NATURAL SCENE STATISTICS OF COLOR AND RANGE

Che-Chun Su, Alan C. Bovik, and Lawrence K. Cormack

Center for Perceptual Systems, The University of Texas at Austin

ABSTRACT

Color and depth play important roles in natural scenes and in vision, and their perception is related. Extensive work has been conducted on studying the luminance statistics of natural scenes; however, there is very little work done on analyzing the statistics between luminance and range in natural scenes, not to mention color and range. In this paper, we present the LIVE Color+3D Database, which contains 12 sets of color images with corresponding ground truth range maps in a high-definition resolution of 1280x720. We examined the statistical distribution of range gradients conditioned on the Gabor responses of the color images, as well as the variations of statistical measures of range gradients with changes in the Gabor responses. The analysis results show that the distributions of range gradients conditioned on the Gabor responses have very similar exponential shapes for both luminance and chrominance channels. Moreover, we also found that the depth difference between neighboring pixels increases as the corresponding magnitudes of the Gabor responses rise.

Index Terms— Natural scene statistics, color, Gabor filters

1. INTRODUCTION

The evolution of human vision systems involves different factors and driving forces, such as natural scene statistics, computational resources in the human brain, and visual tasks humans need to perform [1]. Color and depth play important roles in natural scenes and in vision, and their perception is related [2]. Towards obtaining a better understanding of the statistical relationship between color and range, we studied the joint statistics of color and range using a co-registered database of images [3].

Extensive work has been conducted on studying the luminance statistics of natural scenes [4]. The statistics of natural images are found to exhibit non-Gaussian behavior, and can be represented and modeled by using multi-scale wavelet bases, such as Gaussian scale mixtures [5]. These statistical models have been successfully applied to different image and video applications, such as image de-noising and restoration [6]. Color is one of the most important visual cues in the natural images used by the human vision system to extract more complicated and higher level of information. As with other local image properties such as textures and edges, color can also infer large-scale shape information to help solve visual tasks [7].

However, there is very little work done on analyzing the statistics between luminance and range in natural scenes, not to mention color and range. One major reason for the lack of studies on range statistics has been the dearth of ground truth range data. In this paper, we present the LIVE Color+3D Database [3], in which 12 sets of color images with corresponding ground truth range maps of resolution 1280x720 were collected using an advanced RIEGL-VZ400 laser scanner mounted with a Nikon D700 digital camera [8]. To better approximate color perception in human vision systems, all color images in RGB were transformed into the more perceptually relevant CIELAB color space. We use CIELAB since it is optimized for quantifying perceptual color difference and it better corresponds with human color perception than does the perceptually nonuniform RGB space [9]. Moreover, to emulate how visual neurons respond to visual stimuli, the luminance and chrominance components in natural images were decomposed using Gabor filter banks of diverse different scales and orientations.

The rest of this paper is organized as follows. Section 2 explains the process of acquiring the co-registered LIVE Color+3D Database. The details of analyzing color and range data are described in Section 3, followed by the results in Section 4. Finally, Section 5 gives a discussion.

2. DATA ACQUISITION

The image and range data used in this paper were collected using an advanced range scanner, RIEGL VZ-400, with a Nikon D700 digital camera mounted on top of it [8]. Since there are inevitable translational and rotational shifts when mounting the camera onto the range scanner, calibration is performed before data acquisition. The mounting calibration is done manually using the RIEGL RiSCAN PRO software, which is designed for scanner operation and data processing [10]. Next, to acquire the image and range data in natural scenes, the device obtains distances by lidar reflection and waveform analysis as it rotates, and then the digital camera takes an optical photograph with the same field of view. The acquired range data are exported from the range scanner as point clouds with three-dimensional coordinates and range values, while the image data are stored in the digital camera. Finally, to obtain the aligned 2D range map with the 2D image, the 3D point clouds are projected and transformed into the 2D range map by applying the pinhole camera model with lens distortion [11].

