Engineering Applications of Artificial Intelligence 29 (2014) 152–162

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

# Adaptive finite state machine based visual autonomous navigation system

Daniel O. Sales*, Diogo O. Correa, Leandro C. Fernandes, Denis F. Wolf, Fernando S. Osório

*University of São Paulo, Mobile Robotics Lab, Av. Trabalhador São-Carlense, 400, PO Box 668, São Carlos, Brazil*

ARTICLE INFO

*Article history:*
Received 23 December 2012
Received in revised form
8 August 2013
Accepted 10 December 2013
Available online 14 January 2014

*Keywords:*
Mobile robotics
Autonomous navigation
Kinect
Artificial neural networks
Finite state machines

ABSTRACT

In this paper we present an original approach applied to autonomous mobile robots navigation integrating localization and navigation using a topological map based on the proposed AFSM (adaptive finite state machine) technique. In this approach, the environment is mapped as a graph, and each possible path is represented by a sequence of states controlled by a FSM—finite state machine. An ANN (artificial neural network) is trained to recognize patterns on input data, where each pattern is associated to specific environment features or properties, consequently representing the present context/state of the FSM. When a new input pattern is recognized by the ANN (changing the current context), this allows the FSM to change to the next state and its associated action/behavior. The input features are related to specific local properties of the environment (obtained from sensors data), as for example, straight path, right and left turns, and intersections. This way, the FSM is integrated to a previously trained ANN, which acts as a key component recognizing and indicating the present state and the state changes, allowing the AFSM to select the current/correct action (local reactive behaviors) for each situation. The AFSM allows the mobile robot to autonomously follow a sequence of states/behaviors in order to reach a destination, first choosing an adequate local reactive behavior for each current state, and second detecting the changes in the current context/state, following a sequence of states/actions that codes the topological (global) path into the FSM (sequence of states/actions). The ANN is also a very important component of this system, since it can be trained/adapted to recognize a complex set of situations and state changes. In order to demonstrate the robustness of the proposed approach to different situations and sensors configurations, we evaluated the proposed approach for both indoor and outdoor environments, using a Pioneer P3-AT robot equipped with Kinect sensor for indoor environments, and an automated vehicle equipped with a standard RGB camera for urban roads environments. The proposed method was tested in different situations with success and demonstrated to be a promising approach to autonomous mobile robots control and navigation.

## 1. Introduction

The application of AI techniques in the development of mobile robots applications is a very important research topic. There are several possible applications of robust autonomous intelligent systems, from industry to civilian and military tasks. One of the most desirable features for a mobile robot is the autonomous driving capability. This capability is being studied since the 1980s, and groups such as the CMU NavLab, Stanford AI Lab (SAIL), Google Self-Driving Car Team, AutoNOMOS Labs, among others, have been presenting relevant results on autonomous vehicles navigation.

Nowadays there are many relevant and known researches on autonomous robotics being developed worldwide. Some of them are powered by government initiatives as for example the Darpa Grand Challenge (Thrun et al., 2006; Urmson et al., 2008; Buehler et al., 2007). The first two editions (2004 and 2005) were held in desert (Thrun et al., 2006; Buehler et al., 2007), and 2007 edition in an urban environment (Urmson et al., 2008; Buehler et al., 2009). In Brazil, a government initiative also focuses on the development of autonomous vehicles, being an important research challenge of the Brazilian National Institute of Science and Technology on Embedded Critical Systems (INCT-SEC) (Fernandes et al., 2012).

Autonomous mobile robots usually perform three main tasks: localization, mapping and navigation (Buehler et al., 2007, 2009; Luettel et al., 2012). The robot localization should be known in order to know the starting position and destination of the robot. Mapping is the creation of an environment model using the sensorial data or creating drawings/sketches of the environment, representing its structure. The localization task also occurs simultaneously to navigation control, and is related to robot's position estimation in a previously known environment, using the sensorial data. Navigation task is therefore the ability to obtain enough

---

* Corresponding author. Tel.: +55 1681271671.
*E-mail addresses:* dsales@icmc.usp.br (D.O. Sales), diogosoc@icmc.usp.br (D.O. Correa), lnd@icmc.usp.br (L.C. Fernandes), denis@icmc.usp.br (D.F. Wolf), fosorio@icmc.usp.br (F.S. Osório).

0952-1976/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.engappai.2013.12.006

information about the environment, process it, and act, moving safely through this environment, usually according to a predefined path.

Intelligent autonomous vehicles are able to navigate through environments composed by different sets of streets and/or highways. These roads define the “navigable areas”, that are allowed for the vehicle navigation. If the mobile robot′s approximate position, environment map and path to be followed are known, navigation in this environment lies in following a well-defined path, considering the navigable areas.

This work focuses on topological map localization and navigation tasks, describing the development of vision-based topological navigation systems for both indoor and outdoor applications. A 3D perception module based on a vision sensor (Kinect) was used for indoor environment, while a 2D perception module based on a traditional vision sensor (standard RGB camera) was used in urban environment tests. The developed systems were able to recognize the navigable area of the environments (corridors and streets) processing the input data and classifying them into states which represent the current robot context, allowing the robots to determine their localization into a topological map and also to autonomously navigate through the environment, reaching the desired destination.

