Efficient Recognition of Rotationally Symmetric Surfaces and Straight Homogeneous Generalized Cylinders

Jane Liu*, Joe Mundy*, David Forsyth*, Andrew Zisserman† and Charlie Rothwell†

(*) GE Center for Research and Development, 1 River Rd, Schenectady, NY 12345.

(⋆) Department of Computer Science, University of Iowa, Iowa City, IA 52242.

(†) Robotics Research Group, Department of Engineering Science, University of Oxford, Parks Rd, Oxford OX1 3PJ, UK.

Abstract

It is known that rotationally symmetric surfaces can be recognized from their outlines alone, using cross-ratio's of bitangent intersections. This paper demonstrates a successful implementation of this technique, using a novel bitangent finder, that works on images of real scenes. We report on the stability of the cross-ratio's, and compare this to affine invariants. The recognition technique is shown to extend to the case of straight homogeneous generalised cylinders.

1 Introduction

This paper continues the work described in [9], where it was shown that surfaces of revolution can be recognised from their outline alone, without knowledge of object pose or camera calibration. The recognition method involves commputing *indexes* [8, 11, 17, 21] from the outline which are used to recognise the object via a hash table. In this paper, we demonstrate successful indexing for a range of images of real objects using the constructions of [9]. We show that these constructions also yield indexing functions for straight homogeneous generalised cylinders.

1.1 Description for recognition

This paper concentrates on the problem of recognising curved surfaces from a single outline. Previous approaches include attempts to extend line labelling [10, 12], the development of constraint-based systems [4], the study of how the topology of a surface's outline changes as it is viewed from different points, formalised into a structure known as an *aspect graph* (for example, [13, 15, 16]), and attempts to represent the system of outlines of a curved surface as a linear combination of some small number of outlines (see, for example, [1, 2]).

Much attention has been focussed on particular classes of surface; straight homogenous generalised cylinders have been particularly popular. Relationships between sections of the outline of a straight homogeneous generalised cylinder have been widely studied, and are known to yield a variety of surface parameters in orthographic views [14, 19, 20]. Dhome *et al.* showed that for a class of rotationally symmetric surface, object pose could be recovered for a *known, calibrated camera*, and incorporated this fact into a recognition scheme [5], which was later extended to include straight homogeneous generalised cylinders [6].

1.2 The outline and its geometry

Throughout the paper we assume an idealized pinhole camera, where all rays pass through the focal point. If the focal point is fixed and the image plane is moved, the resulting distortion of the image is a collineation[1]. In what follows, it is assumed that neither the position of the image plane with respect to the focal point nor the size and aspect ratio of the pixels on the camera plane is known[2], so that the image presented to the algorithm is within some arbitrary collineation of the "correct" image. In this abstract model, the image plane makes no contribution to the geometry, and its position in space is ignored. An orthographic view occurs when the pinhole is "at infinity".

The *outline* of a surface is a plane curve in the image, which itself is the projection of a space curve, known as a *contour generator*[3]. The contour generator is given by those points on the surface where the surface turns away from the image plane; formally, the ray through the focal point to the surface is tangent to the surface. As a result, at an outline point, if the

[1] A collineation is a continuous, one-to-one map taking the projective plane to the projective plane *which maps lines to lines*; any collineation is a plane projective transformation.

[2] These quantities can be measured with varying degrees of difficulty; they do not appear to be particularly stable when cameras are moved, shaken or dropped, however.

[3] There are a number of widely used terms for both curves, and no standard terminology has yet emerged.

relevant surface patch is visible, nearby pixels in the image will see vastly different points on the surface, and so outline points usually have sharp changes in image brightness associated with them.

1.3 Indexing rotationally symmetric objects

It is shown in [9] that the intersections of corresponding pairs of lines, bitangent to the outline, are projections of points on the axis of the object. These points are defined solely by the geometry of the object, so that finding these points and forming their cross-ratio's yields a set of geometric invariants of the object which can be measured from image information alone.

