Self-Calibration of a 1D Projective Camera and Its Application to the Self-Calibration of a 2D Projective Camera

Olivier Faugeras, Long Quan, and Peter Strum

Abstract—We introduce the concept of self-calibration of a 1D projective camera from point correspondences, and describe a method for uniquely determining the two internal parameters of a 1D camera, based on the trifocal tensor of three 1D images. The method requires the estimation of the trifocal tensor which can be achieved linearly with no approximation unlike the trifocal tensor of 2D images and solving for the roots of a cubic polynomial in one variable. Interestingly enough, we prove that a 2D camera undergoing planar motion reduces to a 1D camera. From this observation, we deduce a new method for self-calibrating a 2D camera using planar motions. Both the self-calibration method for a 1D camera and its applications for 2D camera calibration are demonstrated on real image sequences.

Index Terms—Vision geometry, camera model, self-calibration, planar motion, 1D camera.

1 Introduction

A CCD camera is commonly modeled as a 2D projective device that projects a point in $\mathcal{P}^3$ (the projective space of dimension 3) to a point in $\mathcal{P}^2$. By analogy, we can consider what we call a 1D projective camera which projects a point in $\mathcal{P}^2$ to a point in $\mathcal{P}^1$. This 1D projective camera may seem very abstract, but many imaging systems using laser beams, infrared, or ultrasound acting only on a source plane can be modeled this way. What is less obvious, but more interesting for our purpose, is that in some situations, the usual 2D camera model is also closely related to this 1D camera model. First, one example might be the case of the 2D affine camera model operating on line segments: The direction vectors of lines in 3D space and in the image correspond to each other via this 1D projective camera model [20]. Other cases will be discussed later.

In this paper, we first introduce the concept of self-calibration of a 1D projective camera by analogy to that of a 2D projective camera which is a very active topic [17], [12], [7], [13], [1], [27], [19] since the pioneering work of [18]. It turns out that the theory of self-calibration of 1D camera is considerably simpler than the corresponding one in 2D. It is essentially determined in a unique way by a linear algorithm using the trifocal tensor of 1D cameras. After establishing this result, we further investigate the relationship between the usual 2D camera and the 1D camera. It turns out that a 2D camera undergoing planar motion can be reduced to a 1D camera on the trifocal plane of the 2D cameras. This remarkable relationship allows us to calibrate a real 2D projective camera using the theory of self-calibration of a 1D camera. The advantage of doing so is evident. Instead of solving complicated Kruppa equations for 2D camera self-calibration, an exact linear algorithm can be used for 1D camera self-calibration. The only constraint is that the motion of the 2D camera should be restricted to planar motions. The other applications, including 2D affine camera calibration, are also briefly discussed. Part of this work was also presented in [10].

The paper is organized as follows: In Section 2, we review the 1D projective camera model and its trifocal tensor. Then, an efficient estimation of the trifocal tensor is discussed in Section 3. The theory of self-calibration of a 1D camera is introduced and developed in Section 4. After pointing out some direct applications of the theory in Section 5, we develop in Section 6 a new method of 2D camera self-calibration by converting a 2D camera undergoing planar motions into a 1D camera. Experimental results on both simulated and real image sequences are presented in Section 7. Finally, some concluding remarks and future directions are given in Section 8.

Throughout the paper, vectors are denoted in lower case boldface, matrices and tensors in upper case boldface. Some basic tensor notation is used: covariant indices as subscripts, contravariant indices as superscripts and the implicit summation convention.

- *O. Faugeras is with INRIA, 2004, route des Lucioles-B.P. 93, 06902 Sophia Antipolis Cedex, France. E-mail: Oliver.Faugeras@sophia.inria.fr.*
- *L. Quan and P. Sturm are with CNRS-INRIA, ZIRST-655 avenue de l'Europe, 38330 Montbonnot, France. E-mail: {Long.Quan, Peter.Sturm}@inrialpes.fr.*

Manuscript received 12 Mar. 1998; revised 01 Mar. 2000; accepted 28 June 2000.
Recommended for acceptance by S. Sarkar.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number 107800.

2 1D Projective Camera and Its Trifocal Tensor

We will first review the one-dimensional camera which was abstracted from the study of the geometry of lines under affine cameras [20]. Also, we can introduce it directly by analogy to a 2D projective camera.

A 1D projective camera projects a point $\mathbf{x} = (x^1, x^2, x^3)^T$ in $\mathcal{P}^2$ (projective plane) to a point $\mathbf{u} = (u^1, u^2)^T$ in $\mathcal{P}^1$ (projective line). This projection may be described by a 2×3 matrix $\mathbf{M}$ as $\lambda\mathbf{u} = \mathbf{M}_{2\times3}\mathbf{x}$. Now, we examine the geometric constraints available for points seen in multiple views similar to the 2D camera case [22], [23], [13], [26], [9]. There is a constraint only in the case of three views, as there is no constraint for two views (two projective lines always intersect in a point in a projective plane).

