# Ambiguous Configurations for the 1D Structure and Motion Problem

Kalle Åström and Fredrik Kahl
Centre for Mathematical Sciences
Lund University
Box 118, S-221 00 Lund, Sweden
E-mail: kalle@maths.lth.se, fredrik@maths.lth.se

## Abstract

*In this paper we investigate, determine and classify the critical configurations for solving structure and motion problems for 1D retina vision. We give a complete categorization of all ambiguous configurations for a 1D (calibrated or uncalibrated) perspective camera irrespective of the number of points and views. It is well-known that the calibrated and uncalibrated case are linked through the circular points. This link enables us to solve for both cases simultaneously. Another important tool is the duality in exchanging points and cameras and corresponding Cremona transformation. These concepts are generalized to the 1D case and used for the investigation of ambiguous configurations. Several examples and illustrations are also provided to explain the results and to provide geometrical insight.*

## 1 Introduction

A key problem in computer vision is to recover the shape of an object from a number of its images. This is known as the *structure and motion problem.* Most work has so far been concentrated on the 2D images of a 3D object [9, 13]. This inverse problem has a number of inherent ambiguities. One well-studied ambiguity is when the visible features lie on a special surface, called a **critical surface**, and the cameras have a certain position relative to the surface. Critical surfaces or "gefährlicher Ort" were already studied by Krames [17] based on a monograph from 1880 on quadrics [23]. In the case of one-dimensional cameras, the situation is less clear. The purpose of this paper is to investigate and classify all critical configurations for the 1D perspective camera.

One-dimensional cameras have proven useful in many applications. In [22, 3] it was shown that the structure and motion problem using lines for affine cameras can be reduced to the structure and motion problem for 1D cameras. Another area of application is vision for planar motion. The ordinary 2D retina vision can be reduced to that of 1D cameras if the motion is planar, i.e. the camera is rotating and translating in one specific plane, cf. [1, 10]. A typical example is the case where a camera is mounted on a vehicle that moves on a flat plane.

The 1D camera may also serve as a good model for the navigation system in *laser guided vehicles*, called LGV. The vision system uses strips of reflector tape which are put on walls or objects along the route of the vehicle, cf. [15]. The *laser scanner* measures the direction from the vehicle to the beacons. This information is used to calculate the position of the vehicle. One example of an LGV is shown in Figure 1.

In this paper the critical configurations for structure and motion estimation with 1D retina cameras are studied and classified. A complete categorization of the different ambiguities is given, both for calibrated and uncalibrated cameras. The paper is primarily based on three mathematical tools:

- The connection between the calibrated and uncalibrated case through the circular points.
- The camera-point duality and the Cremona transformation that makes it possible to switch roles between cameras and points.
- Multiview geometry using multilinear constraints. For the 1D retina case there is only the trilinear tensor and its dual.

The approach taken here is much inspired by the works of Carlsson [7, 8], Hartley and Debunne [12] and Maybank [19].

Prior work on critical configurations has been focused on the 2D perspective camera. For critical surfaces and curves, see [17, 14, 6, 18, 20, 11] and for critical camera motions (in the context of auto-calibration), see [26, 16].

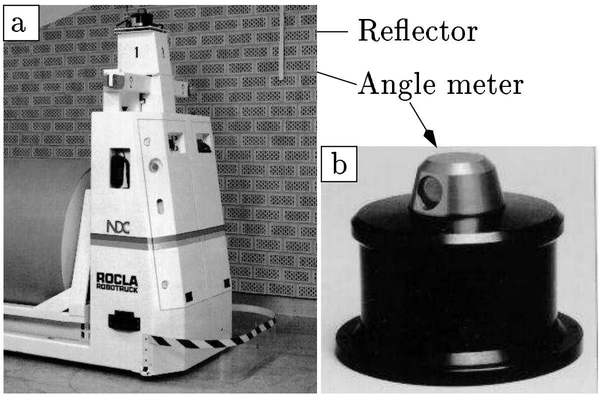


Figure 1: a: A laser guided vehicle. b: A laser scanner or angle meter.

In [2, 22], the 1D camera was studied and it was shown that there are in general two solutions to the structure and motion problem for three views and any number of points. This was further investigated in [4]. It was shown that it is in general possible to solve the structure and motion problem if one has at least 4 points in 4 view or 5 points in 3 views (calibrated case). For the uncalibrated case, two more points are needed. Algorithms for solving the structure and motion problem under different settings were also given.

The results presented here are of both practical and theoretical interest. On a theoretical level, the 1D retina version of camera-point duality and the Cremona transformation can be very useful and provide valuable insight into these problems. A complete classification of critical configurations is useful in practical situations when designing measurement paths for structure and motion estimation. Naturally, one wants to avoid the critical configurations. Although it is quite unlikely that a real life situation is exactly critical, near critical configurations are likely to show up as badly conditioned estimation problems.

## 2 The 1D perspective camera

The camera model considered in this paper is the 1D analogue of the 2D perspective camera, cf. [9, 13].

Introduce an object coordinate system which will be held fixed with respect to the object. Place the camera centre at the origin and let the camera

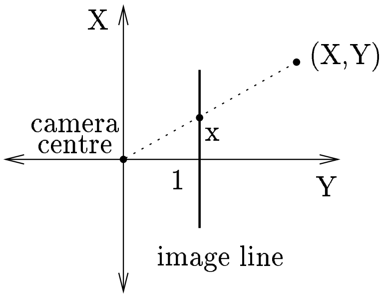


Figure 2: The 1D perspective camera.

axis be aligned with the $Y$-axis. See Figure 2 for an illustration. Then, a point $(X, Y)$ in the plane is projected to a point $x$ on the image line as

$$x = \frac{X}{Y} \,. \tag{1}$$

Using homogeneous coordinates the above projection can be written as a linear equation,

$$\lambda \underbrace{\begin{bmatrix} x \\ y \end{bmatrix}}_{\mathbf{u}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}}_{\mathbf{P}} \underbrace{\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}}_{\mathbf{U}} \quad \Leftrightarrow \quad \lambda \mathbf{u} = \mathbf{P}\mathbf{U} \,, \tag{2}$$

where $\lambda$ is a scalar factor, $\mathbf{u}$ and $\mathbf{U}$ are homogeneous coordinate vectors for points on the line and plane, respectively. If the camera centre is at position $\mathbf{C} = (C_x, C_y)$ and the camera axis is not aligned with the Y-axis, the projection becomes

$$\lambda \mathbf{u} = \begin{bmatrix} \mathbf{R} & -\mathbf{R}\mathbf{C} \end{bmatrix} \mathbf{U} \,, \tag{3}$$

where $\mathbf{R}$ is a $2 \times 2$ rotation matrix encoding the orientation of the camera.

