Solutions to Chapter 2

Prob. 2.1 — What is unusual about vertebrate photoreceptors, compared with most sensory receptors?

Answer (Prob. 2.1) — In a vertebrate photoreceptor, it is the absence of a signal that indicates the absorption of a photon. This behavior is in contrast with invertebrate photoreceptors, as well as other sensory receptors, in which a signal is emitted to indicate the sensation of a stimulus.

Prob. 2.2 — Explain the purpose of an inverted retina in the human eyeball. Give two additional reasons why the inverted retina does not cause significant distortion in the image.

Answer (Prob. 2.2) — The inverted retina allows the photoreceptors to be replenished by the RPE, which is opaque. The inverted retina does not cause significant distortion because the additional cells through which the light passes are, for the most part, transparent, and because in the fovea the elongated cones bypass these additional cells, so that light falls directly on the cones in the fovea.

Prob. 2.3 — List three parts of the human eyeball that refract light.

Answer (Prob. 2.3) — Together, the cornea, aqueous humor, and lens are responsible for focusing light.

Prob. 2.4 — What are the actual names of the three types of cones, which are colloquially called red, green, and blue?

Answer (Prob. 2.4) — The red, green, and blue cones are more accurately called L-, M-, and S-type cones, respectively, for the long, medium, and short wavelengths of light to which they are most sensitive.

Prob. 2.5 — Define horopter.

Answer (Prob. 2.5) — The horopter is the locus of points in the world that yield zero disparity. For idealized spheroidal human eyeballs, the horopter is the Vieth-Müller circle, whereas for rectified planar sensors, the horopter is at infinity.

Prob. 2.6 — Explain what is meant by foveated vision.

Answer (Prob. 2.6) — Foveated vision refers to the fact that some parts of the scene receive higher resolution by the sensor than other parts. In the human eye, for example, those near the optical axis are sensed with higher fidelity than those at the periphery.

Prob. 2.7 — How do we know that the original signal captured by the rods is compressed by subsequent cells before leaving the eyeball?

Answer (Prob. 2.7) — Since the information travels from the rods to the ganglion cells, but there are more ganglion cells than rods, there must be a compression of information.

Prob. 2.8 — True or false: Your right eye is mapped to the left half of your brain, while your left eye is mapped to the right half of your brain.

Answer (Prob. 2.8) — False. In reality, the right half of your visual field is mapped to the left half of your brain, while the left half of your visual field is mapped to the right half of your brain.

Prob. 2.9 — Match each term on the left with with the lighting condition on the right.

scotopic vision	sunlight
photopic vision	moonlight
mesopic vision	starlight

Answer (Prob. 2.9) — Scotopic vision ↔ moonlight, photopic vision ↔ sunlight, and mesopic vision ↔ starlight.

Prob. 2.10 — Cones do not work in the dark, because they are not sensitive enough. What about the converse: Do rods produce meaningful signals in everyday well-lit conditions? Why or why not?

Answer (Prob. 2.10) — In everyday well-lit environments, the rods are saturated and therefore do not provide meaningful information.

Prob. 2.11 — Draw a labeled diagram of the human visual system, including at least ten parts indicated in bold in the text.

Answer (Prob. 2.11) — Answers may vary.

Prob. 2.12 — Why would it be tempting to conclude that short (blue) wavelengths are less important to the human visual system? Why is this conclusion false?

Answer (Prob. 2.12) — There are fewer S-type cones on the retina than L- or M-type cones. Nevertheless, S-type cones are more sensitive to light than the other types, so that humans are able to distinguish between short wavelengths as well as other wavelengths.

Prob. 2.13 — Suppose the following pairs of numbers indicate the luminances of the left and right halves of a piece of paper (ignore units): 100/101, 200/201, 300/301, 150/160, 250/260, 350/360. Which can be discerned?

Answer (Prob. 2.13) — 150/160, 250/260, and 350/360 can all can be discerned since their normalized difference is greater than 0.01 (1%). That is, $(160-150)/150 = 0.067$, $(260-250)/250 = 0.04$, and $(360-350)/350 = 0.029$. 100/101 can also be discerned, since (101-100)/100 = 0.01, although it is close to the threshold. By similar reasoning, 200/201 and 300/301 cannot be discerned.

Prob. 2.14 — What is a receptive field?

