International Journal of Computer Vision 27(3), 227–241 (1998)

Dual Computation of Projective Shape and Camera Positions from Multiple Images

STEFAN CARLSSON
Dept. of Numerical Analysis and Computing Science, Royal Institute of Technology, S-100 44 Stockholm, Sweden
stefanc@nada.kth.se

DAPHNA WEINSHALL
Institute of Computer Science, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel
daphna@cs.huji.ac.il

Received June 10, 1996; Accepted March 3, 1997

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

- What is the structure of the set of points in 3D?
- 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 projectively, these problems are dual because they can be solved using constraint equations where space points and camera positions occur in a reciprocal way. More specifically, by using canonical projective reference frames for all points in space and images, the imaging of point sets in space by multiple cameras can be captured by constraint relations involving three different kinds of parameters only, coordinates of: (1) space points, (2) camera positions (3) image points. The duality implies that the problem of computing camera positions from p points in q views can be solved with the same algorithm as the problem of directly reconstructing $q + 4$ points in $p - 4$ views. This unifies different approaches to projective reconstruction: methods based on external calibration and direct methods exploiting constraints that exist between shape and image invariants.

Keywords: projective shape, reconstruction, positioning, epipolar geometry, duality, multiple views

1. Introduction

The problems of determining the position of cameras relative to a scene, and the 3D reconstruction of the scene from image data, have traditionally been treated as separate problems with relative camera positioning preceding the reconstruction. Given calibrated cameras, their relative orientation can be determined from observations of corresponding points in the images using the epipolar constraint. The Euclidean structure of the set of points in space can then be determined using the relative orientation of the cameras. In the case of uncalibrated cameras, relative camera position and 3D structure of the point set can still be determined, but only up to a linear transformation depending on the projection model of the camera. For parallel projection, position and structure can be determined up to an arbitrary affine transformation (Koenderink and van Doorn, 1991), and for perspective projection cameras, position and structure can be determined up to an arbitrary linear transformation in $\mathcal{P}^3$ (Faugeras et al., 1992; Hartley et al., 1992). In the perspective projection case we therefore use the terms projective reconstruction and projective shape for the computation of 3D shape when we are dealing with uncalibrated cameras.

Camera positioning, or external calibration, in the perspective projection case is based on the determination of the epipolar geometry which for two cameras is captured by the fundamental matrix (Faugeras et al., 1992). This matrix 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: projective shape from projection matrices (Faugeras, 1992; Mohr et al., 1995), cross-ratios (Gros, 1994), or more direct methods expressing 3D invariants directly in terms of the fundamental matrix (Carlsson, 1994; Csurka and Faugeras, 1994). Recently, initiated by the work in (Shashua, 1994, 1995), the generalization of the epipolar constraints to the case of multiple uncalibrated cameras has received widespread attention (Faugeras and Mourrain, 1995; Hartley, 1994; Heyden, 1995a; Luong and Viéville, 1994; Shashua and Werman, 1995; Triggs, 1995). The bilinear image constraints in the case of two images are then generalized to multilinear constraints between multiple images. The case of multiple camera constraints has also been investigated for the case of reconstruction from lines for calibrated cameras (Chen and Huang, 1990; Spetsakis and Aloimonas, 1990) and for uncalibrated cameras (Hartley, 1994).

Interestingly, there are alternative direct ways for projective reconstruction, not relying on the computation of epipolar geometry (Long Quan, 1994; Sparr, 1991, 1994). Especially in the case of constrained scenes, direct methods are powerful (Carlsson, 1995a; Rothwell et al., 1993; Sparr, 1992; Zisserman, 1994) in the sense that fewer points are needed for reconstruction. Constraining the scene also makes possible the computation of epipolar geometry with fewer points (Demey et al., 1992; Mohr, 1992; Zisserman, 1994). The direct methods for reconstruction exploit constraints existing between projective coordinates of coordinates of space points and their image coordinates, not involving camera geometry. Direct methods in unconstrained scenes, but assuming weak perspective projection, were described in (Tomasi and Kanade, 1992; Weinshall, 1993; Weinshall and Tomasi, 1995).

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 exist 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 one 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 of 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.

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 (Heyden, 1995a). In distinction to the work in (Heyden, 1995a) however, we use a *projective* canonical frame which permits us to write all constraint equations in dual form.

In the positioning case, this canonical F-matrix is parameterized by camera positions only. In the reconstruction case, the constraints between space points and image coordinates can be written using a dual of the F-matrix or tensor (denoted below the G-matrix). In the same way as the F-matrix 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 generally, we will show that the problem of computing camera positions from p points in q views is mathematically identical to the problem of reconstructing $q+4$ points in $p-4$ views. For example, we show that the eight-points linear algorithm for computing epipolar geometry between two views can be used for direct projective reconstruction of six points in four views.

The duality between positioning and reconstruction was first reported in the 1995 Workshop on Representations of Visual Scenes: The basic algebraic relations, and the geometrical interpretation in terms of duality between the position of the camera's center of projection and the position of a point in 3D, was described in (Carlsson, 1995b). The basic algebraic duality, which formally exists for any representation of the basis points, was also described in (Weinshall et al., 1995). Weinshall et al. (1995) described the new space-image relations that are obtained using the duality observation, and obtained a low rank factorization of the seven points shape tensor (cf., Shashua and Avidan, 1996). Recently a dual epipolar structure, where the dual epipole is defined by the shape geometry instead of camera geometry, was described in (Irani and Anandan, 1996).

2. Linearly Invariant Representations

The perspective projection of a point P in space to an image plane can be written as:

$$p = \mathbf{M}P \tag{1}$$

where p and P are respectively the homogeneous image and space point coordinates of the point in arbitrarily chosen coordinate systems. $\mathbf{M}$ is a 3×4 matrix depending on the position of the image plane and the camera's projection center.

Given multiple views of a set of points, reconstruction can be achieved by determining the projection matrices $\mathbf{M}_i$ and the 3D point positions P_i. In the general uncalibrated camera case, these can only be determined up to an arbitrary linear transformation (Faugeras et al., 1992; Hartley et al., 1992). Given that $P_1, P_2, \ldots, P_n$ is a reconstruction, it follows that

$$(P_1', P_2', \ldots, P_n') = \mathbf{T}(P_1, P_2, \ldots, P_n)$$

where $\mathbf{T}$ is an arbitrary 4×4 matrix, is an equally valid reconstruction given the observed image data. This equivalence class of reconstructions can be compactly described by the linear invariants of the set $P_1, P_2, \ldots, P_n$ of space points.

The basic relation that will be derived is the constraint relation involving linearly invariant representation of space points, camera positions and image coordinates. We will use five space points as a projective basis and express all other space points and camera positions in this basis with projective coordinates. In each image we choose four 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 center in space only. The specific position and orientation of the image planes does not enter at all in the relations. The reason why the orientation of the image plane is irrelevant when using linearly invariant coordinate representations is illustrated in Fig. 1: for two arbitrary choices of image planes to a given projection center $\check{P}$ the image coordinates are linearly related; the projective coordinate representation is therefore unaffected by this choice.

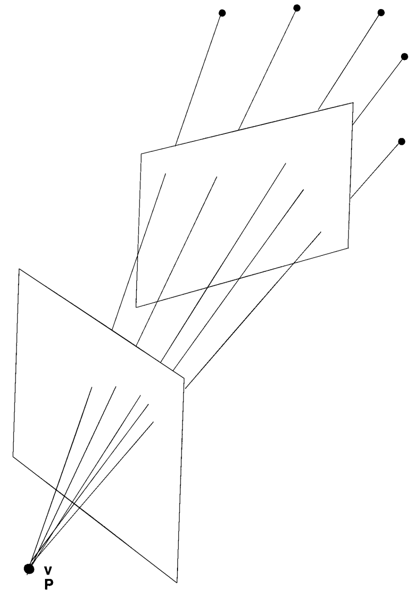

Figure 1. For two arbitrary orientations of the image plane, the image coordinates are related by a projective transformation. Thus choosing a linearly invariant representation for image data implies that the exact position and orientation of the image plane is irrelevant. The camera can therefore be represented by its position $\check{P}$ only.

