碩士論文

宋秉一

指導教授：陳稔 教授

以粒子群最佳化技術重建稠密式物體三維模型

Dense 3D Reconstruction with Particle swarm Optimization

NCTU PAIS Lab

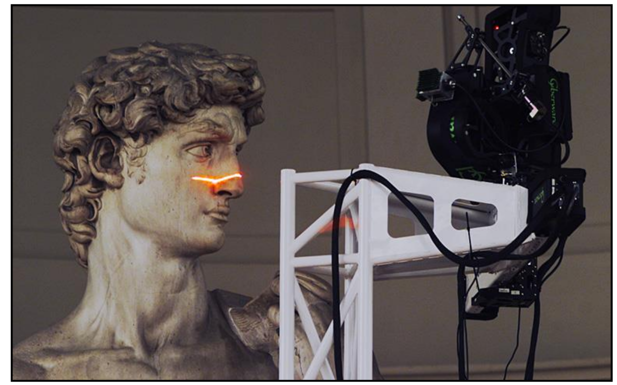
Outline

1. Introduction
2. Patch Optimization
3. Particle swarm Optimization
4. Patch Expansion
5. Patch Filtering
6. Experiment Results
7. Conclusion and Future Work

1. INTRODUCTION

Traditional 3D reconstruction

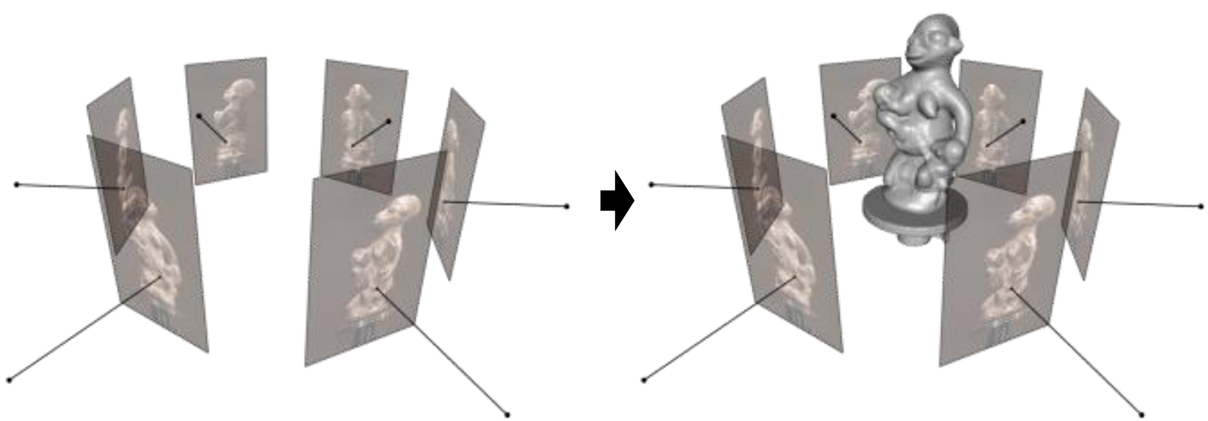
- Laser scanner
- Structured light
  - Very accurate
  - Very expensive
  - Complicated to use



Carlos Hernández, George Vogiatzis, Yasutaka Furukawa.
3d shape reconstruction from photographs: a Multi-View Stereo approach. CVPR2010.

Multi-view stereo

- Photograph
  - Easy to take
  - Very cheap device
  - Inaccurate
  - Sensitive to illumination

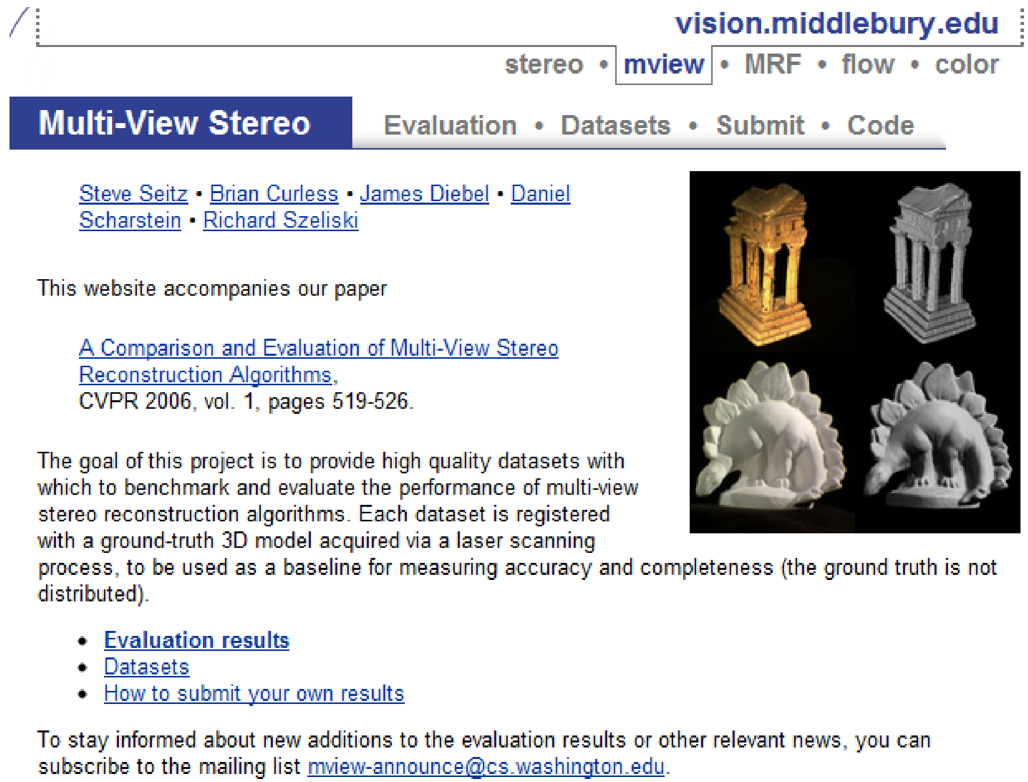


Carlos Hernández, George Vogiatzis, Yasutaka Furukawa.
3d shape reconstruction from photographs: a Multi-View Stereo approach. CVPR2010.

Benchmark

- Middlebury

Ground Truth



vision.middlebury.edu

stereo • mview • MRF • flow • color

Multi-View Stereo Evaluation • Datasets • Submit • Code

Steve Seitz • Brian Curless • James Diebel • Daniel Scharstein • Richard Szeliski

This website accompanies our paper

A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms,
CVPR 2006, vol. 1, pages 519-526.

The goal of this project is to provide high quality datasets with which to benchmark and evaluate the performance of multi-view stereo reconstruction algorithms. Each dataset is registered with a ground-truth 3D model acquired via a laser scanning process, to be used as a baseline for measuring accuracy and completeness (the ground truth is not distributed).

- **Evaluation results**
- Datasets
- How to submit your own results

To stay informed about new additions to the evaluation results or other relevant news, you can subscribe to the mailing list mview-announce@cs.washington.edu.

Competitive Result

Multi-View Stereo | **Evaluation** • Datasets • Submit • Code

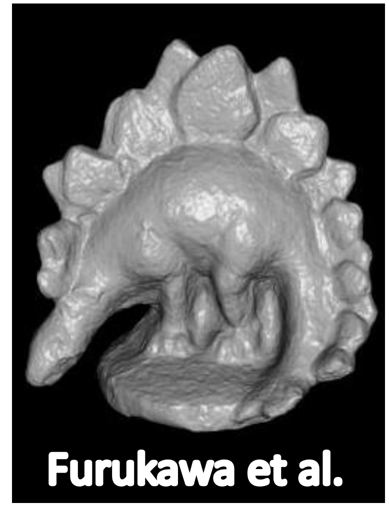
Acc. Threshold: 90% Comp. Threshold: 1.25 mm

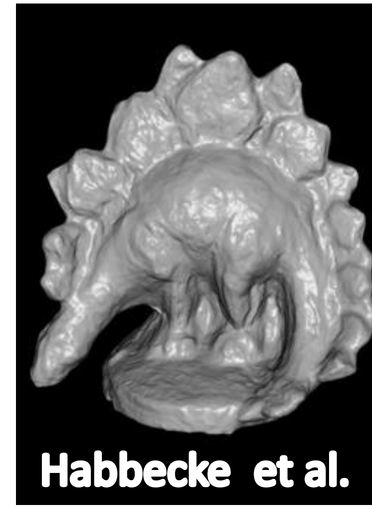
Data in new window [ ] Open Data Window

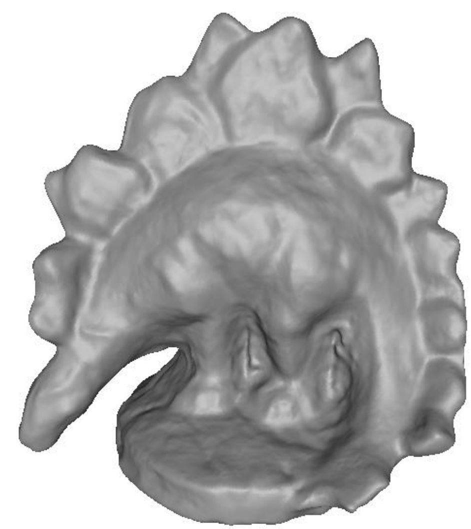
Data: View 1 and Ground Truth Image Size Small

