

Absolute pose and structure from motion for surfaces of revolution: minimal problems using apparent contours

Cody J. Phillips, Kostas Danillidis

GRASP Laboratory

University of Pennsylvania

{codyp, kostas}@cis.upenn.edu

Abstract

The class of objects that can be represented by surfaces of revolution (SoRs) is highly prevalent in human work and living spaces. Due to their prevalence and convenient geometric properties, SoRs have been employed over the past thirty years for single-view camera calibration and pose estimation, and have been studied in terms of SoR object reconstruction and recognition.

Such treatment has provided techniques for the automatic identification and classification of important SoR structures, such as apparent contours, cross sections, bitangent points, creases, and inflections. The presence of these structures are crucial to most SoR-based image metrology algorithms. This paper develops single-view and two-view pose recovery and reconstruction formulations that only require apparent contours, and no other SoR features.

The primary objective of this paper is to present and experimentally validate the minimal problems pertaining to SoR metrology from apparent contours. For a single view with a known reference model, this includes absolute pose recovery. For many views and no reference model this is extended to structure from motion (SfM). Assuming apparent contours as input that have been identified and segmented with reasonable accuracy, the minimal problems are demonstrated to produce accurate SoR pose and shape results when used as part of a RANSAC-based hypothesis generation and evaluation pipeline.

1. Introduction

The geometry of SoRs has received a fair amount of attention in the literature, specifically their projectively invariant properties and resulting utility as calibration objects. The lower-level vision tasks of segmentation, grouping, and classification of SoR structures has been covered extensively [19, 20, 22, 4, 16] as the important first step towards per-

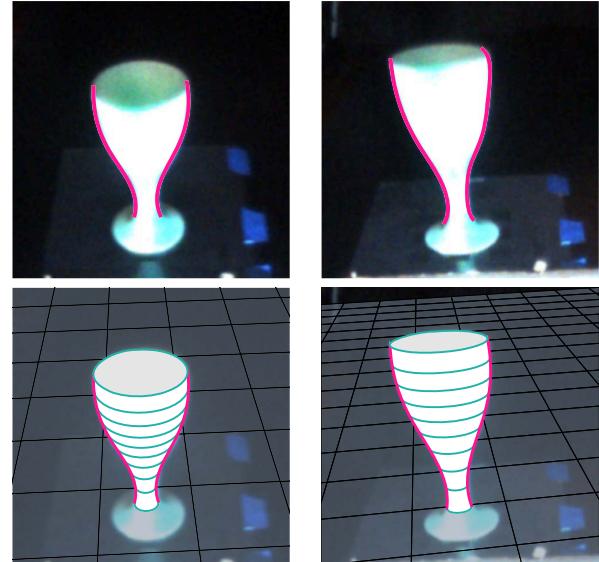


Figure 1. Given two pairs of apparent contours (shown in pink) from two different views (top) with known camera intrinsics, the minimal problems formulated in this paper enables structure from motion pose and shape recovery (bottom). In the case of a known generatrix, the absolute pose can be recovered from a single view.

forming higher-level SoR-metrology tasks.

This paper concentrates on three such high-level tasks with the goal of removing their dependency on special SoR structures required by prior approaches. These tasks are

1. SoR homology axis estimation, the recovery of the projected axis of revolution
2. SoR absolute pose estimation, the 2D-to-3D registration of image contours to a reference shape model
3. SoR shape reconstruction (SfM), the recovery of the shape and pose up to a scale.

This work assumes a calibrated camera, images with the SoR apparent contours segmented and classified to a reasonable degree of accuracy, as well as a reference model

generatrix in the case of absolute pose estimation. It is distinguished from prior work in that it requires no special points or cross-sections.

1.1. Contributions

We present new minimal problems that have application when a segmented surface of revolution is visible in one or two views. As opposed to several articles in the geometry of SoRs we do not assume the visibility of any circular cross-section or bi-tangent points. In particular we prove that for calibrated cameras

- We can find the absolute pose of a known surface of revolution from a single point-tangent pair.
- We can solve structure from motion for unknown surfaces of revolution up to a rotation around the SoR axis and a scale from two corresponding pairs of point-tangents.
- We provide a least-squares solution for inliers with very satisfactory results in a simple specialization of SfM to visual odometry.

1.2. Prior work

The relationship between algebraic surfaces and the deformation of their imaged profile under changes in viewpoint was developed by [10, 1, 2], laying the groundwork for structure from motion. The imaged profile of a smooth surface is called its *apparent contour*, while its corresponding 3D surface boundary called the *contour generator*. Their work assumes that some unknown contour generator has undergone central projection through a calibrated camera with a known center in world coordinates to produce known apparent contours. With a calibrated camera and known camera trajectory, apparent contour points are treated as rays in Euclidean space that must obey a set of geometric constraints imposed by the underlying surface geometry. For a single view, these constraints state that apparent contour rays must be tangent to and intersect the unknown surface, therefore providing surface normal and depth information.

The relationship between the 2D projection of the SoR axis and the apparent contours was initially treated as “pseudo-symmetric” [11, 13, 20, 7], referring to the fact that as a projection becomes closer to orthographic, the apparent contours become closer to being symmetric in the image. The understanding of this relationship has been incrementally refined [11, 22, 14] and was established by [22] as a 4-DoF class of skew-symmetric transforms called *planar harmonic homologies*.

