

# Optical Engineering

OpticalEngineering.SPIEDigitalLibrary.org

## **Operator-based homogeneous coordinates: application in camera document scanning**

Rigoberto Juarez-Salazar  
Victor H. Diaz-Ramirez

# Operator-based homogeneous coordinates: application in camera document scanning

Rigoberto Juarez-Salazar<sup>a,\*</sup> and Victor H. Diaz-Ramirez<sup>b</sup>

<sup>a</sup>CONACYT–Instituto Politécnico Nacional, CITEDI, Nueva Tijuana, Tijuana, Baja California, México

<sup>b</sup>Instituto Politécnico Nacional, CITEDI, Nueva Tijuana, Tijuana, Baja California, México

**Abstract.** An operator-based approach for the study of homogeneous coordinates and projective geometry is proposed. First, some basic geometrical concepts and properties of the operators are investigated in the one- and two-dimensional cases. Then, the pinhole camera model is derived, and a simple method for homography estimation and camera calibration is explained. The usefulness of the analyzed theoretical framework is exemplified by addressing the perspective correction problem for a camera document scanning application. Several experimental results are provided for illustrative purposes. The proposed approach is expected to provide practical insights for inexperienced students on camera calibration, computer vision, and optical metrology among others. © 2017 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.OE.56.7.070801]

Keywords: homogeneous coordinates; pinhole camera; homography; camera calibration; perspective correction; camera document scanning.

Paper 170723T received May 11, 2017; accepted for publication Jul. 7, 2017; published online Jul. 27, 2017.

## 1 Introduction

Projective geometry is an important topic in computer vision because it provides a useful camera imaging model and its fundamental properties.<sup>1</sup> Some applications of this topic are found in camera motion,<sup>2</sup> camera calibration,<sup>3,4</sup> pose estimation for augmented reality,<sup>5</sup> perspective correction,<sup>6</sup> and three-dimensional (3-D) surface imaging<sup>7</sup> among others.

Theoretical concepts of projective geometry are analyzed simply and elegantly using homogeneous coordinates.<sup>8,9</sup> However, projective geometry is commonly presented in abstract form, leaving a gap in how to apply it in computer vision problems.<sup>10</sup> Moreover, homogeneous coordinates are used with a notation that masks basic geometrical aspects and may confuse the inexperienced readers.<sup>11</sup>

In this paper, a simple and intuitive approach to expose some useful concepts of projective geometry is addressed. For this, an alternative notation for homogeneous coordinates based on operators is suggested. To highlight the relevance of this topic in computer vision, the presentation is motivated by a specific problem, namely the perspective correction for a “camera scanner” application.

First, the proposed operators for homogeneous coordinates are defined in Sec. 2. Next, some basic concepts of projective geometry in the one- (1-D) and two-dimensional (2-D) cases are presented in Secs. 3 and 4, respectively. Then, the pinhole camera model is derived in Sec. 5. A perspective correction method, useful for camera document scanning, is described in Sec. 6. Finally, the conclusions of this work are given in Sec. 7. The paper is complemented with two appendices. Appendix A presents the direct linear transformation method for homography matrix estimation. Finally, a simple method to obtain the camera parameters from homographies is explained in Appendix B.

## 2 Definition of Operators

### 2.1 Operators $\mathcal{H}$ and $\mathcal{S}$

A point in an  $n$ -dimensional space will be represented by a vector of the form

$$\mathbf{x} = [x_1 \quad x_2 \quad \cdots \quad x_n]^T, \quad (1)$$

where  $[\cdot]^T$  denotes the transpose. The homogeneous coordinates of the point are obtained by adding an extra entry to  $\mathbf{x}$  with a value equal to the unity. The result is the  $(n + 1)$ -dimensional vector

$$\mathcal{H}[\mathbf{x}] = \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix}, \quad (2)$$

where  $\mathcal{H}$  will be referred to as the homogeneous operator.

The last entry of a homogeneous vector is known as the scale and will be recovered by the scale operator  $\mathcal{S}$ . This operator returns the last entry of any given vector. For instance, for the vectors in Eqs. (1) and (2), we have

$$x_n = \mathcal{S}[\mathbf{x}], \quad \text{and} \quad 1 = \mathcal{S}[\mathcal{H}[\mathbf{x}]]. \quad (3)$$

The operator  $\mathcal{H}$  sets the scale to unity. Another operator that sets the scale to zero is needed. For this, we define the operator

$$\mathcal{H}_0[\mathbf{x}] = \begin{bmatrix} \mathbf{x} \\ 0 \end{bmatrix}. \quad (4)$$

Note that the operator  $\mathcal{H}_0$  does not affect neither the direction nor the norm of  $\mathbf{x}$ . In projective geometry, the points represented by homogeneous coordinates of the form

$$\mathcal{H}_0[\mathbf{x}], \quad \mathbf{x} \neq \mathbf{0}_n, \quad (5)$$

\*Address all correspondence to: Rigoberto Juarez-Salazar, E-mail: [rjuarezsa@conacyt.mx](mailto:rjuarezsa@conacyt.mx)

are known as ideal points, where  $\mathbf{0}_n = [0, \dots, 0]^T$  is the  $n$ -dimensional zero vector.

The operators  $\mathcal{H}$  and  $\mathcal{H}_0$  can be considered as two particular cases of a more general operator defined as

$$\mathcal{H}_s[\mathbf{x}] = \begin{bmatrix} \mathbf{x} \\ s \end{bmatrix}, \tag{6}$$

where  $s$  is any scalar.

The procedure of adding an extra entry to vectors is reverted by returning the given vector except its last entry. For this, we define the inverse operator  $\mathcal{H}_0^{-1}$  as follows.

For any  $(n + 1)$ -dimensional vector

$$\mathbf{y} = [y_1 \ y_2 \ \dots \ y_n \ y_{n+1}]^T, \tag{7}$$

the operator  $\mathcal{H}_0^{-1}$  is defined as

$$\mathcal{H}_0^{-1}[\mathbf{y}] = [y_1 \ y_2 \ \dots \ y_n]^T. \tag{8}$$

Based on the operator  $\mathcal{H}_0^{-1}$ , the inverse of the operator  $\mathcal{H}_s$  for  $s \neq 0$  is defined as

**Table 1** Some useful equalities of the operators  $\mathcal{S}$ ,  $\mathcal{H}_s$ , and  $\mathcal{P}_{M,s}$ . In all cases, we consider  $\lambda \neq 0$ ,  $\gamma_1$  and  $\gamma_2$  are any scalars,  $\mathbf{x}$  is a  $n$ -dimensional vector as given in Eq. (1),  $\Xi_s$  is the matrix defined in Eq. (13),  $\mathbf{y} = \lambda \mathcal{H}_s[\mathbf{x}]$ ,  $M$  is a matrix of size  $(n + 1) \times (n + 1)$ , and  $W$  is a matrix of size  $m \times (n + 1)$ .

Property	Description	
(P1)	$\mathcal{H}_s^{-1}[\mathcal{H}_s[\mathbf{x}]] = \mathbf{x}$	$s \neq 0$
(P2)	$\mathcal{H}_s[\mathcal{H}_s^{-1}[\mathbf{y}]] = \frac{s}{S[\mathbf{y}]} \mathbf{y}$	$s \neq 0$
(P3)	$\mathbf{x} = \mathcal{H}_s^{-1}[\mathbf{y}] \Leftrightarrow \frac{S[\mathbf{y}]}{s} \mathcal{H}_s[\mathbf{x}] = \mathbf{y}$	$s \neq 0$
(P4)	$\mathcal{H}_s[\lambda \mathbf{x}] = \lambda \mathcal{H}_{s/\lambda}[\mathbf{x}]$	
(P5)	$\lambda \mathcal{H}_s[\mathbf{x}] = \mathcal{H}_{\lambda s}[\lambda \mathbf{x}]$	
(P6)	$\mathcal{H}_s[W \mathcal{H}_s[\mathbf{x}]] = \begin{bmatrix} W \\ \mathcal{H}[\mathbf{0}_n]^T \end{bmatrix} \mathcal{H}_s[\mathbf{x}]$	
(P7)	$\mathcal{H}_s[\mathbf{x}_1 \pm \mathbf{x}_2] = \mathcal{H}_{s_1}[\mathbf{x}_1] \pm \mathcal{H}_{s_2}[\mathbf{x}_2]$	$s_1 \pm s_2 = s$
(P8)	$\mathcal{H}_s^{-1}[\lambda \mathbf{y}] = \mathcal{H}_s^{-1}[\mathbf{y}]$	$s \neq 0$
(P9)	$\lambda \mathcal{H}_s^{-1}[\mathbf{y}] = \mathcal{H}_{\lambda s}^{-1}[\mathbf{y}]$	$s \neq 0$
(P10)	$\mathcal{H}_0^{-1}[\gamma_1 \mathbf{y}_1 + \gamma_2 \mathbf{y}_2] = \gamma_1 \mathcal{H}_0^{-1}[\mathbf{y}_1] + \gamma_2 \mathcal{H}_0^{-1}[\mathbf{y}_2]$	
(P11)	$S[\gamma_1 \mathbf{x}_1 + \gamma_2 \mathbf{x}_2] = \gamma_1 S[\mathbf{x}_1] + \gamma_2 S[\mathbf{x}_2]$	
(P12)	$\mathcal{H}_s^{-1}[\mathbf{y}_1 + \mathbf{y}_2] = \frac{1}{S[\mathbf{y}_1 + \mathbf{y}_2]} (S[\mathbf{y}_1] \mathcal{H}_s^{-1}[\mathbf{y}_1] + S[\mathbf{y}_2] \mathcal{H}_s^{-1}[\mathbf{y}_2])$	$s \neq 0$
(P13)	$\mathcal{H}_s^{-1}[\mathbf{y}_1 + \mathbf{y}_2] - \mathcal{H}_s^{-1}[\mathbf{y}_1] = \frac{S[\mathbf{y}_2]}{S[\mathbf{y}_1 + \mathbf{y}_2]} (\mathcal{H}_s^{-1}[\mathbf{y}_2] - \mathcal{H}_s^{-1}[\mathbf{y}_1])$	$s \neq 0$
(P14)	$\mathcal{H}_s[\mathbf{x}] = \Xi_s \mathcal{H}[\mathbf{x}]$	
(P15)	$\mathcal{H}_s^{-1}[\mathbf{y}] = \mathcal{H}^{-1}[\Xi_s^{-1} \mathbf{y}]$	$s \neq 0$
(P16)	$\mathcal{P}_{\lambda M, s}[\mathbf{x}] = \mathcal{P}_{M, s}[\mathbf{x}]$	$s \neq 0$
(P17)	$\mathcal{P}_{M, s}^{-1}[\mathbf{x}] = \mathcal{P}_{M^{-1}, s}[\mathbf{x}]$	$\det M \neq 0$
(P18)	$\mathbf{x}_2 = \mathcal{P}_{M, s}[\mathbf{x}_1] \Leftrightarrow \mathcal{P}_{M^{-1}, s}[\mathbf{x}_2] = \mathbf{x}_1$	$\det M \neq 0$
(P19)	$\mathcal{P}_{M, s}[\lambda \mathbf{x}] = \lambda \mathcal{P}_{M, s/\lambda}[\mathbf{x}]$	
(P20)	$\lambda \mathcal{P}_{M, s}[\mathbf{x}] = \mathcal{P}_{M, \lambda s}[\lambda \mathbf{x}]$	
(P21)	$\mathcal{P}_{M_2, s}[\mathcal{P}_{M_1, s}[\mathbf{x}]] = \mathcal{P}_{M_2 M_1, s}[\mathbf{x}]$	
(P22)	$\mathcal{P}_{\mathbf{I}, s}[\mathbf{x}] = \mathbf{x}$	
(P23)	$\mathcal{P}_{M, s}[\mathbf{x}] = \mathcal{P}_{\Xi_s^{-1} M \Xi_s}[\mathbf{x}]$	$s \neq 0$
(P24)	$\mathcal{H}_s[\mathcal{P}_{M, s}[\mathbf{x}]] = \frac{s}{S[M \mathcal{H}[\mathbf{x}]]} M \mathcal{H}[\mathbf{x}]$	$s \neq 0$
(P25)	$\mathcal{P}_{M, s}[\mathcal{H}_s^{-1}[\mathbf{y}]] = \mathcal{H}_s^{-1}[M \mathbf{y}]$	$s \neq 0$

$$\mathcal{H}_s^{-1}[\mathbf{y}] = \frac{s}{S[\mathbf{y}]} \mathcal{H}_0^{-1}[\mathbf{y}]. \quad (9)$$

In particular, the operator  $\mathcal{H}_1^{-1}$  (written simply as  $\mathcal{H}^{-1}$ ) will be referred to as the inverse homogeneous operator.

