ECE 595: Machine Learning I

Lecture 8.2: Hand-Crafted and Deep Features - SIFT and HOG

Spring 2020

Stanley Chan

School of Electrical and Computer Engineering
Purdue University

PURDUE
UNIVERSITY

Outline

Feature Analysis

- Lecture 7 Principal Component Analysis (PCA)
- Lecture 8 Hand-Crafted and Deep Features

This Lecture

- Little History of Feature Extractions
- Convolution
  - What is convolution (if you don't know what it is yet)?
  - Some interesting facts about convolution
- SIFT and HOG
  - Gaussian derivatives
  - Pyramid
  - Histogram of oriented gradients
- Deep Features
  - What are they?
  - How to use them?

SIFT

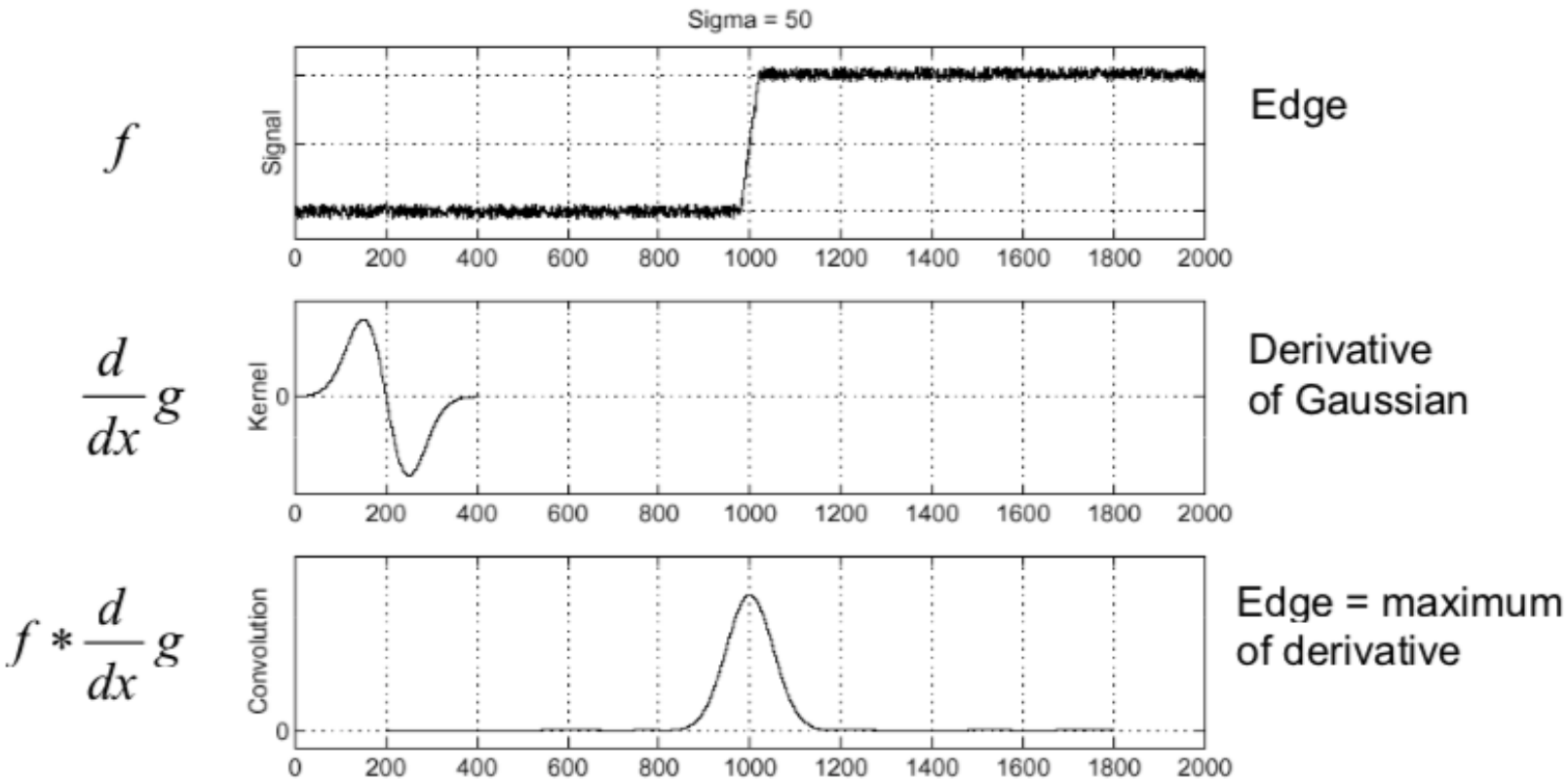
- SIFT = Scale-invariant feature transform.
- Proposed by David Lowe in 1999.
- Idea: Convolve the image with the **2nd order derivative of a Gaussian**.
- Vary the radius of the Gaussian. Locate the radius that maximizes the response.
- What makes SIFT so powerful? The derivative of Gaussian filter extracts the scale:

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

  where $G(\boldsymbol{x}, \sigma) = (1/(2\pi\sigma^2)) \exp\{-\|\boldsymbol{x}\|^2/(2\sigma^2)\}$.
- Output of SIFT: A set of **locations** where there are "**blobs**".
- Can be used for registering images.

