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

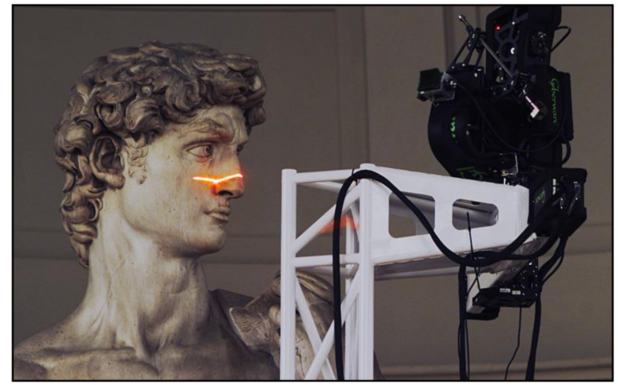
1. INTRODUCTION

Purpose

- 3D reconstruction for
  - Object recognition
  - Movies
  - Virtual reality
  - Heritage preservation
  - Medical research
  - Navigation
  - Etc…

Traditional 3D reconstruction

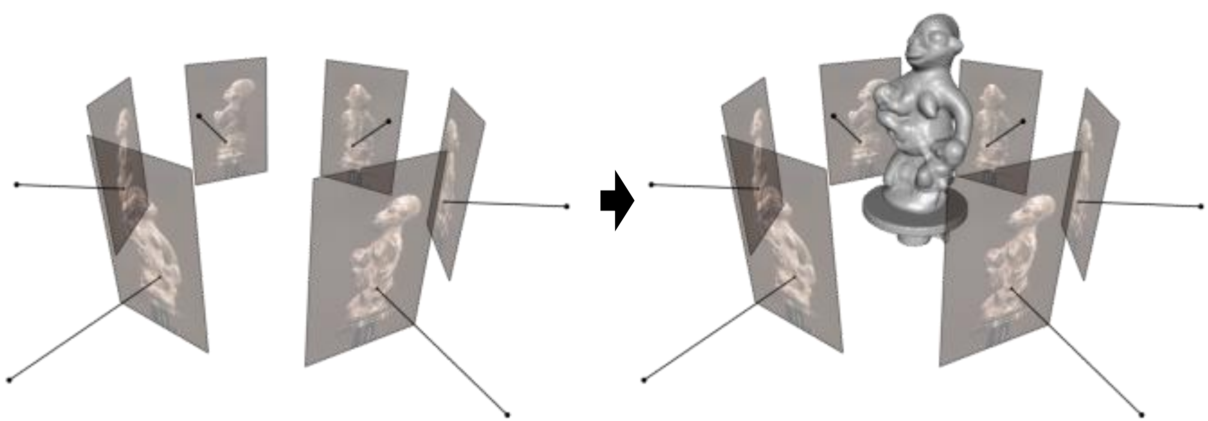
- Laser scanner
- Structured light
  - Very accurate
  - Very expensive
  - Complicated to use



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

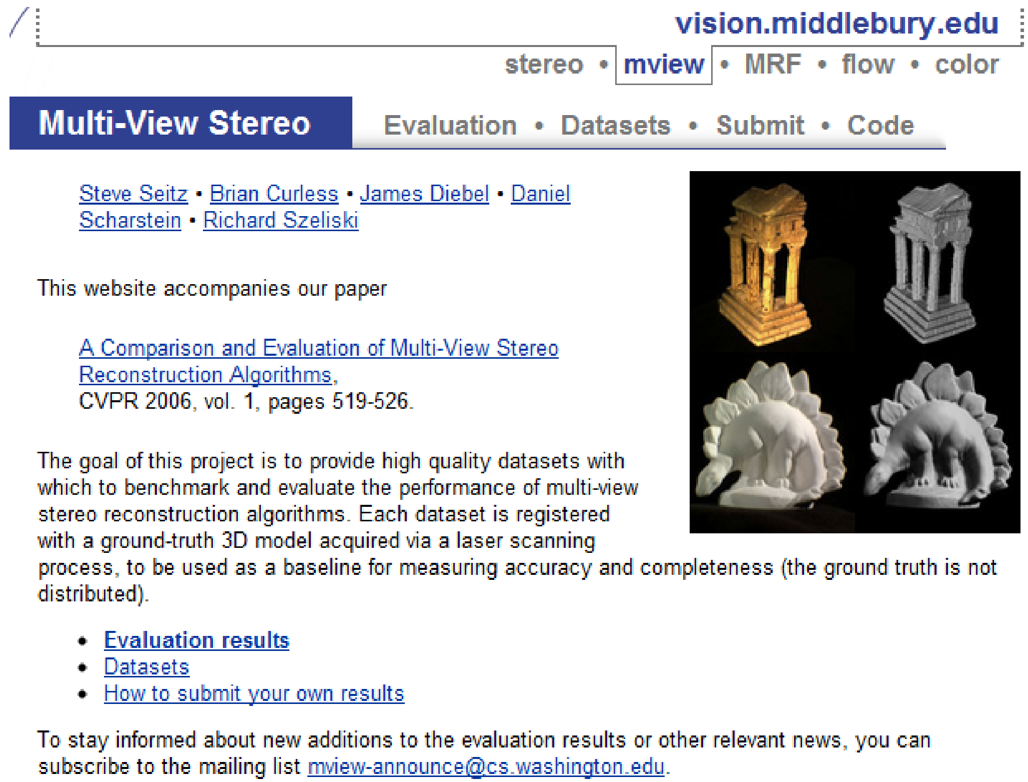


Carlos Hernández, George Vogiatzis, Yasutaka Furukawa.
3d shape reconstruction from photographs: a Multi-View Stereo approach. CVPR2010.

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.

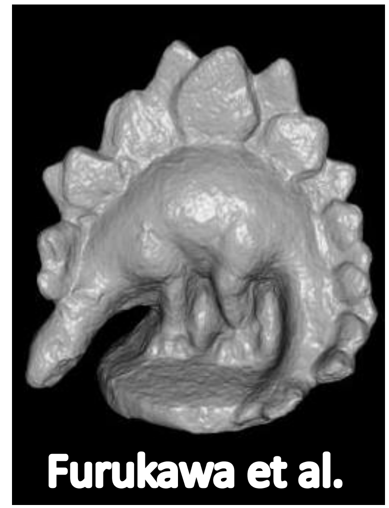
Competitive Result

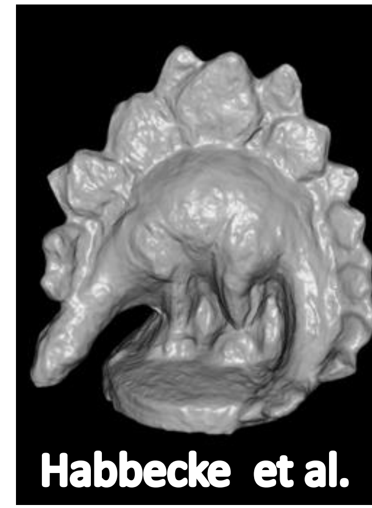
Multi-View Stereo | **Evaluation** • Datasets • Submit • Code

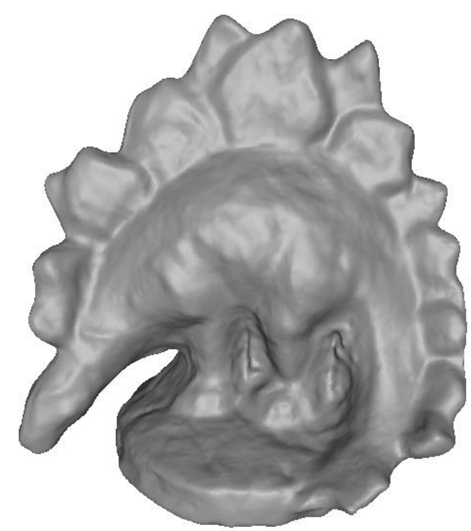
Acc. Threshold: 90% — Comp. Threshold: 1.25 mm — Data in new window — Open Data Window — Data: View 1 and Ground Truth — Image Size Small

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



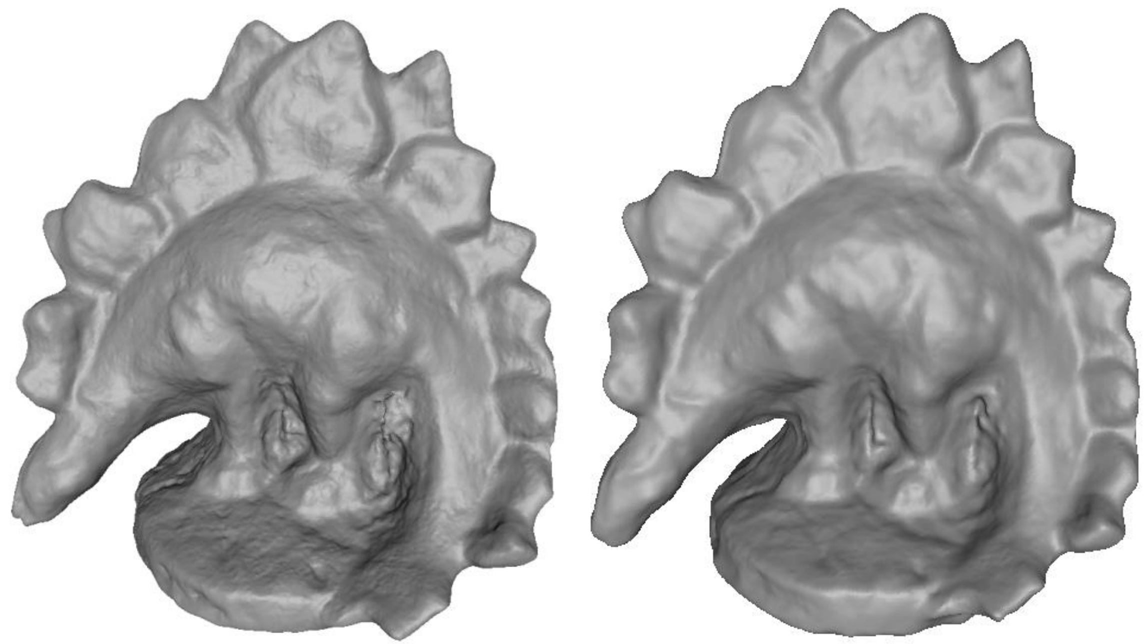




Ours

Competitive Result (Cont.)

Habbecke et al.

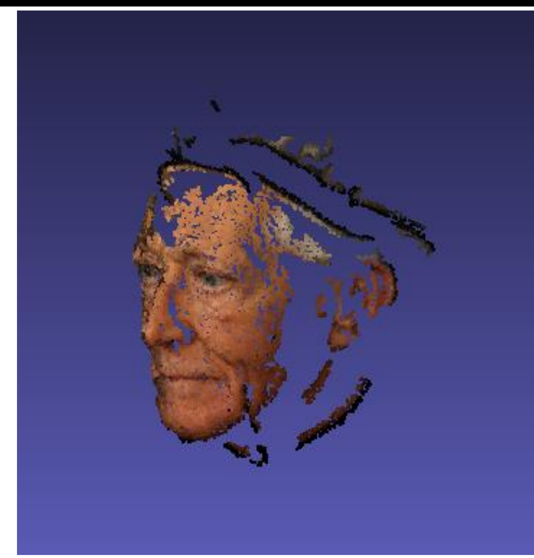


Ours

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

Competitive Result (Cont.)

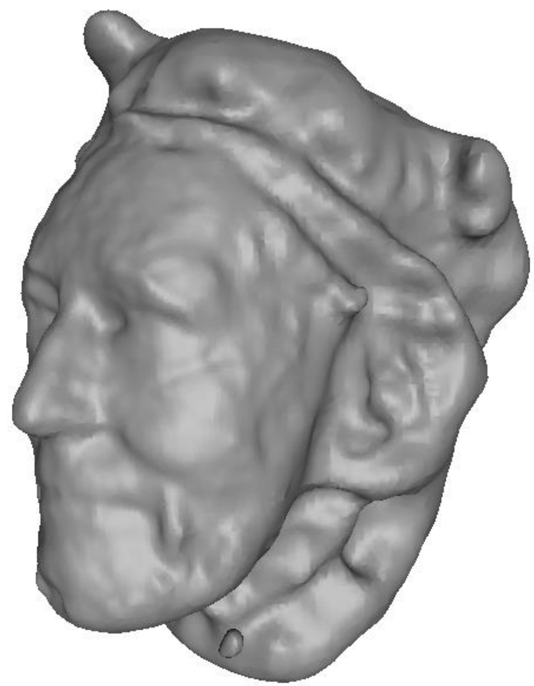
- Completeness

Ours	Furukawa et al.
	

NCTU PAIS Lab

Competitive Result (Cont.)

Ours



Furukawa et al.

Most Related Research

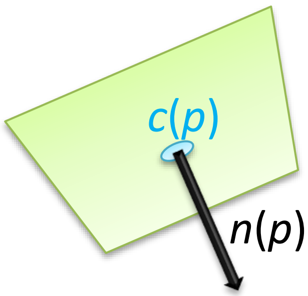
- Patch based MVS & Region growing algorithm
  - Furukawa et al.
    - [8] Accurate, Dense, and Robust Multi-view Stereopsis. PAMI, 2009.
    - [10] Manhattan-world Stereo. CVPR, 2009.
  - Habbecke et al.
    - [6] Iterative Multi-View Plane Fitting. VMV, 2006.
    - [7] A surface-growing approach to multi-view stereo reconstruction. CVPR, 2007.

