

Data Science

Lecture 9-1: Image Data Processing (Image Filtering)



Lecturer: Yen-Chia Hsu

Date: Mar 2025

This lecture covers basic concepts of image filtering and feature extraction.

There are many different types of tasks in [Computer Vision](#). This lecture will only cover the image/video classification, which gives only one label to an image/video.

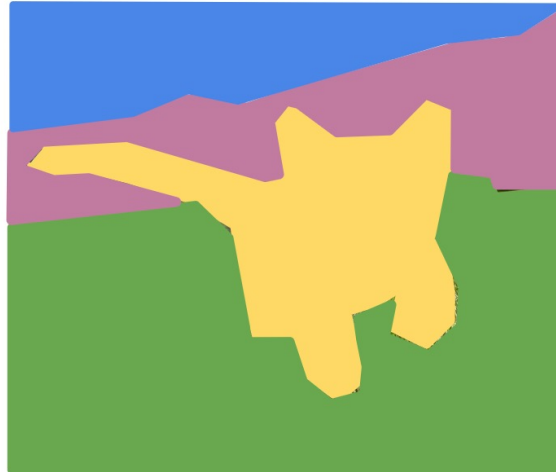
Classification



CAT

No spatial extent

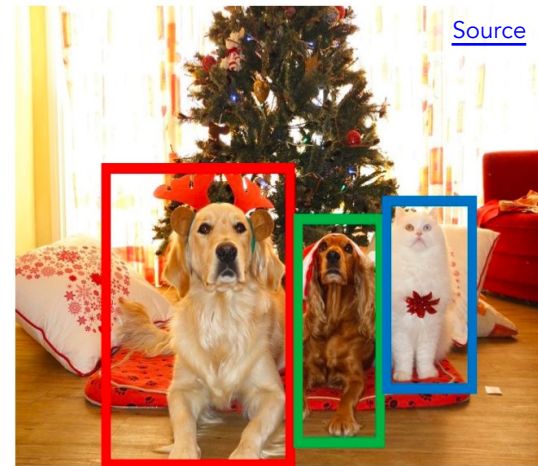
Semantic Segmentation



GRASS, CAT, TREE, SKY

No objects, just pixels

Object Detection



DOG, DOG, CAT

Multiple Object

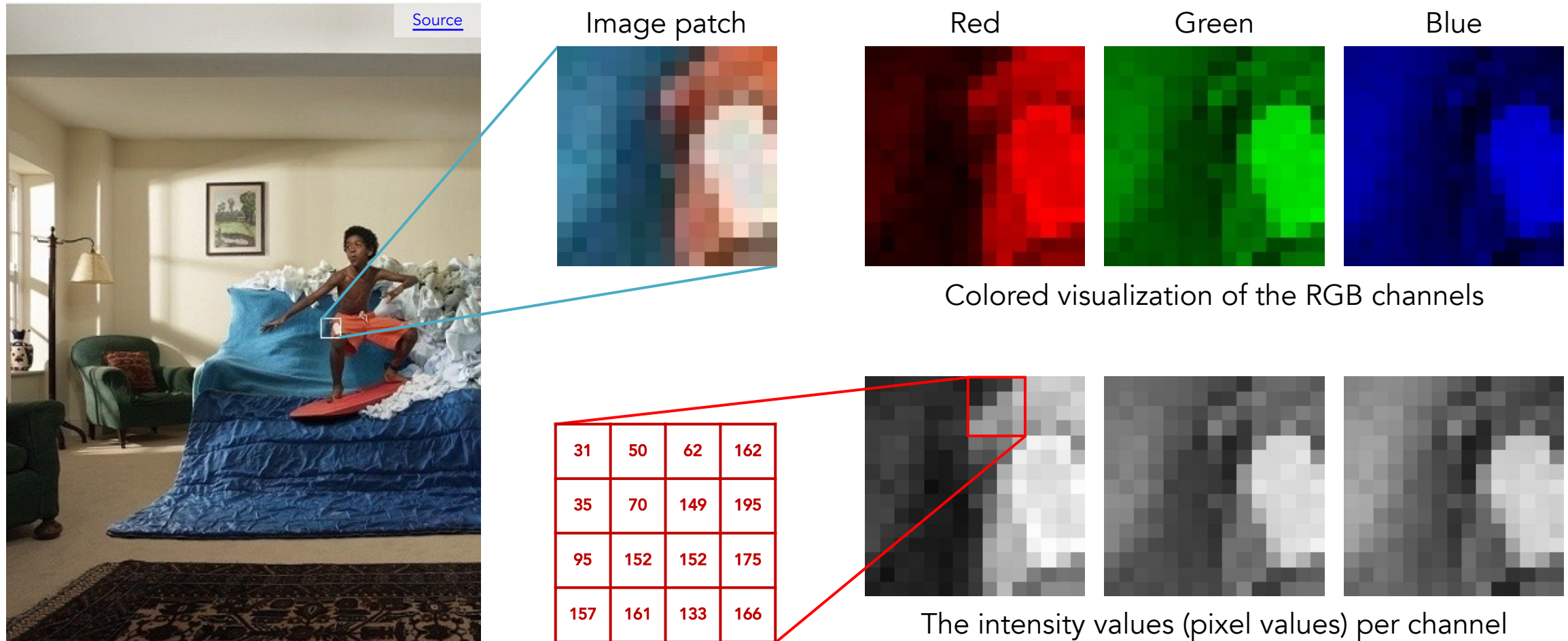
Instance Segmentation



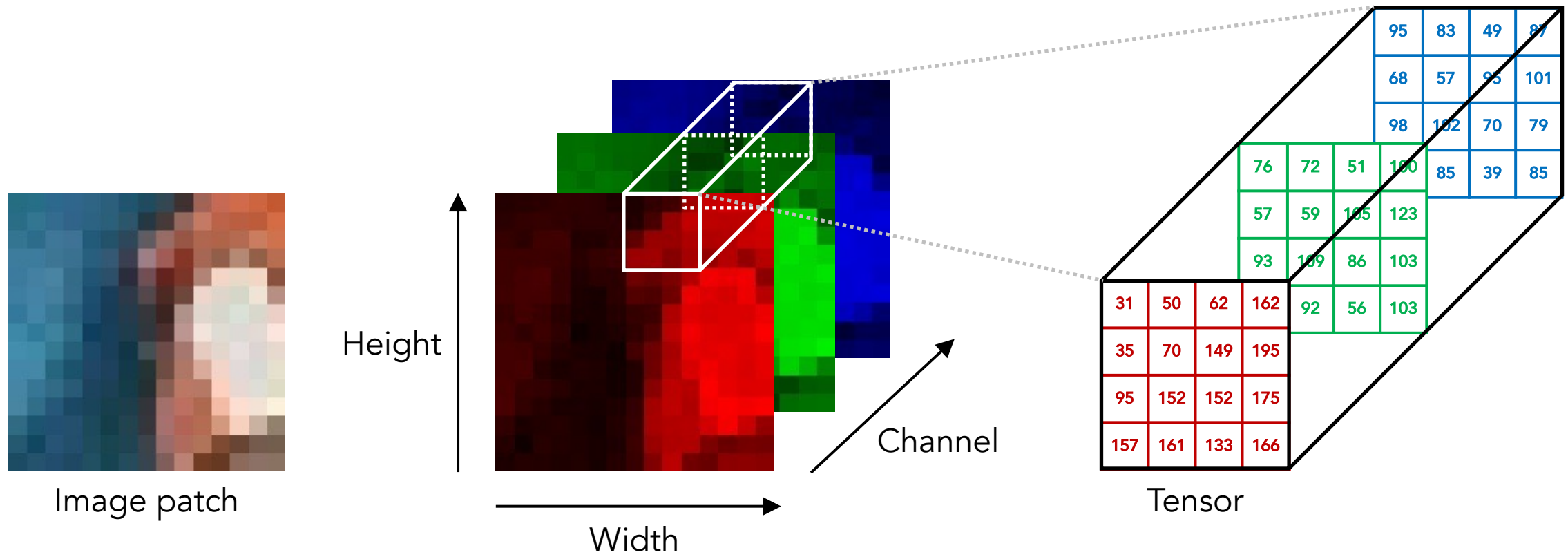
DOG, DOG, CAT

[This image is CC0 public domain](#)

People can see images directly, but computers read only numbers. Typically, computers store images as **pixels** with RGB channels with values ranging from 0 to 255.



Images can be represented as a 3D tensor (width*height*channels). For RGB images, we have 3 channels corresponding to the pixel intensity of red, green, and blue colors.

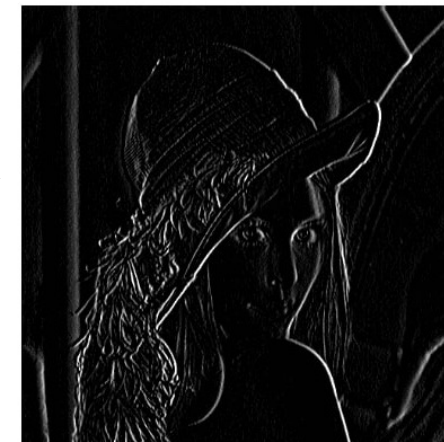
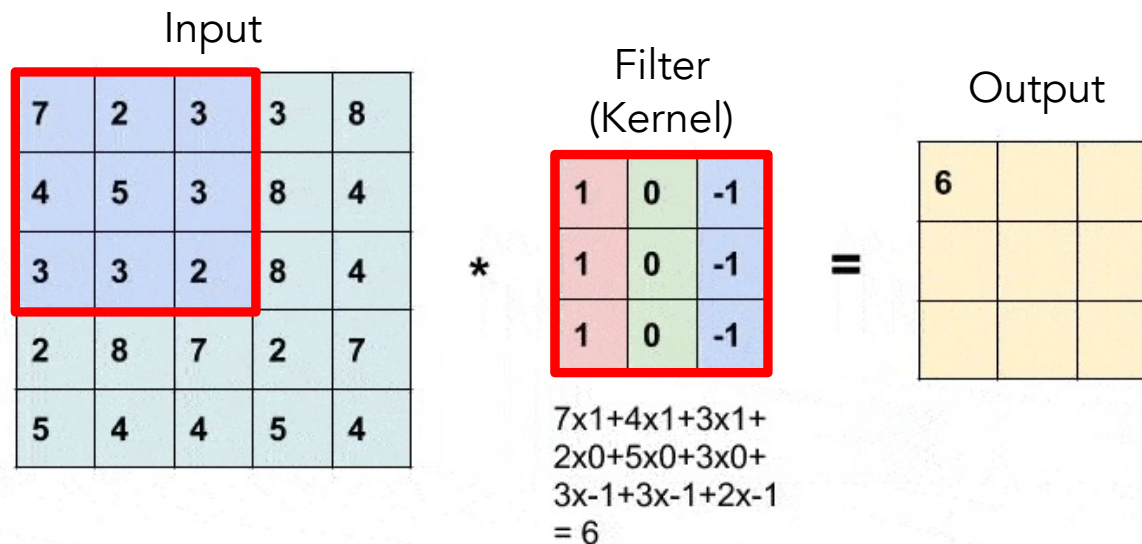


Before the deep learning era, Computer Vision used hand-crafted features. Researchers developed different image filters/kernels to extract features using **convolution**.



