Methods for 3D Reconstruction from Multiple Images

Sylvain Paris
MIT CSAIL

Introduction

- Increasing need for geometric 3D models
  - Movie industry, games, virtual environments…
- Existing solutions are not fully satisfying
  - User-driven modeling: long and error-prone
  - 3D scanners: costly and cumbersome
- Alternative: analyzing image sequences
  - Cameras are cheap and lightweight
  - Cameras are precise (several megapixels)

Outline

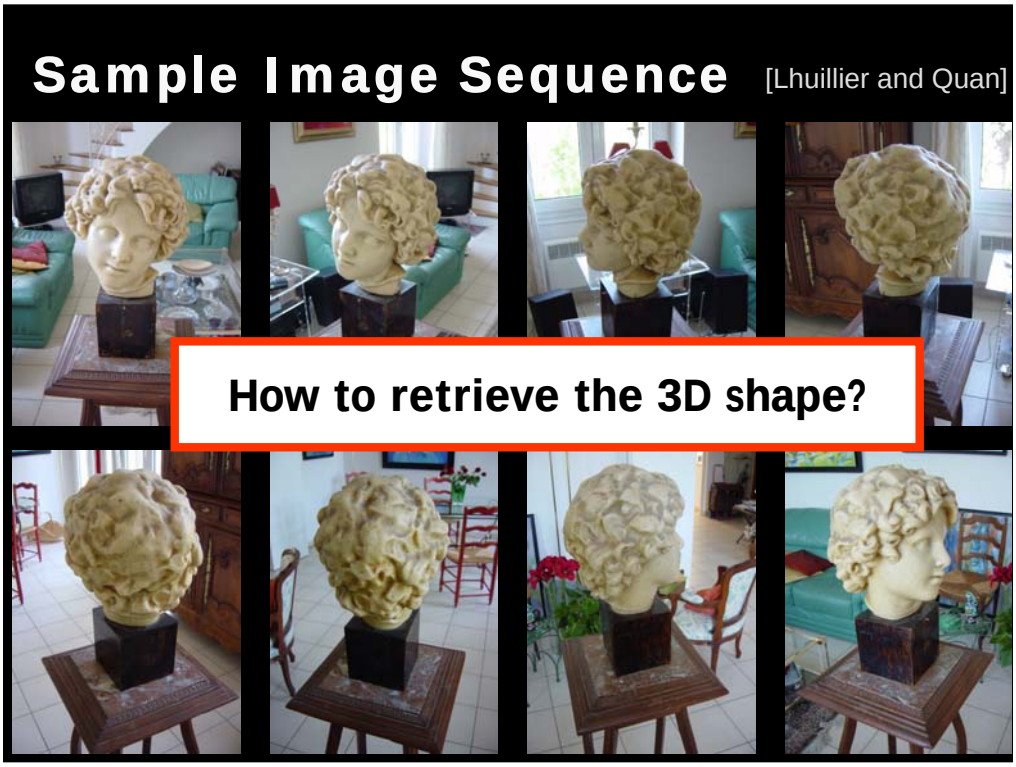
- Context and Basic Ideas
- Consistency and Related Techniques
- Regularized Methods
- Conclusions

Outline

- Context and Basic Ideas
- Consistency and Related Techniques
- Regularized Methods
- Conclusions

Scenario

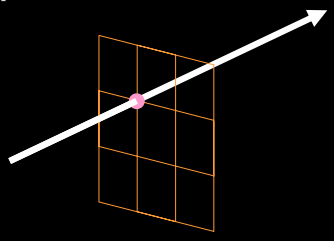
- A scene to reconstruct (unknown a priori)
- Several viewpoints
  - from 4 views up to several hundreds
  - 20~50 on average
- "Over water"
  - non-participating medium



The image sequence is available on Long Quan's webpage:
http://www.cs.ust.hk/~quan/WebPami/pami.html

First Step: Camera Calibration

- Associate a pixel to a ray in space
  - camera position, orientation, focal length…
- Complex problem
  - solutions exist
  - toolboxes on the web
  - commercial software available

