PROJECT 3

ENPM673

Semester:
Spring 2022

Student:
Anish Mitra

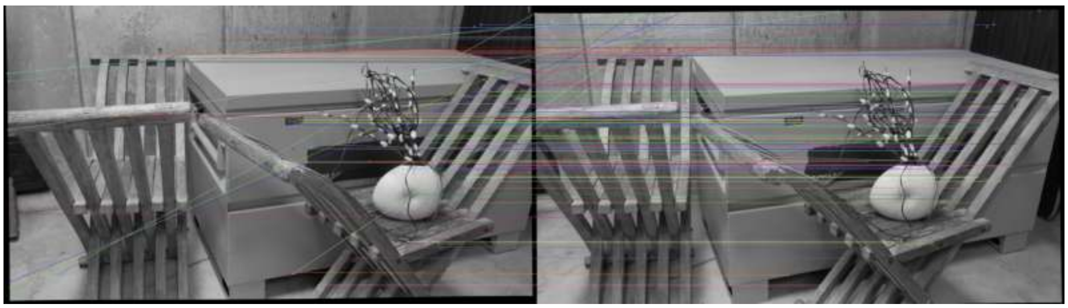
Course code:
ENPM673

Contents

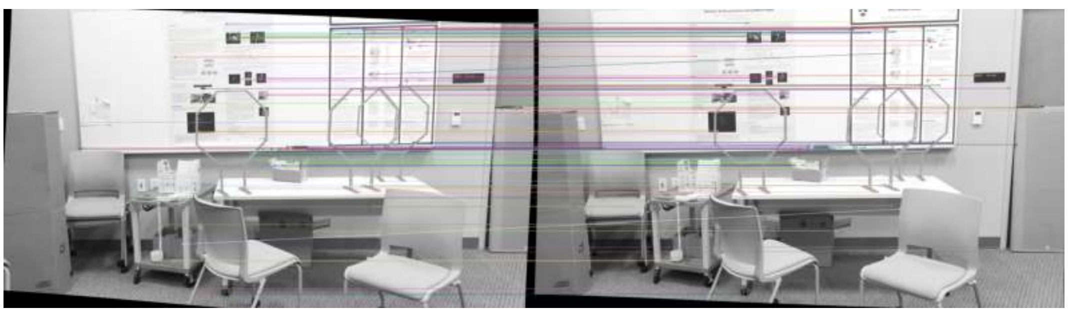
1 Calibration

1.1 Feature Matching

The choice of input image is taken from the user for the required data set. The first 100 matching points are taken into consideration. The first data set will be referred to as 'Curule', the second will be referred to as 'Octagon', and the third 'Pendulum'. The images are read from the respective folder and the required camera parameters are read from the **calib.txt** file. The features are matched between the two images using a SIFT object. Fig. 1 shows the SIFT keypoints matched for different datasets.



(a) Matched keypoints for Curule Dataset



(b) Matched keypoints for Octagon Dataset

(c) Matched keypoints for Pendulum Dataset

Figure 1: Matching SIFT keypoints for datasets

1.2 Fundamental Matrix

The first 100 matched keypoints are taken for further calculation. Fig. 1a shows the 100 matched keypoints for the Curule dataset. Similarly, fig. 1b and fig. 1c shows the matched features for the other two datasets. The fundamental matrix is calculated from the matched keypoints by solving the 'A' matrix shown in fig. 2 using SVD to obtain U,S,V matrices. The last column of V gives the fundamental matrix.

$$\begin{bmatrix} x_1x_1' & x_1y_1' & x_1 & y_1x_1' & y_1y_1' & y_1 & x_1' & y_1' & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_mx_m' & x_my_m' & x_m & y_mx_m' & y_my_m' & y_m & x_m' & y_m' & 1 \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{21} \\ f_{31} \\ f_{12} \\ f_{22} \\ f_{32} \\ f_{13} \\ f_{23} \\ f_{33} \end{bmatrix} = 0$$

