Deriving Line and Surface Orientation by Statistical Methods on the Unit Sphere

Robert T. Collins
Richard S. Weiss

COINS TR90-102

October 1990

Deriving Line and Surface Orientation by Statistical Methods on the Unit Sphere

Robert T. Collins and Richard S. Weiss

Computer and Information Science Department
University of Massachusetts at Amherst
Amherst, MA. 01003 *

October 12, 1990

This paper demonstrates how confidence regions for line and surface orientation can be obtained directly from stereo line correspondences and vanishing point analysis. Since orientations are represented as unit vectors, statistical techniques for estimating the axes and uncertainty of point distributions on the unit sphere are explored. Bingham's distribution is introduced to describe both equatorial and bipolar distributions of unit vectors. Statistical parameter estimation based on Bingham's distribution, as well as a general nonparametric estimation method, are used to solve for the polar axis of a great circle of points and to represent the statistical uncertainty in the resulting orientation estimate.

1 Introduction

Many objects in man-made environments have planar surfaces. For automatic construction of 3D models it is necessary to recover these surfaces together with an estimate of their uncertainty. Line and plane orientations can be computed whenever the images of parallel 3D lines are observed. Particular examples are vanishing point analysis on monocular images of groups of parallel lines, and the analysis of multiple images of a single 3D line undergoing pure translation with respect to the camera, either as part of translational motion sequence, or from a stereo image pair where the camera axes are aligned.

The represention of orientations as unit vectors leads naturally to an examination of distributions on the surface of the unit sphere, and to statistical inferencing techniques over such distributions. One useful tool for deriving orientation information is the estimation of a unit vector perpendicular to a set of derived unit vectors. For instance, the normal vector to a planar surface is perpendicular to the orientations of all the lines in the plane. For two input vectors, this inference is a cross product with an uncertainty that reflects the

*This work was supported in part by DARPA and RADC under contract number F30602-87-C-0140, by DARPA and U.S. Army ETL under contract number DACA76-89-C-0017, and by the National Science Foundation under grant number DCR-8500332.

uncertainty in the original vectors. For more than two vectors, the problem can be viewed as finding the polar axis of a great circle defined by the heads of several uncertain unit vectors. Perpendicularity constraints between orientation vectors also lead to an efficient Hough transform technique on the sphere for clustering vectors that lie in a great circle.

This paper is a compendium of two separate papers [7, 8] in which the theory of directional statistics is applied to the estimation of orientation vectors in computer vision. The first paper focuses on the extraction and statistical description of line and plane orientations from stereo line pair correspondences. The second paper details an application of directional estimation techniques in vanishing point analysis.

2 Orientations from Stereo Line Correspondences

This section demonstrates how confidence regions for line and surface orientation can be obtained directly from stereo line correspondences. Bingham's distribution is introduced to describe both equatorial and bipolar distributions of unit vectors. Statistical parameter estimation based on Bingham's distribution is used to solve for the polar axis of a great circle of points and to represent the statistical uncertainty in the resulting orientation estimate. In addition, the problem of estimating the orientation and uncertainty of the cross product of two uncertain unit vectors is considered, and an approximate solution is given in terms of the "intersection" of two equatorial Bingham distributions.

2.1 Motivation

The orientation of 3D lines can be computed directly from stereo line correspondences without first computing point depths, in contrast to methods that compute depth in order to obtain the orientation of line segments. After computing line directions, it is possible to efficiently discover coplanar lines and thereby recover the orientation and distance of planar surfaces in the scene.

2.1.1 Line Orientations from Stereo

Consider a 3D line segment with unit orientation vector $\boldsymbol{U}$, projecting onto the image plane of a single camera. The focal point of the camera together with the 3D line defines a plane called the *projection plane* of the line. The image projection of the 3D line lies on the intersection of the projection plane and the image plane, thus the projection plane can be computed given a line segment in the image and the focal point. Since a 3D line lies in its projection plane, the plane normal $\boldsymbol{\phi}$ is perpendicular to the orientation $\boldsymbol{U}$. If the same 3D line is imaged from a second camera, oriented the same as the first, but translated by a vector $\boldsymbol{T}$, a second projection plane that is still perpendicular to $\boldsymbol{U}$ will be measured. The 3D line orientation can thus be recovered as a unit vector parallel to the cross product of the two projection plane normals, except when the line image lies along an epipolar line for the two images.

Looking at this another way, translating the coordinate system by $\boldsymbol{T}$ is equivalent to translating lines by $-\boldsymbol{T}$. Pure translation does not change line orientations, so the new line remains parallel to the original. Under perspective projection, parallel 3D lines of orientation

$\boldsymbol{U}$ project to converging image lines which intersect at the vanishing point associated with $\boldsymbol{U}$. Therefore the image of a line in one camera coordinate system intersects the image of the same line in a translated coordinate system at a vanishing point, from which the line direction can be derived.

Unfortunately, the input line data is always imperfect due to noise in its imaging and extraction, and the effects of noise on the computed 3D orientation must be taken into account. When working with real data, any computed quantity should be treated as an estimate only, with an associated measure of uncertainty. In this instance, a method is needed for computing the distribution of the orientation of the cross product of two random vectors. This problem is addressed in Section 2.2.3.

2.1.2 Recovering Planar Surfaces

Having first computed 3D line directions, it is possible to discover coplanar lines and thereby recover the orientation and distance of the planar surfaces that contain them. This is done in two stages. First the lines are broken into groups consistent with a family of parallel planes, then distances are finally computed to partition the lines into sets consistent with individual plane equations.

The normal vector to a planar surface is perpendicular to the orientations of all lines on that surface. Conversely, given a line of orientation $\boldsymbol{U}$ lying on a planar surface, the set of possible surface normals is the set of unit vectors perpendicular to $\boldsymbol{U}$. On the unit sphere, the heads of this set of possible normals trace out a great circle with polar axis $\boldsymbol{U}$. This geometric constraint leads to an efficient Hough transform technique for finding possible surface normals on the sphere, a transform first employed by Barnard in the context of locating potential vanishing points [3]. Each line orientation is mapped onto a 2D histogram representing the surface of the unit sphere, partitioned by azimuth and elevation. Each orientation casts a vote in all buckets along the great circle representing vectors perpendicular to it. Potential surface normals are detected as peaks in the histogram, corresponding to areas where several great circles intersect.

Once potential plane orientations have been identified, the unit normal can be estimated with more accuracy. Although the true line orientations lie on a great circle around the true plane normal, this relationship will not be exact due to errors in the derived line orientations. Instead, line orientations lie scattered in an equatorial belt. In the Section 2.2.2, an estimate of the polar axis of this belt is developed, providing a vector estimate plus an uncertainty region for the 3D surface orientation.

Finally, for lines within a family of parallel planes, a 1D histogram of plane distances is formed by computing for each line a hypothesized distance. The distance is computed as $d_i = \hat{\boldsymbol{n}} \cdot \boldsymbol{p}_i$, where $\boldsymbol{p}_i$ is a 3D point on line i and $\hat{\boldsymbol{n}}$ is the estimated plane normal. Peaks in this 1D histogram represent sets of lines consistent with a single plane equation.

