Solutions and Ambiguities of the Structure and Motion Problem for 1D Retinal Vision

Kalle Åström, Magnus Oskarsson
Centre for Mathematical Sciences,
Lund University,
P.O. Box 118,
S-221 00 Lund, SWEDEN
kalle@maths.lth.se,magnuso@maths.lth.se

Abstract

In this paper we investigate the geometry and algebra of multiple one-dimensional projections of a two-dimensional environment. This is relevant for one-dimensional cameras, e.g. as used in certain autonomous guided vehicles. It is also relevant for understanding the projection of lines in ordinary vision. A third application is on ordinary vision of vehicles undergoing so called planar motion. The structure and motion problem for such cameras is studied and the two possible minimal cases is solved. The technique of solving these questions reveal interesting ambiguities. It is shown that for each solution with three images there is an ambiguous solution. It is also shown that for each solution for four points there is an ambiguous solution. The connection between these two different types of ambiguities are also given. Although the paper deals with calibrated cameras, it is shown that similar results holds for uncalibrated cameras.

1 Introduction

Understanding of one-dimensional cameras is important in several applications. In [7] it was shown that the structure and motion problem using line features in the special case of affine cameras can be reduced to the structure and motion problem for points in one dimension less, i.e. one-dimensional cameras. This was used to solve the problem of three views of seven lines. Two solutions are obtained. However, no geometrical interpretation of these two solutions were given.

Another area of application is vision for planar motion. It is shown that ordinary vision (two-dimensional retina) can be reduced to that of one-dimensional cameras if the motion is planar, i.e. if the camera is rotating and translating in one specific plane only, cf. [1]. A typical example is the case where a camera is mounted on a vehicle that drives on a flat plane or flat road.

Our personal motivation, however, stems from the ap-

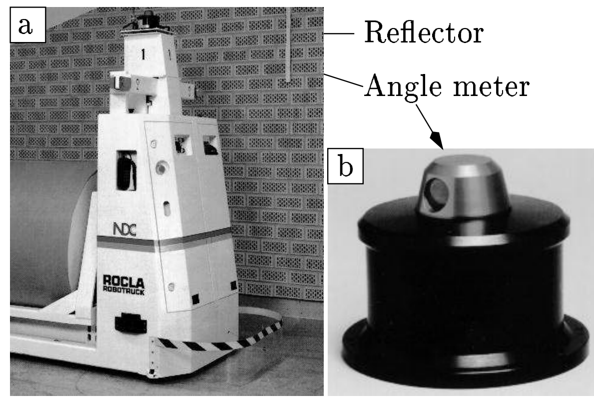

Figure 1: a: A laser guided vehicle. b: A laser scanner or angle meter.

plication of autonomous guided vehicles, which are important components for factory automation. Such vehicles have traditionally been guided by wires buried in the factory floor. This gives a very rigid system. Removal and change of wires is cumbersome and costly. The system can be drastically simplified using navigation methods based on a laser sensor and computer vision algorithms. With such a system the position of the vehicle can be computed instantly. The vehicle can then be guided along any feasible path in the room. This paper deals with some navigation problems for a laser guided vehicle. The navigation system uses strips of inexpensive reflector tape (called **reflectors** or **beacons**) which are put on walls or objects along the route of the vehicle, cf. [5]. The **laser scanner**, also called **angle meter** or **meter**, measures the direction from the vehicle to the beacons, but not the distance. This is the information used to calculate the position of the vehicle.

One interesting problem is the so called **surveying** problem, cf. [2]. This is the procedure to obtain a map of the unknown positions of the beacons using images at unknown position and orientation. This is usually done off-line, once and for all, when the system is installed and then occasionally if there has been changes in the environment. High-accuracy is needed, since the map has to be hard-coded in the system. The performance of the navigation routines depends on the precision of the surveyed map. The *surveying* problem is in essence a *structure and motion* problem, i.e. one tries to solve for both the structure (the map) and the motion of the vehicle.

Note that the discussion here is focused on finding initial estimates of structure and motion estimates. In practice it is necessary to refine these estimates using non-linear optimization or bundle adjustment, cf. [8].

2 Scanner geometry

A laser navigated vehicle is shown in Figure 1.a. 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, at the top of the scanner. 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 α of the laser beam relative to a fixed direction of the scanner is stored. The time t when the reflection occurs is also stored. All beacons are identical. This means that the identity of a beacon cannot be determined from a single measurement.

Introduce an object coordinate system which will be held fixed. The bearing α defined above, depends on the position of the beacon (U_x, U_y) and of the position (P_x, P_y) and orientation P_θ of the scanner according to

$$\alpha(P,U) = \text{atan2}(U_y - P_y, U_x - P_x) - P_\theta \ , \qquad (1)$$

where atan2 is the four quadrant inverse tangent as explained in [6]. The vector (P_x, P_y, P_θ) is called the **camera state**.