The inverse  $\mathcal{H}_0^{-1}$  is a linear operator. That is, for any two scalars  $\lambda_1$  and  $\lambda_2$ , we have

$$\mathcal{H}_0^{-1}[\lambda_1 \mathbf{y}_1 + \lambda_2 \mathbf{y}_2] = \lambda_1 \mathcal{H}_0^{-1}[\mathbf{y}_1] + \lambda_2 \mathcal{H}_0^{-1}[\mathbf{y}_2]. \quad (10)$$

On the other hand, the operator  $\mathcal{H}_s^{-1}$  is invariant to nonzero scalar multiplication of its argument. That is,

$$\mathcal{H}_s^{-1}[\lambda \mathbf{y}] = \mathcal{H}_s^{-1}[\mathbf{y}], \quad s, \lambda \neq 0. \quad (11)$$

The operators  $\mathcal{H}_s$  and  $\mathcal{H}_s^{-1}$  can be expressed in terms of the homogeneous operator and its inverse, namely

$$\begin{aligned} \mathcal{H}_s[\mathbf{x}] &= \Xi_s \mathcal{H}[\mathbf{x}], \quad \text{and} \\ \mathcal{H}_s^{-1}[\mathbf{y}] &= \mathcal{H}^{-1}[\Xi_s^{-1} \mathbf{y}], \quad s \neq 0, \end{aligned} \quad (12)$$

where

$$\Xi_s = \begin{bmatrix} \mathbb{I}_n & \mathbf{0}_n \\ \mathbf{0}_n^T & s \end{bmatrix}, \quad (13)$$

and  $\mathbb{I}_n$  being the  $n \times n$  identity matrix.

## 2.2 Projection Operator $\mathcal{P}$

In general terms, the homogeneous operator carries the representation of a point from  $n$ - to  $(n+1)$ -dimensional vectors while the inverse homogeneous operator returns the representation from  $(n+1)$ - to  $n$ -dimensional vectors. An important transformation emerges when, in the  $(n+1)$ -dimensional space, a linear mapping is applied. Mathematically, we describe this transformation by the projection operator defined as

$$\mathcal{P}_M[\mathbf{x}] = \mathcal{H}^{-1}[M\mathcal{H}[\mathbf{x}]], \quad (14)$$

where  $M$  is an  $(n+1) \times (n+1)$  matrix. A generalized version of the projection operator is obtained using  $\mathcal{H}_s$  and its inverse as

$$\mathcal{P}_{M,s}[\mathbf{x}] = \mathcal{H}_s^{-1}[M\mathcal{H}_s[\mathbf{x}]]. \quad (15)$$

From Eq. (11), it follows that, for  $s \neq 0$ , the operator  $\mathcal{P}_{M,s}$  is invariant to nonzero scalar multiplication of the matrix  $M$ ; that is

$$\mathcal{P}_{\lambda M,s}[\mathbf{x}] = \mathcal{P}_{M,s}[\mathbf{x}], \quad s, \lambda \neq 0. \quad (16)$$

Let  $\mathbf{b} = \mathcal{P}_{M,s}[\mathbf{a}]$  with  $M$  being a nonsingular matrix. From Eq. (15), we have that  $\mathbf{a} = \mathcal{H}_s^{-1}[M^{-1}\mathcal{H}_s[\mathbf{b}]]$ . Therefore, the inverse operator  $\mathcal{P}_{M,s}^{-1}$  is given by

$$\mathcal{P}_{M,s}^{-1}[\mathbf{x}] = \mathcal{P}_{M^{-1},s}[\mathbf{x}], \quad \det M \neq 0, \quad (17)$$

where  $\det M$  denotes the determinant of  $M$ .

Using the equations in Eq. (12), the operator  $\mathcal{P}_{M,s}$  can be expressed in terms of the projection operator, namely

$$\mathcal{P}_{M,s}[\mathbf{x}] = \mathcal{P}_{\Xi_s^{-1}M\Xi_s}[\mathbf{x}], \quad s \neq 0. \quad (18)$$

Note that  $M$  and  $\Xi_s^{-1}M\Xi_s$  are similar matrices. Some useful equalities of the defined operators are summarized in Table 1. For a more comprehensible reading of this paper, the reader is encouraged to demonstrate all the equalities in Table 1.

In the following sections, the defined operators are studied from an intuitive geometrical approach for the 1-D and 2-D cases. Then, the usefulness of this theoretical framework is illustrated by addressing the perspective correction problem for camera document scanning.

## 3 One-Dimensional Space

The 1-D real space can be represented as a line as shown in Fig. 1(a). In this space, a point at a finite distance from the origin is represented by a real number  $x$ ; otherwise, the point is represented by the symbol  $\infty$ .

Alternatively, the 1-D space can be represented by the projective line  $y = 1$  in the  $xy$ -plane as shown in Fig. 1(b). Thus, the coordinate  $x$  of a point in the line becomes the vector

$$\mathbf{y} = \mathcal{H}[x] = \begin{bmatrix} x \\ 1 \end{bmatrix}. \quad (19)$$

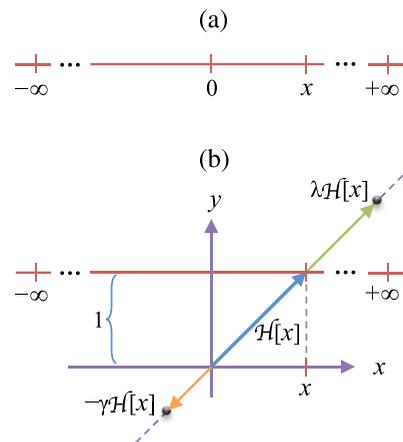
The coordinate  $x$  can be recovered from its homogeneous version  $\mathbf{y}$  as the intersection between the line  $y = 1$  and the line with direction  $\mathcal{H}[x]$  passing through the origin as shown in Fig. 1(b). This is described mathematically as

$$x = \mathcal{H}^{-1}[\mathbf{y}]. \quad (20)$$

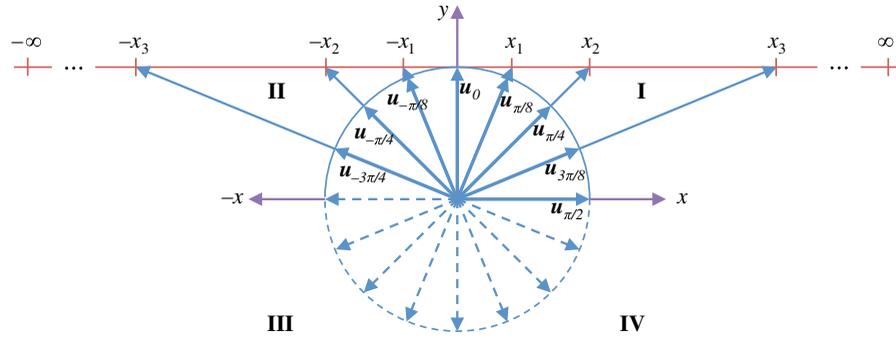
Note that the result is invariant to the scalar multiplication of  $\mathbf{y}$  by a nonzero scalar [e.g.,  $\lambda$  and  $-\gamma$  as shown in Fig. 1(b)] because the intersection between lines is unaltered. In other words,  $\mathcal{H}^{-1}[\lambda \mathbf{y}] = \mathcal{H}^{-1}[-\gamma \mathbf{y}] = \mathcal{H}^{-1}[\mathbf{y}]$  as stated by Eq. (11).

### 3.1 Ideal Point

Homogeneous coordinates provide a different form to identify points of the real line. Consider the unit vector



**Fig. 1** (a) The real line as the 1-D Euclidean space. (b) The 1-D space represented by the projective line ( $y = 1$ ) in a 2-D Euclidean space.



**Fig. 2** Representation of points of the real line using homogeneous coordinates. Opposite homogeneous vectors represent the same point; thus, there is a single point at infinity, given by  $\mathbf{u}(\pi/2) = [1, 0]^T$ .

$$\mathbf{u}(\theta) = \begin{bmatrix} \sin \theta \\ \cos \theta \end{bmatrix}. \quad (21)$$

Thus, the homogeneous representation of  $x$  given by Eq. (19) becomes

$$\mathcal{H}[x] = \lambda \mathbf{u}(\theta), \quad (22)$$

where  $\lambda^2 = 1 + x^2$  and  $\tan \theta = x$ . From Eq. (22), we obtain

$$x = \mathcal{H}^{-1}[\mathbf{u}(\theta)]. \quad (23)$$

Since a vector  $\mathbf{u}$  and its opposite  $-\mathbf{u}$  represent the same point (i.e.,  $\mathcal{H}^{-1}[\mathbf{u}] = \mathcal{H}^{-1}[-\mathbf{u}]$ ), all points of the real line at a finite distance from the origin are associated to a unique angle  $\theta$  in the open interval  $(-\pi/2, \pi/2)$ ; i.e., the vectors  $\mathbf{u}$  different to  $[1, 0]$  and  $[-1, 0]$  in the quadrants I and II, as shown in Fig. 2.

Intuitively, the real line in Euclidean representation has two points at infinity, namely  $-\infty$  and  $+\infty$ . However, in projective geometry, the real line has only a single point at infinity given by the homogeneous coordinates

$$\boldsymbol{\psi} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad (24)$$

which is associated to  $\mathbf{u}(\pi/2)$ , as shown in Fig. 2. It could be argued that  $\pi/2$  corresponds to  $+\infty$  while  $-\pi/2$  to  $-\infty$ . However, note that  $\mathbf{u}(-\pi/2) = -\boldsymbol{\psi}$  is the opposite of  $\boldsymbol{\psi}$ . Hence, they represent the same point.

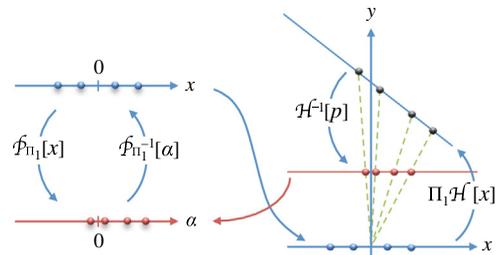
Note that  $\mathcal{H}^{-1}[\boldsymbol{\psi}] = 1/0$  is consistent with the notion that  $\boldsymbol{\psi}$  represents a point at infinity distance from the origin. According to the concepts of projective geometry, the vector  $\boldsymbol{\psi}$  represents an ideal point, see Eq. (5).

### 3.2 One-Dimensional Projection

The line  $y = 0$  can be transformed to any other line by applying a rotation  $\mathbf{Q} = [\mathbf{q}_1, \mathbf{q}_2]$  and a translation  $\mathbf{s}$ . Thus, a point in the line  $y = 0$ , represented by the scalar  $x$ , becomes a point in the  $xy$ -plane given by the vector

$$\begin{aligned} \mathbf{p} &= q\mathcal{H}_0[x] + \mathbf{s} \\ &= \Pi_1 \mathcal{H}[x], \end{aligned} \quad (25)$$

where the matrix  $\Pi_1$  will be referred to as the reference line parameters and has the explicit form



**Fig. 3** The 1-D projection.

$$\Pi_1 = [\mathbf{q}_1 \quad \mathbf{s}]. \quad (26)$$

The first column of  $\Pi_1$  and the determinant  $\det \Pi_1$  provide the direction of the reference line and its distance from the origin, respectively.

If the matrix  $\Pi_1$  is singular, the vectors  $\mathbf{q}_1$  and  $\mathbf{s}$  are collinear. In this case, the origin is a point of the transformed line (the distance of the line from the origin is zero). The matrix  $\Pi_1$  is nonsingular when  $\mathbf{q}_1$  and  $\mathbf{s}$  are linearly independent. In this case, the origin is not a point of the transformed line.

Let  $\mathbf{p}$  in Eq. (25) be the homogeneous coordinates of a point  $\alpha$  in the line. Thus, we obtain the 1-D projection

$$\begin{aligned} \alpha &= \mathcal{H}^{-1}[\mathbf{p}] \\ &= \mathcal{H}^{-1}[\Pi_1 \mathcal{H}[x]] \\ &= \mathcal{P}_{\Pi_1}[x]. \end{aligned} \quad (27)$$

The transformations by  $\mathcal{P}_{\Pi_1}[x]$  and its inverse  $\mathcal{P}_{\Pi_1^{-1}}[\alpha]$  are shown in Fig. 3.

