Reconstruction

EE/CSE 576

Linda Shapiro

3D model

- "Digital copy" of real object
- Allows us to
  - Inspect details of object
  - Measure properties
  - Reproduce in different material
- Many applications
  - Cultural heritage preservation
  - Computer games and movies
  - City modelling
  - E-commerce



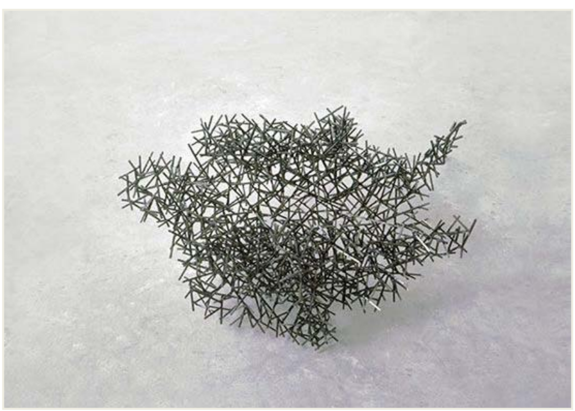


Applications: cultural heritage

SCULPTEUR European project

Applications: art

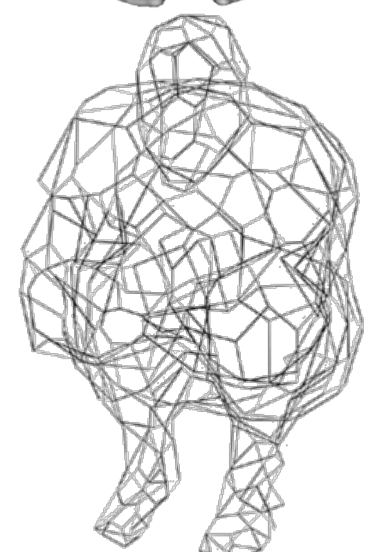
Block Works Precipitate III 2004
Mild steel blocks 80 x 46 x 66 cm



Domain Series Domain VIII Crouching
1999 *Mild steel bar* 81 x 59 x 63 cm

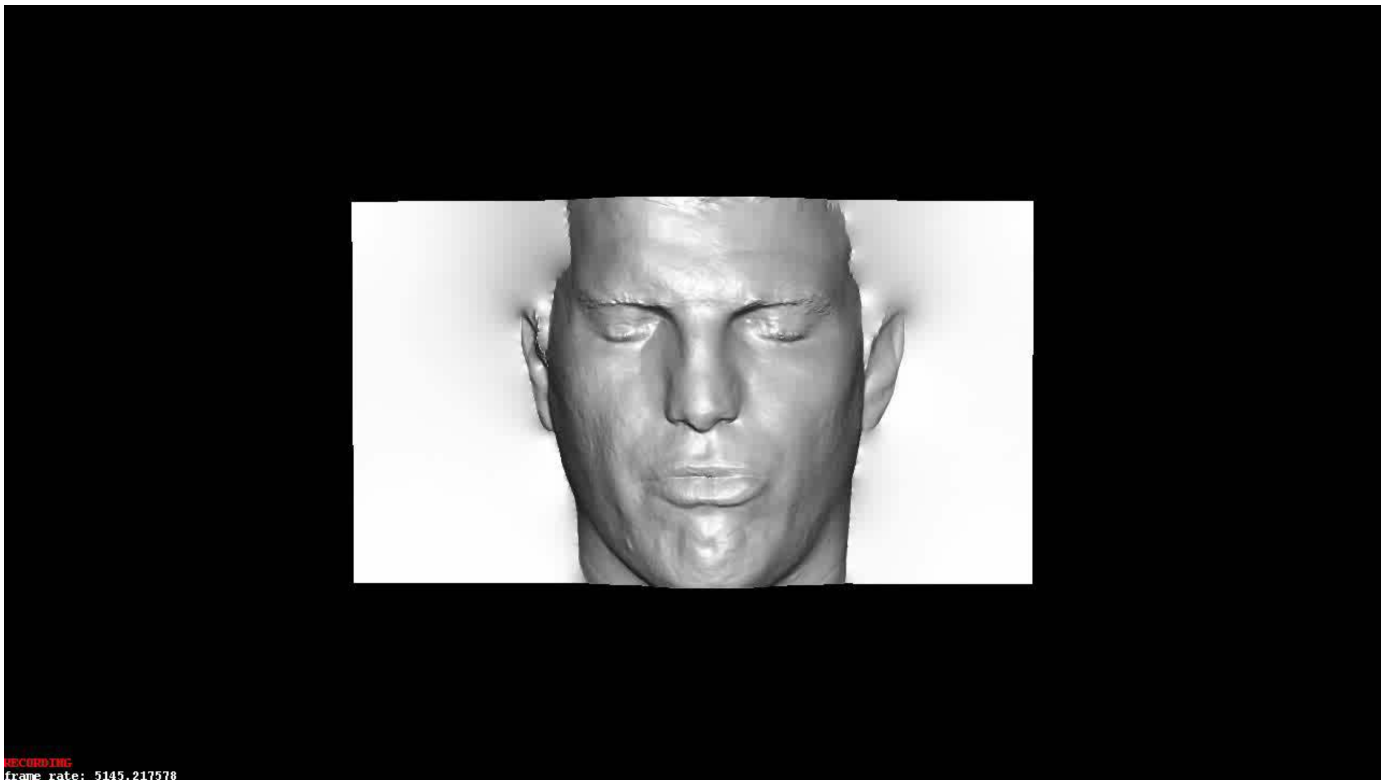
Applications: structure engineering

BODY / SPACE / FRAME, Antony Gormley, Lelystad, Holland

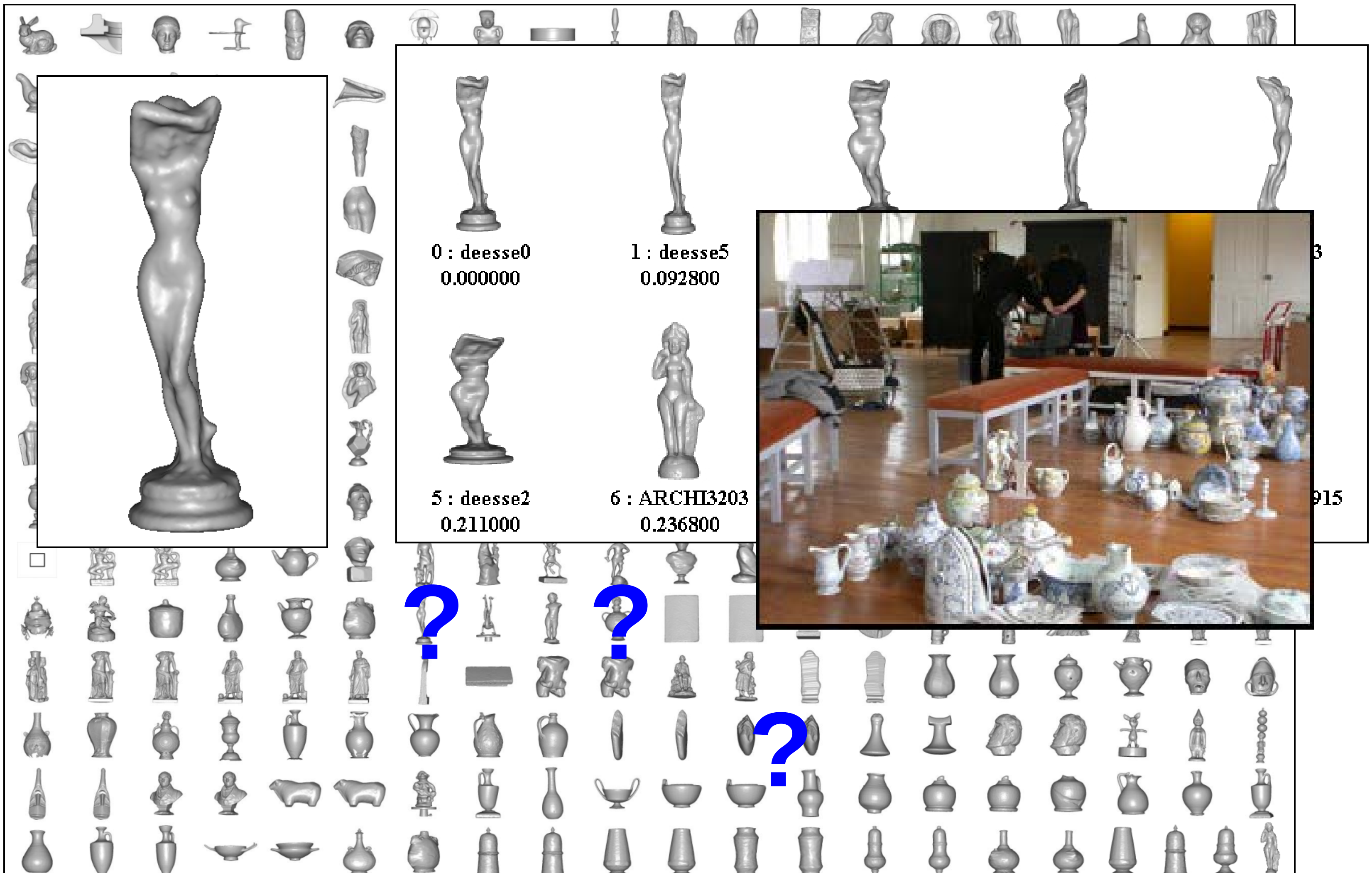


Applications: art

Applications: computer games



Applications: 3D indexation

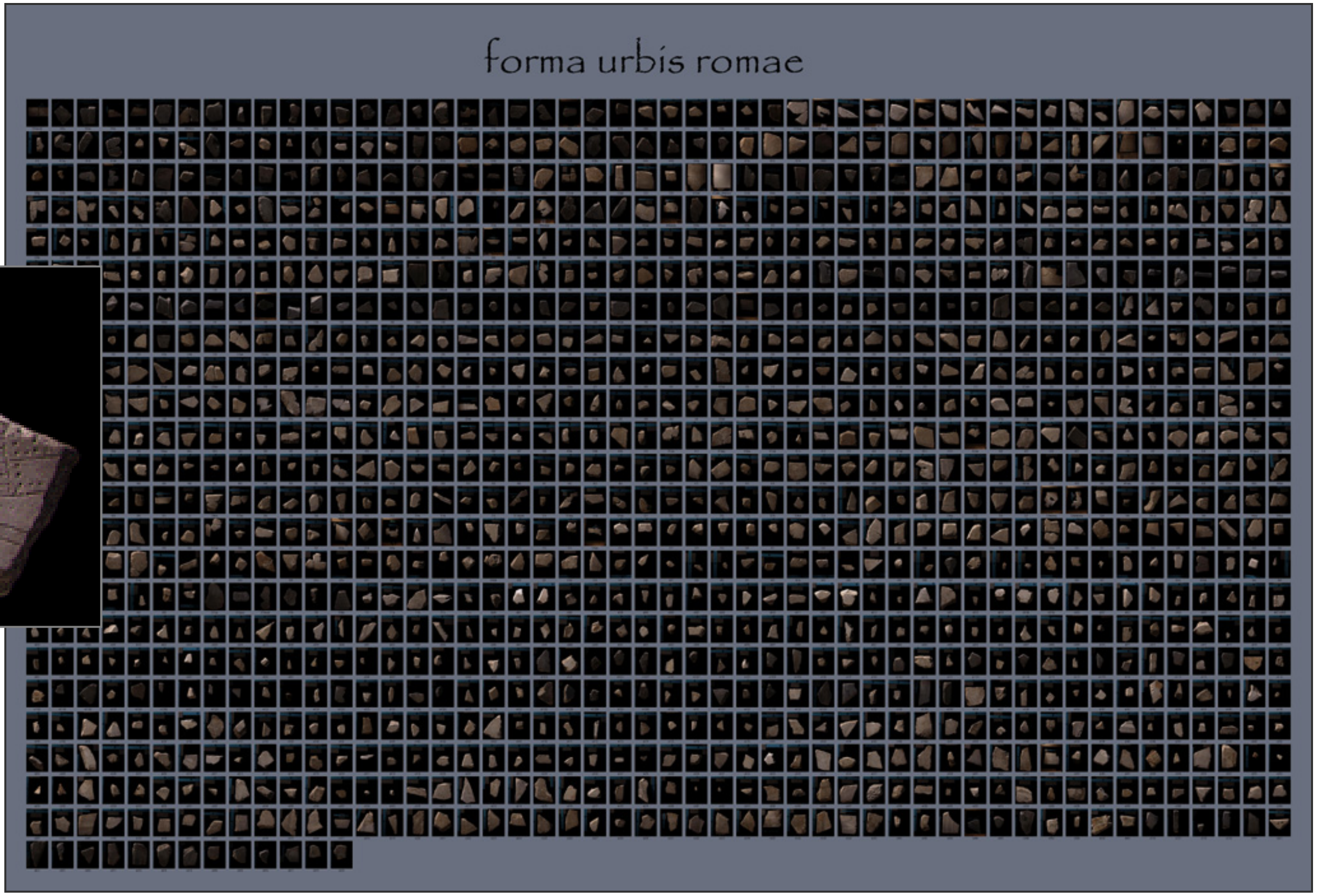


Applications: archaeology

- "forma urbis romae" project

Fragments of the City: Stanford's Digital Forma Urbis Romae Project
David Koller, Jennifer Trimble, Tina Najbjerg, Natasha Gelfand, Marc Levoy
Proc. Third Williams Symposium on Classical Architecture, Journal of Roman Archaeology supplement, 2006.

