Technical University of Denmark

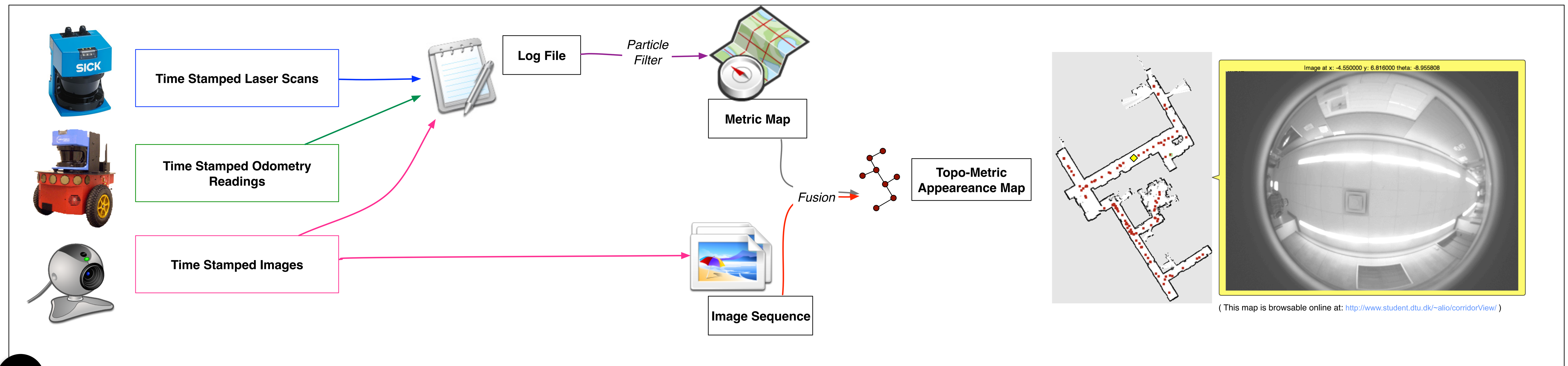
GPU-Boosted Camera-Only Indoor Localization

Özkil, Ali Gürcan; 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):
Özkil, A. G., Fan, Z., Kristensen, J. K., Dawids, S., Christensen, K. H., & Aanæs, H. (2010). GPU-Boosted Camera-Only Indoor Localization. Poster session presented at DTU Visiondays Conference 2010, Kgs.Lyngby, Denmark.



(This map is browsable online at: http://www.student.dtu.dk/~alio/corridorView/)

1 Map Generation

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 1

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 2

A *signature* is the quantitative characteristics 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 adopted *Scale Invariant Feature Transform (SIFT)* [3] for detecting and describing features in *appearances*.

At the end of this stage, a vector of *signatures* is obtained, which correspond to the features of *all* images in the Appearance map.

Camera-only Localization 3

Having the *topo-metric appearance map* and the *signatures* 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, it's *signature* is computed, and matched against the *signatures vector* of *appearances.* Finally, the pointer to the best match(es) is returned from the GPU, and the pose is estimated according to the *topo-metric appearance map.*

Empirical Validation 4

The method is implemented on a mobile platform, and it is evaluated in a typical indoor environment for it's online performance. The 'capable' robot consists of a Pioneer 3-AT robot base, SICK-LMS200 laser range finder, a USB camera and a netbook 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).

