Autonomous Mobile Robots

Module 19: Feature extraction based on visual appearance

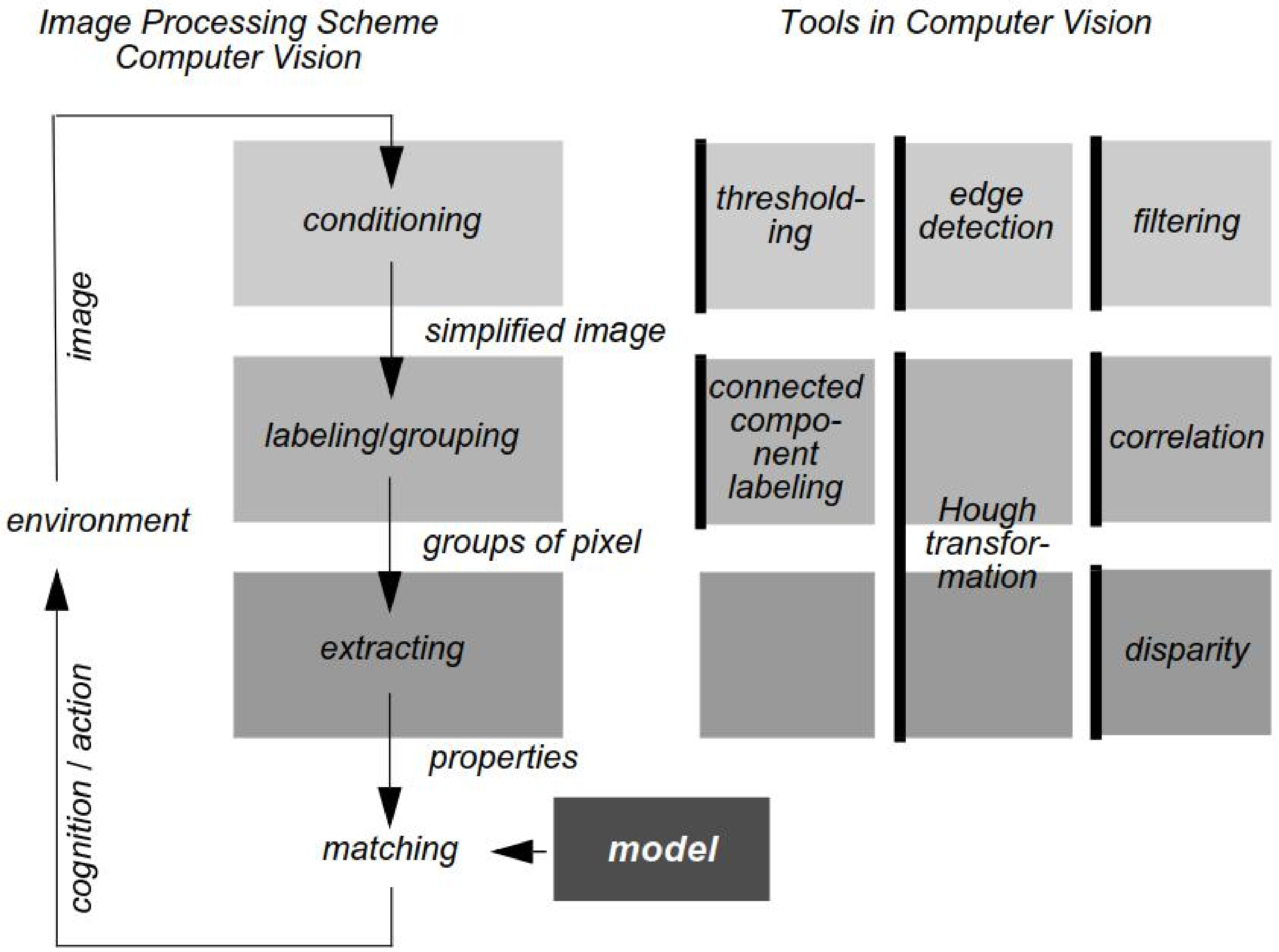
Feature extraction based on visual appearance

