Andrew Johnson

Prof. Daniels

English

Saturday, February 20, 2021

The Perspective Divide

Many modern technologies and ideas have their roots set in the renaissance era. After hundreds of years of adapting and evolving, these technologies may not look the same on the surface; however, when examined closely, they have much in common with innovations made by renaissance scientists and artists. One of the most prevalent of these ideas has been used by everything from these ancient artists, to modern movie and video game studios: perspective projection. Perspective projection is a solution to a much broader problem: three-dimensional rendering - the process of representing three-dimensional scenes on a two-dimensional image.

First imagined in the 16th century by Albrecht Dürer, the most conceptually simple solution to this problem is now known as ray tracing. First, imagine an individual a ray of light as it bounces around the world until it hits a virtual eye. Now repeat this step billions of times, and keep track of how each ray hits a virtual camera, until you have an image. The math to accomplish this is quite simple, but it has one major drawback: time. The ray-tracing method has only very recently become useful with the advancement of modern computers, but even today true ray tracing is only found in large movie studios like Pixar, where supercomputers spend up to a day to render a single still image from one of their movies[1].

To find a more efficient way to do this, modern scientists can turn to the techniques used by renaissance artists. Artists use a variety of techniques to emulate foreshortening, the effect by which distant objects appear smaller; however, one is particularly relevant: single-point perspective. The classic example for single-point perspective is straight railroad tracks going off into the distance, at some point the tracks become so small in your vision that they become a single point on the horizon near the center of the image, known as the vanishing point. If you then include any other object with sides parallel to the railroad tracks, such as a cube, and extend those edges, you will find they also converge on the vanishing point (figure 1).

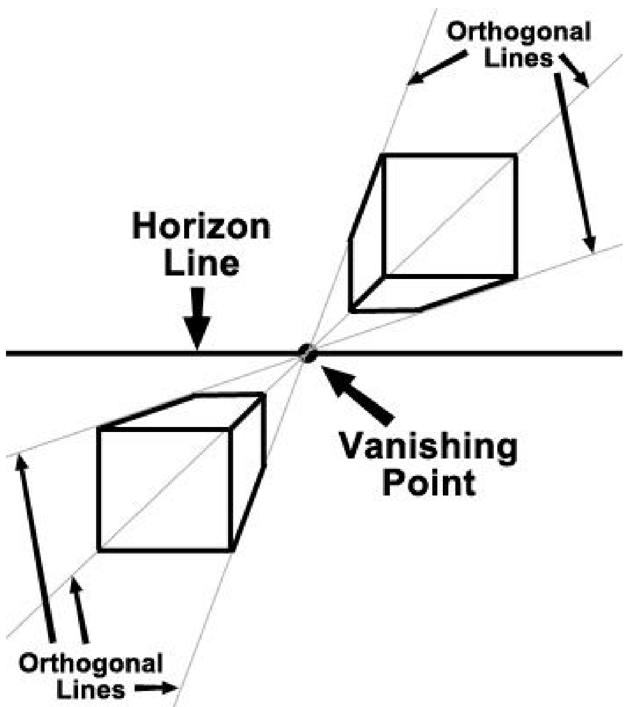

Figure 1. Source: [Wikipedia, One Point Perspective Example]

What this tells us is that any line ending at the vanishing point represents the path a point will take on the paper as it moves away form us. Artists using the technique apply this idea by drawing these perspective lines as guides when giving depth to objects; however all we need to know is that to make a point look further away, we move it in a straight line towards the vanishing point. If we imagine objects as a collection of points, and apply this principle to every point in an object, we can accurately represent its depth in the resulting image.

With this knowledge, the scientists only needed to figure out a way to represent this process mathematically. At first, this may seem like a complicated task; however, a line intersecting the center is very easy to represent mathematically. In fact, to represent such a line mathematically you only need a single point on it, and multiplying or dividing the coordinates of a point on that line will cause it to move along the line. In this case, we divide because we want points to move towards the center as they move further away. The number we will divide by is the distance from the point of view. Represented mathematically, this number is the distance along the Z-axis, where X and Y are the horizontal and vertical axis. Putting all of this information together scientists arrived at two equations that can be used to project any point of 3 dimensions (x,y,z) to an image of 2 dimensions (x',y'): $x' = x/y$ and $y' = y/x$. This algorithm is known as the perspective divide, and it is responsible for most three-dimensional graphics today.

Those familiar with single-point perspective may point out that it can only be used for a head-on view, and this is somewhat true. Artists typically use other methods such as two or three-point perspective for angles looking at the corner of objects; however, the one-point perspective model can work for a rotated view. Imagine looking at a cube at a forty-five-degree angle, instead of viewing this situation as the viewpoint being rotated, we can view it as looking head-on at a rotated cube. If we reframe how we think about the point of view, we can make one-point projection work in any situation; when you look to your left in a video game, mathematically, the entire world is being rotated the opposite way around you. The same concept applies to moving the point of view as well; when you move forward, the entire world moves backward.

Occam's razor states that of two otherwise equal solutions, the simplest one is likely the best, and computer science is no exception. While there have been many solutions to this problem, the perspective divide is by far the most widely used because its simplicity allows it to be flexible and efficient. To me, there is nothing more elegant than an extraordinarily simple solution to a complex problem, and this simple elegancy happens so often when we describe natural phenomena mathematically and is part of the beauty of mathematics. The perspective divide perfectly encapsulates this idea; at first, the problem seems dauntingly complex, but, in the end, the solution is one of the most basic mathematical operations.

1. "Rendering." The Science Behind Pixar, sciencebehindpixar.org/pipeline/rendering.