SRM INSTITUTE OF SCIENCE AND TECHNOLOGY

RAMAPURAM CAMPUS

FACULTY OF ENGINEERING AND TECHNOLOGY

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

ACADEMIC YEAR (2022-2023)

CONTINUOUS LEARNING ASSESMENT-II

ANSWER KEY

Sub Code/Name	18CSE390T – Computer Vision	Set	EVEN
Year/Sem/Branch	III/ V/ B.Tech-CSE-AIML	Date	17.10.22
Max. Marks	50	Duration	90 Mins.

PART A (10 X 1= 10)

ANSWER ALL THE FOLLOWING QUESTIONS

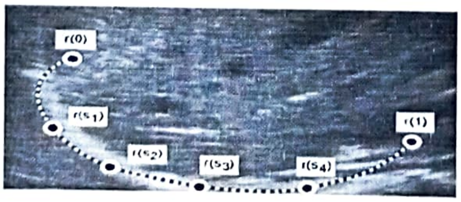
Q.No.	MCQ Questions	Marks	CO	BL	PI
1.	**For edge detection we observe** a) intensity transition b) shape transition c) color transition d) **sign transition**	1	2	1	1.6.1
2	**The direction of angle to the gradient is** a) **Orthogonal** b) Isolated c) Isomorphic d) Isotropic	1	2	1	1.6.1
3	**Edge detection in images is commonly accomplished by** performing a spatial --- ---of the image field. a) Smoothing Filter b) Integration c) **Differentiation** d) Min Filter	1	2	2	1.6.1
4	**Multi-dimensional hashing maps descriptors into ________________ based on some function applied to each descriptor vector.** a) **fixed size buckets** b) variable sized buckets c) table d) Dbms	1	2	2	1.6.1
5	**Isolated edge points can also be grouped into____________________** a) Pixel b) region c) **Longer curves or contours, as well as straight line segments** d) Contour	1	2	1	1.6.1
6	**Techniques like Livewire or Intelligent Scissors are used in** a.Model based segmentation b.**Semi automatic segmentation** c.Threshold segmentation d.Segmentation	1	3	1	1.6.1

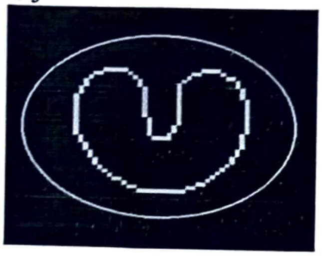
7	**Example of Active Contour** **a.Snakes, intelligent scissors, level set** b. Succesive Approximation c. Hough Transform d.Scissors	1	1	1	1.6.1
8	**An Approach which optimize the contour in real time as the user is drawing** **a) Intelligent Scissors System** b) Gaussian c) Similarity d) Edge	1	3	1	1.6.1
9	**In level set which define the curve** a. Contrast b. Quantization c. Sampling **d. Zero crossing of a characteristic function**	1	3	1	1.6.1
10	**Split and merge technique is** a. Image Restoration Technique **b. an Image Processing Technique Used To Segment An Image** c. Image Enhancement Technique d. Image Acquisition Technique	1	3	1	1.6.1

PART B (4 X 4 = 16)

ANSWER ANY FOUR OUT OF SIX QUESTIONS

Q. No.	Questions	Marks	CO	BL	PI
11	**Discuss about Bias and Gain normalization (MOPS)** For tasks that do not exhibit large amounts of foreshortening, simple normalized intensity patches perform reasonably well and are simple to implement. In order to compensate for slight inaccuracies in the feature point detector (location, orientation, and scale), these multi-scale oriented patches (MOPS) are sampled at a spacing of five pixels relative to the detection scale, using a coarser level of the image pyramid to avoid aliasing. To compensate for affine photometric variations (linear exposure changes or bias and gain, (3.3)), patch intensities are re-scaled so that their mean is zero and their variance is one.	4	2	1	2.5.1
12	**Explain briefly about Vanishing points** In many scenes, structurally important lines have the same vanishing point because they are parallel in 3D. Examples of such lines: horizontal and vertical building edges, zebra crossings, railway tracks, the edges of furniture such as tables and dressers, and the ubiquitous calibration pattern Finding the vanishing points common to such line sets can help refine their position in the image and, in certain cases; help determine the intrinsic and extrinsic orientation of the camera. The first stage in my vanishing point detection algorithm uses a Hough transform to accumulate votes for likely vanishing point candidates. As with line fitting, one possible approach is to have each line vote for all possible vanishing point directions, either using a cube map or a Gaussian sphere, optionally using knowledge about the uncertainty in the vanishing point location to	4	2	2	2.5.2

