Statistics of Range Images

Jinggang Huang
Division of Applied Math
Box F, Brown University
Providence, RI02912
jhuang@cfm.brown.edu

Ann B. Lee
Dept. of Physics
Box 1843, Brown University
Providence, RI02912
besu@cns.brown.edu

David Mumford
Division of Applied Math
Box F, Brown University
Providence, RI02912
David_Mumford@brown.edu

Abstract

The statistics of range images from natural environments is a largely unexplored field of research. It closely relates to the statistical modeling of the scene geometry in natural environments, and the modeling of optical natural images. We have used a 3D laser range-finder to collect range images from mixed forest scenes. The images are here analyzed with respect to different statistics.

1 Introduction

More and more people realize the importance of the statistics of optical images. Good statistical models of optical images are important to applications, such as image compression, noise removal and segmentation. The models may also further our understanding of the encoding of visual information in biological visual systems. Not much research has, however, been done on the statistics of *range images*. There are two main reasons why range images are interesting. First, they lead to a direct understanding of the stochastic nature of the geometry of the world, and give priors for stereo algorithms ([1],[2]). Secondly, they make it possible to construct more realistic models of optical images. For example, authors in [11], [3] and [8] have modeled optical images as a perspective view of the 3D world, with objects of random geometry (size, shape, position) and intensity. The object geometries in these models are usually based on assumptions, which have not been directly verified in real scenes. There is no doubt, that with a fairly large data base of range images, we will better understand the scene geometry of the 3D world, and thus be able to develop more realistic models for optical images.

We have collected 205 panoramic range images from varying environments (both outdoor and indoor scenes) — 54 of these images were taken in different forests in Rhode Island and Massachusetts during August-September. In this paper, we will focus on the forest scenes, because the statistics of these images appear more stable than those of other categories (for example, residential and interior scenes). Figure 1 shows a sample.

Figure 1: A sample image from our range image data base

We used a laser range-finder with a rotating mirror [1] to collect the range images. Each image contains 444×1440 measurements with an angular separation of 0.18 deg. The field of view is thus 80^0 vertically and 259^0 horizontally. Each measurement is calculated from the time of flight of the laser beam. The operational range of the sensor is typically 2-200m. The laser wavelength of the range-finder is $0.9\mu m$, which is in the near infra-red region.

Throughout this paper, we will work with log(range) instead of range directly, because the former statistics is closer to being shape invariant. Figure 2 shows a top view of a laser range-finder (see circle) centered at O, and two homothetic triangles $\triangle ABC$ and $\triangle A'B'C'$ (P_A, P_B and P_C correspond to three pixels in the range image). Assume that the distances between O and the vertices of $\triangle ABC$ are r_A, r_B and r_C, respectively, and the distances between O and the vertices of $\triangle A'B'C'$ are R_A, R_B and R_C, respectively. Let

$$D = range(P_A) - range(P_B)$$

be the difference in range for pixels P_A and P_B, and

[1] 3D imaging sensor LMS-Z210 by Riegl

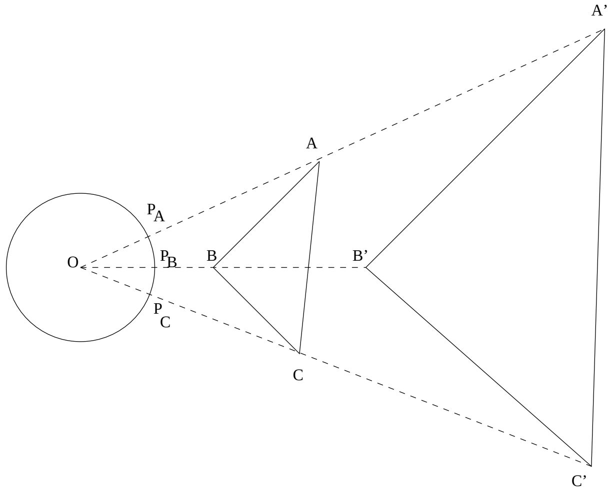

Figure 2: A laser range finder centered at O and two homothetic triangles, $\triangle ABC$ and $\triangle A'B'C'$.