Figure 2: AX=0

To find the best fitting fundamental matrix, RANSAC method is used by taking random 8 matched points at a time and iteratively repeating the process until the best matrix is found. An example of the best fundamental matrices for the 3 data sets are shown below :

$$\text{Curule Dataset Best Fundamental Matrix :} \begin{bmatrix} \text{7.52822663e-11} & 9.25604787e-08 & -4.05459600e-05 \\ -1.01284033e-07 & 8.63656276e-09 & 2.90790261e-03 \\ 4.18684399e-05 & -2.89510636e-03 & -3.02961387e-03 \end{bmatrix}$$

$$\text{Octagon Dataset Best Fundamental Matrix:} \begin{bmatrix} \text{8.21422441e-11} & -3.02888159e-07 & 1.35789539e-04 \\ 3.03286445e-07 & -5.60044748e-10 & -1.96990530e-03 \\ -1.35835282e-04 & 1.96173216e-03 & -5.41602046e-04 \end{bmatrix}$$

$$\text{Pendulum Dataset Best Fundamental Matrix:} \begin{bmatrix} \text{-3.18319662e-11} & 1.28581674e-07 & -6.07727043e-06 \\ -1.26800432e-07 & 4.56333846e-11 & -1.40069580e-03 \\ 5.39548856e-06 & 1.41189128e-03 & -3.26944603e-03 \end{bmatrix}$$

1.3 Essential Matrix

The essential matrix is then found using the fundamental matrix and the camera parameters given in the calibration file (cam0 and cam1) with the formula ($cam1^T.F.cam0$). The essential matrix for the three subsets are given below:

$$\text{Curule Dataset Essential Matrix :} \begin{bmatrix} \text{5.68988306e-05} & 5.80173184e-02 & 3.63793283e-03 \\ -6.31753454e-02 & 3.08256883e-03 & 9.97990401e-01 \\ -5.13130142e-03 & -9.98298590e-01 & 2.98027935e-03 \end{bmatrix}$$

$$\text{Octagon Dataset Essential Matrix:} \begin{bmatrix} \text{9.86827250e-05} & -2.93632052e-01 & -1.56844192e-02 \\ 2.92650766e-01 & -1.00234071e-03 & -9.56084540e-01 \\ 1.56536104e-02 & 9.55795358e-01 & -1.13259276e-03 \end{bmatrix}$$

$$\text{Pendulum Dataset Essential Matrix:} \begin{bmatrix} \text{5.28133857e-06} & 1.60543189e-01 & 4.71719815e-02 \\ -1.59635689e-01 & 6.42996283e-04 & -9.86052707e-01 \\ -4.67685991e-02 & 9.85915401e-01 & 5.48560241e-04 \end{bmatrix}$$

1.4 Rotational and Translational Matrices

The rotational and translational matrices are obtained from the essential matrix using SVD. The camera pose is estimated, from which the rotational and translational matrices are derived. The camera pose contains 4 different rotational and translational matrices, the best one is chosen using the triangulation cheirality condition obtained using least squares method. Decomposing the essential matrix we get the following translational and rotational matrices:

$$\text{Curule Dataset Rotational Matrix :} \begin{bmatrix} 0.99998572 & -0.00130569 & 0.00518273 \\ 0.00132127 & 0.99999461 & -0.00300512 \\ -0.00517878 & 0.00301192 & 0.99998205 \end{bmatrix}$$

$$\text{Curule Dataset Translational Matrix :} \begin{bmatrix} 0.99830895 & -0.00381232 & 0.05800615 \end{bmatrix}$$

$$\text{Octagon Dataset Rotational Matrix :} \begin{bmatrix} \text{9.99999416e-01} & 3.89896667e-04 & -1.00759173e-03 \\ -3.91073552e-04 & 9.99999241e-01 & -1.16808491e-03 \\ 1.00713553e-03 & 1.16847827e-03 & 9.99998810e-01 \end{bmatrix}$$

$$\text{Octagon Dataset Translational Matrix :} \begin{bmatrix} 0.95578983 & -0.0160274 & 0.29361355 \end{bmatrix}$$

$$\text{Pendulum Dataset Rotational Matrix :} \begin{bmatrix} \text{9.99999501e-01} & -3.11727985e-04 & 9.48963078e-04 \\ 3.11157061e-04 & 9.99999771e-01 & 6.01716682e-04 \\ -9.49150432e-04 & -6.01421106e-04 & 9.99999369e-01 \end{bmatrix}$$

$$\text{Pendulum Dataset Translational Matrix :} \begin{bmatrix} 0.98590095 & 0.04707539 & -0.16057154 \end{bmatrix}$$

