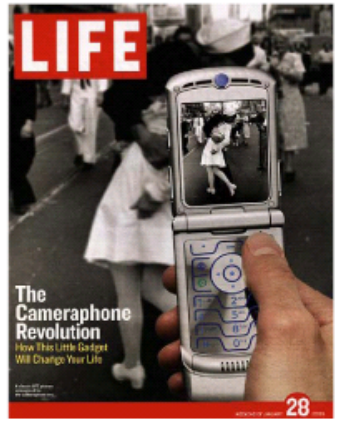
Sensing 4:
Vision

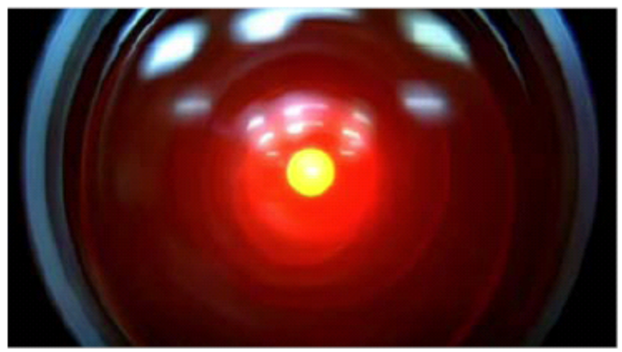
Many slides adapted from slides © R. Siegwart, Steve Seitz

Bookkeeping

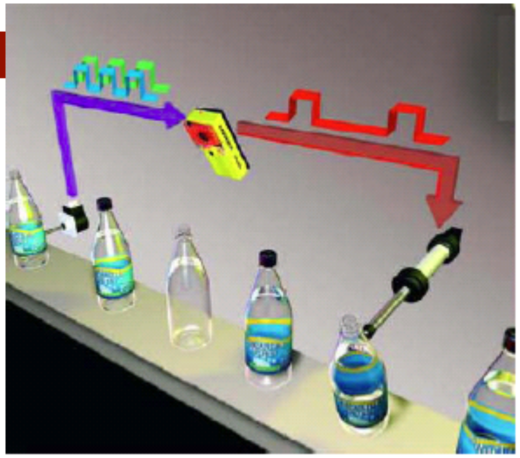
- Assignment 1
  - Blackboard is open, assignment is up
- Projects
  - First deliverable on ~~Tuesday~~ Thursday (Wednesday night)
  - Example posted
  - Pay special attention to goals and milestones
- Reading: SNS – intro to 4.3; 4.4; 4.5-4.5.2; 4.5.5
  - pg.: 195 • 208-212 • 212-215 • 227-234
- Today
  - Cameras, stereo vision, time of flight, structured light

Computer Vision



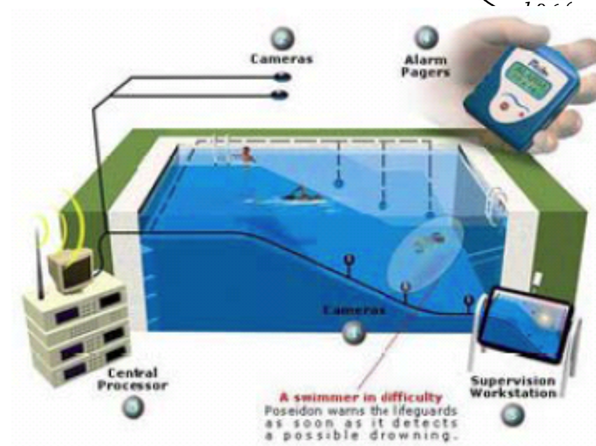


Applications of Computer Vision

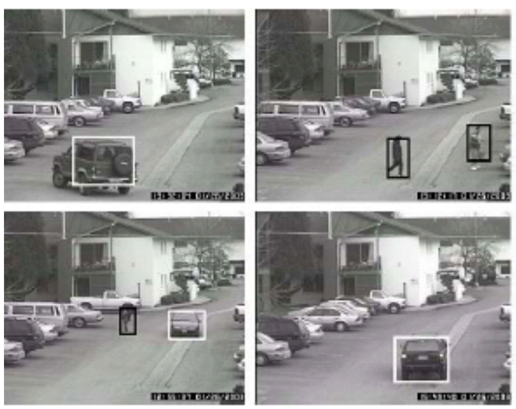


Factory inspection

Reading license plates, checks, ZIP codes



Monitoring for safety (Poseidon)



Surveillance

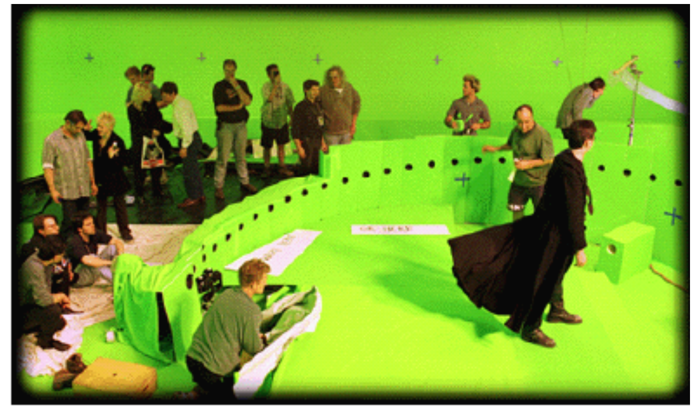
Autonomous driving, robot navigation

Driver assistance (collision warning, lane departure warning, rear object detection)

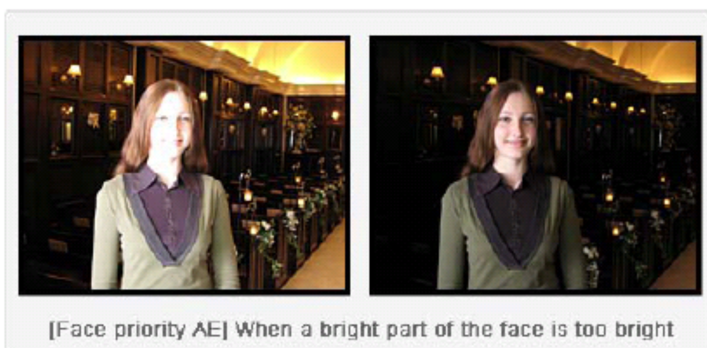
Applications of Computer Vision

Assistive technologies

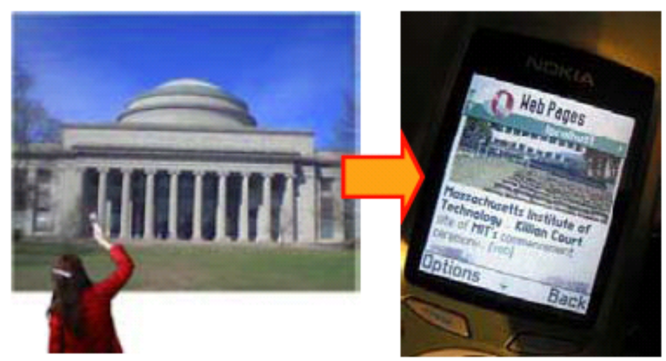
Entertainment (Sony EyeToy)



