Individual Project – CS3010

Interactions with Pepper robot with voice and object Recognition

Project Definition

The concept for this project is to create a software module that allows users to interact with the Softbank’s Aldebaran robot Pepper using speech and object recognition. The basis for this project stems from functionality requirements that my software module should enable a user to perform with the robot.

For example the Pepper robot should be able to recognise a user’s face and that user should be able to issue it commands. In addition, the robot should be able to recognise objects and state the given name of said object when asked. These are the forming foundations of my individual project and the requirements that have been set.

Functional Requirements: TBD

-

-

-

-

Acknowledgements

List of contents

Abstract/executive Summary

Abstract— Object and speech recognition technology is imperative to the software’s application, however the modules that are provided by Softbank’s Pepper are limited in terms of what users can do when interacting with the robot.

This paper uses a combination of Choreographe and Python SDK to create a dynamic software modules that allow users to talk to the robot and enable the user to request the Pepper to identify objects within its vision.

TBD….

3.2 Project Process & Professionalism

requirements specification, use case analysis, domain model, class diagram, sequence diagrams, storyboard, navigation structure, and test plan.

1. Introduction

The motivation behind this individual project is based on the fact that robotic technology is forever advancing at an alarming rate, human interaction with robots is a relatively new field as technology has only recently been able to come to a level where implementation is possible. The concept of a user interacting with a robot based on speech and voice recognition is a potentially powerful tool that can be used in a wide range of areas.

This project aims to look at work and research done previously from other academic literature and incorporate various parts of each practically and theoretically in order to achieve the project definition.

This relates to the broad area of computer science, as computer robotics and artificial intelligence is an area that has much to be explored, and has been looked at recently over the years with exponentially growing expectations. Many real world applications already utilise robots with AI for various jobs such as factory manufacturing of vehicles, medical diagnoses the potential seems endless.

This project aims to hopefully open an avenue of insight to how effective robotics and artificial intelligence can be, with real world uses through creating a software module that allows the Pepper robot to interact with human based on speech and object recognition.

In section 3 we will provide tests for speech recognition, success rates and whether the software module created achieves the set functionalities set out as well as object recognition success rates and how well Pepper is able to identify objects.

Background

[1]The Pepper humanoid robot in front desk application, Arkadiusz Gardecki, Ryszard Beniak ,(2018) , [2018 Progress in Applied Electrical Engineering (PAEE)](https://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=8422067)

This paper proposes using the Pepper robot in a real world application of using the robot as a receptionist at a desk, by creating and extending software modules that have been built into the robot. By using the robot’s vision detection and speech recognition technology to interact with humans.

How this relates? Uses robot’s speech and object recognition to interact with humans performing tasks

# [2]Kinematics Modeling and Trajectory Planning for NAO Robot Object Grasping based on Image Analysis, Yuhan Jiang, Faliang Chang, Fuxin Liang , (2018) 2017 Chinese Automation Congress (CAC)

This paper looks at creating a new way of allowing the NAO robot to calculate distance and trajectory to allow the robot to grasp objects.

This relates? How—Uses robots vision and extends software module to be able to achieve a new functionality.

# [3]Autonomous Task Execution within NAO Robot Scouting Mission Framework, Anja Babić, Nikola Jagodin and Zdenko Kovačić, (2017) 2017 European Conference on Mobile Robots (ECMR) How this relates? It makes use of extension of software modules which in conjunction with robots cameras to move the robot around and complete its objective.

[4]Detecting and tracking objects in HRI: YOLO networks for the NAO “I See You” function, Jiahuan Zhou, Lihang Feng, Member, IEEE, Ryad Chellali, Member, IEEE and Haonan Zhu, (2018), Proceedings of the 27th IEEE International Symposium on Robot and Human Interactive Communication, Nanjing, China, August 27-31, 2018.

How this relates? Makes use of an object detection framework to identify objects.

There has been much work relating to my project done already, for example [2] making use of Pepper’s cameras and extending a software module based on kinetic modelling to enable the robot to detect objects and calculate the trajectory to grasp the object. This area of research can be used to supplement my project’s concept of being able to scan a table of objects be able to point to it. Based on the concepts and use of their bounding box algorithm. It is highly likely that this research will be critical to the development of the project in achieving specific functional requirements.

