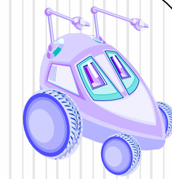
LECTURE 4 - 2

Common Sensing Techniques for Reactive Robots

Introduction to AI Robotics (Sec. 6.6 – 6.9)



Quote of the Week

"Just as some newborn race of superintelligent robots are about to consume all humanity, our dear old species will likely be saved by a Windows crash. The poor robots will linger pathetically, begging us to reboot them, even though they'll know it would do no good."

-Anonymous

ANNOUNCEMENTS

- Lab 4 - Line Following (PI Control) is due on Thursday, 4/1/10
- The lab memo and code is due on Angel by midnight on Thursday, 4/1/10
- Quiz 8 on Sec. 6.6 – 6.9, Lecture 4-2 on Monday, 4/12/10

OBJECTIVES

Upon completion of this lecture the student should be able to:

- Define the following terms in one or two sentences: hue, saturation, image, pixel, image function, computer vision
- Given an RGB image and a range of color values be able to threshold the image on color and construct a color histogram
- Write computer vision code to enable a robot to imprint on and track a color

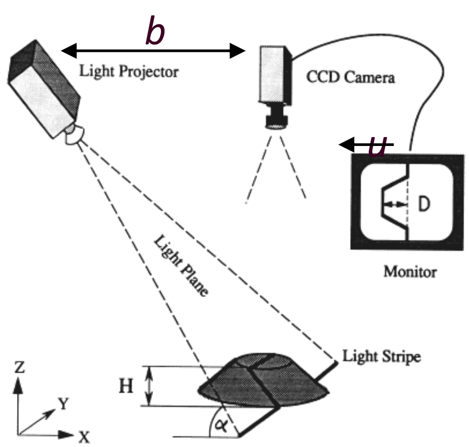
COMPUTER VISION

- *Computer Vision* refers to the processing of data from an image
- An *image* is a representation of data in a picture-like format where there is a physical correspondence to a scene
- The elements in image arrays are called *pixels* and maps into a small region of space
- The most common use in reactive robotics is for *region segmentation* to identify a region in the image with a particular color
- *Color histogramming* is used to identify a region with several colors and math the proportion in a region

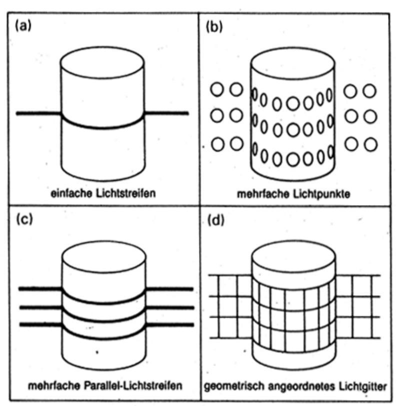
STRUCTURED LIGHT (VISION, 2 OR 3D)

- *Triangulation* can be used to find the distance to a large set of points by replacing a 2D receiver by a CCD or CMOS camera

- The emitter must project a known patter, or *structured light*, onto the environment
  - Light textures
  - Collimated light with a rotating mirror
  - Laser stripe using a prism



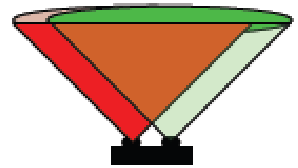
$$H = D \cdot \tan \alpha$$



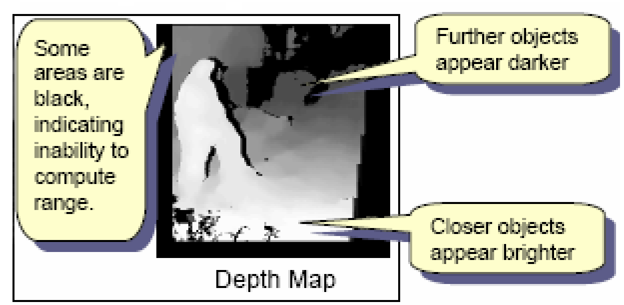
VISION RANGING SENSORS

- In mobile robotics, it is natural to attempt to implement ranging using vision
- Vision collapses the 3D world into a 2D image
- To recover depth information look at several images of a scene
  - The images must be different
  - They should provide differ viewpoints yielding *stereo or motion algorithms*
  - Alternately, do not change the viewpoint but change the camera geometry (i.e. focus or lens iris) yielding *depth from focus algorithms*

STEREO RANGING SYSTEMS:
GOAL

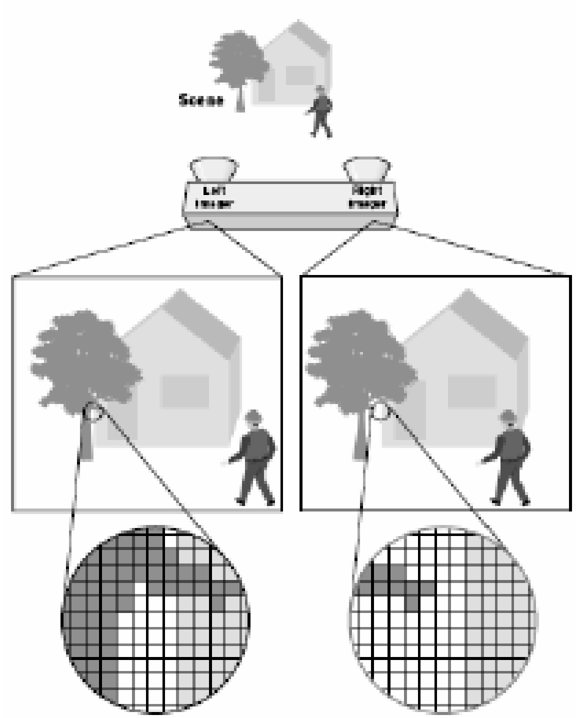


- Calculate the depth or distance of features in an image relative to the sensors (construct a *depth map*)
  - Use images from dual cameras aimed at the same object
  - Locate the same ‘feature’ in both images
  - Use geometric relationships between the 2 cameras and the location of the feature in each image
  - The depth of each feature can be triangulated and a depth map constructed