The advantage of using a linearly invariant representation at the outset is that we get more compact expressions describing the relations between the important variables. The price we pay is that when we want to go back to the standard representation, i.e., choosing Cartesian coordinate systems in image and space, our relations will be expressed as high order polynomials in these coordinates.

3. Constraints between Space Points, Camera Positions and Image Coordinates

Using the projection relation (1) we will now derive invariant relations between points in 3D, camera positions and image coordinates. For that purpose we take five 3D points, with no four of them coplanar, with coordinates $P_1^*, P_2^*, P_3^*, P_4^*, P_5^*$ as a projective basis, where $*$ indicates that we have fixed the scale factor of the homogeneous coordinates. X, Y, Z, and W are the homogeneous projective coordinates of a point P with respect to this basis if

$$P = XP_1^* + YP_2^* + ZP_3^* + WP_4^* \tag{2}$$

The scale factors of the basis are determined by the requirement that points $P_1^*, P_2^*, P_3^*, P_4^*$ be basis vectors in this representation and point P_5^* has coordinates $(1, 1, 1, 1)$.

For the center of projection of the camera we use the notation $\check{P}$ and it can be expressed as

$$\check{P} = \check{X}P_1^* + \check{Y}P_2^* + \check{Z}P_3^* + \check{W}P_4^* \tag{3}$$

Similarly we choose the projective basis for image coordinates as the first four image points, $p_1^*, p_2^*, p_3^*, p_4^*$, corresponding to the space points P_1, P_2, P_3, P_4. The requirement is then that no three image points are collinear. An image point can then be expressed as

$$p = xp_1^* + yp_2^* + wp_3^* \tag{4}$$

where x, y, w are homogeneous projective image coordinates. Points p_1, p_2, p_3 will be the unit vectors and point p_4 will have the projective coordinates $(1, 1, 1)$.

To summarize, if we use the projective coordinates representation, we get

p_1	p_2	p_3	p_4	p
1	0	0	1	x
0	1	0	1	y
0	0	1	1	w

(5)

$\check{P}$	P_1	P_2	P_3	P_4	P_5	P
$\check{X}$	1	0	0	0	1	X
$\check{Y}$	0	1	0	0	1	Y
$\check{Z}$	0	0	1	0	1	Z
$\check{W}$	0	0	0	1	1	W

(6)

It follows that the projection relation (1) must have the form

$$\begin{pmatrix} x \\ y \\ w \end{pmatrix} = \begin{pmatrix} \alpha & 0 & 0 & \delta \\ 0 & \beta & 0 & \delta \\ 0 & 0 & \gamma & \delta \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ W \end{pmatrix}$$

Every space point except the camera's projection center $\check{X}, \check{Y}, \check{Z}, \check{W}$ projects to the image plane, and every point in the image plane except $(0, 0, 0)$ is image of a space point. This implies that we must have

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha & 0 & 0 & \delta \\ 0 & \beta & 0 & \delta \\ 0 & 0 & \gamma & \delta \end{pmatrix} \begin{pmatrix} \check{X} \\ \check{Y} \\ \check{Z} \\ \check{W} \end{pmatrix}$$

Thus we get

$$\begin{aligned} \alpha &= \sigma\check{X}^{-1}, \quad \beta = \sigma\check{Y}^{-1}, \\ \gamma &= \sigma\check{Z}^{-1}, \quad \delta = -\sigma\check{W}^{-1} \end{aligned} \tag{7}$$

We can therefore write the projection equation (1), relating linearly invariant representations of image points, space points, and camera projection centers, as

$$\begin{pmatrix} x \\ y \\ w \end{pmatrix} = \sigma \begin{pmatrix} \check{X}^{-1} & 0 & 0 & -\check{W}^{-1} \\ 0 & \check{Y}^{-1} & 0 & -\check{W}^{-1} \\ 0 & 0 & \check{Z}^{-1} & -\check{W}^{-1} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ W \end{pmatrix}$$

By eliminating the arbitrary scale factor σ we get the following two equations:

$$\frac{x}{w} = \frac{\check{X}^{-1}X - \check{W}^{-1}W}{\check{Z}^{-1}Z - \check{W}^{-1}W} \tag{8}$$

$$\frac{y}{w} = \frac{\check{Y}^{-1}Y - \check{W}^{-1}W}{\check{Z}^{-1}Z - \check{W}^{-1}W} \tag{9}$$

which can be written in the form of constraint relations:

Theorem (Structure position constraint duality). *Projective coordinates X, Y, X, W of a point in 3D and camera position $\check{X}, \check{Y}, \check{X}, \check{W}$ are constrained by projective image coordinates x, y, w according to*

$$w\frac{Y}{\check{Y}} - y\frac{Z}{\check{Z}} + (y - w)\frac{W}{\check{W}} = 0 \tag{10}$$

$$w\frac{X}{\check{X}} - x\frac{Z}{\check{Z}} + (x - w)\frac{W}{\check{W}} = 0 \tag{11}$$

These constraints are dual w.r.t. scene structure and camera positions in the sense that they are unaffected *by the substitutions*

$$(X, Y, Z, W) \Longleftrightarrow (\check{X}^{-1}, \check{Y}^{-1}, \check{Z}^{-1}, \check{W}^{-1})$$

The consequence of this theorem is that the problems of computing positions of points in space (X, Y, Z, W), and camera positions $(\check{X}, \check{Y}, \check{Z}, \check{W})$, from image data have identical mathematical structure.

Note that the actual choice of the fifth point in the 3D projective basis does not affect the basic dual constraint relations that were derived above. In the following chapters we will have reason to choose this point with some liberty, e.g., as one of the camera points.

In the case of multiple points and cameras, we get a pair of constraint equations (10) and (11) for each combination of points cameras. We will use indices $1, 2, 3, 4, n, m, \ldots, i$ to denote space points and indices $a, b, c, \ldots, q$ to denote camera projection points. The image point of space point i in camera q has index ${}_i^q$. Thus we use the following notations:

Space points: $P_1, P_2, P_3, P_4, \ldots, P_n, \ldots, P_m, \ldots, P_i$
Camera projection points: $\check{P}_a, \check{P}_b, \check{P}_c, \ldots, \check{P}_q$
Image point i in camera q: p_i^q

3.1. Projective Reconstruction from Known Camera Positions and Two Images

Using (10) and (11) we can solve for the projective coordinates of point n given the images in two cameras a, b with known positions:

$$\begin{pmatrix} 0 & w_n^a \check{Y}_a^{-1} & -y_n^a \check{Z}_a^{-1} & (y_n^a - w_n^a)\check{W}_a^{-1} \\ w_n^a \check{X}_a^{-1} & 0 & -x_n^a \check{Z}_a^{-1} & (x_n^a - w_n^a)\check{W}_a^{-1} \\ 0 & w_n^b \check{Y}_b^{-1} & -y_n^b \check{Z}_b^{-1} & (y_n^b - w_n^b)\check{W}_b^{-1} \\ w_n^b \check{X}_b^{-1} & 0 & -x_n^b \check{Z}_b^{-1} & (x_n^b - w_n^b)\check{W}_b^{-1} \end{pmatrix} \times \begin{pmatrix} X_n \\ Y_n \\ Z_n \\ W_n \end{pmatrix} = 0 \qquad (12)$$

This equation defines the homogeneous coordinates of the position of point n in space as the null-space of the matrix on the left side, which is determined by cameras a and b.

3.2. Relative Positioning from Known Space Points and One Image

In a very similar way the camera position for camera a can be computed given a known configuration of six points in space $1, \ldots, 4, n, m$. From (10) and (11) we get:

$$\begin{pmatrix} 0 & w_n^a Y_n & -y_n^a Z_n & (y_n^a - w_n^a) W_n \\ w_n^a X_n & 0 & -x_n^a Z_n & (x_n^a - w_n^a) W_n \\ 0 & w_m^a Y_m & -y_m^a Z_m & (y_m^a - w_m^a) W_m \\ w_m^a X_m & 0 & -x_m^a Z_m & (x_m^a - w_m^a) W_m \end{pmatrix} \times \begin{pmatrix} \check{X}_a^{-1} \\ \check{Y}_a^{-1} \\ \check{Z}_a^{-1} \\ \check{W}_a^{-1} \end{pmatrix} = 0 \qquad (13)$$

Using points $m \neq n$ where $m > 4$ and $n > 4$ we can solve for the projective coordinates of the camera position in terms of image and space point coordinates in the same way as in the projective reconstruction above.