## 4 Two-Dimensional Space

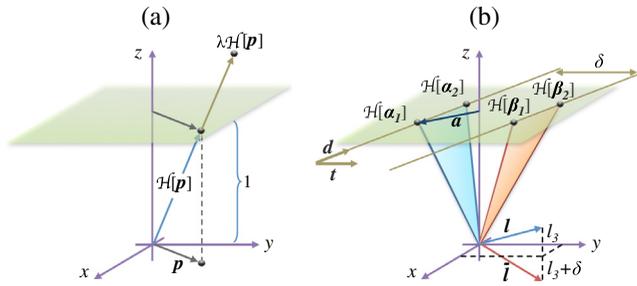
### 4.1 Points and Lines in the Plane

Any point in the 2-D space can be represented as the vector

$$\mathbf{x} = [x_1 \quad x_2]^T. \quad (28)$$

Moreover, the point  $\mathbf{x}$  can be represented by its homogeneous coordinates

$$\mathcal{H}[\mathbf{x}] = \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix}, \quad (29)$$



**Fig. 4** (a) The 2-D space represented by the projective plane ( $z = 1$ ). (b) Parallel lines in the plane.

as shown in Fig. 4(a). Note that  $\mathcal{H}$  takes the 2-D vector  $\mathbf{x}$  (in the plane  $z = 0$ ) and converts it to the 3-D vector  $\mathcal{H}[\mathbf{x}]$ , where  $\mathbf{x}$  is unaltered but now it lies in the projective plane  $z = 1$ . It is worth mentioning that the vector  $\mathbf{x}$  can be recovered from  $\lambda\mathcal{H}[\mathbf{x}]$  as the point of intersection of the line with points  $\mathbf{0}_3$  and  $\lambda\mathcal{H}[\mathbf{x}]$ , see Eq. (11). That is,

$$\mathbf{x} = \mathcal{H}^{-1}[\lambda\mathcal{H}[\mathbf{x}]], \quad \lambda \neq 0. \quad (30)$$

A line in the plane  $xy$ -plane can be written as the homogeneous equation

$$l_1x_1 + l_2x_2 + l_3 = 0, \quad (31)$$

where  $l_1$ ,  $l_2$ , and  $l_3$  are coefficients. Using homogeneous coordinates, Eq. (31) becomes

$$\mathbf{l}^T \mathcal{H}[\mathbf{x}] = 0, \quad (32)$$

where  $\mathbf{x} = [x_1, x_2]^T$  is a point of the line and  $\mathbf{l} = [l_1, l_2, l_3]^T$  is the vector that defines the line. Equation (32) exhibits that  $\mathbf{l}$  and  $\mathcal{H}[\mathbf{x}]$  are orthogonal vectors. Note that the vector  $\mathbf{l}$  is unique up to scale, i.e., the vectors  $\mathbf{l}$  and  $\lambda\mathbf{l}$ , with  $\lambda \neq 0$ , represent the same line.

Let  $\mathbf{x}_1$  and  $\mathbf{x}_2$  be two different points in the  $xy$ -plane. The vector  $\mathbf{l}$  of the line passing through  $\mathbf{x}_1$  and  $\mathbf{x}_2$  can be obtained by the cross product as

$$\mathbf{l} = \mathcal{H}[\mathbf{x}_1] \times \mathcal{H}[\mathbf{x}_2]. \quad (33)$$

By definition of the cross product, the vector  $\mathbf{l}$  is orthogonal to  $\mathcal{H}[\mathbf{x}_1]$  and  $\mathcal{H}[\mathbf{x}_2]$ . Therefore, these vectors satisfy Eq. (32).

Consider two lines defined by the vectors  $\mathbf{l}_1$  and  $\mathbf{l}_2$ . If  $\mathbf{x}$  is the intersection point of these lines, then  $\mathcal{H}[\mathbf{x}]$  is orthogonal to  $\mathbf{l}_1$  and  $\mathbf{l}_2$ . That is

$$\lambda\mathcal{H}[\mathbf{x}] = \mathbf{l}_1 \times \mathbf{l}_2, \quad (34)$$

where  $\lambda = S[\mathbf{l}_1 \times \mathbf{l}_2]$ . Therefore, the intersection point of the lines  $\mathbf{l}_1$  and  $\mathbf{l}_2$  is

$$\mathbf{x} = \mathcal{H}^{-1}[\mathbf{l}_1 \times \mathbf{l}_2]. \quad (35)$$

## 4.2 Parallel Lines

Two different lines are parallel if its defining vectors are of the form

$$\mathbf{l} = \begin{bmatrix} l_1 \\ l_2 \\ l_3 \end{bmatrix}, \quad \bar{\mathbf{l}} = \lambda \begin{bmatrix} l_1 \\ l_2 \\ l_3 + \delta \end{bmatrix}, \quad (36)$$

where  $\lambda, \delta \neq 0$ . This can be verified as follows. Consider two parallel lines in the plane with points  $\alpha$  and  $\beta$  given, respectively, by

$$\alpha = \mathbf{a} + \gamma\mathbf{d}, \quad \text{and} \quad \beta = \mathbf{a} + \gamma\mathbf{d} + \delta\mathbf{t}, \quad (37)$$

where  $\gamma$  is a parameter,  $\delta \neq 0$  is a constant,  $\mathbf{a}$  is a reference point,  $\mathbf{d}$  is a unit vector with direction of the line, and  $\mathbf{t}$  is a unit vector orthogonal to  $\mathbf{d}$ , i.e.,

$$\mathcal{H}_0[\mathbf{t}] \times \mathcal{H}_0[\mathbf{d}] = \mathcal{H}[\mathbf{0}_2], \quad (38)$$

as shown in Fig. 4(b). Two points of each line are

$$\alpha_1 = \mathbf{a} + \gamma_1\mathbf{d}, \quad \beta_1 = \mathbf{a} + \bar{\gamma}_1\mathbf{d} + \delta\mathbf{t},$$

$$\alpha_2 = \mathbf{a} + \gamma_2\mathbf{d}, \quad \beta_2 = \mathbf{a} + \bar{\gamma}_2\mathbf{d} + \delta\mathbf{t}.$$

Thus, the vector of the line with points  $\alpha$  is

$$\begin{aligned} \mathbf{l} &= \mathcal{H}[\alpha_1] \times \mathcal{H}[\alpha_2] \\ &= (\gamma_2 - \gamma_1)\mathcal{H}[\mathbf{a}] \times \mathcal{H}_0[\mathbf{d}], \end{aligned} \quad (39)$$

or, since the line is unaffected by scaling of its vector

$$\mathbf{l} = \mathcal{H}[\mathbf{a}] \times \mathcal{H}_0[\mathbf{d}]. \quad (40)$$

Similarly, the vector of the line with points  $\beta$  is

$$\begin{aligned} \bar{\mathbf{l}} &= \mathcal{H}[\beta_1] \times \mathcal{H}[\beta_2] \\ &= (\bar{\gamma}_2 - \bar{\gamma}_1)(\mathbf{l} + \delta\mathcal{H}[\mathbf{0}_2]) \\ &= \lambda(\mathbf{l} + \delta\mathcal{H}[\mathbf{0}_2]), \end{aligned} \quad (41)$$

where  $\lambda = \bar{\gamma}_2 - \bar{\gamma}_1$ . Therefore, the vectors  $\mathbf{l}$  and  $\bar{\mathbf{l}}$  given in Eq. (36) represent parallel lines.

It is worth mentioning that, if  $\mathbf{l}$  is the vector of a line with direction  $\mathbf{d}$  [see Eq. (40)], then the vector  $\mathcal{H}_0^{-1}[\mathbf{l}]$  is orthogonal to  $\mathbf{d}$ , namely

$$\begin{aligned} \mathbf{d}^T \mathcal{H}_0^{-1}[\mathbf{l}] &= \mathbf{d}^T \mathcal{H}_0^{-1}[\mathcal{H}[\mathbf{a}] \times \mathcal{H}_0[\mathbf{d}]] \\ &= \mathcal{H}_0[\mathbf{d}]^T (\mathcal{H}[\mathbf{a}] \times \mathcal{H}_0[\mathbf{d}]) \\ &= 0. \end{aligned} \quad (42)$$

## 4.3 Ideal Points and the Line at Infinity

In the Euclidean geometry, two parallel lines in the plane do not intersect. However, in the projective geometry, two different lines always intersect at a point. Consider the parallel lines given by the vectors in Eq. (36). Using Eq. (35), the intersection point is

$$\begin{aligned} \mathcal{H}^{-1}[\mathbf{l} \times \bar{\mathbf{l}}] &= \mathcal{H}^{-1}[\mathcal{H}[\mathbf{0}_2] \times \mathbf{l}] \\ &= \mathcal{H}^{-1}[\boldsymbol{\psi}], \end{aligned} \quad (43)$$

where

$$\boldsymbol{\psi} = [-l_2 \quad l_1 \quad 0]^T \quad (44)$$

is the point of intersection in homogeneous coordinates. Note that  $\mathcal{H}^{-1}[\boldsymbol{\psi}] = [-l_2/0, l_1/0]$  provides the insight that

parallel lines intersecting at a point at infinity. As in the 1-D case, the vector  $\boldsymbol{\psi}$  represents ideal points, see Eq. (5).

The vector  $\boldsymbol{\psi}$  is associated with the direction  $\mathbf{d}$  of the line  $\mathbf{l}$ . This is verified by taking into account that  $\mathcal{H}_0^{-1}[\mathbf{l}]$  is orthogonal to  $\mathbf{d}$  [Eq. (42)] as well as to  $\mathcal{H}_0^{-1}[\boldsymbol{\psi}]$  ( $\mathcal{H}_0^{-1}[\boldsymbol{\psi}]^T \mathcal{H}_0^{-1}[\mathbf{l}] = 0$ ), then

$$\mathcal{H}_0^{-1}[\boldsymbol{\psi}] = \lambda \mathbf{d}, \quad (45)$$

where  $\lambda$  is some nonzero scalar.

All ideal points given by Eq. (44) are collinear. The vector of such a line, known as the line at infinity, is

$$\mathbf{l}_\infty = \mathcal{H}[\mathbf{0}_2] = [0 \ 0 \ 1]^T. \quad (46)$$

This can be easily verified by  $\mathbf{l}_\infty^T \boldsymbol{\psi} = 0$  as required by Eq. (32).

The ideal point  $\boldsymbol{\psi}$  in Eq. (44) was obtained as the intersection of two parallel lines  $\mathbf{l}$  and  $\bar{\mathbf{l}}$ . However, the intuition suggests that the same result could be obtained by computing the intersection of the line  $\mathbf{l}$  and the line at infinity  $\mathbf{l}_\infty$ . In fact, we have that

$$\boldsymbol{\psi} = \mathbf{l} \times \mathbf{l}_\infty. \quad (47)$$

Thus, using Eq. (45), the direction  $\mathbf{d}$  of any line  $\mathbf{l}$  is given by

$$\lambda \mathbf{d} = \mathcal{H}_0^{-1}[\mathbf{l} \times \mathbf{l}_\infty], \quad (48)$$

where  $\lambda$  is a nonzero scale factor. For this reason, the line  $\mathbf{l}_\infty$  is interpreted as the set of directions of lines in the plane.

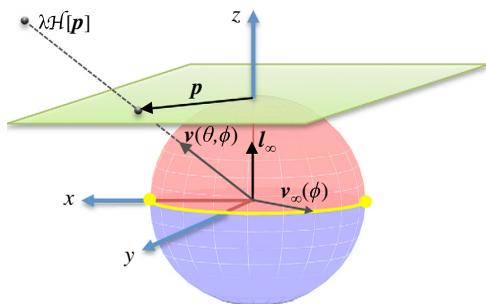
Similar to the 1-D case, homogeneous coordinates provide a different form to identify points of the plane. Consider the unit vector

$$\mathbf{v} = [\sin \theta \cos \phi \ \sin \theta \sin \phi \ \cos \theta]^T, \quad (49)$$

where  $\theta$  and  $\phi$  are polar and azimuth angles, respectively. Thus, the homogeneous coordinates for each point  $\mathbf{x} = [x_1, x_2]^T$  on the plane are given by

$$\mathcal{H}[\mathbf{x}] = \lambda \mathbf{v}, \quad (50)$$

where  $\lambda^2 = 1 + x_1^2 + x_2^2$ . From Eq. (50), the following relation holds



**Fig. 5** Representation of points of the plane using homogeneous coordinates  $\mathbf{v}$ . The upper hemisphere represents points of the plane at a finite distance from the origin, and the half of the equator (yellow semicircle) represents points at infinity.

$$\mathbf{x} = \mathcal{H}^{-1}[\mathbf{v}(\theta, \phi)]. \quad (51)$$

The points of the plane at a finite distance from the origin are given by  $\mathbf{v}(\theta, \phi)$  with  $\theta \in [0, \pi/2)$  and  $\phi \in [-\pi, \pi)$ , i.e., the upper hemisphere of the unit sphere, see Fig. 5. The points of the plane at infinity distance from the origin are parameterized by  $\theta = \pi/2$  and  $\phi \in (-\pi/2, \pi/2]$ . These points have the homogeneous coordinates

$$\mathbf{v}_\infty = [\cos \phi \ \sin \phi \ 0]^T, \quad (52)$$

see Eq. (44). That is, the ideal points are represented by the half equator of the unit sphere, see yellow line in Fig. 5.

