Study guide for Graduate Computer Vision

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

November 23, 2011

Abstract

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. Discuss a continuous distribution with topology. Discuss a discrete distribution with topology.

5. 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.)

6. 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.

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

8. 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.

9. 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?

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

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

12. 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?

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

14. 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.)

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

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

17. 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.)

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

19. Understand the difference between a point source of light (emits a finite amount of light from an infinitessimally small location) and extended source (emits a differentially small amount of light from each position, but, integrated over an area, emits a finite amount of light).

20. Be able to do the "solar panel" calculation showing what percentage of the power of the sun a solar panel absorbs given the relevant parameters (angle of the solar panel to the direction of the sun, size of the solar panel, distance of the solar panel from the sun).

21. Understand pinhole cameras and their pros and cons.

22. Understand the meaning of the (BRDF) bidirectional reflectance distribution function (answer: it is a 4 parameter function (2 parameters for angle of incoming ray, 2 parameters for outgoing ray) that gives the percentage of light in each outgoing direction for each incoming direction).

23. What are two simplest and most commonly discussed BRDFs? Answer: matte surface (also called Lambertian) and perfectly reflective (mirror) surfaces. Lambertian surfaces appear to have the same brightness from every angle. Mirror surfaces reflect all incoming light in a single direction.

24. How does a corner reflector work? Describe its BRDF in relation to a mirror.

25. Describe how one can build a classifier for the material from which a surface is made. Come up with some reasonable features that might help you classify materials.

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

27. Describe two methods for making a low contrast image look better (histogram equalization and color remapping). The first of these moves the data points until they have a uniform distribution, keeping their values in the same relative order. The second one changes the color map (or look up table) so that the colors (or brightness values) of the pixels are spread out further. Discuss pros and cons of these two approaches.

28. A filter is an array which is applied to an image via convolution. That is, each pixel in an image is replaced by a dot product of the filter with the portion of the underlying image centered at that pixel.

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

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

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

32. 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.

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

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

35. 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 (not extermums!) 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.

36. 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?

37. 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?

38. 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

39. 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. They are primarily useful in bright light scenarios (photopic vision). In dim light, we use rods to detect brightness fluctuations, but not color (scotopic vision). Why does watching "black and white" TV seem fairly natural to us?

40. What is the fovea? What does it mean to foveate? What is a saccade? Where is the blind spot and what causes it (it's where the optic nerve attaches to the retina. There are no rods or cones there.) Discuss the different between the retina detecting no light versus the blind spot reporting no data. In lecture, I called this "the difference between 0 and nothing".)

41. 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?

42. 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?

43. Assuming I want to take a picture in 100 milliseconds using a modern camera, when I am designing the CCD array, I have a trade-off between the resolution of the image (number of pixels), and the number of brightness levels that I produce at each pixel.

44. What is the basic idea behind stereo vision? Over what range is stereo vision most useful? Why is stereo vision not useful at large distances? Why can a person with one eye do most of the things we can do with 2 eyes?

45. 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?

46. 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?

47. 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!

48. 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.

49. Matrix multiplication is associative. Thus, instead of applying a rotation, and then a scaling, and then a translation, and then a shear, I can combine matrices first, and apply them all to the image.

50. The preceding item about combining transformations into a single transformation before applying them to images as **TWO** major advantages. One is that it is more efficient. The second is that since you only warp the image once, you don't suffer as much degradation of the image from successive resampling and interpolation, which should be minimized as much as possible.