Introduction to Cognitive Robotics

Module 5: Robot Vision

Lecture 2: Image processing

David Vernon
Carnegie Mellon University Africa

www.vernon.eu

Image Processing

Image processing

- Image to image transformation
- It starts with an image and produces a modified (enhanced) image
- Iconic to iconic transformation

Image analysis

- Image to information transformation
- It starts with an image and produces information representing a description or a decision
- Iconic to symbolic transformation

Image Processing

The image processing phase should :

- facilitate the extraction of information
- compensate for non-uniform illumination
- re-adjust the image to compensate for distortions introduced by the imaging system

Image Processing

There are 3 distinct classes of operations

1. Point Operations

2. Neighbourhood Operations

   - Linear filtering / convolution operations
   - Fourier transform
   - Logical non-linear operations

3. Geometric Operations

Point Operations

Each pixel in the output image is a function of the grey-level of the pixel at the corresponding position in the input image and only of that pixel

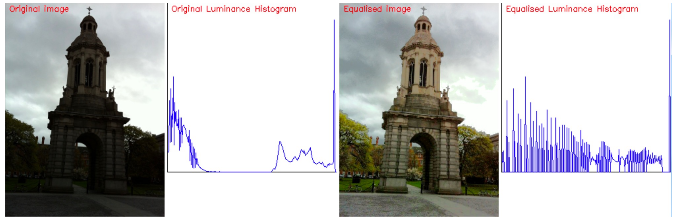
They cannot alter the spatial relationships of the image

Point Operations

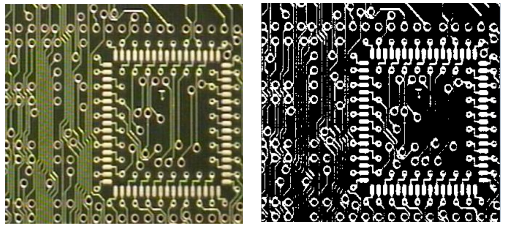
Typical point operations include:

- Photometric Decalibration, to remove the effects of spatial variations in the sensitivity of a camera system

- Contrast Stretching (*e.g.* if a feature or object occupies a relatively small section of the total grey-scale image, these local operations can manipulate the image so that it occupies the entire range)



- Thresholding, in which all the pixels having grey-levels in specified ranges in the input image are assigned a single specific grey-level in the output image



Point Operations

Background Subtraction

- Pixel values of two images are subtracted on a point by point basis.
- Subtraction of a *known* pattern (or image) of super-imposed noise, e.g. photometric decalibration (compensation for vingnetting)
- Motion Detection: stationary objects cancel each other out while moving objects are highlighted.

Point Operations

Thresholding

- Elementary segmentation technique (see later) that
  - assigns a value of 0 or 255 to each pixel
  - depending on whether the image value is less than or greater than the threshold
- This is effectively a labelling process
  - Label pixels background
  - Label pixels foreground / object

Neighbourhood Operations

- Generate an output pixel on the basis of the pixel at the corresponding position in the input image *and on the basis of its neighbouring pixels*

- The size of the neighbourhood may vary :

  3 * 3
  5 * 5
  63 * 63 pixels

- Often referred to as Filtering Operations

Convolution of an image f with a filter kernel or mask h

$$g = f * h$$

Neighbourhood Operations

Other Neighbourhood Operations

- Applying some logical test

  based on, e.g. the presence or absence of object pixels in a local neighbourhood surrounding the pixel in question

- Object Thinning (or Skeletonising)

- Erosion and Dilation (contract/expand an object)

Image Filtering

The 2D Convolution Integral :

$$g(i,j) = f * h = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(i-m, j-n) h(m,n) \, dm dn$$

- The output g of a shift-invariant linear system
- is given by the convolution or application of the input signal f with a function h which is characteristic of the system

Image Filtering

- The function h is normally referred to as the filter
  - It dictates what elements of the input image are allowed to pass through to the output image
  - By choosing an appropriate filter, we can enhance certain aspects of the output and attenuate others
  - A particular filter h is often referred to as a filter kernel
- The form of g depends on
  - The input f
  - The form of the system h through which it is being passed
  - The relationship is given by the convolution integral

Image Filtering

In the discrete domain of digital images the convolution operation is given by:

$$g(i, j) = f * h = \sum_m \sum_n f(i-m, j-n)\, h(m, n)$$

The summation is taken only over the area where *(i-m, j-n)* is defined,
i.e. over the area where *f* and *h* overlap

Image Filtering

$h(-1,-1)$	$h(-1,0)$	$h(-1,+1)$
$h(0,-1)$	$h(0,0)$	$h(0,+1)$
$h(+1,-1)$	$h(+1,0)$	$h(+1,+1)$

