Journal of Physics: Conference Series

PAPER • **OPEN ACCESS**

A Human-Robot Interaction for a Mecanum Wheeled Mobile Robot with Real-Time 3D Two-Hand Gesture Recognition

To cite this article: Xueling Luo *et al* 2019 *J. Phys.: Conf. Ser.* **1267** 012056

View the article online for updates and enhancements.

A Human-Robot Interaction for a Mecanum Wheeled Mobile Robot with Real-Time 3D Two-Hand Gesture Recognition

Xueling Luo[1], Andrea Amighetti[1,2] and Dan Zhang[1,*]

[1]Department of Mechanical Engineering, York University, Toronto, Ontario, Canada
[2]Department of Industrial Mechanical Engineering, University of Brescia, Brescia, Italy

* Email: dzhang99@yorku.ca

Abstract. Human interaction with mobile robot becomes a popular research area and its applications are widely used in industrial, commercial and military fields. A two-hand gesture recognition method with depth camera is presented for real-time controlling the mecanum wheeled mobile robot. Seven different gestures could be recognized from one hand for mobile robot navigation and three gestures could be recognized from the other hand for controlling the gripper installed on the robot. Under the proposed control scheme, the mobile robot system can be navigated and can be operated at the same time for achieving missions by two different groups of hand gestures. The accuracy of the gesture recognition is about 94%. During mobile robot control experiment, the system works timely, accurately and stably for certain tasks such as directional movement, grasping and cleaning obstacles.

1. Introduction

With the development of human computer interaction (HCI), hand gesture recognition based HCI becomes one of the most popular implementation since it is the most similar way as interaction between human and human daily operation activity [1]. In the past decades, there are two main type of sensors for hand gesture recognition: contact sensors such as 'Cyberglove' [2], body-worn inertial sensors [3] and non-contact sensors such as biplanar videoradiography system [4], camera systems [5]- [7]. Those contact sensors are measuring gestures into electrical signals for HCI which could bring accurate detected result and it already been applied into military, industry and commercial fields. However, wearable devices are not always comfortable or welcomed by users and this kind of devices are hard to be manufactured and maintained since they normally consist by several parts and requires complex calibration [5].

While contact sensors for gestures recognition are still widely applied, non-contact sensors got more attention in recent years by its features such as natural interaction, measuring without physical contact and funny user experience. Non-contact sensors show more advantages comparing with traditional sensors in certain applications such as it will not take any risk to damage the object and it has high accuracy. Many application systems with gestures recognition become available in the market for smart surveillance, teleconferencing, HCI, etc. [8]. Besides, gesture recognition also be applied in a visual environment for very large scale biomolecular modelling [9] and works as integrated controller for the virtual environment Battle Field [10]. The most common method for such applications is color spaces recognition which could effectively reach the desired result and could also solve problems like partially occluded [11], [12]. Since vision system was designed by inspiration from human eyes which could be applied for 3-dimension measurement. A few 3D sensors have been published in recent years such as Kinect, Wii, Intel RealSense 3D camera. Among of numerous application, one of the most interesting application is using gesture for mobile robot control.

Zhao [12] presented a gesture-based control method for mobile robot. To apply pattern recognition method such as mean filtering, HSV color space and to use morphological image processing for static gesture recognition. Five different gestures are encoded with their feature information. A library is created by this process and it is also used for later recognition with contour correspondence method. During the control for the robot, a gesture-based state-feedback control strategy is applied. This method only provides five gestures for mobile robot control which is hard to satisfy a complex operation for real application. Besides, the experiment for hand gesture recognition method is in an experimental environment, the detection background is pure white which is easy for the detection. Kundu [13] proposed a hand gesture-based classifier algorithm to control an omnidirectional wheel-chair system. Seven motions are corresponded to seven different gestures to control maneuverability of the wheelchair. Wearable IMU sensor and two EMG sensors are used to extract wrist movement and two forearm muscle activities to encode gestures and for later gestures recognition. The system accuracy is depended on the wearable sensor which need special training to make sure the sensor is placed right for normal user. And the sensor need to be handled carefully to keeping it function. Besides, the training method for the gesture will take a lot of time. Gao's [13] robotic wheelchair control system based on hand gesture control for the disabled is using Kinect sensor for mobile robot operation. The method used Fuzzy control to create a corresponding from gestures to speed and direction for the wheelchair. This method provides a recognition method for gesture movement trend to command the robot's moving trend. During the experiment, there must have some vibration of the hand since users have to hold their hand in the air through the whole time which could affects the control strategy. Besides, during the high-speed model the control method is not working as it desired [13].

