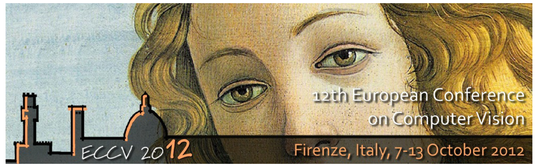
Worldwide Pose Estimation Using 3D Point Clouds

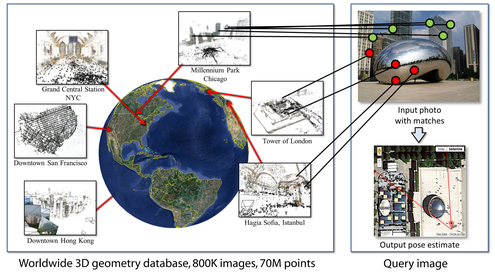
Yunpeng Li — EPFL-CVLab
Noah Snavely — Cornell University
Dan Huttenlocher — Cornell University
Pascal Fua — EPFL-CVLab

This work was supported in part by NSF grants IIS-0713185 and IIS-1111534, Intel Corporation, Amazon.com, Inc., MIT Lincoln Laboratory, and the Swiss National Science Foundation.



Motivation

- Internet photos from all over the world: Flickr, Street View, ...
- Powerful 3D reconstruction techniques ([Agarwal09, Crandall11], etc.)
- This work: Use the above to **register (hence recognize) new images**
- **Full 6-DOF+intrinsics camera pose**



Worldwide 3D geometry database, 800K images, 70M points

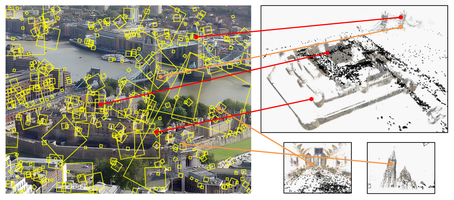
Query image

3D Point Cloud Models

- Created by SfM: SIFT, pairwise image matching, bundle adjustment
- 3D points - position, SIFT descriptors
- Database images -- registered to reconstruction
- **Visibility info** -- bipartite graph between database images and points

Direct 2D-to-3D Registration

- Powerful geometric constraints -- high confidence
- No intermediate retrieval step -- no multiple pairwise image matching
- Full camera pose, thus exact location & orientation
- Image retrieval a by-product if needed



Registration procedure

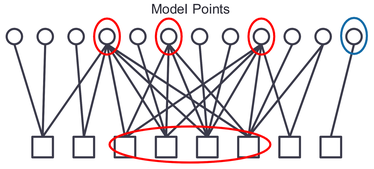
- Extract SIFT for query image
- Match image features to 3D points --> *putative matches*
  - by finding (A)NN in SIFT space, subject to *ratio test*
- RANSAC -- find true matches & pose, accept if at least12 inliers

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 scale --> new techniques needed**

Co-occurrence Prior for RANSAC

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



- A_p = set of images in which point p occurs
- Sample N (=3 or 4) pts with prob.: $\Pr_{\text{select}}(p_1, \ldots, p_N) \propto |A_{p_1} \cap \cdots \cap A_{p_N}|$
- Equiv. to sample sequentially with:

$$\Pr_{\text{select}}(p_i | p_1, \ldots, p_{i-1}) \propto \sum_{p_{i+1}, \ldots, p_N} |A_{p_1} \cap \cdots \cap A_{p_N}|$$

- For speed, approximate with: $\tilde{\Pr}_{\text{select}}(p_i | p_1, \ldots, p_{i-1}) \propto |A_{p_1} \cap \cdots \cap A_{p_i}|$
- Examples of frequently co-occurring points

Bidirectional Matching

- Forward matching: image features to model points
- Inverse matching: model points to image features (prioritized [Li et al. 10])
  - May recover matches too ambiguous for forward matching to find

Query image — Database images

- Combine both: perform inverse matching if forward matching fails

Datasets

- **Landmarks** (images from [Crandall et al. 09], top 1000 landmarks)
  - 200K images and 38M points in reconstructed 3D model, 10K test images

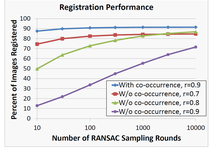
- **San Francisco** (images from [Chen et al. 11], normal perspective only)
  - SF1, upright SIFT w/ histogram eq.: 800K images, 75M points
  - SF0, regular SIFT: 600K images, 30M points. 803 test images (both SF0 & SF1)

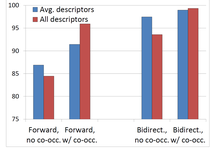
- **Quad** (images from [Crandall et al. 11])
  - 4830 images, 2M points, 348 accurately geotagged (within 1m) test images

Experiments

Quantitative results

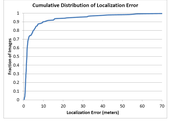
Registration rate for Landmarks, showing effectiveness of co-occurrence RANSAC and bidirectional matching





Building classification performance for San Francisco. Recall rate at 95% precision (as in [Chen et al. 11]). No perspective frontal images (PFI) used

		SF-0	SF-1
Ours (no GPS)	Avg. descriptors	50.2	58.0
	All descriptors	54.2	62.5
Chen *et al.*	No GPS	20	41
	With GPS	-	49



Localization accuracy for Quad

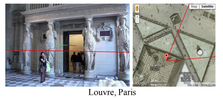
Further evaluate scalability by combining the point cloud model from all 3 datasets (Landmarks + SF0 + Quad). Over 800K images and 70M points in combined model

Query images \ Model	Individual	Combined
Landmarks	98.95	98.90
San Francisco	50.2	47.7
Quad	68.4	61.2

Visualization of registration and localization

Landmarks


Trafalgar Square, London


Louvre, Paris


Reichstag, Berlin


Key West, Florida


Machu Picchu, Peru


Shibuye, Tokyo

San Francisco


Corner of Davis and California


Corner of Clay and Sansome


Pacific Avenue

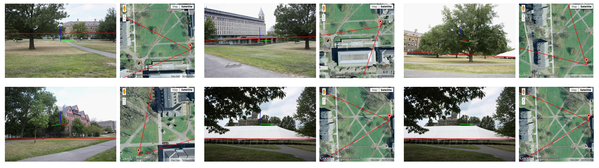

Pine St.


Corner of Pine and Stockton


Corner of Beach and Jones


Columbus Ave.

Quad



Mistakes

Inaccurate


Santa Maria del Mar, Barcelona


Plaça d'Espanya, Barcelona


Birmingham, Victoria Square

Incorrect

Missed

Online Demo:

http://landmark.cs.cornell.edu