Exploring Images in 3D

Visual Recognition and Search

Maysam Moussalem

Outline

- Photo Tourism: Exploring Photo Collections in 3D
- Automatic Popup
- Single-View Metrology

Outline

- Photo Tourism: Exploring Photo Collections in 3D
- Automatic Popup
- Single-View Metrology

Photo Tourism:
Exploring Photo Collections in 3D

Noah Snavely

Steven M. Seitz

University of Washington

Richard Szeliski

Microsoft Research

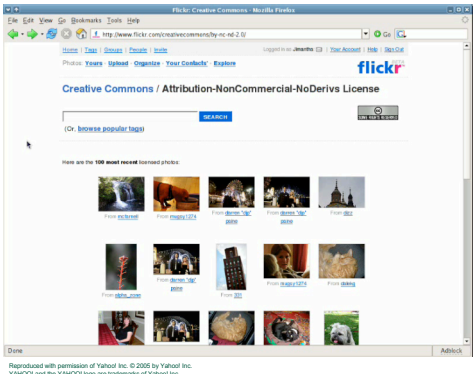


Photo Tourism overview

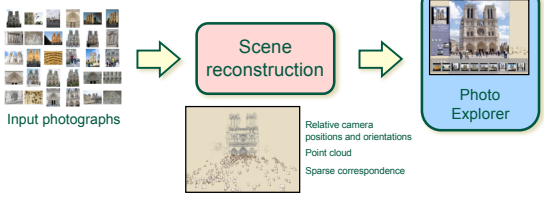
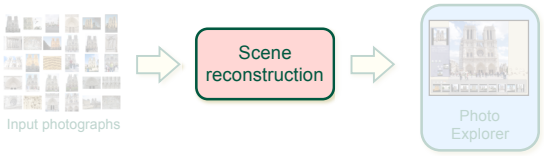
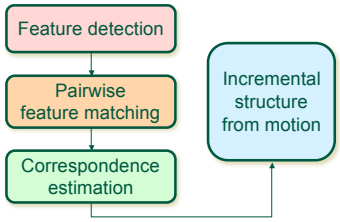


Photo Tourism overview



Scene reconstruction

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

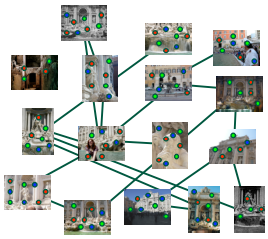


Feature detection

Detect features using SIFT [Lowe, IJCV 2004]