#### 4.4 Two-Dimensional Projection

Any plane in the 3-D space can be obtained as the plane  $z = 0$  after a rotation  $Q = [\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3]$  and a translation  $\mathbf{s}$ . Thus, the points represented by  $\mathbf{x} = [x_1, x_2]^T$ , becomes

$$\begin{aligned} \mathbf{p} &= Q \mathcal{H}_0[\mathbf{x}] + \mathbf{s} \\ &= \Pi_2 \mathcal{H}[\mathbf{x}], \end{aligned} \quad (53)$$

where the matrix  $\Pi_2$  will be referred to as the reference plane parameters and has the explicit form

$$\Pi_2 = [\mathbf{q}_1 \ \mathbf{q}_2 \ \mathbf{s}]. \quad (54)$$

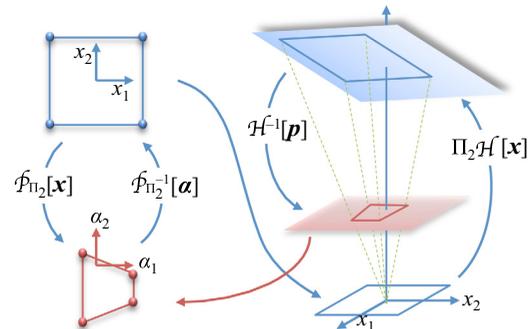
The cross product of the first two columns of  $\Pi_2$  and  $\det \Pi_2$  provides the normal to the reference plane and its distance from the origin, respectively.

The matrix  $\Pi_2$  is singular when  $\mathbf{r}_1$ ,  $\mathbf{r}_2$ , and  $\mathbf{t}$  are coplanar. In this case, the origin  $\mathbf{0}_3$  is a point of the transformed plane (the distance of the reference plane from the origin is zero). Otherwise,  $\Pi_2$  is a nonsingular.

Let  $\mathbf{p}$  in Eq. (53) be the homogeneous coordinates of a point  $\boldsymbol{\alpha}$  in the projective plane. Thus, the relation between the points  $\boldsymbol{\alpha}$  and  $\mathbf{x}$  is given by the 2-D projection  $\mathcal{P}_{\Pi_2}$ , namely

$$\begin{aligned} \boldsymbol{\alpha} &= \mathcal{H}^{-1}[\mathbf{p}] \\ &= \mathcal{H}^{-1}[\Pi_2 \mathcal{H}[\mathbf{x}]] \\ &= \mathcal{P}_{\Pi_2}[\mathbf{x}]. \end{aligned} \quad (55)$$

The projection  $\mathcal{P}_{\Pi_2}[\mathbf{x}]$  and its inverse  $\mathcal{P}_{\Pi_2^{-1}}[\boldsymbol{\alpha}]$  are shown in Fig. 6.



**Fig. 6** The 2-D projection.

#### 4.5 Properties of the Two-Dimensional Projection

As shown in Fig. 6, the 2-D projection  $\mathcal{P}_{\Pi_2}$  excludes several geometrical properties; e.g., shape, angles, lengths, and ratio of lengths. Fortunately, there are some geometrical properties that are preserved. Particularly, we are interested in three of them that are very useful in practice: namely straightness, line–line intersection, and parallelism of the normal and line at infinity vectors.

##### 4.5.1 Straightness property

This property states that a 2-D projection transforms lines to lines.<sup>12</sup> This can be shown as follows. Consider a line with vector  $\mathbf{l}$  and points  $\mathbf{x}$ , that is

$$0 = \mathbf{l}^T \mathcal{H}[\mathbf{x}]. \quad (56)$$

Next, the points  $\mathbf{x}$  are transformed to  $\boldsymbol{\alpha}$  by Eq. (55). Solving Eq. (55) for  $\mathbf{x}$  and substituting in Eq. (56), we obtain

$$0 = \mathbf{l}^T \mathcal{H}[\mathcal{P}_{\Pi_2^{-1}}[\boldsymbol{\alpha}]] = \frac{\mathbf{l}^T \Pi_2^{-1} \mathcal{H}[\boldsymbol{\alpha}]}{\mathcal{S}[\Pi_2^{-1} \mathcal{H}[\boldsymbol{\alpha}]]}, \quad (57)$$

or

$$0 = \mathbf{m}^T \mathcal{H}[\boldsymbol{\alpha}], \quad (58)$$

where

$$\mathbf{m} = \Pi_2^{-T} \mathbf{l}, \quad (59)$$

with  $\Pi_2^{-T}$  being the abbreviation of  $(\Pi_2^{-1})^T$  or  $(\Pi_2^T)^{-1}$ . In summary, the points  $\mathbf{x}$  of a line  $\mathbf{l}$  are transformed by  $\mathcal{P}_{\Pi_2}$  to points  $\boldsymbol{\alpha}$  of a new line  $\mathbf{m}$ .

##### 4.5.2 Line–line intersection

Preservation of the line–line intersection by a 2-D projection refers to the following. If

$$\mathbf{x}_0 = \mathcal{H}^{-1}[\mathbf{l}_1 \times \mathbf{l}_2] \quad (60)$$

is the point where the lines  $\mathbf{l}_1$  and  $\mathbf{l}_2$  intersect, then

$$\boldsymbol{\alpha}_0 = \mathcal{P}_{\Pi_2}[\mathbf{x}_0] \quad (61)$$

is the point where intersect the lines

$$\mathbf{m}_1 = \Pi_2^{-T} \mathbf{l}_1 \quad \text{and} \quad \mathbf{m}_2 = \Pi_2^{-T} \mathbf{l}_2. \quad (62)$$

In fact, the lines  $\mathbf{m}_1$  and  $\mathbf{m}_2$  intersect at the point

$$\begin{aligned} \boldsymbol{\alpha}_0 &= \mathcal{H}^{-1}[\mathbf{m}_1 \times \mathbf{m}_2] \\ &= \mathcal{H}^{-1}[(\Pi_2^{-T} \mathbf{l}_1) \times (\Pi_2^{-T} \mathbf{l}_2)] \\ &= \mathcal{H}^{-1}[\Pi_2(\mathbf{l}_1 \times \mathbf{l}_2)], \end{aligned} \quad (63)$$

where the identity of the cross product

$$(\mathbf{M}\mathbf{u}) \times (\mathbf{M}\mathbf{v}) = (\det \mathbf{M}) \mathbf{M}^{-T}(\mathbf{u} \times \mathbf{v}) \quad (64)$$

was applied. By solving Eq. (60) for  $\mathbf{l}_1 \times \mathbf{l}_2$  and substituting in Eq. (63), we obtain

$$\begin{aligned} \boldsymbol{\alpha}_0 &= \mathcal{H}^{-1}[\mathcal{S}[\mathbf{l}_1 \times \mathbf{l}_2] \Pi_2 \mathcal{H}[\mathbf{x}_0]] \\ &= \mathcal{H}^{-1}[\Pi_2 \mathcal{H}[\mathbf{x}_0]] \\ &= \mathcal{P}_{\Pi_2}[\mathbf{x}_0]. \end{aligned} \quad (65)$$

##### 4.5.3 Parallelism of the normal and line at infinity vectors

The normal of the  $xy$ -plane and the vector  $\mathbf{l}_\infty$  of the line at infinity are parallel. When the projection  $\mathcal{P}_{\Pi_2}$  is applied, the normal  $\mathbf{q}_3$  of the reference plane (with parameters  $\Pi_2$ ) and the new line at infinity  $\mathbf{m}_\infty$  still remain parallel; i.e.,  $\mathbf{m}_\infty = \lambda \mathbf{q}_3$ ,  $\lambda \neq 0$ . Actually, the reference plane has the normal

$$\mathbf{q}_3 = \mathbf{q}_1 \times \mathbf{q}_2, \quad (66)$$

see Eq. (54), whereas the vector of the new line at infinity is

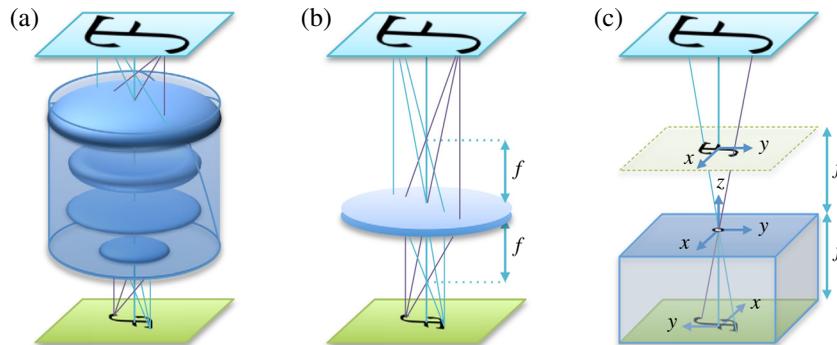
$$\begin{aligned} \mathbf{m}_\infty &= \Pi_2^{-T} \mathbf{l}_\infty \\ &= \lambda (\text{cof } \Pi_2) \mathbf{l}_\infty \\ &= \lambda [\mathbf{q}_2 \times \mathbf{s} \quad -\mathbf{q}_1 \times \mathbf{s} \quad \mathbf{q}_1 \times \mathbf{q}_2] \mathbf{l}_\infty \\ &= \lambda \mathbf{q}_3, \end{aligned} \quad (67)$$

where  $\lambda = 1/\det \Pi_2$ , and  $\text{cof}(\cdot)$  denotes the cofactor matrix.

In the following section, the developed theoretical framework is applied in a real problem.

## 5 Pinhole Camera Model

In practice, the imaging process is performed by a camera lens device as shown in Fig. 7(a). This device produces



**Fig. 7** (a) Illustration of a camera lens. (b) The imaging process modeled using a single thin lens. (c) A pinhole camera. The planes  $z = -f$  and  $z = f$  are the actual and conjugate image planes, respectively.

high quality images because of a complicated system of lenses that minimizes aberration and distortion. However, the imaging process can be modeled using a single thin lens as shown in Fig. 7(b). Moreover, the imaging model can be easily derived using the equivalent pinhole camera as shown in Fig. 7(c).

In the pinhole camera, the origin of a coordinate system is fixed at the pinhole and the  $z$ -axis is parallel to the optical axis. The plane  $z = -f$ , where  $f$  is the focal length, is the actual image plane. Note that the image is inverted; therefore, the  $x$  and  $y$  axes are reverted to describe the image as a magnified version of the object. The inversion of the axes is avoided using the conjugate image plane  $z = f$  as shown in Fig. 7(c).

### 5.1 Centered Pinhole Camera

A typical representation of a pinhole camera is shown in Fig. 8. The coordinate system  $O_c x_c y_c z_c$  is known as the camera reference frame. Let

$$\mathbf{p}_c = [x_c \quad y_c \quad z_c]^T \quad (68)$$

be the coordinates of a point in the camera reference frame. The point  $\mathbf{p}_c$  will be imaged in the plane  $z_c = f$  at the point

$$\boldsymbol{\beta} = [\beta_x \quad \beta_y]^T, \quad (69)$$

where  $\boldsymbol{\beta}$  will be referred to as the physical image coordinates. The pinhole projection model relates the vectors  $\mathbf{p}_c$  and  $\boldsymbol{\beta}$  by

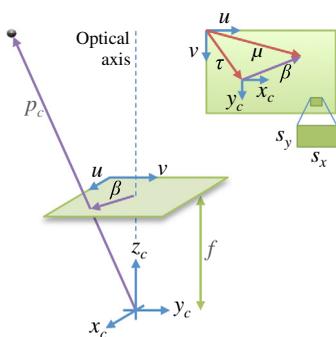
$$\begin{bmatrix} \beta_x \\ \beta_y \\ f \end{bmatrix} = \frac{f}{z_c} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix}. \quad (70)$$

Using homogeneous coordinates for the image point  $\boldsymbol{\beta}$ , Eq. (70) can be rewritten as

$$\begin{aligned} \boldsymbol{\beta} &= \mathcal{H}_f^{-1}[\mathbf{p}_c] \\ &= \mathcal{H}^{-1}[\Xi_f^{-1} \mathbf{p}_c]. \end{aligned} \quad (71)$$

The image formed in the sensor of the camera is sampled as an array of pixels. Then, the physical coordinates  $\boldsymbol{\beta}$  will be transformed to the pixel coordinates

$$\boldsymbol{\mu} = [u \quad v]^T, \quad (72)$$



**Fig. 8** The centered pinhole camera and sampling of the image plane.

which depend on the size of the pixel and skew (diagonal distortion) as shown in Fig. 8. The sampling can be described as

$$\begin{aligned} u &= (\beta_x + \tau_x)/s_x + \sigma\beta_y, \\ v &= (\beta_y + \tau_y)/s_y, \end{aligned} \quad (73)$$

where  $s_x$  and  $s_y$  (with units of length) are the width and height of the pixel, respectively,  $\boldsymbol{\tau} = [\tau_x, \tau_y]^T$  is known as the principal point and represents the point (from the  $uv$ -reference frame) where the optical axis crosses the image plane, and  $\sigma$  is the skew factor ( $\sigma = 0$  for most camera sensors). Equation (73) can be written as