Movie special effects

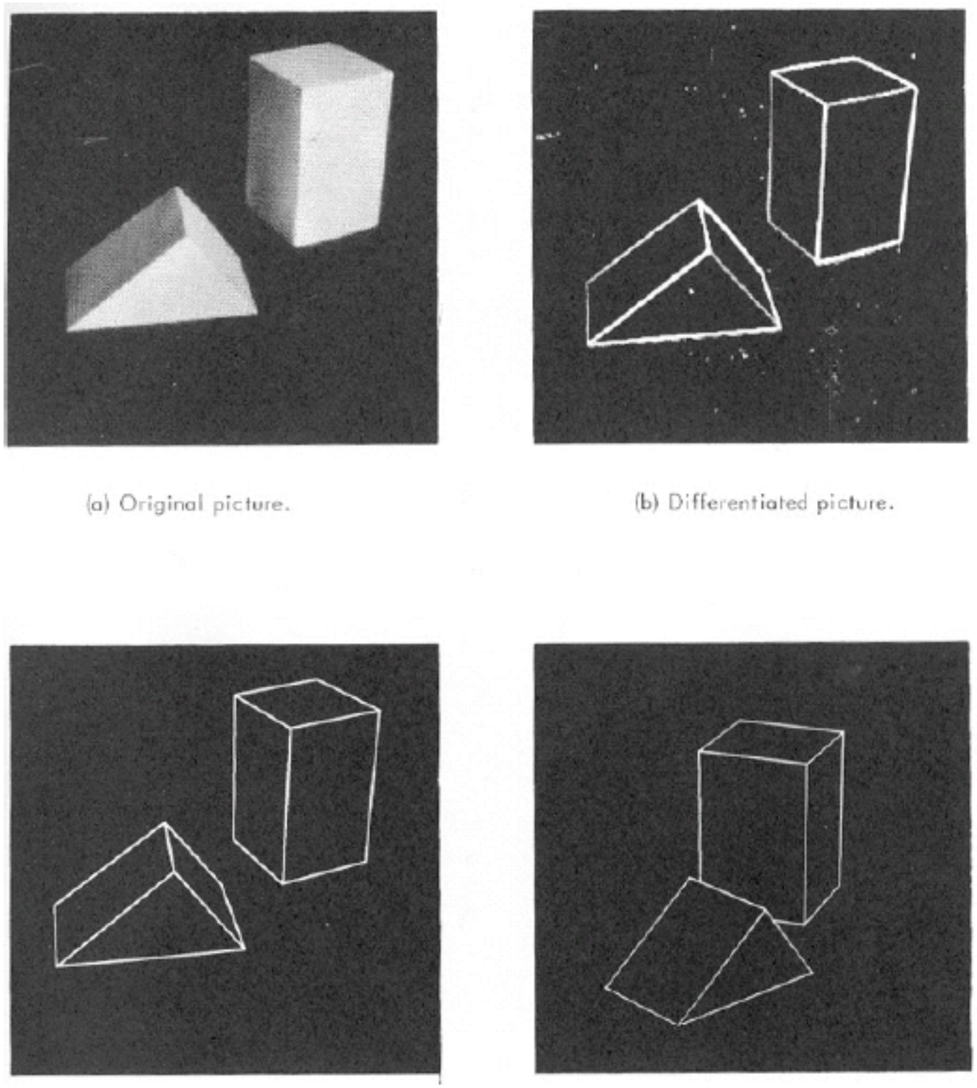


Digital cameras (face detection for setting focus, exposure)



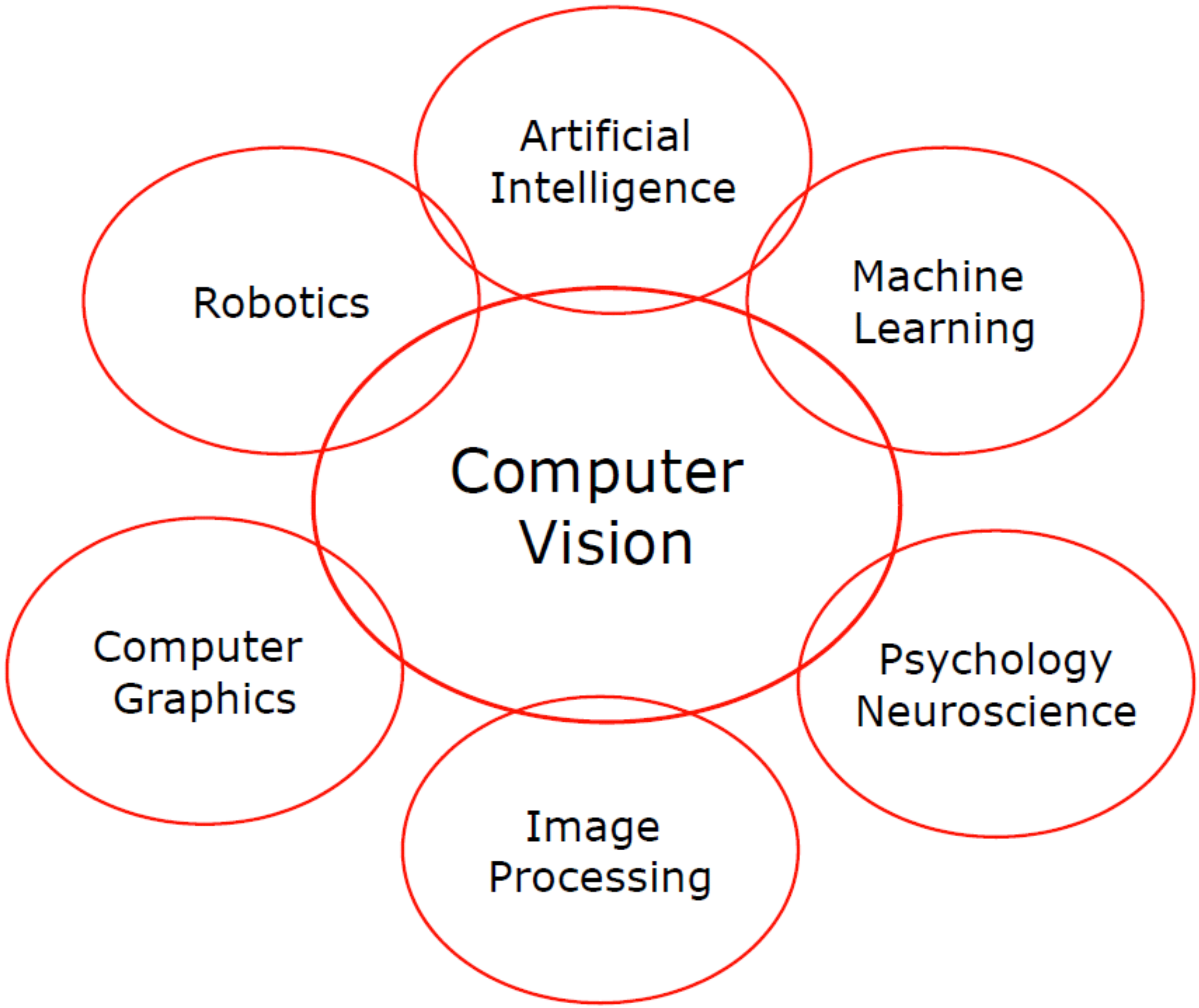
Visual search http://www.kooaba.com/

Origins of Computer Vision



L. G. Roberts, *Machine Perception of Three Dimensional Solids,* Ph.D. thesis, MIT Department of Electrical Engineering, 1963.

