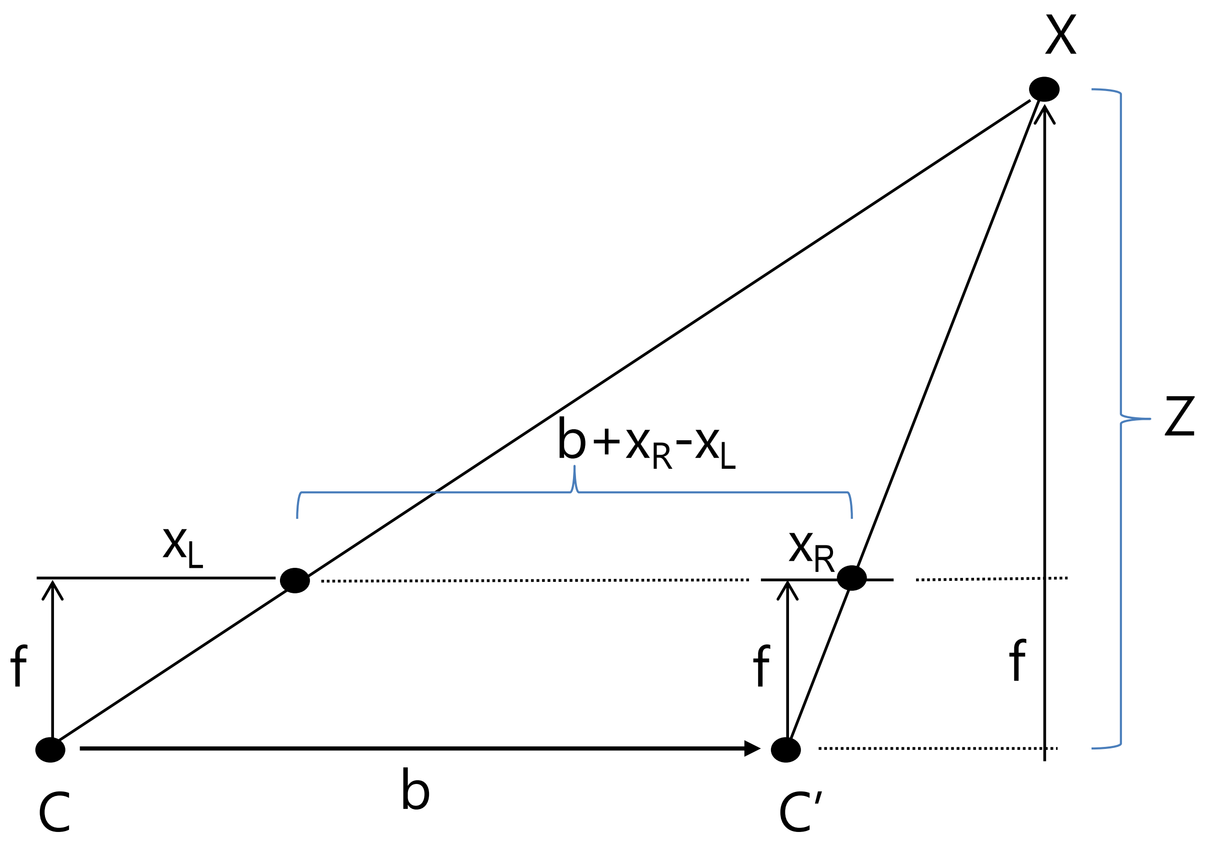
Folien zur Vorlesung am 13.05.2025

3D Computer Vision

TRIANGULATION

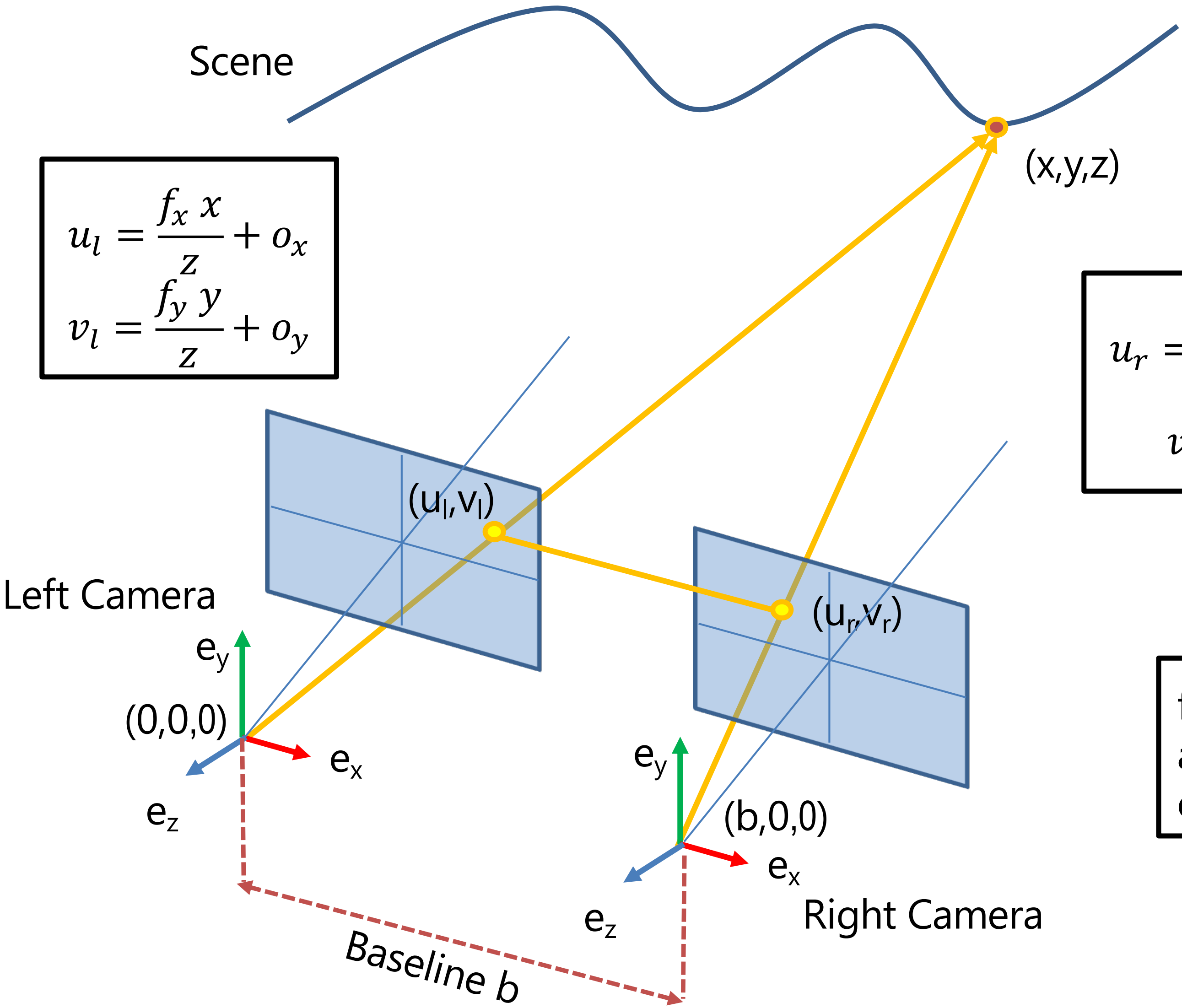
Depth from disparity



disparity d $\quad d = x_L - x_R$

baseline b $\quad d = b*f / Z$

Simple (calibrated) Stereo



$$u_l = \frac{f_x\, x}{z} + o_x$$
$$v_l = \frac{f_y\, y}{z} + o_y$$

$$u_r = \frac{f_x\,(x - b)}{z} + o_x$$
$$v_r = \frac{f_y\, y}{z} + o_y$$

f_x, f_y, b, o_x, o_y in pixel units and known from the camera calibration

Slide inspired by Shree Nayar: Simple Stereo

Simple (calibrated) Stereo

- Solving for (x, y, z)

$$\boxed{x = \frac{b\,(u_l - o_x)}{(u_l - u_r)}} \qquad \boxed{y = \frac{b\,f_x\,(v_l - o_y)}{f_y\,(u_l - u_r)}} \qquad \boxed{z = \frac{b\,f_x}{(u_l - u_r)}}$$

- where $(u_l - u_r)$ is called the disparity

[Slide inspired by Shree Nayar]

Real-time stereo

Nomad robot searched for meteorites in Antartica in 1997

- Used for robot navigation (and other tasks)
  - Several real-time stereo techniques have been developed (most based on simple discrete search)

Stereo reconstruction pipeline

- Steps
  - Calibrate cameras
  - Rectify images
  - Compute disparity
  - Estimate depth

What will cause errors?