Let three views of the point $\mathbf{x}$ be given as follows:

$$\begin{cases} \lambda\mathbf{u} & = & \mathbf{M}\mathbf{x}, \\ \lambda'\mathbf{u}' & = & \mathbf{M}'\mathbf{x}, \\ \lambda''\mathbf{u}'' & = & \mathbf{M}''\mathbf{x}. \end{cases} \tag{1}$$

These can be rewritten in matrix form as

$$\begin{pmatrix} \mathbf{M} & \mathbf{u} & \mathbf{0} & \mathbf{0} \\ \mathbf{M}' & \mathbf{0} & \mathbf{u}' & \mathbf{0} \\ \mathbf{M}'' & \mathbf{0} & \mathbf{0} & \mathbf{u}'' \end{pmatrix} (\mathbf{x}, -\lambda, -\lambda', -\lambda'')^T = \mathbf{0}.$$

The vector $(\mathbf{x}, -\lambda, -\lambda', -\lambda'')^T$ cannot be zero, so

$$\begin{vmatrix} \mathbf{M} & \mathbf{u} & \mathbf{0} & \mathbf{0} \\ \mathbf{M}' & \mathbf{0} & \mathbf{u}' & \mathbf{0} \\ \mathbf{M}'' & \mathbf{0} & \mathbf{0} & \mathbf{u}'' \end{vmatrix} = 0. \tag{2}$$

The expansion of this determinant produces a trifocal constraint for the three views

$$T_{ijk} u^i u'^j u''^k = 0, \tag{3}$$

where T_{ijk} is a $2 \times 2 \times 2$ homogeneous tensor whose components T_{ijk} are 3×3 minors (involving all three views) of the following 6×3 joint projection matrix:

$$\begin{pmatrix} \mathbf{M} \\ \mathbf{M}' \\ \mathbf{M}'' \end{pmatrix} = (1, 2, 1', 2', 1'', 2'')^T.$$

The components of the tensor can be made explicit as $T_{ijk} = [\bar{i}\bar{j'}\bar{k''}]$, for $i, j', k'' = 1, 2$, where the bracket $[ij'k'']$ denotes the 3×3 minor of *i*th, *j*th, and *k*th row vector of the above joint projection matrix and bar "$-$" in $\bar{i}$, $\bar{j}$, and $\bar{k}$ denotes the mapping $(1,2)\mapsto(2,-1)$. It can be easily seen that any constraint obtained by adding further views reduces to a trilinearity. This proves the uniqueness of the trilinear constraint. Moreover, the $2\times2\times2$ homogeneous tensor has $7=2\times2\times2-1$ d.o.f., so it is a minimal parametrization of three views in the uncalibrated setting since three views have exactly $3\times(2\times3-1)-(3\times3-1)=7$ d.o.f., up to a projective transformation in $\mathcal{P}^2$.

This result for the one-dimensional projective camera is very interesting. The trifocal tensor encapsulates exactly the information needed for projective reconstruction in $\mathcal{P}^2$. Namely, it is the unique matching constraint, it minimally parametrizes the three views and it can be estimated linearly. Contrast this to the 2D image case in which the multilinear constraints are algebraically redundant and the linear estimation is only an approximation based on over-parametrization.

3 Estimation of the Trifocal Tensor of a 1D Camera

Each point correspondence in three views $\mathbf{u}\leftrightarrow\mathbf{u}'\leftrightarrow\mathbf{u}''$ yields one homogeneous linear equation for the eight tensor components T_{ijk} for $i,j,k=1,2$:

$$(u^1u'^1u''^1, u^1u'^1u''^2, u^1u'^2u''^1, u^1u'^2u''^2,\\ u^2u'^1u''^1, u^2u'^1u''^2, u^2u'^2u''^1, u^2u'^2u''^2)\mathbf{t}=0,$$

where $\mathbf{t}=(T_{111},T_{112},T_{121},T_{122},T_{211},T_{212},T_{221},T_{222})^T$. With at least seven point correspondences, we can solve for the tensor components linearly.

A careful normalization of the measurement matrix is nevertheless necessary just like that stressed in [11] for the linear estimation of the fundamental matrix. The points in each image are first translated so that their centroid is the origin of the image coordinates, then scaled so that the average distance of the points from the origin is 1. This is achieved by an affine transformation of the image coordinates in each image: $\bar{\mathbf{u}}=\mathbf{A}\mathbf{u}, \bar{\mathbf{u}}'=\mathbf{B}\mathbf{u}'$, and $\bar{\mathbf{u}}''=\mathbf{C}\mathbf{u}''$. With these normalized image coordinates, the normalized tensor components $\bar{T}_{ijk}$ are linearly estimated by SVD from $\bar{T}_{ijk}\bar{u}^i\bar{u}'^j\bar{u}''^k=0$. The original tensor components T_{ijk} are recovered by undoing the normalization transformations: $T_{abc}=\bar{T}_{ijk}A^i_aB^j_bC^k_c$.

4 Self-Calibration of a 1D Camera from Three Views

The concept of camera self-calibration using only point correspondences became popular in the computer vision community following Maybank and Faugeras [18], by solving the so-called Kruppa equations. The basic assumption is that the internal parameters of the camera remain invariant. In the case of the 2D projective camera, the internal calibration (the determination of the five internal parameters) is equivalent to the determination of the image ω of the absolute conic in $\mathcal{P}^3$.

