Advanced Multimedia

Tamara Berg
Places & Photos
Reminders

• Start working on projects!
• Project update presentations – April 20.
• Project help day – April 22
Importance of Places

Tag classification into broad WordNet categories

Importance of Places
Types of Places

Location – GPS coordinates, location on a map.

Named places:
  Specific – Eiffel Tower, Golden Gate Bridge, Taj Mahal, ...
  General – Bakery, Store, Street, Park, ...

Presentation Outline

Locations:
  StreetView – Put Photos on a map from GPS coords.
  Geotagging - Put photos on a map manually.
  ZoneTag - Put photos on a map automatically from cell phones.
  Im2GPS – Use image matching to automatically infer a precise or rough location.

Named Places:
  Photo Tourism – Reconstruct a scene precisely for browsing from multiple views/photographers.
Google Street View displays photos taken from a fleet of Chevrolet Cobalts in United States. Pedestrian areas, narrow streets and park alleys that cannot be accessed by car are not always covered. However, sometimes Google Bikes are used.

On each of these cars (and bikes) 9 directional cameras for the 360° views, GPS units for positioning, Laser Range Finders for the measuring of buildings and 3G/Wi-Fi aerials for whereabouts on 3G and Wi-Fi hotspots are all mounted.

Street View
Geotagging

• The art of adding location information to things like a photo.
• You can geotag your photos using Flickr’s Organizr by dragging and dropping them on to the map where you took them.
• Then browse your or others’ photos according to location.
How might you infer which tags are more likely to be location related? 
ie “baby”, “house”, “empire state building”?
How might you infer which tags are more likely to be location related? ie “baby”, “house”, “empire state building”? How about event related?
Flickr Places

A page on Flickr for every place in the world.

Automatically created based on map and user data.
ZoneTag:  
Putting Your Photos in Context  

Mor Naaman  
Yahoo! Research Berkeley
ZoneTag?

“Everything in the world exists to end up in a photograph” — Susan Sontag

How would we:
Create/store?
Find?
Share?
Discover?
Why Cameraphones?

• Numbers, numbers...
  – 500,000,000
• Programmable
• Context-aware
• Network-connected
• Quality… it’ll be good enough

(Source: Future Image Inc.)
Current Mobile Experience

• Difficult to share (or even save!)
• Hard to find
  – No context
  – No semantic information
ZoneTag Experience

• 2-click upload, smooth experience
• Photo uploaded with location and time metadata
ZoneTag Experience

• Tagging **made** easy!
  – Tag/categorize/annotate your photos from the phone

http://zonetag.research.yahoo.com
Where do locations come from?

• Bluetooth GPS (when available)
• User-contributed cell tower mapping
Where do tags come from?

- You
- Your contacts
- Various datasets
- Customized streams

Served in context!
Suggested Tags

Your Contacts' Local Tags
- Nba
- golden state
- houston
- oakland coliseum
- oakland

Upcoming Venues
- Oakland Arena
My ZoneTag

• Add feeds, customize

ZONETAG™ PHOTOS

Manage your ZoneTag account

For more help, please email us at: zonetag-feedback@yahoo-inc.com

My ZoneTag Photos

Browse your ZoneTag photos

Current RSS Feeds

Upcoming: http://upcoming.org/syndicate/v2/my_even...
Friends: http://upcoming.org/syndicate/v2/friend...
Where2.0: http://www.google.com/calendar/feeds/3jm...
My ZoneTag

My Upcoming Events (Past Events...)

<table>
<thead>
<tr>
<th>Date</th>
<th>Event Name</th>
<th>Status</th>
<th>Comments</th>
<th>People</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jun 13+</td>
<td>O'Reilly Where 2.0 Conference</td>
<td>Attend</td>
<td>2</td>
<td>99</td>
</tr>
<tr>
<td>Jun 13</td>
<td>BayCHI program: Focus on Mobile Phones - 2 talk...</td>
<td>Attend</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>Jul 01</td>
<td>Jolie Holland</td>
<td>Watch</td>
<td>0</td>
<td>10</td>
</tr>
</tbody>
</table>

http://zonetag.research.yahoo.com
I’m Too Lazy (you’re not alone)

• Tagging is the means, not the goal
• Benefits even if you never tagged a single image

(Bradley Horowitz, elatable.com)
Typed (Faceted) Browsing

- http://zonetag.research.yahoo.com/photos

ZoneTag locations this week:

Amherst Belchertown Belmont Berkeley Brisbane Cambridge Chemnitz Conway
El Cerrito Foster City Fremont Half Moon Bay Hardwick Hartford Houston
Kansas City Lawrence Linwood Little Rock Los Altos Menlo Park Mountain
View New Haven New York North Haven Oakland Palo Alto Redwood City
Rockville San Carlos San Francisco Santa Clara Somerville
South San Francisco Stanford Sunnyvale Ware Wheelwright
Social Browsing

• My contacts at Berkeley, yesterday:
ZoneTag For All

Download, sign up, get the full scoop:
http://zonetag.research.yahoo.com
(Research prototype; currently supports Nokia Series 60 phones)
IM2GPS: estimating geographic information from a single image

James Hays and Alexei A. Efros
Carnegie Mellon University
Where?

