ECE 738 – Spring 2025
Projects 3&4 (see Introduction Lecture)
Due March 7 – 6:00pm

This multi-part project involves processing calibrated rectified stereo images of a lab scene:

1) to detect and match features in each stereo pair, preferably of the order of several thousands;
2) achieve subpixel accuracy;
4) reconstruct the scene in 3-D, preferably as dense as possible .

Stereo calibration results, before and after rectification, describing the relative geometry of 2 cameras for original and rectified images are provided. Various tasks of the project are as follows:

Feature Detection and Matching:

- **Task 1 (25%)**: locating several thousand features, ideally distributed over entire image more or less uniformly across the lab scene, by applying 2-3 feature detection methods. These may include Harris detector, SIFT, SURF, thresholding of min eigenvalue, etc.) and are to be selected from through various sources. To be precise: we preferably want several dozen features on various objects in the scene, where actual number depends on strength of surface markings/ texture richness). **Removing duplicate ones**, you can **combine features from more than one method to increase the total number of features, also** remove some to avoid clustered ones within a small region by non-maximum suppression.

- **Task 2 (25%):** applying one of more methods (e.g., SIFT, SURF) to match corresponding feature pairs in various stereo views. The goal again is to identify couple or more dozen matched features on various objects in the scene, according to surface markings/ texture richness of each target.

Scene Reconstruction:

- **Task 3 (25%):** determine disparity d*=x-x' associated with matched features p=(x,y) and p'=(x',y'), and implement a disparity growing scheme to increase the density of disparity values (see lecture slides 32-34).

- **Task 4 (25%):** use dense disparity to determine depth values from: $Z = f\,|B|/d^*$, where f is focal length and B is stereo baseline (see T below). Using these, plot both color-coded disparity and depth maps as one representation of 3-D scene, as well as 3-D point cluster of P=(X,Y,Z) values.

Data: Folder comprising of various original left-right stereo pairs of lab scene, as well as stereo-rectified ones that are associated with the "rectified calibration data" given below.

Original calibration (for original stereo pairs)

$$C_{left} = \begin{pmatrix} 1361 & 0 & 606.4 \\ 0 & 1365.4 & 518.3 \\ 0 & 0 & 1 \end{pmatrix} \quad C_{right} = \begin{pmatrix} 1360 & 0 & 629.0 \\ 0 & 1361.9 & 462.7 \\ 0 & 0 & 1 \end{pmatrix}$$

$$R = \begin{pmatrix} 0.9999 & 0.0112 & -0.0121 \\ -0.0110 & 0.9998 & 0.0158 \\ 0.0123 & -0.0157 & 0.9998 \end{pmatrix} \quad T = \begin{pmatrix} -259.24 \\ 3.95 \\ -1.41 \end{pmatrix} [mm]$$

Calibration after rectification (for rectified stereo pairs):

$$CAM_{left} = \begin{pmatrix} 1277 & 0 & 623.2 \\ 0 & 1277 & 490.4 \\ 0 & 0 & 1 \end{pmatrix} \quad CAM_{right} = \begin{pmatrix} 1277 & 0 & 616.8.0 \\ 0 & 1277 & 490.4 \\ 0 & 0 & 1 \end{pmatrix}$$

$$R = I_{3x3} \qquad T = \begin{pmatrix} -259.3 \\ 0 \\ 0 \end{pmatrix} [mm]$$