A New Hough Transform for the Detection of Arbitrary 3-Dimensional Objects

John Mc Donald and David Vernon

Department of Computer Science, National University of Ireland, Maynooth, Ireland. Email: johnmcd@cs.may.ie

Abstract. The existing Generalised Hough Transform, although altered to cater for scaling and rotation of the object in the plane, fails to detect the object under rotations out of the plane. This is due to the lack of 3-dimensional information contained in the 2-dimensional template image. In this paper we present a new Hough Transform, known as the *Surface Normal Hough Transform* (SNHT), which using a suitable 2-dimensional surface representation, transforms a set of surface normals to a surface parameter space. The effect of the SNHT is to map point sets representing a surface in the input space, to a peak in the parameter space. The coordinates of this peak parameterise the given surface and hence allow for pose invariant object detection and localisation.
Keywords: 3-D Computer Vision; Hough Transform; Pose Invariant Object Detection; Surface Registration.

1 Overview

It is generally accepted that the main aim of computer vision is to realise an adaptive system which is capable of navigation within, and interaction with, the 3-dimensional world. This interaction is facilitated through the interpretation of visual information received from some imaging devices. In order to achieve this goal, some robust method(s) for segmenting objects in images from the perceived background is required.

The *Generalised Hough Transform* (GHT) [1, 5] is a technique used to locate arbitrary curves in images, based on a 2-dimensional template. Given this template, the first stage of the computation is to build an internal representation of the curves it contains. This representation is based on the position and orientation of each *edge point* with respect to some reference point. Using this representation, the GHT then looks for similar configurations of edge points in other input images. If such a configuration is found to exist, the template object is said to be localised, upon which the subsequent segmentation of the object from the background is trivial.

Although this technique has been altered to cater for variation in object scale and object orientation, at present it is not capable of handling variation in object

To be published in *Proc. of the Optical Engineers Society of Ireland and the Irish Machine Vision and Image Processing Joint Conference, 1998*

pose. That is, the rotation of the object out of the image plane. This is due to the fact that our only representation of the object is generated from a single edge-detected 2-dimensional projection, and therefore contains no information regarding 3-dimensional structure.

In this paper we redress this problem. Here it is proposed that by extending the space of the template image from $\mathbb{R}^2$ to $\mathbb{R}^3$, as well as altering the technique for generating the template representation, a new type of representation which takes the 3-dimensional nature of the object into account may be constructed. Since the GHT is a two staged approach, *i.e. (i)* the model extraction and definition stage, and *(ii)*, the Hough transform or localisation stage, it will be necessary to alter the second stage sufficiently to allow for this new representation.

Section 2 gives a formal introduction to the Hough transform and describes the necessary mathematics for what is to follow. In Sect. 3, the stages necessary to compute the existing 2-dimensional GHT are outlined. This is accompanied by a description of techniques for achieving orientation and scale invariance in the GHT (Sect. 3.2). Section 3.3 identifies the limitations of the GHT and justifies the necessity for its extension.

The proposed strategy to be used in extending the technique to cater for 3-dimensional objects is outlined in Sect. 4. This paper concludes with a discussion on the material presented (Sect. 5).

2 The Hough Transform

The *Hough transform* (HT) [2, 4] is a technique which, using some curve representation, transforms a set of points defined over the image space to a set of points defined over some parameter space (known as *Hough space*).

Points in Hough space represent particular instances of the curve in the image. Therefore, the strategy used by the HT is to map sets of points from a particular instance of the considered curve, *i.e.* the template curve, to the point representing the curve in Hough space and, in effect, cause a peak to occur at that point. Once each point in the image plane has been considered, the Hough space is searched for the peak of maximum height. The coordinates of this peak in Hough space give the parameters which define the curve in the image plane.

Equations (1) and (2) define such a mapping. Here $\mathcal{H}()$ denotes the application of the Hough transform, $H(p_1, \ldots, p_m)$ denotes the parameter or Hough space, $I(x_1, \ldots, x_n)$ denotes the image space, and C denotes the set of curve points in I. The δ function evaluates to 1 if the Hough transform of the curve point $(x_1, x_2, \ldots, x_n)$ transforms to the parameter point $(p_1, p_2, \ldots, p_m)$, and the function evaluates to 0 otherwise. So, the Hough space,

$$H(p_1, p_2, \ldots, p_m) = \sum_{(x_1, x_2, \ldots, x_n) \in C} \delta(x_1, x_2, \ldots, x_n;\ p_1, p_2, \ldots, p_m) \qquad (1)$$

where,

$$\delta(x_1, x_2, \ldots, x_n;\ p_1, p_2, \ldots, p_m) = \begin{cases} 1 & \text{if } \mathcal{H}(I(x_1, x_2, ..., x_n)) = \\ & \qquad (p_1, p_2, \ldots, p_m) \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

It can be seen from the above equations that the value of a point in Hough space corresponds to the number of curve points in image space which transform to that position.

