Introduction

We introduce a real-time landmark recognition application for popular Stanford landmarks. The application relies on an initial filtering stage based on GPS location, followed by a Bag-of-visual words [4] histogram matching. After this stage, a homography is estimated between the query and the top result from the histogram matching step using RANSAC algorithm. We use ORB [2] (Oriented and Rotated BRIEF) for fast keypoint detection on the Android platform. Our application does not bank on server-side support and the entire processing happens on the Android device. This is advantageous in terms of lower latency and power requirements. For performance analysis, we used SIFT features [1] followed by vocabulary tree based histogram matching and homography estimation.

The Bag-of-visual words tree was created using 5000 images from San Francisco Landmark Data Set for Mobile Landmark Recognition [3], while testing was performed on 33 images from 6 Stanford landmarks.

screenshots

Our Approach

New query
ORB Feature Extraction
Histogram of visual words
GPS-based Filtering

Match not found

Compare inliers with threshold
RANSAC-based Homography
Top match from histogram matching

Declare match
Match found

Performance Results

The experiments used a dataset collected by us. For cross-device testing, 95 images were used, with leave-one-out matching results reported.

Observations

- When query and database images are from the same device, the accuracy is higher than when cross-device testing is performed.
- As can be inferred from the pie-chart, RANSAC is the most time-intensive step of the algorithm. By restricting the number of RANSAC checks to 1, we manage to keep the total life-cycle of a frame’s processing to <1000 ms.

References