Multi-Operator Gesture Control of Robotic Swarms Using Wearable Devices

Sasanka Nagavalli, Meghan Chandarana, Katia Sycara
Robotics Institute, School of Computer Science
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213
email: snagaval@andrew.cmu.edu, mchandar@cmu.edu
katia@cs.cmu.edu

Michael Lewis
School of Information Sciences
University of Pittsburgh
Pittsburgh, Pennsylvania 15260
email: ml@sis.pitt.edu

***Abstract*—The theory and design of effective interfaces for human interaction with multi-robot systems has recently gained significant interest. Robotic swarms are multi-robot systems where local interactions between robots and neighbors within their spatial neighborhood generate emergent collective behaviors. Most prior work has studied interfaces for human interaction with remote swarms, but swarms also have great potential in applications working alongside humans, motivating the need for interfaces for local interaction. Given the collective nature of swarms, human interaction may occur at many levels of abstraction ranging from swarm behavior selection to teleoperation. Wearable gesture control is an intuitive interaction modality that can meet this requirement while keeping operator hands usually unencumbered. In this paper, we present an interaction method using a gesture-based wearable device with a limited number of gestures for robust control of a complex system: a robotic swarm. Experiments conducted with a real robot swarm compare performance in single and two-operator conditions illustrating the effectiveness of the method. Results show human operators using our interaction method are able to successfully complete the task in all trials, illustrating the effectiveness of the method, with better performance in the two-operator condition, indicating separation of function is beneficial for our method. The primary contribution of our work is the development and demonstration of interaction methods that allow robust control of a difficult to understand multi-robot system using only the noisy inputs typical of smartphones and other on-body sensor driven devices.**

Keywords–Robotic Swarms; Gesture Control; Wearable Devices.

I. INTRODUCTION

There has been significant interest in effective operator interaction methods for human-robot teams consisting of multiple humans and robots. Work in this area has been classified along multiple dimensions including the degree of coordination between individual humans or robots on the team (e.g., independent robots or coordinating robots) [1] [2], the association between robots and humans (i.e., shared pool of robots or robots assigned to individual humans) [3] [4] [5] and the physical distance between the humans and the robots (e.g., remote operators or humans working alongside the robots) [6] [7]. For a multitude of application scenarios ranging from urban search and rescue (USAR) [8] to supply support activities, a variety of operator interfaces for teleoperation [9] have been considered including those based on conventional keyboard, mouse and joystick, voice commands [6], haptic feedback [10] [11] and many more.

In recent years, there has also been significant interest in a particular class of multi-robot system known as a *robotic swarm*. Each individual in the swarm obeys a simple control law that allows interactions only with the robot neighbors. Local interactions between individual robots and their communication and sensing neighbours within a robotic swarm lead to *emergent collective behaviors*. Swarms exhibit various behaviors, such as flocking, rendezvous, dispersion. The swarm control law aims to keep the swarm coherent, meaning that the robots execute the commanded behavior and do not disconnect from the swarm and avoid collisions with one another. One way to control a swarm is via the selection of a *leader* robot. The leader robot is given a command by the human operator to perform a particular motion (e.g., move forward). The other robots align their headings to the leader heading (or an average heading that is dominated by the leader heading) so the flocking behavior is realized. Robot swarms are inherently robust and scalable because robots can fail or be added and removed with minimal system reconfiguration required to keep the swarm operational. Robotic swarms have enormous potential in applications including search and rescue, environmental monitoring and exploration, environmental cleanup. Prior work has considered supervisory control of *remote* swarms and associated issues such as bandwidth constraints and latency [12], control input propagation [13], input timing [14] and intelligibility of swarm motion to humans [15].

This paper considers an interaction method based on a wearable gesture control interface for use in scenarios involving humans working alongside or in close proximity to swarms of coordinating ground robots. Such interaction is very suitable for soldiers in an area of operations where voice or long-range radio communications may not be appropriate due to the presence of adversaries. Gesture control [16] [17] is of specific interest because it enables operators to interact with robots within their spatial vicinity in an intuitive fashion without the use of an additional device (e.g., laptop, joystick) that may prevent the operator from using hands for other tasks. In contrast to prior work that considered collaborative recognition of gestures [18] [19], communication via gestures drawn on tablets [17] or vision-based gesture recognition [16] for control of swarms, we use a wearable gesture recognition device combined with our gesture translation interface, which translates gestures to a variety of swarm commands. Wearable gesture recognition devices, such as the Myo we use in our work, recognize muscle signals of the operator and map them to particular gestures. The advantage of this technique is that it obviates the need to instrument the environment (which is necessary for vision-based recognition), thus making it useful in any environment, such as indoor or outdoor, despite varying visibility (e.g., occlusions) and lighting. The disadvantage and challenge is that wearable gesture recognition must rely on sensors (e.g., accelerometer, electromyography) without direct human interpretation for input. Consequently, discriminability of gestures by both the system and the human, rather than intuitiveness of the gestures, must be the primary criterion for gesture selection. Current wearable gesture devices have only a limited number of gestures that provide good discriminability. Increasing the number of gestures, thus providing a richer command vocabulary, would decrease the discriminability. In this research, we investigate whether wearable gesture control with a limited number of gestures can be made suitable for the complex interactions needed to command a robotic swarm.

The main contributions of this work are (a) the design of an interaction method using a gesture-based wearable device with a limited number of gestures for robust control of a complex system, a robotic swarm, and (b) an experiment on a real robot swarm comparing system performance in a one- vs two-operator scenario. In Section II, we examine the related work. In Section III, we describe the robotic swarm system. In Section IV, we describe our gesture-based interaction method. We outline the experimental design in Section V, followed by the experimental results in Section VI and a discussion in Section VII. Finally, in Section VIII, we state our conclusions and ideas for future work.

II. Related Work

Over the years, researchers have focused on developing multi-operator control of multi-robot teams. These methods typically either assign a subset of the robots to each operator [4] or assign operators jointly to control of the whole team [20]. Results have been equivocal with some researchers reporting advantages for joint control [1] [20] and others [4] [21] finding better performance from operators assigned responsibility for subsets of robots. Separability of function appears to be a key to this difference with studies in which operators performed clearly distinguished tasks [1] [20] benefiting from joint control while those with less well differentiated tasks [2] [4] [21] suffered from diffusion of responsibility.

