Photo Tourism: Exploring Photo Collections in 3D

SIGGRAPH 2006

Noah Snavely
Steven M. Seitz
University of Washington

Richard Szeliski
Microsoft Research

© 2006 Noah Snavely
Photo Tourism
Photo Tourism overview

Input photographs → Scene reconstruction → Photo Explorer

- Relative camera positions and orientations
- Point cloud
- Sparse correspondence
Related work

- Image-based modeling

Debevec, et al.
SIGGRAPH 1996
FAÇADE: Modeling and Rendering Architecture from Photographs

P. Debevec, C. J. Taylor, J. Malik

SIGGRAPH 1997
Overview

• Take a few widely spaced photographs
• Build simple underlying model of scene
• Use correspondences between photos to adjust scene parameters
• Paste photos back onto simple geometry of scene for realistic façade
Photogrammetric Modeling

• User builds a simple geometric model using \textit{blocks}: primitive solid shapes
  – Example: boxes, wedges, prisms, frusta, surfaces of revolution

• User marks correspondences between images and model

• System fits model to images
Photogrammetric Modeling

• The system needs to solve for the parameters of blocks
  – Height, width, translation, rotation, etc.
Photogrammetric Modeling

- Knowns: image segments to block edge correspondences
- Unknowns: block parameters, camera position/orientation
- Architectural constraints reduce the number of unknowns
Photogrammetric Modeling

- Architectural constraints
  - Roof prism lies flush on building block
  - Stacked tower blocks share center axis
Example result

• Three of 12 input photographs for the University High School, Urbana, IL:

Slide credit: D. Luebke
View-Dependent Texture Mapping

• Given the model, treat each camera position as a “slide projector”

• Some images overlap!
  – Idea: pick image taken from viewpoint closest to desired rendering viewpoint
  – Better: use weighted average or do texture mapping on a per-pixel basis
View-Dependent Texture Mapping
Photogrammetric Modeling

- The Campanile, Berkeley

Original photograph with marked edges
Recovered model
Model edges projected onto photograph
Synthetic rendering
Photogrammetric Modeling

• Model of Berkeley campus constructed from 15 photographs:
The Campanile Movie

http://www.debevec.org/Campanile/
Related work

- Image-based modeling
  - Debevec, et al.
    SIGGRAPH 1996
  - Schaffalitzky and Zisserman
    ECCV 2002
  - Brown and Lowe
    3DIM 2005

- Image-based rendering

  Photorealistic IBR:
  - Levoy and Hanrahan, SIGGRAPH 1996
  - Seitz and Dyer, SIGGRAPH 1996
  - Aliaga, et al., SIGGRAPH 2001
  - and many others
Photo Tourism overview

Input photographs → Scene reconstruction → Photo Explorer
Scene reconstruction

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

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

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

Detect features using SIFT [Lowe, IJCV 2004]
Feature matching

Match features between each pair of images
Feature matching

Correspondence estimation

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

Image 1

Image 2

Image 3

Image 4
Structure from motion

minimize $f(R, T, P)$
Incremental structure from motion
Incremental structure from motion
Incremental structure from motion
Reconstruction performance

• For photo sets from the Internet, 20% to 75% of the photos were registered

• Unregistered photos:

• Running time: < 1 hour for 80 photos
  > 1 week for 2600 photos
Photo Tourism overview

Input photographs → Scene reconstruction → Photo Explorer
Photo Explorer
Photo Tourism overview

Input photographs → Scene reconstruction → Photo Explorer

- Navigation
- Rendering
- Annotations

© 2006 Noah Snavely
Object-based browsing
Object-based browsing

- Visibility
- Resolution
- Head-on view
Relation-based browsing

Find all details
Move left
Move right
Zoom in
Zoom out
Find all similar images
Find all zoom outs

Name: 55668857@N00_11...
Added by: 55668857@N00
Date: March 3, 2006, ...
Relation-based browsing

Image A

Image B

© 2006 Noah Snavely
Relation-based browsing

Image A

Image B
Relation-based browsing

Image A

Image B

Image C

to the right of
Relation-based browsing

Image A

to the right of

Image B

Image C
Relation-based browsing

Image A

Image B

to the right of

is detail of

Image D

Image C
Relation-based browsing

Image A

Image B

Image C

Image D

to the right of

is detail of
Relation-based browsing

Image A

is zoom-out of

is detail of

Image B

to the left of

to the right of

is detail of

is zoom-out of

Image C

is detail of

Image D

is zoom-out of

is detail of
Overhead map
Prague Old Town Square
Photo Tourism overview

Input photographs → Scene reconstruction → Photo Explorer

- Navigation
- Rendering
- Annotations
Rendering
Rendering transitions
Rendering transitions
Rendering transitions
Rendering transitions
Photo Tourism overview

Input photographs → Scene reconstruction → Photo Explorer

- Navigation
- Rendering
- Annotations
Annotations
Annotations

Reproduced with permission of Yahoo! Inc. © 2005 by Yahoo! Inc.
YAHOO! and the YAHOO! logo are trademarks of Yahoo! Inc.
Annotations
Microsoft Photosynth Project

- Demo
Limitations / Future work

• Not all photos can be reliably matched
  → Better feature detection / matching
  → Integrating GPS & other localization info.

• Structure from motion scalability
  – ECCV 2008 paper by Li et al.