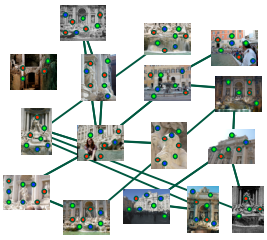
Feature matching

Match features between each pair of images



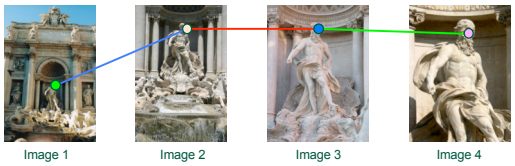
Feature matching

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

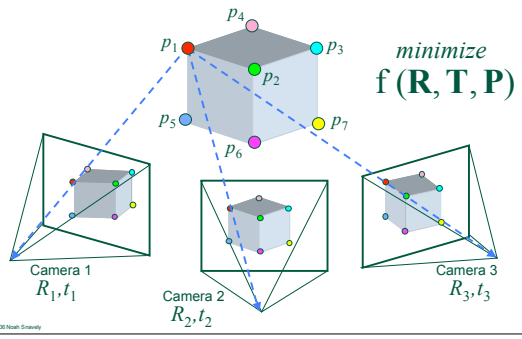


Correspondence estimation

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



Structure from motion



minimize
$f(\mathbf{R}, \mathbf{T}, \mathbf{P})$

Incremental structure from motion

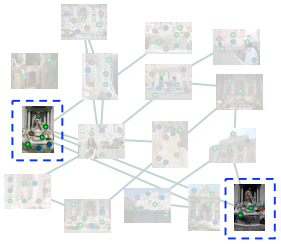


Photo Tourism overview

Input photographs

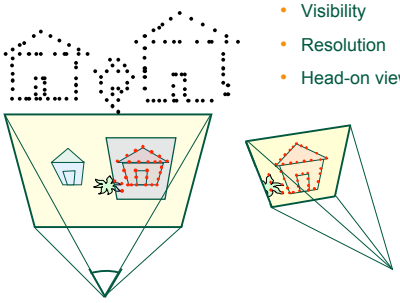
Scene reconstruction

Photo Explorer

- Navigation
- Rendering
- Annotations

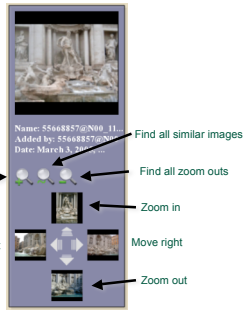
Object-based browsing

Object-based browsing

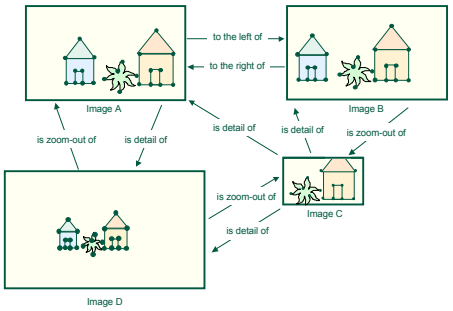


- Visibility
- Resolution
- Head-on view

Relation-based browsing



Relation-based browsing



Overhead map

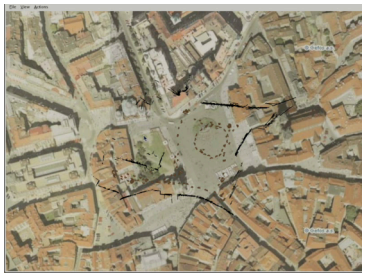


Photo Tourism overview

Input photographs

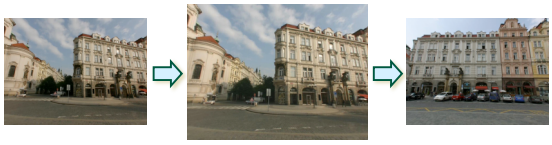
Scene reconstruction

Photo Explorer

- Navigation
- Rendering
- Annotations

Rendering

Rendering transitions



Rendering transitions

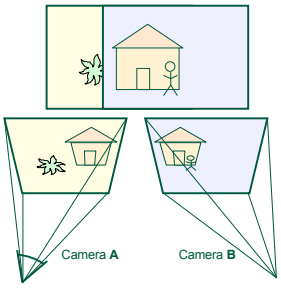


Photo Tourism overview

- Navigation
- Rendering
- Annotations

Annotations

Contributions

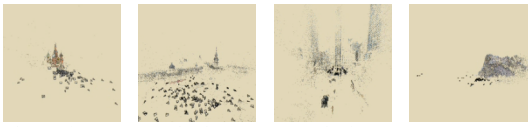
- Automated system for registering photo collections in 3D for interactive exploration
- Structure from motion algorithm demonstrated on hundreds of photos from the Internet
- Photo exploration system combining new image-based rendering and photo navigation techniques

Conclusion

Indexing everyone's photos provides a new way to share and experience our world

To find out more:

- http://phototour.cs.washington.edu
- http://research.microsoft.com/IVM/PhotoTourism
- http://labs.live.com/photosynth



Outline

- Photo Tourism: Exploring Photo Collections in 3D
- Automatic Popup
- Single-View Metrology

Automatic Popup (Hoim, Efros & Hebert)

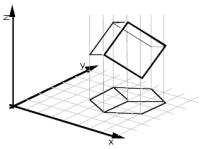
- Method for creating virtual walkthroughs
- Completely automatic
- Requires only a single photograph as input

The World Behind the Image

31

The Problem

- Recovery of 3D geometry from single 2D projection
  - Infinite number of possible solutions!

