

Luminance, disparity and range statistics in 3D natural scenes

Yang Liu^{*a}, Lawrence K. Cormack^b, Alan C. Bovik^a

^aLIVE, Dept. of ECE, Univ. of Texas at Austin, Austin, TX

^bCenter of Perceptual Systems, Dept. of Psychology, University of Texas at Austin,
Austin TX

ABSTRACT

Range maps have been actively studied in the last few years in the context of depth perception in natural scenes. With the availability of co-registered luminance information, we have the ability to examine and model the statistical relationships between luminance, range and disparity. In this study, we find that a one-sided generalized gaussian distribution closely fits the prior of the range gradient. This finding sheds new light on statistical modeling of 2D and 3D image features in natural scenes.

Keywords: depth perception, disparity, natural scene statistics

1. INTRODUCTION

Biological vision systems have evolved by obtaining timely environmental information that is relevant to the organism's survival and reproduction. The statistics of light intensities projected onto the retina are the driving force behind the evolution of vision systems. Studies on natural scene statistics have proved successful in helping understand and predict the functions of cortical cells, and have inspired new image processing methods for engineering purposes.

Most existing studies have focused on the statistics of luminance in natural scenes. However, in the last several years a few researchers have used laser range scanners to obtain 3D range maps of natural scenes. Statistical observations derived from such range maps successfully explains various phenomena in 3D visual perception, such as terrain effects [1], orientation effects [2], and shape from shading [3].

Using the range maps in [1], and assuming a 6.5 cm interocular distance and an empirically measured fixation distance distribution, we previously found that the disparity distribution of those natural scenes covers a range of about 5 to 6 degrees [4], which falls within in the ensemble encoding range of binocular neurons in both V1 and MT+, but the shape of the distribution is more similar to the ensemble tuning of V1. The parameters of the disparity distribution depend on environment, eccentricity and, particularly, elevation.

Although disparity statistics directly derived from distances are important for understanding depth perception, the neurobiological calculation of disparity is actually based on two spatially displaced retinal luminance images, which are not available in previous range datasets [1]. Thus, characterizing the two retinal luminance images is an indispensable step towards understanding the overall process. Using a laser scanner that records luminance and range simultaneously, the luminance, distance, and disparity that are projected onto the retinae from each visual direction can be modeled, yielding the complete chain of information for depth perception. Potetz and Lee's co-registered luminance and range maps are the first dataset available for this purpose [3].

We organize this paper in the following way: In section 2 we present results on luminance, range and disparity at the coordinates of edges detected by a filter bank of LoG (Laplacian of Gaussian) filters. The positive correlation between the mean luminance gradient and mean range gradient is one important observation. Section 3 illustrates the statistics of luminance, range and disparity at the edges of different thresholds using a more localized gradient-based (Roberts) edge operator. We also develop probabilistic models to fit the data. In particular, a one-sided generalized gaussian distribution to the range gradient

samples at luminance edges is found to fit the data well. The parameters of the generalized gaussian distribution are estimated by the moments of the distribution. Some open issues are discussed in Section 4.

2. STATISTICS AT EDGES DETECTED BY LOG FILTERS

2.1 The LoG filters

The LoG (Laplacian of Gaussian) filter is a commonly used edge detector. Assuming $G_\sigma(x, y)$ is a 2D Gaussian function centered at the origin with the standard deviation σ . It takes the form of

$$LoG \square \frac{\partial^2}{\partial x^2} G_\sigma(x, y) + \frac{\partial^2}{\partial y^2} G_\sigma(x, y) = \frac{x^2 + y^2 - 2\sigma^2}{\sigma^4} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)$$

One reason to use the LoG filter is its simplicity and broad usage in image processing, and computer vision. Another reason is that it is widely used to simulate LGN cells, which represents an initial processing of retinal image before entering V1.

2.2 Edges detection using a LoG filter bank

We converted the 24-bit color images into 8-bit grayscale intensity images. The laser scanner does not obtain valid readings on smooth regions such as sky, water, glass, etc. We selected 23 pairs of co-registered range maps and intensity images of good visual quality, containing buildings, vegetation, and sculptures.

One major problem of the dataset is the sky pixels. There are no valid distance readings at sky pixels in the range maps. But the luminance information for the sky is valid. In order to convolve the LoG filters with the images, we assigned a distance of 1000m to all sky pixels.

We denote a range map as $r(x, y)$, where x, y represent the azimuth and elevation of pixels. We then compute the angular subtend of each pixel, and obtain the angular map $\alpha(x, y)$. If an imaginary human observer stood at the exact location of the scanner, and fixated at a certain distance z , the vergent angle δ , which is the angular subtend of the two visual axes, can be decided by z and the interocular distance (usually 6.5cm for human). Usually, the angular difference between $\alpha(x, y)$ and δ is named disparity. We define the disparity map $d(x, y) = \alpha(x, y) - \delta$.

In our analysis, we focus on the magnitude of disparity gradient $\nabla d(x, y) = \left(\frac{\partial d}{\partial x}, \frac{\partial d}{\partial y}\right)$, which is close to the

angular difference between two neighboring pixels. If the angular resolution is 0.05 deg, then the disparity gradient at a pixel represents the relative disparity between two locations with a separation of 0.05 deg.