Thus, cross-ratio's of intersection points of corresponding bitangent lines *yield indexing functions for rotationally symmetric surfaces.* Note, in particular, that these cross-ratio's are invariant to camera calibration, and so can be used with an unknown camera. The rest of the paper shows how these cross-ratio's can be computed from image data using a novel bitangent finder, that they can be used successfully for recognition, and that this technique will work for straight homogenous generalised cylinders as well.

2 A recognition system using cross-ratio's

A recognition system using cross-ratio's as indexing functions works as follows:

- cross-ratio's are constructed for corresponding pairs of bitangents in an image;
- these cross-ratio's are used as keys to a hash-table that contains the correspondence between surfaces and cross-ratio's to yield recognition hypotheses;
- the recognition hypotheses are tallied, verified and accepted or rejected.

In our existing system, we do not verify recognition hypotheses, as edge-based verification for curved surfaces is difficult without pose information, which is not available. The system's model-base contains three surfaces, and it is assumed that there is only one surface in each image to simplify the computation of corresponding bitangents.

The main step is computing cross-ratio's from outlines. This process requires that:

1. all bitangents to the outline be found, and
2. corresponding bitangents identified and intersected.

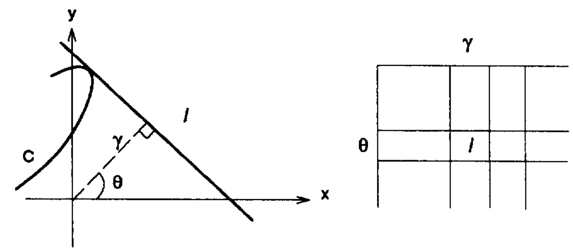

Figure 1: A Hough transformation table at the right is created with θ and γ indices. A tangent line to curve C is represented by θ and γ and stored in the cell (θ, γ) in the Hough transformation table.

2.1 Finding bitangent lines to a curve

A tangent line can be represented by Hough transformation as $l(\theta, \gamma)$, where θ is the orientation of the line and γ is the distance from the image center to the line, as shown in figure 1. Any line in the image can be mapped into a particular cell in the Hough transform table by its location and its orientation. If tangent lines derived from two different points fall into the same cell in the Hough transformation table, then those tangent lines are a bitangent line. The process of finding bitangents proceeds, therefore, by:

1. computing the tangents to the curve and Hough transforming these lines;
2. checking the Hough transformed system for cells containing more than one line, which are bitangents.

2.1.1 Computing and Hough transforming tangents

Tangent lines are computed using an eigenvector line-fitting method [7]. As shown in the left half of figure 2, $\hat{l}$ is the best eigenvector fitting line based on the 7 points around P_i. In the experiments, we used an 11 point neighbourhood. The tangent line at P_i is the line passing through P_i and parallel to $\hat{l}_i$, labelled l_i in the figure.

When the curve has high curvature, the θ and γ values of consecutive tangent lines can be quite different because of the sample spacing, with the result that two consecutive tangents can be marked in cells some way apart in the hough space. As a result, bitangent lines can be missed, because high-curvature segments of curves can lead to widely scattered points in the Hough space, which may not intersect properly (see the right half of figure 2 for an example).

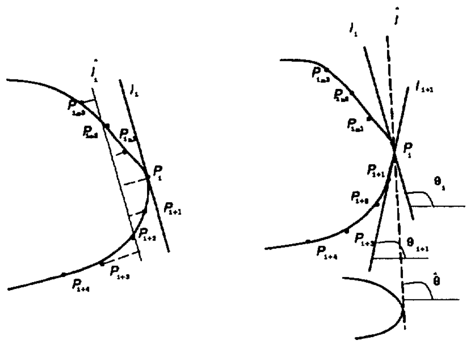

Figure 2: In the left half of the figure, an eigenvector fitting line $\hat{l}_i$ is constructed from seven points $P_{i-3}, P_{i-2}, P_{i-1}, P_i, P_{i+1}, P_{i+2}$, and P_{i+3} and the tangent line at point P_i is l_i, which passes through P_i and is parallel to $\hat{l}$. In the right half, $\hat{l}$ is a bitangent line which is missed because $\hat{l}$ is not booked in the Hough transformation table while scanning P_i and P_{i+1}.