Tip: Mousing over any portion of a method's row will show its reference

Sort By	Temple Full 312 views Acc [mm]	Comp [%]	Temple Ring 47 views Acc [mm]	Comp [%]	Temple Sparse 16 views Acc [mm]	Comp [%]	Dino Full 363 views Acc [mm]	Comp [%]	Dino Ring 48 views Acc [mm]	Comp [%]	Dino Sparse 16 views Acc [mm]	Comp [%]
Furukawa 2	0.54	99.3	0.55	99.1	0.62	99.2	0.32	99.9	0.33	99.6	0.42	99.2
Furukawa 3	0.49	99.6	0.47	99.6	0.63	99.3	0.33	99.8	0.28	99.8	0.37	99.2
DAGM_68	0.57	99.1	0.64	96.4	2.12	62.9	0.33	99.7	0.33	99.7	0.54	98.6
Guillemaut	0.43	99.0	0.71	97.6	0.86	96.2	0.35	100	0.58	99.5	0.68	98.0
Vogiatzis	1.07	90.7	0.76	96.2	2.77	79.4	0.42	99.0	0.49	96.7	1.18	90.8
Habbecke	0.66	98.0					0.43	99.7				
Goesele 2007	0.42	98.2					0.46	96.7				
Lambert3	0.48	99.7					0.48	99.4				
Hernandez	0.36	99.7	0.52	99.5	0.75	95.3	0.49	99.6	0.45	97.9	0.6	98.5
Furukawa	0.65	98.7	0.58	98.5	0.82	94.3	0.52	99.2	0.42	98.8	0.58	96.9
ECCV_279	0.53	99.4	0.81	95.8			0.55	98.1	0.6	96.0		
Zach2	0.51	98.8	0.56	99.0			0.55	98.7	0.51	99.1		
Goesele	0.42	98.0	0.61	86.2	0.87	56.6	0.56	80.0	0.46	57.8	0.56	26.0
Hongxing	0.83	95.7	0.79	96.3	0.97	93.9	0.62	96.3	0.5	99.1	0.52	98.4
Lambert	0.55	99.7					0.76	99.0				
Hornung	0.58	98.7					0.79	95.1				
Jancosek-3DIM09	0.65	85.8	0.7	78.9	0.59	74.9	0.91	73.8	0.71	76.6	0.66	74.9



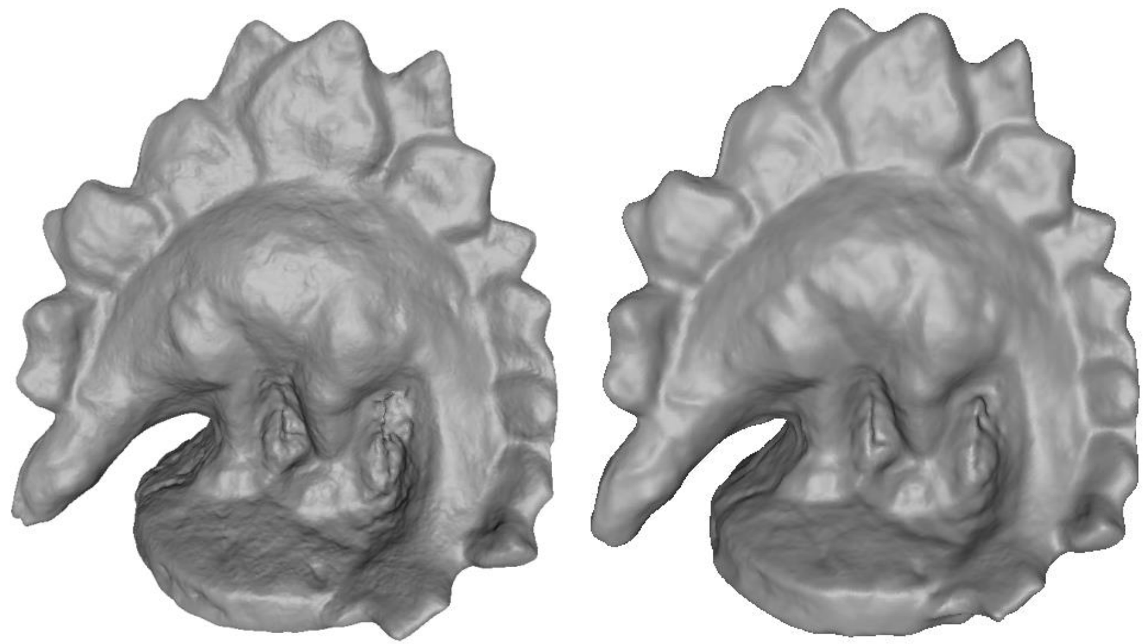




Ours

Competitive Result (Cont.)

Habbecke et al.



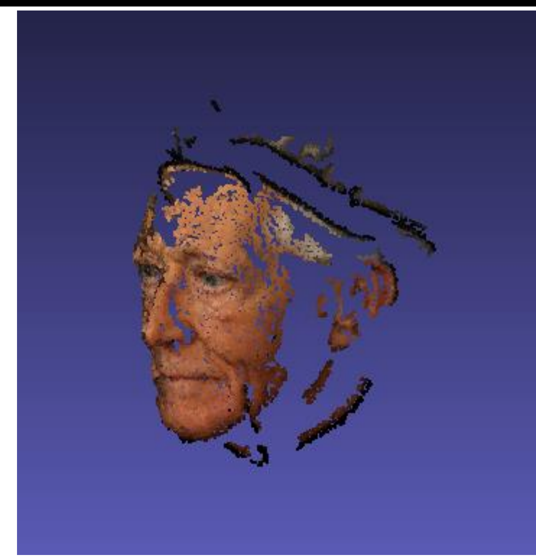
Ours

M. Habbecke and L. Kobbelt, "A surface-growing approach to multi-view stereo reconstruction," in Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2007.

NCTU PAIS Lab

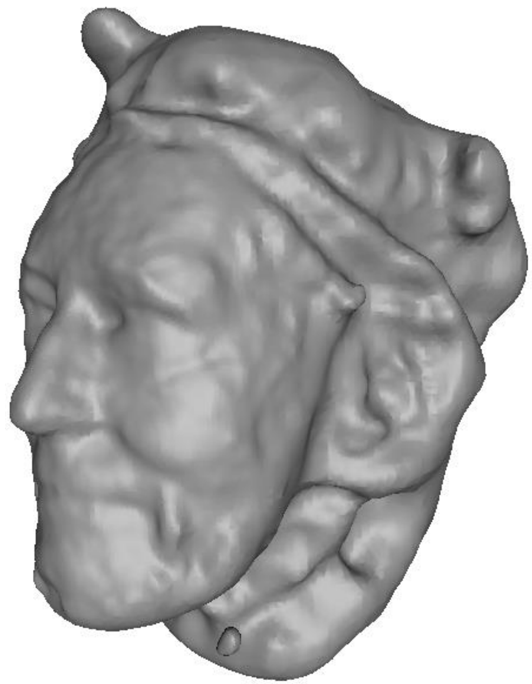
Competitive Result (Cont.)

- Completeness

Ours	Furukawa et al.
	

Competitive Result (Cont.)

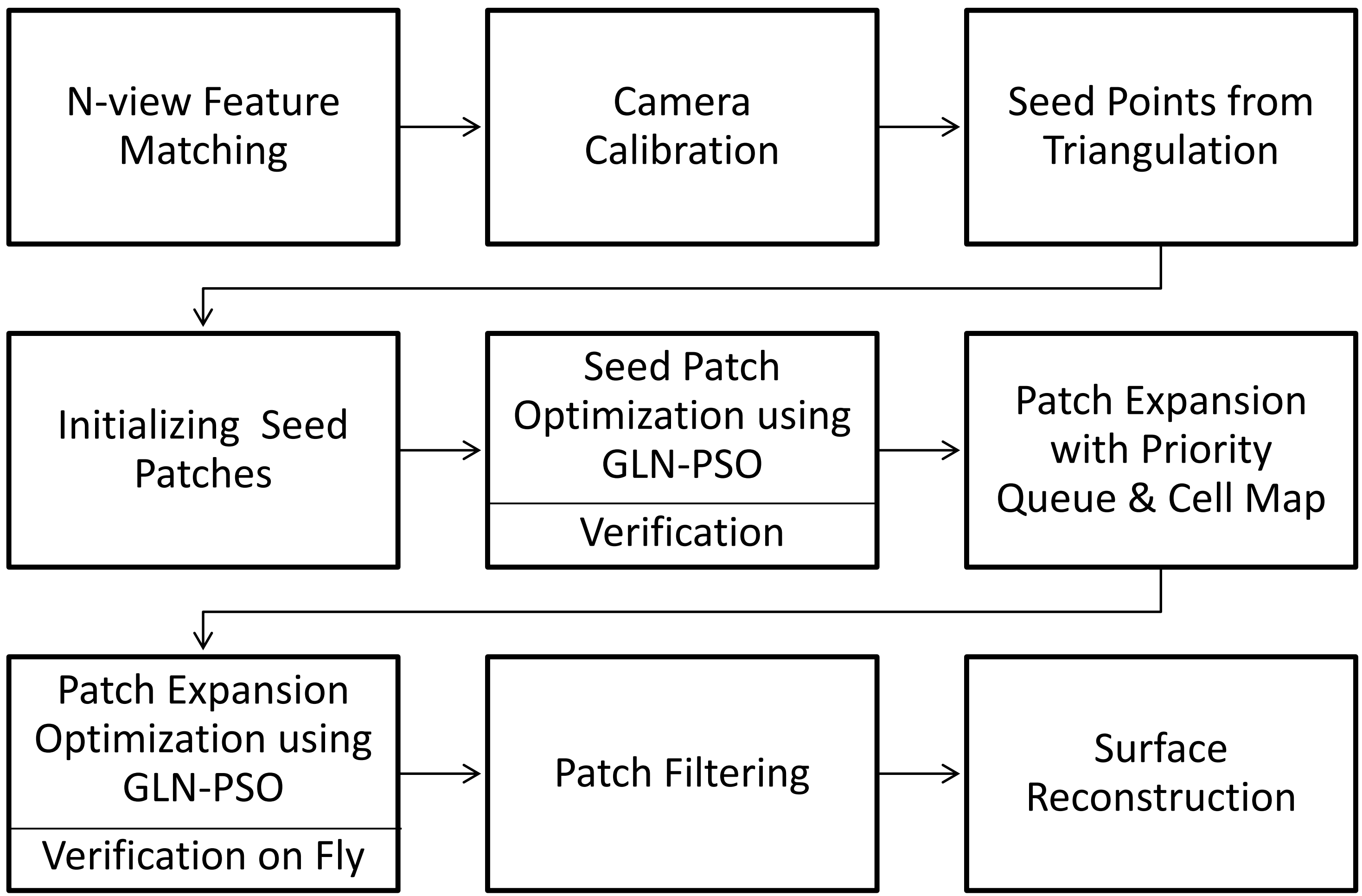
Ours



Furukawa et al.

