Slant and tilt

One familiar perceptual property of an overall surface shape is its depth gradient, or more intuitively, how it slopes away from you. Does it slope to the right, or to the left, or downward (a ceiling) or upward (a floor) ?

Slope depends on the first order properties of a surface, that is, the depth gradient – or equivalently, the surface normal. So, working with camera axes XYZ, let's write the slope of the surface in terms of the depth gradient and normal. We work with spherical coordinates: the north pole is in the $-Z$ direction i.e. opposite to the "straight ahead". We call [1] the latitude σ (also known as *slant*), and we call the longitude τ (known as *tilt*). Slant σ can go from 0 to $\pi/2$ or 90 degrees. Tilt τ can go from 0 to 2π or 360 degrees.

The depth gradient $\nabla Z \equiv (\frac{\partial Z}{\partial X}, \frac{\partial Z}{\partial Y})$ can be written

$$\nabla Z \; = \; |\nabla Z| \; (\cos\tau, \sin\tau)$$

where

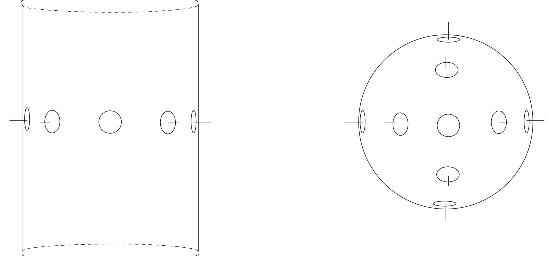
$$|\nabla Z| = \sqrt{(\frac{\partial Z}{\partial X})^2 + (\frac{\partial Z}{\partial Y})^2} \; = \; \tan\sigma$$

and so

$$\nabla Z = \tan\sigma(\cos\tau, \sin\tau).$$

To see why $|\nabla Z| = \tan\sigma$, take the case that tilt is 0, and consider the following figure.

Below are two sketches illustrating how slant and tilt vary on a cylinder and on a sphere. For the cylinder, tilt is either 0 (right half) or π (left half) and slant varies from 0 (midline) to $\pi/2$ (left and right edge). For the sphere, all possible values of slant and tilt are present.



[1]The vision community again has overloaded the symbol σ.

The above definition of slant and tilt make intuitive sense for objects that you are "looking at" that is, objects that are near the optical axis. But the intuition does hold for objects that are not near the optical axis. For example, if I ask you to look straight at a wall, then the slant is 0 degrees. However, note that points on the wall that are off the optical axis will be seen at an oblique angle - though the slant of the wall wouldn't change. If I asked you to "look at" one of these off axis points, you would rotate your eye. But note that this changes the X, Y, Z coordinate system and so the definition of the slant of the wall would change. The point here is is that we can define slant and tilt (relative to our current camera axes XYZ) but we must realize this definition only correspond to our intuition of the direction of the slope τ and the amount of the slope σ in the case for the surface that is intercepted by the optical axis.

Plane under perspective

The sketches on the previous page use little disks dropped on the surface as a way to gauge (measure)[2] the local surface normal. We will explore this technique further to understand the notion of visual slant and how it affected by perspective projection.

Recall from lecture 1 that we can write a plane as

$$Z = AX + BY + C$$

and, multiplying by $\frac{f}{Z}$ we can write it in image plane coordinates as

$$f = Ax + By + C\frac{f}{Z}.$$

Note $\nabla Z = (A, B)$. Also recall that if we let $Z \to \infty$ we get the *vanishing line*

$$f = Ax + By.$$

It is not difficult to show that the distance in the image from any (x, y) to this vanishing line is $|\frac{Cf}{Z\sqrt{A^2+B^2}}|$. (Recall bottom of p. 1 in lecture 13.)

As an example, suppose $\tau = \frac{\pi}{2}$ so $A = 0$ which is the common case that the eye is looking at some point on the ground. In this case,

$$f - By = C\frac{f}{Z}.$$

Writing in terms of the slant and tilt, we have

$$Z = Y \tan\sigma + C$$

and so

$$f = y \tan\sigma + C\frac{f}{Z}.$$

[2]Indeed a recent and powerful method used in visual perception experiments (psychology) for estimating how people perceive shape is to use computer graphics to superimpose little graphically rendered virtual disks on a surface and have users adjust the disks using a computer mouse so that they appear to lie on the surface. See Koenderink et al's fascinating study "Surface Perception in Pictures" (1992) which introduced this technique

Let's drop a disk of diameter D somewhere onto the plane. The farthest and nearest points on the disk define $\Delta Y = D\cos\sigma$, and $\Delta Z = D\sin\sigma$, and so and

$$\begin{aligned}\Delta y \tan\sigma &= Cf\Delta(\frac{1}{Z}) \\ &\approx -C\frac{f}{Z^2}\Delta Z\end{aligned}$$

and so

$$|\Delta y| \approx |CD\frac{f}{Z^2}\cos\sigma|.$$

The width of the disk in the X direction is $\Delta X = D$, and $x = f\frac{X}{Z}$, so we get

$$\Delta x = \frac{f\Delta X}{Z} = \frac{fD}{Z}$$

and so the width of the disk goes like $\frac{1}{Z}$ and the image aspect ratio of the projected disk is

$$\frac{\Delta y}{\Delta x} =\approx |\frac{C\cos\sigma}{Z}|$$

which also goes like $1/Z$. In particular, as we approach the horizon, the aspect ratio goes to 0.