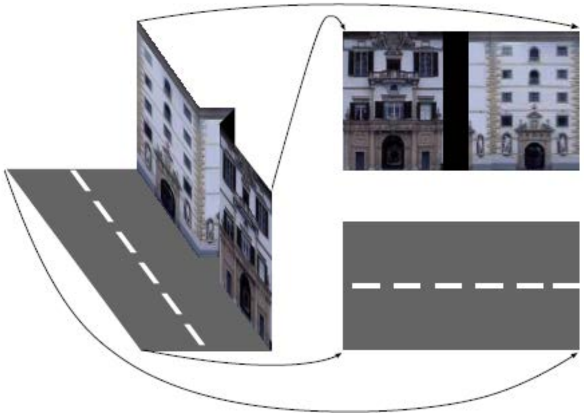
1186 fragments



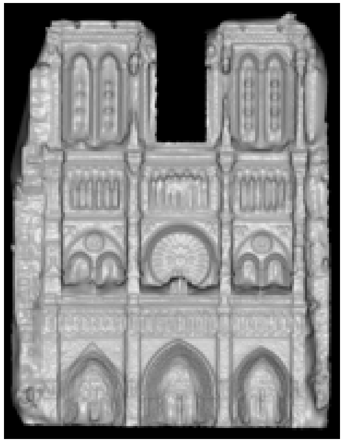
Applications: large scale modelling

[Furukawa10]

[Pollefeys08]

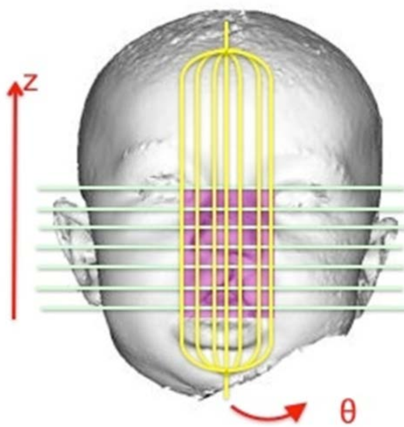


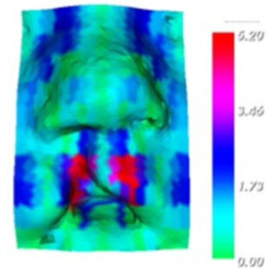
[Cornelis08]



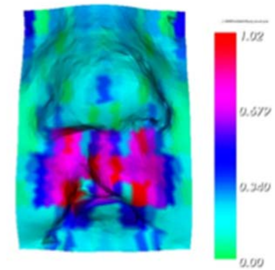
[Goesele07]

Applications: Medicine

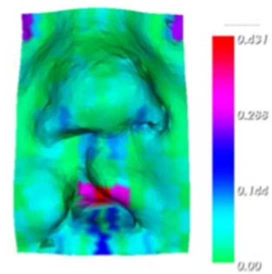




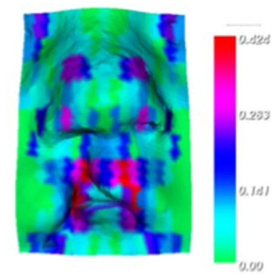
(a) Radius difference



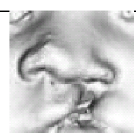
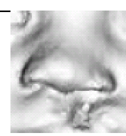
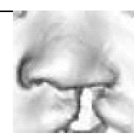
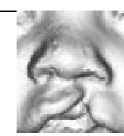
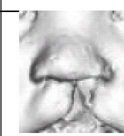
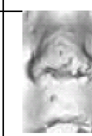
(b) Angle difference



(c) Curvature difference

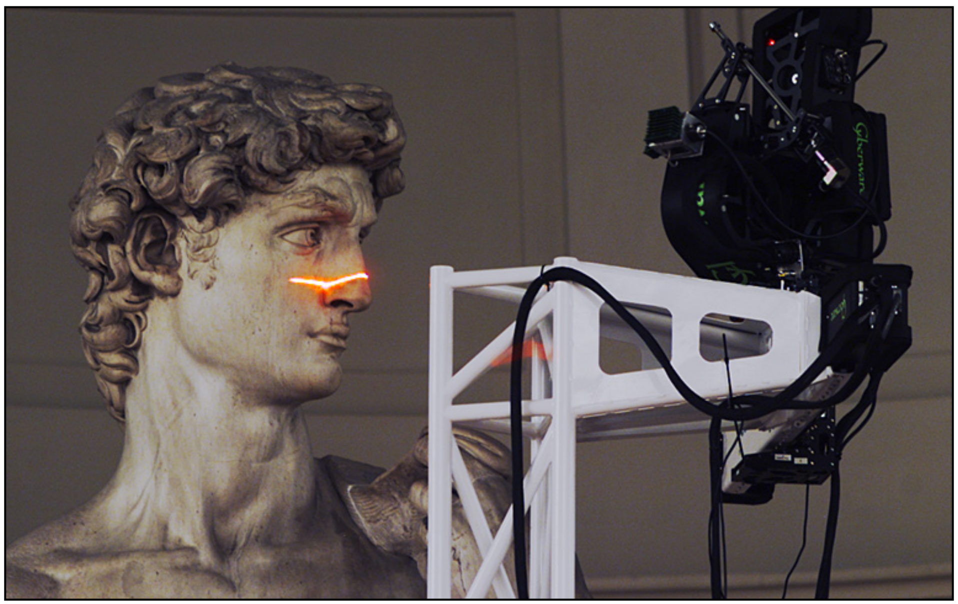


(d) Edge difference

expert's order	1	2	3	4	5	6	7	8	9	10
images										
learning	1	3	2	4	5	6	8	9	7	10
a-lmk	1	2	3	5	6	4	8	7	9	10
mirror	1	2	4	8	5	6	9	3	7	10
m-lmk	1	2	3	4	5	6	9	7	10	8
plane	1	2	3	5	4	6	7	9	10	8

Scanning technologies

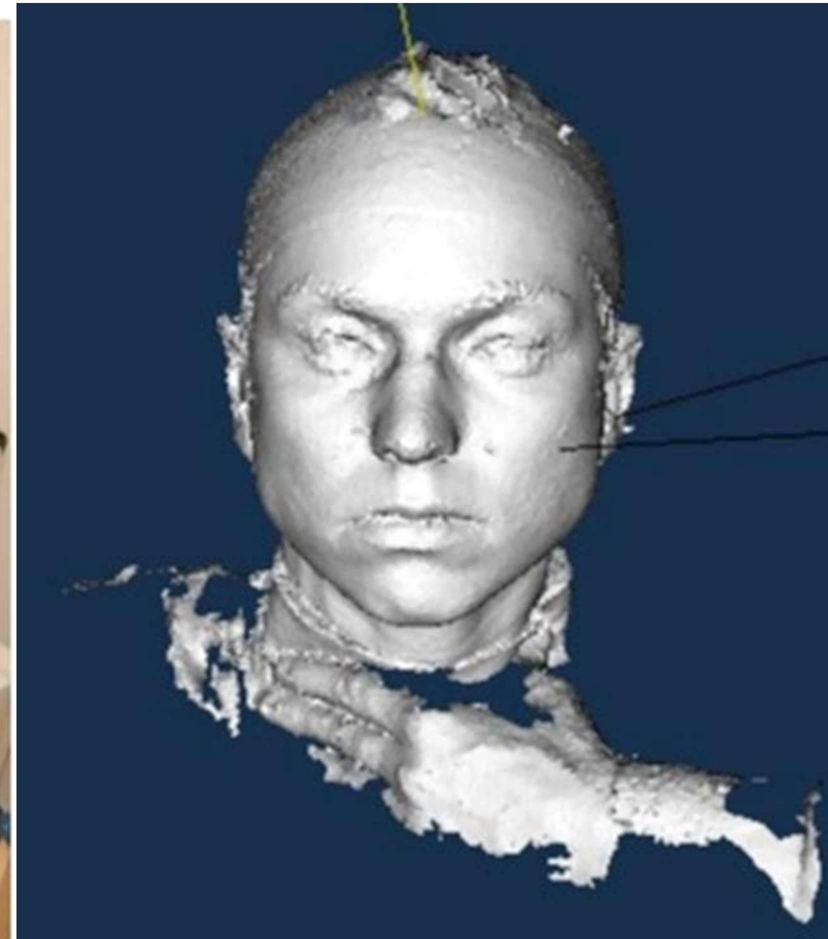
- Laser scanner, coordinate measuring machine
  - Very accurate
  - Very Expensive
  - Complicated to use