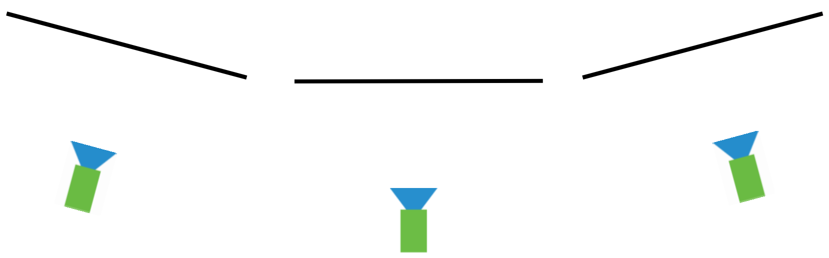
Furukawa et al. (Patch-Based MVS)

Definition of Patch:

- Given initial estimates of
  - Position $c(p)$
  - Normal $n(p)$
  - Visible images $V(p)$
- Refine $c(p)$ and $n(p)$

$$\{c(p), n(p)\} = \underset{\{c(p),n(p)\}}{\arg\max}\, N(p)$$





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

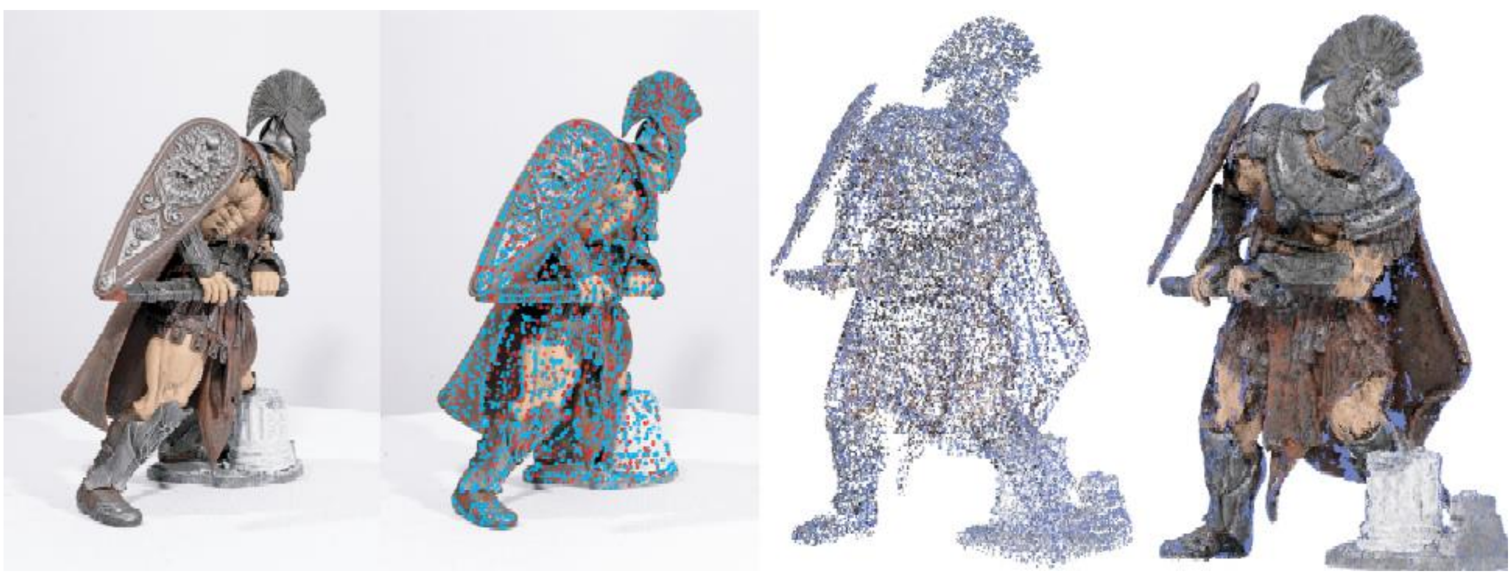
Furukawa et al. (PMVS)

- Patch Iterative Optimization by Conjugate Gradient
  - Photo-consistency, $N(p)$, must be reasonably high after optimization
- Patch Verification process
  - Keep only high photo-consistency images in $V(p)$
  - Accept if $|V(p)| \geq 3$

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

Furukawa et al. (PMVS)

1. Feature detection- Harris and DoG
2. Initial feature matching
3. Iterations of patch expansion and filtering

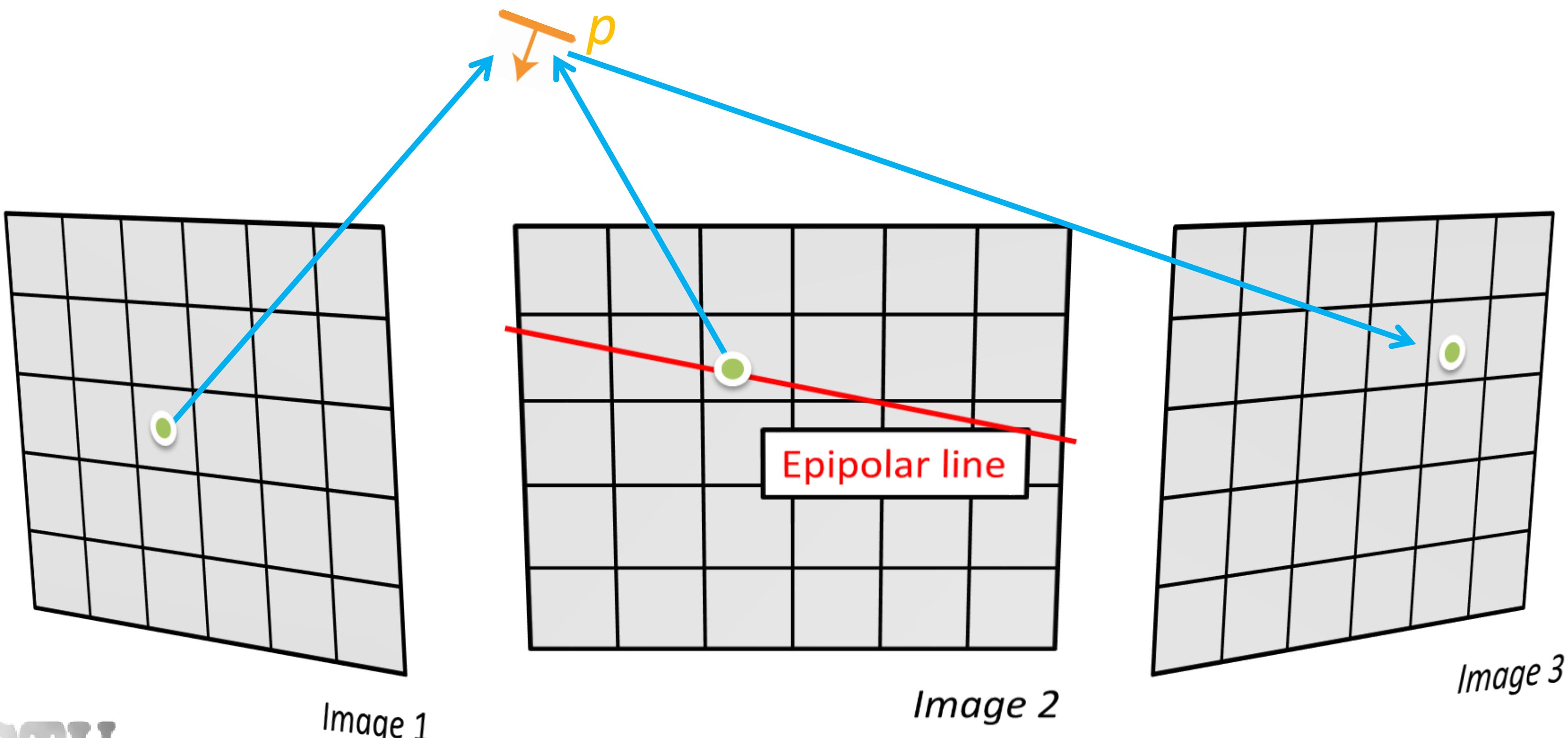


Input image Step 1 Step 2 Step 3

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

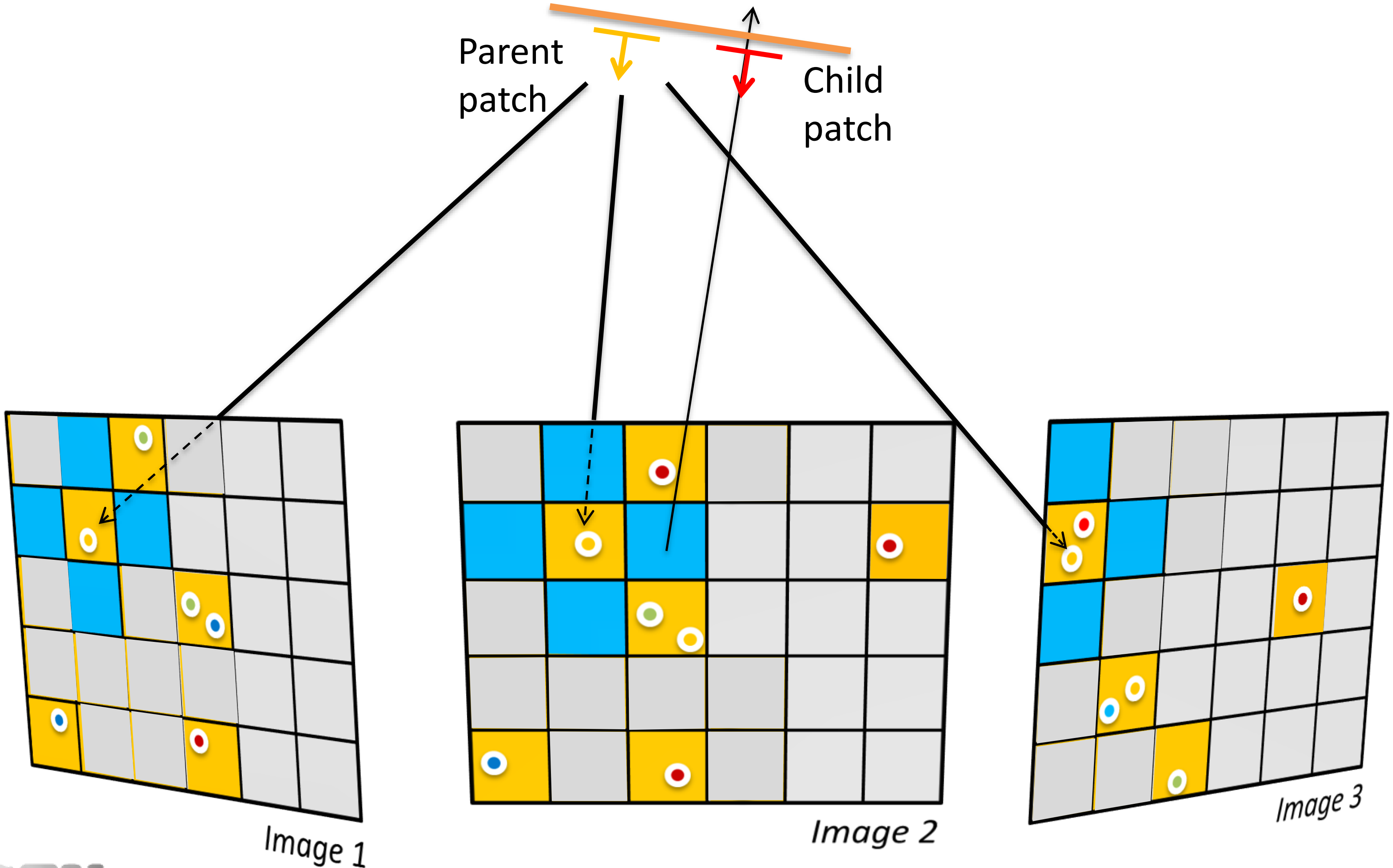
2. Initial feature matching

c(*p*): triangulation
n(*p*): parallel to *Image1*
V(*p*): {*Image1*, *Image2*, *Image3*}



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

3. Iterations of patch expansion and filtering in parent-child fashion



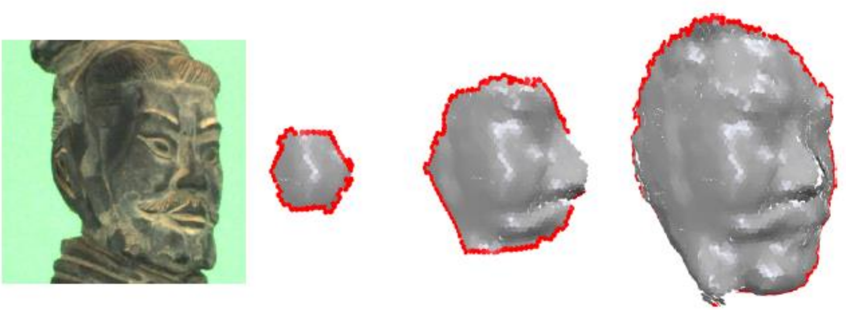
Carlos Hernández, George Vogiatzis, Yasutaka Furukawa.
3d shape reconstruction from photographs: a Multi-View Stereo approach. CVPR2010.

Furukawa et al. (PMVS)

- Surface reconstruction from dense patches obtained

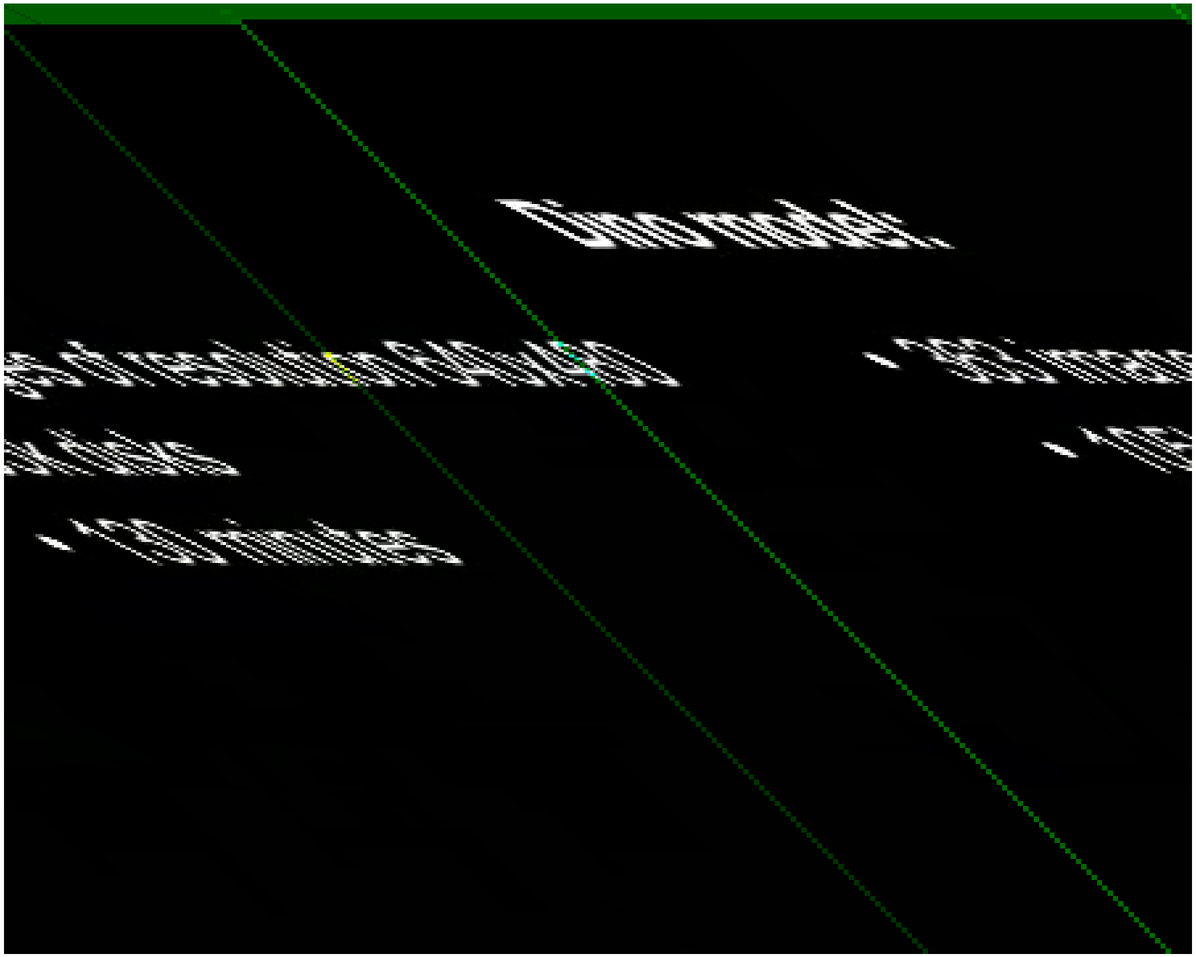
Habbecke et al. (Surface Growing)

