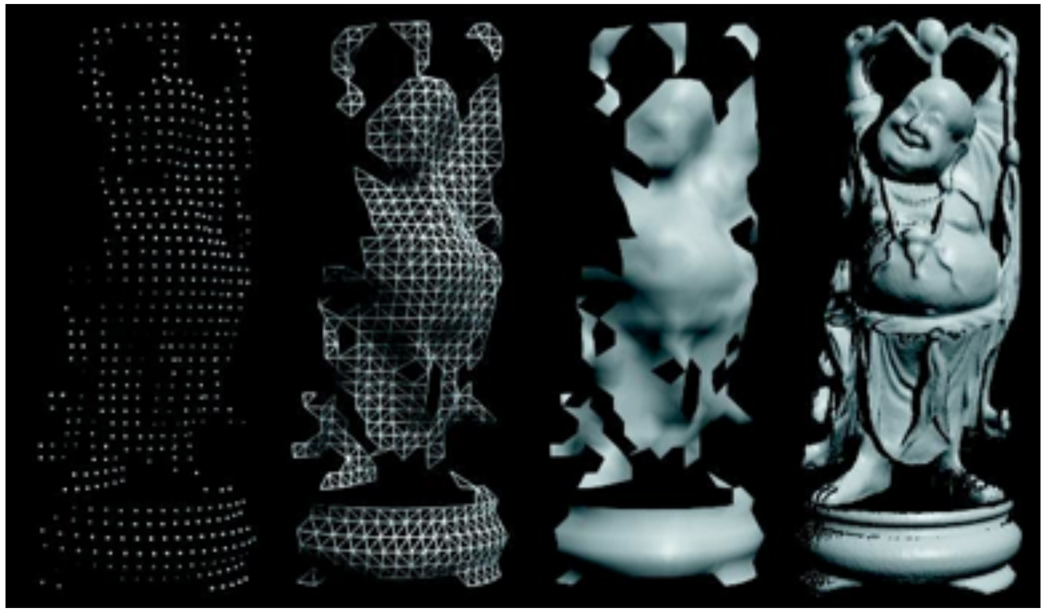
CSCI 599: **Digital Geometry Processing**



4.1 **3D Scanning**

Hao Li

http://cs599.hao-li.com

Administrative

- **Exercise 2:** this thursday after surface registration
- My first office hours later from 2pm to 4pm

