Sensing 4: Vision

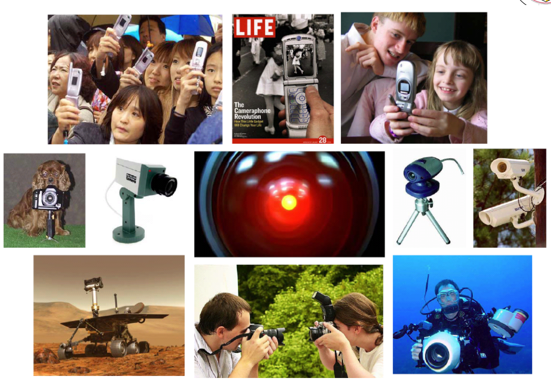

Many slides adapted from slides © R. Siegwart, Steve Seitz, J. Tim Oates

1

UMBC Dining Services · retriever ESSENTIALS · UMBC

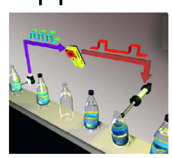
2

Computer Vision

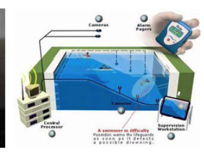


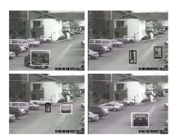
3

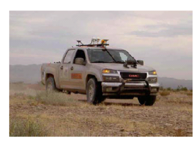
Applications of Computer Vision


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)

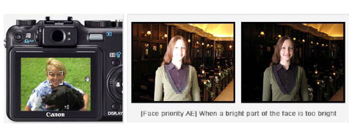
4

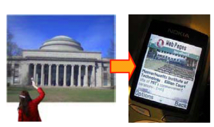
Applications of Computer Vision


Assistive technologies


Entertainment Microsoft Kinect

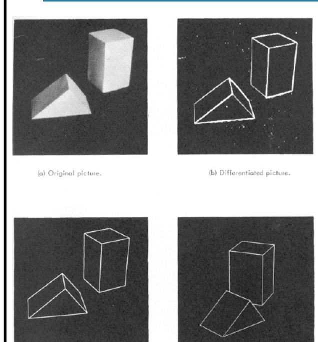

Movie special effects


Digital cameras (face detection for setting focus, exposure)


Visual search

5

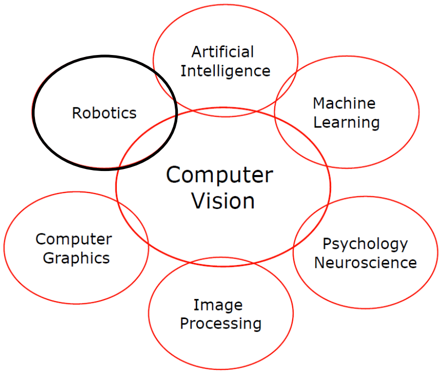
Origins of Computer Vision



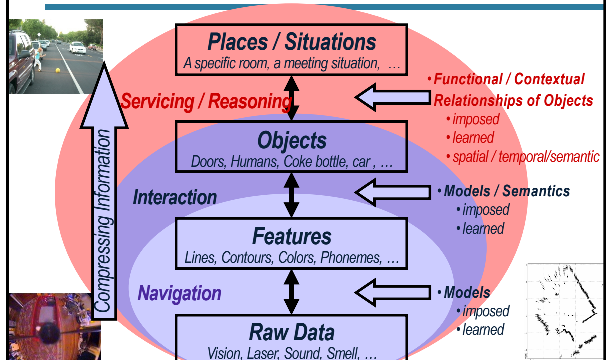
L. G. Roberts, *Machine Perception of Three Dimensional Solids,* Ph.D. thesis, MIT Department of Electrical Engineering, 1963.

