3D model reconstruction from photographs

Paul Bourke

http://paulbourke.net/SriLanka

Contents

- Why
- When
- How
  - Technology
  - Photography
  - Worked example
  - Your turn

Photoscan
Meshlab
Sketchfab

Outcomes

- Familiarity with the state of the technology.
- Understand the terminology.
- Familiarity with the software and tools.
- Some expectations of the limitations.

Why

- Creating richer objects (compared to photographs) for recordings in archaeology and heritage.

- Create geometric models suitable for analysis, eg: in geology or geoscience.

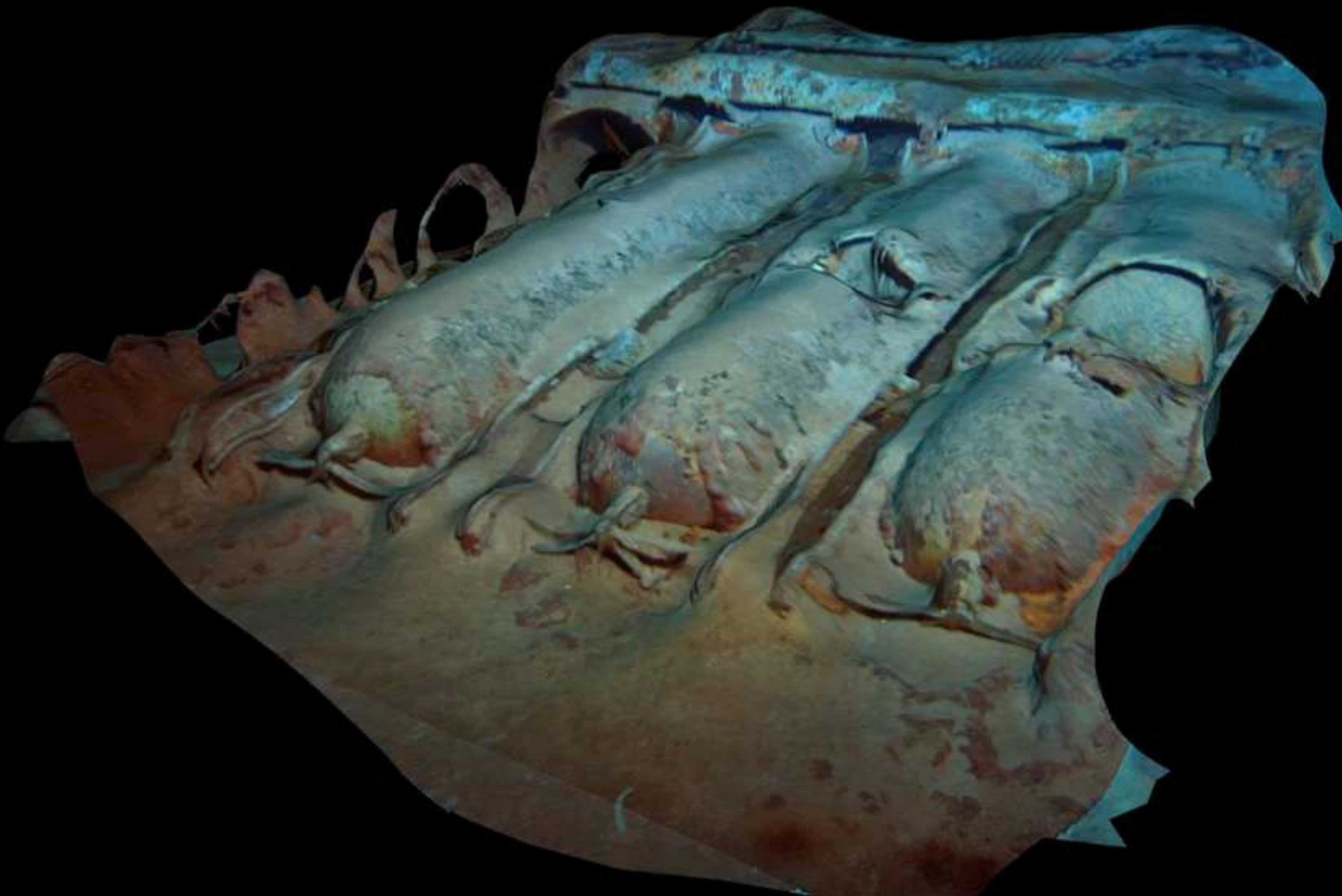
- Create digital assets for virtual environments.

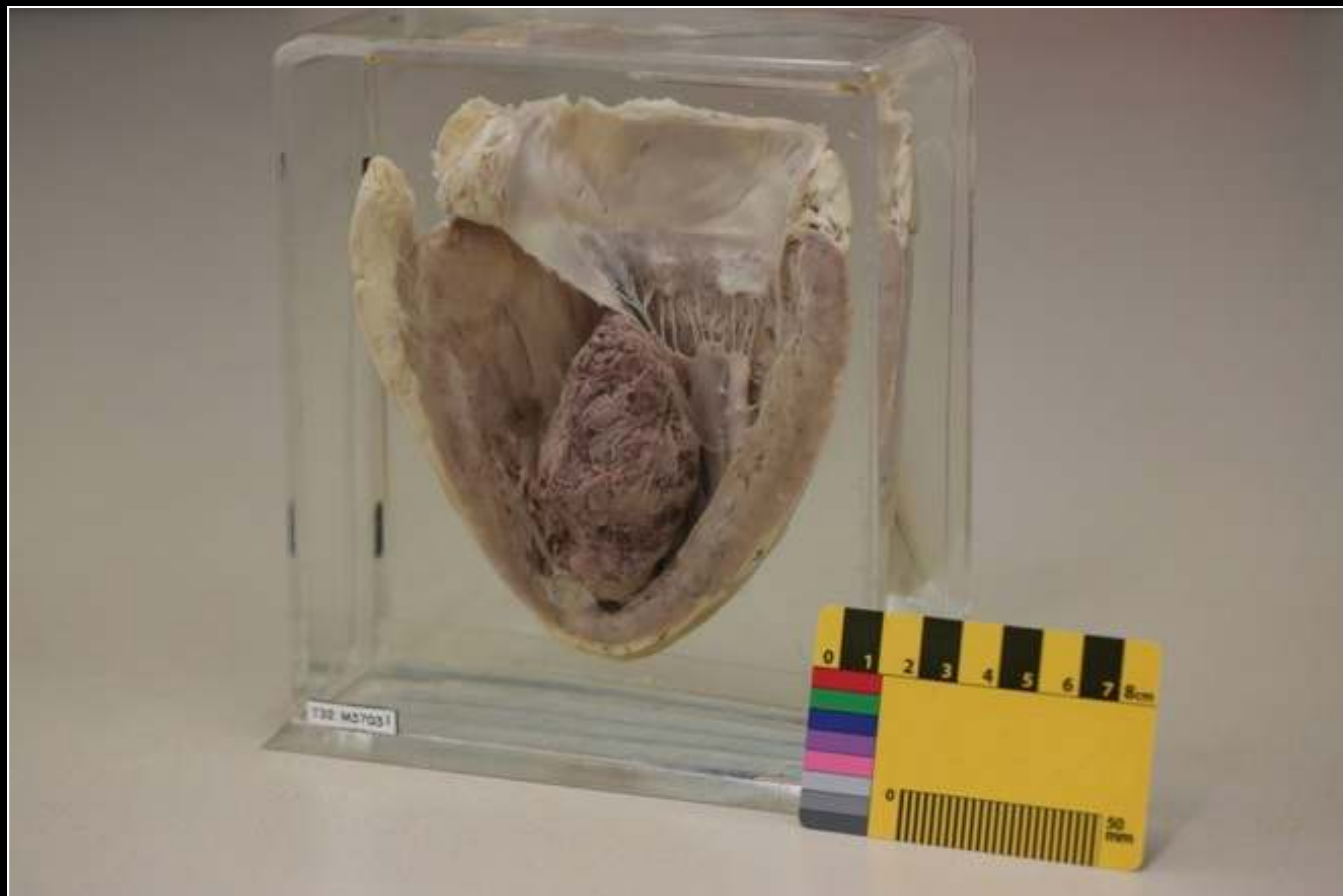
- What can be done with a 3D model is often more interesting than what can be done with a 2D image.

