Duality of Reconstruction and Positioning from Projective Views

Stefan Carlsson*

Computational Vision and Active Perception Laboratory (CVAP)
KTH (Royal Institute of Technology), Stockholm, Sweden

Abstract

Given multiple image data from a set of points in 3-D, there are two fundamental questions that can be addressed:

- *What is the structure of the set of points in 3-D ?*
- *What are the positions of the cameras relative to the points ?*

In this paper we show that for projective views and with structure and position defined modulo linear transformations these problems are are dual in the sense that their solution arises from constraint equations where space point and camera positions occur in a reciprocal way. The problem of computing camera positions from m *points in* n *views can be solved with the same algorithm as the problem of directly reconstructing* $n+4$ *points in* $m-4$ *views. This unifies different approaches for projective reconstruction: methods based on external calibration and direct methods exploiting constraints that exist between space and image invariants.*

1 Introduction

The problems of determining the position of cameras relative a scene and the 3-D reconstruction of the scene from image data have traditionally been treated as separate problems with relative camera positioning preceding the reconstruction. Camera positioning or external calibration for un calibrated cameras is based on the determination of epipolar geometry [7] , [10], [20]. Epipolar geometry of two cameras is captured by the fundamental matrix [7], which can be used to constrain image coordinates of points in two images of the same scene. Projective reconstruction is then achieved using the epipolar information in various alternative ways. Computation of projection matrices [6], [18], cross-ratios [9] or more direct methods expressing 3-D invariants directly in terms of the fundamental matrix [2] ,[4]. Recently, initiated by the work in [21], [23], the generalisation of epipolar constraints to the case of multiple cameras has received widespread attention [8], [11], [12], [15], [22], [27]. The bilinear image constraints in the case of two images are then generalized to multi linear constraints between multiple images.

Interestingly, there are alternative direct ways for projective reconstruction, not relying on the computation of epipolar geometry [13], [24], [26]. Especially in the case of constrained scenes, direct methods are powerful [3] , [19], [25], [28] in the sense that fewer points are needed for reconstruction. Constraining the scene also makes possible the computation of epipolar geometry with fewer points [5], [17], [28]. The direct methods for reconstruction exploit constraints existing between coordinates of space points and their image coordinates, not involving camera geometry.

The existence of basically two alternative methods for achieving projective reconstruction naturally poses the question whether there is a relation between them. In this paper we will demonstrate that this is indeed the case. Essentially what we will show is this:

> There exists joint constraint equations or canonical projection equations, involving projective coordinates of space points, camera positions and image coordinates. In these constraints, space points and camera positions appear symmetrically so that elimination of either gives rise to constraint equations between space points and image coordinates or between camera positions and image coordinates that have exactly the same mathematical structure. The problems computing scene structure and camera positions from image data are therefore dual in the sense that they can be solved with the same algorithm depending on the number of space points and cameras.

*This work was partially performed under the ESPRIT-BRA project VIVA, and with support from the Swedish National Board for Industrial and Technical Development, NUTEK.
Mail address: NADA, KTH, S-100 44 Stockholm, Sweden.
Email: stefanc@bion.kth.se

We will see that the method proposed corresponds to the computation of a fundamental matrix and tensors in a canonical image coordinate system, similar to that presented in [12], . This F-matrix will be parametrized by camera positions only. For the reconstruction case, the constraints between space and image coordinates can be written using a dual of the F-matrix or tensor. In the same way as the F-matrix s constrains points in different images to lie along epipolar lines, the dual of the F- matrix will constrain points in the same image to lie along lines defined by the structure of the space points.
More specifically we will show that the problem of computing camera positions from m points in n views is mathematically identical to the problem of reconstructing $n+4$ points in $m-4$ views. As an example we will see that the constraints used in [13] for computing projective structure from 6 points in 3 images are dual to epipolar constraints for 7 points in 2 images. Similarly the 8-points linear algorithm for computing epipolar geometry between 2 views can be used for direct projective reconstruction of 6 points in 4 views.

2 Linearly invariant representations

The basic relation that will be derived is the constraint relation involving linearly invariant representation of space points, camera positions and image coordinates. This will be done using 5 space points as a projective basis and expressing all other space points and camera positions in this basis with projective coordinates. In each image we choose 4 corresponding image points as a projective basis and express each image point in this basis. The homogeneous projective coordinates are then a linearly invariant representation of space and image points in the sense that ratios of homogeneous coordinate components are absolute linear invariants.
As a consequence of this representation, cameras will be represented by the position of their projection points in space only. Image plane positions do not enter at all in the relations. The reason why the position of image planes are irrelevant when using linearly invariant coordinate representations can be seen from fig. 1. For two arbitrary choices of image planes to a given projection point P_0 the image coordinates will be linearly related. The projective coordinate representation is therefore unaffected by this choice.
The central issue is how to relate the invariant representations chosen for space points , camera positions and image points. In order to do this, we will make use of the very concise relation that exists between an absolute linear invariant of five points in an image and the six point invariant constructed from the five corresponding space points together with the camera position.

