Vignetting

[The first part of this material was presented in lecture 6. See the slides from the lecture 6 lecture.]

Consider a room of height f meters. The ceiling contains a disk-shape hole (skylight) of diameter A, which allows light from the sky to enter the room. What is the irradiance at a point $\mathbf{x} = (x, y, f)$ on the floor of the room ? You may assume that the light from the sky has constant radiance L_{sky}.

Let the center of the skylight be the origin and so $(0, 0, f)$ is the point on the floor of the room that is directly below the center of the hole. The direction $\mathbf{l}(x, y)$ from $\mathbf{x}$ to the center of the hole is

$$\mathbf{l}(x,y) = \frac{(-x,-y,-f)}{\sqrt{x^2+y^2+f^2}}.$$

The surface normal on the floor is $\mathbf{n} = (0, 0, -1)$. Let α be the angle between the normal and the direction to the center of the hole. Then

$$\cos\alpha \;=\; \mathbf{n}(x,y)\cdot\mathbf{l}(x,y) \;=\; \frac{f}{\sqrt{x^2+y^2+f^2}}.$$

The solid angle of the hole as seen from $\mathbf{x}$ is

$$\Omega \;=\; \frac{\text{area of hole}}{\text{distance squared to hole}} \cdot \;\cos\alpha$$

where the $\cos\alpha$ factor is there because the hole is seen obliquely and is thus foreshortened, and so the solid angle is smaller.

The area of the hole is $\pi(\frac{A}{2})^2$ and so

$$\begin{aligned}\Omega &= \frac{\pi A^2}{4}\frac{1}{(x^2+y^2+f^2)}\;\cos\alpha \\ &= \frac{\pi A^2}{4}\frac{(\cos\alpha)^2}{f^2}\;\cos\alpha \\ &= \frac{\pi A^2}{4f^2}(\cos\alpha)^3\end{aligned}$$

Assume that the A is small enough relative to f that we can approximate the directions from (x, y, f) to the different points in the area of the hole with one vector $\mathbf{l}(x, y)$, namely the direction vector to the center of the hold. Thus,

$$\begin{aligned}E(x,y) &= \int L_{src}\;\mathbf{n}(x,y)\cdot\mathbf{l}(x,y)d\Omega \\ &\approx L_{src}\;\mathbf{n}(x,y)\cdot\mathbf{l}(x,y)\;\Omega\end{aligned}$$

Substituting into Ω gives

$$E(x,y) = L_{src}\;(\cos\alpha)^4\;\frac{\pi A^2}{4f^2}\;.$$

As α increases from 0, you can see that the irradiance will fall off. The corners of the room will receive less illumination that the point directly below the skylight.

Vignetting on the sensor plane (new this lecture)

A similar argument as above can be used to derive an expression for the irradiance on the sensor plane behind a lens. Suppose the camera is focussed at infinity and so the sensor plane is at a distance $Z_s = f$ from the center of the lens. Each point (x, y) on the sensor plane receives light from a set of parallel rays in the scene, arriving *at the lens* from direction $\mathbf{l}(x, y)$. These rays are redirected into a cone behind the lens, which is focussed onto the sensor plane. The formula for irradiance $E(x, y)$ is slightly different from above, since each point (x, y, f) on the sensor plane is receiving light from a set of rays in the scene that all came from the same direction $\mathbf{l}(x, y)$. These rays were parallel when they arrived at the lens, and then focussed into a cone with base at the lens and tip/apex at (x, y, f). Thus, rather than the rays having constant radiance L_{src} as on the previous page, they have a radiance that varies with (x, y), and so:

$$E(x, y) = \frac{L(x, y)\pi \ (\cos\alpha)^4}{4} (\frac{A}{f})^2$$

where α is defined as before. Again, the $(\cos\alpha)^4$ term has the effect of darkening the edges of the image, relative to the center of the image. It is called *vignetting*.