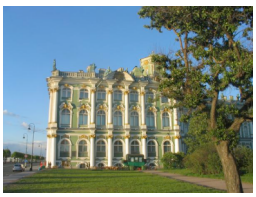


from [Sinha and Adelson 1993]

32

Our World is Structured

Abstract World



Our World

Image Credit (left): F. Cunin and M.J. Sailor, UCSD

33

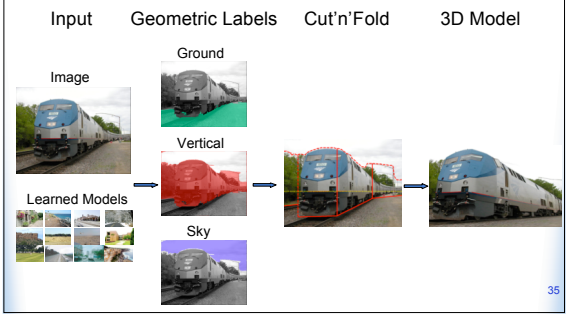
Goals of this Approach

- Simple, piecewise planar models
- Outdoor scenes
- Doesn't need to work all the time (~35%)

Pop-up Book

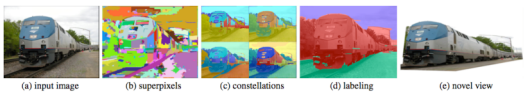
34

Overview



35

Overview of method



- Image to superpixels
- Superpixels to multiple constellations
- Multiple constellations to superpixel labels
- Superpixel labels to 3D models

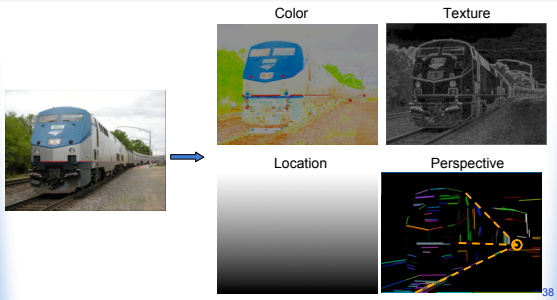
36

Superpixels

- Initially, image represented as 2D array of RGB pixels
- Form superpixels
  - Small, nearly-uniform regions of image
- Improves computational efficiency

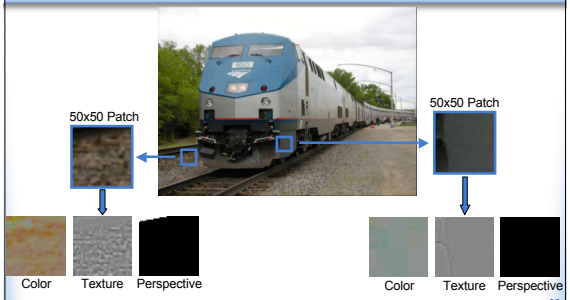
37

Geometric Cues



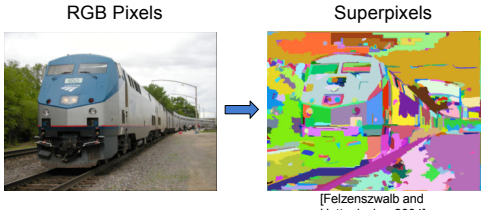
38

Spatial Support



39

Spatial Support



[Felzenszwalb and Huttenlocher 2004]

40

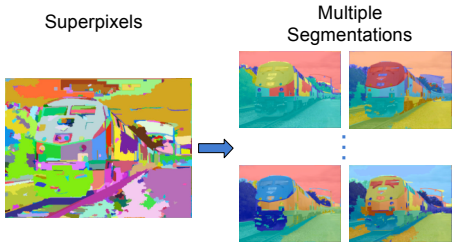
Superpixels to Multiple Constellations

- Group superpixels that are likely to share common geometric model into constellations
  - Assign one randomly selected superpixel to each constellation
  - Iteratively assign remaining superpixels to constellations
    - Maximize average pair-wise log-likelihoods with other superpixels in constellation

$$S(\mathbf{C}) = \sum_{k}^{N_c} \frac{1}{n_k(1-n_k)} \sum_{i,j \in C_k} \log P(y_i = y_j | \mathbf{z}_i - \mathbf{z}_j)$$

