Kristina Poon

Homography relates 3-D coordinates of an object in an image to a corresponding 2-D coordinate pair and vice-versa. This allows objects or scenes found in images to be manipulated in their shape and vantage point by mathematical means.

H is a 3x3 matrix of unknowns:

$$H = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix}$$

The direct linear transform algorithm allows one to solve for the H matrix using $Aq = b$ form.

Direct Linear Transform:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \underset{b}{\begin{bmatrix} \tilde{x}' \\ \tilde{y}' \\ \tilde{w}' \end{bmatrix}} = \underset{q}{H} \underset{A}{\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}}$$

q = unknown

A = known

b = solution

Here, the left-most matrix denotes the new coordinates that the original coordinates have been transformed to. The right-most matrix represents the original coordinates. When the matrix denoted b is solved by taking the dot product of the inverse of the homography, the resulting matrix is divided by its $\tilde{w}'$ component yielding the final homogeneous (x, y) coordinates.

In order to solve for the homography, we use a matrix A that is $2N \times 9$ where N is the number of point pairs between a 2D and 3D image that produces a matrix of the form:

$$A = \begin{bmatrix} -x & -y & -1 & 0 & 0 & 0 & x'x & x'y & x' \\ 0 & 0 & 0 & -x & -y & -1 & y'x & y'y & y' \end{bmatrix}$$

x and y denote the corresponding coordinate values of points taken from the original source image while those with primes denote the new image's corresponding coordinates.

From this matrix A, we compute the singular value decomposition (svd) form $A = UDV^T$.
The solution of the homography turns out to actually be the nullspace of A, which is the last column of V from the computed svd.