The 1D perspective camera has also two intrinsic parameters. One is the focal length, which denotes the distance from the camera centre to the image line, and the other one is the principal point, which denotes the point on the image line where the camera axis intersects. In the projection equation (3), the focal length is set to one and the principal point to zero. For arbitrary values of focal length $f$ and principal point $x_0$, the projection is

$$\lambda \mathbf{u} = \begin{bmatrix} f & x_0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{R} & -\mathbf{R}\mathbf{C} \end{bmatrix} \mathbf{U} = \mathbf{K} \begin{bmatrix} \mathbf{R} & -\mathbf{R}\mathbf{C} \end{bmatrix} \mathbf{U} \,. \tag{4}$$

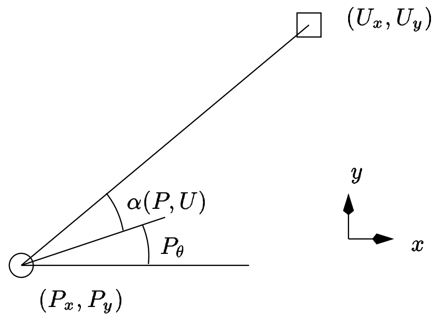


Figure 3: The figure illustrates the measured angle $\alpha$ as a function of scanner position $(P_x, P_y)$, scanner orientation $P_\theta$ and beacon position $(U_x, U_y)$.

When the calibration matrix $\mathbf{K}$ is a priori known the camera is said to be **calibrated**. When $\mathbf{K}$ is unknown, and thus the camera is uncalibrated, it can be regarded as a projection from the projective plane $\mathbf{P}^2$ to the projective line $\mathbf{P}^1$.

An alternative derivation of the camera model originates from the area of laser guided vehicles. See Figure 1.a. for an example. The laser scanner, which is shown in detail in Figure 1.b, is mounted on the top of the vehicle. A laser beam generated by a vertical laser in the scanner is deflected by a rotating mirror. Thus, the laser beam scans the room at a fixed height. When the laser beam hits a beacon (a retroreflective tape, also shown in Figure 1.a), a large part of the light is reflected back to the scanner. The reflected light is processed to find sharp intensity changes. When this happens the bearing $\alpha$ of the laser beam relative to a fixed direction of the scanner is stored.

The bearing $\alpha$ defined above depends on the position of the beacon $(U_x, U_y)$ and of the position $(C_x, C_y)$ and orientation $C_\theta$ of the scanner, cf. Figure 3, according to

$$\alpha = \arg(U_x - C_x + i(U_y - C_y)) - C_\theta\,, \tag{5}$$

where arg is the complex argument (the angle of the vector $(U_x - C_x, U_y - C_y)$ relative to the positive $x$-axis).

The above equation (5) for the measured bearing is non-linear. A somewhat simpler representation of the same equation can be obtained as follows.

The vector between the camera centre and the beacon can be written as

$$\lambda \begin{bmatrix} \cos(\alpha + C_\theta) \\ \sin(\alpha + C_\theta) \end{bmatrix} = \begin{bmatrix} U_x - C_x \\ U_y - C_y \end{bmatrix} = \begin{bmatrix} 1 & 0 & -C_x \\ 0 & 1 & -C_y \end{bmatrix} \begin{bmatrix} U_x \\ U_y \\ 1 \end{bmatrix} .$$

By multiplying each side with a rotation matrix one obtains

$$\lambda \underbrace{\begin{bmatrix} \cos(\alpha) \\ \sin(\alpha) \end{bmatrix}}_{\mathbf{u}} = \underbrace{\begin{bmatrix} \cos(C_\theta) & \sin(C_\theta) \\ -\sin(C_\theta) & \cos(C_\theta) \end{bmatrix} \begin{bmatrix} 1 & 0 & -C_x \\ 0 & 1 & -C_y \end{bmatrix}}_{\mathbf{P}=\begin{bmatrix} \mathbf{R} & -\mathbf{RC} \end{bmatrix}} \underbrace{\begin{bmatrix} U_x \\ U_y \\ 1 \end{bmatrix}}_{\mathbf{U}} .$$

The above equation is of the same form as (3) and thus can be regarded as a calibrated camera. Here we have assumed that the beacon is at a finite position, but the camera equation extends to points at infinity as well.

In conclusion, the 1D perspective camera can be modelled by the equation

$$\lambda \mathbf{u} = \mathbf{PU} , \tag{6}$$

where nothing is assumed about the $2 \times 3$ matrix $\mathbf{P}$ in the uncalibrated case, whereas in the calibrated case the matrix $\mathbf{P}$ is have the additional constraints:

$$\mathbf{P}_{1,1} = \mathbf{P}_{2,2}, \qquad \mathbf{P}_{2,1} = -\mathbf{P}_{1,2} \ . \tag{7}$$

# 3 Problem formulation

Motivated by the previous sections the structure and motion problem will now be defined. We formulate it in an uncalibrated camera setting.

**Problem 3.1.** *Given $n$ image points from $m$ different positions*

$$\mathbf{u}_{I,J}, \quad I = 1, \ldots, m, \ \ J = 1, \ldots, n$$

*the* **structure and motion problem** *is to find the depths $\lambda_{I,J} > 0$, the reconstructed points*

$$\mathbf{U}_J = \begin{pmatrix} X_J \\ Y_J \\ Z_J \end{pmatrix}$$

*and the camera matrices*

$$\mathbf{P}_I = \begin{pmatrix} a_I & b_I & c_I \\ d_I & e_I & f_I \end{pmatrix} ,$$

*such that*

$$\lambda_{I,J}\mathbf{u}_{I,J} = \mathbf{P}_I\mathbf{U}_J\,, \qquad \forall I = 1,\ldots,m, J = 1,\ldots,n\,.$$

It is often convenient to consider things to be equal if they are equal up to scale. The notation $\sim$ will be used to denote equality up to scale. As an example, two camera matrices $\mathbf{P}$ and $\widetilde{\mathbf{P}}$ are considered equal if $\mathbf{P} \sim \widetilde{\mathbf{P}}$. The reason for this is that $\mathbf{P}$ and $\widetilde{\mathbf{P}}$ give the same projections. Only the scale factor $\lambda$ is different.