We designed a Laplacian of Gaussian (LoG) filterbank with center frequencies roughly matching human contrast sensitivity: 3, 4, 5, and 6 cycles/deg, with half-power bandwidths of 1.25 octaves, roughly matching V1 neurons. The LoG filters were used to decompose the grayscale intensity images, the co-registered range maps, and the derived disparity maps.

In the LoG-filtered luminance and range images, a pixel was classified as an edge pixel if it fell between adjacent pixels with different signs of LoG response, and if the difference between the pixel and its neighbors was > 0.75 times the mean of the absolute value of the filtered image. For disparity edge detection, an additional stereoacuity threshold of 10 arcsec was also used: if the relative disparity between the pixel and its neighbors > 10 arcsec, then it is a detectable edge pixel [5]. Psychophysical studies have reported 2-6 arcsec stereoacuity under good viewing condition, and have also showed that the disparity detection threshold changes with spatial frequency [6].

Fig. 1 shows the detected luminance, range, and disparity edges by LoG filters with 4 (middle) and 6 (bottom) cycles/deg center frequency.

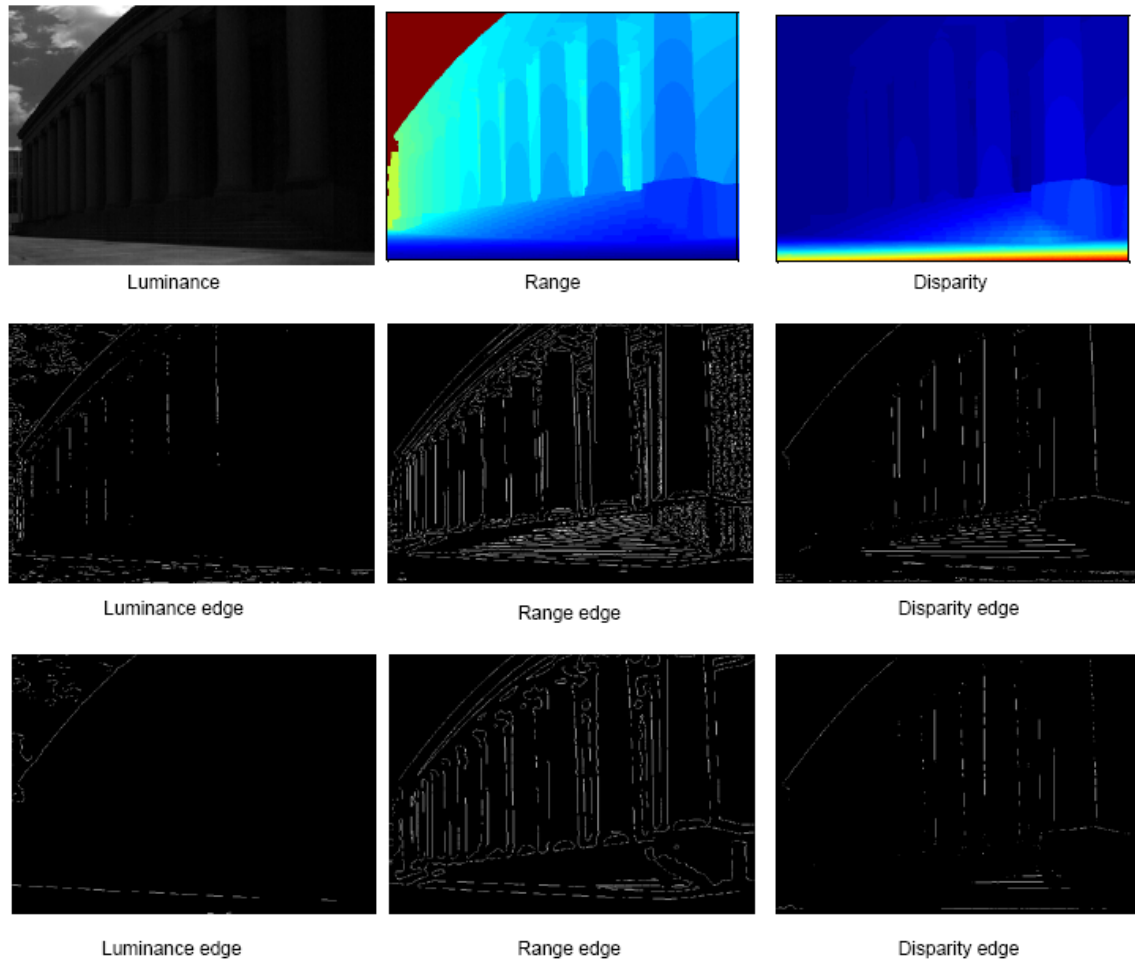


Fig. 1. The first row displays luminance, range and disparity maps of a scene. The luminance, range and disparity edges detected by LoG filters with different center frequencies are shown in the middle panel (6 cycles/deg) and the bottom panel (4 cycles/deg).