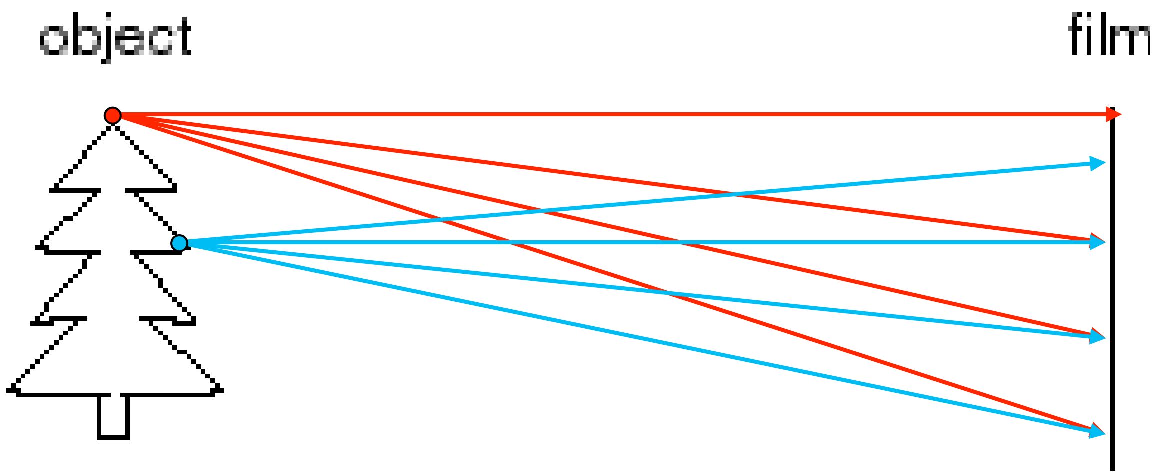
Disciplines Using Vision



The Camera

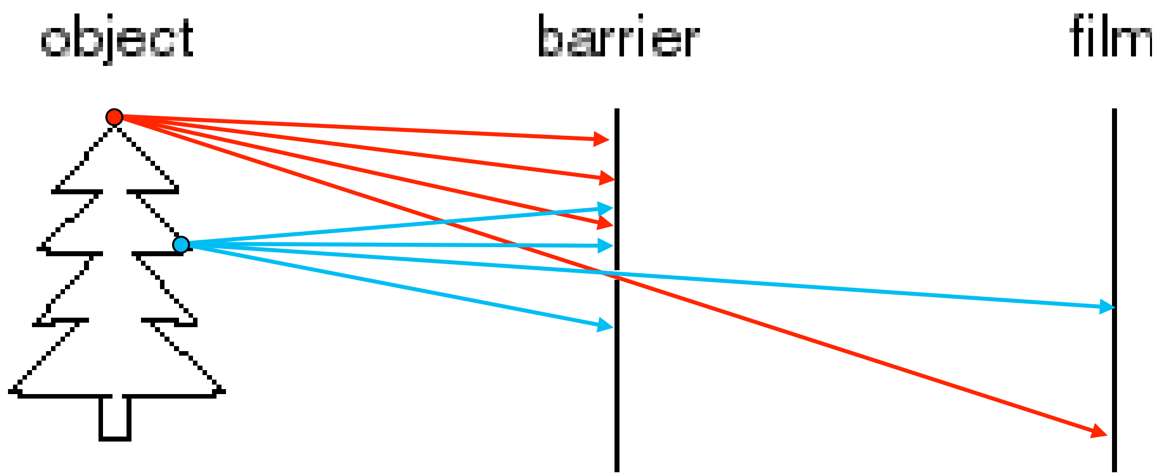
- Parameters
  - Light allowed in (aperture)
  - Shutter speed
  - Resolution
  - Gain/Saturation
  - Focus and focal depth
- Failure modes
  - Blue-to-red sensitivity
  - Cross-sensitivity
  - Dynamic range

How do we see the world?



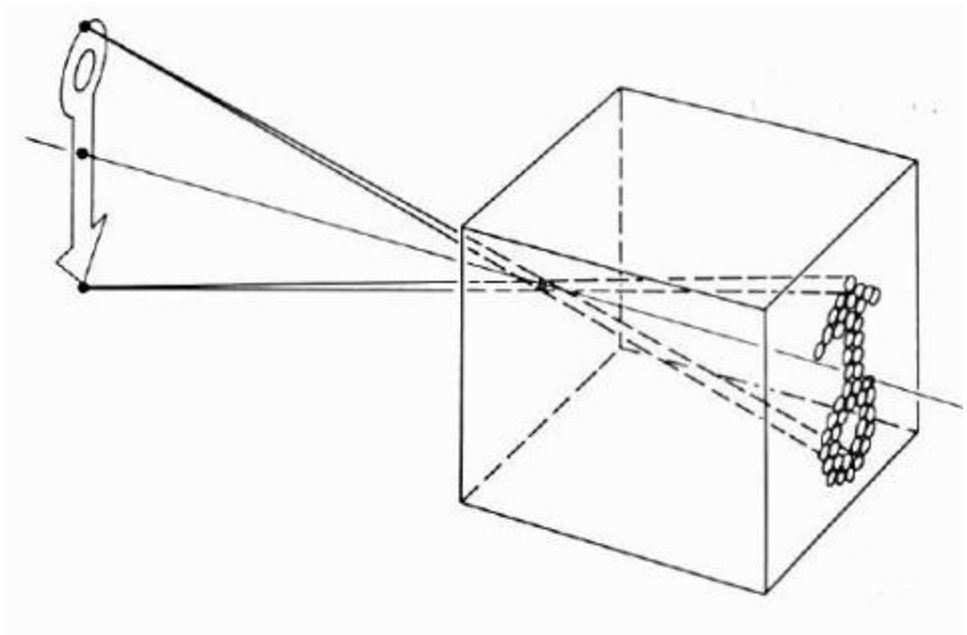
- Designing a camera
- Idea 1: put a piece of film in front of an object
- Do we get a reasonable image?

Pinhole camera



- Add a barrier to block off most of the rays
- This reduces blurring
- The opening is known as the aperture

Pinhole Camera Model



- Pinhole model:
  - Captures pencil of rays – all rays through a single point
  - The point is called Center of Projection
  - The image is formed on the Image Plane

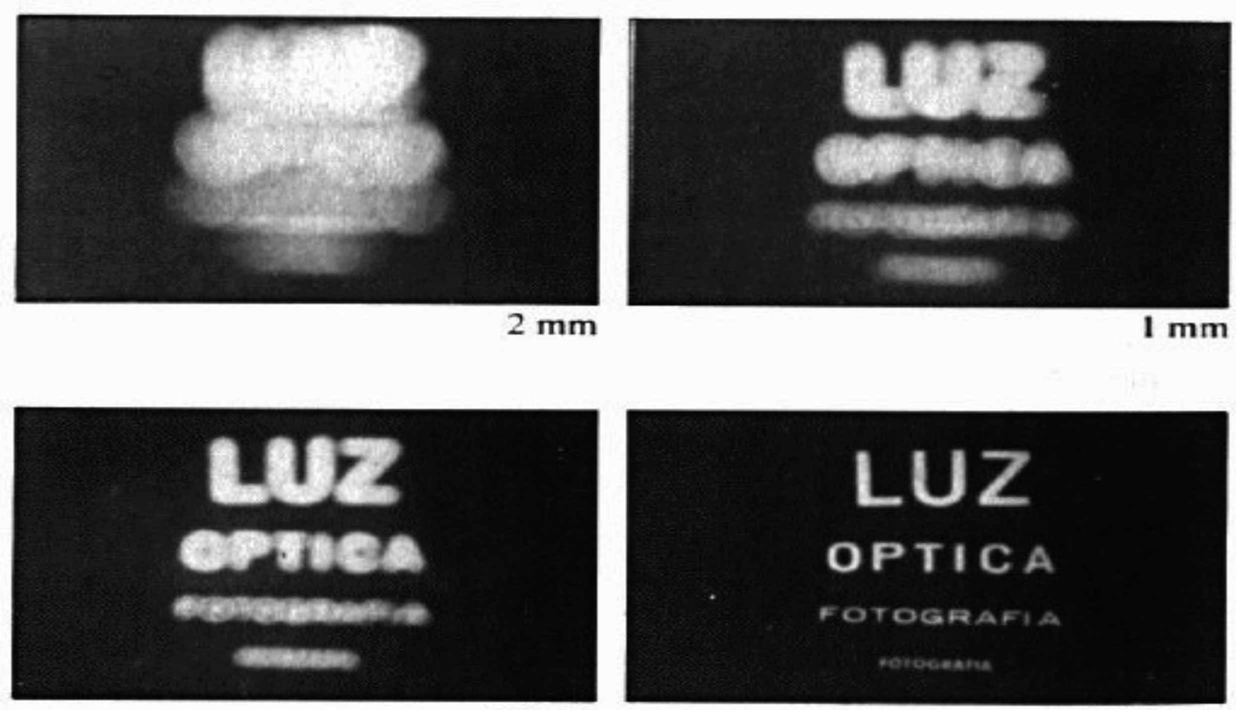
Slide by Steve Seitz

Home-made pinhole camera

Why so blurry?