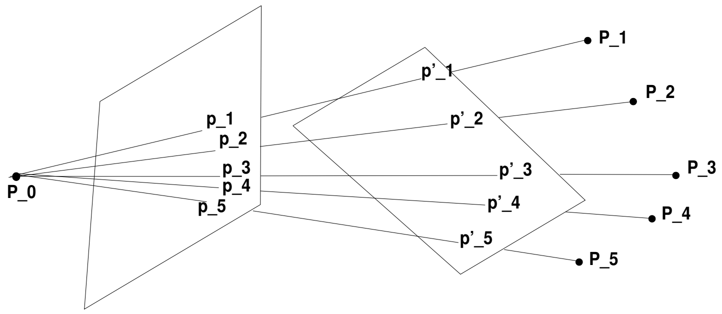


Figure 1: *For two arbitrary image planes image coordinates are related by a projective transformation, Choosing a linearly invariant representation for image data means that the position of the image plane is irrelevant*

$$\frac{[\,p_1\ p_2\ p_4\,]\ \ [\,p_1\ p_3\ p_5\,]}{[\,p_1\ p_3\ p_4\,]\ \ [\,p_1\ p_2\ p_5\,]} = \frac{[\,P_0\ P_1\ P_2\ P_4\,]\ [\,P_0\ P_1\ P_3\ P_5\,]}{[\,P_0\ P_1\ P_3\ P_4\,]\ [\,P_0\ P_1\ P_2\ P_5\,]} \quad (1)$$

The brackets $[\ldots]$ denote determinants and the p_i:s and P_i:s are image and space homogeneous coordinates in arbitrary reference frames respectively. This relation is derived in the appendix. A slightly different but equivalent form of this relation was derived in [1]. This relation can be said to connect linearly invariant structure of camera positions and space points with that of the image points. All relations and constraints that will be derived will be linearly invariant polynomials in image and space coordinates expressed in arbitrary coordinate systems.

3 Constraints between space points, camera positions and image coordinates

Using the relation (1) we will now derive invariant relations between points in 3-D, camera positions and image coordinates. For that purpose we take five 3-D points, with not four of them coplanar, with coordinates $P_1^*, P_2^*, P_3^*, P_4^*, P_5^*$ as a projective basis, where $*$ means that we have fixed the scale factor of the homogeneous coordinates. Any 3-D point P_n can then be expressed in this basis as:

$$P_n = X_n P_1^* \ + \ Y_n P_2^* \ + \ Z_n P_3^* \ + \ W_n P_4^* \quad (2)$$

In the same way, the camera position can be expressed as:

$$P_0 = X_0 P_1^* \ + \ Y_0 P_2^* \ + \ Z_0 P_3^* \ + \ W_0 P_4^* \quad (3)$$

The scale factors of the basis are determined by the requirement that points P_1, P_2, P_3, P_4 be unit vectors

in this representation and point P_5 has coordinates $(1,\ 1,\ 1,\ 1)$
In the same way we take the projective basis for image coordinates in an image using images of the four points P_1, P_2, P_3, P_4 with coordinates $p_1^*, p_2^*, p_3^*, p_4^*$ with no three of them collinear. A point in the image can then be expressed as:

$$p_n \ = \ x_n p_1^* \ + \ y_n p_2^* \ + \ w_n p_3^* \tag{4}$$

Points p_1, p_2, p_3 will then be unit vectors and point p_4 will have projective coordinates $(1,\ 1,\ 1)$.
The relation (1) is valid for arbitrary choices of image and space coordinate system. Writing the relation as:

$$\frac{[p_1\ p_2\ p_4]\ \ [p_1\ p_3\ p_n]}{[p_1\ p_3\ p_4]\ \ [p_1\ p_2\ p_n]} = \frac{[P_0\ P_1\ P_2\ P_4]\ [P_0\ P_1\ P_3\ P_n]}{[P_0\ P_1\ P_3\ P_4]\ [P_0\ P_1\ P_2\ P_n]} \tag{5}$$

and using the projective coordinate representations:

p_1	p_2	p_3	p_4	p_n
1	0	0	1	x_n
0	1	0	1	y_n
0	0	1	1	w_n

P_0	P_1	P_2	P_3	P_4	P_5	P_n
X_0	1	0	0	0	1	X_n
Y_0	0	1	0	0	1	Y_n
Z_0	0	0	1	0	1	Z_n
W_0	0	0	0	1	1	W_n

we get:

$$\frac{y_n}{w_n} \ = \ \frac{Z_0\ (\ W_0 Y_n\ - Y_0 W_n)}{Y_0\ (W_0 Z_n\ -\ Z_0 W_n)} \ = \ \frac{\frac{Y_n}{Y_0} - \frac{W_n}{W_0}}{\frac{Z_n}{Z_0} - \frac{W_n}{W_0}}$$