A projective transformation is a bijective transformation from $\mathbf{P}^n$ to $\mathbf{P}^n$ which can be written, using homogeneous coordinates,

$$\lambda\tilde{\mathbf{U}} = \mathbf{T}\mathbf{U}\,, \ \text{ for some } \ \lambda \neq 0\,,$$

where $\mathbf{T}$ is a non-singular $(n+1)\times(n+1)$ matrix. The projective transformations form a group, which can be obtained from $GL(n+1)$ by identification of proportional matrices. A subgroup of $GL(n+1)$ is the group of similarity transformations, which is defined by restricting the transformation matrices to the following form

$$\begin{bmatrix} \lambda\mathbf{R} & \mathbf{C} \\ 0 & 1 \end{bmatrix} ,$$

where $\mathbf{R}$ is a $n \times n$ rotation matrix, $\lambda$ a scalar, $\mathbf{C}$ a $n$ vector.

We consider two solutions $(\lambda_{I,J}, \mathbf{U}_J, \mathbf{P}_I)$ and $(\tilde{\lambda}_{I,J}, \tilde{\mathbf{U}}_J, \tilde{\mathbf{P}}_I)$ to the structure and motion problem to be the same if they are related by a projective transformation. If there exists a transformation matrix $\mathbf{T}$ such that

$$\widetilde{\mathbf{U}}_J = \mathbf{T}\mathbf{U}_J\,,$$

$$\widetilde{\mathbf{P}}_I = \mu\mathbf{P}_I\mathbf{T}^{-1}\,,$$

$$\tilde{\lambda}_{I,J} = \mu\lambda_{I,J}\,,$$

then both $(\lambda_{I,J}, \mathbf{U}_J, \mathbf{P}_I)$ and $(\tilde{\lambda}_{I,J}, \tilde{\mathbf{U}}_J, \tilde{\mathbf{P}}_I)$ give the same projections $\mathbf{u}_{I,J}$, since

$$\lambda_{I,J}\mathbf{u}_{I,J} = \mathbf{P}_I\mathbf{U}_J, \qquad \forall I = 1,\ldots,m, J = 1,\ldots,n\,.$$

$$\tilde{\lambda}_{I,J}\mathbf{u}_{I,J} = \widetilde{\mathbf{P}}_I \widetilde{\mathbf{U}}_J, \qquad \forall I = 1,\ldots,m,\, J = 1,\ldots,n\,.$$

If the cameras are calibrated, then the camera matrices are of the following form

$$\mathbf{P}_I = \begin{pmatrix} a_I & b_I & c_I \\ -b_I & a_I & d_I \end{pmatrix}\,,$$

and by analogy to the uncalibrated case, two solutions $(\lambda_{I,J}, \mathbf{U}_J, \mathbf{P}_I)$ and $(\tilde{\lambda}_{I,J}, \tilde{\mathbf{U}}_J, \tilde{\mathbf{P}}_I)$ to the structure and motion problem are considered to be the same if they are related by a similarity transformation.

Algorithms for solving the structure and motion problem have previously been presented in, e.g., [2, 22, 4]. Using only two cameras, it is not possible to calculate both structure and motion as any two lines in the plane always intersect. With three measurements of an object point in the plane, there is an addition constraint that the three corresponding lines actually intersect. This can be formulated in the following way.

**Theorem 3.1.** *Let $\mathbf{u}_{1,J}$, $\mathbf{u}_{2,J}$ and $\mathbf{u}_{3,J}$ be the image points of the same object point from three different camera positions. Then the trilinear constraint*

$$\sum_{i,j,k} T_{i,j,k}\mathbf{u}^i_{1,J}\mathbf{u}^j_{2,J}\mathbf{u}^k_{3,J} = 0\,, \tag{8}$$

*is fulfilled for some $2 \times 2 \times 2$ tensor $T$.*

*Proof.* By lining up the camera equations

$$\underbrace{\begin{pmatrix} \mathbf{P}_1 & \mathbf{u}_{1,J} & \mathbf{0} & \mathbf{0} \\ \mathbf{P}_2 & \mathbf{0} & \mathbf{u}_{2,J} & \mathbf{0} \\ \mathbf{P}_3 & \mathbf{0} & \mathbf{0} & \mathbf{u}_{3,J} \end{pmatrix}}_{M} \begin{pmatrix} \mathbf{U}_J \\ -\lambda_{1,J} \\ -\lambda_{2,J} \\ -\lambda_{3,J} \end{pmatrix} = \mathbf{0} \tag{9}$$

we see that the $6 \times 6$ matrix $M$ has a non-trivial right nullspace. Therefore its determinant is zero. Since the determinant is linear in each column it follows that it can be written as

$$\det M = \sum_{i,j,k} T_{i,j,k}\mathbf{u}^i_{1,J}\mathbf{u}^j_{2,J}\mathbf{u}^k_{3,J} = 0\,, \tag{10}$$

for some $2 \times 2 \times 2$ tensor $T$. $\square$

Note that the constraint above only involves the *motion* parameters and the image points. It does not involve the *structure* parameters $\mathbf{U}$. The tensor components can be calculated from the *motion* parameters. If we denote the rows of camera matrix $\mathbf{P}_I$ by $\mathbf{P}_I^1\,\mathbf{P}_I^2$ it is straightforward to see that the tensor components are sub-determinants of the first three columns of the matrix $M$. In fact the components can be obtained as

$$T_{ijk} = \wedge_{ii'}\wedge_{jj'}\wedge_{kk'} \det \begin{bmatrix} \mathbf{P}_1^{i'} \\ \mathbf{P}_2^{j'} \\ \mathbf{P}_3^{k'} \end{bmatrix} .$$

where the tensor $\wedge$ is defined as

$$\wedge_{11} = 0, \qquad \wedge_{12} = -1, \qquad \wedge_{21} = 1, \qquad \wedge_{22} = 0 .$$

It is natural to think of the tensor as being defined only up to scale. Two tensors $T$ and $\widetilde{T}$ are considered equal if they differ only by a scale factor

$$T \sim \widetilde{T} .$$

Let $\mathcal{T}$ denote the set of equivalence classes of trilinear tensors.

As discussed in Section 2 only the relative motion of the camera is important.

**Definition 3.1.** *Let the manifold of* **relative orientation** *of three cameras be defined as the set of equivalence classes of three ordered camera matrices*

$$\mathcal{P} = \left\{ (\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3) \,|\, \mathbf{P}_I = \begin{pmatrix} a_I & b_I & c_I \\ d_I & e_I & f_I \end{pmatrix} \right\} \Big/ \simeq$$

*where the equivalence is defined as*

$$(\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3) \simeq (\widetilde{\mathbf{P}}_1, \widetilde{\mathbf{P}}_2, \widetilde{\mathbf{P}}_3), \;\; \exists \mathbf{T} \in GL(3), \widetilde{\mathbf{P}}_I \sim \mathbf{P}_I \mathbf{T} ,$$

*for* $I = 1, 2, 3$.

