feature extraction and matching assignment

April 13, 2025

1 Introduction

We will first discuss the core idea behind SIFT (Scale-Invariant Feature Transform), which is a feature detection and description algorithm widely used in computer vision for object recognition, image stitching, and tracking. It identifies key points in an image that are invariant to scale, rotation, and illumination changes.

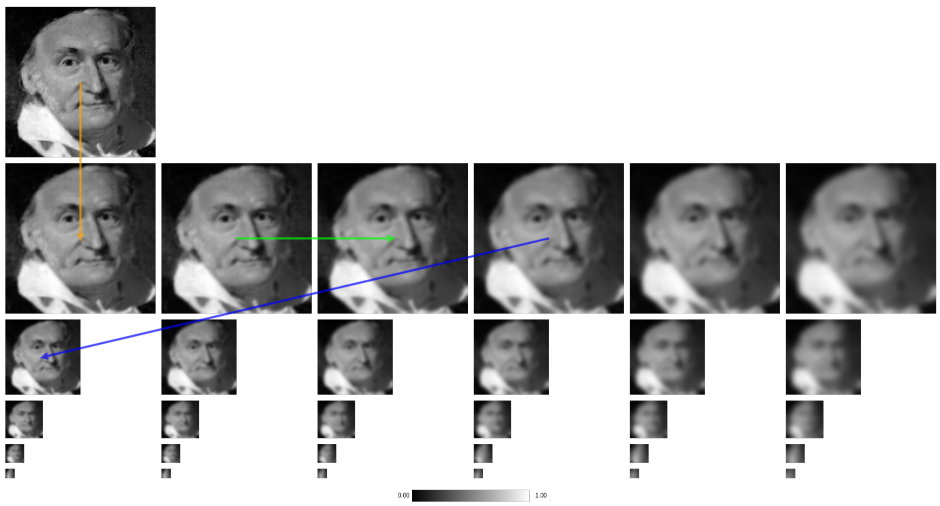

Figure 1: example of SIFT on a portrait. Image generated from https://weitz.de/sift/index.html?size=large

1.1 BLOBS

BLOB stands for Binary Large Objects. Informally, a blob is a region of an image in which some properties like intensity or color are approximately constant. We often use the term 'connected' here. Thus, these regions of densely connected pixels appear to be similar to each other. BLOBS are representative of features in images. So to detect features you may detect BLOBS.

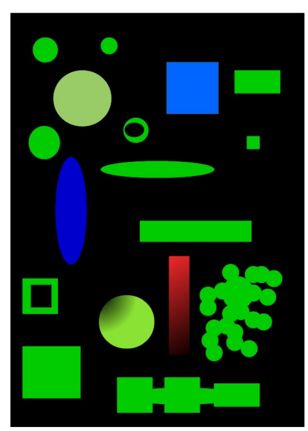

Figure 2: Shapes as examples of blobs

1.2 derivative of gaussian

We learned that to detect edges, we may take the derivative of the image. However, taking the derivative of a noisy signal can lead to many false positives. Thus, we need to smooth the image out a bit using the Gaussian filter, which is just a normal distribution of weights with the highest weight in the middle.

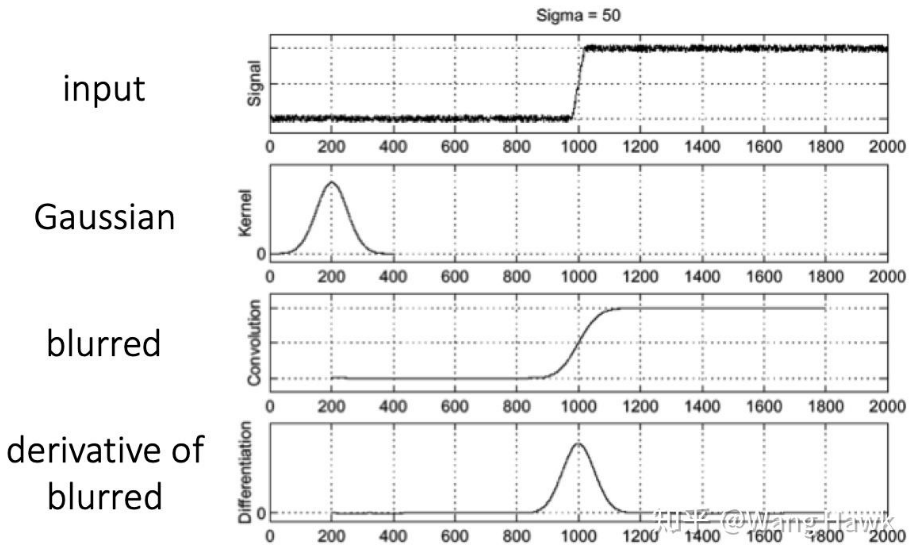

Figure 3: derivative of gaussian.

