An Introduction to Projective Geometry (for computer vision)

Stan Birchfield

1 Introduction

We are all familiar with Euclidean geometry and with the fact that it describes our three-dimensional world so well. In Euclidean geometry, the sides of objects have lengths, intersecting lines determine angles between them, and two lines are said to be parallel if they lie in the same plane and never meet. Moreover, these properties do not change when the Euclidean transformations (translation and rotation) are applied. Since Euclidean geometry describes our world so well, it is at first tempting to think that it is the only type of geometry. (Indeed, the word *geometry* means "measurement of the earth.") However, when we consider the imaging process of a camera, it becomes clear that Euclidean geometry is insufficient: Lengths and angles are no longer preserved, and parallel lines may intersect.

Euclidean geometry is actually a subset of what is known as *projective geometry*. In fact, there are two geometries between them: *similarity* and *affine*. To see the relationships between these different geometries, consult Figure 1. Projective geometry models well the imaging process of a camera because it allows a much larger class of transformations than just translations and rotations, a class which includes perspective projections. Of course, the drawback is that fewer measures are preserved — certainly not lengths, angles, or parallelism. Projective transformations preserve type (that is, points remain points and lines remain lines), incidence (that is, whether a point lies on a line), and a measure known as the *cross ratio*, which will be described in section 2.4.

Projective geometry exists in any number of dimensions, just like Euclidean geometry. For example the projective line, which we denote by $\mathcal{P}^1$, is analogous to a one-dimensional Euclidean world; the projective plane, $\mathcal{P}^2$, corresponds to the Euclidean plane; and projective space, $\mathcal{P}^3$, is related to three-dimensional Euclidean space. The imaging process is a projection from $\mathcal{P}^3$ to $\mathcal{P}^2$, from three-dimensional space to the two-dimensional image plane. Because it is easier to grasp the major concepts in a lower-dimensional space, we will spend the bulk of our effort, indeed all of section 2, studying $\mathcal{P}^2$, the projective plane. That section presents many concepts which are useful in understanding the image plane and which have analogous concepts in $\mathcal{P}^3$. The final section then briefly discusses the relevance of projective geometry to computer vision, including discussions of the image formation equations and the Essential and Fundamental matrices.

	Euclidean	similarity	affine	projective
Transformations				
rotation	X	X	X	X
translation	X	X	X	X
uniform scaling		X	X	X
nonuniform scaling			X	X
shear			X	X
perspective projection				X
composition of projections				X
Invariants				
length	X			
angle	X	X		
ratio of lengths	X	X		
parallelism	X	X	X	
incidence	X	X	X	X
cross ratio	X	X	X	X

Figure 1: The four different geometries, the transformations allowed in each, and the measures that remain invariant under those transformations.

The purpose of this monograph will be to provide a readable introduction to the field of projective geometry and a handy reference for some of the more important equations. The first-time reader may find some of the examples and derivations excessively detailed, but this thoroughness should prove helpful for reading the more advanced texts, where the details are often omitted. For further reading, I suggest the excellent book by Faugeras [2] and appendix by Mundy and Zisserman [5].

2 The Projective Plane

2.1 Four models

There are four ways of thinking about the projective plane [3]. The most important of these for our purposes is homogeneous coordinates, a concept which should be familiar to anyone who has taken an introductory course in robotics or graphics. Starting with homogeneous coordinates, and proceeding to each of the other three models, we will attempt to gain intuition on the nature of the projective plane, whose concise definition will then emerge from the fourth model.

2.1.1 Homogeneous coordinates

Suppose we have a point (x, y) in the Euclidean plane. To represent this same point in the projective plane, we simply add a third coordinate of 1 at the end: $(x, y, 1)$.[1] Overall scaling is unimportant, so the point $(x, y, 1)$ is the same as the point $(\alpha x, \alpha y, \alpha)$, for any nonzero α. In other words,

$$(X, Y, W) = (\alpha X, \alpha Y, \alpha W)$$

for any $\alpha \neq 0$ (Thus the point $(0, 0, 0)$ is disallowed). Because scaling is unimportant, the coordinates (X, Y, W) are called the *homogeneous coordinates* of the point. In our discussion, we will use capital letters to denote homogeneous coordinates of points, and we will use the coordinate notation (X, Y, W) interchangeably with the vector notation $[X, Y, W]^T$.

To represent a line in the projective plane, we begin with a standard Euclidean formula for a line

$$ax + by + c = 0,$$

and use the fact that the equation is unaffected by scaling to arrive at the following:

$$\begin{aligned} aX + bY + cW &= 0 \\ \mathbf{u}^T \mathbf{p} = \mathbf{p}^T \mathbf{u} &= 0, \end{aligned} \tag{1}$$

where $\mathbf{u} = [a, b, c]^T$ is the line and $\mathbf{p} = [X, Y, W]^T$ is a point on the line. Thus we see that points and lines have the same representation in the projective plane. The parameters of a line are easily interpreted: $-a/b$ is the slope, $-c/a$ is the x-intercept, and $-c/b$ is the y-intercept.

To transform a point in the projective plane back into Euclidean coordinates, we simply divide by the third coordinate: $(x, y) = (X/W, Y/W)$. Immediately we see that the projective plane contains more points than the Euclidean plane, that is, points whose third coordinate is zero. These points are called *ideal points*, or *points at infinity*. There is a separate ideal point associated with each direction in the plane; for example, the points $(1, 0, 0)$ and $(0, 1, 0)$ are associated with the horizontal and vertical directions, respectively. Ideal points are considered just like any other point in $\mathcal{P}^2$ and are given no special treatment. All the ideal points lie on a line, called the *ideal line*, or the *line at infinity*, which, once again, is treated just the same as any other line. The ideal line is represented as $(0, 0, 1)$.

Suppose we want to find the intersection of two lines. By elementary algebra, the two lines $\mathbf{u}_1 = (a_1, b_1, c_1)$ and $\mathbf{u}_2 = (a_2, b_2, c_2)$ are found to intersect at the point $\mathbf{p} = (b_1c_2 - b_2c_1, a_2c_1 - a_1c_2, a_1b_2 - a_2b_1)$. This formula is more easily remembered as the cross product: $\mathbf{p} = \mathbf{u}_1 \times \mathbf{u}_2$. If the two lines are parallel, i.e., $-a_1/b_1 = -a_2/b_2$, the point of intersection is simply $(b_1c_2 - b_2c_1, a_2c_1 - a_1c_2, 0)$, which is the ideal point associated with the direction whose slope is $-a_1/b_1$. Similarly, given two points $\mathbf{p}_1$ and $\mathbf{p}_2$, the equation of the line passing through them is given by $\mathbf{u} = \mathbf{p}_1 \times \mathbf{p}_2$.

[1] In general, a point in an n-dimensional Euclidean space is represented as a point in an $(n+1)$-dimensional projective space.

point	$\mathbf{p} = (X, Y, W)$
incidence	$\mathbf{p}^T\mathbf{u} = 0$
collinearity	$\lvert\, \mathbf{p}_1 \quad \mathbf{p}_2 \quad \mathbf{p}_3 \,\rvert = 0$
join of 2 points	$\mathbf{u} = \mathbf{p}_1 \times \mathbf{p}_2$
ideal points	$(X, Y, 0)$

(a)

line	$\mathbf{u} = (a, b, c)$
incidence	$\mathbf{p}^T\mathbf{u} = 0$
concurrence	$\lvert\, \mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3 \,\rvert = 0$
intersection of 2 lines	$\mathbf{p} = \mathbf{u}_1 \times \mathbf{u}_2$
ideal line	$(0, 0, c)$

(b)

Figure 2: Summary of homogeneous coordinates: (a) points, and (b) lines.