The tensor and its relationship to $\mathcal{P}$ has thoroughly been studied in [4]. One key result is the following theorem.

**Theorem 3.2.** *The map*

$$T : \mathcal{P} \longrightarrow \mathcal{T}$$
$$T(\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3)_{ijk} = \wedge_{ii'} \wedge_{jj'} \wedge_{kk'} \det \begin{bmatrix} \mathbf{P}_1^{i'} \\ \mathbf{P}_2^{j'} \\ \mathbf{P}_3^{k'} \end{bmatrix} \tag{11}$$

*is a well defined two-to-one mapping.*

This implies that for three cameras and any number of points seen in these views, there are always two possible solutions and without any further information, one cannot tell which solutions is the correct one. An additional camera will in general lead to a unique solution.

The two solutions are related by a Cremona transformation [21, 2]. The two solutions coincide in the special case where the three camera centers are aligned. In the special case of calibrated cameras the transformation is well-known to be the so called isogonal conjugacy, cf. [5, p. 113].

# 4 The uncalibrated vs the calibrated case

If a camera is calibrated, its camera matrix takes the special form (3), while an uncalibrated camera matrix is allowed to be a general $2 \times 3$ matrix. We have also seen that in the uncalibrated case the solution is in general only determined up to an unknown projective transformation while in the calibrated case the solution is determined up to an unknown similarity.

Another way to characterize the difference between the calibrated and uncalibrated case is through the use of the circular points, see [24].

**Theorem 4.1.** *Knowing that the camera is corrected for internal calibration is equivalent to seeing two extra points (the circular points) in each image.*

*Proof.* In this proof the $i$ will be used for the complex number $i = \sqrt{-1}$. If the camera is corrected for internal calibration and has the following structure

$$\mathbf{P} = \begin{bmatrix} a & b & c \\ -b & a & d \end{bmatrix} ,$$

then the image of the two circular points

$$\mathbf{C}_1 = \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix} \qquad \mathbf{C}_2 = \begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}$$

is known to be

$$\mathbf{c}_1 = \begin{pmatrix} 1 \\ i \end{pmatrix} \qquad \mathbf{c}_2 = \begin{pmatrix} 1 \\ -i \end{pmatrix}$$

since

$$\underbrace{(a+bi)}_{\lambda_1}\mathbf{c}_1 = \mathbf{PC}_1, \qquad \underbrace{(a-bi)}_{\lambda_2}\mathbf{c}_2 = \mathbf{PC}_2\,.$$

On the other hand, in the projective case, if we change the image coordinates so that $\mathbf{u}_{i,1} = \mathbf{c}_1$ and $\mathbf{u}_{i,2} = \mathbf{c}_2$ and also choose object coordinate system so that $\mathbf{U}_1 = \mathbf{C}_1$ and $\mathbf{U}_2 = \mathbf{C}_2$ then the projection matrices must have the form of a calibrated camera matrix since

$$\lambda_1 = \mathbf{P}_{1,1} + \mathbf{P}_{1,2}i, \qquad \lambda_1 i = \mathbf{P}_{2,1} + \mathbf{P}_{2,2}i, \tag{12}$$

$$\lambda_2 = \mathbf{P}_{1,1} - \mathbf{P}_{1,2}i, \qquad -\lambda_2 i = \mathbf{P}_{2,1} - \mathbf{P}_{2,2}i \tag{13}$$

But then we have

$$(\lambda_1 + \lambda_2) = 2\mathbf{P}_{1,1}$$

$$(\lambda_1 + \lambda_2)i = 2\mathbf{P}_{2,2}i$$

so $\mathbf{P}_{1,1} = \mathbf{P}_{2,2}$ And also

$$(\lambda_1 - \lambda_2) = 2\mathbf{P}_{1,2}i$$

$$(\lambda_1 - \lambda_2)i = 2\mathbf{P}_{2,1}$$

so $\mathbf{P}_{1,2} = -\mathbf{P}_{2,1}$ □

An important implication of the the theorem that will be used vividly in the sequel is the following corollary.

**Corollary 4.1.** *The uncalibrated structure and motion problem with $n$ points and $m$ images is equivalent to the calibrated structure and motion problem with $n-2$ points in $m$ images.*

To get the equivalence of the two problems, any 2 of the n points in the uncalibrated problem are chosen and considered to be the circular points. In this setting the uncalibrated problem becomes calibrated, but with two points less.

# 5 The camera-point duality and the Cremona transformation

## 5.1 The uncalibrated case

In [7] Carlsson showed that there is a dual relationship between object points and camera centres for an uncalibrated 2D camera. This duality holds also for the case of uncalibrated projections from 2D to 1D, cf. [4].

Following the terminology introduced by Hartley and Debunne [12], the Carlsson duality can be expressed by a certain Cremona transformation. Cremona transformations are birational transformations of a linear space into another (or the same) space, see [25, p. 20]. Here we will use a specific Cremona transformation, the so called standard quadratic transformation, see [25, p. 47].

**Definition 5.1.** *The mapping* $\Gamma : \mathbf{P}^2 \mapsto \mathbf{P}^2$ *given by*

$$(X, Y, Z) \mapsto (YZ, XZ, XY)$$

*will be called the standard quadratic transformation. The image of a point* $\mathbf{U}$ *under* $\Gamma$ *will be denoted* $\mathbf{U}'$.

The mapping is not defined for any points on the lines joining $(1, 0, 0)$, $(0, 1, 0)$,$(0, 0, 1)$. These three points are called *base points* of the map.

