Abstract

Vision-based robot navigation has long been a fundamental goal in both robotics and computer vision research. While the problem is largely solved for robots equipped with active range-finding devices, for a variety of reasons, the task still remains challenging for robots equipped only with vision sensors. Vision is an attractive sensor as it helps in the design of economically viable systems with simpler sensor limitations. It facilitates passive sensing of the environment and provides valuable semantic information about the scene that is unavailable to other sensors. Two popular paradigms have emerged to analyze this problem, namely Model-based and Model-free algorithms. Model-based approaches demand *apriori* model information to be made available in advance. In case of the latter, required $3D$ information is computed online. Model-free navigation paradigms have gained popularity over model-based approaches due to their simpler assumptions and wider applicability. This thesis discusses a new paradigm to vision-based navigation, namely Image-based navigation. The basic concept is that model-free paradigms involve an unnecessary intermediate depth computation, which is redundant for the purpose of navigation. Rather the motion instruction required to control the robot can be inferred directly from the acquired images. This approach is more attractive as the modeling of objects is now simply substituted by the memorization of views, which is far easier than 3D modeling.

In this thesis, a new image-based navigation architecture is developed, which facilitates online learning about the world by a robot (Chap. 2). The framework capacitates a robot to autonomously explore and navigate a variety of unknown environments, in a way that facilitates path planning and goal-oriented tasks, using visual maps that are contextually built in the process. It also facilitates the incorporation of feedback received from performing specific goal oriented tasks to update the visual representation. Based on this architecture, the design of the individual algorithms required for performing the navigation task (exploration, servoing and learning) is discussed.

In Chap. 3, a novel image-based exploration algorithm based on the idea of frontier-based exploration is proposed. The algorithm infers the frontier boundaries directly from monocular images and uses them to efficiently explore the environment. The frontiers are detected by using a modified segmentation scheme, that separates floor regions from non-floor regions. This method exploits the advantage of the natural structure in the world without involving any 3D reconstruction of the scene. The proposed algorithm can systematically discover an unknown environment and build a visual representation that is most suited for navigation.

Chap. 4 discusses the three major algorithms required for performing the navigation task, namely localization, planning and servoing. First, the idea of qualitative localization is introduced, wherein the localization only identifies the most similar image in the database to the current image. A planning algorithm is then used to infer the right set of intermediate images (from the visual representation) that would lead the robot from its initial position to the required destination. Finally a feed-back based control algorithm that exploits the projective transformations existing between the intermediate views is employed to servo the robot to the goal configuration. The overall algorithm only operates using images and converges to the destination reliably.

An online learning algorithm is described in Chap. 5. The algorithm utilizes additional scene information, gathered over time by the robot, to improve its visual representation. This is done by the way of updating the scene feature descriptors using an incremental learning algorithm. Learning is also employed to exploit the feedback received from previous experiences of the robot for improving the navigation performance. Specifically the task of path-planning for trajectory generation is discussed. An improved reinforcement learning scheme utilizing the potential field method to generate optimal motion trajectories is presented.

The basic advantage of the overall proposed framework is that the robot is no more-limited to a restrictive teach-and-replay scenario. It can now build an internal representation of its workspace autonomously and improve it automatically. Further, the navigation task can be accomplished more efficiently by simply exploiting the constraints existing between the images. The proposed approach enables mobile robots to progress in the direction of increased applicability and robustness. Experimental results on a laboratory set-up confirm the efficacy of the proposed algorithms.

Similar principles had been applied for solving the navigation/servoing problem in non-rigid environments during the earlier phase of this thesis and are presented in the Appendix (Chap. A and B).