3 * 3 Convolution Filter h

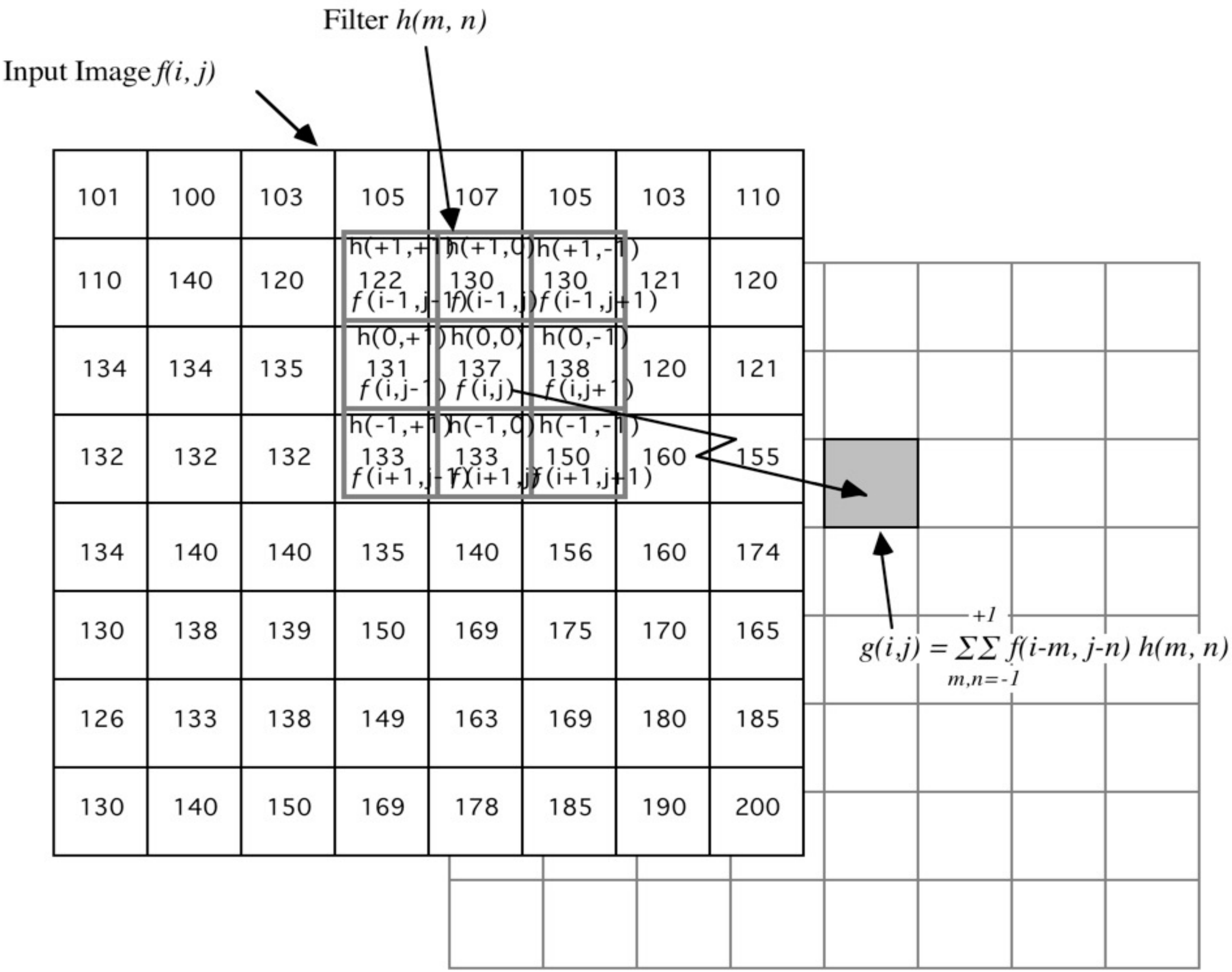

Filter h(m, n)
Input Image f(i, j)
101 100 103 105 107 105 103 110
110 140 120 122 130 130 121 120
134 134 135 131 137 138 120 121
132 132 132 133 133 150 160 155
134 140 140 135 140 156 160 174
130 138 139 150 169 175 170 165
126 133 138 149 163 169 180 185
130 140 150 169 178 185 190 200
h(+1,+1) h(+1,0) h(+1,-1)
f (i-1,j-1) f (i-1,j) f (i-1,j+1)
h(0,+1) h(0,0) h(0,-1)
f (i,j-1) f (i,j) f (i,j+1)
h(-1,+1) h(-1,0) h(-1,-1)
f (i+1,j-1) f (i+1,j) f (i+1,j+1)
g(i,j) = ΣΣ f(i-m, j-n) h(m, n)
+1
m,n=-1