**Theorem 5.1.** *The uncalibrated structure and motion problem with* $n$ *points and* $m$ *images is equivalent to the uncalibrated structure and motion problem with* $m + 3$ *points and* $n - 3$ *images.*

*Proof.* The proof is based on using three (of the n) points as partial basis points, e.g. the first three points. Without loss of generality one may change projective coordinate systems, both in the images and in the scene. Change image coordinate system so that

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \mathbf{u}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad \mathbf{u}_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

and then change object coordinates so that

$$\mathbf{U}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \mathbf{U}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \mathbf{U}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

The projection equation $\lambda_i \mathbf{u}_i = \mathbf{P}\mathbf{U}_i$ for these three points puts linear constraints on the camera matrix. It follows that the camera matrix has the following form

$$\mathbf{P} = \begin{pmatrix} \frac{1}{C_1} & 0 & -\frac{1}{C_3} \\ 0 & \frac{1}{C_2} & -\frac{1}{C_3} \end{pmatrix}, \tag{14}$$

where $\mathbf{C} = (C_1, C_2, C_3)$ is the camera centre. Note that $\mathbf{C}$ is in the right nullspace of $\mathbf{P}$. If $\mathbf{C}$ is any point in $\mathbf{P}^2$, then the matrix of form (14) will be denoted by $P_{\mathbf{C}}$. Now, for any point $\mathbf{U} = (U_1, U_2, U_3)$, let $\mathbf{U}' = \Gamma(\mathbf{U})$ and $\mathbf{C}' = \Gamma(\mathbf{C})$ where $\Gamma(\cdot)$ is the Carlsson map. Note for instance that

$$\mathbf{U}' = \Gamma(\mathbf{U}) = (U_1, U_2, U_3) = (U_2U_3, U_1U_3, U_1U_2) \sim (\frac{1}{U_1}, \frac{1}{U_2}, \frac{1}{U_3}).$$

It follows that

$$\mathbf{P}_{\mathbf{C}}\mathbf{U} = \begin{pmatrix} \frac{U_1}{C_1} - \frac{U_3}{C_3} \\ \frac{U_2}{C_2} - \frac{U_3}{C_3} \end{pmatrix} = \begin{pmatrix} \frac{C_1'}{U_1'} - \frac{C_3'}{U_3'} \\ \frac{C_2'}{U_2'} - \frac{C_3'}{U_3'} \end{pmatrix} = \mathbf{P}_{\mathbf{U}'}\mathbf{C}'. \tag{15}$$

Thus, $\Gamma$ interchanges the roles of object points and camera centres. Thus solving $n$ points in $m$ images is equivalent to solving $m+3$ points in $n-3$ images. □

It will be useful to know how the Carlsson map acts on other geometric objects than object points and camera centres.

**Lemma 5.1.** *The Carlsson map transfers (i) a line passing through two general points $\mathbf{U}_1$ and $\mathbf{U}_2$ to a conic through the dual points $\mathbf{U}_1'$, $\mathbf{U}_2'$ and the three base points and (ii) a cubic curve passing through the three base points to a cubic curve passing through the three base points.*

*Proof.* See Page 49, Theorem V in [25]. □

## 5.2 The calibrated case

In the calibrated case we have two points for "free" (the circular points). These can be used as two of the three base points in forming the camera-point duality. Only one additional points is then needed. For simplicity we will think of this one base point as the origin.

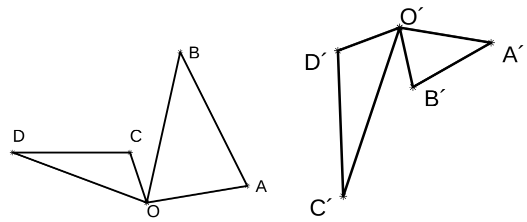


Figure 4: Example of a calibrated Cremona transformation.

**Theorem 5.2.** *The calibrated structure and motion problem with $n$ points and $m$ images is equivalent to the calibrated structure and motion problem with $m+1$ points and $n-1$ images.*

*Proof.* This follows immediately from Theorems 4.1 and 5.1. □

The calibrated Cremona transformation can be derived by changing coordinate systems from the one used in the uncalibrated case, to one where the circular points serve as base points as well.

**Lemma 5.2.** *Consider the calibrated Cremona transformation*

$$(X, Y, Z) \mapsto (XZ, -YZ, X^2 + Y^2) \,. \tag{16}$$

*The transformation has the property that from every point $A$ the angle measured to an arbitrary point $B$ relative to the origin is the same as the angle from the dual point $B'$ to the point $A'$.*

The lemma is illustrated by Figure 4. Notice that the triangle $OAB$ is congruent to $OB'A'$, the triangles $OCD$ is congruent to $OD'C'$ and similarly for any triangle with $O$ as one of the vertices. This means that if we measure bearings from any set of camera positions $(\mathbf{C}_1, \ldots, \mathbf{C}_m)$ to the the points $(\mathbf{X}_1, \ldots, \mathbf{X}_n)$ we will get the same angles as we would if we measured from $(\mathbf{X}'_1, \ldots, \mathbf{X}'_n)$ to $(\mathbf{C}'_1, \ldots, \mathbf{C}'_m)$.

So an algorithm that solves the structure and motion problem for a particular configuration can also be applied to solve for the dual configuration, cf. [12].

1. Dualize image measurements. Subtract from each measured bearing from a given camera the bearing to the first point. Form the dual measurements by transposing the roles of cameras and points for in all views to all points but the first. Add zeros to the first column.

2. Solve the dual structure and motion problem.

3. Dualize the measurements, camera positions and object points.

The calibrated Cremona transformation has several interesting properties.

**Lemma 5.3.** *The dual of a line not passing through the base point is a circle through the base point. The dual of a cubic passing through the base point and the circular points is a cubic passing through the base point and the circular points.*

# 6 Basic ambiguities

We have seen previously that a solution to the structure and motion problem is only determined up to an unknown projective transformation. Also, for three cameras and any number of points, there is a two-fold ambiguity according to Theorem 3.2. Additionally, there are two basic ambiguities that will be discussed here.

The problem of calculating object points using known camera positions is known as *intersection*. There is one critical configuration for which there is not a unique solution.

**Theorem 6.1.** *Consider the case of several views of one point with known camera matrices. The intersection problem is ambiguous if and only if all camera centres and the point lie on a line.*

*Proof.* If all camera centres and the object point are aligned, then any point on that line will generate the same image. Thus it is it easy to see that such a configuration is ambiguous. If the camera centres are not aligned with the point, then there are lines from (at least) two views which have a unique intersection. □

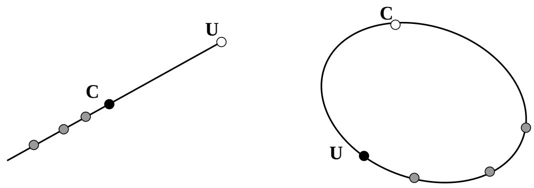


Figure 5: **Left**: The intersection ambiguity, where a point **U** and several camera centres **C** lie on a line. **Right**: The dual resection ambiguity, where a camera centre **C** and several points **U** lie on a conic curve.

*Resection* is the problem of calculating camera positions using image measurements and known object points. In this case, the critical configurations are not obvious.