The main advantages of the HT are that it is particularly resilient to both occlusion and noise. This is due to the fact that each point on the curve is considered independently and therefore still maps to the same point in Hough space, regardless of the presence or absence of auxiliary points.

In general, existing HTs may be divided into two categories: *(i)* The *Standard* or *Classical Hough Transform* (SHT) and, *(ii)* the *Generalised Hough Transform*. The main difference between *(i)* and *(ii)* is the mechanism used in representing the desired curves. The SHT requires a parametric formulation whereas, the GHT uses a look-up table to define the shape. Hence, the SHT is constrained to curves which may be defined analytically, whilst the GHT may be used for the segmentation of arbitrary curves.

3 The Generalised Hough Transform

In the absence of a closed form analytical equation for the considered curve, the GHT defines a new function relating curve point orientations to vector directions and magnitudes. More specifically, the GHT uses an internal representation, $R(\Omega)$, to define the relationship between the orientation of the tangent line,Ω at each curve point to the magnitude, r, and direction, β, of the vector joining that curve point to a predefined reference point (see Fig. 1). Therefore R can be viewed as a one-to-many mapping from orientation-space to (r, β)-space.

The GHT, on the other hand, may be viewed as a mapping from image space, $I(x, y)$, to a Hough space, $H(x_{\mathrm{ref}}, y_{\mathrm{ref}})$, defined over the possible coordinates of the reference point,

$$H(x_{\mathrm{ref}}, y_{\mathrm{ref}}) = \sum_{(x,y) \in C} \delta(x, y;\ x_{\mathrm{ref}}, y_{\mathrm{ref}}) \tag{3}$$

where,

$$\delta(x, y;\ x_{\mathrm{ref}}, y_{\mathrm{ref}}) = \begin{cases} 1 & \text{if } \mathcal{H}(I(x, y)) = \\ & \qquad (x_{\mathrm{ref}}, y_{\mathrm{ref}}) \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Here, the summation is performed for each point contained in the set C, defined as the set of curve points contained in $I(x, y)$.

To compute $\mathcal{H}(I(x, y))$, the orientation, Ω, of the tangent line at $I(x, y)$ is calculated. Using this value, the possible (r, β) pairs for the respective edge

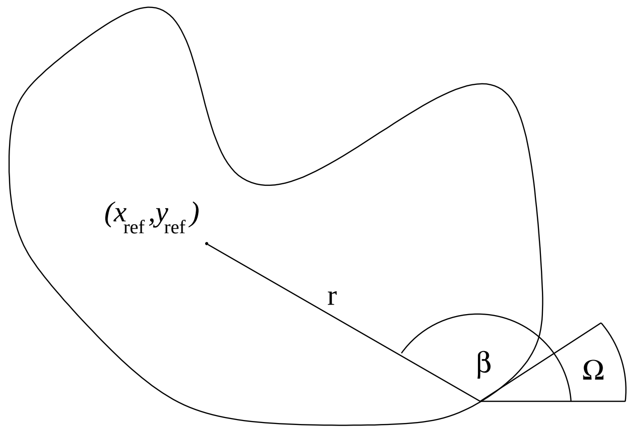

Fig. 1. Computing the required angles

points may be extracted using $R(\Omega)$. Since these pairs give the relationship between the template curve point and the position of the reference point, we may now compute the possible positions of the object reference point in the image. Equations (5) and (6) define this computation, giving us the reference point coordinates $(x_{\mathrm{ref}}, y_{\mathrm{ref}})$, through the mapping,

$$\begin{aligned} \mathcal{H}(I(x,y)) &= (x_{\mathrm{ref}}, y_{\mathrm{ref}}) && (5) \\ &= (x + r\cos(\beta), y + r\sin(\beta)) \ . && (6) \end{aligned}$$

Note that it is possible for more than one curve point in the template image to be at any given orientation, hence this computation may need to be computed more than once for a particular Ω [7].

