

## Abstract

Image matching is a well studied problem in the computer vision community. Starting from template matching techniques, the methods have evolved to achieve robust scale, rotation and translation invariant matching between two similar images. To this end, people have chosen to represent images in the form of a set of descriptors extracted at salient local regions that are detected in a robust, invariant and repeatable manner. For efficient matching, a global descriptor for the image is computed either by quantizing the feature space of local descriptors or using separate techniques to extract global image features. With this, effective indexing mechanisms are employed to perform efficient retrieval on large image databases.

Successful systems have been put in place in desktop and cloud environments to enable image search and retrieval. The retrieval takes fraction of a second on a powerful desktop or a server. However, such techniques are typically not well suited for less powerful computing devices such as mobile phones or tablets. These devices have small storage capacity and the memory usage is also limited. Computer vision algorithms run slower, even when optimized for the architecture of mobile processors. These handheld devices, or so-called smart devices are increasingly used for simple tasks that seem too trivial for a desktop or a laptop and can be easily accessed on a smaller display. Further, they are more popularly used for taking pictures (gradually replacing the space of digital cameras) owing to the improved embedded camera sensors. Hence, a user is more likely to use a query image from the mobile phone, rather than from the desktop. This increases the scope of applications that demand real-time search and retrieval result delivered on a mobile phone.

Many applications (or apps) on mobile smart phones communicate with the cloud to perform tasks that are infeasible on the device. People have attempted to retrieve images in this cloud-based model by either sending the image or its features to the server and receiving back relevant information. We are interested to solve this problem on the device itself with all the necessary computations happening on the mobile processor. It allows a user to not bother for a consistent network connection and the communication overheads associated with the search process. We address the range of applications that need simple text annotations to describe the image queried on the mobile. An interesting use case is a tourist/student/historian visiting a heritage site and can get all information about the monuments and structures on his mobile phone. Once the app is initialized on the device, the camera is opened and just pointing the camera or with a single click all the useful info about the monument is displayed on the screen instantly. The app doesn't use the internet for communicating with any server and should do

all computations on the mobile phone itself. Our methods optimize the process of instance retrieval to enable quick and light-weight processing on a mobile phone or a tablet.

Firstly, we obtain a dataset of images relevant to the application's scope, which is then curated and annotated for the task. This may seem a trivial task, but involves a great deal of manual labour when dealing with datasets of thousands of images. We efficiently address this with a batch annotation approach that results in detailed and part-wise image annotations. Then, we look into the primary problem of performing image retrieval on a mobile phone and propose variants of the Bag-of-words approach. A new indexing mechanism helps us to reduce the memory footprint by a large extent. It also reduces the computational requirements. Finally, we develop the mobile app using existing computer vision libraries for smart phones, keeping in mind a simple and user-friendly design for seamless use by a wide range of users.

We validate our approach by using the standard benchmark datasets of Oxford-5K for our experimental purposes and also show results on Paris-5K and UKBench datasets. We demonstrate a working mobile app designed for a heritage site that successfully retrieves relevant information about a monument captured by the mobile phone.