

## Cylindrical Rectification to Minimize Epipolar Distortion

Sébastien Roy\*†

Jean Meunier†

Ingemar J. Cox\*

NEC Research Institute\*  
4 Independence Way  
Princeton, NJ 08540, U.S.A.

Université de Montréal†  
Département d'informatique et de recherche opérationnelle  
C.P. 6128, Succ. Centre-Ville, Montréal, Québec, H3C 3J7

### Abstract

We propose a new rectification method for aligning epipolar lines of a pair of stereo images taken under any camera geometry. It effectively remaps both images onto the surface of a cylinder instead of a plane, which is used in common rectification methods. For a large set of camera motions, remapping to a plane has the drawback of creating rectified images that are potentially infinitely large and presents a loss of pixel information along epipolar lines. In contrast, cylindrical rectification guarantees that the rectified images are bounded for all possible camera motions and minimizes the loss of pixel information along epipolar line. The processes (eg. stereo matching, etc..) subsequently applied to the rectified images are thus more accurate and general since they can accommodate any camera geometry.

### 1 Introduction

Rectification is a necessary step of stereoscopic analysis. The process extracts epipolar lines and realigns them horizontally into a new *rectified* image. This allows subsequent stereoscopic analysis algorithms to easily take advantage of the *epipolar constraint* and reduce the search space to one dimension, along the horizontal rows of the rectified images.

For different camera motions, the set of matching epipolar lines varies considerably and extracting those lines for the purpose of depth estimation can be quite difficult. The difficulty does not reside in the equations themselves; for a given point, it is straightforward to locate the epipolar line containing that point. The problem is to find a set of epipolar lines that will cover the whole image and introduces a minimum of distortion, for arbitrary camera motions. Since subsequent stereo matching occurs along epipolar lines, it is important that no pixel information is lost along these lines in order to efficiently and accurately recover depth.

Fig. 1 depicts the rectification process. A scene  $S$  is observed by two cameras to create images  $I_1$  and  $I_2$ . In order to align the epipolar lines of this stereo pair, some image transformation must be applied. The most common of such transformations, proposed by Ayache [1] and referred to as *planar rectification*, is a remapping of the original images onto a single plane that is parallel to the line joining the two cameras optical centers (see Fig. 1, images  $P_1$  and  $P_2$ ). This is accomplished by using a linear transformation in projective space applied to each image pixels.

The new rectification method presented in this paper,

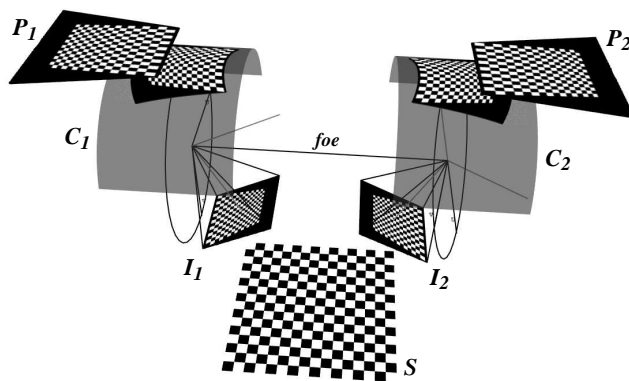


Figure 1: **Rectification. Stereo images ( $I_1, I_2$ ) of scene  $S$  shown with planar rectification ( $P_1, P_2$ ) and cylindrical rectification ( $C_1, C_2$ )**

referred to as *cylindrical rectification*, proposes a transformation that remaps the images onto the surface of a cylinder whose principal axis goes through both cameras optical centers (see Fig. 1, images  $C_1$  and  $C_2$ ). The actual images related for Fig. 1 are shown in Fig. 2.

The line joining the optical centers of the cameras (see Fig. 1) defines the focus of expansion (*foe*). All epipolar lines intersect the focus of expansion. The rectification process applied to an epipolar line always makes that line *parallel* to the *foe*. This allows the creation of a rectified image where the epipolar lines do not intersect and can be placed as separate rows. Obviously, both plane and cylinder remapping satisfy the alignment requirement with the *foe*.

Planar rectification, while being simple and efficient, suffers from a major drawback: it fails for some camera motion, as demonstrated in Sec. 2. As the forward motion component becomes more significant, the image distortion induced by the transformation becomes progressively worse until the image is effectively unbounded. The image distortion induces a loss of pixel information that can only be partly compensated for by making the rectified image size larger<sup>1</sup>. Consequently, this method is useful only for motions with a small forward component, thus lowering the risk of unbounded rectified images. One benefit of planar rectification is that it preserves straight lines, which is an important con-

<sup>1</sup>See Sec. 3.6 for a detailed discussion

sideration if stereo matching is to be performed on edges or lines.

On the other hand, cylindrical rectification is guaranteed to provide a bounded rectified image and significantly reduce pixel distortion, for all possible camera motions. This transformation also preserves epipolar line *length*. For example, an epipolar line 100 pixels long will always be rectified to a line 100 pixels long. This insures a minimal loss of pixel information when resampling the epipolar lines from the original images. However, arbitrary straight lines are no longer preserved, though this may only be a concern for edge based stereo.

Planar rectification uses a single linear transformation matrix applied to the image, making it quite efficient. Cylindrical rectification uses one such linear transformation matrix for *each* epipolar line. In many cases, these matrices can be precomputed so that a similar level of performance can be achieved.

