The perspective transform in the PIL

Bill Casselman
Mathematics Department, University of British Columbia
cass@math.ubc.ca

The Python Image Library includes a method that applies a 2D projective transformation

$$(x, y) \longmapsto \left(\frac{ax + by + c}{gx + hy + 1}, \frac{dx + ey + f}{gx + hy + 1} \right)$$

to an image (and fits it into a rectangle of specified size). The octuple (a, b, c, d, e, f, g, h) is one of the arguments to the transform. But often the more convenient way to describe such a transformation is to specify what the transformation does to the points of the coordinate square

$$e_0 = (0, 0), \quad e_1 = (1, 0), \quad e_2 = (0, 1), \quad e_3 = (1, 1)\,.$$

The question then arises, suppose we know what four points f_i we want to transform the e_i into, what are the corresponding a, b, etc.? It turns out that as long as no three of the f_i belong to a single line, these coefficients are determined uniquely. The rest of this note will develop a formula for them.

We can evaluate the effect of the transform on these points:

$$\begin{aligned} e_0 &\longmapsto (c, f) \\ e_1 &\longmapsto \left(\frac{a + c}{g + 1}, \frac{d + f}{g + 1} \right) \\ e_2 &\longmapsto \left(\frac{b + c}{h + 1}, \frac{e + f}{h + 1} \right) \\ e_3 &\longmapsto \left(\frac{a + b + c}{g + h + 1}, \frac{d + e + f}{g + h + 1} \right)\,. \end{aligned}$$

We want to find out what a, b, etc. are, knowing what the vectors f_0, etc. on the right hand side are. We can see immediately at least what the coefficients c and f have to be, since we have $(c, f) = f_0$. But solving for the rest doesn't look trivial. Those fractions are messy.

Any transformation with fractions in its formula is bound to be awkward, and it is best to think of things in homogeneous coordinates. This amounts to considering 2D points as situated in 3D, and projective transformations as moving things around in 3D. The **projective plane** is by definition the set of lines through the origin in 3D. The usual plane can be embedded into the projective plane by mapping (x, y) to the line through $(0, 0, 0)$ and $(x, y, 1)$. The original sets of points e_i and f_i now become the lines through the sets

$$\begin{aligned} E_0 &= (0, 0, 1) \\ E_1 &= (1, 0, 1) \\ E_2 &= (0, 1, 1) \\ E_3 &= (1, 1, 1) \end{aligned}$$

and

$$\begin{aligned} F_0 &= (x_0, y_0, 1) \\ F_1 &= (x_1, y_1, 1) \\ F_2 &= (x_2, y_2, 1) \\ F_3 &= (x_3, y_3, 1)\,. \end{aligned}$$

But coordinates in 3D are now **homogeneous**—the points (x, y, z) and (cx, cy, cz) determine the same line as long as $c \neq 0$, hence are projectively equivalent.

Every invertible 3×3 matrix gives rise to a linear transformation in 3D, and it is one that takes such lines through the origin to other lines through the origin. It hence acts as a transformation of the projective plane. The effect of the transformation on lines is invariant under scalar multiplication of the matrix. The projective transformation in the PIL may therefore be described as taking

$$(x, y, 1) \longmapsto (ax + by + c, dx + ey + f, gx + hy + 1)\,,$$

and hence characterized by a 3×3 matrix

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{bmatrix}.$$

The problem now is, *given the four points f_i, how to find the matrix taking E_i to F_i?* This is a special case of a slightly more general problem: given any four points in the projective plane, no three of which lie on a projective line (i.e. a plane in 3D), find a projective transformation taking the lines through the E_i to these points (i.e. lines). In Chapter 10 of the book **Mathematical Illustrations** I give an argument for a solution of this problem that is mathematically well motivated, but a little inefficient in practice. (Something similar can be found in the articles by Davis Austin and Jim Blinn in the reference list.) But here I'll explain a less sophisticated one.

The point is to keep in mind that we are really moving lines around, not points. Therefore what I need are non-zero constants α, β and a 3×3 matrix T such that

$$\begin{aligned} TE_0 &= F_0 \\ &= (x_0, y_0, 1) \\ TE_1 &= \alpha F_1 \\ &= (\alpha x_1, \alpha y_1, \alpha) \\ TE_2 &= \beta F_2 \\ &= (\beta x_2, \beta y_2, \beta) \end{aligned}$$

subject to the further condition that

($\bullet$) *the point TE_3 must be be a scalar multiple of F_3.*

The constants α and β will determine the matrix T, since $E_1 - E_0$, $E_2 - E_0$, and E_0 make up the standard basis of $\mathbb{R}^3$. Once we know what T does to these we know the columns of the matrix. Or you can also use more directly the fact that $\alpha = g + 1$ and $\beta = h + 1$.

