Lecture 10: CS677

Sept 21, 2017

Admin and Review

- HW2 due Sept 26; 9AM
- Exam 1, Oct 10, class period
  - Details to follow next week
- Previous class
  - Hough Transform
  - Harris corner detector
  - SIFT features
- Today's objective
  - Homography
  - Stereo geometry
  - Stereo matching

Other Feature Detectors (in OpenCV)

- SURF: similar to SIFT but faster
- BRIEF: shorter descriptors
- FAST: faster corner detector
- ORB (FAST + BRIEF): claims to be like SIFT/SURF but faster, smaller and without patent issues
- More in:
  - https://en.wikipedia.org/wiki/Feature_detection_(computer_vision)
- For our course, knowledge of SIFT is sufficient

Homography

- Consider images of a points, P, in a given plane Π.
- In general, p = 1/z MP, where M is a 3 x 4 matrix
- For points on a plane, we can choose world coordinate system so Z axis is normal to plane, z=0 for all points
- There is still some matrix M, which projects P to p
  - M includes the rotations and translations to transform world coordinates to camera coordinates
  - In computing MP, note that third coordinate of P is 0 so we can drop the third column of M, giving a 3 x 3 matrix, say H, that maps (X,Y,1) to (wx', wy',w)
- Note that H is invertible
  - Geometrically: given image point, construct a ray and intersect with known plane.

Homography Mapping

- Now consider images of the point in two different cameras
- Each is related to P by a homography matrix, hence both are also related by another homography matrix (3 x 3).
  - Alternately, we can invert one image to give P, then project again
- $\boldsymbol{X_2} = \boldsymbol{H}\, \boldsymbol{X_1}$, where $\boldsymbol{X_1}$, $\boldsymbol{X_2}$ are image points in homogeneous coordinates, $\boldsymbol{H}$ is a 3 x 3 matrix.

$$\begin{bmatrix} x_2' \\ y_2' \\ w_2' \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} \frac{x_2'}{w_2'} \\ \frac{y_2'}{w_2'} \end{bmatrix}$$

- Note: homography also relates two views of a *non-planar* scene if only the camera orientation changes (no translation)
  - Second camera essentially maps image points of first camera which are, of course, on a plane (draw figure)
  - In spherical coordinates, just a translation (draw figure)

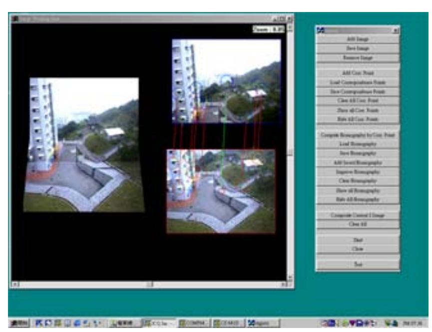
Creating a Mosaic

- Two (or more) images, obtained by camera rotations only are related by a homography
- If we can estimate parameters, we can transform to a common coordinate system and create a panorama (mosaic)
- We can compute homography parameters if we can match sufficient number of points between the two images
- H has nine parameters but scaling is arbitrary, fix $H_{33} = 1$; we are left with 8 unknown parameters
- For each point match between two images, we get two linear equations (from the matrix equation on the previous slide)
- 4 point matches should suffice to compute all parameters.
- How to find four correct matches?

Matching using SIFT Features

- How to match SIFT keypoints?
  - Similarity of description vectors (128-D)
  - Use Euclidean distance or angle between vectors
- Global match
  - Local matches may be ambiguous (a point may match well with several other points in another image)
  - Global matching may help reduce the ambiguities
- Global constraints
  - Assuming points lie in a plane: homography
    - Useful for image mosaicing (assume background is planar)
    - For examples see, http://www.cse.ust.hk/~cstws/research/641D/mosaic/
  - General stereo pairs: epipolar geometry

Example from UST.HK



Global Matching: RANSAC Approach

- Random Sample and Consensus (RANSAC)
- Select set of matches (4 for homography), at random
- Compute the homography matrix
  - A linear solution is possible, in principle, subject to some numerical issues
- Find other matches consistent with this transformation
  - If enough points are found, it is a good estimate, refit the transformation to the expanded set, otherwise discard
- Repeat the above process and choose one that is the "best"
  - Best may be defined by fit error or number of point matches
- RANSAC is a general process, can be used for fitting any model
  - One alternative to Generalized Hough Transform

RANSAC

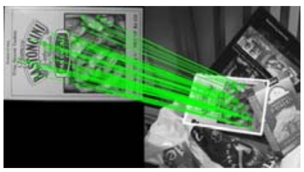
- How many trials are needed before we find the desired model?
- Depends on the number of "outliers"
  - Suppose that w is probability that a match is good
  - If choose n matches randomly (n =4 for homography), probability that all samples are good is (1-w^n)
  - If we repeat k times, probability that all trials result in bad samples is $(1-w^n)^k$
  - Say w= .5 (half the matches are outliers), n =4
    - w^4= .0625 , 1- w^4 = .9375
    - k=1, z = .9375; k=10, z = .52; k = 100, z = .0015
  - Say w= .1 (90% of the matches are outliers), n =4
    - w^4 = .0001 , 1- w^4 = .9999
    - k=1, z = .9999; k=10,000 z = .36; k = 100, 000, z = .00004

Localize Objects, an Example (from Lowe IJCV04)

- Locate Objects on the left, in the right image

A Recognition Example (from Lowe IJCV04)

Example from OpenCV

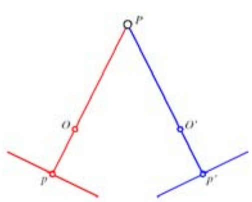


Inference of 3-D

- From a single image
  - Intensity (shape from shading)
  - Texture (requires regular texture, not covered in class/book)
  - 2-D shape (details not covered in class/book)
- From multiple images, static scene
  - From multiple fixed cameras or from camera motion
  - Stereo: Two fixed cameras
  - *Structure from motion*: Estimate camera motion and 3-D scene structure
- From object motion, stationary camera
  - Must infer object motion in addition to 3-D surface
- Most difficult case is when both the sensor and some of the objects in the scene move

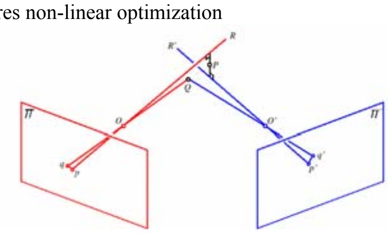
Steps of Stereo Analysis

- Point correspondence between views
  - Difficulties/details covered later
- Triangulation
  - Assuming calibrated cameras, each view provides a ray, say R and R', on which the 3-D point must lie: compute intersection

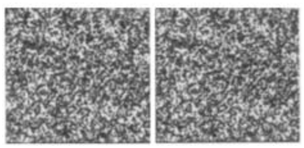


Triangulation: with measurement errors

- Solve for P as mid point of line minimizing *skew* distance between R and R'
- An algebraic method to solve $Zp = M\boldsymbol{P}$ and $Z'p' = M'\boldsymbol{P}$
  - 3 unknowns, 4 linear equations; See FP, p. 201 for details
- Find Q that minimizes $d^2(p,q) + d^2(p',q')$
  - Requires non-linear optimization

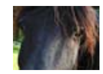


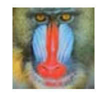
Some Examples



Stereo in Animals

- For humans, effective mostly for a distance of few meters
- Few animals possess it: birds of prey, *brachiating* mammals…





Next Class

- FP: Chapter 7, sections 7.1, 7.2, 7.4, 7.5