CANONICAL REPRESENTATION AND THREE VIEW GEOMETRY OF CYLINDERS

Nassir Navab

Siemens Corporate Research, Imaging and Visualization Department
755 College Road East, Princeton, NJ 08540, USA, navab@scr.siemens.com

ABSTRACT

This paper first introduces a canonical representation for cylinders. The canonical representation introduced here is closely related to the Plücker line respresentation. It then drives the general equation of a cylinder using this representation. In this paper, we also show that this representation is an appropriate one for computer vision applications. In particular, it allows us to easily develop a series of mathematical methods for pose estimation, 3D reconstruction and motion estimation. Finally, one of the main objectives of this paper is to introduce the main equations dominating the three view geometry of cylinders. We show the relationship between cylinders' three-view geometry and that of lines (Spetsakis and Aloimonos, 1990, Weng et al., 1993) and points (Shashua, 1995) defined by the trilinear tensor (Hartley, 1997), and propose a linear method, which uses the correspondences between six cylinders over three views in order to recover the motion and structure.

1 INTRODUCTION

Cylinders are the main components in a large number of industrial sites. Oil platforms, refineries and off-shore installations are almost made of pipes and pipelines. Chemical factories, water treatment plants and food processing factories are also mainly made of cylindrical objects. This is also the case for power generation plants. Figure 1 shows a typical image of an off-shore installation.

Figure 1: Typical image of an off-shore installation. Cylinders are the most common shapes found in such environment.

Majority of these industrial sites does not have a 3D model of the facility. Those who have a 3D model often need to update the model on a regular basis. There is an increasing need for creation of the 3D models based on images. This is the best way for getting an accurate update of 3D model of the factory. This is called as-built reconstruction. Figure 2 shows an as-built reconstruction of a power plant. This models has been built based on thousands of images. Different companies and software products aim at providing as-built reconstruction software or services (AsBuilt,FotoG,Geodelta,Leica) and CyliCon (Navab et al., 1999). The majority of the existing photogrammetric solutions proceed to an off-line image calibration procedure. Often markers are attached on the pipes and other factory installations. Then a series of pictures are taken. The markers are automatically detected, the cameras are calibrated through a bundle adujstment procedure, and the three-dimensional model of the environment including the pipes and pipelines are built. In spite of these
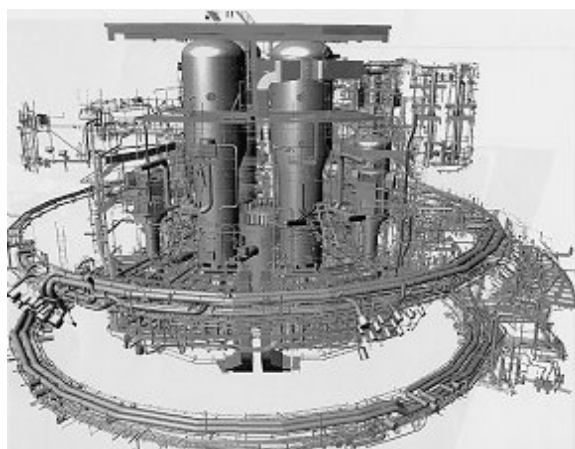

Figure 2: An as-built reconstruction of a power plant based on thousands of images.

activities and the increasing need for calibration and 3D reconstruction using cylindrical objects, there is neither a standard representation for cylinders, nor a standard formulation for motion estimation or 3D reconstruction from cylinders. Some previous work (Chandler and Still, 1994) used points correspondences and tradional photogrammetry methods to model such installations. Sayd et al. (Sayd et al.,1996), Jones et al. (Jones, 1996), Bauer (Bauer, 1997), and Benning and Schwermann (Benning, 1997) used occluding edges of cylinders on two or more calibrated views in order to reconstruct the 3D model of cylinder. Hanek et al. (Hanek et al., 1999) also used the occluding edge of cylinders for pose estimation. Veldhuis and Vosselman (Veldhuis and Vosselman, 1998) and Tangelder et al. (tangelder et al., 2000) propose methods to interactively fit
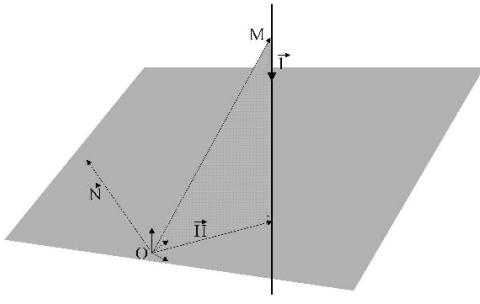
Figure 3: Plücker representation for lines.

the 3D model of cylinders onto their images for pose estimation. However, there was no attempt for defining an appropriate respresentation of cylinder as a geometrical entity. Motion estimation from generalized cylinders has been studied in (Sayd et al.,1996, Sayd et al., 1997). However, in the same way that, in computer vision, lines are not simply considered as particular case of general curves and require their own representation and mathematical framework, straight cylinders need to be studied as a geomertical entity and not as a particular case of generalized cylinders.