• Improved navigation, transitions
  – SIGGRAPH 2008
Finding Paths Through the World’s Photos

N. Snavely, R. Garg, S. Seitz, R. Szeliski

SIGGRAPH 2008
Advances over Photo Tourism

• Better view interpolation for landmarks that are not well approximated by planar facades

• Fluid, freeform 3D navigation
  – Automatic discovery of scene-specific controls: orbits, panoramas
  – Path planning for transitions between different viewpoints
System components

- Input photographs, structure from motion
- Viewpoint scoring function
  - Estimates the quality of rendering a particular input image from a given viewpoint
- Navigation controls
  - Automatically discover orbits, panoramas, representative views
  - Path planning
- Rendering engine
  - For any user position, find best scoring input image and reproject it from the current viewpoint
  - Orbit stabilization, appearance stabilization
Viewpoint Scoring Function

• Ideally, measure the difference between the reprojected view and a real photo of the scene captured at a given location

• Quality criteria:
  – Angular deviation
  – Field of view
  – Resolution
Scene-Specific Navigation Controls

- Orbits
  - An orbit is a distribution of views in a circle converging toward the same object
  - An orbit is specified by an axis and an image
  - Scoring possible orbits: quality, length, convergence, object-centeredness
Scene-Specific Navigation Controls

• Orbits

  – An orbit is a distribution of views in a circle converging toward the same object
  – An orbit is specified by an axis and an image
  – Scoring possible orbits: quality, length, convergence, object-centeredness
  – Orbit enumeration: first find likely axes, then evaluate orbit scoring function for every image and that axis
Detected orbits

Statue of Liberty

Pantheon
Scene-Specific Navigation Controls

• Orbits

• Panoramas
  – A set of images taken from approximately the same viewpoint and covering a good range of viewing directions

• Canonical views
  – Scene summarization algorithm of Simon et al. [ICCV 2007]
  – Choose canonical view for each orbit and panorama, show as thumbnail in scene viewer
Path Planning for Transitions

• Automatically compute path from one image to another, show multiple reprojected images between the two endpoints (unlike Photo Tourism)

• Discrete solution: find shortest path in a transition graph defined by input views
  – Transition cost between two images: computed based on viewpoint scoring function

• The discrete shortest path can be improved by smoothing
Path Planning Example
Orbit Stabilization

- Images are reprojected using a “proxy plane”
- Planes fit to scene points can cause jerky motion, especially for orbits
- For orbit stabilization, choose a proxy plane that passes through the orbit axis and is (roughly) parallel to the source image plane
Appearance Stabilization

- Compute distance between geometrically and photometrically aligned images
- Similarity mode: favor transitions between similar images
- Color compensation: estimate gain and offset between two images, adjust next image to match the previously viewed one
Appearance Stabilization

All images

Similar images

Color compensation
Modeling and Recognition of Landmark Image Collections Using Iconic Scene Graphs

Xiaowei Li, Changchang Wu, Christopher Zach, Svetlana Lazebnik, Jan-Michael Frahm
ECCV 2008
Overview

All images

Appearance-based clustering, geometric verification

Iconic images

Pairwise matching of iconic images

Reconstructed components

SFM

Components of iconic scene graph

Graph cut

Iconic scene graph
The approach

1. Appearance-based clustering
   • Goal: find groups of images with roughly similar viewpoints and scene conditions
The approach

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

1. Appearance-based clustering
2. Geometric verification and iconic image selection

• Perform feature-based geometric matching between a few “top” images from each cluster
• QDEGSAC (Frahm & Pollefeys, 2006) for robust estimation of fundamental matrix or homography
• Select an *iconic image* for each cluster as the image with the most total inliers
The approach

1. Appearance-based clustering
2. Geometric verification and iconic image selection
The approach

1. Appearance-based clustering
2. Geometric verification and iconic image selection
3. Construction of iconic scene graph
   - Perform geometric matching between every pair of iconic images
   - Create a weighted edge for every pair related by a homography or a fundamental matrix
The approach

1. Appearance-based clustering
2. Geometric verification and iconic image selection
3. Construction of iconic scene graph
The approach

1. Appearance-based clustering
2. Geometric verification and iconic image selection
3. Construction of iconic scene graph
4. Finding graph components
   - Use tags to eliminate semantically irrelevant isolated nodes
   - Run normalized cuts to break up the rest of the graph into smaller tightly connected sub-graphs
The approach

1. Appearance-based clustering
2. Geometric verification and iconic image selection
3. Construction of iconic scene graph
4. Finding graph components
The approach

1. Appearance-based clustering
2. Geometric verification and iconic image selection
3. Construction of iconic scene graph
4. Finding graph components
5. Structure from motion
   - Perform SFM separately on each component
   - Maximum-weight spanning tree determines the order of incorporating images into the 3D model
   - If possible, merge component models using geometric relationships along edges that were originally cut
   - Register additional non-iconic images to the models (optional)
Hierarchical 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 dataset
45284 images

Las Vegas

New York

Tokyo

Registered images in largest model: 871
Points visible in 3+ views: 18675
Notre Dame dataset
10840 images

105 iconic images
Notre Dame dataset
10840 images

Registered images in largest model: 337
Points visible in 3+ views: 30802
San Marco dataset

43557 images

134 iconic images

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