Examination questions, subject – B4M33DZO Digital image

Václav Hlaváč
Czech Technical University in Prague
Czech Institute of Informatics, Robotics and Cybernetics
166 36 Prague 6, Jugoslávských partyzánů 3, Czech Republic
`http://people.ciirc.cvut.cz/hlavac, vaclav.hlavac@cvut.cz`

1. října 2018

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* → *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* → *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 int he 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.
8. 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?
9. 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 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 quantization here.

10. Assume that hexagonal raster (resembles a honeycomb) is used for image sampling. What advantages does this sampling has? Why is this sampling not used in most of image digitizers and digital cameras?

11. 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.

12. (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.

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

14. 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.

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

16. Explain three concepts: image contrast, image color saturation, image sharpness.

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

18. Even we have no information about image data interpretation, we can estimate the information content in the image using Shannon entropy. Consider a gray scale image. Calculate Shannon entropy for the image with $2b$ gray levels of the size $N \times N$ from the gray-level histogram $h(i)$, $i = 0, \ldots, 2b - 1$. Which histogram yields the highest entropy?

19. Consider a gray scale image. Write a definition formula for Shannon entropy. Your input is a gray scale histogram h_i, $i = 0, \ldots, 255$. It is known that the image preprocessing does not increase entropy. What do we have to do if we need an image with higher entropy?

20. 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 lenght f, i.e. the distance of the image plane from the center of projection. Derive $x' = \ldots$.

21. 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.

22. 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.

23. Compare properties of a pinhole camera and an (lens) objective composed of individual lenses.

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

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

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

27. 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?

28. 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.

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

30.
    - What reflectance properties does the Lambertian surface have?
    - What is the simplification by a Lambertian reflectance model used for? Give two examples at least, please.

31. 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.)

32. 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.

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

34. Write the formula for smoothing the histgram 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.

35. Which methods could you use if you like to increase the contrast of the gray scale image to improve perception of an observing human. Only this single image is available to you. List two qualitatively different methods. Explain their idea concisely.

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

37. 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.

38. 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?

39. Heisenberg principle gives the relation between the frequency spectrum with in Fourier domain and the time duration of a corresponding signal. Formulate the principle informally and explain its significance for the frequency analysis in image processing.

40. 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 commonly used to overcome this limitation.

41. State the convolution theorem, i.e., how is convolution expressed in the Fourier domain. Consider the 1D case for simplicity.

42. What is the computational complexity of the discrete Fourier transforamtion for a $N \times N$ 2D image if you used the definition formula. I add that the asymptotic estimate of the algorithmic complexity uses $\mathcal{O}(.)$ form. The argument will be an expression containing the value N. The additive and multiplicative constants are not considered.

43. The discrete Fourier transformation was sped up by the Fast Fourier transformation algorithm (FFT). What is the principle of this algorithm?

44. What is the algorithmic complexity of FFT algorithm for 2D image of the size $N \times N$? Use $\mathcal{O}(.)$ notation.

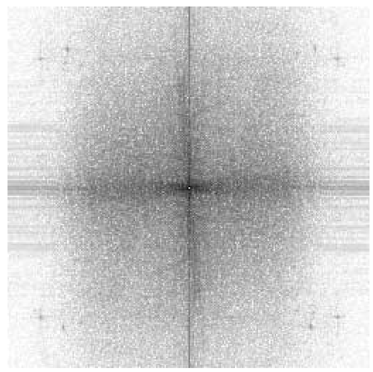
45. Formulate Shannon (also Nyquist, Kotelnikov) sampling theorem for the simpler 1D signal case. Explain (informal explanation suffices, the drawing helps you). What is the idea of sampling theorem proof? (Hint: frequency spectra).

46. Why do analog television norms such as NTSC, PAL (which are several decades old) use interlaced image lines?

47. (a) A television signal, 50 half-frames per second is sampled into the matrix 500×500 pixels in 256 gray levels. Calculate the minimal sampling frequency (in kHz) for its digitization.

    (b) How is the used theorem named? Sketch the idea of the sampling theorem. Described idea is enough. The formulae are not necessary.