In this paper we introduce a canonical representation for cylinders. This is closely related to the line Plücker representation. This representation is suitable for computer vision and photogrammetry applications. We use this representation to define the general equation of a cylinder. We then use this respresentation to describe the motion of cylinders and their viewing geometry. We show that a pair of non parallel cylinders observed by two cameras provide enough constraint for recovery of the motion and structure trough a non-linear minimization. Finally, we introduce the main equations dominating the three view geometry of cylinders. We show the relationship between cylinders' three-view geometry and that of lines and points defined by the trilinear tensor (Hartley, 1997). We also develop a linear method, which uses the correspondences between six cylinders over three views in order to recover the motion and structure.

2 CANONICAL REPRESENTATION OF CYLINDERS

In this section we first review the Plücker representation for lines. Then we introduce a new cylinder representation and its relationship to the Plücker coordinates of lines.

2.1 Plücker representation of 3D lines

A 3D line is represented by two vectors $\mathbf{l}$ and $\mathbf{N}$. $\mathbf{l}$ is the unit vector representing the orientation of the line. The vector $\mathbf{N}$ is orthogonal to the plane defined by the 3D line and the origin of the coordinate system. The lenght of $\mathbf{N}$, $h = \|\mathbf{N}\|$, is the distance from origin to the 3D line. Figure 3 illustrates these entities.

A point $\mathbf{M}$ belongs to the 3D line if and only if its Plücker coordinates $(\mathbf{l}, \mathbf{N})$ satisfy the following equation:

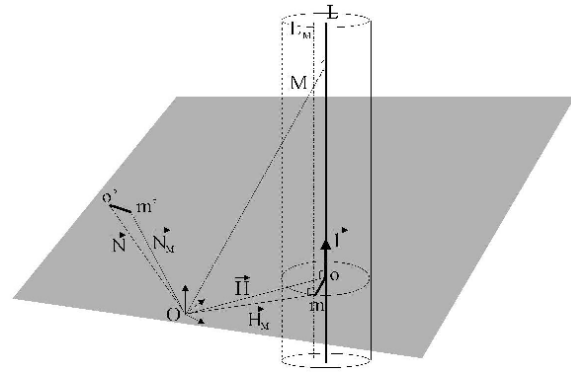$$\mathbf{M} \times \mathbf{l} = \mathbf{N} \qquad (1)$$

Figure 4: Distance of a point $\mathbf{M}$ from a line defined by its Plücker coordinates: illustration for the proof of the Theorem.

The nearest point of a line to the origin, $\mathbf{H} = \mathbf{N} \times \mathbf{l}$, is a particular one and we use it often in this paper. The distance of the line from origin is defined by $h = \|\mathbf{H}\|$.

2.2 Motion of lines

After a coordinate transformation defined by the rotation matrix $\mathbf{R}$ and the translation vector $\mathbf{T}$, a 3D line $(\mathbf{l}_1, \mathbf{N}_1)$ is represented by $(\mathbf{l}_0, \mathbf{N}_0)$ such that (Navab and Faugeras, 1997):

$$\begin{bmatrix} \mathbf{l}_0 \\ \mathbf{N}_0 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{0_{3x3}} \\ \mathbf{E} & \mathbf{R} \end{bmatrix} \begin{bmatrix} \mathbf{l}_1 \\ \mathbf{N}_1 \end{bmatrix} \qquad (2)$$

Where $\mathbf{E} = \tilde{\mathbf{T}}\mathbf{R}$ is known as the essential matrix.

2.3 Image of a line

The 2D image of the 3D line $(\mathbf{l}, \mathbf{N})$ is presented by $\mathbf{n} = \frac{\mathbf{N}}{\|\mathbf{N}\|}$ in the camera coordinate system. There are three more parameters to estimate in order to recover the 3D line, two for the orientation $\mathbf{l}$ of the line and one for the depth h of the line from the camera.

2.4 Canonical representation for Cylinders

In this section we first introduce a canonical representation for cylinders. We then define the equation of the cylinder using this representation.

We represent a cylinder by $(\mathbf{l}, \mathbf{N}, r)$, where $(\mathbf{l}, \mathbf{N})$ is the Plücker representation of the axis of the cylinder and r is the radius of the cylinder.

Theorem: A point belongs to the surface of a cylinder represented by its canonical representation $(\mathbf{l}, \mathbf{N}, r)$ if and only if:

$$\boxed{\|\mathbf{M} \times \mathbf{l} - \mathbf{N}\| = r} \qquad (3)$$

Proof: In order to prove this theorem, we show that $\|\mathbf{M} \times \mathbf{l} - \mathbf{N}\|$ defines the distance of a point $\mathbf{M}$ from a 3D line $\mathbf{L}$ defined by its Plucker coordinate $(\mathbf{l}, \mathbf{N})$. Figure 4 illustrates the geometry. The cross product $\mathbf{M} \times \mathbf{l}$ results in vector $\mathbf{N}_M$, a vector orthogonal to the plane defined by the origin and a line $\mathbf{L}_M$ passing through $\mathbf{M}$ and parallel to the axis of the cylinder. The norm of this vector is nothing but the distance of the line $\mathbf{L}_M$ from the origin. The vectors $\mathbf{H}$ and $\mathbf{H}_M$ are respectively orthogonal to $\mathbf{N}$ and $\mathbf{N}_M$. Therefore the angle between the two pairs of vectors are equal. The two triangles Omo and $Om'o'$ are equal. This means that the third sides of these triangles, om and $o'm'$ are also equal. This means that $\|\mathbf{o}'\mathbf{m}'\| = \|\mathbf{M} \times \mathbf{l} - \mathbf{N}\|$ equals $\|\mathbf{om}\|$, which in turn equals to the distance of $\mathbf{M}$ from the 3D line $\mathbf{L}$. The proof is now complete since by definition "a point $\mathbf{M}$ belongs to the cylinder $(\mathbf{l}, \mathbf{N}, r)$, if and only if its distance to the axis of the cylinder is equal to r"$\square$.

A line can be considered as a cylinder with canonical representation $(\mathbf{l}, \mathbf{N}, r = 0)$. A point $\mathbf{M}$ belongs to this line, or this "very tiny" cylinder, if and only if:

$$\|\mathbf{M} \times \mathbf{l} - \mathbf{N}\| = 0 \tag{4}$$

Therefore, we have $\mathbf{M} \times \mathbf{l} = \mathbf{N}$, which defines the line equation in Plücker coordinates.

2.5 Motion of Cylinders

The motion of lines can easily be extended to the case of cylinders. A cylinder can be represented in a new coordinate system or after a rigid transformation $(\mathbf{R}, \mathbf{T})$ by:

$$\begin{bmatrix} \mathbf{l}_0 \\ \mathbf{N}_0 \\ r_0 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{0}_{3x3} & \mathbf{0}_{3x1} \\ \mathbf{E} & \mathbf{R} & \mathbf{0}_{3x1} \\ \mathbf{0}_{1x3} & \mathbf{0}_{1x3} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{l}_1 \\ \mathbf{N}_1 \\ r_1 \end{bmatrix} \tag{5}$$

Where $\mathbf{E} = \tilde{\mathbf{T}}\mathbf{R}$ is the essential matrix.

2.6 Image of a Cylinder

Image of a cylinder C in general consists of two occluding edges $\mathbf{n}_1$ and $\mathbf{n}_2$. These edges of the cylinder on the image are the intersection of the two planes, which are tangent to the cylinder and pass through the camera optical center. This allows us to recover a subset of parameters of the 3D cylinder. We consider these observable parameters as image of the cylinder. Image of a cylinder $(\mathbf{l}, \mathbf{N}, r)$ is therefore represented in the camera coordinate system as:

- The unit vector $\mathbf{l}$, orientation of the cylinder. This is defined by the intersection of the occluding edges, i.e., their vanishing point:

$$\mathbf{l} = \frac{\mathbf{n}_1 \times \mathbf{n}_2}{\|\mathbf{n}_1 \times \mathbf{n}_2\|} \tag{6}$$

Figure 5: This typical image has been calibrated and used for an actual as-built reconstruction. There are many pipes with one or two markers on them. The 3D position of these markers have been estimated during the camera calibration/orientation process. Section 3.3 proposes a method to use this information for 3D reconstruction of the cylinders.

- The unit vector $\mathbf{n}$, image of the axis of the cylinder, i.e., $\mathbf{n} = \frac{\mathbf{N}}{\|\mathbf{N}\|}$. This is also a function of the two occluding edges:

$$\mathbf{n} = \frac{\mathbf{n}_1 - \mathbf{n}_2}{2} \tag{7}$$

We define $\mathbf{n}_1$ and $\mathbf{n}_2$ such that they both point away from the cylinder. This is a simple rule to fix the signs (or directions) of $\mathbf{n}_1$ and $\mathbf{n}_2$, which are otherwise arbitrary. Note that even if $\mathbf{n}_1$ and $\mathbf{n}_2$, representing the viewing planes of the two occluding edges, are symmetric with respect to $\mathbf{n}$, representing the viewing plane of the axis of cylinder, the occluding edges on the image are in general not symmetric in respect to the image of the axis of cylinder.

- The viewing angle of the cylinder α, which is defined as the angle between the two planes tangent to the cylinder and passing through the camera optical center:

$$\alpha = acos(-\mathbf{n}_1^t \mathbf{n}_2) \tag{8}$$

Note that we have the following relationship between the viewing angle α, the distance of the axis of the cylinder from the origin $h = \|\mathbf{N}\|$, and the radius of the cylinder:

$$r = h\sin(\frac{\alpha}{2}) \tag{9}$$

We therefore take the $\sin(\frac{\alpha}{2})$ instead of viewing angle itself as our observation.

In theory, from one perspective view a cylinder is defined up to one parameter, which can be either its depth or radius. Here, we take the only unknown parameter to be the depth of the axis of the cylinder from optical center $h = \|\mathbf{N}\|$.

