3D Manipulation of 2D Images*

Alexandru Vasile (alexv@ll.mit.edu) & Peter Cho (cho@ll.mit.edu)

MIT Lincoln Laboratory

MIT IAP Course
January 2012

Class 4 notes

*This work was sponsored by the Department of the Air Force under Air Force Contract FA8721-05-C-0002. Opinions, interpretations, conclusions and recommendations are those of the authors and are not necessarily endorsed by the United States Government.

Course Outline

- **Session 1: Single-view geometry overview**
  - Pinhole camera model
  - Projective geometry
  - Camera calibration
- **Session 2: Automatic image feature matching**
  - SIFT algorithm
  - Multiple image matching
  - Image search
- **Session 3: Panorama formation**
  - Homographies
  - Panorama parameter estimation
  - 2D vs 3D mosaic generation
- **Session 4: 3D reconstruction**
  - Structure from Motion
  - 3D geometry from multiple 2D views
  - Photo tourism

Path to 3D Reconstruction

- Obtain reliable image matches and 2 view relations (F matrix). → DONE.

- Compute initial structure and motion (camera poses).

- Refine structure and motion for the two image pair.
  - Bundle adjustment algorithm.

- Keep adding additional images with reliable matches / view relations to the two image reconstruction.

Path to 3D Reconstruction

- Obtain reliable image matches and 2 view relations (F matrix). → DONE.

- Compute initial structure and motion (camera poses).

Structure from Motion

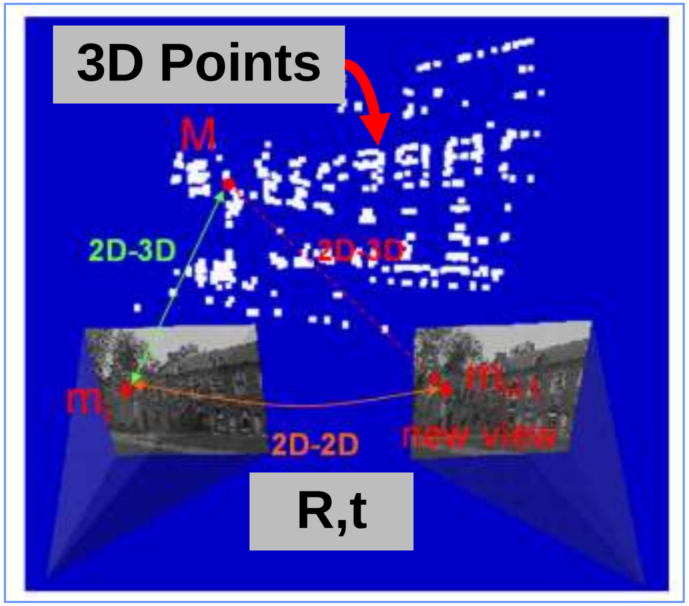
- Refine structure and motion for the two image pair.
  - Bundle adjustment algorithm.

(SfM)

- Keep adding additional images with reliable matches / view relations to the two image reconstruction.