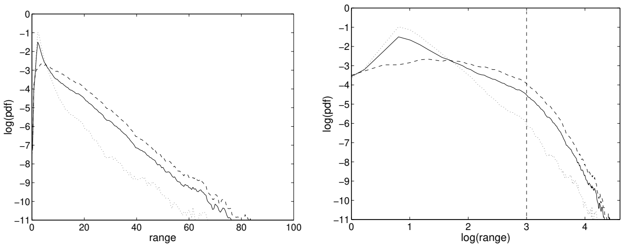

Figure 3: Semi-log (left) and log-log (right) plots of the single pixel statistics (i.e. range statistics). *Solid*: distribution for the whole image. *Dotted*: distribution for the bottom half of a range image. *Dashed*: distribution for the top half of a range image.

let

$$\widehat{D} = \log(range(P_A)) - \log(range(P_B)) = \log\left(\frac{range(P_A)}{range(P_B)}\right)$$

be the difference in *log* range for the same two pixels. Then, a scene with $\triangle ABC$ and a scene with triangle $\triangle A'B'C'$ will lead to different values for D ($r_A - r_B$ vs. $R_A - R_B$) but the *same* value for $\widehat{D}$ ($log(\frac{r_A}{r_B}) = log(\frac{R_A}{R_B})$). Hence, log(range) is appropriate if we want the difference statistics (or any mean-0-filter reaction) to be "shape invariant".

2 Single-Pixel Statistics

The solid line in Figure 3 shows the single-pixel statistics of log(range) images. We observe a sudden change in slope at a range of about 20 meters (or $\log(range) \approx 3$; see vertical line in figure) — this may be related to the accumulation of occlusion effects. In Figure 3, we have also plotted the log(range) histograms for the top half (dashed line) and bottom half (dotted line) of a range image separately. The two halves correspond to different distributions of objects — mainly ground for the top half and mainly trees for the bottom half — and display quite different statistics. The distribution from the top half has an approximately linear tail in a semi-log plot (indicating an exponential decay $\sim e^{-0.12r}$), while the bottom half shows an approximately linear tail in a log-log plot (indicating a power law $\sim r^{-2.6}$). We can qualitatively explain the two different behaviors with the following simplified models:

For the top half, we assume tree trunks (cylinders) uniformly distributed on a plane, according to a Poisson process with density λ. Figure 4 shows a top view of a randomly generated "forest" scene. Each disk represents a cross section of the trunk of a tree. If we assume all disks are of diameter L, a simple calculation will show that the probability that a horizontal beam from the laser range-finder first hits a tree at distance r is given by an *exponential* distribution

$$f(r) \sim \lambda L e^{-\lambda L r} .$$

For the bottom half, we assume flat ground only. Let the height of the sensor from the ground be H, as shown in Figure 5. Then at angle θ, the distance between the sensor and the ground is $r = \frac{H}{sin\theta}$. The laser range-finder samples evenly with respect to the polar coordinate(s) θ (and ϕ), i.e.

$$f(\theta) = constant$$

for $\pi/2 < \theta < \pi$. With respect to r, we then get a *power-law* distribution

$$f(r) = \frac{H}{r^2\sqrt{(1-(H/r)^2)}} \sim \frac{1}{r^2} .$$

3 Derivative Statistics

We now look at the marginal distribution of the horizontal derivative D, which in the discrete case, is simply the difference between two adjacent pixels in the same row. The solid line in Figure 6 shows the log probability density function of D. As in the studies of optical images, this distribution has a high kurtosis with large tails, and a peak at 0. It is closest to

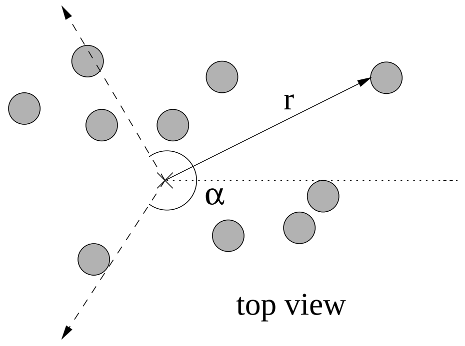

Figure 4: Top view of a randomly generated forest scene. See text.