Although it is assumed throughout this paper that internal camera parameters are known, cylindrical rectification works as well with unknown internal parameters, as it is the case when only the *Fundamental matrix* (described in [2]) is available (See Sec. 3.5).

Many variants of the planar rectification scheme have been proposed [1, 3, 4]. A detailed description based on the *essential matrix* is given in [5]. In [6], a hardware implementation is proposed. In [7], the camera motion is restricted to a *vergent stereo* geometry to simplify computations. It also presents a faster way to compute the transformation by approximating it with a non-projective linear transformation. This eliminates the risk of unbounded images at the expense of potentially severe distortion. In [8], a measure of image distortion is introduced to evaluate the performance of the rectification method. This strictly geometric measure, based on edge orientations, does not address the problem of pixel information loss induced by interpolation (see Sec. 3.6).

Sec. 2 describes planar rectification in more details. The cylindrical rectification method is then presented in Sec. 3. It describes the transformation matrix whose three components are explicitly detailed in Sec. 3.3, 3.2 and 3.1. Sec. 3.4 discuss the practical aspect of finding the set of corresponding epipolar lines in both images to rectify. It is demonstrated in Sec. 3.5 that it is possible to use uncalibrated as well as calibrated cameras. A measure of image distortion is introduced in Sec. 3.6 and used to show how both rectification methods behave for different camera geometries. Examples of rectification for different camera geometries are presented in Sec. 4.

## 2 Linear transformation in projective space

In this section we show how rectification methods based on a single linear transformation in projective space [1, 3, 4] fail for some camera geometries.

As stated earlier, the goal of rectification is to apply a transformation to an image in order to make the epipolar lines parallel to the focus of expansion. The result is a set of images where each row represents one epipolar line and can be used directly for the purpose of stereo matching (see

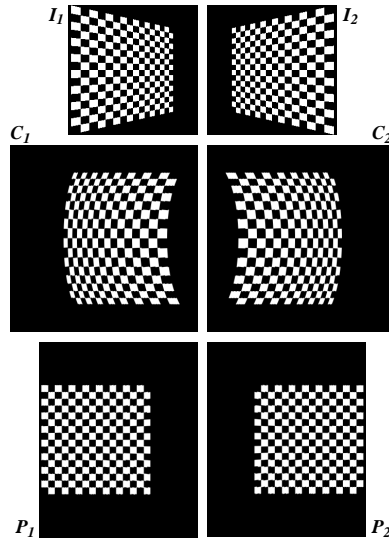


Figure 2: **Images from Fig. 1. Original images ( $I_1, I_2$ ) are shown with cylindrical rectification ( $C_1, C_2$ ) and planar rectification ( $P_1, P_2$ ).**

Fig. 2).

In projective space, an image point is expressed as  $\mathbf{p} = (p_x, p_y, h)^T$  where  $h$  is a scale factor. Thus we can assume these points are projected to  $\mathbf{p} = (p_x, p_y, 1)^T$ .

The linear projective transformation  $\mathbf{F}$  is used to transform an image point  $\mathbf{u}$  into a new point  $\mathbf{v}$  with the relation

$$\mathbf{v} = \mathbf{F} \cdot \mathbf{u} = \begin{bmatrix} F_0 & F_1 & F_2 \\ F_3 & F_4 & F_5 \\ F_6 & F_7 & F_8 \end{bmatrix} \cdot \mathbf{u} \quad (1)$$

where

$$\mathbf{v} = (v_x, v_y, v_h)^T \quad \mathbf{u} = (u_x, u_y, u_h)^T \quad u_h \neq 0$$

The fact that  $u_h \neq 0$  simply implies that the original image has a finite size. Enforcing that the reprojected point is *not* at infinity implies that  $v_h$  must be non-zero, that is

$$v_h = u_x F_6 + u_y F_7 + u_h F_8 \neq 0 \quad (2)$$

Since  $u_x, u_y$  are arbitrary, Eq. 2 has only one possible solution  $(F_6, F_7, F_8) = (0, 0, 1)$  since only  $u_h$  can guarantee  $v_h$  to be non zero and  $\mathbf{F}$  to be homogeneous. Therefore, the transformation  $\mathbf{F}$  must have the form

$$\mathbf{F} = \begin{bmatrix} F_0 & F_1 & F_2 \\ F_3 & F_4 & F_5 \\ 0 & 0 & 1 \end{bmatrix}$$

which corresponds to a camera displacement with *no* forward (or backward) component.

In practice, the rectified image is unbounded only when the *foe* is *inside* the image. Therefore, any camera motion with a large forward component (making the *foe* visible) *cannot* be rectified with this method. Moreover, as soon as the forward component is large enough, the image points are

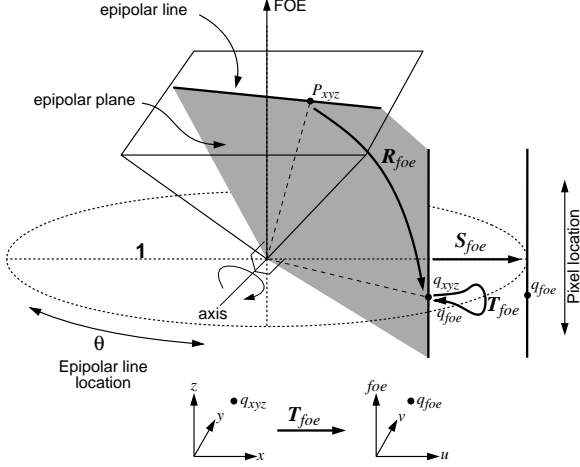


Figure 3: **The basic steps of the cylindrical rectification method. First ( $\mathbf{R}_{foe}$ ), an epipolar line is rotated in the epipolar plane until it is parallel to the foe. Second ( $\mathbf{T}_{foe}$ ), a change of coordinate system is applied. Third ( $\mathbf{S}_{foe}$ ), a projection onto the surface of the unit cylinder is applied.**

mapped so far apart that the rectification becomes unusable due to severe distortion.

