2023 年臺灣國際科學展覽會
優勝作品專輯

作品編號	190031
參展科別	電腦科學與資訊工程
作品名稱	Automatic Destination Coordinating Robot based on Openvino
得獎獎項	

國　　家	Macau
就讀學校	Puiching Middle School, Macao
指導教師	Thomas Lao
作者姓名	Han Pok Man

關鍵詞　automation、navigation、unfixed

作者照片

1. Abstract

In this project, we created a function integrated onto a Lingao Chassis that allows the robot to use Slam and Gmapping to successfully navigate its way to the most convenient destination for the user, while avoiding any obstacles on the way, improving the default Gmapping errors.

2. Introduction

In the past few years, I have experimented with navigation related functions on my robot numerous times. To achieve these functions, I use the Simultaneous Localization and Mapping system, which allows the user to enter a set of coordinates, in which the robot will navigate to. However, as my functions become more and more sophisticated, SLAM's ability has begun to limit the robot's potential, as some functions require the robot to navigate to a non-preset destination, perhaps to navigate to a person whose position is not fixed.

2.1 Current Solutions

Slam(Simultaneous Localization and Mapping) is a technique for robots to scan their surroundings into a map. Then, we can use GMapping to allow the robot to navigate using coordinates offered on the map. However, this creates the problem of the robot not being able to find the best locations, only focusing on the destination the user set manually.

3. Roles

In a nutshell, a robot's roles in society can be summarised into two key factors: automation and assistance.

3.1 Automation

For automation, our robot automates the need to manually remove an obstacle on the user's designated point. As previously mentioned, when this function isn't being used, the robot is unable to avoid obstacles that weren't there during the Slamming process.

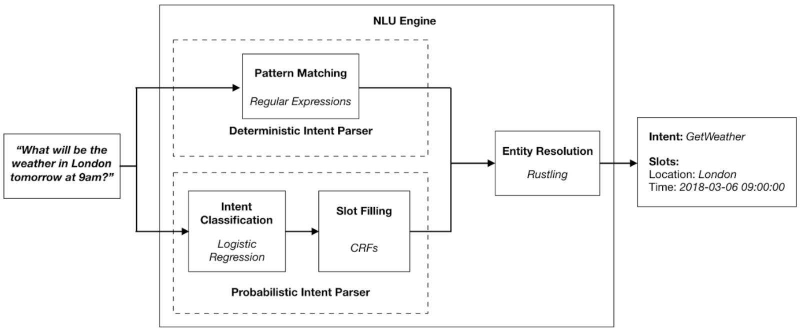
3.2 Assistance

In terms of assistance, our robot is able to assist those who have disabilities or difficulties with reaching or obtaining items. For example, if an elder person wanted to grab an item while sitting on the sofa, this function assists the person by eliminating the need to get up and manually get the item.

4. Software

4.1 Usage of Snips-NLU

In classic text-to-speech, we often use the searching of keywords to determine the user's intentions. However, this brings problems such as the robot's inability to function when a pre-set keyword isn't present. Thus, the Snips-NLU engine was invented to be able to recognize the user's intent without inputting keywords.



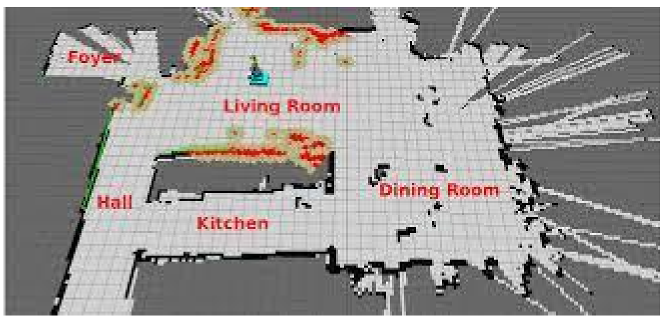
In our function, we created a small dataset that allows the robot to introduce itself, as well as recognize whether the user needs the robot to navigate to him or not.

4.2 OpenVino

Although deep learning is a proven fast way to allow computers to classify or predict things, models trained using deep learning often require a high-end GPU to run. Therefore, Intel released Openvino, a way of running inference on models using CPUs, while achieving the same results.
In this function, we used Openvino's face-detection-adas-0001 model to locate the user, head-pose-estimator to find the best destination, and person-pose-estimator to determine whether the user needs the robot to navigate to him.