- Plane (disk) fitting via photo-consistency measure between visible views under homography
- Disks: active (expandable) and inactive (un-expandable)
- Surface growing initialization with a seed disk
- Surface expansion through active disks
- Visibility test



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

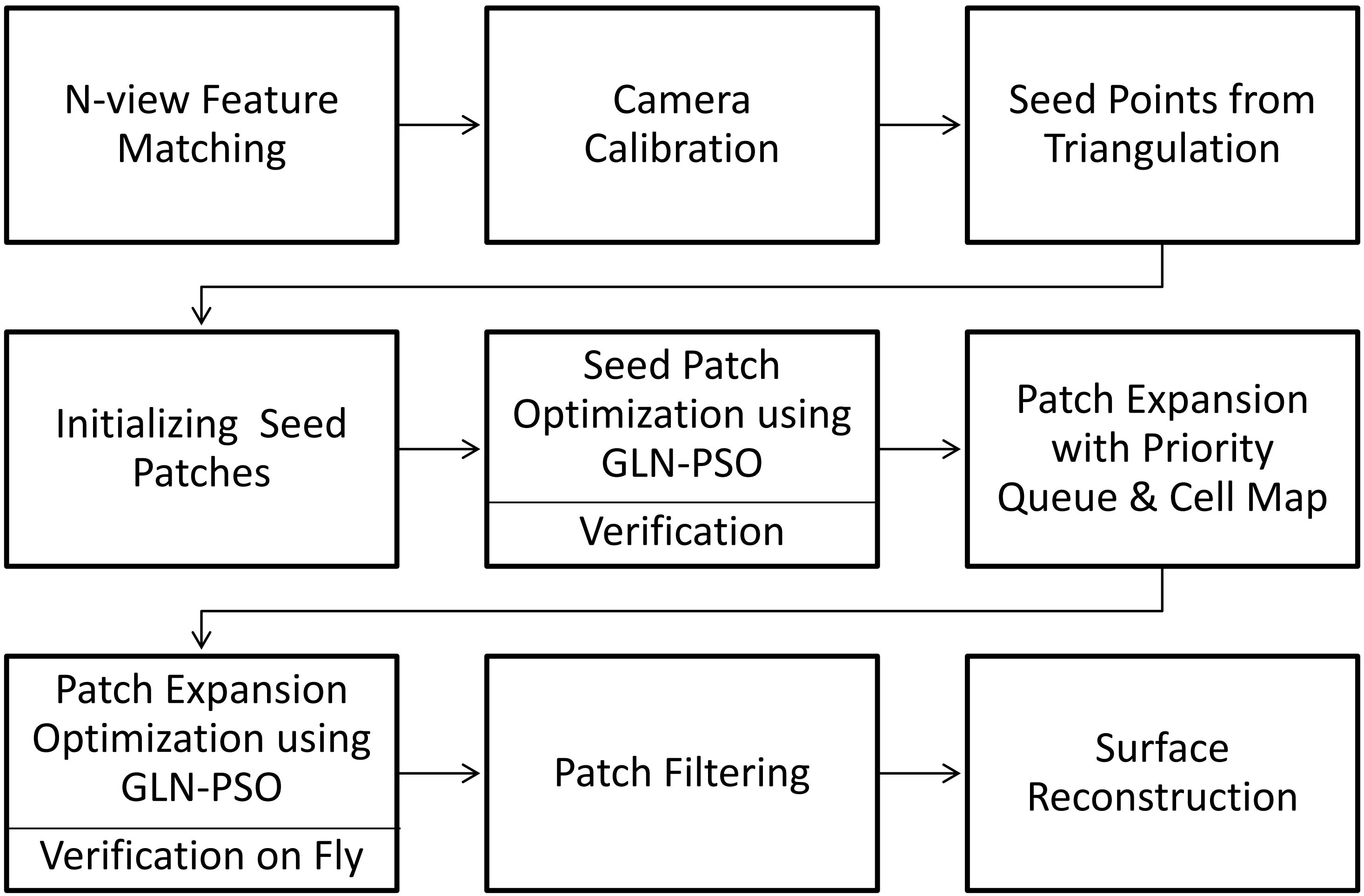
Habbecke et al. (Surface Growing)



http://www.youtube.com/watch?v=f-_d43Bo1R4

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

Proposed system overview

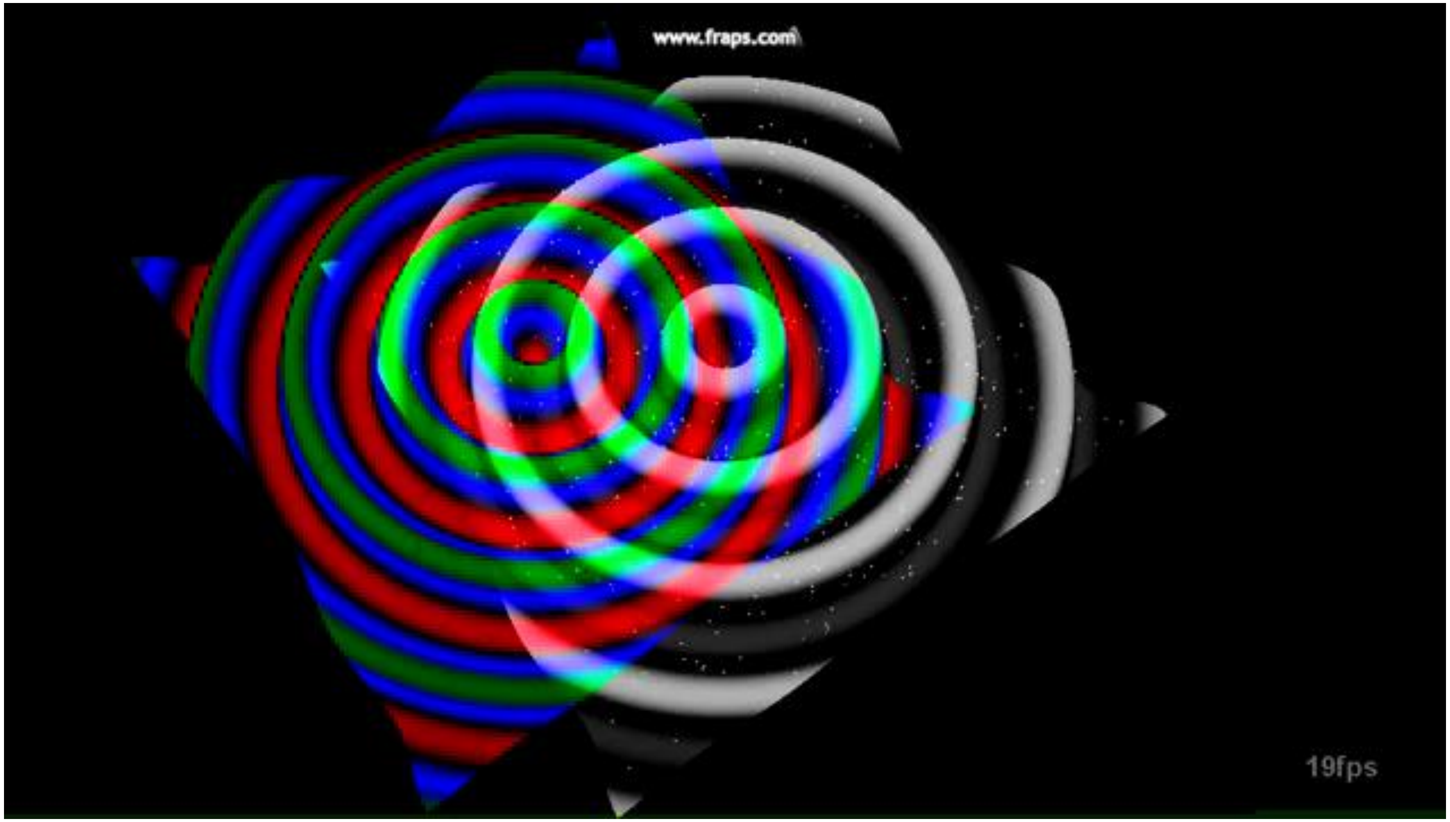


Contributions

- 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

Accurate patches

- GLN-PSO



http://www.youtube.com/watch?v=ML2vuzqw1ok
Particle swarm optimization

Non-uniform view reconstruction

Non-uniform view reconstruction (Cont.)

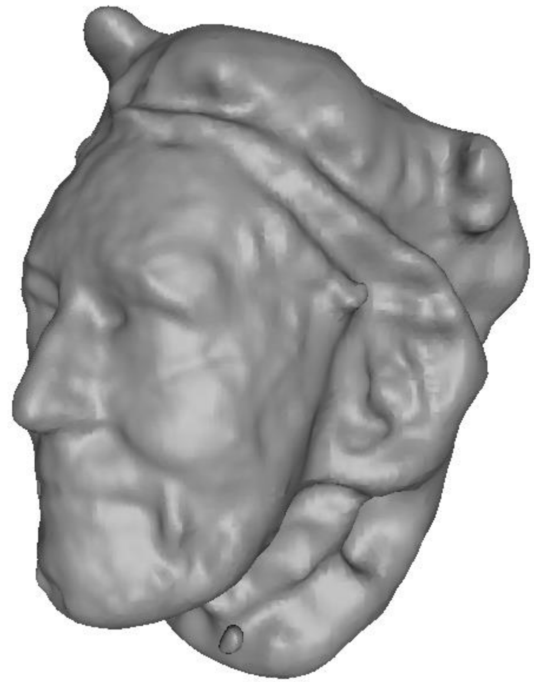
- Completeness

Ours	Furukawa et al.
	

Non-uniform view reconstruction (Cont.)

- Reconstruction result

Ours



Furukawa et al.

Level of detail

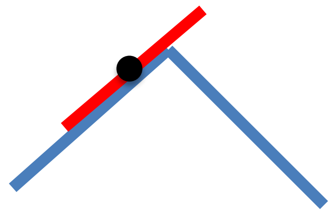
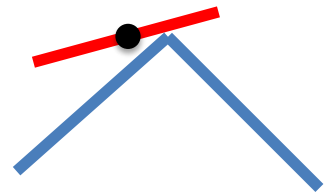
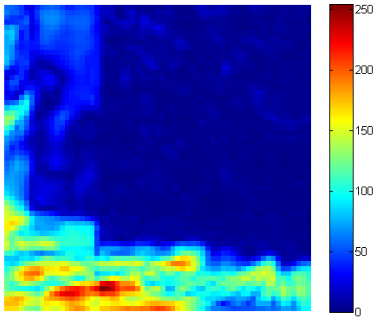
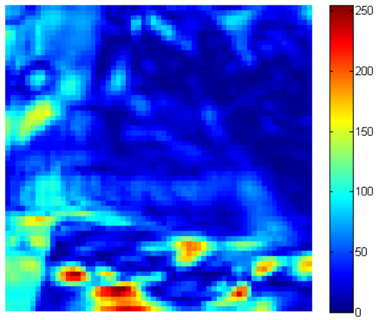
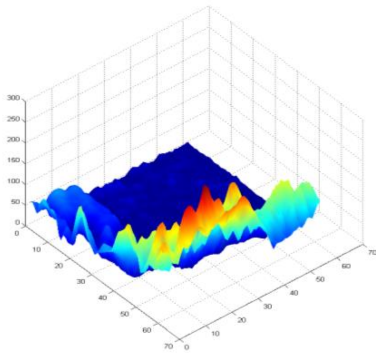
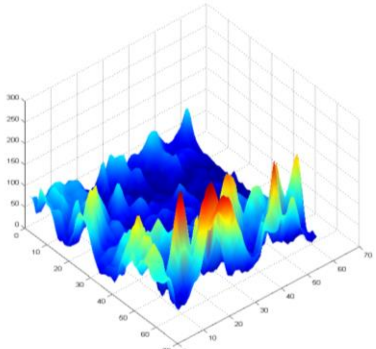
- Textureless model

Scale ratio $\varepsilon = 0.8$

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

NCTU 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		
		

Expansion strategy with priority

- Demo

2. PATCH OPTIMIZATION

N-View feature matching

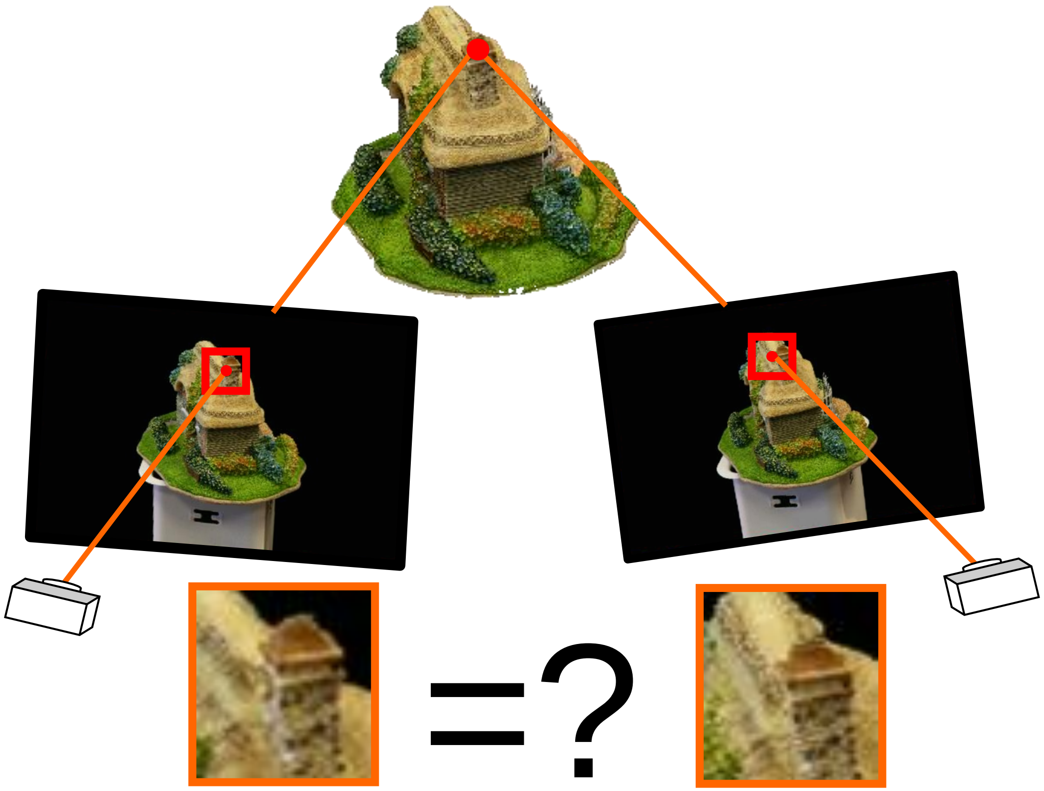
- Ours seed point matching
  - Feature **detector** and **descriptor**
  - Descriptor matching with **Epipolar geometric** constrain
  - With **less erroneous matching** correspondence
  - detectors and descriptors
    - SIFT: weakness of perspective
    - ZM + MSER: strong perspective correction
- Furukawa et al. seed point matching
  - Only **detector** (Harris & DoG) with **Epipolar geometric** constrain
  - **Large number of erroneous** patches
  - Filter out erroneous patch after expansion
  - Iteratively expansion and filtering