Now suppose we want to determine whether three points $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$ lie on the same line. The line joining the first two points is $\mathbf{p}_1 \times \mathbf{p}_2$. The third point then lies on the line if $\mathbf{p}_3^T(\mathbf{p}_1 \times \mathbf{p}_2) = 0$, or, more succinctly, if the determinant of the 3×3 matrix containing the points is zero:

$$\det[\,\mathbf{p}_1 \quad \mathbf{p}_2 \quad \mathbf{p}_3\,] = 0.$$

Similarly, three lines $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$ intersect at the same point (i.e., they are concurrent), if the following equation holds:

$$\det[\,\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3\,] = 0.$$

The concepts of homogeneous coordinates are summarized in Figure 2. For further reading, consult the notes by Guibas [3].

Example 1. Given two lines $\mathbf{u}_1 = (4, 2, 2)$ and $\mathbf{u}_2 = (6, 5, 1)$, the point of intersection is given by:

$$\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 4 & 2 & 2 \\ 6 & 5 & 1 \end{vmatrix} = (2-10)\mathbf{i} + (12-4)\mathbf{j} + (20-12)\mathbf{k} = (-8, 8, 8) = (-1, 1, 1).$$

Example 2. Consider the intersection of the hyperbola $xy = 1$ with the horizontal line $y = 1$. To convert these equations to homogeneous coordinates, recall that $X = Wx$ and $Y = Wy$, yielding $XY = W^2$ for the hyperbola and $Y = W$ for the line. The solution to these two equations is the point (W, W, W), which is the same as the point $(1, 1)$ in the Euclidean plane, the desired result. Now let us consider the intersection of the same hyperbola with the horizontal line $y = 0$, an intersection which does not exist in the Euclidean plane. In homogeneous coordinates the line becomes $Y = 0$ which yields the solution $(X, 0, 0)$, the ideal point associated with the horizontal direction.

2.1.2 Ray space

We have just seen that, in going from Euclidean to projective, a point in $\mathcal{R}^2$ becomes a set of points in $\mathcal{R}^3$ which are related to each other by means of a nonzero scaling factor.

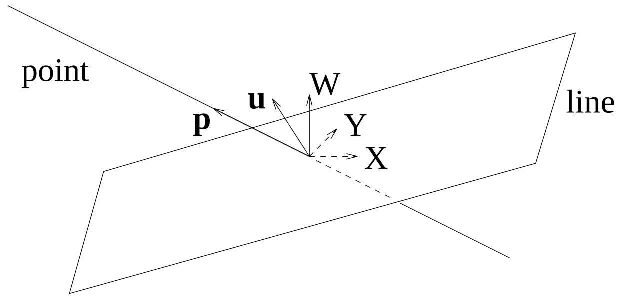


Figure 3: Ray space.

Therefore, a point $\mathbf{p} = (X, Y, W)$ in $\mathcal{P}^2$ can be visualized as a "line" [2] in three-dimensional space passing through the origin and the point $\mathbf{p}$ (Technically speaking, the line does not include the origin). This three-dimensional space is known as the *ray space* (among other names) and is shown in Figure 3. Similarly, a line $\mathbf{u} = (a, b, c)$ in $\mathcal{P}^2$ can be visualized as a "plane" passing through the origin and perpendicular to $\mathbf{u}$. The ideal line is the horizontal $W = 0$ "plane", and the ideal points are "lines" in this "plane."

2.1.3 The unit sphere

Because the coordinates are unaffected by scalar multiplication, $\mathcal{P}^2$ is two-dimensional, even though its points contain three coordinates. In fact, it is topologically equivalent to a sphere. Each point $\mathbf{p} = (X, Y, W)$, represented as a "line" in ray space, can be projected onto the unit sphere to obtain the point $\frac{1}{\sqrt{X^2+Y^2+W^2}}(X, Y, W)$ (Notice that the denominator is never zero, since the point $(0, 0, 0)$ is not allowed). Thus, points in the projective plane can be visualized as points on the unit sphere, as shown in figure 4 (Since each "line" in ray space pierces the sphere twice, both these intersections represent the same point; that is, antipodal points are identified). Similarly, the "planes" that represent lines in ray space intersect the unit sphere along great circles, so lines are visualized as great circles perpendicular to $\mathbf{u}$. The ideal line is the great circle around the horizontal midsection of the sphere, and the ideal points lie on this circle.

2.1.4 Augmented affine plane

To complete our geometrical tour of $\mathcal{P}^2$, let us project the unit sphere onto the plane $W = 1$. Each point (X, Y, W) on the sphere is thus mapped to the point $(\frac{X}{W}, \frac{Y}{W}, 1)$ which lies at the intersection of the $W = 1$ plane with the "line" representing the point. Similarly, lines are mapped to the intersection of the $W = 1$ plane with the "plane" representing the line. Ideal points and the ideal line are projected, respectively, to points at infinity and the line

[2] Since it can become confusing to read statements such as, "A point is represented as a 'line,' " we will always enclose in quotation marks the entities whose sole purpose is visualization in $n + 1$ dimensions.

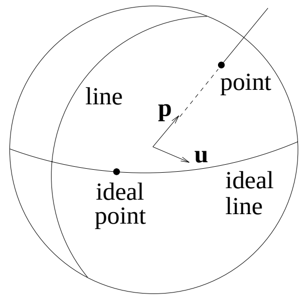


Figure 4: The unit sphere.

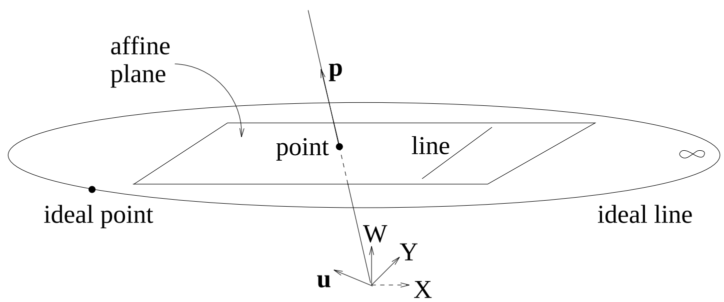


Figure 5: The affine plane plus the ideal line and ideal points.

at infinity, as shown in figure 5. Thus we have returned to a representation in which points are points and lines are lines. A concise definition of the projective plane can now be given:

Definition 1 *The projective plane, $\mathcal{P}^2$, is the affine plane augmented by a single ideal line and a set of ideal points, one for each direction, where the ideal line and ideal points are not distinguishable from regular lines and points.*

The affine plane contains the same points as the Euclidean plane. The only difference is that the former also allows for nonuniform scaling and shear.

2.2 Duality

Looking once again at figure 2, the similarities between points and lines are striking. Their representations, for example, are identical, and the formula for the intersection of two lines is the same as the formula for the connecting line between two points. These observations are not the result of coincidence but are rather a result of the duality that exists between points and lines in the projective plane. In other words, any theorem or statement that is true for the projective plane can be reworded by substituting points for lines and lines for points, and the resulting statement will be true as well.