6

Disciplines Using Vision



Perception for Mobile Robots



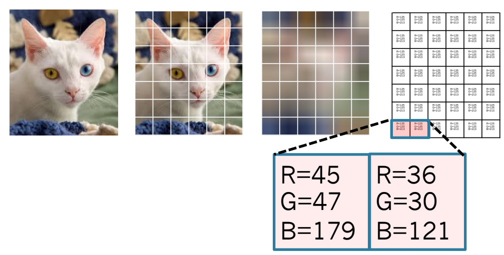
Footery stuffs

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

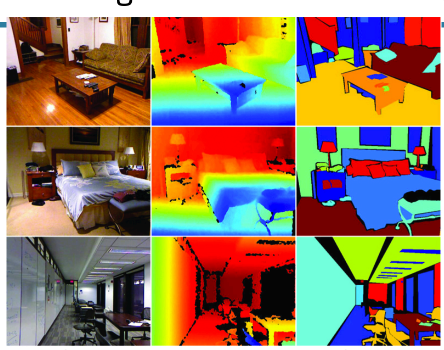
Images

- RGB image: outputs a matrix of RGB values



N. Silberman, P. Kohli, D. Hoiem, R. Fergus. Indoor Segmentation and Support Inference from RGBD Images. ECCV 2012.

Depth Images

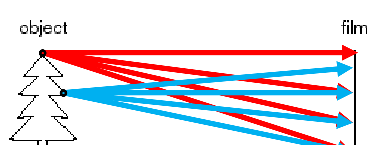


RGB Depth Objects

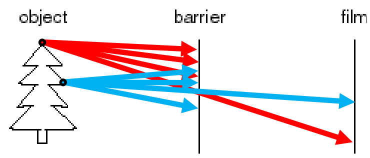
N. Silberman, P. Kohli, D. Hoiem, R. Fergus. Indoor Segmentation and Support Inference from RGBD Images. ECCV 2012.

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



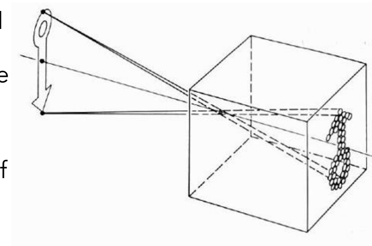
- Add a barrier to block off most of the rays
- This reduces blurring
- The opening is known as the aperture

13

13

Pinhole Camera 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 @ UW 14

14

Home-made pinhole camera

Why so blurry?

Exposure needs light

Smaller hole = darker

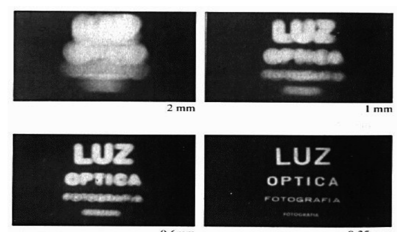
Larger hole = blur

http://www.debevec.org/Pinhole 15

15

Shrinking the aperture

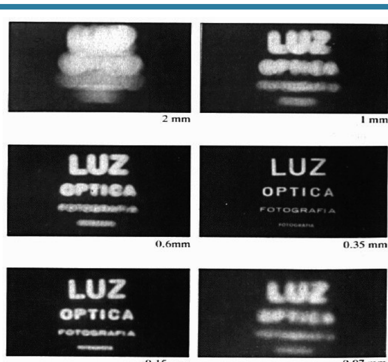
- Why not make the aperture smaller?
  - Less light gets through (must increase the exposure)
  - Diffraction effects…



Slide by Steve Seitz @ UW 16

16

Shrinking the aperture

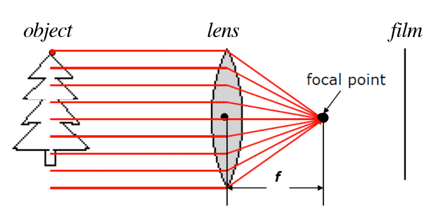


Slide by Steve Seitz @ UW 17

17

Solution: adding a lens

- A lens focuses light onto the film

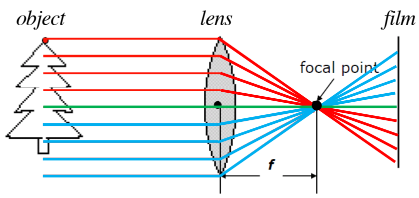


18

18

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*

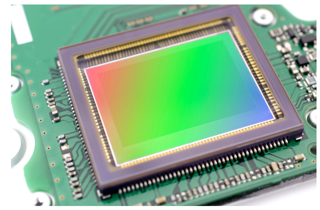


19

19

Digitally: CCDs

- A charge coupled device, or CCD
- Light-sensitive elements called pixels etched on silicom
- Photons hitting this surface generate an *analog* per-element charge that can be read



www.visiononline.org/blog-article.cfm/CCD-vs-CMOS-Image-Sensors-Which-are-Better/82

20

20

Range (Distance) Sensors

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

(Time-of-Flight and Structured Light: Active)

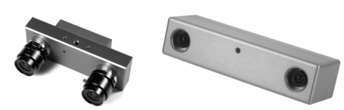
Broadly speaking, *depth cameras* report per-pixel *depth*—how far is something from that single pixel?—as well as (sometimes) RGB