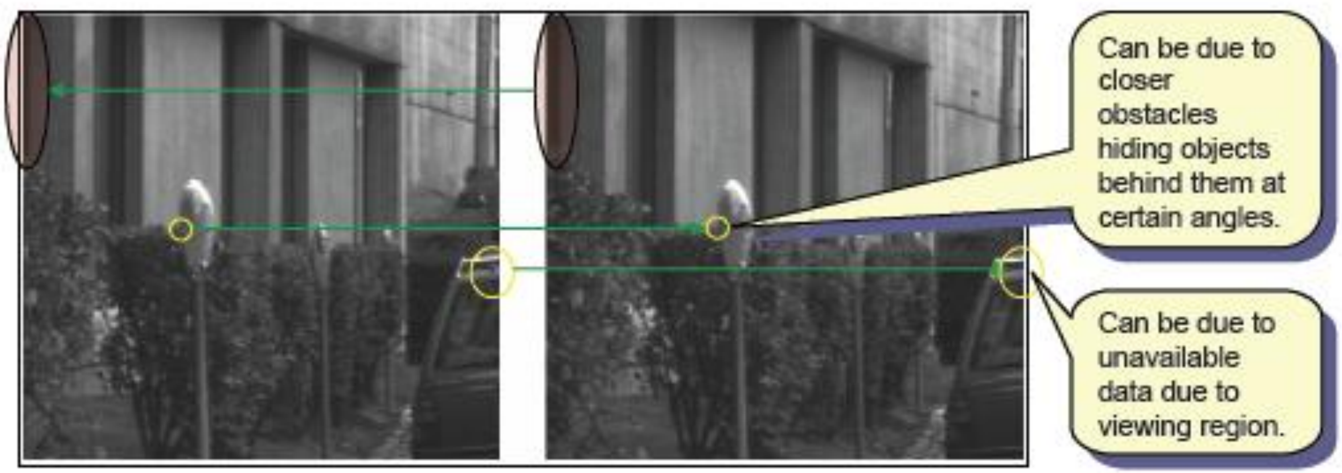
STEREO RANGING SYSTEMS: STEREO VISION

- Objects in left camera appear horizontally shifted from objects seen in right camera
- The size of the shift is the *disparity*
- The ideas is to find a *correspondence* (or match) between points in one image with points in other image



STEREO RANGING SYSTEMS:
STEREO VISION

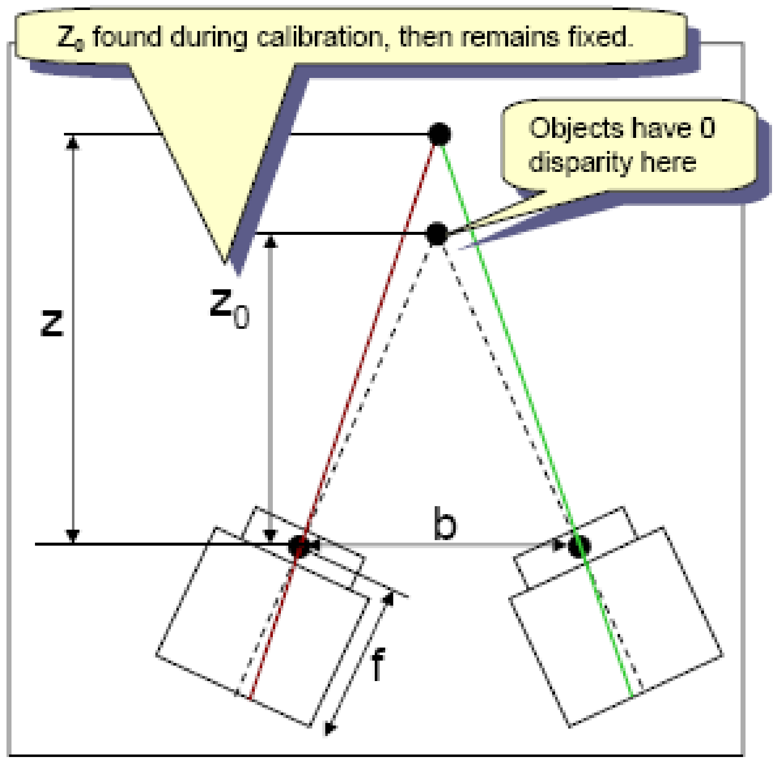
- It is difficult to find corresponding pixels in 2 images
- It is better to find the most likely match
- In some cases, the pixel in one image may not be visible in the other (*occlusion*)