Compute Initial Structure and Motion

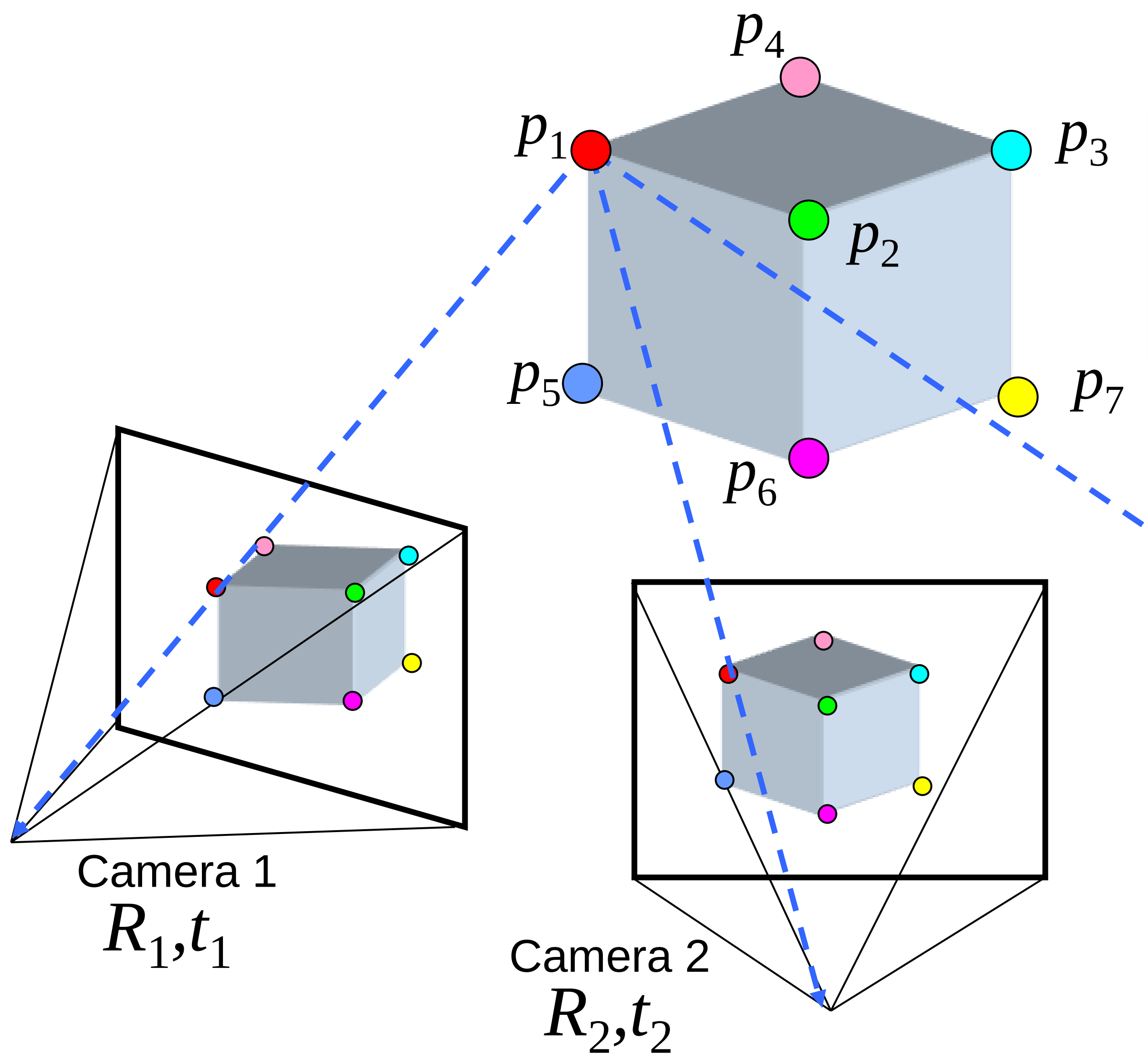
- **Last class, showed how F matrix gives us reliable matches between images.**
- **F also gets us initial camera motion.**
  - **Given good guess of intrinsic matrix K, can get estimate of camera pose (R,t) between the two images.**



- **Get initial 3D reconstruction.**
  - **Have initial 2D matches, camera poses.**
  - **Do triangulation to get 3D points (structure).**
  - **Now have initial 2D matches, 3D points, camera poses.**

- **Still have some errors in initial 3D points and camera pose.**
  - **Refine structure and motion.**

Refine Structure and Motion (Bundle Adjustment)