2D Imaging Pipeline

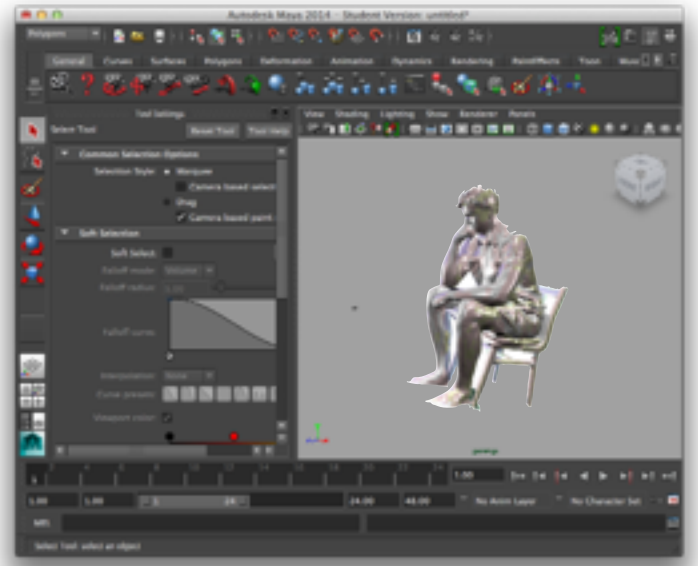
2D capture

2D processing/editing

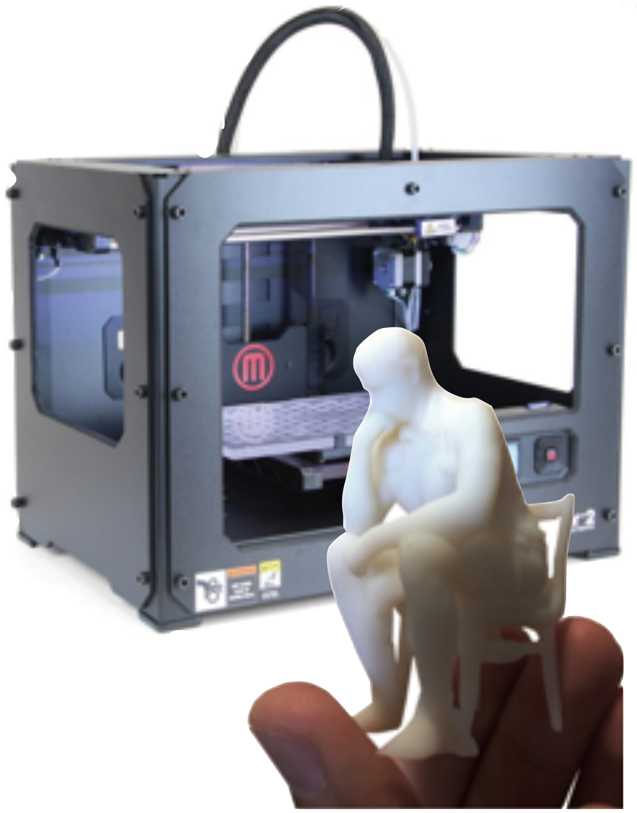
2D printing

3D Scanning Pipeline

3D scanning



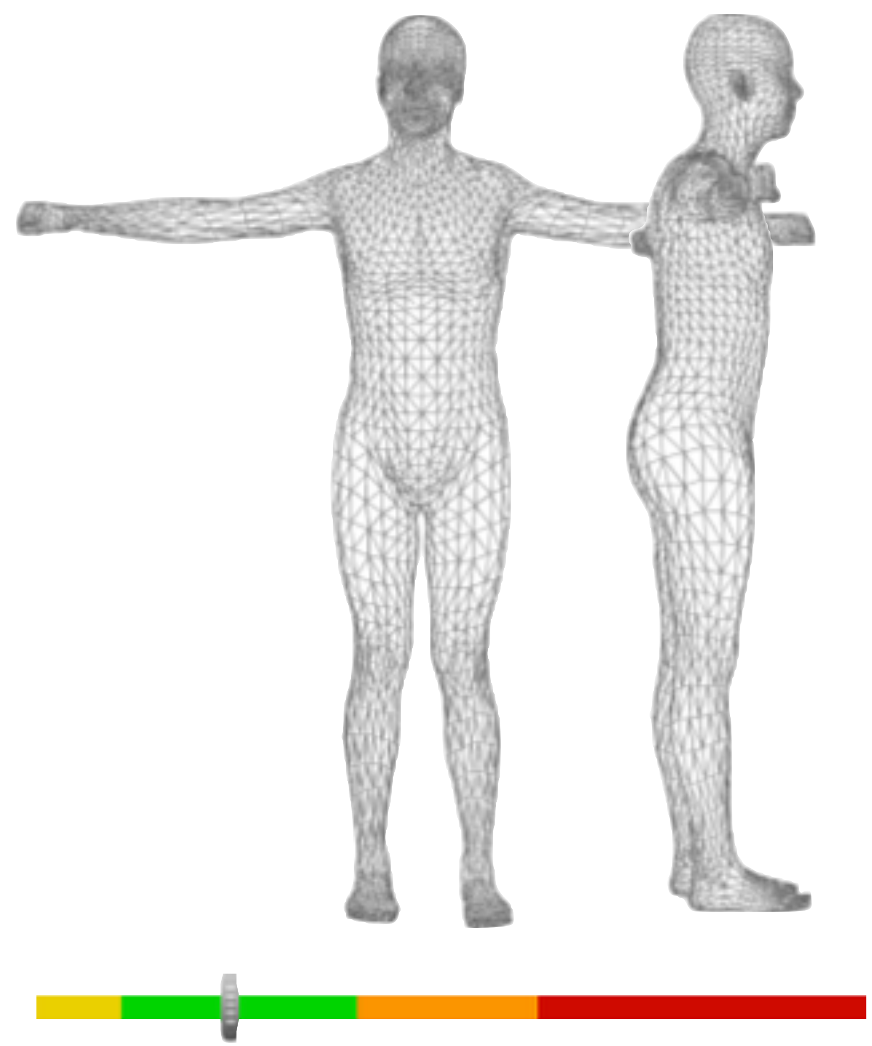
3D processing/editing



3D printing

Applications

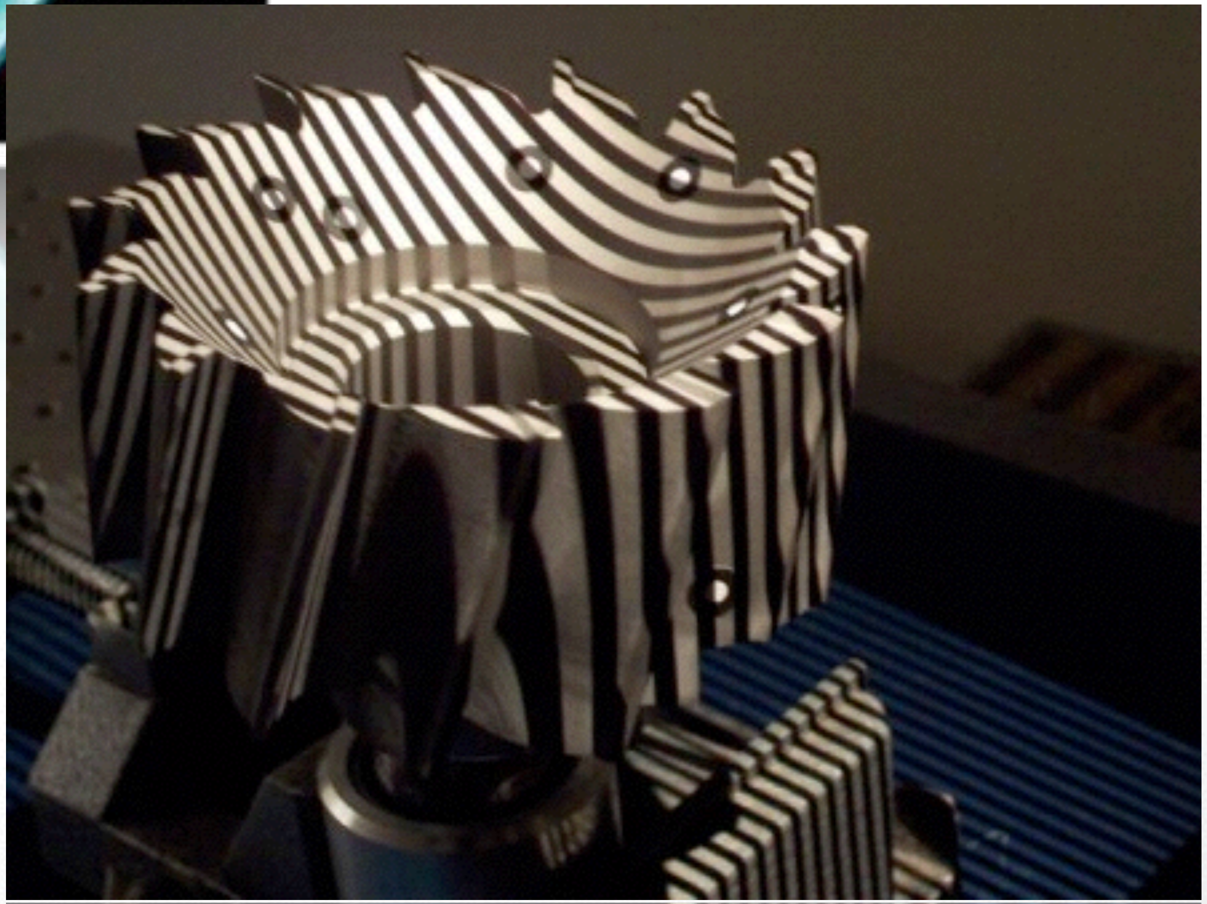
entertainment



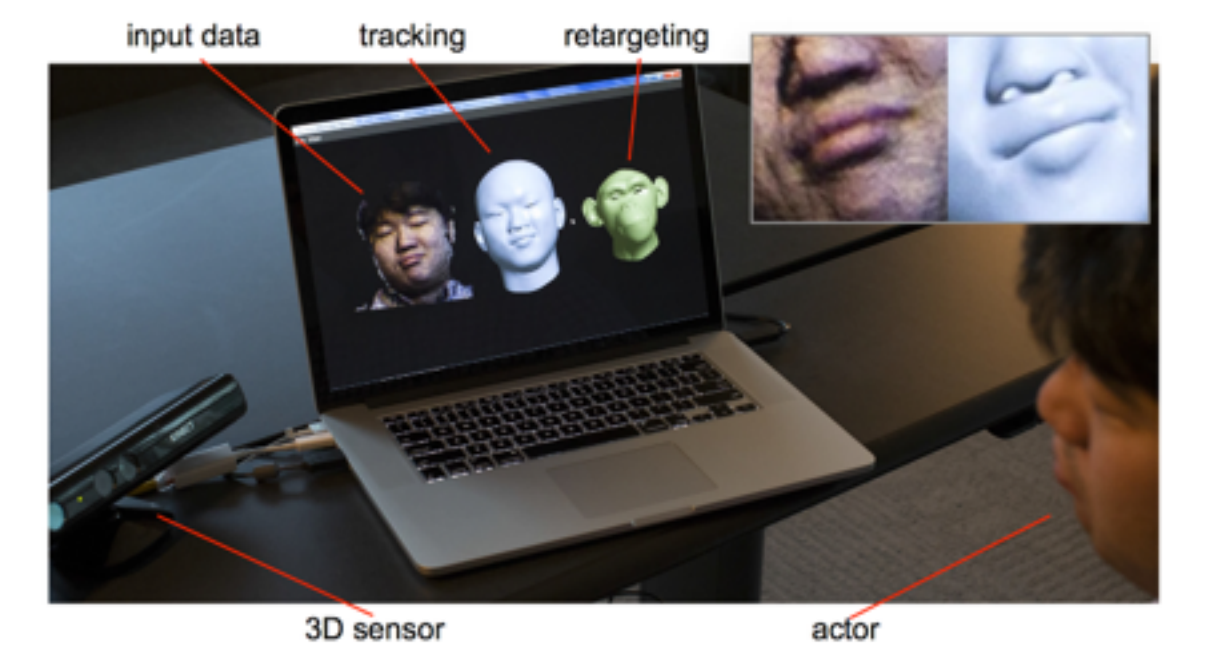
fitness

digital garment

Applications



Applications



Applications: Personalized Games

Digital Michelangelo Project

1G sample points → 8 M triangles

4G sample points → 8 M triangles

Commercialization

Democratization

3D Self-Portraits

Omote3D Shashin Kan

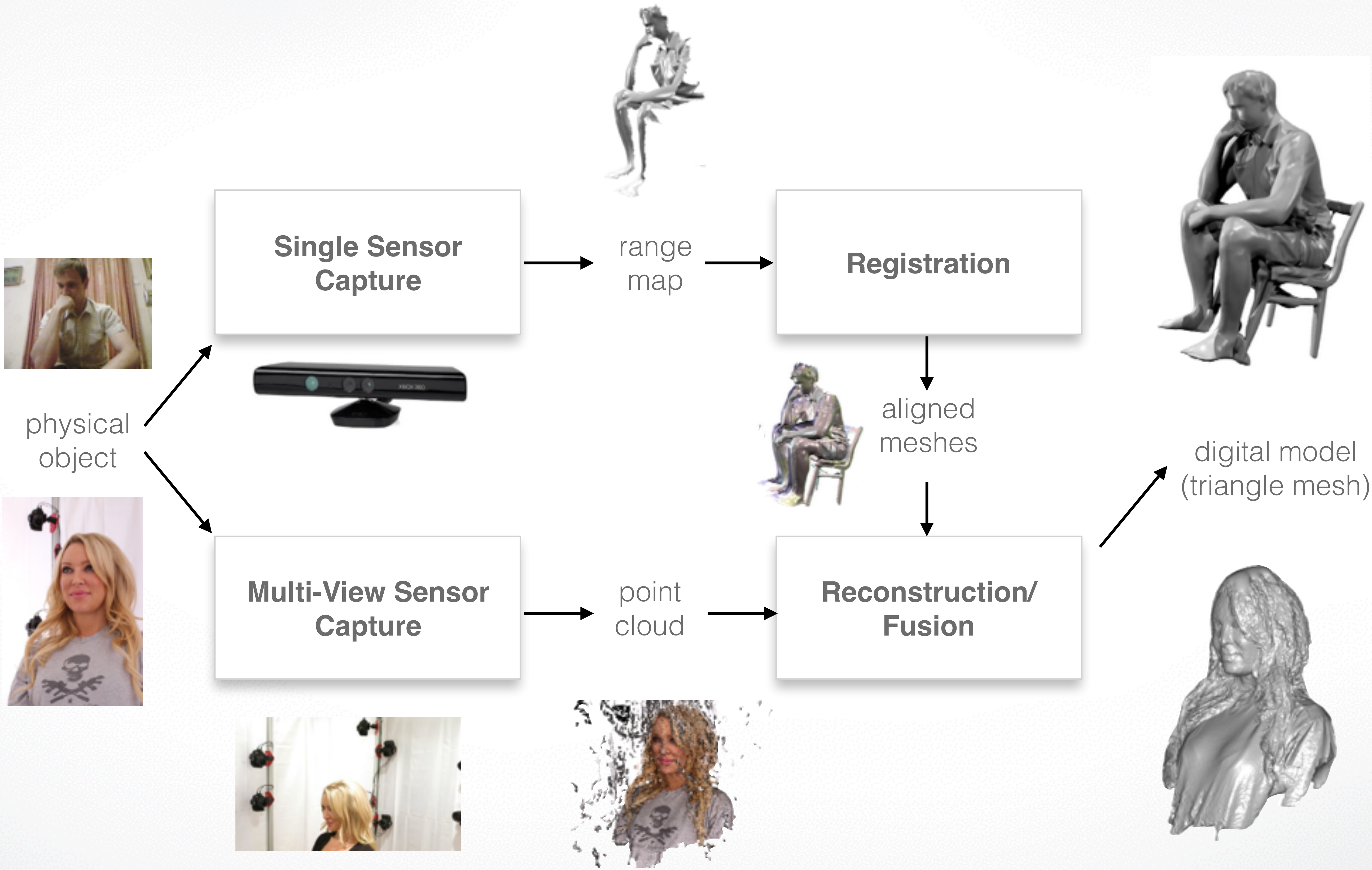
Surface Reconstruction Pipeline

physical
model

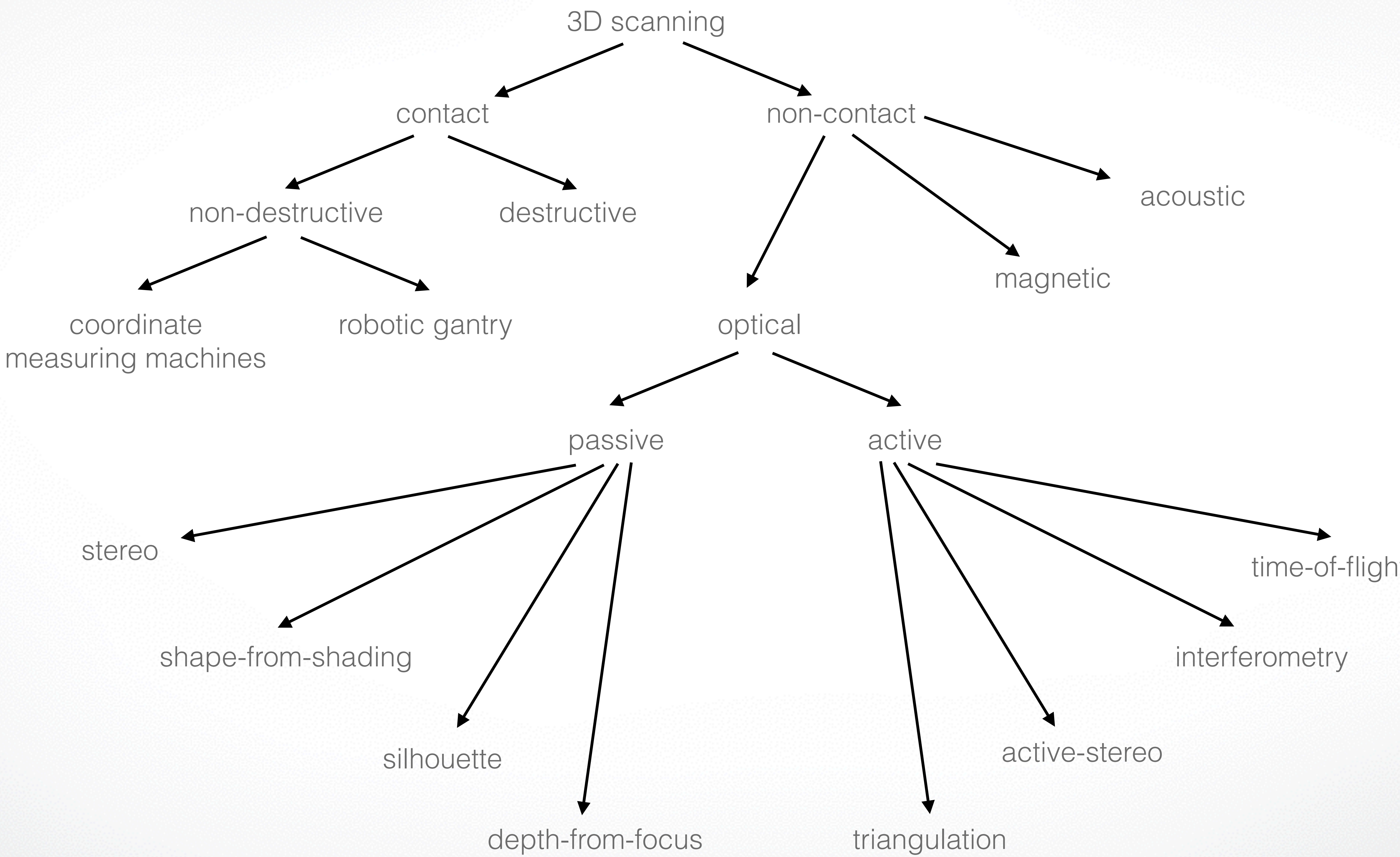
acquired
point cloud

digitized
model

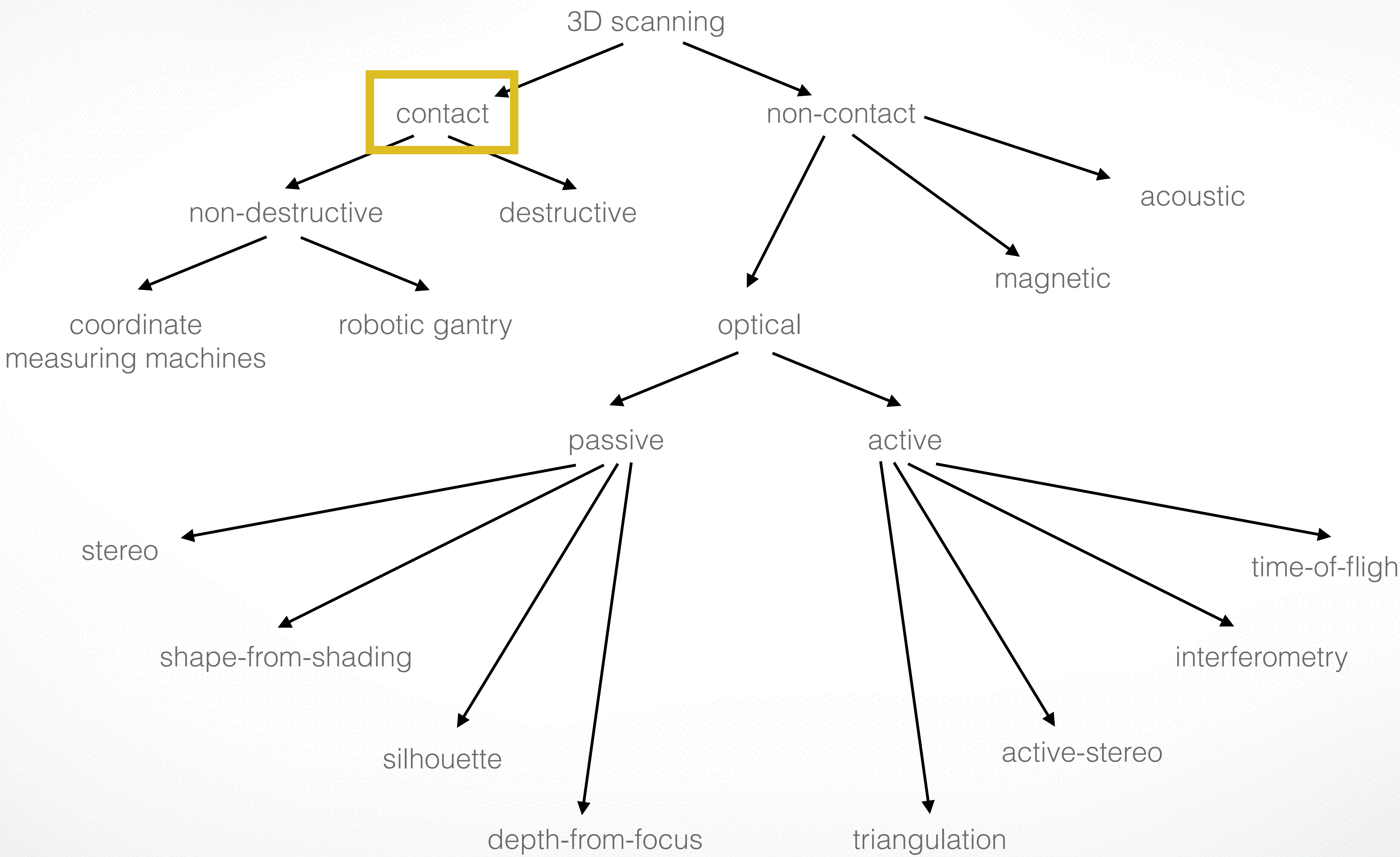
Two **Digitization Approaches**



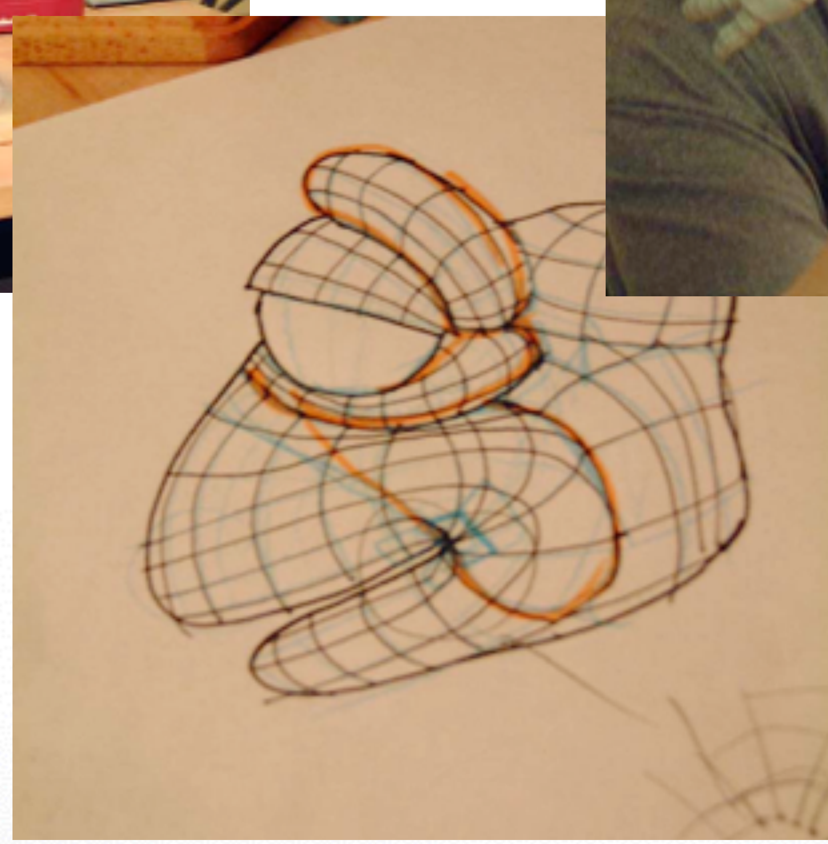
3D Scanning Taxonomy

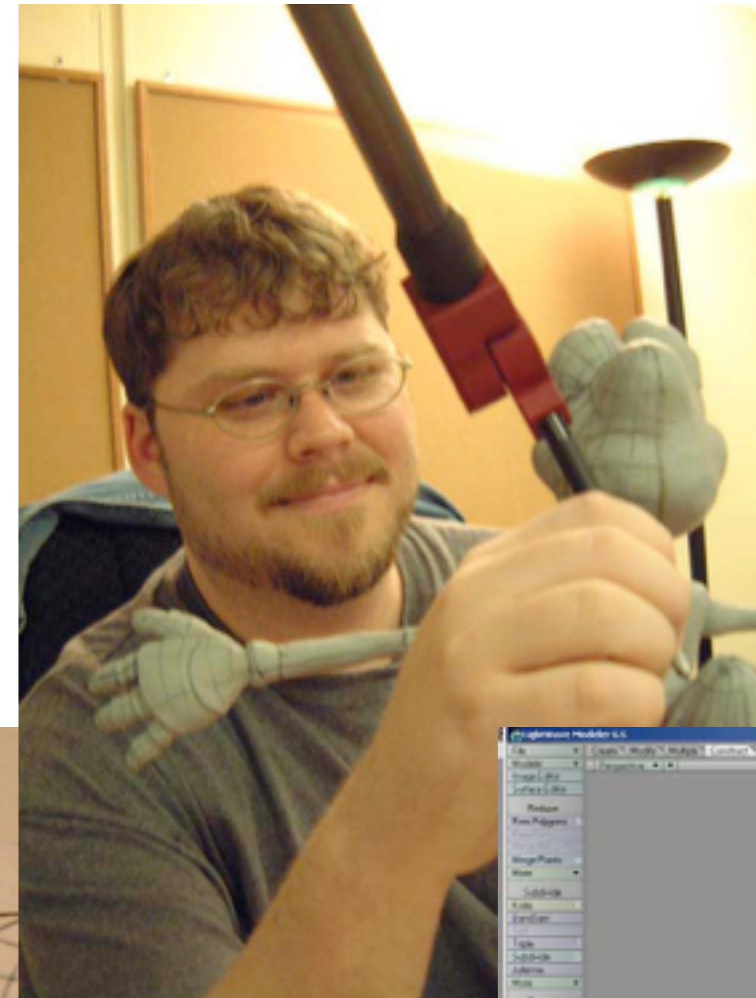


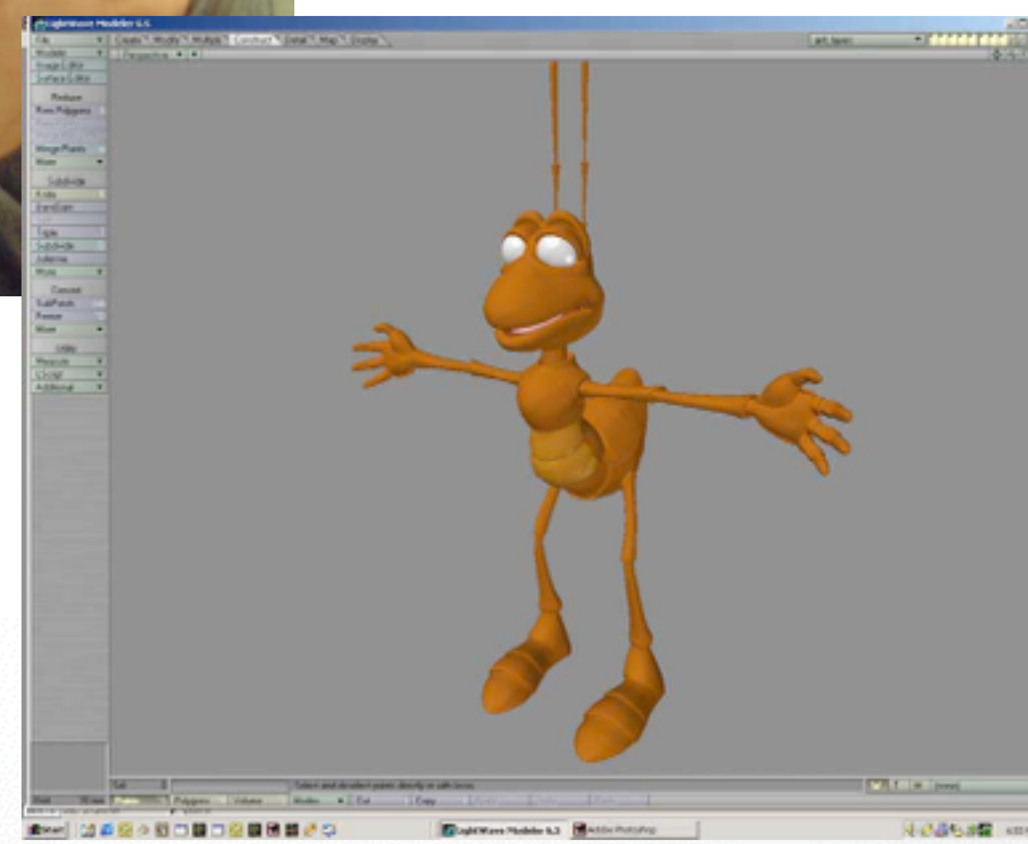
3D Scanning Taxonomy



Contact Scanners







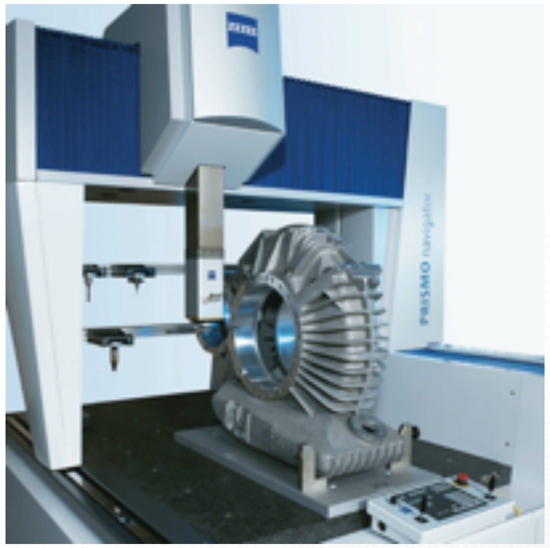
[Immersion Microscribe, Magnetic Dreams]

Contact Scanners

Probe object by physical touch

- used in manufacturing control

- highly accurate
- reflectance independent (transparency!)
- slow scanning, sparse set of samples
- for rigid and non-fragile objects



[Zeiss]

Contact Scanners

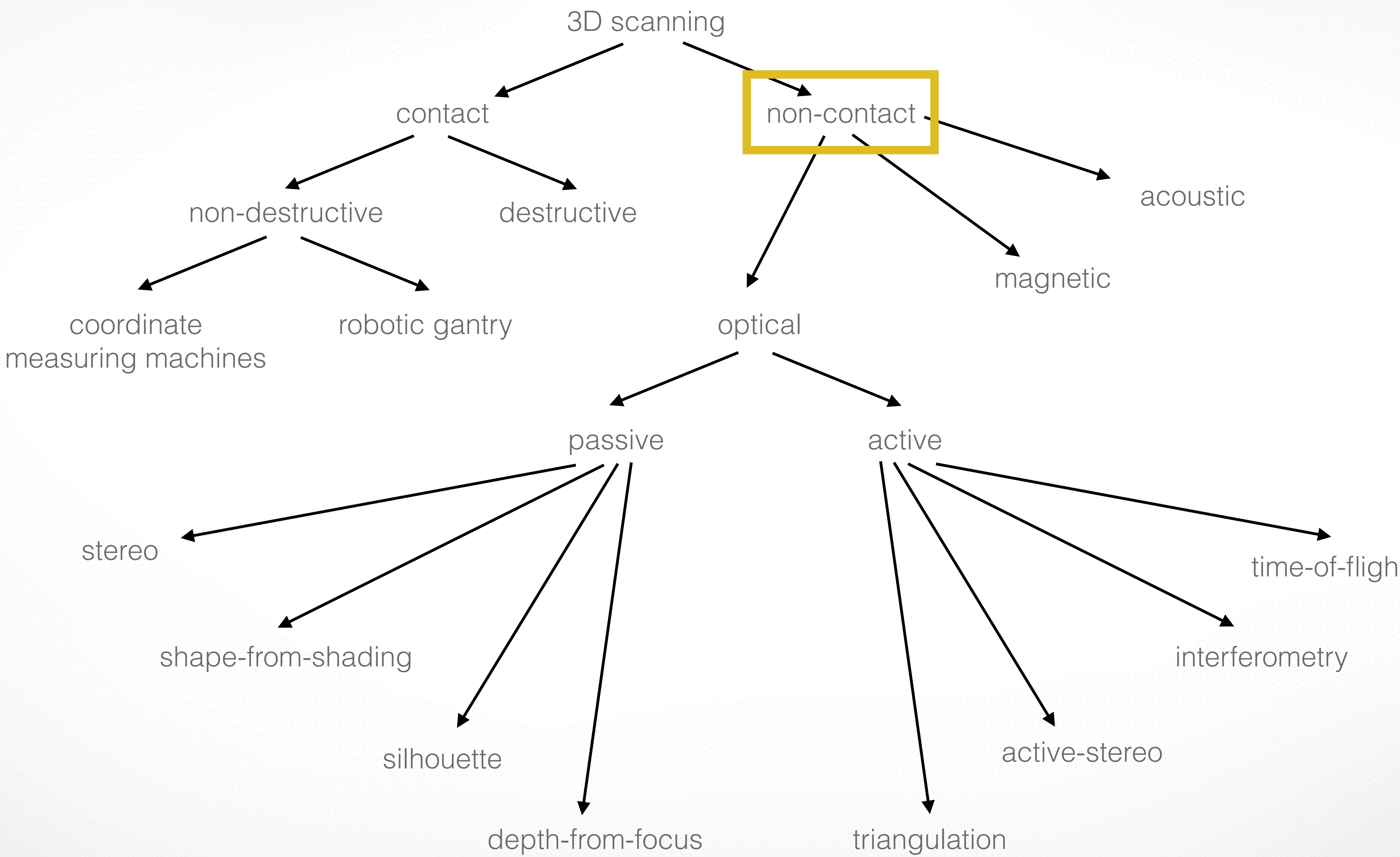
Probe object by physical touch

- hand-held scanners

- less accurate
- slow scanning, sparse set of samples

[Immersion Microscribe]

3D Scanning Taxonomy



Non-Contact

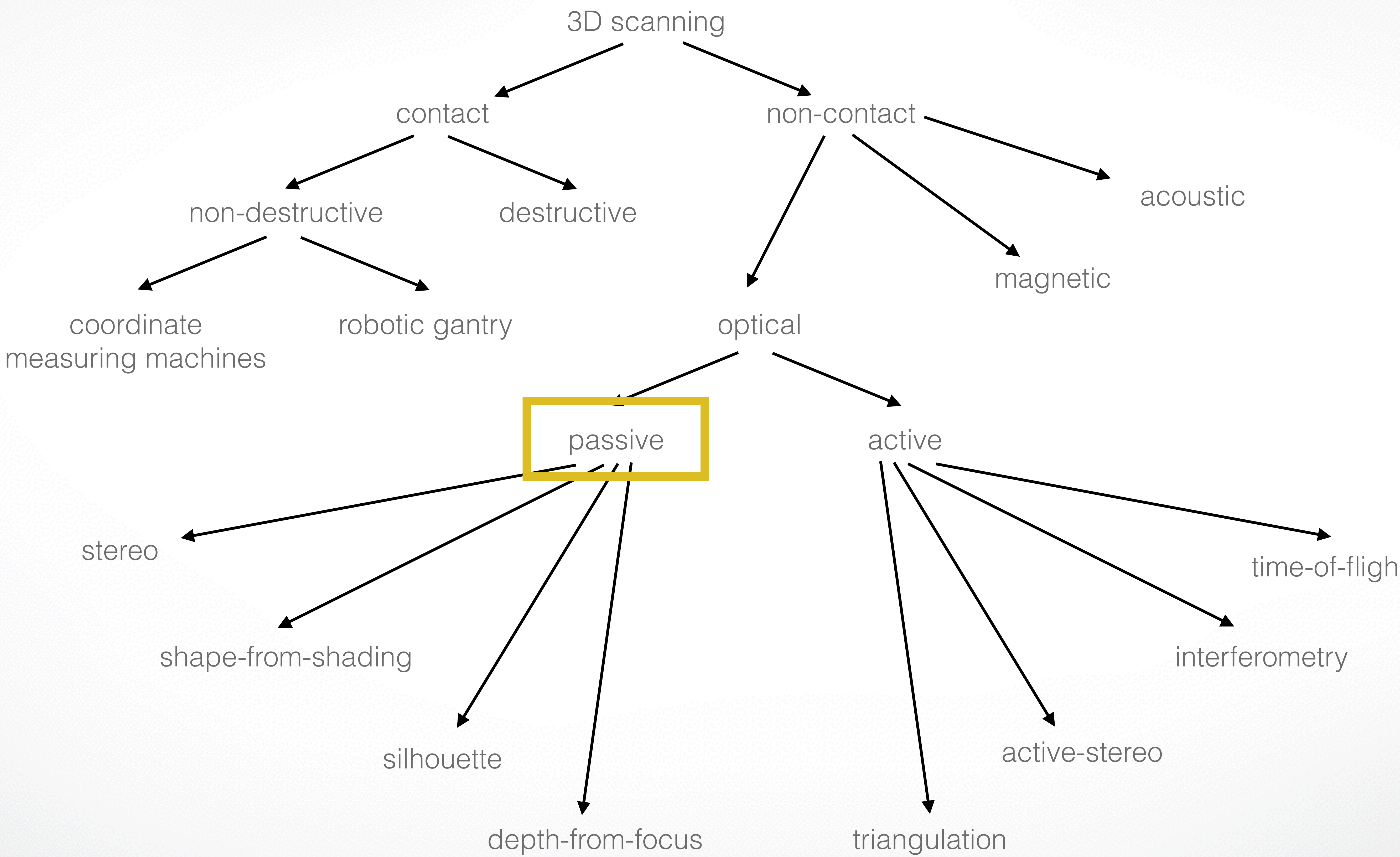
Advantages

- longer and safer distance capture
- potentially faster acquisition
- more automated

Optical Approaches

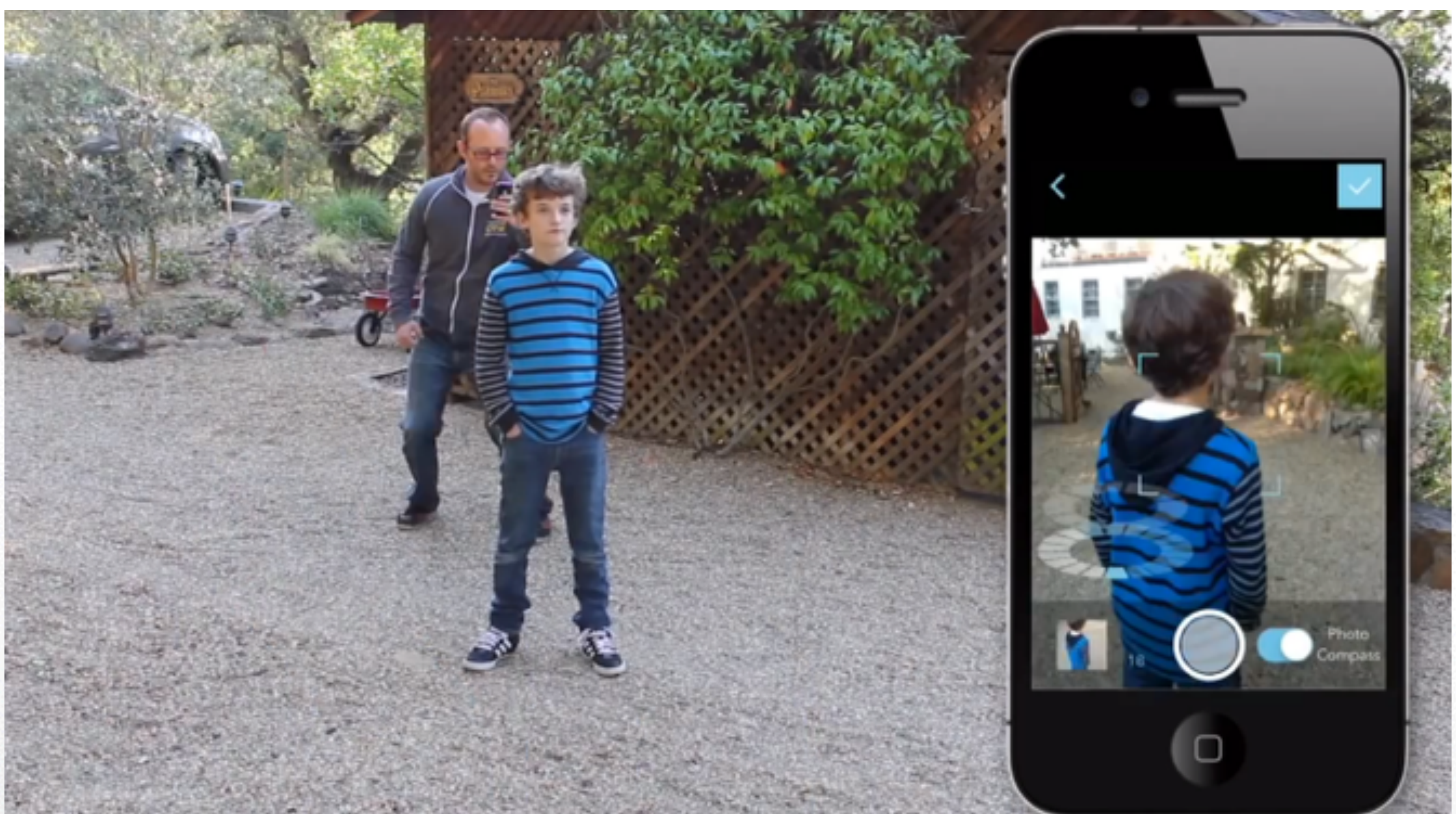
- most relevant and used (no special hardware requirements)
- highly flexible
- most accurate
- **passive** and **active** approaches

