

Computer Vision Notes

Created: 2024-09-06

Updated: 2024-09-08

References:

- Introduction to Image Understanding course at the University of Toronto

Linear Filters (TODO: Tb ch 3.2)

Digital Image: a map $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ or a matrix I of integer intensity values $\in [0,255]$, I is $m \times n$ in a grayscale image, $m \times n \times 3$ in a color image.

Problem: want to locate object in image.

Solution: slide and compare the image of the object.

Problem: noise in image.

Solution: modify pixel by applying function on a neighborhood of pixels e.g., average neighbors (assumes neighbors similar, noise independent) using moving average with (non-)uniform weights.

Correlation (cv2.filter2D, 2D moving average with (non-)uniform weights): Given input I , $G = F \otimes I$ where

$$G(i, j) = \sum_{u=-k}^k \sum_{v=-k}^k F(u, v) \cdot I(i + u, j + v)$$

where size of the weight **kernel/mask** F is $(2k + 1)^2$ and its entries $F(u, v)$ are **filter coefficients**. where $\sum \sum F(u, v) = 1$.

Let $\vec{f} = F(:, \cdot)$, $\vec{t}_{ij} = T_{ij}(:, \cdot)$ where $T_{ij} = I(i - k : i + k, j - k : j + k)$, then

$$G(i, j) = \vec{f}^T \cdot \vec{t}_{ij}$$

Normalized Cross-Correlation: exact match of image crop and filter results in 1.

$$G(i, j) = \frac{\vec{f}^T \cdot \vec{t}_{ij}}{\|\vec{f}\| \|\vec{t}_{ij}\|}$$

Types of Filters

Sharpening Filter:

$$F = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Gaussian Filter: smooth/blur, reduce noise, neighbors closest to a center have the most influence.

$$h(u, v) = \frac{1}{2\pi\sigma^2} e^{-\frac{u^2+v^2}{\sigma^2}}$$

Generic Gaussian Filter: anisotropic (asymmetric), $x \in \mathbb{R}^d$.

$$\mathcal{N}(x; \mu, \Sigma) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1} (x - \mu)\right)$$

Effect of Size of Filter and Variance

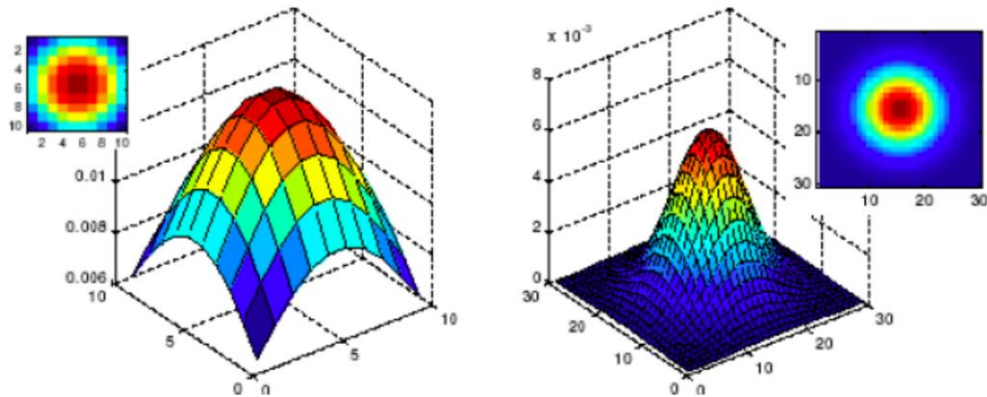


Figure: same $\sigma = 5$ different filter/mask/kernel size 10x10 vs 30x30.

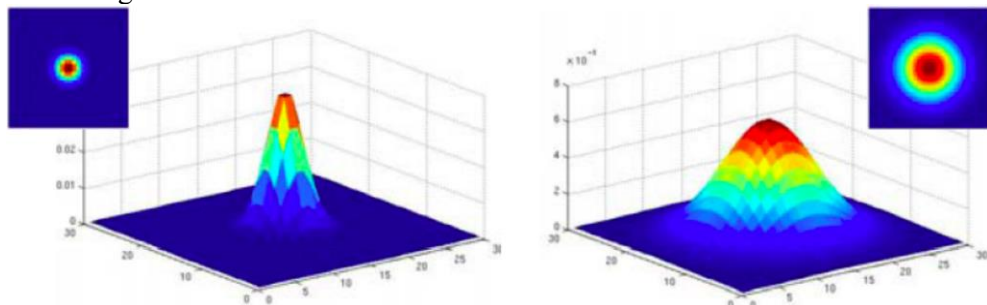


Figure: same size 30x30, $\sigma = 2$ (left) $\sigma = 5$ (right), larger is more smoothing.

Properties of Smoothing

- All values positive
- Sum to 1; prevents rescaling image
- Low-pass filter; removes high frequency (rate of change in pixel intensity values) components which include edges.

Convolution: operator that flips filter horizontally and vertically then applies correlation. Given input I , $G = F * I$ where

$$G(i, j) = \sum_{u=-k}^k \sum_{v=-k}^k F(u, v) \cdot I(i - u, j - v)$$

Properties of Convolution

Commutative	$f * g = g * f$
Associative	$f * (g * h) = (f * g) * h$
Distributive	$f * (g + h) = f * g + f * h$
Associative with scalar multiplier	$\lambda \cdot (f * g) = (\lambda \cdot f) * g$
Convolution Theorem (\mathcal{F} is Fourier Transform)	$\mathcal{F}(f * g) = \mathcal{F}(f) \cdot \mathcal{F}(g)$

Implications of Convolution Theorem

Method 1: convolution $f * g$ runs in N^2 .

Method 2: FFT and IFFT run in $N \log N$ and mult in N .

Separable Filters

[TODO]