PMSCS Program
Department of Computer Science and Engineering
Jahangirnagar University
Final Examination Fall-2021

Answer Script
Total Page# 29

Course Information						
Course Title:	Computer Vision					
Course Code:	PMSCS-664		**Marks:**	30	**Time:**	1 Hour 30 Minutes

Instructions

1. Be present in the corresponding classroom at least 15 minutes prior.
2. Any form of unfair means or cheating (verbal communication, visual copying, unauthorized chit etc.) is a punishable offense.
3. You will not be allowed to use additional sheets. Limit your answer so that it fits within the allocated space.
4. Scan the whole answer script and turn in through Google Classroom.
5. Failing to upload the answer script within the given time span will be considered as disqualification.

Student Information	
Full Name:	ARIK MD ISTHIAQUE
Class Roll:	CSE 2021 0202 4
Date:	19-02-2022

Office Use Only		
Questions	**Marks**	**Remarks**
1.		
2.		
3.		
4.		
Total		

Answer to the question number 1(a):

Computer vision: Computer vision is the interdisciplinary scientific field deals with how computer can gain high level understanding from digital images or videos.

Contrast Between Computer vision & Computer Graphics: Computer vision is the image understanding like AI. It is a sensor modality for robotics. It can emulate human vision or computer. Computer graphics is the exact opposite of this.

images → computer vision → world model

world model → computer graphics → images.

Answer to the question number 1(a):

Computer Graphics change models to images. On the other hand computer vision change images to models.

Computer Graphics examples are -

1. rendering
2. animation
3. UI
4. surface design

computers vision examples are

1. Shape estimation
2. recognition
3. 3d modeling
4. motion estimation

Answer to the question number 1(b):

Harris Detector: Harris Detector is a single-scale image features detector. Harris & Stephens have developed this detector by combining the corner & and edge detectors and to address the limitations of the Moravec's detector. The basic idea of the detectors is there is no edge or corners their will no change in any direction. If their is a edge then no change in the edge direction and if their is edge & corner both are present then their will be change in all direction.

Answer to the question number 1(b):

Here, change of intensity for the shift $[u,v]$:

$$E(u,v) = \sum_{x,y} w(x,y)\,[I(x+u, y+v) - I(x,y)]^2$$

For small ~~change~~ shift $[u,v]$ we have a bilinear approximation

$$E(u,v) \cong [u,v]\, M \begin{bmatrix} u \\ v \end{bmatrix}$$

Here, M is 2×2 matrix

$$M = \sum_{x,y} w(x,y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}$$

Answer to the question number 1(b):

and the measure of corner response

$$R = \det M - k(\text{trace } M)^2$$

here, $\det M = \lambda_2 \lambda_1$

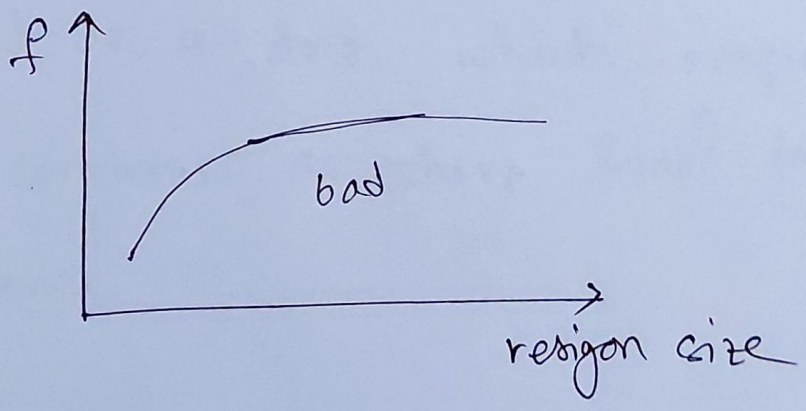
$\text{trace } M = \lambda_1 + \lambda_2$

Answer to the question number 1(c):