Answer (Prob. 2.14) — The visual receptive field of a neuron is the retinal area in which light influences the neuron's response.

Prob. 2.15 — Which cells in the visual pathway transform the signal similar to the Laplacian of Gaussian (LoG)?

Answer (Prob. 2.15) — Ganglion cells provide a center-surround response similar to the LoG.

Prob. 2.16 — The axons of which cells comprise the optic nerve?

Answer (Prob. 2.16) — The optic nerve is formed from the axons of the ganglion cells.

Prob. 2.17 — What is unique about the lobster eye?

Answer (Prob. 2.17) — The lobster eye has no lens but instead focuses by refraction using an array of reflective mirror-lined tubes.

Prob. 2.18 — True or false: The speed of light is constant no matter what medium it is passing through.

Answer (Prob. 2.18) — False. The speed of light depends upon the medium through which it passes, which is what enables a lens to focus incoming rays of light.

Prob. 2.19 — Which has a longer wavelength, red light or blue light?

Answer (Prob. 2.19) — Red has a longer wavelength than blue.

Prob. 2.20 — Suppose a lens is made of a high index plastic whose index of refraction is 1.74. If the speed of light is approximately $3 \cdot 10^8$ m/s in a vacuum, what is the speed of light (phase velocity) as it passes through the lens?

Answer (Prob. 2.20) — The index of refraction of the plastic is $n = \sqrt{\frac{\epsilon}{\epsilon_0}} = 1.74$, and the speed of light in a vacuum is $\frac{1}{\sqrt{\epsilon_0 \mu_0}} = 3 \cdot 10^8$ m/s. Since the material is non-metallic, $\mu \approx \mu_0$, so the speed of light through the plastic is given by $\frac{1}{\sqrt{\epsilon\mu}} = \frac{1}{\sqrt{\epsilon\mu_0}} = \frac{1}{\sqrt{\frac{\epsilon}{\epsilon_0}} \cdot \sqrt{\epsilon_0 \mu_0}} = \frac{3 \cdot 10^8}{1.74} = 1.72$ m/s.

Prob. 2.21 — What is the plenoptic function?

Answer (Prob. 2.21) — The plenoptic function models all the possible images that could be taken by a camera if it were placed at any possible position and orientation within a scene.

Prob. 2.22 — Describe the essential elements of a pinhole camera.

Answer (Prob. 2.22) — A pinhole camera consists of a center of projection and an imaging surface, which is typically a plane.

Prob. 2.23 — What are the wavelengths of visible light?

Answer (Prob. 2.23) — The wavelengths of visible light range from approximately 380 nm to approximately 720 nm.

Prob. 2.24 — Which has a longer wavelength, radio waves or X-rays? Which is more dangerous, and why?

Answer (Prob. 2.24) — Radio waves have a longer wavelength than X-rays. X-rays are more dangerous, because the amount of energy is proportional to the frequency, and the frequency is inversely proportional to the wavelength. Therefore, X-rays contain more energy than radio waves.

Prob. 2.25 — You are sitting at a stoplight listening to 102.1 FM on your old-fashioned radio, getting a weak signal. You wish to roll the car to improve the signal. How far must you roll to move a wavelength? What is the ratio of this wavelength to that of green light?

Answer (Prob. 2.25) — The wavelength of the radio station signal is $\lambda = \frac{c}{f} = \frac{3 \cdot 10^8 \text{ m/s}}{102.1 \cdot 10^6 \text{ Hz}} = 2.94$ m. Therefore, if you roll your call approximately 1.5 m in either direction, you should expect a noticeable change in the signal received. Since green light has a wavelength of approximately 555 nm, the ratio is $\frac{2.94 \text{ m}}{555 \cdot 10^{-9} \text{ m}} = 5.3 \cdot 10^6$, or about 5 million.

Prob. 2.26 — Is scaled orthographic projection more appropriate for a zoom lens or a fisheye lens? Explain your answer.

Answer (Prob. 2.26) — Scaled orthographic projection is a more valid approximation for a zoom lens, since a zoom lens ensures that all light rays are close to the optical axis and traveling in approximately parallel lines.

Prob. 2.27 — Suppose we have a symmetric thin lens composed of two sections of a sphere glued together, where the radii of both sides are equal. Explain why $\left(\frac{1}{r_1} - \frac{1}{r_2}\right)$ in Equation (2.5) is not equal to zero.

