Examination questions from Machine perception and image analysis

Václav Hlaváč
Czech Technical University in Prague
Czech Institute of Informatics, Robotics and Cybernetics
160 00 Prague 6, Jugoslávských partyzánů 3, Czech Republic
http://people.ciirc.cvut.cz/hlavac, vaclav.hlavac@cvut.cz

December 15, 2022

The written test consists of six questions. Answers to each of the questions is assessed by five points maximally. Questions will be randomly selected from the following list. Some of questions may be modified or replaced by other questions. The list of questions may be extended or modified gradually.

---

1. What is the difference between image analysis (also computer vision) on one side and computer graphics on the other side? Give two examples, which illustrate the difference.

2. Interpretation (understanding) of the image can be expressed mathematically using a formal languages theory as *mapping*: *observed image data* $\rightarrow$ *model of a theory*. The model of the theory matches to a specific world, in which the "theory" is valid. A single "theory" may match to more than one world. Interpretation can be understood also as the *mapping*: *syntax* $\rightarrow$ *semantics*. Semantics represents knowledge about a specific world. We understand usually in image analysis that images contain certain objects.

   Give two practical examples of image analysis tasks, in which interpretation is used. With each of the examples, be specific how is the semantics utilized.

3. Signal analysis and low level image processing does not usually interpret analyzed image data. Explain (preferably using mathematical formalism), what the interpretation is. What gain does the interpretation provide on one side and what constraints does it bring on the other side?

4. Why is it difficult to "understand" general 3D scenes in computer vision? Give several reasons and outline them by a short comment. (Six reasons were mentioned in the lecture.)

5. Local and global processing.

   - Discuss briefly the difference between the local and global approach to image analysis. Mention advantages and disadvantages of both approaches.
   - Give two examples od local operations and comment it briefly.
   - Give two examples od global operations and comment it briefly.

6. Explain the notion of the image (continuous image function) $f(x, y)$ or $f(x, y, t)$. What do parameters x, y, t correspond to? Give several examples of 'real life' images captured using different physical principles.

7. What is the image quantization? How and in which device is the quantization performed? Explain the principle. How many quantization levels does a young healthy person distinguishes in a grey level image? What does the person sees in the image, which has less quantization levels than needed?

8. Consider digitization of a 2D image. The distance between equidistant samples is set according to Shannon sampling theorem similarly as in the 1D signal case. In a 2D image case, it is necessary to solve one more issue (relation) besides finding the distance of samples. Which issue is it? How is it solved typically. Which advantages and disadvantages these solutions have? Let me note that I do not talk about intensity quantization here.

9. The contiguity relation between two pixels of a binary image (there is a path between them) defines the image (set) into equivalence classes. Which three properties must such relation fulfil in order to be equivalence? Verify these three properties for the continuity relation in 2D binary image.

10. (a) Define (i) the region and (ii) the convex region in a 2D image. Draw the example of the convex and non-convex region.
    (b) Define the convex hull.
    (c) Draw the convex hull for the example of the non-convex set, which you used in the item (a) of this question.

11. Explain what is meant in relation to images by (a) spatial resolution; (b) spectral resolution; (c) radiometric resolution; and (d) time (also temporal) resolution.

12. Write the definition formula of the Shannon (also information) entropy. Explain variables in the formula. What is the Shannon entropy used for in general? Consider a gray scale image. Give at least two applications of the Shannon entropy in digital image processing.

13. Define distance transform. Definition in words suffices. Sketch the example of a binary image in square raster of the size 6×6. Put two nontouching objects into this image. Each object has the size three pixels. Use 8-neighborhood. Fill the rest of the image by the distance transform values expressing the distance.