2.3 Statistics at Edges detected by LoG filters

The luminance variations of natural scenes can arise from many possible sources. They may be caused by illumination changes without geometrical changes, such as illumination from a TV screen. They may be caused by changes in surface reflectance, such as a painting, or they may be caused by abrupt geometrical changes, such as the boundaries of a building. Some luminance changes are associated with depth change as in the latter example, but some are not as in the former two examples. In this section we show that there exists statistical relationships between luminance gradient and range gradient.

2.3.1 At luminance edges

We binned the pixels at luminance edges according to the magnitudes of the luminance gradients, with each bin containing about 20,000 pixels.

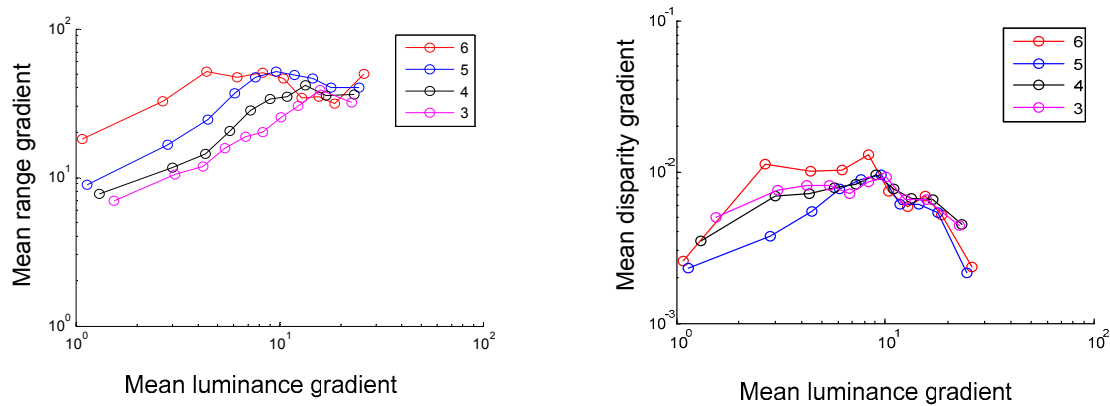


Fig. 2. In the left panel, the mean range gradient and the mean luminance gradient has a nearly monotonic shape at the beginning. The curves saturate for large mean luminance gradient (>10). The right panel shows mean disparity gradient against mean luminance gradient. The relationship between luminance gradient and disparity gradient is not monotonic. Moreover, the values of mean disparity gradient are small (<0.6 arcmin).

In the left panel of Fig. 2, we plot mean range gradients against mean luminance gradients in ten bins. Different colors represent edges detected by LoG filters of different center frequencies. The mean range gradient increases with mean luminance gradient at the beginning, and becomes saturated when the mean luminance gradient is larger than 10. All the curves have similar shape, which indicates that there is a coherent relationship across scales. The right panel of Fig. 2 shows the relationship between mean disparity gradient and mean luminance gradient. Most of the mean disparity gradients are smaller than 0.01 deg in the plots. The curves are not monotonic.

2.3.2 At range edges

Similarly, we binned the pixels at range edges according to the magnitudes of range gradients. The relationship between the mean luminance gradient and the mean range gradient is shown in the left panel of Fig. 3. Similar to what is observed in Fig. 2, the plots of luminance gradient and range gradient are again generally monotonic. Instead of being saturated at larger luminance gradients, the curves increase strongly at large range gradients. It is possibly true, that in natural scenes, a large range gradient may increase the likelihood of a large luminance gradient at the same coordinate. However, the converse is less intuitive, since large luminance gradients occur for many reasons.

The right panel of Fig. 3 shows the relationship between disparity gradient and range gradient. It is interesting to note that these plots exactly illustrate one important property of disparity: the disparity is a non-linear transformation of depth difference. Large depth differences do not necessarily predict large disparity magnitudes. To perceive absolute distance correctly, disparity is not enough.

2.3.3 At disparity edges

A noticeable fact about disparity edges is that they are strongly dependent on distance and elevation. As we can see in the edge maps, the majority of disparity edges are located in the lower halves of the images, which generally have nearer distances too. Relative disparity strongly depends on fixation distance. Two points having a 20 cm depth separation will have a significantly larger disparity when the viewer is 1 m away than 10 m away.

Again, we binned the pixels at disparity edges according to disparity gradient value. Fig. 4 shows the relationship between range gradient and disparity gradient, and the relationship between luminance gradient and disparity gradient, respectively, in the left and right panels.

Fig. 4 shows that a strong disparity edge does not necessarily suggest the presence of a strong luminance edge. Further, the right panel shows that a strong disparity edge is not necessarily associated with a strong range edge either. Unlike the observed positive correlation between luminance and range, the joint distribution of disparity and luminance in natural scenes may be more complex.