Answer (Prob. 2.27) — Even though the two radii are equal, the sign convention used in the formula causes $r_2 = -r_1$.

Prob. 2.28 — Apply the nonlinear transfer function from both Rec. 709 and sRGB in Equations (2.22) and (2.23) to the values $L = 0.2$, 0.4, 0.6, and 0.8. Compute the difference for each value as a percentage of the answer for Rec. 709.

Answer (Prob. 2.28) — The answers are shown in the following table:

L	$\varphi_{709}(L)$	$\varphi_{sRGB}(L)$	δ	$100\delta/\varphi_{709}(L)$
0.2	0.434	0.484	0.051	11.7%
0.4	0.629	0.665	0.036	5.8%
0.6	0.774	0.798	0.023	3.0%
0.8	0.895	0.906	0.011	1.3%

Prob. 2.29 — Will gamma compression become obsolete now that CRT displays are obsolete? Why or why not?

Answer (Prob. 2.29) — Gamma compression will not become obsolete, because it is fundamentally tied to the fact that human visual perception of light is nonlinear.

Prob. 2.30 — Specify the two sections of the modified gamma function with exponent $\gamma = 0.5$ and threshold $\tau = 0.1$.

Answer (Prob. 2.30) — From Equations (2.20) and (2.21), $m = 1.8781$ and $\epsilon = 0.1878$.

Prob. 2.31 — Explain the idea of effective gamma.

Answer (Prob. 2.31) — The effective gamma is the exponent that best approximates the nonlinear transfer function.

Prob. 2.32 — What is the name of the most popular color filter array (CFA)?

Answer (Prob. 2.32) — The Bayer filter is the most popular color filter array.

Prob. 2.33 — Explain the difference between a field and a frame of video.

Answer (Prob. 2.33) — A frame of video is simply an image. In the old days of interlaced analog video, each frame consisted of an odd field (all the pixel data on the odd rows) and an even field (all the pixel data on the even rows).

Prob. 2.34 — List some similarities and differences between CCD and CMOS sensors.

Answer (Prob. 2.34) — CMOS sensors are smaller, lighter, cheaper, less expensive, more flexible, and consume less power than CCD sensors. They both use photodiodes to convert photons to electricity.

Prob. 2.35 — How much is a CMOS sensor affected by blooming?

Answer (Prob. 2.35) — CMOS sensors are not affected by blooming.

Prob. 2.36 — Why does black have the value 16 and not 0 in a digital image?

Answer (Prob. 2.36) — The values 0 through 15, known as the footroom, are used to prevent clipping in case the output of a filter is outside the original range of pixel values.

Prob. 2.37 — Explain what is meant by a Lambertian surface. What is albedo? Which is more likely to be Lambertian: a piece of cloth or a shiny piece of metal?

Answer (Prob. 2.37) — A Lambertian surface is an idealized matte surface that reflects luminance equally in all directions. Albedo is the average reflection coefficient of a surface. A piece of cloth is much closer to being Lambertian than a shiny piece of metal.

Prob. 2.38 — Which radiometric quantity is appropriate for a ray of light? What is the corresponding photometric quantity?

Answer (Prob. 2.38) — Radiance and luminance are the radiometric and photometric quantities, respectively, that do not vary along a ray of light and are therefore appropriate measures of the ray.

Prob. 2.39 — Explain why a thermal infrared camera is able to measure the heat emanating from people and animals.

Answer (Prob. 2.39) — The motions of molecules in the bodies of people and animals due to thermal energy cause infrared wavelengths of light to be emitted, which can then be sensed by a thermal infrared camera.

Prob. 2.40 — Suppose I am standing on the shore looking at a body of water. If the water has an index of refraction of 1.33, at what angle will I experience total internal reflection? How does the answer change if I am underwater looking up? In both cases express the angle with respect to the vertical axis.

Answer (Prob. 2.40) — If I am standing on the shore, total internal reflection is not possible, because the index of refraction of air is less than that of water. However, if I am underwater looking up, the angle is given by using Snell's law of refraction, Equation (2.30), and setting the transmitted angle to 90°:

$$\sin 90^\circ = 1 = \frac{1.33}{1} \sin \theta_i,$$

or $\theta_i = \arcsin(1/1.33) = 48.75^\circ$.

