554SM – COMPUTER VISION AND PATTERN RECOGNITION

Written Examination

Name: .....

Student Number: .....

Instructions:

- Answer the multiple-choice questions (for each question, *only one choice is correct*).
- Answer the essay question.
- Fill in the answers to the multiple-choice questions on the answer sheet (last page).

1. In the bilateral filter, the soft outlier rejection is due to

   (a) the conformal mapping
   (b) the domain kernel
   (c) the range kernel
   (d) none of the above

2. Suppose image B is obtained by applying a homography to image A. Suppose you want to use RANSAC for estimating the homography, based on matched keypoints in the two images. How many correspondences (matched pairs) you need to sample at each RANSAC iteration?

   (a) 2
   (b) 8
   (c) 1
   (d) 4

3. A convolutional layer of a ConvNet has a number of activation maps that depends on

   (a) the size of each filter of that layer
   (b) the number of activation maps of the previous layer
   (c) the number of filters of that layer

4. In the Viola-Jones detector, the early stages of the cascade, typically

   (a) employ less features than the last stages
   (b) employ roughly the same number of features of the last stages
   (c) employ more features than the last stages

5. If we replace one lens on a calibrated stereo pair of cameras with a bigger one, what can we say about the essential matrix, E, and the fundamental matrix, F?

   (a) both are unchanged
   (b) E can change due to a possible change in the lens characteristics, F is unchanged
   (c) both E and F can change due to a possible change in the lens characteristics
   (d) F can change due to a possible change in the lens characteristics, E is unchanged

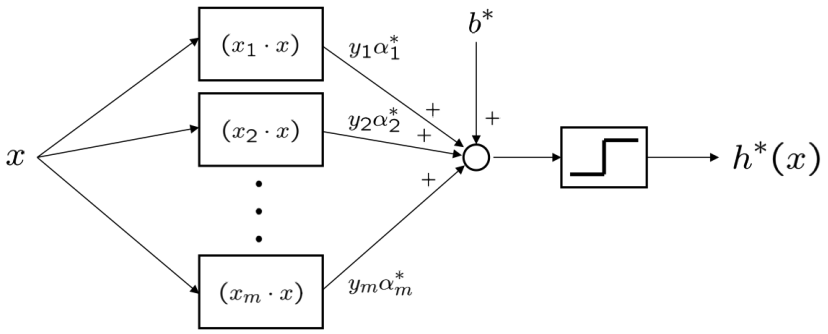
6. In training convolutional neural networks, the stochastic gradient descent consists of

   (a) stochastically perturbing the weights at the beginning of each epoch
   (b) estimating the gradient, at each iteration, based on a subset of randomly chosen examples
   (c) alternating, at each iteration, a step toward the opposite of the gradient and a step in a randomly chosen direction
   (d) none of the above

7. Which of the following statements is **wrong**?

   (a) filtering in the frequency domain results always in faster computation
   (b) convolution in spatial domain is equivalent to multiplication in frequency domain
   (c) using the Fourier transform, an image may be represented as a weighted sum of oriented sinusoids
   (d) none of the above

8. The figure below



(a) represents an SVM in primal form
(b) represents a linear perceptron
(c) represents an SVM in dual form
(d) none of the above

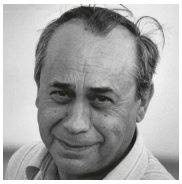
9. Suppose image B is obtained by translating image A of a certain amount in the plane. Suppose you want to use RANSAC for estimating the translation, based on matched keypoints in the two images. How many correspondences (matched pairs) you need to sample at each RANSAC iteration?

(a) 1
(b) 2
(c) 4
(d) 8

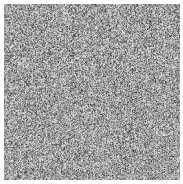
10. A mobile robot is moving laterally and takes two successive pictures (the optical axis of the camera is orthogonal to the moving direction). Then

(a) the epipolar lines are parallel, in both the images
(b) the epipoles are located in the center of the respective image
(c) at least one epipole is not at infinity
(d) none of the above

11. Suppose you have a training set of N face images $\mathcal{S} = \{I_1, \ldots, I_N\}$ of size 256×256. You apply the Eigenfaces approach and you get some performance in recognizing the identity of the faces in the test set $\mathcal{T} = \{J_1, \ldots, J_M\}$. Let define a permutation operator $\phi(\cdot)$ that permutes the pixels of 256×256 images (an example is shown below).



I



$\phi(I)$

Now you apply the same approach to the training set $\tilde{\mathcal{S}}_1 = \{\phi(I_1), \ldots, \phi(I_N)\}$. What can be said about the performance obtained on the test set $\tilde{\mathcal{T}} = \{\phi(J_1), \ldots, \phi(J_M)\}$?

(a) in general, nothing can be said
(b) it is certainly the same
(c) it is likely to be worst
(d) it is likely to be better

12. A training set is not linearly separable. You apply a soft margin SVM with linear kernel and $C = 1$, obtaining a certain number of support vectors, say M. Then, you set $C = 1000$ and train the SVM again, obtaining $\hat{M}$ support vectors. Which of the following is likely to occur?

    (a) $\hat{M} > M$
    (b) $\hat{M} = M$
    (c) $\hat{M} < M$

13. What is the missing value of the following distance transform map?

0	0	0	0	0	0	0
0	0	1	0	0	0	0
0	1	2	1	0	0	0
0	1	2	2	1	1	0
0	1	2		2	1	0
0	0	1	1	1	0	0
0	0	0	0	1	0	0

    (a) 2
    (b) 1
    (c) $\sqrt{2}$
    (d) none of the above

14. Which of the following statements about the fundamental matrix F is **wrong**?

    (a) F is full-rank
    (b) F can be estimated from point matches in pixel coordinates only
    (c) F encodes information on both the intrinsic and the extrinsic parameters
    (d) F defines a relationship between homogeneous pixel coordinates of conjugate points in a pair of images

15. In the scale-space approach, the derivatives need to be normalized because

    (a) otherwise the derivatives could not be computed exploiting the separability
    (b) otherwise the derivatives would depend on the average intensity of the image
    (c) otherwise the derivatives at different scales could not be compared
    (d) none of the above

Essay question: describe the bag-of-words approach for image classification.

Answer (do not exceed the frame below):

Answer Sheet

Question#	Answer
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	

The space below is reserved to the instructor

multiple choice (0 to 15) _____

essay question (0 to 5) _____

oral discussion (0 to 5) _____

project (0 to 5) _____

total _____