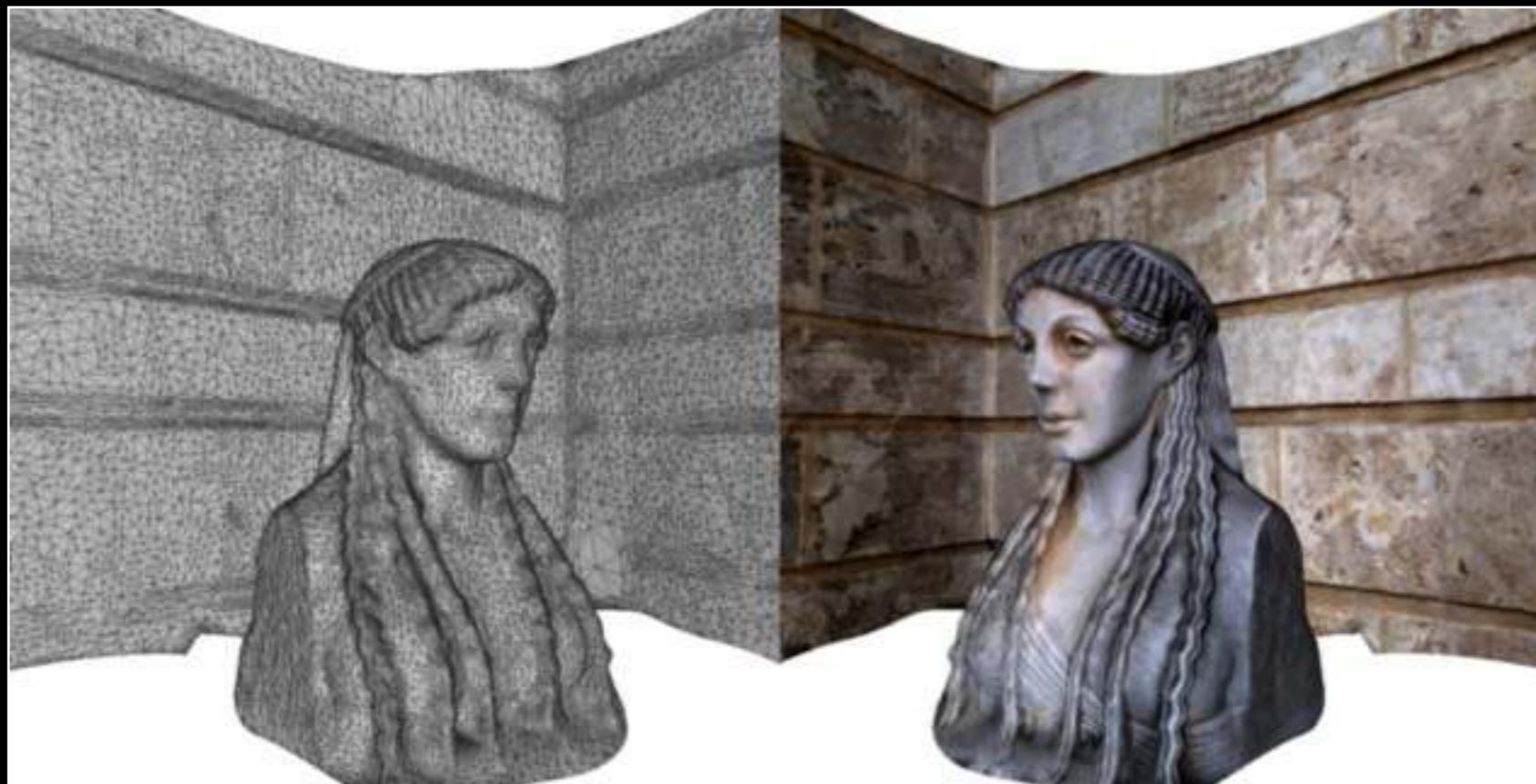


# Automated 3D model reconstruction from photographs

Paul Bourke  
iVEC@UWA



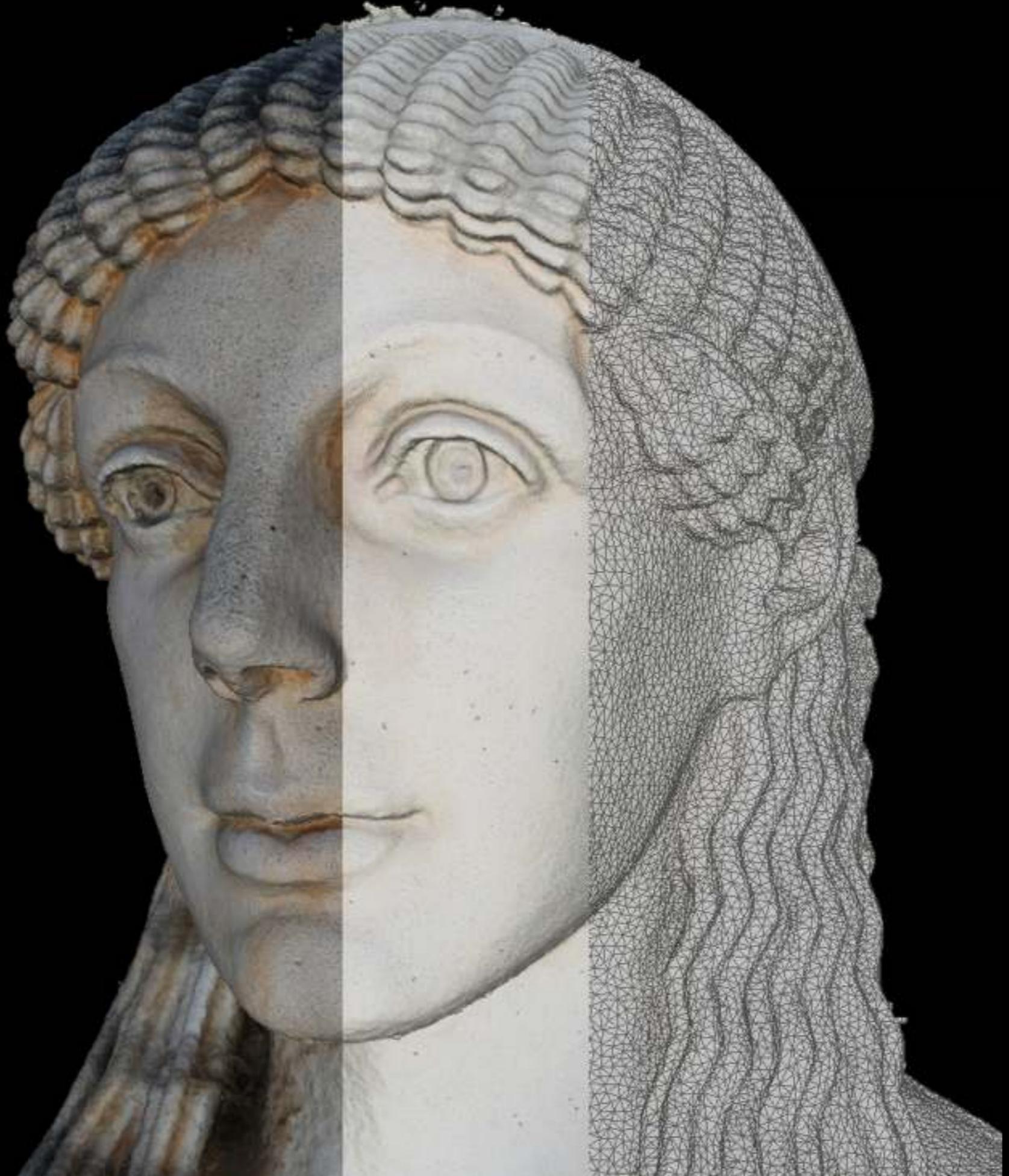
# iVEC : One slide introduction

- Joint venture between the 5 main research institutions in Western Australia.  
UWA, Curtin, Murdoch, ECU, CSIRO.
- Provide researchers at the partners with supercomputing (just hit the top 500 list at number 41), data storage (multiple tens of petabytes), visualisation (three main visualisation nodes).
- I wear three hats
  - Researcher at The University of Western Australia
  - Head of the iVEC visualisation team
  - Director of the iVEC@UWA Centre at The University of Western Australia
- Reconstruction activities driven from mainly Archaeology/Heritage, Geoscience.  
Archaeology: indigenous rock art and marine archaeology.

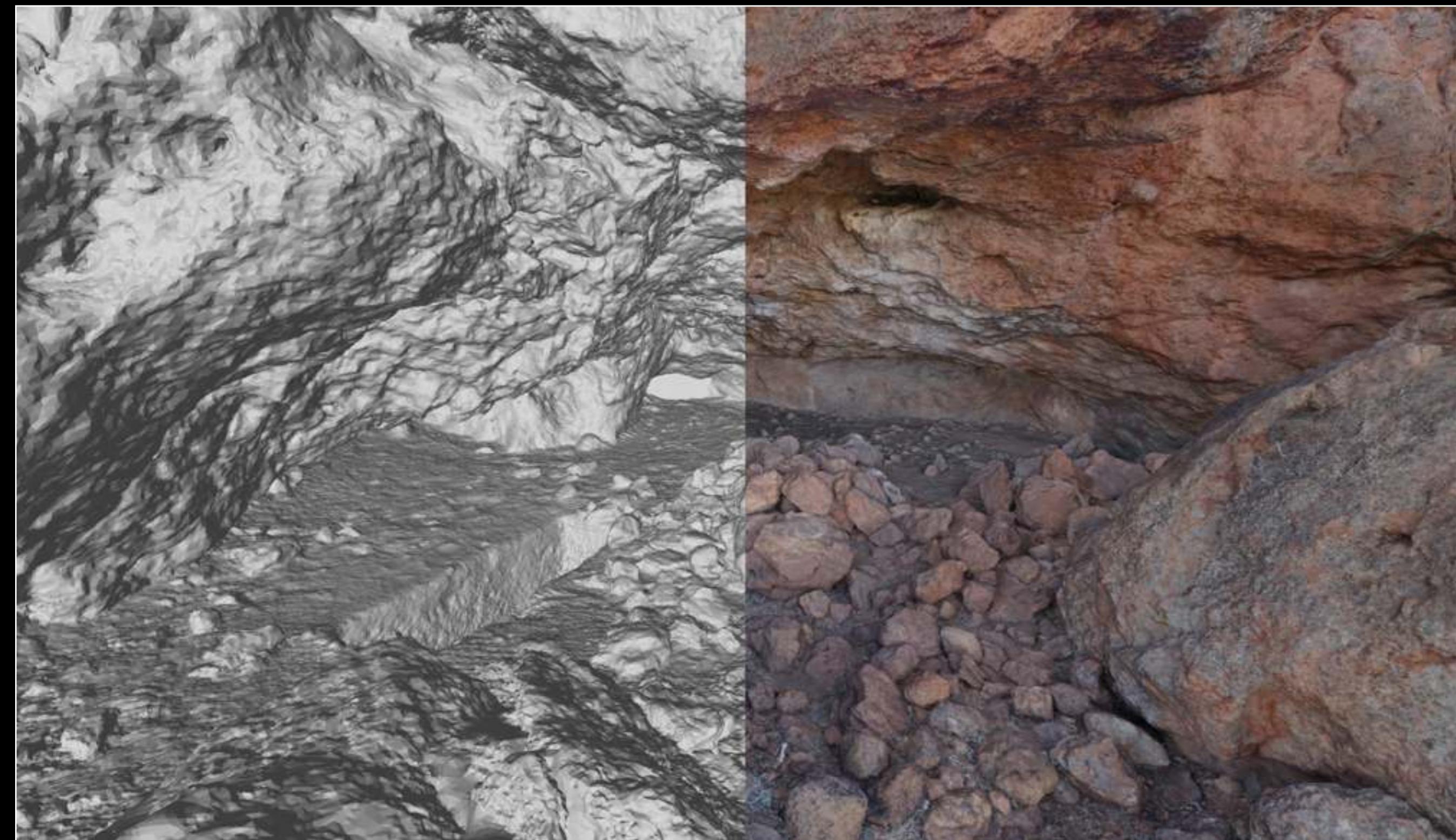








Diotima (Mistress of Pericles)  
Credited with the concept of  
platonic love.  
Statue at UWA



Rock shelter, Weld Range, Western Australia  
Derived from 350 photographs



# Outline

- Introduction, Outcomes, Motivation
- Software
- Photography
- Example 1: 2.5D
- Geometry processing
- Example 2: 3D
- Other topics (Accuracy)

---

Break

---

- Limitations and challenges
- Worked example 3: Grinding stone
- Additional / Advanced topics
- Questions and discussion

These slides will be made available online  
at <http://paulbourke.net/vsmm2014/>

Sample datasets available on request.

# Outcomes

- Familiarity with the state of the technology.
- Knowing what questions to ask.
- Understand the terminology.
- Familiarity with the software and tools.
- Some expectations of the limitations.
- Knowledge of a range of applications/research the technology is being applied to.
- Will not be overly technical but happy to discuss further after the formal part.

# 3D reconstruction from (ad hoc) photographs

- Goal: Automatically construct 3D geometry and texture based solely upon a number of photographs.
- Similar to traditional photogrammetry but employs different algorithms.
- Creating richer objects (compared to photographs) for recordings in archaeology and heritage.
- Create geometric models suitable for analysis, eg: in geology or geoscience.
- Create digital assets for virtual environments.
- Wish to avoid any in-scene markers required by some solutions.  
Often impractical (access) or not allowed (heritage).
- Want to target automated approaches as much as possible.  
[Recent site survey recorded 100's of objects].
- All the examples shown will be applications from the presenters work.

# Applications : Virtual worlds, Serious gaming

- Creating 3D assets for virtual environments, serious games.
- Removes the need for time consuming 3D modelling.
- Removes the interpretation that occurs when modelling organic / complicated shapes.



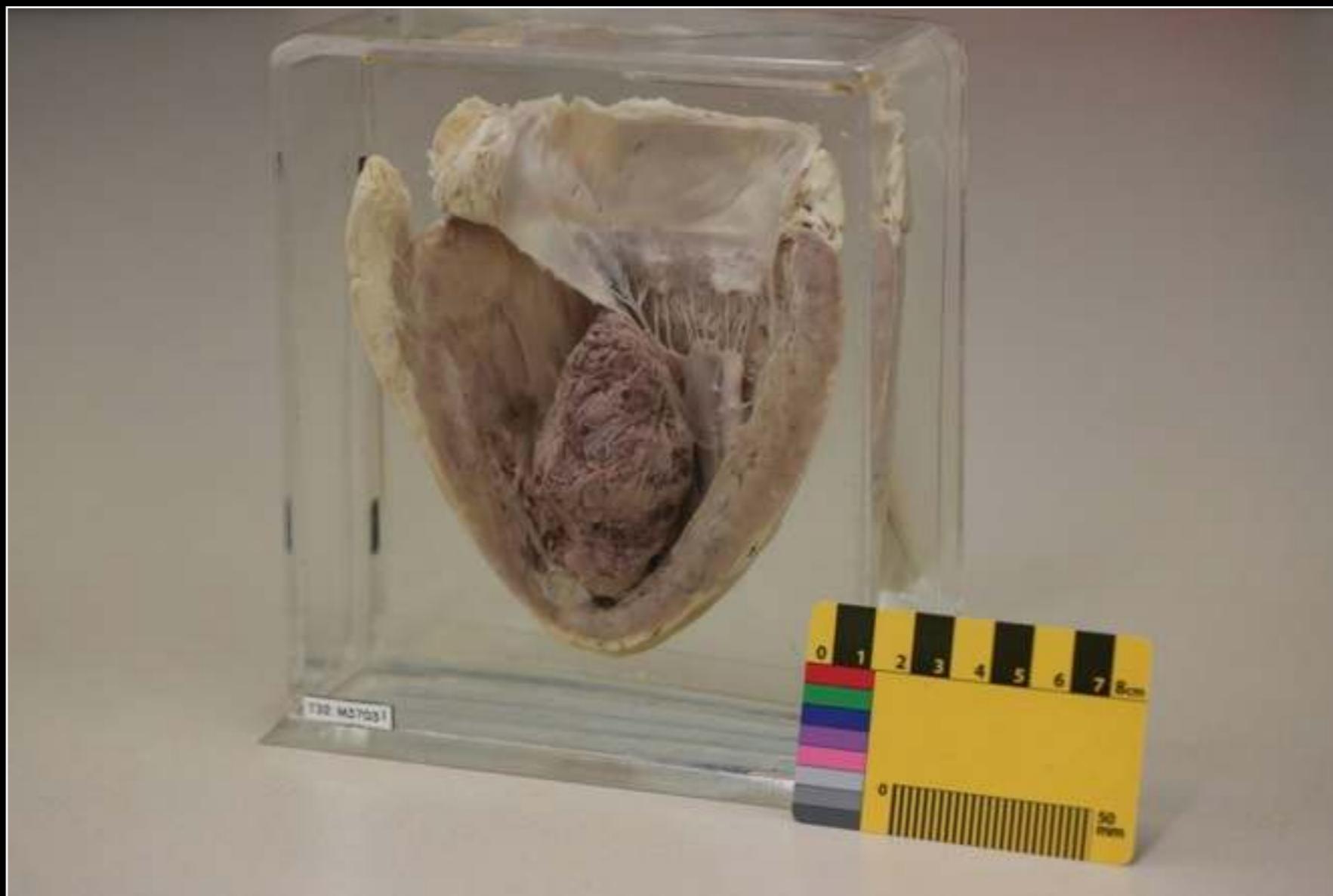
# Applications : Assets for virtual environments

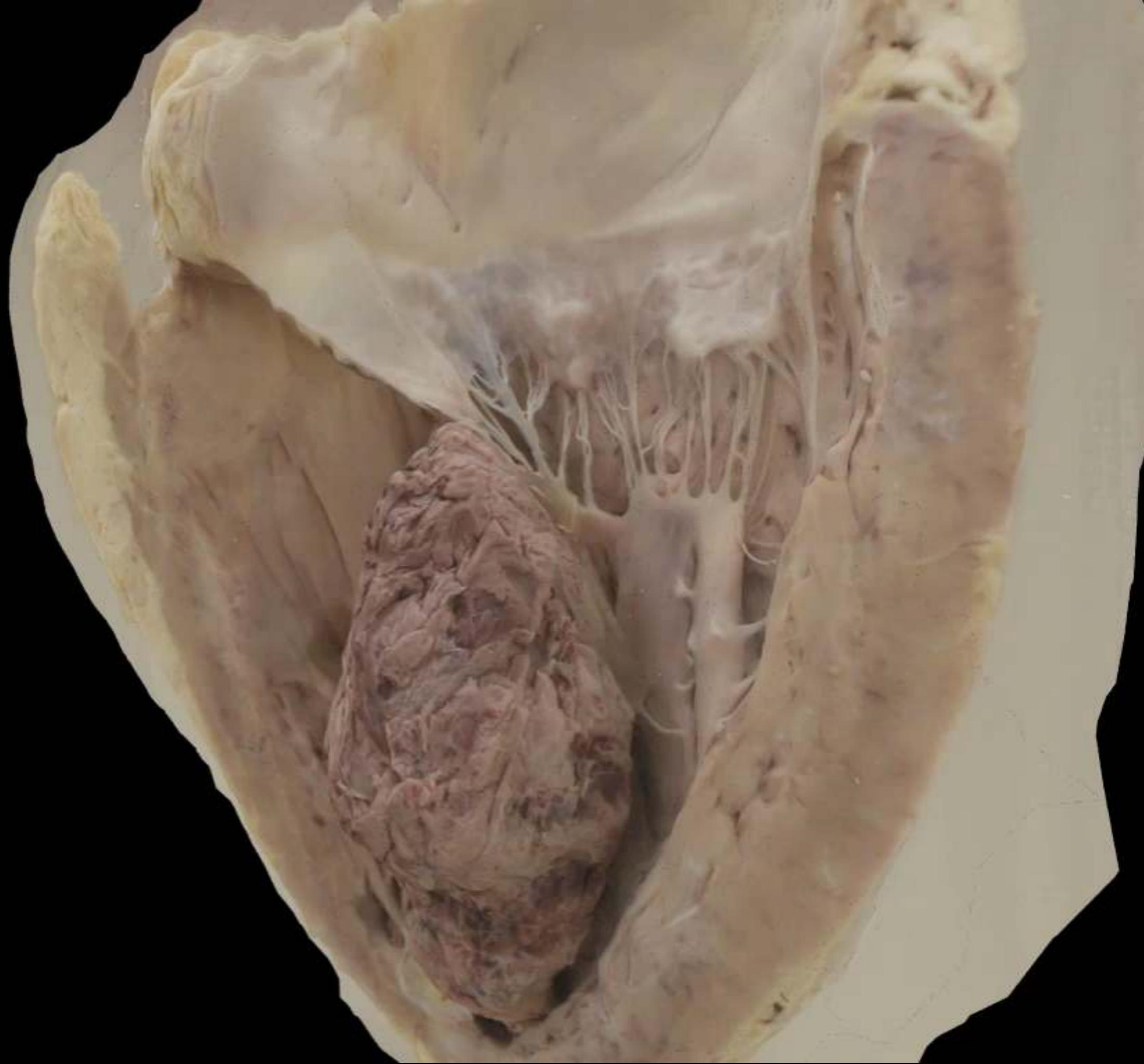


Beacon Island Virtual environment

# Applications : Teaching in medicine

- Medical applications (eg: Feeding mothers breast volume)
- Non intrusive and non contact capture can have advantages.
- Capture of 3D objects for forensic analysis.
- Current project digitising teaching collection of diseased body parts.

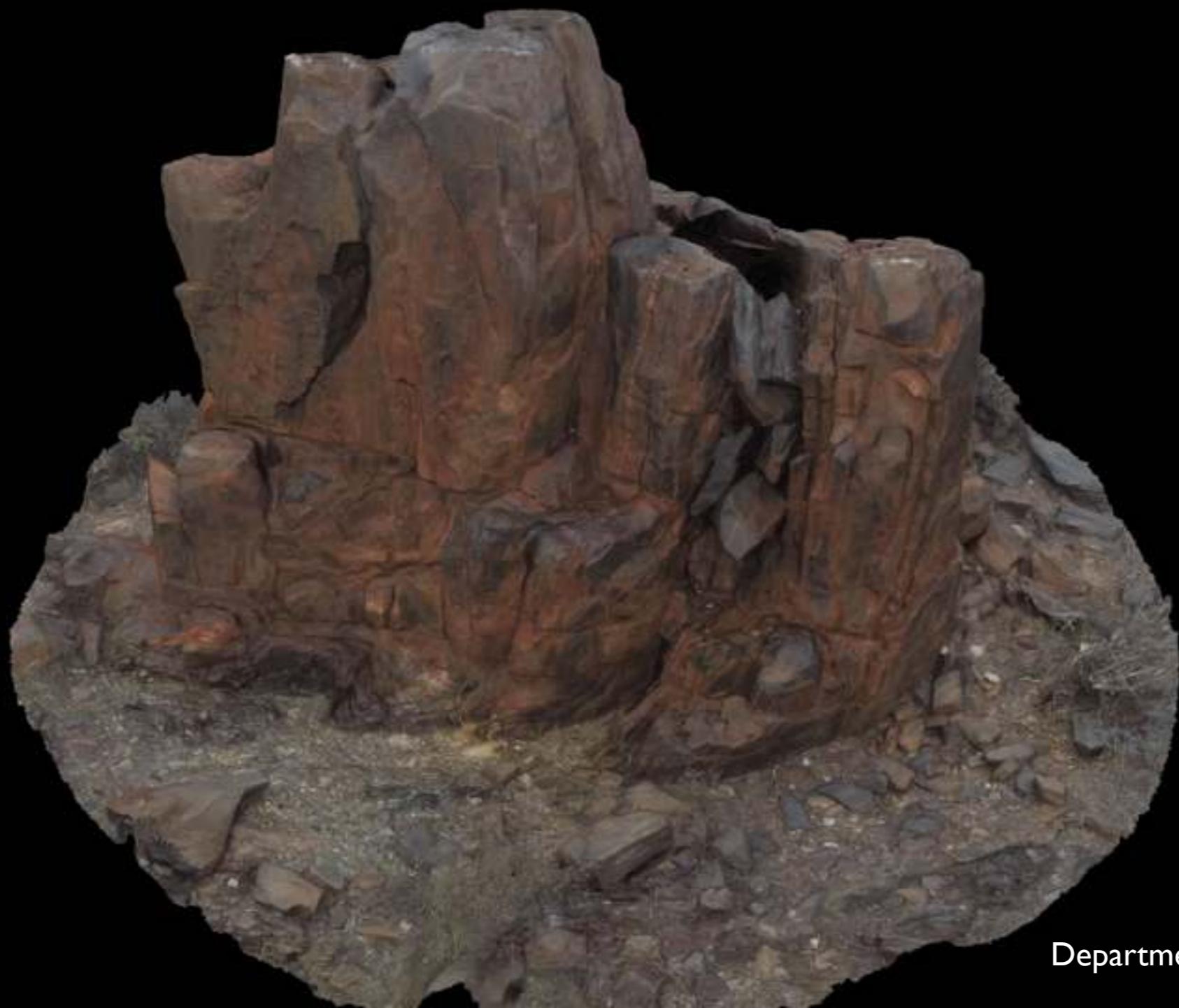






# Applications : Geoscience

- Capturing geological structures for analysis.
- Often in difficult terrain and remote locations.



# Applications : Mining

- Capture rock volume removed in mining operations. Accuracy is critical.
- Advantages from a safety perspective, don't have to close down operations to allow surveyors on site.



# Applications : Artefacts in cultural heritage



Remote in-the-field capture



Ngintaka headdress, South Australia Museum

# Applications : Digital capture in heritage



Dragon Gardens, Hong Kong

# History

- Photogrammetry is the general term given to deriving geometric information from a series of images.
- Initially largely used for aerial surveys, deriving landscape models.  
Originally only used a stereoscopic pair, that is, just two photographs.
- More recently the domain of machine vision, for example: deriving a 3D model of a robots environment.
- Big step forward was the development of SfM algorithms: structure from motion.  
This generally solves the camera parameters and generation of a 3D point cloud.
- Most common implementation is called Bundler: “bundle adjustment algorithm allows the reconstruction of the 3D geometry of the scene by optimizing the 3D location of key points, the location/orientation of the camera, and its intrinsic parameters”.

# Other technologies

- In some areas it is starting to replace technologies such as laser scanning. LIDAR - light detection and ranging.
  - particularly so for capture in difficult locations
  - only requires modest investmentQuestions of accuracy to be discussed later.
- Another technology are so called depth cameras.
  - Primesense (eg: Kinect)
  - Structured light techniques (eg: Artec Scanner)Operate in limited range of lighting conditions, data tends to be quite noisy. Limited range.
- Light field cameras (plenoptic camera).
  - Captures an array of images from a grid of positions
  - Currently resolution is too low.



LIDAR



Structured light

# Software (A selection only)

- Processing pipeline from a number of opensource projects [HPC]

- SiroVision

- PhotoScan [Desktop]

- PhotoSynth

- PhotoModeller Scanner

- 123D Catch

- Visual SfM (Structure from Motion)

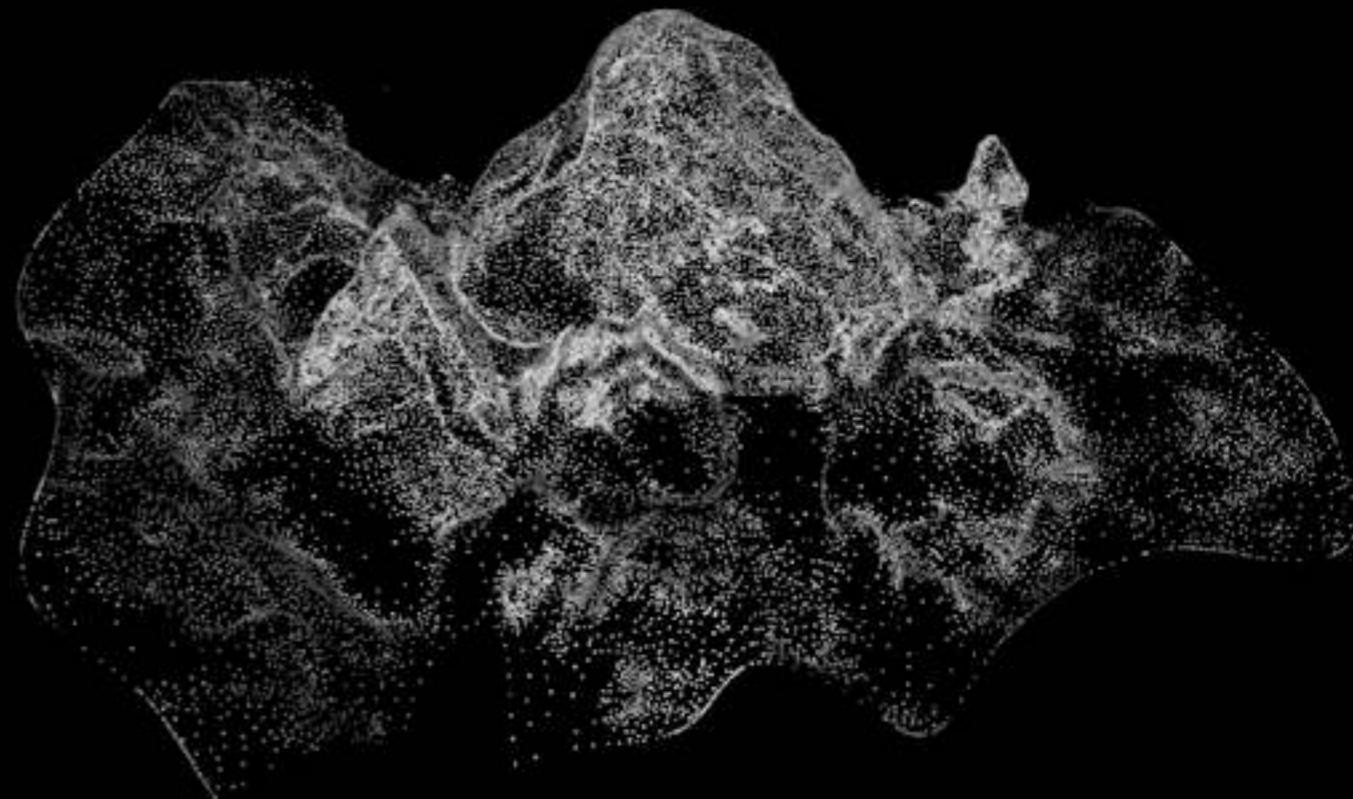
- Apero

- AdamTech solution

- iWitness Pro

- .... lots of others.

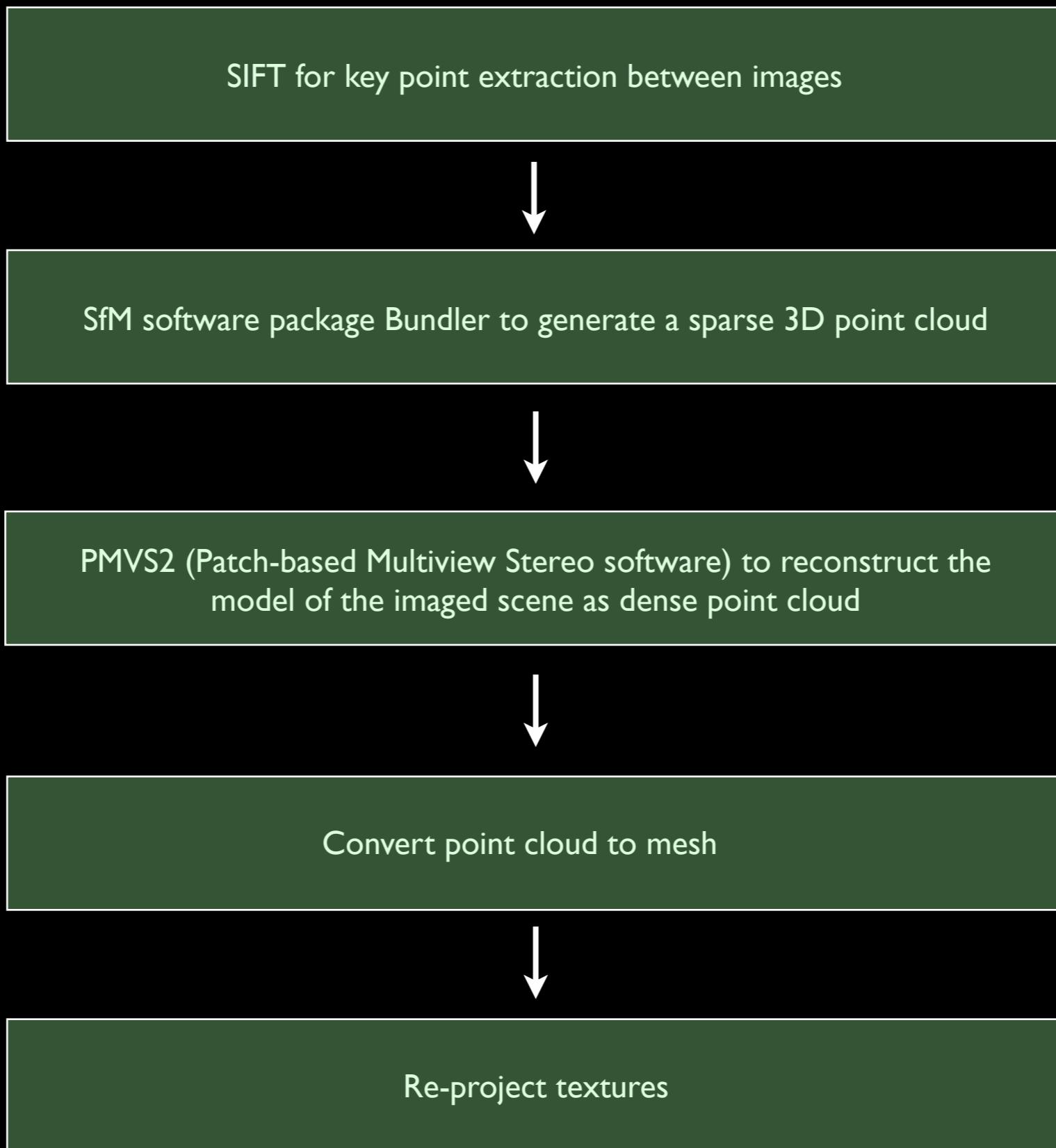
Major aspect of the work we are doing is to evaluate the various tools.



# Software : Pipeline components

- Perform lens calibration (only done once, increasingly optional).
- Read images, correct for lens, and compute feature points between them.  
(eg: SIFT - scale invariant feature transform)
- Compute camera positions and other intrinsic camera parameters.  
(eg: Bundler, SfM - Structure from Motion, <http://phototour.cs.washington.edu/bundler/>)
- Create sparse 3D point cloud, called “bundle adjustment”.  
(eg: PMVS - Patch-based Multi-view Stereo, <http://www.di.ens.fr/pmvs/>)
- Create dense point cloud.  
(eg: CMVS - Clustering Views for Multi-view Stereo, <http://www.di.ens.fr/cmvs/>)
- Form mesh from dense point cloud.  
(eg: ball pivoting, Poisson Surface Reconstruction, Marching Cubes)
- Reproject images from camera positions to derive texture segments.
- Optionally simplify mesh (eg: quadratic edge collapse decimation) and fill holes.
- Export in some suitable format (eg: OBJ files with textures).

# Software : Typical HPC pipeline



# Software : Pipeline - Photographs

- Don't take two photos from the same position.
- Obviously can't reconstruct what is not photographed.
- In general, more is better. Can always process just a subset of the images.



# Software : Pipeline - Sparse point cloud

- Find matching points between photographs, feature point detection.  
SIFT - scale invariant feature transform
- Compute camera positions and other intrinsic camera parameters.  
Bundler, SfM - Structure from Motion

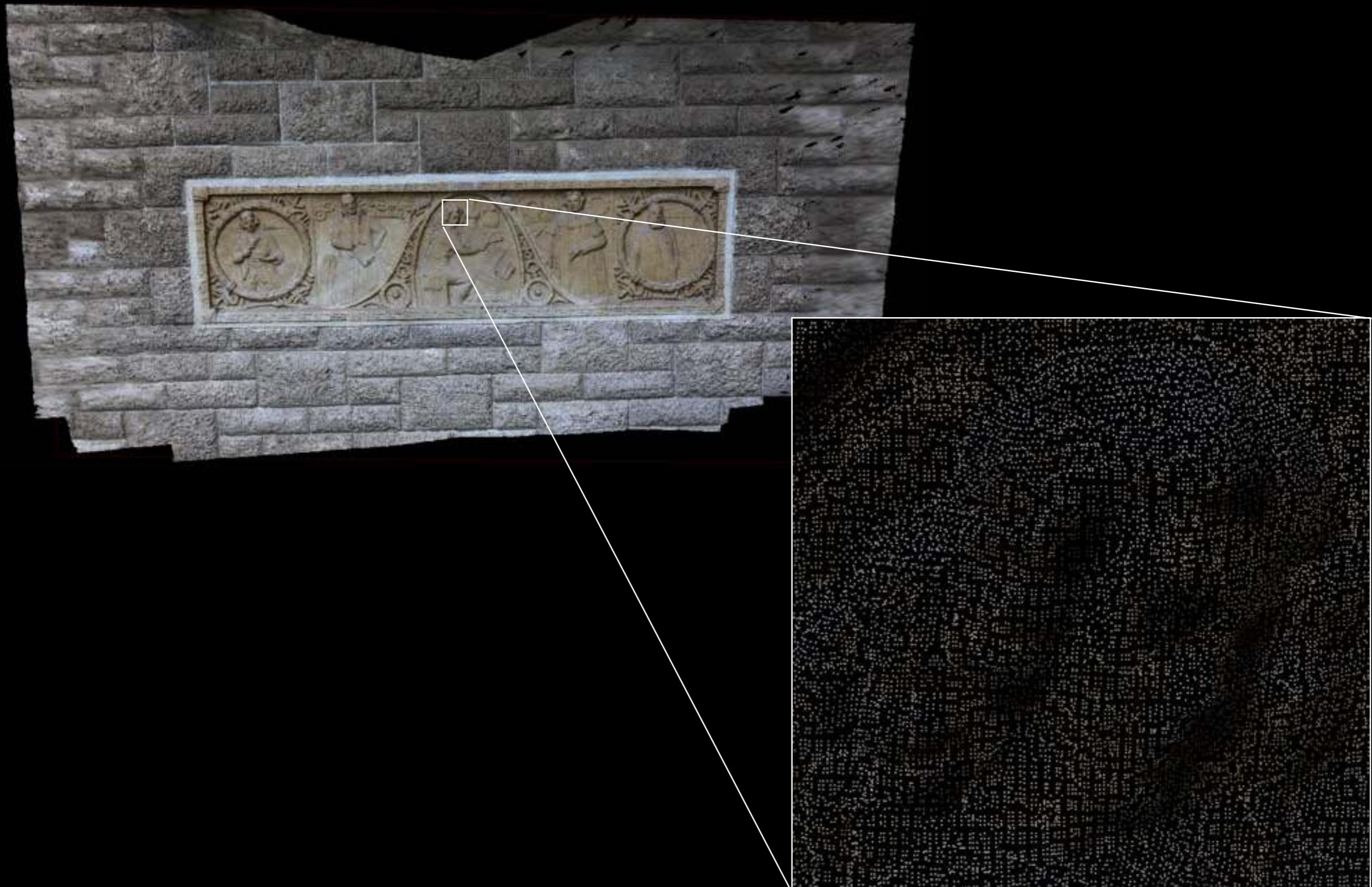


# Software : Pipeline - Dense point cloud

- CMVS - Clustering Views for Multi-view Stereo.



# Software : Pipeline - Dense point cloud



# Software : Pipeline - Mesh generation

- Various algorithms: Ball pivoting, Poisson Surface Reconstruction, Marching Cubes.
- Optionally simplify mesh (eg: quadratic edge collapse decimation) and fill holes.



# Software : Pipeline - Texture mesh

- Re-project photographs from derived camera positions onto mesh.

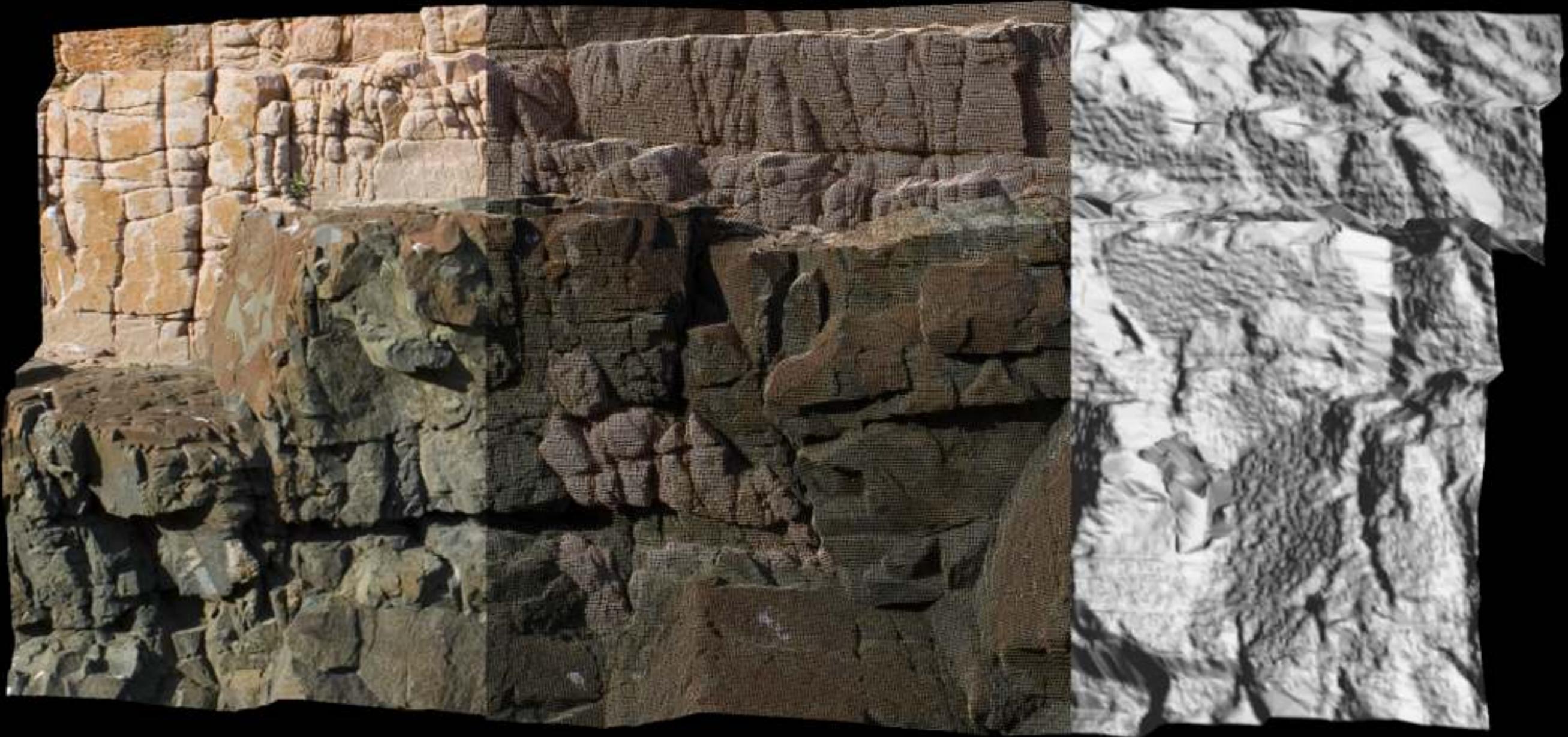


# Software : Pipeline - Export



# Software : Sirovision (<http://sirovision.com>)

- Captured from 2 images only, stereo pairs but with wide base line separation.
- With in-scene markers and calibrated lens claims 3 to 5cm accuracy at 100m distance.
- Targeted mining industry, developed by CSIRO.