The solution to this problem is to interpolate between points in the Hough space, using either a linear or quadratic interpolate, depending on the variation in θ (for our experiments we used a quadratic interpolate if $\Delta\theta > 6^o$, and otherwise a linear interpolate). This strategy leads to continuous curves in the Hough space, and is successful in finding bitangents; examples are shown in figures 3 and 4.

2.2 Determining corresponding bitangents

Once all bitangents have been found, it is necessary to determine which pairs of bitangents correspond (i.e. both come from the same cone of bitangents). This problem can be solved by exploiting the following remarkable symmetry property of rotationally symmetric surfaces:

> **Theorem:** There is a non-trivial plane projectivity which maps the outline of a rotationally symmetric surface to itself. The contour generators corresponding to each half are, in general, space curves, and are related by a mirror symmetry in space.

In effect, this theorem is a stronger way of stating that the outline of a rotationally symmetric surface can be separated into two sides, which are related by plane projectivity. To see that the two sides of the contour in the image are projectively equivalent, for an arbitrary view, construct the plane containing the axis of the surface and the focal point. The surface then has a mirror symmetry in this plane, as does the cone of rays through the focal point and tangent to the surface. This cone yields the outline when it is intersected with the image plane.

If the image plane is perpendicular to the plane of symmetry, then the outline has a mirror symmetry; but the outline in any other image plane is within a projective map, say T of this outline (by construction, with the focal point as the centre of projection), and so we can construct a non-trivial projective mapping that takes the outline to itself as $P = T \circ M \circ T^{-1}$, where M is a mirror symmetry. Since by construction T is a projectivity, and M is a projectivity (it can be given as $diag\ [1, -1, 1]$), P is a projectivity.

This delivers a uniform method for determining points lying on the projection of the 3D symmetry axis. Any projectively covariant construction[4] in the particular (symmetric) image plane which generates points on the image axis, can be used in any image. Examples of such constructions include:

1. the intersections of lines connecting corresponding pairs of distinguished points on each side of the outline. For example, given a corresponding to a', b corresponding to b', the lines ab', $a'b$ intersect on the symmetry axis. Appropriate distinguished points are covariants such as points of contact of bitangents or of inflections.

2. the projective transformation that maps the contour to itself. The projection of the axis will be a line of fixed points of this transformation.

In practice, we use approach 1 (above) with distinguished points derived from a bitangent's contact with the curve. Any pair of corresponding bitangents then generates two points on the axis image: one by the intersection of the bitangents (i.e.the lines ab and $a'b'$), the other by the *cross-construction* above (i.e. the lines ab', $a'b$). This is a simple and successful construction. Note that the order of the points of tangency on each bitangent can be given with reference to their intersection point and so is uniquely defined.

Now, select any two bitangent lines in the image. We give a vote to line c, from both their intersection and cross-construction. The total number of votes for the correct image of the central axis, n, is the number

[4]By this, we mean that we would obtain the same result if we were to perform the construction in one frame, and then project the result to a new frame, or if we were to perform the construction in the new frame on a projection of the original curves; constructions with this property are based around incidence and counting properties. For example, a tangent line is a covariant construction.

angle	cross-ratio	length ratio
45.0	0.486187	1.40862
40.0	0.490561	1.98153
35.0	0.486796	2.14017
25.0	0.486640	2.38409
15.0	0.486260	2.70539
0.0	0.494849	4.13687

Table 1: Stability of invariants, computed to six digits from measured points, with inclination of the axis of a lamp-base to the camera plane. Typical affine and projective invariants are shown. Note that the value of the affine invariant changes at extreme angles, whereas the first two digits of the projective invariant remain wholly unaffected.

of distinguished bitangent cones constructed by the shape of the object. The total number of votes for each incorrect image of the central axis clearly must be 1 or small if the surface is not degenerate, and so the line with maximum number of votes is the image of the real central axis. This voting system is refined further by noting that, for real views, it is extremely hard to arrange the camera such that lines joining corresponding pairs of distinguished points are more than a few degrees away from parallel, as a result of the limited field of view of the camera. Currently, pairs of bitangents where these lines are more than 4^0 off parallel do not contribute to the vote. We return to this restriction in section 4.