Note that image irradiance is inversely proportional to the square of the f-number (recall $N = f/A$). Typical cameras allow users to choose the f-numbers as the following powers of $\sqrt{2}$,

$$\sqrt{2}, \ \sqrt{4}, \ \sqrt{8}, \ \sqrt{16}, \ \sqrt{32}, \ \ldots$$

and these values are presented to the photographer as

$$1.4, \ 2, \ 2.8, \ 4, \ 5.6, \ 8, \ 11, \ 16, \ 22, \ \ldots.$$

Each successive f-number (or "f-stop") *decreases* the image irradiance by a factor of 2 at each pixel.

When a photographer changes the f-stop (often by turning a dial on the camera), the photography is typically manipulating the aperture width A, with f fixed. Thus, it is the aperture that is really going in steps that are a factor $\sqrt{2}$.

The other way to manipulating the f-number is to change the focal length of the camera. This can be done by replacing the lens. Or many cameras have a "zoom" feature that changes the focal length. (All cameras - even ones with a fixed focal length - in fact use multiple lenses. But when you put these lenses togther, you get an effective single lens, which you can think of as a virtual lens. For cameras with zoom, the focal length can be varied by moving the multiple lenses relative to each other. If you want to know how this works, you would need to read an optics book.)

When you change the focal length f, you also change the angular field of view of the camera. Let the sensor width be fixed W_x in the x direction and W_y in the y direction, measured in mm. Then if the camera is focussed at infinity, the angular field of view in the x and y directions satisfy:

$$tan(\frac{\theta_x}{2}) = \frac{W_x}{2f}$$

$$tan(\frac{\theta_y}{2}) = \frac{W_y}{2f}$$

Thus increasing f causes a decrease in θ_x and θ_y. In turn, for a fixed aperture A, increasing f causes a decrease in the image irradiance. This decrease in irradiance is due to the fact that less of the scene (and hence less light) gets imaged on the sensor area.

The argument I just gave glosses over one subtlety, however. When you change f, a different part of the scene will be focussed onto any given pixel (x, y). So, you cannot just say that increasing f will reduce $E(x, y)$ for that pixel since the new part of the scene that becomes focusses on that pixel might have much higher (or lower) radiance $L(x, y)$. To avoid this subtle issue in the argument, consider the case where the value of $L(x, y)$ happens to be the for pixel (x, y) for the two values of f we are comparing. In this case, increasing f leads to a decrease in $E(x, y)$ according to the above argument.

Color and RGB

[ASIDE: In the lecture, I introduced color only at the end of the lecture.]

Digital color images have three intensity values per pixel (RGB) which define a (red,green,blue) color triplet of intensities at each pixel. We would like to understand better what these intensities mean. A good place to start our examination of color is with Isaac Newton and his prism experiment (late 17th century). Newton observed that when a beam of sunlight is passed through a prism, the beam is spread out into a fan of different color lights – like a rainbow. The theory that explains Newton's experiments is now well known. Light is composed of electromagnetic waves, with wavelengths ranging from 400-700 nanometers[1] and these waves are refracted by different amounts when they pass through the prism.[2]

Up to now we have defined radiance as a function of spatial position and direction. But radiance also depends on wavelength. That is, the radiance of light at a point $\mathbf{x}$ and from direction $\mathbf{l}$ and at a wavelength λ is $L(\mathbf{x}, \mathbf{l}, \lambda)$. Often we use the term *spectral power distribution* (SPD) to talk about something that is a function of wavelength. We will see other examples later.

The radiance of light coming from a real source such the sun, a light bulb, or a blue sky each have a different spectra $L(\lambda)$. Some light sources have more power at longer wavelengths, whereas others have more power at shorter wavelengths. A candleflame has relatively more power at long wavelengths than does direct sunlight. A blue sky has relatively more power at short wavelengths than does direct sunlight. Sources from many artificial lights (fluorescents) have most of their power concentrated in a very small range of wavelengths.

