Gesture Controlled Robot

Team 30 — Bofan Yang, Qinlun Luan, Arvind Vijaykumar
ECE 445 Project Proposal — SP19
TA: David Hanley

1. Introduction

Objective

Wireless communication systems form the backbone of modern-day robotics control. Namely, wireless remote control methodologies including IR, RF, and network-based technologies such as WiFi and Bluetooth facilitate communication between a client and a robot such that the robot can successfully actualize its desired functionality. The main advantage it possesses over wired control is that they provide a much broader range for the robot to interact with its environment. External peripherals (game controllers, smartphones, etc.) are usually required in order to wirelessly transmit data to the robot; however, control schemes that don't require the use of an external device have yet to be brought into the mainstream by the robotics community at large. Touchless UI technology based on artificial intelligence theory could form the backbone of a potential solution to this problem, as it draws less attention to the client's method of control and requires no addition device[1]. Moreover, our primary objective in this course is to combine these technologies to explore a new and innovative method of robotic control.

We propose the implementation of a robot that can be remotely controlled via hand gestures to serve as a proof-of-concept for a non-peripheral, machine learning based robotics control scheme. We intend on accomplishing this through a Python-based software application that can identify a hand gesture displayed to the computer's camera, associate it with its corresponding function, and transmit the data over Bluetooth to control the robot. Our robot's functionality lies entirely within the domain of motion control; therefore, the input hand gestures will each correspond with a direction of motion for the robot as well as a gesture to cease movement. We intend to utilize a convolutional neural network (CNN) based architecture for our hand gesture recognition system, the details of which will be elucidated upon in the following sections.

Background

As highlighted above, interfaces used for human-robot interaction often requires using an external device. According to research done by Fujitsu Laboratories[2], a touchless user interface like this hand-gesture based control scheme is an optimal means by which users can control robots, since hand gestures don't require external devices and in general are natural control mechanisms that serve as an instinctual means of communication. In addition, there exists a myriad of different gestures we have at our disposal that can be formed and mapped to our robot's individual functions. In general, the existing research in touchless, specifically, hand gesture recognition based robot control systems often emphasizes on the utilization of external sensors (Leap Motion[3], for example), the accuracy of which is often dependent on the distance between the user and the robot, which is not particularly optimal as a remote control mechanism. In addition, the bulk of the processing is done on the robot's microcontroller itself, meaning that based on the application, the processing unit can be quite costly. In some cases, the systems require the camera module to be mounted to the robot itself for computation, which, again, isn't optimal for remote control, and the number of gestures utilized by the robot for movement is limited to those that are relatively easier to decode via methods in computer vision[4]. More complex computer vision and patterns recognition based methods like template matching and the use of feature vectors are also commonly applied to hand gesture recognition robot control systems[5]. Another main problem of computer vision used in robotic control is the accuracy of the overall motion and according to a smart robotic arm project conducted in Iskenderun Technical University, the use of convolutional neural networks has an 87.8% overall accuracy[6]. Thus, the convolutional neural network based model has the potential to mitigate and improve upon the issues seen in existing designs and research.