Prob. 2.41 — A light field, which is 4D, can be represented as a 2D array of tiny 2D images. Indeed, this is the representation used by a light field camera. Explain how a 2D array of microlenses placed in front of the image sensor might be able to accomplish this.

Answer (Prob. 2.41) — Each microlens causes a small image of the scene to be captured in a portion of the sensor. Since each of these images is captured from a slightly different direction, the representation is that of a light field.

Prob. 2.42 — Mathematically show the two sufficient conditions for scaled orthographic projection to closely approximate perspective projection. (*Hint*: Show from Equations (2.1)–(2.2) that bounding the error $\left| f\frac{x_w}{z_w} - f\frac{x_w}{z_0} \right| < \epsilon$ for some nominal depth z_0 implies $\frac{x}{z_0}\frac{|\delta_z|}{z_0+\delta_z} \le \frac{\epsilon}{f}$, where $z = z_0 + \delta_z$, and ϵ is a constant. Then interpret the result in terms of the two sufficient conditions.)

Answer (Prob. 2.42) — Under perspective projection, a point at location (X, Y, Z) in the world will be imaged at location $(f\frac{X}{Z}, f\frac{Y}{Z})$ in the image plane, where f is the focal length of the camera. If we assume that the object is at a nominal (i.e., approximate) depth Z_0, then scaled orthographic projection will project the same point to location $(f\frac{X}{Z_0}, f\frac{Y}{Z_0})$, where $\frac{f}{Z_0}$ is the scale factor.

The error in the x-direction resulting from the scaled orthographic projection approximation is given by

$$\left| f\frac{X}{Z} - f\frac{X}{Z_0} \right|.$$

Thus, in order to bound the error by a certain amount ϵ, we must have

$$\left| \frac{X}{Z} - \frac{X}{Z_0} \right| \leq \frac{\epsilon}{f}.$$

Now let's assume that $Z = Z_0 + \Delta_Z$. That is, Δ_Z is the difference between the actual depth and the nominal depth. Then,

$$\begin{aligned} X\left| \frac{1}{Z_0 + \Delta_Z} - \frac{1}{Z_0} \right| &\leq \frac{\epsilon}{f} \\ X\left| \frac{-\Delta_Z}{Z_0(Z_0 + \Delta_Z)} \right| &\leq \frac{\epsilon}{f} \\ \frac{X}{Z_0}\frac{|\Delta_Z|}{Z_0 + \Delta_Z} &\leq \frac{\epsilon}{f}, \end{aligned}$$

which can be approximated by

$$\frac{X}{Z_0}\frac{|\Delta_Z|}{Z_0} \leq \frac{\epsilon}{f} \tag{2.31}$$

if $\Delta_Z \ll Z_0$. Similarly, the error in the y-direction will be bounded by ϵ if

$$\frac{Y}{Z_0}\frac{|\Delta_Z|}{Z_0} \leq \frac{\epsilon}{f} \tag{2.32}$$

Combining the two equations yields

$$\frac{R}{Z_0}\frac{|\Delta_Z|}{Z_0} \leq \frac{\epsilon}{f} \tag{2.33}$$

where $R \equiv \sqrt{X^2 + Y^2}$ is the distance from the point to the optical axis. Therefore, we see that two conditions are sufficient (but not necessary) in order for the error to be bounded:

- $R \ll Z_0$, or
- $|\Delta_Z| \ll Z_0$

These conditions state that the image of the object will be well approximated by scaled orthographic projection if no point on the object is far from the optical axis (relative to the nominal depth of the object), and/or if the variation in depth of the object is small (relative to the nominal depth of the object). Notice from the equation above, these two conditions are not independent: the smaller the variation in the object's depth, the farther it is allowed to be from the optical axis, and vice versa. Thus, the projection of an infinite plane which is parallel to the image plane will be modeled well (and indeed, perfectly) by scaled orthographic projection. Similarly, the projection of a very long cylinder whose axis is parallel to the optical axis will be well approximated by scaled orthographic projection as long as its maximum distance to the optical axis is small.

Prob. 2.43 — Suppose a person is standing in front of a pinhole camera so that their face occupies a certain width in the image. If the person moves laterally so that the perpendicular distance to the camera is maintained, does the face width in the image change? Why or why not?

Answer (Prob. 2.43) — The face width does not change, which can be shown using similar triangles.

