Some Previous Projects in COMS 6731, Humanoid Robots

Rob the Robot
is a concierge, a doorman, that greets guests and residents in an apartment building. The benefits of having a doorman are many. Among them are the benefits of having someone to receive and store oversized packages for you so you don't have to pick them up at the post office, and the benefit of having someone notify you when someone comes to visit you. However, given the high cost of hiring a doorman in NYC ($44,000 per year according to New York Post), it makes sense to use a humanoid robot to fulfill these tasks. Our program is implemented on the PR2 robot, which greets guests and delivers stored packages to guests. A similar implementation of this is the butler robot, which can locate, fetch, and deliver drinks.

Our implementation of Rob the Robot, the concierge, is divided into two parts. The first part,"Virtual Assistant," deals with the integration of high level components in order to maximize the amount of service oriented tasks accomplished (for example, mailing/texting/calling, speech recognition and synthesis, facial recognition, motion detection, etc.). The second part, "Object Recognition and Movement," deals with image processing and joint manipulation. The task here is to recognize the right package and perform delivery to residents. They are separated into two different nodes communicating each other through ROS publisher/subscriber platform.

Baymin
Our project focuses on exploring human-robot interaction through speech and vision. We initially obtained the motivation for this project from the Disney movie Big Hero 6 , featuring a robot named Baymax. Baymax interacts with other characters freely, and our goal is to replicate some of these interactions. Our main objective for this project was to implement various humanlike behaviors on a PR2 in increasing complexity and difficulty. Our first milestone was to teach the PR2 to recognize spoken greetings, such as "Hi" and "Good Morning." We trained it to respond with a predetermined action, in our case a hand wave. Our next milestone focused on training the PR2 to respond to a visual stimuli, which is activated by a verbal command. We trained the PR2 to recognize the phrase "Fist Bump" (similar to a command used in Big Hero 6). The PR2 will then position its gripper to mimic a human fist. Upon determining how far (i.e. depth using PointCloud) the human fist is, the PR2 will make a "fist bump" motion, moving its gripper back and forth. Our final milestone was to implement a continuous visual tracking behavior that is activated by the verbal command "Follow the girl/guy/person (or any object) in the red (or any specific color) shirt." For the purpose of our project, the PR2 will recognize the colors: red, blue, yellow, or green, and will follow the identified person. For the rest of the report, we will refer to these milestones as Task 1 (Greeting), Task 2 (Fist Bump), and Task 3

Mr. Geppetto: a PR2 Remote Interactive Control System Using Gesture/Body Recognition

Recent advancements in human-computer interfaces also make it possible to capture human gestures in detail. Microsoft Kinect sensor, for one, has been proven to be one of the most effective and affordable motion capture device and used in many HCI research. In addition, virtual reality and augmented reality technologies becomes mature to support real time or subrealtime environmental feedback from a remote place. For example, Oculus is widely adopted to provide integrated visual information to the user. It also captures the head position of the user. Furthermore, biological signals attracts people's eyes recently. The Myo armband is one of the commercial product using human muscle electrical signals to control various devices. In this project, we create a remote interactive control system to control PR2 using intuitive human actions such as gesture and voice, and receiving the environment data from the humanoid.

RobotArt

Much of human expression involves visual art and paintings. Many artists use a lot of subtle movements and make a lot of little decisions in order to create beautiful works. Humans are also very vision oriented; we do a lot of internal visual analysis whenever we see drawings and paintings, and from that we can derive the mood and thought processes of the artist creating it. All of this is a challenge to emulate with a robot. In this project, we explored the different ways humanoid robots could create visual art using computer vision and trajectory planning with the Baxter robot system.

Humanoid robot writes words

The main purpose of our project is to enable PR2 robot to write words using pen just like what human does in real life. The simulation is under ROS system with motion planning for arm using MoveIt!.

Robot Chess:

One long term goal of game robots is that robots can play games with human like human. Furthermore, robots can beat experts in games like AlphaGo which is the most successful go robot. Chexter is a chess robot which have approximately the same cognitive and manipulative abilities like human. It consists of a chess engine, computer vision and robot control. First, chess engine is the brain of chess robot, which computes next moves. Second, computer vision is the eye of chess robot, which receives the positions of chess pieces. Last, robot control is the body of chess robot, which executes the moves by the data from chess engine and computer vision. In this work, we focus on robot control. Our goals are moving arm without collision, locating chess pieces positions precisely, picking up chess pieces stably, and putting down the chess pieces successfully.

EMO-INFO-BOT: Emotional Information Provider Robot

Emo-Info-Bot is an emotional information provider social robot which provides relevant information about a landmark based on person's emotion. This social robot (software system) has the capability to understand the emotion of the person and interact with him/her accordingly. It takes as input an audio-video signal and a landmark (one of the 16 neighbourhoods of Manhattan) and deliver srelevant information about that
landmark depending on the emotion of the person in the audio-video signal. The system supports both offline and online (real-time) emotion recognition.

PR2 with xylophone
In this project, PR2 can
1) pick up stick
2) move head to look for xylophone
3) move to standby position to play xylophone
4) play xylophone
5) vocalize xylophone sound in Gazebo

TURIS: The Ultimate Robotic Ironing System
In this project, we built a intelligent system that can perform fully autonomous clothes ironing using a off-the-shelf hand-held iron. This task involves the following sub-tasks:

1. Recognize wrinkles on a piece of garment.
2. Plan a trajectory that the iron should move in order to effectively remove the wrinkles.
3. Based on the trajectory of the iron, plan a sequence of motions for the robot to move the iron.
4. Perform the planned motions.
5. Peform recognition again and evaluate if there are still wrinkles on the garment, and decide to stop the task of continue one more iteration.

