On the Perspective Projection of 3-D Planar-Faced Junctions

R.Consales, A. Del Bimbo, P. Nesi
Dipartimento di Sistemi e Informatica, Facoltà di Ingegneria
Università degli Studi di Firenze
Via S. Marta 3, 50139 Firenze, Italy

Abstract

In this paper, a method to determine the spatial orientation of a 3-D planar-faced object from a single perspective view is presented. In this method the junctions, that is the concurrences of three object edges in a vertex, are considered as key features. A new geometric constraint, related to the perspective projection of the object junctions, is shown and exploited to make the process faster and more efficient. In this way, the knowledge of only two parameters is sufficient to verify the object orientation in 3-D space. A new system of equations is proposed which define the orientations space for a junction as a subset of $\Re^2$ *, instead of* $\Re^3$ *as in the previous literature.*

Introduction

Recently, early vision processes, as "shape from" techniques, binocular stereo matching, structure from motion and surface reconstruction, have been intensively investigated in order to comprehend human vision mechanisms and to develop image understanding or object recognition systems.

The primary purpose of early vision processes is to retrieve physical properties of a 3-D object, like its structure, motion or orientation, from one or more images. Unfortunately, the projection transform of a 3-D object in a 2-D plane is not an information preserving mapping, since the dimension along the optical axis is not recoverable. As a consequence a single image can be arisen from infinite objects, and is not enough to infer the 3-D shape of an imaged object. In other words, the problem of retrieving the 3-D object structure from a single 2-D view has not an unique solution and is heavily ill-posed, unless additional data are considered. In the several approach proposed in literature, this supplementary information frequently comes from constraints on the projection transform or from heuristic rules.

In order to recognize a 3-D object from a single 2-D view and to estimate its spatial orientation, its six degrees of freedom must be determined. Generally this task is accomplished considering the matching between three image lines and three model lines.

In this framework both orthographic [1], [2] and perspective [3-10] projection model have been adopted, the latter being of major interest as it resembles the way human vision operates, and because it is more constraining [3], [4].

Using perspective projection, Barnard [5] describes a computational method to backproject image features, such as angles and curvatures, into 3-D space. Referring to the "right-angle illusion" [6], he interprets a triplet of image lines as the projection of a right-angled vertex, since he asserts three lines are always seen as a rectangular trihedral vertex unless they form two very acute angles.

Shakunaga and Kaneko [7] introduce a new relation, named Perspective Angle Transform (PAT), to recover a 3-D object configuration from a single image. Although using this transform several non-linear constraints are reduced to algebraic equations, this approach imposes unduly constraining assumptions on the orthogonality of certain model lines.

Dhome et al. [3] present an analytical solution to the problem of determining the attitude of a 3-D object from a single perspective view, but they still need Lowe's logical rules [11] to prune the set of multiple solutions the find analytically.

Horaud [8] proposes a backprojection method to determine the possible orientations of a 3-D object, by analyzing the object junctions, that is the intersections of three non-coplanar straight edges in an object vertex, and considering three space angles as known.

A similar approach is presented by Kanatani [9] who studies the constraints on the spatial orientation of line segments, assuming lengths and angles as known.

In this paper, a backprojection method to determine the spatial orientation of a 3-D planar-faced object from a single perspective view is presented. In this method, as in Horaud's approach [8], the junctions of three object edges are considered. A new geometric constraint, related to the perspective projection of the junctions, is shown and exploited to make the process faster and more efficient. In this way, the knowledge of only two parameters, instead of three, allows to verify the object orientation in 3-D space.

0-8186-2910-X/92 $3.00 © 1992 IEEE

Backprojection of a space junction

Hereafter, the perspective projection model is used and for the sake of clarity, lines and planes are referred to in italics and their orientations in boldface. A left-handed Cartesian coordinate system is adopted, centered in the camera focus F, with the z-axis along the optical axis and the image plane coinciding with the plane $z = f$ where f is the focal length (see Fig.1).