- Appearance-based feature extraction relevant to mobile robotics
- The method must operate in real-time
- Method must be robust to real-world conditions outside laboratory
- Cannot make a lot of assumptions on the illumination and color
- Vision-based interpretation tries to reduce the information as much as possible
- A sonar sensor can produce data at around 50 bits per second, a camera can output 240 million bits per second - hence large volume of data is to be dealt with
- Hence vision-based feature extraction is essential to reduce volume

Feature extraction based on visual appearance

- This large volume of information needs to be reduced to a smaller size by keeping most important things and discarding relatively less important things
- Vision-based feature extraction can be classified into two types
  - **Spatially localized features** – features found in regions of image, at specific locations
  - **Whole-image features** – functions of entire image or large area of image
- All vision sensors provide images with noise, need a filtering or preprocessing step to reduce noise

Feature extraction based on visual appearance



Reproduced from: Siegwart, Roland, Illah Reza Nourbakhsh, and Davide Scaramuzza. Introduction to autonomous mobile robots. MIT press, 2011.

Feature extraction based on visual appearance

- **Image Preprocessing**
- Before extracting features, image needs to be smoothed
- If feature extraction is done without this step, there will be a lot of noisy results and misleading features
- To smooth the images, a standard process adopted is to perform convolution with Gaussian distribution function

Smooth → low pass filter : removes high frequency.

Feature extraction based on visual appearance

Smoothing : convolution with Gaussian dist. fn.

$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau)\, g(t-\tau)\, d\tau$$

For images: $I' = G \otimes I$

I : original image

G : Gaussian : discrete kernel.

Eg: 3x3 $\quad G = \frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$

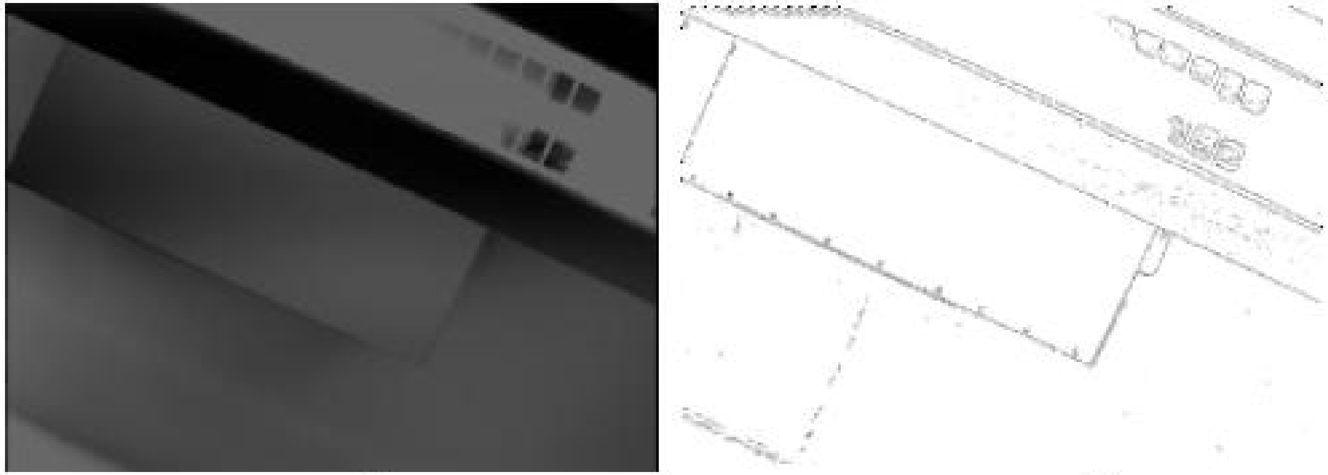
5	7	21	11	.	18	29
9	11	14	13	.	71	33
3	8	5	14	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
19	15	75	.	.	51	33
61	49	87	.	.	36	43

This operation is essentially a weighted averaging.