STEREO RANGING SYSTEMS: STEREO VISION

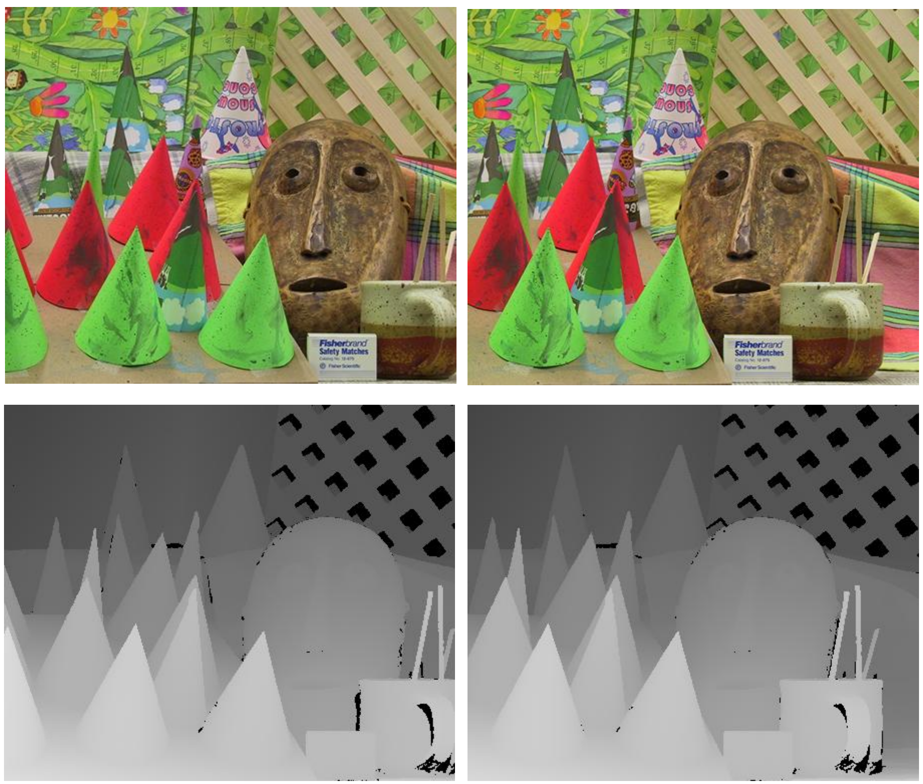
A more realistic scenario is when the cameras do not lie on the same plane

$$z = (f \cdot b)/(d + (f \cdot b)/z_o)$$



STEREO VISION

- 3D information can be computed from two images



- Compute *disparity*
  - displacement of a point in 2D between the two images
- Disparity is inverse proportional with actual distance in 3D
- Compute relative positions of cameras

STEREO VISION

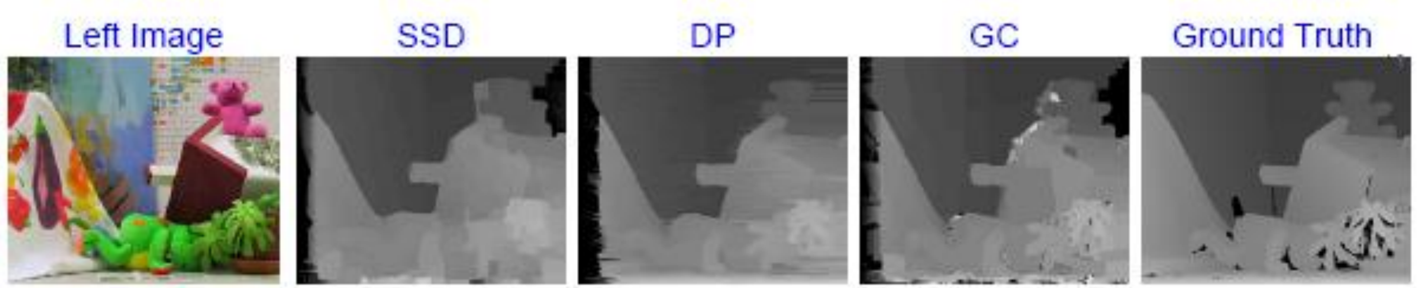
1. Distance is inversely proportional to *disparity*
   - closer objects can be measured more accurately
2. Disparity is proportional to b.
   - For a given disparity error, the accuracy of the depth estimate increases with increasing baseline b.
   - However, as b is increased, some objects may appear in one camera, but not in the other.
3. A point visible from both cameras produces *a conjugate pair*
   - Conjugate pairs lie on *epipolar line*

STEREO RANGING SYSTEMS: CORRESPONDENCE

- Desired characteristics
  - Corresponding image regions are similar
  - Each point matches a single point in the other image (unlikely)
- Two main matching methods
  - *Feature-based*
    - Start from image structure (e.g. edges)
  - *Correlation-based*
    - Start from grey levels

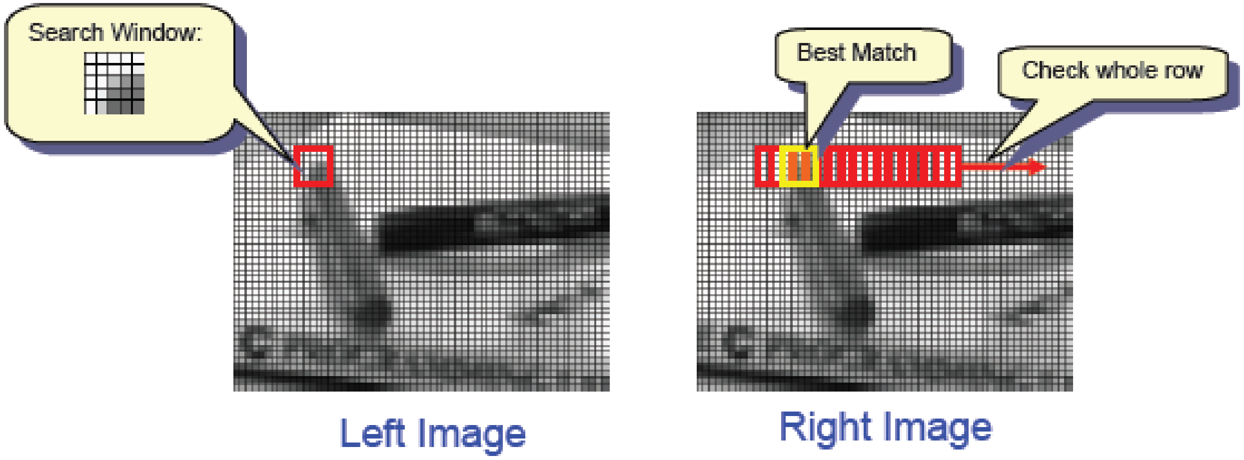
STEREO RANGING SYSTEMS: CORRELATION

- There are several methods
  - Sum of Squared Difference (SSD)
  - Dynamic Programming (DP)
  - Graph Cut (GC)
  - Belief Propagation (BP)
  - Markov Random Fields (MRF)



