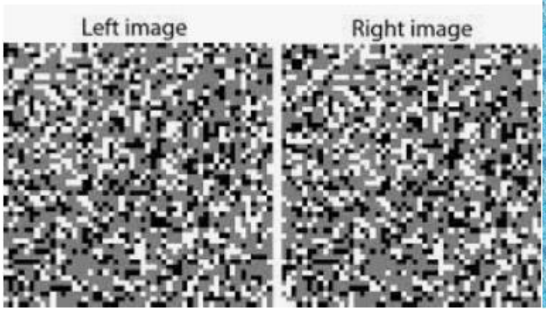
<u>**Stereovision**</u>

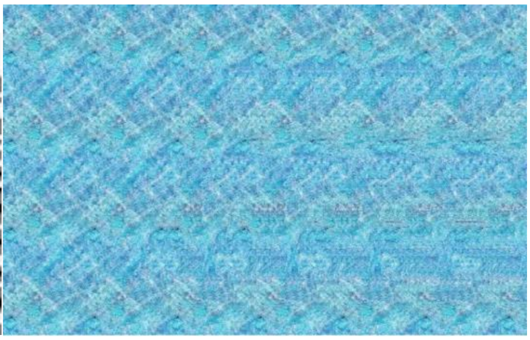
"**Computer stereo vision** is the extraction of 3D information from digital images, such as those obtained by a [CCD camera](). By comparing information about a scene from two vantage points, 3D information can be extracted by examining the relative positions of objects in the two panels. This is similar to the biological process of [stereopsis]()." (https://en.wikipedia.org/wiki/Computer_stereo_vision).

The active stereo vision is a form of stereo vision, which actively employs a light such as a laser or a [structured light]() to simplify the stereo matching problem. The opposite term is passive stereo vision.

Many animals use stereovision in order to have a sense of depth. Research in psychology of the vision system has made huge progress, especially from Bela Julesz, author of the Random-dot stereograms (https://en.wikipedia.org/wiki/Random_dot_stereogram). In recent times, auto stereograms, also called many times stereograms, allows one to see 3D from a single image (https://en.wikipedia.org/wiki/Autostereogram). "Most people with normal binocular vision are capable of seeing the depth in auto-stereograms, but to do so they must overcome the normally automatic coordination between accommodation (focus of the eyes) and horizontal vergence (angle of the eyes). "

Brewster-type stereoscope, 1870





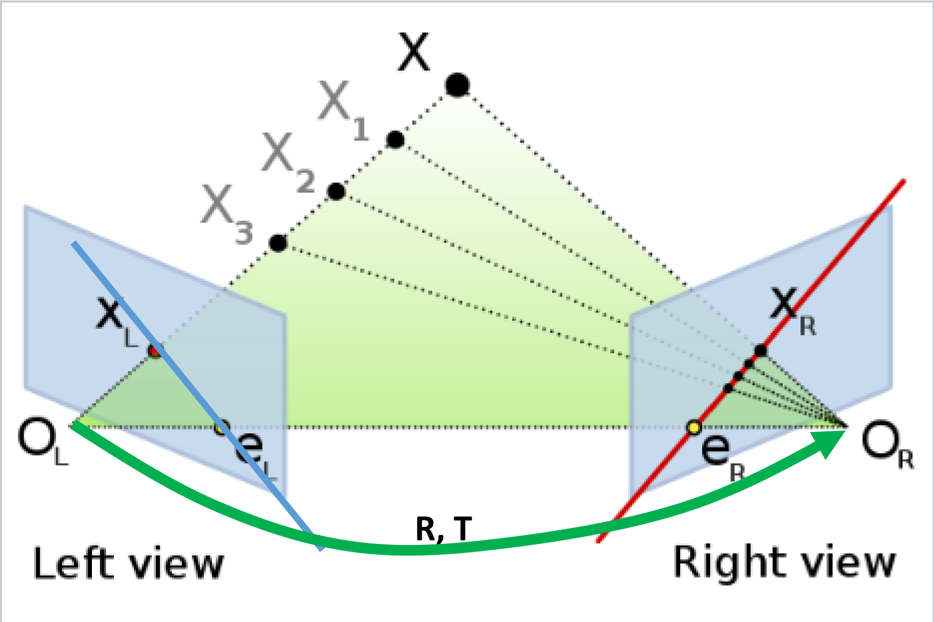
Normal stereogram Auto-Stereogram.