Figure 1 shows an example of the partition created for a stereo hallway image. The algorithm forms hypotheses of all three visible wall planes, and correctly identifies that one plane orientation is shared by two parallel planes at different depths.

The above method for hypothesizing planes from stereo line correspondences has some unique features. The method typically employed to solve the same problem is to compute 3D line segments or points, then cluster them into planar patches. We believe our approach

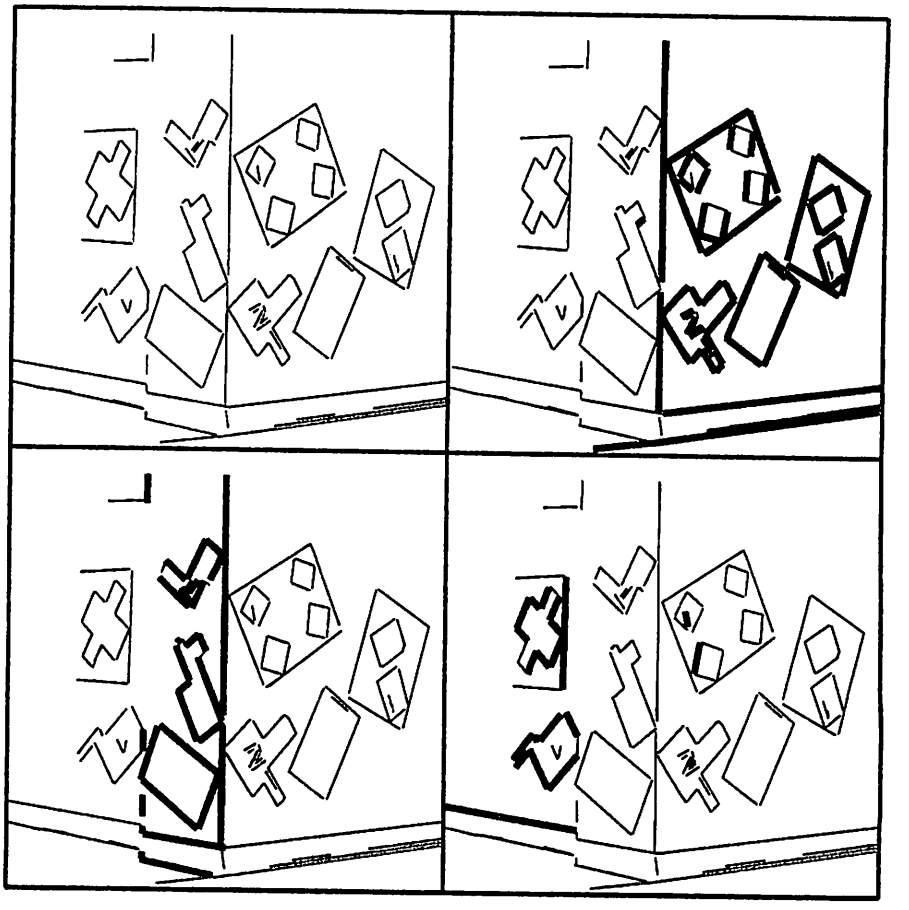


Figure 1: Matched input lines from the left image of a stereo pair are shown upper left; also shown are 3 sets of lines consistent with hypothesized surface planes

has the following advantages:

- The computation of plane distance is decoupled from computing plane orientation, thus the Hough transform for detecting surface normals is 2-dimensional instead of 3, and the entire process is roughly $\mathcal{O}(N)$ in the number of lines [6].
- Parallel planes are immediately identified, and their shared orientation is computed from all of the lines on them.
- When the depth to a point on a line is finally computed, the line is assumed to lie in a plane with a given orientation that has been estimated from several lines. The depth computed under this constraint is presumably more accurate (see [19]).

2.2 Estimates of Orientation

In this section, the statistical problem of representing and estimating orientation vectors and deriving confidence regions for the resulting estimates is examined. Two particular cases are explored: estimating the polar axis of a great circle of unit vectors, and estimating the cross product of two uncertain unit vectors.

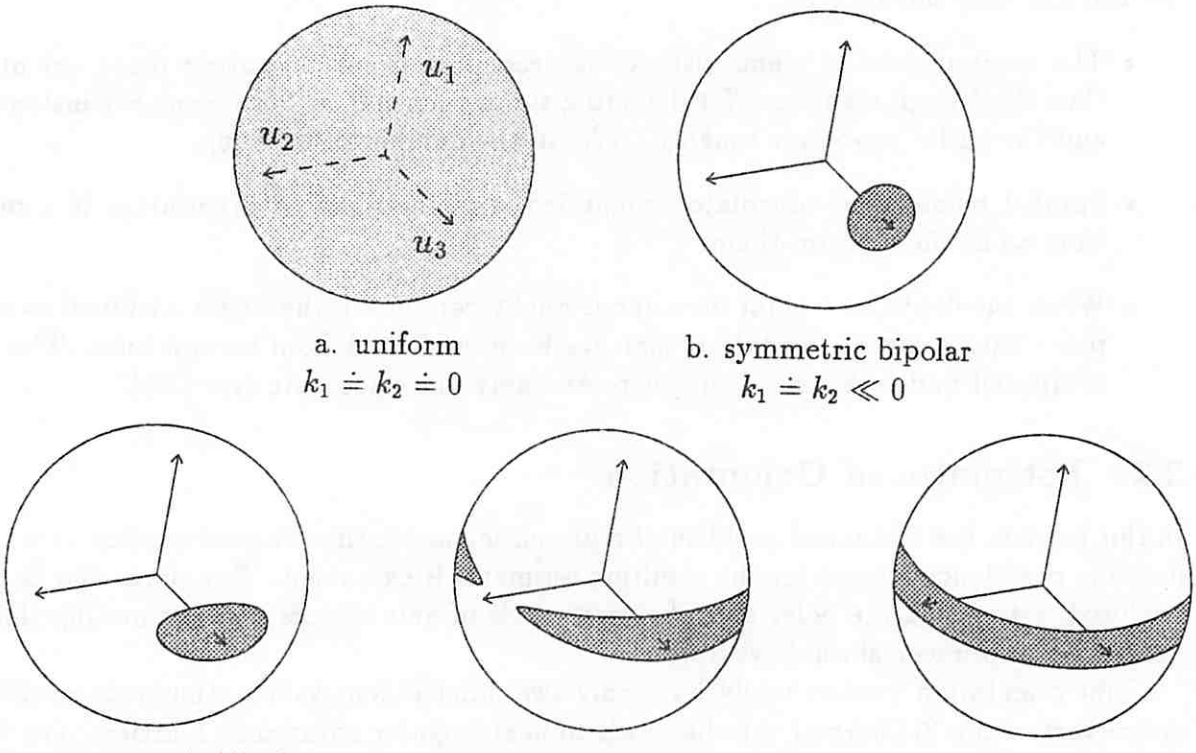
Unit orientation vectors really have only two independent values, thus representing the entire vector as a 3D normal variable leads to near-singular covariance matrices. One possibility is to represent unit vector uncertainties as bivariate normal vectors in a local coordinate system on the tangent plane or in spherical cordinates on the sphere surface. However, this approach leads to some messy coordinate system bookkeeping when trying to combine the uncertainties of two or more vectors. We choose instead to represent the uncertainty as a 3D normal with zero mean, constrained to lie on the surface of the sphere.