We define the image of the cylinder as $(\mathbf{l}, \mathbf{n}, \sin(\frac{\alpha}{2}))$. We have:

$$\begin{bmatrix} \mathbf{l} \\ \mathbf{N} \\ r \end{bmatrix} = \begin{bmatrix} \mathbf{l} \\ h\mathbf{n} \\ h\sin(\frac{\alpha}{2}) \end{bmatrix} \tag{10}$$

3 3D RECONSTRUCTION OF CYLINDERS

In this section we use the canonical representation to propose different methods for the 3D reconstruction of cylinders from two, three or more calibrated views. Here we suppose that both intrinsic and extrinsic parameters of the cameras are known. We also discuss some interesting scenarios, where only one calibrated view of the cylinder and some additional information such as the 3D coordinates of a point on the cylinder are known. This is often the case when markers are attached to cylinders for camera calibration, see Fig. 5. We propose a method to bring this additional information into the 3D reconstruction framework.

3.1 Stereo Reconstruction

The vanishing point is quite noise sensitive. The estimation of the orientation of the cylinder l_0 from one view is often not reliable. In particular when the diameter of the cylinder is small compared to its distance from the camera. When calibrated stereo views are available, it is more reliable to take:

$$l_0 = n_0 \times Rn_1 \tag{11}$$

Where n_0 and n_1 define the viewing plane of the axis of the cylinder, i.e the plane which contains the axis of the cylinder and the optical center, respectively in the first and second camera. The viewing plane of the axis of cylinder in the second camera can also be defined by its orientation Rn_1 and its distance from origin $n_1^t T$ in the first camera coordinate system. The depth of the cylinder from the first camera h_0 is then obtained as:

$$h_0 = \frac{n_1^t T}{(l_0 \times n_0)^t Rn_1} \tag{12}$$

This defines the 3D cylinder, i.e, its canonical coordinates as $(l_0, N_0 = h_0 n_0, r = h_0 \sin(\frac{\alpha}{2}))$.

3.2 3D Reconstruction from Three or more Views

Here we describe the 3D reconstruction of cylinders from three calibrated views. This framework easily extends to more than three views. In the rest of this paper, we suppose that orientation and position of the third camera relative to the first one is always defined by rotation S and translation U. We first recover the orientation of the cylinder using all observations:

$$\begin{bmatrix} n_0 \\ Rn_1 \\ Sn_2 \end{bmatrix} l_0 = 0 \tag{13}$$

This linear system of equations is of the form $AX = 0$ and can be easily solved for $X(= l_0)$. The solution is the eigenvector associated to the smallest eigenvalue of $A^t A$. This can also be solved using Singular Value Decomposition (SVD).

The depth of the cylinder h_0 is then estimated as the least squares solution to the following system of equations:

$$\begin{bmatrix} (l_0 \times n_0)^t Rn_1 \\ (l_0 \times n_0)^t Sn_2 \end{bmatrix} h_0 = \begin{bmatrix} n_1^t T \\ n_2^t U \end{bmatrix} \tag{14}$$

3.2.1 Using a single occluding edge of a cylinder
If only one occluding edge of the cylinder is visible in an image, we can still use it for 3D reconstruction from multiple views. In fact, in theory we only need to view three occluding edges in order to reconstruct the cylinder. Let us write the equations for the case, where we see two occluding edges in one image and only one in the second one. In this case we have the image of the axis of the cylinder n_0 in the first view, but only one of the occluding edges of the cylinder, n_{11} in the second view. We still have:

$$l_0 = n_0 \times Rn_{11} \tag{15}$$

If we also observe a single occluding edge n_{21} in the third image, we will have:

$$\begin{bmatrix} n_0 \\ Rn_{11} \\ Sn_{21} \end{bmatrix} l_0 = 0 \tag{16}$$

Finally, if we see only one occluding edge in each of the three images, we have:

$$\begin{bmatrix} n_{01} \\ Rn_{11} \\ Sn_{21} \end{bmatrix} l_0 = 0 \tag{17}$$

In order to compute the depth of the cylinder and its radius, we use the constraints that the tangent plane defined by the occluding edge is at distance r to the axis of the cylinder. In the first case, where two edges are seen in the first view and one in the second one, we have:

$$h_0 = \frac{n_{11}^t T}{(l_0 \times n_0)^t Rn_{11} \pm \sin(\frac{\alpha}{2})} \tag{18}$$

This defines two cylinders with canonical coordinates $(l_0, N_0 = h_0 n_0, r = h_0 \sin(\frac{\alpha}{2}))$, one for each value of h_0 in Eq. 18. This ambiguity could only be removed if the user has additional information about the position of the cylinder relative to the tangent plane. This information is available on the image and could be used to obtain a unique solution. If there are multiple images with single occluding edges, there will be no more ambiguity. This means that only one of the two solutions for h_0 would satisfy the constraints in all views.

