



HAL
open science

Single-Image Interactive 3D Generalized Cylinder Modeling

Murat Yirci, Mathieu Brédif, Nicolas Papanoditis

► **To cite this version:**

Murat Yirci, Mathieu Brédif, Nicolas Papanoditis. Single-Image Interactive 3D Generalized Cylinder Modeling. Journées Françaises d'Informatique Graphique, Nov 2016, Grenoble, France. hal-02551493

HAL Id: hal-02551493

<https://hal.science/hal-02551493v1>

Submitted on 23 Apr 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Single-Image Interactive 3D Generalized Cylinder Modeling

Murat Yirci, Mathieu Brédif and Nicolas Papanoditis

Univ. Paris-Est, LASTIG MATIS, IGN, ENSG, F-94160 Saint-Mande, France

Abstract

In this work, we present a new direct approach to reconstruct 3D models of objects in the form of homogeneous, circular generalized cylinders, featuring a planar axis. The input of the proposed method is a single image on which the user strokes guide the modeling process with a user selected prior. The main idea is the combination of human perception (via user strokes) with the computational power (reconstructing the 3D circle sections constituting the generalized cylinders). The proposed method is then successfully applied to generate models of objects in the public spaces (e.g. lamp posts, tree trunks, and pipes).

Keywords : geometric reconstruction from single images, generalized cylinders, 3D circles, perspective projection

1. Introduction

Realistic 3D city models are commonly required by virtual reality, computer games and 3D GIS applications. Different techniques have been utilized to produce these models at various levels of detail. Recently, sensor-equipped vehicles (namely mobile mapping systems) have been collecting vast amounts of close range data (images and/or lidar point clouds) in cities, which can be used to reconstruct detailed, photo-realistic models of the public spaces. In this work, we focus on modeling of the public-space objects that are in the form of generalized cylinders such as poles, lamp posts, tree trunks and pipes. The input to our modeling system is a single image on which the user strokes along the axis of the objects are used to reconstruct the model. The main idea here is to leverage both the cognitive power of human perception and the computational power of the computer.

A generalized cylinder is defined by a space curve, referred to as its axis, and a set of cross-sections centered along this axis [Bin71]. Generalized cylinders can be classified into different groups according to the properties of their cross-sections and axes [NE95]. For instance, the axis can be a straight line, a planar curve or a general space curve. Similarly, the cross-sections can be circles, ellipses or polygons. Furthermore, a generalized cylinder is called *right* (rather than *oblique*) if the normals of its cross-sections are tangent to its axis. The transformations between the cross-sections also define different types of generalized cylinders : a uniform scaling (cross-sections with the same shape but varying in size) yields homogeneous generalized cylinders, whereas enforcing all the cross-sections to be identical in size and shape results in uniform generalized cylinders. Please see Naeve and Eklundh [NE95] for the details of the taxonomy.

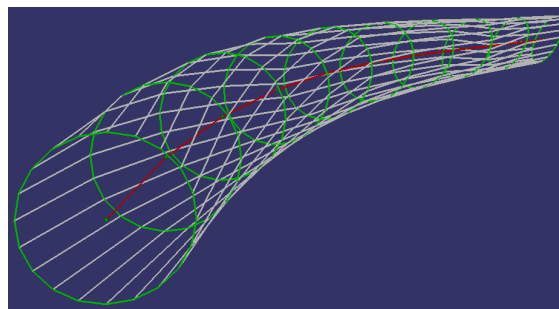


Figure 1: Homogeneous, circular generalized cylinder defined by 10 circular cross-sections and a piecewise linear axis.

In this article, we are interested in the reconstruction of right, homogeneous, circular generalized cylinders which are defined by a set of 3D circles connected to each other by 3D linear segments (Fig. 1).

1.1. Previous Works

Modeling objects from single images is an ill-posed problem due to the loss of depth information during the image formation. Therefore, all of the works in this field proposed various assumptions and constraints to infer the 3D geometry of a scene or an object. Interested readers may refer to Oswald et al. [OTNC13] for a survey of single-view modeling beyond generalized cylinder modeling.