We developed a two-hand gesture recognition method with a depth camera from the Kincet sensor for controlling the mecanum wheeled mobile robot. A gripper was designed and equipped for the mobile robot. A model based on 3D geometry is designed for gesture recognition. Seven different gestures could be recognized from right hand for mobile robot movement operation and three gestures could be recognized from left hand to control the gripper. In the experiment, certain tasks like moving in a designed map, gripping an item in the path and put the object in certain place. The system works timely, accurately and stably.

2. Proposed method

We proposed a mobile robot operational control method with gesture recognition which can be described as figure 1. The system consisted of two main processes: image processing and human-robot interaction. During image processing, there are two different groups of gestures defined respectively for robot navigating and robot operational control. In this paper, we presented a hand-finger poses estimation method based on pose-geometry and depth information.

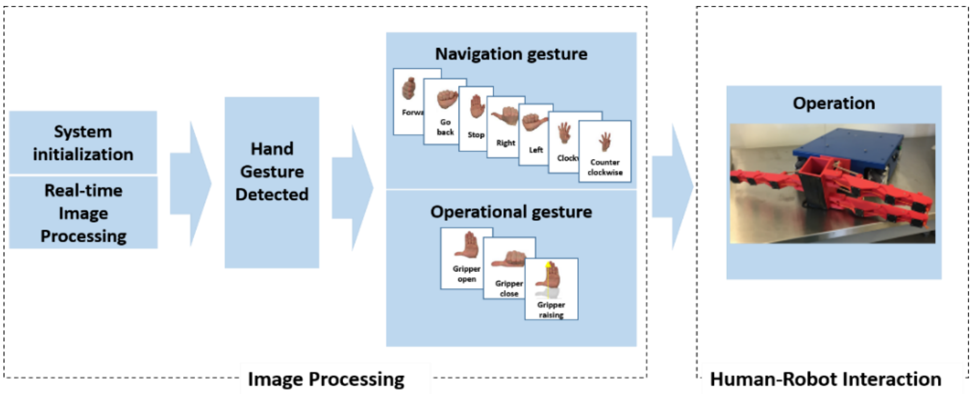
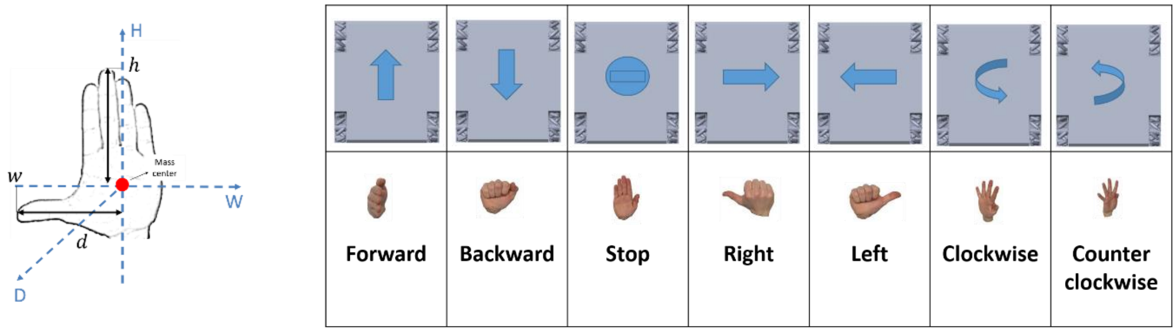


Figure 1. The system working flowchart

IOP Conf. Series: Journal of Physics: Conf. Series **1267** (2019) 012056 doi:10.1088/1742-6596/1267/1/012056

In order to discriminate 7 different gestures, a coordinate system of a hand was defined which is shown in figure 2. In figure 2(a), h, w and d are the length of the mass centre to the middle finger in vertical direction, the length of the mass centre to the thumb in horizontal direction and the length of the mass centre to the thumb in depth direction.