In naturalistic settings, where tasks are clearly separable, responsibility is often allocated by function. Flocks of sheep and other domesticated animals are commonly controlled by a shepherd through use of collaborating agents, such as sheep dogs, who maintain coherence within the herd while the shepherd determines the overall direction of the flock [22]. Whether this division of labor is inherent to the task and extends to control of robotic swarms as well or is simply an artifact of the shepherd's inability to control the periphery of the swarm is an empirical question. The selection of a leader and choice of swarm behaviors are less clearly separable tasks and might either benefit from division of labor or impose coordination and communication overheads outweighing the advantages of a second operator. In this study, we investigated this issue by comparing two experimental conditions: (1) where swarms were controlled in both heading and coherence (coherently performing the commanded behavior) by a single operator with (2) swarms controlled in heading by a "shepherd" and for coherence by a second operator playing the role of the sheep dog.

III. Robotic Swarm System

Our robot swarm consists of five TurtleBots running the Robot Operating System (ROS Indigo) under Ubuntu 14.04 LTS. Each TurtleBot is outfitted with a USB-controlled LED (BlinkStick) and with an AprilTag in a known location on its body so that the TurtleBot can be identified and visually tracked via a set of overhead cameras. AprilTags [23] are a visual fiducial system that encode data into a pattern of white and black squares on a grid. In our case, we encode the unique identifier (UID) for the robot into the tag. The use of this fiducial system enables us to identify and track each TurtleBot in the swarm using a set of overhead monocular cameras. Note that the AprilTags and overhead localization for the TurtleBots were used to collect experimental data. In a real world deployment, such instrumentation of the environment would not be present.

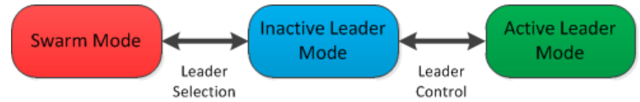


Figure 1. Each robot may be in one of three modes of operation: Swarm Mode, Inactive Leader Mode or Active Leader Mode. Only one robot may be leader (inactive or active) at any time.

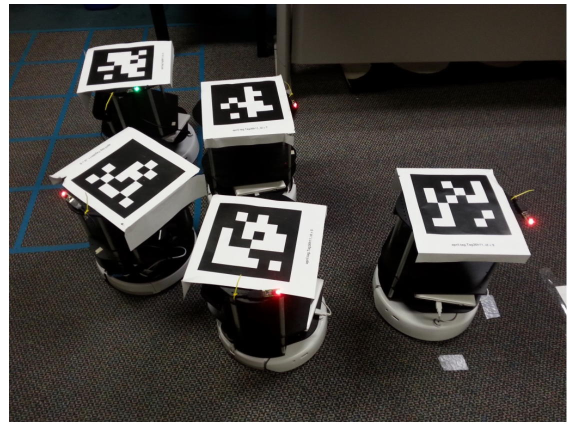


Figure 2. The TurtleBot swarm always has one leader. The leader may be an Active Leader (green LED, shown) or an Inactive Leader (blue LED).

The USB-controlled LED enables the human operator(s) working alongside the robot to identify its current mode of operation (Figure 1). Each individual robot in the swarm may be in one of three modes of operation. When the LED is red, the robot is in Swarm Mode, obeying local control laws based on the selected swarm behavior (discussed later) and the poses of other robots within its spatial neighborhood. When the LED is green (Figure 2), the robot is in Active Leader Mode and can only be directly controlled by a human operator (i.e., it ignores the behavior of other swarm robots). When the LED is blue, the robot is in Inactive Leader Mode and behaves similarly to a robot in Swarm Mode, but may be switched to an Active Leader when desired by the operator.

For our swarm, only one robot may be a leader (whether inactive or active) at any given time, but the operator may select a different leader robot during operation. The distinction between Inactive Leader Mode and Active Leader Mode is one that is practically useful to the operator. For example, consider the situation where the operator would like to select a different leader before making the decision to directly influence the leader. In our system, the operator would (1) switch the current Active Leader to Inactive Leader Mode, (2) select a new Inactive Leader and (3) only switch the new Inactive Leader to

Active Leader when desired. Until then, the operator has visual confirmation (blue LED) that the desired Inactive Leader has been selected without actually influencing that robot.

A. Robot Dynamic Model

Each robot in the swarm has the following dynamic model, where x^i, y^i and θ^i represent the position and orientation of robot i. The control inputs to the robot are given by u_v^i and u_ω^i, which represent the commanded linear velocity and commanded angular velocity respectively.

$$\dot{x}^i = u_v^i \cos\left(\theta^i\right), \quad \dot{y}^i = u_v^i \sin\left(\theta^i\right), \quad \dot{\theta}^i = u_\omega^i \tag{1}$$

For convenience, we define the position vector $\mathbf{p}^i \in \mathbb{R}^2$, which is the xy-coordinates of the robot as given above, and bearing vector $\mathbf{b}^i \in \mathbb{R}^2 : \|\mathbf{b}^i\| = 1$, which is a unit vector in the heading direction θ^i. The function $\phi(\mathbf{v}_1, \mathbf{v}_2)$ finds the smallest angle required to rotate from $\mathbf{v}_1$ to $\mathbf{v}_2$. It is assumed that the sign is adjusted accordingly (deliberately omitted from equation below for simplicity of exposition).

$$\begin{aligned} \mathbf{b}^{ij} &= \frac{\mathbf{p}^j - \mathbf{p}^i}{\|(\mathbf{p}^j - \mathbf{p}^i)\|_2} \\ \phi(\mathbf{v}_1, \mathbf{v}_2) &= \cos^{-1}\left(\frac{\mathbf{v}_1^T \mathbf{v}_2}{\|\mathbf{v}_1\|_2 \|\mathbf{v}_2\|_2}\right) \end{aligned} \tag{2}$$

B. Swarm Behaviors

Our swarm obeys a variety of behaviors chosen by the operator. These behaviors are outlined below. Note that the control law given in [14] was used for the formation behaviors.

