

Single Plane Model Extension using Projective Transformations and Data Fusion

Robert T. Collins

Computer Science Department,
University of Massachusetts, Amherst, MA. 01003

Given at least four reference points or lines on a planar surface and their correspondences in an image, the relative positions of other structures on the plane can be derived without solving for camera position or intrinsic calibration parameters. A new framework for data fusion in the projective plane is presented to merge the position estimates of coplanar points and lines derived in this way.

1 Introduction

Acquiring 3D models of the environment is an important current research problem in computer vision. Since modeling the world in all its complexity is a daunting task, many researchers have focused on man-made domains where planar surfaces and linear surface markings predominate. The relevant geometric entities – points, lines and planes – are easily represented as linear subspaces. Furthermore, a rich set of results from the field of projective geometry become available.

The relevance of projective geometry to the visual acquisition of planar surface models cannot be overstressed: *Projective geometry provides a mathematical foundation for characterizing and representing the relationships between linear subspaces that remain invariant under the imaging process.*

This paper describes an approach to model extension using properties of projective mappings between planes. In particular, a priori knowledge of the relative positions of four or more coplanar points or lines is used to derive the positions of other points and lines on the same plane in a manner invariant to camera location and intrinsic camera parameters. Model extension is just one application of a general framework being developed for geometric inference in projective space. One of the main contributions of this work is the development of an appropriate methodology for fusing geometric information in the projective plane. An expanded version of this paper appears in [3].

*This work was funded by DARPA and TACOM under contract number DAAE07-91-C-R035 and by the National Science Foundation under grant number CDA-8922572.

2 Projective Transformations

This section briefly summarizes properties of projective mappings between planes. For a more comprehensive discussion the reader is invited to consult a projective geometry text such as [8].

Using homogeneous coordinates, a general projective transformation between planes can be written as

$$k [x', y', s'] = [x, y, s] H^t$$

where k is a nonzero scalar, s and s' are 1 for finite points in the plane and 0 for infinite points (see below), and H is a nonsingular 3×3 matrix of transformation parameters. Since homogeneous coordinates are equivalent up to scalar multiples, the transformation matrix can be multiplied by any nonzero constant and still represent the same mapping, and therefore has only 8 independent parameters. A nonsingular projective mapping that is linear in homogeneous coordinates is called a *homography*.

Because they are linear, invertible and closed under composition, homographies greatly simplify the analysis of projective mappings. However, in order to make a homography bijective a line of points at infinity must be explicitly added to each plane to correspond to the cases where s and s' go to zero. A plane that has been augmented in this way is a new geometric entity called the *projective plane*. The projective plane has a different global topology than the Euclidean plane, and this has implications for the representation of observed points and their uncertainty. This topic is explored in Section 4.

The fundamental theorem of projective geometry states that a plane to plane homography is completely determined by the correspondences of 4 coplanar points or lines in general position. In practice it is better to use as many point and line correspondences as possible to reduce errors in the estimated transformation caused by noise in the observed image data. Faugeras and Lustman present a least squares approach to estimating a homography by solving an overconstrained linear system of equations[6].

3 Planar Surface Model Extension

This section outlines an approach to planar model extension that assumes the positions of four or more points or lines on the object plane are already known. The transformation mapping known object features to their corresponding images is a homography containing information about the camera location and imaging parameters. The parameters of this homography are estimated from the correspondences of at least four known object features and their image projections. By inverting the homography other image features can be backprojected onto the object plane, allowing the positions of new points and lines on the plane to be determined without having to solve for camera location or intrinsic calibration parameters. Good experimental results have been achieved using this method, as reported in [3].

This approach is similar to one used by Mohr [7], who locates object points using pairs of cross ratios between the point and four known object locations. When exactly four point correspondences are used, the homography estimated using the least squares approach reduces to the mapping affected by Mohr's cross ratio pair; when more correspondences are known the least squares method should be more accurate. Furthermore, a new approach is now presented that allows positions estimated from several images to be merged to derive more accurate point and line positions from noisy observations.