STEREO VISION:
SSD CORRELATION

- Take a small area of data in left image and compare it with similar-size area in the right image along the same *epipolar line* (i.e. same height in the image if the cameras are horizontally level)

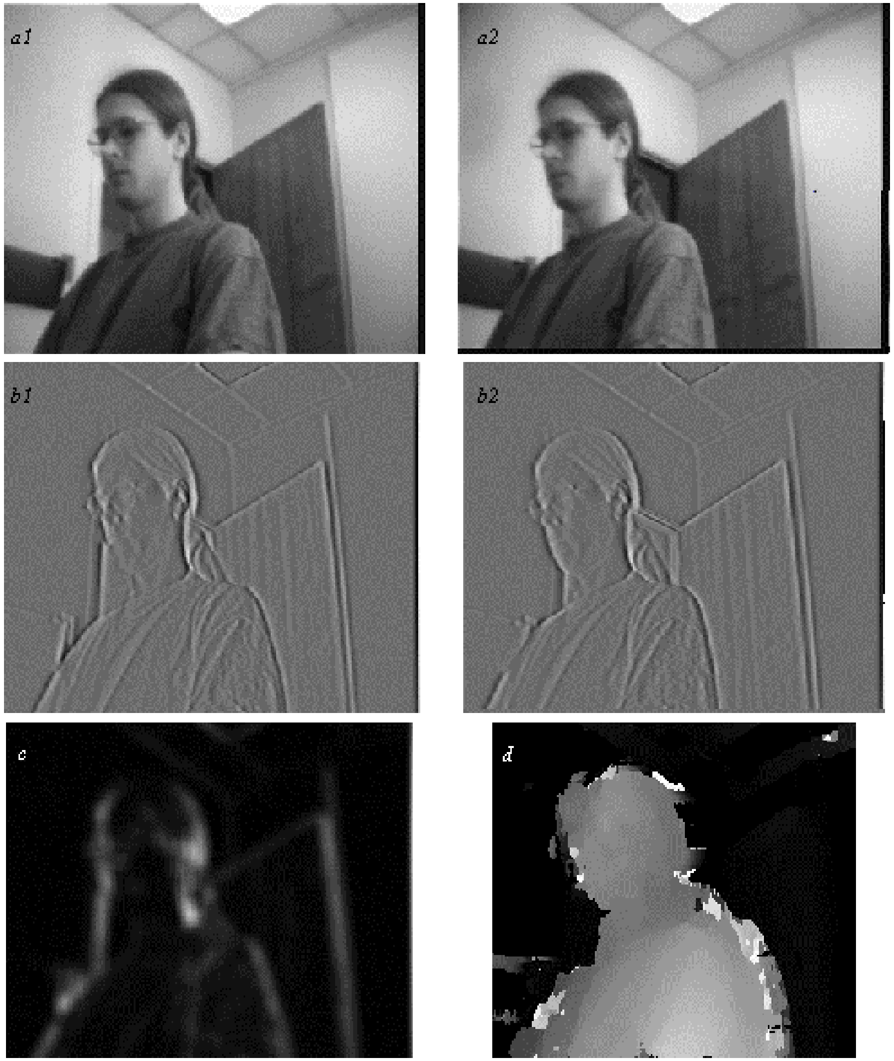


STEREO VISION:
CORRELATION

- To improve matching
  - Apply image filters before and after processing
  - Identify corners and edges to help fill in areas with no data available
  - Use sensor fusion (i.e. data from other sensors) to fill in missing gaps
  - Project structure light onto objects to improve matches

STEREO VISION EXAMPLE

- Extracting depth information from a stereo image
  - a1 and a2: left and right image
  - b1 and b2: vertical edge filtered left and right image;
    filter = [1 2 4 -2 -10 -2 4 2 1]
  - c: confidence image:
    bright = high confidence (good texture)
  - d: depth image:
    bright = close; dark = far