1) Stop Moving (Default): Robots stop moving.

$$\forall i : u_v^i = 0, \quad \forall i : u_\omega^i = 0 \tag{3}$$

2) Move Forward: Robots move forward (in their local coordinate frame) with a constant linear velocity V.

$$\forall i : u_v^i = V, \quad \forall i : u_\omega^i = 0 \tag{4}$$

3) Move Backward: Robots move backward (in their local coordinate frame) with a constant linear velocity V.

$$\forall i : u_v^i = -V, \quad \forall i : u_\omega^i = 0 \tag{5}$$

4) Turn Clockwise: Robots rotate clockwise with a constant angular velocity Ω.

$$\forall i : u_v^i = 0, \quad \forall i : u_\omega^i = -\Omega \tag{6}$$

5) Turn Anti-Clockwise: Robots rotate anti-clockwise with a constant angular velocity Ω.

$$\forall i : u_v^i = 0, \quad \forall i : u_\omega^i = \Omega \tag{7}$$

6) Flocking: Regions around each robot are divided into three zones for flocking as shown in Figure 3. Each robot experiences a virtual attractive force toward robots in its "attraction" zone, a virtual torque to align heading with neighbour robots in the "alignment" zone and a virtual repulsive force away from robots in the "repulsion" zone. Strictly speaking, we use a first-order model, so the virtual forces and torques are actually virtual biases in the linear and angular velocity control inputs u_v and u_ω. Note that all three zones were

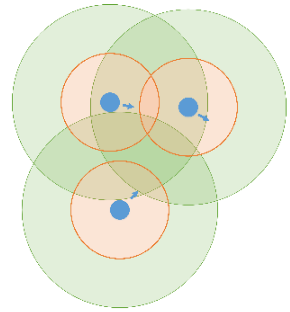


Figure 3. Zones of attraction (all white), heading alignment (green) and repulsion (red) for each robot (blue).

(a) Double Tap (b) Fist (c) Spread Fingers (d) Wave In (e) Wave Out

Figure 4. The five gestures recognized by the Myo as shown on the Thalmic Labs website [24].

implemented on the robots, but for simplicity of exposition, only the attraction portion of the control law is given below.

$$\begin{aligned} \forall i : u_v^i &= \frac{K_v}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \|(\mathbf{p}^j - \mathbf{p}^i)\|_2 \\ \forall i : u_\omega^i &= K_\omega \phi\left(\frac{\sum_{j \in \mathcal{N}(i)} \mathbf{b}^{ij}}{\|\sum_{j \in \mathcal{N}(i)} \mathbf{b}^{ij}\|_2}, \mathbf{b}^i\right) \end{aligned} \tag{8}$$

7) Align Heading: Robots align heading with other robots within their spatial neighborhood.

$$\forall i : u_v^i = 0, \quad \forall i : u_\omega^i = \frac{K_\omega}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \phi(\mathbf{b}^i, \mathbf{b}^j) \tag{9}$$

8) Line Formation - X Direction: Robots move into a line formation along the world frame's X direction.

9) Line Formation - Y Direction: Robots move into a line formation along the world frame's Y direction.

10) Circle Formation: Robots move into a circle formation.

IV. Gestures for Robot Swarm Control

In this section, we describe the wearable gesture recognition devices we used and our gesture-based interaction method.

A. Wearable Gesture Recognition Devices

In our system, each of the human operators for the robot swarm wears one or more gesture recognition devices and receives feedback about swarm state via LEDs on the robots. For the gesture recognition device, we chose to use an off-the-shelf device called the Myo developed by Thalmic Labs [24]. The Myo is an armband that combines eight EMG sensors for measuring myoelectric muscle signals with a nine-axis IMU (three-axis gyroscope, three-axis accelerometer and three-axis magnetometer) to detect hand gestures. There are five standard hand gestures shown in Figure 4 that were used to develop the gesture-based interaction for our experiment: *Double Tap*, *Fist*, *Spread Fingers*, *Wave In* and *Wave Out*. The device can also provide simple haptic feedback via vibration, but that feature was not used in this work. Data is received from the device via Bluetooth.

B. Gesture-Based Interface

Our design for robust interaction of robotic swarms, given a limited number of gestures whose recognition is error prone, relies on three key ideas: (a) constructing a rich vocabulary of commands, a command grammar, out of a small number of gestures, (b) providing safeguards against errors in gesture recognition and (c) a gesture-based "virtual menu" that allows selection of robots as leaders. These ideas were operationalized. The first idea was implemented by giving different semantic mappings to the same gesture (semantic overloading) depending on context (role). The second idea was operationalized by a multi-step process, rather than letting the operator switch from one behavior to another on the fly, thus running the risk of poor signal recognition. For example, in the two-operator case, where the heading is given, the process is stop, select robot (and mode), give heading command. The third idea is implemented by mapping the "wave in" and "wave out" operations to a selection action for selecting the next behavior or the next robot depending on the role and whether there was a single or multiple operators (see swarm behavior selection section for examples). Albeit complicated in the single operator case, this design allowed for robust operation as shown by our experimental results. Note that in this work, we deliberately made the control task challenging (small crowded room, box obstacles) so the operator would be continuously occupied with guiding the robots through the environment. In a real outdoor environment, the operator would spend less time engaged in controlling the swarm with gestures.

Our gesture-based interface for operating the swarm is divided into two distinct roles: (1) Swarm Behavior Selection and (2) Swarm Leader Selection. Each role is allocated to a different Myo armband and all gestures recognized by that armband are interpreted in the context of the associated role. In this way, one operator can wear both armbands and try to fulfill both roles or two operators can each wear an armband and only fulfill their own role.