Early work on SoRs [18, 9] explore shape profile properties that are viewpoint and projectively invariant, such as bi-tangent points, creases, and inflections. These features,

together with circular cross sections, can provide a great deal of shape and pose information, especially if the camera is uncalibrated.

Bi-tangent points [22, 14] and visible cross sections [7, 5, 4, 6, 3] have been used to solve for the SoR axis homology. With known camera calibration, the homology reduces [21, 5, 4, 6, 3] to having 2-DoF, from which a rotation matrix that encodes the camera’s roll and yaw can be recovered [21]. In the case that cross sections are used, the camera pitch attitude and position are constrained in addition to the roll and yaw.

In [15], pose estimation and 3D reconstruction from apparent contours are addressed for multiple views, but the solutions are applicable only to cylindrical objects. Apparent contours alone without cross sections or special points are used in [17, 16] to perform metric SoR pose recovery and reconstruction, but the relative transform between the two cameras is assumed to be known.

To the authors’ knowledge, all prior methods for SoR pose recovery, and therefore Euclidean shape reconstruction, assume bi-tangent points, cross sections, “pseudo-symmetry” (a near-orthographic camera), or other additional assumptions and constraints.

1.3. Surfaces of Revolution

Surfaces of revolution are generated by sweeping a planar curve or *generatrix function* about an axis of revolution. Cylindrical coordinates are therefore a natural way of expressing SoR points and normals, with height extending along the z -axis, the axis of revolution. The generatrix function $r(h)$ determines the radius for a given height and has a corresponding derivative function $\rho(h)$ that yields the derivative $\frac{dr(h)}{dh}$.

Using these two functions, surface points and normals are parameterized by height h and azimuth θ as

$$\mathbf{x}_o(\theta, h) = (r(h) \cos \theta, r(h) \sin \theta, h)^T \quad (1)$$

and

$$\mathbf{n}_o(\theta, h) \propto (\cos \theta, \sin \theta, -\rho)^T \quad (2)$$

respectively. Points of the same azimuth θ are said to belong to the same *meridian*, which is the half plane π_m extends from $\hat{\mathbf{z}}$ outwards. The normal of any SoR point must also be contained in the same meridian, yielding the meridian constraint

$$(\hat{\mathbf{z}}_o \times \mathbf{n}_o)^T \mathbf{x}_o = 0. \quad (3)$$

2. Forward projection

The contour generator of an SoR is only dependant on the generatrix and the location of the 3D center of projection. Due to rotational symmetry, the center point has only 2-DoF and can be viewed as the cameras height \hat{h} above the

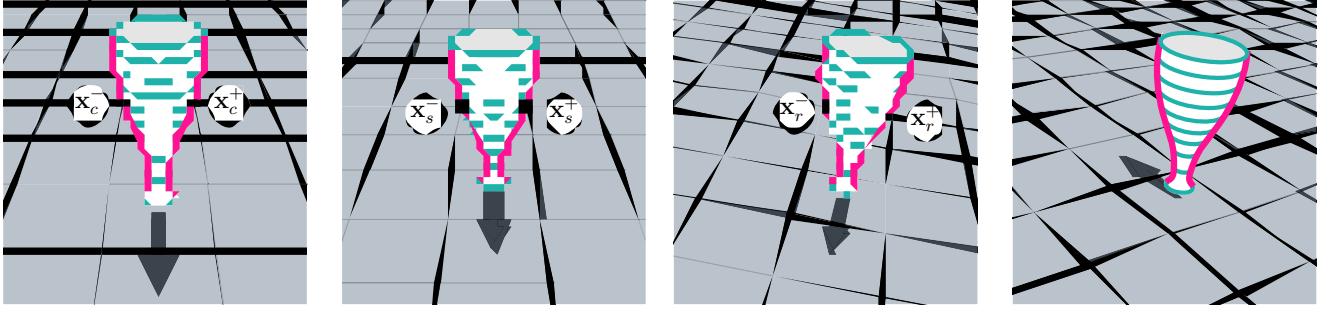


Figure 2. The forward projection of an SoR is composed of a sequence of rotations. Pose recovery proceeds backwards from right to left.

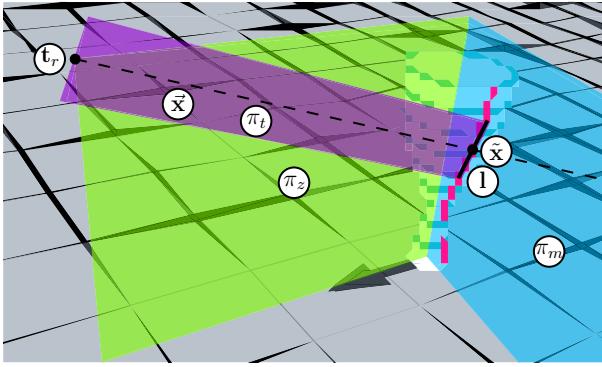


Figure 3. SoR relationships: The image ray \tilde{x} travels from camera center $\mathbf{t}_r = (\hat{d}, 0, \hat{h})^\top$ along purple tangent plane π_t (containing I) and strikes the SoR at contour generator point \tilde{x} . It is perpendicular to the blue meridian plane π_m (containing \tilde{x}).

ground plane and distance \hat{d} along the ground plane from the SoR axis. The formation of the contour generator, and by extension apparent contour, can be factorized into 5 distinct steps that each encode various geometric parameters.

