# Generating Legible and Glanceable Swarm Robot Motion through Trajectory, Collective Behavior, and Pre-attentive Processing Features

LAWRENCE H. KIM and SEAN FOLLMER, Stanford University, United States

As swarm robots begin to share the same space with people, it is critical to design *legible* swarm robot motion that clearly and rapidly communicates the intent of the robots to nearby users. To address this, we apply concepts from intent-expressive robotics, swarm intelligence, and vision science. Specifically, we leverage the trajectory, collective behavior, and density of swarm robots to generate motion that implicitly guides people's attention toward the goal of the robots. Through online evaluations, we compared different types of intent-expressive motions both in terms of legibility as well as glanceability, a measure we introduce to gauge an observer's ability to predict robots' intent pre-attentively. The results show that the collective behavior-based motion has the best legibility performance overall, whereas, for glanceability, trajectory-based legible motion is most effective. These results suggest that the optimal solution may involve a combination of these legibility cues based on the scenario and the desired properties of the motion.

CCS Concepts: • **Human-centered computing** → *Empirical studies in HCI*; *Empirical studies in collaborative and social computing*;

Additional Key Words and Phrases: Legible swarm robot motion, human-multirobot interaction, glanceability

## 1 INTRODUCTION

Swarms of autonomous robots are beginning to occupy our environment. They are being used for many tasks such as transportation of goods and people [4], search-and-rescue [31], firefighting [35], and agriculture [39]. These swarms of robots allow people to monitor large areas through distributed sensing and manipulate objects both in a distributed and collective manner. Regardless of the level of the robots' autonomy, people will still play a significant role in human-swarm interaction ranging from tele-operator [2] to supervisor [11] to end-user [29, 38].

To facilitate *in situ* interaction with a large number of mobile robots, it is important to enable human observers to quickly "read" and predict what the robots are going to do. Researchers have generated legible motion for a single robot to help users feel safer [45], accomplish tasks
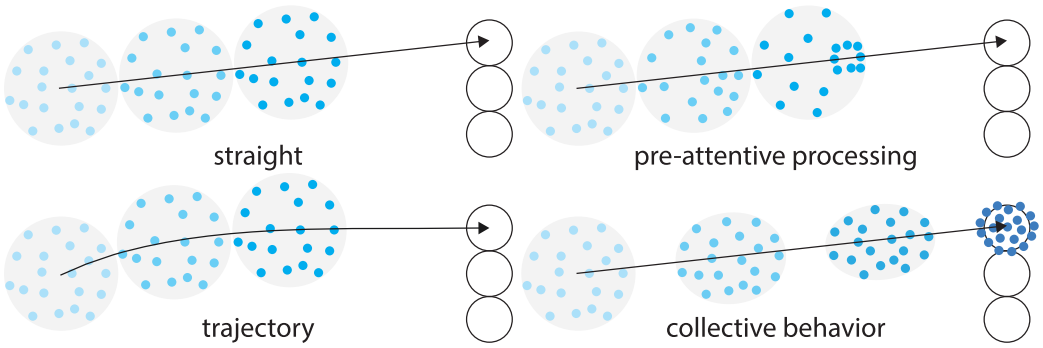
Fig. 1. We propose and evaluate different types of legible swarm robot motion based on trajectory, collective behavior, and pre-attentive processing features. The blue circles represent robots that are moving toward their goal represented by the white circles.

more efficiently [6], and potentially build trust and human-robot rapport. As we approach the the future of ubiquitous robots where people of various backgrounds will frequently encounter, collaborate, and interact with groups of robots in public and personal spaces, we need to study the notion of legibility not only for interaction with a single robot but also with groups of robots.

In this article, we introduce the concept of *glanceability* for robot motion and explore how we can improve both the *legibility* and *glanceability* for a swarm of robots. Given a large number of robots, it is unlikely that every robot will work on the same task simultaneously or that people will be solely focusing on the robots. Instead, each robot will be optimally assigned to different tasks and people will be attending to other tasks as well. In such circumstance, supervisors of these robots face a daunting job of simultaneously overseeing many groups. To lessen the complexity, it is important to generate a *glanceable* motion, which we define as being pre-attentively legible, to expedite the robot monitoring process for observers. A glanceable motion differs from a legible motion in that the former is designed to convey intent rapidly across different points along the trajectory, while the latter conveys intent early but not necessarily rapidly throughout the entire trajectory.

To design legible and glanceable swarm robot motion, we leverage findings from the legibility of a single robot's motion [16], swarm intelligence literature [8], and visual perception [47]. Specifically, we investigate how the trajectory-based legible motion that has been used for a single robot [6, 17] can be applied to a group of robots, how collective behavior (e.g., rendezvous) [41] could be used to inform the user about the goal, and how pre-attentive processing features (e.g., density) [47] could be used in designing swarm motion that rapidly directs user's attention toward the desired goal as shown in Figure 1.

The underlying mechanisms behind each of these legible motion strategies are different. Thus, we expect that certain task and robot parameters will influence these motions in a different manner. For instance, the algorithm behind trajectory-based legible motion is heavily dependent on the relative locations of the targets, while other motions, that are based on collective behavior or pre-attentive processing feature, are only dependent on the absolute location. Thus, trajectory-based motion may be more affected by how close the targets are compared to the other conditions. The size of the initial spread of the robots may also have varying impact as the rendezvous condition might benefit as the merging rate is increased, whereas for other conditions, having larger spread may only confuse observers, as it is less clear where the center of the robots is.

Thus, we evaluated and compared the legibility and glanceability performance of the different legible motions as well as the effects of target difficulty and initial spread. To investigate this,

we ran the two between-subject studies online using the Amazon Mechanical Turk platform. The study results serve as guidelines on how to generate legible and/or glanceable swarm robot motion.

In summary, the contributions of this article are:

- introduction and design of collective behavior-based and pre-attentive processing feature-based legible motions,
- introduction of the concept of a glanceable robot motion,
- two crowdsourced between-subject studies to evaluate the legibility and glanceability performances of different types of legible motion, and
- guidelines for generating legible and glanceable swarm robot motion

## 2  RELATED WORK

### 2.1  Legible Robot Motion

To interact seamlessly with people, robots need to behave in a manner that is legible and comprehensible to people. Dragan et al. introduced a mathematical approach to generate legible motion based on theories from psychology [17]. They also differentiated the concepts of predictability and legibility depending on whether the direction of the inference is "action-to-goal" or "goal-to-action," respectively. They proposed an algorithm based on the principle of rational action [15], Bayes' Rule, and optimization to generate the legible trajectory. Many researchers have followed the basic principle with adaptations to improve the performance of the algorithm [6, 51], increase legibility taking into account of different viewpoints [36], and demonstrate that its performance depends on the type of the manipulator [5]. Zhao et al. investigated the effects of different gripper orientations and found that gripper pointing towards the target in a straight trajectory has the best performance [51]. However, Bodden et al. [6] found their "point position" heuristic-based algorithm outperformed that from Dragan et al. Instead of conveying the robot's destination, Mavrogiannis et al. designed a planning framework that aims at generating motion that clearly communicates an agent's intended collision avoidance strategy [33]. While other methods have been proposed to convey intent of robots, such as Augmented Reality [42, 48], lights [3, 46], and gaze [19], these approaches either require additional equipment or humanoid robots. Thus, we chose to use an adapted version of the algorithm from Bodden et al. [6] to generate the trajectory-based legible motion for simple non-humanoid robots in our studies. For a more thorough survey of the existing work in legibility of robot behavior, refer to Reference [30].

### 2.2  Legibility of Multi-robot Motion

Recently, researchers have extended the concept of legibility to multi-robot systems. Capelli et al. investigated the effects of three different motion variables (trajectory, dispersion, and stiffness) on the legibility of a multi-robot system [10]. In a **Virtual Reality (VR)** setting, 20 virtual robots were used to conduct the study. The results demonstrated that dispersion, stiffness, and both in conjunction have significant effects on the response time. Higher dispersion and harder stiffness led to faster response times, while the minimum-jerk trajectory was more accurate than the arc-trapezoidal trajectory. Moreover, Capelli et al. ran a similar study but with multiple multi-robot systems [9]. The results indicated that the trajectory and dispersion significantly impact the prediction accuracy, while harder stiffness increased the prediction time, a result in contrast to that of their previous study.

In this article, we explore different types of legible swarm robot motion in terms of both legibility as well as glanceability, a new measure of legibility that tests whether an observer can read the intent of the robots pre-attentively. While we also adopt the trajectory-based legible motion similar to Capelli et al., we use a different algorithm. Instead of relying on concepts from animation [27]

as done in Reference [9, 10], we use an algorithm that has been proven to increase legibility [6]. We also introduce and compare the performance of additional methods to enhance legibility, such as collective behavior-based and pre-attentive processing feature-based algorithms.

### 2.3 Glanceability

Glanceable designs allow users to grasp information with only a quick look [7, 21, 32]. There are many elements that can affect glanceability such as quantity and type of information as well as design elements including color, shape, position, and size. Depending on the application, designers have to determine how and where information is presented to strike a balance between these elements without distracting user's attention. For instance, Matthews et al. designed an email display that used abstract representations, consisted of visually distinct components, and maintained consistency [32].

One particularly relevant property of our visual system for glanceable design is pre-attentive processing. While focused attention is only possible within a small portion of the visual field, pre-attentive processing features such as color and shape are detected rapidly and in parallel within the brief period of a single fixation [23]. The exact reason as to why certain information is processed pre-attentively is unknown, but it is generally accepted that the selection is influenced by the interaction of the salience of a stimulus and the observer's current intentions and/or goals [18]. For instance, the speed and efficiency of pre-attentive processing is contingent on the observer's current intentions and/or goals [20]. In this article, we primarily focus on controlling the salience of the robot motion to improve its glanceability.

### 3 LEGIBILITY AND GLANCEABILITY OF A ROBOT MOTION

We first define the legibility and glanceability of a swarm robot motion. For legibility, we follow the same definition as prior work in legibility of a single robot motion [5, 6, 17]. Legibility of a robot motion entails how well and quickly an observer can predict the intent or goal of the robot without any prior knowledge. It has been measured by how quickly observers are able to make a prediction and by the self-reported prediction confidence rating [16, 36].

However, the concept of glanceability has not been introduced for a robotic motion. Glanceability is also a measure of legibility but with an added constraint of exposure time. Instead of seeing the entire robot trajectory, an observer sees only a short segment of the motion and needs to predict the intent of the robot based on such limited information. Legible robot motion may not always translate to glanceable motion, because it is designed to convey intent early but not necessarily rapidly across different points along the trajectory.