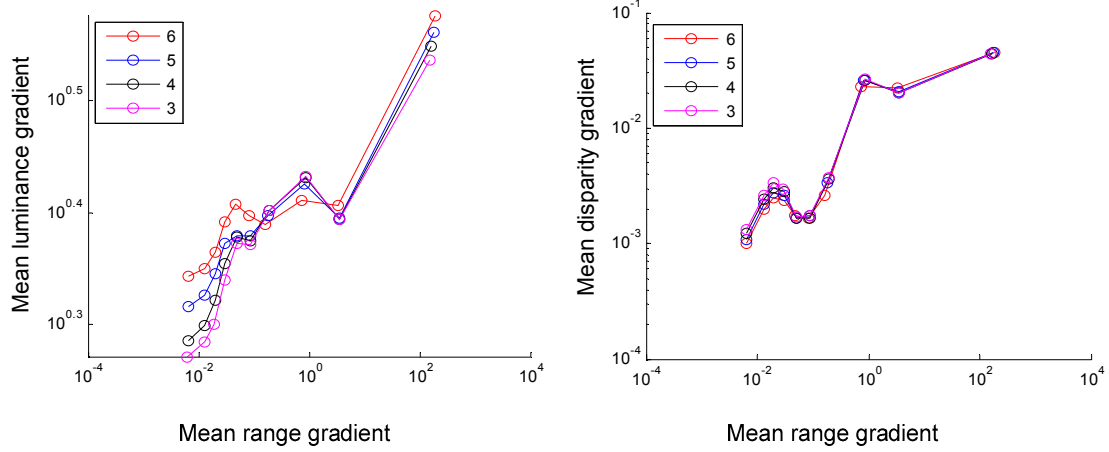


Fig. 3. In the left panel, the mean luminance gradient has a roughly monotonic relationship with the mean luminance gradient. The right panel shows mean disparity gradient against mean range gradient. The relationship between range gradient and disparity gradient is more complex.

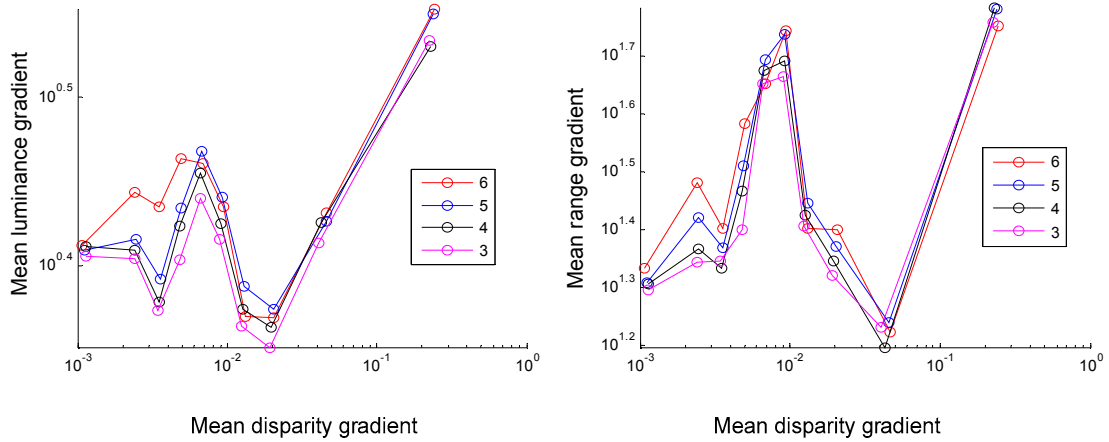


Fig. 4. The left panel shows mean luminance gradient against mean disparity gradient at disparity edges. The right panel shows mean range gradient against mean disparity gradient at disparity edges.

Several existing problems with the dataset should be addressed: First of all, we assigned an arbitrary value (1000 m) to all sky pixels and other invalid pixels (water, glass, etc.), this may affect the curves for larger range gradients in Fig. 2 and Fig. 3. Another problem is that the LoG filter has limited resolution in edge localization, especially when its center frequency is low. To more accurately probe the luminance, range and disparity gradients at a higher resolution, we developed another method as described in the following.

3 STATISTICS AT EDGES DETECTED BY ROBERTS EDGE DETECTOR

First of all, as Potetz and Lee did in [3], we detected, and removed the invalid regions. Then, we applied a highly localized edge detection method: the Roberts edge detector estimates the gradient magnitude using directional templates oriented at 45 deg and 135 deg. Those locations whose gradients exceed a certain threshold are recognized as edges. We chose the threshold as a percentage of the maximum value of each gradient magnitude image.

We manipulate the threshold and obtain associated edges. When we set the threshold to be 0.9, only very strong edges remained. Fig. 5 shows the edges for different types of images.