4.1 The Internal Parameters of a 1D Camera and the Circular Points

For a 1D camera represented by a 2×3 projection matrix $\mathbf{M}_{2\times3}$, this projection matrix can always be decomposed into

$$\mathbf{M}_{2\times3}=\mathbf{K}_{2\times2}(\mathbf{R}_{2\times2}\ \ \mathbf{t}_{2\times1}),$$

where

$$\mathbf{K}_{2\times2}=\begin{pmatrix}\alpha & u_0\\ 0 & 1\end{pmatrix}$$

represents the two internal parameters: α, the focal length in pixels and u_0, the position of the principal point; the external parameters are represented by a 2×2 rotation matrix $\mathbf{R}_{2\times2}$,

$$\mathbf{R}_{2\times2}=\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}$$

and the translation vector $\mathbf{t}_{2\times1}$.

The object space for a 1D camera is a projective plane, and any rigid motion on the plane leaves the two circular points I and J invariant (a pair of complex conjugate points on the line at infinity of the plane). Similarly to the 2D camera case where the knowledge of the internal parameters is equivalent to that of the image of the absolute conic, the knowledge of the internal parameters of a 1D camera is equivalent to that of the image points $\mathbf{i}$ and $\mathbf{j}$ of the circular points in $\mathcal{P}^2$.

The relationship between the image of the circular points and the internal parameters of the 1D camera follows directly by projecting one of the circular points $I=(i,1,0)^T$, where $i=\sqrt{-1}$, by the camera $\mathbf{M}_{2\times3}$:

$$\lambda\mathbf{i}=e^{-i\theta}\begin{pmatrix}u_0+i\alpha\\ 1\end{pmatrix}=\begin{pmatrix}\alpha & u_0\\ 0 & 1\end{pmatrix}(\mathbf{R}_{2\times2}\quad\mathbf{t})\begin{pmatrix}i\\1\\0\end{pmatrix}.$$

It clearly appears that the real part of the ratio of the projective coordinates of the image of the circular point $\mathbf{i}$ is the position of the principal point u_0 and the imaginary part is the focal length α.

4.2 Determination of the Images of the Circular Points

Our next task is to locate the circular points in the images. Let us consider one of the circular points, say I. This circular point is projected onto $\mathbf{i}$, $\mathbf{i}'$, and $\mathbf{i}''$ in the three views. As they should be invariant because of our assumption that the internal parameters of the camera are constant, we have:

$$\lambda\mathbf{i}=\lambda'\mathbf{i}'=\lambda''\mathbf{i}''\equiv\mathbf{u},$$

where $\mathbf{u}=(u^1,u^2)^T=\rho(a+ib,1)^T$ for $\lambda,\lambda',\lambda'',\rho\in\mathcal{C}$.

The triplet of corresponding points $\mathbf{i}\leftrightarrow\mathbf{i}'\leftrightarrow\mathbf{i}''$ satisfies the trilinear constraint (3) as all corresponding points do, therefore, $T_{ijk}i^ii'^ji''^k=0$, i.e., $T_{ijk}u^iu^ju^k=0$. This yields the following cubic equation in the unknown $x=u^1/u^2$:

$$T_{111}x^3+(T_{211}+T_{112}+T_{121})x^2+(T_{212}+T_{221}+T_{122})x+T_{222}=0. \quad (4)$$

A cubic polynomial in one unknown with real coefficients has in general either three real roots or one real root and a pair of complex conjugate roots. The latter case of one real and a pair of complex conjugates is obviously the case of interest here. In fact, (4) characterizes all the points of the projective plane which have the same coordinates in three views. This is reminiscent of the 3D case where one is interested in the locus of all points in space that project onto the same point in two views (see Section 6). The result that we have just obtained is that, in the case where the internal parameters of the camera are constant, there are in general three such points: the two circular points which are complex conjugate, and a real point with the following geometric interpretation.

First, consider the case of two views and let us ask the question, what is the set of points such that their images in the two views are the same? This set of points can be called the 2D horopter (h) of the two 1D views. Since the two cameras have the same internal parameters, we can ignore them and assume that we work with the calibrated pixel coordinates. In that case, a camera can be identified to an orthonormal system of coordinates centered at the optical center, one axis is parallel to the retina, the other one is the optical axis. The two views correspond to each other via a rotation

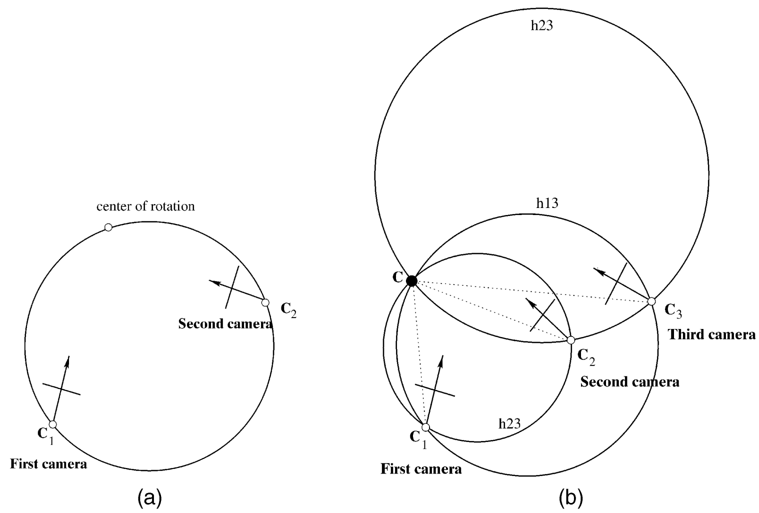

Fig. 1. (a) The two-dimensional horopter which is set of points having the same coordinates in the two views (see text). (b) The geometric interpretation of the real point C which has the same images in all three views (see text).