http://www.debevec.org/Pinhole/

Shrinking the aperture

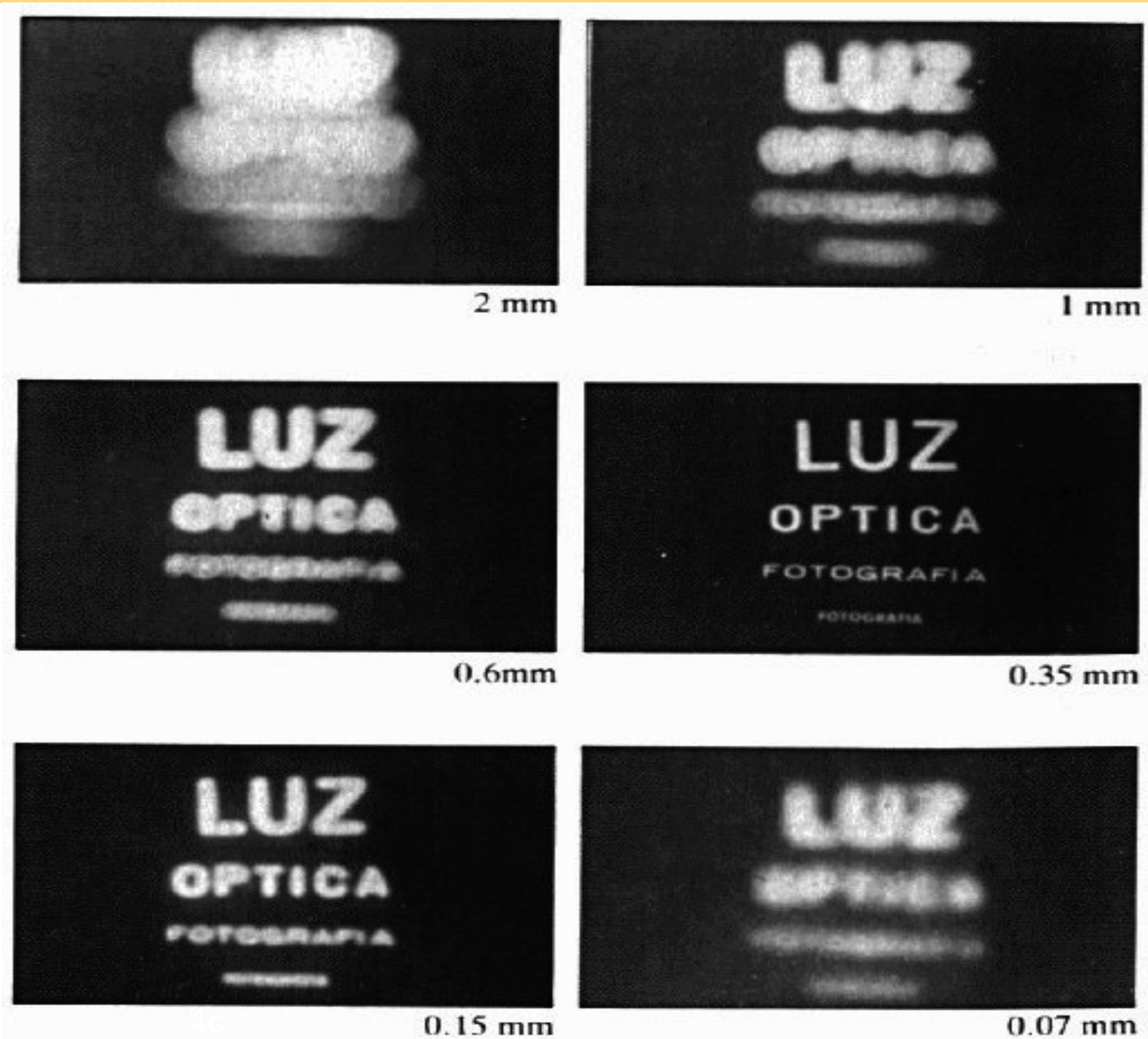


2 mm 1 mm 0.6mm 0.35 mm

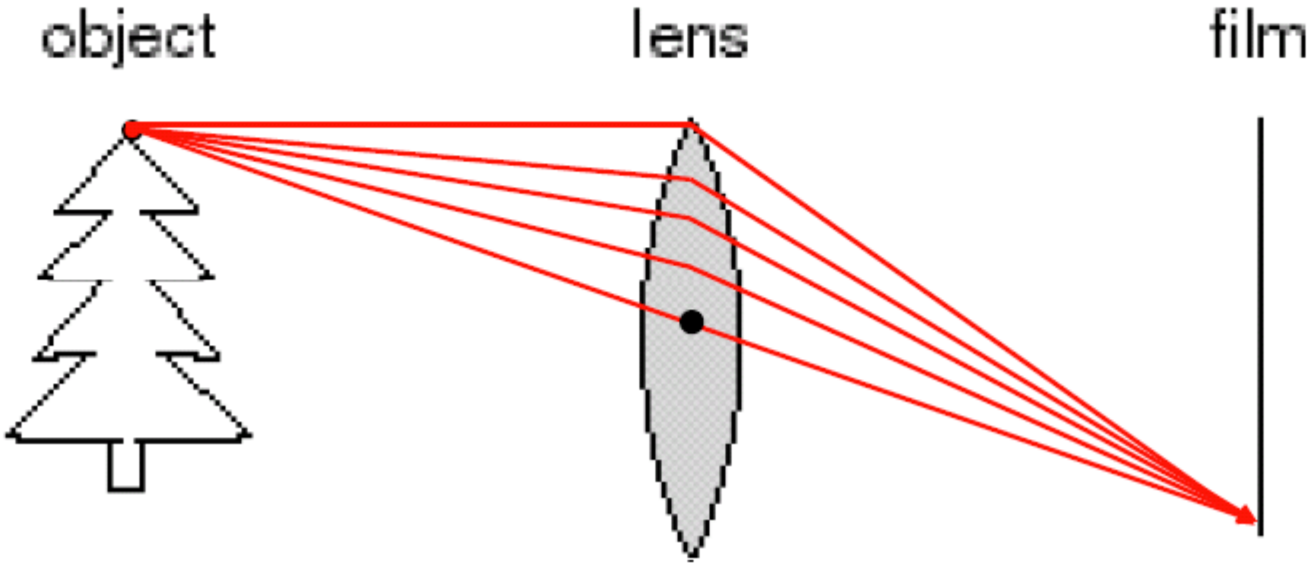
- Why not make the aperture as small as possible?
  - Less light gets through (must increase the exposure)
  - Diffraction effects…

Slide by Steve Seitz

Shrinking the aperture

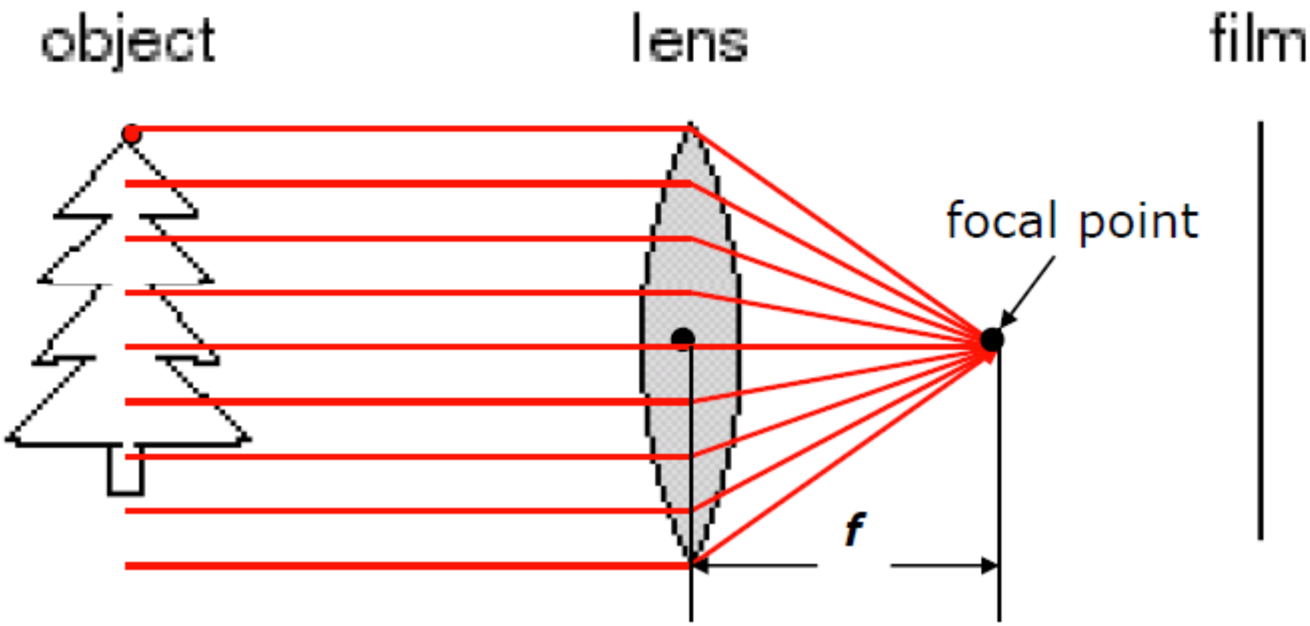


Solution: adding a lens



- A lens focuses light onto the film
  - Rays passing through the center are not deviated

