The National Higher School of Artificial Intelligence
المدرسة الوطنية العليا للذكاء الاصطناعي

Course « Computer Vision»

Sid-Ahmed Berrani

2024-2025

Correlation and convolution

Linear spatial filtering can be described in terms of correlation and convolution.

- **Correlation:**

The process of moving a filter mask over a signal (the image in our case) and computing the sum of products at each location.

- **Convolution:**

Similar to correlation but the filter mask is first rotated by 180°.

Correlation

An example:

Suppose that we want to compute the correlation of the 1D signal:

f(x) = 0 0 0 1 0 0 0 0

with the mask:

w(x) = 1 2 3 2 8

0 0 0 0 **0 0 0 1 0 0 0 0** 0 0 0 0

1 2 3 2 8

0 0 0 8 2 3 2 1 0 0 0 0

Correlation vs. Convolution

- Correlation **is a function of displacement** of the filter. The first value of correlation corresponds to zero displacement, the second corresponds to one unit displacement, and so on
- Correlating a filter w with a function that contains all 0s and a single 1 yields a result that is a copy of w, but rotated by 180°
- Convolution works exactly the same way, but the filter is rotated by 180° before the shift operations.
- A fundamental property of convolution is that convolving a function with a unit impulse yields a copy of the mask at the location of the impulse.

2D Correlation/Convolution

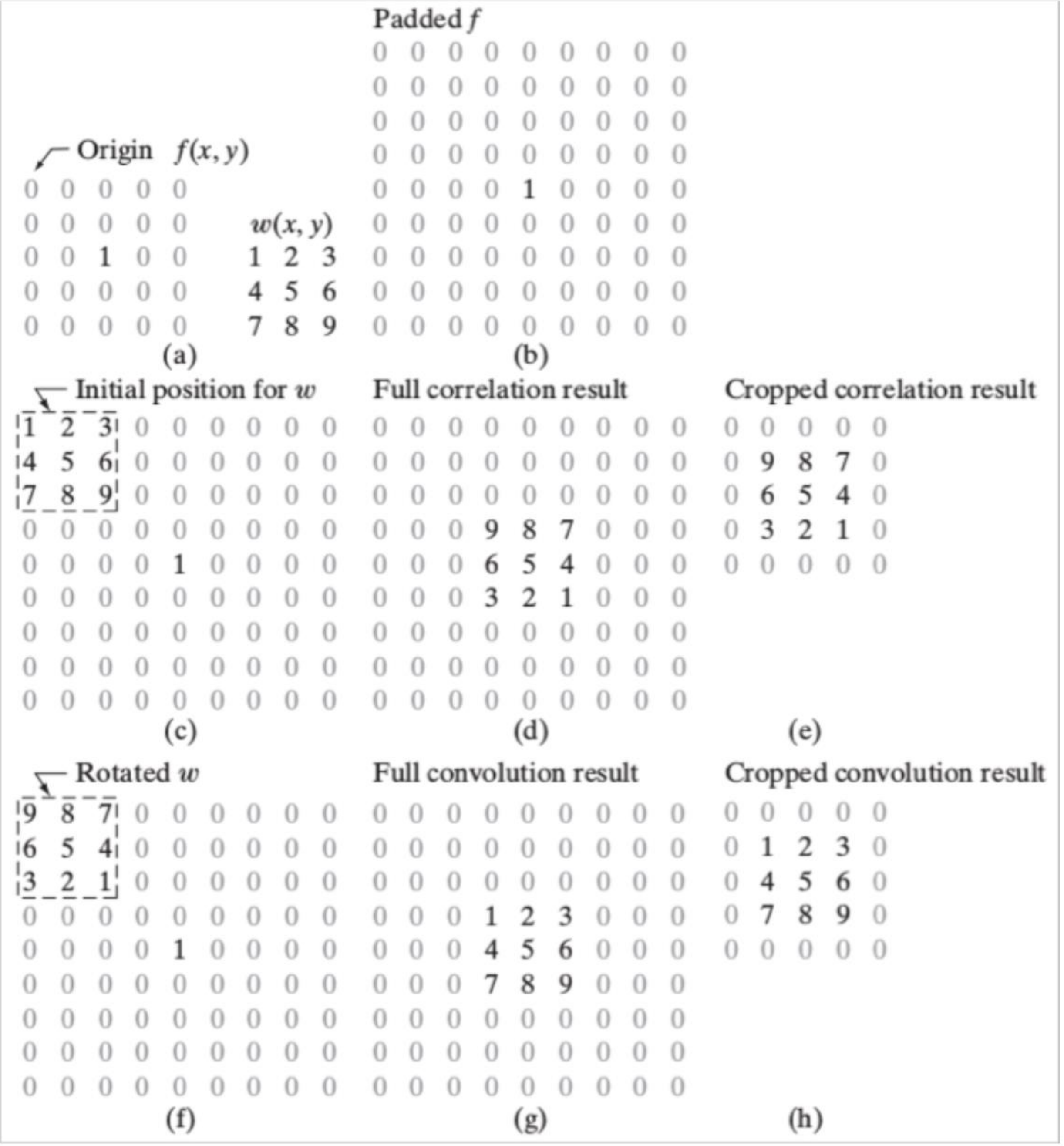
In case of 2D functions, like images, the correlation/convolution works in a similar manner:

- For a filter of size MxN we first pad the image with a minimum of:
  - M-1 rows at top and M-1 rows at bottom (filled with 0s)
  - N-1 cols at left and N-1 cols at right (filled with 0s)
- We shift the filter at each vertical and horizontal shift to perform the correlation/convolution operation:

Correlation: $$(w \otimes f)(x,y) = \sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s,t) f(x+s, y+t)$$

Convolution: $$(w * f)(x,y) = \sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s,t) f(x-s, y-t)$$

2D Correlation
Convolution



Template Matching

Image *(f)*

Template *(w)*

- How do we locate the template in the image?

=> Minimize:

$$E[i,j] = \sum_{s=-a}^{a} \sum_{t=-b}^{b} [w(s,t) - f(i+s, j+t)]^2$$

Template Matching

$$E[i,j] = \sum_{s=-a}^{a} \sum_{t=-b}^{b} [w(s,t) - f(i+s, j+t)]^2$$

$$E[i,j] = \sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s,t)^2 + f(i+s, j+t)^2 - 2\, w(s,t)\, f(i+s, j+t)$$

Minimizing $E[i,j]$ => Maximizing

$$\sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s,t)\, f(i+s, j+t) = w \otimes f(i,j)$$

Filters as Templates

- Filters offer a natural mechanism for finding simple patterns.
- Filters respond most strongly to pattern elements that look like the filter.

Correlation as a Dot Product

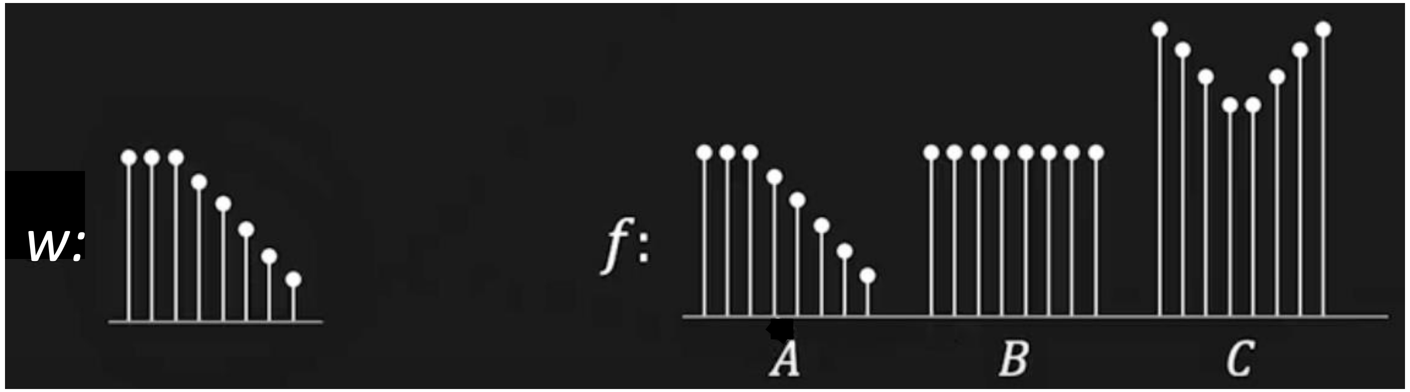
When performing a correlation (or a convolution):