2.1. Contour generator

The first step in apparent contour formation encodes the generatrix, along with the camera height \hat{h} and distance \hat{d} along the x - y SoR ground plane. At this point the SoR camera is the *canonical rotation*,

$$\mathbf{R}_c^o \stackrel{\text{def}}{=} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ -1 & 0 & 0 \end{bmatrix}, \quad (4)$$

with its z -axis parallel to the x - y SoR ground plane and contained within the x - z SoR plane, represented in green in fig. 3. The contour generator in the canonical pose is

parameterized by height h as

$$\tilde{\mathbf{x}}_o(h) = (\tilde{x}_o, \tilde{y}_o, \tilde{z}_o)^\top \quad (5)$$

$$\tilde{x}_o = r_h^2 + r_h \rho_h(\hat{h} - h)/\hat{d} \quad (6)$$

$$\tilde{y}_o = \pm \sqrt{r_h^2 - \tilde{x}_o^2} \quad (7)$$

$$\tilde{z}_o = h. \quad (8)$$

This form is similar to the parameterization from [7, 12]. It is derived from the surface-tangency constraint and surface-intersection constraints explored in [10, 1, 2], specialized for SoRs. The surface-intersection constraint requires that $\tilde{\mathbf{x}}_o$ is a SoR surface point, respecting eq. 1. The surface-tangency constraint requires that the camera ray \mathbf{x}_o be tangent to the surface, respecting

$$\tilde{\mathbf{x}}_o^\top \mathbf{x}_o = 0. \quad (9)$$

This is represented by the purple plane in fig. 3, with normal \mathbf{n}_o .

Notice that \tilde{y}_o in eq. 7 has two solutions with opposite signs. This implies that there are two symmetric contour generators for an SoR in the canonical rotation.

2.2. Transformation to general view

The remaining 4 steps in apparent contour formation transform the contour generator points from the canonical pose to a general pose with arbitrary camera rotation and translation with respect to the SoR coordinate system. In expanded form, this transformation is accomplished by

$$\tilde{\mathbf{x}}_g = \mathbf{R}_g^r \mathbf{R}_r^s \mathbf{R}_s^c \mathbf{R}_c^o \tilde{\mathbf{x}}_o + \mathbf{R}_g^r \mathbf{t}_r. \quad (10)$$

Each term encodes a specific aspect of the pose and allows for the pose recovery process to be similarly factorized into subproblems. The effect of each step on the resulting image is visualized in fig. 2.

The first rotation term \mathbf{R}_c^o was already presented in eq. 4 as the relationship between canonical SoR and camera poses, transforming SoR points to the camera basis (fig. 2a). The second term

$$\mathbf{R}_s^c = \mathbf{R}_x(\phi_s^c) \quad (11)$$

encodes the camera's pitch attitude (fig. 2b), while the third term \mathbf{R}_r^s encodes the camera roll and yaw (fig. 2c).

The contour generators are symmetric through the camera's y - z plane in the canonical pose, and remains so under a rotation about the x -axis. The application of \mathbf{R}_r^s breaks this symmetry, while its inverse \mathbf{R}_s^r restores it. Therefore this is called the *symmetry rotation*.

The final transformation \mathbf{R}_g^r applied both to the rotation and translation components represents a single-axis motion about the SoRs z -axis. Due to the SoRs rotational symmetry, this motion has no effect on the location of the contour generators with respect to the camera, as can be seen by comparing fig. 2c and fig. 2d. This invariance produces an equivalence relation defined over all poses, with an equivalence class having a representative translation of $\mathbf{t}_r = (\hat{d}, 0, \hat{h})^\top$. Accordingly, it is only possible to recover the pose equivalence class from an SoR view.

3. Back projection

3.1. Generatrix recovery

With a known representative pose defined by $(\mathbf{R}_o^r, \hat{d}, \hat{h})$, it is possible to recover the generatrix values $(\mathbf{r}_h, \rho_h, h)$ corresponding to a given representative apparent contour homogeneous point \mathbf{x}_r and tangent homogeneous line \mathbf{l}_r . Rotating these homogeneous vectors into the SoR coordinate system, we have the direction vector

$$\mathbf{x}_o = (x_o, y_o, z_o)^\top \quad (12)$$

of the camera ray

$$\tilde{\mathbf{x}}_o = \lambda \mathbf{x}_o + \mathbf{t}_o \quad (13)$$

that must be tangent to and intersect the SoR surface. The rotated homogeneous tangent line

$$\mathbf{l}_o = (u'_o, v_o, w_o)^\top \quad (14)$$

encodes the surface normal, which is recovered as

$$\mathbf{n}_o \propto \begin{pmatrix} u_o \\ v_o \\ w_o \end{pmatrix}, u_o = -\frac{y_o v_o + z_o w_o}{x_o} \quad (15)$$

using the tangent constraint in eq. 9. The depth value λ of eq. 13 is recovered as

$$\lambda = \frac{v_o \hat{d}}{u_o y_o - v_o x_o} \quad (16)$$

by applying the meridian constraint of eq. 3. With a known distance λ , the 3D contour generator point $\tilde{\mathbf{x}}_o$ is recovered directly, along with the generatrix radius

$$\mathbf{r}(h) = \sqrt{x_o^2 + y_o^2}. \quad (17)$$

From eq. 2 relating ρ to \mathbf{n}_o , it is recovered by normalizing the first two components as

$$\rho(h) = \frac{w_o}{\sqrt{u_o^2 + v_o^2}}. \quad (18)$$

These equations are derived similarly to the forms found in [21].