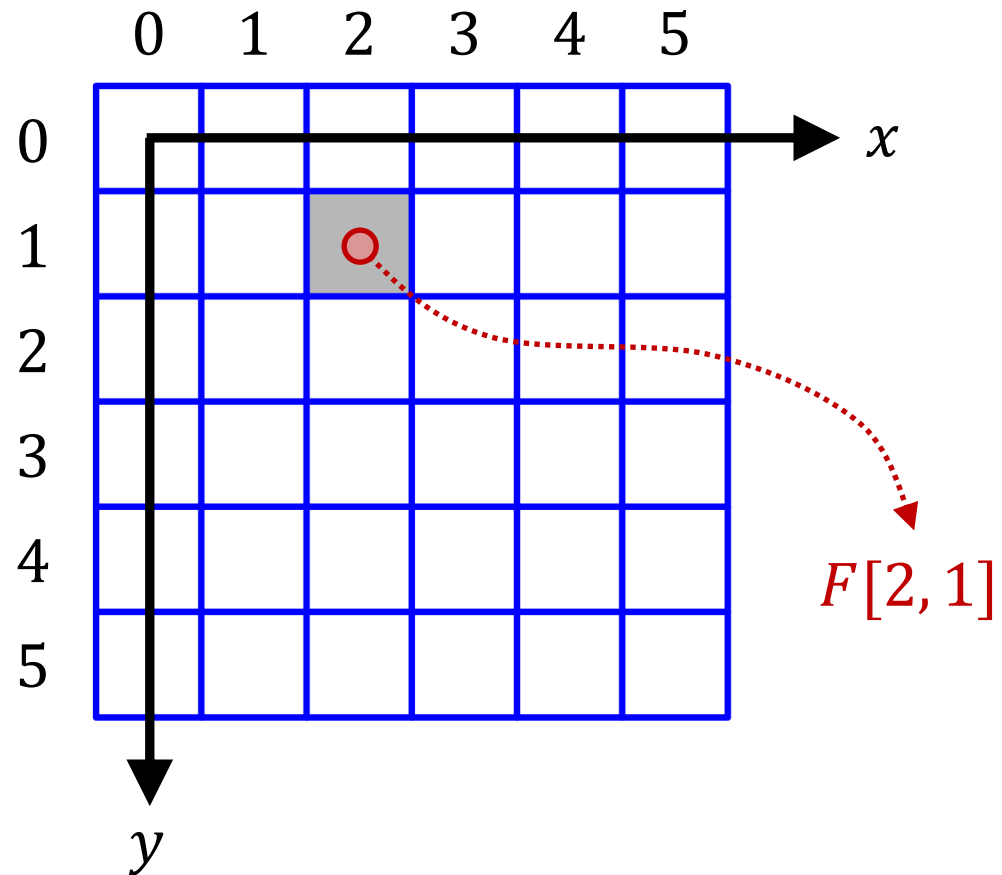
Input image

Discrete convolution: sum of the element-wise multiplication



Output image

About image coordinates, the origin is at the top-left pixel. Notation $F[x, y]$ means the value of the center of the pixel for image F at location $[x, y]$ in the 2D array.

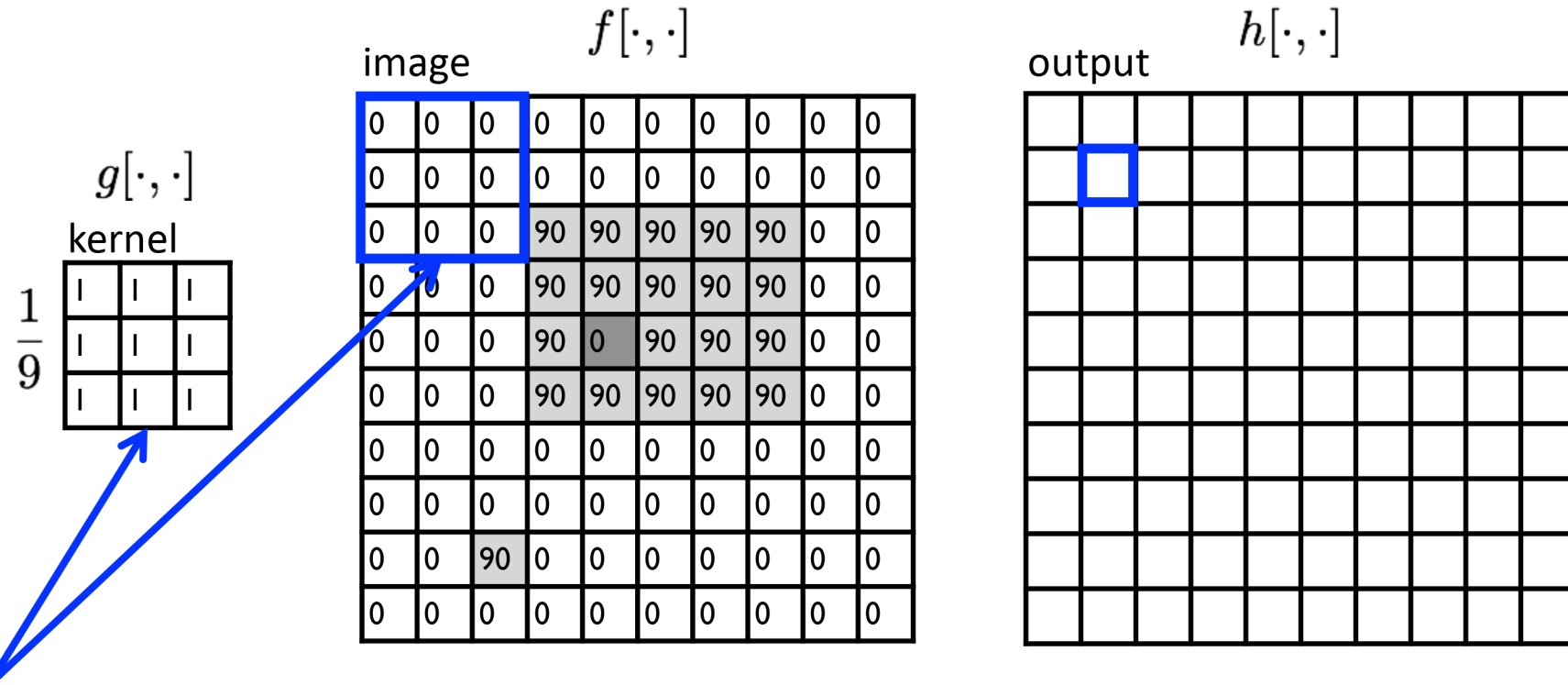


$$F[x, y]$$

F is an array of numbers (pixels)

x, y are integers (column/row indices)

Let's run the box filter

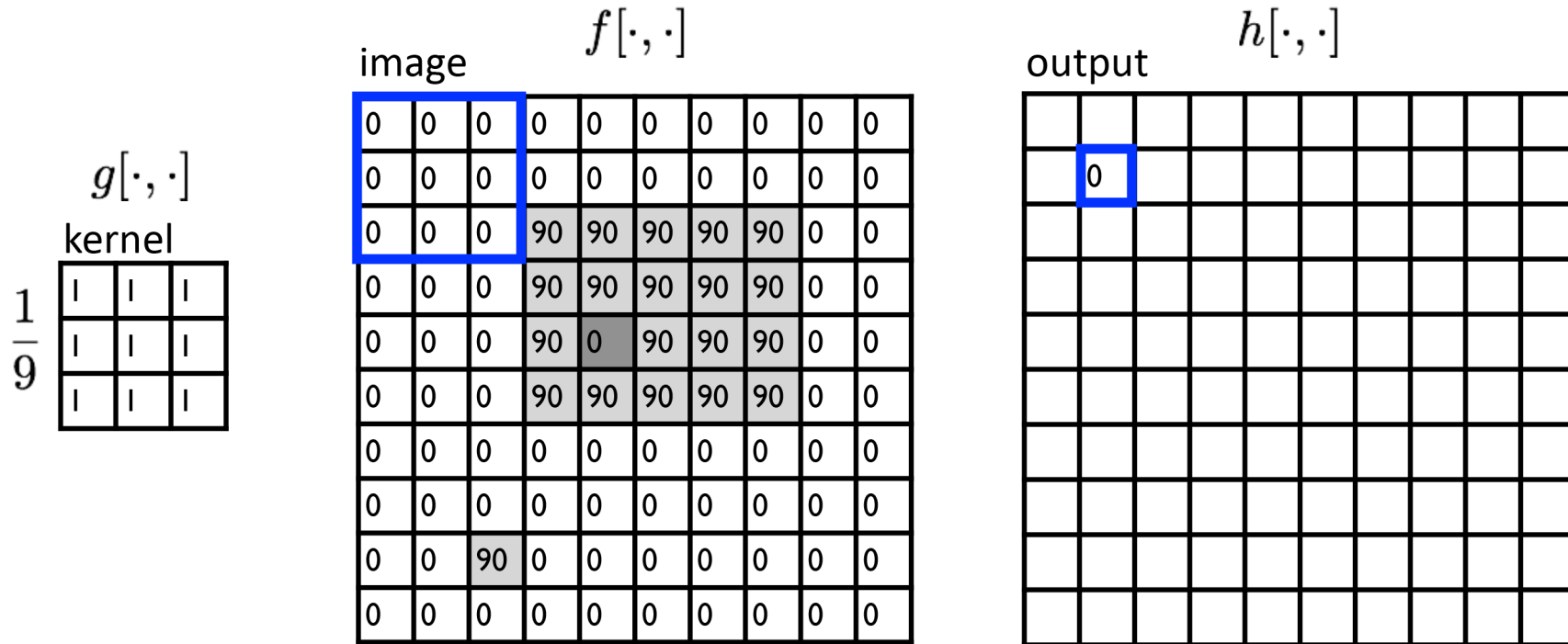


note that we assume that the kernel coordinates are centered

$$h[m, n] = \sum_{k, l} g[k, l] f[m + k, n + l]$$

output
filter
image (signal)

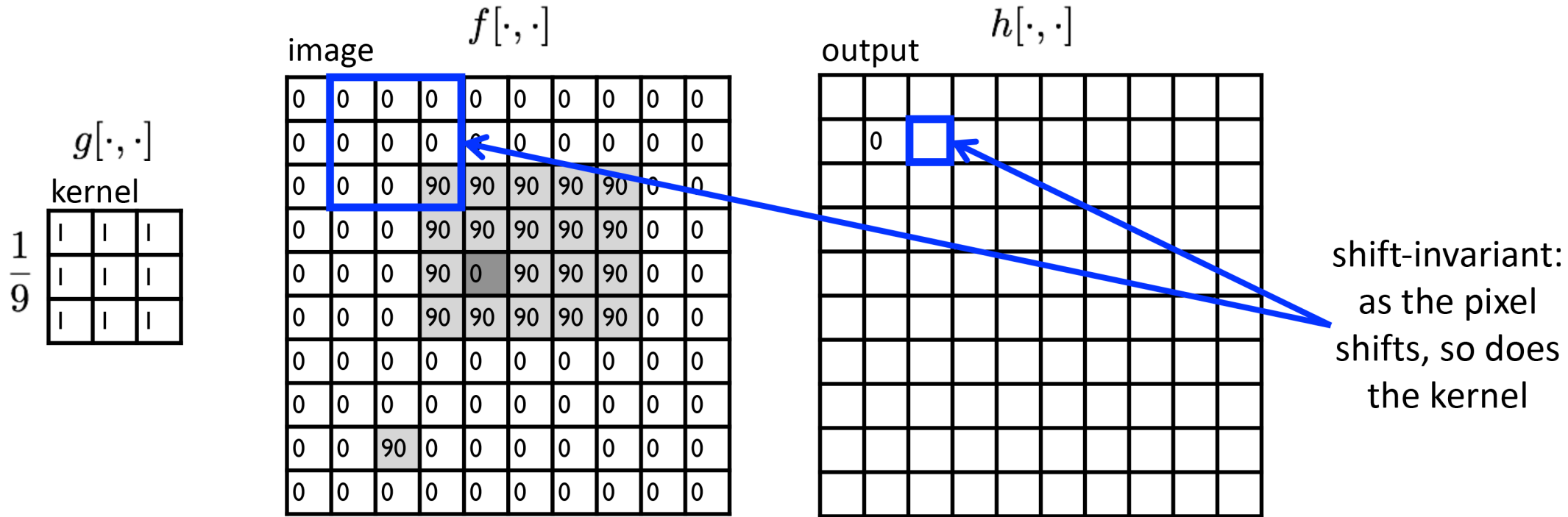
Let's run the box filter



$$h[m, n] = \sum_{k, l} g[k, l] f[m + k, n + l]$$

output
 k, l
filter
image (signal)