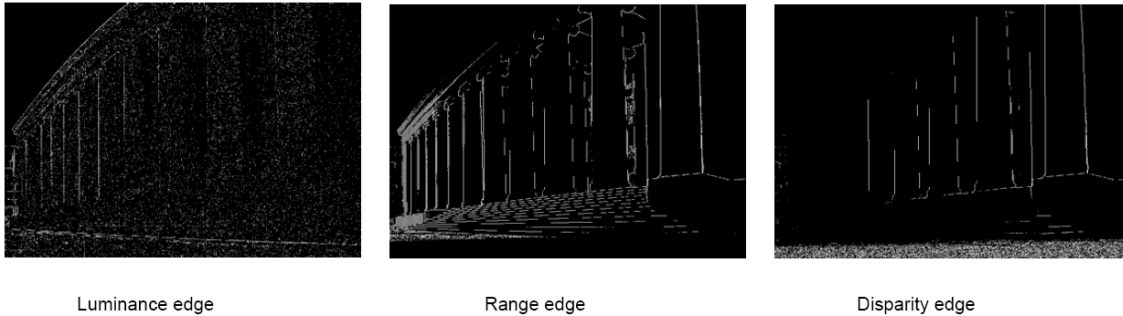


Fig. 5. The luminance, range and disparity edges detected by a Roberts operator using threshold 0.9. The sky pixels have been removed before edge detection.

One observation that can be made by looking at Fig. 5 is the spatial distribution of disparity edges. In this particular scene, there is not much visible luminance and range variation on the ground, but that the predominant disparity edges are located in the lower half of the image, where the distances are nearer too.

3.1 Luminance and range gradients at luminance edges

Our aim is to find statistical models that fit the data. A recurring property of natural scene statistics is the non-gaussian shape of the marginal distributions of image features. For example, we plotted the histograms of the range gradient at luminance edges with different strengths in Fig. 6.

We set thresholds to be 0.9, 0.6, 0.5 and 0.1. Those pixels whose gradient magnitude exceeds 0.9 times the maximum value of the filtered image are deemed to be strong luminance edges. The pixels whose magnitudes are between 0.5 and 0.6 times the maximum gradient magnitude are called weak luminance edges. All the pixels below 0.1 times the maximum gradient magnitude are non-edges.

We obtained 1,041,726 pixels at strong luminance edges, 1,066,268 pixels at weak luminance edges, and randomly picked 1,041,726 pixels (to match the pixel number of strong edges) from regions without luminance edges (regions without luminance edges contained much more pixels than strong or weak edges).

The maximum range gradient is 226, and the minimum range gradient is 0. We then set up 200 bins between zero and 226, with each bin width 2.26, and plotted the normalized frequency in each bin for strong luminance edges, weak luminance edges, and non-edges in Fig. 6. They share a similar high-peaked and heavy-tailed shape, which have been observed in many image properties of natural scenes, but the difference among them is phenomenal. Peaks near zero grow lower and tails grow heavier as the luminance edges become stronger, which indicates that larger range gradients occur more frequently with larger luminance gradients.

To explore the relationship between luminance gradient and range gradient, we binned the strong luminance pixels into 20 bins (about 50,000 in each bin), and showed the mean range gradient against the mean luminance gradient in each bin in Fig. 7. The result is very intriguing and provocative: it is not only monotonic, it is almost linear.

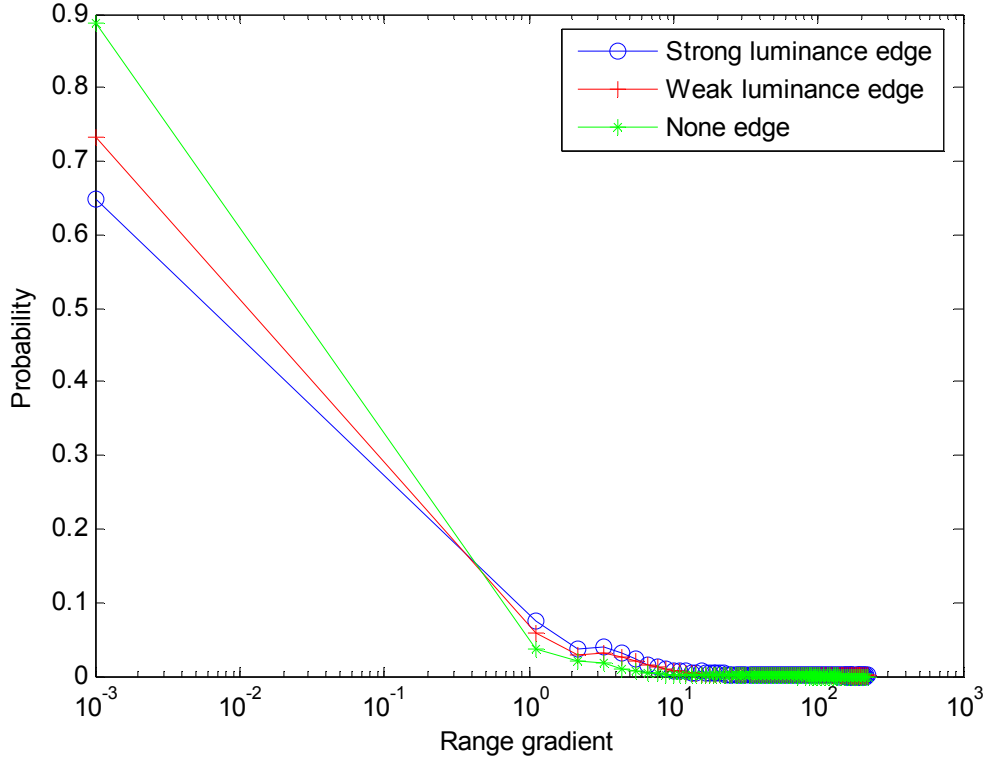


Fig. 6. The distributions of range gradient at luminance edges of different strengths. The blue curve shows the histogram of range gradient at strong luminance edges (threshold 0.9), the red curve shows the histogram at weak luminance edges (between threshold 0.6 and 0.5), and the green one shows the histogram at regions without luminance edges (threshold <0.1).

