Disparity statistics in natural scenes

Yang Liu Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX, USA

Alan C. Bovik Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX, USA

Lawrence K. Cormack Department of Psychology, The University of Texas at Austin, Austin, TX, USA

Binocular disparity is the input to stereopsis, which is a very strong depth cue in humans. However, the distribution of binocular disparities in natural environments has not been quantitatively measured. In this study, we converted distances from accurate range maps of forest scenes and indoor scenes into the disparities that an observer would encounter, given an eye model and fixation distances (which we measured for the forest environment, and simulated for the indoor environment). We found that the distributions of natural disparities in these two kinds of scenes are centered at zero, have high peaks, and span about 5 deg, which closely matches the macaque MT cells' disparity tuning range. These ranges are fully within the operational range of human stereopsis determined psychophysically. Suprathreshold disparities (>10 arcsec) are common rather than exceptional. There is a prevailing notion that stereopsis only operates within a few meters, but our finding suggests that we should rethink the role of stereopsis at far viewing distances because of the abundance of suprathreshold disparities.

Keywords: stereopsis, binocular disparity, stereoacuity, stereo upper limit, Panum's fusion area

Citation: Liu, Y., Bovik, A. C., & Cormack, L. K. (2008). Disparity statistics in natural scenes. *Journal of Vision, 8*(11):19, 1–14, http://journalofvision.org/8/11/19/, doi:10.1167/8.11.19.

Introduction

Stereopsis is the ability to see depth using retinal disparities, which are the differences in relative image position of scene points on the two retinae. It is widely thought (Palmer, 1999) that stereopsis does not operate beyond a few meters because the retinal disparity produced by a given depth interval decreases dramatically with distance. Further, binocular disparity is a relative cue and must be scaled by absolute distance information. The most obvious sources of this information, convergence and accommodation, cease to be effective beyond a few meters, so it is not surprising that stereopsis is assumed to behave similarly. Nobody, however, has measured the disparities actually occurring outdoors and examined their relationship to the visual system. It is easy to show that suprathreshold disparities can indeed occur at far viewing distances (e.g., von Helmholtz, 1962; Tyler, 1991), but if they are abundant or even in the majority rather than being a curious exception, then we should perhaps rethink the role of stereopsis in both ordinary vision and in primate evolution.

To determine disparities in the natural world, we must have two pieces of information: the real distance from each direction where the incoming light strikes the retina and the point of binocular fixation. For the former, laser range scanning technology suits this task well. Several researchers collected range maps in the real world (Huang, Lee, & Mumford, 2000; Potetz & Lee, 2003; Yang & Purves, 2003a). Yang and Purves (2003b) placed their laser scanner at a height of 1.65 m, which is a typical eye height of a male adult. We computed disparity distributions in natural scenes using available range maps from Yang and Purves (2003b) and distributions of binocular fixation points that we either measured empirically or simulated. In this paper, we concentrated our analysis on the horizontal plane at eye height because the vast majority of our observers' fixations were near this plane.

Range map data

We formatted the range data into a right hand world coordinate system with metric unit, where the x axis goes from left to right, the y axis goes from down to up, and the z axis points toward the viewer. In each range map, the origin of the coordinate system is the location of the laser scanner.

To calculate exact disparities, we must first have an optical model of the human eye. We approximated the human eye as a perfect sphere with its center at its nodal point (Le Grand, 1980). With an interpupillary distance of

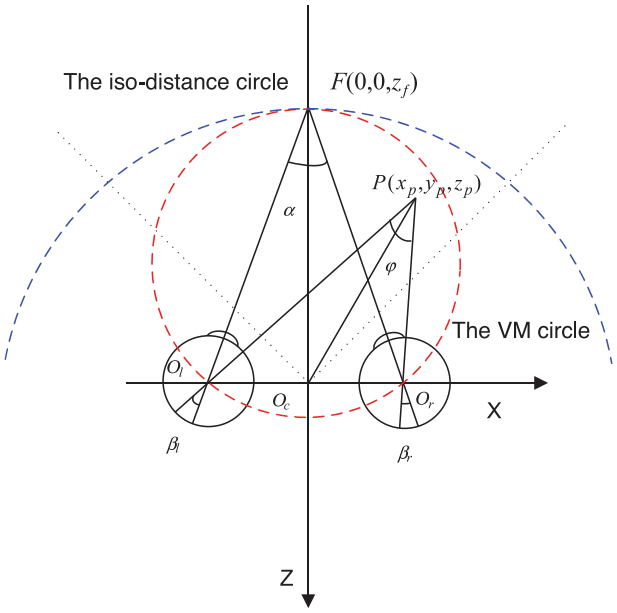

Figure 1. The geometry of binocular disparity computation. When the observer is fixating at point *F*, the vergence angle is α. Another point *P* has two projections on two retinae. The angle between the left visual axis FO_l and the left projection line PO_l is β_l, and the angle between the right visual axis FO_r and the right projection line PO_r is β_r. We define the binocular disparity to be $d = \beta_r - \beta_l$. Theoretically, if the point *P* is inside the ideal horopter defined by the current fixation *F*, the disparity of *P* is negative (crossed). Otherwise it is positive (uncrossed). When *P* is exactly located on the horopter, its two projections are on the corresponding retinal points, which yield 0 disparity.

0.065 m, we assume that the nodal points of the two identical eye balls are located at (−0.0325, 0, 0) and (0.0325, 0, 0), and the observer is looking at the negative direction of the z axis.

Instead of computing the 3D projection on the retinal sphere, we only consider the projection in a particular 2D situation, which is the horizontal great circle of the retinal sphere. All the scene points on the XZ plane are projected on the horizontal great circle with different horizontal disparities but zero vertical disparities. A scene point's horizontal disparity can be accurately calculated when we know the coordinates of the current fixation and its own position.

Figure 1 is a top view of the XZ plane when the two eyes are fixating on a point F (x_f, y_f, z_f). We define the fixation distance to be the distance between the fixation point F and O_c, which is also the midpoint between the two eyes. To simplify the computation, we assume the two eyes have a midsagittal fixation, which means $x_f = 0$ and $y_f = 0$. By doing so, we can use the fixation distance z_f to represent a fixation. The disparity of point P can be specified by:

$$d = \beta_r - \beta_l = \alpha - \varphi \tag{1}$$

where

$$\alpha = 2\,\mathrm{atan}(-0.0325/z_f), \tag{2}$$

$$\varphi = \mathrm{atan}\left(\frac{-x_p - 0.0325}{z_p}\right) - \mathrm{atan}\left(\frac{-x_p + 0.0325}{z_p}\right). \tag{3}$$

Each range map is a geometrical description of the world by a laser scanner centered at O_c. If we think of the laser scanner as the cyclopean eye, then its retina is hit by laser beams at all directions with the same angular sampling step size. Of course, this geometrical description of the world may not be the same when the laser scanner is located at the left eye's position O_l, or the right eye's position O_r. But since the interpupillary distance is much smaller than the distances in the range maps, it is reasonable to assume that the range maps are valid for two eyes at O_l and O_r.

From Equation 1 we can see that the disparity is decided by z_f, x_p, y_p, and z_p. The range map gives x_p, y_p, and z_p. We must decide the position of the fixation point $F(0, 0, z_f)$ beforehand to calculate the disparity. In order to simulate the human fixation in the forest scenes, we conducted the following experiment to collect human fixation distances in an environment similar to the one in which the range data were collected.

Empirical fixation distance distribution

Three observers, one female and two male, with normal or corrected to normal vision, were asked to walk in a wooded park near the campus of the University of Texas at Austin. (Pilot observations showed that some people look at the ground a great deal of the time. Our three observers did not. Obviously, the effects of the ground plan are very interesting and relevant to stereopsis and binocular vision, but in this paper we wanted to concentrate on disparities at large viewing distances along the horizontal meridian.)

