6. APPLICATION OF DEPTH ESTIMATION

In computer vision, depth-map is one of the most advantageous features of an image. In an image, there are different objects of various size and shape placed at a distance. If the knowledge of depth is known in prior then, task such as image segmentation, de-blurring, bokeh effect, and object recognition can be improved with greater accuracy. Depth-map of an RGB image has a various applications, among all we have build three most interesting and applicable applications.

6.1. Anaglyph 3d Stereo Imaging

Anaglyph is famous for 3d movies. Anaglyph is a stereoscopic photograph that has the two images of different colours which are superimposed. In this project, we have predicted depth map from a single image and calculated disparity using equation 3.1. We have assumed original image as left image and derived right image by applying disparity shift to the original image. We extracted red channel from the left image and extracted blue, green channel from derived right channel. The colour channels were blended to from anaglyph. Figure: 6.1 shows, a constructed anaglyph generated from a single RGB image.

(a) RGB-Image (b) Anaglyph

Figure 6.1: Anaglyphs 3d stereo imaging

6.2. 3d-Reconstruction

3d-reconstruction is the process of building the real shape, size and appearance of an real object. In general, 3d reconstruction of an environment is performed by an active or passive method. In this project, we have constructed 3d point cloud from a single image and with a depth map predicted from our proposed model 4.8. In compare to active method our method is much easier and doesn't not require heavier, larger and expensive devices like Lidar, Kinect and stereo camera setup. There are also different techniques which uses a single camera motion from video to reconstruct a 3d point cloud. But the from this technique sparse depth map is generated and requires multiple frames from a single output. Our model predicts depth-map from a single image and using the depth map, we have calculated 3d location (X, Y, Z) for each pixel (u, v) points.
Given depth value d at (u, v) image coordinate, the corresponding 3d point is:

$$x = \frac{(u - cx) * z}{fx}, y = \frac{(v - cy) * z}{fy}, z = \frac{d}{depth\_scale} \tag{6.1}$$

