First newsletter for the EUROPA2 project

This newsletter is about the EU-funded research project EUROPA2 (European Robotic Pedestrian Assistant 2.0).

The Project

In recent years, there was a growing interest in robots operating in urban environments such as city centers. Such urban areas are highly dynamic and complex environments which introduce numerous challenges to autonomous robots. This covers aspects of representing the environment including semantics as the vehicle needs to be able to effectively store and extend maps of its environment, to plan its actions, and to localize itself in a highly accurate way even if GPS signals are missing. Furthermore, reliable navigation requires solutions to several complex problems regarding the perception of the environment, state estimation, and interpretation of the gathered information. All of this considering the uncertainty associated with the data and thus the resulting decisions. In addition to the task of estimating their own position in the environment, robots deployed in urban areas need to identify and estimate the positions and velocities of other relevant agents in their vicinity such as those of potential users, but also those of other traffic participants like pedestrians, bicycles and cars.

The EUROPA2-project is concerned with several aspects of mobile robotics:

- Online semantic scene interpretation for more efficient navigation.
- New means to combining publicly available prior map models with the perceptions of the robot, enabling the system to maintain and exploit map knowledge without having mapped a city beforehand.
- Bootstrapping the navigation system with publicly available maps to avoid requiring to map the environment with the robot in advance.
- Life-Long operation of the whole robotic system, including map management, navigation, and calibration.
- Perception under different weather conditions and seasons.
- Introspective classification for robust decision making, i.e., considering the certainty of classification results appropriately.
- To extend and improve existing online methods for detecting and tracking dynamic objects towards close-range and partially visible pedestrians and other traffic participants.
- Develop the foundations for representations of dynamic urban environments for autonomous robots. This includes better models for the interaction between people and the robot as well as with other traffic participants such as cars.

The Consortium

To achieve these goals, the EUROPA2-consortium consists of leading researchers in the fields of autonomous robots, autonomous cars, navigation, vision and perception, and mobile mapping. EUROPA2 is targeted at developing novel technologies that will open new perspectives for commercial applications of service robots in the future.

- Albert-Ludwigs-University of Freiburg (Germany) - Autonomous Intelligent Systems Lab
- University of Oxford (England) - Mobile Robotics Group
- Katholieke Universiteit Leuven (Belgium) - Center for the Processing of Speech and Images
- Eidgenössische Technische Hochschule Zürich (Switzerland) - Autonomous Systems Lab
- GeoAutomation (Industry partner)
- University of Bonn - Department for Photogrammetry

Progress

In the beginning of the project the robot platforms that existed already from the previous project needed to be adapted with additional sensors, like fast 3D range scanners. They are mostly identical apart from the robot in Oxford that is also equipped with an experimental radar to detect far away obstacles.

Freiburg Zürich Oxford

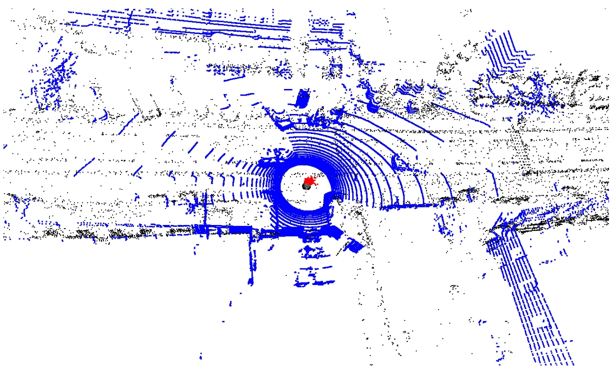
The consortium performed and published research regarding many different aspects of the project. A full list of the publications is available on the EUROPA2 homepage at `http://europa2.informatik.uni-freiburg.de`.

In the following we will give a few details about some of the performed work.