The natural scenes where the image and range data were collected include the campus at The University of Texas at Austin, and the Texas State Capitol. Figure 1 shows two examples of the natural scenes with aligned 2D range maps and 2D images.

3. DATA ANALYSIS

3.1. Pre-processing

Before analyzing the range and color data, some pre-processing must be performed. First, all color images are transformed into the more perceptually relevant CIELAB color space. Then, the luminance and chrominance components are filtered by Gabor filter banks with different scales and orientations to emulate how visual neurons respond to visual stimuli. The Gabor filter bank is adopted to extract features from the luminance and chrominance channels because it closely models image decompositions in primary visual cortex [12]. In gen-
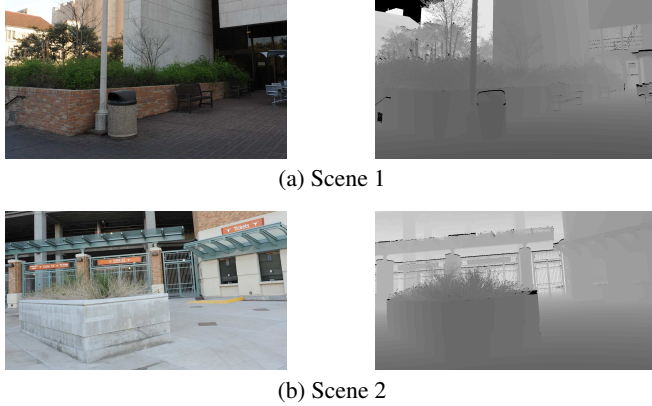
(a) Scene 1

(b) Scene 2

Fig. 1. Two examples of the natural scenes, 2D images on the left and aligned 2D range maps on the right.

eral, the complex 2-D Gabor filter is defined as

$$G(x, y, \sigma_x, \sigma_y, \zeta_x, \zeta_y, \theta)$$
$$= \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{1}{2}\left[\left(\frac{R_1}{\sigma_x}\right)^2 + \left(\frac{R_2}{\sigma_y}\right)^2\right]} e^{i(x\zeta_x + y\zeta_y)} \quad (1)$$

where $R_1 = x\cos\theta + y\sin\theta$ and $R_2 = -x\sin\theta + y\cos\theta$, σ_x and σ_y are the standard deviations of an elliptical Gaussian envelope along x and y axes, ζ_x and ζ_y are the spatial center frequencies of the complex sinusoidal grating, and θ is the orientation.

Since physiological evidence shows that vision neurons in primary visual cortex usually have an elliptical Gaussian envelope with an aspect ratio of 1.5-2.0, with propagating direction along the short axis of the elliptical Gaussian envelope, the 2-D Gabor filters take the form

$$G(x, y, \gamma, \sigma, \omega, \theta) = \frac{1}{2\pi\gamma\sigma^2} e^{-\frac{1}{2}\left[\left(\frac{R_1}{\sigma}\right)^2 + \left(\frac{R_2}{\gamma\sigma}\right)^2\right]} e^{i\omega R_1} \quad (2)$$

where $\gamma = \sigma_y/\sigma_x$ is the aspect ratio of the elliptical Gaussian envelope, $\sigma = \sigma_x$, and $\omega = \sqrt{\zeta_x^2 + \zeta_y^2}$ is the radial center frequency. To adopt a suitable set of Gabor filter banks which can cover most of the frequency domain, the two parameters of the elliptical Gaussian envelope need to be chosen properly, including the aspect ratio, γ, and the standard deviation, σ [13].

Since the simple and complex cells in primary visual cortex have receptive fields at different scales, we utilize a multi-scale set of Gabor filter banks with different spatial frequencies and orientations. Six spatial frequencies, 0.84, 1.37, 2.22, 3.61, 5.87, and 9.53, with the unit of cycles/degree are used, and four different sinusoidal grating orientations are chosen for each spatial frequency: horizontal (0 degree), diagonal-45 (45 degree), vertical (90 degree), and diagonal-135 (135 degree) [14].