- Camera calibration errors
- Poor image resolution
- Occlusions
- Violations of brightness constancy (specular reflections)
- Large motions
- **Low-contrast image regions**

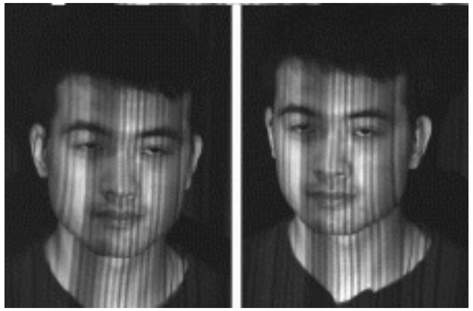
Active stereo with structured light

https://ios.gadgethacks.com/news/watch-iphone-xs-30k-ir-dots-scan-your-face-0180944/

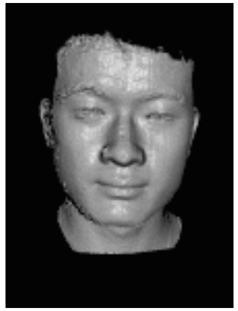
Active stereo with structured light

https://www.sick.com/de/en/machine-vision/3d-machine-vision/visionary-s/c/g507251

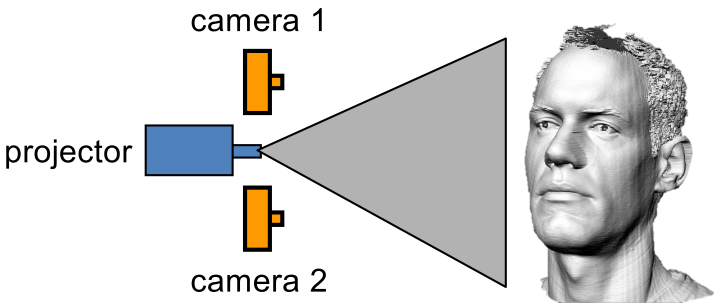
Active stereo with structured light



→

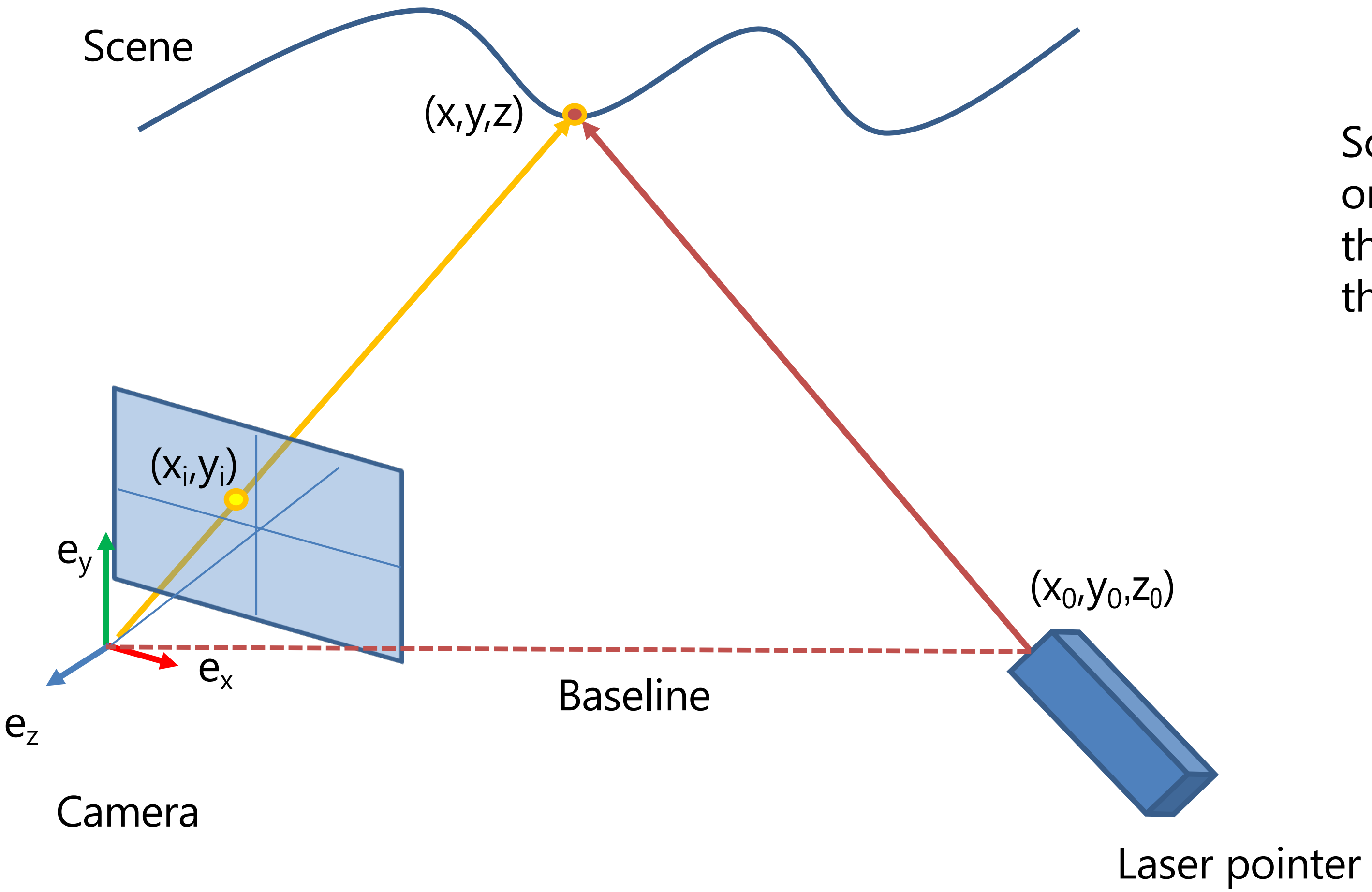


[Li Zhang's Spacetime Stereo]



- Project "structured" light patterns onto the object
  - simplifies the correspondence problem
  - basis for active depth sensors, such as Kinect and iPhone X (using IR)

Using a light beam instead of a second camera



Scene point (x,y,z) lies on the intersection of the Camera ray and the light ray.

Slide inspired by Shree Nayar

Using a light beam instead of a second camera

Capture background image and images with laser point and subtract them to find the coordinates (x_i, y_i).