14. Why is the distance transformation algorithm so important in image analysis? What is it used for in applications?

15. When capturing an image of a 3D world using a camera, the involved geometry is represented by a pin-hole camera model (also called perspective projection). 3D point (x, y, z) projects into the image plane as (x', y'). Draw a corresponding figure (a planar one suffices, do not draw, e.g. y-coordinate). Assume that you know 3D coordinates (x, y, z), the focal length f, i.e. the distance of the image plane from the center of projection. Derive $x' = \ldots$.

16. What is the main role of the lens in a (photo)-camera? Describe the role of the lens from the physical point of view informally.

17. The lens of a camera is often explained by a simplified geometric optics model (theory) in practice. What are the preconditions enabling to use this model?. I note that the more complex model is the wave optics model.

18. Compare properties of a pin-hole camera and an (lens) objective composed of individual lenses.

19. Explain what is the natural vignetting. Is the natural vignetting more pronounced for wide-angle lenses or for tele-lenses. Explain (derive in the better case), what is the cause of natural vignetting.

20. Explain what is the radial distortion of a lens. How the radial distortion manifests itself. Can it be corrected? If so, how.

21. Explain the telecentric lens principle. What is it used for?

22. Explain what is depth of focus (or depth of field) for the optical lens. Which parameter (usually controllable) parameter of the lens allows to change it?

23. Imagine capturing of a 3D scene. The elementary surface patch in the scene reflects certain radiation L into a camera. The irradiation E on the light sensitive camera chip in the pixel x, y is directly proportional to the image function $f(x, y)$ value ($\approx$ brightness). On which elementary patch and light sources properties does the value of $f(x, y)$ depends for the chosen x, y? Formulate the preconditions of this radiometric model and sketch the corresponding figure.

24. Explain the concept "bidirectional distribution function" of an image abbreviated as BRDF. Why and for what is BRDF good for?

25. - What reflectance properties has the Lambertian surface?
    - What is the simplification by a Lambertian reflectance model used for? Give two examples at least, please.

26. What solves the irradiance equation in radiometry? Formulate the task and the basic thought of its derivation. The formulas are not necessary. (It might help you if you draw a sketch and denote variables in it.)

27. Characterize the image preprocessing. What constitute its input and output? What is the image preprocessing good for? List three examples of image preprocessing methods. Names of methods suffice.

28. Characterize 2D convolution. For what is 2D convolution used in digital image processing?

29. Write the formula for smoothing the histogram h_i, $i = 0, \ldots, 255$ using the sliding average. Use the square window of the size $2K + 1$ pixels with its representative pixel in the window center.

30. Give the formulae for the one-dimensional Fourier transform and its inverse. Explain informally the underlying ideas.

31. What is the asymptotic computational complexity of a one-dimensional Fourier transformation as a function of the length n of a discrete input signal (sequence)? Use the 'capital $\mathcal{O}$' notation.

32. Explain what is a two-dimensional Fourier transformation, its difference to one-dimensional Fourier transformation. (Use either the definition formulas or informal explanation.) How is 2D Fourier transformation utilized in image processing?

33. Fourier transform is defined for periodic signals. Many practical signals, we work commonly with, are not periodic. Name and explain informally two principles which are often used to overcome this limitation.

34. Formulate Shannon (also Nyquist, Kotelnikov) sampling theorem for the simpler 1D signal case. Explain (informal explanation suffices, the drawing helps you). What is the idea of sampling theorem proof? (Hint: frequency spectra).

35. Taxonomize image preprocessing methods into four groups according to the used neighborhood expressed with respect to its representative pixel. Give example of a preprocessing method for each of the groups.

36. Explain the principle of brightness corrections (usually used to remove systematic defects in image acquisition) when considering a multiplicative disturbance model. Express mathematically.

37. Homogeneous coordinates are often used when expressing geometric transformation of an image. Explain what are the homogeneous coordinates. What advantage homogeneous coordinates bring when expressing affine geometric transformation. (Hint: Recall the language describing a page named PostScript).