4 Merging Geometric Information

A method for merging geometric information derived from multiple views is proposed in this section, based on fusing data points in the projective plane. From each viewpoint a homography is estimated that backprojects image points onto the object plane, thereby providing an estimate for the location of each object point. Over multiple images, multiple location estimates are obtained. Each point location estimate in homogeneous coordinates represents a point in the projective plane; multiple location estimates for each object point form a sample of points in the projective plane, clustered around the point in the projective plane representing the homogeneous coordinates of the true object point location. This section describes a method for estimating the true point position from its sample cluster.

In Section 2 the projective plane is described as the Euclidean plane augmented with a line of points at infinity. This is not the best way to visualize the projective plane, however, since the Euclidean plane is topologically open, while the projective plane is topologically closed. To see why, consider a hypothetical traveler following a ray starting at the origin and continuing out infinitely far. After "arriving" at infinity,

the traveler is located at some point $(x, y, 0)$ in homogeneous coordinates. But in homogeneous coordinates this is the same point as $(-x, -y, 0)$, so the explorer can keep traveling "past infinity" and eventually return to the origin still facing in the same direction.

As a result of this wraparound effect, if the topology of the projective plane is ignored and it is treated as a Euclidean plane a single cluster of points centered around a point at infinity will appear as two clusters infinitely far apart. Any estimation technique based on "averaging" these points using a Gaussian distribution in the plane will produce bad results, because the unimodal Gaussian distribution is a terrible approximation to the underlying bimodal distribution. Proper handling of points at infinity is not just of theoretical interest. Such points do arise in practice [4].

Since the projective plane is topologically closed, it is better to think of it as a closed 2D space like the surface of a sphere. More formally, consider $R^3 - 0$ (three space with the origin removed), and define an equivalence relation $(x_1, x_2, x_3) \sim (ky_1, ky_2, ky_3)$ for nonzero k . The projective plane P^2 is then defined as the quotient space $(R^3 - 0) / \sim$. Viewing R^3 geometrically as Euclidean 3-space, each member of the quotient space is an equivalence class of points along an infinite line through the origin (excluding the origin itself). Consider now the surface of the unit sphere $S^2 = \{(x_1, x_2, x_3) | x_1^2 + x_2^2 + x_3^2 = 1\}$, and form the quotient space S^2 / \sim . Each equivalence class now contains one pair of diametrically opposite points. Equating these equivalence classes with those of the projective plane in the obvious way shows that the surface of the unit sphere with antipodal points equated is isomorphic to the projective plane.

The most important benefit to come from this isomorphism is that it allows probability distributions on the sphere to be reinterpreted as distributions in the projective plane. Since diametrically opposite points on the sphere must be treated as equivalent in order to represent the projective plane, an appropriate distribution must possess the property of antipodal symmetry, i.e. the probability value at any point on the sphere must be the same as the probability at the diametrically opposite point.

A useful characterization of distributions on the sphere is presented in Beran [1]. Beran considers exponential distributions on the sphere, that is, distributions of the form $\exp\{P\}$ where P is a polynomial evaluated over the surface of the sphere. This is not as restrictive as it seems, since any strictly positive function F on the sphere can be represented as $\exp\{\ln\{F\}\}$. Exponential forms are considered due to their ease of use in maximum likelihood estimation.

Assuming a distribution of the form $\exp\{P\}$, the polynomial P can be decomposed using spherical harmonics, analogous to the way polynomials in Euclidean

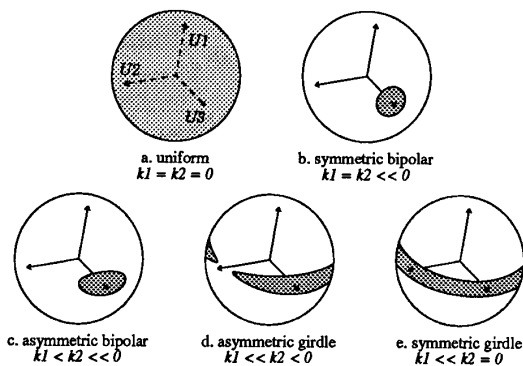


Figure 1: Bingham's Distribution Contours.

space are decomposed using Fourier analysis. If the distribution is required to have antipodal symmetry, all odd order harmonics are identically zero. This leaves an expression $\exp\{Y_0 + Y_2 + Y_4 + \dots\}$. The zeroth harmonic is a constant, so the $\exp\{Y_0\}$ term can be factored out and absorbed into the distribution's normalization constant. Therefore, the low order approximation to any antipodally symmetric exponential distribution on the sphere is of the form $\exp\{Y_2\}$. A distribution having this form has already been studied in the statistical literature, where it is called Bingham's distribution [2].