3.2. Conditional distributions and statistics

The perception of three-dimensional geometry occurs as a cyclopean 3D image in the brain [15]; however, the human vision system is able to sense depth from a single static image. This capability implies that statistical relationships may exist between the luminance/chrominance information in two-dimensional images and the three-dimensional range information in the natural environments. In

Table 1. Pearson's chi-square value for the fitted exponential distributions

	L*	a*	b*
χ^2	5.9652	5.7511	7.0583

addition, the depth information acquired by the human vision system is more relative than absolute, which means that we know which objects are further and which are closer, but we are not sure about the exact distance of each object from us. Therefore, to explore the statistical relationship between color and range, we examined the conditional distribution of the magnitude of range gradients over the magnitude of Gabor responses to the three CIELAB color channels, L*, a*, and b*.

4. RESULTS

Figure 2 shows the conditional distribution of the magnitude of range gradients on the magnitude of image Gabor responses for L*, a*, and b* channels at the sub-band with frequency = 5.87 (cycle/degree) and orientation = horizontal (0 degree). We can see that for both luminance and chrominance channels, the conditional distributions are quite similar. Also shown in Figure 2 are the fitted Weibull distributions represented by the dotted lines, and we can see that these conditional distributions are well fitted by the Weibull functions. Table 1 shows the corresponding Pearsons chi-square values to test the goodness-of-fit for the fitted Weibull distributions. The chi-square values for all three channels are much smaller than the upper 5% cutoff value of the chi-square distribution, $\chi^2_{(.05)} = 101.88$ with degree of freedom = 80, which means that by more than chance these conditional distributions are drawn from the fitted Weibull distributions. Note that the conditional distributions remain Weibull-shaped at other Gabor response values and similar results exist for all frequencies and orientations.

Figure 3 (a) shows the conditional distribution of magnitude of range gradients as a function of magnitude of image Gabor responses for L* channel and Figure 3 (b) plots the corresponding shape and scale parameters of the fitted Weibull distributions. We can see that given different values of Gabor responses, the conditional distribution of range gradients has the same shape which can be well fitted by the Weibull function. In addition, the conditional distribution has heavier tails with larger Gabor responses, which can be demonstrated by the monotonically decreasing shape parameters as a function of Gabor responses shown in Figure 3 (b). Note that the same findings exist for chromatic channels which are not shown here due to the limited space.