- The response is obtained by associating image elements with filter kernel elements, multiplying the associated elements and summing
- It is the same process as a dot product.
- The dot product achieves its largest value when the vector representing the image is parallel to the vector representing the kernel.

$\Rightarrow$ A filter responds most strongly when it encounters an image pattern that looks like the filter.

$\Rightarrow$ But this dot product is a poor way to find patterns because the response might be large just because the image region is bright.

Correlation for Pattern Matching



$$R_{wf}[i,j] = w \otimes f(i,j) = \sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s,t)\, f(i+s, j+t)$$

$$R_{wf}(C) > R_{wf}(B) > R_{wf}(A)$$

However, we need $\boldsymbol{R_{wf}(A)}$ to be the maximum.

Normalized Correlation for Pattern Matching

<u>**Solution:**</u> ***Normalizing the correlation***

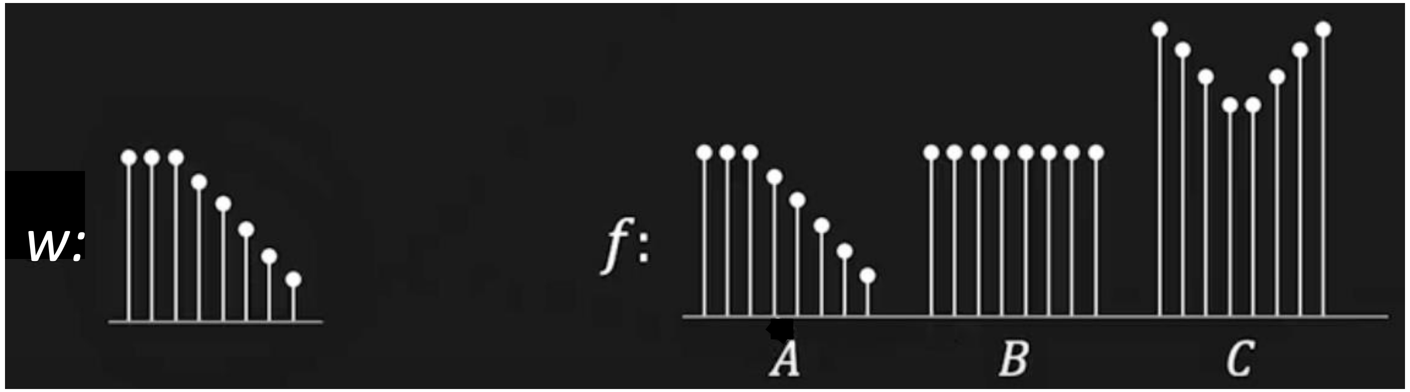
$$N_{wf}[i,j] = \frac{\sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s,t)\, f(i+s, j+t)}{\sqrt{\sum_s \sum_t w(s,t)^2} \cdot \sqrt{\sum_s \sum_t f(i+s, i+t)^2}}$$

Energy of the template

Energy of the image area covered by the template

$\Rightarrow$ This make the correlation insensitive to brightness.

