IMAGE INTENSITIES AS RANDOM VARIABLES

You may find it useful to consult the tutorials section in the book website for a brief review of probability.

We treat image intensities as random quantities in numerous places in the book. For example, let z_i, $i = 0,1,2,\ldots,L-1$, denote the values of all possible intensities in an $M \times N$ digital image. The probability, $p(z_k)$, of intensity level z_k occurring in the image is estimated as

$$p(z_k) = \frac{n_k}{MN} \tag{2-67}$$

where n_k is the number of times that intensity z_k occurs in the image and MN is the total number of pixels. Clearly,

$$\sum_{k=0}^{L-1} p(z_k) = 1 \tag{2-68}$$

Once we have $p(z_k)$, we can determine a number of important image characteristics. For example, the mean (average) intensity is given by

$$m = \sum_{k=0}^{L-1} z_k p(z_k) \tag{2-69}$$

Similarly, the variance of the intensities is

$$\sigma^2 = \sum_{k=0}^{L-1} (z_k - m)^2 p(z_k) \tag{2-70}$$

The variance is a measure of the spread of the values of z about the mean, so it is a useful measure of image contrast. In general, the nth central moment of random variable z about the mean is defined as

$$\mu_n(z) = \sum_{k=0}^{L-1} (z_k - m)^n p(z_k) \tag{2-71}$$

We see that $\mu_0(z) = 1$, $\mu_1(z) = 0$, and $\mu_2(z) = \sigma^2$. Whereas the mean and variance have an immediately obvious relationship to visual properties of an image, higher-order moments are more subtle. For example, a positive third moment indicates that the intensities are biased to values higher than the mean, a negative third moment would indicate the opposite condition, and a zero third moment would tell us that the intensities are distributed approximately equally on both sides of the mean. These features are useful for computational purposes, but they do not tell us much about the appearance of an image in general.

As you will see in subsequent chapters, concepts from probability play a central role in a broad range of image processing applications. For example, Eq. (2-67) is utilized in Chapter 3 as the basis for image enhancement techniques based on histograms. In Chapter 5, we use probability to develop image restoration algorithms, in Chapter 10 we use probability for image segmentation, in Chapter 11 we use it to describe texture, and in Chapter 12 we use probability as the basis for deriving optimum pattern recognition algorithms.

Summary, References, and Further Reading

The material in this chapter is the foundation for the remainder of the book. For additional reading on visual perception, see Snowden et al. [2012], and the classic book by Cornsweet [1970]. Born and Wolf [1999] discuss light in terms of electromagnetic theory. A basic source for further reading on image sensing is Trussell and Vrhel [2008]. The image formation model discussed in Section 2.3 is from Oppenheim et al. [1968]. The IES Lighting Handbook [2011] is a reference for the illumination and reflectance values used in that section. The concepts of image sampling introduced in Section 2.4 will be covered in detail in Chapter 4. The discussion on experiments dealing with the relationship between image quality and sampling is based on results from Huang [1965]. For further reading on the topics discussed in Section 2.5, see Rosenfeld and Kak [1982], and Klette and Rosenfeld [2004].

See Castleman [1996] for additional reading on linear systems in the context of image processing. The method of noise reduction by image averaging was first proposed by Kohler and Howell [1963]. See Ross [2014] regarding the expected value of the mean and variance of the sum of random variables. See Schröder [2010] for additional reading on logic and sets. For additional reading on geometric spatial transformations see Wolberg [1990] and Hughes and Andries [2013]. For further reading on image registration see Goshtasby [2012]. Bronson and Costa [2009] is a good reference for additional reading on vectors and matrices. See Chapter 4 for a detailed treatment of the Fourier transform, and Chapters 7, 8, and 11 for details on other image transforms. For details on the software aspects of many of the examples in this chapter, see Gonzalez, Woods, and Eddins [2009].

Problems

Solutions to the problems marked with an asterisk (*) are in the DIP4E Student Support Package (consult the book website: www.ImageProcessingPlace.com).

