MOBILE PLATFORM MOTION CONTROL SYSTEM BASED ON HUMAN GESTURES

Rui Vilaça¹, Jorge Ramos², Vinicius Silva³, João Sepúlveda⁴, João Sena Esteves⁵
¹,² Department of Industrial Electronics, University of Minho, Guimarães, Portugal
³,⁴,⁵ R&D Centre Algoritmi, University of Minho, Guimarães, Portugal
a71431@alunos.uminho.pt, a71681@alunos.uminho.pt
a65312@alunos.uminho.pt, mjs@dei.uminho.pt, sena@dei.uminho.pt

Abstract - Controllers based on human gestures recognition may be a suitable and user-friendly solution for the interaction between machines and humans. The academic research on this topic may benefit the industry and the society in general, deserving a proper support from governmental institutions. A challenging field of application of such systems is the control of mobile platforms. Traditional mobile platform controllers may be limiting regarding freedom of movements and are inadequate in some applications. Mobile platforms with four Mecanum wheels allow the improvement of pre-existing solutions, such as wheelchairs or forklifts, endowing them with higher manoeuvrability, but they require a complex motion control because of the wide variety of movements allowed. The control of all possible movements of such a platform cannot be achieved with a simple joystick or steering wheel. This paper describes a motion control system for a mobile platform with four Mecanum wheels. The system is based on hand tracking and gestures recognition using an Intel RealSense 3D sensor.

Keywords: Mobile Robotics, Human-Robot Interaction, Intel RealSense 3D, Computer Vision, Mecanum Wheels.

1. Introduction

Robotics is an area of considerable interest in the scientific community, especially in Engineering. Although their great impact in the Industry, the first robots could not move away from the workplace, which was a limitation in some situations [1]. In order to overcome this limitation, different solutions have been developed, allowing robots to travel inside buildings and/or between them.

The most common and intuitive ways of controlling the movements of mobile robots use pedals, steering wheels or joysticks, among others [2]. These solutions are very robust and reliable but not generic. Mecanum wheeled platforms are omnidirectional, which is defined by the ability to move instantaneously in any direction from any configuration [3].

Because of this ability, those who are used to conventional mobile platforms may find omnidirectional platforms somewhat hard to control because some of the movements are not intuitive for them.

Rotating over itself or making any translational movement are examples of such movements.

Traditional controllers, based on joysticks or steering wheels, among others, may not be well suited for this type of platforms because: (i) the controller may not have all control elements required in order to produce all the movements allowed by the platform; (ii) a controller enabling the use of all steering capabilities of the platform may be complicated to operate.

New attempts to solve the control problem were tried, including the use of human gestures as a remote control. To do the tracking of the human body, solutions with Inertial Movement Units (IMU) or cameras were tried [4][5].

This paper describes a motion control system for a mobile platform with four Mecanum wheels based on human gestures, using Intel RealSense 3D F200 as a tracking device. The system is being developed in the context of Project II, a curricular unit from the Integrated Master on Industrial Electronics and Computers, at the University of Minho (Guimarães, Portugal).

The remainder of the paper is organized as follows: Section 2 and Section 3 present the theoretical background related to Mecanum wheels and Human Gestures Recognition. Section 4 describes the work that has been developed. Some preliminary results are presented in Section 5.
Section 6 presents the conclusions and some ideas regarding future developments of the work.

2. Mecanum wheel

The Mecanum wheel, invented in 1973 by the Swedish engineer Bengt Ilon, consists of a set of rolls symmetrically placed around the wheel body. The angle the axis of each roll makes with the rotation axis of the wheel is, usually, ±45° [6][7] (Figure 1).

![Mecanum Wheel](image)

Figure 1: Mecanum Wheel

In a two-dimension space, a body has three Degrees of Freedom (DOF), being able to perform translation movements in any direction and rotate over itself. However, most mobile platforms, built with conventional wheels, do not have full and independent control over the three DOFs. These platforms can produce a wide variety of trajectories in a two-dimensional space but they do not allow all movements, such as moving in the direction of the rotation axis of the wheels [7].

Omnidirectional platforms present advantages over conventional ones since they can produce all possible trajectories in a two-dimensional space. Figure 2 shows some of the movements that a mobile platform with four Mecanum wheels can produce.

Although they may present advantages over conventional platforms, four wheeled Mecanum platforms, like the one used in this work, have a major limitation over the type of surfaces they can work. When the wheel is rotating, at most two rollers and least one roller are/is in contact with the ground. This causes a small area of contact between the wheel and the ground, making the wheels susceptible to slip in [8][3].