Textured

Mesh

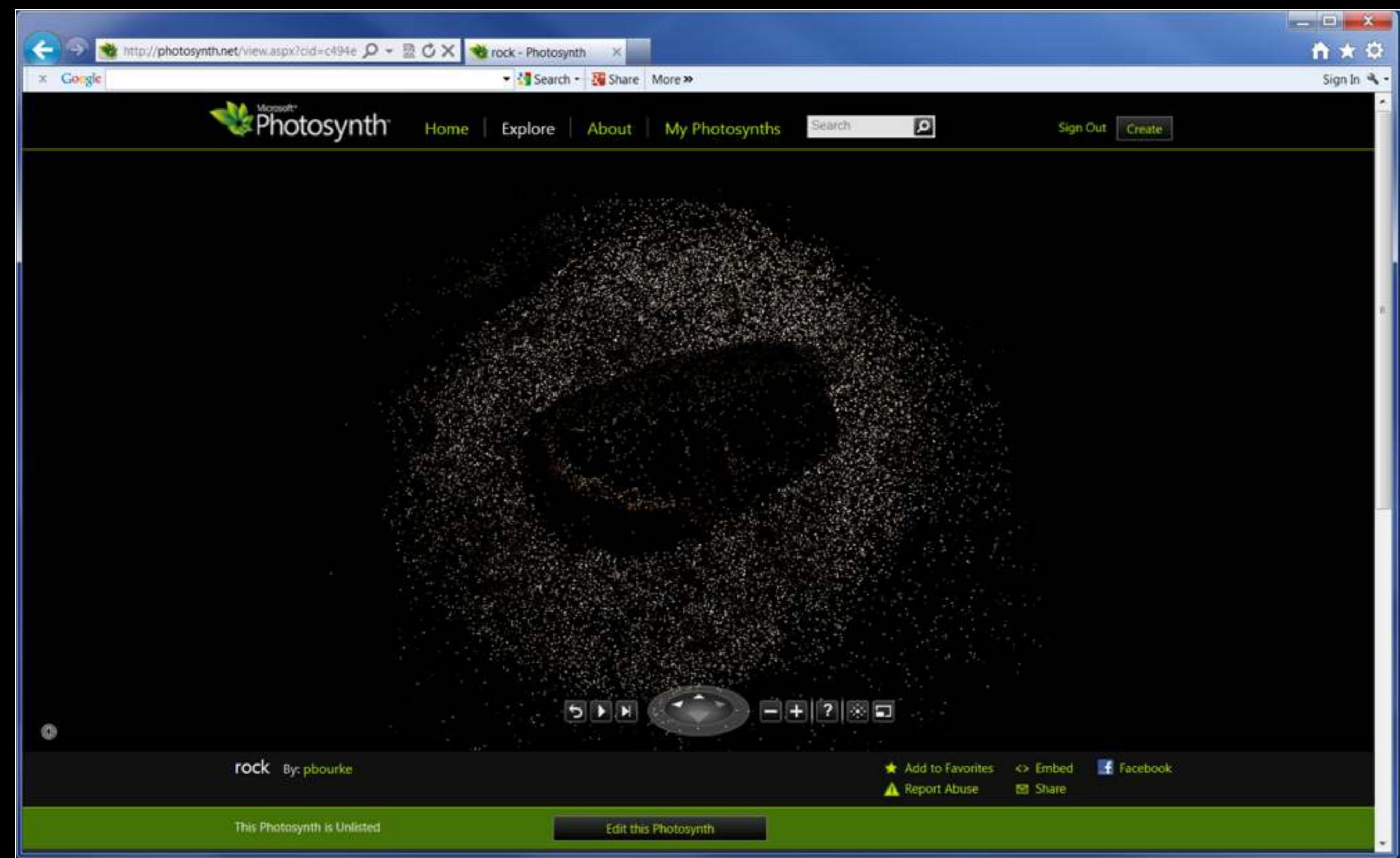
Surface

# Software : PhotoSynth

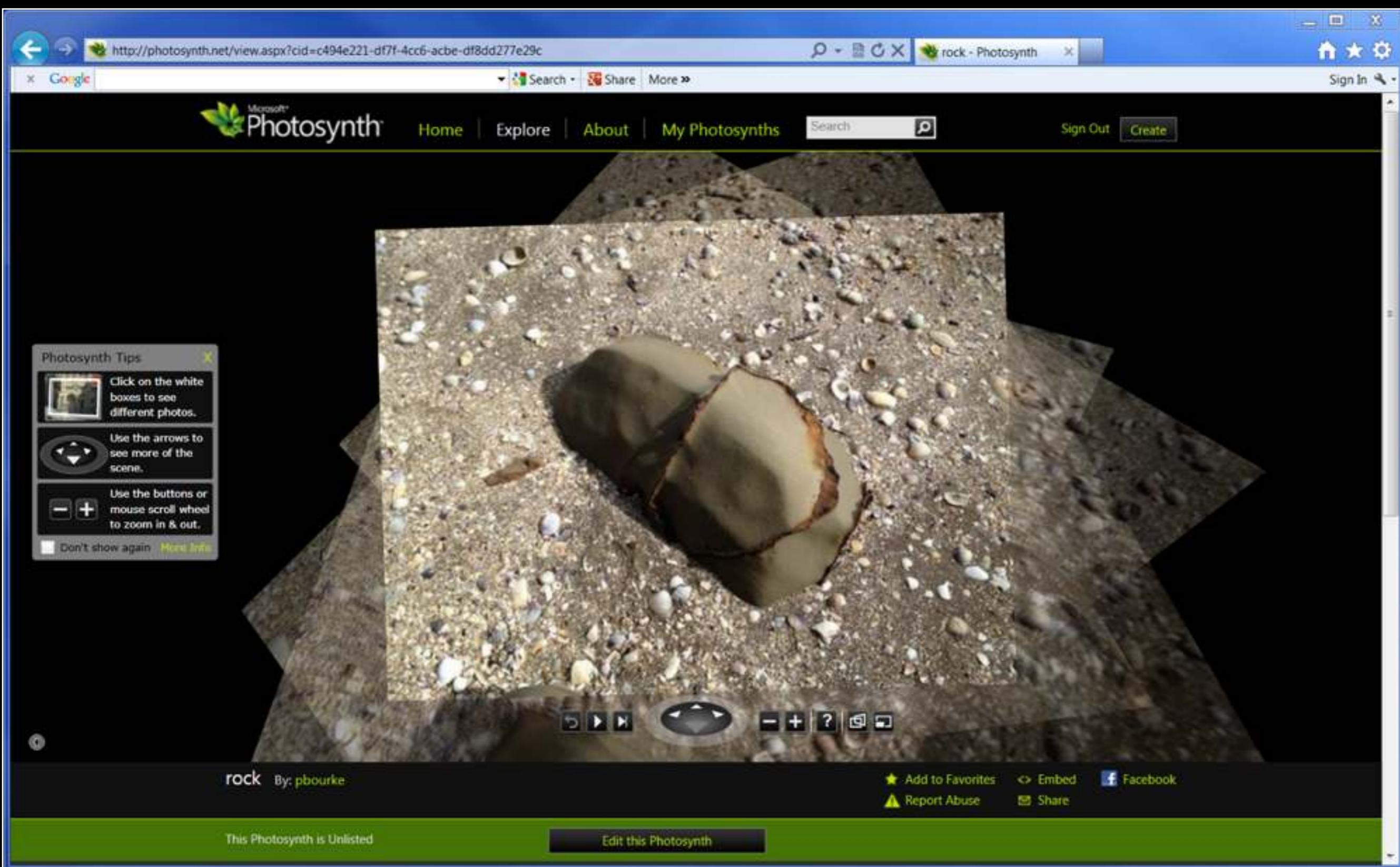
- Microsoft, MSWindows only (obviously)  
<http://photosynth.net>
- Based upon Bundler. GUI front end, computed remotely.
- Provides a “image effect” based upon reconstructed surface.
- Can be useful for identifying image sets for other pipelines.
- Not possible to extract the mesh/texture data from within the online software.
- Synth Export  
<http://synthexport.codeplex.com/>  
Provides point cloud and camera parameter export. Would need to reconstruct mesh by other means.
- Not a leading edge tool any more.



# Software : PhotoSynth



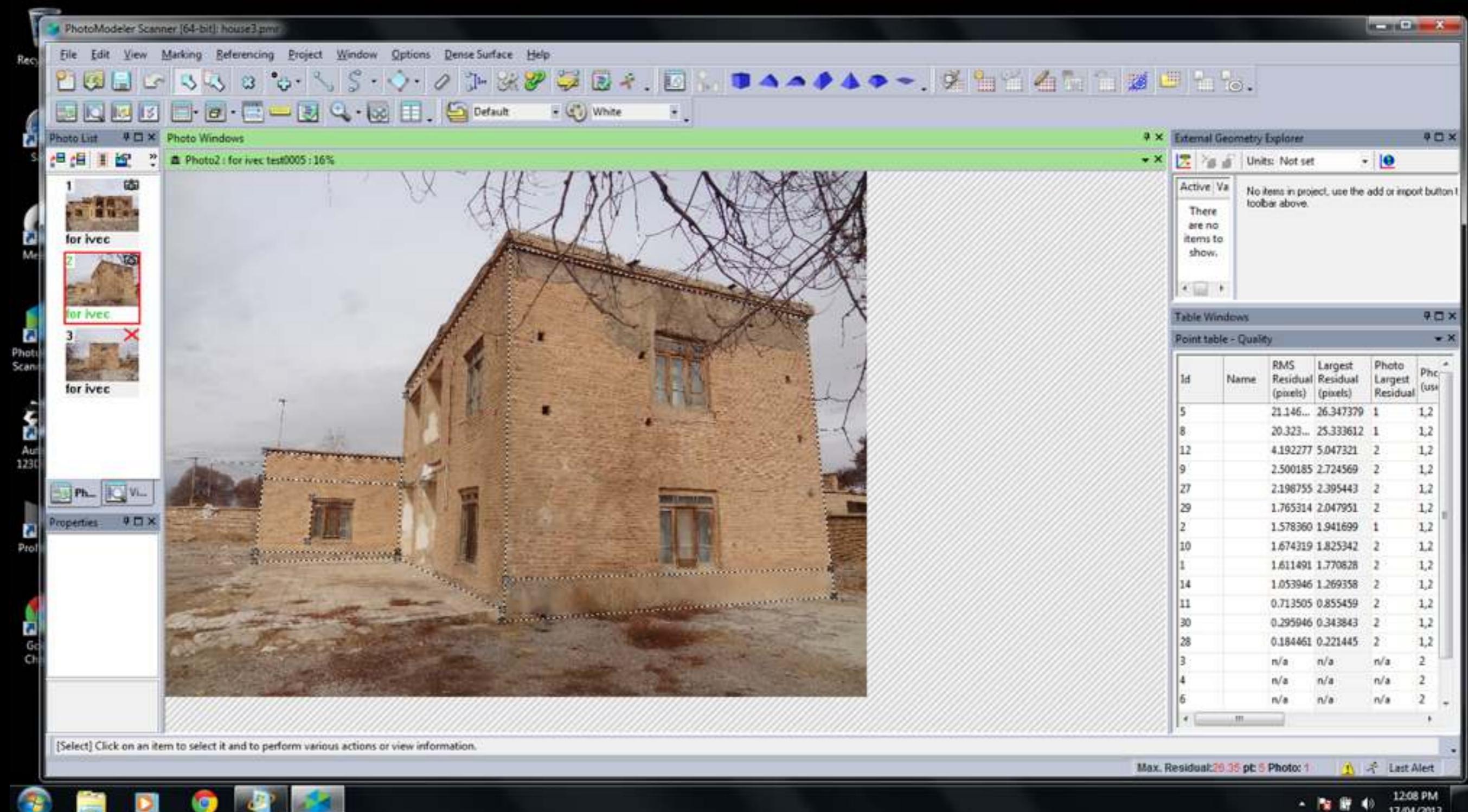
# Software : PhotoSynth



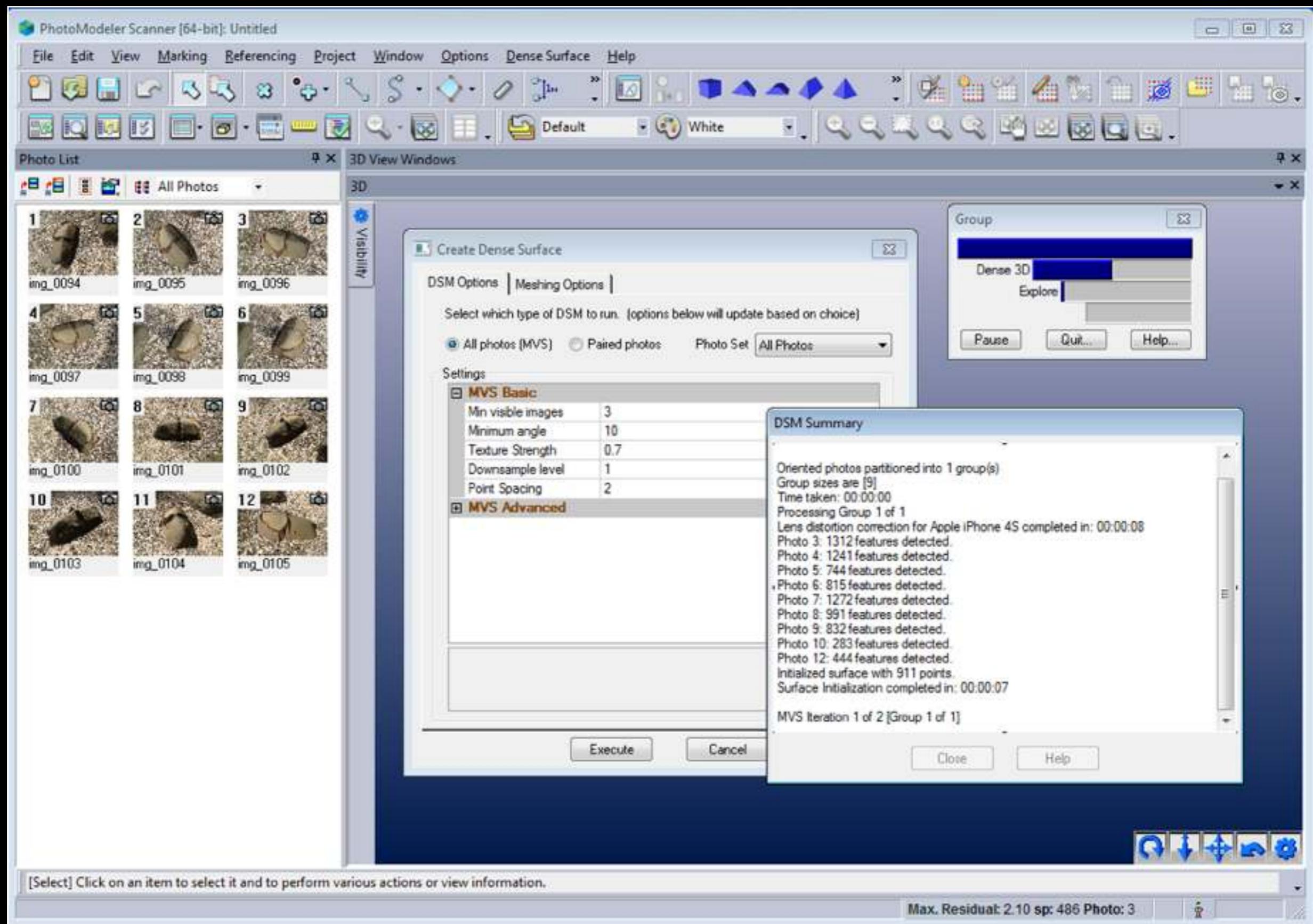
# Software : PhotoModeller Scanner

- From EOS systems.
- <http://www.photomodeler.com/>
- Comes in two flavours, the standard package is for human driven extraction of rectangular objects such as building facades.
- PhotoModeller Scanner is for more organic shapes.
- Claims to be capable of very accurate results, generally has a more rigorous procedure.
- Requires manual interaction.
- MSWindows only.
- A contender to PhotoScan but to date have not had better results.
- VERY slow compared to almost everything else.

# Software : PhotoModeller



# Software : PhotoModeller Scanner

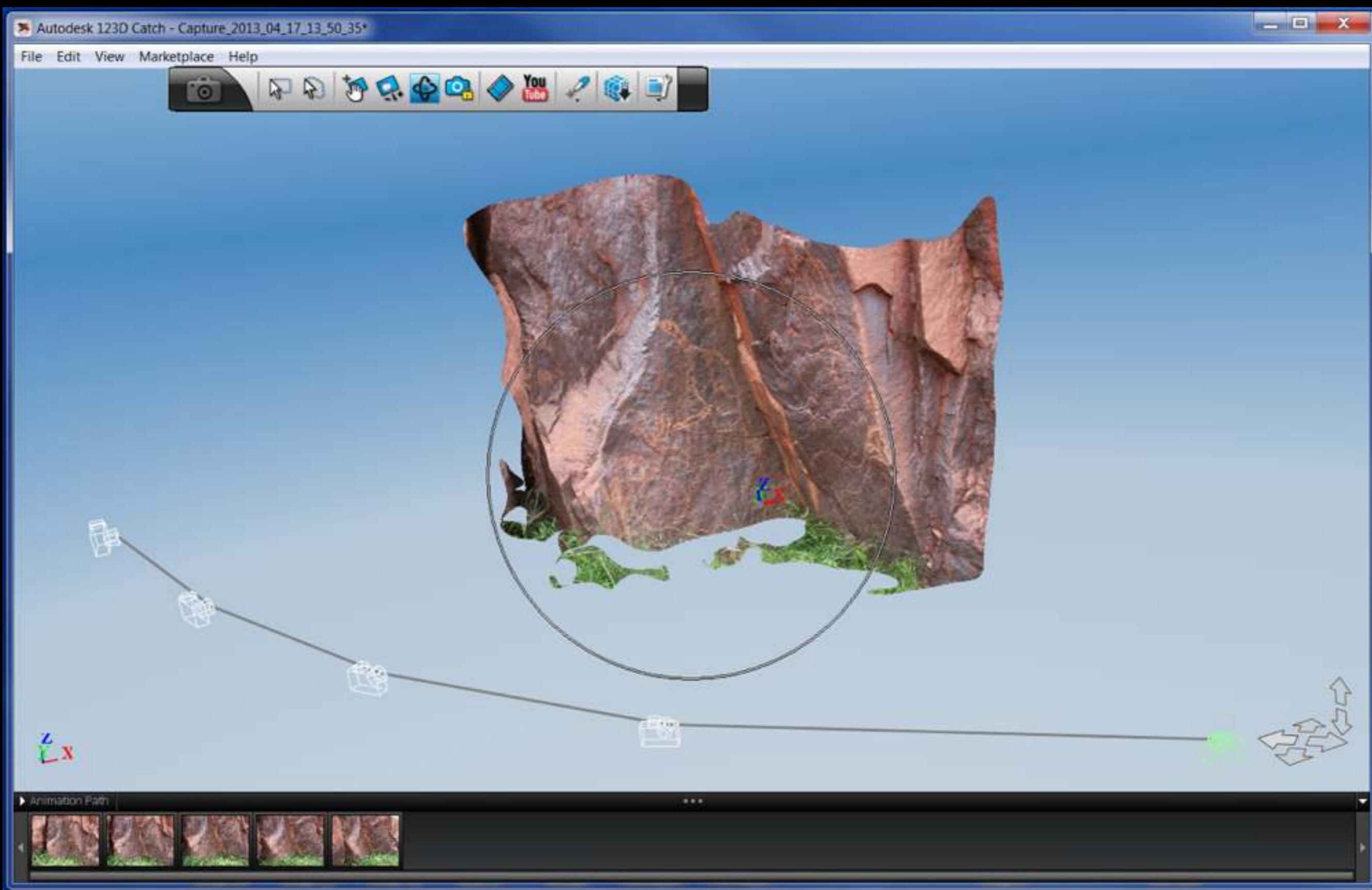


# Software : 123D Catch

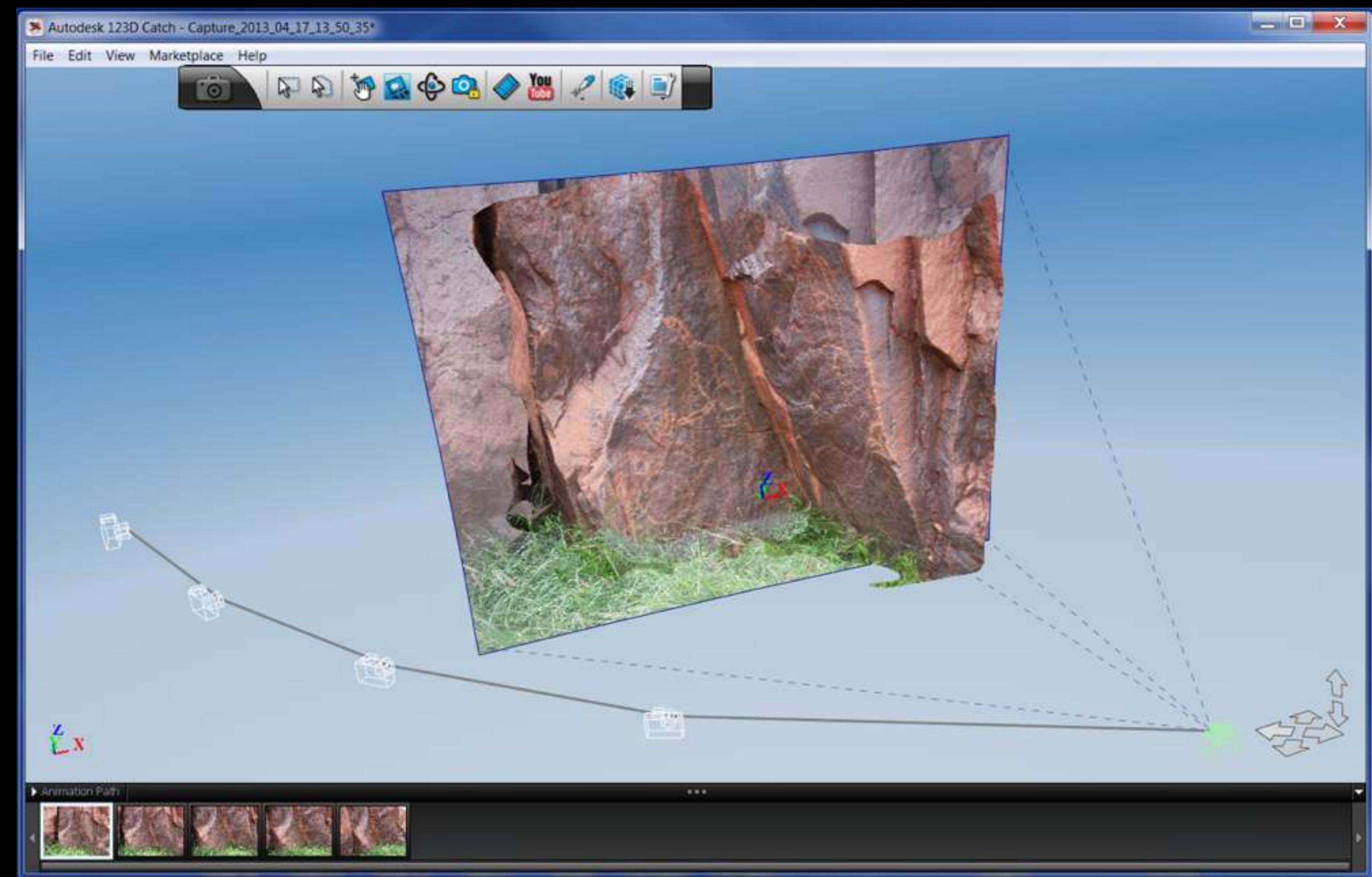
- From AutoDesk.
- Free.
- Cloud based so requires an internet connection.
- Reasonable rate of success but no option to change algorithm parameters if things don't work.
- Does not provide access to intermediate data, such as the point cloud.
- No option for (standard) camera calibration.
- MSWindows only GUI.
- No longer a leading edge solution.



# Software : 123D Catch



# Software : 123D Catch

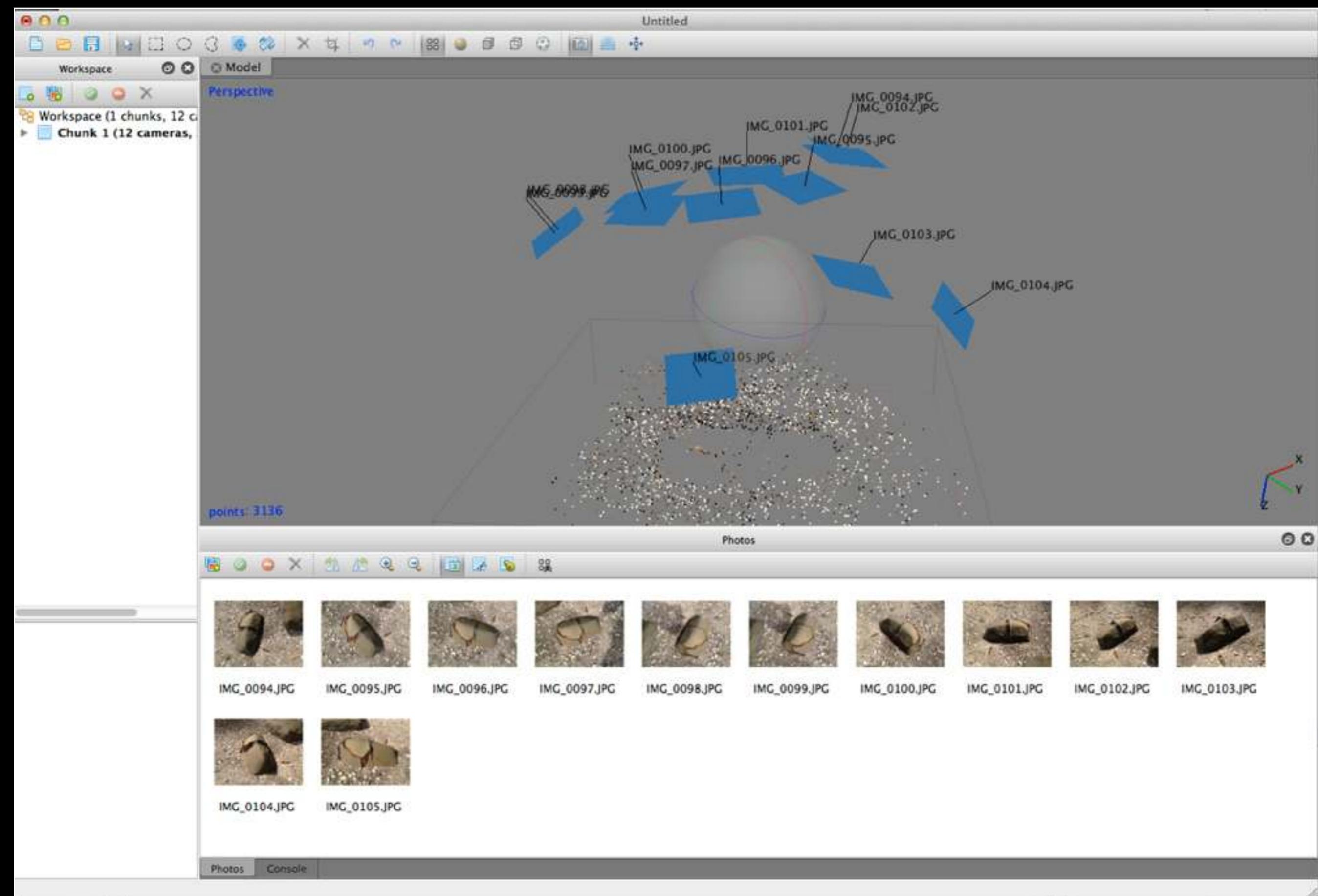


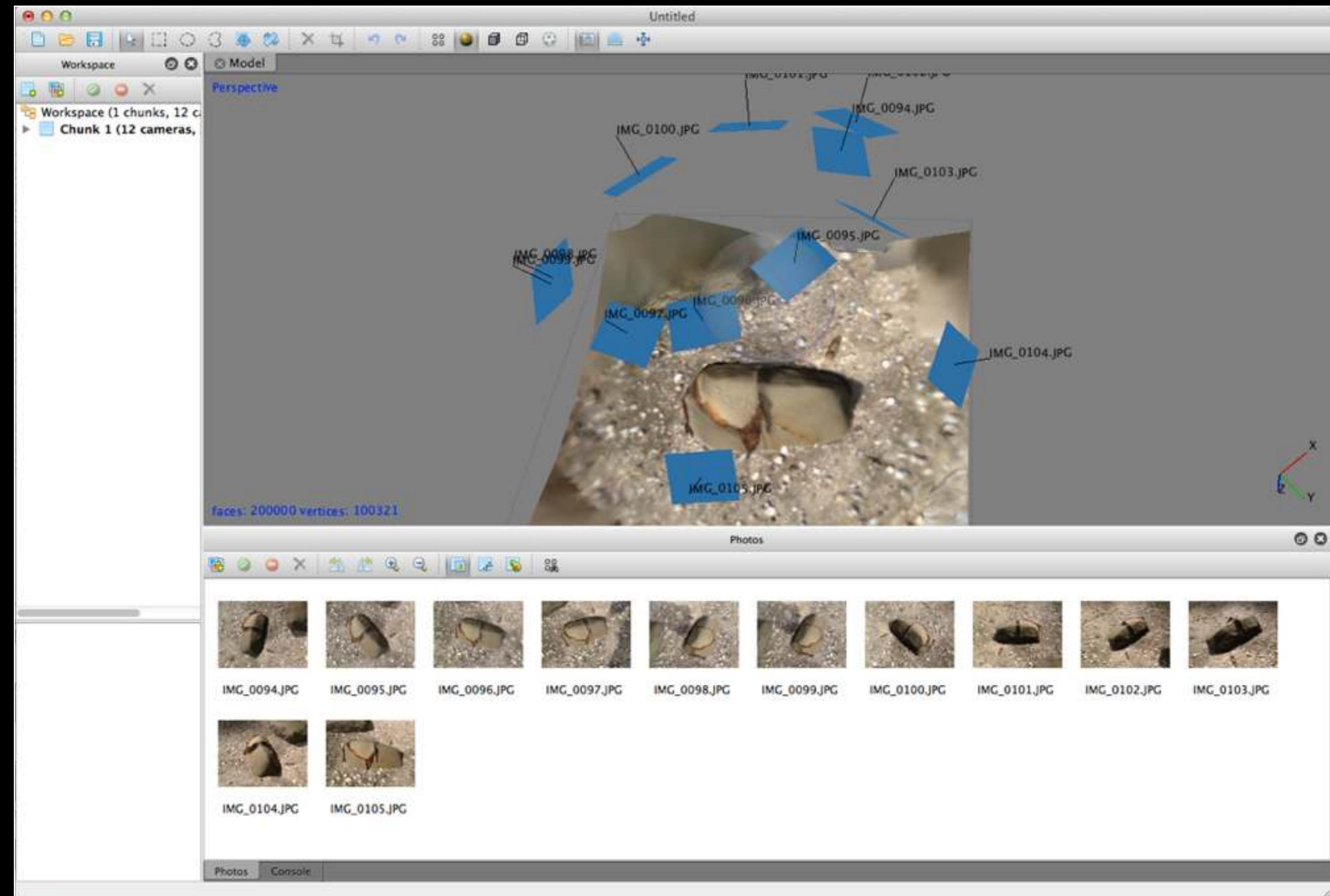
# Software : Visual SfM - Bundler

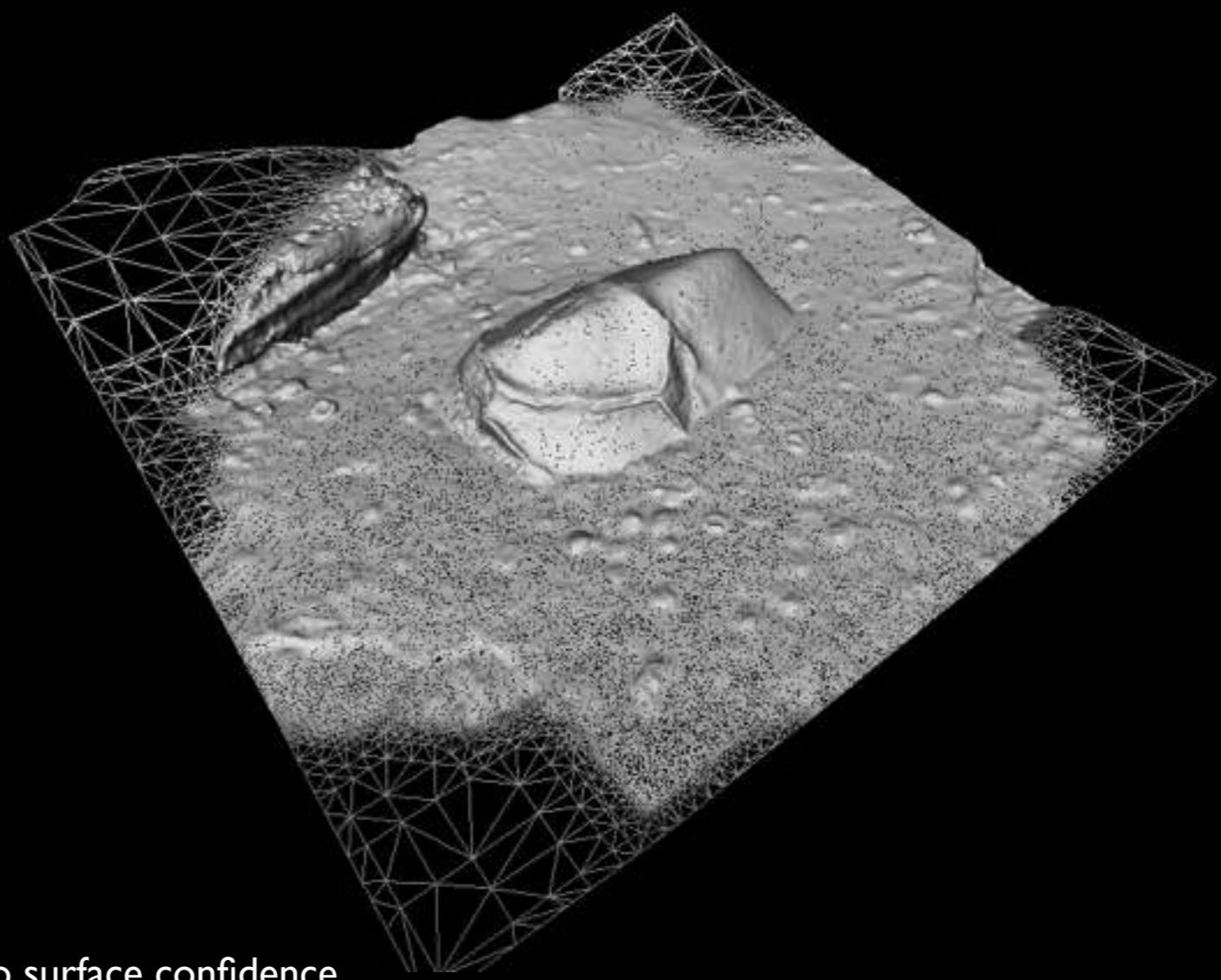
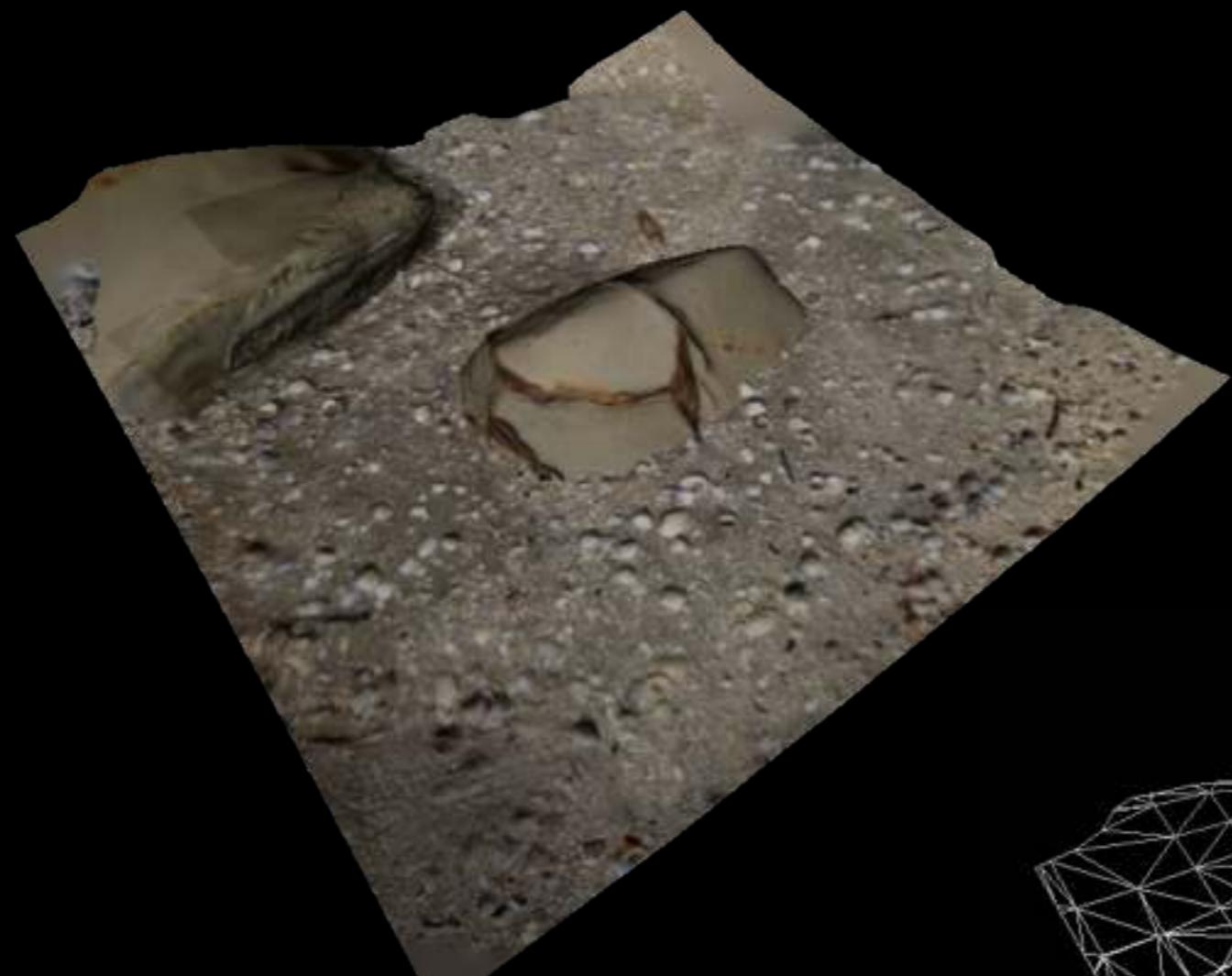
- From the University of Washington.
- An open source distribution of Bundler (MSWindows, Mac, Linux).
- Includes a GPU accelerated implementation.
- Matches images, derives camera attributes, and computes a point cloud.
- Dense point cloud and mesh generation needs to be performed elsewhere.
- <http://www.cs.washington.edu/homes/ccwu/vsfm/>
- Bundler on Mac OS X called easyBundler.
- <http://openendedgroup.com/field/ReconstructionDistribution>
- Variations of this are what we run on our HPC machines, scales well.  
Most stages run in parallel except the texture projection.
- A good place to start if interested in OpenSource pipelines.

# Software : PhotoScan

- From AgiSoft. <http://www.agisoft.ru/products/photoscan>.
- A series of individual steps (pipeline) one follows.  
Offers a batch mode for processing large collections.
- Good mixture between low level control and automation.  
Generally “just works” but can tuned for problematic cases.
- Available for Mac and MSWindows.
- Two versions
  - Standard is quite affordable
  - Pro version largely for georeferencing and other features important for the surveying community.
- Under rapid development ... regularly improving.
- Very stable.
- Fast, all parts of the pipeline seem to load balance well over cores.







Mesh density related to surface confidence.

# Software : Distinguishing features

- Degree of human guidedness and interaction required.
- Degree of control over the process, options that support fixing errors.
- Big difference between the need to reconstruct one object vs hundreds.  
My bias is towards largely automated processes.
- Requirement or opportunity for camera calibration. Becoming less necessary.  
Should result in higher accuracy, questionable for a single fixed focal lens.
- Sensitivity to the order the photographs are presented.
- The number of photographs and resolution that can be handled.
- Degree to which one needs to become an “expert”, learning the tricks to get good results.
  - There are potentially a large number of variables.
  - Trade off between simplicity and control.
  - 123D Catch is at one end of the scale, PhotoModeller Scanner at the other end.
- Ability to create high resolution textures, larger than 4Kx4K, or multiple textures.

# Photography : Lenses

- Preferred: fixed focal length lens, also referred to as a “prime lens”.
  - Depends on the software, but generally recommended as it removes one variable to be solved in the Bundler stage.
- EXIF: generally software reads EXIF data from images to determine focal length, sensor size, and in some cases lens make/model for calibration curves.
- Most “point and click” cameras have a fixed focal lenses because they require no moving parts, don’t require electronics (not drawing extra power).
- I use Canon 5D MkIII with prime lenses: 28mm, 50mm, 100mm macro.