	Preferred approach is to use pairs of detected line segments to form candidate vanishing point locations. Let ˆmi and ˆmj be the (unit norm) line equations for a pair of line segments and li and lj be their corresponding segment lengths.The location of the corresponding vanishing point hypothesis can be computed as $v_{ij} = \hat{m}_i \times \hat{m}_j$ (4.28) and the corresponding weight set to $w_{ij} = \|v_{ij}\| l_i l_j.$ (4.29) This has the desirable effect of down weighting (near-)collinear line segments and short line segments. The Hough space itself can either be represented using spherical coordinates or as a cube map				
13	**Write short notes on Edge Linking** **Applications:** line detection and sparse stereo matching .If the edges have been detected using *zero crossings* of some function, linking them up is straightforward, since *adjacent edgels share common endpoints.* Linking the edgels into chains involves picking up an unlinked edgel and following its neighbors in the same direction. Either a *sorted list of edgels* (sorted first by x coordinates and then by y coordinates, for example) or a *2D array* can be used to accelerate the neighbor finding. If edges were not detected using *zero crossings*, finding the continuation of an edgel can be tricky. In this case, *comparing the orientation* (and, optionally, phase) of adjacent edgels can be used for disambiguation. Ideas from *connected component computation* can also sometimes be used to make the edge linking process even faster. Once the edgels have been linked into chains, we can apply an optional thresholding with hysteresis to remove low-strength contour segments	4	2	2	2.5.4
14	**Discuss in detail about Snakes** • Represents an object boundary or some other salient image feature as a parametric curve. An energy functional E is associated with the curve.The problem of finding object boundary is cast as an energy minimization problem **A Snake is a parametric**  • The course of the snake smoothly follows high	4	3	2	2.5.1

	object boundary. • A smooth boundary is generated bridging regions of noisy data or missing gradients. • Active contour is particularly well suited to segment an object instance in an image where the data are distorted by noise • A higher level process or a user initializes any curve *close to the object boundary.* • The snake then starts *deforming* and moving towards the desired object boundary. • In the end it completely "shrink-wraps" around the object • 				
15	**Difference between Divisive and Agglomerative algorithms in Cluster analysis.** **Region splitting (divisive clustering)** Splitting the image into successively finer regions is one of the oldest techniques in computer vision. First computes a histogram for the whole image and then finds a threshold that best separates the large peaks in the histogram. This process is repeated until regions are either fairly uniform or below a certain size. More recent splitting algorithms often optimize some metric of intra-region similarity and inter-region dissimilarity. **Region merging (agglomerative clustering)** Region merging techniques also date back to the beginnings of computer vision. Use a dual grid for representing boundaries between pixels and merge re-gions based on their relative boundary lengths and the strength of the visible edges at these boundaries. In data clustering, algorithms can link clusters together based on the distance between their closest points (single-link clustering), their farthest points (complete-link clustering), or something in between provide a probabilistic interpretation of these algorithms and show how additional models can be incorporated within this framework. A very simple version of pixel-based merging combines adjacent regions whose average color difference is below a threshold or whose regions are too small.	4	3	2	2.6.4
16	**Write short note on Pose Estimation.** • Estimating an object's 3D pose from a set of 2D point projection • Pose estimation problem is also known as extrinsic calibration • Problem of recovering pose from three correspondences, which is the minimal amount of	4	3	2	2.6.2

is known as the perspective-3-point-problem (P3P).

