Explorations in digital capture 2D and 3D

Paul Bourke

iVEC @UWA

Summary

- High resolution scanning is about two realisation:

  If we are digitally capturing an object then should try to do so at a resolution and quality we won't regret in the future.

  We cannot buy an arbitrary high resolution sensor.

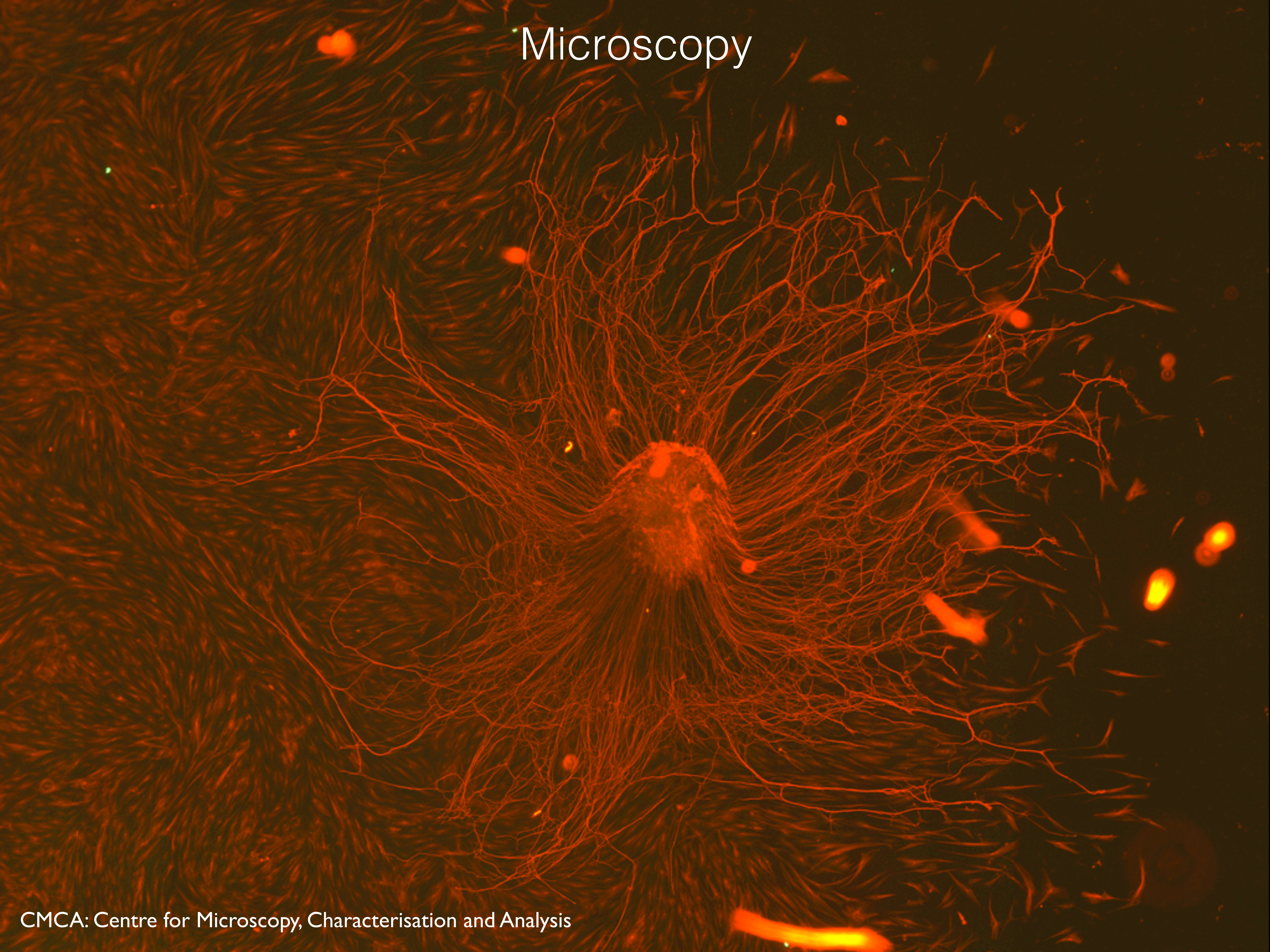
- The solution is to use available sensors to contribute to part of the final image.

  Line scanning

  Area scanning: panorama, mosaics

- This is not just for document archives, it is becoming pervasive across many disciplines.