38. Explain the rationale of histogram equalization.

39. Consider an affine planar geometric transform (comprising the scale change, rotation, translation and sheer) given by the equation

$$\begin{aligned} x' &= a_0 + a_1 x + a_2 y \,, \\ y' &= b_0 + b_1 x + b_2 y \,. \end{aligned} \tag{1}$$

    (a) How many control points do you need to know at least if you like to calculate coefficients of the affine transformation in (1)?
    (b) More control points are used often practically, which corresponds to the overdetermined system of equations (1). Why the redundant number of control points is used?

(c) What is the name of the method used commonly to solve the overdetermined system of equations?

40. When using geometric transforms for discrete images there is need to approximate the value of the image function $f(x, y)$. Why? List two approximation methods at least (suggestion: draw a figure, provide a formula, . . .).

41. Explain the image function value (brightness) interpolation after applying the nearest geometric transform. Consider the nearest neighbor and linear interpolation methods.

42. Explain the principle of the bicubic brightness interpolation after a geometric transformation. How many samples of the image function will you need? It will be probably easier for you and me if you sketch a simple drawing.

43. Consider filtration of a random additive noise in the image. The correct value can be estimated as the mean (arithmetic average) of n noisy values. How many times is the amplitude of the noise dumped by the filtration if the standard deviation of the noise was σ originally? Explain what is the statistical principle of such noise dumping. (A hint: the central limit theorem.)

44.

1	0	0	15	15
0	1	0	15	14
0	1	1	4	15
1	0	0	15	15
2	3	0	15	14

The image on the left depicts a cut-out from an image function. Bold border delimits a neighborhood, the domain of a convolution kernel, from which a new filtered value has to be calculated. The representative pixel of the neighborhood is in the center of the convolution mask and is denoted by a cross. Calculate the filtrated value by (a) smoothing using averaging; and (b) median filtration.

45. Consider linear image filtration performed by a convolution with the 11×11, which approximates Gaussian filter. Is this operation linear? Explain your decision mathematically.

46. Explain concepts in the image domain and their relations: edge, edge element (abbrev. edgel) and boundary of a region.

47. Formulate the 2D image segmentation task. What constitutes the input and what the output? Name two different segmentation methods and explain briefly the algorithm they use.

48. In the 2D image segmentation context, explain concepts: region and object. Select a specific segmentation method and write what is the relation between the region and the object in this particular method.

49. Segmentation is based on semantics of a specific task, i.e. on the ability to explore the prior knowledge leading to the image interpretation. Explain on the example how semantics simplifies image analysis.

50. Explain concepts entire and partial segmentation using a mathematical formalism. Give one example for entire and one for partial segmentation.

51. The back illumination is often used to simplify segmentation, e.g., in digital profile projectors in industry. It is possible to gauge dimensions or deviations from correct shape more precise than it follows from the Shannon sampling theorem. A sub-pixel accuracy can be achieved. Explain the principle and demonstrate it on the example.

52. Divide image segmentation methods into categories, say four ones. Name and briefly describe each of them.

53. When segmenting by intensity thresholding, it is desirable to set the threshold value automatically. It is often possible when sought objects differ significantly in intensity from the background. How is the optimal threshold set in this case? When the method fails?

54. It is possible to use the outcome of the edge detector for segmentation (for finding region boundaries). Describe the idea of such an approach. What problems this approach has? How are these problems mitigated?

55. Explain the k-means method and its principal properties. How can this method be used for image segmentation?

56. Explain the image segmentation using mean-shift principle.

57. Explain what statistical pattern recognition (aka machine learning) is. What is the goal? What is input? What is the output?

58. Briefly explain the terms object (pattern), object description, features, feature vector, hidden state (more specifically, classification class), classifier.

59. Explain why in statistical pattern recognition (aka machine learning) classifiers are most often taught empirically.

60. Consider the statistical recognition (machine learning) task. Explain classifier teaching with teacher and teaching without teacher. Give at least one example of use for each.