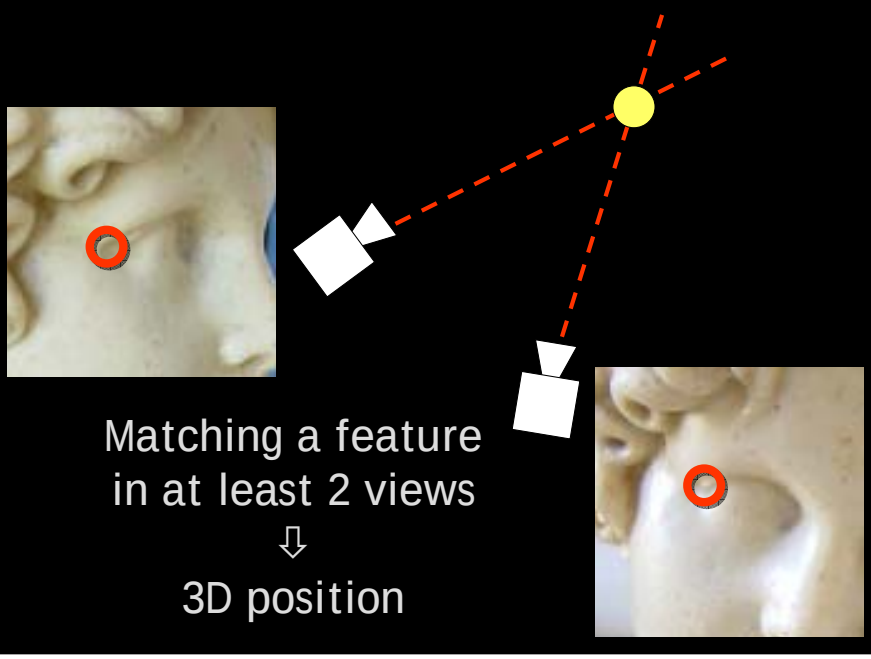


2D pixel ⇔ 3D ray

Outline

- Context and Basic Ideas
- Consistency and Related Techniques
- Regularized Methods
- Conclusions

General Strategy: Triangulation



Matching a feature
in at least 2 views

⇩

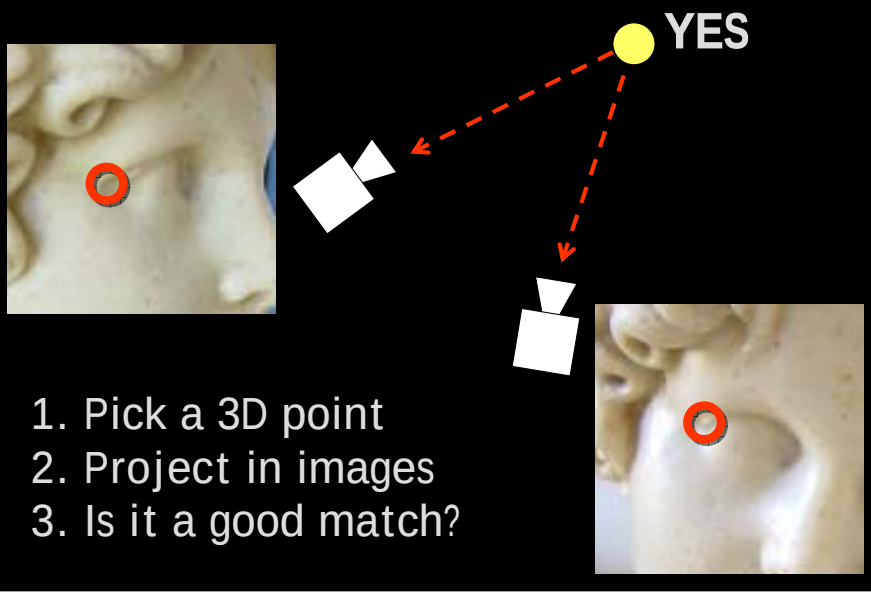
3D position

Matching First

Which points are the same?

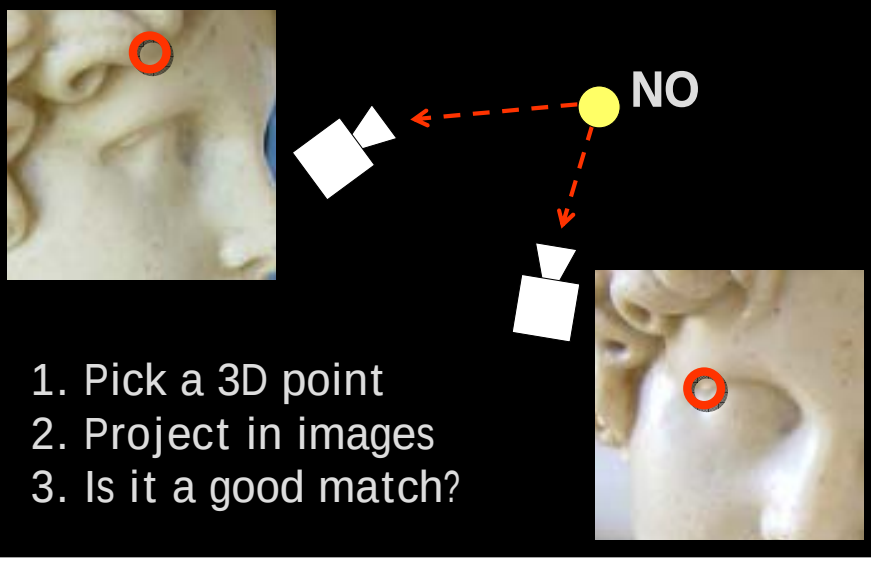
Impossible to match all points ⇨ holes.
Not suitable for dense reconstruction.

