Visualization Methods for Computer Vision Analysis

IDAV Institute for Data Analysis and Visualization

Mauricio Hess-Flores[1], Shawn Recker[1], Kenneth I. Joy[1] and Mark A. Duchaineau[2]

[1]Institute for Data Analysis and Visualization, University of California, Davis, USA,
{strecker, mhessf}@ucdavis.edu, kenneth.i.joy@gmail.com
[2]Currently at Google, Inc., duchaineau@cognigraph.com

Fifth International Conferences on Pervasive Patterns and Applications (PATTERNS 2013)
Valencia, Spain
May 27 – June 1, 2013

Outline

- Motivation
- Applications
  - Visualization of scene structure uncertainty
  - Feature tracking summaries
- Conclusions and future work

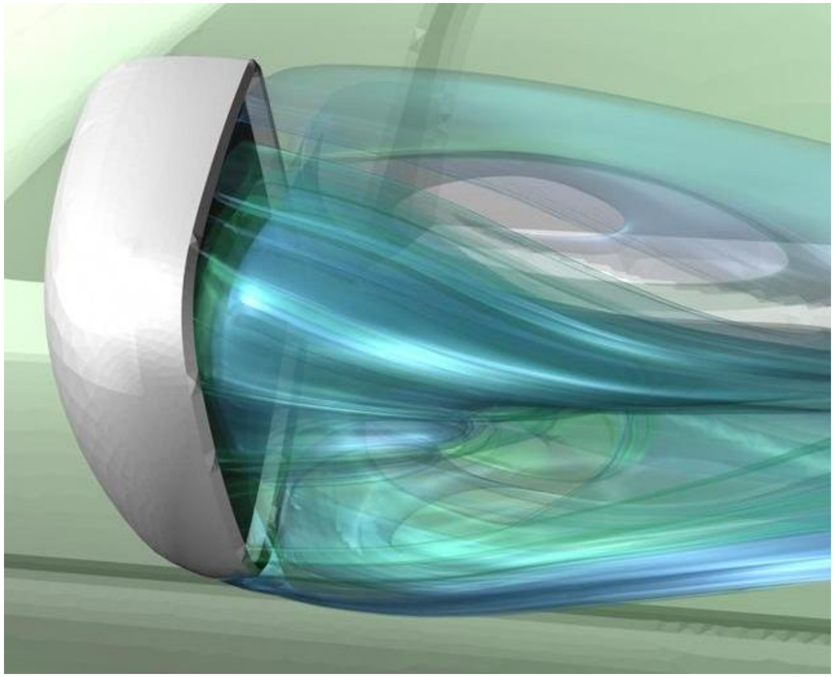
Motivation

- **Computer vision**:
  - A field that includes methods for acquiring, processing, analyzing, and understanding images in order to produce numerical or symbolic information.
  - Common applications: scene reconstruction, navigation, visual surveillance, object detection, recognition and tracking, human-computer interaction, automatic inspection

Motivation

Scientific visualization:

- A field which aims to provide renderings of volumes and surfaces, including those that are time-dependent, to graphically illustrate scientific data for its understanding.



Flow about a BMW mirror can be seen through the use of transparent integral surfaces, surfaces that follow the flow through time

Christoph Garth, Han Krishnan, Xavier Tricoche, Tom Tricoche, Kenneth I. Joy, "Generation of Accurate Integral Surfaces in Time-Dependent Vector Fields," IEEE Transactions on Visualization and Computer Graphics, vol. 14, no. 6, pp. 1404-1411, November/December, 2008.

Motivation

- **Many computer vision methods are based on mathematical optimization of initial parameter estimates to achieve accurate results**
  - *Bundle adjustment* non-linear optimization [1] in 3D scene reconstruction

- **Little interest has been given as to how <u>individual</u> values and <u>patterns</u> in parameter space affect the total cost**

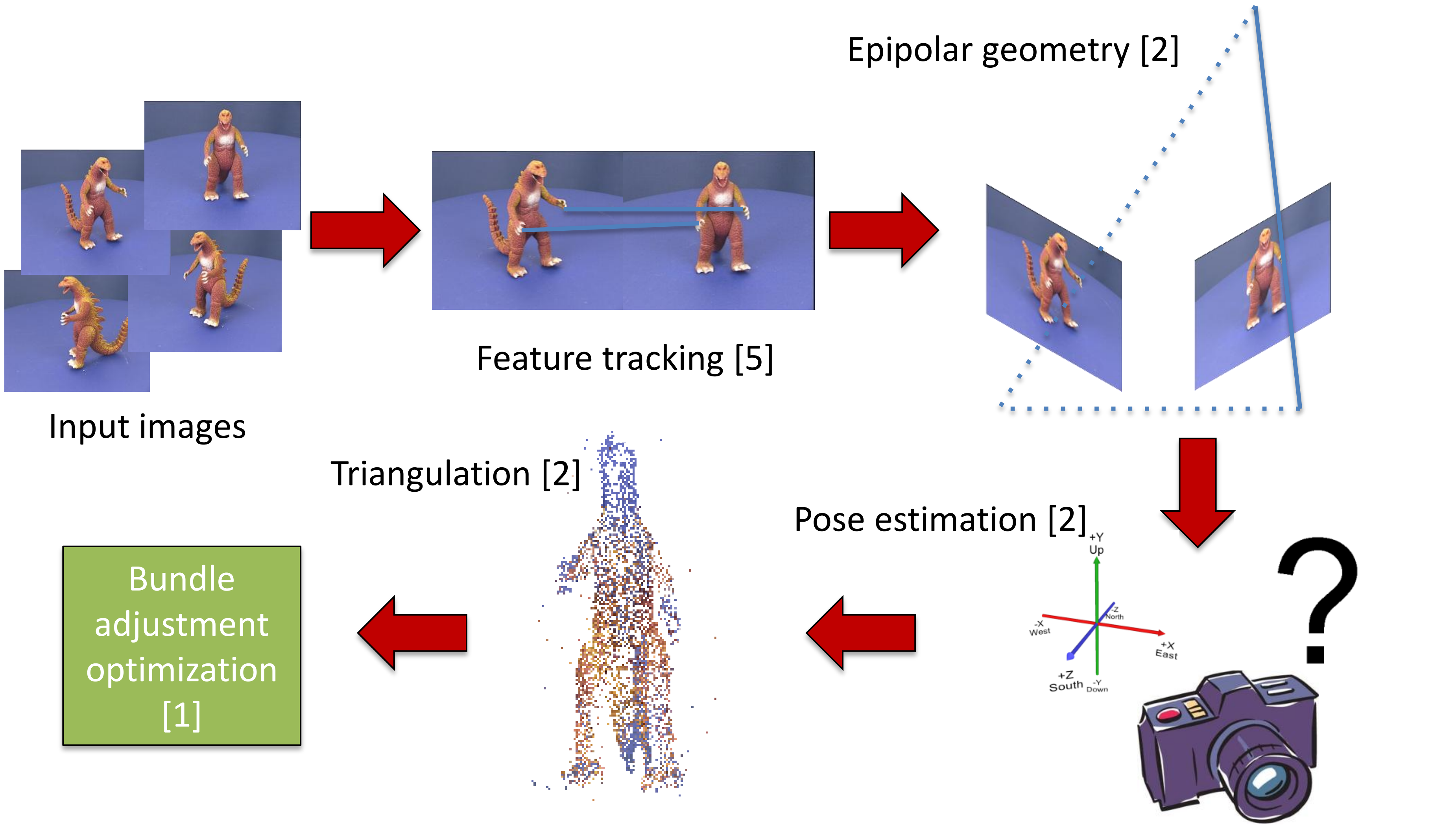
- **The main objective of our work is to introduce *scientific visualization* techniques to the *computer vision* community**:
  - Allows to determine patterns in the data, which aids in developing better algorithms
  - Allows for unique visually-aided numerical exploration of the solution space in a number of applications
  - As a very useful educational and algorithm design/test tool

- **Most of our work so far has been focused on such an analysis applied to the multi-view reconstruction of scenes [2,3]**

Motivation

- **Current work has focused on**:
  - Visualization of scene structure uncertainty
  - Feature tracking summaries
  - Covariance analysis (very initial)