Rosie the Delivery Robot
Our goal is for Rosie to transport medical equipment and tools in hospitals, so doctors and nurses can focus exclusively on taking care of patients. For Rosie to navigate the environment autonomously, we began by performing SLAM to create a map of the 6th floor of CEPSR. We then used the navigation stack provided by ROS for Rosie to move around the hallways autonomously.

Robotic Gomoku Player
In our project, our main goal is to make the PR2 robot play Gomoku game on a chessboard. Gomoku is an abstract strategy board game which needs two players to compete. The black side plays first, and players alternate in placing a chess piece of their color on an empty intersection. The winner is the first player to get an unbroken row of five stones horizontally, vertically, or diagonally [1]. So far, we have achieved two main goals, one is that we make the PR2 recognize the chessboard and grasp a chess piece from a specified position to a target position on the chessboard. The other one is that we make the PR2 to analyze the current situation of the game and use an AI to compute the next move.

3D Convnet: Grasp Energy Estimation for EigenGrasp Planner
The purpose of this project is to create an alternate grasp quality evaluation method within the GraspIt! eigengrasp planner that uses 3D convolutional neural networks to quickly and efficiently calculate the energy function of novel grasps. The current grasp quality methods

within GraspIt! use metrics which are largely determined by either the position and orientation of contact points or the configuration of a hand/gripper and the normals projected from the fingers. While the first category places a premium on the geometry of the object being handled, the second category mainly address the force configuration and limits of the hand joints. We believe there is a correlation and agreement between these broad categories as they address the same challenge from two different angles. Using deep learning, we couple these two methods by predicting grasp quality based on a full object model and a given hand configuration. Since the same network filters are used to predict multiple quality metrics, we hypothesize that the learned model would give a smoother model for generating grasp quality that is potentially more robust than current energy functions.

Baxter Learns it's A,B,C's

Given alphabet blocks displayed on a table, our robot (Baxter) will be capable of understanding speech commands, recognizing the letters, and picking the appropriate blocks up. Each block will have one letter facing up. The entire alphabet can be represented with the blocks.

Affective Emotional Characteristics in Speech for Humanoids

Humanoid behavior especially in the sense of spoken terms varies on our ability to understand how to teach emotions. From the moment we are born, we have expressions and emotions that are inherent as survival features according to some. Other emotions and expressions are learned from watching and perceiving the environment around us. Therefore, there lies in an inherent importance to being able to teach humanoids how to learn and react to these types of emotional traits.

In this study, I would like to do a human behavior study to understand how humans perceive different affective emotions, namely whether a person speaking is powerful, convincing, confident, charming, enthusiastic, believable or boring. To conduct the survey, I will use provide a human subject with a snippet of a student presentation video and ask for a rating between 1 and 5 to qualify whether a speaker exudes that attribute. I will conduct the study in two parts: in one version the human subject will be asked to judge the affective emotions based only on an audio stream and in the second version the user will be provided with both an audio and video stream.

The hypothesis is that perception of these affective emotions vary based on the type of input provided (audio only versus audio/visual cues) and thus the creation of both types of classifiers would be necessary to properly train a humanoid to decipher between the emotions. Using these results, I would like to create classifiers for these specific emotions so that a humanoid can utilize them to determine proper responses to interactions with a user. This type of human behavior is one that would need to be taught to the humanoid since a humanoid does not intrinsically have the ability to judge this.

EMOTIONS IN HUMANOID ROBOTS AND IMPROVING HUMAN-ROBOT INTERACTION

This is a study on the still nascent field of Emotions in Humanoid Robots and Improving Human-Robot Interaction. These fields have rich potential for future applications in a variety of fields, such as healthcare, customer service, etc. Robotic interaction with humans is limited today, with only a few industries using them to a major extent, but has great potential to be very useful. Incorporating emotions in robots is a major technical and cognitive challenge. In order to introduce emotions into humanoid robots, we must first have an understanding of basic human emotions and their nuances and combinations. Joy, anger, sadness, etc. are the fundamental emotions that human beings exhibit [1], and these emotions are the basis for more complex emotions such as optimism, contempt, etc.[1] We look at past work and examples of human-robot interaction and implementation of emotions in humanoid robots in various social settings. We then look to answer the ‘Appearance versus Reality’ question – a fundamental question which asks if robot behaviour in social settings is actually an indication of the social sophistication of the robot, or if the robot merely appears to be socially intelligent and is not actually so [2]. The Turing Test [3] is taken as a case of this – and we analyse the effectiveness of this test and look at alternative methods of answering this question and variants of the Turing Test. We look at both variants of the Turing Test – the cognitive aspect and the emotional aspect. Many computers have been known to be cognitively superior but not even close on an emotional comparison – the paper cites a few examples (IBM Watson, Google AlphaGo, etc.) in this regard [4] [5]. We then look into the nuances and caveats of empathy, perhaps the most complex emotion, and research how humans exhibit empathy and how it can be ported to robots. [6] Furthermore,