Gaussian Filter for Edge Detection

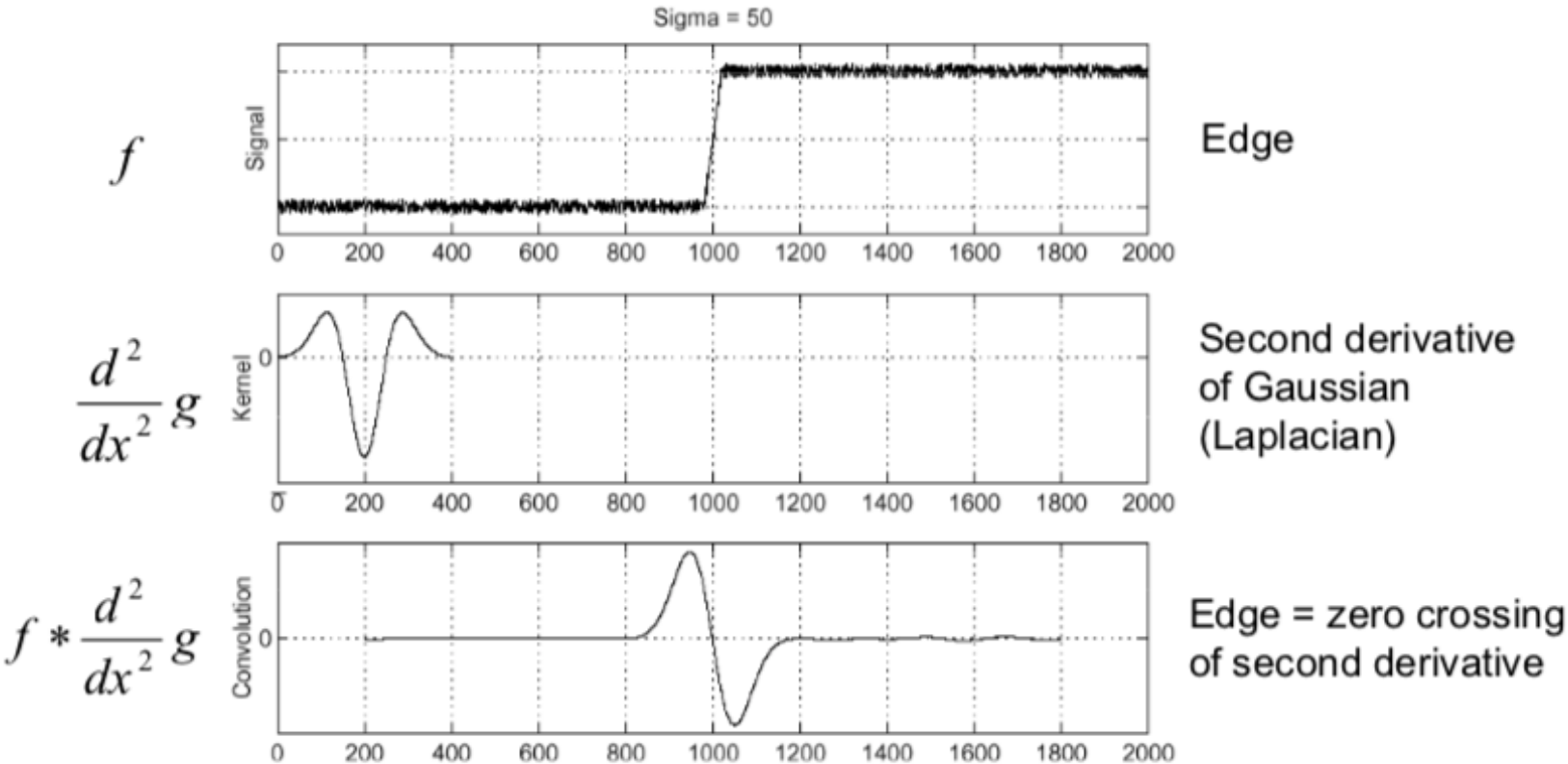
Recall: Edge detection



https://towardsdatascience.com/sift-scale-invariant-feature-transform-c7233dc60f37

