Motivation
- Internet photos from all over the world (Flickr, Street View, ...) allowing powerful 3D reconstruction techniques (Agarwal09, Crandall11, etc.)
- This work: Use the above to develop new techniques
  - Full 6-DOF intrinsic camera pose
  - Image retrieval as a by-product if needed
  - Powerful geometric constraints -- high confidence

Co-occurrence Prior for RANSAC
- Co-visibility gives powerful new prior for RANSAC sampling:
  - Choose points co-occurring often in database images

Direct 2D-to-3D Registration
- Powerful geometric constraints -- high confidence
- No intermediate retrieval step -- multiple pairwise image matching
- Full camera pose, thus exact location & orientation
- Image retrieval by product if needed

Challenges at Large Scale
- Noisy putative matches as database grows
- Harder to find good matches
- Harder to distinguish good matches from bad ones
- Much larger search space -> new techniques needed

Datasets
- **Landmarks**: (images from [Crandall et al. 09], top 1000 landmarks)
  - 20K images and 38M points in reconstructed 3D model
  - 10K test images from [Crandall et al. 11]
- **San Francisco**: (images from [Chen et al. 11], normal perspective only)
  - 5F, upright SIFT w/ histogram eq.: 800K images, 75M points
  - 5F, regular SIFT: 600K images, 30M points. 803 test images (both 5F & 5F)
- **Quad**: (images from [Crandall et al. 11])
  - 4830 images, 2M points, 348 accurately geotagged (within 1m) test images

Experiments

**Quantitative results**

Bidirectional Matching
- Forward matching: image features to model points
- Reverse matching: model points to image features
- Examples of frequently co-occurring points

Registration procedure
- Extract SIFT for query image
- Match image features to 3D points -> putative matches
  - by finding AANN in SIFT space, subject to ratio test
- RANSAC -- find true matches & pose, accept if at least 12 inliers

Online Demo: http://landmark.cs.cornell.edu