2.2.1 Bingham's Distribution

Bingham's distribution is a standard statistical distribution for representing both bipolar and equatorial clusters of points on the sphere [4, 15]. It describes a trivariate normal vector with zero mean and arbitrary covariance matrix, conditioned on the length of the vector being unity. Bingham's distribution thus represents the portion of a trivariate normal distribution $N_3(\boldsymbol{x}; 0, \boldsymbol{\Sigma})$ that intersects the surface of the unit sphere. The covariance matrix $\boldsymbol{\Sigma}$ can be written as $\boldsymbol{U}\boldsymbol{S}\boldsymbol{U}^t$, where $\boldsymbol{U} = [\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3]$ is a 3×3 orthogonal matrix and $\boldsymbol{S} = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \sigma_3^2)$ is a diagonal matrix of variances. Without loss of generality, assume that $\sigma_1^2 \leq \sigma_2^2 \leq \sigma_3^2$. Letting $\boldsymbol{K} = \mathrm{diag}(\kappa_1, \kappa_2, \kappa_3) = -\frac{1}{2}\boldsymbol{S}^{-1}$ and constraining $\boldsymbol{x}$ to have unit length yields the Bingham distribution

$$\begin{aligned} \mathcal{B}(\boldsymbol{x}; \boldsymbol{K}, \boldsymbol{U}) &= B(\boldsymbol{K}) \exp\{\boldsymbol{x}^t \boldsymbol{U} \boldsymbol{K} \boldsymbol{U}^t \boldsymbol{x}\} \\ &= B(\boldsymbol{K}) \exp\{\mathrm{tr}(\boldsymbol{K}\boldsymbol{U}^t \boldsymbol{x}\boldsymbol{x}^t \boldsymbol{U})\} \\ &= B(\boldsymbol{K}) \exp\{\kappa_1 (\boldsymbol{u}_1^t \boldsymbol{x})^2 + \kappa_2 (\boldsymbol{u}_2^t \boldsymbol{x})^2 + \kappa_3 (\boldsymbol{u}_3^t \boldsymbol{x})^2\}, \end{aligned} \tag{1}$$

where $B(\boldsymbol{K})$ is the normalizing constant required to make $\int_{\boldsymbol{x}^t\boldsymbol{x}=1} \mathcal{B}(\boldsymbol{x}; \boldsymbol{K}, \boldsymbol{U}) = 1$, and $\mathrm{tr}(\cdot)$ denotes the trace of a matrix.



a. uniform
$k_1 \doteq k_2 \doteq 0$

b. symmetric bipolar
$k_1 \doteq k_2 \ll 0$

c. asymmetric bipolar
$k_1 < k_2 \ll 0$

d. asymmetric girdle
$k_1 \ll k_2 < 0$

e. symmetric girdle
$k_1 \ll k_2 \doteq 0$

Figure 2: Bingham's Distribution – representative contours for varying shape parameter magnitudes

It is easy to verify that $\mathcal{B}(-\boldsymbol{x}; \boldsymbol{K}, \boldsymbol{U}) = \mathcal{B}(\boldsymbol{x}; \boldsymbol{K}, \boldsymbol{U})$, thus the distribution is antipodally symmetric, and it is appropriate to interpret $\boldsymbol{x}$ as an undirected axis, or as an observation confined to the unit hemisphere. Further analysis shows that the "shape" parameters κ_i are determined only up to an additive constant. For uniqueness it is customary to subtract out the largest κ_i, in this case κ_3, leaving $k_i = \kappa_i - \kappa_3$ with $k_1 \leq k_2 \leq k_3 = 0$. Figure 2 displays the Bingham distribution for varying values of k_1 and k_2.

Given a set of n unit vectors $\boldsymbol{\phi}_i = \langle x_i, y_i, z_i \rangle$ assumed to be distributed according to Bingham's distribution, a sufficient statistic for the orientation and shape parameters $\boldsymbol{U}$ and $\boldsymbol{K}$ is the sample second moment or *scatter* matrix

$$\boldsymbol{M} = \frac{1}{n} \sum_{i=1}^{n} (\boldsymbol{\phi}_i \boldsymbol{\phi}_i^t) = \frac{1}{n} \begin{bmatrix} \sum x_i^2 & \sum x_i y_i & \sum x_i z_i \\ \sum x_i y_i & \sum y_i^2 & \sum y_i z_i \\ \sum x_i z_i & \sum y_i z_i & \sum z_i^2 \end{bmatrix}. \tag{2}$$

Since the scatter matrix is a symmetric real matrix, it can be decomposed into $\boldsymbol{M} = \boldsymbol{A}\Lambda\boldsymbol{A}^t$, where $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3]$ is an orthogonal matrix of eigenvectors, and $\Lambda = \text{diag}(\lambda_1, \lambda_2, \lambda_3)$ is a diagonal matrix of corresponding eigenvalues with $\lambda_1 \leq \lambda_2 \leq \lambda_3$ summing up to 1. It can be shown [4, 15] that the maximum likelihood estimate of the Bingham

orientation matrix $\boldsymbol{U}$ is the matrix of eigenvectors $\boldsymbol{A}$. Maximum likelihood estimates of the shape parameters k_1 and k_2 are nontrivial functions of $\boldsymbol{\Lambda}$ [13].

Recall from Section 2.1.2 that a set of line orientation vectors perpendicular to the surface normal formed an equatorial belt around the surface normal direction. We will assume the distribution of line orientation vectors can be described as an asymmetric Bingham distribution, elongated along a great circle (Figure 2c-e). Finding the planar surface normal reduces then to estimating the polar axis of the distribution. If the convention $\lambda_1 \leq \lambda_2 \leq \lambda_3$ is adhered to, the polar axis of an equatorial Bingham distribution will be $\boldsymbol{u}_1$, the eigenvector associated with the smallest eigenvalue λ_1 of the second moment matrix.