2nd Order Gaussian Derivative

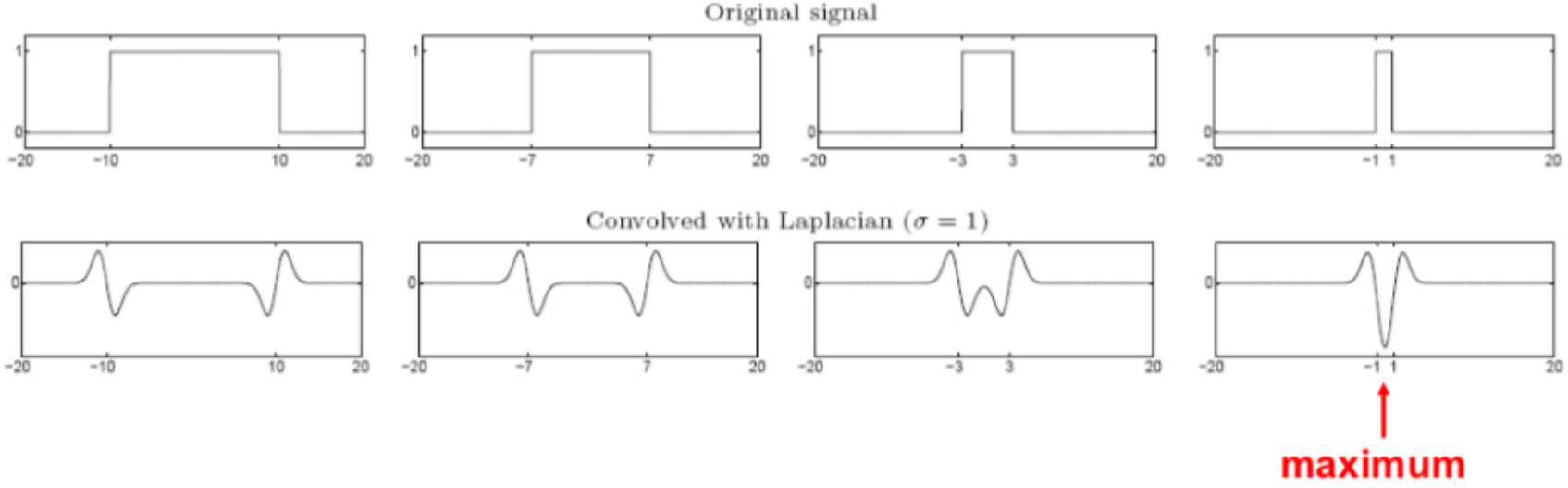
Edge detection, Take 2



https://towardsdatascience.com/sift-scale-invariant-feature-transform-c7233dc60f37

Locating Blobs

- Edge = ripple
- Blob = superposition of two ripples



Spatial selection: the magnitude of the Laplacian response will achieve a maximum at the center of the blob, provided the scale of the Laplacian is "matched" to the scale of the blob

https://towardsdatascience.com/sift-scale-invariant-feature-transform-c7233dc60f37