Finally, Figure 4 plots the mean magnitude of range gradients as functions of magnitude of Gabor responses for luminance and chrominance channels. The three panels in both Figure 4 (a) and (b) show the responses of horizontal Gabors at three different spatial frequencies. For the luminance channel, the gradient of range increases monotonically with small Gabor responses and seems to get saturated with larger Gabor responses at all frequencies. On the other hand, the gradient of range increases monotonically with Gabor responses at all frequencies for the chromatic channels, which implies that the chrominance information can also possibly be utilized in depth perception.
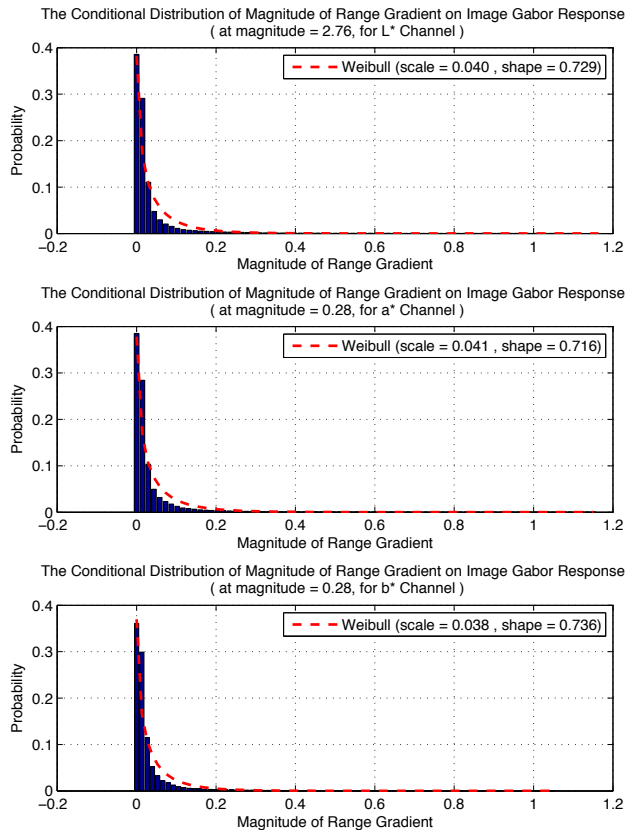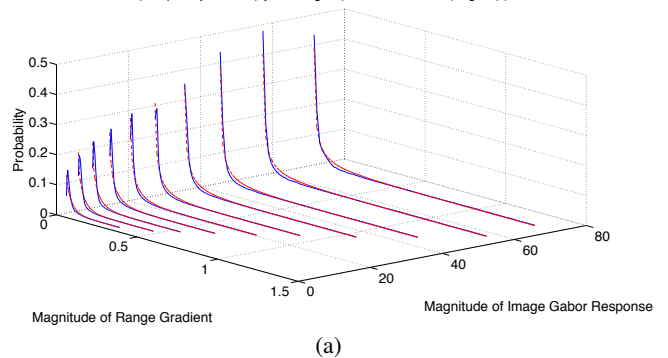
Fig. 2. The conditional distribution of the magnitude of range gradients on the magnitude of image Gabor responses for the three CIELAB color channels, at frequency = 5.87 (cycle/degree) and orientation = horizontal (0 degree). From top to bottom: L*, a*, and b*. The dotted line represents the fitted Weibull distribution.

5. DISCUSSION

In this paper, we have examined the conditional distributions relating the luminance or chrominance information with range gradients. Of more relevance to perception and image processing, the distributions of range gradients conditioned on the Gabor responses, whether luminance or chrominance, have very similar Weibull shapes. We have also examined the variations of statistical measures, e.g. mean, of range gradients with variations in the Gabor responses. Most importantly, we found that the depth difference between neighboring pixels increases as the corresponding magnitudes of the Gabor responses rise. Therefore, the way these range distributions vary with the Gabor responses indicates that the visual system could, in principle, use these conditional statistics to help recover depth information from the environment. Moreover, these statistical relationships not only yield insight into how 3D structure in the environment might be recovered from image data, but may also be applied to various image and video engineering applications, e.g. image de-noising and restoration, quality assessment of 3D images and video, and stereo correspondence algorithms [16]. In our on-going research, we are finding that the inclusion of chrominance cues can augment the performance of Bayesian stereo algorithms, in agreement with similar conclusions regarding human stereopsis and early stereo algorithms [2, 17, 18].
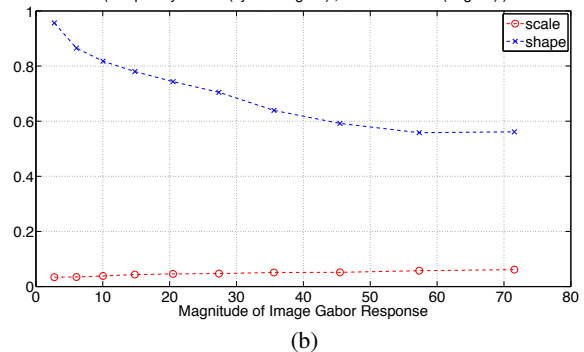

(a)

(b)

Fig. 3. (a) The conditional distribution of magnitude of range gradients as a function of magnitude of image Gabor responses, and (b) the corresponding shape and scale parameters of the fitted Weibull distributions for L* channel at frequency = 5.87 (cycle/degree) and orientation = horizontal (0 degree).