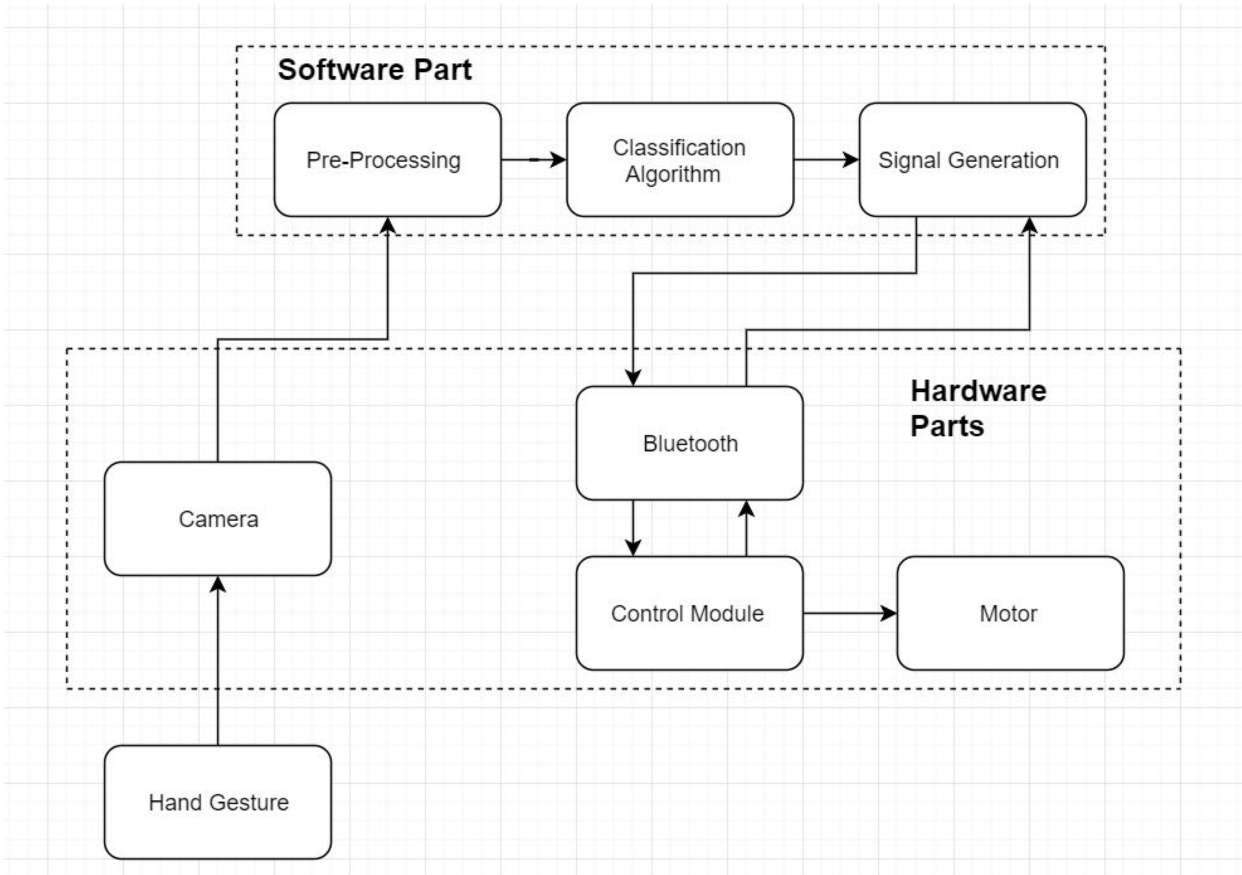
High-Level Requirements

- Computer vision based hand gesture recognition system must be able to identify gesture inputs successfully at a high degree of accuracy.
- Corresponding robot motion must occur in real time without any noticeable delay.
- Robot design must be functionally sound with respect to its operating environment, power consumption, speed, and component placement.

2. Design:

Block Diagram

The three primary elements in our modular design include the software component, the interface between software and hardware, and the hardware component. The software system takes a hand gesture as input from the client's laptop camera and identifies it through the CNN based gesture recognition system. Once an identification is made, the interface links the detected gesture with its equivalent robot motor control function and sends the necessary data via Bluetooth to the Bluetooth module. Finally, in hardware, the Bluetooth data transmits the received data to the microcontroller, which then carries out motor control.



Functional Overview

Hand Gesture: This block serves as input to our system. The user can input a gesture from a predetermined set of hand gestures that can be used to control our robot.

Camera: The system must acquire a real-time video feed in which the client is able to display hand gestures for processing. For that purpose, the application would require access to the computing device’s camera.

Pre-Processing Module: This module is required to convert the gesture data into a format recognizable by the neural network by reading and resizing the image, removing noises, and creating smooth segmentation. Specifically, light and background invariance are required in order to separate the hand from its environment, regardless of the environment in question. Thresholding would be also be required such that the application is also skin-complexion invariant. Finally, the hand region would be detected and the necessary data required for the CNN to identify and distinguish the hand gesture would be obtained through various topological and geometric transformations including contour extraction and detecting convexity defects and sent to the next stage of the software component.

Classification Algorithm Module: This module determines the hand gesture based on data from the previous module. By feeding the data obtained in the previous module to the CNN classifier, CNN will naturally be able to learn and identify the features associated with each individual gesture for identification. Once a gesture is identified to a high degree of accuracy, the next step would be to map it to the corresponding direction of movement.

Signal Generation Module: At a high level, this module is reminiscent of a switch box. The module receives a signal informing it that a valid gesture has been identified by the CNN classifier and based on that gesture, the module sends the corresponding control signals, through the Bluetooth module and to our robot for further processing. The module would also be responsible for detecting acknowledgment signals sent as feedback from the microcontroller and displaying them to our application.

Bluetooth Module: This module controls the interaction between the software component and the hardware component. It is mounted on the vehicle and wirelessly receives the aforementioned control signals over Bluetooth to be sent to the microcontroller.

Control Module: This module determines the rotational velocity and direction of the motors based on the input control signals. It mainly manipulates the voltages applied to each motor. The power applied to the motor changes along with the changes in voltage magnitudes and the direction and speed of movement also depends on the sign and magnitude of the voltages. After the voltages are set, an ACK signal is sent back to the client to notify that the data has been received and processed.

Motors: The motors provide power for the rotating wheels and are connected to a common power source.