Two experimenters followed the observer. The first experimenter, who did not look at the observer while walking, timed random intervals between 2 and 30 seconds. At the end of each random interval, "time" was called and the observer reported the current locus of fixation (The observers found it to be easy after a few trials). The line-of-sight distance to the indicated region was then measured using a laser rangefinder for distances larger than 15 meters. A measuring wheel was used for distances smaller than 15 meters. We collected 100 fixation distances for each observer for a

total of 300 binocular fixation samples. We selected 255 out of the 300 fixation distances by omitting all distances smaller than 2 m, which served to eliminate the rare occasions when the observers were looking at the ground at their feet. We denote the 255 fixation distances as d_1, d_2, … d_{255}. These fixation distances range from 2.7 m to 123.8 m.

Figure 2 shows the distribution of fixation distances we obtained (top right), along with the distribution of distances present in the range data (top left). Both distributions are roughly log-normal, which can be seen as a linear relationship when the quantiles of the log of the data are plotted against equal quantiles from the standard normal distribution (middle). Similarly, when the quantiles of the range data are plotted against equal quantiles of the fixation data, the trend is roughly linear indicating that the two distributions are very similar in shape. This indicates that, effectively, observers are selecting fixation distances in accord with their prevalence in the environment. If this is true, then such an observer could be easily modeled for the purposes of computing disparity distributions, a point to which we will return below.

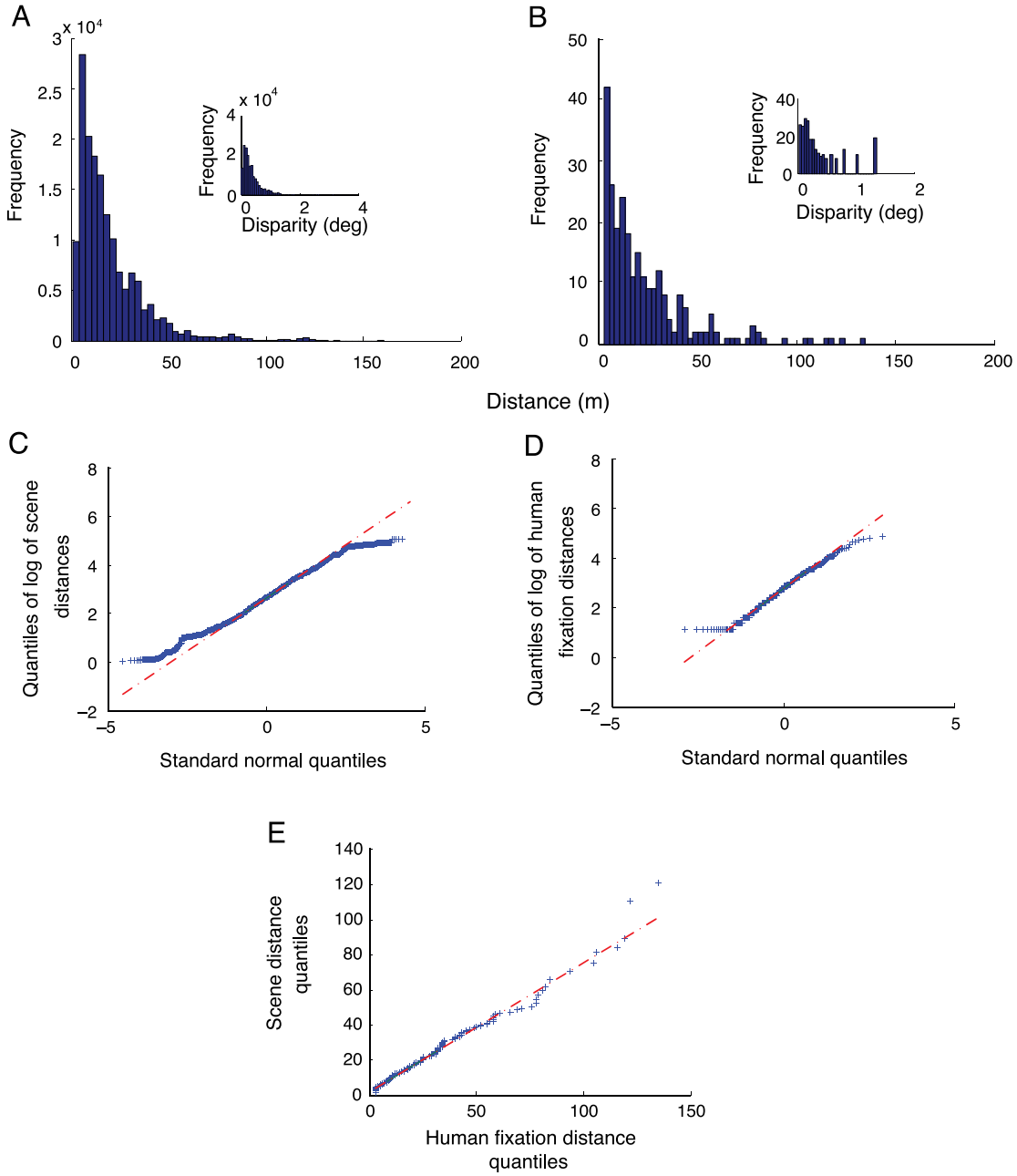

Figure 2. The human fixation distance (A) are roughly log normal distributed and matched well with the scene distances (B) at eye level. See text for details. The insets in the upper row show the scene distances converted to angular units that, for some readers, may facilitate comparisons with the subsequent figures.

The computation of natural binocular disparity

Previous studies show humans generally fixate on relevant objects during daily visual tasks, such as making a sandwich (Hayhoe, Shrivastava, Mruczek, & Pelz, 2003) and making tea (Land, Mennie, & Rusted, 1999). Similarly, our observers always reported fixating on a specific object, generally on or near the horizontal plane. We used the following method to combine our human fixation distances with the range maps.

For each fixation distance d_i, we went through the 23 forest range maps to find a matched distance d on the horizontal plane, with a criterion of the error $(d_i - d) / d_i < 5\%$. This matching process gave us (hopefully) a good intuitive association between the range maps and our fixation data, the idea being "Given that observers generally fixate on objects, and given that this observer was fixating at a distance d, where could he or she have been plausibly looking in the range data?" After we found the matched distance d, we placed a virtual observer with the midpoint between two eyes located at the origin O_c and directly facing the direction of the matched distance d.

Figure 3 shows this process schematically for fixations 1, 2, and 255. The top row shows each fixation along with the 631 scene points within a horizontal 90-deg angle centered on fixation (the horizontal resolution of the range maps is about 0.143 deg). The next row shows the disparities computed for each of these points, and the third row shows the histograms of these disparities for each fixation point (column). Finally, the data were combined across fixations to yield the overall disparity distribution shown at the bottom.

This distribution of disparities is shown in greater detail in Figure 4A. This plot represents our best estimate of the distribution of angular disparities available to an observer looking straight ahead in a natural outdoor wooded environment. The peak is very near zero, but the span is over 4.5 deg. The maximum of disparity is 1.24 deg, which corresponds to the nearest fixation at about 2.66 m and a farthest scene point at about 160 m; the minimum is −3.30 deg, which corresponds to the farthest fixation at