As already discussed, there is a trade-off between the intuitiveness of the interface and the discriminability of the gestures. Although work can be done to develop an interface that includes more intuitive gestures for the various robot leader and behavior mappings, there are many cases where these gestures are more difficult for the sensor to characterize. Our gesture-based interaction was designed to be as intuitive as possible while still keeping the limited number of five standard gestures, leveraging the higher discriminability of these gestures.

1) Swarm Behavior Selection: In this role, the operator selects a swarm behavior that will be followed by all robots in Swarm Mode. The operator is given a circular list (the virtual menu) of the swarm behaviors described in Section III-B. The *Wave In* and *Wave Out* gestures allowed the operator to select the previous or next behavior in the list respectively. This mapping leverages most operators' prior experience with swiping left or right on a touch screen device to switch between screens. Once the operator moves to the desired

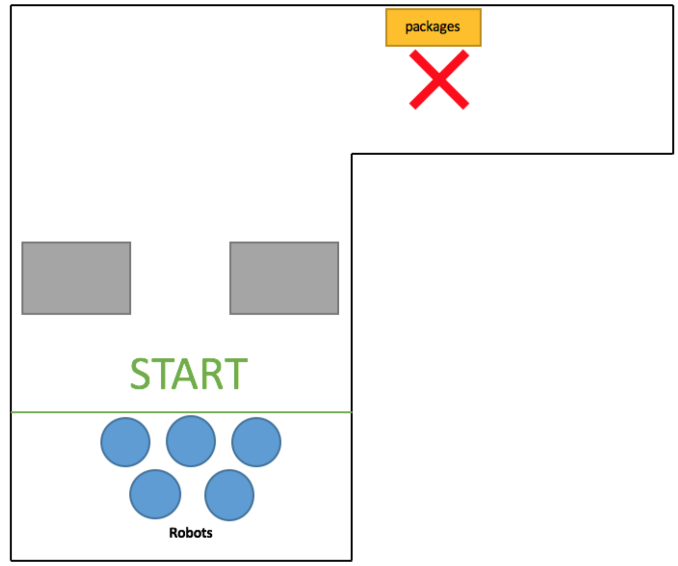


Figure 5. Diagram of environment used for trials. Robots begin in the region indicated by Start and packages are located in the region indicated by red X.

swarm behavior in the menu they can activate it by performing the *Spread Fingers* gesture. Unless the behavior is activated, the swarm will not follow the selected behavior. The *Spread Fingers* gesture is intended to correlate the fingers moving away from the palm to the robots moving away from their starting position via the activated swarm behavior. The *Fist* gesture is commonly used for signifying someone or something to hold or wait. Therefore, it is mapped to the default behavior ('Stop Moving'). For this role the operator ignores the *Double Tap* gesture.

2) Swarm Leader Selection: In this role, the operator selects an Inactive Leader (indicated by blue LED) or controls an Active Leader (indicated by green LED) within the swarm. Assume there is an implicit circular list of robot UIDs ordered numerically. When the selected robot is in Inactive Leader Mode, the *Wave In* gesture selects the previous robot on the list and the *Wave Out* gesture selects the next robot on the list. The *Wave In* and *Wave Out* gestures were chosen for moving through the list of UIDs for the same reason they were chosen to select a swarm behavior. The *Double Tap* gesture switches the selected robot from Inactive Leader Mode to Active Leader Mode or vice versa. A *Double Tap* gesture was used instead of a *Spread Fingers* gesture so as to distinguish between choosing a robot and choosing a movement. Once a new leader robot is chosen with the *Double Tap* gesture and is placed in Active Leader Mode, the *Wave In* gesture causes the chosen leader robot to rotate counterclockwise, the *Wave Out* gesture causes the robot to rotate clockwise. These mappings are chosen such that the direction of the operator's hand when they perform the gesture corresponds to the direction of rotation of the leader robot. The *Spread Fingers* gesture once again signifies a robot movement away from its current location – in this case a forward movement. A *Fist* gesture causes the robot to stop moving for the same reason as in the swarm behavior selection.

V. Experimental Design

Our experiment investigates the resulting performance in two conditions: one-operator condition where the operator

Figure 6. Actual environment used for trials with robots in their initial locations. The large boxes are obstacles.

performs both roles versus two-operator condition where each operator has a single role, but the two must coordinate. In other words, we want to compare the difficulty and workload resulting from these two conditions. This investigation is conducted in a package retrieval scenario that requires operators to guide the swarm to a target location, retrieve packages and return to the start location. Two conditions were tested: two operator and one operator. In the two operator condition, the subject would wear one armband and fulfill the Swarm Behavior Selection role and an experienced human operator (same person for all two operator experiments) would wear another armband and fulfill the Swarm Leader Selection role. Both operators wore their respective armbands on their right arm. For the second condition (one operator condition), the subject would wear both armbands and fulfill both the Behavior Selection and Leader Selection roles simultaneously. The armband associated with the Leader Selection role was always worn on the right arm of the subject and the armband associated with the Swarm Behavior Selection was worn on the left arm during the one operator trials.

Each subject participated in both the two operator and one operator trials. All subjects participated in the two operator trial first (with the experienced operator partner) before participating in the one operator condition trial. This experimental design was used to mitigate any bias due to possible learning effects. In this setup, if the subjects did get better over time, it would only help their performance in the one operator trial. The order was also chosen to reduce any bias the two operator condition trial had because of the experienced operator's performance.

Before each trial the subject was instructed to wear the Myo armband(s) and allow the sensors to warm up. When the armband(s) was/were warm enough (5-10 minutes) the subject would then sync the armband(s) with the system using the *Wave Out* gesture. Once the armband(s) was/were properly synced the subjects were allowed to practice the 5 standard gestures. As soon as the subject was comfortable performing all gestures with each armband they were wearing, the researchers explained the complete mapping of the gestures to the robot behavior and leader selection. No additional training was done. The subjects did not try the behavior and/or leader selection before the trial began.

Figure 7. Leader selection participant (left) and the behavior selection participant (right) during a double operator trial. Each participant is wearing a Myo on their right arm.

