[19] R. Nevatia and T.O. Binford, “Description and recognition of complex curved objects,” *Artificial Intelligence*, 8(1):77-98, 1977.

[20] N. Pillow, S. Utcke and A. Zisserman, “Viewpoint-Invariant Representation of Generalized Cylinders Us ing the Symmetry Set”, Proc. British Machine Vision Conference, 1994.

[21] J. Ponce and D. Chelberg, “Finding the limbs and cusps of generalized cylinders,” *International of Computer Vision*, 1:195-210, 1987.

[22] J. Ponce, “Ribbons, Symmetries and Skewed Symmetries,” In *Proceedings of the Image Understanding Workshop*, pages 1074-1079, Massachusetts, 1988.

[23] J. Ponce, D Chelberg and W.B. Mann, “Invariant properties of straight homogeneous generalized cylinders and their contours,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(9):951-966, 1989.

[24] K. Rao and R. Nevatia, “Description of complex objects from incomplete and imperfect data,” In *International Journal of Computer Vision*.

[25] M. Richetin, M. Dhome, J.T. Lapestre and G. Rives, “Inverse Perspective Transform Using Zero-Curvature Contours Points: Applications to the Localization of Some Generalized Cylinders from a Single View,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(2):185-192, 1991.

[26] P. Saint-Marc and G. Medioni, “B-spline contour representation and symmetry detection,” In *First European Conference on Computer Vision*, pages 604-606, Antibes, France, April 1990.

[27] H. Sato and T.O. Binford, “Finding and recovering SHGC objects in an edge image,” *Computer Vision Graphics and Image Processing*, 57(3), pages 346-356, 1993.

[28] S.A. Shafer and T. Kanade, “The theory of straight homogeneous generalized cylinders,” Technical Report CS-083-105, Carnegie Mellon University, 1983.

[29] F. Ulupinar and R. Nevatia, “Shape from contours: SHGCs,” In *Proceedings of IEEE International Conference on Computer Vision*, pages 582-582, Osaka, Japan, 1990.

[30] F. Ulupinar and R. Nevatia, “Perception of 3-D Surfaces from 2-D Contours,” *IEEE Transactions PAMI*, pages 3-18, 15, 1993.

[31] F. Ulupinar and R. Nevatia, “Recovering Shape from Contour for Constant Cross Section Generalized Cylinders,” In *Proceedings of Computer Vision and Pattern Recognition*, pages 674-676. 1991. Maui, Hawaii.

[32] M. Zerroug and R. Nevatia, “Volumetric descriptions from a single intensity image”, to appear in *International Journal of Computer Vision*.

[33] M. Zerroug and R. Nevatia, “Segmentation and 3-D Recovery of SHGCs from a Single Intensity Image,” in *Proceedings of the European Conference on Computer Vision*, pages 319-330, Stockholm, 1994.

[34] M. Zerroug and R. Nevatia, “Quasi-invariant properties and 3-D shape recovery of non-straight, non-constant generalized cylinders”, In *Proceedings of Computer Vision and Pattern Recognition*, pages 96-103. 1993. New York.

[35] M. Zerroug and R. Nevatia, “Using invariance and quasi-invariance for the segmentation and recovery of curved objects”, In Lecture Notes in Computer Scienc 825, Eds J.L. Mundy, A. Zisserman and D. Forsyth, Springer-Verlag, 1994.

[36] M. Zerroug and R. Nevatia, “Segmentation and 3-D recovery of curved axis generalized cylinders in an intensity image”, *IEEE ICPR*, October 1994 (to appear).

[37] M. Zerroug and R. Nevatia, “From an intensity image to 3-D segmented descriptions,” *IEEE ICPR*, October 1994 (to appear).

[38] Mourad Zerroug, “Segmentation and Inference of 3-D Descriptions from an Intensity Image,” Ph.D. dissertation, Institute for Robotics and Intelligent Systems, University of Southern California, 1994.

This work, in conjunction with past one on straight-axis generalized cylinders [33], serve as the basis for the development of a more general system for handling complex composite objects which are made up of several parts. Work on this system is currently being pursued by the authors who have already achieved new results [37]. The results of this work have several applications, most notably in recognition of complex curved objects from a 2-D image.

References

[1] J. Ben-Arie, The probabilistic peaking effect of viewed angles and distances with application to 3-D object recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12, pages 760-774, 1990.

[2] I. Biederman, "Recognition by components: A theory of human image understanding," *Psychological Review*, 94(2):115-147, 1987.

[3] T.O. Binford, "Visual perception by computer," *IEEE Conference on Systems and Controls*, December 1971, Miami.

[4] T.O. Binford, T.S. Levitt and W.B. Mann, "Bayesian inference in model-based machine vision," Proceedings of *AAAI Uncertainty Workshop*, 1987.

[5] T.O. Binford and T.S. Levitt, "Quasi-invariants: Theory and Exploitation," In *Proceedings of the Image Understanding Workshop*, pages 819-829, Washington DC, 1993.

[6] R.A. Brooks, "Model-based three dimensional interpretation of two dimensional images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 5(2):140-150, 1983.

[7] J. B. Burns, R.S. Weiss and E.M. Riseman, "View Variation of Point-Set and Line Segment Features", *IEEE Transactions PAMI*, 15, pages 51-68, 1993.

[8] A. Gross and T. Boult, "Recovery of generalized cylinders from a single intensity view," In *Proceedings of the Image Understanding Workshop*, pages 557-564, Pennsylvania, 1990.

[9] R. Horaud and M. Brady, "On the geometric interpretation of image contours," *Artificial Intelligence*, 37:333-353, 1988

[10] Q. Huang and G. C. Stockman, Generalized Tube Model: Recognizing 3D Elongated Objects from 2D Intensity Images, In *Proceedings of IEEE Computer Vision and Pattern Recognition*, pages 104-109, 1993.

[11] D.A. Huffman. Impossible Objects as Nonsence Sentences. Machine Intelligence 6, B. Meltzer and D. Michie (eds.), pages 295-323, Edinburgh University Press, Edinburgh 1971.

[12] J.J. Koenderink. *Solid Shape*. M.I.T. Press, Cambridge, MA. 1990.

[13] J. Liu, J. Mundy, D. Forsyth, A. Zisserman and C. Rothwell, "Efficient recognition of rotationally symmetric surfaces and straight homogeneous generalized cylinders," In *Proceedings of IEEE Computer Vision and Pattern Recognition*, pages 123-128, 1993.

[14] J. Malik, "Interpreting Line Drawings of Curved Objects," *International Journal of Computer Vision*, 1(1):73-103, 1987.

[15] R.S. Millman and G.D. Parker, "Elements *of differential geometry,*" Prentice Hall. 1977.

[16] R. Mohan and R. Nevatia, Perceptual Organization for Scene Segmentation, *IEEE Transactions PAMI*. 1992.

[17] *Geometric Invariance in Computer Vision*. J.L. Mundy and A. Zisserman editors, MIT Press, 1992.

[18] V. Nalwa, "Line drawing interpretation: Bilateral symmetry," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:1117-1120, 1989.

not affected the final results of the system. The only effect was a larger search space, and longer run times, for looser parameter values.

scene	boundaries	parallel symmetries	local patches	parts hypotheses	verified parts
fig. 21	424	1237	43	32	3
fig. 27	831	1634	48	41	2
fig. 28	1661	7567	132	110	2

Table 6.1Evolution of the number of hypotheses of the PRGC module

scene	parallel symmetry	local patch detection	grouping of local patches	verification of parts
fig. 21	0:30	0:33	0:07	0:39
fig. 27	0:46	0:29	0:09	1:00
fig. 28	9:05	2:39	0:30	5:51

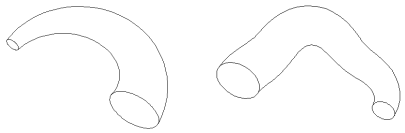
Table 6.2Run times of the PRGC module (min:sec.)

7 Conclusion

This work has derived important properties of the projection of some curved-axis generalized cylinders and a method for their detection and recovery from a single intensity image. The properties derived were shown to be either invariant or quasi-invariant to orthographic projection. The invariant properties of circular PRCGCs extend the set of properties previously derived in the literature for PRCGCs. Circular PRGCs have not been addressed in the past and their quasi-invariant properties derived in this paper constitute a progress towards handling more complex shapes. They show that in the absence of strict invariance, much can be done by exploiting less strict but well-behaving and equally effective properties.

The method described exploits the properties to effect segmentation and recover descriptions of these generic shapes from a single intensity image which may include noise, surface markings and partial occlusion. Satisfactory results were obtained. Unlike past work on ribbon-based description, our method takes explicit account of the 3-dimensionality of the desired descriptions. We then have shown that the 3-D descriptions do not vary much (up to some variations for degenerate regions of the viewing sphere and object shapes), an important characteristic for 3-D scene analysis. The resulting 3-D descriptions can be directly used for recognition based on their symbolic attributes. Although we have addressed only PRCGCs and circular PRGCs, we believe that the methodology is promising to handling other shapes as well.

ones by our method from the 2-D boundaries of figure 32.a. The results do “look” similar to the ground truth. A quantification of the “similarity” is possible and we omit it here (a detailed analysis is given in [38]).



a. original (synthetic) boundaries

b. ground truth axes and scaling functions

c. recovered axes and scaling functions

Figure 32 *Synthetic objects (a), their 3-D descriptions (b) and the recovered 3-D descriptions (c).*

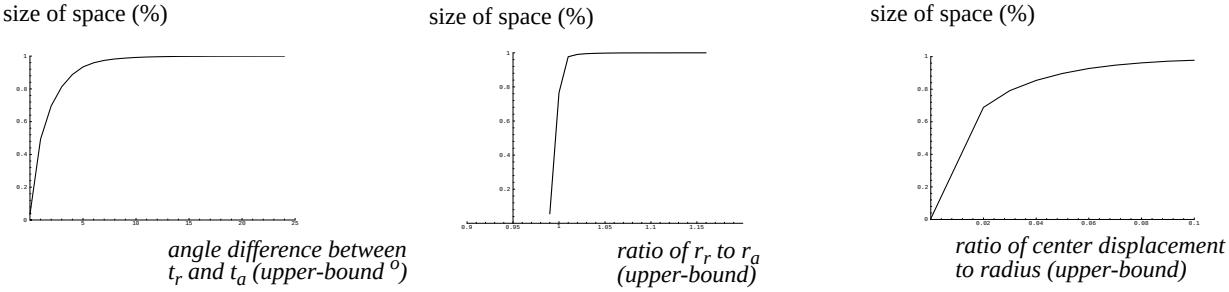
However, these measures cannot always be used because ground truth is in general not known and the images could, in principle, have projected from an infinite number of 3-D objects. The evaluation of the interpretation is a problem with all perceptual tasks where the problems being solved are usually under-constrained. Humans do make assumptions about viewed objects and sometimes make mistakes. In our system, the results appear to be consistent with our (human) perception of shape from the same input contours. A more adequate measure would be one which compares the results with human perception.

6.2 Performance

The tables below summarize the performance of our system (implemented in Sun Common LISP, under the CME environment on a Sparc 10/30) on the images given. Table 6.1 shows the evolution of the number of hypotheses and table 6.2 the corresponding run times. The implementation uses a number of parameters (thresholds) to measure the different projective properties. For all the images, these parameters have been *constant*. However, the performance of the system has been tested by changing the values of key parameters by up to 50% of their default values. Those changes have

directly indicate how the inferred 3-D descriptions vary (as functions of the viewing and shape parameters) due to these approximations.

We can analyze the above variation due to the approximation of the projection of the 3-D axis orientation by a vector orthogonal to the cross-section segment (this analysis can also be thought of as an analysis of the sensitivity of 3-D shape to small perturbations in image measurements of the axis orientation). We can analyze the variation of the 3-D orientation of the axis, the position of the 3-D axis point on the axis plane and the radius (scaling) of the recovered cross-section, assuming some known values of these elements. Figure 31 shows the three plots that resulted. Figure 31.a shows the plot of the size of the parameter space for which the angle between the "actual" axis orientation, t_a, and the recovered one, t_r, is within several upper-bounds (for example, the 3-D angular variation is 5^o or less over 93.4% of the space). Figure 31.b shows the plot of the size of the parameter space for which the ratio of the recovered cross-section radius r_r to the "actual" cross-section radius r_a is within different bounds (for example, the ration is within 2% of r_a over 99% of the parameter space). Figure 31.c shows the plot of size of the parameter space for which the ratio of the distance between the recovered 3-D axis point p_r and the "actual" 3-D axis point p_a over the cross-section radius is within different bounds (for example, this distance is within 5% of the cross-section radius over 89.6% of the space; the ratio measure the relative closeness of the 3-D points).



a. axis (cross-section) orientation

b.cross-section radius

c. axis position

Figure 31 Plots of size of the parameter space for which the recovered 3-D description is within different bounds of the actual one.