3.2. Symmetry rotation recovery minimal problem

Given two corresponding apparent contour points \mathbf{a} and \mathbf{b} from a representative camera view, the symmetry rotation \mathbf{R}_s^r is recovered as

$$\mathbf{R}_s^r = (\hat{\mathbf{x}} \quad \hat{\mathbf{y}} \quad \hat{\mathbf{z}})^\top \quad (19)$$

$$\hat{\mathbf{x}} = \frac{\mathbf{b} - \mathbf{a}}{\|\mathbf{b} - \mathbf{a}\|}, \hat{\mathbf{y}} = \frac{\mathbf{a} \times \mathbf{b}}{\|\mathbf{a} \times \mathbf{b}\|}, \hat{\mathbf{z}} = \frac{\hat{\mathbf{x}} \times \hat{\mathbf{y}}}{\|\hat{\mathbf{x}} \times \hat{\mathbf{y}}\|} \quad (20)$$

Additionally, the two symmetric apparent contours are related by

$$\mathbf{H}_s = \mathbf{R}_r^s \text{diag}(1, -1, 1) \mathbf{R}_s^r \quad (21)$$

Such a recovery is possible because it is known that \mathbf{a} and \mathbf{b} must be symmetric in the canonical view, that $\mathbf{a} - \mathbf{b}$ must be in the direction of the camera x -axis, and $\mathbf{a} \times \mathbf{b}$ must lie in the camera's y - z plane.

4. Absolute pose minimal problems

4.1. Attitude recovery from generatrix derivative

Given an apparent contour point \mathbf{x}_s and tangent line \mathbf{l}_s in the symmetric view, along with its corresponding generatrix derivative ρ , it is possible to solve for the attitude ϕ that relates the symmetric and canonical view. If the corresponding values of \mathbf{x}_c and \mathbf{l}_c in the canonical view were known, the computation of ϕ would be trivial.

From the forward projection eq. 10, symmetric-view direction vectors transform to the SoR system by

$$\begin{pmatrix} a_o \\ b_o \\ c_o \end{pmatrix} = \begin{pmatrix} -b_s \sin \phi - c_s \cos \phi \\ a_s \\ -b_s \cos \phi + c_s \sin \phi \end{pmatrix}. \quad (22)$$

Applying this to \mathbf{x}_s and \mathbf{l}_s along with the Pythagorean identity allows the ρ -recovery parameterization from eq. 18 to be written as

$$\rho = \frac{w_o}{\sqrt{u_s^2 + v_s^2 + w_s^2 - w_o^2}}, \quad (23)$$

which simplifies to

$$\rho = \frac{w_o}{\sqrt{-w_o^2 + 1}}, \quad (24)$$

by normalizing \mathbf{l}_s . Solving for w_o and representing it terms of the rotated line \mathbf{l}_s provides the relationship between ρ and ϕ ,

$$(p^{-2} + 1)^{-1/2} = -v_s \cos \phi + w_s \sin \phi, \quad (25)$$

expressed as a single sinusoid by the phase shift identity

$$(\rho^{-2} + 1)^{-1/2} = k \sin(\phi + \psi) \quad (26)$$

with constants

$$k = \sqrt{v_s^2 + w_s^2}, \quad \psi = \arctan2(w_s, v_s).$$

The solution is constrained to lie in front of the camera with

$$-\pi/2 \leq \phi \leq \pi/2.$$

4.2. Translation recovery from generatrix point

Given an apparent contour point \mathbf{x}_c and tangent line \mathbf{l}_c in the symmetric view along with its corresponding generatrix point (r, h) , it is possible to solve for the representative camera distance \hat{d} and height \hat{h} . By the definition of SoR points and normals in eq. 1 and 2, the contour generator point is recovered by scaling the normal \mathbf{n}_o recovered from \mathbf{x}_c and \mathbf{l}_c by r as

$$\begin{pmatrix} \tilde{x}_o \\ \tilde{y}_o \end{pmatrix} = r \begin{pmatrix} u_o \\ v_o \end{pmatrix}. \quad (27)$$

From these values and the camera ray eq. 13, the camera translation is constrained to be

$$\begin{pmatrix} \hat{d} \\ \hat{h} \end{pmatrix} = \begin{pmatrix} \tilde{x}_o \\ h \end{pmatrix} - \frac{rv_o}{\tilde{y}_o} \begin{pmatrix} u_o \\ z_o \end{pmatrix}. \quad (28)$$

4.3. Absolute pose refinement

Assuming initial pose parameters $\Omega_0 = (\phi, \hat{d}, \hat{h})$, the contour generator $\tilde{X}^a(\Omega_0)$ is recovered from the measured image data and the synthesized contour generator $\tilde{X}^b(\Omega_0)$ is created from the generatrix. These contour generators are used to provide n point correspondences

$$(\mathbf{x}_i^a, \mathbf{l}_i^a) \Leftrightarrow (\mathbf{x}_i^b, \mathbf{l}_i^b).$$

With the synthetic contour generator assumed to be correct, an update rule for Ω is desired that minimizes the sum squared contour generator reconstruction error

$$\operatorname{argmin}_{\Omega} \Omega E = \sum_i^n |(\tilde{\mathbf{x}}_i^a(\Omega) - \tilde{\mathbf{x}}_i^b)^2|_1. \quad (29)$$

The attitude parameter ϕ is updated alternatively with the translation parameters (\hat{d}, \hat{h}) , while all other parameters held constant.

