situations of what is often called "cue-conflict"—when two or more sources of information are presented yielding discrepant ordinal information about depth. If we assume that our threshold analyses were properly carried out and that they can be used to generalize to suprathreshold considerations, then certain predictions can be made.

Consider four examples. First, occlusion would appear to dominate all other sources of information for layout (see Bruno & Cutting, 1988, Figure 7; Krech & Krutchfield, 1958) and in all contexts. In fact in situations of conflict, these functions suggest only binocular disparity (Nakayama et al, 1989) and perhaps height in the visual field would seem to come close to the power of occlusion. Second, we also find evidence as to why accommodation and convergence are generally so weak, and in situations of conflict will generally dominate no other information. Their role in such situations seems simply to modify perceived layout through the use of other sources (Wallach & Frey, 1972a, 1972b). Third, our threshold functions also suggest that height in the visual field would dominate relative size information out to about 500 m, but that size would dominate beyond. We know of no data on this matter.

Fourth, consider a potential problem for our hierarchy. The data shown in Figure 1 would seem to predict that height in the visual array would always dominate binocular disparity, at least beyond about 2 m. However, this seems implausible; and within 1 m Gilchrist (1977, 1980) has provided empirical proof. One possible rationale for this reversal in our system is through consideration, not simply of the functions shown, but also the validity of the assumptions underlying each source. The assumptions of height in the visual field entail objects resting on a flat ground plane as seen from one's usual eye height—and all seem almost "cognitive." In contrast, the assumptions for binocular disparity are physiologically given. It seems likely, then, that certain classes of disparities on certain objects could easily invalidate an assumption made for height in the visual field; thus, disparity differences at the base of an object compared to those of the ground immediately in front of the object would render invalid the assumption of the object resting on the support surface—and binocular disparity would then dominate height in the visual field. Such an analysis offers some empirical force to the otherwise weak notion of "cooperativity" among sources of information about layout.

**Three Categories of Source-Distance Relations**

In addition, the relations among the various functions in Figure 1 suggest there are three categories of information sources for layout—those that do not vary with the logarithm of distance, those that dissipate with distance, and one that increases with the logarithm distance.

***Sources Invariant with Distance***. Three sources of information are unaffected by logarithm of distance from the observer—occlusion, relative size, and relative density—and all are pictorial sources of information. Such effective consistency across distances suggests that in experimental situations information conflict occlusion will always dominate size, which in turn will always dominate density. Moreover, in natural situations these three sources can be used as anchors across different depth ranges to coordinate perception of layout from arm's length to the horizon along a logarithmic scale. Indeed, elaborating Birnbaum's (1983) terms, the similarity of function shape of these sources within the same measurement scale should well reinforce one another.

***Sources Dissipating with Distance***. Members of the second category generally decrease in effectiveness with distance, and would do so whether measured logarithmically or linearly. These are height in the visual field and motion perspective (both assuming objects are not too close), and the physiologically coordinated cluster of ocular sources—binocular disparity, convergence, and accommodation. Throughout the relevant range of distances these five sources maintain their relative ranking and, given the same general function shape and scale, we would expect them to act in consort reinforcing each other's measurements.

***A Source Increasing with Distance.*** Finally, the only source of information that generally increases its effectiveness with the logarithm of distance is aerial perspective. In typical situations, however, this source is likely to become effective only with great distance. It may even help compensate for ("cooperate" with) the declining effectiveness of the information sources in the previous group by enhancing occlusion information.

Given three classes of ordinal depth threshold functions—those constant with the logarithm of distance, those decreasing, and one increasing—we speculate that one can go beyond Birnbaum (1983) and, given the shared ordinal scale, postulate the utility of function-shape contrast. That is, in addition to the possibility of members of each category reinforcing one another, the existence in optimal situations of different function shapes can serve to diversify and solidify judgments of distance. We claim these functions serve to separate at least three different types of distance.

**Segmenting the Space Around Us**

As an attempt to satisfy the primary goal of this chapter, the relations in Figure 1 suggest a framework for a theory of contextual use of the various sources of information about layout. From the pattern of threshold results and calculations in Figure 1, we suggest that the layout around a moving perceiver can be divided into three circular, egocentric regions which grade into one another. For the purposes of this discussion we will eliminate relative density from further consideration since it appears to be on the margin of utility throughout the visible range.

***Personal space***. The zone immediately surrounding the observer's head, generally within arm's reach and slightly beyond, is quite personal. Typically, others are allowed to enter it only in situations of some intimacy, or in situations of public necessity, such as when one rides in a crowded subway car (e.g. Hall, 1966; Sommer, 1969). The perception of people and of objects in this space is served by a number of sources of information, but this number is smaller and experimentally more tractable than the complete array we have been discussing. Somewhat arbitrarily we will delimit this space to be within 2 m, and typically this is the space worked within by a stationary individual. Thus, motion perspective is not typically generated by the observer through his or her own motion, but instead motion parallax (and perhaps kinetic depth information) is generated by observer manipulation.

Within this region, and according to our framework as we have laid it out, six sources of information are generally effective. Five are as they appear in Figure 1—occlusion, retinal disparity, relative size, convergence, and accommodation. We suggest that, when each is available, these five roughly dominate each other in that order. Motion information is also used, but it is not typically the information from motion perspective since this is a space typically perused by a stationary observer, not a pedestrian. Indeed, the useful motion within this domain will most typically be the motion generated by observer head motion, or observer manipulation of objects. We also suggest that in most circumstances this collection of these six sources of information are selected from the background of the other three (or more) sources (which have generally null value), integrated, and combined—each constraining the other—to produce the ability of human observers to accurately manipulate objects around them. Notice that four of these sources decline in effectiveness with distance in near space, with the declining effectiveness of accommodation and convergence (and the intrusion of height in plane) helping to delimit it from more distance spaces; two other functions (those for occlusion and size) are constant.

***Action space***. In the circular region just beyond personal space is a space of an individual's public action. Relatively speaking, we move quickly within this space, we can talk (even teach) within it without too much difficulty, and if needs be we could toss something to a compatriot within it or throw a projectile at an object or animal. This space is also served by a different collection and ranking of sources of information. There are five: Occlusion, height in the visual field, binocular disparity, motion perspective (here for a pedestrian), and relative size. These are selected from the background of the other four sources, which have essentially no substantive value. The first and the last of these sources are constant with the logarithm of distance and the other three decline. Because the utility of disparity and motion perspective decline to our effective threshold value of 10% at about 30 m, we suggest this effectively delimits space at 30 m.

***Vista space.*** Beyond about 30 m very little changes for the binocular pedestrian except over a period of several seconds. For example, the motion of an object is considerably less salient than its displacement, and the benefits to perceiving layout from two eyes are also greatly diminished. The only effective sources of information are four sources that have been traditionally called the "pictorial cues"—occlusion, height in the visual field, relative size, and aerial perspective. We suggest these are selected from the background of other five, uninformative sources. Since the effectiveness of binocular disparity and motion perspective have diminished substantially in this space, we regard their inputs as relatively modest; what lies beyond about 30 m for a pedestrian, we claim, is the layout of a vista, generally unperturbed by the motions of the observer. But let us be clear, this is not to say we do not see objects in depth beyond 30 m; clearly we do. What we are claiming is that only the monocular and static sources of information are typically available in any quality to the pedestrian. Vista space is also the region in which very large paintings are most effective in either revealing layout, or deceiving the eye so that extended layout is seen—such as in the Pozzo ceiling in the Church of St. Ignazio in Rome (Pirenne, 1970).

***Source potency ranked within spaces***. With egocentric space segmented in this manner, we can now rank order the relative importance of the various sources of information within the three spaces (again, given the assumption that we can generalize from threshold to suprathreshold considerations). Such a set of rankings is shown in Table 1, done by integrating the area under each depth threshold function within each spatial region, then comparing relative areas. Notice several things. First, we ranked convergence and accommodation as tied throughout; and second, within personal space we ranked them as superior to relative density, since although all three might appear to be tied, we have chosen to emphasize the fact that the logarithmic scale of distance is truncated at the near-end of the scale. Third, we ranked height in the visual field above relative density in vista space to emphasize the near field (it is not everywhere that one can see farther than, say, 200 m into the distance), but we did not include it at all in personal space (unless one is crawling, as an infant might, on the floor). Finally, we ranked aerial perspective as tied with density in vista space since under many atmospheric conditions the function is slid leftward and subsumes more area in the nearer part of vista space than does density.

We assume these rankings by area beneath each curve reflect the general importance of the sources of information within each segment of space. With these rankings in place, we wish to apply our scheme in three ways: looking at the relational development of their use representative art, looking at the ontogeny of the use of the various sources of information, and looking at scientific research.

