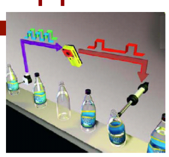
Sensing 4:
Vision

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

Computer Vision

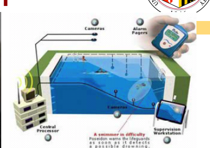
Applications of Computer Vision

3

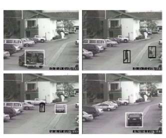


Factory inspection

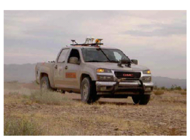
Reading license plates, checks, ZIP codes



Monitoring for safety (Poseidon)



Surveillance

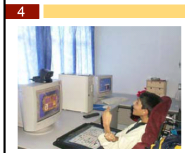


Autonomous driving, robot navigation

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

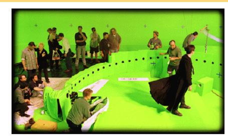
Applications of Computer Vision

4

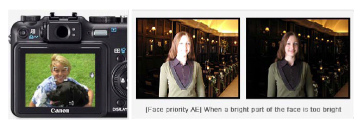


Assistive technologies

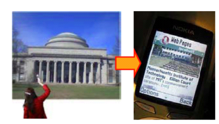
Entertainment Microsoft Kinect



Movie special effects



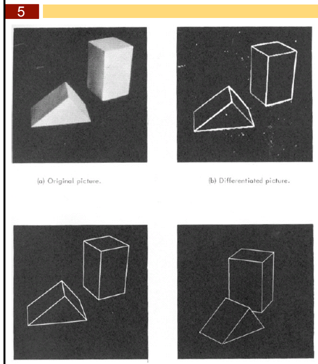
Digital cameras (face detection for setting focus, exposure)



Visual search

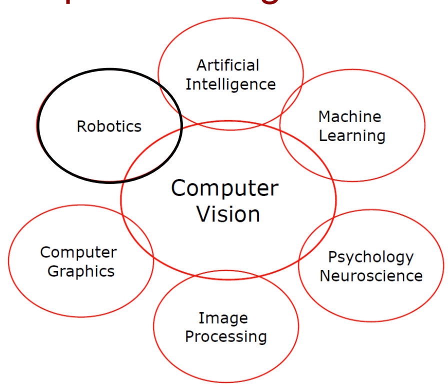
Origins of Computer Vision

5



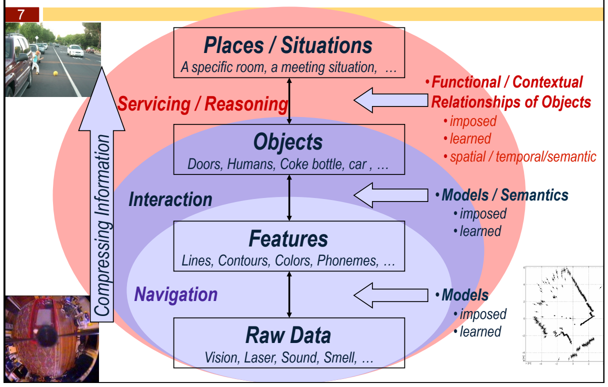
L. G. Roberts, *Machine Perception of Three Dimensional Solids*, Ph.D. thesis, MIT Department of Electrical Engineering, 1963.

Disciplines Using Vision



Perception for Mobile Robots

7



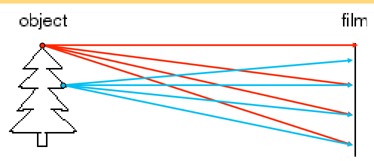
The Camera

8

- 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?

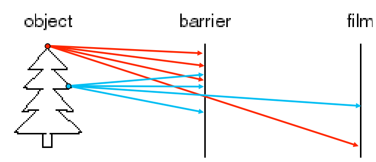
9



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

Pinhole camera

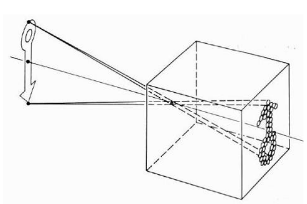
10



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

Pinhole Camera Model

11



- 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

12

Why so blurry?

Exposure needs light

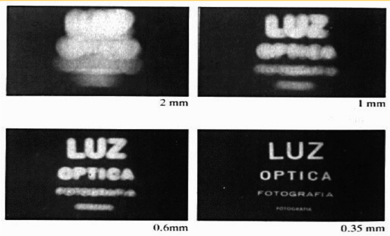
Smaller hole = darker

Larger hole = blur

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

Shrinking the aperture

13

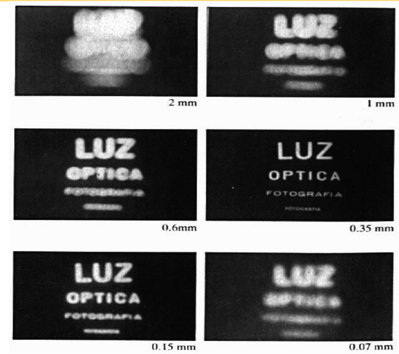


- Why not make the aperture as small as possible?
  - Less light gets through (must increase the exposure)
  - Diffraction effects…

Slide by Steve Seitz

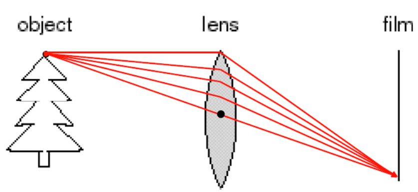
Shrinking the aperture

14



Solution: adding a lens

15



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

Range (Distance) Sensors

17

- 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

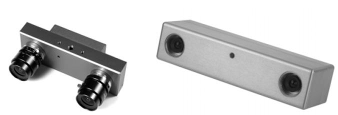
Active

…and that's a camera!