48. The image on the left side provides the input intensity image. The image on the right side shows the Fourier power spectrum displayed as the intensity image. The dark pixels denote higher values. Two distinct crosses are visible in the spectrum. The first cross that is more pronounced matches vertical and horizontal directions. The less pronounced cross is slightly rotated in the counter clockwise direction. Explain to which phenomena the crosses correspond in the original intensity image.

49. Linear orthogonal integral transformations are useful for representing and processing signals and images (as e.g. Fourier, cosine transformation, principal component method). Explain what is the priciple of these methods. Give two distinct exaples of their use.

50. Assume that we like to restore a distorted image, e.g. by motion blur. Let the degradation be modeled as a convolution with the known kernel $H(u, v)$, where u, v are spatial frequencies in Fourier transformation. In such a case an inverse filter can be used for the restoration. Formulate the method, write the formula and explain its principle informally. What are the limitation of the method for practical use?

51. Assume that we like to restore a distorted image, e.g. by motion blur. Let the degradation be modeled as a convolution with the known kernel $H(u, v)$, where u, v are spatial frequencies in Fourier transformation. Assume that statistical properties of the involved additive noise can be estimated. In such a case Wiener filter can be used for the restoration. Formulate the method and explain informally its principle. The formula may simplify the answer but it is not necessary. What is the advantage of Wiener filtering as compared to its more specialized version named inverse filter?

52.
    - Explain the principle of Wiener filter (text desription suffice). Which noise parameters have to be estimated before Wiener filter can be used?
    - Is Wiener filter linear?
    - Let $N(u, v)$ be Fourier image of the additive noise. The inverse filter, which is a special case of Wiener filter, restores the undegraded image $G(u, v)$ according the formula

    $$F(u, v) = G(u, v)H^{-1}(u, v) - N(u, v)H^{-1}(u, v) .$$

    The expression subtracted on the right side of the equation brings troubles, which forces to use a more general Wiener filter in many practical cases. Explain what troubles are meant here and what causes them.

53. Consider image restauration methods for specific degradation (as motion blur, lens defocus, atmospheric turbulence) and additive noise. The degradation can be modeled linearly. The degradation is expressed by a convolution kernel $h(x, y)$ spanning the whole image, where x, y are image coordinates. When expressing in frequency domain after Fourier transformation is applied, the degradation is given by the filter $H(u, v)$, where u, v are spatial frequences. Let $G(u, v)$ be a Fourier image of the observable degraded image. Let $F(u, v)$ be the estimate of the degradation free image, the desired outcome of restauration. Let $\nu(x, y)$ be the additive nosie, which is independent statistically on the image. Its couterpart in the frequency domain is $N(u, v)$.

    (a) Write the formula for calculating $F(u, v)$ using the inverse filtration method in the frequency domain.(b) Troubles with inverse filtration occur under certain condition. What are the conditions when inverse filtration malfunctions and a more complicated Wiener filtration has to be used?

54. 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.

55. Explain the principle of brightness corrections for a multiplicative model of the degradation. Express mathematically. Give example of the use. Methods are often used to correct a systematic degradation in an image capturing process.

56. 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).

57. Explain the rationale of histogram equalization.

58. Explain why the histogram of a discrete image is not flat after histogram equalization, what we would aim at in the ideal case.

59. A general intensity transformation T essentially replaces each intensity p by a new one $q = T(p)$. Assume a standard 8-bit grayscale image. Will the number of output intensities always the same as the number of input intensities? Explain and give examples.

60. Consider a grayscale image. Histogram equalization is used to better utilize the grayscale and enhance contrast of the image. Does histogram equalization increase the amount of information contained in image data when assessed by entropy? Explain.

61. 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?

62. 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 ...).

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

64. 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 originally σ? Explain what is the statistical priciple of noise dumping (A hint: the central limit theorem).

65. Is it possible to filter additive noise using ordinary filtering, e.g. from 21 samples, while obtaining the image which is not blurred? If yes, how?

66.

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.

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

68. If recursive filters (IIR – infinite impulse response filters) creates difficulties for 2D images, which do not appear for 1D signals. That is the reason why recursive filters are used rarely for images. What causes these problems? (Hint: causality).

69. What is the reason why separable filters are used while filtering images linearly? Explain the basic idea of separable filters. What conditions the filter has to fulfil in order to be implemented as a separable filter?

70. Consider that you like to increase the contrast of the image. What methods can you use if only the single input image is available. Name two such methods.