--Insert Table 1 about here--

## THREE APPLICATIONS FOR THE SYSTEM OF PERSONAL, ACTION, AND VISTA SPACES

### Layout, Pictorial Art, and a Hierarchy of Information Sources

Let us consider the general development of pictorial art as it has represented space. Before beginning, however, several caveats are in order. First, this is not the place to give a full treatment of the psychological study of space in pictorial art; only an extremely brief sketch can be given. Second, we will consider pictorial art only through the Renaissance and certainly not through the development of photography. This constraint eliminates consideration of the last half of the 19th and all of the 20th century, removing motion perspective from consideration, as well as the three ocular and binocular sources (accommodation, convergence, and disparity), and leaves us with only with the five pictorial sources of information—occlusion, relative size, relative density, height in the visual field, and aerial perspective. Third, although some interesting things have been said by psychologists about the development of art—sometimes briefly (e.g. Gregory, 1966), sometimes in detail (e.g. Blatt, 1984; Hagen, 1986; Kubovy, 1986)—there has also been considerable mischief (e.g. Gablik, 1976), which we do not wish to replicate. In our context, perhaps the worst sin would be to suggest an invariant order in the application of pictorial sources of information across cultures and time.

***Method.*** We sampled widely from images in many traditions of pictorial art, we focused on single pictures, and within those images we noted the sources of information about layout from our list that are used. In this manner, we are definitely *not* attempting an analysis of the *history* of art. Instead, we are deliberately ahistorical, focusing sequentially on single artists (or schools) at a particular time, who composed a particular piece of art for a particular purpose. If within this image a single source of information is used alone to depict depth, then it is at the top of the hierarchy; if a second source is used only in the context of the first and without any others, then it is ranked second; and so forth.

***Results and discussion.*** If only one source of information about three-dimensional layout is used by an artist, that source is almost certainly occlusion. For example, occlusion is the only such information seen in paleolithic art from about 12,000 years ago—it is used in the Hall of the Bulls cave paintings in Lascaux, those in Niaux, and those in Altamira (see, the images in Biederman, 1948; and Hobbs, 1991). Occlusion is also typically the only depth information used in Egyptian art of 2500 B.C (see, for example, the images in Hagen, 1986; and Hobbs, 1991). We take this finding as evidence that occlusion is at the head of the hierarchy for information about layout.

If a second source of information is used to portray three-dimensional layout, that information is typically height in the visual field. The conjunction of occlusion and height, with no other sources, can be seen in High Classical Greek Art (4th century B.C.); in Roman wall paintings (1st century B.C.); in traditional Chinese landscapes (10-13th centuries); in Japanese art (12-15th centuries); in the works of Cimabue (13th), Duccio di Buoninsegna and Simone Martini (14th), and Giovanni di Paolo (15th century) in Western art; and in Persian drawings (15th century) (see, for example, the images in Blatt, 1984; Cole, 1992; Hagen, 1986; Hobbs, 1991; and Wright, 1983). This widespread array of sources suggests that height in the visual field is second in the hierarchy.[15]

If a third source of information is used, that information is typically relative size. Before the systematization of linear perspective, many of the images of Giotto, Gaddi, and Lorenzetti (14th century), for example, use relative size information coupled with occlusion and height. Some traditional Chinese and Japanese art also adds size variation to occlusion and height (see, for example, the images in Cole, 1992; and Hagen, 1986). The layout in these images suggests that relative size is the third source in the hierarchy.

After these three sources, however, the ordering in our hierarchy is less clear. In the Western tradition, it would appear that relative density is next. Coupled with occlusion, height, and size, the use of density first appeared with the use of local perspective, particularly in the tilings of floors such as those of Lorenzetti; and was complete with the global perspective of Alberti, Donatello, Massachio, and Uccello (15th century) (see, for example, the images in Cole, 1992). Only later did Leonardo (15-16th centuries) systematically employ the blueness of aerial perspective. This ordering makes sense within our system. Linear perspective took occlusion as granted and made a coherent system out of height and size, creating a tool to represent density information in a system of parallel lines. This feat makes the registration of information about layout in action space quite clear, and allows for the separation of action space from extreme vista space through the use of aerial perspective.

On the other hand, in Roman art, in Chinese landscapes, and in many Romanesque works, artists often used a dimming and fuzziness of contour of shapes in the distance which can be said to mimic aerial perspective. These artists did this without any clear use of relative density, and sometimes without clear variation in size.[16] Nonetheless, because the systematic use of aerial perspective (with Leonardo) followed the systematicity of linear perspective (with Alberti), we rank density fourth in our hierarchy and aerial perspective fifth. The complete orderings are given in Table 2.

--Insert Table 2 about here--

With this ordering in place, let us discuss our segmentation of space in the context of pictorial art. The space in pictures is traditionally neither too far nor too near, in action space. Few pictures (even portraits) are composed to portray objects too near, in personal space; and only with the full development of Renaissance art is vista space clearly represented. If we consider the region of action space shown in Figure 1, from about

2 to 30 m, and consider the rankings of their importance given in Table 1, we can then compare those with the rankings within our hierarchy given in Table 2. Interestingly, that rank- order correlation is perfect ($\underline{r}_s$ = 1.00, $\underline{p}$ < 0.01). Moreover, in vista space the correlation is high as well ($\underline{r}_s$ = 0.87). These results are also shown in Table 3.

Thus, there is a sense in which the depth-threshold data of the various sources of information shown in Figure 1 and the rankings generated from them given in Table 2, predict well the hierarchical use of the sources in pictorial art. Why would this be so? We follow the suggestion of Blatt (1984, p. 26) "The history of art is the history of the artist's overcoming limitations and widening the range of representational possibilities." We suggest that the discovery and the use of the various sources of information about depth is reflected in a concomitance of their obviousness and their ease of implementation. Occlusion is relatively obvious and easy to implement, height and relative size are perhaps a bit less obvious (requiring more of the "artist's perspective" of registering information as it is projected to the eye) and certainly a bit more difficult to implement in a systematic manner, and relative density and aerial perspective not very obvious and technically quite difficult to implement.[17]

**Layout and the Ordering of Successful Source Use in Ontogeny**

A second arena in which to apply our system of spatial segmentation and rankings of information sources concerns the perceptual development of infants and children. Considerable research has been devoted to the discovery of the onset of the use of various types of information about depth and layout, and Yonas and Granrud (1985), in particular, discuss this evidence directly.[18] Thus our method is a straightforward literature search. The onset values for the use of the various sources of information about layout are given in Table 2. The age values for relative size and motion perspective stem from work reported by Yonas and Granrud (1985), that for convergence from von Hofsten (1977), that for stereopsis from Fox, Aslin, Shea, and Dumais (1980), and the value for occlusion is from Baillargeon (1987) for situations of kinetic occlusion.[19] We assume the value for accommodation is the same as that for convergence, since they are reflexively linked, although Banks (1980) and Haynes, White, and Held (1965) have shown full accommodative responses in infants by 3.5 months. Finally, the values for relative density, height in the visual field, and aerial perspective are unknown, but they clearly arise after the others.

***Results and discussion.*** We wish to consider three rank-order correlations, those between ranks of the onset of source use by infants and the rank of their importance as determined by the area under their ordinal depth-threshold functions within each region of space around the observer. Despite the relatively coarse information about the use of some sources, the rank-order correlation between onset and potency within personal space is impressive ($\underline{r}_s$ = 0.80, $\underline{p}$ < 0.01), but those between for action and vista space are not ($\underline{r}_s$s = 0.09 and -0.14, respectively). These results are shown in Table 3.

--Insert Table 3 about here--

This pattern of correlations makes sense. For the young infant, personal space is the only space that matters; outside of a meter or so little interaction of import occurs. Only later, with toddlerdom and beyond, will action space become important. And indeed, our own experience is that many young children have little, if any, appreciation of vista space. A memorable example of this can be found at Grand Canyon National Park. There, in various book shops in concession areas, one can peruse numerous, beautifully produced photographic books of the vistas around the Colorado River basin. But these books are for adults. The children's sections carry books of photographs of indigenous animals and plants—all composed within 10 or so meters of the camera. Apparently, young children cannot easily appreciate the absence of motion perspective and binocular disparity information.

**Layout and its Corpus of Scientific Research**

A third arena in which to apply our system of depth segmentation and rankings is scientific research. However crudely, one can reflect on the overall study of the perception of layout by simply counting the number of articles published on each or several of these sources. The number of reports concerned with each source of information might then be taken as a rough estimate of scientific interest in it, and indirectly as an estimate of its importance in understanding how human beings appreciate depth and layout.

