1 DLT for computing Homography

Previously we discussed the use of homogeneous coordinates in computing projective transformations. The homogeneous coordinate representation enables us to map these projective transformations as linear mappings, with the 3×3 homography matrix transforming the homogeneous vectors. Recall that

- The Homography matrix H is non-singular and has 8 degrees of freedom (DoF). Since it is homogeneous matrix, the scaling factor can be arbitrarily set to a constant (e.g., you can think of the last element of the 9-element matrix set to 1). Hoever, proper care should be taken that this particular element is not close to 0, else you will have stability issues in estimating the H numerically.
- Any non-singular 3×3 matrix H is a homography; Similarly any homography can be represented by such a 3×3 matrix H.
- A homography transformation maps lines in one space to lines in the transformed space.

In this lecture notes, we will explore how to go about estimating the homography matrix given 2 images and a set of corresponding points. The first step is to write down the equations that relate the homogeneous coordinates of the corresponding points.

Setting up the Homography equations

Objective: Given $n \geq 4$ 2D to 2D point correspondences $\{\mathbf{x_i} \leftrightarrow \mathbf{x'_i}\}$, determine the 2D homography matrix H such that $\mathbf{x'_i} = H\mathbf{x_i}$.

It involves Homogeneous vectors. The vectors $\mathbf{x'_i}$ and $H\mathbf{x_i}$ are not equal as these are homogeneous vectors, have the same direction but may differ in magnitude by a non-zero scale factor. Hence, we can write:

$$\mathbf{x'_i} \times H\mathbf{x_i} = \mathbf{0} \tag{1}$$

$$H\mathbf{x_i} = \begin{pmatrix} \mathbf{h}^{1^T}\mathbf{x}_i \\ \mathbf{h}^{2^T}\mathbf{x}_i \\ \mathbf{h}^{3^T}\mathbf{x}_i \end{pmatrix} \tag{2}$$

where $\mathbf{h}^{j^T}$ is the j-th row of H. Then, with $\mathbf{x'}_i = (x'_i, y'_i, w'_i)^T$,

$$\mathbf{x'_i} \times H\mathbf{x_i} = \begin{pmatrix} y'_i\mathbf{h}^{3^T}\mathbf{x}_i - w'_i\mathbf{h}^{2^T}\mathbf{x}_i \\ w'_i\mathbf{h}^{1^T}\mathbf{x}_i - x'_i\mathbf{h}^{3^T}\mathbf{x}_i \\ x'_i\mathbf{h}^{2^T}\mathbf{x}_i - y'_i\mathbf{h}^{1^T}\mathbf{x}_i \end{pmatrix} \tag{3}$$

Noting that $\mathbf{h}^{j^T}\mathbf{x}_i = \mathbf{x}_i^T\mathbf{h}^j$, $j = \{1, 2, 3\}$, we can write:

$$\begin{bmatrix} \mathbf{0}^T & -w'_i\mathbf{x}_i^T & y'_i\mathbf{x}_i^T \\ w'_i\mathbf{x}_i^T & \mathbf{0}^T & -x'_i\mathbf{x}_i^T \\ -y'_i\mathbf{x}_i^T & x'_i\mathbf{x}_i^T & \mathbf{0}^T \end{bmatrix} \begin{pmatrix} \mathbf{h}^1 \\ \mathbf{h}^2 \\ \mathbf{h}^3 \end{pmatrix} = \mathbf{0} \tag{4}$$

- These are of the form $\mathbf{A}_i\mathbf{h} = \mathbf{0}$, where $\mathbf{A}_i$ is a 3×9 matrix, and $\mathbf{h}$ is a 9-vector of the entries of matrix H.

$$\mathbf{h} = \begin{pmatrix} \mathbf{h}^1 \\ \mathbf{h}^2 \\ \mathbf{h}^3 \end{pmatrix}, \quad H = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix} \tag{5}$$

  with h_i the i-th element of $\mathbf{h}$.