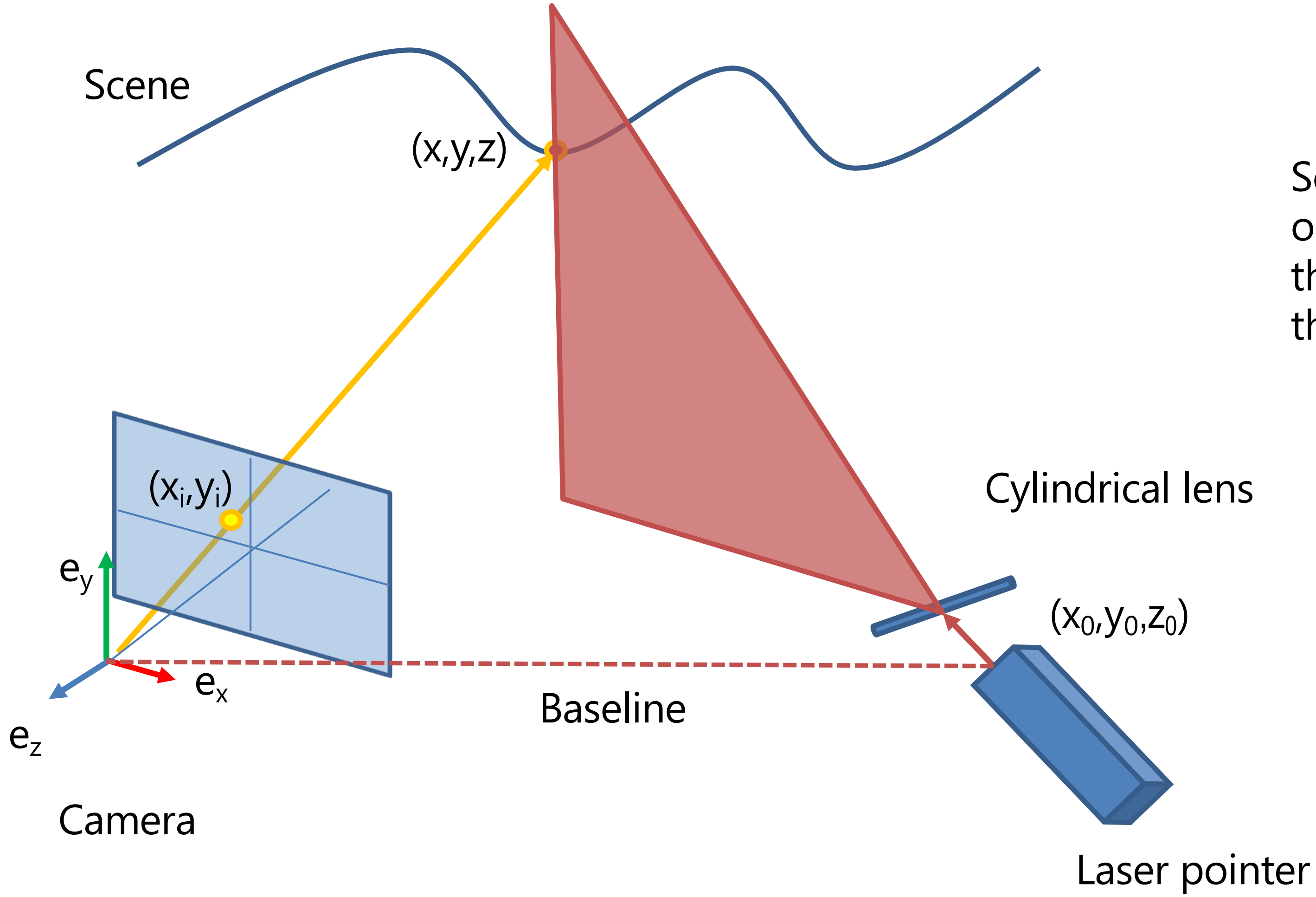
However, for each point in the depth map, one image is needed.

For 640x480 @ 30Hz → ~3 hours acquisition time!

Image: University of Tsukuba

[Slide inspired by Shree Nayar]

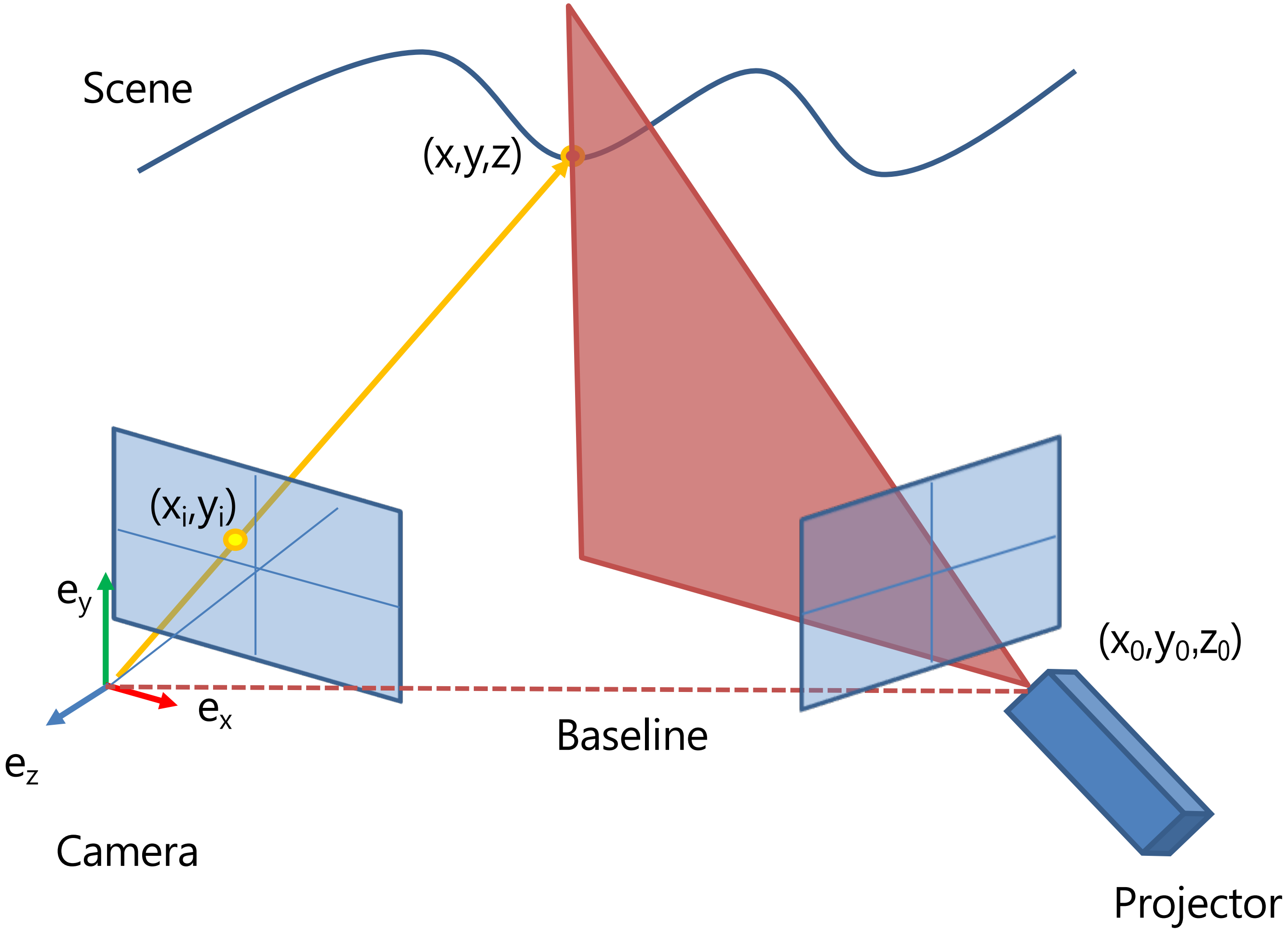
Using a light plane instead of a second camera



Scene point (x,y,z) lies on the intersection of the Camera ray and the light plane.