Solution: adding a lens



- A lens focuses light onto the film
  - Rays passing through the center are not deviated
  - All parallel rays converge to one point on a plane located at the *focal length f*

Range (Distance) Sensors

- Range sensors – how far is robot from something?
  - Key element for localization and environment modeling
- Stereo
  - Humans; Bumblebee/Bb2
- Time-of-Flight
  - Laser
  - Sonar
  - Kinect 2
- Structured Light
  - Kinect

Active

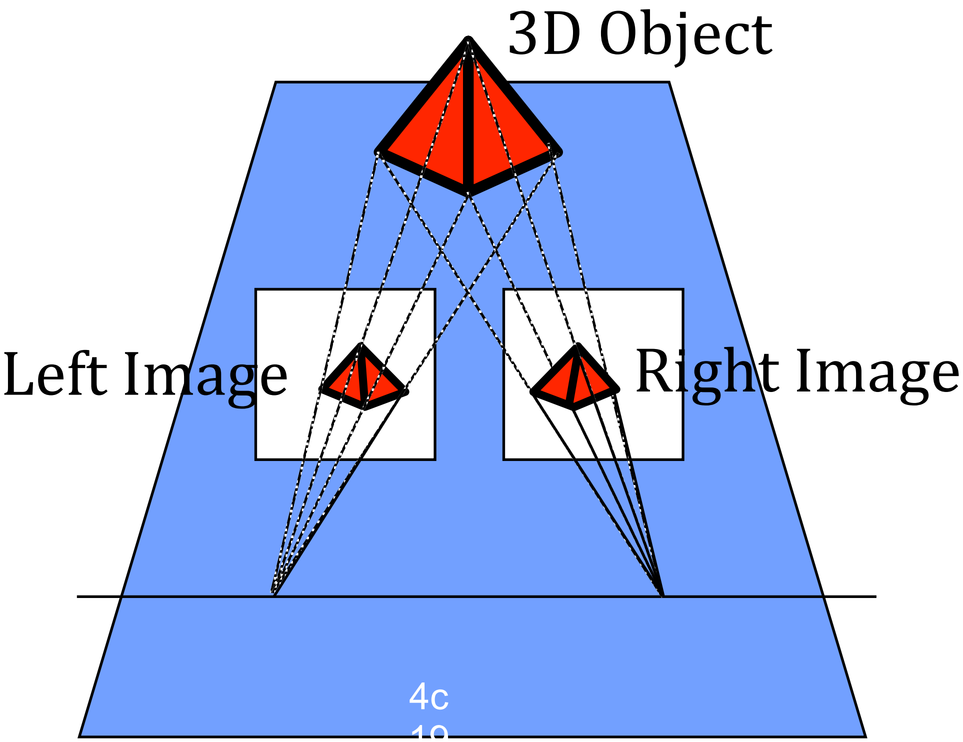
Distance Using Vision

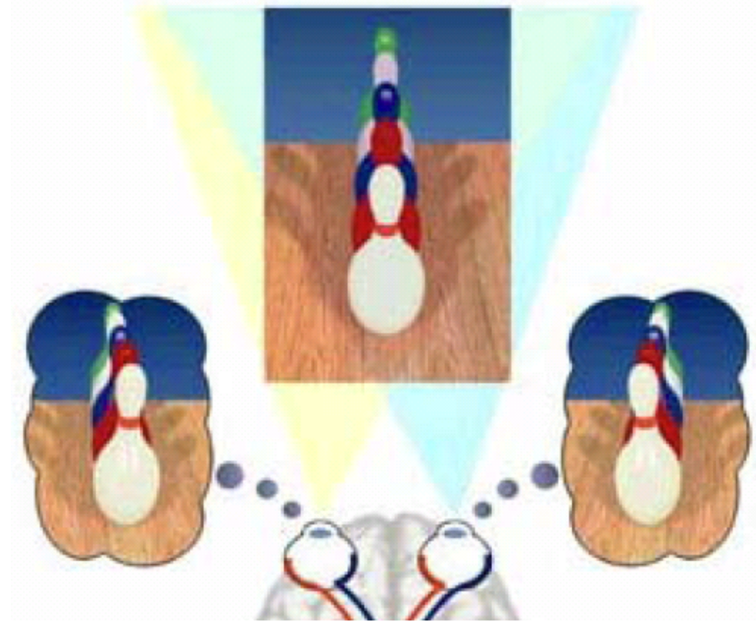
- Stereo Vision
  - Two sensors (cameras)
  - Known relative position and orientation


- Structure from motion:
  - Use a single moving camera
  - 3D structure and camera motion can be estimated

Stereo Vision

- Reconstruct a 3D scene
- Two images, two points of view





b = baseline: distance between optical centers of cameras
f = focal length
v-v' = disparity between views

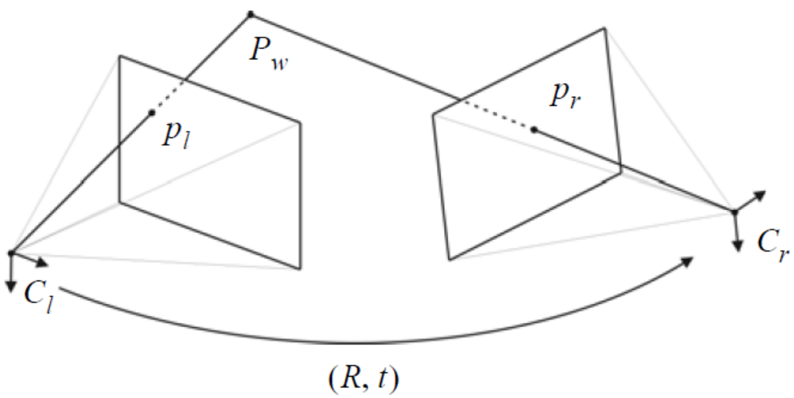
Stereo Vision Accuracy

b = baseline: distance between optical centers of cameras
f = focal length
v-v' = disparity between views

- Simplified: assume cameras are
  - Identical
  - Aligned on a horizontal axis
- Distance is *inversely* proportional to disparity
  - Closer objects can be measured more accurately
- Disparity is proportional to b
  - For a given disparity error, the accuracy of the depth estimate increases with increasing baseline b
  - However, as b is increased, some objects may appear in one camera, but not in the other
- Increasing image resolution improves accuracy

Calibration and Alignment

- Two identical cameras do not exist in nature
- Aligning cameras perfectly on a horizontal axis is hard



- Need to estimate relative pose between cameras
  - *Rotation* and *translation* – and since cameras are not identical, also
    - focal length, image center, radial distortion
- Epipolar rectification: compare two feature-rich images

Correspondence

- Two cameras see slightly different scenes
  - What points in one correspond to points in the other?
  - Compare all points in image to all points in other image
  - This image search can be computationally expensive, imperfect

Output: Disparity Map

- Find the correspondent points of all image pixels of the original images
- For each pair of conjugate points compute the disparity $d = v - v'$
- Output: disparity map

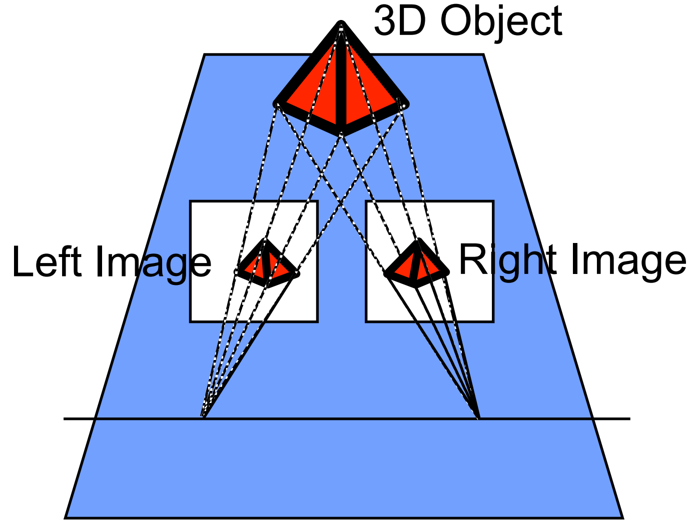
Left image

Right image

Disparity map

Disparity maps are usually visualized as grey-scale images. Objects that are closer to the camera appear lighter, those who are further appear darker.