Sampling 3D Space



1. Pick a 3D point
2. Project in images
3. Is it a good match?

Sampling 3D Space



1. Pick a 3D point
2. Project in images
3. Is it a good match?

Consistency Function

"Is this 3D model consistent with the input images?"

- No binary answer
  - noise, imperfect calibration…
- Scalar function
  - low values: good match
  - high values: poor match

Examples of Consistency Functions

- Color: variance [Seitz 97]
  - *Do the cameras see the same color?*
  - Valid for matte (Lambertian) objects only.
- Texture: correlation
  - *Is the texture around the points the same?*
  - Robust to glossy materials.
  - Problems with shiny objects and grazing angles.
- More advanced models [Yang 03, Jin 05]
  - Shiny and transparent materials.

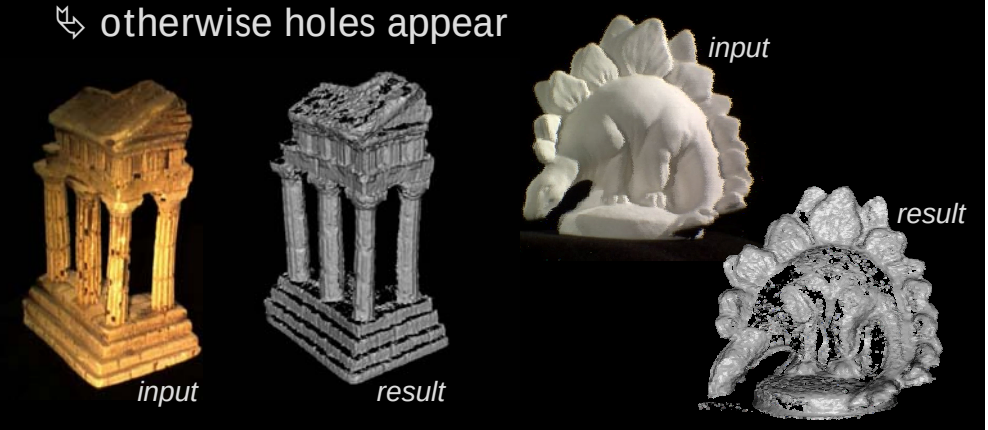
[Seitz 97] Photorealistic Scene Reconstruction by Voxel Coloring S. M. Seitz and C. R. Dyer, Proc. Computer Vision and Pattern Recognition Conf., 1997, 1067-1073.

[Yang 03] R. Yang, M. Pollefeys, and G. Welch. Dealing with Textureless Regions and Specular Highlight: A Progressive Space Carving Scheme Using a Novel Photo-consistency Measure, Proc. of the International Conference on Computer Vision, pp. 576-584, 2003

[Jin 05] H. Jin, S. Soatto and A. Yezzi. Multi-view stereo reconstruction of dense shape and complex appearance
Intl. J. of Computer Vision 63(3), p. 175-189, 2005.

Reconstruction from Consistency Only

- Gather the good points [Lhuillier 02, Goesele 06]
  - requires many views
  - otherwise holes appear



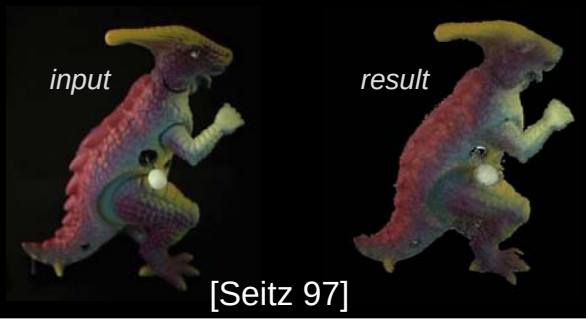
[Goesele 06]

[Lhuillier 02] ECCV'02, Quasi-Dense Reconstruction from Image Sequence. M. Lhuillier and L. Quan, Proceedings of the 7th European Conference on Computer Vision, Copenhagen, Denmark, Volume 2, pages 125-139, May 2002

[Goesele 06] Michael Goesele, Steven M. Seitz and Brian Curless. Multi-View Stereo Revisited, Proceedings of CVPR 2006, New York, NY, USA, June 2006.