2.3 Results

Table 1 demonstrates the stability of invariants as the inclination of the object axis to the viewer is increased. In the case of cross-ratio's, which are a projective invariant, the values only alter significantly when the image segmentation fails, and it is no longer possible to determine tangents. In contrast, a simple ratio of distances along the axis, which is an affine invariant, varies systematically with inclination angle. This is the expected behaviour since affine approximation to projection deteriorates as object depth increases relative to distance from the camera. However, this indicates that affine invariants are not acceptable as indexing functions for unrestricted object pose.

A recognition system has been built that can identify one of three possible objects: a doorknob, a lamp and a stand, using projective invariants alone. Each image contains only one object, and the task is to identify which.

Recognition proceeds by computing all possible

view	"doorknob"	ambiguous	other	miss
1	3	0	1	0
2	4	0	1	0
3	4	0	1	0
4	4	0	1	0

Table 2: Hash-table returns for four views of a doorknob. A hash table was preloaded using a fifth view of the doorknob, and with views of two other surfaces, with the cross-ratio's acting as keys, and the name of the relevant surface inserted into the table. The hash-table was then accessed using the cross-ratio's measured in each image; each return is counted as a vote for the surface or surfaces retrieved. The columns show the label returned from the hash-table; the alternatives are the correct label alone, a number of labels including the correct label ("ambiguous"), a collection that does not include the correct label, and nothing retrieved at all. The single persistently incorrect return occurs because not every cross-ratio could be measured from the view used to load the hash-table. Nonetheless, the surface is clearly identified in each case by choosing the return with the most votes.

cross-ratio's of bitangent intersections from an image, rounding these values to a single digit, and using them as a key to a hash-table, which was preloaded with the names of the surfaces, using cross-ratio's computed from one image of each surface. In particular, for the stand and the doorknob, a number of cross-ratio's could be computed from each image, and the final identification was made by voting for the object with the greatest number of returns. Table 2 gives typical results. Note that the technique described is showing a degree of robustness, as surfaces are correctly identified despite the differing number of cross-ratio's computed for each image as a result of noise-related difficulties in obtaining all bitangents. In no views, of a total of 15, was the final identification incorrect.

3 Indexing straight homogeneous generalised cylinders

A straight homogeneous generalised cylinder (SHGC) can be defined as a surface that, in some Euclidean frame, can be parametrised as:

$$(f_1(t)g_1(s), f_1(t)g_2(s), f_2(t))$$

Thus, in the appropriate frame, the sections of this surface corresponding to planes $z = constant$ are uniformly scaled copies of the plane curve $(g_1(s), g_2(s))$.

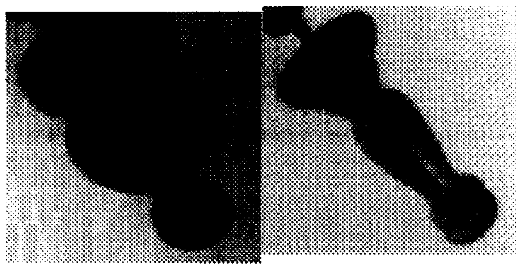

Figure 3: Typical images of real rotationally symmetric objects, used to obtain the recognition results; the left figure shows the knob, the right figure shows the stand.

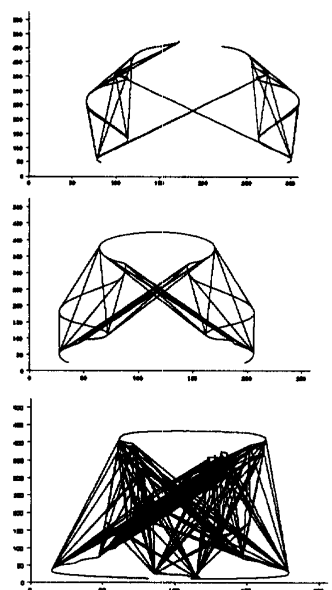

Figure 4: This figure shows all the constructed bitangent lines and the outlines of the images of three samples: (a) a lamp, (b) a knob and (c) a candle stand.

