Questions to answer:

What is an intrinsic camera?

- Intrinsic camera parameters are a set of values that describe internal characteristics of a camera. These parameters are important as they provide information about how a camera transforms a 3D scene into a 2D scene. These parameters is crucial for various tasks such as camera calibrations, 3D reconstruction and image rectification. Some key components of the camera is:
    - **Focal length:** Expressed in pixels, defines the distance between the camera sensor and the center of the lens. Determines the field of view of the camera and affects the scale of the projected image
    - **Principal point:** The point where the optical axis interacts the image sensor. It is located at the center of the image but due to manufacturing imperfections, it might always be the case.
    - **Lens distortion:** Real-world lenses often introduce distortions into the captured image, typically radial and tangential distortion. Radial distortion causes straight lines to appear curved while tangential distortion occurs when the lens and the image sensor are not perfectly parallel.
    - **Skew Coefficient:** This parameter accounts for the angle between the x and y pixel axes. In an ideal camera, these axes are perpendicular and the skew is zero. However, in real cameras, there might be a slight skewness
    - **Pixel expect ratio:** This is the ratio of the physical unit size in the x direction to the size in the y direction. Most camera has square pixels so this ratio is typically close to one.

What is an extrinsic camera?

- Extrinsic camera parameter refer to the parameter that define the position and orientation of a camera in the world coordinate system. Unlike intrinsic parameters, which are specific to the camera’s internal properties, extrinsic parameters relate the camera to it’s surroundings.. Here are key aspects of extrinsic camera parameters:
    - **Position(translation vector):** This specifies the camera’s position in the world coordinates. Typically represented as a vector that shows how far and in what direction the camera is from the origin of the world coordinate system

- **Orientation (Rotation matrix)** Defines the direction in which the camera is pointing. It's often represented as a rotation matrix. This matrix transforms coordinates from the world coordinate system to the camera coordinate system.

Understanding extrinsic parameters of a camera is crucial for tasks like 3D reconstruction where you need to know how different views of an object or scene to each other in the real world. **Basically provides information about the camera's placement and orientation in a larger environment.**

How does a 3D point get projected to a pixel with a perspective projection?

- To project a 3D point to a pixel using perspective projection:
    - **Transform the 3D point:** Convert the point from the world coordinates to camera coordinates using the camer'as extrinsic parameters (position and orientation)
    - **Apply perspective projection:** Use the cameras intrinsic parameters to project the 3D point onto the camera's image plane. This results in a 2D point in a 3D point (x,y,z)
    - **Convert to 2D coordinates:** Divide x and y by Z (the depth) to get 2D coordinates maintaining perspective effects
    - **Adjust for camera intrinsics:** Transform these 2D coordinates into actual pixel coordinates in the image, considering the camera's intrinsic parameters like the principal point and pixel aspect ratio. Apply lens distortion correction if necessary.

What are homogeneous coordinates and what are they good for?

- Homogeneous coordinates are an extension of the usual Cartesian coordinates used in mathematics and computer graphics. Useful in projective geometry. Projections include vision which makes it non-linear to use Euclinean and you cannot intersect two lines that are parallel → you get zero.

  You get an extra definition to the conventional coordinates so instead of (x,y) you get (wx, wy, w) where w is a non-zero scaler. The 2D vector (x,y,1) represents the same point as (2x,2y,2). Homogeneous coordinates allow faster computation for translation, rotation, scaling and perspective projection. All can be performed using matrix multiplication. They can handle points at infinity. For example in 2D the point at infinity along the x-axis is represented as (1,0,0). Simplites perspective projection and facilitates complex transformation

How is a perspective projection expressed in homogenous coordinates

$$\begin{bmatrix} f_x & 0 & x_0 & 0 \\ 0 & f_y & y_0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1/f & 0 \end{bmatrix}$$

- The fx and fy are local lengths repressed in pixel units. X0 and y0 are the coordinates of the principal points and *f* is the distance from the camera center to the image plane. The matrix accounts for the camera's intrinsic parameters and performs the perspective transformation. After multiplication 3D point by this matrix, you get a 3D point in the homogenous coordinates which can be converted to 2D pixel coordinates by dividing by the third component (Depth)

What is a vanishing point and how do you find it

- A vanishing point is a point in an image where parallel lines in 3d space appear to converge. This makes far away objects appear smaller and giving a sense of depth

1. In an image, identify set of lines that are parallel
2. Extend the lines within the image. In a proper perspective projection, these lines will appear to converge
3. Intersection point: The point where these lines appear to meet is the vanishing point. For a set of parallel lines there's typically one vanishing point

What is the perspective projection?

What is a weak projection?

- Scaled orthographic projection, simplification of the perspective projection model used in computer vision and graphics. It
    - Instead of using varying depth for each point, weak projection assumes a constant or average depth for all points in the scene
    - Makes calculations simpler than full perspective projection as it avoids the complexity of dividing by the z - coordinates for each point
    - It’s used in scenariones where objects are at relatively uniform indstance from the camera or when the perspective effect is minimal and can be apprixmated.

What is an affine camera model?

- Simplified version of perspective model. Preserves parallelism, lines that are parallel in the 3D world remain parallel in 2D image. No convergence of lines at a vanish point and simpler maths as does not require division of depth.

  Useful for small field of view: Commonly used in tasks like image stitching or registrations where preserving parallelism amd simplifying the computation is more critical than accounting for perspective distortion.

When choosing a perspective model versus an Affine one

- Choose perspective model when accurate representation of depth and convergence is importance - Wide field of view. Use affine for simpler calculations when parallel lines mst remain parallel , typically for small field of view or flat scenes