4.4. Attitude update rule

The attitude parameter ϕ is updated to minimize the mean angular error between corresponding synthesized and measured contour generator points. This is accomplished by subtracted off the mean angular error to yield the updated attitude

$$\phi_{t+1} = \phi_t - \frac{1}{n} \sum_i^n [\arctan2(z_i^b, y_i^b) - \arctan2(z_i^a, y_i^a)]. \quad (30)$$

4.5. Translation update rule

Combining the depth recovery eq. 16 and camera ray eq. 13, a data-based contour generator point is expressed as a function of Ω as

$$\tilde{\mathbf{x}}_i^a = \mu_i \hat{d} \begin{pmatrix} x_o \\ y_o \\ z_o \end{pmatrix} + \begin{pmatrix} \hat{d} \\ 0 \\ \hat{h} \end{pmatrix}, \quad (31)$$

$$\mu_i = \frac{v_i}{u_i y_i - x_i v_i}. \quad (32)$$

With fixed synthesized model-based contour generator points $\tilde{\mathbf{x}}_i^b = (\tilde{x}_i, \tilde{y}_i, \tilde{z}_i)^\top$, the reconstruction error eq. 29 is expressed as the sum of three components,

$$E = E_x + E_y + E_z \quad (33)$$

$$E_x = \sum_i \left(\mu_i x_i \hat{d} + \hat{d} - \tilde{x}_i \right)^2 \quad (34)$$

$$E_y = \sum_i \left(\mu_i y_i \hat{d} - \tilde{y}_i \right)^2 \quad (35)$$

$$E_z = \sum_i \left(\mu_i z_i \hat{d} + \hat{h} - \tilde{z}_i \right)^2. \quad (36)$$

This error function is minimized by the values of \hat{d} and \hat{h} that satisfy

$$\frac{\partial E}{\partial \hat{d}} = 0, \quad \frac{\partial E}{\partial \hat{h}} = 0 \quad (37)$$

4.5.1 Minimizing error with respect to \hat{h}

The optimal value for \hat{h} is solved directly from $\frac{\partial E}{\partial \hat{h}} = 0$ as

$$0 = \cancel{\frac{\partial E_x}{\partial \hat{h}}}^0 + \cancel{\frac{\partial E_y}{\partial \hat{h}}}^0 + \frac{\partial E_z}{\partial \hat{h}} \quad (38)$$

$$\hat{h} = \hat{d} H_1 + H_2 \quad (39)$$

$$H_1 = -\frac{1}{n} \sum_i \mu_i z_i \quad (40)$$

$$H_2 = \frac{1}{n} \sum_i \tilde{z}_i \quad (41)$$

4.5.2 Minimizing error with respect to \hat{d}

Substituting in \hat{h} , \hat{d} is then solved as

$$0 = \frac{\partial E_x}{\partial \hat{d}} + \frac{\partial E_y}{\partial \hat{d}} + \frac{\partial E_z}{\partial \hat{d}} \quad (42)$$

$$0 = \hat{d}D_1 + \hat{h}D_2 + D_3 \quad (43)$$

$$D_1 = 2 \sum_i (a_i^2 + b_i^2 + c_i^2) \quad (44)$$

$$D_2 = 2 \sum_i c_i \quad (45)$$

$$D_3 = -2 \sum_i (a_i \tilde{x}_i + b_i \tilde{y}_i + c_i \tilde{z}_i) \quad (46)$$

$$a_i = (\mu_i x_i + 1), b_i = \mu_i y_i, c_i = \mu_i z_i \quad (47)$$

4.5.3 Summarized update rules

The optimal update values for (\hat{d}, \hat{h}) are summarized as

$$\hat{d} = -\frac{H_2 D_2 + D_3}{D_1 + H_1 D_2} \quad (48)$$

$$\hat{h} = \hat{d} H_1 + H_2, \quad (49)$$

with the value of \hat{h} dependant on \hat{d} .

5. Structure from motion minimal problems

5.1. Two-view attitude recovery

In sec. 4.1 the attitude ϕ is recovered by solving it in terms of a known generatrix derivative ρ by eq. 26. In the SfM formulation, there are two unknown attitudes, ϕ_a and ϕ_b for views a and b . Additionally, the generatrix and therefore derivative ρ is also unknown.

Equation 26 is still applicable in SfM without explicitly knowing any ρ values if two pairs of contour generator correspondences,

$$\begin{aligned} (\mathbf{l}_{a1}, \mathbf{x}_{a1}) &\Leftrightarrow (\mathbf{l}_{b1}, \mathbf{x}_{b1}) \\ (\mathbf{l}_{a2}, \mathbf{x}_{a2}) &\Leftrightarrow (\mathbf{l}_{b2}, \mathbf{x}_{b2}), \end{aligned}$$

across views a and b .

Since corresponding points must have the same derivative ρ , each pair creates an equality when substituted into eq. 26, yielding two equations,

$$k_{a1} \sin(\phi_a + \psi_{a1}) = k_{b1} \sin(\phi_b + \psi_{b1}) \quad (50)$$

$$k_{a2} \sin(\phi_a + \psi_{a2}) = k_{b2} \sin(\phi_b + \psi_{b2}), \quad (51)$$

with two unknowns. The attitudes are recovered by solving this system of equations.