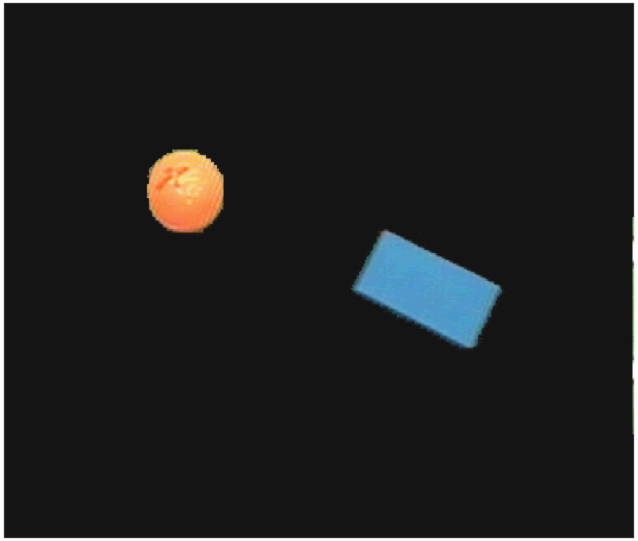


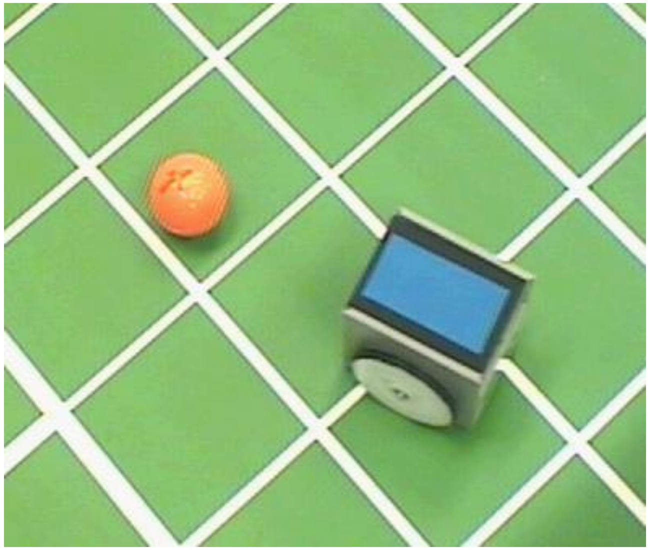
VISION FROM MOTION

- Take advantage of motion to facilitate vision
- Static system can detect moving objects
  - Subtract two consecutive images from each other $\Rightarrow$ the *movement* between frames
- Moving system can detect static objects
  - At consecutive time steps continuous objects move as one
  - Exact movement of the camera should be known
- Robots are typically moving themselves
  - Need to consider the movement of the robot

COLOR TRACKING SENSORS

- Unlike ultrasonic and infrared range finders, vision systems can also detect and track color in the environment





COLOR-TRACKING SENSORS

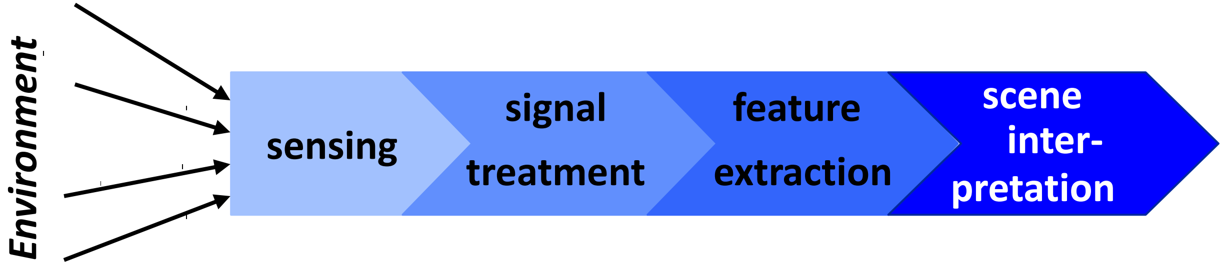
- There is no correspondence problem to be solved in such algorithms (it only requires one image)
- By using sensor fusion, color tracking can produce significant information gains

STEREO RANGING SYSTEMS

- Advantages
  - Better resolution than ultrasonic and infrared
  - Very reliable when environment is sufficiently cluttered
  - Often packaged with software to calculate depth
- Disadvantages
  - Cannot identify mirrors and/or glass
  - Sensitive to lighting conditions
  - Poor performance when environment lacks features
  - More expensive than ultrasonic and infrared
  - Larger than ultrasonic and infrared
  - Difficult to calibrate

FEATURE EXTRACTION: SCENE INTERPRETATION

- A mobile robot must be able to determine its relationship to the environment by sensing and interpreting the measured signals.
  - A wide variety of sensing technologies are available
  - However, the main difficulty lies in interpreting these data, that is, in deciding what the sensor signals tell us about the environment.
  - To extract information from one or more sensor readings to generate a higher level ***percept*** to inform the robot's environment model and action is ***feature extraction***



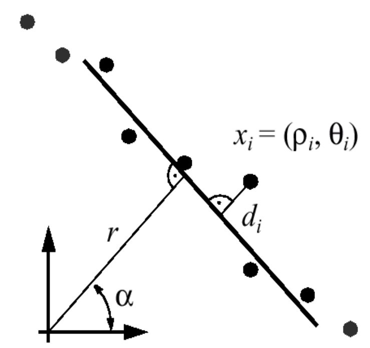
FEATURE EXTRACTION:
FEATURES

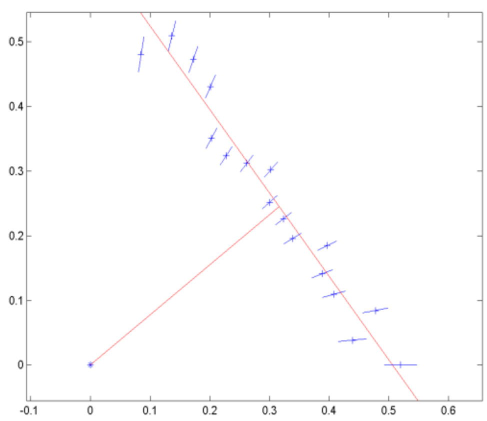
- Features are distinctive elements or geometric primitives of the environment.
- Good features are always perceivable and easily detectable form the environment
- They usually can be extracted from measurements and mathematically described.
  - *low-level features* include *geometric primitives* like lines, circles
  - *high-level features* include edges, doors, tables or trash cans.

In mobile robotics, features help for localization and map building.