3.1 Algorithmic and Representational Issues

Typically, a look-up table known as an *r-table*, is used to represent R (see Fig. 2). Hence, the first stage of processing involves computation of the r-table. This is achieved by first choosing a reference point $(x_{\mathrm{ref}}, y_{\mathrm{ref}})$ within the template, normally taken to be the centroid of the curve. Next each curve point c_i is visited, at which the orientation of tangent line Ω_i, the angle β_i, and magnitude r_i, of the vector joining the reference point to the edge point are calculated. These values are then stored as a pair, (r_i, β_i), in the r-table at the position indexed by the edge point's orientation, Ω_i.

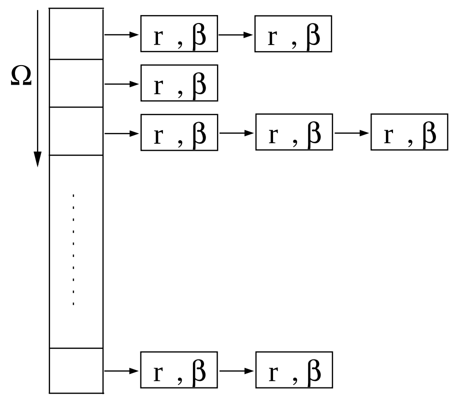

Fig. 2. The GHT r-Table

Equations (7), (8) and (9) define the calculations for the required values detailed above,

$$c_i = (x_i, y_i) \tag{7}$$

$$r_i = \sqrt{(x_{\mathrm{ref}} - x_i)^2 + (y_{\mathrm{ref}} - y_i)^2} \tag{8}$$

$$\beta_i = \tan^{-1}\left(\frac{y_{\mathrm{ref}} - y_i}{x_{\mathrm{ref}} - x_i}\right) \quad . \tag{9}$$

During the second, or transform, stage a 2-dimensional array known as an *accumulator array* is used to store each $(x_{\mathrm{ref}}, y_{\mathrm{ref}})$ as they are computed. The accumulator array then represents the Hough space for the given transform. The sampling density of the accumulator array is normally chosen to be equivalent that of the unknown image. This allows for the maximum degree of accuracy required in specifying the reference point in the image. Once the second stage is completed, *i.e.* after each point in the image has been considered, the final stage is to search the accumulator for the point at which the peak of maximum height has occurred. This point is taken to be the position of the object reference point in the unknown image.

3.2 Incorporating Orientation and Scale Invariance

From the description of the GHT presented here, it can be seen that the technique does not cater for rotation or scaling of the object. That is, when the object is rotated by some offset, ϕ, the computed orientation, Ω', of each point differs from the template orientation, Ω, as in (12). This results in the mapping to (r, β)-space being computed for $R(\Omega')$, as opposed to $R(\Omega)$.

Scaling of the object by a factor of s results in the correct mapping, $R(\Omega)$, but invalid magnitudes, r, of the extracted (r, β)-vectors, *i.e.* the distance between the curve points and reference point in the image space has increased, but the corresponding r's defined by $R(\Omega)$ have not been scaled accordingly.

To overcome these problems, instead of computing the GHT for the curve at a specific scale and orientation, the GHT is computed for each of the possible values of s and ϕ, over a selected range of orientations and scales [1]. We refer to this strategy as the orientation and scale-invariant generalised Hough transform (OSIGHT). The Hough space for the OSIGHT is defined as

$$H(x_{\mathrm{ref}}, y_{\mathrm{ref}}, s, \phi) = \sum_{(x,y)\in C} \delta(x, y;\ x_{\mathrm{ref}}, y_{\mathrm{ref}}, s, \phi) \tag{10}$$

where,

$$\delta(x, y;\ x_{\mathrm{ref}}, y_{\mathrm{ref}}) = \begin{cases} 1 & \text{if } \mathcal{H}(I(x,y)) = \\ & \quad (x_{\mathrm{ref}}, y_{\mathrm{ref}}, s, \phi) \\ 0 & \text{otherwise.} \end{cases} \tag{11}$$

As can be seen from (10) and (11), specification of a curve now requires a 4-dimensional vector, $(x_{\mathrm{ref}}, y_{\mathrm{ref}}, s, \phi)$, where s and ϕ denote scale and orientation, respectively, and the template curve is assumed to have $s = 1$ and $\phi = 0$. Hence, we must extend the Hough space from two-dimensions to four-dimensions (*i.e.* the Hough space is now defined over $x_{\mathrm{ref}}, y_{\mathrm{ref}}, s$ and ϕ).