What can you say about where these photos were taken?
How?

Collect a large collection of geo-tagged photos

6.5 million images with both GPS coordinates and geographic keywords, removing images with keywords like birthday, concert, abstract, ...

Test set – 400 randomly sampled images from this collection. Manually removed abstract photos and photos with recognizable people – 237 test photos.
How?

Features
Tiny images – 16x16 color images
Color histograms
Texton histograms – clustered responses to a bank of filters.
Line features – histogram describing statistics of straight lines in image
Gist descriptor + color
Geometric context (ground, sky, vertical)
How?

Data-driven geolocation:

For each input image compute features
Compute distance in feature space to all 6 million images in the database (each feature contributes equally).
Label the image with GPS coordinates of:
  1 nearest neighbor
k=120 nearest neighbors – probability map over entire globe.
Test Images
Distribution of photos
Results
Results
Results
Performance across database size

![Graph showing performance across database size with a log scale for database size and percentage of geolocations within 200km. The graph includes two lines: a green line labeled "First Nearest Neighbor Scene Match" and a red dotted line labeled "Chance - Random Scenes." The x-axis represents database size in thousands of images, and the y-axis represents the percentage of geolocations within 200km. The graph shows an increasing trend with larger database sizes.]
Performance across test set

- Chance: Random Scenes
- First Nearest Neighbor Scene Match
- Mean Shift Mode, Largest Cluster
- Best Possible Mean Shift Mode
- Best Possible of 120 Nearest Neighbors

Geolocation Error, km (exp scale)

Percentage of Test Set
Error with multiple guesses

![Graph showing the median distance from best guess to query location as a function of the number of guesses allowed (N). The graph compares different methods: Chance—N Random Scenes, N Nearest Neighbors, N Mean Shift Modes, Ordered by Cluster Size, Best Possible Mean Shift Mode, and Best Possible of 120 Nearest Neighbors.](image)
Estimating population density

Given a population density map of the world, estimate the population of an image by sampling at the estimated location. Images ranked by predicted population density.
Given an elevation map of the world, they can predict the elevation of an image according to its location. Images ranked by their estimated elevation.
Given landcover map, predict which images are most likely to be examples of each category
Urban/Rural Classification

![Diagram showing Urban/Rural Classification](image)
Photo Tourism: Exploring Photo Collections in 3D

Noah Snavely
Steven M. Seitz
University of Washington

Richard Szeliski
Microsoft Research

© 2006 Noah Snavely

Thanks to the authors for making slides available!
Photo Tourism
Photo Tourism overview

System for interactive browsing and exploring large collections of photos of a scene. Computes viewpoint of each photo as well as a sparse 3d model of the scene.
Input Photos
Photo Tourism overview

Input photographs → Scene reconstruction → Photo Explorer

[Note: Change to Trevi for]
Scene reconstruction

- Automatically estimate
  - position, orientation, and focal length of cameras
  - 3D positions of feature points

Diagram:
- Feature detection
- Pairwise feature matching
- Correspondence estimation
- Incremental structure from motion
Feature detection

- Detect features using SIFT [Lowe, IJCV 2004]
Feature detection

- Detect features using SIFT [Lowe, IJCV 2004]
SIFT Reminder

Image gradients → Keypoint descriptor
Pairwise feature matching

- Match features between each pair of images
Pairwise feature matching

- Refine matching using RANSAC [Fischler & Bolles 1987] to estimate fundamental matrices between pairs

Fundamental matrix –

$F$ is a 3x3 matrix with rank 2 such that for:

Corresponding points in stereo pair $y_1$ and $y_2$

$y_2^T F y_1 = 0.$
RANSAC

“Random Sample Consensus” – iterative method to estimate model parameters from observed data with outliers.

Algorithm:
1.) Pick 8 points, assume they are in correspondence.
2.) Estimate F.
3.) Calculate percentage of inliers.
4.) Iterate 1-3, k times.
5.) Output the best model found
Correspondence estimation

- Link up pairwise matches to form connected components of matches across several images
Structure from motion

\[
\text{minimize } f(R, T, P)
\]
Incremental structure from motion

- Optimize parameters for two cameras and common points
- Find new image with most matches to existing points
- Initialize new camera using pose estimation
- Bundle adjust
- Add new points
- Bundle adjust
Incremental structure from motion
Reconstruction performance

- For photo sets from the Internet, 20% to 75% of the photos were registered
- Most unregistered photos belonged to different connected components
- Running time: < 1 hour for 80 photos
  > 1 week for 2600 photos
Photo Tourism overview
Navigation controls

- Free-flight navigation
- Object-based browsing
- Relation-based browsing
- Overhead map
Free Flight Navigation
Object-based browsing
Relation-based browsing

Name: 55668857@N00_11...
Added by: 55668857@N00
Date: March 3, 2005, ...

Find all details

Find all similar images

Find all zoom outs

Zoom in

Move left

Move right

Zoom out
Rendering
Rendering
Annotations
Saint Basil's Cathedral
Trafalgar Square
Rockefeller Center
Mount Rushmore