**Theorem 6.2.** *Consider $m > 4$ object points with known positions and one unknown camera. The resection problem is ambiguous if and only if all points and the camera centre lie on a conic curve.*

*Proof.* Consider first the critical configuration for the intersection problem, i.e. one point and $n > 1$ camera centres lying on a single line. The configuration is still ambiguous if more object points are added. Four points (where no three are on a line) and several $n > 1$ cameras are ambiguous if and only if one of the points and all camera centres are on a single line. If the other three points are taken as base points, the dual statement is (using Lemma 5.1 for the line): $m > 4$ object points and one camera centre are ambiguous if and only if all points and the camera centre lie on a conic curve. $\square$

The intersection and resection ambiguities are illustrated in Figure 5. The calibrated version follows from Lemma 5.3.

**Corollary 6.1.** *Consider $m > 2$ object points with known positions and one unknown camera. The calibrated resection problem is ambiguous if and only if all points and the camera centre lie on a circle.*

# 7 Three view ambiguities

A structure and motion problem with three views can be ambiguous in three ways.

1. The alternative reconstructions have the same relative camera motion.
2. The alternative reconstruction have different relative camera motion, but the corresponding trilinear tensor is the same.
3. The alternative reconstructions have different relative camera motion and the corresponding trilinear tensor is different.

For case one there is a unique relative motion, so one can without loss of generality assume that the camera positions are known. The alternative reconstruction differs in at least one of the object points. This can only happen if the camera centres and that point is collinear, see Theorem 6.1. For case two, Theorem 3.2 shows that for each trilinear tensor there are two possible relative orientations. Thus any three view problem is critical in the sense that there are at least two possible solutions. For the third case, we ask if there are cases where there might be more than two solutions to the structure and motion problem, i.e. when the tensor is not uniquely defined. We will call this case a *three view ambiguity.*

We are now ready to state the theorem describing exactly when there are three view ambiguities. For an example, see Figure 6.

**Theorem 7.1.** *The structure and motion problem for three views and arbitrary number of points is ambiguous if and only if the three camera centres and all the object points lie on a cubic curve.*

There is an interesting special case when all the points and at least one of the camera centres lie on a conic. It fits into the theorem since there is a cubic consisting of the conic through the points and one camera centre and a line through the remaining camera centres. The cubic thus covers all points and camera centres. The problem is then critical in the sense that the resection problem for the first camera is critical, cf. Theorem 6.2.

*Proof.* Consider a situation where there is an ambiguity. Consider one of the solutions to the problem. For this solution there is a placement of cameras, $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$. The condition that there is an ambiguous solution is equivalent to saying that there is an alternative tensor $T_{ijk}$ such that

$$\sum T_{ijk}\mathbf{a}^i\mathbf{b}^j\mathbf{c}^k = 0\,, \tag{17}$$

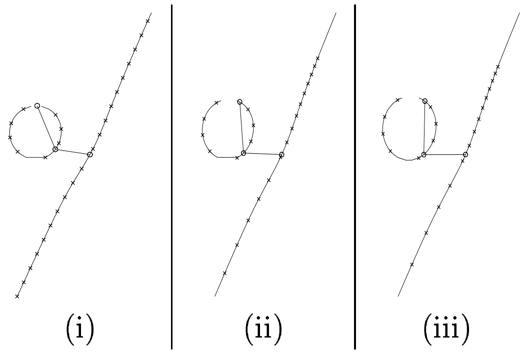

Figure 6: Three cameras (*circles*) are viewing 22 points (*crosses*). All three configurations (out of a one-parameter family) are consistent with the 1D image points. The 25 plane points lie on a cubic.

where $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ are image points in the three images respectively. Introducing the notation $\mathbf{A}^i$ for row $i$ of camera matrix $A$ and similarly for camera matrices $B$ and $C$, the image coordinates are given by:

$$\mathbf{a}^i = \mathbf{A}^i\mathbf{X}, \quad \mathbf{b}^i = \mathbf{B}^i\mathbf{X}, \quad \mathbf{c}^i = \mathbf{C}^i\mathbf{X}\,.$$

By inserting this into (17) we see obtain the following constraint on possible object points:

$$p(\mathbf{X}) = \sum T_{ijk}(\mathbf{A}^i\mathbf{X})(\mathbf{B}^j\mathbf{X})(\mathbf{C}^k\mathbf{X}) = 0\,,$$

The constraint on the object point $\mathbf{X}$ is a third degree polynomial in $\mathbf{X} \in \mathbf{P}^2$. This shows that all object points pass through this cubic curve. To see that the camera centres lie on the same curve it is sufficient to observe that $\mathbf{AF} = 0$, when $\mathbf{F}$ is the camera centre for camera one. This gives directly that $p(\mathbf{F}) = 0$. Notice that the structure and motion problem is in general well defined for 7 points in 3 views. With 6 object points and 3 camera centres, there is in general a unique cubic curve passing through these points. That the 7th object point also lie on this curve is exceptional. To show the only if part we consider an object where all camera centres and object points lie on an arbitrary third degree polynomial. Without loss of generality we may change both object coordinate system and image coordinate system so that

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

as long as the three cameras are not on a line.

The mapping from ambiguous tensors to cubic curves is a linear mapping. Each ambiguous tensor which has eight parameters

$$T = (T_{111}, T_{112}, T_{121}, T_{122}, T_{211}, T_{212}, T_{221}, T_{222})^T$$

corresponds to a cubic curve

$$p(\mathbf{X}) = \sum T_{ijk} (\mathbf{A}^i \mathbf{X})(\mathbf{B}^j \mathbf{X})(\mathbf{C}^k \mathbf{X}) = 0 \ ,$$

where the coefficients

$$c = \left(c_{x^3}, c_{x^2y}, c_{x^2z}, c_{xy^2}, c_{xyz}, c_{xz^2}, c_{y^3}, c_{y^2z}, c_{yz^2}, c_{z^3}\right)^T$$

of the polynomial $p(\mathbf{X})$ depend linearly on the tensor coefficients

$$c = MT \, . \tag{18}$$

For this particular choice of coordinates the matrix $M$ becomes

$$M = \begin{bmatrix} 0&0&0&0&0&0&0&0 \\ 1&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0 \\ 0&0&0&0&1&0&0&0 \\ 0&0&1&0&0&1&0&0 \\ 0&0&0&1&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&0&1 \\ 0&0&0&0&0&0&0&0 \end{bmatrix} \, .$$

It is straightforward to see that the matrix $M$ has rank 7. Notice that the true tensor is a null vector to the matrix, so $M$ must have rank less than or equal to 7. If the three camera centres happen to be on a line, it is easy to check that the corresponding mapping is also linear with rank 7. The space of tensors is a projective space of dimension 7. Since the true tensor $T_{\mathrm{true}}$ maps onto the null vector, the mapping (18) is in fact constant on linear combinations of a tensor $T$ and $T_{\mathrm{true}}$, i.e.

$$M(\lambda T + \mu T_{\mathrm{true}}) \sim MT$$

. In other words lines in the tensor space through $T_{\mathrm{true}}$ are mapped onto the same cubic. This space of lines is a projective space of dimension 6. The mapping (18) is in fact a bijective mapping from this space (which can be identified with $\mathbf{P}^6$) to the manifold of cubic curves that pass through the three camera centres (also $\mathbf{P}^6$).