Block Requirements

Camera:

1. The camera needs to be able to take obtain the real-time video feed at a sufficient resolution (probably 1280x720 minimum).
2. Software support must be implemented to ensure the camera works successfully and sends real-time data to the next module.

Pre-Processing Module:

1. This module must segregate the user's hand from the environment, filter out excessive lighting that could potentially alter the hand region, and provide an appropriate threshold for skin color filtering such that the hand region can be accurately determined.

2. This module must be able to distinguish invalid and valid data for input into the CNN classifier. It shouldn’t send information to the next module if the images are considered to be unqualified.
3. It must find the exact and correct information in the accepted images; in particular the significant contour extractions, convexity defects, and so forth. The error rate (including the rate unqualified images regarded as accepted images and the rate of incorrect data in the output) should be at most 10%.

Classification Algorithm Module:

1. The processing time of each gesture should be fast enough such that there is no significant delay between the time the gesture is displayed to the screen and when the gesture is identified.
2. An acceptable degree of accuracy must be determined before the system send the required control signals to the next module (verified through testing).

Signal Transmission Module:

1. Compatibility must be ensured with the Bluetooth module such that it can transmit and receive signals sent to and from the module respectively.

Bluetooth Module:

1. This module must be appropriately placed onto our robotic cart.
2. The range of transmission should be within an acceptable radius (verified through testing).
3. Must be compatible with our control module (likely a TI microcontroller)

Control Module:

1. This module must be appropriately placed onto our robotic cart.
2. Must accurately decode input data and convert them to correct motor voltages (programmed in C).
3. Able to generate feedback to indicate if the order is successfully carried out.

Motors:

1. A 9V power source is used to provide the energy and each of them must be connected to a 330-ohm resistor for motor control.

Risk Analysis

The bulk of our project lies in the successful operation of our hand gesture recognition system. The neural network’s complicatedness can grow very fast without ensuring a high degree of modularity in our project design and debugging the system would be a chore. We intend on mitigating this through having our design be highly modular so that we can verify the operation of each individual subsystem before connecting them together. In addition, our design will largely utilize open source resources with regards to computer vision and pattern recognition algorithms as well as the CNN itself in order to mitigate its complexity.

Aside from that, the hardware modules are relatively straightforward to implement. However, another potential risk could result from the time complexity of the combined resources in translating the gesture to the direction of motion, especially putting our second high-level requirement into consideration.

3. Ethics and Safety

The main ethical concern in our project is data safety. We need a large training set consisting of clear hand gestures for our CNN to identify relevant features and use them to distinguish between gestures. Some of these photos may contain surrounding information that may be considered sensitive by their owners. It is our responsibility to obtain open source training data that is viable to utilize for this project.

In addition, as we detailed in our original project pitch, we're modeling our robot based on the cart from ECE 110 lab and as such, there shouldn't be any severe safety issues in the design. While power dissipation is an inevitability in motor control based robotics, we need to ensure that our specified voltage and currents don't exceed the thresholds of the motors and the cart design itself should have all components placed and shielded optimally to avoid damage in the case of an accident, including the motors, microcontroller, Bluetooth module and power strip. Also, locomotive robots, in general, need to keep several design considerations in mind, including terrain understanding, limitations in speed and size, the materials used in the design, and operating distance from the client.

References

[1] EE Times, 'March of the Touchless User Interfaces', 2017. [Online]. Available: https://www.eetimes.com/author.asp?section_id=36&doc_id=1331859.

[2] Semantic Scholar, 'Touchless User Interface Utilizing Several Types of Sensing Technology', 2014. [Online]. Available: https://pdfs.semanticscholar.org/cf18/3eab584d0bd8971afa81f8ab5da40ca9913f.pdf.

[3] Chen, Sunjie, et al. "Hand Gesture Based Robot Control System Using Leap Motion." *Intelligent Robotics and Applications Lecture Notes in Computer Science*, 2015, pp. 581–591., doi:10.1007/978-3-319-22879-2_53.

[4] Thinh, Nguyen Truong, et al. "Gait of Quadruped Robot and Interaction Based on Gesture

Recognition.” *Journal of Automation and Control Engineering*, vol. 3, no. 6, 2015, pp. 53–58., doi:10.12720/joace.4.1.53-58.

[5] Wang, Mei, et al. “Hand Gesture Recognition Using Valley Circle Feature and Hu’s Moments Technique for Robot Movement Control.” *Measurement*, vol. 94, 2016, pp. 734–744., doi:10.1016/j.measurement.2016.09.018.

[6] Iscimen, Atasoy, Kutlu, et al. “Smart Robot Arm Motion Using Computer Vision.” [Online]. Available: http://www.eejournal.ktu.lt/index.php/elt/article/viewFile/13749/7276.