MECSE 1997-5

Geometric Motion Constraints for Different Camera Models and for Continuous and Discrete Motion: Rigid Motion

D. Suter

Dec. 11 1997

TECHNICAL REPORT
DEPARTMENT OF ELECTRICAL AND COMPUTER SYSTEMS ENGINEERING,
FACULTY OF ENGINEERING,
MONASH UNIVERSITY, CLAYTON 3168,
VIC, AUSTRALIA

Geometric Motion Constraints for Different Camera Models and for Continuous and Discrete Motion: Rigid Motion

D. Suter
Dept. Electrical and Computer Systems Engineering
Monash University, Clayton 3168
Australia

Dec. 11 1997

1 Introduction

Solving for motion, both for image plane motion (which is often loosely termed "optic flow") and for "real world" or object motion, using image sequence data, involves some form of constraint. We can divide the constraints into three classes:

- Photometric (brightness) constraints - i.e., the constraints on the image plane (2-D) motion derived from photometric principles and expressed through the photometric quantities. e.g., the optic flow constraint derived from conservation of image brightness:

$$I_x u + I_y v = I_t \tag{1}$$

- Geometric constraints - i.e., the constraints on image plane motion derived from geometric principles of the projection process. e.g, the discrete epi-polar displacement constraints expressed through the fundamental or essential matrix (when the motion is rigid and the projection is perspective)
- Combined Geometric and Photometric Constraints e.g., the constraints of [HJ88] [SS97].

From these classes, one can derive a whole plethora of constraints: for example, by varying whether the motion is discrete or continuous, whether the motion is rigid (general) or rigid planar, and the type of projection model.

This paper is concerned solely with constraints from the second category. Moreover, we restrict ourselves to rigid (general) motion, that is, we do not investigate special cases of rigid motion (for example rigid motion of planar surfaces) for which the constraints may degenerate to alternative expressions. We cover both the case of discrete motion and the case of continuous motion. In terms of camera models, we cover (in order of generality) orthographic, weak perspective, paraperspective, affine, and perspective.

The main result of this paper is that it is shown that for all affine camera models (including those of: orthographic, weak perspective, paraperspective) and for calibrated and uncalibrated versions (the latter even allowing the calibration parameters to vary with time), and for both continuous rigid and discrete rigid motion; we find that the motion obeys a simple affine constraint of the form:

$$ax' + by' + cx + dy + e = 0 \tag{2}$$

in the discrete motion case, and,

$$au + bv + cx + dy + e = 0 \tag{3}$$

in the continuous motion case.

Such a general result does not appear to be stated in the literature. Moreover, we explicitly derive general expressions for the constants a, b, c, d, e in terms of the intrinsic and extrinsic parameters (the former including calibration parameters and the latter including motion parameters).

Weber and Malik [WM97] appear to be aware of the special case of discrete motion calibrated weak perspective. In that same paper, they use a optic flow measurements as if they were discrete displacements to employ their constraint for motion segmentation and for structure from motion. Our results show that, if the optic flow measurements really were what they are supposed to be - *velocities rather than displacements*, the form of the constraint Weber and Malik should have used is, fortuitously for them perhaps, of the same form as that they used. However, the relationship between the parameters of the constraint and the intrinsic and extrinsic parameters of the camera/motion *has a different expression to that they employed.* Of course, one may argue pragmatically that the difference will vanish in the limit of small motion and/or that one cannot truly capture infinitesimals such as actual velocities, rather one always works with small displacements. Be that is it may: the present paper puts their work, and others past or future, that employ these camera models, on a firmer theoretical footing and with clear expressions for the difference between discrete and continuous motion parameters.

A minor contribution of this paper is that we derive, for completeness, the corresponding motion constraints for (rigid) motion under (uncalibrated or calibrated) perspective. The discrete case is very well known. The continuous case has recently been put on a firmer footing [BCB97] after a previous derivation that appealed to various approximations [VF95]. The form of the constraint derived here (and in the previously cited works) is essentially equal to that known for over a decade (e.g., [Sub88]) in the calibrated case, in that it is a relatively simple algebraic manipulation to convert one to the other. However, the generality of the uncalibrated version and the easily remembered form of the constraint presented here (and in [BCB97] [VF95]) have a much greater appeal. Our interest here, in covering this established ground once more, is not only for completeness but to highlight the similarity of the constraint form (now bilinear rather than affine) to that we derive for the general affine camera models.

The structure of the papers is as follows. We first include direct derivations of the calibrated camera versions of the motion constraint. Though these results are merely special cases of our general result, we include the derivation mainly for two reasons. Firstly, the reader that is interested in either the method of derivation of the constraint, or verification of such, but does not need the generality of the affine camera, may simply consult this work for these special cases where the algebra is simple. Secondly, for those calibrated versions, which are, despite recent interest in uncalibrated cameras, still of great interest and utility, our simple expressions for the parameters of the constraint, allow ready application of our results without evaluating the relevant expressions derived for the uncalibrated affine camera. In section 2 we derive the constraint for the calibrated orthographic camera in the discrete and continuous case. In section 3 we derive the constraint for the (calibrated) weak perspective camera and continuous motion. For the discrete case, we refer the reader to [WM97] and the cited references therein, or our later general affine exposition. Similarly, in section 4, since the algebra involved is very tedious, we simply give an expression for the calibrated camera projection process and then refer to the general result. Finally, in section 5 we present the main result of this paper - the general affine camera motion constraint for uncalibrated cameras and both for continuous and discrete motion. We also round off the discussion by illustrating the derivation of, and the form of, the corresponding constraint for a perspective camera (with continuous motion).

