Exam

02506
Spring 2021

General information

- The exam consists of 20 questions. You only need to submit the answers to the questions. You don't need to uplaod any notes or calculations.
- There is one correct answer for each question. Some of the numeric answers have been rounded, and might slightly deviate from your result. This should not prevent you from being able to pick the correct answer.
- Each page contains one question. If there are illustrations and images, those refer to the question on that page.
- The notation in the questions is the same as in the course note.
- Data and images are provided for some questions. The question text states which data or image should be used.
- To load arrays saved as text files you can use `numpy.loadtxt` in python and `dlmread` in matlab.

Relevant links

- Course note
  `http://www2.imm.dtu.dk/courses/02506/LECTURE_NOTES_2021.pdf`
- Data for questions 4, 6, 11, 12, 14 and 17
  `http://www2.imm.dtu.dk/courses/02506/EXAM_DATA_2021.zip`

Question 1

Scale-normalized Laplacian When using scale-space for blob detection, you have computed the values given in the table below.

t	L_{xx}	L_{yy}
10	6.24	7.11
11	6.04	7.12
12	5.12	7.16
13	4.16	6.16

What is the scale-normalized Laplacian response at $t = 12$?

(a) 10.32
(b) 12.28
(c) 13.16
(d) 36.66
(e) 43.00
(f) 134.16
(g) 147.36
(h) 439.91

Question 2

Image scale-space Laplacian scale-space may be used for detecting blobs in an image I.

Which of the following statements is correct?

(a) Since the second order Gaussian derivative is separable, we just need to convolve the image with the 1D kernel $\frac{\partial^2 g}{\partial u^2} = \frac{u^2-t}{\sqrt{2\pi t^5}} \exp(-\frac{u^2}{2t})$ in the x-direction where t is the scale.

(b) To compute $\frac{\partial^2 I}{\partial x^2}$ at scale t, we can convolve the image with a 1D kernel $g = \frac{1}{\sqrt{2\pi t}} \exp(-\frac{u^2}{2t})$ in the y-direction and with a 1D kernel $\frac{\partial^2 g}{\partial u^2} = \frac{u^2-t}{\sqrt{2\pi t^5}} \exp(-\frac{u^2}{2t})$ in the x-direction.

(c) Since the image is 2D, we must convolve with a 2D kernel that is the outer products of the two orthogonal 1D kernels $\frac{\partial^2 I}{\partial u^2} = \frac{u^2-t}{\sqrt{2\pi t^5}} \exp(-\frac{u^2}{2t})$ and $g = \frac{1}{\sqrt{2\pi t}} \exp(-\frac{u^2}{2t})$ where t is the scale.

(d) Scale-space is obtained by smoothing, so the Laplacian scale-space is obtained by smoothing with two orthogonal 1D kernels $g = \frac{1}{\sqrt{2\pi t}} \exp(-\frac{u^2}{2t})$ where t is the scale.

Question 3

Blob detection In scale-space blob detection, we use the Laplacian scale space. For a small image region we have computed the linear scale-space $\nabla^2 L$ shown below using scale steps of $\Delta t = 1$. We focus on a pixel in position $(r, c) = (4, 2)$.

$\nabla^2 L$ at scale $t = 1$:

	1	2	3	4	5
1	0.3	2.3	−0.7	−5.1	−3.9
2	6.4	16.8	17.3	4.9	−4.9
3	11.5	24.8	26.8	9.0	−8.4
4	18.1	28.9	27.3	8.1	−9.0
5	15.9	25.7	23.2	7.2	−7.0

$\nabla^2 L$ at scale $t = 2$:

	1	2	3	4	5
1	0.7	2.4	1.4	−1.2	−2.7
2	6.1	10.8	9.9	3.8	−1.7
3	11.1	17.6	16.4	7.4	−1.5
4	13.3	19.9	18.1	8.1	−1.7
5	10.2	16.2	14.7	6.4	−1.8