Glanceability is especially relevant in the context of human-multirobot interaction, as observers will have to monitor swarms of robots, which cannot be done simultaneously. Instead of continuously watching a robot, users will have to constantly shift focus from one group to another, spending a limited amount of time on each group [14]. Thus, a glanceable swarm robot motion will be more suitable for either sporadic interactions that require user's attention only intermittently or interaction with multiple groups of robots.

To define the time limit for glanceability, we rely on the visual perception literature, as users need to visually process the robot motion and infer its intent. Specifically, pre-attentive processing has traits desirable for glanceable robot motion such as being processed rapidly and in parallel without the need for focused attention [23, 47]. Hence, we propose a time limit of 250 ms, the upper threshold for pre-attentive processing [23], because our goal is to generate a swarm robot motion that can be processed pre-attentively. Mathematical formulations for both legible and glanceable motions are defined in the Sections 5.4 and 6.4, respectively.
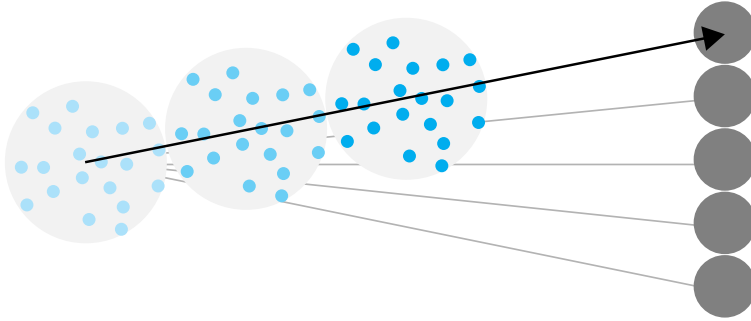
Fig. 2. One of the common path planning algorithms minimizes the overall trajectory length resulting in a **straight motion**. For a group of robots, this results in robots synchronously following a straight trajectory toward the goal.

## 4 DESIGN AND GENERATION OF LEGIBLE SWARM ROBOT MOTION

To design a legible swarm robot motion, we first formulate the problem as a non-linear constrained optimization problem. Then, we introduce the commonly used optimization setup to plan the shortest trajectory from point A to point B and use this trajectory as the control condition. To design the other methods, we explain how we leverage findings from prior works in single robot legible motion [6, 17], swarm intelligence [34], and visual perception [23, 47]. The objectives and design criteria are explained for trajectory, swarm behavior, and pre-attentive processing feature-based legible motions.

### 4.1 Problem Formulation

As introduced by Witkin and Kass [50] and used by Bodden et al. to generate intent-expressive motion [6], we use a trajectory optimization or *spacetime constraints* as the framework to construct different types of legible swarm robot motion. An optimization problem is used to generate a trajectory that minimizes the given specified objectives. The constraints are used to enforce requirements for the trajectory, such as the initial and final locations.

A trajectory, $X$, is a function that maps time to configurations of the group of 2-D mobile robots, $X : R- > R^{2n}$, where n is the number of robots. $X(t)$ is used to denote the configurations at time $t$, and $g(X)$ and $c_i(X)$ are the objective function and the set of constraints, respectively. The resulting optimization for the duration of the trajectory, $t_0$ to $t_f$ is set up as below:

$$X^* = \operatorname{argmin} g(X) \text{ subject to } c_i(X),$$

where $X^*$ denotes the optimal trajectory. We use the following constraints:

(1) begin in the designated positions with no overlap among the robots, $X(t_0) = X_0$.
(2) end in the designated positions, $X(t_f) = X_f$.

### 4.2 Straight Motion

A common path planning algorithm computes a trajectory that goes from the initial position to the final position as quickly as possible without colliding with obstacles [24, 40] or is minimal in the required energy [50]. Here, we use the setup below to compute the straight trajectory that minimizes the overall length resulting in trajectories. as depicted in Figure 2, and use it as the
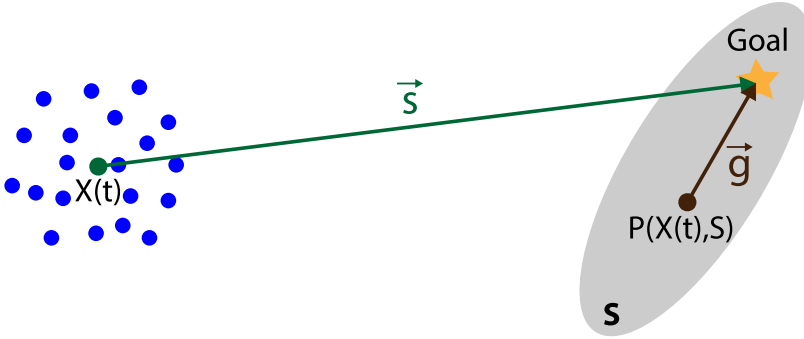
Fig. 3. The trajectory-based legible motion is a result of the optimization of the distance from the predicted goal to the actual goal (i.e., $\vec{g}(X(t))$), the distance from the center of the robots to the goal (i.e., $\vec{s}(X(t))$), and the overall trajectory length (i.e., $\|X(t)'\|^2$).

control condition for the experiments.

$$Straight = \int_{t_o}^{t_f} \|X(t)'\|^2 \, dt \tag{1}$$

### 4.3 Trajectory-based Legible Motion

Leveraging and modifying the trajectory has been the state-of-the-art to improve the legibility of a single robot motion [5, 6, 17, 36]. One of its main benefits is that it can be generalized to all types of robots both with or without anthropomorphic features. While the exact weights and formula for the legibility measure are not the same, the overall approach to generating a legible trajectory is similar across different prior works in that they all optimize a sum of costs related to both the legibility and functional costs (e.g., length and smoothness of the trajectory) to ensure reasonable and smooth detour toward the target. A similar approach can be applied to the motion of a swarm of robots in a centralized manner where they follow the same trajectory while maintaining equal distance among the robots. Although some researchers have explored the idea of modifying trajectory to generate legible multi-robot motion [9, 10], both papers employed an arc-trapezoidal trajectory based on the animation principles of "Slow in and out" and "Arcs" from Reference [27]. They did not use trajectory specifically designed and proven to improve legibility of a single robot motion such as References [6, 17]. Thus, we propose using and evaluating the adapted algorithm from [6] to generate the legible trajectory for a swarm of robots.

We use the same optimization framework as shown below:

$$Legible = \int_{t_o}^{t_f} \alpha \|\vec{g}(X(t))\|^2 + \beta \|\vec{s}(X(t))\|^2 + \epsilon \|X(t)'\|^2 dt. \tag{2}$$

This minimizes the following components:

$$\vec{g}(X(t)) = Goal - P(X(t), S),$$

$$\vec{s}(X(t)) = Goal - X(t).$$

As shown in Figure 3, $\vec{g}(X(t))$ is the vector from the observer's predicted goal out of the set S of possible goals to the actual goal, and $\vec{s}(X(t))$ is the vector from the current position $X(t)$ to the actual goal. For the heuristic $P(X(t), S)$, we use the "point position" formula from Reference [6], where the predicted goal will be the member of the goal set that is closest to the current position of the center of the robots based on Euclidean distance.
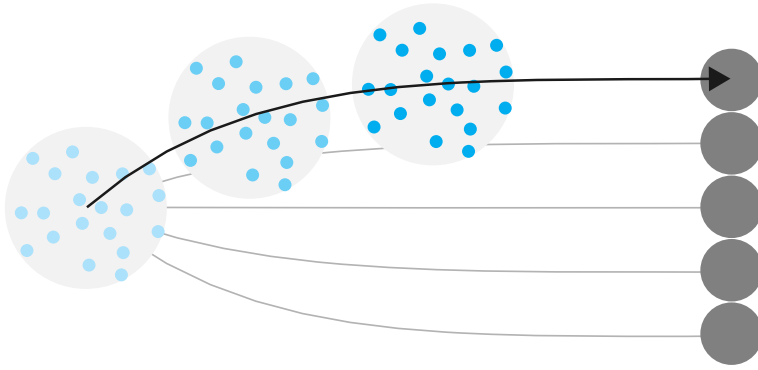
Fig. 4. An example of **trajectory-based legible motion** is shown. The trajectories for each goal are shown in gray lines but due to space constraints, only one instance is completely visualized where the robots follow a curved trajectory toward the topmost goal while maintaining a constant formation during the trajectory.

The optimization weights $\alpha$ and $\beta$ determine the tradeoff between direct/energy-efficient motions and legible ones. All figures and experiments in this work use $\beta = \alpha/10$ and $\epsilon = \alpha/10$, the same ratio used in Reference [6]. The third term in Equation (3) is a regularization term (as mentioned in Section 3.1) to ensure continuity in joint angles. The resulting trajectories are applied such that the center of the robots follow the trajectory while maintaining the same formation as shown in Figure 4.

## 4.4 Collective Behavior-based Legible Motion

In contrast to the trajectory-based legible motion where we optimized an objective for the entire group of robots, collective behavior in swarms found in nature often emerges from a set of simple distributed rules. Interestingly, these different swarm behaviors elicit different perceptions from people even though they are not explicitly designed to do so [25]. For instance, rendezvous behavior (i.e., agents moving toward the same location) is perceived as being more urgent and arousing than torus behavior (i.e., agents moving in a circle) even without any context [25]. In addition, prior work has shown the humans can perceive and recognize different swarm behaviors accurately (~80%) even with high levels of noise (>80%) [43, 49]. Thus, we leverage one of the collective behaviors, rendezvous, to generate a legible motion and compare it with other types of legible motion.

The rendezvous behavior involves a large number of agents that move toward the same destination. In reality, while robots can move toward the same destination, they cannot occupy the exact same space. Thus, we use the existing literature on circle packing to derive the final configurations for the robots that are closely packed without any overlap [22]. This set of configurations is used to define the final position constraint. To match the robots with the closest final configuration that will result in the shortest overall length of the sum of the trajectories, we use the Hungarian algorithm [26]. Note that the above two steps require some centralized coordination. Finally, the straight trajectory for each robot is generated with Equation (4.2). The resulting motions for each target are shown in Figure 5.