2.3 Pencil of lines

A set of concurrent lines in $\mathcal{P}^2$, that is, a set of lines passing through the same point, is a one-dimensional projective space called a *pencil of lines*. That the space is one-dimensional is the obvious result of applying the principle of duality: a set of concurrent lines is the same as a set of collinear points. We will say no more about a pencil of lines other than to mention that it exists and that it has several applications in computer vision.

2.4 The cross ratio

As mentioned before, projective geometry preserves neither distances nor ratios of distances. However, the *cross ratio*, which is a ratio of ratios of distances, is preserved and is therefore a useful concept. Given four collinear points $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$, and $\mathbf{p}_4$ in $\mathcal{P}^2$, denote the Euclidean distance between two points $\mathbf{p}_i$ and $\mathbf{p}_j$ as Δ_{ij}. Then, one definition of the cross ratio is the following:

$$Cr(\mathbf{p}_1, \mathbf{p}_2; \mathbf{p}_3, \mathbf{p}_4) = \frac{\Delta_{13}\Delta_{24}}{\Delta_{14}\Delta_{23}}. \tag{2}$$

In other words, select one of the points, say $\mathbf{p}_1$, to be a reference point. Compute the ratio of distances from that point to two other points, say $\mathbf{p}_3$ and $\mathbf{p}_4$. Then compute the ratio of distances from the remaining point, in this case $\mathbf{p}_2$, to the same two points. The ratio of these ratios is invariant under projective transformations.

The Euclidean distance between two points $\mathbf{p}_i = [X_i, Y_i, W_i]^T$ and $\mathbf{p}_j = [X_j, Y_j, W_j]^T$ is computed from the 2D Euclidean points obtained by dividing by the third coordinate, as mentioned in section 2.1.1:

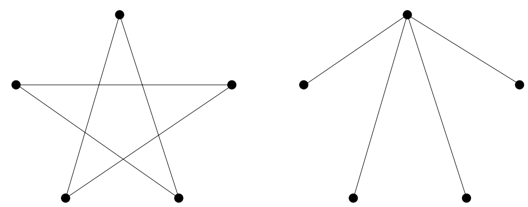
$$\Delta_{ij} = \sqrt{\left(\frac{X_i}{W_i} - \frac{X_j}{W_j}\right)^2 + \left(\frac{Y_i}{W_i} - \frac{Y_j}{W_j}\right)^2}.$$

Actually, the cross ratio is the same no matter which coordinate is used as the divisor (as long as the same coordinate is used for all the points); thus, if all the points lie on the ideal line ($W_i = 0$ for all i), then we can divide by X_i or Y_i instead. For a set of collinear points, we can always select a coordinate such that at least three of the points have nonzero entries for that coordinate. If one of the points has a zero entry, simply cancel the terms containing the point (because it lies at infinity); for example, if the second point is the culprit ($W_2 = 0$; $W_1, W_3, W_4 \neq 0$), then $\Delta_{23} = \Delta_{24} = \infty$, which cancel each other:

$$Cr(\mathbf{p}_1, \mathbf{p}_2; \mathbf{p}_3, \mathbf{p}_4) = \frac{\Delta_{13}}{\Delta_{14}}.$$

Although the cross ratio is invariant once what the order of the points has been chosen, its value is different depending on that order. Four points can be chosen $4! = 24$ ways, but in fact only six distinct values are produced, which are related by the set

$$\{\tau, \frac{1}{\tau}, 1-\tau, \frac{1}{1-\tau}, \frac{\tau-1}{\tau}, \frac{\tau}{\tau-1}\}.$$



(a) (b)

Figure 6: The cross ratio can be used with five noncollinear points.

As we hinted before, there are other measures of the cross ratio, all of which are also invariant under projective transformations. Not surprisingly, duality leads to a cross ratio for four concurrent lines by replacing the Euclidean distance between two points with the sine of the angle between two lines (I have not confirmed whether the cosine also works). Another less obvious way to measure the cross ratio between four concurrent lines is to use a new, arbitrary line that intersects them; the cross ratio of the lines is then defined as the cross ratio of the four points of intersection (The cross ratio will be the same no matter which line is used).

As a final comment on the cross ratio, it is worth noting that it does not require that the original points be collinear. For example, given five points in a star configuration, as shown in figure 6, we can connect the dots as shown in (a) to yield lines containing four collinear points, the points of intersection, whose cross ratio can be used. Another possibility is to draw lines from one of the points to the other four, as shown in (b), thus yielding four concurrent lines whose cross ratio can be used.

2.5 Conics

In Euclidean geometry, the second-order conic sections (ellipses, parabolas, and hyperbolas) are important phenomena, beyond the first-order curves such as lines and planes. Ellipses, parabolas, and hyperbolas lose their distinction in projective geometry because they are all projectively equivalent, that is, any form can be projected into any other form. Collectively, these curves are referred to as conics, with no distinction between the different forms.

Just as a circle in Euclidean geometry is defined as a locus of points with a constant distance from the center, so a conic in projective geometry is defined as a locus of points with a constant cross ratio to four fixed points, no three of which are collinear. Note that in both cases the shape of the curve is defined with respect to an invariant of the particular geometry, distance in the case of Euclidean, and cross ratio in the case of projective.

The equation of a conic is given by:

$$\mathbf{p}^T C \mathbf{p} = 0,$$

or

$$c_{11}X^2 + c_{22}Y^2 + c_{33}W^2 + 2c_{12}XY + 2c_{13}XW + 2c_{23}YW = 0,$$

where $\mathbf{p}$ is a 3×1 vector and C is a symmetric 3×3 matrix.

Since points and lines are dual concepts, it is not surprising that a conic is a self-dual figure. That is, it can be considered as a locus of points (as we have just done), or it can be considered as an envelope of tangent lines (the set of lines that are tangent to the conic). The equation for the envelope of lines is $\mathbf{u}^T|C|C^{-1}\mathbf{u}$.

2.6 Absolute points

A surprising property of conics is that every circle intersects the ideal line, $W = 0$, at two fixed points. To see this, note that a circle is a conic with all off-diagonal elements (c_{12}, c_{13}, and c_{23}) set to zero and all diagonal elements equal:

$$X^2 + Y^2 + W^2 = 0,$$

which therefore intersects the ideal line $W = 0$ at

$$X^2 + Y^2 = 0.$$

This equation has two complex roots, known as the *absolute points*: $\mathbf{i} = (1, i, 0)$ and $\mathbf{j} = (1, -i, 0)$. (Although we have, for simplicity, assumed that homogeneous coordinates are real, they can in general be the elements of any commutative field in which $1 + 1 \neq 0$ [1, p. 112].) It will be shown in the next two subsections that the absolute points remain invariant under similarity transformations, which makes them useful for determining the angle between two lines.

2.7 Collineations

A *collineation* of $\mathcal{P}^2$ is defined as a mapping from the plane to itself such that the collinearity of any set of points is preserved. Such a mapping can be achieved with matrix multiplication by a 3×3 matrix T. Each point $\mathbf{p}$ is transformed into a point $\mathbf{p}'$:

$$\mathbf{p}' = T\mathbf{p}.$$

We will use the terms *transformation* and collineation interchangeably. Since scaling is unimportant, only eight elements of T are independent. Therefore, since each point contains two independent values, four pairs of corresponding points are necessary to determine T.

