Readings
- [Moeslund](#) Chapter 10   Geometric Transformations
- [Grauman](#) 5.1, 5.2   Direct Linear Transform & RANSAC
- [Szeliski](#) 6.1, 6.1.1, 6.1.2, 6.1.4   2D Feature-Based Alignment

Summary

Motivation
- simply stitching images aligned via shifting will not produce a seamless mosaic or panorama; the two images need to first be projected onto a common

Homographies
- think of image as a function, then filtering changes the range (i.e. pixel values or intensity); warping changes the domain (i.e. pixel locations or indices)
- 2D transformations: translation, rotation, scaling, affine, projective
- image mosaicking needs a project transformation, also known as a homography
- a homography can be expressed as a matrix H; multiplying H with the original locations gives the transformed locations
- H can be solved from matching locations via the direct linear transform (DLT), which gives the linear least squares solution
- Homographies are applicable when scene is (approximately) planar, or when scene is captured under camera rotation only (as is the case for panoramas)
- Warp images by taking all source pixel intensities and sending them to their new locations
- A source pixel that ends up between multiple target pixel locations will distribute its intensity or colour to multiple neighbours;
- Target pixels that receive intensity from more than one source will be averaged

RANSAC
- DLT not robust since it finds an "average" transform, so solution gets corrupted by outliers
- Random sample consensus limits the impact of outliers
- Iteratively sample a subset of points, solve for model parameters, then count the number of other inliers based on some threshold; stopping condition either via
  - An N based on p (probability at least one iteration with no inliers in sampled set)
  - Early stopping criteria, when number of found inliers equal to expected proportion
- 2 RANSAC parameters:
  - Distance threshold for counting inliers, set by trial and error
  - Number of iterations to loop (N), either based on p or early stopping
- Robust feature alignment integrates solving of RANSAC loop of solving for homography
  - based on a small group of possible matches, estimate a homography H
  - search for more matches consistent with the hypothesized H

FAQs

Q: It is mentioned that scaling of the pixel space influences DLT. L9 part 2 video interpreted this statement as the resolution. The video also mentioned that small image or large image will probably result in different results. Is this referring to spatial or intensity resolution?

As the normalised DLT performs a whitening step to the two images, I assume that the small/large refer to intensity resolution?

A: It is referring to spatial resolution. The input the DLT are the spatial coordinates. We whiten the coordinate locations so that it is no longer dependent on the image resolution and position of the matches within the image.

Please refer to Slide 7 of Lecture 9 to note the key difference between our previous lectures e.g. on filtering which are operating on the range of the image function vs. L9 - 11 where we are looking at the domain of the image function.