low pass filter

Feature extraction based on visual appearance

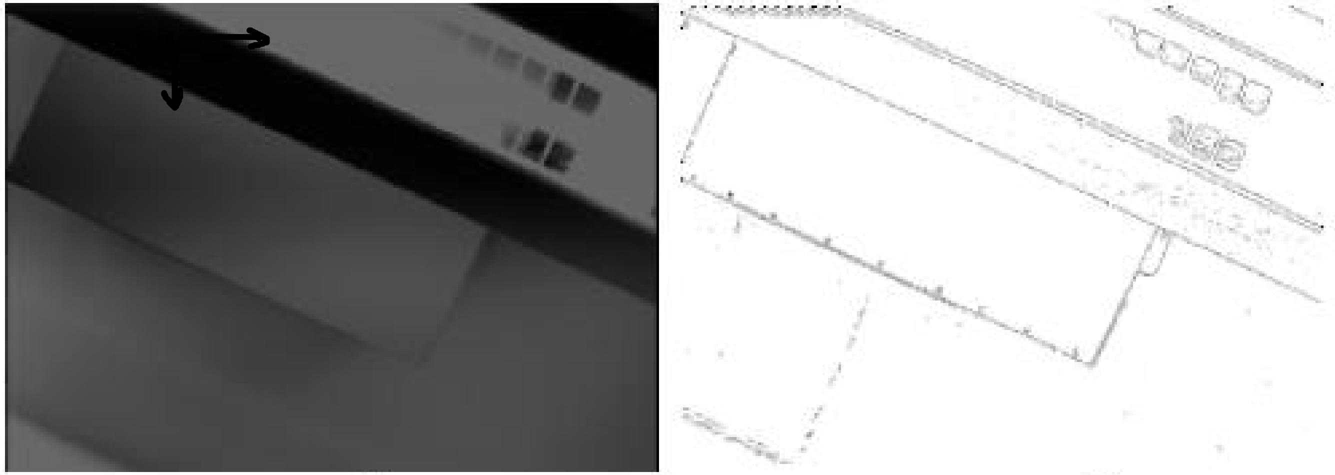
- We will focus only on spatially localized features
- Within spatially localized features, we will focus only on edge detection
- Edge detection is a very commonly used local feature extractor



a) b)

Reproduced from: Siegwart, Roland, Illah Reza Nourbakhsh, and Davide Scaramuzza. Introduction to autonomous mobile robots. MIT press, 2011.

Feature extraction based on visual appearance


a) b)

Differentiating const = 0

$\frac{\partial}{\partial x}$ $\frac{\partial}{\partial y}$

$\frac{\partial}{\partial t}$

Threshold T · edge pixel needs to exceed

- We can understand edge detection as a differentiation process, or a gradient operation
- Edges define regions in the image plane where significant change in image brightness takes place
- The edge image has significantly lesser volume of information than original image

Reproduced from: Siegwart, Roland, Illah Reza Nourbakhsh, and Davide Scaramuzza. Introduction to autonomous mobile robots. MIT press, 2011.

Feature extraction based on visual appearance

- Since edge is a sharp transition, it is important that the image is smoothed first before performing edge detection
- Canny edge detection is one example of edge detection algorithm

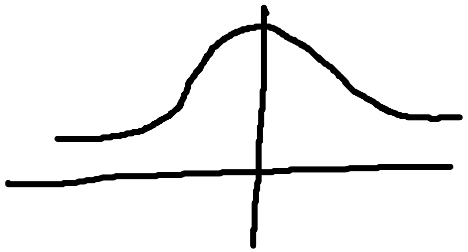
John Canny 1983

Gaussian smoothing $I_1 = G \otimes I$

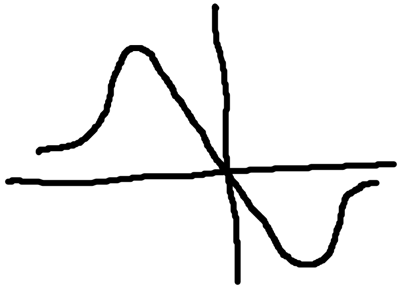
Derivative of smoothed image $I_2 = f \otimes I_1 = (I_1)' = (G \otimes I)'$

