Homography Estimation

Jonathan Ventura

September 9, 2016

A *homography* relates image projections of points on a plane. Corresponding points (x_1, y_1) and (x_2, y_2) are related by

$$w \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = \mathsf{H} \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} \tag{1}$$

for some scalar w. H is the 3×3 homography matrix.

$$\mathsf{H} \equiv \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \tag{2}$$

Expanding Equation 1 gives three equations:

$$wx_1 = h_{11}x_2 + h_{12}y_2 + h_{13} \tag{3}$$

$$wy_1 = h_{21}x_2 + h_{22}y_2 + h_{23} \tag{4}$$

$$w = h_{31}x_2 + h_{32}y_2 + h_{33} \tag{5}$$

Plugging the expression for w into the previous two equations leaves two equations:

$$(h_{31}x_2 + h_{32}y_2 + h_{33})x_1 - (h_{11}x_2 + h_{12}y_2 + h_{13}) = 0 \tag{6}$$

$$(h_{31}x_2 + h_{32}y_2 + h_{33})y_1 - (h_{21}x_2 + h_{22}y_2 + h_{23}) = 0 \tag{7}$$

Re-arranged as a matrix equation:

$$\begin{bmatrix} -x_2 & 0 & x_1x_2 & -y_2 & 0 & x_1y_2 & -1 & 0 & x_1 \\ 0 & -x_2 & x_2y_1 & 0 & -y_2 & y_1y_2 & 0 & -1 & y_1 \end{bmatrix} \begin{bmatrix} h_{11} \\ h_{21} \\ h_{31} \\ h_{12} \\ h_{22} \\ h_{32} \\ h_{13} \\ h_{23} \\ h_{33} \end{bmatrix} = \mathbf{0} \tag{8}$$

Stacking N pairs of such equations (from N point correspondences) gives $2N$ equations here denoted as a matrix-vector product

$$\mathsf{A}\mathbf{h} = \mathbf{0} \tag{9}$$

which can be solved by finding a null vector of A. Solving for the variables governing the similarity relationships in this way is often called the Direct Linear Transform (DLT).

The *re-projection error* gives a geometric measure of how well the estimated homography fits a given correspondence.

$$\text{err} = \sqrt{(x_1 - \frac{(h_{11}x_2 + h_{12}y_2 + h_{13})}{(h_{31}x_2 + h_{32}y_2 + h_{33})})^2 + (y_1 - \frac{(h_{21}x_2 + h_{22}y_2 + h_{23})}{(h_{31}x_2 + h_{32}y_2 + h_{33})})^2} \tag{10}$$

Because the DLT minimizes (in a least-squares sense) an algebraic error rather than this geometric error, it is often not the best solution when faced with noisy measurements. However, DLT is typically used to provide an initialization to non-linear refinement which minimizes a geometric error. The combination of DLT followed by non-linear refinement is called the Gold Standard method [1].

References

[1] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision.* Cambridge University Press, 2003.