2 Rectification

2.1 Homography Matrices

The images need to be rectified as they are taken from different perspectives. The homography matrices for both the images of the dataset (H1 and H2) are found to apply perspective transformation on the image. The homography matrices found for each dataset are included below:

$$\text{Curule Dataset H1 Matrix :} \begin{bmatrix} \text{2.81508991e-03} & -2.04969924e-04 & -1.41885247e-01 \\ -4.32740250e-05 & 2.89566791e-03 & 4.32013325e-02 \\ -1.04504164e-07 & 9.23734007e-09 & 3.00021852e-03 \end{bmatrix}$$

$$\text{Curule Dataset H2 Matrix :} \begin{bmatrix} \text{9.68287573e-01} & -3.30697898e-03 & 3.22296986e+01 \\ -1.44196936e-02 & 1.00005508e+00 & 1.38131630e+01 \\ -3.30277032e-05 & 1.12799052e-07 & 1.03164568e+00 \end{bmatrix}$$

$$\text{Octagon Dataset H1 Matrix :} \begin{bmatrix} \text{-1.62442245e-03} & 3.66832327e-05 & -2.84428419e-01 \\ 1.59711269e-04 & -1.96941535e-03 & -1.51099457e-01 \\ 3.56813053e-07 & -6.72351356e-09 & -2.31484162e-03 \end{bmatrix}$$

$$\text{Octagon Dataset H2 Matrix :} \begin{bmatrix} \text{8.25864214e-01} & -1.40554564e-02 & 1.74760301e+02 \\ -8.08533090e-02 & 1.00152086e+00 & 7.67979101e+01 \\ -1.81240618e-04 & 3.08455017e-06 & 1.17232534e+00 \end{bmatrix}$$

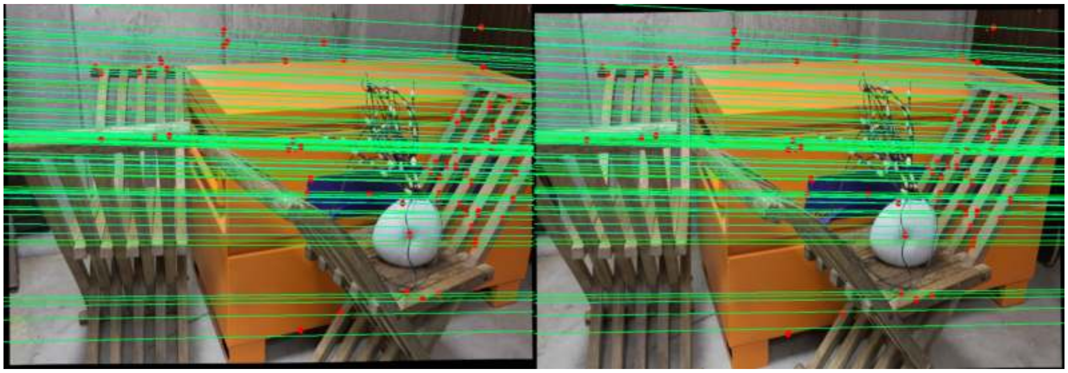
$$\text{Pendulum Dataset H1 Matrix :} \begin{bmatrix} \text{1.47886149e-03} & -2.57692071e-05 & -1.70007409e-01 \\ 4.87483315e-06 & 1.41333419e-03 & -9.08508794e-03 \\ 1.16405843e-07 & 4.84856737e-09 & 1.28565463e-03 \end{bmatrix}$$

$$\text{Pendulum Dataset H2 Matrix :} \begin{bmatrix} \text{1.07940365e+00} & 4.47194019e-02 & -1.00375985e+02 \\ 3.75245869e-03 & 1.00101331e+00 & -4.14954594e+00 \\ 8.36049609e-05 & 3.46373095e-06 & 9.17868823e-01 \end{bmatrix}$$