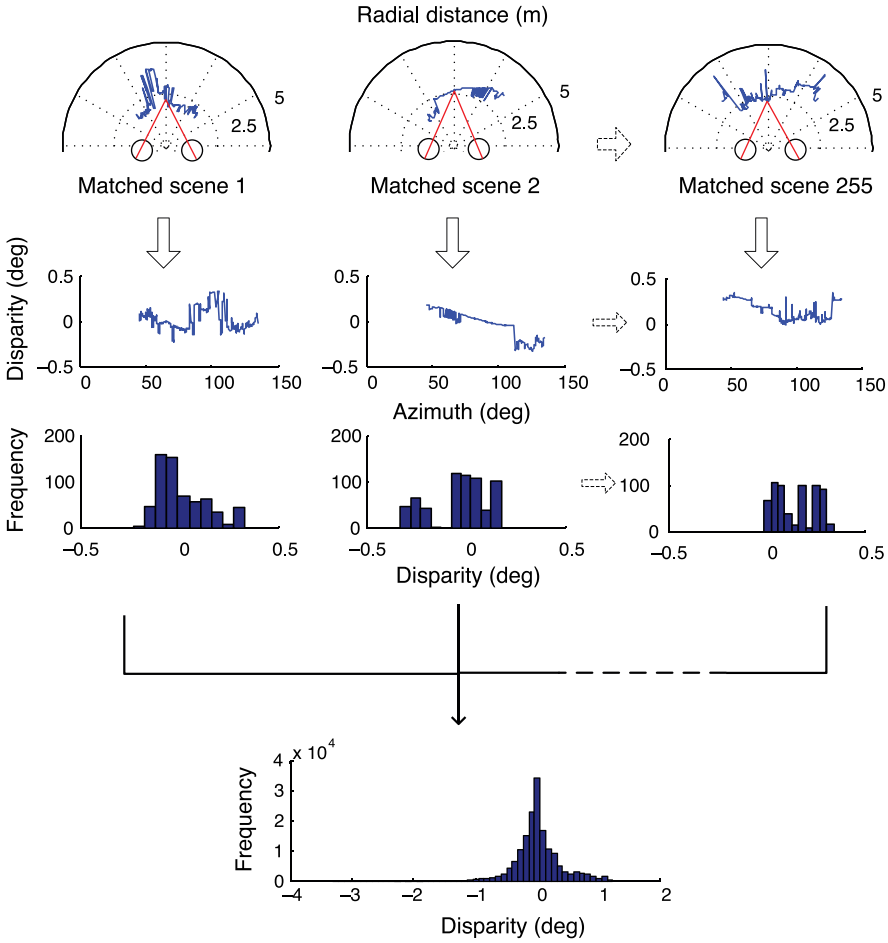

Figure 3. This flowchart shows the process of disparity computation. We find a scene distance matching every human fixation distance and chose its ±45-deg neighborhood to compute disparities. Each 90-deg neighborhood contributes 631 disparities to the overall histogram. The total histogram contains 255 × 631 = 160905 disparities.

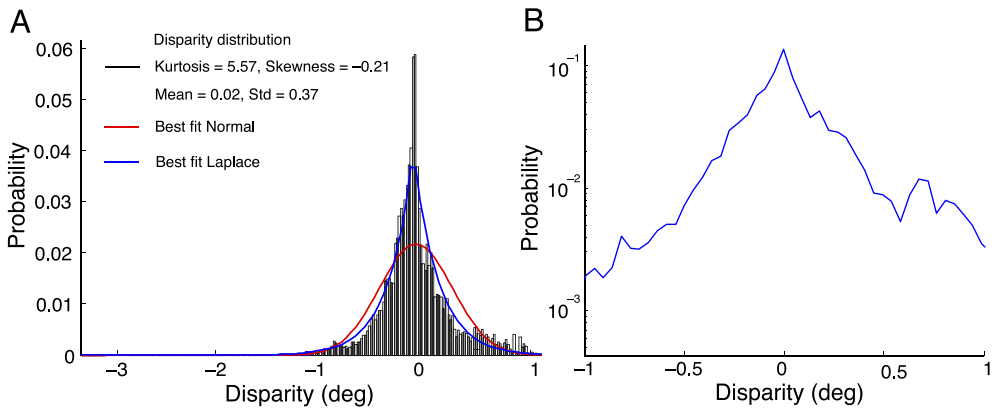

Figure 4. (A) The distribution of binocular disparity derived from 23 forest range maps. The red curve is the best fitting Gaussian on the data. It is visually apparent that the distribution is non-Gaussian because of the higher peak and heavier tails. (B) The distribution between ±1 deg shows a good linearity using semi-log plot.

about 124 m and a nearest scene point at about 1.2 m. The mean is roughly zero deg (4.07 arcmin), and the standard deviation is about 0.37 deg (22.2 arcmin). The kurtosis is 5.57, which is much larger than that of the Gaussian (kurtosis = 0), indicating that there are too many disparities near zero to be consistent with a Gaussian disparity distribution. Indeed, it is apparent that the Gaussian red curve is not a satisfactory fit by eye. We also fitted the Laplace distribution, whose probability density function is $f(x \mid \mu, b) = \exp(-\mid x - \mu \mid /b)/2b$, to the disparity data using the MLE (maximum likelihood estimation) method. The fit is clearly better. Moreover, a plot of log-density vs. disparity is shown in Figure 4B, and the roughly linear flanks are consistent with the Laplace distribution (though the kurtosis Laplace distribution is 3, which is still substantially lower than our obtained value of close to 6). Crucially, however, 98% of the distribution comprises disparity values greater than ±10 arcsec (a large but not unreasonable stereothreshold).

As mentioned above, the distributions of environmental distance and fixation distance are strikingly similar, which is what would be obtained if, for our purposes, observers were selecting a random visual direction, and fixating the nearest object along that direction (in other words, this is the simplest behavior that would lead to a match between the range and fixation distributions shown in Figure 2). To test this, we recomputed the distribution of disparities from the range maps using a simulated observer instead of our empirical distribution of fixation distances. This simulated observer first selected a random (uniform) azimuth for head orientation over ±120 deg (i.e., half the 330-deg field of view of the rangefinder minus the ±45-deg field over which we compute disparities) and then selected the coordinate corresponding to the nearest object in that direction (i.e., the range returned by the rangefinder in that direction) as the fixation point. We did this a total of 255 times (selecting a random range map each time) to yield the same number of disparity estimates as above (255 × 631). The resulting disparity distribution is shown in Figure 5 and is qualitatively very similar to that in Figure 4. The dark blue line in Figure 5 shows the best fitting Laplace distribution for these disparities, while the dashed blue line shows the best fitting Laplace for the disparities estimated using the human fixation data (i.e., re-plotted from Figure 4A). The inset shows the quantile–quantile (Q–Q) plot of the two estimated disparity distributions. The striking agreement (the linearity of the Q–Q plot) indicates that, for the present purposes at least, we can use a model that selects the nearest point along a random visual direction as a fixation point to generate a reasonable approximation of human fixation distances. Obviously, this is not what human observers do, but it produces the same distribution of fixation distances over the long run.

A main point of this study is that the distribution of disparities in large-scale (outdoor) environments is largely suprathreshold, and therefore a potential stimulus to

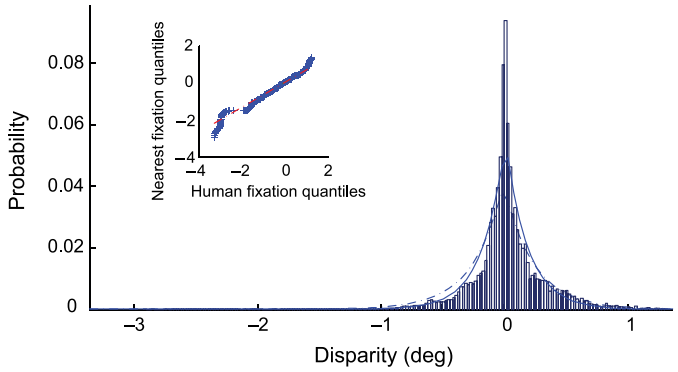

Figure 5. The disparity distribution from nearest fixations with uniformly distributed orientations, the solid blue line shows the best fitting Laplace, and the dashed blue line is the re-plot of the best fitting Laplace of the disparity distribution from human fixation distances. The inset shows the quantile–quantile plot of the two distributions. The good linearity is prominent.