FEATURE EXTRACTION:
RANGE DATA

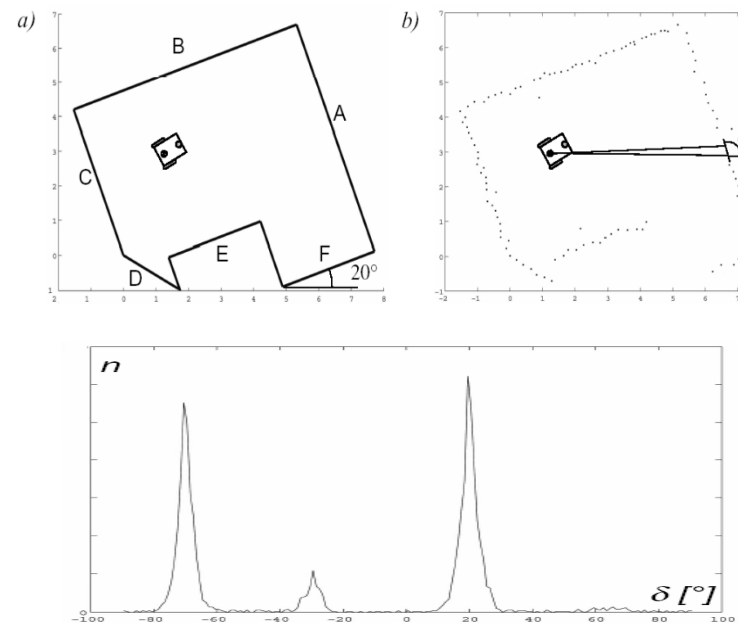
- Laser, Ultrasonic and vision-based ranging extract features that are geometric primitives such as line segments, circles, corners, edges
- Most other geometric primitives are too complex and no closed form solutions exist.
- However, lines segments are very often sufficient to model the environment, especially for indoor applications.





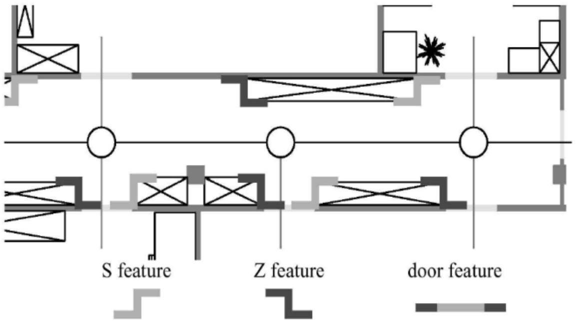
RANGE HISTOGRAM FEATURES
ANGULAR HISTOGRAM

- An *angular histogram* is a simple way of combining characteristic elements of an image
  - A 360 degree range can is performed
  - The hits are recorded on a map
  - An algorithm measures the relative angle between adjacent hits



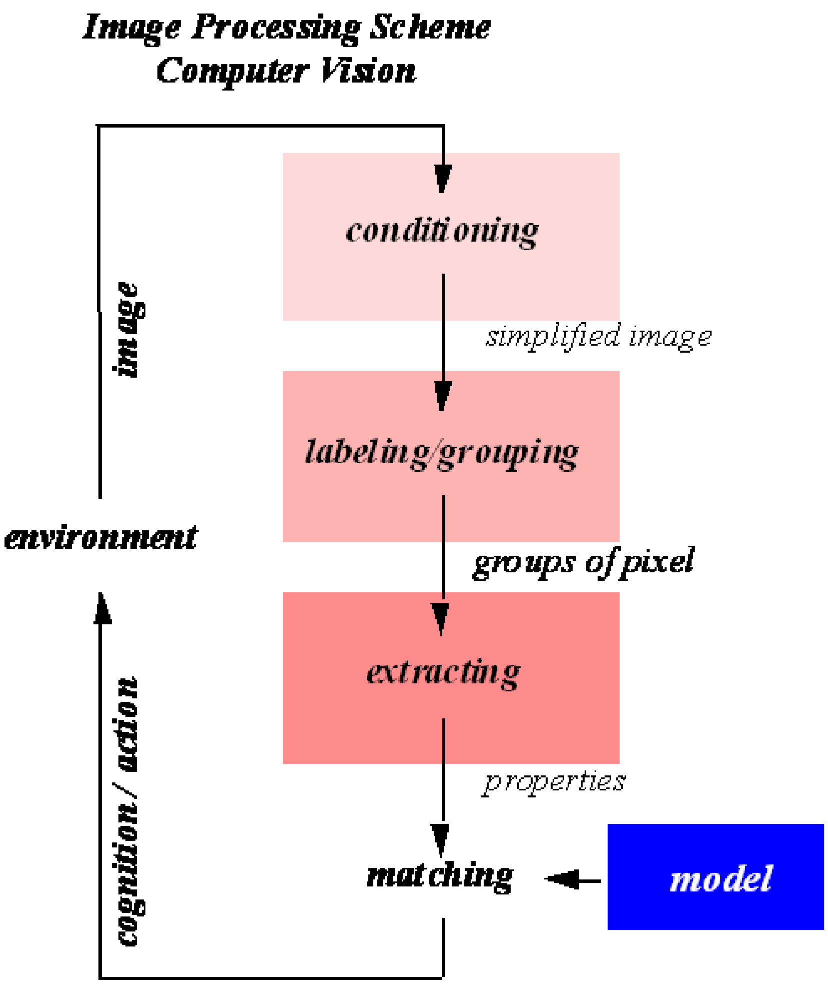
EXTRACTING OTHER GEOMETRIC FEATURES

- A robot must make use of multiple features simultaneously, comprising a *feature set* appropriate for its operating environment
- *Corner features* are defined as a point feature with an orientation
- *Step discontinuities* are a step change perpendicular to the direction of travel (concave or convex)
- *Doorways* are opening of the appropriate dimension in the wall, characterized by their width



VISUAL APPEARANCE: IMAGE PREPROCESSING

- Conditioning
  - Suppresses noise
  - Implemented with
    - gray-scale modification (e.g. thresholding)
    - (low pass) filtering
- Labeling
  - Determination of the spatial arrangement of the events, i.e. searching for a structure
- Grouping
  - Identification of the events by collecting together pixel participating in the same kind of event
- Extracting
  - Compute a list of properties for each group
- Matching

