uc3m | Universidad Carlos III de Madrid

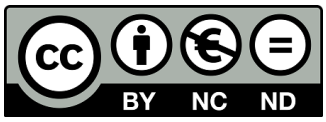
Master Degree in Telecommunication Engineering
Academic Year (e.g. 2019-2020)

Master Thesis

"Embedded solution for person identification and tracking with a robot"

Ignacio Condés Menchén

Fernando Díaz de María
Eduardo Perdices García
Leganés, pending date



SUMMARY

This project describes the development process of an embedded system capable of performing a reactive following of a person. It makes use of convolutional neural networks and probabilistic tracking for processing the perception acquired by a RGB-D camera. This input is processed in a NVIDIA Jetson TX2, an embedded System-on-Module (SoM). This device is capable of performing computationally demanding tasks onboard, coping with the complexity required to run a robust tracking and following algorithm. The full design is implemented on a robotic mobile base, which receives velocity commands from the board, intended to move towards the desired person.

Keywords: deep learning, robotics, person following

DEDICATION

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Motivation

This work is focused on exploring the synergy between two science fields, which are outstanding nowadays: *robotics* and *deep learning*. These are combined for obtaining a robust system capable of following a certain person navigating towards it on a reactive behavioral. This behavioral is composed of two main components: the *perception block*, responsible of processing the images from an RGB-D sensor placed on the system, and the *actuation block*, which moves the robotic base accordingly to the relative position of the person to be followed.

The original idea was proposed on [1], where a neural following system was developed to be run in a standard laptop into which the camera and the robot were plugged. In the following dissertation, we will revisit this work and describe the points of interest which have allowed to enhance the previous version of this work.

The key aspects of this project can be brought in as follows:

Embedded solution the system is mounted on a battery-powered robot, on a *mobile base* form factor. This robot features a high-performance GPU embedded on a System-on-Module. Thus, this ensemble can work on its own, without requiring an external computer to perform the inferences or running algorithms in parallel. A remote monitoring of the behavioral is available as well, but it is not required for the system to work.

Person identification the proposed system runs 3 neural networks. These networks perform inferences over the images perceived by the RGB-D sensor, which is attached to the system as the *point-of-view* of the robot. The inferences are devoted to detect the different persons present in the scene, as well as to distinguish them by means of an identity sign: their face.

Tracking the full system includes a probabilistic tracker, based on dynamic modeling. This leverages the *trajectories* followed by the persons while they wander on the visual field of the robot, as well as the relative distance to the person, obtained by the depth sensor included in the camera. As a result, we can have a gain on the robustness of the system, compared to a version governed exclusively by the neural inferences, which are sensitive to visual occlusions. Trusting just on these inferences could easily result on an unsteady behavioral. However, this can be avoided introducing the probabilistic modeling, as it will be explained later.

1.2. State of the art

TODO

1.3. Objectives

This work has been carried out in order to fulfill certain requirements in a particular person following application:

1. Achieve a real-time following behavioral using embedded low-power hardware and a low-complexity educational robot.
2. Build the inference pipeline using exclusively concurrent CNNs (*convolutional neural networks*).
3. Combine a neural system with probabilistic filtering to carry out a robust multimodal tracking of the persons in front of the robot. This will provide the system with extra endurance and robustness against detection losses/occlusions.

These objectives allow to summarize the starting point for the development of this project: the available materials are an educational robot equipped with a battery, an embedded *SoM* and a RGB-D sensor.

The result will be an autonomous robot which will follow a specific person, whose face has to be known beforehand (using a *reference face* image).

2. MATERIALS AND METHODS

This chapter is devoted to describe the process followed to develop the system.

BIBLIOGRAPHY

[1] I. Condés and J. Cañas, "Person Following Robot Behavior Using Deep Learning: Proceedings of the 19th International Workshop of Physical Agents (WAF 2018), November 22-23, 2018, Madrid, Spain," in. Jan. 2019, pp. 147–161. DOI: 10.1007/978-3-319-99885-5_11.