***Method***. The relevant searches can be done through using PsycINFO, a database published and updated annually by the American Psychological Association which includes 1300 scholarly journals, plus dissertations and other reports, collected since 1984. Thus, at the time this analysis was done, ten years of research—1984-1993—could be combed for each source of information, as listed in Table 2.

Searches are not necessarily straightforward. First, one must discern the appropriate set of keywords. Since some of these sources go by multiple names—for example, motion parallax, and motion perspective—and all must be searched. Second, one must be sure that each source of information is studied in the context of the perception of layout, which the APA defines as "depth perception." For example, a keyword search for "binocular" turned up over 700 studies, but a conjunctive search of "binocular" and "depth" yielded only 149: similarly, a search for "stereoscopic" turned up over 300 studies, but a search of "stereoscopic" and "depth" yielded only 137. A search for "diplopia" and "depth" found only one study. Since a conjunctive search for "stereoscopic" and "binocular" and "depth" turned up 58 studies, the total number of relevant entries is 149 + 137 + 1 - 58 = 229. Conjunctive searches of each source with "depth" are also needed to disambiguate a keyword. For example, a search for "occlusion" by itself brought up over 200 entries, but most of these are about teeth; a search for "accommodation" found nearly 650 studies, but not surprisingly most of these are Piagetian in character; and a search for "convergence" yielded almost 1000 studies but these are about everything from economics to mathematics to psychoanalysis. Finally, once the search were complete, each entry was perused to make sure it was a relevant study.

***Results and discussion***. Again we wish to consider three rank order correlations, those between ranks of the scientific articles produced on each source of information and the rank of importance within each region of space around the observer. Interestingly, we find a pattern similar to that with infants. Whereas the rank-order correlation between scientific articles and personal space is impressive ($r_S = 0.92$, $p < 0.01$), those for farther spaces declines rapidly ($r_S = 0.49$ for action space and $r_S = 0.02$ for vista space). These results are shown in Table 3. If the across-source publishing record of our discipline can be taken as an indication of what we researchers regard as important about the perception of layout, then we are indeed a very near-sighted lot.

## CONCLUSIONS

In the beginning of this chapter we asked three questions: Why are there so many sources of information about layout? How is it that we perceive layout with near-metric accuracy when none of these sources yield metric information about it? And can we not do better, theoretically, in understand our perception of layout than simply make a list.

Our answer to the first question begins with Todd's (1985) answer given above; perceiving layout is extremely important to human beings, so important that it must be redundantly specified, so that the redundancy can guard against the failure of any given source, or the failure of any of the assumptions on which a given source is based. But information redundancy is only part of the answer. Different sources of information about layout metrically reinforce and contrast with each other, providing a powerful network of constraints.

Our answer to the second proceeds from this idea. Through the analysis of depth-threshold functions for nine different sources of information about layout one can begin to understand how those sources of information sharing the same-shaped functions across distances can help ramify judgments of layout by serving to correct measurement errors in each. Equally important, those sources differing in function shape across distance serve to enhance resolution generally at the near-edge of some distance domains—those immediately around us, those within which we can act, and those that we simply observe. Together, same-shaped and different-shaped functions constrain each other, providing a better global metric of egocentric distance than any single source provides on its own.

Third, on the basis of our analyses and the pattern of functions shown in Figure 1, we suggest that list making has mislead us about space and layout. Psychologists and other vision scientists have generally considered layout, space, and distance as a uniform commodity in which observers carry out their day-to-day activities. Moreover, as noted by our analyses of scientific studies of the sources of information about layout, our discipline has not done justice to our perception of the far field, of landscapes, and the like. We claim there are at least three kinds of space around each of us, each perceived via different combinations of different sources of information, each combination dependent on the context in which the information is presented. Researchers, we hope, can now focus on particularly combinations of sources within each type of space in an effort to understand how they constrain one another and help provide the near-metric quality of perceived space.

## REFERENCES


Allik, J. & Tuulmets, T. (1991). Occupancy model of perceived numerosity. *Perception & Psychophysics, 49*, 303-314.

Ames, A. (1955). *An interpretive manual for the demonstrations in the Psychology Research Center, Princeton University*. Princeton, NJ: Princeton University Press.

Andersen, G.J. & Braunstein, M.L. (1983). Dynamic occlusion in the perception of rotation in depth. *Perception & Psychophysics, 34*, 356-362.

Anderson, N.H. (1981). *Foundations of information integration theory*. San Diego, CA: Academic Press.

Anderson, N.H. (1982). *Methods of information integration theory*. San Diego, CA:

Arditi, A. (1986). Binocular vision. In K.R. Boff, L. Kaufman, & J.P. Thomas (eds.) *Handbook of Perception and Human Performance,* Vol 1, (pp. 23:1-41). New York: Wiley.

Baillargeon, R. (1987). Object permanence in 3.5 and 4.5-month-old infants. *Developmental Psychology, 23* 655-664.

Baird, J.C. & Biersdorf, W.R. (1967). Quantitative functions for size and distance. *Perception & Psychophysics, 2*, 161-166.

Baird, J.C., Merrill, A.A., Tannenbaum, J. (1979). Studies of the cognitive representation of spatial relations: II. A familiar environment. *Journal of Experimental Psychology: General, 108*, 92-98.

Baird, J.C. & Wagner, M. (1983). Modeling the creation of cognitive maps. In H.L. Pick, Jr. & L.P. Acredolo (ed.) *Spatial orientation: Theory, research, and application* (pp. 321-366). New York: Plenum.

Baird, J.C. (1970). *Psychophysical analysis of visual space*. Oxford: Pergamon Press.

Banks, M.S. (1980). The development of visual accommodation during early infancy. *Child Development, 51*, 646-666.

Bardy, B.G., Baumberger, B., Flückiger, M., & Laurent, M. (1992). On the role of global and local visual information in goal-directed walking. *Acta Psychologica, 81*, 199-210.

Barlow, H.B. (1978). The efficiency of detecting changes of density in random dot patterns. *Vision Research, 18*, 637-650.

Battro, A.M., Netto, S.P. & Rozestraten, R.J.A. (1976). Riemannian geometries of variable curvature in visual space: Visual alleys, horopters, and triangles in big opens fields. *Perception, 5,* 9-23.

Berbaum, K., Tharp, D. & Mroczek, K. (1983). Depth perception of surfaces in pictures: Looking for conventions of depiction in Pandora's box. *Perception, 12*, 5-20.

Berkeley, G. (1709). An essay towards a new theory of vision. Reprinted *in The works of George Berkeley* (1837, pp. 86-116). London: Thomas Tegg & Son.

Berry, R.N. (1948). Quantitative relations among vernier, real, and stereoscopic depth acuities. *Journal of Experimental Psychology, 38*, 708-721.