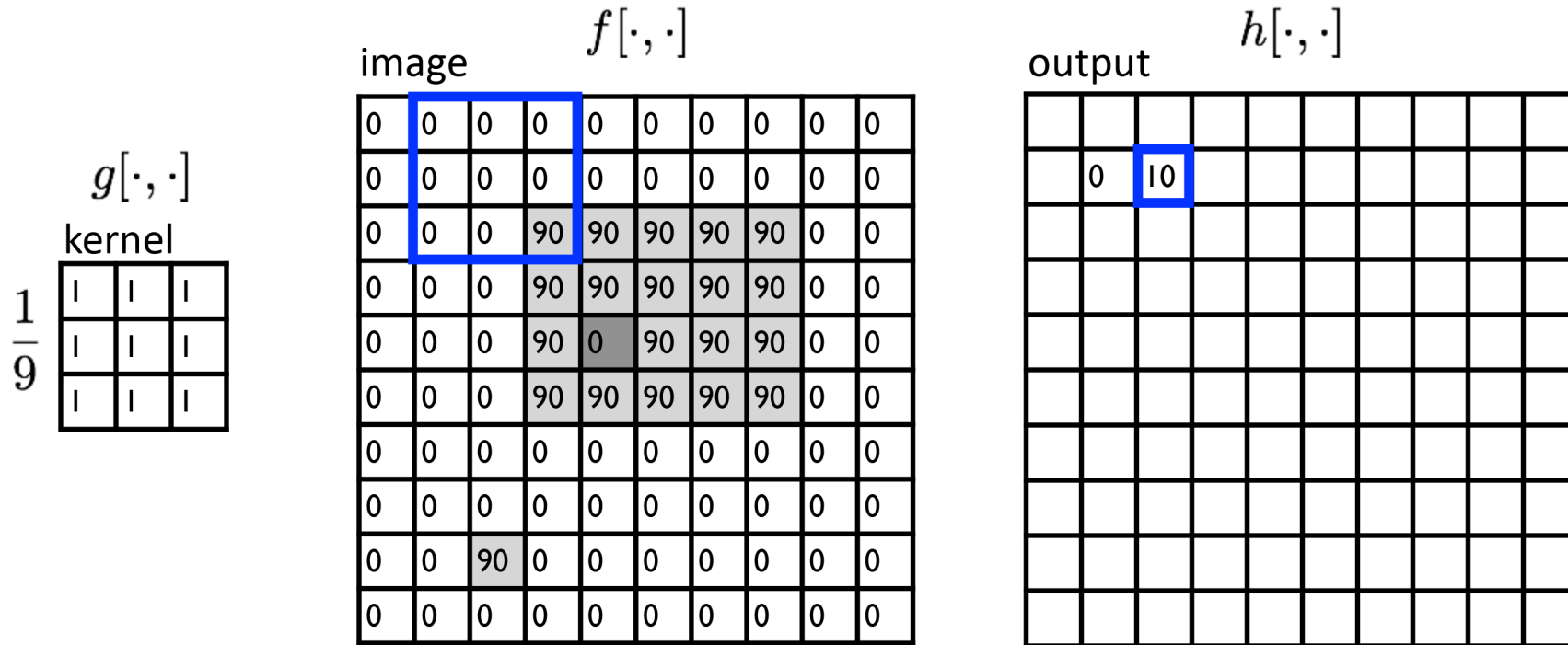
Let's run the box filter



$$h[m, n] = \sum_{k, l} g[k, l] f[m + k, n + l]$$

output
 k, l
filter
image (signal)

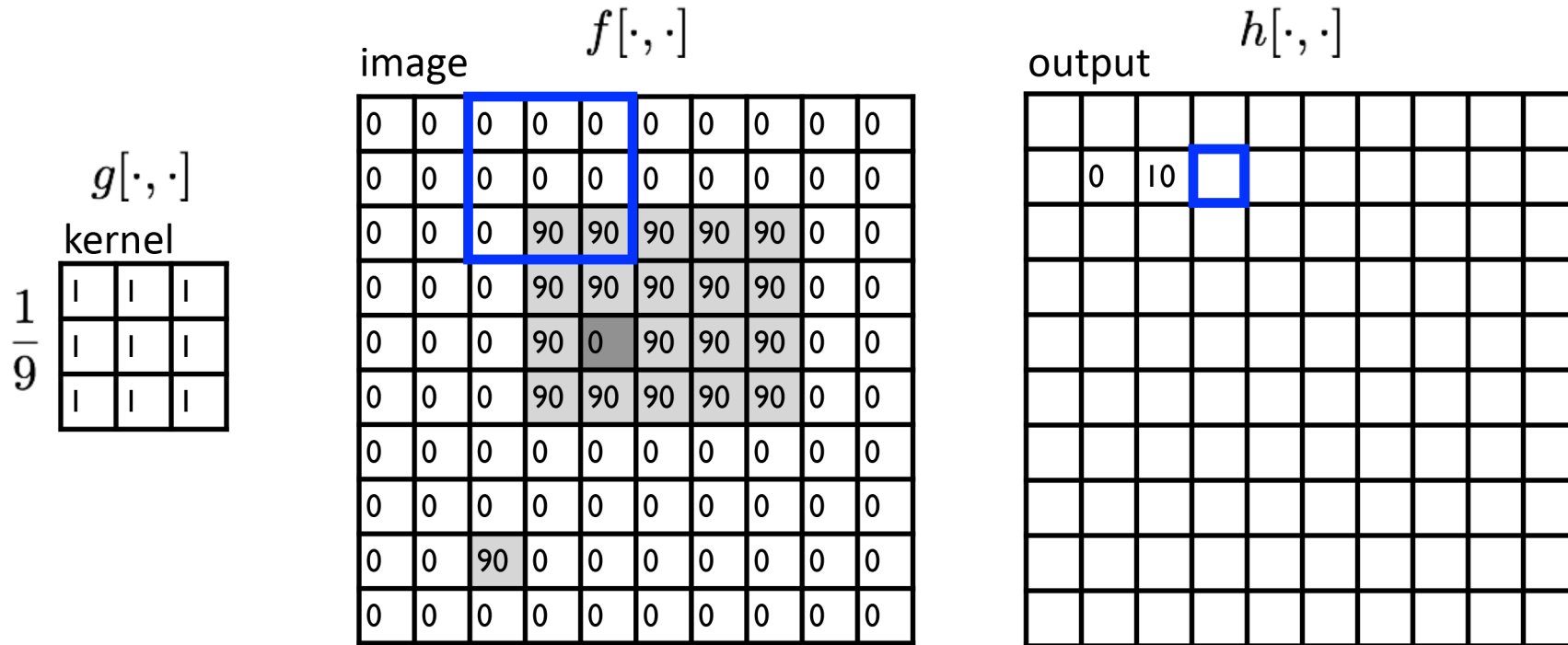
Let's run the box filter



$$h[m, n] = \sum_{k, l} g[k, l] f[m + k, n + l]$$

output
 k, l
filter
image (signal)

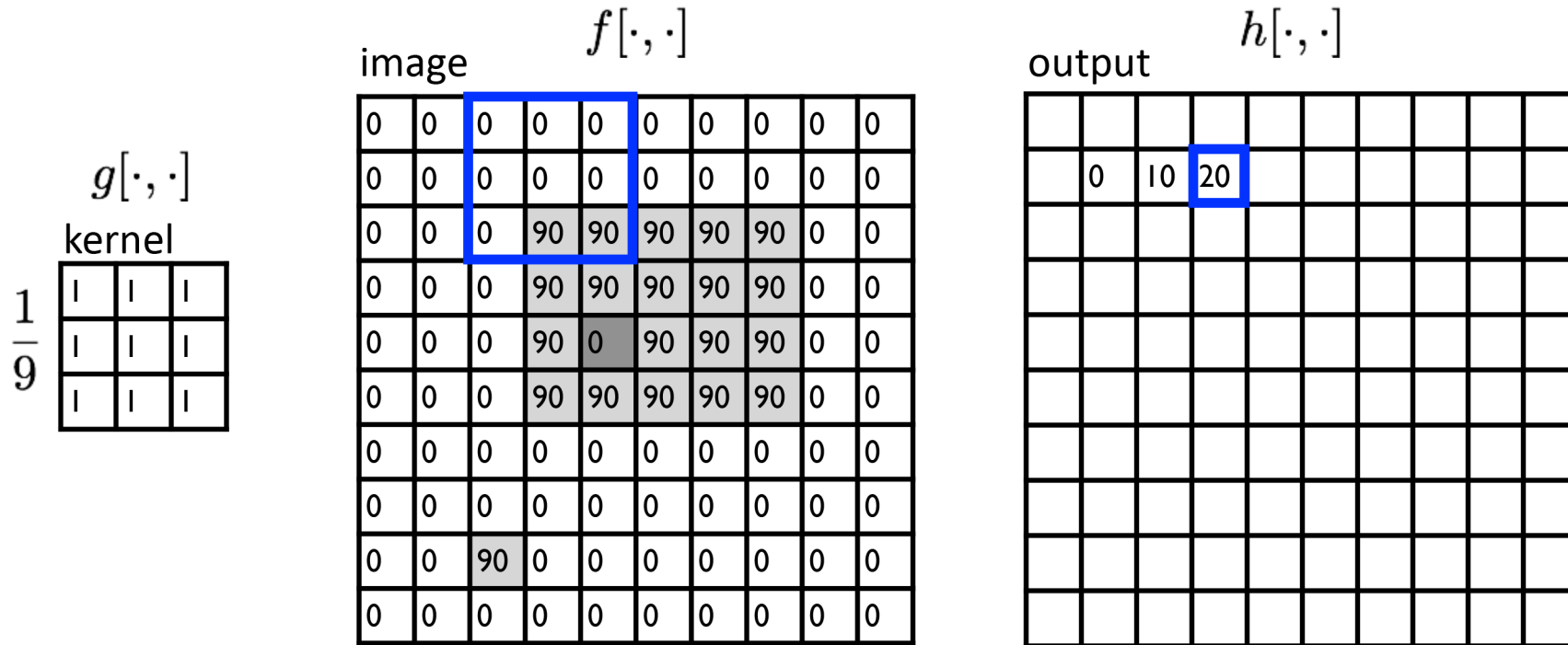
Let's run the box filter



$$h[m, n] = \sum_{k, l} g[k, l] f[m + k, n + l]$$

output
 k, l
filter
image (signal)

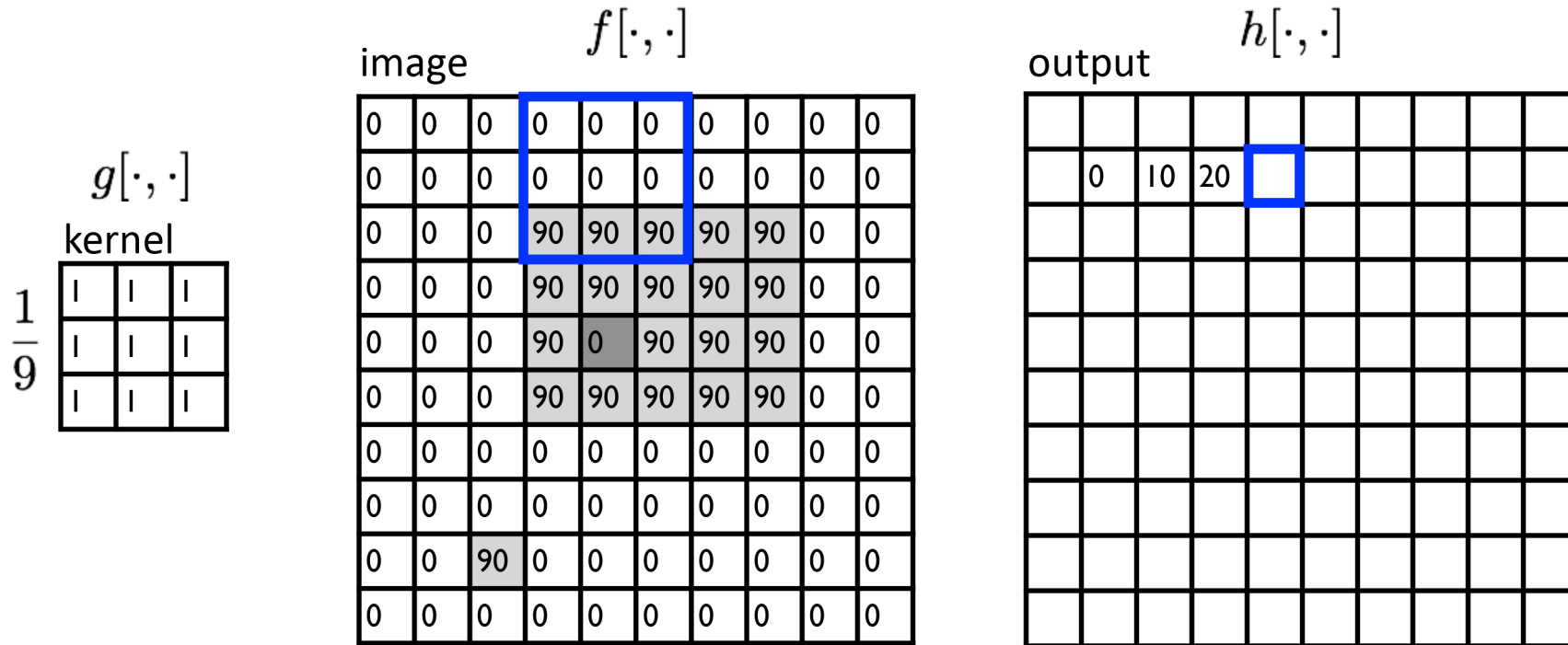
Let's run the box filter



$$h[m, n] = \sum_{k, l} g[k, l] f[m + k, n + l]$$

output
 k, l
filter
image (signal)

