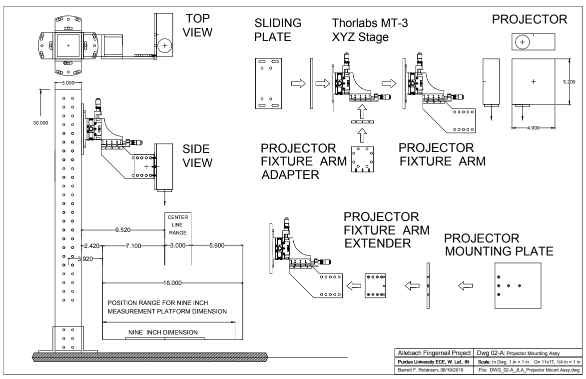
This will prevent deformation of the image during the calibration process. So we use a 3M MPro 150 VGA [9] (640 × 480) pocket projector (Fig. 2) in the experiment. The camera we are using is a Logitech C920S Webcam [10] (Fig. 3). This camera has different aspect ratio options. In the experiment, we are using 960 × 720, which is a 4 : 3 aspect ratio. This is the same as the projector's aspect ratio.

The engineering drawing for the system setup is given in Fig. 4. The picture for the actual mechanical setup can be found in Fig. 5. From now on, all the experiments and the results in this paper will be based on this experimental setup.



(a) Engineering drawing for the projector part

(b) Engineering drawing for the camera part

Figure 4: Engineering drawing for the system.

(a) Front view

(b) Top view

Figure 5: System Layout.

3. Camera and Projector Calibration

3.1 Camera Calibration and Verification

The camera can be viewed as a device that maps a 3D world point onto the 2D camera image plane. The basic camera model is the pinhole model. The goal of the camera calibration is to estimate the intrinsic parameters and the extrinsic parameters based on a pinhole camera model, but take the lens distortions into consideration, as well. The intrinsic parameters are: focal length ($f_x^c; f_y^c$), principal point ($u_o^c; v_o^c$), and skew coefficient f_{sq}^c. The extrinsic parameters are the 3 × 3 rotation matrix $\mathbf{R}_c$ and the 3 × 1 translation vector $\mathbf{T}_c$. The function of the extrinsic parameters is to map a point in the 3D world coordinate system to the 2D image plane. The intrinsic parameters are used to describe the camera's physical properties [4].

The camera model can be understood as follows: Denote the point in the world coordinates as $P_w = \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix}$. Given such a point, we first convert it to the camera coordinates, $P_c = \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix}$ by using the extrinsic parameters:

$$P_c = \mathbf{R}_c P_w + \mathbf{T}_c \quad (1)$$

As we stated before, the matrix $\mathbf{R}_c$ is a 3 × 3 matrix that describes the rotation property; and $\mathbf{T}_c$ is a 3 × 1 matrix that denotes the point's translation. Note that P_c here is a 3D coordinates point that represents the point in the camera coordinate system, and that the origin is the camera focal point. Then, we can use the property of the pinhole camera to convert the 3D point in the camera coordinate system to the normalized camera image plane, which is 2D