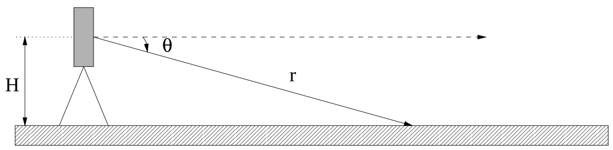

Figure 5: Ground model. See text.

the statistics for optical images of man-made environments [8], but has an even higher peak at 0. This indicates that the structure of range images is simpler than for optical images — an issue that we will explore further in the next section. The derivative statistic has also been used to test the scale invariance property (see Sinai [13], Field [5], Ruderman [12], and for scale invariance) in natural images. Given an image I, we define a scaled-down image $I^{(N)}$ by computing the average of $N \times N$ disjoint blocks. In [12], [14], [3] and [8], it has been shown that, for optical natural images, the distribution of D is the same for I and $I^{(N)}$. In this paper, we scale down the images by taking the minimum, instead of the average of $N \times N$ blocks. This is the appropriate renormalization for range images, because laser range finders measure the distance to the nearest object in each solid angle. Figure 6 shows the distribution of D at different scales. The results indicate that range images scale well.

4 Bivariate Statistics

Below, we study the co-occurrence statistics of two pixels with different separation distances. Let

$$K(a, b|x) = Pr\{I(\mathbf{x_1}) = a, I(\mathbf{x_2}) = b \mid ||\mathbf{x_1} - \mathbf{x_2}|| = x\}$$

where $I(\mathbf{x_1})$ and $I(\mathbf{x_2})$ represent the log(range) at pixels $\mathbf{x_1}$ and $\mathbf{x_2}$. The left column of Figure 7 shows

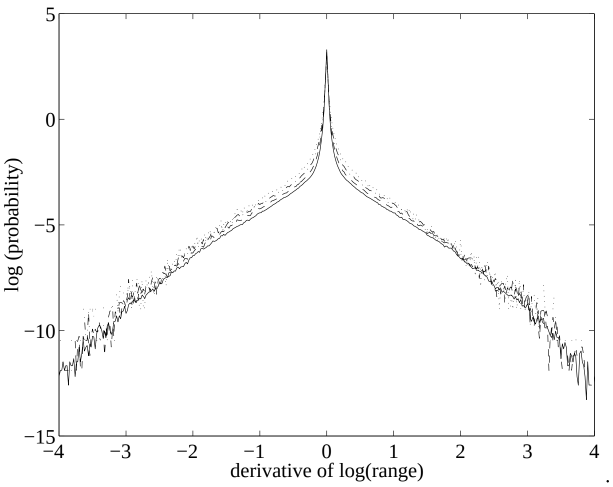

Figure 6: Derivative statistics at different scales.

the contour plots of $K(a, b|x)$ for separation distances $x = 1, 16, 256$.

The authors in [8] propose the following model for the bivariate statistics of optical images:

$$K(a, b|x) = [1 - \lambda(x)]q(a)q(b) + \\ 2\lambda(x)h_x(a + b)g_x(b - a)$$

where q is the marginal distribution for a single pixel, h_x are distributions similar to q, and g_x are distributions highly concentrated at 0. The first term models the case where the two pixels belong to different objects (we assume that different object are statistically independent), the second term represents the case where they are on the same object (assume that the sum and the difference of the pixel values are independent), and $\lambda(x)$ is the probability of their being on the same object.

The right column of Figure 7 shows a fit of the model above to the empirical bivariate statistics (left column) of range images; Figure 8 shows the functions used in the fit. The mixture nature — "same" versus "different" objects — of both of data and the model, is seen in the changing shape of the contour levels in the center versus the tails. We also find that the random collage model fits *better* to the bivariate statistics of range images than to that for optical images (see [8]). Again, this indicates that range images present a simpler, cleaner problem than optical images. For example, the concept of objects is better defined for range images where we do not need to take lighting, color variations, texture etc. into account.