6. REFERENCES

[1] B.A. Olshausen and D.J. Field, "Vision and the coding of natural images," *American Scientist*, vol. 88, pp. 238, 2000.

[2] J.R. Jordan, W.S. Geisler, and A.C. Bovik, "Color as a source of information in the stereo correspondence process," *Vision research*, vol. 30, no. 12, pp. 1955–1970, 1990.

[3] LIVE, "LIVE Color+3D Database," http://live.ece.utexas.edu.

[4] D.J. Field, "Wavelets, vision and the statistics of natural scenes," *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 357, no. 1760, pp. 2527, 1999.

[5] S. Lyu and E.P. Simoncelli, "Statistical modeling of images with fields of Gaussian scale mixtures," *Advances in Neural Information Processing Systems*, vol. 19, pp. 945, 2007.

[6] J. Portilla, V. Strela, M.J. Wainwright, and E.P. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain," *Image Processing, IEEE Transactions on*, vol. 12, no. 11, pp. 1338–1351, nov. 2003.

[7] I. Fine, D.I.A. MacLeod, and G.M. Boynton, "Surface segmentation based on the luminance and color statistics of natural scenes," *JOSA A*, vol. 20, no. 7, pp. 1283–1291, 2003.
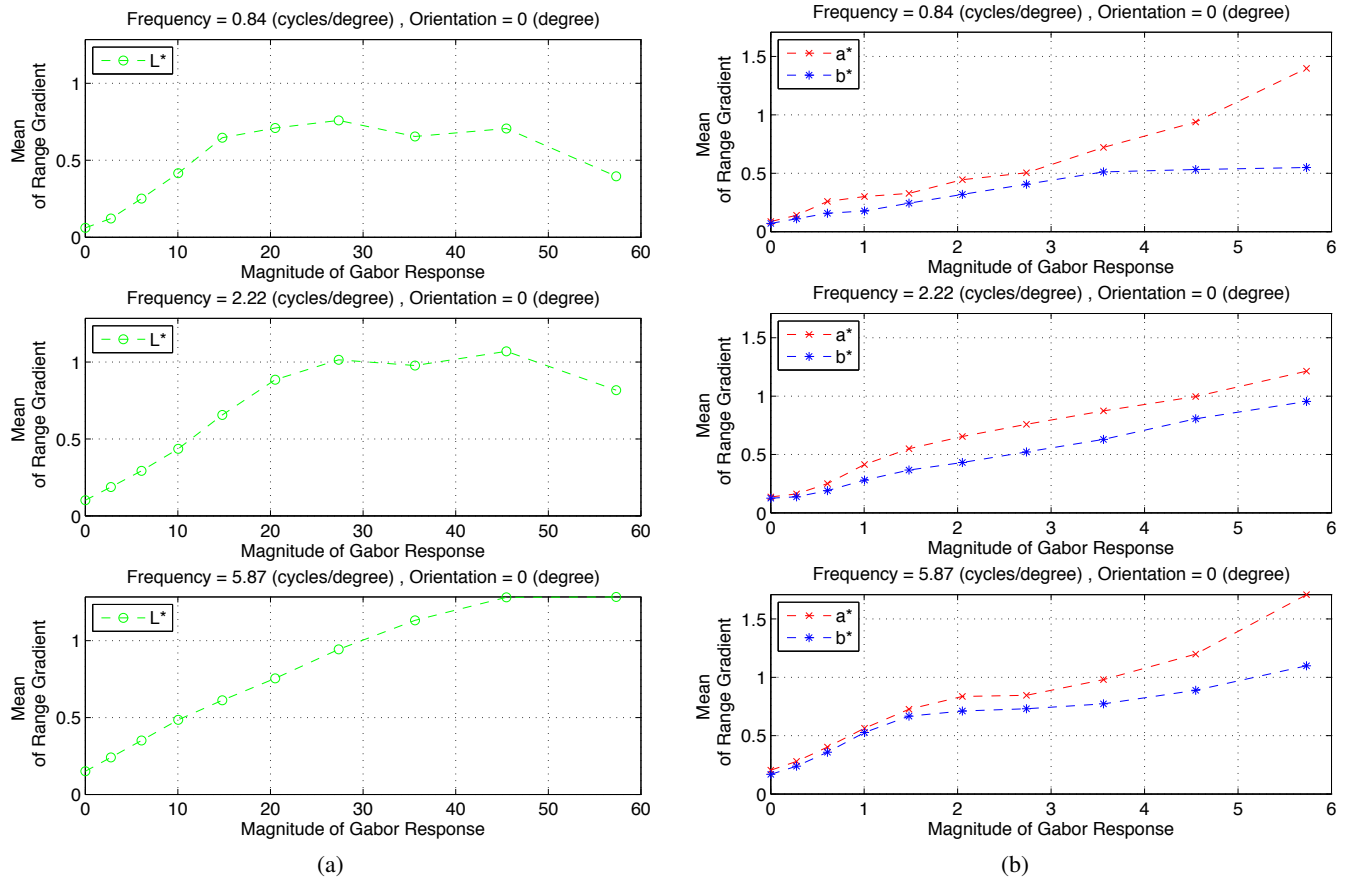
Fig. 4. The mean of the magnitude of range gradients as functions of the magnitude of Gabor responses, at three different spatial frequencies at the same horizontal orientation for: (a) L* channel and (b) a* and b* channels.