The above equation (1) for the measured bearing is non-linear. A somewhat simpler represention of the same equation can be obtained as follows. Introduce alternative representations for the **bearing**

$$\alpha \longleftrightarrow \mathbf{u} = \begin{bmatrix} \cos(\alpha) \\ \sin(\alpha) \end{bmatrix} \ ,$$

for the **beacon position**

$$(U_x, U_y) \longleftrightarrow \mathbf{U} = \begin{bmatrix} U_x \\ U_y \\ 1 \end{bmatrix} \ ,$$

and for the **camera state**

$$(P_x, P_y, P_\theta) \leftrightarrow \mathbf{P} = \begin{bmatrix} \cos(P_\theta) & \sin(P_\theta) \\ -\sin(P_\theta) & \cos(P_\theta) \end{bmatrix} \begin{bmatrix} 1 & 0 & -P_x \\ 0 & 1 & -P_y \end{bmatrix} \ . \qquad (2)$$

Using these notations equation (1) can be written

$$\lambda \mathbf{u} = \mathbf{PU} \ . \qquad (3)$$

The alternative representation for the camera state above will be called the **camera matrix**. Notice that the structure of this 2×3 matrix is

$$\mathbf{P} = \begin{bmatrix} a & b & c \\ -b & a & d \end{bmatrix} \ , \qquad (4)$$

with $a^2 + b^2 = 1$. It is straightforward to obtain the elements (a, b, c, d) of the camera matrix from the meter state (P_x, P_y, P_θ) and vice versa.

It is sometimes useful to consider dual image coordinates

$$\alpha \longleftrightarrow \mathbf{v} = \begin{bmatrix} \sin(\alpha) & \cos(\alpha) \end{bmatrix} \ , \qquad (5)$$

so that $\mathbf{vu} = 0$. This is particularly useful since it simplifies the camera constraint as

$$\lambda \mathbf{vu} = 0 = \mathbf{vPU} \ . \qquad (6)$$

Problem formulation

Motivated by the previous section the structure and motion problem will now be defined.

Problem 2.1. *Given n bearings from m different positions*

$$\mathbf{u}_{i,j}, \quad i = 1, \ldots, m, j = 1, \ldots, n$$

the **surveying problem** *is to find the depths $\lambda_{i,j} > 0$, the reconstructed points $\mathbf{U}_j$ and the camera matrices $\mathbf{P}_i$ such that*

$$\lambda_{i,j} \mathbf{u}_{i,j} = \mathbf{P}_i \mathbf{U}_j, \qquad \forall i = 1, \ldots, m, j = 1, \ldots, n \ .$$

We consider two solutions $(\lambda_{i,j}, U_j, P_i)$ and $(\tilde{\lambda}_{i,j}, \tilde{U}_j, \tilde{P}_i)$ to the surveying problem to be the same if they are related by a similarity transformation, i.e. if there exists a transformation matrix

$$\mathbf{s} = \begin{pmatrix} \mu\cos(\theta) & -\mu\sin(\theta) & A \\ \mu\sin(\theta) & \mu\cos(\theta) & B \\ 0 & 0 & 1 \end{pmatrix}$$

such that $\tilde{U}_j = \mathbf{s} U_j$, $\tilde{P}_i = \mu P_i \mathbf{s}^{-1}$, $\tilde{\lambda}_{i,j} = \mu \lambda_{i,j}$.

In order to understand how much information is needed in order to solve the structure and motion problem, it is useful to calculate the number of degrees of freedom of the problem and the number of constraints

m	n=1	2	3	4	5	6	7
1	0	-1	-2	-3	-4	-5	-6
2	-2	-2	-2	-2	-2	-2	-2
3	-4	-3	-2	-1	0	1	2
4	-6	-4	-2	0	2	4	6
5	-8	-5	-2	1	4	7	10
6	-10	-6	-2	2	6	10	14

Table 1: The number of excess constraints $mn - (2n + 3m - 4)$ for the structure and motion problem with m images of n points.

given by the projection equation. Each object point has two degrees of freedom and each camera state has three. Since it is only possible to obtain solutions to the problem up to a change of euclidean coordinate system and change of scale four degrees of freedom are removed. Each measured bearing gives one constraint on the estimated parameters. Assuming that each point is visible in every camera we get mn constraints. The number of excess constraints $mn-(2n+3m-4)$ is given in Table 1. Disregarding the case of 1 point in 1 image, there are two interesting cases where the number of constraints is exactly equal to the number of degrees of freedom in the estimated parameters. These two cases

1. Three images of five points (m=3, n=5).
2. Four images of four points (m=4, n=4)

will be called the minimal cases of the structure and motion problem.