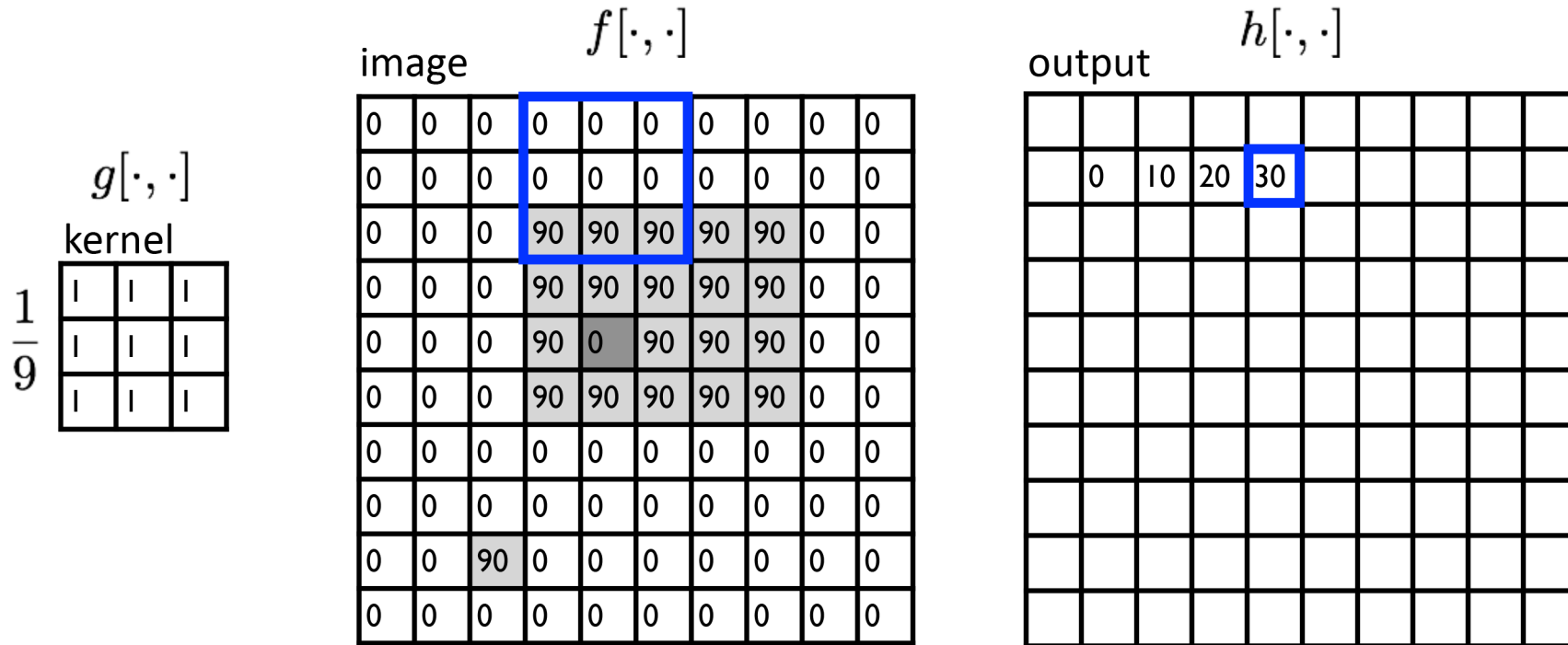
Let's run the box filter



$$h[m, n] = \sum_{k, l} g[k, l] f[m + k, n + l]$$

output
 k, l
filter
image (signal)

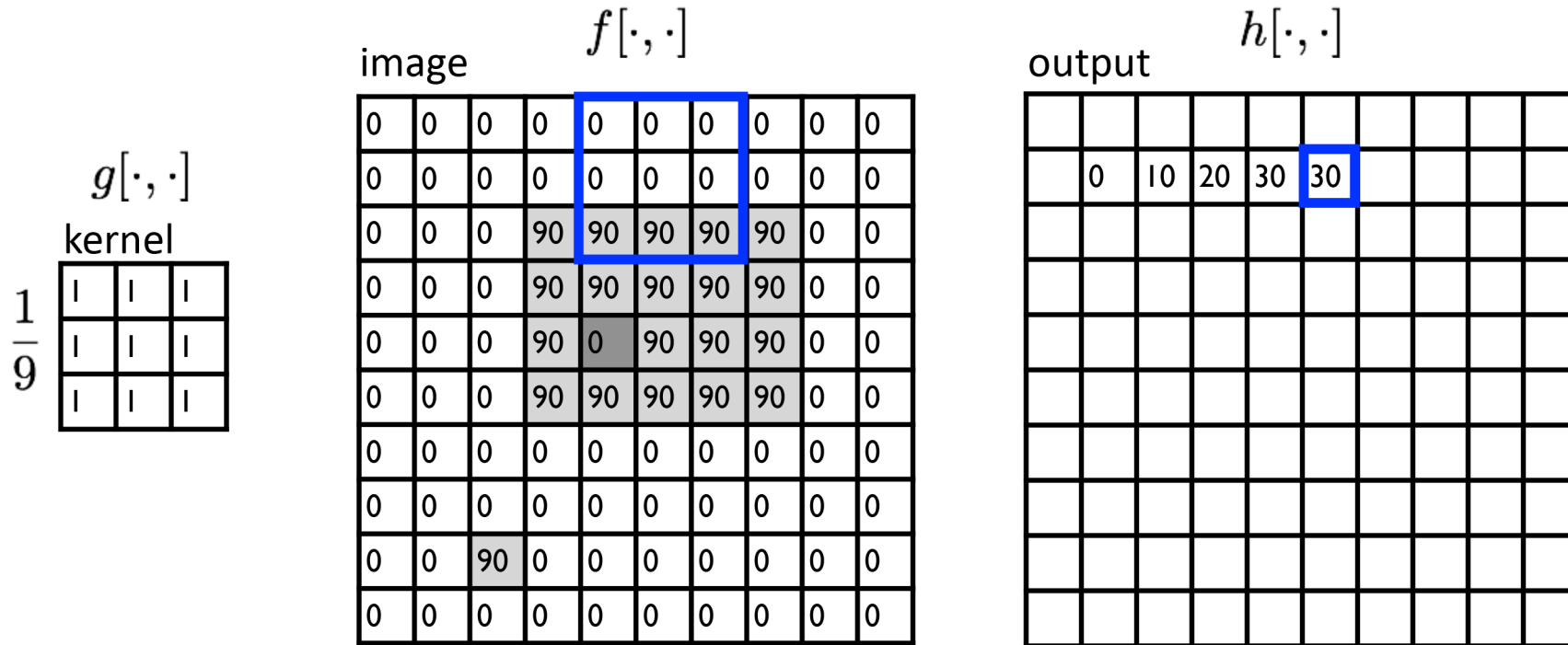
Let's run the box filter



$$h[m, n] = \sum_{k, l} g[k, l] f[m + k, n + l]$$

output
 k, l
filter
image (signal)

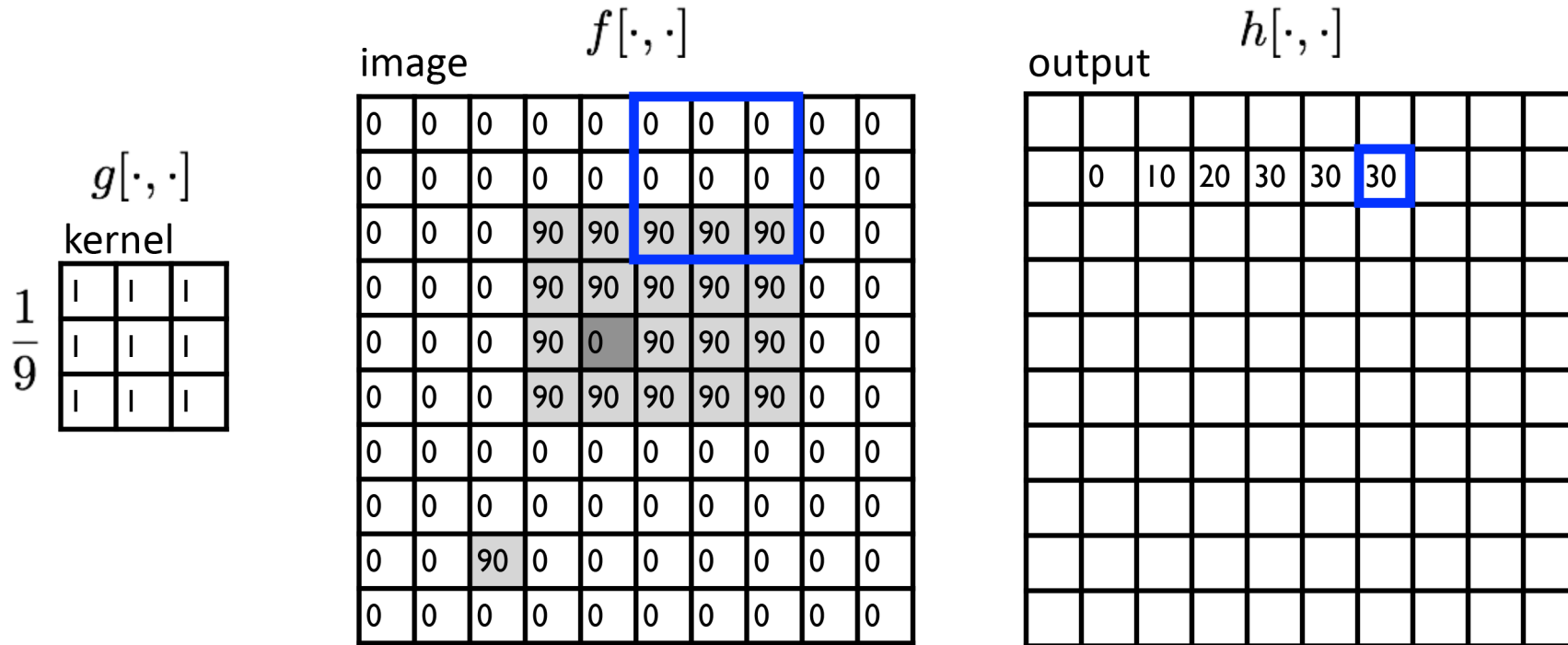
Let's run the box filter



$$h[m, n] = \sum_{k, l} g[k, l] f[m + k, n + l]$$

output
 k, l
filter
image (signal)