The results of this paper showcase that using kinetic modelling for grasping of objects via image analysis can be very effective due to the experimentation and results given. However, success rate will decline the smaller the object is therefore one thing that must be taken into account when adopting this paper’s methods is how this will affect the definition of the requirements or achieve them E.G. what kind of objects will the software module be able to identify and point to? Should the objects be of a specific range of sizes, or should the robot be able to identify and locate any object of any size etc.

Additionally, another noteworthy paper [4] which uses a framework called “YOLO” (You only look once) that is able to identify objects based on a weighting and confidence level, which I can use to implement into my software module allowing Pepper to identify a wide range of objects. This in conjunction with [2] may allow me to achieve the projects objective of scanning for an object and being able to identify it and point to it using the kinetic modelling method [2], and instead of the image analysis use YOLO framework [4]. Which is one of the definitive functionalities I have set for this project.

The main aspects that should be utilised and taken into account in developing the deliverable is how they made use of the YOLO network framework for their given application. In this case it was made so that the robot was able to detect and track objects within its environment. This research indicates that it can be used in the development of the project and help shape the functional requirements.

As, results from this paper show that the robot is able to localize and track an object within its vision. However, it is limited by many factors namely – being unable to localize(identify) small and adjacent objects as well as environments where there are multiple objects confidence detection declines.

[1] Also adds relevance as it emphasises the kind of real world job application that Pepper can be used in depending on the software module created. Indicating my claim that research into this area of computer science has relevant significance, and my project aims to do something similar by showing how humans are able to interact with Pepper based on the set functionalities I have defined.

Another academic piece of research [3] shows that you can create a software module for autonomous task execution for a robot based purely on a navigation and the robot’s cameras. Which is similar to my project as I will have to issue the robot a command namely, look for and Identify a given object i.e. cup etc.

All of the research papers within this topic area and scope are heavily relied on for the basis of future developments allowing other

Lots of research within this area is indicative of the future prevalence that robotic artificial intelligence will have in modern applications in the future. However, my project differs from previous work as it will make use of a combination of different theoretical ideas and practical implementations done by other papers to be able to achieve my project definition and allow users to interact with robots in a simple straight-forward way.

REVIEW ALL THE PAPERS IN THIS SECTION I.E. TALK ABOUT IT WHAT IT IS, THEN LOOK AT EXPERIMENTS AND RESULTS—WHAT DO THEY MEAN—AND WHAT RELEVANCE DOES IT HAVE?

Preparation

There was a need to look at previous work, and analyse what approaches they took to achieve their specific goals, and take the relevant parts of those papers and incorporate some of their ideas into mine. Additionally, learnt the hardware aspects of the Pepper robot such as what kind of hardware it has that I could utilise, in this case the eye cameras for object detection and microphones for sound input (speech detection).

Additionally, I had to learn how to use Choreographe which is a software application that Softbank’s provide with their robots that allow users to program the robot with a 10 unit module tutorial, which taught all the basic aspects of coding and using the robot.

The entirety of my requirements analysis Involved looking at the project definition defined, and what my supervisor has advised in terms of what is expected by my software project and the sort of functionalities it should provide for users as well as many project management factors that had to be taken into account i.e. hardware factors, time-constraints, complexity etc.

There was much refinement needed before the actual implementation of the software project itself, as functionalities of the project definition of what the software project should be able to do changed overtime as realisation of what was and, was not feasible occurred. For example, tracking and moving the robot towards an object in the environment is not impossible but would prove difficult due to the robot’s hardware limits (infrared sensors lacked precision and accuracy) so requirements were adjusted accordingly.

In addition taking into account the time-constraint, shaped what sort of requirements that was set for the project due to the fact that if overly-ambitious and complex functionalities being chosen may hinder the development of the project entirely.

Later learnt how to code the software module in a python SDK which is a lot more advanced and difficult that Choreographe, but offers much more flexibility and potential to what is possible. Which allowed me to change my requi

Also, had to understand how to structure the software project in order to keep it maintained and manageable. Because Python is a new language to me, and it is not a typical OOP language like Java, I had to research into what kind of system architectures were available to use.