5 Joint Statistics of Haar Wavelet Coefficients

It has been shown that, the statistics in the wavelet domain can be very informative ([4], [7], [8]). We here choose the Haar wavelet for its simplicity: Any structure in the statistics can be directly related to pixel values. Note also that the central limit theorem can easily make the statistics of larger support filters appear misleadingly Gaussian.

Figure 9 shows the different Haar filters. To describe how the wavelet coefficients in the subbands are related to each other, we use the same definitions as in [4]: Coefficients at adjacent spatial locations in the same subband are *brothers*, and coefficients in the same level and the same position, but different orientations are *cousins*.

Figure 10 shows contour plots of the joint density functions for different wavelet coefficient pairs. We observe strong cusps in all contour level curves; most of which lie on the lines $x = 0$, $y = 0$, $x = \pm y$ and $x = \pm 2y$. As a comparison, look at Figure 11 where the corresponding contour plots are calculated for optical images in van Haterens's image database [6]. We see that many cusps occur along the same lines, but are not as peaked.

Furthermore, we did the following experiments on the range images:

1. We scaled the images down by taking the block minimum and calculated the joint histograms. Figure 12 shows the result.

2. We scaled the images down by taking the block average and calculated the joint histograms, Figure 13 shows the result.

When we scale a range image down by block minimum, the resulting image is approximately the same as the image taken in the same environment at half resolution (as mentioned before, range images are almost scale invariant under block minimum). The second method (block averaging) is, however, closer to how a digital camera for optical images works: The intensity at each pixel is the average (or some weighted mean) of the intensity at points covered by that pixel. This explains why the statistics shown in Figure 12 is similar to that in Figure 10, and the statistics in Figure 13 is similar to that in Figure 11.

Note that the observed cusps in Figure 10 are not caused by noise, but correspond to real structures in the images. It is important to see what these structures are:

For the cousin pairs (horizontal and vertical; horizontal and diagonal) this is relatively straight-forward — because the joint distribution of horizontal, vertical and diagonal wavelet coefficients is a sufficient statistics for 2×2 blocks modulo mean (3 variables). Figure 14 shows an equiprobable surface of the 3D joint distribution (horizontal, vertical, diagonal). We see vertices along the lines $y = z = 0$, $x = z = 0$, $x = y = 0$ and $x = \pm y = \pm z$. Simple calculations show that these vertices correspond to the following 2×2 blocks and their rotations:

$$\begin{pmatrix} a & a \\ b & b \end{pmatrix}, \begin{pmatrix} a & b \\ b & a \end{pmatrix}, \begin{pmatrix} a & b \\ b & b \end{pmatrix}$$

For the *horizontal filter and left brother* pair, we see cusps along $y = 0$, $x = 0$, $y = x$, $y = 2x$ and $y = \frac{1}{2}x$. To find the 2×4 patches that correspond to these vertices, we sample randomly from our database for patches with strong filter reactions along $y = 0$, $y = x$ and $y = \frac{1}{2}x$. The 8 range values in the typical patches fall clearly into 2 and occasionally 3 tight clusters. Calling the clusters a,b, and c, we get for example:

direction	*typical patches*
$y = 0$	$\begin{pmatrix} a & b & b & c \\ b & b & b & c \end{pmatrix}, \begin{pmatrix} a & a & b & c \\ b & b & b & c \end{pmatrix}, \begin{pmatrix} a & b & a & a \\ a & a & a & a \end{pmatrix}$
$y = x$	$\begin{pmatrix} a & b & b & b \\ a & a & a & b \end{pmatrix}, \begin{pmatrix} a & b & b & a \\ a & a & a & a \end{pmatrix}, \begin{pmatrix} a & a & a & a \\ b & b & b & b \end{pmatrix}, \begin{pmatrix} a & a & a & a \\ b & b & b & b \end{pmatrix},$
$y = \frac{1}{2}x$	$\begin{pmatrix} a & a & a & b \\ b & b & b & b \end{pmatrix}$

