Minimal Structure and Motion problems for 1D vision

Kalle Åström, Magnus Oskarsson
Centre for Mathematical Sciences,
Mathematical Imaging Group,
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 undegoing so called planar motion. The structure and motion problem for such cameras is studied and the two possible minimal cases is solved. For each solution with three images it can be shown tat there is an ambiguous solution. For each solution for four points there is also an ambiguous solution.

1 Introduction

Understanding of one-dimensional cameras is important in several applications. In [5] 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, 3]. 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 application of autonomous guided vehicles. 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. [4]. 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** or the structure and motion problem. This is the procedure to obtain a map of the unknown positions of the beacons using images at unknown position and orientation.

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. [6].

2 Scanner geometry

A laser navigated vehicle measures the direction (bearing) in a horizontal plane to beacons on the walls.

Introduce the following notation for the **bearing**

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

for the **beacon position** in the plane

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

and for the **camera state**, i.e. the position and orientation of the camera:

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

Using these notations the projection equation is

$$\lambda \mathbf{u} = \mathbf{P}\mathbf{U} . \tag{2}$$

It is sometimes useful to consider dual image coordinates

$$\alpha \longleftrightarrow \mathbf{v} = \begin{bmatrix} -\sin(\alpha) & \cos(\alpha) \end{bmatrix} , \tag{3}$$

so that $\mathbf{vu} = 0$. This is particularly useful since it simplifies the camera constraint (2) as

$$\lambda \mathbf{vu} = 0 = \mathbf{vPU} . \tag{4}$$

Structure and Motion

Motivated by the previous sections 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*

$$\mathbf{U}_J = \begin{pmatrix} X_J \\ Y_J \\ 1 \end{pmatrix}$$

and the camera matrices

$$\mathbf{P}_I = \begin{pmatrix} a_I & b_I & c_I \\ -b_I & a_I & d_I \end{pmatrix}, \tag{5}$$

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 λ is different.

Definition 2.1. The group of similarity transformations is defined as

$$\mathcal{S} = \left\{ \mathbf{s} \sim \begin{pmatrix} C\cos(\theta) & -C\sin(\theta) & A \\ C\sin(\theta) & C\cos(\theta) & B \\ 0 & 0 & 1 \end{pmatrix} \right\} \tag{6}$$

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 surveying problem to be the same if they are related by a similarity transformation. If there exists a transformation matrix $\mathbf{s}$ such that

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

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

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

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

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

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

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 given by the projection equation. Each object point has two degrees of freedom and each camera state has three. The solution is only defined up to a similarity transformation (6). This manifold $\mathcal{S}$ has dimension 4. Using n points and m cameras we thus have $2n+3m-4$ degrees of freedom in the parameters. 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:

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

	n						
m	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

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 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*

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

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

Some properties of the calibrated trilinear tensor $T = T_{i,j,k}$ in (7) will now be analysed in more detail. The tensor components can be calculated from the *motion* parameters alone. If we denote the rows of camera matrix $\mathbf{P}_I$ by $\mathbf{P}^1_I$ $\mathbf{P}^2_I$ it can be shown that

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

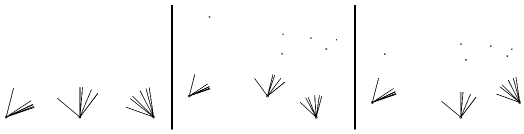

Figure 1: a: Three images. b: The first solution. c: The second solution.

where the tensor $\wedge$ is defined as

$$\wedge = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} . \qquad (8)$$

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

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

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

$$\begin{aligned} -T_{111} + T_{122} + T_{212} + T_{221} = 0, \\ T_{112} + T_{121} + T_{211} - T_{222} = 0\,. \end{aligned} \qquad (9)$$

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

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

Theorem 3.2. *Let $\mathcal{T}$ and $\mathcal{P}$ be defined as above then the map*

$$T : \mathcal{P} \longrightarrow \mathcal{T}$$

$$T(\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3)_{ijk} = \wedge_{ii'} \wedge_{jj'} \wedge_{kk'} \det \begin{bmatrix} \mathbf{P}_1^{i'} \\ \bar{\mathbf{P}}_2^{j'} \\ \hat{\mathbf{P}}_3^{k'} \end{bmatrix} \qquad (10)$$

is a well defined two-to-one mapping.