2 Orthographic Projection

2.1 Orthographic - Continuous Motion

If the object undergoes, relative to the camera, rigid motion, we can express the object position, relative to the camera by:

$$\mathbf{P}(t) = \mathbf{R}(t)\mathbf{P}(0) + \mathbf{D}(t) \tag{4}$$

where $\mathbf{R}(t)$ is the 3 rotation matrix and $\mathbf{D}(t)$ is the displacement of the rigid motion between time t and time 0. Orthographic projection involves simply neglecting the 3rd coordinate from this expression.

If we differentiate this, evaluated at $t = 0$ we arrive at

$$\dot{\mathbf{P}}(t) = \mathbf{W}(t)\mathbf{P}(0) + \mathbf{V}(t) \tag{5}$$

If we extract the first two rows and then eliminate the unknown Z and replace X and Y by their projections x and y, we arrive at:

$$W_{23}(u - W_{12}y - V_x) = W_{13}(v - W_{21}x - V_y) \tag{6}$$

or using

$$a = W_{23} \tag{7}$$
$$b = -W_{13} \tag{8}$$
$$c = -W_{13}W_{12} \tag{9}$$
$$d = -W_{12}W_{23} \tag{10}$$
$$e = (-W_{23}V_x + W_{13}V_y) \tag{11}$$

we have

$$au + bv + cx + dy + e = 0 \tag{12}$$

We can also write the constraint in the form:

$$\begin{pmatrix} u & v & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & a \\ 0 & 0 & b \\ c & d & e \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0 \tag{13}$$

2.2 Orthographic - Discrete Motion

If the object undergoes, relative to the camera, rigid motion, we can express the new object position P' object position, relative to the camera by:

$$\mathbf{P}' = \mathbf{RP} + \mathbf{D} \tag{14}$$

Orthographic projection involves simply neglecting the 3rd coordinate from this expression.

If we extract the first two rows and then eliminate the unknown Z and replace X and Y by their projections x and y, we arrive at:

$$R_{23}(x' - R_{11}x - R_{12}y - D_x) = R_{13}(v - R_{21}x - R_{22}y - D_y) \tag{15}$$

or using

$$a = R_{23} \tag{16}$$
$$b = -R_{13} \tag{17}$$
$$c = R_{13}R_{21} - R_{11}R_{23} \tag{18}$$
$$d = R_{13}R_{22} - R_{12}R_{23} \tag{19}$$
$$e = (-R_{23}D_x + R_{13}D_y) \tag{20}$$

we have

$$ax' + by' + cx + dy + e = 0 \tag{21}$$

We can also write the constraint in the form:

$$\begin{pmatrix} x' & y' & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & a \\ 0 & 0 & b \\ c & d & e \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0 \tag{22}$$

3 Weak Perspective

3.1 Weak Perspective - Continuous Motion

Weber and Malik [WM97] derive a constraint on optic flow for images under weak perspective projection and where the camera/object motion is rigid. However, their derivation is not strictly correct in that they appeal to the equivalence between discrete displacements and optic flow. Since the later ought to be a true velocity the equivalence is only true in the limit as the time interval shrinks to zero.

Here, we re-derive the constraint following a more logically consistent (in that we work in the continuous/infinitesimal setting throughout) manner. The form of the constraint we derive is identical to that of Weber and Malik *but the parameters involved now have slightly different interpretation or relationship to the quantities underlying the real world motion and the projection.* That is, instead of a logically inconsistent (except in special cases such as constant motion) relationship between image *flow* at *an instant* and finite parameters such as camera rotation *between a pair of images*, we arrive at a relationship between the constraint parameters flow and the *instantaneous* linear and angular velocity of the object.

Assuming weak perspective projection, we have the following relationship between the image points $(x(t), y(t))$ and the object point $\mathbf{P}(t) = (X(t), Y(t), Z(t))$ projected onto that point.

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \frac{f}{Z_{ave}(t)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X(t) \\ Y(t) \\ Z(t) \end{pmatrix} \tag{23}$$

where $Z_{ave}(t)$ is the average of $Z(t)$ (over the object) and we have assumed calibration to the extent that we have ignored intrinsic camera parameters (except f) that designate shear or aspect ratio, and we have also ignored extrinsic parameters for the attitude of the camera w.r.t. t the object coordinate frame.

It is now a simple (but perhaps tedious) matter to differentiate the expression 23 (w.r.t. time and evaluated at a time instant which we can arbitrarily choose as $t = 0$). Since all quantities are evaluated at the chosen time ($t = 0$), we drop this part of the notation to give, with the obvious interpretation,

$$\begin{aligned} \begin{pmatrix} u \\ v \end{pmatrix} &= \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} \qquad (24) \\ &= \frac{-f\dot{Z}_{ave}}{Z_{ave}^2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + \\ &\quad \frac{f}{Z_{ave}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \left\{ \begin{pmatrix} 0 & W_{12} & W_{13} \\ -W_{12} & 0 & W_{23} \\ -W_{13} & -W_{23} & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + \begin{pmatrix} V_x \\ V_y \\ V_z \end{pmatrix} \right\} \end{aligned}$$

where we have used the fact that $\dot{\mathbf{R}} = \mathbf{W}$ an anti-symmetric angular velocity matrix.