Bingham's distribution can be described as a trivariate Gaussian 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 Gaussian distribution that intersects the surface of the unit sphere, with varying ellipsoidal shapes of the underlying Gaussian contours producing a variety of distributional forms on the sphere (see Figure 1). Bingham's distribution has been used previously in a computer vision setting to represent uncertainties in line and plane orientations estimated from vanishing point analysis and stereo line correspondences [4, 5].

A point or line location estimate in homogeneous coordinates represents a point in the projective plane; multiple estimates form a sample of data points in the plane. To fuse data points in the projective plane each point is assumed to be a noisy observation of the true point location. The previous analysis shows that Bingham's distribution is an approximation to any noise process in the projective plane, therefore it is assumed that observed points are corrupted by a Bingham noise process centered about the true point location.

Normalizing the homogeneous coordinates of each sample point yields an antipodal pair of points on the unit sphere. Assuming a manageable level of noise,

the normalized sample points form a cluster on the sphere distributed according to a bipolar Bingham distribution (Figures 1b and 1c). An estimate of the homogeneous coordinates of the true point position can therefore be obtained as an estimate of the polar axis of the Bingham distribution that best fits the normalized sample points. The most common method for estimating a distribution's parameters from a sample of observations is maximum likelihood estimation.

Relevant statistics for Bingham's distribution can be summarized as follows; a more detailed presentation can be found in [2, 5]. Given a set of n unit vectors $\phi_i = \langle x_i, y_i, z_i \rangle$ assumed to be distributed according to Bingham's distribution, a sufficient statistic for the orientation and shape parameters of the distribution is the sample second moment or *scatter matrix* $\phi_i \phi_i^t$. It can be shown that the maximum likelihood estimate of the pole of a bipolar Bingham distribution is the eigenvector associated with the largest eigenvalue of this sample scatter matrix. Equations for computing confidence regions on the sphere can be found in [2, 5]. Once again, for a look at experimental results the reader is directed to [3].

It must be noted that this maximum likelihood approach for data fusion implicitly assumes that all points in the sample are independent and identically distributed. While the independence assumption may be a necessary evil, points in the sample will probably *not* be identically distributed, since some extracted image features are more accurate than others. Future work will address ways of combining point estimates of different, but estimable, accuracy.

References

- [1] R. Beran, "Exponential Models for Directional Data," *Annals of Statistics*, V.7, No.6, 1979, pp. 1162-1178.
- [2] C. Bingham, "An Antipodally Symmetric Distribution on the Sphere," *The Annals of Statistics*, Vol. 2, 1974, pp. 1201-1225.
- [3] R.T. Collins, "Single Plane Model Extension using Projective Transformations," Proc. Darpa I.U. Workshop, San Diego, CA., Jan. 1992, pp. 917-923.
- [4] R.T. Collins and R.S. Weiss, "Vanishing Point Calculation as a Statistical Inference on the Unit Sphere," *Proc. 3rd ICCV*, Osaka, Dec. 1990, pp. 400-403.
- [5] R.T. Collins and R.S. Weiss, "Deriving Line and Surface Orientation by Statistical Methods," *Proc. Darpa I.U. Workshop*, Pittsburgh, Sept. 1990, pp. 433-438.
- [6] O.D. Faugeras and F. Lustman, "Motion and Structure from Motion in a Piecewise Planar Environment," *IJPRAI*, Vol. 2, 1988, pp. 485-508.
- [7] R. Mohr and E. Arbogast, "It Can be Done Without Camera Calibration," *PRL*, V. 12, 1990, pp. 39-43.
- [8] C.E. Springer, *Geometry and Analysis of Projective Spaces*, W.H. Freeman & Co., San Francisco., 1964.