3.2 Modeling of range gradient at luminance edges

We focused on modeling the conditional distribution of range gradient. As we observed in Fig. 2, 3, 6 and 7, there are undoubtedly positive correlations between the luminance gradient and range gradient at luminance edges. A simple, parameterized probability model of the posterior distribution of range gradient given luminance edges may prove very useful in studies on depth perception and computational stereopsis. There may be functional relationships between the parameters of the distributions of luminance and range gradient. Furthermore, these relationships may be embedded in the neural circuits of the brain during the evolution.

The observed non-gaussian, peaked, and apparently heavy-tailed shape of the histograms of the conditional data suggests that we use a heavy-tailed distribution model, like those used in robust statistics and nonlinear filter theory [7]. In studies of bandpass natural scenes [8], others have found that the marginal distribution of wavelet coefficients are well described by the generalized gaussian distribution:

$$f(x) = A(\sigma, \theta)e^{-(x/\sigma)^\theta}$$

$$A(\sigma, \theta) = \frac{\theta}{2\sigma\Gamma(\frac{1}{\theta})},$$

where $A(\sigma, \theta)$ is a normalization constant, and σ and θ are controlling the dispersion and shape, respectively. When $\theta = 2$, the distribution is gaussian, while $\theta = 1$ corresponds to the laplace distribution. In general, reducing θ makes the distribution more peaked near zero and with heavier tails.

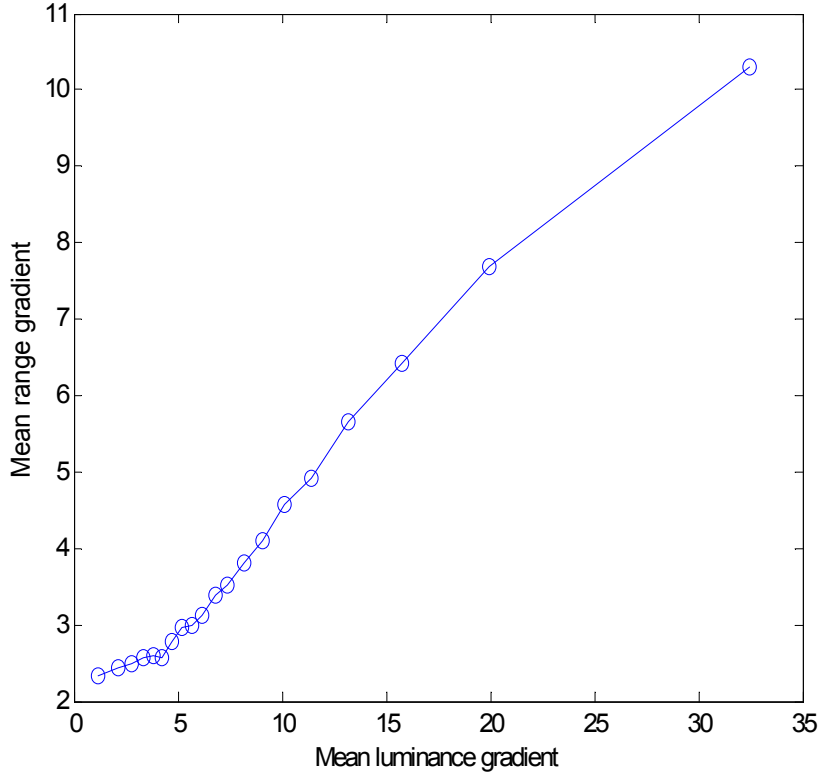


Fig. 7. We binned the luminance edge pixels (threshold=0.9) according to their luminance gradient magnitudes, and plotted the mean range gradient against the mean luminance gradient in each bin. The relationship is almost linear.

One significant difference between our conditional range gradient samples and the wavelet coefficient of natural scenes is that range gradient magnitudes are non-negative. Thus the distribution is one-sided, rather than symmetrical. We thus define a one-sided generalized gaussian:

$$f(x) = \begin{cases} A(\sigma, \theta)e^{-(x/\sigma)^\theta}, & x \geq 0 \\ 0, & x < 0 \end{cases}, A(\sigma, \theta) = \frac{\theta}{\sigma\Gamma(\frac{1}{\theta})}$$

We suggest that the posterior probability distribution of conditional range gradients can be well-modeled using a one-sided generalized gaussian distribution. The parameters of the generalized gaussian can be estimated in a variety of ways. In [9], an equation for the mean and standard deviation of the distribution

was used to derive the parameters. We adopted a similar approach, and estimated the parameters using the sample mean and sample variance of range gradients (See Appendix).