$\nabla^2 L$ at scale $t = 3$:

	1	2	3	4	5
1	1.3	2.5	1.9	0.0	−1.5
2	5.2	7.8	7.0	3.4	−0.3
3	8.8	12.4	11.3	6.1	0.6
4	9.9	13.7	12.5	6.8	0.8
5	7.6	11.1	10.3	5.5	0.3

$\nabla^2 L$ at scale $t = 4$:

	1	2	3	4	5
1	1.4	2.3	1.9	0.5	−0.8
2	4.3	5.9	5.3	2.9	0.3
3	6.7	8.9	8.2	4.9	1.3
4	7.4	9.8	9.0	5.5	1.5
5	5.9	8.0	7.5	4.5	1.1

$\nabla^2 L$ at scale $t = 5$:

	1	2	3	4	5
1	1.2	1.9	1.6	0.6	−0.5
2	3.3	4.4	4.0	2.4	0.4
3	5.0	6.5	6.0	3.8	1.3
4	5.5	7.0	6.5	4.3	1.5
5	4.4	5.8	5.5	3.6	1.2

Which of the following statements is correct?

(a) It cannot be determined if there is a blob in the image region.
(b) A dark blob is detected in position $(r, c) = (4, 2)$ and it has a size corresponding to a circle with a radius of approximately $\sqrt{2}$.
(c) A dark blob is detected in position $(r, c) = (4, 2)$ and it has a size corresponding to a circle with a radius of approximately $\sqrt{6}$.
(d) A bright blob is detected in position $(r, c) = (4, 2)$ and it has a size corresponding to a circle with a radius of approximately $\sqrt{6}$.

Question 4

Dictionary clustering We use probabilistic clustering-based segmentation as in the lecture notes Section 3.1.

We want to compute a dictionary to prepare for the pixel-wise segmentation of images. We have four dictionary clusters with cluster centers c_1, c_2, c_3 and c_4. Furthermore we have 17 pixels labeled with either 1 or 2.

In the table below we listed pixel labels and the distances of the pixels to the four cluster centers. The arrays with distances and labels are available in the data folder as `distances.txt` and `labels.txt`.

i	1	2	3	4	5	6	7	8	
$\ell(i)$	1	2	1	2	2	2	2	2	
$d(c_1, i)$	4.4	5.9	4.9	0.4	3.3	4.4	2.5	2.6	
$d(c_2, i)$	1.2	0.4	0.4	2.5	2.6	2.9	2.2	1.8	...
$d(c_3, i)$	4.9	5.0	8.9	5.0	9.0	1.4	5.5	7.6	
$d(c_4, i)$	4.9	6.7	2.0	5.5	6.7	7.6	6.7	6.0	

	9	10	11	12	13	14	15	16	17
	1	1	1	1	1	2	2	1	1
	4.7	5.8	1.1	3.9	0.3	5.3	2.4	5.4	4.2
...	1.6	2.3	0.9	0.6	1.6	2.3	0.8	2.5	0.0
	8.5	7.1	0.7	9.5	5.4	8.8	10.1	4.3	0.6
	7.8	4.1	8.9	6.9	4.7	3.2	2.1	2.8	0.5

Which value is the closes to the probability of label 1 ($\ell = 1$) for cluster 2?

(a) 0.00
(b) 0.21
(c) 0.32
(d) 0.46
(e) 0.50
(f) 0.54
(g) 0.82
(h) 1.00

Question 5

Smooth segmentation Below we show the probability image for three labels obtained using dictionary-based segmentation.