Camera calibration
Structure from motion

- Camera calibration with pattern
  - Controllable scene
  - Usually indoor
  - Accurate even with few views

- Structure from motion without pattern
  - Large scene
  - Usually outdoor
  - Robust with large number of views

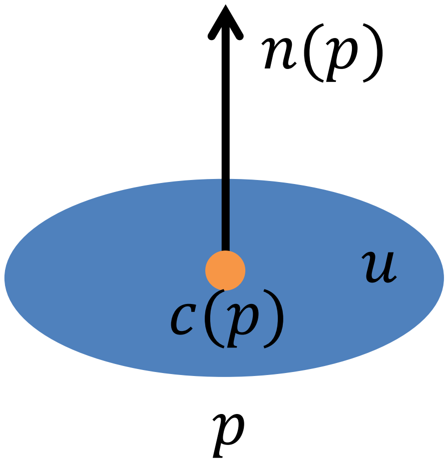
Seed point from feature matching



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

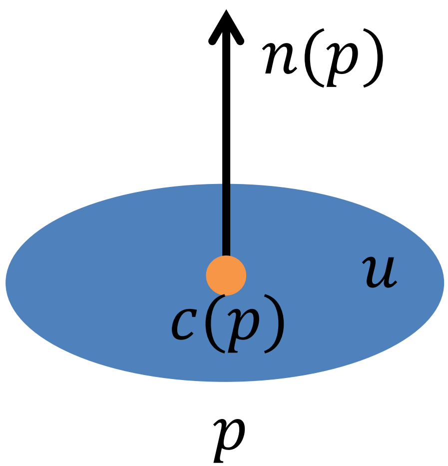
Patch Definition

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



Patch Center

- Patch center $c(p)$
  - seed point from triangulation



Patch Visible Cameras

- Patch visible cameras $V(p)$
  - From feature matching

Patch Visible Cameras (Cont.)

- Invisible camera filtering
  - Expand visible camera included invisible cameras.
  - Filtering after optimization.
  - Re-optimization when $V(p)$ is changed.
- Reference correlation camera $V_\gamma(p)$

$$V_\gamma(p) = arg \max_{i \in V(p)} \sum_{j \in V(p), i \neq j} t_i \cdot t_j$$

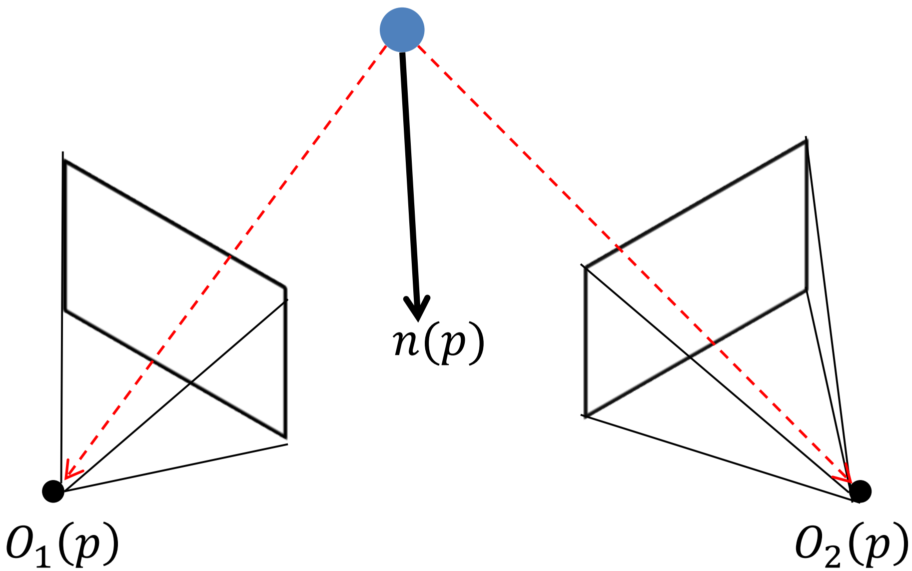
- Filter out invisible camera with correlation threshold α

$$V(p) \leftarrow \{i \in V(p), t_i \cdot t_{V_\gamma(p)} > \alpha\}$$

NCTU PAIS Lab

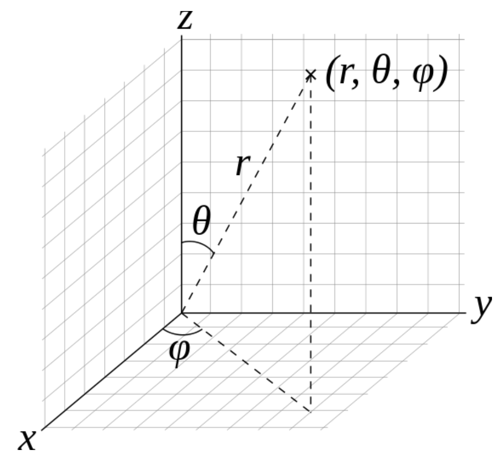
Patch Normal

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



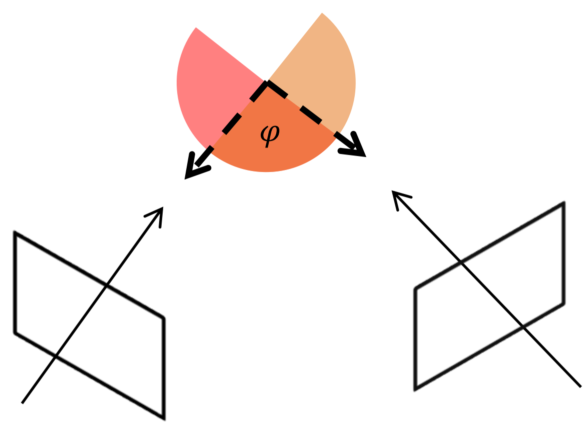
Patch Normal Range

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



Wikipedia: Spherical coordinate system

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



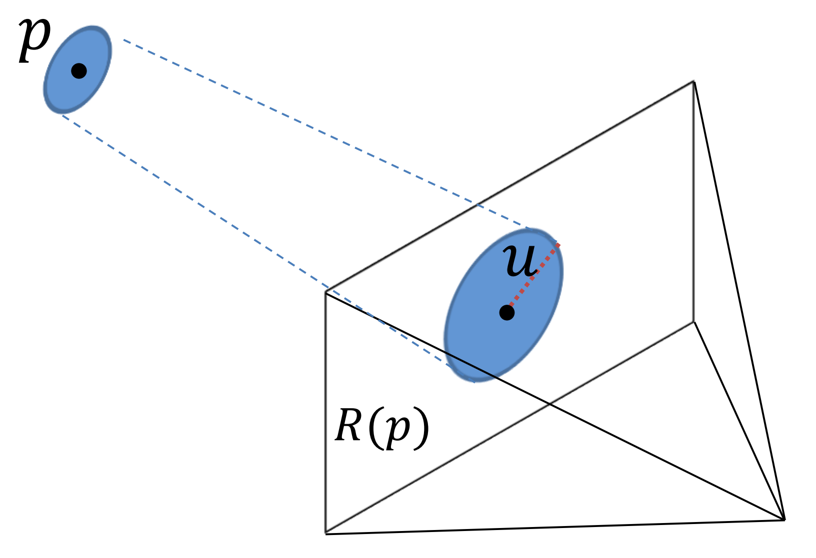
Patch Reference Camera

- Patch reference camera $R(p)$
  - $n(p)$
  - Most perpendicular view in $V(p)$

- Update $R(p)$ after patch optimization
  - Update by optimized $n(p)$
  - Re-optimization when $R(p)$ is updated.
    - Speedup: only seed patch.

Patch Extent

- Patch Extent u
  - Projected region on $R(p)$
- Texture sampling
  - Bilinear interpolation



Patch Depth

- Patch Depth $d(p)$
  - Distance from reference camera center to patch center

$$d(p) = \left\| c(p) - O_{R(p)} \right\|$$

- Depth unit ray

$$r(p) = \frac{c(p) - O_{R(p)}}{d(p)}$$

Patch Depth Range

- Patch Depth Range $d_{range}(p)$

$$d_{range}(p) = d(p) \pm \max_{i \in V(p), i \neq R(p)} \beta \left\| P_i\big(c(p) + r(p)\big) - P_i\big(c(p)\big) \right\|^{-1}$$

  - Nature range with other visible camera view

$$i \in V(p), i \neq R(p)$$

  - Transform pixel to world unit

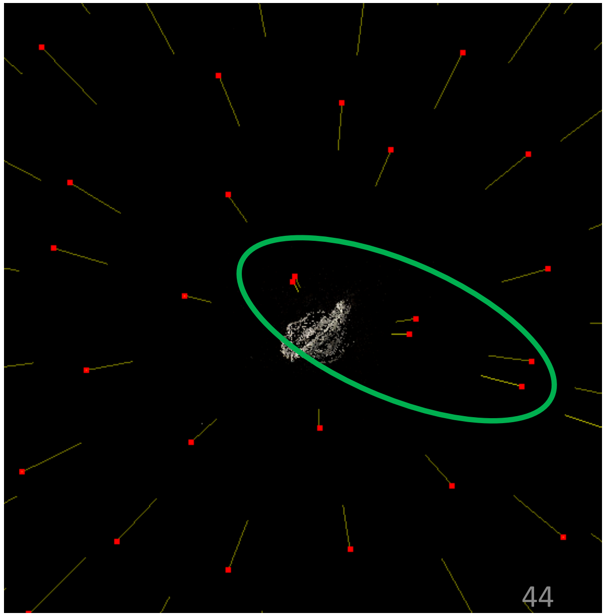
$$\left\| P_i\big(c(p) + r(p)\big) - P_i\big(c(p)\big) \right\|^{-1}$$

  - Displacement scalar β (pixel)
    - Related with cells size

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

NCTU PAIS Lab

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

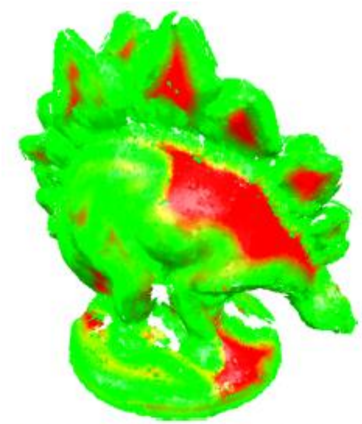


Figure 2. Variance-driven determination of disk size. Red colored parts of the surface are approximated by large disks due to little image texture, while green color depicts small disks.

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

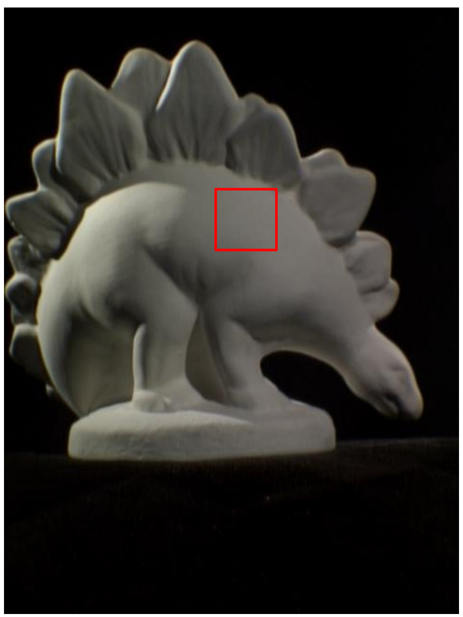
Level of Detail (cont.)

- 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

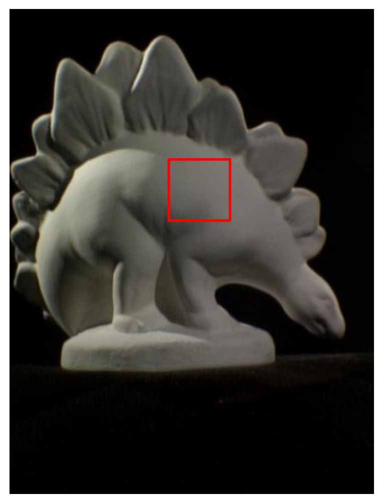
Level of Detail (cont.)

- Level of Detail $l(p)$

Scale ratio $\varepsilon = 0.8$



$l(p) = 0$



$l(p) = 1$

$l(p) = 2$

Level of Detail (cont.)

- Choose minimum $l(p)$ to satisfy the intensity variation threshold v in extent region

$$l(p) = arg\,\min_{l} \left(|\Omega_l(p)|^{-1} \sum_{i \epsilon \Omega_l(p)} \left(I_{R(p)}(i) - |\Omega_l(p)|^{-1} \sum_{i \epsilon \Omega_l(p)} I_{R(p)}(i) \right)^2 \right) > v$$

- Scale ratio strategy ε
  - Large with more pyramid layers (more memory usage), but with critical enough texture in extent.
  - Small with less pyramid layers, but may cause including erroneous region in extent.

Patch optimization

- Optimization (DOF 3)
  - Depth $d(p)$
    - 1D distance to reference camera center
    - Optimized patch center

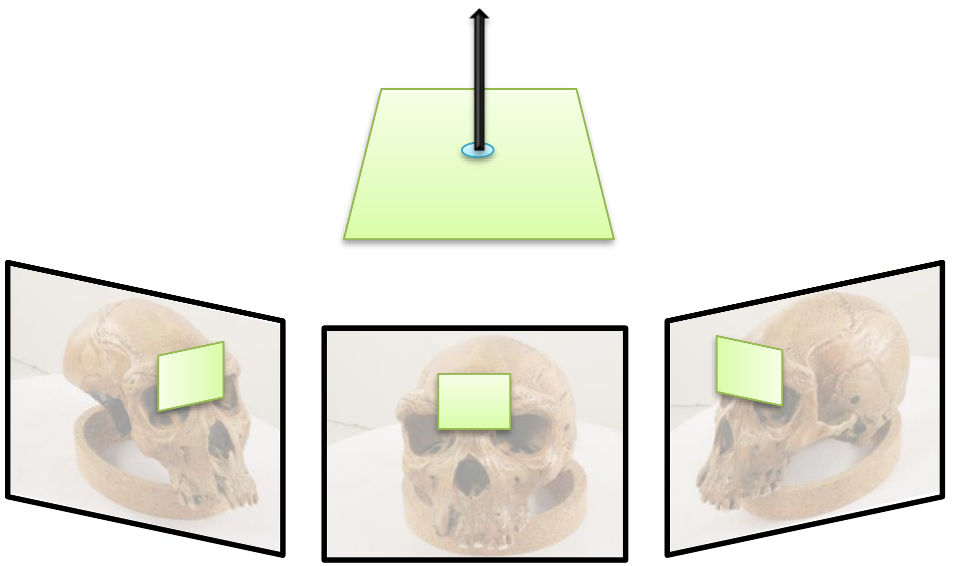
$$c(p) = d(p) * r(p) + O_{R(p)}$$

  - 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

Patch optimization (cont.)