In the transform stage, as each point in the unknown image is visited, its orientation, Ω, is calculated. We then calculate $(x_{\mathrm{ref}}, y_{\mathrm{ref}}, s, \phi)$ for each possible s and ϕ in a brute force manner, as outlined below.

For each possible orientation offset, ϕ, the corresponding curve point orientation, Ω', is calculated as,

$$\Omega' = \Omega - \phi\ . \tag{12}$$

This value is then mapped to a set of points in (r, β) space using $R(\Omega')$. Equations (13) and (14) may be used in conjunction with each element in this (r, β) set, to map the image point to the corresponding Hough point,

$$\begin{aligned} \mathcal{H}(I(x,y)) &= (x_{\mathrm{ref}}, y_{\mathrm{ref}}, s, \phi) &(13)\\ &= (x + (s.r)\cos(\beta + \phi), y + (s.r)\sin(\beta + \phi), s, \phi)\ . &(14) \end{aligned}$$

Again, as each (r, β)-pair is considered, a set of reference points is calculated, where each element of this set corresponds to a specific scale s.

3.3 Limitations of the GHT

A significant problem encountered with the GHT, with regard to 3-dimensional vision, is variation due to object pose. Since the template is a 1-dimensional curve defined over 2-dimensions, the technique is restricted to an single edge-detected two-dimensional projection of the desired object. Subsequently, rotation of the object *out* of the image plane results in a new 2-dimensional projection

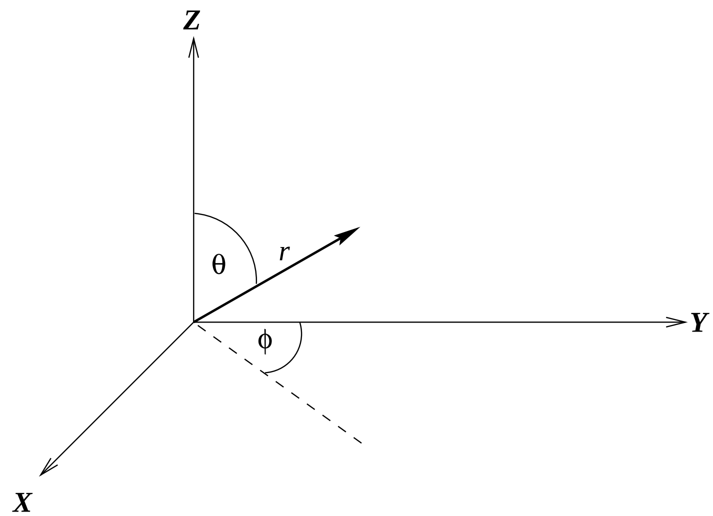

Fig. 3. Spherical Polar Coordinate System

and consequently an unknown edge image which will not be localised using the GHT. Note, symmetry is not considered here.

It should also be noted that this is a problem of *representation*, that is to say that the computational model (Sect. 2) of the Hough Transform is capable of catering for 3-dimensions if it is provided with the necessary 2-dimensional surface representation and $\mathcal{H}$ equations.

In the following section we present a new technique known as the *Surface Normal Hough Transform* (SNHT) for generating this required representation in conjunction with the necessary $\mathcal{H}$ equations for performing the Hough transform stage.

4 The Surface Normal Hough Transform

The Surface Normal Hough Transform (SNHT) is a new technique for computing the 3-dimensional position of a surface having a specified pose and, by extension, a technique for computing the 3-dimensional position *and* pose of a surface with respect to the pose of a prototypical exemplar of that surface. This technique, therefore, overcomes the problem discussed in Sect. 3.3 and may also be used as a solution to the familiar registration problem.

As with the GHT, in the absence of a closed form analytical equation for the considered surface the SNHT defines a new function relating *surface point orientations*, to vector directions and magnitudes. To specify the orientation of a given surface point, the orientation of the unit normal $\hat{\mathbf{n}}$ at that point is used.