We estimated σ and θ for three different edge conditions: strong edges, weak edges, and non-edges, and plotted the probability function together with the histograms in Figs 8(a), (b), and (c). To compare with Fig. 6, we re-plotted the three probability functions together in Fig. 8(d). Generally, the estimated probability functions fit the data well, and the peaks and tails show similar relationship as the data in Fig. 8. The parameter θ lies in the range $[0.179, 0.253]$. This is lower than the typical range $[0.4, 0.8]$ of θ of the wavelet coefficients of 2D natural scenes [10], suggesting that the range data is more heavy-tailed.

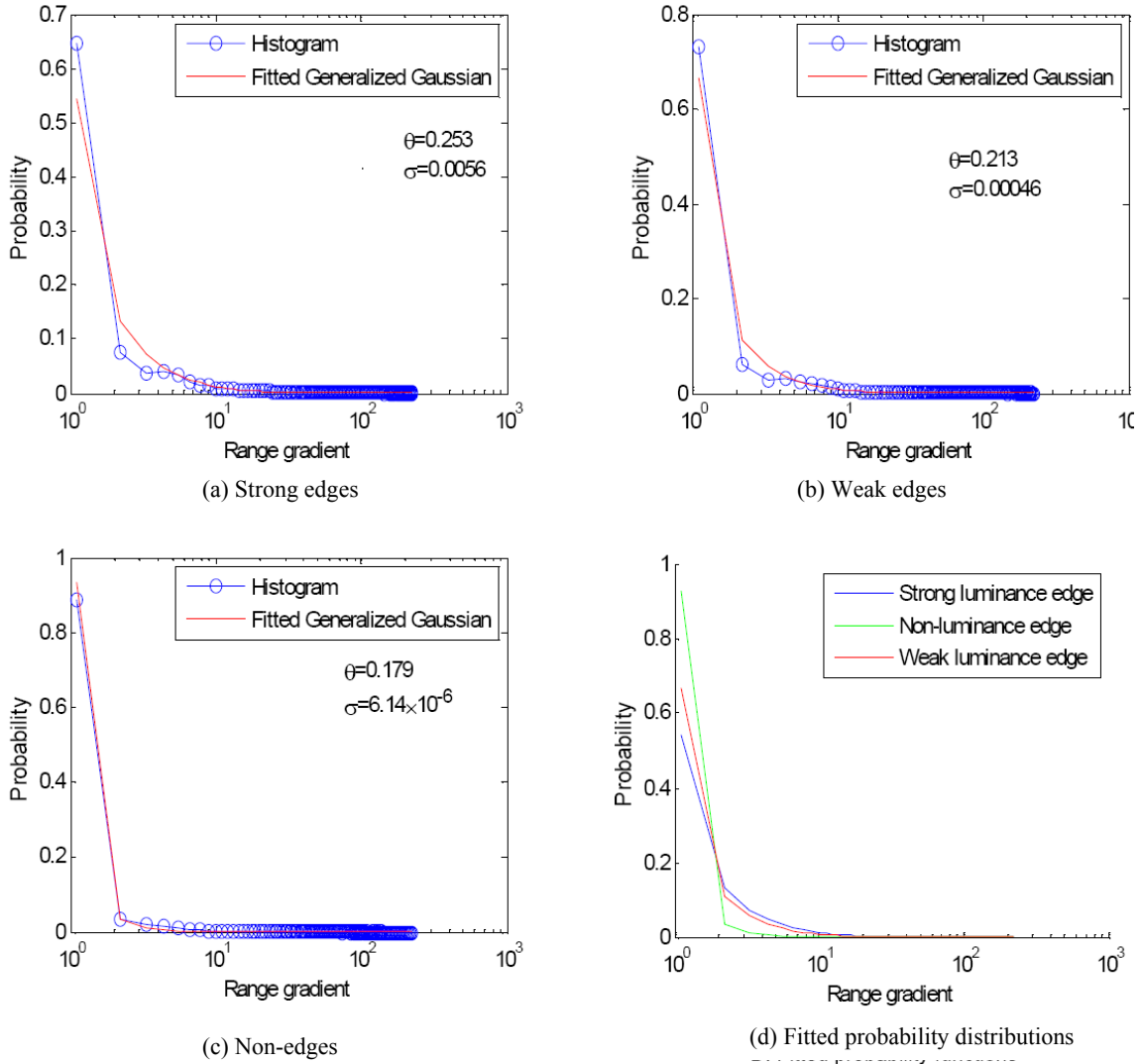


Fig. 8. (a), (b), and (c) show histograms of range gradient and the estimated one-sided generalized gaussian distribution under different luminance edge conditions. (d) shows the overlapped plots of the estimated probability functions.

4. CONCLUSION

In this study, we analyzed the statistical relationships among luminance, range and disparity in natural scenes. We found that luminance gradient and range gradient at luminance edges exhibit a strong relationship. Furthermore, the distribution of range gradient is nicely modeled by a one-sided generalized gaussian distribution. The parameters of the pdfs can be estimated from range samples.