Visualization of Scene Structure Uncertainty

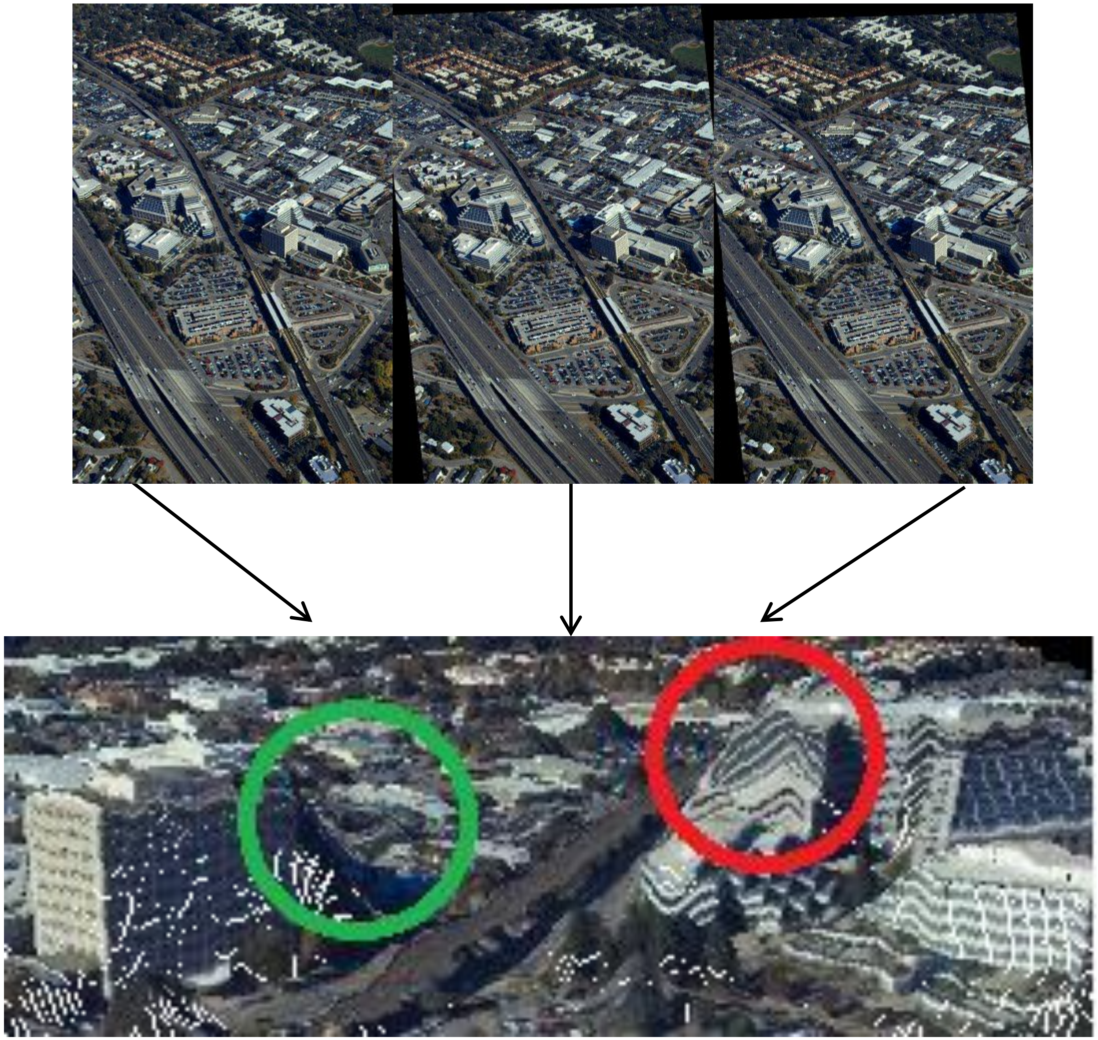


Multi-view scene reconstruction process: from images to point cloud

Dinosaur dataset images from [4].

Visualization of Scene Structure Uncertainty

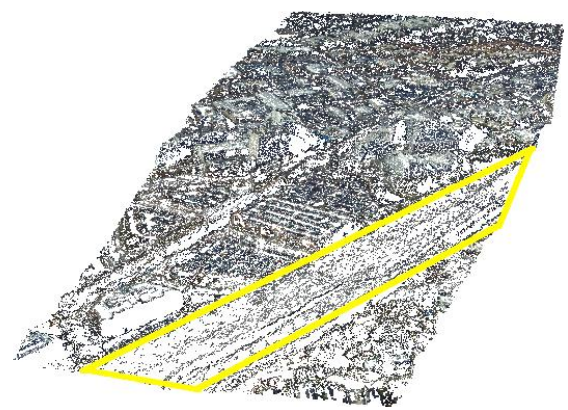
Warped input images



Occlusions Repetitive patterns

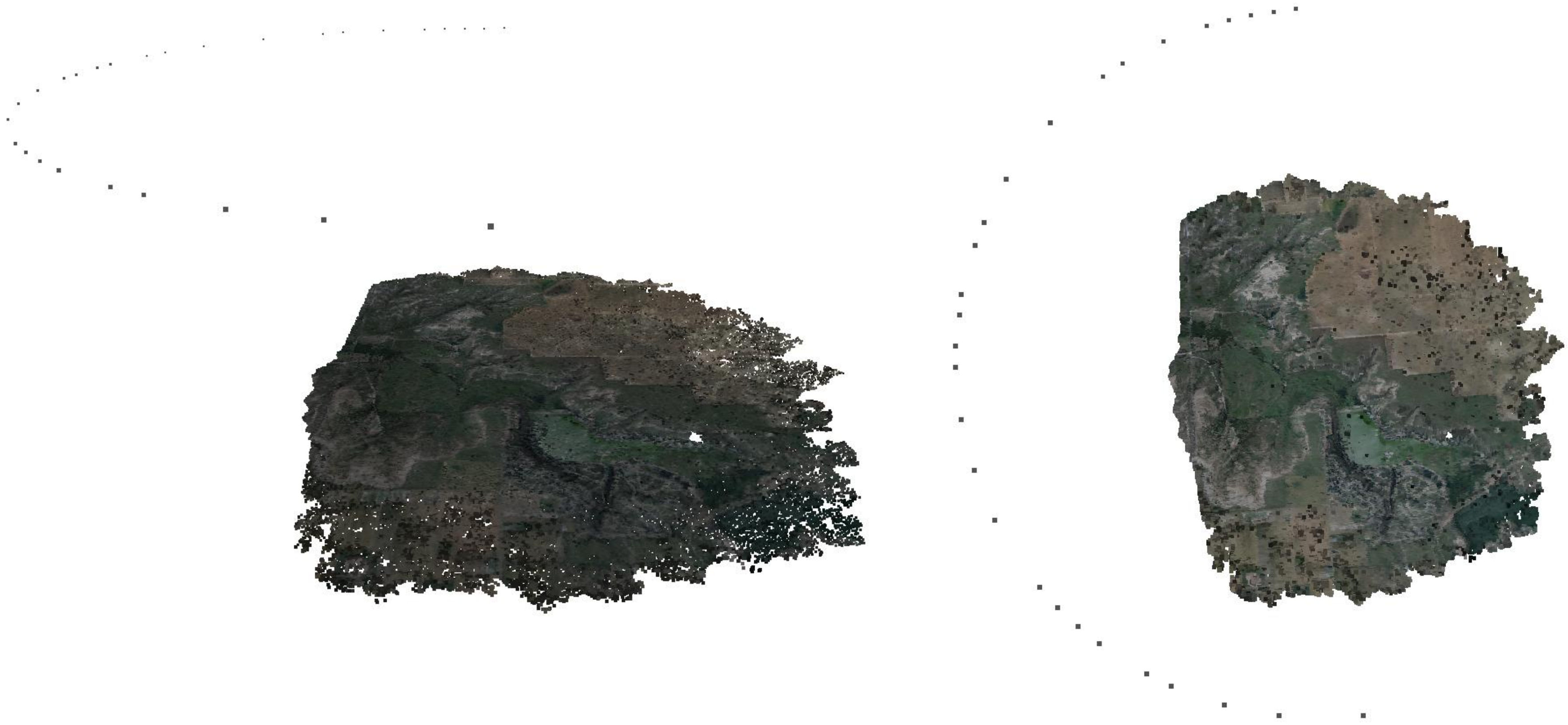
<u>Scene reconstruction pipeline</u>:

- **Feature tracking**
- Epipolar geometry and camera calibration
- Structure computation
- Bundle adjustment



Texture-less regions

Visualization of Scene Structure Uncertainty



Quasi-dense reconstruction of an aerial scene, 175866 points

Camera positions are rendered as blue dots

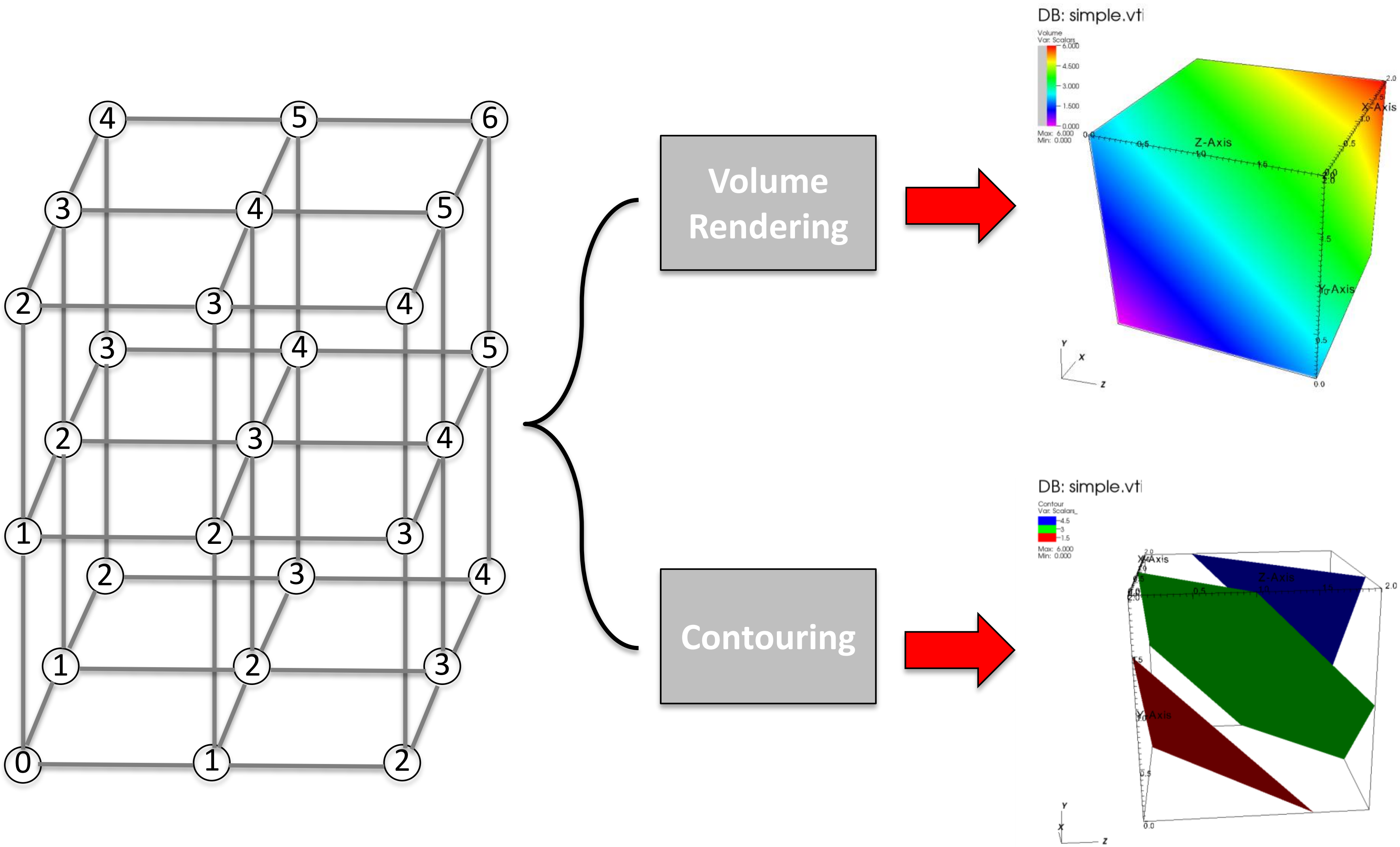
Visualization of Scene Structure Uncertainty

- We seek to investigate which factors affect the quality of multi-view scene reconstruction, in ways that are more visual and less mathematical than those of the current literature.

- To do so, we perform an analysis of scene structure **uncertainty** and its **sensitivity** to different multi-view scene reconstruction parameters, by borrowing techniques from scientific visualization.

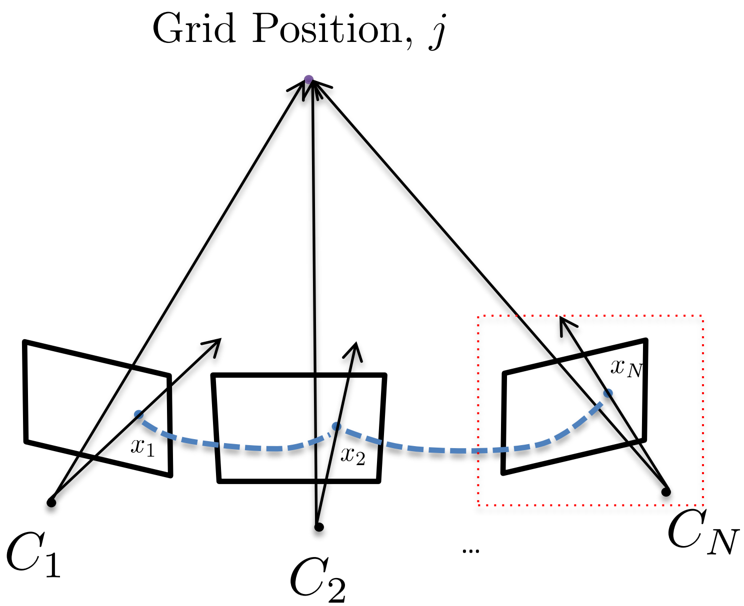
- We have created a tool based on **scalar field volume rendering** [6], which provides visual and numerical insight into structural uncertainty for a given 3D point cloud position, by analyzing the likelihood of positional error in its near vicinity.

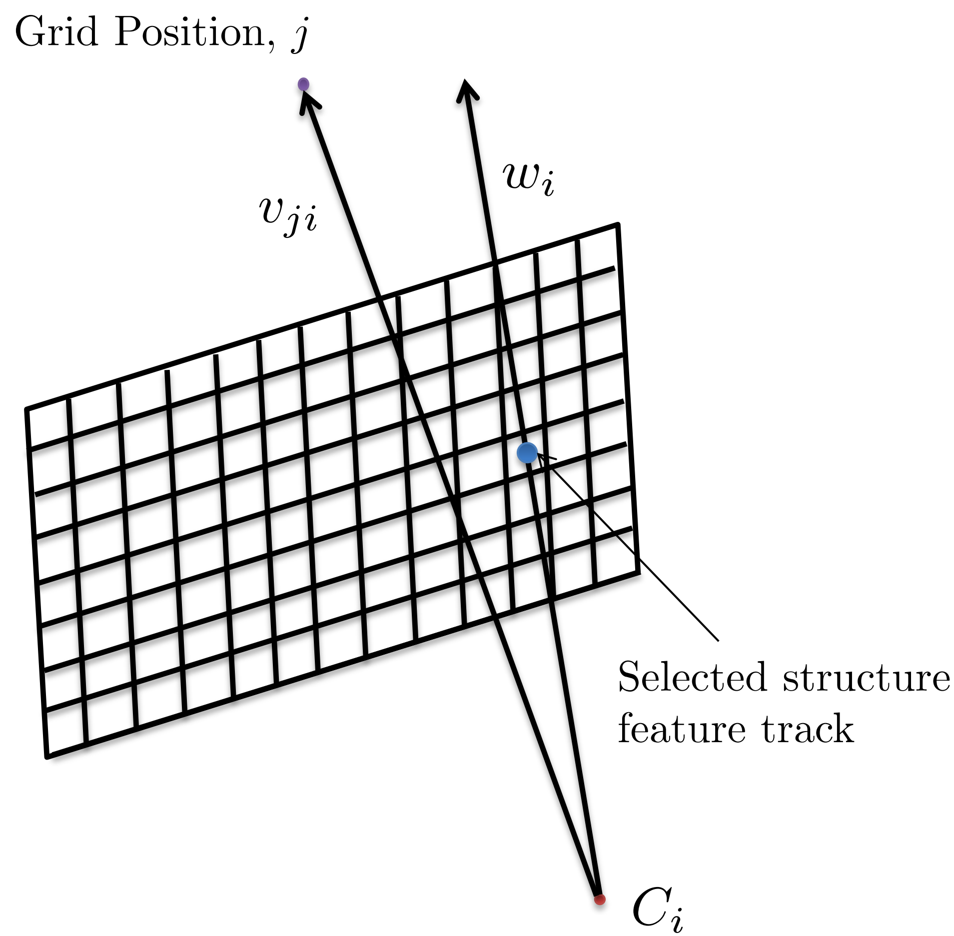
Visualization of Scene Structure Uncertainty