Below is a photograph I took of some upside-down dessert plates on my dining room table. The plates are all the same size in 3D and they are spaced evenly on the table. You can see that both the image size and aspect ratio of the plates varies. The image width Δx of the near plate is about twice as great as the far plate, and the aspect ratio $\Delta y/\Delta x$ of the far plate is about 1/3 whereas that the of the near plate is just under 1.

Shape from texture

Surfaces in the world are not covered with little disks, obviously, such as in the figures above. In man-made environments though, surface often *do* have such little markers on them which are equivalent to disks, for example, tiled floors and ceilings, and buildings with regularly placed windows or bricks. For these surfaces, it is possible to use the systematic distorsion of the tile elements to infer the overall surface shape.[3] This observation leads us to the problem known as *shape from texture*.

Consider a surface with lots of reflectance variations and/or shading and shadows on it. The reflectance markings could be like wood grain, fur, rust, scratches, dents, dirt, stains, etc. Or it could be like what you seen on a grassy lawn partly covered in leaves. Note that although we don't have a inform grid of plates or tiles, the same type of deformations described above are present, namely the individual texture elements (say leaves) are smaller and more foreshortened toward the horizon.

How could you infer the shape of the surface (say the slant and tile of the ground) from such an image? One idea is to examine the second moment matrix. Since the texture elements are more foreshorted near the horizon, one might imagine that there would be more vertical gradients near the horizon. Thus, a system change in the ratio of the eigenvalues of the second moment matrix might be an indication of the tilt of the surface. And if you could measure the size of the elements (using a blob detector) then you might use the variations in the size as well. This is just a sketch of an idea. For more details, see e.g.

- Witkin, "Recovering surface shape and orientation from texture" Artificial Intelligence, 1991.
- Lindeberg and Garding, "Shape from texture from a multi-scale perspective", ICCV 1993.

[3]Indeed such surfaces also give you vanishing points!

Depth from defocus

One last "Shape from X" problem to be discussed involves optical blur, and is called *depth from defocus*. Recall from lecture 4 the expression for image irradiance from a thin lens.

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

where the expression in brackets in the area of the aperture. To derive this expression we assumed that all rays that pass through the lens and arrive at the image point (x, y) have the same radiance $L(x, y)$. For surfaces that have significant texture and that do not lie exactly on the image plane, however, this assumption will typically *not* be true.

To obtain a proper expression for image irradiance in the presence of blur, we could consider each point (x, y) on the sensor plane and consider where the rays that reached (x, y) started out, and then average the radiances of those rays. (Think of this as reverse projection.) Alternatively, we could forward project, by considering the rays that all come from the same point in the scene, and follow those rays until they spread out on the sensor plane. We will take the latter approach.

Consider the image irradiance $E_p(x, y)$ for a very small aperture (p is for *pinhole*). We can think of this irradiance function as being in focus since the aperture is small by assumption and the blur width is (recall lecture 4):

$$\Delta X_i = A \mid Z_s \ (\ \frac{1}{f} - \frac{1}{Z_o} \ ) \ - 1 \mid . \tag{1}$$

Of course, the irradiance will be very low when the aperture is so small. So, we are interested in what happens with a larger aperture.

To keep the notation simple, let's suppose that the blur that arises has a Gaussian distribution over space, with standard deviation σ which is proportional to ΔX_i. We model image irradiance by

$$E(x,y) = \int E_p(x',y') \ \sigma^2 G_\sigma(x-x', y-y')dx'dy'.$$

We use the blur function $\sigma^2 G_\sigma(x, y)$ rather than $G_\sigma(x, y)$ since, as we open the aperture, we allow more light in. i.e. we are not merely averaging the irradiance over the blur width neighborhood, but we are also increasing the irradiance as A increases. (When taking a photograph, if one increases the aperture, one typically also increases the shutter speed to counteract the increase in irradiance, and thus maintain exposure.)

If the scene has constant depth Z, then the blur width will be constant and so will σ and the integral will be a convolution. A weaker assumption is that, in local image region, the depth is approximately constant, and so the above integral can be locally approximated as a convolution

$$E(x,y) = (E_p * \sigma G_\sigma)(x,y).$$

Various methods have been invented for estimating the blur width in a local region in an image. Note that since blur width is related to depth (and various camera parameters), it is possibly to use the amount of blur in a local region for estimating depth.

For a few examples of papers related to focus (including a few classics, and a very recent paper of mine that contains many citations to other focus papers, see
`http://www.cim.mcgill.ca/~langer/558/TermPaperTopics.txt`