We see that the problems of projective reconstruction from known camera positions and camera positioning from known space points are mathematically identical. This is a consequence of the structure of the basic constraint relations (10) and (11). We will also see that, using the constraint relations we get by eliminating either space points or camera positions, we get mathematically identical expressions.

4. Position and Structure Constraints from Image Measurements

From (12) and (13) we can derive relations between camera positions and image measurements, or between space coordinates and image measurements, by noting that the 4×4 matrices in general have rank 3.

4.1. Position Constraints—Camera Positions and Image Coordinates

The geometric basis for the space-image and camera-image constraint relations developed below can be seen in Fig. 2. The camera-image (or epipolar) constraint follows from the four point coplanarity of the camera positions and image points in 3D (a_0, b_0, a_1, b_1 and

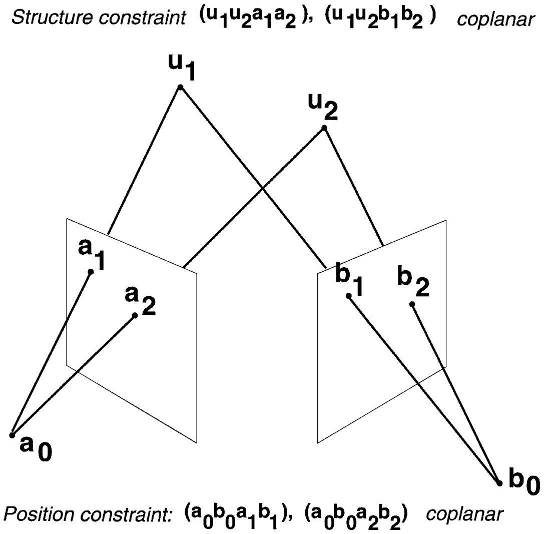

Figure 2. Geometric interpretation of epipolar and space-image constraints in terms of four point coplanarities, see text.

a_0, b_0, a_2, b_2 in the figure). The space-image constraint follows from coplanarities of space points and image points (u_1, u_2, a_1, a_2 and u_1, u_2, b_1, b_2 in the figure).

From (12) we get

$$\det\begin{pmatrix} 0 & w_n^a\check{Y}_a^{-1} & -y_n^a\check{Z}_a^{-1} & (y_n^a - w_n^a)\check{W}_a^{-1} \\ w_n^a\check{X}_a^{-1} & 0 & -x_n^a\check{Z}_a^{-1} & (x_n^a - w_n^a)\check{W}_a^{-1} \\ 0 & w_n^b\check{Y}_b^{-1} & -y_n^b\check{Z}_b^{-1} & (y_n^b - w_n^b)\check{W}_b^{-1} \\ w_n^b\check{X}_b^{-1} & 0 & -x_n^b\check{Z}_b^{-1} & (x_n^b - w_n^b)\check{W}_b^{-1} \end{pmatrix} = 0 \qquad (14)$$

This can be written in the following bilinear form

$$\begin{pmatrix} x_n^a \\ y_n^a \\ w_n^a \end{pmatrix}^T \begin{pmatrix} 0 & f_2 - f_1 & f_3 - f_2 \\ f_4 - f_5 & 0 & f_5 - f_3 \\ f_6 - f_4 & f_1 - f_6 & 0 \end{pmatrix} \begin{pmatrix} x_n^b \\ y_n^b \\ w_n^b \end{pmatrix} = 0 \qquad (15)$$

where

$$f_1 = \check{X}_b^{-1}\check{Y}_a^{-1}\check{Z}_b^{-1}\check{W}_a^{-1}$$
$$f_2 = \check{X}_b^{-1}\check{Y}_a^{-1}\check{Z}_a^{-1}\check{W}_b^{-1}$$
$$f_3 = \check{X}_b^{-1}\check{Y}_b^{-1}\check{Z}_a^{-1}\check{W}_a^{-1}$$
$$f_4 = \check{X}_a^{-1}\check{Y}_b^{-1}\check{Z}_b^{-1}\check{W}_a^{-1}$$
$$f_5 = \check{X}_a^{-1}\check{Y}_b^{-1}\check{Z}_a^{-1}\check{W}_b^{-1}$$
$$f_6 = \check{X}_a^{-1}\check{Y}_a^{-1}\check{Z}_b^{-1}\check{W}_b^{-1}$$

This is the epipolar constraint on the projective image coordinates in image a and image b. The 3×3 matrix F in (15) is the fundamental matrix parameterized 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 imaging parameters. The structure of this F-matrix is similar to the one derived in (Heyden, 1995a) where a similar but not identical canonical framework is used.

It can be readily shown that the determinant

$$(f_2 - f_1)(f_5 - f_3)(f_6 - f_4) + (f_3 - f_2)(f_4 - f_5)(f_1 - f_6) = 0 \qquad (16)$$

The rank of the F-matrix is therefore ≤ 2.

4.2. *Structure Constraints—Space Points and Image Coordinates*

For two points in one image we get from (13)

$$\det\begin{pmatrix} 0 & w_n^a Y_n & -y_n^a Z_n & (y_n^a - w_n^a)W_n \\ w_n^a X_n & 0 & -x_n^a Z_n & (x_n^a - w_n^a)W_n \\ 0 & w_m^a Y_m & -y_m^a Z_m & (y_m^a - w_m^a)W_m \\ w_m^a X_m & 0 & -x_m^a Z_m & (x_m^a - w_m^a)W_m \end{pmatrix} = 0 \qquad (17)$$

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

$$\begin{pmatrix} x_n^a \\ y_n^a \\ w_n^a \end{pmatrix}^T \begin{pmatrix} 0 & g_2 - g_1 & g_3 - g_2 \\ g_4 - g_5 & 0 & g_5 - g_3 \\ g_6 - g_4 & g_1 - g_6 & 0 \end{pmatrix} \begin{pmatrix} x_m^a \\ y_m^a \\ w_m^a \end{pmatrix} = 0 \qquad (18)$$

where

$$g_1 = X_m Y_n Z_m W_n$$
$$g_2 = X_m Y_n Z_n W_m$$

$$g_3 = X_m Y_m Z_n W_n$$
$$g_4 = X_n Y_m Z_m W_n$$
$$g_5 = X_n Y_m Z_n W_m$$
$$g_6 = X_n Y_n Z_m W_m$$

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 satisfied.

It can be readily shown that the determinant

$$(g_2 - g_1)(g_5 - g_3)(g_6 - g_4) + (g_3 - g_2)(g_4 - g_5)(g_1 - g_6) = 0 \quad (19)$$

The rank of the G-matrix is therefore ≤ 2.

5. Computing Camera and Space Projective Coordinates from the Constraint Relations

5.1. Linear Computation of the F- and G-Matrices

The constraints (15) and (18) have exactly the same form if we make the identifications $a \leftrightarrow n$ and $b \leftrightarrow m$. The problem of computing the projective coordinates of the camera's projection center from the fundamental matrix (15) is therefore equivalent to the problem of computing projective coordinates of space points from the space-image constraint matrix (18).

Equations (15) and (18) provide linear constraints on the six unknown elements of the F-matrix and the G-matrix, respectively. Due to the structure of the F- and G-matrices their elements sum to 0.

$$\sum f_{ij} = \sum g_{ij} = 0$$

This gives one additional constraint in both cases.

Since there are six homogeneous non-zero elements of the matrices F and G, and we have the constraint that the elements sum to 0, we need four constraints of the type (15) and (18) for a linear computation of the matrix elements in both cases.

In order to get a more compact notation we write $\hat{p}_n^{a^T} = (x_n^a, y_n^a, w_n^a)$ for the projective coordinates of point n in image a and F_{ab} and G_{mn} to indicate that the F- and G-matrices are associated with the pair of cameras a, b and the pair of points m, n, respectively.