Biederman, C. (1948). *Art as the evolution of visual knowledge.* Red Wing, Minnesota: Charles Biederman.
Bingham, G. (1993). Form as information about scale: Perceiving the size of trees. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 1139-1161.
Blank, A.A. (1978). Metric geometry in human binocular vision: Theory and fact. In E.L.J. Leeuwenberg & H.F.J.M. Buffart (eds.) *Formal theories of visual perception*, (pp. 82-102). Chichester, England: Wiley.
Blatt, S. J. (1984*). Continuity and change in art.* Hillsdale, N.J.: Erlbaum.
Birnbaum, M.H. (1983). Scale convergence as a principle for the study of perception. In H. Geissler (ed.) *Modern issues in perception* (pp. 319-335). Amsterdam: North-Holland.
Boring, E.G. (1942). *Sensation and perception in the history of experimental psychology*. New York: Appleton-Century-Crofts.
Braunstein, M.L. (1962). Depth perception in rotating dot patterns: Effects of numerosity and perspective. *Journal of Experimental Psychology, 64,* 415-420.
Braunstein, M.L. (1976). *Depth perception through motion*. New York: Academic Press.
Braunstein, M.L. & Stern, K.R. (1980). Static and dynamics factors in the perception of rotary motion. *Perception & Psychophysics, 27,* 313-320.
Bridgeman, B., Van der Heijden, A.H.C., & Velichkovsky, B.M. (1994). A theory of visual stability across saccadic eyemovements. *Behavioral and Brain Sciences, 17*, 247-292.
Bruno, N. & Cutting, J.E. (1988). Minimodularity and the perception of layout. *Journal of Experimental Psychology: General, 117,* 161-170.
Brunswik, E. (1956). *Perception and the representative design of psychological experiments.* Berkeley: University of California Press.
Brunswik, E. & Kamiya, J. (1953). Ecological cue-validity of "proximal" and other Gestalt factors. *American Journal of Psychology, 66,* 20-32.
Bülthoff, H.H., Fahle, M., & Wegmann, M. (1991). Perceived depth scales with disparity gradient. *Perception, 20,* 145-153.
Bülthoff, H.H., & Mallot, H.A. (1988). Interpreting depth modules: Stereo and shading. *Journal of the Optical Society of America, A, 5,* 1749-1758.
Burton, G. & Turvey, M.T. (1990). Perceiving lengths of rods that are held but wielded. *Ecological Psychology, 2,* 295-324.
Burton, H.E. (1945). The optics of Euclid. *Journal of the Optical Society of America, 35,* 357-372.
Carlson, V.R. (1977). Instructions and perceptual constancy. In W. Epstein (ed.*) Stability and constancy in visual perception: Mechanisms and processes* (pp. 217-254). New York: Wiley.
Carr, H. (1935). *An introduction to space perception*. New York: Longmans.
Cavanagh, P. & Leclerc, Y.G. (1990). Shape from shadows. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 3-27.
Chapanis, A., & McCleary, R.A. (1955). Interposition as a cue for the perception of relative distance. *American Journal of Psychology, 48,* 113-132.
Cole, A. (1992*). Perspective*. London: Dorling Kindersley.
Cook, M. (1978). Judgment of distance on a plane surface. *Perception & Psychophysics, 23*, 85-90.
Cutting, J.E. (1986). *Perception with an eye for motion.* Cambridge, MA: MIT Press/Bradford Books.
Cutting, J.E. (1991a). Four ways to reject directed perception. *Ecological Psychology, 3,* 25-34.
Cutting, J.E. (1991b). Why our stimuli look as they do. In G. Lockhead & J.R. Pomerantz (eds.), *Perception of structure: Essays in honor of Wendell R. Garner* (pp. 41-52). Washington, D.C.: American Psychological Association.
Cutting, J.E. (1993). Perceptual artifacts and phenomena: Gibson's role in the 20th century. In S. Massin (ed.) *Foundations of perceptual theory* (pp. 231-260). Amsterdam: North-Holland.
Cutting, J.E. & Bruno, N. (1988). Additivity, subadditivity, and the use of visual information: A reply to Massaro (1988). *Journal of Experimental Psychology: General, 117,* 422-424.
Cutting, J.E., Bruno, N., Brady, N.P., & Moore, C. (1992a). Selectivity, scope, and simplicity of models: A lesson from fitting judgments of perceived depth. *Journal of Experimental Psychology: General, 121*, 364-381.
Cutting, J.E. & Millard, R.M. (1984). Three gradients and the perception of flat and curved surfaces. *Journal of Experimental Psychology: General, 113* 198-216.

Cutting, J.E., Springer, K., Braren, P., & Johnson, S. (1992b). Wayfinding on foot from information in retinal, not optical, flow. *Journal of Experimental Psychology: General, 121*, 41-72.

Daniels, N. (1974). *Thomas Reid's Inquiry*. New York: Burt Franklin.

Da Silva, J.A. (1985). Scales of perceived egocentric distance in a large open field: Comparison of three psychophysical methods. *American Journal of Psychology, 98,* 119-144.

Dees, J.W. (1966). Accuracy of absolute visual distance and size estimation in space as a function of stereopsis and motion parallax. *Journal of Experimental Psychology, 72*, 466-476.

Denton, G.G. (1980). The influence of visual pattern on perceived speed. *Perception, 9*, 393-402.

Dosher, B.A., Sperling, G., & Wurst, S.A. (1986). Tradeoffs between stereopsis and proximity luminance covariation as determinants of perceived 3D structure. *Vision Research, 26*, 973-990.

Dunn, B.E., Gray, G.C., & Thompson, D. (1965). Relative height on the picture-plane and depth perception. *Perceptual and Motor Skills, 21*, 227-236.

Durgin, F.H. (1994). Texture density adaptation and the perceived numerosity and distribution of texture. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 149-169.

Duwaer, A.L. & van den Brink, G. (1981). What is the diplopia threshold*? Perception & Psychophysics, 29*, 295-309.

Ebenholtz, S.M. & Walchli, R.M. (1965). Stereoscopic thresholds as a function of head- and object orientation. *Vision Research, 5*, 455-461.

Edgerton, S.Y. (1975). *The Renaissance discovery of linear perspective*. New York: Basic Books.

Elkins, J. (1992). Renaissance perspectives. *Journal of the History of Ideas, 53*, 209-230.

Epstein, W. (1963). The influence of assumed size on apparent distance. *American Journal of Psychology, 76,* 257-265.

Epstein, W. (1965). Nonrelational judgments of size and distance. *American Journal of Psychology, 78*, 120-123.

Ferris, S.H. (1972). Motion parallax and absolute distance. *Journal of Experimental Psychology, 95*, 258-263.

Fincham, E.F. & Walton, J. (1957). The reciprocal actions of accommodation and convergence. *Journal of Physiology (London), 137,* 488-508.

Fisher, S.K. & Ciuffreda, K.J. (1988). Accommodation and apparent distance. *Perception, 17,* 609-612.

Flückiger, M. (1991). La perception d'objets lointains. In M. Flückiger & K. Klaue (eds.*) La perception de l'environnement* (pp. 221-238). Lausanne, Switzerland: Delachaux et Niestlè.

Flückiger, M. & Baumberger, B. (1988). The perception of an optical flow projection on the ground surface. *Perception, 17,* 633-645.

Fodor, J.A. (1983). *Modularity of mind*. Cambridge, MA: MIT Press/Bradford Books.

Foley, J.M. (1964). Desarguesian property in visual space. *Journal of the Optical Society of America, 54,* 684-692.

Foley, J.M. (1966). Locus of perceived equidistance as a function of viewing distance. *Journal of the Optical Society of America, 56,* 822-827.

Foley, J.M. (1991). Stereoscopic distance perception. In S.R. Ellis, Kaiser, M.K., & A.C. Grunwald (ed.), *Pictorial communication in virtual and real environments* (pp. 559-566). London: Taylor & Francis.

Forel, F.A. (1892). *Le Lèman* (Vol. 2). Geneva, Switzerland: Slatkine.

Fox, R., Aslin, R.N., Shea, S.L., & Dumais, S.T. (1980). Stereopsis in human infants. *Science, 207*, 323-324.

Freeman, R.B. (1966). Absolute threshold for visual slant: The effect of size and retinal perspective. *Journal of Experimental Psychology, 71*, 170-176

Gablik, S. (1976). *Progress in art*. London: Thames and Hudson.

Gehringer, W.L. & Engel, E. (1986). The effect of ecological viewing conditions on the Ames distortion room illusion. *Journal of Experimental Psychology: Human Perception and Performance, 12*, 181-185.

Gerbino, W., Stultiens, C.J., Troost, J.M. (1990). Transparent layer constancy. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 3-20.

Gibson, E.J. & Bergman, R. (1954). The effect of training on absolute estimation of distance over the ground. *Journal of Experimental Psychology, 48,* 473-482.

Gibson, J.J. (1948). Studying perceptual phenomena. In T.G. Andrews (ed.) *Methods of psychology* (pp. 158-188). New York: Wiley.

Gibson, J.J. (1950). *Perception of the visual world.* Boston: Houghton Mifflin.

Gibson, J.J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.

