Study guide for Graduate Computer Vision

Erik G. Learned-Miller
Department of Computer Science
University of Massachusetts, Amherst
Amherst, MA 01003

November 21, 2013

Abstract

You should know the answers to all of these questions. Almost all of them were answered in class at some time, but if you do not know the answers, you can look them up on line. Also, it is fine if you work with friends to figure out answers to questions you don't know. The goal is for you to learn this material before the test, however you can.

1. Know Bayes' rule. What are likelihoods, priors, and posteriors? Be able to use these terms properly. Be able to compute marginal distributions and conditional marginal distributions.

2. Estimate probability distributions from data.

3. Discuss a discrete random variable that has no "topology". This means that the presence of one value of the random variable tells you noting about the presence of another value of the random variable. Example: drawing a red ball from an urn tells me nothing about how many orange balls are in the urn, even though red and orange are "close". (I want you to come up with your own example of this.)

4. Know how to estimate a continuous distribution with a Parzen density estimator (also known as kernel density estimates, Parzen window estimates, etc.).

5. Know the concept of "leave-one-out" estimation. Why is this important for setting the kernel width (or the standard deviation for a Gaussian kernel) in Parzen density estimation? What happens if you include a sample in its own estimate and then maximize the likelihood of the sample with respect to the kernel size?

6. How can you estimate a discrete distribution (with topology) in which the distribution can take on a very large number of values (say, brightness values) when you have only a small amount of data? (Use fewer bins, use a "soft binning" strategy, or spread a sample among multiple bins.)

7. If you are told that certain variables are independent, this should enable you to get more accurate estimates of the joint distribution of those variables from data.

8. Know the definition for the entropy of a discrete random variable in terms of its probability distribution.

9. Know the definition of the mutual information of two random variables. Note that in order to compute the mutual information of two random variables, you must have access to the joint distribution of these variables, not just the marginal distributions.

10. What does the mutual information of two random variables tell me about the statistical dependence of the random variables. What is the mutual information of two random variables that have 0 bits of mutual information?

11. Why is mutual information a good criterion for the Prokudin-Gorsky alignment problem, while correlation might not be good?

12. Know the definition of correlation. Are uncorrelated random variables statistically independent? Are statistically independent random variables uncorrelated?

13. Know the basic setup of supervised learning. If you use training data to modify the parameters of a classifier, such as using training data to estimate the probability distributions of each class, do you expect a classifier built on these parameters to have a higher error on the training set or test set? Why?

14. Be able to rapidly calculate the number of distinct NxN images in which each pixel can take on K values.

15. What's the problem with the naive approach to building an object classification problem by estimating the distribution of images in each class? (Answer: too many bins in the probability distribution to estimate from any practical size sample.)

16. Know the rough range of wavelengths of visible light. (Answer: 400-700 nanometers).

17. Which wavelengths are just longer than visible? Which are just shorter?

18. If I give you the response of a light detector of each wavelength of light (as a continuous function), and I give you the relative amount of energy at each wavelength of light in a light source S, how do you determine the response of the detector to the light source S. (Answer: Integrate the product of response function and the light power distribution function.)

19. Understand the definition of a linear function. Be able to show algebraically why $f(x) = 3x + 7$ is not a linear function.

20. Describe how look up tables works for displaying a scalar-valued image with a particular color map.

21. Image filtering is the process of taking an array (a filter) and applying it to an image by overlaying it on the image and replacing the center pixel with the dot product of the filter and the underlying image. Convolution is exactly like filtering, except that before doing the filtering, we flip the filter vertically and horizontally. Convolution models the process of image formation in which each pixel of some "original" image gets perturbed by the point spread function. A good example of this is what happens to an image of stars in the night sky when it comes through the atmosphere.

22. Give two different filters (show the array of numbers) that blur an image, and understand why they work. (Answers: box filter and Gaussian filter).

23. If the values of a matrix add to 1, what can you say about an image that is filtered with that matrix (answer: the sum of the image will remain fixed). Understand why this is true.

24. Give a filter that will enhance vertical edges in an image. A horizontal edge.

25. Discuss the filters used to make a Gaussian pyramid (as in the SIFT paper). Discuss the filters used to make a Difference of Gaussian (DOG) pyramid.

26. How do I estimate the derivative of an image with respect to the x direction ($\frac{\partial I}{\partial x}$). The y-direction?

27. Express the estimated local image gradient as a function of these estimated partial derivatives.

28. Understand the basic sequence of steps in producing a SIFT descriptor representation for an image:

    (a) Build a DoG image pyramid.