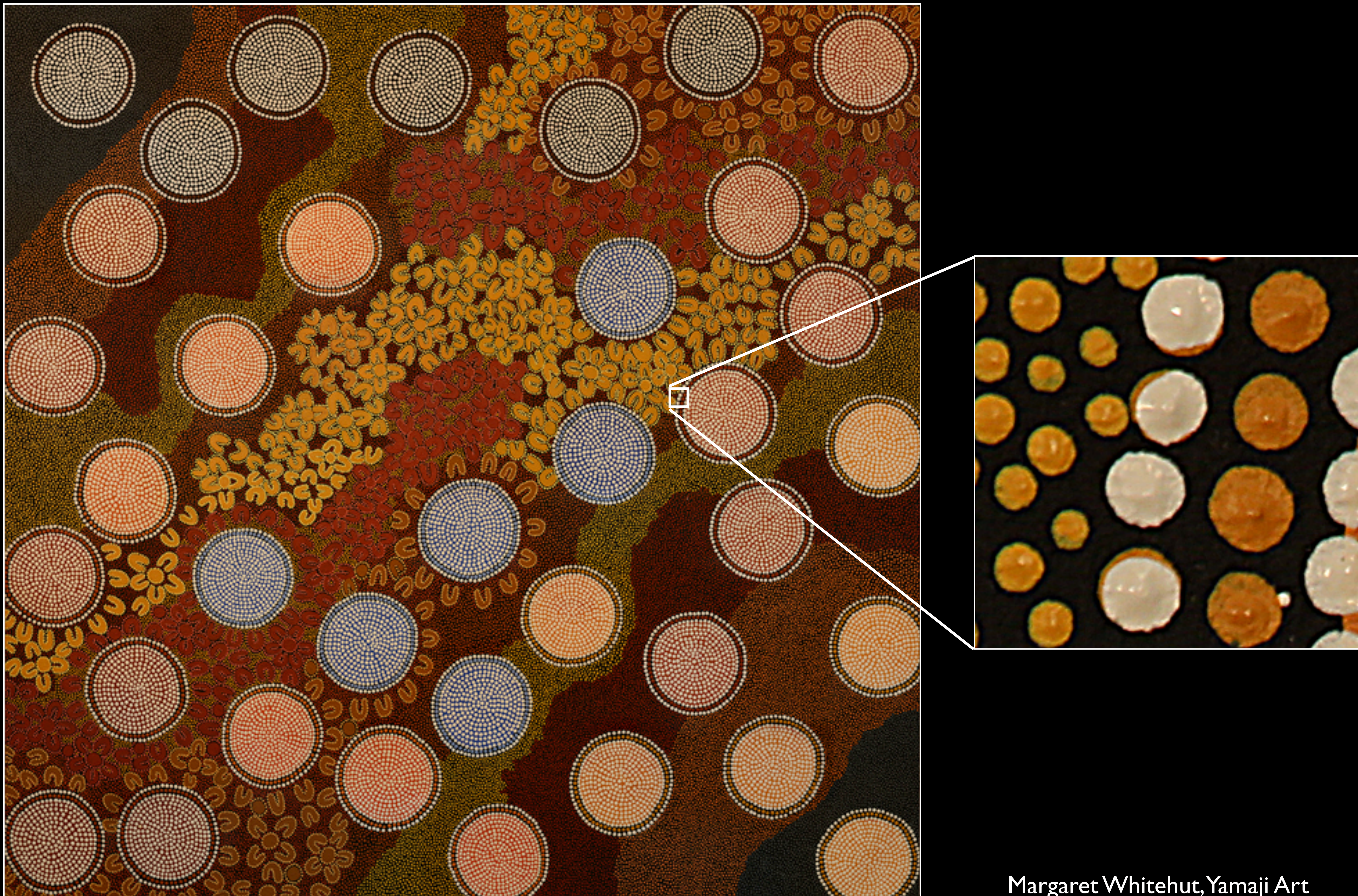
Microscopy



CMCA: Centre for Microscopy, Characterisation and Analysis

Astrophysics

Dot painting : Forensics



Margaret Whitehut, Yamaji Art

Rock art

Rock art site recording

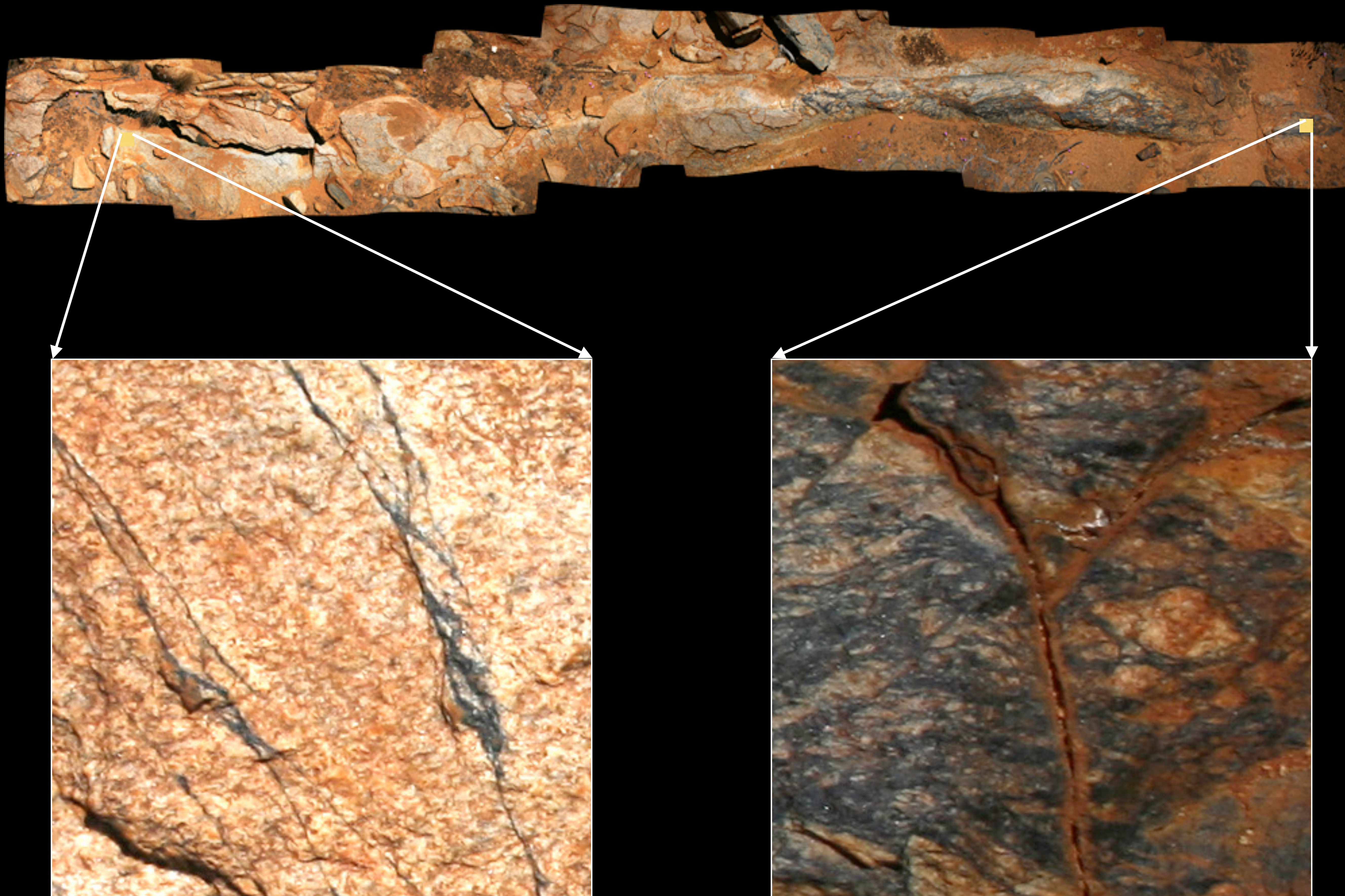
Manmanna, Archaeology, UWA

Manmanna, Archaeology, UWA

Site recording

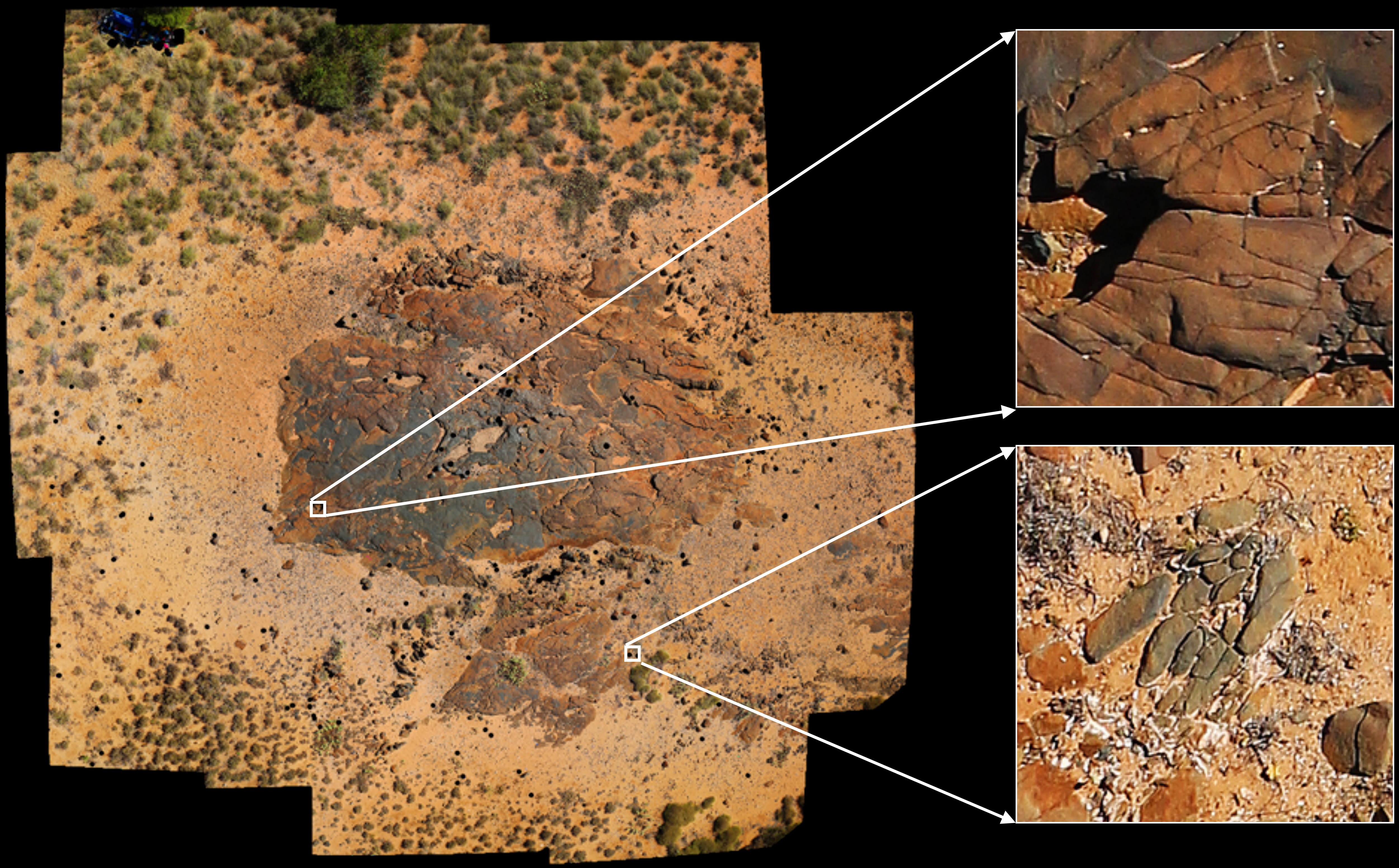
Beacon Island, Maritime Museum of Western Australia

Geoscience



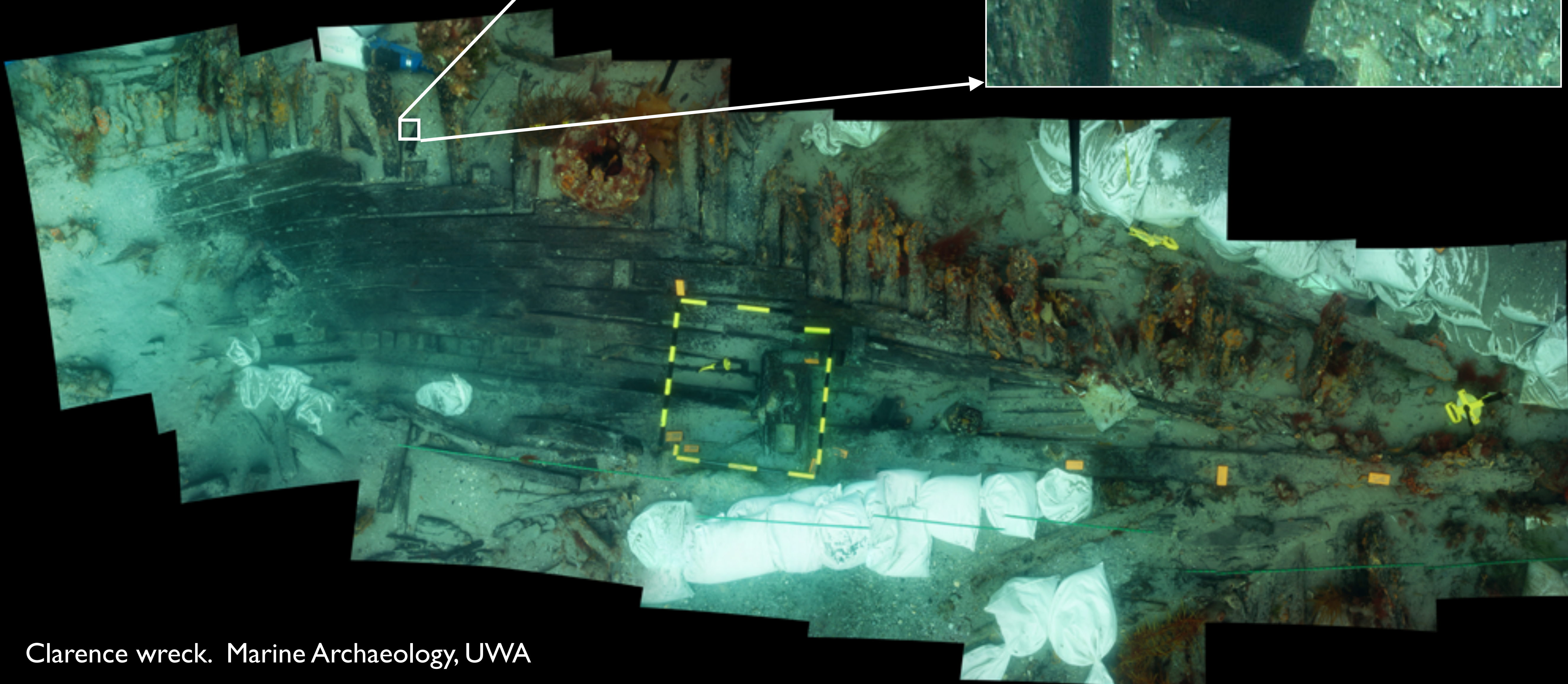
Department of Mines and Petroleum

Aerial scanning



Centre for Rock Art Archaeology + Management, UWA

Marine Archaeology



Clarence wreck. Marine Archaeology, UWA

Challenges

- Main challenge is not the capture technology, it will steadily improve.