## 4.5 Pre-attentive Processing Feature-based Legible Motion

Observing a swarm of round mobile robots resembles viewing an array of simple circles on a screen. Thus, we explore utilizing existing knowledge in the vision perception field, specifically about the pre-attentive visual processing features [47]. Researchers have identified visual param-
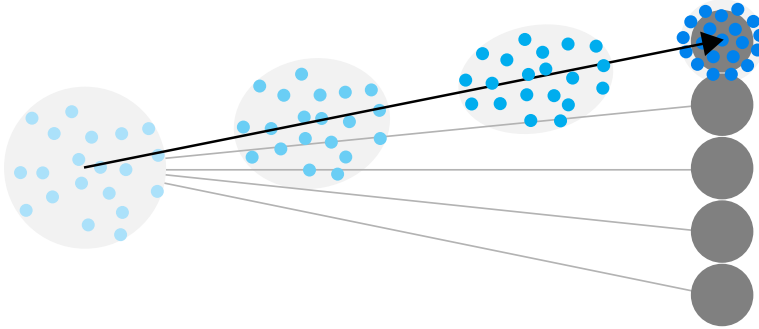
Fig. 5. An example of **collective behavior-based legible motion** is visualized where the robots are merging toward the goal but the center of robots follow a straight trajectory. The gray lines indicate the trajectories of the center of the robots for other goals.
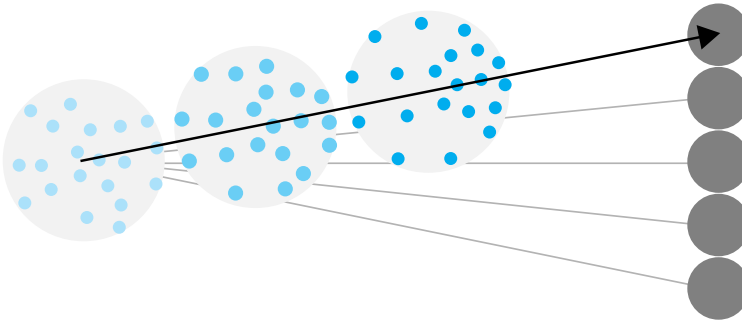


Fig. 6. An example of **pre-attentive processing feature-based legible motion** is shown. The robots change formation while following a straight trajectory overall. Due to the space constraints, the transition from the initial randomly distributed formation to the final formation is stretched.

eters that are processed subconsciously within approximately 250 ms and these factors include color, shape, orientation, and density [23]. Since legible swarm robot motion needs to provide hints about where the robots are headed, we saw potential in using the pre-attentive processing features, as they could rapidly direct user's attention toward a desired region within the collection of the robots. Out of the potential candidates (e.g., color, shape, size), we are particularly interested in using density, as it only requires the ability to control the distribution of the robots, which is possible for any mobile robots. Other pre-attentive processing features will require additional resources and capabilities from the robots, such as external light sources, and shape or size-changing ability.

To uniformly distribute the robots within the two regions of different densities, we precomputed this desired distribution by using the **Centroidal Voronoi Tessellation (CVT)** [37]. This method has been used to uniformly distribute robots for animation display with multiple mobile robots [1] and mobile sensing network [12]. While an online approach such as bio-inspired swarm algorithm [13] is desired and necessary for real-time applications, these approaches may not yield the exact pattern that we would like to create (i.e., the pattern shown in Figure 6). Thus, in this article, we focus on designing the desired motion and evaluating the legibility of this type of motion before optimizing the computational performance for real-time interaction.

Unlike the previous types of legible motion, this motion requires rearrangement within the robots to transition from the initial configuration to the desired distribution. Thus, to reach the desired formation within a reasonable time, the speed of each robot is adjusted proportional to

the distance from each robot's target similar to the method described in Reference [1]. For this article, we tuned the parameters to ensure that the robots reached the desired distribution within 1 second. An example of the resulting motion is shown in Figure 6.

## 5 EVALUATION 1: LEGIBILITY

To evaluate and compare the different legible motions, we ran two online studies to measure their legibility and glanceability performances. Both experiments involved participants watching a video of the robots moving toward one of the possible targets and making a prediction on the goal of the robots. In the first study, we investigate how different parameters (i.e., types of legible motion, initial spread of the robots, and target difficulty) impact the *legibility* of the swarm robot motion. In addition to the different legibility cue conditions as discussed in the previous section, we investigate the effects of initial spread of the robots and target difficulty, as we expect these parameters to have a different impact on the performance of different types of legible motion. For our task, the target difficulty is determined by how many targets are adjacent to the goal, as explained in Section 4.3.

### 5.1 Hypotheses

The trajectory-based multi-robot legible motion is a direct adaptation of the prior work on single robot legible motion [6, 17]. As prior work have shown the effectiveness of leveraging trajectory to increase legibility of a motion, we also hypothesize that trajectory-based multi-robot legible motion will also enhance legibility compared to the control condition. However, the heuristic used to generate legible trajectory heavily depends on where the targets are relative to each other. Thus, we hypothesize that trajectory-based legible motion will have better legibility performance for easier targets.

While rendezvous behavior has not been used to enhance legibility of robot motion, robots that merge toward a destination may provide additional information about the location of the destination. Thus, we hypothesize rendezvous-based legible motion will have better legibility performance than the control condition. Since the rendezvous-based legible motion only depends on the absolute position of the goal, we hypothesize that the target difficulty will not affect its legibility performance.

Similar to rendezvous behavior, density also has not been used to generate legible swarm robot motion. However, it is one of the pre-attentive processing features [23, 47] that can be detected rapidly by human visual processing pathways. Thus, we expect a trajectory that features a denser region pointing toward the goal will improve the legibility at least compared to the control condition without any additional cues about where the goal is. Similar to the rendezvous condition, density condition also mostly depends on the absolute position of the goal with some dependence on the initial robot distribution and thus should have similar legibility performance regardless of the target difficulty.

Contrary to other task parameters such as target radius, inter-target distance, and distance between the robots and targets, the initial radius of the circle encompassing the robots is a parameter is a controllable parameter. We also expect it to have an impact on the legibility in general as well as varying influence on each type of legibility cue. Prior work has shown that increasing dispersion level decreases the prediction time [10], and thus we also expect decrease in prediction time. Since the nature of different legibility cues are different, we expect different responses when the initial radius is changed. For instance, rendezvous-based legible motion may benefit from larger initial radius, as that increases the merging rate of the motion; whereas for the control and trajectory conditions, a larger radius may only confuse observers, as it is less clear where the center of the robots is.
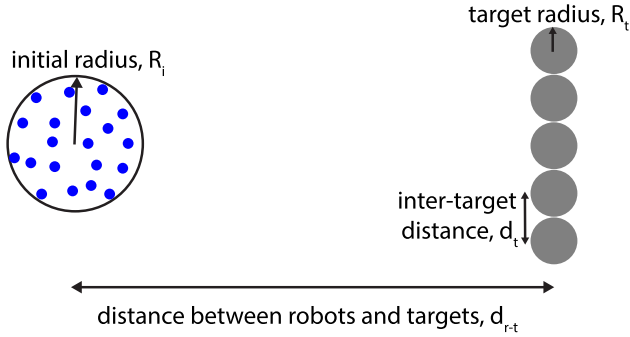
Fig. 7. Task setup for user study with robots initially on the left side and five targets on the right side.

- *Effects of Legibility Cue*

**H1.1** Trajectory-based legible motion will improve legibility compared to the control condition.
**H1.2** Rendezvous-based legible motion will improve legibility compared to the control condition.
**H1.3** Density-based legible motion will improve legibility compared to the control condition.

- *Effects of Target Difficulty*

**H1.4** Trajectory-based legible motion will have better legibility performance for easier targets.
**H1.5** Rendezvous-based legible motion will have similar legibility performance regardless of the target difficulty
**H1.6** Density-based legible motion will have similar legibility performance regardless of the target difficulty.

- *Effects of Initial Radius*

**H1.7** Larger initial radius will decrease prediction time.
**H1.8** Initial radius will have different effects based on the type of legibility cue.

## 5.2 Task

Participants were asked to watch a set of videos in which a set of simulated robots begin on the left side with five potential targets on the right side, as shown in Figure 7. Robots moved towards one of the targets and the objective for the participants was to stop the video when they were confident about where the robots are going to. They then predicted where the robots are heading to by selecting the goal and rated their confidence in their prediction on a Likert scale from 1 to 7. Three parameters determine the difficulty of the task: the radius of the circle that initially encompasses the robots ($R_i$), the distance between the initial position and the targets, and the distance between the targets. The values of the latter three parameters were determined experimentally to ensure adequate difficulty, while the $R_i$ was one of the independent variables, as we expected it to affect how well each legibility condition performs. The size of the workspace and the robots were based on realistic tabletop settings with centimeter-scale robots such as the Zooids [28]. The values of the constant task parameters used and the initial radius are listed in Table 1. The different robot motions were programmed in C++ in Visual Studio using the Zooids Simulator [28] which uses the **Hybrid Reciprocal Velocity Obstacle (HRVO)** [44] for a collision-free and oscillation-free navigation of multiple mobile robots.

Table 1. Task Parameters for Both Evaluations

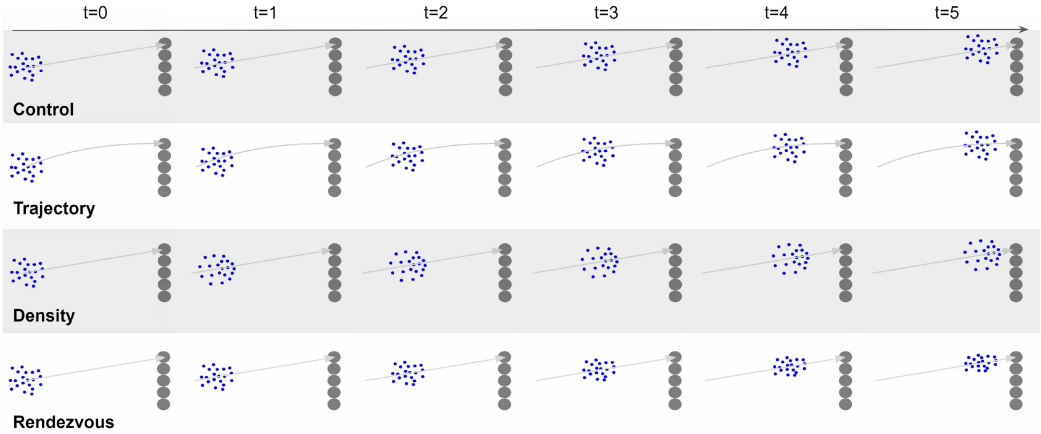| Variable | Units | Value | Description |
|---|---|---|---|
| $d_{r-t}$ | [m] | 1.0539 | Distance between the robots and targets |
| $R_r$ | [m] | 0.013 | Radius of each robot |
| $R_t$ | [m] | 0.05 | Radius of each target |
| $d_t$ | [m] | 0.1 | Distance between adjacent targets |
| $R_i$ | [m] | 0.125, 0.15, 0.175 (small, medium, large) | Initial radius values used for the studies |



Fig. 8. Motion trajectories for different legibility cues. The gray arrows are only drawn to help differentiate each legibility cue and were not visible to the study participants.