We extract from this the equations for u and v (replacing X and Y by their projections according to 23):

$$u = -\frac{\dot{Z}_{ave}}{Z_{ave}}x + W_{12}y + W_{13}\frac{fZ}{Z_{ave}} + \frac{f}{Z_{ave}}V_x \tag{25}$$

$$v = -\frac{\dot{Z}_{ave}}{Z_{ave}}y - W_{12}x + W_{23}\frac{fZ}{Z_{ave}} + \frac{f}{Z_{ave}}V_y \tag{26}$$

$$\tag{27}$$

If we now eliminate the unknown Z from these two equations, we arrive at our constraint:

$$au + bv + cx + dy + e = 0 \tag{28}$$

where

$$a \;=\; W_{23} \tag{29}$$

$$b = -W_{13} \quad (30)$$

$$c = \frac{\dot{Z}_{ave}}{Z_{ave}}W_{23} - W_{13}W_{12} \quad (31)$$

$$d = -\frac{\dot{Z}_{ave}}{Z_{ave}}W_{13} - W_{12}W_{23} \quad (32)$$

$$e = \frac{f}{Z_{ave}}(-W_{23}V_x + W_{13}V_y) \quad (33)$$

We can also write the constraint in the form:

$$\begin{pmatrix} u & v & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & a \\ 0 & 0 & b \\ c & d & e \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0 \quad (34)$$

3.2 Weak Perspective - Discrete Motion

Since this is covered in the literature [WM97] and also by our general results (later) we defer further discussion except to stress that the general form of the constraint is as we have seen before.

4 Paraperspective

The paraperspective model can be defined by:

$$x = \frac{f}{Z_{ave}}X + \frac{f}{Z_{ave}}X_{ave} - \frac{f}{Z_{ave}^2}X_{ave}Z \quad (35)$$

(for object and camera coordinate systems coincident and calibrated) and the corresponding equation for the y-coordinate projection.

With such a model, even this calibrated case becomes tedious to solve. We defer such in preference to the general solution for all affine models see section 5.1.

As an example of the complexity of the relationship between the constraint parameters and the intrinsic and extrinsic camera parameters, we quote one result (for the discrete motion case):

$$b = -\frac{fX_{ave}}{Z_{ave}^2}R_{11} - \frac{fY_{ave}}{Z_{ave}^2}R_{12} + \frac{fX_{ave}}{Z_{ave}^2}R_{33} + \frac{fX_{ave}^2}{Z_{ave}^3}R_{31} + \frac{fX_{ave}Y_{ave}}{Z_{ave}^3}R_{32} \quad (36)$$

5 Uncalibrated Camera Versions

5.1 Affine Camera Models

Looking at the strategy used in deriving the geometric motion constraint we see that it can be summarised as:

Discrete

1. write down two (vector) equations for the projections of the same world point (X, Y, Z) onto two cameras, $(x, y, 1)$ and $(x', y', 1)$, assuming (w.l.o.g.) that the coordinate system of the first camera and of the world coincide.
2. eliminate X and Y from the second equation by relating them to their projected values in the first image.
3. select the two (scalar) equations for x' and y'
4. Re-arrange both equations to as to eliminate Z

Continuous

1. write down an equation for the (time varying) projection of a same world point (X, Y, Z) onto a camera $(x(t), y(t), 1)$ assuming (w.l.o.g.) that the coordinate system of the camera and of the world coincide at $t = 0$.

2. differentiate this equation
3. eliminate X and Y from the second equation by relating them to their projected values at $t = 0$.
4. select the two (scalar) equations for u and v
5. Re-arrange both equations to as to eliminate Z

of course, as the above clearly illustrates, the coefficients of Z in both equations must be non-zero for this process to work.

We show here, that so long as the projection equations are affine in the relevant variables (X, Y, Z), then the process above can be carried out and will yield an expression of the form already encountered (e.g., equation 28). Of course, in different cases the parameters of that constraint will have different values - somehow related to the intrinsic and extrinsic parameters of the projection model and/or rigid motion. Since we work in such generality, we cover all cases of orthographic, weak perspective, paraperspective and general affine (which also includes calibrated and uncalibrated cases in the sense often used in the literature where the calibration is a 5 parameter affine matrix).

Since we are "preserving" the "Z" component in our characterisation of the projection process, we adopt a slightly non-standard representation of the calibration matrix A.

$$A = \begin{pmatrix} A_s & A_t^p & A_t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{37}$$

where A_s is a 2×2 invertible submatrix (essentially scales and or shears the image plane) and A_t is a 2×1 translation component to align the image plane coordinate system with the optical centre. A_s and A_t cover the usual components of a calibrated camera that caters for linear distortions of the lens and for correction for the optical centre. The other component A_t^p is a 2×1 vector that is needed, for our purposes, only for the paraperspective camera models. In the orthographic and weak perspective formulations, this matrix can be taken as $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

5.2 Discrete Motion Case

We adopt the convention (w.l.o.g.) that the object is stationary and the that the camera moves. Moreover, we choose the original camera coordinate frame to coincide with the object coordinate frame. The the first projection process is just the affine (2D) calibration process, and the second projection process is that of an affine (3D) camera attitude and position transformation followed by an affine (2) calibration process.