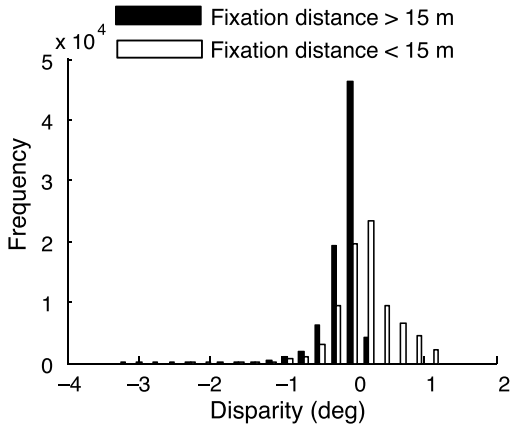

Figure 6. The shape of disparity distribution changes with fixation distance.

stereopsis. It might be argued that the suprathreshold disparities arise largely from very distant objects when fixation is very near, or from very near objects when fixation is at infinity. In Figure 6, we separate the histogram into two sub-histograms, which correspond to fixation distances larger than the median fixation distance of 15 m (red) and those smaller than 15 m (blue). For far fixations, there is a very high peak of near zero disparities corresponding to continuous local far neighborhood around the fixations and also a considerably long tail of crossed disparities from nearer scene points. For nearer fixations, the peak around zero is much lower. This is because the relative disparity between two points of same distance interval is larger when the absolute distances of the two points are closer to an observer. Because of the piecewise continuous structure of natural scenes, a point in the 90-deg scene matched with a nearer fixation distance is more likely to reside in a nearer local neighborhood containing the fixation point, and vice versa with farther fixation distances. After these scene distances have been converted to disparities, the disparities in the nearer scenes differ from those in the farther scenes significantly as Figure 6 shows. But we should stress that the difference between the red and blue plots in Figure 6 is ultimately caused by different fixation distances. It is the upper (>15 m) and lower (<15 m) fixation distances that partition the matched scenes into two different groups and yield the different disparity distributions in Figure 6.

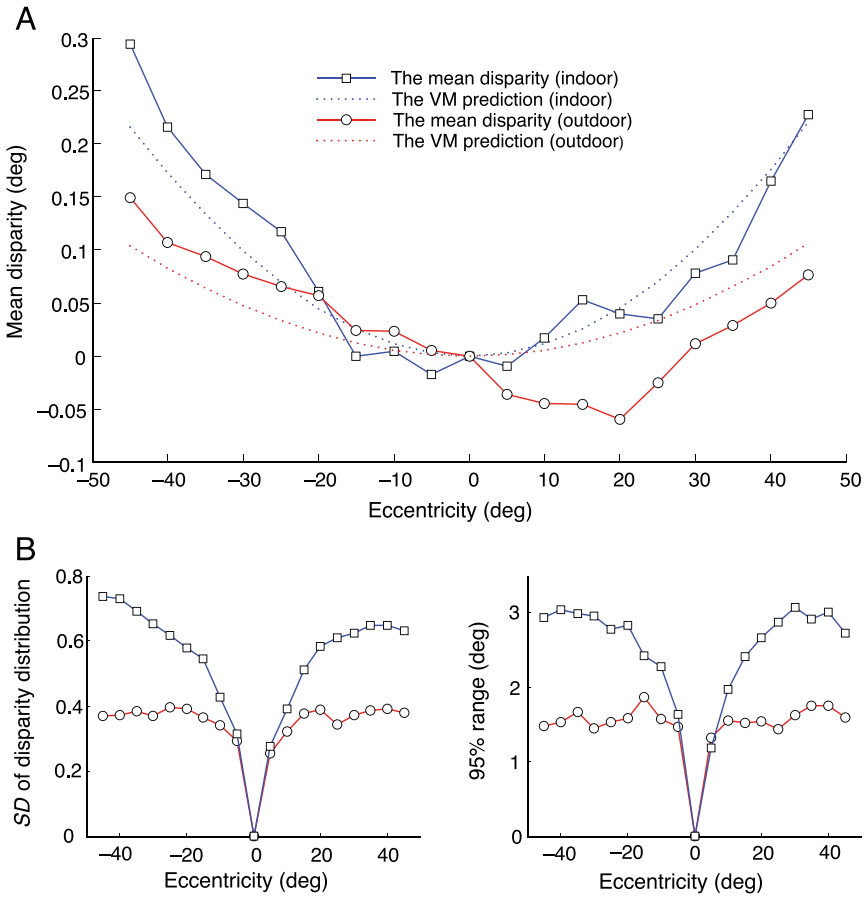

Figure 7. (A) The mean of disparity distributions at different eccentricities from forest scenes (circles) and indoor scenes (squares). (B) The standard deviations of two disparity distributions at different eccentricity. (C) The 95% ranges of two disparity distributions at different eccentricities.

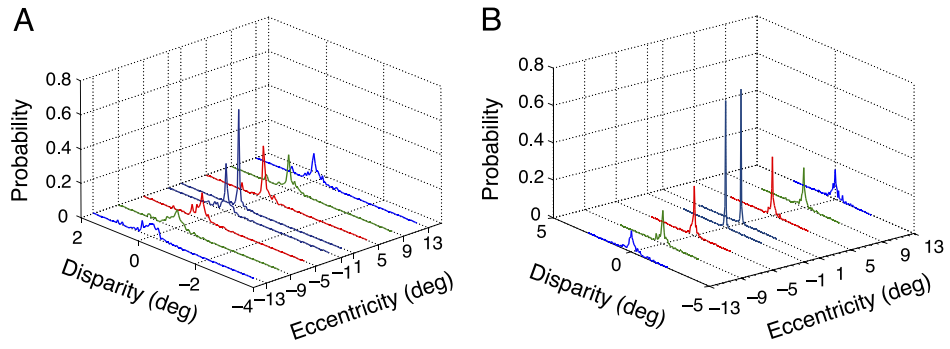

Figure 8. Disparity distributions at different eccentricities in (A) forest scenes and (B) indoor scenes.

We also examined the relationship between sampled disparities and horizontal eccentricity. Figures 7A and 7B show the mean (circles and red line) and standard deviation (circles and red line) of the disparity distribution as a function of horizontal eccentricities (using 5-deg eccentricity bins). Figure 7C (circles and red line) shows the 95% range of the distribution (disparities between the 97.5% percentile and 2.5% percentile) against eccentricities. The monotonic increase of the mean and standard deviation is prominent. Figure 8A shows the pdfs of the disparities at ±13, ±9, ±5, and ±1 deg. It is obvious to see the distributions are all unimodal, roughly peaked near zero, and become broader as the eccentricity increases.

The increase of standard deviation and span of the disparity distribution is easy to understand since the world is piecewise continuous. With eccentricity growing, peripheral scene points are less likely to be located on the same object or surface as the fixation point, which results a larger range of disparity values. Another interesting observation is that the mean disparity also seems to increase with eccentricity, which indicates that the peripheral points tend to have positive (far) disparities. Does this also indicate that peripheral locations are generally farther away than straight-ahead fixations in terms of absolute distance?[1] It is unclear how this could be the case in the current study given that, even if there were some systematic orientation bias in the original range data, it is difficult to imagine how this could propagate to the disparity data given the randomness inherent in our simulations. Nevertheless, we verified the stationarity of the range data (as a whole) with respect to orientation by computing the range distributions (and the means and standard deviations) at many orientations in the original data and, indeed, the mean range, the standard deviation, and (as far as we could tell) the shape of the range distribution were independent of orientation. The only thing remaining that could have produced the orientation dependence was the nature of the disparity calculation, specifically, the VM circle. Referring to Figure 1, assume that the fixation F is such that the mean disparity along OcF is zero. Given the independence of the range data, this condition will hold (at least approximately) for any other fixation point on the iso-distance circle. Thus, if fixation were to remain at F, then the mean disparity at any eccentricity will be shifted by the disparity between the VM circle and the iso-distance circle at that eccentricity. This disparity shift is plotted as the red dashed line in Figure 7A. The agreement is good given the standard deviations associated with the data, and the remaining discrepancy is probably due to the particular sample of ranges we acquired. Note also that the subjective horopter measured by psychophysical methods is a little different from the VM circle (Howard & Rogers, 1995) and is generally between the VM circle and the frontal parallel plane located at the fixation point. Thus, the effective dependence of mean disparity on eccentricity should be less in practice than estimated here.