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} 1/s_x & \sigma & \tau_x/s_x \\ 0 & 1/s_y & \tau_y/s_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_x \\ \beta_y \\ 1 \end{bmatrix}, \quad (74)$$

or using a compact notation

$$\boldsymbol{\mu} = \mathcal{H}^{-1}[\mathcal{S}\mathcal{H}[\boldsymbol{\beta}]] = \mathcal{P}_S[\boldsymbol{\beta}], \quad (75)$$

where  $S$  is the sampling matrix given as

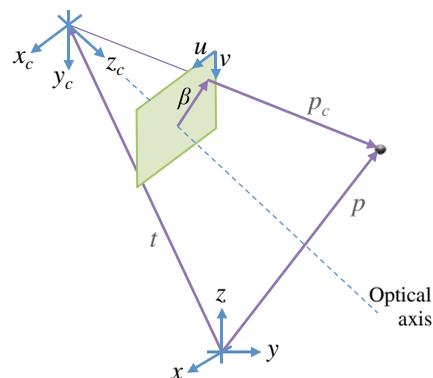
$$S = \begin{bmatrix} 1/s_x & \sigma & \tau_x/s_x \\ 0 & 1/s_y & \tau_y/s_y \\ 0 & 0 & 1 \end{bmatrix}. \quad (76)$$

Substituting Eq. (71) into Eq. (75), we obtain the image  $\boldsymbol{\mu}$  (in pixel coordinates) of the point  $\mathbf{p}_c$  as

$$\boldsymbol{\mu} = \mathcal{P}_S[\mathcal{H}^{-1}[\Xi_f^{-1} \mathbf{p}_c]] = \mathcal{H}^{-1}[\mathcal{S}\Xi_f^{-1} \mathbf{p}_c] = \mathcal{H}^{-1}[K\mathbf{p}_c], \quad (77)$$

where  $K = \mathcal{S}\Xi_f^{-1}$  is known as the matrix of intrinsic camera parameters having the explicit form

$$K = \begin{bmatrix} 1/s_x & \sigma & \tau_x/f s_x \\ 0 & 1/s_y & \tau_y/f s_y \\ 0 & 0 & 1/f \end{bmatrix}. \quad (78)$$



**Fig. 9** The noncentered pinhole camera.

Since  $\det K = (s_x s_y f)^{-1}$ , the matrix  $K$  is nonsingular for any experimental case.

Given a point  $\boldsymbol{\mu}$  (in pixel coordinates), the actual coordinates  $\mathcal{H}_f[\boldsymbol{\beta}]$  of an image point (physical coordinates on the image plane  $z = f$ ) can be obtained from Eq. (75) as

$$\begin{aligned}\mathcal{H}_f[\boldsymbol{\beta}] &= \mathcal{H}_f[\mathcal{P}_{S^{-1}}[\boldsymbol{\mu}]] \\ &= \Xi_f \mathcal{H}[\mathcal{P}_{S^{-1}}[\boldsymbol{\mu}]] \\ &= \Xi_f S^{-1} \mathcal{H}[\boldsymbol{\mu}] / \mathcal{S}[S^{-1} \mathcal{H}[\boldsymbol{\mu}]] \\ &= K^{-1} \mathcal{H}[\boldsymbol{\mu}],\end{aligned}\quad (79)$$

where the equality  $\mathcal{S}[S^{-1} \mathcal{H}[\boldsymbol{\mu}]] = 1$  was used.

### 5.2 Noncentered Pinhole Camera

Let us consider that the pinhole camera is at an arbitrary position and orientation with respect to a world coordinate system  $Oxyz$  as shown in Fig. 9. The position and orientation of the camera are defined by the vector  $\mathbf{t}$  and the rotation matrix  $R$ , respectively. Let

$$\mathbf{p} = [x \quad y \quad z]^T \quad (80)$$

be a point in the world coordinate system. Then, the point  $\mathbf{p}$  is seen from the camera reference frame as

$$\mathbf{p}_c = R^T(\mathbf{p} - \mathbf{t}) = L\mathcal{H}[\mathbf{p}], \quad (81)$$

where  $L$  is known as the matrix of extrinsic camera parameters having the explicit form

$$L = [R^T \quad -R^T \mathbf{t}]. \quad (82)$$

By substituting Eq. (81) into Eq. (77), the complete imaging process by a noncentered pinhole camera is given as

$$\boldsymbol{\mu} = \mathcal{H}^{-1}[KL\mathcal{H}[\mathbf{p}]] = \mathcal{H}^{-1}[C\mathcal{H}[\mathbf{p}]], \quad (83)$$

where  $C = KL$  is the matrix of the camera.

### 5.3 Homography Matrix

In general terms, Eq. (83) describes a transformation of points  $\mathbf{p}$  of the 3-D space to points of the 2-D one. A very useful transformation is obtained when  $\mathbf{p}$  represents points of a plane in the 3-D space. In this case, Eq. (83) is reduced to a transformation from the 2-D space to itself.

Consider that  $\mathbf{p}$  represents the points of a plane in the 3-D space; mathematically, see Eq. (53)

$$\mathbf{p} = \Pi \mathcal{H}[\boldsymbol{\rho}], \quad (84)$$

where  $\boldsymbol{\rho} = [\rho_x, \rho_y]^T$  parameterizes the plane,  $\Pi = [\mathbf{q}_1, \mathbf{q}_2, \mathbf{s}]$  is the matrix of the plane,  $\mathbf{q}_1$  and  $\mathbf{q}_2$  are columns of the rotation matrix  $Q = [\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3]^T$ , and  $\mathbf{s}$  is a translation vector. Next, the points  $\mathbf{p}$  are transformed to  $\boldsymbol{\mu}$  by Eq. (83) as

$$\begin{aligned}\boldsymbol{\mu} &= \mathcal{H}^{-1}[C\mathcal{H}[\Pi\mathcal{H}[\boldsymbol{\rho}]]] \\ &= \mathcal{H}^{-1}\left[C\begin{bmatrix} \Pi \\ \mathcal{H}[\mathbf{0}_2]^T \end{bmatrix}\mathcal{H}[\boldsymbol{\rho}]\right] \\ &= \mathcal{H}^{-1}[G\mathcal{H}[\boldsymbol{\rho}]] \\ &= \mathcal{P}_G[\boldsymbol{\rho}],\end{aligned}\quad (85)$$

where  $G$  is known as the homography matrix and has the explicit form

$$\begin{aligned}G &= KL\begin{bmatrix} \Pi \\ \mathcal{H}[\mathbf{0}_2]^T \end{bmatrix} \\ &= KR^T[\Pi \quad -\mathbf{t}\mathcal{H}[\mathbf{0}]^T] \\ &= KR^T\bar{\Pi},\end{aligned}\quad (86)$$

where

$$\begin{aligned}\bar{\Pi} &= [\Pi \quad -\mathbf{t}\mathcal{H}[\mathbf{0}]^T] \\ &= [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \mathbf{s} - \mathbf{t}].\end{aligned}\quad (87)$$

From Eqs. (79) and (85), a point  $\boldsymbol{\rho}$  is imaged at the point with actual image coordinates

$$\begin{aligned}\mathcal{H}_f[\boldsymbol{\beta}] &= K^{-1}\mathcal{H}[\boldsymbol{\mu}] \\ &= K^{-1}\mathcal{H}[\mathcal{P}_G[\boldsymbol{\rho}]] \\ &= K^{-1}G\mathcal{H}[\boldsymbol{\rho}]/\lambda \\ &= R^T\bar{\Pi}\mathcal{H}[\boldsymbol{\rho}]/\lambda,\end{aligned}\quad (88)$$

where  $\lambda = \mathcal{S}[G\mathcal{H}[\boldsymbol{\rho}]]$ .

The homography matrix is singular when the pinhole is at a point of the reference plane. For any other case,  $\det G = \mathbf{q}_3^T(\mathbf{s} - \mathbf{t})/(s_x s_y f)$  and Eq. (85) can be inverted as

$$\boldsymbol{\rho} = \mathcal{P}_{G^{-1}}[\boldsymbol{\mu}]. \quad (89)$$

The homography matrix is very useful for many computer vision tasks. In Appendix A, the direct linear transformation method for homography estimation is described.

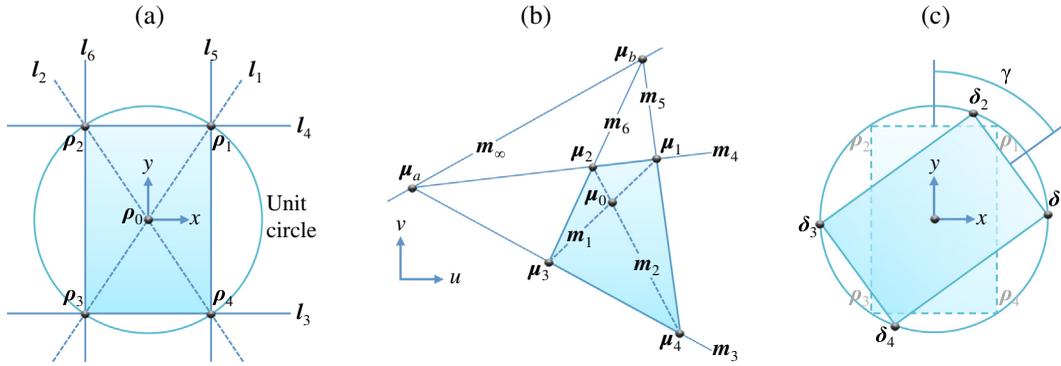
## 6 Perspective Correction for Document Scanning

A camera document scanning application performs several image processing tasks, such as quadrilateral detection, perspective correction, resampling, and image enhancement. In this section, the perspective correction task is addressed to illustrate the application of the proposed approach.

### 6.1 Assumptions

In Appendix A, we show that the perspective of a flat object can be easily corrected using the associated homography. For this, at least four correspondences  $(\boldsymbol{\mu}_k, \boldsymbol{\rho}_k)$  must be provided. However, for practical document scanning, the coordinates  $\boldsymbol{\rho}_k$  are unknown. Instead, it is assumed that the document to be digitized is rectangular and the orthogonality and parallelism properties of its edges are exploited.

The estimation of the homography is greatly simplified by assuming a centered pinhole camera with known intrinsic parameters; e.g., by a previous camera calibration, see Appendix B. Thus, we only require to estimate the reference



**Fig. 10** (a) Depict of the rectangular paper to be digitized. (b) The quadrilateral obtained by pinhole imaging of a rectangular paper. (c) The remaining rotation after perspective correction of the quadrilateral shown in (b).

plane parameters  $\Pi$ , i.e., the rotation matrix  $Q$  and the translation vector  $\mathbf{s}$ , see Eq. (84).

## 6.2 Estimation of the Reference Plane Parameters

Consider a coordinate system in the reference plane with origin at the center of the document to be scanned as shown in Fig. 10(a). The  $x$ - and  $y$ -axes of this coordinate system are parallel with the upper/lower and left/right sides of the paper, respectively. The corners of the document to be digitized have coordinates given by the vectors

$$\boldsymbol{\rho}_k, \quad k = 1, \dots, 4. \quad (90)$$

In this configuration, the vectors  $\boldsymbol{\rho}_k$  are symmetric about the  $y$ -axis; that is

$$\boldsymbol{\rho}_2 = T\boldsymbol{\rho}_1, \quad \boldsymbol{\rho}_4 = T\boldsymbol{\rho}_3, \quad (91)$$

where

$$T = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (92)$$

When the document is imaged by the camera, the original rectangle is transformed to a quadrilateral because of the perspective distortion. The corners of the imaged document have coordinates given by the vectors

$$\boldsymbol{\mu}_k, \quad k = 1, \dots, 4, \quad (93)$$

as shown in Fig. 10(b). The vectors  $\boldsymbol{\mu}_k$  and  $\boldsymbol{\rho}_k$  are related by Eqs. (85) and (89); however, the vectors  $\boldsymbol{\rho}_k$  and the homography  $G$  are unavailable. Only the vectors  $\boldsymbol{\mu}_k$  are available, which are easily obtained from the image by pointing the vertices of the imaged document.