- Patch Extent Homography with level of detail
    1. Homography from $I_{R(p)}$ to patch plane.

$$H_1 = (D(p)L(p)M_{R(p)} - L(p)m_{R(p)}n(p)^T)^{-1}$$

    2. Homography from patch plane to $I_{V(p)}$

$$H_2 = (D(p)L(p)M_i - L(p)m_i n(p)^T), i \in V(p)$$

Patch optimization (cont.)

- Homography

$$\boldsymbol{H_i(p) = H_2 H_1, i \in V(p)\ , where}$$

$$P_i = K_i[R_i\ T_i] = [M_i\ m_i]$$

$$D(p) = \frac{-n(p)^T c(p)}{\|n(p)\|}$$

$$L(p) = \begin{bmatrix} \varepsilon^{-l(p)} & 0 & 0 \\ 0 & \varepsilon^{-l(p)} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

NCTU PAIS Lab

Adaptive Fitness Weighting

- Cross planar patch
  - Distance weighting

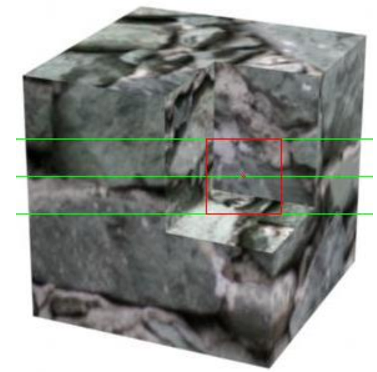
$$g(x,y)=\frac{e^{-\left(\frac{(x-x_0)^2}{2\sigma_g}+\frac{(y-y_0)^2}{2\sigma_g}\right)}}{\sum_{u,v\in p} e^{-\left(\frac{(u-x_0)^2}{2\sigma_g}+\frac{(v-y_0)^2}{2\sigma_g}\right)}}$$

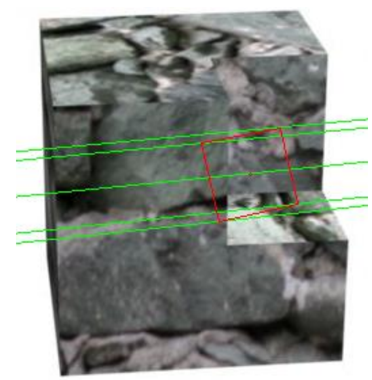
  - Difference weighting

$$h(x,y)=e^{-\frac{d(x,y)^2}{2\sigma_h}}$$

- Textureless
  - Normalized Edge gradient magnitude

$$G_i=\left\|I_i*G_x+I_i*G_y\right\|*\left\|I_i*G_x+I_i*G_y\right\|_{max}^{-1}$$
$$k(x,y)=G_i(x_I,y_I)$$





Adaptive Fitness Weighting (cont.)

- Adaptive weighted Fitness
  - Minimize sum of weighted absolute mean difference $\Delta T(p)$

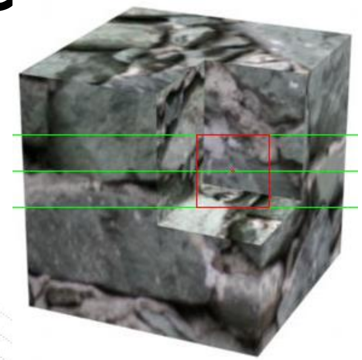
$$\left(n(p), d(p)\right) = arg \min_{n(p),d(p)} \sum_{i \in V(p)} w\left(\Omega_{l(p)}(p)\right) \Delta T(p)$$

$$\bar{T}(p) = |V(p)|^{-1} \sum_{j \in V(p)} I_j(H_j(p)\Omega_{l(p)}(p))$$

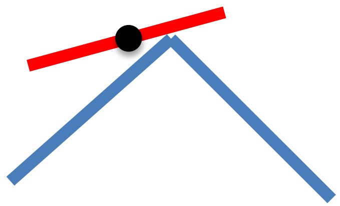
$$\Delta T(p) = |V(p)|^{-1} \sum_{i \in V(p)} \left| I_i(H_i(p)\Omega_{l(p)}(p)) - \bar{T}(p) \right|$$

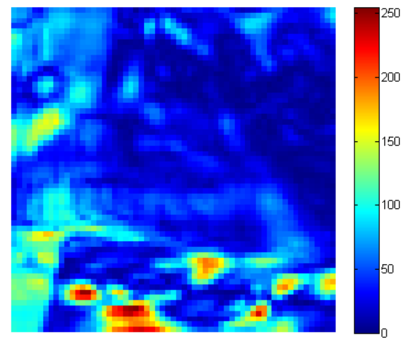
NCTU PAIS Lab

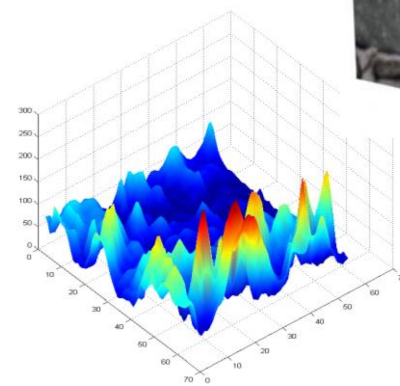
Adaptive Fitness Weighting (cont.)



- Cross planar patch extent
  - Averaging error

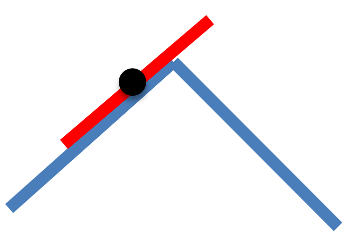


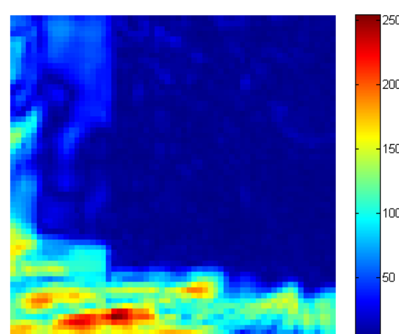


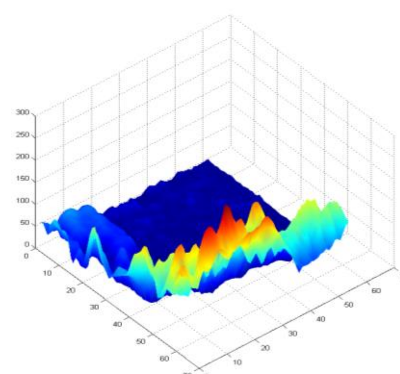


- After adapted fitness weighting

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







Patch Normalized Correlation

- Evaluation of patch quality
  - Used in filtering and priority

$$\gamma(p) = \frac{\sum_{i,j \in V(p), i \neq j} t_i \cdot t_j}{|V(p)|^2 - |V(p)|}$$

NCTU PAIS Lab

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

Introduction to PSO

- PSO (Particle swarm Optimization)
  - Flexible initialization
    - Give initial range instead of initial point
  - Parallel particles (fitness function).
  - Quality controllable result
    - Control by particle and iteration number
    - Control by convergence threshold

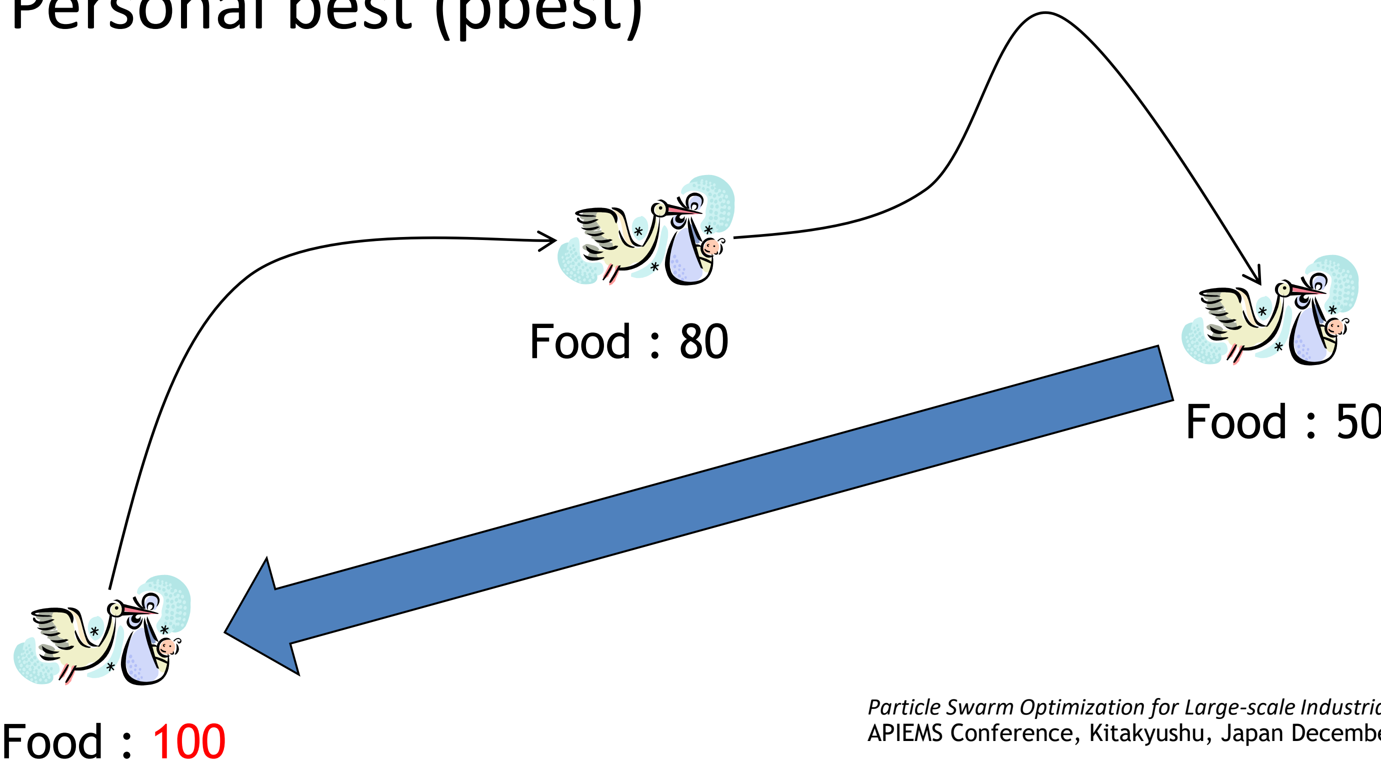
Introduction to PSO (cont.)

- PSO Algorithm

1. Initial particles (initial parameter & velocity).
2. Set particle pbest from initial parameter.
3. Loop until convergence criteria
   1. Update particle velocity.
   2. Move particles.
   3. Update pbest and gbest.
4. Return gbest.

Particle swarm Optimization

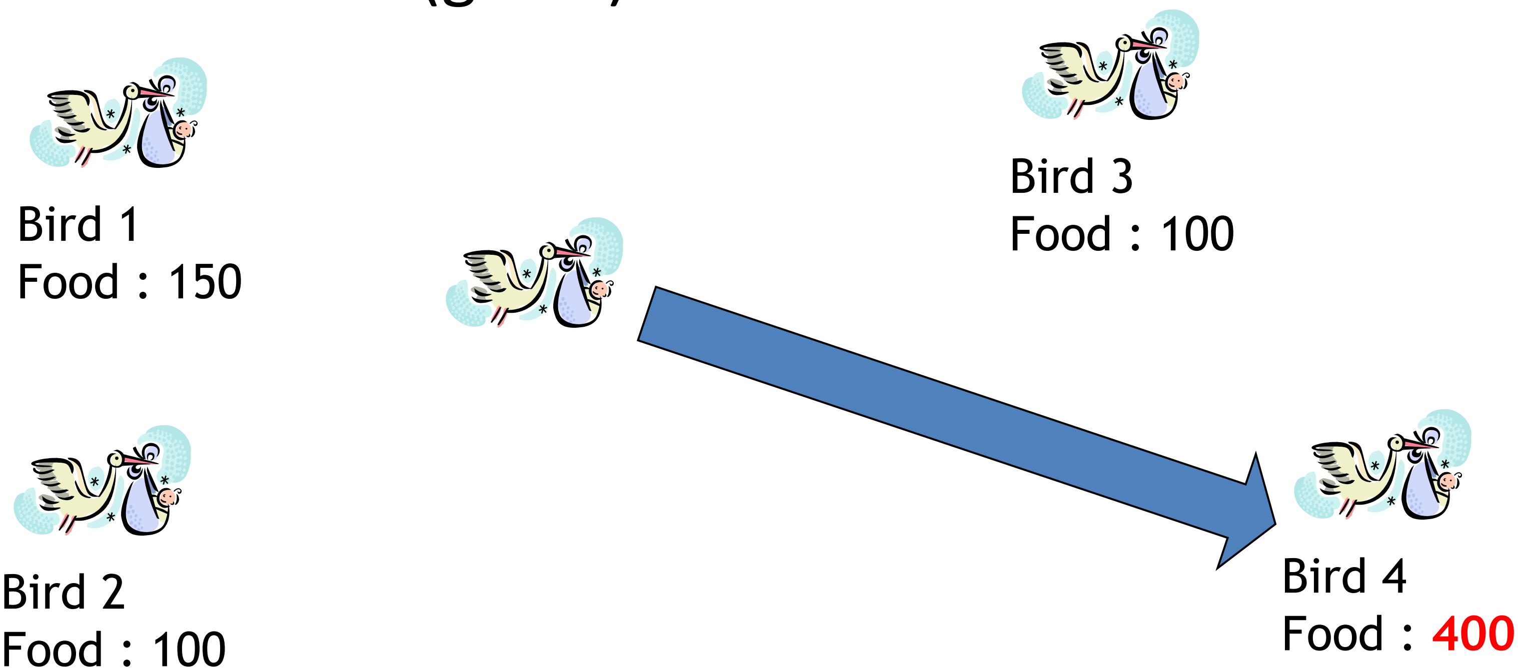
- Particle swarm Optimization
  - Personal best (pbest)



Particle Swarm Optimization for Large-scale Industrial Applications.
APIEMS Conference, Kitakyushu, Japan December, 14-16, 2009

Particle swarm Optimization (cont.)