Linear F-matrix Computation from Eight Points and Two Images. Given two images, image points with index $i \geq 5$ provide constraints of type (15) on the F-matrix. For $i = 1, 2, 3$ we get trivial identities due to the structure of the F-matrix and $i = 4$ just gives the summation constraint. Thus eight points observed in two images a, b will give four linear constraints on the elements of the matrix F_{ab}

$$\hat{p}_i^{a^T} F_{ab} \hat{p}_i^b = 0, \quad i = 5, 6, 7, 8 \quad (20)$$

Linear G-matrix Computation from Six Points and Four Images. Similarly, given six points 1, 2, 3, 4, m, n and multiple images, we get one constraint from each image on the G-matrix for points mn. Note that the elements of the matrix G_{mn} are trivially 0 for $m \leq 4$ or $n \leq 4$. Thus six points observed in four images will give four linear constraints on the elements of the matrix G_{mn}

$$\hat{p}_m^{q^T} G_{mn} \hat{p}_n^q = 0 \quad q = a, b, c, d \quad (21)$$

5.2. Computing Relative Projective Coordinates from the F- and G-Matrices

The projective coordinates of space and camera projection points can be computed from the elements of the F- and G-matrices. These can be given a simpler structure by introducing the relative projective coordinates

$$\chi = \frac{\check{X}_b^{-1}}{\check{X}_a^{-1}} \quad \psi = \frac{\check{Y}_b^{-1}}{\check{Y}_a^{-1}} \quad \zeta = \frac{\check{Z}_b^{-1}}{\check{Z}_a^{-1}} \quad \omega = \frac{\check{W}_b^{-1}}{\check{W}_a^{-1}}$$

and (22)

$$\chi = \frac{X_m}{X_n} \quad \psi = \frac{Y_m}{Y_n} \quad \zeta = \frac{Z_m}{Z_n} \quad \omega = \frac{W_m}{W_n},$$

respectively.

After factoring out common factors, the F- and G-matrices in (15) and (18):

$$F = \begin{pmatrix} 0 & f_{12} & f_{13} \\ f_{21} & 0 & f_{23} \\ f_{31} & f_{32} & 0 \end{pmatrix} \quad G = \begin{pmatrix} 0 & g_{12} & g_{13} \\ g_{21} & 0 & g_{23} \\ g_{31} & g_{32} & 0 \end{pmatrix} \quad (23)$$

can now be written as

$$\propto \begin{pmatrix} 0 & \chi(\omega-\zeta) & -\chi(\omega-\psi) \\ -\psi(\omega-\zeta) & 0 & \psi(\omega-\chi) \\ \zeta(\omega-\psi) & -\zeta(\omega-\chi) & 0 \end{pmatrix} \tag{24}$$

Given the elements of the F- and G-matrices, we want to compute the coordinates $\check{X}_a, \ldots, \check{W}_a, \check{X}_b, \ldots, \check{W}_b$ of the camera positions, and $X_m, \ldots, W_m, X_n, \ldots, W_n$ of points m and n, respectively. As an intermediate step we compute the parameters $\chi, \psi, \zeta, \omega$. If we let q_{ij} denote either f_{ij} or g_{ij} we get from (23) and (24) the explicit solutions for the normalized variables:

$$\begin{aligned} \frac{\chi}{\omega} &= -\frac{q_{1,2}+q_{1,3}}{q_{3,1}+q_{2,1}} \\ \frac{\psi}{\omega} &= -\frac{q_{2,3}+q_{2,1}}{q_{1,2}+q_{3,2}} \\ \frac{\zeta}{\omega} &= -\frac{q_{3,1}+q_{3,2}}{q_{1,3}+q_{2,3}} \end{aligned} \tag{25}$$

5.3. *Nonlinear Computation of F- and G-Matrices*

The linear computation of the elements of the F- and G-matrices does not make use of the fact that these matrices are rank ≤ 2. This imposes an extra constraint on the matrix elements, which means that we need only seven points and two images in the F-matrix case and six points and three images for the G-matrix computation. However, since the rank constraint is non-linear, we will in general have multiple solutions in these cases. Alternatively, the computation can be made more efficient by directly computing the substitution variables χ, ψ, ζ and ω. Note that the form of the matrix (24) automatically imposes the rank constraint. Now every constraint relation of the form:

$$\begin{pmatrix} x_{ip} \\ y_{ip} \\ w_{ip} \end{pmatrix}^T \begin{pmatrix} 0 & \chi(\omega-\zeta) & -\chi(\omega-\psi) \\ -\psi(\omega-\zeta) & 0 & \psi(\omega-\chi) \\ \zeta(\omega-\psi) & -\zeta(\omega-\chi) & 0 \end{pmatrix} \times \begin{pmatrix} x_{jq} \\ y_{jq} \\ w_{jq} \end{pmatrix} = 0 \tag{26}$$

can be considered as a nonlinear constraint on the three normalized unknowns $\frac{\chi}{\omega}, \frac{\psi}{\omega}, \frac{\zeta}{\omega}$. This means that we need a minimum of three equations to get a solution.

Nonlinear F-matrix Computation for Seven Points and Two Images. In the F-matrix case we can take $p = a, q = b$, and $i = j = 5, 6, 7$ in order to get three equations for computing camera positions from seven points and two cameras.

Nonlinear G-matrix Computation for Six Points and Three Images. In the G-matrix case we can take $p = q = a, b, c$ and $i = 5$, $j = 6$ in order to get three equations for computing space point positions from six points and three cameras.

These equations are exactly those treated and solved in (Long Quan, 1994) for the structure computation case.

5.4. *Computing Projective Coordinates from the F- and G-Matrices*

The parameters $\chi, \psi, \zeta, \omega$ give ratios of homogeneous projective coordinates for either camera positions or space points. The computation of either camera positions or space point positions can be made especially simple by making a particular choice of the fifth point in the projective 3D basis.

Choosing the Fifth Basis Point as a Space Point. Suppose we choose the fifth 3D basis point as the fifth space point P_5. We then have

$$(X_5, Y_5, Z_5, W_5) = (1, 1, 1, 1) \tag{27}$$

For $n = 5$, $m > 5$ in (22) we then get

$$\chi = X_m, \quad \psi = Y_m, \quad \zeta = Z_m, \quad \omega = W_m \tag{28}$$

In other words, we get the position of point m directly from the substituted variables.

For the camera positions we still get ratios of projective coordinates for two cameras. However, we now have auxiliary constraints on the camera positions from the original constraint equations (12).

$$\begin{aligned} w_5^a \check{Y}_a^{-1} - y_5^a \check{Z}_a^{-1} + \left(y_5^a - w_5^a\right) \check{W}_a^{-1} &= 0 \\ w_5^a \check{X}_a^{-1} - x_5^a \check{Z}_a^{-1} + \left(x_5^a - w_5^a\right) \check{W}_a^{-1} &= 0 \\ w_5^b \check{Y}_b^{-1} - y_5^b \check{Z}_b^{-1} + \left(y_5^b - w_5^b\right) \check{W}_b^{-1} &= 0 \\ w_5^b \check{X}_b^{-1} - x_5^b \check{Z}_b^{-1} + \left(x_5^b - w_5^b\right) \check{W}_b^{-1} &= 0 \end{aligned} \tag{29}$$

If we use these together with (22)

$$\chi = \frac{\check{X}_b^{-1}}{\check{X}_a^{-1}}, \quad \psi = \frac{\check{Y}_b^{-1}}{\check{Y}_a^{-1}}, \quad \zeta = \frac{\check{Z}_b^{-1}}{\check{Z}_a^{-1}}, \quad \omega = \frac{\check{W}_b^{-1}}{\check{W}_a^{-1}} \tag{30}$$

we get a linear system for computing camera positions from $\chi, \psi, \zeta, \omega$ (which are computed linearly from the F-matrix, as described in Section 5.2).

