CS231m Project Proposal

Building a landmark recognition system without server-side processing

Anirban Chatterjee
achatter, 05883022

Introduction

Landmark recognition on mobile devices is a well-researched topic and extensive publications exist demonstrating real-time recognition on mobile devices. However, even though the data acquisition is performed on the device, typically, almost all of the processing is carried out on dedicated servers. This ensures fast processing and energy efficient query processing, however, network bandwidth places significant latency on recognition and information retrieval. Typically for tourists who have visited a new place, network access is not guaranteed, rendering server based approaches useless. If we decided to port this design completely to client side, we would suffer from data storage issues. We aim to address this using a modular approach, discussed in the following sections.

Plan of approach

We plan to use a vocabulary tree to match query images with the database. Since storage on portable devices is limited, we will use a "module" based approach. Tourists can download these packets on their smartphones, which can then be used offline to match queries. This is similar to several "Maps" apps, which follow a similar approach. The modules themselves are histograms created by "pushing" the descriptors down a previously created vocabulary tree. The vocabulary tree is ported with the application too.

Technical details

The target platform is Android 4.4.3. We will test our application on the Tegra Note device we have access to. This will also allow us to detect performance metrics accurately.
The steps we have in mind are as follows:

1. Create a database of landmarks. We will stick to landmarks on the Stanford campus for demonstration purposes.
2. Create a vocabulary tree with descriptors collected from several images.
3. For each image in the database, "push" the descriptors down the vocabulary tree and store the computed histogram.
4. Port these histograms as well as the vocabulary tree as downloadable "module"s to the device.
5. On receiving a query, use histogram-based method to recognize the landmark.
6. Display the information regarding that landmark on the screen.

Milestones

(i) Collect image database: This step has been completed. We have a collection of images of major buildings on campus. These have been snapped from different viewpoints, under different illumination and with different devices (Tegra Note, iPhone and Moto Droid).

(ii) Run algorithm on a stand-alone computer and check performance: Initial SIFT based matching has provided good recognition results. Vocabulary tree based implementation has to be tried out. -> This week

(iii) Porting the application to device: This application has to be ported to the device. -> Next week.

References

Publicly available databases:
http://www.stanford.edu/~dmchen/mvs.html#sflandmarkdataset
http://www.robots.ox.ac.uk/~vgg/data/parisbuildings/
http://www.vision.rwth-aachen.de/data
Previous Work:
https://stacks.stanford.edu/file/druid:yt916dh6570/Alterman_Jaffey_Ho_Location_Reviews.pdf