Scalar field volume rendering [6]

Visualization of Scene Structure Uncertainty





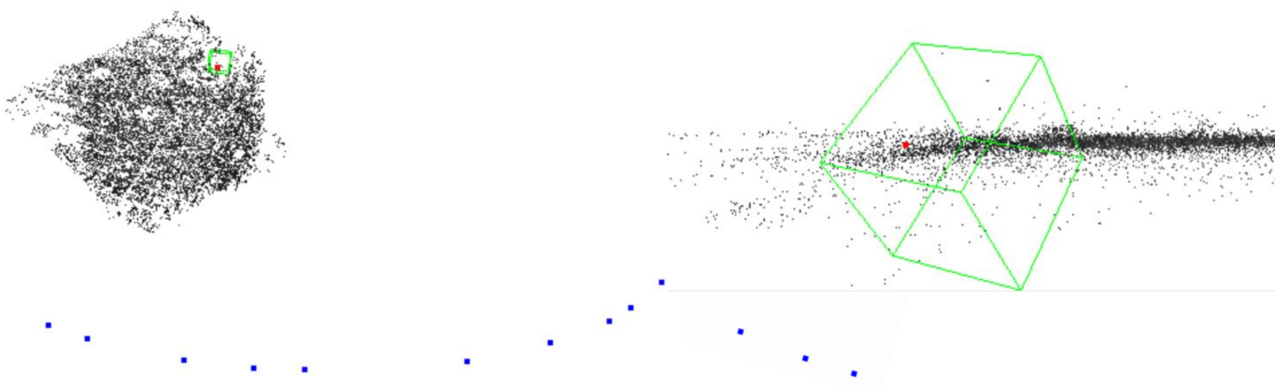
$$S_{ji} = cos^{-1}(v_{ji} \cdot w_i)$$

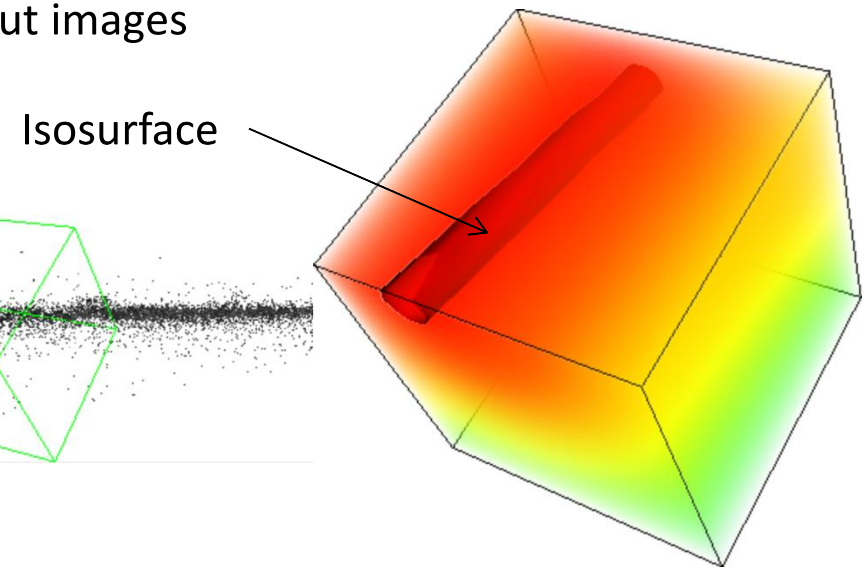
$$S_{j,ave} = \frac{\sum_{i=1}^{N} S_{ji}}{N}$$

$$S_{j,range} = S_{ji,max} - S_{ji,min}$$

Visualization of Scene Structure Uncertainty

Sample input images





Point cloud scene reconstruction and estimated camera positions (blue dots)

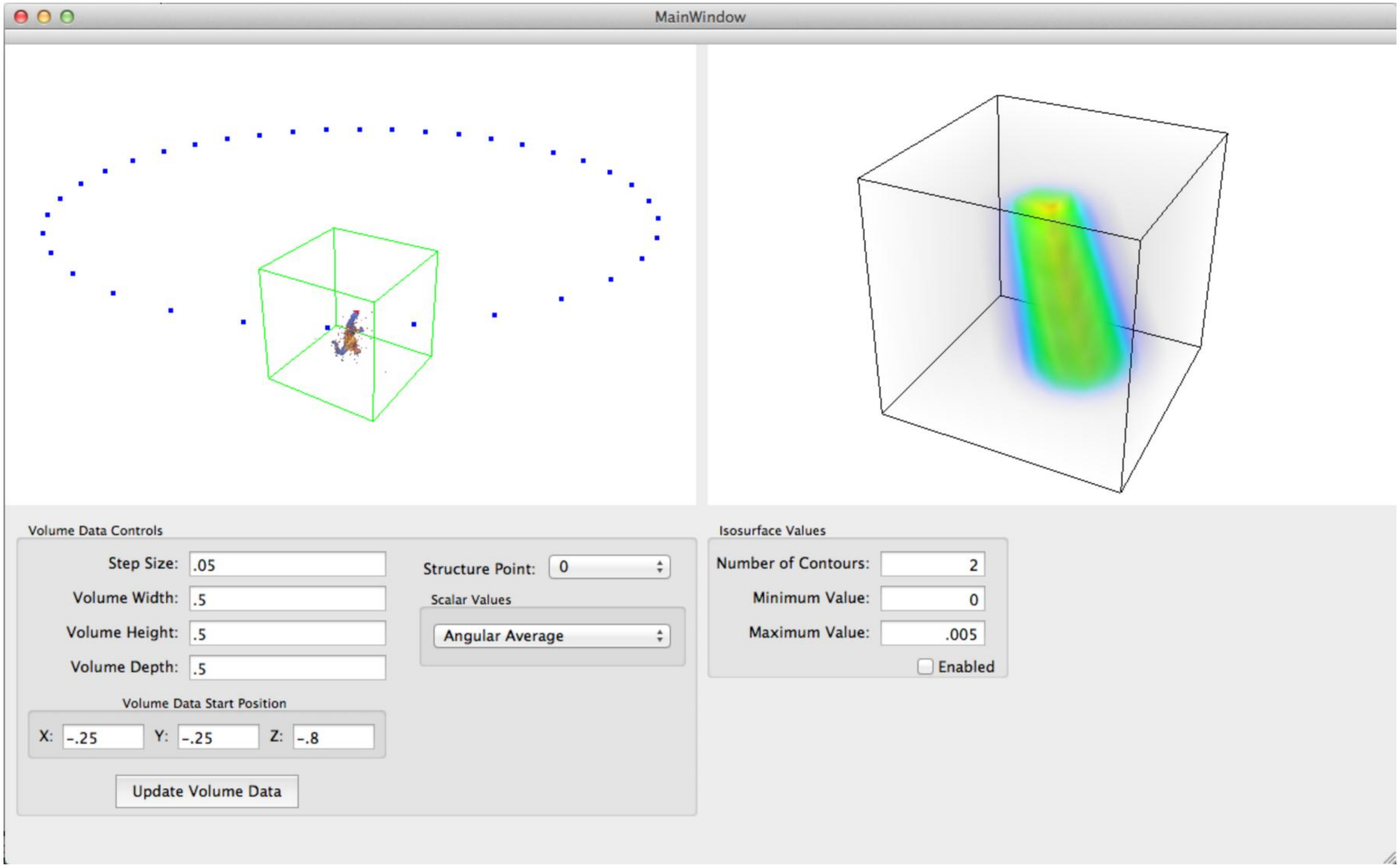
Bounding box (green) for analyzed position in the point cloud (red dot)

Computed scalar field in the bounding box (red = lower uncertainty)

Shawn Recker, Mauricio Hess-Flores, Mark A. Duchaineau, Kenneth I. Joy, "Visualization of Scene Structure Uncertainty in a Multi-View Reconstruction Pipeline", in "Vision, Modeling and Visualization Workshop (VMV 2012)", pp 183--190, 2012.