Gibson, J.J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
Gibson, J.J., Olum, P., & Rosenblatt, F. (1955). Parallax and perspective during aircraft landings. *American Journal of Psychology, 68*, 40-51.
Gilchrist, A.L. (1977). Perceived lightness depends on perceived spatial layout. *Science, 195*, 185-187.
Gilchrist, A.L. (1980). When does perceived lightness on perceived spatial arrangement. *Perception & Psychophysics, 28*, 527-538.
Gogel, W.C. (1961). Convergence as a cue to absolute distance. *Journal of Psychology, 52*, 287-301.
Gogel, W.C. (1974). Cognitive factors in spatial responses. *Psychologia, 17*, 213-225.
Gogel, W.C. (1993). The analysis of perceived space. In S. Masin (ed.) *Foundations of perceptual theory* (p. 113-182). Amsterdam: North-Holland.
Gogel, W.C. & Tietz, J.D. (1973). Absolute motion parallax and the specific distance tendency. *Perception & Psychophysics, 13*, 284-292.
Graham, C.H. (1951). Visual perception. In S.S. Stevens (ed.) *Handbook of experimental psychology* (pp. 868-920). New York: Wiley.
Graham, C.H., Baker, K.E., Hecht, M., & Lloyd, V.V. (1948). Factors influencing thresholds for monocular movement parallax. *Journal of Experimental Psychology, 38*, 205-223.
Gregory, R.L. (1966). *Eye and brain*. New York: McGraw-Hill.
Grünbaum, A. (1973). *Philosophical problems of space and time* (2nd ed.). Boston: Reidel.
Gulick, W.L. & Lawson, R.B. (1976). *Human stereopsis*. New York: Oxford.
Hagen, M.A. (1986). *Varieties of Realism: Geometries of representational art*. Cambridge, England: Cambridge University Press.
Hall, E.T. (1966). *The hidden dimension*. New York: Doubleday.
Haynes, H., White, B.L., Held, R. (1965). Visual accommodation in human infants. *Science, 148*, 528-530.
Helmholtz, H. von. (1866). *Physiological optics*, Vol. 3 (J.P Southall, trans., 3rd ed, 1925). Menasha, WI: The Optical Society of America.
Hirst, R.J. (1967). Illusions. *The encyclopedia of philosophy* (Vol. 4, pp 130-133). New York: Macmillan Publishing & The Free Press.
Hobbs, J.A. (1991). *Art in context* (4th ed.) Fort Worth, TX: Harcourt Brace Jovanich.
Hochberg, J. (1966). Representative sampling and the purposes of perceptual research: Pictures of the world, and the world of pictures. In K.R. Hammond (ed.), *The psychology of Egon Brunswik* (pp. 361-381). New York: Holt, Rinehart, & Winston.
Hochberg, J. (1971). Perception. In J.W. Kling & L.A. Riggs (eds.) *Handbook of Experimental Psychology* (3rd ed.) (pp 396-550). New York: Holt, Rinehart, & Winston.
Hofsten, C. von (1976). The role of convergence in space perception. *Vision Research, 16*, 193-198.
Hofsten, C. von (1977). Binocular convergence as a determinant of reaching behavior in infancy. *Perception, 6*, 139-144.
Hopkins, J. (1973). Visual geometry. *Philosophical Review, 82*, 3-34.
Jameson, D. & Hurvich, L.M. (1959) Note on the factors influencing the relation between stereoscopic acuity and observation distance. *Journal of the Optical Society of America, 49*, p. 639.
Indow, T. (1982). An approach to geometry of visual space with no apriori mapping functions: Multidimensional mapping according to Riemannian metrics. *Journal of Mathematical Psychology, 26, 205-236.*
Indow, T. & Watanabe, T. (1984). Parallel- and distance-alleys with moving points in the horizontal plane. *Perception & Psychophysics, 26*, 144-154.
Ittelson, W.H. (1952). *The Ames demonstrations in perception*. Princeton, N.J.: Princeton University Press.
Johansson, G. (1973). Monocular movement parallax and near-space perception. *Perception, 2*, 136-145.
Julesz, B. (1971). *Foundations of cyclopean perception*. Chicago: University of Chicago Press.
Kaplan, G.A. (1969) Kinetic disruption of optical texture: The perception of depth at an edge. *Perception & Psychophysics, 6*, 193-198.
Kaufman, L. (1974). *Sight and mind*. New York: Oxford University Press.
Kellman, P.J. & Shipley, T.F. (1991). A theory of visual interpolation of object perception. *Cognitive Psychology, 23*, 141-221.
Kellman, P.J. & Spelke, E.S. (1983). Perception of partly occluded objects in infancy. *Cognitive Psychology, 15*, 483-524.

Kersten, D., Bülthoff, H.H., Schwartz, B.L., Kurtz, K.J. (1992). Interaction between transparency and structure from motion. *Neural Computation, 4*, 573-589.

Kersten, D., & Legge, G. (1983). Convergence accommodation. *Journal of the Optical Society of America, 73*, 322-388.

Krech, D. & Crutchfield, R.S. (1958). *Elements of psychology*. New York: Knopf.

Koenderink, J.J. (1986). Optic flow. *Vision Research, 26*, 161-180.

Kosslyn, S.M., Pick, H.L., & Fariello, G.R. (1974). Cognitive maps in children and men. *Child Development, 45*, 707-716.

Kruskal, J.B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika, 29*, 707-719.

Kubovy, M. (1986). *The psychology of perspective and Renaissance art.* Cambridge, England: Cambridge University Press.

Künnapas, T. (1968). Distance perception as a function of available visual cues. *Journal of Experimental Psychology, 77*, 523-529.

Landy, M.S., Maloney, L.T., & Young, M.J. (1991). Psychophysical estimation of the human depth combination rule. In P.S. Shenker (ed.) *Sensor fusion III: 3-D perception and recognition, Proceedings of the SPIE, 1383*, 247-254.

Larish, J.F. & Flach, J.M. (1990). Sources of optical information useful for perception of speed of rectilinear-linear self-motion. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 295-302.

Laurent, M. & Cavallo, V. (1985). Roles des modalities de prise d'informations visuelles dans un pointage locomoteur [The role of visual inputs in a locomotion task.] *L'Année Psychologique, 85*, 41-48.

Lee, D.N. (1980). The optic flow field. *Philosophical Transactions of the Royal Society, London, B, 290,* 169-179.

Leeuwenberg, E. (1971). A perceptual coding language for visual and auditory patterns. *American Journal of Psychology, 84,* 307-347.

Leibowitz, H.W., Johnson, C.A., & Isabelle, E. (1972). Peripheral motion detection and refractive error. *Science, 177,* 1207-1208.

Leibowitz, H.W., Shina, K., & Hennessy, H.R. (1972). Oculomotor adjustments and size constancy. *Perception & Psychophysics, 12*, 497-500.

Lie, I. (1965). Convergence as a cue to perceived size and distance. *Scandinavian Journal of Psychology, 6,* 109-116.

Loomis, J.M., DaSilva, J.A., Fujita, N., & Fukushima, S.S. (1992). Visual space perception and visually directed action. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 906-921.

Luneburg, R.K. (1947). *Mathematical analysis of binocular vision*. Princeton, N.J.: Princeton University Press.

Luo, R.C. & Kay, M.G. (1989). Multisensor integration and fusion in intelligent systems. *IEEE Transactions on Systems, Man, and Cybernetics, 19,* 901-931.

Luo, R.C. & Kay, M.G. (1992). Data fusion and sensor integration: State-of-the-art 1990s. M.A. Abidi & R.C. Gonzalez (eds.) *Data fusion in robotics and machine intelligence* (pp.7-135). San Diego: Academic Press.

Maloney, L.T. & Landy, M.S. (1989). A statistical framework for robust fusion of depth information. In W.A. Perlman (ed.) *Visual communication & image processing, IV, Proceedings of the SPIE, 1199*, 1154-1163.

Marr, D. (1981). *Vision*. San Francisco: Freeman.

Massaro, D.W. (1987). *Speech perception by ear and eye: A paradigm for psychological research.* Hillsdale, N.J.: Erlbaum Associates.

Massaro, D.W. (1988). Ambiguity in perception and experimentation. *Journal of Experimental Psychology: General, 117,* 417-421.

Massaro, D.W. (1989). Multiple book review of *Speech perception by ear and eye: A paradigm for psychological inquiry. Behavioral and Brain Sciences, 12*, 741-794.

Massaro, D.W. & Cohen, M.M. (1993). The paradigm and the fuzzy logical model of perception are alive and well. *Journal of Experimental Psychology: General*, 122, 115-124.