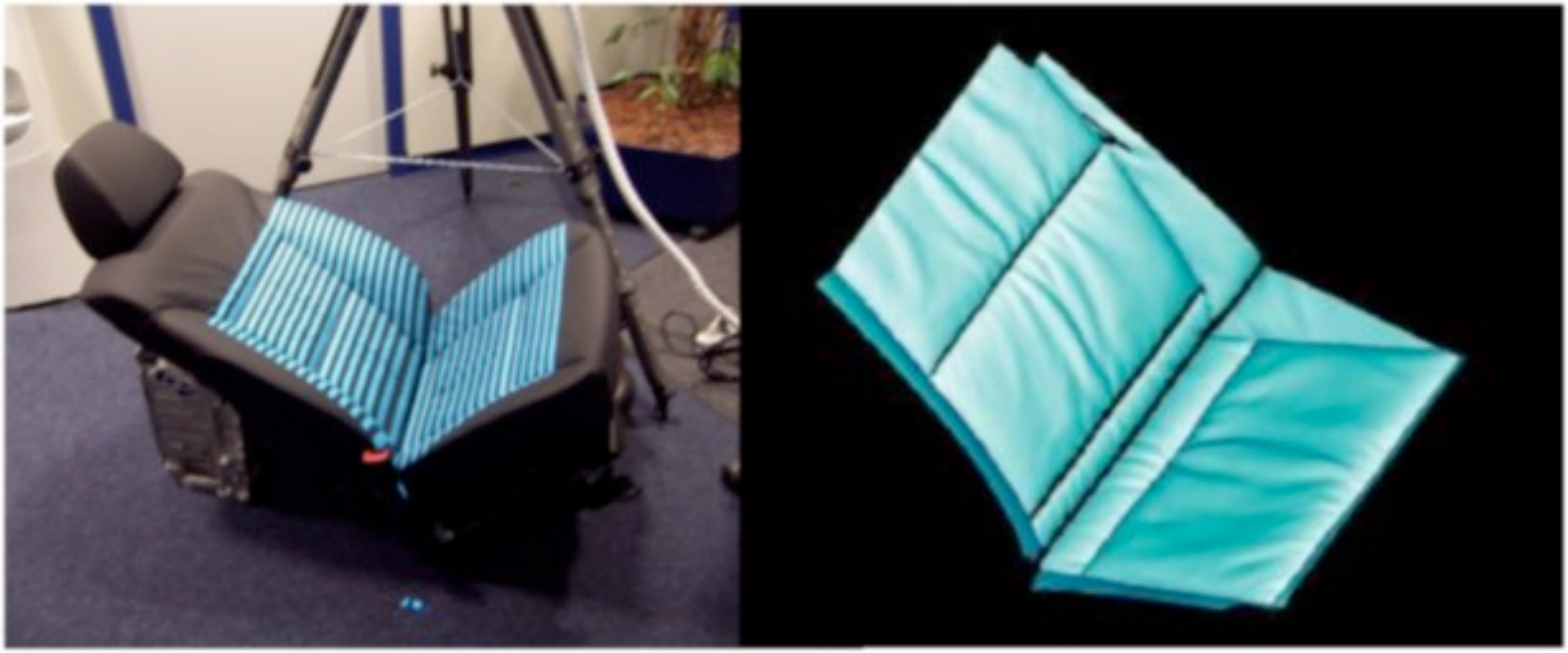
Summary



1. Stereo camera calibration → compute camera relative pose
   - Epipolar rectification → align images
2. Search correspondences
3. Output: compute stereo triangulation or disparity map
4. Consider baseline and image resolution to compute accuracy!

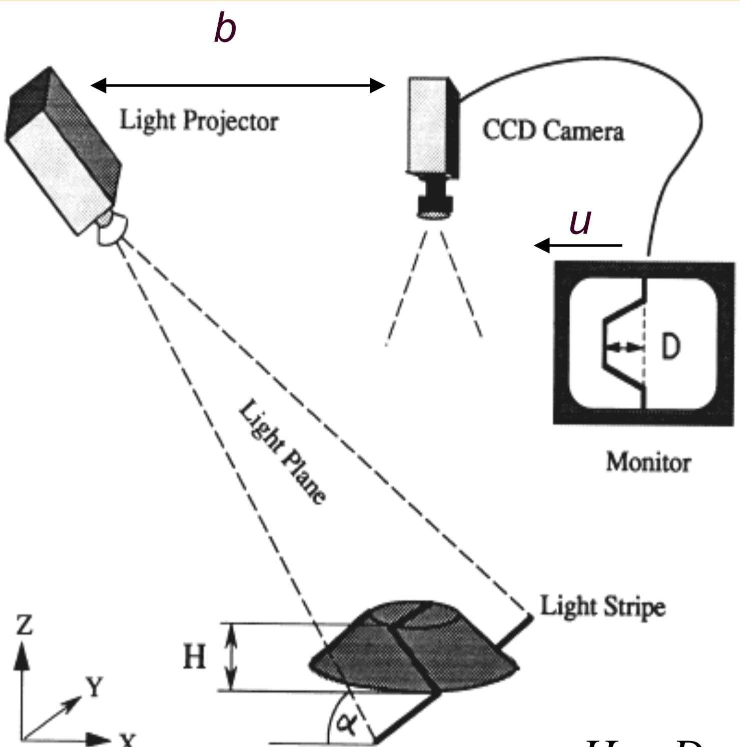
Structured Light

- What if you know what the light should look like?