minimize

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

Putting All Together: An SfM Reconstruction Pipeline

Feature detection

Feature matching
Find scene points seen by multiple cameras

Initialization
Robustly estimate camera poses and scene points

Bundle adjustment
Refine camera poses **R, T** and scene structure **P**

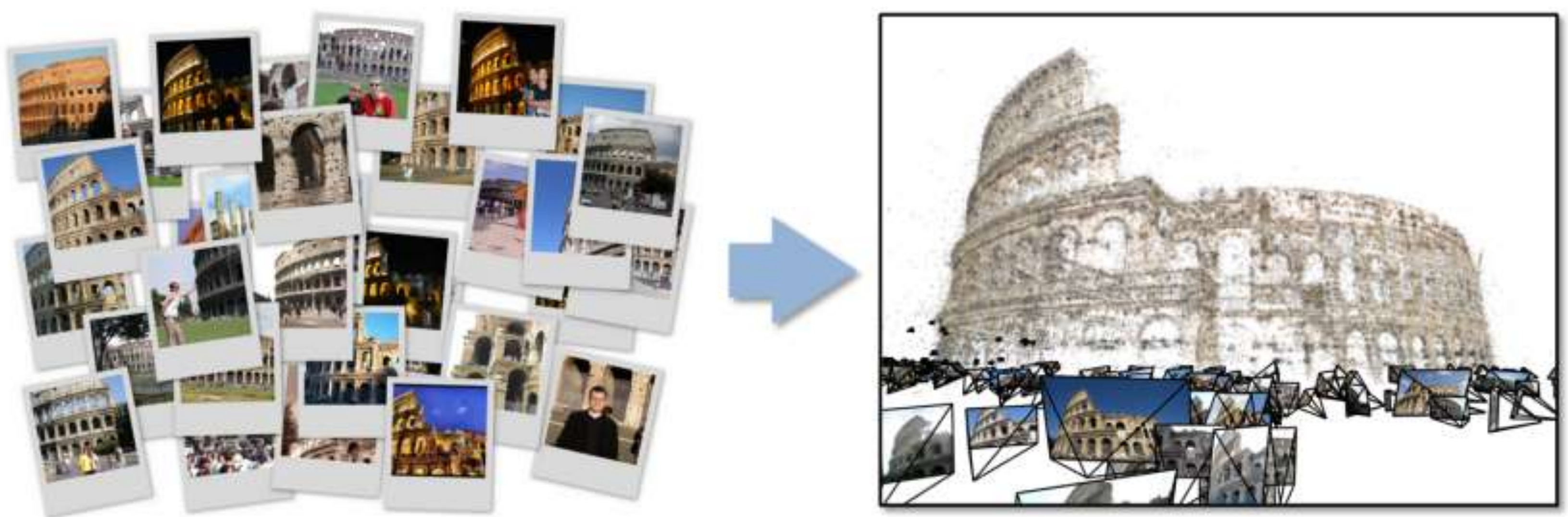
3D Reconstruction

Bundle Adjustment (BA)
Start with seed model
- Run bundle adjustment
- Remove outliers
- Add another image
- Repeat until no more images can be added

[Snavely06], [Li08], [Agarwal09], [Frahm10] …

Example: SfM from Internet Images

- Recent work has built 3D models from large, unstructured online image collections
  - [Snavely06], [Li08], [Agarwal09], [Frahm10], Microsoft's PhotoSynth, …