5.1.1 Converting to quadratic form

Relating the two rotation parameters ϕ_a and ϕ_b by a new variable $\Delta\phi$ as

$$\Delta\phi = \phi_b - \phi_a$$

and substituting ϕ_b with this form yields

$$k_{a1} \sin(\phi_a + \psi_{a1}) = k_{b1} \sin(\phi_a + \Delta + \psi_{b1}) \quad (52)$$

$$k_{a2} \sin(\phi_a + \psi_{a2}) = k_{b2} \sin(\phi_a + \Delta + \psi_{b2}). \quad (53)$$

Collecting the scaling coefficients on the left-hand side,

$$k_1 \sin(\phi_a) = \sin(\phi_a + \Delta) \quad (54)$$

$$k_2 \sin(\phi_a + \psi_a) = \sin(\phi_a + \Delta + \psi_b), \quad (55)$$

applying the Pythagorean trigonometric identity to eq. 54, and the angle addition identity to eq. 55 produces the new system:

$$\sqrt{1 - k_1^2 \sin^2(\phi_a)} = \cos(\phi_a + \Delta) \quad (56)$$

$$k_2 \sin(\phi_a + \psi_a) = \sin(\phi_a + \Delta) \cos(\psi_b) + \cos(\phi_a + \Delta) \sin(\psi_b) \quad (57)$$

Substituting in eq. 54 and eq. 56 into eq. 57 and rearranging terms yields

$$a \sin(\phi_a) + b \cos(\phi_a) = \sqrt{1 - c \sin^2(\phi_a)} \quad (58)$$

$$a = k_2 \cos(\psi_a) - k_1 \cos(\psi_b), b = k_2 \sin(\psi_a), c = k_1^2,$$

which is able to be written as a form quadratic in $x = \cot(\phi_a)$:

$$(b^2 - 1)x^2 + 2abx + (a^2 + c - 1) = 0.$$

Finally, ϕ_a is recovered as $\arctan x^{-1}$ and ϕ_b is recovered by back-substitution.

5.2. SfM translation recovery

The absolute translation is recovered in sec. 4.2 using the recovered attitude ϕ in conjunction with known generatrix values (r, h) corresponding to the apparent contour point. With structure from motion, the generatrix values are unknown, however the scale and generatrix height offset are free parameters that can be arbitrarily chosen. For a two-view correspondence pair, we can exploit this ambiguity in scale and assume they have a shared radius $r = 1$. Similarly, we can exploit the arbitrary system height offset and assume the generatrix height as $h = 0$. Using translation recovery eq. 27 parameterized in this way with ϕ_a and ϕ_b yields translations (\hat{d}_a, \hat{h}_a) and (\hat{d}_b, \hat{h}_b) . Since the height offset is arbitrary, it is reasonable to compute $\Delta h = \hat{d}_b - \hat{d}_a$ and let $\hat{d}_a = 0$ and $\hat{d}_b = \Delta h$.

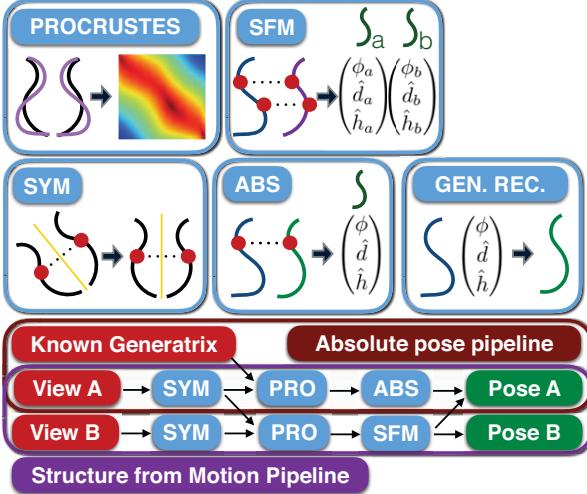


Figure 4. Algorithm overview: The two pipelines show the flow of the red inputs through their respective processes to yield the green outputs. The five blue boxes represent the processes used by the pipelines, with Procrustes matching two contours to produce a sampling probability map, and SYM, ABS and SFM representing the inputs and outputs for the papers three minimal problems. The generatrix reconstruction inputs and outputs are represented in GEN. REC.

5.3. SfM pose refinement

In sec. 4.3, the absolute pose refinement technique minimizes the error between the recovered contour generator and the contour generator synthesized from the known generatrix. In the case of SfM, the generatrix is not known a priori, however the pose hypothesis produces a reconstructed generatrix for each view. The SfM pose refinement procedure refines pose a using the generatrix from view b , and then repeats this procedure with a and b swapped until the stop criteria has been reached.

6. Algorithm overview

Previous sections have presented the minimal problems that allow hypothesis generation for the tasks of symmetry rotation recovery, absolute pose recovery, and structure from motion in addition to estimation refinement. This section describes the various pipelines for achieving pose recovery and structure from motion starting from pairs of segmented SoR apparent contours, and resulting in pose and generatrix recovery. These pipelines are summarized in fig. 4.

6.1. Hypothesis generation and scoring

A standard approach to utilizing minimal problems for the purpose of parameter estimation is to use a RANSAC-based [8] procedure for generating and scoring hypotheses. Such a procedure randomly samples points to create

a hypothesis from the corresponding minimal problem. It scores and accepts the hypothesis as a valid candidate if it has sufficient inliers. All three estimation tasks in this paper involve some form of curve matching and are solvable by correspondence-based minimal problems. If the sampling process could be biased towards points that are more likely to correspond and yield an acceptable hypothesis, then fewer evaluations would be required to terminate successfully.

