Applications in EVs, Mars Rovers, even in drones

LiDAR uses a similar principle as sonar

This will be an overview

We all are familiar with the stereo imaging capability that our eyes give us. And in computational systems? Computers accomplish this task by finding correspondences between points that are seen by both the cameras form different angles

TRIANGULATION

Assume that we have two cameras whose image planes are exactly coplanar with each other, with exactly parallel optical axes that are a known distance apart, and with equal focal lengths $f_l = f_r$. Also, assume for now that the principal *points* c_x^{left} and c_x^{right} have been calibrated to have the same pixel coordinates in their respective left and right images.

So here the images are row-aligned and that every pixel row of one camera aligns exactly with the corresponding row in the other camera. We will also assume that we can find a point P in the physical world in the left and the right image views at p_l and p_r, which will have the respective horizontal coordinates x^l and x^r. (Not the centre of image)

In this simplified case, taking x^l and x^r to be the horizontal positions of the points in the left and right imager (respectively) allows us to show that the depth is inversely proportional to the disparity(difference) between these views, where the disparity is defined simply by $d = x^l - x^r$.

FORMULA

When disparity is near 0, small disparity differences make for large depth differences. When disparity is large, small disparity differences do not change the depth by much. The consequence is that stereo vision systems have high depth resolution only for objects relatively near the camera

PREPROCESSING

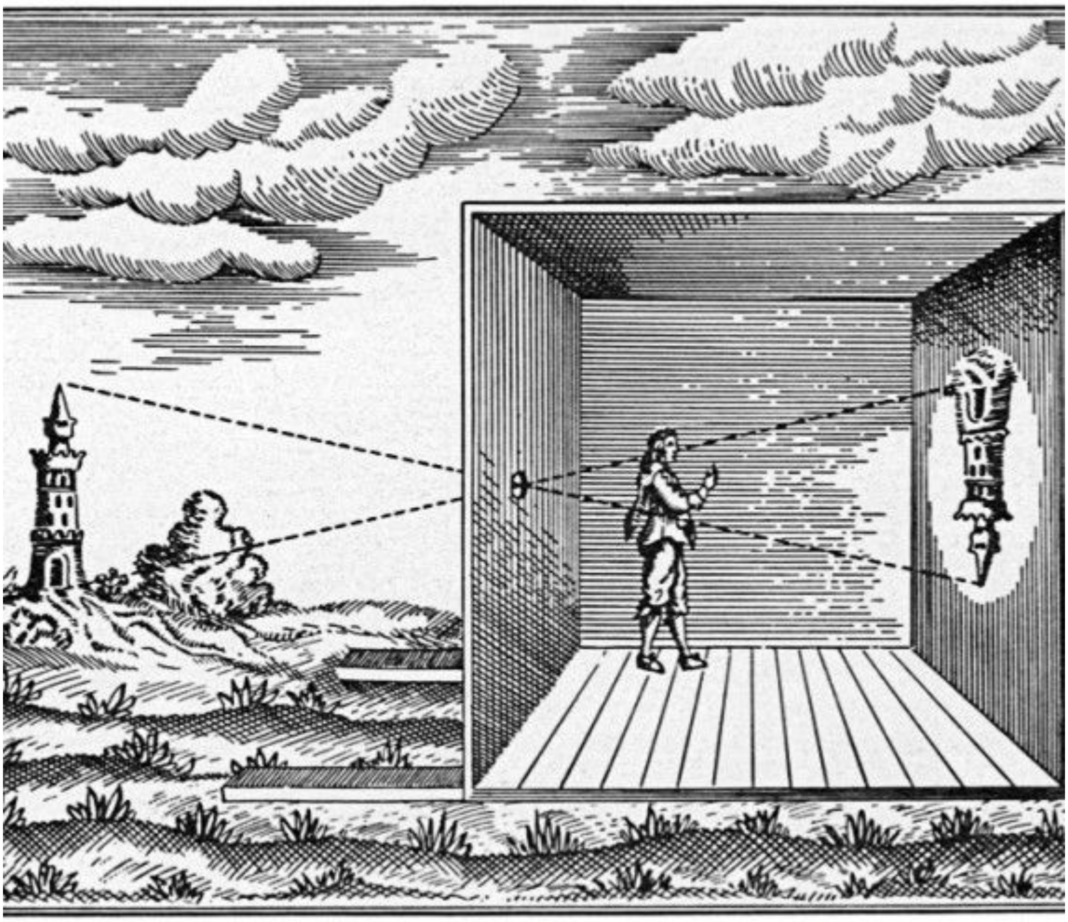
1. Mathematically remove radial and tangential lens distortion.
2. Adjust for the angles and distances between cameras, a process called rectification.

For which the main ingredients are the essential matrix E, which has info he translation and rotation

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{f} * \frac{f}{\lambda} k[R|t] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

that relate the two cameras in physical space about and the fundamental matrix F, the internal information about the camera.

We will get that by Stereo calibration is the process of computing the geometrical relationship between the two cameras in space.



A 17th century camera obscura illustration

The focal length tells us about a pinhole camera: how far away from the aperture do the light rays converge to create a clear image.

In this picture, the focal length is the distance between the wall with the hole and the wall with the image appearing. In modern camera modeling, we just flip this upside down image back around the origin back onto the positive Z axis, but it would still be on the plane Z = f.