Weld Range

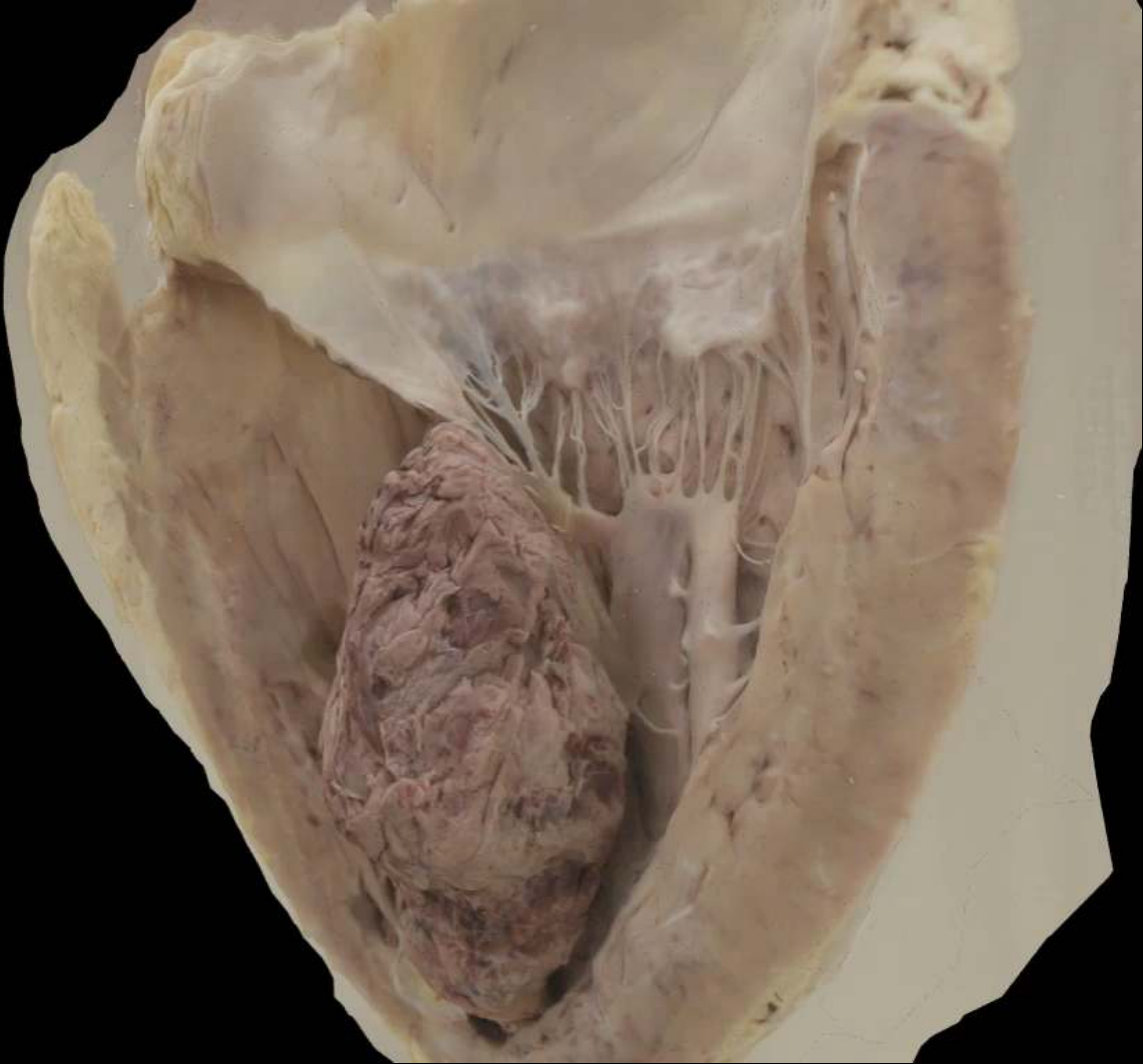
Beacon Island

Beacon Island Virtual environment





Pathology, QE II



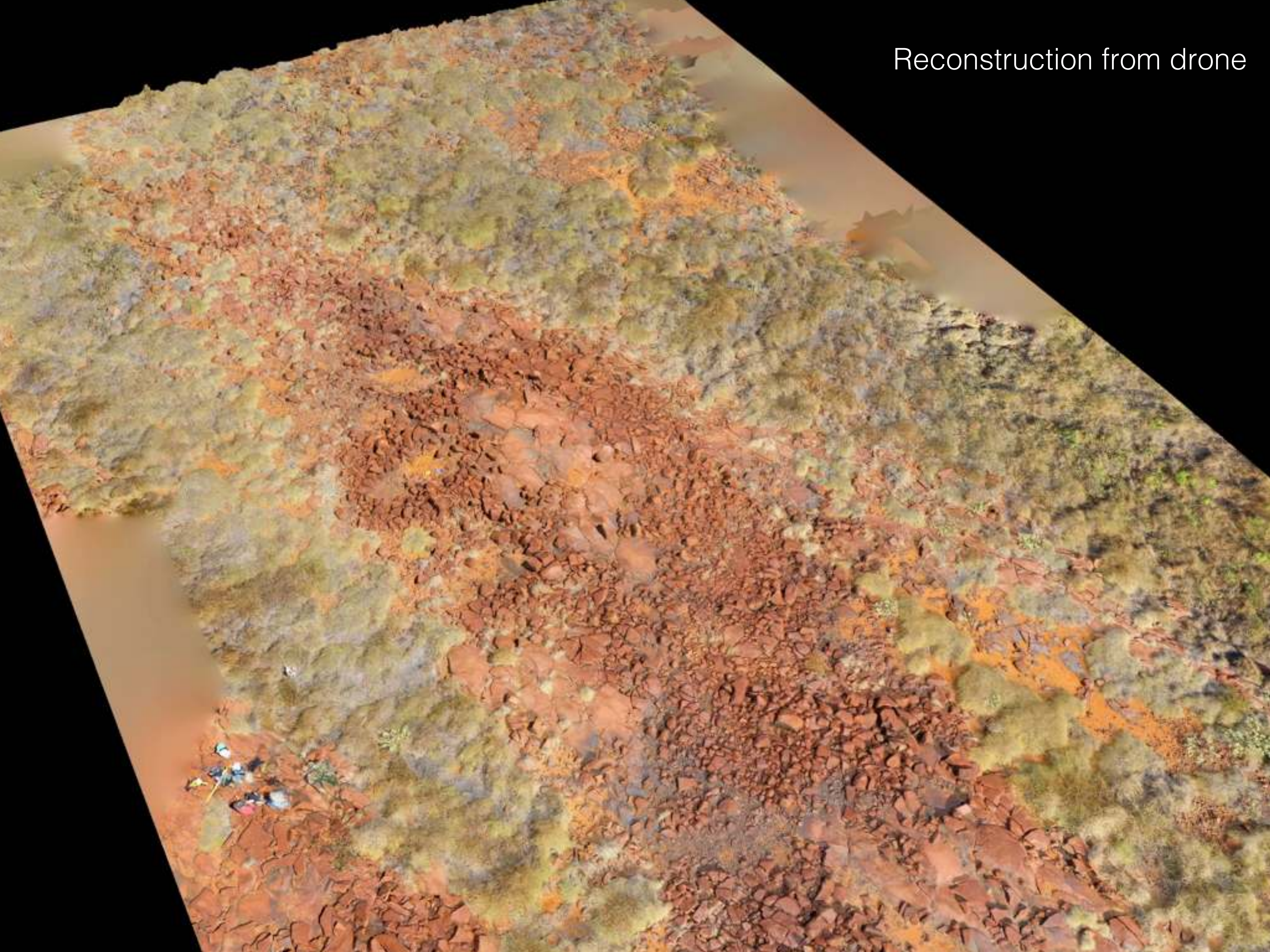
Department of Mines and Petroleum

Centre for Exploration Targeting, UWA

Remote in-the-field capture

Ngintaka headdress, South Australia Museum

Reconstruction from drone



When

- Best suited to organic shapes that are otherwise hard to model.
- Provides an unbiased model compared to human modeller.
- Some models the technique is less suited to
  - Plain flat surfaces
  - Highly reflective or mirror surfaces

History

- Photogrammetry is the general term given to deriving 3D 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

LIDAR

- 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 investment
  Questions 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.

Structured light

- Light field cameras (plenoptic camera).
  - Captures an array of images from a grid of positions
  - Currently resolution is too low.

Software

- Processing pipeline from a number of opensource projects
- SiroVision
- PhotoScan
- PhotoSynth
- PhotoModeller Scanner