The points  $\boldsymbol{\mu}_k$  are used to compute the following lines, see Fig. 10(b),

$$\begin{aligned} \mathbf{m}_1 &= \mathcal{H}[\boldsymbol{\mu}_3] \times \mathcal{H}[\boldsymbol{\mu}_1], & \mathbf{m}_4 &= \mathcal{H}[\boldsymbol{\mu}_1] \times \mathcal{H}[\boldsymbol{\mu}_2], \\ \mathbf{m}_2 &= \mathcal{H}[\boldsymbol{\mu}_2] \times \mathcal{H}[\boldsymbol{\mu}_4], & \mathbf{m}_5 &= \mathcal{H}[\boldsymbol{\mu}_4] \times \mathcal{H}[\boldsymbol{\mu}_1], \\ \mathbf{m}_3 &= \mathcal{H}[\boldsymbol{\mu}_3] \times \mathcal{H}[\boldsymbol{\mu}_4], & \mathbf{m}_6 &= \mathcal{H}[\boldsymbol{\mu}_2] \times \mathcal{H}[\boldsymbol{\mu}_3]. \end{aligned} \quad (94)$$

Next, with the lines  $\mathbf{m}_k$ , the following three intersection points are computed

$$\begin{aligned} \boldsymbol{\mu}_0 &= \mathcal{H}^{-1}[\mathbf{m}_1 \times \mathbf{m}_2], \\ \boldsymbol{\mu}_a &= \mathcal{H}^{-1}[\mathbf{m}_3 \times \mathbf{m}_4], \\ \boldsymbol{\mu}_b &= \mathcal{H}^{-1}[\mathbf{m}_5 \times \mathbf{m}_6]. \end{aligned} \quad (95)$$

Since the intrinsic camera parameters are assumed to be known, the actual image coordinates of the points  $\boldsymbol{\mu}_i$  can be obtained by Eq. (79) as

$$\mathcal{H}_f[\boldsymbol{\beta}_i] = K^{-1}\mathcal{H}[\boldsymbol{\mu}_i], \quad (96)$$

with  $i = a, b, 0, 1, 2, 3, 4$ .

### 6.2.1 Normal vector

Note that the points  $\boldsymbol{\mu}_a$  and  $\boldsymbol{\mu}_b$  are the projections of the ideal points

$$\mathcal{H}[\boldsymbol{\rho}_a] = [1 \ 0 \ 0]^T, \quad \mathcal{H}[\boldsymbol{\rho}_b] = [0 \ 1 \ 0]^T, \quad (97)$$

respectively. Thus, the line  $\mathbf{m}_\infty$  (in physical coordinates on the image plane  $z = f$ ) is parallel to the normal  $\mathbf{q}_3$ , see Eq. (67). That is,

$$\mathbf{n} = \mathcal{H}_f[\boldsymbol{\beta}_a] \times \mathcal{H}_f[\boldsymbol{\beta}_b] = \mathbf{q}_3/\lambda, \quad (98)$$

for some  $\lambda$ . Thus, the normal of the reference plane is obtained as the normalization of the vector  $\mathbf{n}$ , namely

$$\mathbf{q}_3 = \mathbf{n}/\|\mathbf{n}\| = \frac{K^T[(\mathbf{m}_3 \times \mathbf{m}_4) \times (\mathbf{m}_5 \times \mathbf{m}_6)]}{\|K^T[(\mathbf{m}_3 \times \mathbf{m}_4) \times (\mathbf{m}_5 \times \mathbf{m}_6)]\|}. \quad (99)$$

### 6.2.2 Translation vector

The translation vector  $\mathbf{s}$  is obtained by taking into account that  $\mathcal{P}_G$  preserves the line–line intersection. Thus, from Eq. (65) we have  $\boldsymbol{\mu}_0 = \mathcal{P}_G[\boldsymbol{\rho}_0]$  with  $\boldsymbol{\rho}_0 = \mathbf{0}_2$ . Therefore, from Eq. (88) we have

$$\mathcal{H}_f[\boldsymbol{\beta}_0] = \bar{\Pi}\mathcal{H}[\mathbf{0}_2]/\xi = \mathbf{s}/\xi, \quad (100)$$

where  $\xi$  is a scalar to be determined. For this, note that the vectors

$$\mathbf{p}_k = \zeta_k \mathcal{H}_f[\boldsymbol{\beta}_k], \quad k = 1, \dots, 4, \quad (101)$$

are points of the reference for values  $\zeta_k$  and  $\xi$  such that the equation of the plane  $\mathbf{q}_3(\mathbf{p}_k - \mathbf{s}) = 0$  is satisfied. This leads to

$$\zeta_k = \xi \frac{\mathbf{q}_3^T \mathcal{H}_f[\boldsymbol{\beta}_0]}{\mathbf{q}_3^T \mathcal{H}_f[\boldsymbol{\beta}_k]}. \quad (102)$$

Since the points  $\boldsymbol{\rho}_k$  are on the unit circumference, see Fig. 10(a), then  $\|\boldsymbol{\rho}_k\| = \|\mathbf{p}_k - \mathbf{s}\| = 1$ , which leads to

$$\xi_k = \left\| \frac{\mathbf{q}_3^T \mathcal{H}_f[\boldsymbol{\beta}_0]}{\mathbf{q}_3^T \mathcal{H}_f[\boldsymbol{\beta}_k]} \mathcal{H}_f[\boldsymbol{\beta}_k] - \mathcal{H}_f[\boldsymbol{\beta}_0] \right\|^{-1}, \quad (103)$$

where the subindex  $k$  in  $\xi$  emphasizes the fact that a different value could be obtained for each  $\mathcal{H}_f[\boldsymbol{\beta}_k]$  due to inaccuracies of  $\boldsymbol{\mu}_k$ . Therefore, the value  $\xi$  is computed as

$$\xi = \text{mean}\{\xi_1, \xi_2, \xi_3, \xi_4\}. \quad (104)$$

The result is used in Eq. (100), and the translation of the reference plane is now available.

### 6.2.3 Euler angles

The reference plane is fully characterized by six degrees of freedom (DOF), namely position (three coordinates) and orientation (three angles). The vectors  $\mathbf{q}_3$  and  $\mathbf{s}$  provide five DOFs. Specifically, the vector  $\mathbf{s}$  provides three DOFs that fix the position while  $\mathbf{q}_3$  provides two DOFs defining the orientation by the azimuth and polar angles given, respectively, by

$$\tan \phi = q_{23}/q_{13}, \quad \cos \theta = q_{33}, \quad (105)$$

where  $[q_{13}, q_{23}, q_{33}]^T = \mathbf{q}_3$  is the third column of the rotation matrix  $Q$ . The remaining angle  $\gamma$  (the angle around the normal  $\mathbf{q}_3$ ) can be obtained as follows.

From Eqs. (84) and (53), we have

$$\mathbf{p}_k = Q\mathcal{H}_0[\boldsymbol{\rho}_k] + \mathbf{s}, \quad (106)$$

where the matrix  $Q$  is defined as the Euler sequence

$$Q = Q_z(\phi)Q_y(\theta)Q_z(\gamma), \quad (107)$$

with  $Q_z$  and  $Q_y$  being the rotation matrices around the  $z$ - and  $y$ -axes, respectively. Thus, using Eq. (101), the estimated vector  $\mathbf{s}$ , and the angles  $\theta$  and  $\phi$ , we compute the (perspective corrected) points

$$\boldsymbol{\delta}_k = \mathcal{H}_0^{-1}[Q_y^T(\theta)Q_z^T(\phi)^T(\mathbf{p}_k - \mathbf{s})] = \begin{bmatrix} \delta_{xk} \\ \delta_{yk} \end{bmatrix}, \quad (108)$$

with  $k = 1, \dots, 4$ , see Fig. 10(c). The vectors  $\boldsymbol{\delta}_k$  and  $\boldsymbol{\rho}_k$  are related by

$$\boldsymbol{\rho}_k = \bar{Q}_z^T(\gamma)\boldsymbol{\delta}_k, \quad (109)$$

where

$$\bar{Q}_z^T(\gamma) = \begin{bmatrix} \cos \gamma & \sin \gamma \\ -\sin \gamma & \cos \gamma \end{bmatrix}. \quad (110)$$

The vectors  $\boldsymbol{\rho}_k$  are unavailable, but we use their symmetry properties given in Eq. (91) to obtain

$$\bar{Q}_z^T(\gamma)\boldsymbol{\delta}_2 = T\bar{Q}_z^T(\gamma)\boldsymbol{\delta}_1, \quad \bar{Q}_z^T(\gamma)\boldsymbol{\delta}_4 = T\bar{Q}_z^T(\gamma)\boldsymbol{\delta}_3. \quad (111)$$

The product  $\bar{Q}_z^T(\gamma)\boldsymbol{\delta}_k$  can be written as

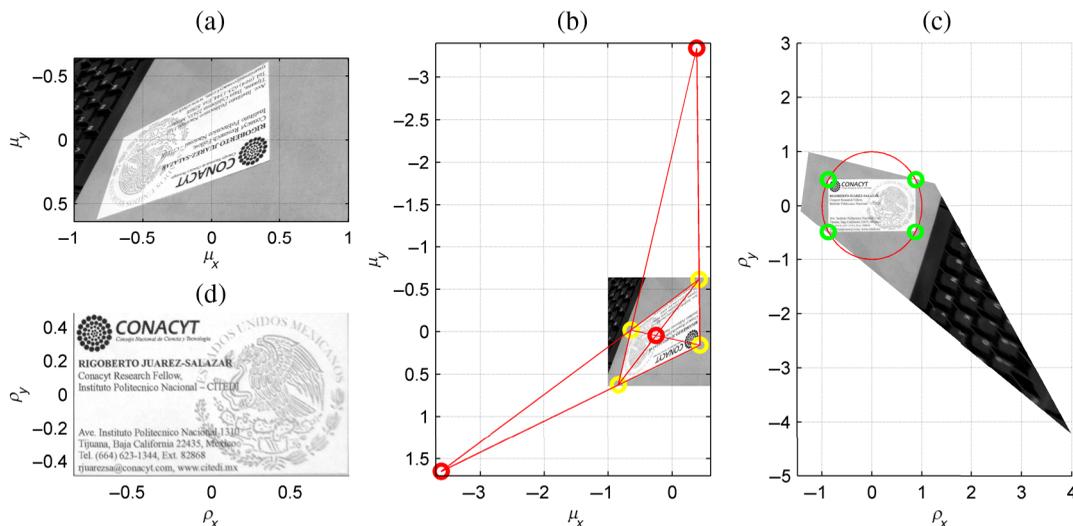
$$\bar{R}^T\boldsymbol{\delta}_k = \mathcal{B}[\boldsymbol{\delta}_k]\Gamma, \quad (112)$$

where  $\Gamma = [\sin \gamma, \cos \gamma]^T$  and

$$\mathcal{B}[\boldsymbol{\delta}_k] = \begin{bmatrix} \delta_{yk} & \delta_{xk} \\ -\delta_{xk} & \delta_{yk} \end{bmatrix}. \quad (113)$$

Thus, Eq. (111) can be rewritten as

$$\mathbb{B}\Gamma = \mathbf{0}_4, \quad (114)$$



**Fig. 11** (a) An input image with a rectangular object in scene. (b) The corners of the rectangle are marked by yellow circles. (c) Corrected image. (d) Zoom of (c) highlighting the region of interest.

where

$$\mathbb{B} = \begin{bmatrix} \mathcal{B}[\delta_2] - T\mathcal{B}[\delta_1] \\ \mathcal{B}[\delta_4] - T\mathcal{B}[\delta_3] \end{bmatrix}. \quad (115)$$

The nontrivial solution of Eq. (114) for  $\Gamma$  is obtained as the right-singular vector corresponding to the smallest singular value of  $\mathbb{B}$ . Finally, the angle  $\gamma$  is obtained from  $\Gamma$  by

$$\tan \gamma = \mathcal{H}^{-1}[\Gamma]. \quad (116)$$

The estimated parameters are used to create the matrix  $\Pi$ . Then, the required homography  $G$  is obtained by Eq. (86) (with  $\mathbf{t} = \mathbf{0}_3$  and  $R = \mathbb{I}$  because of the centered pinhole camera configuration). Finally, the perspective distortion of the image is corrected by displaying the intensity of each pixel of the image at the point  $\boldsymbol{\rho}$  computed by Eq. (89).

### 6.3 Illustrative Example

The functionality of the presented algorithm is illustrated by the following example. The camera described in Appendix B and the estimated intrinsic parameters  $K$  given in Eq. (156) are used here.

Figure 11(a) shows the image of a rectangular object acquired by the camera. Then, the four corners of the quadrilateral are marked from the image as shown by the yellow circles in Fig. 11(b). The points  $\boldsymbol{\mu}_0$ ,  $\boldsymbol{\mu}_a$ , and  $\boldsymbol{\mu}_b$  are indicated by the red circles in Fig. 11(b). It is worth mentioning that  $\boldsymbol{\mu}_a$ , or  $\boldsymbol{\mu}_b$ , or both could be points at infinity. Even in these cases, the presented methodology is valid.

The information estimated with the four corners are

$$\begin{aligned} \mathbf{s} &= [-0.2289 \quad 0.0561 \quad 2.9236]^T, \\ \phi &= 0.6776, \\ \theta &= 0.9879, \\ \gamma &= 2.3041. \end{aligned} \quad (117)$$

With these parameters, the matrix of the reference plane is

$$\Pi = \begin{bmatrix} -0.7528 & 0.1010 & -0.2289 \\ 0.3478 & -0.7779 & 0.0561 \\ 0.5588 & 0.6202 & 2.9236 \end{bmatrix}. \quad (118)$$

Thus, the resulting homography is

$$G = \begin{bmatrix} -2.0219 & 0.2429 & -0.7272 \\ 0.9237 & -2.0786 & 0.1298 \\ 0.5588 & 0.6202 & 2.9236 \end{bmatrix}. \quad (119)$$

All points  $\boldsymbol{\mu}$  of the image are transformed to points  $\boldsymbol{\rho}$  of the reference plane by Eq. (89). Next, the pixels of the image are displayed at the points  $\boldsymbol{\rho}$  as shown in Fig. 11(c).