To transform a line $\mathbf{u}$ into a line $\mathbf{u}'$, we note that collinearity must be preserved, that is, if a point $\mathbf{p}$ lies on the line $\mathbf{u}$, then the corresponding point $\mathbf{p}'$ must lie on the corresponding line $\mathbf{u}'$. Therefore,

$$\mathbf{p}^T\mathbf{u} = 0 = (T^{-1}\mathbf{p}')^T\mathbf{u} = (\mathbf{p}')^T(T^{-T}\mathbf{u}),$$

which indicates that

$$\mathbf{u}' = T^{-T}\mathbf{u}.$$

From these results, it is not hard to show that a point conic C transforms to $T^{-T}CT^{-1}$, and a line conic $|C|C^{-1}$ transforms to $T|C|C^{-1}T^T$.

Regarding transformations, recall that *projective* $\supset$ *affine* $\supset$ *similarity* $\supset$ *Euclidean.* Let's study the matrix T to uncover the relationships between these various geometries. First we will write out the elements of T, for reference:

$$T_{projective} = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix}.$$

The affine plane is just the projective plane minus the ideal line. Therefore, affine transformations must preserve the ideal line and the ideal points, that is, any point $[X, Y, 0]^T$ must be transformed into $[\alpha X, \alpha Y, 0]^T$ for some arbitrary scaling α:

$$\alpha \begin{bmatrix} X \\ Y \\ 0 \end{bmatrix} = T \begin{bmatrix} X \\ Y \\ 0 \end{bmatrix},$$

which implies that $t_{31} = t_{32} = 0$. The matrix for affine transformation, then, is

$$T_{affine} = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ 0 & 0 & t_{33} \end{bmatrix},$$

where once again only six of these parameters are independent, since scale is unimportant.

Unlike affine transformations, similarity transformations preserve angles and ratios of lengths. Delaying the derivation for a moment, we simply state the result:

$$T_{similarity} = \begin{bmatrix} \cos\theta & \sin\theta & t_{13} \\ -\sin\theta & \cos\theta & t_{23} \\ 0 & 0 & t_{33} \end{bmatrix}, \tag{3}$$

where θ is an arbitrary angle.

Under Euclidean transformation, scale is important, and therefore the point $\mathbf{p}$ must first be converted to Euclidean coordinates by dividing by its third element. The transformation then is

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}.$$

In closing this section, we offer one final proposition, along with its proof:

Proposition 1 *A transformation is a similarity transformation if and only if it preserves the absolute points,* $[1, \pm i, 0]$.

The "only if" is rather easy to see: The absolute point $[1, \pm i, 0]^T$ is transformed through equation (3) to the point $e^{\pm i\theta}[1, \pm i, 0]^T$, which is equivalent because the scale factor is ignored.

The "if" is a little more complicated, but still rather straightforward. Starting with the unrestricted equation for T,

$$T = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix},$$

the fact that $[1, i, 0]^T$ is preserved yields the following two equations:

$$\begin{aligned} \frac{t_{11} + it_{12}}{t_{21} + it_{22}} &= \frac{1}{i} \\ t_{31} + it_{32} &= 0. \end{aligned}$$

Since the elements of T are constrained to be real, this leads to the following three constraints on the elements of T:

$$\begin{aligned} t_{11} &= t_{22} \\ t_{12} &= -t_{21} \\ t_{31} &= t_{32} = 0. \end{aligned}$$

So the matrix of T looks like this:

$$T = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ -t_{12} & t_{11} & t_{23} \\ 0 & 0 & t_{33} \end{bmatrix}.$$

Given two arbitrary numbers t_{11} and t_{12}, we can always reparameterize them as $t_{11} = k \cos\theta$ and $t_{12} = k \sin\theta$, where θ is an angle and k is a scalar. Multiplying the previous equation by $1/k$ (which is legal because we are working in homogeneous coordinates), we then get

$$T = \begin{bmatrix} \cos\theta & \sin\theta & t_{13}/k \\ -\sin\theta & \cos\theta & t_{23}/k \\ 0 & 0 & t_{33}/k \end{bmatrix},$$

which is seen to be the equation of a similarity transformation when compared with equation (3). (NOTE: Using the point $[1, -i, 0]^T$ yields the same result.)

2.8 Laguerre formula

Absolute points have a surprising but important application: they can be used to determine the angle between two lines. To see how this works, let us assume that we have two lines $\mathbf{u}_1$ and $\mathbf{u}_2$ which intersect the ideal line at two points, say $\mathbf{p}_1$ and $\mathbf{p}_2$. Then, the cross ratio between these two points and the two absolute points $\mathbf{i}$ and $\mathbf{j}$ yields the directed angle θ from the second line to the first:

$$\theta = \frac{1}{2i} \log Cr(\mathbf{p}_1, \mathbf{p}_2; \mathbf{i}, \mathbf{j}),$$

which is known as the *Laguerre formula.*

To gain some intuition on why this formula is true, let us consider a simple example. Suppose we have two lines

$$\begin{aligned} a_1 x - y &= 0 \\ a_2 x - y &= 0 \end{aligned}$$

in the affine plane. It is clear that these two lines can be represented as two vectors $\mathbf{v}_1 = [1, a_1]^T$ and $\mathbf{v}_2 = [1, a_2]^T$ in the Euclidean plane. The directed angle between the two lines is the directed angle between the two vectors and is given by:

$$\tan\theta = \frac{\mathbf{v}_2 \times \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_2} = \frac{a_1 - a_2}{1 + a_1 a_2}.$$

Now in the projective plane these lines are represented as $[a_1, -1, 0]^T$ and $[a_2, -1, 0]^T$, which are found by mapping points $[x, y]^T$ in the affine plane to points $[x, y, 1]^T$ in the projective plane. The ideal line passing through $\mathbf{i} = [1, i, 0]^T$ and $\mathbf{j} = [1, -i, 0]^T$ is given by $\mathbf{i} \times \mathbf{j} = [0, 0, 1]^T$. The two points of intersection between this line and the two original lines are given by $[1, a_1, 0]^T$ and $[1, a_2, 0]^T$. The cross ratio of the four points is then given by:

$$\begin{aligned} Cr(\mathbf{p}_1, \mathbf{p}_2; \mathbf{i}, \mathbf{j}) &= \frac{a_1 - i}{a_1 + i} \cdot \frac{a_2 + i}{a_2 - i} \\ &= \frac{1 + a_1 a_2 + i(a_1 - a_2)}{1 + a_1 a_2 + i(a_2 - a_1)}. \end{aligned}$$

Converting the complex numbers from rectangular to polar coordinates yields:

$$\begin{aligned} &= \frac{e^{i \tan^{-1} \frac{(a_1 - a_2)}{1 + a_1 a_2}}}{e^{i \tan^{-1} \frac{(a_2 - a_1)}{1 + a_1 a_2}}} \\ &= e^{2i \tan^{-1} \frac{(a_1 - a_2)}{1 + a_1 a_2}}, \end{aligned}$$

from which it follows that

$$\frac{1}{2i} \log Cr(\mathbf{p}_1, \mathbf{p}_2; \mathbf{i}, \mathbf{j}) = \tan^{-1} \frac{a_1 - a_2}{1 + a_1 a_2} = \tan^{-1} \frac{\mathbf{v}_2 \times \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_2},$$

which is the desired result.

