Human-Blimp Interaction

Emilie Baker
School of Mechanical and Aerospace Engineering
Cornell University
Ithaca, New York
Email: erb259@cornell.edu

Nialah Wilson
School of Mechanical and Aerospace Engineering
Cornell University
Ithaca, New York
Email: njw68@cornell.edu

***Abstract*—The focus of this project is on human interaction with autonomous blimps. Blimps have been shown to be less intimidating than quadrotors due to their reduced noise level and appearance. We examine if humans can recognize gestures from a blimp and if the blimp can infer how well the human understands its commands based on the human's body positions. We do the ladder with a decision tree. Two experiments were conducted. The first was to gather attributes for the decision tree. The second was to test the predicting power of the decision tree and determine the human participant's success rate. The blimp gives a single command to a human participant. Then, based on the participants actions, the blimp continues with additional commands or repeats the command. The corrected decision tree predicted human success/failure with 73 percent accuracy and humans were able to infer the correct command 83 percent of the time.**

I. Introduction

Gesture recognition by robots has been studied previously [1]-[2]. A large focus has been on identifying human gestures and controlling a robot based on these gestures. We want to understand the reverse; if, and to what extend, can humans understand nonverbal commands given by robots. One area where this is useful is in crowd control and evacuation. An ongoing project at Cornell between the Collective Embodied Intelligence Lab, the Human-Robot Collaboration and Companionship Lab, and the Verifiable Robotics Lab focuses on ad-hoc collaboration of human-robot swarms in evacuation scenarios. In this study, mobile robots will be used to communicate information and give instructions to multiple humans. One of the mobile robots is an autonomous blimp, which will be used to project images onto the ground. The authors of [3] found that people are less intimidated by blimps than quadrotors due to their noise level and appearance. This prompted us to believe that blimps might be able to communicate commands to humans nonverbally via means other than image projection. In a disaster or evacuation scenario, it is especially important for the blimp to infer if people understand the given command. We infer this is based on human body pose, and design a decision tree so the blimp can understand how well people understand its commands. This study is exploratory work to understand how well blimps can communicate instructions to humans in an evacuation scenario and how well a blimp can predict the actions. Fig. 1 shows a participant during one of the preliminary experiments interacting with the blimp.

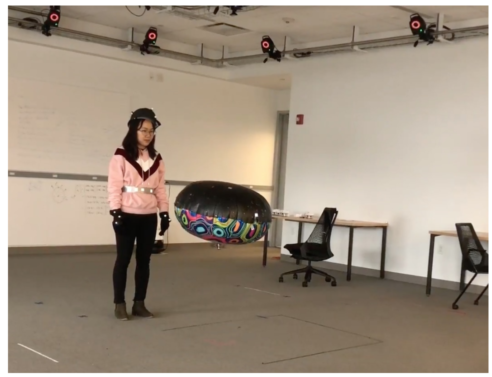

Fig. 1. Participant is wearing a helmet, gloves, and a belt which are all being tracked with a Vicon motion tracking system. In front of the participant, is the blimp. This was the starting state for each trial.

II. System

A. Preliminary Experiment

A commercially purchased blimp, Plantraco MicroBlimp, was used for the study. The blimp consists of two main components, a balloon and a gondola. The gondola has a microcontroller and 3 propellers each of which can be controlled separately and proportionally. For the preliminary experiment, the blimp's motion was controlled by an operator using the iPhone App via bluetooth. A larger balloon was necessary so that the balloon could be painted black and still support the weight of the MicroBlimp gondola. The balloon had to be painted black so that markers could be attached to the balloon and the blimp could be tracked as an object by a Vicon motion tracking system. In addition to the blimp being tracked by the Vicon motion tracking system, the participants wore a helmet, belt, and gloves (both left and right) that were also being tracked. The ROS package, vicon_bridge, was used to collect data from the Vicon motion tracking system and a python file was created for all the objects being tracked (including the blimp, helmet, gloves, and belt).