NCTU PAIS Lab

Proposed system overview



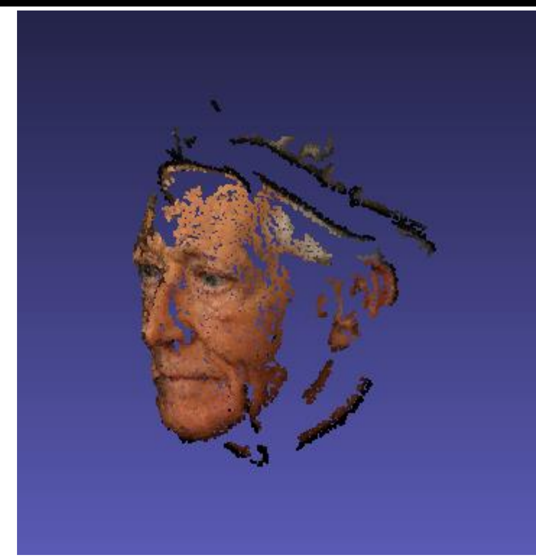
Contributions

- Accurate patches
  - Global optimization with GLN-PSO
  - Stabilization without derivative computation
- Non-uniform view reconstruction
  - More general object reconstruction
- Level of details
  - Texture variation based pyramid image scales
  - Textured/textureless/repetitive patterns
- Adaptive fitness weighting
  - Sharp patch slope change
  - Bilateral or trilateral filtering (pixel distance, texture difference, and edge magnitude)
- Expansion strategy with priority
  - Best-first vs. breadth-first
  - Reliable surface growing

Non-uniform view reconstruction

Non-uniform view reconstruction (Cont.)

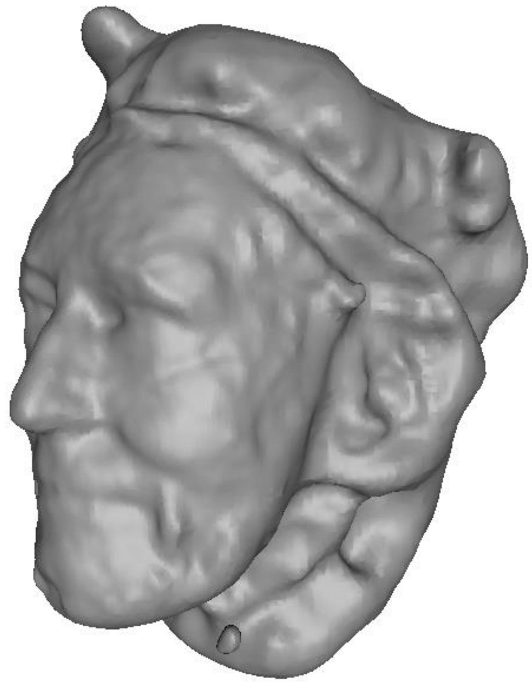
- Completeness

Ours	Furukawa et al.
	

Non-uniform view reconstruction (Cont.)

- Reconstruction result

Ours



Furukawa et al.

Level of detail

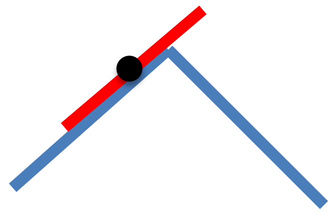
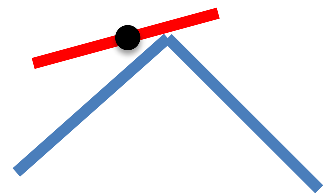
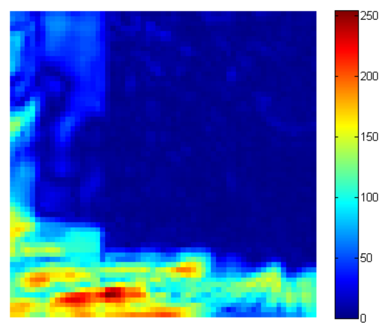
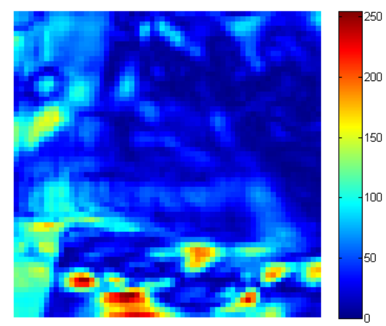
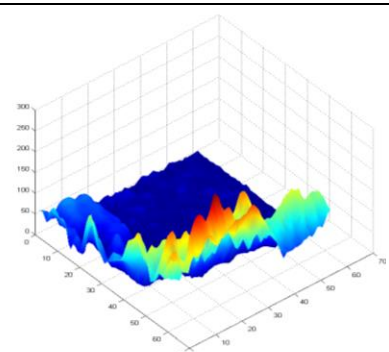
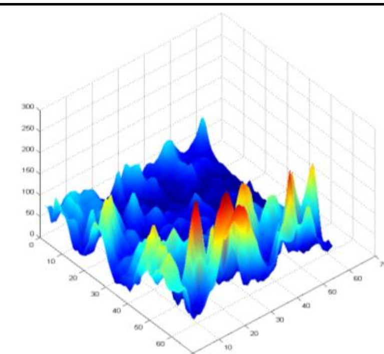
- Textureless model

Scale ratio $\varepsilon = 0.8$

$l(p) = 0$ $l(p) = 1$ $l(p) = 2$

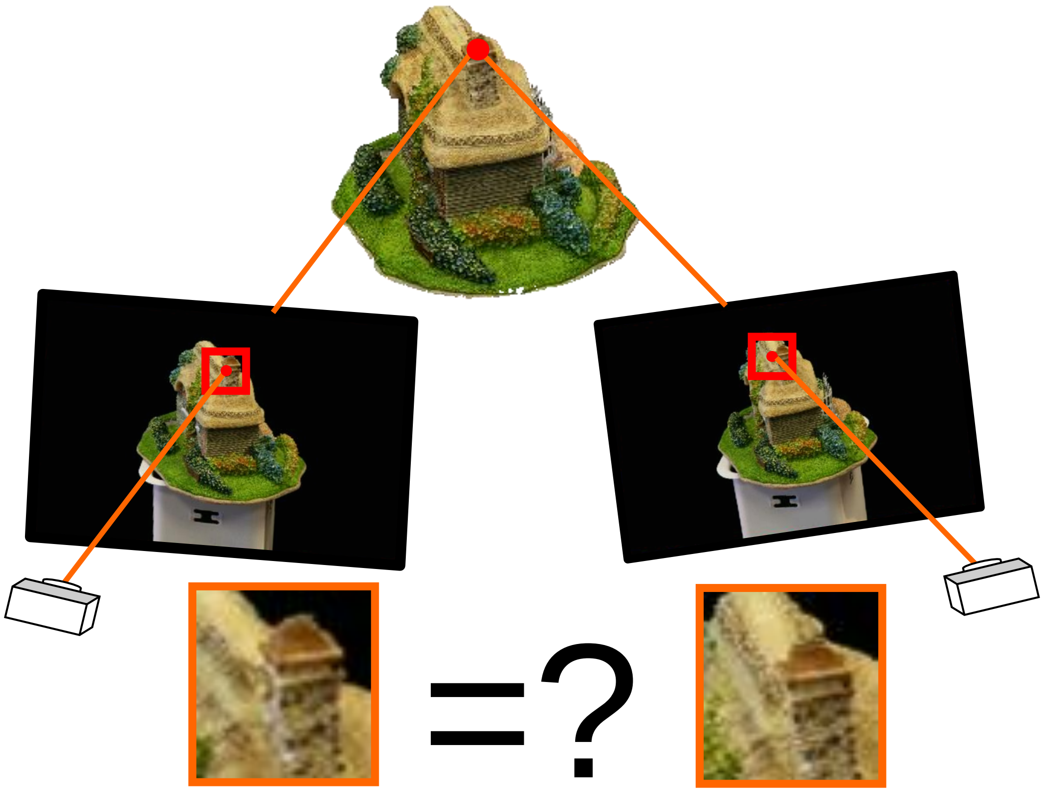
PAIS Lab

Adaptive fitness weighting

	With adaptive weighting (ours)	Without adaptive weighting (Furukawa's)
重建結果示意圖		
normalized pixel-wised sum of mean difference		
		

2. PATCH OPTIMIZATION

Seed point from feature matching

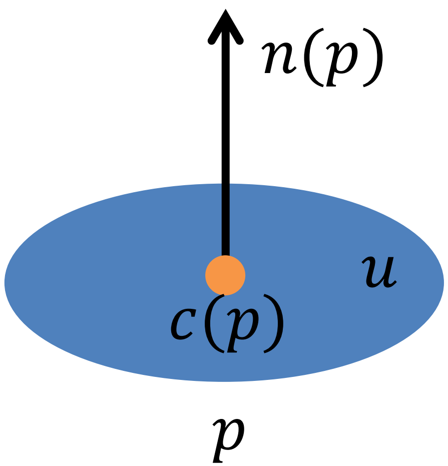


Carlos Hernández, George Vogiatzis, Yasutaka Furukawa.
3d shape reconstruction from photographs: a Multi-View Stereo approach. CVPR2010.

