# Vanishing Point Calculation as a Statistical Inference on the Unit Sphere

**Robert T. Collins and Richard S. Weiss**

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

*In this paper 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.*

## 1 Introduction

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 [6]. 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.

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 approaches [1, 8, 13]. In Barnard [1], 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 [8]). 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 [13]. Yet, other problems inherent to the Hough transform remain [5], 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 [10] 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 [3] 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 point. This method combines the $\mathcal{O}(n)$ efficiency of the Hough transform approach with the accuracy of the cross product approach.

In this paper 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 presents an investigation into the sensitivity of the estimates to the number of lines in a convergent cluster. See [4] for a companion pa-

*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.

CH2934-8/90/0000/0400$01.00 © 1990 IEEE

per describing an application of directional statistics to another computer vision problem, determining 3D line and plane orientations from stereo line correspondences.

## 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. 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 [7, 11]. 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.

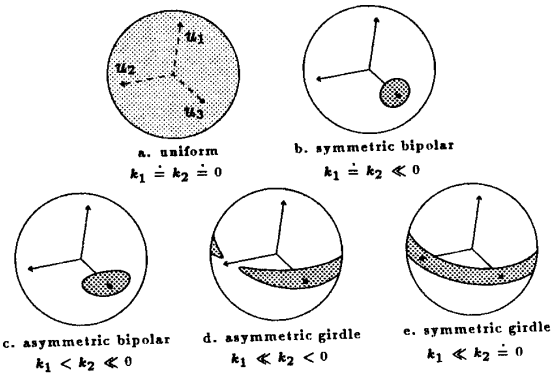



Figure 1: Bingham's Distribution – representative contours for varying shape parameter magnitudes

### 2.1 Case 1 : Bingham's Distribution

Bingham's distribution is a standard statistical distribution for representing both bipolar and equatorial clusters of points on the sphere [2, 11]. Bingham's distribution describes the intersection of a trivariate normal distribution $N_3(\boldsymbol{x}; 0, \boldsymbol{\Sigma})$ with the surface of the unit sphere. The covariance matrix $\boldsymbol{\Sigma}$ can be decomposed into $\boldsymbol{U}\boldsymbol{S}\boldsymbol{U}^t$, where $\boldsymbol{U} = [\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3]$ is a $3 \times 3$ 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

$$\mathcal{B}(\boldsymbol{x}; \boldsymbol{K}, \boldsymbol{U}) = B(\boldsymbol{K}) \exp\{\mathrm{tr}(\boldsymbol{K}\boldsymbol{U}^t\boldsymbol{x}\boldsymbol{x}^t\boldsymbol{U})\},$$

where $B(\boldsymbol{K})$ is a normalizing constant and $\mathrm{tr}(\cdot)$ denotes the trace of a matrix.

It is easy to verify that this distribution is antipodally symmetric. Further analysis shows that the "shape" parameters $\kappa_i$ are determined only up to an additive constant. For uniqueness it is customary to subtract out the largest $\kappa_i$, yielding $k_i = \kappa_i - \kappa_3$, $k_1 \leq k_2 \leq k_3 = 0$. Figure 1 displays Bingham's distribution for varying values of $k_1$ and $k_2$.

From Figure 1d-e we see that the polar axis of an equatorial Bingham distribution is $\boldsymbol{u}_1$. A sufficient statistic for the orientation and shape parameters $\boldsymbol{U}$ and $\boldsymbol{K}$ is the sample second moment matrix

$$\boldsymbol{T} = \frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{x}_i\boldsymbol{x}_i^t) = \boldsymbol{A}\boldsymbol{\Lambda}\boldsymbol{A}^t, \qquad (1)$$

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 eigenvalues with $\lambda_1 \leq \lambda_2 \leq \lambda_3$ summing up to 1. It can be shown [2, 11] that the maximum likelihood estimate of the orientation matrix $\boldsymbol{U}$ is the matrix of eigenvectors $\boldsymbol{A}$. Maximum



Figure 2: Straight line segments from an outdoor building scene and a cluster of converging lines.

likelihood estimates of the nonzero shape parameters $k_1$ and $k_2$ are nontrivial functions of $\Lambda$ [9].