For the *horizontal filter and upper brother* pair, we see cusps along $y = 0$, $x = 0$, $y = \pm x$, $y = -2x$ and $y = -\frac{1}{2}x$. The typical patches in the database, that lead to high filter reactions along these lines are:

direction	*typical patches*
$y=0$	$\begin{pmatrix} a & a \\ b & b \\ b & b \\ b & b \end{pmatrix}$
$y=-x$	$\begin{pmatrix} a & a \\ b & a \\ b & a \\ a & a \end{pmatrix}, \begin{pmatrix} a & a \\ b & a \\ b & a \\ a & a \end{pmatrix}, \begin{pmatrix} a & a \\ b & b \\ b & b \\ a & a \end{pmatrix}, \begin{pmatrix} a & b \\ b & b \\ b & b \\ b & a \end{pmatrix}$
$y=-2x$	$\begin{pmatrix} a & b \\ b & b \\ b & b \\ a & a \end{pmatrix}, \begin{pmatrix} a & a \\ b & a \\ b & b \\ a & a \end{pmatrix}$
$y=x$	$\begin{pmatrix} a & a \\ b & b \\ a & a \\ b & b \end{pmatrix}, \begin{pmatrix} a & a \\ a & b \\ a & b \\ b & b \end{pmatrix}, \begin{pmatrix} a & b \\ b & b \\ a & a \\ a & b \end{pmatrix}$

We see that all the striking cusps in the contour plots in the Haar wavelet domain relate to the piecewise constant geometric structure in range images.

6 Conclusions

We have investigated several simple statistics of range images of forest scenes. The bivariate statistics of two pixels verify the basic assumption of the random collage model — that the world can be broken down into piecewise smooth regions (in, for example, range or intensity) that depend little on each other. The 2D and 3D joint distributions of Haar filter reactions show some striking features, which also indicate the presence of piecewise constant geometric structures and sharp discontinuities in range images. The above results all point to the fact that range images are much simpler to analyze than optical images: The concept of "objects" is often better defined in terms of changes in range, than in terms of changes in intensity, color, texture, lighting etc. We believe that segmenting range images from natural scenes and a thorough analysis of the results — for example, the structure *within* objects, and the sizes, positions, and dependencies *between* different objects — will lead to a better understanding of the scene geometry of the 3D world, as well as more realistic statistical models for optical images.

References

[1] P. N. Belhumeur and D. Mumford. "A bayesian treatment of the stereo correspondence problem using half-occluded regions". *Proc. IEEE Conference Computer Vision and Pattern Recognition, 1992* 506-512

[2] P. N. Belhumeur. "A Binocular Stereo Algorithm for Reconstructing Sloping, Creased, and Broken Surfaces in the Presence of Half-Occlusion". *Proc. Fourth International Conference on Computer Vision* 431-438

[3] Z. Chi. "Probability Models for Complex Systems". Ph.D. Thesis, Div. of Applied Math., Brown Univ. 1998

[4] R. W. Buccigrossi and E. P. Simoncelli. "Image Compression via Joint Statistical Characterization in the Wavelet Domain". *GRASP Laboratory Technical Report #414, Univ. of Pennsylvania*, available at: http://www.cns.nyu.edu/~eero/publications.html

[5] D. J. Field. "Scale-invariance and self-similar 'wavelet' transforms: An analysis of natural scenes and mammalian visual systems". In M. Farge, J. C. R. Hunt, and J. C. Vassilicos, editors, *Wavelets, fractals, and Fourier transforms* 151-193. Clarendon Press, Oxford, 1993

[6] J. H. van Hateren and A. van der Schaaf. "Independent Component Filters of Natural Images Compared with Simple Cells in Primary Visual Cortex" Proc.R.Soc.Lond. B 265:359-366,1998

[7] J. Huang and D. Mumford. "Statistics of Natural Images and Models". *Proc. IEEE Conference Computer Vision and Pattern Recognition, 1999* 541-547

[8] A. B. Lee, J. Huang and D. Mumford. "Random-Collage Model for Natural Images". Submitted to *International Journal of Computer Vision*

[9] A. B. Lee and D. Mumford. "Scale-Invariant Random-Collage Model for Natural Images". In *Proc. IEEE Workshop on Statistical and Computational Theories of Vision*, Fort Collins, CO, 1999

[10] D. Mumford and B. Gidas. "Stochastic Models for Generic Images". Submitted to *Quarterly Journal of Mathematics*.