![Vehicle motion according to the direction](image)

Figure 2: Vehicle motion according to the direction [7]

These platforms are not a generic solution and face some limitations but still have good applications on the world, such as:

(i) in the military field, the high manoeuvrability is suitable for rescue missions, military activities, planetary explorations, mine operations, among others – NASA developed OmniBot for exploration purposes [7];

(ii) in the industry, once again, the high manoeuvrability allows industrial forklifts to move through narrow aisle or standard doors – Airtrax Sidewinder is an example of a forklift equipped with a four wheeled Mecanum platform [9];

(iii) in the medical field, powered wheelchairs are proven to be a plus for handicapped people, providing them a more independent lifestyle – four wheeled Mecanum wheelchairs facilitates the locomotion of these people [10];

(iv) in the educational field, promote the study of mathematic, mechanic and electronic. One of the many examples is Uranus, a four wheeled Mecanum robot, developed in Carnegie Melon University [11][7].

3. Human Gestures Recognition

Gestures recognition is a subject of great interest, since it allows a more natural interaction between the human and the machine, by only using our body, being defined as "the tracking and interpretation as semantically meaningful movement" [4].

Approaches with cameras and IMUs were already explored, both with advantages and disadvantages. IMUs, which are usually employed in physical controllers, are a reliable solution, however, it is necessary to carry the controller all the time, being an intrusive solution and, sometimes, not practical. Some examples of such devices are JPL BioSleeve and the Acceleglove [4][12].

Regarding cameras, computer vision is a subject of major importance. It can be applied to a wide variety of situations and improve some of the solutions already implemented, leaving the text-based ones behind. Computer vision can be applied to various domains, such as, Virtual Reality, Robotics, Desktop and Tablet applications, Games, Sign Languages, among others [13]. Solutions based on ordinary webcams [14], infrared sensors (Leap Motion) [15] or depth sensors (Kinect and RealSense 3D) [16][17] were already implemented in the literature. Cameras represent a less intrusive solution, since the user is completely equipment free but, once data is obtained from the cameras, the quality and reliability of this data have a great dependency over the environment light, the capture distance/perspective and even the camera quality.
4. Work Developed

In this chapter, it is described all the software developed, as well as, the mathematical models and hardware used in the present work.

1) Hardware

Since the main purpose of this work is to recognize hand gestures to later control a mobile platform, the Intel RealSense 3D camera (model F200) was chosen. Compared to the Kinect camera and its Software Development Kit (SDK), the Intel RealSense camera is smaller and its SDK has a built-in model of the human hand. When compared to Leap Motion, the RealSense offers a longer range (150cm) to detect and analyse the hand [17].

![Figure 3: Mobile platform with four Mecanum wheels fully assembled](image)

The platform has four infrared sensos built-in that are used to prevent the platform to shock with objects, providing some security to the user and for the surrounding people. The user cannot overcome the command from the sensors. If the platform is at 20cm of an object, all four motors stops and that direction is no longer allowed.

About the direction of rotation of each motor that allows it to perform a specific trajectory, it is applied the following matrix (eq. 1), which can be defined as the kinematic matrix of a mobile platform with four Mecanum wheels:

\[
\begin{bmatrix}
  v1 \\
  v2 \\
  v3 \\
  v4
\end{bmatrix}
= \begin{bmatrix}
  1 & 1 & -\frac{C+L}{2} \\
  -1 & 1 & \frac{C+L}{2} \\
  -1 & 1 & \frac{C+L}{2} \\
  1 & 1 & \frac{C+L}{2}
\end{bmatrix}
\begin{bmatrix}
  Vx \\
  Vy \\
  Wz
\end{bmatrix}
\]  

(1)

In the respective matrix, \(v1, v2, v3\) and \(v4\) represents the linear velocity of each motor; \(Vx\), the linear velocity of the platform in X axis; \(Vy\), the linear velocity of the platform in Y axis; \(Wz\), the angular velocity of the platform in Z axis; \(C\) the distance between wheel 1 and 3 or 2 and 4 (Figure 2); \(L\) the distance between wheel 1 and 2 or 3 and 4 (Figure 2) [18][19].

To ensure a fluid movement, it is used a PID controller to assure that each wheel rotates at the applied speed and that sudden changes in the wheels speed are not perceived.