The polar axis of this equatorial distribution, and hence the desired 3D orientation vector, has maximum likelihood estimate $\hat{u} = \hat{u}_1 = a_1$. Bingham suggests [2] that for large sample sizes, an approximate $1-\alpha$ confidence region for $u_1$ is the set of vectors

$$\left\{ v : \; v^t \, [a_2, \, a_3] \, D \, [a_2, \, a_3]^t \, v \;\; \leq \;\; \frac{\chi^2_{2,\alpha}}{2\,n} \right\}, \quad (2)$$

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 $\alpha$ critical point of the $\chi^2$ distribution with 2 degrees of freedom.

### 2.2 Case 2 : General Antipodal Symmetry

Prentice develops a nonparametric estimation procedure for antipodally symmetric distributions on the sphere [12]. His argument, based on the method of moments, is that the population second moment matrix $E(xx^t) = UKU^t$ can be estimated by the sample second moment matrix (1). Equating matrix factors in the obvious way leads to estimators $\hat{u}_i = a_i$, so that again the desired polar axis estimate is $a_1$.

Variances of the computed parameters are estimated from the quantities

$$c_{jk} \;=\; \frac{1}{n} \sum_{i=1}^{n} (\hat{u}_j^t \, x_i)^2 \, (\hat{u}_k^t \, x_i)^2, \qquad 1 \leq i, j \leq 3$$

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 $\hat{u}_1 = a_1$ is found to be

$$\left\{ v : \; v^t \, [a_2, \, a_3] \, D \, [a_2, \, a_3]^t \, v \;\; \leq \;\; \frac{\chi^2_{2,\alpha}}{n} \right\}, \quad (3)$$

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

### 2.3 A Numerical Example

The left half of Figure 2 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. The right half of the figure shows one cluster of converging line segments found using the histogramming technique of Barnard [1]. The cluster contains 536 line segments. Its sample second moment matrix is (all quantities rounded to four decimal places)

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

which yields eigenvector and eigenvalue matrices

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

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

ordered so that the eigenvalues appear in increasing magnitude. Both methods yield $a_1$, the first column of $A$, as the estimate of the polar axis. Unit vector (0.0305, 0.9588, 0.2820) is therefore the estimate for the direction $u$ 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 [9]). An approximate 95% confidence region (2) for $u$ is thus the ellipse centered at $a_1$, with axes of half length $b_{12} = \sqrt{5.991/15607.822}\,\mathrm{rad} = 1.12\,\mathrm{deg}$ and $b_{13} = \sqrt{5.991/668631.96}\,\mathrm{rad} = 0.17\,\mathrm{deg}$, directed along great circles towards axes $a_2$ and $a_3$ respectively.

To compute an approximate confidence region assuming only general antipodal symmetry, the sample fourth moment matrix $y_i = A^t x_i$ is needed. This turns out to be

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

The confidence region (3) is located and oriented the same as in the Bingham case, and has axes of half length $p_{12} = \sqrt{5.991/12740.426}\,\mathrm{rad} = 1.24\,\mathrm{deg}$ and $p_{13} = \sqrt{5.991/672178.85}\,\mathrm{rad} = 0.17\,\mathrm{deg}$, quite similar to the Bingham half lengths.

## 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 eigenvalue associated with the smallest eigenvector 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 [3], and showing this vector to be a natural estimator of 3D line orientation.

Both confidence region equations (2) and (3) are based on asymptotic distributions of the eigenvectors of the second moment matrix, and are thus only valid

| 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

| 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

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 $512 \times 512$ 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 $\chi^2$ 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 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] S.T. Barnard, "Interpreting Perspective Images," *AI Journal*, Vol. 21, No. 4, November 1983, pp. 435-462.

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

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

[4] R.T. Collins and R.S. Weiss, "Deriving Line and Surface Orientation by Statistical Methods," to appear in *Proc. Darpa I.U. Workshop*, September, 1990.

[5] 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.

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

[7] 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.

[8] 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.

[9] J.T. Kent, "Asymptotic Expansions for the Bingham Distribution," *Appl. Stats.*, Vol. 36, 1987, pp. 139–144.

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

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

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

[13] L. Quan and R. Mohr, "Determining Perspective Structures Using Hierarchical Hough Transform," *P.R. Letters*, Vol. 9, 1989, pp. 279–286.