[11] D. L. Ruderman. "Origins of Scaling in Natural Images". *Vision Research* vol 37, No23, 3385-3395, 1997

[12] D. L. Ruderman. "The Statistics of Natural Images". *Network*, **5**(4):517-548, 1994

[13] Y. G. Sinai. "Self_Similar Probability Distributions". *Theory of Probability and Its Applications* Volume XXI, 64-80, 1976

[14] S. C. Zhu, and D. Mumford. "Prior Learning and Gibbs Reaction-Diffusion". *IEEE Transactions on PAMI*, vol 19, no 11, pp 1236-1250, 1997.

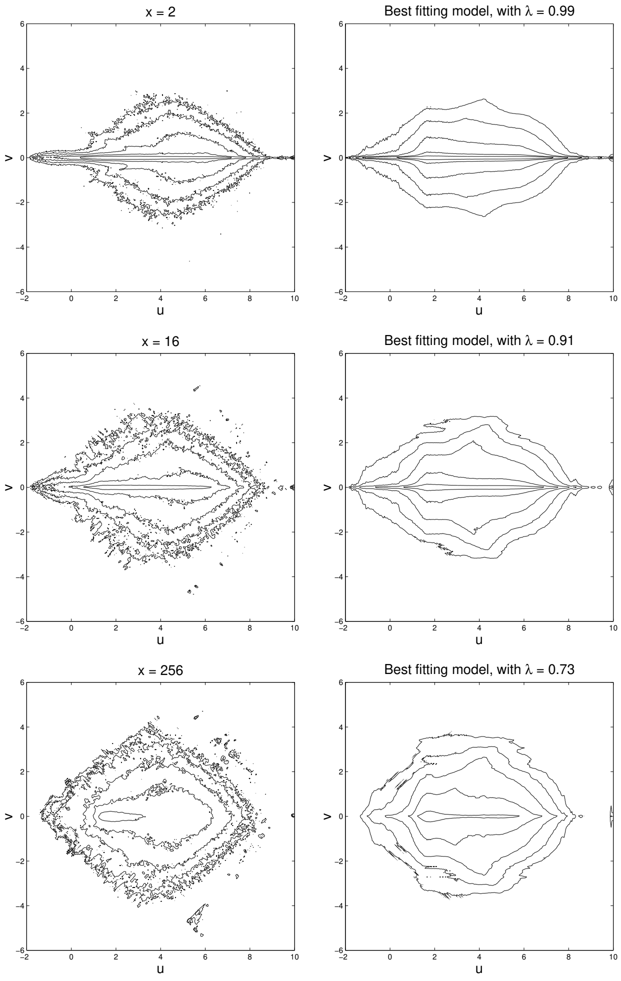

Figure 7: Contour plots of the log histograms of pixel pairs for range images (left column) and the best bivariate fit to the random collage model (right column). x: distance between the two pixels, u: sum of the pixel values, v: difference of the two pixel values.

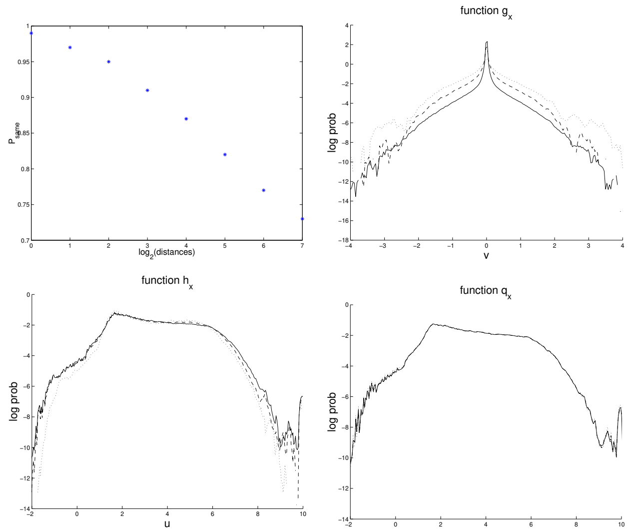

Figure 8: The values for λ and the 1D functions g_x, h_x and q_x in the best bivariate fit at pixel separation distances $x = 2$ (solid), 16 (dashed), and 256 (dotted).

