CV End Sem Exam - Solutions

Q1)

$$\begin{bmatrix} 12 & 11 & 10 & 9 \\ 8 & 7 & 6 & 5 \\ 4 & 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 50 \\ 6 \\ 7 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 50\times 12 + 6\times 11 + 7\times 10 \\ 8\times 50 + 6\times 7 + 7\times 6 \\ 4\times 50 + 6\times 3 + 2\times 7 \end{bmatrix} = \begin{bmatrix} 736 \\ 484 \\ 232 \end{bmatrix}$$

Thus, $\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \frac{736}{232} \\ \frac{484}{232} \end{bmatrix} = \begin{bmatrix} 3.17 \\ 2.09 \end{bmatrix}$

row = round(v) ; col = round(u)

row = round(2.09) = 2 ; col = round(3.17) = 3

Thus final location = (2, 3)

Rubric:

1. **Full marks for correct answer.**
2. **Deduct 0.5 marks for change in the order, i.e., the student reports the answer as (3,2)**
3. **Deduct 0.5 marks for not rounding off.**
4. **Deduct 0.5 marks for not converting to cartesian coordinates and reporting the location in homogeneous coordinate system**

Q2) The transposed convolution without cropping

```
fin =

     6    25    34    15
    30   101   101    26
    42   127   131    52
    12    56   103    63
```

Cropping requires that the output size be equal to the product of input size and stride. Thus, 2*1 = 2 and hence the final output should be 2x2. Cropping is done such that the transposed output without cropping is cropped equally from all 4 directions (up, down, left and right). Thus we remove the top & bottom rows, left and right columns to produce the output:

```
   101   101
   127   131
```

Rubric:

1. **2 marks for correct transposed convolution output without cropping.**
2. **1 mark for correct answer after cropping.**

Q3) The three most popular techniques for background subtraction are

1) **Frame - Initial Frame**
   We subtract the initial frame from every frame. This is done under the assumption that the first frame doesn't contain the object, and that only the object is moving.
2) **Frame - Average Frame**
   We subtract the average of all frames from each frame. This is done under the assumption that the objects are mostly moving, so the average comprises mainly of the background. We also assume that the object is sufficiently different from the background.
3) **Frame Difference**
   We subtract the preceding frame from every frame. This method is useful for identifying changes that take place from one frame to another. This method works well for heterogeneous regions but not for homogeneous ones.

Rubric:

1. **0.75 marks for each correct way of background subtraction. The process must be clearly explained: What is subtracted from what. If that's not done, deduct 0.5 marks.**
2. **0.25 marks for at least one assumption for each.**

Q4) The binary representation for 171 is **10101011**.
Starting in a clockwise direction from the top-left pixel, i.e. 82, we get the following inequalities for $\mathcal{X}$.

- $\mathcal{X} \leqslant 82$
- $\mathcal{X} > 57$
- $\mathcal{X} \leqslant 90$
- $\mathcal{X} > 53$
- $\mathcal{X} \leqslant 92$
- $\mathcal{X} > 56$
- $\mathcal{X} \leqslant 99$
- $\mathcal{X} \leqslant 78$

Here we assume that the LBP vector takes in 1 when $\mathcal{X} \leqslant$ val and 0 when $\mathcal{X} >$ val.
For this case, the range is $\mathcal{X} \in$ **(57, 78].**

If we assume that the LBP vector takes in 1 when $\mathcal{X} <$ val, and 0 when $\mathcal{X} \geqslant$ val.
Following similar steps as above,
For this case, the range is $\mathcal{X} \in$ **[57, 78).**

Rubric:

1. **Give full marks (i.e. 3 marks) for either of the correct answers, 0 otherwise.**
2. **Deduct 0.15 marks for each number missing from the range [58,77].**

Q5)

$KA = 0, 90, 180, -90$

$$M = \begin{bmatrix} 84 & 54 & 90 \\ 78 & 66 & 54 \\ 96 & 54 & 90 \end{bmatrix} \quad A = \begin{bmatrix} -60 & 30 & 120 \\ -150 & 60 & 150 \\ -30 & -120 & 90 \end{bmatrix}$$

For, $M = 84, A = -60$

Possible key angles: $0, -90$

Thus, $\frac{2}{3}$ of 84 goes to -90, $\frac{1}{3}$ of 84 goes to 0

$\Rightarrow$ 56 goes to -90, 28 goes to 0.