71. What is the edge in the image $f(x, y)$? How the edge expressed mathematically? Please, give the defining fomulas both for the continuous and the discrete images..

72. What is the edge element (abbreviated edgel)? What is the edgel used for in image analysis?

73. Edges are sometimes sought in the image function $f(x, y)$ by Laplace operator $\nabla^2 f(x, y)$. Write the formula defining Laplace operator for a continuous image function $f(x, y)$. Are properties of Laplacean operator direction dependent?

74. What advantage brings the second derivative of the image function when finding the edge location? Write in which edge detector is the second derivative as used an advantage and how.

75. Marr's approach to edge detection explores the zero crossing of the image function second derivative. A convolution (smoothing) with a Gaussian filter g is often used while calculating image derivatives. Let the second derivative of this smoothing operation is denoted $\nabla^2 d = \nabla^2 (f * g) = \nabla^2 f * g = \ldots$. The method uses a smart trick, which bypasses the derivation of the image function f.

    (a) Please, use this trick and continue the above derivation.

    (b) This trick uses special properties of involved operations. Which properties are explored here?

76. Consider that you have and captured grayscale digital image. Explain the image sharpening operation. I do not mean the global modification of the gray scale, e.g. by modifying the image histogram. What is the aim of image sharpening? When is such sharpening used?

77. Explain the 'palette color image' concept. Explain why and for what are pallete color images used?

78. What is color? What co-play of three phenomena does a human perceive the color?

79. Why do we perceive (see) some object as color ones? Consider, for instance, a freshly cut red rose. Why do we see the green stem, leaves and the red flower (blossom) of the rose?

80. When we describe color from a physics point of view, we imagine the visible part of the frequency spectrum obtained e.g. by decomposing the incoming light by an optical prism (I. Newton's experiment). Write the range of wavelengths (from, to) in namometers [nm], which a human eye perceives. Mention four visible colors in an ascending order according to their wavelength. (Hint: Recall the rainbow colors).

81. What sensors are used for seeing color in human vision? Draw the approximate sensitivity of the sensors as a graph. The horizontal axis gives the wavelength in nanometers[nm]. The vertical axis gives the relative sensitivity of the sensor in the range between zero and one.

82. Explain the concept – the color matamerism. What it means for human color perception?

83. What is the color space? How is it defined? Consider for simplicity the color space corresponding to color sensors in a human eye .

84. How and why the color space CIE XYZ was established? Explain the color triangle and sketch it. What do coordinates x, y? of a color mean? What are spectral colors and where are they in the color triangle?

85. What does the color gamut of a particular capturing or display mean? How does the color gamut correspond to a color triangle? Compare the colog gamut of a good quality film to a cheap computer printer.

86. Data compression (including image compression) relies on lowering the data redundance or as the case can be on lowering the data irrelevance. Explain both concepts. Give one example of lowering redundance and one example of lowering irrelevance in image compression. Explain briefly the compression methods you considered.

87. Explain what is lossy and lossless image compression while using concepts data redundance and data irrelevance.

88. Shannon (also information) entropy is used for estimating redundance while compressing image data. Consider a gray scale image with the histogram $h(i)$, $i = 1 \dots 255$. Calculate the entropy estimate. How do you calculate redundance if each pixel value is represented by n bits?

89. Huffman coding is used in data compression for reducing redundance in coding. Explain its principle. Is Huffman coding optimal? If so, under which conditions?

90. Huffman coding is used in data compression for reducing redundance in coding. Under which condition is Huffman coding optimal? When these conditions are not met the method can be improved by an improved method. What is its name? How does this method differ from Huffman coding?

91. (a) Explain the principle of the very often used lossy image compression method JPEG?

    (b) When using high compression ratios, 8×8 squares are visible in the decompressed image. What causes this blocking effect? Why such a solution was selected?

92. What is the difference between lossy and lossless image compression methods? Explain the principle of boty lossy and lossless methods. Give one example for each method.

93. Compression methods are also used for one-dimensional (1D) signals. An image can be represented as 1D signal too, e.g. if it is traversed line by line. This happens when we 'ZIP' the image, i.e. LZW algorithm working with a dictionary is used. If a compression method dedicated for images is used, we can achieve higher compression. Why? Use the term redundancy in your explanation. (The answer to this question answers also a natural question: Why the 2D image compression differs from 1D signals compression?).

