

Computer Vision Project: Object Detection and Localization

Conti, D'Auria, Tangaro

April 28, 2025

1 Pipeline

In this project, we developed a pipeline to detect and localize objects in test images, as an application of keypoint-based feature matching against a set of model views. Our approach consists of the following stages:

1.1 Feature Extraction

Run SIFT to detect keypoints and compute descriptors on the test image. For each model view of each object, run SIFT to detect keypoints and compute descriptors within the object (after applying binary mask, if needed).

1.2 Feature Matching

Use a brute-force matcher to compare its descriptors with those of the test image. For each descriptor, find its two nearest neighbors. Apply Nearest Neighbor Distance Ratio (NNDR) test: only keep a match if the best match is significantly better than the second one.

1.3 Object Presence Decision

Track the model view with the highest number of good matches. If said number exceeds a set threshold, consider the object as present in the image.

1.4 Object Localization

Select top- N ($N = 4$ in our case) best model views. Collect coordinates of matched keypoints in the test image from each selected model view. Merge all matched points in a single set. Find dense regions corresponding to the object by applying Mean Shift Clustering. Draw a minimum area rotated rectangle over the points of the largest cluster to identify the object.

2 Project Structure

The project is organized with the following directory and file structure:

```
object-detect-cv25/  
  data/  
    object_detection_dataset/...  
  results/  
    004_sugard_box/  
    006_mustard_bottle/
```

```
    035_power_drill/  
include/  
    dataloader.hpp  
    detection.hpp  
    matching.hpp  
    object_localizer.hpp  
    preprocessing.hpp  
    utils.hpp  
src/  
    dataloader.cpp  
    detection.cpp  
    main.cpp  
    matching.cpp  
    object_localizer.cpp  
    preprocessing.cpp  
    utils.cpp
```