A. *Experimental Task*

The experimental task was a package retrieval task and the same experimental task was used across all trials in both the one operator and two operator conditions. There were three distinct parts to each trial. First, the operator(s) was/were required to use the gestures to guide the swarm of TurtleBots from the start region to the region indicated by the red X in Figure 5. Second, the operator(s) was/were required to load packages (small boxes) onto each swarm robot. Finally, the operator(s) was/were required to guide the swarm back to the start region. The trials began when the operators began using the gestures to move the TurtleBots and ended when the TurtleBots made it back to the start area. Large boxes were setup in the environment as obstacles through which the operators guided the robots in order to reach the package pickup location. For a sense of scale, the actual environment used for trials is shown in Figure 6. The traversable area of the main portion of the environment containing obstacles was approximately 20 feet long and 14 feet wide. The traversable area of the narrow corridor was approximately 4 feet wide and 10 feet long. The obstacles were placed 4.25 feet apart. For comparison, an individual TurtleBot is 1 foot in diameter. Both operators were located within visual range of the robots and audible range of each other such that they could coordinate commands sent to the robots (Figure 7).

The human operators faced several challenges navigating through the environment including boxes which narrow the passage, as well as turning into the narrow corridor. In addition, there were several "blind spots" on the map where overhead cameras could not detect AprilTags, so TurtleBots in Swarm Mode would behave erratically due to incorrect localization. TurtleBot wheel odometry and gyroscope were intentionally not used to correct for this effect, so that operator(s) would be forced to intervene to correct for the effect.

B. *Participants and Data Collection*

There were a total of 8 participants in the experiment (excluding the experienced operator). Participants were all graduate students (five male, three female). Only one of the subjects had previous experience using the Myo armband and none of the participants had previously used the armband to control our robotic swarm. Data was recorded using the 'rosbag' program included with ROS. Recorded data included AprilTag poses detected by overhead cameras, gestures detected by the armbands and the active behavior and leader for the swarm. Videos of each trial were also recorded.

VI. Results

The data collected are presented in terms of performance measures and operator measures below. Performance measures include (1) the distance traveled by the robots – maximum of any robot and total of all robots – and (2) the average dispersion from the calculated centroid of the robots' location. Operator measures are presented for both the behavior selection armband and the leader control armband. The measures included are (1) the total number of gestures recognized by the armband, (2) the number of extraneous gestures performed, (3) the count of activated behaviors and leader robots, (4) the total time spent on each behavior and leader robot, and (5) the average time spent each time a behavior or leader robot is chosen. Data were analyzed using the number of roles simultaneously fulfilled by the subject as a factor.

A. Performance Measures

1) Completion Time: Figure 8 shows a plot of trial completion times for each subject under the two operator and one operator conditions. Most subjects were able to complete trials in the two operator condition significantly more quickly than in the one operator condition. The sample means for the two operator and one operator condition trial completion times were 12.115 minutes and 23.806 minutes, respectively. The difference in means was significant at the $p < 0.05$ level ($F_{(1,17)} = 8.082$, $p = 0.011$).

2) Distance Traveled: The sample means for the total distance traveled by all TurtleBots was 127.21 meters and 190.12 meters for the two operator and one operator trials respectively. Figure 9a shows the total distance traveled by all robots in each trial. The sample means for the maximum distance traveled by any TurtleBots was 27.75 meters for the two operator trials and 49.44 meters for the one operator trials. Figure 9b shows the maximum distance traveled by a TurtleBot in all trials. Both differences were significant at the $p < 0.05$ level ($F_{(1,17)} = 6.652$, $p = 0.02$; $F_{(1,17)} = 4.525$, $p = 0.048$ respectively). The average dispersion of the TurtleBots throughout the trials was calculated by summing the average squared euclidean distance between each TurtleBot and the centroid at each time point. The two operator trials had an average dispersion of 66.4 centimeters while the one operator trials had an average of 73.4 centimeters.

B. Operator Measures

Figures 9c and 9d show the total number of recognized gestures performed using the behavior selection armband and leader selection armband respectively. The two operator trial results are shown in blue and the one operator trial results are shown in red. The number of gestures performed with the leader armband was significant at the $p < 0.01$ level ($F_{(1,17)} = 8.599$, $p = 0.009$).

The number of extraneous gestures made by subjects using the Behavior and Leader Selection armbands were calculated. In Behavior Selection, an extraneous gesture was characterized as one of the following two cases: (1) the gesture repeated a gesture immediately before it, or (2) the operator performed an unmapped gesture, which in this case was the *Double Tap* gesture. A gesture was characterized as extraneous when using the Leader Selection armband if one of the following two cases

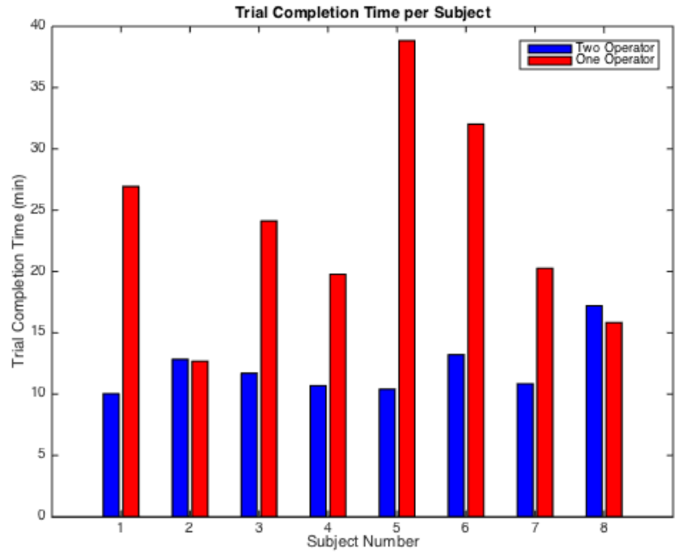


Figure 8. Trial completion times for each subject under both the two operator and one operator conditions.