B. Final Experiment

The final experiment setup was identical to the preliminary experiment, except the belt was not used and the blimp was autonomous. The blimp was adapted to be controlled via a computer rather than the intended iPhone App. The decision tree was used to determine whether or not the participant

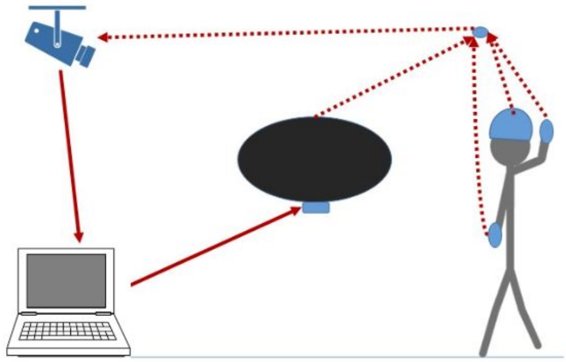

Fig. 2. Motion is tracked via Vicon and a decision tree is used to determine the blimps next command, based on body positions.

succeeded or failed to execute the correct command. This was output to the operator. Fig 2 is a diagram of the final experiment setup.

III. Method

A. Training Data

A preliminary study was first conducted in order to collect data to create a model that would be implemented in the final experiment. For the preliminary study, 1 female Cornell student and 2 male Cornell students were recruited to participate in the study without compensation. None of the students were permitted to watch each other to prevent the participants from figuring out what the study was trying to analyze. Each of the participants were asked to wear a helmet, gloves (both left and right), and a belt. Each of these items had markers attached to them and were being tracked by a Vicon motion tracking system.

For the preliminary study, the blimp was controlled by an operator using the iPhone App, Microblimp, as commercially intended. The commands given varied from turning in a circle, to moving in a straight direction, or some combination of the two. The commands were chosen at random by the operator. Each of the participants were given 3 different commands. The students were told that the blimp would be trying to communicate with them via it's motion. The students were told when the blimp started and stopped communicating an instruction to the participant. The participants were asked to wait until the command stopped before trying to execute the instructions given by the blimp. After the students attempted to execute the instructions given by the blimp, they were told whether or not they succeeded or failed to follow the instructions that the blimp tried to convey. If the participant failed to correctly mimic the instructions given by the blimp, the motion was repeated and the students were asked to make another attempt at trying to execute the instructions conveyed by the blimp. The participants were given up to three chances to understand and mimic the blimps instructions. After three failed attempts, a different command was given by the blimp.

Data was collected using vicon_bridge ROS package that permitted the collection of data from the Vicon motion tracking system. The data being collected included the x, y, and z translation coordinates as well as roll, pitch, and yaw for all the objects being tracked which included the blimp, helmet, left glove, right glove, and belt. In addition to tracking the motion of the blimp and the participants, all of the trials were video recorded with a phone.

B. Machine Learning

Using the data collected from the preliminary experiment, analysis was done to find patterns that would determine whether or not a participant would fail or succeed in following the instructions given by a blimp. For example, some of the features included the distance between the blimp and the participant, the distance between participant's hands, and the angle between the participant's head and the blimp.

The distance between the participant and the blimp predicted success correctly 100 percent of the time and correctly predicted failure 75 percent of the time. After analyzing the videos, this attribute was thrown out because the data was highly biased by the fact that the participant who succeeded in all their trials generally stood further away from the blimp, and the participant who failed in many of their attempts in general stood closer to the blimp. There would need to be more participant's in the trial data to know whether or not distance between the blimp and the participant can be used as a reliable feature. One of the expectations of this experiment was that the distances between arms and certain arm configurations would be good predictors. The reasoning behind this is that participants might shrug or raise their hand when they did not understand what the blimp was trying to communicate to them. However, none of the features regarding hands proved to be good indicators. After looking at the videos, some of the participants held their hands really close to their body where as others were more expressive by shrugging. Due to the variations, all of the features that were looked at for hand motion ended up not being good predictors.

The angle between the blimp and the participant's head was calculated using the Pythagorean theorem:

$$\theta = atan(\frac{z_{helmet} - z_{blimp}}{y_{helmet} - y_{blimp}}) \tag{1}$$

where $theta$ is the angle formed between the blimp and the helmet, z_{helmet} is the z coordinate of the helmet, z_{blimp} is the z coordinate of the blimp, y_{helmet} is the y coordinate of the helmet, and y_{blimp} is the y coordinate of the blimp. The y coordinates were used in this equation over the x coordinated because the blimp remained directly in front of the participant in all of the trials which is the y-axis. This angle corresponds to how much the participants were looking at the blimp.