For five image points we can compute two independent linear invariants. By a simple permutation of the coordinates of (5) we get for the second invariant:

$$\frac{x_n}{w_n} \ = \ \frac{Z_0\ (\ W_0 X_n\ - X_0 W_n)}{X_0\ (W_0 Z_n\ -\ Z_0 W_n)} \ = \ \frac{\frac{X_n}{X_0} - \frac{W_n}{W_0}}{\frac{Z_n}{Z_0} - \frac{W_n}{W_0}}$$

These two equations actually contain all linear invariant information relating 3-D coordinates of points and viewpoint and image coordinates in arbitrary coordinate systems. They can also be written as "projection" equations:

$$\begin{pmatrix} x_n \\ y_n \\ w_n \end{pmatrix} = \sigma \begin{pmatrix} X_0^{-1} & 0 & 0 & W_0^{-1} \\ 0 & Y_0^{-1} & 0 & W_0^{-1} \\ 0 & 0 & Z_0^{-1} & W_0^{-1} \end{pmatrix} \begin{pmatrix} X_n \\ Y_n \\ Z_n \\ W_n \end{pmatrix}$$

A simple rewriting allows us to write these relations in the form of constraint equations as:

$$w_n \frac{Y_n}{Y_0} \ - \ y_n \frac{Z_n}{Z_0} \ + \ (y_n - w_n) \frac{W_n}{W_0} \ = \ 0 \tag{6}$$

In a similar way we can derive:

$$w_n \frac{X_n}{X_0} \ - \ x_n \frac{Z_n}{Z_0} \ + \ (x_n - w_n) \frac{W_n}{W_0} \ = \ 0 \tag{7}$$

The very important observation we can make is that these relations are unaffected by the substitutions:

$$(X_n, Y_n, Z_n, W_n) \Leftrightarrow (X_0^{-1}, Y_0^{-1}, Z_0^{-1}, W_0^{-1})$$

This is the basis of the duality of the problems of computing positions of points in space, $X_n \ldots$ and camera positions $X_0 \ldots$ from image data.

3.1 Projective reconstruction from known camera positions and two images

Using eq. (6) and (7) we can solve for the projective coordinates of a point given the images in two cameras with known positions. From eq. (6) and (7) we get:

$$\begin{pmatrix} 0 & w_n^a Y_{0,a}^{-1} & -y_n^a Z_{0,a}^{-1} & (y_n^a - w_n^a) W_{0,a}^{-1} \\ w_n^a X_{0,a}^{-1} & 0 & -x_n^a Z_{0,a}^{-1} & (x_n^a - w_n^a) W_{0,a}^{-1} \\ 0 & w_n^b Y_{0,b}^{-1} & -y_n^b Z_{0,b}^{-1} & (y_n^b - w_n^b) W_{0,b}^{-1} \\ w_n^b X_{0,b}^{-1} & 0 & -x_n^b Z_{0,b}^{-1} & (x_n^b - w_n^b) W_{0,b}^{-1} \end{pmatrix} \begin{pmatrix} X_n \\ Y_n \\ Z_n \\ W_n \end{pmatrix} = 0 \tag{8}$$

where we have used indexes a and b to denote data from the two images.

3.2 Relative positioning from known space points and one image

In a very similar way the camera position for one camera can be computed given a known configuration of six points in space. From eq. (6) and (7) we get:

$$\begin{pmatrix} 0 & w_m Y_m & -y_m Z_m & (y_m - w_m) W_m \\ w_m X_m & 0 & -x_m Z_m & (x_m - w_m) W_m \\ 0 & w_n Y_n & -y_n Z_n & (y_n - w_n) W_n \\ w_n X_n & 0 & -x_n Z_n & (x_n - w_n) W_n \end{pmatrix} \begin{pmatrix} X_0^{-1} \\ Y_0^{-1} \\ Z_0^{-1} \\ W_0^{-1} \end{pmatrix} = 0 \quad (9)$$

Using points $m \neq n$ where $m > 4$ and $n > 4$ we can solve for the projective coordinates of the camera positions in terms of image and space point coordinates in the same way as in the projective reconstruction above.
We se that the problems of projective reconstruction from known camera positions and camera positioning from known space points are mathematically identical which of course is a consequence of the structure of the basic constraint relations of (6) and (7). We will se that also for the constraint relations we get by eliminating space points or camera positions respectively, we get mathematically identical expressions.

4 Camera position and space point constraints from image measurements

From eq. (8) and (9) we can derive camera-image point and space-image point constraints by noting that the 4×4 matrices in general have rank 3.