3 Intersection and the discrete multilinear constraint

In this section the simpler problem of determining the position of an object point using bearings from several known locations is studied. This problem is usually referred to as **intesection** or **reconstruction** in the literature.

Problem 3.1. *Given m bearing directions*

$$\mathbf{u}_i, \quad i = 1, \ldots, m$$

from m known meters states

$$\mathbf{P}_i, \quad i = 1, \ldots, m$$

to one object point $\mathbf{U}$ in unknown position the **intersection problem** *is to find the depths $\lambda_i > 0$ and the object point $\mathbf{U}$ such that*

$$\lambda_i \mathbf{u}_i = \mathbf{P}_i \mathbf{U}, \qquad \forall i = 1, \ldots, m \ .$$

Each measured bearing from known position constrains the location of the object point to the line of sight. The equation for this line is simplest to derive using dual image coordinates $\mathbf{v}$. According to (6) $\mathbf{v}_i \mathbf{P}_i \mathbf{U} = 0$, thus $\mathbf{l}_i = \mathbf{v}_i \mathbf{P}_i$ is the line of sight. The geometric interpretation of the intersection problem is to intesect these m lines $(\mathbf{l}_1, \ldots, \mathbf{l}_m)$ at a point. The problem has no solution using only one measurement, but using two bearings the solution is in general unique.

Algorithm 3.1. (Intersection).

1. *Given m bearing directions $\mathbf{u}_i$, $i = 1, \ldots, m$ from m known meters states $\mathbf{P}_i$, $i = 1, \ldots, m$ form a matrix*
$$M = \begin{pmatrix} \mathbf{v}_1 \mathbf{P}_1 \\ \cdots \\ \mathbf{v}_m \mathbf{P}_m \end{pmatrix}$$
2. *Perform a singular value decomposition of M,*
$$M = U \Sigma V$$
3. *the last column of V is the vector $\mathbf{U}$ that minimises $M\mathbf{U}$, i.e. the solution to the intersection problem that minimises the errors in $\mathbf{l}_i \mathbf{U} = 0$.*

The calibrated trilinear tensor

The case of three cameras is of particular importance. Using three measured bearings from three different known location, the object point is found by intersecting three lines. This is only possible if the three lines actually do intersect. This gives an additional constraint, which 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 bearing directions to the same object point from three different camera states. Then the trilinear constraint*

$$\sum_{i,j,k} T_{i,j,k} \mathbf{u}_{1,J}^i \mathbf{u}_{2,J}^j \mathbf{u}_{3,J}^k = 0 \ , \tag{7}$$

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

Proof. By lining up the camera equations

$$\underbrace{\begin{pmatrix} \mathbf{P}_1 & \mathbf{u}_{1,J} & \mathbf{0} & \mathbf{0} \\ \mathbf{P}_2 & \mathbf{0} & \mathbf{u}_{2,J} & \mathbf{0} \\ \mathbf{P}_3 & \mathbf{0} & \mathbf{0} & \mathbf{u}_{3,J} \end{pmatrix}}_{M} \begin{pmatrix} \mathbf{U}_J \\ -\lambda_{1,J} \\ -\lambda_{2,J} \\ -\lambda_{3,J} \end{pmatrix} = \mathbf{0} \tag{8}$$

we see that the 6×6 matrix M has a non-trivial right-nullspace. Therefore its determinant is zero. Since the determinant is linear in each column it follows that

$$\det M = \sum_{i,j,k} T_{i,j,k} \mathbf{u}_{1,J}^i \mathbf{u}_{2,J}^j \mathbf{u}_{3,J}^k = 0 \ , \tag{9}$$

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

The calibrated trilinear tensor $T = T_{i,j,k}$ in (7) will now be analysed in more detail. Note that the constraint above only involves the *motion* parameters and the bearing directions. It does not involve the *structure* parameters $\mathbf{U}$. The tensor components can be calculated from the *motion* parameters. In fact they are subdeterminants of the first three columns of the matrix M. If we denote the rows of camera matrix $\mathbf{P}_I$ by $\mathbf{P}_I^1\ \mathbf{P}_I^2$, the components can be obtained as

$$T_{ijk} = \wedge_{ii'} \wedge_{jj'} \wedge_{kk'} \det \begin{bmatrix} \mathbf{P}_1^{i'} \\ \mathbf{P}_2^{j'} \\ \mathbf{P}_3^{k'} \end{bmatrix} \quad . \qquad (10)$$