Sigma 28mm, f1.8



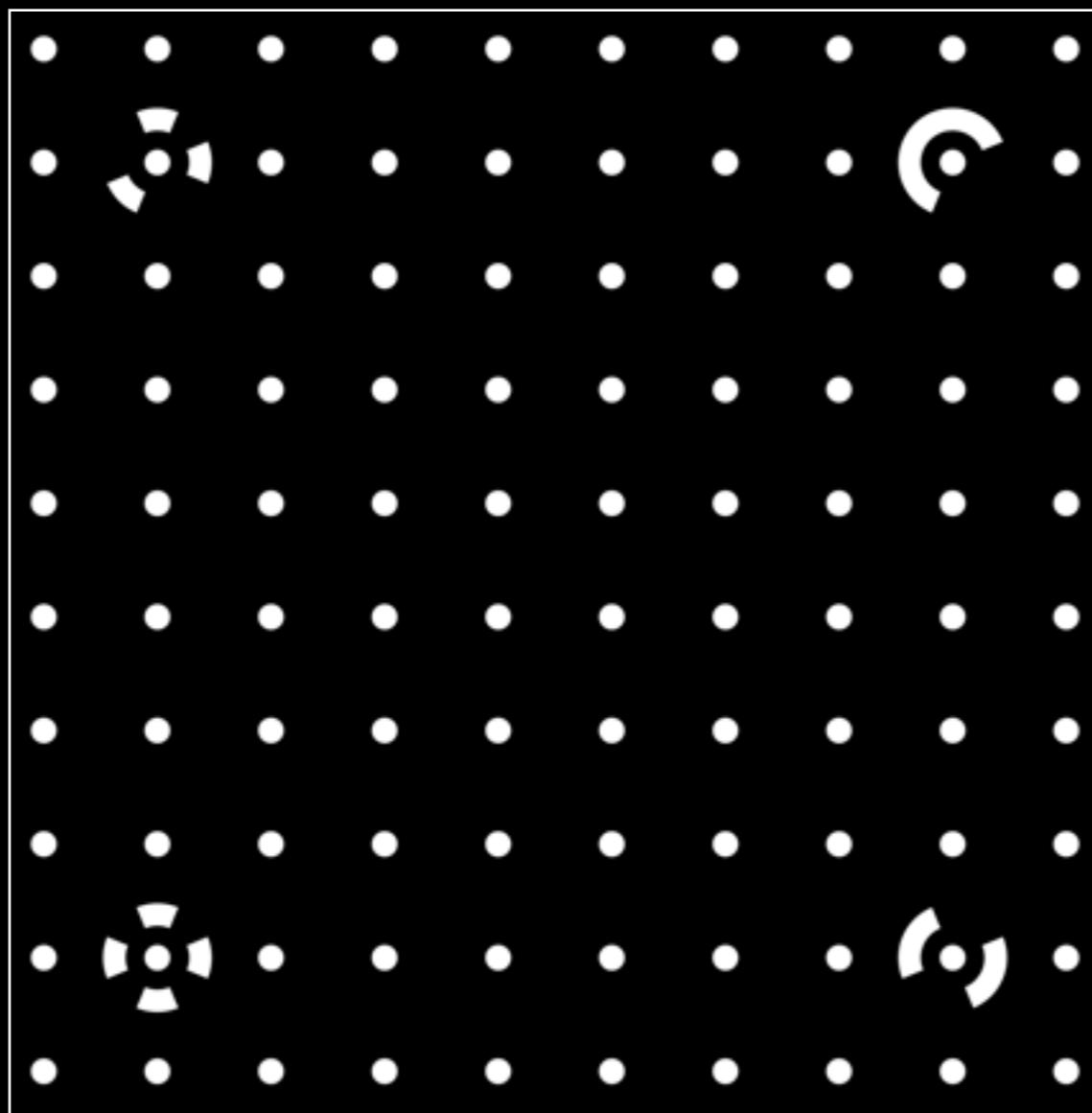
Sigma 50mm, f1.4

# Photography : shooting guide

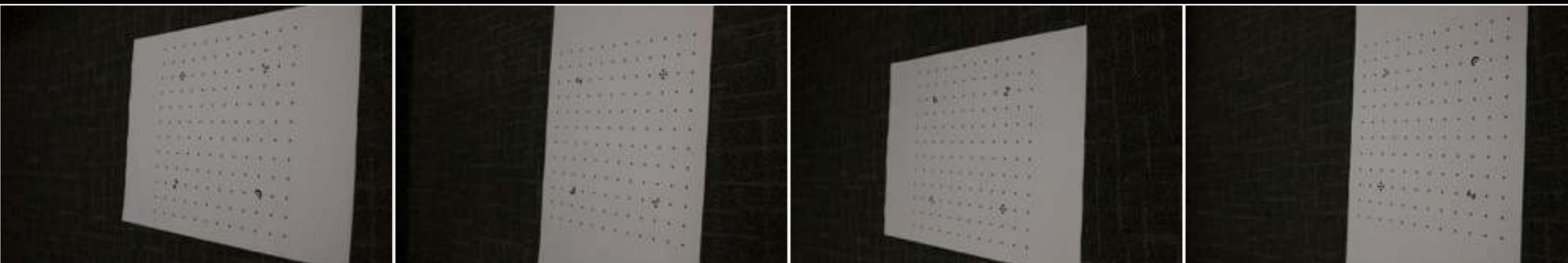
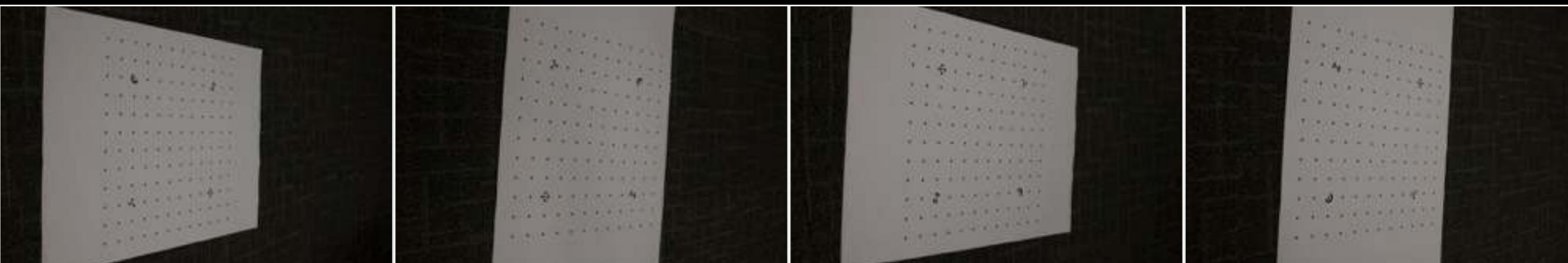
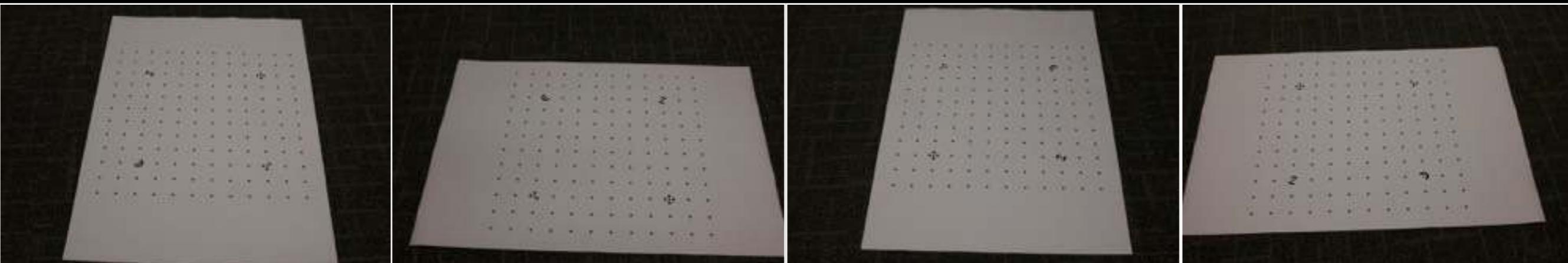
- Obviously one cannot reconstruct what one does not capture.
- Aim for plenty of overlap between photographs (Can always remove images).
- For 2.5D surfaces as few as 2 shots are required, more generally 6 - 20.
- For 3D objects typically 20 or more. ~ 10 degree steps.  
Repeat at one or more levels if the object is concave vertically.
- For extended objects and overlapping photographs perhaps hundreds.  
1/3 to 1/2 image overlap ideal.
- Historically worked better for the images to be captured in order moving around the object, this is being relaxed in the latest algorithms.
- Generally no point capturing multiple images from the same position!  
The opposite of panoramic photography for example.
- Camera orientation typically doesn't matter, this is solved for when computing camera parameters in the Bundle processing.
- Lens calibration is becoming less important as standard lens have online published curves. Still necessary for accurate results from unusual lenses.

# Photography : Camera calibration

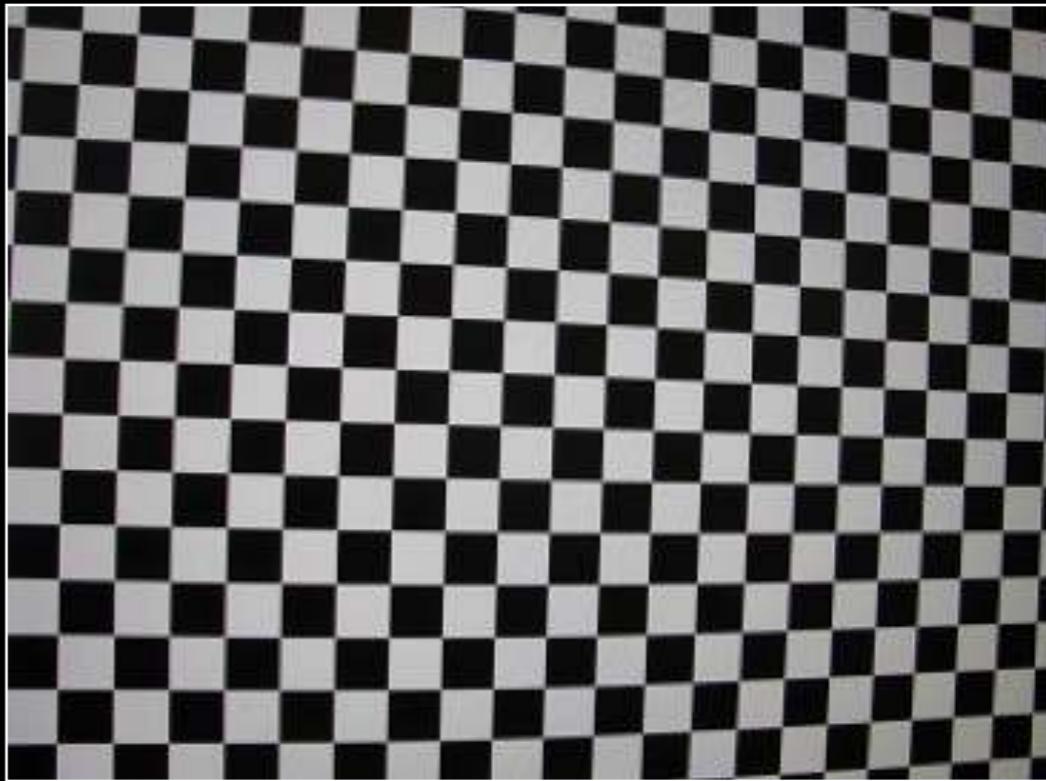
- Camera/lens characteristics derived photographs of a known pattern.
- Different procedures depending on the software.
- Calibration pattern used by PhotoModeller shown here: printed A1 sheet.



- 4 photographs captured (one from each direction).
- Repeated with the camera in three orientations (rotated 90, 0, -90).



- PhotoScan provides a separate utility called “lens”.
- Estimates
  - focal length in both directions
  - principle point components in both directions
  - radial and tangential distortion coefficients
- $f_x, f_y, c_x, c_y, K_1, K_2, K_3, P_1, P_2$
- Produces a display on screen to photograph from different directions.
- Generally doesn't solve for focal length, reads from EXIF.



EXIF focal length: 50  
 $f_x = 8026.46 \pm 1.5152$   
 $f_y = 8027.75 \pm 1.42957$   
 $c_x = 2877.05 \pm 1.13418$   
 $c_y = 1906.64 \pm 0.814478$   
 $\text{skew} = -0.806401 \pm 0.151285$   
 $k_1 = -0.176187 \pm 0.00377854$   
 $k_2 = 0.285354 \pm 0.0770751$   
 $k_3 = 0.300547 \pm 0.619451$   
 $p_1 = 0.000219219 \pm 2.64764e-05$   
 $p_2 = -0.000172641 \pm 3.58682e-05$

## Untitled

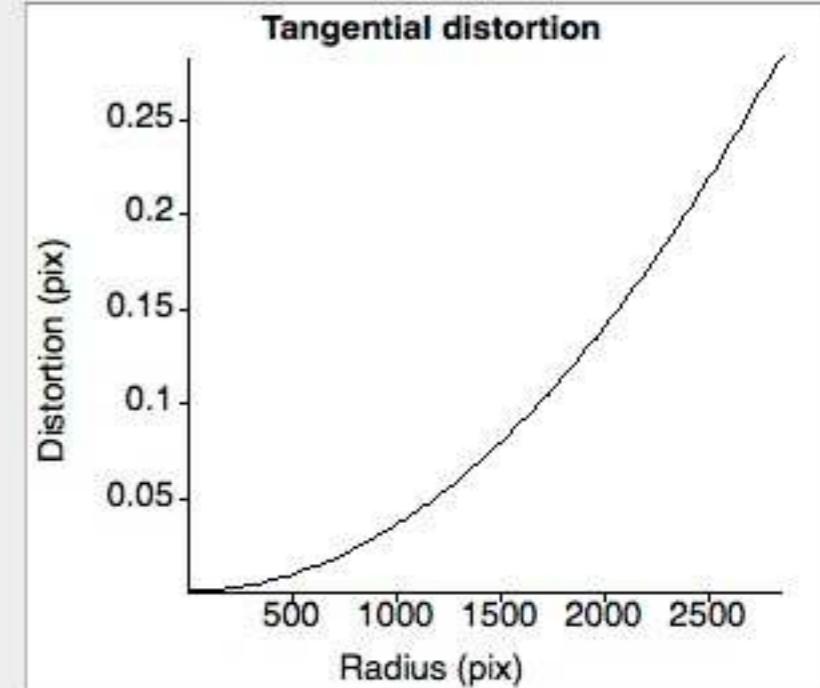
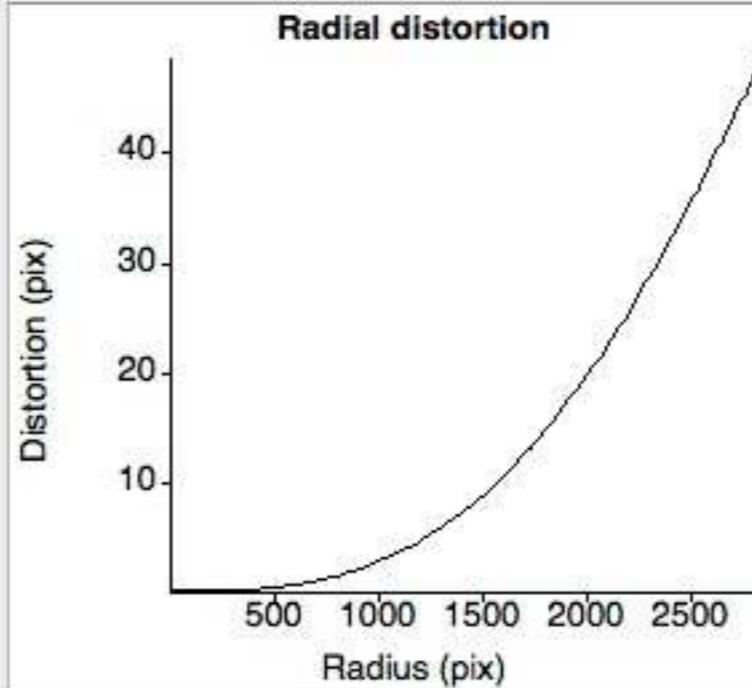
Photos Report

X

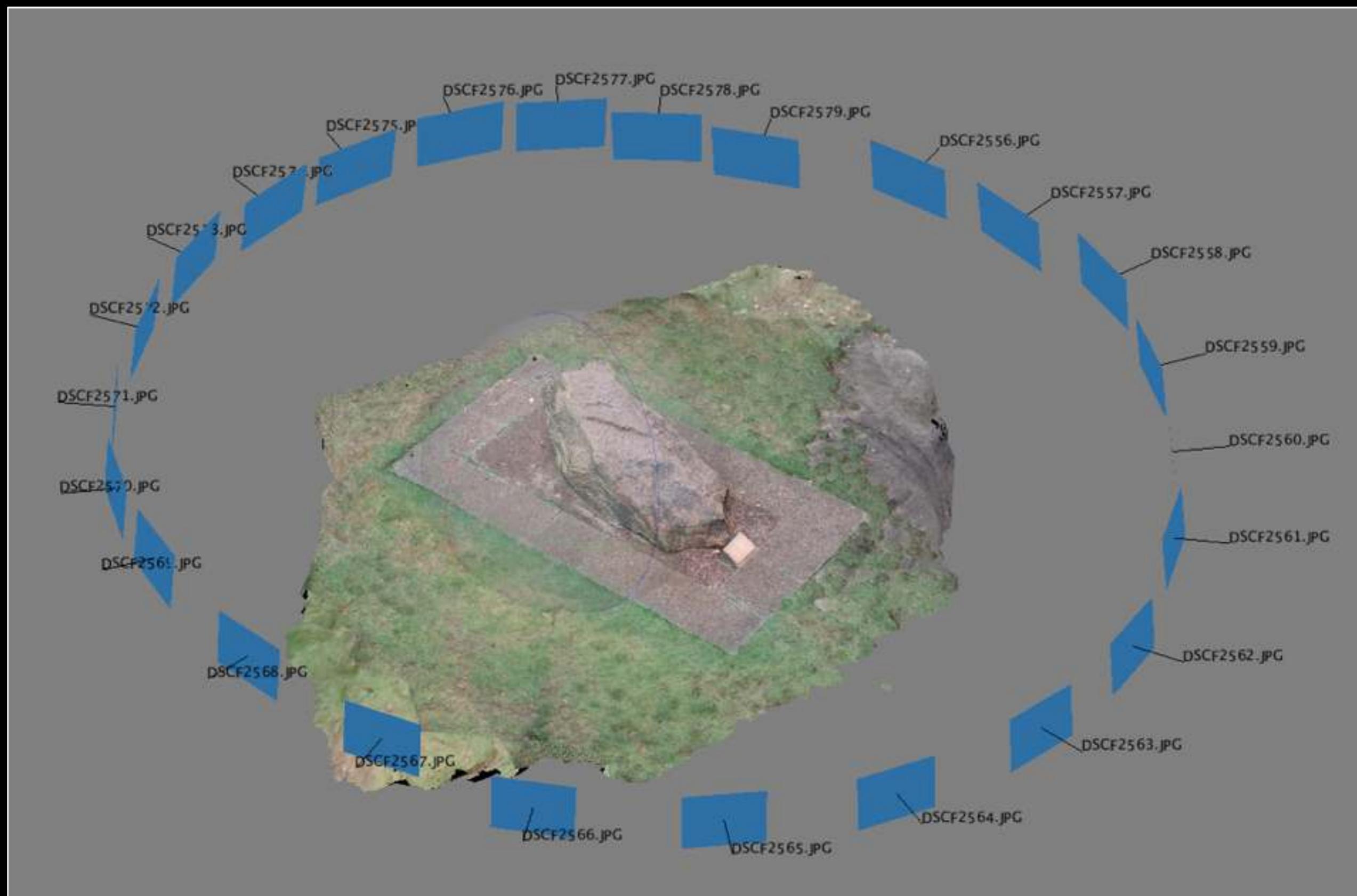
File	Size	Points
▼ f 50		
IVEC1859.JPG	5760x3840	1434
IVEC1860.JPG	5760x3840	1320
IVEC1861.JPG	5760x3840	1208
IVEC1862.JPG	5760x3840	986
IVEC1863.JPG	5760x3840	978
IVEC1864.JPG	5760x3840	1191
IVEC1865.JPG	5760x3840	953

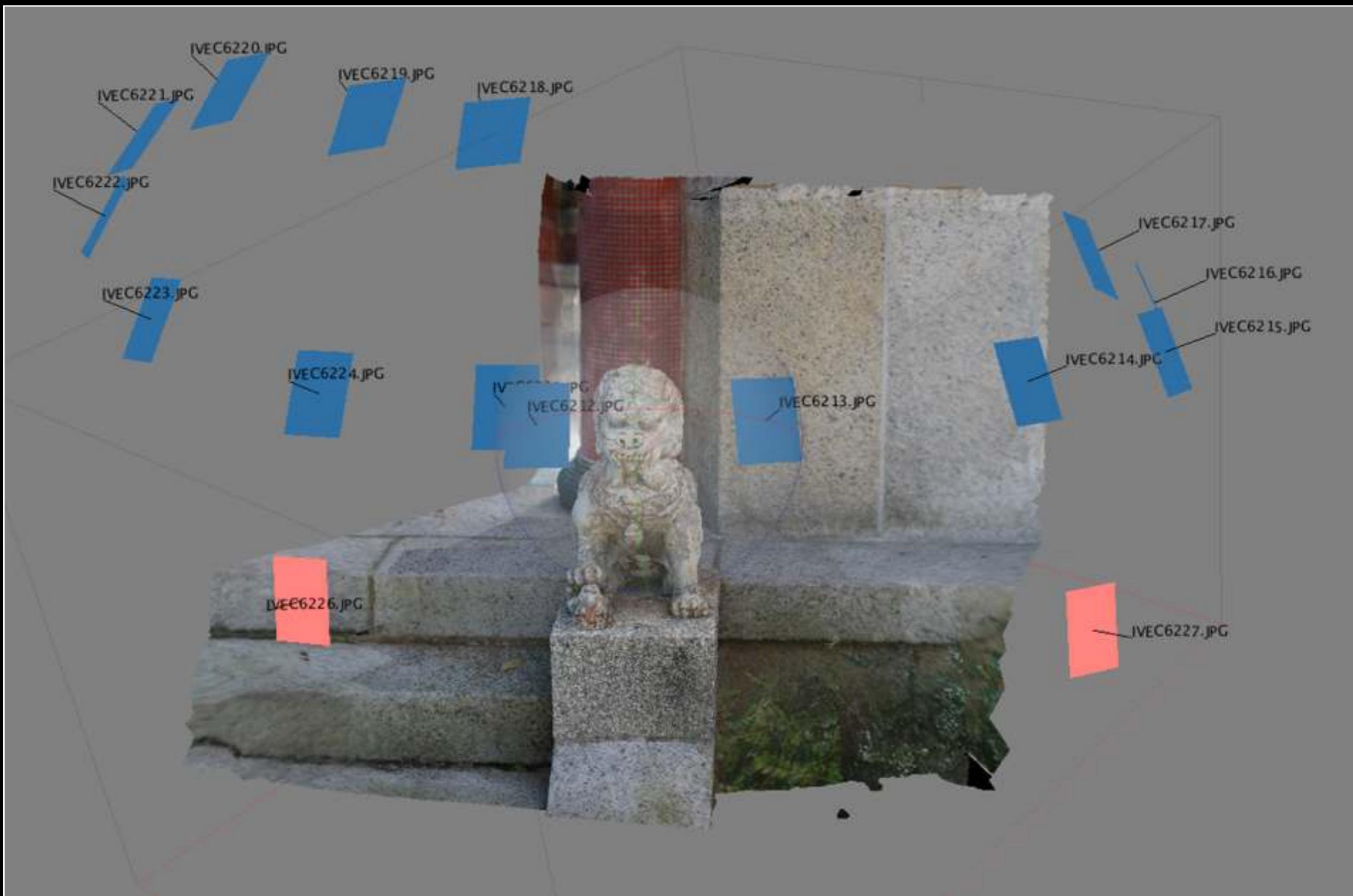
EXIF focal length: 50

Parameter	Value	Std Error
Image width	5760	
Image height	3840	
Focal length (x)	8026.46	1.5152
Focal length (y)	8027.75	1.42957
Principal point (x)	2877.05	1.13418
Principal point (y)	1906.64	0.814478
Skew	-0.806401	0.151285
Radial K1	-0.176187	0.00377854
Radial K2	0.285354	0.0770751
Radial K3	0.300547	0.619451
Radial K4	-1.09108	2.89591
Tangential P1	0.000219219	2.64764e-05
Tangential P2	-0.000172641	3.58682e-05



# Photography : shooting guide

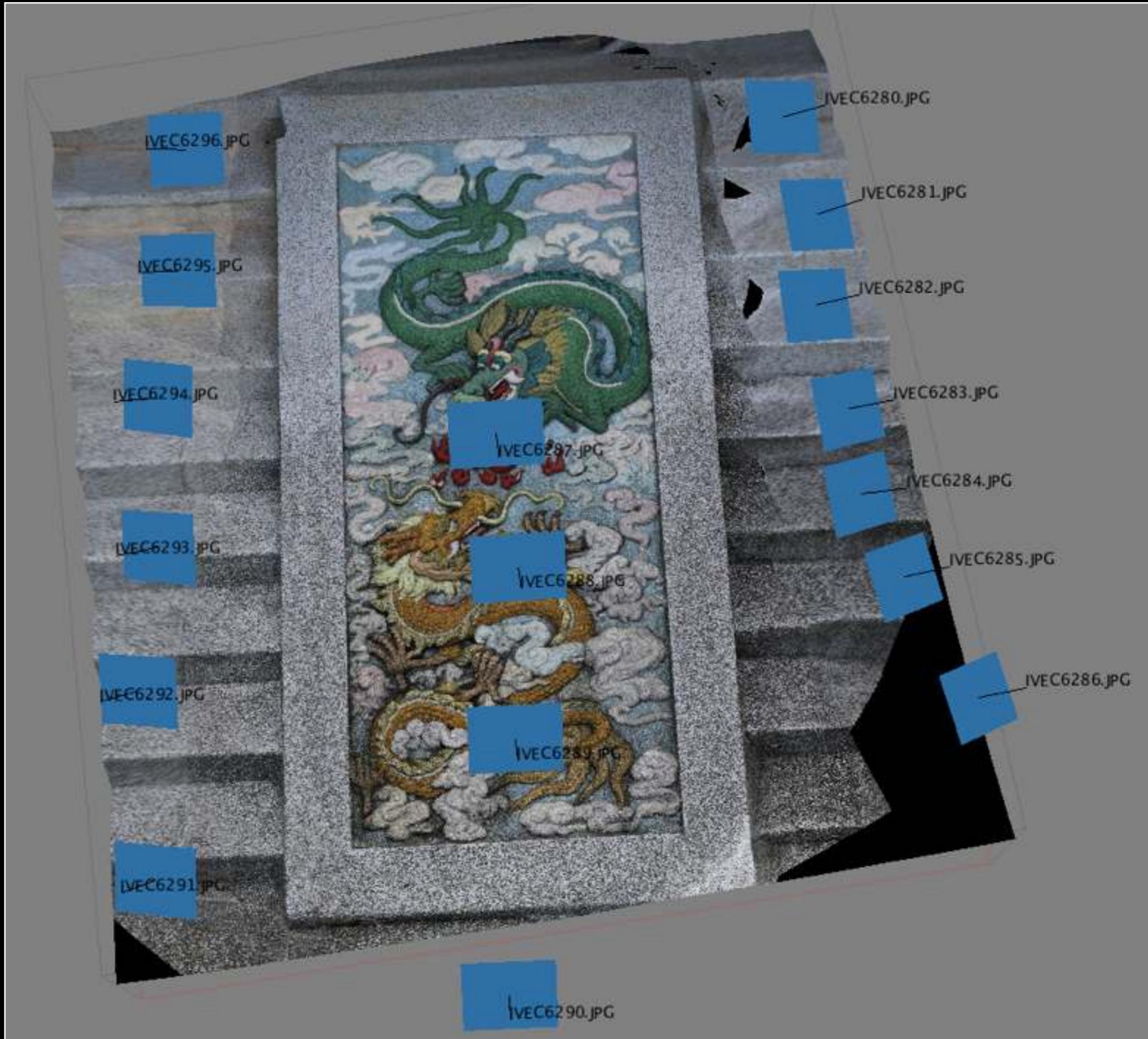




Dragon Gardens, Hong Kong

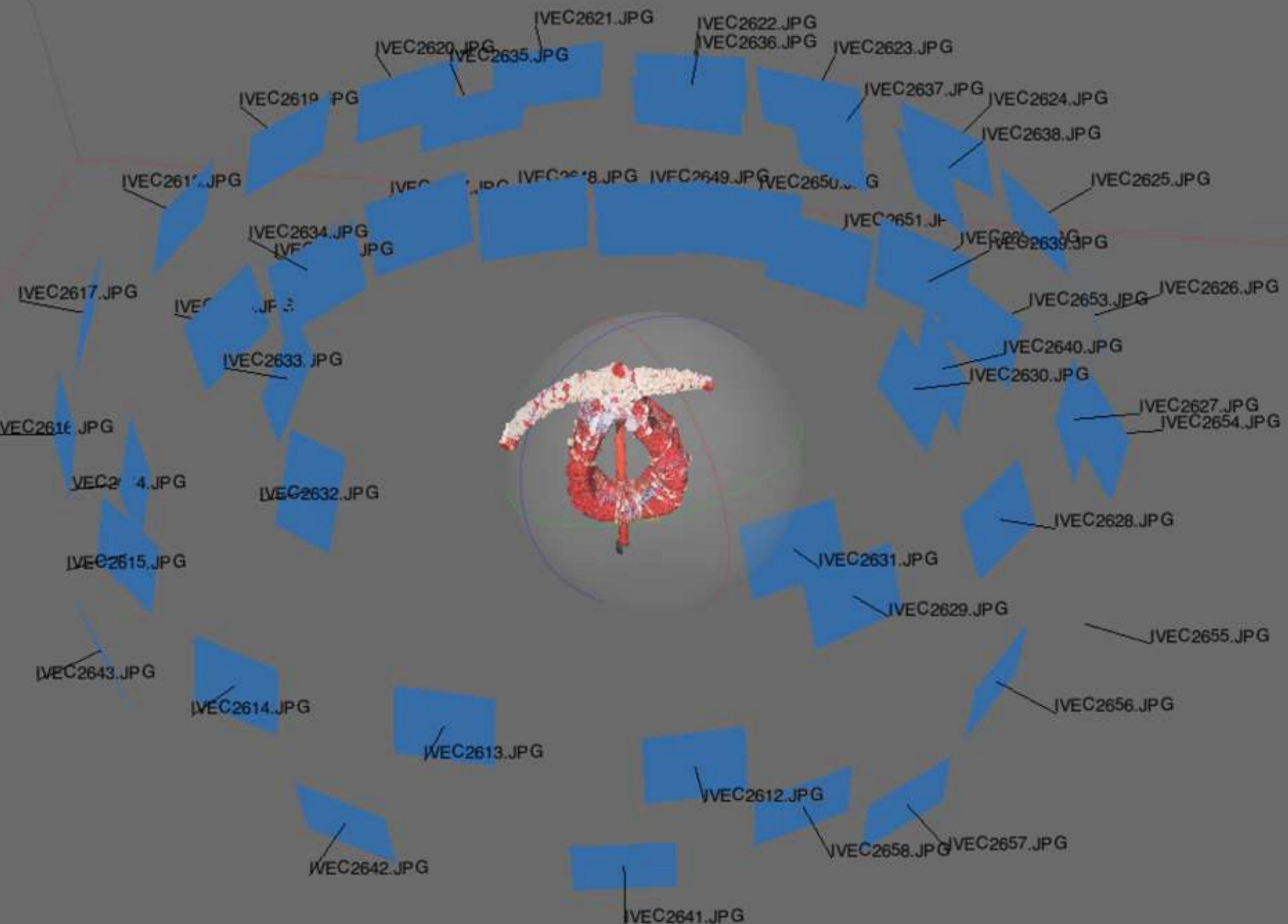
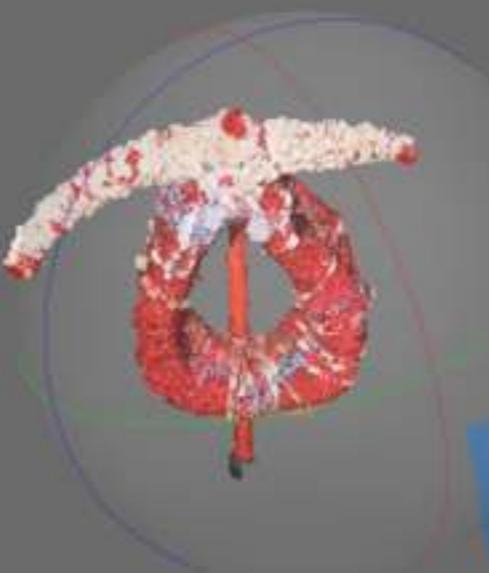


Manipal, India

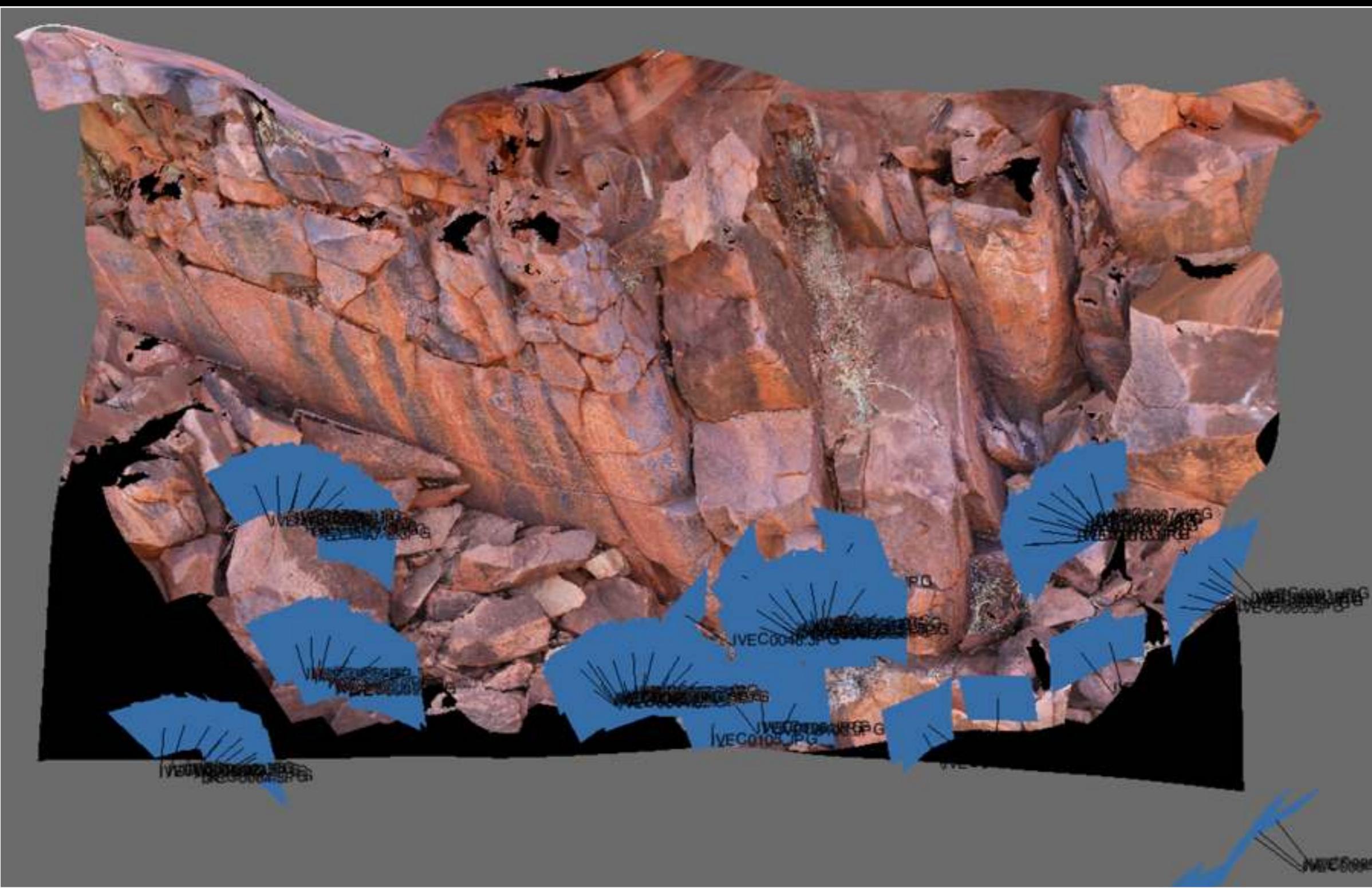


Dragon Gardens, Hong Kong





Ngintaka headress



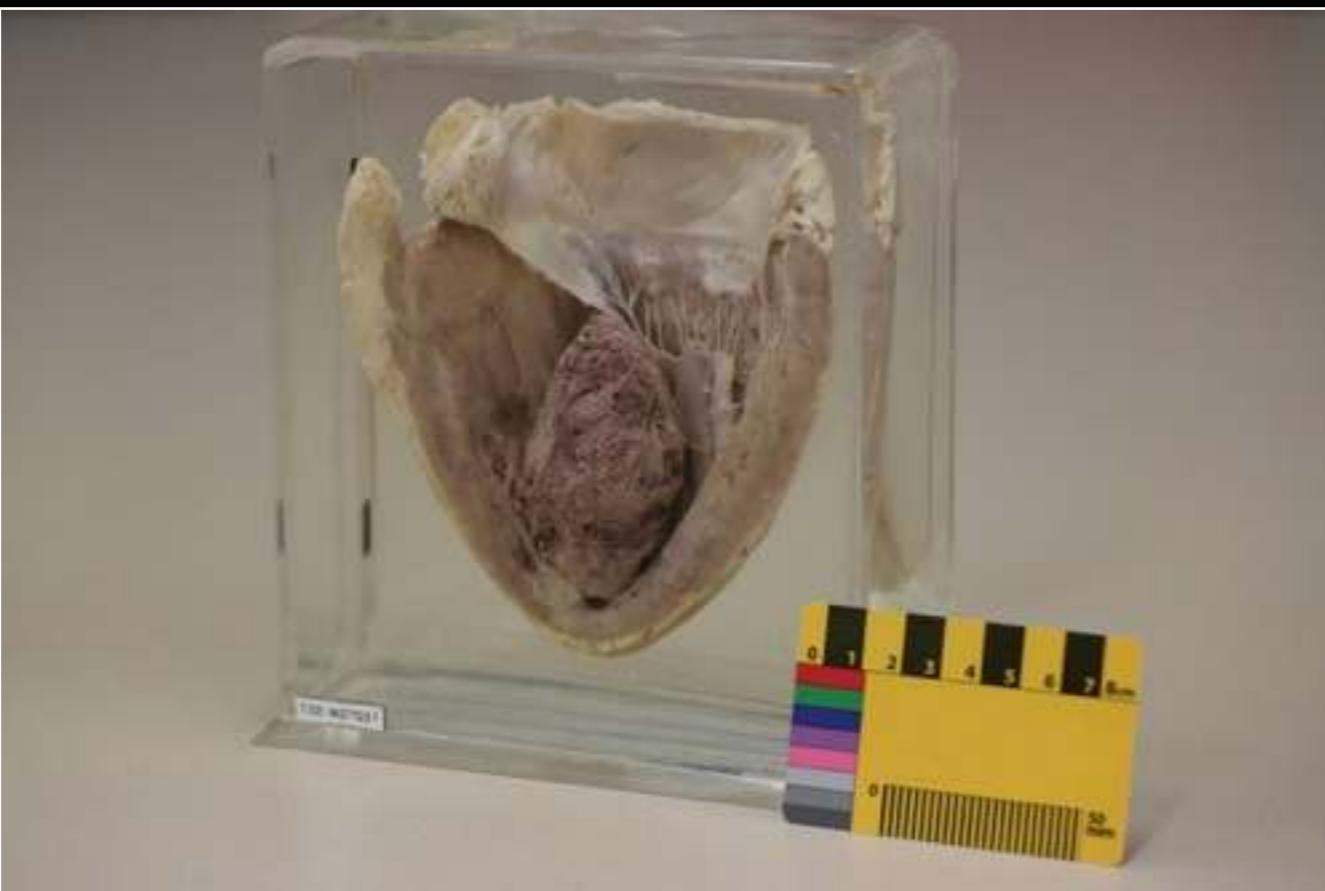
Manmanna



West Angeles rock shelter

# Photography : Linear reference objects

- Improved results if there is a linear reference object in the scene.
- They need not be part of the final reconstruction if slightly outside the object of interest.
- Reference colour bars also useful if colour correction is important.



QE II Pathology



Long Island

# Photography : Photograph quality

- Clearly improved results with crisp, high resolution images.
- Improvement with increased photograph count, see later in accuracy discussion.
- Easy area to explore, capture a photograph set and progressively subsample or blur images.



Dragon Gardens, Hong Kong

## Comparing dense point clouds



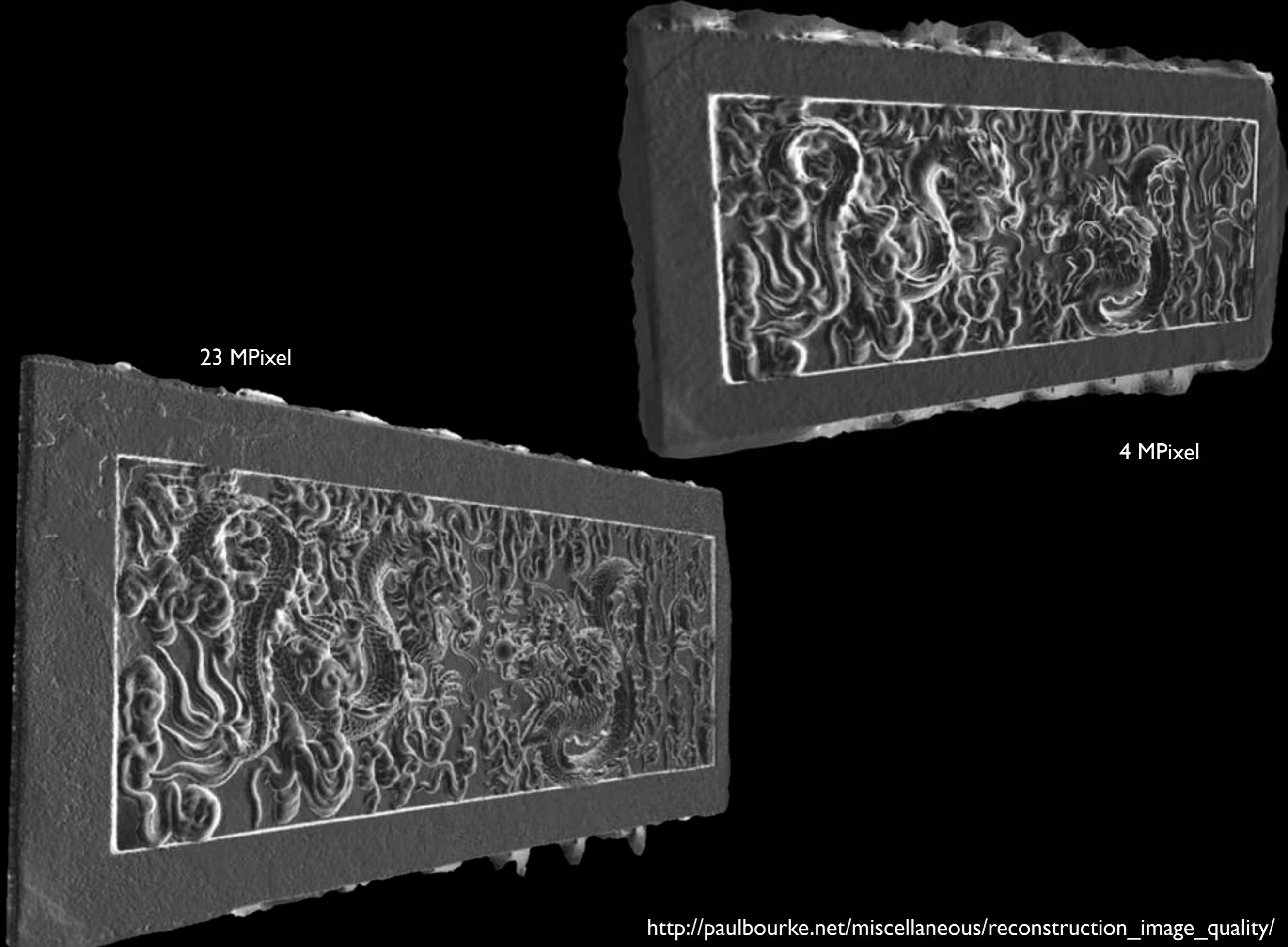
4 MPixel



23 MPixel



Dense point cloud for greyscale photographs



# Example 1 : Motifs, Indian Temple

- A relatively low number of photographs are required for 2.5D surfaces.
- Degree of concavity determines the number of photographs required.
- Facades and engravings (low concavity) can require as few as 3 to 6 images.
- 20cm high engraving on doors, 200+ engravings to capture.
- Photographs can be orientated at any angle.
- Each object takes perhaps 15 sec to capture, 10 minutes (on average) to process.
- Able to process in the field and redo any that failed.
- This example uses a simple point and shoot camera.