When a surface reflects light, it changes the spectrum, namely it absorbs some wavelengths more than others. The BRDF thus also has a wavelength dependence and so we would write:

$$L(\mathbf{x}, \mathbf{l}_{out}, \lambda) = \int \rho(\mathbf{x}, \mathbf{l}_{in}, \mathbf{l}_{out}, \lambda) \; L_{in}(\mathbf{x}, \mathbf{l}_{in}, \lambda) \; \mathbf{n}(\mathbf{x}) \cdot \mathbf{l}_{in} \; d\Omega_{in}.$$

If the surface is Lambertian, then this equation reduces to

$$L(\mathbf{x}, \lambda) = \int \rho(\mathbf{x}, \lambda) \; L_{in}(\mathbf{x}, \mathbf{l}_{in}, \lambda) \; \mathbf{n}(\mathbf{x}) \cdot \mathbf{l}_{in} \; d\Omega_{in}.$$

Notice that there is no interaction between different wavelengths. The power is reweighted wavelength by wavelength. The function $\rho(\mathbf{x}, \lambda)$ can (and typically does) vary over points on a surface

[1] A nanometer is 10^{-9} meters. e.g. for a wavelength of 500 nm, you need about 2000 "waves" to cover one millimeter.

[2] The same rainbow effect occurs in a lens too, but typically the spread is very small and anyhow lens designers correct for it – so its not a problem.

$\mathbf{x}$. This just corresponds to our common experience that surface materials are non-uniform – they have reflectance (color) variations across them.

What can we say about $\rho(\mathbf{x}, \lambda)$ for the case of a Lambertian surface? If the values of $\rho(\mathbf{x}, \lambda)$ are very small, than the radiance leaving the surface will be small. In this case, the surface appears *dark* colored – it reflects relatively little light over all wavelengths. On the other hand, if the values of $\rho(\mathbf{x}, \lambda)$ are large, then the surface will be light colored – the lightest surface being white. Notice that there are limits on how large $\rho()$ can be, since a surface cannot reflect more light than arrives at it!

What if $\rho(\mathbf{x}, \lambda)$ has smaller values at short wavelengths but relatively larger values at long wavelengths? In this case, there will be relatively more long wavelength light reflected, and the reflected light will be more "reddish" (long wavelength) than the incoming light.

Trichromatic images (RGB)

Digital cameras make images with red, green, and blue components (RGB). This is done by dividing each pixel into four sub-pixels, and putting a partially transparent filter in front of each sub-pixel. The filters preferentially remove the energy at different wavelengths. The R filter allows the longer wavelengths through but absorbs the short and medium. The G filter allows the medium wavelengths through, but absorbs the short and long. The B filter allows the short through, but absorbs the medium and long. There are typically two G subpixels and only one R and B. (See the Bayer filter `http://en.wikipedia.org/wiki/Bayer_filter`.)

How can we relate this filtering idea to the terminology we introduced earlier? Let $E(\mathbf{x}, \lambda)$ be the (spectral) irradiance of the light reaching a sub-pixel. When we place an R,G, or B filter in front of that sub-pixel, some fraction $C_R(\lambda), C_G(\lambda)$, or $C_B(\lambda)$ of the light is allowed through. The spectral irradiance reaching the pixel at wavelength λ would be $C_{R,G,B}(\lambda)\ E(x, y, \lambda)$. Adding up the irradiance over all λ gives an RGB triplet at each pixel:

$$E_{R,G,B}(x, y) = \int C_{R,G,B}(\lambda)\ E(x, y, \lambda) d\lambda$$

Note that the (sub)pixel itself cannot distinguish different wavelengths and has no "memory" of what the distribution $E(x, y, \lambda)$ was before it was filtered.

Exposure and shutter speed

Another important factor in determining the amount of light reaching each sensor is the time duration t over which the sensor is exposed to the light. Recall that we have been talking about the power (energy per unit time). To calculate the total light energy, we need to multiply the irradiance by the time t that the sensor is exposed to the light.

When you expose a (sub-)pixel to a certain irradiance for a certain time t, the pixel receives a certain power. This is called the *exposure*, $E_{R,G,B}\ t$.