where the tensor $\wedge$ is defined as $\wedge = \begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}$. If the object coordinate system is changed

$$\mathbf{P}_1 \mapsto \tilde{\mathbf{P}}_1 = \mathbf{P}_1\mathbf{s},\ \mathbf{P}_2 \mapsto \tilde{\mathbf{P}}_2 = \mathbf{P}_2\mathbf{s},\ \mathbf{P}_3 \mapsto \tilde{\mathbf{P}}_3 = \mathbf{P}_3\mathbf{s}\,,$$

where $\mathbf{s} \in \mathcal{S}$ denotes a 3×3 transformation matrix, the tensor components change according to

$$\tilde{T}_{i,j,k} = \wedge_{ii'} \wedge_{jj'} \wedge_{kk'} \det \begin{bmatrix} \tilde{\mathbf{P}}_1^{i'}\mathbf{s} \\ \tilde{\mathbf{P}}_2^{j'}\mathbf{s} \\ \tilde{\mathbf{P}}_3^{k'}\mathbf{s} \end{bmatrix} = \wedge_{ii'} \wedge_{jj'} \wedge_{kk'} \det \begin{bmatrix} \mathbf{P}_1^{i'} \\ \mathbf{P}_2^{j'} \\ \mathbf{P}_3^{k'} \end{bmatrix} \det \mathbf{s} = \mu T_{i,j,k} \ , \qquad (11)$$

with $\mu = \det \mathbf{s}$. A change of coordinate system only changes the magnitude of the tensor.

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

$$T \sim \widetilde{T}, \quad \exists \mu \neq 0, T = \mu \widetilde{T}$$

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

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

where the equivalence is defined as

$$(\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3) \simeq (\widetilde{\mathbf{P}_1}, \widetilde{\mathbf{P}_2}, \widetilde{\mathbf{P}_3}), \quad \exists \mathbf{s} \in \mathcal{S}, \widetilde{\mathbf{P}_I} \sim \mathbf{P}_I\mathbf{s}, I = 1,2,3 \ . \qquad (12)$$

Thus the above discussion states that the map $(\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3) \mapsto T$ is in fact a well defined map from the manifold of equivalence classes $\mathcal{P}$ to $\mathcal{T}_u$.

Theorem 3.2. *A tensor $T_{i,j,k}$ is a calibrated trilinear tensor if and only if*

$$\begin{aligned} -T_{111} + T_{122} + T_{212} + T_{221} &= 0, \\ T_{112} + T_{121} + T_{211} - T_{222} &= 0 \ . \end{aligned} \qquad (13)$$

When these constraints are fulfilled it is possible to solve (10) for $(\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3)$. There are in general two solutions, possibly non-real. The map (10) is a well defined two-to-one map from $\mathcal{P}$ to tensors in $\mathcal{T}_u$ fulfilling (13).

The proof can be found in [3].

4 The surveying problem for three images

The previous section on the calibrated trilinear tensor has provided us with the tool for solving the structure and motion problem for three cameras of at least five points.

Algorithm 4.1. (Structure and motion from three images).

1. *Given three images of at least five points,*

   $$\mathbf{u}_{i,j}, \quad i = 1, \dots, 3, j = 1, \dots, n, n \geq 5 \ .$$

   Denote by $u^j = u_{1,j}$ by $u^j = \tilde{u}_{1,j}$ and by $\hat{u}^j = u_{1,j}$.

2. *Calculate the trilinear tensor T that fulfills the linear constraints (13) and $\sum_{i,j,k} T_{i,j,k}\mathbf{u}_{1,J}^i\mathbf{u}_{2,J}^j\mathbf{u}_{3,J}^k = 0$, $\forall\, J = 1, \dots n$.*

3. *Calculate the vector $\mathbf{T}$ that minimises $MT = 0$ and fullfills the constraints (13).*

4. *Calculate the two possible solutions to the relative orientation (P_1, P_2, P_3) from $\mathbf{T}$ according to the proof of Theorem 3.2.*

5. *For each solution to the motion calculate structure according using intersection (Algorithm 3.1).*

Note that there is a two-fold ambiguity in the solution of the structure and motion problem, irrespective of the number of corresponding points. Additional points will help in estimating the tensor more accurately but will not be able to rule out one of the two possible solutions to structure and motion problem. The geometric interpretation of the second solutions is based on isogonal conjugacy, c.f. [3]. The calculations above do, however, not take the sign of the directions into account. Thus some of the reconstructed points sometimes have negative depth. This does not, however, guarantee uniqueness in general.

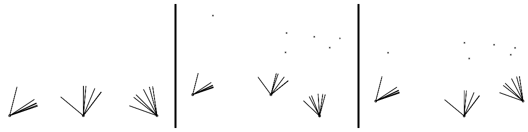

Figure 2: a: Three images. b: The first solution. c: The second solution.