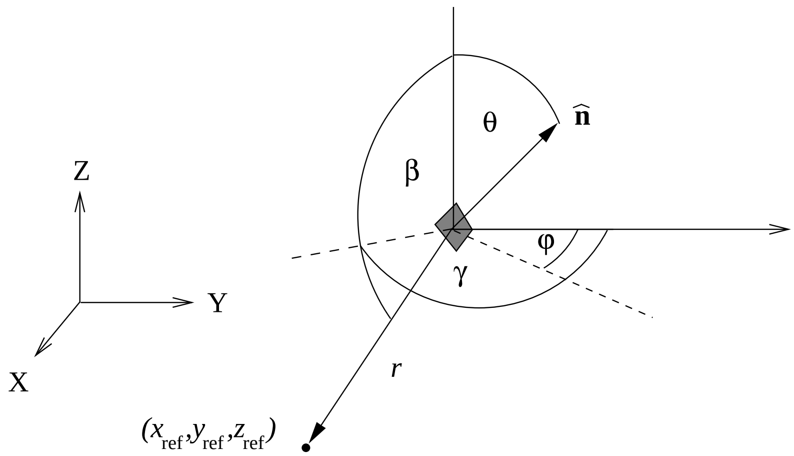

Fig. 4. 3-dimensional vector representations.

This may be defined as,

$$\hat{\mathbf{n}} = (\theta, \phi) \tag{15}$$

where, θ and ϕ denote polar and azimuth angle, respectively (see Fig. 3).

In a manner analogous to the GHT, the SNHT also uses an R-mapping to represent the considered surface, although in this instance it is a function of the two variables θ and ϕ described above. Also, because the surface is defined over 3-dimensions, the reference point must in turn be 3-dimensional *i.e.* the reference point is now chosen to be some predefined point $(x_{\mathrm{ref}}, y_{\mathrm{ref}}, z_{\mathrm{ref}})$. Hence, the vector joining the surface point to the reference point is represented in spherical polar coordinates as an (r, β, γ)-*triple* (see Fig. 4). R is therefore defined as a mapping from (θ, ϕ)-space to (r, β, γ)-space.

In order to allow for this new surface representation the Hough transform equations of the second stage must also be altered accordingly. Extending the computational model described in Sect. 2, the SNHT may now be mathematically defined as

$$H(x_{\mathrm{ref}}, y_{\mathrm{ref}}, z_{\mathrm{ref}}) = \sum_{(x,y,z) \in S} \delta(x, y, z;\ x_{\mathrm{ref}}, y_{\mathrm{ref}}, z_{\mathrm{ref}}) \tag{16}$$

where,

$$\delta(x, y, z;\ x_{\mathrm{ref}}, y_{\mathrm{ref}}, z_{\mathrm{ref}}) = \begin{cases} 1 & \text{if } \mathcal{H}(I(x, y, z)) = \\ & \quad (x_{\mathrm{ref}}, y_{\mathrm{ref}}, z_{\mathrm{ref}}) \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

Here, S is defined as the set of surface points in the input image space, $I(x, y, z)$.

To compute $\mathcal{H}$, the orientation, (θ, ϕ), of the surface normal at each point in the image $I(x, y, z)$ is first calculated. These orientations then allow the corresponding surface point's (r, β, γ)-triples to be computed from $R(\theta, \phi)$. Since an (r, β, γ)-triple represents a 3-dimensional vector from a surface point to a possible reference point, this reference point may now be recovered using the $\mathcal{H}$ equation,

$$\mathcal{H}(I(x,y,z)) = (x_{\mathrm{ref}}, y_{\mathrm{ref}}, z_{\mathrm{ref}}) \tag{18}$$

$$= (x + r\sin(\beta)\cos(\gamma), y + r\sin(\beta)\sin(\gamma), z + r\cos(\gamma)) \ . \tag{19}$$

4.1 Algorithmic and Representational Issues

We represent the R-mapping through the use of an r-table, similar to that described in Sect. 3.1,although since the R-mapping is now defined as a function of two variables, a 2-dimensional array is used to represent it (see Fig.5). Also, since the vectors joining the surface points to the reference point are of the form (r, β, γ), the r-table now stores triples instead of 2-tuples.

Computation of the r-table is achieved by first choosing a reference point $(x_{\mathrm{ref}}, y_{\mathrm{ref}}, z_{\mathrm{ref}})$ for the surface, again typically taken to be the *center-of-mass*. Next, each surface point, s_i, is visited at which the orientation of the unit normal, $\hat{\mathbf{n}}_i$, is calculated. Using this orientation to index into the r-table, the vector (r_i, β_i, γ_i), joining s_i to the reference point is added to that position.