1.3 Laplacian of Gaussian as a BLOB detector

Moreover, the second derivative of Gaussian, AKA Laplacian of Gaussian can act as a blob detector by combining smoothing and second-order differentiation. Take the diagrams in Figure 4 as an example. Convolving with the Laplacian of Gaussian will result in a maximum when the σ matches the size of the BLOB.

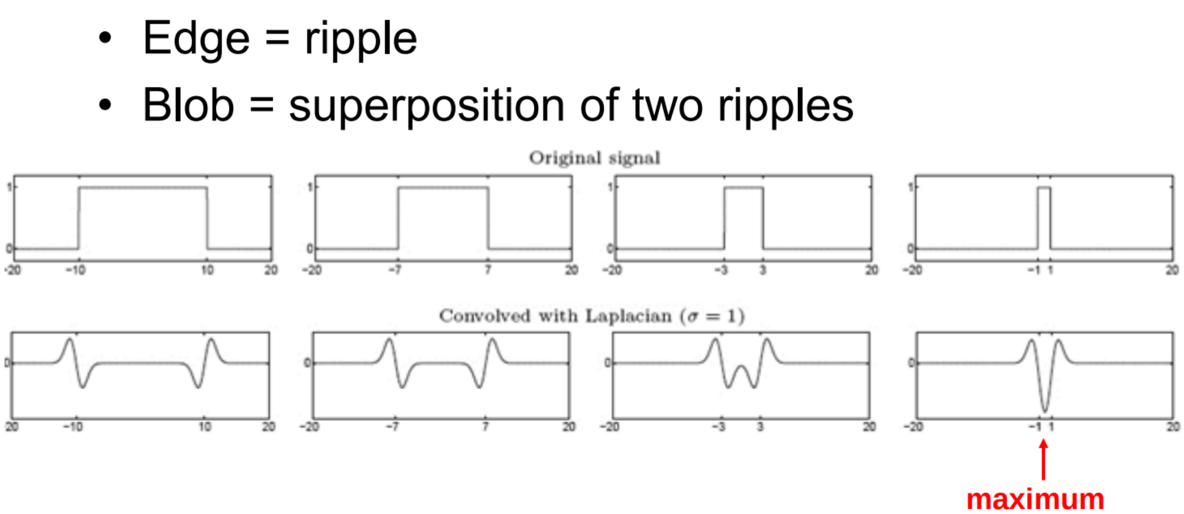

Figure 4: Laplacian of gaussian. Image taken from here

1.4 Difference of Gaussian (DoG)

The Difference of Gaussian (DoG) is an efficient approximation of the Laplacian of Gaussian (LoG) for blob detection. Instead of computing the computationally expensive second derivative directly, DoG subtracts two Gaussian-blurred versions of the image at slightly different scales. This approximation closely resembles the LoG response while being faster to compute. By varying the Gaussian standard deviation σ, DoG creates a multi-scale representation where blobs appear as extrema. This method is widely used in feature detection algorithms like SIFT, offering a balance between accuracy and computational efficiency.

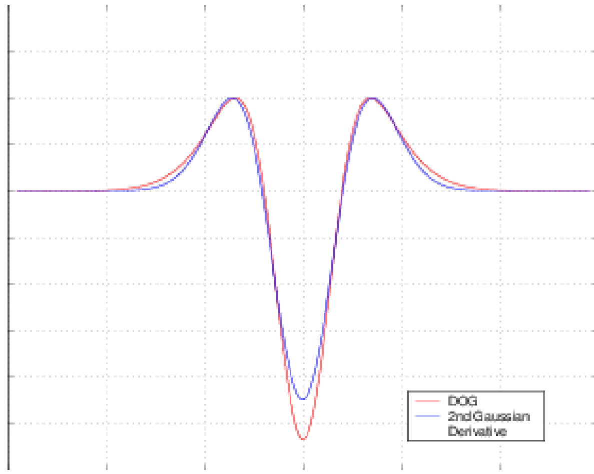

Figure 5: difference of Gaussian approximation. Image taken from here

The Difference of Gaussian (DoG) approximation to the Laplacian of Gaussian (LoG) is given by:

$$\text{DoG}(x, y, \sigma) = G(x, y, k\sigma) - G(x, y, \sigma)$$

where:

- $G(x, y, \sigma)$ is a Gaussian-blurred image with standard deviation σ, - k is a scaling factor (typically $k \approx \sqrt{2}$).