In such a representation, the overall imaging process is

$$\begin{pmatrix} x \\ y \\ \mathcal{Z} \\ 1 \end{pmatrix} = \begin{pmatrix} A_s & A_t^p & A_t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} R & D \\ 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \tag{38}$$

It is easy to show/verify that

$$A^{-1} = \begin{pmatrix} A_s^{-1} & -A_s^{-1} A_t^p & -A_s^{-1} A_t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{39}$$

and, similarly,

$$\begin{pmatrix} R & D \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} R^{-1} & -R^{-1} D \\ 0 & 1 \end{pmatrix} \tag{40}$$

The equations for the two projected points (one in each camera) are:

$$\begin{pmatrix} x \\ y \\ Z \\ 1 \end{pmatrix} = \begin{pmatrix} A_s & A_t^p & A_t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \tag{41}$$

$$\begin{pmatrix} x' \\ y' \\ Z' \\ 1 \end{pmatrix} = \begin{pmatrix} A'_s & A_t^{p\prime} & A'_t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} R & D \\ 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \tag{42}$$

Using the inversion of equation 41, we can eliminate X and Y in equation 42 to arrive at:

$$\begin{pmatrix} x' \\ y' \\ Z' \\ 1 \end{pmatrix} = \begin{pmatrix} A'_s & A_t^{p\prime} & A'_t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} R & D \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A_s^{-1} \begin{pmatrix} x \\ y \end{pmatrix} - ZA_s^{-1}A_t^p - A_s^{-1}A_t \\ Z \\ 1 \end{pmatrix} \tag{43}$$

Expanding the right hand side we have:

$$\begin{pmatrix} A'_s & A_t^{p\prime} & A'_t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} R \begin{pmatrix} A_s^{-1} \begin{pmatrix} x \\ y \end{pmatrix} - ZA_s^{-1}A_t^p - A_s^{-1}A_t \\ Z \end{pmatrix} + D \\ 1 \end{pmatrix} \tag{44}$$

and

$$\begin{pmatrix} A'_s & A_t^{p\prime} & A'_t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \begin{pmatrix} R_{1[1:2]} \cdot \left(A_s^{-1} \begin{pmatrix} x \\ y \end{pmatrix} - ZA_s^{-1}A_t^p - A_s^{-1}A_t \right) + R_{13}Z \\ R_{2[1:2]} \cdot \left(A_s^{-1} \begin{pmatrix} x \\ y \end{pmatrix} - ZA_s^{-1}A_t^p - A_s^{-1}A_t \right) + R_{23}Z \\ R_{3[1:2]} \cdot \left(A_s^{-1} \begin{pmatrix} x \\ y \end{pmatrix} - ZA_s^{-1}A_t^p - A_s^{-1}A_t \right) + R_{33}Z \end{pmatrix} + D \\ 1 \end{pmatrix} \tag{45}$$

where $R_{3[1:2]}$ is the 2-vector composed of the first two elements of the third row of R.

Finally we obtain:

$$\begin{pmatrix} A'_{s\,1} \cdot \begin{pmatrix} R_{1[1:2]} \cdot \left(A_s^{-1} \begin{pmatrix} x \\ y \end{pmatrix} - ZA_s^{-1}A_t^p - A_s^{-1}A_t \right) + ZR_{13} + D_1 \\ R_{2[1:2]} \cdot \left(A_s^{-1} \begin{pmatrix} x \\ y \end{pmatrix} - ZA_s^{-1}A_t^p - A_s^{-1}A_t \right) + ZR_{23} + D_2 \end{pmatrix} + A_{t\,1}^{p\prime}\Gamma + A'_{t\,1} \\ A'_{s\,2} \cdot \begin{pmatrix} R_{1[1:2]} \cdot \left(A_s^{-1} \begin{pmatrix} x \\ y \end{pmatrix} - ZA_s^{-1}A_t^p - A_s^{-1}A_t \right) + ZR_{13} + D_1 \\ R_{2[1:2]} \cdot \left(A_s^{-1} \begin{pmatrix} x \\ y \end{pmatrix} - ZA_s^{-1}A_t^p - A_s^{-1}A_t \right) + ZR_{13} + D_2 \end{pmatrix} + A_{t\,2}^{p\prime}\Gamma + A'_{t\,2} \\ \Gamma \\ 1 \end{pmatrix} \tag{46}$$

Where we have used the abbreviation:

$$\Gamma = R_{3[1:2]} \cdot \left(A_s^{-1} \begin{pmatrix} x \\ y \end{pmatrix} - ZA_s^{-1}A_t^p - A_s^{-1}A_t \right) + D_3 + R_{33}Z \tag{47}$$

We now have two equations (the first two rows giving expressions for x' and y') with a common unknown Z, and, since this unknown enters into the equations by simple multiplication with a coefficient, we can easily eliminate Z from the equations. Indeed, the equations are of the form:

$$x' = \alpha x + \beta y + \gamma Z + \delta \tag{48}$$

$$y' = \eta x + \phi y + \tau Z + \rho \tag{49}$$

so that if $\tau \neq 0$ and $\gamma \neq 0$ we have

$$\gamma(y' - \eta x - \phi y - \rho) = \tau(x' - \alpha x - \beta y - \delta) \tag{50}$$

or

$$\tau x' - \gamma y' + (\gamma\eta - \tau\alpha)x + (\gamma\phi - \tau\beta)y + (\gamma\rho - \tau\delta) = 0 \tag{51}$$