3D Scanning Taxonomy

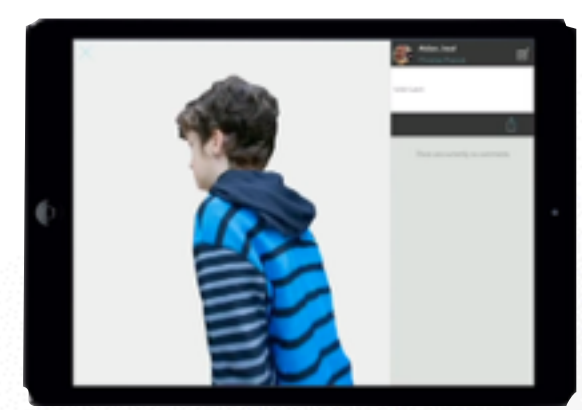


Passive

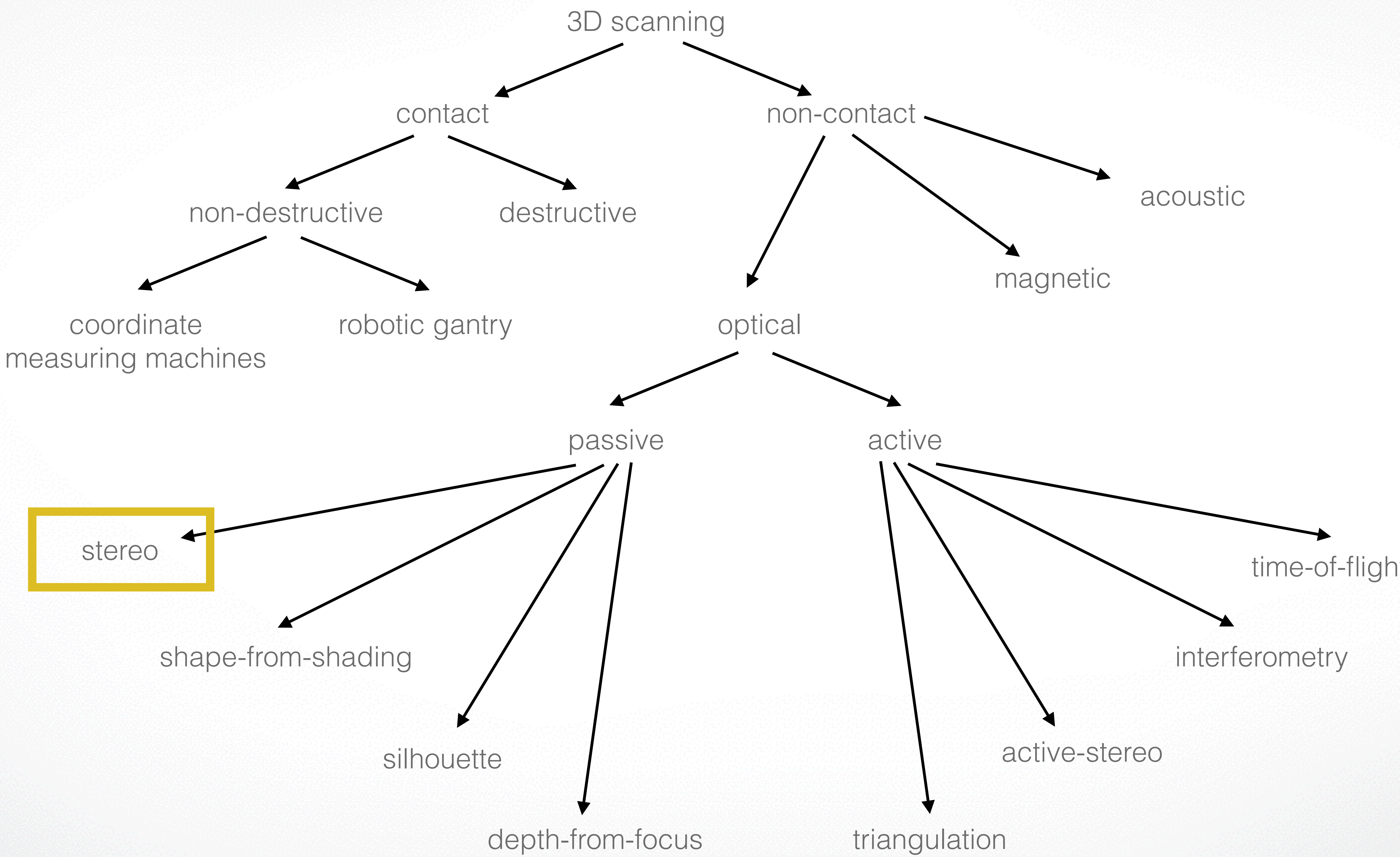
- exclusively based on **sensor(s)**
- computer vision-driven (stereo, multi-view stereo, structure from motion, scene understanding, etc.)
- main challenges: **occlusions** and **correspondences**
- typically assumes a **2D manifold** with **Lambertian reflectance**



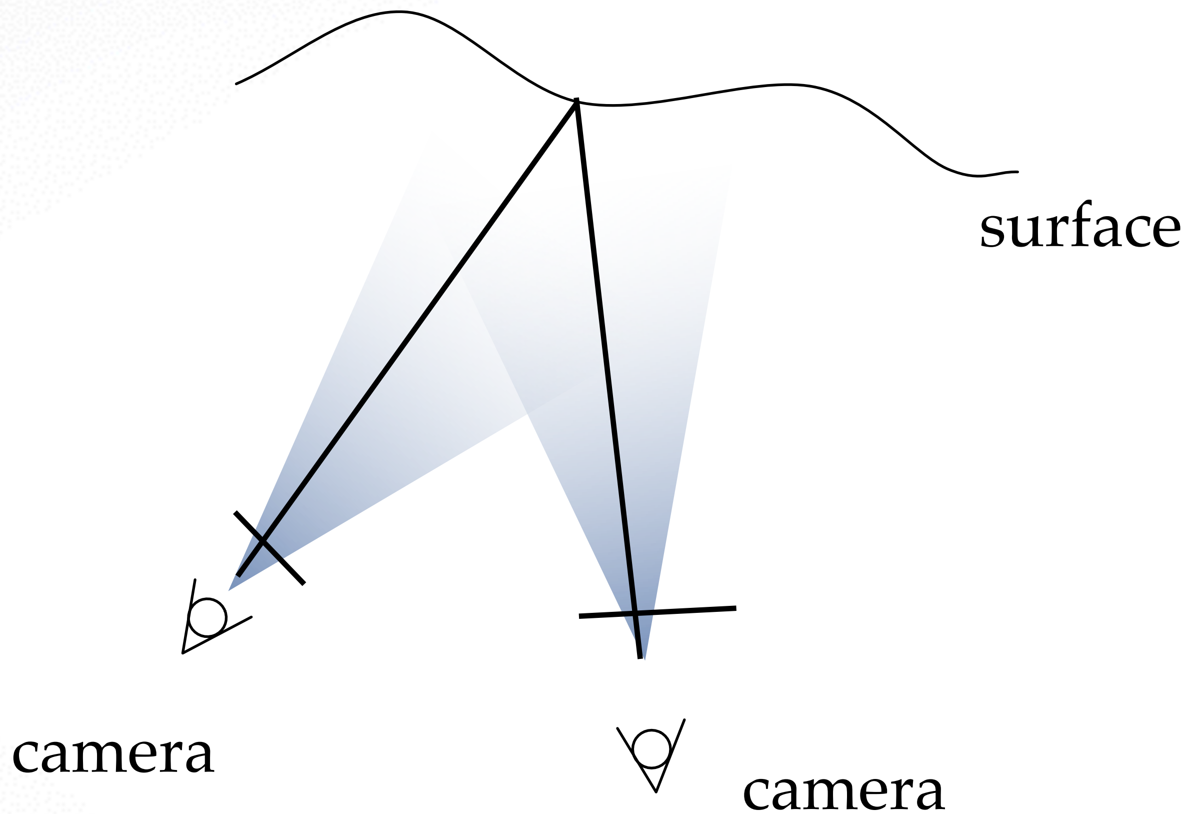
Autodesk 123D Catch



3D Scanning Taxonomy



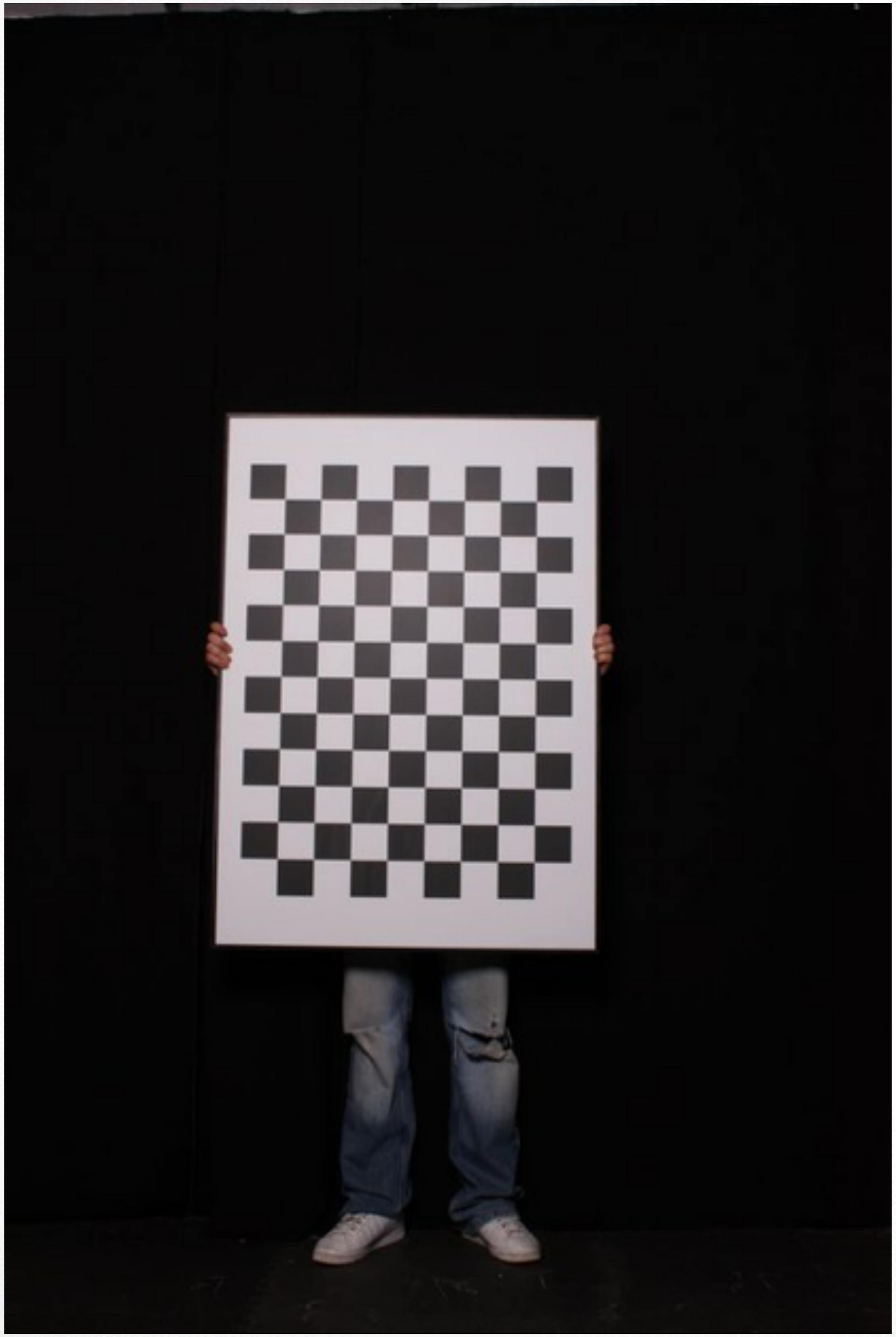
Stereo

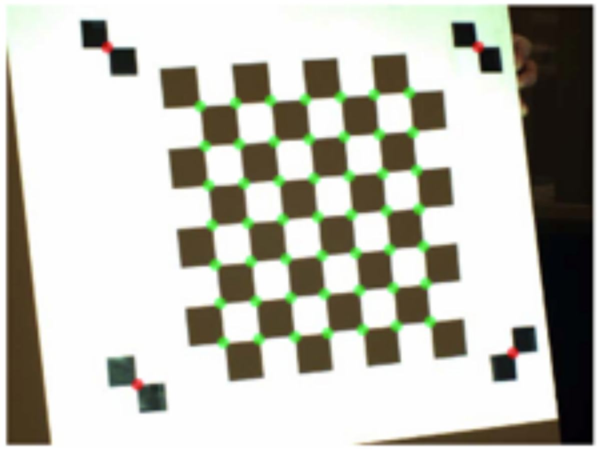


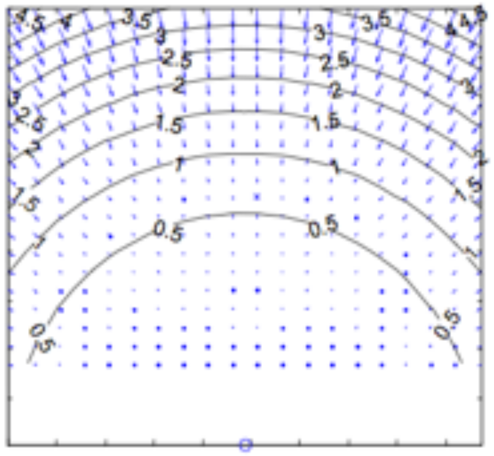
triangulation

image rectification

Calibration







extrinsics and intrisics (pinhole model)