Probability of label 1:

	1	2	3	4	5
1	0.00	0.00	0.00	0.00	0.00
2	0.00	0.00	0.04	0.00	0.00
3	0.00	0.00	0.00	0.00	0.00
4	0.00	0.02	0.02	0.02	0.00
5	0.00	0.00	0.00	0.00	0.00

Probability of label 2:

	1	2	3	4	5
1	0.02	0.02	0.01	0.01	0.01
2	0.15	0.15	0.90	0.01	0.01
3	0.15	0.15	0.02	0.02	0.02
4	0.15	0.95	0.95	0.95	0.70
5	1.00	0.97	0.97	0.97	0.70

Probability of label 3:

	1	2	3	4	5
1	0.98	0.98	0.98	0.98	0.98
2	0.85	0.85	0.06	0.98	0.98
3	0.85	0.85	0.98	0.98	0.98
4	0.85	0.03	0.03	0.03	0.30
5	0.00	0.03	0.03	0.03	0.30

To reduce noise, each probability pixel will be re-computed as the average of the four nearest pixels and the pixel itself. We focus on the pixel position $(r, c) = (2, 3)$.

Which of the following statements is correct?

(a) Without noise reduction, a pixel $(r, c) = (2, 3)$ has label 1 and with noise reduction it will have label 1.
(b) Without noise reduction, a pixel $(r, c) = (2, 3)$ has label 2 and with noise reduction it will have label 2.
(c) Without noise reduction, a pixel $(r, c) = (2, 3)$ has label 2 and with noise reduction it will have label 3.
(d) Based on the given information, it is not possible to determine the label of a pixel $(r, c) = (2, 3)$.

Question 6

Feature-based registration Two matching point sets are used for feature-based registration of two images. A point set $\mathbf{q}_i$ is from image I_q and a point set $\mathbf{p}_i$ is from image I_p, with $i = 1, ..., 30$. Points are available in data folder as `points_p.txt` and `points_q.txt`.

The model for point correspondence is

$$\mathbf{q}_i = s\mathbf{R}\mathbf{p}_i + \mathbf{t} \ ,$$

where the rotation matrix is

$$\mathbf{R} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

with angle $\theta = 140°$, the translation vector is $\mathbf{t} = \begin{bmatrix} 36 & 13 \end{bmatrix}^T$, and the scale is $s = 1.7$.

Point sets have been matched using SIFT features and they might contain outliers. We define an outlier as a point where the residual in image I_p is larger than 2 pixels. That is, an outlier is a point i where the distance between $\mathbf{p}_i$ and the transformation of the point $\mathbf{q}_i$ into image I_p is larger than 2 pixels.

How many outlier points are there?

(a) 0
(b) 1
(c) 2
(d) 3
(e) 4
(f) 5
(g) 6
(h) 7
(i) 8
(j) 10
(k) 23
(l) 30

Question 7

SIFT features Which of the following statements best describes the SIFT features?

(a) A SIFT feature is computed to be invariant to scale, rotation, and change in illumination of the image. SIFT is composed of an interest point and a descriptor where the descriptor is used for finding correspondence between images.
(b) A SIFT feature is computed to be invariant to scale but neither to rotation or the change in illumination of the image. SIFT is composed of an interest point and a descriptor where the descriptor is used for finding the scale of the feature.
(c) A SIFT feature is invariant to scale, rotation, and change in illumination of the image, only if these vary with less than $\pm 10\%$. SIFT is composed of an interest point and a descriptor where the descriptor is used for finding the orientation of the feature.
(d) A SIFT feature is computed to be invariant to scale, rotation, mirroring, and change in illumination of the image. SIFT is composed of an interest point and a descriptor where the descriptor is used for finding the orientation of the feature and the correspondence between images.

Question 8

Gaussian features We consider features from a Gaussian and its derivatives as described in the lecture note Section 3.1. Pixel features are computed for an image I of size 1024×1024. The features consist of Gaussian derivatives up to the third order at scales [1,2,4].

What is the size of the resulting feature image?