After finding the maximum likelihood estimate $\hat{\boldsymbol{u}} = \hat{\boldsymbol{u}}_1 = \boldsymbol{a}_1$ of the desired 3D orientation vector, the construction of a confidence region can begin. Bingham suggests [4] that for large sample sizes, an approximate $1-\alpha$ confidence region for $\boldsymbol{u}_i$ is an ellipse centered at $\hat{\boldsymbol{u}}_i$, with axes of lengths $\{\chi^2_{2,\alpha}/(2\,n\,(k_i - k_j)(\lambda_i - \lambda_j))\}^{\frac{1}{2}}$ directed along great circles towards $\hat{\boldsymbol{u}}_j, j \neq i$, where $\chi^2_{2,\alpha}$ is the upper α critical point of the χ^2 distribution with 2 degrees of freedom. For small errors this is approximately the set of vectors

$$\left\{ \boldsymbol{v} : \ \boldsymbol{v}^t \, \boldsymbol{R} \, \boldsymbol{v} \ \leq \ \frac{\chi^2_{2,\alpha}}{2\,n} \right\}, \tag{3}$$

where $\boldsymbol{R}$ is the matrix

$$[\boldsymbol{a}_2, \boldsymbol{a}_3] \begin{bmatrix} (k_1 - k_2)(\lambda_1 - \lambda_2) & 0 \\ 0 & k_1(\lambda_1 - \lambda_3) \end{bmatrix} [\boldsymbol{a}_2, \boldsymbol{a}_3]^t.$$

2.2.2 Approximate Confidence Regions

In addition to decomposing the scatter matrix into principle components (eigenvalues and eigenvectors), construction of a Bingham confidence region requires computation of the shape parameters k_1 and k_2. This process is computationally expensive, however, due to the complexity of the distribution normalization constant [13]. To find a more convenient approximation, first note that a decomposition of the sample scatter into principle components is also found at the heart of the least-squares perpendicular error plane fit [2]. In particular, the eigenvector associated with the smallest eigenvalue of the sample scatter matrix is an estimate of both the pole of the equatorial Bingham distribution, and the normal of the least-squares best-fit plane.

To see this, let the heads of unit vectors $\boldsymbol{\phi}_i = \langle x_i, y_i, z_i \rangle$ be the data points for the plane fit. The normal vector to the desired plane will be the unit vector $\boldsymbol{v}$ which minimizes the sum of squared distances

$$\sum_{i=1}^{n} (\boldsymbol{v}^t \boldsymbol{\phi}_i)^2 \ = \ \sum_{i=1}^{n} \boldsymbol{v}^t \boldsymbol{\phi}_i \boldsymbol{\phi}_i^t \boldsymbol{v} \ = \ n \boldsymbol{v}^t \boldsymbol{M} \boldsymbol{v} \tag{4}$$

where $\boldsymbol{M}$ is the scatter matrix from equation (2). It is clear that the constant positive factor n can be ignored, that is, the vector $\boldsymbol{v}$ minimizing $\boldsymbol{v}^t \boldsymbol{M} \boldsymbol{v}$ also minimizes (4). Adding a Lagrange multiplier λ to include the constraint of unit length, the quantity to be minimized becomes

$$E = \boldsymbol{v}^t \boldsymbol{M} \boldsymbol{v} + \lambda(1 - \boldsymbol{v}^t \boldsymbol{v}).$$

After differentiating with respect to $\boldsymbol{v}$ and λ and setting the results to zero, we find that $\boldsymbol{M}\boldsymbol{v} = \lambda\boldsymbol{v}$, subject to $\boldsymbol{v}^t\boldsymbol{v} = 1$. The multiplier λ must be an eigenvalue of $\boldsymbol{M}$, and $\boldsymbol{v}$ its corresponding eigenvector. To minimize the quantity (4) we choose the unit vector $\boldsymbol{v}_{min}$ associated with the smallest eigenvalue λ_{min}, because the variance in the normal direction should be the smallest.

To develop an approximate confidence region, further note that the linear least squares plane fit $z = \alpha x + \beta y$ yields the same plane estimate as the perpendicular error plane fit when the estimated plane is perpendicular to the z axis, since then the distance parallel to the z-axis from a point to the plane is the orthogonal distance [1]. This means that if the original coordinate system is rotated such that the normal of the orthogonal plane fit coincides with the z-axis, a linear least squares plane fit will also yield the z-axis as the rotated plane normal vector. More importantly, the confidence region associated with this linear least-squares plane fit is an approximate rotated confidence region.

Without detailing the intermediate steps in building a confidence region for a linear least squares solution (see [5], for example), the following approximation to the Bingham confidence region is derived

$$\left\{\boldsymbol{v}: \quad \boldsymbol{v}^t\,\boldsymbol{P}\,\boldsymbol{v} \quad \leq \quad \frac{2\,F_\alpha(2, n-2)}{n-2}\right\}, \tag{5}$$

where $F(2, n-2)$ is an F variable with 2 and $n-2$ degrees of freedom, and $\boldsymbol{P}$ is the matrix

$$[\boldsymbol{a}_2, \boldsymbol{a}_3] \begin{bmatrix} \lambda_2/\lambda_1 & 0 \\ 0 & \lambda_3/\lambda_1 \end{bmatrix} [\boldsymbol{a}_2, \boldsymbol{a}_3]^t.$$

This region most closely approximates equation (3) for rotationally symmetric, low variance equatorial distributions, as in Figure 2e. Furthermore, approximating Bingham polar axis estimation as a least squares plane fit allows us to include information about non-equal variances among the input vectors simply by substituting the formula for a weighted least squares fit.

2.2.3 Cross Product Confidence

Recall that the 3D orientation of parallel lines in the world is computed from the cross product of the projection plane normals of their images. A cross product is a great circle pole estimate for two vectors. Unfortunately, the formula for estimating polar axis confidence regions is not valid for less than three vectors. In this section a formula to compute an approximate confidence region for the cross product is derived, based on intersecting equatorial Bingham distributions.

A geometrical intuition for computing an uncertain cross product is to intersect belts on the sphere, each belt representing all the points perpendicular to the uncertainty region at its pole [18]. In the current formalism, if two equatorial Bingham distributions exist, with distributions $P(\boldsymbol{x} \sim N(0, \boldsymbol{\Sigma}_1) \mid \boldsymbol{x}^t\boldsymbol{x} = 1)$ and $P(\boldsymbol{x} \sim N(0, \boldsymbol{\Sigma}_2) \mid \boldsymbol{x}^t\boldsymbol{x} = 1)$, the probability that a vector lies in the intersection of the two is their joint probability, which assuming independence is the product of the individual probabilities. It can be shown that this joint probability has the Bingham distribution $P(\boldsymbol{x} \sim N(0, (\boldsymbol{\Sigma}_1^{-1} + \boldsymbol{\Sigma}_2^{-1})^{-1}) \mid \boldsymbol{x}^t\boldsymbol{x} = 1)$.

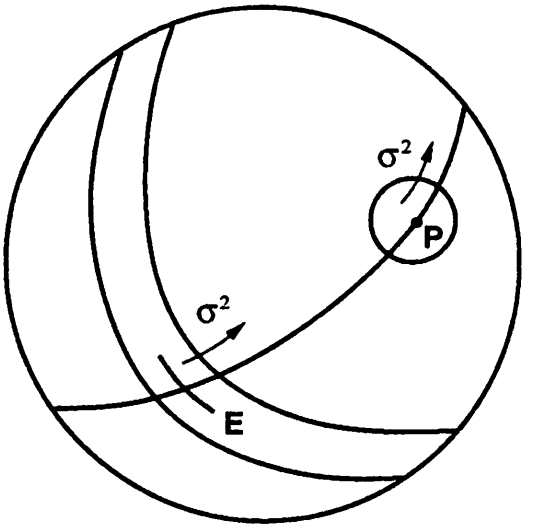


Figure 3: An equatorial Bingham distribution E is the "dual" of a bipolar one P if the variance along a line of longitude across the equator of E is the same as the variance along that line of longitude at the pole of the P.