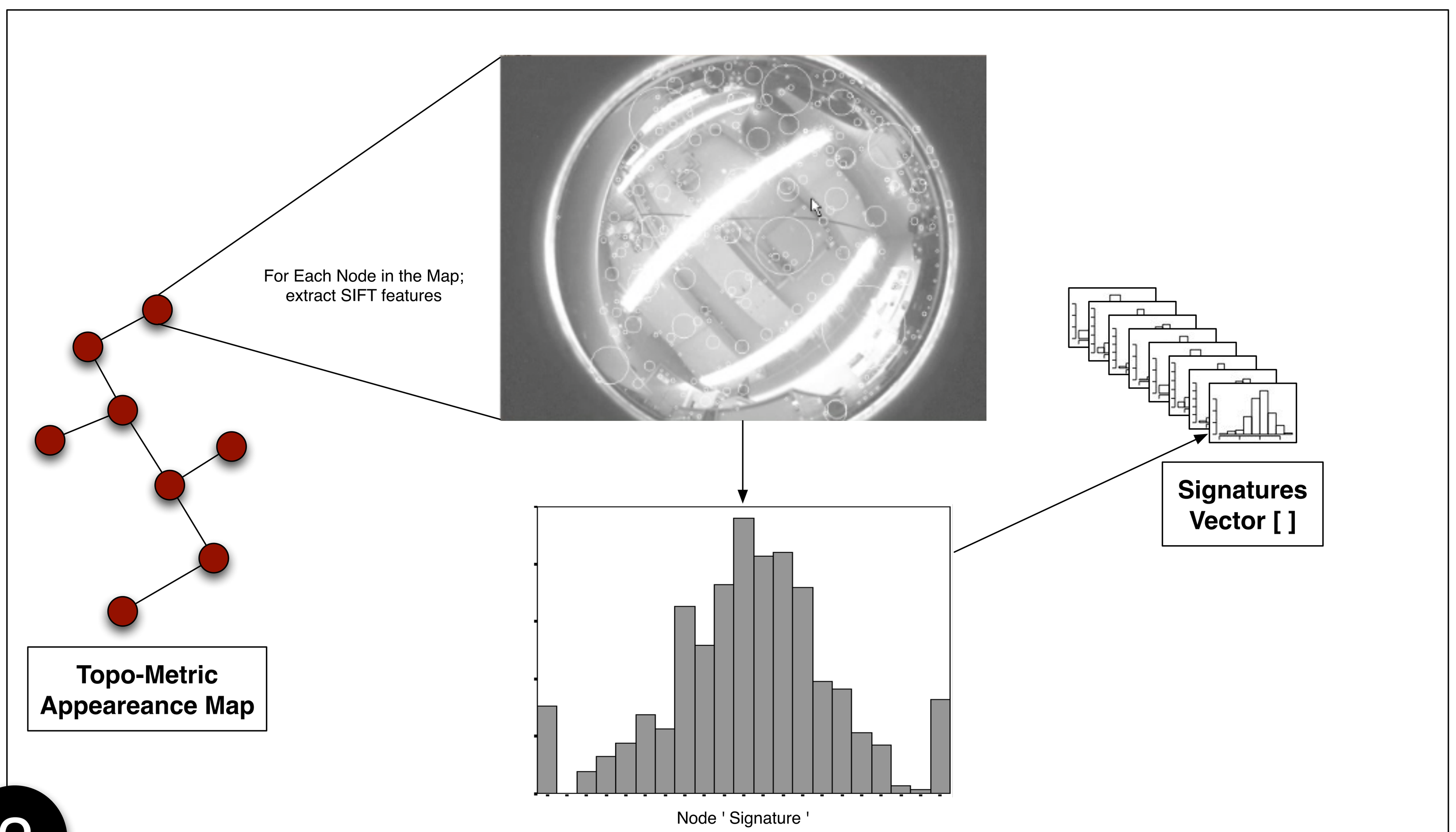
Localization performance is evaluated using a 'less capable' setup; which simply consisted of a USB camera connected to a MacbookPro laptop with discrete Nvidia graphics chipset (gt-9600). The setup is placed on a wheeled push cart, and manually moved inside the environment with fast walking speed (~1m/s). 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

[1] A. Ozkil et.al, Service robots for hospitals: A case study of transportation tasks in a hospital, In proceedings of International Conference on Robotics and Applications, 2009
[2] G. Grisetti, et.al., Fast and Accurate Slam with Rao-Blackwellized Partocle Filters, Robotics and Autonomous Systems, 2007
[3] D.G. Loewe, Distinctive Image Features From Scale-Invatiant Keypoints, International journal of Computer Vision, 2004
[4] C. Wu, SIFT-GPU: A GPU Implementation of Scale-Invariant Feature Transform, http://cs.unc.edu/~ccwu/siftgpu, 2007