Straight, homogeneous generalized cylinders (esp. solids of revolution) have been studied extensively. Ponce et al. [PCM89] analyzed the orthographic projections of these cylinders using differential geometry and identified several invariants. Later, his results were extended for perspective camera models [RDLR91] and then other researchers exploited these works to generate models of objects from single images [SB93, UN95].

We have been inspired by 3-Sweep method [CZS*13] which is a part-based modeling system to model man-made objects from single images. A part of their work is related to generalized cylinder modeling in which the user strokes are first used to reconstruct a sketchy model which is later enhanced in an optimization process. The initial model is composed of a set of 3D circles (constructed with a ratio-preserving projection model) whose centers are positioned on the user drawn axis in the image plane. The projection of the initial model approximates the object's outline / silhouette in the image but it does not construct a coherent model in 3D space. The task of the optimization process is to correct the geometry of the model (i.e. the orientation and position of the 3D circles and the axis of the generalized cylinder) while trying to keep the model's projection consistent with the object's outline in the image space. However, in order to achieve this, the circular cross-section constraint is relaxed which results in elliptic generalized cylinder models.

In this work, we propose an approach similar to [CZS*13] for the user input, but reconstruct the model directly from the perspective camera model based on an overall axis prior, which does not need further optimization and is guaranteed to generate right, circular and homogeneous generalized cylinder models.

2. Modeling Generalized Cylinders

2.1. Human Computer Interaction - Input

In our modeling system, the user initiates the modeling by drawing an ellipse on the input image which is the projection of the first 3D circle located at the beginning of the generalized cylinder. Then, the user piecewisely draws the projection of the axis by dragging the mouse pointer over the generalized cylinder image. During the dragging, at every user clicked point, a copy of the previous major axis is fit to the outline of the generalized cylinder. The fit is performed by first rotating the copied major axis so that it becomes perpendicular to last axis segment and then shrinking or stretching its length until an edge of the outline is encountered (for this purpose, the edges in the input image are precomputed as a preprocess). (Fig. 2).

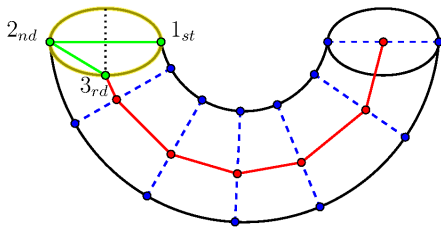


Figure 2: User strokes on the input image : the initial ellipse (yellow) is drawn by three clicks (green), the projection of the axis (red) is piecewisely drawn, and the set of the major axes (blue) fit to the outline of the generalized cylinder at the user-clicked points.

2.2. Reconstructing the Initial 3D Circle Under Perspective Projection

The initial problem to be solved is the reconstruction of the first 3D circle whose projection is fully drawn by the

user on the input image. We adopted the analytic method of Safaee-Rad et al. [SRTSB92] to compute two possible 3D circles (except degenerate cases) in the camera coordinate frame (Fig. 3). Note that, there exists an infinite number of 3D circles with the same projection, lying on the two pencils of parallel planes. A chosen radius determines two of these 3D circles. In order to eliminate one of the two possible 3D circle pencils, we propose to ask the user to finalize the ellipse drawing by clicking the closest end of the minor axis w.r.t. the viewing point or camera center. This is used to identify whether the generalized cylinder is seen from bottom to top or vice-versa (Fig. 3, bottom left).

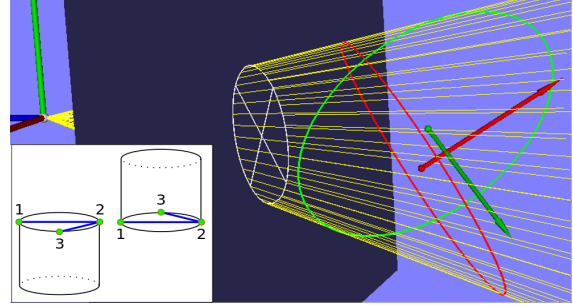


Figure 3: Two possible 3D circles with a fixed radius whose perspective projections are the same ellipse on the image plane. To differentiate between the two possible 3D circles, the user is asked to finalize the ellipse drawing at the closest end of the minor axis w.r.t. the camera center (bottom left).