To calculate the required surface normal we proceed as follows. Let $s(x, y)$ denote the considered surface and, let p and q denote the partial derivatives of this surface in the x and y directions, respectively, *i.e.* $p = \frac{\partial}{\partial x}s(x_i, y_i)$ and $p = \frac{\partial}{\partial y}s(x_i, y_i)$. Then in Cartesian coordinates the normal vector may be calculated as [3],

$$\hat{\mathbf{n}}_i = \frac{\mathbf{n}_i}{||\mathbf{n}_i||} \tag{20}$$

$$= \frac{(-p_i, -q_i, 1)^T}{\sqrt{1 + {p_i}^2 + {q_i}^2}} \tag{21}$$

where, $||\mathbf{n}_i||$ denotes the magnitude of the vector $\mathbf{n}_i$. In keeping with convention, and to allow easier manipulation in what follows, we will specify $\hat{\mathbf{n}}_i$ in terms of a spherical polar coordinate system, with origin at the feature point, azimuth angle ϕ parallel to the XY-plane, and polar angle θ taken to be the angle between the vector and the Z-axis (see Fig. (4)).

These angles may be specified, in terms of the Cartesian formulation, as follows,

$$\phi_i = \tan^{-1}\left(\frac{\hat{\mathbf{n}}_i.\hat{\mathbf{e}}_x}{\hat{\mathbf{n}}_i.\hat{\mathbf{e}}_y}\right) \tag{22}$$

$$= \tan^{-1}\left(\frac{-p_i}{-q_i}\right) \tag{23}$$

$$\theta_i = \cos^{-1}\left(\hat{\mathbf{n}}_i.\hat{\mathbf{e}}_z\right) \tag{24}$$

$$= \cos^{-1}\left(\frac{1}{\sqrt{1+{p_i}^2+{q_i}^2}}\right) \tag{25}$$

where, $\hat{\mathbf{n}}_i.\hat{\mathbf{e}}_j$ denotes the inner product of the vectors $\hat{\mathbf{n}}_i$ and $\hat{\mathbf{e}}_j$, and $\hat{\mathbf{e}}_x$,$\hat{\mathbf{e}}_y$ and $\hat{\mathbf{e}}_z$ denote unit vectors along the X, Y, and Z axes, respectively.

Equations (26), (27), (28) and, (29) may be used to calculate (r, β, γ), the vector joining the surface point, s_i, to the reference point, $(x_{\text{ref}}, y_{\text{ref}}, z_{\text{ref}})$,

$$s_i = (x_i, y_i, z_i) \tag{26}$$

$$r_i = \sqrt{(x_{\text{ref}} - x_i)^2 + (y_{\text{ref}} - y_i)^2 + (z_{\text{ref}} - z_i)^2} \tag{27}$$

$$\gamma_i = \cos^{-1}\left(\frac{z_{\text{ref}}}{r_i}\right) \tag{28}$$

$$\beta_i = \cos^{-1}\left(\frac{(x_i)}{r_i \sin(\gamma_i)}\right) \quad . \tag{29}$$

As with the GHT, the transform stage is simply a reversal of the above process. That is, given a surface $s(x, y)$, for each point, $s(x_i, y_i)$, contained in this surface the orientation of the unit normal, $\hat{\mathbf{n}}_i = (\theta_i, \phi_i)$, is calculated. Using this value to index into the r-table, the (r, β, γ)-triples corresponding to this orientation are extracted. Recovery of the reference point for each triple is carried out as detailed in (18) and (19). Representation of the Hough space may again be achieved through the use of an accumulator array, which, due to the dimensions of the reference point, must in this instance be 3-dimensional.

4.2 Introducing Invariance

Extension of the SNHT to allow invariance to both scaling and orientation, in 3-dimensions, is also possible. Again this pertains to extending the dimensionality of the accumulator array and altering the transform equations sufficiently for each of the invariants required. We refer to this strategy as the orientation and scale invariant SNHT (OSISNHT) and mathematically describe it as

$$H(x_{\text{ref}}, y_{\text{ref}}, z_{\text{ref}}, s, \delta\eta, \delta\zeta, \delta\psi) = \sum_{(x,y,z)\in S} \delta(x, y, z;\ x_{\text{ref}}, y_{\text{ref}}, z_{\text{ref}}, s, \delta\eta, \delta\zeta, \delta\phi) \tag{30}$$