Massaro, D.W. & Friedman, D. (1990). Models of integration given multiple sources of information. *Psychological Review, 97*, 225-252.
Messaris, P. (1994). *Visual literacy: Image, mind, & reality*. Boulder, CO: Westview Press.
Metelli, F. (1974). The perception of transparency. *Scientific American, 230*(4), 90-98.
Minnaert, M. (1993). *Light and color in the outdoors*. (Originally published in Dutch in 1973). New York: Springer Verlag.
Morgan, M.W. (1968). Accommodation and convergence. *American Journal of Optometry and Archives of American Academy of Optometry, 45,* 417-454.
Nagata, S. (1991). How to reinforce perception of depth in single two-dimensional pictures. In S.R. Ellis, Kaiser, M.K., & A.C. Grunwald (ed.), *Pictorial communication in virtual and real environments*, (pp. 527-545). London: Taylor & Francis.
Nakayama, K. (1977). Geometric and physiological aspects of depth perception. *SPIE: Three-Dimensional Image Processing, 120,* 2-9.
Nakayama, K., Shimojo, S. & Ramachandran, V.S. (199). Transparency: Relation to depth, subjective contours, luminance, and neon color spreading. *Perception, 19*, 497-513.
Nakayama, K., Shimojo, S., & Silverman, G.H. (1989). Stereoscopic depth: Its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception, 18,* 55-68.
Nawrot, M. & Blake, R. (1991.) The interplay between stereopsis and structure from motion. *Perception & Psychophysics, 49*, 230-244.
Ogle, K.O. (1952). On the limits of stereoscopic vision. *Journal of Experimental Psychology, 48*, 50-60.
Ogle, K.O. (1958). Note on stereoscopic acuity and observation distance. *Journal of the Optical Society of America, 48*, 794-798.
Ono, H., Rogers, B.J., & Ohmi, M., & Ono, M.E. (1988). Dynamic occlusion and motion parallax in depth perception. *Perception, 17,* 255-266.
Pirenne, M.H. (1970). *Optics, painting, & photography*. Cambridge, England: Cambridge University Press.
Prazdny, K. (1983). Information in optic flows. *Computer Vision, Graphics, and Image Processing, 22*, 235-259.
Purdy, J. & Gibson, E.J. (1955). Distance judgments by the method of fractionation. *Journal of Experimental Psychology, 50*, 374-380.
Putnam, H. (1963). An examination of Grünbaum's philosophy of geometry. In S. Baumrin (ed.) *Philosophy of Science* (pp. 205-215). New York: Interscience.
Ratoosh, P. (1949). On interposition as cue for the perception of distance. *Proceedings of the National Academy of Sciences, Washington, 35*, 257-259.
Richards, W. & Miller, J.F. (1969). Convergence as a cue to depth. *Perception & Psychophysics, 5*, 317-320.
Rieser, J.J., Ashmead, D.H., Talor, C.R., & Youngquist, G.A. (1990). Visual perception and the guidance of locomotion without vision to previously seen targets. *Perception, 19*, 675-689.
Rogers, B.J. & Collett, T.S. (1989). The appearance of surfaces specified by motion parallax and binocular disparity. *Quarterly Journal of Experimental Psychology, 41*, 697-717.
Runeson, S. (1988). The distorted-room illusion, equivalent configurations, and the specificity of static optic arrays. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 295-304.
Schwartz, & Sperling, G. (1983). Luminance controls the perceived 3-D structure of dynamic 2-D displays. *Bulletin of the Psychonomic Society, 21*, 456-458.
Sedgwick, H. (1973). The visible horizon. Ph.D. Dissertation, Cornell University. *Dissertation Abstracts International*, 1973-1974, *34*, 1301-1302B (University Microfilms, 73-22,530).
Sedgwick, H. (1983). Environment-centered representation of spatial layout: Available visual information from texture and perspective. In J. Beck, B. Hope, & A. Rosenfeld (eds.) *Human and machine vision*, (pp. 425-458). New York Academic Press.
Sedgwick, H. (1986). Space perception. In K.R. Boff, L. Kaufman, & J.P. Thomas (eds.*) Handbook of Perception and Human Performance*, Vol 1, pp. 21:1-57. New York: Wiley.
Shepard, R.N. (1980). Multidimensional scaling, tree-fitting, and clustering. *Science, 210,* 390-398.
Simon, H.A. (1969). *The sciences of the artificial*. Cambridge, MA: MIT Press.
Sommer, R. (1969). *Personal Space*. Englewood Cliffs, N.J.: Prentice Hall.
Stevens, K.A. (1981). The information content of texture gradients. *Biological Cybernetics*, 42, 95-105.

Stevens, K.A. & Brookes, A. (1988). Integrating stereopsis with monocular interpretations of planar surfaces. *Vision Research, 28*, 371-386.
Suppes, P. (1977). Is visual space Euclidean? *Synthese, 35*, 397-421.
Taylor, P. (ed. & trans.) *The notebooks of Leonardo da Vinci*. New York: The New American Library.
Teghtsoonian, M. & Teghtsoonian, R. (1969). Scaling apparent distance in a natural outdoor setting. *Psychonomic Science, 21*, 215-216.
Teichner, W.H., Kobrick, J.L., & Wehrkamp, R.F. (1955). The effects of terrain and observation distance on relative depth perception. *American Journal of Psychology, 68*, 193-208.
Terzopoulos, D. (1986). Integrating visual information from multiple sources. In A. Pentland (ed.) *From pixels to predicates* (pp. 111-142). Norwood, NJ: Ablex.
Thomson, J. (1980). How do we use visual information to control locomotion? *Trends in Neuroscience, 3*, 247-250.
Todd, J.T. (1985). Perception of structure from motion: Is projective correspondence of moving elements a necessary condition? *Journal of Experimental Psychology: Human Perception and Performance, 11*, 689-710.
Toye, R.C. (1986). The effect of viewing position on the perceived layout of scenes. *Perception & Psychophysics, 40*, 85-92.
Ullman, S. (1979). *The interpretation of motion*. Cambridge, Mass.: MIT Press.
Ullman, S. (1983). Recent computational studies in the interpretation of structure from motion. In J. Beck, B. Hope, & A. Rosenfeld (eds.) *Human and machine vision* (pp. 459-480). New York: Academic Press.
Uomori, K., & Nishida, S. (1994). The dynamics of the visual system in combining conflicting KDE and binocular stereopsis cues. *Perception & Psychophysics, 55*, 526-536.
van der Meer, H.C. (1979). Interrelation of the effects of binocular disparity and perspective cues on judgments of depth and height. *Perception & Psychophysics, 26*, 481-488.
Vishton, P.M., Nijhawan, R., & Cutting, J.E. (1994). Moving observers utilize static depth cues in determining their direction of motion. *Investigative Opthamology & Visual Science, 35*, p. 2000.
Wagner, M. (1985). The metric of visual space. *Perception & Psychophysics*, 38, 483-495.
Wallach, H. & Frey, K.J. (1972a). Adaptation in distance perception based on oculomotor cues. *Perception & Psychophysics, 11*,77-83.
Wallach, H. & Frey, K.J. (1972b). The nature of adaptation in distance perception based on oculomotor cues. *Perception & Psychophysics, 11*, 110-116.
Wallach, H. & Karsh, E.B. (1963a). Why the modification of stereoscopic depth-perception is so rapid. *American Journal of Psychology, 76*, 413-420.
Wallach, H. & Karsh, E.B. (1963b). The modification of stereoscopic depth-perception and the kinetic-depth effect. *American Journal of Psychology, 76*, 429-435.
Wallach, H., Moore, M.E., Davidson, L. (1963). Modification of stereoscopic depth-perception. *American Journal of Psychology, 76*, 191-204.
Wallach, H. & Norris, C.M. (1963). Accommodation as a distance cue. *American Journal of Psychology, 76*, 659-664.
Wallach, H. & O'Connell, D.N. (1953). The kinetic depth effect. *Journal of Experimental Psychology, 45*, 205-217.
Wanger, L.R., Ferwerda, J.A., & Greenberg, D.P. (May, 1992). Perceiving the spatial relationships in computer-generated images. *IEEE Computer Graphics, 12*, 44-59. TITLE?
Watson, A.B. (1978). A Riemannian geometric explanation of the visual illusions and figural aftereffects. In E.L.J. Leeuwenberg & H.F.J.M. Buffart (eds.) *Formal theories of visual perception*, (pp. 139-169). Chichester, England: Wiley.
Watson, J.S., Banks, M.S., von Hofsten, C., Royden, C.S. (1992).Gravity as a monocular cue for perception of absolute distance and/or absolute size. *Perception, 12*, 259-266.
Wheatstone, C. (1838). Contributions to the physiology of vision: I. On some remarkable and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society, London, 128*, 371-394.
Witkin, A.P. & Tenenbaum, J.M. (1983). On the role of structure in vision. In J. Beck, B. Hope, & A. Rosenfeld (eds.) *Human and machine vision* (pp. 481-543). New York: Academic.

Wohlwill, J.F. (1964). Changes in distance judgments as a function of corrected and uncorrected practice. *Perceptual and Motor Skills, 19,* 403-413.
Woodworth, R.S. & Schlosberg, H. (1954). *Experimental Psychology* (2nd ed.). New York: Wiley.
Wright, L. (1983). *Perspective in perspective*. London: Routledge & Kegan Paul.
Yonas, A., Craton, L.G., & Thompson, W.B. (1987). Relative motion: Kinetic information for the order of depth at an edge. *Perception & Psychophysics, 41*, 53-59.
Yonas, A., & Granrud, C.E. (1985). The development of sensitivity to kinetic, binocular, and pictorial depth information in human infants. In D.J. Ingle, M. Jeannerod, & D.N. Lee (eds.) *Brain mechanisms and spatial vision* (pp. 113-145). Dordrecht, The Netherlands: Martinus Nijhoff.
Zegers, R.T. (1948). Monocular movement parallax thresholds as functions of field size, field position, and speed of stimulus movement. *Journal of Psychology, 26*, 477-498.

## AUTHOR NOTE

Both authors are at the Department of Psychology, Cornell University. Correspondence should be addressed to James E. Cutting, Department of Psychology, Uris Hall, Cornell University, Ithaca, NY 14853-7601 USA; electronic mail may be sent to jec7@Cornell.edu.

