Modeling and Recognition of Landmark Image Collections Using Iconic Scene Graphs

Xiaowei Li, Changchang Wu, Christopher Zach, Svetlana Lazebnik, Jan-Michael Frahm
Motivation

- Target problem: organizing community photo collections of famous landmark sites such as the Statue of Liberty

- We present a unified system for dataset collection, scene summarization, 3D reconstruction, and recognition for landmark images

- Approach: integrate 2D recognition and 3D structure-from-motion techniques for an efficient and scalable solution
Summary of approach

1. Appearance-based clustering
   • Run k-means clustering with gist descriptors (Oliva & Torralba, 2001) to find groups of images with roughly similar viewpoints and scene conditions

2. Geometric verification of clusters
   • Perform feature-based geometric matching between a few “top” images from each cluster
   • Select an iconic image for each cluster as the image with the most inliers

3. Construction of iconic scene graph
   • Perform geometric matching between every pair of iconic images
   • Create an edge for every pair related by a fundamental matrix or a homography
4. **Tag-based filtering**
   - Eliminate semantically irrelevant isolated nodes of the iconic scene graph

5. **Structure from motion**
   - Run graph cuts to break iconic scene graph into smaller components
   - Perform SFM separately on each component. Use a maximum-weight spanning tree to determine the order of incorporating images into the 3D model
   - Merge component models using geometric relationships along edges that were originally cut
   - Enlarge models by registering non-iconic images

6. **Recognition**
   - Register a new test image to the iconics using gist or vocabulary tree matching (Nister & Stewenius, 2006) followed by geometric verification
Overview

All images

Clustering with gist, intra-cluster verification

Iconic images

Pairwise matching of iconic images

Reconstructed components

SFM

Components of iconic scene graph

Iconic scene graph
Iconic scene graph for browsing

- **Level 1**: components of iconic scene graph
- **Level 2**: iconic images belonging to each component
- **Level 3**: images inside the gist cluster of each iconic
Statue of Liberty results

Originally: 45284 images

196 iconic images

New York

Registered images in largest model: 871
Points visible in 3+ views: 18675

Las Vegas

Tokyo
Statue of Liberty evaluation

Modeling
Unlabeled images: 42983  
Labeled images: 2301

Testing
1092 images

Stage 1: gist clustering  
Stage 2: per-cluster geometric verification  
Stage 3: per-image geometric verification  
Stage 4: tag-based filtering
Notre Dame results

Originally: 10840 images
105 iconic images

Registered images in largest model: 337
Points visible in 3+ views: 30802
Notre Dame evaluation

Modeling
Unlabeled images: 9760
Labeled images: 1080

Stage 1: gist clustering
Stage 2: per-cluster geometric verification
Stage 3: per-image geometric verification
Stage 4: tag-based filtering

Testing
1044 images
San Marco results

Originally: 43557 images

134 iconic images

Registered images in largest model: 749
Points visible in 3+ views: 39307
San Marco evaluation

Modeling
Unlabeled images: 38332
Labeled images: 5225

Stage 1: gist clustering
Stage 2: per-cluster geometric verification
Stage 3: per-image geometric verification
Stage 4: tag-based filtering

Testing
1094 images
Computing Iconic Summaries for General Visual Categories

Rahul Raguram and Svetlana Lazebnik

To appear at the First IEEE Workshop on Internet Vision (in conjunction with CVPR 2008)
Motivation

• We want to obtain complete, concise, and visually compelling summaries of image query results for general (and possibly abstract) categories

• At present, photo sharing websites such as Flickr don’t do a very good job of this
Summary of approach

- **Our definition:** an *iconic image* is a high-quality representative of a group of images consistent both in terms of appearance and semantics

- **Finding iconic images:**
  - Cluster appearance with gist (Oliva & Torralba, 2001)
  - Cluster tags with pLSA (Hofmann, 1999)
  - Form joint clusters by intersecting the two clusterings; retain only joint clusters that are large enough
  - Find representative iconic image for each joint cluster as the image with the highest quality score (Ke et al., 2006)

- **Displaying iconic summaries:** group iconic images by pLSA cluster (theme) and compute layout of pLSA clusters with multidimensional scaling
Interesting effect of joint clustering: “Visual rhymes”
Apple summary
18

Apple details

A  nyc applestore
    newyork newyorkcity

B  apple
    ipod
    nano
    mac

C  mac macintosh
    mini macmini

D  logo macintosh mac macbook

E  green apples red tree

F  fruit red macro food
Beauty summary

A. portrait woman beautiful girl
   - woman girl portrait nude
   - fashion model glamor studio

B. water nature beach ocean
   - california northern california nature beautiful

C. sky clouds nature sunset
   - flowers nature flower macro

D. nature flower macro rose
   - japan girls beautiful nippon
Beauty details

A portrait woman beautiful girl
B water nature beach ocean
C sky clouds nature sunset
D nature flower macro rose
Closeup summary

A. Macro drop splash water
   - blue abstract
   - water white

B. Bird nature flight gull
   - insect macro dof nature

C. Insect butterfly macro dragonfly
   - bee flower yellow nature

D. Eye macro eyes close

E. Portrait face bw macro
   - baby cute pink boy

F. Lips lip macro pink
   - strawberry strawberries berries red

- Macro drop splash water
- Blue abstract water white
- Bird nature flight gull
- Insect macro dof nature
- Eye macro eyes close
- Portrait face bw macro
- Baby cute pink boy
- Lips lip macro pink
- Strawberry strawberries berries red
Closeup details

A  macro drop splash water

B  bird nature flight gull

C  insect butterfly macro dragonfly

D  eye macro eyes close

E  portrait face bw macro

F  lips lip macro pink

Explore / Tags / closeup / clusters

Comparison: Flickr clusters
Love summary
Love details