These plots show that the recovered 3-D descriptions do not vary much; the displacements tend to be larger only close to degenerate regions of the parameter space, namely where limbs do not exist or high values for both the thickness ratio and the sweep derivative (non-common shapes).

We have also compared the results of the 3-D recovery of circular PRGCs on objects with known ground truth. Figure 32.a shows the objects used for the analysis. Figure 32.b shows the ground truth 3-D axis and scaling functions of the two objects and figure 32.c shows the recovered

constructed shapes in terms of cross-sections, meridians and 3-D axes. The gaps in some parts are due to the image discontinuities such as (self)occlusion and contour breaks.

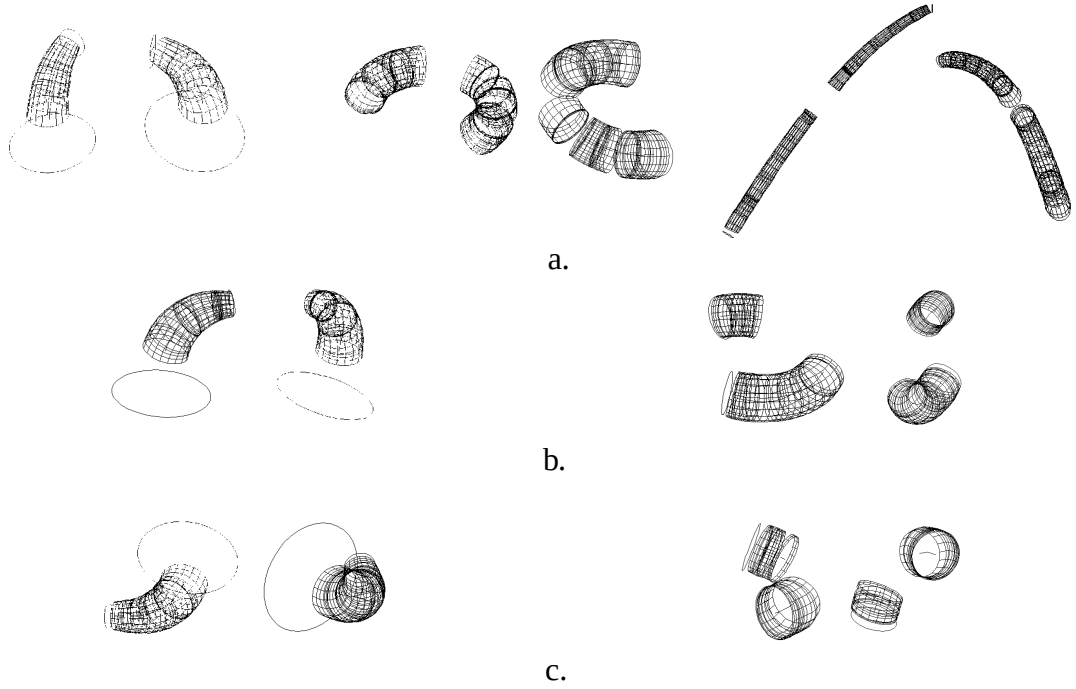


Figure 30 Recovered 3-D primitives shown for their original poses and different ones. a. from the detected objects of figure 26; b. from the detected objects of figure 27; c. from the detected objects of figure 28.

6 Analysis of the Results

6.1 How good are the resulting 3-D descriptions?

An issue of interest is how to evaluate the “goodness” of the resulting 3-D descriptions. There are two aspects of this “goodness” measure. First, we would like to know whether the reconstruction is sensitive or not to the image appearance of the part, or the viewing direction. The second, aspect has to do with how close the resulting descriptions are to the “actual” ones; i.e. the ground truth.

If the part is a PRCGC then the use of the invariant Property 1 guarantees the non-sensitivity of the resulting 3-D descriptions to the viewing direction (up to projection, noise and quantization errors). For a circular PRGC, the quasi-invariant Properties 5 and 6 indicate that the image descriptions are “good approximations” of the projective ones over “most viewing directions” but do not

5.2 Inference of the 3-D axis and the scaling function

The method we describe below has some similarities with the method of [31] and consists of finding for each pair of co-cross-sectional points the position and orientation of the 3-D cross-section which passes through them while being internal to the surface.

Let p_1 and p_2 be two co-cross-sectional points of the part's surface and $\vec{t}_1$ and $\vec{t}_2$ be the respective tangents to the part's boundaries at those points. Let $\vec{w}$ be the 2-D axis tangent for the cross-section segment $p_1 p_2$ (see figure 29). The 3-D cross-section which passes through p_1 and p_2 and which is tangential to the part's (limb) boundary is found as follows:

Step 1. Back-project $\vec{w}$ on the 3-D axis plane to obtain the orientation of the 3-D cross-section, say $\vec{W}$.

Step 2. In 3-D, rotate a reference end cross-section, say C_1, by the rotation $R(\vec{n}_1, \vec{W})$ which rotates $\vec{n}_1$ to $\vec{W}$. Call the resulting cross-section C_r and its projection c_r.

Step 3. Find, on c_r, the two points whose tangents are the same as the tangents $\vec{t}_1$ and $\vec{t}_2$ to the projected limbs at p_1 and p_2. Call them q_1 and q_2.

Step 4. Compute the scale r of the desired cross-section with respect to the reference one as $r = dist\ (p_1, p_2)\ /\ dist\ (q_1, q_2)$. This comes from the fact that lengths ratio of parallel segments is an orthographic invariant.

Step 5. Scale c_r by r and translate it so that p_1 and q_1 coincide. Call the resulting cross-section c_p.

Step 6. Back-project c_p so that its center lies on the 3-D axis plane. This results in the desired 3-D cross-section C_p whose radius is $r\ r_1$ where r_1 is the radius of C_1.

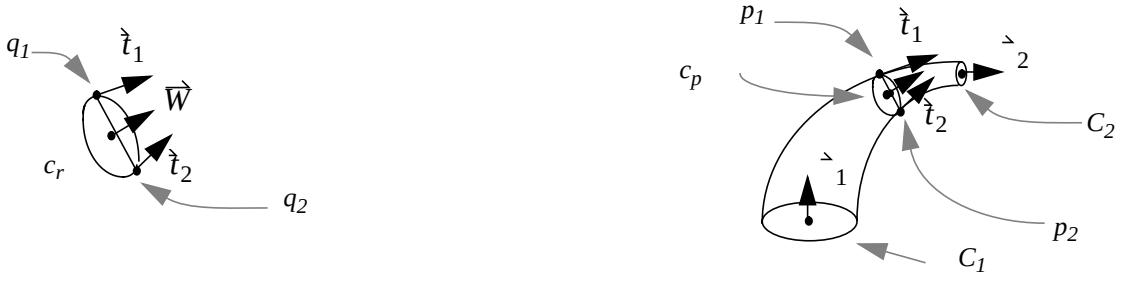


Figure 29 Inferring 3-D descriptions of curved-axis objects.

This method is applied to all component local surface patches of a curved-axis part. Its results on the part's hypotheses of figures 26, 27 and 28 are shown in figure 30. The figure shows the re-

positions with respect to the viewer (facing towards or away, if visible). This step attempts to recover the 3-D descriptions for such parts. The 3-D inference method we use is inspired from the method of [31]. This latter uses constraints on the surface orientations at the visible portion of the surface of a PRCGC; i.e. computes a viewer-centered description. The 3-D reconstruction of circular PRGCs has *not* been previously addressed in the literature. The method we discuss here directly computes the 3-D intrinsic (object-centered) descriptions, which consist of the cross-section, its axis and its scaling function of an object, and applies to both PRCGCs and circular PRGCs.

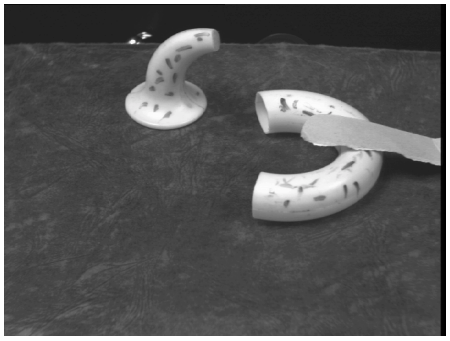
5.1 Inference of the 3-D end cross-sections and the axis plane

To infer the 3-D end cross-sections requires estimating the orientations of their supporting planes (we consider the general case of an unknown cross-section shape). For this, the method of [29,31] based on ellipse fitting is used. That method consists of selecting among two elliptic fits, the least squares and the scatter fits, the one which minimizes the error of fit. The orientation of the 3-D circle(s) that project(s) onto the fitted ellipse is used as an estimate of the cross-section plane orientation. For a given curve, there are in principle two possible orientations (one the *necker* reverse of the other). The selection among these uses the classification of the cross-section orientation as facing towards the image plane of away from it, which was made at the part closure verification step. This results in a unique solution for each cross-section.

Calling $\hat{n}_1$ and $\hat{n}_2$ the resulting plane orientations of the two extremal cross-sections C_1 and C_2, the axis plane orientation is given by $\hat{n}_a = \hat{n}_1 \times \hat{n}_2$, because the axis is orthogonal to the cross-sections. Its position (depth) can be arbitrarily fixed. The 3-D positions of the two end cross-sections (their planes) are automatically fixed since their centers belong to the axis plane.

Each end cross-section is back-projected onto its plane to obtain its 3-D shape. If at least one of the end cross-sections is complete (facing towards the camera for e.g) the 3-D cross-section shape is complete and a full 3-D recovery is feasible (discussed next). If both end cross-sections are incomplete (facing away from the camera, for example) then the 3-D recovery may not be possible. For this, some way of inferring the cross-section shape is needed. This could be done by assuming particular shapes, such as circular, whenever consistent. An instance, is the use of the ellipse fit error to classify whether the 3-D cross-section is (likely to be) circular or not (using normalized errors and fixed thresholds, the details of which we omit). In case it is circular (i.e. the visible portion is elliptical) then the recovered 3-D circle is used as the complete 3-D cross-section.

Figure 26 Verified part hypotheses from the surface patches of figure 21.



a. b. c.

Figure 27 Additional example of resulting object hypotheses for a different scene of similar objects. a. initial intensity image; b. edge image; c. verified part hypotheses.

a. b. c.

Figure 28 Additional example of results of the segmentation method. a. initial intensity image; b. edge image; c. verified part hypotheses.

5 Recovery of 3-D shape

The above method results in verified objects to which are associated *projective descriptions* giving the co-cross-sectional correspondences and information about the cross-sections and their relative

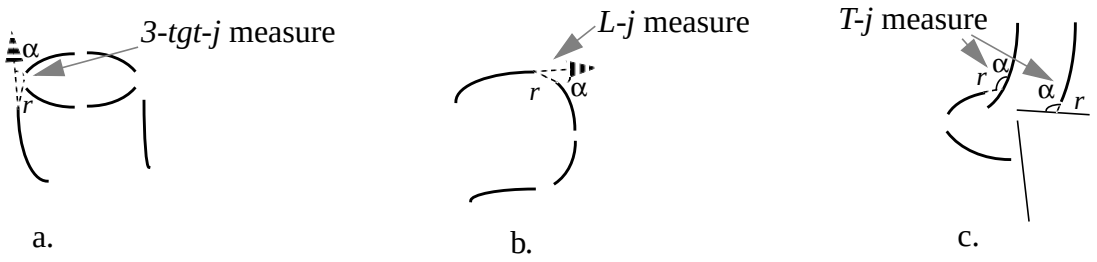


Figure 25 Junction measures.

pairing is one which results in boundary links between the junctions (the cross-section if it is, at least partially, visible), except for the occlusion junctions. Wrong junctions (or their pairings) when found, typically result in no connectivity at all. The search for boundary links uses proximity and orientation continuity constraints between the boundaries, and is organized as a depth first search with a depth bound (of 5 in our experiments). Part hypotheses for which no valid closure is found at either end are rejected.