In the next section, we described how *cylindrical rectification* can alleviate these problems by making a different use of linear transformations in projective space.

### 3 Cylindrical rectification

The goal of cylindrical rectification is to apply a transformation of an original image to remap on the surface of a carefully selected cylinder instead of a plane. By using the line joining the cameras optical centers as the cylinder axis (Fig. 1), all straight lines on the cylinder surface are necessarily parallel to the cylinder axis and focus of expansion, making them suitable to be used as epipolar lines.

The transformation from image to cylinder, illustrated in Fig. 3, is performed in three stages. First, a rotation is applied to a selected epipolar line (step  $\mathbf{R}_{foe}$ ). This rotation is in the epipolar plane and makes the epipolar line parallel to the foe. Then, a change of coordinate system is applied (step  $\mathbf{T}_{foe}$ ) to the rotated epipolar line from the image system to the cylinder system (with foe as principal axis). Finally, (step  $\mathbf{S}_{foe}$ ), this line is *normalized* or *reprojected* onto the surface of a cylinder of unit diameter. Since the line is already parallel to the cylinder, it is simply scaled along the direction perpendicular to the axis until it lies at unit distance from the axis. A particular epipolar line is referenced by its angle  $\theta$  around the cylinder axis, while a particular pixel on the epipolar line is referenced by its angle and position along the cylinder axis (see Fig. 3).

Even if the surface of the cylinder is infinite, it can be shown that the image on that surface is always bounded. Since the transformation aligns an epipolar line with the axis of the cylinder, it is possible to remap a pixel to infinity only if its epipolar line is originally infinite. Since the original image is finite, all the visible parts of the epipolar lines are

also of finite length and therefore the rectified image cannot extend to infinity.

The rectification process transforms an image point  $\mathbf{p}_{xyz}$  into a new point  $\mathbf{q}_{foe}$  which is expressed in the coordinate system foe of the cylinder. The transformation matrix  $\mathbf{L}_{foe}$  is defined so that the epipolar line containing  $\mathbf{p}_{xyz}$  will become parallel to the cylinder axis, the foe. Since all possible epipolar lines will be parallel to the foe, they will also be parallel to one another and thus form the desired *parallel aligned epipolar geometry*.

We have the linear rectification relations between  $\mathbf{q}_{foe}$  and  $\mathbf{p}_{xyz}$  stated as

$$\begin{aligned} \mathbf{q}_{foe} &= \mathbf{L}_{foe} \mathbf{p}_{xyz} \\ &= (\mathbf{S}_{foe} \mathbf{T}_{foe} \mathbf{R}_{foe}) \mathbf{p}_{xyz} \end{aligned} \quad (3)$$

and inversely

$$\begin{aligned} \mathbf{p}_{xyz} &= \mathbf{L}_{foe}^{-1} \mathbf{q}_{foe} \\ &= (\mathbf{R}_{foe}^T \mathbf{T}_{foe}^T \mathbf{S}_{foe}^{-1}) \mathbf{q}_{foe} \end{aligned} \quad (4)$$

where

$$\mathbf{S}_{foe} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{k} & 0 \\ 0 & 0 & \frac{1}{k} \end{bmatrix} \quad \text{and} \quad \mathbf{T}_{foe} = \begin{bmatrix} \text{foe} \\ \mathbf{u} \\ \mathbf{v} \end{bmatrix}$$

These relations are completely invertible (except for the special case  $\mathbf{p}_{xyz} = \text{foe}$ , which is quite easily handled). The matrix  $\mathbf{R}_{foe}$  represents the rotation of the image point in projective space. The matrix  $\mathbf{T}_{foe}$  represents the change from the camera coordinate system to the cylinder system. The matrix  $\mathbf{S}_{foe}$  represents the projective scaling used to project rectified point onto the surface of the unit cylinder.

The next three subsections will describe how to compute the coordinate transformation  $\mathbf{T}_{foe}$ , the rotation  $\mathbf{R}_{foe}$ , and the scaling  $\mathbf{S}_{foe}$ .

#### 3.1 Determining transformation $\mathbf{T}$

The matrix  $\mathbf{T}_{foe}$  is the coordinate transformation matrix from system  $(\mathbf{x}; \mathbf{y}; \mathbf{z})$  to system  $(\text{foe}; \mathbf{u}; \mathbf{v})$  such that

$$\begin{aligned} \mathbf{q}_{foe} &= \mathbf{T}_{foe} \mathbf{q}_{xyz} \\ \mathbf{q}_{xyz} &= \mathbf{T}_{foe}^T \mathbf{q}_{foe} \end{aligned} \quad (5)$$

and is uniquely determined by the position and motion of the cameras (see Fig. 3).