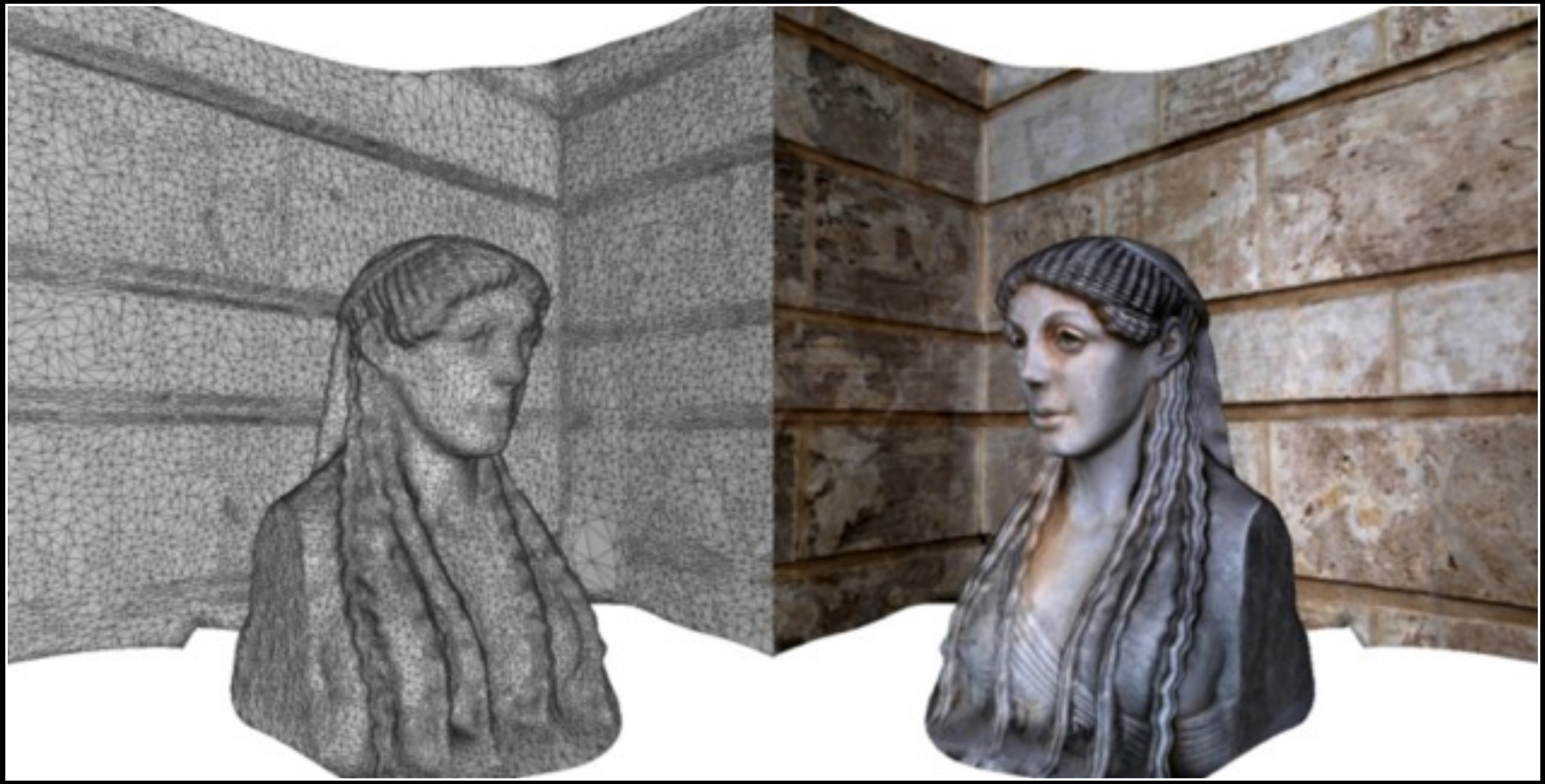
- Lack of good standards for annotating ultra large images: location/area/direction meta data containers. Subsequent lack of high level searching.

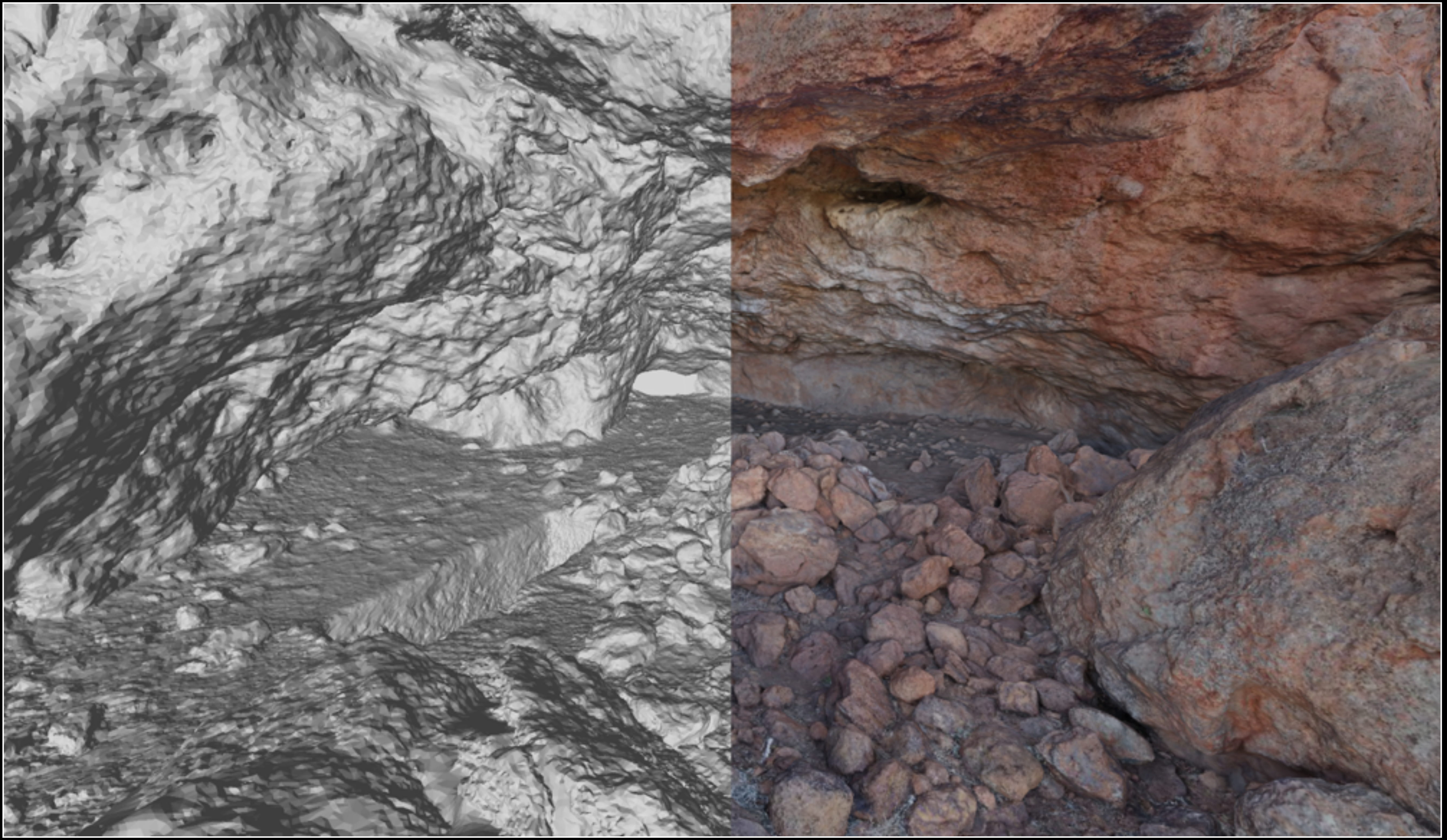
- Lack of implementation of good standards for interactively navigating within these images.

- Lack of support in databases and archives. Be warned those who are trying to archive these though LiveARC for the national (RDSI) archives.

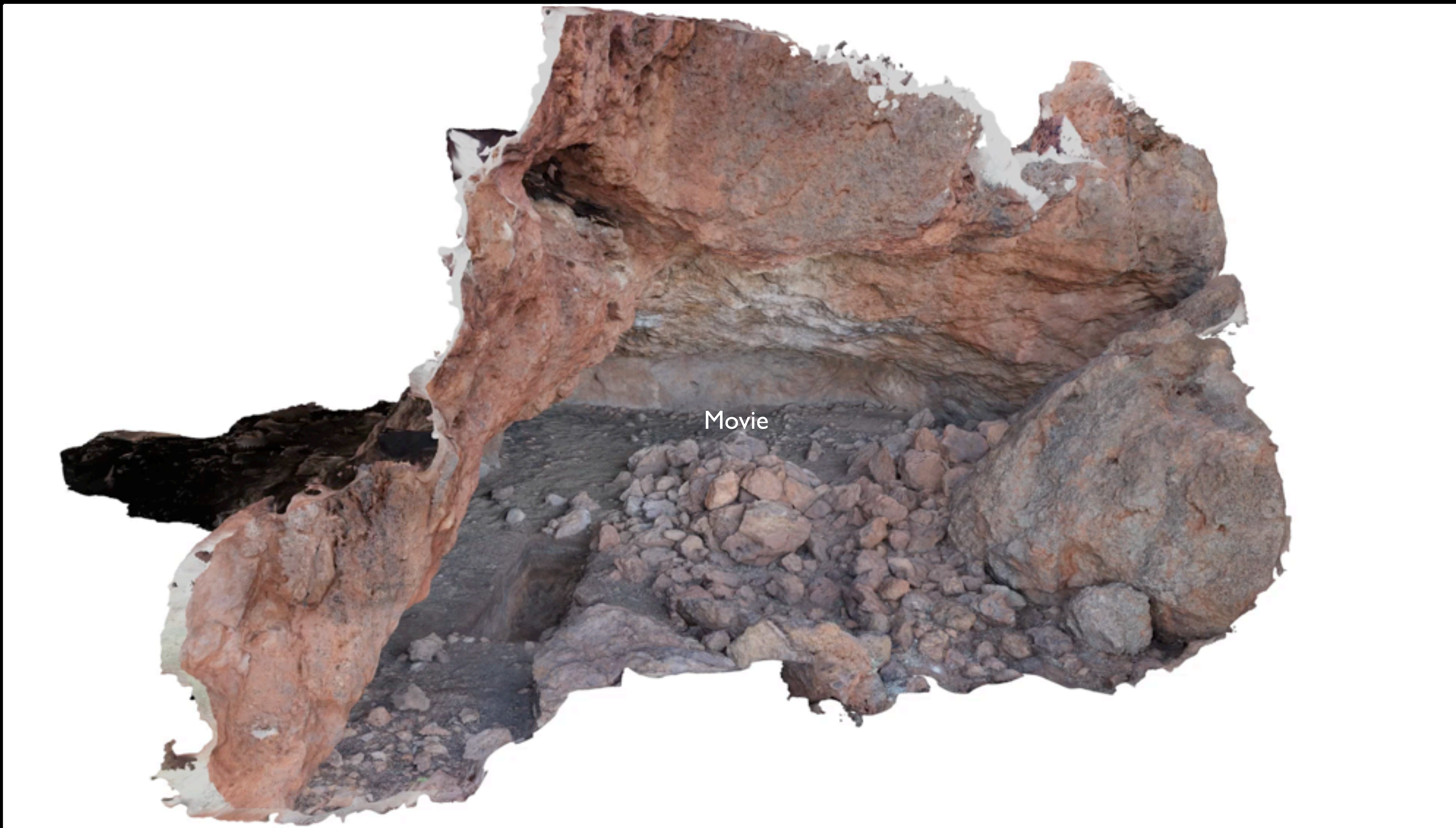
Automated 3D model reconstruction from photographs

Paul Bourke
iVEC@UWA





Rock shelter, Weld Range
Derived from 350 photographs



Rock shelter, Weld Range
Derived from 350 photographs

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.
- Creating assets for games and 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.
  [Current site surveys recorded 100's of objects].

Applications : Virtual worlds

- 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.