6.1.1 Procrustes sampling probability

Using a coarse Procrustes alignment between two curves, the inter-curve point distances can be used as the sampling probability for point correspondences, with closer points sampled with higher frequency. In this way, the expected number of required hypothesis evaluations can be minimized, without being too reliant on the quality of the alignment. If the alignment fails, the system gracefully degrades towards a uniform sampling. An example Procrustes matching and score map is presented in fig. 4.

6.1.2 Hypothesis scoring

For all three problems, the mean geometric error between curve correspondences is used as the score. Correspondences are computed between the observed contours and the expected contours via the hypothesized parameters. The RANSAC procedures used a minimum acceptable model count stopping criteria, terminating when 50 models with sufficient inliers were found.

6.2. Absolute pose recovery

The absolute pose pipeline takes an SoR apparent contour pair from a single view, expressed in natural camera coordinates, as well as the generatrix for pose matching. After symmetric rotation is estimated and the contours are transformed to the symmetric view (see fig. 2b). The Procrustes correspondence probabilities are computed between the apparent contours and the generatrix, giving the sampling probability between each apparent contour point and generatrix height. The absolute pose RANSAC procedure is performed, with each pose hypothesis generated from a single contour point/generatrix height pair using the minimal problem defined in sec. 4.1 and refined as discussed in sec. 4.3.

6.3. Structure from motion

Structure from motion is achieved using just two SoR apparent contour pairs from two views, a and b , both expressed in natural camera coordinates. The symmetry rotation for each view is estimated, and the contours are transformed into the symmetric view (see fig. 2b). The Procrustes correspondence probabilities are computed between

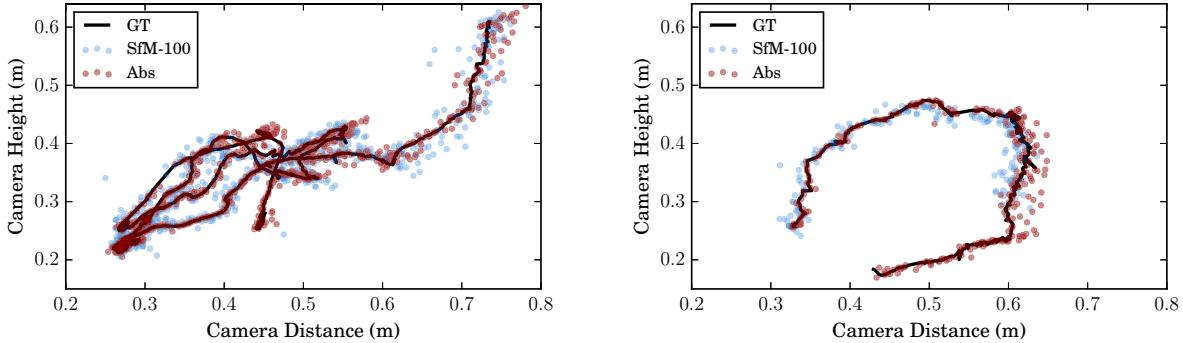


Figure 5. The estimated SfM and absolute pose trajectories for two image sequences are plotted in blue and red over the black ground truth trajectory.

the contours of view *a* and view *b*, giving the sampling probability between each apparent contour point across the two views. The SfM pose RANSAC procedure is performed, with each pose hypothesis generated by two pairs of *a/b* contour correspondences using the minimal problem defined in sec. 5 and refined as discussed in sec. 5.3.

7. Evaluation

Setup	Pitch ϕ	Distance \hat{d}	Height \hat{h}
Abs	0.04°	5.7mm (1.2%)	3.11mm (0.9%)
SfM-100	0.71°	7.8mm (1.8%)	5.31mm (1.6%)
SfM-75	0.69°	7.9mm (1.8%)	5.67mm (1.7%)
SfM-50	0.74°	8.6mm (1.9%)	5.51mm (1.7%)
SfM-25	0.74°	8.3mm (1.9%)	5.81mm (1.7%)
SfM-5	0.73°	8.5mm (1.9%)	6.08mm (1.8%)
SfM-2	0.74°	8.8mm (2.0%)	5.94mm (1.8%)
SfM-1	0.73°	9.0mm (2.0%)	6.44mm (1.9%)

The above table shows the average pose errors in terms of attitude ϕ and camera translation (\hat{d}, \hat{h}) for the absolute pose estimation algorithm (Abs) as well as SfM run with various baselines (measured in terms of video frames) over two different video motion sequences of approximately 1500 total frames in length. The scale and height offset ambiguity of the SfM output was resolved by registering the reconstructed generatrices to the ground truth generatrix to allow for metric performance comparison. The accuracy of the absolute and SfM pose results are shown in fig. 5, where they mirror the ground truth trajectories with few outliers.

All setups had an average angle error of less than 1°, and translation error of 2% or less of the relative distance. Absolute pose from a known generatrix results in the lowest error, which is reasonable as there fewer unknown parameters to estimate and therefore less uncertainty. Examining the effect of increasing baseline, the results show the expected trend of increasing error as the baseline distance decreases. Notice however, that the average inter-frame displacement was large enough that the algorithm does not fail with suc-

cessive frames.