From the angle calculations, we computed the minimum, maximum, and average angle for the duration of each trial. The average angle computed for each trail provided the most regularities in the data. The average angles for all the trails ranged from 0.023 to 0.723 radians, had an average of 0.477 radians and a standard deviation of 0.200 radians. From observing the average angles calculated across all trial and

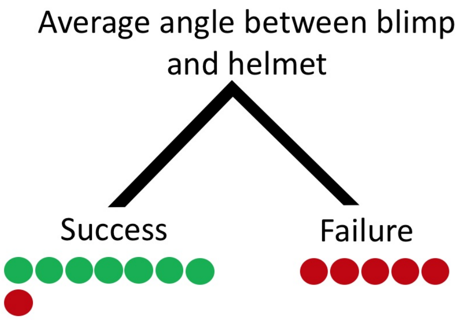

Fig. 3. Rule 1 decision tree for determining success or failure based on the average angle between the blimp and the helmet. This tree was based on preliminary data.

for all participants, a threshold value of 0.56 radians was established. This means that if the calculated average angle was less than 0.56 radians, the model predicts success. On the other hand, if the calculated average angle is greater than or equal to 0.56 radians, the model predicts failure. The threshold value was chosen based on its ability to correctly identify failure and success amongst all the trials.

Given the fact that many of the features we calculated ended up not being valid predictors, we decided to implement a 1R algorithm, one rule algorithm. This means that the most predictive single feature, which in this case was the angle between the blimp and the participant, is the only feature the algorithm uses to determine success or failure. An image of the decision tree is shown in Fig. 3. The feature, average angle between blimp and helmet, is able to predict all of the trials except one correctly. There is one trial that is a failure that the feature predicts as a success. An algorithm was wrote and implemented as part of the ROS package and python files that collect the data from the Vicon tracking system.

C. Final Experiments

The final experiments were conducted to test how well the decision tree predicted success/failure of the human participants. The Vicon system was used to track the blimp, and each participant's left hand, right hand, and head. There were a total of 6 participants, 2 female and 4 male, all Cornell students in the College of Engineering. They were all volunteers, had no previous knowledge of the experiment, and received no compensation. The participants were brought into the experiment room one at a time. They were instructed to put on the Vicon tracking helmet and gloves, then given instructions. They were told the blimp would try to communicate with them via its motion, and once it was done giving its instruction to infer what action the blimp intended for them to do.

The blimp was controlled by an operator via a computer. Depending on whether the participant succeeded or failed, the operator sent a predetermined command pattern. Each participant received 3 separate commands from the blimp. After 3 failed attempts, the next instruction was given. If the participant succeeded before 3 attempts was reached, the next command was given. Each participant received the same 3 commands, rotate in a circle, walk forward, and crouch down while rotating, but they were given in a random order.

	Predicts success correctly	Predicts failure correctly	Overall ability to predict correctly
Preliminary Experiment (threshold 0.56 radians)	87.5%	100%	92.3%
Final Experiment (threshold 0.56 radians)	100%	0%	50%
Final Experiment (threshold 1.7 radians)	86.7%	60%	73.3%

TABLE I
ABILITY OF THE DATA TO PREDICT SUCCESS AND FAILURE FOR DIFFERENT THRESHOLDS.

	Command								
	Rotate Down			Rotate			Forward		
Participant	Attempt 1	Attempt 2	Attempt 3	Attempt 1	Attempt 2	Attempt 3	Attempt 1	Attempt 2	Attempt 3
1	F	F	S	S	-	-	S	-	-
2	S	-	-	F	F	S	S	-	-
3	F	F	F	S	-	-	S	-	-
4	F	S	-	S	-	-	S	-	-
5	F	F	F	S	-	-	S	-	-
6	F	S	-	F	F	F	S	-	-

first second third

TABLE II
ABILITY OF THE PARTICIPANTS TO CORRECTLY EXECUTE THE GIVEN COMMAND.