followed by a translation. This can always be described in general as a pure rotation around a point A whose coordinates can easily be computed from the cameras' projection matrices. A simple computation then shows that the horopter (h) is the circle going through the two optical centers and A, as illustrated in Fig. 1a. In fact, it is the circle minus the two optical centers. Note that since all circles go through the circular points (hence their name), they also belong to the horopter curve, as expected.

In the case of three views, the real point, when it exists, must be at the intersection of the horopter (h_{12}) of the first two views and the horopter (h_{23}) of the last two views. The first one is a circle going through the optical centers C_1 and C_2, the second one is a circle going through the optical centers C_2 and C_3. Those two circles intersect in general at a second point C which is the real point we were discussing, and the third circle (h_{13}) corresponding to the first and third views must also go through the real point C, see Fig. 1b.

We have therefore established the interesting result that the internal parameters of a 1D camera can be uniquely determined through at least seven point correspondences in three views: The seven points yield the trifocal tensor and (4) yields the internal parameters.

5 Applications

The theory of self-calibration of 1D cameras is considerably simpler than the corresponding one in 2D [18] and can be directly used whenever a 1D projective camera model occurs; for instance, self-calibration of some active systems using laser beams, infrared [3], or ultrasound whose imaging system is basically reduced to a 1D camera on the source plane; and partial/full self-calibration of 2D projective cameras using planar motions.

The first type of applications is straightforward. The interesting observation is that the 1D calibration procedure can also be used for self-calibrating a real 2D projective camera if the camera motion is restricted to planar motions. This is discussed in detail in the remainder of this paper.

6 Calibrating a 2D Projective Camera using Planar Motions

A planar motion consists of a translation in a plane and a rotation about an axis perpendicular to that plane. Planar motion is often performed by a vehicle moving on the ground, and has been used for camera self-calibration by Beardsley and Zisserman [4] and by Armstrong et al. [1].

Recall that the self-calibration of a 2D projective camera [8], [18] consists of determining the five unchanging internal parameters of a 2D camera, represented by a 3×3 upper triangular matrix

$$\mathbf{K} = \begin{pmatrix} \alpha_u & s & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix}.$$

This is mathematically equivalent to the determination of the image of the absolute conic ω, which is a plane conic described by $\mathbf{x}^T(\mathbf{K}^{-1})^T(\mathbf{K}^{-1})\mathbf{x} = 0$ for image points $\mathbf{x}$. Given the image of the absolute conic $\mathbf{x}^T\mathbf{C}\mathbf{x} = 0$, the calibration matrix $\mathbf{K}$ can be found from $\mathbf{C}$ using the Choleski decomposition.

6.1 Converting 2D images into 1D images

For a given planar motion, the trifocal plane—the plane through the camera centers—of the camera is coincident with the motion plane as the camera is moving on it. Therefore, the image location of the motion plane is the same as the trifocal line which could be determined from fundamental matrices. The determination of the image location of the motion plane has been reported in [1], [4]. Obviously, if restricting the working space to the trifocal plane, we have a perfect 1D projective camera model which projects the points of the trifocal plane onto the trifocal line in the 2D image plane, as the trifocal line is the image of the trifocal plane. In practice, very few or no point at all really lies on the trifocal plane.

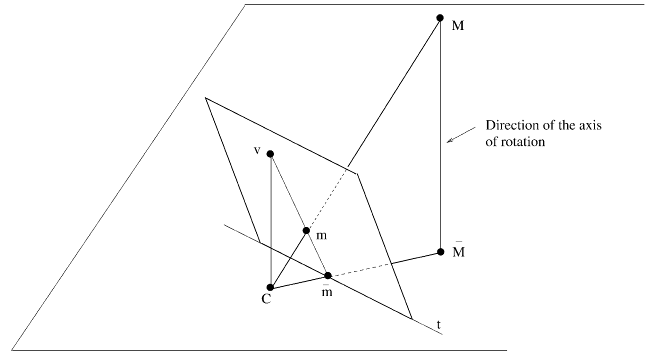

Fig. 2. Creating a 1D image from a 2D image from the vanishing point of the rotation axis and the trifocal line (see text).

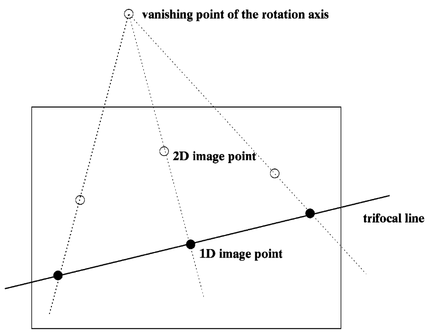

Fig. 3. Converting 2D image points into 1D image points in the image plane is equivalent to a projective projection from the image plane to the trifocal line with the vanishing point of the rotation axis as the projection center.