In Figure 5 is illustrated the wheel velocity reference (setpoint), the measured wheel velocity (measured) and the error (setpoint – measured). The PID algorithm outputs are computed for each wheel every 10ms with \(Kp = 2.5\), \(Ki = 0.3\) and \(Kd = 0.01\). These parameters were obtained empirically.
3) **Software - Camera**

For the hand tracking and gestures recognition is used the *Intel RealSense SDK* and *OpenCV* library. The *RealSense* Application Programming Interface (API) returns, in real time, a skeleton model of the human hand containing the positions of twenty-two hand joints in a 3D space.

The camera software will function in two modes: (i) hand tracking (Figure 6), which is already implemented, with or without velocity control; (ii) gestures recognition, which is still under development.

In the first mode, the camera provides the position and orientation of the hand. For this purpose, and since it does not influence the main outcome, it is only considered two of the twenty-two joints, for performance reasons. The two joints considered are the center of the wrist and the middle finger tip (Figure 7).
...
For example, to determine if the index finger is stretched out, the sequence should be 10000, where the sequence represents the index of the middle, ring, pinky, and thumb fingers (Figure 9). The code was developed to facilitate the introduction of new gestures.

\[ \text{dist}_j = \sqrt{(F_{tx} - W_x)^2 + (F_{ty} - W_y)^2} \]  

(5)

5. Results

About the goals of this project, they were not all reached yet, since the project is still under development. However, from what could be already tested, the movements of the platform are satisfactorily accurate. It is very important to maintain present that the data gathered from the camera is very influenced by the environment conditions, such as, the light, inducing some errors.

In the hand tracking mode, the results are very positive. The platform describes all singular trajectories (forward, backward, left, right and rotation over itself) with extreme accuracy, since it is the platform that replicates the hand movement (verified empirically). When doing combined movements, such as, forward and turning left at the same time, the platform movement is good, however not perfect.

In the gestures recognition mode, until now, there were only made some basic tests, but it is already possible to see some positive reactions from the platform, since motors reference signals do not come directly from the camera. These values are predefined, making them immune to noise and have a higher reliability regarding the first mode. Once the values are static and predefined, the platform reacts with more accuracy in the combined trajectories and with the same accuracy in the singular trajectories.

Gestures recognition presents some errors in the detection. This validation was made by comparing some characters displayed in the PC with the related gestures performed by the user. At the moment, it is only possible to recognize six gestures (table 1).

Regarding Arduino, the communication with the computer works well and the libraries developed to the motors work as expected to ensure a fluid movement. As illustrated in Figure 5, the measured wheel velocity is always close to the reference wheel velocity, minimizing the error. This ensure that the velocity provided by camera/user is the mobile platform velocity.

<table>
<thead>
<tr>
<th>Sequence of trues</th>
<th>Gestures</th>
</tr>
</thead>
<tbody>
<tr>
<td>00000</td>
<td>All five fingers shrunk</td>
</tr>
<tr>
<td>10000</td>
<td>Index finger stretched out</td>
</tr>
<tr>
<td>11000</td>
<td>Index and Middle finger stretched out</td>
</tr>
<tr>
<td>11100</td>
<td>Index, Middle and Ring finger stretched out</td>
</tr>
<tr>
<td>11110</td>
<td>Index, Middle, Ring finger and Pinkie stretched out</td>
</tr>
<tr>
<td>11111</td>
<td>All five fingers stretched out</td>
</tr>
</tbody>
</table>
6. Conclusions and Future Work

Mobile platforms with four Mecanum wheels can be an asset to some solutions, improving them. Computer Vision solutions offer a way of freeing the user of physical controllers. These technologies can be an asset to improve solutions in an industrial environment, among others.

The main goal of this project was to develop a motion control system based on human gestures for a mobile platform. The solution presented in this paper aims to control the four Mecanum wheels of the mobile platform, by tracking the user hand and recognizing the hand gestures. The Intel RealSense 3D camera (model F200) was used as a tracking device.

Although the project is not finished, the preliminary experimental results show a good reaction from the platform to the hand movement, in both modes, providing a good interaction between the platform and the user.

As future work, it is suggested improving the number of recognizable gestures, developing an autonomous navigation functionality connected to the gesture recognition mode, testing the system with the new Intel RealSense 3D SR300 (it presents a higher resolution and cleaner image), integrating a vision system in the platform and developing a Graphical User Interface (GUI) to facilitate the Human-Robot interaction.

Acknowledgements

The authors would like to express their acknowledgments to COMPETE: POCI-01-0145-FEDER-007043 and FCT – Fundação para a Ciência e Tecnologia within the Project Scope: UID/CEC/00319/2013.

References