3 Projective Space

All of the concepts that we have discussed for the projective plane, $\mathcal{P}^2$, have analogies in projective space, $\mathcal{P}^3$. For example, there is a duality between points and planes, lines are self-dual, a pencil of planes is a two-dimensional projective space,[3] the cross ratio

[3] Do not be confused by the term *space*, which can refer to either three dimensions or an arbitrary number of dimensions.

between planes is invariant, quadrics play the same role as conics, the absolute conic remains invariant under similarity transformations, and the Laguerre formula can be used to find the angle between two projection rays. For more details, see [2].

A point in $\mathcal{P}^3$ is represented by a 4-tuple $\mathbf{p} = (X, Y, Z, W)$, and similarly for a plane $\mathbf{n}$. Not surprisingly, a point lies in a plane if and only if $\mathbf{p}^T\mathbf{n} = 0$. Slightly more difficult are the tasks of finding the plane which passes through three given points or of finding the intersections of planes. To answer these questions, we must first define a representation for lines.

3.1 Representing lines: The Plücker relations

Recall from the Section 2 that the coordinates of the line passing through two points $\mathbf{p}_1 = (X_1, Y_1, W_1)$ and $\mathbf{p}_2 = (X_2, Y_2, W_2)$ is given by

$$\mathbf{u} = (Y_1W_2 - W_1Y_2, W_1X_2 - X_1W_2, X_1Y_2 - Y_1X_2).$$

Notice that these three coordinates are just the determinants of the three 2×2 submatrices of the following matrix:

$$[\,\mathbf{p}_1 \quad \mathbf{p}_2\,] = \begin{bmatrix} X_1 & X_2 \\ Y_1 & Y_2 \\ W_1 & W_2 \end{bmatrix},$$

taken in the appropriate order and given the appropriate sign.

The procedure is similar in $\mathcal{P}^3$. The coordinates of the line $\mathbf{u}$ passing through two points $\mathbf{p}_1 = (X_1, Y_1, Z_1, W_1)$ and $\mathbf{p}_2 = (X_2, Y_2, Z_2, W_2)$ is given by the determinants of the six 3×3 submatrices of the following matrix:

$$[\,\mathbf{p}_1 \quad \mathbf{p}_2\,] = \begin{bmatrix} X_1 & X_2 \\ Y_1 & Y_2 \\ Z_1 & Z_2 \\ W_1 & W_2 \end{bmatrix}.$$

In other words, $\mathbf{u} = (l_{41}, l_{42}, l_{43}, l_{23}, l_{31}, l_{12})$, where

$$\begin{aligned} l_{41} &= W_1X_2 - X_1W_2 \\ l_{42} &= W_1Y_2 - Y_1W_2 \\ l_{43} &= W_1Z_2 - Z_1W_2 \\ l_{23} &= Y_1Z_2 - Z_1Y_2 \\ l_{31} &= Z_1X_2 - X_1Z_2 \\ l_{12} &= X_1Y_2 - Y_1X_2. \end{aligned}$$

These coordinates l_{ij} are called the *Plücker coordinates* of the line. It is fairly easy to show that, if the points $\mathbf{p}_1$ and $\mathbf{p}_2$ are not ideal (that is, W_1 and W_2 are not zero), then the coordinates have a nice Euclidean interpretation:

$$\begin{aligned} (l_{41}, l_{42}, l_{43}) &= \bar{\mathbf{p}}_2 - \bar{\mathbf{p}}_1 \\ (l_{23}, l_{31}, l_{12}) &= \bar{\mathbf{p}}_1 \times \bar{\mathbf{p}}_2, \end{aligned}$$

where $\bar{\mathbf{p}}_i = \frac{1}{W_i}(X_i, Y_i, Z_i)$, $i = 1, 2$, are the coordinates of the corresponding Euclidean points. That is, the first three Plücker coordinates describe the direction of the line, and the last three coordinates describe the plane containing the line and the origin and the distance from the origin to the line. Therefore the six Plücker coordinates are sufficient to describe the line. The coordinates are not independent, however, because they always satisfy

$$l_{41}l_{23} + l_{42}l_{31} + l_{43}l_{12} = 0,$$

which can be derived by noting that the 4×4 determinant $|\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_1, \mathbf{p}_2|$ is identically zero.

Where does the magic number six come from? That is, why do we need six parameters to represent a line in $\mathcal{P}^3$? Interestingly, it turns out that it takes $\binom{n+1}{k}$ parameters to represent an entity defined by k points in a space requiring $n+1$ parameters for each point (To see this, count the number of submatrices in the matrix above). For example, in $\mathcal{P}^2$ a point requires $\binom{3}{1} = 3$ parameters, and a line (which is defined by two points) also requires $\binom{3}{2} = 3$ parameters. In $\mathcal{P}^3$, a point requires $\binom{4}{1} = 4$ parameters, a line $\binom{4}{2} = 6$ parameters, and a plane $\binom{4}{3} = 4$ parameters.

3.2 Intersections and unions of points, lines, and planes

Now that we have a representation for lines, we can proceed to more complex relations. It should not surprise the reader that the coordinates of the plane passing through the three points $\mathbf{p}_i = (X_i, Y_i, Z_i, W_i)$, $i = 1, 2, 3$, is obtained from the determinants of the four 3×3 submatrices of

$$\begin{bmatrix} X_1 & X_2 & X_3 \\ Y_1 & Y_2 & Y_3 \\ Z_1 & Z_2 & Z_3 \\ W_1 & W_2 & W_3 \end{bmatrix}.$$

Paying careful attention to the order of the submatrices then yields the plane's coordinates:

$$\mathbf{n} = \left(\begin{vmatrix} Y_1 & Y_2 & Y_3 \\ W_1 & W_2 & W_3 \\ Z_1 & Z_2 & Z_3 \end{vmatrix}, \begin{vmatrix} X_1 & X_2 & X_3 \\ Z_1 & Z_2 & Z_3 \\ W_1 & W_2 & W_3 \end{vmatrix}, \begin{vmatrix} X_1 & X_2 & X_3 \\ W_1 & W_2 & W_3 \\ Y_1 & Y_2 & Y_3 \end{vmatrix}, \begin{vmatrix} X_1 & X_2 & X_3 \\ Y_1 & Y_2 & Y_3 \\ Z_1 & Z_2 & Z_3 \end{vmatrix} \right).$$

Note that, as a result of duality, the same formula can be used to find the point defining the intersection of three planes.

A point $\mathbf{p} = (X, Y, Z, W)$ lies on a line $\mathbf{u} = (l_{41}, l_{42}, l_{43}, l_{23}, l_{31}, l_{12})$ if and only if the vectors $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}$ are collinear, that is, the determinants of the four 3×3 submatrices of the following matrix

$$\begin{bmatrix} X_1 & X_2 & X \\ Y_1 & Y_2 & Y \\ Z_1 & Z_2 & Z \\ W_1 & W_2 & W \end{bmatrix}$$

are zero. In terms of the Plücker coordinates, this concurrence can be expressed as follows:

$$A\mathbf{p} = \mathbf{0},$$

where

$$A = \begin{bmatrix} l_{23} & l_{31} & l_{12} & 0 \\ 0 & -l_{43} & l_{42} & l_{23} \\ -l_{43} & 0 & l_{41} & -l_{31} \\ -l_{42} & l_{41} & 0 & l_{12} \end{bmatrix}.$$

An intuitive way to think about A is to realize that a line can be defined as the intersection of two planes. Therefore, a point lies on the line if it lies in the two planes. The equation above says that a point lies on the line if it lies in four planes. Only two of A's rows are important for any given line (indeed, A is of rank two), but all four rows are necessary to ensure that degenerate cases are handled properly.