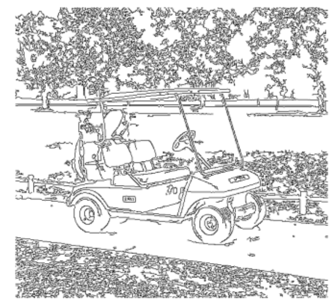


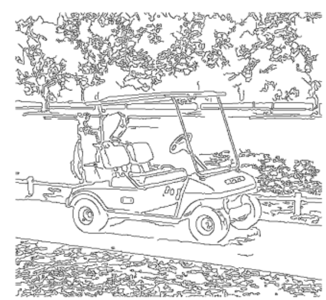
FEATURE EXTRACTION:
FILTERING AND EDGE DETECTION

- The single most popular spatially localized feature is *edge detection*
- Edges
  - Locations where the brightness undergoes a sharp change,
  - Differentiate one or two times the image
  - Look for places where the magnitude of the derivative is large.
  - Noise, thus first filtering/smoothing required before edge detection
- Gaussian Smoothing
  - Removes high-frequency noise
  - Convolution of intensity image I with G

EDGE DETECTION

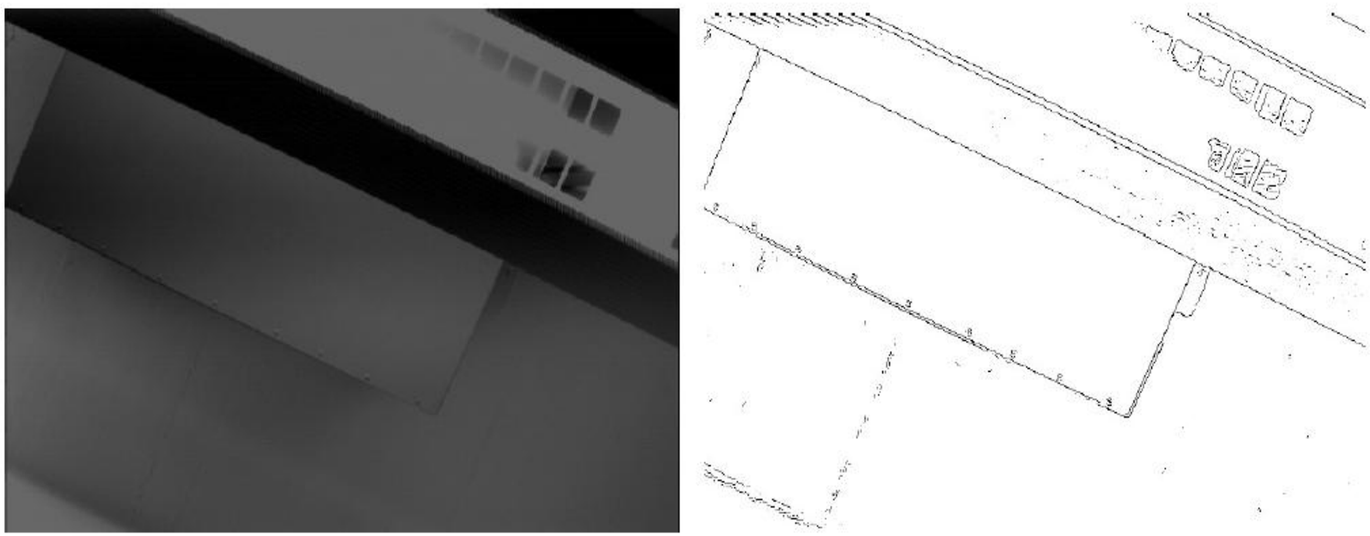
- *Edge* :a curve in the image across which there is a change in brightness
- Finding edges
  - Differentiate the image and look for areas where the magnitude of the derivative is large
- Difficulties
  - Not only edges produce changes in brightness: shadows, noise
- Smoothing
  - Filter the image using *convolution*
  - Use filters of various orientations
- *Segmentation*: get objects out of the lines





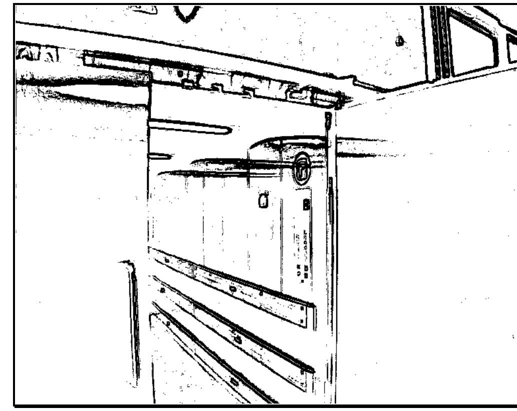
FEATURE EXTRACTION:
EDGE DETECTION

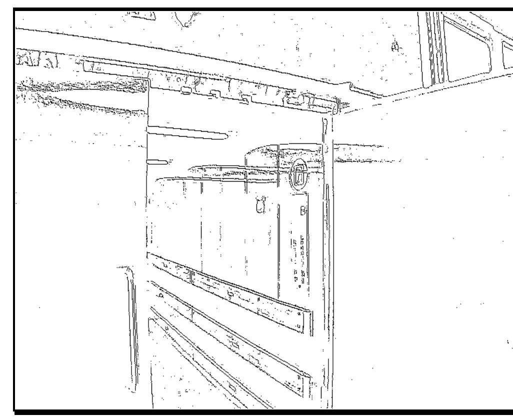
- Ultimate goal of edge detection
  - an idealized line drawing.
- Edge contours in the image correspond to important scene contours.