(a) $1012 \times 1012 \times 1$
(b) $1024 \times 1024 \times 3$
(c) $1012 \times 1012 \times 10$
(d) $1024 \times 1024 \times 9$
(e) $1018 \times 1018 \times 12$
(f) $1020 \times 1020 \times 15$
(g) $1024 \times 1024 \times 30$
(h) $1024 \times 1024 \times 45$

Question 9

Image gradient Image gradients may be approximated as pixel differences, e.g. by filtering using a central difference filters

$$\begin{array}{|c|c|c|}\hline 0 & 0 & 0 \\ \hline -1 & 0 & 1 \\ \hline 0 & 0 & 0 \\ \hline \end{array} \quad \text{and} \quad \begin{array}{|c|c|c|}\hline 0 & -1 & 0 \\ \hline 0 & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline \end{array} \quad .$$

Consider the 10×10 intensity image below.

	1	2	3	4	5	6	7	8	9	10
1	231	232	178	146	122	144	167	154	144	122
2	217	214	181	158	145	130	111	108	147	139
3	207	211	209	175	147	140	113	120	190	193
4	204	208	210	215	211	215	186	181	196	168
5	207	207	203	204	209	229	210	171	184	185
6	204	201	206	204	208	220	212	174	178	181
7	205	210	204	205	201	215	229	179	177	166
8	200	205	205	203	206	211	228	182	169	156
9	210	205	205	202	208	209	229	188	172	180
10	207	203	202	203	204	207	222	207	168	159

Which value is closest to the magnitude of the gradient in pixel position $(r, c) = (3, 4)$?

(a) -3534
(b) -5
(c) 59
(d) 62
(e) 84
(f) 119
(g) 3534
(h) 7093

Question 10

MRF energy change We consider a binary segmentation using MRF as in the course notes Section 5.2.

The current configuration is shown in the illustration using blue-and-red shading. The part of the image predominantly left is assigned to class blue while the part of the image predominantly right is assigned to class red.

The two classes are characterized by $\mu(\mathrm{blue}) = 30$ and $\mu(\mathrm{red}) = 20$. Smoothness parameter $\beta = 125$.

14	40	30	45	32	40
33	32	40	50	23	20
30	35	34	45	25	21
31	25	40	52	20	27
30	38	32	40	25	20

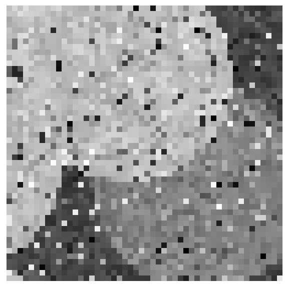
The new configuration is obtained by changing the label of one pixel, the pixel with the value 52, circled in the illustration. In the new configuration this pixel gets assigned to the class blue.

What is the change in segmentation energy $\Delta E = E_{\mathrm{new}} - E_{\mathrm{current}}$?

(a) -1258
(b) -915
(c) -790
(d) -290
(e) -165
(f) 165
(g) 275
(h) 859
(i) 1149
(j) 1758

Question 11

Posterior energy An MRF framework is used to segment an image into three segments. The image has with pixel values in the range $[0, 255]$ and is available in the data folder as `circly.png`.



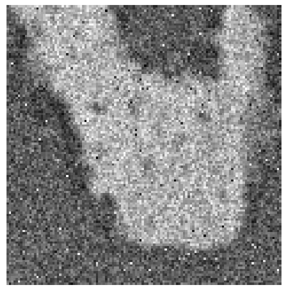
We use the same energy formulation as in the lecture notes Section 5.2. The parameters are set to $\mu(1) = 70$, $\mu(2) = 120$, $\mu(3) = 180$ and $\beta = 100$.

What is the posterior energy of the maximum-likelihood segmentation?

(a) 100100
(b) 104300
(c) 389338
(d) 883802
(e) 885862
(f) 987262
(g) 985962
(h) 988102

Question 12

Binary MRF An MRF is used to obtain a binary segmentation as in the lecture notes Section 5.4. The image is available in the data folder as `bony.png`. The parameters are $\mu(1) = 130$, $\mu(2) = 190$ and $\beta = 3000$.