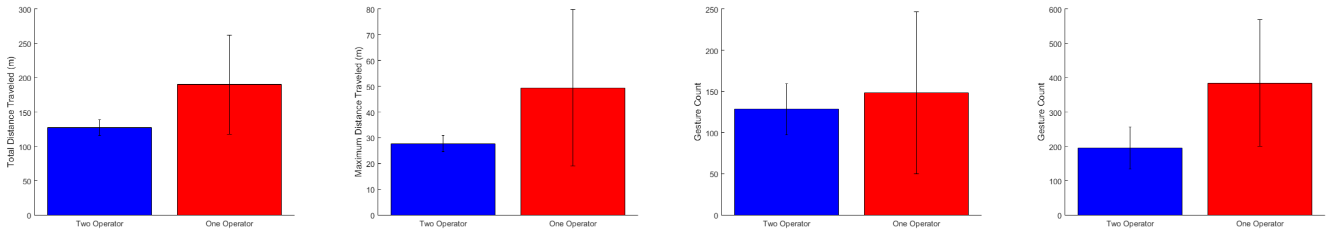
was true: (1) the gesture repeats a gesture immediately before it or (2) the gesture is an unmapped gesture in the Inactive Leader mode. In the Inactive Leader mode (while an operator is switching to a new leader robot), the *Fist* and *Spread Fingers* gestures are unmapped. The two operator results are from the expert user while the one operator's are from the subject. Subjects using the behavior armband performed an average of 48.67 and 55.40 extraneous gestures for the two operator and one operator conditions respectively. An average of 126.89 extraneous gestures were performed using the leader armband in the two operator condition and 251.70 for the one operator condition. The number of extraneous gestures performed with the leader armband was significant at the $p < 0.05$ level ($F_{(1,17)} = 6.157$, $p = 0.024$).

On average subjects selected the 'Stop Moving' behavior the most followed by 'Flocking' (Figure 11). Subjects spent an average total time of 7.49 minutes and 15.01 minutes on the 'Stop Moving' behavior in the two operator and one operator condition respectively (Figure 12). The average time spent on the 'Stop Moving' behavior each time it was selected was 21.54 seconds in the two operator condition and 38.75 seconds in the one operator condition. Both the total time and average time spent on the 'Stop Moving' behavior were significant at the $p < 0.05$ level ($F_{(1,17)} = 7.292$, $p = 0.015$; $F_{(1,17)} = 4.971$, $p = 0.04$ respectively).

C. Correlations

High correlations were found between operator behavior and performance measures at the $p < 0.01$ significance level. Figure 10 shows that completion time was found to correlate to the 'Stop Moving' behavior, the number of gestures performed using the Behavior Selection armband, the total distance traveled by all TurtleBots, and the maximum distance traveled by a single TurtleBot ($r = 0.987$, $r = 0.709$, $r = 0.949$, and $r = 0.886$ respectively). The number of gestures performed with the Behavior Selection armband and number of errors occurring with the Leader Selection armbands were highly correlated with the time spent in the 'Stop Moving' behavior ($r = 0.703$ and $r = 0.935$ respectively). The total time spent

 ISBN: 978-1-61208-538-8



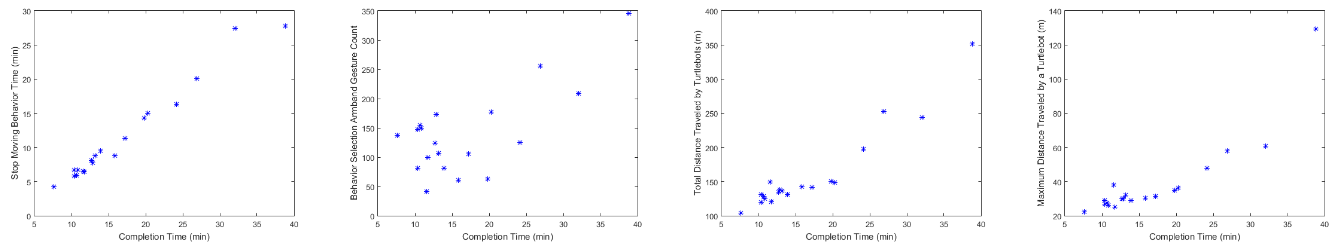
(a) Average total distance traveled by all TurtleBots per trial.

(b) Average maximum distance traveled by any TurtleBot per trial.

(c) Average total number of recognized gestures performed with behavior armband per trial.

(d) Average total number of recognized gestures performed with the leader armband per trial.

Figure 9. Average measures for single and double operator conditions per trial.



(a) Correlation between time spent on the 'Stop Moving' behavior vs. trial completion time.

(b) Correlation between total number of gestures performed using Behavior Selection armband vs. trial completion time.

(c) Correlation between total distance traveled by the TurtleBots vs. trial completion time.

(d) Correlation between the maximum distance traveled by a single TurtleBot vs. trial completion time.

Figure 10. Correlations

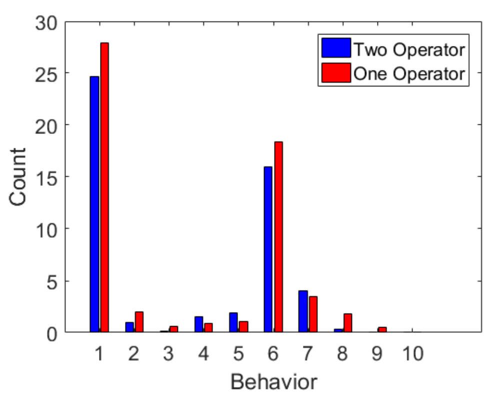


Figure 11. The average number of behavior selections. They are numbered according to Section III-B.

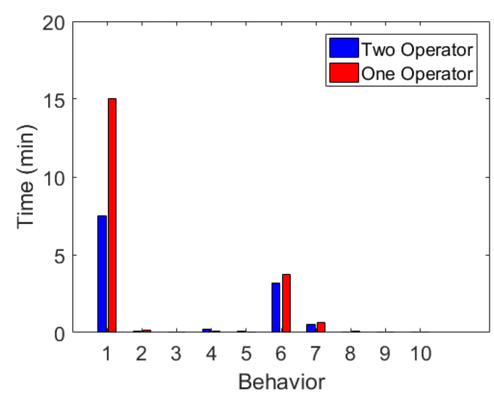


Figure 12. The average total time spent on each behavior. They are numbered according to Section III-B.

on the 'Stop Moving' behavior was also highly correlated with the total distance traveled by all TurtleBot and the maximum distance traveled by a single TurtleBot ($r = 0.737$ and $r = 0.728$ respectively).