3.3 Reconstruction from One View

In the case we have only one view of the cylinder, as we described in the previous sections, the cylinder is defined up to one parameter, its depth from the camera or its radius. If there are feature points on the cylinder that are reconstructed or can be reconstructed from multiple views, see Fig. 5, the canonical representation and the associated equation of the cylinder allow an elegant and simple solution for cylinder reconstruction. For each point M on the cylinder, we have:

$$\|M \times l - N\| = r \tag{19}$$

From the single view observation of the cylinder, $(\mathbf{l}_0, \mathbf{n}_0, \sin(\frac{\alpha}{2}))$ we can deduce:

$$\|\mathbf{M} \times \mathbf{l}_0 - h_0 \mathbf{n}_0\| = h_0 \sin(\frac{\alpha}{2}) \qquad (20)$$

This results in a second degree equation in h_0:

$$(1-\sin^2(\frac{\alpha}{2})) \, h_0^2 - 2 \, \mathbf{n}_0^t(\mathbf{M}\times\mathbf{l}_0) \, h_0 + \|\mathbf{M}\times\mathbf{l}_0\|^2 = 0 \quad (21)$$

Solving this equation results in two possible values for h_0. The larger one is the correct depth. This is because the point is visible in the image. The smaller solution would correspond to a point on the other side of a transparent cylinder.

If two 3D points are measured on the cylinder, we can compute the depth of cylinder without using the viewing angle α. This could result in a more accurate estimate since the viewing angle is also sensitive to noise. If we have two points $\mathbf{M}_1$ and $\mathbf{M}_2$, h_0 will be defined uniquely and explicitly as:

$$h_0 = \frac{\|\mathbf{M}_2 \times \mathbf{l}_0\|^2 - \|\mathbf{M}_1 \times \mathbf{l}_0\|^2}{2 \, (\mathbf{n}_0^t(\mathbf{M}_1 \times \mathbf{l}_0) - \mathbf{n}_0^t(\mathbf{M}_2 \times \mathbf{l}_0))} \qquad (22)$$

4   CYLINDER BASED POSE ESTIMATION

In the case we have a single view of a set of known cylinders, the canonical representation makes it easy to formulate the pose estimation problem. Let us suppose that we observe a few cylinders in the scene. We represent them in the world coordinate system by $(\mathbf{l}_i, \mathbf{N}_i, r_i)$. Their images could be respectively represented by: $(\mathbf{l}_i^c, \mathbf{n}_i, \sin(\frac{\alpha_i}{2}))$. We have:

$$\begin{bmatrix} \mathbf{R} & \mathbf{0}_{3x3} & \mathbf{0}_{3x1} \\ \mathbf{E} & \mathbf{R} & \mathbf{0}_{3x1} \\ \mathbf{0}_{1x3} & \mathbf{0}_{1x3} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{l}_i \\ \mathbf{N}_i \\ r_i \end{bmatrix} = \begin{bmatrix} \mathbf{l}_i^c \\ h_i \mathbf{n}_i \\ h_i \sin(\frac{\alpha_i}{2}) \end{bmatrix} \tag{23}$$

The unknowns are the motion parameters $(\mathbf{R}, \mathbf{T})$, and the depth of different cylinders h_i. We could solve this in two consecutive steps. First we could recover the camera orientation by solving:

$$\mathbf{R}\mathbf{l}_i = \mathbf{l}_i^c \qquad (24)$$

The optimum estimation of orientation is the unit quaternion, which is the eigenvector of the following 4×4 matrix $\mathbf{Q}$ associated to its largest eigenvalue (Horn, 1987, Besl and McKay, 1992):

$$\mathbf{Q} = \sum_i \begin{bmatrix} \mathbf{l}_i^t\mathbf{l}_i^c & \mathbf{l}_i \times \mathbf{l}_i^c \\ (\mathbf{l}_i \times \mathbf{l}_i^c)^t & \mathbf{l}_i\mathbf{l}_i^{ct} + \mathbf{l}_i^c\mathbf{l}_i^t \end{bmatrix} \qquad (25)$$

We need at least two independent, i.e, non parallel cylinders, 3D/2D cylinder correspondences in order to recover the orientation of the camera. The more correspondences we have, the more accurate the result will be.

The depth of the pipes can also be recovered directly: $h_i = \frac{r_i}{\sin(\frac{\alpha_i}{2})}$. The translation can then be estimated by solving:

$$\mathbf{T} \times \mathbf{R}\mathbf{l}_i + h_i\mathbf{R}\mathbf{n}_i = \mathbf{N}_i \qquad (26)$$

This provides two equations per cylinder correspondences. We therefore need two independent pairs of cylinder correspondences to recover all the pose parameters. Note that we can also solve for h_i and $\mathbf{T}_i$ simultaneously by solving one system of linear equations combining the last two constraints. This would provide more accurate pose estimates.