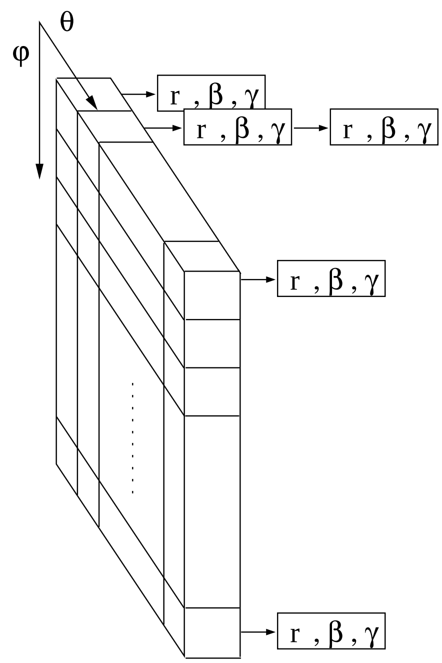

Fig. 5. SNHT r-Table.

where,

$$\delta(x,y,z;\ x_{\mathrm{ref}},y_{\mathrm{ref}},z_{\mathrm{ref}},s,\delta\eta,\delta\zeta,\delta\psi) = \begin{cases} 1 & \text{if } \mathcal{H}(I(x,y,z)) = \\ & (x_{\mathrm{ref}},y_{\mathrm{ref}},z_{\mathrm{ref}},s,\eta,\zeta,\psi) \\ 0 & \text{otherwise.} \end{cases} \tag{31}$$

It can be seen from (31) that three independent rotations are possible. Here, we use the familiar roll-pitch-yaw conventions [6] (see Fig. 6), where roll corresponds to a rotation about the Z-axis, pitch corresponds to a rotation about the Y-axis, and yaw corresponds to a rotation about the X-axis.

In order to compute the OSISNHT, the relationships between the angles described in (22), (24), (28) and (29), and these three rotations described above, are required. Using these relationships, the change in the unit normals for a given surface, resultant from any of the rotations described above, may be calculated. For example, if we have a unit normal $\hat{\mathbf{n}}$, and we rotate it about the X-axis by an arbitrary angle, the resultant vector $\hat{\mathbf{n}}'$ may be computed.

To describe the relationships outlined above, the angles η, ζ, and ψ, are defined as those angles resulting from the projection of $\hat{\mathbf{n}}$ onto the YZ-,XZ-, and XY-planes, respectively (see Fig. 6). Using these angles we may now define roll, pitch, and yaw as the changes $\delta\eta$, $\delta\zeta$, and $\delta\psi$, in the angles η, ζ, and ψ, respectively.

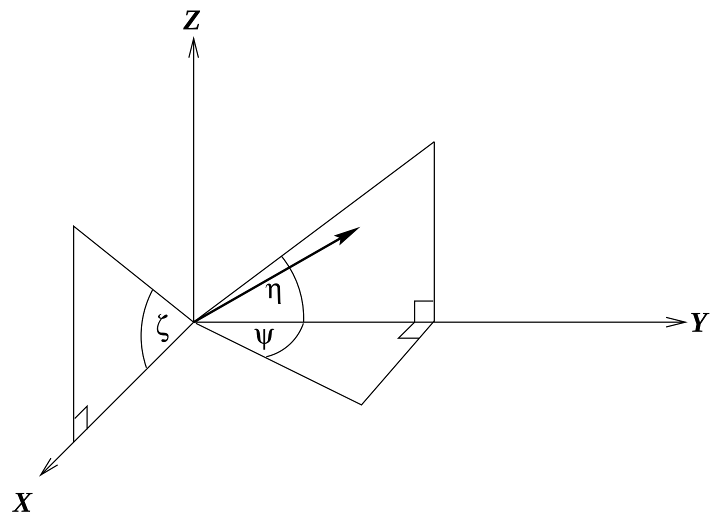

Fig. 6. Roll, Pitch, and Yaw angles

For a given unit normal, $\hat{\mathbf{n}} = (\theta, \phi)$, calculation of the three angles in Fig. 6 may be carried out as,

$$\eta = \tan^{-1}\left(\frac{1}{\tan(\theta)\cos(\phi)}\right) \tag{32}$$

$$\zeta = \tan^{-1}\left(\frac{1}{\tan(\theta)\sin(\phi)}\right) \tag{33}$$

$$\psi = \phi \ . \tag{34}$$

Therefore, the unit vector $\hat{\mathbf{n}}' = (\theta', \phi')$, resulting from a rotation of $\delta\eta$, about the X-axis, may be calculated using,

$$\theta' = \cos^{-1}\left(\cos(\phi)\sin(\eta + \delta\eta)\right) \tag{35}$$

$$\phi' = \tan^{-1}\left(\frac{\tan(\phi)}{\cos(\eta + \delta\eta)}\right) \ . \tag{36}$$