Practice stereograms in:

https://puzzlewocky.com/optical-illusions/3d-illusions/stereograms/.

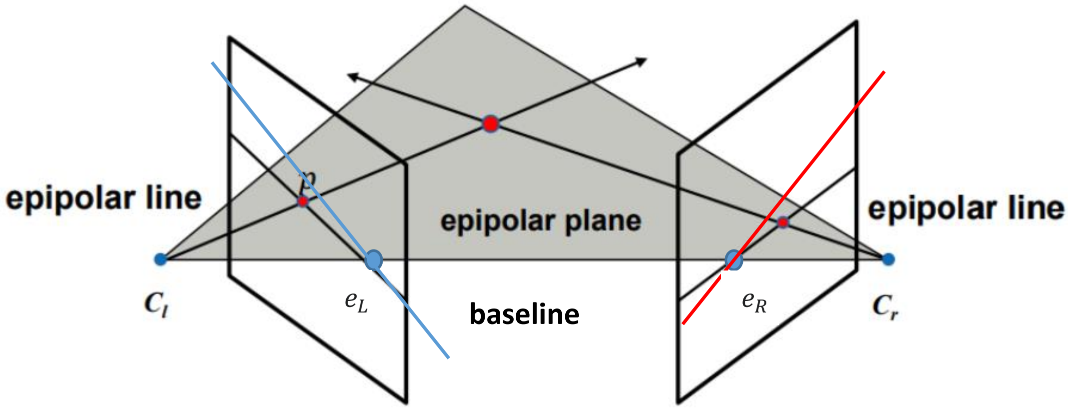
Epipolar Geometry

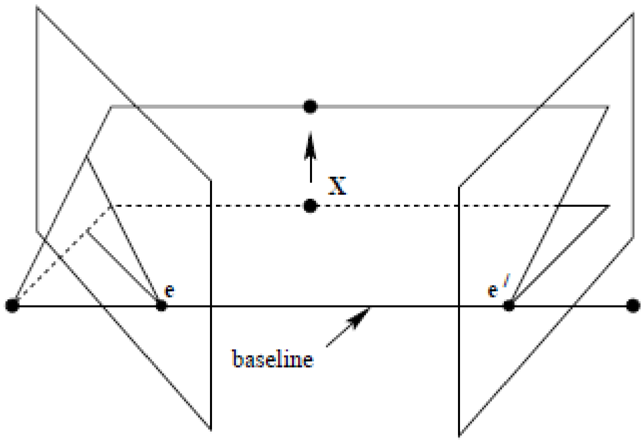
In computer Vision, passive stereo has been one of the most studied fields. In order to understand it, one of the most important concepts is the “epipolar constraint” (https://en.wikipedia.org/wiki/Epipolar_geometry).



Epipolar geometry

Suppose that we have two cameras, left and wright. We calibrate both cameras, and take out their distortions. For each point of the left camera (XL), there is a straight line X, whose projection is on a line on the right image (red line that goes from eR to XR). The red line is called the **right epipolar line**. The plane between the optical centers and the line X is called the **epipolar plane**. The intersection of the epipolar plane and the left image plane is called the **left epipolar line** (in blue in the left image). Finally, the intersection of the **baseline** (line between the left and right optical centers) with the left and right image, correspond to the **left and right epipoles** (eL, eR).