5   CYLINDER BASED STRUCTURE FROM MOTION

In this section we try to recover motion and structure of cylinders from two or more views. In the case two views are available, the equations are quite similar to the pose estimation case. The only difference is that in the first frame the pipe is defined up to one parameter, i.e, depth or radius. The main constraint is the following:

$$\begin{bmatrix} \mathbf{l}_0 \\ h_0\mathbf{n}_0 \\ h_0 \sin(\frac{\alpha_0}{2}) \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{0}_{3x3} & \mathbf{0}_{3x1} \\ \mathbf{E} & \mathbf{R} & \mathbf{0}_{3x1} \\ \mathbf{0}_{1x3} & \mathbf{0}_{1x3} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{l}_1 \\ h_1\mathbf{n}_1 \\ h_1 \sin(\frac{\alpha_1}{2}) \end{bmatrix}$$
$$(27)$$

The relative orientation can be computed in the same way as in the pose estimation case. Once again we need at least two independent cylinders viewed in two images. Replacing h_1 by $h_0 \frac{\sin(\frac{\alpha_0}{2})}{\sin(\frac{\alpha_1}{2})}$, we have:

$$\mathbf{T} \times \mathbf{R}\mathbf{l}_1 + h_0(\frac{\sin(\frac{\alpha_0}{2})}{\sin(\frac{\alpha_1}{2})}\mathbf{R}\mathbf{n}_1 - \mathbf{n}_0) = 0 \qquad (28)$$

This equation is a homogeneous one. It is known that in structure from motion shape or translational motion can be only recovered up to scale. Once again we need two independent cylinders in order to recover the translational motion up to scale.

In the same way if we have three views, we can write the following constraints:

$$\mathbf{R}\,\mathbf{l}_1 = \mathbf{l}_0 \qquad (29)$$
$$\mathbf{S}\,\mathbf{l}_2 = \mathbf{l}_0 \qquad (30)$$
$$\mathbf{T} \times \mathbf{R}\mathbf{l}_1 + h_0(\frac{\sin(\frac{\alpha_0}{2})}{\sin(\frac{\alpha_1}{2})}\mathbf{R}\mathbf{n}_1 - \mathbf{n}_0) = 0 \qquad (31)$$
$$\mathbf{U} \times \mathbf{S}\mathbf{l}_2 + h_0(\frac{\sin(\frac{\alpha_0}{2})}{\sin(\frac{\alpha_2}{2})}\mathbf{S}\mathbf{n}_2 - \mathbf{n}_0) = 0 \qquad (32)$$

The h_0 can be eliminated from the last two constraints. This leaves us with a set of equations defining the relationship between the three perspective views. One could estimate the motion parameters followed by the depth of cylinders using the correspondences between two independent cylinders.

6   CYLINDERS AND TRILINEAR TENSORS

In this section we present a new formulation for the projective geometry of cylinders in three views. This also
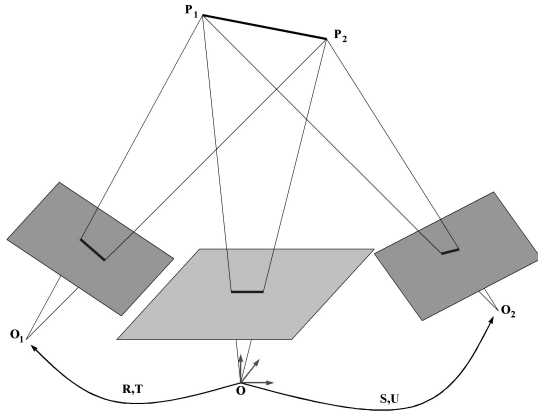
Figure 6: Three perspective cameras are observing a line segment. The first camera is aligned with the world coordinate system.

provides a three-view analysis of cylinders and its relationship to the three-view formulation of points (Shashua, 1995) and lines (Spetsakis and Aloimonos, 1990, Weng et al., 1993) in terms of trilinear tensors (Hartley, 1997).

Let us first review the linear algorithm (Spetsakis and Aloimonos, 1990, Weng et al., 1993) to recover the motion from three perspective views using line correspondences.

Once again let us consider that three internally calibrated perspective cameras are observing a scene composed of rigidly attached lines. Without loss of generality we set the world coordinate frame as the coordinate frame of the first camera (see Fig. 6). Furthermore, we assume that the motion of the first camera is given by the rotation R and the translation T and the motion of the second camera is given by S and U respectively.

We note that for each image, a plane (the projection plane) is formed by the 3D line and the center of projection. For example, the projection plane for the first camera and the line in Fig. 6 is formed by the two points P_1, P_2 on the line and the center of projection O_1. Let us denote the normals of the projection planes for the cameras with n_0, n_1 and n_2 for the first, the second and the third cameras respectively in their local coordinate frames. In the world coordinate system the normals become n_0, $R^T n_1$ and $S^T n_2$. The signed distance between the origin and the projection planes are 0, $T^T n_1$ and $U^T n_2$. We can represent these planes with four-dimensional homogeneous vectors as follows:

$$\Pi_0 = \begin{pmatrix} n_0^T \\ 0 \end{pmatrix}, \Pi_1 = \begin{pmatrix} R^T n_1 \\ T^T n_1 \end{pmatrix} \text{ and } \Pi_2 = \begin{pmatrix} S^T n_2 \\ U^T n_2 \end{pmatrix}.$$