Shawn Recker, Mauricio Hess-Flores, Mark A. Duchaineau, Kenneth I. Joy, "Visualization of Scene Structure Uncertainty in Multi-View Reconstruction", in "Applied Imagery Pattern Recognition (AIPR) Workshop", 2012.

Visualization of Scene Structure Uncertainty



Inputs: camera projection matrices
feature tracks
scene structure

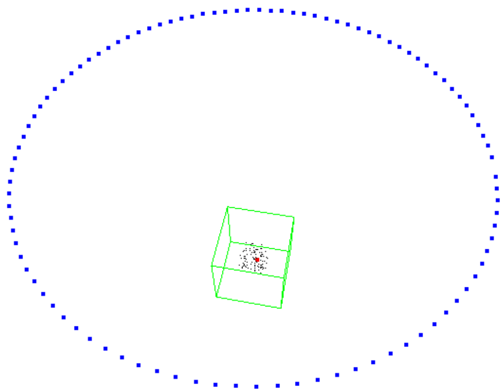
Output: statistical, visual, and isosurface information

Visualization of Scene Structure Uncertainty

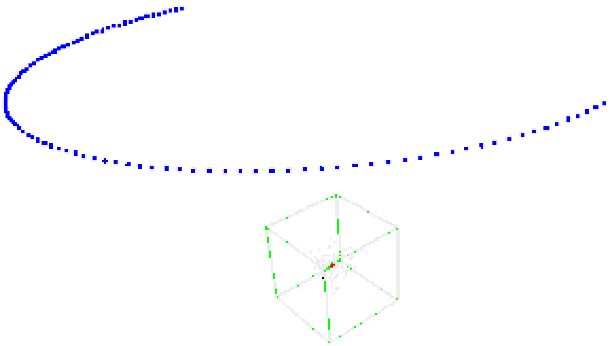
- **Uncertainty**:
  - Arises from **error accumulation** in the stages leading up to structure computation, such as **frame decimation**, **feature tracking**, and **camera-calibration**, where larger regions of low uncertainty, and hence larger isosurfaces, indicate robustness of the computed structure.
- **Sensitivity**:
  - Defined as the **change in scalar field values** as a specific reconstruction parameter's value changes.

S. Recker, M. Hess-Flores, M. A. Duchaineau, and K. I. Joy, "Visualization of Scene Structure Uncertainty in a Multi-View Reconstruction Pipeline," in Proceedings of the Vision, Modeling, and Visualization Workshop, 2012, pp. 183–190.

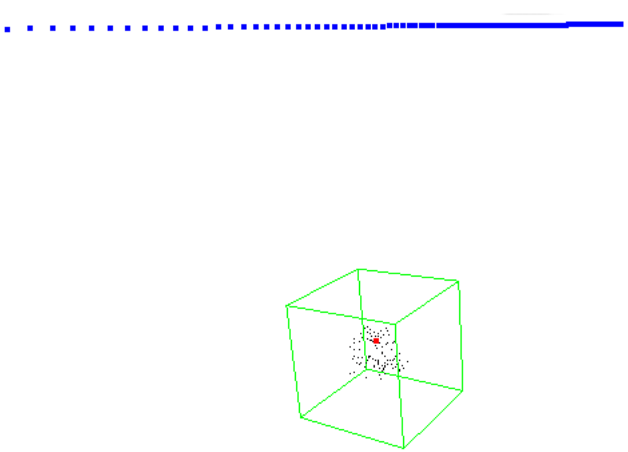
Visualization of Scene Structure Uncertainty



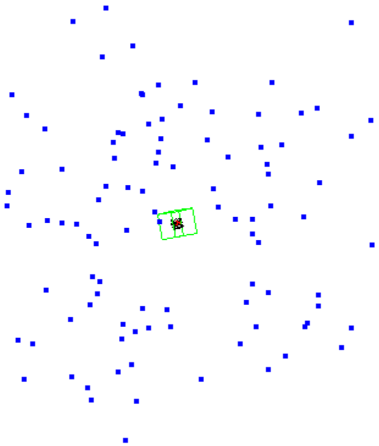
Circle



Semi-circle



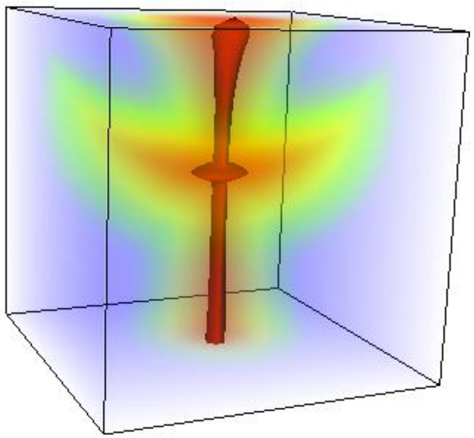
Line



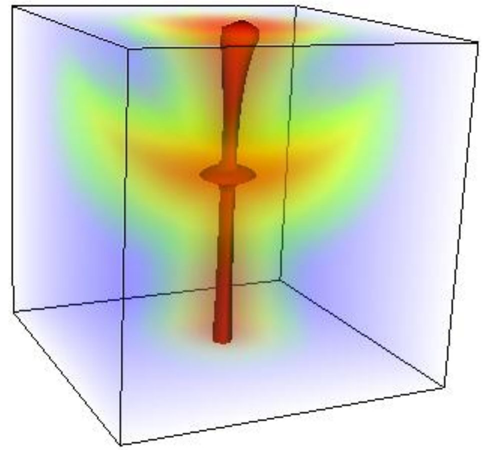
Random

Tested camera configurations

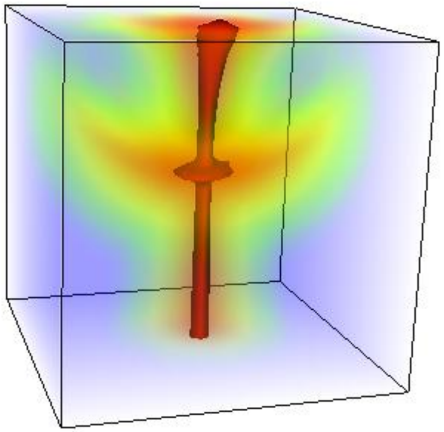
Visualization of Scene Structure Uncertainty