Discussion

Relationship between different environments and natural disparity distribution

While environments such as room and car interiors probably had a minimal impact on primate evolution, they undoubtedly provide a large amount of input to both the mature and developing human binocular system. Fortunately, Yang and Purves (2003b) also collected range data from 27 indoor scenes at Duke University. The scanner setting was the same as in the scanning of the forest scenes. These indoor range maps represent descriptions of the distances of a 330 × 80 field of view from the perspective of a person with an eye-height of 1.65 m in typical indoor scenes like classrooms, halls, corridors, etc.

In order to get a rough approximation of the disparity distribution of indoor scenes, we took advantage of the observation (above) that a simple algorithm that selects the nearest point along a random direction as the fixation point produces a very human-like distribution of fixation distances. Again, we are certainly not claiming that this is what people actually do; we only note that, over many

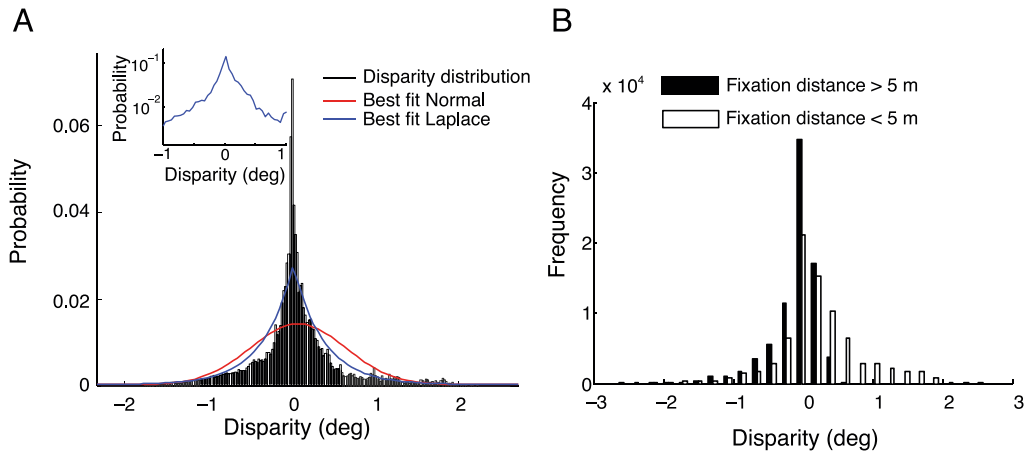

Figure 9. (A) The distribution of disparity in indoor scenes. The inset shows the distribution between ±1 deg using semi-log plot. (B) The shape of distribution in indoor scenes varies with different fixation distances.

fixations, this algorithm matches the distribution of actual egocentric distances in the outdoor environment, as does the distribution of human fixation distances shown above. For the present, then, we will assume that the same is true for indoor environments.[2]

In order to compare with the results we obtained from the forest range maps, we used the identical procedure on the indoor range maps (255 fixations and 90-deg binocular visual field) and derived 255 × 631 disparities from the 27 indoor environments. The resulting distribution is shown in the left panel of Figure 9, and it is qualitatively very similar to both of the estimated outdoor distributions shown above. The distribution spans 5.18 deg. The mean is 0.13 deg, the standard deviation is 0.50, and the kurtosis is 4.07. Roughly 97.3% disparities are beyond a disparity threshold of 10 arcsec. Compared with the forest disparity distribution, we found that the spanned range in indoor scenes is slightly larger. It also contains more far disparities than the forest scenes. Since we have not actually measured human fixation distances in indoor scenes, this comparison is tenuous. But still, the two distributions are alike: they both are non-Gaussian and have high peaks and heavy tails.

As above, we also partitioned the histogram into two sub-histograms according to the fixation distances. The blue one is the histogram when fixation distances are less than 5 m, and the red one is the histogram when fixation distances are larger than 5 m. Five meters is the median of all the 255 "fixation" distances with uniformly distributed orientations from the indoor scenes. Clearly, the results are very similar to those from the outdoor data.

Figure 7 (the blue line with squares) shows the eccentricity analysis for the indoor data. The VM circle calculation fits the indoor data quite well as well (the dashed blue line). As expected, both the standard deviation and range of the indoor data are a bit larger than the outdoor data. Figure 8B shows the distributions plotted separately at eccentricities of ±13, ±9, ±5, and ±1 deg.

Relationship between neurophysiological findings and the natural disparity distribution

The neurophysiological mechanisms of disparity processing in different cortical areas of cats and primates has been under extensive study for the past 40 years (Barlow, Blakemore, & Pettigrew, 1967; Bishop & Pettigrew, 1986; Cumming, 2002; DeAngelis, Ohzawa, & Freeman, 1991, 1995; Ohzawa, DeAngelis, & Freeman, 1990; Ohzawa & Freeman, 1986; Parker & Cumming, 2001; Poggio & Fischer, 1977; Poggio, Gonzalez, & Krause, 1988; Prince, Cumming, & Parker, 2002; Prince, Pointon, Cumming, & Parker, 2002; Tsao, Conway, & Livingstone, 2003). The binocular simple cells are the first stage of disparity processing in the brain. The receptive field of a simple cell can be modeled as a Gabor function, which is a sinusoid multiplied by a Gaussian envelope. Strong evidence suggest that positional shifts between the two Gaussian envelopes, and phase shifts between the two sinusoids of the left and right Gabor functions (receptive fields) of a simple cell are the entities responsible for disparity coding (for a review, see Cumming & DeAngelis, 2001). A quadrature pair of the simple cells composes the inputs to a complex cell, which is widely accepted to be a robust disparity detector.

Poggio and Fischer (1977) classified 119 binocular sensitive neurons at the A17 (foveal striate) and A18 (prestriate) cortex of monkeys into 4 categories: tuned excitatory (TE), tuned inhibitory (TI), near (NE) and FA (Far). The TE cells are activated, and TI cells are inhibited by stimuli close to fixation (disparities within ±0.1 deg). NE cells are activated by stimuli nearer than fixation and suppressed by farther stimuli, while the FA cells have the opposite behavior, when the stimuli's disparities are beyond ±0.1 deg.

Among 119 neurons, 66 are TE cells and 14 are TI cells. Thus, the total number of neurons sensitive to near zero disparity is 80, which composes 67% (80/119) of all

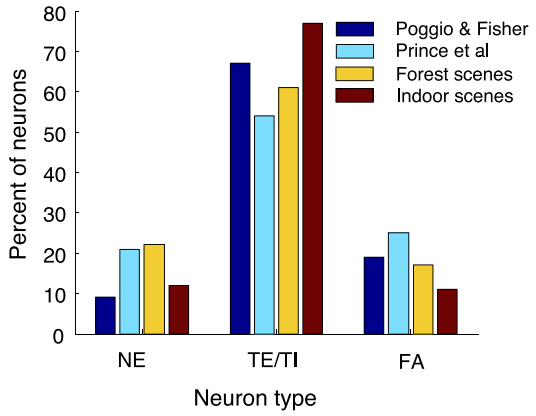

Figure 10. The percentage of binocularly tuned V1 Neurons from Poggio and Fisher (1977), Prince, Cumming, et al. (2002), the percentage of natural disparities within ±5-deg eccentricity from the forest scenes and from the indoor scenes. NE represents the neurons tuned to near disparities smaller than −0.1 deg. TE and TI represent neurons tuned to disparities within ±0.1 deg. FA represent far disparities larger than 0.1 deg. The yellow bars and red bars show the proportions of the disparities in our disparity distributions that can excite the NE, TE and TI, and FA cells, respectively.

cells. There are 22 (19%) FA cells and 11 (9%) NE cells. They also found that all neurons' preferred disparities are limited within ±1 deg. Although the 4 category classification of neurons is rather rough and has been challenged by newer findings in psychophysics and physiology (Cormack, Stevenson, & Schor, 1993; Stevenson, Cormack, Schor, & Tyler, 1992), it still yields a quantitative measurement of the proportions of disparity tuning in V1.