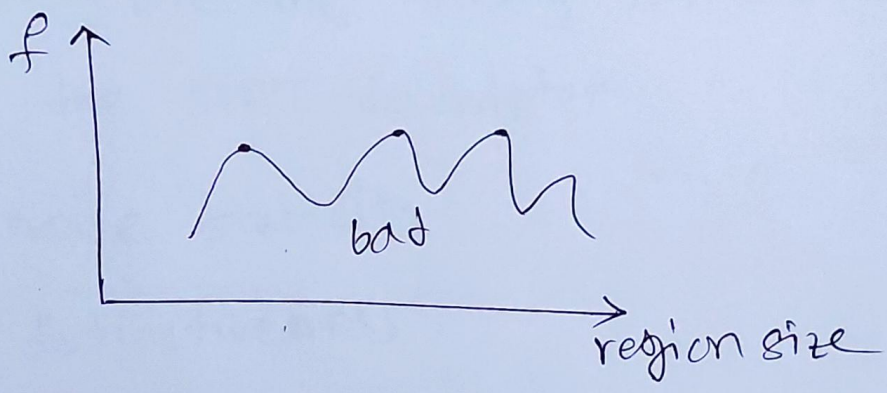
Scale Invariant Features Transform:

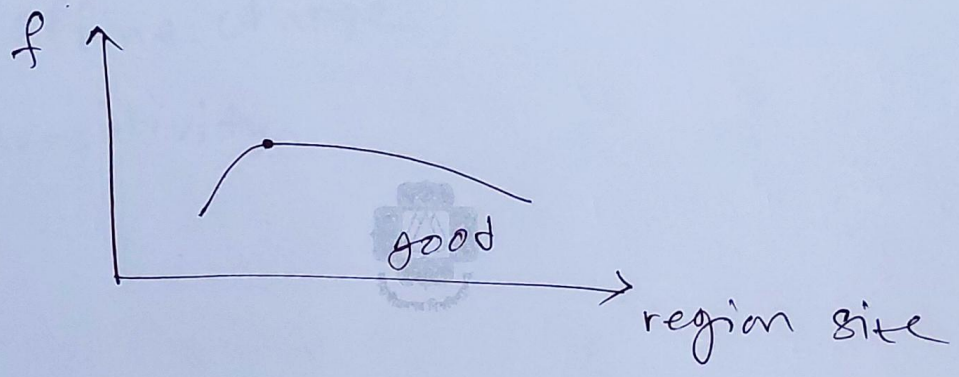
It also know as SIFT. Here a number of interest points are detected in the images using the DOG operation. The points are selected as local extrema of the DOG function. At each interest point, a feature vector is extracted.

A good function for scale detection has one stable sharp peak.



Answer to the question number 1(c):





For common images a good function would be a one which responds to contrast as sharp local intesity changes

Answer to the question number 1(c):

Following are the tuning and evaluatings for the SIFT descriptor

① noise stability

② distinctiveness

③ affine change

④ sensitivity.

Answer to the question number 2(a):

clustering: A cluster is represented by a single point know as the centroid. Cluster boundary is decided by the farthest data point in the cluster. As there is no pre trained data process in this system, so the clustering is called a unsupervised learning.

Two goals of supervised learning:

Supervised learning has two main goal.

① **predictive:** make new predictions for a new sample describe by its attributes.

Answer to the question number 2(a):

② Informative: Help to understand the relationship between the inputs and the outputs. Also find the most relevant output.

Answer to the question number 2(b):

Clustering Techniques: Task of grouping a set of data points such that data points in the same group are more similar to each others than data points in another group.

There are two types of clustering-

① Hierarchical

② Partitional

Two of the techniques have their own methods to do clustering.

