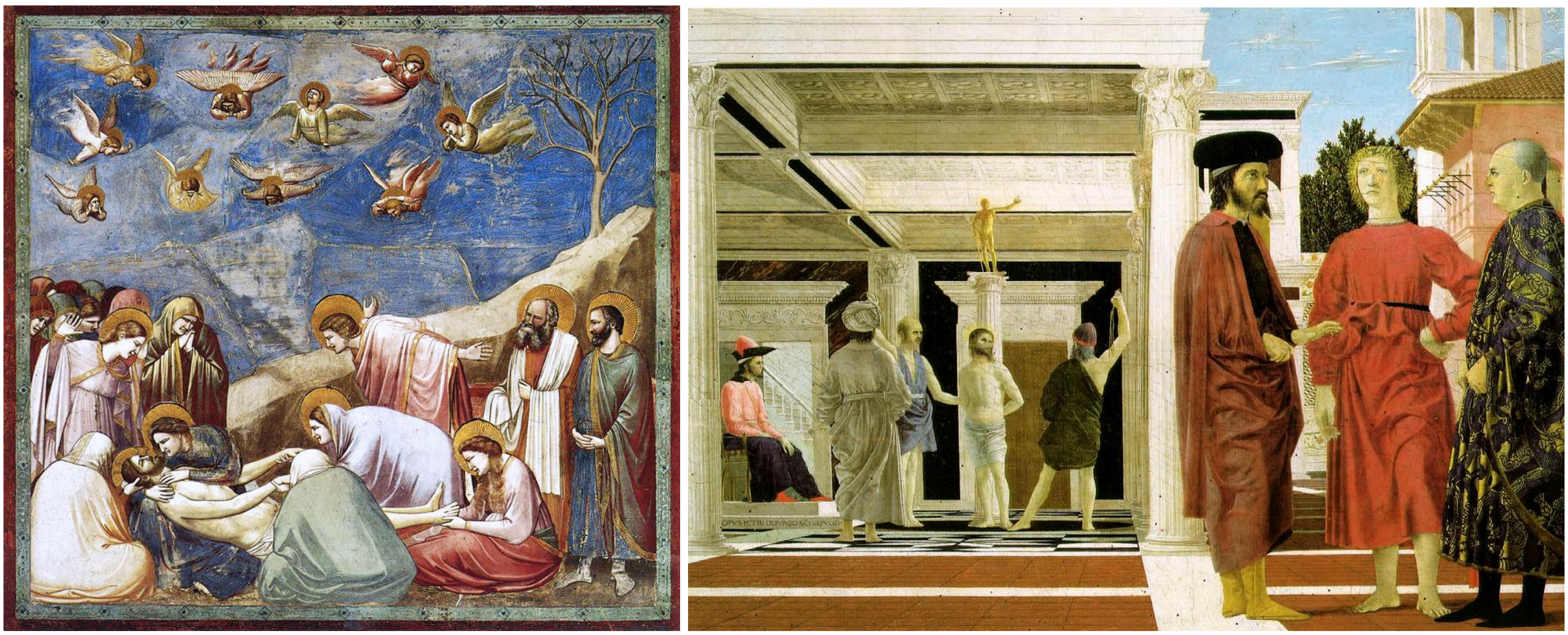
Matrices in Linear Perspective

Atharv Chowdhary

Introduction

For thousands of years, human art was horrendously bad at accurately portraying the world around us. This is not to say art was necessarily bad, but rather that it seemed beautiful in defiance of the natural order. All of this changed, however, during the Renaissance. In the 14th and 15th centuries, artists (especially those in Italy, the heart of the Renaissance), began working to make their art as realistic as possible. Renaissance artists desired not only to create images that were appealing to the eye, but images that actively deceived the eye into thinking it was looking at something real.