NCTU PAIS Lab

Patch Definition

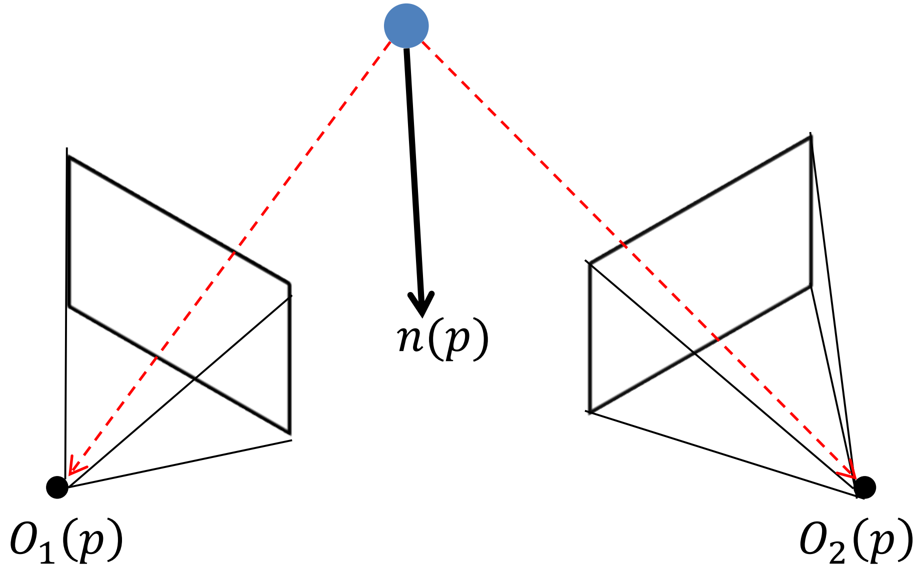
- Center $c(p)$
- Normal $n(p)$
- Extent u



NCTU PAIS Lab

Patch Normal

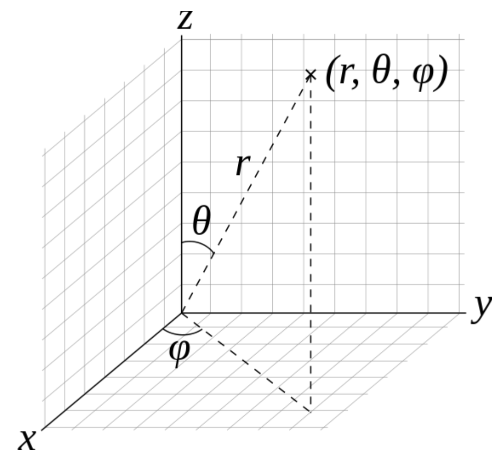
- Patch Normal $n(p)$
  - $n(p) = \frac{1}{|V(p)|} \sum_{i \in V(p)} \frac{O_i(p) - c(p)}{\|O_i(p) - c(p)\|}$



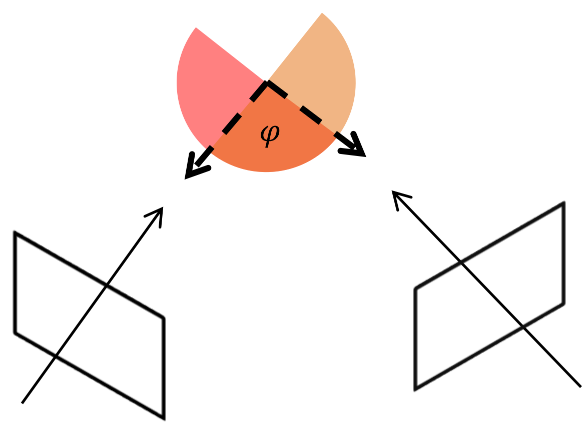
Patch Normal Range

- Spherical Coordinate
  - θ, φ
- Range of θ: $[0, \pi]$
- Range of φ
  - Intersection of viewing cone of image in $V(p)$

$$\varphi_{range} = \bigcap_{i \in V(p)} \varphi_i(p)$$



Wikipedia: Spherical coordinate system

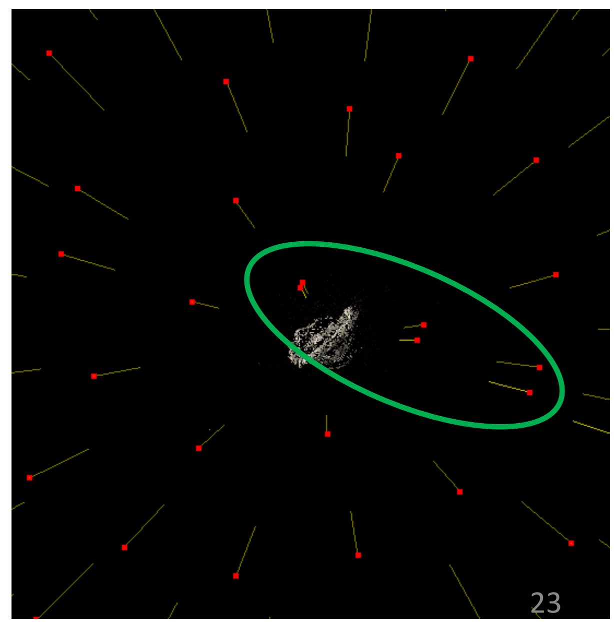


NCTU PAIS Lab

Patch Depth Range (cont.)

- Deal with narrow baseline views
  - Erroneous huge range (-DBL_MAX ~ DBL_MAX)
  - Difficult to convergence with a correct result
  - Remove views with small projection distance

$$\left\| P_i\big(c(p) + r(p)\big) - P_i\big(c(p)\big) \right\| < 1$$



Top view from Middlebury dino dataset

Level of Detail

- Deal with
  - Textureless images.
  - High-quality images. (over 10M pixels)

Middlebury dino dataset
(640x480) pixels

T. Beeler, B. Bickel, P. Beardsley, R. Sumner, M. Gross. High-Quality Single-Shot Capture of Facial Geometry. Proceedings of ACM SIGGRAPH, July 25-29. 2010.
(4000x3000) pixels

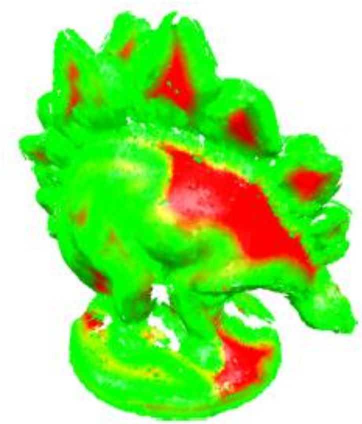


Figure 2. Variance-driven determination of disk size. Red colored parts of the surface are approximated by large disks due to little image texture, while green color depicts small disks.

M. Habbecke and L. Kobbelt, "A surface-growing approach to multi-view stereo reconstruction," in Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2007.

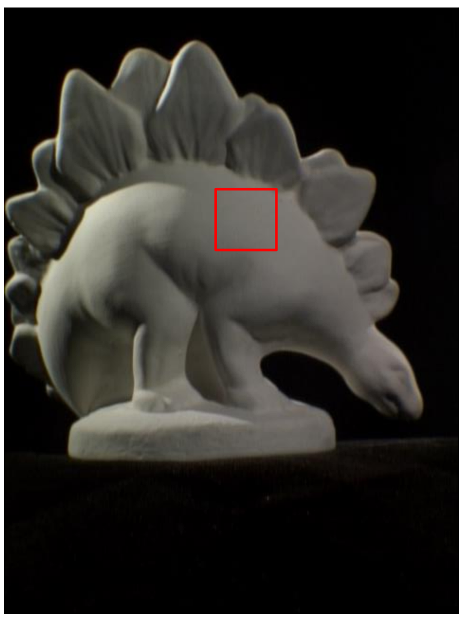
Level of Detail (cont.)

- Patch Extent size:
  - Furukawa et al: small(7 or 9) & trapped
    - filtering & re-expansion
  - Habbecke et al: large(100 ~ 2000) & inefficiency
  - Ours: fixed size (15 or 31) with pyramid sampling

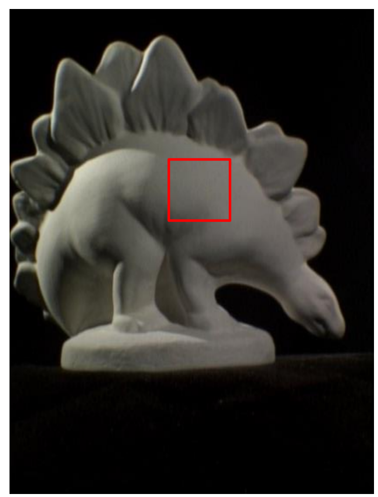
Level of Detail (cont.)