- Particle swarm Optimization
  - Global best (gbest)



Particle Swarm Optimization for Large-scale Industrial Applications.
APIEMS Conference, Kitakyushu, Japan December, 14-16, 2009

5.1.3 Particle swarm Optimization (cont.)

- Particle swarm Optimization
  - Moving Velocity

$$v_{id} = wv_{id} + c_p u(p_{id} - x_{id}) + c_g u(g_{id} - x_{id})$$

Velocity = Inertia Velocity + Cognitive Learning + Social Learning

5.2 Introduction to GLN-PSO

- GLN-PSO (**G**lobal-**L**ocal-**N**ear Neighbor **P**article **swarm O**ptimization)
  - An advanced version of standard PSO
    - Local best (K-nearest neighbor pbest)
    - Near Neighbor best (pbest of each dimension)
  - Improve local search
    - Good for patch expansion. Slightly adjust from parent patch.

GLN-PSO

- GLN-PSO
  - More social learning.
    - **Local Best (lbest)**
    - **Near Neighbor Best (nbest)**

$$v_{id} = wv_{id} + c_p u(p_{id} - x_{id}) + c_g u(g_{id} - x_{id}) + \boldsymbol{c_l u(l_{id} - x_{id})} + \boldsymbol{c_n u(n_{id} - x_{id})}$$

GLN-PSO

- Local Best (lbest)
  - Local search.
  - The best K-nearest neighbor pbest.

- Problem: Sorting distance.
  - Slow for large number of particles/dimensions.
  - **Worth it.**

GLN-PSO

- Near Neighbor Best (nbest)
  - pbest maximize FDR of **each dimension**.
  - Fitness distance ratio (FDR).

$$n_{id} = \arg\max_{p_{jd}} \left\{ FDR = \frac{\xi(X_i) - \xi\left(P_j\right)}{\left|x_{id} - p_{jd}\right|} \, which \; i \neq j \right\}$$

NCTU PAIS Lab

GLN-PSO

- Convergence
  - Dispersion index

$$\bar{\delta} = \frac{\sum_{i=1}^{I} \sum_{d=1}^{D} \left| x_{id} - p_{gd} \right|}{I \cdot D}$$

  - Velocity index

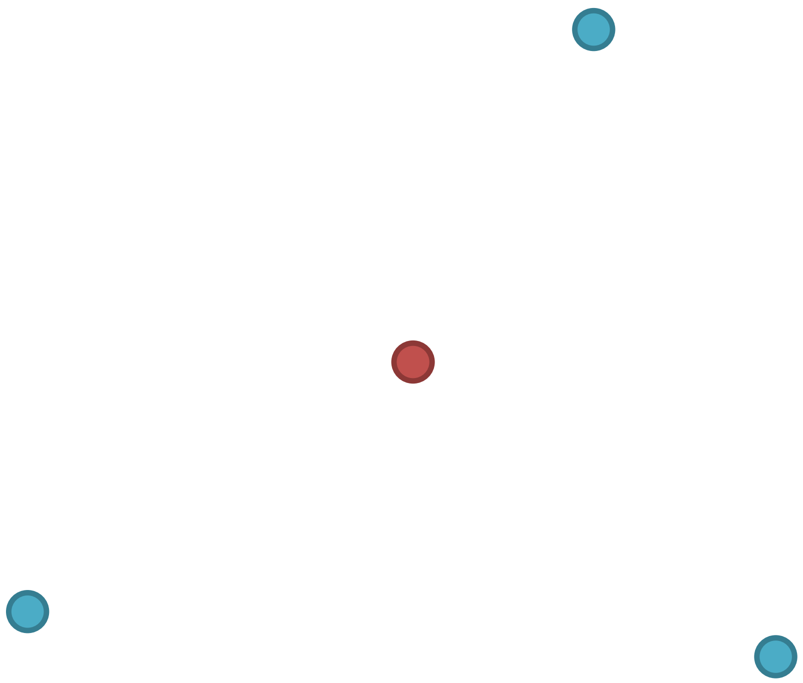
$$\bar{v} = \frac{\sum_{i=1}^{I} \sum_{d=1}^{D} \left| v_{id} \right|}{I \cdot D}$$

GLN-PSO

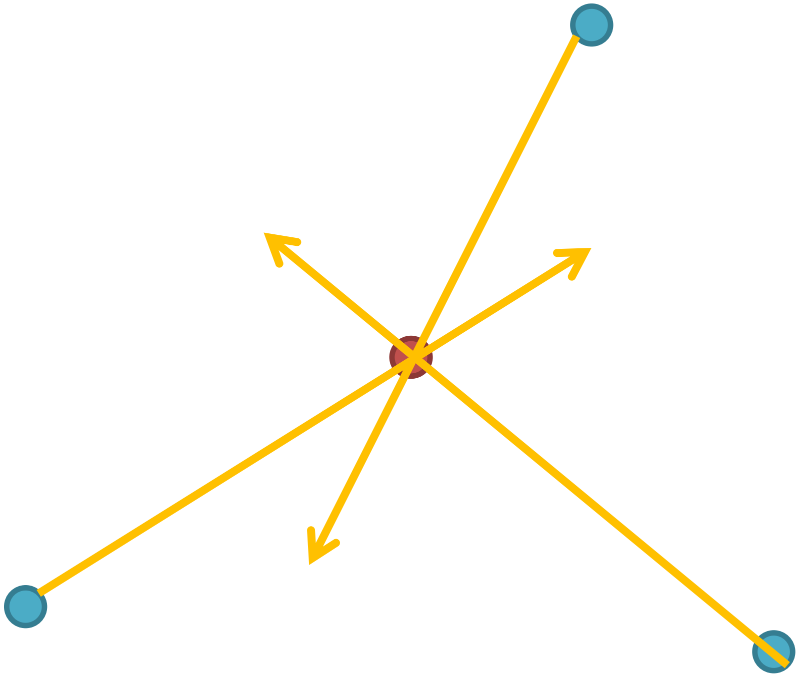
- Acceleration Constant Strategy
  - Convergence faster
  - $w < 1$, decreased every iteration
    - $w \geq 1$ will *not* convergence until the end of iteration.
  - $c_g > 1$, overshooting but looks well ($c_g = 1.2\sim1.5$)

$$v_{id} = wv_{id} + c_p u(p_{id} - x_{id}) + c_g u(g_{id} - x_{id}) + c_l u(l_{id} - x_{id}) + c_n u(n_{id} - x_{id})$$

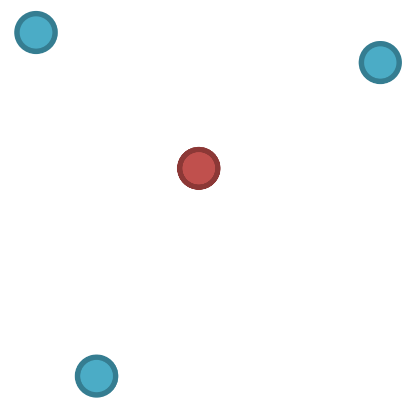
PSO Example



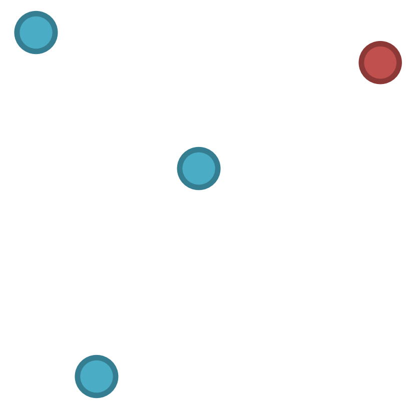
PSO Example (cont.)



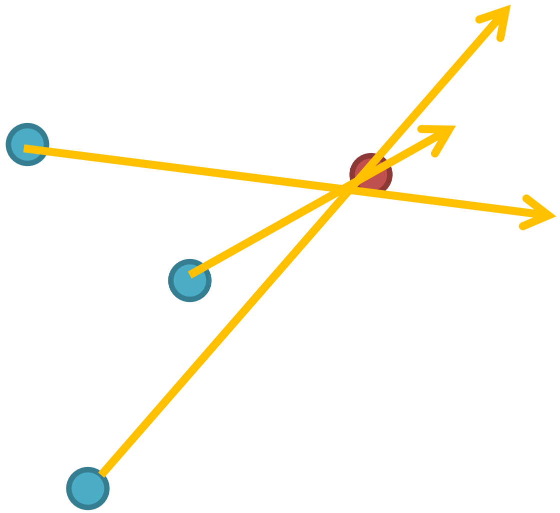
PSO Example (cont.)



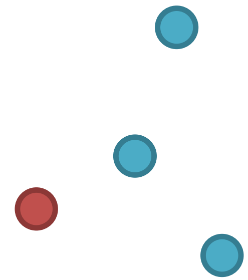
PSO Example (cont.)



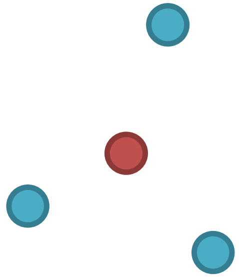
PSO Example (cont.)



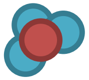
PSO Example (cont.)



PSO Example (cont.)



PSO Example (cont.)

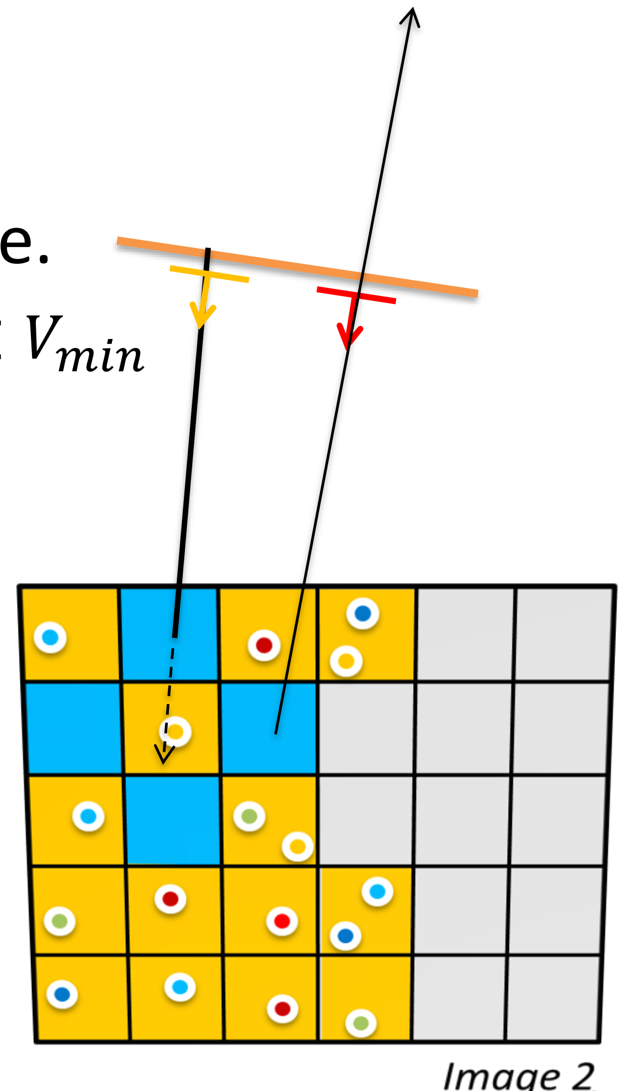


convergence

4. PATCH EXPANSION

Expansion Patch

- Expansion patch
  - $V(p)$ is estimated by view angle.
    - From parent patch when $V(p) < V_{min}$
    - Invisible camera filtering
  - $n(p)$ from parent patch
  - $c(p)$ extent from parent patch



Patch Priority

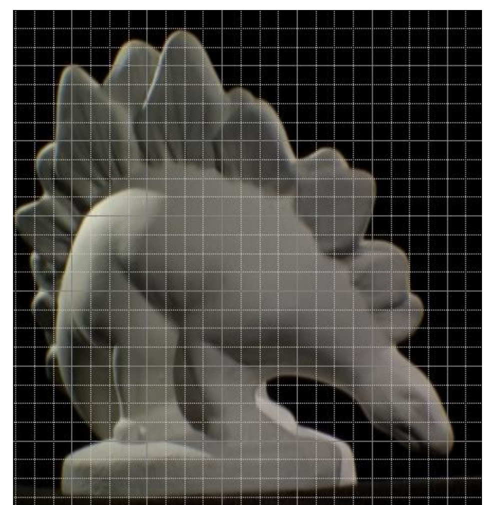
- Patch Priority $q(p)$

$$q(p) = (l(p) + 1)\xi(p)e^{-\gamma(p) - \frac{|V(p)|}{|V|}}$$

  - Adaptive weighted fitness $\xi(p)$
  - Normalized correlation $\gamma(p)$
  - Number of visible camera $|V(p)|$
  - Level of detail $l(p)$

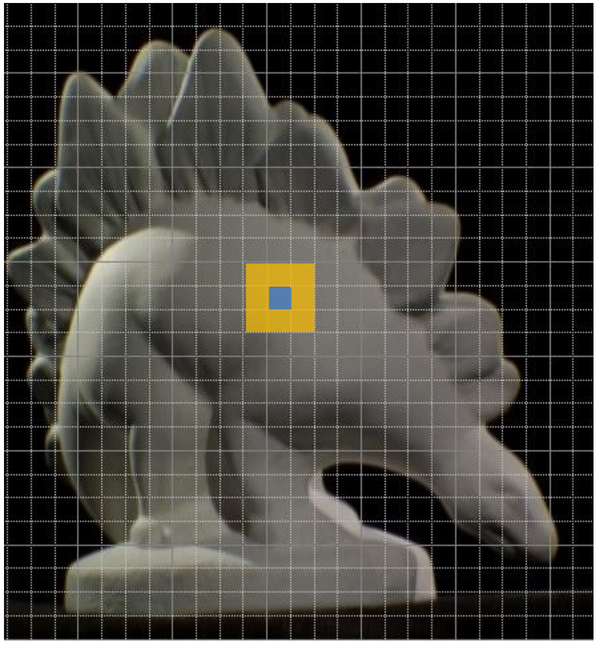
Image Cell Map

- Completeness evaluation
  - $\tau \times \tau$ pixel cells
  - Small with more surface detail and reconstruction time
  - Large with rough surface detail and faster reconstruction time



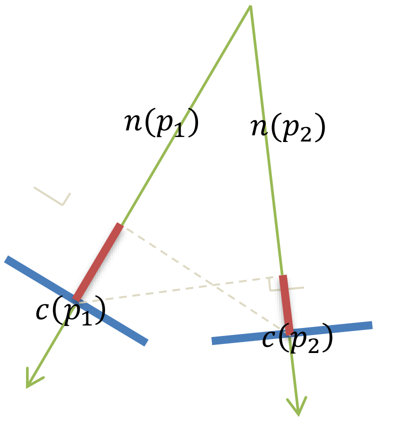
Expansion of Neighboring Cells

- Project p to all visible cell maps $Q_{V(p)}$
- Expand one-ring neighbor cells $N_i(p), i \epsilon V(p)$
  - Speedup: only expand $N_{R(p)}(p)$



NCTU PAIS Lab

Expansion of Neighboring Cells (cont.)



