ECE 661
Purdue U.

Avi Kak
Lecture 11

Estimating Homographies with Linear Least-Squares Minimization

Reference: "Multiple View Geometry in Computer Vision" by Hartley and Zisserman

- We want to estimate the 3x3 homography H that relates points in two different planes through $\boxed{x' = Hx}$. We are given a set of correspondences (x, x') that may include sampling and other measurement errors. Some of these correspondences may also be just plain wrong.
- We now express $H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$ as $H = \begin{bmatrix} h^{1T} \\ h^{2T} \\ h^{3T} \end{bmatrix}$ where $h^1 = \begin{pmatrix} h_{11} \\ h_{12} \\ h_{13} \end{pmatrix}$, $h^2 = \begin{pmatrix} h_{21} \\ h_{22} \\ h_{23} \end{pmatrix}$, $h^3 = \begin{pmatrix} h_{31} \\ h_{32} \\ h_{33} \end{pmatrix}$
- With this new way of writing the matrix H, we have $\boxed{Hx = \begin{pmatrix} h^{1T}x \\ h^{2T}x \\ h^{3T}x \end{pmatrix}}$
- Under the assumption that our correspondences are error-free (this assumption will be relaxed later), it must be the case that for a given correspondence (x, x'), the vectors x' and Hx are identical. It must therefore follow that $\boxed{x' \times Hx = 0}$ where '$\times$' means cross-product
- Writing $x' = \begin{pmatrix} x' \\ y' \\ w' \end{pmatrix}$, we can therefore say $\boxed{\begin{pmatrix} x' \\ y' \\ w' \end{pmatrix} \times \begin{pmatrix} h^{1T}x \\ h^{2T}x \\ h^{3T}x \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}}$.
- The vector cross product shown above yields: →
$$\boxed{\begin{aligned} \vec{0}^T h^1 - w'x^T h^2 + y'x^T h^3 &= 0 \\ w'x^T h^1 + \vec{0}^T h^2 - x'x^T h^3 &= 0 \\ -y'x^T h^1 + x'x^T h^2 + \vec{0}^T x^T h^3 &= 0 \end{aligned}}$$
These three equations are NOT linearly independent. If you add x' times the first equation with y' times the second equation, you get w' times the third equation.
- Therefore, retaining only the first two equations:
$$2 \updownarrow \underbrace{\begin{bmatrix} \underbrace{\vec{0}^T}_{3} & \underbrace{-w'x^T}_{3} & \underbrace{y'x^T}_{3} \\ w'x^T & \vec{0}^T & -x'x^T \end{bmatrix}}_{9} \underbrace{\begin{pmatrix} h^1 \\ h^2 \\ h^3 \end{pmatrix}}_{9\times1} = \underbrace{\vec{0}}_{2\times1}$$
- Each (x, x') correspondence gives us a pair of equations expressed by the form $\boxed{A\vec{h} = \vec{0}}$ where A is the 2x9 matrix shown in the previous bullet and $\vec{h}$ as: $\vec{h} = \begin{pmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \\ h_{33} \end{pmatrix}$
- Let's say we are given N correspondences $\boxed{(x_i, x_i') \text{ for } i = 1, 2, \ldots, N}$. We write $\boxed{A_i h = \vec{0}}$ for the two equations of the i^{th} correspondence. We can now stack all the resulting $2N$ equations in $\boxed{A\vec{h} = \vec{0}}$ with
$$A = 2N \updownarrow \underbrace{\begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_N \end{bmatrix}}_{9}$$
- So the homography we seek is the solution to the $2N$ homogeneous equations $\boxed{A\vec{h} = \vec{0}}$. (Note that 'homogeneous' in 'homogeneous equations' is NOT related to 'homogeneous' in 'homogeneous coordinates'.)
- More generally, let's say we have a system of homogeneous equations $\boxed{A\vec{h} = \vec{0}}$ with $\boxed{A}$ as a $\boxed{m \times n}$ matrix of known values and $\boxed{\vec{h}}$ as an n-element vector of unknowns. What are the conditions on A for there to exist a solution for $\vec{h}$?