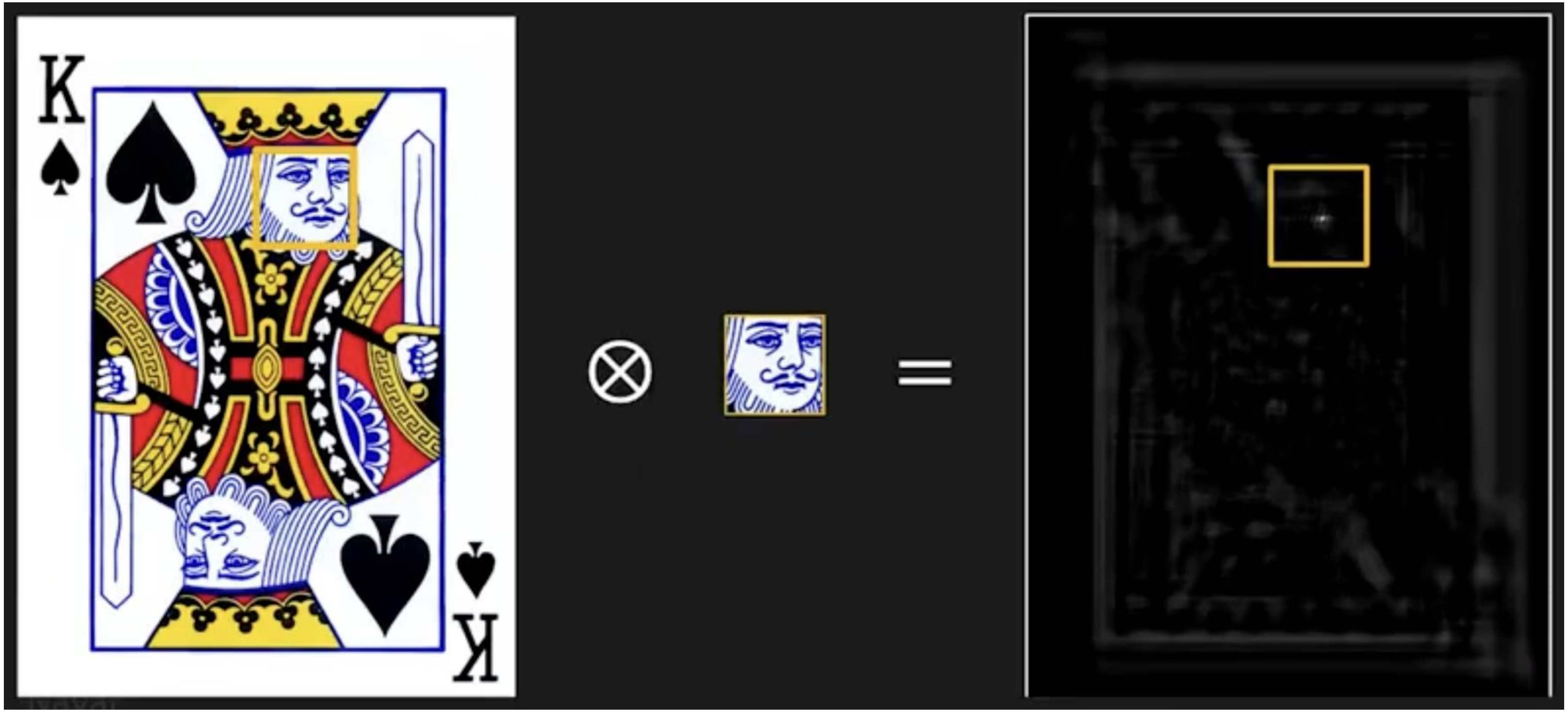
Normalized Correlation for Pattern Matching



$$N_{wf}[i,j] = \frac{\sum_{s=-a}^{a}\sum_{t=-b}^{b} w(s,t)\, f(i+s, j+t)}{\sqrt{\sum_s \sum_t w(s,t)^2} \cdot \sqrt{\sum_s \sum_t f(i+s, i+t)^2}}$$

$$N_{wf} A > N_{wf}(B) > N_{wf}(C)$$

Normalized Correlation for Pattern Matching



Normalized Correlation for Pattern Matching

$\otimes$

=