Prince, Cumming, et al. (2002) studied 180 neurons in V1 of macaque monkeys with strong binocular disparity selectivity. Their data support a continuum of disparity tuning instead of 4 discrete categories. However, they still applied the same simple criterion by Poggio and Fischer (1977) to the 180 neurons. Their results were in good agreement with Poggio and Fisher: 69 (38%) are TE neurons, 29 (16%) are TI neurons, 38 (21%) are NE neurons and 44 (25%) are FA neurons.

Now consider the disparity distribution in the 23 forest range maps (Figure 4). Using the 0.1-deg criterion and only considering the disparities within the ±5 deg eccentricity limit found in both results (Poggio & Fischer, 1977; Prince, Cumming, et al., 2002), 17% of the natural disparities are larger than 0.1 deg (FA), 22% are smaller than −0.1 deg (NE), and 61% are within ±0.1 deg (TE/TI). In the 27 indoor range maps, 12% of the disparities are larger than 0.1 deg, 11% are smaller than −0.1 deg, and 77% are within ±0.1 deg.

The general agreement between the natural disparity distribution and the physiological data of V1 are very interesting. As Figures 4 and 9 show, 97.5% of the disparities in forest scenes and 93% of the disparities in indoor scenes are within ±1 deg. This means that most of the natural disparities are within the encoding range of the V1 cells. Figure 10 shows that more than half of natural disparities within ±5 deg eccentricity are between ±0.1 deg, which is consistent with the V1 cells. The forest scenes contain a slightly greater number of near disparities than far disparities. However, the indoor scenes contain almost same number of far disparities and near disparities. Of course, as our previous results show, the distribution of disparities has a dependency on fixation distances. Near fixation distances produces more far (uncrossed, positive) disparities, and far fixation distances produces more near (crossed, negative) disparities. It seems that the disparities in the forest and indoor scenes are roughly symmetric, not the same as V1, which has more far tuned neurons than near tuned neurons. However, it is debatable to conclude that more V1 neurons are tuned to far than to near given only two studies (Poggio & Fisher, 1977; Prince, Cumming, et al., 2002), and it is also debatable that the 23 forest scenes and 27 indoor scenes are the most relevant environments in V1 evolution. Perhaps those elements of the environment that are within reach of the primate arm contribute more to the development of V1 tuning, but we do not know what the disparity distribution in small distances looks like without available range maps.

It is known that the size of receptive fields increases with retinal eccentricity, and the spatial frequency of receptive fields decreases with retinal eccentricity (Wilson & Sherman, 1976). In the well-accepted binocular energy model proposed by Ohzawa and Freeman (1986), a complex cell receives inputs from a quadrature pair of simple cells. The complex cell's disparity tuning curve is decided by the size and spatial frequency of the simple cells (Ohzawa & Freeman, 1986; Qian & Zhu, 1997; Zhu & Qian, 1996). Peripheral neurons with larger receptive fields and lower spatial frequencies can encode disparities at a wider range, but with coarser resolution. As is shown by a recent report on V1 (see Figure 11 in Prince, Cumming, et al., 2002), the disparity tuning curves do become continuously coarser and wider as the eccentricity grows.

Another study on binocular receptive fields and eccentricity (Joshua & Bishop, 1970) on anesthetized cats showed that the standard deviation of receptive field disparity, which is a term similar to the preferred disparity of the neuron, also increases with eccentricity. Table 1 lists the standard deviation of the receptive field disparity from Joshua and Bishop (1970), and the standard deviation of the natural disparity distribution in the forest scenes and the indoor scenes in different eccentricity groups. The qualitative agreement among the two columns is obvious. Again, this clearly shows that the property of disparity sensitivity in V1 may match the distribution of natural disparities.

Recent physiological findings suggest that V1 is not directly responsible for high level depth perception (Cumming & Parker, 2000). DeAngelis and Uka (2003)

Eccentricity	Standard deviation of spread of receptive field disparities	Standard deviation of the natural binocular disparities in the 23 forest scenes	Standard deviation of the natural binocular disparities in the 27 indoor scenes
0–4 deg	0.50 deg (90 cells)	0.23 deg	0.29 deg
4–8 deg	0.76 deg (74 cells)	0.26 deg	0.32 deg
8–12 deg	0.79 deg (39 cells)	0.35 deg	0.40 deg
12–16 deg	0.90 deg (10 cells)	0.41 deg	0.46 deg

Table 1. Changes in horizontal receptive field disparities with horizontal retinal eccentricities, adapted from Joshua and Bishop (1970).

studied the MT (V5) area of awake macaque monkeys. They found that MT neurons are more significantly tuned to horizontal disparities than V1 cells and have a broader range of disparity tuning. Unlike the ±1 deg limit of the V1 cells reported by Poggio and Fischer (1977) and Prince, Cumming, et al. (2002), the range of MT's disparity tuning covers about 6 deg centered at 0. The disparity tuning range of MT, shown in Figure 11A, is more closely matched with natural disparity distributions, which also spans 4–5 deg, although it would appear that MT oversamples large disparities relative to their prevalence in our distributions. This is also shown by the Q–Q plots in Figures 11C and 11D; there is excellent agreement over a roughly 2-deg central range, but the MT data show heavier tails. An interesting property of the MT disparity tuning is the preference of negative (near) disparities. Although the majority of the MT cells are still tuned to smaller disparities near zero, there are noticeably more (61%) neurons tuned to negative (near) disparities than positive (far) disparities, which is just the reverse of the V1 cells, and is also different from the roughly symmetric natural disparity distribution. Figure 11B shows the proportions the disparities in the forest scenes and indoor scenes respectively, compared with the MT

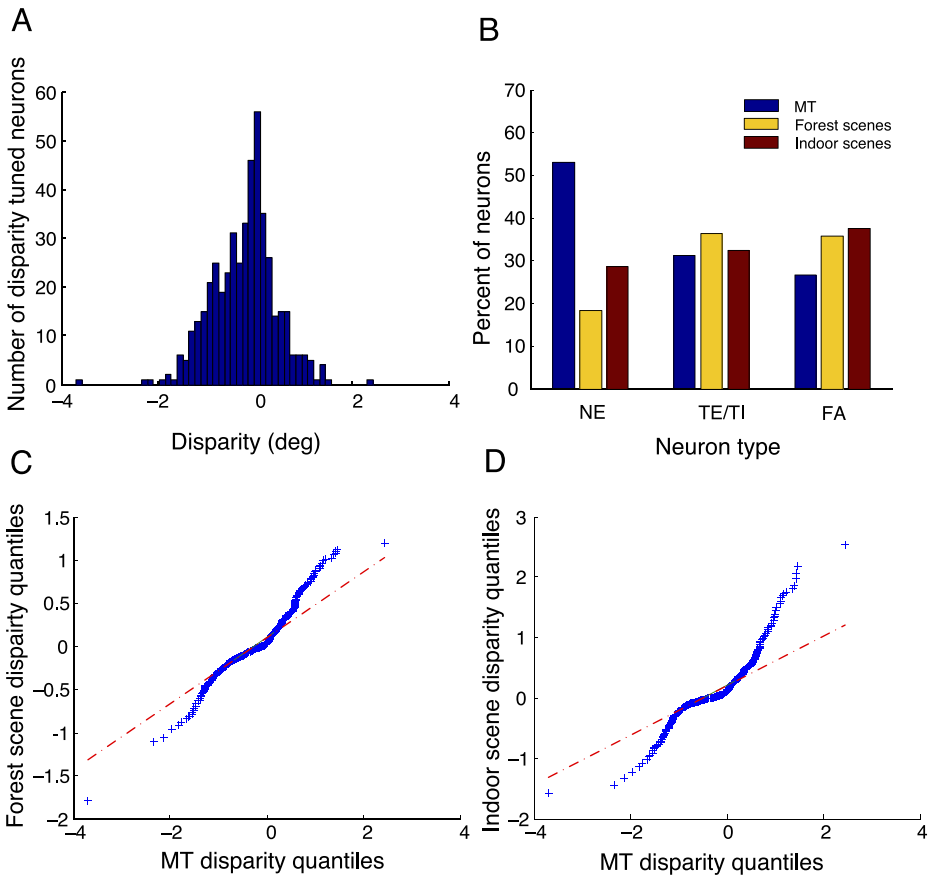

Figure 11. (A) The distribution of the preferred disparities of 471 MT neurons. (B) The disparity percentage of the MT neurons, the forest scenes, and the indoor scenes. We used the same ±0.1-deg criterion here as in Figure 10. You may notice that the shape of natural disparity bar plot is different from that of Figure 10. It is because that we chose the disparities within a 30-deg eccentricity to match the MT neurons' eccentricity range instead of the 5-deg eccentricity range of V1. (C) The Q–Q plot of the disparity distribution in the 23 forest range maps against the preferred disparities of 471 MT neurons. (D) The Q–Q plot of the disparity distribution in the 27 indoor range maps against the preferred disparities of 471 MT neurons.