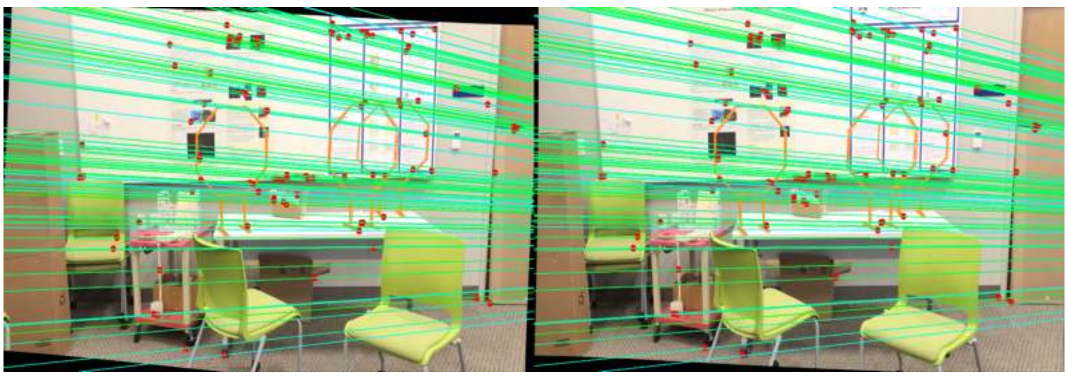
2.2 Before Rectification

To better visualize the effect of rectification on the image, epipolar lines were drawn on both the images of the data set. Fig. 3 shows the drawing of epipolar lines on each image before rectification.

Fig. 3a shows the unrectified epipolar lines for the Curule dataset. Similarly, the unrectified epipolar lines are drawn for the other two datasets as shown in fig. 3b and fig. 3c.



(a) Unrectified epipolar lines for Curule Dataset



(b) Unrectified epipolar lines for Octagon Dataset

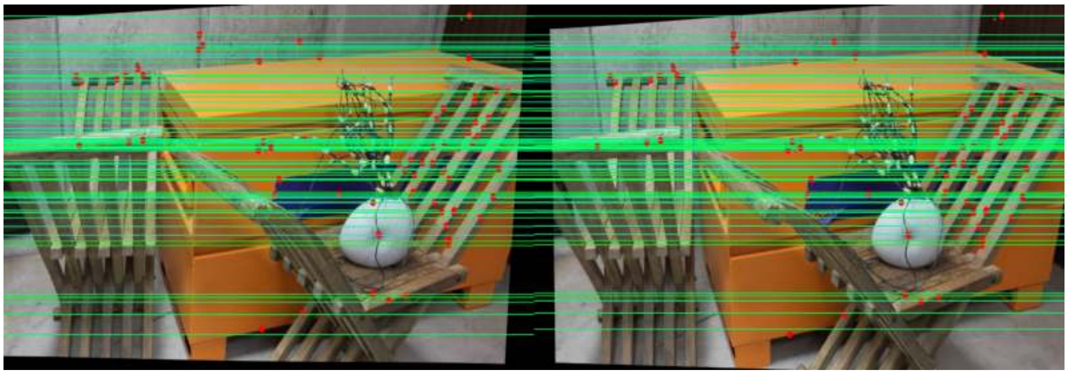
(c) Unrectified epipolar lines for Pendulum Dataset