As a result, in this frame the z-coordinate axis forms an "axis" for the surface, which has a similar role to the axis of a rotationally symmetric surface.

Now consider the family of planes through this axis; an arbitrary plane from this family is given by $ax + by = 0$, for some a, b. In coordinates in this plane, the intersection between the surface and the plane can be given by:

$$(\lambda f_1(t), f_2(t))$$

where λ is a function of s. In particular, only λ changes as we move from plane to plane in the family. We have:

> **Lemma:** The envelope of the family of planes tangent to the surface along a curve of fixed t (a "parallel"), is a cone or a cylinder.

The lemma is proven by noting that every tangent plane in this family intersects the z-axis in the same point; this, in turn is proven by showing that the y-intercept of a line tangent to a curve of the form $(\lambda f_1(t), f_2(t))$ is the same for any $\lambda \neq 0$. Note that the cones or cylinders are also SHGC's, with the z-axis as their "axis" and with the same cross-section as the surface.

From this lemma, we have immediately:

> **Lemma:** Families of planes bitangent to SHGC's form cones, with their vertices on the axis of the surface.

Thus, we can construct indexing functions for SHGC's in exactly the same way as we constructed indexing functions for rotationally symmetric surfaces.

4 Invariants and Quasi-Invariants

There is an apparent inconsistency between the method employed in section 2.2 to match corresponding points on each side of the outline, and the results in section 2.3. In the former it is assumed that an affine property, parallelism, is approximately preserved during projection, yet in the latter it is demonstrated that affine invariants degrade systematically as the inclination of the 3D symmetry axis increases.

Figure 5 illustrates that the significant angle in terms of the affine approximation to projection is not the inclination of the axis, but the angle between the plane containing the axis and focal point, and the image plane normal, α. It can be shown that Δ/D is proportional to $\tan\alpha$. For a finite image plane, α is bounded by the finite field of view. Consequently, this limits the "extent" of non-parallelism of lines joining corresponding points on the outline. This approximate parallelism for finite image planes may be thought of as a *quasi*-invariant [3].

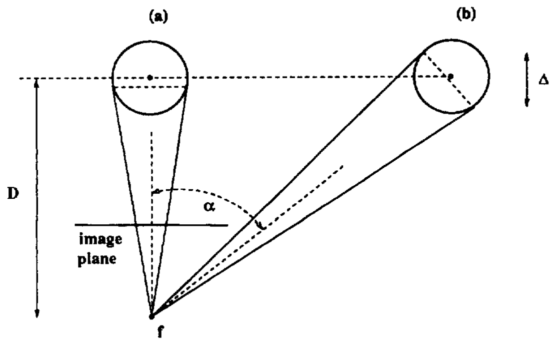

Figure 5: Schematic plan view of a surface of revolution and tangents from the optical centre. α is the angle between the plane containing the axis of the surface and the focal point, and the image plane normal. In (a) lines joining corresponding points on the contour generator are parallel to the image plane. Consequently, their image projections are parallel. This holds even when the axis of the surface is *not* parallel to the image plane. In (b) the object has the same depth but is translated parallel to the image plane. In this case lines joining corresponding points do not project to parallel lines.

5 Discussion

We have demonstrated that rotationally symmetric surfaces can be successfully indexed using bitangents computed automatically from image edges by a bitangent finding algorithm which we have described. We have shown that this approach can be extended to recognise straight homogeneous generalised cylinders.

We are currently investigating the second approach to bitangent finding, namely determining the projective transformation between "sides" of the outline. This is providing an efficient filter for restricting correspondences.

Acknowledgements

Supported in part by the National Science Foundation under award no. IRI-9209729, in part by General Electric, and in part by Esprit Project VIVA.