Movie

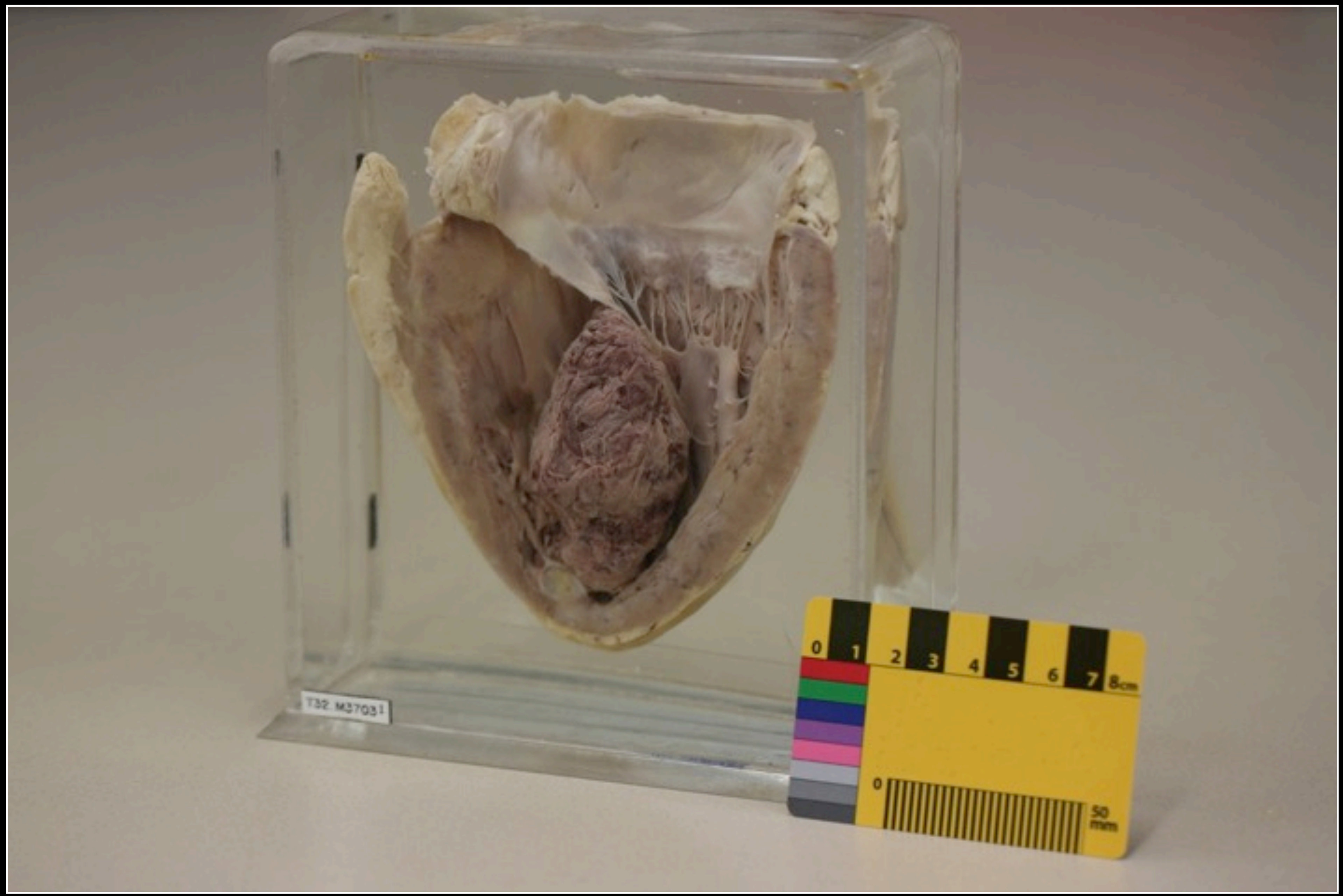
Beacon Island

Applications : Assets for virtual heritage

Beacon Island

Applications : Teaching in medicine

- Medical applications.
- Non intrusive capture can have advantages.
- Capture of 3D objects for forensic analysis.