Chaturmukha, India





Chaturmukha, India

# Photography : Coverage

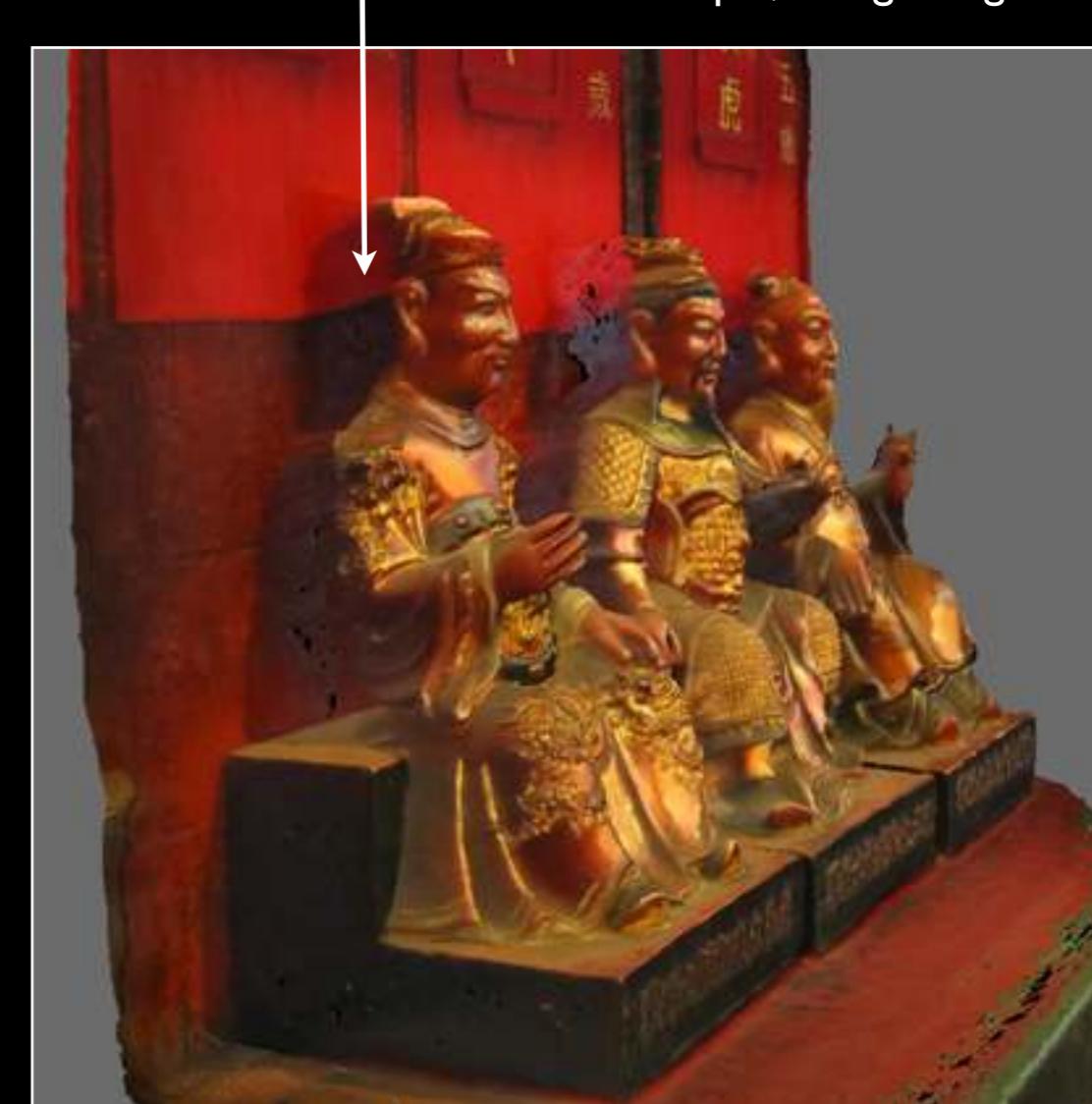
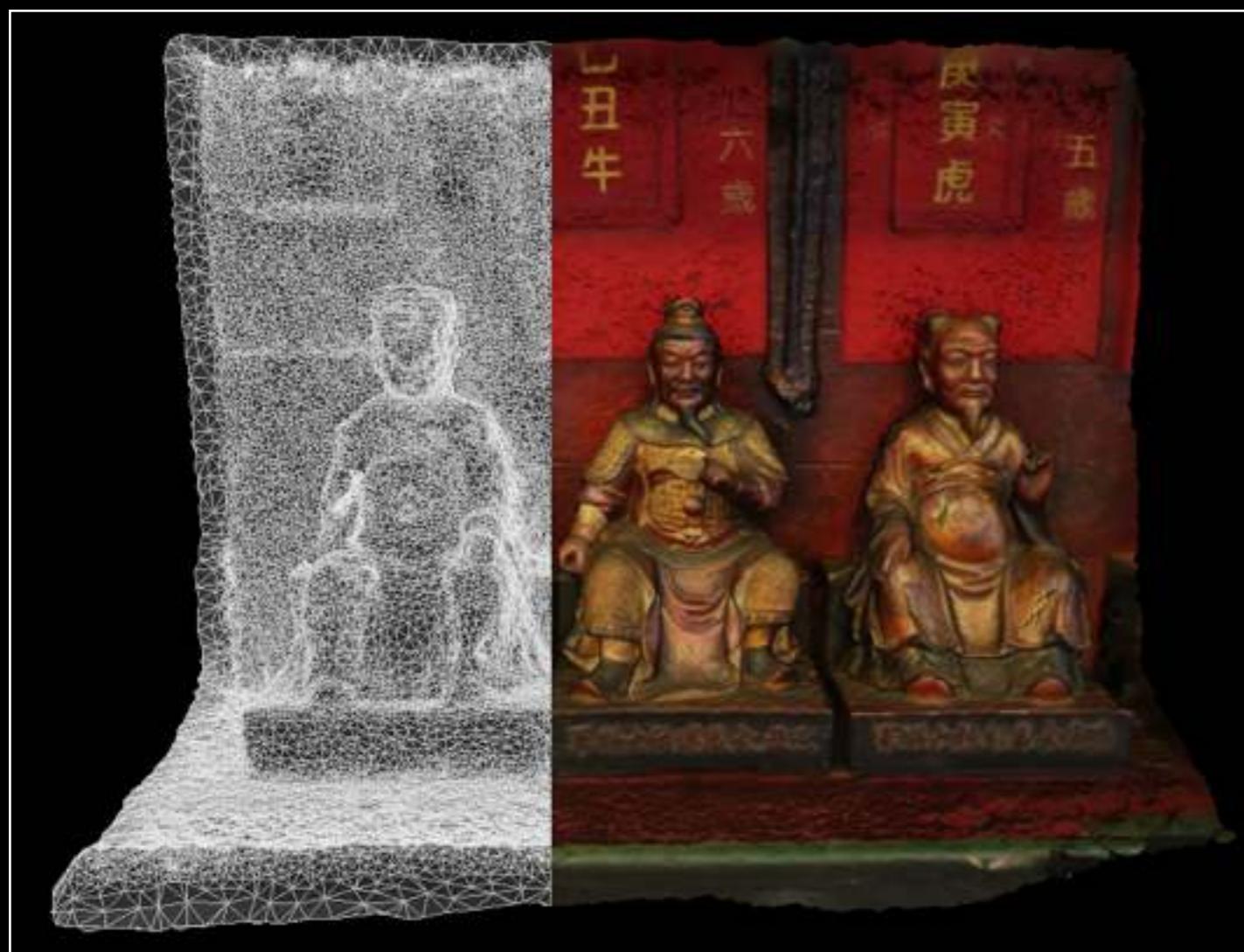
- Pretty obvious, one is never going to reconstruct what isn't captured.



~30 photographs

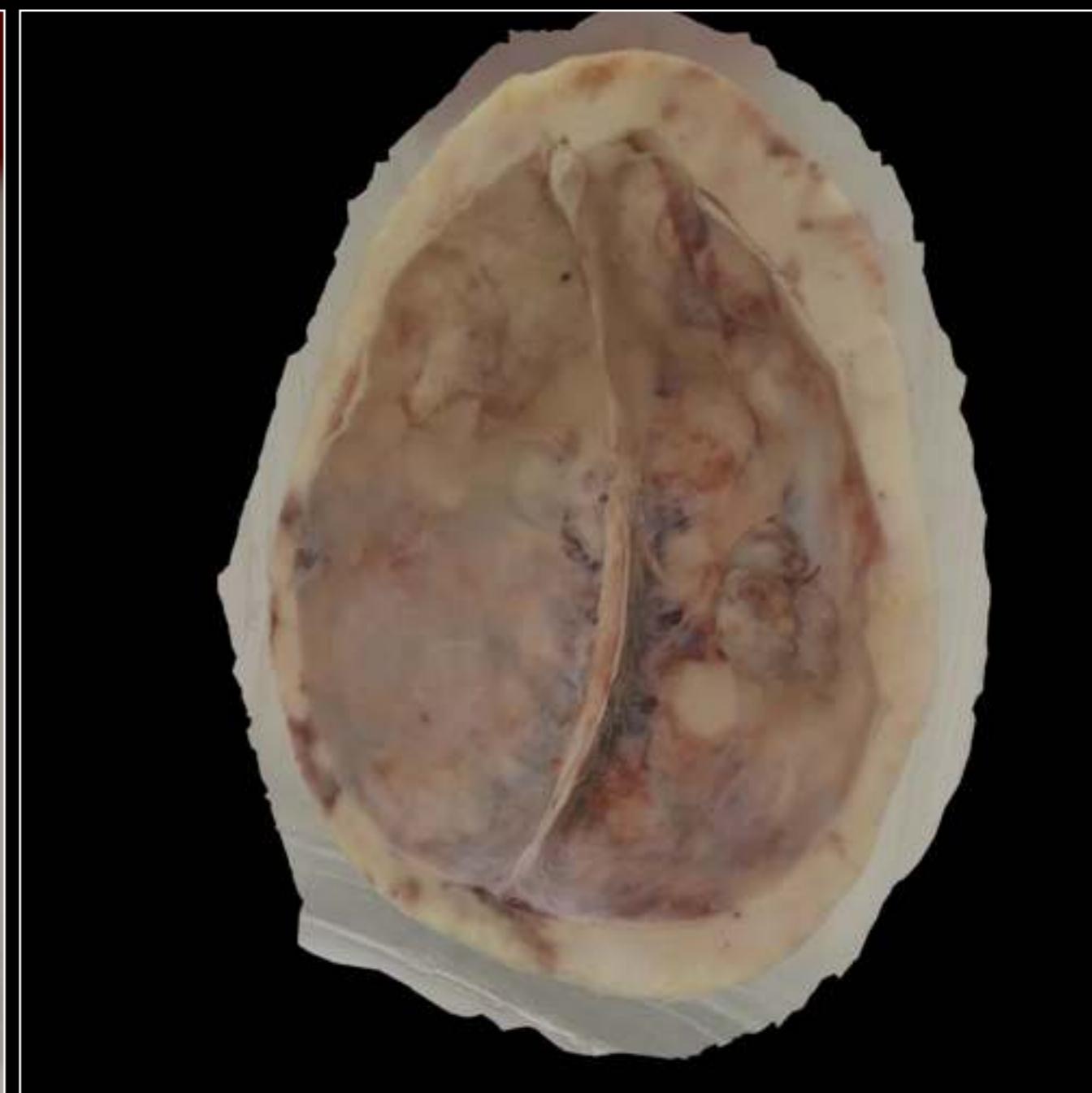
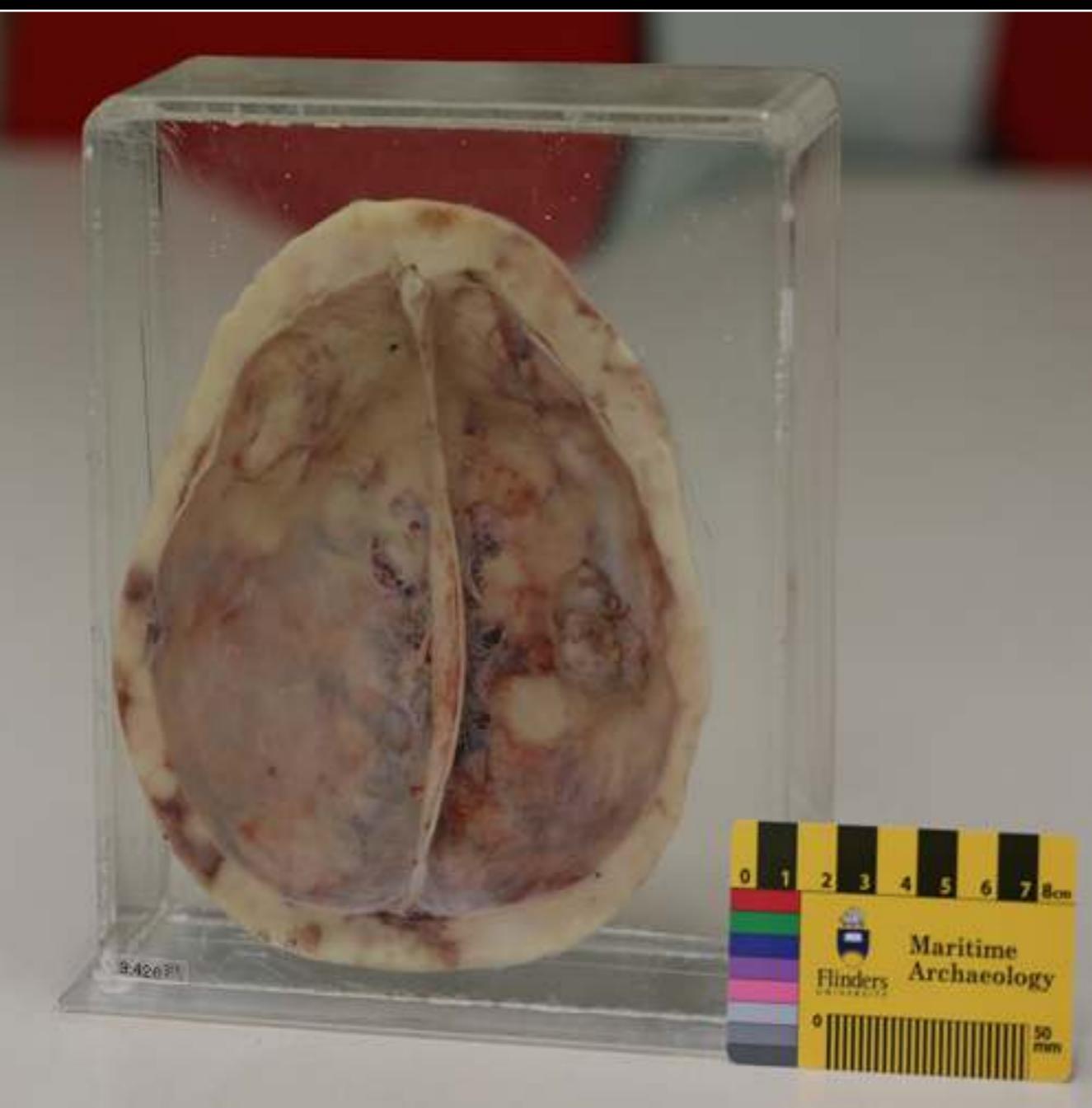
Limited by left most photograph

Tung Wah hospital  
temple, Hong Kong



# Photograph : depth of focus

- Feature point detection benefits from sharp / crisp photographs.
- Shallow depth of focus can be used to ensure feature points are only found at the depth of interest.

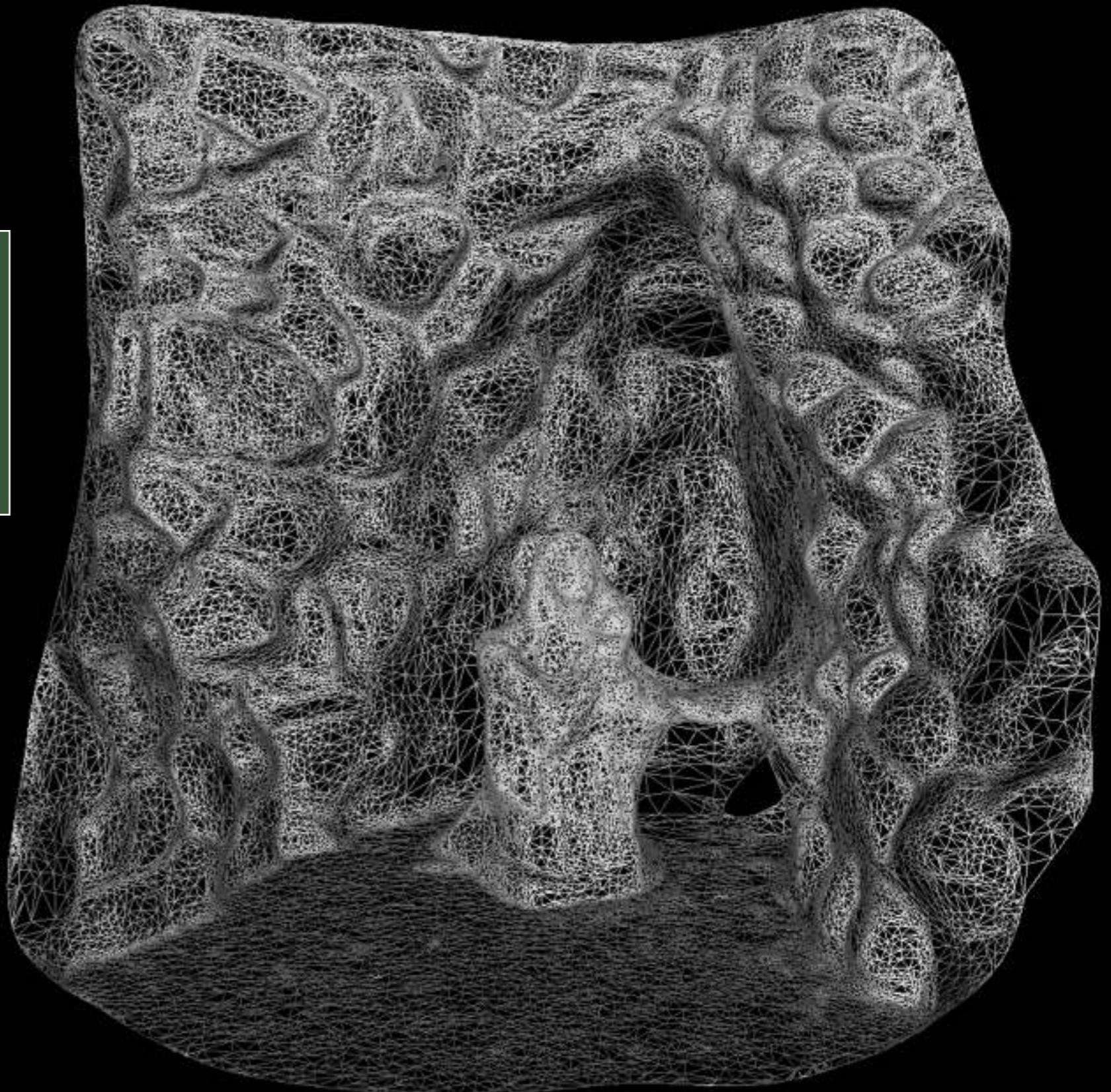


# Geometry post processing

- Generally dealing with unstructured meshes

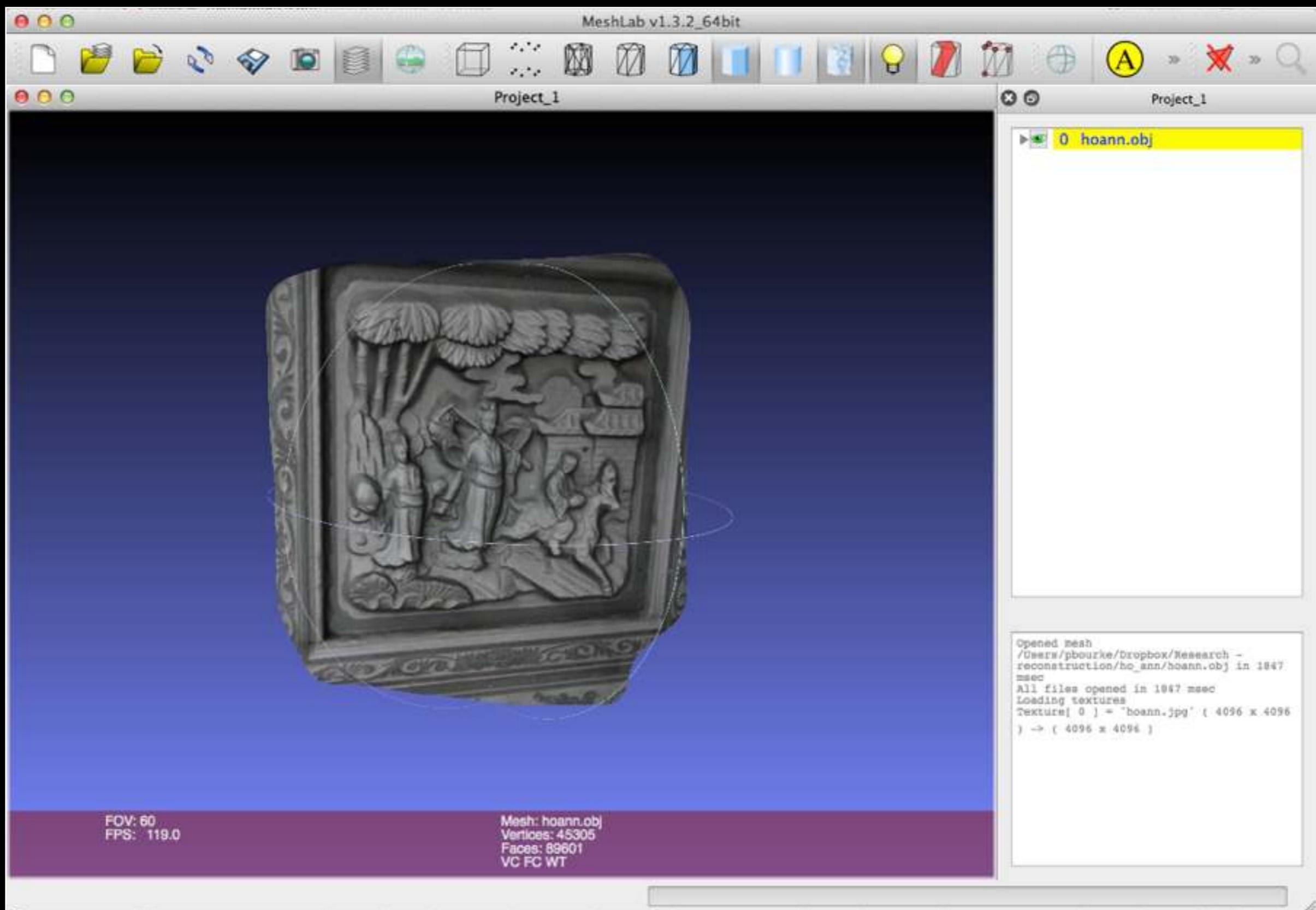
- Mesh simplification
- Hole closing
- Removing shrapnel

- Per vertex editing
- Mesh thickening
- Meshlab
- Blender
- File formats



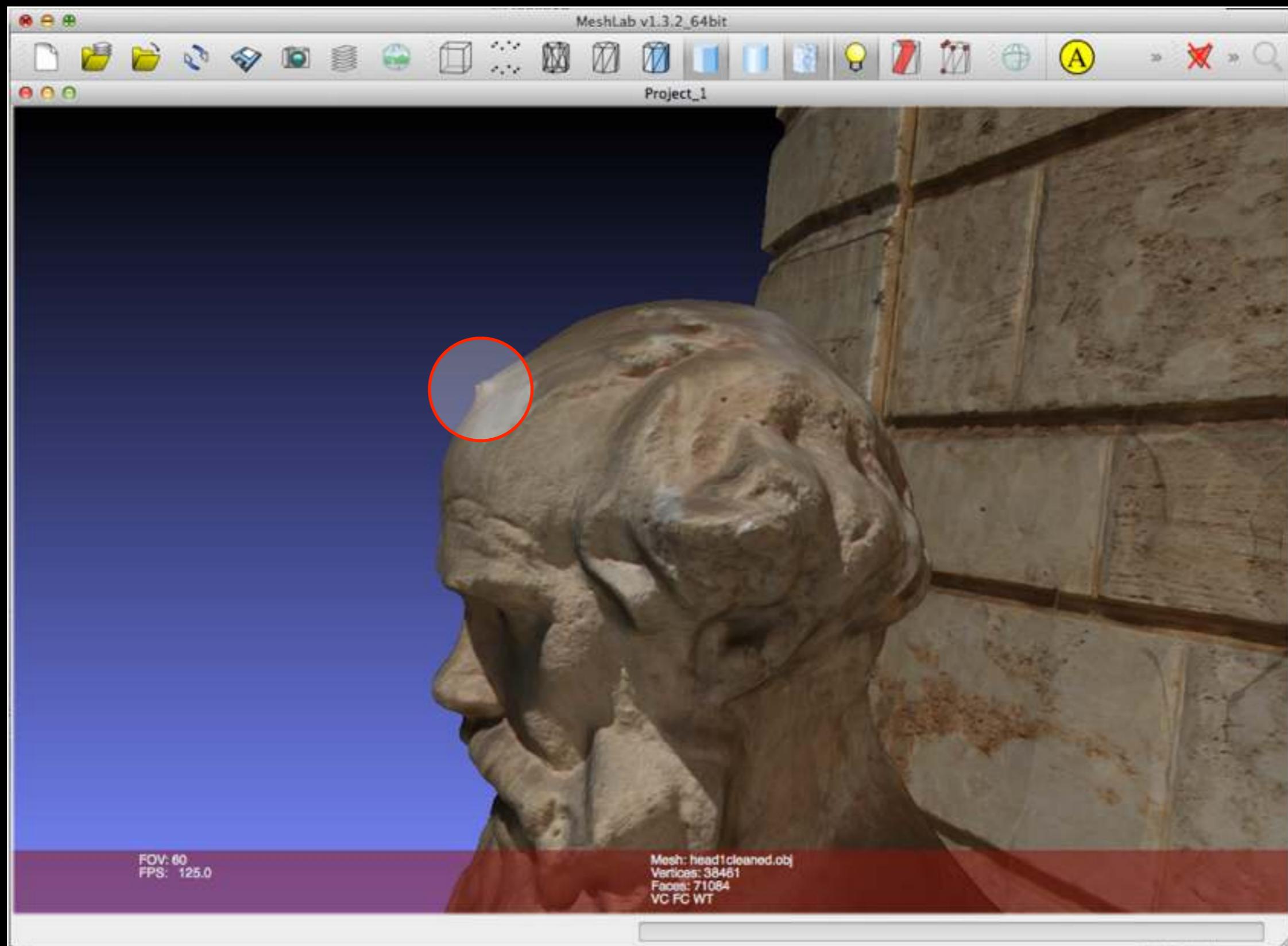
# Geometry processing : MeshLab

- There are a number of packages that can be used to manipulate the resulting textured mesh files.
- Meshlab is the free package of choice.
- It is cross platform with a high degree of compatibility.
- Very general tool for dealing with textured meshes.
- Has a large collection of algorithms and is extensible.
- Unfortunately not all algorithms are “reliable”.
- In cases where raw Bundler is used to create a point cloud, Meshlab can be used to construct the mesh using one of a number of algorithms.
  - Ball pivot (my general choice)
  - Marching Cubes
  - Poisson surface reconstruction

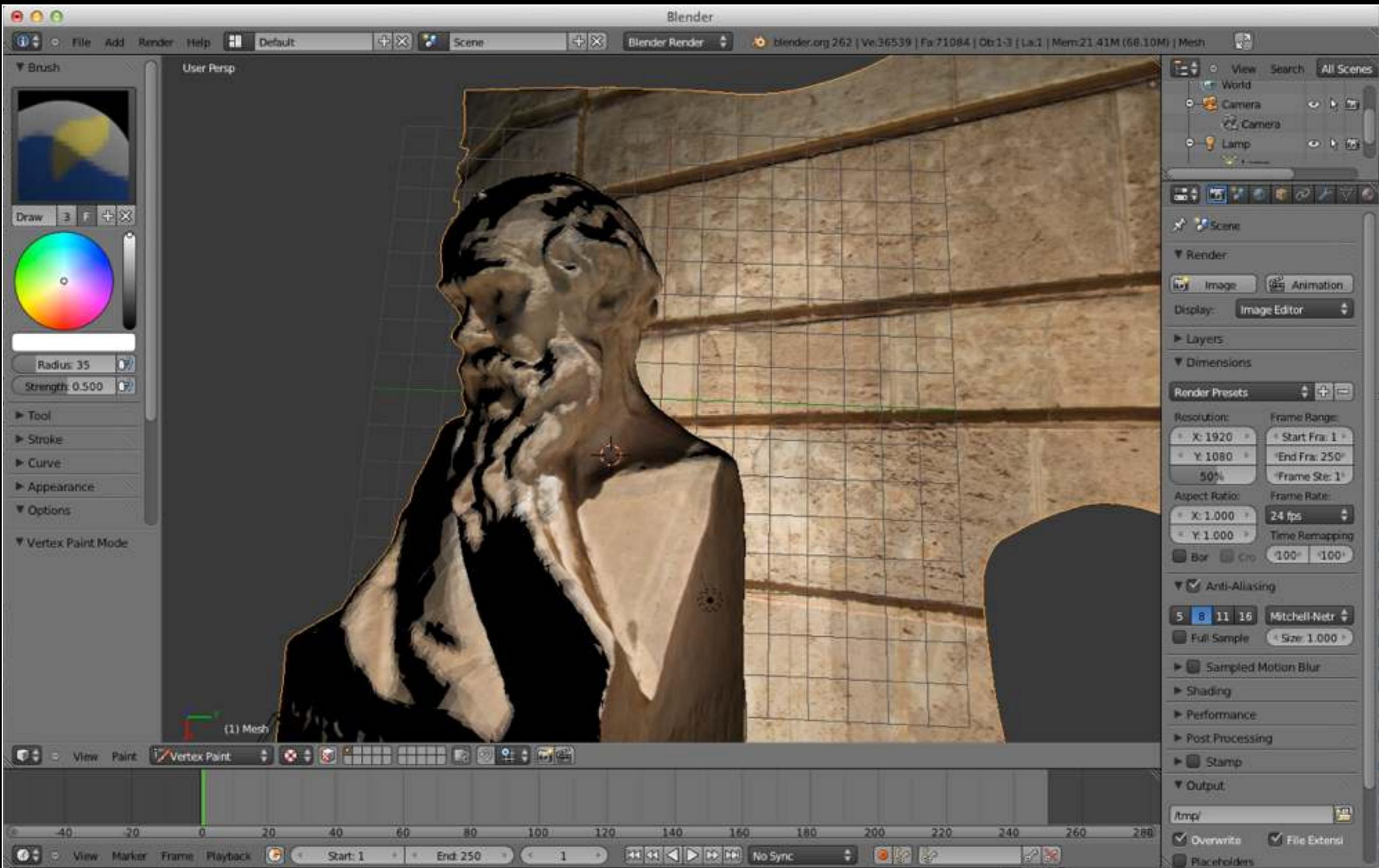


# Geometry processing : Blender

- Largely used for per vertex editing.
- “Big hammer to crack a small nut”, takes some time to learn the interface.
- For example, not uncommon to get single vertex “spikes” especially of highly specular grains.
- Contains it’s own mesh simplification and thickening algorithms.
- Also used to export in a myriad of additional formats.  
For example fbx for Unity3D, not available in MeshLab.



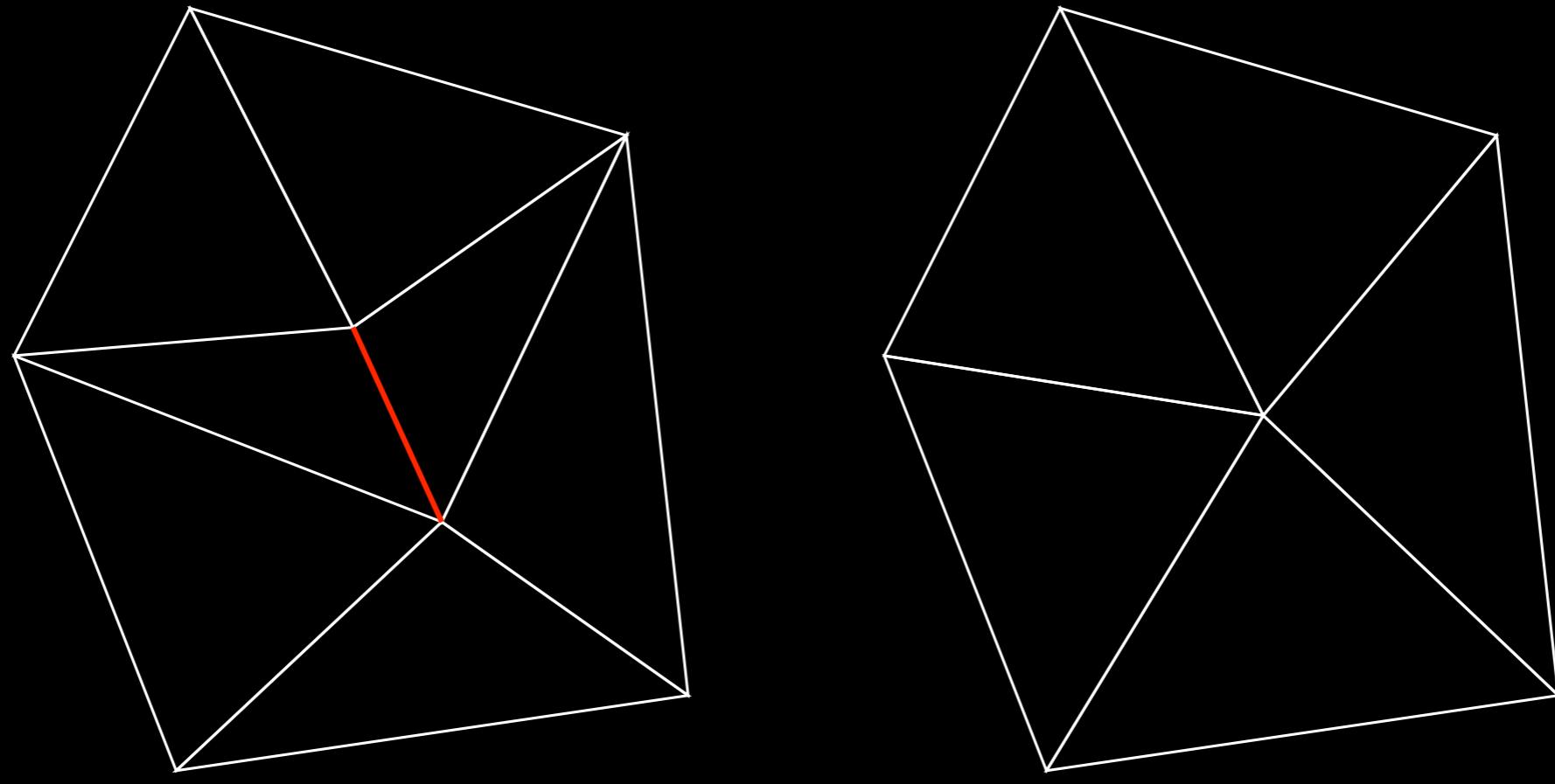
Socrates, UWA



# Geometry processing : Mesh simplification

- Meshes directly from the reconstruction (generated from the dense point cloud) are generally inefficient. Often need to reduce them for realtime applications and/or web based delivery.
- Also used to create multiple levels of details (LOD) for gaming and other realtime applications.
- The goal is easy to understand: remove mesh density where it will make minimal impact on the mesh appearance. For example, don't need high mesh density in regions of low curvature.
- Most common class of algorithm is referred to as “edge collapse”, replace an edge with a vertex.
- A texture and geometry approximation ... need to estimate new texture coordinate at new vertices.
- Need to preserve the boundary.
- This has been a common topic in computer graphics research and is still an active topic in computer graphics.

- Most edge collapse algorithms involve replacing an edge with a vertex
  - How to choose the edges to remove is the “trick”.
  - Where to locate the new vertex so as to minimise the effect on the surface.
  - How to estimate the new texture coordinate.
- Number of triangles reduced by 2 on each iteration.
- Can calculate the deviation of the surface for any particular edge collapse.  
Choose edges that result in the smallest deviation.  
Remove edges on flat regions, retain edges in regions of high curvature.



Red edge removed, results in two fewer triangles



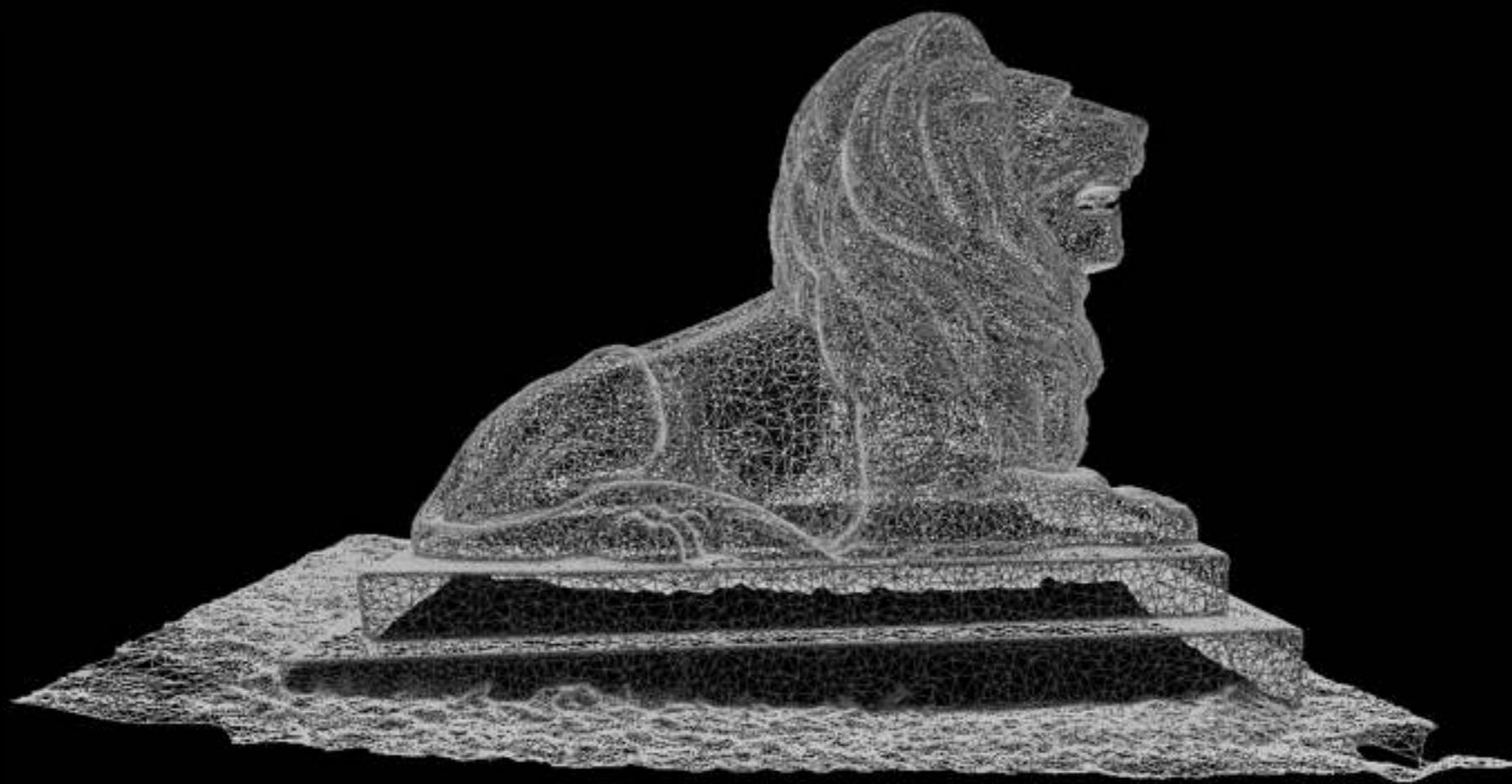
1,000,000 triangles



100,000 triangles



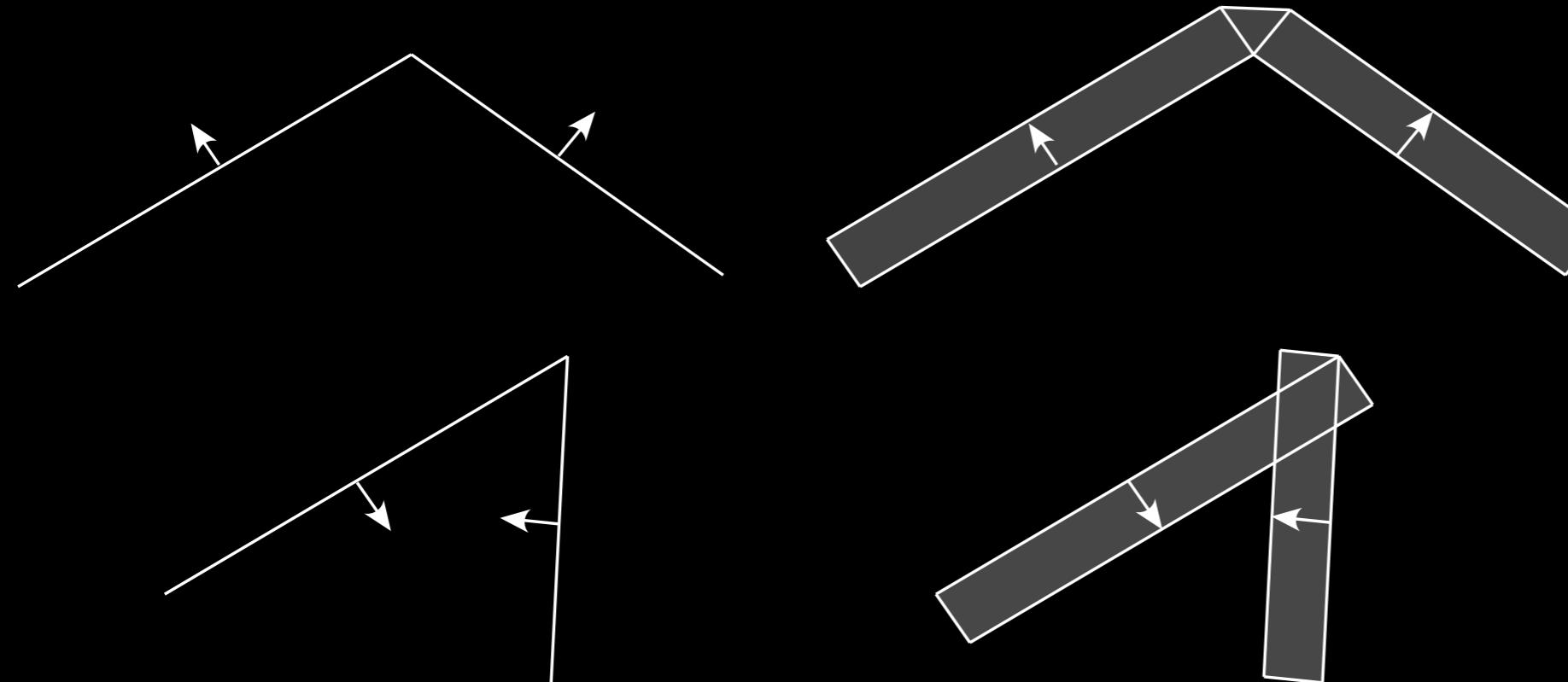
1,000,000 triangles



100,000 triangles

# Geometry processing : Mesh thickening

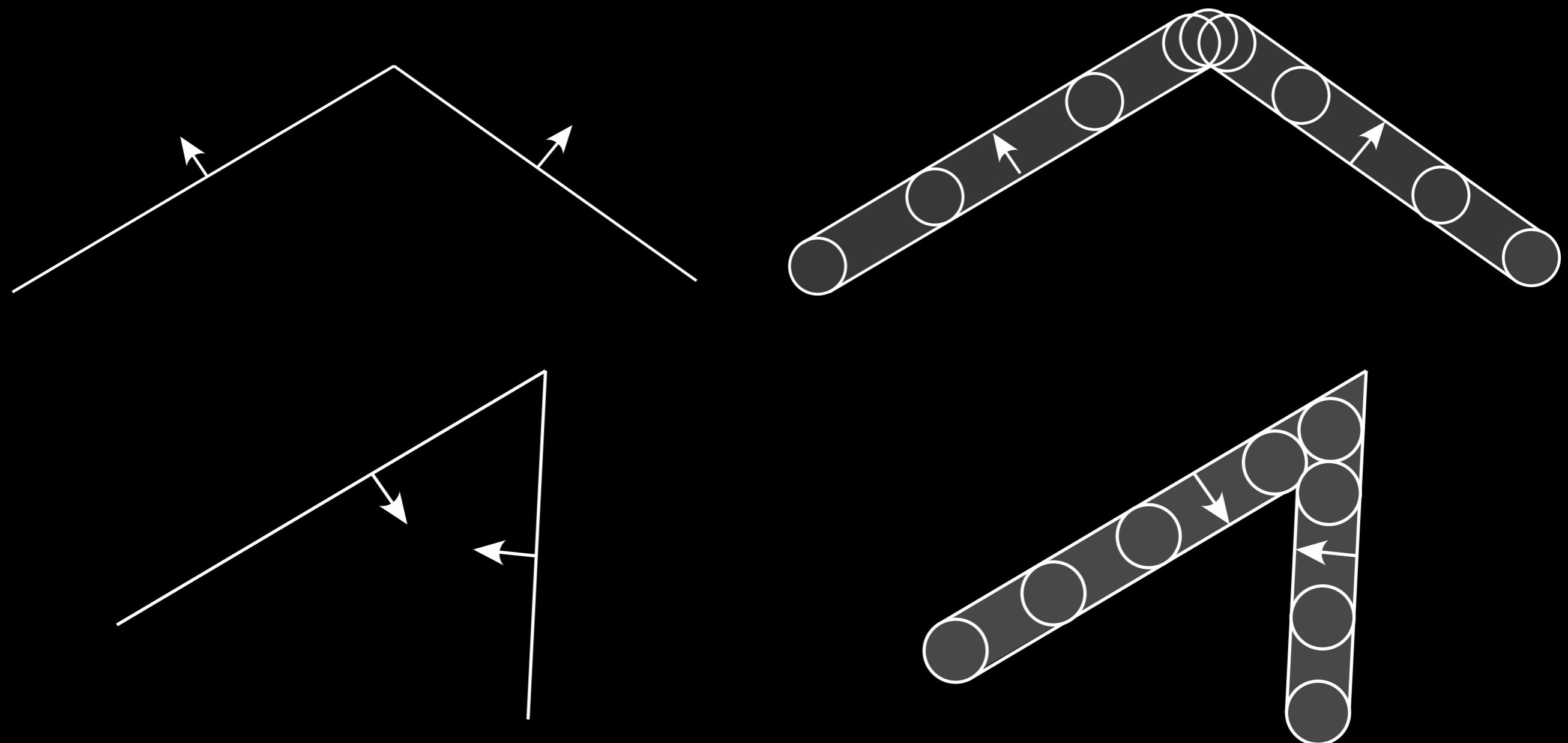
- Cases exist where one does not want idealised “infinitely thin” surfaces.
- Double sided rendering in realtime APIs is not quite the same visual effect as physical thickness.
- Required to create physical models, see rapid prototyping later.
- Perhaps the most common algorithm is known as “rolling ball”.