2.1 If you use a sheet of white paper to shield your eyes when looking directly at the sun, the side of the sheet facing you appears black. Which of the visual processes discussed in Section 2.1 is responsible for this?

2.2* Using the background information provided in Section 2.1, and thinking purely in geometrical terms, estimate the diameter of the smallest printed dot that the eye can discern if the page on which the dot is printed is 0.2 m away from the eyes. Assume for simplicity that the visual system ceases to detect the dot when the image of the dot on the fovea becomes smaller than the diameter of one receptor (cone) in that area of the retina. Assume further that the fovea can be modeled as a square array of dimension 1.5 mm on the side, and that the cones and spaces between the cones are distributed uniformly throughout this array.

2.3 Although it is not shown in Fig. 2.10, alternating current is part of the electromagnetic spectrum. Commercial alternating current in the United States has a frequency of 60 Hz. What is the wavelength in kilometers of this component of the spectrum?

2.4 You are hired to design the front end of an imaging system for studying the shapes of cells, bacteria, viruses, and proteins. The front end consists in this case of the illumination source(s) and corresponding imaging camera(s). The diameters of circles required to fully enclose individual specimens in each of these categories are 50, 1, 0.1, and 0.01 μm, respectively. In order to perform automated analysis, the smallest detail discernible on a specimen must be 0.001 μm.

(a)* Can you solve the imaging aspects of this problem with a single sensor and camera? If your answer is yes, specify the illumination wavelength band and the type of camera needed. By "type," we mean the band of the electromagnetic spectrum to which the camera is most sensitive (e.g., infrared).

(b) If your answer in (a) is no, what type of illumination sources and corresponding imaging sensors would you recommend? Specify the light sources and cameras as requested in part (a). Use the minimum number of illumination sources and cameras needed to solve the problem. (*Hint:* From the discussion in

Section 2.2, the illumination required to "see" an object must have a wavelength the same size or smaller than the object.)

2.5 You are preparing a report and have to insert in it an image of size 2048×2048 pixels.

(a)* Assuming no limitations on the printer, what would the resolution in line pairs per mm have to be for the image to fit in a space of size 5×5 cm?

(b) What would the resolution have to be in dpi for the image to fit in 2×2 inches?

2.6* A CCD camera chip of dimensions 7×7 mm and 1024×1024 sensing elements, is focused on a square, flat area, located 0.5 m away. The camera is equipped with a 35-mm lens. How many line pairs per mm will this camera be able to resolve? (*Hint:* Model the imaging process as in Fig. 2.3, with the focal length of the camera lens substituting for the focal length of the eye.)

2.7 An automobile manufacturer is automating the placement of certain components on the bumpers of a limited-edition line of sports cars. The components are color-coordinated, so the assembly robots need to know the color of each car in order to select the appropriate bumper component. Models come in only four colors: blue, green, red, and white. You are hired to propose a solution based on imaging. How would you solve the problem of determining the color of each car, keeping in mind that cost is the most important consideration in your choice of components?

2.8* Suppose that a given automated imaging application requires a minimum resolution of 5 line pairs per mm to be able to detect features of interest in objects viewed by the camera. The distance between the focal center of the camera lens and the area to be imaged is 1 m. The area being imaged is 0.5×0.5 m. You have available a 200 mm lens, and your job is to pick an appropriate CCD imaging chip. What is the minimum number of sensing elements and square size, $d \times d$, of the CCD chip that will meet the requirements of this application? (*Hint:* Model the imaging process as in Fig. 2.3, and assume for simplicity that the imaged area is square.)

2.9 A common measure of transmission for digital data is the *baud rate*, defined as symbols (bits in our case) per second. As a minimum, transmission is accomplished in packets consisting of a start bit, a byte (8 bits) of information, and a stop bit. Using these facts, answer the following:

(a)* How many seconds would it take to transmit a sequence of 500 images of size 1024×1024 pixels with 256 intensity levels using a 3 M-baud (10^6 bits/sec) baud modem? (This is a representative medium speed for a DSL (Digital Subscriber Line) residential line.

(b) What would the time be using a 30 G-baud (10^9 bits/sec) modem? (This is a representative medium speed for a commercial line.)