Since the mapping is bijective, our arbitrary third degree curve on which the object points lie, correspond to an ambiguous tensor (in fact a one parameter family of tensors). Thus the structure and motion problem for that case is critical. This concludes the proof. □

From the principle of duality, the following theorem is obtained.

**Theorem 7.2.** *The structure and motion problem for any number of views of 6 points is ambiguous if and only if the camera centres and the object points lie on a cubic curve.*

*Proof.* The image under the Carlsson map of a cubic curve through the base points is again a cubic curve through the base point, cf. Lemma 5.1. The dual of 3 cameras and $n$ points is $m = n - 3$ cameras and 6 points. So by the principle of duality and Theorem 7.1, the statement is proved. □

# 8 General n points in m views ambiguity

Up to now, we have limited either the number of cameras or the number of points considered. Based on the previous results, the general problem will now be solved. A natural generalization of the three view case for the word "ambiguous" is that the alternative reconstructions have different relative camera motion and for each triplet of cameras theree exists ambiguous solution with different trilinear tensor. A measurement situation still admits alternative solutions if all cameras and at least one point are colinear or if all points and at least one camera lies on a conic. If for some triplet of cameras the relative motion is determined by measurements, then all the points are also uniquely determined. The ambiguity must then lie in one of the other cameras. This is a resection ambiguity. For the remaining of this section it will thus be assumed that for each triplet of cameras the ambiguous solution admits a different trilinear tensor.

**Theorem 8.1.** *A 1D structure and motion problem is ambiguous (in the above sense) regardless of the number of cameras and points if and only if all the camera centres and the object points lie on a common third degree curve.*

*Proof.* We begin by showing that a problem is ambiguous if all points lie on a third degree curve. Assume that camera centres and object points lie on a third degree curve $c$. By first restricting the problem to only 6 points, we know from Theorem 7.2 that the configuration is ambiguous and there is (at least) one-parameter family of solutions. Now, consider a 7th point on the curve $c$. We need to show that the constraints generated by the projection equation for this extra point does not break the ambiguity. However, all these constraints reduce to trilinear constraints, as there are no higher order

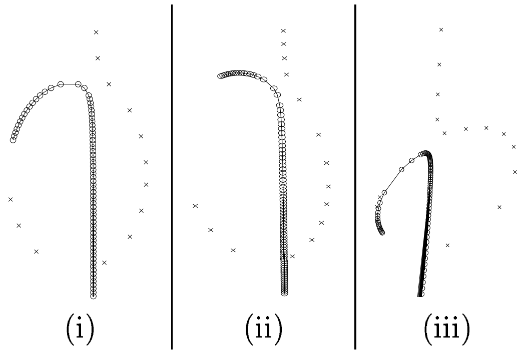


Figure 7: The figure illustrates three solutions (out of a one-parameter family) to the same structure and motion problem. There are 82 cameras (*circles*) viewing 15 points (*crosses*). It is critical because all 97 points lie on a cubic.

constraints for 1D camera motion, cf. [4]. Thus, it suffices to consider three arbitrary cameras $\mathbf{P}_i$, $\mathbf{P}_j$ and $\mathbf{P}_k$. In the proof Theorem 7.1, we showed that the map from stars of tensors to cubic curves (through the camera centres) is bijective. So, from $c$ and the three cameras centres, a star of tensors $\lambda T_1 + \mu T_2$, where $(\lambda, \mu) \in \mathbf{P}^1$, is obtained. But according to the proof of Theorem 7.1, as long as the 7th point is on $c$, all tensors in $\lambda T_1 + \mu T_2$ are still valid solutions.

To show that each ambiguous problem has the property that all points lie on a third degree curve we use a proof by contradiction. Assume, thus that there exist ambiguous problems with $m$ views of $n$ points such that the $m+n$ points do not lie on a common third degree curve. Such problems must have $m > 3$ and $n > 6$ because of Theorems 7.1 and 7.2 respectively. Study such a problem where $m + n$ is minimal. If we remove one point or one camera, we obtain an ambiguous problem with one point less. By the assumption these $m + n - 1$ points must lie on a third degree curve. In particular, this means that all 10 point subconfigurations must lie on a third degree curve. According to Lemma A.1, all $m + n$ points then lie on a cubic curve. Thus all ambiguous configurations have the property that all $m + n$ points lie on a third degree curve. $\square$

In Figure 7 an example of a critical configuration is illustrated. Even though there are 82 views of 15 points, the $1D$ images alone cannot disambiguate between a one-parameter family of solutions. Another example is

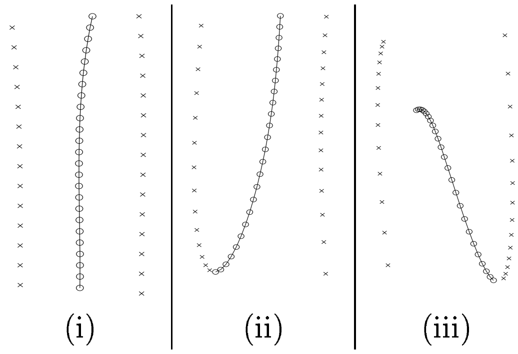

Figure 8: Three solutions (out of a one-parameter family) to the same structure and motion problem. In the example, the camera moves in a corridor with scene points on both walls, which is quite common in robot navigation. There are 25 cameras (*circles*) viewing 29 points (*crosses*). It is critical because all 54 points lie on a cubic.

| Ambiguity | projective case | Euclidean case |
|---|---|---|
| Intersection | Line | Line |
| Resection | Conic | Circle |
| Structure and motion | Cubic | Cubic through circular points |

Table 1: Table illustrating the three different types of ambiguities for both the projective and the Euclidean case.

illustrated in Figure 8, where a camera moves along corridor which is frequently occurring in practical situations.