- Level of Detail $l(p)$

Scale ratio $\varepsilon = 0.8$



$l(p) = 0$



$l(p) = 1$

$l(p) = 2$

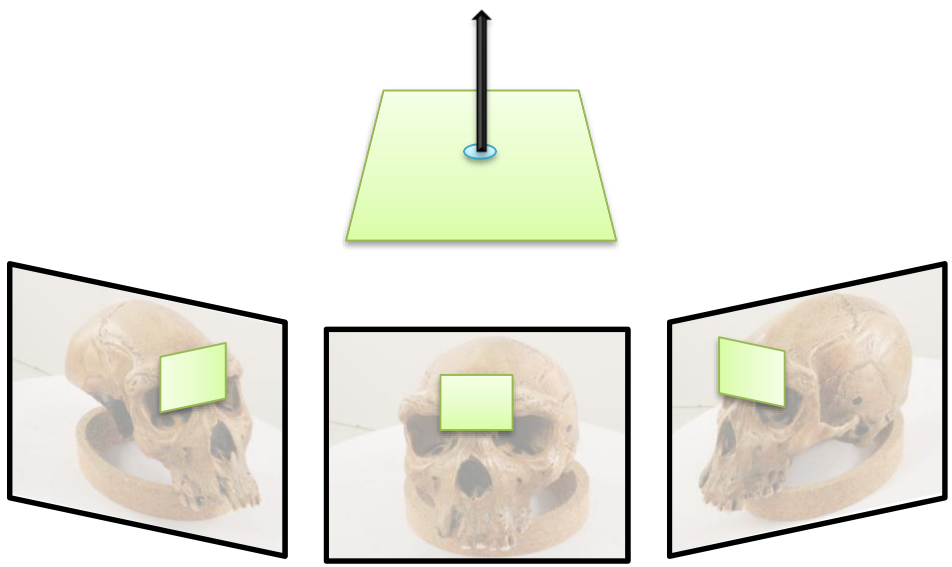
Patch optimization

- Optimization (DOF 3)
  - Depth $d(p)$
    - 1D distance to reference camera center
    - Optimized patch center
      $$c(p) = d(p) * r(p) + O_{R(p)}$$
  - Normal θ, φ
    - 2D spherical normal

NCTU PAIS Lab

Patch optimization (cont.)

- Minimize projected texture difference

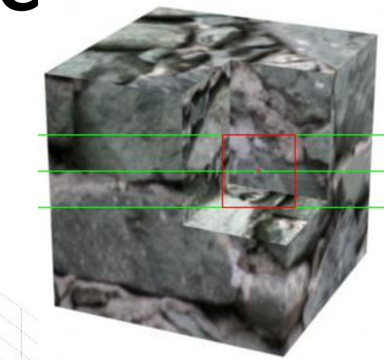


Visible images $V(p)$

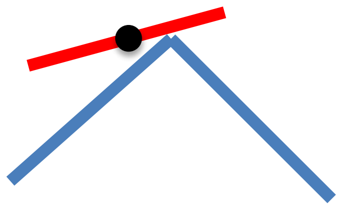
Carlos Hernández, George Vogiatzis, Yasutaka Furukawa.
3d shape reconstruction from photographs: a Multi-View Stereo approach. CVPR2010.

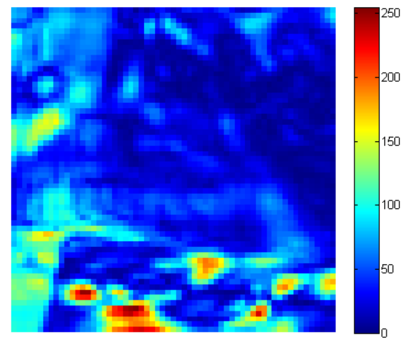
NCTU PAIS Lab

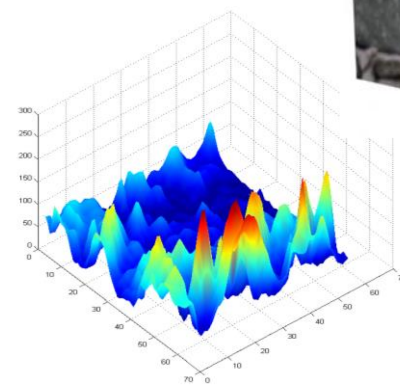
Adaptive Fitness Weighting (cont.)



- Cross planar patch extent
  - Averaging error

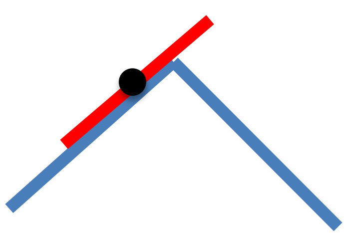


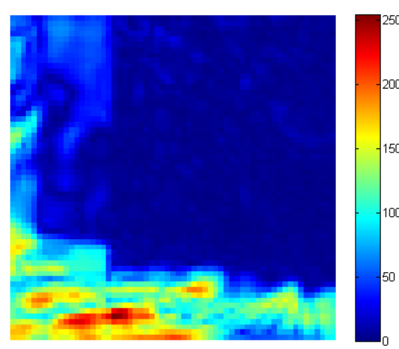


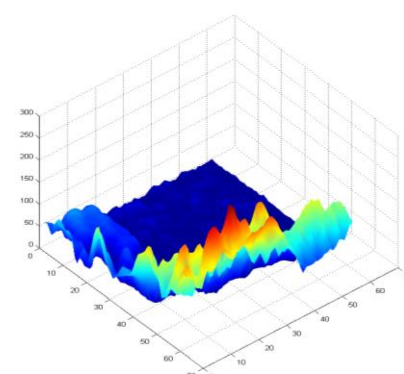


- After adapted fitness weighting

$$w(x,y) = \frac{g(x,y)k(x,y)h(x,y)}{\sum_{u,v \in p} g(u,v)k(x,y)h(u,v)}, (x,y) \in \Omega_l(p)$$







Optimization Strategy

- Inserting initial patch parameter to particle pool
  - Estimated $n(p)$
  - Estimated $d(p)$

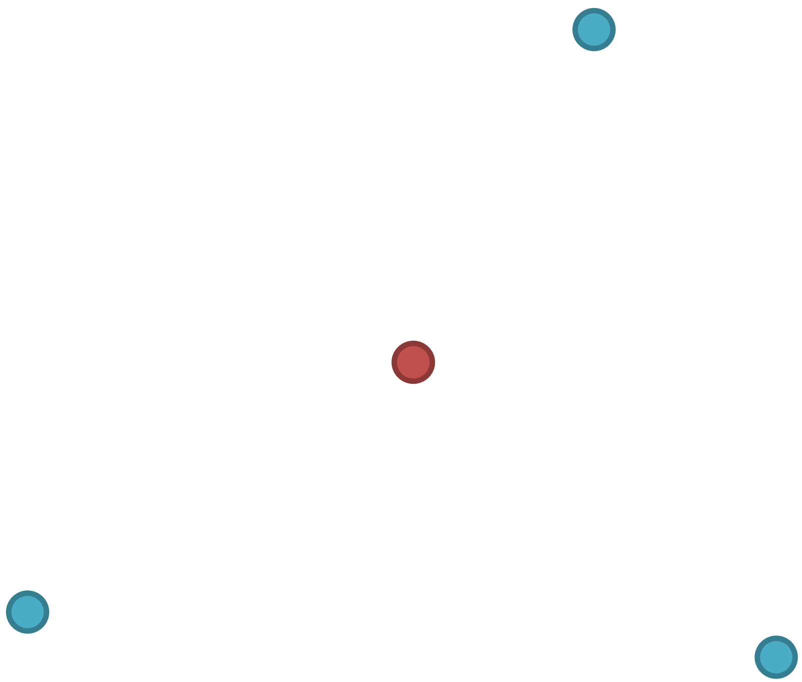
Type	Average Fitness
GLN-PSO w/o initial particle	5.0581
GLN-PSO with initial particle	**4.1242**
PSO w/o initial particle	5.5116
PSO with initial particle	4.3865

Pawn dataset **10266** seed patch optimization with **60** iteration

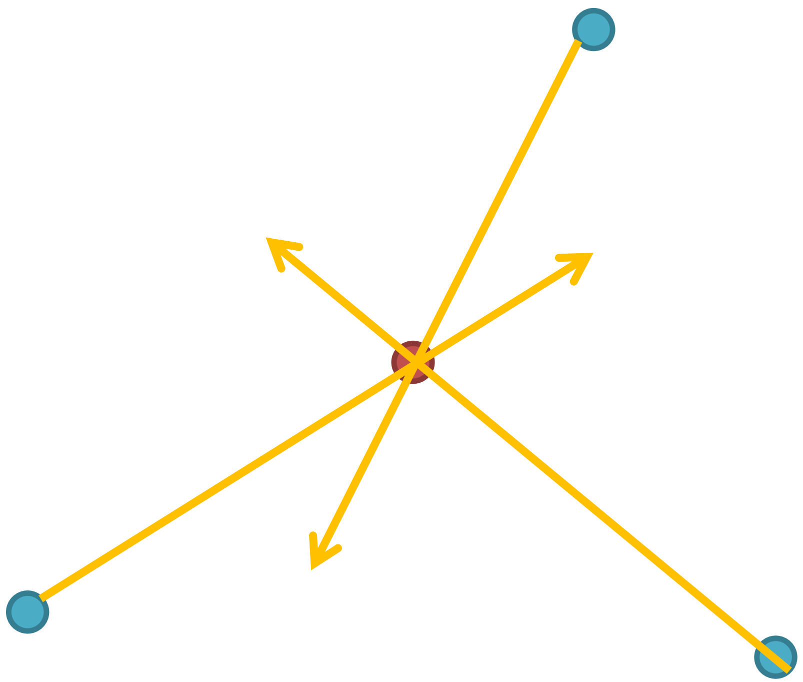
NCTU PAIS Lab

3. PARTICLE SWARM OPTIMIZATION

PSO Example

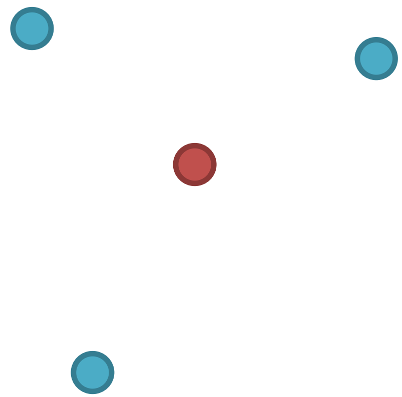


PSO Example (cont.)

