Computer Vision
Homework 03: Image Formation
CS 670, Fall 2019

Name: Kunjal Panchal
Student ID: 32126469
Email: kpanchal@umass.edu

Sept 29, 2019

Contents

1 Problem 1: Equation of a Line in 3D passing through a Point and in a Direction

1.1 General Equation about 3D lines

The vector equation of a line in three-dimensional geometry is given in following way:

Suppose, there are two points given for that line -
point 1 $P_1 = (x_1, y_1, z_1)$ and point 2 $P_2 = (x_2, y_2, z_2)$.

For a $\lambda \in [0, 1]$, the line segment equation joining P_1 and P_2 is written by [1]

$$r = (1 - \lambda)P_1 + (\lambda)P_2 \tag{1}$$

See Figure 1

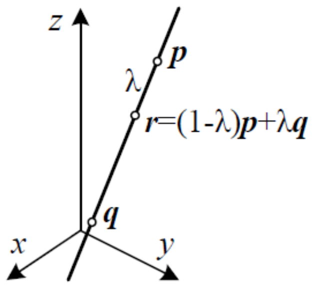

Figure 1: 3D Line equation, $r = (1 - \lambda)p + \lambda q$

Here, r represents any point on the line made by points P_1 and P_2.
A special case is where the point P_2 would lie at infinity, i.e.,

$$P_2 = (\hat{d}_x, \hat{d}_y, \hat{d}_z) = \hat{d}$$

where, d̂ is the direction of the line which passes through point P_1.

We can **rewrite** the Eqaution 1 as,

$$\boxed{r = P_1 + \lambda \hat{d}} \tag{2}$$

1.2 Representation 1 for the line with a Point and a Direction

If we write Equation 6 in separate x, y and z components, we get;

$$(r_x, r_y, r_z) = (P_x, P_y, P_z) + \lambda(\hat{d}_x, \hat{d}_y, \hat{d}_z) \tag{3}$$

$$r_x = P_x + \lambda\hat{d}_x; \qquad r_y = P_y + \lambda\hat{d}_y; \qquad r_z = P_z + \lambda\hat{d}_z \tag{4}$$

$$\frac{r_x - P_x}{\hat{d}_x} = \frac{r_y - P_y}{\hat{d}_y} = \frac{r_z - P_z}{\hat{d}_z} = \lambda \tag{5}$$

In the problem definition, point $P_1 = \mathbf{p} = (\mathbf{x}, \mathbf{y}, \mathbf{z})$, direction $\hat{\mathbf{d}} = (\mathbf{a}, \mathbf{b}, \mathbf{c})$ and $\mathbf{r} = (\mathbf{r_x}, \mathbf{r_y}, \mathbf{r_z})$ is any other point on the line made by point P_1 and direction vector $\hat{d}$

Referring to Equation 5,
The final equation will be:

$$\boxed{\frac{r_x - x}{a} = \frac{r_y - y}{b} = \frac{r_z - z}{c}} \tag{6}$$

1.3 Representation 2 for the line with a Point and a Direction

We just use Equation 6,

$$\boxed{(r_x, r_y, r_z) = (x, y, z) + \lambda(a, b, c)} \tag{7}$$

1.4 Closing Remarks

The direction $\hat{d}$ is just a vector value we get from subtracting one point co-ordinates which is on the line, from another point co-ordinates which is on the same line.

If we make $\lambda = 1$ in Equation 6, it will be $r = P_1 + \hat{d}$

Thus, $\hat{d} = r - P_1$

So, given two points of the line we may write down a parametric representation:

$$P_2 = P_1 + \lambda(r - P_1) \tag{8}$$

That's how we have a simple geometric interpretation.