Left—Lamentation of Christ (Giotto); Pre-Renaissance
Right—The Flagellation of Christ (Piero Della Francesca); Renaissance

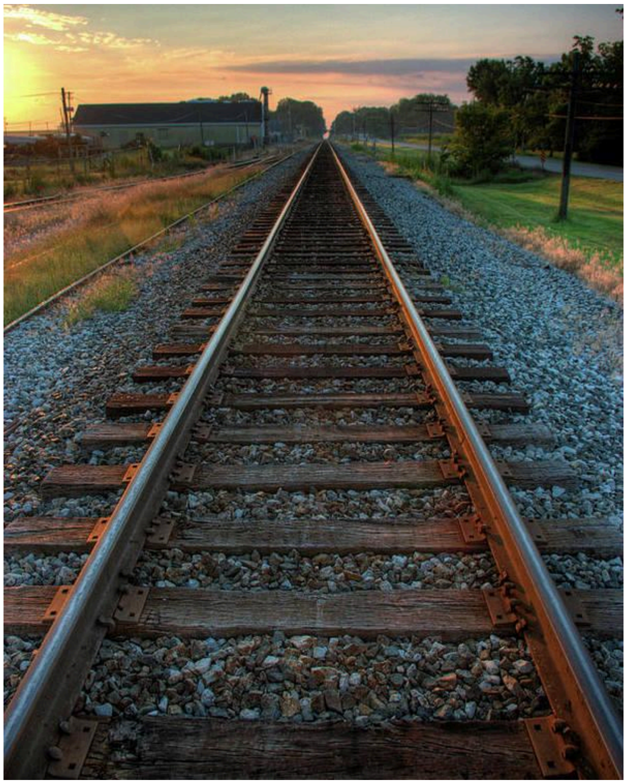
This was accomplished by using one key technique: linear perspective. The fundamentals of this procedure involve drawing orthogonal lines to a vanishing point (often on one or more of a series of parallel lines). The painting is then modelled around the orthogonal lines, which gives it the impression of having depth.

It may seem as though linear perspective is grounded in math, and it is. While the technique was applied artistically, the great artists of the Renaissance—the very artists that pioneered the technique of linear perspective—were also great mathematicians. Fillippo Brunelleschi, Leonardo da Vinci, and several other artists wrote mathematical treatises on why linear perspective works and how to apply it. In other words, linear perspective has a clear mathematical basis, and that is what this paper seeks to explore.

Below, linear perspective will be examined and explained mathematically. The paper will then explore applications of linear perspective in 3D and 2D art.

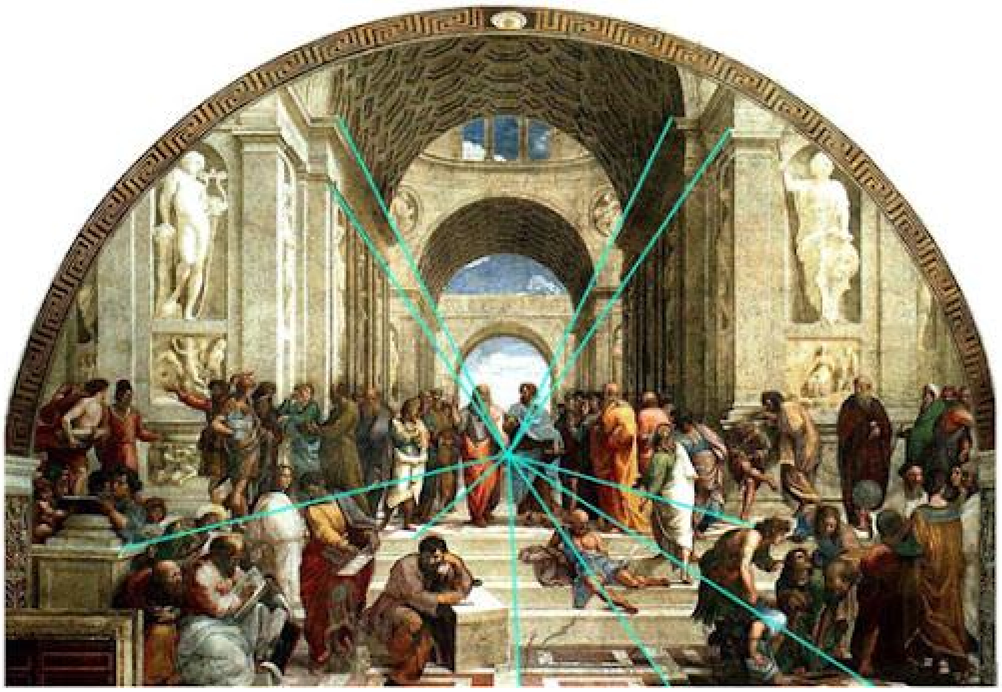
Basic Methodology

Linear perspective is grounded in a few basic principles, the most central of which is known as the "vanishing point". Formally, this is the point at which parallel lines in an image converge. Though the definition sounds like something impossible to achieve, it is necessary to consider how our eyes perceive three dimensional space. Consider the image below:



The railroad tracks are parallel—they simply never converge. However, when standing at a fixed point in three dimensional space, our eyes (and objects like cameras) perceive parallel lines as converging at some infinitely distant point in the horizon. This is the vanishing point of an image.