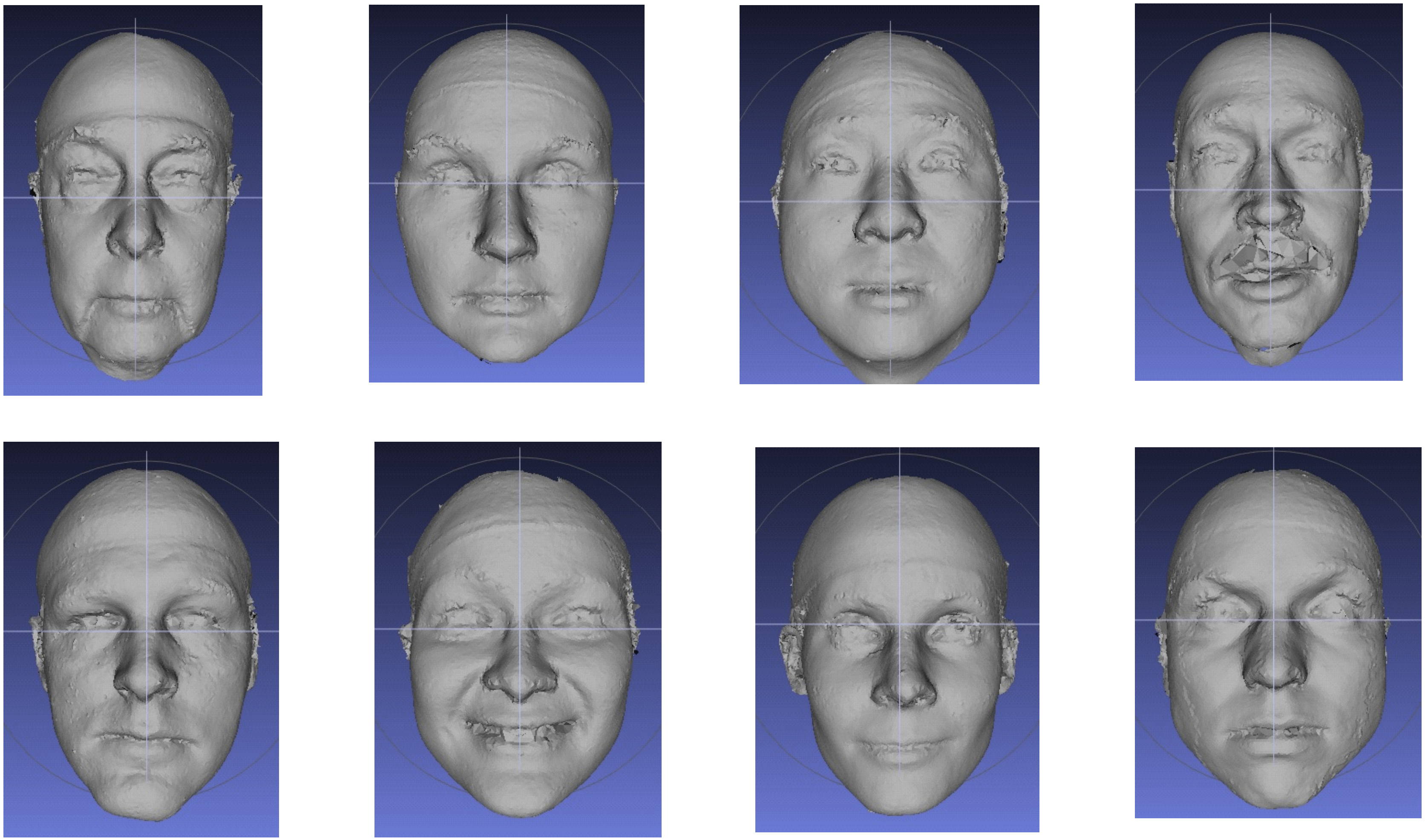
Minolta

Contura CMM

Medical Scanning System

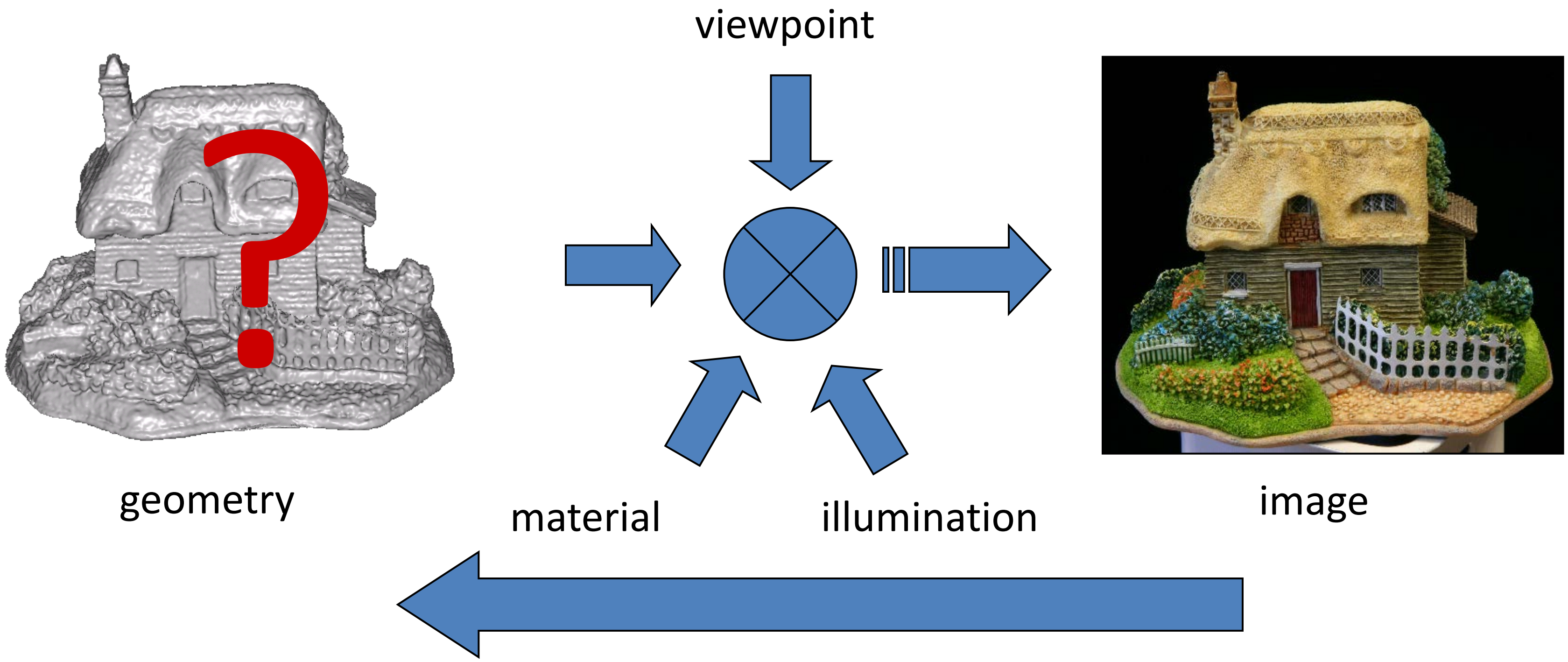


The "Us" Data Set (subset)



3d shape from photographs

“Estimate a 3d shape that would generate the input photographs given the same material, viewpoints and illumination”



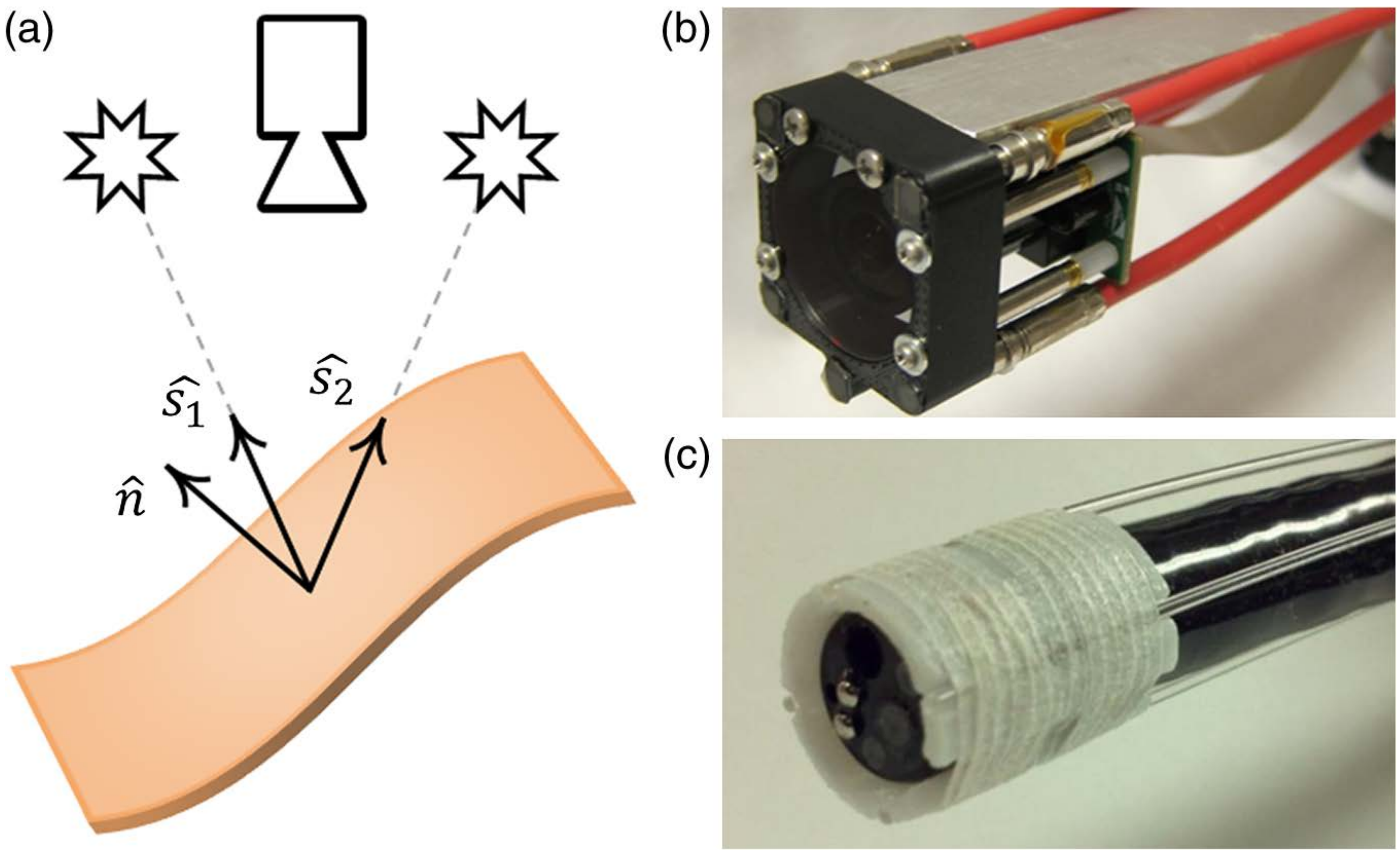
Photometric Stereo

- Estimate the surface normals of a given scene given multiple 2D images taken from the *same* viewpoint, but under *different lighting* conditions.
- Basic photometric stereo required a Lambertian reflectance model:

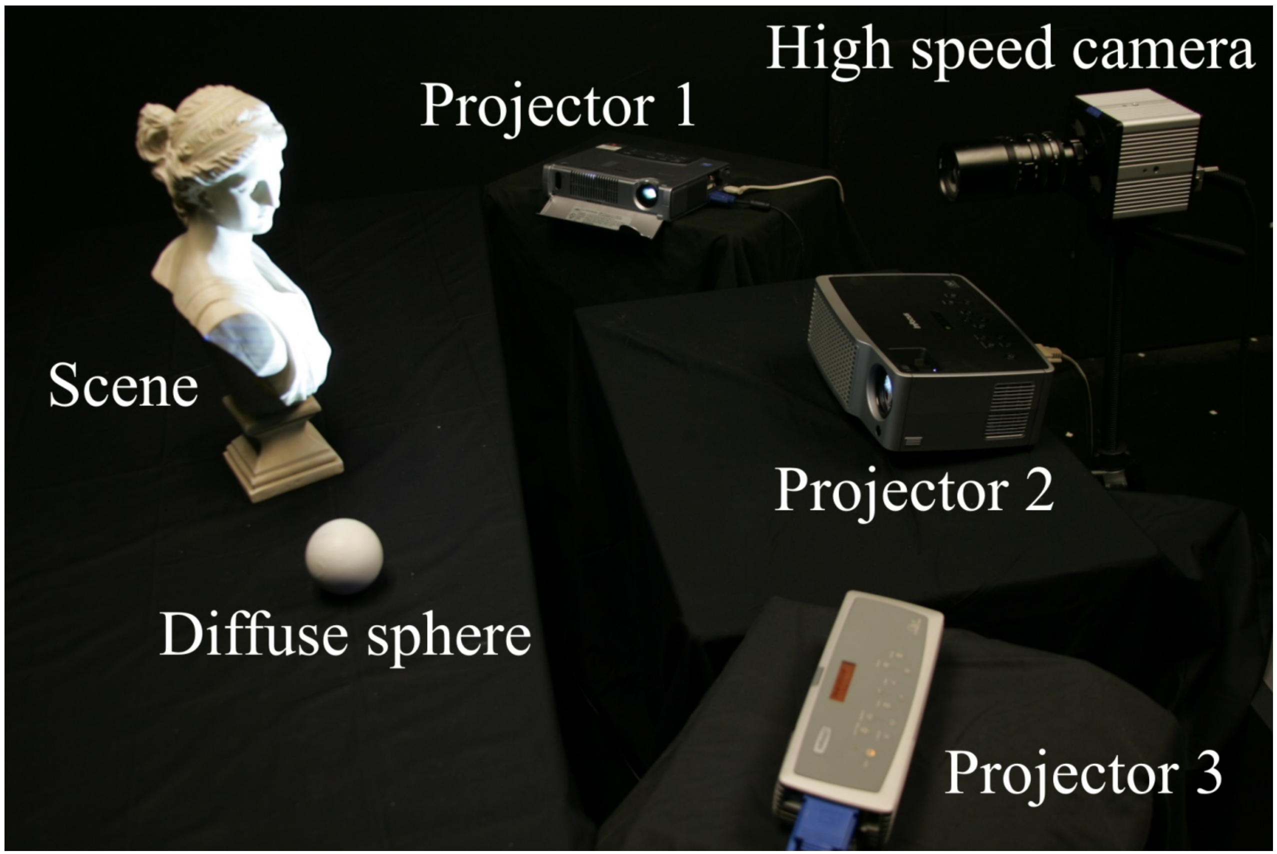
$$I = \rho \, \mathbf{n} \cdot \mathbf{v}$$

where I is pixel intensity, **n** is the normal, **v** is the lighting direction, and ρ is diffuse albedo constant, which is a reflection coefficient.

Basic Photometric Stereo



Basic Photometric Stereo



Basic Photometric Stereo

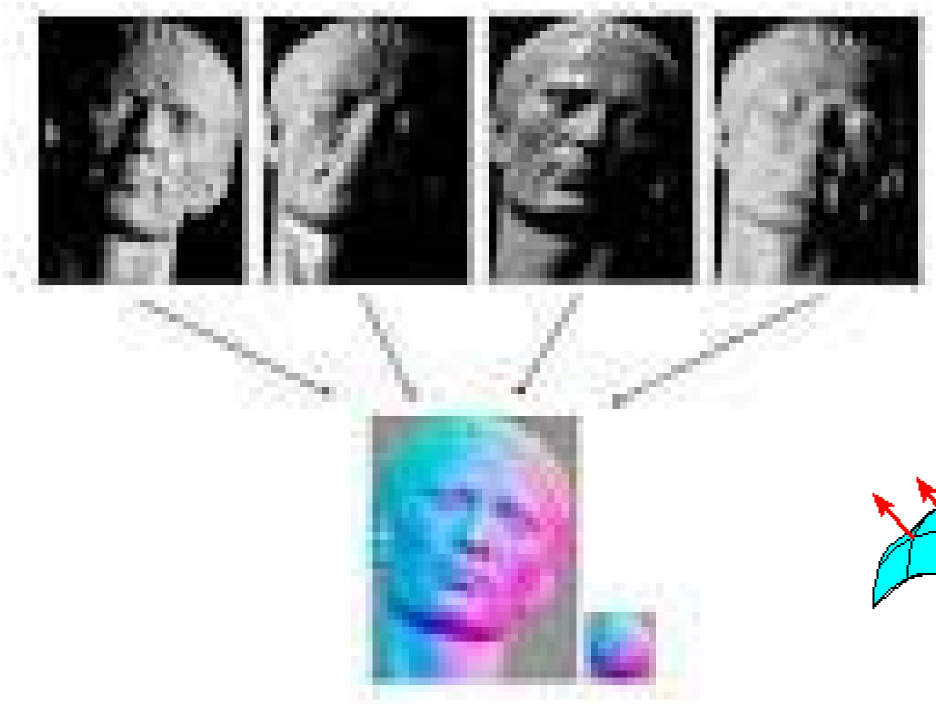
- K light sources
- Lead to K images $R_1(p,q)$, ...,$R_K(p,q)$ each from just one of the light sources being on
- For any (p,q), we get K intensities $I_1,...I_K$
- Leads to a set of linear equations of the form

$$I_k = \rho \mathbf{n} \bullet \mathbf{v}_k$$

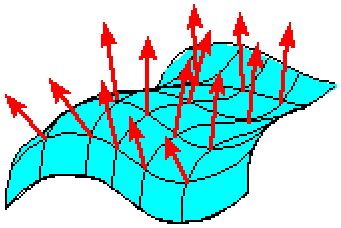
- Solving leads to a surface normal map.