Slide inspired by Shree Nayar

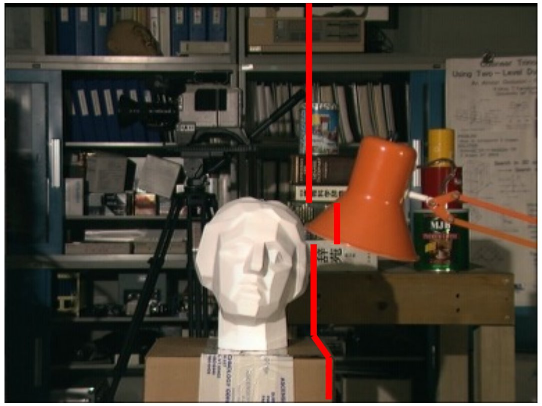
Using a projector instead of a second camera



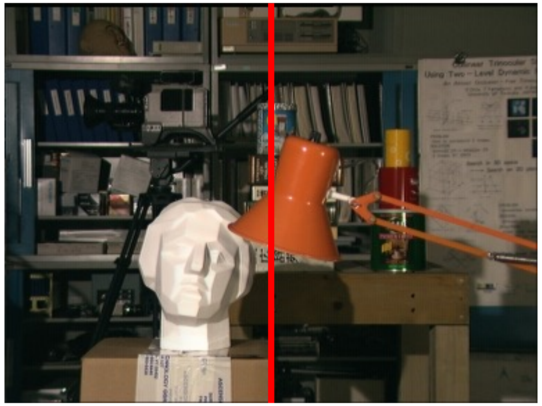
Scene point (x,y,z) lies on the intersection of the Camera ray and the light plane.

[Slide inspired by Shree Nayar]()

Using a projector instead of a second camera



What the camera sees



What the projector „sees“

Capture images with laser lines and find all the coordinates (x_i, y_i) for the whole column.

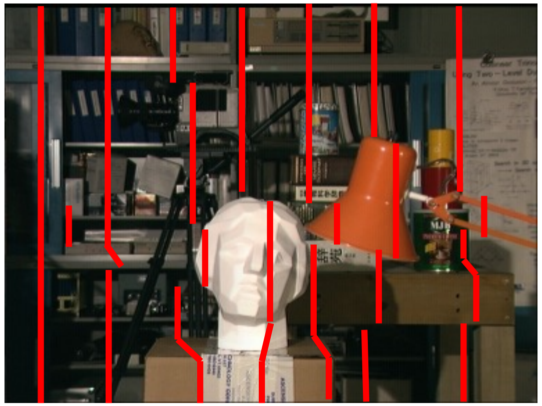
Now only for each column in the depth map, one image is needed.

Still 640 @ 30Hz → ~21 seconds acquisition time!

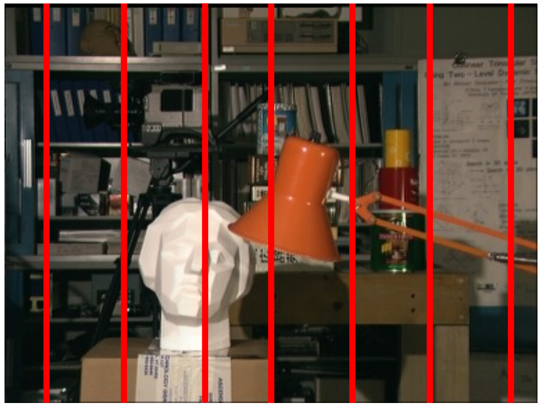
Image: University of Tsukuba

[Slide inspired by Shree Nayar]

Using a projector instead of a second camera



What the camera sees



What the projector „sees“

Can we capture less images with multiple laser lines and find all the coordinates (x_i, y_i) for all the columns?

No! There is no guarantee that the order is preserved.

Image: University of Tsukuba

Slide inspired by Shree Nayar

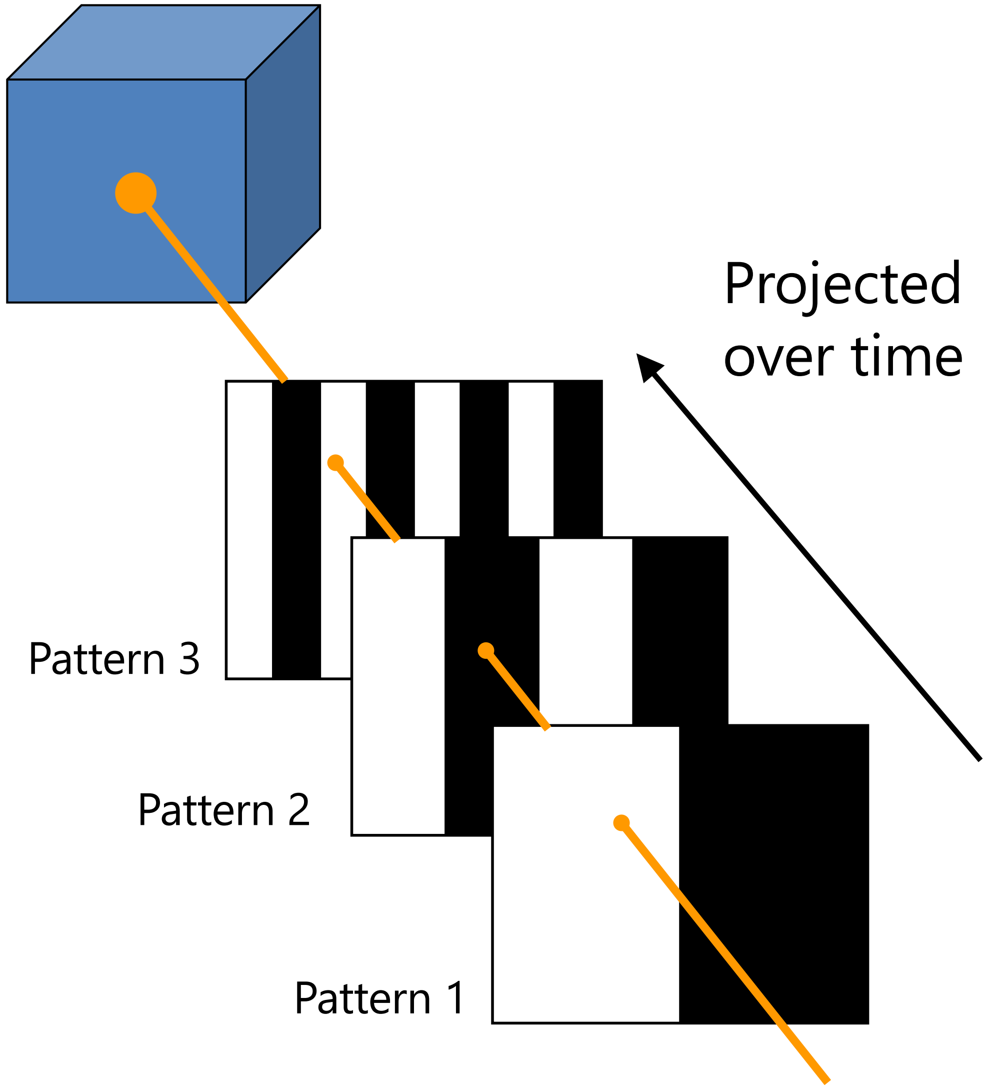
Binary Coding

Faster:

$2^n - 1$ stripes in n images.