## 5.3 Independent Variables

To test our hypotheses, we evaluate the performance of different legibility cues including control, trajectory, rendezvous, and density, which are shown in Figure 8. As we expect the initial radius $R_i$ to have an impact on performances of different legibility conditions, we also tested three different values of initial radius that were determined experimentally through pilot studies to ensure adequate difficulty. Finally, we study the effects of the target location, as it affects the difficulty of the task by showing the participants videos of the robots moving toward each of the five targets. As the three middle targets have more potential targets that could confuse observers, we labelled these three targets as difficult and the two outer targets as easy. There is a total of $4 \times 3 \times 2 = 24$ conditions with two between-subject factors (i.e., legibility condition and initial radius) and one within-subject factor (i.e., target difficulty).

## 5.4 Measures

For each trial, we recorded the participant's target prediction, prediction time, and prediction confidence on a Likert Scale from 1–7 as done in Reference [36]. We also combined these three answers into a single legibility score metric similar to prior work [16, 36] as below: If the prediction is correct, then

$$LEGIBILITY\_SCORE = \frac{T_{total} - T_{predict}}{T_{total}} \times \frac{R_{conf}}{7}, \tag{3}$$

where $T_{total}$ is the total time duration from the initial position to the goal, $T_{predict}$ is the time of the prediction, and $R_{conf}$ is the Likert scale rating of the prediction confidence. If the prediction is incorrect, then we assign a score of 0.

We also collected subjective ratings of the robots in terms of legibility and predictability using a modified questionnaire from prior work [6]. On a seven-point scale, participants rated the statements 1–2 for legibility and statements 3–4 for predictability.

(1) It was easy to predict which target the robots were moving toward.
(2) The robots moved in a manner that made their intention clear.
(3) The robots' motion matched what I would have expected if I knew the target beforehand.
(4) The robots' motion was not surprising.

## 5.5 Procedure

For the legibility task, participants were asked to watch and stop a seven-second-long video either by pressing space bar on the keyboard or clicking on the pause button on the screen when they felt confident enough to make a prediction. Once the participants paused the video, they made their prediction by clicking on one of the targets presented as gray circles and rated their prediction confidence on a scale of 1–7. To familiarize with the task, participants went through a practice trial and then began the experiment with five videos where they were shown robots moving toward each of the five targets. Afterwards, participants filled out a questionnaire on the perception of the robots in terms of legibility and predictability, and provided demographic information.

## 5.6 Participants

We recruited 283 participants through Amazon Mechanical Turk. For each of 12 conditions (4 legibility cue conditions × 3 initial radius), approximately 20 participants viewed and rated the corresponding video. For quality control, only participants that satisfy the following requirements were included in the analysis:

(1) Their Location is in the United States of America.
(2) Their HIT approval rate is greater than 90.
(3) Their Number of HITs approved is greater than 50.
(4) They have not previously participated in any of our pilot studies.
(5) They are not experiencing any symptoms that may affect performance in the experiment.
(6) They have participated only once.

For the analysis, we removed 52 participants with duplicate IP addresses as well as those who failed to follow the instructions correctly (i.e., did not pause or paused even before the robots started moving).

Participants reported ages ranging from 18 to 81 with mean = 37.8 and SD = 12.4. A total of 51% identified as men and 49% as women, and 21%, 63%, and 16% of participants reported education levels of middle/high school, college, and advanced degrees, respectively. After completing the experiment in 2.75 minutes on average, each participant received compensation at a hourly rate of approximately $15.00 US dollars.

## 5.7 Analysis

To investigate both main and interaction effects for the overall legibility score, prediction time, prediction confidence, and accuracy, we ran a $4 \times 3 \times 2$ mixed-design ANOVA with two between-subject factors (i.e., legibility condition and initial radius) and one within-subject factor (i.e., target

Table 2. Descriptive Statistics of Study 1

| | | Control | | | | | | Trajectory | | | | | | Density | | | | | | Rendezvous | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Radius1 | | Radius2 | | Radius3 | | Radius1 | | Radius2 | | Radius3 | | Radius1 | | Radius2 | | Radius3 | | Radius1 | | Radius2 | | Radius3 | |
| | | M | SD | M | SD | M | SD | M | SD | M | SD | M | SD | M | SD | M | SD | M | SD | M | SD | M | SD | M | SD |
| *Objective* | Time | 5.9 | 1.7 | 5.9 | 1.5 | 5.4 | 1.4 | 6.1 | 1.7 | 5.4 | 1.5 | 5.6 | 1.6 | 6.2 | 1.5 | 5.9 | 1.5 | 5.6 | 1.5 | 6.6 | 1.5 | 5.6 | 1.4 | 5.7 | 1.4 |
| *Measures* | Accuracy | 1.0 | 0.2 | 0.9 | 0.4 | 0.9 | 0.3 | 0.9 | 0.3 | 0.9 | 0.4 | 0.9 | 0.3 | 0.9 | 0.2 | 0.9 | 0.2 | 0.8 | 0.4 | 1.0 | 0.2 | 1.0 | 0.2 | 1.0 | 0.2 |
| | Confidence | 5.6 | 1.0 | 5.8 | 0.9 | 5.6 | 1.2 | 6.0 | 0.8 | 5.8 | 1.1 | 5.7 | 1.1 | 6.1 | 0.8 | 6.1 | 0.9 | 5.3 | 1.1 | 6.5 | 0.8 | 6.0 | 0.9 | 6.5 | 0.7 |
| | Leg_Score | 0.3 | 0.2 | 0.3 | 0.2 | 0.4 | 0.2 | 0.3 | 0.2 | 0.4 | 0.2 | 0.4 | 0.2 | 0.3 | 0.2 | 0.3 | 0.2 | 0.3 | 0.2 | 0.3 | 0.2 | 0.4 | 0.2 | 0.4 | 0.2 |
| *Subjective* | Legibility | 5.3 | 1.0 | 5.4 | 1.2 | 5.3 | 1.4 | 5.6 | 0.7 | 5.5 | 1.4 | 5.3 | 0.8 | 5.5 | 0.9 | 5.3 | 1.5 | 5.1 | 1.1 | 5.9 | 1.4 | 5.8 | 1.0 | 6.1 | 0.8 |
| *Measures* | Predictability | 5.5 | 1.0 | 5.5 | 1.5 | 5.2 | 1.5 | 5.3 | 1.0 | 5.5 | 1.4 | 5.4 | 1.1 | 5.3 | 1.2 | 5.8 | 1.2 | 4.8 | 1.5 | 5.8 | 0.9 | 5.8 | 1.0 | 5.9 | 0.7 |

Color depicts the ordering of results from dark blue (best performing) to light blue (worst performing).
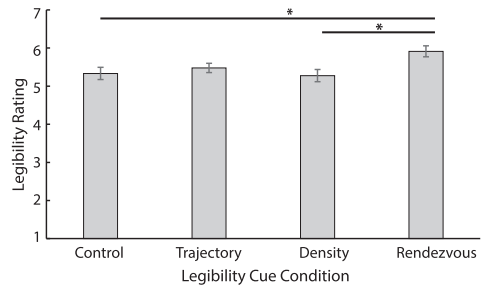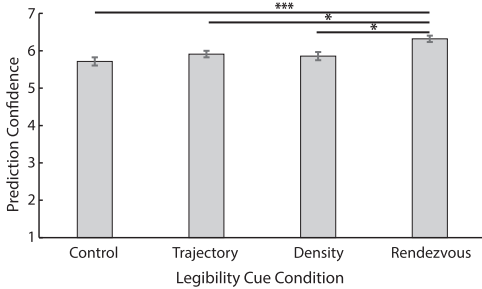


(a) Effects of legibility cue condition on self-reported prediction confidence

(b) Effect of legibility cue condition on self-reported legibility rating

Fig. 9. (a) Rendezvous had the highest confidence ratings among all conditions. (b) Rendezvous condition is rated higher than density and control conditions.

difficulty). Then, we performed the Bonferroni-corrected post hoc tests on the statistically significant effects.

Similarly, we ran a $4 \times 3$ between-subjects ANOVA with with two between-subject factors (i.e., condition and initial radius) for the subjective ratings. Then, we performed the Bonferroni-corrected post hoc tests on the statistically significant effects.

## 5.8 Results & Discussion

The descriptive statistics of Study 1 are shown in Table 2. All the bar graphs (Figures 9–12) plot the mean and standard error. Significantly different pairs indicated by a bar accompanied by varying numbers of ∗, where ∗: $.01 \le p < .05$, ∗∗: $.001 \le p < .01$, and ∗ ∗ ∗: $p < .001$. Only the statistically significant results are described below.

*5.8.1 Effects of Legibility Cue.* Legibility cue had significant effects on prediction confidence ($F(3, 214) = 6.9, p < 0.001, \eta^2 = 0.088$) and self-reported legibility ($F(3, 223) = 3.7, p = 0.013, \eta^2 = 0.047$) as shown in Figures 9(a) and 9(b), respectively. Out of the four different legibility cues, the rendezvous behavior-based legible motion had the best legibility performance in terms of self-reported prediction confidence and self-reported legibility rating. It was rated significantly higher in prediction confidence and self-reported legibility ($F(3, 223) = 3.7, p = 0.013, \eta^2 = 0.047$). These results support our hypothesis H1.2. However, density and trajectory-based legible motions did not perform significantly better than the control condition, failing to support our hypotheses H1.1 and H1.3.