Choosing the Fifth Basis Point as a Camera Point. In a completely dual way we can get a simple solution for the camera positions if we choose the fifth 3D basis point as a camera position.

$$(\check{X}_a, \check{Y}_a, \check{Z}_a, \check{W}_a) = (1, 1, 1, 1) \tag{31}$$

We then get from (22)

$$\chi = X_b^{-1}, \quad \psi = Y_b^{-1}, \quad \zeta = Z_b^{-1}, \quad \omega = W_b^{-1} \tag{32}$$

In other words, we get the position of camera b directly from the substituted variables. This position is now computed in a system spanned by the first four space points and the position of camera a. In order to express the space points in this system, we have to use auxiliary constraints that are dual to those in (29). From (13) we get:

$$\begin{aligned} w_n^a Y_n - y_n^a Z_n + \left(y_n^a - w_n^a\right) W_n &= 0 \\ w_n^a X_n - x_n^a Z_n + \left(x_n^a - w_n^a\right) W_n &= 0 \\ w_m^a Y_m - y_m^a Z_m + \left(y_m^a - w_m^a\right) W_m &= 0 \\ w_m^a X_m - x_m^a Z_m + \left(x_m^a - w_m^a\right) W_m &= 0 \end{aligned} \tag{33}$$

Which can be used together with (22)

$$\chi = \frac{X_n}{X_m}, \quad \psi = \frac{Y_n}{Y_m}, \quad \zeta = \frac{Z_n}{Z_m}, \quad \omega = \frac{W_n}{W_m} \tag{34}$$

to compute the positions of space points m and n linearly from $\chi, \psi, \zeta, \omega$.

6. Duality of Reconstruction and Positioning from Multiple Data

6.1. Multilinear Constraints

The constraints (12) 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 whose rank ≤ 3.

For one point and many images (12) becomes:

$$\begin{pmatrix} 0 & \cdots {}^a\check{Y}_a^{-1} & \cdots {}^a\check{Z}_a^{-1} & \cdots {}^a\check{W}_a^{-1} \\ \cdots {}^a\check{X}_a^{-1} & 0 & \cdots {}^a\check{Z}_a^{-1} & \cdots {}^a\check{W}_a^{-1} \\ 0 & \cdots {}^b\check{Y}_b^{-1} & \cdots {}^b\check{Z}_b^{-1} & \cdots {}^b\check{W}_b^{-1} \\ \cdots {}^b\check{X}_b^{-1} & 0 & \cdots {}^b\check{Z}_b^{-1} & \cdots {}^b\check{W}_b^{-1} \\ 0 & \cdots {}^c\check{Y}_c^{-1} & \cdots {}^c\check{Z}_c^{-1} & \cdots {}^c\check{W}_c^{-1} \\ \cdots {}^c\check{X}_c^{-1} & 0 & \cdots {}^c\check{Z}_c^{-1} & \cdots {}^c\check{W}_c^{-1} \\ & & \vdots & \end{pmatrix} \times \begin{pmatrix} X_n \\ Y_n \\ Z_n \\ W_n \end{pmatrix} = 0 \tag{35}$$

Similarly, for many points in one image (13) becomes:

$$\begin{pmatrix} 0 & \cdots {}_5 Y_5 & \cdots {}_5 Z_5 & \cdots {}_5 W_5 \\ \cdots {}_5 X_5 & 0 & \cdots {}_5 Z_5 & \cdots {}_5 W_5 \\ 0 & \cdots {}_6 Y_6 & \cdots {}_6 Z_6 & \cdots {}_6 W_6 \\ \cdots {}_6 X_6 & 0 & \cdots {}_6 Z_6 & \cdots {}_6 W_6 \\ 0 & \cdots {}_7 Y_7 & \cdots {}_7 Z_7 & \cdots {}_7 W_7 \\ \cdots {}_7 X_7 & 0 & \cdots {}_7 Z_7 & \cdots {}_7 W_7 \\ & & \vdots & \end{pmatrix} \begin{pmatrix} \check{X}_a^{-1} \\ \check{Y}_a^{-1} \\ \check{Z}_a^{-1} \\ \check{W}_a^{-1} \end{pmatrix} = 0 \tag{36}$$

Since both rectangular matrices have rank ≤ 3 all their 4×4 minors must vanish, giving us constraints among up to four cameras and image measurements, or up to four space points and image measurements (Faugeras and Mourrain, 1995). These constraints are bi- tri- or quadri-linear in the projective image coordinates. They are related to the multiple image epipolar constraints previously derived and discussed in (Faugeras and Mourrain, 1995; Heyden, 1995a; Shashua, 1994; Triggs, 1995).

The problems of computing scene structure and camera positions from multiple points and views are mathematically identical:

Theorem (Reconstruction positioning duality). *For multiple cameras and points, the camera position, or epipolar constraints have exact dual counterparts in constraints on space points and image coordinates depending on the choice of number of points and frames. In general the constraints we get for camera positions from p points in q views are mathematically identical to the constraints for space points from $q+4$ points in $p-4$ views.*

The multilinear constraints in the position and reconstruction cases can be treated in a unified way, similar to the bilinear constraints. We first recall that they are obtained from the fact that the rank of some rectangular matrix is 3. This remains true when the matrix is multiplied on the right by

$$\begin{pmatrix} X & 0 & 0 & 0 \\ 0 & Y & 0 & 0 \\ 0 & 0 & Z & 0 \\ 0 & 0 & 0 & W \end{pmatrix} \tag{37}$$

for (almost) any X, Y, Z, W.

We first multiply the rectangular matrices in (35) and (36) by

$$\begin{pmatrix} \check{X}_a & 0 & 0 & 0 \\ 0 & \check{Y}_a & 0 & 0 \\ 0 & 0 & \check{Z}_a & 0 \\ 0 & 0 & 0 & \check{W}_a \end{pmatrix} \tag{38}$$

and

$$\begin{pmatrix} X_5^{-1} & 0 & 0 & 0 \\ 0 & Y_5^{-1} & 0 & 0 \\ 0 & 0 & Z_5^{-1} & 0 \\ 0 & 0 & 0 & W_5^{-1} \end{pmatrix} \tag{39}$$

respectively. Next we generalize (22) and substitute for the camera positions and image coordinates in (35):

$$\begin{aligned}
\chi' &= \frac{\check{X}_b^{-1}}{\check{X}_a^{-1}}, \quad \psi' = \frac{\check{Y}_b^{-1}}{\check{Y}_a^{-1}}, \quad \zeta' = \frac{\check{Z}_b^{-1}}{\check{Z}_a^{-1}}, \quad \omega' = \frac{\check{W}_b^{-1}}{\check{W}_a^{-1}}, \\
\chi'' &= \frac{\check{X}_c^{-1}}{\check{X}_a^{-1}}, \quad \psi'' = \frac{\check{Y}_c^{-1}}{\check{Y}_a^{-1}}, \quad \zeta'' = \frac{\check{Z}_c^{-1}}{\check{Z}_a^{-1}}, \quad \omega'' = \frac{\check{W}_c^{-1}}{\check{W}_a^{-1}}, \\
\chi''' &= \frac{\check{X}_d^{-1}}{\check{X}_a^{-1}}, \quad \psi''' = \frac{\check{Y}_d^{-1}}{\check{Y}_a^{-1}}, \quad \zeta''' = \frac{\check{Z}_d^{-1}}{\check{Z}_a^{-1}}, \quad \omega''' = \frac{\check{W}_d^{-1}}{\check{W}_a^{-1}}.
\end{aligned} \tag{40}$$

$$\begin{aligned}
x &= x_n^a, \quad y = y_n^a, \quad z = z_n^a, \quad w = w_n^a, \\
x' &= x_n^b, \quad y' = y_n^b, \quad z' = z_n^b, \quad w' = w_n^b, \\
x'' &= x_n^c, \quad y'' = y_n^c, \quad z'' = z_n^c, \quad w'' = w_n^c, \\
x''' &= x_n^d, \quad y''' = y_n^d, \quad z''' = z_n^d, \quad w''' = w_n^d.
\end{aligned} \tag{41}$$

and for the space point positions and image coordinates in (36)