lens distortion

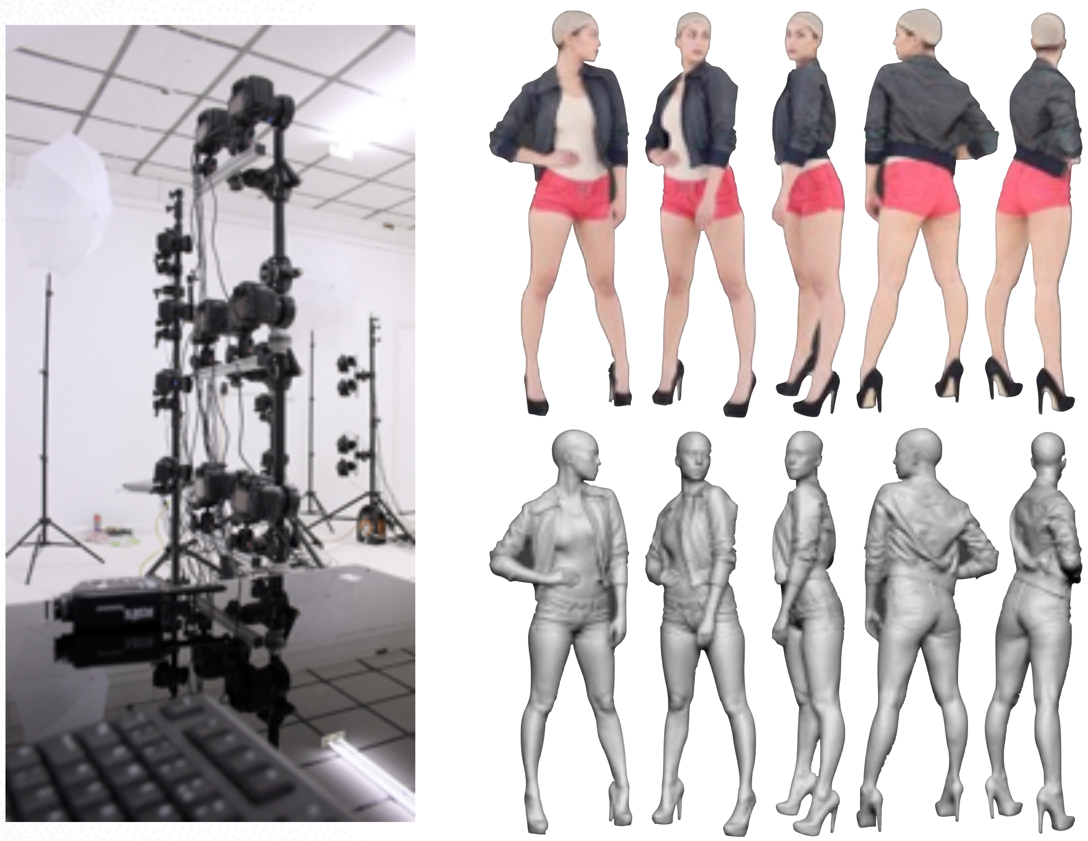
camera calibration toolbox

Stereo

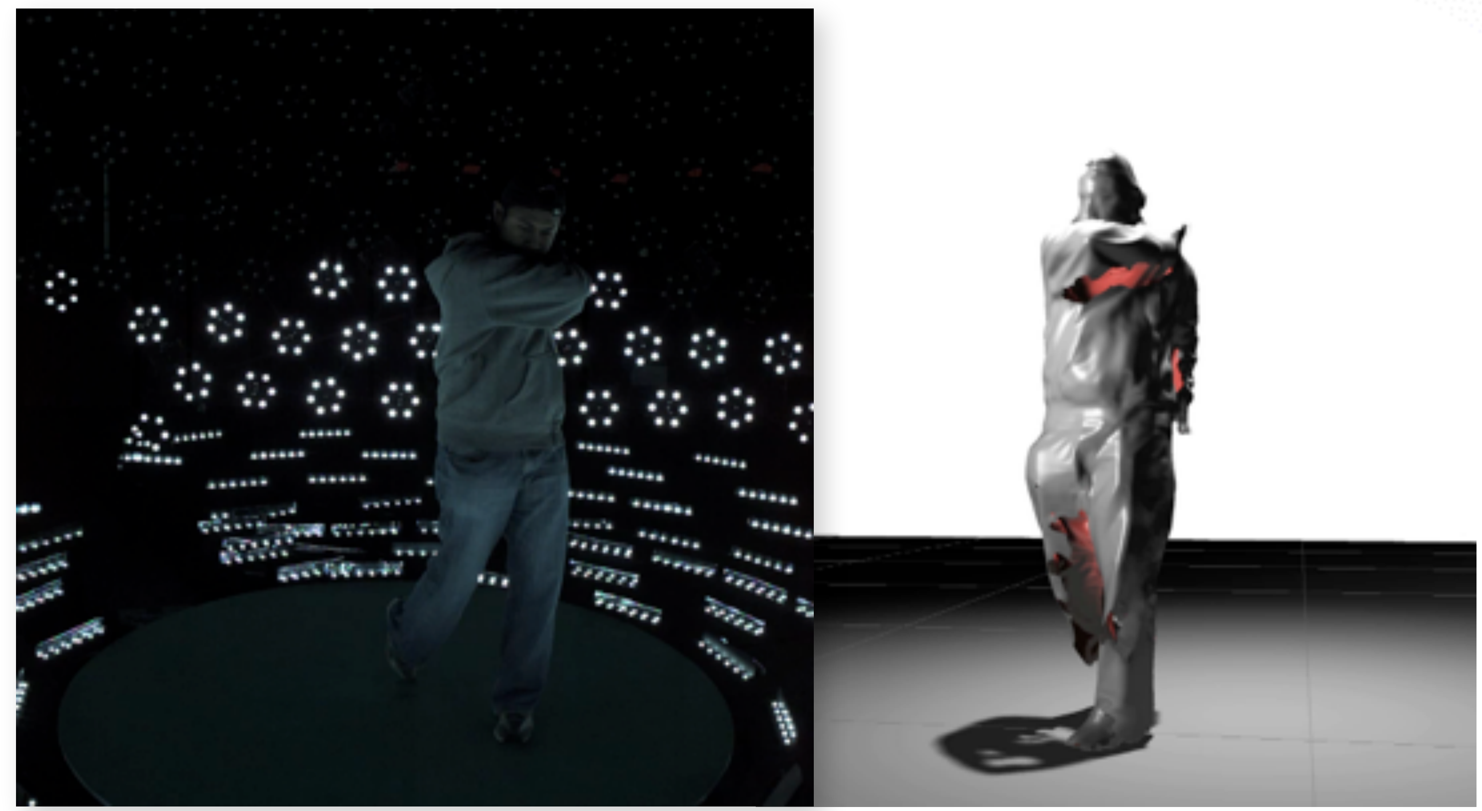
input

output

Multi-View Stereo



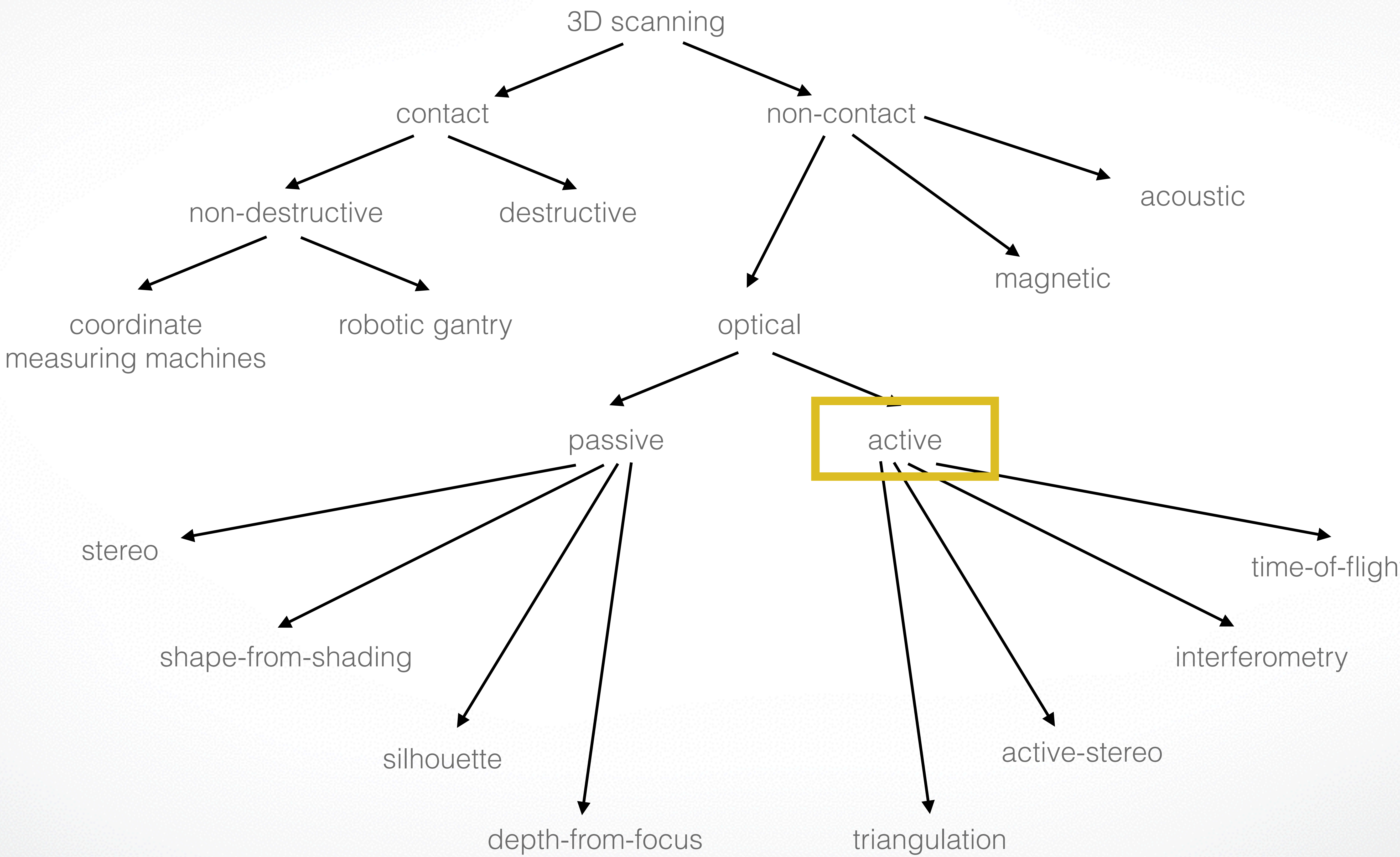
multi-view stereo



multi-view photometric stereo

Multi-View Stereo

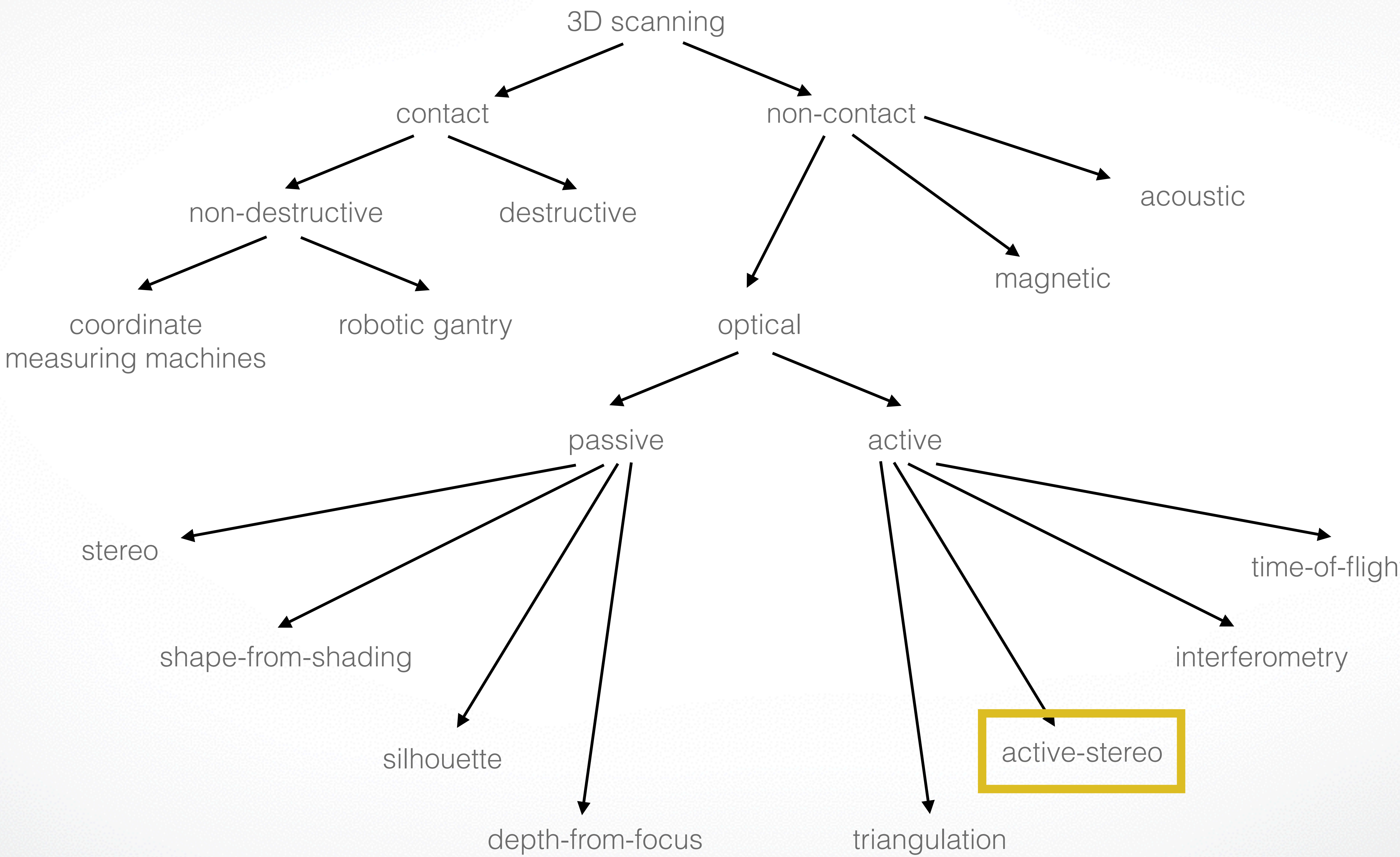
3D Scanning Taxonomy



Active

- based on **sensor** and **emitter** (controlled EM wave)
- influence of surface reflectance to emitted signal
- **correspondence** problem simplified (via known signal) → less computation (realtime?)
- examples (laser, structured light, photometric stereo)
- high resolution and dense capture possible, even for texture poor regions
- more sensitive to surface reflection properties (mirrors?)

3D Scanning Taxonomy



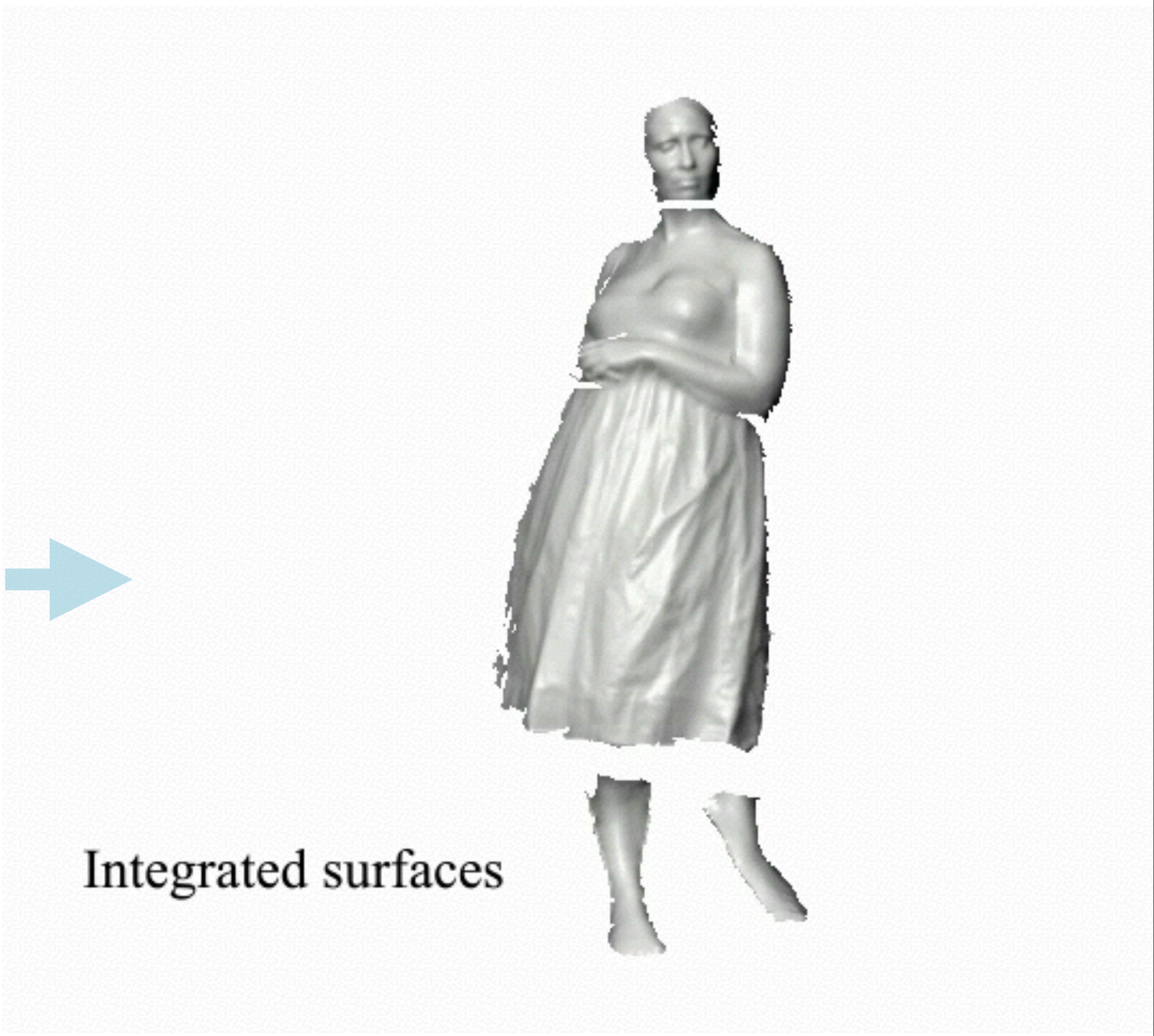
Active Stereo

Photometric Stereo

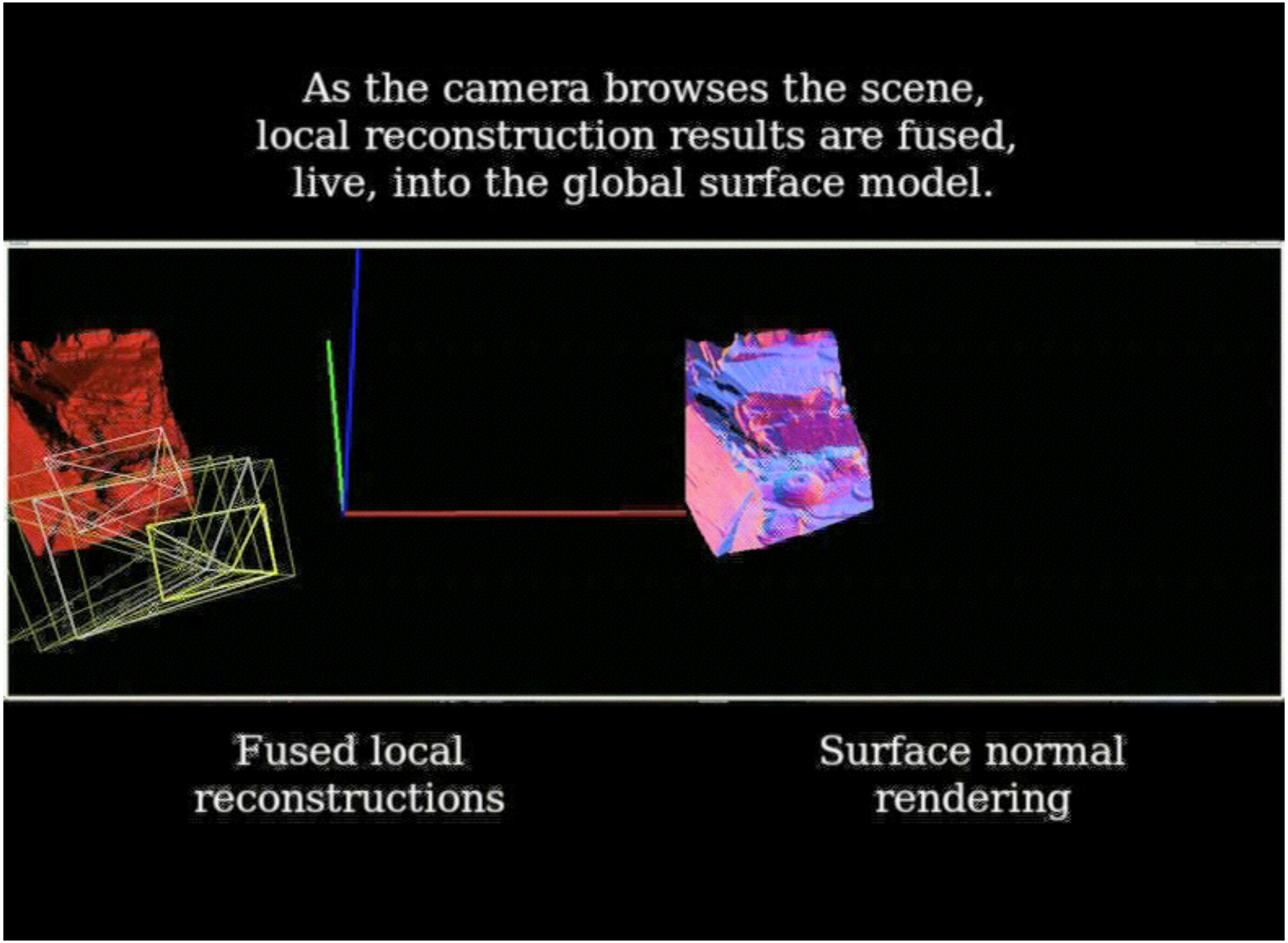
Lightstage 6 (USC-ICT)

8 Normal Maps / Frame

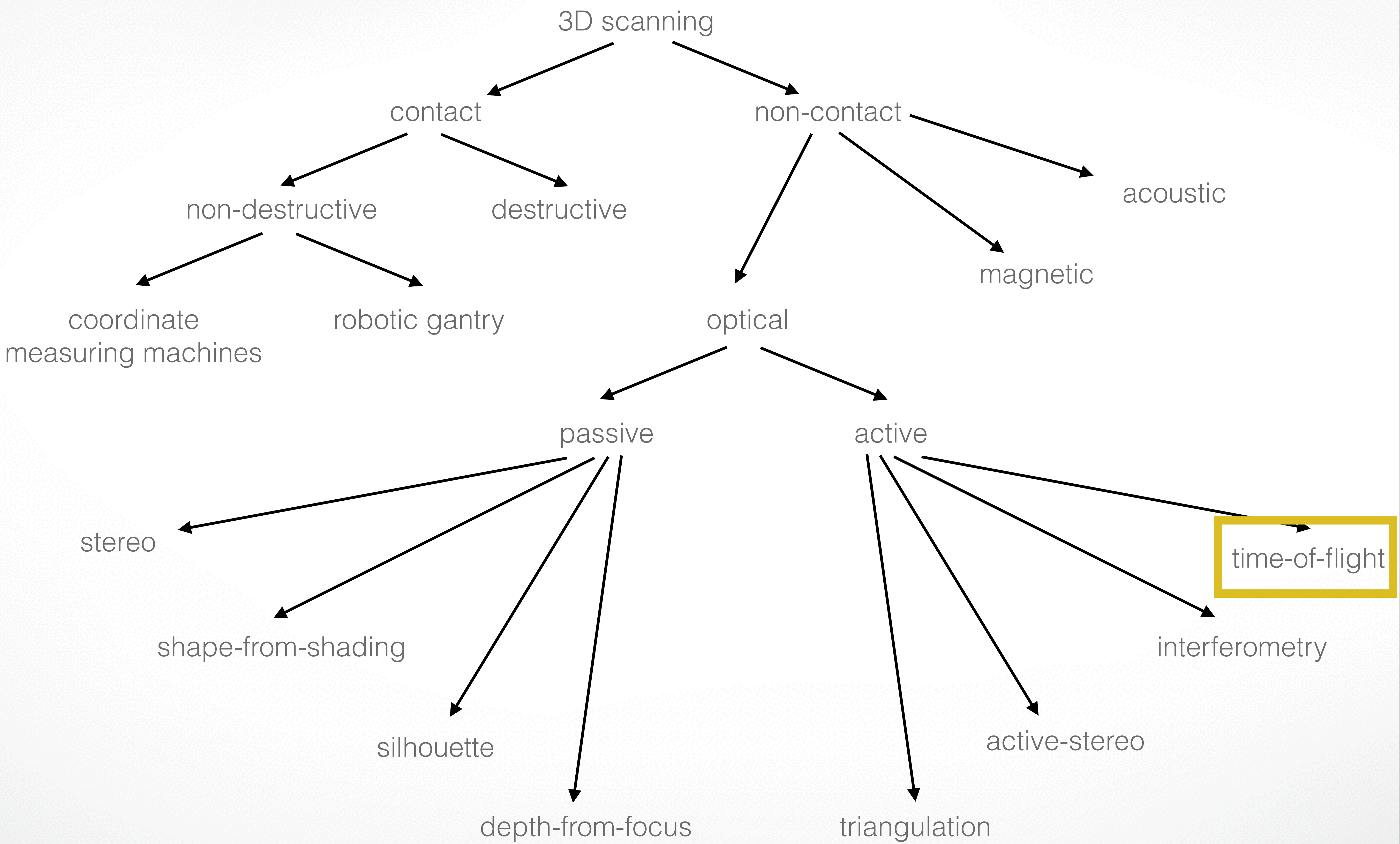
Photometric Stereo



Dense Structure from Motion



3D Scanning Taxonomy



Time-of-Flight Cameras

Probe object by laser or infrared light

- Emit pulse of light, measure time till reflection from surface is seen by a detector
- Known speed of light & round-trip time allows to compute distance to surface

Laser LIDAR

- **Li**ght **D**ectection **a**nd **R**anging
- Good for long distance scans
- 6mm accuracy at 50 m distance

[Leica]

Time-of-Flight Cameras

Probe object by laser or infrared light

- Emit pulse of light, measure time till reflection from surface is seen by a detector
- Known speed of light & round-trip time allows to compute distance to surface

Infrared light

- 176x144 pixels, up to 50 fps
- 30 cm to 5 m distance
- 1 cm accuracy
- technology is improving drastically

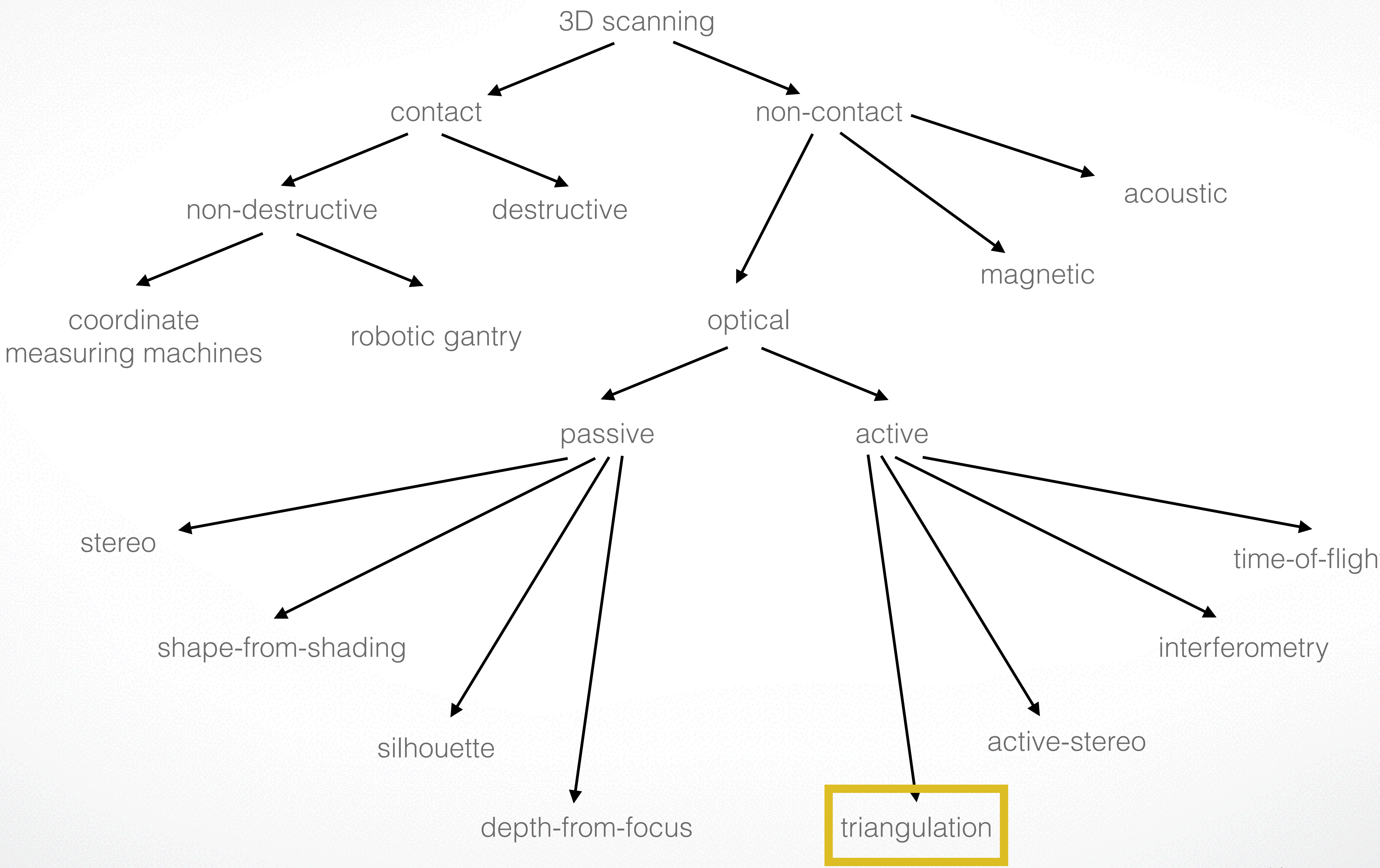
[Mesa Imaging]

Kinect One

Kinect One (= second gen Kinect)