2.10* High-definition television (HDTV) generates images with 1125 horizontal TV lines interlaced (i.e., where every other line is "painted" on the screen in each of two fields, each field being 1/60th of a second in duration). The width-to-height aspect ratio of the images is 16:9. The fact that the number of horizontal lines is fixed determines the vertical resolution of the images. A company has designed a system that extracts digital images from HDTV video. The resolution of each horizontal line in their system is proportional to vertical resolution of HDTV, with the proportion being the width-to-height ratio of the images. Each pixel in the color image has 24 bits of intensity, 8 bits each for a red, a green, and a blue component image. These three "primary" images form a color image. How many bits would it take to store the images extracted from a two-hour HDTV movie?

2.11 When discussing linear indexing in Section 2.4, we arrived at the linear index in Eq. (2-14) by inspection. The same argument used there can be extended to a 3-D array with coordinates x, y, and z, and corresponding dimensions M, N, and P. The linear index for any (x, y, z) is

$$s = x + M(y + Nz)$$

Start with this expression and

(a)* Derive Eq. (2-15).

(b) Derive Eq. (2-16).

2.12* Suppose that a flat area with center at (x_0, y_0) is illuminated by a light source with intensity distribution

$$i(x,y) = Ke^{-[(x-x_0)^2 + (y-y_0)^2]}$$

Assume for simplicity that the reflectance of the area is constant and equal to 1.0, and let $K = 255$. If the intensity of the resulting image is quantized using k bits, and the eye can detect an abrupt change of eight intensity levels between adjacent pixels, what is the highest value of k that will cause visible false contouring?

2.13 Sketch the image in Problem 2.12 for $k = 2$.

2.14 Consider the two image subsets, S_1 and S_2 in the following figure. With reference to Section 2.5, and assuming that $V = \{1\}$, determine whether these two subsets are:

(a)* 4-adjacent.

(b) 8-adjacent.

(c) m-adjacent.

$$\begin{array}{c|cccc|cccc|c}
\multicolumn{1}{c}{} & \multicolumn{4}{c}{S_1} & \multicolumn{4}{c}{S_2} & \multicolumn{1}{c}{} \\
\cline{2-9}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\
1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
\cline{2-9}
\multicolumn{1}{c}{0} & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & \multicolumn{1}{c}{1}
\end{array}$$

2.15* Develop an algorithm for converting a one-pixel-thick 8-path to a 4-path.

2.16 Develop an algorithm for converting a one-pixel-thick m-path to a 4-path.

2.17 Refer to the discussion toward the end of Section 2.5, where we defined the background of an image as $(R_u)^c$, the complement of the union of all the regions in the image. In some applications, it is advantageous to define the background as the subset of pixels of $(R_u)^c$ that are not *hole* pixels (informally, think of holes as sets of background pixels surrounded by foreground pixels). How would you modify the definition to exclude hole pixels from $(R_u)^c$? An answer such as "the background is the subset of pixels of $(R_u)^c$ that are not hole pixels" is not acceptable. (*Hint:* Use the concept of connectivity.)

2.18 Consider the image segment shown in the figure that follows.

(a)* As in Section 2.5, let $V = \{0,1\}$ be the set of intensity values used to define adjacency. Compute the lengths of the shortest 4-, 8-, and m-path between p and q in the following image. If a particular path does not exist between these two points, explain why.

$$\begin{array}{rcccl}
 & 3 & 1 & 2 & 1\,(q) \\
 & 2 & 2 & 0 & 2 \\
 & 1 & 2 & 1 & 1 \\
(p) & 1 & 0 & 1 & 2
\end{array}$$

(b) Repeat (a) but using $V = \{1,2\}$.

2.19 Consider two points p and q.

(a)* State the condition(s) under which the D_4 distance between p and q is equal to the shortest 4-path between these points.

(b) Is this path unique?

2.20 Repeat problem 2.19 for the D_8 distance.

2.21 Consider two *one-dimensional* images f and g of the same size. What has to be true about the orientation of these images for the elementwise and matrix products discussed in Section 2.6 to make sense? Either of the two images can be first in forming the product.

