Brian Maguire

CIS-475

Dr. Zhang

HR: <https://docs.opencv.org/2.4/modules/flann/doc/flann_fast_approximate_nearest_neighbor_search.html>

<https://docs.opencv.org/3.3.0/da/df5/tutorial_py_sift_intro.html>

<https://docs.opencv.org/2.4.1/modules/calib3d/doc/camera_calibration_and_3d_reconstruction.html#stereobm-stereobm>

<https://docs.opencv.org/3.0-beta/doc/py_tutorials/py_calib3d/py_depthmap/py_depthmap.html>

**Structure from Motion**

**Introduction:**

There are many uses for computer vision in the modern-day world. In an ever-expanding technical industrial age, the need for a machine to have a pair of eyes is paramount. A part of giving a machine vision is being able to identify objects in that are in motion. Objects move all the time, so being able to tell where and how far away an object is while still identifying it is an impressive feat for machines. To do this, researchers created structure from motion. The goal of structure from motion is being able to identify an object that has changed positions in different ways. Whether an object be shown at a different angle or a different depth, openCV has several methods that are available to help accomplish this goal.

**Methods:**

In this project, two different methods were utilized to detect an object in different positions. The first was just to detect characteristics common to a specific object then find them in another picture, side-by-side. To go through the code: first all of the images were inserted then SIFTed, which stands for “scale-invariant feature transform.” The SIFT method breaks down a picture to different scales then detects key points within these. Once this was complete, the nearest neighbor algorithm had to be utilized in order to find harsh differences in gradients in the images. These differences indicate where objects start and end. After this step, three arrays are created that hold the good points, the points from the first image, then the points from the second image. Then, ratios are created between the images and compared. When the ratios are created, lines are drawn from a specific point in each picture and compared in both. The results are shown in the “results” section.

For the second method, which measures the depth of an object, the stereo method was used. The images were both imported into the file and the disparity between the two images was calculated to show the difference between them. Once, that is shown we can see the edges on the speaker as well as the raised surfaces. This image is shown in the results section.

**Results:**

A screenshot of a cell phone

Description generated with very high confidenceFor the first program, similar features of the same object are shown in the image.As one can see, the lines are traced to several identifiable features like the corners or the center logo in the upper center of the speaker. Each line is traced in both images to the same location within the image. Even though the speaker is shown at different angles, it can be inferred that the speaker is the same in image because the points have the same features. The results of this method being used were fairly successful as the identification of the different points were accurate. If this were used for a video rather than an image, its results may not be as successful because the calculation would take a considerable amount of time.

In the second image, the depth and edges can be seen in the image below as the light areas are raised and the dark areas are the background. Given this, the speaker can be seen in the foreground as well as the corner of the wall to its right. At the top left of the image, the light on the ceiling can be seen as well. This program was not as successful as the previous program. This may be due to the lighting of the image or the similarity in color from the different sections of the wall or the ceiling. If this were to be used with a video, it may yield much better results than just an image.