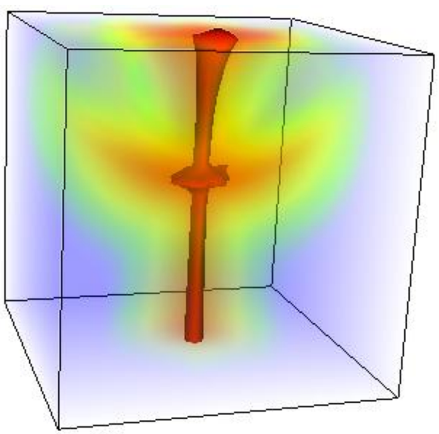
30 cameras



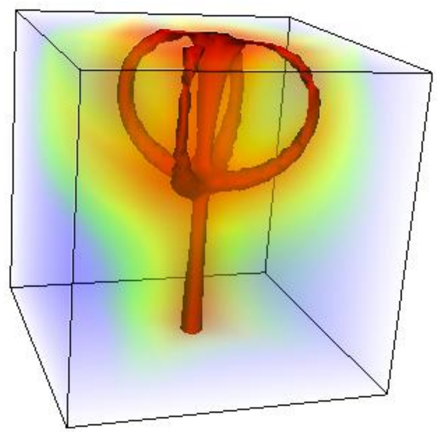
15 cameras



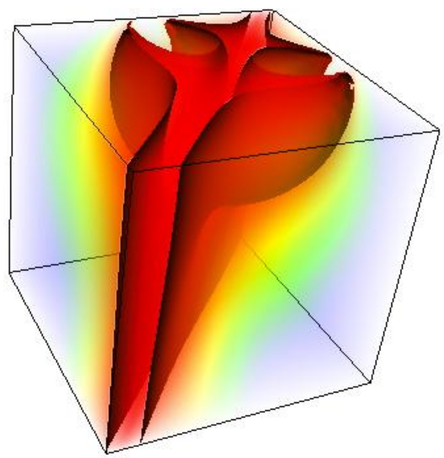
10 cameras



8 cameras



4 cameras

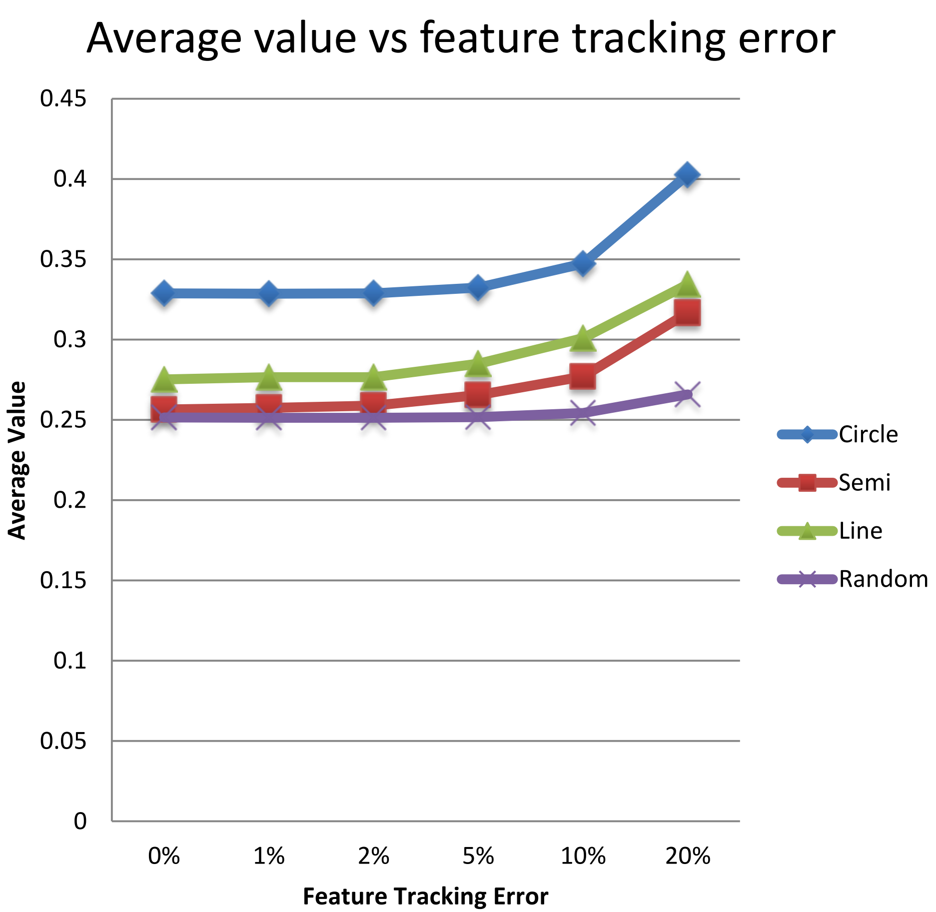


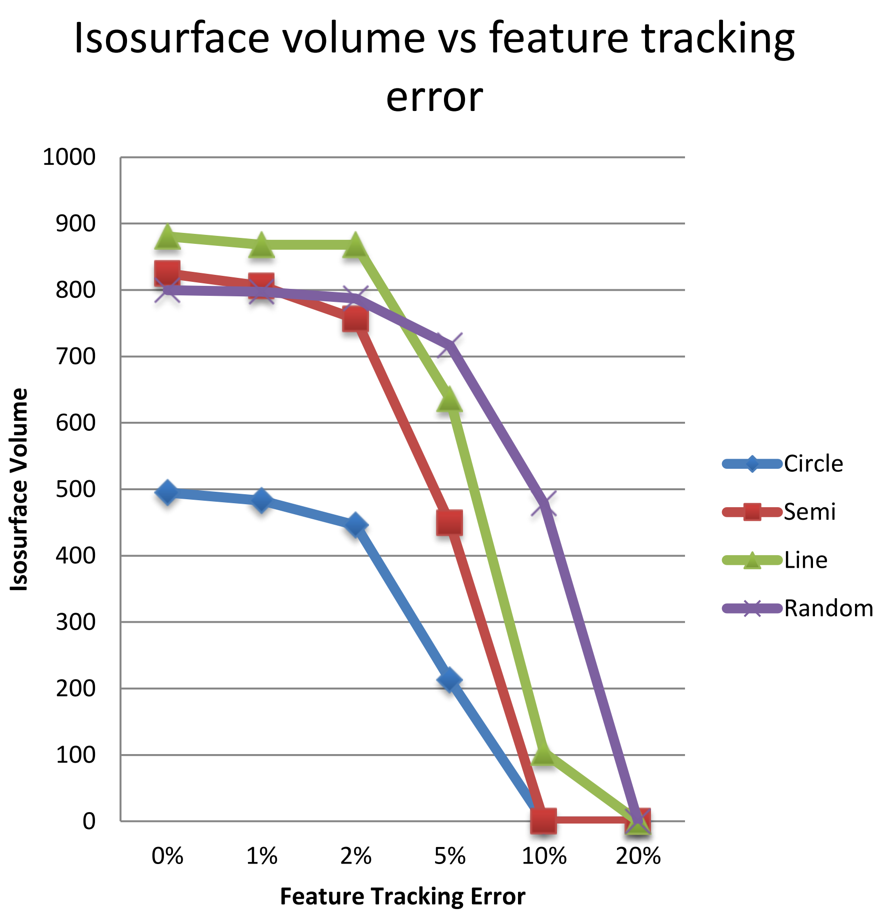
2 cameras

Frame decimation [8] simulation – circle configuration

Isovalue = 0.05

Visualization of Scene Structure Uncertainty



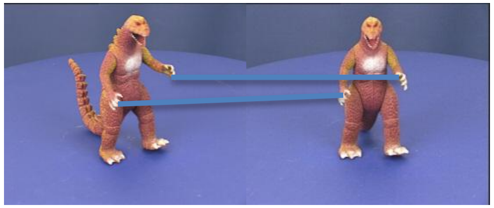


Feature tracking error simulation - scalar field values

Isovalue = 0.05

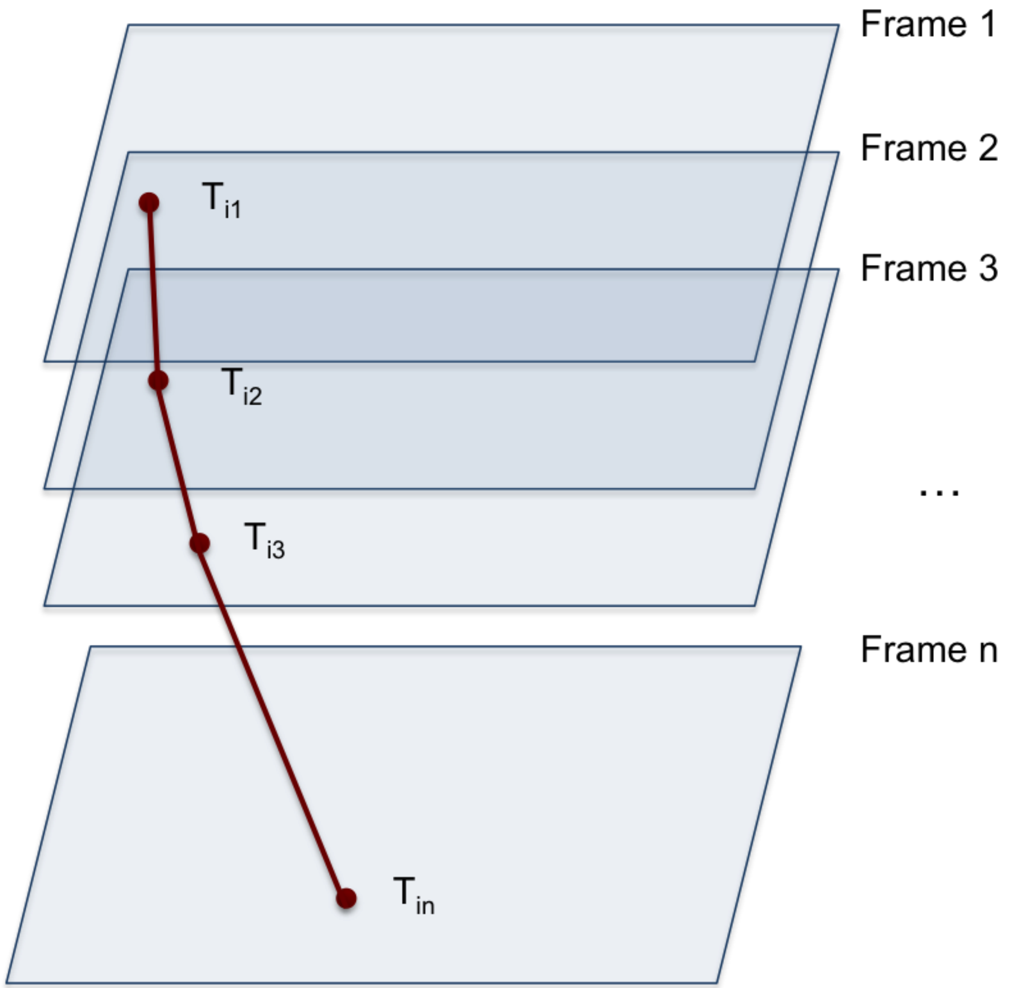
Feature tracking summaries

- **Feature track**:
  - A feature track is a set of pixel positions representing a scene point tracked over a set of images.



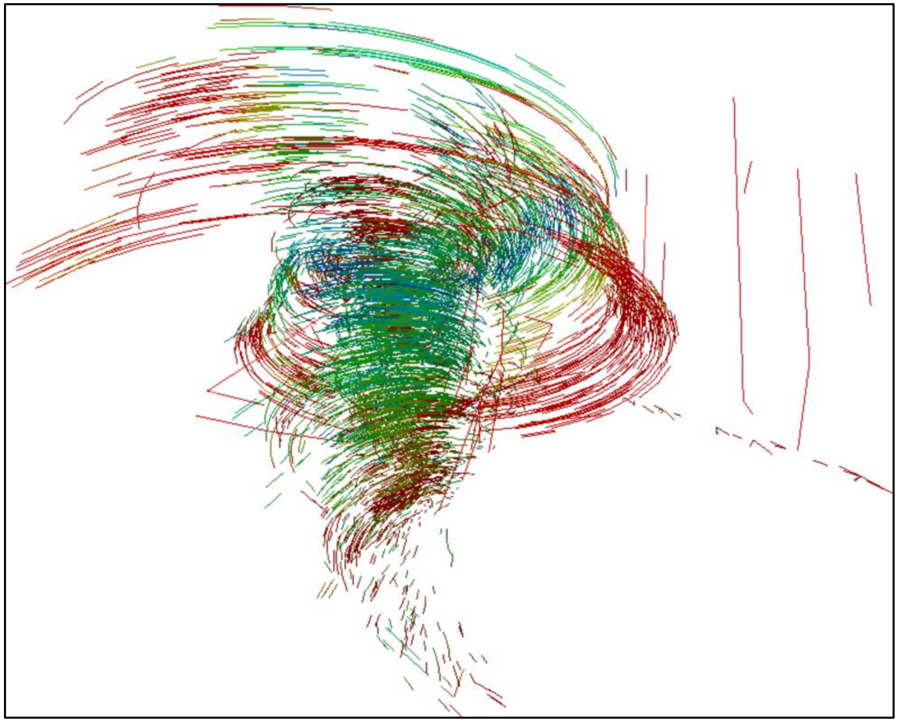
- **Feature track summary**:
  - Created by stacking frames vertically and observing the 'path' taken by a track, which we call *tracklines*.
  - Reconstruction parameter values can be encoded along the trackline, for example individual **reprojection errors** at each track position.
  - This provides insight into **track degeneration patterns** over time, as well as information about highly inaccurate individual track positions, all of which adversely affect subsequent camera pose estimation and structure computation.