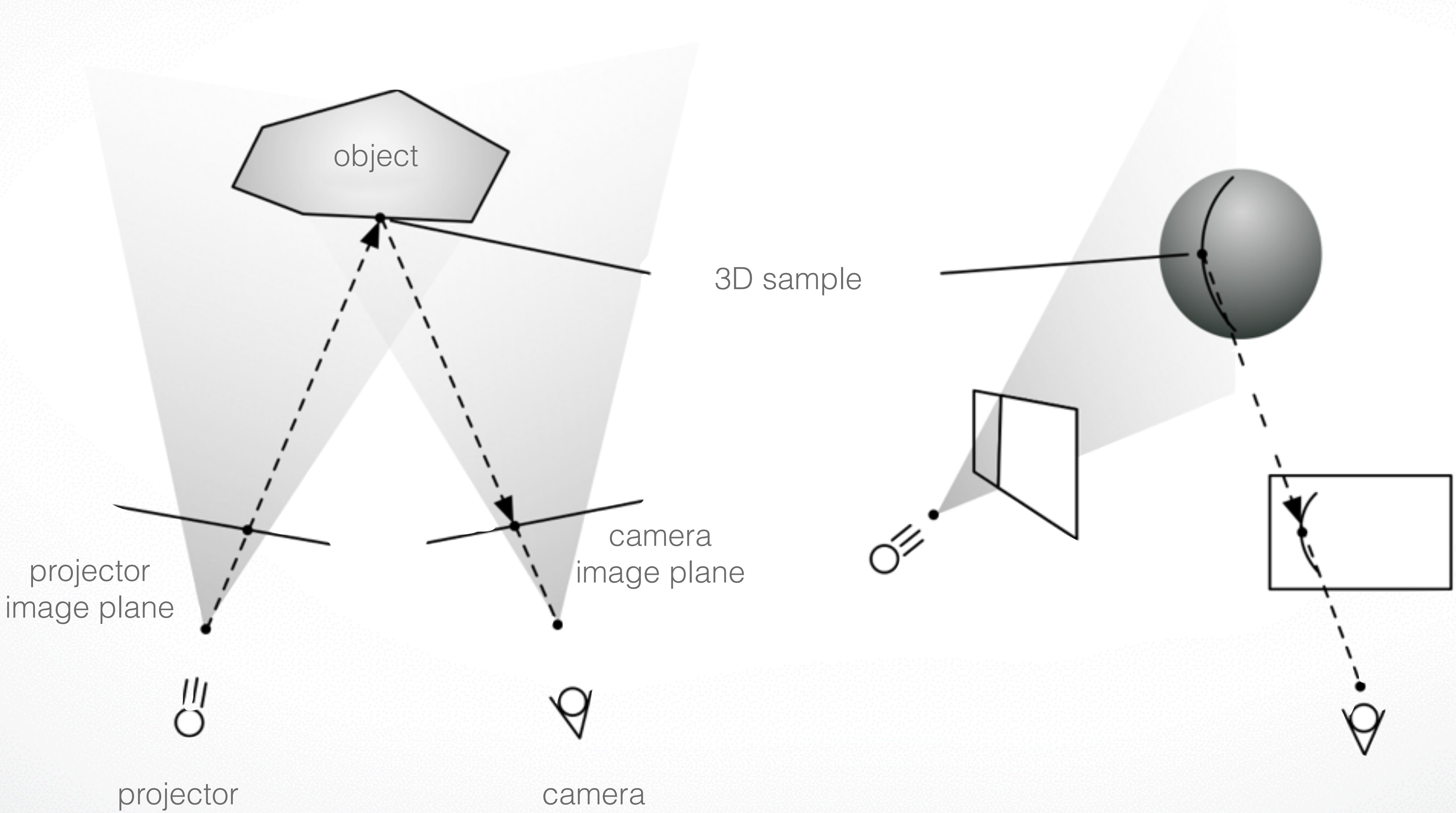
- Time-of-Flight Technology
- 30 fps
- Depth map x/y resolution: 512 x 424
- z-resolution 1 mm & accuracy:
  - <1.5 mm (depth < 50 cm)
  - < 3.9 mm (depth < 180 cm)
  - < 17.6 mm (depth < 450 cm)
- 1080 HD for RGB input
- uses Kinect2 SDK

3D Scanning Taxonomy

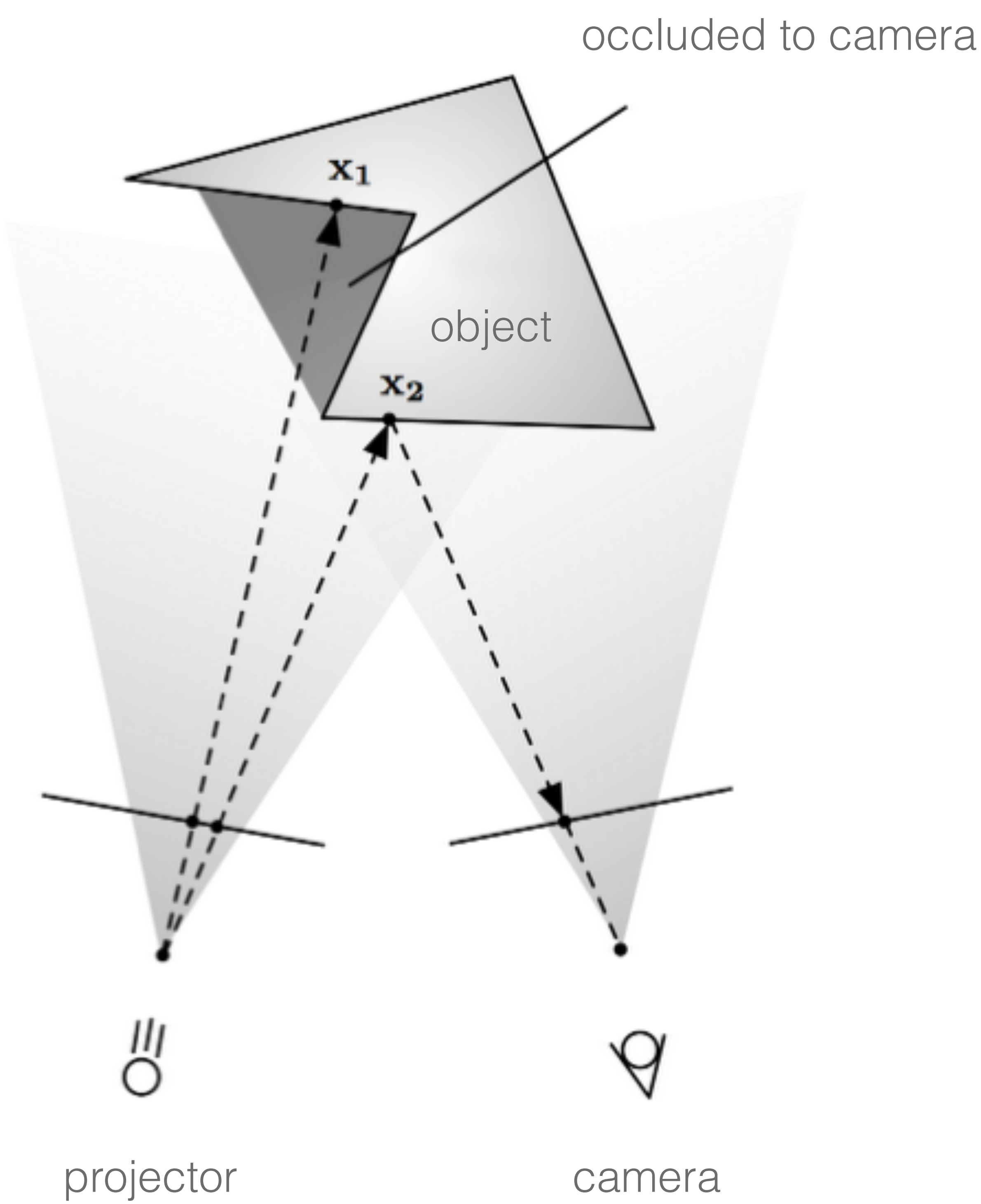


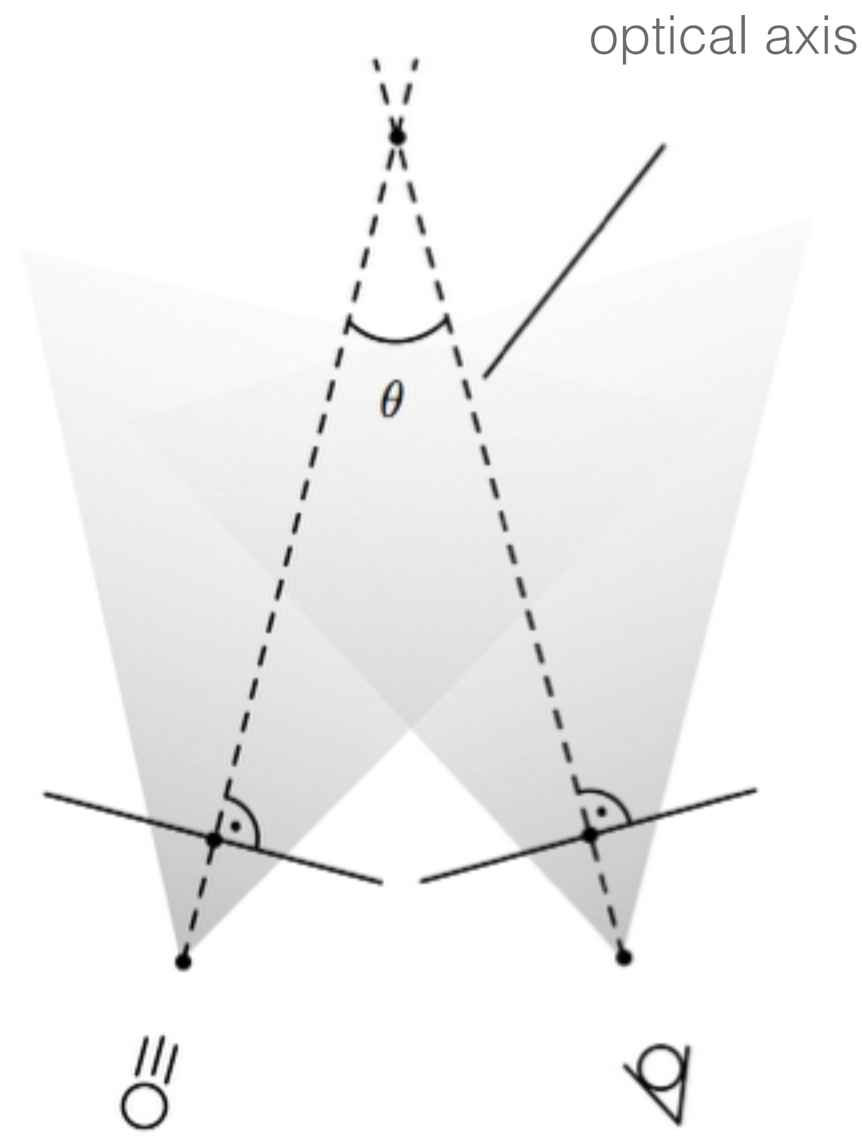
Optical Triangulation

2D View **3D View**

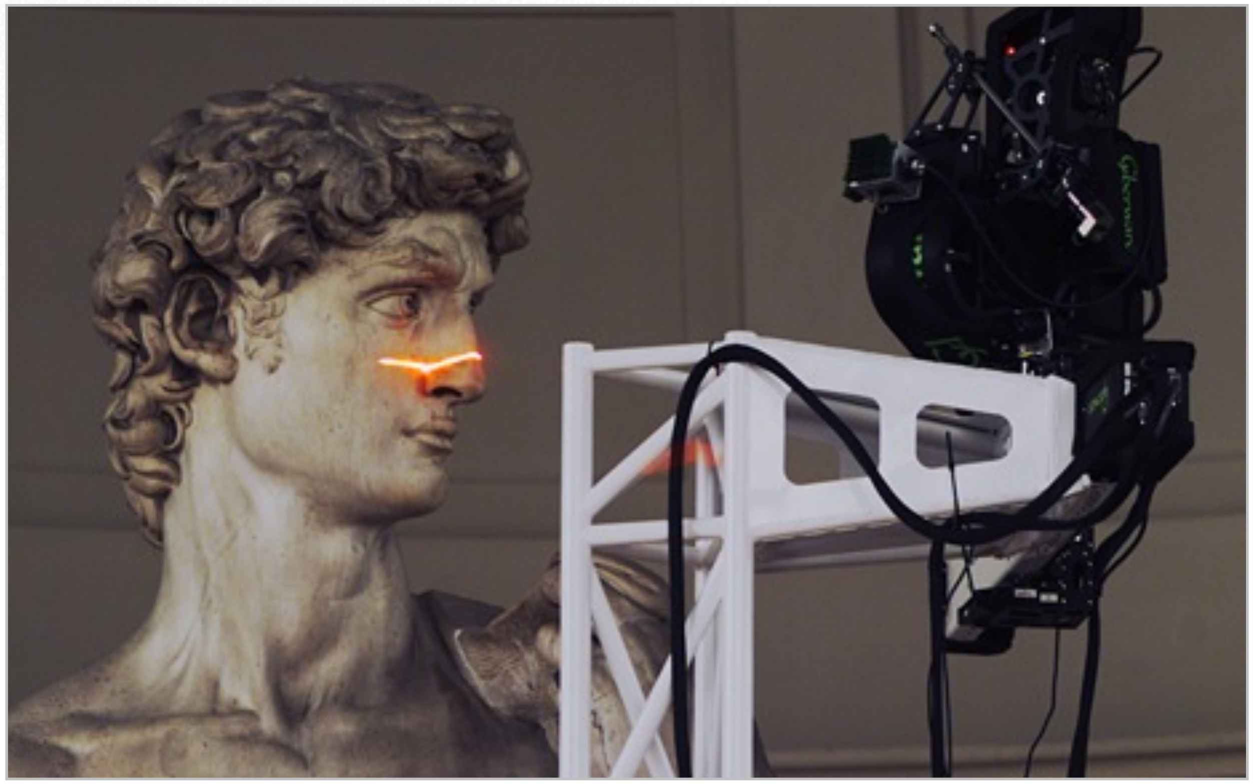


Geometric **Constraints**





Laser-Scanning



Digital Michelangelo Project

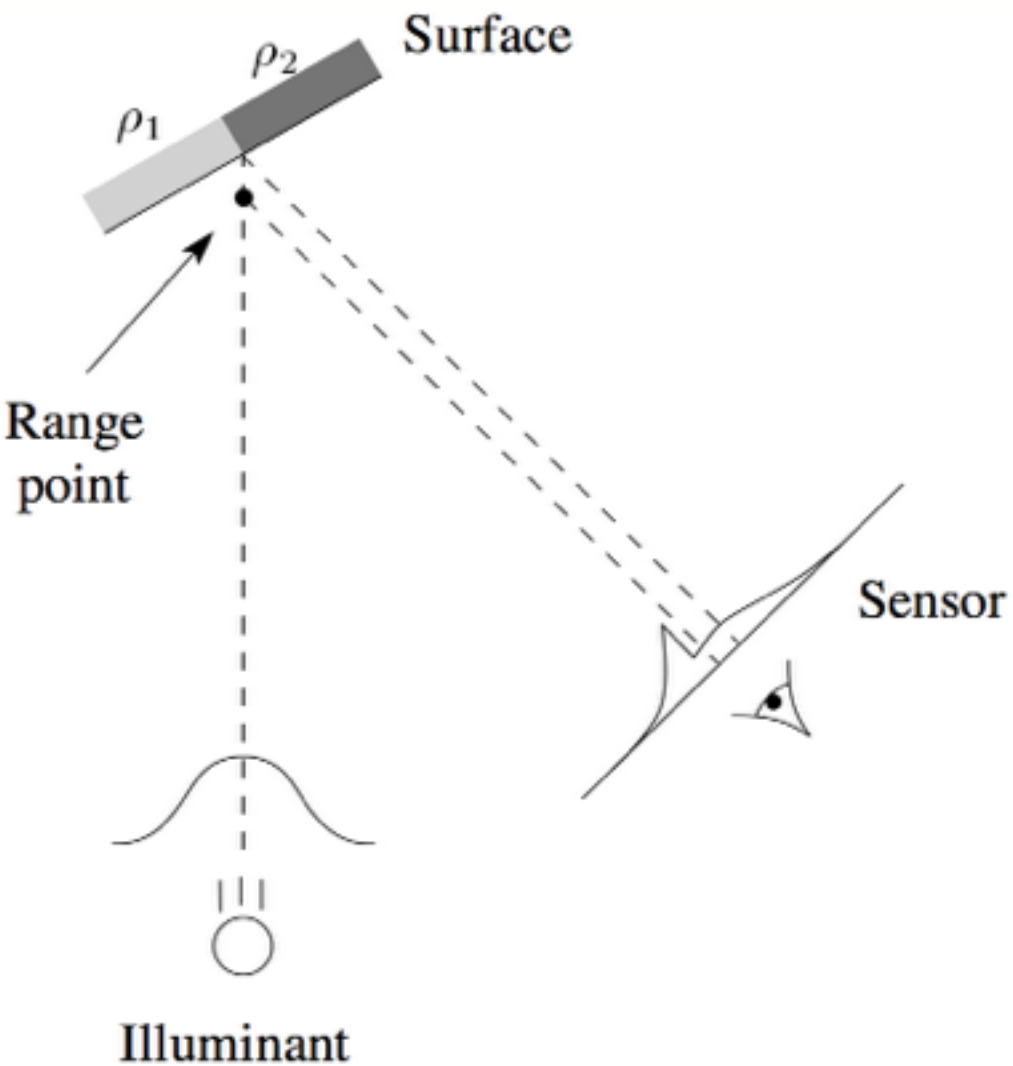
Cyberware

Konica Minolta

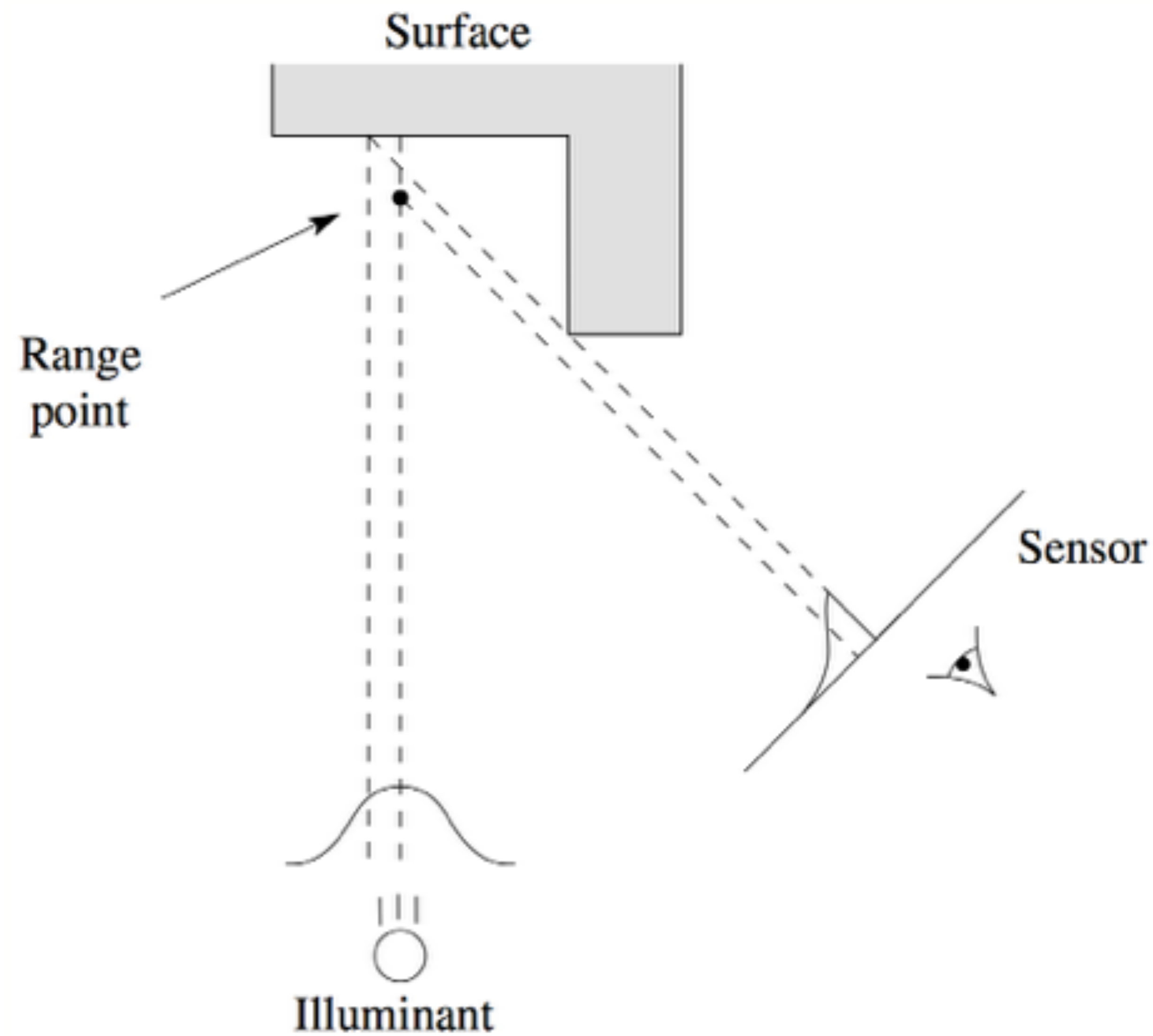
Laser-Based Optical Triangulation

- gained popularity for high accuracy capture (< 1mm)
- professional solutions are still expensive
- long range
- very insensitive to object's color (e.g. black) and lighting conditions
- may lead to laser speckle on rough surface → space time analysis
- slow process (plane-sweep) → no suitable for dynamic objects

Surface Perturbs Laser Shape

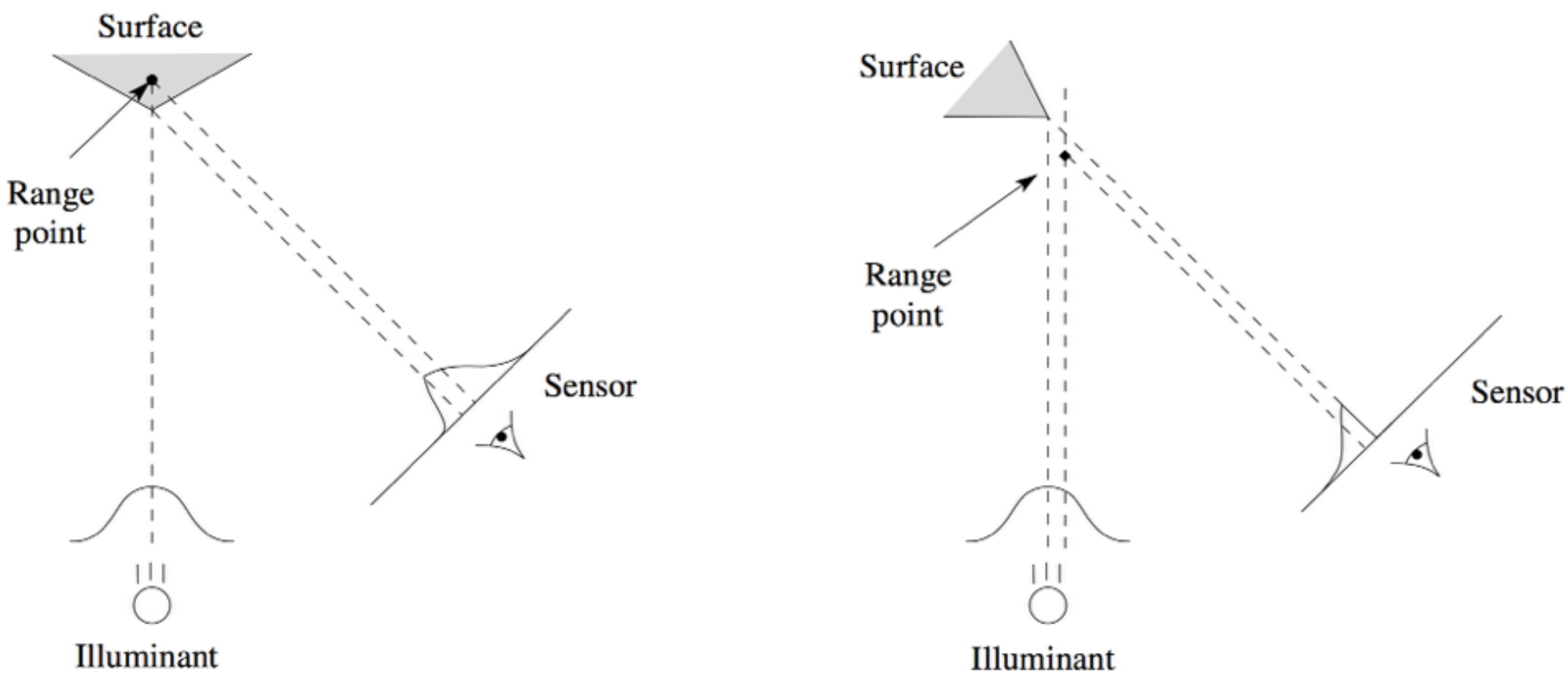


reflectance discontinuity



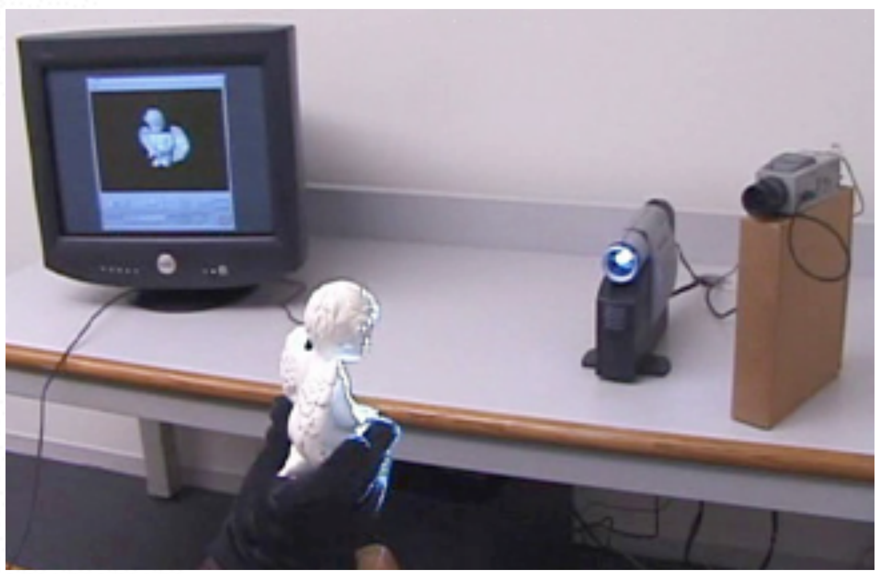
sensor occlusion

Surface Perturbs Laser Shape



shape variation

Single-View Structure Light Scanning



[Rusinkiewicz et al. ‘02]