Example:

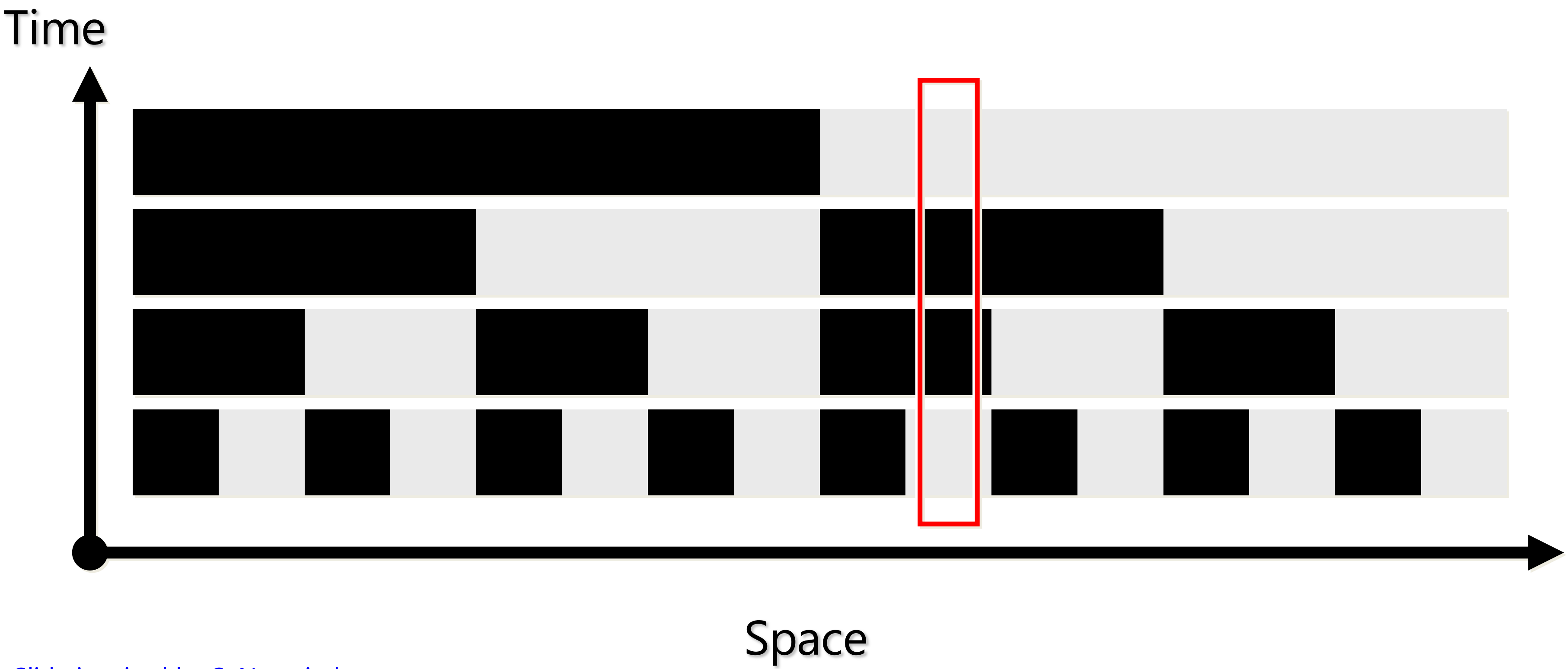
3 binary-encoded patterns which allows the measuring surface to be divided in 8 sub-regions



Slide inspired by S. Narasimhan

Binary Coding

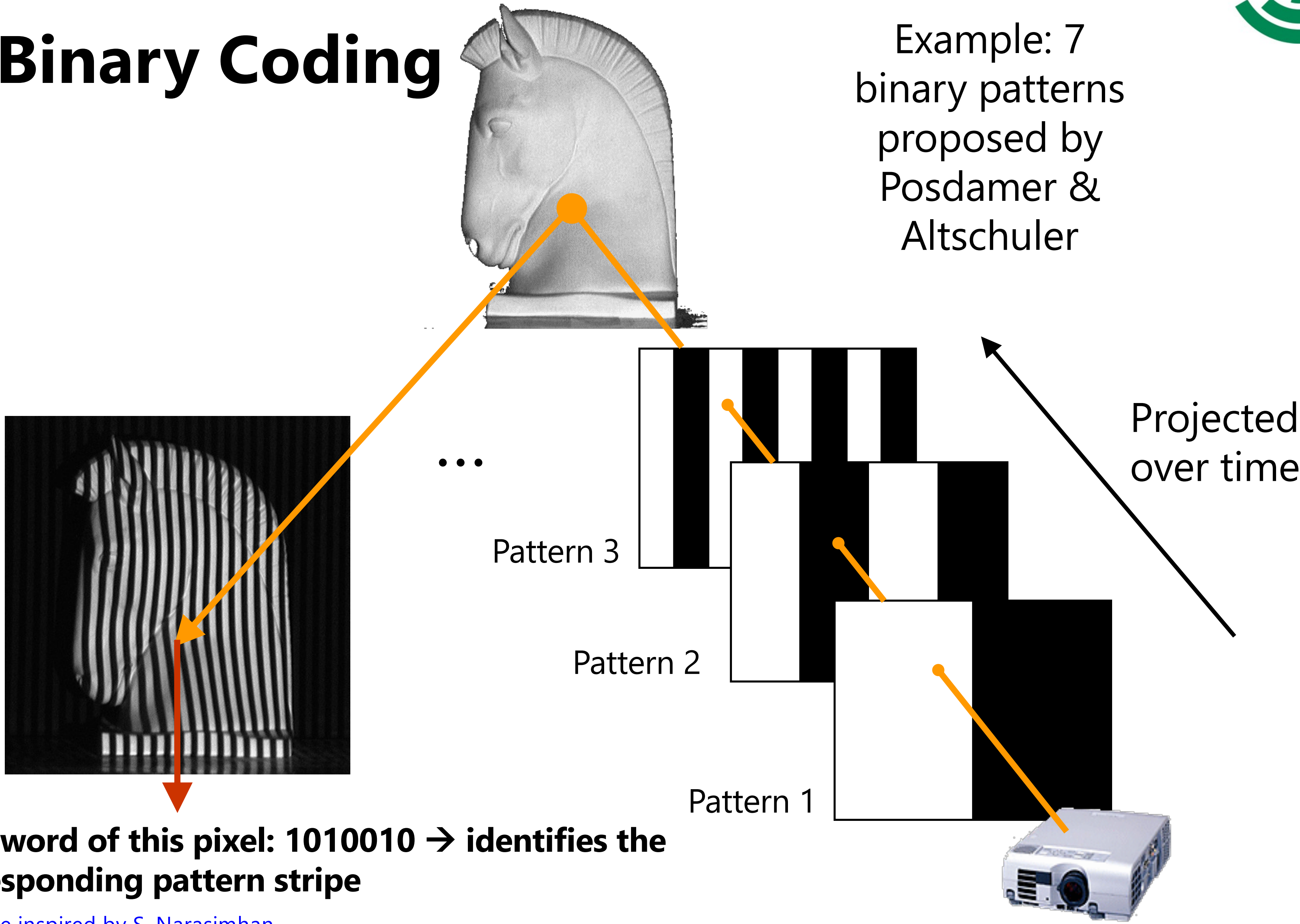
- Assign each stripe a unique illumination code over time [Posdamer 82]



Slide inspired by S. Narasimhan

Binary Coding

Example: 7 binary patterns proposed by Posdamer & Altschuler



Codeword of this pixel: 1010010 → identifies the corresponding pattern stripe

Slide inspired by S. Narasimhan

Strutured light scanning

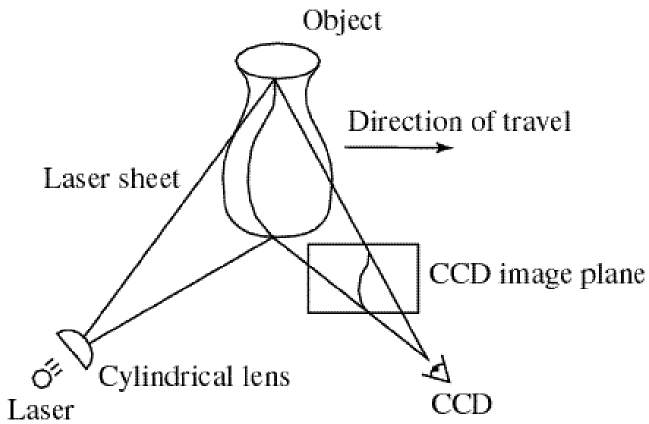
Further improvements developed:

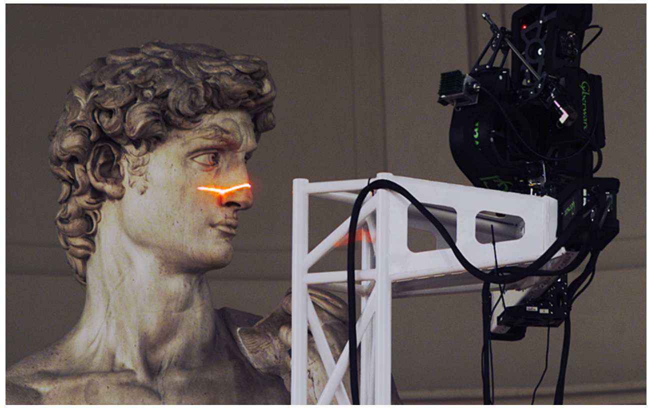
- „Gray code“, avoiding light bleeding artefacts.
- Using color projection to get k-ary coding
- Using intensity variation functions (aka phase shifting method) allows a reduction to only three images in principle.

Major drawback:

- Projector-camera calibration is needed and precision highly depends on that.

Laser scanning





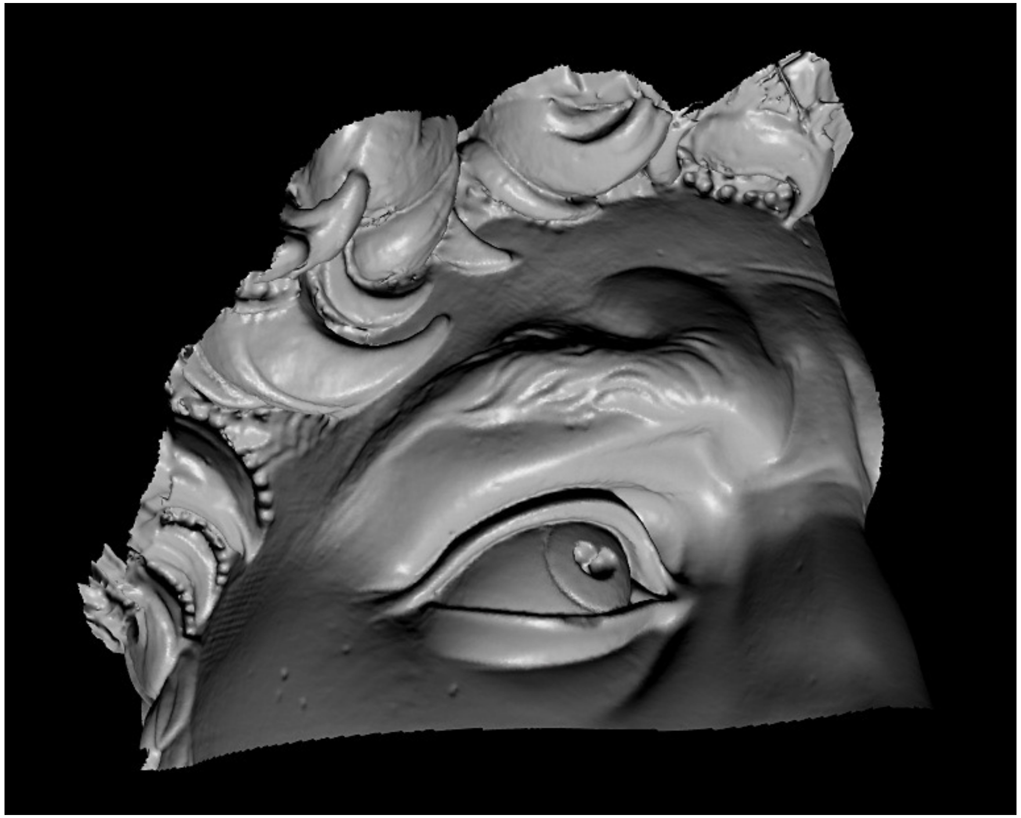
Digital Michelangelo Project (2000)
http://graphics.stanford.edu/projects/mich/

- Optical triangulation
  - Project a single stripe of laser light
  - Scan it across the surface of the object
  - This is a very precise version of structured light scanning

Laser scanned models

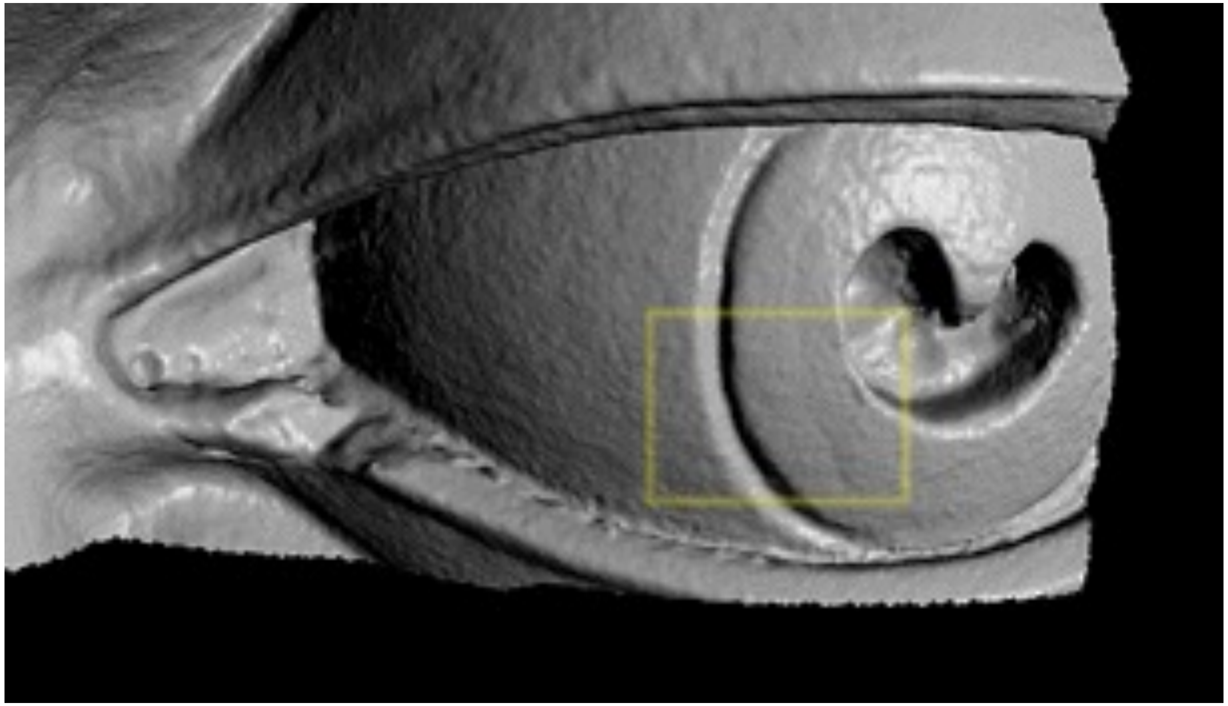
The Digital Michelangelo Project, Levoy et al.