Prob. 2.44 — Consider two thin lenses, both symmetric and made of the same material. If one lens has twice the focal length of the other, what is the relationship between their radii?

Answer (Prob. 2.44) — If we let f and f' be the two focal lengths, then from Equation (2.5),

$$\begin{aligned}\frac{1}{f} &= (n-1)\left(\frac{2}{r}\right)\\ \frac{1}{f'} &= (n-1)\left(\frac{2}{r'}\right).\end{aligned}$$

Since $f' = 2f$,

$$\frac{r'}{2(n-1)} = 2\frac{r}{2(n-1)},$$

or $r' = 2r$. Therefore, if the focal lengths are related by a factor of two, the radii are also related by a factor of two.

Prob. 2.45 — In the case of a thin lens, what condition is necessary in order for a distinction between the focal points and nodal points to be important?

Answer (Prob. 2.45) — The surrounding medium must be different on the two sides; otherwise they coincide.

Prob. 2.46 — Suppose a camera has an f-number of 8. What is the aperture, expressed as a function of f? What is the aperture of a camera whose light gathering ability is twice as great, also expressed as a function of f?

Answer (Prob. 2.46) — The f-number is defined as f/d. Since $f/d = 8$, then $d = f/8$, or the aperture of the camera is $f/8$. To gather twice as much light, the area of the aperture should be twice as large. Since the area is proportional to the square of the diameter, this means that the diameter must increase by a factor of $\sqrt{2}$. The aperture of such a camera is therefore $f\sqrt{2}/8 = f/5.7$.

Prob. 2.47 — List the four types of vignetting. Under what conditions are they important?

Answer (Prob. 2.47) — Natural vignetting is only important if the lens is not covered by a graduated neutral density filter. Optical vignetting, mechanical vignetting, and pixel vignetting can be important near the boundaries of the image when a large aperture is used.

Prob. 2.48 — List the three ways of transferring energy. Of these, which one can travel through a vacuum?

Answer (Prob. 2.48) — Energy is transferred by conduction, convection, and radiation. Only the latter can travel through a vacuum.

Prob. 2.49 — What is the name of the set of equations that underlie all applications using electromagnetism?

Answer (Prob. 2.49) — All applications using electromagnetism are governed by Maxwell's equations.

Prob. 2.50 — Derive the homogeneous vector wave equations in Equations (2.26)–(2.27) from Maxwell's equations in Equations (2.24)–(2.25). (Hint: It is *not* necessary that you understand what the divergence and curl operators actually do. Simply take the curl of (2.24), and apply the fact that the curl operator is linear. You will need the vector identity $\nabla \times \nabla \times E = \nabla(\nabla \cdot E) - \nabla^2 E$, and similarly for B.)

Answer (Prob. 2.50) — Take the curl of (2.24) and (2.25), respectively. Since both differentiation and curl are linear, we can swap their order, then substitute (2.25) and (2.24), respectively, to get

$$\begin{aligned}
\nabla \times (\nabla \times E) &= \nabla \times \left(-\frac{\partial B}{\partial t}\right) = -\frac{\partial}{\partial t}(\nabla \times B) = -\frac{\partial}{\partial t}\left(\epsilon_0\mu_0\frac{\partial E}{\partial t}\right) = -\epsilon_0\mu_0\frac{\partial^2 E}{\partial t^2} \\
\nabla \times (\nabla \times B) &= \nabla \times \left(\epsilon_0\mu_0\frac{\partial E}{\partial t}\right) = \epsilon_0\mu_0\frac{\partial}{\partial t}(\nabla \times E) = \epsilon_0\mu_0\frac{\partial}{\partial t}\left(-\frac{\partial B}{\partial t}\right) = -\epsilon_0\mu_0\frac{\partial^2 B}{\partial t^2}
\end{aligned}$$

Applying the identity $\nabla \times \nabla \times E = \nabla(\nabla \cdot E) - \nabla^2 E$ and recognizing that $\nabla \cdot E = 0$ from (2.24), and similarly for B, we have

$$\begin{aligned}
\nabla^2 E &= \epsilon_0\mu_0\frac{\partial^2 E}{\partial t^2} \\
\nabla^2 B &= \epsilon_0\mu_0\frac{\partial^2 B}{\partial t^2},
\end{aligned}$$

which is the desired result.