Panoramic Image Stitching

By Alfredo Ramirez

UCSD Jacobs | Electrical and Computer Engineering

CVRR Computer Vision and Robotics Research Laboratory

Selected Research

- "Automatic Panoramic Image Stitching using Invariant Features," by M. Brown & D. Lowe
- "Object Recognition from Local Scale-Invariant Features," by David G. Lowe
- "Image Alignment and Stitching: A Tutorial," by Richard Szeliski
- "Homography Estimation, " by David Kriegman
  - 4-point algorithm for homography estimation

Approach

1. Feature extraction using SIFT
   - SIFT is scale, translation, rotation and partially to illumination changes and affine projection
   - Scale invariant localization of feature by finding local minima and maxima of Difference of Gaussians
   - Gradient magnitude and orientation of feature candidates are calculated
   - Candidates are reduced by thresholding the gradient magnitude and limiting to corner points
   - Each feature is assigned an orientation

2. Calculate corresponding features in images
   - Nearest neighbor of SIFT features is calculated by finding the feature with shortest Euclidian distance
   - If multiple images involved, only the k nearest neighbors are found
   - Process is sped up using K-D tree

3. Estimate homographies using RANSAC
   - Homography is a 3x3 transformation matrix for converting coordinates in one image plane into coordinates in another image plane.
   - Estimation needs at least 4 corresponding points to calculate using linear least squares
   - RANSAC autonomously calculates best homography using feature candidates
     - Randomly choose 4 features and calculate homography. Test homography by comparing corresponding points. Repeat for certain number of trials
   - Estimate homography for m best matching images (images with most corresponding feature candidates)

(a) Image 1

(b) Image 2

(c) SIFT matches 1

(d) SIFT matches 2

(e) RANSAC inliers 1

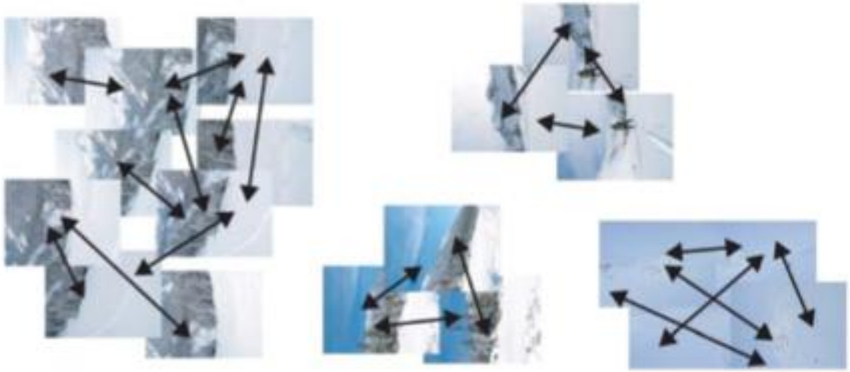
(f) RANSAC inliers 2

(g) Images aligned according to a homography

Figure 1. SIFT features are extracted from all of the images. After matching all of the features using a k-d tree, the m images with the greatest number of feature matches to a given image are checked for an image match. First RANSAC is performed to compute the homography, then a probabilistic model is invoked to verify the image match based on the number of inliers. In this example the input images are 517×374 pixels and there are 247 correct feature matches.

- 4. Image match verification
  - Compare number of RANSAC inliers to total number of features using probabilistic model
  - If total inliers is more than a certain threshold, then pair of images is matched and assumed to overlap in panorama
  - This technique allows the automated finding of panorama sequence by finding sets of pair image matches

(a) Image matches



(b) Connected components of image matches

(c) Output panoramas

Figure 2. Recognising panoramas. Given a noisy set of feature matches, we use RANSAC and a probabilistic verification procedure to find consistent image matches (a). Each arrow between a pair of images indicates that a consistent set of feature matches was found between that pair. Connected components of image matches are detected (b) and stitched into panoramas (c). Note that the algorithm is insensitive to noise images that do not belong to a panorama (connected components of size 1 image).

- 5. Project images onto panorama
  - Need to solve for rotation and focal length of all cameras in scene
  - Bundle adjustment is nonlinear least squares problem to find optimum estimation of parameters
  - Use rotations and focal lengths to project images onto one another to form panorama
- 6. Gain compensation
  - Reduces effect of differences in intensity between overlapping images
  - Solve least squares problem that finds optimum intensity gain for images that will reduce the error

- 7. Use Multi-Band Blending
  - Reduces effects around edges from unmodeled errors
  - Intensity at each pixel is a weighted sum of all overlapping pixels
  - Images are divided into frequency bands and weight maps for each band is created
  - Maps are created by Gaussian filtering the previous bands map
    - Lower frequency bands will be blended over wider range than higher frequencies

(a) Half of the images registered

(b) Without gain compensation

(c) With gain compensation

(d) With gain compensation and multi-band blending

Figure 5. Gain compensation. Note that large changes in brightness between the images are visible if gain compensation is not applied (a)–(b). After gain compensation, some image edges are still visible due to unmodelled effects such as vignetting (c). These can be effectively smoothed out using multi-band blending (d).

Questions

- 1. What are some potentialuses for image stitching
- 2. Briefly describe why SIFT was used for feature extraction.
- 3. How does RANSAC find the optimum homography and how does it know it’s the optimum one
- 4. Why is bundle adjustment needed in the algorithm
- 5. How is multiband blending better than linear blending