In photography, the inverse of the time t that the sensor is exposed to the light is called the *shutter speed*. A shutter speed of 60, for example, means that the sensor is exposed for 1/60 second. It is standard in photography to allow the user to vary the shutter speed ($\frac{1}{t}$) roughly in powers of 2, namely,

$$1\ , 2\ , 4\ , 8\ , 15\ , 30\ , 60\ , 125\ , 250\ , 500,\ 1000,\ \ldots$$

Of course, the actual values are powers of 2, i.e. e.g. 1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024.

Recall that typical cameras allow users to vary the f-numbers in powers of $\sqrt{2}$, which steps the irradiance by factors of 2. This suggests that if a user decreases the f-number by one step (a factor of 2 increase in irradiance), and increases the shutter speed by one step (a factor of 2 decrease), then the exposure should stay the same.

Camera response function, and dynamic range

This exposure $E_{RGB}(x, y)t$ determines the RGB intensity values for the pixel at (x, y). Typically, one has 8 bits for each of RGB and so the values are in a range 0 to 255.

The RGB values $I_{RGB}(x, y)$ of the image produced by the camera depend on the exposure $E_{RGB}\ t$ via some non-linear response function T which is sometimes called a *tone mapping function.* Typically this function $T()$ is approximately the same for R, G, and B. Then,

$$I_{RGB} = T\ (E_{RGB}\ t)$$

where

$$T : [0, \infty) \rightarrow \{0, 1, 2, \ldots, 255\}.$$

The function $T()$ takes the value 0 for low exposures (underexposed) then ramps up over some range of exposures, and then maxes out at 255 beyond some maximum exposure (overexposed). See the slides.

In many scenes that you wish to photograph, the range of image irradiances is great. The result is that some regions of the image will either be overexposed or underexposed. One of the big challenges of shooting a good photograph is avoiding areas that are under or over exposed. This requires choosing the camera position and orientation in 3D, and also setting the camera parameters: the focal length, the aperture, and the shutter speed.

The *dynamic range* of a tone mapping function (or camera) is the range of exposures values that it can distinguish. This is roughly the range of exposure values that do not under or overexpose the pixels. Typically one defines dynamic range as a ratio of max:min exposures that the camera can measure without over- or under-exposing.

A similar concept is the *dynamic range of a scene* to be the max:min ratio of exposures that this scene produces. Notice that since we are considering dynamic range to be a ratio, we can define it as a ratio of max:min image irradiances, or max:min scene radiances too.

High dynamic range imaging (HDR)

Photographers are often faced with scenes have a dynamic range that is greater than the dynamic range of the camera. There is no way to shoot such a scene without over or underexposing some pixels. Methods have been developed to expand the dynamic range of images which require taking multiple photos with different shutter speeds.[3] Here I will sketch out the very basic idea.

This method is called *high dynamic range imaging.*

To avoid underexposing the darker regions of the scene, you need a slower shutter speed (large t, or small $\frac{1}{t}$). To avoid overexposing very bright regions of the scene, you need a very fast shutter

[3] "Recovering High Dynamic Range Radiance Maps from Photographs" by Paul Debevec and Jitendra Malik, SIGGRAPH 1996. `http://www.debevec.org/Research/HDR/`

speed (small t, or large $\frac{1}{t}$). The idea of high dynamic range imaging is to take a sequence of images with many shutter speeds, so that pixels will be properly exposed (not near 0 or 255) in at least one of the images. Since any pixel will be properly exposed for *some* shutter speed $1/t$, one can use that image to estimate the exposure for that pixel. Since the shutter speed and sensor area size is known, the exposure gives the irradiance.

To understand better what is happening here, consider the log (base 2) of the exposure,

$$\log(T^{-1}(I_{RGB})) \;=\; \log(E_{RGB}t) \;=\; \log E_{R,G,B}(x,y) - \log(\frac{1}{t})$$

and note that changing the shutter speed just shifts the log of the exposure. If the shutter speeds are consecutive factors of 2 (which is typical with cameras), then doubling the shutter speed $1/t$ decrements the log exposure (at each pixel) by 1. See the slides for another illustration using a real image example.