The presence of a vanishing point is often crucial to making something look realistic, as our eyes are naturally primed to perceive them. In the photo above, the vanishing point is extremely clear because of the presence of the railroad tracks, which effectively act as lines that guide the eyes toward the vanishing point. Renaissance artists used a similar trick to make it easier for them to locate and draw around vanishing points. In the early stages of painting, they traced the 'orthogonals' (the lines present in the painting/drawing that would have been parallel in three dimensional space) until they met at a singular vanishing point. Artists could then use these orthogonals and the vanishing point to model the rest of their painting in greater detail.

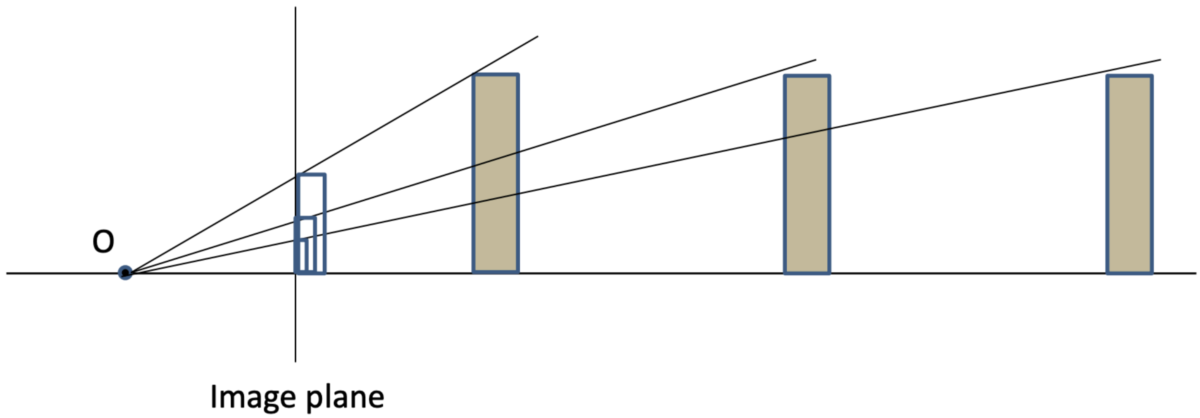


"The School of Athens" - Raphael
The orthogonal lines are highlighted here, converging at a singular vanishing point. As seen, this gives the painting greater depth.

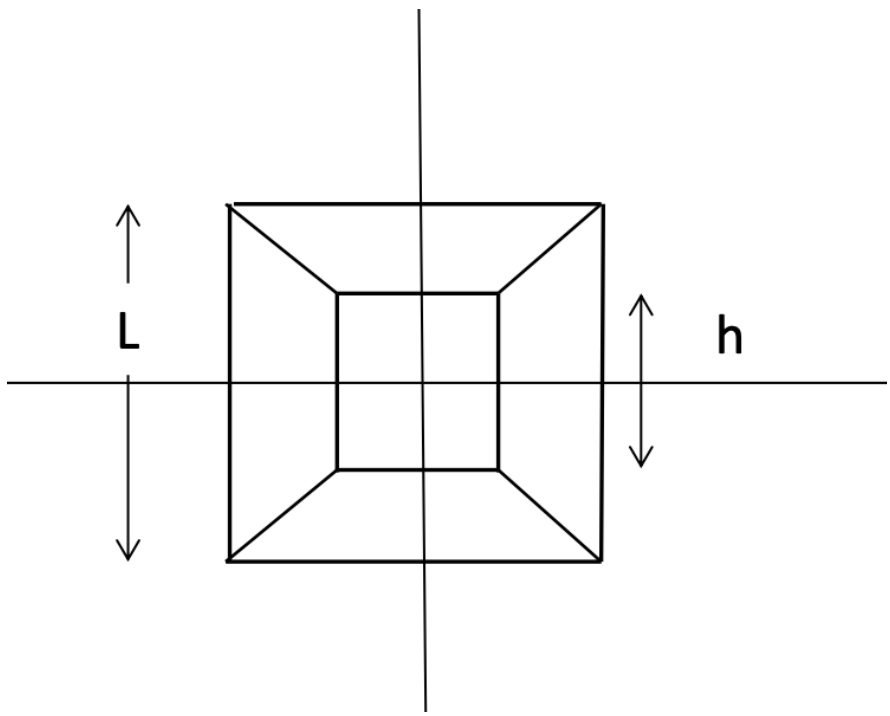
To explain the techniques behind linear perspective, we can start from one of its primary users and mathematical formulators: Albrecht Durer.