neurons, based on the same ±0.1-deg criteria but within a 30-deg eccentricity to match the MT neurons' eccentricity range. The MT's deviation from the natural disparity distribution and the V1 neurons is obvious.

What causes the deviation of MT's disparity tuning from V1's? We assume that V1 is specialized for disparity encoding instead of depth perception. If V1 is simply an encoder, then its job should be loyally conveying the disparity information to higher cortical areas for further disparity processing. Thus, the disparity tuning of V1 cells should exactly reflect the disparity distribution in the natural environment, nothing more, nothing less. However, not every disparity is equally important to the well-being of the creature. From an ecological and evolutional point of view, a nearer object is more important to the creature, since it means either a reachable fruit or an approaching danger. It is imaginable that higher cortical areas such as MT weight the input from V1 differently according to their importance.

Relationship between psychophysical findings and the natural disparity distribution

The operational range of stereopsis is determined by two limits:

1. depth discrimination threshold (stereoacuity) and
2. upper disparity limit. Various methods have been used to measure the two limits.

The obtained results are different, given different experiment designs. But it is commonly believed that the depth discrimination is the best at the foveal center and with zero disparity and decreases with growing eccentricity and pedestal disparity. The upper disparity limit increases with eccentricity and stimulus size.

Ogle (1952) studied the limit of stereopsis by moving one rod in depth away from the fixation. There are three stereo perception levels:

1. strong depth perception in the Panum's fusional area (±5 arcmin at the fovea), where objects are fused,
2. strong depth perception with diplopia, and
3. vague depth perception with diplopia.

He called the first two *patent stereopsis,* and the third *qualitative stereopsis.*

Ogle found the range of *patent stereopsis* extends to about ±10 arcmin at the fovea. There is a clear disparity–depth relationship in this area, which makes fine depth judgment possible. Beyond the ±10 arcmin limit, the performance of judging the stimulus to be nearer or further than the fixation is still above chance, but the fine disparity–depth relationship is lost. The range of *qualitative stereopsis* at fovea is about ±15 arcmin. There is no reliable depth perception beyond the *qualitative stereopsis* range. At the periphery, the range of reliable depth perception is much larger. For example, *patent stereopsis* extends to about 70 arcmin and *qualitative stereopsis* about 2 deg at a 6-deg eccentricity.

Westheimer and Tanzman (1956) found that the limit of convergent (uncrossed) disparity is about 6 deg and divergent (uncrossed) disparity about 10 deg. Blakemore (1970) studied the range of stereopsis using a small fixation point and briefly exposed, vertical slit-shaped targets at different eccentricities and pedestal disparities. He found that the upper limit for reliable qualitative localization of a slit as nearer or further than the fixation is 4–7 deg of absolute disparity in a convergent direction and 9–12 deg in a divergent direction near the fovea. Even larger absolute disparities can be recognized as the eccentricity of the slits increases. Landers and Cormack (1997) also found that reliable disparity discrimination is well beyond 1 deg.

The width of the disparity distribution in the 23 forest scenes is about 4.5 deg, which is covered by the range found by Blakemore (1970). Apparently, the distribution of disparities depends on the distance distributions in the environments. Blakemore's experiments had a view distance of only 43.7 cm, which is significantly smaller than the distances in a forest. This could be one reason that the disparity limit he measured is larger than the range of the disparity distribution.

Blakemore (1970) studied the disparity discrimination thresholds at different eccentricities and different pedestal disparities. He found that the best stereoacuity is achieved at the zero eccentricity and at the fixation. There is a loss of stereoacuity as the eccentricity grows, where the disparities are more likely to be large. Our findings show that binocular disparities in the natural world have a larger standard deviation as the eccentricity grows.

Tyler (1973) measured the sensitivity function for the disparity modulation and found that the stereoacuity peaked at a disparity modulation frequency about 1 cpd. For lower disparity modulation frequency, both the lower and higher disparity limits were elevated. Prince and Rogers (1998) measured the sensitivity function for disparity modulation at different eccentricities. They found that the sensitivity decreased with eccentricity, but interestingly, the peak sensitivity to disparity modulation always happened at the same frequency at all eccentricities, with cortical magnification factor taken into account.

In a very recent study, Hibbard (2007) also modeled the statistics of disparity in a similar way as we did, but he used a collage sphere model (spheres randomly distributed at different spatial locations) instead of actual range data to calculate the disparity distribution. He also found the disparity-eccentricity dependency and the Laplace-like shape of the disparity distribution, which are very similar to our results. Our results thus indicate that Hibbard's model produces an accurate statistical description of disparity information for at least two common environments.

Conclusion

By computing disparities from actual range data, we found that the distribution of binocular disparities in forest scenes and indoor scenes from Yang and Purves's (2003b) range map database is highly peaked at 0 deg and spans several degrees. The range of the disparity distribution is fully covered by the range of macaque MT cells' (DeAngelis & Uka, 2003) disparity tuning. The proportion of the distribution is qualitatively consistent with the proportion of the disparity tuned V1 neurons (Figure 10). We are fully aware that our distribution is not the canonical measurement of the disparities across all tasks and environments. But it does seem to be a robust estimate of the disparities that an observer is likely to encounter in both large scale (outdoor) and small scale (indoor) environments when gaze is roughly horizontal.

Simoncelli and Olshausen (2001) pointed out that the evolution and development of a biological visual system is driven by three fundamental factors: (1) the tasks that the visual system must perform, (2) the computational capabilities and limitations of the neurons, and (3) the living environment of the organism. In this paper, we tried to correlate the disparity distribution in the natural environments with other research areas in stereopsis in an attempt to integrate these points. Examples of this integrated perspective are biologically compatible stereopsis models (Read, Parker, & Cumming, 2002; Tsai & Victor, 2003). These models tried to solve depth perception with current computational models (DeAngelis et al., 1991; Ohzawa et al., 1990; Qian & Zhu, 1997) of binocular neurons. Read et al. (2002) assumed a non-Gaussian, zero-favoring distribution of the disparity as a Bayesian prior. Our disparity distribution supports her assumption quite well and can also serve as a more accurate priori probability in a Bayesian framework of stereo correspondence.

Finally, we point out that the most up-to-date laser scanning technology offers co-registered distance and luminance maps (Potetz & Lee, 2003). It is even more interesting to look at the joint distribution of luminance features and disparity features in the natural world if this technology is combined with a fixation selection scheme, which will greatly enhance our understanding of the stereo correspondence and disparity processing.

Acknowledgments

This research was supported by NSF Grant ITR-0427372. We thank Zhiyong Yang and Dale Purves for sharing their range maps. We also thank Kalpana Seshadrinathan for her useful comments and proofreading.

Commercial relationships: none.
Corresponding author: Yang Liu.
Email: young76@mail.utexas.edu.
Address: Department of Electrical and Computer Engineering, The University of Texas at Austin, TX 78712.

Footnotes

[1]If anything, the opposite is probably true when navigating, considering the nature of paths, roads, and hallways.