The only remaining task is to construct a "dual" equatorial Bingham distribution from a bipolar one, where dual means that the longitudinal variance across the equator of the equatorial distribution is the same as the longitudinal variance at the pole of the bipolar distribution (see Figure 3). For general asymmetric bipolar distributions this is difficult, but an approximate dual for rotationally symmetric forms exhibiting small variances can be constructed. Under those conditions $k_1 = k_2 \doteq -1/(2\sigma^2)$, where σ^2 is the variance across the pole [4]. If the axis of the pole is $\boldsymbol{u}_3$, the matrix of shape parameters $\boldsymbol{K}$ becomes $\mathrm{diag}(0, 0, -1/(2\sigma^2))$, and $\boldsymbol{\Sigma}^{-1} = -2\boldsymbol{U}\boldsymbol{K}\boldsymbol{U}^t$. Decomposition into principle components of the sum of $\boldsymbol{\Sigma}_1^{-1}$ and $\boldsymbol{\Sigma}_2^{-1}$ yields the orientation and shape of the resulting joint Bingham.

2.3 Numerical Results

The stereo line correspondence example from the last section was taken with a two-camera setup with parallel focal axes and a baseline of 20 inches. The distance from the left camera to the corner was approximately 18 feet. As described in Section 2.1, each line orientation was computed as the cross product of two corresponding projection plane normals. The variance in each input normal was taken to be circular, and inversely proportional to the average length of the matched lines in the image. From these two uncertain projection plane normals, the line orientation was computed as their cross product, and an associated confidence region was computed as described in Section 2.2.3. Although this generally gave an elliptical confidence region, uncertainty in the line direction was taken to be a circular distribution circumscribing the elliptical one on the sphere, summarizing the region description with a single variance.

Two plane orientations were discovered from the Hough transform array (two planes of

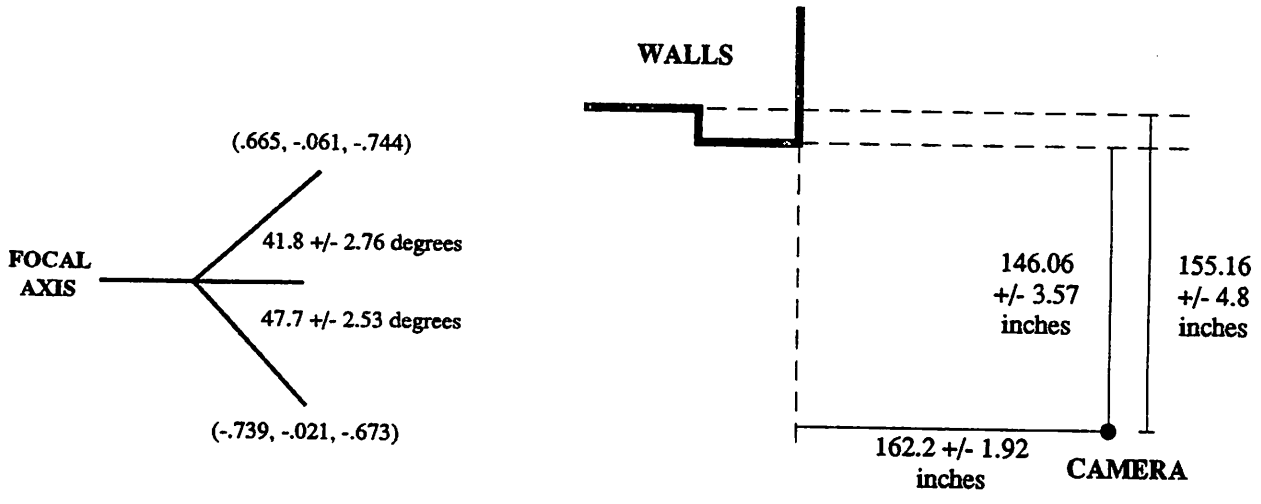


Figure 4: Numerical results for planar surface hypotheses, at the 95% confidence level. Left depicts the orientation and uncertainty in surface normals. Right shows perpendicular distance estimates to each plane.

the three share the same orientation). To compute plane orientations from each line group found, a weighted least-squares plane fit was applied to the line directions in the group, using a weight inversely proportional to the line direction variance. Approximate confidence regions were computed using equation 5.

The mean orientation vector for the right wall was computed to be $(.665, -.061, -.744)$, as depicted in Figure 4a. A 95% confidence region around the right wall orientation is approximately an ellipse, with half-lengths on the sphere of 2.48 and 2.76 degrees. The shared orientation of the two left walls was estimated to be $(-.739, -.021 - .673)$, with a 95% confidence region of half-lengths 2.53 and 2.0 degrees. The relative orientation between the two estimated orientations is 89.4 degrees, and the actual walls are in fact perpendicular within the usual limits of construction accuracy.

In partitioning the lines into individual plane equations, the two parallel left walls were separated. The perpendicular distance from the camera to the extended plane of the far left wall was computed as 155.16 ± 4.8 inches, at the 95% confidence level, and that of the closer left wall as 146.06 ± 3.57 inches (see Figure 4b). This gives a nominal difference of 9.1 inches, whereas the faces of these two walls are in fact 8.5 inches apart, or a 7% error in depth for the estimated value. A 95% confidence region for the depth of the wall on the right hand side of the image was estimated to be 162.2 ± 1.92 inches.

3 Vanishing Point Analysis

Under perspective projection, parallel lines in three-space project to converging lines in the image plane. The common point of intersection, perhaps at infinity, is called the *vanishing point*. Vanishing point analysis provides strong cues for inferring the 3D structure of a scene from only a single view [10]. Under known camera geometry, the vanishing point of a group

of parallel lines determines their orientation in three-space. Two or more vanishing points from lines known to lie in a single 3D plane establish a *vanishing line,* which completely determines the orientation of the plane.

In computer vision, the pure geometry of vanishing point analysis must be applied to the physical world. Linear structures in the world have finite extent and project to line segments in the image plane. The slopes of these segments are perturbed by noise both in the imaging process and in the method of line segment extraction. Finding the vanishing point from a set of lines thus becomes a problem of estimation.

In this section vanishing point computation is characterized as a statistical estimation problem on the unit sphere; in particular as the estimation of the polar axis of an equatorial distribution. This framework facilitates the construction of confidence regions for 3D line orientation.

3.1 Previous Work

A practical algorithm for finding vanishing points from a set of line segments in an image must address two issues: how to cluster line segments going to a single vanishing point, and how to estimate an accurate vanishing point from a given line cluster. The former is handled elegantly by Hough transform approachs [3, 12, 17]. In Barnard [3], line segments are mapped onto great circles in a histogram representing the surface of a unit sphere. Potential vanishing points are detected as peaks in the histogram, corresponding to areas where several great circles intersect.