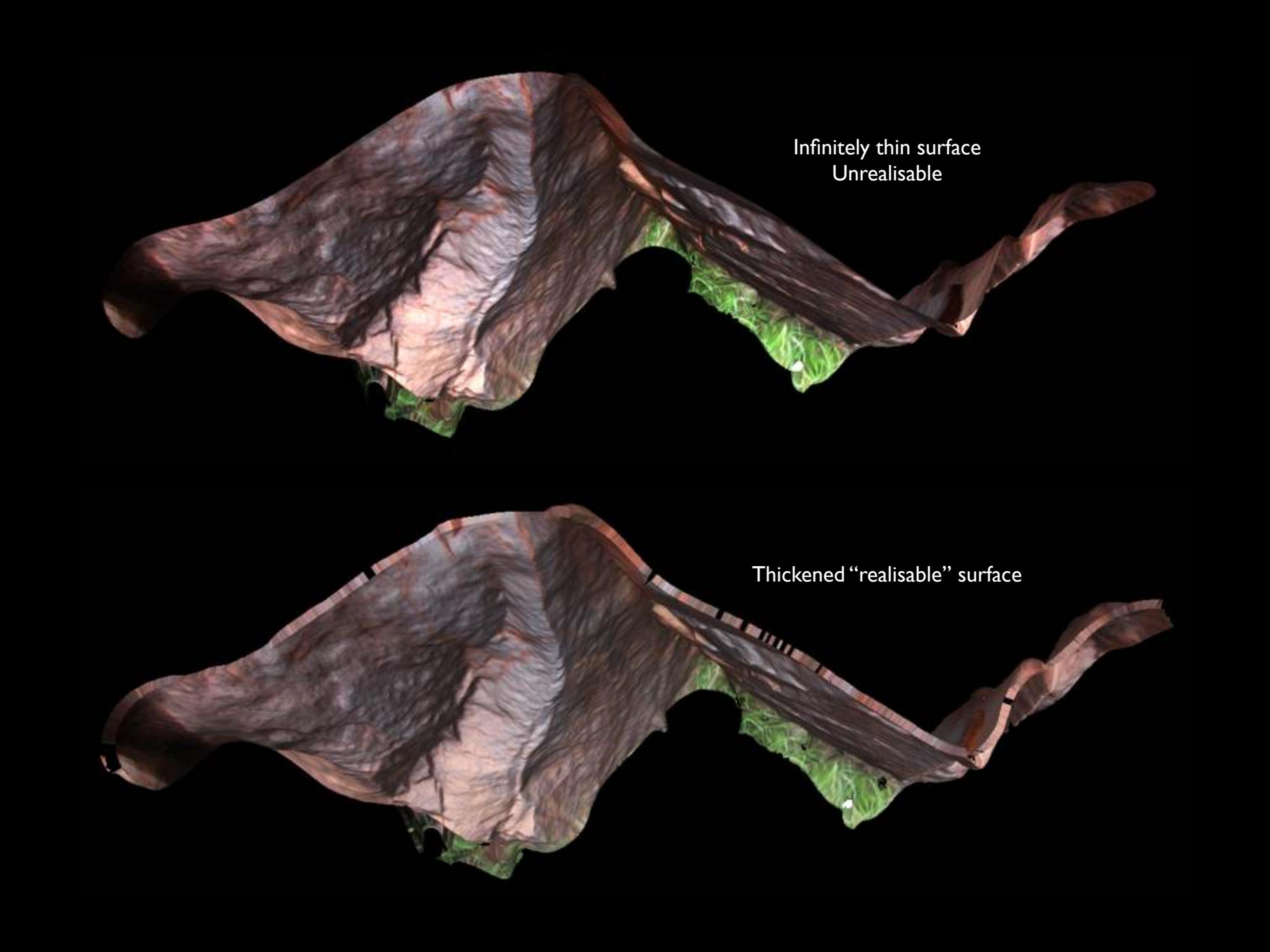


Thin joints arise at regions of high curvature

Get “poke-through” with sharp concave interiors

- Solution is called “rolling ball” thickening.
- Imagine a ball rolling across the surface, form an external mesh along the ball path.
- Implemented in Blender as a modifier called “solidify”.

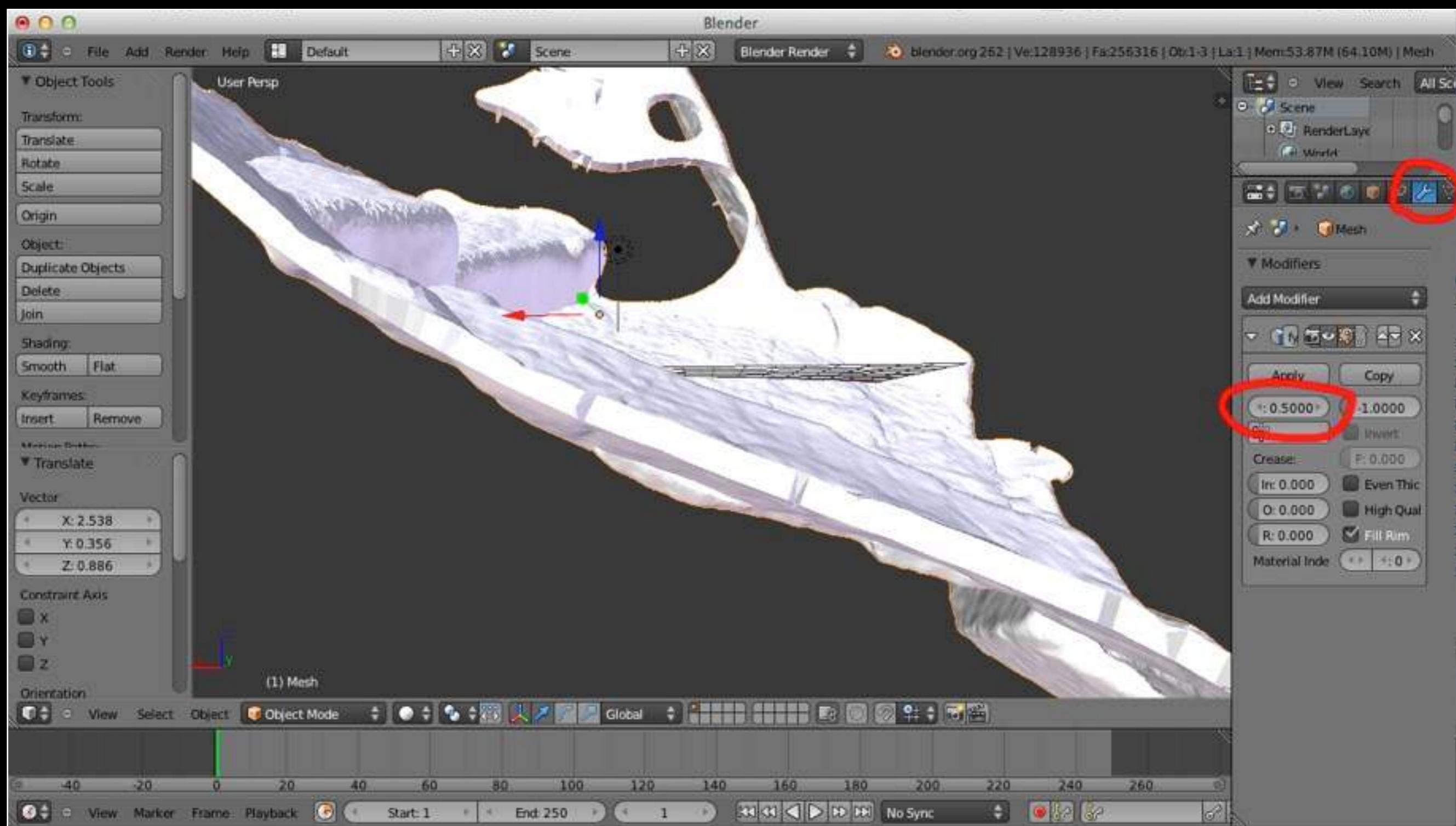




Ininitely thin surface  
Unrealisable

Thickened “realisable” surface

- “Solidify” modifier in Blender.
- Modifiers are elegant since they don’t permanently affect the geometry, can change later.



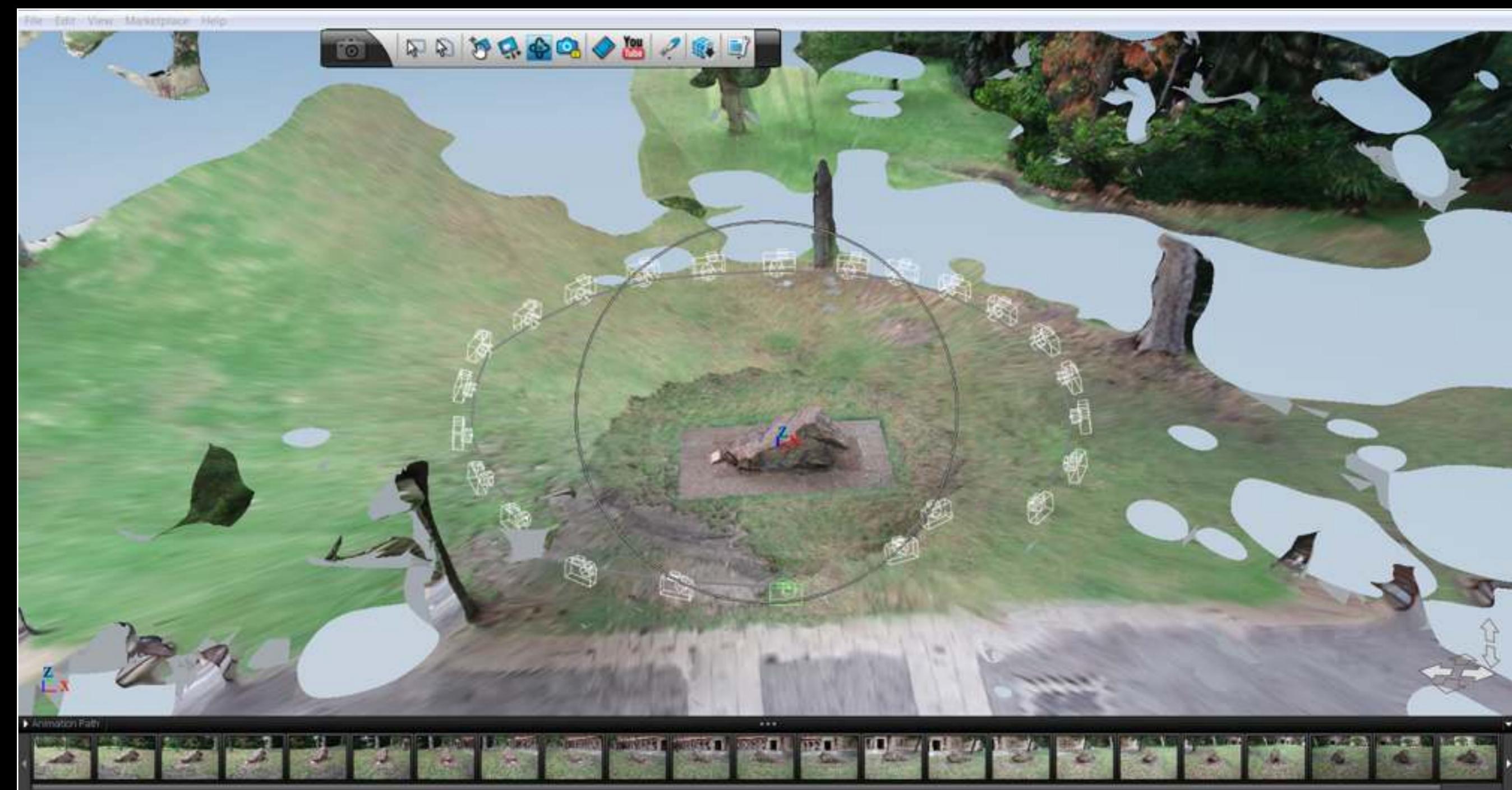
# Geometry processing : Removing shrapnel and hole closing

- Very common for there to be extraneous geometry.
- Remove reconstructed parts of the scene that are not of interest.
- Not uncommon for meshes to contain small holes, may be closed automatically by some reconstruction packages.
- Typically use MeshLab for hole closing.
- Also supported in some reconstruction packages, for example: PhotoScan.
- Don't usually contain texture information, holes usually due to regions not visible in any photograph.



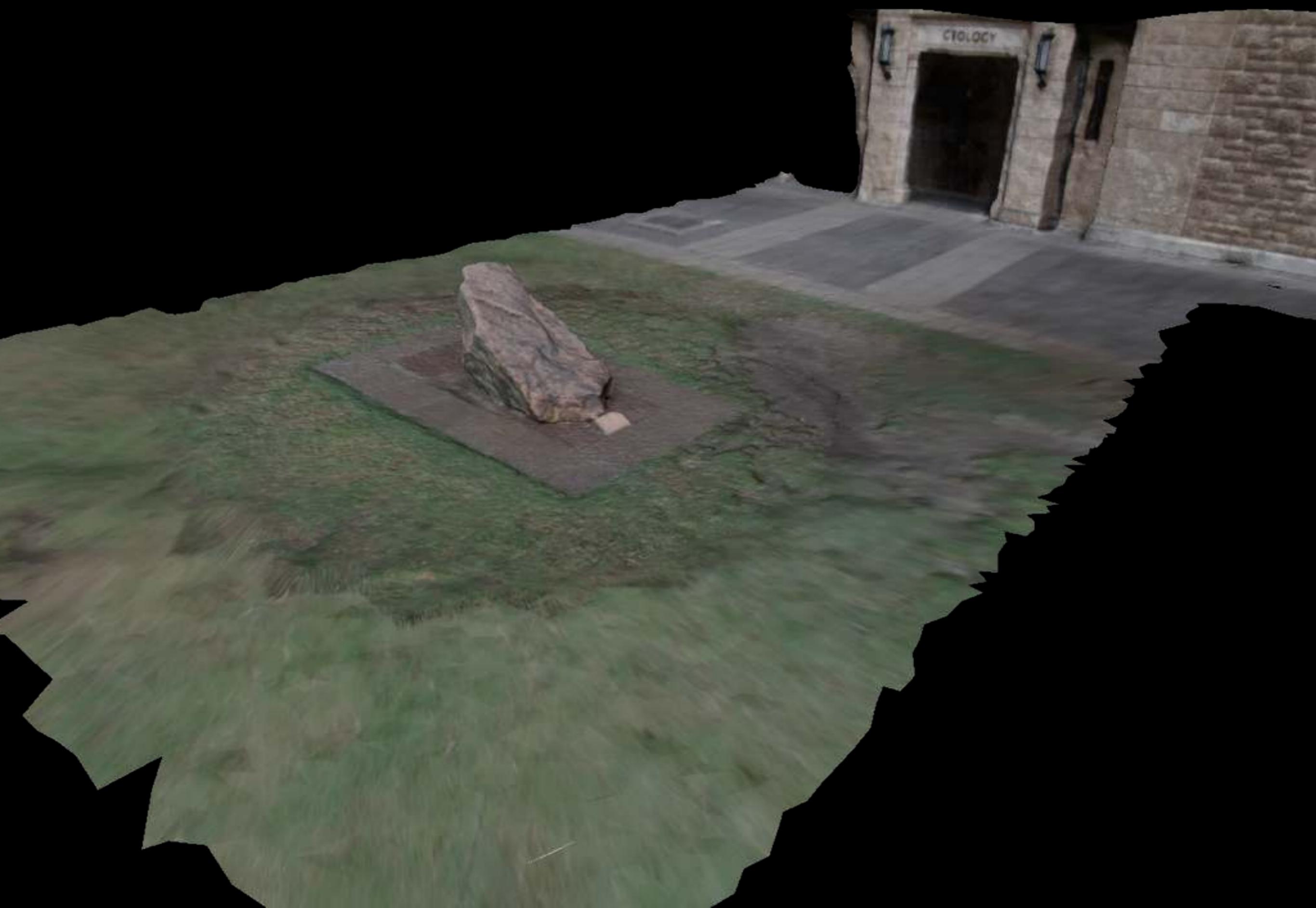
Indigenous marking stones





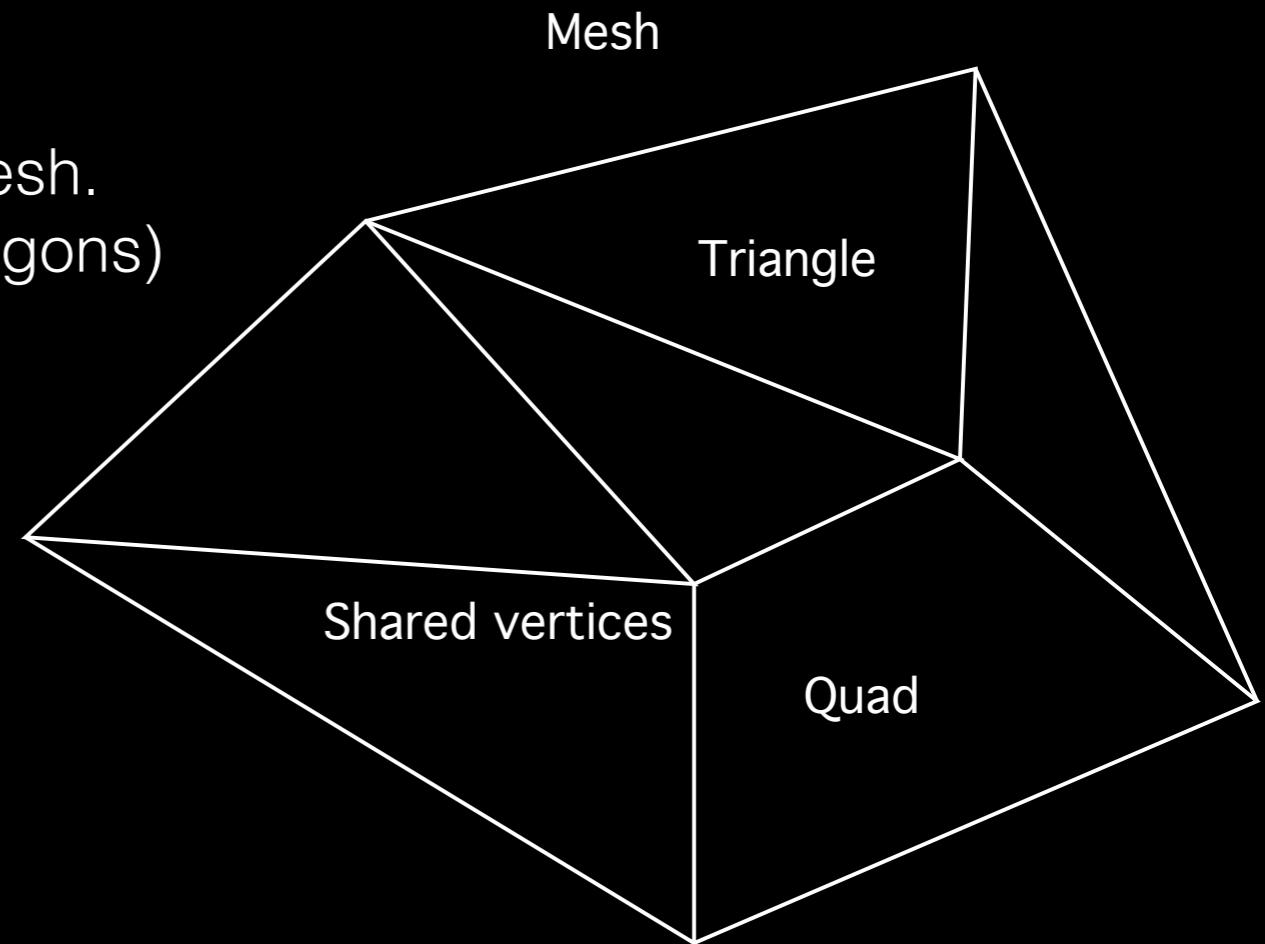
File Edit View Marketplace Help





# Geometry processing : File formats

- Requirements: unstructured triangular mesh.
  - mesh (vertices - edges - triangles - polygons)
  - texture coordinates
  - image based textures
- Common options
  - 3ds (3DStudioMax)
  - vrml, x3d
  - obj (Wavefront)
  - dae (collada)
- Pretty much standardised on obj, desirable characteristics.  $(x,y,z,u,v)$ 
  - text only so human readable
  - relatively easy to parse by software for post processing or custom utilities
  - well supported by commercial 3D applications (import/export)
  - shared vertices so no chance of numerical holes
  - supports multiple texture materials and images
- [Poorly formed obj files by 123D Catch]



- Anatomy of an OBJ file. Consists of 3 parts
  - vertex, face, normals, texture coordinates
  - materials file
  - texture image files



newmtl material\_0

Ka 0.2 0.2 0.2

Kd 0.752941 0.752941 0.752941

Ks 1.000000 1.000000 1.000000

Tr 1.000000

illum 2

Ns 0.000000

map\_Kd stone\_tx\_0.jpg

filename

material name

mtllib ./stone.obj.mtl

v 7.980470 5.627900 3.764240

v 8.476580 2.132000 3.392570

v 8.514860 2.182000 3.396990

:

:

vn -0.502475 -1.595313 -2.429116

vn 1.770880 -2.076491 -5.336680

vn -0.718451 -4.758880 -3.222428

:

:

vt 0.214445 0.283779

vt 0.213670 0.287044

vt 0.211291 0.287318

:

:

usemtl material\_0

f 5439/4403/5439 5416/4380/5416 7144/6002/7144

f 5048/4013/5048 6581/5437/6581 5436/4400/5436

f 5435/4399/5435 5049/4014/5049 5436/4400/5436

:

:

vertex  
index

normal  
index

triangles  
texture  
coordinate  
index

vertices

normals

texture  
coordinates

## Example 2 : Diotima (UWA)

- Require significantly more images ... a full 3D object.
- 16 images in this case, a relatively low number for a full 3D object.
- Some algorithms perform better if the images are captured in sequence with the best matches at the start of the bundle adjustment.
- Depends on whether the software does a compare between all images.
- Diffuse lighting conditions so no strong shadows, see later on limitations.
- “Bald” spot because no photographs from above, see later on limitations on access.
- My test subject for comparing algorithms and capture.





Diotima (Mistress of Pericles)  
16 images

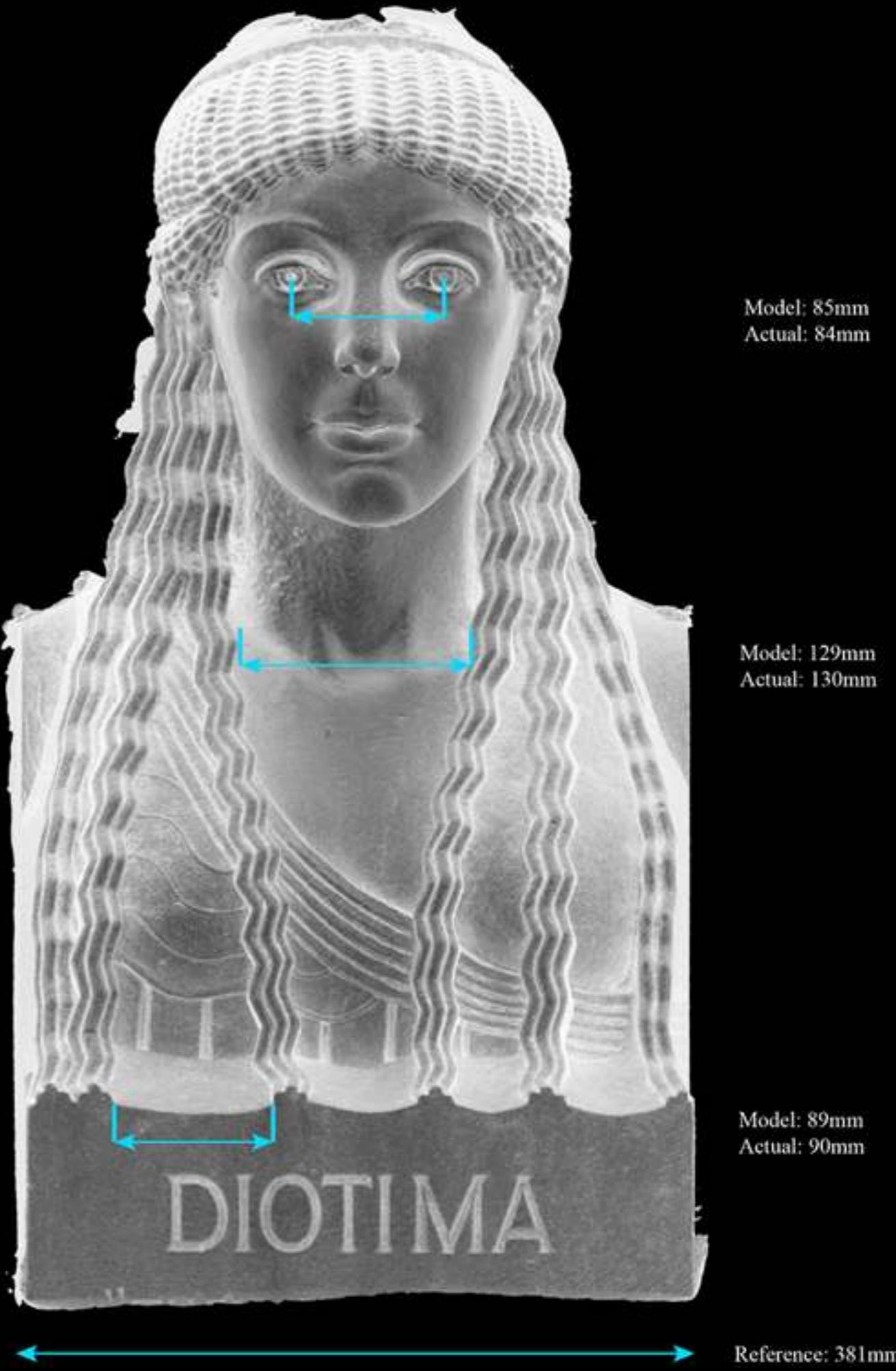
# Other topics

- Accuracy
- Resolution: Real vs apparent, Geometric vs texture
- Relighting
- Rendering
- Annotation
- Texture editing

# Other topics: Accuracy

- The first question many people ask is how accurate is it?  
Especially surveyors.
- Not easy to quantify
  1. Not all parts of a model are equally accurate.
  2. How does one get a ground truth to compare to?
  3. Accuracy can depend on characteristics of the model itself.
  4. How can you be sure best practice and best technology have been used?
- We have tested three methods to determine accuracy
  1. Reconstruct and compare key measures with known object.
  2. Perform ensemble reconstructions from large image sets, compare variation.
  3. Compare with other scanning techniques: laser scanning, CT, structured light.
  4. Visual comparison of zoomed in photographs of real and reconstructed.

- No absolute scale but use one length as reference.
- Model from 60 images.
- Subsequent measurements accurate to 2mm, most 1mm.





Original photograph



Reconstructed model



Shaded to emphasise surface variation



Original photograph



Reconstructed model

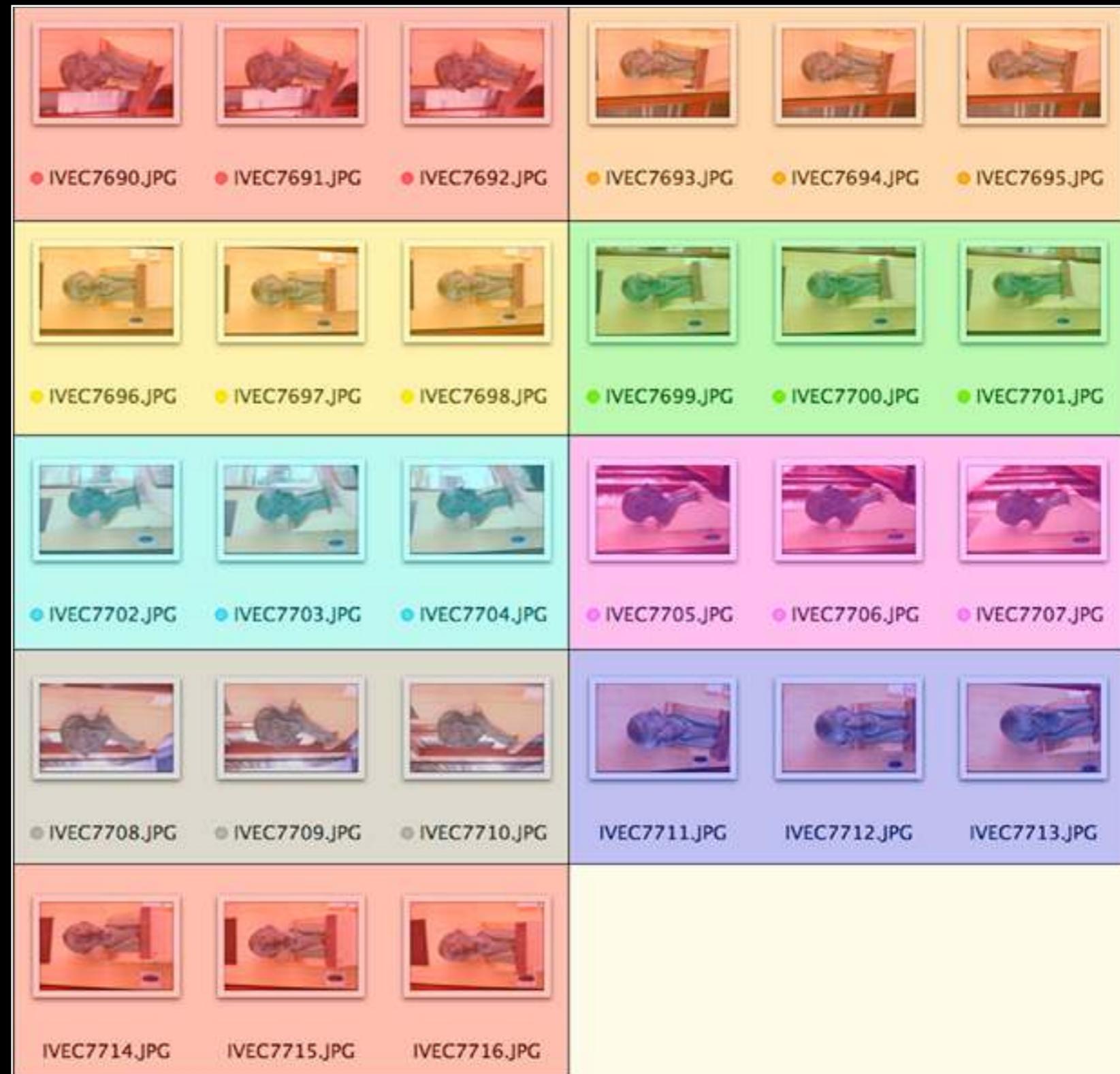


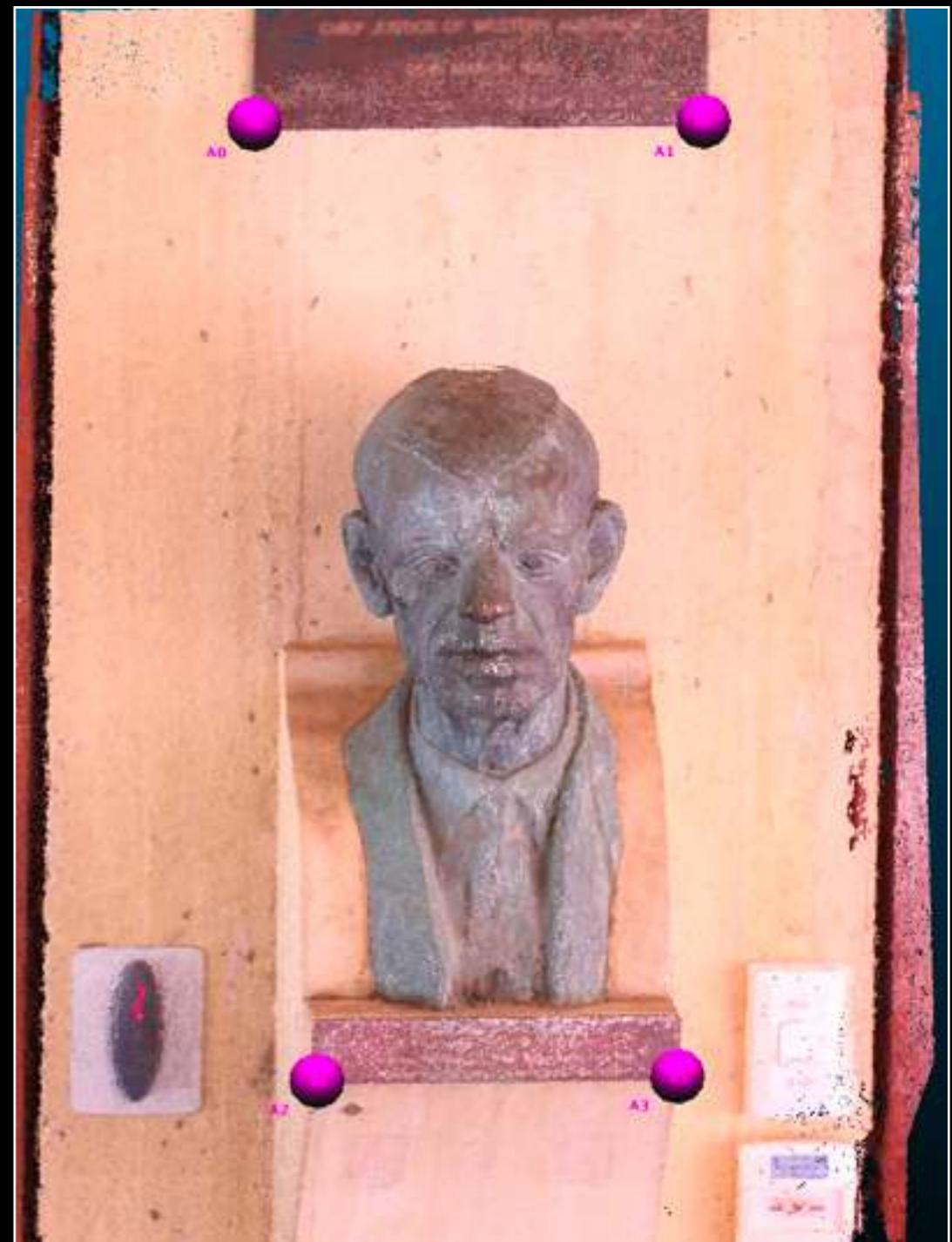
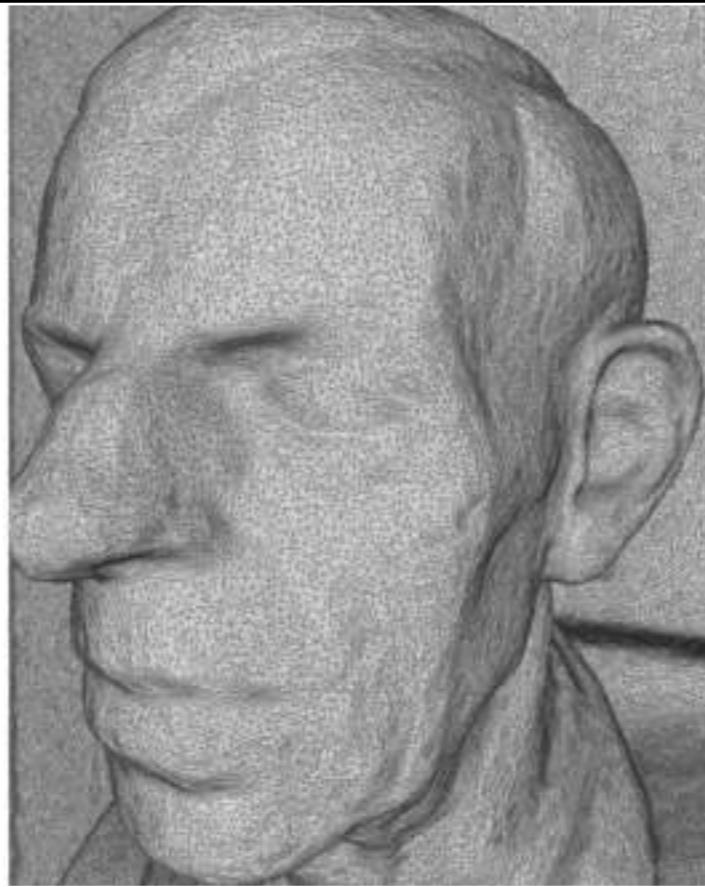
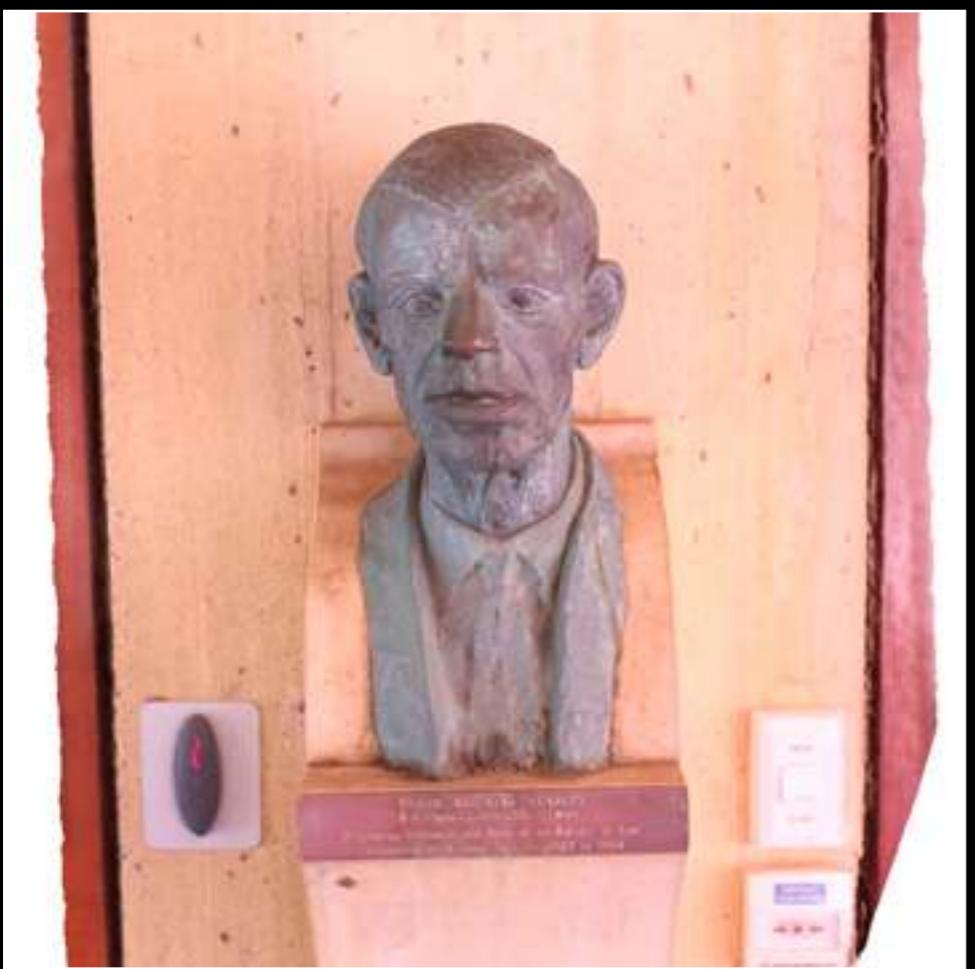
Shaded to emphasise surface variation

Original

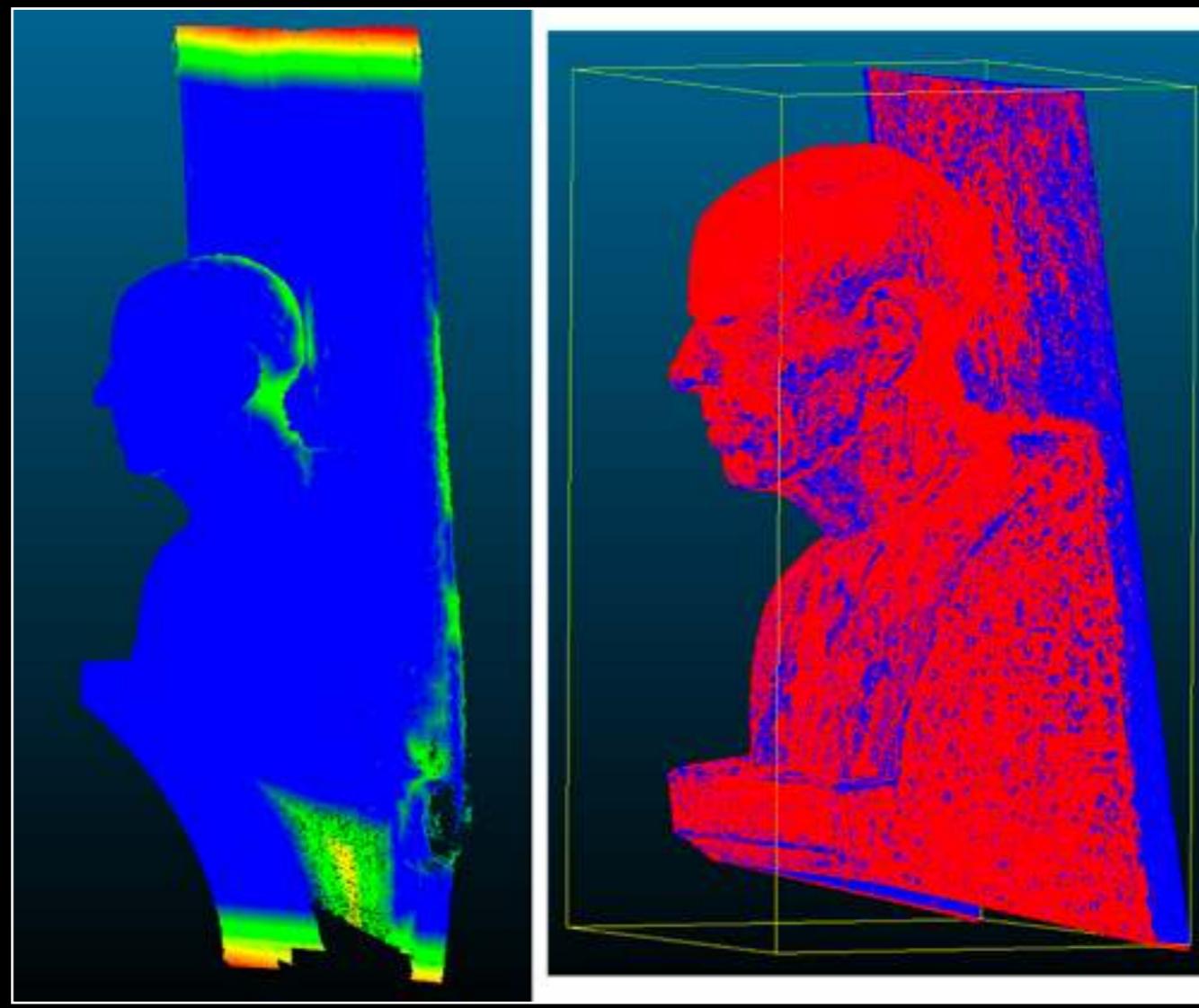


- Ensemble tests of repeatability.
- Over photograph a model. Choose random selections of photographs.
- Or, take multiple photographs from expected optimal positions, each photograph slightly perturbed. Choose random selection from each position.
- Reconstruct and compare variations in the final models.