Figure 3: Unrectified epipolar lines drawn for each dataset

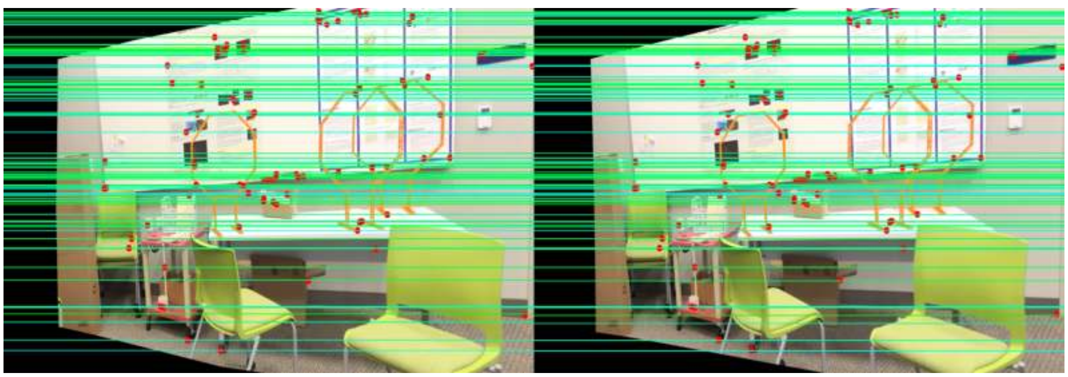
The green lines show the lines that are taken into consideration for further processing stages and the red dots indicate the standout features points of the images. As we can see, the epipolar lines are not horizontal.

2.3 After rectification

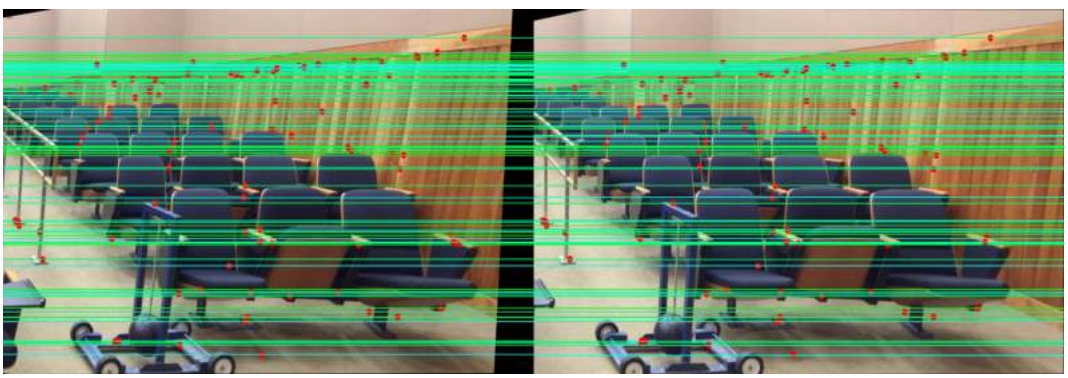
As we saw above, the epipolar lines are not horizontal before rectification of the image, and thus the homography matrix mentioned in the previous section is used to warp the perspective of the image to obtain horizontal epipolar lines. Fig. 4 shows the epipolar lines drawn on the images after rectification. There is a clear difference as the epipolar lines are all horizontal now. These horizontal lines help for further processing in the stereo vision pipeline. Fig. 4a, fig. 4b, fig. 4c show the epipolar lines after rectification for each dataset.



(a) Rectified epipolar lines for Curule Dataset



(b) Rectified epipolar lines for Octagon Dataset



(c) Rectified epipolar lines for Pendulum Dataset

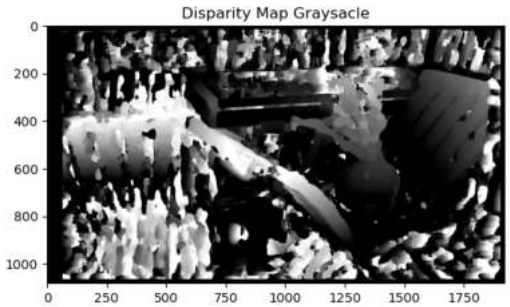
Figure 4: Rectified epipolar lines drawn for each dataset

3 Correspondence

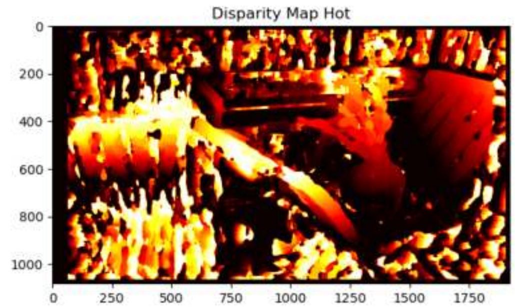
The disparity between the two images is calculated using the epipolar lines obtained from rectification. Fig. 5 shows the disparity map for each data set.