*5.8.2 Effects of Target Difficulty.* As expected, target difficulty had significant effects on all measures including overall legibility score ($F(1, 218) = 31.2, p < 0.001, \eta^2 = 0.125$), prediction confidence ($F(1, 214) = 34.24, p < 0.001, \eta^2 = 0.138$), prediction accuracy ($F(1, 218) = 3.96$,
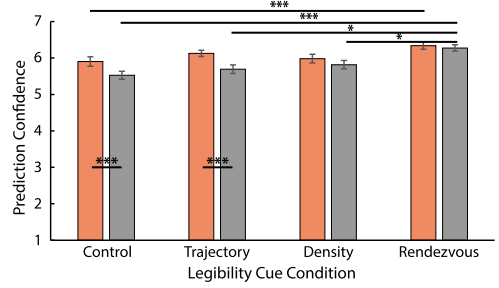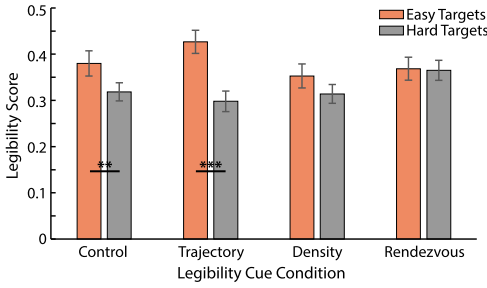
(a) Interaction effects between target difficulty and legibility cue conditions on the overall legibility score

(b) Interaction effects between target difficulty and legibility cue conditions on prediction confidence

Fig. 10. (a) There are significant differences between the target difficulty levels for the control and trajectory conditions but not for the density and rendezvous conditions. (b) For hard targets, rendezvous outperformed all other conditions, while it only outperformed the control condition for easy targets. The confidence ratings are significantly different between the two target difficulty levels for control and trajectory conditions.
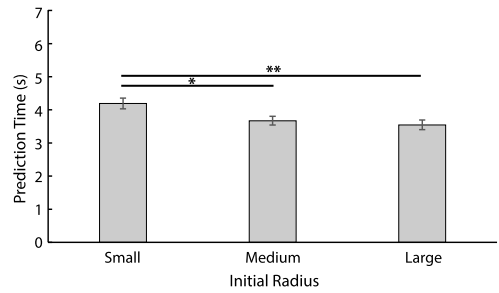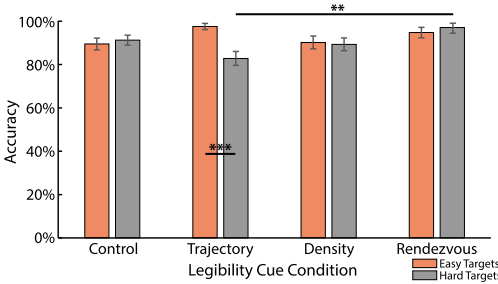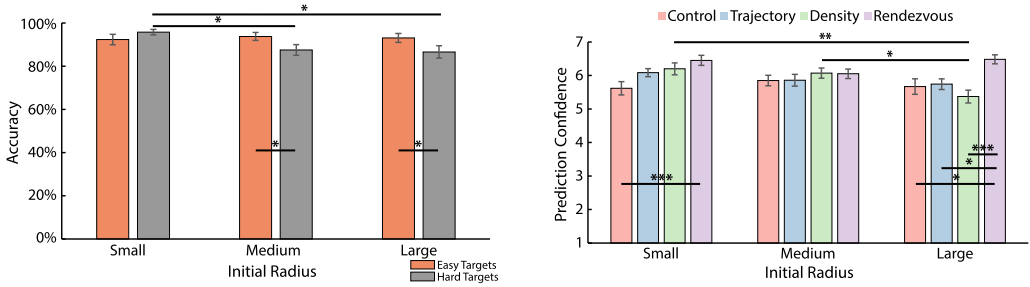


(a) Interaction effects between target difficulty and legibility cue conditions for prediction accuracy

(b) Effects of initial radius on prediction time

Fig. 11. (a) There was a significant difference between the target difficulty levels for trajectory condition, and the accuracy was significantly higher for rendezvous condition than trajectory condition when the target difficulty was hard. (b) With larger initial radius values, the prediction time significantly decreased.

$p = 0.048, \eta^2 = 0.018$), and prediction time ($F(1, 214) = 16.95, p < 0.001, \eta^2 = 0.073$). Harder targets (i.e., three targets in the middle) had lower legibility score, confidence ratings, and prediction accuracy, while it had higher prediction time.

As shown in Figures 10(a), 10(b), and 11(a), the interaction effects between legibility cue and target difficulty demonstrated that both the control and trajectory conditions do not perform consistently across different targets, while density and rendezvous conditions do. We observed that the control condition has lower overall legibility score ($F(1, 56) = 9.0, p = 0.004, \eta^2 = 0.139$) and prediction confidence ($F(1, 56) = 16.1, p < 0.001, \eta^2 = 0.223$) for harder targets, but similar prediction accuracy for both target difficulties. For trajectory-based legible motions, we saw a significant difference across different target difficulties for overall legibility score ($F(1, 59) = 27.1, p < 0.001, \eta^2 = 0.315$), accuracy ($F(1, 59) = 16.82, p < 0.001, \eta^2 = 0.222$), and prediction confidence ($F(1, 58) = 18.99, p < 0.001, \eta^2 = 0.247$). It is surprising to find that the control condition is not rated consistently in terms of prediction confidence, since its trajectories do not alter based on where the target is. In contrast, the trajectories for the trajectory condition heavily depend on number and location of other targets relative to the goal. These results support H1.4–H1.6.

(a) Interaction effects between target difficulty and initial radius for prediction accuracy

(b) Interaction effects between initial radius and legibility cue conditions for prediction confidence

Fig. 12. (a) For hard targets, medium and large radius trials had lower accuracy than small radius trials. For medium and large radius trials, there were significant differences between the two difficulty levels. (b) For large initial radius trials, rendezvous had better scores than all other conditions, while it only had better score than the control condition for small radius trials. The confidence ratings decreases with the large radius.

While the study results suggest that collective behavior-based motion is the most legible motion in terms of self-reported prediction confidence and self-reported legibility rating, different legible cues had varying performances based on the context or difficulty of the task. For instance, the trajectory condition had the highest overall legibility score when the target difficulty was easy (mean = 0.429), albeit not significantly, as shown in Figure 10(a). Thus, this suggests that it may be best to use a combination of different legibility cues based on the context. For instance, when moving toward targets that are easier, trajectory-based legible motion can be used; whereas for harder targets, robots could rendezvous toward the goal.

*5.8.3 Effects of Initial Radius.* The effects of initial radius are mixed. For prediction time, there was a significant main effect ($F(2, 214) = 5.53, p = 0.005, \eta^2 = 0.049$), as shown in Figure 11(b), where the participants made the prediction more quickly when the radius was larger. However, there are also interaction effects with target difficulty on prediction accuracy ($F(2, 218) = 5.23, p = 0.006, \eta^2 = 0.046$), where the accuracy decreases with larger initial radius, as shown in Figure 12(a). This trend of faster prediction agrees with prior work [10], while the decrease in accuracy has not been shown before. This could be due to the difference in task difficulty, as the targets in our task are much closer to each other than those from Reference [10]. These results support H1.7 but also provide evidence that accuracy is decreased with larger radii, especially for hard targets.

As hypothesized in H1.8, initial radius had different effects based on the type of legibility cue. As shown in Figure 12(b), there were interaction effects between legibility cue and initial radius on prediction confidence ($F(6, 214) = 2.44, p = 0.026, \eta^2 = 0.064$). In particular, only the density-based legible motion was significantly affected by the change in initial radius ($F(2, 53) = 6.47, p = 0.003, \eta^2 = 0.196$), while others were unaffected. We saw a decrease in prediction confidence with an increase of initial radius for density conditions. This suggests that the density-based legible motion could be further optimized especially for different radii. In addition, there were significant effects of conditions for small ($F(3, 74) = 4.69, p = 0.005, \eta^2 = 0.160$) and large initial radius values ($F(3, 68) = 6.39, p < 0.001, \eta^2 = 0.22$). When the initial radius is small, rendezvous condition has higher confidence score than the control condition. For middle initial radius trials, there was no significant effect, while rendezvous condition had higher ratings than all other conditions for largest initial radius trials. This indicates that rendezvous-based motion is most legible motion when the robots are more widely spread out.

## 6 EVALUATION 2: GLANCEABILITY

With a swarm of robots, it is unlikely that all of the robots will be performing the same task simultaneously or that people will be solely focusing on the robots. Rather, groups of robots will be assigned to different tasks to optimize the overall efficiency, and people will be attending to other tasks. In such scenarios, supervisors of these swarm robot systems face an exhausting and complex task of trying to monitor multiple groups at the same time. To alleviate some of the burden, we aim to design glanceable swarm robot motion that can be understood by observers with a quick glance. While researchers have used different thresholds for glanceable design [21], in this article our goal is to generate robot motion that can be processed pre-attentively and thus use the threshold of 250 ms [47]. Thus, the task remains the same as the previous study except that the participants are shown 250 ms long videos of different segments along the trajectory. Using the set of motions described in Section 3, we evaluate their glanceability and the effects of task parameters, such as target difficulty and initial radius.

### 6.1 Hypotheses

The trajectory-based legible motion plateaus after the initial leap toward the goal. Thus, we expect trajectory condition to perform better than the control condition only in the beginning but not during the later half. On the contrary, both density and rendezvous conditions provide constant cue toward the goal throughout the trajectory and thus we expect both to improve the glanceability consistently compared to the control condition.

As the algorithm for trajectory condition depends on the relative positions of the targets, we expect that the trajectory-based legible motion will have different glanceability performance based on the target difficulty, whereas both density and rendezvous-based legible motion will not as their algorithm only depend on the absolute location of the goal.

For the initial radius, we hypothesize that it will have a different impact on each of the legibility cue similar to the prior experiment. Larger radius may be beneficial for rendezvous condition as the angle of rendezvous becomes larger, whereas it may decrease the glanceability for trajectory and density conditions, as the center of the robots may be less clear to the observers.

- *Effects of Legibility Cue*

**H2.1** Trajectory-based legible motion will only improve glanceability in the beginning but not during the later half of the trajectory compared to the control condition.
**H2.2** Rendezvous behavior-based legible motion will constantly improve the glanceability throughout the trajectory compared to the control condition.
**H2.3** Density-based legible motion will constantly improve the glanceability throughout the trajectory compared to the control condition.

- *Effects of Target Difficulty*

**H2.4** Trajectory-based legible motion will have better glanceability performance for easier targets.
**H2.5** Rendezvous-based legible motion will have similar glanceability performance regardless of the target difficulty.
**H2.6** Density-based legible motion will have similar glanceability performance regardless of the target difficulty.