However, we may virtually project any 3D point onto the trifocal plane, therefore, here comes the central idea of our method: *the 2D images of a camera undergoing planar motion reduce to 1D images by projecting the 2D image points onto the trifocal line*. This can be achieved in at least two ways.

First, if the vanishing point $\mathbf{v}$ of the rotation axis is well-defined. This vanishing point of the rotation axis being the direction perpendicular to the common plane of motion can be determined from fundamental matrices by noticing that the image of the horopter for planar motion degenerates to two lines [1], one of which goes through the vanishing point of the rotation axis; we may refer to [1] for more details.

Given a 3D point M with image m, we mentally project it to $\bar{M}$ in the plane of motion, the projection being parallel to the direction of rotation. The image $\bar{m}$ of this virtual point can be obtained in the image as the intersection of the line through v and m with the trifocal line t. Since the vanishing point v of the rotation axis and the trifocal line t are well defined, this construction, illustrated in Fig. 2, is a well-defined geometric operation.

Note that this is also a perspective projection from $\mathcal{P}^2$ (image plane) to $\mathcal{P}^1$ (trifocal line): $m \mapsto \bar{m}$ as illustrated in Fig. 3.

Alternatively, if the vanishing point is not available, we can nonetheless create the virtual points in the trifocal plane. Given two points M and M' with images m and m', the line (M, M') intersects the plane of motion in $\bar{M}$. The image $\bar{m}$ of this virtual point can be obtained in the image as the intersection of the line (m, m') with the trifocal line t, see Fig. 4.

Another important consequence of this construction is that *2D image line segments can also be converted into 1D image points!* The construction is even simpler, as the resulting 1D image point is just the intersection of the line segment with the trifocal line.

6.1.1 1D Self-Calibration

At this point, we have obtained the interesting result that a 1D projective camera model is obtained by considering only the reprojected points on the trifocal line for a planar motion. The

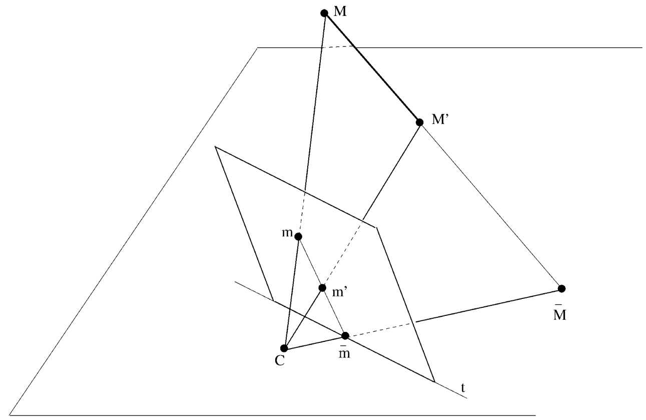

Fig. 4. Creating a 1D image from any pair of points or any line segment (see text).

1D self-calibration method described in Section 4 will allow us to locate the image of the circular points common to all planes parallel to the motion plane.

6.1.2 Estimation of the Image of the Absolute Conic for the 2D Camera

Each planar motion generally gives us two points on the absolute conic, together with the vanishing point of the rotation axes as the pole of the trifocal line w.r.t. the absolute conic. The pole/polar relation between the vanishing point of the rotation axes and the trifocal line was introduced in [1]. As a whole, this provides four constraints on the absolute conic. Since a conic has five d.o.f., at least two different planar motions, yielding eight linear constraints on the absolute conic, will be sufficient to determine the full set of five internal parameters of a general 2D camera by fitting a general conic of the form $\mathbf{x}^T\mathbf{C}\mathbf{x} = au^2 + bv^2 + cuv + du + ev + f = 0$. If we assume a four-parameter model for camera calibration with no image skew (i.e., $s = 0$), one planar motion yielding four constraints is generally sufficient to determine the four internal parameters of the 2D camera. However, this is not true for some very common planar motions such as purely horizontal or vertical motions with the image plane perpendicular to the motion plane. It can be easily proven that there are only three instead of four independent constraints on the absolute conic in these configurations. We need at least two different planar motions for determining the four internal parameters.

Also, this suggests that even if the planar motion is not purely horizontal or vertical, but close, the vanishing point of the rotation axes only constrains loosely the absolute conic. Using only the circular points located on the absolute conic is preferable and numerically stable, but we may need at least three planar motions to determine the five internal parameters of the 2D camera. Note that the numerical instability of the vanishing point for a nearly horizontal trifocal line was already reported by Armstrong in [2].

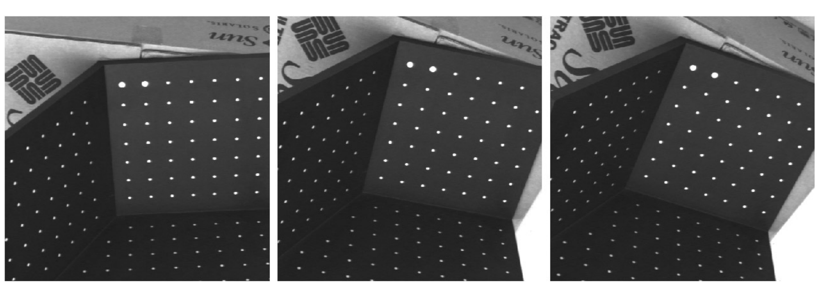

Fig. 5. Three images of the first planar motion.

