554SM – COMPUTER VISION AND PATTERN RECOGNITION

Written Examination

July 01, 2019

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).

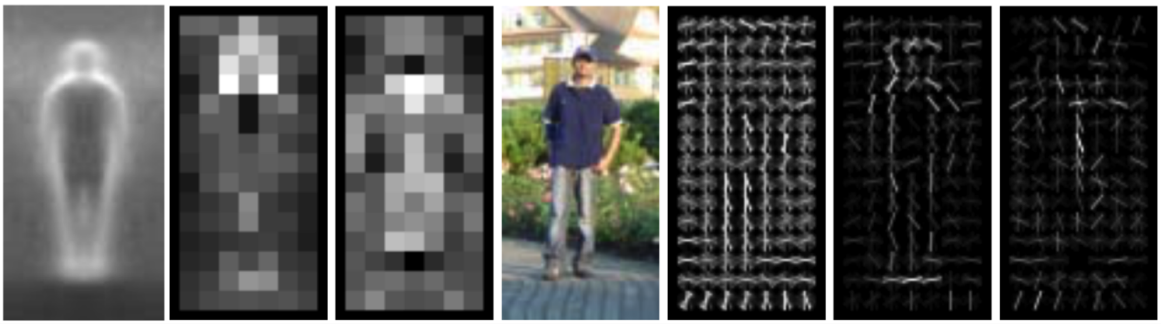
1. The number of decision variables of the optimization problem of soft margin SVM in *dual* form is

   (a) equal to the number of examples
   (b) equal to the dimension of the input space plus the number of examples plus one
   (c) equal to the dimension of the feature space
   (d) none of the above

2. In applying the empirical risk minimization principle, when the capacity of the hypothesis set increases

   (a) the bias is likely to increase
   (b) the variance is likely to decrease
   (c) none of the above

3. Consider the following figure, taken from the original Histogram of Oriented Gradients paper.



(1) (2) (3) (4) (5) (6) (7)

   Which of the following captions refers to image (7)?

   (a) average gradient image over training examples
   (b) HoG descriptor weighted by negative SVM weights
   (c) HoG descriptor
   (d) HoG descriptor weighted by positive SVM weights
   (e) none of the above

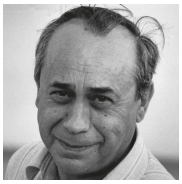
4. Which of the following statements is **wrong**?

   (a) telecentric optics perform orthographic projections
   (b) in images produced by telecentric optics, the apparent size of the objects is independent of their distance from the sensor
   (c) telecentric optics block all the light beams, except those parallel to the optical axis
   (d) none of the previous is wrong

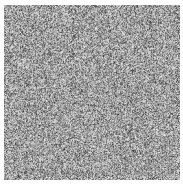
5. Suppose that, by applying the Harris detector to a specific location of an image, you get a second moment matrix having eigenvalues $\lambda_1 \approx 0$ and $\lambda_2 \gg 0$. Then, that location corresponds to

   (a) an edge
   (b) a corner
   (c) a flat zone
   (d) none of the above

6. Suppose you have a training set of N labeled face/non-face images $\mathcal{S} = \{I_1, \ldots, I_N\}$ of the same size, say 256×256. You train a linear SVM (the input vectors being the raw pixel intensities) to distinguish between face and non-face and you get some performance on 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 train a linear SVM on 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) it is likely to be the same
(b) in general, nothing can be said
(c) it is likely to be better
(d) it is likely to be worst

7. Which of the following statements about instance recognition from local features is **wrong**?

(a) its performance drops dramatically in the presence of even small occlusions
(b) it is usually accompanied by a geometric consistency check
(c) it is best suitable for planar objects, or deformations explained by affine transformation
(d) none of the above

8. A mobile robot is moving forward and takes two successive pictures (the optical axis of the camera is parallel to the moving direction). Then

(a) the epipoles (thought of as image points) have the same location in both the images
(b) the epipolar lines are parallel, in both the images
(c) since it's a pure forward motion, both the epipoles are at infinity
(d) none of the above

9. The depth of field of a camera

(a) increases as the lens diameter decreases
(b) increases as the lens diameter increases
(c) is independent of the lens diameter

10. You are implementing an algorithm that requires to compute, from point matches, the essential matrix E of a pair of cameras and the fundamental matrix F of another pair of cameras. Suppose you are forced to use the normalized eight-point algorithm for just one of the computations, and the standard eight-point algorithm for the other. What would you choose?

(a) using the normalized eight-point algorithm for computing F
(b) using the normalized eight-point algorithm for computing E
(c) there is no reason to prefer any of the previous choices

11. Suppose we want to describe a keypoint using a basis of 14 steerable filters. The length of the descriptor will be

(a) it depends on the size of the filters.
(b) 14
(c) 14^2

12. Under affine transformations

    (a) lengths are preserved
    (b) parallelism is preserved
    (c) angles are preserved
    (d) none of the above

13. The number of degrees of freedom of a perspective projection matrix is

    (a) 11
    (b) 12
    (c) 8
    (d) none of the above

14. Which of the following statements about the essential matrix E is **wrong**?

    (a) E has rank 2
    (b) E expresses a constraint on the normalized coordinates of conjugate points as a function of the relative pose of the two cameras
    (c) E encodes information on the extrinsic parameters only
    (d) E has six degrees of freedom

15. Suppose you are applying RANSAC to estimate a transformation between two images and the model requires $n = 1$ points. Suppose that the fraction of inliers is $w = 0.5$. How many iterations are needed to get a probability of failure of 0.125?

    (a) 8
    (b) 2
    (c) 3
    (d) 1

Essay question: describe the scale-space approach to keypoints detection.

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 _____