Recently, some works have described and proposed (Luettel et al., 2012; Mueller et al., 2011) the use of hybrid approaches based on topological maps, landmarks detection and topological navigation applied to autonomous vehicles. This topological navigation approach does not require a very detailed environment map (metric map). The environment is mapped as a graph containing the main elements (topological nodes), in a simpler path representation. Furthermore, accurate pose estimation is also not really necessary. The mobile robot can estimate it′s approximate position and safely navigate choosing adequate reactive controls. So, the main task of this system is to detect the current node in a Topological Map, which allows to autonomously decide when and how to go straight, turn left or right, even when more than one possibility is detected simultaneously (at an intersection, for example).

An artificial neural network (ANN) (Haykin, 1998) is used to classify the input data obtained from the sensorial system, recognizing patterns on this data (representing the current context/state). An adaptive finite state machine (AFSM), integrating the ANN classifier with a traditional FSM, is used to identify the current state and to represent the state sequence which describes any path at the environment. The ANN is trained to recognize all possible states according to a specific application, so a FSM generator converts any predefined path into a sequence of well-known states to be recognized and followed using the AFSM.

The mobile robot motion control is based on a hybrid control approach. This way, the navigation system combines the high-level deliberative control (path planning) with different reactive behaviors (each state has its own set of associated behaviors), allowing a safe motion. For example, when the robot is following a straight forward road, the reactive behavior just keeps the mobile robot moving forward in the correct lane and restricted to the navigable/safe area. As the mobile robot moves along the road, there is a sequence of “local behaviors” defined by the FSM and selected according to each new situation identified, as for example, if there is an intersection, the robot selects a local behavior/action (e.g. crossing the intersection, turn to the left or turn to the right), according to the predefined path described by the FSM.

The initial ideas of the AFSM-based approach were already successfully applied in previous authors′ works (Sales et al., 2011, 2012a, 2012c) with different sets of sensors and states, showing the feasibility of this implementation and motivating studies concerning the application of this technique with vision-based sensorial systems. These previous works allowed to consolidate the proposition of this new approach, resulting in the AFSM approach as described in this paper.

The main objective of the applications presented in this paper is to demonstrate the AFSM vision-based navigation process, which starts detecting specific local properties of the environment (edges or discontinuities) characterizing the different types of environment features (context/state). The ANN classification results are then used with a high-level path planner (FSM sequencer) to select the local reactive behavior/actions, assuring safe motion by choosing an adequate reactive control for each different state.

The next sections of this paper are organized as follows: Section 2 presents some previous related work; Section 3 presents the Topological Navigation System overview, Section 4 presents the system setup for experiments and obtained results for indoor and outdoor environments, and Section 5 presents the conclusion and future work.

## 2. Related work

There are many different approaches for autonomous mobile robot navigation, each one using a different sensorial system (for example laser, sonar, GPS, IMU, compass). These sensors can be used solely or combined (Buehler et al., 2007; Siciliano and Khatib, 2008; Goebl et al., 2008). More recently, special attention has been done to vision-based navigation systems (Sales et al., 2010; Souza et al., 2013; Liu et al., 2013; Bonin-Font et al., 2008; Bertozzi et al., 2000). These systems use video cameras as the main sensor in order to navigate and avoid obstacles. There are several different types of cameras as monocular, stereo, omnidirectional and thermal (Infrared), for example. Cameras are very suitable for navigation and obstacle avoidance tasks for all sizes of robots due to its low weight and energy consumption (Zingg et al., 2010). Furthermore, one single image can provide different types of information about the environment simultaneously. It is also possible to reduce costs by using cameras rather than other types of sensors (Scaramuzza and Siegwart, 2008).

Vision-based navigation approaches are already usual in navigation systems for structured or semi-structured environments (Shinzato and Wolf, 2010; Nefian and Bradski, 2006; Álvarez et al., 2008). These systems allow classifying the image elements, with pathway segmentation for safe navigable area identification, resulting in reactive models for navigation control. Works such as ALVINN (Pomerleau, 1989) and RALPH (Pomerleau, 1995) were some of the first to apply Artificial Neural Networks for reactive control in outdoor environments.

The robot navigation through a path represented by sequences of images is not new, once we found in the literature different references to visual navigation methods for mobile robots (Jones et al., 1997), as for example, the work developed by Matsumoto et al. (1996, 2003), known as the VSRR (View-Sequenced Route Representation). However, this approach is based mainly in the correlation of pairs of images, being limited to the application of this image matching algorithm (with the implicit limitations and restrictions of this method) in order to determine the present location of the robot, and then applying a simple set of direct actions (e.g. forward, turn left turn right). Other works proposed by Kortenkamp et al. (1998), Huber and Kortenkamp (1998), are also based on the idea that multiple proximity spaces can be chained together using kinematic models, where each proximity space is controlled by a set of behaviors that combine to determine its location. Although this is an interesting proposition, integrating proximity spaces with associated local actions, this method requires special sensors (stereo camera), and also works with the