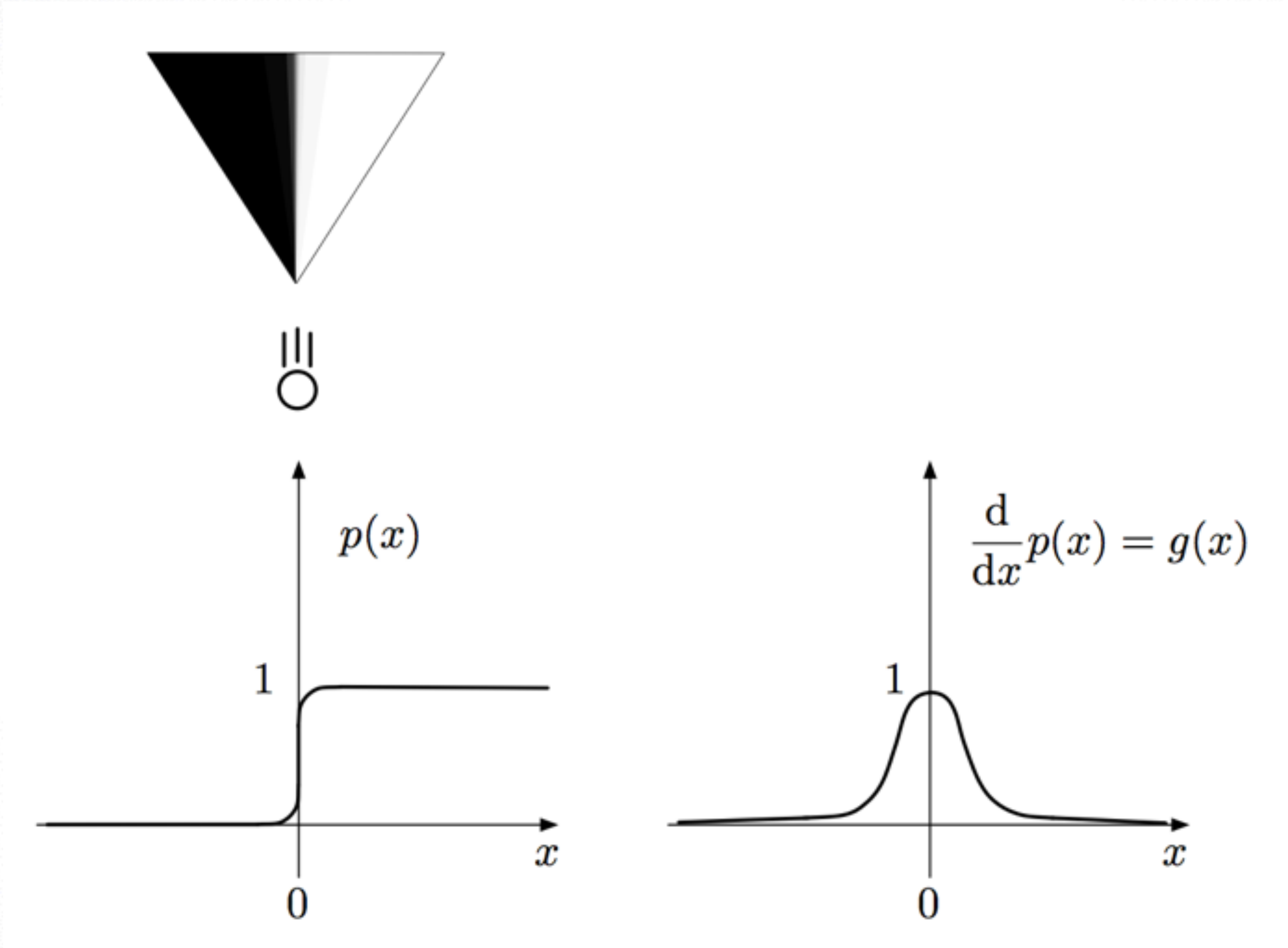
Artec Group

[Newcombe et al. ’11]
KinectFusion

Structured Light Scanning

- developed to increase capture speed by simultaneously projecting multiple **stripes or dots** at once
- increase accuracy using edge detection
- due to cost and flexibility, based on a **video projector**
- challenge: recognize projected patterns (**correspondence**)
  - under occlusions
  - different surface reflection properties (furry object?)
  - less projections → faster but correspondence harder
- typically assumes a **2D manifold** with **Lambertian reflectance**

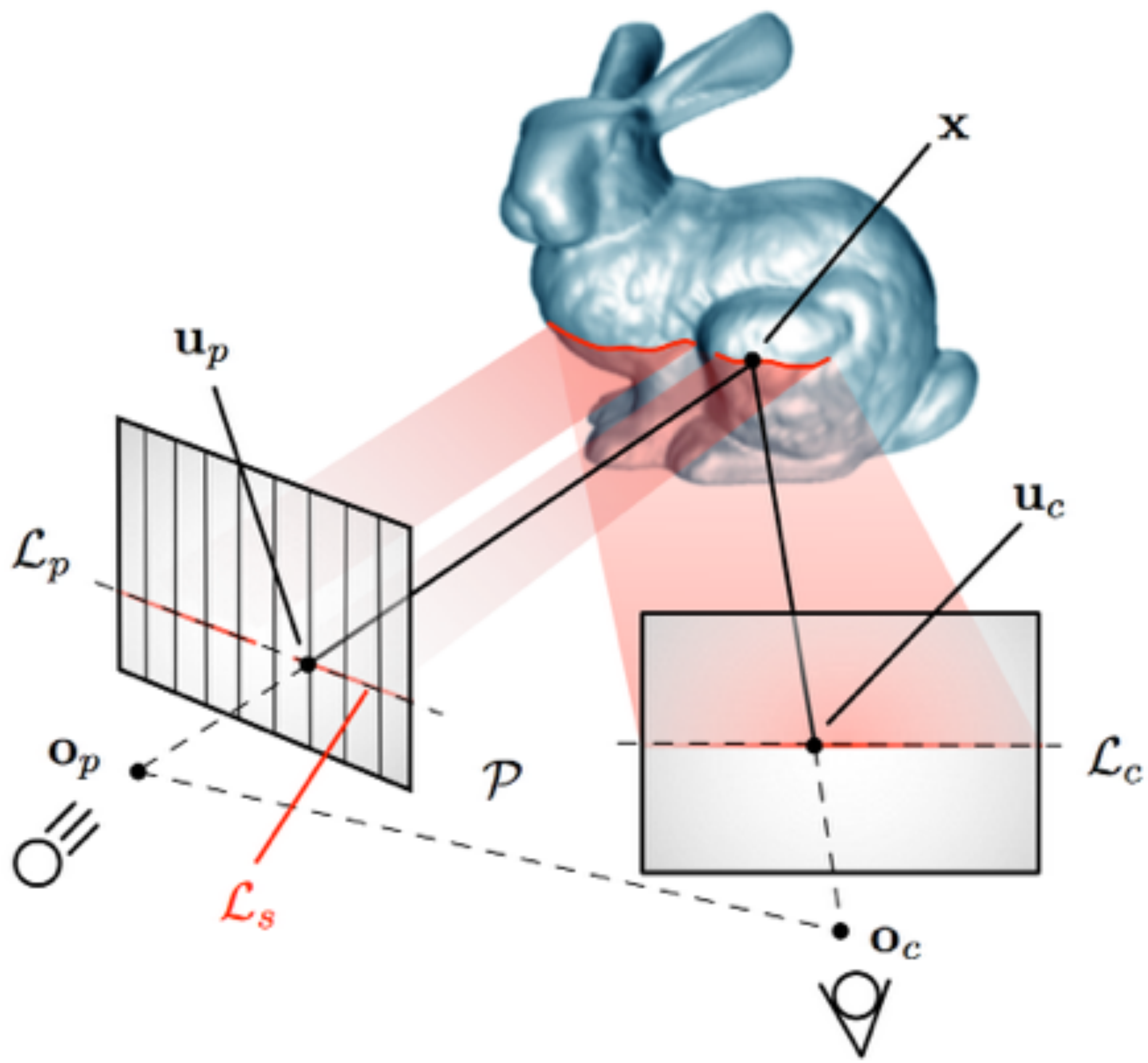
Stripe Edge Detection



Epipolar Geometry

correspondence is a 1D search

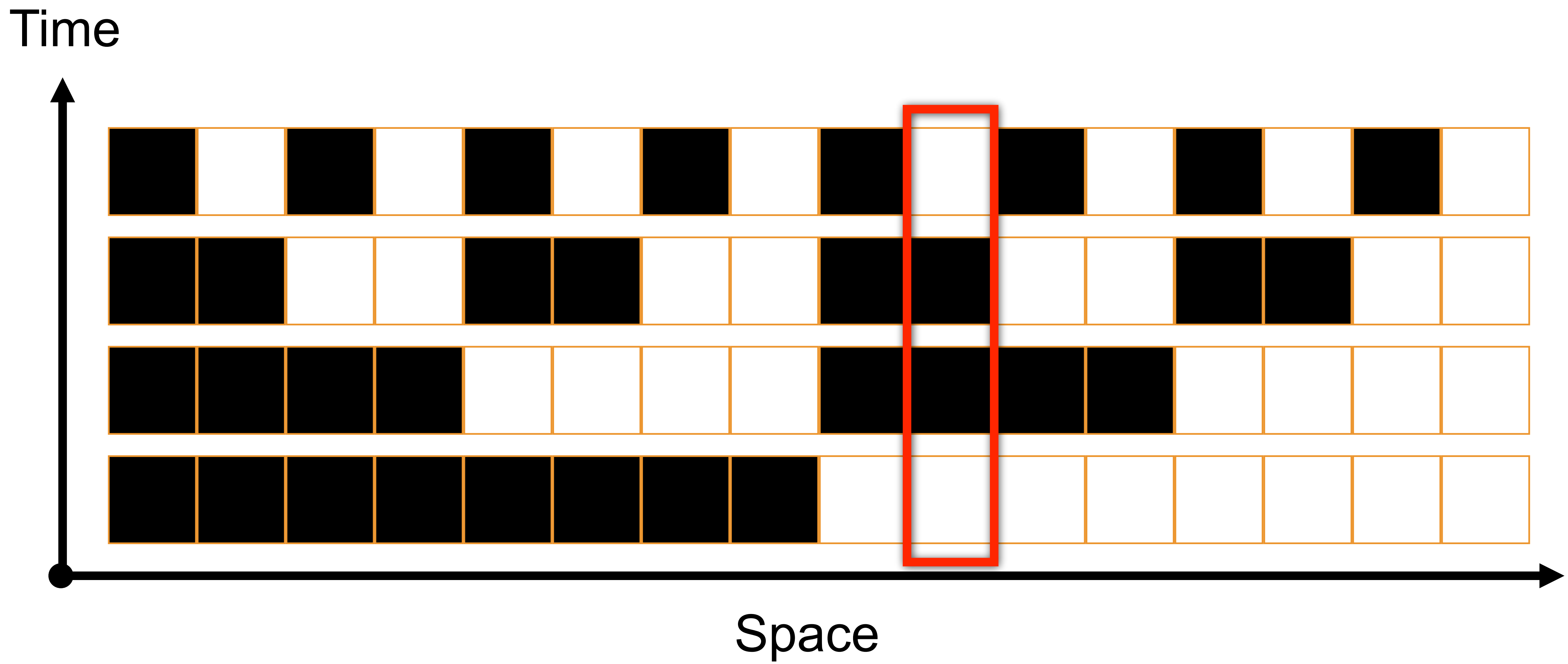
- same for passive stereo (but with rectification)



Time-Coded Light Patterns

Binary coded pattern

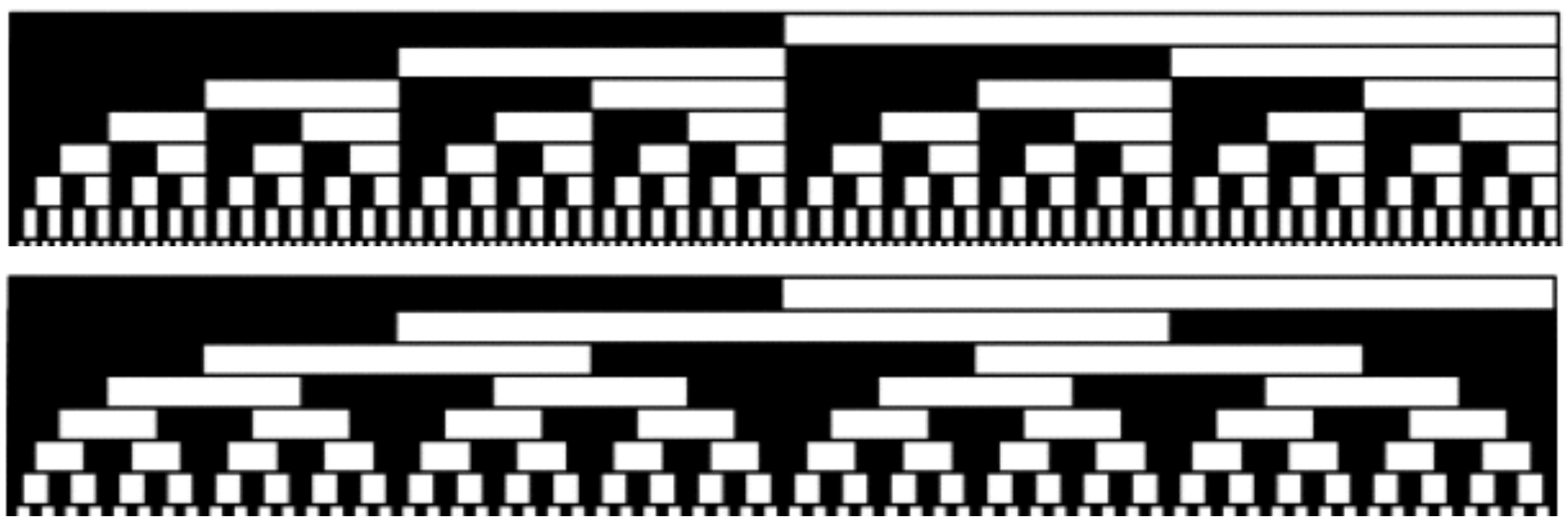
- project several b/w patterns over time
- color patterns identify row/column



Time-Coded Light Patterns

Gray Code Pattern

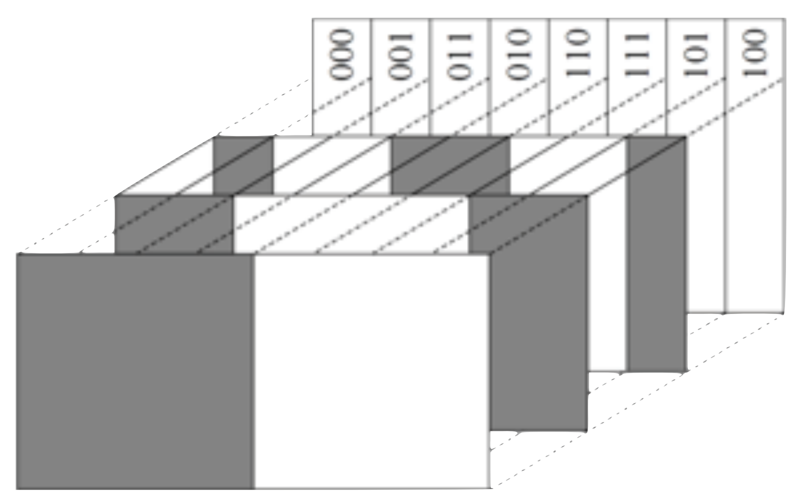
- Wider stripes than naive binary coding
- While same number of patterns, it performs better



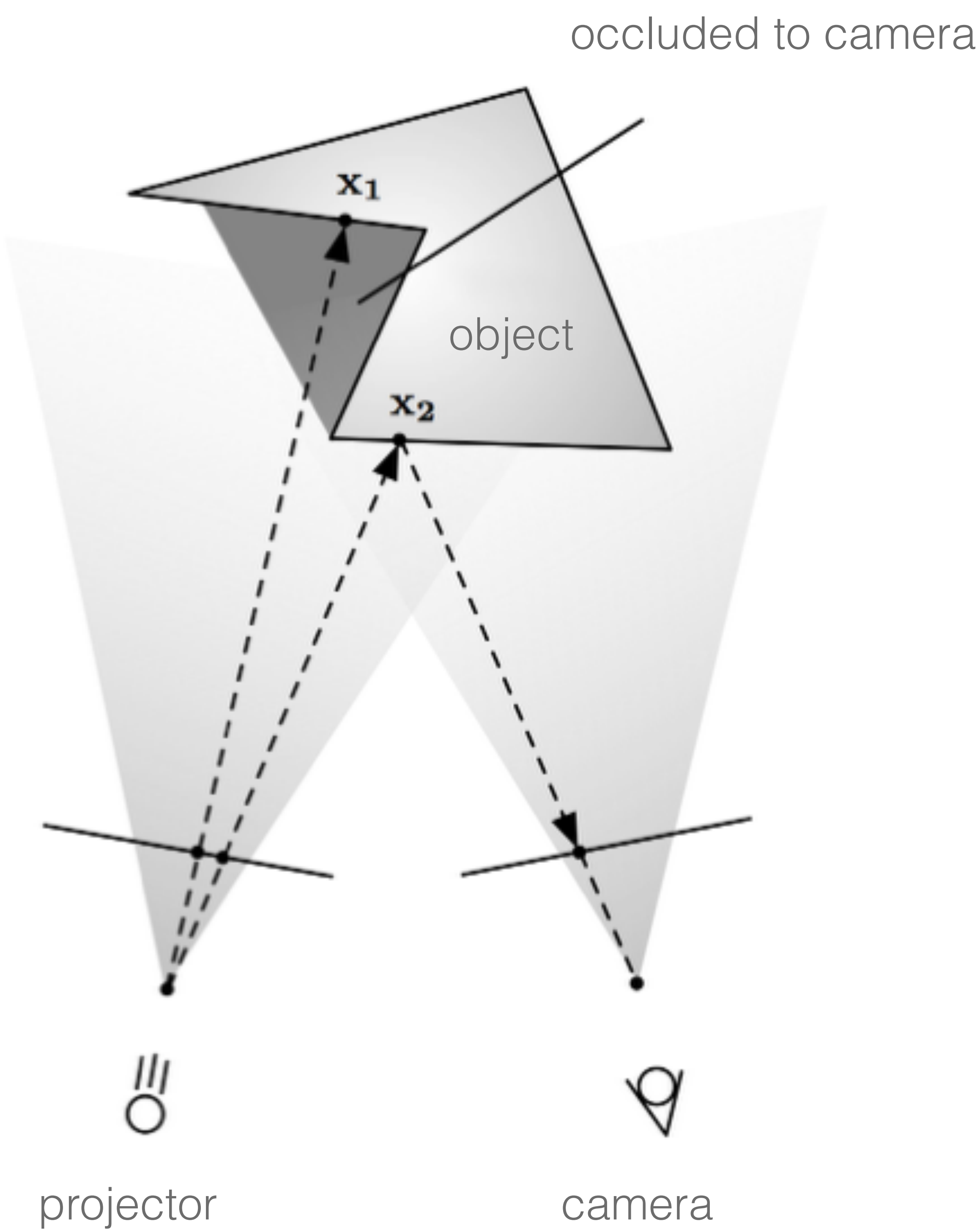
Binary Code

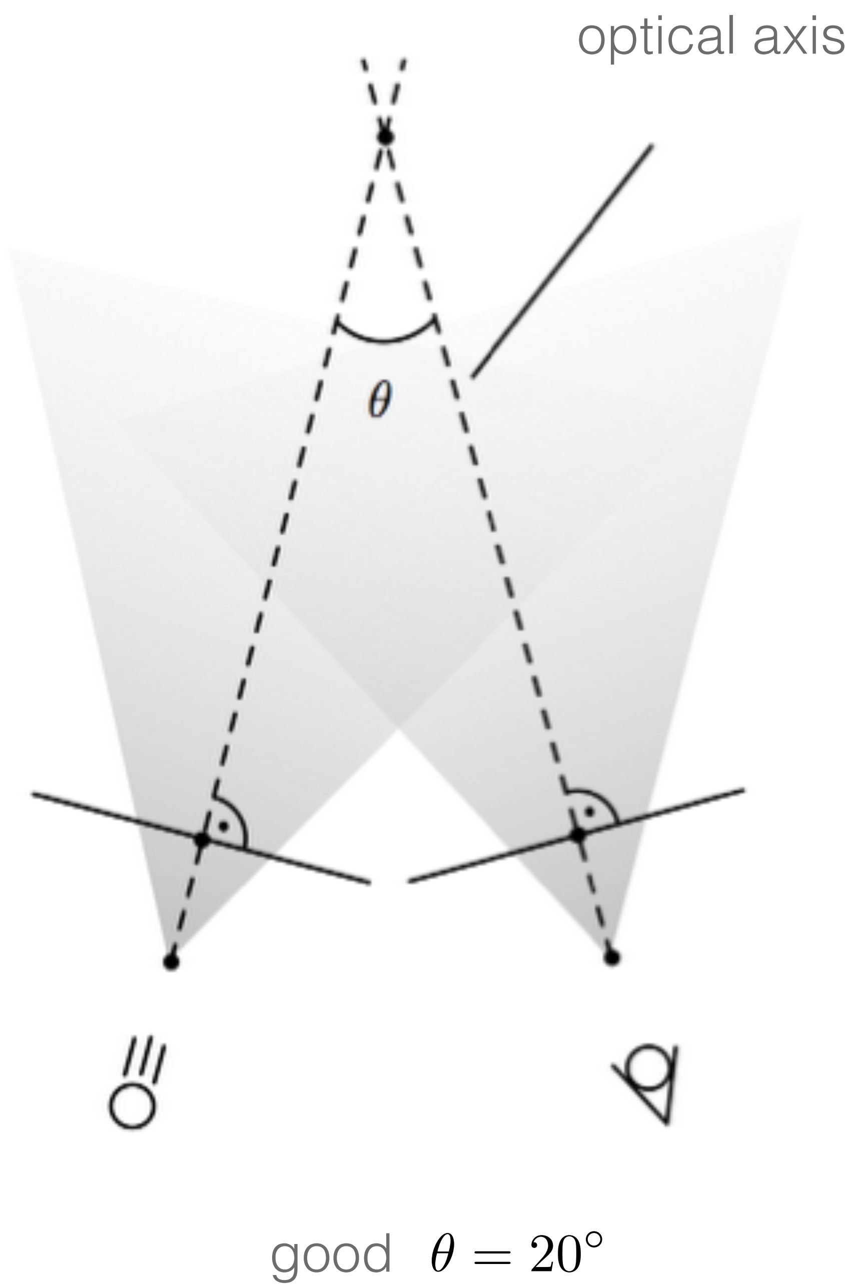
Gray Code

```
Bin2Gray(B,G)
G ← B
for i ← n-1 downto 0
    G[i] ← B[i+1] xor B[i]
```