Scale Pyramid

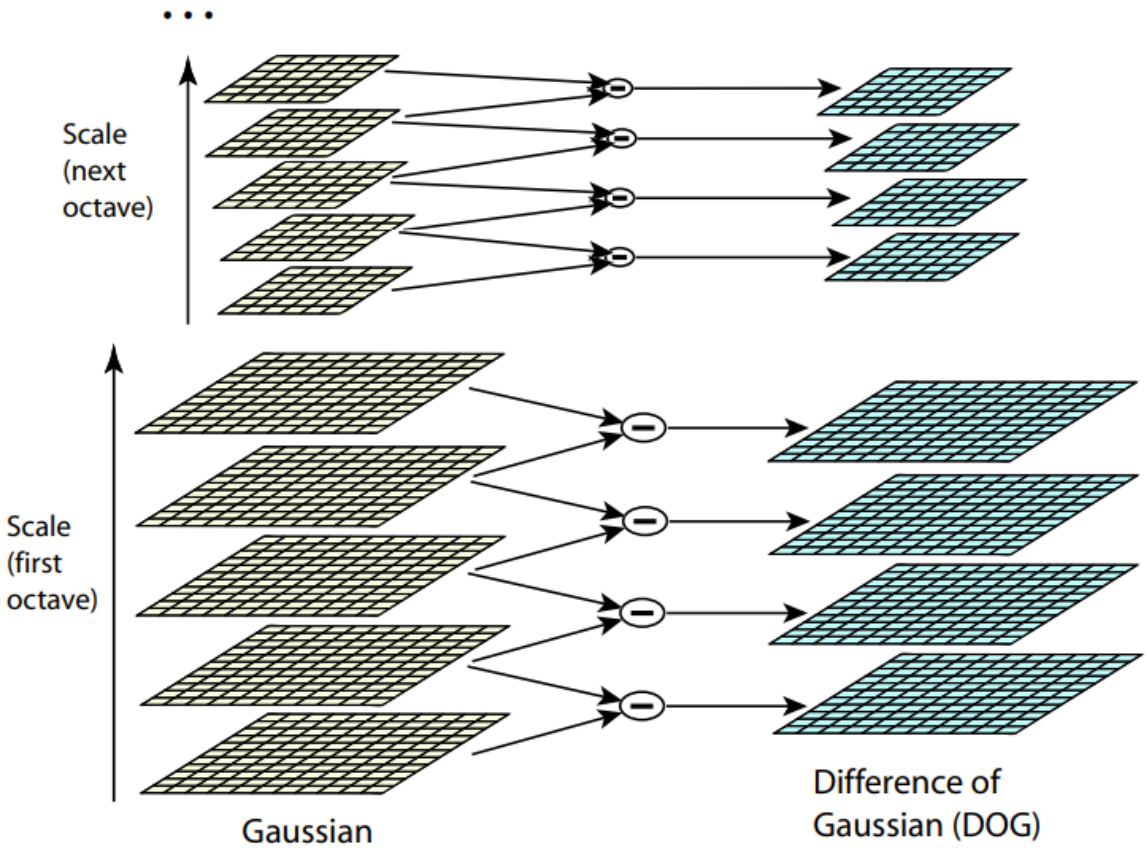
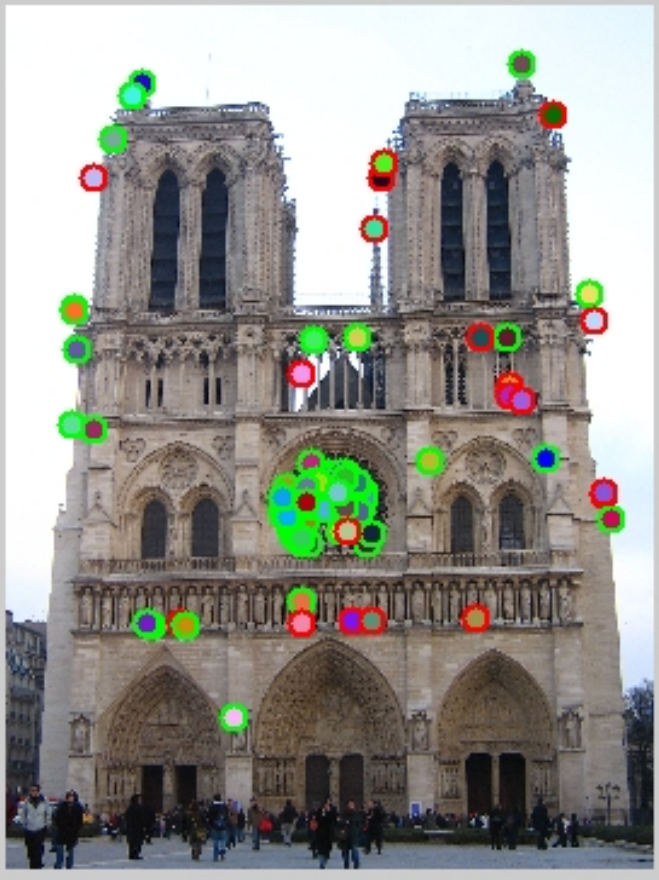
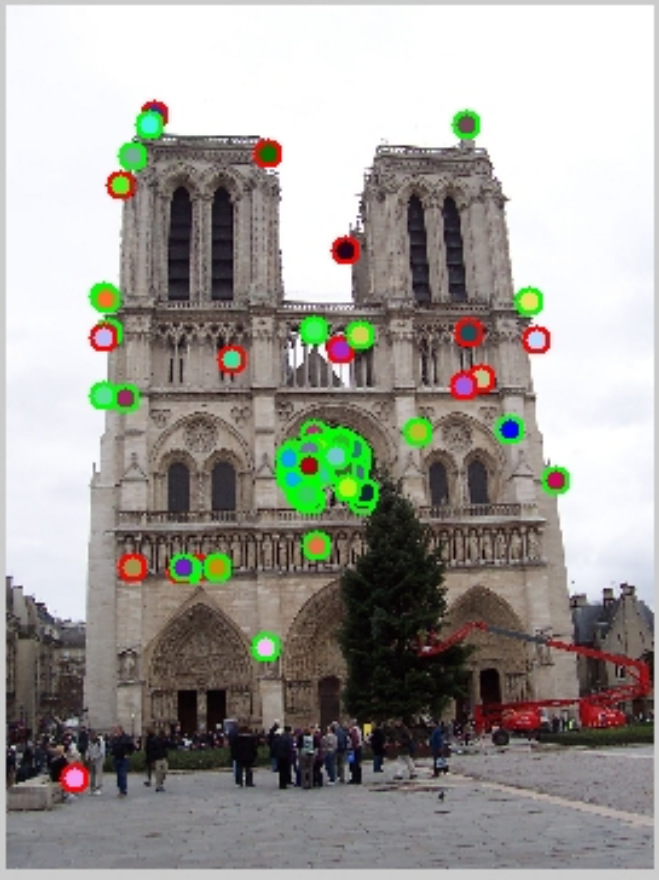