# 9 Conclusions

We have given a complete categorization of all ambiguous configurations for the structure and motion problem in 1D retina vision. The main ambiguity is when all object points (regardless of how many) and all camera centres (again, regardless of the number of cameras) lie on a cubic curve.

## Acknowledgements

The authors thank Stefan Carlsson for raising the question on which structure and motion problems for 1D retina are degenerate. The authors would also like to thank Magnus Oskarsson. This work has been supported by the Swedish Research Council (Vetenskapsrådet), project 221-2000-476.

## Appendix

**Lemma A.1.** *Let* $\mathbf{Q}_1, \mathbf{Q}_2, ..., \mathbf{Q}_N$ *denote* $N > 10$ *arbitrary points in* $\mathbf{P}^2$*. If, for each combination of 10 points, there exists a cubic curve through these 10 points, then there exists a cubic curve through all* $N$ *points.*

*Proof.* For each point $\mathbf{Q}_i = (X_i, Y_i, Z_i)^T$, let $\tilde{\mathbf{Q}}_i = (X_i^3, X_i^2 Y_i, ..., Z_i^3)^T$, i.e. a 10 vector containing all cubic monomials. Then, $\mathbf{Q}_i$ lies on a cubic with coefficients $c = (c_{x^3}, c_{x^2y}, ..., c_{z^3})^T$ if and only if $c^T \tilde{\mathbf{Q}}_i = 0$. Further, let

$$M = \begin{bmatrix} \tilde{\mathbf{Q}}_1 & \tilde{\mathbf{Q}}_2 & ... & \tilde{\mathbf{Q}}_N \end{bmatrix},$$

which is a $10 \times N$ matrix. As each combination of 10 points lies on a cubic, it follows that each $10 \times 10$ submatrix of $M$ has a non-empty left nullspace. This implies that $\operatorname{rank} M \leq 9$ and there is non-empty left nullspace of $M$. Let $c_M$ be a vector in that nullspace. Thus $c_M^T \tilde{\mathbf{Q}}_i = 0$ and all points lie on the cubic corresponding to $c_M$. □

## References

[1] M. Armstrong, A. Zisserman, and R. Hartley. Self-calibration from image triplets. In *European Conf. Computer Vision*, pages 3–16, Cambridge, UK, 1996.

[2] K. Åström. *Invariancy Methods for Points, Curves and Surfaces in Computational Vision.* PhD thesis, Dept of Mathematics, Lund University, Sweden, 1996.

[3] K. Åström, A. Heyden, F. Kahl, and M. Oskarsson. Structure and motion from lines under affine projections. In *Int. Conf. Computer Vision*, pages 285–292, Kerkyra, Greece, 1999.

[4] K. Åström and M. Oskarsson. Solutions and ambiguities of the structure and motion problem for 1d retina vision. *Journal of Mathematical Imaging and Vision*, 12(2):121–135, 2000.

[5] H. F. Baker. *An Introduction to Plane Geometry.* Cambridge University Press, 1943.

[6] T. Buchanan. The twisted cubic and camera calibration. *Computer Vision, Graphics and Image Processing*, 42:130–132, 1988.

[7] S. Carlsson. Duality of reconstruction and positioning from projective views. In *IEEE Workshop on Representation of Visual Scenes*, pages 85–92, Cambridge Ma, USA, 1995.

[8] S. Carlsson and D. Weinshall. Dual computation of projective shape and camera positions from multiple images. *Int. Journal Computer Vision*, 27(3):227–241, 1998.

[9] O. Faugeras. *Three-Dimensional Computer Vision.* MIT Press, Cambridge, Mass, 1993.

[10] O. D. 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. In *European Conf. Computer Vision*, volume I, pages 36–52, Freiburg, Germany, 1998.

[11] R. Hartley. Ambiguous configurations for 3-view projective reconstruction. In *European Conf. Computer Vision*, volume I, pages 922–935, Dublin, Ireland, 2000.

[12] R. Hartley and G. Debunne. Dualizing scene reconstruction algorithms. In *3D Structure from Multiple Image of Large-Scale Environments, European Workshop, SMILE*, pages 14–31, Freiburg, Germany, 1998.

[13] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision.* Cambridge University Press, 2000.

[14] B. K. P. Horn. Motion fields are hardly ever ambiguous. *Int. Journal Computer Vision*, 1(3):239–258, 1987.

[15] K. Hyyppä. Optical navigation system using passive identical beacons. In *Proceedings intelligent autonomous systems, Amsterdam*, 1987.

[16] F. Kahl, B. Triggs, and K. Åström. Critical motions for auto-calibration when some intrinsic parameters can vary. *Journal of Mathematical Imaging and Vision*, 13(2):131–146, 2000.

[17] J. Krames. Zur Ermittlung eines Objectes aus zwei Perspectiven (Ein Beitrag zur Theorie der gefährlichen Örter). *Monatsh. Math. Phys.*, 49:327–354, 1940.

[18] S. Maybank. The projective geometry of ambiguous surfaces. *Phil. Trans. Royal Soc. London A*, 332:1–47, 1990.

[19] S. Maybank. *Theory of Reconstruction from Image Motion.* Springer-Verlag, Berlin, Heidelberg, New York, 1993.

[20] S. Maybank and A. Shashua. Ambiguity in reconstruction from images of six points. In *Int. Conf. Computer Vision*, pages 703–708, Mumbai, India, 1998.

[21] L. Quan. Inherent two-way ambiguity in 2d projective reconstruction from three uncalibrated 1d images. In *Int. Conf. Computer Vision*, pages 344–349, Kerkyra, Greece, 1999.

[22] L. Quan and T. Kanade. Affine structure from line correspondences with uncalibrated affine cameras. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 19(8):834–845, 1997.

[23] H. Schröter. *Theorie der Oberflächen zweiter Ordnung und der Raumkurve dritter Ordnung als Erzeugnisse projektvischer Gebilde.* Leibzig: Teubner, 1880.

[24] J. G. Semple and G. T. Kneebone. *Algebraic Projective Geometry.* Clarendon Press, Oxford, 1952.

[25] J. G. Semple and L. Roth. *Introduction to Algebraic Geometry.* Clarendon Press, Oxford, 1949.

[26] P. Sturm. Critical motion sequences for monocular self-calibration and uncalibrated Euclidean reconstruction. In *Conf. Computer Vision and Pattern Recognition*, pages 1100–1105, San Juan, Puerto Rico, 1997.