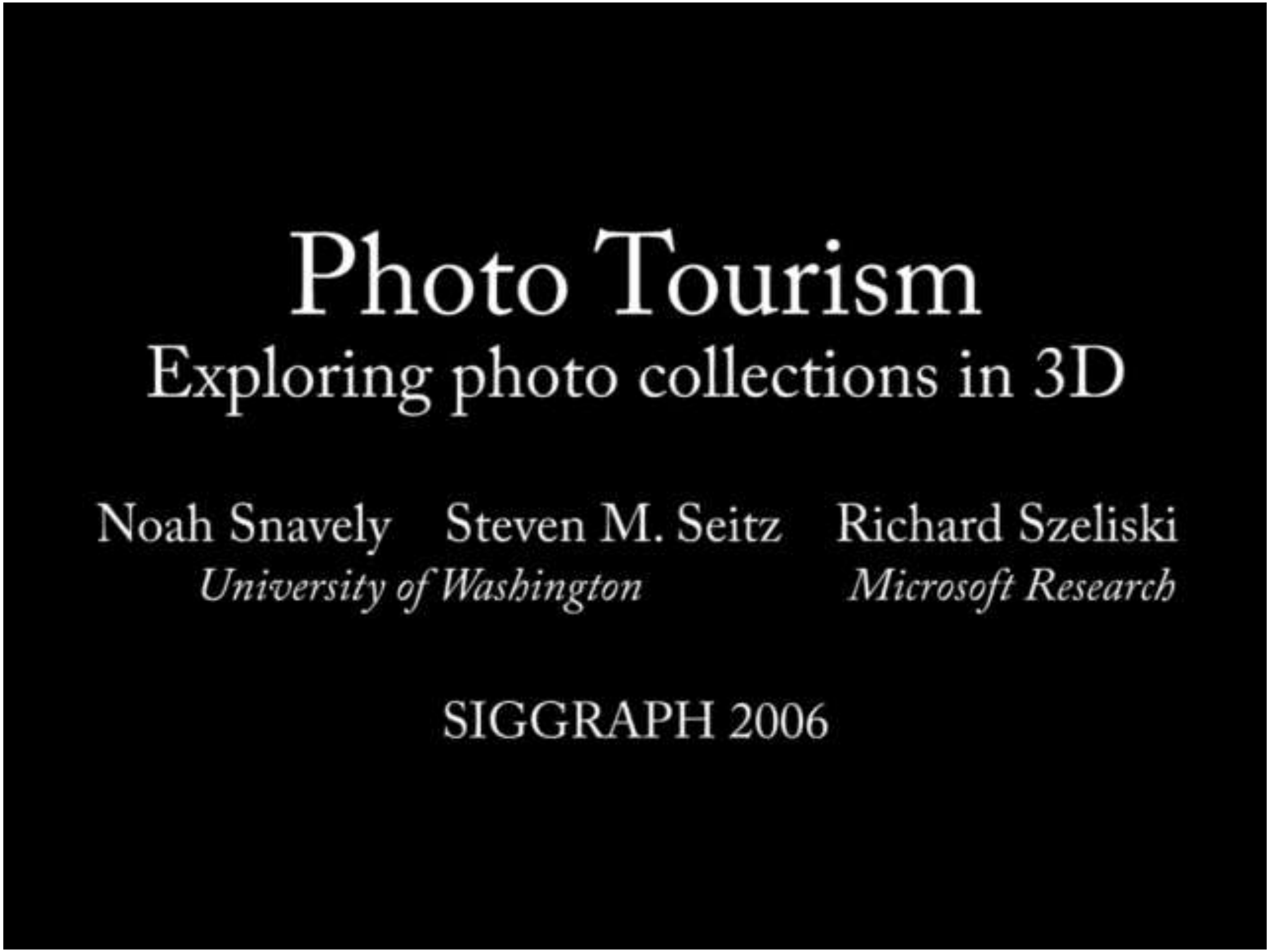


- SfM is a key part of these reconstruction pipelines

Example: SfM from Internet Images Colosseum, Rome

Example: SfM from Internet Images
Piazza San Marco, Venice, Italy

Photo-Tourism



Example: Semi-Cooperative Urban Photo Collection

- **MIT campus chosen as a representative "small city"**
- **30K+ digital ground photos shot in July 2009 during 5 fieldtrips each lasting for a few hours**
  - **Vast majority of pictures taken outdoors during sunny weather conditions**
  - **Qualitative effort made to densely gather imagery from multiple vantage points**

LL adventurers shooting digital photos outside MIT's media lab

Example: Semi-Cooperative Urban Photo Collection

- **MIT campus chosen as a representative "small city"**
- **30K+ digital ground photos shot in July 2009 during 5 fieldtrips each lasting for a few hours**
  - **Vast majority of pictures taken outdoors during sunny weather conditions**
  - **Qualitative effort made to densely gather imagery from multiple vantage points**
  - **Photographing proceeded faster as pointing grew less careful over time**

Doubling data collection rate…

Example: Semi-Cooperative Urban Photo Collection

- **MIT campus chosen as a representative "small city"**

- **30K+ digital ground photos shot in July 2009 during 5 fieldtrips each lasting for a few hours**
  - **Vast majority of pictures taken outdoors during sunny weather conditions**
  - **Qualitative effort made to densely gather imagery from multiple vantage points**
  - **Photographing proceeded faster as pointing grew less careful over time**