Photometric Stereo

Inputs



3D normals



3d shape from photographs

Photograph based 3d reconstruction is:

- ✓ practical
- ✓ fast
- ✓ non-intrusive
- ✓ low cost
- ✓ Easily deployable outdoors
- ✗ “low” accuracy
- ✗ Results depend on material properties

Reconstruction

- Generic problem formulation: given several images of the same object or scene, compute a representation of its 3D shape

Reconstruction

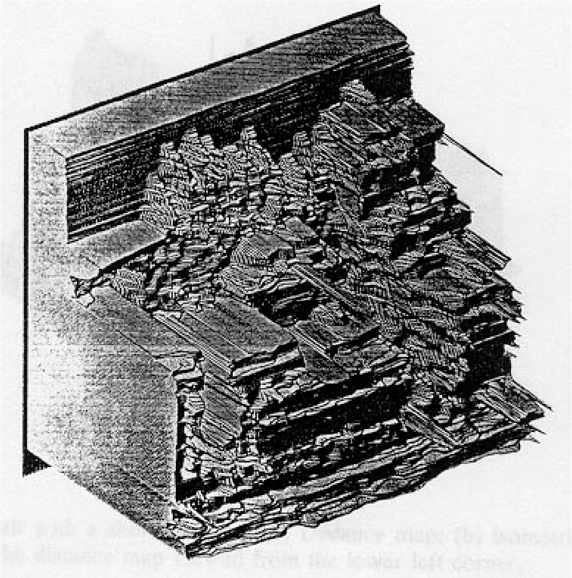
- Generic problem formulation: given several images of the same object or scene, compute a representation of its 3D shape
- “Images of the same object or scene”
  - Arbitrary number of images (from two to thousands)
  - Arbitrary camera positions (camera network or video sequence)
  - Calibration may be initially unknown
- “Representation of 3D shape”
  - Depth maps
  - Meshes
  - Point clouds
  - Patch clouds
  - Volumetric models
  - Layered models

Multiple-baseline stereo

I1

I2

I10

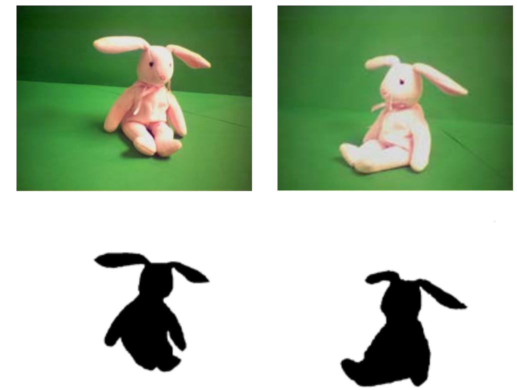


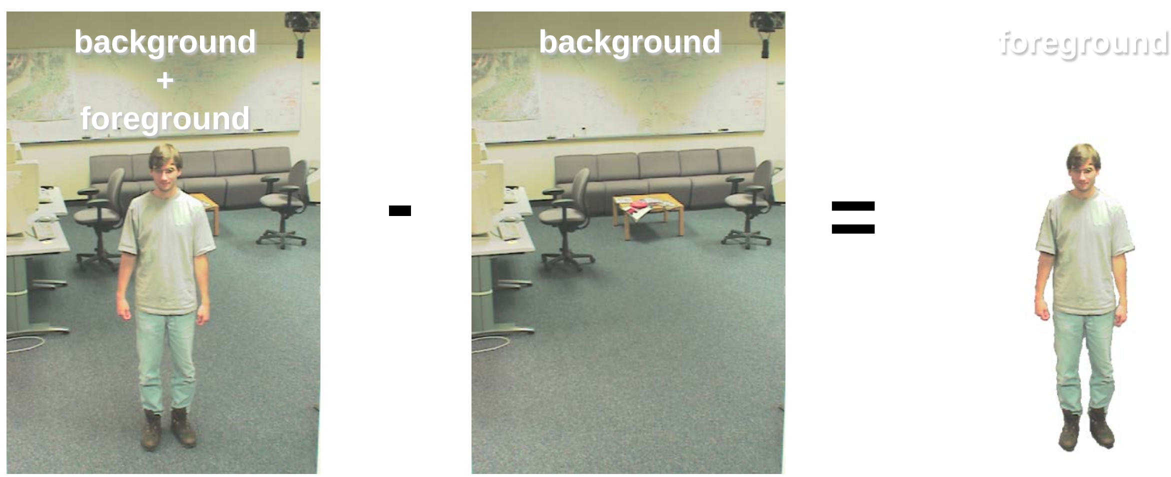
M. Okutomi and T. Kanade, "A Multiple-Baseline Stereo System," IEEE Trans. on Pattern Analysis and Machine Intelligence, 15(4):353-363 (1993).

Reconstruction from silhouettes

Can be computed robustly

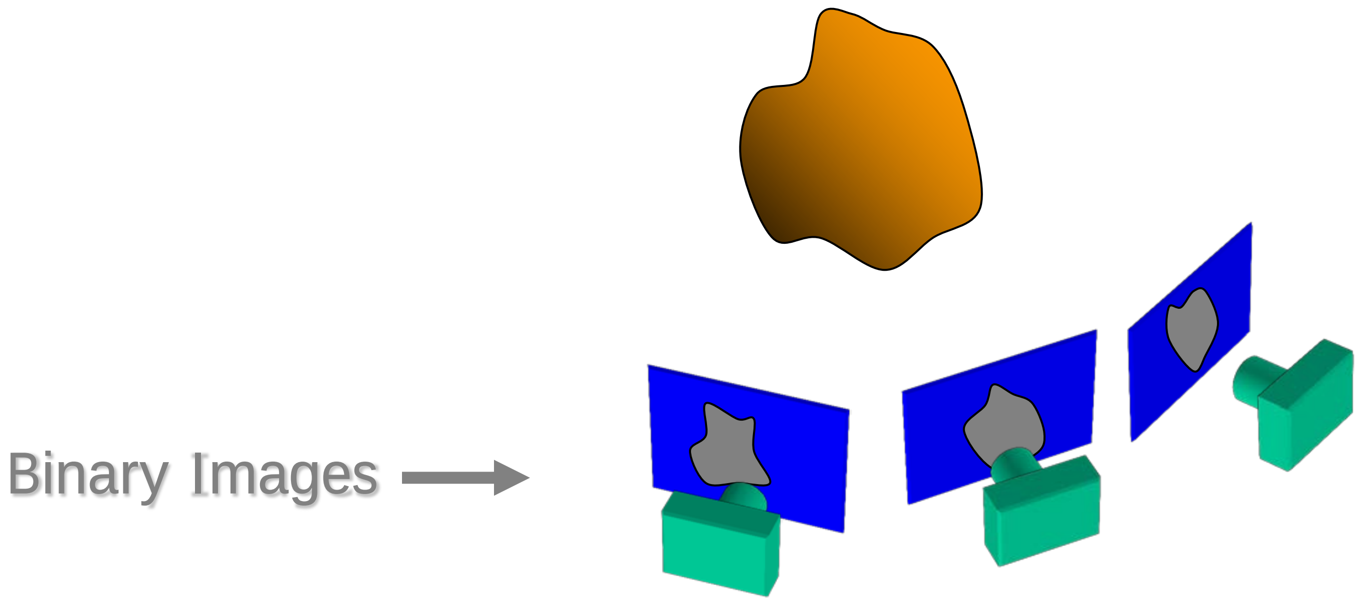
Can be computed efficiently





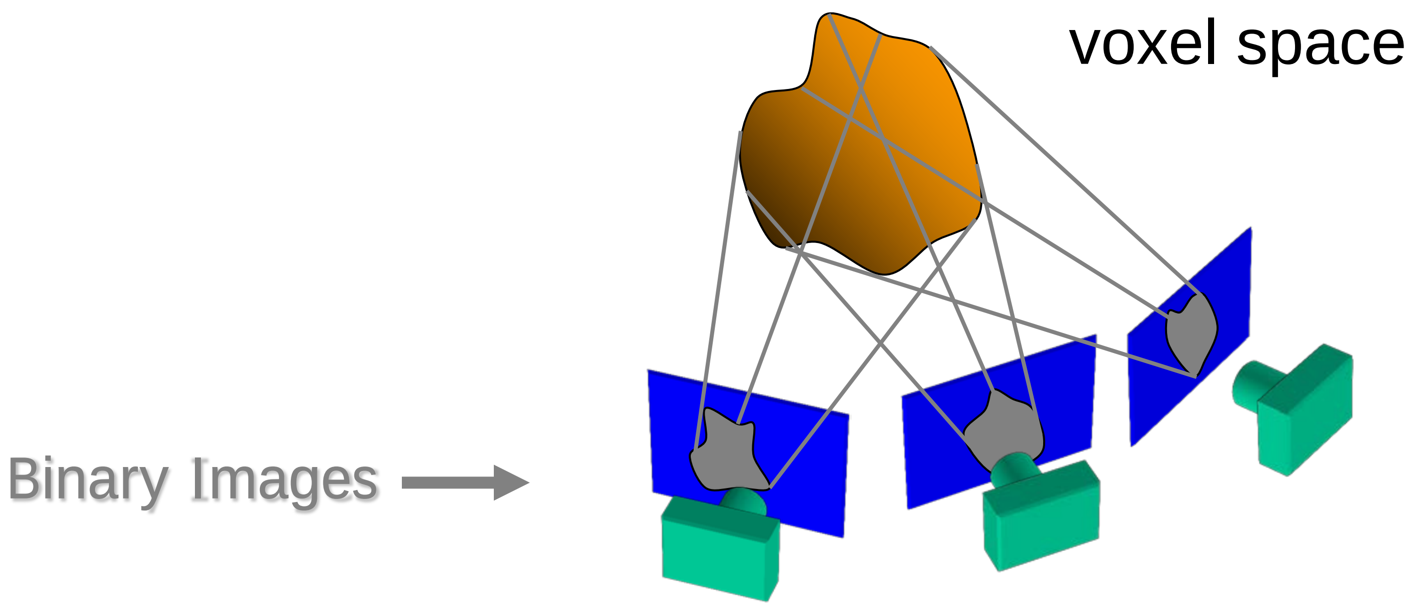
Reconstruction from Silhouettes

- The case of binary images: a voxel is photo-consistent if it lies inside the object's silhouette in **all** views



Reconstruction from Silhouettes

- The case of binary images: a voxel is photo-consistent if it lies inside the object's silhouette in all views

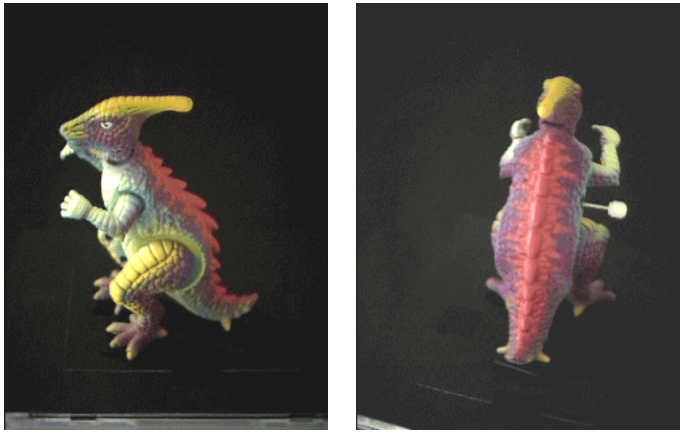


Finding the silhouette-consistent shape (*visual hull*):

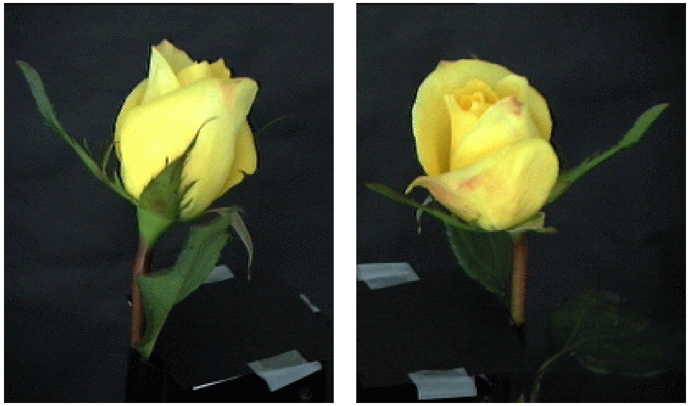
- *Backproject* each silhouette
- Intersect backprojected volumes