For, $M = 54, A = 30$

Possible key angles: $0, 90$

$\Rightarrow$ $\frac{2}{3}$ of 54 goes to 0, $\frac{1}{3}$ of 54 goes to 90

$\Rightarrow$ 36 " " 0, 18 " " 90

For, $M = 90, A = 120$

Possible $KA \Rightarrow 90, 180$

$\Rightarrow$ 60 goes to 90, 30 goes to 180

For, $M = 78, A = -150$

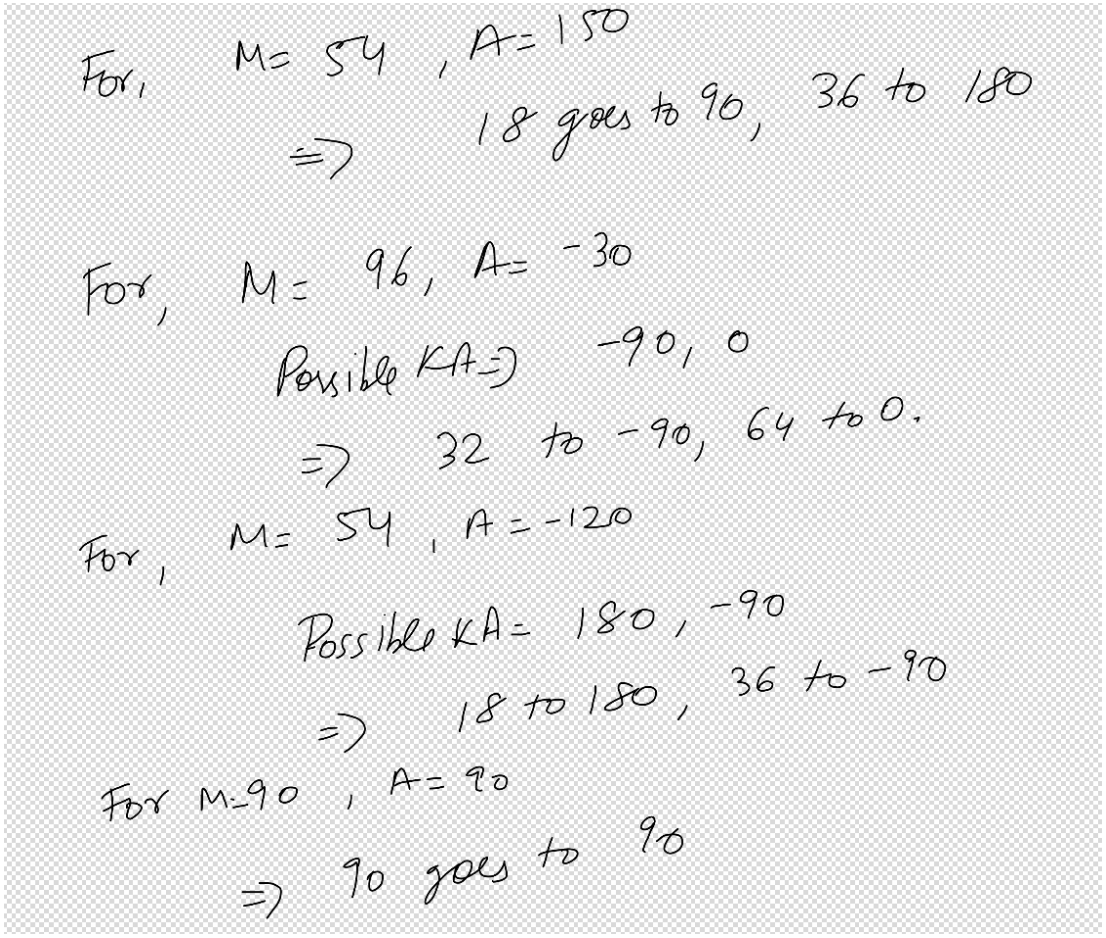
Possible $KA \Rightarrow 180, -90$

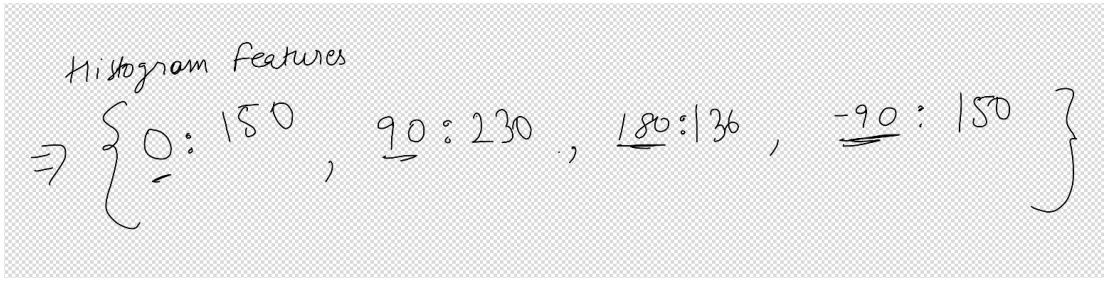
$\Rightarrow$ 52 goes to 180, 26 goes to -90

For, $M = 66, A = -60$

Possible $KA \Rightarrow 0, 90$

$\Rightarrow$ 22 goes to 0, 44 goes to 90





Rubric:

1. **For each particular pixel value, its contributions to forming the histogram is correct. 0.25 marks for each pixel. (0.25*9 = 2.25 marks) [0.1 for correctly identifying the key angles; 0.15 for correct distribution of magnitude. Apply strict binary marking]**
2. **The final histogram is correct. (0.75 mark)**

Q6) Similarity Score for any bin b is given by:

$$S(b) = \frac{Min\left(H_1(b), H_2(b)\right)}{Max\left(H_1(b), H_2(b)\right)}$$

We first create a Histogram according to the colour bins (according to colour frequencies)

For I_1 we have H_1 as follows,
{0: 4,
1: 5,
2: 5,
3: 2}

For I_2 we have H_2 as follows,
{0: 1,
1: 10,