2.3. Parametric Solution to 3D Circle Center

The perspective projection of a 3D circle might be unintuitive at first since perspective projection does not preserve the ratio of lengths. Therefore, for non-degenerate cases, neither the 3D circle centers are projected as the ellipse centers nor the ellipse axes are the projections of the circle diameters which complicate the computations (Fig. 4).

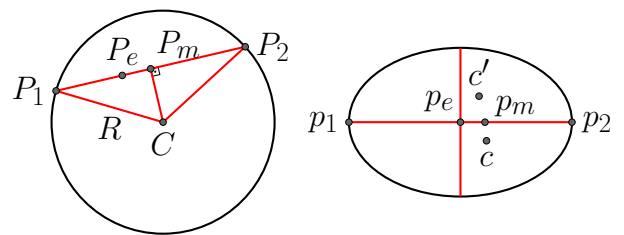


Figure 4: The perspective projection of a 3D circle : the points with lower-case labels on the ellipse are the projections of the upper-case points on the 3D circle. The circle center is C and R is the radius. P_m is the midpoint of P_1 and P_2 . p_1 and p_2 are the major axis endpoints, p_e is the ellipse center and c' is the projected center of the 3D circles of the other pencil that share the same ellipse projection.

We have seen in section 2.2 that the six parameters of the first 3D circle are computed analytically from the given ellipse (up to an overall depth and radius scaling) : unit normal (2D), center position (3D), and radius (1D). However, we only partially observe the projections of the remaining 3D circles : only the major axes of the projected ellipses are

known (Fig. 2). In order to reconstruct these 3D circles, we first define the following problem : compute the family of 3D circles whose perspective projected ellipses have the given major axis. The solution to this problem defines a parametric representation for the centers of the circles in terms of the unit normal (\hat{n}) and the radius (R). To solve the defined problem, we identified three distinct planes (Π_1, Π_2, Π_3) on which the center of the circle lies. Interested readers may refer to Yirci [Yir16] for the details of the derived parametric solution. Here, we only present the outlines of the derivations. For the pinhole camera model, the coordinates of the points p_i and P_i for $i = 1, 2$ (Fig. 4) can be written as $p_i = [x_i \ y_i]$ and $P_i = Z_i [x_i/f \ y_i/f \ 1]$ where x_i and y_i are the known pixel coordinates on the input image, f is the given focal length of the camera, and Z_i is the unknown depth of the point P_i in the camera coordinate frame.

Offset planes, $\Pi_1(\hat{n}, R), \Pi_2(\hat{n}, R)$: Let l_i ($i = 1, 2$) be the two 2D tangent lines on the image plane passing through p_i and perpendicular to the ellipse major axis. Then the two planes (T_i) passing through the camera center and l_i are tangent to the 3D circles at the points P_i . In 3D projective space, these tangent planes can be computed by $T_i = M^T l_i$, where M is the 3x4 camera projection matrix. Furthermore, we know from 3D projective geometry that a tangent plane to a dual quadric has to satisfy the following equation : $T_i^T Q^* T_i = 0$. We applied this equation together with the dual quadric representation of a 3D circle [SB14] which yields two planes Π_i passing through the center C of the 3D circle. These planes are called offset planes since they come out to be offset versions of the tangent planes T_i . Offset planes are parametrized by R and \hat{n} .

Bisector plane, Π_3 : The plane passing through C and P_m (Fig. 4) that is perpendicular to the supporting plane of the circle. The equation of the bisector plane can be obtained by $(C - P_m) \cdot (P_1 - P_2) = 0$. Notice that the bisector plane is the locus of the points whose distance to P_1 and P_2 are equal. Bisector plane is parametrized by \hat{n} and Z_2 (note that Z_1 could also be used equally).