41

Multiple Segmentations



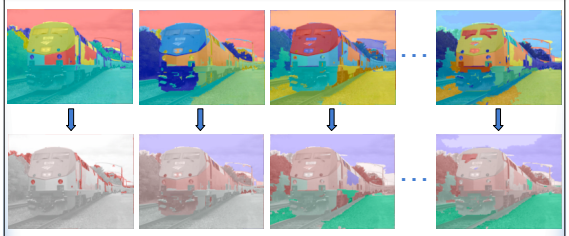
42

Training Procedure

- For each training image
  - Compute superpixels
  - Compute superpixel features
- Estimate pairwise-likelihood function
- For each training image
  - Form multiple sets of constellations for varying N_c
  - Label each constellation according to superpixel ground truth
  - Compute constellation features
- Estimate constellation label and homogeneity likelihood functions

43

Labeling Segments



For each segment:

- Get $P(y_j = v|\mathbf{x}, \mathbf{h}_{ji})P(\mathbf{h}_{ji}|\mathbf{x})$

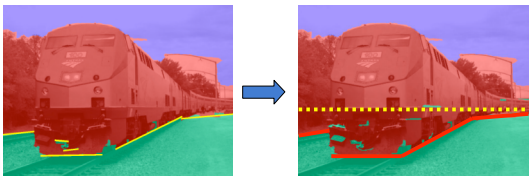
44

3D Model

- Need to determine
  - camera parameters
  - where each vertical region intersects the ground

45

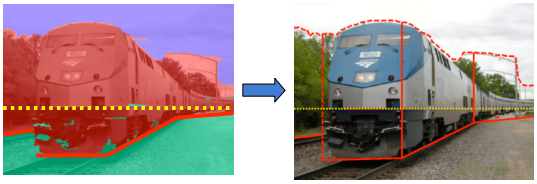
Cutting and Folding



- Form polylines from boundary segments
  - Join segments that intersect at slight angles
  - Remove small overlapping polylines
- Estimate horizon position from perspective cues

46

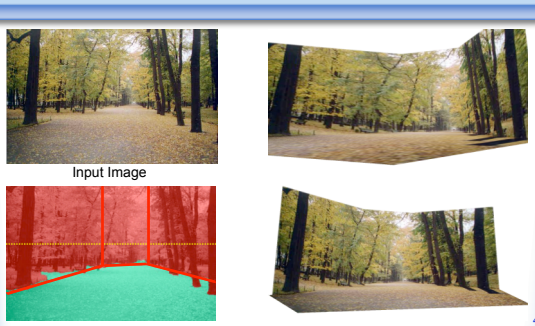
Cutting and Folding



- "Fold" along polylines and at corners
- "Cut" at ends of polylines and along vertical-sky boundary

47

Results



Input Image

Cut and Fold

Automatic Photo Pop-up

48

Results




Input Image

Automatic Photo Pop-up

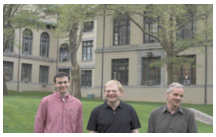
Failures

Labeling Errors



Failures

Foreground Objects




Conclusion

- First system to automatically recover 3D scene from single image!
- *Learn* statistics of our world from training images

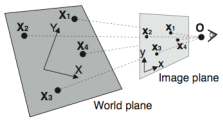

Outline

- Photo Tourism: Exploring Photo Collections in 3D
- Automatic Popup
- Single-View Metrology

Single-View Metrology (Criminisi)

- Simple and effective algorithms for
  - Extracting geometric information
    - Lengths of segments on planar surfaces
    - Distances of points from planes
  - Constructing 3D model from single perspective of a scene

(1) Planar Measurements

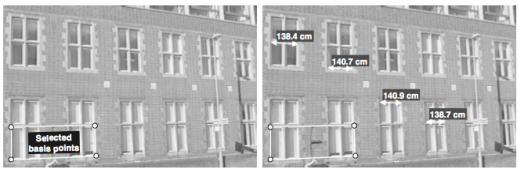


Algorithm 1: planar measurements.

1. Given an image of a planar surface estimate the image-to-world homography matrix $\mathbf{H}$;
2. Repeat
   (a) Select two points $\mathbf{x}_1$ and $\mathbf{x}_2$ on the image plane;
   (b) Back-project each image point into the world plane via (1) to obtain the two world points $\mathbf{X}_1$ and $\mathbf{X}_2$;
   (c) Compute the Euclidean distance $d(\mathbf{X}_1, \mathbf{X}_2)$.