- 123D Catch
- ReMake (was Memento)
- Apero
- AdamTech solution
- iWitness Pro
- .... lots of others.
  Major aspect of the work we are doing is to evaluate the various tools.

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.
- 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 a potentially a large number of variables.
  - Trade off between simplicity and control.
- Ability to create high resolution textures, larger than 4Kx4K, or multiple textures.

Typical pipeline

SIFT for key point extraction between images

↓

SfM software package Bundler to generate a sparse 3D point cloud and estimate camera poses.

↓

PMVS2 (Patch-based Multiview Stereo software) to reconstruct the model of the imaged scene as dense point cloud

↓

Convert point cloud to mesh

↓

Re-project textures

Sparse point cloud

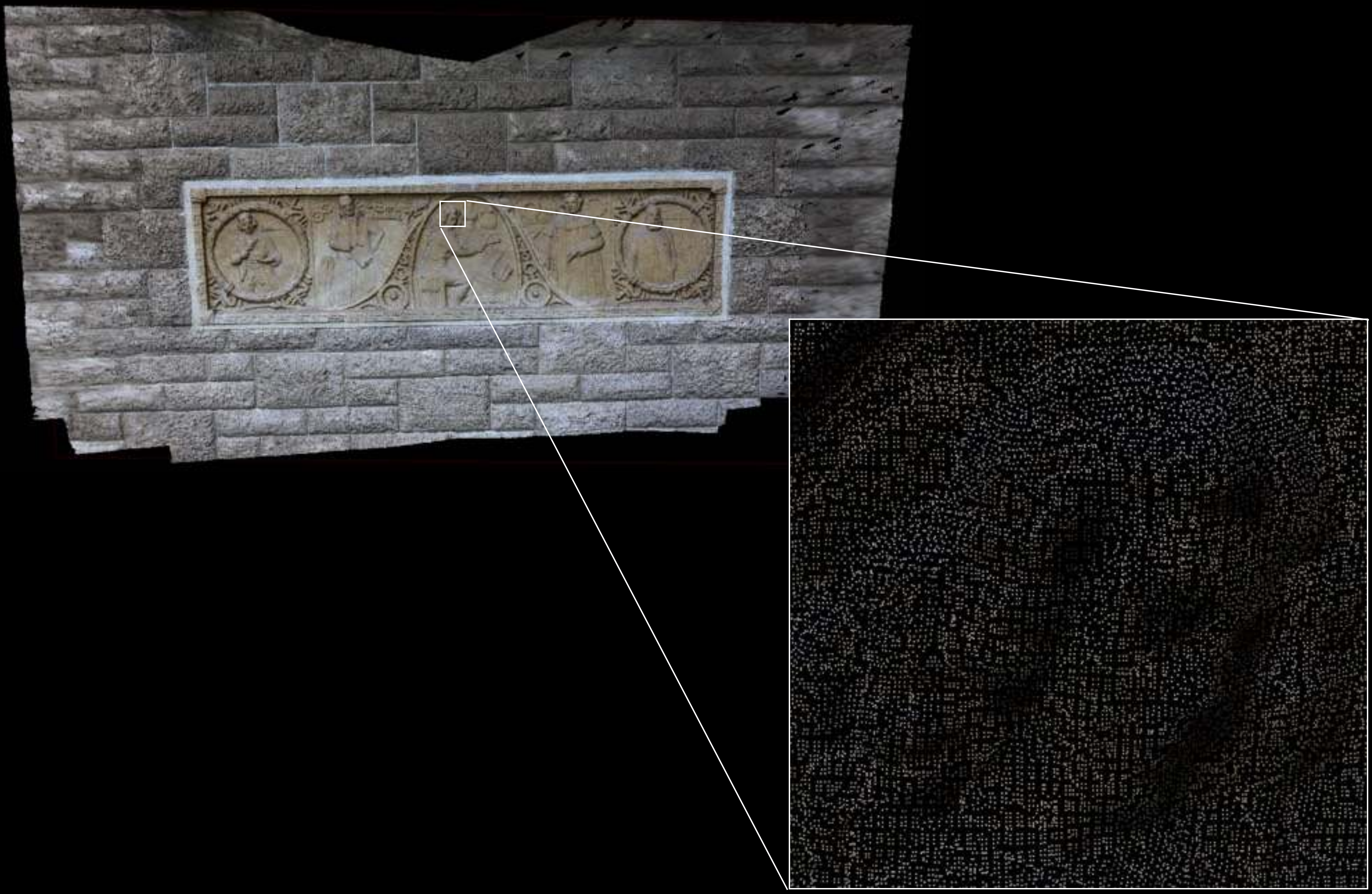
- Find matching points between photographs, feature point detection. SIFT - scale invariant feature transform
- Compute camera positions and other intrinsic camera parameters. Bundler, SfM - Structure from Motion

Dense point cloud

- CMVS - Clustering Views for Multi-view Stereo.