This research was supported by National Science Foundation Grant SBR-9212786 and by a John Simon Guggenheim Memorial Fellowship during 1993-1994 both to James E. Cutting. James Cutting thanks Claude Bonnet, Juan Segui, and particularly Jean Lorenceau at the Laboratoire de Psychologie Expérimentale of the Université René Descartes (Paris V) for their graciousness during a sabbatical leave; to Walter Gerbino and the Dipartimento di Psicologia of the Università degli studi di Trieste during a brief tenure as a visiting professor; and to Nan Karwan for long-term discussions on this topic.

## FOOTNOTES

[1] Consistent with Gibson (1966, 1979) we prefer to use the term perception of layout rather than the perception of depth or of space. Strictly speaking observers do not perceive depth, but objects in depth; and they do not perceive space, but objects in space.

[2] Not all psychologists accept the existence of multiple sources of information for a given object or event (see Burton & Turvey, 1990, vs. Cutting, 1991a, 1992a). We claim this belief separates those who espouse direct perception from directed perception (Cutting, 1986).

[3] The term cue comes from 16th century theater documentation and the abbreviation q, for *quando*, Latin for "when." As appropriated by psychology, but continuing theatrical play, a cue entails a knowledgeable observer, one with foreknowledge about when and how to act upon layout information. The term *cue*, then, is heavily aligned with an empiricist position in any nativist/empiricist debates about layout. We prefer the term *source of information*, despite its length, so as to remain neutral in this debate.

[4] Indeed, many would suggest that such judgments would be foreshortened still further. Direct scaling and related methods are often criticized as being open to "cognitive correction;" most adults know that distances foreshorten as they increase and could easily compensate judgments with this knowledge (Baird, 1970; Carlson, 1977; Gogel, 1974). However, in the context of this chapter we see no reason to separate cognition and perception.

[5] Despite Helmholtz's formulation, and the more technical version of it offered by Ratoosh (1949), Chapanis and McCleary (1955) found more to occlusion than good continuation. Figural properties entered as well, an idea well captured by Leeuwenberg (1971) and Kellman and Shipley (1991).

[6] Where is the angular subtense, d is the distance to the object, and h is the height of the object, and where that object has a planar face at right angles to the observer, and where the observer is looking at the middle of the measured surface extent. Under these constraints the formula is:

$$\partial = 2 * \arctan (h / [d*2])$$

[7] The assumption of similarity, and not identity, is also important when considering texture. A texture surface, by conventional definition, is one peppered with similar texture elements; these typically lie flat on the object surface and will be at various orientations to the observer, yielding nonidentical retinal projections.

[8] The experiments of Teichner et al (1955) appear to have covaried relative size and height in the visual field. Our analysis suggests height is the more powerful sources of information than relative size in the near field. Nonetheless, it makes sense that relative size should not attenuate with distance and that its threshold value should be near 3%.

[9] However, even without a metric assumption size can still be scaled in terms of eye height. There is, for example, information in the horizon-ratio about size (e.g. Sedgwick, 1973). It is useful to assume that the projected height of the horizon is at the same height as the eye, which Edgerton (1975) called *the horizon-line isocephaly*. In the flattest terrestrial situations it is at an angle of about $89.98^{o}$ with respect to the gravitational vertical (Cutting, 1986). Thus, anything intersecting the projection of the horizon line at the point of observation (the eye) is at the same height as the eye; anything or anyone whose height is have the distance from its based to the horizon line is half the height of the point of observation; anything which has the horizon line bisect its vertical extent is twice the height of the point of observation. In this manner, scaled information is available in the optical array relative to the point of observation; again, if eye height is known, then absolute size and absolute distance information is available, although will not necessarily used.

[10] It is said that Israeli pilots flying over the Negev desert often have problems maintaining altitude. The Negev has stochastically regular texture (dunes) but the frequencies and amplitudes gradually change north to south. In maintaining edge rates, the pilots must increase altitude as the texture density decreases (Shimon Ullman, personal communication).

[11] Data are available for relative depth-through-motion threshold up to the equivalent of 10 m for a pedestrian (Zegers, 1948). Beyond 10 m there are no data, but the motion generated by the pedestrian is so slow that any object could be easily fixated during pursuit motions. Once fixated, the absolute motion thresholds can be used.

[12] Of course, convergence is also yoked to stereopsis, and when the vergence system runs awry, rendering it difficult or impossible to foveate the same part of an object with both eyes, strabismus results.

[13] It is clear our visual system makes some roughly accurate assumption about interocular distance. In this vein, it is interesting that early stereoscopic pictures—showing cityscapes and other scenes of large expanse—enhanced the distance between the eyes. In turn, these diminished the effective size of the objects seen. Thus, many early American stereoscopic pictures were visual oxymorons—presenting sequoias (which are large, but then diminished to near normal size in appearance by inter-camera distance) in conjunction with people. Rather than impressing the (late-20th century) observer with the size of sequoias they impress one with the diminutive nature of the people next to them.

[14] Size gradient here was called the perspective gradient by Cutting and Millard (1984), and the scaling gradient by Stevens (1981).

[15] It is sometimes difficult to be sure that in height in the visual field is being used in pre-Romanesque art or art outside of the Western tradition. In Greek vases and in Egyptian wall paintings, smaller people and objects are often found in the interstices between major figures, but it is not clear whether this is done for reasons of balance and other esthetic concerns or whether depth and distance layout are being portrayed.

[16] Related, but separate, from aerial perspective in Japanese art is the cloud convention (Hagen, 1986), where clouds are inserted in the image so that a story line can be told. The clouds separate one story line from another.

[17] Yonas and Granrud (1985), like many others, compose a list of information sources which is different than ours. We have not included reference to their discussion of shading, familiar size, and the accretion and deletion of texture.

[18] Stereopsis and motion, of course, are technically quite difficult and could only develop artistically with the development of photography and cinema. Blur (the photographic analog to misaccommodation) is also generally not found in art until after photography.

[19] We also have evidence of the second author's working in the laboratory of Elizabeth Spelke that dynamic occlusion information is used by infants of 3 mo and younger in many situations.

**Table 1: Rankings of information sources by the areas under their curves in Figure 1 within the three kinds of space**

| Source of Information | Personal Space | Action Space: all sources | Action Space: pictorial sources | Vista Space |
|---|---|---|---|---|
| 1. occlusion, interposition | 1 | 1 | 1 | 1 |
| 2. relative size | 4 | 3.5 | 3 | 2 |
| 3. relative density | 7 | 6 | 4 | 4.5 |
| 4. height in visual field, height in the picture plane | --- | 2 | 2 | 3 |
| 5. aerial perspective, atmospheric perspective | 8 | 7 | 5 | 4.5 |
| 6. motion perspective, motion parallax | 3 | 3.5 | --- | 6 |
| 7. convergence | 5.5 | 8.5 | --- | 8.5 |
| 8. accommodation | 5.5 | 8.5 | --- | 8.5 |
| 9. binocular disparity, stereopsis, diplopia | 2 | 5 | --- | 7 |

**Table 2: Sources of Information about depth, hierarchical development in pictures, age of their onset for use by children, and citations from PsycINFO (1984+).**

| Source | Development in Art | Onset Age in Infants (mo) | Number of Scientific Articles |
|---|---|---|---|
| 1. occlusion, interposition | 1. used alone | 3.0 | 33 |
| 2. relative size | 3. used with occlusion, and typically with height in the visual field | 5.5 | 29 |
| 3. relative density | 4. used only with full development of linear perspective, and typically before systematic aerial perspective in Western tradition | 7+ | 4 |
| 4. height in plane, height in visual field | 2. used with occlusion | 7+ | 19 |
| 5. aerial perspective, atmospheric perspective | 5. perfected after linear perspective, but also used before. | 7+ | 0 |
| 6. motion parallax, motion perspective | -- | 5.0 | 38 |
| 7. convergence | -- | 4.5 | 24 |
| 8. accommodation | -- | 4.5 | 15 |
| 9. binocular disparity, stereopsis, diplopia | -- | 3.5 | 229 |

**Table 3:**

**Rank-order correlations of the prominence of the sources of information within each space with ordered use of sources across the history of art, with ordered development in infants, and with ordered popularity in Scientific Research.**

| | Regions of Egocentric Space | | |
|---|---|---|---|
| | Personal | Action | Vista |
| Hierarchy in art | ---[a] | 1.00[b] | 0.87 |
| Onset of use by infants | 0.80[b] | 0.09 | -0.14 |
| Scientific articles | 0.92[b] | 0.49 | 0.02 |