4.3.2 The inter-part filtering

False part hypotheses may survive the intra-part filtering (regular markings on an object's surface, for example). The inter-part filtering consists of filtering out part hypotheses which are in conflict with "more regular" interpretations. Two parts are in conflict if they share boundaries or the area of one is completely included in the other's. The first filtering we use, exploits the fact that 3-*tgt-j* closure is a strong indicator of regularity as this type of junction is unlikely to have occurred by chance and is a good indicator of the presence of a volumetric shape [2]. Thus, a part with a closure other than 3-*tgt-j* and which conflicts with another part with a 3-*tgt-j* closure is rejected. The second filtering rejects part hypotheses with closures other than 3-*tgt-j* and which are completely included (subsumed) by other part hypotheses.

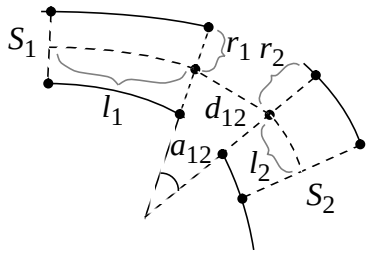
The results of the verification step on the parts hypotheses of figure 21 are shown in figure 26. The figure shows the parts with their surface boundaries and the closures (cross-sections) found for each part. The gaps in the parts are due to missing boundaries either in low contrast regions (e.g. the long straw) or occlusion regions (e.g. the truncated torus occluded by the long straw). Additional results of the method on other images are shown in figures 27 and 28.

Figure 24 Examples of grouping hypotheses from the surface patches of figure 21.b.

4.3.1 The intra-part filtering

The intra-part filtering step consists of verifying that a part hypothesis has one of the closure patterns of section 3.2. The closure verification consists of searching, at the ends of a part, valid junction pairings such as given from the catalogs of figures 15, 16, 17, 18 and 19 (since these catalogs do not account for occlusion by other bodies, the combinations are augmented to include *T*-junctions where the tops of the *T*s are boundaries other than the part's boundaries). This is done by searching for junctions at each end of a part hypothesis (between the part's boundaries and other boundaries in its vicinity) and for connectivity relationships between them. Junctions are found using "junction measures" in order to account for image imperfections. Figure 25 illustrates the measures for the junctions of the catalog we use. The measures account for both proximity and angular variation between boundaries. Proximity is between the extremity of a part's boundary and the junction point (formed with the other boundaries) and the angles are between the tangent at the extremity a part's boundary and the line joining that extremity to the junction point. A junction is one for which the interacting boundaries form proximity and angular variations smaller than fixed thresholds. To detect cusps at the end of a part, we use the property that they are usually accompanied by *T*-junctions [12] where the top of the *T*s are the side boundaries of the part (unlike the other occlusion *T*-junctions where the part's side boundaries are the stems of the *T*s).

To find the object closure, the method first searches for evidence of the cross-section, otherwise for evidence of its absence; occlusion for example. To handle "false junctions" and control the search explosion, the method uses heuristics to select the "best" closures. They include the selection of one junction of each type based on proximity to the part's end, and a ranking of the junctions from the "strongest" to the "weakest". For example, a 3-*tgt-j* is stronger than the *L-j* because the *L-j* structural arrangement is a sub-case of the 3-*tgt-j* arrangement where one branch is removed. We omit further details of this process (which are described in [38]). A successful junction



$$m'_{12} = |r_2 - r_1| / d_{12}$$

$$g_{12} = d_{12} / (l_1 + d_{12} + l_2)$$

Figure 22 Geometric constraints between local curved-axis surface patches.

The structural compatibility constraints impose the structural arrangements between the local surface patches to be consistent with the inner-surface Property 9 of section 3.2; i.e. they either do not spatially overlap or form valid *T*-junctions (such that the top of the Ts, "cusps", lie inside the area bounded by the surface patches and belong to the same patch, if more than one).; figure 23.

a. non-overlapping patches b. self-occluding patches

Figure 23 Structural constraints between local curved-axis surface patches.

At the end of a local surface patch S_1, the selected compatible local surface patch S_2 is the one such that the following measure is *minimized* between both patches among all neighboring structurally compatible ones

$$M_{12} = g_{12}\, a_{12}\, m'_{12}$$

i.e. $M_{12} < M_{1j}\ \forall j\ j \neq 2$ *and* $M_{12} < M_{i2}\ \forall i\ i \neq 1$; where a_{12} is the angle between the extremal cross-section segments as shown in figure 22. Minimizing M_{12} corresponds to favoring small values of the local shape measures (e, $\dot{r}$).

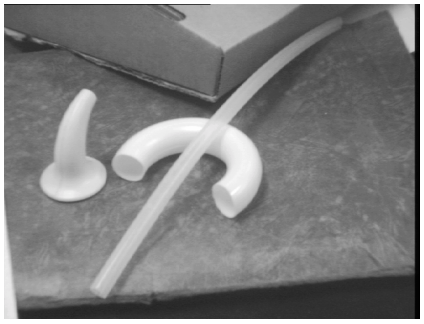
The grouping hypotheses, obtained by our method, for the local surface patches of figure 21.b are shown in figure 24 where compatible patches are labeled by the same number.

4.3 Verification of object hypotheses

Aggregates of local curved-axis surface patches constitute parts hypotheses that are to be verified for global consistency in a two-step process: the *intra-part filtering* and the *inter-part filtering*. Both steps are discussed below.

Minimizing E_1 forces the distance between the local axis and the mid-points of correspondence segments $p_i\ q_i$ to be small; i.e. the local axis should be the 2-D axis as defined in section 3.1. Minimizing E_2 avoids high curvature axes which would correspond to highly curved and thick shapes that are close to self-intersect; λ_1 and λ_2 are real scalars (whose sum is 1). The method gives more weight to E_1 than E_2 as the former *should* be zero for a right ribbon axis and the latter is only a regularity measure (in our system λ_1 is fixed to 0.8 and λ_2 to 0.2). This method has proved much faster, even including the parallel symmetry detection, than the original projection method.

Figure 21 shows some the local curved-axis surface patches detected from a real intensity image. The figure displays cross-section segments and axes (in bold lines). Grouping compatible surface patches together and filtering out false hypotheses are the subject of the next two steps.



a. b. c.

Figure 21 Resulting local surface patches; a. original intensity image; b. edge image; c. detected local surface patches.

4.2 Grouping of local surface patches

The grouping aims at merging local surface patch hypotheses, which are likely to belong to the same object, into a single object hypothesis. It uses a combination of geometric and structural compatibility constraints.

The geometric compatibility constraints are similar in spirit to the methods of [10,24] used for grouping ribbons. Basically, a pair of local surface patches are considered for grouping if their local surface descriptions are “continuous”. More specifically, two local surface patches S_1 and S_2 are geometrically compatible if the relative difference m'_{12} between the sizes of their extremal cross-section segments (see figure 22) and the relative distance g_{12} between their extremities are smaller than fixed thresholds.

sists of searching around the initial correspondences ($p_1 q_1^0$ and $p_2 q_2^0$ in figure 20.*a*) for the "best" right ribbon segments between two successive points (p_1 and p_2). The details are as follows:

- form initial correspondences by intersecting the normal lines at the extremities of each B-spline segment of the parallel symmetry axis with both boundaries (figure 20.*a)*. Select an arbitrary boundary, say B_1, as a reference. Call such initial correspondences $p_i\ q_i^0$

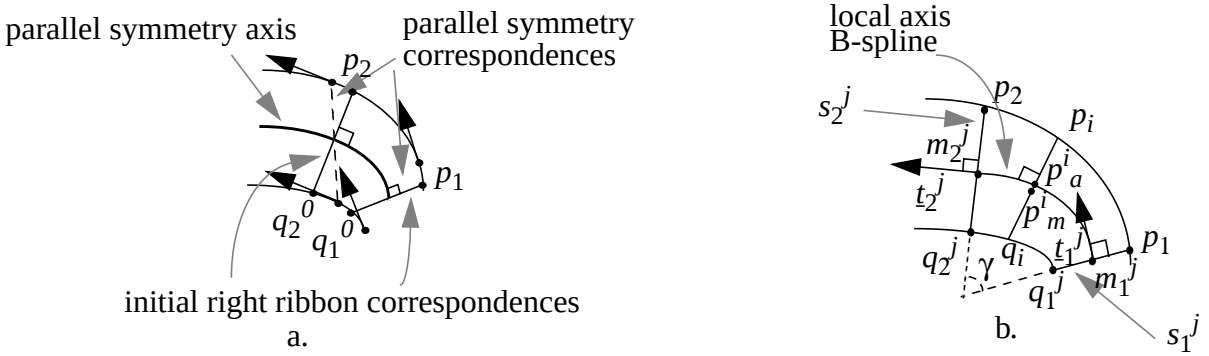


Figure 20 Right ribbon detection. a. parallel symmetry axis as initial estimate of right ribbon axis; b. finding right ribbon correspondences.

- find the "best" corresponding points, q_1 and q_2, for the extremal points p_1 and p_2. This is done by searching around the initial correspondences $p_1\ q_1^0$ and $p_2\ q_2^0$, over an angular region of a selected range $\Delta\theta$ and a step $d\theta$. At each step j of the search, the correspondences are the segments $s_1^j = p_1\ q_1^j$ and $s_2^j = p_2\ q_2^j$. Let us call *local axis* the quadratic B-spline segment (if any) defined by m_1^j and m_2^j, the midpoints of s_1^j and s_2^j, and orientations $\underline{t}_1^j$ and $\underline{t}_2^j$ which are the normals to s_1^j and s_2^j (see figure 20.*b*). The "best" right ribbon segment mapping the points p_1 and p_2 is defined as the one which minimizes the following measure (see figure 20.*b*):

$$E = \lambda_1 E_1 + \lambda_2 E_2 \tag{4.1}$$

where $E_1 = \dfrac{\sum_i^n dist(p^i_a, p^i_m)/dist(p_i, q_i)}{n}$; p^i_a is an ith point on the local axis; p_i, q_i corresponding points (intersections of the line orthogonal to the axis at p^i_a with the two boundaries) and p^i_m the midpoint of $p_i\ q_i$ and $E_2 = \gamma / dist(m_1^j, m_2^j)$; γ being the angle between the segments s_1^j and s_2^j.

- for each subsequent unmatched point p_k of the reference boundary, use the previously found right ribbon correspondence segment $p_{k-1}q_{k-1}$ and search around $p_k q_k^0$ for the corresponding point q_k that minimizes E of equation 4.1.

fer and Kanade [28] have derived the conditions on the viewing direction and SOR parameters which cause self-occlusion to be observed in the image.

The purpose of the above analysis of the structural properties is to derive *verification* tests for object hypotheses made on the basis of the geometric properties. They are also used to infer useful shape information such as whether the cross-section points away or towards the viewer. The use of the projective properties (both geometric and structural) for the detection of relevant objects is discussed in the following section.

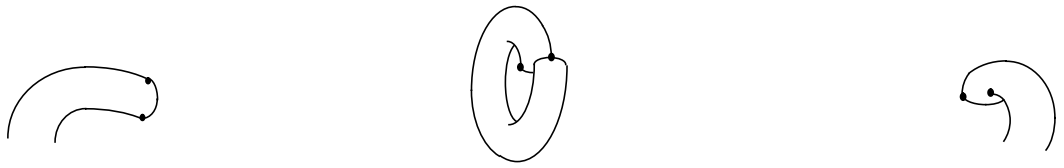
4 The segmentation method in real images

In this section we describe the method used for solving the segmentation and shape recovery of PRCGCs and circular PRGCs from a single real intensity image that includes imperfections such as noise, markings, shadows and occlusion. It consists of a hypothesize-verify process of three steps: the *detection* of local surface fragments (henceforth *local surface patches*), *grouping* of compatible ones and *verification* of object hypotheses. The constraints used in those steps are based on the projective properties discussed in section 3. The different steps are discussed below. For lack of space we will skip many details of the method and concentrate on its main steps.

