Chapter 20

ROBOT LOCALIZATION USING VISION

Fredrik Linåker
Accenture Technology Labs
Sophia Antipolis, France *
fredrik.linaker@accenture.com

Masumi Ishikawa
Department of Brain Science and Engineering
Kyushu Institute of Technology, Japan
ishikawa@brain.kyutech.ac.jp

Abstract State-of-the-art vision-based techniques for robot localization are reviewed and a novel technique based on local computation is presented. The technique extracts compact feature vectors from omnidirectional camera images, allowing real-time processing of the image stream for accurate Sequential Monte Carlo localization. The proposed technique allows for a graceful degradation against noise as it is not based on global image features which easily can be disrupted by occlusions. The technique also seems to be a promising candidate for GPGPU implementation as it involves only simple local autocorrelations and a single integration step.

Keywords: appearance-based localization, omnidirectional vision, particle filters

1. Introduction

Robot localization in indoor environments, using long-range distance sensors like laser range finders [1], millimeter-wave radar [2] or sonars [3–5], is now generally considered as a solved problem. Localization using vision is however still an open problem. Besides being an interesting area

*This research was partially supported by the 21st Century COE Program of the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan.


www.springerlink.com © Springer-Verlag Berlin Heidelberg 2007
of research relating to neuroscience and cognition, vision as a primary sensor has a number of advantages. Cameras have a virtually unlimited range and can cover large field-of-views at high update rates. Due to the passive nature, multiple cameras do not interfere with each other when operating in the same area. Information like color and texture is readily available in the images, and camera systems are available at relatively low costs and have a limited power consumption.

In this article, which is an extended version of [6], we present a novel technique for image processing which enables a mobile robot equipped with an omnidirectional camera to perform localization in real-time. As the localization space is continuous, but computation is to take place on-line using limited resources, Sequential Monte Carlo (particle filters) are here used for estimating the approximate position and orientation of the robot. As shown, our vision-based localization system lets a mobile robot perform global localization and recover from kidnappings based on only image data and odometry readings.

2. Vision-Based Localization

3D-model based localization

A camera-based equivalent of a typical distance-based localization system would be to acquire a detailed three-dimensional model of the environment [7]. This 3D model can then be used during localization to generate the expected 2D projections (camera images) at different locations. Creating accurate expected camera images is difficult, as lighting, texturing, reflections, etc. all influence the outcome, but at the same time, only this approach has the potential to simulate arbitrary re-lighting and re-texturing of the environment.

The computational and resource costs can be reduced by keeping a less detailed map of the environment. Instead of being able to internally generate complete camera images, the system could settle for being able to predict what features a camera image would contain at different locations. Generally speaking, features found in the camera images can be considered as corresponding to structures in the 3D environment as the images are 2D projections thereof. As such, features can be 3D points or regions which are reliably detected from different viewpoints [8]. The features in this case make up a sort of sparse 3D map of the environment.

The downside of an approach using features anchored in the 3D-environment is that a single camera image does not contain all information necessary to fully match the features because the depth component is missing. Therefore, the 3D model based approaches regularly depend on auxiliary distance sensors like stereo [9] and trinocular [10] cameras.