While the complexity of Barnard's algorithm is $\mathcal{O}(n)$ in the number of lines processed, vanishing points are located only to within the boundaries of a histogram bucket (similar comments apply to Kender's involute aggregation transform [12]). This makes the approach ill-suited for accurately estimating the true vanishing point location. If a low variance estimate is required, the histogram must be partitioned very finely, negating some of the computational benefits of the approach. This drawback can be overcome by using a more sophisticated hierarchical apparatus, as in Quan and Mohr [17]. Yet, other problems inherent to the Hough transform remain [9], namely problems occurring when the true vanishing point falls near a histogram boundary, causing candidates which should be grouped together to fall into separate buckets. In summary, while algorithms based on the Hough transform excel at quickly clustering line segments into convergent groups, the final estimate of vanishing point location and variance should be based on the line segments themselves rather than the arbitrary bucket boundaries of a histogram data structure.

In contrast to the Hough transform approach, Magee and Aggarwal [14] compute intersections of pairs line segments directly, using cross product operations. Vanishing points are detected as clusters of intersection points on the sphere. Since computed points are maintained and distances compared uniformly over the sphere, vanishing points can be estimated with greater accuracy. The drawback is that examining all pairs of line segments yields a complexity of $\mathcal{O}(n^2)$.

The camera focal point and the endpoints $\boldsymbol{p}$ and $\boldsymbol{q}$ of an image line segment form a *projection plane* with unit normal $\boldsymbol{\phi} = (\boldsymbol{p} \times \boldsymbol{q}) / \|\boldsymbol{p} \times \boldsymbol{q}\|$. In [6] a least-squares perpendicular error plane is fit to the projection plane normals of line segments in convergent line clusters. The normal vector of a cluster's best-fit plane provides an estimate of the cluster's vanishing

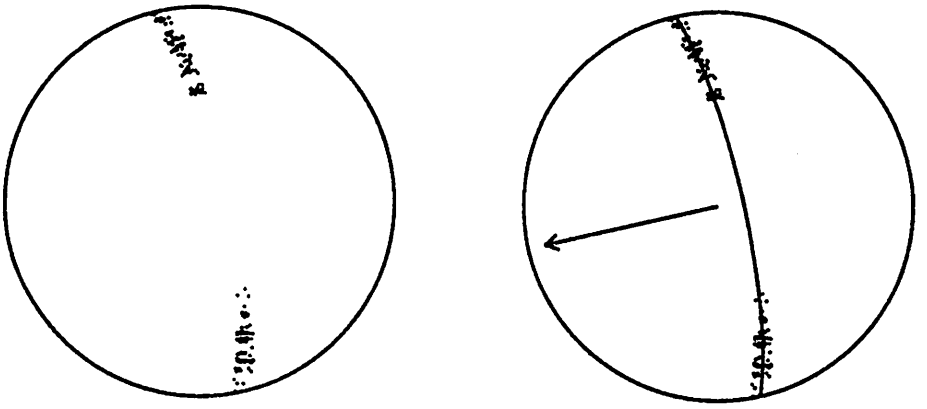


Figure 5: Projection plane normals of 3D parallel lines cluster around a great circle on the unit sphere. The normal of the best fitting plane is an estimate of the 3D line orientation.

point. This method combines the $\mathcal{O}(n)$ efficiency of the Hough transform approach with the accuracy of the cross product approach. In particular, the Hough transform is applied as an initial clustering method and as an efficient spatial access mechanism, while the final analysis of vanishing point location is performed on the underlying data.

In this section we take a fresh look at vanishing point calculation as a statistical estimation problem. It is assumed that image line segments have been previously clustered into groups of convergent lines. For each group, the vanishing point location is estimated as the polar axis of an equatorial distribution on the unit sphere, and the statistical error of the estimate is determined. Section 3.3 presents an investigation into the sensitivity of the estimates to the number of lines in a convergent cluster.

3.2 Vanishing Point Estimation

Consider a collection of 3D lines with unit orientation $\boldsymbol{u}$. Their projections in the image plane converge to a common vanishing point, perhaps at infinity. A vector from the origin pointing towards this vanishing point also has orientation $\boldsymbol{u}$. Since the vanishing point associated with a 3D line lies somewhere along the infinite extension of its projection, it follows that a vector pointing towards the vanishing point lies in the projection plane of the line, and is thus perpendicular to the projection plane normal. Conversely, the projection plane normals of parallel 3D line segments lie in a single plane passing through the origin perpendicular to orientation $\boldsymbol{u}$.

A unit vector can be represented as the point on a unit sphere where the vector head touches the sphere's surface. The heads of the above coplanar unit vectors lie on a great circle with polar axis $\boldsymbol{u}$. However, due to noise in the extracted image line segments, coplanarity of the projection plane normals will not be exactly satisfied. The corresponding points on the sphere tend to cluster about a great circle, forming an *equatorial* distribution (Figure 5). Computing the polar axis of this distribution, and hence line orientation $\boldsymbol{u}$, can be treated as a statistical estimation problem on the sphere.

Let X be a random variable over the surface of the unit sphere, and $f(X)$ its probability

density function. The set of unit projection plane normals forms a random sample $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ from $f(X)$. We are dealing with orientations, or unsigned unit vectors, so it is appropriate to restrict attention to a single hemisphere to avoid ambiguity. By convention we choose the unit hemisphere $z > 0$ facing towards the image plane, and any vector with a negative z component is negated so that it falls within the positive hemisphere. Rather than manipulate distributions defined on the positive hemisphere, with "wrap around" conditions at the equator $z = 0$, it is convenient to consider distributions with *antipodal symmetry*, i.e. with $f(-X) = f(X)$.

Vanishing point calculation can thus be characterized as estimating the true polar axis of an antipodally symmetric equatorial distribution given a random sample of size n. Posed in this way, the problem falls squarely within the domain of directional statistics [11, 15]. In this section two methods for estimating the pole of an equatorial distribution $f(X)$ are examined. The first is a parametric method based on the assumption that $f(X)$ is a Bingham distribution; the second is a nonparametric method for general antipodally symmetric distributions.

3.2.1 Case 1 : Bingham's Distribution

Bingham's distribution was described in Section 2.2.1. To restate the relevant results, a sufficient statistic for the orientation and shape parameters $\boldsymbol{U}$ and $\boldsymbol{K}$ is the sample second moment matrix

$$\boldsymbol{M} \;=\; \frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{x}_i\boldsymbol{x}_i^t) \;=\; \boldsymbol{A}\boldsymbol{\Lambda}\boldsymbol{A}^t, \tag{6}$$

factored into $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3]$, an orthogonal matrix of eigenvectors, and $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)$, a diagonal matrix of corresponding nondecreasing eigenvalues. The maximum likelihood estimate of the orientation matrix $\boldsymbol{U}$ is the matrix of eigenvectors $\boldsymbol{A}$. In particular, the maximum likelihood estimate of the polar axis of both the equatorial and asymmetric forms of Bingham's distribution is $\hat{\boldsymbol{u}} = \hat{\boldsymbol{u}}_1 = \boldsymbol{a}_1$. Maximum likelihood estimates of the nonzero shape parameters k_1 and k_2 are nontrivial functions of Λ.