For each epipolar line, matching window technique is found where the surroundings of a pixel in the left image is compared to the slightly translated pixels in the right image without taking the full image context into account. This is achieved using sum of squared differences method (or SSD). The disparity is then rescaled to a range of 0-255 and viewed in the form of grayscale and heat map.

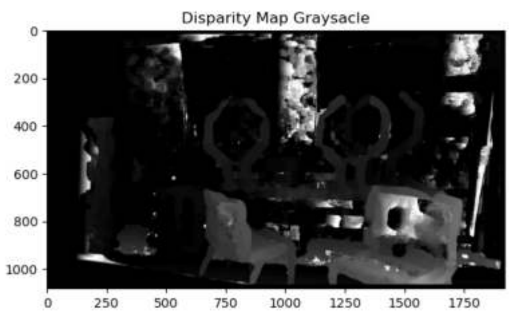
Fig. 5a shows the disparity map of the curule data set in gray scale and fig. 5b shows the disparity in heat. In both these cases, the lighter tones mean that the feature is closer to the camera, and the darker tones mean the object is further away from the camera. Similarly, fig. 5c, fig. 5d, fig. 5e and fig. 5f show the disparity maps for the other data sets.



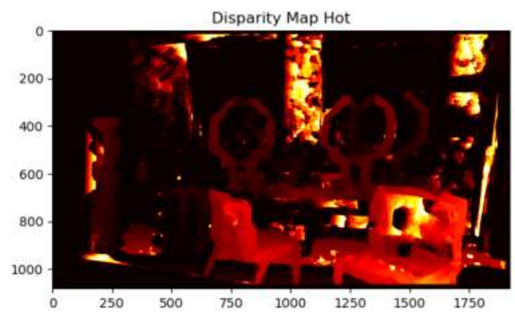
(a) Curule disparity map (grayscale)



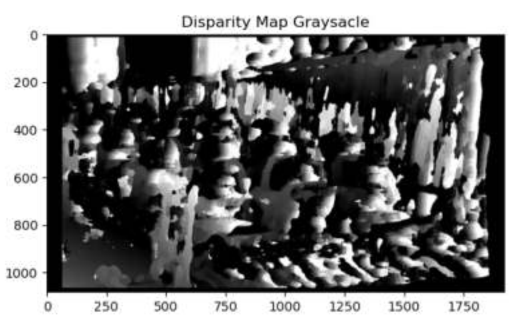
(b) Curule disparity map (Heatmap)



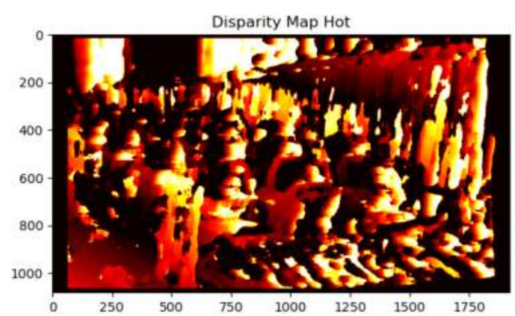
(c) Octagon disparity map (grayscale)



(d) Octagon disparity map (Heatmap)



(e) Pendulum disparity map (grayscale)



(f) Pendulum disparity map (Heatmap)

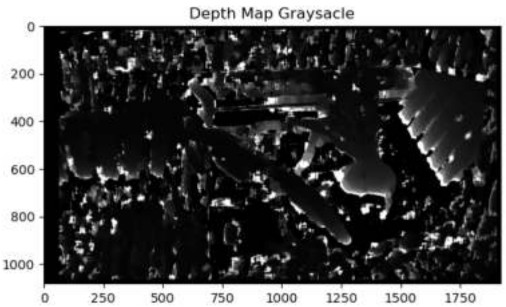
Figure 5: Disparity maps in gray scale and heat maps