Inspection of the equations reveals that

$$\gamma = -A'_{s\,1} \cdot \begin{pmatrix} R_{1[1:2]} \cdot A_s^{-1} A_t^p - R_{13} \\ R_{2[1:2]} \cdot A_s^{-1} A_t^p - R_{23} \end{pmatrix} + A^{p\prime}_{t\,1}(R_{33} - R_{3[1:2]} \cdot A_s^{-1} A_t^p) \tag{52}$$

$$\tau = -A'_{s\,2} \cdot \begin{pmatrix} R_{1[1:2]} \cdot A_s^{-1} A_t^p - R_{13} \\ R_{2[1:2]} \cdot A_s^{-1} A_t^p - R_{23} \end{pmatrix} + A^{p\prime}_{t\,2}(R_{33} - R_{3[1:2]} \cdot A_s^{-1} A_t^p) \tag{53}$$

$$\delta = A'_{s\,1} \cdot \begin{pmatrix} R_{1[1:2]} \cdot \left(-A_s^{-1} A_t \right) + D_1 \\ R_{2[1:2]} \cdot \left(-A_s^{-1} A_t \right) + D_2 \end{pmatrix} + A'_{t\,1} + A^{p\prime}_{t\,1}\left(-R_{3[1:2]} \cdot A_s^{-1} A_t + D_3\right) \tag{54}$$

$$\rho = A'_{s\,2} \cdot \begin{pmatrix} R_{1[1:2]} \cdot \left(-A_s^{-1} A_t \right) + D_1 \\ R_{2[1:2]} \cdot \left(-A_s^{-1} A_t \right) + D_2 \end{pmatrix} + A'_{t\,2} + A^{p\prime}_{t\,2}\left(-R_{3[1:2]} \cdot A_s^{-1} A_t + D_3\right) \tag{55}$$

$$\alpha = A'_{s\,1} \cdot \begin{pmatrix} R_{1[1:2]} \cdot {A_s^{-1}}_{[1:2]1} \\ R_{2[1:2]} \cdot {A_s^{-1}}_{[1:2]1} \end{pmatrix} + A^{p\prime}_{t\,1} R_{3[1:2]} \cdot {A_s^{-1}}_{[1:2]1} \tag{56}$$

$$\eta = A'_{s\,2} \cdot \begin{pmatrix} R_{1[1:2]} \cdot {A_s^{-1}}_{[1:2]1} \\ R_{2[1:2]} \cdot {A_s^{-1}}_{[1:2]1} \end{pmatrix} + A^{p\prime}_{t\,2} R_{3[1:2]} \cdot {A_s^{-1}}_{[1:2]1} \tag{57}$$

$$\beta = A'_{s\,1} \cdot \begin{pmatrix} R_{1[1:2]} \cdot {A_s^{-1}}_{[1:2]2} \\ R_{2[1:2]} \cdot {A_s^{-1}}_{[1:2]2} \end{pmatrix} + A^{p\prime}_{t\,1} R_{3[1:2]} \cdot {A_s^{-1}}_{[1:2]2} \tag{58}$$

$$\phi = A'_{s\,2} \cdot \begin{pmatrix} R_{1[1:2]} \cdot {A_s^{-1}}_{[1:2]2} \\ R_{2[1:2]} \cdot {A_s^{-1}}_{[1:2]2} \end{pmatrix} + A^{p\prime}_{t\,2} R_{3[1:2]} \cdot {A_s^{-1}}_{[1:2]2} \tag{59}$$

For the discrete motion calibrated **orthographic** camera, one has:

$$A = A' = I \tag{60}$$

and the above relationships can be seen to agree with those derived before.

For the discrete motion calibrated **weak perspective** camera, one has:

$$A = A' = \begin{pmatrix} \frac{f}{Z_{ave}} & 0 & 0 & 0 \\ 0 & \frac{f}{Z_{ave}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{61}$$

$$A^{-1} = A^{-1\prime} = \begin{pmatrix} \frac{Z_{ave}}{f} & 0 & 0 & 0 \\ 0 & \frac{Z_{ave}}{f} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{62}$$

$$\tag{63}$$

and the above relationships can again be seen to agree with those derived before.

For the discrete motion calibrated **paraperspective** camera, one has:

$$A = A' = \begin{pmatrix} \frac{f}{Z_{ave}} & 0 & -\frac{fX_{ave}}{Z_{ave}^2} & \frac{fX_{ave}}{Z_{ave}} \\ 0 & \frac{f}{Z_{ave}} & -\frac{fY_{ave}}{Z_{ave}^2} & \frac{fY_{ave}}{Z_{ave}} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{64}$$

$$A^{-1} = A^{-1'} = \begin{pmatrix} \frac{Z_{ave}}{f} & 0 & \frac{X_{ave}}{Z_{ave}} & -X_{ave} \\ 0 & \frac{Z_{ave}}{f} & \frac{Y_{ave}}{Z_{ave}} & -Y_{ave} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{65}$$

$$\tag{66}$$

5.3 Continuous Motion Case

We find it convenient to assume that object point is moving and the camera is stationary. In such a representation, the overall imaging process is

$$\begin{pmatrix} x(t) \\ y(t) \\ Z(t) \\ 1 \end{pmatrix} = \begin{pmatrix} A_s(t) & A_t^p(t) & A_t(t) \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X(t) \\ Y(t) \\ Z(t) \\ 1 \end{pmatrix} \tag{67}$$

The inverse of the A matrix has the corresponding form to that given in the discrete case.