Can represent with single convolution operation

$$I' = G' \otimes I$$

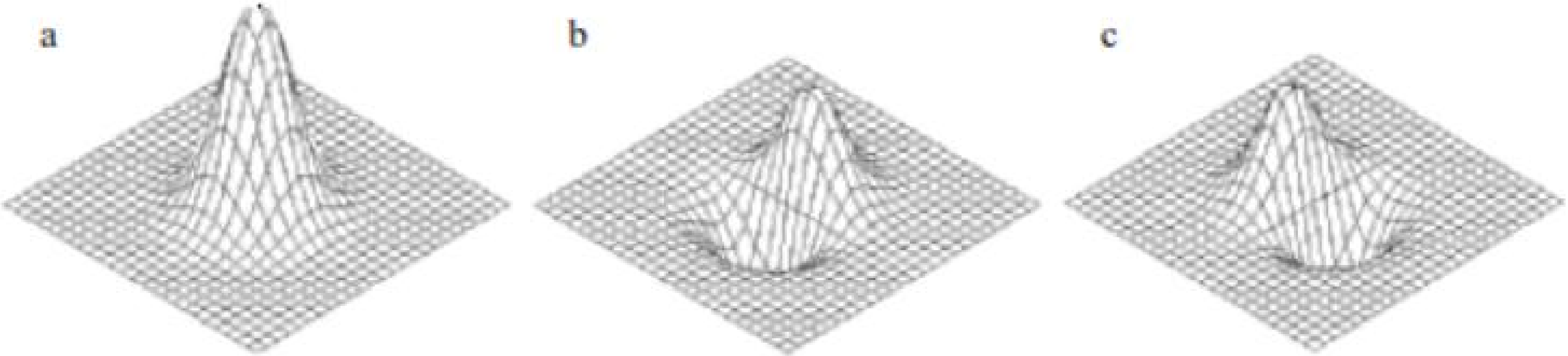


$$\frac{1}{\sqrt{2\pi}\,\sigma} e^{-x^2/2\sigma^2}$$



$$\frac{-x}{\sqrt{2\pi}\,\sigma^3} e^{-x^2/2\sigma^2}$$

Feature extraction based on visual appearance



$G(x)G(y)$ $\qquad$ $G'(x)\,G(y)$ $\qquad$ $G'(y)\,G(x)$

$$f_V(x,y) = G'(x)\,G(y)$$

$$f_H(x,y) = G'(y)\,G(x).$$

$$R_V = f_V \otimes I$$

$$R_H = f_H \otimes I$$

Other features:
- Floor plane

$$R(x,y) = \sqrt{R_V^2 + R_H^2}$$

If $R(x,y) > T$ (threshold)
then mark as an edge.

Reproduced from: Siegwart, Roland, Illah Reza Nourbakhsh, and Davide Scaramuzza. Introduction to autonomous mobile robots. MIT press, 2011.

Autonomous Mobile Robots

Module 20: Unit III - Localization & Mapping Introduction

Unit-III

LOCALIZATION AND MAPPING: Introduction - Bayes filter – Kalman Filter - Extended Kalman Filter - Information Filter - Histogram Filter - Particle Filter – Challenges of Localization- Map Representation- Probabilistic Map based Localization-Monte carlo localization- Landmark based navigation-Globally unique localization- Positioning beacon systems- Route based localization – Mapping - Metrical maps - Grid maps - Sector maps – Hybrid Maps – SLAM.

SLAM - Simultaneous Localization and Mapping

- Navigation : to go from one place to another autonomously
- Has 4 building blocks
- Perception: interpret sensor data
- Localization: determine position in environment
- Cognition: decide how to act to achieve goals
- Motion Control: modulate motor output

Recursive State Estimation

State: Collection of all aspects of a robot and its environment that can impact the future course of the robot

- State can change over time: dynamic state - position, velocity of moving robot, car
- State can be static - position of wall in a room
- Typical states we will deal with:
  - Mobile robot pose: position and orientation
  - For planar robot: pose is given by x,y,Ψ
  - Drone robot pose: (x,y,z) : 3 positions, (Φ,Θ,Ψ) : 3 orientations
  - Locations and features of surrounding objects in environment
  - Locations and velocities of moving objects and people
- State variable is represented by x

Environment Interaction

- Robot can influence the state of its environment through its actuators **- Control actions**
- Robot can gather information about the state through its sensors - **Sensor measurements**

- In practice, a robot continuously executes control action and measurements concurrently
- Control actions are represented by u
- Sensor measurements are represented by z
- Sensing measurements provide information about environment's state, hence it increases robot's knowledge
- Motion leads to a loss of knowledge due to uncertainty in actuation