The unit vector, $\hat{\mathbf{n}}' = (\theta', \phi')$, resulting from a rotation of $\delta\zeta$, about the Y-axis, may be calculated using,

$$\theta' = \cos^{-1}\left(\sin(\theta)\sin(\zeta + \delta\zeta)\right) \tag{37}$$

$$\phi' = \tan^{-1}\left(\tan(\theta)\cos(\zeta + \delta\zeta)\right) \ . \tag{38}$$

The unit vector, $\hat{\mathbf{n}}' = (\theta', \phi')$, resulting from a rotation of $\delta\zeta$, about the Z-axis, may be calculated using,

$$\theta' = \theta \tag{39}$$

$$\phi' = \phi + \delta\psi \ . \tag{40}$$

Therefore, computation of the OSISNHT corresponds to calculating the $\mathcal{H}$-mapping over a range of s,$\delta\eta$, $\delta\zeta$, and $\delta\psi$ in a brute force manner similar to that of the OSIGHT described in Sect. 3.2. Again we extend the dimensions of the Hough space accordingly to cater for the the chosen invariance.

Therefore, given a surface point, $s(x_i, y_i, z_i)$, we compute the unit normal, $\hat{\mathbf{n}}_i$, at that point. Using the equations defined above, we then calculate the rotated unit normal $\hat{\mathbf{n}}'_j$ for each possible $\delta\eta_j$, $\delta\zeta_j$, $\delta\psi_j$ combination. Next, for each $\hat{\mathbf{n}}'_j$, the corresponding (r_j, β_j, γ_j) triples are extracted using $R({\theta_j}', {\phi_j}')$.

$$\begin{aligned}\mathcal{H}(I(x,y,z)) &= (x_{\mathrm{ref}}, y_{\mathrm{ref}}, z_{\mathrm{ref}}) \\ &= (x + (s_k.r)\sin(\beta)\cos(\gamma), \\ &\qquad y + (s_k.r)\sin(\beta)\sin(\gamma), \\ &\qquad z + (s_k.r)\cos(\gamma)) \ .\end{aligned} \tag{41}$$

For each triple we also rotate it by $\delta\eta_j$, $\delta\zeta_j$, and $\delta\psi_j$. Then, using the resultant values and the $\mathcal{H}$ equation defined in (41) we calculate the corresponding reference point, $(x_{\mathrm{ref}}, y_{\mathrm{ref}}, z_{\mathrm{ref}}, s, \delta\eta, \delta\zeta, \delta\psi)$ for each s_k over a specified range.

5 Conclusion

Although modifications to the Generalised Hough Transform cater for scaling and rotation of the object in the plane, the GHT fails to detect the object under rotations out of the plane. This is due to the lack of 3-dimensional information contained in the 2-dimensional template used in generating the R-mapping. In this paper our computational model for the Hough transform shows that this is a problem of representation, and that the Hough transform *is* capable of catering for three-dimensions if it is provided with the necessary 2-dimensional surface representation and $\mathcal{H}$ equations. Accordingly, we have proposed a new Hough transform, the Surface Normal Hough Transform, which overcomes the out-of-plane rotation problem by extending the dimensionality of both the input space and transform equations to directly incorporate surface information. Our current work involves enabling the SNHT to detect patterns in 2-dimensional input images, by extracting 3-dimensional surface normals through the use of occluding contours.

References

1. D.H. Ballard. Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognition*, 13:111–122, 1981.
2. R.O. Duda and P.E. Hart. Use of the Hough transform to detect lines and curves in pictures. *Communications of ACM*, 15:11–15, Jan 1972.
3. Berthold Klaus Paul Horn. *Robot Vision*. The MIT Press, 1987.
4. P.V.C. Hough. Method and means of recognizing complex patterns. U.S. Patent 3069654, 1962.

5. P.M. Merlin and D.J. Farber. A parallel mechanism for detecting curves in pictures. *IEEE Transactions on Computing*, C-24:96–98, 1975.
6. R.P. Paul. *Robot Manipulators*, chapter 2, pages 45–47. The MIT Press, 1982.
7. David Vernon. *Machine Vision : Automated Visual Inspection and Robot Vision*. Prentice Hall, 1991.