These three planes meet at a common line. A necessary condition for this is that the 4×3 matrix formed by the vectors representing the planes has rank 2. After some algebraic manipulation, this results in the following linear system of equations:

$$n_0 \times \begin{pmatrix} n_1^T E n_2 \\ n_1^T F n_2 \\ n_1^T G n_2 \end{pmatrix} = 0, \text{ where } \begin{cases} E \stackrel{\text{def}}{=} R_1 U^T - T S_1^T \\ F \stackrel{\text{def}}{=} R_2 U^T - T S_2^T \\ G \stackrel{\text{def}}{=} R_3 U^T - T S_3^T \end{cases}$$

This result has earlier been obtained by Spetsakis and Aloimonos (Spetsakis and Aloimonos, 1990) and by Weng, Huang and Ahuja (Weng et al., 1993). Note that these three matrices, $\mathbf{E}$, $\mathbf{F}$ and $\mathbf{G}$, are a particular contraction of the more general trifocal tensor described by Hartley in (Hartley, 1997) for the projective case. When 13 or more line correspondences are available, we can recover the three matrices from which the actual motion parameters can be extracted using the method described in (Weng et al., 1993).

In order to study the three view-geometry of cylinders, we first describe the axis of a cylinder in terms of the intersection of two orthogonal planes, such that one of them also contains the origin of the coordinate system. These two planes are defined by their normals and their distances to the origin $(\mathbf{n}_i, 0)$ and $(\mathbf{n}_i \times \mathbf{l}_i, h_i)$. The three-view constraints forces the six planes defined by the triplet of cylinder correspondences to all intersect in one line. This means that the following 4×6 matrix is of rank two:

$$\begin{pmatrix} & n_1^T R_1 & S_1^T n_2 & & h_1^T R_1 & h_2^T S_1 \\ n_0 & n_1^T R_2 & S_2^T n_2 & h_0 & h_1^T R_2 & h_2^T S_2 \\ & n_1^T R_3 & S_3^T n_2 & & h_1^T R_3 & h_2^T S_3 \\ 0 & n_1^T T & U^T n_2 & h_0 & h_1 + h_1^T T & h_2 + h_2^T U \end{pmatrix}, \tag{33}$$

where $\mathbf{h}_i = \mathbf{n}_i \times \mathbf{l}_i$, $i = 0..2$. Eliminating the remaining shape parameters, i.e., h_0, h_1 and h_2, using the fact that the radius of the cylinder $r = h_0 \sin(\frac{\alpha_0}{2}) = h_1 \sin(\frac{\alpha_1}{2}) = h_1 \sin(\frac{\alpha_1}{2})$, we have five planes intersecting in one line.

$$\begin{pmatrix} n_0 & R^T n_1 & S^T n_2 & R^T h_1 - a_1 h_0 & S^T h_2 - a_2 h_0 \\ 0 & n_1^T T & U^T n_2 & h_1^T T & h_2^T U \end{pmatrix}, \tag{34}$$

where $a_1 = \frac{\sin(\frac{\alpha_0}{2})}{\sin(\frac{\alpha_1}{2})}$ and $a_2 = \frac{\sin(\frac{\alpha_0}{2})}{\sin(\frac{\alpha_2}{2})}$. The five plane, Π_0, Π_1, Π_2,

$$\Pi_3 = \begin{pmatrix} R^T h_1 - a_1 h_0 \\ h_1^T T \end{pmatrix}, \text{ and } \Pi_4 = \begin{pmatrix} S^T h_2 - a_2 h_0 \\ h_2^T U \end{pmatrix}$$

intersect in one line if and only if three triplets of planes have a unique intersection. For example,

$$\begin{cases} \text{rank} \begin{pmatrix} \Pi_0 & \Pi_1 & \Pi_2 \end{pmatrix} = 2 \\ \text{rank} \begin{pmatrix} \Pi_0 & \Pi_3 & \Pi_4 \end{pmatrix} = 2 \\ \text{rank} \begin{pmatrix} \Pi_0 & \Pi_1 & \Pi_4 \end{pmatrix} = 2 \end{cases}$$

Noting that $h_0 \times n_0 = l_0$, these rank constraints result in the following sets of equations:

$$n_0 \times \begin{pmatrix} n_1^T E n_2 \\ n_1^T F n_2 \\ n_1^T G n_2 \end{pmatrix} = 0$$

$$n_0 \times \begin{pmatrix} h_1^T E h_2 \\ h_1^T F h_2 \\ h_1^T G h_2 \end{pmatrix} + (a_1 h_2^T U - a_2 h_1^T T) l_0 = 0$$

$$n_0 \times \begin{pmatrix} n_1^T E h_2 \\ n_1^T F h_2 \\ n_1^T G h_2 \end{pmatrix} - (a_2 n_1^T T) l_0 = 0$$

These equations provide six independent homogenous equations in 33 variables. These variables include the trinilear tensor and the two translation vectors. Therefore, in general six cylinder correspondences over three frames provide us with a linear solution for structure from motion problem. Once the translational vectors T and U are estimated, The rotation matrices R and S can be easily recovered from E, F and G.