4.1 Constraints between camera positions and image coordinates

From (8) we get:

$$det \begin{pmatrix} 0 & w_n^a Y_{0,a}^{-1} & -y_n^a Z_{0,a}^{-1} & (y_n^a - w_n^a) W_{0,a}^{-1} \\ w_n^a X_{0,a}^{-1} & 0 & -x_n^a Z_{0,a}^{-1} & (x_n^a - w_n^a) W_{0,a}^{-1} \\ 0 & w_n^b Y_{0,b}^{-1} & -y_n^b Z_{0,b}^{-1} & (y_n^b - w_n^b) W_{0,b}^{-1} \\ w_n^b X_{0,b}^{-1} & 0 & -x_n^b Z_{0,b}^{-1} & (x_n^b - w_n^b) W_{0,b}^{-1} \end{pmatrix} = 0 \quad (10)$$

After factoring out common factors this can be written as the bilinear form:

$$\begin{pmatrix} x_n^a \\ y_n^a \\ w_n^a \end{pmatrix}^T \begin{pmatrix} 0 & f_2 - f_1 & f_3 - f_2 \\ f_4 - f_5 & 0 & f_5 - f_3 \\ f_6 - f_4 & f_1 - f_6 & 0 \end{pmatrix} \begin{pmatrix} x_n^b \\ y_n^b \\ w_n^b \end{pmatrix} = 0 \quad (11)$$

where:

$$f_1 = X_{0,a}\, Y_{0,b}\, Z_{0,a}\, W_{0,b} \quad f_2 = X_{0,a}\, Y_{0,b}\, Z_{0,b}\, W_{0,a}$$
$$f_3 = X_{0,a}\, Y_{0,a}\, Z_{0,b}\, W_{0,b} \quad f_4 = X_{0,b}\, Y_{0,a}\, Z_{0,a}\, W_{0,b}$$
$$f_5 = X_{0,b}\, Y_{0,a}\, Z_{0,b}\, W_{0,a} \quad f_6 = X_{0,b}\, Y_{0,b}\, Z_{0,a}\, W_{0,a}$$

This is the epipolar constraint on the projective image coordinates in image a and image b. The 3×3 matrix F, is the fundamental matrix parametrized by the positions of cameras a and b. The choice of the basis for the projective image coordinates ensures that the fundamental matrix is determined completely by the camera positions and no other camera parameters.
Using the fact that we know that $(X_5,\ Y_5,\ Z_5,\ W_5) = (1,\ 1,\ 1,\ 1)$ we also get the constraints:

$$w_n^a\ Y_{0,a}^{-1} -\ y_n^a\ Z_{0,a}^{-1} +\ (y_n^a - w_n^a)\ W_{0,a}^{-1} =\ 0$$
$$w_n^a\ X_{0,a}^{-1} -\ x_n^a\ Z_{0,a}^{-1} +\ (x_n^a - w_n^a)\ W_{0,a}^{-1} =\ 0$$
$$w_n^b\ Y_{0,b}^{-1} -\ y_n^b\ Z_{0,b}^{-1} +\ (y_n^b - w_n^b)\ W_{0,b}^{-1} =\ 0$$
$$w_n^b\ X_{0,b}^{-1} -\ x_n^b\ Z_{0,b}^{-1} +\ (x_n^b - w_n^b)\ W_{0,b}^{-1} =\ 0 \quad (12)$$

These constraints can be used to obtain camera positions from the elements of the F-matrix as will be shown in the next section.

4.2 Constraints between space points and image coordinates

For two points in one image we get from (9):

$$det \begin{pmatrix} 0 & w_m Y_m & -y_m Z_m & (y_m - w_m) W_m \\ w_m X_m & 0 & -x_m Z_m & (x_m - w_m) W_m \\ 0 & w_n Y_n & -y_n Z_n & (y_n - w_n) W_n \\ w_n X_n & 0 & -x_n Z_n & (x_n - w_n) W_n \end{pmatrix} = 0 \quad (13)$$

Because of the duality in the structure of the equations for space points and camera positions we can write the space - image point constraint in a form similar to the epipolar constraint

$$\begin{pmatrix} x_m \\ y_m \\ w_m \end{pmatrix}^T \begin{pmatrix} 0 & g_2 - g_1 & g_3 - g_2 \\ g_4 - g_5 & 0 & g_5 - g_3 \\ g_6 - g_4 & g_1 - g_6 & 0 \end{pmatrix} \begin{pmatrix} x_n \\ y_n \\ w_n \end{pmatrix} = 0 \quad (14)$$

where:

$$g_1 \;=\; X_n\ Y_m\ Z_n\ W_m\ g_2 \;=\; X_n\ Y_m\ Z_m\ W_n$$

$$g_3 \;=\; X_n\ Y_n\ Z_m\ W_m\ g_4 \;=\; X_m\ Y_n\ Z_n\ W_m$$

$$g_5 \;=\; X_m\ Y_n\ Z_m\ W_n\ g_6 \;=\; X_m\ Y_m\ Z_n\ W_n$$

Given known space points, this equation constrains an image point to lie on a line determined by the space points and five image points in the very same way as epipolar lines constrain the position of image points in multiple images. Note that for $m, n < 5$ the constraints are trivially fulfilled.
For $m = 5$ we have $(X_5,\ Y_5,\ Z_5,\ W_5) \;=\; (1,\ 1,\ 1,\ 1)$. Writing out the bilinear form of eq. (14) we get:

$$(w_m y_n - x_m y_n)X_n Z_n + \;(x_m y_n - x_m w_n)X_n W_n +$$
$$+\;(x_m w_n - y_m w_n)X_n Y_n + \;(y_m x_n - w_m x_n)Y_n Z_n +$$
$$+\;(y_m w_n - y_m x_n)Y_n W_n + \;(w_m x_n - w_m y_n)Z_n W_n = 0$$

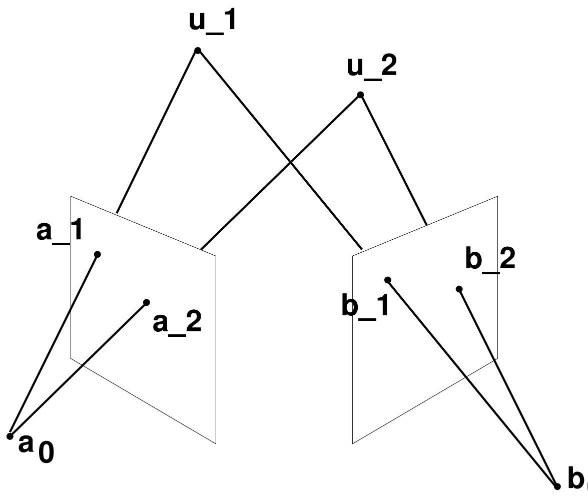
This is exactly the same relation as that in [13] where it was used to compute absolute invariants for 6 points from 3 images.
The geometric basis for the constraint relations on space-image points and camera-image points can bee seen from fig. 2. The camera-image point or epipolar constraint follows from the four point coplanarity of the camera positions and image points in 3-D, while the space-image point constraints result from coplanarities of space points and image points.

5 Computing camera and space point projective coordinates from constraint matrices F and G

The constraints (11) and (14) have exactly the same form if we make the identifications $a \leftrightarrow m$ and $b \leftrightarrow n$. The problem of computing the projective coordinates for the camera positions from the fundamental matrix (11) is therefore equivalent to the problem of computing projective coordinates for space points from the space image point constraint matrix (14).
The elements of the fundamental matrix can be factored and by making the substitutions:

$$X \;=\; \frac{X_{0,a}\ W_{0,b}}{X_{0,b}\ W_{0,a}}\ \ Y \;=\; \frac{Y_{0,a}\ W_{0,b}}{Y_{0,b}\ W_{0,a}}\ \ Z \;=\; \frac{Z_{0,a}\ W_{0,b}}{Z_{0,b}\ W_{0,a}}$$

it can be written:



Epipolar constraint: [a_0 b_0 a_1 b_1] = [a_0 b_0 a_2 b_2] = 0
Space–image point constraint: [u_1 u_2 a_1 a_2] = [u_1 u_2 b_1 b_2] = 0

Figure 2: *Geometric interpretation of epipolar and space-image point constraints in terms of four point coplanarities*

$$F \;=\; \begin{pmatrix} 0 & X(1-Z) & -X(1-Y) \\ -Y(1-Z) & 0 & Y(1-X) \\ Z(1-Y) & -Z(1-X) & 0 \end{pmatrix} \;=\; 0$$

Image points with index ≥ 5 will provide constraints for computations of the elements of this matrix. Given 7 points in 2 images or 6 points in 3 images we get 3 nonlinear equations for X, Y and Z for the camera position and space point problem respectively. These non-linear equations are exactly those used in [13], where it was shown that multiple solutions can be obtained by solving an equation of third degree. We see that this technique which was developed for the space point reconstruction case can equally well be applied to the camera positioning problem.
If we want linear solutions we can take 8 points in 2 images or 6 points in 4 images for camera position and space points respectively. Together with the linear constraint:

$$\sum f_{i,j} \;=\; 0$$

we get 5 linear constraint equations for the 6 F-matrix elements $f_{i,j}$. After solving these equation X, Y and Z can be expressed in closed form in the elements $f_{i,j}$. The parameters X, Y and Z give ratios of homogeneous coordinates for either camera positions or space points.For the camera positions we have the auxiliary constraints (12), from which we can obtain projective camera coordinates. For the space point reconstruction we note that X, Y and Z will give the projective coordinates for the 6:th point directly since $(X_5,\ Y_5,\ Z_5,\ W_5) = (1,\ 1,\ 1,\ 1)$