Reconstruction from Consistency Only

- Remove the bad points [Seitz 97, Kutulakos 00]
  1. start from bounding volume
  2. carve away inconsistent points
  ↳ requires texture
     - otherwise incorrect geometry



[Seitz 97]

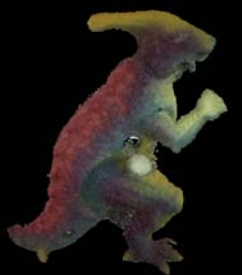
[Seitz 97] Photorealistic Scene Reconstruction by Voxel Coloring S. M. Seitz and C. R. Dyer, Proc. Computer Vision and Pattern Recognition Conf., 1997, 1067-1073.

[Kutulakos 00] A Theory of Shape by Space Carving. K. N. Kutulakos and S. M. Seitz, International Journal of Computer Vision, 2000, 38(3), pp. 199-218

Summary of "Consistency Only Strategy"

- With high resolution data
  - mostly ok (except textureless areas)
  - sufficient in many cases
- Advice: try a simple technique first.
- More sophisticated approach
  - fill holes
  - more robust (noise, few images…)

[Goesele 06]



[Seitz 97]

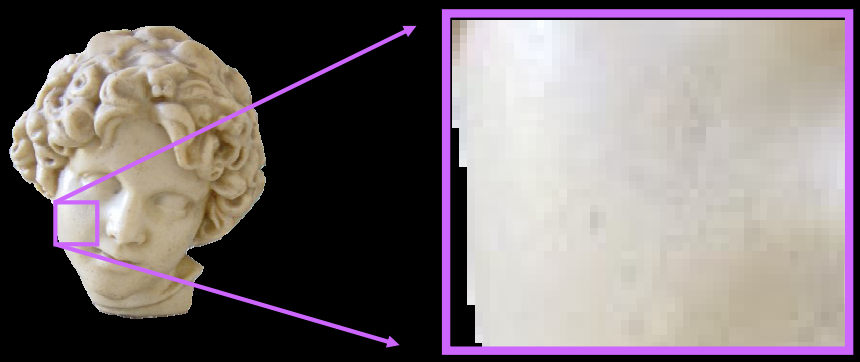
[Seitz 97] Photorealistic Scene Reconstruction by Voxel Coloring S. M. Seitz and C. R. Dyer, Proc. Computer Vision and Pattern Recognition Conf., 1997, 1067-1073.

[Goesele 06] Michael Goesele, Steven M. Seitz and Brian Curless. Multi-View Stereo Revisited, Proceedings of CVPR 2006, New York, NY, USA, June 2006.

Outline

- Context and Basic Ideas
- Consistency and Related Techniques
- Regularized Methods
- Conclusions

Consistency is not Enough



- Textureless regions
  - Everything matches.
  - No salient points.

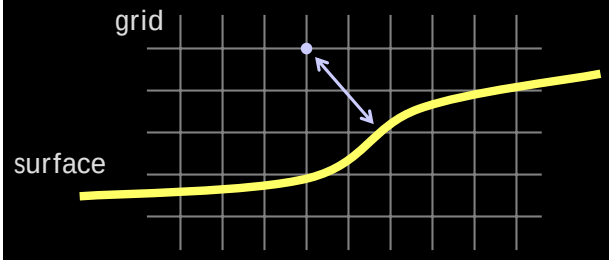
An Ill-posed Problem

There are several different 3D models consistent with an image sequence.

- More information is needed.
  - User provides a priori knowledge.
  - Classical assumption: Objects are "smooth."
  - Also know as *regularizing the problem*.
- Optimization problem:
  - Find the "best" smooth consistent object.

Minimal Surfaces with Level Sets

- Smooth surfaces have small areas.
  - ↳ "smoothest" translates into "minimal area."
- Level Sets to search for minimal area solution.
  - ↳ surface represented by its "distance" function

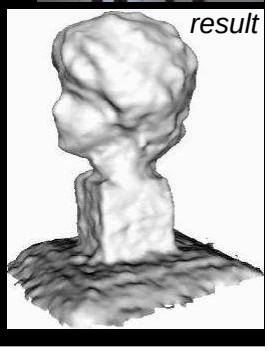


Each grid node stores its distance to the surface.