$$\begin{aligned}
\chi' &= \frac{X_6}{X_5}, \quad \psi' = \frac{Y_6}{Y_5}, \quad \zeta' = \frac{Z_6}{Z_5}, \quad \omega' = \frac{W_6}{W_5}, \\
\chi'' &= \frac{X_7}{X_5}, \quad \psi'' = \frac{Y_7}{Y_5}, \quad \zeta'' = \frac{Z_7}{Z_5}, \quad \omega'' = \frac{W_7}{W_5}, \\
\chi''' &= \frac{X_8}{X_5}, \quad \psi''' = \frac{Y_8}{Y_5}, \quad \zeta''' = \frac{Z_8}{Z_5}, \quad \omega''' = \frac{W_8}{W_5},
\end{aligned} \tag{42}$$

$$\begin{aligned}
x &= x_5^a, \quad y = y_5^a, \quad z = z_5^a, \quad w = w_5^a, \\
x' &= x_6^a, \quad y' = y_6^a, \quad z' = z_6^a, \quad w' = w_6^a, \\
x'' &= x_7^a, \quad y'' = y_7^a, \quad z'' = z_7^a, \quad w'' = w_7^a, \\
x''' &= x_8^a, \quad y''' = y_8^a, \quad z''' = z_8^a, \quad w''' = w_8^a.
\end{aligned} \tag{43}$$

Instead of both (35) and (36) we get the following constraint matrix whose rank is also 3:

$$C = \begin{pmatrix}
w & 0 & -x & (x-w) \\
0 & w & -y & (y-w) \\
w'\chi' & 0 & -x'\zeta' & (x'-w')\omega' \\
0 & w'\psi' & -y'\zeta' & (y'-w')\omega' \\
w''\chi'' & 0 & -x''\zeta'' & (x''-w'')\omega'' \\
0 & w''\psi'' & -y''\zeta'' & (y''-w'')\omega'' \\
w'''\chi''' & 0 & -x'''\zeta''' & (x'''-w''')\omega''' \\
0 & w'''\psi''' & -y'''\zeta''' & (y'''-w''')\omega''' \\
 & & \vdots &
\end{pmatrix} \tag{44}$$

All the multilinear constraints are obtained from the equations describing the fact that the determinant of each 4×4 minor of C is 0; these equations are derived below.

6.2. Bilinear Equations

We are given a constraint matrix composed of the first four rows of matrix C from (44); thus there is a single constraint—the determinant of C should vanish. This analysis was carried out in detail, for the structure and positioning cases, in Sections 4 and 5.

6.3. Trilinear Equations

We are given a constraint matrix composed of the first six rows of matrix C from (44). There are $\binom{6}{4} = 15$ constraints when considering all 4×4 minors of C. However, using algebraic tools we found that generically there are only seven independent equations: three bilinear equations involving subsets of six points (similar to those given in (15), (18)), and four new trilinear equations. The four new equations give us the following set of constraints:

$$\begin{pmatrix} 0 & A & B & 0 & 0 & 0 & C & 0 & D & E & 0 \\ 0 & 0 & 0 & F & G & H & I & 0 & J & K & 0 \\ L & M & N & 0 & 0 & 0 & 0 & O & P & 0 & Q \\ 0 & 0 & 0 & 0 & R & S & 0 & T & U & 0 & V \end{pmatrix} \times \begin{pmatrix} \chi'\psi'' - \omega'\zeta'' \\ \chi'\zeta'' - \omega'\zeta'' \\ \chi'\omega'' - \omega'\zeta'' \\ \psi'\chi'' - \omega'\zeta'' \\ \psi'\zeta'' - \omega'\zeta'' \\ \psi'\omega'' - \omega'\zeta'' \\ \zeta'\chi'' - \omega'\zeta'' \\ \zeta'\psi'' - \omega'\zeta'' \\ \zeta'\omega'' - \omega'\zeta'' \\ \omega'\chi'' - \omega'\zeta'' \\ \omega'\psi'' - \omega'\zeta'' \end{pmatrix} = 0 \tag{45}$$

where the elements of the constraint matrix above are trilinear functions of the image measurements:

$$\begin{aligned}
A &= -x\,x''\,w' + x''\,w\,w' \\
B &= x\,x''\,w' - x\,w'\,w'' \\
C &= x\,x'\,w'' - x'\,w\,w'' \\
D &= -x'\,x''\,w + x'\,w\,w'' \\
E &= -x\,x'\,w'' + x\,w'\,w'' \\
F &= -x\,w'\,w'' + y\,w'\,w'' \\
G &= -x''\,y\,w' + x''\,w\,w' \\
H &= x''\,y\,w' - y\,w'\,w'' \\
I &= x\,y'\,w'' - y'\,w\,w'' \\
J &= -x''\,y'\,w + y'\,w\,w'' \\
K &= -x\,y'\,w'' + x\,w'\,w'' \\
L &= -x\,w'\,w'' + y\,w'\,w'' \\
M &= x\,y''\,w' - y''\,w\,w' \\
N &= -x\,y''\,w' + x\,w'\,w'' \\
O &= -x'\,y\,w'' + x'\,w\,w'' \\
P &= x'\,y''\,w - x'\,w\,w'' \\
Q &= x'\,y\,w'' - y\,w'\,w'' \\
R &= -y\,y''\,w' + y''\,w\,w' \\
S &= y\,y''\,w' - y\,w'\,w'' \\
T &= y\,y'\,w'' - y'\,w\,w'' \\
U &= -y'\,y''\,w + y'\,w\,w'' \\
V &= -y\,y'\,w'' + y\,w'\,w''
\end{aligned} \tag{46}$$

We can write each new trilinear constraint in a tensor form, using $3 \times 3 \times 3$ tensors T^l. More specifically, we have

$$\sum_{i,j,k} \mathsf{T}^l_{ijk}(p)_i(p')_j(p'')_k = 0, \quad l = 1, \ldots, 4 \tag{47}$$

Each of the four constraints has a different shape tensor. For example, the $3 \times 3 \times 3$ tensor of the first constraint (the first row in (45)) is composed of the 3×3 matrices:

$$T^1_{ij1} = \begin{pmatrix} 0 & 0 & \chi'(\omega'' - \zeta'') \\ 0 & 0 & 0 \\ \omega'\zeta'' - \zeta'\omega'' & 0 & \chi'\zeta'' - \omega'\zeta'' \end{pmatrix} \tag{48}$$

$$T^1_{ij2} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{49}$$

$$T^1_{ij3} = \begin{pmatrix} \chi''(\zeta' - \omega') & 0 & \omega'\chi'' - \chi'\omega'' \\ 0 & 0 & 0 \\ \zeta'(\omega'' - \chi'') & 0 & 0 \end{pmatrix} \quad (50)$$

Linear F-tensor and G-tensor Computation from Seven Points and Three Images. There are only 11 homogeneous variables (or 10 unknowns) in all four tensors. Writing the constraints as in (45), and since there are four constraints with 10 unknowns, we can compute the elements of the shape tensors from three images or more, or the camera tensors from seven points or more. In either case, we are left with an overconstrained linear system of equations for the 10 unknowns using at least three images or at least seven points.

Given the elements of the F- and G-tensors, we want to compute the coordinates of the camera positions and the space points. As an intermediate step we compute the parameters $\chi', \psi', \zeta', \omega'$ and $\chi'', \psi'', \zeta'', \omega''$. If we let $\{q_i\}_{i=1}^{11}$ denote the 11 different elements of the 3D F- or G-tensors from (45), we get from (45) the explicit solutions for the normalized variables

$$\frac{\chi'}{\omega'} = \frac{q_1 - q_2}{q_{11}}, \quad \frac{\psi'}{\omega'} = \frac{q_4 - q_5}{q_{10}}, \quad \frac{\zeta'}{\omega'} = \frac{q_7 - q_8}{q_{10} - q_{11}},$$
$$\frac{\chi''}{\omega''} = \frac{q_4 - q_7}{q_6 - q_9}, \quad \frac{\psi''}{\omega''} = \frac{q_1 - q_8}{q_3 - q_9}, \quad \frac{\zeta''}{\omega''} = \frac{q_2 - q_5}{q_3 - q_6}. \quad (51)$$

Nonlinear F-tensor Computation from Six Points and Three Images and G-tensor from Seven Points and Two Images. We can normalize the homogeneous four-vectors $\chi', \psi', \zeta', \omega'$ and $\chi'', \psi'', \zeta'', \omega''$ by setting $\omega' = \omega'' = 1$, which leaves us with six unknowns. Thus it is possible to *nonlinearly* compute the shape parameters from two images of seven points, and the camera parameters from three images of six points.