Pathology, QE II


Movie

Applications : Geoscience

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

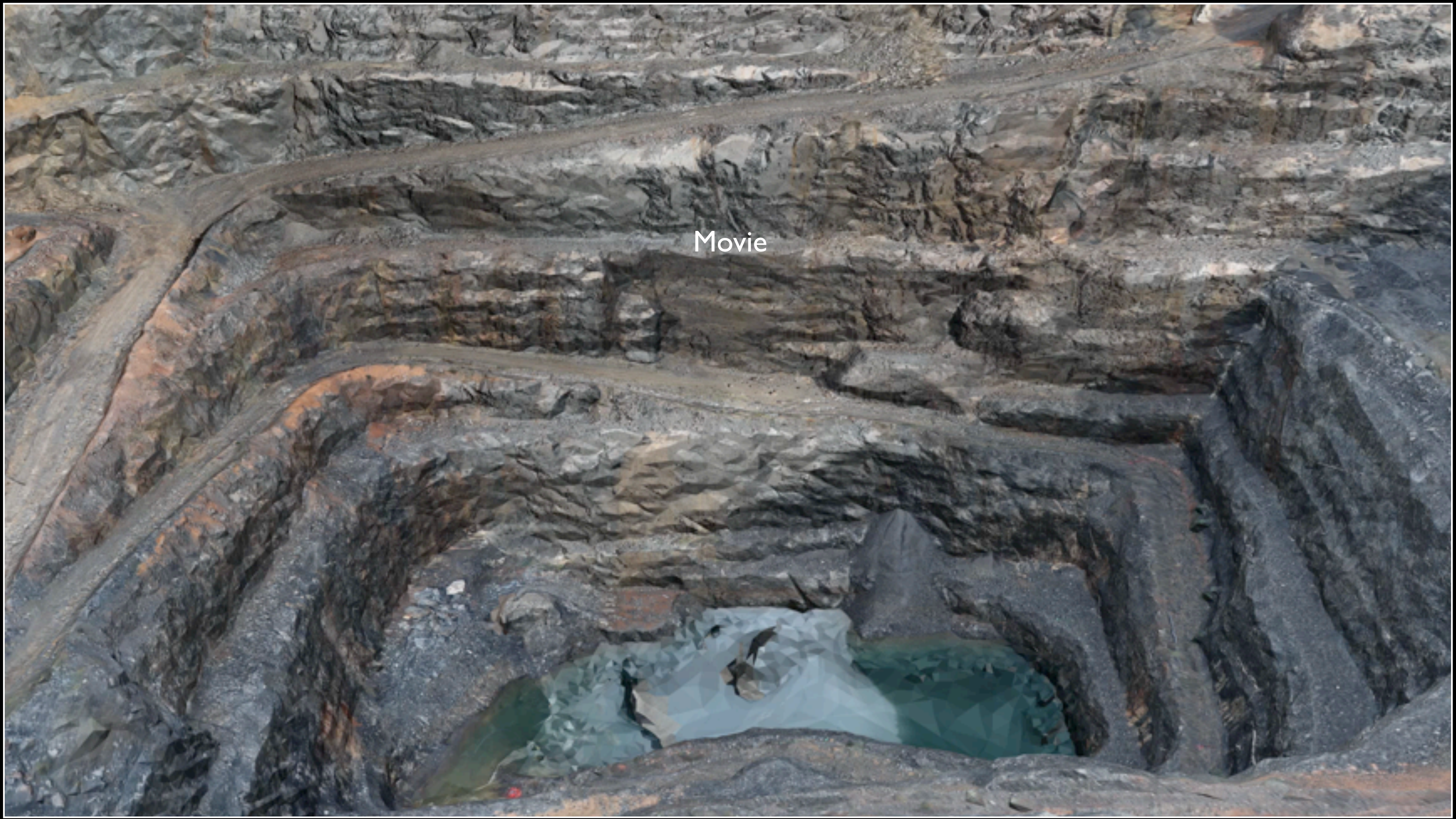
Department of Mines and Petroleum

Geoscience

Movie

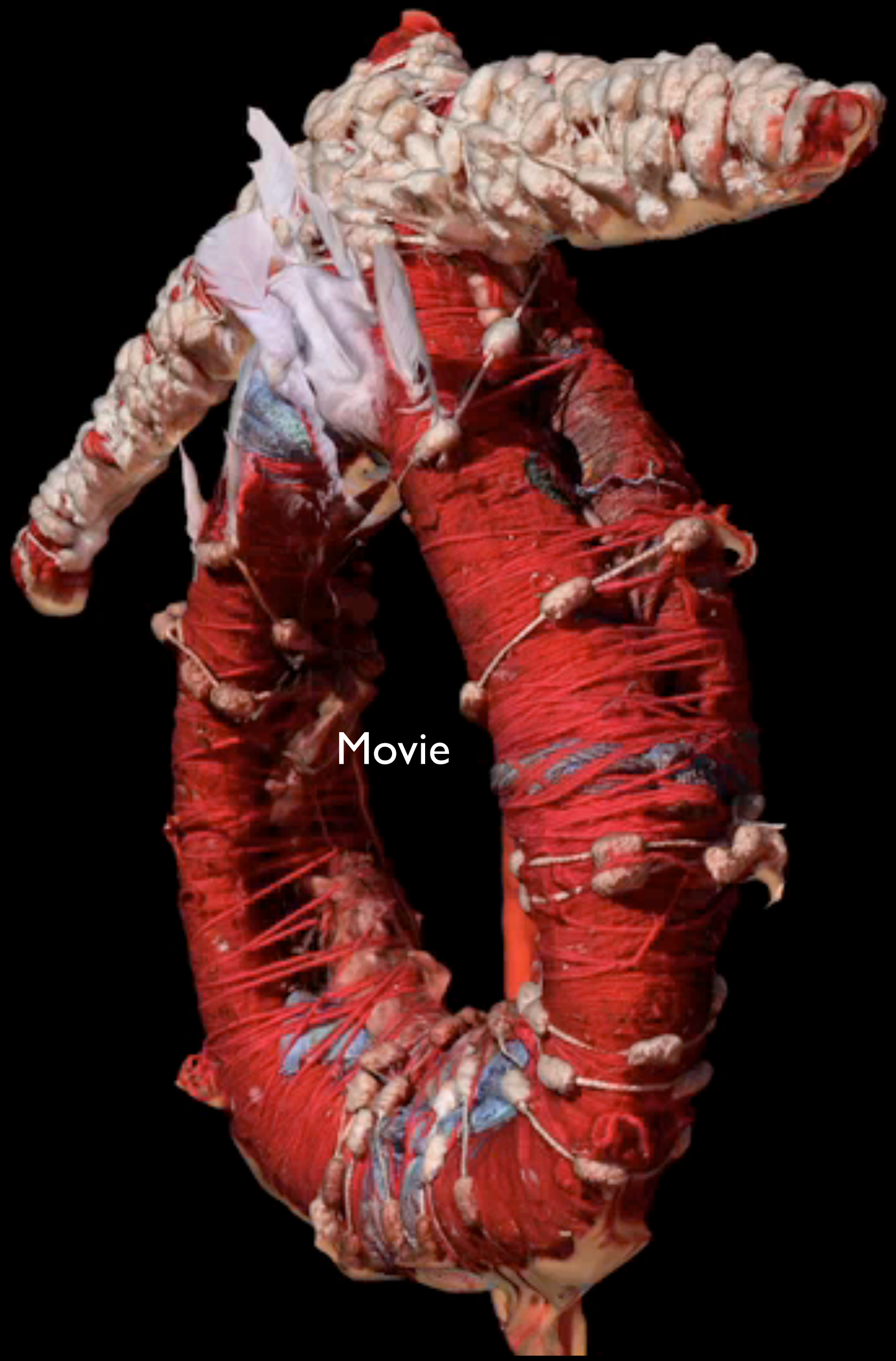
Applications : Mining

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



Centre for Exploration Targeting, UWA

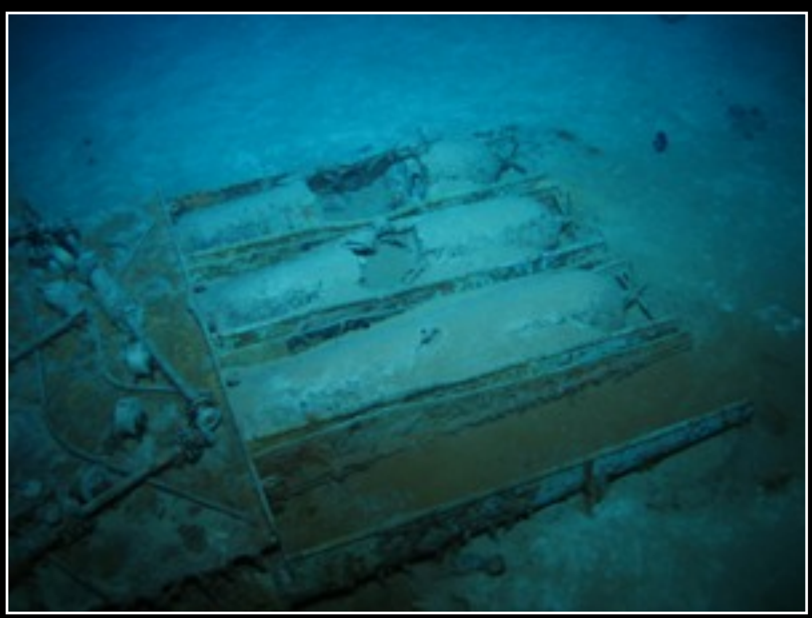
Applications : Artefacts in cultural heritage



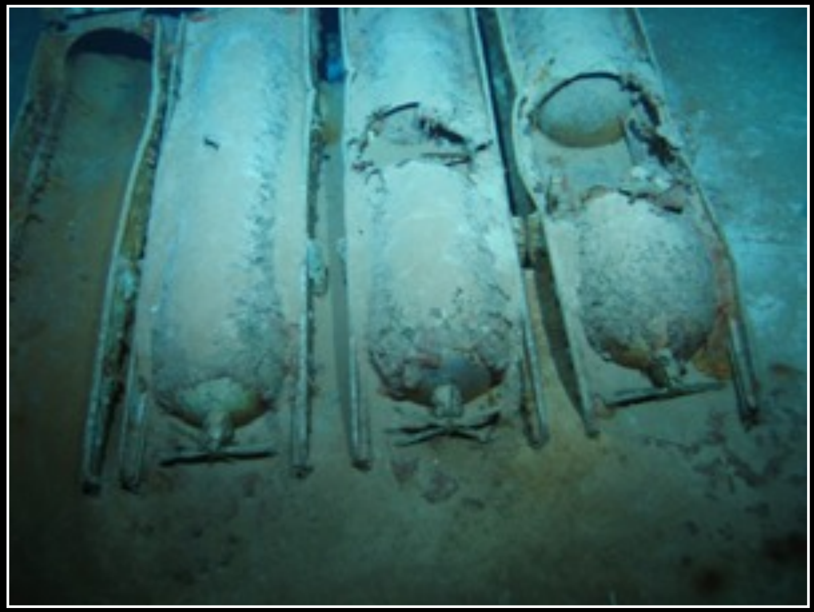
Ngintaka, South Australia Museum

Applications : Marine archaeology

- 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.

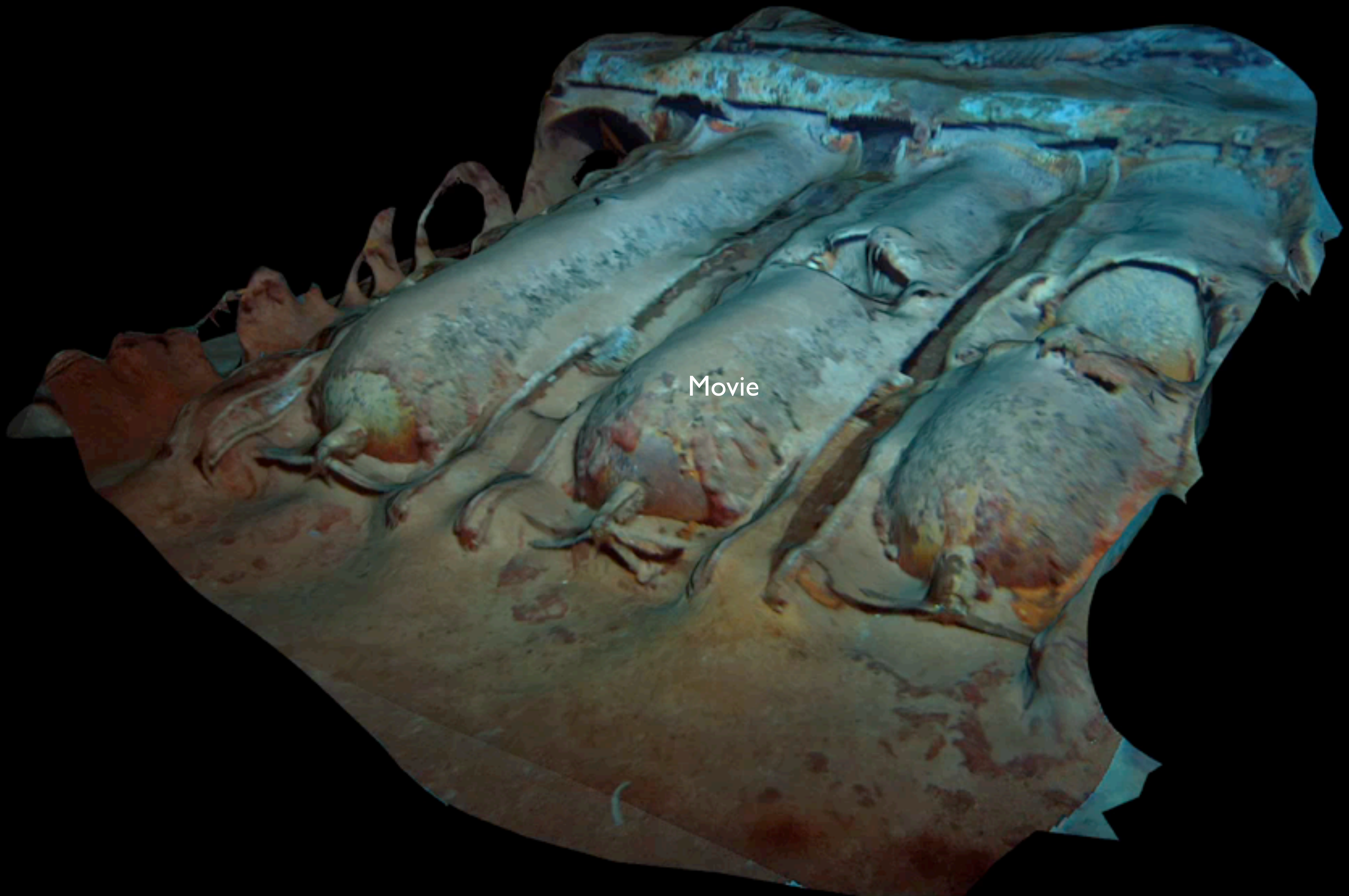






HMAS Sydney

Additional applications : Underwater



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 investment

- Another technology are so called depth cameras.
  - Primesense (eg: Kinect)
  - Structured light techniques (eg: Artec Scanner)

- Both in theory can give more accurate results. Subject to debate.

- Both also have limitations around lighting conditions and range.

- Future: Light field cameras (plenoptic camera).
  - Captures an array of images from a grid of positions

LIDAR

Structured light

Pipeline components

- Perform lens calibration (only done once, increasingly optional 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 pipeline

SIFT for key point extraction between images

↓

SfM software package Bundler to generate a sparse 3D point cloud

↓

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

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 analyse just a subset of the images.

Software : Pipeline - Sparse point cloud

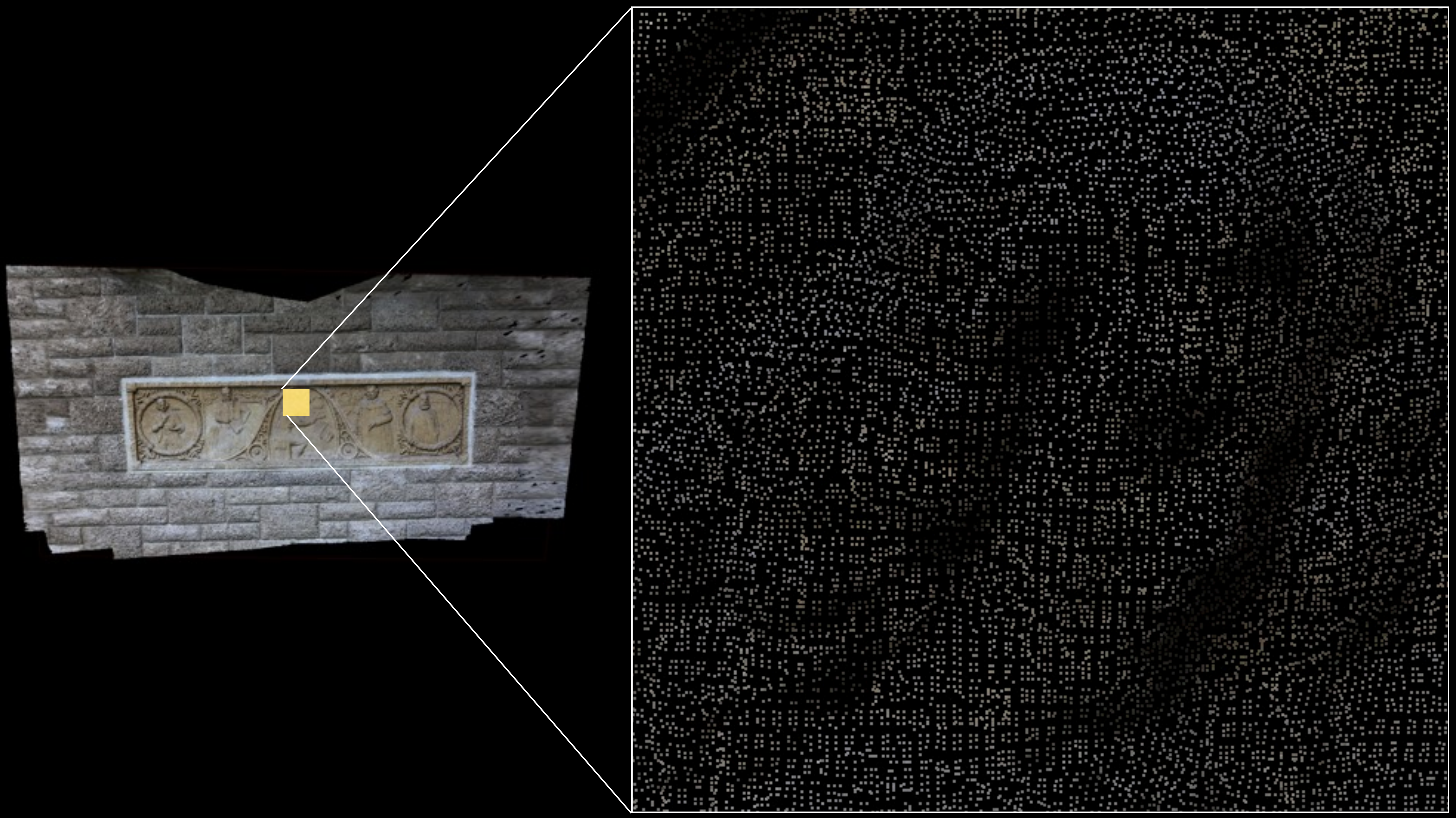
- Find matching points between photographs, feature point detection.
  SIFT - scale invariant feature transform
- Compute camera positions and other intrinsic camera parameters.
  Bundler, SfM - Structure from Motion

Software : Pipeline - Dense point cloud

- CMVS - Clustering Views for Multi-view Stereo.