4.1 Detection of local surface patches

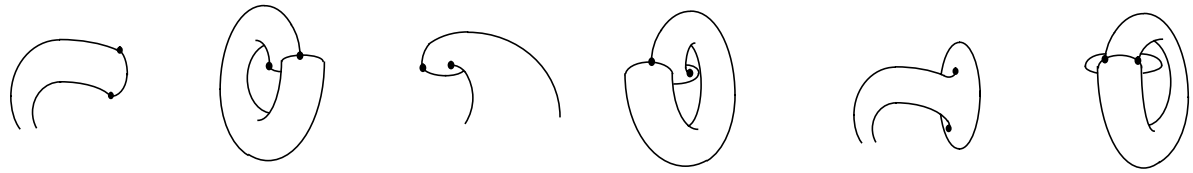
To handle occlusion and other sources of discontinuity, a system must be able to detect fragments of objects so that those that project from the same object will later be merged. From the properties of section 3.1, two types of geometric primitives relate the boundaries of a curved-axis part: parallel symmetries (invariant Property 1) and right ribbons (quasi-invariant Properties 5 and 6). Thus, the method to detect local surface patches must detect and use parallel symmetries and right ribbons. The detection of parallel symmetries used in our system uses the method described in [26,33] based on quadratic B-spline representation. We will thus omit its description.

To detect right ribbons we have used an improvement of the projection method of [19] (which was also later used by [24]) which searches for axis segments at several orientations. Based on the analysis at the end of section 3.1, our method uses the parallel symmetry axis (if any) between a pair of boundaries, as an initial estimate for the right ribbon axis in a minimization process. Given two boundaries B_1 and B_2, and the parallel symmetry axis between them, the method con-



a. *L-j L-j* b. *L-j T-j* c. *L-j cusp*

Figure 17 Closure patterns for a PRCGC with a cross-section pointing away from the camera

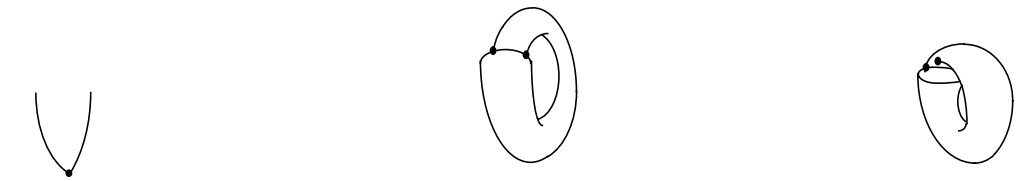


a. *L-j L-j* b. *L-j T-j* c. *L-j cusp* d. *T-j cusp* e. *cusp cusp* f. *T-j T-j*

Figure 18 Closure patterns for a circular PRGC with a cross-sections pointing away from the camera.

3.2.3 Non-visible cross-section

When the cross-section is not visible, then this is either due to a vanishing cross-section radius (a case we call *convergent closure*) or to self-occlusion causing *T-j*s to occur. In the latter case, the different combinations are shown in figure 19 and include (*T-j, T-j*) and (*T-j, cusp*).



a. *convergent* b. *T-j T-j* c. *T-j cusp*

Figure 19 Closure patterns for an invisible cross-section.

The analysis we have given is of a qualitative nature using results from previous work (the junction catalog, for example). Related analyses have been carried out in the past. Koenderink [12] has provided a mathematical analysis of the conditions under which certain *generic events* of general curved surfaces occur, and Ponce and Chelberg [21] have provided a mathematical derivation of limbs and cusps of GCs for the purposes of rendering and aspect graphs computation. Also, Sha-

Circular PRGCs

The possible combinations of circular PRGCs are more complex than PRCGCs (they a super-set) and are shown in figure 16. The only combination that does not occur is the (*cusp, cusp*) which, as mentioned above, corresponds to the case when the cross-section faces away from the camera. Notice that unlike for PRCGCs the combination (*T-j, cusp*) occurs here because the cross-section can flare (case f. of figure 16).

a. *3-tgt-j 3-tgt-j* b. *3-tgt-j T-j* c. *3-tgt-j cusp* d. *T-j T-j* f. *T-j cusp*

Figure 16 Closure patterns of a circular PRGC with a cross-section pointing towards the camera.

3.2.2 Visible cross-section pointing away from the camera

When the cross-section is visible and points away from the camera the possible junctions are pair wise combinations of the set (*L-j, T-j, cusp)*. Combinations involving *3-tgt-js* cannot be observed because *3-tgt-j*s indicate a cross-section pointing towards the camera. As before, we separate the cases depending on the primitive.

PRCGCs

The possible combinations for PRCGCs are shown in figure 17 and include (*L-j, L-j*), (*L-j, T-j*), (*L-j, cusp*). The combination (*T-j, cusp*) cannot occur here because it would imply that the cross-section is not visible (the cusp would have to be a curvature *cusp* and the *T-j* would have to be curvature *T-j*, both making the cross-section invisible). The combination (*T-j, T-j*) cannot occur also because the cross-section would not be visible and the combination (*cusp, cusp*) occurs only when at least one of the cusps is a flaring cusp, which cannot happen for PRCGCs.

Circular PRGCs

In the case of circular PRGCs, the different combinations are shown in figure 18 and include (*L-j, L-j*), (*L-j, T-j*), (*L-j, cusp*), (*T-j, cusp*), (*cusp, cusp*) and (*T-j, T-j*). Notice that the combination (*T-j, cusp*) can occur in this case unlike for PRCGCs, because we can have a *flaring cusp* here (figure 18.d). Also, the combination (*T-j, T-j*), when both *T-j*s are *curvature T-j*s while the cross-section is (partially) visible, can occur because of the possibility of flarings (case f).

below. Note that our purpose is *not* to give all the different aspects of the objects (in the aspect graph sense). Rather we want to illustrate the different *junction combinations* (which could be obtained for different aspects). The given lists are believed to be exhaustive because they consider all valid junction pairings; arguments are given when appropriate.

3.2.1 Visible cross-section pointing towards the camera

When the cross-section is visible and points towards the camera then the possible patterns are pair wise combinations of the set (*3-tgt-j, T-j, cusp)*. Combinations involving *L-j*s cannot occur in this case because *L-j*s correspond to a cross-section pointing away from the camera. The observed combinations depending on the primitive are discussed below.

PRCGCs

The possible combinations for PRCGCs are shown in figure 15 and include (*3-tgt-j, 3-tgt-j*), (*3-tgt-j, cusp*) and (*3-tgt-j, T-j*). *T-j*s are due to two reasons: when the cross-section points towards the camera and flares (*flaring T-j*) as in figure 16.d, and when the part self-occludes due to the curvature of the axis (*curvature T-j*) as in figure 15.c (left junction). The (*T-j, T-j*) combination cannot occur here because there is no flaring for PRCGCs (the scaling function is constant) and when both are curvature *T-j*s (part self-occludes) the cross-section would be completely invisible whereas we are considering a visible cross-section. Also, the (*T-j, cusp*) combination cannot occur because both junctions would be curvature junctions (curvature *T-j* and curvature *cusp*) implying a non-visible cross-section. The (*cusp, cusp*) combination holds only when the cross-section faces away from the camera.

a. *3-tgt-j 3-tgt-j* b. *3-tgt-j cusp* c. *3-tgt-j T-j*

Figure 15 Closure patterns of a PRCGC with a cross-section pointing towards the camera.

Notice that the above arguments hold because the axis is planar. For non-planar axes other combinations may occur.

8. the analysis can be easily extended to multiple surface objects.

to self-occlusion. This has the effect of producing *cusps* and *T-j*s in the image of the surface of a part. The existence of such self-occlusion junctions in the image depends on both the primitive and the viewing direction. Figure 13 gives examples. The inner-surface property can be stated as follows.

Property 9: In the image, the surface of an object can be delimited by continuous boundaries or by boundaries forming (self-occlusion) *T-j*s such that the cusps (ending parts of the tops of the Ts) lie inside the surface. Further, if more than one self-occlusion *T-j* is visible the cusps belong to the same portion of the surface.

The last part of the property is quite natural because the *cusping* boundaries belong to the occluding portion of a PRGC surface and a portion of such a regular surface is either occluding or occluded, not both (there is no twist in the sweep). This is best explained by the examples of figure 14 illustrating violations of the inner-surface property.

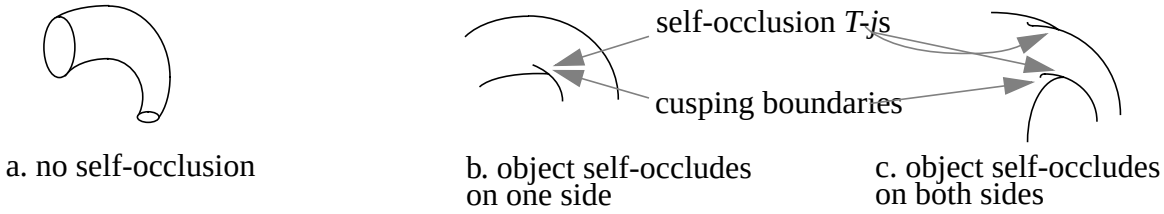


Figure 13 Inner-surface structural properties.

Figure 14 Impossible structural arrangements of self-occluding surface patches.

The closure properties will be given as junction combinations and connectivity relationships between them at the end of an object. To analyze the possible arrangements, it is useful to consider the cases when the cross-section is visible and pointing *towards* the camera, when it is visible and pointing *away* from the camera and when it is not visible. In the following analysis we assume that junctions are observed in pairs at each end of a part; i.e. that the surface is bounded by a pair of “side” boundaries.[8] In this case, the different pair wise junction combinations at the ends of the surface boundaries of an object and their connectivity, depending on the primitive, are discussed

3.2 Structural Properties

Among the properties of the image of an object which can be described as a PRGC[5] are also those characterizing the junction patterns (and their relationships) between its boundaries. Such junctions depend on both the part geometry and the viewpoint. Analyzing the possible junction patterns that can be observed in the image for different classes of parts and viewing directions is another way of using the 3-dimensionality of the scene objects.

Junctions of polyhedral and curved objects have been previously addressed (in [11] and [14] respectively, for example). An important conclusion of those efforts is that the relationships between the image boundaries of a 3-D object belong to a finite, small, catalog of junctions (*junction catalog*). For piecewise smooth objects, the junction catalog includes so-called *three-tangent* junctions (henceforth *3-tgt-j*) *L*-junctions (henceforth *L-j*), *T*-junctions (henceforth *T-j*) and *terminal* (henceforth *cusp*[6]) (see figure 12). For polyhedral objects or non-smooth cross-sections, other junctions such as *arrow* and *Y*-junctions occur [14]. For reasons of simplicity, in the analysis of this section we will use smooth (convex) cross-sections to illustrate the properties; i.e. the junctions we use in the analysis are *3-tgt-j*, *L-j*, *T-j* and *cusps*.[7] Other aspects of intensity images such as effects of occlusion by another object and of breaks in boundaries are not discussed here but in section 4 where we describe the segmentation and shape description methods. Also, the analysis assumes that parts are isolated and ignores junctions that occur due to joints between parts of a compound object (which are beyond the scope of this paper).

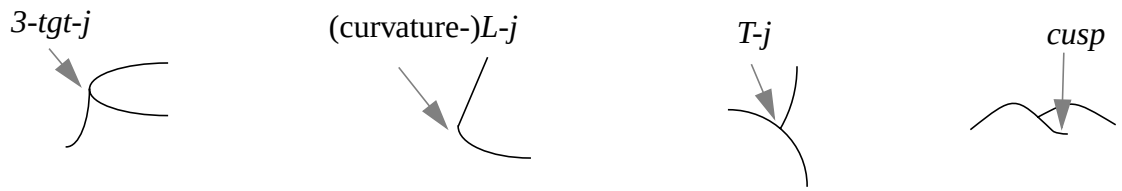


Figure 12 Junctions used for the structural properties.

It is useful to distinguish structural properties along the surface of a part (*inner-surface* properties) from those at its extremities (*closure* properties). Both types of properties give a catalog of possible junctions and connectivity relationships between them. Inner-surface junctions occur due

5. we will use interchangeably the terms “object”, “part” and PRGC to refer to objects addressed in this paper.
6. cusps are limb points where the viewing direction is tangential to the limb curve. See [12] for a detailed analysis of similar and other visual generic events.
7. extension to polyhedral junctions is straightforward and only requires augmenting the set of combinations we give.