6 Duality of reconstruction and positioning for multiple images

The constraints (10) and (13) can be written for arbitrary number of cameras $a, b, c \ldots$ and arbitrary number of points $5, 6, 7, \ldots$ respectively. Instead of 4×4 matrices we get rectangular four column matrices with rank ≤ 3. For one point we get:

$$\begin{pmatrix} 0 & \ldots_a Y_{0,a}^{-1} & \ldots_a Z_{0,a}^{-1} & \ldots_a W_{0,a}^{-1} \\ \ldots_a X_{0,a}^{-1} & 0 & \ldots_a Z_{0,a}^{-1} & \ldots_a W_{0,a}^{-1} \\ 0 & \ldots_b Y_{0,b}^{-1} & \ldots_b Z_{0,b}^{-1} & \ldots_b W_{0,b}^{-1} \\ \ldots_b X_{0,b}^{-1} & 0 & \ldots_b Z_{0,b}^{-1} & \ldots_b W_{0,b}^{-1} \\ 0 & \ldots_c Y_{0,c}^{-1} & \ldots_c Z_{0,c}^{-1} & \ldots_c W_{0,c}^{-1} \\ \ldots_c X_{0,c}^{-1} & 0 & \ldots_c Z_{0,c}^{-1} & \ldots_c W_{0,c}^{-1} \\ & & \cdot & \\ & & \cdot & \\ & & \cdot & \end{pmatrix} \begin{pmatrix} X_n \\ Y_n \\ Z_n \\ W_n \end{pmatrix} = 0$$

Similarly, for one camera we get:

$$\begin{pmatrix} 0 & \ldots_a Y_5 & \ldots_a Z_5 & \ldots_a W_5 \\ \ldots_a X_5 & 0 & \ldots_a Z_5 & \ldots_a W_5 \\ 0 & \ldots_a Y_6 & \ldots_a Z_6 & \ldots_a W_6 \\ \ldots_a X_6 & 0 & \ldots_a Z_6 & \ldots_a W_6 \\ 0 & \ldots_a Y_7 & \ldots_a Z_7 & \ldots_a W_7 \\ \ldots_a X_7 & 0 & \ldots_a Z_7 & \ldots_a W_7 \\ & & \cdot & \\ & & \cdot & \\ & & \cdot & \end{pmatrix} \begin{pmatrix} X_{0,a}^{-1} \\ Y_{0,a}^{-1} \\ Z_{0,a}^{-1} \\ W_{0,a}^{-1} \end{pmatrix} = 0$$

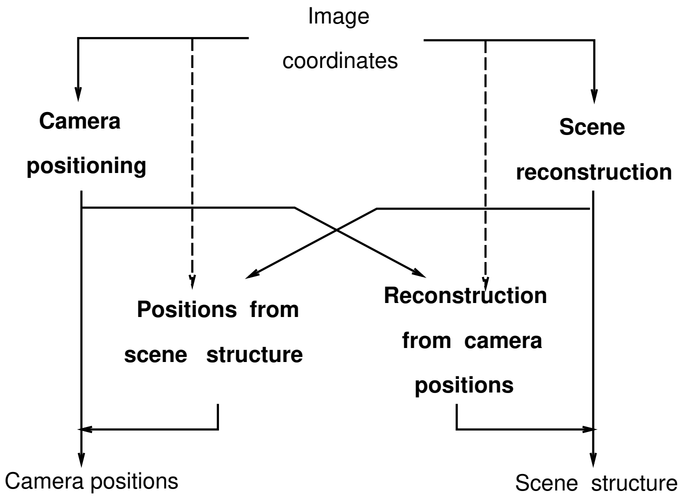


Figure 3: *Dual pathways for camera positioning and scene reconstruction from image coordinates*

Since all rectangular matrices have rank ≤ 3 all their 4×4 minors will vanish which means that we get constraints among up to four cameras and image points or up to four space points and image points [8]. These constraints will be bi- tri- or quadrilinear in the projective image coordinates. The constraints for camera positions are just the multi linear constraints derived in [8],[12], [21], [27]. These constraints have exact dual counterparts in constraints on space points and image coordinates depending on the choice of number of points and views. E.g., If we take 7 points and 2 views the space point constraints we get will correspond to the case of 6 points in 3 views for camera position constraints. In general he constraints we get for camera positions from m points in n views are mathematically identical to the constraints for space points from $n + 4$ points in $m - 4$ views.