With the correction of perspective, the yellow circles in Fig. 11(b) become the green ones in Fig. 11(c). The region of interest is the rectangle with corners marked by green circles in Fig. 11(c). Finally, a zoom of the region of interest is shown in Fig. 11(d).

## 7 Conclusions

An operator-based approach for homogeneous coordinates was proposed. Several basic geometrical concepts and properties of the operators were investigated. With the proposed approach, the pinhole camera model and a simple camera calibration method were described. The study of this work was motivated by developing a perspective correction method useful for a camera document scanning application. Several experimental results illustrate the analyzed theoretical aspects. The proposed approach could be a good starting point to introduce inexperienced students in the scientific discipline of computer vision.

## Appendix A: Estimation of the Homography Matrix

In this appendix, we illustrate the method known as direct linear transformation for homography matrix estimation. This method is very useful for illustration purposes because of its simplicity. However, the highest accuracy and robustness are reached with other advanced methods available in the literature.<sup>9,13</sup>

Let  $G$  be the homography matrix defined in Eq. (86). Consider that the matrix  $G$  is row partitioned as follows:

$$G = \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{g}}_1^T \\ \bar{\mathbf{g}}_2^T \\ \bar{\mathbf{g}}_3^T \end{bmatrix}, \quad (120)$$

where

$$\begin{aligned} \bar{\mathbf{g}}_1^T &= [g_{11} \quad g_{12} \quad g_{13}], \\ \bar{\mathbf{g}}_2^T &= [g_{21} \quad g_{22} \quad g_{23}], \\ \bar{\mathbf{g}}_3^T &= [g_{31} \quad g_{32} \quad g_{33}]. \end{aligned} \quad (121)$$

Equation (85), which relates points of the reference and image planes, can be rewritten as

$$\boldsymbol{\mu} = \begin{bmatrix} u \\ v \end{bmatrix} = \frac{1}{\bar{\mathbf{g}}_3^T \mathcal{H}[\boldsymbol{\rho}]} \begin{bmatrix} \bar{\mathbf{g}}_1^T \mathcal{H}[\boldsymbol{\rho}] \\ \bar{\mathbf{g}}_2^T \mathcal{H}[\boldsymbol{\rho}] \end{bmatrix}, \quad (122)$$

or

$$\begin{bmatrix} u \bar{\mathbf{g}}_3^T \mathcal{H}[\boldsymbol{\rho}] \\ v \bar{\mathbf{g}}_3^T \mathcal{H}[\boldsymbol{\rho}] \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{g}}_1^T \mathcal{H}[\boldsymbol{\rho}] \\ \bar{\mathbf{g}}_2^T \mathcal{H}[\boldsymbol{\rho}] \end{bmatrix}. \quad (123)$$

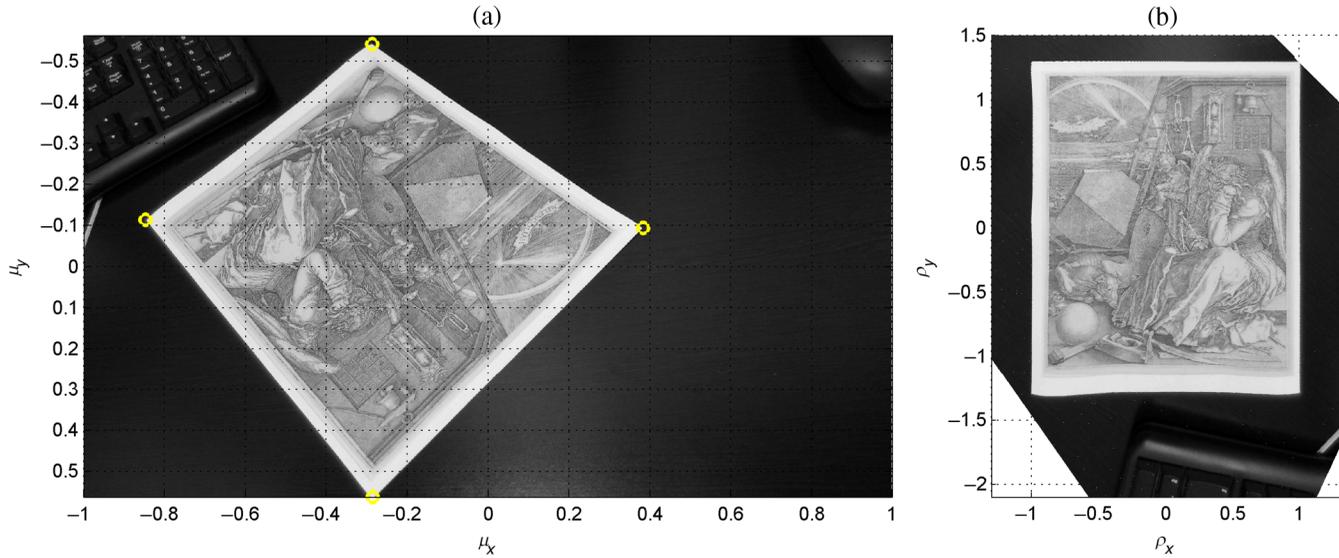
Furthermore, Eq. (123) can be written in matrix form as

$$A \bar{\mathbf{g}} = \mathbf{0}_2, \quad (124)$$

where

$$\begin{aligned} A &= \begin{bmatrix} \mathcal{H}[\boldsymbol{\rho}]^T & \mathbf{0}_3^T & -u \mathcal{H}[\boldsymbol{\rho}]^T \\ \mathbf{0}_3^T & \mathcal{H}[\boldsymbol{\rho}]^T & -v \mathcal{H}[\boldsymbol{\rho}]^T \end{bmatrix}, \\ \bar{\mathbf{g}} &= [\bar{\mathbf{g}}_1^T \quad \bar{\mathbf{g}}_2^T \quad \bar{\mathbf{g}}_3^T]^T. \end{aligned} \quad (125)$$

Equation (124) relates a single point  $\boldsymbol{\rho}$  on the reference plane with the corresponding point  $\boldsymbol{\mu}$  on the image plane.



**Fig. 12** (a) Image of  $1168 \times 2080$  pixel capturing a scene with the *Melencolia I* printed on a letter size paper. (b) Perspective corrected Image.

If  $n$  pairs  $(\boldsymbol{\rho}_k, \boldsymbol{\mu}_k)$ , with  $k = 1, 2, \dots, n$ , are available, the  $n$  corresponding equations of the form Eq. (124) can be written as

$$\mathbb{A}\bar{\mathbf{g}} = \mathbf{0}_{2n}, \quad (126)$$

where

$$\mathbb{A} = [A_1^T \ A_2^T \ \dots \ A_n^T]^T, \quad (127)$$

$$A_k = \begin{bmatrix} \mathcal{H}[\boldsymbol{\rho}_k]^T & \mathbf{0}_3^T & -u_k \mathcal{H}[\boldsymbol{\rho}_k]^T \\ \mathbf{0}_3^T & \mathcal{H}[\boldsymbol{\rho}_k]^T & -v_k \mathcal{H}[\boldsymbol{\rho}_k]^T \end{bmatrix}.$$

The nontrivial solution  $\bar{\mathbf{g}}$  of Eq. (126) can be obtained using the constraint  $\|\bar{\mathbf{g}}\| = 1$ . Thus, by using the singular value decomposition of  $\mathbb{A}$ , the solution for  $\bar{\mathbf{g}}$  is the right-singular vector corresponding to the smallest singular value of  $\mathbb{A}$ , see Appendix C of Ref. 14.

The application of this method is illustrated as follows. Consider the image shown in Fig. 12(a). A letter size paper printed with the *Melencolia I* by Albrecht Dürer is in the scene. Using the aspect ratio 1:1.2941 of the letter paper, the coordinates of the corners are fixed to

$$\begin{aligned} \boldsymbol{\rho}_1 &= [1, 1.2941]^T, & \boldsymbol{\rho}_3 &= -\boldsymbol{\rho}_1, \\ \boldsymbol{\rho}_2 &= [-1, 1.2941]^T, & \boldsymbol{\rho}_4 &= -\boldsymbol{\rho}_2. \end{aligned} \quad (128)$$

The coordinates of the imaged corners are

$$\begin{aligned} \boldsymbol{\mu}_1 &= [-0.2858, 0.5661]^T, \\ \boldsymbol{\mu}_2 &= [0.3826, -0.0938]^T, \\ \boldsymbol{\mu}_3 &= [-0.2884, -0.5403]^T, \\ \boldsymbol{\mu}_4 &= [-0.8479, -0.1135]^T, \end{aligned} \quad (129)$$

see yellow circles in Fig. 12(a). With these four pairs  $(\boldsymbol{\rho}_k, \boldsymbol{\mu}_k)$ , we obtain the homography

$$G = \begin{bmatrix} -0.2437 & 0.2292 & -0.2442 \\ 0.2258 & 0.1870 & -0.0888 \\ -0.0524 & -0.0989 & 0.8497 \end{bmatrix}. \quad (130)$$

The homography  $G$  fully defines a pinhole imaging process. Thus, it can be inverted to obtain an undistorted view of the reference plane from its perspective distorted image. Specifically, using Eq. (89) all points  $\boldsymbol{\mu}$  of the image are transformed to points  $\boldsymbol{\rho}$  of the reference plane. Then, the pixels of the image are displayed at the points  $\boldsymbol{\rho}$  as shown in Fig. 12(b). Note that corners of the paper in the corrected image are at the coordinates specified by Eq. (128).

The least number of point correspondences for two-dimensional homography estimation is four. However, the accuracy of the estimation is improved when more than four point correspondences are provided. For this reason, checkerboard patterns<sup>15</sup> and gratings<sup>16,17</sup> are useful target objects. In this appendix, the corner points of the imaged rectangle where obtained manually from the image. However, the corner points can be obtained automatically using checkerboard patterns or gratings along with grid detection<sup>18</sup> or phase demodulation,<sup>19</sup> respectively.

## Appendix B: Camera Parameters from Homographies

The homography matrix involves both intrinsic  $K$  and extrinsic  $L$  camera parameters as well as the reference plane parameters  $\Pi$ . In this appendix, we show how to obtain the intrinsic and extrinsic camera parameters from several homographies.

### B.1 Intrinsic Camera Parameters

Consider that the reference plane is the  $xy$ -plane of the world coordinate system; i.e.,  $\mathbf{s} = \mathbf{0}_3$  and

$$\bar{Q} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}. \quad (131)$$

In this case, the homography  $G$ , defined in Eq. (86), is reduced to

$$G = K[\bar{\mathbf{r}}_1 \quad \bar{\mathbf{r}}_2 \quad -R^T \mathbf{t}], \quad (132)$$

where  $\bar{\mathbf{r}}_1^T$  and  $\bar{\mathbf{r}}_2^T$  are the first and second rows of the rotation matrix  $R$ , respectively. Consider that the matrix  $G$  is column partitioned as follows:

$$G = \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{bmatrix} = [\mathbf{g}_1 \quad \mathbf{g}_2 \quad \mathbf{g}_3], \quad (133)$$

where

$$\mathbf{g}_1 = \begin{bmatrix} g_{11} \\ g_{21} \\ g_{31} \end{bmatrix}, \quad \mathbf{g}_2 = \begin{bmatrix} g_{12} \\ g_{22} \\ g_{32} \end{bmatrix}, \quad \mathbf{g}_3 = \begin{bmatrix} g_{13} \\ g_{23} \\ g_{33} \end{bmatrix}. \quad (134)$$

Thus, Eq. (132) can be written as

$$[\bar{\mathbf{r}}_1 \quad \bar{\mathbf{r}}_2 \quad -R^T \mathbf{t}] = K^{-1}[\mathbf{g}_1 \quad \mathbf{g}_2 \quad \mathbf{g}_3]. \quad (135)$$

Since  $\bar{\mathbf{r}}_1$  and  $\bar{\mathbf{r}}_2$  are orthonormal vectors ( $\bar{\mathbf{r}}_1$  and  $\bar{\mathbf{r}}_2$  are rows of a rotation matrix), we have the following two constraints  $\bar{\mathbf{r}}_1^T \bar{\mathbf{r}}_2 = 0$  and  $\|\bar{\mathbf{r}}_1\|^2 = \|\bar{\mathbf{r}}_2\|^2$ , which can be written as

$$\mathbf{g}_1^T W \mathbf{g}_2 = 0, \quad \mathbf{g}_1^T W \mathbf{g}_1 = \mathbf{g}_2^T W \mathbf{g}_2, \quad (136)$$

where the symmetric matrix  $W$  is defined as

$$W = K^{-T} K^{-1} = \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{12} & w_{22} & w_{23} \\ w_{13} & w_{23} & w_{33} \end{bmatrix}. \quad (137)$$