Two lines $\mathbf{u}$ and $\mathbf{u}'$ intersect if and only if their Plücker coordinates satisfy the equation

$$(l_{41}l'_{23} + l'_{41}l_{23}) + (l_{42}l'_{31} + l'_{42}l_{31}) + (l_{43}l'_{12} + l'_{43}l_{12}) = 0,$$

which arises from the fact that the 4×4 determinant $|\,\mathbf{p}_1 \quad \mathbf{p}_2 \quad \mathbf{p}'_1 \quad \mathbf{p}'_2\,|$ is zero, where $\mathbf{p}_1$ and $\mathbf{p}_2$ lie on $\mathbf{u}$ and $\mathbf{p}'_1$ and $\mathbf{p}'_2$ lie on $\mathbf{u}'$.

In this section we have shown how to find the line containing two points and the plane containing three points. Symmetrically, we have shown how to find the point at the intersection of three planes but have not shown how to find the line at the intersection of two planes. This latter problem can be solved by computing the determinants of the submatrices, but it appears according to my calculations that a different choice of submatrices will have to be made. We have shown how to determine whether a point lies on a line or in a plane, and we have shown whether two lines intersect, though we have not calculated their intersection point. Other remaining problems include finding the intersection of a line and a plane, calculating the plane defined by two lines, and computing the plane defined by a point and a line.

4 Projective Geometry Applied to Computer Vision

Projective geometry is a mathematical framework in which to view computer vision in general, and especially image formation in particular. The main areas of application are those in which image formation and/or invariant descriptions between images are important, such as camera calibration, stereo, object recognition, scene reconstruction, mosaicing, image synthesis, and the analysis of shadows. This latter application arises from the fact that the composition of two perspective projections is not necessarily a perspective projection but is definitely a projective transformation; that is, projective transformations form a group, whereas perspective projections do not. Many areas of computer vision have little to do with projective geometry, such as texture analysis, color segmentation, and edge detection. And even in a field such as motion analysis, projective geometry offers little help when the rigidity assumption is lost because the relationship between projection rays in successive images cannot be described by such simple and elegant mathematics.

The following three sections contain the image formation equations, detailed derivations of the Essential and Fundamental matrices, and an interesting discussion of the interpretation of vanishing points.

4.1 Image formation

Image formation involves the projection of points in $\mathcal{P}^3$ (the world) to points in $\mathcal{P}^2$ (the image plane). The perspective projection equations with which we are familiar,

$$x = -f\frac{X}{Z}$$
$$y = -f\frac{Y}{Z}$$

where the point (X, Y, Z) in the world is projected to the point (x, y) on the image plane, are inherently nonlinear. Converting to homogeneous coordinates, however, makes them linear:

$$\mathbf{p}' = T_{perspective}\mathbf{p},$$

where $\mathbf{p}' = [x, y, w]^T$, $\mathbf{p} = [X, Y, Z, W]^T$, and the perspective projection matrix T is given by:

$$T_{perspective} = \begin{bmatrix} -f & 0 & 0 & 0 \\ 0 & -f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

The entire image formation process includes perspective projection, along with matrices for internal and external calibration:

$$\begin{aligned} \tilde{P} = T_{internal}T_{perspective}T_{external} &= \begin{bmatrix} k_u & k_c & u_0 \\ 0 & k_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -f & 0 & 0 \\ 0 & -f & 0 \\ 0 & 0 & 1 \end{bmatrix} [\, R \quad \mathbf{t} \,] \\ &= \begin{bmatrix} \alpha_u & -\alpha_u \cot\theta & u_0 \\ 0 & \alpha_v / \sin\theta & v_0 \\ 0 & 0 & 1 \end{bmatrix} [\, R \quad \mathbf{t} \,] \\ &= AD, \end{aligned} \tag{4}$$

where α_u and α_v are the scale factors of the image plane (in units of the focal length f), θ is the skew ($\theta = \pi/2$ for most real cameras), the point (u_0, v_0) is the principal point, R is the 3×3 rotation matrix, and $\mathbf{t}$ is the 3×1 translation vector. The matrix A contains the internal parameters and perspective projection, while D contains the external parameters.

It is sometimes convenient to decompose the 3×4 projection matrix $\tilde{P}$ into a 3×3 matrix P and a 3×1 vector p

$$\tilde{P} = [\, P \quad p \,]$$

so that

$$P = AR \;\text{ and }\; p = A\mathbf{t}. \tag{5}$$

4.2 Essential and fundamental matrices

Suppose we have a stereo pair of cameras viewing a point $\mathbf{M}$ in the world which projects onto the two image planes at $\bar{\mathbf{m}}_1$ and $\bar{\mathbf{m}}_2$ (Since we are dealing with homogeneous coordinates, $\mathbf{M}$ is 4×1, and $\bar{\mathbf{m}}_1$ and $\bar{\mathbf{m}}_2$ are each 3×1). If we assume the cameras are calibrated, then $\bar{\mathbf{m}}_1$ and $\bar{\mathbf{m}}_2$ are given in *normalized coordinates*, that is, each is given with respect to its camera's coordinate frame. The epipolar constraint says that the vector from the first camera's optical center to the first imaged point, the vector from the second optical center to the second imaged point, and the vector from one optical center to the other are all coplanar. In normalized coordinates, this constraint can be expressed simply as

$$\bar{\mathbf{m}}_2^T(\mathbf{t} \times R\bar{\mathbf{m}}_1) = 0,$$

where R and $\mathbf{t}$ capture the rotation and translation between the two cameras' coordinate frames. The multiplication by R is necessary to transform $\bar{\mathbf{m}}_1$ into the second camera's coordinate frame. By defining $[\mathbf{t}]_x$ as the matrix such that $[\mathbf{t}]_x\mathbf{y} = \mathbf{t} \times \mathbf{y}$ for any vector $\mathbf{y}$,[4] we can rewrite the equation as a linear equation:

$$\bar{\mathbf{m}}_2^T([\mathbf{t}]_x R\bar{\mathbf{m}}_1) = \bar{\mathbf{m}}_2^T E\bar{\mathbf{m}}_1 = 0,$$

where $E = [\mathbf{t}]_x R$ is called the *Essential matrix* and has been studied extensively over the last two decades.