"Draughtsman Making a Perspective Drawing of a Reclining Woman" - Albrecht Durer

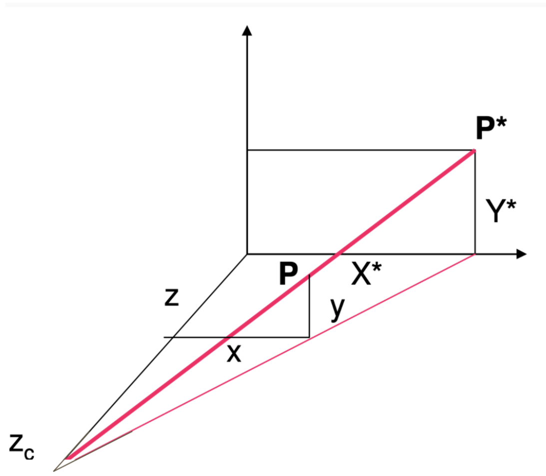
In one of his many sketches, Durer depicted himself painting his subject after looking at them through a window. This seems a little confusing, but becomes more clear if we take a glance at what exactly is happening here geometrically:



(Source: SFU)



The O is the point at which the observer's eye is located, and the image plane is the plane at which they perceive the objects (this would be the little window Durer looks through in his sketch). Each object is projected onto the image plane with a different size, with the smallest being the object farthest away. Orthogonals are thus used to trace connections from the largest to smallest objects in a manner that conveys depth. The next step is thus to find a way to relate coordinates in 3D space to those in 2D space that employ linear perspective to retain the illusion of depth. Consider the following diagram:



(Source IITD)

In this case, the observer is located at z_c (a distance of z_c away from the ray Y). The observer is looking at the point **P***, which is projected onto the image plane as **P**. All other lengths are labelled as per the diagram. Using similar triangles, we can determine

$$\frac{x*}{z_C} = \frac{x}{z_C - z} \Rightarrow x* = \frac{x}{1 - \frac{z}{z_c}}$$

Via a similar process (this time also involving the Pythagorean Theorem), we find an identical equation for y*:

$$y* = \frac{y}{1 - \frac{z}{z_c}}$$

If we take $r = -\frac{1}{Zc}$ for some arbitrary r, then we can express the two equations as

$$x* = \frac{x}{zr + 1} \qquad y* = \frac{y}{zr + 1}$$

The above form becomes useful later. Notice that this gives us linear perspective in <u>one dimension only</u>. In other words, the linear perspective attained by this method of coordinate searching will only give us only one vanishing point; as such, we will have only one set of converging orthogonals. In this case, we will have linear perspective with respect to the z-dimension, as we took similar triangles along the z-axis. However, the same calculations work in all dimensions—the form of the equations obtained will be identical to those above.

Now we can finally transfer this knowledge to linear algebra. In order to efficiently convert x, y, and z coordinates into the desired form, we can use matrices. In particular, we are using matrix transformations here.