Both the absolute pose recovery and SfM rely on the symmetry rotation recovery procedure, which had an average axis angular reprojection error of 0.64°, and an average pixel reprojection error of 1.1px. Comparing the recovered generatrices to the groundtruth generatrix, ther was a radius reconstruction error of 5.2mm.

8. Conclusion

This paper has presented three minimal problems that enable absolute pose recovery and structure from motion from the calibrated apparent contours of a surface of revolution, without the use of any other special geometric structures such as cross sections or bi-tangent points. An algorithm for applying these minimal problems to real-world data has been developed, and the performance has been evaluated over two image sequences to show a high level of accuracy with less than 1° orientation error and less than 2% translation error. Additionally, the average reconstruction error from structure from motion was shown to around 5mm.

8.1. Caveats and degeneracies

In both the single and two-view case, the estimation of the camera attitude is dependent upon the accuracy of the tangent lines selected for the minimal problems. In ideal data, a correct correspondence precisely defines the attitudes. Attitude errors generally have greater effect on SoR ground plane height estimation than distance. If the two tangent lines are collinear (as in a cylinder or cone), the attitude system of equation has infinite solutions.

References

- [1] R. Cipolla. *Active Visual Inference of Surface Shape*. PhD thesis, Oxford, UK, UK, 1991. UMI Order No. GAXD-96213. 2, 3

- [2] R. Cipolla and A. Blake. Surface shape from the deformation of apparent contours. *Int. J. Comput. Vision*, 9(2):83–112, Nov. 1992. 2, 3
- [3] C. Colombo, D. Com, and A. D. Bimbo. Camera calibration with two arbitrary coaxial circles. pages 265–276. Springer, 2006. 2
- [4] C. Colombo, D. Comanducci, A. Del Bimbo, and F. Pernici. Accurate automatic localization of surfaces of revolution for self-calibration and metric reconstruction. pages 55–55, June 2004. 1, 2
- [5] C. Colombo, A. Del Bimbo, and F. Pernici. Shape reconstruction from a single photograph for 3D object retrieval and visualization. volume 1, pages 897–9001, 2002. 2
- [6] C. Colombo, A. Del Bimbo, and F. Pernici. Metric 3D reconstruction and texture acquisition of surfaces of revolution from a single uncalibrated view. 27(1):99–114, Jan. 2005. 2
- [7] M. Dhome, J.-T. Lapresté, G. Rives, and M. Richetin. Spatial localization of modelled objects of revolution in monocular perspective vision. In *Proceedings of the First European Conference on Computer Vision, ECCV '90*, pages 475–485, London, UK, UK, 1990. Springer-Verlag. 2, 3
- [8] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6):381–395, June 1981. 7
- [9] D. A. Forsyth, J. L. Mundy, A. Zisserman, and C. A. Rothwell. Recognising rotationally symmetric surfaces from their outlines. pages 639–647, Berlin, Heidelberg, 1992. Springer Berlin Heidelberg. 2
- [10] P. Giblin and R. Weiss. Reconstruction of surfaces from profiles. Technical report, Amherst, MA, USA, 1987. 2, 3
- [11] R. Glachet, M. Dhome, and J. T. Lapresté. Finding the perspective projection of an axis of revolution. *Pattern Recogn. Lett.*, 12(11):693–700, Nov. 1991. 2
- [12] R. Glachet, M. Dhome, and J. T. Lapreste. *Computer Vision — ECCV'92: Second European Conference on Computer Vision Santa Margherita Ligure, Italy, May 19–22, 1992 Proceedings*, chapter Finding the pose of an object of revolution, pages 681–686. Springer Berlin Heidelberg, Berlin, Heidelberg, 1992. 3
- [13] A. Gross and T. Boult. Recovery of shgcs from a single intensity view. 18(2):161–180, Feb. 1996. 2
- [14] P. R. S. Mendonça, K. Y. K. Wong, and R. Cipolla. Epipolar geometry from profiles under circular motion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6):604–616, June 2001. 2
- [15] N. Navab and M. Appel. Canonical representation and multi-view geometry of cylinders. *International Journal of Computer Vision*, 70(2):133–149, 2006. 2
- [16] C. Phillips, M. Lecce, and K. Daniilidis. Seeing glassware: from edge detection to pose estimation and shape recovery. 2016. 1, 2
- [17] C. Phillips, M. Lecce, C. Davis, and K. Daniilidis. Grasping surfaces of revolution: Simultaneous pose and shape recovery from two views. pages 1352–1359, May 2015. 2
- [18] J. Ponce, D. Chelberg, and W. Mann. Invariant properties of straight homogeneous generalized cylinders and their contours. 11(9):951–966, Sept. 1989. 2
- [19] H. Sato and T. Binford. On finding the ends of straight homogeneous generalized cylinders. pages 695–698, June 1992. 1
- [20] H. Sato and T. O. Binford. Finding and recovering shgc objects in an edge image. 57(3):346–358, May 1993. 1, 2
- [21] K.-Y. K. Wong, P. R. S. Mendonça, and R. Cipolla. Reconstruction of surfaces of revolution from single uncalibrated views. In *Proc. British Machine Vision Conference 2002*, pages 93–102, 2002. 2, 4
- [22] A. Zisserman, J. Mundy, D. Forsyth, J. Liu, N. Pillow, C. Rothwell, and S. Utcke. Class-based grouping in perspective images. In *Computer Vision, 1995. Proceedings., Fifth International Conference on*, pages 183–188, June 1995. 1, 2