x and x' are the distance between points in image plane corresponding to the scene point 3D and their camera center. B is the distance between two cameras (which we know) and f is the focal length of camera (already known). So in short, above equation says that the depth of a point in a scene is inversely proportional to the difference in distance of corresponding image points and their camera centers. So with this information, we can derive the depth of all pixels in an image.

So it finds corresponding matches between two images. We have already seen how epiline constraint make this operation faster and accurate. Once it finds matches, it finds the disparity. Let’s see how we can do it with OpenCV.

Q2. By rectifying both images we reduce the number of unknown parameters in the E matrix. (we have no more rotation, only translation). This way, if the calibration matrixes for the cameras are known, this implies that the corresponding points have same “y” coordinate.