Note, however, that there are no ordinary matrix transformations on a 3xn matrix of x, y, and z coordinates that would yield the desired form. What we can do, however, is turn to 4xn matrices for help. We **'homogenize'** the 3xn matrix by adding a row of 1s at the bottom. Then, we can perform the following calculation:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & r & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

One quick matrix multiplication later, we find the solution to the above expression:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & r & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \\ rz + 1 \end{bmatrix}$$

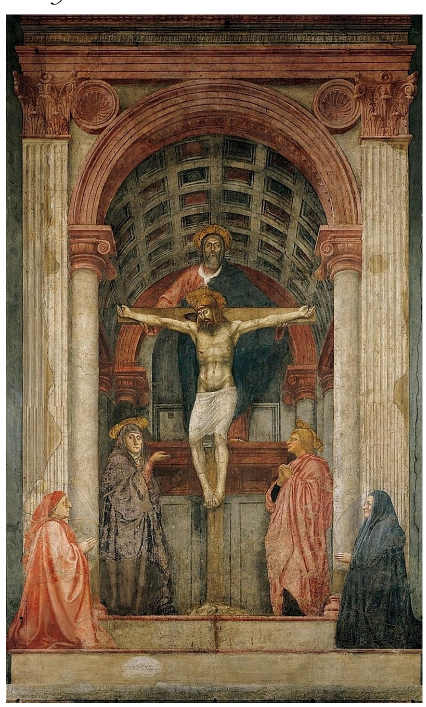
The rz+1 term that appears at the bottom of the solution vector is the exact term that appears in the denominator of our other x* and y* values. To obtain our desired coordinates, we can then divide the entire vector by the scalar rz+1. In other words, we multiply the entire vector by the scalar $\frac{1}{rz+1}$. This yields:

$$\begin{bmatrix} \frac{x}{rz+1} \\ \frac{y}{rz+1} \\ \frac{z}{rz+1} \\ 1 \end{bmatrix}$$

The top two terms are looking the way we want now. There are now two ways to proceed: we either shift back into two dimensional space, or stay in three dimensional space. If we choose to stay in three dimensional space, then we simply take the top three values and treat them as our coordinates for x, y, and z respectively. There will unfortunately be distortion in the z-direction, which we did not want, but, from the right angle, this is negated and linear perspective in only the x and y directions will remain. If we choose to map the coordinates back into two dimensional space, then we simply take the two top values and treat them as our x and y coordinates respectively. This creates a perfect linear perspective in two dimensional space.

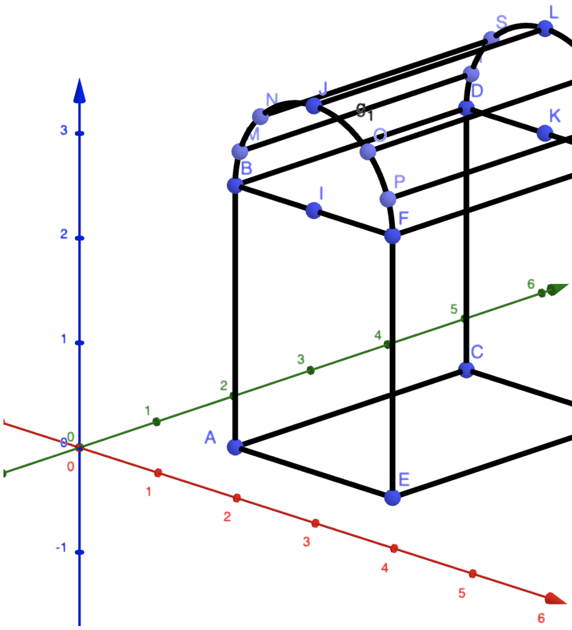
Extension #1: Replicating Art

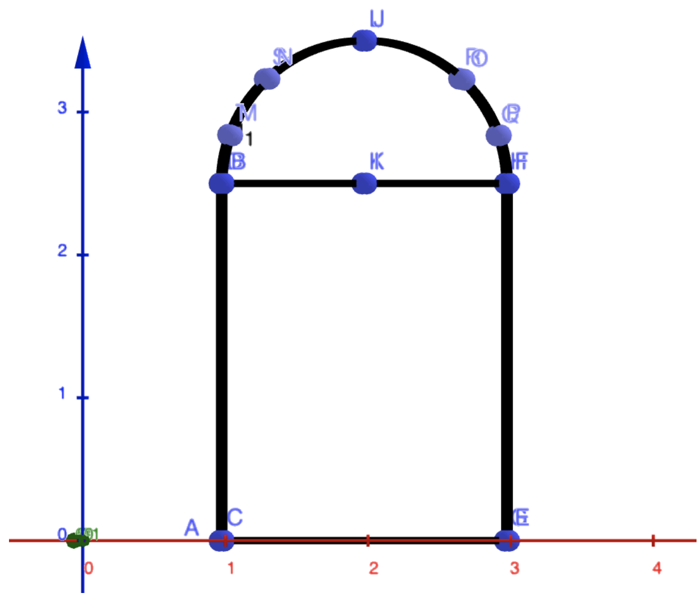
For my first extension, I really wanted to create my own examples of linear perspective using matrices. I chose to be inspired by one of the most famous pieces of all time to use linear perspective: Masaccio's *The Holy Trinity*:



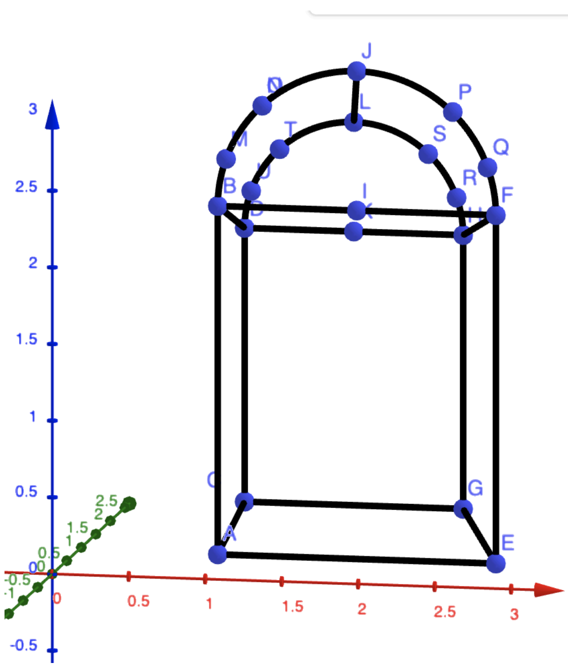
The Holy Trinity - Masaccio

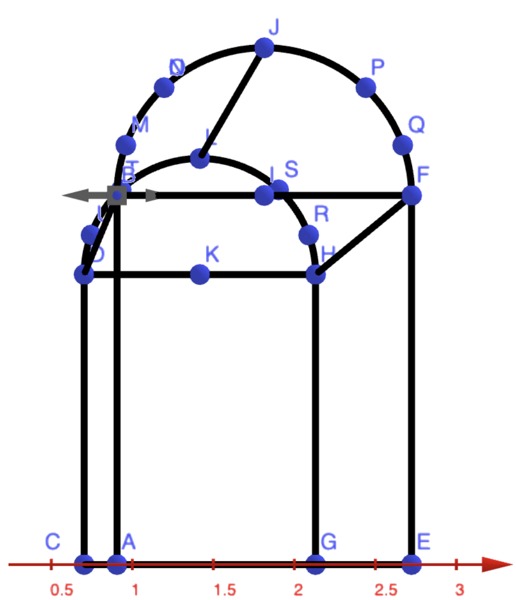
I decided to try and replicate the center hall of the structure, featuring a rectangular base and semi-circular roof. The 3D model appeared as follows:



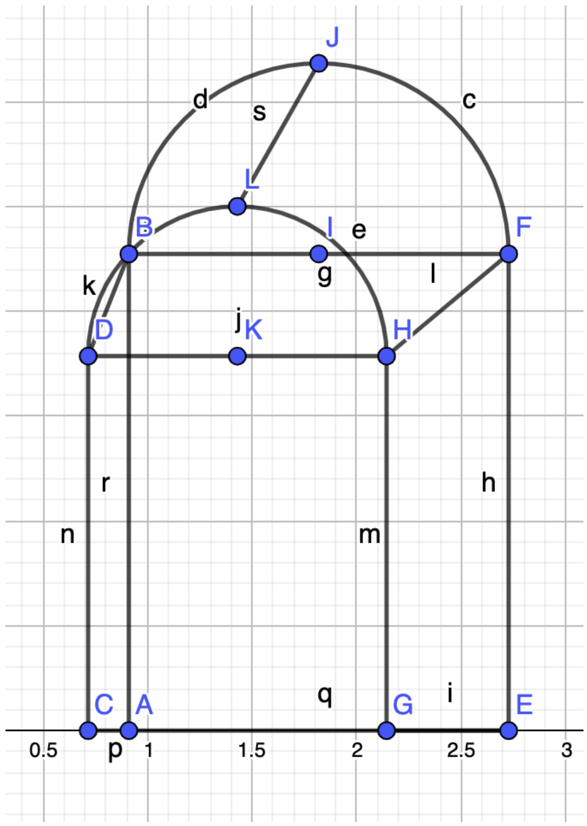


As is evident from a front view, this model lacks linear perspective—it simply appears flat and two dimensional if looking at it directly from the front. However, by taking all the coordinates, plugging them into a matrix and setting r=0.1, we get the following models:





Notice the stark contrast between the two sets of images. While they are both three dimensional models of an object, the former images simply cannot convey depth along the z-dimension nearly as well as the latter set. What is even more interesting is that the depth along the z-dimension is preserved on a two dimensional plane, as mentioned previously. Even when there *is* no z-dimension, linear perspective retains the illusion of there being one:



The above image is entirely two dimensional. The z-coordinate was eliminated and the figure was plotted on the x-y plane using the same x and y coordinates as it had in the previous set of images. This image demonstrates how linear perspective preserves the illusion of depth in even two dimensions.

Extension #2: Cubes using Linear Perspective in 3 Dimensions

So far, we have only employed linear perspective in one dimension—namely, the z-dimension. In other words, we are using linear perspective to convey depth, as the orthogonals only converge when extended in the z-dimension. For this extension, I wanted to explore applying linear perspective in more than just one dimension. What if we applied linear perspective in all three dimensions?

Recall the following matrix:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & r & 1 \end{bmatrix}$$

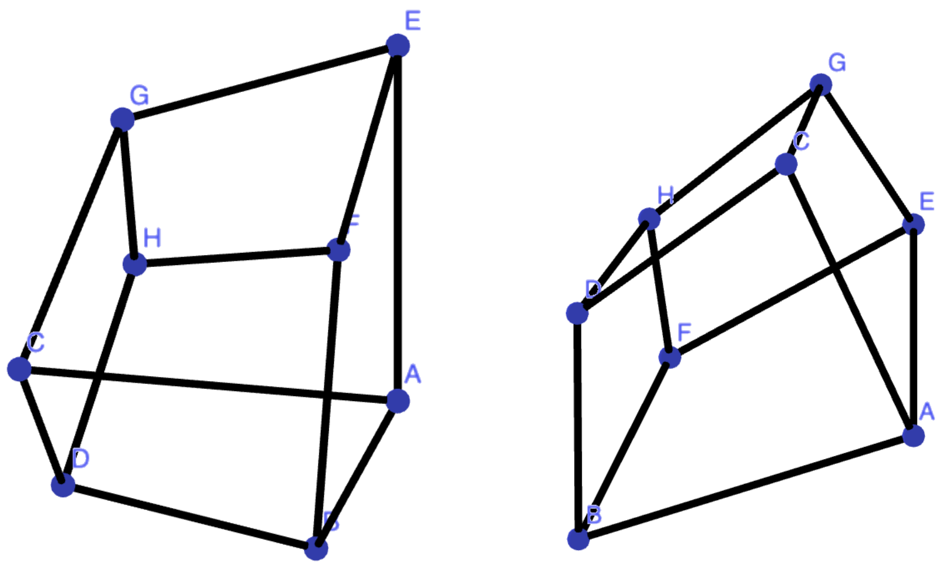
The r-value lets us impose a transformation that applies linear perspective in the z-dimension. If we shifted r to the second column instead of the third, then we would be applying linear perspective in the y-dimension (we would be dividing all coordinates by ry+1). Similarly, if we shifted r to the first column, then we would be applying linear perspective in the x-dimension. Therefore:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ a & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & b & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & c & 1 \end{bmatrix}$$