How do we choose α and β so as to satisfy ($\bullet$)? we know that $E_3 = E_1 + E_2 - E_0$ so

$$TE_3 = (-x_0 + \alpha x_1 + \beta x_2, -y_0 + \alpha y_1 + \beta y_2, -1 + \alpha + \beta)\,.$$

This is supposed to be a multiple of F_3, so we must have

$$(-x_0 + \alpha x_1 + \beta x_2, -y_0 + \alpha y_1 + \beta y_2, -1 + \alpha + \beta) = \text{ a multiple of } (x_3, y_3, 1)\,.$$

The multiple must be $-1 + \alpha + \beta$, so we must solve

$$\begin{aligned} (-1 + \alpha + \beta)x_3 &= -x_0 + \alpha x_1 + \beta x_2 \\ (-1 + \alpha + \beta)y_3 &= -y_0 + \alpha y_1 + \beta y_2 \end{aligned}.$$

This gives us a 2×2 system of equations

$$\begin{aligned} (x_1 - x_3)\alpha + (x_2 - x_3)\beta &= x_0 - x_3 \\ (y_1 - y_3)\alpha + (y_2 - y_3)\beta &= y_0 - y_3\,. \end{aligned}$$

which we can solve as long as the determinant

$$\begin{vmatrix} x_1 - x_3 & x_2 - x_3 \\ y_1 - y_3 & y_2 - y_3 \end{vmatrix} = (x_1 - x_3)(y_2 - y_3) - (x_2 - x_3)(y_1 - y_3) \neq 0\,.$$

This is equivalent to the condition that f_1, f_2, and f_3 do not lie on a line.

For an explicit solution, let

$$A = \begin{bmatrix} x_1 - x_3 & x_2 - x_3 \\ y_1 - y_3 & y_2 - y_3 \end{bmatrix}.$$

Then

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} y_2 - y_3 & x_3 - x_2 \\ y_3 - y_1 & x_1 - x_3 \end{bmatrix},$$

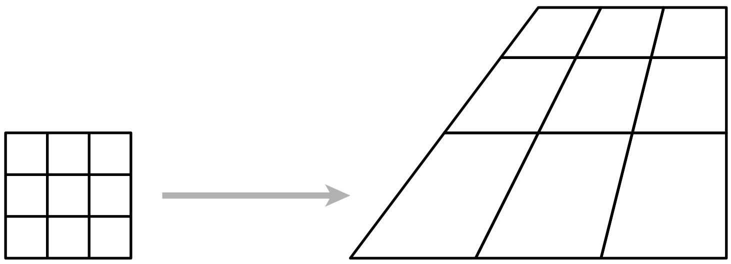
and

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = A^{-1} \begin{bmatrix} x_0 - 1 \\ y_0 - 1 \end{bmatrix}.$$

Finally, once we know α and β we set

$$T = \begin{bmatrix} \alpha x_1 - x_0 & \beta x_2 - x_0 & x_0 \\ \alpha y_1 - y_0 & \beta y_2 - y_0 & y_0 \\ \alpha - 1 & \beta - 1 & 1 \end{bmatrix}.$$

Incidentally, when applying a perspective transform, you have to be careful. The transform is not defined on the line $gx + hy + 1 = 0$, which is said to pass off to infinity. You might want to check this condition before getting an error from an attempt to divide by zero. One simple way to deal with thsi problem is to work up until the last step in terms of homogeneous 3D coordinates.



References

David Austin, **Using projective geometry to correct a camera**, the March, 2013 AMS Feature Column, available at

`http://www.ams.org/samplings/feature-column/fc-2013-03`

Jim Blinn, *Inferring Transforms*, Chapter 13 in **Notation, Notation, Notation** (extracts from his column in *Computer Graphics and Applications*), Morgan Kaufmann, 2003.

Bill Casselman, **Mathematical Illustrations**, Cambridge University Press, 2004. Also available at

`http://www.math.ubc.ca/~cass/graphics/text/www/index.html#main`