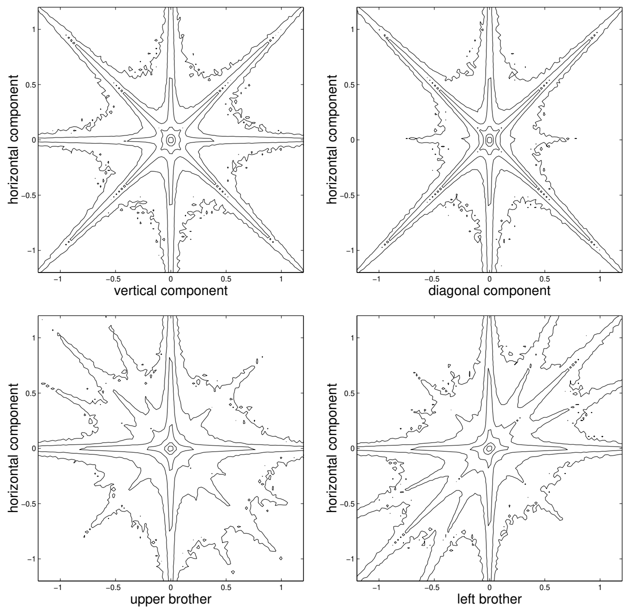

Figure 10: Contour plots of the log histograms of wavelet coefficient pairs for range images.

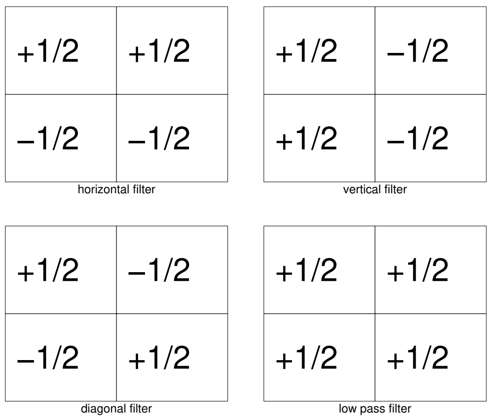

Figure 9: Haar filters

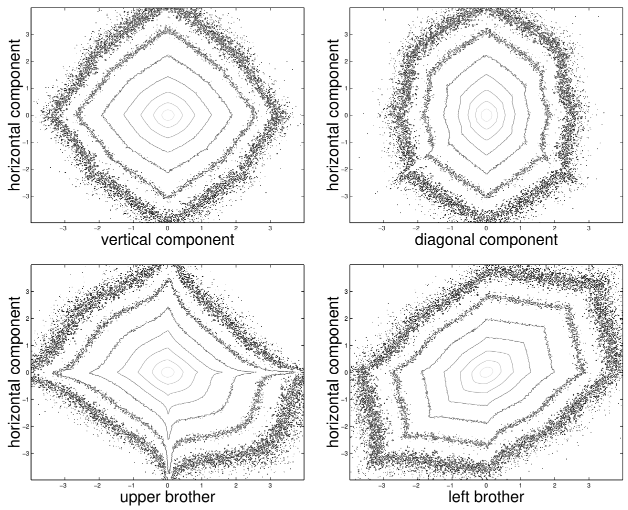

Figure 11: Contour plot of the log histogram of wavelet coefficient pairs for natural optical images.