Example 4.1. We illustrate the discussion above with a simple example. Figure 2 shows three images of the same object points. Algorithm 4 above is used to find the two possible solutions to the structure and motion problem. These are also shown in Figure 2. Note that in this example all points in both reconstruction have the correct orientation. ■

5 Resection

In this section the problem of determining the camera state using the bearings to reflectors with known coordinates is studied. This problem is usually referred to as **resection** or **absolute orientation** in the literature.

Problem 5.1. *Given n bearing directions*

$$\mathbf{u}_J, \quad J = 1, \dots, n$$

from an unknown meters state to n object points in known position $\mathbf{U}_j$ *the* **resection problem** *is to find the depths* $\lambda_J > 0$ *and the camera matrix* $\mathbf{P}$ *such that*

$$\lambda_J \mathbf{u}_J = \mathbf{P}\mathbf{U}_J, \qquad \forall J = 1, \dots, n \ .$$

Using the dual coordinates for image bearings (cf. (6)): $0 = \mathbf{v}_J\mathbf{P}\mathbf{U}_J$, a constraint is obtained which is linear in the camera matrix. Denote by $\mathbf{p} = (a, b, c, d)^T$ the vector containing the four parameters of the camera matrix, cf. (4): Introduce the operator $D : \mathbb{R}^3 \to \mathbb{R}^{2\times4}$ according to

$$D\left(\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}\right) = \begin{pmatrix} X & Y & Z & 0 \\ Y & -X & 0 & Z \end{pmatrix}$$

Observe that $\mathbf{PU} = D(\mathbf{U})\mathbf{p}$. Thus it is possible to rearrange the equations so to obtain a linear constraint of the following type. $0 = \mathbf{v}_J\mathbf{P}\mathbf{U}_J = \mathbf{v}_J D(\mathbf{U}_j)\mathbf{p}$. Each measured bearing gives one linear constraint on the elements $\mathbf{p}$ of the camera matrix. Since $P_a^2 + P_b^2 = 1$, the unknown camera matrix has three degrees of freedom. With two bearings the position of the camera is constrained to that of a circle through the two object points. With at least three measured bearings to known positions the solution is in general unique.

Algorithm 5.1. (Resection).

1. *Given n bearing directions* $\mathbf{u}_J$, $J = 1, \dots, n$ *to n known object points* $\mathbf{U}_J$, $J = 1, \dots, n$ *form a matrix*

$$M = \begin{pmatrix} \mathbf{v}_1 D(\mathbf{U}_1) \\ \dots \\ \mathbf{v}_n D(\mathbf{U}_n) \end{pmatrix} \tag{14}$$

2. *Perform a singular value decomposition of M,*

$$M = U\Sigma V$$

where U and V are orthogonal matrices and Σ *is a diagonal matrix with decreasing non-negative numbers on the diagonal.*

3. *The last column of V is the vector* $\mathbf{p}$ *that minimises the constraint* $\mathbf{v}_J\mathbf{P}\mathbf{U}_J = \mathbf{v}_J D(\mathbf{U}_j)\mathbf{p} = 0$, $J = 1, \dots, n$.

6 Solution to the problem of four points in four images

Resection using three points has in general a unique solution. It does not put any extra constraints on the images. However, resection using four points is only possible if some additional constraints are fulfilled. The projection equations

$$\lambda_{IJ}\mathbf{u}_{IJ} = \mathbf{P}_I\mathbf{U}_J = D(\mathbf{U}_J)\mathbf{p}_I$$

for a specific image I can be rewritten as

$$\underbrace{\begin{pmatrix} D(\mathbf{U}_1) & \mathbf{u}_{I,1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ D(\mathbf{U}_2) & \mathbf{0} & \mathbf{u}_{I,2} & \mathbf{0} & \mathbf{0} \\ D(\mathbf{U}_3) & \mathbf{0} & \mathbf{0} & \mathbf{u}_{I,3} & \mathbf{0} \\ D(\mathbf{U}_4) & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{u}_{I,4} \end{pmatrix}}_{N} \begin{pmatrix} \mathbf{p}_I \\ -\lambda_{I,1} \\ -\lambda_{I,2} \\ -\lambda_{I,3} \\ -\lambda_{I,4} \end{pmatrix} = \mathbf{0} \tag{15}$$