Minimal Surfaces with Level Sets

- Distance function evolves towards best tradeoff consistency vs area.
  [Keriven 98, Jin 05, Lhuillier 05]
- Advantages
  - match arbitrary topology
  - exact visibility
- Limitations
  - no edges, no corners
  - convergence unclear (ok in practice)

[Lhuillier 05]



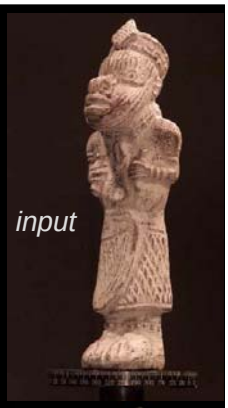
[Keriven 98] R. Keriven and O. Faugeras. Complete dense stereovision using level set methods. In Hans Burkhardt and Bernd Neumann, editors, Proceedings of the 5th European Conference on Computer Vision, volume 1406 of Lecture Notes on Computer Science, pages 379-393. Springer-Verlag, 1998.

[Jin 05] H. Jin, S. Soatto and A. Yezzi. Multi-view stereo reconstruction of dense shape and complex appearance
Intl. J. of Computer Vision 63(3), p. 175-189, 2005.

[Lhuillier 05] A Quasi-Dense Approach to Surface Reconstruction from Uncalibrated Images. Maxime Lhuillier and Long Quan. Trans. On Pattern Analysis and Machine Intelligence, vol 27, no. 3, pp. 418--433, March 2005

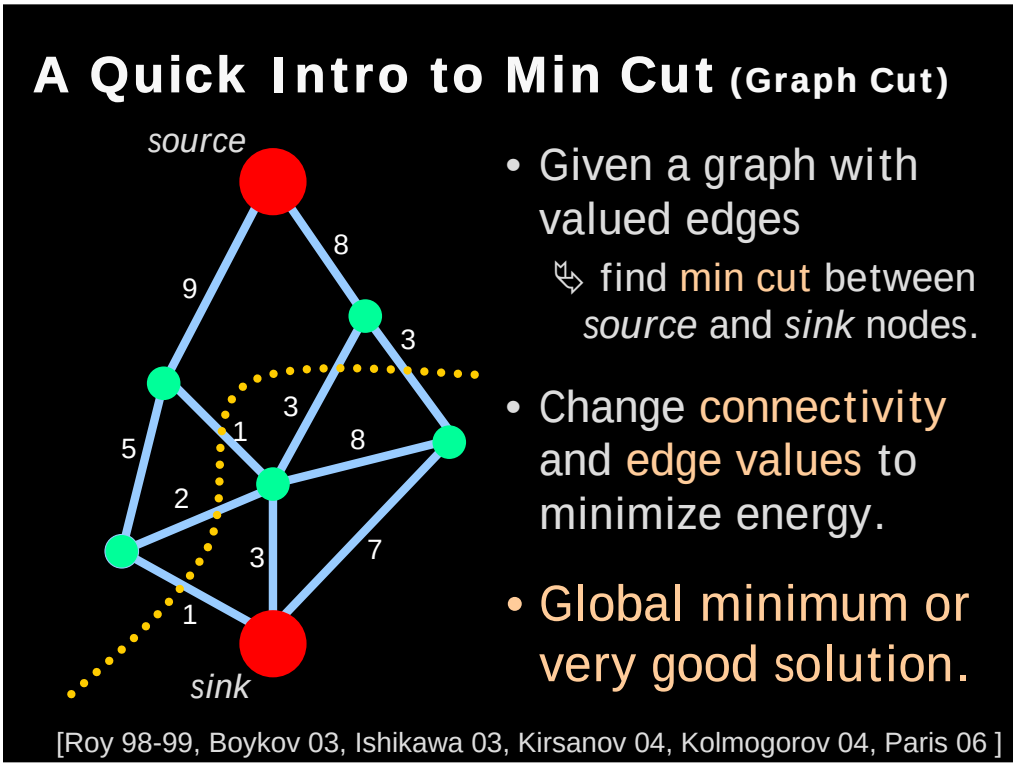
Snakes [Hernández 04]

- Explicit surface representation
  - ↳ triangle mesh
- Controlled setup
- Robust matching scheme
  - ↳ precise
  - ↳ handles very glossy material
  - ↳ computationally expensive



[Hernández 04]