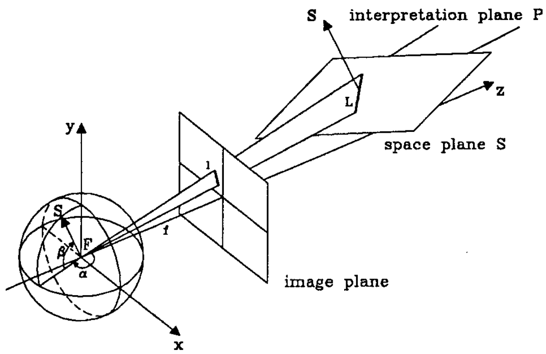

Fig.1 - Cartesian coordinate system and Gaussian sphere

The direction of a line and the orientation of a plane are represented as points on the unit *Gaussian sphere* centered at the camera focus F. In practice, only one hemisphere is sufficient to map the space of all possible orientations [6]. In this paper the hemisphere oriented toward the viewer, determined by the points with spheric coordinates (α, β) with $\alpha \in (\pi/2, 3\pi/2)$ and $\beta \in (-\pi/2, \pi/2)$, is considered and represented as an $\alpha-\beta$ plane. Any image entity is backprojected on a space plane S (see Fig.1), whose orientation $\mathbf{S}$ must satisfy the geometric constraints due to the perspective projection. These constraints force $\mathbf{S}$ to belong to loci of points on the considered hemisphere and hence on the $\alpha-\beta$ plane.

A space junction is the concurrence of three non-coplanar straight lines in a point in space, called vertex. Let us now consider the generic space junction in Fig.2. L_1, L_2, and L_3 give rise to three image lines, l_1, l_2, l_3 in the image plane. For any image line l_i, i=1,2,3, it is possible to determine the *interpretation plane* P_i that passes through the camera focus F and contains both l_i and L_i. Let $\mathbf{S}$ be the normal unit vector of the plane S formed by L_1 and L_2, θ and ψ the angles formed by the projection of L_3 on S with L_1 and L_3, respectively, and φ the angle between L_1 and L_2 (see Fig.2). In order to determine the junction orientation, two items must be considered: L_1 and L_2 form the angle φ on the space plane S:

$$\mathbf{L}_1 \cdot \mathbf{L}_2 = \| \mathbf{L}_1 \| \| \mathbf{L}_2 \| \cos(\varphi) \qquad (1)$$

and L_3 belongs to the interpretation plane P_3:

$$\mathbf{L}_3 \cdot \mathbf{P}_3 = 0 \qquad (2)$$

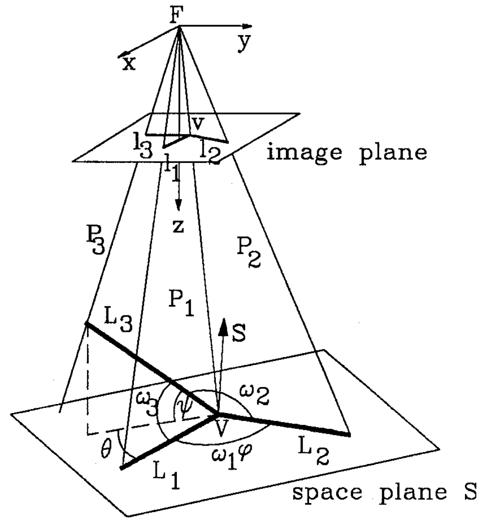

Fig.2 - Backprojection of a space junction

Since L_1 and L_2 belong to the interpretation planes P_1, P_2, and lie on the space plane S, their unit direction vectors can be expressed as:

$$\mathbf{L}_i = \frac{\mathbf{S} \times \mathbf{P}_i}{\| \mathbf{S} \times \mathbf{P}_i \|} \quad i = 1,2 \qquad (3)$$

and a constraint on the orientation of the space plane S is obtained as follows:

$$(\mathbf{S} \times \mathbf{P}_1) \cdot (\mathbf{S} \times \mathbf{P}_2) = \| \mathbf{S} \times \mathbf{P}_1 \| \| \mathbf{S} \times \mathbf{P}_2 \| \cos(\varphi) . \qquad (4)$$

Equation (2) represents the constraint on orientation $\mathbf{L}_3$. With respect to the local Cartesian coordinate system defined by $\mathbf{S}, \mathbf{L}_1$ and $(\mathbf{S} \times \mathbf{L}_1)$, the direction vector $\mathbf{L}_3$ can be expressed as:

$$\mathbf{L}_3 = \mathbf{S} \sin(\psi) + \mathbf{L}_1 \cos(\theta) \cos(\psi) + (\mathbf{S} \times \mathbf{L}_1) \sin(\theta) \cos(\psi)$$