Observe that the 8×8 matrix N 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

$$\det N = \sum_{i,j,k,l} Q_{i,j,k,l}\mathbf{u}^i_{I,1}\mathbf{u}^j_{I,2}\mathbf{u}^k_{I,3}\mathbf{u}^l_{I,4} = 0 \ , \tag{16}$$

for some $2\times2\times2\times2$ tensor Q. It is relatively straightforward to see that the tensor components can be obtained as determinants of the first four columns of N. The notation $D(\mathbf{U}_J)^1$ and $D(\mathbf{U}_J)^2$ will be used for the two rows of the matrix $D(\mathbf{U}_J)$. The tensor components can be computed from the object points $(\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3, \mathbf{U}_4)$ as

$$Q_{ijkl} = \wedge_{ii'} \wedge_{jj'} \wedge_{kk'} \wedge_{ll'} \det \begin{bmatrix} D(\mathbf{U}_1)^{i'} \\ D(\mathbf{U}_2)^{j'} \\ D(\mathbf{U}_3)^{k'} \\ D(\mathbf{U}_4)^{l'} \end{bmatrix} . \qquad (17)$$

This proves the following theorem.

Theorem 6.1. *Let $\mathbf{u}_{I,1}$, $\mathbf{u}_{I,2}$, $\mathbf{u}_{I,3}$ and $\mathbf{u}_{I,4}$ be the bearing directions to four* **different points** *in the* **same** *image. Then the dual quadrilinear constraint (16) is fulfilled for some $2 \times 2 \times 2 \times 2$ tensor Q, defined by (17).*

Instead of having a number of constraints involving motion parameters, structure parameters, image measurements and depths we obtain a constraint involving image measurements and structure parameters only.

The tensor components $Q_{i,j,k,l}$ depend on the structure of the four points: $(\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3, \mathbf{U}_4)$. Similar to the calibrated trilinear tensor it can be shown that the mapping from structure parameters to tensor components is a two-to-one mapping.

Definition 6.1. Let the manifold $\mathcal{U}$ of **Euclidean shape** of four points be defined as the set of equivalence classes of four ordered points

$$\mathcal{U} = \left\{ (\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3, \mathbf{U}_4) \,|\, \mathbf{U}_J \sim \begin{pmatrix} X_J \\ Y_J \\ 1 \end{pmatrix} \right\} \Big/ \simeq$$

where the equivalence is defined as

$$(\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3, \mathbf{U}_4) \simeq (\widetilde{\mathbf{U}_1}, \widetilde{\mathbf{U}_2}, \widetilde{\mathbf{U}_3}, \widetilde{\mathbf{U}_4}),$$
$$\exists \mathbf{s} \in \mathcal{S}, \forall I \widetilde{\mathbf{U}_I} \sim \mathbf{s}\mathbf{U}_I \ . \qquad (18)$$

Definition 6.2. Let the manifold $\mathcal{Q}_c$ of dual calibrated quadrilinear tensors be defined as all $2 \times 2 \times 2 \times 2$ tensors Q_{ijkl} defined up to scale fullfilling

$$\begin{aligned} &Q_{1111} = 0, \quad Q_{2222} = 0, \\ &Q_{1112} = -Q_{2221}, \quad Q_{1121} = -Q_{2212}, \\ &Q_{1122} = Q_{2211}, \quad Q_{1211} = -Q_{2122}, \\ &Q_{1212} = Q_{2121}, \quad Q_{1221} = Q_{2112}, \\ &Q_{1222} = -Q_{2111}, \\ &Q_{2122} - Q_{2111} + Q_{2221} + Q_{2212} = 0, \\ &Q_{2211} + Q_{2121} + Q_{2112} = 0 \ . \end{aligned} \qquad (19)$$

Theorem 6.2. *A tensor $Q_{i,j,k,l}$ is a calibrated dual quadrilinear tensor if and only if (19) are fulfilled. When these constraints are fulfilled it is possible to solve (17) for the structure. There are two solutions. The map (17) is a well defined two-to-one map from $\mathcal{U}$ to tensors in $\mathcal{Q}_c$.*

7 Solution to the problem of four points in four images

The previous section on the dual calibrated quadrilinear tensor has provided us with the tool for solving the structure and motion problem for at least four images of four points.

Algorithm 7.1. (Structure and motion from four points).

1. *Given at least four images of four points,*
2. *Calculate the dual quadrilinear tensor fullfilling (19) and the quadrilinear constraints*

$$\sum_{i,j,k,l} Q_{i,j,k,l} \mathbf{u}^i_{I,1} \mathbf{u}^j_{I,2} \mathbf{u}^k_{I,3} \mathbf{u}^l_{I,4} = 0, \quad I = 1, ..., m.$$

3. *Find the two solutions to structure from the tensor components, c.f. Theorem 6.2.*
4. *For each solution to the structure calculate motion using resection (Algorithm 5.1).*

Example 7.1. Using the following bearing measurements

$\alpha_{I,J}$	J			
I	1	2	3	4
1	-2.3562	2.3562	1.1844	-0.76025
2	-2.1588	2.6779	0.78328	-0.99563
3	-2.5536	2.5536	1.7567	-1.1651
4	-2.3562	2.8198	2.0054	-1.312

we obtain two possible solutions on the meter states and the object positions. These are illustrated in Figure 3. Notice that additional images can only be used for estimating the tensor components with greater accuracy. They cannot be used to distinguish between the two solutions.