- **Approximately 2000 photos shot indoors**

Photographing interior of MIT's health center

Urban Photos Shot Around MIT


DSC_0009-1.rd.jpg
729.3 KB


DSC_0018-1.rd.jpg
845.9 KB


DSC_0064-1.rd.jpg
806.3 KB


DSC_0232.rd.jpg
713.9 KB


DSC_5703.rd.jpg
711.8 KB


DSC_5830.rd.jpg
917.0 KB


DSC_5865.rd.jpg
780.2 KB


DSC_5958.rd.jpg
825.2 KB


DSC_5988.rd.jpg
678.0 KB


DSCF9179.rd.jpg
1.4 MB

DSCF9198.rd.jpg
806.2 KB


IMG_0381.rd.jpg
856.9 KB


IMG_0434.rd.jpg
807.0 KB


IMG_0500-1.rd.jpg
870.4 KB


IMG_0523.rd.jpg
800.4 KB


IMG_0635.rd.jpg
574.7 KB


IMG_0642-1.rd.jpg
1.0 MB


IMG_0646.rd.jpg
525.3 KB


IMG_0654.rd.jpg
521.0 KB


IMG_0680-1.rd.jpg
813.6 KB


IMG_0688-1.rd.jpg
963.7 KB


IMG_0725.rd.jpg
677.2 KB


IMG_0730-1.rd.jpg
1.5 MB


IMG_0733-1.rd.jpg
1.1 MB


IMG_0747.rd.jpg
521.1 KB


IMG_0786-1.rd.jpg
828.0 KB


IMG_0822.rd.jpg
455.3 KB


IMG_0888.rd.jpg
549.8 KB


IMG_0918-1.rd.jpg
877.7 KB


IMG_0939.rd.jpg
524.8 KB

Structured Output from Unstructured Input

Lowe, “Distinctive Image Features from Scale-Invariant Keypoints”, 2004

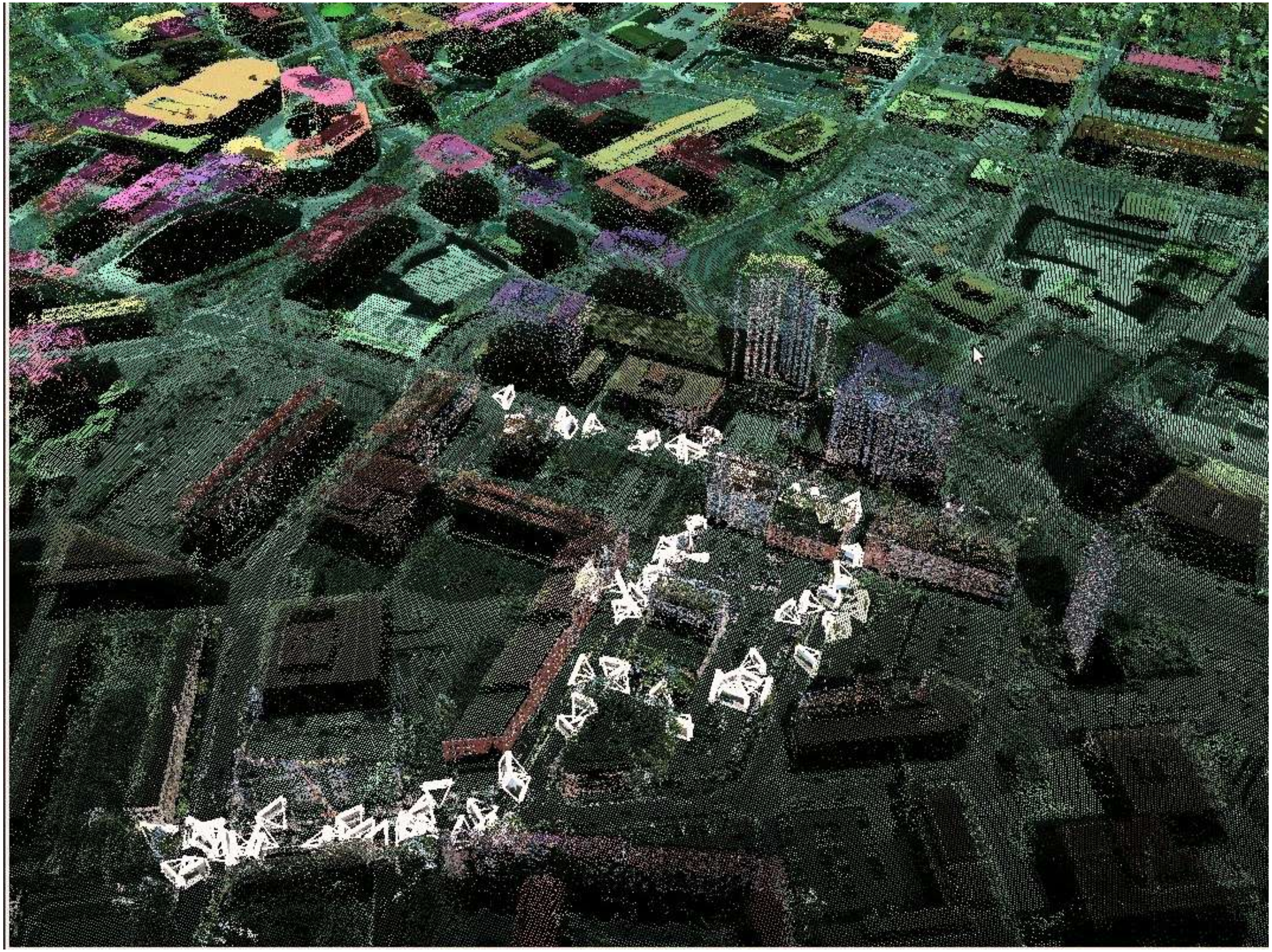
Snavely, Seitz & Szeliski, “Photo tourism: Exploring Photo Collections in 3D”, 2006