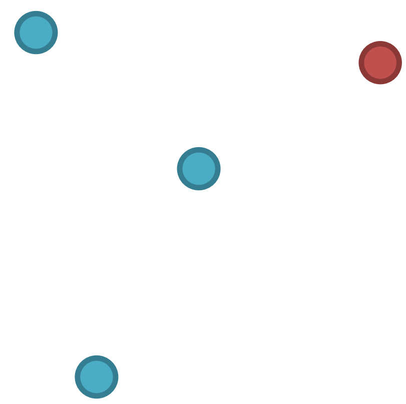


NCTU PAIS Lab

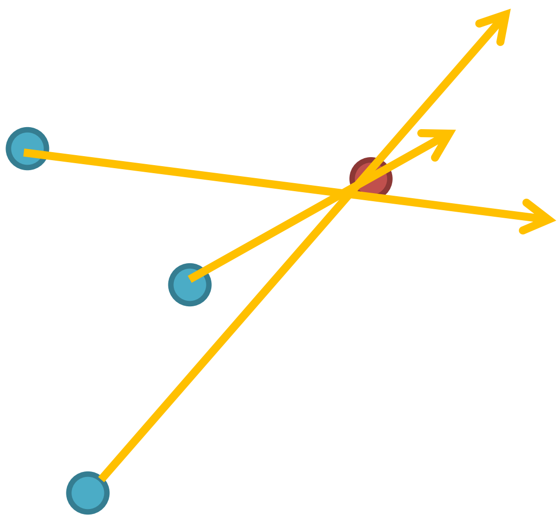
PSO Example (cont.)



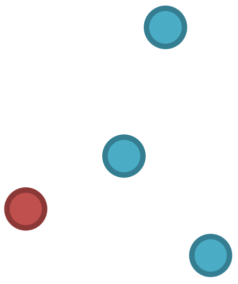
PSO Example (cont.)



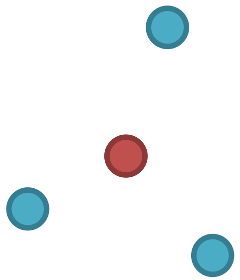
PSO Example (cont.)



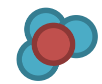
PSO Example (cont.)



PSO Example (cont.)



PSO Example (cont.)



convergence

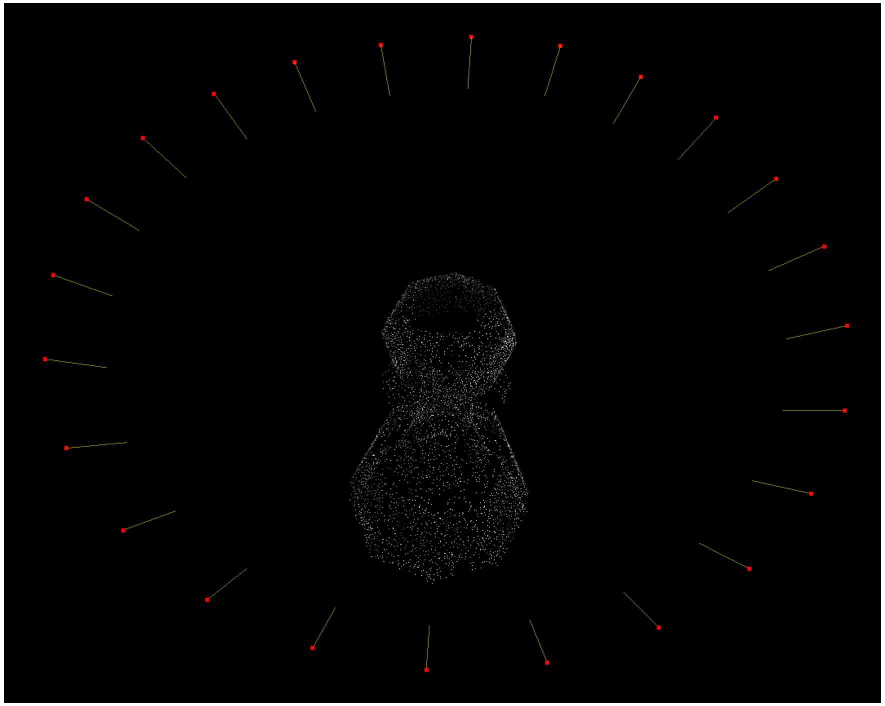
6. EXPERIMENTS

Experiments

- Reconstruction of Synthesis Images
  - Maya rendering images
- Reconstruction of Real Images
  - Middlebury dataset
    - Dino
    - Template
  - Human Face dataset

Reconstruction of Synthesis Images

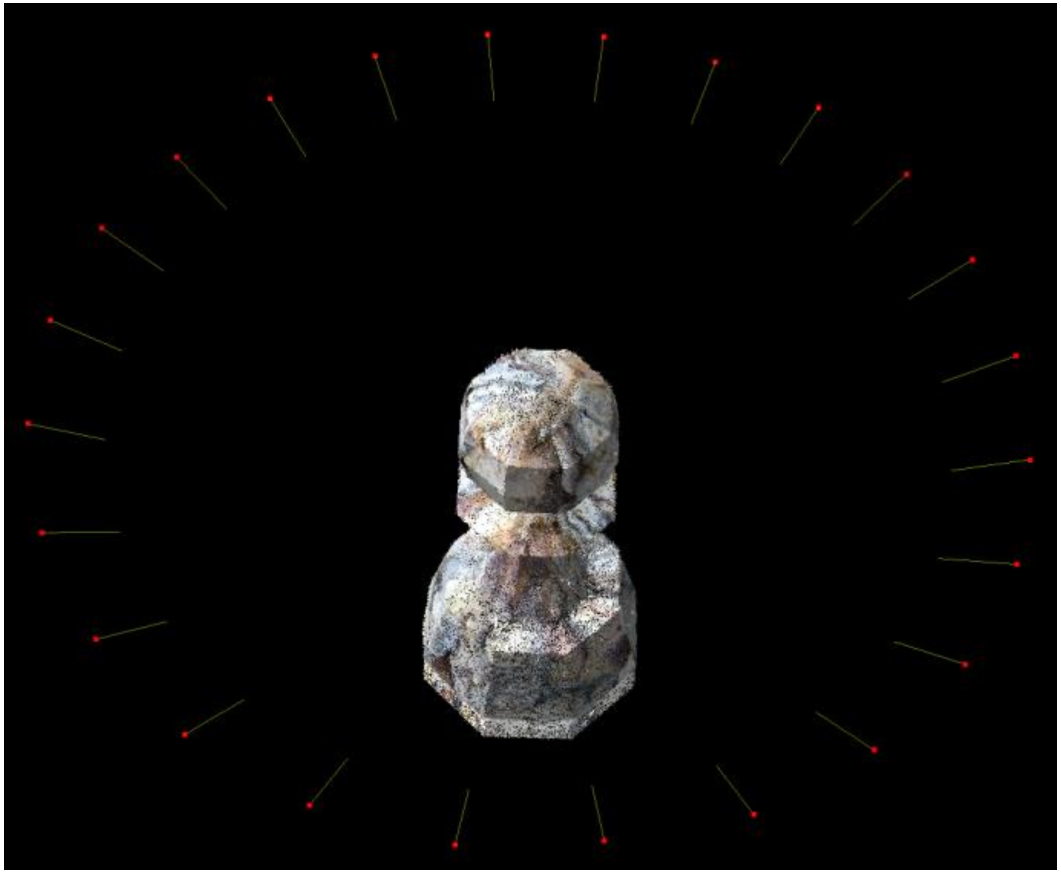
- Reconstruction pawn model
  - 24 ring view images with silhouette mask
  - 10266 seed points
  - 6490 seed patches after verification



PAIS Lab

Reconstruction of Synthesis Images (Cont.)

- Reconstruction patch model (95068 patches)



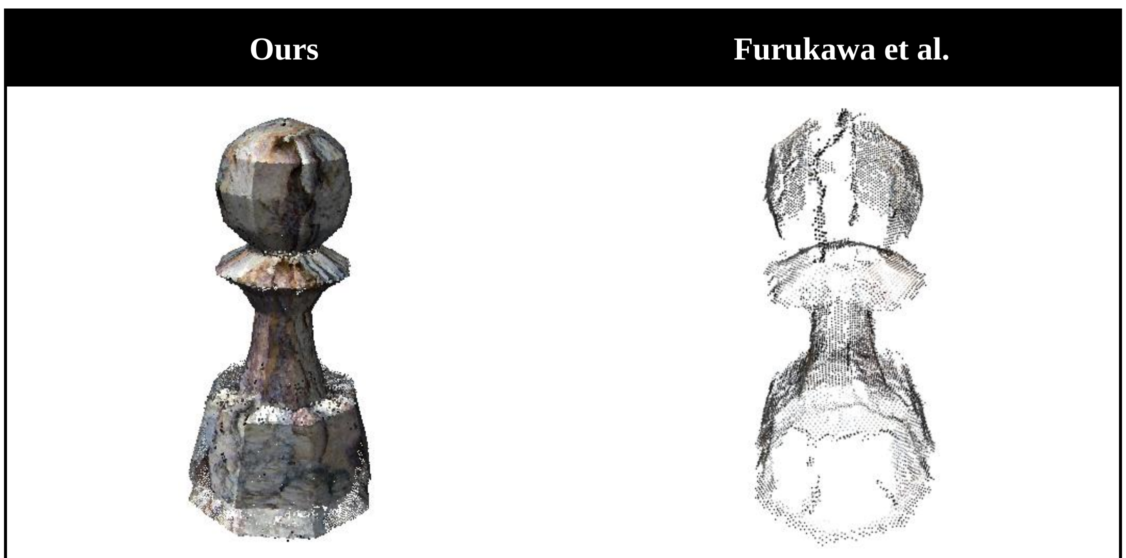
8.1 Reconstruction of Synthesis Images (Cont.)