UWA Geography Building

Dense point cloud



Mesh generation

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

Texture mesh

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

Export

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.

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.

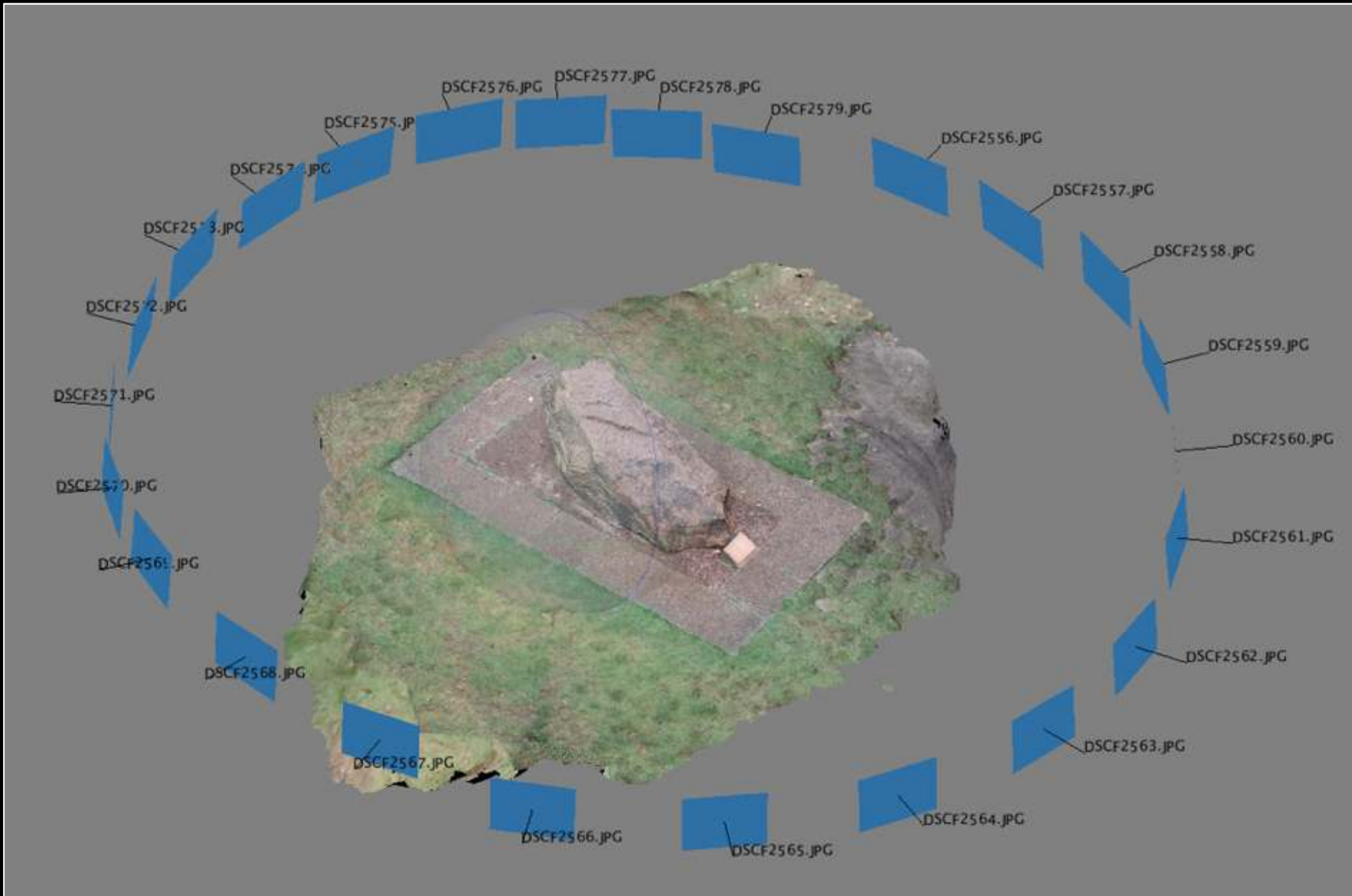
Sigma 28mm, f1.8

Sigma 50mm, f1.4

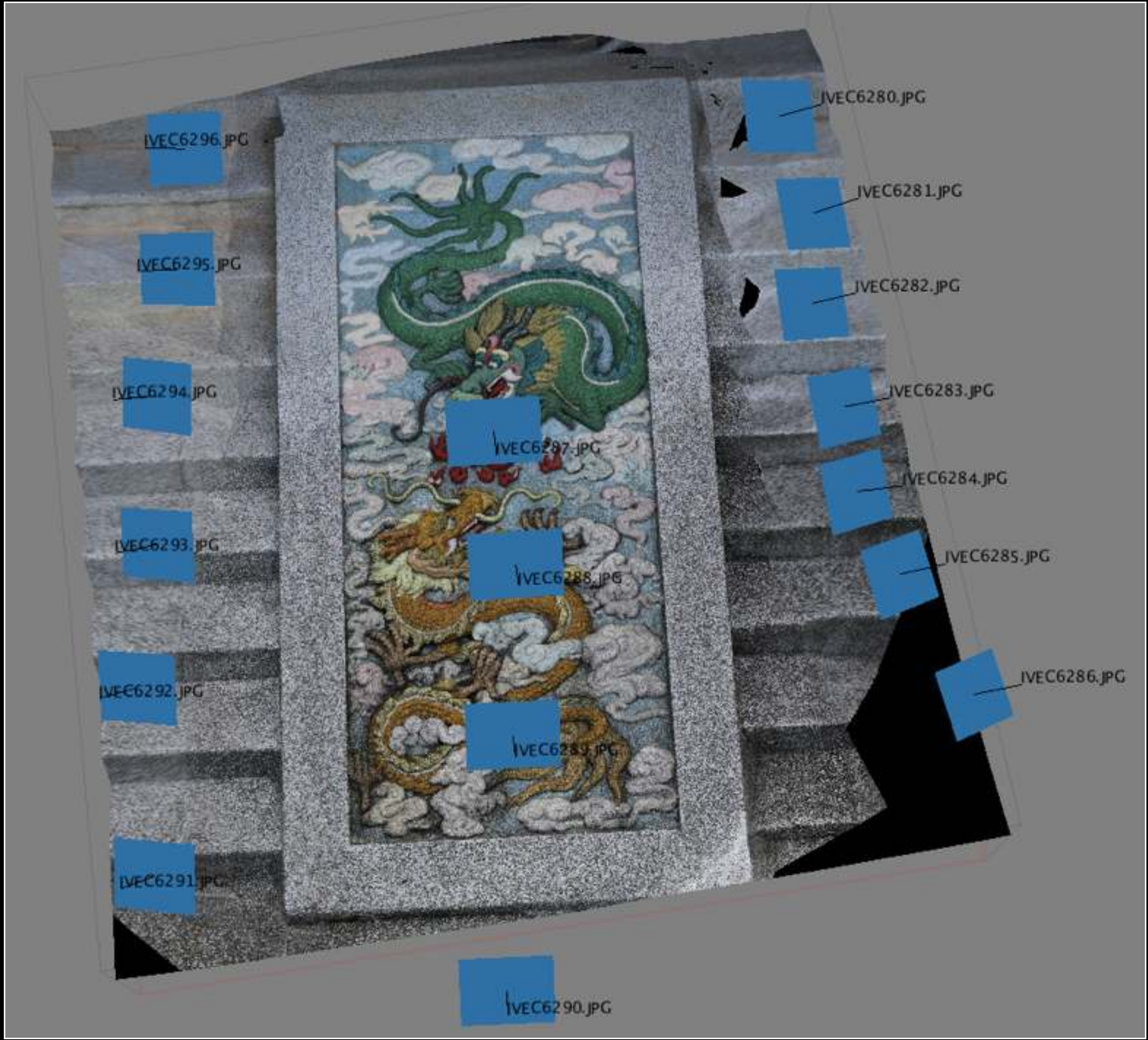
Photography : shooting guide

- 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 to1/2 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.