- The conditions on the $m \times n$ matrix A for there to exist a solution for the n-element vector $\vec{h}$ are:
  - First note that the rank of A will never exceed n.
  - If the rank of A is exactly n, then the only solution is the trivial solution $\vec{h} = \vec{0}$.
  - If the rank of A is $n-1$ and $m = n-1$, then there exists a unique solution for $\vec{h}$ (unique up to a scale factor). This is the situation you have when you use 4 non-collinear correspondences (x, x') to calculate a homography. You will have 8 equations for the 9 unknowns in $\vec{h}$. That is, $m = 8$ and $n = 9$ and $\text{rank}(A) = 8$.
  - If the rank of A is less than $n-1$, then there exist infinitely many solutions for $\vec{h}$ (and, generally, we do NOT want any of them). In this case, we have an underdetermined system of equations.
  - If the rank of A equals $n-1$ and $\boxed{m \geq n}$, we have an overdetermined system of equations. Now, strictly speaking, there exists no solution at all — but this is exactly the situation we are interested in. In this case, we seek that solution $\vec{h}$ which minimizes $\|A\vec{h}\|$, which makes sense considering our starting constraint $A\vec{h} = \vec{0}$. Since unconstrained minimization of $\|A\vec{h}\|$ will result in the trivial solution $\vec{h} = \vec{0}$, the minimization of $\|A\vec{h}\|$ must be subject to the constraint $\boxed{\|\vec{h}\| = 1}$. This is referred to as Linear Least-Squares Minimization with homogeneous equations.
  - In summary, the rank of the $m \times n$ matrix A must be exactly $n-1$ for there to exist a solution. With $\boxed{\text{rank}(A) = n-1}$, if $m = n-1$, we have a unique solution up to scale. In this case, A has a 1-D null space and the solution vector lies in this space. On the other hand, when $\boxed{m \geq n}$, we must resort to Linear Least-Squares Minimization to construct a solution.

- The solution of the linear least-squares minimization problem $\boxed{\min \|A\vec{h}\|}$ $\boxed{\text{subject to } \|\vec{h}\| = 1}$ is given by that eigenvector of A^TA which corresponds to its smallest eigenvalue. Proof: By Singular Value Decomposition $\boxed{A_{m\times n} = U_{m\times n} \cdot D_{n\times n} \cdot V^T_{n\times n}}$ wher U is orthogonal $(U^TU = I_{n\times n})$ and norm-preserving $(\|U\vec{x}\| = \|\vec{x}\|)$, D is a diagonal matrix of singular values, and V is orthonormal $(V^TV = I_{n\times n})$. Additionally, the column vectors of V are the eigenvectors of A^TA. Since U is norm-preserving, $\boxed{\|A\vec{h}\| = \|UDV^T\vec{h}\| = \|DV^T\vec{h}\|}$. Now set $\vec{y} = V^T\vec{h}$. Since V amounts to pure rotation, $\|\vec{y}\| = \|\vec{h}\|$. Therefore, our original problem $\boxed{\min \|A\vec{h}\| \text{ subject to } \|\vec{h}\|=1}$ becomes $\boxed{\min \|D\vec{y}\| \text{ subject to } \|\vec{y}\|=1}$ Assuming that the singular values on the diagonal in D are in descending order (which is always the case), solving $\boxed{\min_{\vec{y}} \|D\vec{y}\| \text{ subject to } \|\vec{y}\| = 1}$ is trivial since the solution is given by $\boxed{\vec{y} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}}$. That is, in the solution vector $\vec{y}$, all the elements are zero except for the last one, which is 1. Now we can write the solution for $\vec{h}$ as $\boxed{\vec{h} = V^{-T}\vec{y}}$. Since $V^TV = I$, we have $V^T = V^{-1}$, and $V^{-T} = V$. Therefore, we have for the solution $\vec{h} = V\vec{y}$. This implies that $\vec{h}$ should be set to the last column vector in V, which will be the eigenvector of A^TA corresponding to the smallest singular value in D.