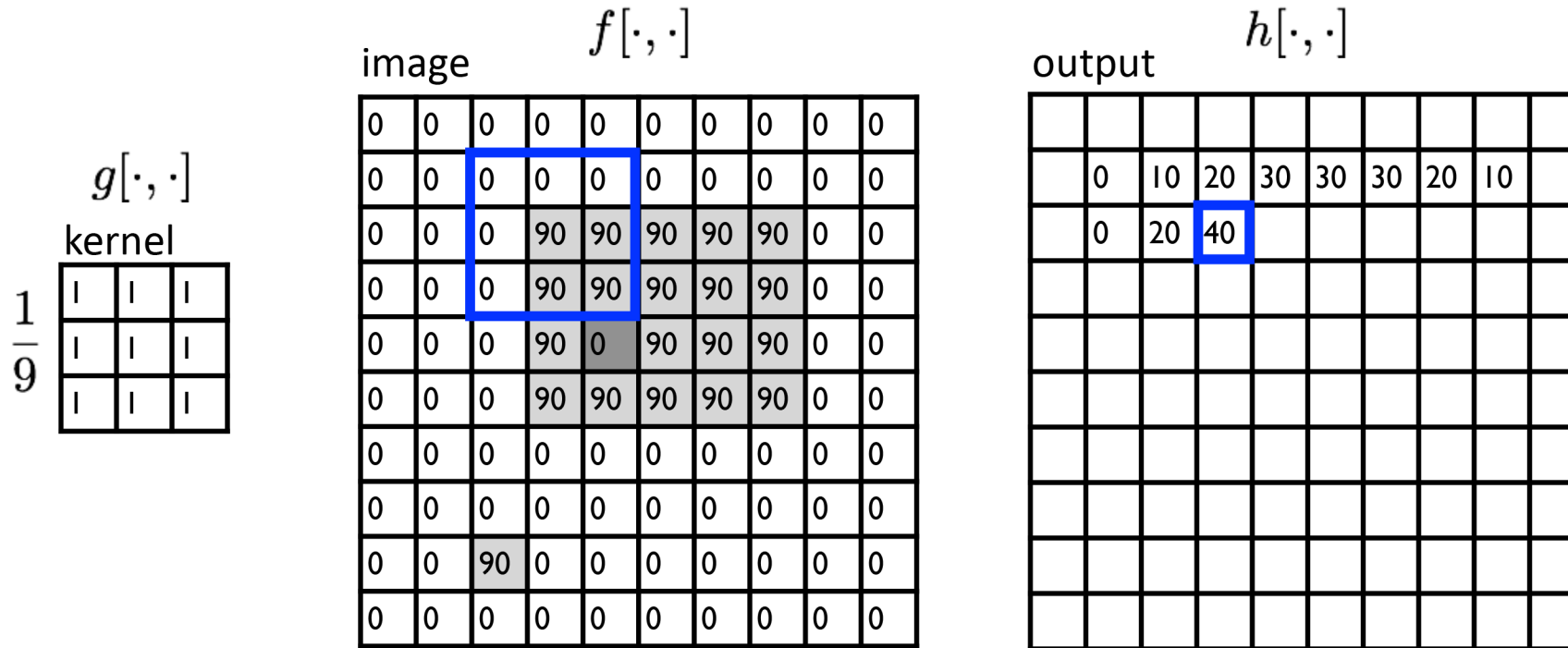
Let's run the box filter



$$h[m, n] = \sum_{k, l} g[k, l] f[m + k, n + l]$$

output
 k, l
filter
image (signal)

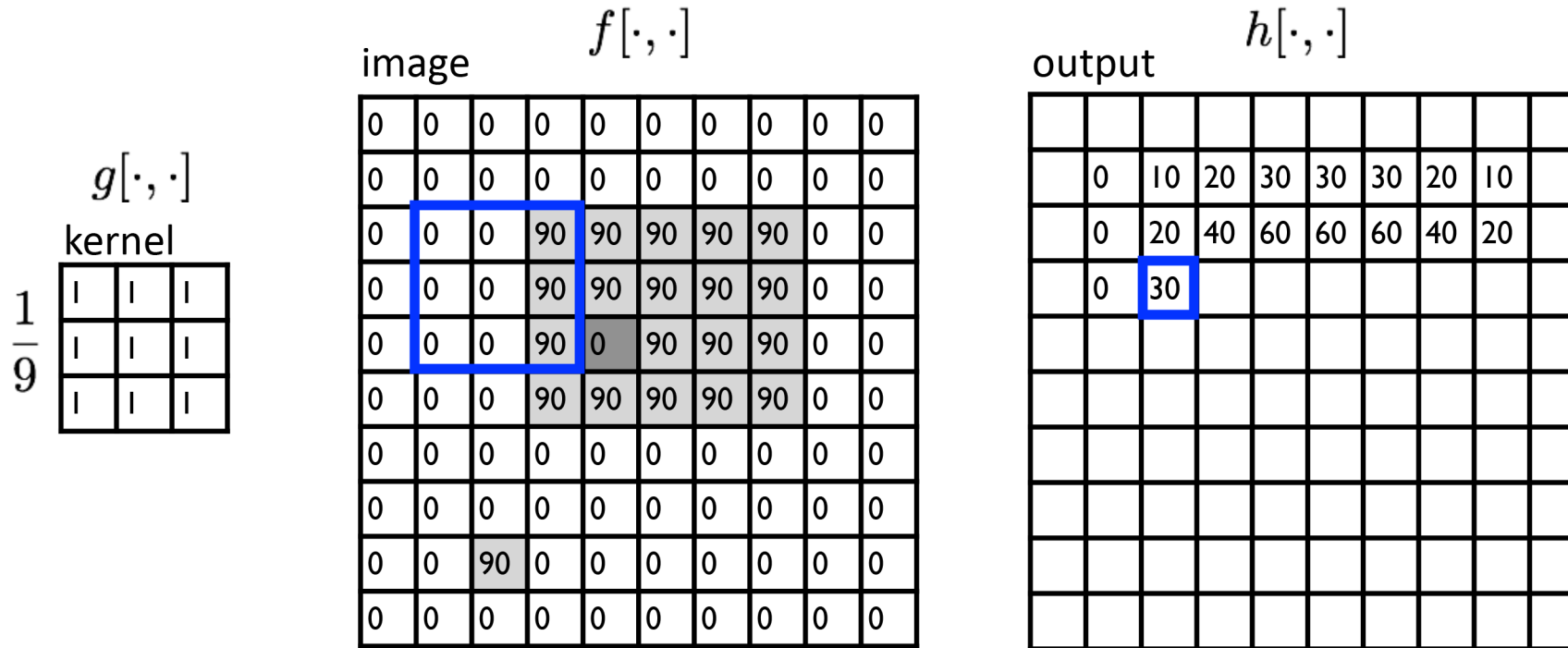
Let's run the box filter



$$h[m, n] = \sum_{k, l} g[k, l] f[m + k, n + l]$$

output
 k, l
filter
image (signal)

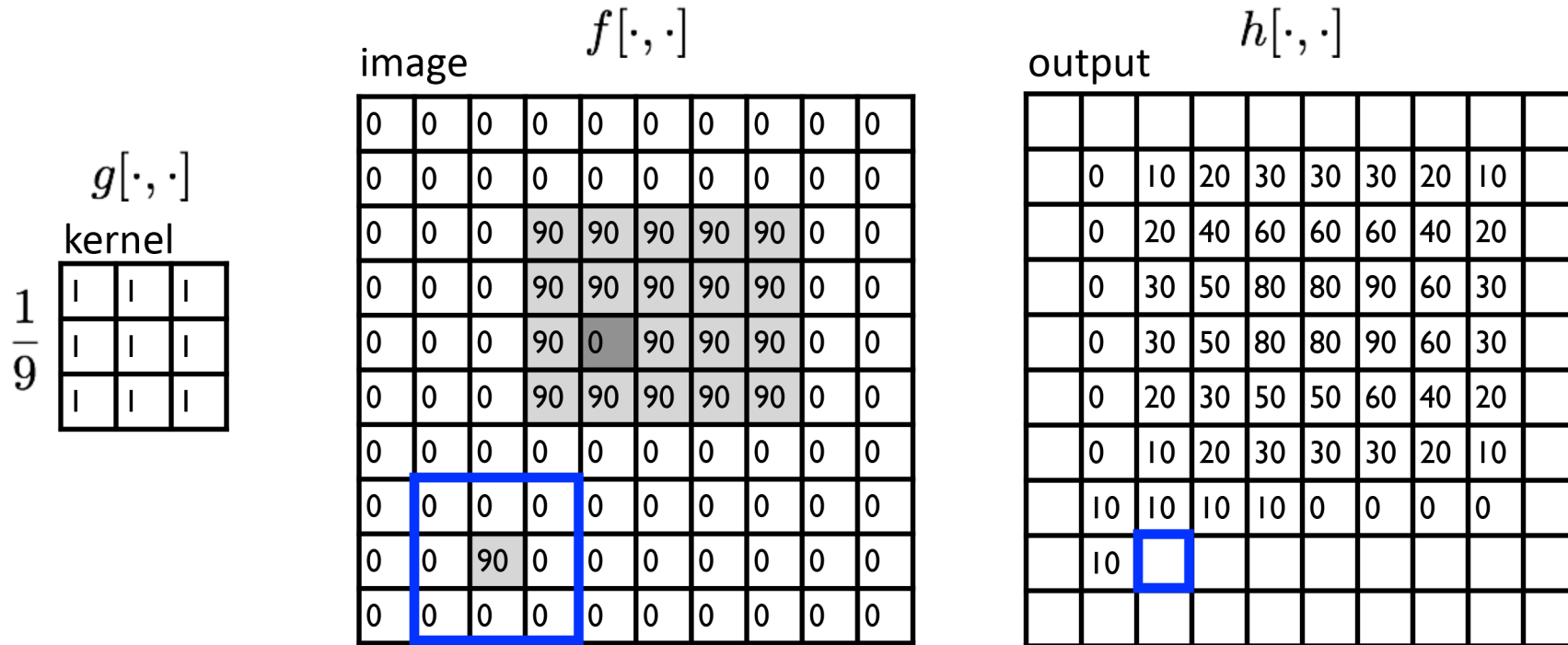
Let's run the box filter



$$h[m, n] = \sum_{k, l} g[k, l] f[m + k, n + l]$$

output
 k, l filter
image (signal)

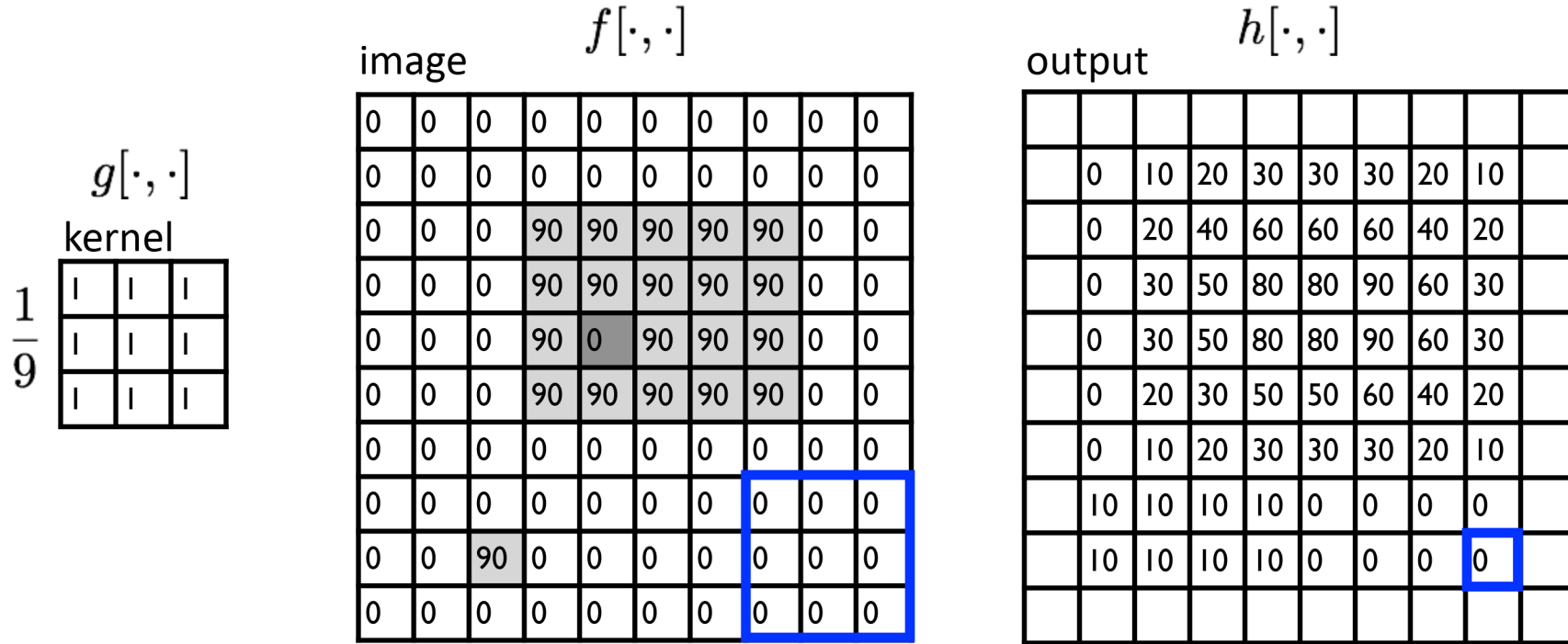
Let's run the box filter



$$h[m, n] = \sum_{k, l} g[k, l] f[m + k, n + l]$$

output
 k, l filter
image (signal)