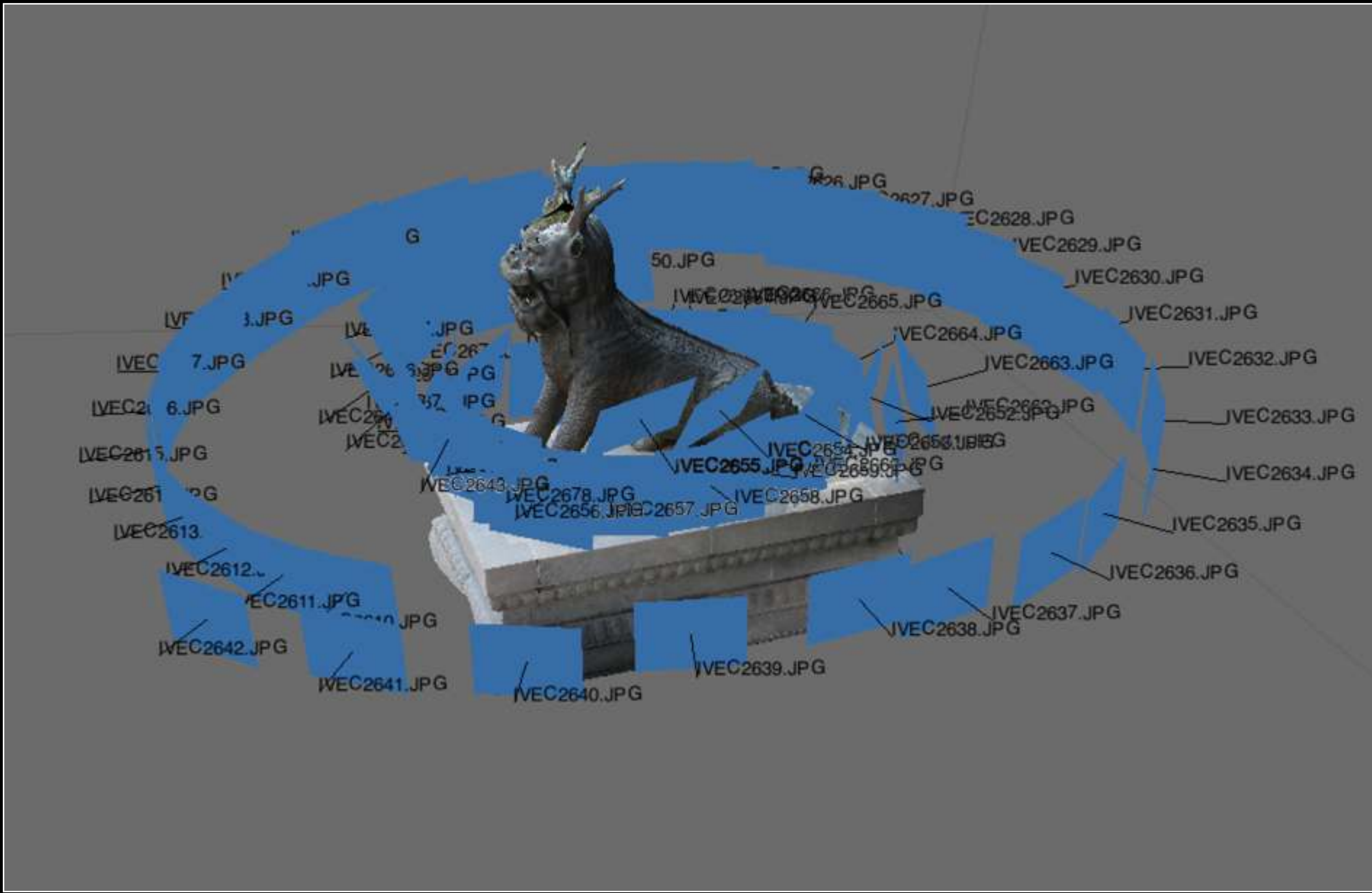
Photography



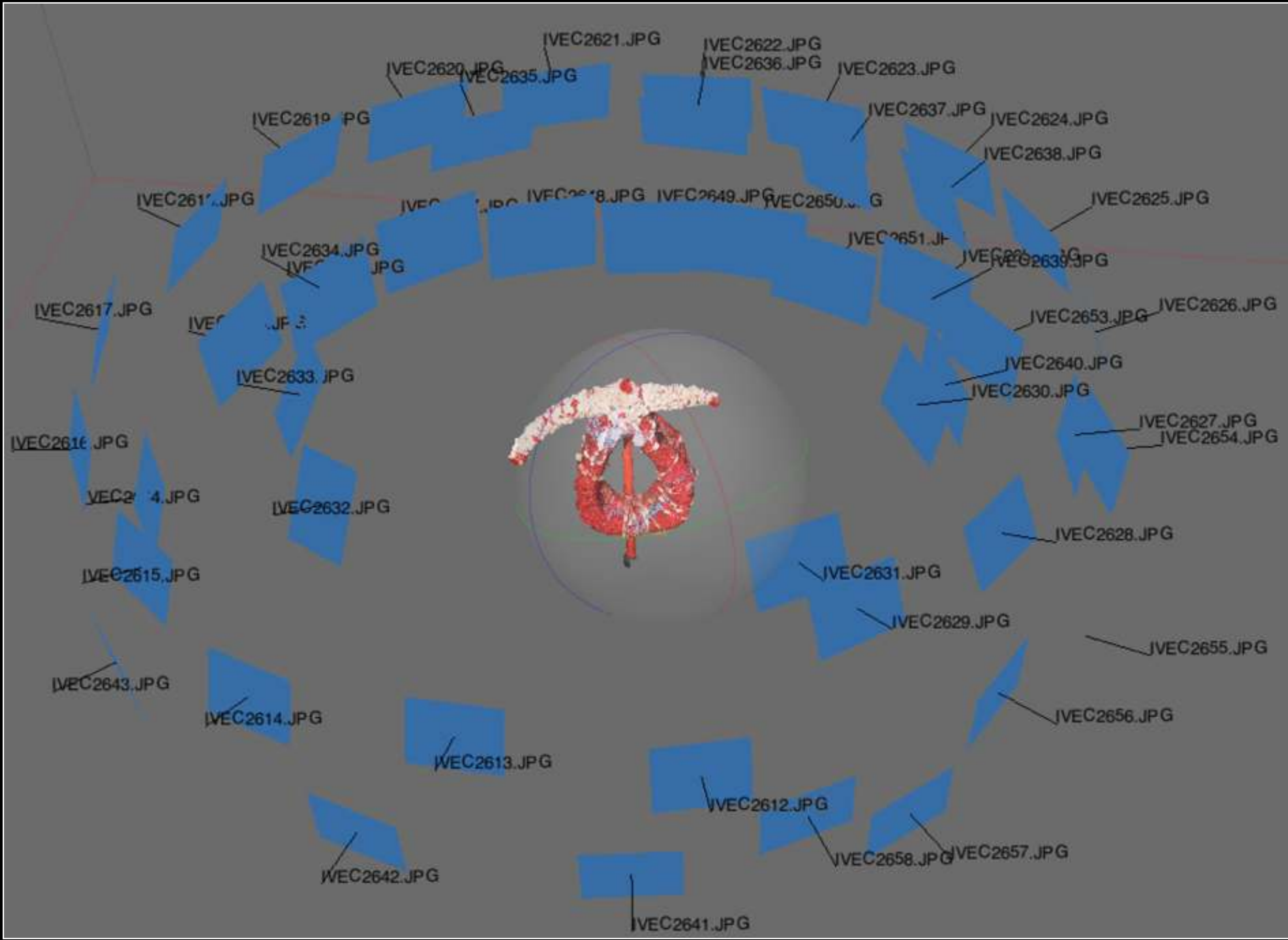
Manipal, India



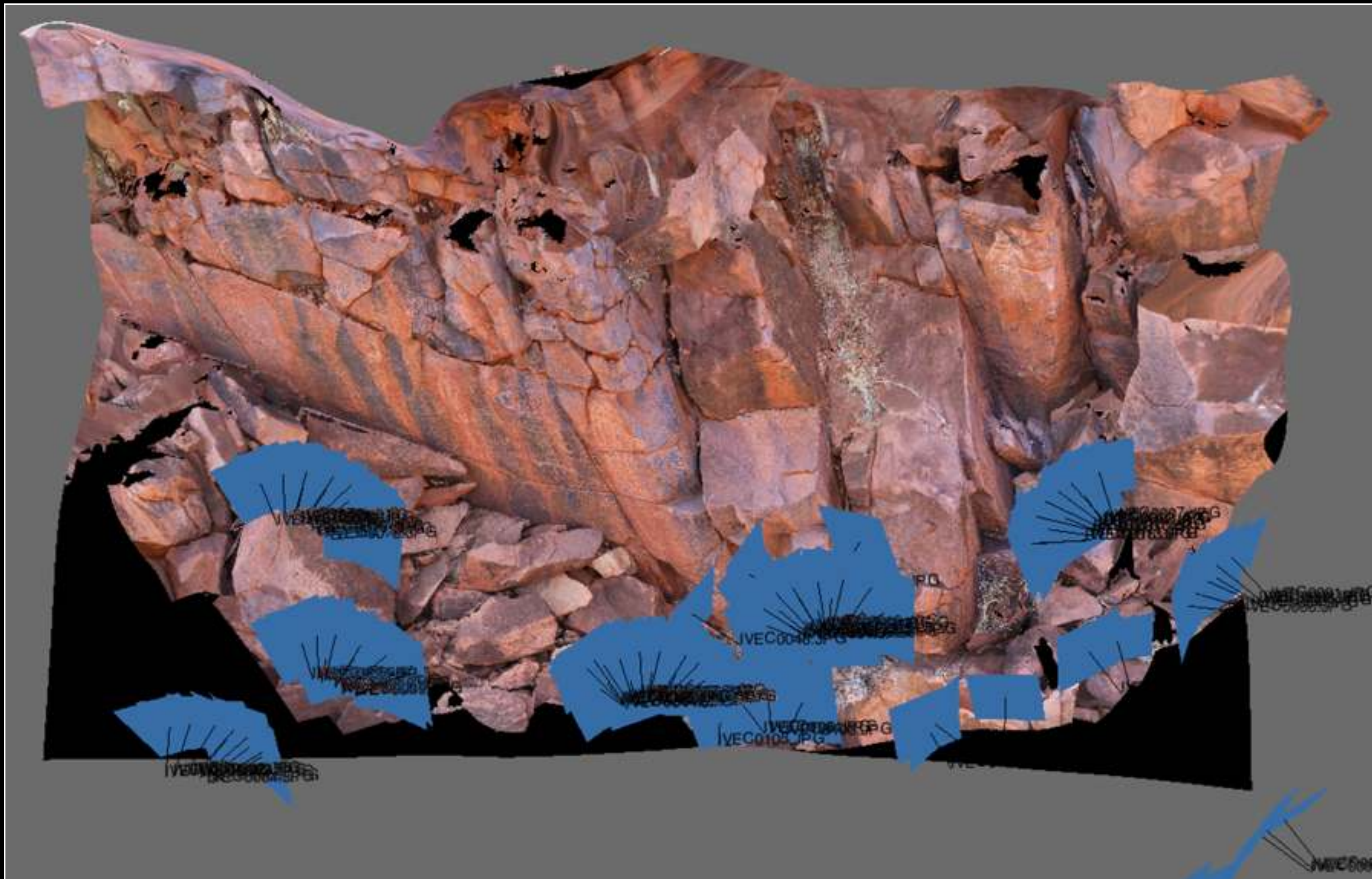
Dragon Gardens, Hong Kong



City University, Hong Kong



Ngintaka headress



Wanmanna

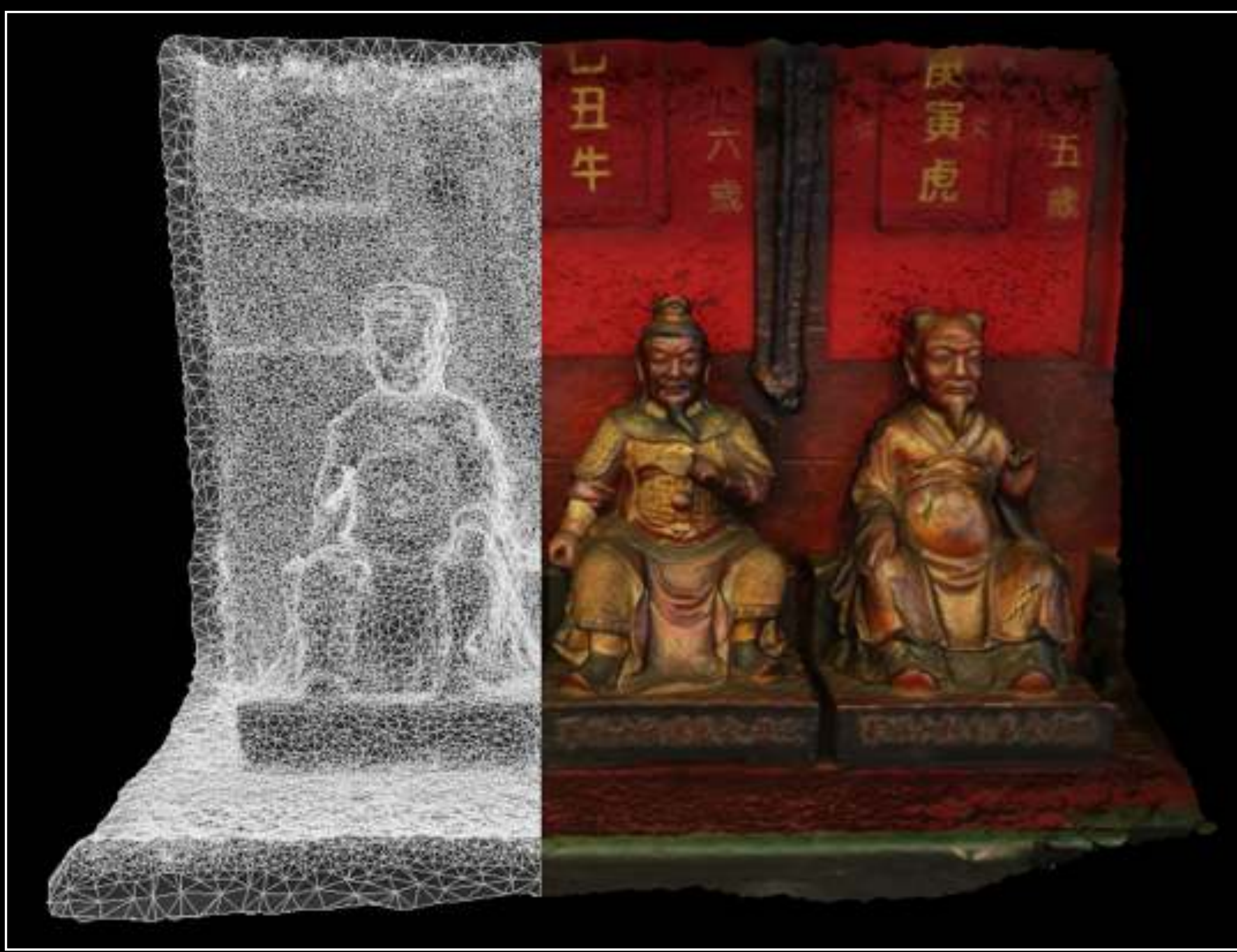
Coverage

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

~30 photographs

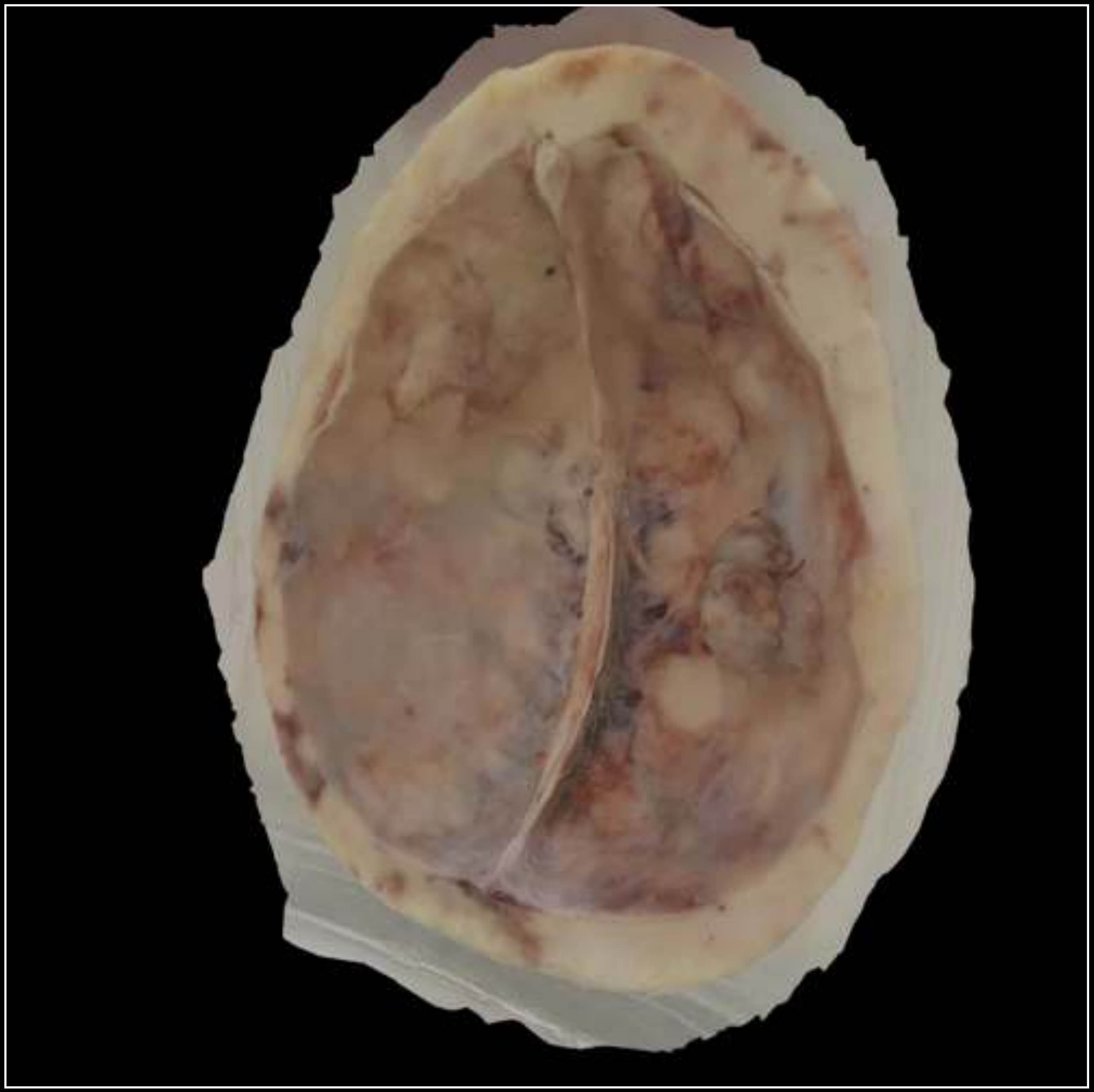
Limited by left most photograph

Tung Wah hospital temple, Hong Kong



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
- ... lots of others

