ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Computer Vision
AS 2023

Prof. Marc Pollefeys *Prof. Siyu Tang* *Prof. Fisher Yu*

Image Segmentation Exercise

18th of November, 2023

We have discussed 3 image segmentation approaches in this course: (1) K-Means algorithm, (2) Expectation-Maximization algorithm, (3) Mean-Shift.

a) Compared to the K-Means algorithm, the Expectation-Maximization (EM) algorithm can model more diverse cluster shapes than K-Means, what limits K-Means in its capability to model the cluster shape, and what gives EM more flexibility than K-Means? Besides modeling more diverse cluster shapes, what is the other major advantage of the EM algorithm over the K-Means algorithm? **3 pts.**

ANSWER ANSWER

(1) K-Means's limitation comes from the fact that it uses squared distance for measuring error of each point to its cluster centroid. (1-pt)

(2) EM models each cluster as a multivariate Gaussian distribution, thus it's more flexible to handle cluster shapes such as ellipse. (1-pt)

(3) EM explicitly models the probability of a point belonging to a cluster, while K-Means must strictly assign a point to a single cluster. (1-pt)

ANSWER ANSWER

b) Given the number of clusters as K, and the dimension of a data-point as d, what is the space complexity (i.e. big O) for storing an EM model and a K-Means model, respectively? **2 pts.**

ANSWER ANSWER

$O(Kd^2)$ for EM, as each cluster is represented by mean and covariance, covariance matrix takes $O(d^2)$ space (1-pt). $O(Kd)$ for k-means, as we only need to storage d dimensional cluster center for each cluster. (1-pt)

ANSWER ANSWER

c) In the Mean-Shift segmentation algorithm, we do not need to pre-define the number of clusters k at initialization. Rather, we merge the mean and combine the clusters during the algorithm.

Explain:

(1) How does this decision relate to our assumption of the data? Compare it to the assumption we made for the EM algorithm.

(2) What will happen when the data has outliers? Compare it to k-means and EM algorithm. **4 pts.**

ANSWER ANSWER

(1) Mean-Shift segmentation algorithm is model-free and does not assume any prior about the data. EM algorithm assumes that each cluster can be represented by a Gaussian distribution (spherical / elliptic) (2-pt).

(2) Mean-Shift segmentation algorithm is robust to outliers as it will be too far away to affect other clusters. In k-means algorithm, an outlier needs to belong to some group and it will affect the mean of that group. This problem is less pronounced in EM as the outliers will always have low probability in each Gaussian (2-pt).

ANSWER ANSWER

We have discussed Hough transforms in the lecture:

d) For line detection using Hough transform, we are given a set of N detected 2D edge points $\{(x, y)^{(i)}\}_{i=1}^{N}$ and we want to map each point into a parameter space of lines. Given the definition of a line as $y = ax + b$ where (a, b) are parameters of the line, a naive implementation would be using (a, b) directly as the parameter space. Explain:

(1) What is the problem of defining parameter space with (a, b)?

(2) What is the alternative parameter space used in Hough transform?

(3) What are the bounds of the alternative parameter space? **3 pts.**

ANSWER ANSWER

(1) The parameter domain is unbounded, a vertical line can result in infinite value of a. (1-pt)

(2) Parameter space (ρ, θ) where $\rho = x \cos\theta + y \sin\theta$. (1-pt)

(3) θ is bounded in $[-\pi, \pi]$ while ρ is also bounded by the largest distance between any line composed from detected edge points and the origin (1-pt)

ANSWER ANSWER

In the lecture, we have discussed dilated convolution.

e) Given two 3x3 consecutive convolution layers with stride 1, what is the receptive field of a pixel after applying these two layers? What if the **second** convolution layer is replaced with dilated convolution with dilation=2? **4 pts.**

ANSWER ANSWER

(1) Receptive field is 5 or 5x5 (2-pt).

(2) Receptive field is 7 or 7x7 (2-pt).

ANSWER ANSWER