UWA Geography Building

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.

UWA Geography Building

Software : Pipeline - Export

Movie

UWA Geography Building

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.
  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
  - 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.
- Generally have some minimum focus distance and small aperture.
- 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 111 with prime lenses: 28mm, 50mm, 100mm macro.

Sigma 28mm, Canon mount

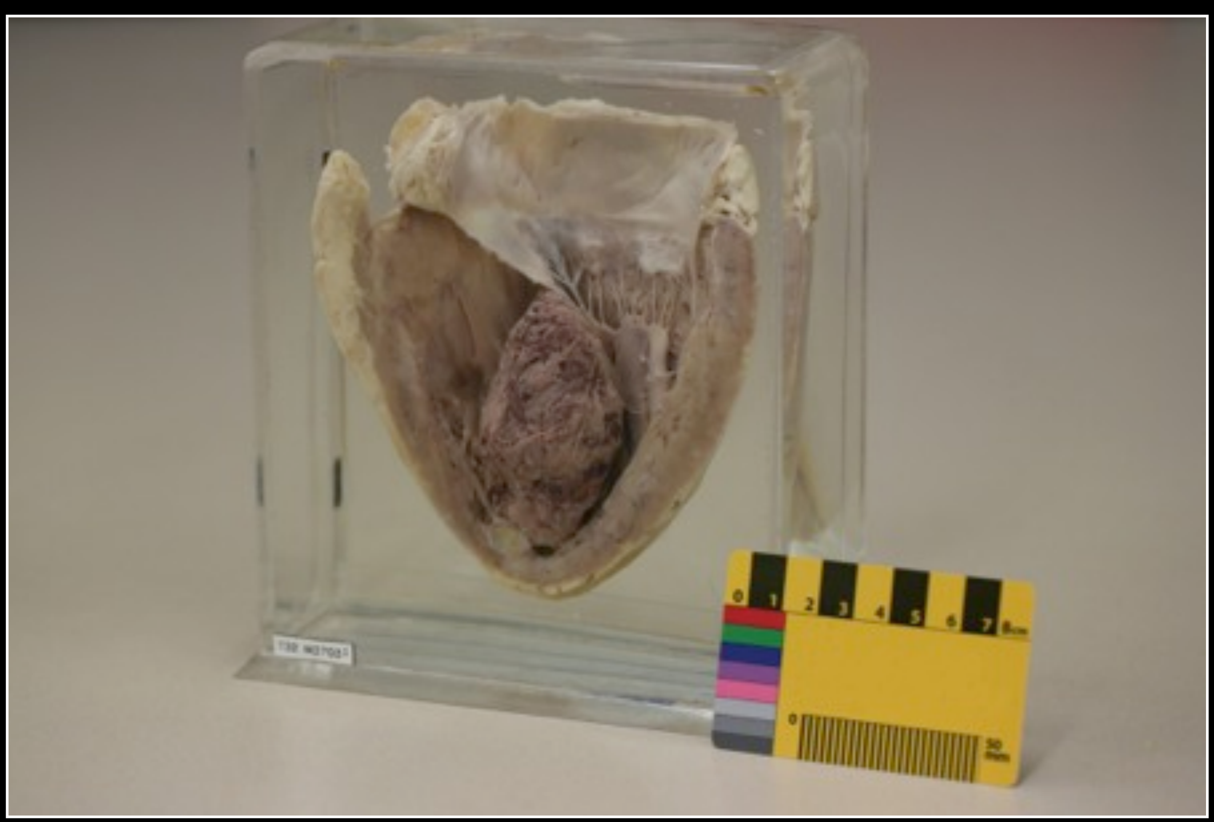
Sigma 50mm, Canon mount

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.
- 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.
- Generally works better for the images to be captured in order moving around the object (may no longer be the case for 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.
- Calibration: Most of the packages that include accuracy metrics will assume a camera calibration.

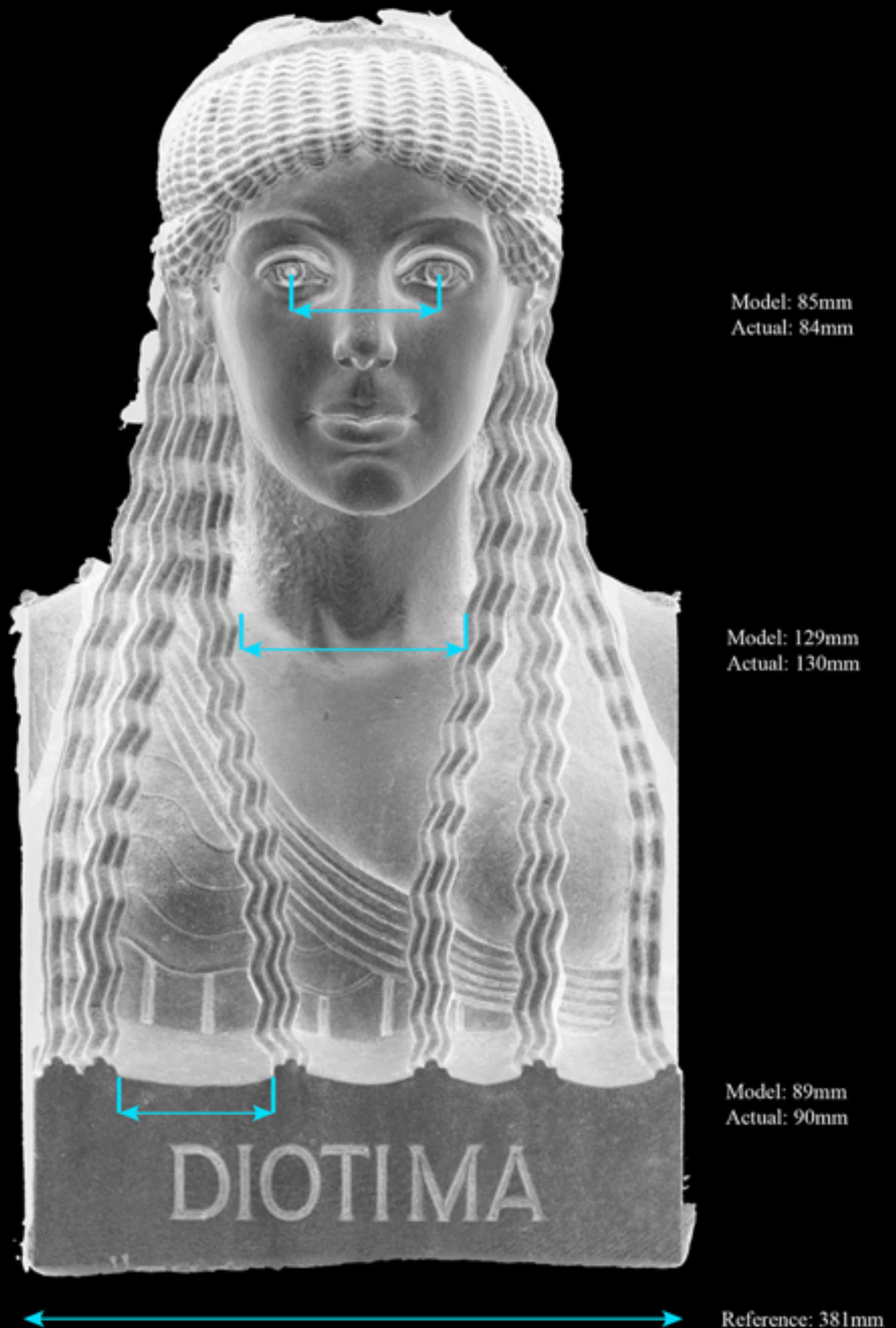
Photogaphy : Linear reference objects

- Assists processing 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 representation is important.