It is clear that these two vectors are not only mutually parallel but parallel to the axis tangent as well. Furthermore, the right ribbon axis point (the mid-point of the segment connecting corresponding points) is also, by virtue of Definition 4, the parallel symmetry axis point ❑.

Figure 11.a illustrates this property. In the general case of a non-constant right ribbon, this property does not hold. This is the subject of the following stronger property.

Property 8: The only case where the right ribbon correspondences are exactly parallel symmetry correspondences is when the derivative of the right ribbon sweep function vanishes.

Proof:

We show that tangents to the two sides of such a right ribbon are only parallel when the sweep derivative vanishes. Using equation (3.19), the tangent to the C_+ curve is given by

$$\underline{C}'_+(t) = [\, 1 - m(t)\, k(t) \,]\, \underline{a}'(t) + m'(t)\, \underline{n}(t)$$

and using equation (3.20), the tangent to the C_- curve is given by

$$\underline{C}_-'(t) = [\, 1 + m(t)\, k(t) \,]\, \underline{a}'(t) - m'(t)\, \underline{n}(t).$$

These two vectors are parallel *if and only if* their determinant vanishes. It is easy to show that this latter is $-2\, m'(t)$. Therefore, the only condition under which these two vectors are parallel is $m'(t) = 0$; i.e. only at extrema of the ribbon scaling function or in the case of a constant sweep ❑.

There is in general an "offset" between the parallel symmetry correspondences and the correspondences of a right ribbon with non-constant sweep (figure 11.b). This offset is proportional to the derivative of the right ribbon scaling function.

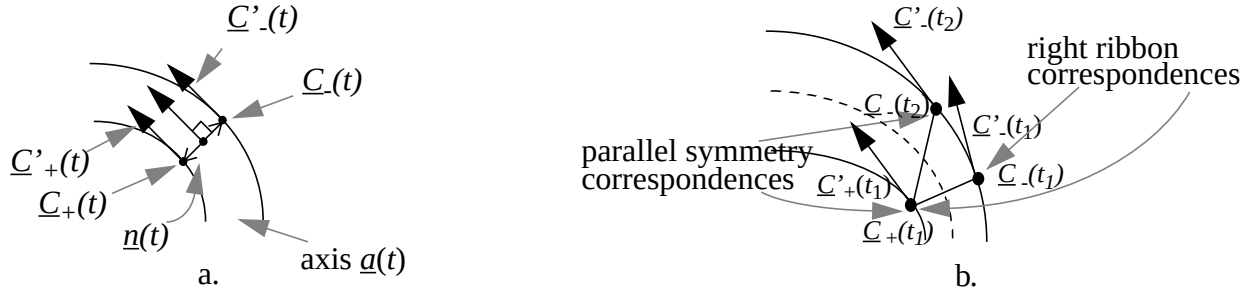


Figure 11 Right ribbon and parallel symmetry correspondences. Coincident in the constant sweep case (a), offset in the non-constant sweep case (b).

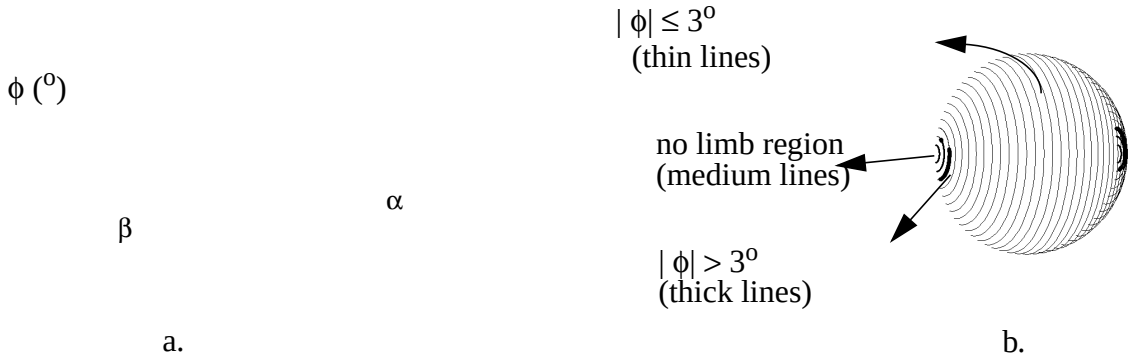


Figure 10 Property 6 for $(e, \dot{r}) = (0.2, 0.2)$ *. a. 3-D plot; b. half viewing sphere.*

So far, the analysis has indicated that two types of relationships characterize curved-axis primitives: parallel symmetry and right ribbons. The former give exactly the image of co-cross-sectional points of a PRCGC while the latter give "good" approximations of co-cross-sectional points of a circular PRGC. Because, both types of symmetry will be used to find curved-axis objects in the image, it is useful to analyze the relationship between parallel symmetry and right ribbons. This analysis will later be used to derive efficient methods for curved-axis object detection. The relationships are the subject of the following properties (see figure 11).

Property 7: A right ribbon with *constant* cross-section produces parallel symmetric side curves. Furthermore, the axis of parallel symmetry coincides with the right ribbon axis.

Proof:

Consider a right ribbon whose two sides are parameterized as follows:

$$\underline{C}_+ (t) = \underline{a} (t) + m (t)\, \underline{n} (t) \tag{3.19}$$

$$\underline{C}_- (t) = \underline{a} (t) - m (t)\, \underline{n} (t) \tag{3.20}$$

where $\underline{a} (t)$ is a 2-D curve (axis) parameterized by arclength t, $\underline{n} (t)$ the normal vector to the axis and $m (t)$ the 2-D scaling function ($m (t) = m$ is constant in this case). Corresponding points of such a right ribbon are $\underline{C}_+ (t)$ and $\underline{C}_- (t)$ for the same value of t. The tangent at the $\underline{C}_+ (t)$ side is given by

$$\underline{C}'_+ (t) = [\, 1 - m\, k (t)\,]\, \underline{a}' (t)$$

since $m'(t) = 0$; $\underline{a}' (t)$ is the unit tangent vector of the axis $\underline{a} (t)$ and $k (t)$ its curvature. Similarly, the tangent at the $\underline{C}_- (t)$ side is given by

$$\underline{C}_-' (t) = [\, 1 + m\, k (t)\,]\, \underline{a}' (t)$$

iv) Figure 8.b shows the graph of the size of the previous space of observation where ϕ is within different (upper-bound) angular values of 0^o.

v) Table 3.2 gives the sizes of the regions on the viewing sphere where ϕ is 3^o or less for some values of $(e, \dot{r})$. The size is with respect to the space region where limbs exist.

vi) Figure 10.a gives the 3-D plot of ϕ as a function of (α, β) for $(e, \dot{r}) = (0.2, 0.2)$ and figure 10.b the corresponding half viewing sphere, such as discussed for figure 9.b.

	$e = 0.1$	$e = 0.2$	$e = 0.3$	$e = 0.4$
$\dot{r} = 0.1$	96.80	95.04	92.26	91.76
$\dot{r} = 0.2$	99.21	95.45	93.01	90.78
$\dot{r} = 0.3$	100	97.31	94.27	92.82
$\dot{r} = 0.4$	100	99.00	96.30	94.07

Table 3.2 Viewing sets sizes (percent) where the image angle ϕ is 3^o or less.

The behavior of this quasi-invariant property is similar to the previous one. The graph of figure 8.*b* shows that the two tangents are within 3^o of each other over 94.48% of the previous space of observation and within 5^o over 96.90% of that space. Table 3.2 shows that the degradation is gradual and slow. Note that because the size of the regions on the viewing sphere where limbs do not exist is mainly influenced by $\dot{r}$, the relative size of the region where the property holds may tend to increase as $\dot{r}$ increases (the ratio is over a smaller region, where limbs exist at all). Also, from figure 10, it can be seen that the angle difference tends to increase only close to regions where limbs do not exist.

Note that it can be verified that the properties hold exactly (perfect orthogonality and coincidence of axes) for general circular PRGCs for the two viewpoints, where the viewing direction $\vec{V}$ is orthogonal to the axis plane (i.e. side view, $\cos\alpha = \cos\beta = 0$, $\theta_1 = 0$, $\theta_2 = \pi$) and where it is in the axis plane (i.e. frontal view, $\sin\alpha = 0$, $\cos\theta_1 = \cos\theta_2$, $\sin\theta_1 = -\sin\theta_2$).

These two properties are quasi-invariant with respect to transforms parameterized over the 4-D space of observation. Thus they are orthographic quasi-invariants (with respect to (α, β)) *and* object-shape parameters quasi-invariants (with respect to $(e, \dot{r})$). They indicate that the image of a general circular PRGC can be well approximated by a right ribbon in the sense that the *right ribbon correspondences are close to co-cross-sectional points correspondences over most of the parameter space*.

to the 3-D axis tangent; i.e. an unlikely viewpoint. Therefore, even if $\vec{V}$ is close to being in the axis plane, the property would degrade only at points where the axis tangent is in the direction of $\vec{V}$ (a small set of points) and it would still hold for the rest of the surface (a much larger set).

	$e = 0.1$	$e = 0.2$	$e = 0.3$	$e = 0.4$
$\dot{r} = 0.1$	96.72	92.44	88.50	85.57
$\dot{r} = 0.2$	96.86	89.44	83.62	78.55
$\dot{r} = 0.3$	97.11	88.25	80.05	74.83
$\dot{r} = 0.4$	97.29	87.57	78.72	72.86

Table 3.1 Viewing sets sizes (percent) where the observed image angle γ is within 5^o of 90^o

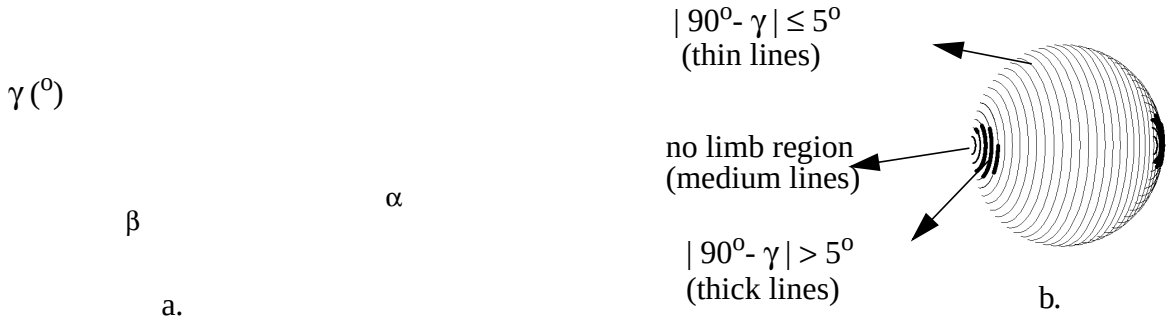


Figure 9 Property 5 for $(e, \dot{r}) = (0.2, 0.2)$. a. 3-D plot of γ as a function of the viewing angles; b.corresponding half viewing sphere.

This property can be thought of as the projective counterpart of the 3-D orthogonality between the segment connecting limb points and the tangent to the 3-D axis (in 3-D, the cross-section is planar and orthogonal to the axis).

Property 6: In an orthographic projection of a circular PRGC, tangents to the 2-D axis and the projection of the tangents to the 3-D axis at the corresponding points are "almost parallel over most of the parameter space of observation." Furthermore, the (angular) measure of parallelism (ϕ) tends to degrade only close to degenerate regions of the parameter space for which limbs do not exist and which include non-general viewing directions or non-common shapes such as those close to self-intersect themselves.

To show this quasi-invariant property, we have used the method discussed previously for Property 5. The results of the analysis of the image angle ϕ between these two tangent vectors are summarized in analogous ways.

- using equation (3.15) determine the projection of the two pairs of points $P(s, \theta_{11})$, $P(s, \theta_{12})$ and $P(s + ds, \theta_{21})$, $P(s + ds, \theta_{22})$ (say p_{11}, p_{12}, p_{21} and p_{22})
- determine the angle between the cross-section segment given by p_{11} and p_{12} and the 2-D axis tangent given by the line joining the midpoints of $p_{11}\,p_{12}$ and $p_{21}\,p_{22}$.

We derive the angles in the image between cross-section segments and 2-D axis tangents over the space of observation defined by $\alpha \in [0, \pi]$, $\beta \in (0, \pi)$, $e < 0.5$ and $|\dot{r}| < 0.5$. To avoid long tables of numbers, we attempt to summarize the results in various ways.

i) Figure 8.a shows the graph of the size (in percent) of the space of observation for which the image angle γ is within different (upper-bound) angle values of 90^o.