Calibrated Image Acquisition

Calibrated Turntable

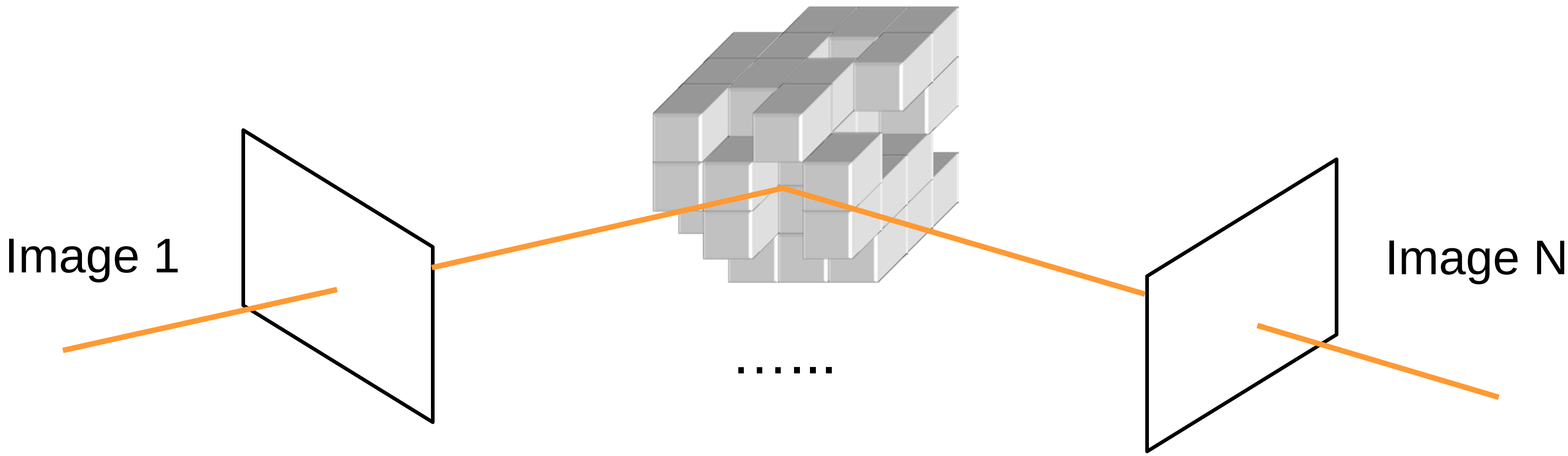


Selected Dinosaur Images



Selected Flower Images

Space Carving in General



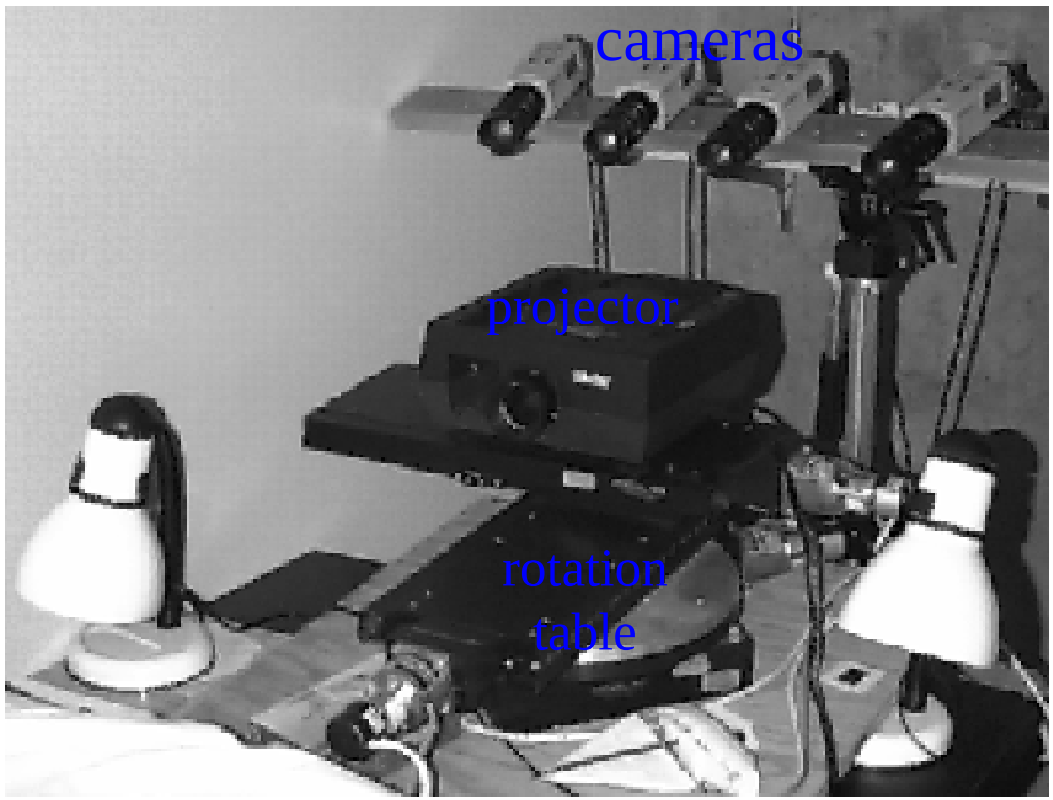
Space Carving Algorithm

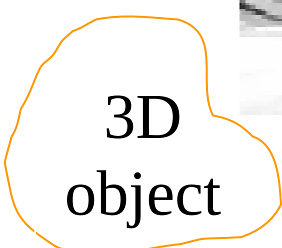
- Initialize to a volume V containing the true scene
- Choose a voxel on the outside of the volume
- Project to visible input images
- Carve if not photo-consistent (inside object's silhouette)
- Repeat until convergence

K. N. Kutulakos and S. M. Seitz, **A Theory of Shape by Space Carving**, *ICCV* 1999

Our 4-camera light-striping stereo system

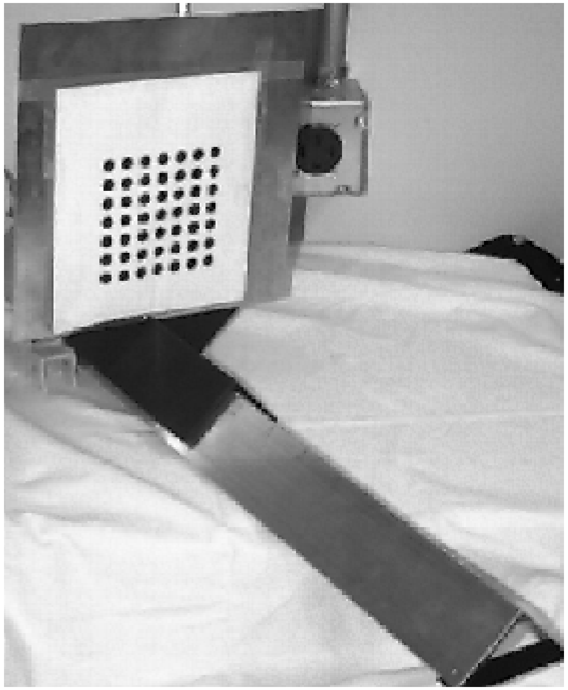
(now deceased)





Calibration Object

The idea is to snap images at different depths and get a lot of 2D-3D point correspondences.



Surface Modeling and Display from Range and Color Data

Kari	Pulli	UW
Michael	Cohen	MSR
Tom	Duchamp	UW
Hugues	Hoppe	MSR
John	McDonald	UW
Linda	Shapiro	UW
Werner	Stuetzle	UW

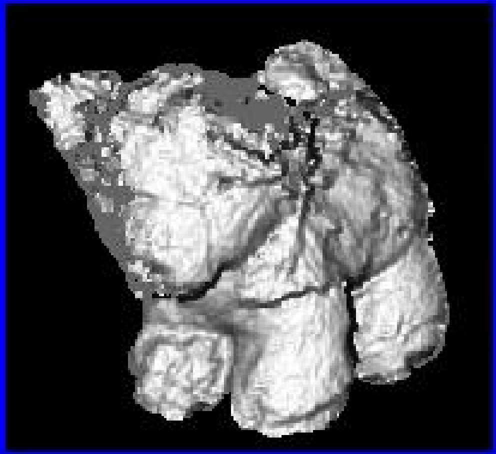
UW = University of Washington
Seattle, WA USA

MSR = Microsoft Research
Redmond, WA USA

Introduction

Goal

- develop robust algorithms for constructing 3D models from range & color data
- use those models to produce realistic renderings of the scanned objects



Surface Reconstuction

Step 1: Data acquisition

Obtain range data that covers the object. Filter, remove background.

Step 2: Registration

Register the range maps into a common coordinate system.

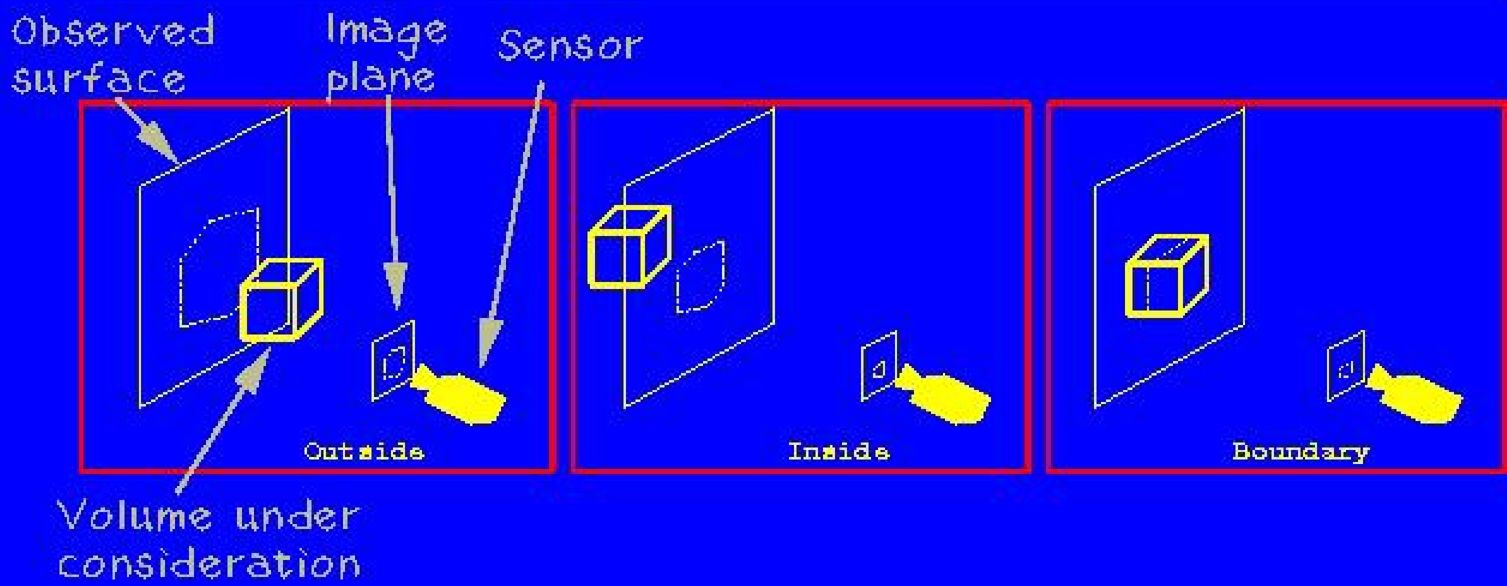
Step 3: Integration

Integrate the registered range data into a single surface representation.

Step 4: Optimization

Fit the surface more accurately to the data, simplify the representation.

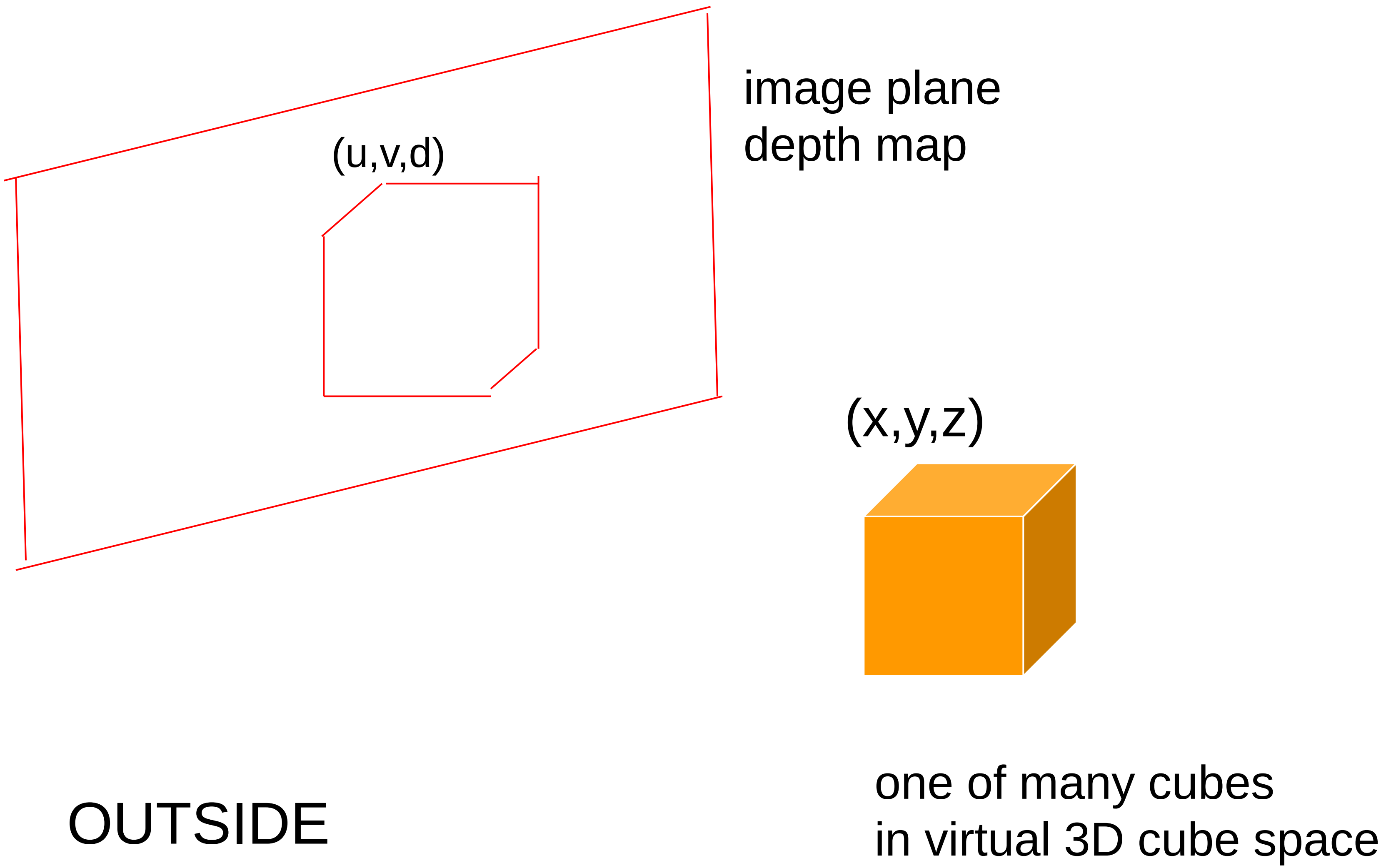
Carve space in cubes



Label cubes

- Project cube to image plane (hexagon)
- Test against data in the hexagon

3D space is made up of many cubes.