What is the area of the segment 2 (the brighter part) given by the maximum-a-posteriori solution?

(a) 4337
(b) 4429
(c) 5810
(d) 5962
(e) 7320
(f) 8830
(g) 10211
(h) 10303

Question 13

Deformable models Which of the following statement is *not* correct with regards to deformable models?

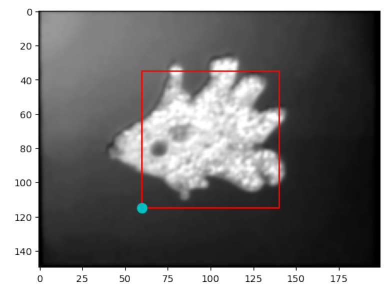
(a) Internal forces are computed based on the curve and the image intensities.
(b) For an outward pointing normal, the scalar component of the force must be negative for the curve to move inwards.
(c) External forces are computed based on the curve and image intensities.
(d) A parametric curve can be represented by points connected by straight line segments.

Question 14

External force We consider image segmentation using a deformable curve with the external force as in the lecture notes Section 6.1.

The image is available in the data folder as `frame.png`. Image values should be normalized to be in the range $[0, 1]$ by dividing the pixel values by 255 after loading.

A curve is initiated as the square placed in the center of the image, with side lengths of 80 pixels, as shown in the illustration.



What is the scalar component of the external force acting on the curve point at the bottom left corner (highlighted) of this square?

(a) -12446.85
(b) -0.48
(c) -0.42
(d) -0.34
(e) -0.25
(f) -0.19
(g) 0.25
(h) 0.39

Question 15

Curve smoothing Consider a closed curve represented by the sequence of points

$$S = ((0.1, 2.9),\ (1.2, 5.4),\ (3.3, 7.1),\ \dots,(6.2, 0.9),\ (3.5, 0.2),\ (1.4, 1.1)) \ .$$

This curve will be smoothed using an approach described in lecture notes. That is, curve points will be displaced, where point displacement is obtained by filtering the curve with a kernel

$$\alpha \begin{bmatrix} 0 & 1 & -2 & 1 & 0 \end{bmatrix} + \beta \begin{bmatrix} -1 & 4 & -6 & 4 & -1 \end{bmatrix} \ .$$

Here, we use an explicit (forward Euler) approach, with $\alpha = 0.05$ and $\beta = 0.1$.

What is the position of the first point in the sequence, originally at $(0.1, 2.9)$, after one smoothing step?

(a) $(0.280, 2.995)$
(b) $(0.490, 3.035)$
(c) $(0.340, 2.950)$
(d) $(0.520, 3.065)$
(e) $(0.640, 3.100)$
(f) $(0.100, 2.900)$
(g) $(0.250, 2.948)$
(h) $(1.600, 3.380)$

Question 16

Surface detection Consider an image I and the result of the optimal surface detection using an on-surface cost c_{on}.

6	4	3	7	6	4
6	5	9	8	6	5
7	7	3	**3**	**2**	2
6	**3**	**0**	7	5	**1**
2	4	4	6	3	5
4	5	4	5	6	6
3	5	4	4	2	7

Which settings (constraint and cost) were used to obtain this result?

(a) $\Delta_x = 0$ and $c_{\text{on}} = \text{abs}(I - 2)$
(b) $\Delta_x = 1$ and $c_{\text{on}} = \text{abs}(I - 2)$
(c) $\Delta_x = 2$ and $c_{\text{on}} = 10 - I$
(d) $\Delta_x = 3$ and $c_{\text{on}} = I$

Question 17

Optimal cost Consider an image I with pixel intensities given in the illustration (note the coordinate system with z axis pointing upwards). The array with pixel values is available in the data folder as `layers.txt`.