This means that algorithms developed for multiple camera geometry for various sets of points and images can be used for direct computation of projective structure of space points and in the same way, the algorithms for direct computation of projective structure can be used for obtaining relative camera positions. Fig 4 illustrates this for various published algorithms. As was mentioned, the linear 8-points algorithm for computing the F-matrix [14], can be used for direct computation of projective structure for 6 points in 4 images. Likewise, as pointed out in [21], there is a linear algorithm for camera geometry for 7 points in 3 views. This algorithm could equally well be applied to the projective reconstruction problem with the same number of points and views. There is also a linear positioning algorithm for 6 points in 4 views [12], which

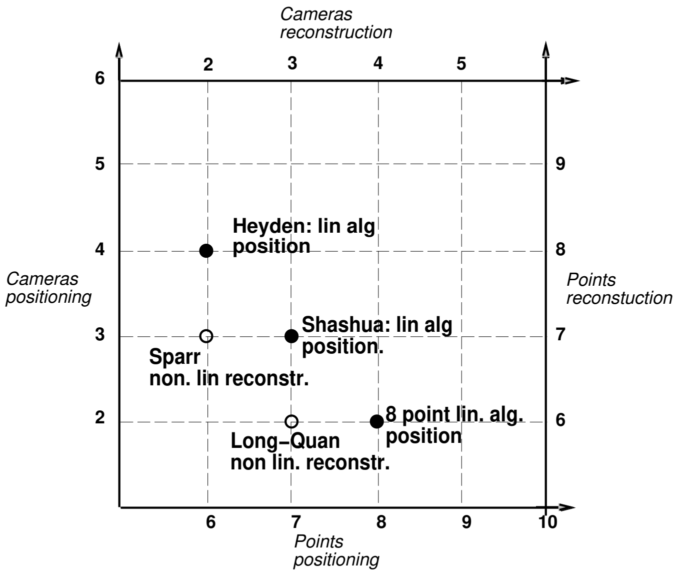


Figure 4: *The problem of positioning cameras from m points in n views is mathematically identical to the problem of reconstructing $n+4$ points in $m-4$ views*

could be used for direct reconstruction of 8 points in 2 views.

The fact that the problems of projective reconstruction and camera positioning have this dual structure poses challenging problems for computationally efficient algorithms. The problems of reconstruction and positioning are very intimately connected and should not be solved separately. Since for a given number of points and views there are alternative ways to achieve reconstruction and positioning this poses the question of the best choice in a given situation. Constraints on scene structure and constraints on camera geometry can be used to reduce the number of unknowns to be solved for and should be exploited in similar ways. The existence of ambiguous configurations of points [16] should have a dual counterpart in ambiguous camera configurations.

7 Summary

We have derived relations between linearly invariant representations for points in 3-D, camera positions and image points for a projective camera. These relations have a dual structure with respect to camera and 3-D points in the sense that the problems of computation of camera positions and point scene structure can be solved with exactly the same algorithm depending on the number of 3-D points and views

A View variation of linear image invariants

If we define the eucledian point coordinates in 3-D $\bar{P}$ and homogeneous coordinates P:

$$\bar{P} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \qquad P = \begin{pmatrix} \bar{P} \\ 1 \end{pmatrix} \tag{15}$$

the relation between the homogeneous space and image coordinates of a point can be written as:

$$p = w \, Q \, (I \mid -\bar{P}_0 \,) P \tag{16}$$

where Q is a 3×3 matrix contains intrinsic camera parameters, $\bar{P}_0$ is the eucledian coordinate vector of camera position and w is an arbitrary scale factor. Using this and the relations (15) we get:

$$p_i = w_i \, Q \, (\bar{P}_i - \bar{P}_0) \tag{17}$$

Taking 3 points we have:

$$[p_i \; p_j \; p_k \,] = w_i w_j w_k \, [Q] \, [\bar{P}_i - \bar{P}_0 \; \bar{P}_j - \bar{P}_0 \; \bar{P}_k - \bar{P}_0]$$

If we develop the determinant of the right side we have:

$$[\bar{P}_i - \bar{P}_0 \; \bar{P}_j - \bar{P}_0 \; \bar{P}_k - \bar{P}_0] =$$

$$= [\bar{P}_i \; \bar{P}_j \; \bar{P}_k] - [\bar{P}_0 \; \bar{P}_j \; \bar{P}_k] - [\bar{P}_i \; \bar{P}_0 \; \bar{P}_k] - [\bar{P}_i \; \bar{P}_j \; \bar{P}_0] =$$

$$= \begin{bmatrix} \bar{P}_0 & \bar{P}_i & \bar{P}_j & \bar{P}_k \\ 1 & 1 & 1 & 1 \end{bmatrix} = [\, P_0 \; P_i \; P_j \; P_k \,] \tag{18}$$

Inserting this into eq. (18) we get:

$$[p_i \; p_j \; p_k \,] = w_i w_j w_k \, [Q] \, [\, P_0 \; P_i \; P_j \; P_k \,] \tag{19}$$

This is the fundamental relation that will be used in order to derive relations between image and 3-D invariants. We immediately get the relation between the image invariant of 5 points and the linear 3-D invariant of the 5 points in space together with the camera position:

$$\frac{[\, p_1 \; p_2 \; p_4 \,] \;\; [\, p_1 \; p_3 \; p_5 \,]}{[\, p_1 \; p_3 \; p_4 \,] \;\; [\, p_1 \; p_2 \; p_5 \,]} = \frac{[\, P_0 \; P_1 \; P_2 \; P_4 \,] \;\; [\, P_0 \; P_1 \; P_3 \; P_5 \,]}{[\, P_0 \; P_1 \; P_3 \; P_4 \,] \;\; [\, P_0 \; P_1 \; P_2 \; P_5 \,]}$$

Note that the scale factors w_i cancel due to the choice of combinations of points in the brackets. The linear invariance properties of these expressions follow from the product rule for determinants whereby the determinant of the transformation matrix applied to the homogeneous coordinates will cancel.