This can be used to solve the structure and motion problem for three images of at least five points.

4 The dual quadrilinear tensor

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

$$\underbrace{\begin{pmatrix} D(\mathbf{U}_1) & \mathbf{u}_{I,1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ D(\mathbf{U}_2) & \mathbf{0} & \mathbf{u}_{I,2} & \mathbf{0} & \mathbf{0} \\ D(\mathbf{U}_3) & \mathbf{0} & \mathbf{0} & \mathbf{u}_{I,3} & \mathbf{0} \\ D(\mathbf{U}_4) & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{u}_{I,4} \end{pmatrix}}_{M} \begin{pmatrix} \mathbf{p}_I \\ -\lambda_{I,1} \\ -\lambda_{I,2} \\ -\lambda_{I,3} \\ -\lambda_{I,4} \end{pmatrix} = \mathbf{0} \qquad (11)$$

Observe that the 8×8 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

$$\det M = \sum_{i,j,k,l} Q_{i,j,k,l} \mathbf{u}^i_{I,1} \mathbf{u}^j_{I,2} \mathbf{u}^k_{I,3} \mathbf{u}^l_{I,4} = 0\,, \qquad (12)$$

for some $2 \times 2 \times 2 \times 2$ tensor Q. It is relatively straightforward to see that the tensor components can be obtained as subdeterminants of the first four rows of M. 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

$$Q((\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3, \mathbf{U}_4))_{ijkl} = \wedge_{ii'} \wedge_{jj'} \wedge_{kk'} \wedge_{ll'} \det \begin{bmatrix} D(\mathbf{U}_1)^{i'} \\ D(\mathbf{U}_2)^{j'} \\ D(\mathbf{U}_3)^{k'} \\ D(\mathbf{U}_4)^{l'} \end{bmatrix} . \qquad (13)$$

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 we will show that the mapping from structure parameters to tensor components is a two-to-one mapping.

Definition 4.1. 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

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

Table 2: The number of solutions to the surveying problem with m images of n points. Superscripted stars indicate overdetermined situations.

	n				
m	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	∞	$2^\star$	$1^\star$	$1^\star$	$1^\star$

Definition 4.2. 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

$$\begin{aligned}
&Q_{1111} = 0, \\
&Q_{2222} = 0, \\
&Q_{1112} = -Q_{2221}, \\
&Q_{1121} = -Q_{2212}, \\
&Q_{1122} = Q_{2211}, \\
&Q_{1211} = -Q_{2122}, \\
&Q_{1212} = Q_{2121}, \\
&Q_{1221} = Q_{2112}, \\
&Q_{1222} = -Q_{2111}, \\
&Q_{2122} - Q_{2111} + Q_{2221} + Q_{2212} = 0, \\
&Q_{2211} + Q_{2121} + Q_{2112} = 0 \ .
\end{aligned} \tag{14}$$

Theorem 4.1. *A tensor $Q^{i,j,k,l}$ is a calibrated dual quadrilinear tensor if and only if (14) are fulfilled. When these constraints are fulfilled it is possible to solve (13) for the structure. There are two solutions.*

The Table 2 has a symmetric appereance. This can be shown using a duality that was developed by Carlsson in [2]. According to this duality the 4-points-in-4-images-problem is equivalent to the 5-points-in-3-images-problem. This explains the symmetry in Table 2.

5 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.

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] S. Carlsson. Duality of reconstruction and positioning from projective views. In *IEEE Workshop on Representation of Visual Scenes*. IEEE, 1995.

[3] 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 *Proc. 5th European Conf. on Computer Vision, Freiburg, Germany*, pages 36–52. Springer-Verlag, 1998.

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

[5] 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.

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