[Hernández 04] Silhouette and Stereo Fusion for 3D Object Modeling. C. Hernández and F. Schmitt. Computer Vision and Image Understanding, Special issue on "Model-based and image-based 3D Scene Representation for Interactive Visualization", vol. 96, no. 3, pp. 367-392, December 2004



[Roy 98] A Maximum-Flow Formulation of the N-Camera Stereo Correspondence Problem. Proceedings of the Sixth International Conference on Computer Vision. 1998. Sébastien Roy Ingemar J. Cox

[Roy 99] Stereo Without Epipolar Lines: A Maximum-Flow Formulation. S Roy - International Journal of Computer Vision, 1999

[Boykov 03] Computing Geodesics and Minimal Surfaces via Graph Cuts. Yuri Boykov and Vladimir Kolmogorov. In International Conference on Computer Vision, (ICCV), vol. I, pp. 26-33, 2003.

[Ishikawa 03] Exact Optimization for Markov Random Fields with Convex Priors. Hiroshi Ishikawa IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, No. 10, pp. 1333-1336. October 2003

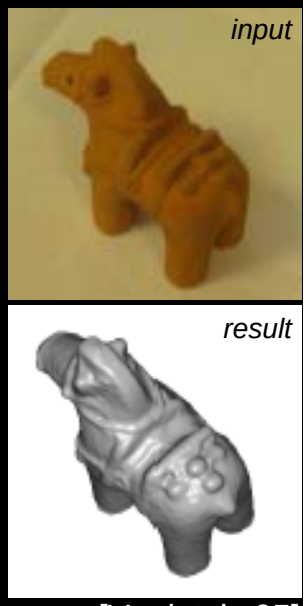
[Kirsanov 04] "A Discrete Global Minimization Algorithm for Continuous Variational Problems“ D. Kirasanov and S. J. Gortler. Harvard Computer Science Technical Report: TR-14-04, July 2004

[Kolmogorov 04] What Energy Functions can be Minimized via Graph Cuts? Vladimir Kolmogorov and Ramin Zabih. In IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 26(2):147-159, February 2004.

[Paris 06] A Surface Reconstruction Method Using Global Graph Cut Optimization. Sylvain Paris, François Sillion, and Long Quan. International Journal on Computer Vision (IJCV'06)

Minimal Surfaces with Graph Cut

- Graphs can be used to compute min surfaces [Boykov 03]
- Visibility must be known
  - requires silhouettes [Vogiatzis 05]
- Advantages
  - high accuracy
  - capture edges, corners
  - convergence guaranteed



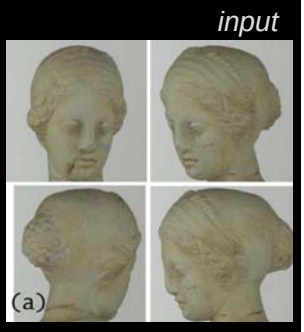
[Vogiatzis 05]

[Boykov 03] Computing Geodesics and Minimal Surfaces via Graph Cuts. Yuri Boykov and Vladimir Kolmogorov. In International Conference on Computer Vision, (ICCV), vol. I, pp. 26-33, 2003.

[Vogiatzis 05] Multi-view stereo via Volumetric Graph-cuts. G. Vogiatzis, P.H.S. Torr and R. Cipolla. In Proceedings IEEE Conference on Computer Vision and Pattern Recognition 2005

Exploiting Silhouettes

- Traditional techniques
  - 3D model only inside silhouettes
- Exact silhouettes [Sinha 05]
  - coherent framework
  - high accuracy at silhouettes
  - robust
  - but computationally expensive
    - (4D graph)
  - lacks detail (can be improved)

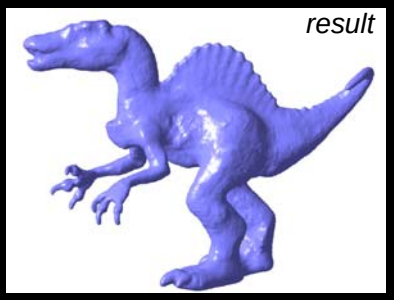


[Sinha 05]

[Sinha 05] S. Sinha, M. Pollefeys, Multi-view Reconstruction using Photo-consistency and Exact Silhouette Constraints: A Maximum-Flow Formulation, Proc. ICCV'05, Vol. 1, pp. 349-356, 2005.