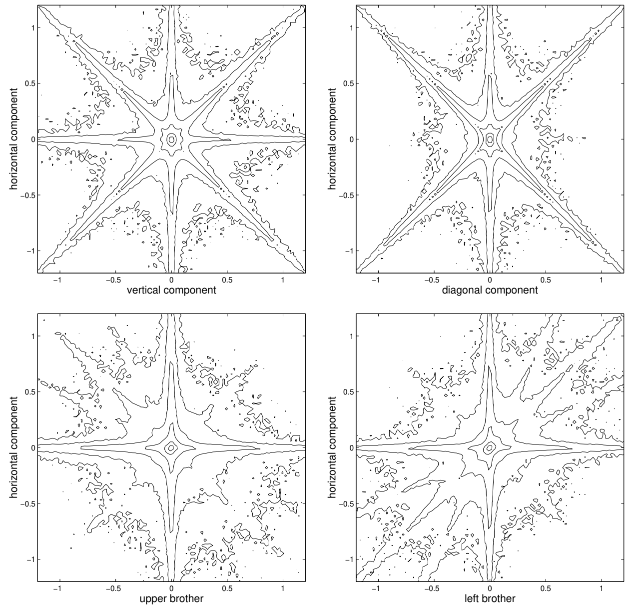

Figure 12: Contour plots of the log histogram of wavelet coefficient pairs, calculated from range images scaled down by taking the minimum of $2X2$ blocks.

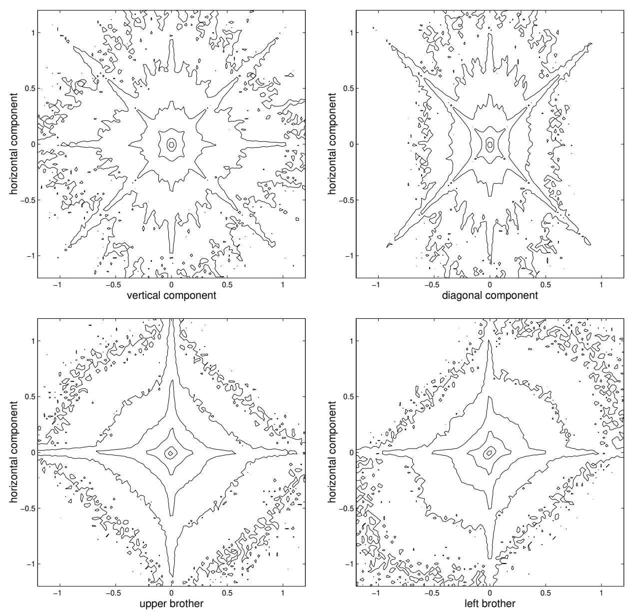

Figure 13: Contour plots of the log histogram of wavelet coefficient pairs, calculated from range images scaled down by taking the average of 2×2 blocks.

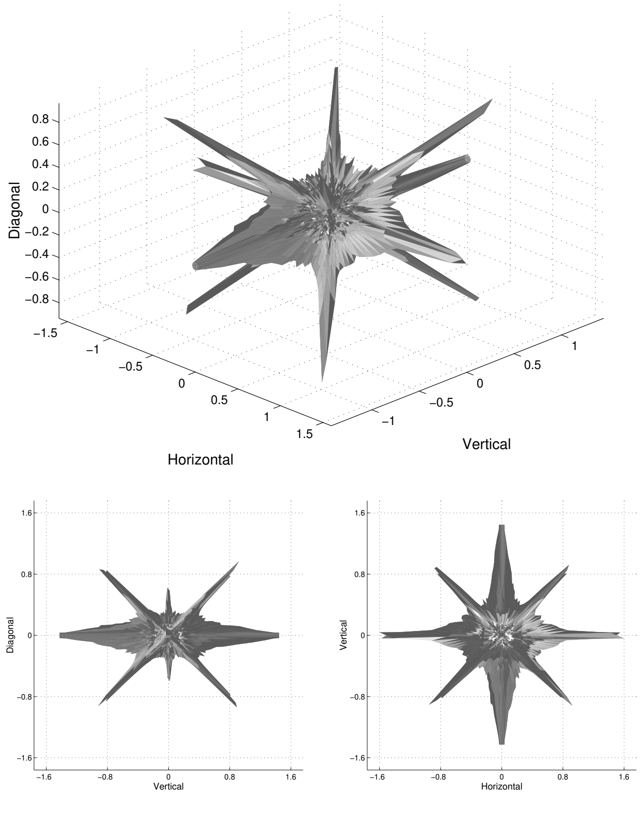

Figure 14: An equi-surface of a 3D joint histogram of horizontal, vertical and diagonal wavelet coefficients in range images, viewed from three different angles.