[2]Of course people deploy their fixations differently (in some sense) for different tasks in different environments, and this may influence the shape of the distribution of fixation distances relative to the distribution of environmental distances, but until we can simultaneously measure ranges and fixations in the same environment and co-register the data, we will proceed under the above assumption.

References

Barlow, H. B., Blakemore, C., & Pettigrew, J. D. (1967). The neural mechanism of binocular depth discrimination. *The Journal of Physiology, 193,* 327–342. [PubMed] [Article]

Bishop, P. O., & Pettigrew, J. D. (1986). Neural mechanisms of binocular vision. *Vision Research, 26,* 1587–1600. [PubMed]

Blakemore, C. (1970). The range and scope of binocular depth discrimination in man. *The Journal of Physiology, 211,* 599–622. [PubMed] [Article]

Cormack, L. K., Stevenson, S. B., & Schor, C. M. (1993). Disparity-tuned channels of the human visual system. *Visual Neuroscience, 10,* 585–596. [PubMed]

Cumming, B. G. (2002). An unexpected specialization for horizontal disparity in primate primary visual cortex. *Nature, 418,* 633–636. [PubMed]

Cumming, B. G., & DeAngelis, G. C. (2001). The physiology of stereopsis. *Annual Review of Neuroscience, 24,* 203–238. [PubMed]

Cumming, B. G., & Parker, A. J. (2000). Local disparity not perceived depth is signaled by binocular neurons in cortical area V1 of the Macaque. *Journal of Neuroscience, 20,* 4758–4767. [PubMed] [Article]

DeAngelis, G. C., Ohzawa, I., & Freeman, R. D. (1991). Depth is encoded in the visual cortex by a specialized receptive field structure. *Nature, 352,* 156–159. [PubMed]

DeAngelis, G. C., Ohzawa, I., & Freeman, R. D. (1995). Neuronal mechanisms underlying stereopsis: How do

simple cells in the visual cortex encode binocular disparity? *Perception, 24,* 3–31. [PubMed]

DeAngelis, G. C., & Uka, T. (2003). Coding of horizontal disparity and velocity by MT neurons in the alert macaque. *Journal of Neurophysiology, 89,* 1094–1111. [PubMed] [Article]

Hayhoe, M. M., Shrivastava, A., Mruczek, R., & Pelz, J. B. (2003). Visual memory and motor planning in a natural task. *Journal of Vision, 3*(1):6, 49–63, http://journalofvision.org/3/1/6/, doi:10.1167/3.1.6. [PubMed] [Article]

Hibbard, P. B. (2007). A statistical model of binocular disparity. *Visual Cognition, 15,* 149–165.

Howard, I. P., & Rogers, B. J. (1995). *Binocular vision and stereopsis*. New York: Oxford University Press.

Huang, J., Lee, A. B., & Mumford, D. (2000). *Statistics of range images*. Paper presented at CVPR, 2000.

Joshua, D. E., & Bishop, P. O. (1970). Binocular single vision and depth discrimination. Receptive field disparities for central and peripheral vision and binocular interaction on peripheral single units in cat striate cortex. *Experimental Brain Research, 10,* 389–416. [PubMed]

Land, M., Mennie, N., & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception, 28,* 1311–1328. [PubMed]

Landers, D. D., & Cormack, L. K. (1997). Asymmetries and errors in perception of depth from disparity suggest a multicomponent model of disparity processing. *Perception & Psychophysics, 59,* 219–231. [PubMed]

Le Grand, Y. (1980). *Physiological optics*. New York: Springer-Verlag.

Ogle, K. N. (1952). On the limits of stereoscopic vision. *Journal of Experimental Psychology, 44,* 253–259. [PubMed]

Ohzawa, I., DeAngelis, G. C., & Freeman, R. D. (1990). Stereoscopic depth discrimination in the visual cortex: Neurons ideally suited as disparity detectors. *Science, 249,* 1037–1041. [PubMed]

Ohzawa, I., & Freeman, R. D. (1986). The binocular organization of complex cells in the cat's visual cortex. *Journal of Neurophysiology, 56,* 243–259. [PubMed]

Palmer, S. E. (1999). *Vision science—photons to phenomenology*. Cambridge, MA: MIT Press.

Parker, A. J., & Cumming, B. G. (2001). Cortical mechanisms of binocular stereoscopic vision. *Progress in Brain Research, 134,* 205–216. [PubMed]

Poggio, G. F., & Fischer, B. (1977). Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. *Journal of Neurophysiology, 40,* 1392–1405. [PubMed]

Poggio, G. F., Gonzalez, F., & Krause, F. (1988). Stereoscopic mechanisms in monkey visual cortex: Binocular correlation and disparity selectivity. *Journal of Neuroscience, 8,* 4531–4550. [PubMed] [Article]

Potetz, B., & Lee, T. S. (2003). Statistical correlations between two-dimensional images and three-dimensional structures in natural scenes. *Journal of the Optical Society of America A, Optics, Image Science, and Vision, 20,* 1292–1303. [PubMed]

Prince, S. J., Cumming, B. G., & Parker, A. J. (2002). Range and mechanism of encoding of horizontal disparity in macaque V1. *Journal of Neurophysiology, 87,* 209–221. [PubMed] [Article]

Prince, S. J., Pointon, A. D., Cumming, B. G., & Parker, A. J. (2002). Quantitative analysis of the responses of V1 neurons to horizontal disparity in dynamic random-dot stereograms. *Journal of Neurophysiology, 87,* 191–208. [PubMed] [Article]

Prince, S. J., & Rogers, B. J. (1998). Sensitivity to disparity corrugations in peripheral vision. *Vision Research, 38,* 2533–2537. [PubMed]

Qian, N., & Zhu, Y. (1997). Physiological computation of binocular disparity. *Vision Research, 37,* 1811–1827. [PubMed]

Read, J. C., Parker, A. J., & Cumming, B. G. (2002). A simple model accounts for the response of disparity-tuned V1 neurons to anticorrelated images. *Visual Neuroscience, 19,* 735–753. [PubMed]

Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience, 24,* 1193–1216. [PubMed]

Stevenson, S. B., Cormack, L. K., Schor, C. M., & Tyler, C. W. (1992). Disparity tuning in mechanisms of human stereopsis. *Vision Research, 32,* 1685–1694. [PubMed]

Tsai, J. J., & Victor, J. D. (2003). Reading a population code: A multi-scale neural model for representing binocular disparity. *Vision Research, 43,* 445–466. [PubMed]

Tsao, D. Y., Conway, B. R., & Livingstone, M. S. (2003). Receptive fields of disparity-tuned simple cells in macaque V1. *Neuron, 38,* 103–114. [PubMed] [Article]

Tyler, C. W. (1973). Stereoscopic vision: Cortical limitations and a disparity scaling effect. *Science, 181,* 276–278. [PubMed]

Tyler, C. W. (1991). Cyclopean vision. In D. Regan (Ed.), *Vision and visual dysfunction. Binocular vision* (vol. 9, pp. 38–74). London: Macmillan.

von Helmholtz, H. (1962). *Treatise on physiological optics* (vol. III). New York: Dover.

Westheimer, G., & Tanzman, I. J. (1956). Qualitative depth localization with diplopic images. *Journal of the Optical Society of America, 46,* 116–117. [PubMed]

Wilson, J. R., & Sherman, S. M. (1976). Receptive-field characteristics of neurons in cat striate cortex: Changes with visual field eccentricity. *Journal of Neurophysiology, 39,* 512–533. [PubMed]

Yang, Z., & Purves, D. (2003a). A statistical explanation of visual space. *Nature Neuroscience, 6,* 632–640. [PubMed]

Yang, Z., & Purves, D. (2003b). Image/source statistics of surfaces in natural scenes. *Network, 14,* 371–390. [PubMed]

Zhu, Y. D., & Qian, N. (1996). Binocular receptive field models, disparity tuning, and characteristic disparity. *Neural Computation, 8,* 1611–1641. [PubMed]