61. Consider statistical pattern recognition (machine learning). State the difference between classification and regression.

62. In statistical pattern recognition, it is considered that the observed objects can be expressed as a point in a multidimensional vector space. What are the coordinates of this point which correspond to an object?

63. Formulate a Bayesian decision task. What are the prerequisites for using a Bayesian approach? What is the statistical model in the Bayesian formulation of a considered decision making task?

64. Why is the Bayesian task of statistical decision making, despite its generality and beauty, not immediately practical? Hint: it has to do with a statistical model.

65. What tasks are known in statistical pattern recognition? Formulate basic tasks.

66. What is the difference between a set and a multiset? Explain why the training data is a multiset and not a set?

67. Experimental evaluation of the classifier. What is the confusion matrix for the case of classification into two classes (also dichotomous classification)?

68. Experimental evaluation of the classifier. What is the confusion matrix for the case of classification into more than two classes? Give an example of a simple task and the error matrix for it.

69. Experimental evaluation of the classifier. ROC-curve: explanation of concept and use.

70. Explain the principle of a linear classifier that is learned by a perceptron algorithm. When does the algorithm converge?

71. Consider a linear classifier and its learning by a perceptron algorithm. What are the ideas of perceptron algorithm convergence proof (Novikoff theorem informally)?

72. Explain the basic idea of a linear classifier based on the principle of support vectors (SVM - Support Vector Machine) for the simple case of a linearly separable training set. Perhaps an illustrative image will help. How to deal with the situation when the data in the training set is not separable?

73. Unsupervised learning and the EM algorithm. Formulate the task and explain the principle of the algorithm. Does the algorithm converge to a global solution and why?

74. 3D počítačové vidění. Formulujte úlohu (případně několik úloh). Co je vstupem, co je výstupem úloh 3D počítačového vidění?

75. David Marr (1945-1980, author of a new model of image visual perception, influential in the field of computational neuroscience) formulated the task of 3D computer vision, which allowed it to be investigated by scientific methods. Explain his basic ideas. Rely on the concepts: input iconic image, primal sketch, two and a half dimensional sketch, 3D representation related to the object.

76. Marr's model of understanding the visual image of a 3D scene has remained the only widely accepted theory since 1980s, although its applicability is very limited. Marr's stimulus led to the development of methods collectively called Shape from X. List at least five such methods.

77. One of the widely used results of Marr's model of understanding the visual image of a 3D scene is the shape from shading, also the shape from motion (SfM). Formulate what task SfM solves.

78. Explain what projective space is.

79. What are homogeneous coordinates? What is the main advantage of using homogeneous coordinates when using them for projection?

80. What is homography? What are the two most common uses of homography in computer vision? Formulate these two tasks and write down what they are used for.

81. 2D homography maps a plane to a plane.

$$\alpha \begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} = H \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}, \quad \alpha \neq 0,$$

where $H - [3 \times 3]$ is the homography matrix. Explain how you can calculate homography matrix entries $h_{11}, \ldots, h_{33}$ knowing corresponding points in two images related by homography. Describe the idea. Do not derive. Formulas are not needed.

82. The projective transformation can be expressed in matrix form as M having three-row and four columns. What is the input and output of the projective transformation expressed in this way. Write a relationship.

83. The projective matrix M can be factorized (i.e., expressed as the product of partial matrices).

$$\mathbf{u} \simeq M\,\mathbf{X} = K \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & -R\,\mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix} \mathbf{X}\,.$$

Name the sub-matrices and write what they are for.

84. What is the matrix of internal calibration parameters K? Write it and explain its parameters.

85. The external calibration parameters are described by the rotation matrix R and the vector $\mathbf{t}$. Explain what these parameters are for. How are external calibration parameters determined?