Image Filtering

Note that the mask is first rotated by 180° since

$$g(i,j) = f * h = \sum_m \sum_n f(i-m, j-n)\, h(m,n)$$

- $f(i\text{-}1, j\text{-}1)$ must be multiplied by $h(1, 1)$
- $f(i\text{-}1, j)$ must be multiplied by $h(1,0)$
- ……. ,
- and $f(i+1, j+1)$ must be multiplied by $h(\text{-}1,\text{-}1)$

Often the rotation by 180° is omitted if the mask is symmetric

Image Filtering

Gaussian filter

Image smoothing

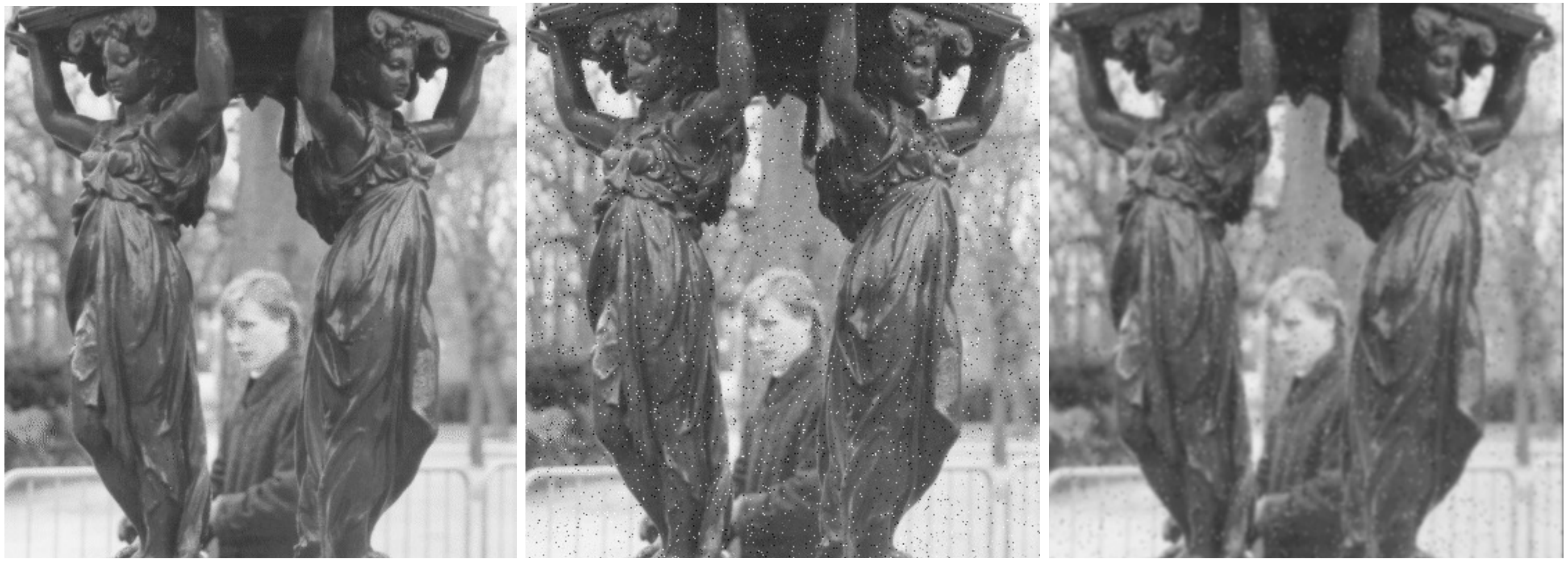
$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{\frac{-(x^2 + y^2)}{2\sigma^2}}$$

- σ defines the effective spread of the function
  - Gaussian functions with a small value for σ are narrow
  - Those with with a large value for σ are broad
  - Note that, since the Gaussian function is defined over an infinite support, i.e. it has non-zero (but very small) values at we must truncate the function

