Homework2

March 2025

Hi, everyone. This is the second homework for our computer vision class. This assignment includes two parts: short answer and programming problems. Here are the **requirements**:

- You are required to use Python for all the programming tasks.
- We recommend you typeset your report using LATEX and submit the PDF. You are asked to submit a single zip file on WebLearning. The zip file should consist of the report and your code.
- Referring to the public code on the internet is allowed, but copying is absolutely forbidden.
- Please finish the programming and report independently.
- We provide the code framework for you (see in src.zip attached). Zip all your files and name them as: name_id_hw_2.zip,
  eg: zhangsan_2020123456_hw_2.zip, and the structure of the directory should be like this:

```
zhangsan_2020123456_hw_2
├── report.pdf:  your report file
└── src:  put your code
    ├── README.md:  more details for code implementations
    ├── problem1_sift:  solution of SIFT problem
    └── etc...
```

1 Implement SIFT

Please implement the Scale-Invariant Feature Transform (SIFT) algorithm by yourself. Test your code on *school_gate.jpeg* and visualize the keypoints with their scales and orientations as shown in Figure 1 (to make the figure not too messy, you can visualize a subset of them). Furthermore, you are asked to verify the translation, rotation, and scale invariance of SIFT, i.e. drawing lines between the corresponding keypoints before and after translation, rotation, and changing scale.

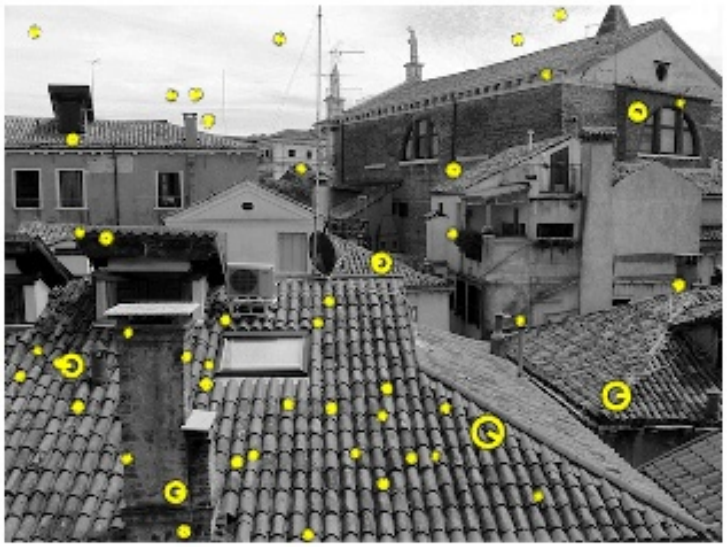

Figure 1: An example of SIFT visualization.

2 Matching

Use your own SIFT to find the best robust feature matches (correspondences) for the features in *book_reference.jpeg* and features in image *books.jpeg*. The "robust" means that you should use RANSAC to remove the outliers. Visualize the matching results as shown in Figure 2.

Figure 2: The visualization example of matching.

3 Homography

In this exercise, you are given a screenshot *football.png* of a realistic football game where the player is taking a free kick. The size of the goal area, the yellow area shown in Figure 3, is $18.32m \times 5.5m$. You are asked to estimate: (1) the distance of the football from the goal line (2) the distance of the referee from the football (3) the horizontal distance of the goalkeeper's left foot from the left goalpost, i.e. the three red lines in Figure 3. (Error less than 5% is allowed)

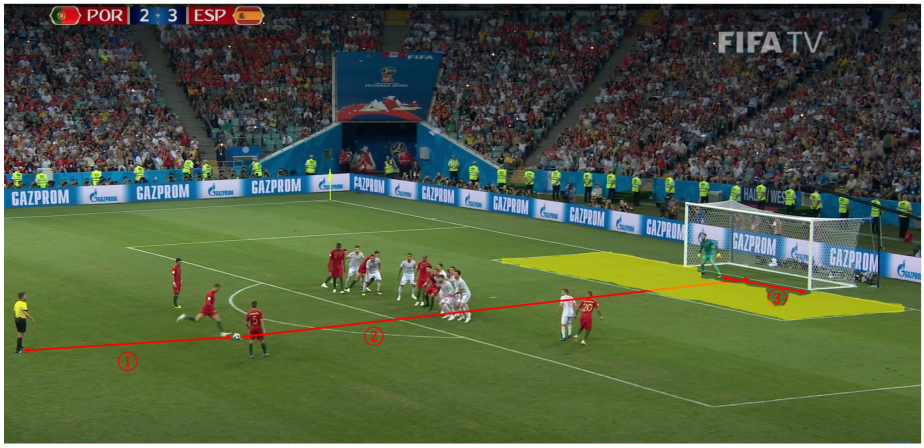

Figure 3: The football game screenshot.

4 Panorama Stitching

You are given a set of photos. The goal is to take the photos, *building_image/building_1~4.png* and stitch them into one photograph. You can do this by extracting SIFT features from both photos, matching them, and estimating the homography of one photo with respect to the other. Use RANSAC to find the best homography. Once you compute the homography, "stitch" the photos together, forming a panorama.