Pairs of models are aligned by choosing 4 reference points. Translation and rotation alignment, no differential scaling permitted.



- Comparing scanning modalities
  - Photogrammetry
  - Laser scanner (NextEngine 2020I HD)
  - CT scan (Siemens SOMATOM Definition AS)
  - Structured light scanner (Artec) discounted
  - Depth camera (Kinect) discounted
- Geoscience sample



## Model A

Photos: 516

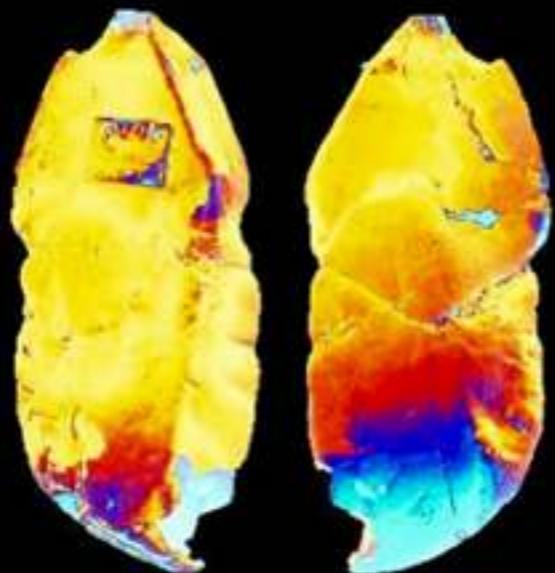
Vertices: 501226

Faces: 1000000

MAE: 0.6mm

Percent Error:

0.22%



## Model B

Photos: 307

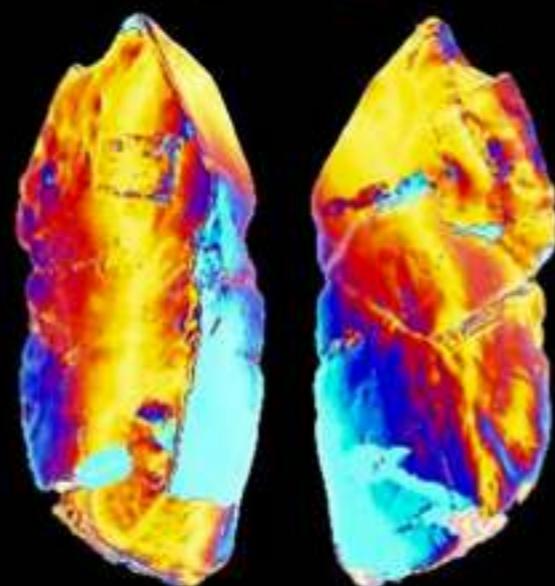
Vertices: 494004

Faces: 985143

MAE: 1.0mm

Percent Error:

0.37%



## Model C

Photos: 52

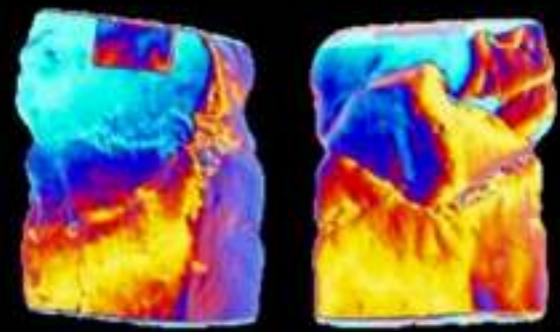
Vertices: 247455

Faces: 493195

MAE: 1.0mm

Percent Error:

0.48%



## Model D

Photos: 87

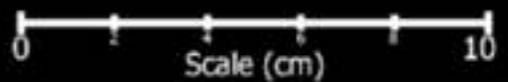
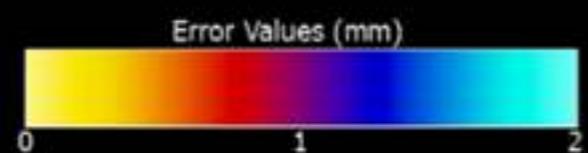
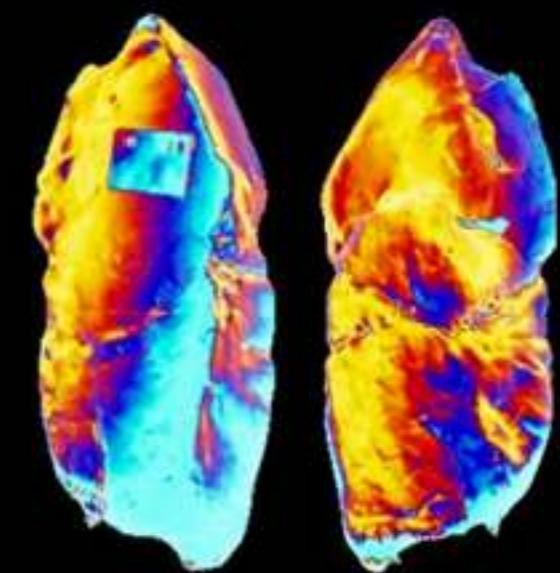
Vertices: 98803

Faces: 197215

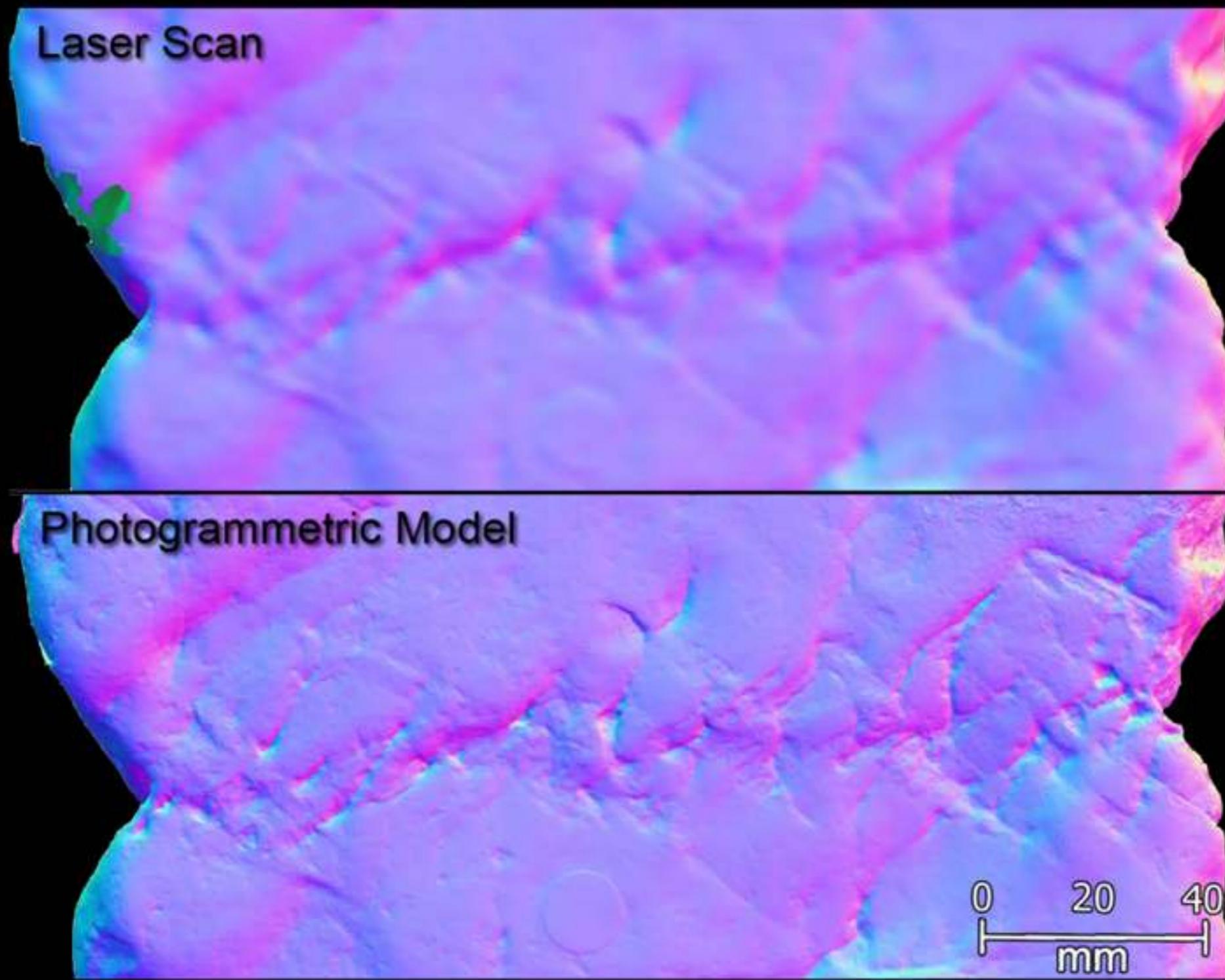
MAE: 1.1mm

Percent Error:

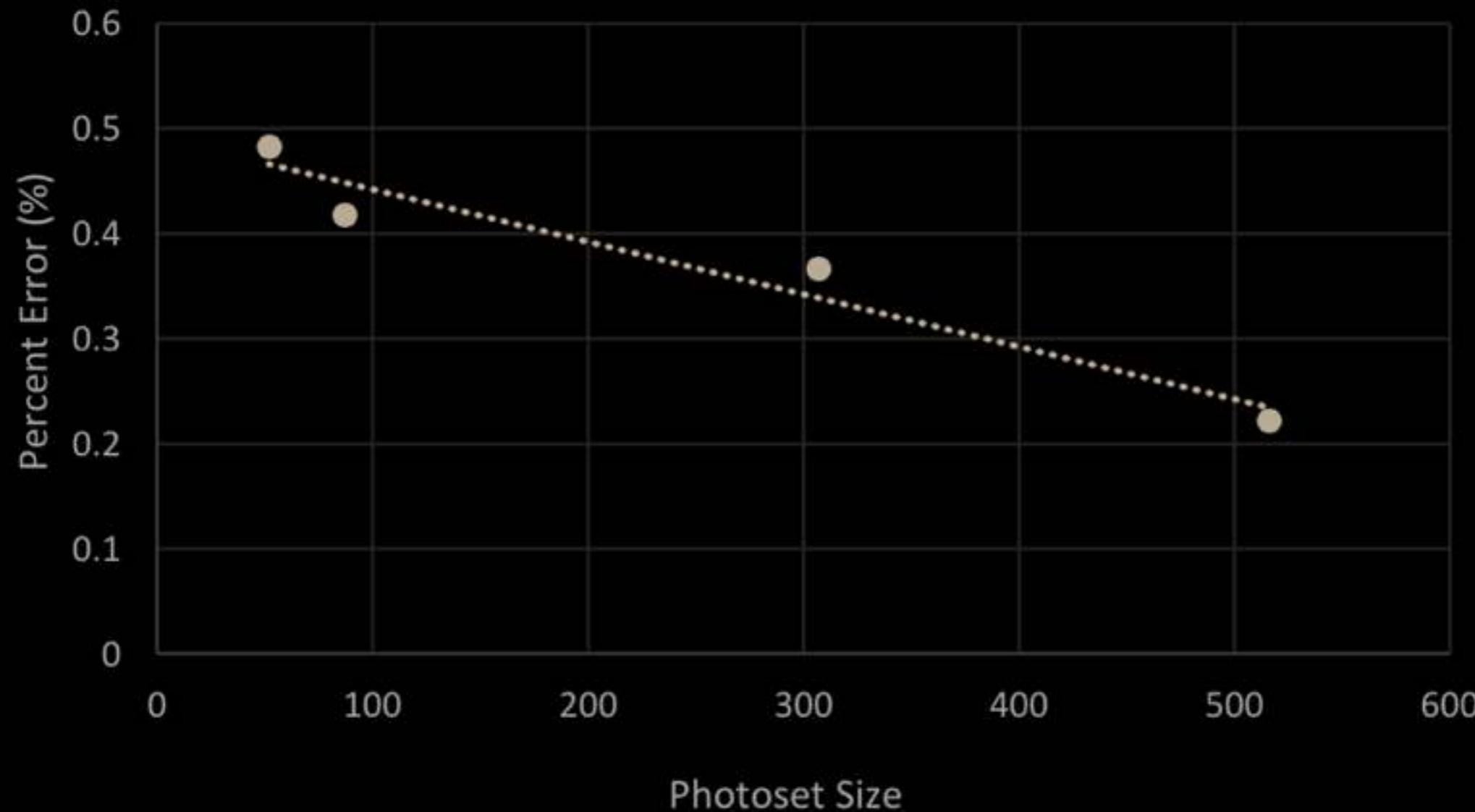
0.42%



- Is the laser scanner in this case being used optimally?
- A major problem comparing capture modalities is ensuring it is a fair test.

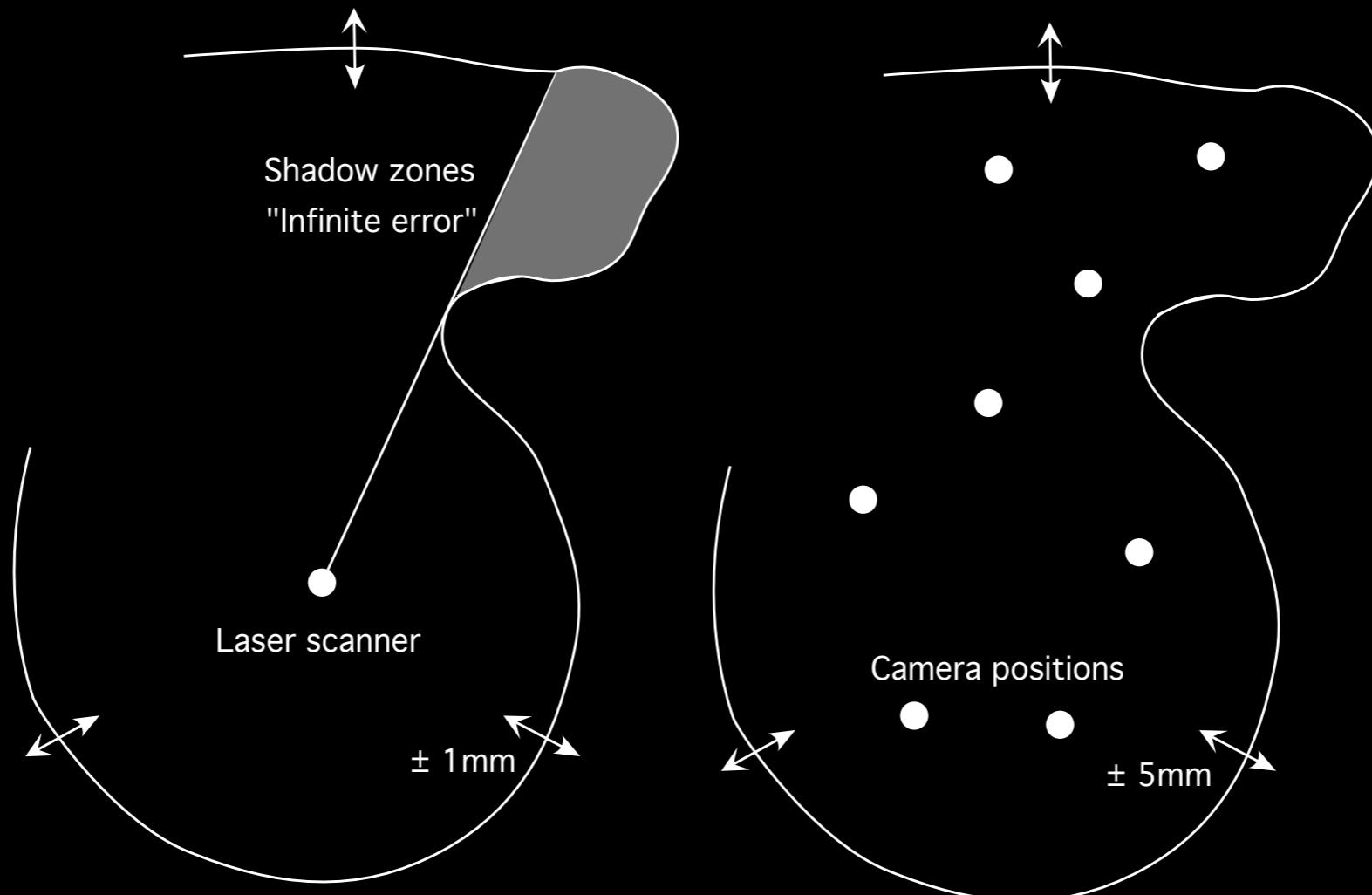


## Photoset Size vs Percent Error

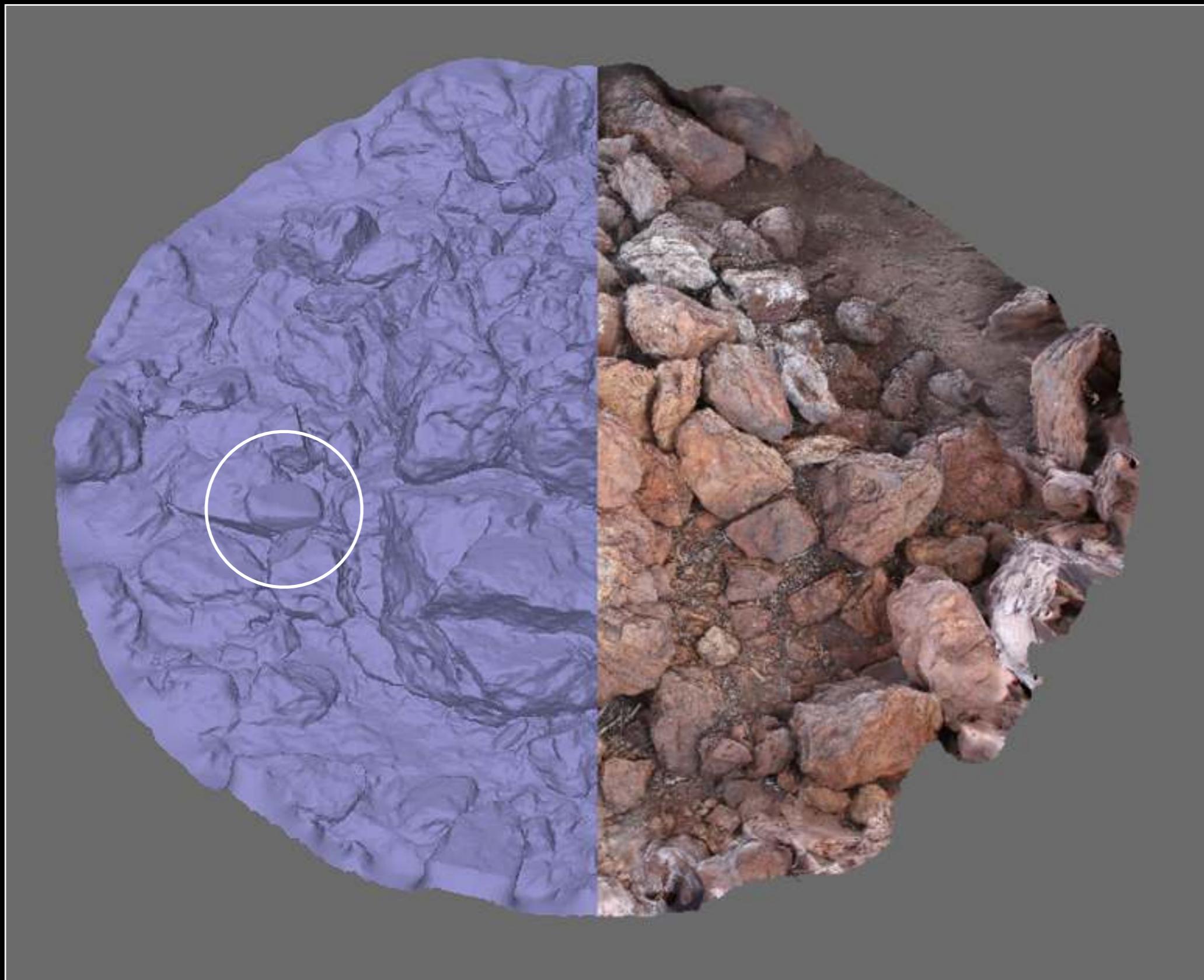


Relationship between photoset size and percentage error. While four points is not nearly enough to define a correlation, there does appear to be a linear relationship. The relationship does not seem to approach a limiting percentage error, suggesting that even greater accuracies are possible with larger photosets.

- Comparing accuracy with laser scanners.
- Not a competition, laser scanners are clearly more accurate in situations where they are suited.
- Indeed the capture modality used should be chosen based upon the desired outcomes and the nature of the scanning technology, no clear winner technology.
- Note that accuracy of laser scanners is not all it is often claimed to be.
- What does it mean for a large model to be 1mm accurate in some places but hugely inaccurate in others, occlusion zones for example?



- How does one laser scan a pile of rocks?



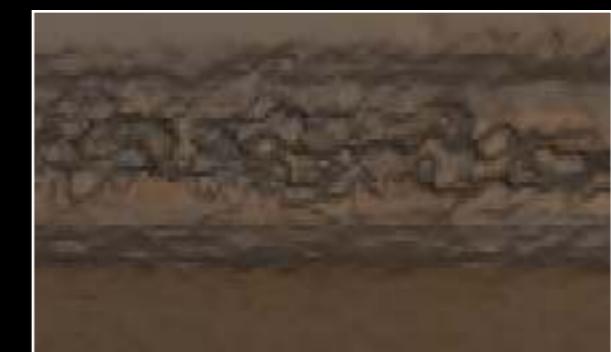
# Reconstruction vs Laser Scan summary

- Not a competition!
- One should be choosing the right tool for the job.
- Consideration also given to the purpose the models will be used for.  
Trusted dimensions for visible parts of the model?  
A high fidelity sense of the place?  
High quality textures?

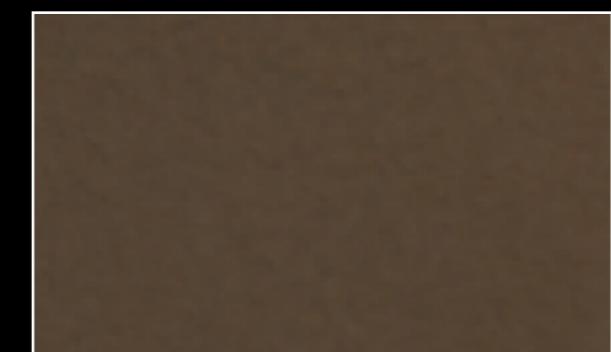
	3D reconstruction	Laser scanning
Geometric accuracy	Improving	High
Effort	Low	High
Time	Fast	Often long
Visual quality	Potentially high	Average
Occlusion issues	Less problematic	More problematic

- Laser scanners generally superior for
  - Rectangular objects
  - Featureless surfaces

Took multiple photographs to reconstruct the plain white ceiling.



Borders are plain white, like ceiling and failed to reproduce smooth surface



Wall has a texture and reconstructed well

- 20 minutes to take photographs for 3D reconstruction
- 5 hours for 6 laser scan positions ... still lots of occluded areas



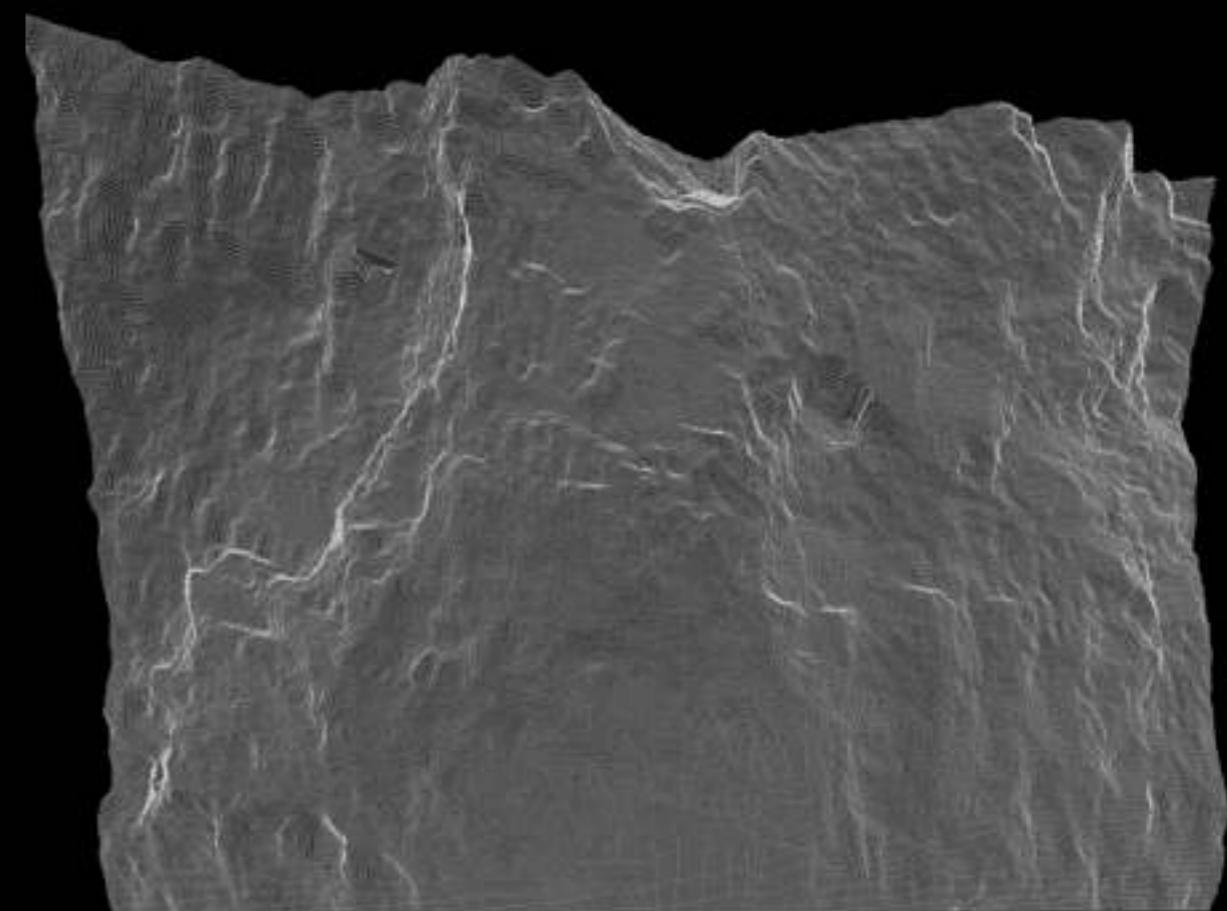
# Other topics : Resolution

- Actual mesh resolution vs apparent mesh resolution.
- Texture resolution rather than geometric resolution.
- Requirements vary depending on the end application.
  - Realtime environments require low geometric complexity and high texture detail
  - Analysis generally requires high geometric detail
  - Digital record wants high geometric and texture detail

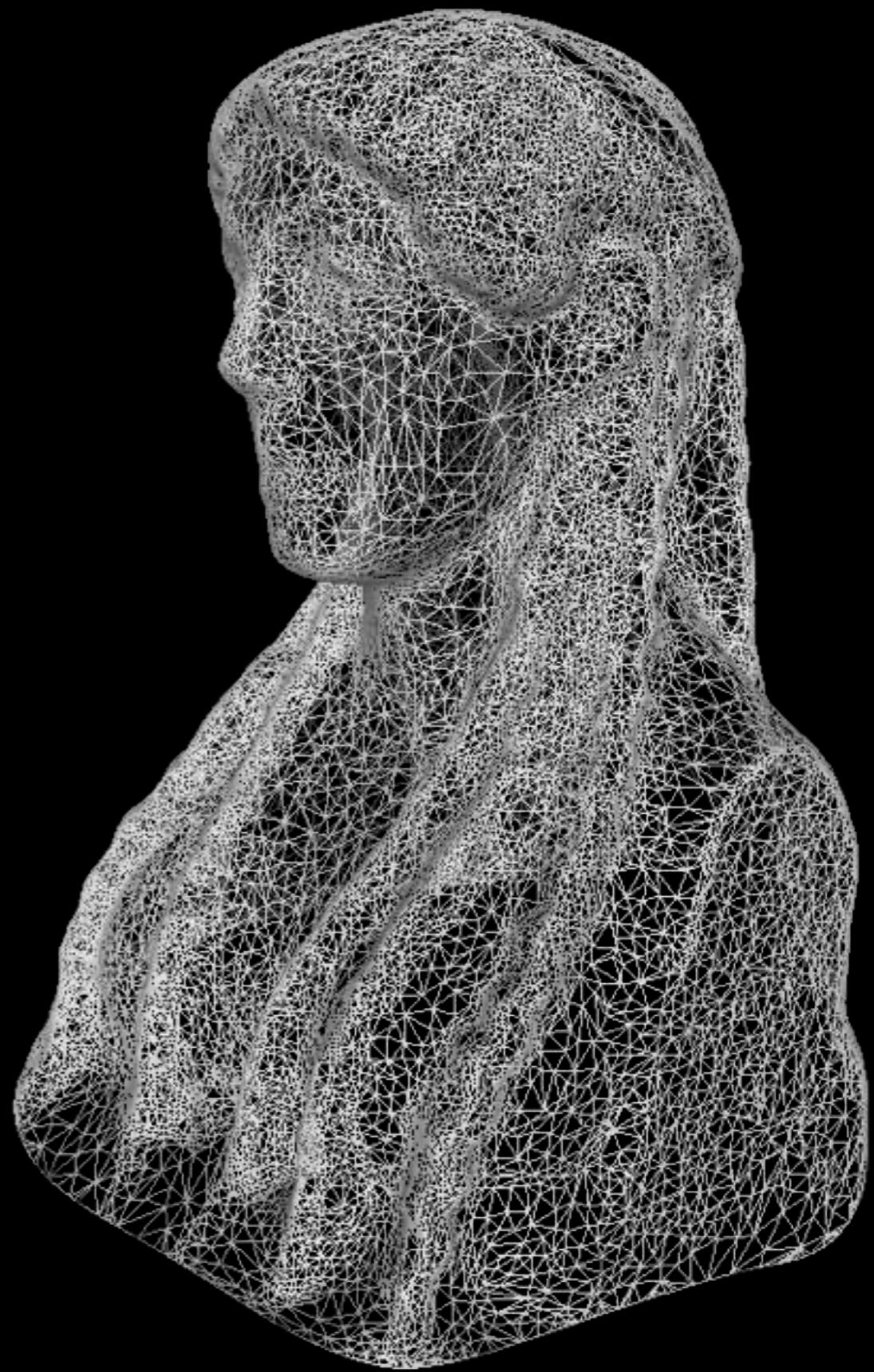
	Geometric resolution	Texture resolution
Gaming	Low	High
Analysis	High	Don't care
Education	Medium	High
Archive/heritage	High	High
Online	Low/Average	Low/average



Apparent high resolution



Low resolution mesh



Example from 2010



Example from 2014

# Other topics : Relighting

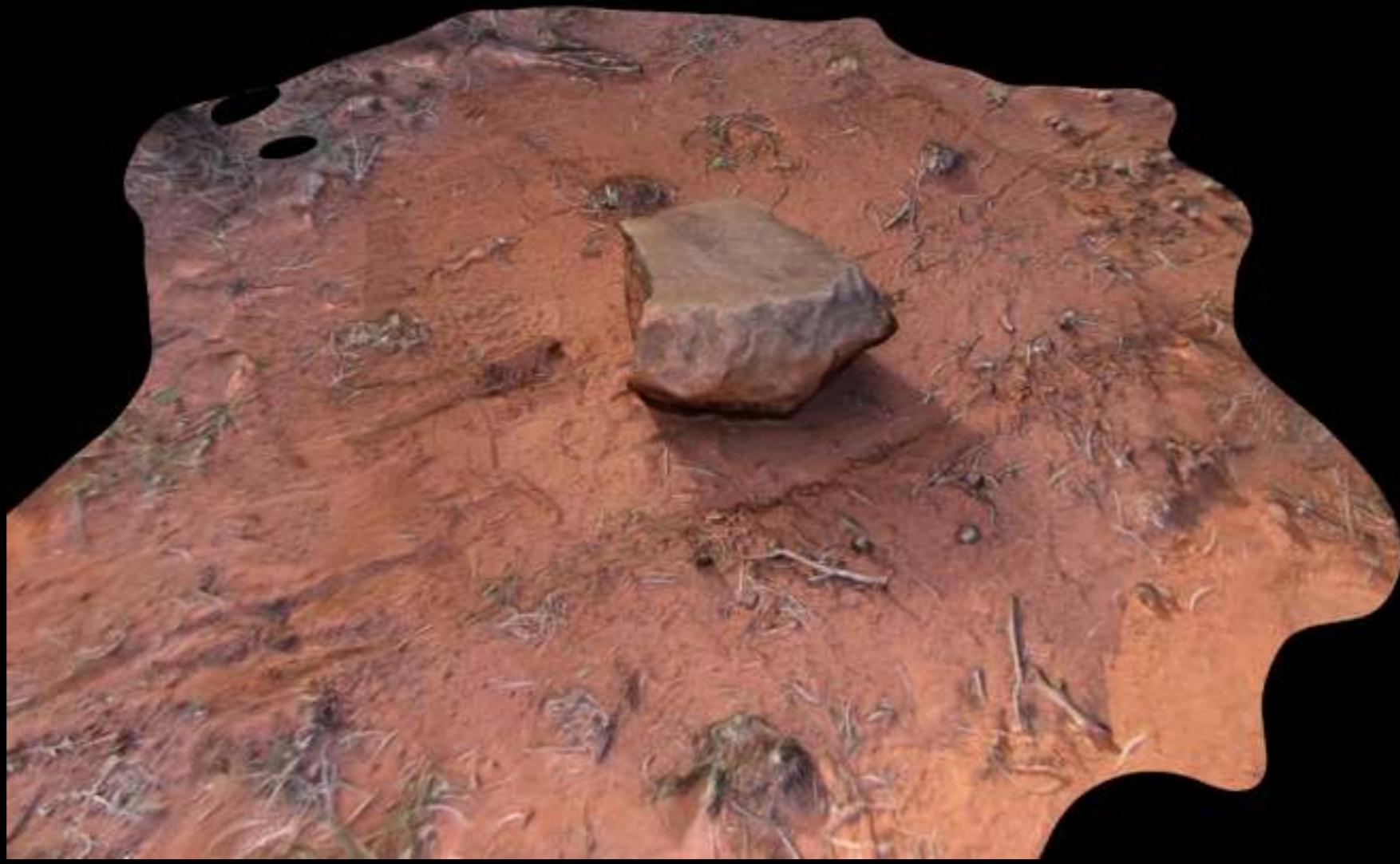
- We have a 3D model, can “relight” it.  
For example: cast shadows, adjust diffuse/specular shading.
- Obviously works best with diffuse lit models.
- See later for baked on texture limitations.
- Interesting in the archaeology context since it is well known that some features are “revealed” in different lighting conditions.
- Cannot replicate effects of dyes but can replicate effects due to shading/shadowing of fine details.





Wanmanna

# Other topics : Rendering



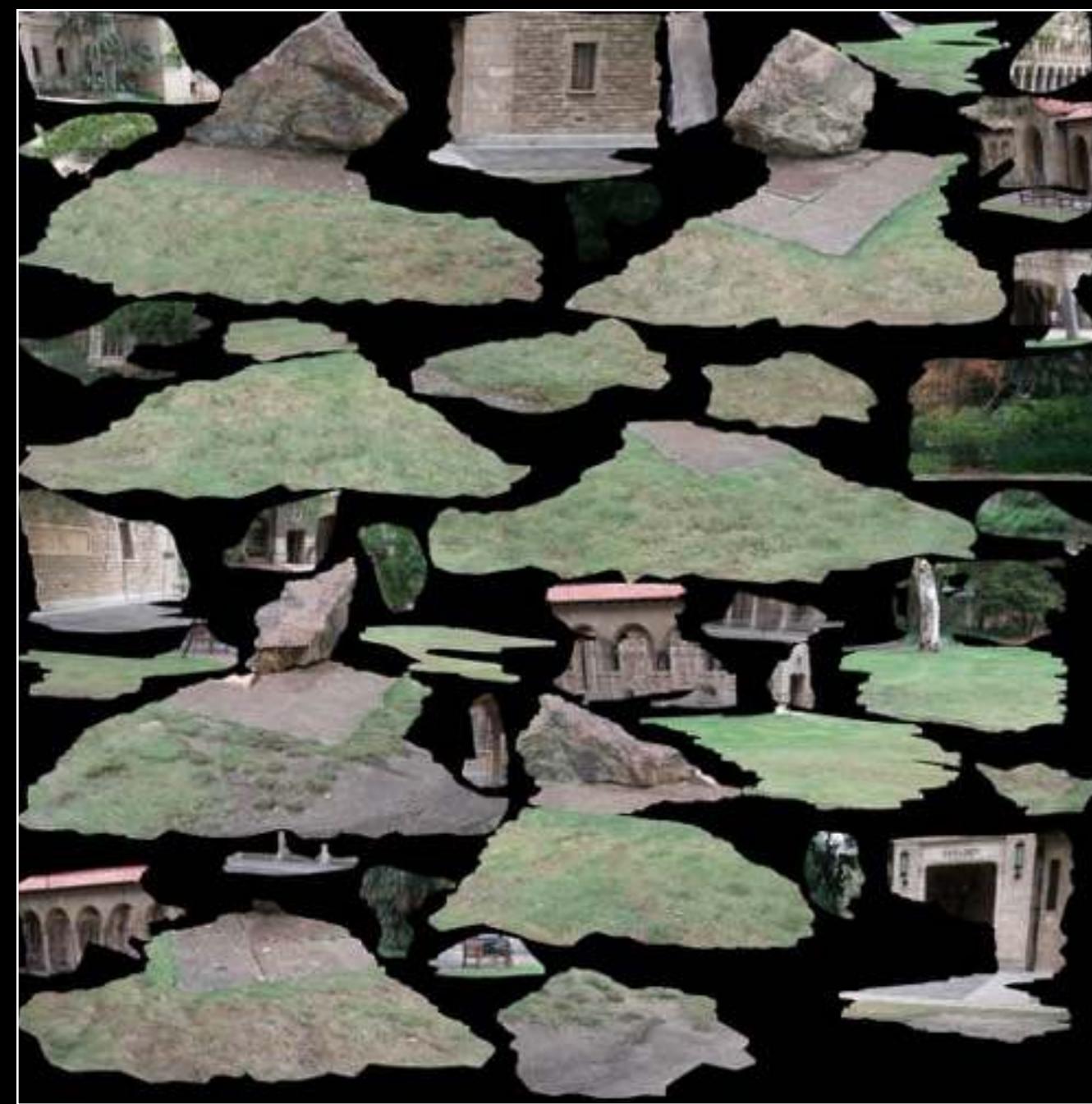
# Other topics : Analysis



1000 Buddha temple, Manipal, India

# Other topics : Annotating

- Textures from the reconstruction algorithms are often “interesting”.
- Exact form of the texture depends to some extent on the software being used  
Can often identify the software based upon the appearance of the texture maps.
- They are derived from re-projection of the image from the derived camera position onto the reconstructed mesh, hence potentially very high quality (perceived resolution).
- Can generally still be drawn on, treated as an image for image processing in PhotoShop, etc.



Texture map 1



Texture map 2



v

u

l

0 0



Textured mesh



0  
0

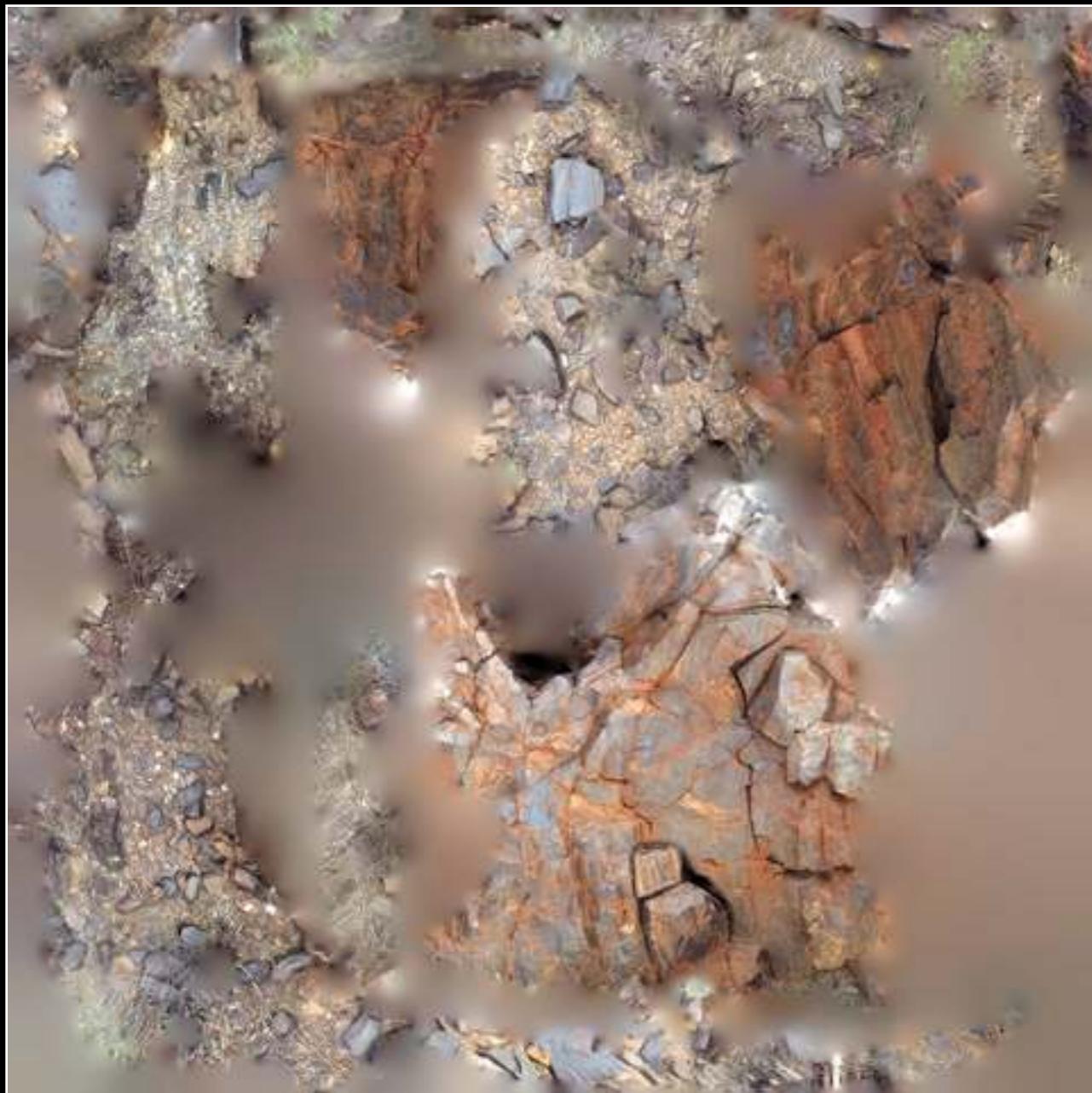
u





# Other topics : Texture editing

- Some texture mapping modes are easier to edit than others
- Can be difficult for per camera reprojected textures (left)
- Easier for orthographic texture maps (right), but not always a supported option.