Accuracy

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



Comparisons

Original photograph

Reconstructed model

Shaded to emphasise surface variation

Comparisons

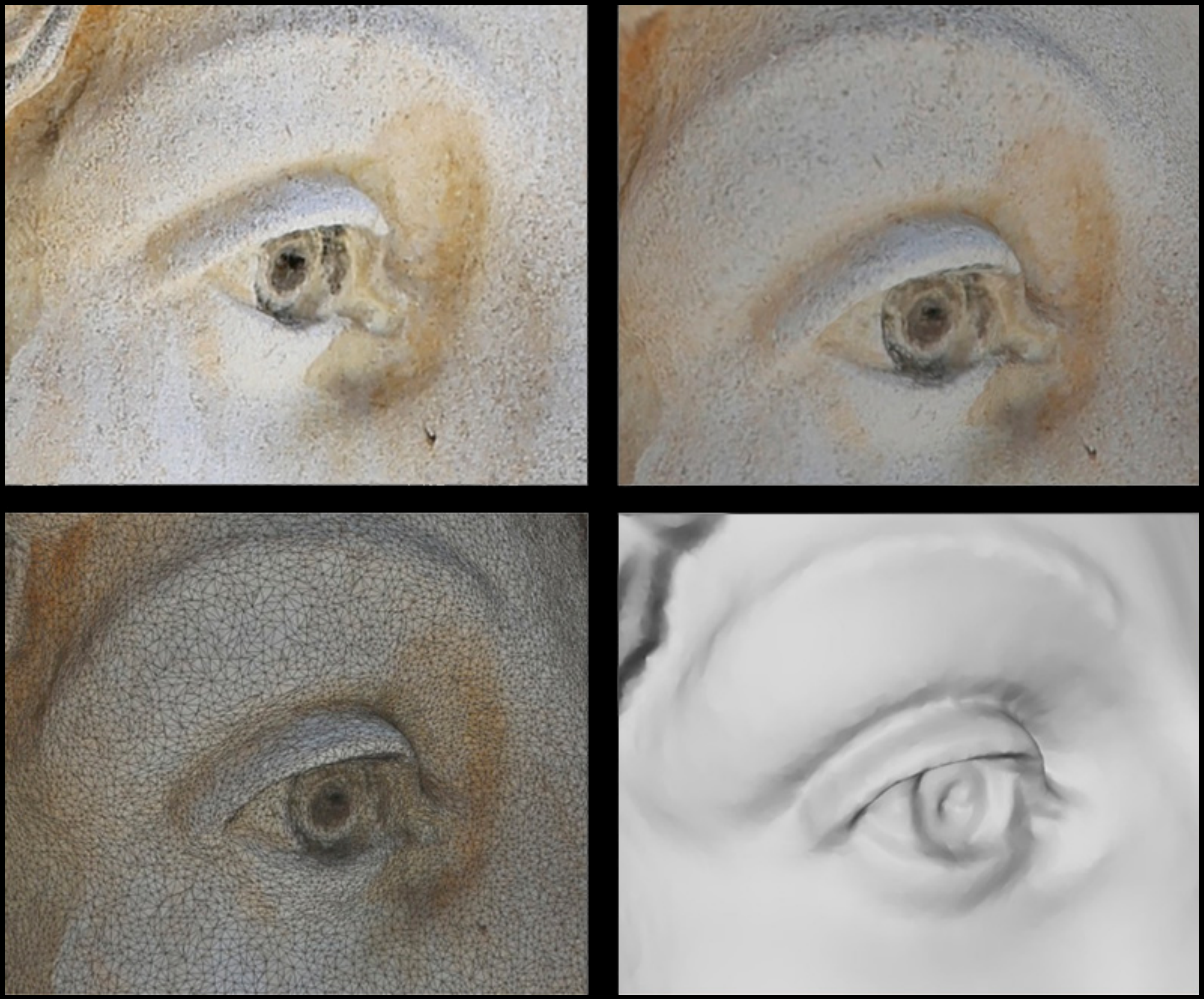
Original photograph

Reconstructed model

Shaded to emphasise surface variation

Comparisons

Original

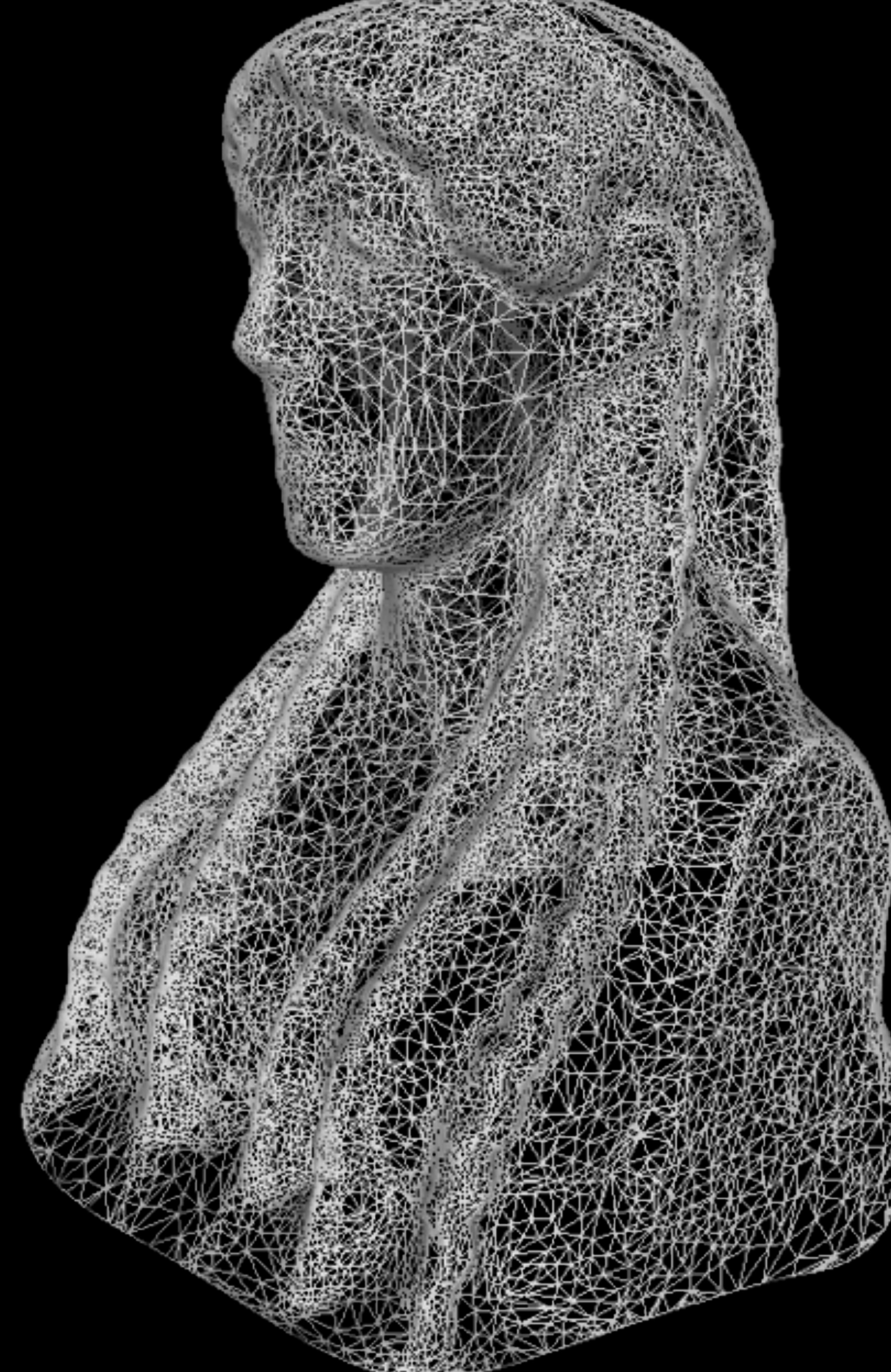


Resolution

- 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

Resolution



Example from 2009

Resolution

Example from 2014

Real vs apparent detail

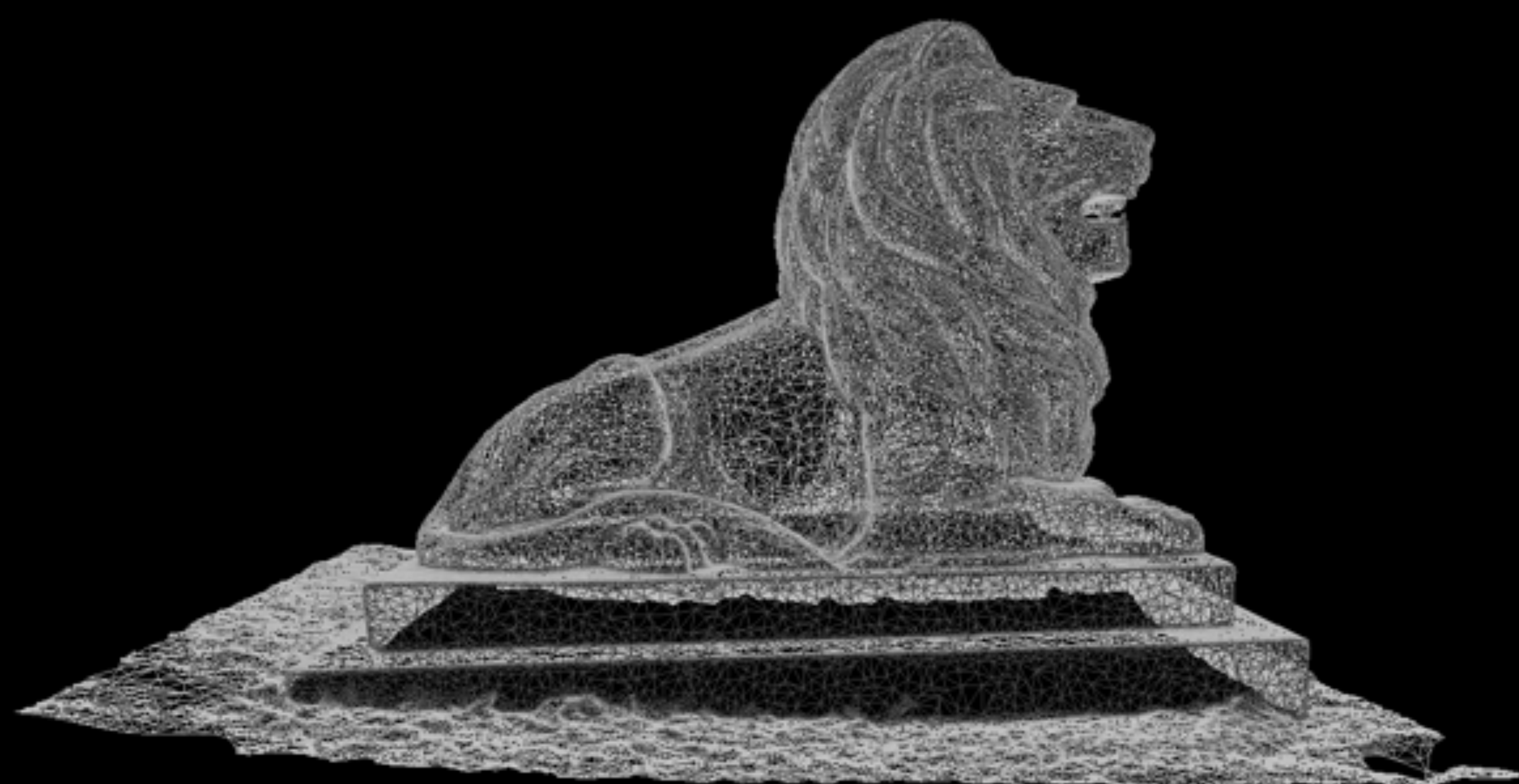
Real vs apparent detail

1,000,000 triangles

100,00 triangles

Real vs apparent detail

1,000,000 triangles



100,00 triangles

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.

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.