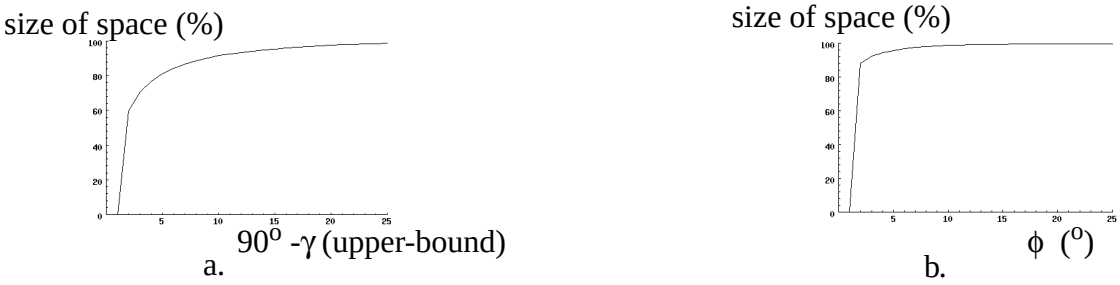


Figure 8 Plots of the size of space of observation for different upper-bounds of the image angles γ (a) and ϕ (b).

ii) Table 3.1 summarizes the parts (in percent) of the regions on the viewing sphere where γ is within $5^o$of 90^o for certain values of $(e, \dot{r})$. The size is with respect to the space region where limbs exist.

iii) Figure 9.a shows a 3-D plot of γ as a function of (α, β) for $(e, \dot{r}) = (0.2, 0.2)$ and figure 9.b shows the corresponding display of the half viewing sphere ((α, β) sub-space). This figure shows where the property holds (thin lines), where it does not hold (thick lines) and where limbs do not exist (medium lines). In the display, vertical circles correspond to constant β values with a 5^o step.

The graph of figure 8.a shows that over 84.30% of that 4-D space (excluding special values $e = 0$ or $\dot{r} = 0$; i.e. SORs and circular PRCGCs) the 2-D angle is within 5^o of 90^o and for over 92.63% of the space is within 10^o of 90^o. It can be seen, from table 3.1, that the size of the space where the property holds (2-D angle within 5^o of being right) gradually decreases as both e and $\dot{r}$ take higher values. Also, notice from figure 9 that the space where the property holds is connected and the property is well behaved. It tends to gradually degrade for small values of $|\tan\beta|$, that is close to regions where limbs do not exist. Notice also that a small value of $|\tan\beta|$ requires a very specific viewing direction $\vec{V}$ that is not only close to being in the axis plane but also almost parallel

animal limbs or industrial parts, appear not to have, at the same time, high values of e and $\dot{r}$; i.e.when e is high $|\dot{r}|$ is small (otherwise the thickness ratio would rapidly increase, which would cause self-intersection) and when $|\dot{r}|$ is high e is small (such as is the case for some musical wind instruments, a saxophone for example, where the outlet, which has a high value of $|\dot{r}|$, has locally a linear axis, $e = 0$).

We address objects whose limbs are visible over a wide range of views. For this reason and for the purpose of avoiding large tables of results, we consider e < 0.5 and $\dot{r} < 0.5$; i.e. for a unit radius cross-section, sweep rates less than half the current radius per unit arclength along the axis, an already high-slope sweep. Curved-axis objects with a higher thickness ratio are rare in our environment, so do those with a higher sweep derivative.

2) Quasi-Invariant Properties

Property 5: In the projection of a circular PRGC, the cross-section segments and the 2-D axis tangents at their mid-points are "almost orthogonal over most of the parameter space." Furthermore, the (angular) measure of orthogonality (γ) tends to degrade only close to degenerate regions of the parameter space for which limbs do not exist and which include non-general viewing directions or non-common shapes such as those close to self-intersect themselves.

To illustrate this quasi-invariant property, we analyze the behavior of the angle γ between a cross-section segment and the tangent to the 2-D axis as a function of the parameters in our space of observation. Algebraic analysis of this quasi-invariant property is difficult since it requires knowledge of how θ varies with respect to s at limb points, which is not known in general. Instead, we analyze it numerically by discretizing the space of observation and, for each point $(\alpha, \beta, e, \dot{r})$ of the space, solving the limb equation and deriving the projections of co-cross-sectional limb points and the tangent to the 2-D axis. Note that the analysis is not made on particular object instances but uses the whole parameter space of observation described above for the class of circular PRGCs. The method is as follows (we omit the details):

for each set of parameters $(\alpha_1, \beta_1, e_1, \dot{r}_1)$ (at some s) do

- select an arbitrary 3-D frame $F_1 = (\hat{t}_1, \hat{n}_1, \hat{b}_1)$
- solve the limb equation (3.18) to obtain the two limb points θ_{11} and θ_{12}
- determine $(\alpha_2, \beta_2, e_2, \dot{r}_2)$ at $s + ds$ (for some small ds)
- solve equation (3.18) for the second pair of limb points θ_{21} and θ_{22} (at $s + ds$) and express their coordinates in F_1

Below, we show that there are *quasi-invariant* properties which are as useful as the previous invariant properties. For this we have to

a) identify the relevant parameter space of observation, giving the parameters and their ranges of useful values

b) show that the properties of interest constitute quasi-invariant mappings in that their values remain within a "small" set over "most" of the parameter space.

1) Space of Observation

Let us rewrite the limb equation (3.14) in the following form

$$(1- e\cos\theta)\cos(\theta - \alpha) = \dot{r}\cot\beta \qquad (3.18)$$

where $e = \kappa r = r / R$ (R being the radius of curvature of the axis); e is a local measure of the relative *thickness* of the shape of the circular PRGC. Smaller values indicate rather elongated surfaces (small cross-section radius compared to axis curvature radius) whereas larger values thick and highly bent surfaces. We will call e the *thickness ratio*; $\dot{r}$ is a measure of how "fast" the cross-section changes its size (radius).

Equation (3.18) indicates that two pairs of parameters affect the behavior of the contours of a general circular PRGC:

- (α, β) corresponding to the viewing direction
- $(e, \dot{r})$ corresponding to the object parameters (local shape measures)

The whole space of observation is a 4-D space $(\alpha, \beta, e, \dot{r})$. For the same object, those parameters vary as s varies; α and β can take any values on the viewing sphere for which the limb equation (3.18) admits a finite (non empty) set of solutions; i.e. $\alpha \in [0, 2\pi)$ and $\beta \in (0, \pi) \cup (\pi, 2\pi)$. Letting $\alpha \in [0, \pi]$ and $\beta \in (0, \pi)$ is sufficient as the limb equation (3.18) is symmetric with respect to π for α and β.

The 2-D subspace $(e, \dot{r})$ is also constrained. We have $|e| \leq 1$ (the cross-section radius r is smaller than the radius of curvature, R, of the axis), otherwise the surface self-intersects. $\dot{r}$ is also constrained, since $|1- e\cos\theta| \leq 2$ (as $|e| \leq 1$) and thus, from equation (3.18), $|\dot{r}| \leq 2|\tan\beta|$. This implies that, at each point, the closer the viewing direction is to the axis tangent direction (i.e. smaller $|\tan\beta|$), the smaller $|\dot{r}|$ has to be (otherwise there would be no visible limbs). Thus, for an axis point where $\beta = 15^o$, for example, the cross-section has limbs if $|\dot{r}| < 0.53$. This implies that, for a given circular PRGC, in order for the surface to have limb points (and its projection to have visible points) over a wide range of views, $|\dot{r}|$ has to be small. Objects seen in daily environments, such as

where $\kappa(s)$ is the curvature of the axis $A(s)$ and $\dot{r}(s)$ the derivative of $r(s)$.

Assuming $\sin(\beta(s)) \neq 0$,[4] and dropping the arguments θ and s, this equation can be written as follows

$$[1 - \kappa r \cos\theta] \cos(\theta - \alpha) = \dot{r} \cot\beta \tag{3.14}$$

Using the projection equation (3.7), a circular PRGC point $P(s, \theta)$ can be shown to project as

$$p = r \begin{pmatrix} \cos\beta \cos(\theta - \alpha) \\ \sin(\theta - \alpha) \end{pmatrix} \tag{3.15}$$

in the image frame $(\hat{u}, \hat{v})$. The cross-section segment between projections of co-cross-sectional limb points $P(s, \theta_1)$ and $P(s, \theta_2)$, can be expressed in its local 2-D frame as

$$\vec{C}_s = r \begin{pmatrix} \cos\beta(\cos(\theta_2 - \alpha) - \cos(\theta_1 - \alpha)) \\ \sin(\theta_2 - \alpha) - \sin(\theta_1 - \alpha) \end{pmatrix} \tag{3.16}$$

and its midpoint (2-D axis point) is given by

$$\vec{M}_{2D} = \frac{r}{2} \begin{pmatrix} \cos\beta(\cos(\theta_2 - \alpha) + \cos(\theta_1 - \alpha)) \\ \sin(\theta_2 - \alpha) + \sin(\theta_1 - \alpha) \end{pmatrix} \tag{3.17}$$

Properties of circular PRGCs

The properties we give below also give relationships between projections of co-cross-sectional limb points and the 2-D axis and between the 2-D axis and the projection of the 3-D axis. We will not consider the special cases of circular PRGCs which include circular PRCGCs (constant cross-section; $\dot{r} = 0$) as they have been discussed in section 3.1.1, or surfaces of revolution (SORs; $\kappa = 0$) which are a sub-set of SHGCs whose properties have been studied in the past [23,27,29,33]. Rather, we consider the general case where $\dot{r} \neq 0$ *and* $\kappa \neq 0$. Properties of general circular PRGCs have not been previously addressed in the literature.

So far, all the given properties were *invariant* with respect to the viewing direction (angles). It can easily be shown that they are *not* invariant for general circular PRGCs. For example, the orthogonality of the 2-D axis tangents and the cross-section segments is not a property of general circular PRGCs. Equations to demonstrate this are given in [34] and are omitted here for lack of space.

4. which excludes $\vec{V}$ being parallel to $\vec{t}$, a non general viewing direction for which the limb equation has an infinite or zero number of solutions.

cross-section segment is *not* constant. In [38] an analysis is given to demonstrate this. It is omitted here for lack of space.

3.1.2 Geometric properties of general circular PRGCs

Definition 5: A circular PRGC (*circular planar right generalized cylinder*) is obtained by sweeping a circle *C* along a planar axis curve *A* while scaling it by a smooth function *r* and maintaining it orthogonal to *A* (restricted to pass through the center of *C*).

Letting $C(\theta) = (\rho \cos\theta, \rho \sin\theta)$ be a parametrization of *C* (ρ is the constant radius, fixed without loss of generality to 1), $A(s)$ an arclength parameterization of *A* (assumed to be C^1 and at least piecewise C^3 for the same reason as for PRCGCs) and $r(s)$ a C^1 function, the surface of a circular PRGC can be parameterized as follows:

$$P(s, \theta) = A(s) + r(s) \cos\theta \; \hat{n}(s) + r(s) \sin\theta \; \hat{b}(s) \tag{3.12}$$

where as before $\hat{n}(s)$ and $\hat{b}(s)$ are respectively the normal and binormal to the axis curve *A*(s). Figure 7 illustrates this representation. The *cross-sections* are the curves of constant *s* and the *meridians* are the curves of constant θ.

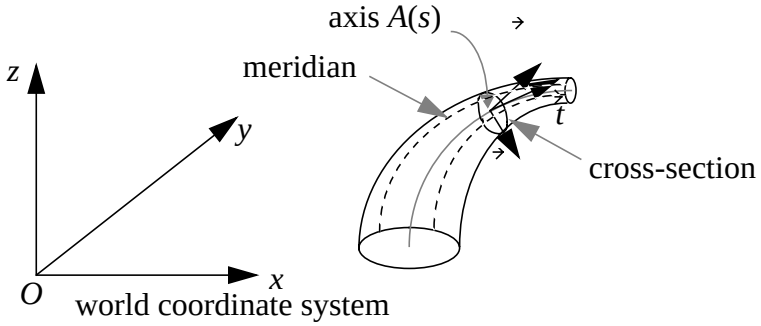


Figure 7 Sample circular PRGC and related representation.