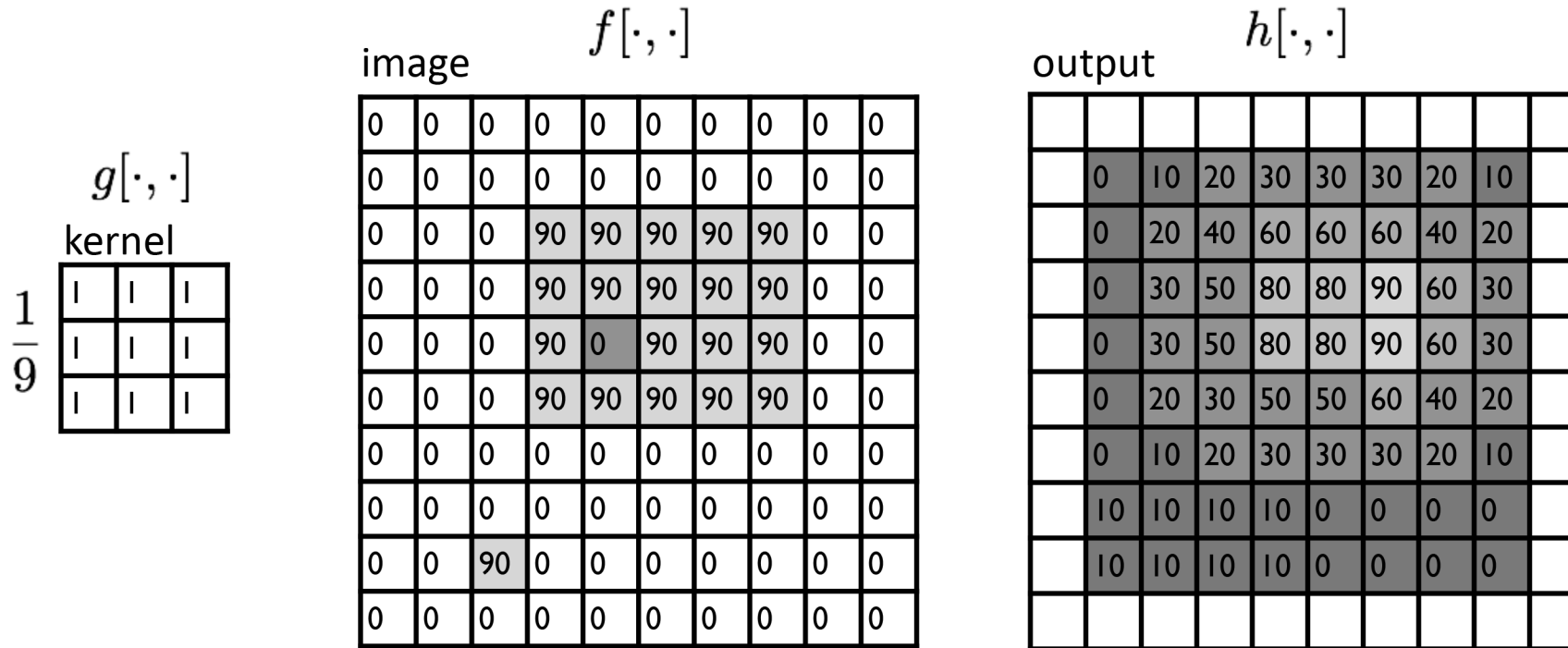
Let's run the box filter



$$h[m, n] = \sum_{k, l} g[k, l] f[m + k, n + l]$$

output
 k, l
filter
image (signal)

... and the result is



$$h[m, n] = \sum_{k, l} g[k, l] f[m + k, n + l]$$

output
 k, l filter
image (signal)

Exercise 9.1: Compute the result of the following convolution operations.

0	0	0	0
0	1	2	0
0	3	4	0
0	0	0	0

Image

$$\begin{matrix} * \\ \begin{matrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} \\ =? \end{matrix}$$

Kernel A

Hint:

	3

$$\begin{matrix} * \\ \begin{matrix} 1 & -1 \\ 1 & -1 \end{matrix} \\ =? \end{matrix}$$

Kernel B

Hint:

		4

What will be the result of the following convolution operation?



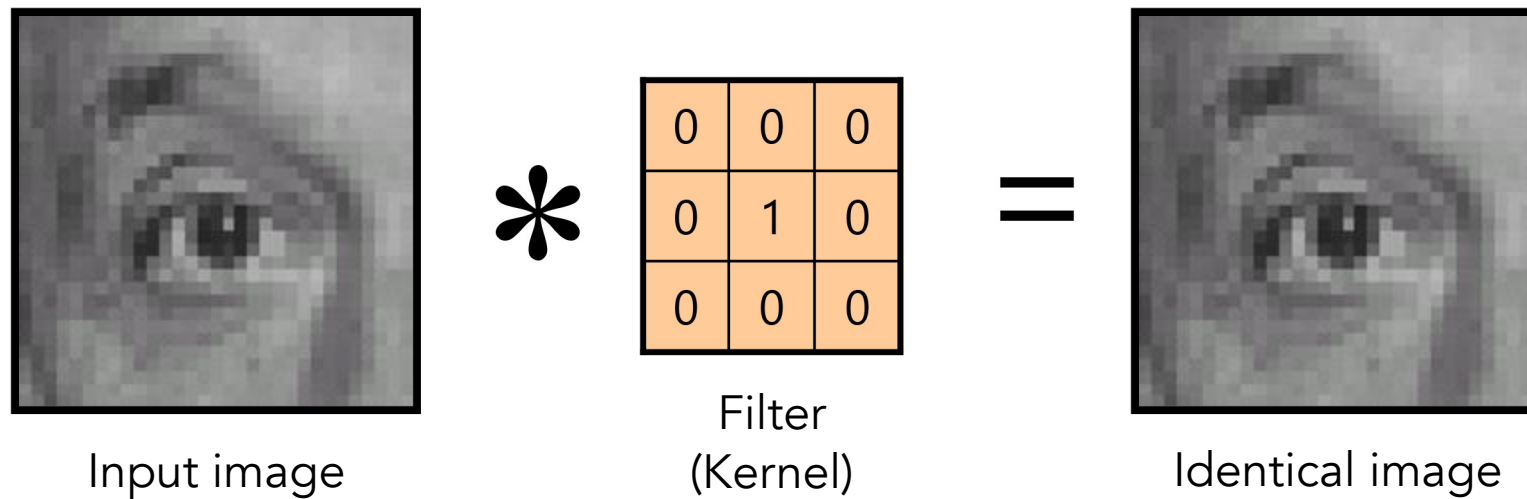
Input image



0	0	0
0	1	0
0	0	0

Filter
(Kernel)

The following convolution operation produces an identical image.



What will be the result of the following convolution operation?



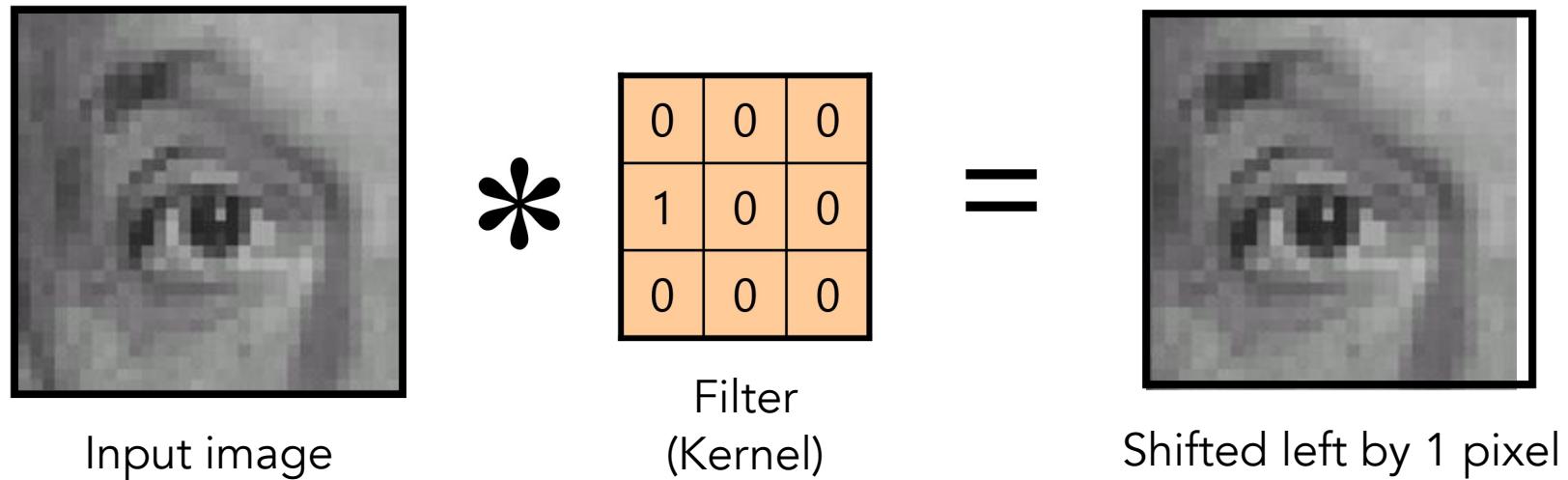
Input image



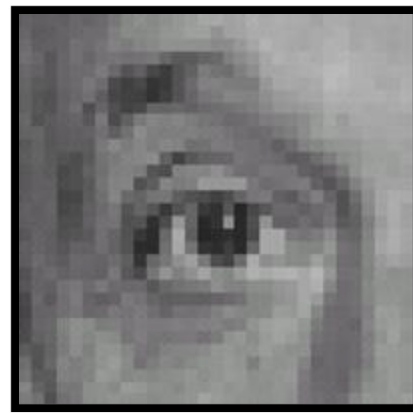
0	0	0
1	0	0
0	0	0

Filter
(Kernel)

The following convolution operation shifts the image left by one pixel.



What will be the result of the following convolution operation?