TABLE 1
Estimated Positions of the Images of the Circular Points by Self-Calibration with Different Triplets of Images of the First Sequence

Image triplet	Fixed point	Circular points by self-calibration	Circular points by calibration
$(16, 19, 22)$	493.7	$290.7 \pm i2779.1$	$310.3 \pm i2650.3$
$(16, 20, 22)$	421.8	$250.1 \pm i2146.3$	$273.9 \pm i2153.5$
$(17, 19, 21)$	533.1	$291.3 \pm i2932.4$	$241.3 \pm i2823.1$
$(16, 18, 20)$	617.8	$238.5 \pm i2597.6$	$238.1 \pm i2791.5$
$(18, 20, 22)$	368.3	$230.6 \pm i2208.2$	$272.1 \pm i2126.2$

The quantities are expressed in the first image pixel coordinate system. The location of circular points by calibration vary as the trifocal line location varies.

TABLE 2
Estimated Positions of the Image of Circular Points with Different Triplets of Images

Image triplet	Circular points	Fixed point
$(16, 18, 20)$	$245.5 \pm i2490.5$	590.0
$(18, 20, 22)$	$221.4 \pm i2717.8$	384.4
$(16, 20, 22)$	$236.2 \pm i2617.3$	452.9
$(16, 19, 22)$	$240.0 \pm i2693.4$	488.0
$(17, 19, 21)$	$304.7 \pm i2722.7$	516.6
known position by calibration	$262.1 \pm i2590.6$	

These quantities vary because the 1D trifocal tensor varies. The trifocal line and the vanishing point of the rotation axes are estimated using seven images of the sequence instead of the minimum of three images.

TABLE 3
Estimated Position of the Image of Circular Points with One Triplet of the Second Image Sequence

Image triplet	Fixed point	Circular points by self-calibration	Circular points by calibration
$(8, 11, 15)$	927.2	$269.7 + i1875.5$	$276.5 + i1540.1$

Obviously, if we work with a three-parameter model with known aspect ratio and without skew, one planar motion is sufficient [1].

As we have mentioned at the beginning of this section, the method described in this section is related to the work of Armstrong et al. [1], but there are some important differences which we now explain.

1. First, our approach gives an elegant insight of the intricate relationship between 2D and 1D cameras for a special kind of motion, called planar motion.
2. Second, it allows us to only use the fundamental matrices of the 2D images and the trifocal tensor of 1D images to self-calibrate the camera instead of the trifocal tensor of 2D images. It is now well-known that fundamental matrices can be very efficiently and robustly estimated [29], [25]. The same is true of the estimation of the 1D trifocal tensor [20] which is a *linear* process. Armstrong et al., on the other hand, use the trifocal tensor of 2D images which, so far, has been hard to estimate due to complicated algebraic constraints to our knowledge. Also, the trifocal tensor of 2D images takes a special form in the planar motion case [1] and the new constraints have to be included in the estimation process.

   It may be worth mentioning that in the case of interest here, planar motion of the cameras, the Kruppa equations become degenerate [28] and the recovery of the internal parameters is impossible from the Kruppa equations. Since it is known that the trifocal tensor of 2D images is algebraically equivalent to the three fundamental matrices plus the restriction of the trifocal tensor to the trifocal plane [14], [15], [9], our method can be seen as an inexpensive way of estimating the full trifocal tensor of 2D images: First, estimate the three fundamental matrices (nonlinear but simple and well-understood), then estimate the trifocal tensor in the trifocal plane (linear).

Although it looks superficially that both the 1D and 2D trifocal tensors can be estimated linearly with at least seven image correspondences, this is misleading since the estimation of the 1D trifocal tensor is exactly linear for seven d.o.f., whereas the linear estimation of the 2D trifocal tensor is only a rough approximation based on a set of 26 auxiliary parameters for its 18 d.o.f. and obtained by neglecting eight complicated algebraic constraints.

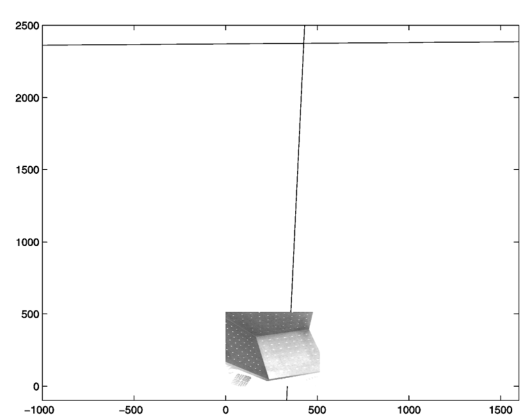

Fig. 6. The image of the motion planes of the two planar motions.