Deriving limbs and their projections

Using equation (3.12) and the expression of the surface normal given in equation (3.2) and of the viewing direction given in equation (3.3), the limb equation of a circular PRGC can be shown to be given by

$$\sin\beta(s) [1 - \kappa(s)\rho \, r(s) \cos\theta] \cos(\theta - \alpha(s)) - \rho \; \dot{r}(s) \; \cos\beta(s) = 0 \tag{3.13}$$

Property 3: In the projection of a circular PRCGC, the angle γ between cross-section segments and 2-D axis tangent is $\pi / 2$ regardless of the viewing direction.

Proof:

Reporting the results of the previous proof in equation (3.9), we obtain the expression of the cross-section segment

$$r\,(0, \pm 2)^t \tag{3.11}$$

The 2-D axis shown to be the projection of the 3-D axis, its tangent vector is the projection of the 3-D tangent vector $\grave{t}$. This latter, from the projection equation (3.7), is given by

$$(-\sin \beta, 0)^t$$

which is orthogonal the previous cross-section segment ❑.

Property 4: In the projection of a circular PRCGC, the length of the cross-section segments is constant (it is equal to the diameter of the 3-D cross-section).

Proof:

Using equation (3.11), the length of the cross-section is 2 *r* where *r* is the constant 3-D cross-section radius ❑.

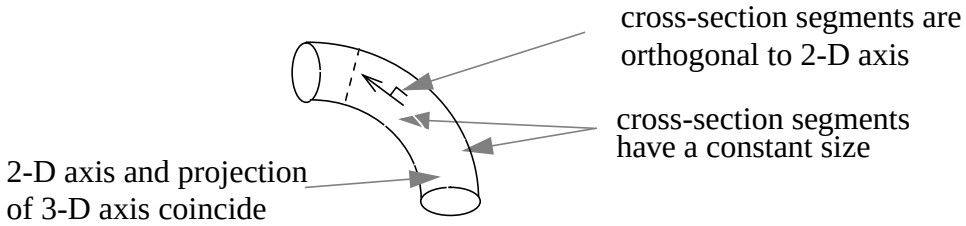


Figure 6 Invariant properties of circular PRCGCs

Properties 2, 3 and 4 say that in the image plane, the projection of a circular PRCGC can be described by a *right ribbon* which has a constant sweep function such that the ribbon corresponding points are the projections of co-cross-sectional points and the ribbon axis the projection of the 3-D axis.

For a PRCGC with a *non-circular* cross-section, it can be shown that the cross-section segment and the 2-D axis are *not* mutually orthogonal for arbitrary cross-section shapes, that the 2-D axis and the 3-D axis do *not* coincide in general (although they are parallel symmetric) and the

$$p = \rho\begin{pmatrix}\cos\beta\cos(\theta-\alpha)\\ \sin(\theta-\alpha)\end{pmatrix} \tag{3.8}$$

From the terminology given in Definition 3, a cross-section segment of a PRCGC, between the projections of co-cross-sectional limb points $P(s, \theta_1)$ and $P(s, \theta_2)$, can be expressed in its local 2-D frame as

$$\vec{C}_s = \begin{pmatrix}\cos\beta(\rho_2\cos(\theta_2-\alpha)-\rho_1\cos(\theta_1-\alpha))\\ \rho_2\sin(\theta_2-\alpha)-\rho_1\sin(\theta_1-\alpha)\end{pmatrix} \tag{3.9}$$

where $\rho_i = \rho(\theta_i)$; and its midpoint (2-D axis point) is given by

$$\vec{M}_{2D} = \frac{1}{2}\begin{pmatrix}\cos\beta(\rho_2\cos(\theta_2-\alpha)+\rho_1\cos(\theta_1-\alpha))\\ \rho_2\sin(\theta_2-\alpha)+\rho_1\sin(\theta_1-\alpha)\end{pmatrix} \tag{3.10}$$

Properties of PRCGCs

The following property is due to [31]. It relates the image of co-cross-sectional points of a PRCGC.

Property 1: The limb (or meridian) projections of a PRCGC are parallel symmetric, such that symmetric points are co-cross-sectional, regardless of the viewing direction. Furthermore, the image of the 3-D axis is also parallel symmetric to the limb (or meridian) projections such that symmetric points are on the same cross-section.

In the special case of a circular PRCGC (circular cross-section), we can derive additional invariant properties. They establish further relationships between the 2-D axis and the projection of the 3-D axis and between cross-section segments and the 2-D axis of a circular PRCGC. Three new invariant properties are given below with their proofs. Figure 6 illustrates the new properties.

Property 2: In an orthographic projection of a circular PRCGC, the 2-D axis and the projection of the 3-D axis coincide regardless of the viewing direction.

Proof:

From equation (3.5), at limb points, we have

$\cos(\theta_1 - \alpha) = \cos(\theta_2 - \alpha) = 0$, and thus, $\sin(\theta_1 - \alpha) = -\sin(\theta_2 - \alpha) = \pm 1$

for each s along the axis. Therefore, from equation (3.10), the 2-D axis point is given by $(0, 0)^t$, the origin of the local 2-D frame which, as previously discussed, is the projection of the axis point $A(s)$ no matter what the viewing direction (given locally by α and β) is ❑.

Let $(\alpha(s), \beta(s))$ be the angular coordinates of the unit viewing vector $\vec{V}$ in the Frenet-Serret frame $(\grave{t}(s), \grave{n}(s), \grave{b}(s))$, then

$$\vec{V} = \cos\beta(s)\ \grave{t}(s) + \sin\beta(s)\cos\alpha(s)\ \grave{n}(s) + \sin\beta(s)\sin\alpha(s)\ \grave{b}(s) \tag{3.3}$$

Using the Frenet-Serret theorem [15] relating the above basis vectors and their derivatives and writing the orthogonality of $\vec{N}(s, \theta)$ and $\vec{V}$ yields the following limb equation:

$$\cot(\theta - \alpha) = -\dot{\rho} / \rho \tag{3.4}$$

where $\dot{\rho}(\theta)$ is the derivative of $\rho(\theta)$ and $\alpha = \alpha(s)$ (we will henceforth omit the arguments). In the special case of a circular cross-section (i.e. a circular PRCGC, where $\dot{\rho} = 0$) this equation becomes

$$\cos(\theta - \alpha) = 0 \tag{3.5}$$

yielding two limb points determined by $\theta = \pm\pi / 2 + \alpha(s)$; *i.e.* two diametrically opposite points (figure 5).

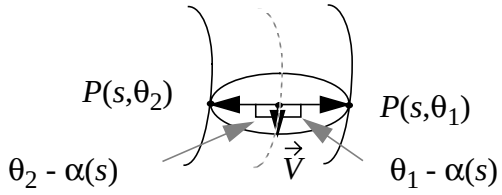


Figure 5 Limb points properties for a circular PRCGC.

The orthographic projection of the (3-D) Frenet-Serret frame $(\grave{t}, \grave{n}, \grave{b})$ on a plane orthogonal to$\vec{V}$ gives a "moving" local 2-D frame $(\grave{u}, \grave{v})$ in the image for each value of *s*. The relationship between the 3-D and the 2-D frames is as follows:

Let $P = t\grave{t} + n\grave{n} + b\grave{b}$ be a point expressed in the 3-D frame, then its projection, *p*, on a plane orthogonal to $\vec{V}$ is given by

$$p = (-\sin\beta\ t + \cos\beta\cos\alpha\ n + \cos\beta\sin\alpha\ b)\ \grave{u} + (-\sin\alpha\ n + \cos\alpha\ b)\ \grave{v} \tag{3.6}$$

where $\grave{u} = \grave{u}(s)$ is the projection of $\grave{t}$ and $\grave{v} = \grave{v}(s)$ is orthogonal to $\grave{u}$ following the right hand rule. Written in vector form, local 3-D coordinates $(t, n, b)^t$ project as local 2-D coordinates

$$p = (-\sin\beta\ t + \cos\beta\ \cos\alpha\ n + \cos\beta\ \sin\alpha\ b, -\sin\alpha\ n + \cos\alpha\ b)^t \tag{3.7}$$

The projection of the axis point $A(s)$ is thus the origin of the local 2-D frame. Using equations (3.1) and (3.7), a PRCGC point $P(s, \theta)$ can be shown to project, in the image frame $(\grave{u}, \grave{v})$, as:

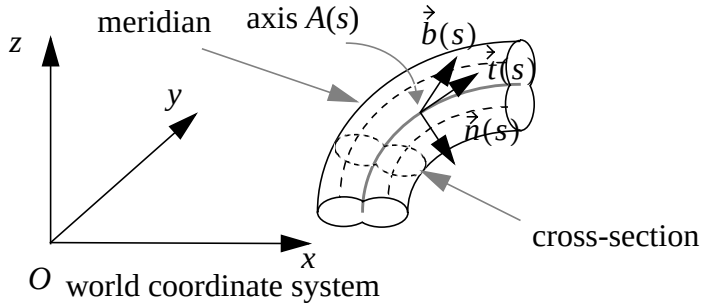


Figure 3 Sample PRCGC and related representation

Definition 3: The image line segment joining the projection of co-cross-sectional points is called a *cross-section segment*. The mid-point of a cross-section segment is called a *2-D axis point*. The locus of 2-D axis points is called the *2-D axis*[3] (see figure 4).

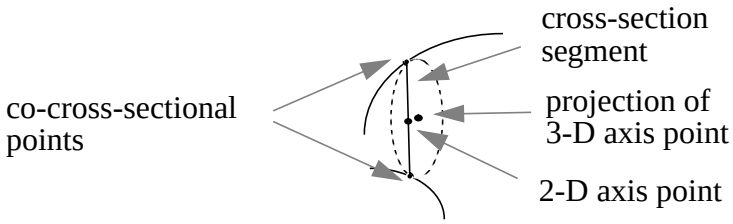


Figure 4 Terminology of (the projection of) a GC.

Definition 4: Two curves are said to be *parallel symmetric* [30] if there exists a continuous and monotonic, point wise, correspondence function such that corresponding points have parallel tangent vectors. The axis of parallel symmetry is the locus of mid-points of segments connecting parallel symmetric points (lines of parallel symmetry).

Deriving limbs and their projections

A detailed description of the limb derivation process for GCs is given in [21]. Here, we only give some basic steps to define the limb boundaries and express their equation. A point on the GC surface is a limb point *iff* the viewing direction $\vec{V}$ is orthogonal to the surface normal $\vec{N}(s, \theta)$ at that point.The surface normal at a point $P(s, \theta)$ is given by

$$\vec{N}(s) = \frac{\partial P}{\partial s} \times \frac{\partial P}{\partial \theta} \Big/ \left\| \frac{\partial P}{\partial s} \times \frac{\partial P}{\partial \theta} \right\| \quad \text{, where } P = P(s, \theta) \tag{3.2}$$

3. note that the 2-D axis is *not* necessarily the projection of the 3-D axis.

3.1 Geometric properties

Geometric projective properties are those which characterize the intrinsic description of a GC, namely its cross-section, its axis and its scaling function. They consist of (symmetry) relationships between the boundaries of an object and fall in two categories: *invariant* properties and *quasi-invariant* properties. The viewing geometry is approximated by orthographic projection in this work and thus, the properties are invariant or quasi-invariant with respect to orthographic transforms (parameterized by the angles on the viewing sphere). To derive such properties, we give a useful mathematical formulation of the limbs of PRGCs and their projections. A number of definitions are also given to clarify the terminology. Some of the properties we discuss have been derived by others in previous work and will be given without proofs. Those we derive are given with their proofs. We first discuss PRCGCs then circular PRGCs.

3.1.1 Geometric properties of PRCGCs

Definition1: A PRCGC (*planar right constant generalized cylinder*) is obtained by sweeping a planar cross-section curve C along a planar smooth axis curve A while maintaining C constant and orthogonal to A.