2 Problem 2: Vanishing Point of Parallel Lines

2.1 The Geometry of Perspective Projection

- PERSPECTIVE PROJECTION Under this model, the depth (Z- value/ co-ordinate) of a 3D object is devided from the 3D point co-ordinates (x, y, z) to get 2D representation point co-ordinates $(f\frac{x}{z}, f\frac{y}{z})$ where f is the distance from optimal center to image plane
- POSITION of OPTIMAL CENTER At (0, 0, 0) of the co-ordinate system
- POSITION of IMAGE PLANE Parallel to $x-y$ plane of co-ordinate system
- 3D to 2D $(x, y, z) \to (f\frac{x}{z}, f\frac{y}{z})$

Hence, perspective projection is not an affine transformation; it **does not map parallel lines to parallel lines**, for instance. Unlike the orthographic and parallel projections, the **projection vectors are not uniform** for all points and vectors; rather, there is a projection point or perspective point, and the line of projection is defined by the vector between each point and the perspective point.

2.2 Perspective Projection of Two Parallel Lines

As shown in Figure 2, consider one line which is getting projected according to the co-ordinates shown in the previous subsection.

Let's say the first line is passing through a point $\mathbf{p_1}$ in r direction.

That will project points on image plane according to the optimal center distance from the image plane.

Now, if there's another line with same direction r, but passing through a point p_2 which is not on the first line, will have it's 2D representation line intersecting with the 2D representation of the first line.

There will be a point where all the lines in same direction converges/ intersects to a point which will make a parallel line to the others, if connected with the optical center. **That's called Vanishing Point**.

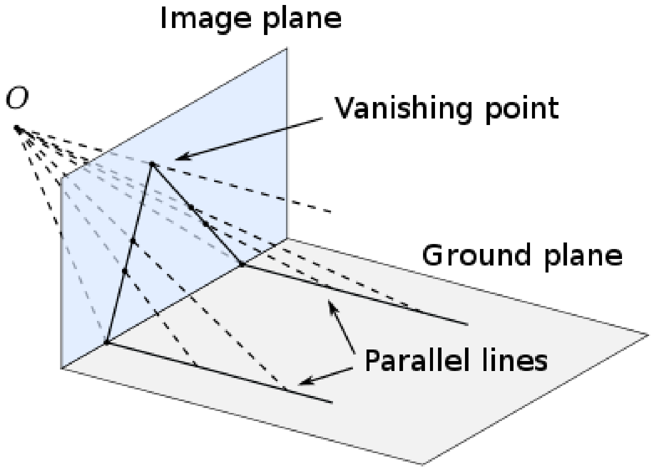

Figure 2: A 2D construction of perspective viewing, showing the formation of a vanishing point [2]

E.g., horizontal and vertical building edges, zebra crossings, railway tracks, the edges of furniture such as tables and dressers.

It is necessary for the directions of the lines to be the same (lines being parallel), in order for them to converge to a single point. If the **directions are different**/ if the lines are not parallel, the line passing through optimal center, parallel to those lines, will intersect with those lines 2D projections at different points, thus, **not converging**.

3 Problem 3: Vanishing Line of a Plane

If we think about a plane as a group all the possible lines passing through points on that plane; we can get group of *vanishing points* from these groups of lines, all parallel lines being in the same group.

The set of vanishing points we get will be on a same plane as well i.e., we can imagine a plane consisting of all the vanishing points formed by groups of parallel lines of a ground plane.

That plane also include the camera optical center.

The important thing to remember here is that all points be on the same plane, to get a vanishing line which includes all the vanishing points derived from all the lines passing through them.

This means, all points should have the same surface normal, i.e., the perpendicular direction vector ($\hat{n}$) of a line passing through the point ($p_1 \in P$ [Set of points of the plane]) concerned any other point of the plane ($p_i \in P$) should be the same as the perpendicular vector of any line which is actually on the plane. See Figure 3

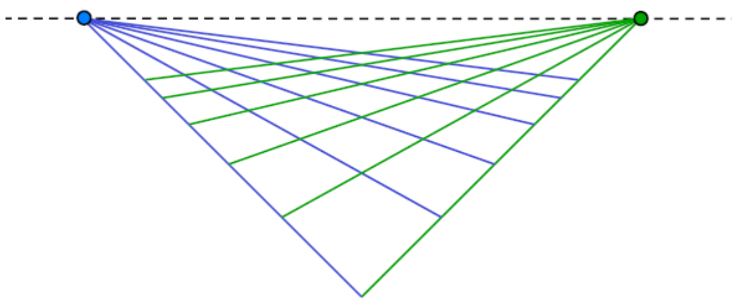

Figure 3: A 2D construction of perspective viewing, showing the formation of a vanishing line [3]

3.1 The Plane

At every point $p(x, y, z) \in P$, the z co-ordinates will be the same.