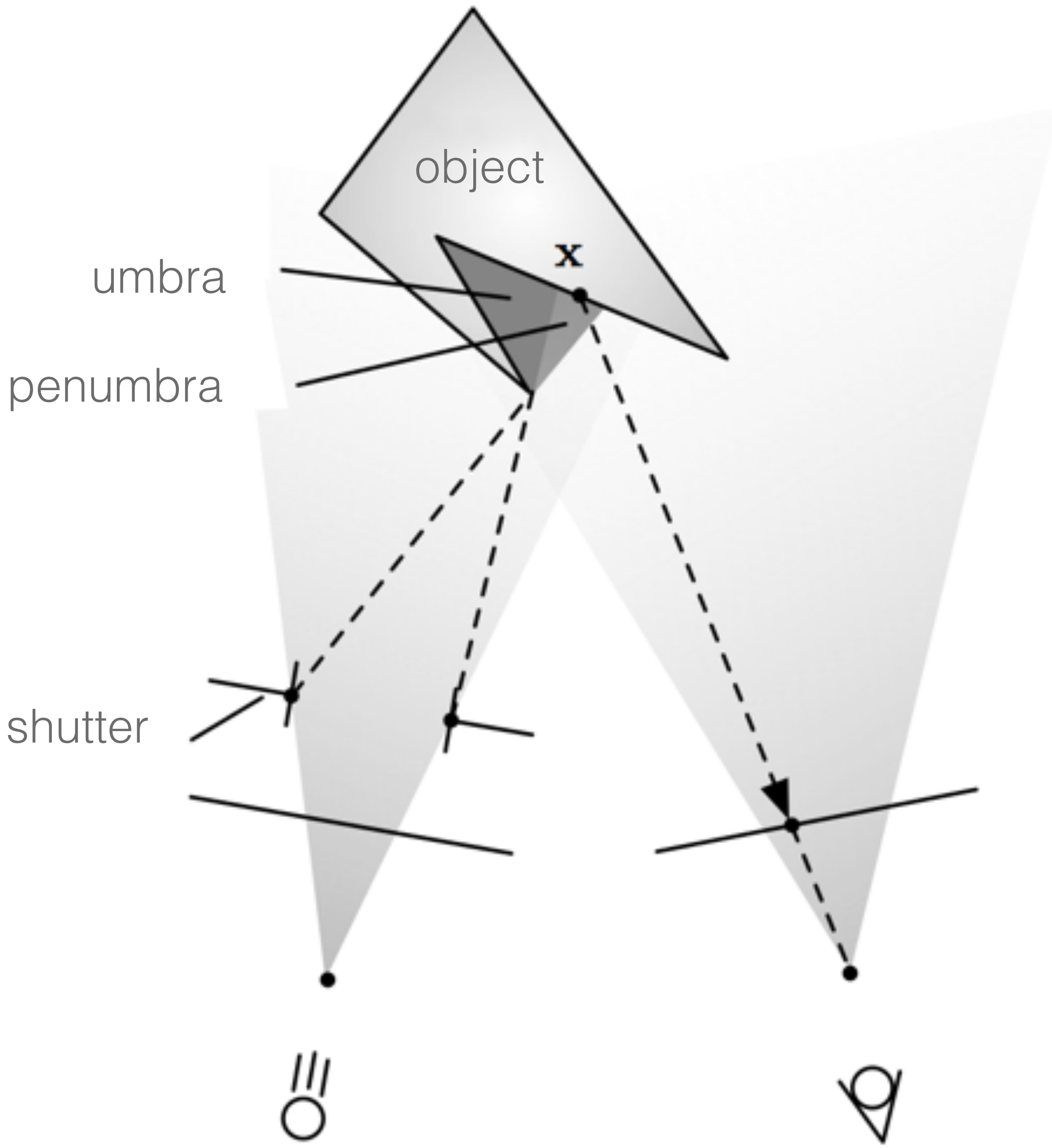


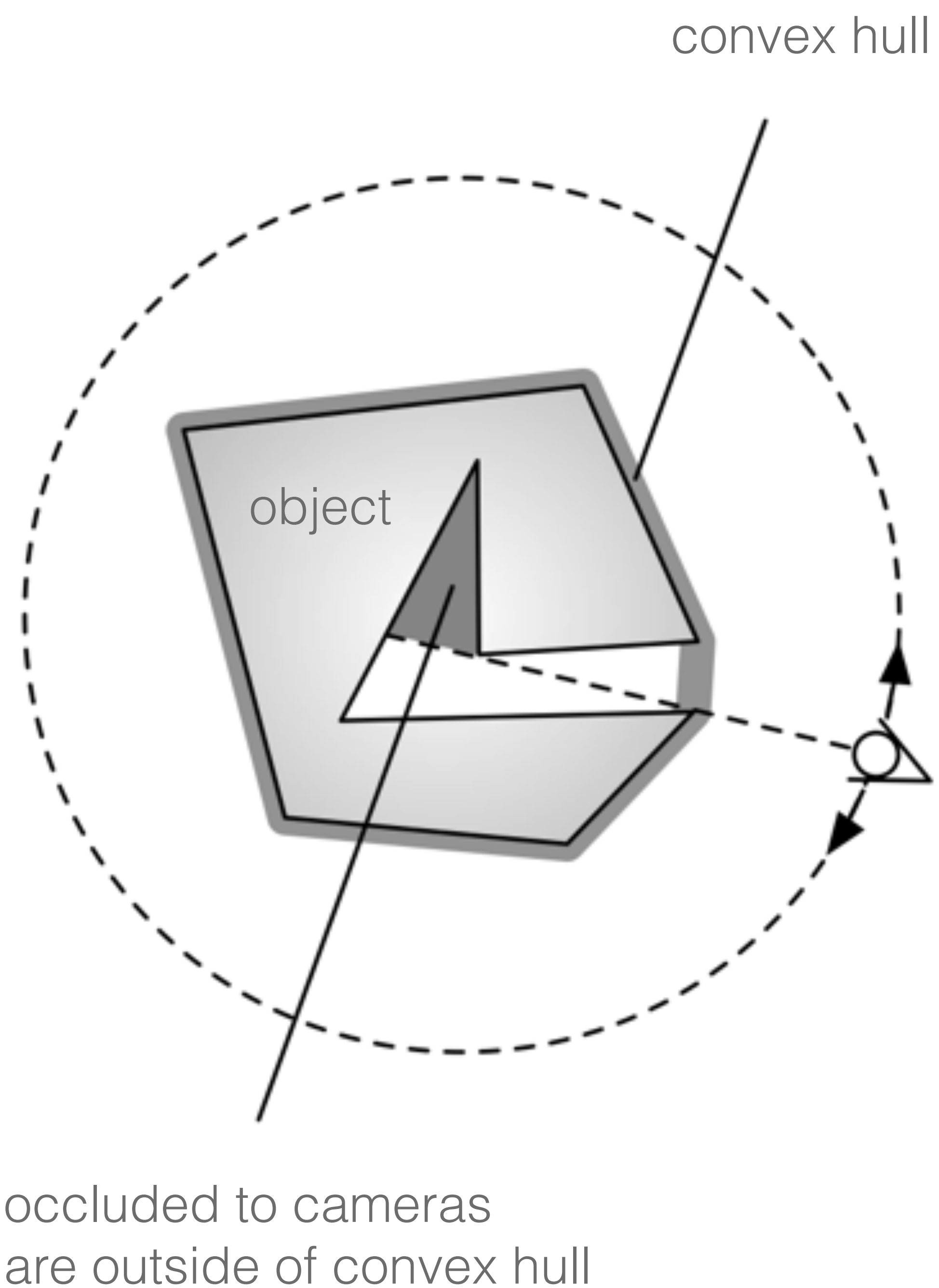
Geometric **Constraints**





Geometric **Constraints**

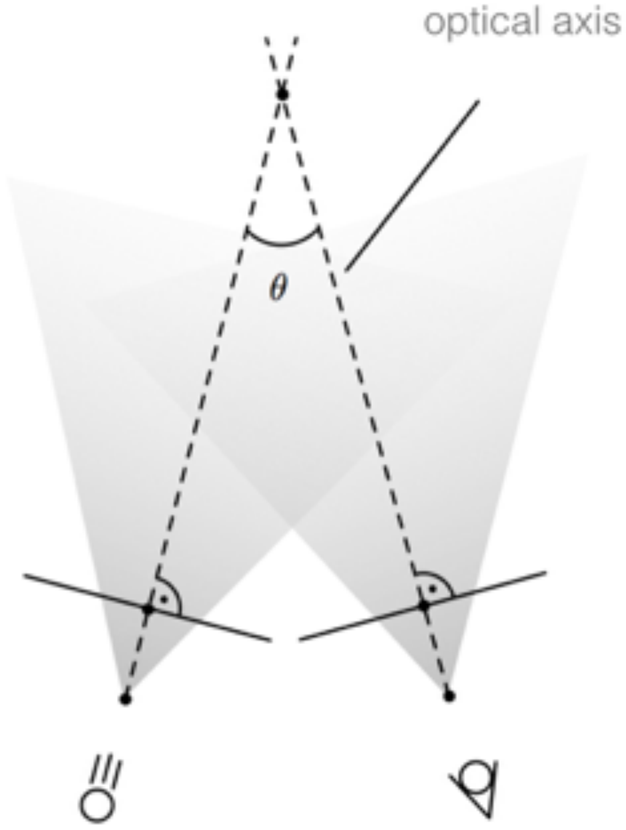




Take Home Message

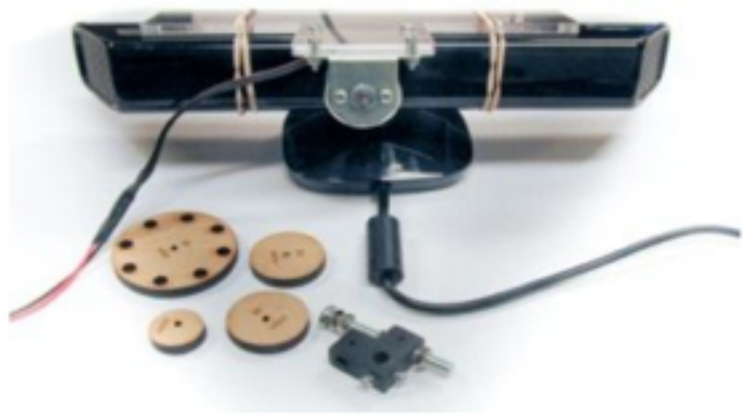
Occlusions in Concave Regions

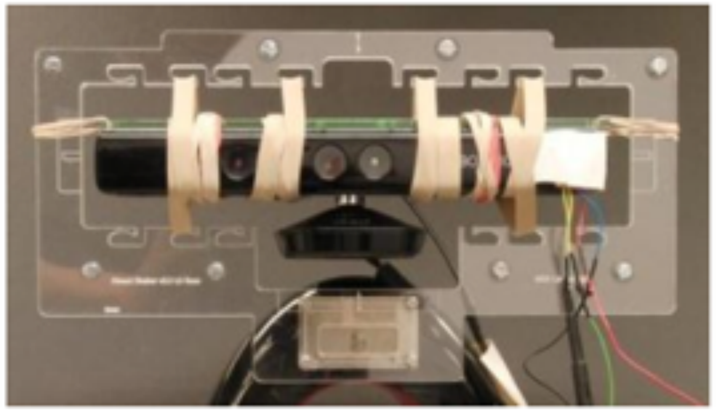
- Longer baseline: more shadowing
- Shorter baseline: less precision
- In practice: $\theta = 20^\circ$



Interference of Patterns

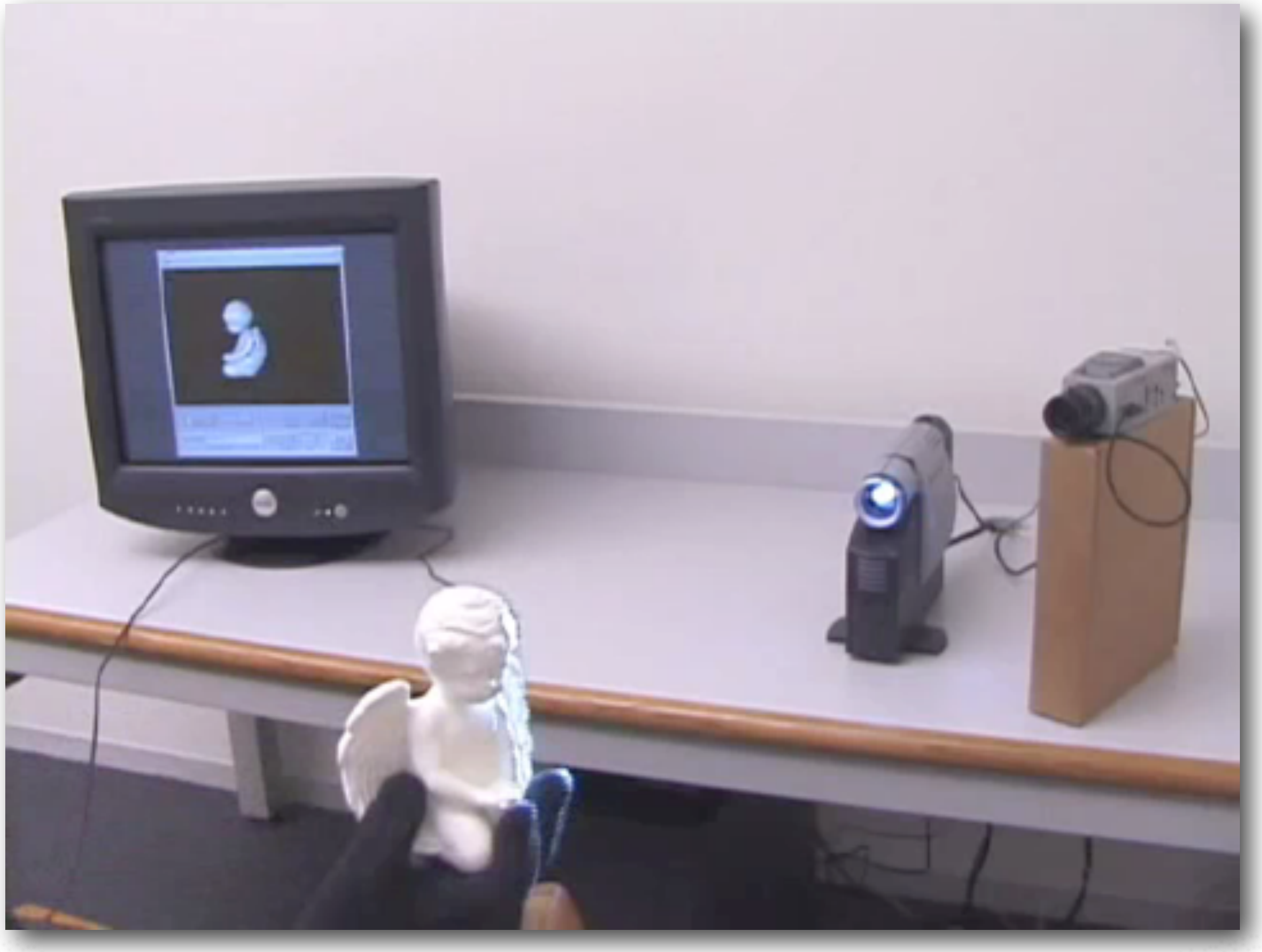
- Challenges for multi-view capture





Shake'n'Sense [MSR 2012]

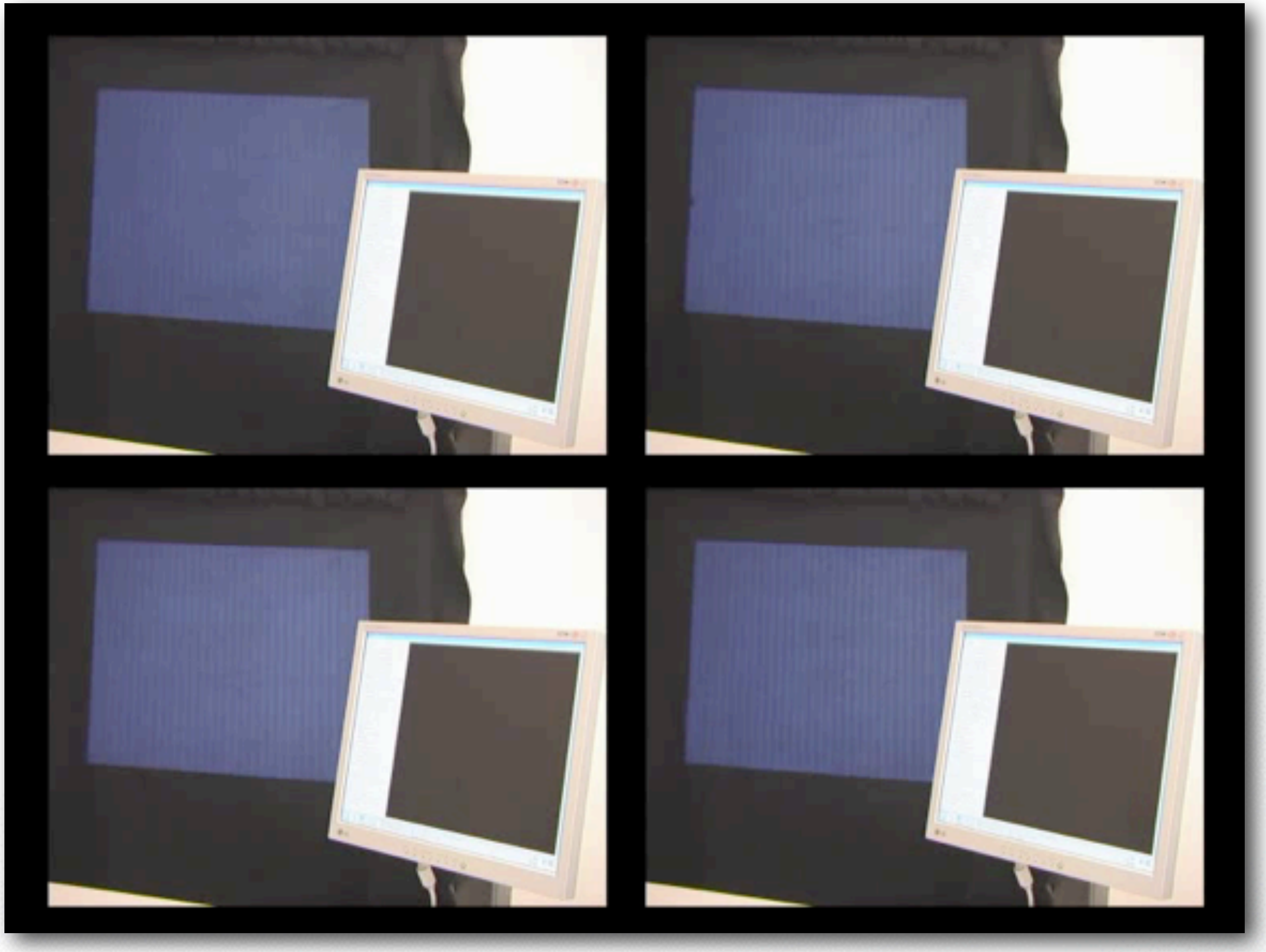
Realtime Structured Light



Phase Shift Patterns

Realtime Depth Capture

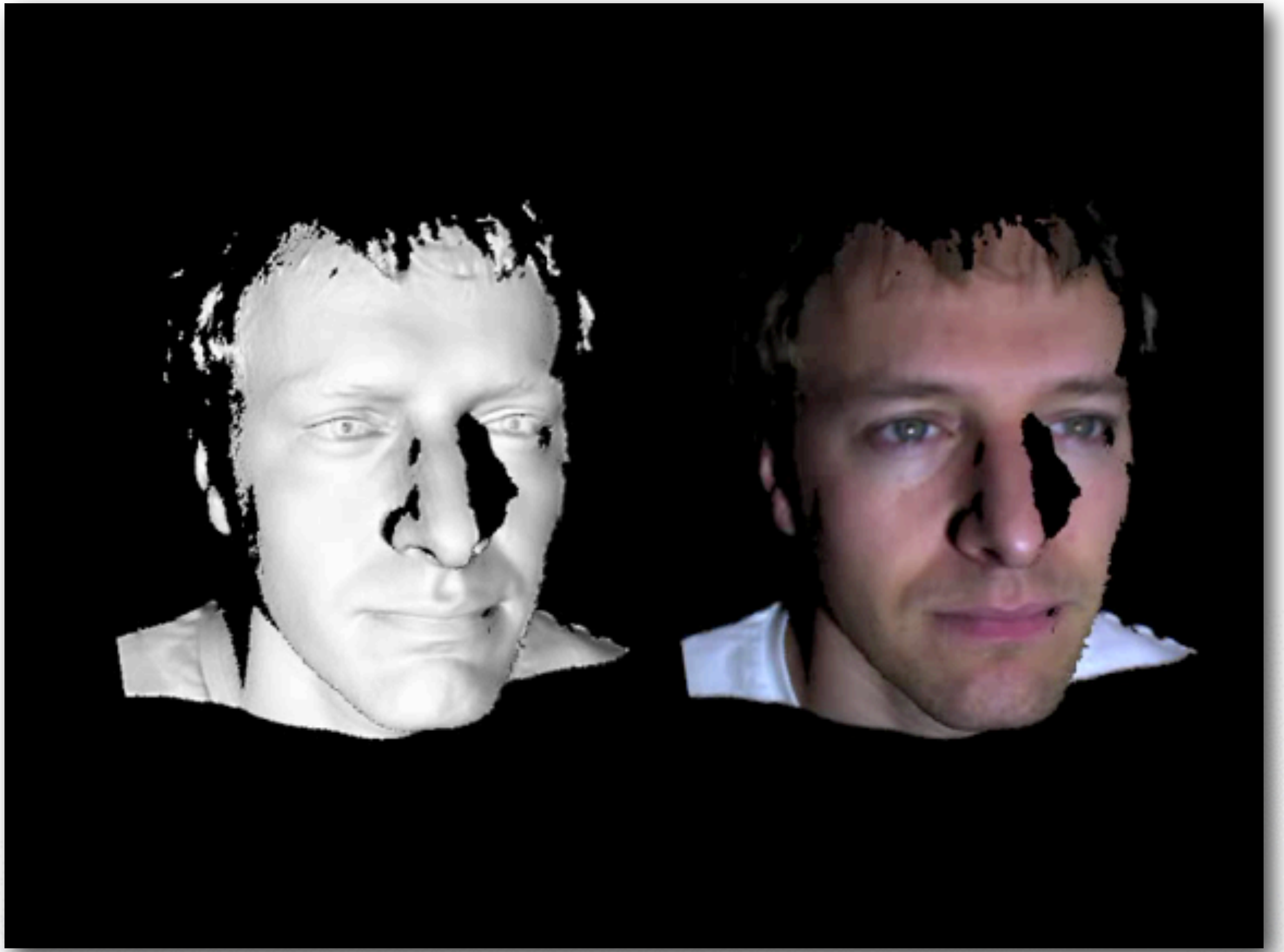
- Moving and Deforming Scenes
- Subpixel accuracy
- High resolution