Letting $C(\theta) = (\rho(\theta)\cos\theta, \rho(\theta)\sin\theta)$ be a parameterization[2] of C, $A(s)$ an arclength parameterization of A, the surface of a PRCGC can be parameterized as follows (with respect to some world coordinate system):

$$P(s, \theta) = A(s) + \rho(\theta)\cos\theta\ \hat{n}(s)\ + \rho(\theta)\sin\theta\ \hat{b}(s) \tag{3.1}$$

where $A(s)$ is an arclength parameterization of the axis curve A, $\hat{n}$ and $\hat{b}$ respectively the normal and the (constant) binormal to the axis curve. Figure 3 illustrates this parameterization. The function $A(s)$ is assumed to be C^1, and at least piecewise C^3, so that edges on the surface are only due to vertices of the cross-section function $\rho(\theta)$ which is assumed to be *piecewise* C^1. The *cross-sections* are the curves of constant s and the *meridians* are the curves of constant θ.

Definition 2: Two points of a GC are called *co-cross-sectional* if they belong to the same cross-section (i.e. they can be expressed as $P(s, \theta_1)$ and $P(s, \theta_2)$ for some s where $A(s)$ is defined; see figure 4).

2. the cross-section curve C is assumed to be *star-shaped* to admit such a parameterization, chosen only to simplify the mathematical analysis. Extension to non-star shapes only requires a different parameterization.

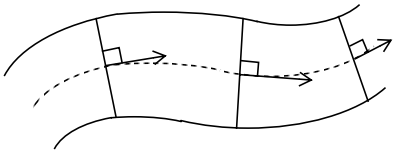


Figure 2 Sample right ribbon.

such ribbons in an image. The ribbons so detected correspond to different components of a complex object.

In [6], Brooks used similar ribbons in a model based approach to detection of airplane instances in an image. Right ribbons with straight axis, together with other shapes such as ellipses constitute the basic generic shapes for matching model-predicted and image-observed features.

In [24], Rao and Nevatia also used right ribbons in a system for segmentation and description of complex objects from imperfect contours. Ribbons were used to detect surface fragments and form aggregates of those fragments that correspond to object parts descriptions. Detection of right ribbons is based on the projection method of [19].

In the above methods for parts detection and description, ribbons were used as intuitive descriptors which were justified on a heuristic and empirical basis. Although they proved useful for detection of surfaces and parts, the descriptions they provide were inherently 2-dimensional. Recovery of information about 3-D shape, from ribbons, was not addressed, neither was the relationship between the descriptions that can be obtained for different viewpoints or occlusion patterns.

For the purposes of 3-D analysis, it is important to derive 3-D descriptions whenever possible or otherwise projective descriptions which rigorously relate to some 3-D shape scheme. In doing so, the recovered shape descriptions become less dependent (ideally independent) of the viewpoint, an important ability in 3-D vision. Consequently, the descriptions are very useful for 3-D object recognition. In the remaining sections of this paper, we discuss means (projective properties and methods) to achieve such descriptions for PRGCs.

3 Projective properties of PRCGCs and circular PRGCs

The projective properties of the shapes we address are important because from them we can derive features and their relationships that can be used for object detection and for 3-D reconstruction. We first discuss the geometric properties of the primitives and then their structural ones.

2.1 Recovery of curved-axis generalized cylinders

Most past work has addressed straight axis primitives such as surfaces of revolution (SORs) [18] and SHGCs [8,23,27,28,29,33]. Curved axis primitives have been addressed by very few researchers. In [21], derivation of limbs and cusps of GCs is addressed. However, the derivation although useful for computation of aspect graphs and rendering, does not address the recovery of GCs.

Ulupinar and Nevatia [31] have addressed properties and recovery of PRCGCs. The main property is that orthographic projections of the limbs of PRCGCs are *parallel symmetric* (generalization of straight line parallelism to curved ones; the notion will be formally defined in section 3.1.1), and that symmetric points belong to the same cross-section, regardless of the viewing direction. This property has been used by the authors to rule the surface of a PRCGC and recover surface normals along of the visible parts of the surface. Their 3-D recovery method uses a differential geometric analysis of the relationship between surface normals and image symmetries and structure.

More recently, Pillow et al. [20] have studied certain perspective invariants of *canal surfaces*, which are PRCGCs with circular cross-sections, under the constraint that their axes have inflections. The properties are used for recognition by using invariant representations of the axial shapes.

To our knowledge, there has been no previous work on derivation of projective properties and recovery of primitives with *both* a curved axis *and* a non-constant size cross-section such as circular PRGCs, even from perfect boundaries. Also, there is no previous work on segmentation and 3-D recovery of curved-axis GCs from a single real intensity image.

2.2 Ribbon detection methods

Ribbons have traditionally been used as 2-dimensional counterparts of GCs. They provide an axial representation scheme that captures useful (image) shape properties such as elongation, size variation and curvature and thus provide descriptions that can be used for recognition. There have been a number of different types of ribbons (see [22] for a classification) and the ones which were mostly used are those characterized by a (2-D) segment orthogonal to the axis curve (which is not restricted to be straight). We will henceforth call them *right ribbons*. Figure 2 shows an example.

Early work on ribbon finding includes Nevatia and Binford's [19] method for description and recognition of complex curved objects. Given a range image and its extracted contours, their method was to detect ribbons then match the resulting (graphical) representations to stored models. The ribbons used were *right ribbons*. They introduced a method, called the *projection method*, to detect

a. b.

Figure 1 Sample PRCGC (a) and circular PRGC (b).

for the automatic segmentation and descriptions of complex curved objects which are made up of both straight and curved-axis parts. The significance of such a part-based approach has been acknowledged since the early days of the field. It is even a psychologically plausible approach for human visual perception as demonstrated by Biederman's work [2]. We expect, within the next few years, to develop systems for part-based description and recognition of 3-D compound objects from a 2-D image (a system for the segmentation and descriptions of such objects is currently under development by the authors).

We organize the discussion in this paper as follows. In section 2, we discuss related previous work. In section 3 we discuss the projective properties of the contours of PRCGCs and circular PRGCs. In this section, we introduce new orthographic invariant and quasi-invariant properties of these classes of shapes. In section 4, we describe an implemented system for their segmentation and description in a real intensity image. In section 5 we describe a method for the recovery of complete 3-D volumetric descriptions of PRCGCs and circular PRGCs. Results of the methods on real images are given throughout the discussion. In section 6, we discuss issues related to the evaluation of the resulting descriptions and the performance of the implemented system. We conclude this paper in section 7.

2 Relation to previous work

Related previous work spans efforts on the derivation and use of invariant properties of generic shapes, 3-D shape recovery from 2-D contours and methods of segmentation of ribbons and generalized cylinders. These areas are not independent and we will concentrate the discussion on the most closely related work on recovery of curved axis generalized cylinders and on ribbon detection methods.

refer to functions which are not strictly invariant to their parameterizations such as the viewing angles. Rather, their values are allowed to vary but do so slowly and remain restricted to a relatively small range of values over most of the parameter space. (In [1,7], similar notions have been used to characterize the variation of certain image features as functions of the viewing parameters.) Quasi-invariance relaxes the strong independence constraints imposed by invariance and which may not be satisfied for all geometric entities such as complex 3-D shapes.

The class of geometric invariant and quasi-invariant properties we are interested in are those which characterize object-level shapes, not point sets. That is, they provide means to relate computable image descriptions to 3-D ones based on GCs. Thus they take explicit account of the three-dimensionality of the objects and of their desired descriptions. We have derived and used such properties in the past to segment and recover SHGCs from a real, imperfect, intensity image [32,33,35]. Here, we would like to extend that approach to handle PRGCs (derive new properties and methods) so as to be able to address even more complex, compound, objects which consist of the arrangement of several parts with both straight and curved axes.

The set of curved axis GCs is too general to be rigorously addressed. We address two broad classes of PRGCs, namely *planar right constant generalized cylinders* (PRCGCs) and *circular planar right generalized cylinders* (circular PRGCs). PRCGCs are characterized by a constant cross-section (such as tori-like objects) whereas circular PRGCs by a possibly varying size, but circular, cross-section (such as a horn). See figure 1.[1] We derive in this paper a number of rigorous geometric invariant and quasi-invariant properties of these generic shapes and describe a system that automatically detects their projections and recovers their 3-D shapes in the presence of noise, surface markings, shadows and partial occlusion. The segmentation method exploits the geometric as well as the structural properties of the objects in a multi-level perceptual grouping framework based on a hypothesize-verify paradigm. The method approximates the projection geometry by orthography.

We believe that this work constitutes a significant progress in two relevant ways. First, it extends the class of shapes which can be recovered from a single real intensity image. Second, it shows that in the absence of strict invariance, quasi-invariance can be rigorously analyzed and successfully exploited for segmentation and shape description.

One may question the generality of a method that addresses PRCGCs and circular PRGCs. This work is not meant to be independent. Rather, it is part of our larger effort to develop a system

1. Although the term PRGCs is more general, we will use it in this paper to refer to the union of PRCGCs and circular PRGCs.

1 Introduction

This paper addresses the recovery of object-level 3-D shape descriptions from a single, real, intensity image. The interest of such descriptions lies in that they are important in recognition and learning because they are rich and provide powerful and discriminatory indexing keys, in accessing large data bases of objects, and symbolic shape attributes for comparison between objects. Such 3-D descriptions also have applications in robotic grasping where they can, for example, be used to determine stable grasping modes. Our interest is to develop a system capable of *generic* 3-D descriptions from a 2-D image. To reach this objective, we need to solve a number of classical vision problems: scene segmentation, 3-D inference and shape description.

This work is an attempt to provide one of the links in realization of such a system. More specifically, we address the recovery of 3-D descriptions based on sub-classes of generalized cylinders (GCs) termed *planar right generalized cylinders* (PRGCs). These are GCs with curved, planar, axes orthogonal to their cross-sections. Generalized cylinders are known to be powerful shape descriptors, most notably for the purposes of part/whole (part-based) descriptions [2,3,6,19]. However, most past work on the segmentation and description of GCs has either addressed purely 2-dimensional descriptions based on so-called ribbons [16,19,24] or has addressed 3-D descriptions of generalized cylinders with straight axes, *straight homogeneous generalized cylinders* (SHGCs) [8,9,13,23,27,29,33]. Although some work can be found on recovery of 3-D descriptions of certain curved-axis GCs from perfect and segmented boundaries [31], no previous work has addressed the segmentation and 3-D recovery of curved-axis generalized cylinders from a real intensity image.

To do this, requires the development of methods that are able to detect objects and recover their shapes in the presence of noise, surface markings, shadows and partial occlusion. Our approach to this problem is to derive and exploit rigorous properties of the projection of the target 3-D descriptions (henceforth *projective properties*). Two types of properties are of interest: those characterizing the projection of the intrinsic GC description, which we call *geometric properties*, and those characterizing the lawful boundary interactions of the image of an object (such as junctions), which we call *structural properties*. Because it is crucial to derive shape descriptions in a viewpoint insensitive manner, we are interested in properties that are either *invariant* or *quasi-invariant* with respect to the viewing transformation.

Invariance has recently received much interest in the vision community. Most of the work has derived projective invariants based on algebraic or differential functions, typically computed on a selected set of points (or lines) [17]. Quasi-invariance has been introduced by Binford [4,5] to

Three-dimensional descriptions based on the analysis of the invariant and quasi-invariant properties of some curved-axis generalized cylinders

Mourad Zerroug and Ramakant Nevatia

PAMI Ref. No. 94-09-02

Institute for Robotics and Intelligent Systems
Powell Hall 204
University of Southern California
Los Angeles, CA 90089-0273
zerroug@iris.usc.edu

Abstract

We address the recovery of object-level 3-D descriptions of some classes of curved-axis generalized cylinders. For this, the first part of the paper analyzes the projective properties of two common generic shapes, *planar right constant generalized cylinders* (PRCGCs) and *circular planar right generalized cylinders* (circular PRGCs). The properties we analyze include new geometric invariant and quasi-invariant properties of the orthographic projection of the above shapes and a useful classification of their structural properties as functions of their pose. The second part of the paper describes an implemented system which detects and recovers PRCGCs and circular PRGCs from an intensity image in the presence of noise, surface markings, shadows and partial occlusion. The methods exploit the projective properties to hypothesize and verify relevant curved-axis objects, thus explicitly using the three-dimensionality of the objects and of the desired descriptions. This work extends past work on the recovery of volumetric shapes from an intensity image by addressing new primitives, deriving new properties and by developing a system that recovers them from an intensity image. We demonstrate our method on several real intensity images.

Keywords: Shape description, generalized cylinders, invariants, quasi-invariants, segmentation, grouping.