Input image

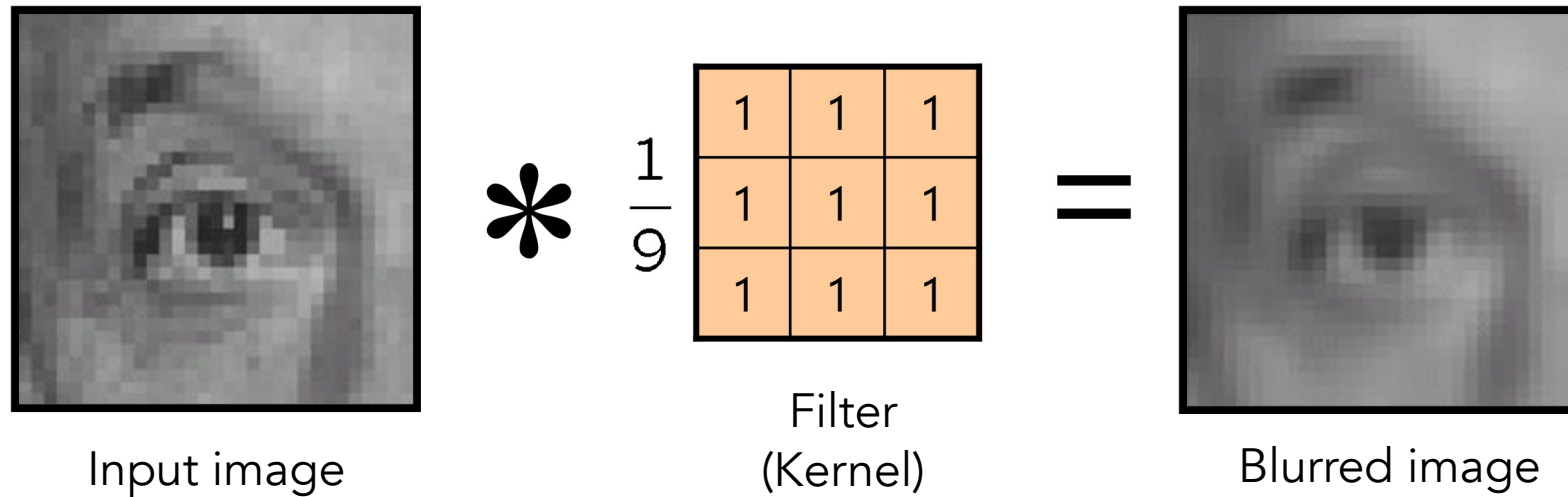


$\frac{1}{9}$

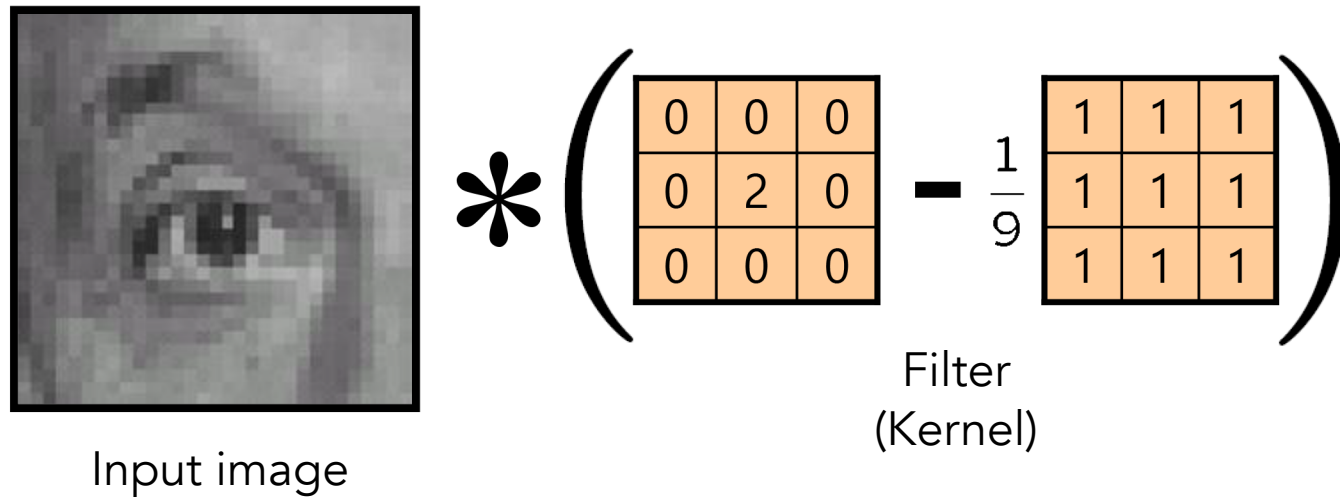
1	1	1
1	1	1
1	1	1

Filter
(Kernel)

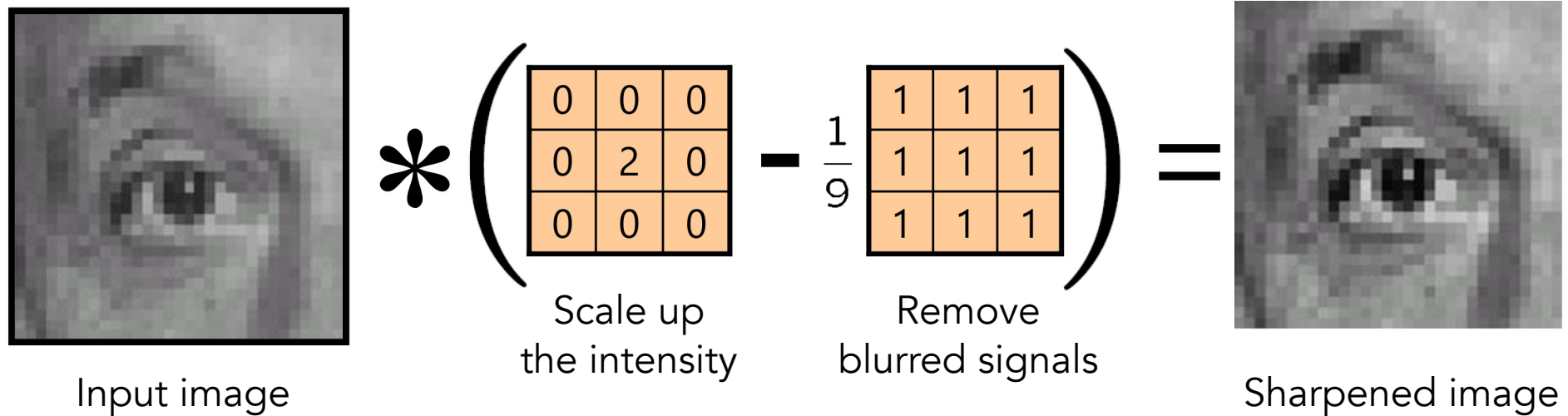
The following convolution operation blurs the image.



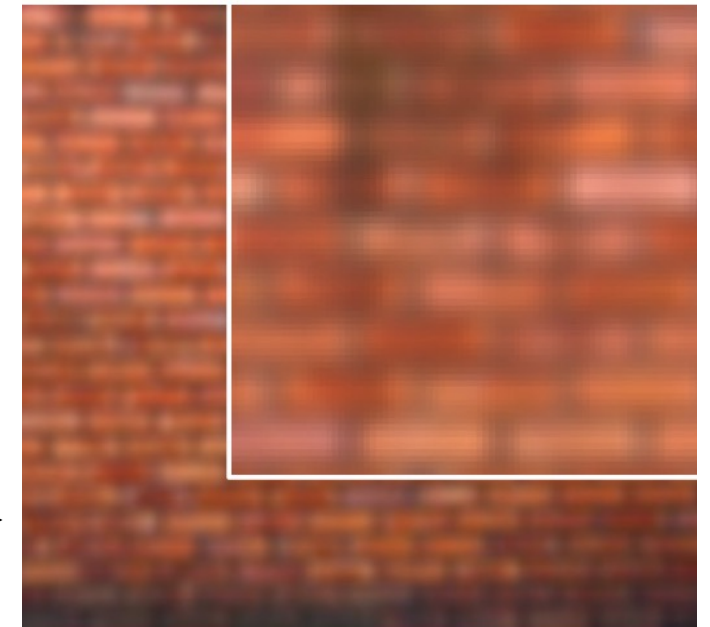
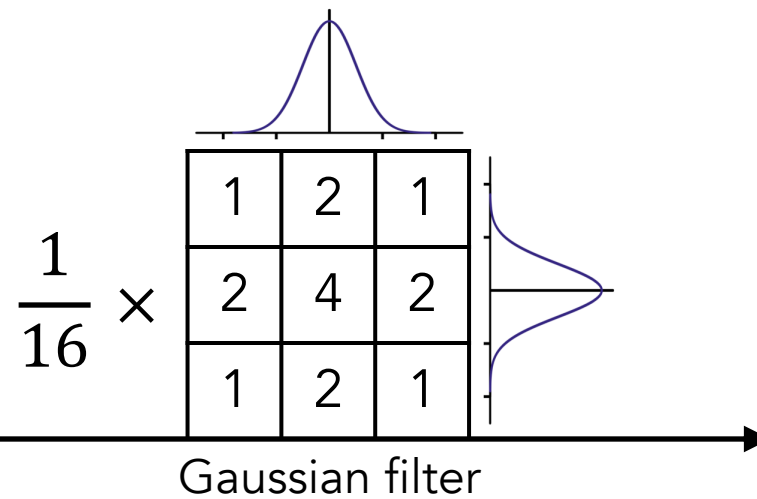
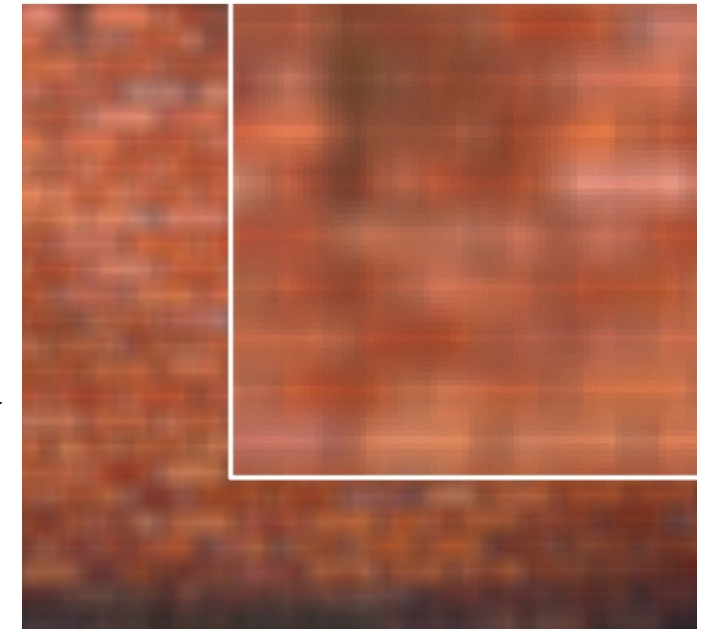
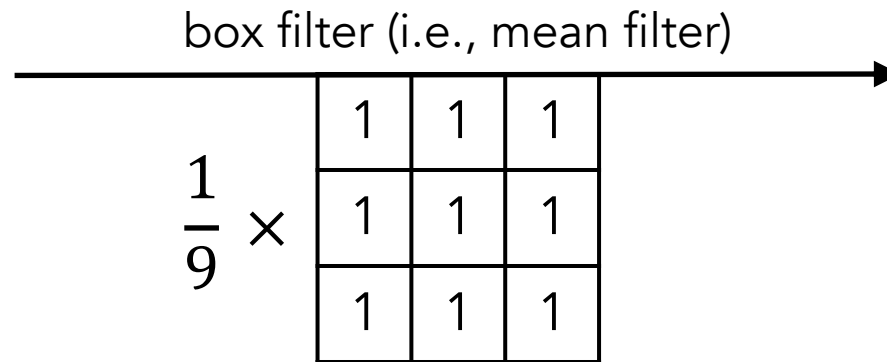
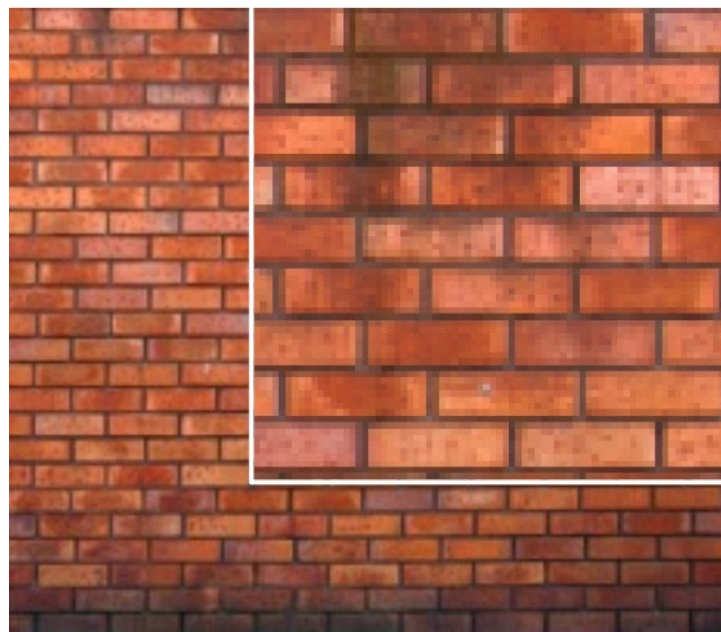
What will be the result of the following convolution operation?



The following convolution operation **sharpens the image**.



We can use image filters to **blur images**, such as using the box filter and the Gaussian filter (2D Gaussian distribution).



We can use image filters to **detect edges**, such as using the Sobel filter to compute the derivative of signals.