The intersection of Π_1, Π_2 and Π_3 can be computed by the Cramer's rule which yields a parametric solution to the circle centers (Equation 1).

$$C(\hat{n}, R, Z_2) = R \begin{bmatrix} s_1 \\ s_3 \\ s_5 \end{bmatrix} + Z_2 \begin{bmatrix} s_2 \\ s_4 \\ 0 \end{bmatrix} \quad (1)$$

where s_i (for $i = 1 \dots 5$) are functions of \hat{n}, p_1, p_2 , and f [Yir16]. Note that p_1, p_2 , and f are known for all 3D circles contrary to \hat{n}, R , and Z_2 . However, another relation $R = \|C - P_2\|$ is used to eliminate Z_2 from the Equation 1 which results in a 3D family of 3D circle centers $C(\hat{n}, R)$. On the contrary to the initial 3D circle estimation, this problem is not ambiguous : there is only a single R -parametrized family of circle centres, given the normal \hat{n}, p_1, p_2, f .

2.4. Different Priors to Reconstruct Generalized Cylinders

To sum up, the initial circle is reconstructed upto a scale by the methods described in section 2.2 and in order to compute the remaining 3D circles that constitute the generalized

cylinder, a parametric solution to circle centers is derived in 2.3. However, to be able to use this parametric equation, we still need to estimate the scale (radius) or depth (the z -coordinate of the circle center w.r.t. camera frame) and the normals of the 3D circles. We propose to utilize a prior selected by the user according to the properties of the generalized cylinders to be modeled.

Planar axis prior : If the centers of the 3D circles forming the generalized cylinder are on the same plane (i.e. generalized cylinder has a planar axis), then the user can use this prior. However, the user input described in section 2.1 is not sufficient to compute the plane of the axis. Therefore, for this prior, the user is asked to finish the modeling on the other end of the generalized cylinder by also drawing the projection of the last 3D circle. As a result, the two 3D circles (first and last) can be computed analytically upto a scale (section 2.2). Since the normals of these circles lie on the axis plane, the cross product of their normals yields the normal of the axis plane. The next step is the selection of an arbitrary depth or scale for the first 3D circle, which defines a point on the axis plane. Using the normal and the point, the axis plane is computed and the scale (or depth) of the last 3D circle is chosen such that its center lies on the axis plane. Finally, the normals of the intermediate 3D circles (between the first and the last 3D circles) are taken as parallel to the user drawn axis in the image plane after it is projected onto the computed axis plane. These normals are first used to estimate unit circles using the parametric center equation (section 2.3), and then they are scaled so that their centers lie on the axis plane. As a result, a coherent model of the generalized cylinder can be reconstructed out of the computed 3D circles.

Linear axis prior : If the generalized cylinder to be modeled is a straight cylinder, then the planar axis method cannot be applied successfully since the normals of the first and the last 3D circles become parallel (the cross product of their normals become zero). In fact, the reconstruction problem is easier for the straight generalized cylinders because the whole axis lies on a 3D line whose direction is defined by the normal of the first 3D circle. Once an arbitrary depth or scale is selected for the first 3D circle, the axis line can be constructed. The normals of the intermediate 3D circles are equal to the axis direction and their scales are computed so that the circle centers are aligned on the axis with minimum least squares error.

Figure 5 displays examples of reconstructed models. Before starting to model, the user selects the appropriate prior for the object to be modeled. If the selected prior does not fit well to the object to be modeled, then the generated model might not be plausible.

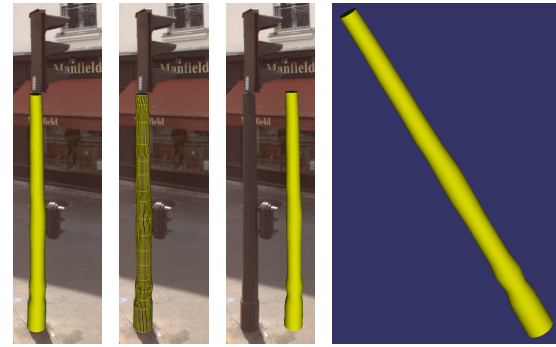
3. Conclusion

We have presented a new direct approach to model right, homogeneous, circular generalized cylinders with planar axis. The user helps the modeling system to reconstruct the models by performing the cognitive tasks in which the computers are not yet good enough. Then, the system quickly computes the 3D circles in a split second that forms the generalized cylinder.