- Can obviously do colour correction/grading on the texture post reconstruction.
- Need to be careful with filtering which may cause artefacts between texture seams.





# Limitations and Challenges

- Occluders - Problematic
- Movement in the scene
- Thin structures
- Baked on shadows
- Lighting changes during capture
- Access to ideal vantage points
- Online and database access
- High level queries for geometric
- Reflective surfaces

# Limitations : Occluders

- Algorithms seem to be generally poor at handling foreground occluders.
- For example: columns in front of a building.
- Reason: a small change in camera position results in a large difference in visible objects.
- Capturing the backdrop behind an object.
  - Often better, assuming possible, to capture them separately





St Lawrence, Manipal, India

# Limitations : Movement

- Objects to be reconstructed obviously need to be stationary across photographs.
- Grass moving in the wind is a common problem for field work.
- Solution is to create a camera array for time simultaneous photography.



# Limitations : Thin structures

- Difficult to reconstruct objects approaching a few pixels in the images (sampling theory).
- Example of grasses in the rock art reconstruction.



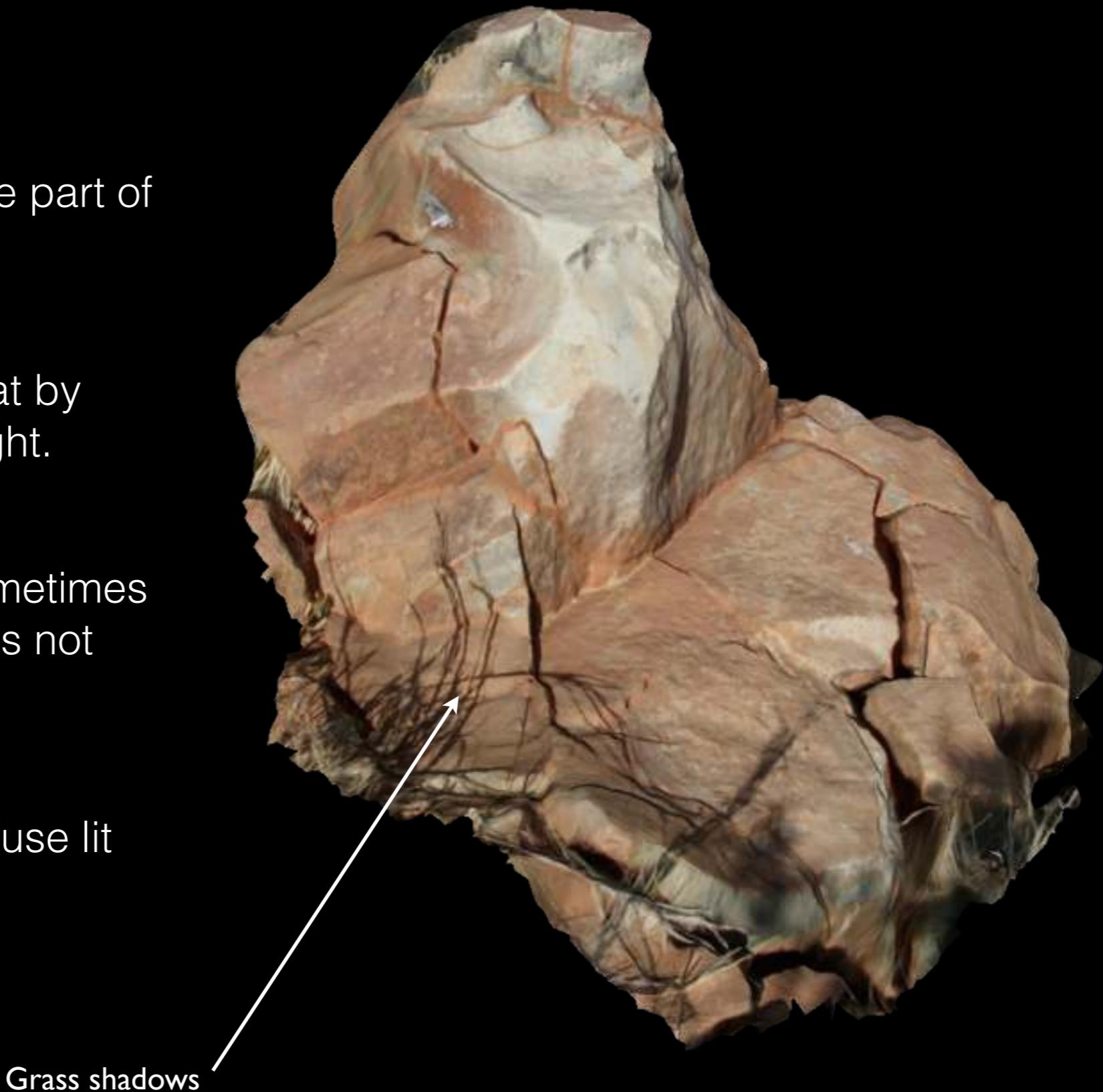
Not 3D structure but grass texture on rock face

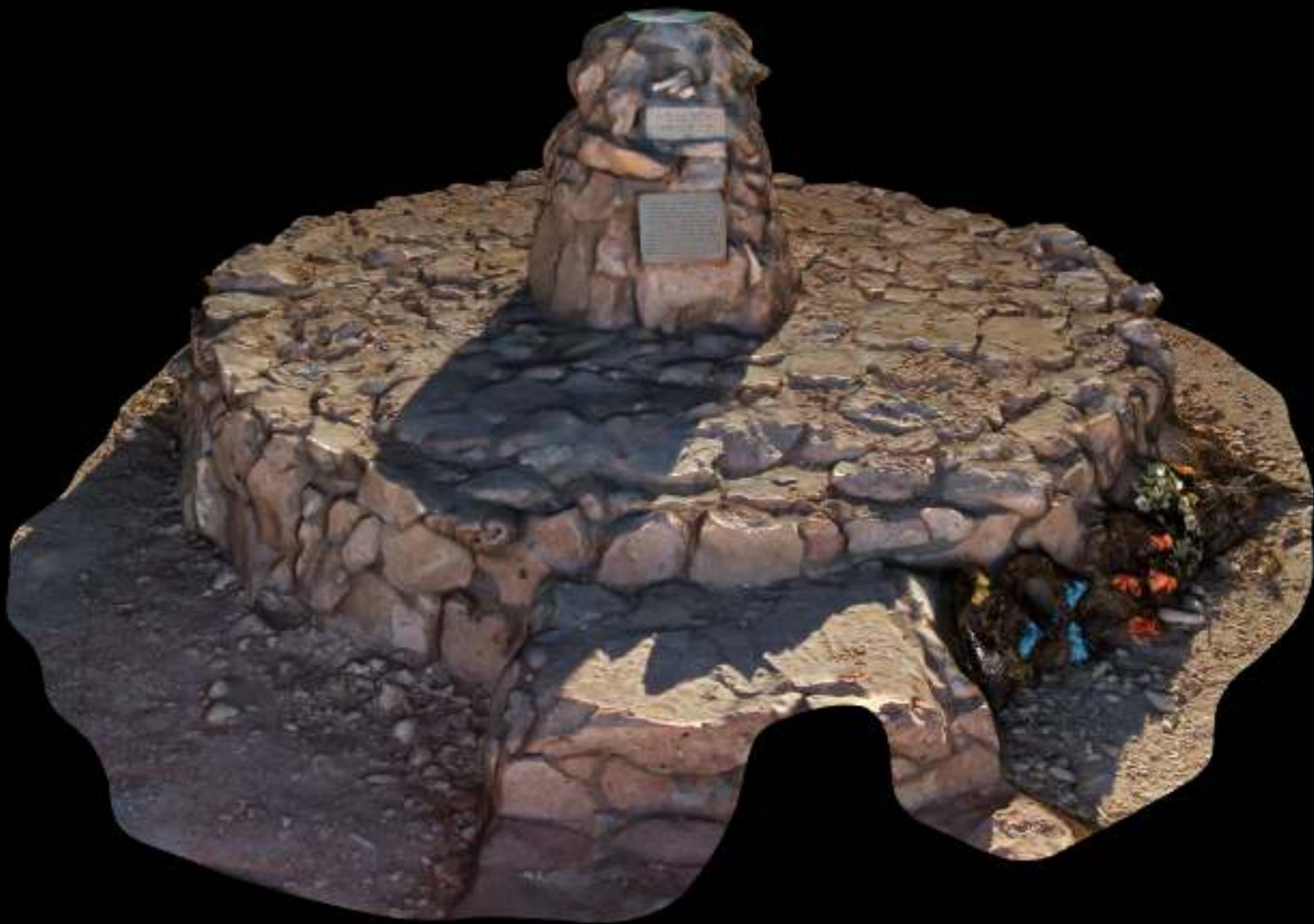


Grass not resolved

# Limitations : Baked on shadows

- Shadows obviously become part of the texture maps.
- Can be alleviated somewhat by photographing in diffuse light.
- For outside objects can sometimes choose times when object is not directly lit.
- Can sometimes choose diffuse lit days, cloudy.





HMAS Sydney Cairn, Canarvon

# Limitations : Lighting changes and access

- For field work access to preferred positions for photographs may be problematic.
- Similarly capturing photographs from above the object, elevated positions.
- When capturing 30+ photographs for 3D objects the lighting conditions may change  
eg: clouds passing overhead.  
Processes generally insensitive to this except for variations in resulting textures.
- Shadows of the photographer.

# Limitations : Reflective surfaces

- Mirror surface fold the world about the mirror plane.
- Not unexpected then than reconstructions can build the world “behind” the mirror.



- Mirror surfaces can provide a non-linear reflection of the world that will influence the feature point detection.
- Gives rise to a new art form.
  - Photogrammetry that goes wrong in “interesting” ways.
- Bug highly specular surface can also be an issue.



Fort Canning, Singapore

# Limitations : Database/online representations

- Claim that the need to store these higher level forms of data capture will increase.
- Will this replace the need for storing photographic data?
- Surprisingly (depressingly) even after all these years of online delivery there are still no entirely satisfactory ways of distributing 3D data.
- Options
  - VRML, x3d : very poor cross platform support
  - 3D PDF : dropped by Adobe some years back
  - WebGL? HTML5 / Canvas?
- Key missing components:
  - progressive texture
  - progressive geometry



# Example 3 : Grinding stone

- Will do a full worked example based upon grinding stone from the Ngintaka story
- 22 photographs around the stone
- Example of light/colour changes due to polarising filter and angle to sun direction



DSCF5040.JPG

DSCF5041.JPG

DSCF5042.JPG

DSCF5043.JPG

DSCF5044.JPG

DSCF5045.JPG

DSCF5046.JPG



DSCF5047.JPG

DSCF5048.JPG

DSCF5049.JPG

DSCF5050.JPG

DSCF5051.JPG

DSCF5052.JPG

DSCF5053.JPG



DSCF5054.JPG

DSCF5055.JPG

DSCF5056.JPG

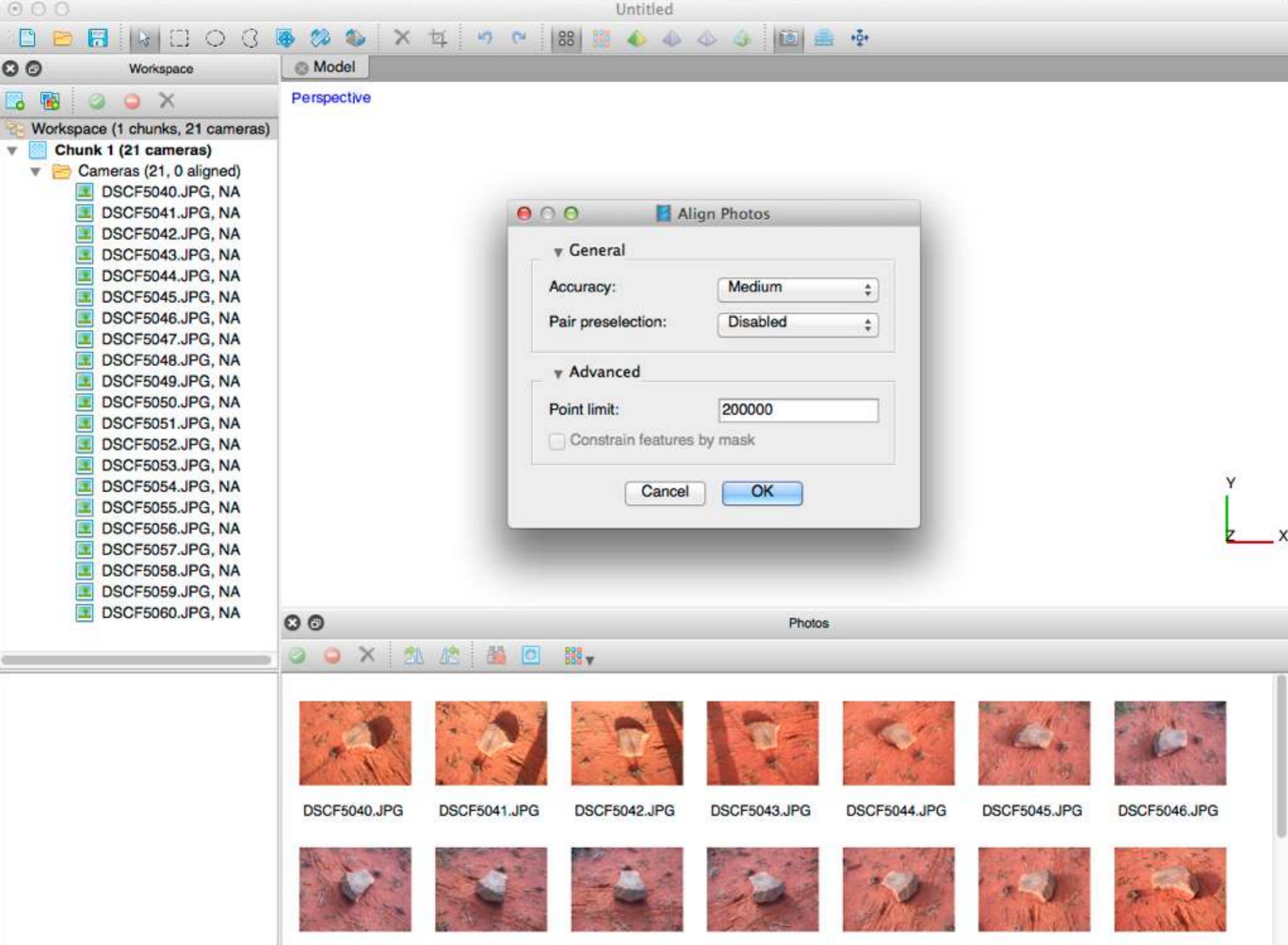
DSCF5057.JPG

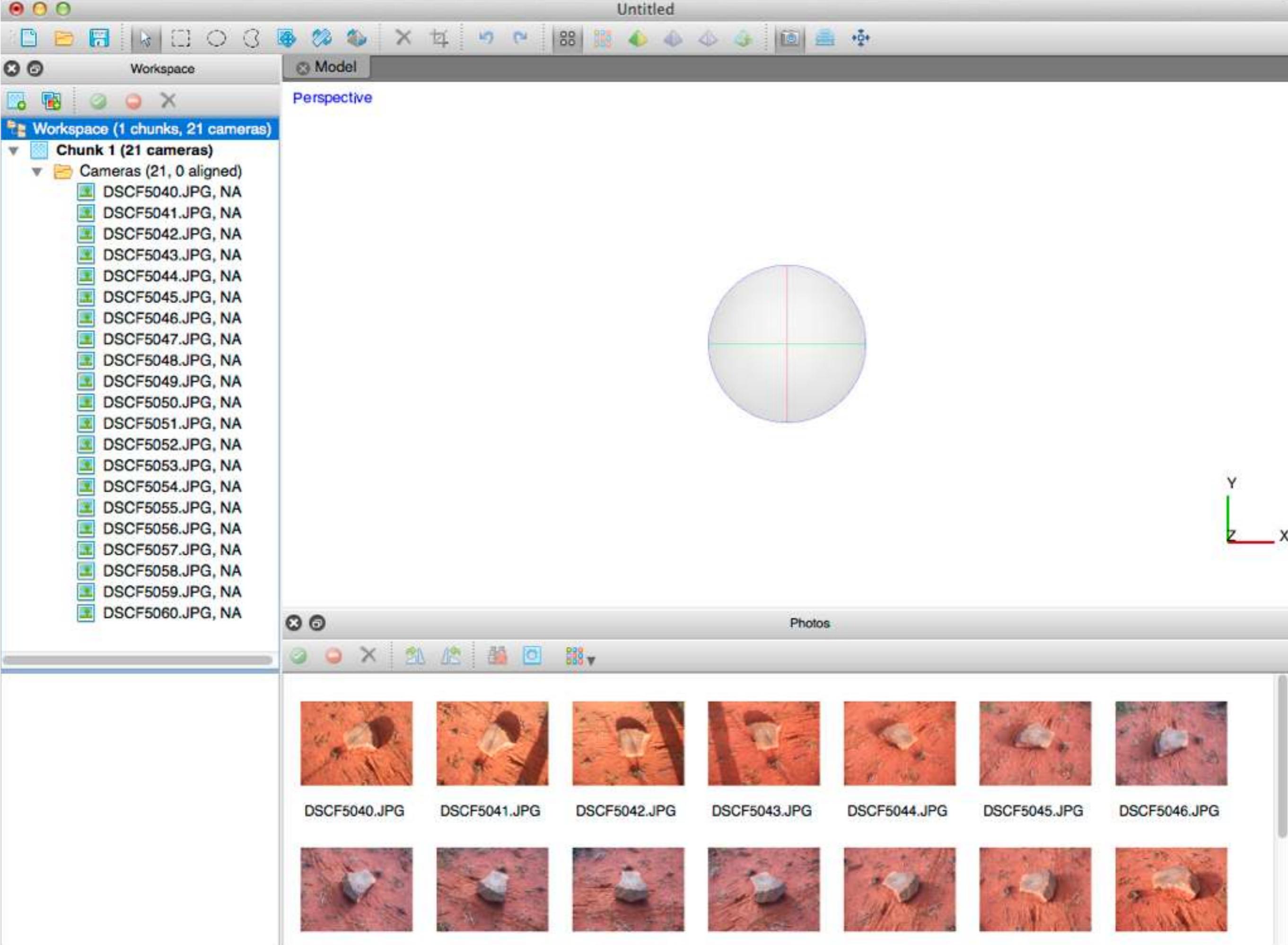
DSCF5058.JPG

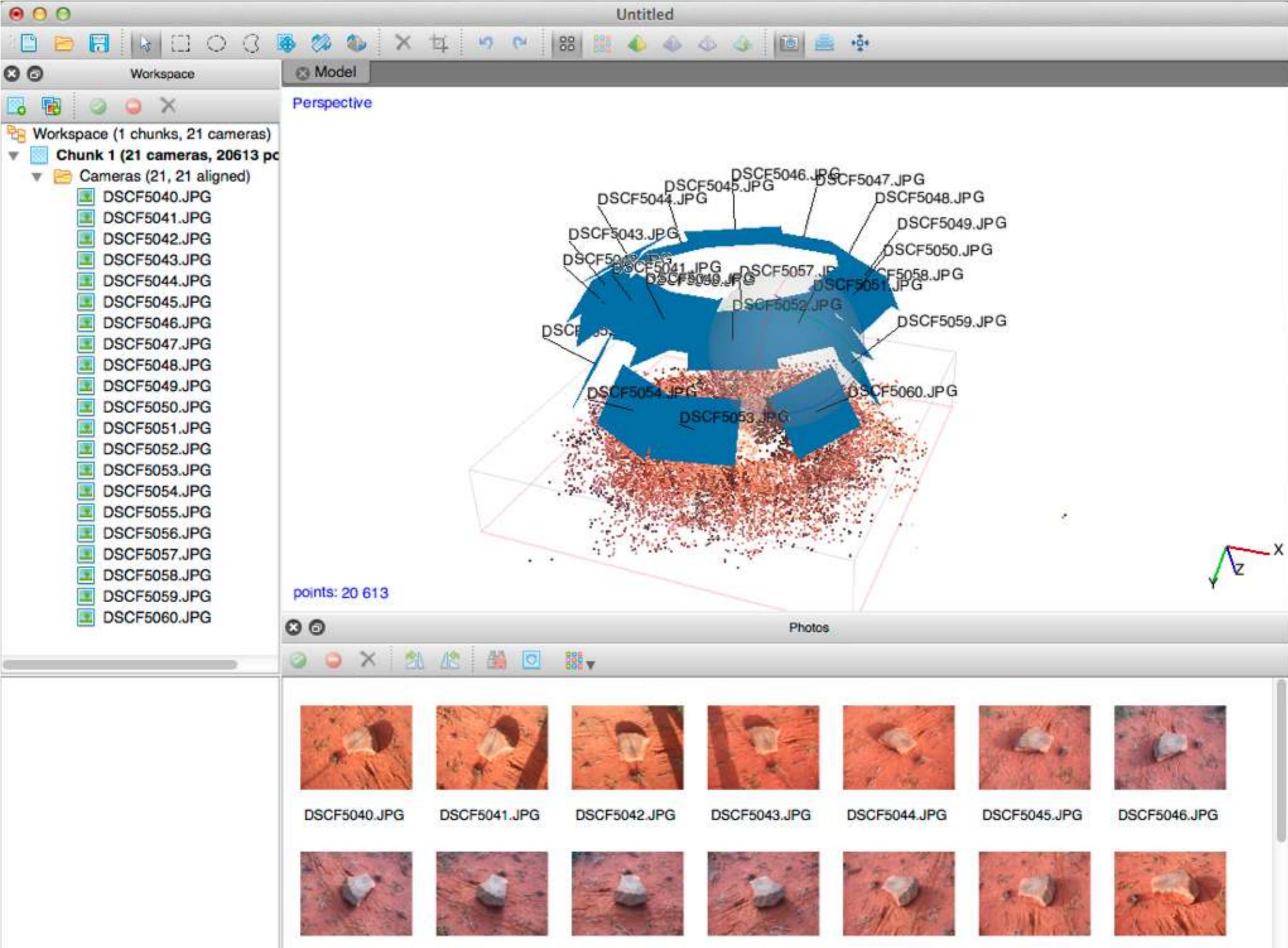
DSCF5059.JPG

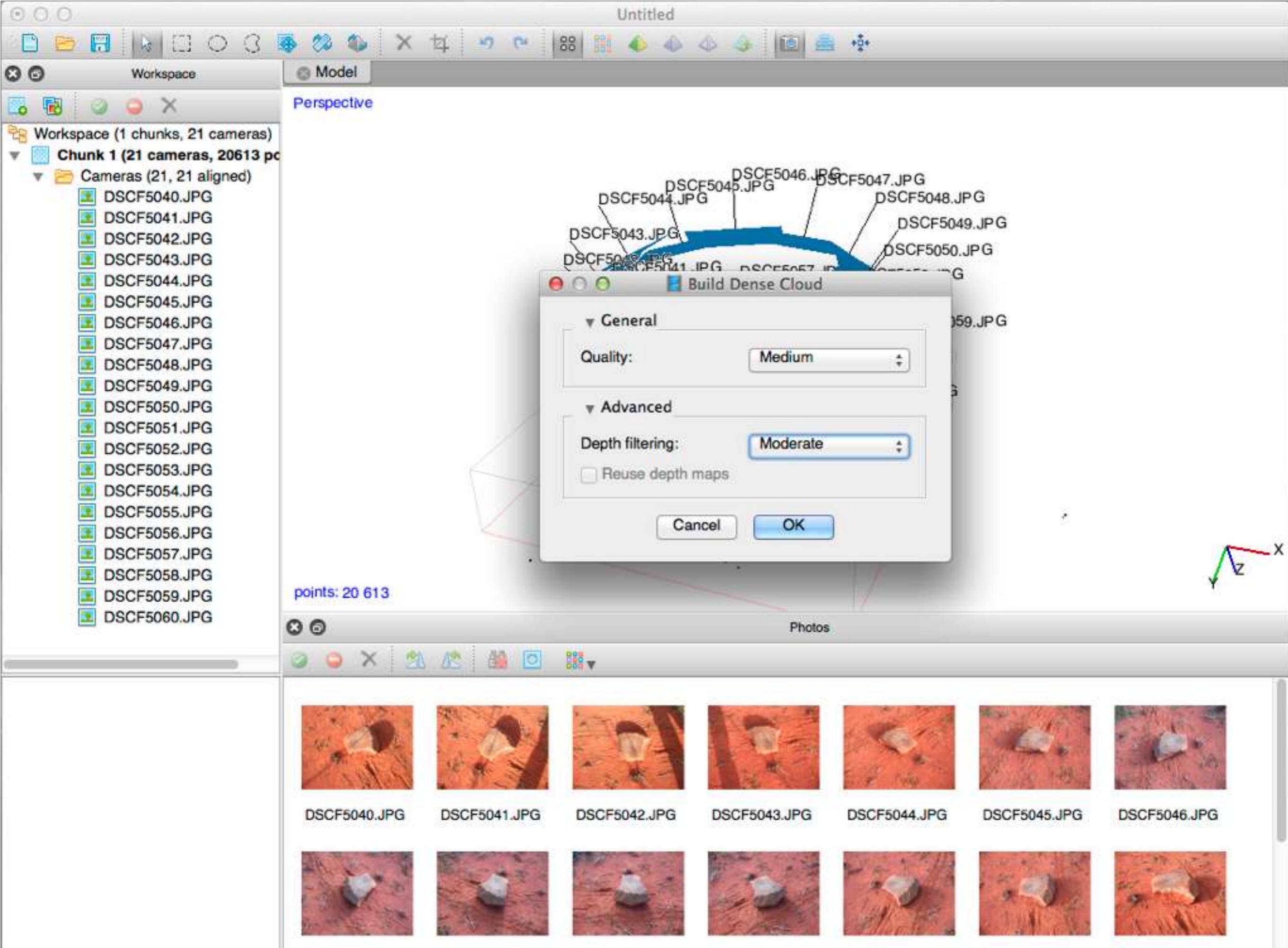
DSCF5060.JPG











Untitled

Navigation Model Perspective

Workspace (1 chunks, 21 cameras)

Chunk 1 (21 cameras, 20613 pc)

Cameras (21, 21 aligned)

- DSCF5040.JPG
- DSCF5041.JPG
- DSCF5042.JPG
- DSCF5043.JPG
- DSCF5044.JPG
- DSCF5045.JPG
- DSCF5046.JPG
- DSCF5047.JPG
- DSCF5048.JPG
- DSCF5049.JPG
- DSCF5050.JPG
- DSCF5051.JPG
- DSCF5052.JPG
- DSCF5053.JPG
- DSCF5054.JPG
- DSCF5055.JPG
- DSCF5056.JPG
- DSCF5057.JPG
- DSCF5058.JPG
- DSCF5059.JPG
- DSCF5060.JPG

Dense Cloud (1719334 points)

Perspective

5057.JPG

DSCF5045.JPG

DSCF5056.JPG

DSCF5044.JPG

DSCF5042.JPG

DSCF5043.JPG

points: 733 951 selection: 115

DSCF5048.JPG

DSCF5059.JPG

DSCF5049.JPG

DSCF5050.JPG

DSCF5060.JPG

Z X Y

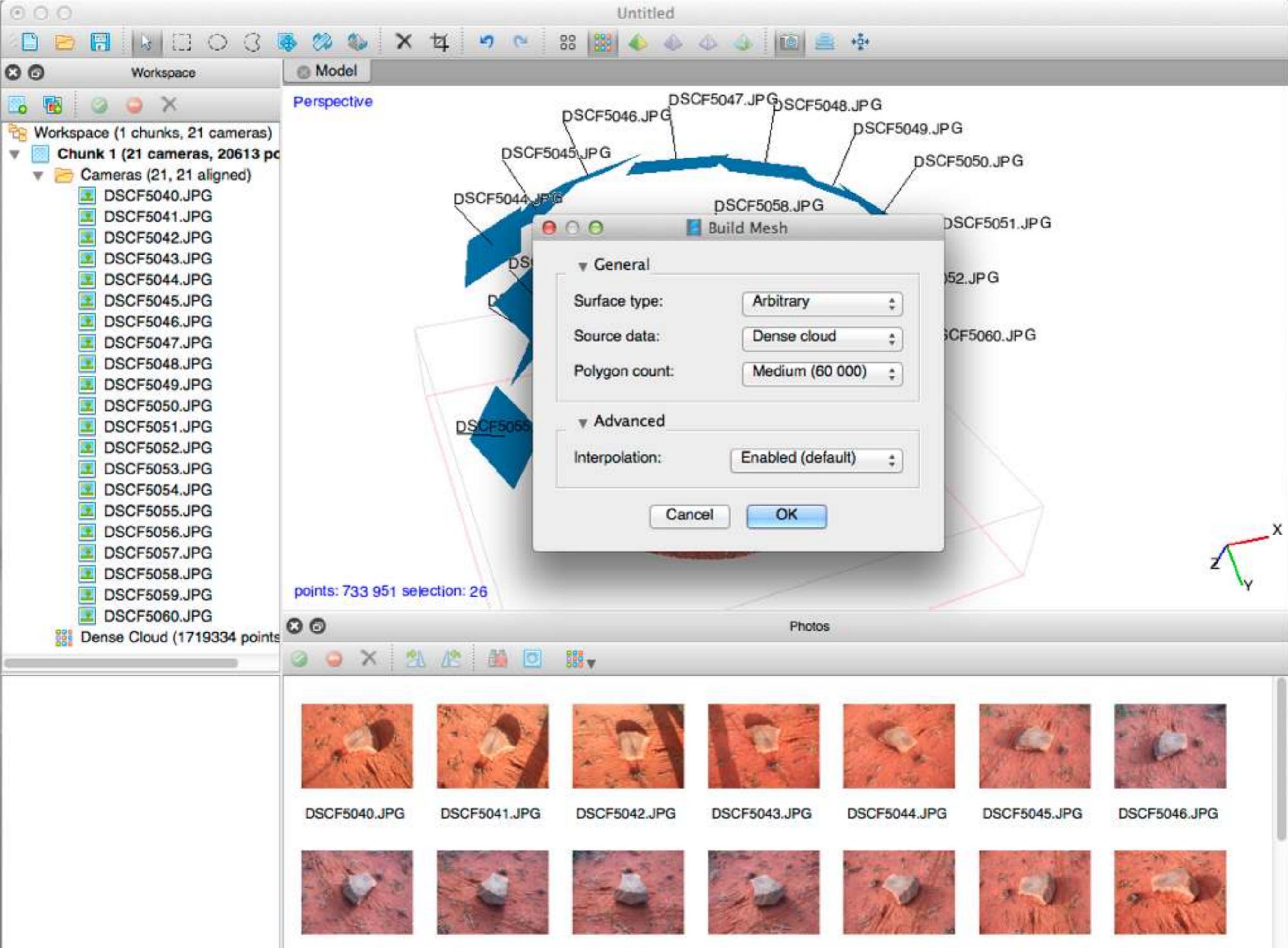
Photos

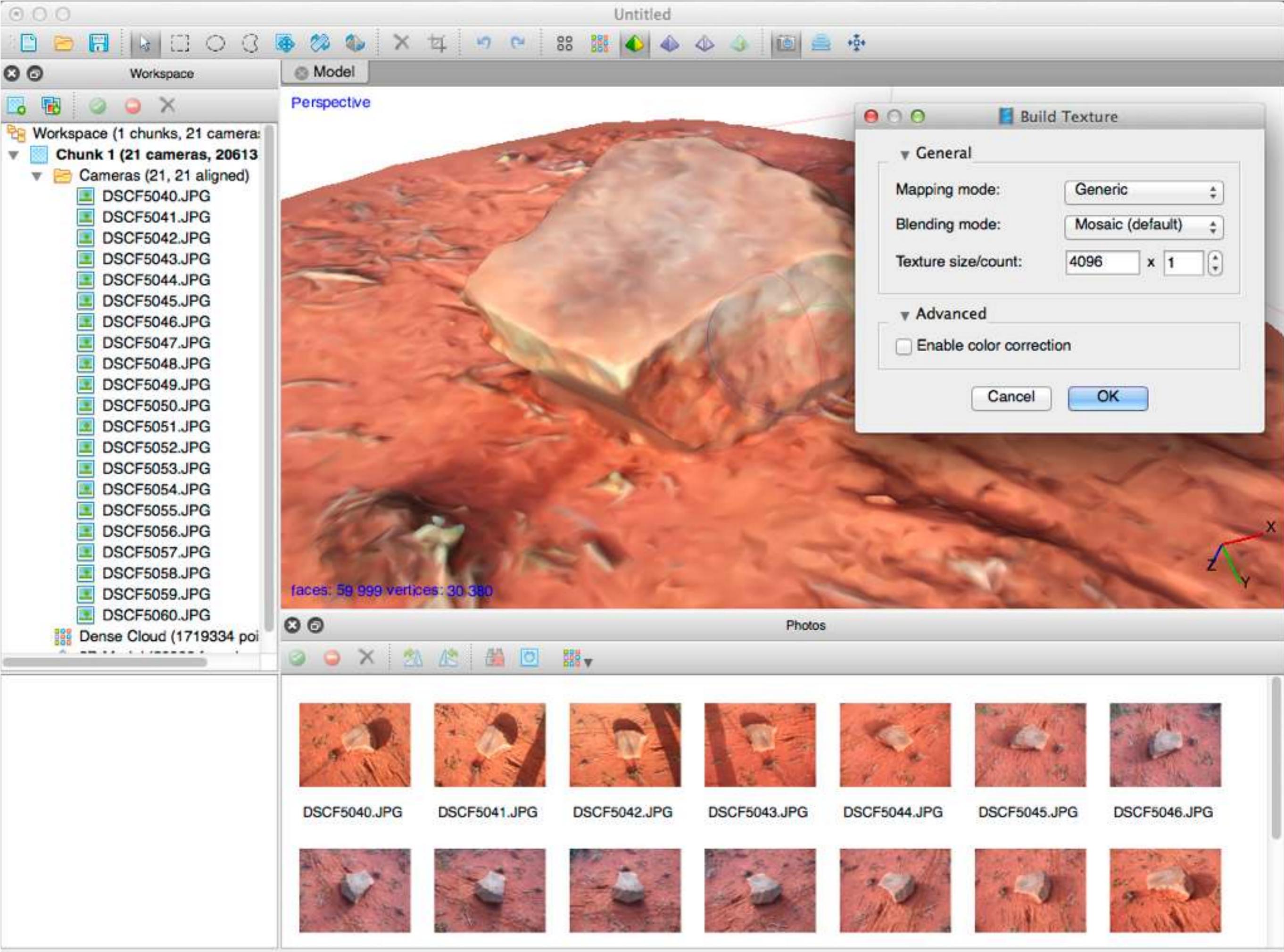
DSCF5040.JPG DSCF5041.JPG DSCF5042.JPG DSCF5043.JPG DSCF5044.JPG DSCF5045.JPG DSCF5046.JPG

DSCF5047.JPG DSCF5048.JPG DSCF5049.JPG DSCF5050.JPG DSCF5051.JPG DSCF5052.JPG DSCF5053.JPG

DSCF5054.JPG DSCF5055.JPG DSCF5056.JPG DSCF5057.JPG DSCF5058.JPG DSCF5059.JPG DSCF5060.JPG

Navigation tool





Untitled

Workspace Model Perspective

Workspace (1 chunks, 21 cameras)

Chunk 1 (21 cameras, 20613)

Cameras (21, 21 aligned)

- DSCF5040.JPG
- DSCF5041.JPG
- DSCF5042.JPG
- DSCF5043.JPG
- DSCF5044.JPG
- DSCF5045.JPG
- DSCF5046.JPG
- DSCF5047.JPG
- DSCF5048.JPG
- DSCF5049.JPG
- DSCF5050.JPG
- DSCF5051.JPG
- DSCF5052.JPG
- DSCF5053.JPG
- DSCF5054.JPG
- DSCF5055.JPG
- DSCF5056.JPG
- DSCF5057.JPG
- DSCF5058.JPG
- DSCF5059.JPG
- DSCF5060.JPG

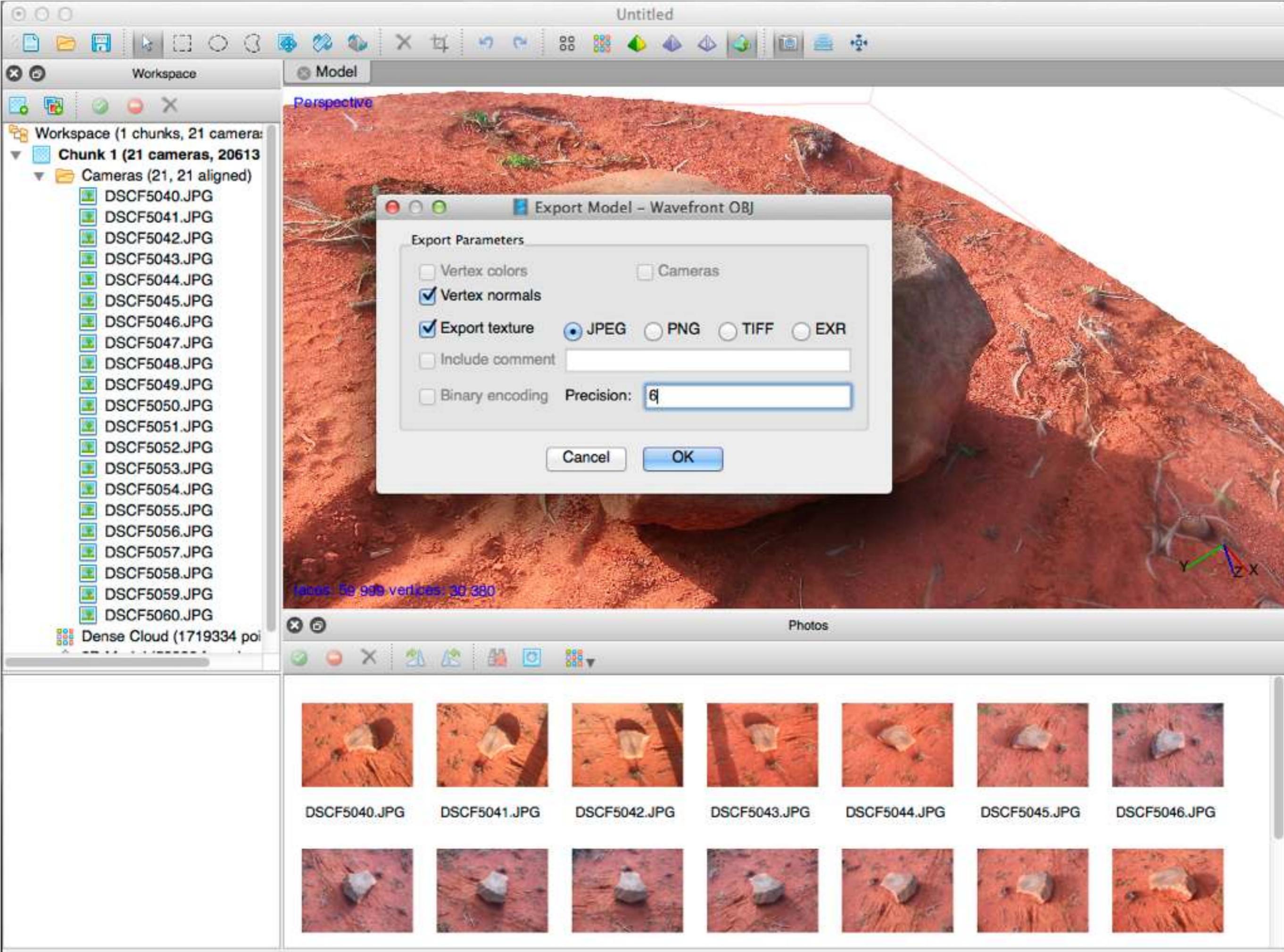
Dense Cloud (1719334 poi)

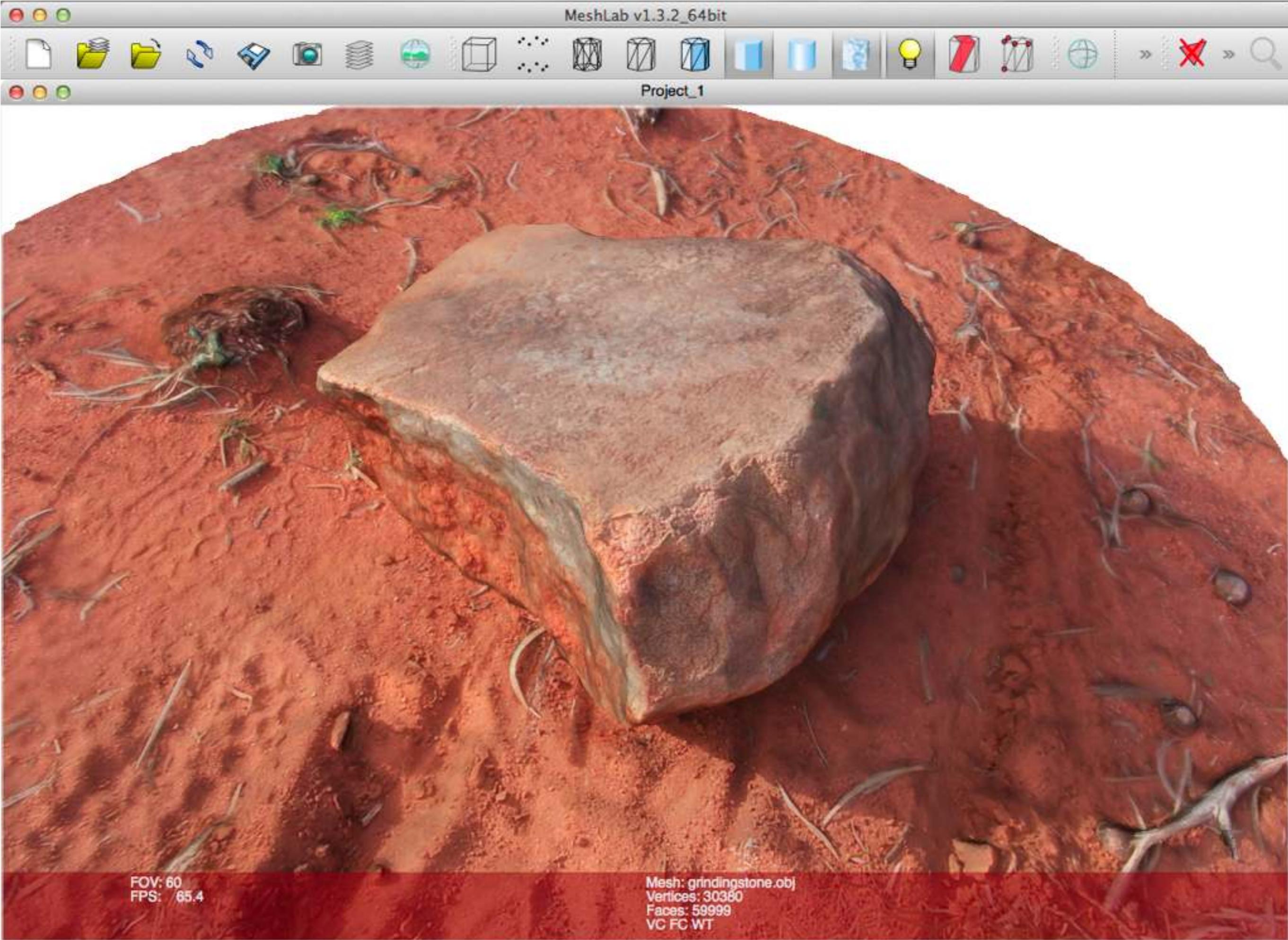
Perspective View: A 3D reconstruction of a large, irregularly shaped rock or object resting on a reddish-brown, textured surface. The model is semi-transparent, showing its internal structure. A coordinate system (X, Y, Z) is visible in the bottom right corner. Text at the bottom left of the view indicates: faces: 59 999 vertices: 39 380.