Any camera has a position  $\mathbf{pos}$  and a rotation of  $\phi$  degrees around the axis  $\mathbf{axis}$  relative to the world coordinate system. A homogeneous world point  $\mathbf{p}_w$  is expressed in the system of camera  $a$  (with  $\mathbf{pos}_a$ ,  $\mathbf{axis}_a$ , and  $\phi_a$ ) as

$$\mathbf{p}_a = \mathbf{R}_{aw} \mathbf{p}_w$$

where  $\mathbf{R}_{aw}$  is the  $4 \times 4$  homogeneous coordinate transformation matrix obtained as

$$\begin{aligned} \mathbf{R}_{aw} &= \begin{bmatrix} \text{rot}(\mathbf{axis}_a, -\phi_a) & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} I & -\mathbf{pos}_a \\ \mathbf{0} & 1 \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{r}_{aw} & -\mathbf{r}_{aw} \cdot \mathbf{pos}_a \\ \mathbf{0} & 1 \end{bmatrix} \end{aligned}$$

where

$$\mathbf{r}_{aw} = \text{rot}(\mathbf{axis}_a, -\phi_a)$$

and  $\text{rot}(\mathbf{A}, \theta)$  is a  $3 \times 3$  rotation matrix of angle  $\theta$  around axis  $\mathbf{A}$ . The corresponding matrix  $\mathbf{R}_{bw}$  for camera  $b$  with  $\mathbf{pos}_b$ ,  $\mathbf{axis}_b$ , and  $\phi_b$  is defined in a similar way.

The direct coordinate transformation matrices for camera  $a$  and  $b$  such that

$$\begin{aligned} \mathbf{p}_a &= \mathbf{R}_{ab} \mathbf{p}_b \\ \mathbf{p}_b &= \mathbf{R}_{ba} \mathbf{p}_a \end{aligned}$$

are defined as

$$\begin{aligned} \mathbf{R}_{ab} &= \mathbf{R}_{aw} \cdot \mathbf{R}_{bw}^{-1} = \begin{bmatrix} \mathbf{r}_{ab} & \mathbf{foe}_a \\ \mathbf{0} & 1 \end{bmatrix} \\ \mathbf{R}_{ba} &= \mathbf{R}_{bw} \cdot \mathbf{R}_{aw}^{-1} = \begin{bmatrix} \mathbf{r}_{ba} & \mathbf{foe}_b \\ \mathbf{0} & 1 \end{bmatrix} \end{aligned}$$

where

$$\begin{aligned} \mathbf{r}_{ab} &= \mathbf{r}_{aw} \mathbf{r}_{bw}^T \\ \mathbf{r}_{ba} &= \mathbf{r}_{bw} \mathbf{r}_{aw}^T \\ \mathbf{foe}_a &= \mathbf{r}_{aw} \cdot (\mathbf{pos}_b - \mathbf{pos}_a) \\ \mathbf{foe}_b &= \mathbf{r}_{bw} \cdot (\mathbf{pos}_a - \mathbf{pos}_b) \end{aligned}$$

from which we can derive the matrix  $\mathbf{T}_{foe;a}$  for rectifying the image of camera  $a$  as

$$\mathbf{T}_{foe;a} = \begin{bmatrix} n(\mathbf{foe}_a) \\ n(\mathbf{z} \times \mathbf{foe}_a) \\ n(\mathbf{foe}_a \times (\mathbf{z} \times \mathbf{foe}_a)) \end{bmatrix} \quad (6)$$

where  $n(\mathbf{v}) = \mathbf{v}/\|\mathbf{v}\|$  is a normalizing function. The corresponding matrix  $\mathbf{T}_{foe;b}$  for rectifying the image of camera  $b$  can be derived similarly or more simply by the relation

$$\mathbf{T}_{foe;b} = -\mathbf{T}_{foe;a} \cdot \mathbf{r}_{ab} \quad (7)$$

For the case where  $\mathbf{foe}_a = \mathbf{z}$ , the last two rows of  $\mathbf{T}_{foe;a}$  can be any two orthonormal vectors perpendicular to  $\mathbf{z}$ .

### 3.2 Determining rotation $\mathbf{R}$

The epipolar line containing a point  $\mathbf{p}_{xyz}$  will be rotated around the origin (the camera's optical center) and along the epipolar plane until it becomes parallel to the  $\mathbf{foe}$ . The epipolar plane containing  $\mathbf{p}_{xyz}$  also contain the  $\mathbf{foe}$  (by definition) and the origin. The normal to that plane is

$$\mathbf{axis} = \mathbf{foe} \times \mathbf{p}_{xyz} \quad (8)$$

and will be the axis of rotation (see Fig. 3), thus ensuring that  $\mathbf{p}_{xyz}$  remains in the epipolar plane. In the case  $\mathbf{p}_{xyz} = \mathbf{foe}$ , the axis can be any vector normal to the  $\mathbf{foe}$  vector.