Linear Least-Squares Minimization with Inhomogeneous Equations

- Since only the ratios of the elements of the homography H are important, we may assume arbitrarily that $\boxed{h_{33}=1}$ (although later we will see that this is not always a safe assumption).

- Let's substitute $h_{33}=1$ in $\boxed{\begin{bmatrix} \vec{0}^T & -w'\mathbf{x}^T & y'\mathbf{x}^T \\ w'\mathbf{x}^T & \vec{0}^T & -x'\mathbf{x}^T \end{bmatrix}\begin{pmatrix} h^1 \\ h^2 \\ h^3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}}$. The resulting form can be expressed as $\boxed{\begin{bmatrix} 0 & 0 & 0 & -w'x & -w'y & -w'w & y'x & y'y \\ w'x & w'y & w'w & 0 & 0 & 0 & -x'x & -x'y \end{bmatrix}\begin{pmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \end{pmatrix} = \begin{pmatrix} -y'w \\ x'w \end{pmatrix}}$ where we use $\mathbf{x} = \begin{pmatrix} x \\ y \\ w \end{pmatrix}$ to parallel $\mathbf{x}' = \begin{pmatrix} x' \\ y' \\ w' \end{pmatrix}$.

- Therefore, a single $(\mathbf{x}, \mathbf{x}')$ correspondence gives us $\boxed{\underbrace{A}_{2\times8}\,\underbrace{\vec{h}}_{8\times1} = \underbrace{\vec{b}}_{2\times1}}$

- Given N correspondences $(\mathbf{x}_i, \mathbf{x}_i')$, $i = 1, 2, \ldots, N$, we can write $\boxed{A_i\vec{h} = \vec{b}_i}$ for the i^{th} correspondence. Stacking together the equations for all N correspondences gives ~~the~~ us 2N inhomogeneous equations that can be expressed as $\boxed{A\vec{h} = \vec{b}}$ where A is a $\boxed{2N\times8}$ matrix of known values and $\vec{b}$ a 2N-element vector of known values. And $\vec{h}$ has 8 unknowns

- To gain insights into when there exists a solution to the system of 2N inhomogeneous equations derived above, let's consider the general problem of solving $\boxed{A\vec{h} = \vec{b}}$ with A as an $m\times n$ matrix of knowns, $\vec{b}$ as an m-vector of knowns, and $\vec{h}$ as an n-vector of unknowns:
  - If the rank of A is less than n, we have infinitely many solutions, but we do not want any of them. (This is the underdetermined case.)
  - If $\boxed{\text{rank}(A) = n}$ and also $\boxed{m = n}$, we have a unique solution for $\vec{h}$. (Despite its uniqueness for the solution, we are NOT interested in this case because of possible noise in the values in A and $\vec{b}$.)

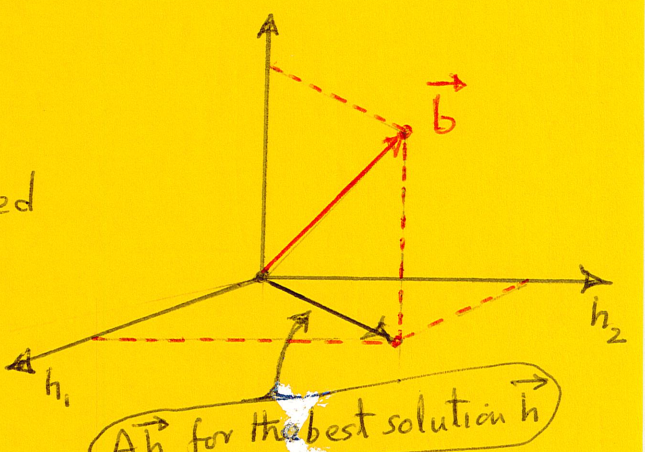
— If $\boxed{\text{rank}(A)=n}$ and $\boxed{m>n}$, there does NOT exist a solution for $\vec{h}$, but this is the case we are interested in, as explained below.