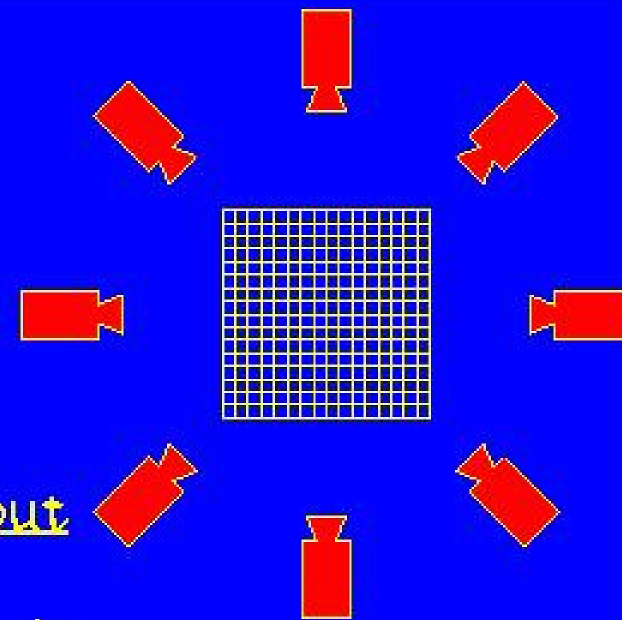


Several views

Processing order:

FOR EACH cube

FOR EACH view



Rules:

any view thinks cube's out
=> it's out

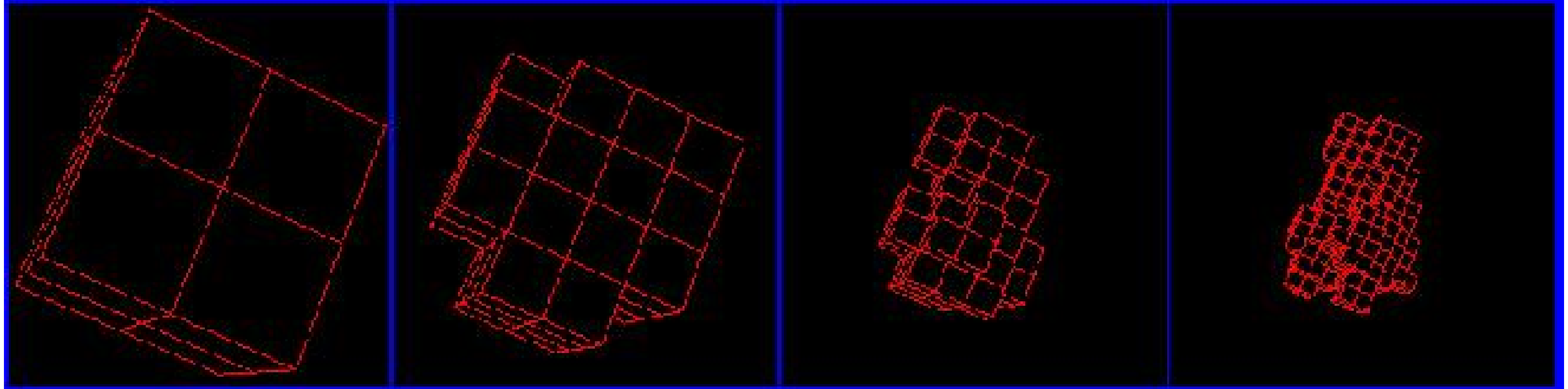
every view thinks cube's in
=> it's in

else
=> it's at boundary

Hierarchical space carving

- Big cubes => fast, poor results
- Small cubes => slow, more accurate results
- Combination = octrees

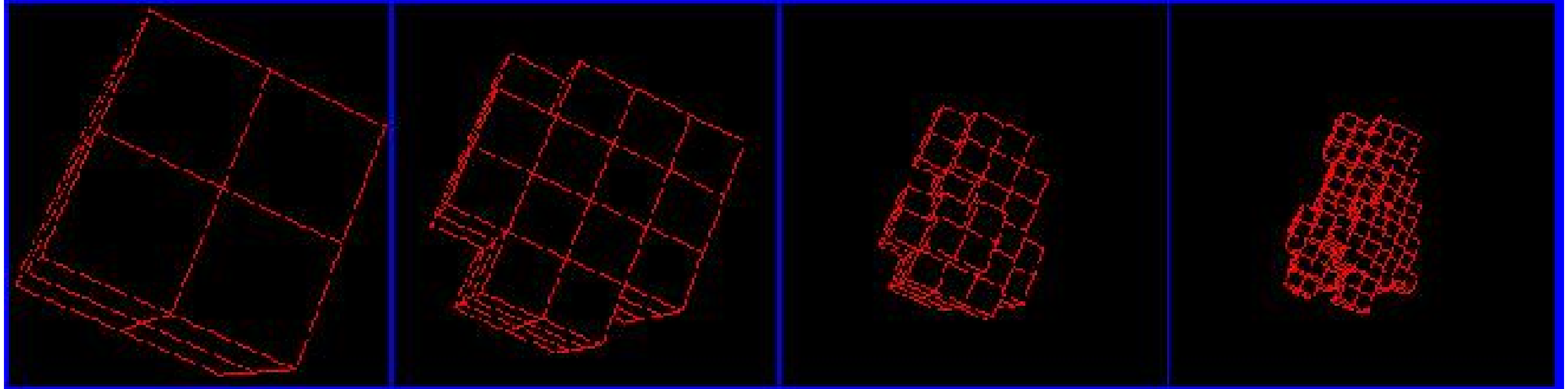
RULES:
- cube's out => done
- cube's in => done
- else => recurse



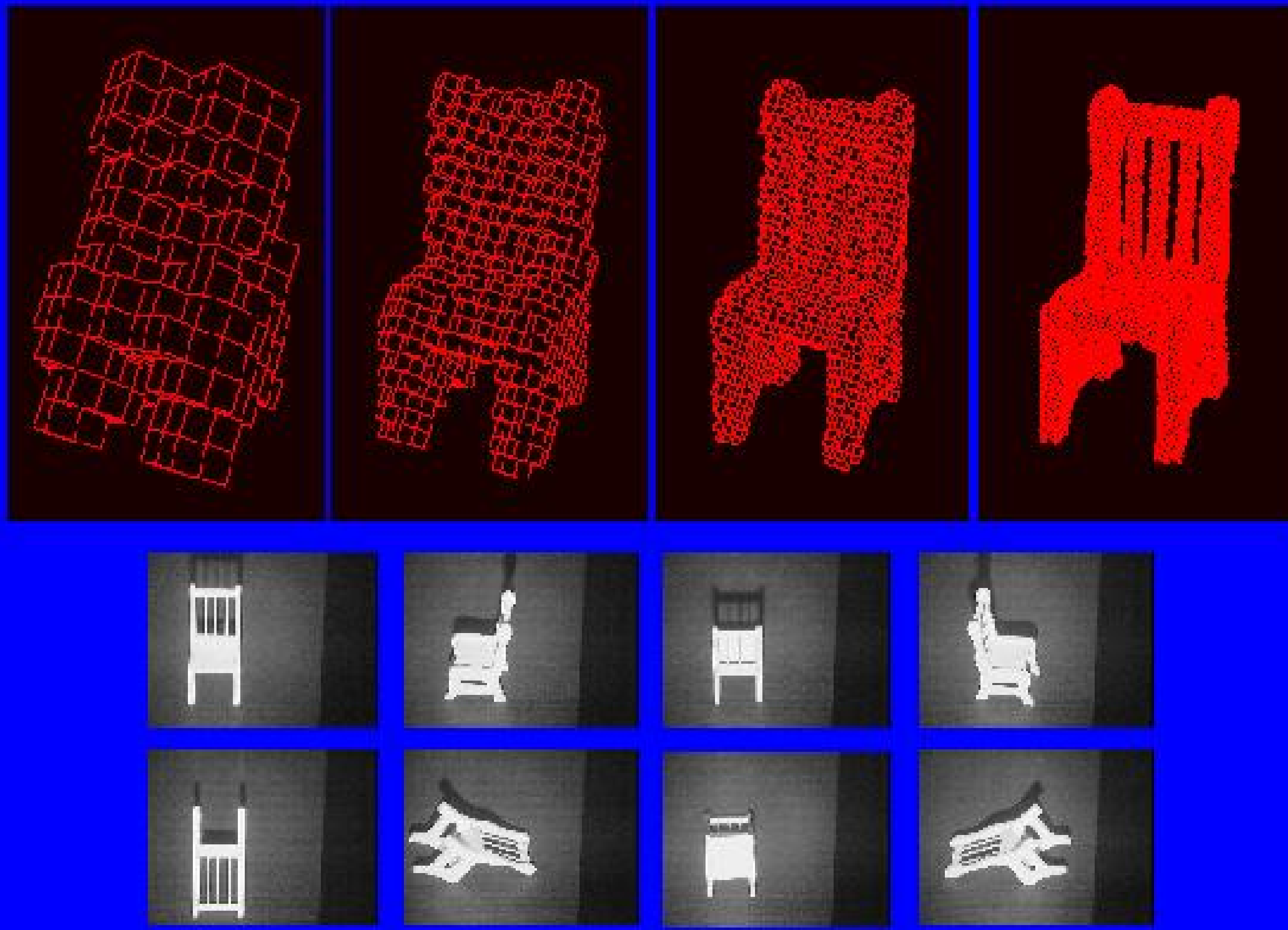
Hierarchical space carving

- Big cubes => fast, poor results
- Small cubes => slow, more accurate results
- Combination = octrees

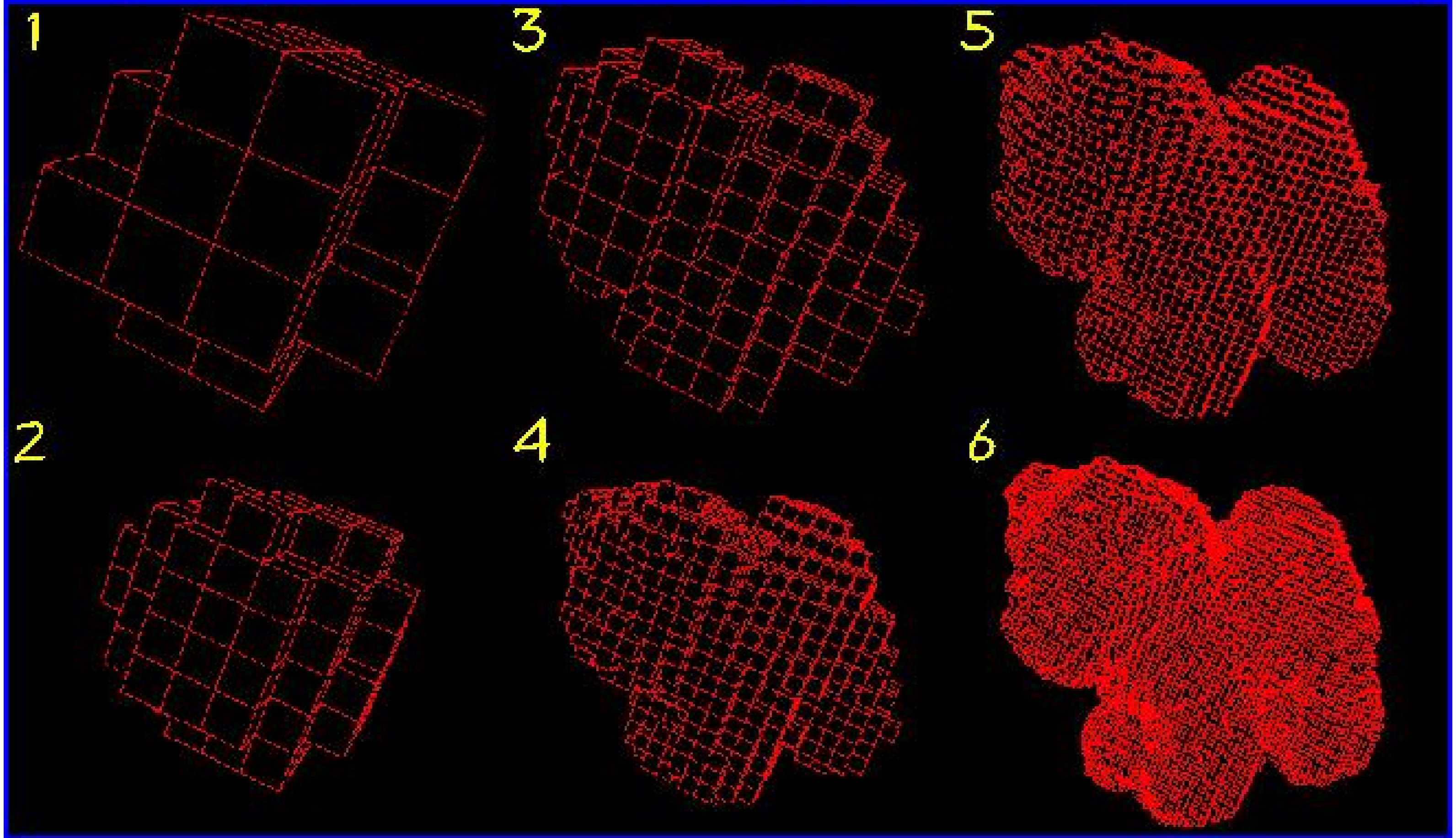
RULES:
- cube's out => done
- cube's in => done
- else => recurse