Figure 1: For each octave of scale space, the initial image is repeatedly convolved with Gaussians to produce the set of scale space images shown on the left. Adjacent Gaussian images are subtracted to produce the difference-of-Gaussian images on the right. After each octave, the Gaussian image is down-sampled by a factor of 2, and the process repeated.

https://people.eecs.berkeley.edu/~malik/cs294/lowe-ijcv04.pdf

Example

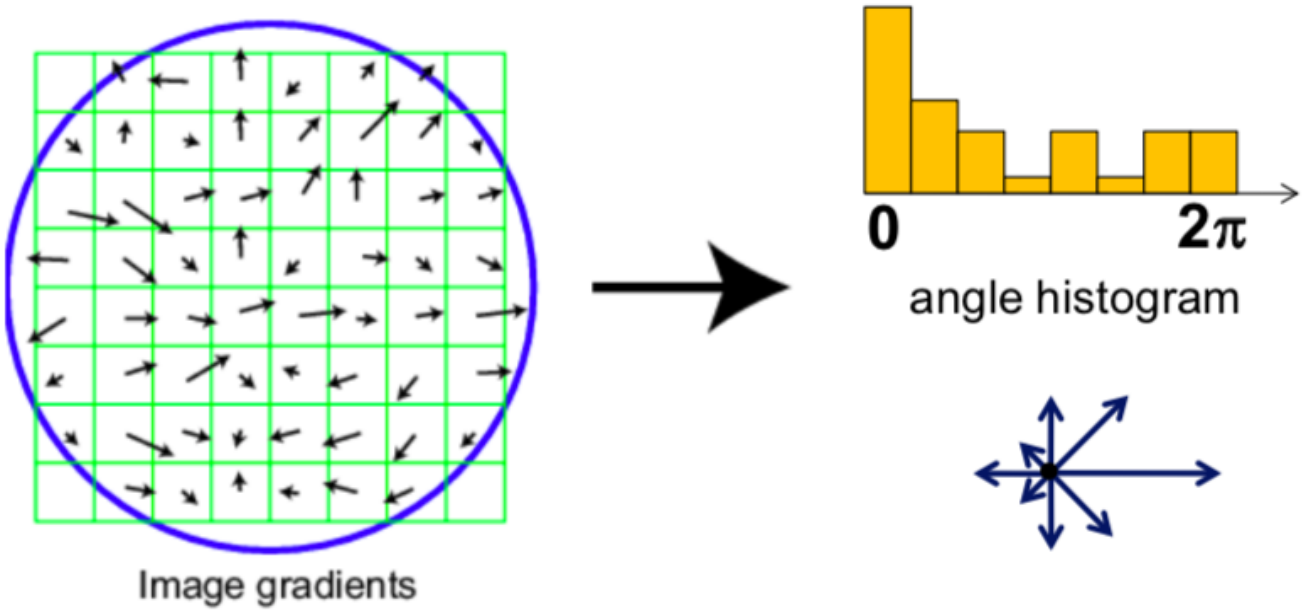




http://cs.brown.edu/courses/cs143/2013/results/proj2/rroelke/

SIFT + HOG

- HOG = Histogram of Oriented Gradient
- Used to "encode" the detected blobs
  - Take 16x16 square window around detected feature
  - Compute edge orientation (angle of the gradient - 90°) for each pixel
  - Throw out weak edges (threshold gradient magnitude)
  - Create histogram of surviving edge orientations



https://towardsdatascience.com/sift-scale-invariant-feature-transform-c7233dc60f37