The DoG approximates the Laplacian of Gaussian as:

$$\nabla^2 G(x, y, \sigma) \approx \frac{G(x, y, k\sigma) - G(x, y, \sigma)}{(k-1)\sigma^2}$$

This provides an efficient alternative for blob detection and scale-space analysis.

2 Instructions

2.1 What is required?

Simply use the SIFT or ORB or any pre-implemented feature extractor in OpenCV and use a pre-implemented feature matcher as well to locate an object presented alone in a query image in a bigger image containing the object in a smaller size with a different background or just containing multiple objects (target image). You do not need to draw a rectangle on the detected object but locate it and draw points on it in the target image You can do this in groups of 2 and the deadline is next week.

2.2 Bonus

Try to use this object detection process on a video and constantly draw a rectangle on it.

2.3 example

Figure 6: query example

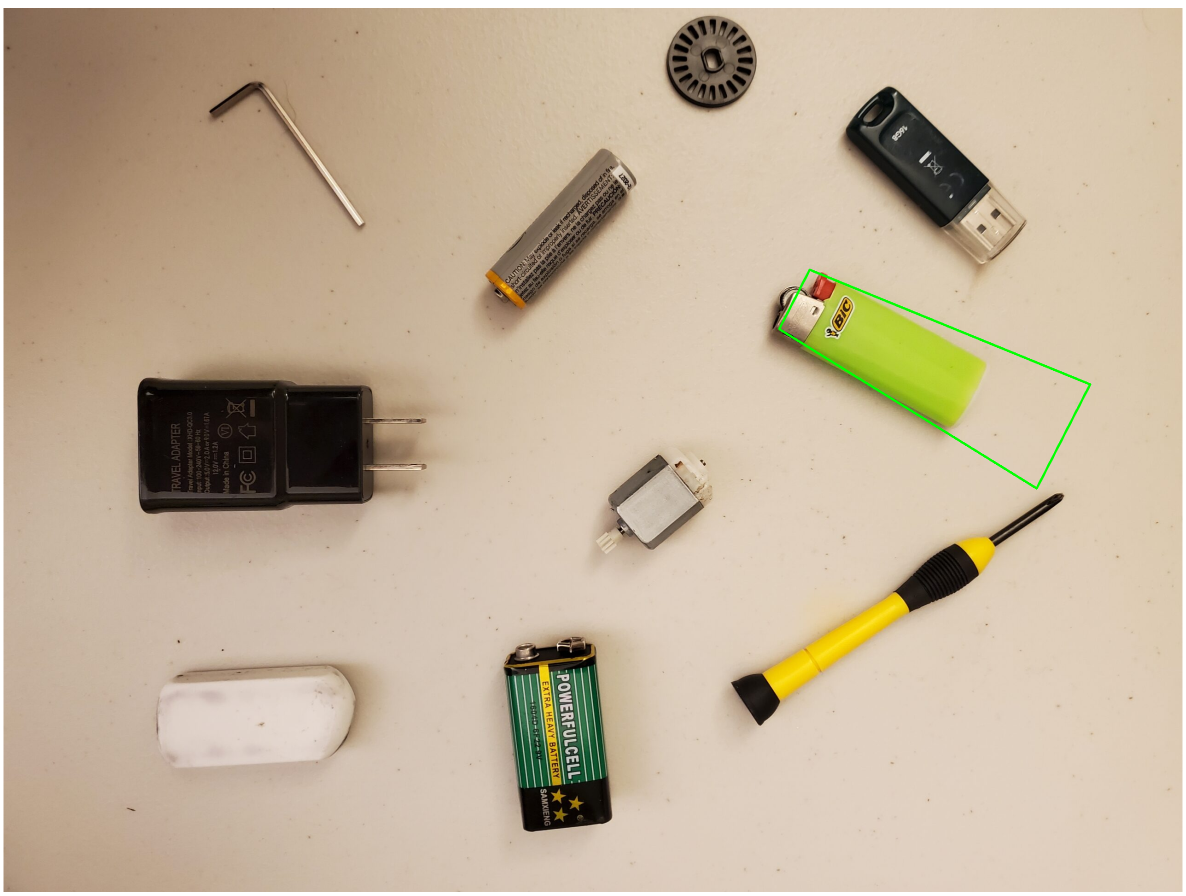

Figure 7: target image with the object detected