8 Duality between number of points and number of images

In Section 4 it was shown that the problem of 5 points in 3 images have in general two solutions. It is also shown that if there is a solution to the problem of more than 5 points in 3 images, then there are 2 solutions. Similarly in Section 7 it is shown that the problem of 4 points in 4 images has two solutions. If there is a solution to the problem of 4 points in more than four images then there are 2 solutions. The problem of at least five points in at least 4 images is overdetermined and if there is a solution it is in general unique. The situation is illustrated in Table 2.

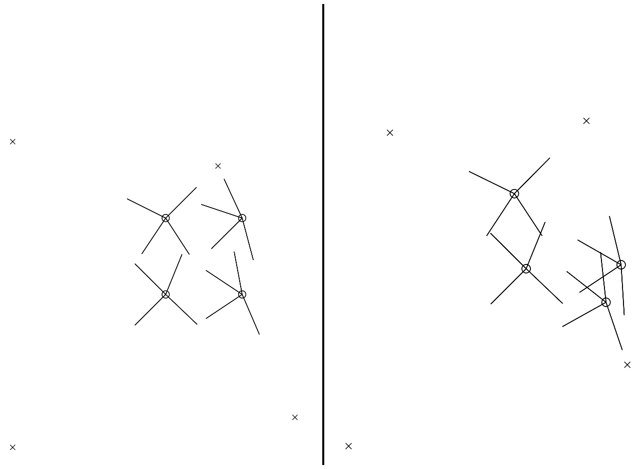

Figure 3: Two different solutions to the structure and motion problem with four images of four points. The image directions are shown as unit vectors from the center of the camera. The object points are shown as small stars.

m \ n	3	4	5	6	7
2	∞	∞	∞	∞	∞
3	∞	∞	2	$2^\star$	$2^\star$
4	∞	2	$1^\star$	$1^\star$	$1^\star$
5	∞	$2^\star$	$1^\star$	$1^\star$	$1^\star$
6	∞	$2^\star$	$1^\star$	$1^\star$	$1^\star$
7	∞	$2^\star$	$1^\star$	$1^\star$	$1^\star$

Table 2: The number of solutions to the surveying problem with m images of n points.

Connection to uncalibrated cameras

In this paper it has been assumed that bearings are measured and therefore that the camera matrix has the special form given by (4). In many situations it can be of interest to study the structure and motion problem for so called uncalibrated cameras. This is identical to the surveying problem, except that the camera matrix is allowed to be a general 2×3 matrix. The difference in the study of minimal cases is however slight, due to the following theorem.

Theorem 8.1. *Knowing that the camera is corrected for internal calibration is equivalent to seeing two extra points (the circular points) in each image.*

Proof. 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

$$C_1 = \begin{pmatrix} 1 \\ I \\ 0 \end{pmatrix} \qquad C_2 = \begin{pmatrix} 1 \\ -I \\ 0 \end{pmatrix}$$

is known to be

$$c_1 = \begin{pmatrix} 1 \\ I \end{pmatrix} \qquad c_2 = \begin{pmatrix} 1 \\ -I \end{pmatrix}$$

since

$$\underbrace{(P_a + P_b I)}_{\lambda_1} c_1 = PC_1, \qquad \underbrace{(P_a - P_b I)}_{\lambda_2} c_2 = PC_2 \ .$$

On the other hand, in the projective case, if we change the image coordinates so that $u_{i,1} = c_1$ and $u_{i,2} = c_2$ and also choose object coordinate system so that $U_1 = C_1$ and $U_2 = C_2$ then the projection matrices must have the form of a calibrated camera matrix since

$$\lambda_1 = P_{1,1} + P_{1,2} I, \qquad \lambda_1 I \quad = P_{2,1} + P_{2,2} I, \qquad (20)$$
$$\lambda_2 = P_{1,1} - P_{1,2} I, \qquad -\lambda_2 I \quad = P_{2,1} - P_{2,2} I \qquad (21)$$

But then we have $(\lambda_1 + \lambda_2) = 2P_{1,1}$ and $(\lambda_1 + \lambda_2) I = 2P_{2,2} I$ so $P_{1,1} = P_{2,2}$. Furthermore, $(\lambda_1 - \lambda_2) = 2P_{1,2} I$ and $(\lambda_1 - \lambda_2) I = 2P_{2,1}$ so $P_{1,2} = -P_{2,1}$. ∎

Corollary 8.1. *The uncalibrated surveying problem with n points and m images is equivalent to the calibrated surveying problem with $n-2$ points in m images.*

Thus it follows that for the uncalibrated surveying problem, the two minimal cases are three images of seven points and four images of six points. In both of these two situations there is an two-fold ambiguity in the solution. The ambiguity is not resolved by adding points in the three image case or adding images in the six point case.