- Extensions to larger numbers of points collectively known as PnP
- Simplest way to recover the pose of the camera is to form a set of linear equations analogous to those used for 2D motion estimation from the camera matrix form of perspective projection

$$x_i = \frac{p_{00}X_i + p_{01}Y_i + p_{02}Z_i + p_{03}}{p_{20}X_i + p_{21}Y_i + p_{22}Z_i + p_{23}} \quad (6.33)$$

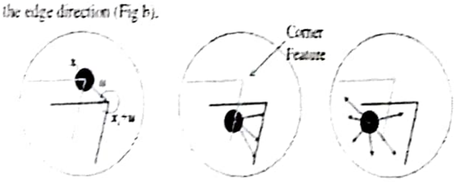
$$y_i = \frac{p_{10}X_i + p_{11}Y_i + p_{12}Z_i + p_{13}}{p_{20}X_i + p_{21}Y_i + p_{22}Z_i + p_{23}}. \quad (6.34)$$

where (x_i, y_i) are the measured 2D feature locations and (X_i, Y_i, Z_i) are the known 3D feature locations

- System of equations can be solved in a linear fashion for the unknowns in the camera matrix P by multiplying the denominator on both sides of the equation.
- The resulting algorithm is called the direct linear transform

PART C (2 X 12 = 24)

ANSWER EITHER OF OR IN EACH UNIT

Q. No.	Questions	Marks	CO	BL	PI
17	a) **Explain in detail about Feature Detection techniques with relevant examples and diagrams.** **Feature detectors** **Figure shows aperture problem for various images.** • The two images I_0 (yellow) and I_1 (red) are overlaid. • The red vector u indicates the displacement between the patch centers • $w(x_i)$ weighting function (patch window) is shown as a dark circle. • Patches with gradients in at least two (significantly) different orientations are the easiest to localize.(Fig a). • Although straight line segments at a single orientation suffer from the aperture problem i.e., it is only possible to align the patches along the direction normal to the edge direction (Fig b).  (a) Patch with stable (point-like) flow (b) Classic aperture problem (barber-pole illusion) (c) Textureless region	12	2	3	2.6.4

Feature detection

Shifting the window W by (u, v)

- Compare each pixel before and after shift image by Summing the Squared Differences (SSD)
- I_0 and I_1 are the two images being compared, (u, v) is the displacement vector, $w(x, y)$ is a spatially varying weighting (or window) function, the summation i is over all the pixels in the patch
- This defines an Weighted SSD "Error" of $E(u, v)$

$$E_{WSSD}(u,v)=\sum_i w(x_i,y_i)[I_0(x_i+u,y_i+v)-I_0(x_i,y_i)]^2$$

Auto-Correlaton function or surface - Compute the stable of image with respect to small variation in position *Δu* by comparing an image patch against itself.

$$E_{AC}(\Delta u)=\sum_i w(x_i,y_i)[I_0(x_i+\Delta u,y_i+\Delta v)-I_0(x_i,y_i)]^2$$

OR

b) What are Feature Descriptors? Explain the following Feature Descriptors:
ii) SIFT
iii)GLOH.

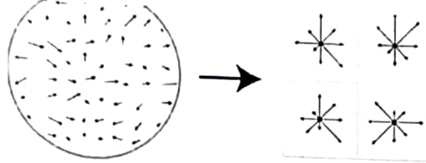
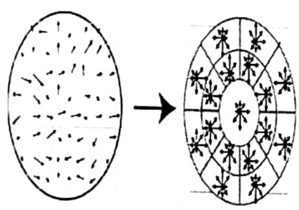
SIFT

Feature descriptors are descriptions of the visual features of the contents in images, or videos.