2.22* In the next chapter, we will deal with operators whose function is to compute the sum of pixel values in a small subimage area, S_{xy}, as in Eq. (2-43). Show that these are linear operators.

2.23 Refer to Eq. (2-24) in answering the following:

(a)* Show that image summation is a linear operation.

(b) Show that image subtraction is a linear operation.

(c)* Show that image multiplication in a nonlinear operation.

(d) Show that image division is a nonlinear operation.

2.24 The median, ζ, of a set of numbers is such that half the values in the set are below ζ and the other half are above it. For example, the median of the set of values $\{2,3,8,20,21,25,31\}$ is 20. Show that an operator that computes the median of a subimage area, S, is nonlinear. (*Hint:* It is sufficient to show that ζ fails the linearity test for a simple numerical example.)

2.25* Show that image averaging can be done recursively. That is, show that if $a(k)$is the average of k images, then the average of $k+1$ images can be obtained from the already-computed average, $a(k)$, and the new image, f_{k+1}.

2.26 With reference to Example 2.5:

(a)* Prove the validity of Eq. (2-27).

(b) Prove the validity of Eq. (2-28).

For part (b) you will need the following facts from probability: (1) the variance of a constant times a random variable is equal to the constant squared times the variance of the random variable. (2) The variance of the sum of uncorrelated random variables is equal to the sum of the variances of the individual random variables.

2.27 Consider two 8-bit images whose intensity levels span the full range from 0 to 255.

(a)* Discuss the limiting effect of repeatedly subtracting image (2) from image (1). Assume that the results have to be represented also in eight bits.

(b) Would reversing the order of the images yield a different result?

2.28* Image subtraction is used often in industrial applications for detecting missing components in product assembly. The approach is to store a "golden" image that corresponds to a correct assembly; this image is then subtracted from incoming images of the same product. *Ideally*, the differences would be zero if the new products are assembled correctly. Difference images for products with missing components would be nonzero in the area where they differ from the golden image. What conditions do you think have to be met in practice for this method to work?

2.29 With reference to Eq. (2-32),

(a)* Give a general formula for the value of K as a function of the number of bits, k, in an image, such that K results in a scaled image whose intensities span the full k-bit range.

(b) Find K for 16- and 32-bit images.

2.30 Give Venn diagrams for the following expressions:

(a)* $(A \cap C) - (A \cap B \cap C)$.

(b) $(A \cap C) \cup (B \cap C)$.

(c) $B - [(A \cap B) - (A \cap B \cap C)]$

(d) $B - B \cap (A \cup C)$; Given that $A \cap C = \varnothing$.

2.31 Use Venn diagrams to prove the validity of the following expressions:

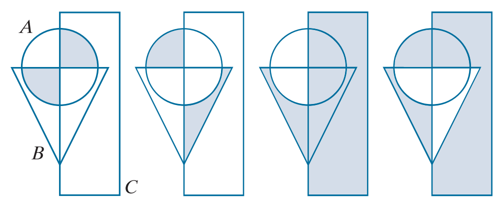
(a)* $(A \cap B) \cup [(A \cap C) - A \cap B \cap C] = A \cap (B \cup C)$

(b) $(A \cup B \cup C)^c = A^c \cap B^c \cap C^c$

(c) $(A \cup C)^c \cap B = (B - A) - C$

(d) $(A \cap B \cap C)^c = A^c \cup B^c \cup C^c$

2.32 Give expressions (in terms of sets A, B, and C) for the sets shown shaded in the following figures. The shaded areas in each figure constitute one set, so give only one expression for each of the four figures.



(a)* (b) (c) (d)

2.33 With reference to the discussion on sets in Section 2.6, do the following:

(a)* Let S be a set of real numbers ordered by the relation "less than or equal to" $(\leq)$. Show that S is a partially ordered set; that is, show that the reflexive, transitive, and antisymmetric properties hold.

(b)* Show that changing the relation "less than or equal to" to "less than" $(<)$ produces a strict ordered set.