IV. EVALUATION/FINDINGS

To test how well our model predicts success and failure, our model outputted it's prediction and this was compared to whether or not we observed the participant actually succeed or fail. Our model predicts success for every trial in the final experiment which results in the model having an overall ability to predict success/failure correctly 50 percent of the time as seen in Table 1. Due to the poor performance of the model, the data from the final experiment was analyzed. Specifically, the average angle between the blimp and the participant for each trial was computed and checked for patterns. The range for the average angle amongst all trials was between 0.003 to 0.400 radians with an average of 0.162 radians and a standard deviation of 0.092 radians. Given the original established threshold of 0.56 radians, and the fact that all of the average angle values were smaller than 0.400 radians, all of the data was correctly predicted as successful based on the model implemented. Next, the data was checked for regularities. A new threshold of 1.7 radians was established for final experiment data which resulted in the model correctly predicting success 86.7 percent of the time, correctly predicting failure 60 percent of the time, and an overall 73.3 percent ability to predict the outcome correctly.

To understand how well the participants understood the commands we recorded how many attempts each participant needed to successfully execute the command. This is detailed in Table 2. The commands were given in random order. The order in which the command was given is indicated by the color of the cell, where blue, yellow, or light blue means the participant received this command first, second, or third respectively. F indicates a failed attempt and S indicates a successful attempt.

Participants successfully completed the given command within three attempts 83% of the time. In three attempts, the rotate down command was successfully executed 67% of the time, the rotate command 83%, and walk forward 100%. To contrast, on the first attempt only 16.7% correctly executed rotate down, 67% correctly executed rotate, and 100% correctly executed walk forward. When rotate down was the first command given, success on the first attempt was 0%. When it was given third, there was a 33.3% success rate for the first attempt. Rotate down was never given second. For rotate, when it was first: 0% success, second: 100% success, and third: 100%, all for the first attempt. Walk forward was correctly executed 100% of the time regardless of order given.

V. Discussion

In order for the feature, average angle between the blimp and the participant's head, to correctly predict success or failure with an accuracy higher than 50 percent, a new threshold had to be established for the final experiment. This suggests that there is still a possible pattern in the data relating average angle between the blimp and the participant's head to success and failure. One possible reason why the threshold for the preliminary experiment and the final experiment were different is due to the change from an operator controlling the blimp to autonomous control. This is because we faced difficulties in being able to reliably repeat the same motions to all the participants in the preliminary experiment. Also, we should have measured and consistently placed the blimp a certain distance in front of the participant each time. Lastly, there was a longer pause between starting the data collection from the Vicon motion capture system and the blimp starting to give its instructions to the participants. Thus, participants may have been looking away from the blimp while the data started collecting which could result in needing a larger threshold value for the final experiments.

At first glance, it would appear that some commands, walk forward, are more intuitive to people than others, rotate down; however it is not as straight forward as this. Because the commands were given at random, the data is not evenly distributed between how many times a command was given first, second, or third. Therefore it is hard to differentiate between whether or not a command actually was harder to understand than another, or if it was just difficult in general for participants to understand the first command given.

VI. Conclusion

This exploratory work examined if people can correctly understand a nonverbal command given by an autonomous MicroBlimp, and if a MicroBlimp can infer humans' level of understanding based on their gestures. We found that the average angle between the participant's head and the blimp (indicating looking at the blimp) was the best predictor that the person had successfully executed a command. We also found that people were able to understand commands given by the blimp, some commands better than others. Our data is inconclusive however because we did not have enough participants for our data to be statistically significant nor was there enough consistency between participant trials.

Future work will include conducting more experiments to get a larger sample size, using a feedback loop to completely autonomously control the blimp, and use the output of the decision tree alone to determine whether or not the blimp should repeat its command or move to the next one.

VII. Acknowledgements

We'd like to thank Professor Guy Hoffman for his guidance, members of the Verifiable Robotics Lab for assistance with the Vicon system and the use of their lab space, as well as Professor Kirstin Petersen for the use of the blimp and other lab resources.

References

[1] Cicirelli et al. *A kinect-based gesture recognition approach for a natural human robot interface*, International Journal of Advanced Robotic Systems, 2015

[2] T. Naseer, J. Sturm, and D. Cremers *FollowMe: Person following and gesture recognition with a quadrocopter*, IEEE/RSJ International Conference on Intelligent Robots and Systems, 2013

[3] Chun Fui Liew and Takehisa Yairi. 2013. *Quadrotor or blimp?: noise and appearance considerations in designing social aerial robot*, In Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction (HRI '13). IEEE Press, Piscataway, NJ, USA, 183-184.