4.3 SLAM (Simultaneous Localization and Mapping) with G-Mapping

In our function, the chassis is a very important technique. In order to achieve functions that involve the usage of the robot chassis, Slam and G-Mapping are an unavoidable technique. As can be seen on the right, G-Mapping allows us to scan a map of the robot's surroundings. Using SLAM, we can now allow the robot to navigate around in points that are on the map. SLAM is crucial in our function because it allows us to calculate the user's position and also navigate to the closest destination.

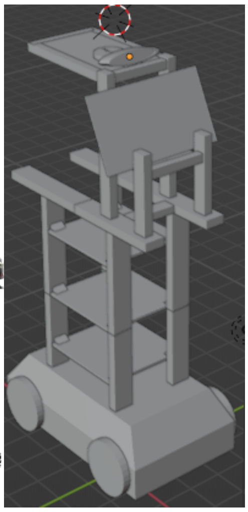


5. Hardware

5.1 Development History

Compared to last year, this year, we created a platform on top of a compartment for the robotic arm(for the respeaker, so that the function could be activated in more ways compared to the older way of using hand raising detection.

5.1.1 Previous Design Vs Current Design

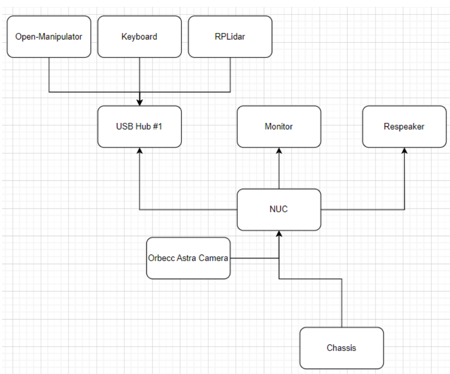


In the older version of our robot, we had a compartment for an astra camera and atop of it was a platform for a respeaker. However, we realised that this design has two big flaws.

Firstly, we realised that our previous design had a lack of space for the robotic arm to store items, therefore, in our second design, we created a compartment for our robotic arm behind the monitor.

Furthermore, we encountered a problem with the previous astra camera. Since the previous Astra camera had its own operating system, the files from the camera were being mixed with the files in the NUC, causing a lot of problems for us. Therefore, we created a platform that is able to both store the respeaker and hold the camera.

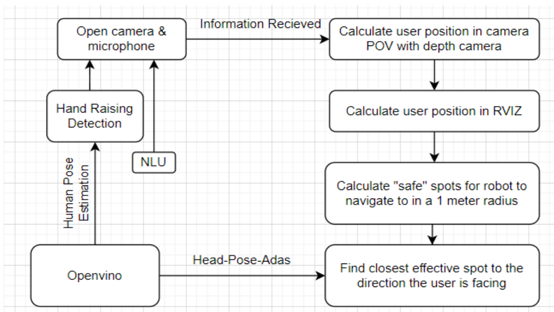
5.2 Cable Management



In our previous years of attending robotic competitions with similar robots, cable management has always been a problem for us, as they are usually messy. Thus, we created a map for a more clear view of the cables.

6. Methodology

The process of the function can be split into three parts, the activation system, the calculation system, and the navigation system.



6.1 The Activation System

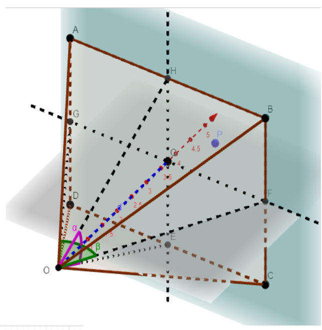
To make the function more accessible, I created two methods to activate the function. First, I imported the Openvino model “human-pose-estimator” to calculate the user’s posture. If the user is raising his/her hand, the model will detect it and activate the function

Apart from activation using visual sensors, the function can also be activated via voice recognition. If the user wants to summon the robot, the user can just tell the robot, and the robot will then use Snips-NLU to recognize the intent of the user. If the intent is to summon the robot, the robot will initiate the next step, which is to calculate the user.

6.2 The Calculation System

To calculate the user’s coordinates, the robot has to do two steps, being to calculate the person’s coordinates within the camera’s perspective, and to add the aforementioned coordinates to the robot’s coordinates in RVIZ.