Taking the time derivative, (and evaluated at the time $t = 0$) we arrive at:

$$\begin{pmatrix} u \\ v \\ \dot{Z} \\ 0 \end{pmatrix} = \begin{pmatrix} \dot{A}_s & \dot{A}_t^p & \dot{A}_t \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} + \begin{pmatrix} A_s & A_t^p & A_t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{pmatrix} W & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} + \begin{pmatrix} V_y \\ V_y \\ V_z \\ 0 \end{pmatrix} \right) \tag{68}$$

Using the inversion of equation 41, we can eliminate X and Y to arrive at:

$$\begin{pmatrix} u \\ v \\ \dot{Z} \\ 0 \end{pmatrix} = \begin{pmatrix} \dot{A}_s & \dot{A}_t^p & \dot{A}_t \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} A_s^{-1} \begin{pmatrix} x \\ y \end{pmatrix} - ZA_s^{-1}A_t^p - A_s^{-1}A_t \\ Z \\ 1 \end{pmatrix} \tag{69}$$

$$+ \begin{pmatrix} A_s & A_t^p & A_t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{pmatrix} W & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} A_s^{-1} \begin{pmatrix} x \\ y \end{pmatrix} - ZA_s^{-1}A_t^p - A_s^{-1}A_t \\ Z \\ 1 \end{pmatrix} + \begin{pmatrix} V_x \\ V_y \\ V_z \\ 0 \end{pmatrix} \right) \tag{70}$$

Expanding:

$$\begin{pmatrix} u \\ v \\ \dot{Z} \\ 0 \end{pmatrix} = \begin{pmatrix} \dot{A}_s \left(A_s^{-1} \begin{pmatrix} x \\ y \end{pmatrix} - ZA_s^{-1}A_t^p - A_s^{-1}A_t \right) + Z\dot{A}_t^p + \dot{A}_t \\ 0 \\ 0 \end{pmatrix} \tag{71}$$

$$+ \begin{pmatrix} A_s & A_t^p & A_t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} W_{1[1:2]} \cdot \left(A_s^{-1} \begin{pmatrix} x \\ y \end{pmatrix} - ZA_s^{-1}A_t^p - A_s^{-1}A_t \right) + W_{13}Z + V_x \\ W_{2[1:2]} \cdot \left(A_s^{-1} \begin{pmatrix} x \\ y \end{pmatrix} - ZA_s^{-1}A_t^p - A_s^{-1}A_t \right) + W_{23}Z + V_y \\ W_{3[1:2]} \cdot \left(A_s^{-1} \begin{pmatrix} x \\ y \end{pmatrix} - ZA_s^{-1}A_t^p - A_s^{-1}A_t \right) + W_{33}Z + V_z \\ 0 \end{pmatrix} \tag{72}$$

And, finally,

$$\begin{pmatrix} u \\ v \\ \dot{Z} \\ 0 \end{pmatrix} = \begin{pmatrix} \dot{A}_s \left(A_s^{-1} \begin{pmatrix} x \\ y \end{pmatrix} - ZA_s^{-1}A_t^p - A_s^{-1}A_t\right) + Z\dot{A}_t^p + \dot{A}_t \\ 0 \\ 0 \end{pmatrix} \tag{73}$$

$$+ \begin{pmatrix} A_{s\,1} \cdot \begin{pmatrix} W_{1[1:2]} \cdot \left(A_s^{-1} \begin{pmatrix} x \\ y \end{pmatrix} - ZA_s^{-1}A_t^p - A_s^{-1}A_t\right) + W_{13}Z + V_x \\ W_{2[1:2]} \cdot \left(A_s^{-1} \begin{pmatrix} x \\ y \end{pmatrix} - ZA_s^{-1}A_t^p - A_s^{-1}A_t\right) + W_{23}Z + V_y \end{pmatrix} + \Gamma A_{t\,1}^p \\ A_{s\,2} \cdot \begin{pmatrix} W_{1[1:2]} \cdot \left(A_s^{-1} \begin{pmatrix} x \\ y \end{pmatrix} - ZA_s^{-1}A_t^p - A_s^{-1}A_t\right) + W_{13}Z + V_x \\ W_{2[1:2]} \cdot \left(A_s^{-1} \begin{pmatrix} x \\ y \end{pmatrix} - ZA_s^{-1}A_t^p - A_s^{-1}A_t\right) + W_{23}Z + V_y \end{pmatrix} + \Gamma A_{t\,2}^p \\ \Gamma \\ 0 \end{pmatrix} \tag{74}$$

where we have used the abbreviation:

$$\Gamma = W_{3[1:2]} \cdot \left(A_s^{-1} \begin{pmatrix} x \\ y \end{pmatrix} - ZA_s^{-1}A_t^p - A_s^{-1}A_t\right) + W_{33}Z + V_z \tag{75}$$

We now have two equations (the first two rows giving expressions for u and v) with a common unknown Z, and, since this unknown enters into the equations by simple multiplication with a coefficient, we can easily eliminate Z from the equations. Indeed, the equations are of the form:

$$u = \alpha x + \beta y + \gamma Z + \delta \tag{76}$$
$$v = \eta x + \phi y + \tau Z + \rho \tag{77}$$

so that if $\tau \neq 0$ and $\gamma \neq 0$ we have

$$\gamma(y' - \eta x - \phi y - \rho) = \tau(x' - \alpha x - \beta y - \delta) \tag{78}$$

or

$$\tau x' - \gamma y' + (\gamma\eta - \tau\alpha)x + (\gamma\phi - \tau\beta)y + (\gamma\rho - \tau\delta) = 0 \tag{79}$$