Agarwal et al, “Building Rome in a Day”, 2009

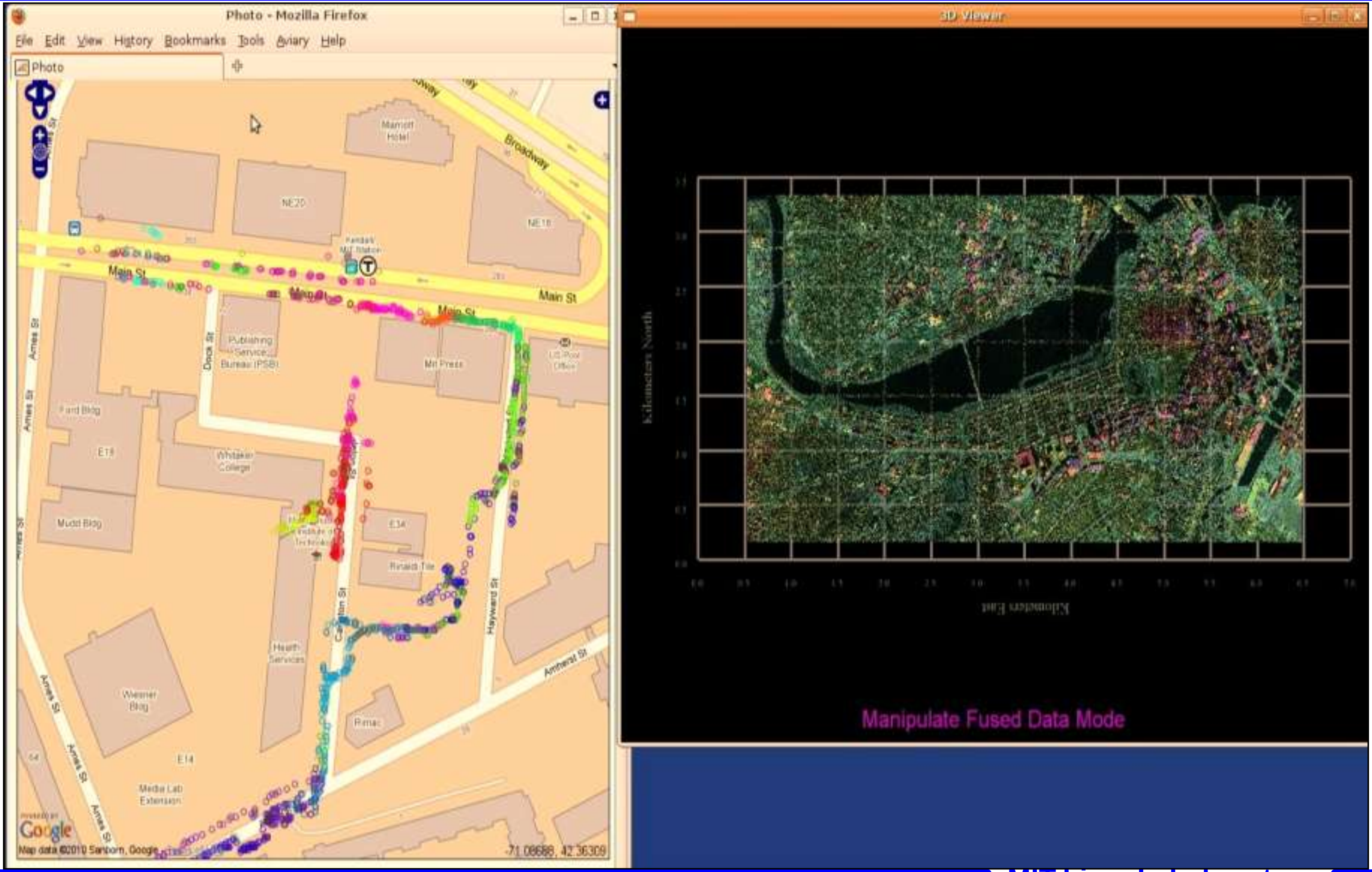
East MIT Campus Reconstruction

Relative position & poses for 2317 photos automatically determined

East MIT Campus Reconstruction Aligned to 3D Lidar Map



Google Map Display



Web browser interface developed by Jennifer Drexler

Hand-Launched Glider Setup

- **Synchronize camera & GPS clocks by taking picture of latter with former**

- **Mount camera & GPS to glider's underside prior to hand launching**
  - **Electric motor assists initial glider ascent**
  - **Pilot finds & rides air thermals to remain aloft for hours**