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

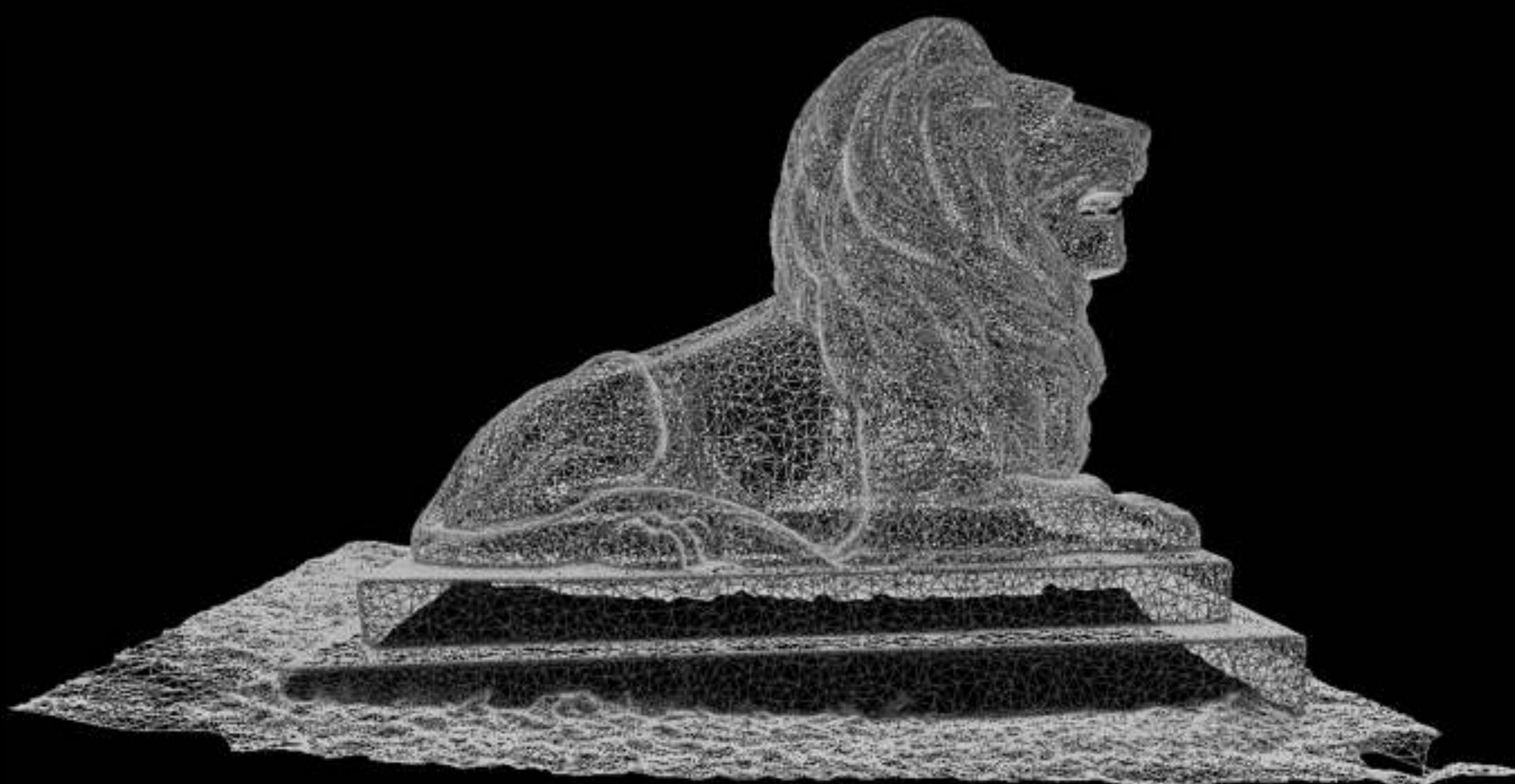
Mesh simplification

- 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.

1,000,000 triangles

100,00 triangles

1,000,000 triangles



100,00 triangles

Considerations

- Different applications have different requirements.
- For archiving one may store the highest resolution version.
- For online one may create lower resolution versions.

- Same general considerations one has for images/photographs.

Other topics

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