Exploiting Silhouettes

- Exact silhouettes [Furukawa 06]
  - more detail
  - slightly less robust
    - silhouettes handled separately
  - better tradeoff
  - but computationally expensive (2 hours +)

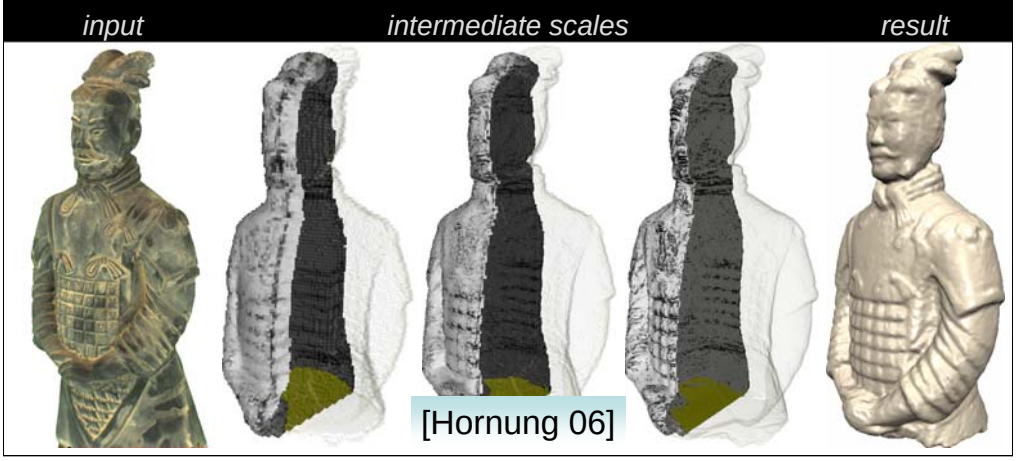


[Furukawa 06]

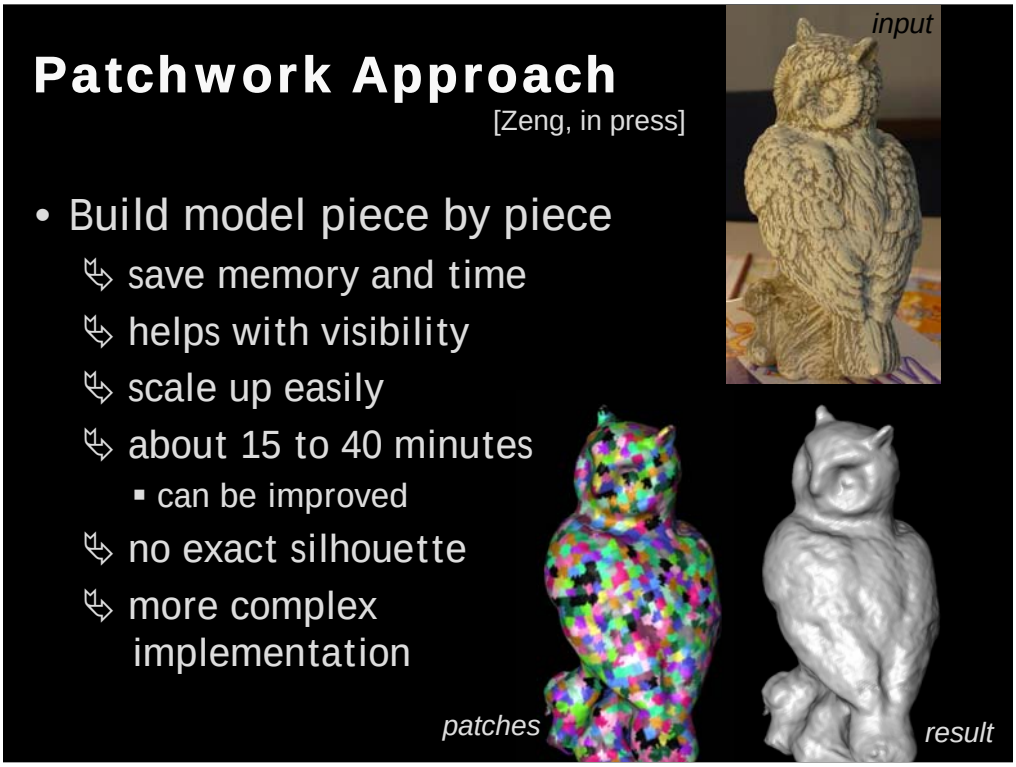
[Furukawa 06] Carved Visual Hulls for Image-Based Modeling. Yasutaka Furukawa and Jean Ponce. European Conference on Computer Vision, Graz, Austria, May 2006

Multi-scale Approach [Hornung 06]