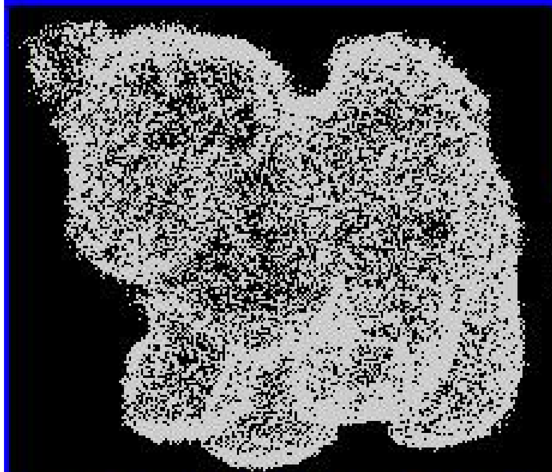
The rest of the chair



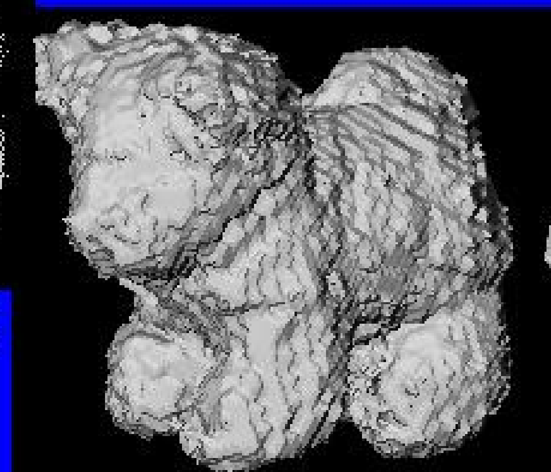
Same for a husky pup



Optimizing the dog mesh



Registered points

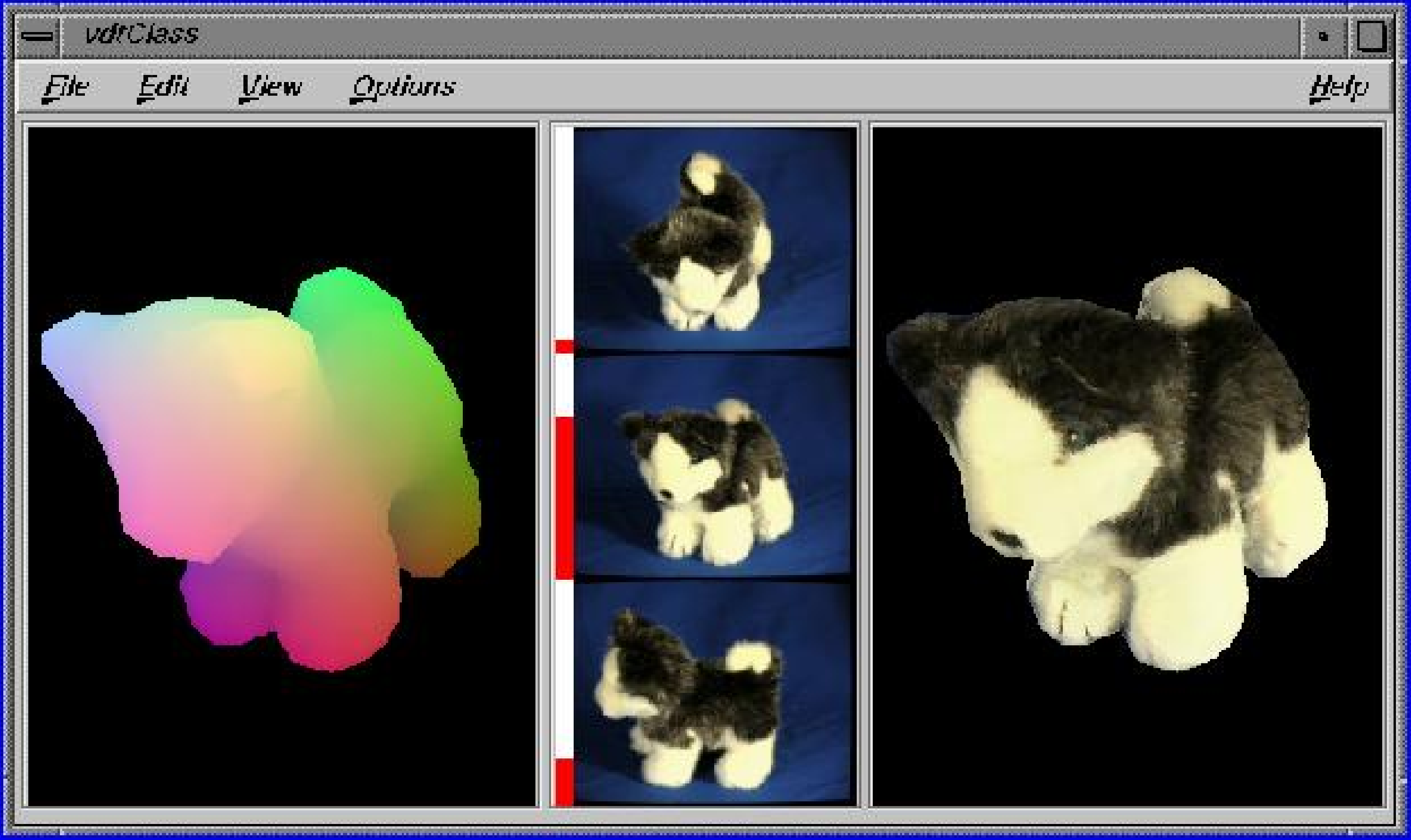


Initial mesh

Optimized mesh

View dependent texturing

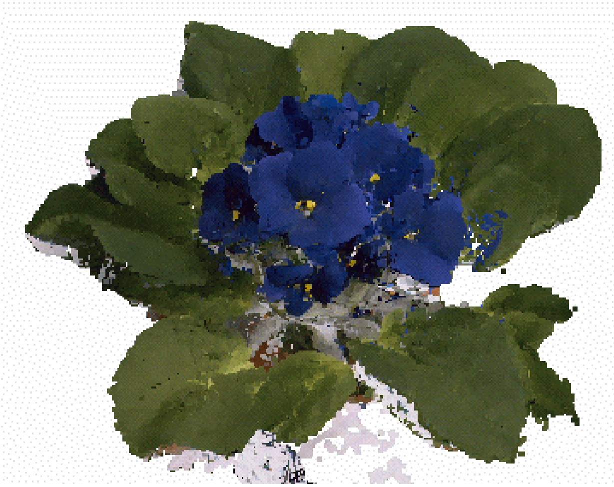
Our viewer



More: Space Carving Results: African Violet

Input Image (1 of 45)

Reconstruction



Reconstruction

Reconstruction

Source: S. Seitz

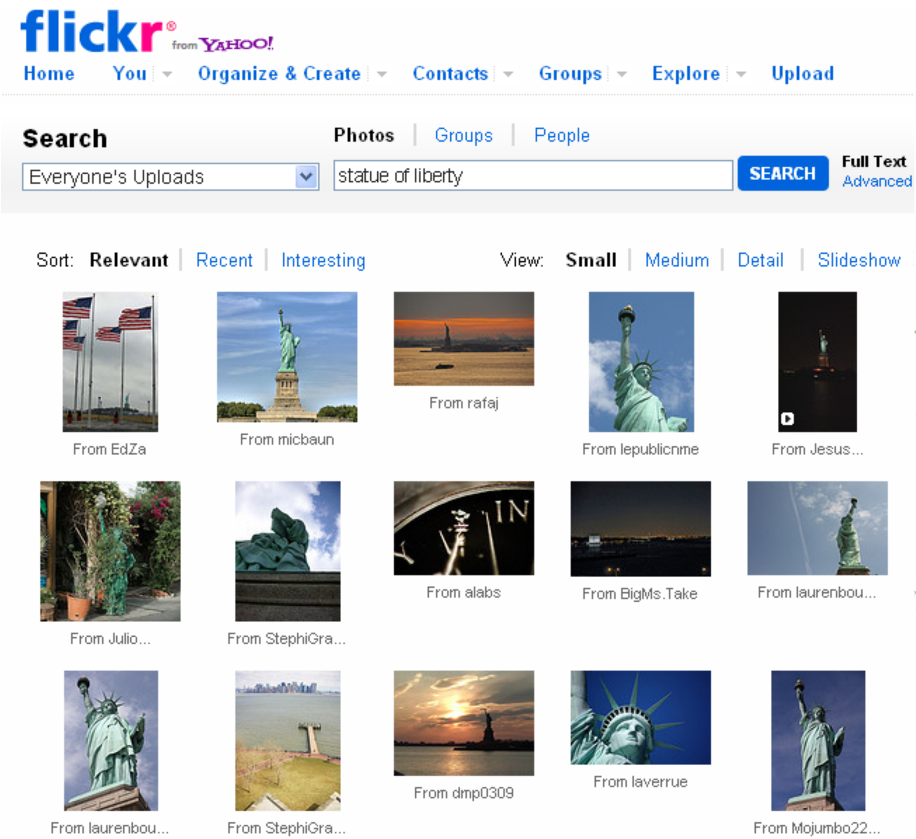
More: Space Carving Results: Hand

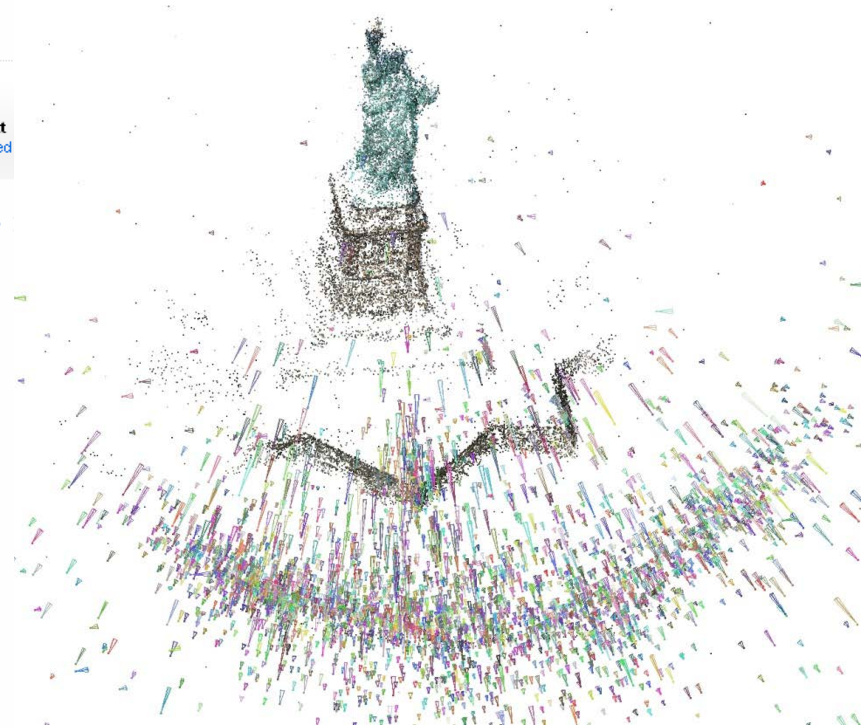
Input Image
(1 of 100)