Image Filtering

Salt and pepper noise: 1% of pixels either black or white

Smoothing: outliers are spread but not eliminated

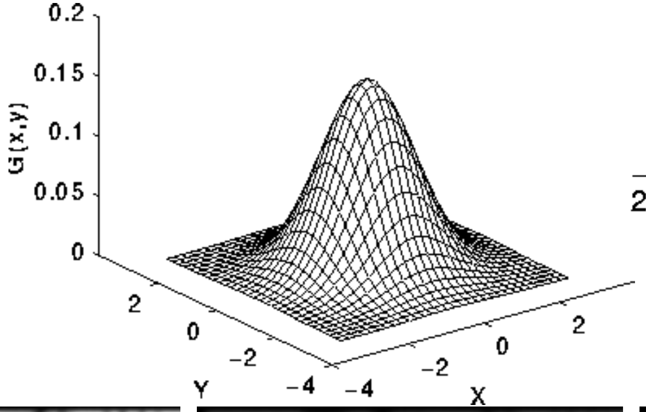


Original image   Salt and Pepper   Gaussian filter σ =8, 5x5

Credit: Markus Vincze, Technische Universität Wien

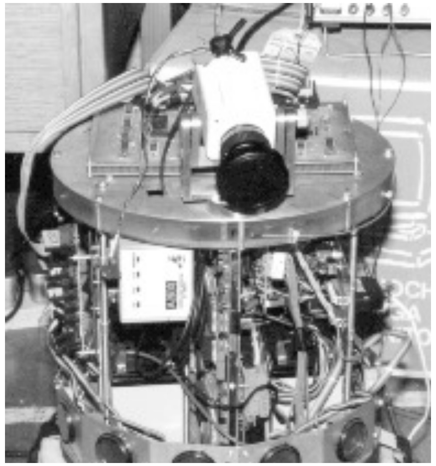
Image Filtering

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{\frac{-x^2+y^2}{\sigma^2}}$$

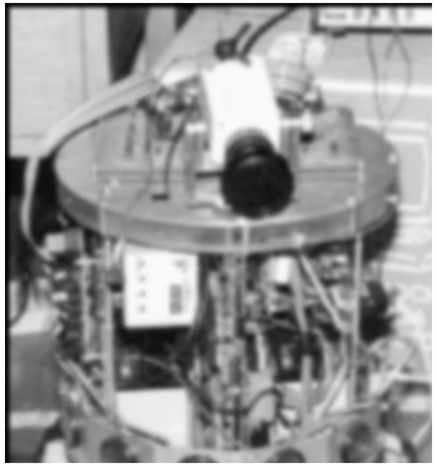


$\frac{1}{273}$

1	4	7	4	1
4	16	26	16	4
7	26	41	26	7
4	16	26	16	4
1	4	7	4	1



Original image



σ =1, 5x5

σ =2, 9x9

σ =4, 15x15

Credit: Markus Vincze, Technische Universität Wien

Image Filtering

Why is the Gaussian such a popular smoothing function?

It can be adjusted to control the level of detail in the image

- Small σ will retain a significant amount of detail
- Large σ will retain only the gross structure

Image Filtering

The Gaussian function optimizes the trade-off between two conflicting requirements

- Localization in the spatial frequency domain
- Localization of the space domain

- The more concentrated $f(x, y)$ is, the more spread out its Fourier transform $F(\omega_x, \omega_y)$
  - If we squeeze a function in x or y, its Fourier transform stretches out in ω_x or ω_y
  - It is not possible to arbitrarily concentrate both a function and its Fourier transform

- The Gaussian function is smooth and localized in both the spatial and frequency domains
  - The Fourier Transform of a Gaussian is a Gaussian

- "It is the unique distribution that is simultaneously optimally localized in both domains"

David Marr, Vision, 1982.

Image Filtering

Median filter

Noise suppression technique

- A pixel is assigned the value of the median of pixel values in some local neighbourhood
- The size of the neighbourhood is arbitrary
- Median filter is superior to the mean filter in that image blurring is minimised

Image Filtering

- Median Filter: Salt and Pepper (5%)
- Eliminates outliers (up to 50%)

123	125	126	130	140
122	124	126	127	135
118	120	150	125	134
119	115	119	123	133
111	116	110	120	130

Neighbourhood values:

115, 119, 120, 123, 124,
125, 126, 127, 150

Median value: 124

Median filter 3x3

Median filter 7x7

3x Median filter 3x3

Credit: Markus Vincze, Technische Universität Wien

Reading

R. Szeliski, *Computer Vision: Algorithms and Applications*, Springer, 2010.

- Section 2.2.3 Optics
- Section 2.3 The digital camera
  - Section 2.3.1 Sampling and aliasing
  - Section 2.3.2 Colour
- Section 3.1 Point operations
  - Section 3.1.1 Pixel transform
  - Section 3.1.4 Histogram equalization
- Section 3.2 Linear Filtering

D. Vernon, *Machine Vision*, 1991.

- Section 2.2.1 Image formation: elementary optics
- Section 2.2.2 Camera sensors
- Section 3.1 Sampling and quantization
- Section 4.1.4 Background subtraction
- Section 4.2.1 Convolution