Distance Using Vision

18

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

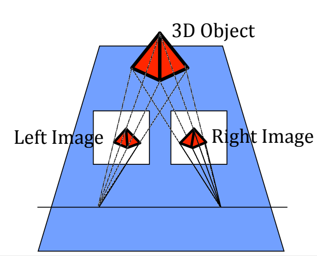


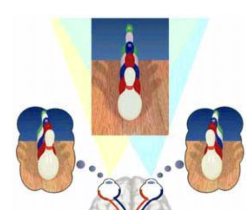
- Structure from motion:
  - Use a single moving camera
  - 3D structure and camera motion can be estimated

Stereo Vision

19

- 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

20

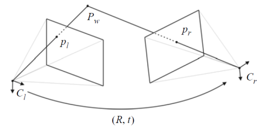
- Simplified: assume cameras are
  - Identical
  - Aligned on a horizontal axis
- Distance is *inversely* proportional to disparity
  - Closer objects can be measured more accurately
- Disparity is proportional to b
  - For a given disparity error, the accuracy of the depth estimate increases with increasing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

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

Calibration and Alignment

21

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

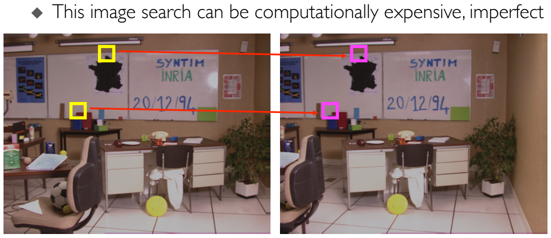



- Need to estimate relative pose between cameras
  - *Rotation* and *translation* – and since cameras are not identical, also
    - focal length, image center, radial distortion
- Epipolar rectification: compare two feature-rich images

Correspondence

22

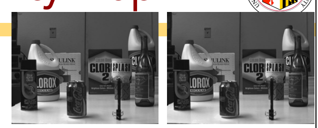
- Two cameras see slightly different scenes
  - What points in one correspond to points in the other?
  - Compare all points in image to all points in other image
  - This image search can be computationally expensive, imperfect

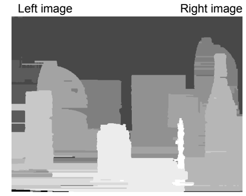


Output: Disparity Map

23

- 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

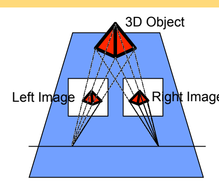
Distance images usually visualized as grey-scale images. Objects that are closer to the camera appear lighter, those who are further appear darker.

Disparity

24


Summary

25

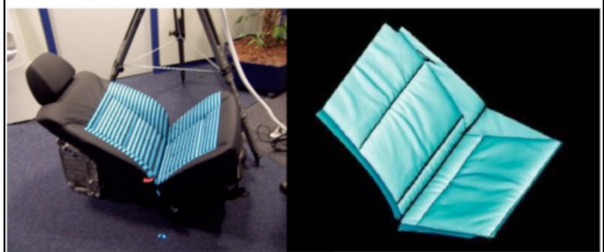


1. Stereo camera calibration → compute camera relative pose
   - Epipolar rectification → align images
2. Search correspondences
3. Output: compute stereo triangulation or disparity map
4. Consider baseline and image resolution to compute accuracy!

Structured Light

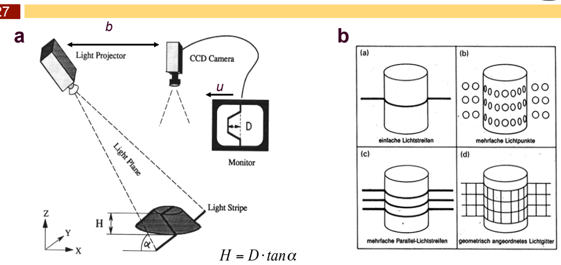
26

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



Structured Light

27

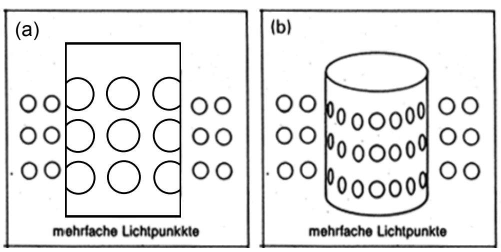


- Eliminate correspondence problem by projecting known light on the scene
- Light perceived by camera
- Range to an illuminated point can then be determined from geometry

Structured Light

28

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



Microsoft Kinect

29