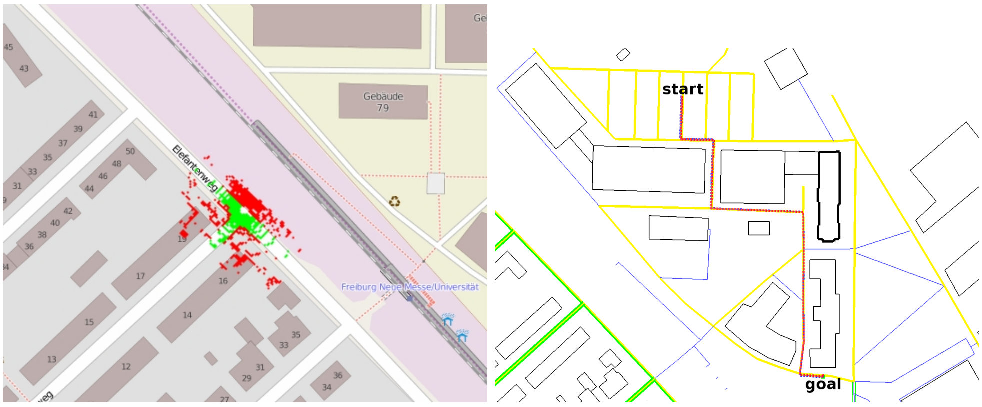
Using publicly available maps

One essential prerequisite for autonomous operation of a mobile robot is to know the position of the platform in the world. This enables the robot to plan efficient paths from its current position to a desired goal location. The challenges in the context of the EUROPA2 project are the scale of the environments the robot is supposed to operate in. To be able to make use of publicly available maps for this purpose, we need to relate the perceptions of the robot to the map material, thereby requiring accurate position estimates within the map frame. GPS measurements help to reduce the search space but are not sufficient by themselves since they are not accurate enough and not always available.

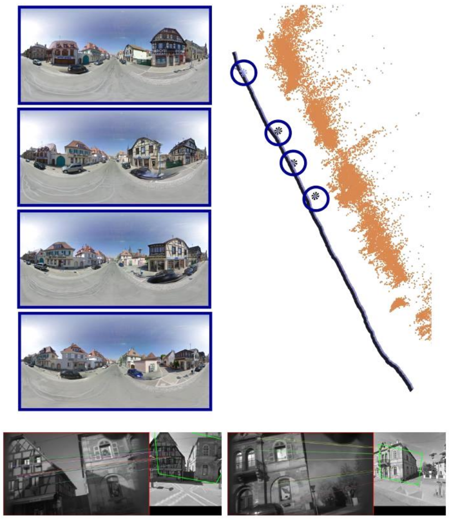
As a consequence, we require methods to determine the robot pose within the map frame based on the on-board sensors. In our research we considered different types of existing maps. First, GeoAutomation, our industry partner in the project, creates accurate maps of the environment based on sparse feature positions from camera images. To localize our robot within those maps we developed a laser-based method to track the robot within these maps. The main challenges here are that the used Velodyne scanner has a substantially different field of views and data densities and that the huge amount of data needs to be processed online. Our system is based on Monte-Carlo-Localization and can seamlessly switch between localization in GeoAutomation maps and 3D maps collected by the robot itself. The following image shows a Velodyne 3D scan overlay (blue) with the corresponding pose in the GeoAutomation map (black) of Zürich, estimated with our 3D localization module. As can be seen, the blue ring like sampling pattern of the Velodyne is denser than the GIS map, especially close to the robot.



In addition we also implemented a method to localize in the road network provided by Open-StreetMaps. This module also uses a modified Monte-Carlo localization. During the update step, the algorithm consider the robot odometry information and the IMU data to obtain a low variance prediction. This step is equivalent in spirit to the original MCL approach. During the update step, however, the information from the laser cannot be directly used, since no obstacle information is present in the map. To overcome this, we exploit the road information present in OpenStreetMap and use the current laser data to reconstruct the local road structure and compare it to the one stored in OpenStreeMap. The following figure (left) shows an example situation for the algorithm using a Velodyne scan classified as road (green) and non-road (red). To enable large scale planning we can then use the knowledge of the robot pose in the road network to plan large-scale trajectories for the robot. On the right of the following figure is an example trajectory planned like this.