- From this point on, we are only considering the case $\boxed{m>n}$ and $\boxed{\text{rank}(A)=n}$ for solving $A\vec{h}=\vec{b}$ with A as an $m\times n$ matrix and $\vec{b}$ as an m-vector. The n elements of $\vec{h}$ are the unknowns.

- For this case, obviously, $\vec{b}\in\mathbb{R}^m$. But what space does $A\vec{h}$ belong to? Nominally, $A\vec{h}$ also resides in $\mathbb{R}^m$, since $A\vec{h}$ is an m-dimensional vector. More precisely speaking, though, $\boxed{A\vec{h}\in\mathbb{R}^n\subset\mathbb{R}^m}$. The reason for that is that the space spanned by the m row vectors of A is only n-dimensional since each row vector is just n-dimensional. So the mapping $A\vec{h}$ gives us nominally m-dimensional vectors but they reside in an n-dimensional subspace of $\mathbb{R}^m$. [It is interesting to reflect on the fact that the dimensionality of the column space is also n since we have only n column vectors.]

What this tells us is that the rank of a matrix determines the intrinsic dimensionality of both the row space and the column space.

- For illustration, consider the case $m=3$ and $n=2$. So we have 2 unknowns in $\vec{h}=\begin{pmatrix}h_1\\h_2\end{pmatrix}$ and 3 inhomogeneous equations. Additionally, we have $\vec{b}\in R^3$ and $A\vec{h}\in\mathbb{R}^2$. This can be visualized as shown at right. (Note that this is a oversimplified depiction of reality. The $\mathbb{R}^2$ space spanned by $A\vec{h}$ does not have to coincide with the (h_1,h_2) plane.) We are now faced with the following question: Given that $A\vec{h}$ spans some $\mathbb{R}^2$ in the $\mathbb{R}^3$ space in which $\vec{b}$ resides, what is the best solution for $\vec{h}$?



- The best choice for $\vec{h}$ will bring $A\vec{h}$ as close as possible to $\vec{b}$. In other words, the best solution will minimize $\|\vec{b}-A\vec{h}\|$. Such a solution would correspond to dropping a perpendicular from $\vec{b}\in\mathbb{R}^m$ to the subspace spanned by $A\vec{h}$, as shown in the figure above. Algebraically, the point $\vec{h}$ where the perpendicular from $\vec{b}$ meets the subspace spanned by $A\vec{h}$ is given by $\boxed{A^T(\vec{b}-A\vec{h})=\vec{0}}$. In other words, the linear least squares solution for $\vec{h}$ must satisfy $\boxed{(A^TA)\vec{h}=A^T\vec{b}}$, which yields $\boxed{\vec{h}=(A^TA)^{-1}A^T\vec{b}=A^+\vec{b}}$ where $\boxed{A^+=(A^TA)^{-1}A^T}$ is known as the pseudoinverse of A. (Note that $A^+A=(A^TA)^{-1}A^TA=I$.)

- Note that A^TA is a square $n\times n$ matrix whose rank is exactly n by virtue of the fact that $\text{rank}(A)=n$. A^TA being a square matrix of full rank implies that its inverse $(A^TA)^{-1}$ is guaranteed to exist. (In general, A^TA is symmetric and positive semidefinite. So its eigenvalues are real and nonnegative.)

- Before leaving the subject of linear least-squares minimization with inhomogeneous equations, let's go back to the beginning of this section and re-examine the $\boxed{h_{33}=1}$ assumption that was used to convert a system of 2N homogeneous equations into a system of 2N inhomogeneous equations.

- For most computations of homography between two images recorded under "ordinary" conditions (regarding the relationship between the camera and the scene), you would be safe with the $\boxed{h_{33}=1}$ assumption. However, if you are into, say, computational cameras, you may not want to make this assumption.