2: 3,
3: 2}

We now calculate the similarity of each of the bins according to the above histogram values.

S(0) = $\frac{1}{4}$= 0.25

S(1) = $\frac{5}{10}$ = 0.50

S(2) = $\frac{3}{5}$ = 0.6

S(3) = $\frac{2}{2}$ = 1

Now, we compute the similarity maps (SI) for both I_1 and I_2.

$SI(I_1)$. =

$\frac{1}{4}$	$\frac{5}{10}$	$\frac{3}{5}$	1
$\frac{5}{10}$	$\frac{5}{10}$	$\frac{5}{10}$	$\frac{3}{5}$
1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
$\frac{3}{5}$	$\frac{3}{5}$	$\frac{3}{5}$	$\frac{5}{10}$

$SI(I_2)$.=

$\frac{5}{10}$	$\frac{5}{10}$	$\frac{5}{10}$	1
$\frac{5}{10}$	$\frac{5}{10}$	$\frac{5}{10}$	$\frac{3}{5}$
$\frac{3}{5}$	$\frac{3}{5}$	$\frac{5}{10}$	$\frac{5}{10}$
$\frac{1}{4}$	1	$\frac{5}{10}$	$\frac{5}{10}$

Rubric:

1. **Similarity score is correctly calculated for each bin. 0.5 mark for each correct bin. (0.5*4 = 2 marks)**
2. **Similarity maps for I1 and I2 are correct. 0.5 mark for each of I1 and I2. (0.5*2 = 1 mark). Binary marking**

Q7) We get the following sobel filter outputs without padding:

Ix=

536	729	213
132	722	616
-68	720	816

Iy=

-192	-401	-203
-188	-390	-192
0	0	0

We need Ix and Iy in 3x3 window around the centra pixel, and we already have them above.