94. Write two definitions of the compression ratio using redundance and memory saving.

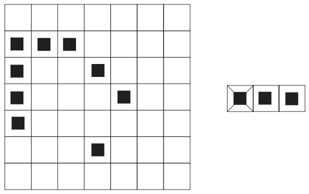
95. Explain the image compression methods principle which use linear integral transformations. Name two such methods and sketch their principle.

    Why different compression methods are used for 2D images and 1D signals (sequences)?

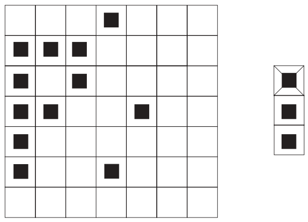
96. JPEG compression explores cosine transformation. Let the image havs n rows and n columns. What is the time complexity of the cosine transformation from definition and what for the fast variant (its principle is similar to FFT)? (Use the notation $\mathcal{O}(.)$ for expressing the complexity).

97. Is the erosion of a binary image a commutative operation? Explain why? (Hint: start wit one definition of erosion)

98. Write the formula defining the erosion of binary images. What is the binary image erosion good for?

99. 

A binary mathematical morphology example. The left image gives the point set A; the right image gives the structuring element B. Its representative pixel is denoted by a cross. Draw the result of the dilation $A \oplus B$.

100. 

A binary mathematical morphology example. The left image gives the point set A; the right image gives the structuring element B. Its representative pixel is denoted by a cross. Draw the result of the erosion $A \ominus B$.

101. Explain what it means if the operation is named idempotent. For what is the idempotence property used in mathematical morphology?

102. What does it mean if the operation is called idempotent? Is closing (a mathematical morphology operation) idempotent?

103. Consider a continuous image function. Draw skeletons corresponding to the inner part of an equilateral triangle. We dealt with skeletons when we learned mathematical morphology.

104. Consider a continuous image function. Draw skeletons corresponding to the pair of filled circles of different diameters touching one the other. We dealt with skeletons when we learned mathematical morphology.

105. 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.

106. 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 partuclar method.

107. 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.

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

109. 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.

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

111. While segmenting based on intensity threshold, 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?

112. While segmenting by thresholding (with more than one threshold in a more general case too), it is possible to approximate the histogram (probability of intensities) by several Gaussians. The segmentation task is converted to separation of a probability distribution mixture. This approach is common in statistical pattern recognition. Write how constituting probability distributions are obtained and which method is used for separating them. You can use the example from the lecture as an example.

113. 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?

114. While seeking edge elements (significant edges, edgels), two procedures are used: (a) Thresholding with hysteresis; (b) suppressing local gradient magnitude maxima (called non-maximal suppression). Explain both these approaches.

115. The edge detector may yield edge elements (edgels), which are not contiguous and which are noisy. Connecting edgels into closed boundaries can be uneasy. Sometimes a postprocessing step named the edge relaxation is used. Consider a its simple, pragmatic instance of a general task named relaxation labeling. Explain it briefly.

116. Partial border segments can be, e.g. created by fusing edge elements (edgels) using the Gestalt grouping principle. Explain it briefly.

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

118. Objects in images can be described using either global or local methods. Compare them and give two examples for each of them.

119. Explain two basic approaches to region coloring. It is a method taking as the input the outcome of binary image segmentation. Regions in the output are marked by a unique label, e.g. by a number or by color. Labels/colors for distinct regions differ.

120. To which number does the region area converges (in m^2) and the region perimeter (in m) in a discrete image when the image resolution increases? Hing: Imagine a sequence of Earth images taken by satellites of a growing resolution and a specific object with a curvilinear boundary, e.g. a rocky sea coast.

121. The popular feature representing shape of the region (e.g. the object in @D image after segmentation) is the compactness (sometimes called non-compactness). Define the compactness. How does the compactness changes when the image resolution grows? Why?

122. Explain the idea of the chain code used for object boundary description. Draw a region of seven pixels at least which is not a rectangle. Consider 8-neighborhood. Write the corresponding chain code and the derivation of the chain code.

123. Consider a region in 2D image, outcome of segmentation. Explain the principle of moment descriptors for a region. How is invariance to various geometric transformation achieved?