- They describe elementary characteristics such as
  - the shape
  - the color
  - the texture or the motion

12	2	2	2.7.1

- SIFT features are formed by computing the gradient at each pixel in a 16×16 window around the detected keypoint, using the appropriate level of the Gaussian pyramid at which the keypoint was detected.
- The gradient magnitudes are downweighted by a Gaussian fall-off function (shown as a blue circle in (Figure 4.18a) in order to reduce the influence of gradients far from the center, as these are more affected by small misregistrations.
- In each 4 × 4 quadrant, a gradient orientation histogram is formed by (conceptually) adding the weighted gradient value to one of eight orientation histogram bins

	 (a) image gradients (b) keypoint descriptor Figure 4.18 A schematic representation of Lowe's (2004) scale invariant feature transform (SIFT). (a) Gradient orientations and magnitudes are computed at each pixel and weighted by a Gaussian fall-off function (blue circle). (b) A weighted gradient orientation histogram is then computed in each subregion, using trilinear interpolation. While this figure shows an 8 × 8 pixel patch and a 2 × 2 descriptor array, Lowe's actual implementation uses 16 × 16 patches and a 4 × 4 array of eight-bin histograms. • **Gradient location-orientation histogram (GLOH)** • This descriptor, is a variant on SIFT that uses a log-polar binning structure instead of the four quadrants The spatial bins are of radius 6, 11, and 15, with eight angular bins (except for the central region), for a total of 17 spatial bins and 16 orientation bins. • The 272-dimensional histogram is then projected onto a 128-dimensional descriptor using PCA trained on a large database.  (a) image gradients (b) keypoint descriptor Figure 4.19 The gradient location-orientation histogram (GLOH) descriptor uses log-polar bins instead of square bins to compute orientation histograms (Mikolajczyk and Schmid 2005). –				
18	**a) List the approaches used to locate Boundary Curves in Images. Explain Intelligent Scissors and Level Set in detail.** **Intelligent Scissors** • Intelligent scissors system developed by Mortensen and Barrett • User draws a rough outline (the white curve in the system computes and draws a better curve that clings to high-contrast edges • To compute the optimal curve path (live-wire), the image is first pre-processed to associate low costs with edges (links between neighboring horizontal, vertical, and diagonal, i.e., N8 neighbors) that are likely to be boundary elements. • system uses a combination of zero-crossing, gradient magnitudes, and gradient orientations to compute these cost • The user traces a rough curve, the system continuously recomputes the lowest-cost path between the starting seed point and the current mouse location using Dijkstra's al-gorithm, • Breadth-first dynamic programming algorithm that terminates at the current target location • In order to keep the system from jumping around unpredictably, the system will "freeze"	12	3	1	2.7.1

inactivity **Level Set** • A limitation of active contours based on parametric curves of the form f(s) (snakes, b-snakes,…) is that it is challenging to change the topology of the curve as it evolves. • If the shape changes dramatically, curve reparameterization may also be required. • An alternative representation for such closed contours is to use ***level sets (LS).*** • – LS evolve to fit and track objects of interest by modifying the underlying embedding function instead of curve function f(s) • Level sets for closed contours • Zero-crossing(s) of a characteristic function define the curve • Fit and track objects of interest by modifying the underlying embedding function $f(x, y)$ instead of the curve f(s) • Efficient algorithm • A small strip around the locations of the current zero-crossing needs to updated at each step				
OR				

b) Illustrate the Expectation Maximization algorithm in K-means and Mixture of Gaussians..
• k-means and mixtures of Gaussians
• Model the feature vectors associated with each pixel (e.g., color and position) as samples from an unknown probability density function and then try to find clusters (modes) in this distribution.
• use a parametric model of the den-sity function
• Density is the superposition of a small number of simpler distributions (e.g., Gaussians) whose locations (centers) and shape (covariance) can be estimated
  • Meanshift is falling under the category of a clustering algorithm in contrast of Unsupervised learning
  • Assigns the data points to the clusters iteratively by shifting points towards the mode
  • Mode is the highest density of data points in the region, in the context of the Meanshift
  • Given a set of data points, the algorithm iteratively assigns each data point towards the closest cluster centroid
  • Direction to the closest cluster centroid is determined by where most of the points nearby
	12	3	3	2.7.1

PREPARED BY
Dr.J.JOSPIN JEYA

HOD/CSE
Dr.K.RAJA