References

[1] Basri, R. and Ullman, S. "The alignment of objects with smooth surfaces," *Proc ICCV 2*, Tarpon Springs, 1988.

[2] Basri, R. and Ullman, S. "Recognition by linear combination of models," *IEEE PAMI*,

[3] Binford, T.O., Levitt, T.S., and Mann, W.B., "Bayesian inference in model-based machine vision," in Kanal, L.N., Levitt, T.S., and Lemmer, J.F. (ed.s), *Uncertainty in A.I. 3*, Elsevier, 1989.

[4] Brooks, R. A., "Model-Based Three-Dimensional Interpretations of Two Dimensional Images," *IEEE PAMI*, **5**, 2, p. 140, 1983.

[5] Dhome, M., LaPreste, J.T, Rives, G., and Richetin, M. "Spatial localisation of modelled objects in monocular perspective vision," *Proc. First European Conference on Computer Vision*, O.D. Faugeras (ed.), Springer LNCS-x, 1990.

[6] Dhome, M., Glachet, R. and LaPreste, J.T., "Recovering the scaling function of an SHGC from a single perspective view," *Proc. CVPR-92*, 1992.

[7] Duda, R. O. and Hart, P. E., "Pattern Recognition Classification and Scene Analysis", a Wiley-interscience publication 1973.

[8] Forsyth, D.A., Mundy, J.L., Zisserman, A.P., Heller, A., Coehlo, C. and Rothwell, C.A., "Invariant Descriptors for 3D Recognition and Pose," *IEEE Trans. Patt. Anal. and Mach. Intelligence*, **13**, 10, 1991.

[9] Forsyth, D.A., Mundy, J.L., Zisserman, A. and Rothwell, C.A., "Recognising rotationally symmetric surfaces from their outlines," *Proc. Second European Conference on Computer Vision*, G. Sandini (ed.), Springer LNCS-x, 1992.

[10] Freeman, H. and Shapira, R., "Computer Recognition of Bodies Bounded by Quadric Surfaces from a set of Imperfect Projections," IEEE Trans. Computers, **C27**, 9, 819-854, 1978.

[11] Lamdan, Y., Schwartz, J.T. and Wolfson, H.J. "Object Recognition by Affine Invariant Matching," Proceedings CVPR, p.335-344, 1988.

[12] Malik, J., "Interpreting line drawings of curved objects," *IJCV*, **1**, 1987.

[13] Plantinga, H. and Dyer, C. " Visibility, Occlusion and the Aspect Graph," *CS TR 736*, U. Wisconsin, 1987.

[14] Ponce, J. "Invariant properties of straight homogeneous generalised cylinders," *IEEE Trans. Patt. Anal. Mach. Intelligence*, **11**, 9, 951-965, 1989.

[15] Ponce, J. and Kriegman, D.J., "Toward 3D curved object recognition from image contours," in J.L. Mundy and A.P. Zisserman (ed.s) *Geometric Invariance in Computer Vision*, MIT Press, 1992.

[16] Rieger, J. "Global Bifurcation Sets and Stable Projections of Non-Singular Algebraic Surfaces," *Int. J. Computer Vision*, **7** *3*, 1992.

[17] Rothwell, C.A., Zisserman, A.P., Forsyth, D.A. and Mundy, J.L., "Using Projective Invariants for constant time library indexing in model based vision," *Proc. British Machine Vision Conference* , 1991.

[18] Taubin, G. and Cooper, D.B., "Object recognition based on moment (or algebraic) invariants," in J.L. Mundy and A.P. Zisserman (ed.s) *Geometric Invariance in Computer Vision*, MIT Press, 1992.

[19] Ulupinar, F, and Nevatia, R. "Shape from Contour using SHGCs," *Proc. ICCV*, Osaka, 1990.

[20] Ulupinar, F, and Nevatia, R. "Recovering shape from contour for constant cross-section generalised cylinders," *Proc. CVPR*, Mauii, 1991.

[21] Wayner, P.C. "Efficiently Using Invariant Theory for Model-based Matching," Proceedings CVPR, p.473-478, 1991.