[8] RIEGL, "RIEGL VZ-400 3D Terrestrial Laser Scanner," http://rieglusa.com/products/terrestrial/vz-400/index.shtml.

[9] U. Rajashekar, Z. Wang, and E.P. Simoncelli, "Perceptual quality assessment of color images using adaptive signal representation," *Human Vision and Electronic Imaging XV*, vol. 7527, jan. 2010.

[10] RIEGL, "RIEGL RiSCAN PRO Software for 3D Terrestrial Laser Scanner," http://rieglusa.com/products/terrestrial/vz-400/software.shtml.

[11] Intel Corporation, "OpenCV: Camera Calibration and 3D Reconstruction," http://opencv.willowgarage.com/documentation/camera_calibration_and_3d_reconstruction.html.

[12] D.J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," *Journal of the Optical Society of America A*, vol. 4, no. 12, pp. 2379–2394, 1987.

[13] K. Seshadrinathan and A.C. Bovik, "Motion tuned spatio-temporal quality assessment of natural videos," *Image Processing, IEEE Transactions on*, vol. 19, no. 2, pp. 335–350, feb. 2010.

[14] S. Bhagavathy, J. Tesic, and BS Manjunath, "On the Rayleigh nature of Gabor filter outputs," in *Image Processing, 2003.*

ICIP 2003. Proceedings. 2003 International Conference on. IEEE, sept. 2003, vol. 3.

[15] B. Julesz, *Foundations of cyclopean perception.*, University of Chicago Press, 1971.

[16] Y. Liu, L. Cormack, and A. Bovik, "Statistical modeling of 3d natural scenes with application to bayesian stereopsis," *Image Processing, IEEE Transactions on*, to be published.

[17] John R. Jordan and Alan C. Bovik, "Using chromatic information in edge-based stereo correspondence," *Computer Vision, Graphics, and Image Processing: Image Understanding*, vol. 54, no. 1, pp. 98–118, july 1991.

[18] John R. Jordan and Alan C. Bovik, "Using chromatic information in dense stereo correspondence," *Pattern Recognition*, vol. 25, no. 4, pp. 367–383, april 1992.