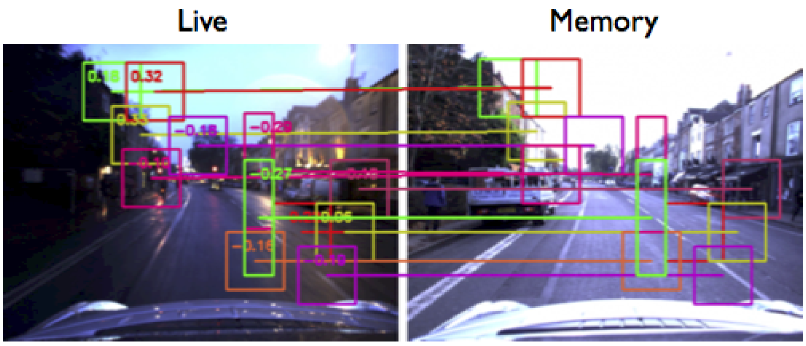
As a third option we exploited Google StreetView geotagged imagery as an accurate source of global positioning. To make this approach as general as possible, we only make use of a monocular camera and an odometry estimate. We model the problem as a non-linear least squares estimation in two phases. The first estimates the 3D position of tracked feature points from short camera sequences. The second computes the rigid body transformation between the panoramic Street View images and the estimated points. This approach can be considered as a complement to GPS systems that computes accurate positioning without an expensive or dedicated hardware setup. The following figure shows some elements of the localization in Google StreetView images. The top left shows example panoramic images from StreetView, the right extracted localization poses and estimated 3D points. The bottom two images show feature matching between the user's camera images and rectilinear views of the panorama.



Experience-based Navigation

Navigation is a key aspect for the operation of the mobile robot. In the EUROPA2 project, we aim for life-long operation in urban environments relying mainly on vision. Accordingly, we use a universal framework for large-scale, life-long localization despite changes in weather, lighting, season and scene structure. Experience-based Navigation (*EBN*) is a key tool in this direction.

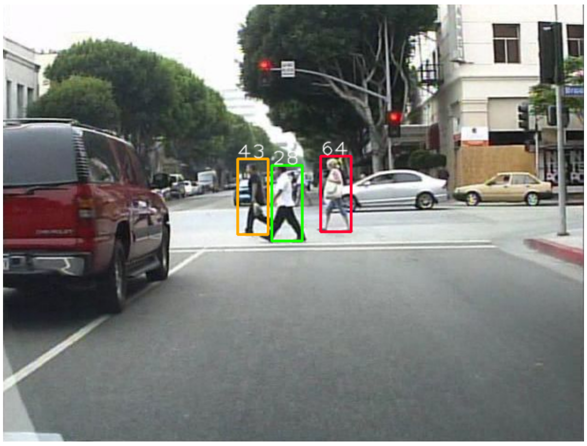
In *EBN*, the world is modeled using several representations of the same place, which are called experiences. Each experience represents the environment under particular conditions. The goal is to grow and curate a visual map of the world capable of capturing the full spectrum of appearance changes that characterize each place. The figure below shows an example matching. On the left is the image captured during the robot's operation. The image on the right is from a previously captured experience (memory). The different colored rectangles highlight matching features between the live image and the memory.



EBN makes use of stereo Visual Odometry (*VO*) to estimate the robot's motion while moving in the environment and to estimate the transformation between the robot and the experience saved in memory. Our implementation of VO operates on a stereo image pair to produce a set of observed 3D landmarks, where each landmark is defined relative to the coordinate frame of the camera. To estimate the robot's trajectory, VO calculates the 6DoF transformation between the current image and the previous image pair. A graph is built from the landmarks and the 6DoF transformations calculated by VO. This graph is referred to as the experience graph. When visiting a place for the first time, the output of the VO is saved and a new experience is created. When returning to the same place, the system will first try to localize against the experience; if localization fails, a new experience is created.

People Detection

Safety is a major concern in the operation of autonomous mobile robots. In the context of the EUROPA2 project, the robot will be operating in an urban environment which is characterized by the presence of pedestrians. As such it is crucial for the robot to be able to detect, and track people so as to avoid collisions during motion. We developed a robust detector with a strong focus on efficiency. The detector is able to process a video feed in real time while running on a single CPU core. Considering the energy constraints and the limited amount of computing resources available on a mobile platform, an efficient detector offers significant advantages over power-hungry GPU based implementations. Below is are example images of running the detector. Rectangles highlight the detected pedestrians.



The detector is based on the Aggregated Channel Features (*ACF*) framework. In ACF, image features can be considered as sums of pixels in fixed-size squares instead of in variable-sized rectangles. Additionally, ACF uses fast feature pyramids for detecting objects at multiple image scales: instead of calculating the image features explicitly at each scale, they are approximated via extrapolation from nearby scales. These two extensions offer a significant speed improvement over other feature-based frameworks, and they allow for a fast CPU based implementation.

Thank you for your interest in the EUROPA2-project. In the next newsletter we will provide information about more aspects of the project and will give updates regarding new developments.