Now suppose the cameras are uncalibrated. Then the matrices A_1 and A_2 (from (4)) containing the internal parameters of the two cameras are needed to transform the normalized coordinates into pixel coordinates:

$$\begin{aligned} \mathbf{m}_1 &= A_1\bar{\mathbf{m}}_1 \\ \mathbf{m}_2 &= A_2\bar{\mathbf{m}}_2. \end{aligned}$$

This yields the following equation:

$$\begin{aligned} (A_2^{-1}\mathbf{m}_2)^T(\mathbf{t} \times RA_1^{-1}\mathbf{m}_1) &= 0 \\ \mathbf{m}_2^T A_2^{-T}(\mathbf{t} \times RA_1^{-1}\mathbf{m}_1) &= 0 \qquad (6) \\ \mathbf{m}_2^T F\mathbf{m}_1 &= 0, \qquad (7) \end{aligned}$$

where $F = A_2^{-T} E A_1^{-1}$ is the more recently discovered *Fundamental matrix*.

Thus both the Essential and Fundamental matrices completely describe the geometric relationship between corresponding points of a stereo pair of cameras. The only difference between the two is that the former deals with calibrated cameras, while the latter deals with uncalibrated cameras. The Essential matrix contains five parameters (three for rotation and two for the direction of translation — the magnitude of translation cannot be recovered due

[4] If $\mathbf{t} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$, then $[\mathbf{t}]_x = \begin{bmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{bmatrix}$.

to the depth/speed ambiguity) and has two constraints: (1) its determinant is zero, and (2) its two non-zero singular values are equal. The Fundamental matrix contains seven parameters (two for each of the epipoles and three for the homography between the two pencils of epipolar lines) and its rank is always two [4].

There are several other ways to derive the Essential and Fundamental Matrices, each of which presents a little more insight into their nature. In the next few subsections, we will look at these methods and then summarize our findings.

4.2.1 Alternate derivation: algebraic

To describe the relationship between R, $\mathbf{t}$, A_1, and A_2 more exactly, and to connect the above equations with those found in [6], we offer the following algebraic derivation.

Recall that a point $\mathbf{M}$ produces an image $\mathbf{m}$ through the equation $\mathbf{m} = \tilde{P}\mathbf{M}$. Without loss of generality, we can assume that $\mathbf{M}$ is given with respect to the first camera's coordinate frame to yield the following two imaging equations:

$$\begin{aligned} \lambda_1 \mathbf{m}_1 &= A_1 [\, \mathbf{I} \quad \mathbf{0} \,] \mathbf{M} \\ \lambda_2 \mathbf{m}_2 &= A_2 [\, R \quad \mathbf{t} \,] \mathbf{M}, \end{aligned}$$

where λ_1 and λ_2 are scale factors, $\mathbf{I}$ is the 3×3 identity matrix and $\mathbf{0}$ is the 3×1 null vector. By letting $\mathbf{M} = [\, \hat{\mathbf{M}}^T \quad 1 \,]^T$ ($\hat{\mathbf{M}}$ is 3×1), we achieve the following relation:

$$\begin{aligned} \lambda_2 \mathbf{m}_2 &= A_2 [\, R \quad \mathbf{t} \,] \mathbf{M} \\ &= A_2 (R\hat{\mathbf{M}} + \mathbf{t}) \\ &= \lambda_1 A_2 R A_1^{-1} \mathbf{m}_1 + A_2 \mathbf{t} && (8) \\ \lambda_2 A_2^{-1} \mathbf{m}_2 &= \lambda_1 R A_1^{-1} \mathbf{m}_1 + \mathbf{t}. && (9) \end{aligned}$$

Geometrically, this equation says that the vector on the left is a linear combination of the two vectors on the right. Therefore, they are all coplanar, and the vector $\mathbf{v} = \mathbf{t} \times RA_1^{-1}\mathbf{m}_1$ is perpendicular to that plane:

$$\begin{aligned} \lambda_2 (A_2^{-1} \mathbf{m}_2)^T \mathbf{v} &= \lambda_1 (R A_1^{-1} \mathbf{m}_1)^T \mathbf{v} + \mathbf{t}^T \mathbf{v} \\ &= 0 \\ \mathbf{m}_2^T (A_2^{-T} (\mathbf{t} \times R) A_1^{-1}) \mathbf{m}_1 &= 0, \end{aligned}$$

which is identical to (6).

Similarly, the vector $\mathbf{w} = A_2\mathbf{t} \times A_2 R A_1^{-1}\mathbf{m}_1$ is perpendicular to the vectors in (8):

$$\begin{aligned} \lambda_2 \mathbf{m}_2^T \mathbf{w} &= \lambda_1 (A_2 R A_1^{-1} \mathbf{m}_1)^T \mathbf{w} + (A_2 \mathbf{t})^T \mathbf{w} \\ &= 0 \\ \mathbf{m}_2^T (A_2 \mathbf{t} \times A_2 R A_1^{-1}) \mathbf{m}_1 &= 0. \end{aligned}$$

This is a surprising result because it gives us a new and equivalent expression for F:

$$F = [A_2 \mathbf{t}]_x A_2 R A_1^{-1}, \tag{10}$$

which shows that F can be written as the product of an anti-symmetric matrix $[A_2\mathbf{t}]_x$ and an invertible matrix $A_2RA_1^{-1}$ [4].

4.2.2 Alternate derivation: from the epipolar line

Faugeras [2] approaches the problem from a slightly different direction by using the fact that the point $\mathbf{m}_2$ must lie on the epipolar line corresponding to $\mathbf{m}_1$:

$$\mathbf{m}_2^T\mathbf{l} = 0. \tag{11}$$

That line contains two points, the epipole $\mathbf{e}$ (the projection of the first camera's optical center into the second camera) and the point at infinity associated with $\mathbf{m}_1$:

$$\mathbf{l} = \mathbf{e} \times \mathbf{m}_\infty.$$

In [2, pp. 40-41] it is shown that the epipole is given by

$$\mathbf{e} = \tilde{P}_2 \begin{bmatrix} \mathbf{C} \\ 1 \end{bmatrix} = \tilde{P}_2 \begin{bmatrix} -P_1^{-1}p_1 \\ 1 \end{bmatrix},$$

and the point at infinity by

$$\mathbf{m}_\infty = P_2P_1^{-1}\mathbf{m}_1.$$

Therefore, the epipolar line is:

$$\begin{aligned} \mathbf{l} &= \mathbf{e} \times \mathbf{m}_\infty \\ &= \tilde{P}_2 \begin{bmatrix} -P_1^{-1}p_1 \\ 1 \end{bmatrix} \times P_2P_1^{-1}\mathbf{m}_1 \\ &= [\, A_2R \quad A_2\mathbf{t}\,] \begin{bmatrix} 0 \\ 1 \end{bmatrix} \times (A_2R)A_1^{-1}\mathbf{m}_1 \\ &= A_2\mathbf{t} \times A_2RA_1^{-1} \\ &= A_2\mathbf{t} \times (A_2RA_1^{-1})\mathbf{m}_1, \end{aligned}$$

where we have used the substitutions in (5). Combining with (11), we get the desired result:

$$F = [A_2\mathbf{t}]_x A_2RA_1^{-1}.$$

4.2.3 Summary

For reference, we now summarize the equation for the Essential matrix and the two equations for the Fundamental matrix:

$$\boxed{\begin{aligned} E &= [\mathbf{t}]_x R \\ F &= A_2^{-T}EA_1^{-1} \\ &= [A_2\mathbf{t}]_x A_2RA_1^{-1}. \end{aligned}}$$

4.3 Vanishing points

Anyone who has taken a course in perspective drawing is familiar with the notion that lines on the paper which represent parallel lines in the world intersect on the paper at a point known as the vanishing point. Each set of lines has a different vanishing point. But, just what is a vanishing point? Projective geometry sheds light on this issue.