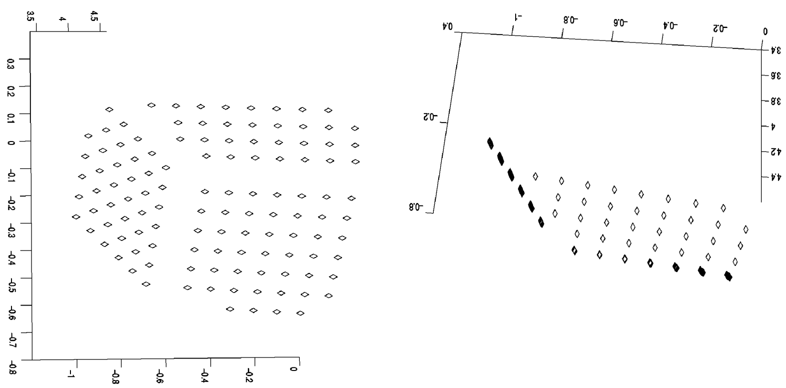

Fig. 7. Two views of the resulting 3D reconstruction by self-calibration.

3. Third, but this is a minor point, our method may not require the estimation of the vanishing point of the rotation axes.

7 Experimental Results

The theoretical results for 1D camera self-calibration and its applications to 2D camera calibration have been implemented and experimented on synthetic and real images. Due to space limitation, we do not present the results on synthetic data, the algorithms generally perform very well. We only show some real examples. Here, we consider a scenario of a real camera mounted on a robot arm. Two sequences of images are acquired by the camera moving in two different planes. The first sequence contains seven (indexed from 16 to 22) images (cf. Fig. 5) and the second contains eight (indexed from 8 to 15).

The calibration grid was used to have the ground truth for the internal camera parameters which have been measured as $\alpha_u = 1534.7$, $\alpha_v = 1539.7$, $u_0 = 281.3$, and $v_0 = 279.0$ using a standard calibration method [6].

We take triplets of images from the first sequence and, for each triplet, we estimate the trifocal line and the vanishing point of the rotation axes using the three fundamental matrices of the triplet. The 1D self-calibration is applied for estimating the images of the circular points along the trifocal lines. To evaluate the accuracy of the estimation, the images of the circular points of the trifocal plane are recomputed in the image plane from the known internal parameters by intersecting the image of the absolute conic with the trifocal line. Table 1 shows the results for different triplets of images of the first sequence.

Since we have more than three images for the same planar motion of the camera, we could also estimate the trifocal line and the vanishing point of the rotation axes by using all the available fundamental matrices of the seven images of the sequence. The results using redundant images are presented for different triplets in Table 2. We note the slight improvement of the results compared with those presented in Table 1.

The same experiment was carried out for the other sequence of images where the camera underwent a different planar motion. Similar results to the first image sequence are obtained. We only give the result for one triplet of images in Table 3 for this sequence.

Now, two sequences of images, each corresponding to a different planar motion, yield four distinct imaginary points on the image plane which must be on the image ω of the absolute conic. Assuming that there is no camera skew, we could fit to those four points an imaginary ellipse using standard techniques and compute the resulting internal parameters. Note that we did not use the pole/polar constraint of the vanishing point of the rotation axes on the absolute conic as it was discussed in Section 6. This constraint is not numerically reliable.

To have an intuitive idea of the planar motions, the two trifocal lines together with one image are shown in Fig. 6.

The ultimate goal of self-calibration is to get 3D metric reconstruction. 3D reconstruction from two images of the sequence is performed by using the estimated internal parameters as illustrated in Fig. 7. To evaluate the reconstruction quality, we did the same reconstruction using the known internal parameters. Two such reconstructions differ merely by a 3D similarity transformation which could be easily estimated. The resulting relative error for normalized 3D coordinates by similarity between the reconstruction from self-calibration and offline calibration is 3.4 percent.

8 Conclusions and Other Applications

First, we have established that the two internal parameters of a 1D camera can be uniquely determined through the trifocal tensor of three 1D images. Since the trifocal tensor can be estimated linearly from at least seven points in three 1D images, the method of the 1D self-calibration is a real linear method (modulo the fact that we have to find the roots of a third degree polynomial in one variable), no over-parameterization was introduced.

Second, we have proven that if a 2D camera undergoes a planar motion, the 2D camera reduces to a 1D camera in the plane of motion. The reduction of a 2D image to a 1D image can be efficiently performed by using only the fundamental matrices of 2D images. Based on this relation between 2D and 1D images, the self-calibration method for 1D cameras can be applied for self-calibrating a 2D camera. Our experimental results based on real image sequences show the good stability of the solutions yielded by the 1D self-calibration method and the accurate 3D metric reconstruction that can be obtained from the internal parameters of the 2D camera estimated by the 1D self-calibration method. The camera motions that defeat the self-calibration method developed in Section 4 are described in [24].

References

[1] M. Armstrong, A. Zisserman, and R. Hartley, "Self-Calibration from Image Triplets," *Proc. Fourth European Conf. Computer Vision,* B. Buxton and R. Cipolla, eds., pp. 3–16, Apr. 1996.

[2] M. Armstrong, "Self-Calibration from Image Sequences," PhD thesis, Univ. of Oxford, 1996.