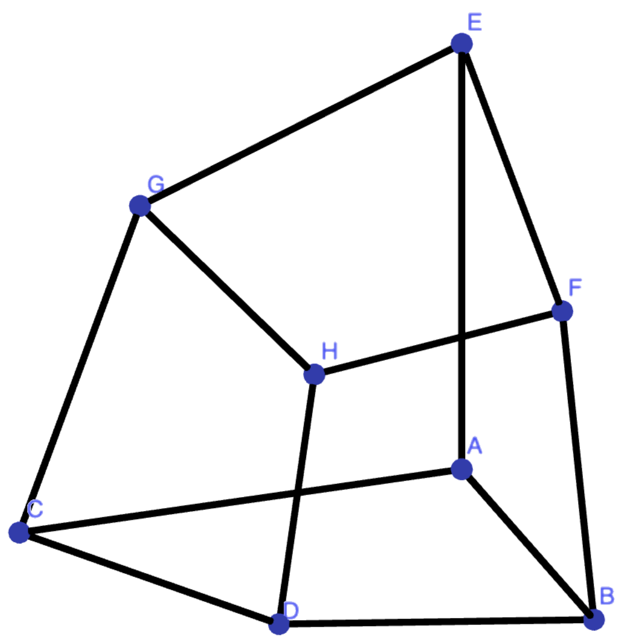
The leftmost matrix applies the linear perspective transformation in the x-dimension, the middle matrix applies the linear perspective transformation in the y-dimension, and the rightmost matrix applies the linear perspective transformation in the z-dimension. In order to find the transformation matrix that does all three of these things, we multiply the above matrices to find:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ a & b & c & 1 \end{bmatrix}$$

The desired transformation is therefore rather easy. We simply substitute values for a, b, and c that correspond to our desired viewing point (recall that r was the negative reciprocal of z_c; the a, b, and c values are similarly tied to x_c, y_c, and z_c). I used this transformation on a simple 6x6x6 cube, and the results were fairly surprising.



The above image certainly obeys linear perspective in all three dimensions (one can see the orthogonals meeting if extended in any direction). However, it looks nothing like a cube—that is, until we view the object from the viewing point. Doing show reveals the following image:



The object now seems to resemble a cube rather closely, and seems to be a rather accurate three dimensional cube at that. This was incredibly fascinating to watch in real time as I rotated the object—it seemed almost unreal how the viewing perspective changed the entire look of the object!

Critical Friends Feedback

As always, getting feedback from my critical friends was a critical step in helping me improve the quality of the product as a whole. This time around, a lot of the feedback I got was fairly positive—many people seemed to really enjoy the visually engaging/interesting explanations and extensions. However, there was valuable feedback to be gained from the entire experience as well:

Yussef suggested delving more into the modern applications of the math that I described. I thought this was an excellent point and something I have since incorporated into my conclusion, especially since the math I described had many fascinating uses in modern technology.

Matthew also recommended trying to make my project as a whole a little clearer, which is something else I took into consideration. I can see how it might have gotten confusing—especially when being explained verbally. I have touched up the paper where possible to make it as clear cut as possible.

Conclusion

Art history is a subject that has been deeply fascinating to me for a while now. Renaissance art is one of my favorite areas of study in the field, and, upon learning about linear perspective for the first time, I knew there had to be fascinating mathematics behind it. Turns out, I was right. The mathematics behind linear perspective is incredibly fascinating, and genuinely ahead of its time. Linear perspective is a crucial part of many systems today, and it has recently become popular in the game development industry as part of projects on ray tracing and linear algebra. Game developers and even film animators are simply using more advanced versions of this math to make their 3D environments map to a 2D screen in a way that feels realistic. There is no doubt the applications of linear algebra in conjunction with linear perspective is crucial to art today, just as linear perspective shaped the art of decades prior.

It was a genuinely very enriching, enjoyable, and insightful experience being able to apply my knowledge of linear algebra to art history. If given the opportunity to keep working, I would like to take this theme even further. There are so many opportunities for exploration—for example, could I take a 2D work of art that doesn't employ linear perspective (e.g. The Birth of Venus as a notable example) and apply linear perspective to it? Or potentially I could stick to the theme of linear algebra in art history without necessarily applying linear perspective; maybe I could explore the role of color versus form in Venetian vs Florentian art, or correct the distortion on surrealist paintings. There is a much deeper connection between math and art than I realised before doing this project, and I would love to keep exploring these ideas at a later date.

Works Cited

http://new.math.uiuc.edu/public403/perspective/alberti/alberti.html
https://web.iitd.ac.in/~hegde/cad/lecture/L9_persproj.pdf
https://www.sfu.ca/~rpyke/perspective.pdf
https://www.math.utah.edu/~treiberg/Perspect/Perspect.htm#top