The sum of Ix^2 for all the pixels inside the window= 2971150
The sum of Iy^2 for all the pixels inside the window= 463182
The sum of IxIy for the the pixels inside the window=-863148

Now M is given by (Assuming w = 1) using

2971150	-863148
-863148	463182

The two eigen values are:
$\boldsymbol{\lambda}_1$ = 0.195x10^6, $\boldsymbol{\lambda}_2$ = 3.239x10^6
For **Harris**
$R = \boldsymbol{\lambda}_1 * \boldsymbol{\lambda}_2 - k(\boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2)^2$
Taking k as 0.05 (Values between 0.04 and 0.06 are allowed)
R = 4.2x10^10
For **Shi-Tomasi**
$R = \min(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$
R = 0.195x10^6

Rubric:

1. **1.5 marks for correct Ix and Iy matrices to be used. Reduce 0.5 mark if some other appropriate filer is used instead of Sobel.**
2. **1 marks for correctly creating the M matrix.**
3. **1 marks for computing eigen values or det(M) and trace(M).**
4. **1 mark for Harris's R score [Note: students may take different k]**
5. **0.5 marks for Shi Thomsi's R score.**

Q8)

```
a1, multi class:
 [4.658886145103375e-15, 0.9999999999999953, 6.375869581278964e-59, 1.0112214926104438e-43]
a2, multi label:
 [1.0, 1.0, 2.8625185805493937e-20, 4.5397868702434395e-05]
```

a1 and a2 denote activations for parts 1 and part 2, respectively. Students can aslo write it as follows:
a1=[0 1 0 0]
a2=[1 1 0 0]
Or leave them as exponentials.

o1: {" dog "}

o2 : {"cat", " dog "}

o1 and o2 denote outputs for parts 1 and part 2, respectively.

Groundtruth for first part will be cat. Groundtruth for second part will be cat and mouse.

```
multi class loss:
 20.72326583694641
multi label loss:
 7.680822301655608
```

multi class loss and multi label loss denote losses incurred in parts 1 and part 2, respectively. Students can also leave these losses as infinities or in exponential form.

Formula used for multiclass loss:

$$Loss = -\sum_{i=1}^{4} y_i \log(\hat{y}_i + eps)$$

Formula used for multilabel loss:

$$Loss = -\frac{1}{4} \sum_{i=1}^{4} \left(y_i \log(\hat{y}_i + eps) + (1 - y_i) \log((1 - \hat{y}_i) + eps) \right)$$

In both eqns,

$y \rightarrow$ Real label

$\hat{y} \rightarrow$ Predicted label

$eps \rightarrow$ Very small value. (Took 10^{-8})

Rubric:

1. **1 mark each for correct a1 and a2 activation values. (1*2=2 marks)**
2. **0.5 mark each for correct o1 and o2 output classes. (0.5*2=1mark)**
3. **1 mark each for correct loss values. (1 * 2=2 marks)**

Q9)

Coordinates of the center pixel: (41,41)

After scaling by factor of two,
It becomes (82,82)
After translating by (50,50),
It becomes (132,132)

By rotating it 45 degrees clockwise,
It becomes (132*sqrt(2), 0)

Rubric:

1. **0.5 marks for correctly locating the center initially.**
2. **1.5 marks each for mentioning the correct location after every transformation. (We have three transformations, so 4.5 marks in total)**

Answers within an error of 1 should be accepted as correct.

Q10)

Original Image

7	1	2	2	2
7	1	2	2	2
6	6	6	2	2
6	6	6	2	3
6	6	6	5	3

Otsu Thresholded Image

1	0	0	0	0
1	0	0	0	0
1	1	1	0	0
1	1	1	0	0
1	1	1	1	0

GT Image

0	0	0	0	0
1	0	0	0	0
1	1	1	0	0
1	1	1	0	0
1	1	1	0	0

IoU score of segmentation: **10/12=0.833**

Now we create a tight bounding box around the segmentation result as follows.

1	0	0	0	0

0	0	0	0	0

1	0	0	0	0
1	1	1	0	0
1	1	1	0	0
1	1	1	1	0

1	0	0	0	0
1	1	1	0	0
1	1	1	0	0
1	1	1	0	0

Bounding Box is represented by the coloured pixels.
Now calculating the required values
IoU is **12/20=0.6**
CorLoc score is **1** as IoU >0.5

Rubric:

1. **2 marks for Correct Otsu Segmentation.**
2. **1 mark for correct IoU for segmentation.**
3. **1 mark for correct IoU for bounding box.**
4. **1 mark for correct CorLoc score.**