- Skip neighbor cells
  - Already has a neighbor patch

$$\mathrm{n}(p_1, p_2) = \frac{|(c(p_1) - c(p_2)) \cdot n(p_1)| + |(c(p_1) - c(p_2)) \cdot n(p_2)|}{2} < \rho$$

  - Already has a high confidence patch: $\gamma(p) > \alpha$
  - Full cells $|N_i(p)| \geq C_{max}$

NCTU PAIS Lab

Expansion Algorithm

- While priority queue is not empty
  1. Pop the top priority patch $\boldsymbol{p}$ form priority queue
  2. Initialize $\boldsymbol{C_i}(\boldsymbol{p}), \boldsymbol{N_i}(\boldsymbol{p}), \boldsymbol{N_i^*}(\boldsymbol{p})\ \boldsymbol{where\ i} \in \boldsymbol{V}(\boldsymbol{p})$
  3. For each $\boldsymbol{p'}$ in $\boldsymbol{N_i^*}(\boldsymbol{p})$
     1. Initialize $\boldsymbol{V}(\boldsymbol{p'}), \boldsymbol{R}(\boldsymbol{p'}), \boldsymbol{n}(\boldsymbol{p'}), \boldsymbol{c}(\boldsymbol{p'}), \boldsymbol{d}(\boldsymbol{p'}), \boldsymbol{r}(\boldsymbol{p'})$
     2. Optimize $\boldsymbol{p'}$
     3. Filter out invisible views $\boldsymbol{V}(\boldsymbol{p'}) \leftarrow \{\boldsymbol{i} \in \boldsymbol{V}(\boldsymbol{p'}), \boldsymbol{t_i} \cdot \boldsymbol{t}_{\boldsymbol{V_\gamma(p)}} > \boldsymbol{\alpha}\}$
     4. If $|\boldsymbol{V}(\boldsymbol{p'})| < \boldsymbol{V_{min}}$, drop $\boldsymbol{p'}$
     5. If all $|\boldsymbol{C_i}(\boldsymbol{p'})| \geq \boldsymbol{C_{max}}, \boldsymbol{i} \in \boldsymbol{V}(\boldsymbol{p'})$, drop $\boldsymbol{p'}$
     6. Compute $\boldsymbol{q}(\boldsymbol{p'})$
     7. Push $\boldsymbol{p'}$ into $\boldsymbol{Q}_{\boldsymbol{V(p')}}$ and priority queue.

NCTU PAIS Lab

5. PATCH FILTERING

Patch Filtering

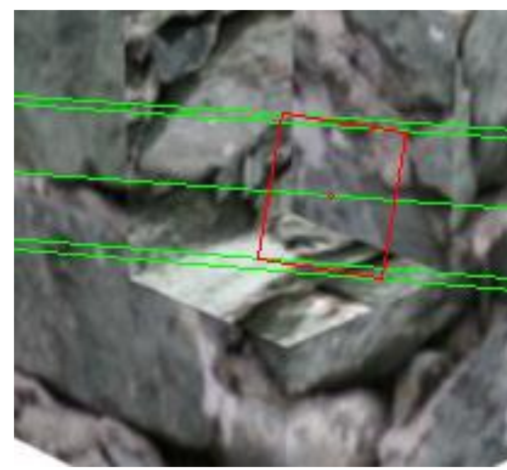
- Patch Verification
- Post-Processing Patch Filtering

Patch Verification

- Patch Verification
  - Visible camera number filtering

$$|V(p)| < V_{min}$$

  - Homography window ratio filtering
  - Silhouette filtering (optional)



Post-Processing Patch Filtering

- Depth Test Filtering
- Patch Correlation Filtering
- Neighboring Cells Filtering

Depth Test Filtering

- Depth Test Filtering
  - Filter out view i when $d(p)$ is not smallest in $C_i(p), i \in V(p)$.
  - Filter out p if $|V(p)| < V_{min}$

Patch Correlation Filtering

- Patch Correlation Filtering
  - Filter out patch with less $|V(p)|$ and $\gamma(p)$ in same cell $C_i(p)$

$$|V(p)|\gamma(p) < \sum_{p' \in C_i(p), i \in V(p)} \gamma(p')$$

NCTU PAIS Lab

Neighboring Cells Filtering

- Neighboring Cells Filtering
  - Consider patches in $N_i(p), i \in V(p)$
  - Neighbor cell includes neighbor patch $\mathrm{n}(p_1, p_2)$
  - Filter out p by neighbor support ratio

$$\frac{\left|\sum_{p_j \in N_i(p)} \mathrm{n}(p, p_j) < \rho\right|}{\sum_{i \in V(p)} |N_i(p)|} < \boldsymbol{\Lambda}$$

NCTU PAIS Lab

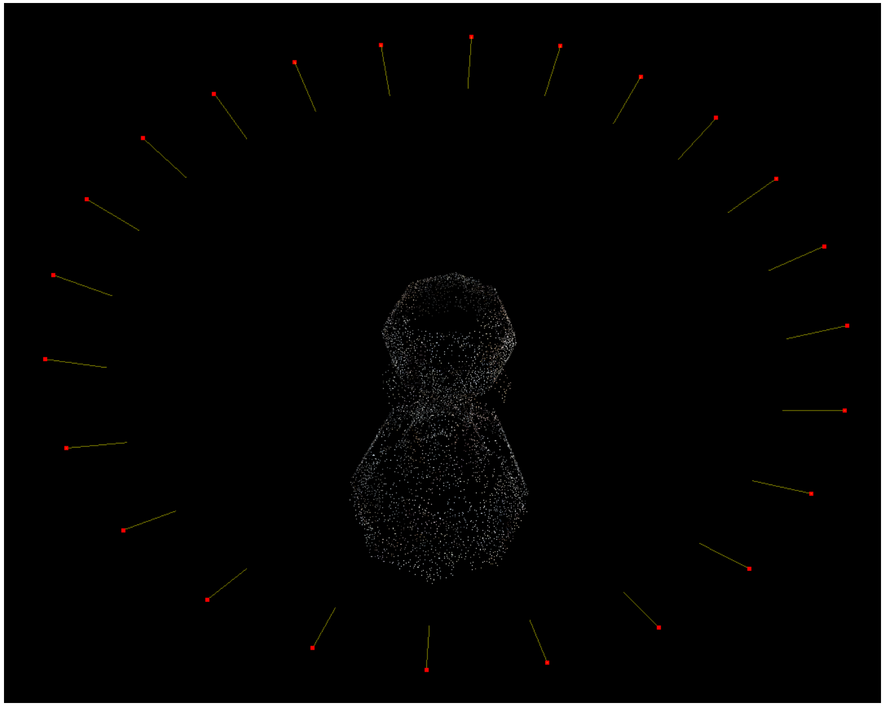
6. EXPERIMENTS

Experiments

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

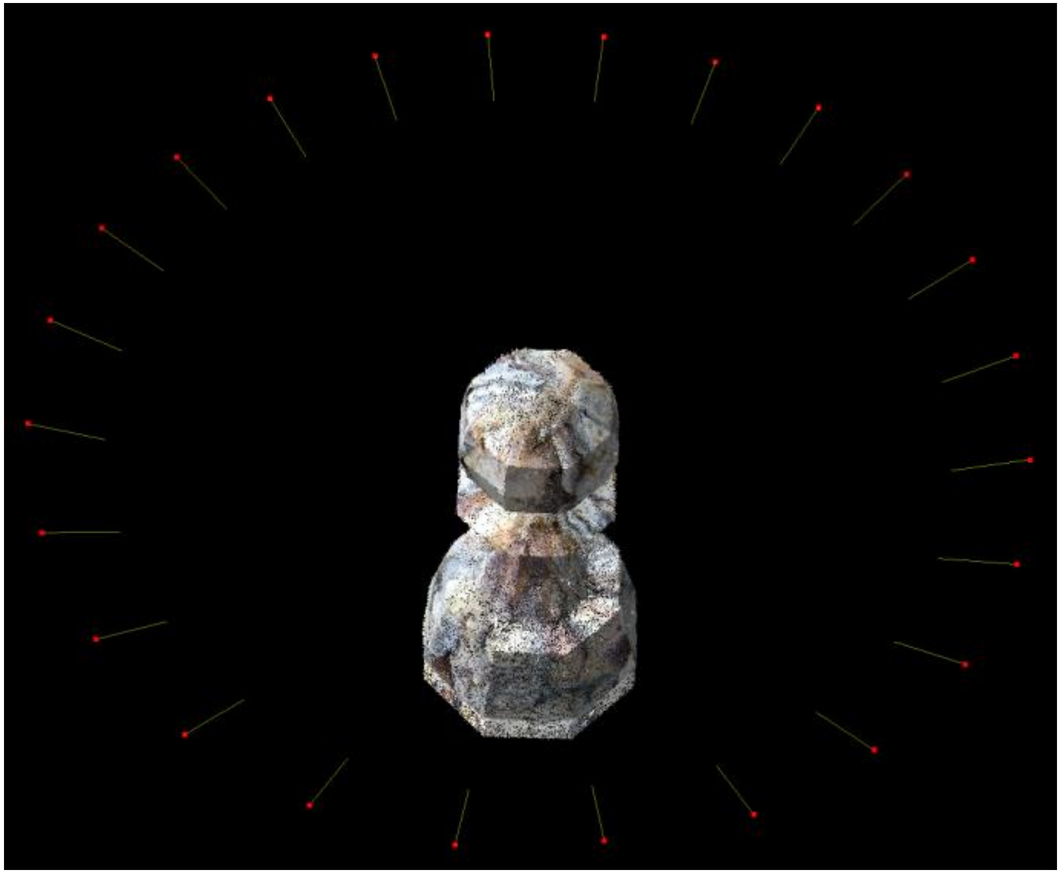
Reconstruction of Synthesis Images

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



Reconstruction of Synthesis Images (Cont.)

- Reconstruction patch model (95068 patches)



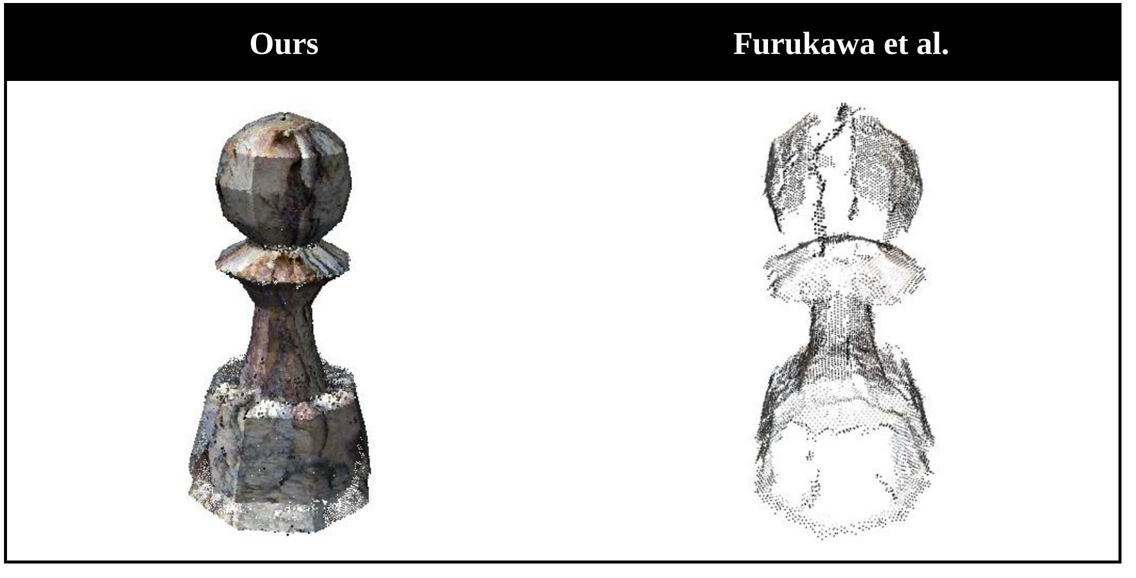
8.1 Reconstruction of Synthesis Images (Cont.)

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

$$\tau = 2$$

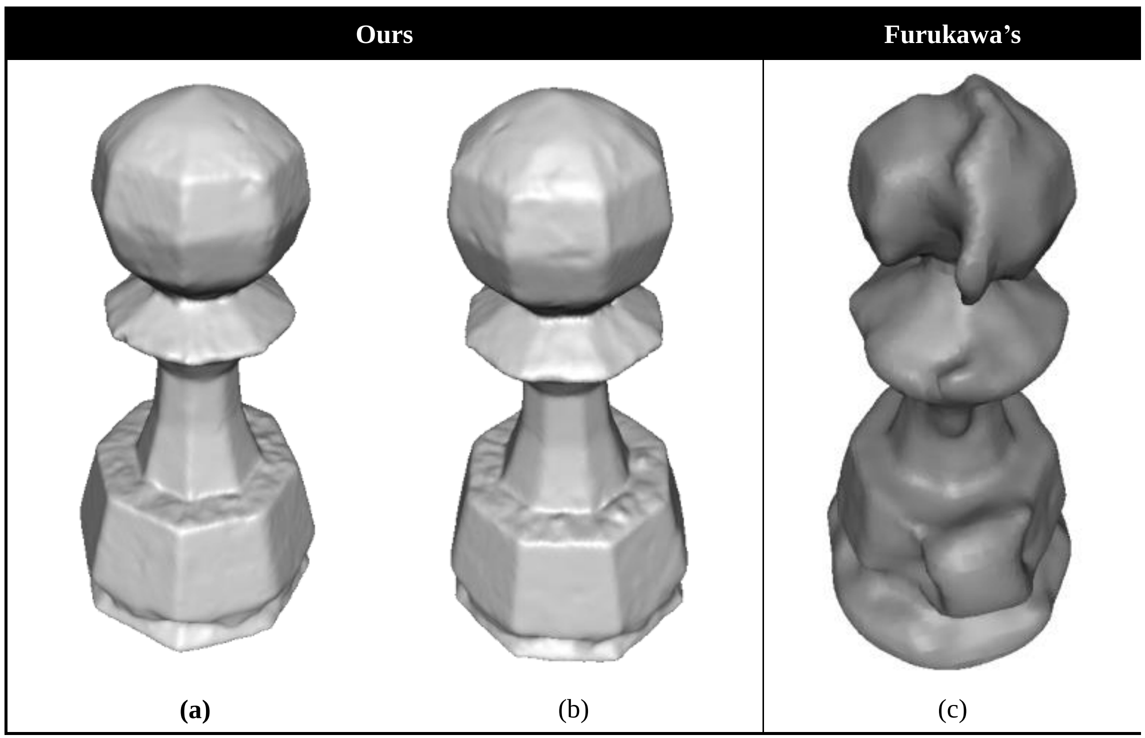
  - Same minimum visible camera number

$$V_{min} = 3$$



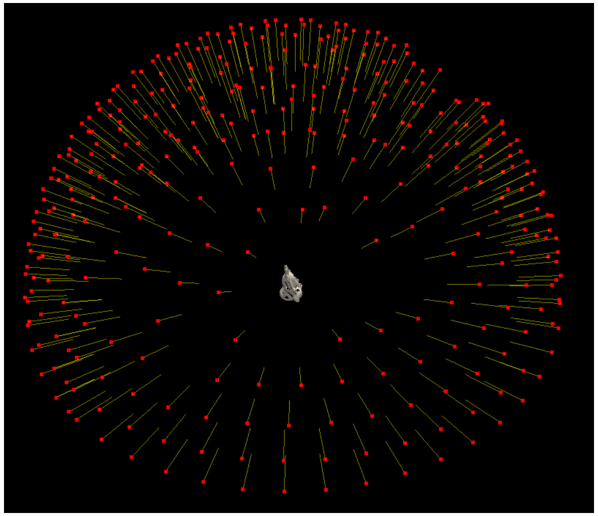
Reconstruction of Synthesis Images (Cont.)