Phase Shift Patterns

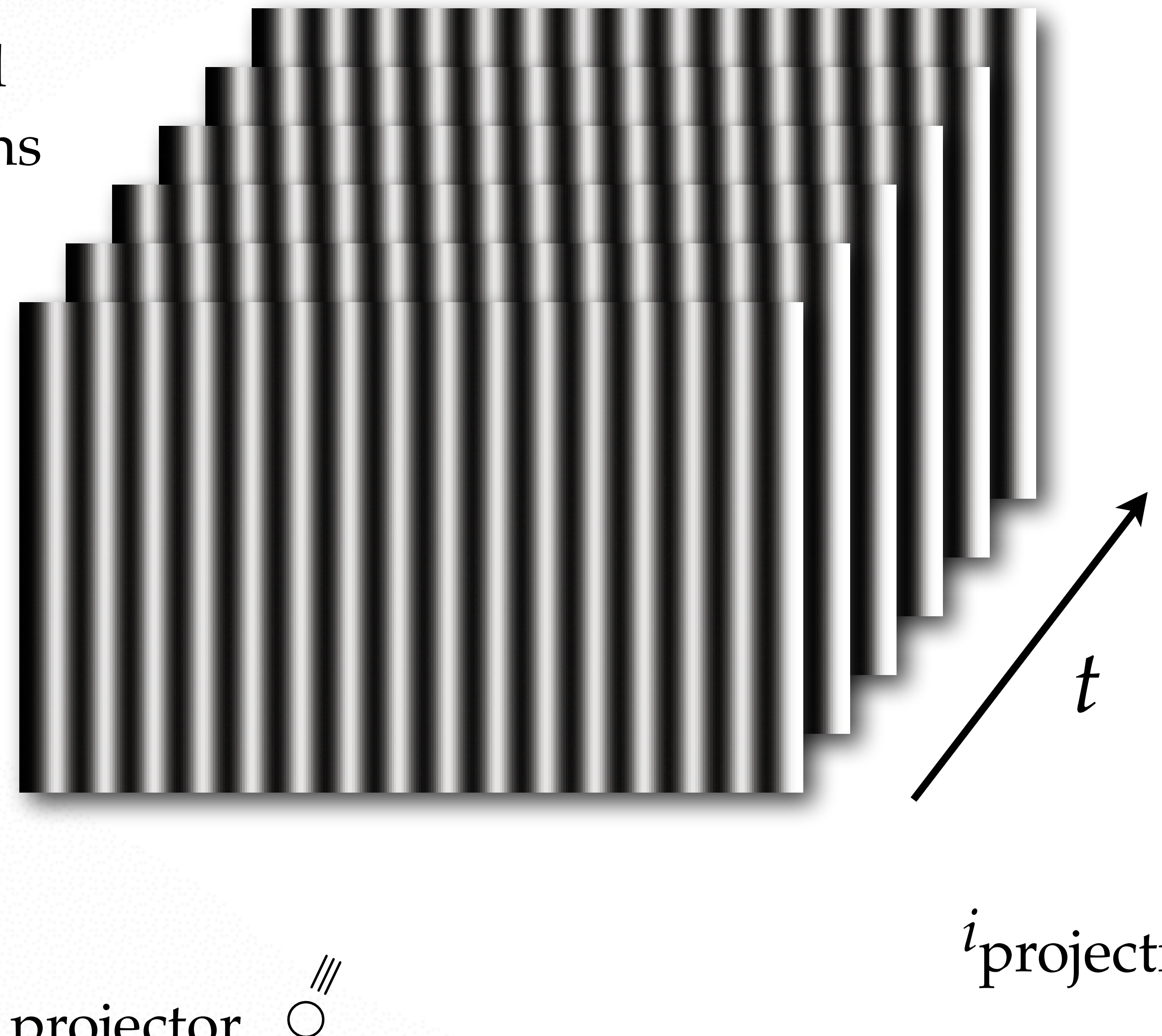
Realtime Depth Capture

- Moving and Deforming Scenes
- Subpixe accuracy
- High resolution



Phase Shift Patterns

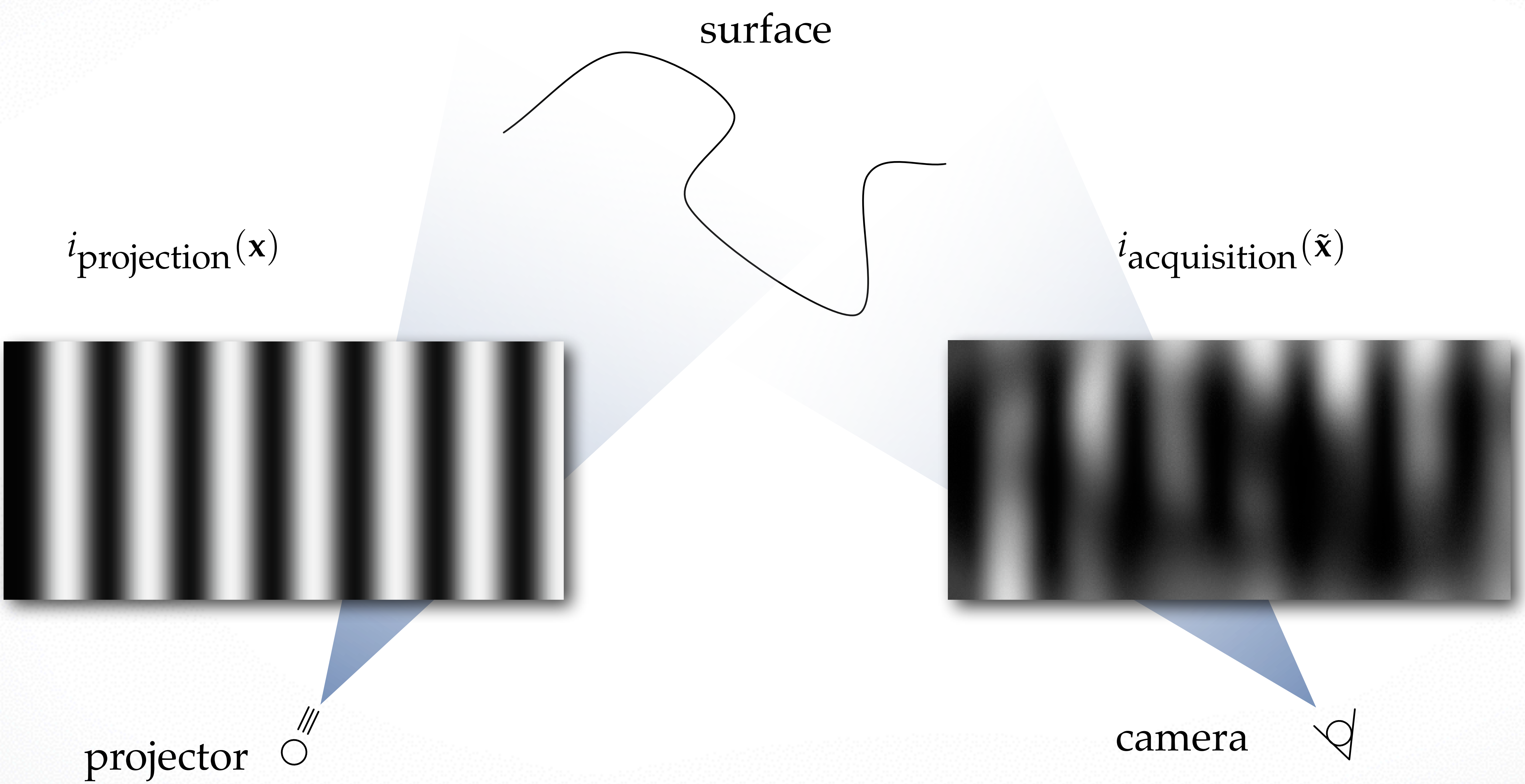
phase
shifted
patterns



$$i_{\text{projection}}(\mathbf{x}, t) = \frac{1}{2}(1 + \cos(\theta(\mathbf{x}) - \phi(t)))$$

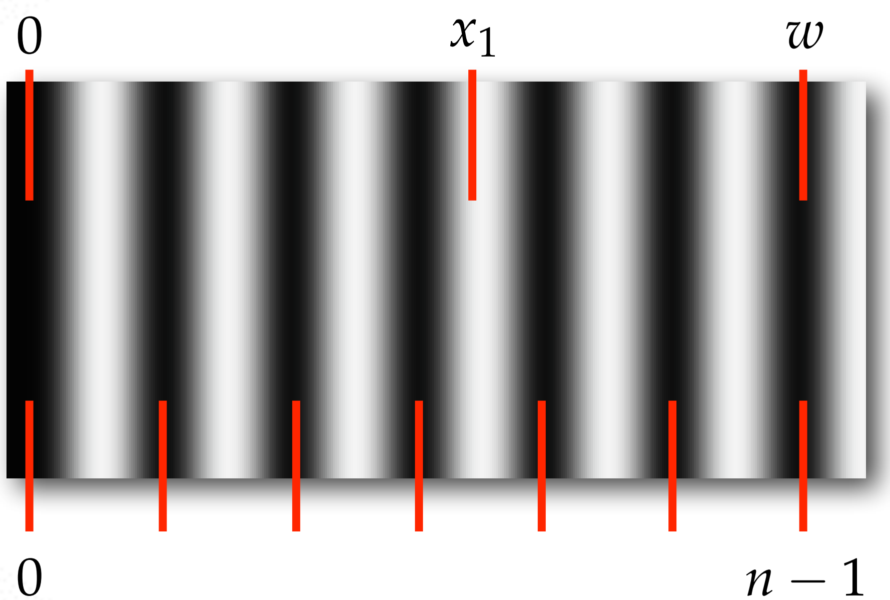
$\mathbf{x} \in \mathbb{R}^2$

Phase Shift Patterns

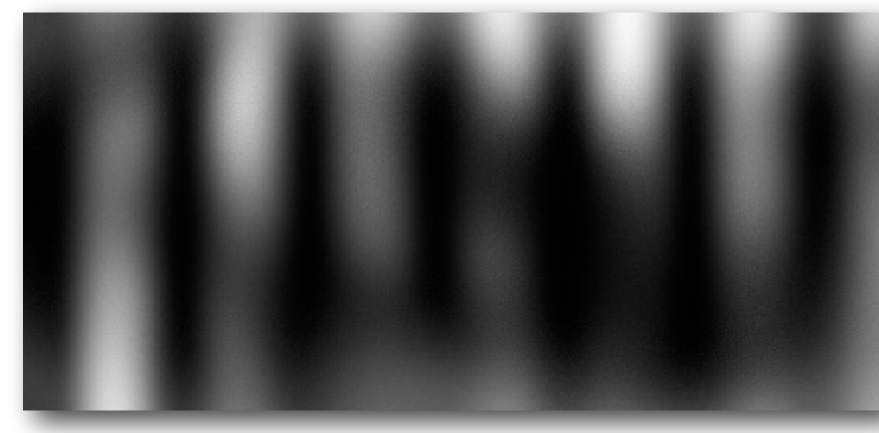


Phase Shift Patterns

$i_{\text{projection}}(\mathbf{x})$



$i_{\text{acquisition}}(\tilde{\mathbf{x}})$



$$i_{\text{acquisition}}(\tilde{\mathbf{x}}) = i_{\text{albedo}}(\tilde{\mathbf{x}}) + i_{\text{amplitude}}(\tilde{\mathbf{x}})\cos(2\pi n \frac{x_1}{w} - \phi)$$

Phase Unwrapping

$$i_{\text{acquisition}}(\tilde{\mathbf{x}}) = i_{\text{albedo}}(\tilde{\mathbf{x}}) + i_{\text{amplitude}}(\tilde{\mathbf{x}})\cos(2\pi n\frac{x_1}{w} - \phi)$$

$$\downarrow$$

$$i_{\text{acquisition}}(\tilde{\mathbf{x}}) = i_{\text{albedo}}(\tilde{\mathbf{x}}) + i_{\text{amplitude}}(\tilde{\mathbf{x}})\cos(\theta - \phi)$$

$$\theta \in [0, 2\pi]$$

Phase Unwrapping

$$i_{\text{acquisition}}(\tilde{\mathbf{x}}) = i_{\text{albedo}}(\tilde{\mathbf{x}}) + i_{\text{amplitude}}(\tilde{\mathbf{x}})\cos(\theta - \phi)$$

$$\downarrow$$

$$i^{(t)}_{\text{acquisition}}(\tilde{\mathbf{x}}) = i_{\text{albedo}}(\tilde{\mathbf{x}}) + i_{\text{amplitude}}(\tilde{\mathbf{x}})\cos(\theta - 2\pi\frac{t}{m})$$

Three unknowns:

$\theta \in [0, 2\pi]$

$i_{\text{albedo}}(\tilde{\mathbf{x}})$

$i_{\text{amplitude}}(\tilde{\mathbf{x}})$

Phase Unwrapping

$$i^{(t)}_{\text{acquisition}}(\tilde{\mathbf{x}}) = i_{\text{albedo}}(\tilde{\mathbf{x}}) + i_{\text{amplitude}}(\tilde{\mathbf{x}})\cos(\theta - 2\pi\frac{t}{m})$$

$$\theta = \arctan(\frac{3^{\frac{1}{2}}(i^1_{\text{acquisition}} - i^3_{\text{acquisition}})}{2i^2_{\text{acquisition}} - i^1_{\text{acquisition}} - i^3_{\text{acquisition}}})$$

$$i_{\text{albedo}}(\tilde{\mathbf{x}}) = \frac{1}{3}\sum_{t=1}^{3} i^t_{\text{acquisition}}(\tilde{\mathbf{x}})$$

$$i_{\text{amplitude}}(\tilde{\mathbf{x}}) = (\frac{(i^3_{\text{acquisition}} - i^1_{\text{acquisition}})^2}{3} + \frac{(2i^2_{\text{acquisition}} - i^1_{\text{acquisition}} - i^3_{\text{acquisition}})^2}{9})^{\frac{1}{2}}$$

Phase Unwrapping

phase solution is unique only up to period...

$$\downarrow$$

phase "unwrapping"

$$\downarrow$$

$$\tilde{\theta}(\tilde{\mathbf{x}}) = \theta(\tilde{\mathbf{x}}) + 2\pi k(\tilde{\mathbf{x}})$$

$$k \in [0, n-1]$$

Kinect for XBOX 360

Kinect (= 1st gen Kinect)

- Structured Light Technology (Primesense Sensor)
- 640 x 680 @ 30 fps
- 1280x960 @ 12 fps
- accuracy:
  - < a few mm (depth < 50 cm)
  - < 4 cm (depth < 500 cm)
- VGA for RGB input
- uses Kinect1.x SDK

Summary

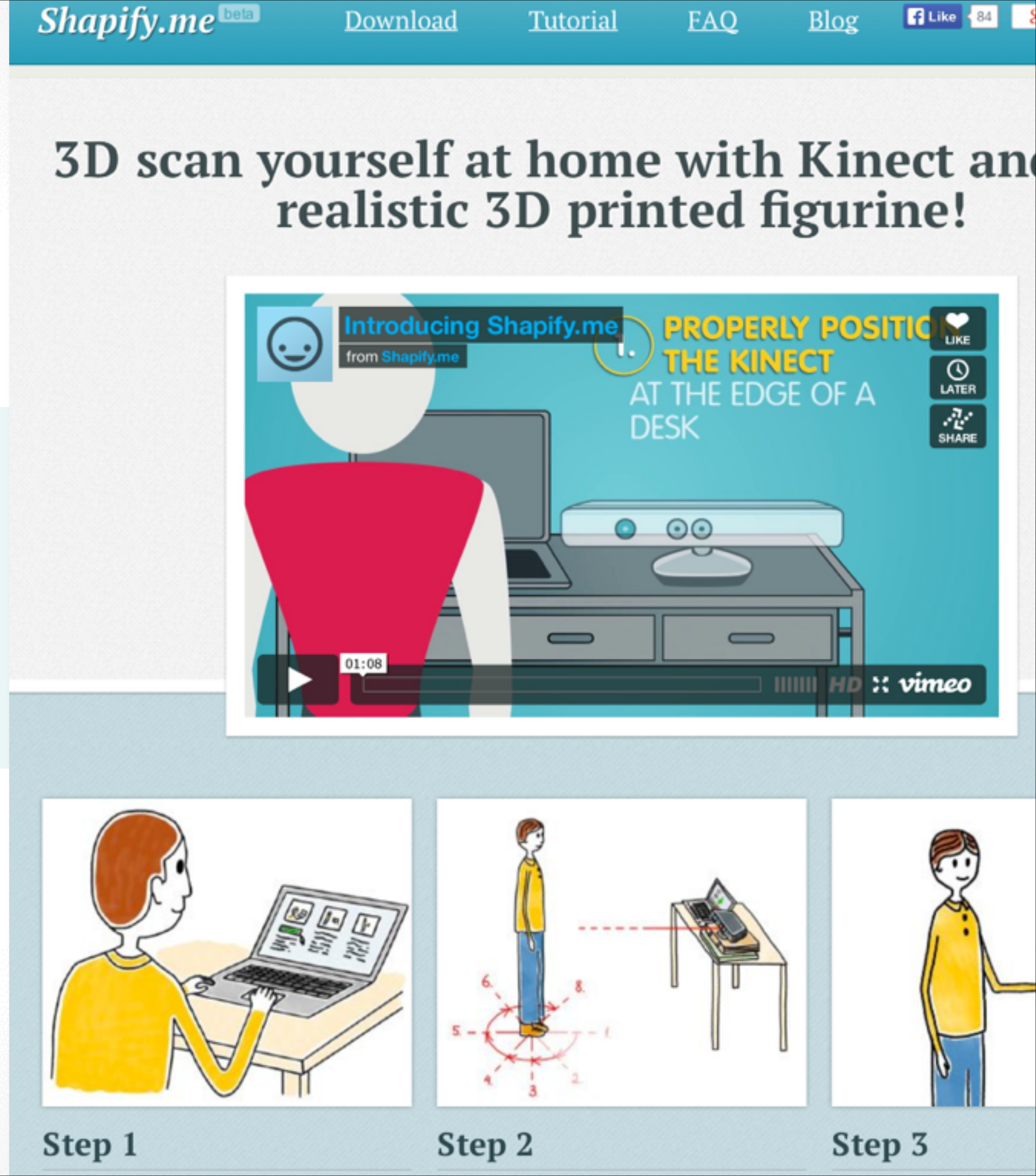
The Future will be more accessible

- Real-time depth sensors (smaller, more accurate, higher resolution, less noise, larger working volume, portable)
  - TOF, structured Light, camera Arrays
- Multi-view stereo capture (sparser, better algorithms, real-time, very large working volume, high speed, portable)
  - Robotic camera tracking

tracking a ping pong ball

http://**shapify.me**

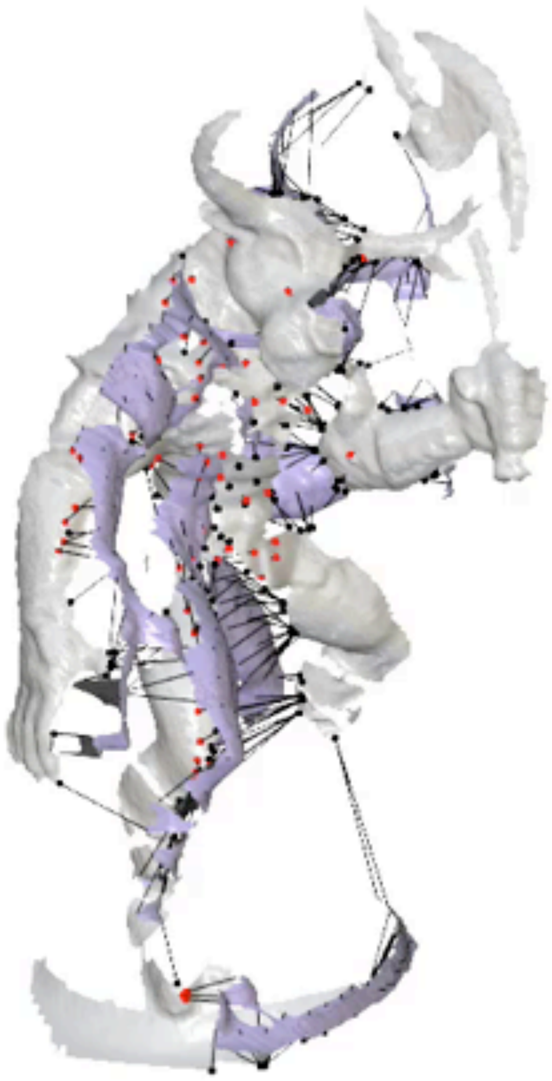
presented by Artec Group



Literature

- Lanman and Taubin, “Build Your Own 3D Scanner: Optical Triangulation for Beginners”, SIGGRAPH 2009 Courses
- Curless, “New Methods for Surface Reconstruction from Range Images”, PhD Thesis, Stanford University 1997
- Levoy et al., “Digital Michelangelo Project”, Stanford 1997 - 2000
- Zhang, “www.me.iastate.edu/directory/faculty/song-zhang/“
- Newcombe & Davison, “Live Dense Reconstruction with a Single Moving Camera”, CVPR 2010

Next Time



Surface Registration

http://cs599.hao-li.com

Thanks!