7 CONCLUSION

In this paper, we introduce a canonical representation of cylinders based on the Plücker line respresentation. The basic motion equations for points and lines have been well described in the computer vision literature. This paper completes this picture by describing the motion equations for cylinders. One of the originality of this paper is in considering the cylinder as an observable geometrical entity instead of only describing the equations provided by the geometrical constraints of its occluding edges. Based on this simple and elegant representation all the necessary structure and motion equations are derived.

This paper also provides a link between the trilinear tensor, defined for points and lines, and the three view geometry of cylinders. It finally proposes a linear algorithm, which allows the recovery of structure from motion based on a minimum of six cylinder correspondences over three views.

REFERENCES

http://www.asbuilt.com/,
http://www.geodelta.com/,
http://www.leica-geosystems.com/
http://www.vexcel.com/fotog/index.html.

Bauer, S. Semi-automatic reconstruction of multi-section cylindrical pipelines from sets of calibrated images. Diploma Thesis, Friedrich Alexander-Universität Erlangen-Nürnberg, 1997.

Benning, W. PHIDIAS-MS - Eine digitale Photogrammetrieapplikation unter Microstation für Nahbereichsanwendungen. In *Allgemeine Vermessungsnachrichten*, Vol. 1997 (1), pp. 16-25.

Besl, P.J. and McKay, N.D. A method for registration of 3-D shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1992, 14 (2), 239–256.

Chandler, J.H. and Still, A.L. Analytical photogrammetry and the refurbishment of the CAT Reformer, BP Refinery, Grangemouth, Scotland - Lessons learnt. In *International Archives of Photogrammetry and Remote Sensing*, Vol. 30, Part 5, pp. 28-34, Melbourne, 1994.

Hanek, R., Navab, N. and Appel, Mirko. Yet another Method for Pose Estimation: A Probabilistic Approach using Points, Lines, and Cylinders. In Proc. *Conference on Computer Vision and Pattern Recognition*, Fort Collins, Colorado, June 1999.

Hartley, R.I. Lines and points in three views and the trifocal tensor. Int. J. of Computer Vision, 22(2):125– 140, 1997.

Horn, B. Closed-form solution of absolute orientation using unit quaternions. Journal of the Optical Society of America, April 1987, A 7, 629–642.

Jones, M.A., Chapman, D.P. and Hamid, A.A. Close-range photogrammetry using geometric primitives for efficient CAD modelling of industrial plant. In *International Archives of Photogrammetry and Remote Sensing*, Vol. 31, part B5, pp. 284-289, Vienna, 1996.

Navab, N. and Faugeras, O.D. The critical sets of lines for camera displacement estimation: a mixed Euclidean-projective and constructive approach International Journal of Computer Vision; Vol.23, no.1; 1997; p.17-44.

Navab, N., Cubillo, E. , Bascle, B., Lockau, J., Kamsties, K. and Neuberger, M. CYLICON: A Software Platform for the creation & update of virtual factories. In Proc. *7th IEEE International Conference on Emerging Technologies and Factory Automation* (ETFA'99), Barcelona Spain, October 1999.

Sayd, P., Dhome, M. and Lavest, J.M. Axial Stereovision for recovering straight cylinders In Proc. of IEEE Conference on Robotics and Cybernetics CESA'96, pp 560-564, Lille, July 1996.

Sayd, P., Dhome, M., Lavest, J.M. and Lapreste, J.T. Non Uniform Circular Generalized Cylinder Reconstruction From Multiples Perspectives Views In Proc. of the 10th Scandinavian Conference on Image Analysis, pp 955-962, Lappeenranta Finland, June 1997.

Shashua, A. Algebraic functions for recognition. *PAMI*, 17(8):779–789, 1995.

Spetsakis, M.E. and Aloimonos, Y. Structure from motion using line correspondences. *IJCV*, 4(3):171–183, 1990.

Shiu, Y.C. and Zhuang, H. Pose determination of cylinders, cones, and spheres. In *Applications of Artificial Intelligence X: Machine Vision and Robotics*, Orlando, Florida, April 1992.

Tangelder, J.W.H, Vosselman, G. and van den Heuvel, F.A. Object Oriented measurement of Pipe Systems Using Edge Matching adn CSG Models with Constrains In Proceedings of the XIXth congress of ISPRS, Amsterdam 2000,"International Archives of Photogrammetry and Remote Sensing", vol XXXIII, part B 5/2 commission V, ISSN 0256-1840.

Veldhuis, H. and Vosselman, G. The 3D rescronstruction of straight and curved pipes using digital line photogrammetry. In *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 53 (1998), pp. 6-16.

Weng, J., Ahuja, N. and Huang, T.S. "Optimal motion and structure estimation," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 15, pp. 864–884, September 1993.