- The equation $\mathbf{A}_i\mathbf{h} = \mathbf{0}$ is linear in the unknowns $\mathbf{h}$. The matrix elements of $\mathbf{A}_i$ are quadratic in the known coordinates of the points.
- Out of the three equations, only two of them are linearly independent (recall: homogeneous equations, determined only upto a scaling factor.)
- We can rewrite the equations as

$$\begin{bmatrix} \mathbf{0}^T & -w'_i\mathbf{x}_i^T & y'_i\mathbf{x}_i^T \\ w'_i\mathbf{x}_i^T & \mathbf{0}^T & -x'_i\mathbf{x}_i^T \end{bmatrix} \begin{pmatrix} \mathbf{h}^1 \\ \mathbf{h}^2 \\ \mathbf{h}^3 \end{pmatrix} = \mathbf{0} \tag{6}$$

- This can be written as $\mathbf{A}_i\mathbf{h} = \mathbf{0}$, where $\mathbf{A}_i$ is a 2×9 matrix.
- These equations hold for any homogeneous coordinate representation (x'_i, y'_i, w'_i) of the point $\mathbf{x}'_i$, and in particular, $w'_i = 1$. Note that the (x'_i, y'_i) are the image coordinates of the point.
- *Q: What happens when $w'_i = 0$? what should you do in this case?*

Solving $A\mathbf{h} = 0$

- Homography matrix computations require solving a set of linear equations of the form $A\mathbf{h} = 0$.
  - For homography, each point correspondence results in 2 equations, hence with n point correspondences we have $2n$ equations. A is of dimension $2n \times 9$.
  - If $w'_i = 0$, the the pair of equations in 6 reduces to a single equation. You may then need the third equation that
- If A has a rank exactly equal to 8, then there is a unique solution.
- However, in practice, you collect many more equations (overdetermined system) than the number of free parameters, requiring you to minimize some metric rather than solving for an exact solution. This leads us to a solution that minimizes the Euclidean 2-norm, $||A\mathbf{h}||$.
- Note that a trivial solution to $A\mathbf{h} = 0$ is $\mathbf{h} = 0$, which is of no interest to us.

Use of Singular Value Decomposition

- Recall that $A\mathbf{h} = 0$ results in a homogeneous solution, the scale of $\mathbf{h}$ does not matter. We can choose $||\mathbf{h}|| = 1$.
- Minimizing $||A\mathbf{h}||$ subject to $||\mathbf{h}|| = 1$ can be efficiently solved using the *singular value decomposition (SVD)*

SVD Solution to $A\mathbf{h}h = 0$
The solution is the unit singular vector corresponding to the smallest singular value of A

The resulting algorithm is called the **Direct Linear Transform**

1.1 Algorithm 1. DLT for Homography

DLT#1 for Homography

Objective: Given $n \geq 4$ 2D to 2D point correspondences $\{\mathbf{x_i} \leftrightarrow \mathbf{x'_i}\}$, determine the 2D homography matrix H such that $\mathbf{x'_i} = H\mathbf{x_i}$.

DLT algorithm #1:

1. For each correspondence $\mathbf{x_i} \leftrightarrow \mathbf{x'_i}$, compute the matrix A_i.
2. Assemble the n, 2×9 matrices A_i into a single $2n \times 9$ matrix A.
3. Obtain the SVD of A. The unit singular vector corresponding to the smallest singular value is the solution $\mathbf{h}$. Specifically, if $A = UDV^T$ with D diagonal with positive diagonal entries, arranged in descending order down the diagonal, then $\mathbf{h}h$ is the last column of V.

1.2 Normalization of coordinates needed for DLT

Choice of image coordinates

- The choice of image coordinates is often arbitrary; You can choose the origin to be at the center of the image or at the corners of the image.
- It is undesirable to have the end results dependent on the choice of the coordinate system
- Unfortunately, *the DLT method is not invariant to similarity transformations!*
- Hence proper normalization of the coordinates is needed
- see figure on the next slide from Hartley and Zisserman that compares un-normalized computations with normalized computations

Un-normalized coordinates (Hartley 2006, Ch. 4)

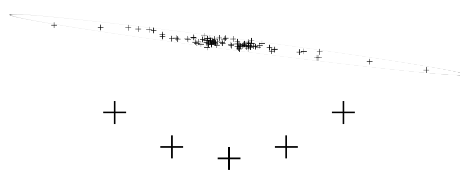

Figure 1: Results of Monte Carlo simulation of computation of 2D homographies. A set of 5 points (denoted by large crosses) was used to compute a 2D homography. Each of the 5 points is mapped (in the noise-free case) to the point with the same coordinates, so that homography H is the identity mapping. Now, 100 trials were made with each point being subject to 0.1 pixel Gaussian noise in one image. (For reference, the large crosses are 4 pixels across.) The mapping H computed using the DLT algorithm was then applied to transfer a further point into the second image. The 100 projections of this point are shown with small crosses and the 95% ellipse computed from their scatter matrix is also shown. These are the results without data normalization.