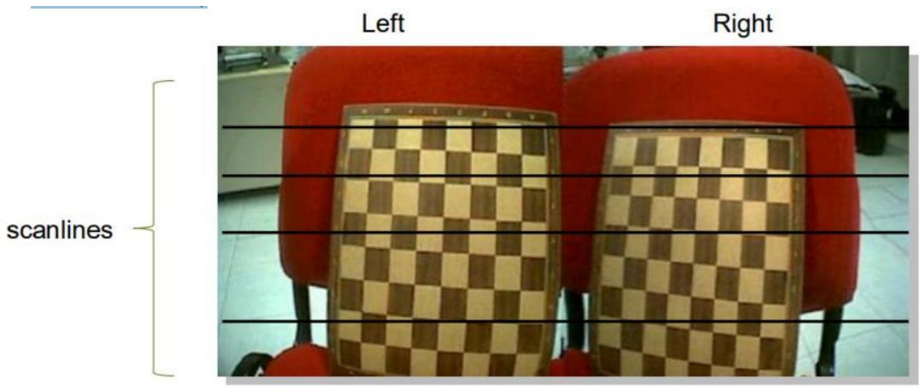
All epipolar planes pass though the baseline, also called **epipolar pincel**. Therefore**, all epipolar lines** pass though the **eipoles**, in other word, the epipoles are the intersection of all epipolar lines.

Thanks to the epipolar constraint, corresponding points can be searched for, along epipolar lines ⇒ computational cost reduced to 1 dimension!

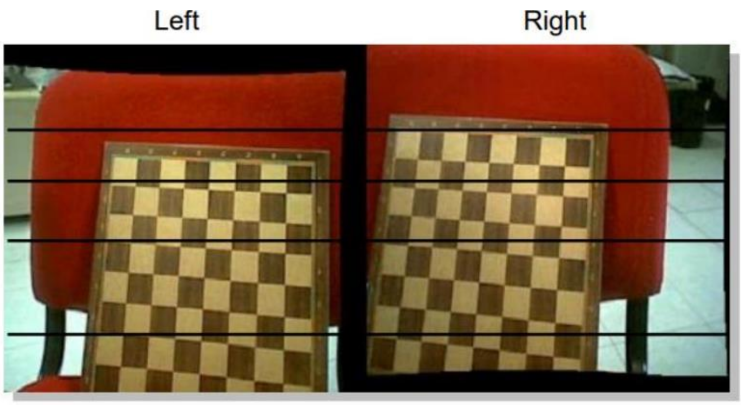
The problem with this constraint is that each pair of epipolar lines have a different direction in both the left and right images, so that it makes the searching more complicated. The solution, called **image rectification**, consists on modify the images, so that the epipolar lines are on the same horizontal lines.

Image rectification

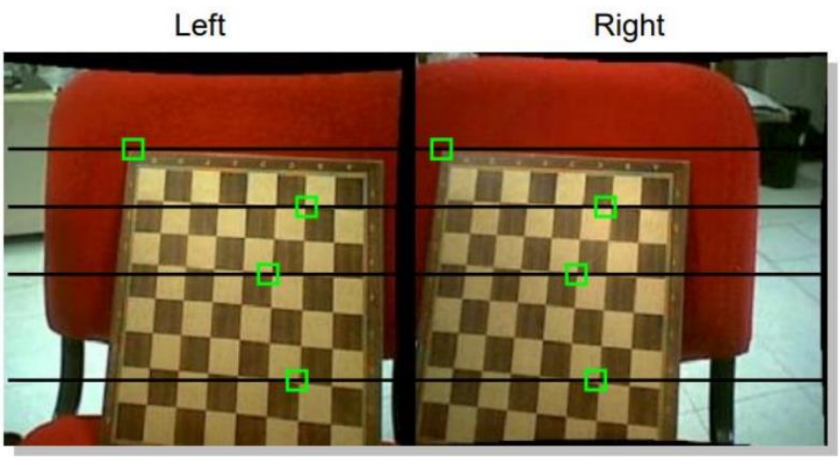
Goal: transform the left and right image so that pairs of conjugate epipolar lines become collinear and parallel to one of the image axes (usually the horizontal one)



In the above Figure, we see a pair of stereo-images with distortion and non-rectified. The first step is to remove the radial distortion.