The angle of rotation needed can be computed by using the fact that the normal  $\mathbf{z} = (0, 0, 1)^T$  to the image plane has to be rotated until it is perpendicular to the  $\mathbf{foe}$ . This is because the new epipolar line has to be parallel to the  $\mathbf{foe}$ . The rotation angle is the angle between the normal  $\mathbf{z} = (0, 0, 1)^T$  projected on the epipolar plane (perpendicular

to the rotation axis) and the plane normal to the  $\mathbf{foe}$  also containing the origin. By projecting the point  $\mathbf{p}_{xyz}$  onto that plane, we can directly compute the angle. We have  $\mathbf{z}'$ , the normal  $\mathbf{z}$  projected on the epipolar plane defined as

$$\mathbf{z}' = \mathbf{axis} \times (\mathbf{z} \times \mathbf{axis}) = \begin{bmatrix} -\mathbf{axis}_x \mathbf{axis}_z \\ -\mathbf{axis}_y \mathbf{axis}_z \\ \mathbf{axis}_x^2 + \mathbf{axis}_y^2 \end{bmatrix}$$

and  $\mathbf{p}'$ , the projected  $\mathbf{p}_{xyz}$  on the plane normal to the  $\mathbf{foe}$ , defined as

$$\mathbf{p}' = \mathbf{T}_{foe}^T \mathbf{B} \mathbf{T}_{foe} \mathbf{p}_{xyz} \quad (9)$$

where  $\mathbf{T}_{foe}$  was previously defined in Eq. 6.

The rotation matrix  $\mathbf{R}_{foe}$  rotates the vector  $\mathbf{z}'$  onto the vector  $\mathbf{p}'$  around the axis of Eq. 8 and is defined as

$$\mathbf{R}_{foe} = \text{rot}_{\mathbf{p}', \mathbf{z}'} \quad (10)$$

where  $\text{rot}_{\mathbf{a}, \mathbf{b}}$  rotates vector  $\mathbf{b}$  onto vector  $\mathbf{a}$  such that

$$\text{rot}_{\mathbf{a}, \mathbf{b}} = \begin{bmatrix} n(\mathbf{a}) \\ n(\mathbf{a} \times \mathbf{b}) \\ n((\mathbf{a} \times \mathbf{b}) \times \mathbf{a}) \end{bmatrix}^T \begin{bmatrix} n(\mathbf{b}) \\ n(\mathbf{a} \times \mathbf{b}) \\ n((\mathbf{a} \times \mathbf{b}) \times \mathbf{b}) \end{bmatrix}$$

If the point  $\mathbf{q}_{foe}$  is available instead of point  $\mathbf{p}_{xyz}$ , (as would be the case for the inverse transformation of Eq. 4) we can still compute  $\mathbf{R}_{foe}$  from Eq. 10 by substituting  $\mathbf{q}_{xyz}$  for  $\mathbf{p}_{xyz}$  in Eq. 8 and 9 where  $\mathbf{q}_{xyz}$  is derived from  $\mathbf{q}_{foe}$  using Eq. 5. Notice that because  $\mathbf{p}_{xyz}$  and  $\mathbf{q}_{xyz}$  are in the same epipolar plane, the rotation axis will be the same. Also, the angle of rotation will also be the same since their projection onto the plane normal to the  $\mathbf{foe}$  is the same (modulo a scale factor).

### 3.3 Determining the scaling $\mathbf{S}$

The matrix  $\mathbf{S}_{foe}$  is used to project the epipolar line from the unit image plane (i.e. located at  $z = 1$ ) onto the cylinder of unit radius. To simplify notation in the following equation, we define

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

As shown in Eq. 3 and 4,  $\mathbf{S}_{foe}$  has one scalar parameter  $k$ . This parameter can be computed for a known point  $\mathbf{p}_{xyz}$  (Eq. 3) by enforcing unit radius and solving the resulting equation

$$\begin{aligned} \|\mathbf{B} \mathbf{q}_{foe}\| &= 1 \quad (11) \\ \left\| \mathbf{B} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{k} & 0 \\ 0 & 0 & \frac{1}{k} \end{bmatrix} \mathbf{T}_{foe} \mathbf{R}_{foe} \mathbf{p}_{xyz} \right\| &= 1 \end{aligned}$$

which yields the solution

$$k = \|\mathbf{B} \mathbf{T}_{foe} \mathbf{R}_{foe} \mathbf{p}_{xyz}\|$$

For the case of a known point  $\mathbf{q}_{foe}$  (Eq. 4), enforcing that the epipolar lines all have their  $z$  coordinates equal to 1 gives the equation

$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \cdot \mathbf{p}_{xyz} = 1$$

$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \cdot (\mathbf{R}_{foe}^T \mathbf{T}_{foe}^T \begin{bmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & k \end{bmatrix} \mathbf{q}_{foe}) = 1$$

which can be simplified to

$$(\mathbf{T}_{foe} \mathbf{c}_3) \cdot (\mathbf{A} \mathbf{q}_{foe}) + k(\mathbf{T}_{foe} \mathbf{c}_3) \cdot (\mathbf{B} \mathbf{q}_{foe}) = 1$$

where  $\mathbf{c}_3$  is the third column of rotation matrix  $\mathbf{R}_{foe}$ . The solution is then

$$k = \frac{1 - (\mathbf{T}_{foe} \mathbf{c}_3) \cdot (\mathbf{A} \mathbf{q}_{foe})}{(\mathbf{T}_{foe} \mathbf{c}_3) \cdot (\mathbf{B} \mathbf{q}_{foe})}$$

It should be noted that the denominator can never be zero because of Eq. 11 and the fact that  $\mathbf{T}_{foe} \mathbf{c}_3$  can never be zero or orthogonal to  $\mathbf{B} \mathbf{q}_{foe}$ .