Answer to the question number 2(b):

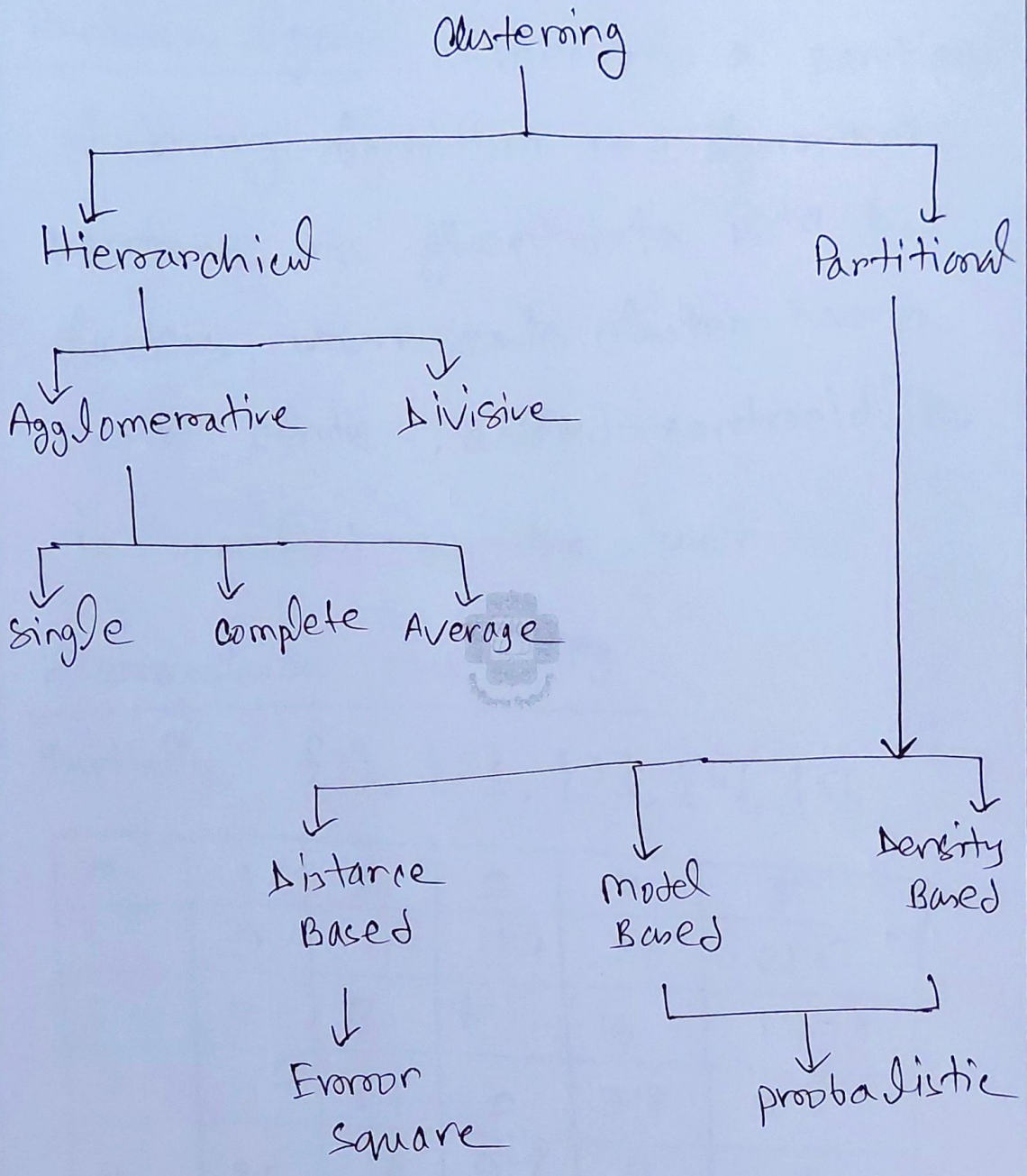


fig: clustering techniques.

Answer to the question number 2(c):

K-means algo: K-means is a partional clustering algorithm. The algorithm partions the given data into K clusters. where each cluster has a cluster center, called centroid. K is specified by the user.