To calculate the first step, we first used the formula to the right to calculate the person’s coordinates in the camera’s perspective, mainly using some trigonometry-based calculations and information from the robot’s depth camera.



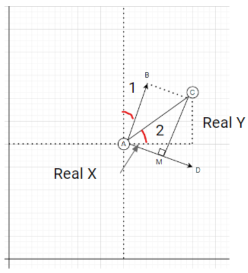
In photo:
AB = CD = w'
AD = BC = h'
P = (x' , y')

In world:
AB = CD = w
AD = BC = h
P = (x , y)
OQ = d

$$\because \frac{y}{h} = \frac{y'}{h'}$$
$$\Rightarrow h = 2 \cdot d \cdot tan\frac{\alpha}{2}$$
$$\therefore y = \frac{y'}{h'} \cdot 2 \cdot d \cdot tan\frac{\alpha}{2}$$

$$\because \frac{x}{w} = \frac{x'}{w'}$$
$$\Rightarrow w = 2 \cdot d \cdot tan\frac{\beta}{2}$$
$$\therefore x = \frac{x'}{w'} \cdot 2 \cdot d \cdot tan\frac{\beta}{2}$$

To convert the information we get from the first step, we used the formula on the right, which consists of some trigonometry, to calculate the Real_X and Real_Y in the Image. By adding this to the robot’s real coordinates, we can calculate the person’s coordinates.




```
Known: ∠1, CM, AM, AC
Therefore，
Tan-1 (AM / MC) = ∠ACM

Since the X axis is perpendicular to the Y axis，
Therefore ∠1 + ∠BAC + ∠2 = 90

SinceBA // CM，
Therefore∠BAC = ∠ACM

So ∠2 = 90 - ∠1 - ∠BAC

Thus，
Real_Y = sin(∠2) * AC
Real_X = cos(∠2) * AC

Real_X, Real_Y += Robot Coordinates
```

7. Research Data

During the creation of this function, we collected data on several different aspects of software to find the most efficient and optimal things to use on our model.

7.1 Proving the problem

When developing the navigation section, we realised that we couldn’t tell the robot to go directly to the user’s coordinates, as SLAM would recognize the user as an obstacle. To prove that this wasn’t a one-time bug, we conducted tests on the Slam function implemented in the chassis. As a result, we found out that while all tests with no obstacles succeeded, only one test out of those who had an obstacle succeeded, proving that our objective was, in fact, not a bug.

Test Number	With obstacle	Success?
1	No	Success
2	Yes	Failed
3	Yes	Failed
4	No	Success
5	Yes	Success
6	No	Success
7	Yes	Failed

7.2 Finding the best Face Detection Model

Face detection is seen as a beginner project in the deep-learning journey. Therefore, a vast

Model Name	Consistent Time tracking face
CvLib	~ 2 seconds
HaarCascades	~ 1 seconds
DLib	~ 7 seconds
Openvino Face Detection Adas 0001	~ 10 seconds

amount of camera-related libraries have developed their own face detection model. In order to find the most efficient model for our function, we used these models to track our faces, and found the model that could track a face continuously. At the end, we realised that the best model to use for our function was Openvino's Face-Detection-Adas-0001 model, able to track a face for roughly 10 seconds.

8. Future

8.1 Robotic Dog

Although our robot is able to navigate flat terrains, if we were to use our robot in real-life scenarios, the robot would have to be able to navigate rougher terrain. Our laboratory has recently bought a robotic dog from Unitree. In the future, we hope to implement our function onto the robot dog so that it would be able to help more people and navigate through rough terrains, ultimately benefiting humanity.

8.2 3D-LiDar

Currently, we use a 2D-Lidar for G-Mapping. However, 2D-Lidars are not able to map hollow objects such as tables and chairs. In the future, we hope to use 3D-Lidar to combat these problems, so that our robot will be able to navigate more efficiently without us having to manually cover the bottoms of chairs during the G-Mapping process.

9. Conclusion

In general, the point of robots is to assist, automate. We believe that this function can not only assist us and other robotics enjoyers in their future studies, but also to automate the task of manually removing obstacles so that the robot will be able to navigate its way to the destination.

【評語】190031

This project proposes an algorithm to navigate a robot while avoiding obstacles on the way. The report is described clearly. The methodology is detailed. Some comments are given below.

The topic is not a new one. It is suggested to have a literature survey for the existing works and a comparison.

More experiments are suggested for performance evaluation.