### 3.4 Common angle interval

In general, a rectified image does not span the whole cylinder. The common angle interval is the interval that yields all common epipolar lines between two views. In order to control the number of epipolar lines extracted, it is important to determine this interval for each image.

Notice that the rectification process implicitly guarantees that a pair of corresponding epipolar lines have the *same* angle on their respective cylinder, and therefore the same row in the rectified images. The concern here is to determine the angle interval of epipolar lines effectively present in both images.

It can be shown that if a rectified image does *not* span the whole cylinder, then the extremum angles are given by two corners of the image. Based on this fact, it is sufficient to compute the angle of the four corners and one point between each pair of adjacent corners. By observing the ordering of these angles and taking into account the periodicity of angle measurements, it is possible to determine the angle interval for one image.

Given the angle intervals computed for each image separately, their intersection is the common angle interval sought. The subsequent stereo matching process has only to consider epipolar lines in that interval.

### 3.5 The case of uncalibrated cameras

Until now, it was always assumed that the cameras were calibrated, i.e. their internal parameters are known. The parameters are the principal point (optical axis), focal lengths and aspect ratio. More generally, we can represent all these parameters by a  $3 \times 3$  upper triangular matrix. In this section, we assume that only the *fundamental matrix* is available. This matrix effectively hides the internal parameters with the camera motion (external parameters) in a single matrix.

The fundamental matrix  $\mathbf{F}$  defines the epipolar relation between points  $\mathbf{p}_a$  and  $\mathbf{p}_b$  of the images as

$$\mathbf{p}_b^T \cdot \mathbf{F} \cdot \mathbf{p}_a = 0 \quad (12)$$

It is straightforward to extract the focus of expansion for each image by noticing that all points of one image must satisfy Eq. 12 when the point selected in the other image is its *foe*. More precisely, the relations for  $\mathbf{foe}_a$  and  $\mathbf{foe}_b$  are

$$\begin{aligned} \mathbf{p}_b^T \cdot \mathbf{F} \cdot \mathbf{foe}_a &= 0 \quad \forall \mathbf{p}_b \\ \mathbf{foe}_b^T \cdot \mathbf{F} \cdot \mathbf{p}_a &= 0 \quad \forall \mathbf{p}_a \end{aligned}$$

which yield the homogeneous linear equation systems

$$\mathbf{F} \cdot \mathbf{foe}_a = 0 \quad (13)$$

$$\mathbf{F}^T \cdot \mathbf{foe}_b = 0 \quad (14)$$

which are easily solved.

At this point, it remains to show how to derive the constituent of matrix  $\mathbf{L}_{foe}$  of Eq. 3 from the fundamental matrix  $\mathbf{F}$ . These are the matrices  $\mathbf{S}_{foe}$ ,  $\mathbf{R}_{foe}$ , and  $\mathbf{T}_{foe}$ .

The transformation  $\mathbf{T}_{foe;a}$  can be directly obtained from Eq. 6, using  $\mathbf{foe}_a$  obtained in Eq. 13. Symmetrically (using Eq. 14) we obtain

$$T_{foe;b} = \begin{bmatrix} n(\mathbf{foe}_b) \\ n(\mathbf{z} \times \mathbf{foe}_b) \\ n(\mathbf{foe}_b \times (\mathbf{z} \times \mathbf{foe}_b)) \end{bmatrix}$$

The rotation matrix  $\mathbf{R}_{foe}$  is computed from the *foe* (which is readily available from the fundamental matrix  $\mathbf{F}$ ) and the transform matrix  $\mathbf{T}_{foe}$ , exactly as described in Sec. 3.2.

Since the scaling matrix  $\mathbf{S}_{foe}$  is directly computed from the value of rotation matrix  $\mathbf{R}_{foe}$  and transform  $\mathbf{T}_{foe}$ , it is computed exactly as described in Sec. 3.3.

The rectification method is applicable regardless of the availability of the internal camera parameters. However, without these parameters, it is impossible to determine the minimum and maximum disparity interval which is of great utility in an subsequent stereo matching. In this paper, all the results obtained performed with known internal parameters.

### 3.6 Epipolar distortion and image size

The distortion induced by the rectification process in conjunction with the resampling of the original image can create a loss of pixel information, i.e. pixels in the original image are not accounted for and the information they carry is simply discarded during resampling. We measure this loss along epipolar lines, since it is along these lines that a subsequent stereo process will be carried out. To establish a measure of pixel information loss, we consider a rectified epipolar line segments of a length of one pixel and compute the length  $L$  of the original line segment that is remapped to it. For a given length  $L$ , we define the loss as

$$\text{loss}(L) = \begin{cases} 0 & L \leq 1 \\ 1 - 1/L & L > 1 \end{cases}$$

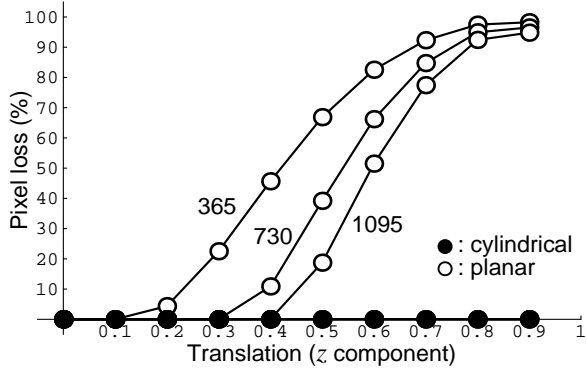


Figure 4: **Pixel loss as a function of camera translation  $T = (1, 0, z)$ . Rectified image width is 365, 730 and 1095 pixels for an original width of 256 pixels.**

A shrinking of original pixels (i.e.  $L > 1$ ) creates pixel information loss while a stretching (i.e.  $L < 1$ ) simply reduce the density of the rectified image. For a whole image, the measure is the expected loss over all rectified epipolar lines, broken down into individual one pixel segments.