From the Problem 1, we can say the direction of a point/ a line joining two points is $\hat{d} = (a, b, c)$;

Hence, for each point $p_i = (x_i, y_i, z_i)$, the c of the direction will be the same.

Normal vector is perpendicular to the plane;
therefore, $\hat{n} = (0, 0, \hat{n_z}) = \hat{k} \cdot n_z$

$$n \cdot p = \hat{k}n_z \cdot (\hat{i}, \hat{j}b, \hat{k}c) = n_z \cdot c$$

and since normal vector magnitude is 1 in every co-ordinate; $n_z = 1$

$$\boxed{n \cdot p = 1 \cdot c = c} \tag{9}$$

For each point p of the plane.

3.2 The Vanishing Line

- The horizon line is a theoretical line that represents the eye level of the observer
- If the object is below the horizon line, its vanishing lines angle up to the horizon line
- If the object is above, they slope down
- All vanishing lines end at the horizon line

4 Problem 4: Why Cameras Need Lenses

4.1 Pinhole Camera Model

- In the beginning, to capture a 2D representation of the 3D world; people used a very tiny aperture to let only very choice rays pass through a barrier/ aperture.

- See Figure 4. The key idea behind this tiny aperture or a "pinhole" was to do away with the problem encountered of blurriness when we just put a film in front of an object.

- If there was just a film in front of an object, rays from all directions will be captured by the film, resulting in a point at object getting put on all the points of the film, where the rays from that 3D point are getting reflected on all points of the film. That results in totally illegible, blurry 2D representation.

- With the introduction of a pinhole model. We created a very tiny aperture which made sure than only one ray from a 3D point is allowed to pass through the barrier, and thus, only one point was represented on the film as in Figure 4.

- That made the image clearer. It just captures/ lets through pencil of rays.

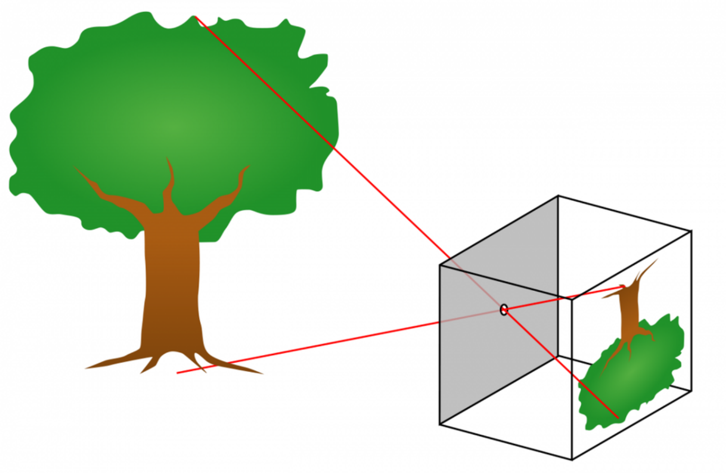

Figure 4: Schematic of a pinhole camera [4]

4.2 Why Pinhole Camera-like Aperture is not Sufficient

One might think that making the aperture smaller would give us shaper/ more accurate results. But there are limitations to this model:

- When we are allowing only one ray through for each point, naturally, the intensity will be much lower, resulting in darker images.
- As shown in Figure 5, it only works on certain aperture size; above that, we will get blurry images as lot many rays are passing through it for a single point, resulting in the point getting represents in 2D image plane by more than one point. It is called **Circle of Confusion**. If the aperture is too small, very little light gets through it, not only resulting in dark images but also in diffraction.