Feature tracking summaries

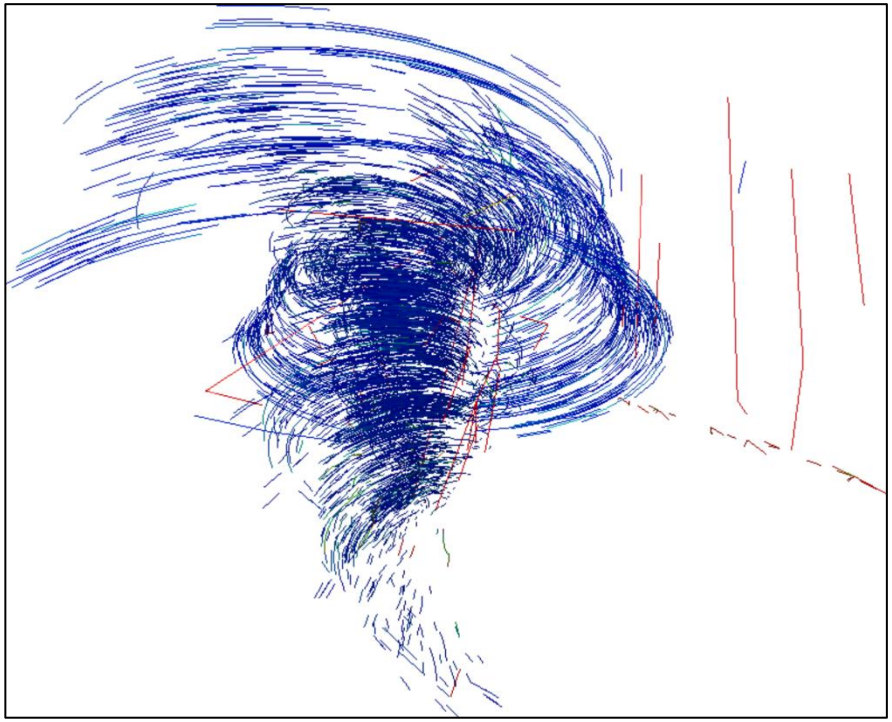


Creation of a *trackline* for a feature track summary

Feature tracking summaries



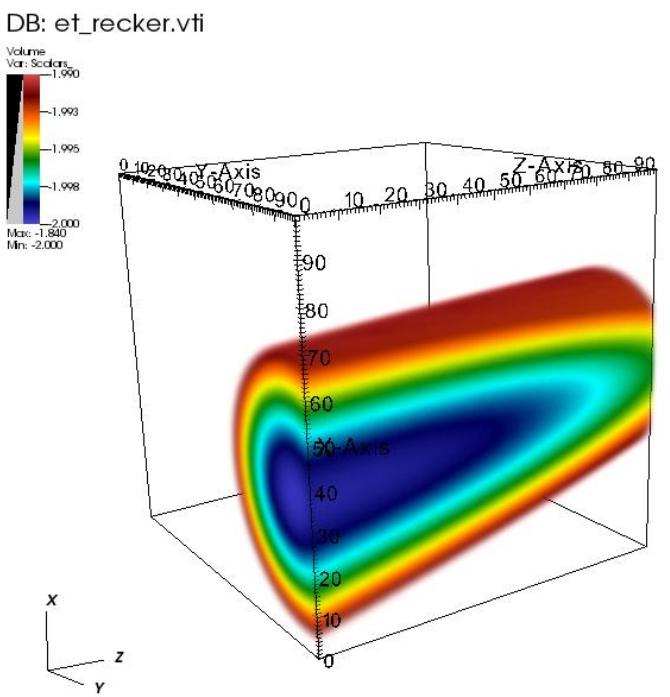
Before bundle adjustment optimization

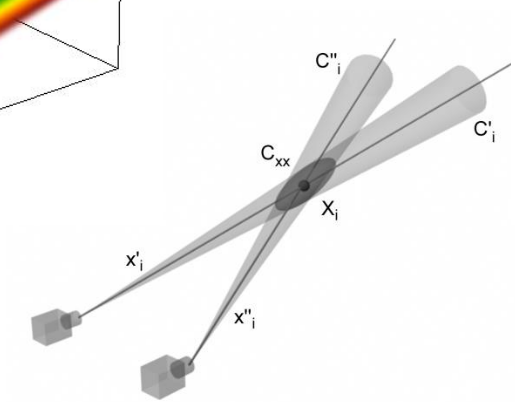


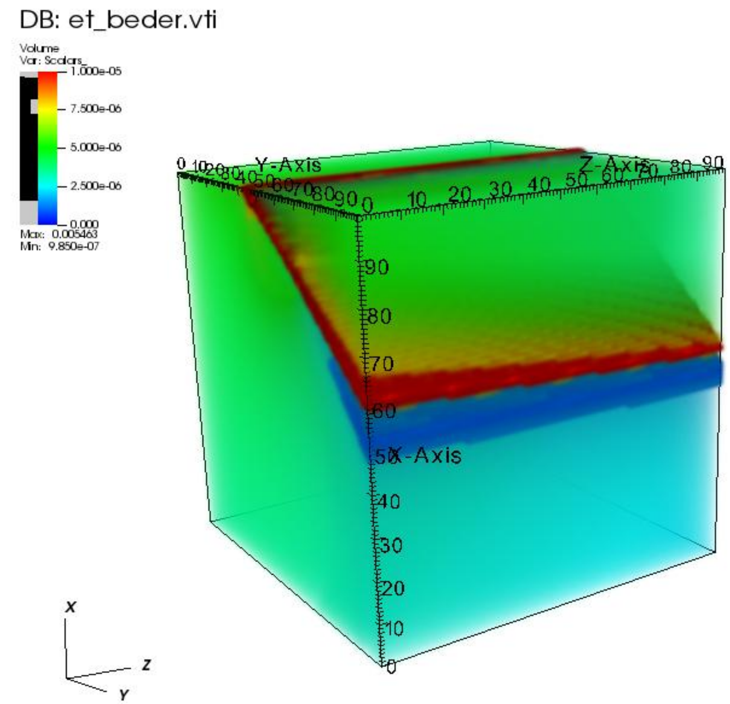
After bundle adjustment optimization

Feature track summary for the *Dinosaur* dataset. Each individual track represents the evolution over time, from top to bottom, of a given feature track's pixel position at each image of a sequential image stream.

Covariance analysis







Scalar field with our cost function (left) vs. Beder’s reconstruction uncertainty ellipsoid roundness [9] (right)

Conclusions and future work

- General idea of applying scientific visualization as a strong tool for computer vision research and analysis.

- Our framework, when used in tools with user interaction, allow for a visual uncertainty and sensitivity analysis, as opposed to purely numerical-based traditional methods.

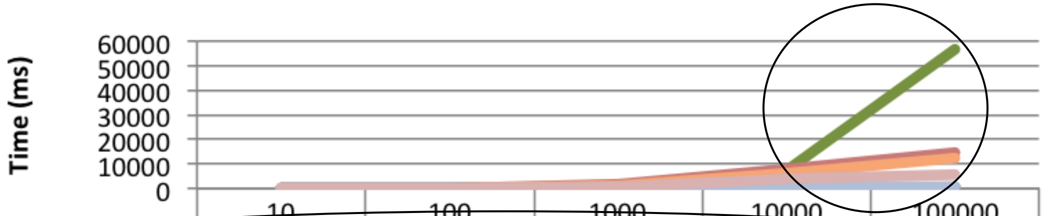
- Visualization of patterns in parameter estimates helps shed light into performance analysis for the underlying estimation algorithms.

- Initial results:
  - Multi-view scene structure uncertainty
  - Feature tracking summaries

- **We are very excited about many other possibilities stemming from this framework!**

Conclusions and future work

Our framework has directly let us design a state-of-the-art triangulation algorithm.



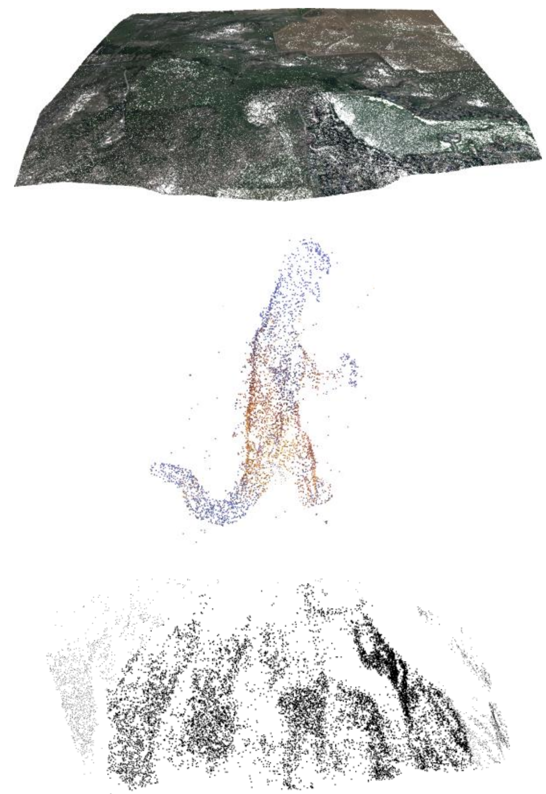
	10	100	1000	10000	100000
LT	30	114	744	7262	56900
MP	2	1	2	6	29
SGTL75	48	52	57	63	241
SGTL95	50	184	544	607	794
SGTL90	48	147	203	198	370
SGTL99	41	219	1666	7649	14632
SGTM75	15	31	34	54	225
SGTM95	20	126	430	496	651
SGTM90	21	97	161	140	315
SGTM99	24	151	1263	6026	12204
SLT95	28	79	226	333	433
SLT99	28	111	669	3795	5479

Feature Track Length

Standard linear triangulation (→ LT)

Our chosen triangulator (→ SGTM95)

Feature track length versus computation time



Sample reconstructions

Our method is *150x* faster than standard linear triangulation with long feature tracks, and is up to four orders of magnitude faster than optimal methods.

Shawn Recker, Mauricio Hess-Flores, Kenneth I. Joy, "Statistical Angular Error-Based Triangulation for Efficient and Accurate Multi-View Scene Reconstruction", in "Workshop on the Applications of Computer Vision (WACV)", 2013.

Conclusions and future work

- **Next steps and other applications we are targeting**:
- Further work on visualization of covariance matrix information
- Scalar field summaries for bundle adjustment optimization and other optimization cost functions
- Video processing
  - Object tracking results summaries for performance analysis
  - Video content summaries

Conclusions and future work

Thank you for your time and attention!

Contact information:
Mauricio Hess-Flores
Postdoctoral scholar
Institute for Data Analysis and Visualization
University of California, Davis
mhessf@ucdavis.edu

References

[1] M. Lourakis and A. Argyros, “The design and implementation of a generic sparse bundle adjustment software package based on the Levenberg-Marquardt algorithm,” Institute of Computer Science - FORTH, Heraklion, Crete, Greece, Tech. Rep. 340, August 2000.

[2] R. I. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision, 2nd ed. Cambridge University Press, 2004.

[3] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, “A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms,” in CVPR ’06: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC, USA: IEEE Computer Society, 2006, pp. 519–528.

[4] Visual Geometry Group, University of Oxford, “Multi-view and Oxford Colleges building reconstruction,” http://www.robots.ox.ac.uk/~vgg/data/data-mview.html.

[5] D. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints,” International Journal On Computer Vision, vol. 60, no. 2, pp. 91–110, 2004.

[6] INC. K.: Vtk: Visualization toolkit, 2012.

[7] S. Recker, M. Hess-Flores, M. A. Duchaineau, and K. I. Joy, “Visualization of Scene Structure Uncertainty in a Multi-View Reconstruction Pipeline,” in Proceedings of the Vision, Modeling, and Visualization Workshop, 2012, pp. 183–190.

[8] Daniel Knoblauch, Mauricio Hess-Flores, Mark A. Duchaineau, Kenneth I. Joy, Falko Kuester. Non-Parametric Sequential Frame Decimation for Scene Reconstruction in Low-Memory Streaming Environments. ISVC 2011, Part I, LNCS 6938, pp. 363-374. Springer, Heidelberg (2011)

[9] Beder, C., Steffen, R.: Determining an Initial Image Pair for Fixing the Scale of a 3D Reconstruction from an Image Sequence. In: DAGM-Symposium’06. (2006) 657–666