Considering equation (3) for the unit vector $\mathbf{L}_1$, and the cross product's rules, equation (2) can be written as:

$$(\mathbf{S} \cdot \mathbf{P}_3) \| \mathbf{S} \times \mathbf{P}_1 \| \sin(\psi) + (\mathbf{S} \cdot (\mathbf{P}_1 \times \mathbf{P}_3)) \cos(\theta) \cos(\psi) + \\ + \left[(\mathbf{S} \cdot \mathbf{P}_1)(\mathbf{S} \cdot \mathbf{P}_3) - (\mathbf{P}_1 \cdot \mathbf{P}_3) \right] \sin(\theta) \cos(\psi) = 0 \qquad (5)$$

Equations (4) and (5) are defined in terms of five unknowns: the space angles φ, θ, ψ, and the spheric coordinates (α, β) of $\mathbf{S}$, hence the solution space of these equations is in $\Re^5$. These equations are non-linear and solutions cannot easily be determined analytically. According to Horaud [8], a constructive method can be implemented on a computer system. By setting specific values for the angles φ, θ, ψ, and making the orientation $\mathbf{S}$ vary over the Gaussian hemisphere, two curves can be plotted in the $\alpha-\beta$ plane, which intersect in 0, 1 or 2 points, the latter case being due to the Necker's cube illusion. The intersections, if any, correspond to the orientations of plane S for the space angles chosen. In Fig.3, the curves C_1 and C_2, corresponding to equations (4) and (5) respectively, have been built for a sample junction with $\varphi = 109.4^o$, $\theta = 35.7^o$, $\psi = 45^o$.

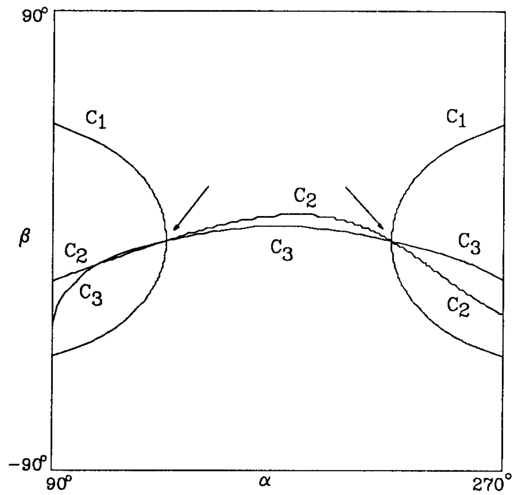

Fig.3 - Backprojection of a sample space junction

Now, considering the "vertex angles" ω_1, ω_2 and ω_3 formed by the three space lines (see Fig.2), the relationships:

$$\omega_1 = \varphi \qquad (6a)$$
$$\cos(\omega_2) = \cos(\psi)\cos(\varphi + \theta) \qquad (6b)$$
$$\cos(\omega_3) = \cos(\psi)\cos(\theta) \qquad (6c)$$

and their inverse:

$$\varphi = \omega_1 \qquad (7a)$$
$$\theta = arctg\left[\frac{\cos(\omega_3)\cos(\omega_1) - \cos(\omega_2)}{\cos(\omega_3)\sin(\omega_1)}\right] \qquad (7b)$$
$$\psi = arcos\left[\frac{\cos(\omega_3)}{\cos(\theta)}\right] \qquad (7c)$$

can be stated. These allow to express equations (4) and (5) in terms of ω_1, ω_2, ω_3, respectively in the form:

$$(\boldsymbol{S} \times \boldsymbol{P}_1) \cdot (\boldsymbol{S} \times \boldsymbol{P}_2) = \| \boldsymbol{S} \times \boldsymbol{P}_1 \| \, \| \boldsymbol{S} \times \boldsymbol{P}_2 \| \cos(\omega_1) \qquad (8)$$

$$(\boldsymbol{S} \cdot \boldsymbol{P}_3)\|\boldsymbol{S} \times \boldsymbol{P}_1\| h(\omega_1, \omega_2, \omega_3) + (\boldsymbol{S} \cdot (\boldsymbol{P}_1 \times \boldsymbol{P}_3))\cos(\omega_3) + \left[(\boldsymbol{S} \cdot \boldsymbol{P}_1)(\boldsymbol{S} \cdot \boldsymbol{P}_3) - (\boldsymbol{P}_1 \cdot \boldsymbol{P}_3)\right] g(\omega_1, \omega_2, \omega_3) = 0 \qquad (9)$$

where:

$$g(\omega_1, \omega_2, \omega_3) = \frac{\cos(\omega_3)\cos(\omega_1) - \cos(\omega_2)}{\sin(\omega_1)}$$
$$h(\omega_1, \omega_2, \omega_3) = \sin\left[arcos\left[\frac{\cos(\omega_3)}{\cos\left[arctg\left(g(\omega_1, \omega_2, \omega_3)\right)\right]}\right]\right]$$