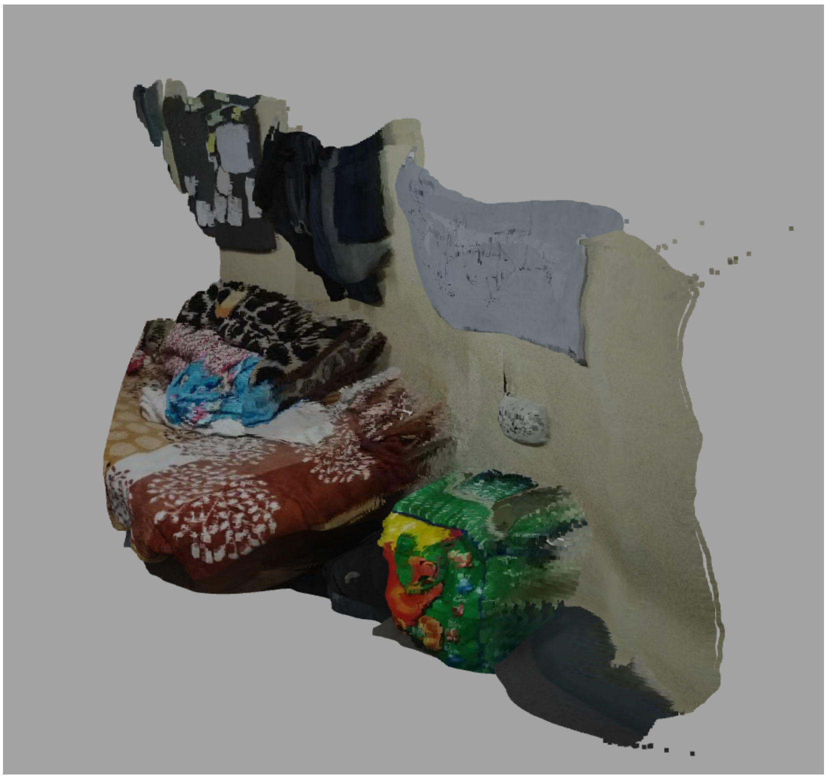


Figure 6.2: 3D reconstruction of a room

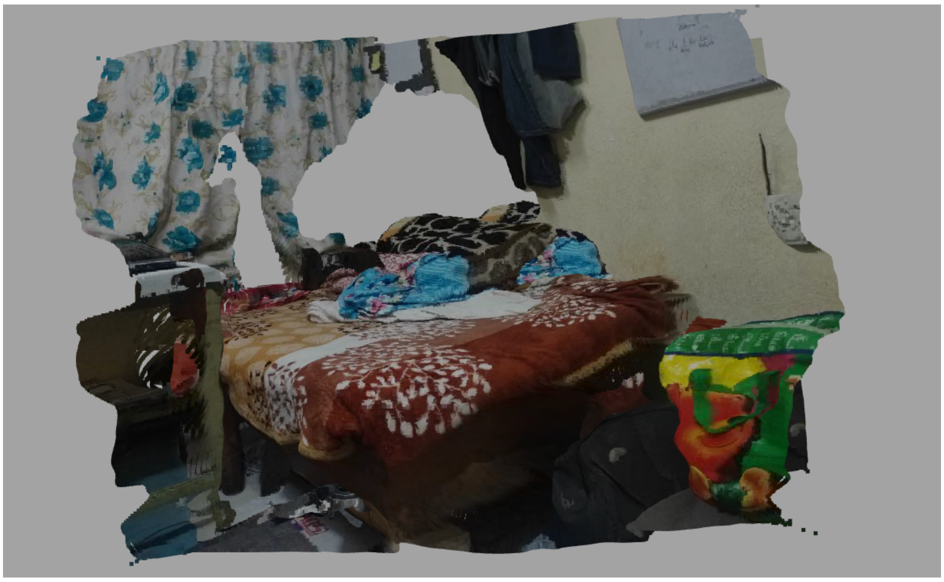


Figure 6.3: 3D reconstruction of a room

6.3. Dimension Measurement

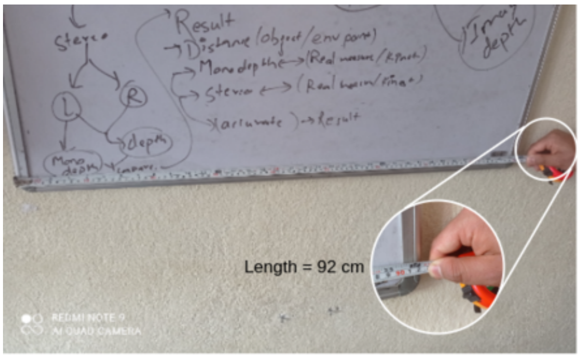
It has always been a concern matter to measure the dimension of any object in an image. One of the most interesting application of our proposed model 4.8 is to measure the dimension such as length, angle, area and volume of an object without any requirement of measuring tools and active devices. For this application, we have predicted depth map from our proposed model 4.8 and using the depth information we calculated 3d location (X, Y, Z) for each pixel (u, v) points using equation 6.1. After calculating the point clouds, each pair of a point in 3d space is assumed to form a vector. Magnitude of a vector gives the distance between any two points and the slope of a vectors is used to derive the angles between two vectors.
Used equation to calculate distance for given two points:
$X1 = [x_1, y_1, z_1], X2 = [x_2, y_2, z_2]$

$$distance = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \tag{6.2}$$

Used equation to calculate angle for given lines

$$\theta = \arccos\left(\frac{a^2 + b^2 - c^2}{2ab}\right) \tag{6.3}$$

(a) Measuring dimension and angle of whiteboard



(b) Measurement with measuring tape

Figure 6.5: Measuring dimension of a bag

In figure:6.4a, we have calculated the length of the whiteboard as 0.926 meter and angle $89.43°$ which is approximately $90°$ We have also compared our predicted result with the real measurement. Figure:6.4b shows, the measuring of whiteboard with measuring tape. We can clearly observe that the measurements are nearly equal. In figure: 6.5, we have calculated the height, width and breadth of a bag as 0.316m, 0.409m and 0.257m respectively, from this we can easily calculate volume and surface area of the bag.

6.4. Disaster Surveillance and 3-D Mapping

It is not always possible to visit and survey the disaster-hit to study about the prevailing impacts of the scenario since there may be life-risk. So the alternative can be drone surveillance. Drone captures images of a disaster area, which can be further processed. Since, it is not appropriate to mount a Lidar, Kinect or stereo camera with a drone due to its bulky dimension which make the monocular technique the best method for 3D-reconstruction of the scene. Video based reconstruction might be the solution but, our method can produce 3d reconstruction with a single frame so it requires less number of frames to construct dense 3d cloud of disaster areas. Our method predicts the depth map of a input image which is converted into 3d point cloud using above mentioned technique.

The point clouds thus generated with this method can be shared among various researchers for the study purposes which will make the workflow faster and reliable. Therefore researchers need not be present at the disaster-stricken areas.

An example of a 3-D reconstructed scene from an earthquake-hit area from Nepal is shown in figure 6.6.

Figure 6.6: 3-D reconstruction of a disaster-stricken area