References

[1] S.Carlsson, Relative positioning from model indexing, Image and Vision Computing Vol 12 No 3 pp 179 - 186 April 1994

[2] S.Carlsson, The Double Algebra: An Effective Tool for Computing Invariants in Computer Vision "Applications of Invariance in Computer Vision" Springer LNCS 825 pp. 145 - 164 (1994)

[3] S. Carlsson View Variation and Linear Invariants in 2-D and 3-D Tech. rep. Royal Institute of Technology (1995)

[4] G. Csurka and O.D. Faugeras Computing three-dimensional projective invariants from a pair of images using the Grassmann-Cayley algebra Tech. report INRIA-Sophia Antipolis Nov. 1994

[5] S. Demey, A. Ziserman and P. Beardley, Affine and projective structure from motion, Proc. BMVC-92, (1992)

[6] O.D. Faugeras, (1992), What can be seen in three dimensions with an un calibrated stereo rig?, Proc. 2:nd ECCV, pp. 563-578.

[7] O.D. Faugeras, Q.T. Luong, and S.J. Maybank, (1992), Camera self- calibration: Theory and experiments, Proc. 2:nd ECCV, pp. 321-334.

[8] O.D. Faugeras, and B. Mourrain Algebraic and geometric properties of point correspondences between N images, Proc 5:th ICCV, (1995)

[9] P. Gros , How to use the cross ratio to compute invariants from two images "Applications of Invariance in Computer Vision" Springer LNCS 825 pp. 107 - 126 (1994)

[10] R.I. Hartley, Estimation of relative camera positions for un-calibrated cameras, Proc. 2:nd ECCV, pp. 579-587.

[11] R.I. Hartley Lines and Points in Three Views - An Integrated Approach IUWS (1994)

[12] A. Heyden Reconstruction from image sequences by means of relative depths. Proc 5:th ICCV, (1995)

[13] Long-QUAN, Invariants of 6 points from 3 uncalibrated images, Proc. 3:rd ECCV, pp. Vol. II 459 - 470 (1994)

[14] H.C. Longuet-Higgins and K. Prazdny,"The interpretation of a moving retinal image", Proc. Roy. Soc. Lond. B-208, pp. 385-397, 1980

[15] Luong, Q.-T. and Viéville, T. Canonic representations for the geometries of multiple projective views", Proc. 3:rd ECCV (1994)

[16] S.J. Maybank, The projective geometry of ambiguous surfaces, Proc. Royal Soc. A 322: 1- 47, (1990)

[17] R. Mohr, Projective Geometry and Computer Vision, Handbook of Pattern Recognition and Computer Vision, (1992)

[18] R.Mohr, L. Quan, F. Veillon , Relative 3D Reconstruction using Multiple Uncalibrated Images International Journal of Robotics research (to appear)

[19] C.A. Rothwell, D.A. Forsyth, A.P. Zisserman, J.L. Mundy (1993) Extracting Projective Structure from Single Perspective Views of 3-D Point Sets Proc. of 4:th ICCV, pp. 573-582.

[20] A. Shashua, (1993), Projective depth: A Geometric Invariant for 3-D Reconstruction From Two Perspective/Orthographic Views and for Visual recognition, Proc 4:th ICCV, pp. 583-590.

[21] A. Shashua, Trilinearity in visual recognition by alignment, Proc. ECCV A, 479-484. (1994)

[22] A. Shashua and M. Werman Trilinearity of Three Perspective Views and its Associated Tensor, Proc 5:th ICCV, (1995)

[23] A. Shashua, Algebraic functions for recognition, IEEE-PAMI (To appear)

[24] G. Sparr, (1991), Proc. 1:st ESPRIT-DARPA workshop on Invariants in Computer Vision, Reykjavik, (1991)

[25] G. Sparr, (1992), Depth Computations from Polyhedral Images, Image and Vision Computing, vol. 10, no. 10, pp. 683-688.

[26] G. Sparr, A common framework for kinetic depth, reconstruction and motion for deformable objects, Proc. ECCV B, 471-482. (1994)

[27] B. Triggs Matching Constraints and the Joint Image Proc 5:th ICCV, (1995)

[28] A. Zisserman, A Case Against Epipolar Geometry, "Applications of Invariance in Computer Vision" Springer LNCS 825 pp. 69 - 88 (1994)