On Environment Modeling for Visual Navigation

Imaging, Modeling and Representation of Real Scenes

Zhu Zhigang

Ph.D. Dissertation

Department of Computer Science and Technology

Tsinghua University

1997

Abstract

Visual navigation of a mobile robot in a natural environment has always been a very interesting but challenging problem. It involves almost every aspect of computer vision issues - from visual sensors, robust algorithms to visual representation. The fundamental tasks of visual navigation include global localization, road following and obstacle detection. Visual environment modeling is the foundation of visual navigation - and to an extension, most of the real world problems in computer vision. Three key components – smart visual sensing, robust visual algorithms and hierarchical visual representations are equally important in real scene modeling. This book is based on the author's Ph.D. thesis at Tsinghua University. The book presents a systems approach to visual modeling of a natural scene for robot navigation and in large to scene understanding and representation. There are four main contributions in the thesis:

1. A purposive, multi-scale and full-view visual scene modeling approach is proposed for visual navigation in a natural environment. It combines three novel modules into an integrated system, the **POST**: ***P**anoramic vision for scene modeling, **O**mnidirectional vision for road understanding and **ST**ereo vision for obstacle detection*. This approach tries to overcome the drawbacks of traditional visual navigation methods that have mostly depended on local and/or single view visual information. However, the proposed approach is not just a simple combination of the three novel sensors and methods, but rather a systematic integration under the strategy of purposive vision ("the right way for the right work"), and under the philosophy of systems approach which emphases that "the whole is more than sum of its components". Thus, right sensor design, adequate level of scene representation and corresponding robust algorithms are specifically explored for each given task while the interconnection among the vision sub-systems are taken into consideration under the overall goal of autonomous navigation.

2. A two-stage method is presented for 3D panoramic scene modeling from image

sequences captured by a video camera subject to vibration on a common road surface. First, a 3D image stabilization method is proposed which eliminates vibration from vehicle's motion so that "seamless" epipolar plane images (EPIs) and panoramic view images (PVIs) can be generated. Second, an efficient panoramic EPI analysis method is proposed to combine the advantages of both PVIs and EPIs efficiently in three important steps: locus orientation detection, motion boundary localization, and occlusion/resolution recovery. The two-stage method not only combines Zheng-Tsuji's PVI method with Baker-Bolle's EPI analysis, resulting in the so-called panoramic EPI method, but also generalizes them to handle image sequence subject to unpredictable camera vibration. Finally, a compact layered representation for a large-scale scene has been proposed, which can be used in both visual navigation and image-based rendering.. Since camera calibration, image segmentation, feature extraction and matching are avoided, all the proposed algorithms are fully automatic and rather general.

3. A new road following approach, the ***R**oad **O**mni-View **I**mage **N**eural **N**etworks* (**ROVINN**), has been proposed. It combines the omni-directional image sensing technique with neural networks in such a manner that the robot is capable to learn recognition and steering knowledge from the omnidirectional road images which in turn guarantee the robot never to get lost. The ROVINN approach brings Yagi's COPIS (conic omnidirectional projection image sensor) method to the outdoor road scene and provides an alternative solution different from the CMU's ALVINN. Compact and rotation-invariant image features are extracted by integrating the principle component analysis (PCA) and the Fourier transform (DFT). The modular neural networks of the ROVINN estimate road orientations more robustly and efficiently by classifying the roads as a first step, which enables the robot to adapt to various road types automatically.

4. A novel method called *Image Reprojection Transformation (***IRT***)* is presented for stereo-vision-based road obstacle detection. Obstacle detection is modeled as a reflexive behavior of detecting anything that is different from the planar road surface. Dynamic

reprojection transformation algorithms are developed so that the algorithms can work in rough road surface. The novelty of the (dynamic) reprojection transformation method, which ensembles the gaze control of the human vision, lies in the fact that it brings the road surface to zero disparity so that the feature extraction and matching procedures of the traditional stereo vision are avoided in the proposed obstacle detection algorithms. The progressive processing strategy from yes/no verification, focus of attention, to 3D measurement based on the reprojection transformation make the obstacle detection efficient, fast and robust.

To validate the proposed strategies and methods, we have implemented the following algorithms and systems.

(1) 3D **Scene modeling system**. In the 3D panoramic scene modeling, the algorithms of motion filtering and image stabilization, kinetic occlusion detection and depth layering, have been developed, and the novel depth layered setting (DLS) models have been constructed for many image sequences. These efforts found a ground base for landmark selection of global localization and image synthesis of photo-realistic image-based rendering.

(2) **Designs of novel sensors**. An omni-view image sensor is designed and realized, and its properties for the outdoor road understanding are thoroughly studied. A patented single camera binocular vision system is also designed and constructed, and has been put into real use of obstacle detection. An integrated full view smart sensor POST (Panoramic, Omnidirectional and STereo vision) is proposed, which integrates 360-degree omnidirectional view, binocular forward view as well as both left and right side views by using a single camera.

(3) **Real scene experiments.** Experimental results of training and testing the ROVINN using real road images have shown that the proposed method for road following is quite

promising. A real-time visual obstacle detection system has been set up and tested in the outdoor road scene extensively.

Index Terms: Visual navigation; Scene modeling; Image stabilization; Reprojection transformation; Panoramic vision; Omnidirectional Vision; Neural network; Smart sensors; Landmark localization; Road understanding; Obstacle detection.