The fundamental property of cylindrical rectification is the conservation of the length of epipolar lines. Since pixels do not stretch or shrink on these lines, no pixel information is lost during resampling, except for the unavoidable loss introduced by the interpolation process itself. For planar rectification, the length of epipolar lines is not preserved. This implies that some pixel loss will occur if the rectified image size is not large enough. In Fig. 4, three different rectified image width (365, 730, 1095 pixels) were used with both methods, for a range of camera translations  $T = (1, 0, z)$  with a  $z$  component in the range  $z \in [0, 1]$ . Cylindrical rectification shows no loss for any camera motion and any rectified image width<sup>2</sup>. However, planar rectification induces a pixel loss that depends on the camera geometry. To compensate for such a loss, the rectified images have to be enlarged, sometimes to the point where they become useless for subsequent stereo processing. For a  $z$  component equal to 1 (i.e.  $T = (1, 0, 1)$ ), all pixels are lost, regardless of image size.

## 4 Experiments and results

Some examples of rectification applied to different camera geometries are illustrated in this section. Fig. 5 presents an image plane and the rectification cylinder with the reprojected image, for an horizontal camera motion. In this case, the epipolar lines are already aligned. The rows represent different angles around the cylinder, from  $0^\circ$  to  $360^\circ$ . The image always appears twice since every cylinder point is *projective* across the cylinder axis. The number of rows determine the number of epipolar lines that are extracted from the image.

Fig. 6 depicts a camera geometry with forward motion. The original and rectified images are shown in Fig. 7 (planar rectification can not be used in this case). Notice how

<sup>2</sup>The minimum image width that guarantees no pixel loss is equal to  $\sqrt{w^2 + h^2}$  for an original image of size  $(w, h)$

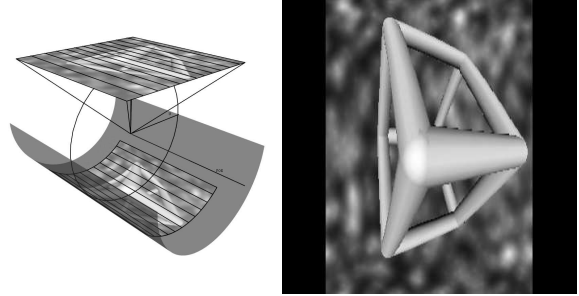


Figure 5: **Image "cube" rectified. Horizontal camera motion (foe =  $(1, 0, 0)$ ). A row represent an individual epipolar line.**

the rectified displacement of the sphere and cone is purely horizontal, as expected.

Fig. 8 depicts a typical camera geometry, suitable for planar rectification, with rectified images shown in Fig. 9. While the cylindrical rectification (images  $C_1, C_2$  in Fig. 9) introduces little distortion, planar rectification (images  $P_1, P_2$ ) significantly distorts the images, which are also larger to compensate for pixel information loss.

Examples where the foe is inside the image are obtained when the forward component of the motion is large enough with respect to the focal length (as in Fig. 7). It is important to note that planar rectification always yields an unbounded image (i.e. infinite size) for these cases and thus can not be applied.

The execution time for both methods is very similar. For many camera geometries, the slight advantage of planar rectification relating to the number of matrix computation is overcome by the extra burden of resampling larger rectified images to reduce pixel loss.

## 5 Conclusion

We presented a new method, called *cylindrical rectification*, for rectifying stereoscopic images under arbitrary camera geometry. It effectively remaps the images onto the surface of a unit cylinder whose axis goes through both cameras optical centers. It applies a transformation in projective space to each image point. A single linear transformation is required per epipolar line to rectify. While it does not preserves arbitrary straight lines, it preserves epipolar line lengths, thus insuring minimal loss of pixel information. As a consequence of allowing arbitrary camera motions, the rectified images are always bounded, with a size independent of camera motion.

The approach has been implemented and used successfully in the context of stereo matching [9], ego-motion estimation [10] and tridimensional reconstruction and has proved to provide added flexibility and accuracy at no significant cost in performance.

## References

- [1] N. Ayache and C. Hansen. Rectification of images for binocular and trinocular stereovision. In *Proc. of Int. Conf. on Pattern Recognition*, pages 11–16, Washington, D.C., 1988.
- [2] Q.-T. Luong and O. D. Faugeras. The fundamental matrix: Theory, algorithms, and stability analysis. *Int. J. Computer Vision*, 17:43–75, 1996.