2 Signature Generation

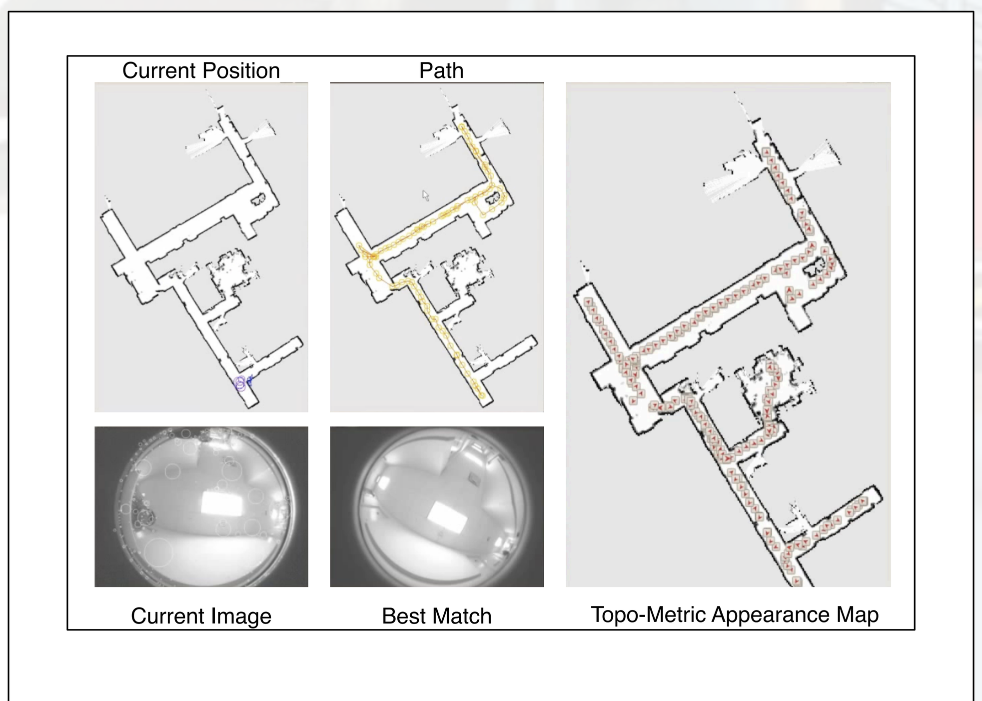
'capable' setup for mapping

'less capable' setup for localization

After mapping the environment with a 'capable' robot, the localization experiment is done by moving the laptop-camera setup using a push-cart. (i.e. *no odometry, no range-sensors)*

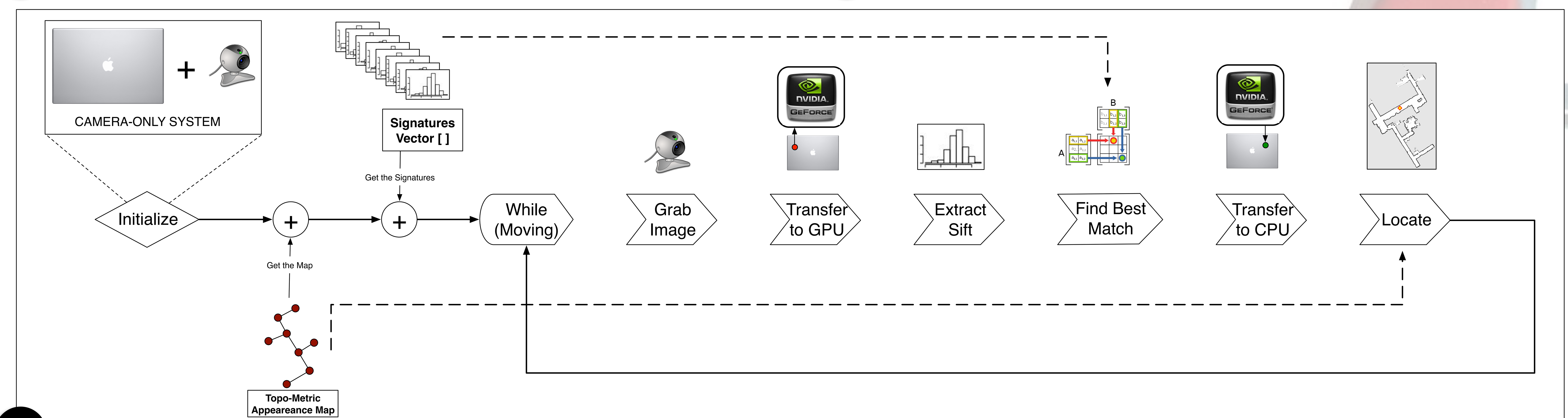
The cart is moved inside the environment and the best match was tracked over the time to obtain the path followed.

The system is able to re-estimate its position correctly at a rate of 1 Hz. At a fast walking speed (~1m/s) it is possible to track the path online.



A screenshot from the experiment

4 Empirical Validation



3 Camera-only Localization

Ali Gurcan Ozkil [1,2] [alio@man.dtu.dk], Zhun Fan [1] [zhfa@man.dtu.dk], Jens Klæstrup Kristensen [2] [jek@force.dk], Steen Dawids [2,3] [sdawids@man.dtu.dk], Kim Hardam Christensen [2] [kmc@force.dk], Henrik Aanæs [4] [haa@imm.dtu.dk]

1: Engineering Design & Product Development, Department of Management Engineering, Technical University of Denmark
2: Welding & Production Innovation, Force Technology, Denmark
3: Bispebjerg Hospital, Hovedregionen, Denmark
4: Image Analysis & Computer Graphics, Institute for Mathematical Modelling, Technical University of Denmark