- The basic problem with the $h_{33}=1$ assumption is that it is incompatible with the possibility that a homography may map the origin in one plane to a point on ℓ_∞ in the other plane. Let H be a homography from a domain plane to a range plane. The homogeneous coordinate representation of the origin in the domain plane is $\begin{pmatrix}0\\0\\1\end{pmatrix}$. In the range plane, this point is mapped to $\boxed{H\cdot\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}h_{13}\\h_{23}\\h_{33}\end{pmatrix}}$. If for some reason we wanted the mapped point to be on ℓ_∞ in the range plane, we must allow h_{33} to become zero. Obviously, when $\boxed{h_{33}=0}$, we have $\boxed{H=\begin{bmatrix}A & \vec{t}\\ \vec{v}^T & 0\end{bmatrix}}$. This is obviously not in the affine group or any of its subgroups. This type of H belongs only at the root of our transformation hierarchy. In other words, the "distortion" created by this sort of H is purely projective. Note also that we cannot allow $\vec{v}$ to become zero since otherwise H will become singular.

Because last row is NOT (0 0 1)

- It's good to keep in mind that just because the origin is being mapped to an ideal point does not imply that other points in the vicinity of the origin in the domain plane will also be mapped to infinity in the range plane. Let's consider the point at a unit distance from the origin on the x-axis in the domain plane. Its image point in the range plane is given by $\boxed{\begin{bmatrix}h_{11} & h_{12} & h_{13}\\ h_{21} & h_{22} & h_{23}\\ h_{31} & h_{32} & 0\end{bmatrix}\begin{pmatrix}1\\0\\1\end{pmatrix} = \begin{pmatrix}h_{11}+h_{13}\\ h_{21}+h_{23}\\ h_{31}\end{pmatrix}}$ which, if $\boxed{h_{31}\neq 0}$, is some point at finite coordinates in the range plane. As mentioned earlier, at least one of h_{31} and h_{32} must be nonzero.

Data Normalization

- Sometimes you may be able to calculate a more accurate homography between two given images if you first precondition them by one or more of the following three steps: (1) shift the origin — possibly to the center of the images. Since A involves squares of pixel coordinates, A^TA involves pixel coordinates to the power of 4. Depending on how large the images are, and also depending on how the images are represented in the memory (integers vs. floating points), fourth power of the pixel coordinates may result in much too large a dynamic range for the values of the elements of A^TA. If that causes any implementation issues, you might be able to mitigate the problem by shifting the origin to the center of the images. ② Rescale the images if they are of very different scales to begin with and if that causes a problem with the numerical stability of the calculations. ③ Reorient the coordinate frame in one or both of the images.

- Fortunately, data preconditioning of the sort mentioned above is easily incorporated in the overall homography calculation as shown below.

- Following the notation used earlier, let x and x' represent the coordinates in the two given images. Let the preconditioning of the domain image be represented by a known transformation T, which itself is a homography. Similarly, let the preconditioning of the range image be represented by the known transformation T'. And let $\tilde{x}$ and $\tilde{x}'$ represent the coordinates in the preconditioned images. We have $\boxed{\hat{x} = Tx}$ and $\boxed{\tilde{x}' = T'x'}$. What we really want to calculate is H in $\boxed{x' = Hx}$. Since $\boxed{x = T^{-1}\tilde{x}}$ and $\boxed{x' = T'^{-1}\tilde{x}'}$, the transformation $\boxed{x' = Hx}$ turns into $\boxed{T'^{-1}\tilde{x}' = HT^{-1}\tilde{x}}$, which yields $\boxed{\tilde{x}' = T'HT^{-1}\tilde{x}}$. Therefore if $\tilde{H}$ is the homography that is actually computed using the preconditioned images, it is related to the original homography H by $\boxed{\tilde{H} = T'HT^{-1}}$.

- We can therefore recover the original homography H from the computed homography $\tilde{H}$ by $\boxed{H = T'^{-1}\tilde{H}T}$.