4 Depth Maps

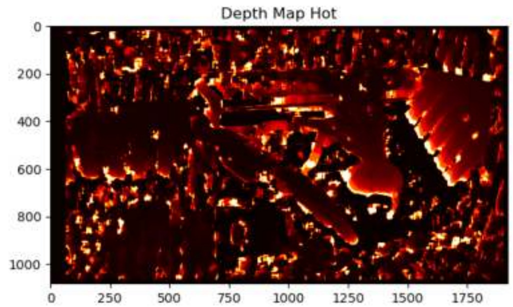
The depth image is calculated from the disparity using the baseline length and the focal length of the camera using the formula:

$$z = \frac{bf}{x_l^{'-x_r'}}$$

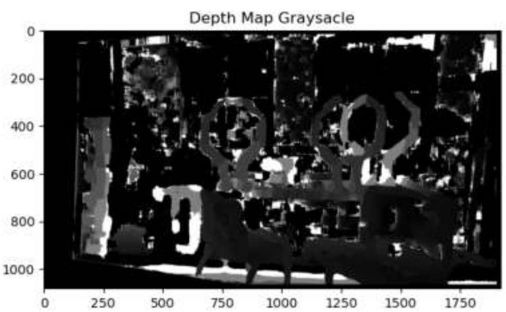
where b:baseline, f:focal length, $x_l^{'}$: left image disparity, $x_r^{'}$: right image disparity. The depth maps are shown in fig. 6. Each data set has a depth map generated for the image both in grayscale and heatmap color. Fig. 6a shows the depth map of the curule data set in gray scale and fig. 6b shows the depth in heat. In contradiction to the disparity maps, the lighter tones indicated that the feature is further away from the camera, and the darker tones are nearer to the camera. Similarly, fig. 6c, fig. 6d, fig. 6e and fig. 6f show the disparity maps for the other data sets.



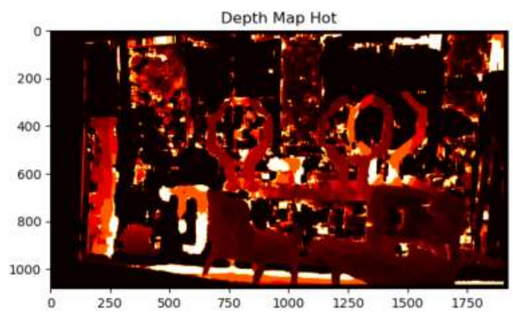
(a) Curule depth map (grayscale)



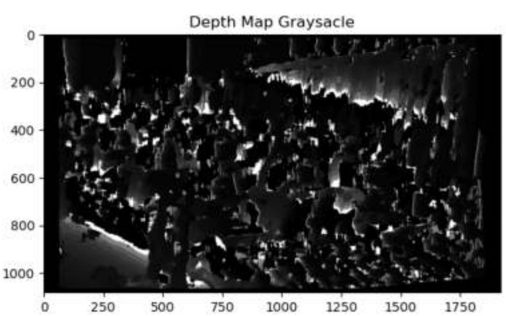
(b) Curule depth map (Heatmap)



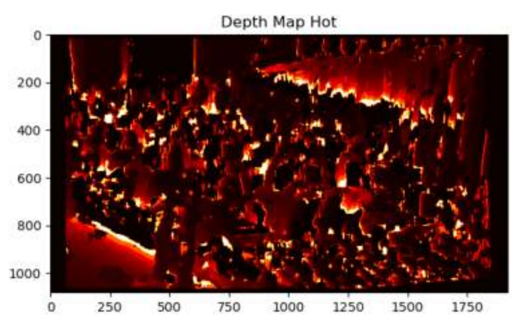
(c) Octagon depth map (grayscale)



(d) Octagon depth map (Heatmap)



(e) Pendulum depth map (grayscale)



(f) Pendulum depth map (Heatmap)

Figure 6: Depth maps in gray scale and heat maps