6.4. Quadrilinear Equations

We are given a constraint matrix composed of the first eight rows of matrix C from (44). There are $\binom{8}{4} = 70$ constraints when considering all 4×4 minors of C.

Using Gröbner bases in the Computer Algebra System SINGULAR, we computed an algebraic basis for the space spanned by the 15 constraints. This basis has seven equations: three bilinear equations involving subsets of six points, and 12 trilinear equations involving seven points. Thus there are no algebraically new constraints, and we get no new equations on the original variables $\chi', \psi', \zeta', \omega', \chi'', \psi'', \zeta'', \omega''$.

However, there are new quadrilinear equations that define 22 independent linear constraints on 41 unknowns. We can write each quadrilinear constraint in a tensor form, using $3 \times 3 \times 3 \times 3$ tensors Q^l:

$$\sum_{i,j,k,l} \mathsf{Q}^l_{ijkl}(p)_i(p')_j(p'')_k(p''')_l = 0, \quad l = 1, \ldots, 22 \quad (52)$$

Linear F-tensor Computation from Six Points and Four Images and G-tensor from Eight Points and Two Images. Each of the 22 constraints has a different tensor. However, there are only 41 unknowns in all 22 tensors, and writing the constraints as in 45 allows us to compute all F-tensors from two images or more. Dually we can compute all the G-tensors using six or more points in four images. Given the elements of the F- and G-tensors, we once again compute the parameters $\chi', \psi', \zeta', \omega', \chi'', \psi'', \zeta'', \omega''$ and $\chi''', \psi''', \zeta''', \omega'''$. If we let $\{q_i\}_{i=1}^{41}$ denote the 41 different elements of the 3D F- or G-tensors, we get the explicit solutions for the normalized variables:

$$\frac{\chi'}{\omega'} = \frac{q_{35} - q_{38}}{q_{15} - q_{11}}, \quad \frac{\psi'}{\omega'} = \frac{q_{28} - q_{21}}{q_3 - q_{11}}, \quad \frac{\zeta'}{\omega'} = \frac{q_9 - q_{33}}{q_3 - q_{15}},$$
$$\frac{\chi''}{\omega''} = \frac{q_8 - q_2}{q_{13} - q_1}, \quad \frac{\psi''}{\omega''} = \frac{q_4 - q_{14}}{q_{12} - q_1}, \quad \frac{\zeta''}{\omega''} = \frac{q_{16} - q_{17}}{q_{12} - q_{13}},$$
$$\frac{\chi'''}{\omega'''} = \frac{q_{39} - q_{40}}{q_{24} - q_7}, \quad \frac{\psi'''}{\omega'''} = \frac{q_{22} - q_{23}}{q_{41} - q_7}, \quad \frac{\zeta'''}{\omega'''} = \frac{q_{31} - q_5}{q_{41} - q_{24}}. \quad (53)$$

Nonlinear F-tensor and G-tensor Computation. No new nonlinear algorithm can be described because no new independent constraint equation is obtained from the quadrilinear equations.

6.5. Space-Point and Camera-Point Relations

The tri-linear constraints for external calibration described in (Hartley, 1994; Shashua, 1994) are obtained from (45), and the quadrilinear constraints described in (Faugeras and Mourrain, 1995; Shashua and Werman, 1995; Triggs, 1995) are obtained from the corresponding quadrilinear system. This is done by making the substitutions defined in (40) and (41). We can use these equations to find camera positions. If we choose the coordinates of the center of camera a to be at $[1, 1, 1, 1]$, then χ', ψ', ζ', ω' give the center of camera b (inverse), χ'', ψ'', ζ'', ω'' give the center of camera c (inverse), and χ''', ψ''', ζ''', ω''' give the center of camera d (inverse), in the basis defined by the first four space points and the camera position a.

Similarly, the dual trilinear relations between seven points in three images are obtained from (45), and the dual quadrilinear relations between eight points in two images are obtained from the corresponding quadrilinear system. This is done by substituting (42) and (43), for eight points $P_1, \ldots, P_8$ observed in the image of camera a. If we choose the coordinates of the fifth space point to be at [1, 1, 1, 1], then χ', ψ', ζ', ω' give the coordinates of point 6, χ'', ψ'', ζ'', ω'' give the coordinates of point 7, and χ''', ψ''', ζ''', ω''' give the coordinates of point 8 in the basis defined by the first five space points.

In order to relate the alternative representations for different choices of basis points we may use the transformations described in Section 5.4.

7. Discussion

7.1. The Duality between Reconstruction and Positioning

We showed that the constraints we get for camera positions from p points in q views are mathematically identical to the constraints for space points from $q + 4$ points in $p - 4$ 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. Similarly, the algorithms for direct computation of projective structure can be used to obtain relative camera positions. This is illustrated in Fig. 3. Figure 4 illustrates this for various published algorithms.

As was discussed in Section 5, the linear eight-points algorithm for computing the F-matrix (Longuet-Higgins and Prazdny, 1980) can be used for direct computation of projective structure for six points in four images. Likewise, as pointed out in (Shashua, 1994), there is a linear algorithm for camera geometry for seven points in three views. This algorithm could equally well be applied to the projective reconstruction problem with the same number of points and views, as discussed in Section 6. A systematic description of direct reconstruction algorithms has been presented in (Heyden, 1995b).

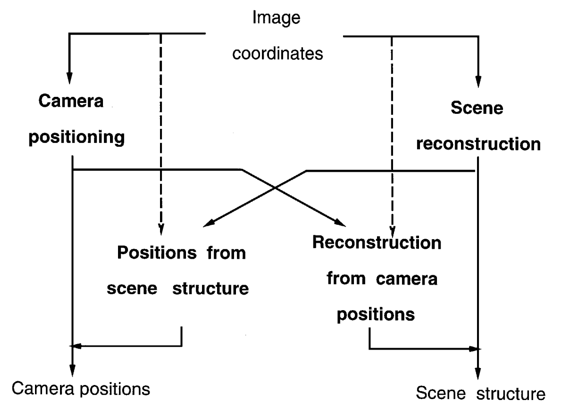

Figure 3. Dual pathways for camera positioning and scene reconstruction from image coordinates.

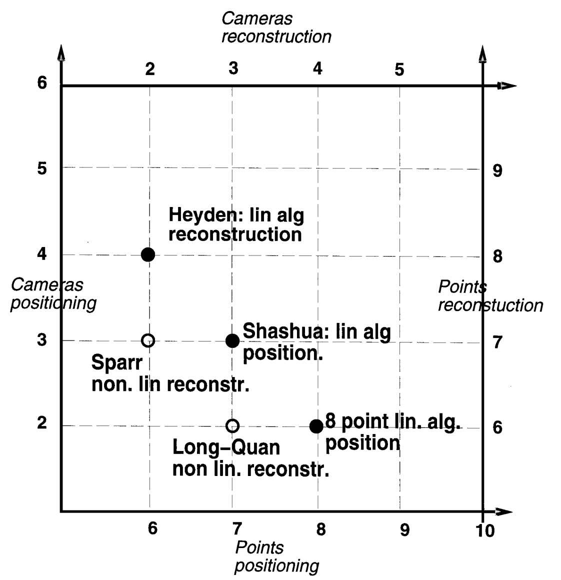

Figure 4. The problem of positioning cameras from p points in n views is mathematically identical to the problem of reconstructing $n + 4$ points in $p - 4$ views (The points indicate work in which algorithms were discussed for the first time).

Table 1 gives a summary of the dual algorithms, and how they can be obtained from the above analysis:

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

Table 1. Summary of dual algorithms for reconstruction and positioning.