- *Effects of Initial Radius*

**H2.7** Initial radius will have different effects based on the type of legibility cue.

## 6.2 Task

Similar to the previous study, participants watched videos and were asked to predict which target the robots are heading to. The setup of the task was exactly the same as before except for one part. Instead of having them watch the full trajectory and asking them to pause the video when they are confident about their prediction, participants watched a short 250 ms clip of the robots moving toward one of the targets with 2 seconds of blank screen with only the five targets shown. They were then asked to predict the goal target as well as their prediction confidence on a Likert scale from 1 to 7.

## 6.3 Independent Variables

In addition to the set of independent variables used in the prior study, we varied the timing of video segment shown to the participants. As we wanted to evaluate the glanceability of the motion, we only displayed 250 ms segment of the 7-second-long trajectory at specific moments. As it is desirable to convey the intent of the robots as early as possible, we chose to show participants 250 ms clips beginning at 1 second (i.e., time 1) and 2 seconds (i.e., time 2) of the trajectory. As prior study indicated that most participants make the prediction around 4 seconds (i.e., time 4), we decided to show 250 ms at 4 seconds as well.

## 6.4 Measures

There currently is no measure or instrument for glanceability. Hence, we propose a glanceability score metric that is defined as a weighted sum of the product between the prediction accuracy (0 if incorrect, and 1 if correct) and self-reported prediction confidence at different times along the path. This metric is similar to that for legibility score as defined in previous study and prior work [16, 36], where the legibility scores at each timestamp are aggregated to generate a single legibility score for the entire path. Higher weights are assigned for correct predictions earlier in the trajectory, as it is desirable to convey the intent of the robots as early as possible. In our experiment, we recorded the participant's goal prediction and prediction confidence for each trial on a Likert Scale from 1 to 7. Then, we summed the predictions and prediction confidence ratings from 1, 2, and 4 seconds into a single glanceability score metric as below:

$$GLANCEABILITY\_SCORE = \sum_T \frac{T_{total} - T}{T_{total}} \times \frac{R_{conf}}{7} \times \mathbf{1}[Prediction == Correct], \qquad (4)$$

where T is the time of the segment shown to the participants (i.e., 1, 2, and 4 for this study), $T_{total}$ is the total time duration from the initial position to the goal, and $R_{conf}$ is the Likert scale rating of the prediction confidence.

## 6.5 Procedure

Participants were first instructed about what their task is. They then watched 250 ms long clips of the swarm robot motion a total of 15 times with five different targets and three segments in a random order. After watching all the videos, participants filled out a questionnaire on the perception of the robots and provided demographic information.

## 6.6 Participants

We recruited 240 participants through Amazon Mechanical Turk. For each of the 12 conditions (4 conditions × 3 initial radius values), approximately 20 participants were randomly assigned to a condition. They then viewed and rated 15 randomized videos (5 different targets × 3 time segments) corresponding to that condition. For quality control, we had the same requirements as in the

Table 3. Descriptive Statistics of Study 2

| | | Control | | | | | | Trajectory | | | | | | Density | | | | | | Rendezvous | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Radius1 | | Radius2 | | Radius3 | | Radius1 | | Radius2 | | Radius3 | | Radius1 | | Radius2 | | Radius3 | | Radius1 | | Radius2 | | Radius3 | |
| | | M | SD | M | SD | M | SD | M | SD | M | SD | M | SD | M | SD | M | SD | M | SD | M | SD | M | SD | M | SD |
| *Objective Measures* | Accuracy_1 | 0.6 | 0.2 | 0.5 | 0.3 | 0.5 | 0.2 | 0.7 | 0.3 | 0.7 | 0.2 | 0.6 | 0.2 | 0.4 | 0.3 | 0.6 | 0.3 | 0.5 | 0.3 | 0.5 | 0.3 | 0.4 | 0.3 | 0.5 | 0.3 |
| | Accuracy_2 | 0.7 | 0.3 | 0.7 | 0.3 | 0.6 | 0.3 | 0.7 | 0.3 | 0.7 | 0.2 | 0.7 | 0.3 | 0.6 | 0.3 | 0.7 | 0.2 | 0.6 | 0.2 | 0.6 | 0.3 | 0.6 | 0.3 | 0.6 | 0.3 |
| | Accuracy_4 | 0.8 | 0.2 | 0.9 | 0.2 | 0.8 | 0.3 | 0.9 | 0.2 | 0.8 | 0.2 | 0.9 | 0.2 | 0.8 | 0.3 | 0.9 | 0.1 | 0.7 | 0.3 | 0.9 | 0.2 | 0.9 | 0.2 | 0.8 | 0.3 |
| | Confidence_1 | 4.9 | 1.5 | 5.3 | 1.1 | 4.9 | 1.2 | 5.1 | 1.3 | 4.6 | 1.2 | 5.4 | 1.0 | 5.1 | 1.0 | 4.8 | 0.9 | 5.1 | 0.8 | 5.2 | 1.0 | 5.4 | 0.9 | 4.8 | 1.0 |
| | Confidence_2 | 5.3 | 1.5 | 5.5 | 0.9 | 5.2 | 1.0 | 5.2 | 1.2 | 5.2 | 0.9 | 5.4 | 1.0 | 5.4 | 0.9 | 5.2 | 0.8 | 5.3 | 0.8 | 5.3 | 1.0 | 5.5 | 0.8 | 5.0 | 0.9 |
| | Confidence_4 | 5.5 | 1.2 | 5.8 | 0.7 | 5.6 | 0.8 | 5.8 | 1.1 | 5.5 | 0.8 | 5.8 | 0.8 | 5.6 | 0.8 | 5.8 | 0.8 | 5.6 | 0.9 | 6.0 | 0.5 | 5.9 | 0.6 | 6.0 | 0.6 |
| | Glance_Score | 1.0 | 0.4 | 1.1 | 0.4 | 0.9 | 0.3 | 1.1 | 0.4 | 1.1 | 0.3 | 1.1 | 0.3 | 0.9 | 0.4 | 1.1 | 0.3 | 0.9 | 0.4 | 1.0 | 0.3 | 1.0 | 0.3 | 0.9 | 0.3 |
| *Subjective Measures* | Legibility | 4.8 | 1.3 | 5.6 | 0.9 | 4.4 | 1.5 | 5.2 | 1.2 | 4.8 | 1.5 | 5.1 | 1.0 | 4.9 | 1.0 | 5.3 | 1.2 | 5.1 | 1.3 | 5.1 | 1.0 | 5.1 | 1.2 | 4.8 | 1.0 |
| | Predictability | 4.9 | 1.3 | 5.2 | 1.2 | 4.8 | 1.1 | 5.2 | 1.4 | 5.2 | 1.3 | 4.7 | 1.1 | 4.7 | 1.1 | 5.3 | 1.2 | 5.4 | 1.0 | 5.2 | 1.0 | 5.1 | 1.6 | 4.8 | 1.4 |

Color depicts the ordering of results from dark blue (best performing) to light blue (worst performing).
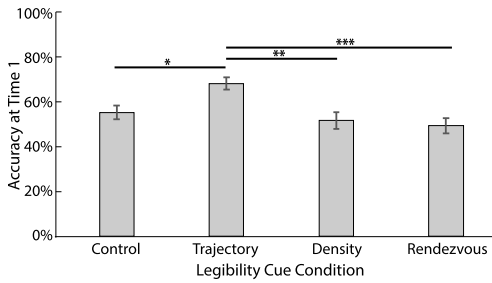


(a) Effects of legibility cue on the prediction accuracy at time 1
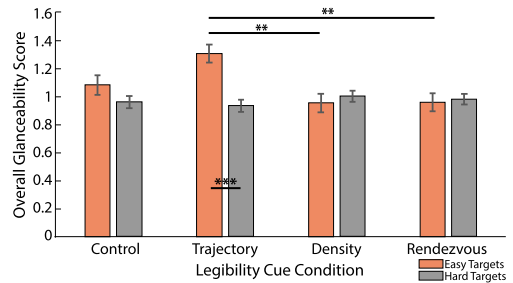


(b) Interaction effects between target difficulty and legibility cue conditions for glanceability score

Fig. 13. (a) The trajectory condition had significantly higher prediction accuracy than other conditions at time 1. (b) In terms of overall glanceability score, there is a significant difference between the target difficulty levels for trajectory condition but not for other conditions.

previous legibility study. For the analysis, we removed 9 participants with duplicate IP addresses as well as those who failed to follow the instructions correctly (i.e., did not fill out demographics information appropriately).

Participants reported ages ranging from 20 to 72 with mean = 35.7 and SD = 10.6. A total of 55.4% identified as men, 44.2% as women, and 0.4% as non-binary, while 23.4%, 63.2%, and 13.4% of participants reported education levels of middle/high school, college, and advanced degrees, respectively. After completing the experiment in 6.6 minutes on average, each participant received compensation at an hourly rate of approximately $15.00 US dollars.

## 6.7 Analysis

To investigate both main and interaction effects for the overall glanceability score, prediction confidences, and accuracies at times 1, 2, and 4, we ran a $4 \times 3 \times 2$ mixed-design ANOVA with two between-subject factors (i.e., condition and initial radius) and one within-subject factor (i.e., target difficulty). Then, we performed the Bonferroni-corrected post hoc tests on the statistically significant effects.

Similarly, we ran a $4 \times 3$ between-subjects ANOVA with with two between-subject factors (i.e., condition and initial radius) for the subjective ratings. Then, we performed the Bonferroni-corrected post hoc tests on the statistically significant effects.
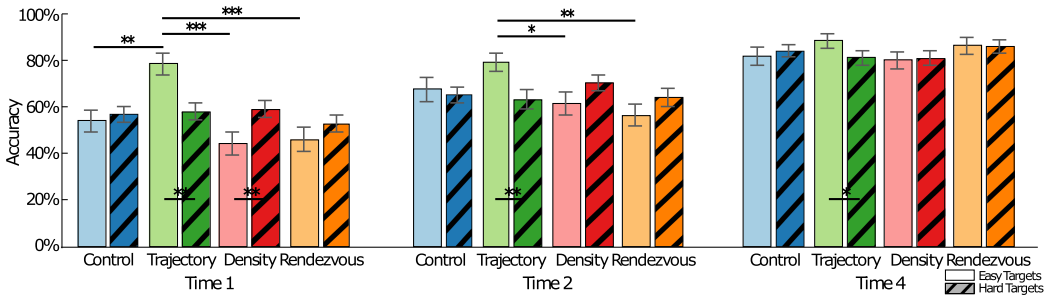
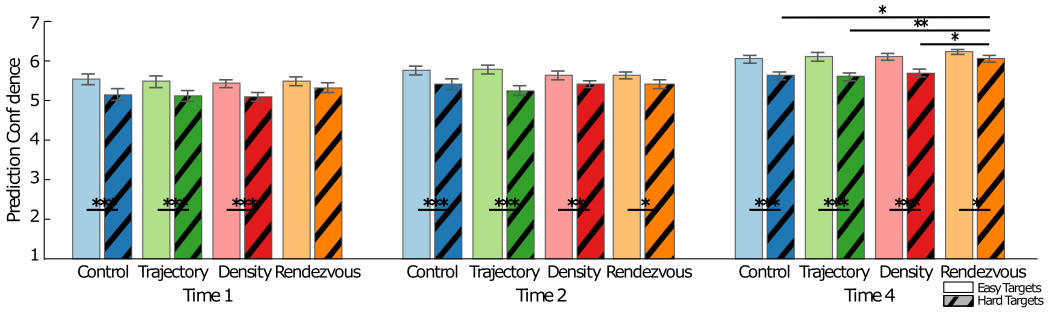Fig. 14. Interaction effects between target difficulty and legibility cue on prediction accuracy across time.