(a) **(b)**

Figure 2. (a) Hand gesture coordinate, (b) Hand gesture for mobile robot operation

There are 7 different hand gesture defined for mobile robot operation which are shown in figure 2(b). Based on gesture geometry, a fast and accurate method could be presented as function (1)

$$G(w,h,d)=\begin{cases} w(T_{WN}) \quad \&\& \quad h(T_{HN}) & \ldots Backward \\ w(T_{WX}) \quad \&\& \quad h(T_{HN}) & \ldots Right \\ w(-T_{WX}) \quad \&\& \quad h(T_{HN}) & \ldots Left \\ w(T_{WN}) \quad \&\& \quad h(T_{HX}) & \ldots Stop \end{cases} \tag{1}$$

where, $G(w,h,d)$ is the current gesture geometry. When there is no rotation trend detected, the gesture would be classified in logic equation (1). There are two gestures is required for initialization which are the gesture “Left” with thumb open and the other four fingers close and the gesture “Stop” with thumb close and the other four fingers upright. This process is to determine ranges of thresholds that thumb close - T_{WN}, other four fingers close - T_{HN}, thumb open - T_{WX} and the other four fingers open - T_{HX}, respectively. When both $T_W < T_{WN}$ and $T_H < T_{HN}$ are detected, the state of the detection is “Forward”. During the operation, if the depth of the mass centre is changed from the initial value, all thresholds will change with the same ratio.

When the rotation trend appearing in the system, the rotation direction of the system can determinate by checking the movement of thumb in depth direction. For discriminating the rotational signal, we use the famous method “convexity defects” [14] [15] for detection. When T_{WN} is detected, the command signal for navigation is clockwise rotation. When $-T_{WN}$ is detected, the command is counter clockwise rotation.

Figure 3. Hand gesture for mobile robot operation

Operation for gripper control is less complex than movement control. There are only three gestures for the gripper's control. In the figure 3, the map between those three gestures and the gripper commands is shown. The detection logic of the operational gestures is similar as the navigation gestures detection. When "Gripper open" and "Gripper close" are detected, if additional vertical direction movement of the gesture centre is also detected, the system can respond by sending gripper "lifting" or "putting-down" signal.

3. Mobile robot kinematic

Mecanum wheels have been designed with passive rollers around the wheel circumference at $45^{\circ}$of angle. In the figure 4, the configuration of the robot is shown.

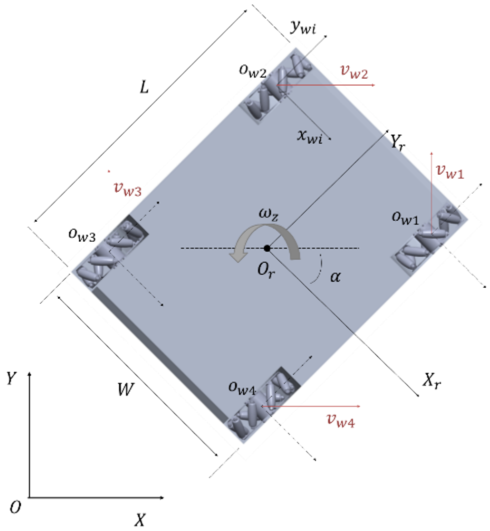


Figure 4. The configuration of the robot

Then, the inverse kinematics of the robot could be expressed as:

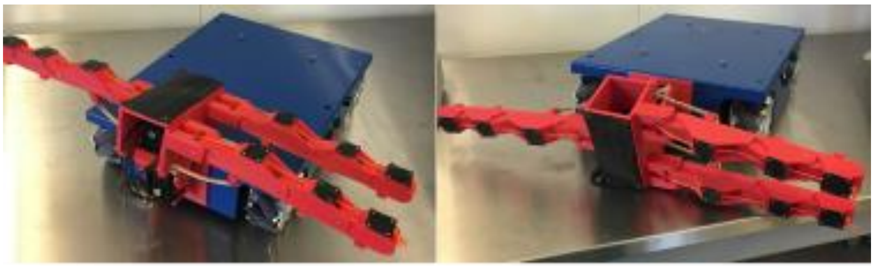
$$\begin{bmatrix} \dot{x}_r \\ \dot{y}_r \\ \dot{\alpha}_r \end{bmatrix} = \frac{R}{4} \begin{bmatrix} -1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 \\ \frac{W+L}{2} & -\frac{W+L}{2} & -\frac{W+L}{2} & \frac{W+L}{2} \end{bmatrix} \begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \\ \dot{\theta}_3 \\ \dot{\theta}_4 \end{bmatrix} \tag{2}$$

where, the coordinates OXY is the fixed coordinate and the $0_r X_r Y_r$ is the moving coordinates for the robot. The position of the robot is $S_r = [x_r,\ y_r,\ \alpha_r]^T$, R is the radius of the wheel, W and L are wide and length of the robot, $[\dot{\theta}_1, \dot{\theta}_2, \dot{\theta}_3, \dot{\theta}_4]$T is the angular velocity of each wheel. In the control of the robot, the signal could be sent in as the coordinate$[\dot{x}_r, \dot{y}_r, \dot{\alpha}_r]$.

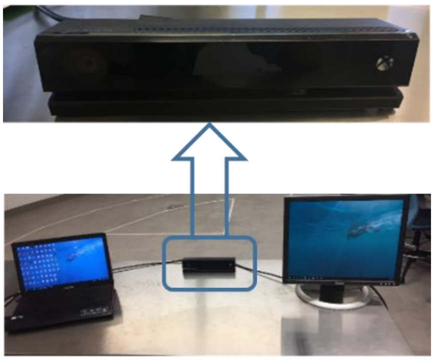
4. Experiment

For hand gesture recognition, a laptop with windows 10 system, Inter I7-6700 CPU is working as the main processor. In the figure 5(a), the mobile robot with mecanum wheels is shown and a gripper is installed on it. Inside of the robot there is an Arduino chip working as the main processor for the robot. The gripper is an underactuated gripper with two motors with 2-DOF: open - close and lifting - putting down. The communication between the laptop and the robot is based on a pair of Bluetooth chips. The set up for the gesture recognition system is shown in figure 5(b). And a Kinect is used as the visual sensor of the system. The figure 5(c) is the experiment area for the mobile robot to verify the proposed method could order the robot to act as it has been commanded. The experimental field is label as a map with straight line and curve which is for verifying flexibility of the proposed control method for the mobile robot system. During the experiment, there are several obstacles placed on the map for robot to pick up and remove.

IOP Conf. Series: Journal of Physics: Conf. Series **1267** (2019) 012056 doi:10.1088/1742-6596/1267/1/012056



(a) The mecanum wheeled robot with a gripper system



(b) The visual system for gesture recognition

(c) The experimental field

Figure 5. The experiment setup

In the table 1, the accuracy of the gestures recognition for both hands are show. It is a continue experiment with a series of gesture changing in turn. The number of detected gestures is shown in the table and the number of correctly detected gestures is also shown in the table. The accuracy of the navigation gestures is 94.39%. And the accuracy of the operational gestures is 96.34% since the gesture group is less complex than the navigation gesture group.

Table 1. Accuracy of hand recognition algorithm

	Detected (Number)	Correctly detected (Number)	Accuracy (%)
Forward	217	206	94.93
Backward	194	185	95.36
Right	226	213	94.25
Left	207	196	94.69
Stop	191	183	95.81
Clockwise rotation	213	197	92.49
CCW rotation	192	179	93.23
Gripper close	230	221	96.09
Gripper open	217	209	96.31
Gripper lifting	208	201	96.63

For the initial value setup, we detected the 'stop' and the 'left' gesture from the right-hand. And the 'gripper close' gesture was detected for operational gesture initial value. This process must be done manually to make sure those gestures could be detected clearly. In the figure 6, a few photos took from the right monitor of the experiment setup in figure 5(a). Figure 6 are different gestures corresponding to operation forward, backward, right and counter clockwise rotation, respectively. The robot could respond to the gesture command accurately and timely.