21

21

Distance Using Vision

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



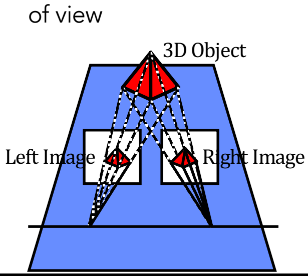
- Structure from motion:
  - Use a single moving camera
  - 3D structure and camera motion can be estimated

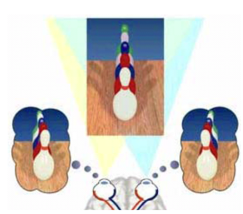
22

22

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

23

23

Stereo Vision Accuracy

- Simplified: assume cameras are
  - Identical
  - Aligned on a horizontal axis
- Distance is inversely proportional to disparity
- 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

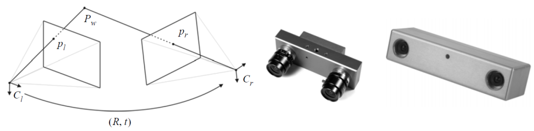
b = baseline: distance between optical centers of cameras
f = focal length
v-v' = disparity between views

24

24

Calibration and Alignment

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

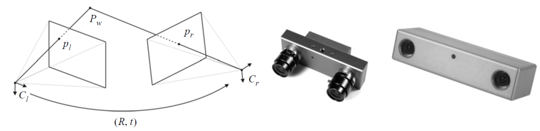


25

25

Calibration and Alignment

- 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



26

26

Correspondence

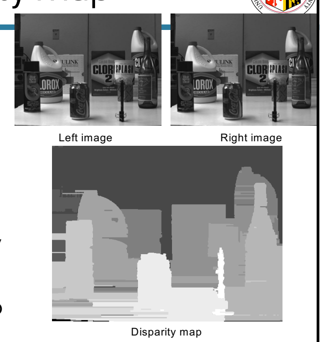
- Two cameras see slightly different scenes
  - What points in one correspond to points in the other?
  - Compare all points in image to all points in other image
  - This image search can be slow, imperfect

27

27

Output: Disparity Map

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

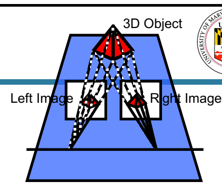


Distance maps, visualized as grey-scale images. Objects closer to the camera appear lighter.

28

28

Summary



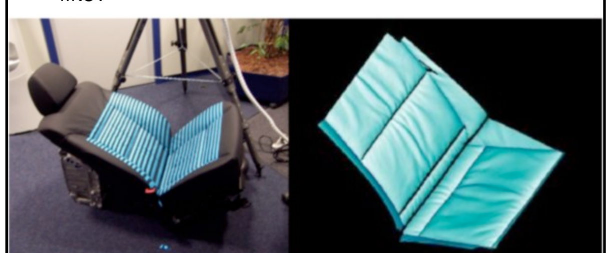
- Stereo camera calibration: compute camera relative pose
  - Epipolar rectification → align images
- Search correspondences
- Output: compute stereo triangulation or disparity map
- Compare to baseline and consider image resolution to compute accuracy

30

30

Structured Light

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

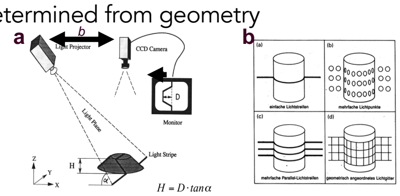


31

31

Structured Light

- 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

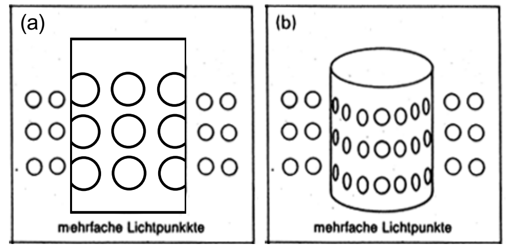


32

32

Structured Light

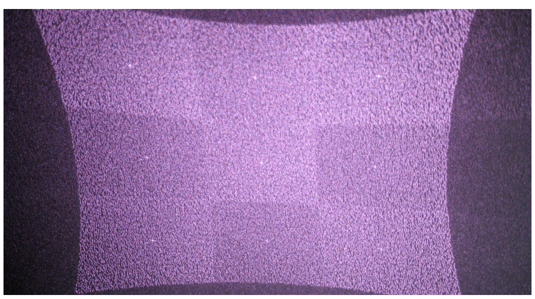
- Light is distorted by object it is falling on
- Two kinds of distortion: size and shape