Fig. 15. Interaction effects between target difficulty and legibility cue on prediction confidence across time.

## 6.8 Results & Discussion

The descriptive statistics of Study 1 are shown in Table 3. All the bar graphs (Figure 13–16) plot the mean and standard error. Significantly different pairs indicated by a bar accompanied by varying numbers of $*$, where $*$: $.01 \leq p < .05$, $**$: $.001 \leq p < .01$, and $***$: $p < .001$. Only the statistically significant results are described below.

*6.8.1 Effects of Legibility Cue.* The legibility cue had significant effects on prediction accuracy at time 1 ($F(3, 248) = 6.5, p < 0.001, \eta^2 = 0.073$), as shown in Figure 13(a). At time 1, trajectory condition had significantly higher accuracy than other conditions. Diving deeper, we see from Figure 14 that the trajectory condition led to higher prediction accuracy than other conditions when the targets were hard at times 1 and 2. These results support H2.1, as the trajectory condition performs better than other conditions in the beginning but the performance becomes similar at time 4. However, both density and rendezvous conditions did not significantly improve accuracy or confidence at all times except at time 4, when rendezvous condition had higher confidence ratings than others for harder targets. These results fail to support H2.2–H2.3.

*6.8.2 Effects of Target Difficulty.* The target difficulty had significant effects on the overall glanceability score ($F(1, 218) = 11.3, p < 0.001, \eta^2 = 0.049$) as well as on prediction confidence at time 1 ($F(1, 248) = 46.6, p < 0.001, \eta^2 = 0.158$), time 2 ($F(1, 248) = 56.6, p < 0.001, \eta^2 = 0.186$), and time 4 ($F(1, 248) = 91.6, p < 0.001, \eta^2 = 0.27$). As expected, the overall glanceability score and the confidence ratings at all times were lower for harder targets.

As shown in Figures 13(b), 14, and 15, there were significant interaction effects between target difficulty and legibility cue on overall glanceability score ($F(3, 218) = 7.6, p < 0.001, \eta^2 = 0.095$), prediction accuracy at time 1 ($F(3, 248) = 7.8, p < 0.001, \eta^2 = 0.087$), and time 2 ($F(3, 248) = 4.9, p = 0.003, \eta^2 = 0.056$), as well as prediction confidence at time 4 ($F(3, 248) = 3.1$,
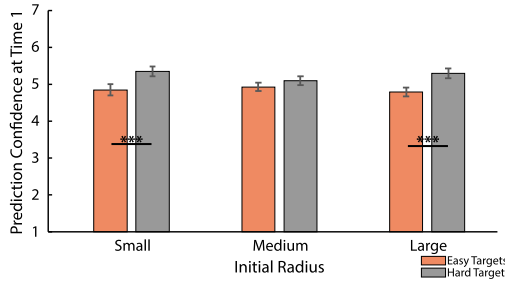
Fig. 16. Interaction effects between target difficulty and initial radius for prediction confidence at time 1.

$p = 0.027, \eta^2 = 0.036$). The results demonstrated that when the targets were easier, the trajectory-based legible motion had significantly higher prediction accuracy at time 1 ($F(1, 64) = 11.3, p = 0.001, \eta^2 = 0.151$), time 2 ($F(1, 64) = 10.3, p = 0.002, \eta^2 = 0.138$), and time 4 ($F(1, 64) = 5.4, p = 0.023, \eta^2 = 0.078$), as well as higher confidence ratings at time 1 ($F(1, 64) = 12.6, p < 0.001, \eta^2 = 0.164$), time 2 ($F(1, 64) = 31.3, p < 0.001, \eta^2 = 0.328$), and time 4 ($F(1, 64) = 34.7, p < 0.001, \eta^2 = 0.352$) supporting H2.4.

In contrast to the trajectory-based legible motion, the rendezvous-based legible motion had similar prediction accuracies across two target difficulties at all times while having significantly higher confidence ratings for easier targets at time 2 ($F(1, 61) = 6.1, p = 0.016, \eta^2 = 0.091$) and time 4 ($F(1, 61) = 5.1, p = 0.027, \eta^2 = 0.077$). In terms of the accuracy, H2.5 is supported but not in terms of the prediction confidence. However, it is worth noting that even for the confidence ratings, the p-values for rendezvous condition are only marginally lower than 0.05, whereas the p-values for the other conditions are much lower (i.e., $p < 0.001$), as shown in Figure 15. This suggests that the rendezvous-based legible motion is less affected by the target difficulty than other types of legible motion.

Last, the density-based legible motion had significantly higher prediction accuracy at time 1 with harder targets and had significantly higher confidence ratings for easier targets at all times. These results reject H2.6 in terms of both accuracy and confidence.

*6.8.3 Effects of Initial Radius.* The initial radius only had one significant interaction effect with the target difficulty on prediction confidence at time 1 ($F(2, 248) = 4.5, p = 0.012, \eta^2 = 0.035$), as shown in Figure 16. Specifically, the target difficulty had significant effects for small ($F(1, 79) = 21.3, p < 0.001, \eta^2 = 0.213$) and large initial radius conditions ($F(1, 78) = 25.8, p = 0.012, \eta^2 = 0.249$). However, as we did not observe any interaction effects between radius and legibility cue, we cannot accept H2.7.

## 7 DISCUSSION

From the two evaluations on the legibility and glanceability, we observed both similar and different trends. In general, we found that the rendezvous-based legible motion has the highest legibility score, whereas the trajectory-based motion is the most glanceable motion. However, a closer look at the results reveals that the trajectory condition performs especially well when the robots are heading toward the easier targets both in terms of glanceability and legibility, as shown in Figures 10(a) and 13(b). Thus, one interpretation of these results is that we should employ trajectory-based legible motion for the easier targets while using rendezvous behavior for the more difficult targets.

The legibility study results suggest that the trajectory-based legible motion does not perform as well when applied to a group of robots rather than a single robot. While prior work in legibility

of a single robot motion showed significant difference in legibility between straight and legible trajectory conditions [6, 17], we did not see such significant results between the two. This may suggest that the number of robots itself impacts the legibility, as there are more robots heading toward the goal. In contrast, we see significant improvement in glanceability, especially in the earlier segments of the trajectory for difficult targets. This may be due to the logarithmic nature of the trajectory-based motion where the robots initially move drastically away from non-targets compared to other types of legible motion. This exaggeration seems to help convey intent of the robots to the observers, especially when exposed to short snippets of the motion.

As we hypothesized, the target difficulty had significant impact on both the legibility and glanceability of the trajectory-based motion. For all measures, there was a significant difference between the easy and hard targets. This may be one of the reasons why we did not see a significant improvement in legibility, as our task may have been more difficult than those from prior work [6, 17]. On the contrary, for rendezvous and density conditions, the target difficulty had no effect on legibility and minor effects on glanceability mostly in terms of the prediction confidence. These results suggest that if consistency in legibility or glanceability (i.e., small variance in legibiliy/glanceability across different targets) is the main objective, the rendezvous or density-based motions are more suitable solutions than the trajectory-based motion.

Initial radius parameter had significant impacts on legibility in terms of prediction time, accuracy, and confidence. When the initial radius was smaller, participants were able to make the prediction earlier for both easy and hard targets, and more accurately for hard targets. When the robots are widely spread (i.e., larger initial radius value), rendezvous-based motion significantly outperforms all other types of motion potentially due to the more drastic merging effect. In contrast, for glanceability, initial radius parameter only had a significant interaction effect with target difficulty on prediction confidence at time 1. This is surprising, as the effects of initial radius are quite different from those for legibility. These discrepancies between the results from the two studies suggest that legibility and glanceability need to be both explicitly taken into account when designing swarm robot motion, because the most legible motion does not necessarily translate to the most glanceable motion.

## 8 LIMITATIONS & FUTURE WORK

This work provides preliminary investigation of different ways to enhance legibility of a multi-robot system. While we varied some parameters (e.g., the initial radius and the target difficulty) in addition to the legibility cue, we did not explore the effects of any other task parameters, such as target locations, which may affect the performance of different types of legible motions. The task setup is also quite simplistic with a top-down view and a row of aligned targets. Thus, the study results may not generalize for real-life scenarios where people will be viewing these robots at different angles and with various unstructured target compositions. Future studies should explore whether and how these different task parameters alter the trends observed from this work.

For the density-based legible motion, we used circular shapes for both the denser and sparser regions. Different shapes, such as a denser line inside a sparse circle pointing toward the goal, could be used to better direct user's attention toward the goal. In addition, the radius ratio between the denser and sparser regions is kept the same for different initial radius values. As we see that even changing the absolute size impacts the prediction confidence, we expect that increasing this ratio could further improve the legibility of the motion.

For the rendezvous-based legible motion, further investigation on the effects of initial radius values could be beneficial. While we did not observe any significant effect of radius on the performance of swarm behavior-based legible motion, we did not test a wide range of initial radii. With

a drastically larger radius, we expect to see an increase in performance for rendezvous conditions while other conditions will perform worse. Other parameters could be explored as well. For instance, instead of having robots rendezvous straight toward the goal, they could merge earlier in the trajectory and then move straight toward the goal as a more compact unit.

For glanceability study, we used a threshold of 250 ms, as we wanted to evaluate which motion can be processed pre-attentively. However, the study results may have been different if the threshold was longer or shorter. We found that when given the freedom to observe the entire trajectory, the result was different in that the rendezvous condition had the best performance instead of the trajectory condition. While our study results may serve as a starting point, further studies should be conducted to verify if our glanceability results still hold for the desired application.