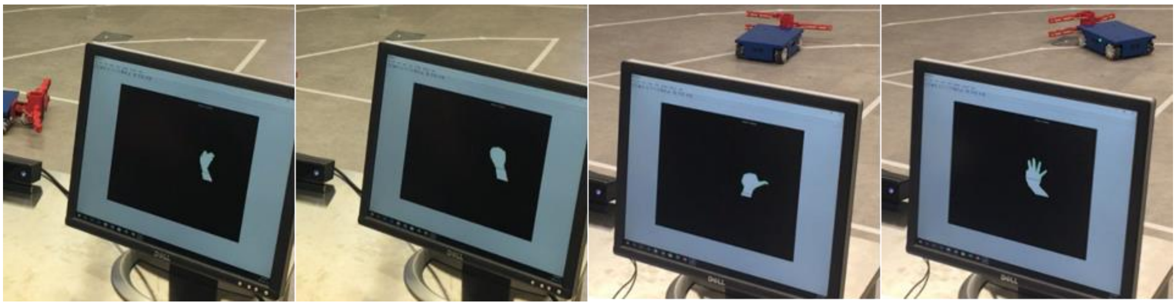
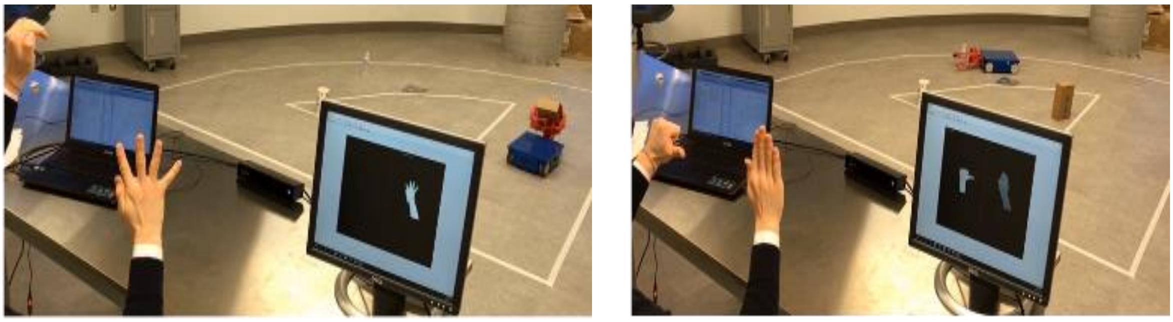


Figure 6. The result of mobile robot navigation with hand gesture

To verify the system could work for certain tasks, we put a few objects in the map and use two-hand to operate the robot for directional movement, picking up the object and putting the object into a certain place. In figure 7, photos of navigation and operational processes are shown. In figure 7(a), the robot is responded to command lifting gripper and counter clockwise rotation in the same time. In current moment, the robot tries to move the obstacle away from its path. In figure 7(b), the robot is responded to command closing gripper and stop moving in the same time. Under this stage, the robot met an obstacle and it had to stop and grip the obstacle in front of it.



(a) To rotate the robot CCW and to lift the gripper **(b)** To stop the robot and to close the gripper

Figure 7. The result of mobile robot navigation and operation with hand gesture

In the experiment, for different users we must initialize the system for each of them to make sure the high accuracy of this method. The user should try to keep each gesture as stable as possible. However, small rotation angle of each gesture could be adjusted by the threshold ranges and the slow change in depth direction could also be accepted by the algorithm.