Among measures of command usage only the average times spent stopped and flocking were significantly correlated ($r = 0.68, p = 0.001$). As shown in Figure 13, counts of commands and errors were intercorrelated with errors in behavior selection closely paralleling the number of behavior selections while leader selection errors followed the number of leader selections. Behavior and Leader Selection errors were uncorrelated, however, each was correlated with the number of commands of the other type. This pattern shows strong interaction between the behavior selection and leader control task suggesting that they may not be fully separable.

The time spent stopped was the only aspect of a command showing strong correlations with performance. It was correlated with Total Distance ($r = 0.916, p < 0.001$), Maximum Distance ($r = 0.836, p < 0.001$), Completion Time ($r = 0.987, p < 0.001$) and marginally correlated with Average Dispersion ($r = 0.43, p = 0.066$). The Average Time Stopped was correlated only with the Average Dispersion

ISBN: 978-1-61208-538-8

		Behavior		Leader	
		Errors	**Total**	**Errors**	**Total**
Behavior	**Errors**				
	Total	0.853			
Leader	**Errors**	-	0.622		
	Total	0.464	0.610	0.984	
$R = 0.456, p = 0.05, r = 0.693, p = 0.001$					

Figure 13. Correlations between errors and total commands for behavior selection and leader selection.

	Behavior		Leader	
	Errors	**Total**	**Errors**	**Total**
Total Distance	0.687	0.788	0.866	0.870
Max Distance	0.602	0.779	0.817	0.786
Completion Time	0.598	0.709	0.906	0.930
$R = 0.575, p = 0.01, r = 0.693, p = 0.001$				

Figure 14. Correlations between commands issued for behavior selection or leader selection and subject performance according to total distance traveled by all robots, maximum distance traveled by any robot, trial completion time.

($r = 0.602, p < 0.01$). These relations suggest that longer times spent stopped either in selecting new behaviors or deciding on a course of action was strongly predictive of poor performance on other aspects of the task. Figure 14 shows a pattern of strong correlations between aggregate measures of command usage and performance. As with use of the Stop command, larger numbers of commands and errors led to poorer performance on each of the measures.

VII. Discussion

Subjects performed significantly better in the two operator condition, where they were only required to fulfill the Behavior Selection role, than in one operator condition, where they were required to fulfill both Behavior Selection and Leader Selection roles. This suggests that for our interface, the responsibilities for a role have been selected appropriately to match the level of effort required for a human to perform effectively in that role. Attempting to perform multiple roles seems to have a significantly detrimental effect on human operator performance. In addition, the results demonstrate the viability of the implemented gesture-based interface in controlling a robotic swarm since all operator(s) successfully completed the experimental task without driving the robotic swarm into an unrecoverable state, as was frequently observed during initial informal tests with only the Behavior Selection role.

Although the subjects participated in the two operator trial before the one operator trial, they performed better in the two operator trial. This could be attributed to the experience of the expert outweighing learning effects seen in the subjects' performance, although it cannot be said for certain that this is the sole cause. The operators in the two operator trials seemed to effectively communicate throughout the trial which was apparent in the lower total and maximum distances traveled, and the number of extraneous gestures performed with the behavior armband. For most subjects, the total number of gestures performed with the behavior armband in the one operator trial was close to or lower than the number performed in the two operator trials. This can be attributed to the learned effects seen from the subjects first participating in the two operator trial. However, the number of gestures was less in the two operator condition when the subjects were not required to split their attention between the behavior and leader selection.

The 'Stop Moving' behavior was selected most by subjects. Observations saw that subjects made this behavior the prerequisite for selecting all other movement behaviors. In many instances subjects selected the 'Stop Moving' behavior while they walked around and inspected the TurtleBots' current positions so as to determine the next command(s) required to accomplish their goal. This in between behavior planning stages were almost double in the one operator trials than in the two operator trials as seen by the average time spent on the 'Stop Moving' behavior each time it was selected. The second most used behavior was the 'Flocking' behavior. Subjects seemed to prefer the efficiency of coordinated movement seen when the TurtleBots followed the direction and movements of the leader robot controlled by the leader armband. This significantly reduced the commands that would be necessary to send to the TurtleBots with more primitive commands like 'Move Forward', 'Move Backward', 'Turn Clockwise', and 'Turn Anti-Clockwise.'

The high correlations between the completion time and (1) the 'Stop Moving' behavior, (2) the number of gesture performed using the Behavior Selection armband, (3) the total distance traveled by all TurtleBots, and (4) the maximum distance traveled by any TurtleBot suggests that excessive time spent in the stopped state generating erroneous gestures coincided with poor control and excessive travel for the TurtleBots. The one operator trials sent commands less efficiently to the TurtleBots and stopped more often to make decisions. The lower control resulted in larger total distances and maximum distances traveled in the one operator condition than the two operator condition.

Although the current interface seems to effectively allow operators to control the current swarm size, as additional robots are added to the swarm the interface as it stands now may cause the operators to reach their workload maximum. A more intuitive interface would allow the operators to send commands to the robots with added efficiency. By further expanding the interface to provide a way to control a subset group of robots, the system will be able to provide a more scalable solution for robot swarm control.

VIII. Conclusion and Future Work

In this paper, we investigated an interaction method using a gesture-based interface that used a very limited set of gestures for robust control of a complex system, namely a robotic swarm. The approach was based on 3 key ideas: (a) constructing a rich vocabulary of commands, a command grammar, out of a small number of gestures, (b) providing safeguards against errors in gesture recognition and (c) a gesture-based "virtual menu" that allows selection of robots as leaders. The gesture-based interface incorporated multiple roles (i.e., Behavior Selection, Leader Selection) and each gesture recognition device was associated with a different role. Our experiments indicated that human operators performed significantly better when each operator was fulfilling one role than with one operator fulfilling both roles. Single operators performing both the Behavior and Leader Selection roles tended to be less efficient with their control of the Turtle-Bots, which resulted in larger total distances traveled by all TurtleBots and maximum distance traveled by any TurtleBot.