Laser scanned models



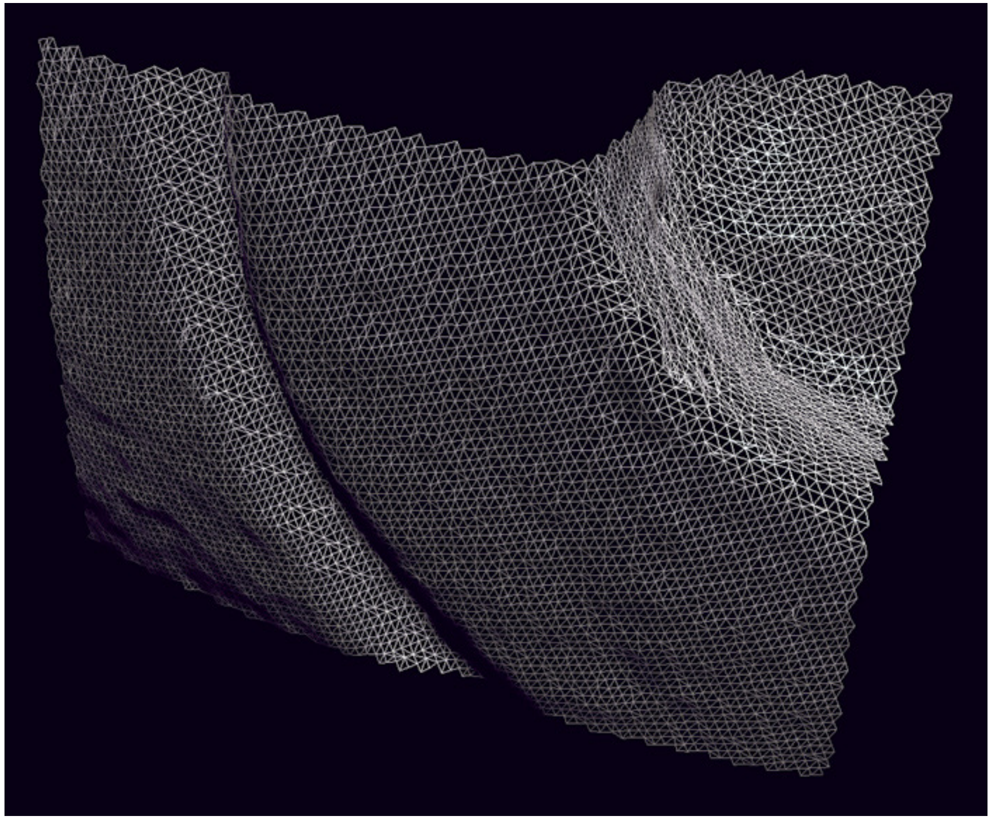
The Digital Michelangelo Project, Levoy et al.

Laser scanned models



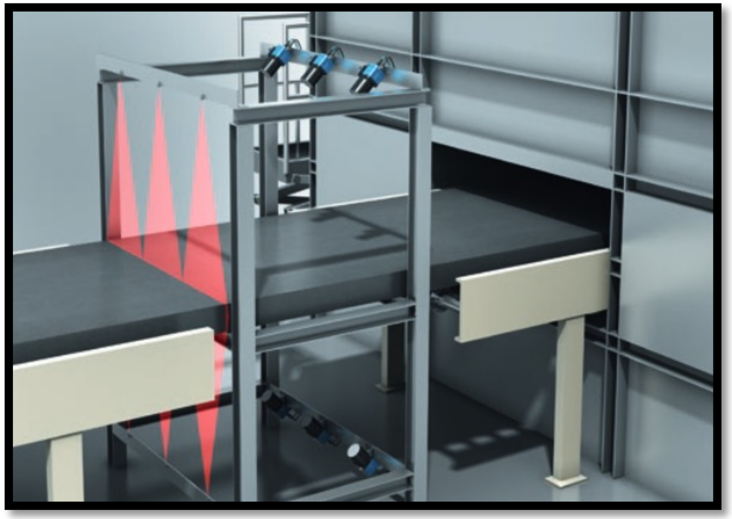
The Digital Michelangelo Project, Levoy et al.

Laser scanned models



The Digital Michelangelo Project, Levoy et al.

Industrial triangulation – SICK Ranger 3



Images from https://cdn.sick.com/media/docs/8/48/448/product_information_ranger3_3d_vision_en_im0080448.pdf

Industrial triangulation – SICK Ranger 3

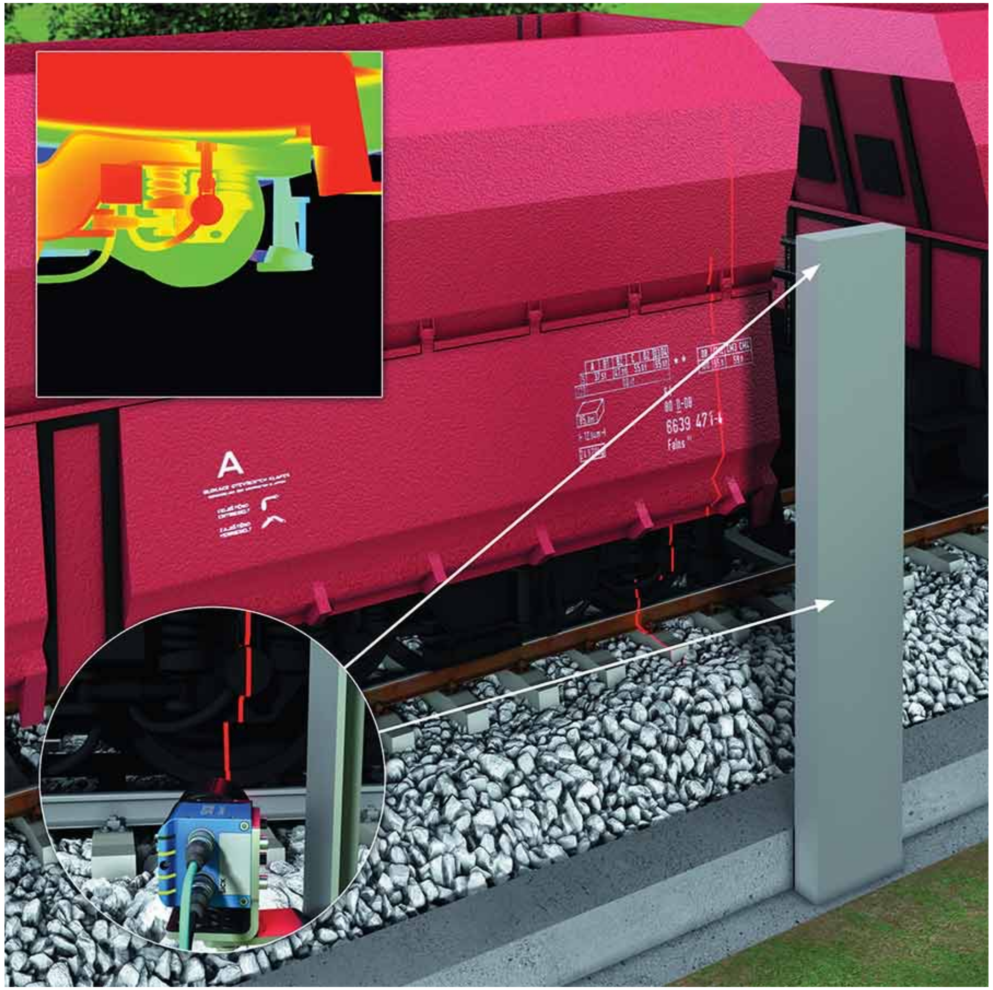


Image from Mattias Johannesson, Laser Triangulation Tackles Imaging Tasks Big and Small, https://www.photonics.com/Articles/Laser_Triangulation_Tackles_Imaging_Tasks_Big_and/a64616 (2019)