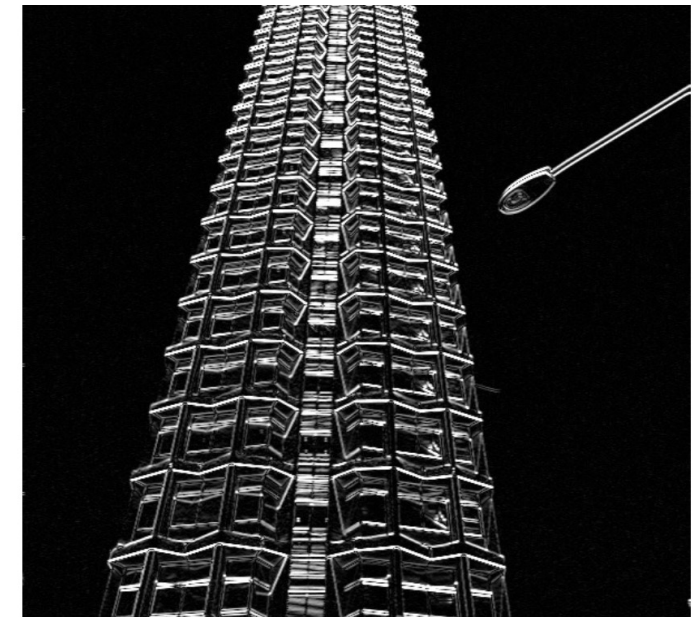
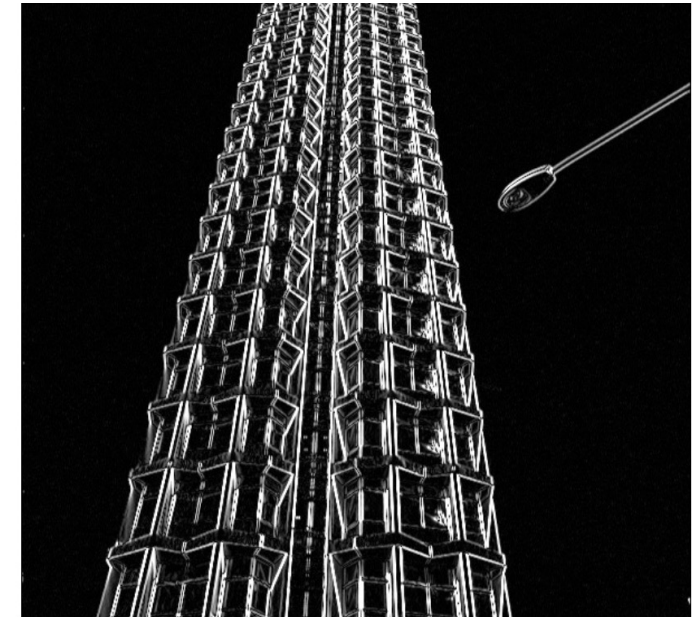


Horizontal Sobel filter

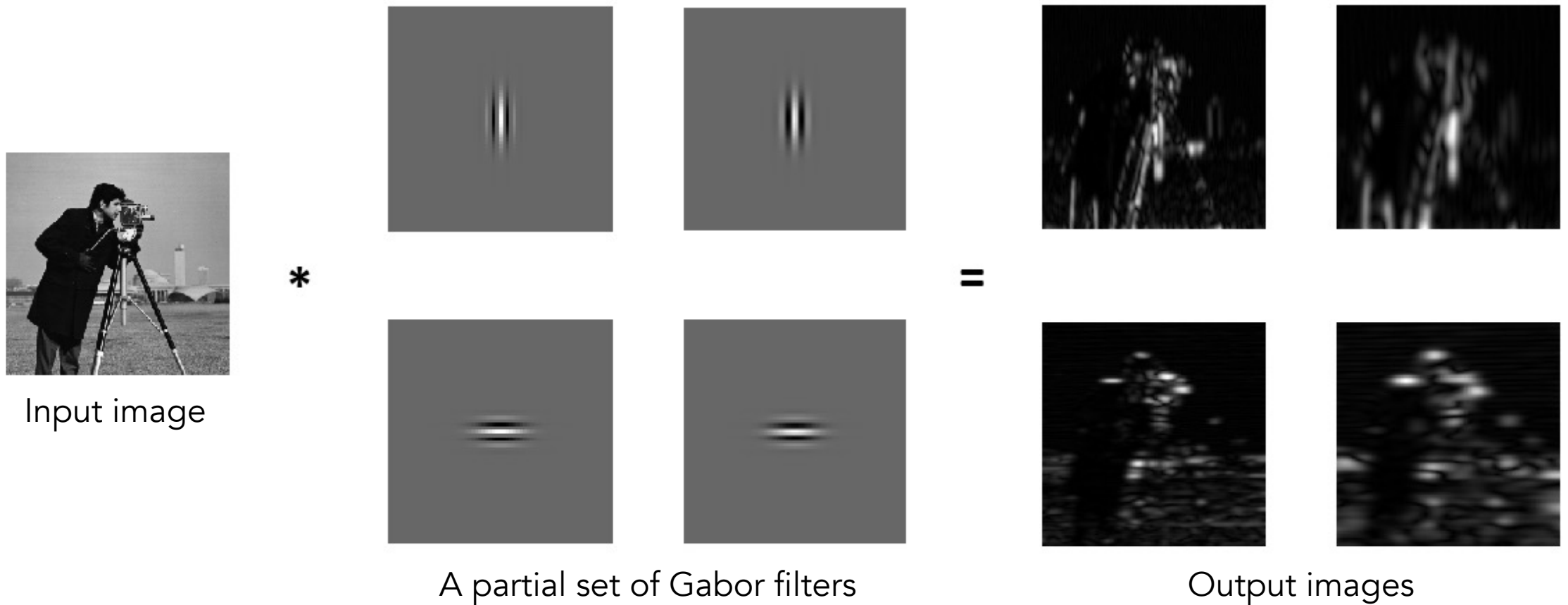
1	0	-1
2	0	-2
1	0	-1

1	2	1
0	0	0
-1	-2	-1

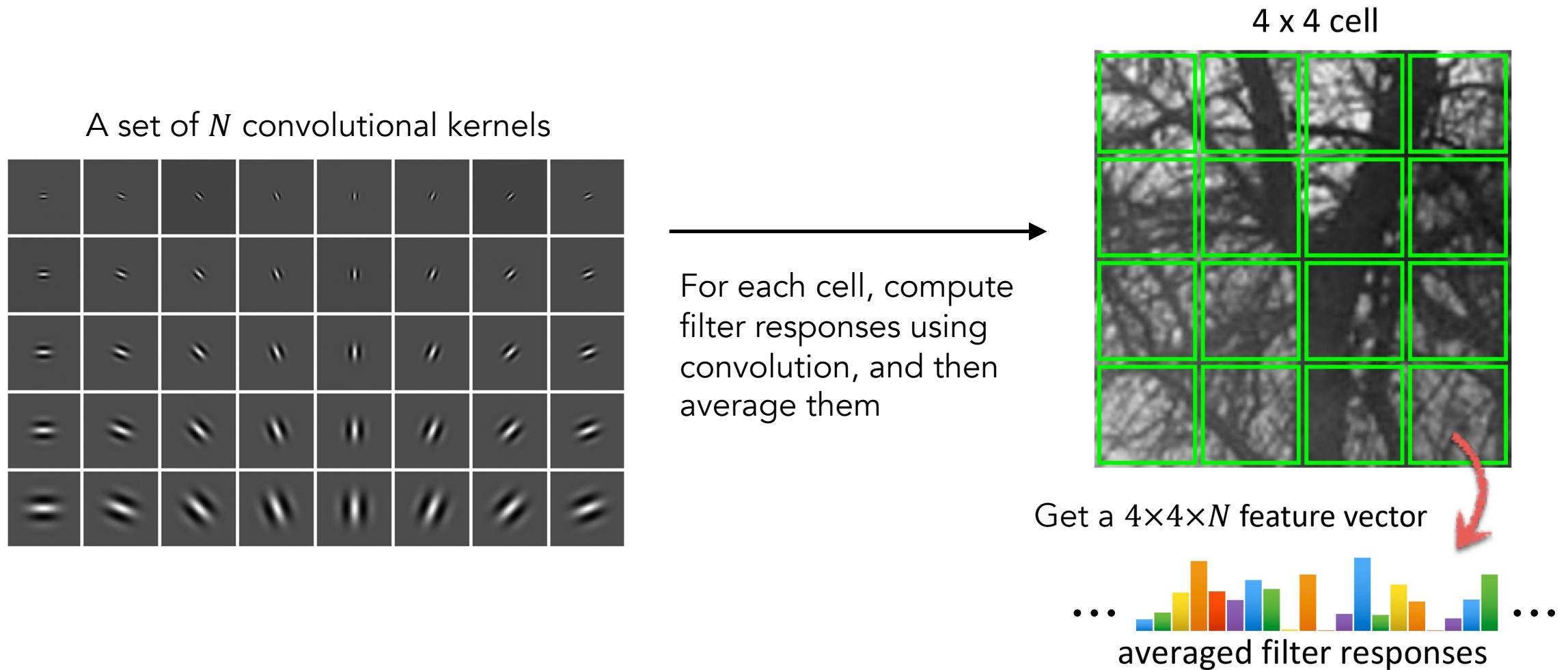
Vertical Sobel filter



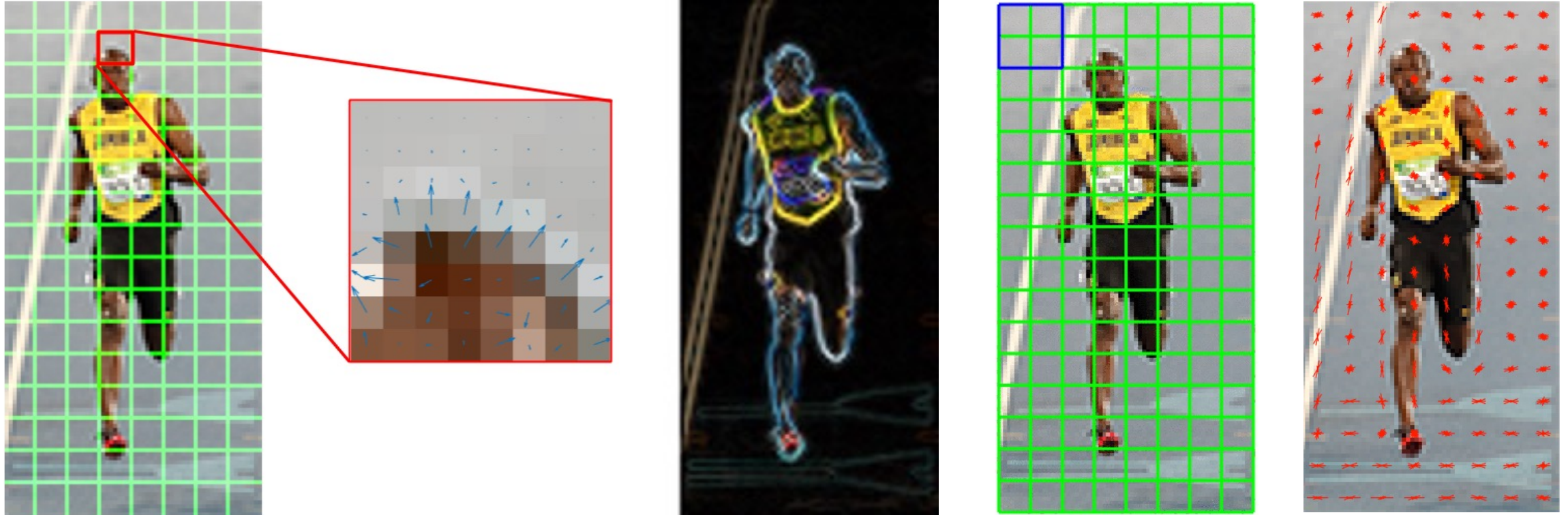
We can use various hand-crafted convolutional kernels (i.e., filter banks) to extract different kinds of information in the image, such as using the Gabor filters.



To construct features, we can perform convolution on image patches using a set of kernels (i.e., filter banks) and then aggregate them into a feature vector.



There are many ways of constructing feature vectors, such as Histogram of Gradients.



Compute gradients using Sobel filter



Aggregate gradients into histograms

Take-Away Messages

- Computers represent images as tensors with numbers in several channels (e.g., red, green, and blue).
- The origin (i.e., index $[0,0]$) of the image coordinate is at the top-left corner of the image.
- Convolution (the discrete version) means sliding a kernel/filter over the image to compute the sum of element-wise multiplication, which can also be seen as a dot product.
- We can use convolution to perform many different image-filtering operations, such as shifting pixels, blurring/sharpening images, and detecting edges.
- We can use a filter bank (i.e., a collection of kernels) to perform convolution on image patches, and the result (i.e., a set of filter responses) can be used as features.



Questions?