55

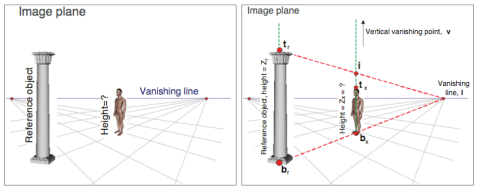
(1) Planar Measurements



56

(2) Measuring Distances from Planes

- Aim = compute height of an object (ex. man in figure) relative to reference (ex. column)

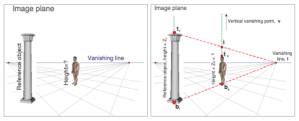


57

(2) Measuring Distances from Planes

Algorithm 2: computing heights of objects in single views.

1. Estimate the vanishing point $\mathbf{v}$ for the vertical direction;
2. Estimate the vanishing line $\mathbf{l}$ of the reference plane;
3. Select top and base points of the reference segment (points $\mathbf{t}_r$ and $\mathbf{b}_r$, respectively);
4. Compute the metric factor α by applying: $\alpha = -\frac{||\mathbf{b}_r \times \mathbf{t}_r||}{Z_r(\mathbf{l}\cdot\mathbf{b}_r)||\mathbf{v}\times\mathbf{t}_r||}$;
5. Repeat
   (a) Select top and base of the object to measure (points $\mathbf{t}_x$ and $\mathbf{b}_x$, respectively);
   (b) Compute the height Z_x by applying: $Z_x = -\frac{||\mathbf{b}_x \times \mathbf{t}_x||}{\alpha(\mathbf{l}\cdot\mathbf{b}_x)||\mathbf{v}\times\mathbf{t}_x||}$;



58

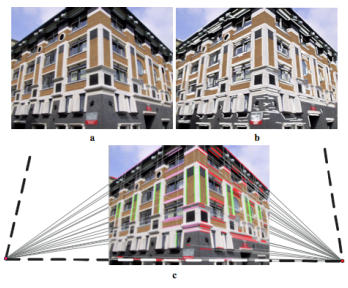
Estimating Vanishing Points and Lines

- Can be computed from
  - Image-to-world homography H (if known)
  - Automatic/semi-automatic techniques that work directly on image plane
    - Example: simple RANSAC-based algorithm

1. Automatic Canny edge detection and straight line fitting to obtain the set of straight edge segments $\mathcal{E}$ (fig. 5b) [4];
2. Repeat
   (a) Randomly select two segments $\mathbf{s}_1, \mathbf{s}_2 \in \mathcal{E}$ and intersect them to give the point $\mathbf{p}$;
   (b) The support set $\mathcal{S}_p$ is the set of straight edges in $\mathcal{E}$ going through the point $\mathbf{p}$;
3. Set the dominant vanishing point as the point $\mathbf{p}$ with the largest support $\mathcal{S}_p$;
4. Remove all edges in $\mathcal{S}_p$ from $\mathcal{E}$ and goto 2 for the computation of the next vanishing point

59

Estimating Vanishing Points and Lines



60

(3) Constructing 3D Models

- Segmentation of scene objects
  - Achieved by interactive silhouette cut-out
  - Normalized cross-correlation
    - Countour thought of as 1-dimensional curve separating 2 dissimilar regions (texture, color, …)

61

(3) Constructing 3D Models

- Filling of occluded areas in an "undetectable" way
- Can be done by
  - Exploiting symmetries and patterns
  - Non-parametric texture synthesis

62

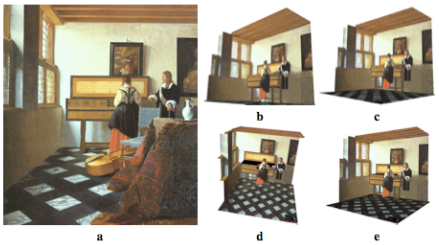
Complete Reconstruction

The complete algorithm

1) Select reference plane and estimate H
2) Select reference height and compute metric factor
3) Repeat
   1) Segment object and measure its height and position
   2) Fill in areas occluded by selected objects
   3) Insert selected object in output 3D model

63

Example



The Music Lesson, by Vermeer

64

Outline

- Photo Tourism: Exploring Photo Collections in 3D
- Automatic Popup
- Single-View Metrology

65