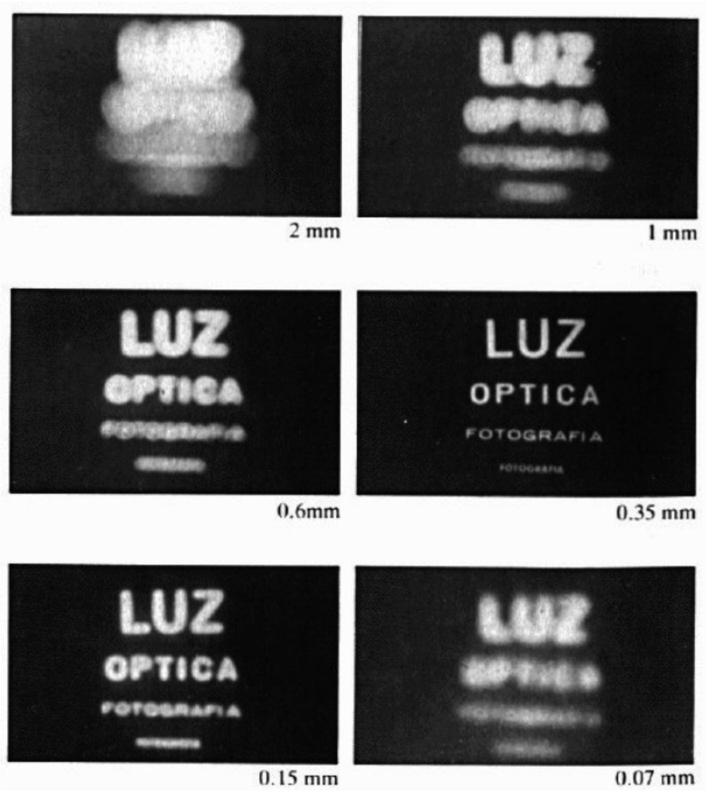

Figure 5: Result of shrinking the pinhole camera aperture[3]

- We need long exposure to get enough light to pass through camera.

4.3 Cameras with Lenses

A **LENS** is an optical device which focuses [convex See Figure 6] or disperses [concave] light rays through refraction.

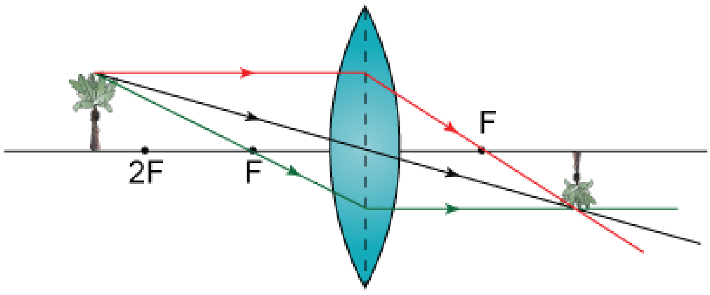

Figure 6: A convex lens focusing all the light rays coinciding on it from an infinite distance

With a lens:

- Rays passing through center won't get deviated [acts like a pinhole model].
- Other rays will get bent to get focused on a single point.
- Which results in a point in 3D be represented by only one point in 2D, but with greater intensity and sharpness as we are focusing multiple rays getting reflected from that point.
- THUS, camera use lenses to direct and focus the light rays to get an accurate representation of the 3D counterparts of the object.
- Also; the film, sensor, or back wall of our eyeball is usually much smaller than the view we are trying to capture. Therefore, we need to bend the light to reduce the size of the image.
- A camera lens also **slows down the light speed**. The speed of light changes when it passes through translucent materials. So, light is bending and slowing as it enters and exits a lens (depending on the design of the lens). Giving it more exposure.

4.4 Conclusion and Bottom Line

Without a lens, light from all directions reaches every point on the film or sensor, it isn't a picture just a gray blur.

In order to get a high resolution/ detailed image in either case, and to get it fast enough, you need a device **LENS** that will focus and amplify the effectiveness of light gathering to one small area - call that a retina, a piece of photosensitive film or a digital photo-diode.

References

[1] Richard Szeliski, *Computer Vision: Algorithms and Applications* ISBN 978-1-84882-935-0

[2] *Vanishing Point* https://en.wikipedia.org/wiki/Vanishing_point

[3] Subhransu Maji, *Image Formation* https://www.dropbox.com/s/80uqet17ajm8zz0/lec05\%2B06_image_formation.pdf?dl=0

[4] Larry D. Kirkpatrick, Gregory E. Francis, *"Light". Physics: A World View (6 ed.) p 339* ISBN 978-0-495-01088-3