The Table 2 has a symmetric appereance. This can be showed using a technique that Carlsson developed in [4]. The technique that he used for uncalibrated projections from 3D to 2D is here exploited for uncalibrated projections from 2D to 1D:

Theorem 8.2 (Carlsson Duality). *The uncalibrated surveying problem with n points and m images is equivalent to the uncalibrated surveying problem with $m+3$ points and $n-3$ images.*

Proof. This is simplest seen by choosing the image coordinates of the first three points according to

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \mathbf{u}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad \mathbf{u}_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad (22)$$

and the object coordinates of the first three points according to

$$\mathbf{U}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \mathbf{U}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \mathbf{U}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \tag{23}$$

Since $\lambda_i \mathbf{u}_i = \mathbf{P}_i \mathbf{U}_i$ it follows that the camera matrix has the following form

$$\mathbf{P} = \begin{pmatrix} V_1 & 0 & V_3 \\ 0 & V_2 & V_3 \end{pmatrix} .$$

The camera equation for the remaining points

$$\mathbf{u}_{i,j} = \begin{pmatrix} U_1 V_1 + U_3 V_3 \\ U_2 V_2 + U_3 V_3 \end{pmatrix} \tag{24}$$

is the symmetric in camera parameters (V_1, V_2, V_3) and structure parameters (U_1, U_2, U_3). Thus any algorithm for solving n points in m images can be used to solve the $m + 3$ points in $n - 3$ images. ■

Theorem 8.3. *The calibrated surveying problem with n points and m images is equivalent to the calibrated surveying problem with $m + 1$ points and $n - 1$ images.*

Proof. This follows immediately from Theorem 8.1 and Theorem 8.2. ■

According to the theorem the 4-points-in-4-images-problem is equivalent to the 5-points-in-3-images-problem. This explains the symmetry in Table 2.

9 Conclusions

In this paper we have introduced the minimal conditions for solving the structure and motion problem for cameras with one-dimensional retina. The emphasis has been on calibrated cameras.

For the minimal case of three images with five points it was shown how to solve the problem using the calibrated trilinear tensor. It was shown that the there is a two-to-one map from the relative orientation of three cameras to the calibrated trilinear tensor. This explains why there are two solutions to the structure and motion problem for three cameras.

For the minimal case of four images with four points it was shown how to solve the problem using the dual calibrated quadrilinear tensor. It was shown that there is a two-to-one map from the shape of four planar points to this tensor. This explains why there are two solutions to the structure and motion problem for four points.

Finally the connection with the calibrated and the uncalibrated cameras are given. From this it follows that similar results holds for three images with seven points and for four images with six points.

Using the duality of Carlsson it is then shown that the above two types of ambiguities are in fact dual to each other.

Acknowledgments

This work has been supported by, Netzler Dahlgren Co, the ESPRIT Reactive LTR project 21914, CUMULI and the Swedish Research Council for Engineering Sciences (TFR), project 95-64-222

References

[1] M. Armstrong, A. Zisserman, and R. Hartley. Self-calibration from image triplets. In *Proc. 4th European Conf. on Computer Vision, Cambridge, UK*, pages 3–16. Springer-Verlag, 1996.

[2] K. Åström. Automatic mapmaking. In D. Charnley, editor, *Selected Papers from the 1st IFAC International Workshop on Intelligent Autonomous Vehicles, Southampton, UK, 18-21 april 1993*, pages 181–186. Pergamon Press, 1993.

[3] K. Åström. Multilinear constraints in two-dimensional vision and isogonal conjugacy. In *Proc. Symposium on Image Analysis, Lund, Sweden*, pages 34–38, 1996.

[4] S. Carlsson and D. Weinshall. Dual computation of projective shape and camera positions from multiple images. *Int. Journal of Computer Vision*, 1998.

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

[6] The Math Works Inc. *MATLAB Reference Guide.* The Math Works, Inc., 1992.

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

[8] C.C. Slama, editor. *Manual of Photogrammetry.* American Society of Photogrammetry, Falls Church, VA, 4:th edition, 1984.