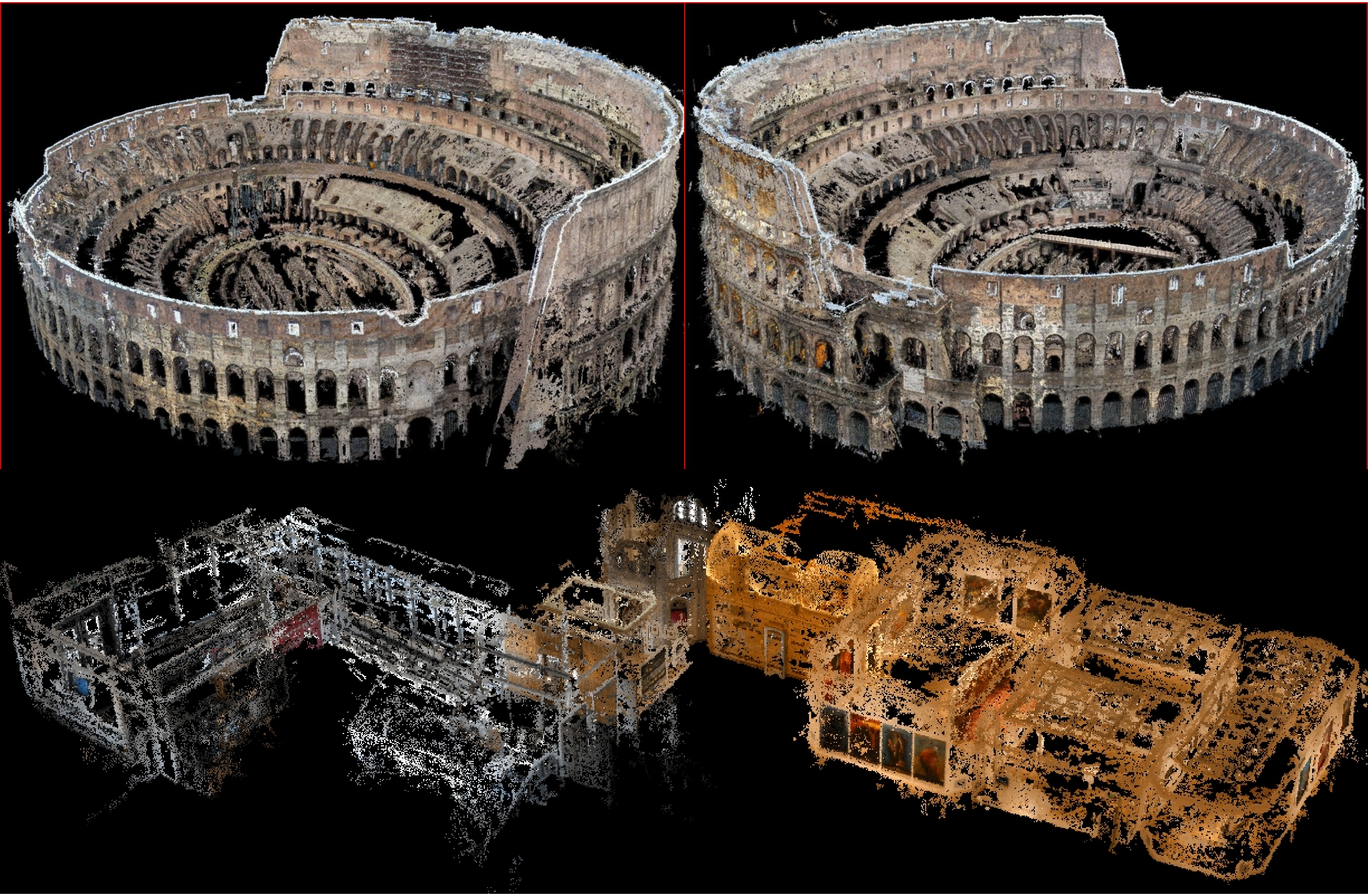
Views of Reconstruction

Stereo from community photo collections

- Up to now, we've always assumed that camera calibration is known
- For photos taken from the Internet, we need *structure from motion* techniques to reconstruct both camera positions and 3D points



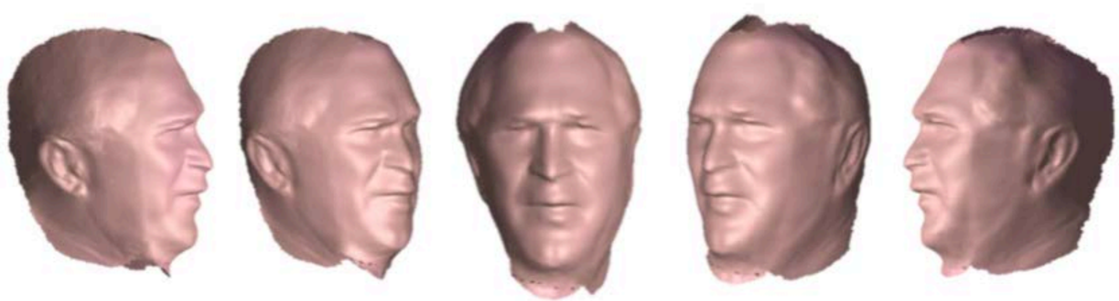




Head Reconstruction from Uncalibrated Internet Photos

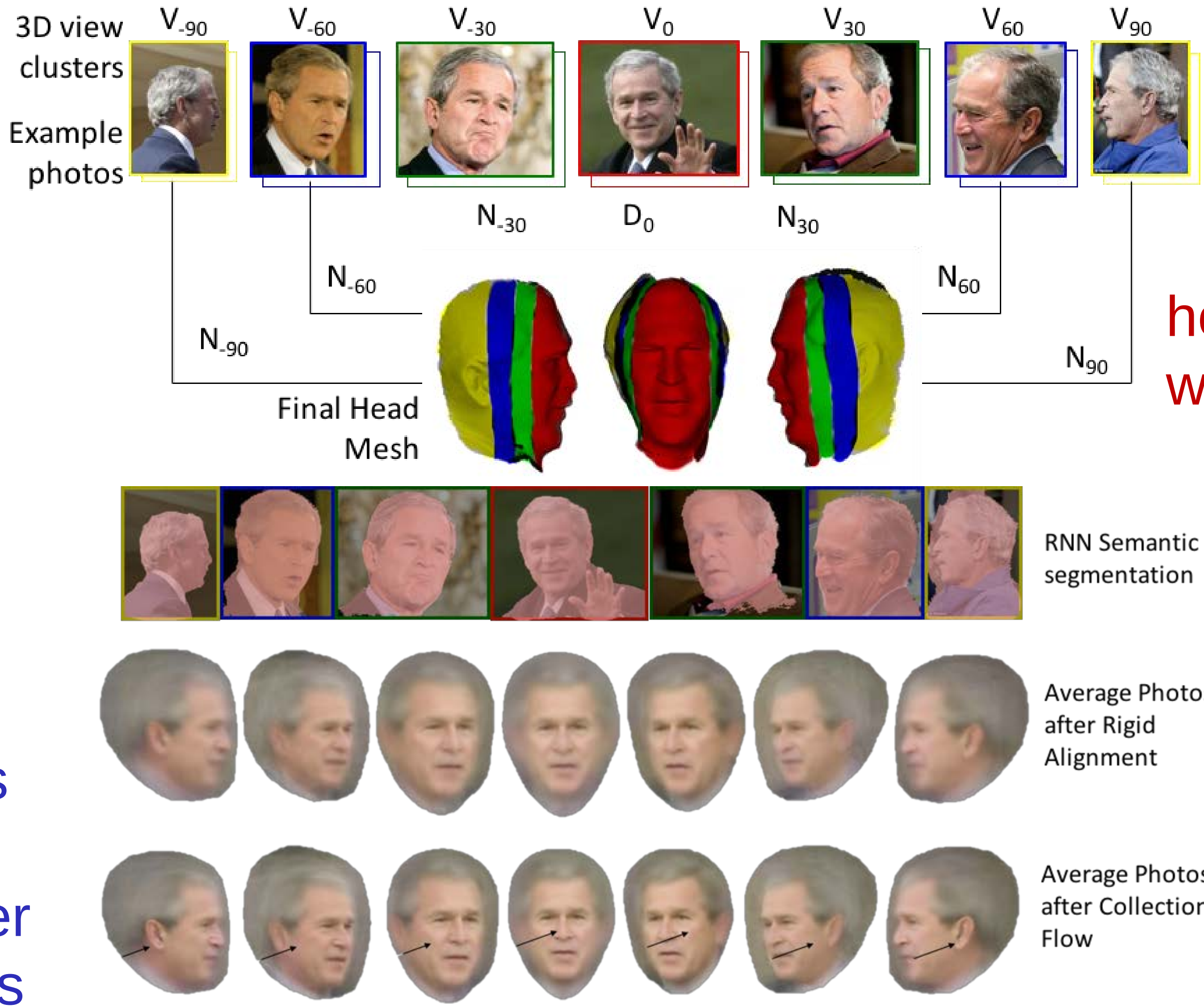
Input: Internet photos in different poses and expressions

Output: 3D model of the head



work of
Shu Liang

Pose Clusters



clustered in 2D

how they will fit in 3D

find the person

rigid head alignment

pixel level alignment

arrows show sharper regions

Photometric Stereo on frontal pose

Frontal view photos to n x p matrix Q. n: number of photos, p: number of pixels of the facial mask.

Using Singular Valued Decomposition, Q = LN, where L is the lighting and N is the surface normals. Only the 1st 4 components are used.

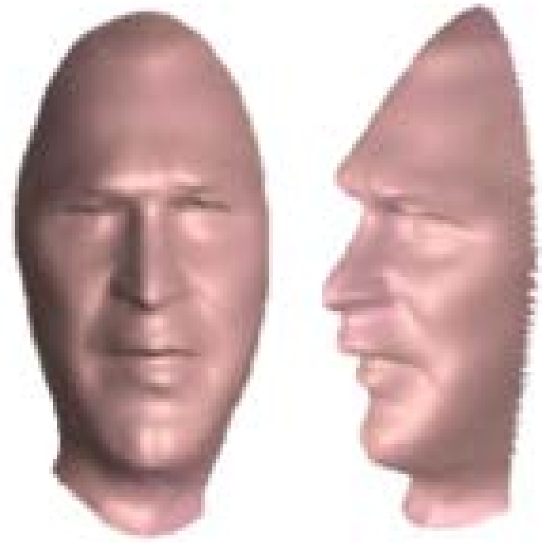
Expression normalization by low-rank approximation. The first row shows the warped images, the 2nd row shows the the low rank approximated images. Note how the lighting is mostly preserved, but the facial expression is normalized.

Kemelmacher-Shlizerman I. et al., Face Reconstruction in the Wild, ICCV 2011

Photometric Stereo on frontal pose

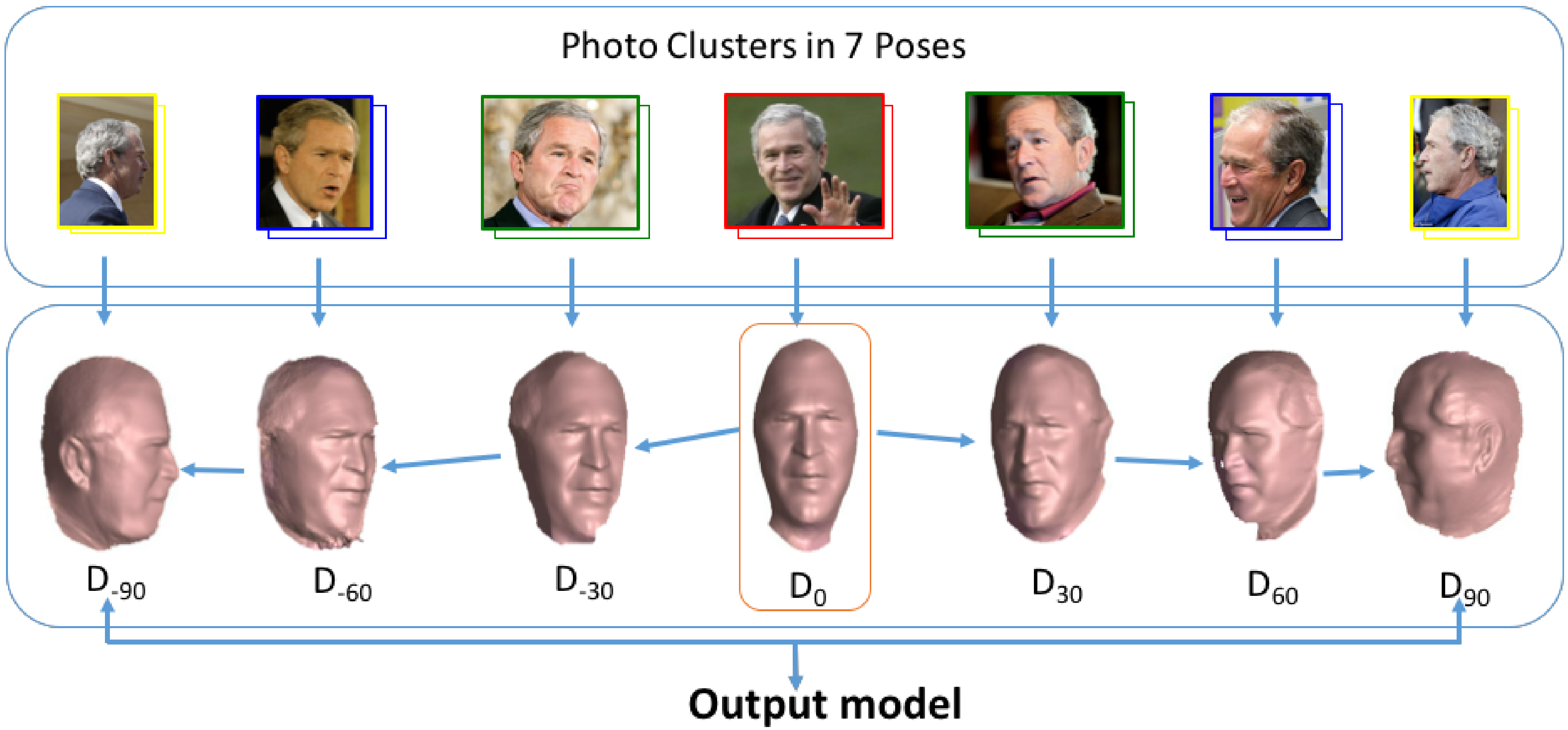
1. Use SVD to compute 3D surface normals

2. Compute 3d depth map from the surface normals



3D shapes generated from frontal pose shown from 2 different views

Boundary-Value Growing



- There are ambiguities in scale between the poses.
- These have to be resolved by a process that uses neighboring poses to solve ambiguities.

Results

Image [redacted]