

Part 2:

SIFT (Scale Invariant Feature Transform) is a method for obtaining distinct and invariant features from photos. It can be used for a variety of computer vision tasks, such as image matching and object detection. The SIFT method consists of several stages:

Scale-space extrema detection: In this step, potential interest points in the image that are both orientation- and scale-invariant are sought after. This is achieved by identifying regions where intensity changes noticeably using a difference-of-Gaussian function.

Key point localization: A comprehensive model is fitted to each possible point of interest to ascertain its exact location and scale. The selection of key points is based on stability metrics.

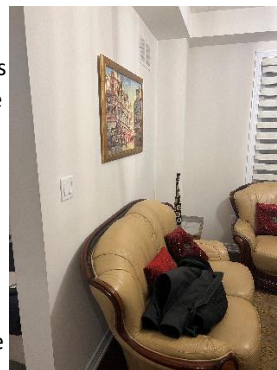
Orientation assignment: Every key point is given one or more orientations based on the local image gradient directions. This step ensures that the image rotation won't affect it.

Key point descriptor: Each key point's surrounding local image gradients are measured at the selected scale. These gradients are transformed into a form of representation that allows significant variations in illumination and local shape distortion. This descriptor captures the special characteristics of the key point.

For image matching, SIFT features are extracted from a set of reference photos and stored in a database. Euclidean distance is used to compare each feature individually to the features in the database whenever a new image is displayed. To find candidate matching features, their feature vectors are compared for similarity. This process allows for the reliable matching of different perspectives of an object or scene.

When it comes to image stitching, SIFT features can be used to align and combine multiple images into a panoramic view. The key point descriptors between the images are matched, and the transformation parameters (such as translation and rotation) are estimated, to properly align the images. By using the matched key points to blend the overlapping areas of the images, a seamless panoramic image is created.

When all is said and done, SIFT features provide resistance to image rotation, scale, affine distortion, noise, and changes in lighting. Precise matching against a large feature database is possible due to their extreme distinctiveness. Standard hardware can achieve real-time performance because SIFT features are computed efficiently.





My camera could be rotated horizontally, and the algorithm recognized this and stitched the images appropriately. There are two black "circles" on top and bottom because of the camera's rigidly horizontal movement. If I wanted to eliminate this blank space, I could simply take pictures above and below the horizontal while maintaining an overlap.