FEATURE EXTRACTION: NONMAXIMA SUPPRESSION

- Output of a Canny edge detector is usually a black and white image where the pixels with gradient magnitude above a predefined threshold are black and all the others are white
- *Nonmaxima suppression* sets all pixels to zero that do not represent the local maxima
- *Nonmaxima suppression* generates contours described with only one pixel thinness



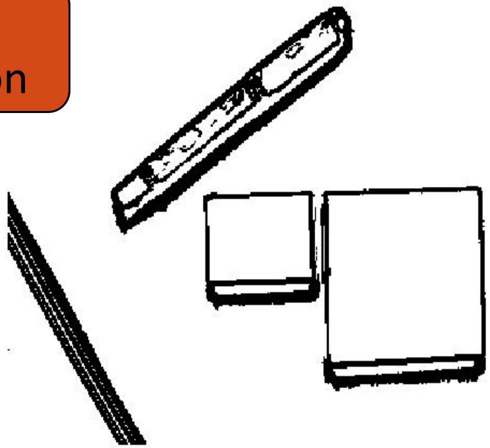


FEATURE EXTRACTION EXAMPLE

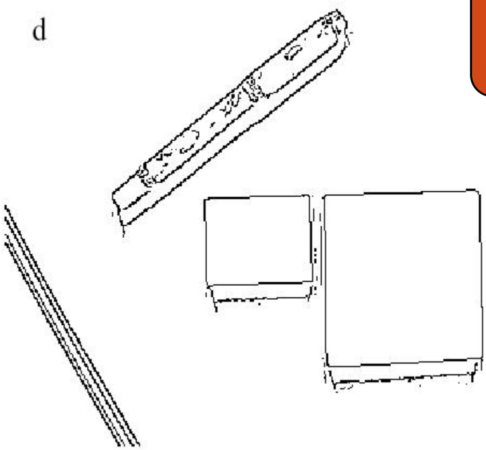
Raw Image

Sobel Filter

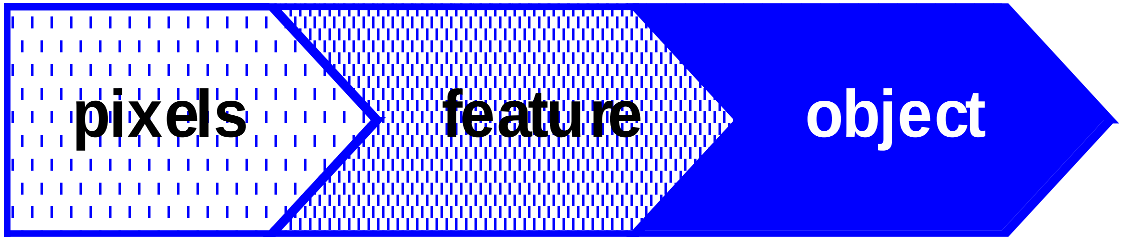
Edge Detection



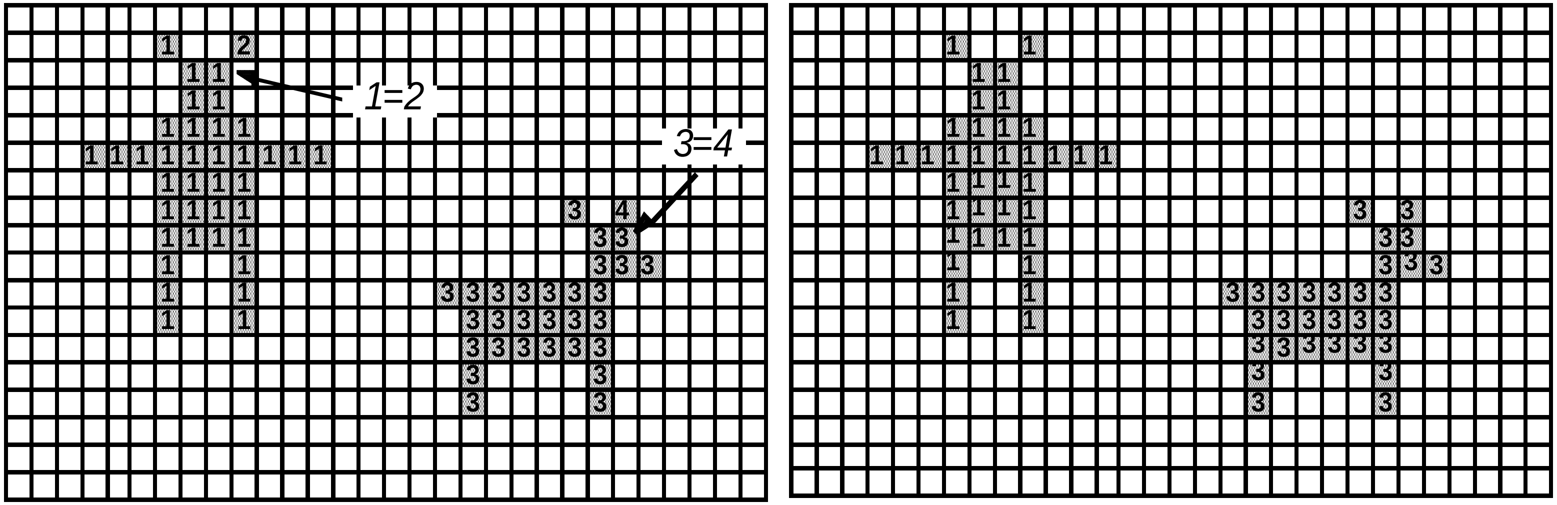
Nonmaxima Suppression



GROUPING, CLUSTERING: ASSIGNING FEATURES TO FEATURES



Connected Component Labeling



FEATURE EXTRACTION:
FLOOR PLANE EXTRACTION

- Vision based identification of a traversable path
- The processing steps
  - As pre-processing, smooth I_f using a Gaussian smoothing operator
  - Initialize a histogram array H with *n* intensity values
  - For every pixel (x,y) in I_f increment the histogram:

FEATURE EXTRACTION:
WHOLE-IMAGE FEATURES

- Whole-Image features are not designed to identify specific spatial structures
- They sever as a compact representation of the entire local region
- Extract one or more features that are correlated with the robot's position for localization