z						
5	2	4	3	6	2	7
4	1	5	3	2	2	6
3	10	12	14	9	15	7
2	16	16	15	11	10	17
1	6	11	5	12	11	8
	1	2	3	4	5	6 x

Layered surface detection is used to find a curve s which divides the image into the brighter lower part of the image and the darker upper part of the image. The in-region cost is given by

$$c(s, c_{\text{bright}}, c_{\text{dark}}) = \sum_{x=1}^{X} \left(\sum_{z=1}^{s(x)} c_{\text{bright}}(x,z) + \sum_{z=s(x)+1}^{Z} c_{\text{dark}}(x,z) \right) .$$

with $c_{\text{bright}} = 20 - I$ and $c_{\text{dark}} = I$.

What is the cost of the optimal solution constrained by $\Delta_x = 0$?

(a) 0
(b) 72
(c) 184
(d) 198
(e) 212
(f) 248
(g) 298
(h) 352

Question 18

Loss value A neural network is trained for classifying images into 5 classes (class 1 to class 5). The network uses softmax activation in the last layer and a cross-entropy loss. The target class for a certain image is the class 2. The output of the linear transformation in the last layer is

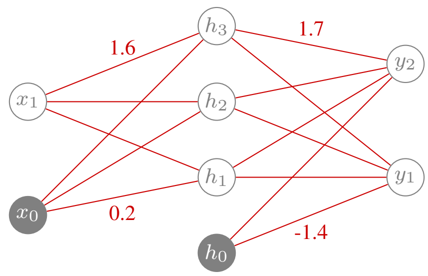
$$\hat{\mathbf{y}} = \begin{bmatrix} 0.5 & 8.2 & 6.9 & -0.1 & 0.3 \end{bmatrix}.$$

What is the loss for this one image?

(a) 0.21398
(b) 0.24185
(c) 0.65587
(d) 1.54185
(e) 1.78370
(f) 2.10413
(g) 7.94185
(h) 26.40925

Question 19

MLP output Consider a multilayer perceptron (MLP) as in the illustration.



The network uses ReLU activation in the hidden layer and softmax in the last layer. A few network weight values can be seen in the illustration, and the full network weights are given by the matrices

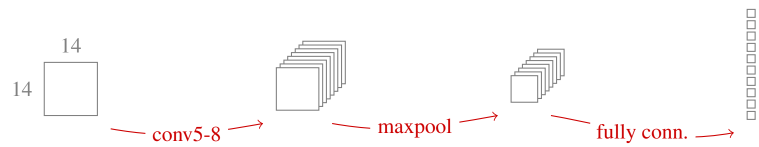
$$W^{(1)} = \begin{bmatrix} 0.2 & -1.3 \\ -0.3 & 1.8 \\ -1.7 & 1.6 \end{bmatrix}, \quad W^{(2)} = \begin{bmatrix} -1.4 & 1.5 & -0.5 & 0.9 \\ 0.2 & 1.2 & -0.9 & 1.7 \end{bmatrix}.$$

Setting $x_1 = 2.5$, what is the value of the neuron y_2 after activation?

(a) -6.0050
(b) -3.3333
(c) 0.0645
(d) 0.3300
(e) 0.5000
(f) 0.8532
(g) 0.9355
(h) 2.0775

Question 20

Network size A neural networks is used to classify images. Network architecture is sketched in the illustration (activation is not drawn).



The network is used for grayscale images of size 14×14 pixels. First is a convolutional layer using 5×5 kernels with 8 channels. No spatial padding of the input is used, and the output is restricted to the voxels where the kernel lies entirely withing the image. Second is max pooling and downscaling over a 2×2 window. Third is a fully connected layer with the output size 10.

How many learnable paremeters does the network have, including the biases?

(a) 2200
(b) 2202
(c) 2209
(d) 2210
(e) 2211
(f) 2218
(g) 3081
(h) 3082
(i) 3089
(j) 3098
(k) 4138
(l) 161530