Movie

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

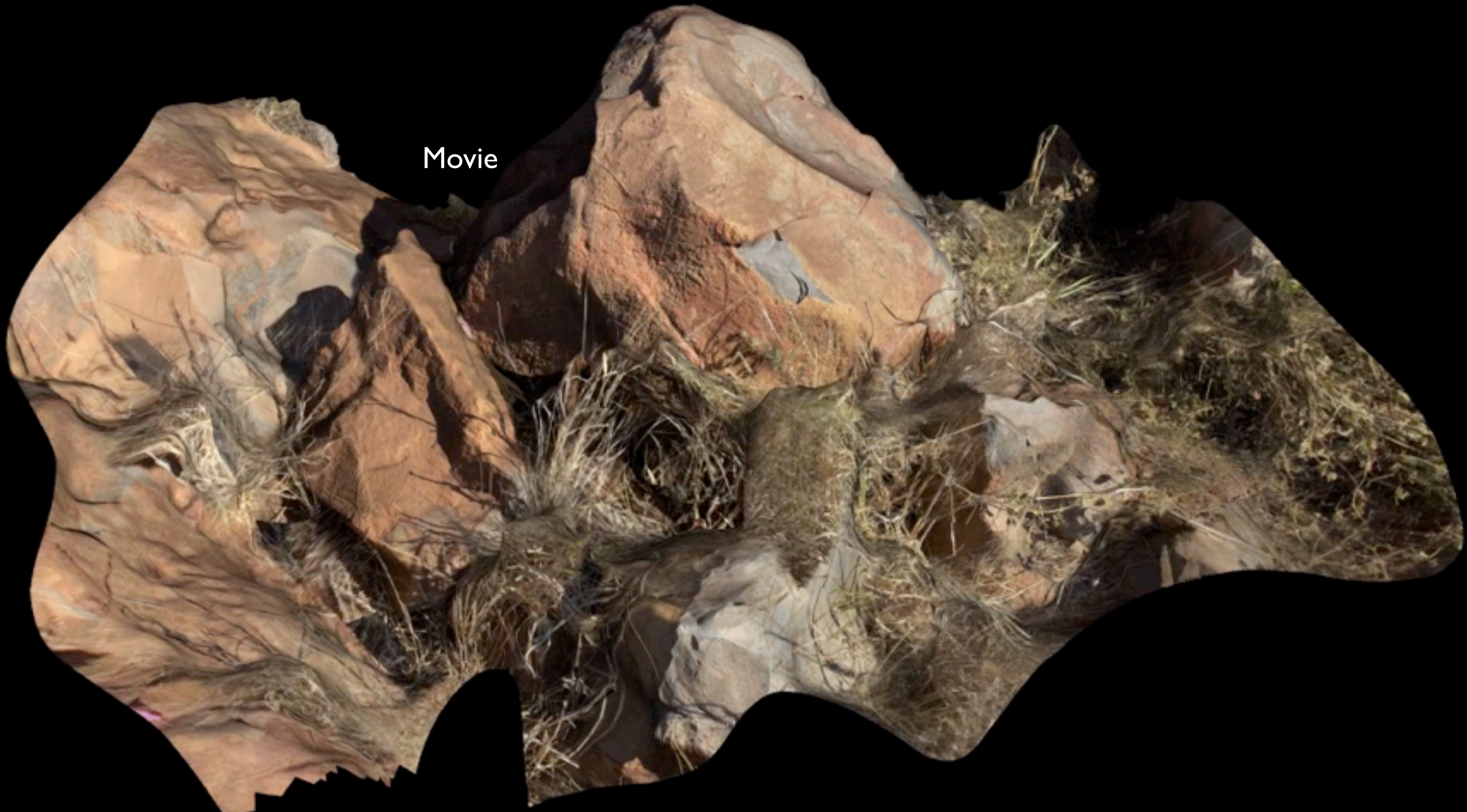
Limitations : Occluders

- Algorithms seem to be generally poor at handling foreground occluders. For example: columns in front of a building.
- Reason: a small change ins camera position results in a large difference in visible objects.

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.

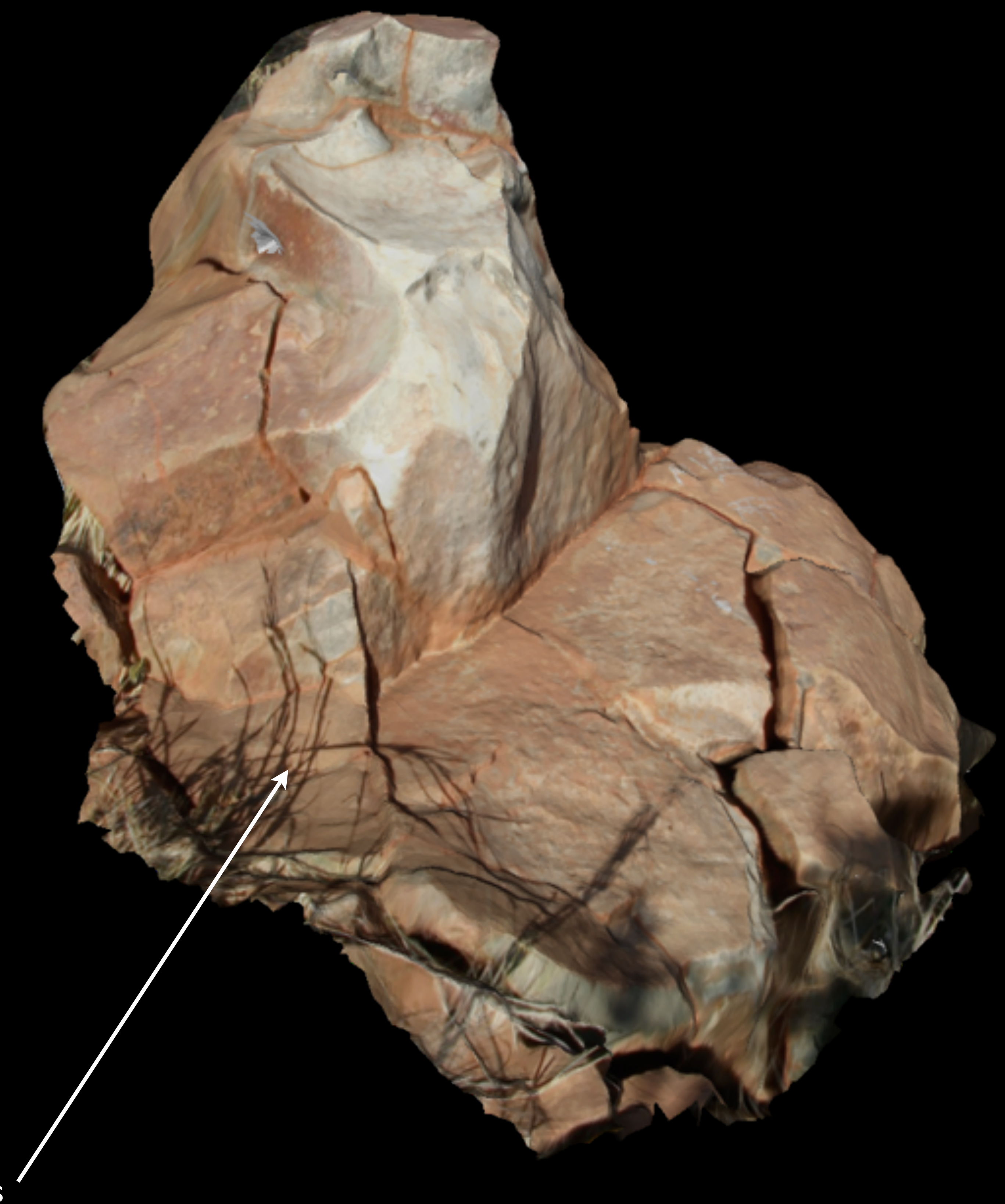


Limitations : Thin structures

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

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.



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

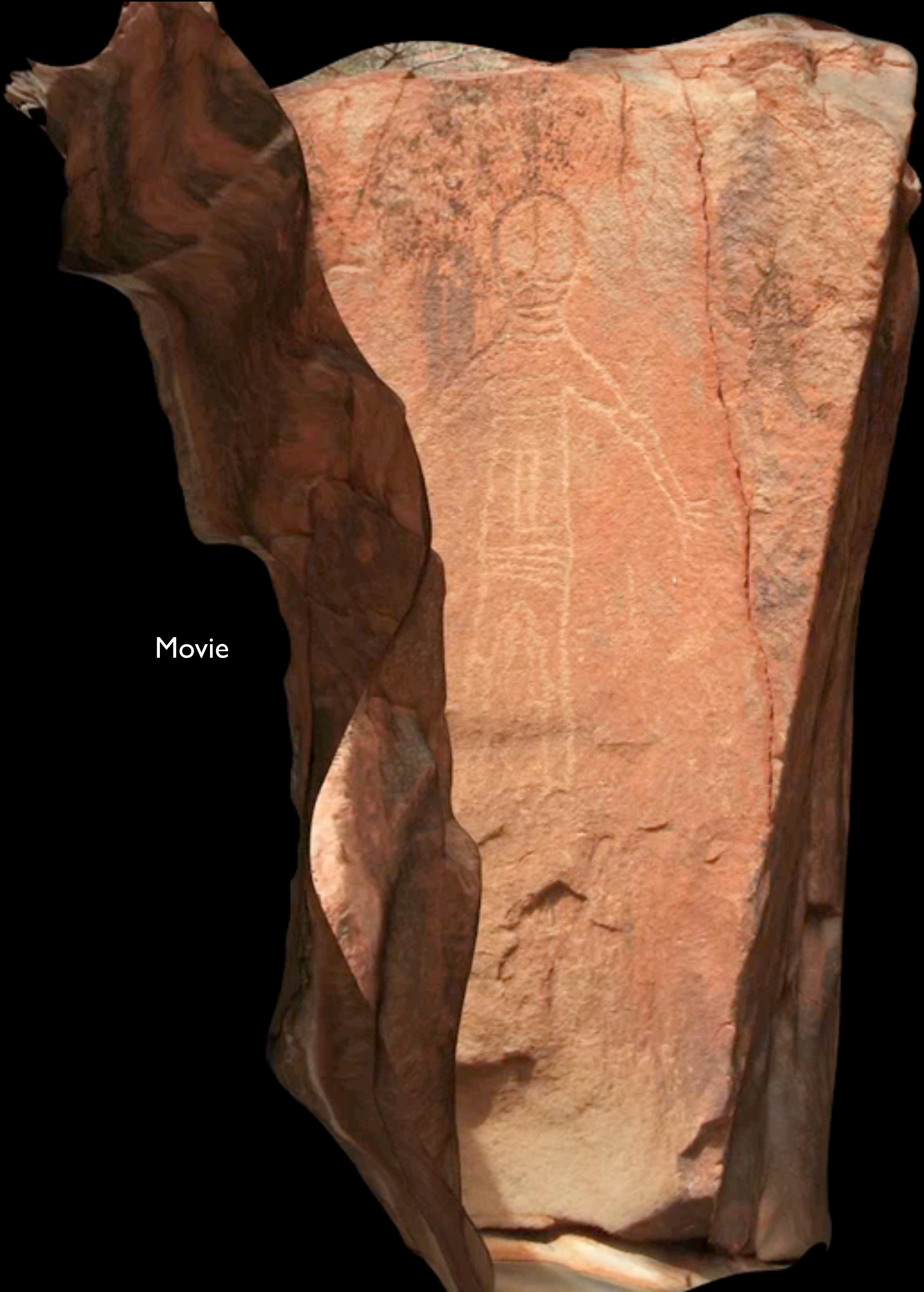
- Mirror surfaces can provide a non-linear reflection of the world that will influence the feature point detection.
- Gives rise to a new art form.
  - Photogrammetry that goes wrong in “interesting” ways.

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

Gommateswara, Manipal, India



Movie

Cave Yallabilli Mindi - Weld Range