- **Gather imagery (3 Hz) & GPS readings (1 Hz) over 20 minute aerial missions that fly up to 430 meters above ground**

Radian sailplane glider (< $400)

Canon powershot camera (< $300)

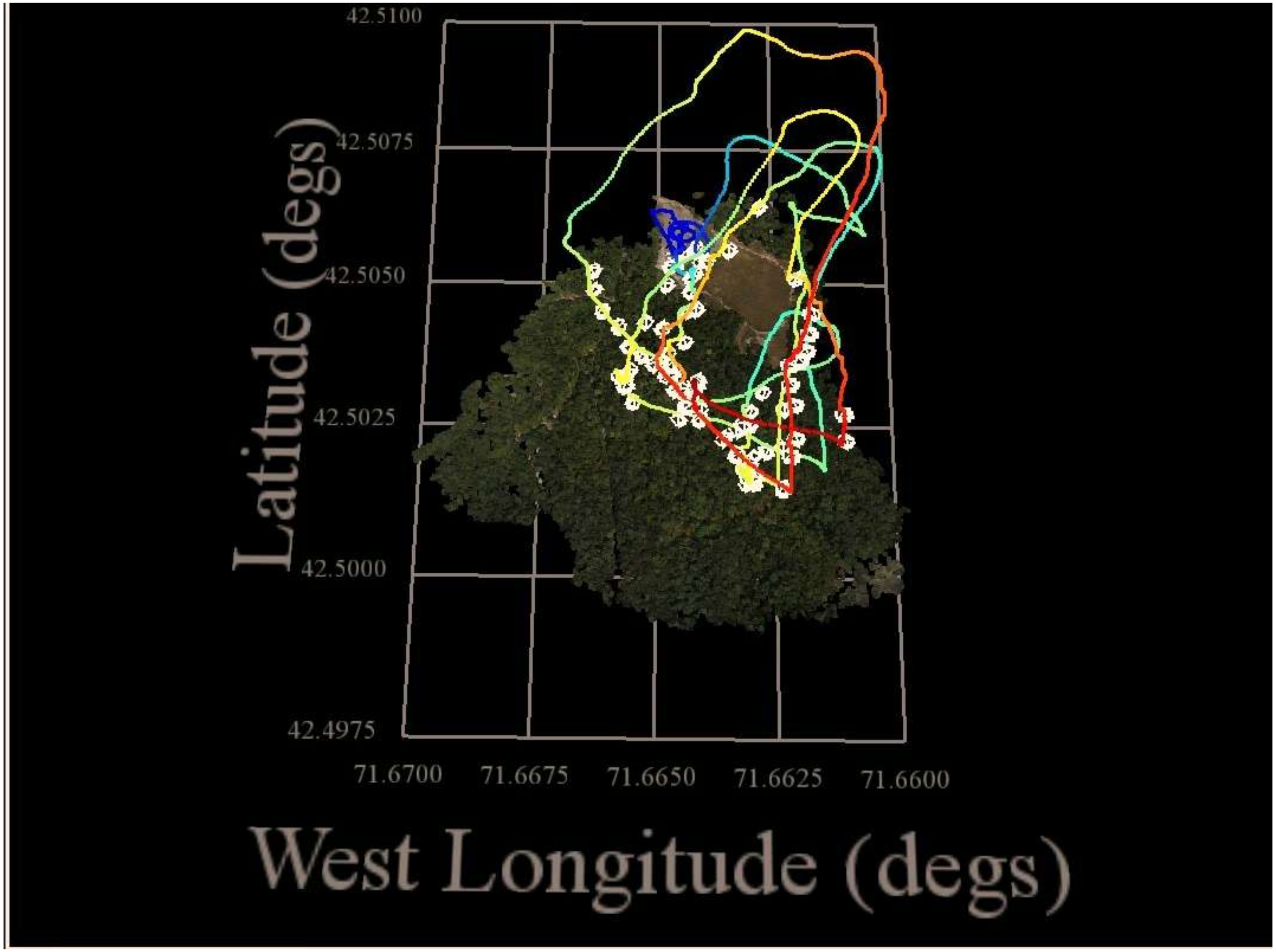
Garmin GPS (< $100)

Raw Video Imagery from Aerial Glider

3000 video frames collected during sailplane flight #3 on 22 Sep 2010

Camera & Dense Point Cloud Reconstruction vs GPS track

74 of 1464 reconstructed video frames displayed as 3D frusta vs glider's measured GPS track



Texas-Tech Lubbock Campus 3D Reconstruction

Visual SfM

- **Many tools available on the internet to do your own 3D reconstructions from your very own images.**
  - **All attempt to reconstruct static background from images. (moving cars, people get erased Soviet-style).**

- **Best one out there right now is VisualSFM (GUI based).**
  - **Uses Prof. Snavely's Bundler code (Sparse 3D reconstruction) + Prof. Furukawa's PMVS (Dense 3D reconstruction).**

LIVE
DEMO !