33

33

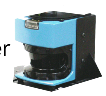
Microsoft Kinect

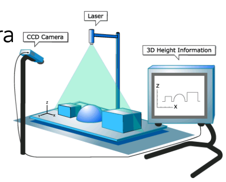


34

34

Range sensors

- Sonar 
- Laser range finder 
- Time of Flight Camera
- Structured light

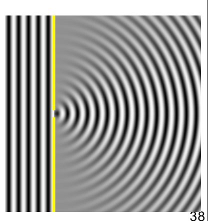


37

37

Range: Time-of-Flight

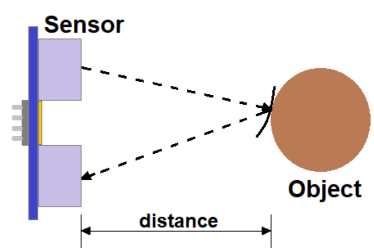
- Time-of-flight uses propagation speed of waves
  - Sound or electromagnetic
- Distance traveled by a wave is:
- d = c • t
  - d = distance traveled (round-trip)
  - c = speed of wave propagation
  - t = time of flight



38

38

Time-of-Flight



circuitdigest.com/microcontroller-projects/arduino-vl6180-tof-range-finder-sensor-for-distance-measurement

39

39

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)

41

41

Time-of-Flight: Accuracy

- 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

42

42

Scanning Range Sensing

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

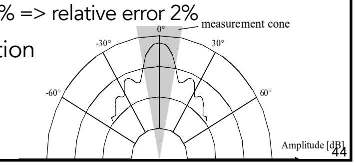
Figure 4.11
(a) Schematic drawing of laser range sensor with rotating mirror; (b) Scanning range sensor from EPS Technologies Inc.; (c) Industrial 180 degree laser range sensor from Sick Inc., Germany

43

43

Sonar

- Typical frequency: 40kHz - 180 kHz
  - Lower frequencies correspond to longer range
  - Sound from piezo transducer
  - Transmitter and receiver separated or not separated
- Range between 12 cm up to 5 m
  - (Ideal) resolution of ~2 cm
  - (Ideal) Accuracy 98% => relative error 2%
- ~Conical propagation
- Typical intensity distribution:



44

44

Sonar: Speed

- Transmit a packet of (ultrasonic) pressure waves
- Distance d of the echoing object can be found from propagation speed of sound c and the time of flight t. $d = \frac{c \cdot t}{2}$
- Speed of sound c (340 m/s) in air is: $c = \sqrt{\gamma \cdot R \cdot T}$
- Where
  - γ: adiabatic index (isentropic expansion factor)
  - R: gas constant
  - T: temperature in degree Kelvin

45

45

Sonar: Bandwidth

- An object that is 3 m away will take 20 ms, limiting its operating speed to 50 Hz.
- This update rate can affect maximum speed possible while still sensing and avoiding obstacles safely.
- But if the robot has a ring of 20 ultrasonic sensors, each firing sequentially and measuring to minimize interference between the sensors, then the ring's cycle time becomes 0.4 seconds => frequency of each one sensor = 2.5 Hz.

46

46

Laser Range Sensor

- Slightly deprecated term: LIDAR
- Similar to sonar
  - Without the signal speed issues
- More accessible, robust and cheaper than ever before


SICK

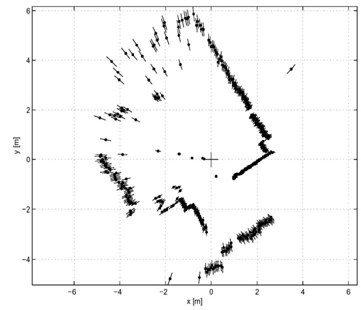

Hokuyo

48

48

Example of Scanning

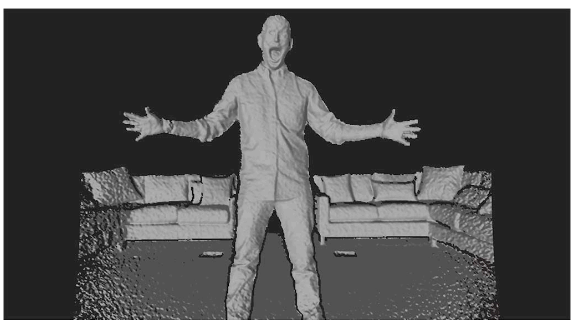
- Length of the lines through measurement points show uncertainty



54

54

Modern Time-of-Flight



60

60