    (b) Find the local extrema of the DoG representation.

    (c) Eliminate points with low contrast or low values of the minimum local curvature.

    (d) Put a local coordinate system down at a keypoint based upon the locally most common gradient direction and the scale at which the keypoint was found in the DoG representation.

    (e) Build a set of 16 (4x4) local histograms of gradient orientations.

29. Know the definition of the entropy of discrete probability distribution ($H(X) = -\sum_x p(x) \log p(x)$). Understand the essential difference between entropy and variance. Think about these questions about a four-sided die.

    (a) If the die is fair, and numbered 1-4. What is the variance? What is the entropy?

    (b) If the numbers on the die are multiplied by 10, what happens to the variance of the die? What happens to the entropy?

30. Describe congealing, the joint alignment of images. What criterion does congealing use for alignment? Does it minimize it or maximize it? What advantages does congealing have over the alignment of 2 images at a time? What disadvantages?

31. What the key elements of an effective feature for classification? Answer: discriminativeness and repeatability. Discuss examples of features for face recognition that might be

    (a) discriminative but not repeatable

    (b) repeatable but not discriminative

    (c) both

(d) neither

32. The light-detecting cells in the retina are rods and cones. Cones come in 3 varieties, “red”, “green” and “blue”. They all have different, but broad spectrum responses.

33. Understand the two basic methods for capturing color images with a modern CCD camera (use a Bayer pattern with a single CCD array or use a beam-splitter and 3 CCDs.). How can I produce a normal RGB image from the output of a Bayer pattern?

34. Why can a television reproduce most of the colors we’re familiar with using only 3 colors of phospors for each position on the screen?

35. How good is the human visual system at judging absolute brightnesses? Give an example showing it’s not good at this. Is this a bug or a feature of the human visual system? What *are we* good at judging?

36. What is a distribution field? How can you build one from 100 images? How can you build one from a single image? If you build a distribution field from a single image, how can you “spread” the information about a particular color in the image dimensions? How can you spread it in the feature dimension?

37. A distribution field, by definition, has a probability distribution at each position in the image. If I convolve my distribution field with a discretized version of a 2-dimension Gaussian kernel, why do I still have a distribution field (ignoring problems that occur at the edge of the image)? Answer: Because when I convolve a distribution field with a Gaussian kernel, each new pixel column can be written as the weighted sum of previous probability distributions. Furthermore, the weights in this weighted sum add up to one (since they are defined by a Gaussian distribution). That means this weighted sum is a *convex combination*, i.e. a weighted sum in which the weights are positive and add to 1. Any convex combination of probability distributions is still a probabilty distribution since its total mass must still be one, and there is no way any of the values could have become 0. That’s the definition of a probability distribution!

38. Convolution is both commutative and associative. This means that instead of convolving and image with a Gaussian and then convolving it with a derivative filter, and can first combine the filters and apply them both to the image at the same time, achieving a more efficient result.

39. Understand these families of image transformations: translations, rigid (Euclidean), similarity, affine, linear. What does each of them preserve?

40. How can we formally define alignment in terms of a matching criterion and a set of transformations? (Answer: find the transformation of one image which maximizes the similarity between that image and another image.)

41. Why is exhaustive search an unsatisfying solution to image alignment?

42. Explain how image alignment can be used to build a basic tracking algorithm.

43. Discuss a serious problem with pixelwise alignment, i.e., a function which compares two images by only comparing corresponding pixel pairs. (Answer: small shifts of an image can result in large differences between the images.)

44. An image can be represented as a histogram over brightness values or colors. What is wrong with this representation of an image?

45. Discuss compromises between pixelwise image representations and representing an image with a single histogram. Answer: arrays of histograms.

46. What 3 elements define a specific congealing algorithm? Answer: a set of arrays (for exmaple, images, volumes, time-series), a set of transformations (2-d affine, 3-d rigid, or brightness transformations, a criterion of joint alignment: sum of pixelstack entropies.

47. Know how to derive the maximum likelihood estimator for the mean of a distribution given a set of samples $x_1, x_2, ..., x_n$.

48. Discuss the idea of making a "funnel" for aligning a set of images by using the intermediate results from congealing to produce a series of distribution fields to which new images can be aligned. Why is it called a funnel?

49. Describe the two properties that I discussed in class that are good for an image similarity function to have. ("norm-like" and weak invariance to position). Know what these are, and examples that satisfy these conditions and that do not satisfy these conditions.