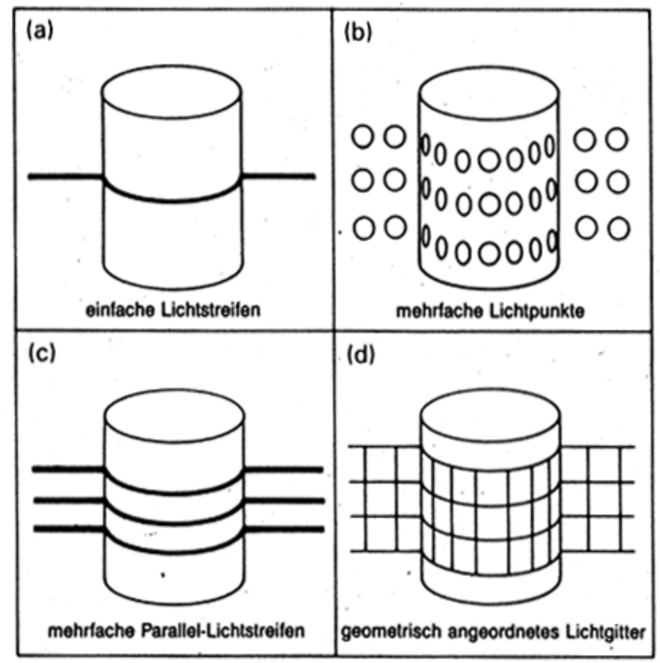
Structured Light

a



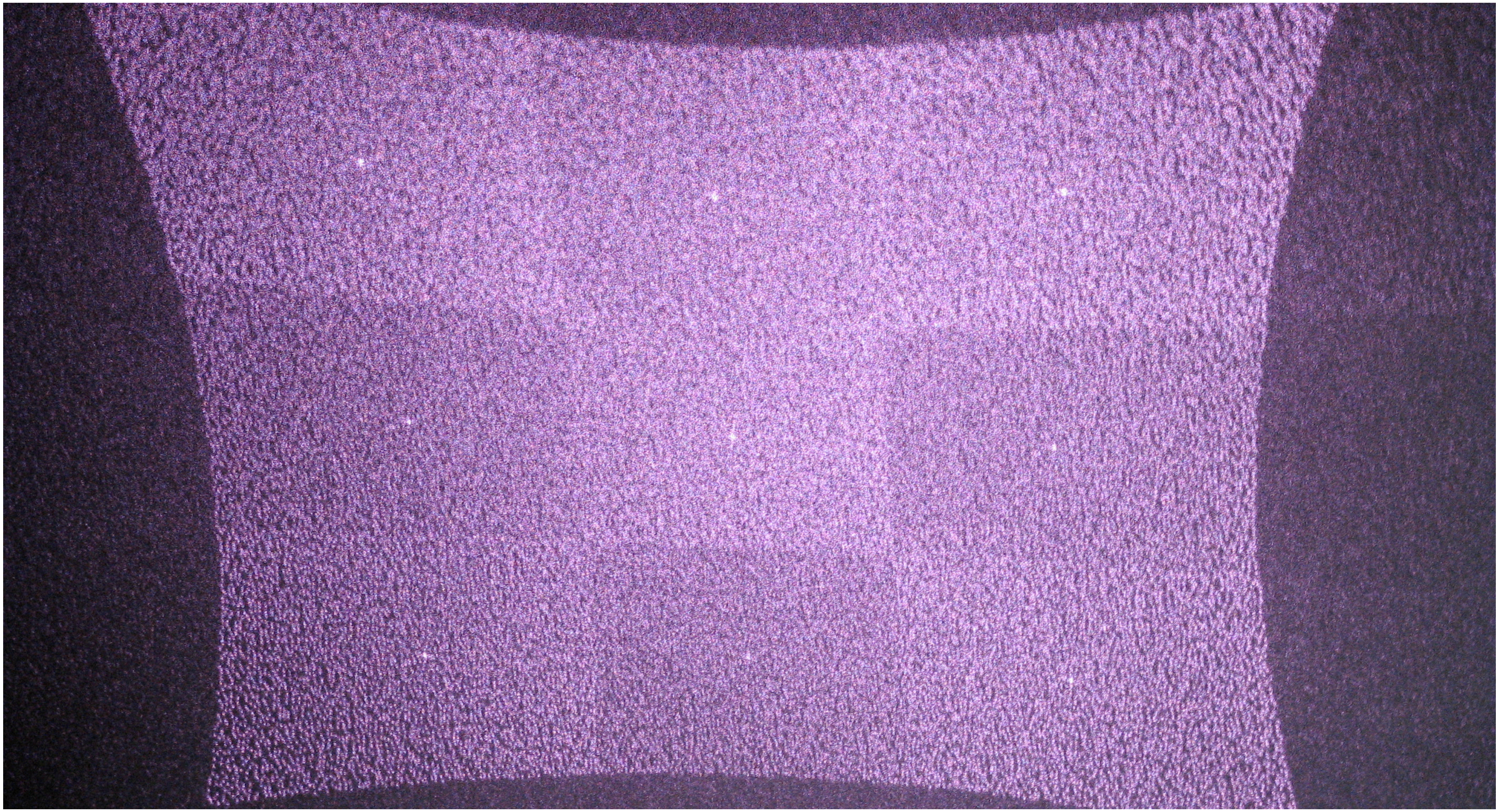
$$H = D \cdot \tan\alpha$$

b



- Eliminate correspondence problem by projecting known light on the scene
- Light perceived by camera
- Range to an illuminated point can then be determined from geometry

Microsoft Kinect

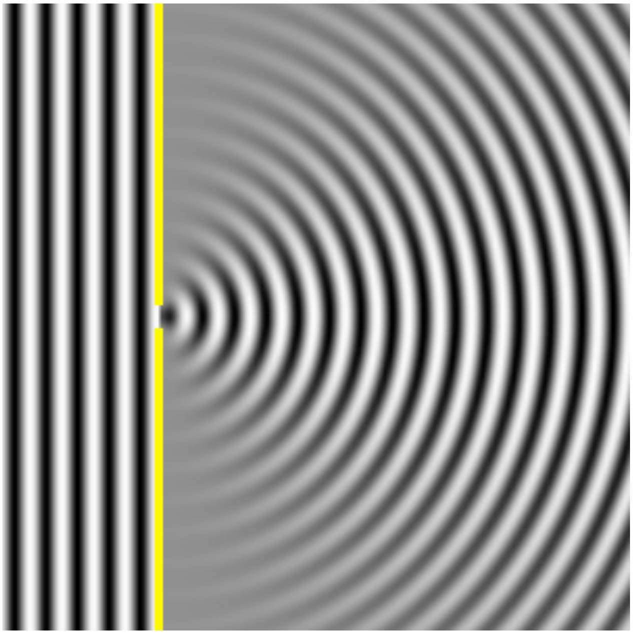


Range: Time-of-Flight

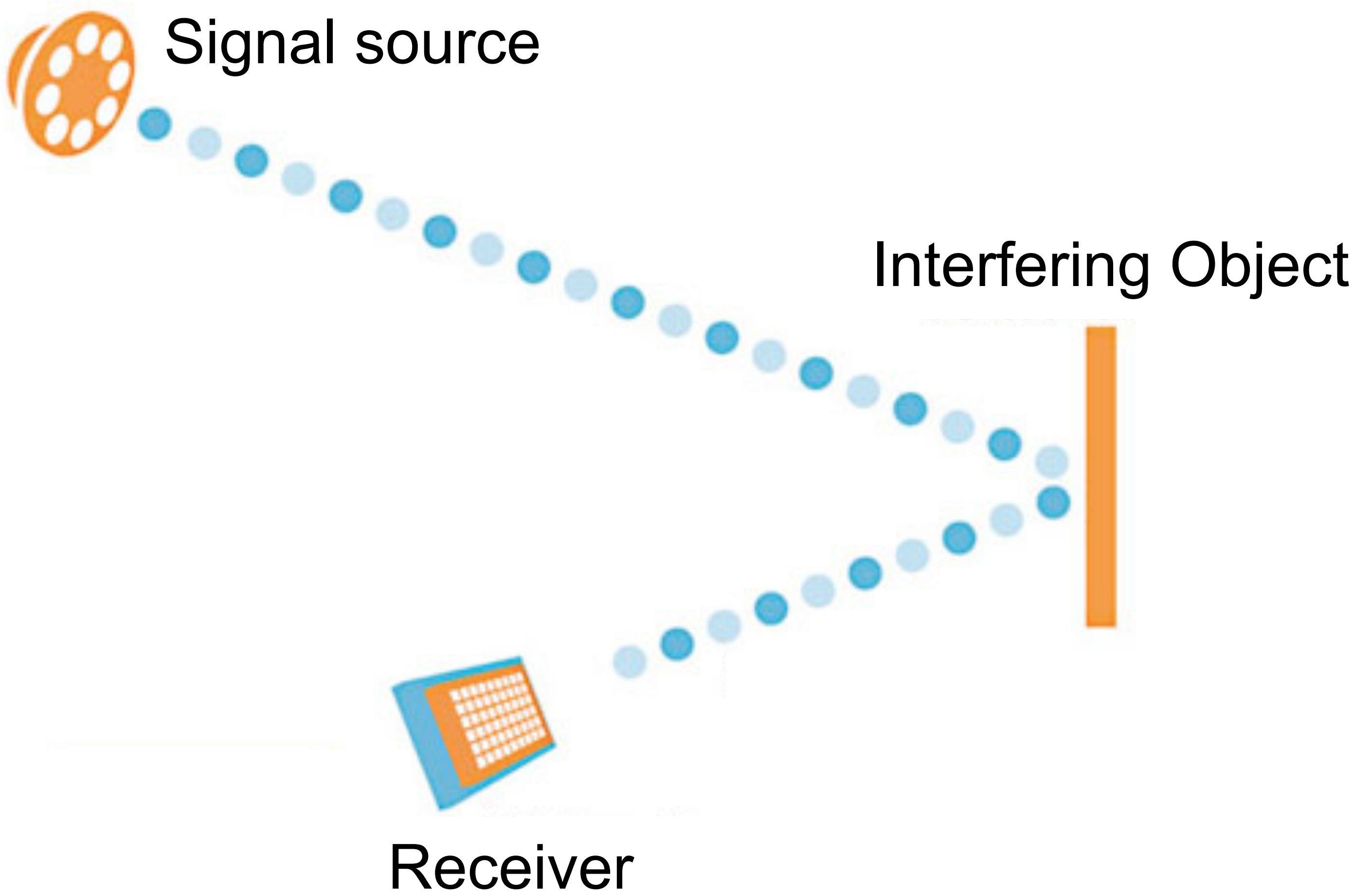
- Time-of-flight uses *propagation speed* of waves
  - Sound or electromagnetic
- Distance traveled by a wave is:

$$d = c \cdot t$$

d = distance traveled (round-trip)
c = speed of wave propagation
t = time of flight



Time-of-Flight



Time-of-Flight: Accuracy

- Sources of inaccuracy:
  - Uncertainties about exact time of arrival of the reflected signal
  - Inaccuracies in the time of flight measure (laser range sensors)
  - Opening angle of transmitted beam (ultrasonic range sensors)
  - Interaction with the target (surface, specular reflections)
- Variation of propagation speed
  - Propagation speed of sound: 0.3 m/ms
  - Propagation speed of of electromagnetic signals: 0.3 m/ns
    - One million times faster.
    - Laser range sensors expensive and delicate.
- Speed of mobile robot and target

Scanning Range Sensing

- Confidence in the range (phase estimate) is inversely proportional to the square of the received signal amplitude.