Accuracy

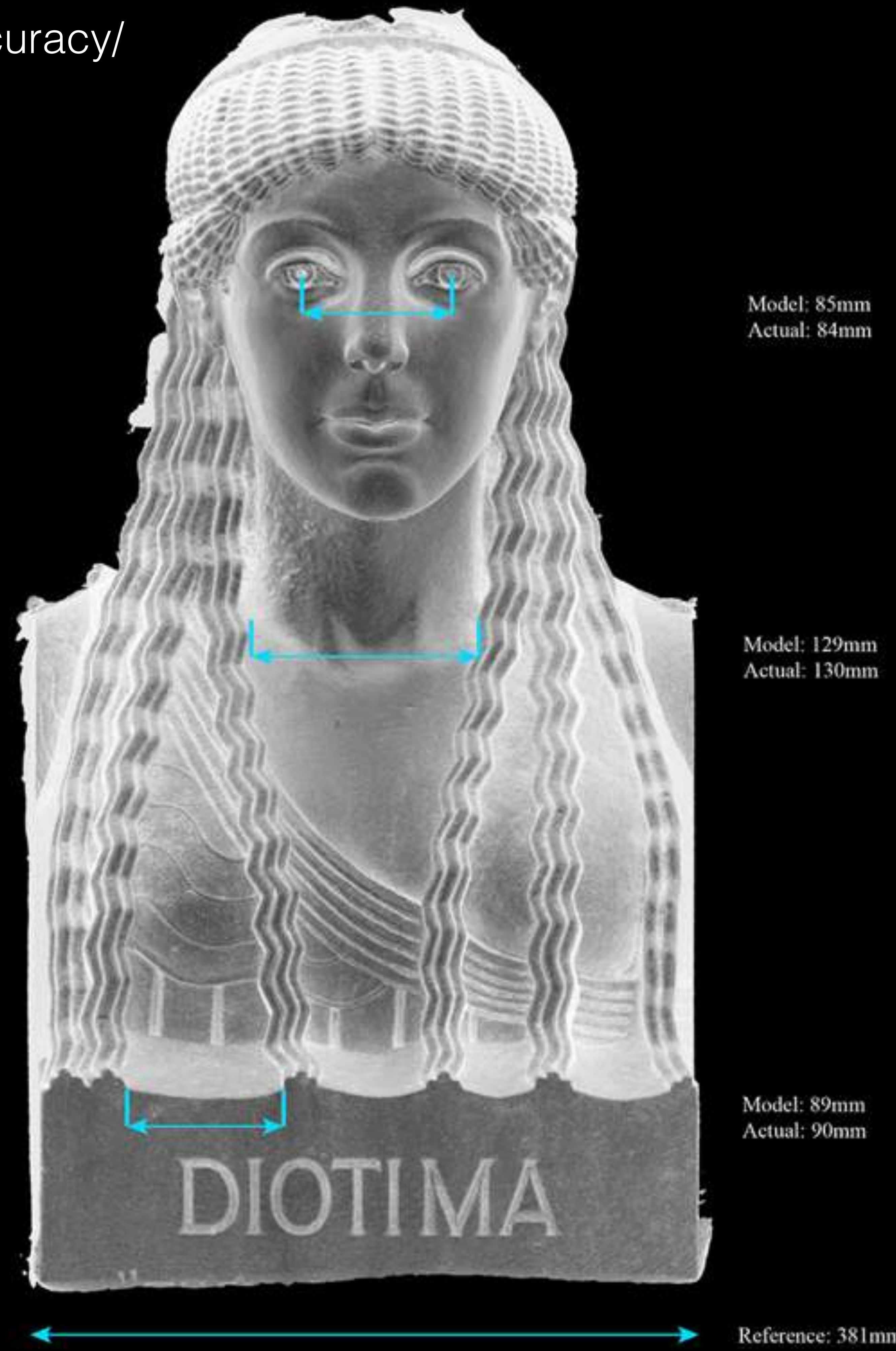
- The first question many people ask is how accurate is it?

- 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.

http://paulbourke.net/miscellaneous/reconaccuracy/

- 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

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


DIOTIMA


DIOTIMA

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.

Relighting

Relighting

Analysis

1000 Buddha temple, Manipal, India

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.

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

Occluders

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

The University of Western Australia

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.

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

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.

Grass shadows

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.

Reflective surfaces

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

Worked example 1
Guard Stone, Sri Lanka Museum

Worked example 2
Lion from Fort Canning entrance