Lab 4
Do Your Own 3D Reconstructions!

- **Install Visual SfM + dependencies …**
  - **http://www.cs.washington.edu/homes/ccwu/vsfm/**
- **Full Win/Mac/Linux Install Instructions:**
  - **http://www.cs.washington.edu/homes/ccwu/vsfm/install.html**

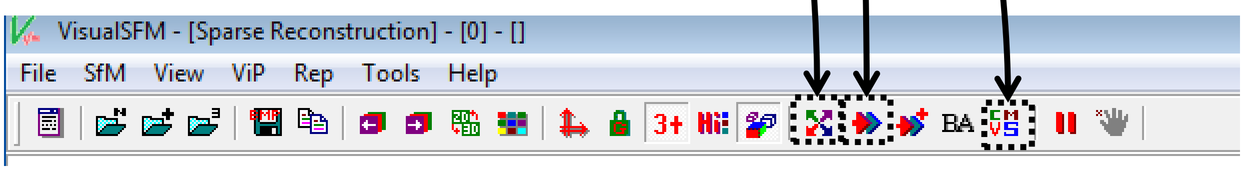
…

- **Or grab directly Win32/64 packaged code from here:**

  http://web.mit.edu/alexv/Public/IAP_2012/class_04/Lab04/VisualSFM_win32+ALL.zip
  http://web.mit.edu/alexv/Public/IAP_2012/class_04/Lab04/VisualSFM_win64+ALL.zip
- **Compressed Win32/64 Instructions:**
  - **Unzip VisualSFM to some folder. Try launch VisualSFM.exe, if it fails, you need to install MSVC2010 runtime (Run vcredist_*.exe located in same directory)**

Lab 4
Do your own 3D Reconstructions!

- **Learn how to use Visual SFM:**
  - See 2 min instructional video at http://www.cs.washington.edu/homes/ccwu/vsfm/
  - Five Steps:
    1. Load images (File → Open + Multi Images)
       Go to dir with images, do shortcut CTRL-A to grab all images.
       Should see the images displayed on main GUI window.
    2. Do SIFT + SIFT Match (Button press)
    3. Do sparse reconstruction (Button press)
    4. Do dense reconstruction (Button press)
       Input name to save reconstruction (eg. Kermit)
       3D data will be saved to kermit.ply in same dir.
    5. Go to View → Dense 3D Reconstructions



- **Grab a test data set and check your installation:**
  - http://web.mit.edu/alexv/Public/IAP_2012/class_04/Lab04/data_sets.zip
  - Two small test data sets are available: Kermit and E.T.

Lab 4
Do your own 3D Reconstructions!

Suggestions for doing/viewing your own reconstructions:

- **Take a bunch of pictures of around a target of interest from different viewpoints, moving around the object. Try to keep the object in the center of camera.**

- **Avoid shiny objects, objects with lots of repeated textures, symmetry.**

- **SfM reconstructions can take a while to process. Start out small, with 10-15 images, at no more than 1 Mpixel resolution (can use IrfanView to batch downsample your images).**

- **Can use Meshlab to better manipulate the created .PLY 3D data files.**
  - **http://meshlab.sourceforge.net/**

References

- Richard I. Hartley and Andrew Zisserman. Multiple View Geometry. Cambridge University Press, Cambridge, UK, 2004
- Sameer Agarwal, Noah Snavely, Ian Simon, Steven M. Seitz, Richard Szeliski: Building Rome in a Day. ICCV 2009
  - http://grail.cs.washington.edu/projects/rome/
- Yasutaka Furukawa and Jean Ponce: Patch-based Multi-View Stereo Software.
  - http://grail.cs.washington.edu/software/pmvs/
- M.I. A. Lourakis and A.A. Argyros, BA: A Software Package for Generic Sparse Bundle Adjustment.
  - http://www.ics.forth.gr/~lourakis/sba/
- Changchang Wu, VisualSFM : A Visual Structure from Motion System.
  - http://www.cs.washington.edu/homes/ccwu/vsfm/index.html