- Compare with Furukawa et al. (PMVS)
  - Same cell size

$$\tau = 2$$

  - Same minimum visible camera number

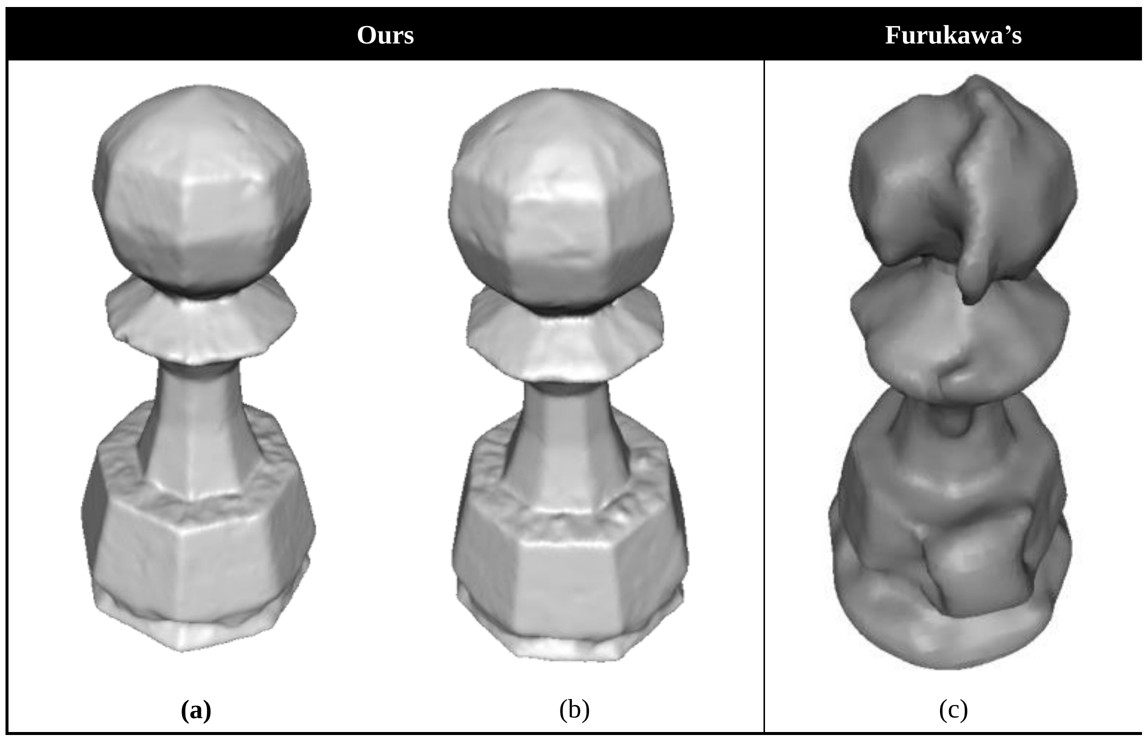
$$V_{min} = 3$$



NCTU PAIS Lab

Reconstruction of Synthesis Images (Cont.)

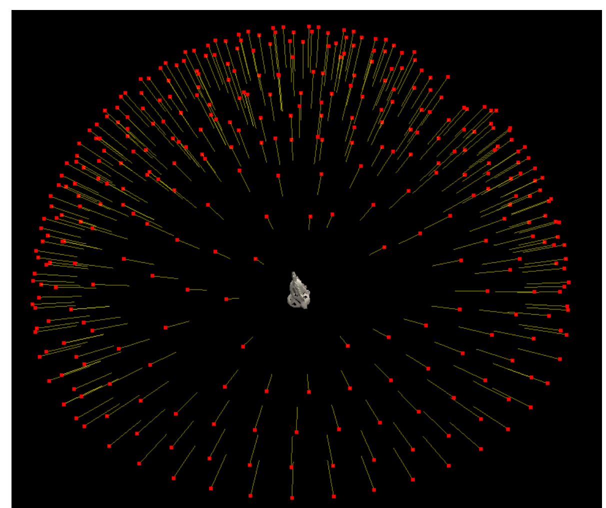
- Poisson Surface Reconstruction result



NCTU PAIS Lab

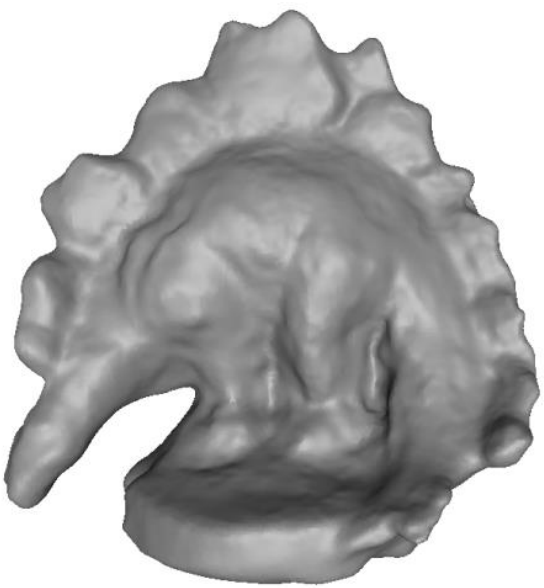
Reconstruction of Middlebury Dinosaur Model

- Middlebury Dino dataset
  - 363 domed views
  - Automatic silhouette extraction
  - 6057 seed points
  - 1872 seed patches after verification

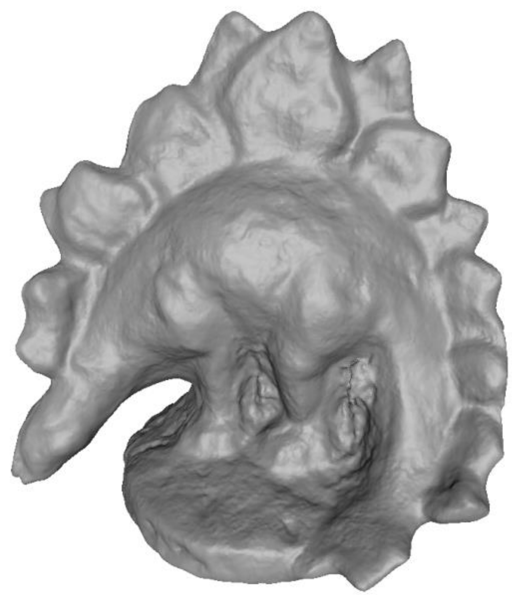


Reconstruction of Middlebury Dinosaur Model (Cont.)

- Poisson Surface Reconstruction result



Cell size 4 (41922 patches)

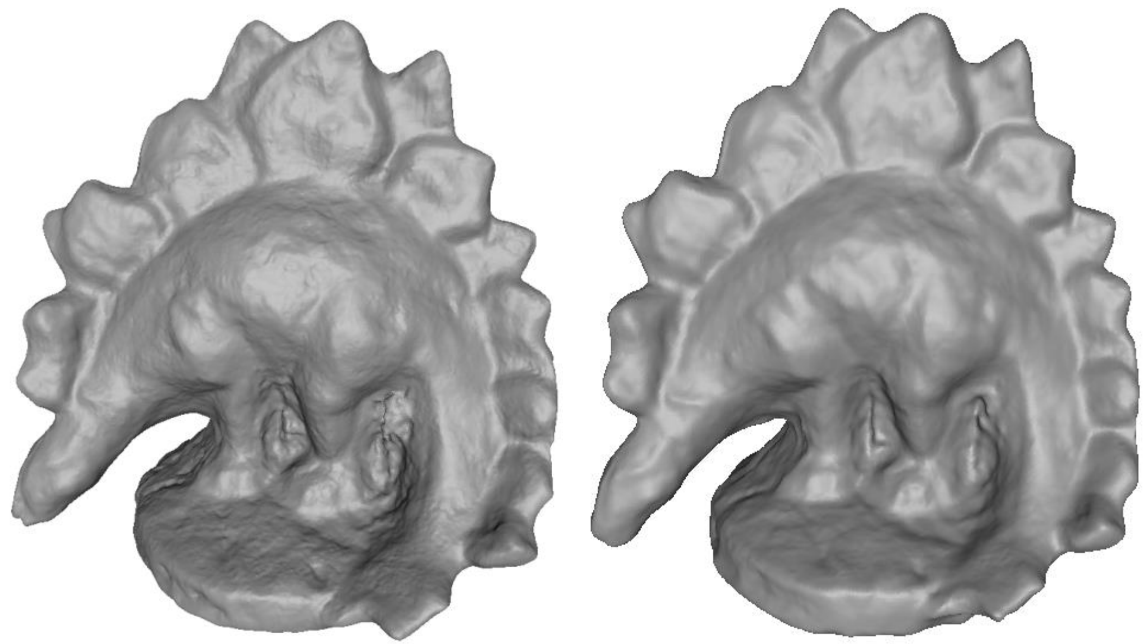


Cell size 2 (210269 patches)