For large sample sizes, an approximate $1 - \alpha$ confidence region for $\boldsymbol{u}_1$ is the set of vectors

$$\left\{\boldsymbol{v} :\;\; \boldsymbol{v}^t\,[\boldsymbol{a}_2,\, \boldsymbol{a}_3]\; D\; [\boldsymbol{a}_2,\, \boldsymbol{a}_3]^t\,\boldsymbol{v} \;\;\leq\;\; \frac{\chi^2_{2,\alpha}}{2n}\right\}, \tag{7}$$

where $D = \mathrm{diag}(\,(k_1 - k_2)(\lambda_1 - \lambda_2), k_1(\lambda_1 - \lambda_3)\,)$ and $\chi^2_{2,\alpha}$ is the upper α critical point of the χ^2 distribution with 2 degrees of freedom.

3.2.2 Case 2 : General Antipodal Symmetry

Prentice develops a nonparametric estimation procedure for antipodally symmetric distributions on the sphere [16]. His argument, based on the method of moments, is that the population second moment matrix $E(\boldsymbol{x}\boldsymbol{x}^t) = \boldsymbol{U}\boldsymbol{K}\boldsymbol{U}^t$ can be estimated by the sample second moment matrix (6). Equating matrix factors in the obvious way leads to estimators $\hat{\boldsymbol{u}}_i = \boldsymbol{a}_i$, so that again the desired polar axis estimate is $\boldsymbol{a}_1$.

Figure 6: Straight line segments from an outdoor building scene (upper left), and three clusters of converging lines found using Barnard's histogram method [3].

Variances of the computed parameters are estimated from the quantities

$$c_{jk} = \frac{1}{n} \sum_{i=1}^{n} (\hat{\boldsymbol{u}}_j^t \boldsymbol{x}_i)^2 (\hat{\boldsymbol{u}}_k^t \boldsymbol{x}_i)^2, \quad 1 \le i, j \le 3$$

which are the nonzero sample fourth moments in the coordinate system of the principle axes. Again invoking the method of moments, and assuming large sample sizes, an approximate confidence region for $\hat{\boldsymbol{u}}_1 = \boldsymbol{a}_1$ is found to be

$$\left\{ \boldsymbol{v} : \ \boldsymbol{v}^t \, [\boldsymbol{a}_2, \boldsymbol{a}_3] \, D \, [\boldsymbol{a}_2, \boldsymbol{a}_3]^t \, \boldsymbol{v} \ \le \ \frac{\chi^2_{2,\alpha}}{n} \right\}, \tag{8}$$

where $D = \text{diag}((\lambda_1 - \lambda_2)^2/c_{12}, (\lambda_1 - \lambda_3)^2/c_{13})$.

3.2.3 A Numerical Example

The upper left quadrant of Figure 6 shows a set of straight line segments extracted from an outdoor building scene. This 512 × 512 image was taken using a camera with a field of view of roughly 30 degrees. Also shown are three clusters of converging line segments found using the histogramming technique of Barnard [3]. The cluster in the upper right quadrant contains 536 line segments. Its sample second moment matrix is (all quantities rounded to four decimal places)

$$\boldsymbol{M} = \frac{1}{536} \begin{vmatrix} 521.8195 & -19.7644 & 2.3812 \\ -19.7644 & 2.1703 & -3.4969 \\ 2.3812 & -3.4969 & 12.0102 \end{vmatrix},$$

which yields eigenvector and eigenvalue matrices

$$\boldsymbol{A} = \begin{vmatrix} 0.0305 & -0.0154 & 0.9993 \\ 0.9588 & -0.2817 & -0.0380 \\ 0.2820 & 0.9594 & 0.0049 \end{vmatrix}$$

$$\boldsymbol{\Lambda} = \frac{1}{536} \begin{vmatrix} 0.4189 & 0 & 0 \\ 0 & 12.9985 & 0 \\ 0 & 0 & 522.5825 \end{vmatrix},$$

ordered so that the eigenvalues appear in increasing magnitude. Both methods yield $\boldsymbol{a}_1$, the first column of $\boldsymbol{A}$, as the estimate of the polar axis. Unit vector (0.0305, 0.9588, 0.2820) is therefore the estimate for the direction $\boldsymbol{u}$ of the original 3D line segments.

Under the Bingham assumption, eigenvalues $\lambda_1 = 7.816E^{-4}$, $\lambda_2 = 0.0243$, and $\lambda_3 = 0.975$ yield maximum likelihood estimates $k_1 = -640.23$ and $k_2 = -21.16$ (see [13]). An approximate 95% confidence region (7) for $\boldsymbol{u}$ is thus the ellipse centered at $\boldsymbol{a}_1$, with axes of half length

$$b_{12} = \sqrt{\frac{5.991}{15607.822}}\text{rad} = 1.12\text{deg} \quad b_{13} = \sqrt{\frac{5.991}{668631.96}}\text{rad} = 0.17\text{deg},$$

directed along great circles towards axes $\boldsymbol{a}_2$ and $\boldsymbol{a}_3$ respectively.

To compute an approximate confidence region assuming only general antipodal symmetry, the sample fourth moment matrix $\boldsymbol{y}_i = \boldsymbol{A}^t\boldsymbol{x}_i$ is needed. This turns out to be

$$\boldsymbol{C} = \begin{vmatrix} 1.5048E^{-6} & 2.327E^{-5} & 7.5682E^{-4} \\ 2.327E^{-5} & 1.0098E^{-3} & 2.3218E^{-2} \\ 7.5682E^{-4} & 2.3218E^{-2} & 9.5099E^{-1} \end{vmatrix}.$$

The confidence region (8) is located and oriented the same as in the Bingham case, and has axes of half length

$$p_{12} = \sqrt{\frac{5.991}{12740.426}}\text{rad} = 1.24\text{deg} \quad p_{13} = \sqrt{\frac{5.991}{672178.85}}\text{rad} = 0.17\text{deg}.$$

quite similar to the Bingham half lengths.

Line orientation estimates and associated confidence regions were similarly computed for the other two convergent line clusters in Figure 6. For the cluster in the lower left quadrant of the figure, the estimated line orientation vector was (0.775, -0.1762, 0.6069). An approximate 95% confidence ellipse on the sphere for the Bingham assumption has major and minor axis half lengths of 1.35 deg and 0.23 deg. The Prentice nonparametric method yields corresponding half lengths of 1.41 deg and 0.23 deg. For the cluster in the lower right quadrant an estimated 3D line orientation of (-0.6, -0.234, 0.765) is computed. The Bingham 95% confidence ellipse has half lengths of 1.25 deg and 0.2 deg, while the corresponding Prentice half lengths are 1.14 deg and 0.2 deg.

number of lines	expected number of members					
	900	925	950	975	990	995
100	827	862	900	938	969	986
80	845	876	917	953	970	985
60	856	885	914	952	981	990
40	874	904	932	955	979	987
20	835	863	899	934	971	982
10	797	837	878	906	937	959
5	650	693	737	783	832	862

Table 1: Bingham Confidence Region Counts

3.3 Discussion

We have shown that the projection plane normals of converging line segments in the image plane can be thought of as a random sample from an equatorial distribution on the unit sphere. The method of maximum likelihood under the Bingham assumption, and the method of moments under the assumption of antipodal symmetry, both yield the eigenvector associated with the smallest eigenvalue of the sample second moment matrix as an estimator of the polar axis. It is interesting to note that this is the same result produced by a least-squares perpendicular error plane fit to the heads of the projection plane normals, corroborating the intuition behind [6], and showing this vector to be a natural estimator of 3D line orientation.