(c) Now let S be the set of lower-case letters in the English alphabet. Show that, under $(<)$, S is a strict ordered set.

2.34 For any nonzero integers m and n, we say that m

is divisible by n, written m/n, if there exists an integer k such that $kn = m$. For example, 42 (m) is divisible by 7 (n) because there exists an integer $k = 6$ such that $kn = m$. Show that the set of positive integers is a partially ordered set under the relation "divisible by." In other words, do the following:

(a)* Show that the property of reflectivity holds under this relation.

(b) Show that the property of transitivity holds.

(c) Show that anti symmetry holds.

2.35 In general, what would the resulting image, $g(x, y)$, look like if we modified Eq. (2-43), as follows:

$$g(x, y) = \frac{1}{mn} \sum_{(r,c) \in S_{xy}} T\left[f(r, c)\right]$$

where T is the intensity transformation function in Fig. 2.38(b)?

2.36 With reference to Table 2.3, provide single, composite transformation functions for performing the following operations:

(a)* Scaling and translation.

(b)* Scaling, translation, and rotation.

(c) Vertical shear, scaling, translation, and rotation.

(d) Does the order of multiplication of the individual matrices to produce a single transformations make a difference? Give an example based on a scaling/translation transformation to support your answer.

2.37 We know from Eq. (2-45) that an affine transformation of coordinates is given by

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

where (x', y') are the transformed coordinates, (x, y) are the original coordinates, and the elements of $\mathbf{A}$ are given in Table 2.3 for various types of transformations. The inverse transformation, $\mathbf{A}^{-1}$, to go from the transformed back to the original coordinates is just as important for performing inverse mappings.

(a)* Find the inverse scaling transformation.

(b) Find the inverse translation transformation.

(c) Find the inverse vertical and horizontal shearing transformations.

(d)* Find the inverse rotation transformation.

(e)* Show a composite inverse translation/rotation transformation.

2.38 What are the equations, analogous to Eqs. (2-46) and (2-47), that would result from using triangular instead of quadrilateral regions?

2.39 Do the following.

(a)* Prove that the Fourier kernel in Eq. (2-59) is separable and symmetric.

(b) Repeat (a) for the kernel in Eq. (2-60).

2.40* Show that 2-D transforms with separable, symmetric kernels can be computed by: (1) computing 1-D transforms along the individual rows (columns) of the input image; and (2) computing 1-D transforms along the columns (rows) of the result from step (1).

2.41 A plant produces miniature polymer squares that have to undergo 100% visual inspection. Inspection is semi-automated. At each inspection station, a robot places each polymer square over an optical system that produces a magnified image of the square. The image completely fills a viewing screen of size 80×80 mm. Defects appear as dark circular blobs, and the human inspector's job is to look at the screen and reject any sample that has one or more dark blobs with a diameter of 0.8 mm or greater, as measured on the scale of the screen. The manufacturing manager believes that if she can find a way to fully automate the process, profits will increase by 50%, and success in this project will aid her climb up the corporate ladder. After extensive investigation, the manager decides that the way to solve the problem is to view each inspection screen with a CCD TV camera and feed the output of the camera into an image processing system capable of detecting the blobs, measuring their diameter, and activating the accept/reject button previously operated by a human inspector. She is able to find a suitable system, provided that the smallest defect occupies an area of at least 2×2 pixels in the digital image. The manager hires you to help her specify the camera and lens system to satisfy this requirement, using off-the-shelf components. Available off-the-shelf lenses have focal lengths that are integer multiples of 25 mm or 35 mm, up to 200 mm. Available cameras yield image sizes of 512×512, 1024×1024, or 2048×2048 pixels. The *individual* imaging elements in these cameras are squares measuring $8 \times 8\ \mu$m, and the spaces between imaging elements are 2 μm. For this application, the cameras cost much more than the lenses, so you should use the lowest-resolution camera possible, consistent with a suitable lens. As a consultant, you have to provide a written recommendation, showing in reasonable detail the analysis that led to your choice of components. Use the imaging geometry suggested in Problem 2.6.