[a] too few sources to assess by correlation

[b] $\underline{p} < 0.01$

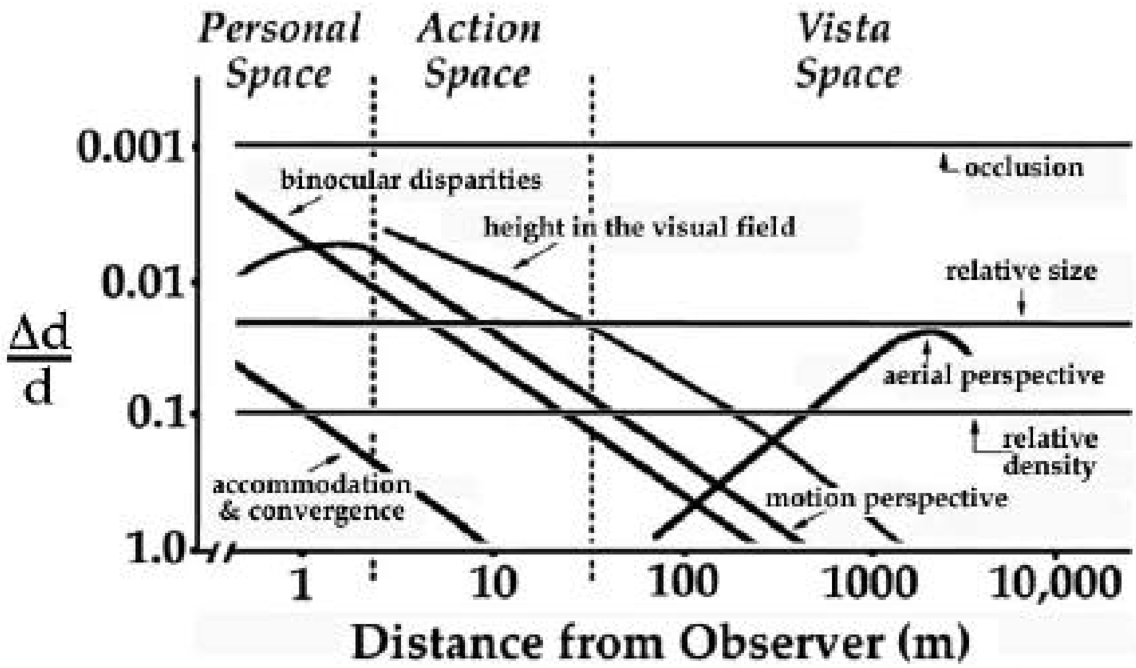
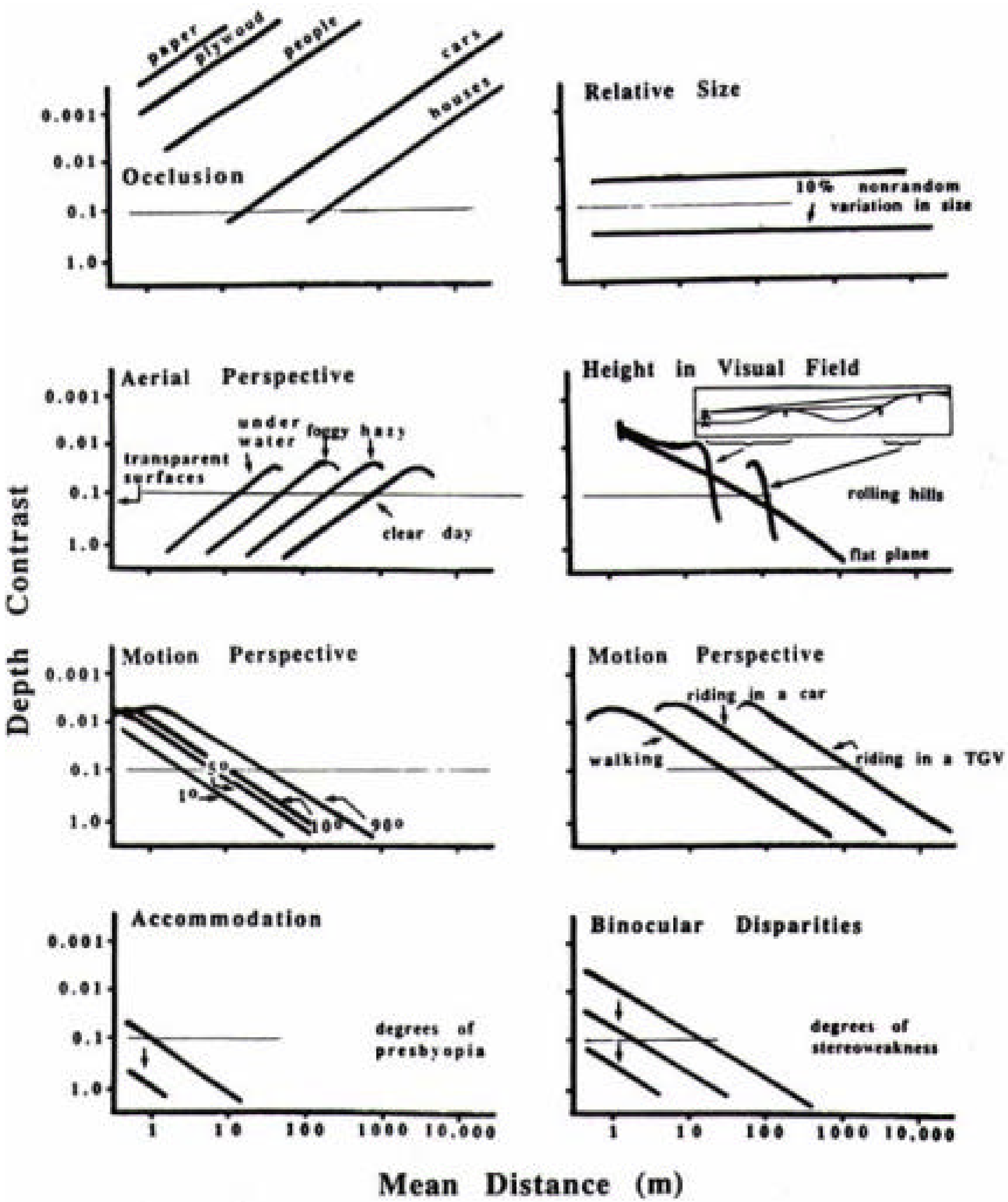



**Figure 1**: Just-discriminable depth thresholds as a function of the log of distance from the observer, from 0.5 to 5000 meters, for nine different sources of information about layout. Such plots were originated by Nagata (1981), and are extensively modified and elaborated here; they are plotted with analogy to contrast sensitivity functions. Our assumption is that more potent sources of information are associated with smaller depth-discrimination thresholds; and that these threshold functions reflect suprathreshold utility. These functions, in turn, delimit three types of space around the moving observer—personal space, action space, and vista space—each served by different sources of information, with different weights. This array of functions, however, is idealized; Figure 2 shows variations on the themes shown here.

**Figure 2 (next page)**: Eight variations in the depth-threshold functions for various sources of information under conditions which are either nonideal, or which violate some assumptions made in Figure 1. The first panel shows occlusion depth threshold functions for five different substances—paper, plywood, people, cars, and houses. The function in Figure 1 is flat because of the parallel cascade of these rising functions. The second panel shows the decrease in effectiveness of relative size, when there is a 10% variation in size. When there is 10% random variability in size across the whole layout, it is not clear there would be any diminuation of this function. The third panel shows the change in threshold function of height in the visual field, when the surface of support is nonplanar. The fourth shows the functions for aerial perspective under different atmospheric conditions, plus the function underwater and a suggestion about transparency. The fifth panel shows four functions for motion perspective; that for 90° to the observer's linear path (as shown in Figure 1), and those for 10, 5, and 1° from the path; and the sixth the 90° function for a pedestrian, a passenger in an automobile, and a passenger in a fast train (the TGV in France). The seventh panel shows the decrease of the efficacy of accommodation with age, and the eighth the possibilities of different degrees of stereoweakness. We claim that these types of variations in these functions occur across contexts and individuals, and modify the orderings of importance (the weights) in the integration of information about layout.


Occlusion
paper
plywood
people
cars
houses
Relative Size
10% nonrandom
variation in size
Aerial Perspective
transparent
surfaces
under
water
foggy
hazy
clear day
Height in Visual Field
rolling hills
flat plane
Motion Perspective
1°
5°
10°
90°
Motion Perspective
walking
riding in a car
riding in a TGV
Accommodation
degrees of
presbyopia
Binocular Disparities
degrees of
stereoweakness
0.001
0.01
0.1
1.0
1
10
100
1000
10,000
Depth Contrast
Mean Distance (m)