Hierarchical clustering

Initially, {1}, {2}, {3}, {4}, {5}

x	1	2	3	4	5
1	0	4	11.7	20	21.5
2	4	0	8.1	16	17.9
3	11.7	8.1	0	9.8	9.8
4	20	16	9.8	0	8
5	21.5	17.9	9.8	8	0

Answer to the question number 2(c):

Now, from the distance table

{1,2}, {3}, {4}, {5} and d = 4

	1,2	3	4	5
1,2	0	8.1	16	17
3	8.1	0	9.8	9.8
4	16	9.8	0	8
5	17	9.8	8	0

Now, {1,2}, {3}, {4,5} and d = 8

	1,2	3	4,5
1,2	0	8.1	16
3	8.1	0	9.8
4,5	16	9.8	0

Answer to the question number 2(c):

Now $\{1, 2, 3\}, \{4, 5\}$ and = 8.1

	1, 2, 3	4, 5
1, 2, 3	0	9.8
4, 5	9.8	0

So, the final ~~clusters~~ clusters $\{1, 2, 3, 4, 5\}$ and $d = 9.8$

This page is kept blank intentionally...

Answer to the question number 3(a):

P.T.O

Answer to the question number 3(a):

P.T.O

Answer to the question number 3(a):

P.T.O

Answer to the question number 3(b):

P.T.O

Answer to the question number 3(b):

P.T.O

Answer to the question number 3(c):

P.T.O

Answer to the question number 3(c):

P.T.O

Answer to the question number 4(a):

Multi-class Confusion matrix: Confusion matrix is used to know the performance of machine learning classifications. It is represented in matrix form. The matrix is N×N matrix, N is the number of classes. If it is greater then 1 then we call it multiclass matrix.

given,

class	+	−
+	85 (TP)	7 (FN)
−	18 (FP)	9 (TN)

$$\therefore \text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN}$$

$$= .78$$

$$= 78\%$$

Answer to the question number 4(a):

$$\text{Sensitivity} = \frac{TP}{TP+FN} = \frac{85}{85+7}$$

$$= .92$$

$$\text{Specificity} = \frac{TN}{TN+FP} = \frac{9}{9+18} = .33$$

$$\text{precision} = \frac{TP}{TP+FP} = \frac{85}{85+18} = .82$$

$$\text{Recall} = \frac{TP}{TP+FN} = \frac{85}{85+7} = .92$$

$$\text{F1 score} = \frac{2TP}{2TP+FP+FN}$$

$$= \frac{2 \times 85}{2 \times 85 + 18 + 7}$$

$$= .87$$

Answer to the question number 4(b):

Perceptron train for logical OR operation.

Epoch 1 $\alpha = .1$ $th = .2$

x_0	x_1	y_d	w_1	w_2	y	e	w_1	w_2
0	0	0	.3	-.1	0	0	.3	-.1
0	1	1	.3	-.1	0	1	.3	0
1	0	1	.3	0	1	0	.3	0
1	1	1	.3	0	1	0	.3	0

Epoch 2

x_1	x_2	y_d	w_1	w_2	y	e	w_1	w_2
0	0	0	.3	0	0	0	.3	0
0	1	1	.3	0	0	1	.3	.1
1	0	1	.3	.1	1	0	.3	.1
1	1	1	.3	.1	1	0	.3	.1

Answer to the question number 4(b):

As the w_1, w_2 final value remain same: epoch 3

x_1	x_2	y_d	w_1	w_2	y	e	w_1	w_2
0	0	0	.3	.1	0	0	.3	.1
0	1	1	.3	.1	0	-1	.3	.2
1	0	1	.3	.2	1	0	.3	.2
1	1	1	.3	.2	1	0	.3	.2
epoch 4								
0	0	0	.3	.2	0	0	.3	.2
0	1	1	.3	.2	1	0	.3	.2
1	0	1	.3	.2	1	0	.3	.2
1	1	1	.3	.2	1	0	.3	.2

As the w_1, w_2 final value remain sam we can say the final value is $w_1 = .3$, $w_2 = .2$

Answer to the question number 4(b):