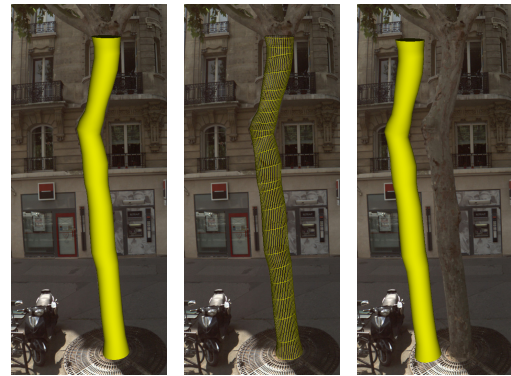
An object can only be modeled up to a scale factor from a single-view when there is not any prior information about the depth or size of the object. On the other hand, correctly scaled and geo-localized models can be obtained with multiple views which can be very important for 3D GIS applications. Another limitation of the single-view modeling is that a single image cannot capture the dimensions of the objects along the viewing direction, therefore those features cannot be modeled. Also note that, integrating other views into the modeling system should improve the accuracy of the reconstructed models. Thus, multi-view modeling is the first future work to be done which can be followed by automatic texture extraction from the images and extending the type of objects that can be modeled such as polygonal generalized cylinders.

References

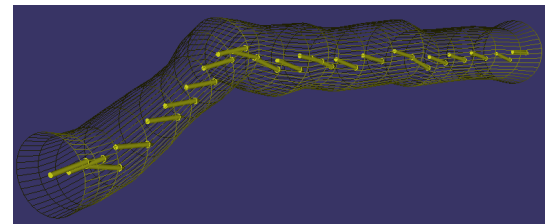
- [Bin71] BINFORD T. O. : Visual perception by computer. In *IEEE Conference on Systems and Controls* (1971).
- [CZS*13] CHEN T., ZHU Z., SHAMIR A., HU S., COHEN-OR D. : 3-sweep, extracting editable objects from a single photo. *ACM TOG* (2013).
- [NE95] NAEVE A., EKLUNDH J. O. : Representing generalized cylinders. In *Proceedings of Europe-China Workshop on Geometrical Modelling and Invariants for Computer Vision* (1995), pp. 63–70.
- [OTNC13] OSWALD M. R., TÖPPE E., NIEUWENHUIS C., CREMERS D. : A review of geometry recovery from a single image focusing on curved object reconstruction. In *Innovations for Shape Analysis*. Springer Berlin Heidelberg, 2013, pp. 343–378.
- [PCM89] PONCE J., CHELBERG D., MANN W. B. : Invariant properties of straight homogeneous generalized cylinders and their contours. *IEEE Trans. Pattern Anal. Mach. Intell* (1989).
- [RDLR91] RICHTIN M., DHOME M., LAPRESTE J. T., RIVES G. : Inverse perspective transform using zero-curvature contour points : application to the localization of some generalized cylinders from a single view. *IEEE Trans. Pattern Anal. Mach. Intell* (1991).
- [SB93] SATO H., BINFORD T. O. : Finding and recovering shgc objects in an edge image. *CVGIP : Image Understanding* (1993).
- [SB14] SOHEILIAN B., BRÉDIF M. : Multi-view 3d circular target reconstruction with uncertainty analysis. *ISPRS Annals. Vol. Volume II-3* (2014), 143–148.
- [SRTSB92] SAFAEE-RAD R., TCHOUKANOV I., SMITH K., BENHABIB B. : Three-dimensional location estimation of circular features for machine vision. *IEEE Trans. Robot. Autom* (1992).
- [UN95] ULUPINAR F., NEVATIA R. : Shape from contour : Straight homogeneous generalized cylinders and constant cross section generalized cylinders. *IEEE Trans. Pattern Anal. Mach. Intell* (1995).
- [Yir16] YIRCI M. : *2D Arrangements for Public Space Mapping and Transportation*. PhD thesis, Université Paris-Est, 2016.



(a), (b), (c) and (d) : Lamp post model.



(e), (f), (g) and (h) : Tree trunk model.



(i), (j), (k) and (l) : Pipe model.

Figure 5: Three models are generated using different priors : a lamp post from the linear axis prior, a tree trunk and a pipe from the planar axis prior. (a), (b), (e), (f) and (j) display the models right after the user finishes the modeling, (c), (d), (g), (h), (k) and (l) represent the models at different positions and orientations and (i) displays the input image.