- Poisson Surface Reconstruction result



Reconstruction of Middlebury Dinosaur Model

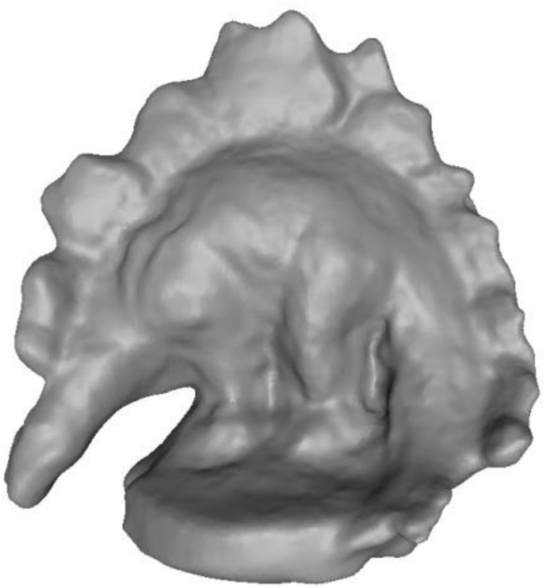
- Middlebury Dino dataset
  - 363 domed views
  - Automatic silhouette extraction
  - 6057 seed points
  - 1872 seed patches after verification



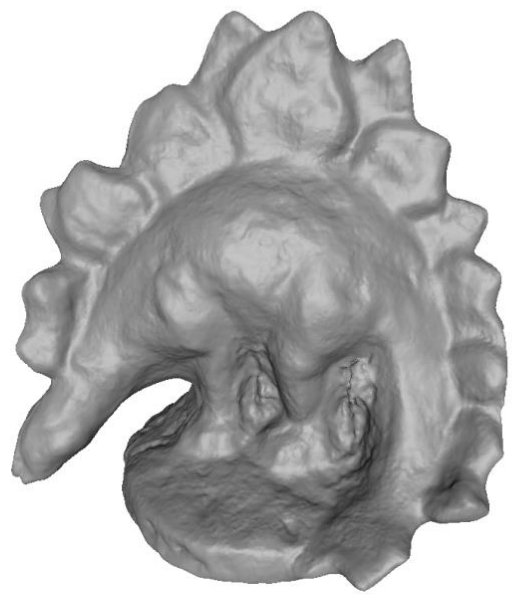
NCTU PAIS Lab

Reconstruction of Middlebury Dinosaur Model (Cont.)

- Poisson Surface Reconstruction result



Cell size 4 (41922 patches)

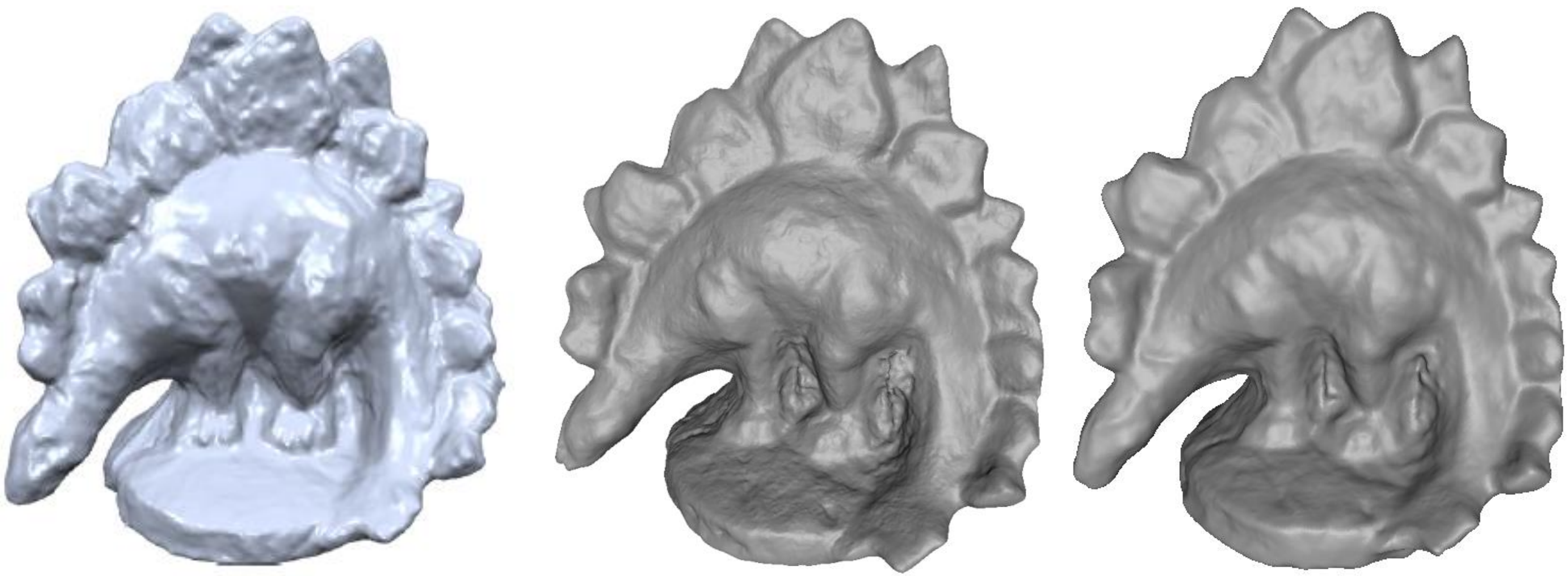


Cell size 2 (210269 patches)

Ground Truth

NCTU PAIS Lab

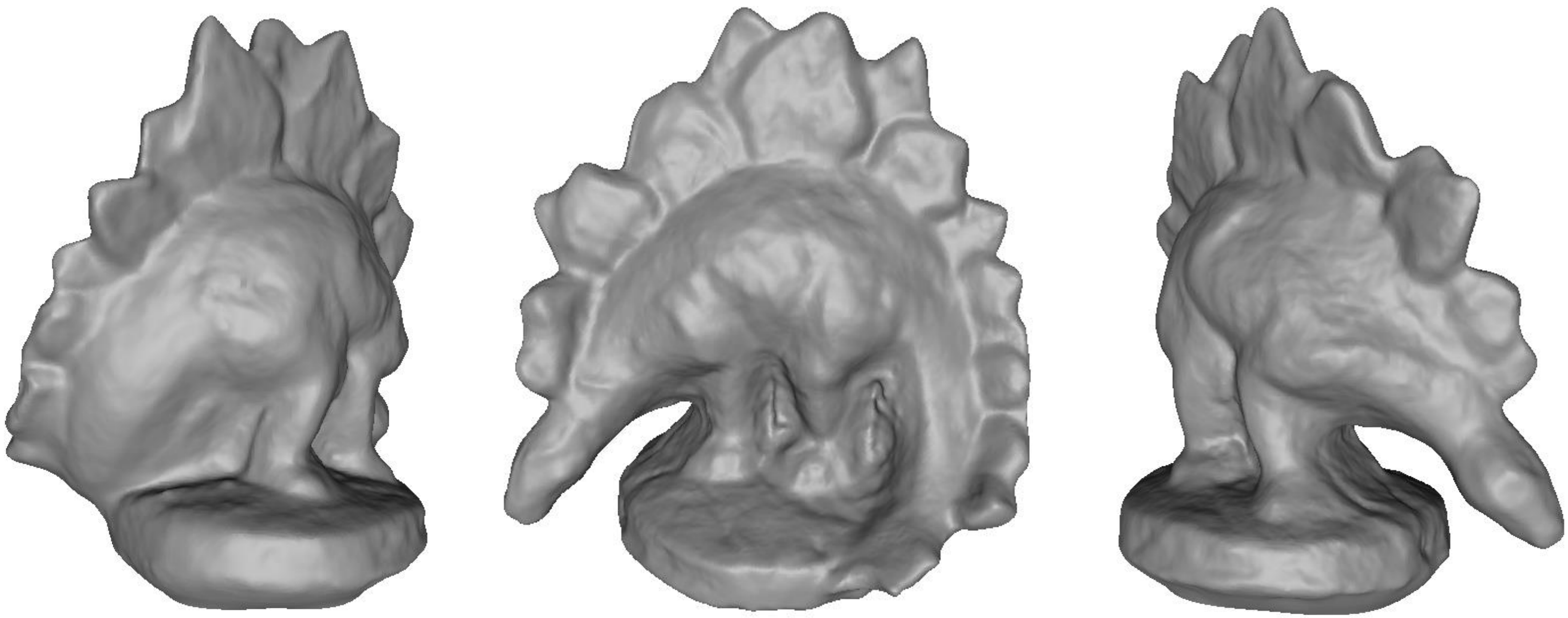
Reconstruction of Middlebury Dinosaur Model (Cont.)



Habbecke's

Ours

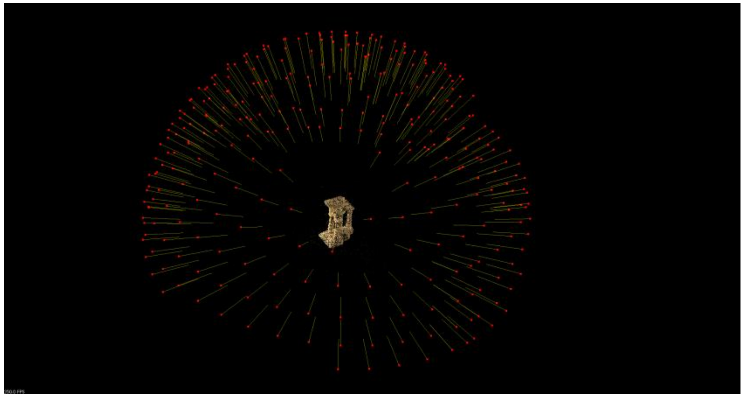
Reconstruction of Middlebury Dinosaur Model (Cont.)



Reconstruction of Middlebury Temple Model

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



PAIS Lab

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

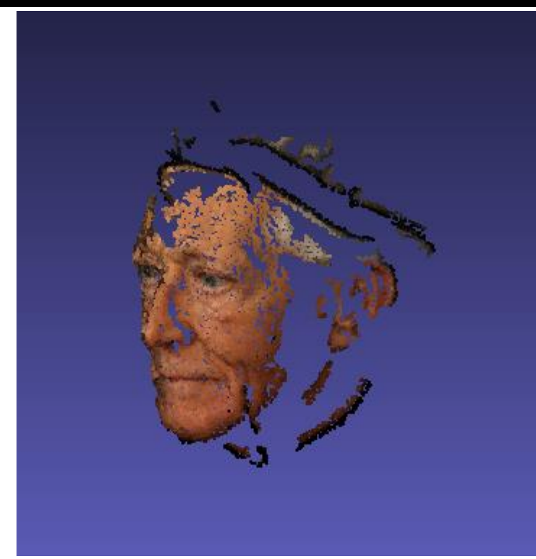
- Human Face dataset
  - 7 views
  - 1687 seed points
  - 1003 seed patches after verification

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.

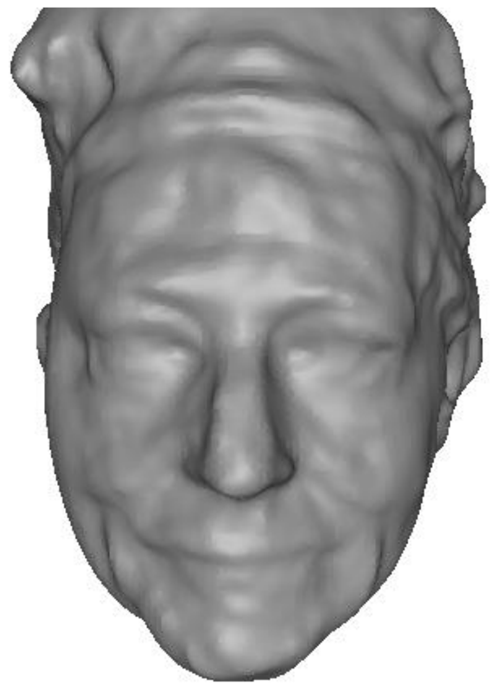
Reconstruction of Human Face Model (Cont.)

- Reconstruction patch model (104115 patches)

Ours	Furukawa et al.
	

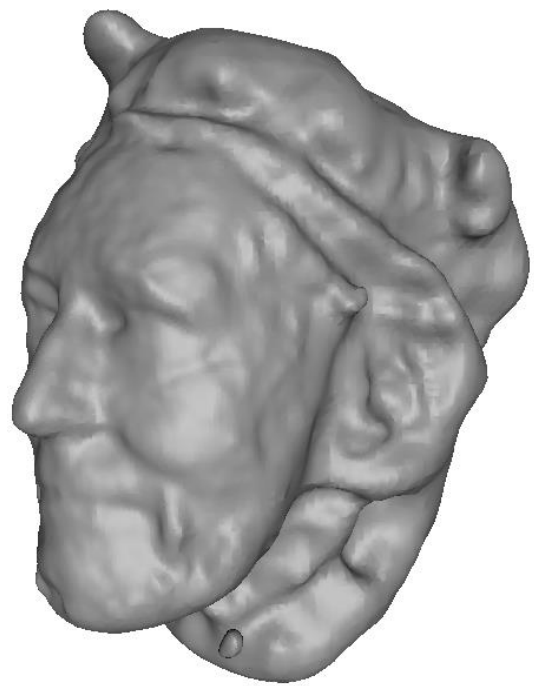
Reconstruction of Human Face Model (Cont.)

- 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