- Dark or distant objects → worse estimates than closer brighter objects

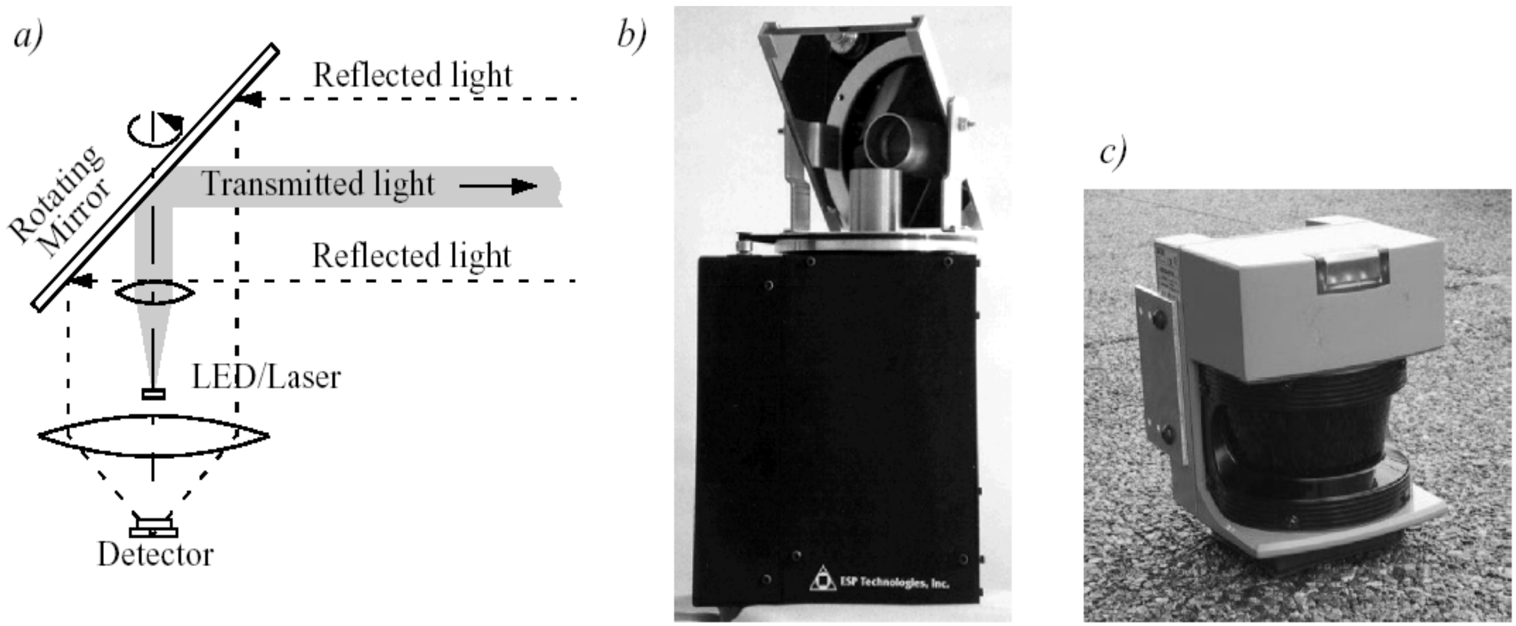
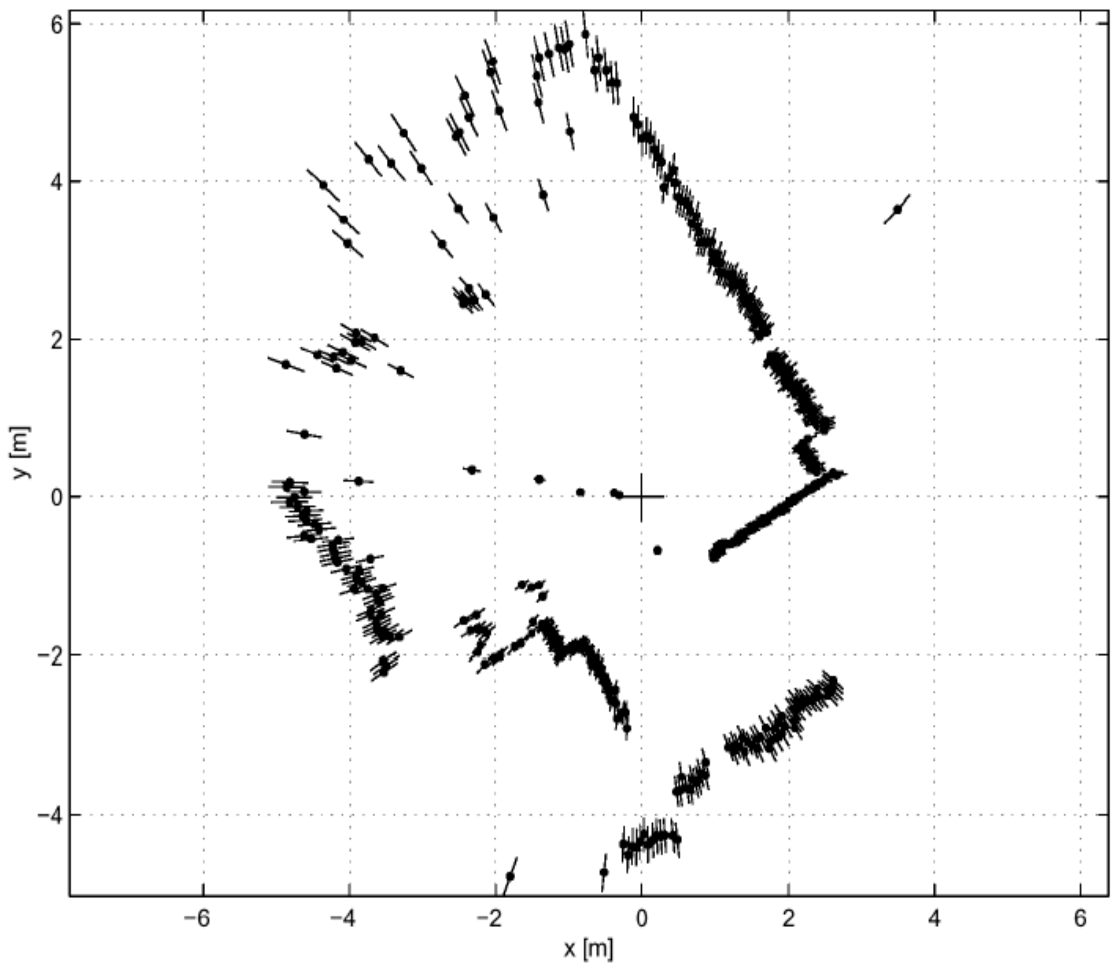


Figure 4.11
(a) Schematic drawing of laser range sensor with rotating mirror; (b) Scanning range sensor from EPS Technologies Inc.; (c) Industrial 180 degree laser range sensor from Sick Inc., Germany

Example of Scanning

- Length of the lines through measurement points indicate uncertainty



Modern Time-of-Flight