Inspection of the equations reveals that

$$\gamma = -A_{s\,1} \cdot \begin{pmatrix} W_{1[1:2]} \cdot A_s^{-1}A_t^p - W_{13} \\ W_{2[1:2]} \cdot A_s^{-1}A_t^p - W_{23} \end{pmatrix} + A_{t\,1}^p(-_{3[1:2]} \cdot A_s^{-1}A_t^p + W_{33}) + -\dot{A}_{s\,1} \cdot A_s^{-1}A_t^p + \dot{A}_{t\,1}^p \tag{80}$$

$$\tau = -A_{s\,2} \cdot \begin{pmatrix} W_{1[1:2]} \cdot A_s^{-1}A_t^p - W_{13} \\ W_{2[1:2]} \cdot A_s^{-1}A_t^p - W_{23} \end{pmatrix} + A_{t\,2}^p(-W_{3[1:2]} \cdot A_s^{-1}A_t^p + W_{33}) + -\dot{A}_{s\,2} \cdot A_s^{-1}A_t^p + \dot{A}_{t\,2}^p \tag{81}$$

$$\delta = A_{s\,1} \cdot \begin{pmatrix} W_{1[1:2]} \cdot \left(-A_s^{-1}A_t \right) + V_x \\ W_{2[1:2]} \cdot \left(-A_s^{-1}A_t \right) + V_y \end{pmatrix} - \dot{A}_{s\,1} \cdot A_s^{-1}A_t + \dot{A}_{t1} + A_{t\,1}^p\left(-W_{3[1:2]} \cdot A_s^{-1}A_t + V_z\right) \tag{82}$$

$$\rho = A_{s\,2} \cdot \begin{pmatrix} W_{1[1:2]} \cdot \left(-A_s^{-1}A_t \right) + V_x \\ W_{2[1:2]} \cdot \left(-A_s^{-1}A_t \right) + V_y \end{pmatrix} - \dot{A}_{s\,2} \cdot A_s^{-1}A_t + \dot{A}_{t\,2} + A_{t\,2}^p\left(-W_{3[1:2]} \cdot A_s^{-1}A_t + V_z\right) \tag{83}$$

$$\alpha = A_{s\,1} \cdot \begin{pmatrix} W_{1[1:2]} \cdot {A_s^{-1}}_{[1:2]1} \\ W_{2[1:2]} \cdot {A_s^{-1}}_{[1:2]1} \end{pmatrix} + A_{t\,1}^p W_{3[1:2]} \cdot {A_s^{-1}}_{[1:2]1} + \dot{A}_{s\,1} \cdot {A_s^{-1}}_{[1:2]1} \tag{84}$$

$$\eta = A_{s2} \cdot \begin{pmatrix} W_{1[1:2]} \cdot {A_s^{-1}}_{[1:2]1} \\ W_{2[1:2]} \cdot {A_s^{-1}}_{[1:2]1} \end{pmatrix} + A^p_{t\,2} W_{3[1:2]} \cdot {A_s^{-1}}_{[1:2]1} + \dot{A}_{s2} \cdot {A_s^{-1}}_{[1:2]1} \tag{85}$$

$$\beta = A_{s1} \cdot \begin{pmatrix} W_{1[1:2]} \cdot {A_s^{-1}}_{[1:2]2} \\ W_{2[1:2]} \cdot {A_s^{-1}}_{[1:2]2} \end{pmatrix} + A^p_{t\,1} W_{3[1:2]} \cdot {A_s^{-1}}_{[1:2]2} + \dot{A}_{s1} \cdot {A_s^{-1}}_{[1:2]2} \tag{86}$$

$$\phi = A_{s2} \cdot \begin{pmatrix} W_{1[1:2]} \cdot {A_s^{-1}}_{[1:2]2} \\ W_{2[1:2]} \cdot {A_s^{-1}}_{[1:2]2} \end{pmatrix} + A^p_{t\,2} W_{3[1:2]} \cdot {A_s^{-1}}_{[1:2]2} + \dot{A}_{s2} \cdot {A_s^{-1}}_{[1:2]2} \tag{87}$$

For the continuous motion calibrated **orthographic** camera, one has:

$$A = A' = I \tag{88}$$

and the above relationships can be seen to agree with those derived before.

For the continuous motion calibrated **weak perspective** camera, one has:

$$A = A' = \begin{pmatrix} \frac{f}{Z_{ave}} & 0 & 0 & 0 \\ 0 & \frac{f}{Z_{ave}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{89}$$

$$A^{-1} = A^{-1'} = \begin{pmatrix} \frac{Z_{ave}}{f} & 0 & 0 & 0 \\ 0 & \frac{Z_{ave}}{f} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{90}$$

$$\tag{91}$$

and the above relationships can again be seen to agree with those derived before.

For the continuous motion calibrated **paraperspective** camera, one has:

$$A = A' = \begin{pmatrix} \frac{f}{Z_{ave}} & 0 & -\frac{fX_{ave}}{Z^2_{ave}} & \frac{fX_{ave}}{Z_{ave}} \\ 0 & \frac{f}{Z_{ave}} & -\frac{fY_{ave}}{Z^2_{ave}} & \frac{fY_{ave}}{Z_{ave}} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{92}$$

$$A^{-1} = A^{-1'} = \begin{pmatrix} \frac{Z_{ave}}{f} & 0 & \frac{X_{ave}}{Z_{ave}} & -X_{ave} \\ 0 & \frac{Z_{ave}}{f} & \frac{Y_{ave}}{Z_{ave}} & -Y_{ave} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{93}$$

$$\tag{94}$$

5.4 Perspective

We re-derive here the optic flow geometric constraint of Vieville and Faugeras [VF95]. We also note that Brooks et. al. [BCB97] have also derived this constraint in the exact manner (as opposed to the first order approximation manner of [VF95]).