It will be shown that the knowledge of the interpretation plane P_3 allows to verify the orientation of a space junction by setting only two of the three vertex angles, thus reducing the solution space of the equations (8) and (9) from $\Re^3$ to $\Re^2$. Let us assume angles ω_1 and ω_3 are known. If the orientation $\boldsymbol{S}$ is set, the direction vectors $\boldsymbol{L}_1$, $\boldsymbol{L}_2$ are derived with equation (3). Considering the cone C obtained from the rotation of a generic straight line L, starting from the junction vertex V, around L_1 with aperture angle at ω_3, the straight line L_3 corresponds to the intersection of C with the interpretation plane P_3. If the junction vertex V coincides with the origin of the coordinate system, as in Fig.4, cone C can be expressed as:

$$\boldsymbol{L}_1 \cdot \boldsymbol{L} = \| \boldsymbol{L} \| \cos(\omega_3) \qquad (10)$$

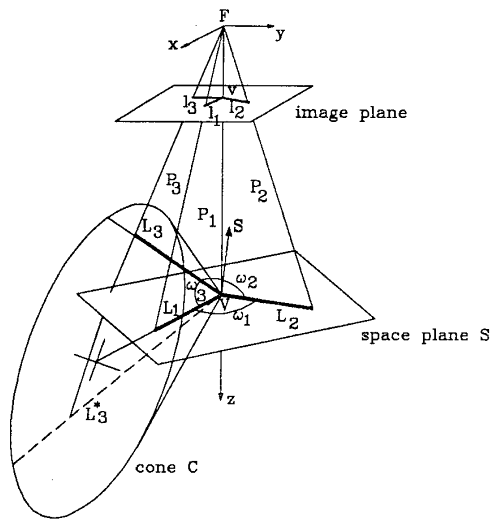

Fig.4 - Cone $\boldsymbol{C}$ formed by L_3 around L_1

Since L_3 belongs both to the cone surface and the plane P_3, the unit vector $\boldsymbol{L}_3$ satisfies the following constraints:

$$\| \boldsymbol{L}_3 \| = 1 \qquad (11a)$$
$$\boldsymbol{L}_1 \cdot \boldsymbol{L}_3 = \cos(\omega_3) \qquad (11b)$$
$$\boldsymbol{L}_3 \cdot \boldsymbol{P}_3 = 0 \qquad (11c)$$

Equations (11) form a system of three equations in the three unknown components of $\boldsymbol{L}_3$, that generally admits two solutions, considering the inherent ambiguity of the 2-D projection of 3-D objects. Once $\boldsymbol{L}_3$ is obtained, angle ω_2 can be directly determined from the dot product:

$$\boldsymbol{L}_2 \cdot \boldsymbol{L}_3 = \cos(\omega_2) \qquad (15)$$

These results are true if the vertex V of the space junction corresponds to the origin of the coordinate system. Actually, the direction vector $\boldsymbol{L}_3$ of the intersection between C and P_3, is invariant under a spatial shift of the junction along the optical axis (z-axis), provided that V lies on this axis. This implies that the vertex v of the image junction coincides with the center of the image plane (the origin of the image coordinate system). In general, this assumption is not always verified, and the vertex v must be moved to the center of the image plane. A simple translation would fake the projected image, therefore a rotation is needed. In line with Haralick [10], a representation in homogeneous coordinates can be used to describe this transformation. Obviously if the camera is rotated by a rotation matrix $R = (r_{ij})$, the scene will rotate by the inverse $R^{-1} = (s_{ij})$, therefore the homogeneous coordinates $(x, y, z, 1)$ of a generic point in the scene must be rotated by R^{-1}, translated to the image plane, and then perspectively projected. Considering the Cartesian coordinate system here adopted, the perspective projection and the translation can be represented by the transformation matrices M_p and M_t respectively, where:

$$M_p = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 1 \end{bmatrix}, \quad M_t = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -f \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Hence, the rotated homogeneous coordinates (x^*, y^*, t^*) of (x, y, z) are:

$$\begin{bmatrix} x^* \\ y^* \\ t^* \end{bmatrix} = M_p M_t \begin{bmatrix} s_{11} & s_{12} & s_{13} & 0 \\ s_{21} & s_{22} & s_{23} & 0 \\ s_{31} & s_{32} & s_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

and the rotated image coordinates are:

$$x' = \frac{x^*}{t^*} = f\frac{x s_{11} + y s_{12} + z s_{13}}{x s_{31} + y s_{32} + z s_{33}}, \quad y' = \frac{y^*}{t^*} = f\frac{x s_{21} + y s_{22} + z s_{23}}{x s_{31} + y s_{32} + z s_{33}}$$

Further explanations of these equations can be found in [10]. In order to move the vertex v to the center of the image plane the *standard transformation* reported in [9] can be applied. The associated *standard rotation* $R(a,b)$ represents a camera rotation that maps a given point (a,b,f) to the origin $(0,0,f)$ on the image plane. Since the matrix $R(a,b)$ is orthogonal with unit determinant [9], the inverse R^{-1} coincides with the transpose R^T, therefore the image point (x,y) is transformed into the image point (x', y') with coordinates:

$$x' = f\frac{x r_{11} + y r_{21} + f r_{31}}{x r_{13} + y r_{23} + f r_{33}}, \quad y' = f\frac{x r_{12} + y r_{22} + f r_{32}}{x r_{13} + y r_{23} + f r_{33}}.$$

By using the appropriate *standard rotation* $R(x_v,y_v)$, the vertex v of the image junction can be moved to the image origin and the vertex V of the space junction to the z-axis, so that equations (11) still hold. As a consequence, the equations (8) and (9) can be verified by imposing a value only for the vertex angles ω_1 and ω_3.

In order to solve equations (8) and (9), a constructive method, similar to the procedure sketched above, can be adopted. In this case, specific values for only two angles, instead of three as in Horaud's approach, must be set. With respect to the same sample junction (vertex angles $\omega_1 = 109.4^o$, $\omega_3 = 54.7^o$), the curves C_1 and C_3 in Fig.3 have been obtained computing equations (8) and (9), respectively, for each point in the $\alpha-\beta$ space. Curve C_3 is slightly different from curve C_2 corresponding to equation (5). Nevertheless, the solutions of equations (8) and (9) (the intersections between C_1 and C_3) and the solutions of equations (4) and (5) (the intersections between C_1 and C_2) are the same. This reveals that, taking into account the constraints related to the perspective projection of a space junction, the knowledge of only two vertex angles, instead of three, is sufficient to evaluate the possible orientations of that junction, thus reducing the solution space from $\Re^3$ to $\Re^2$.

References

[1] T. Kanade, "Recovery of the Three-Dimensional Shape of an Object from a Single View", Artificial Intelligence, Vol.17, pp. 409-460, 1981.

[2] A.P. Witkin, "Recovering Surface Shape and Orientation from Texture", Artificial Intelligence, Vol.17, pp. 17-45, 1981.

[3] M. Dhome, M. Richetin, J. Lapresté and G. Rives, "Determination of the Attitude of 3-D Objects from a Single Perspective Views", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.11, N.12, pp. 1265-1278, Dec 1989.

[4] F. Ulupinar and R. Nevatia, "Constraints for Interpretation of Line Drawings under Perspective Projection", Computer Vision, Graphics, and Image Processing: Image Understanding, Vol.53, N.1, pp. 88-96, Jan. 1991.

[5] S.T. Barnard, "Choosing a Basis for Perceptual Space", Computer Vision, Graphics, and Image Processing, Vol.29, pp. 87-99, 1985.

[6] S.T. Barnard, "Interpreting Perspective Images", Artificial Intelligence, Vol.21, pp.435-462, 1983.

[7] T. Shakunaga and H. Kaneko, "Perspective Angle Transform: Principle of Shape from Angles ", International Journal of Computer Vision, Vol.3, pp.239-254, 1989.

[8] R. Horaud, "New Methods for Matching 3-D Object with Single Perspective Views", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.9, N.3, pp. 401-412, May 1987.

[9] K. Kanatani, "Constraints on Length and Angle", Computer Vision, Graphics, and Image Processing, Vol.41, pp. 28- 42, 1988.

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

[11] D.G. Lowe, "Perceptual Organization and Visual Recognition", Kluwer, Boston, 1985, chp.7.