The results also demonstrated the viability of the gesture-based interface in enabling human operators to robustly control robotic swarms in proximal interactions.

In future work, we plan to run a larger number of subjects in multiple trials for each subject to study performance improvement with experience. We also plan to perform sensitivity analysis, varying system parameters (e.g., number of robots, course layout). Additionally, we will explore methods for gesture-based control of subsets of swarms to provide a more scalable solution.

ACKNOWLEDGEMENTS

This research has been sponsored in part by AFOSR Grant FA9550-15-1-0442 and an NSERC PGS D scholarship.

REFERENCES

[1] J. M. Whetten, M. A. Goodrich, and Y. Guo, "Beyond robot fan-out: Towards multi-operator supervisory control," in *Systems Man and Cybernetics (SMC), 2010 IEEE International Conference on*. IEEE, 2010, pp. 2008–2015.

[2] F. Gao, M. L. Cummings, and L. F. Bertuccelli, "Teamwork in controlling multiple robots," in *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*. ACM, 2012, pp. 81–88.

[3] M. Lewis, H. Wang, S. Y. Chien, P. Velagapudi, P. Scerri, and K. Sycara, "Process and performance in human-robot teams," *Journal of Cognitive Engineering and Decision Making*, vol. 5, no. 2, pp. 186–208, 2011.

[4] M. Lewis *et al.*, "Teams organization and performance in multi-human/multi-robot teams," in *Systems Man and Cybernetics (SMC), 2010 IEEE International Conference on*. IEEE, 2010, pp. 1617–1623.

[5] F. Gao, M. L. Cummings, and E. T. Solovey, "Modeling teamwork in supervisory control of multiple robots," *Human-Machine Systems, IEEE Transactions on*, vol. 44, no. 4, pp. 441–453, 2014.

[6] S. Teller *et al.*, "A voice-commandable robotic forklift working alongside humans in minimally-prepared outdoor environments," in *2010 IEEE International Conference on Robotics and Automation*, May 2010, pp. 526–533.

[7] M. Lewis, "Human interaction with multiple remote robots," *Reviews of Human Factors and Ergonomics*, vol. 9, no. 1, pp. 131–174, 2013.

[8] E. Olson *et al.*, "Progress toward multi-robot reconnaissance and the magic 2010 competition," *Journal of Field Robotics*, vol. 29, no. 5, pp. 762–792, 2012.

[9] J. Y. Chen, E. C. Haas, and M. J. Barnes, "Human performance issues and user interface design for teleoperated robots," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 37, no. 6, pp. 1231–1245, 2007.

[10] S. Nunnally, P. Walker, M. Lewis, N. Chakraborty, and K. Sycara, "Using haptic feedback in human robotic swarms interaction," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 57, no. 1. SAGE Publications, 2013, pp. 1047–1051.

[11] T. Setter, A. Fouraker, H. Kawashima, and M. Egerstedt, "Haptic interactions with multi-robot swarms using manipulability," *Journal of Human-Robot Interaction*, vol. 4, no. 1, pp. 60–74, 2015.

[12] P. Walker, S. Nunnally, M. Lewis, A. Kolling, N. Chakraborty, and K. Sycara, "Neglect benevolence in human control of swarms in the presence of latency," in *Systems, Man, and Cybernetics (SMC), 2012 IEEE International Conference on*. IEEE, 2012, pp. 3009–3014.

[13] S. A. Amraii, P. Walker, M. Lewis, N. Chakraborty, and K. Sycara, "Explicit vs. tacit leadership in influencing the behavior of swarms," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, 2014, pp. 2209–2214.

[14] S. Nagavalli, L. Luo, N. Chakraborty, and K. Sycara, "Neglect benevolence in human control of robotic swarms," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, 2014, pp. 6047–6053.

[15] S. Nagavalli, S.-Y. Chien, M. Lewis, N. Chakraborty, and K. Sycara, "Bounds of neglect benevolence in input timing for human interaction with robotic swarms," in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 2015, pp. 197–204.

[16] G. Podevijn, R. O'Grady, Y. S. Nashed, and M. Dorigo, "Gesturing at subswarms: Towards direct human control of robot swarms," in *Towards Autonomous Robotic Systems*. Springer, 2014, pp. 390–403.

[17] D. Perzanowski *et al.*, *Communicating with teams of cooperative robots*. Springer, 2002.

[18] A. Giusti, J. Nagi, L. Gambardella, and G. A. Di Caro, "Cooperative sensing and recognition by a swarm of mobile robots," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*. IEEE, 2012, pp. 551–558.

[19] J. Nagi, A. Giusti, L. M. Gambardella, and G. A. Di Caro, "Human-swarm interaction using spatial gestures," in *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*. IEEE, 2014, pp. 3834–3841.

[20] B. Mekdeci and M. Cummings, "Modeling multiple human operators in the supervisory control of heterogeneous unmanned vehicles," in *Proceedings of the 9th Workshop on Performance Metrics for Intelligent Systems*. ACM, 2009, pp. 1–8.

[21] T. D. Fincannon, A. W. Evans, F. Jentsch, E. Phillips, and J. Keebler, "Effects of sharing control of unmanned vehicles on backup behavior and workload in distributed operator teams," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 53, no. 18. Sage Publications, 2009, pp. 1300–1303.

[22] B. Bennett and M. Trafankowski, "A comparative investigation of herding algorithms," in *Proc. Symp. on Understanding and Modelling Collective Phenomena (UMoCoP)*, 2012, pp. 33–38.

[23] E. Olson, "AprilTag: A robust and flexible visual fiducial system," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, May 2011, pp. 3400–3407.

[24] T. Labs, "Myo gesture control armband - wearable technology by thalmic labs," 2017, [Retrieved: February, 2017]. [Online]. Available: https://www.myo.com/techspecs