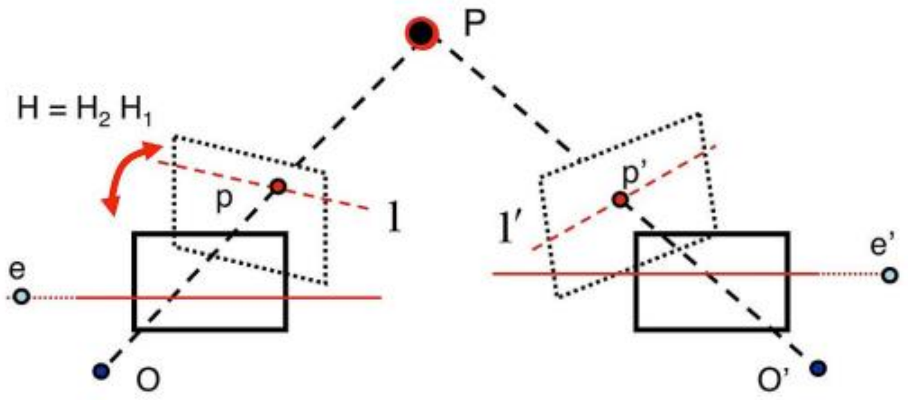
After taking out the distortion, the process of image **stereo-rectification** calculates some image transformations, based on **homographies** and some matrix properties of stereo (**fundamental matrix and essential matrix**), that transform the images so that the epipolar lines are parallel.

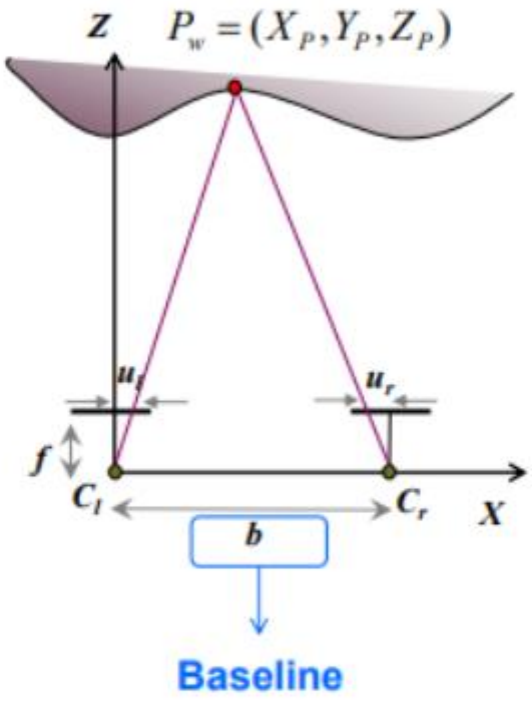


Correspondence is made line by line in the same horizontal line.

Calculation of the Z depth based on disparities

The rectification problem setup: we compute two **homographies** that we can apply to the image planes to make the resulting planes parallel.





After rectification, both cameras are parallel to each other, and with the same focal length f.

Disparity is the difference in image location of the projection of a 3D point in two image planes.

Baseline b is the distance between the two cameras.

From similar triangles:

$$\frac{f}{Z_P} = \frac{u_l}{X_P}$$

$$\frac{f}{Z_P} = \frac{-u_r}{b - X_P}$$

$$\Rightarrow \quad Z_P = \frac{bf}{\underbrace{\boxed{u_l - u_r}}_{\text{Disparity}}}$$

Many times this Z depth information is represented as a grey scale in the image, called Depth image or **D-image**. Also, the rgb image with depth information is called **RGB-D image**.

Example of a Depth image from stereo.

In the previous image, the points with higher depth are darker, while **the ones that are closer are brighter**. Sometimes the depth calculated from the rectified images is again transformed into one of the camera frames (usually the left camera frame).

Conclusion

Stereovision, active and passive, allows one to calculate the depth of the observed points relative to the cameras.

Stereo rectification allows the transformation of the images, so that the correspondence points are in the same horizontal lines, which can be parallized.

Several methods of calculating corresponding points in the same image lines have been developed based on different methods, e.g. correlation of images.

The mathematical formulation of stereovision from the geometrical point of view has been very well developed, and there are many free packages to calculate them (including Matlab packages).

In the following, we will explain the mathematical formulations of stereovision.