Because image formation is the projection from a 3D world to a 2D surface, each point on the image plane is the projection of an infinite number of points in the world. Usually, the closest point is opaque and therefore we think of the point on the image plane as being the projection of only one point in the world. However, ideal points in the world (i.e., ideal points in $\mathcal{P}^3$), always project onto the image plane regardless of the opacity of other points. Each point on the image plane is the projection of an ideal point. To see this, consider the following perspective projection equation (a general projective transformation is used for simplicity) from an ideal point $[X, Y, Z, 0]^T$ in the world to a point $[x, y, w]^T$ in the image plane:

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & t_{13} & t_{14} \\ t_{21} & t_{22} & t_{23} & t_{24} \\ t_{31} & t_{32} & t_{33} & t_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 0 \end{bmatrix}.$$

Because the matrix is full rank, each ideal point projects to a different point on the image plane. Since parallel lines in 3D space intersect at an ideal point in $\mathcal{P}^3$, their projections in the image plane must still intersect at a point. But now, through the projection matrix, the ideal point has become a "real" point, in the sense that it is no longer ideal. However, some ideal points do not become "real" points. If the ideal point represents a direction that is parallel to the image plane, then the dot product of the ideal point with the third row of the matrix (which is the z axis of the plane) is zero, and the projected point is still ideal. Just as the ideal points of $\mathcal{P}^2$ have a one-to-one correspondence with all the points in $\mathcal{P}^1$, so the ideal points of $\mathcal{P}^3$ have a one-to-one correspondence with all the points in $\mathcal{P}^2$. Thus, we see that the fact that parallel lines in the world intersect when drawn on a piece of paper follows naturally from projective geometry.

A Demonstration of Cross Ratio in $\mathcal{P}^1$

Let $\mathbf{p}_i = (X_i, 1), i = 1, \ldots, 4$ be four points on the projective line. (In this demonstration, we will only consider finite points, although the cross ratio holds for infinite points as well.) Define the distance Δ_{ij} between two points i and j as $\Delta_{ij} = |X_i - X_j|$. What we want to show is that the cross ratio

$$Cr(\mathbf{p}_1, \mathbf{p}_2; \mathbf{p}_3, \mathbf{p}_4) = \frac{\Delta_{13}\Delta_{24}}{\Delta_{14}\Delta_{23}}$$

is preserved under projective projection of the points.

A point $\mathbf{p}_i$ is projected through a 2×2 transformation matrix T to a new point $\mathbf{p}_i' = (t_{11}X_i + t_{12}, t_{21}X_i + t_{22})$. Therefore, the new coordinate $\mathbf{p}_i' = (X_i', 1)$ is defined by:

$$X_i' = \frac{t_{11}X_i + t_{12}}{t_{21}X_i + t_{22}}.$$

Then, the distance between two points X_i' and X_j' is

$$\begin{aligned}
\Delta_{ij}' &= |X_i' - X_j'| \\
&= \left| \frac{t_{11}X_i + t_{12}}{t_{21}X_i + t_{22}} - \frac{t_{11}X_j + t_{12}}{t_{21}X_j + t_{22}} \right| \\
&= \left| \frac{\det(T)(X_i - X_j)}{(t_{21}X_i + t_{22})(t_{21}X_j + t_{22})} \right|,
\end{aligned}$$

where $\det(T) = t_{11}t_{22} - t_{12}t_{21}$. The ratio between two distances, one from a point X_i' to another point X_j', and another from the same point X_i' to a third point X_k', is

$$\begin{aligned}
\frac{\Delta_{ij}'}{\Delta_{ik}'} &= \frac{|X_i' - X_j'|}{|X_i' - X_k'|} \\
&= \left| \frac{\det(T)(X_i - X_j)}{(t_{21}X_i + t_{22})(t_{21}X_j + t_{22})} \cdot \frac{(t_{21}X_i + t_{22})(t_{21}X_k + t_{22})}{\det(T)(X_i - X_k)} \right| \\
&= \left| \frac{X_i - X_j}{X_i - X_k} \cdot \frac{t_{21}X_k + t_{22}}{t_{21}X_j + t_{22}} \right|,
\end{aligned}$$

which is the original ratio Δ_{ij}/Δ_{ik}, multiplied by a constant that is dependent only upon the coordinates X_j and X_k. A similar ratio Δ_{lj}/Δ_{lk}, taken with respect to another point X_l, has this same constant, and therefore dividing the two ratios causes the constants to cancel:

$$\begin{aligned}
Cr(\mathbf{p}_1, \mathbf{p}_2; \mathbf{p}_3, \mathbf{p}_4) &= \frac{\Delta_{13}'\Delta_{24}'}{\Delta_{14}'\Delta_{23}'} \\
&= \left| \frac{X_1 - X_3}{X_1 - X_4} \cdot \frac{t_{21}X_4 + t_{22}}{t_{21}X_3 + t_{22}} \cdot \frac{X_2 - X_4}{X_2 - X_3} \cdot \frac{t_{21}X_3 + t_{22}}{t_{21}X_4 + t_{22}} \right| \\
&= \left| \frac{X_1 - X_3}{X_1 - X_4} \cdot \frac{X_2 - X_4}{X_2 - X_3} \cdot \right| \\
&= \frac{\Delta_{13}\Delta_{24}}{\Delta_{14}\Delta_{23}},
\end{aligned}$$

showing that the cross ratio is unaffected by projection.

Suppose that one of the points, say $\mathbf{p}_1$, is at infinity (i.e., its second coordinate is zero). Then, dividing by the second coordinate (which is what we normally do to transform the point into the required form) yields $X_1 = \infty$. Substituting into the above formula yields:

$$Cr(\mathbf{p}_1, \mathbf{p}_2; \mathbf{p}_3, \mathbf{p}_4) = \frac{\Delta_{24}}{\Delta_{23}},$$

since the terms with ∞ cancel each other. (Technically speaking, we must take the limit of the equation as X_1 tends to ∞, but the result is the same.) Similarly, if any of the other points are at infinity, we simply cancel the terms containing the point, and the result is the cross ratio. Remember that at most one point may be at infinity, because the points must be distinct, and there is only one point at infinity on the projective line.

Acknowledgment

Many thanks to Olaf Hall-Holt for helping me understand Plücker coordinates.

References

[1] H. S. M. Coxeter. *Projective Geometry.* Toronto: University of Toronto, second edition, 1974.

[2] O. Faugeras. *Three-Dimensional Computer Vision.* Cambridge, MA: MIT Press, 1993.

[3] L. Guibas. *Lecture notes for CS348a: Computer Graphics – Mathematical Foundations.* Stanford University, Autumn 1996.

[4] Q.-T. Luong and O. D. Faugeras. Fundamental matrix: Theory, algorithms, and stability analysis. *International Journal of Computer Vision*, 17(1):43–75, 1996.

[5] J. L. Mundy and A. Zisserman. *Geometric Invariance in Computer Vision.* Cambridge, MA: MIT Press, 1992.

[6] Z. Zhang and G. Xu. *Epipolar Geometry in Stereo, Motion and Object Recognition: A Unified Approach.* Kluwer Academic Publishers, 1996.