Photos View: A grid of thumbnail images showing various perspectives of the same scene. The images are labeled below them:

- DSCF5040.JPG
- DSCF5041.JPG
- DSCF5042.JPG
- DSCF5043.JPG
- DSCF5044.JPG
- DSCF5045.JPG
- DSCF5046.JPG

The Photos view also contains two additional rows of thumbnails, though they are not explicitly labeled with file names.





FOV: 60  
FPS: 65.4

Mesh: grindingstone.obj  
Vertices: 30380  
Faces: 59999  
VC FC WT

grindstone

Back View Arrange Share Edit Tags Dropbox Quick Look Action >>

FAVORITES

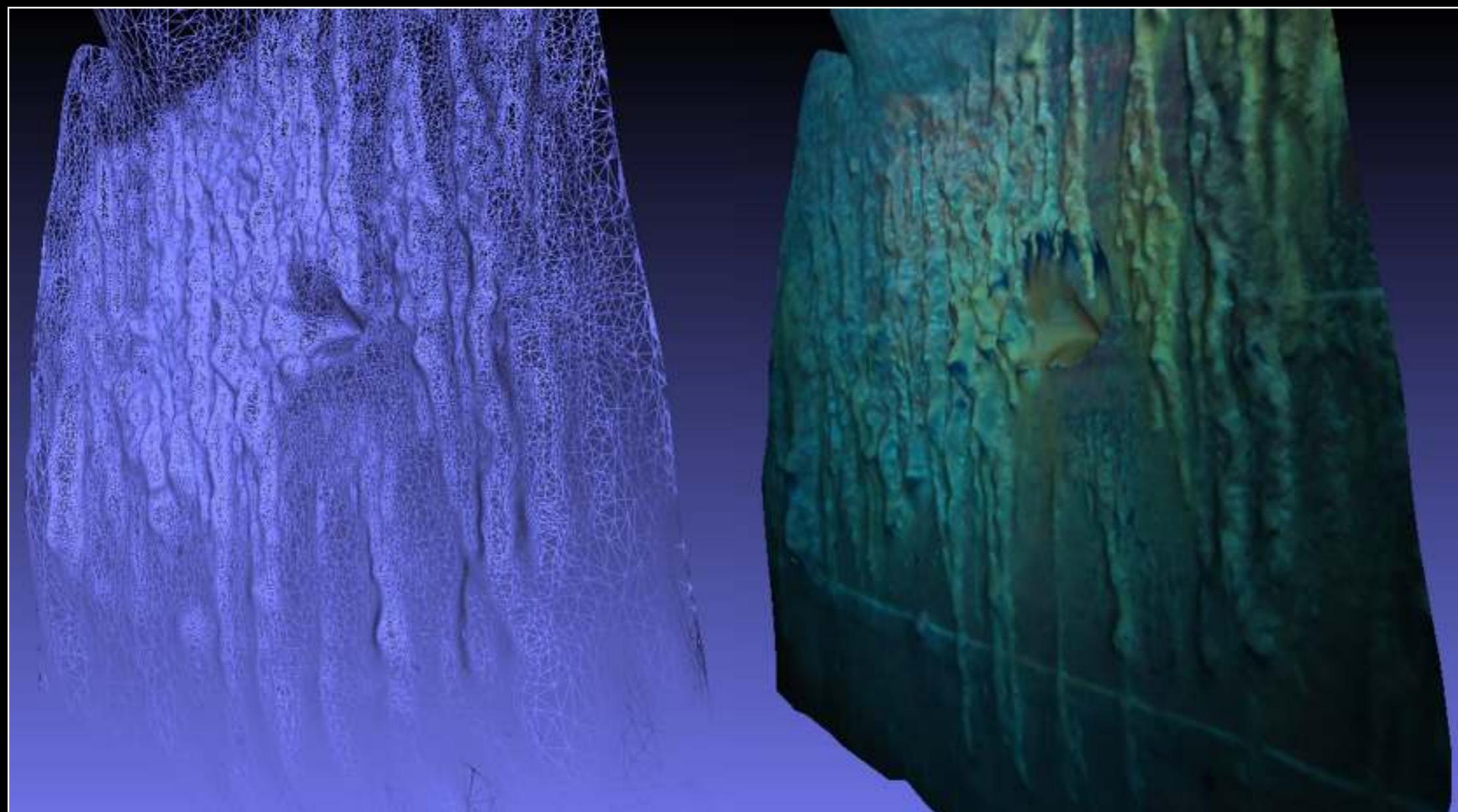
- Dropbox
- All My Files
- AirDrop
- Applications
- pbourke
- Desktop
- Downloads

Name	Date Modified	Size
grinding stone.psz	Today 3:48 pm	37.9 MB
grindingstone.jpg	Today 3:46 pm	4.7 MB
grindingstone.mtl	Today 3:46 pm	203 bytes
grindingstone.obj	Today 3:46 pm	5.7 MB
photos	31 Aug 2013 2:56 pm	149.8 MB



# Additional applications

- Underwater
- Aerial photography
- Rapid Prototypes



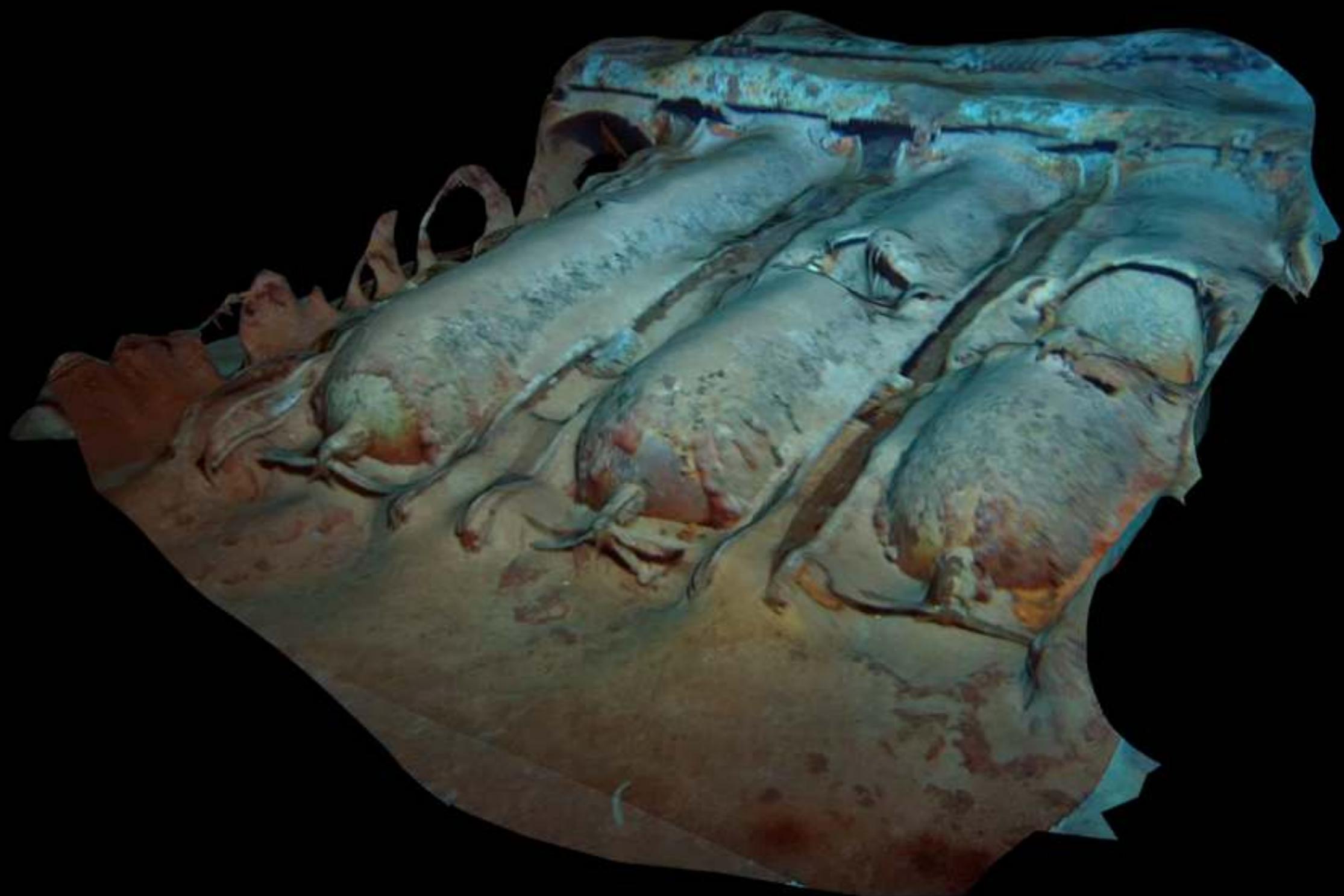
Kormoran

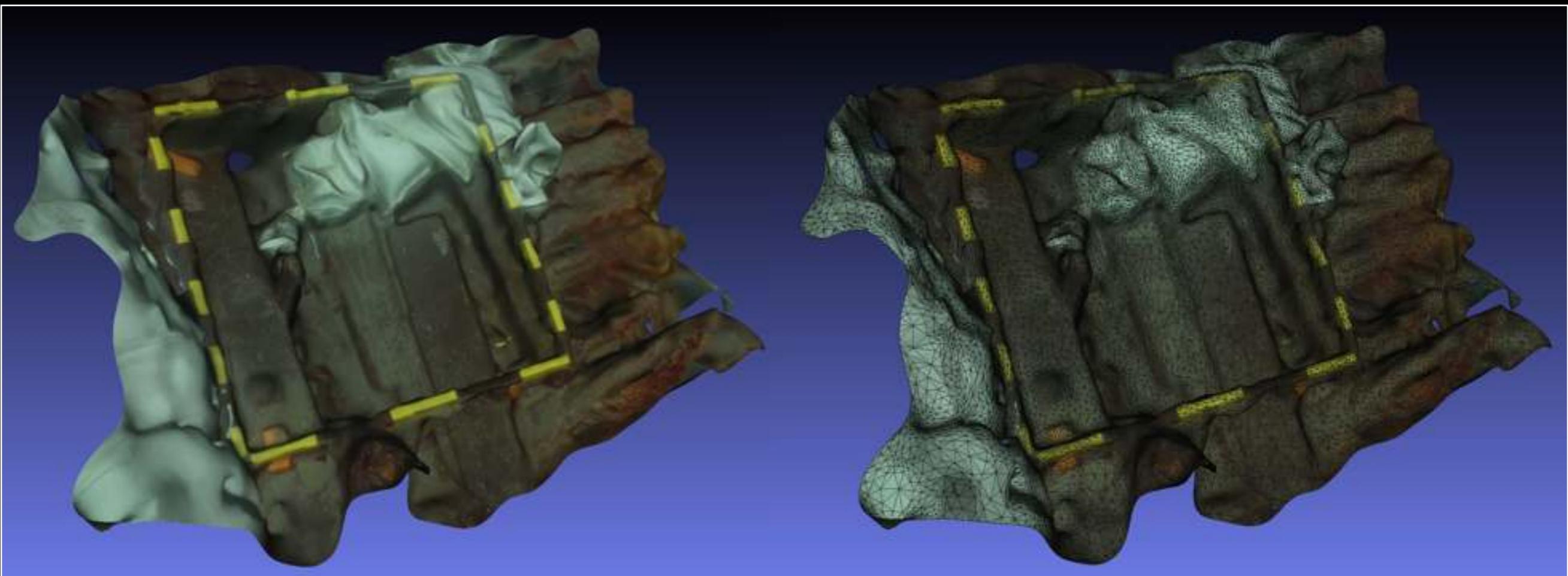
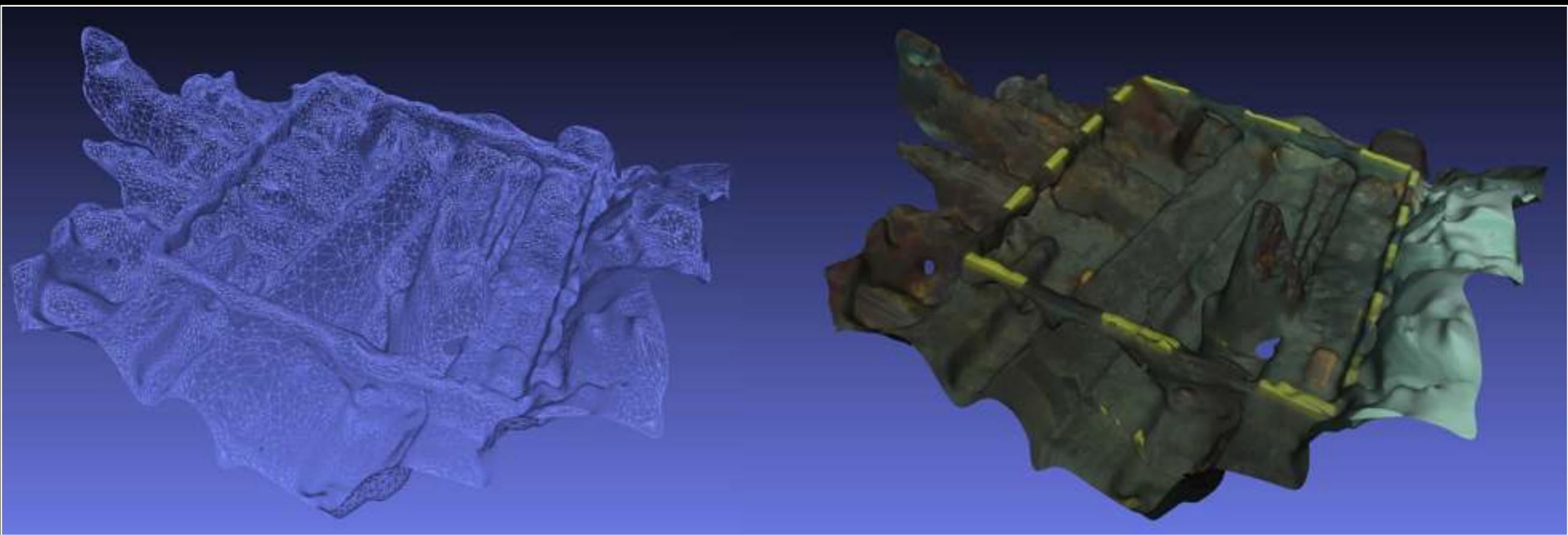
# Additional applications : Underwater

- Capture of underwater object more challenging.
- How to compensate for the light absorption through a column of water.
- Example: HMAS Sydney in 2.5KM of water.



HMAS Sydney





# Additional applications : Aerial photography

- Capturing inaccessible geological formations.
- Also building structures out of reach.
- Vibration and rolling shutter issues.





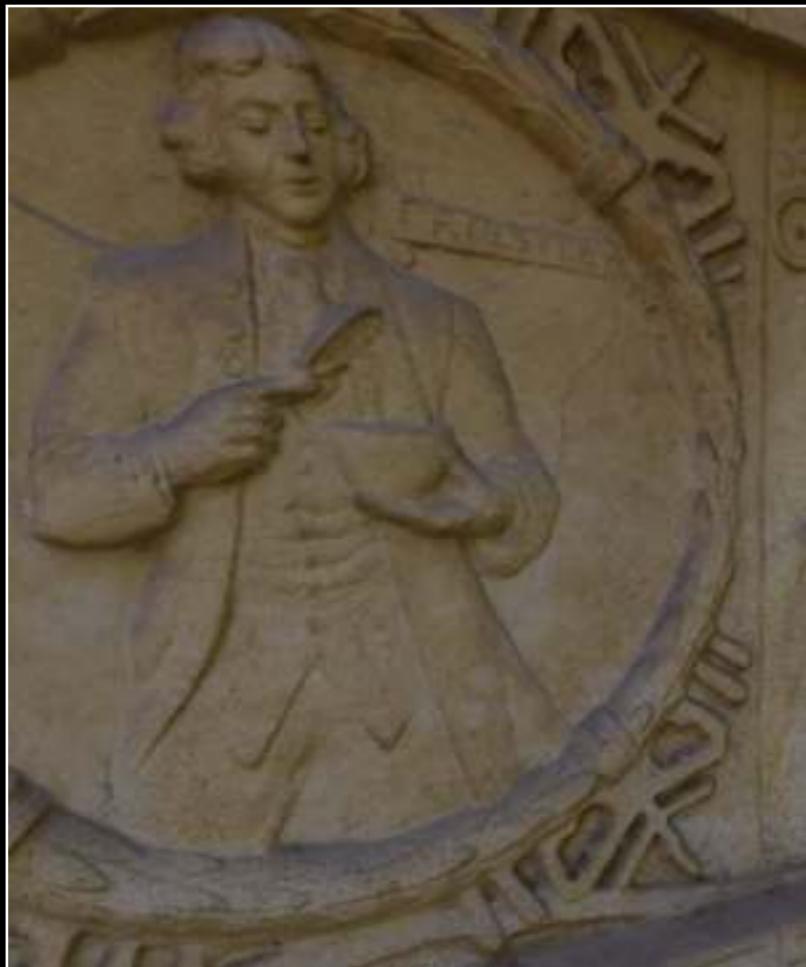
# Additional applications : Rapid prototypes

- Can complete the loop:  
capture a real object photographically - reconstruct it - generate a real object.
- Requires a solid object (thickened), with enough structural integrity.
- Models need to be “watertight”, hence hole closing algorithms.

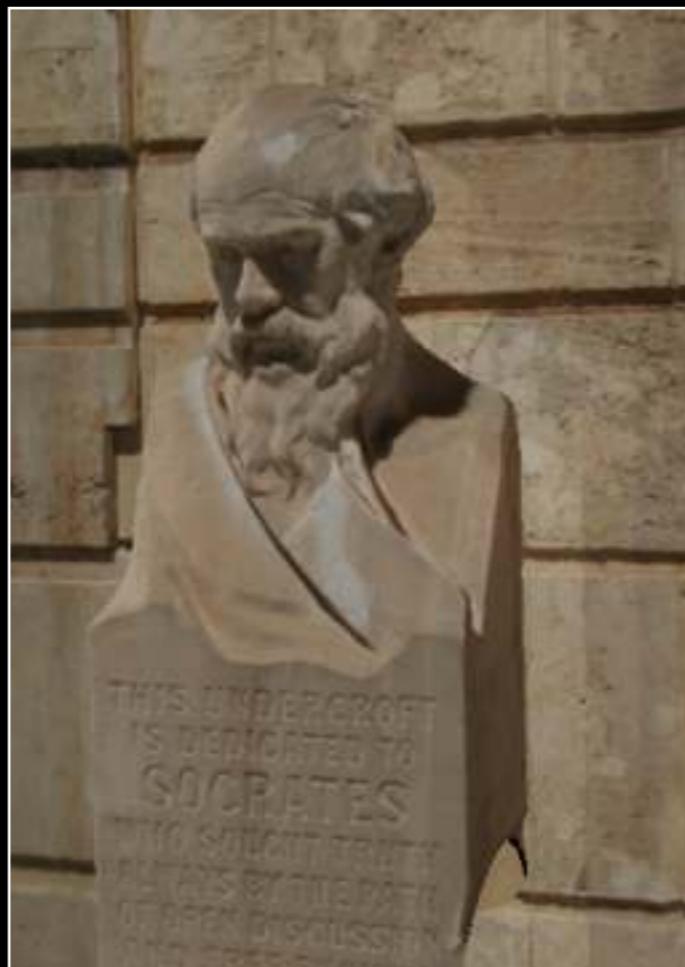


# Advanced topics : Reference images

- Creating a library of reference photographs by which to compare algorithms.
- Not just “good” photographic sets but also poorly captured photographs.
- Distinguish three categories: 2.5D, 3D simple object (topological sphere), 3D complex object.



2.5D



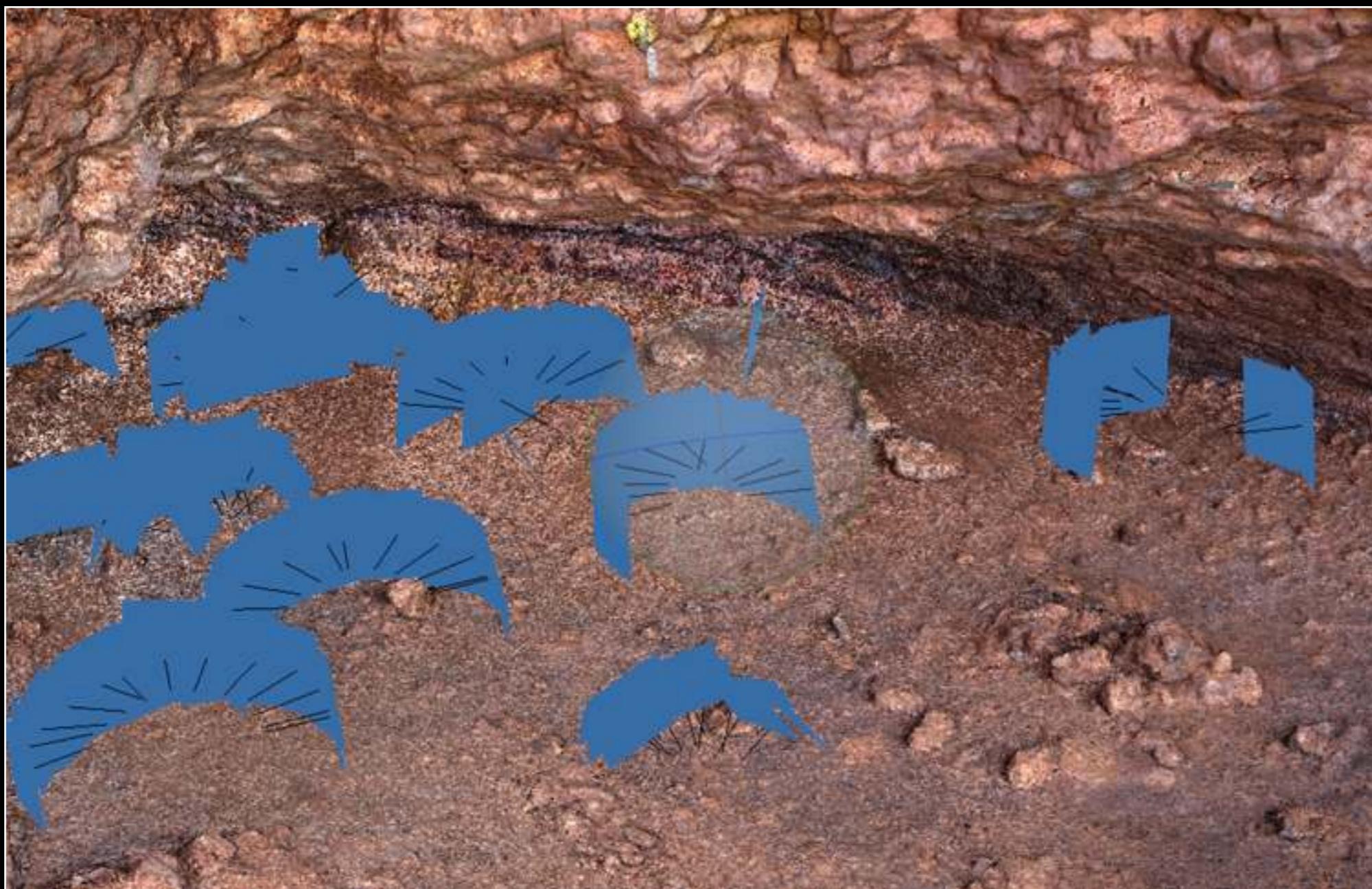
topological sphere



topological N-sphere ( $N > 2$ )

# Advanced topics : Interior spaces

- These can be difficult due to the lack of space to move for different camera shots.
- Multiple shots can be taken from a single position in order to get coverage, but they should be viewed a single perspective view.



# Advanced topics : Multi-resolution

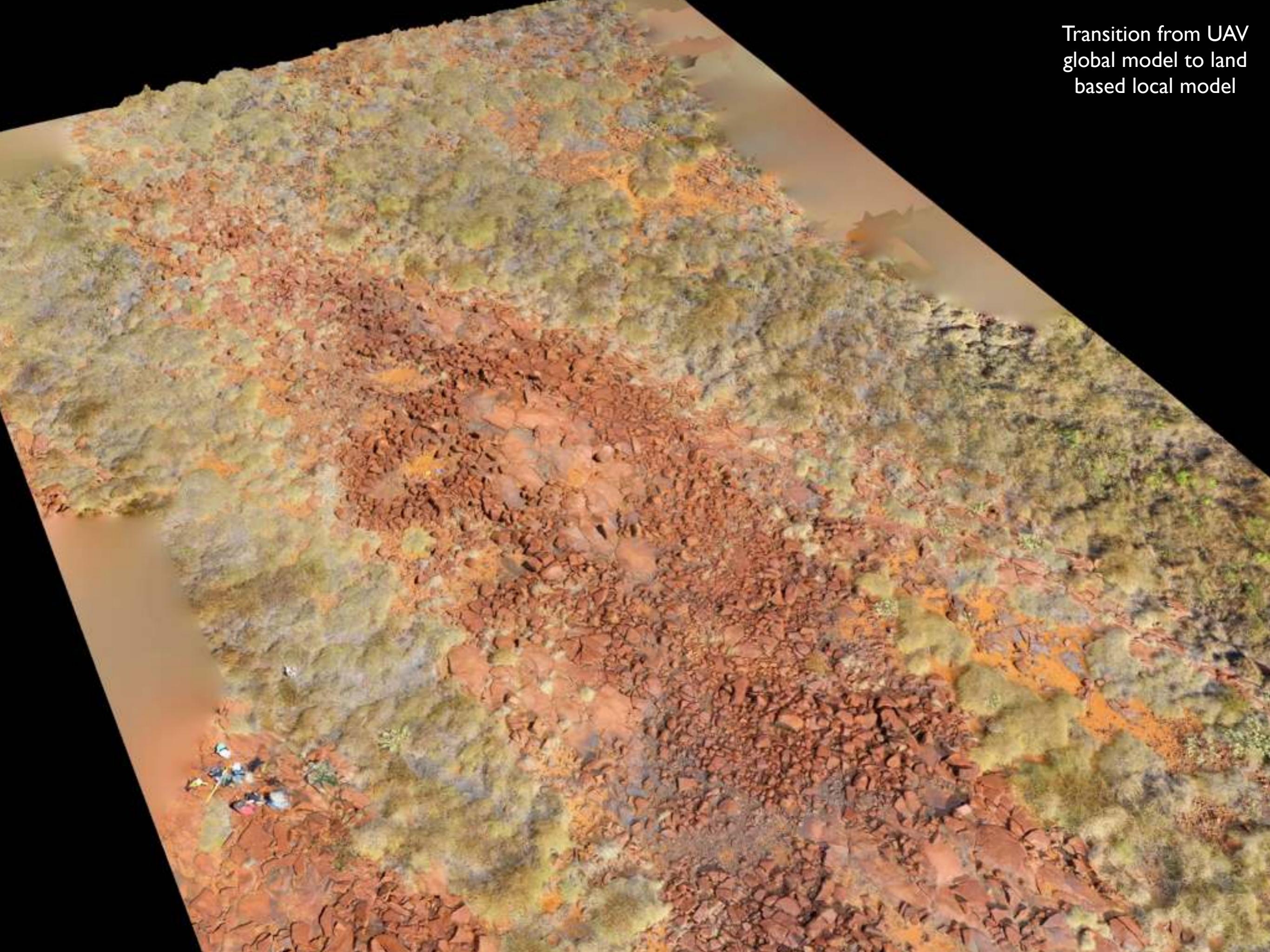
- Similar to multi-resolution or hierarchical images, can we do the same for 3D meshes.
- Large scale reconstruction at one resolution, smaller higher resolution parts of interest.







Transition from UAV  
global model to land  
based local model



# Reading / references

- Barazzetti, L., Scaioni, M. and Remondino, F. (2010), Orientation and 3D modelling from markerless terrestrial images: combining accuracy with automation. *The Photogrammetric Record*, 25: 356–381. doi: 10.1111/j.1477-9730.2010.00599.x
- Percoco, G. (2011), Digital close range photogrammetry for 3D body scanning for custom-made garments. *The Photogrammetric Record*, 26: 73–90. doi: 10.1111/j.1477-9730.2010.00605.x
- Remondino, F., Rizzi, A., Girardi, S., Petti, F. M. and Avanzini, M. (2010), 3D Ichnology—recovering digital 3D models of dinosaur footprints. *The Photogrammetric Record*, 25: 266–282. doi: 10.1111/j.1477-9730.2010.00587.x
- Bill Jeffery. (2006), From Seabed to Computer Screen—digital mapping of submerged and shipwreck sites. *Bulletin of the Australian Institute for Maritime Archaeology*, 23:86-94
- Bryan, P. G. and Clowes, M. (1997), Surveying Stonehenge By Photogrammetry. *The Photogrammetric Record*, 15: 739–751. doi: 10.1111/0031-868X.00082
- Barazzetti, L., Remondino, F. and Scaioni, M. (2009) Combined use of photogrammetric and computater vision techniques for fully automated and accurate 3D modelling of terrestrial objects. SPIE Optics+Photonics, 7447, 2-3 August, San Diego, CA, USA.
- Barthelsen, J., Mayer, H., Hirschmuller, H., Kuhn, A, Michelini, M. 92012) orientation and dense reconstruction from bordered wide baseline image sets. PFG 2012.
- Besl, P. McKay, N. (1992) A method for registration of 3D shapes. *IEEE Transactions on pattern analyse and machine intelligence*. PAMI 14 (2).
- Cignoni, P., Callieri, M., Corsini, M., Dellepaine, M., Ganovelli, F., Ranzuglia, G., (2008). MeshLab: an opensource mesh processing tool. *Eurographics Italian Chapter Conference*, The Eurographics Association, 129-136.
- C. Wu, 2011. "VisualSfM: A Visual Structure from Motion System", <http://homes.cs.washington.edu/~ccwu/vsfm/> (31 Jan. 2013)
- Courchay, J., Pons, J.P., Monasse, P., Keriven, R., (2010). Dense and accurate spatio-temporal multiview stereovision. *Computer Vision ACCV 2009*, Lecture notes in computer Science, 5995, 11-22

# Reading / references

- Noah Snavely, Steven M. Seitz, Richard Szeliski. Photo Tourism: Exploring image collections in 3D. ACM Transactions on Graphics (Proceedings of SIGGRAPH 2006), 2006.
- Noah Snavely, Steven M. Seitz, Richard Szeliski. Modeling the World from Internet Photo Collections. International Journal of Computer Vision, 2007.
- M. Favalli n, A. Fornaciai, I. Isola, S. Tarquini, L. Nannipieri. Multiview 3D reconstruction in geosciences. Computers & Geosciences, 44 (2012) 168–176
- Besl, P.J., McKay, N.D., 1992. A method for registration of 3-D shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence 14, 239–256.
- de Matí ´as, J., de Sanjose ´, J.J., Lo ´pez-Nicola ´s, G., Sagu ´e ´s, C., Guerrero, J.J., 2009. Photogrammetric methodology for the production of geomorphologic maps: application to the Veleta Rock Glacier (Sierra Nevada, Granada, Spain). Remote Sensing 1, 829–841.
- Dowling, T.I., Read, A.M., Gallant, J.C., 2009. Very high resolution DEM acquisition at low cost using a digital camera and free software. In: Proceedings of the 18th World IMACS/MODSIM Congress, Cairns, Australia.
- Furukawa, Y., Ponce, J., 2007. Accurate, dense, and robust multi-view stereopsis. In: Proceedings, IEEE Conference on Computer Vision and Pattern Recognition CVPR 2007, pp. 1–8.
- Furukawa, Y., Ponce, J., 2009. Accurate camera calibration from multi-view stereo and bundle adjustment. International Journal of Computer Vision 84, 257–268.
- Hartley, R.I., Zisserman, A., 2004. Multiple View Geometry in Computer Vision. Cambridge University Press.
- Levoy, M., Pulli, K., Curless, B., Rusinkiewicz, S., Koller, D., Pereira, L., Ginzton, M., Anderson, S., Davis, J., Ginsberg, J., Shade, J., Fulk, D., 2000. The Digital Miche- langelo Project: 3D scanning of large statues. Computer Graphics (SIGGRAPH 2000 Proceedings).

# Reading / references

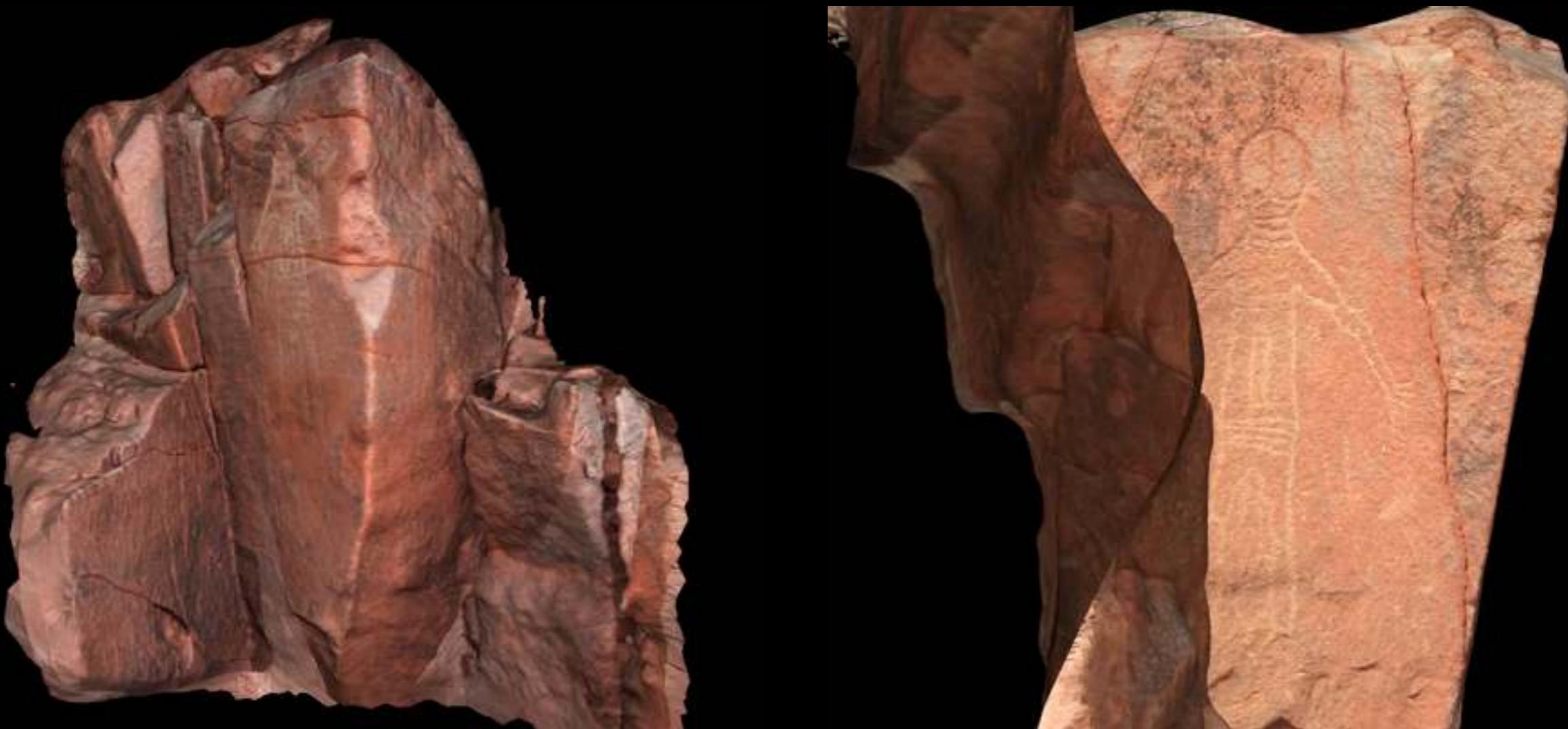
- Lourakis, M., Argyros, A., 2008. SBA: A generic sparse bundle adjustment C/C++ package based on the Levenberg-Marquardt algorithm. <http://www.ics.forth.gr/lourakis/sbaS>.
- Mikhail, E.M., Bethel, J.S., McGlone, J.C., 2001. Introduction to Modern Photogrammetry. John Wiley & Sons, Inc., New York.
- Nister, D., 2004. Automatic passive recovery of 3D from images and video. In: Proceeding of the Second IEEE International Symposium on 3D Data Processing, Visualization and Transmission, pp. 438–445.
- Snavely, N., Seitz, D., Szeliski, R., 2007. Modeling the world from internet photo collections. International Journal of Computer Vision 80, 189–210.
- Triggs, B., McLauchlan, P., Hartley, R., Fitzgibbon, A., 2000. Bundle adjustment—A modern synthesis. in: Triggs, W., Zisserman, A., Szeliski, R. (Eds.), Vision Algorithms: Theory and Practice, LNCS, Springer-Verlag, pp. 298–375.
- T. P. Kersten and M. Lindstaedt, Automatic 3D Object Reconstruction from Multiple Images for Architectural, Cultural Heritage and Archaeological Applications Using Open-Source Software and Web Services. Photogrammetrie - Fernerkundung - Geoinformation, Heft 6, pp. 727-740.
- S. Lowe. 2004. Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision. 60 (2), pp 91-110.
- Agarwal, N. Snavely, I. Simon, S. M. Seitz and R. Szeliski, Building Rome in a Day. International Conference on Computer Vision, 2009, Kyoto, Japan.
- M. Jancosek and T. Pajdla, 2011. Multi-View Reconstruction Preserving Weakly-Supported Surfaces, IEEE Conference on Computer Vision and Pattern Recognition 2011
- M. Hanif and A. Seghouane, 2012. Blurred Image Deconvolution Using Gaussian Scale Mixtures Model in Wavelet Domain. IEEE 2012 International Conference on Digital Image Computing Techniques and Applications

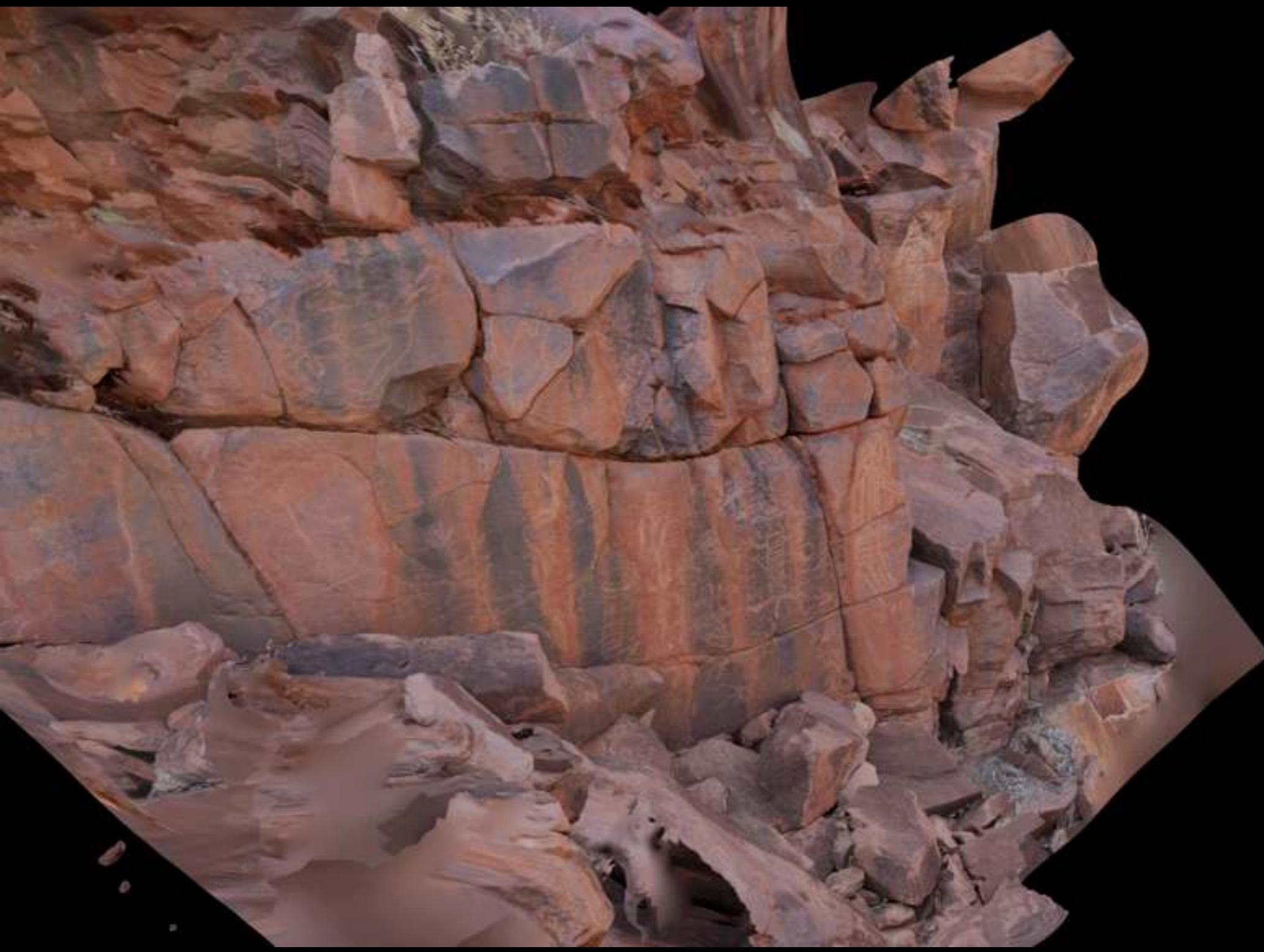
# Summary for high quality reconstruction

- High quality SLR camera (and know how to use it)
- Good quality prime lens
- Perform lens calibration for wide angle or otherwise extreme lenses
- Err on the side of taking more images
- Strong linear reference objects in shot can assist reconstruction
- Select best software currently on the market  
(PhotoScan is hard to beat at time of writing)
- Results benefit from crisp high resolution photographs  
Not particularly sensitive to colour detail

# Summary : An exciting time to be in this game

- Algorithms are improving at a steady pace.
- Our expectations are steadily rising.





# Questions / discussion

The new digital tourist?