- Optimizing only a narrow band
- Progressive refinement
  - About 10 to 30 minutes (and no exact silhouettes)



[Hornung 06] A. Hornung and L. Kobbelt. Hierarchical volumetric multi-view stereo reconstruction of manifold surfaces based on dual graph embedding. CVPR 2006



[Zeng, in press] Accurate and Scalable Surface Representation and Reconstruction from Images. Gang Zeng, Sylvain Paris, Long Quan, and Francois Sillion. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)

Challenges for the Future

- Shinny materials: metal, porcelain… [Vogiatzis 06]
- Choice of the parameters
  - Controlled setup is ok.
  - Difficulties: handheld camera, outdoor,…
- Visibility and graph cut
  - Restricted setup [Kolmogorov 02]
  - Only at "large scale" [Vogiatzis 05, Zeng in press]
  - Promising direction: iterative graph cuts [Boykov 06]

[Kolmogorov 02] Multi-camera Scene Reconstruction via Graph Cuts. Vladimir Kolmogorov and Ramin Zabih. European Conference on Computer Vision (ECCV), May 2002

[Vogiatzis 05] Multi-view stereo via Volumetric Graph-cuts. G. Vogiatzis, P.H.S. Torr and R. Cipolla. In Proceedings IEEE Conference on Computer Vision and Pattern Recognition 2005

[Boykov 06] An Integral Solution to Surface Evolution PDEs via Geo-Cuts. Yuri Boykov, Vladimir Kolmogorov, Daniel Cremers, Andrew Delong. In *European Conference on Computer Vision*, (*ECCV*), LNCS 3953, vol.III, pp.409-422, May 2006.

[Vogiatzis 06] Reconstruction in the Round Using Photometric Normals and Silhouettes. G. Vogiatzis, C. Hernández and R. Cipolla. CVPR 2006, New York, vol. 2, pp. 1847-1854.

[Zeng, in press] Accurate and Scalable Surface Representation and Reconstruction from Images. Gang Zeng, Sylvain Paris, Long Quan, and Francois Sillion. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)

Outline

- Context and Basic Ideas
- Consistency and Related Techniques
- Regularized Methods
- Conclusions

Going Underwater

- Main point to adapt: consistency function
  - More robust matching [Zhang, to appear]
  - "Inverting" perturbations [Hermosillo 01, Kim 03]
- Thin features (plants, seaweed…)
- Objects in motion [Pons 05]

[Hermosillo 01] Gerardo Hermosillo,Christophe Chefd'Hotel, Olivier Faugeras, A Variational Approach to Multi-Modal Image Matching. International Journal of Computer Vision (IJCV), volume 50, number 3, November 2002, pages 329-343.

[Kim 03] Visual Correspondence Using Energy Minimization and Mutual Information. Junhwan Kim, Vladimir Kolmogorov and Ramin Zabih. In IEEE International Conference on Computer Vision (ICCV), October 2003

[Pons 05] J-P. Pons, R. Keriven and O. Faugeras. Modelling Dynamic Scenes by Registering Multi-View Image Sequences. In IEEE Conference in Computer Vision and Pattern Recognition, 2005.

[Zhang, to appear] H. Zhang, S. Negahdaripour. Integrating BG and GC models in dense stereo reconstruction with Markov Random Fields. Journal of Multimedia

Conclusions

- 3D reconstruction is a hard problem.
- Solutions exist.
  - Need to be adapted to specific environment.
- Consistency carries information and adds detail.
  - Regularization removes noise and fills holes.
- Start with a simple solution.
  - A complete failure is not a good sign.

References

- These slides: full-length refs in comments
  - ↳ available on my webpage soon
    http://people.csail.mit.edu/sparis/
- This talk has been inspired by
  - ↳ my PhD dissertation [Paris 04]
  - ↳ a recent survey [Seitz 06]

[Paris 04] Extraction of Three-dimensional Information from Images -- Application to Computer Graphics. Sylvain Paris. Ph.D. thesis from Université Joseph Fourier (Grenoble, France) 2004

[Seitz 06] Steve Seitz, Brian Curless, James Diebel, Daniel Scharstein, Rick Szeliski, "A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms", CVPR 2006, vol. 1, pages 519-526.

Thank you

I am grateful to Shahriar Negahdaripour and Donna Kocak for inviting me.

My work on 3D reconstruction has been made in collaboration with Zeng Gang, Long Quan, and François Sillion.

I am thankful to Frédo Durand, my host at MIT. My research at MIT is supported by a grant from Royal Dutch/Shell Group.