Both confidence region equations (7) and (8) are based on asymptotic distributions of the eigenvectors of the second moment matrix, and are thus only valid for large sample sizes. How do these constructions behave under small sample sizes? As an initial step toward answering this question, both confidence region construction methods were tested on simulated data. Line orientation estimates and confidence regions were computed for linesets of size $n = 5$, 10, 20, 40, 60, 80 and 100. One thousand trials were run for each size n. For each trial, the unit hemisphere was sampled uniformly to chose a random 3D line orientation, and a sample of n lines was generated in a 512×512 image, assuming a camera with a 45 deg field of view. For each line, a uniform random point in the image was chosen, and a length between 5 and 50 pixels long was selected. A line of the selected length was placed with its center on the chosen image point, and oriented towards the vanishing point defined by the given 3D line orientation. After placement, each endpoint of the line segment was perturbed independently by a bivariate normal noise process centered at the endpoint, having standard deviations of 1.5 pixels along the line segment, and 0.5 pixels perpendicular to the line.

For each of the 1000 trials of size n, a 3D line orientation was estimated from the sample line segments, and both Bingham and Prentice confidence regions were constructed for 6 χ^2 values, corresponding to 90%, 92.5%, 95%, 97.5%, 99% and 99.5% levels of confidence. The original 3D line orientation was then tested for inclusion in each of these regions. The resulting counts are an indication of the actual level of accuracy of each confidence region, for varying sample sizes (Tables 1 and 2).

From these tables, the derived confidence regions appear to be on the small side; the true orientation was contained in a confidence region fewer times than expected, almost uniformly

number of lines	expected number of members					
	900	925	950	975	990	995
100	817	852	896	932	974	987
80	840	874	910	935	982	988
60	845	886	913	952	984	992
40	876	907	930	957	981	988
20	826	860	896	942	977	987
10	765	808	854	893	931	942
5	608	639	681	734	785	811

Table 2: Prentice Confidence Region Counts

across all confidence levels and lineset sizes. This suggests that the simulated sample distributions did not exactly conform to either distributional assumption. Furthermore, accuracy of the confidence regions in the Prentice table do not differ markedly from those in the Bingham table, a feature already noted in the numerical example of the last section. Since the maximum likelihood estimation of the Bingham shape parameters k_1 and k_2 is computationally nontrivial, and the distributional assumption more restrictive, the nonparametric alternative of Prentice is an attractive approximation.

Finally, the accuracy levels of all columns stays roughly the same as the number of lines in the sample drops from 100 down to 20. Below 20 the accuracy degrades, dropping notably for small sample sizes of 5. This suggests that the asymptotic arguments involved in the construction of the confidence regions hold for clusters of at least 20 lines.

References

[1] L. Ammann and J. Van Ness, "A Routine for Converting Regression Algorithms into Corresponding Orthogonal Regression Algorithms," *ACM Transactions on Mathematical Software,* vol. 14, no. 1, March 1988, pp. 76–87.

[2] D.H. Ballard and C.M. Brown, *Computer Vision,* Prentice-Hall, Inc, New Jersey, 1982, pp. 486–487.

[3] S.T. Barnard, "Interpreting Perspective Images," *AI Journal,* Vol. 21, No. 4, November 1983, pp. 435-462.

[4] C. Bingham, "An Antipodally Symmetric Distribution on the Sphere," *The Annals of Statistics,* Vol. 2, No. 6, 1974, pp. 1201–1225.

[5] G.E.P. Box and N.R.Draper, *Empirical Model-Building and Response Surfaces,* John Wiley and Sons, New York, 1987.

[6] R.T. Collins and R.S. Weiss, "An Efficient and Accurate Method for Computing Vanishing Points," Topical Meeting on Image Understanding and Machine Vision, 1989 Technical Digest Series, Vol. 14, (Optical Society of America, Washington, D.C.) pp. 92–94.

[7] R.T. Collins and R.S. Weiss, "Deriving Line and Surface Orientation by Statistical Methods," *Proceedings Darpa I.U. Workshop,* Pittsburgh, PA., September 1990, pp. 433–438.

[8] R.T. Collins and R.S. Weiss, "Vanishing Point Calculation as a Statistical Inference on the Unit Sphere," to appear in *Proceedings International Conference on Computer Vision,* Osaka, Japan, December 1990.

[9] W.E.L. Grimson and D.P. Huttenlocher, "On the Sensitivity of the Hough Transform for Object Recognition," *Proc. of the International Conference on Computer Vision,* 1988, pp. 700–706.

[10] R.M. Haralick, "Using Perspective Transformations in Scene Analysis," *Computer Graphics and Image Processing,* Vol. 13, 1980, pp. 191–221.

[11] P.E. Jupp and K.V. Mardia, "A Unified View of the Theory of Directional Statistics, 1975–1988," *International Statistical Review,* Vol. 57, 1989, pp. 261–294.

[12] J.R. Kender, "Shape From Texture: An Aggregation Transform that Maps a Class of Textures into Surface Orientation," Proceedings IJCAI 6, Tokyo, Japan, August 20-23, 1979, pp. 475-480.

[13] J.T. Kent, "Asymptotic Expansions for the Bingham Distribution," *Applied Statistics,* Vol. 36, No. 2, 1987, pp. 139–144.

[14] M.J. Magee and J.K. Aggarwal, "Determining Vanishing Points from Perspective Images," *CVGIP,* vol. 26, 1984, pp. 256-267.

[15] K.V. Mardia, *Statistics of Directional Data,* Academic Press, New York, 1972.

[16] M.J. Prentice, "A Distribution-Free Method of Interval Estimation for Unsigned Directional Data," *Biometrika,* Vol. 71, 1984, pp. 147–154.

[17] L. Quan, "Determining Perspective Structures Using Hierarchical Hough Transform," *Pattern Recognition Letters,* Vol. 9, 1989, pp. 279–286.

[18] R.S. Weiss, H. Nakatani, and E.M. Riseman, "An Error Analysis for Surface Orientation from Vanishing Points," to appear in *Pattern Analysis and Machine Intelligence.*

[19] L.B. Wolff and T.E. Boult, "Using Line Correspondence Stereo to Measure Surface Orientation," *Proceedings Eleventh International Joint Conference on Artificial Intelligence,* Detroit, Michigan, August 1989, pp. 1655–1660.