GPU-Boosted Camera-Only Indoor Localization

Özkil, Ali Gürçan; Fan, Zhun; Kristensen, Jens Klæstrup; Dawids, Steen; Christensen, Kim Hardam; Aanæs, Henrik

Publication date:
2010

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
WHERE AM I?

GPU-BOOSTED CAMERA-ONLY LOCALIZATION FOR MOBILE ROBOTS

Abstract
Localization can be defined as the process of estimating the pose of an agent, given a representation of the environment and sensor input. In this work, we use Topo-metric Appearance Maps to represent the environment and introduce a new method for localization using only a camera. The method relies on local image features detection, description and matching; by parallelizing these computationally intensive tasks on the graphical processing unit (GPU), it is possible to do online localization using a Topo-metric Appearance Map. The method is developed as an integral part of a mobile service robot system [1], and empirically evaluated using a real robot in a typical indoor environment.

Map Generation
Building a Topo-metric Appearance Map is an offline process, and it requires a ‘capable’ robot, which is equipped with wheel encoders to estimate the path of the robot, a range sensor to detect objects in the environment and to correct the path estimation, and a camera to capture appearances. The path estimated by wheel encoders are prone to system noise and integration errors, and it needs to be corrected using a particle filter [2]. This process is usually referred as Simultaneous Localization and Mapping (SLAM), and it has two outputs: A metric grid-map and the corrected path. Using the corrected path and the metric map of the robot, it is possible to select a sparse set of appearances from the dataset of collected images, and correlate them to metric poses based on their timestamps. The final result of this process is the Topo-metric Appearance Map, a sparse set of appearances with known positions and orientations.

Signature Generation
A signature is the descriptive characteristic of an appearance, whereas an appearance is a representative image of a scene in the environment. One way of creating a signature is to use local image features. In this work, we adopt Scale Invariant Feature Transform (SIFT) [3] for detecting and describing features in appearances.

Camera-only Localization
Having the Topo-metric Appearance map and the signature vector, it is possible to do pose estimation, using a ‘less capable’ system which only uses a camera as the only sensor. The system simply compares every taken image to the images in the appearance map. In the majority of applications, localization has to be an online process. Conventional methods for feature detection and matching are computationally intensive and time-consuming; rendering this method impractical. However, they are also massively parallel, which are suitable to run on graphical processing units (GPU) to achieve significant speed-ups over conventional (CPU based) approaches. This method relies on a GPU based approach [4]: live camera frame is transferred to GPU. Its signature is computed and matched against the signature vector of appearances. Finally, the position to the best match(es) is returned from the GPU, and the pose is estimated according to the topo-metric appearance map.

Empirical Validation
The method is implemented as a mobile platform, and it is evaluated in a typical indoor environment for its online performance. The ‘capable’ robot consists of a Pioneer 3-AT robot base, SICK-LMS100 laser range finder, a USB camera and a laptop computer. The robot is manually guided in an indoor environment (~30x40 meters), and the collected data is post-processed to obtain the topo-metric appearance map, which consist of 204 images (out of 915 collected).

Localization performance is evaluated using a ‘less capable’ setup which simply consisted of a USB camera connected to a MacBookPro laptop with discrete Nvidia graphics chipset (gt-9600). The setup is placed on a wheeled push-cart, and manually moved inside the environment with a walking speed (~1m/s). Best five matches from the appearance map was tracked, and it was observed that the system was able to consistently estimate its correct position in the map at a speed of ~1Hz, despite the changes in illumination, partial occlusions and elevation of the camera.

Conclusion
Localization is an essential requirement for the majority of mobile robot applications, and in this work, we introduce a gpu-boosted camera-only method for online localization and evaluate it in real-life settings. With much room for improvement, the potential of the method is eminent. Eliminating the need for additional sensors (such as laser range-finders) for localization can significantly decrease costs, power needs and mechanical/electrical complexities of autonomous robots.

References