Ground Truth

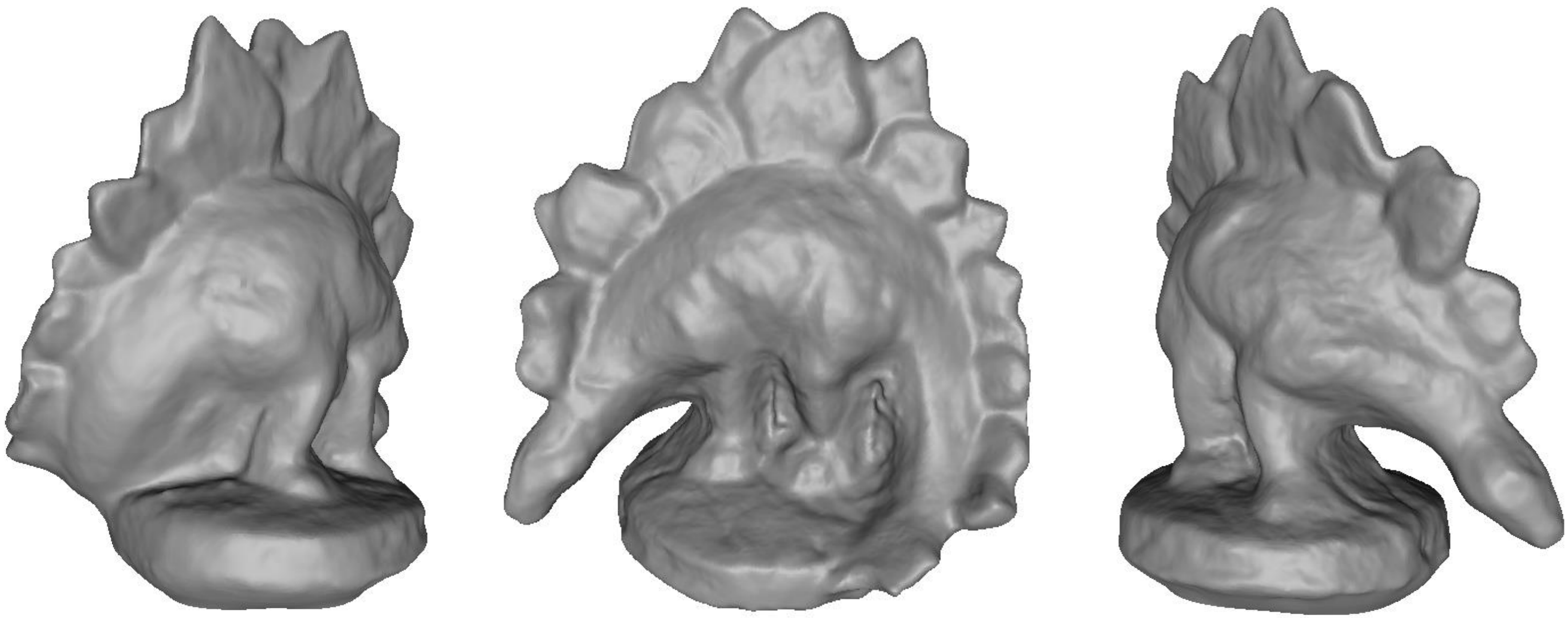
Reconstruction of Middlebury Dinosaur Model (Cont.)

Habbecke's



Ours

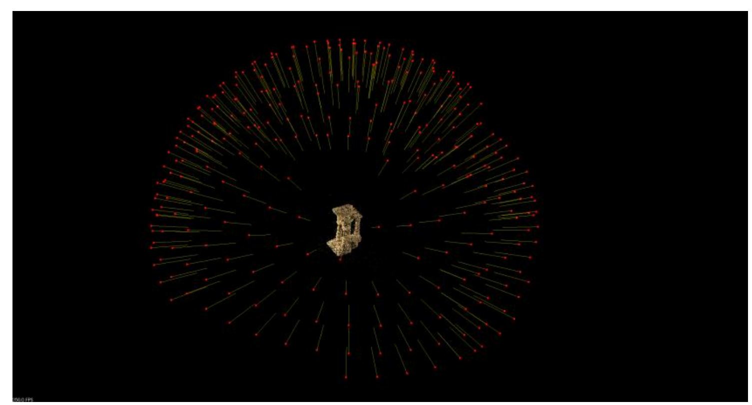
Reconstruction of Middlebury Dinosaur Model (Cont.)



Reconstruction of Middlebury Temple Model

- Middlebury Temple dataset
  - 312 domed views
  - Automatic silhouette extraction
  - 40100 seed points
  - 12634 seed patches after verification



Reconstruction of Middlebury Temple Model (Cont.)

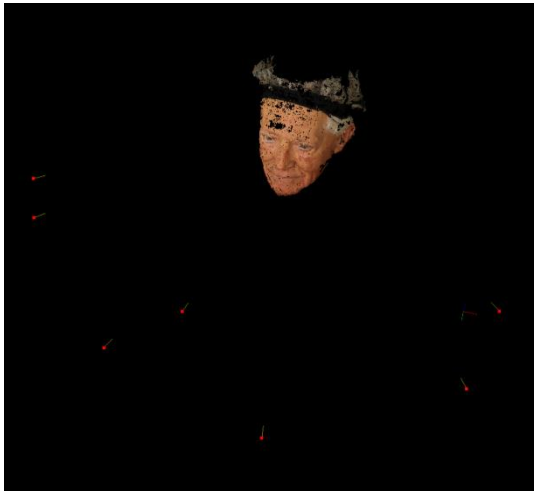
- Poisson Surface Reconstruction result

Cell size 4 (29489 patches)

Cell size 2 (138760 patches)

Reconstruction of Human Face Model

- Human Face dataset
  - 7 views
  - 1687 seed points
  - 1003 seed patches after verification

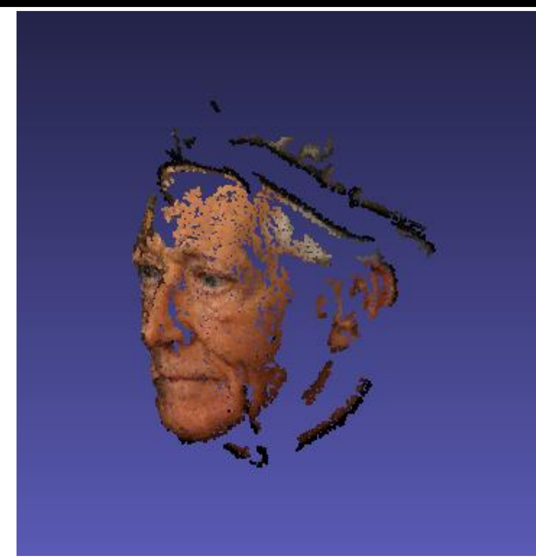


NCTU PAIS Lab

T. Beeler, B. Bickel, P. Beardsley, R. Sumner, M. Gross. High-Quality Single-Shot Capture of Facial Geometry. Proceedings of ACM SIGGRAPH, July 25-29. 2010.

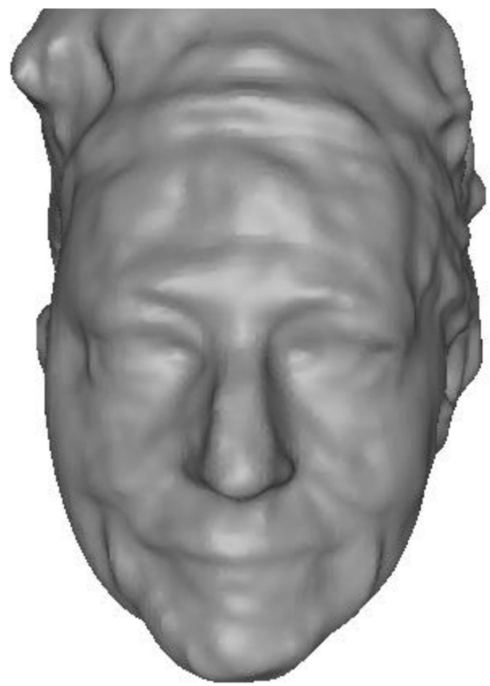
Reconstruction of Human Face Model (Cont.)

- Reconstruction patch model (104115 patches)

Ours	Furukawa et al.
	

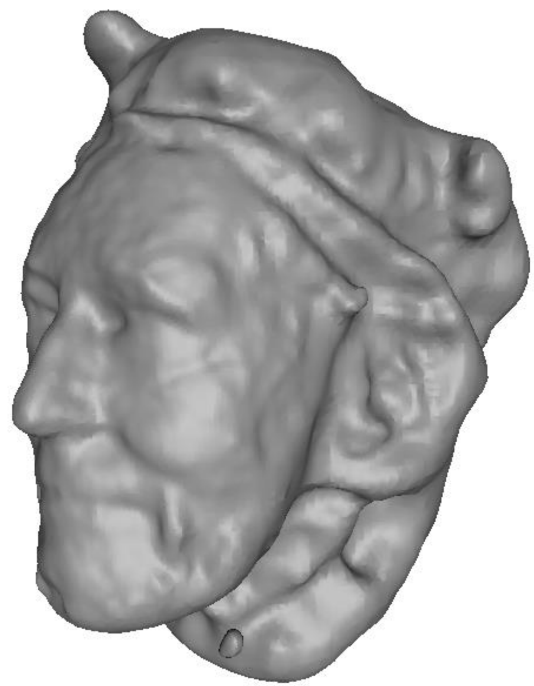
Reconstruction of Human Face Model (Cont.)

- Poisson Surface Reconstruction result



Reconstruction of Human Face Model (Cont.)

Ours



Furukawa’s

Algorithm Comparison

- Comparison

Method	Furukawa et al.	Habbecke et al.	Ours
Image Seed Point	Harris DoG	Free pixel	**Feature descriptor**
Seed Point Matching	Epipolar line	2D Image Homography	**Descriptor & Epipolar line**
Patch Fitness Weighting	Average	Average	**Adapted weighting**
Camera View Baseline	Uniform	Uniform	**un-uniform**
Optimization	Conjugate Gradient	Gradient	**GLN-PSO**
Image sampling	Scaled image	Original	**Pyramid image (LOD)**
Window size	Fixed size small window	Adapted size window	Fixed size small window

7. FUTURE WORKS

Future works

- Too many parameters
  - Over 20 parameters
  - Difficult to estimate the effect for single parameter
- Better feature detector and descriptor
  - Strong perspective correction
- PSO is so slow.
  - CUDA implementation.

Thank You

- Check our MVS source code in the website
  - http://code.google.com/p/pais-mvs/
  - MVS_Viewer
  - MVS_Animation
  - TMVS