Positioning		Reconstruction				
Points	Frames	Points	Frames	Constraint	Algorithm	Discussion
8	2	6	4	Bilinear	Linear	Section 5.1
7	3	7	3	Trilinear	Linear	Section 6.3
6	4	8	2	Quadrilinear	Linear	Section 6.4
7	2	6	3	Bilinear	Nonlinear	Section 5.3
6	3	7	2	Trilinear	Nonlinear	Section 6.3

of what the best choice in a given situation is. 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 (Maybank, 1990) should have a dual counterpart in ambiguous camera configurations.

7.2. Conditions for Duality

In the discussion above we used linear invariant representation of points in the image and in space, which made the exact orientation of the image plane irrelevant (see Fig. 1). Thus the camera was fully characterized by its position—a vector in $\mathcal{P}^3$. The general case can be handled in a similar way, only that it will be necessary to substitute the vector of image measurements by:

$$\begin{pmatrix} x_n^a \\ y_n^a \\ w_n^a \end{pmatrix} \Longrightarrow Q \begin{pmatrix} x_n^a \\ y_n^a \\ w_n^a \end{pmatrix} \tag{54}$$

where Q is some nonsingular 3×3 matrix. Now the parameterization of the camera includes the center of projection, *and* the orientation of the image plane represented by the 8 numbers in matrix Q.

Substituting (54) into (35) *would not* change significantly the essence of the following derivations. Indeed the fundamental matrix and the multilinear camera tensors will now depend on 12 numbers (the camera position and the image orientation) instead of 4 (only the camera position) from each camera, but it is still possible to derive multilinear camera-image relations which are independent of shape. This is the essence of the analysis described in, e.g., (Faugeras and Mourrain, 1995; Hartley, 1994; Heyden, 1995a; Luong and Viéville, 1994; Shashua and Werman, 1995, Triggs, 1995).

On the other hand, substituting (54) into (36) *would* significantly change the essence of the following derivations. Now the G-matrix and the multilinear shape tensors depend on the shape and the camera—they now depend on the orientation of the image via matrix Q. It is not possible anymore to derive multilinear space-image relations which are independent of the camera. Thus the normalization of the image plane is essential for the derivation of space-image relations, and for the usefulness of the duality observation.

Acknowledgments

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. Vision research at the Hebrew University is supported by the U.S. Office of Naval Research under Grant N00014-93-1-1202, R&T Project Code 4424341—01. Both authors acknowledge the support of the EC-Israel Exploratory Collaboration Activity, EC-IS-003, SAM, Shape and Motion

References

Carlsson, S. 1994. The double algebra: An effective tool for computing invariants in computer vision. *Applications of Invariance in Computer Vision*, Springer, LNCS 825, pp. 145–164.

Carlsson, S. 1995a. View variation and linear invariants in 2-D and 3-D. Tech. rep., Royal Institute of Technology, ISRN KTH/NA/P–95/22–SE.

Carlsson, S. 1995b. Duality of reconstruction and positioning from projective views. In *IEEE Workshop on Representations of Visual Scenes*, Cambridge, MA.

Chen, H.H. and Huang, T.S. 1990. Matching 3-D line segments with applications to multiple-object motion estimation. *T-PAMI 12*, pp. 1002–1008.

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

Demey, S., Ziserman, A., and Beardley, P. 1992. Affine and projective structure from motion. In *Proc. BMVC-92*.

Faugeras, O.D. 1992. What can be seen in three dimensions with an uncalibrated stereo rig? In *Proc. 2nd ECCV*, pp. 563–578.

Faugeras, O.D., Luong, Q.T., and Maybank, S.J. 1992. Camera self-calibration: Theory and experiments. In *Proc. 2nd ECCV*, pp. 321–334.

Faugeras, O.D. and Mourrain, B. 1995. Algebraic and geometric properties of point correspondences between N images. In *Proc 5th ICCV*, pp. 951–956.

Gros, P. 1994. How to use the cross ratio to compute invariants from two images. *Applications of Invariance in Computer Vision*, Springer, LNCS 825, pp. 107–126.

Hartley, R.I. 1994. *Lines and Points in Three Views—An Integrated Approach IUWS*.

Hartley, R., Gupta, R., and Chang, T. 1992. Stereo from uncalibrated cameras. In *Proc. CVPR*, pp. 761–764.

Heyden, A. 1995a. Reconstruction from image sequences by means of relative depths. In *Proc 5th ICCV*, pp. 1058–1063.

Heyden, A. 1995b. Reconstruction from multiple images using kinetic depths. Tech. Report. Dep. of Mathematics, Lund University.

Irani, M. and Anandan, P. 1996. Parallax geometry of pairs of points for 3D scence analysis. In *Proc. 4th ECCV*, vol. I, pp. 17–30.

Koenderink, J.J. and van Doorn, A.J. 1991. Affine structure from motion, *J. Optic. Soc. Am. A*, 2:377–385.

Long Quan. 1994. Invariants of 6 points from 3 uncalibrated images. In *Proc. 3rd ECCV*, vol. 2, pp. 459–470.

Longuet-Higgins, H.C. and Prazdny, K. 1980. The interpretation of a moving retinal image. In *Proc. Roy. Soc. Lond. B-208*, pp. 385–397.

Luong, Q.-T. and Viéville, T. 1994. Canonic representations for the geometries of multiple projective views. In *Proc. 3rd ECCV*.

Maybank, S.J. 1990. The projective geometry of ambiguous surfaces. In *Proc. Royal Soc. A*, vol. 322, pp. 1–47.

Mohr, R. 1992. Projective geometry and computer vision. *Handbook of Pattern Recognition and Computer Vision*.

Mohr, R., Quan, L., and Veillon, F. 1995. Relative 3-D reconstruction using multiple uncalibrated images international. *Journal of Robotics Research*, 6:619–632.

Rothwell, C.A., Forsyth, D.A., Zisserman, A.P., and Mundy, J.L. 1993. Extracting projective structure from single perspective views of 3-D point sets. In *Proc. of 4th ICCV*, pp. 573–582.

Shashua, A. 1993. Projective depth: A geometric invariant for 3D reconstruction from two perspective/orthographic views and for visual recognition. In *Proc 4th ICCV*, pp. 583–590.

Shashua, A. 1994. Trilinearity in visual recognition by alignment. In *Proc. 3rd ECCV A*, pp. 479–484.

Shashua, A. 1995. Algebraic functions for recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8):779–789.

Shashua, A. and Werman, M. 1995. Trilinearity of three perspective views and its associated tensor. In *Proc 5th ICCV*, pp. 920–925.

Shashua, A. and Avidan, S. 1996. The rank4 constraint in multiple view geometry. In *Proc. 4th ECCV*, pp. 196–206.

Sparr, G. 1991. In *Proc. 1st ESPRIT-DARPA Workshop on Invariants in Computer Vision*, Reykjavik.

Sparr, G. 1992. Depth Computations from polyhedral images. *Image and Vision Computing*, 10(10):683–688.

Sparr, G. 1994. A common framework for kinetic depth, reconstruction and motion for deformable objects. In *Proc. ECCV B*, pp. 471–482.

Spetsakis, M.E. and Aloimonos, J.(Y.). 1990. Structure from motion using line correspondences. *International Journal of Computer Vision*, 4:171–183.

Tomasi, C. and Kanade, T. 1992. Shape and motion from image streams under orthography: A factorization method. *International Journal of Computer Vision*, 9(2):137–154.

Triggs, B. 1995. Matching constraints and the joint image. In *Proc 5th ICCV*, pp. 338–343.

Weinshall, D. 1993. Model-based invariants for 3D vision. *International Journal of Computer Vision*, 10(1):27–42.

Weinshall, D. and Tomasi, C. 1995. Linear and incremental acquisition of invariant shape models from image sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(5):512–517.

Weinshall, D., Werman, M., and Shashua, A. 1995. Shape tensors for efficient and learnable indexing. In *Proc. of the IEEE Workshop on Representations of Visual Scenes*, Cambridge, MA.

Zisserman, A. 1994. A case against epipolar geometry. *Applications of Invariance in Computer Vision*, Springer, LNCS 825, pp. 69–88.