86. Consider two perspective cameras. What is epipolar constraint? Shat is its significance? Express with a picture.

87. How is epipolar constraint expressed mathematically? (hint: bilinear transformation) Write it.

88. Formulate the role of correspondence in the analysis of (intensity) images. What is the role of correspondence looking for? What is its input? What is its output?

89. What are correspondence search algorithms used for (intensity) image analysis and 3D computer vision?

90. The role of correspondence of (intensity) images in 3D vision tasks is in principle ill-conditioned. Write down three possible causes for this ill-conditionality.

91. What algorithms are used to search for correspondences between two intensity images? Sketch the principle of two qualitatively different algorithms. The first algorithm assumes nothing more than that some of the images overlap. The second algorithm will use certain (discussed in the subject) knowledge to simplify the correspondence search algorithm.

92. What sensors used in robotics, such as in self-driving cars, while measuring depth in the image and working on the principle of energy flight time?

93. The result of capturing a depth map with, for example, a stereo camera, lidar or radar is a point cloud in the observer's coordinate system. A depth map is a point cloud with (x, y, z) coordinates. What data, relations are we missing from such a depth map in order to work with it further?
94. The Iterative Closest Points (ICP) algorithm is used to register overlapping point clouds (or depth maps). Explain the assumptions of its use, the basic idea, and the properties. What prerequisites must the data meet in order to use the ICP algorithm?
95. What sensors used in robotics, such as in self-driving cars, while measuring depth in the image and working on the principle of energy flight time?
96. Explain briefly the principle of operation of radar.
97. Explain briefly the principle of operation of lidar.
98. Consider this robot definition: "A physically-embodied, artificially intelligent autonomous device, which can sense its environment and can act in it to achieve some goals." Explain briefly the following concepts: physically-embodied, artificially intelligent and autonomous.
99. In terms of artificial intelligence and robotics, what skills should intelligence provide? Just write the names of the skills. (Hint: I listed eight in the lecture.)
100. Robots influence their world with effectors. Explain the difference between an actuator and an effector. Give two examples of action members.
101. Three types of drives are most often used in industrial robots: electric, hydraulic and pneumatic. Write the main advantages and disadvantages of each.
102. There are many types of electric drives (motors). Write the basic breakdown of electric motors in the form of a tree. For the most important engines in your classification, briefly write its advantages and disadvantages.
103. Explain the term direct kinematic task. What is its input and output? Can it be calculated by the formula for an open kinematic chain (usual industrial manipulator)?
104. Explain the term inverse kinematic problem. What is its input and output? Can it be calculated by the formula for an open kinematic chain (a common industrial manipulator) with six degrees of freedom? Comment on the circumstances of this calculation.
105. When we need to plan the movement of the robot, we distinguish between path planning (global plan) and trajectory (local plan). What algorithms are used for path planning? List and briefly describe the principle.
106. When we need to plan the movement of the robot, we distinguish between path planning (global plan) and trajectory (local plan). What algorithms are used to create the trajectory? List them and briefly describe the principle.
107. Pro reprezentaci pohybu robotu v jeho okolí se používá konfigurační prostor ($\mathcal{C}-$space). Vysvětlete, co je to konfigurační prostor. Jaké výhody a nevýhody má jeho použití?
108. Draw a block diagram of a cognitive (deliberative) robot and feedback loops in it.
109. Robot control architectures. Deliberative, also classical vs. reflexive. Briefly explain both and give advantages and disadvantages of both.
110. Feedback and its importace in general. Then, more specifically, what does feedback enable in the control of dynamic systems? How does the introduction of feedback affect the knowledge requirements of the controlled system model? Give an example.
111. What is a force/torque compliant robot (manipulator)? What is force/moment compliance achieved? Which new force/torque compliant applications do they bring?
112. What is mechanical impedance? How is it similar to electrical impedance?
113. How to create an active compliant actuator (hint: e.g., one degree of freedom of a force/torque compliant robot)? Consider only mechanical binding.
114. State the basic principles of tactile sensors and indicate their principle.