This is the first effort we are aware of to quantitatively analyze and model the relationship between luminance gradients and range gradients, and the distributions conditioned on supra-threshold gradients. Whether or not the human vision system uses this statistical relationship remains an open question. Newer psychophysical experiments may be designed to test it. The models we have developed and are studying may prove useful for computational stereopsis, where new algorithms may be developed under a Bayesian framework, given successful modeling of the likelihood and prior distributions.

There are some inherent limitations of the dataset we used. According to [3], the scanning resolution is about 0.05 deg/pixel. We already know that the human vision system has a much wider bandwidth than the dataset's sampling frequency allows. Another disadvantage is the quality of the luminance images is not always satisfying. Many images lack contrast details, but contain notable high frequency noise. However, we were reluctant to do any preprocessing, since it may introduce unknown influences into the data. With technical advances, newer laser scanners can provide much higher scanning resolution (0.001deg/pixel). We believe the improved resolution will significantly help our understanding of depth perception since it matches the bandwidth of human vision system better.

APPENDIX

The probability density function of one-sided generalized Gaussian distribution is defined as:

$$f(x) = \begin{cases} \frac{1}{\sigma\Gamma(\frac{1}{\theta}+1)} e^{-(x/\sigma)^\theta}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

The mean of the distribution can be computed by:

$$\begin{aligned} E(x) &= \int_0^\infty xf(x)dx = \int_0^\infty \frac{x}{\sigma\Gamma(\frac{1}{\theta}+1)} e^{-(x/\sigma)^\theta} dx \\ &= \frac{\theta}{\sigma\Gamma(\frac{1}{\theta})} \int_0^\infty xe^{-(x/\sigma)^\theta} dx \\ &= \frac{\sigma\Gamma(\frac{2}{\theta})}{\Gamma(\frac{1}{\theta})} \end{aligned}$$

The second order moment of the distribution can be computed by:

$$\begin{aligned}
E(x^2) &= \int_0^{\infty} x^2 f(x) dx = \int_0^{\infty} \frac{x^2}{\sigma \Gamma(\frac{1}{\theta} + 1)} e^{-(x/\sigma)^\theta} dx \\
&= \frac{\theta}{\sigma \Gamma(\frac{1}{\theta})} \int_0^{\infty} x^2 e^{-(x/\sigma)^\theta} dx \\
&= \frac{\sigma^2 \Gamma(\frac{3}{\theta})}{\Gamma(\frac{1}{\theta})}
\end{aligned}$$

We have the following relationship by taking the ratio between $E(x^2)$ and $E(x)$:

$$\frac{E(x^2)}{E(x)} = \frac{\Gamma(\frac{1}{\theta})\Gamma(\frac{3}{\theta})}{\Gamma(\frac{2}{\theta})^2}$$

We can use the sample mean $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$, and $\overline{x^2} = \frac{1}{n} \sum_{i=1}^n x_i^2$ to estimate $E(x)$ and $E(x^2)$ respectively:

$$\frac{\overline{x^2}}{\bar{x}^2} \approx \frac{\Gamma(\frac{1}{\theta})\Gamma(\frac{3}{\theta})}{\Gamma(\frac{2}{\theta})^2}$$

We define the right part of the above equation as a ratio function $\rho(\theta) = \frac{\Gamma(\frac{1}{\theta})\Gamma(\frac{3}{\theta})}{\Gamma(\frac{2}{\theta})^2}$. A precise lookup

table can be constructed by evaluate $\rho(\theta)$ at very small intervals. θ and σ can be estimated from the lookup table and the above equations.

REFERENCES

- [1] Yang, Z., and Purves, D. "A statistical explanation of visual space", *Nat. Neurosci*, vol. 6, no. 6, pp. 632-640 (2003).
- [2] Howe C.Q., and Purves D. "Range image statistics can explain the anomalous perception of length," *Proc. Natl Acad Sci* vol. 99, pp. 13184-13188 (2002).
- [3] Potetz, B., and Lee, T.S. "Statistical correlations between two-dimensional images and three-dimensional structures in natural scenes," *J Opt Soc Am A Opt Image Sci Vis*, vol. 20, no. 7, pp. 1292--1303 (2003).
- [4] Liu, Y., Cormack, L.K., and Bovik A.C., "Disparity statistics in natural scenes," *Journal of Vision*, vol. 8, no. 11, (2008).
- [5] Howard, I.P. and Rogers, B.J. *Binocular vision and stereopsis*. New York: Oxford University Press (1995).
- [6] Tyler, C. W. "Stereoscopic vision: cortical limitations and a disparity scaling effect," *Science*, vol. 181, no. 96, pp. 276-278 (1973).

- [7] Bovik, A.C., Huang, T.S. and Munson, D.C., "A generalization of median filtering using linear combinations of order statistics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 31, no. 6, pp. 1342-1350, (1983).
- [8] Mallat, S. "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674-693 (1989).
- [9] Sharifi, K. and Leon-Garcia, A. "Estimation of shape parameter for generalized Gaussian distribution in subband decomposition of video," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 5, no. 1, pp. 52-56 (1995).
- [10] Simoncelli, E. P., "Statistical modeling of photographic images", in *The Handbook of Image and Video Processing*, A.C. Bovik (Ed.), Academic Press, pp. 431—441, (2005).