Lecturer: Prof. Pascal Fua
Course: CS442 Computer Vision
Date:
Duration : 90 minutes

1

Student One

SCIPER : **111111**

Do not turn the page before the start of the exam. This document is double-sided, has 16 pages, the last ones possibly blank. Do not unstaple.

- Place your student card on your table.
- **A one page two-sided hand-written cheat-sheet** is allowed to be used during the exam.
- Using any **electronic device** is not permitted during the exam.
- All questions have one or more correct answers.
- The grading scheme is such that random answering is discouraged:
  - Each answer of a multiple choice question is awarded +1 point if correct and −1 point if incorrect. If the **whole** question is left unanswered no points (positive nor negative) are awarded. Note that "correct" means that a true answer should be ticked and that a false one should be left unticked.

	Correct answers:	Student's answers:	Grading:
a)	■	■	+1
b)	□	■	-1
c)	■	□	-1
d)	□	□	+1

  - The scores for separate questions are **not clipped to 0**, that is, you can get negative score for a question.
  - The multiple choice questions contribute to X points.
  - The full-text question at the end of the exam contributes X points and you cannot get negative points for answering it incorrectly.
- Use a **black or dark blue ballpen** and clearly erase with **correction fluid** if necessary.
- If a question is wrong, the teacher may decide to nullify it.

Respectez les consignes suivantes | Observe this guidelines | Beachten Sie bitte die unten stehenden Richtlinien

choisir une réponse | select an answer
Antwort auswählen

ne PAS choisir une réponse | NOT select an answer
NICHT Antwort auswählen

Corriger une réponse | Correct an answer
Antwort korrigieren

ce qu'il ne faut **PAS** faire | what should **NOT** be done | was man **NICHT** tun sollte

Pour votre examen, imprimez de préférence les documents compilés à l'aide de auto-multiple-choice.

First part, multiple choice questions

Question 1 Five image patches are shown in Fig. 1. In them, black represent zero intensity and white represent maximum intensity. Intermediate intensities are represented using shades of gray. Image gradient directions of these patches are indicated using red arrows. Select the answers where the directions are indicated accurately.

(a) (b) (c) (d) (e)

Figure 1: Image patches with gradient direction indicated with a red arrow

- [ ] b
- [ ] d
- [ ] e
- [ ] c
- [x] a

Question 2 Initialization of the live wire algorithm requires the user to select

- [ ] three edge pixels and one additional pixel closer to the edge.
- [ ] no selection is required.
- [ ] three edge pixels.
- [ ] one edge pixel.
- [x] two edge pixels.

Question 3 You are given a binary image segmentation task with input image shown in Fig. 2. You are also given a set of unordered pixels labeled as background (blue) and foreground (red), as shown in Fig. 2. Which segmentation algorithm will use this information *the best*?

(I) (II)

Figure 2: (I) Input image and set of pixels (II) Segmented image

- [ ] Simple Thresholding
- [x] Graph-cut
- [ ] 5D Mean-shift (features: R,G,B,X,Y)
- [ ] Adaptive Thresholding
- [ ] 3D Mean-shift (features: R,G,B)

Question 4 You are applying 5D Mean shift algorithm for image segmentation, that is, a feature vector consist of color (R,G,B) and spatial information (x,y). The final segmentation result depends on which following factors, assuming all others factors (listed or not) remain constant?

- [x] Color space (i.e. LAB, RGB)
- [x] Kernel width (Bandwith)
- [x] Number of iterations
- [ ] RGB to BGR conversion, applied both on input image and initial cluster centers

Question 5 In an image of a Lambertian surface illuminated by a distant point light source and without cast shadows, the intensity value of a pixel can be zero when

- [ ] the angle between surface normal and light source is less than 90^o.
- [ ] albedo is greater than a threshold specific to the surface.
- [x] the angle between surface normal and light source is greater than 90^o.
- [x] the albedo is zero.

Question 6 You are performing depth prediction from stereo cameras. Which statements regarding the narrow and wide baselines for stereo are correct?

- [ ] Wide baseline stereo is less precise.
- [x] Narrow baseline stereo makes matches easier to establish.
- [x] Wide baseline stereo usually produces more occlusions.
- [ ] Narrow baseline stereo is more precise.
- [ ] Narrow baseline stereo usually produces more occlusions.

Question 7 You are given two images I_1 and I_2 taken from a stereo setup. Additionally, you are given a small patch x on image I_1. You want to calculate the disparity between I_1 and I_2 around the location of patch x. For this, you need to search for the patch x in image I_2, with the normalized cross correlation metric. Where should you search for the correspondence in I_2, without missing any matches, at the same time speeding up the search?

- [ ] on the epipoles
- [ ] in the same column as that of x
- [ ] in the local neighborhood of x
- [x] along the epipolar line
- [ ] in the same row as that of x

Question 8 Which of the following camera parameters should be explicitly known to project a 3D mesh whose vertices are expressed in the camera coordinate system onto an image plane?

- [ ] Camera coordinate center in world coordinates.
- [x] Camera intrinsic including focal distance and principal point.
- [ ] Camera rotation angle.
- [ ] Camera translation.
- [ ] The matrix to find the epipolar line in another camera view.

Question 9 The visual hull algorithm

- [x] can be used to correct the crude silhouette image of one camera using the other available cameras.
- [ ] can capture the texture of the object.
- [ ] can capture concavities of the true object shape.
- [ ] can generate a bigger 3D volume for the object using multiple cameras than using a single camera.

Question 10 Which type of layers/operations are least likely to be found in a standard deep neural network used for image segmentation?

- [ ] Average-pooling
- [x] Fully connected layers
- [ ] Convolutional layers
- [ ] Max-pooling

Question 11 Which of the following are true about a fully convolutional neural network with a standard bottleneck architecture used for segmentation?

- [x] They can have skip connections to recover the spatial information lost during downsampling.
- [ ] They cannot be applied to images of any size.
- [ ] The initial layers tend to detect global features (like objects or parts of objects) whereas the higher layers detect local features (like edges).
- [ ] The input and output spatial feature dimensions (height and width) of a convolutional layer cannot be the same.