1.3 Normalization makes a big difference

Normalized coordinates (Hartley 2006, Ch. 4)

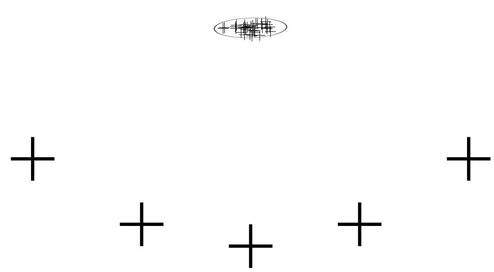

Figure 2: These are the results with normalization. The leftand rightmost reference points have (unnormalized) coordinates (130, 108) and (170, 108).

2 Data Normalization

Data Normalization is essential for DLT

5. The Direct Linear Transform

- DLT uses SVD of $A = UDV^T$ to obtain a solution to $A\mathbf{h} = 0$. Note that this system of equations is overderdetermined, no exact solution.
- Solution is obtained by setting $\mathbf{h}$ to the last column of V (corresponding to the smallest singular value), that minimizes $\|A\mathbf{h} = 0\|$ subject to $\|\mathbf{h}\| = 1$.
- If you go back and look at the elements of A, it consists of products of image pixel coordinates, which can be like $(x, y, w) = (128, 100, 1)$; Hence the scaling is vastly different, resulting in numerical instability. In matrix terminology, we say matrix A is ill-conditioned. Small changes in values could result in large changes in the computed parameters.

A sensible solution.
Move the origin and scale the coordinates

Coordinate normalization

The steps are simple.

(1) Center the origin
Compute the translation that takes takes points x_i and moves them (translation) such that the centroid of the points is the coordinate origin $(0, 0)$

(2) scale the coordinates
Now scale the coordinates that the average distance of the points to the centroid is $\sqrt{2}$

Overall, a similarity transformation T
Note that the steps (1) and (2) above, together, is a similarity transformation T that maps the original points $\{x_i\}$ to the transformed set of points $\{\tilde{x_i}\}$.

Normalization and Homography

Suppose T is the similarity transformation applied to $\mathbf{x}$ such that $\tilde{\mathbf{x}} = T\mathbf{x}$, and $T' : \tilde{\mathbf{x}'} = T'\mathbf{x}'$. Note T and T' are 3×3 homographies. We are computing the transformation H between $\mathbf{x}$ and $\mathbf{x}'$.

$$\mathbf{x}' = H\mathbf{x} \tag{7}$$

$$\tilde{\mathbf{x}'} = T'H\mathbf{x} = T'HT^{-1}\tilde{\mathbf{x}} \tag{8}$$

Hence, for $\tilde{\mathbf{x}} \leftrightarrow \tilde{\mathbf{x}'}$, $\tilde{H} = T'HT^{-1}$.

From this we get the desired transformation in the original pixel coordinates $\mathbf{x} \leftrightarrow \mathbf{x}'$ as $H = T'^{-1}\tilde{H}T$

Normalized DLT for Homography

Objective: Given $n \geq 4$ 2D to 2D point correspondences $\{\mathbf{x_i} \leftrightarrow \mathbf{x'_i}\}$, determine the 2D homography matrix H such that $\mathbf{x'_i} = H\mathbf{x_i}$.

DLT algorithm for Homography:

1. *Normalization of x.* Compute a similarity transformation T, consisting of a translation and scaling, that takes points x_i to a new set of points $\tilde{x}_i$ such that the centroid of the points $\tilde{x}_i$ is the coordinate origin $(0,0)^T$, and their average distance from the origin is $\sqrt{2}$

2. *Normalization of x'.* Compute a similar transformation T' for the points in the second image, transforming points x'_i to $\tilde{x'_i}$

3. *DLT:* Apply DLT#1 to the correspondences $\tilde{\mathbf{x}} \leftrightarrow \tilde{\mathbf{x'}}$ to obtain a homography $\tilde{H}$

4. *Denormailization.* Set $H = {T'}^{-1}\tilde{H}T$