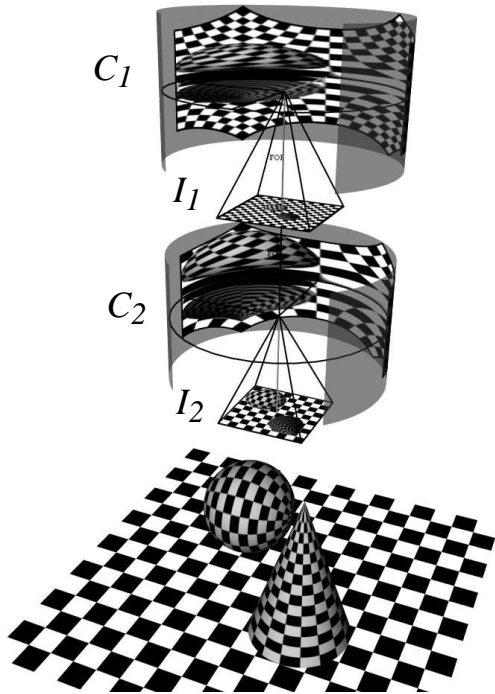


Figure 6: **Forward motion.** A sphere and cone are observed from two cameras displaced along their optical axis. The original images  $I_1, I_2$  are remapped onto the cylinder as  $C_1, C_2$ .

- [3] S. B. Kang, J. A. Webb, C. L. Zitnick, and T. Kanade. An active multibaseline stereo system with real-time image acquisition. Technical Report CMU-CS-94-167, School of Computer Science, Carnegie Mellon University, 1994.
- [4] O. Faugeras. *Three-dimensional computer vision*. MIT Press, Cambridge, 1993.
- [5] R. Hartley and R. Gupta. Computing matched-epipolar projections. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages 549–555, New York, N.Y., 1993.
- [6] P. Courtney, N. A. Thacker, and C. R. Brown. A hardware architecture for image rectification and ground plane obstacle detection. In *Proc. of Int. Conf. on Pattern Recognition*, pages 23–26, The Hague, Netherlands, 1992.
- [7] D. V. Papadimitriou and T. J. Dennis. Epipolar line estimation and rectification for stereo image pairs. *IEEE Trans. Image Processing*, 5(4):672–676, 1996.
- [8] L. Robert, M. Buffa, and M. Hébert. Weakly-calibrated stereo perception for rover navigation. In *Proc. 5th Int. Conference on Computer Vision*, pages 46–51, Cambridge, 1995.
- [9] I. J. Cox, S. Hingorani, B. M. Maggs, and S. B. Rao. A maximum likelihood stereo algorithm. *Computer Vision and Image Understanding*, 63(3):542–567, 1996.
- [10] S. Roy and I. J. Cox. Motion without structure. In *Proc. of Int. Conf. on Pattern Recognition*, volume 1, pages 728–734, Vienna, Austria, 1996.

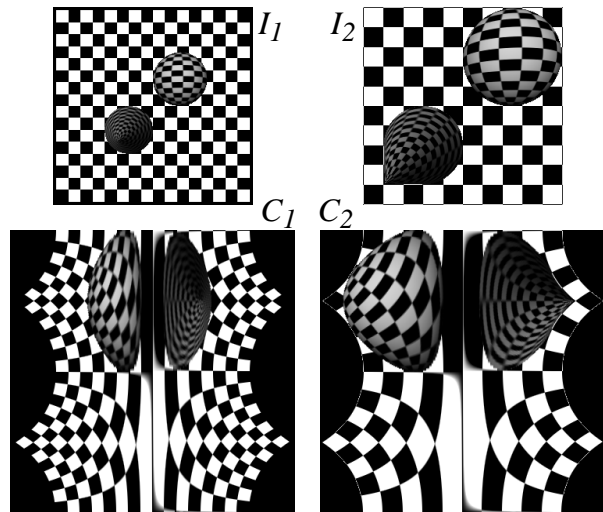


Figure 7: **Rectification of forward camera motion.** The images  $I_1, I_2$  are shown with their cylindrical rectification  $C_1, C_2$ . The rectified image displacements are all horizontal.

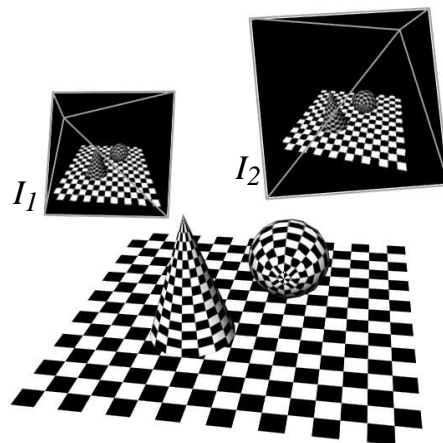


Figure 8: **Camera geometry suitable for planar rectification.**  $I_1, I_2$  are the original images.

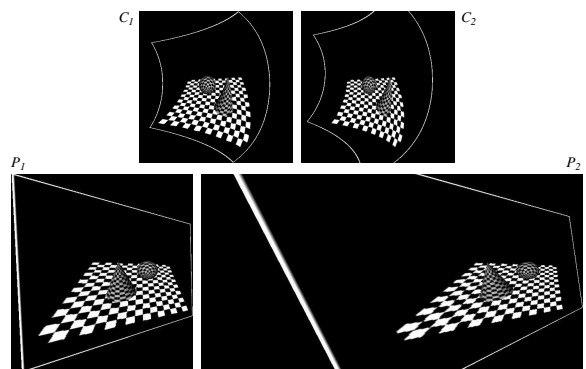


Figure 9: **Rectified images.** Cylindrical rectification ( $C_1, C_2$ ) and planar rectification ( $P_1, P_2$ )