For an uncalibrated camera with varying intrinsic parameters $\mathbf{A}(t)$, and relative motion between object points and camera given by $\mathbf{P}(t)$, we have:

$$Z(t)\mathbf{m}(t) = \mathbf{A}(t)\mathbf{P}(\mathbf{t}) \tag{95}$$

where $\mathbf{m}(t)$ is the imaged position of $\mathbf{P}(\mathbf{t})$. Note that we are representing the projected position in homogeneous coordinates $\mathbf{m}(t) = (x, y, 1)$, on the other hand we represent the object point in non-homogeneous coordinates. This fact together with the notion of separating the Z scaling from the projection matrix, and "preserving the third component" has the advantage of leaving the "projection matrix" $\mathbf{A}(t)$ invertible. Substituting from the equation 4 to 95 we arrive at

$$Z(t)\mathbf{m}(t) = \mathbf{A}(t)\mathbf{R}(t)\mathbf{P(0)} + \mathbf{A}(t)\mathbf{D(t)} \tag{96}$$

We now differentiate both sides, w.r.t. and evaluate at $t = 0$ (we then drop the explicit reference to the time instant):

$$\dot{Z}\mathbf{m} + Z\dot{\mathbf{m}} = \dot{\mathbf{A}}\mathbf{P} + \mathbf{AWP} + \mathbf{AV} \tag{97}$$

- note that $\dot{\mathbf{m}} = (u, v, 0)$.

Using the fact that $\mathbf{A}(t)$ is invertible we may write $\mathbf{P} = Z\mathbf{A^{-1}m}$:

$$\dot{Z}\mathbf{m} + Z\dot{\mathbf{m}} = \dot{\mathbf{A}}Z\mathbf{A^{-1}m} + \mathbf{AW}Z\mathbf{A^{-1}m} + \mathbf{AV} \tag{98}$$

Now we form the anti-symmetric matrix $\hat{\mathbf{AV}}$ from the vector $\mathbf{AV}$ (such that $\hat{\mathbf{AV}} \times \mathbf{a}$ gives the vector cross product for any vector $\mathbf{a}$). Multiplying both sides by this gives us:

$$\dot{Z}\hat{\mathbf{AV}}\mathbf{m} + Z\hat{\mathbf{AV}}\dot{\mathbf{m}} = \hat{\mathbf{AV}}\dot{\mathbf{A}}Z\mathbf{A^{-1}m} + \hat{\mathbf{AV}}\mathbf{AW}Z\mathbf{A^{-1}m} \tag{99}$$

Now multiplying both sides by $\mathbf{m^T}$ gives:

$$Z\mathbf{m^T}\hat{\mathbf{AV}}\dot{\mathbf{m}} = \mathbf{m^T}\hat{\mathbf{AV}}\dot{\mathbf{A}}Z\mathbf{A^{-1}m} + \mathbf{m^T}\hat{\mathbf{AV}}\mathbf{AW}Z\mathbf{A^{-1}m} \tag{100}$$

where we have used the fact that $\mathbf{a^T Aa} = 0$ for any vector $\mathbf{a}$ and anti-symmetric matrix $\mathbf{A}$. Finally, rearranging, we arrive at

$$\mathbf{m^T}\hat{\mathbf{AV}}\dot{\mathbf{m}} - \mathbf{m^T}(\hat{\mathbf{AV}}(\dot{\mathbf{A}}\mathbf{A^{-1}} + \mathbf{AWA^{-1}}))\mathbf{m} = 0 \tag{101}$$

The above is the required constraint.

We note that one can easily derive special cases. For example, if the camera is calibrated $\mathbf{A} = \mathbf{I}$ and so:

$$\mathbf{m^T}\hat{\mathbf{V}}\dot{\mathbf{m}} - \mathbf{m^T}(\hat{\mathbf{V}}(I + \mathbf{W}))\mathbf{m} = 0 \tag{102}$$

or

$$\mathbf{m^T}\hat{\mathbf{V}}\dot{\mathbf{m}} - \mathbf{m^T}\hat{\mathbf{V}}\mathbf{Wm} = 0 \tag{103}$$

We also note that, since we multiplied throughout by $\hat{AV}$, that if $V = 0$, then we have no constraint. However, in the case $V = 0$ we can multiply throughout by *any* non-zero antisymmetric matrix, $\hat{B}$, say, and we still obtain a constraint of the same form! However, of course, now the parameters in the constraint no longer have the same real world meaning.

In the calibrated case, if $W = 0$ and $V_x = V_y = 0$, then we obtain

$$-V_z vx + V_z yu = 0 \tag{104}$$

or, in other words, (u, v) is parallel to (x, y) - the flow is purely divergent.

References

[BCB97] M. J. Brooks, W. Chojnacki, and L. Baumela. Determining the egomotion of an uncalibrated camera from instantaneous optic flow. *J. Opt. Soc. Am. A*, 14(10):2670–2677, October 1997.

[HJ88] B. K. P. Horn and E. J. Weldon Jr. Direct methods for recovering motion. *Int. Journal Computer Vision*, 2(1):51–76, June 1988.

[SS97] G.P. Stein and A. Shashua. Model-based brightness constraints: On direct estimation of structure and motion. In *Proceedings CVPR'97*, pages 400–406, New York, June 1997.

[Sub88] M. Subbarao. *Interpretation of Visual Motion: A Computational Study*. Pitman, London, 1988.

[VF95] T. Vieville and O. D. Faugeras. Motion analysis with a camera with unknown and possibly varying intrinsic parameters. In *Proc. ICCV'95*, pages 750–756, New York, 1995. IEEE.

[WM97] J. Weber and J. Malik. Rigid body segmentation and shape description from dense optical flow under weak perspective. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 19(2):139–143, February 1997.