[3] K. Åström, "Invariancy Methods for Points, Curves, and Surfaces in Computational Vision," PhD thesis, Lund Univ., 1996.

[4] P.A. Beardsley and A. Zisserman, "Affine Calibration of Mobile Vehicles," *Proc. Europe-China Workshop Geometrical Modeling and Invariants for Computer Vision,* R. Mohr and C. Wu, eds., pp. 214–221, Apr. 1995.

[5] T. Buchanan, "The Twisted Cubic and Camera Calibration," *Computer Vision, Graphics, and Image Processing,* vol. 42, no. 1, pp. 130–132, Apr. 1988.

[6] O. Faugeras and G. Toscani, "Camera Calibration for 3D Computer Vision," *Proc. Int'l Workshop Machine Vision and Machine Intelligence,* 1987.

[7] O. Faugeras, "Stratification of Three-Dimensional Vision: Projective, Affine, and Metric Representations," *J. Optical Soc. Am.,* vol. 12, pp. 465–484, 1995.

[8] O. Faugeras and S. Maybank, "Motion from Point Matches: Multiplicity of Solutions," *Int'l J. Computer Vision,* vol. 3, no. 4, pp. 225–246, 1990.

[9] O. Faugeras and B. Mourrain, "About the Correspondences of Points Between n Images," *Proc. Workshop Representation of Visual Scenes,* pp. 37–44, June 1995.

[10] O. Faugeras, L. Quan, and P. Sturm, "Self-Calibration of a 1D Projective Camera and Its Application to the Self-Calibration of a 2D Projective Camera," *Proc. European Conf. Computer Vision,* June 1998.

[11] R. Hartley, "In Defence of the 8-Point Algorithm," *Proc. Fifth Int'l Conf. Computer Vision,* pp. 1,064–1,070, June 1995.

[12] R.I. Hartley, "Euclidean Reconstruction from Uncalibrated Views," *Proc. Darpa-Esprit Workshop Applications of Invariants in Computer Vision,* pp. 187–202, Oct. 1993.

[13] R.I. Hartley, "A Linear Method for Reconstruction from Lines and Points," *Proc. Fifth Int'l Conf. Computer Vision,* E. Grimson, ed., pp. 882–887, June 1995.

[14] A. Heyden, "Geometry and Algebra of Multiple Projective Transformations," PhD thesis, Lund Univ., 1995.

[15] A. Heyden and K. Astrom, "Algebraic Properties of Multilinear Constraints," *Proc. Fourth European Conf. Computer Vision,* B. Buxton and R. Cipolla, eds., pp. 671–682, Apr. 1996.

[16] A. Heyden, "A Common Framework for Multiple-View Tensors," *Proc. Fifth European Conf. Computer Vision,* pp. 3–19, June 1998.

[17] Q.-T. Luong and O. Faugeras, "Self-Calibration of a Moving Camera from Point Correspondences and Fundamental Matrices," *The Int'l J. Computer Vision,* vol. 22, no. 3, pp. 261–289, 1997.

[18] S.J. Maybank and O.D. Faugeras, "A Theory of Self Calibration of a Moving Camera," *Int'l J. Computer Vision,* vol. 8, no. 2, pp. 123–151, 1992.

[19] M. Pollefeys, R. Koch, and L. Van Gool, "Self-Calibration and Metric Reconstruction in Spite of Varying and Unknown Internal Camera Parameters," *Int'l Conf. Computer Vision,* pp. 90–95, 1998.

[20] L. Quan and T. Kanade, "Affine Structure from Line Correspondences with Uncalibrated Affine Cameras," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 19, no. 8, pp. 834–845, Aug. 1997.

[21] J.G. Semple and G.T. Kneebone, *Algebraic Projective Geometry.* Oxford Science Publication, 1952.

[22] A. Shashua, "Algebraic Functions for Recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 17, no. 8, pp. 779–789, Aug. 1995.

[23] M. Spetsakis and J. Aloimonos, "A Unified Theory of Structure from Motion," *Proc. DARPA Image Understanding Workshop,* pp. 271–283, 1990.

[24] P. Sturm, "Vision 3D Non Calibrée: Contributions à la Reconstruction Projective et étude des Mouvements Critiques Pour L'Auto-Calibrage," PhD Thesis, INPG, Dec. 1997.

[25] P.H.S. Torr and A. Zissermann, "Performance Characterization of Fundamental Matrix Estimation Under Image Degradation," *Machine Vision and Applications,* vol. 9, pp. 321–333, 1997.

[26] B. Triggs, "Matching Constraints and the Joint Image," E. Grimson, ed., *Proc. Fifth Int'l Conf. Computer Vision,* pp. 338–343, June 1995.

[27] B. Triggs, "Autocalibration and the Absolute Quadric," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 609–614, June 1997.

[28] Cyril Zeller and Olivier Faugeras, "Camera Self-Calibration from Video Sequences: The Kruppa Equations revisited," Research report 2793, INRIA, Feb. 1996.

[29] Z. Zhang, R. Deriche, O. Faugeras, and Q.T. Luong, "A Robust Technique for Matching Two Uncalibrated Images Through the Recovery of the Unknown Epipolar Geometry," *Artificial Intelligence,* vol. 78, pp. 87–119, 1995.