Industrial triangulation – SICK Ranger 3

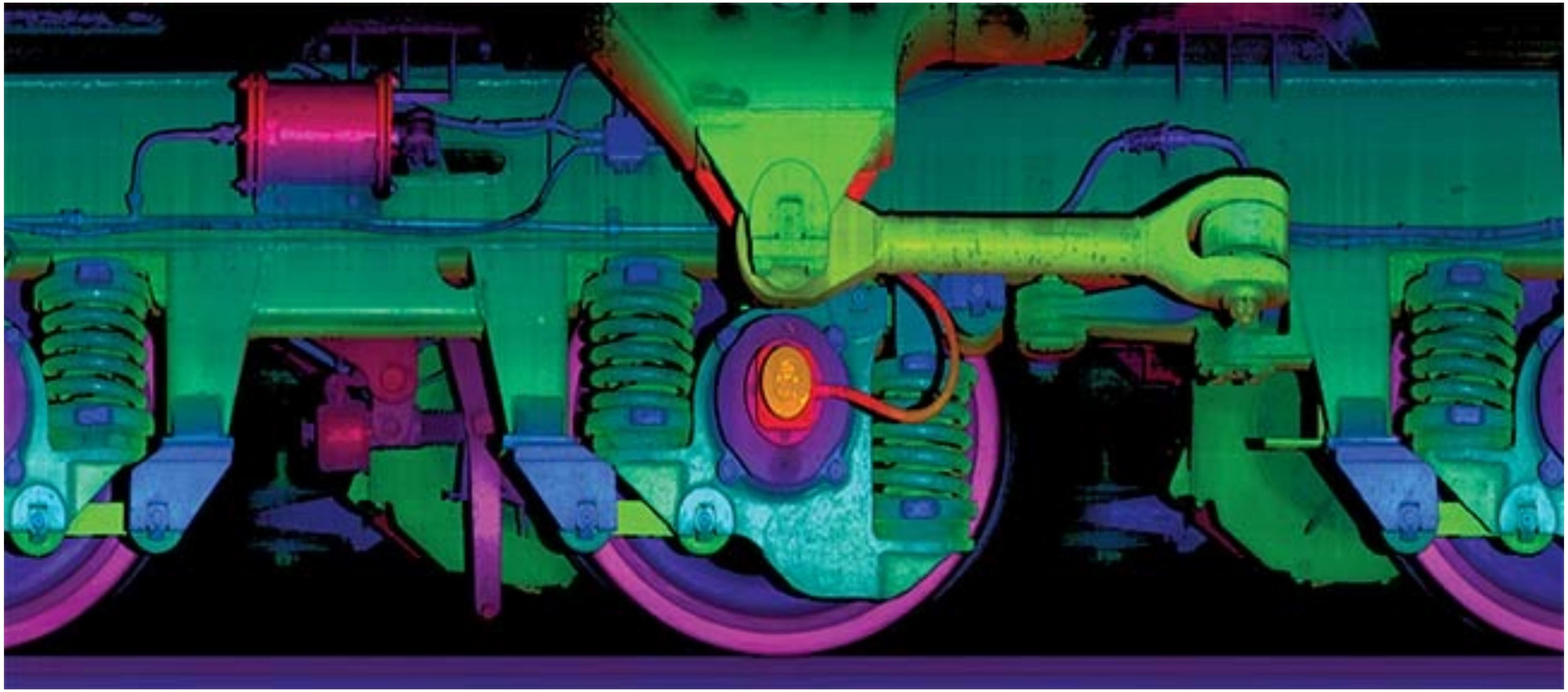
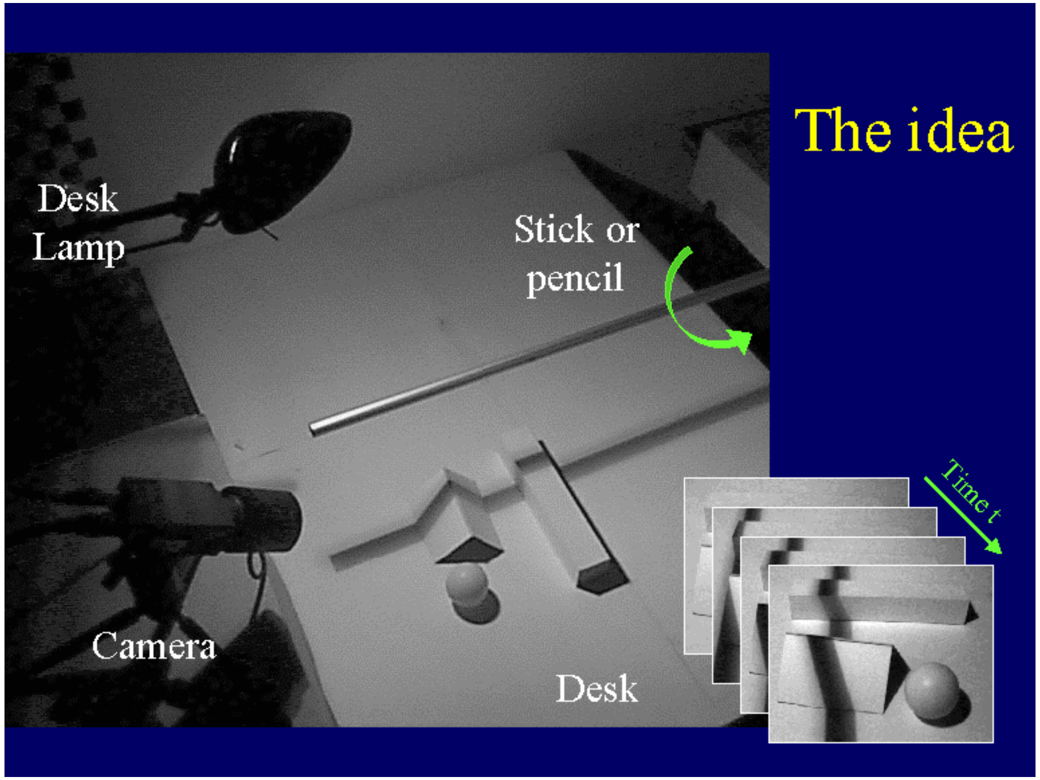
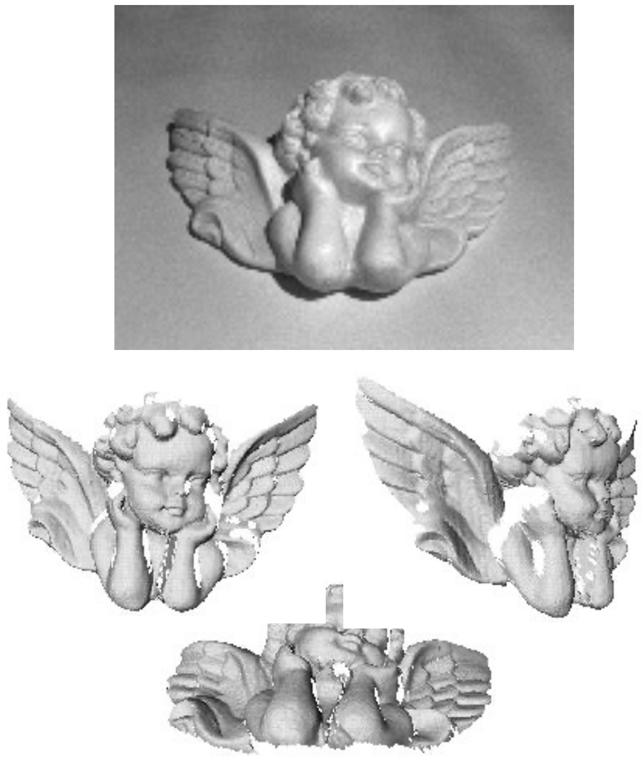


Image from Mattias Johannesson, Laser Triangulation Tackles Imaging Tasks Big and Small, https://www.photonics.com/Articles/Laser_Triangulation_Tackles_Imaging_Tasks_Big_and/a64616 (2019)

3D Photography on your Desk





http://www.vision.caltech.edu/bouguetj/ICCV98/ (link is dead, Bouguet is at Cruise now)