The bilinear form  $\mathbf{g}_i^T W \mathbf{g}_j$  can be rewritten as

$$\mathbf{g}_i^T W \mathbf{g}_j = \mathcal{V}_{ij}[G] \mathbf{w}, \quad (138)$$

where

$$\mathcal{V}_{ij}[G] = \begin{bmatrix} g_{1i}g_{1j} \\ g_{2i}g_{2j} \\ g_{3i}g_{3j} \\ g_{2i}g_{1j} + g_{1i}g_{2j} \\ g_{3i}g_{1j} + g_{1i}g_{3j} \\ g_{3i}g_{2j} + g_{2i}g_{3j} \end{bmatrix}^T, \quad (139)$$

and

$$\mathbf{w} = [w_{11} \quad w_{22} \quad w_{33} \quad w_{12} \quad w_{13} \quad w_{23}]^T. \quad (140)$$

Then, the constraints given by Eq. (136) become

$$V[G] \mathbf{w} = \mathbf{0}_2, \quad (141)$$

where  $V[G]$  is the following  $2 \times 6$  matrix:

$$V[G] = \begin{bmatrix} \mathcal{V}_{12}[G] \\ \mathcal{V}_{11}[G] - \mathcal{V}_{22}[G] \end{bmatrix}. \quad (142)$$

A nontrivial solution of Eq. (141) for  $\mathbf{w}$  can be obtained using several homographies  $G_k, k = 1, 2, \dots, m$ . For this, we compute the homographies of different images where the position and orientation of the reference plane (or the camera, or both) are varying (in an unknown manner) while the intrinsic camera parameters remain constant. Thus, we solve the new matrix equation

$$\mathbb{V} \mathbf{w} = \mathbf{0}_{2m}, \quad (143)$$

where

$$\mathbb{V} = [V[G_1]^T \quad V[G_2]^T \quad \dots \quad V[G_m]^T]^T. \quad (144)$$

In general, at least three homographies ( $m = 3$ ) are required. However, two homographies are sufficient assuming zero-skew.

Equation (143) can be solved for  $\mathbf{w}$  using the singular value decomposition method, see Appendix C of Ref. 14. Since the obtained solution, labeled as  $\tilde{\mathbf{w}}$ , is unique up to scale, the associated matrix  $\tilde{W}$  is related to  $W$  by

$$\tilde{W} = \lambda W = \lambda K^{-T} K^{-1}, \quad (145)$$

where  $\lambda \neq 0$  is an unknown constant. With the estimated matrix  $\tilde{W}$ , the unknown scalar  $\lambda$  and the entries  $k_{ij}$  of the intrinsic parameter matrix

$$K = \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ 0 & k_{22} & k_{23} \\ 0 & 0 & 1 \end{bmatrix}, \quad (146)$$

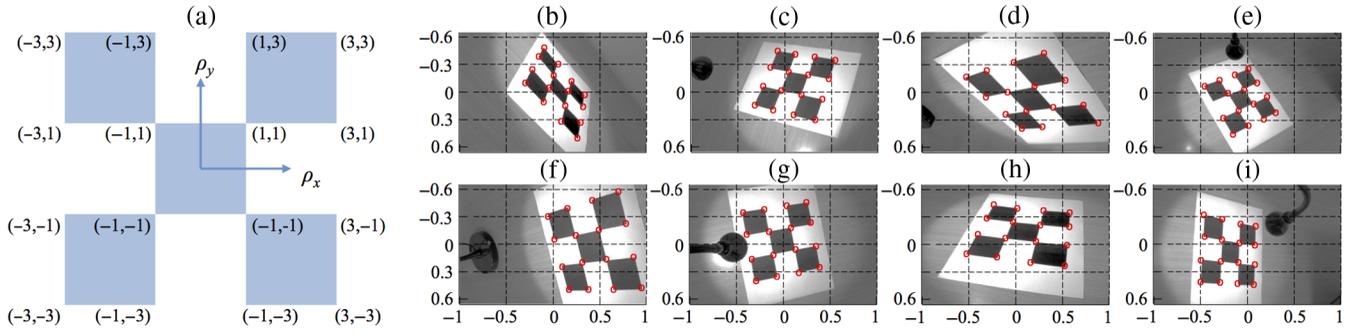
are given in closed form as

$$\begin{aligned} \lambda &= (\det \tilde{W})/d, \\ k_{11} &= \sqrt{\lambda/\tilde{w}_{11}}, \\ k_{22} &= \sqrt{\lambda\tilde{w}_{11}/d}, \\ k_{12} &= -\tilde{w}_{12}\sqrt{\lambda/\tilde{w}_{11}d}, \\ k_{13} &= (\tilde{w}_{12}\tilde{w}_{23} - \tilde{w}_{22}\tilde{w}_{13})/d, \\ k_{23} &= (\tilde{w}_{12}\tilde{w}_{13} - \tilde{w}_{11}\tilde{w}_{23})/d, \end{aligned} \quad (147)$$

where  $d = \tilde{w}_{11}\tilde{w}_{22} - \tilde{w}_{12}^2$ .

It is worth mentioning that the intrinsic camera parameters ( $f, s_x, s_y, \tau_x, \tau_y$ , and  $\sigma$ ) cannot be obtained using only the matrix  $K$ . Fortunately, the matrix  $K$  is sufficient for many computer vision tasks. For the case where the intrinsic camera parameters are required explicitly, we can assume that the skew and size of the pixel are known (e.g.,  $s_x, s_y$ , and  $\sigma$  are consulted in the datasheet of the camera sensor). Thus, the estimation of the remaining intrinsic parameters is a linear problem with the least-squares solution

$$f = s_x s_y \frac{s_x k_{22} + s_y k_{11} + s_x s_y \sigma k_{12}}{s_x^2 + s_y^2 + s_x^2 s_y^2 \sigma^2}, \quad (148)$$



**Fig. 13** (a) A 3 × 3 checkerboard pattern and the coordinates of the corners. (b)–(i) 8 of 15 images of the reference plane acquired by a camera from different unknown viewpoints.

$$\tau_x = s_x k_{13}, \quad (149)$$

$$\tau_y = s_y k_{23}. \quad (150)$$

## B.2 Extrinsic Camera Parameters

Once the matrix  $K$  is available, the rotation matrix  $R$  and the translation vector  $\mathbf{t}$  can be estimated for each provided homography as follows. First, we compute the estimate  $\tilde{R}^T$  of the matrix  $R^T$  as

$$\tilde{R}^T = [\mathbf{h}_1 \quad \mathbf{h}_2 \quad \mathbf{h}_1 \times \mathbf{h}_2], \quad (151)$$

where using Eq. (135), the vectors  $\mathbf{h}_1$  and  $\mathbf{h}_2$  are given as

$$\mathbf{h}_1 = K^{-1} \mathbf{g}_1, \quad \mathbf{h}_2 = K^{-1} \mathbf{g}_2. \quad (152)$$

Then, the rotation matrix  $R$  is obtained from  $\tilde{R}$  ensuring the orthogonality condition of rotation matrices. For this, the singular value decomposition  $\tilde{R} = U\Sigma V^T$  is obtained and the required rotation matrix is determined as

$$R = UV^T. \quad (153)$$

Finally, the translation vector  $\mathbf{t}$  is computed as

$$\mathbf{t} = -RK^{-1} \mathbf{g}_3. \quad (154)$$

## B.3 Illustrative Example

As an example, we describe a simple experiment to obtain the intrinsic parameters of a camera. A camera with a pixel size of  $6 \mu\text{m}$  (square pixel), resolution of  $752 \times 480$  pixel, and imaging lens with focal length of 6 mm was used. The  $3 \times 3$  checkerboard pattern shown in Fig. 13(a) was printed on a letter paper. Then, 15 images of the printed pattern lying on the reference plane were captured from different unknown viewpoints, see Figs. 13(b)–13(i).

We use the coordinates of the corners shown in Fig. 13(a) as the known points  $\boldsymbol{\rho}_k$  on the reference plane. The corresponding points  $\boldsymbol{\mu}_k$  in the image plane were obtained by marking the corners of the checkerboard pattern in the image. Then, with the pairs  $(\boldsymbol{\rho}_k, \boldsymbol{\mu}_k)$ , an homography matrix  $G_k$  was computed for each acquired image. With these homographies, the matrix  $\mathbb{V}$  defined in Eq. (144) was created. Then, Eq. (143) was solved for  $\mathbf{w}$ , the resulting matrix  $\tilde{W}$  is

$$\tilde{W} = \begin{bmatrix} -0.1389 & 0.0005 & -0.0058 \\ 0.0005 & -0.1378 & -0.0008 \\ -0.0058 & -0.0008 & -0.9806 \end{bmatrix}. \quad (155)$$

From this, the intrinsic parameter matrix  $K$  was recovered as

$$K = \begin{bmatrix} 2.6563 & -0.0103 & -0.0419 \\ 0 & 2.6674 & -0.0059 \\ 0 & 0 & 1 \end{bmatrix}. \quad (156)$$

For validation purposes, we estimate the focal length using the known information  $s_x = s_y = 6 \mu\text{m}$ , and  $\sigma = 0$ . The reader should note that the quantities  $s_x$  and  $s_y$  are defined in this experiment as

$$s_x = s_y = \frac{752}{2} 6 \times 10^{-3} \text{ mm} \quad (157)$$

because the points  $\boldsymbol{\mu}$  were obtained in a coordinate system with a unit of length equal to a half of the image width, see Figs. 13(b)–13(i). From the matrix  $K$  in Eq. (156), the focal length was estimated using Eq. (148). The result is  $f = 6.0032$  mm, which is very close to the nominal focal length (6 mm) of the employed camera lens.

## Acknowledgments

This work was supported by CONACyT México through the project Cátedras/880.

## References

1. O. Faugeras, Q.-T. Luong, and T. Papadopoulos, *The Geometry of Multiple Images: the Laws that Govern the Formation of Multiple Images of a Scene and Some of Their Applications*, MIT Press, Cambridge (2004).
2. O. D. Faugeras and S. Maybank, "Motion from point matches: multiplicity of solutions," *Int. J. Comput. Vision* **4**(3), 225–246 (1990).
3. Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.* **22**, 1330–1334 (2000).
4. Y. Zhao and Y. Li, "Camera self-calibration from projection silhouettes of an object in double planar mirrors," *J. Opt. Soc. Am. A* **34**, 696–707 (2017).
5. T. Taketomi et al., "Camera pose estimation under dynamic intrinsic parameter change for augmented reality," *Comput. Graphics* **44**, 11–19 (2014).
6. H. H. Ip and Y. Chen, "Planar rectification by solving the intersection of two circles under 2D homography," *Pattern Recognit.* **38**(7), 1117–1120 (2005).
7. B. Cyganek and J. P. Siebert, *An Introduction to 3D Computer Vision Techniques and Algorithms*, John Wiley & Sons Ltd., Chichester, West Sussex (2009).

8. O. Faugeras, *Three-Dimensional Computer Vision: a Geometric Viewpoint*, Artificial Intelligence, MIT Press, Cambridge (1993).
9. R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, Cambridge (2003).
10. J. L. Mundy and A. Zisserman, *Appendix—Projective Geometry for Machine Vision*, pp. 463–519, MIT Press, Cambridge (1992).
11. W. Burger, “Zhang’s camera calibration algorithm: in-depth tutorial and implementation,” Technical Report HGB16-05, University of Applied Sciences Upper Austria, School of Informatics, Communications and Media, Department of Digital Media, Hagenberg, Austria (2016).
12. F. Devernay and O. Faugeras, “Straight lines have to be straight,” *Mach. Vision Appl.* **13**(1), 14–24 (2001).
13. H. Zeng, X. Deng, and Z. Hu, “A new normalized method on line-based homography estimation,” *Pattern Recognit. Lett.* **29**(9), 1236–1244 (2008).
14. Z. Zhang, “A flexible new technique for camera calibration,” Technical Report MSR-TR-98-71, Microsoft Research (1998).
15. L. Kr “Accurate chequerboard corner localisation for camera calibration,” *Pattern Recognit. Lett.* **32**(10), 1428–1435 (2011).
16. R. Juarez-Salazar et al., “Camera calibration by multiplexed phase encoding of coordinate information,” *Appl. Opt.* **54**, 4895–4906 (2015).
17. R. Juarez-Salazar, L. N. Gaxiola, and V. H. Diaz-Ramirez, “Single-shot camera position estimation by crossed grating imaging,” *Opt. Commun.* **382**, 585–594 (2017).
18. A. Herout, M. Dubska, and J. Havel, *Vanishing Points, Parallel Lines, Grids*, pp. 41–54, Springer, London (2013).
19. R. Juarez-Salazar, F. Guerrero-Sanchez, and C. Robledo-Sanchez, “Theory and algorithms of an efficient fringe analysis technology for automatic measurement applications,” *Appl. Opt.* **54**, 5364–5374 (2015).

Biographies for the authors are not available.