Homework 4 OpenCV

Computer Vision 2021/22

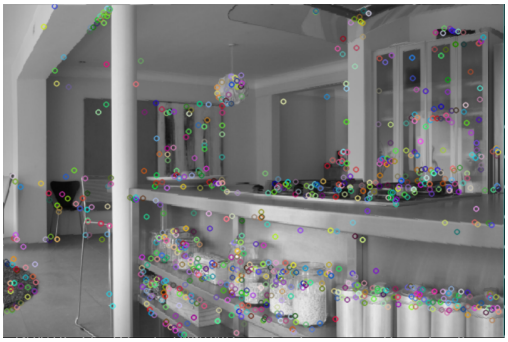
Alessandro Manfè 2041603

May 16th 2022

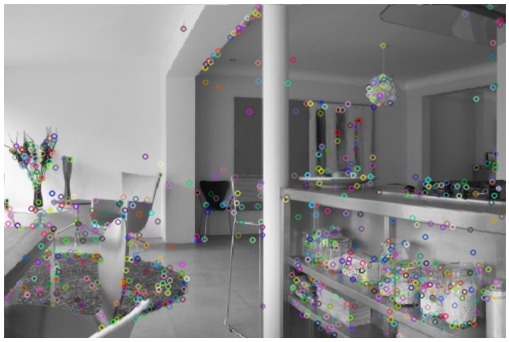
This homework's report purpose is to provide the experimental results of some application of OpenCV library. In particular we're going to see some examples of image stitching, using feature detection and matching techniques.

Task 1

The goal of this task is to merge a dataset of similar images taken in temporal succession in order to create what is called in photography a "Panorama". In order to do so we're using feature detection to get keypoints of each image. The method we're using it's called SIFT, which allows us to compute description of these features as well.
N.B. we first need to compensate the FOV angle of the camera performing a cylindrical projection of the image.



Featuers: Kitchen 1



Featuers: Kitchen 2

Using brute force matcher (BFMatcher class) we then compute the matched instances of features for each pair of images. The resulting list of matches is then filtered by a threshold obtained by the multiplication of the minimum distance among all DMatch and an arbitrarily defined ratio.
We can visualize the matched fatures by keypoint representation of the previous images:

Feature matching

The next step is using those informations to get a transformation matrix, precisely a perspective transformation matrix, that describes exactly how much an image is misplaced with respect to the other one. In particular this kind of matrix is in the form:

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{bmatrix} \tag{1.1}$$

For the purpose of this task we want to derive the translation component along the x coordinate (width). To obtain the correct value we need to divide the result of the perspective transofrmation by a term w, function of g and h and different for each point value $(gx + hy + 1)$.
Most of the time the transformation is mainly given by translation and rotation so 'g' and 'h' components are close to 0. We can simplify the matrix and obtain an affine transformation matrix, whose 'c' term represent the translation distance we need for the next part:

$$\begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix} \tag{1.2}$$

From here we just have to arrange the images one next to the other in order and with the proper alignment. We can simply create a larger frame and copy each image into it using an incremental value for the x coordinate.

Some of the images in the datasets are not really aligned along the y axis and perfectly orientated (some rotation is present) so due to our previous semplification there might be incorrect edge matchings. Notice that also brightness and contrast might differ between adjacent images.

The results of stitching for each dataset are the following:

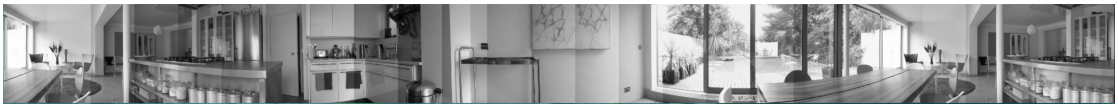


Image 1: kitchen

Image 2: dolomites

Image 3: lab

Image 4: lab automatic

Image 5: lab manual