5. Conclusion

The hand gesture method we proposed could recognize seven commands from right-hand and three commands from left-hand for the mecanum wheeled mobile robot control. The system works timely, accurately and stably for the desired tasks such as directional movement, grasping the item and carrying it to the certain place. The accuracy of the gesture recognition is 94.4% for this application. The experiment result proved that the designed system could be control remotely and it has well performance on those certain tasks. This kind of system could be useful in industry workshop or for certain risky task such as carry hot, sharp, chemical or radioactive items. The camera system is easy to install or maintain. The training of this method only needs three well placed gesture which would easy for users to learn. Besides, the algorithm is using the depth camera which would not affect by the light condition. This feature of the proposed method will allow it to be used in special working space where the light condition is strictly controlled.

However, this work still could be improved. The hand recognition method is designed for a relatively clean background which is not allowed much interference. If the human body parts such as main body or human face is too closed to the hand or their position are too close to the camera than the hand. The algorithm would not detect the gesture accurately.

Acknowledgments
The authors would like to thank the financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC) and gratefully acknowledge the financial support from the York Research Chairs (YRC) program.

References
[1] Ren Z and Yuan J. "Depth Camera Based Hand Gesture Recognition and its Applications in Human-Computer-Interaction", *IEEE 8th Int. Conf. on Information, Communications and Signal Processing*, Singapore, 2011.
[2] Lee C and Xu Y, "Online, interactive learning of gestures for human/robot interfaces", *IEEE Int. Conf. on Robotics and Automation*, Minneapolis, 1996.
[3] Junker H, Amft O, Lukowicz P and et. al. "Gesture spotting with body-worn inertial sensors to detect user activities", *Pattern Recognition*, 41: pp. 2010-2024, 2008
[4] Miranda D.L., Rainbow M.J., Crisco J.J. and et. al. "Kinematic differences between optical motion capture and biplanar videoradiography during a jump–cut maneuver", *Journal of Biomechanics*, 43: pp. 567-573, 2013
[5] Ren Z, Yuan J, Meng J and et. al. "Robust part-based hand gesture recognition using Kinect sensor", *IEEE Trans. on Multimedia*, 15(5): pp. 1110-1120, 2013.
[6] Shimada N, Shirai Y, Kuno Y, and et. al. "Hand gesture estimation and model refinement using monocular camera-ambiguity limitation by inequality constraints", *In Processing of Third IEEE Int. Conf. on Face and Gesture Recognition*, Nara, 1998.
[7] Chua C, Guan H, and Ho Y. "Model-based 3d hand posture estimation from a single 2d image", *Image and Vision Computing*, 20: pp.191 – 202, 2002.
[8] Hsieh C, Liou D and Lee D. "A real time hand gesture recognition system using motion history image", *The 2nd Int. Conf. on Signal Processing Systems*, pp.394-439, Dalian, 2010.
[9] Zeller M, Phillips J.C., Dalke A and et. al. "A visual computing environment for very large scale biomolecular modeling", *IEEE 24th Int. Conf. on Application-Specific Systems, Architectures and Processors*, pp. 3-12, Zurich, 1997.
[10] Wu C, Chen W and Lin C. "Depth-based hand gesture recognition", *Multimedia Tools and Applications*, 75(12), pp. 7065 – 7086, 2016.
[11] Soleimanizadeh S and Mohamad D, Saba T, Rehman A. "Recognition of partially occluded objects based on the three different color spaces (RGB, YCbCr, HSV)", *3D Research*, 6(3): pp.1–10, 2015.
[12] Zhao H, Hu J and et. al. "Hand Gesture Based Control Strategy for Mobile Robots", *2017 29th Chinese Control and Decision Conf.*, pp: 5868-5872, Chongqing, 2017.
[13] Kundu S, Mazumder O and et. al. "Hand Gesture Recognition Based Omnidirectional Wheelchair Control Using IMU and EMG Sensors", *Journal of Intelligent & Robotic Systems*, pp. 1-13, 2017.
[14] Manresa C, Varona J, Mas R and et. al. "Perales Hand tracking and gesture recognition for humancomputer interaction", *Electron Letters Computer Vision and Image Analysis*, 5(3), pp. 96–104, 2005
[15] Yeo H, Lee B and Lim H. "Hand Tracking and Gesture Recognition System for Human-computer Interaction using Low-cost Hardware", *Multimedia Tools and Application*, 74, pp. 2687-2715, 2015.