While this work investigates these different types of legible motions independently, we see an opportunity to combine them in a novel way. The most straightforward way would be to have the robots rendezvous or form patterns that pre-attentively direct the user's attention toward the goal, while following the legible trajectory. Another method could involve manipulating when each cue is triggered to guide users toward the goal. For example, robots could first follow the legible trajectory and then rendezvous toward the goal. Future work could focus on optimizing these combinations.

Another way to combine these motions is to change the type of legible motion based on the context of the task, such as the difficulty or location of the target. As we observe significant effects of task parameters (i.e., target difficulty and initial radius) on legibility and glanceability, we could use the motion optimal for each setting. For instance, we would deploy trajectory-based motion for easy targets and rendezvous-based motion for hard targets when optimizing based on the legibility score. One potential issue with this approach is that users may find the mixture of motions confusing, as there is a lack of consistency across trials. Further studies are needed to evaluate whether people can adapt to this hybrid approach or prefer a more consistent approach.

This work focuses on conditions where all of the robots are moving toward only one static goal. In the real world, there will be cases when robots will need to reach and manipulate multiple moving objects. While Capelli et al. explored this problem of multiple objects, they applied and used the same motion parameters (i.e., dispersion, trajectory, and stiffness) that they used for one target instead of exploring new ways to tackle the multi-target problem [9]. In the future, we plan to explore different combinations of the features, both from this article and prior works, to address the multiple moving target problem.

## 9 CONCLUSION

Given the potential for multi-robot systems that interact and collaborate with people, it is important to equip these robots with intent-expressive movements that facilitate such experiences. In this work, we explored the use of trajectory, collective behavior, and pre-attentive processing features to generate legible and glanceable swarm robot motion. We ran two online studies to compare the legibility and glanceability of the different legibility cues. The study results suggest that the rendezvous behavior-based motion is the most legible with the highest confidence, whereas the trajectory-based motion has the highest glanceability. We also observed significant effects of task parameters, such as the radius of the initial circle that surrounds the robots and the location of the targets, which determine the difficulty level. Rather than a one-size-fits-all solution, generating legible and glanceable swarm robot motion may require a more complex solution that consists of a combination of these different legibility cues based on the context of the task.

## ACKNOWLEDGMENTS

# REFERENCES

[1] Javier Alonso-Mora, Andreas Breitenmoser, Martin Rufli, Roland Siegwart, and Paul Beardsley. 2012. Image and animation display with multiple mobile robots. *Int. J. Robot. Res.* 31, 6 (2012), 753–773.

[2] C. Chace Ashcraft, Michael A. Goodrich, and Jacob W. Crandall. 2019. Moderating operator influence in human-swarm systems. In *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (SMC'19)*. IEEE, 4275–4282.

[3] Alexandra Bacula, Jason Mercer, and Heather Knight. 2020. Legible light communications for factory robots. In *Proceedings of the ACM/IEEE International Conference on Human-robot Interaction*. 119–121.

[4] Aaron Becker, Golnaz Habibi, Justin Werfel, Michael Rubenstein, and James McLurkin. 2013. Massive uniform manipulation: Controlling large populations of simple robots with a common input signal. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'13)*. IEEE, 520–527.

[5] Christopher Bodden, Daniel Rakita, Bilge Mutlu, and Michael Gleicher. 2016. Evaluating intent-expressive robot arm motion. In *Proceedings of the 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN'16)*. IEEE, 658–663.

[6] Christopher Bodden, Daniel Rakita, Bilge Mutlu, and Michael Gleicher. 2018. A flexible optimization-based method for synthesizing intent-expressive robot arm motion. *Int. J. Robot. Res.* 37, 11 (2018), 1376–1394.

[7] S. Boerema, Randy Klaassen, Hendrikus J. A. op den Akker, and Hermanus J. Hermens. 2012. Glanceability evaluation of a physical activity feedback system for office workers. In *Proceedings of the6th International Symposium on eHealth Services and Technologies*. 52–57.

[8] Eric Bonabeau, Marco Dorigo, and Guy Theraulaz. 1999. *Swarm Intelligence: From Natural to Artificial Systems*. Oxford University Press.

[9] Beatrice Capelli, Cristian Secchi, and Lorenzo Sabattini. 2019. Communication through motion: Legibility of multi-robot systems. In *Proceedings of the International Symposium on Multi-robot and Multi-Agent Systems (MRS'19)*. IEEE, 126–132.

[10] Beatrice Capelli, Valeria Villani, Cristian Secchi, and Lorenzo Sabattini. 2019. Understanding multi-robot systems: On the concept of legibility. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'19)*. IEEE, 7355–7361.

[11] Jessie Y. C. Chen, Michael J. Barnes, and Michelle Harper-Sciarini. 2010. Supervisory control of multiple robots: Human-performance issues and user-interface design. *IEEE Trans. Syst., Man, Cyber., Part C (Applic. Rev.)* 41, 4 (2010), 435–454.

[12] Jorge Cortes, Sonia Martinez, Timur Karatas, and Francesco Bullo. 2004. Coverage control for mobile sensing networks. *IEEE Trans. Robot. Autom.* 20, 2 (2004), 243–255.

[13] Iain D. Couzin, Jens Krause, Richard James, Graeme D. Ruxton, and Nigel R. Franks. 2002. Collective memory and spatial sorting in animal groups. *J. Theoret. Biol.* 218, 1 (2002), 1–12.

[14] Jacob W. Crandall, Michael A. Goodrich, Dan R. Olsen, and Curtis W. Nielsen. 2005. Validating human-robot interaction schemes in multitasking environments. *IEEE Trans. Syst., Man, Cyber.-Part A: Syst. Hum.* 35, 4 (2005), 438–449.

[15] Gergely Csibra and György Gergely. 2007. "Obsessed with goals": Functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychologica* 124, 1 (2007), 60–78.

[16] Anca Dragan and Siddhartha Srinivasa. 2013. Generating legible motion. In *Proceedings of the Robotics: Science and Systems Conference*. DOI:https://doi.org/10.15607/RSS.2013.IX.024

[17] Anca D. Dragan, Kenton C. T. Lee, and Siddhartha S. Srinivasa. 2013. Legibility and predictability of robot motion. In *Proceedings of the 8th ACM/IEEE International Conference on Human-robot Interaction*. IEEE Press, 301–308.

[18] Howard E. Egeth and Steven Yantis. 1997. Visual attention: Control, representation, and time course. *Ann. Rev. of Psychol.* 48, 1 (1997), 269–297.

[19] Kerstin Fischer, Lars C. Jensen, Stefan-Daniel Suvei, and Leon Bodenhagen. 2016. Between legibility and contact: The role of gaze in robot approach. In *Proceedings of the 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN'16)*. IEEE, 646–651.

[20] Charles L. Folk and Roger Remington. 2006. Top-down modulation of preattentive processing: Testing the recovery account of contingent capture. *Vis. Cog.* 14, 4–8 (2006), 445–465.

[21] Rúben Gouveia, Fábio Pereira, Evangelos Karapanos, Sean A. Munson, and Marc Hassenzahl. 2016. Exploring the design space of glanceable feedback for physical activity trackers. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 144–155.

[22] Ronald L. Graham, Boris D. Lubachevsky, Kari J. Nurmela, and Patric R. J. Östergård. 1998. Dense packings of congruent circles in a circle. *Disc. Math.* 181, 1–3 (1998), 139–154.

[23] Christopher Healey and James Enns. 2011. Attention and visual memory in visualization and computer graphics. *IEEE Trans. Vis. Comput. Graph.* 18, 7 (2011), 1170–1188.

[24] Mrinal Kalakrishnan, Sachin Chitta, Evangelos Theodorou, Peter Pastor, and Stefan Schaal. 2011. STOMP: Stochastic trajectory optimization for motion planning. In *Proceedings of the IEEE International Conference on Robotics and Automation*. IEEE, 4569–4574.

[25] Lawrence H. Kim and Sean Follmer. 2017. Ubiswarm: Ubiquitous robotic interfaces and investigation of abstract motion as a display. *Proc. ACM Interact., Mob., Wear. Ubiq. Technol.* 1, 3 (2017), 66.

[26] Harold W. Kuhn. 1955. The Hungarian method for the assignment problem. *Nav. Res. Log. Quart.* 2, 1–2 (1955), 83–97.

[27] John Lasseter. 1987. Principles of traditional animation applied to 3D computer animation. In *ACM Siggraph Computer Graphics*, Vol. 21. ACM, 35–44.

[28] Mathieu Le Goc, Lawrence H. Kim, Ali Parsaei, Jean-Daniel Fekete, Pierre Dragicevic, and Sean Follmer. 2016. Zooids: Building blocks for swarm user interfaces. In *Proceedings of the 29th Symposium on User Interface Software and Technology*. ACM, 97–109.

[29] Mathieu Le Goc, Charles Perin, Sean Follmer, Jean-Daniel Fekete, and Pierre Dragicevic. 2018. Dynamic composite data physicalization using wheeled micro-robots. *IEEE Trans. Vis. Comput. Graph.* 25, 1 (2018), 737–747.

[30] Christina Lichtenthäler and Alexandra Kirsch. 2016. Legibility of robot behavior: A literature review. https://hal.archives-ouvertes.fr/hal-01306977.

[31] Yugang Liu and Goldie Nejat. 2013. Robotic urban search and rescue: A survey from the control perspective. *J. Intell. Robot. Syst.* 72, 2 (2013), 147–165.

[32] Tara Matthews, Jodi Forlizzi, and Stacie Rohrbach. 2006. Designing glanceable peripheral displays. In *Proceedings of the 6th Conference on Designing Interactive Systems*.

[33] Christoforos I. Mavrogiannis, Wil B. Thomason, and Ross A. Knepper. 2018. Social momentum: A framework for legible navigation in dynamic multi-agent environments. In *Proceedings of the ACM/IEEE International Conference on Human-robot Interaction*. 361–369.

[34] Francesco Mondada, Giovanni C. Pettinaro, Andre Guignard, Ivo W. Kwee, Dario Floreano, Jean-Louis Deneubourg, Stefano Nolfi, Luca Maria Gambardella, and Marco Dorigo. 2004. SWARM-BOT: A new distributed robotic concept. *Auton. Robots* 17, 2–3 (2004), 193–221.

[35] Amir M. Naghsh, Jeremi Gancet, Andry Tanoto, and Chris Roast. 2008. Analysis and design of human-robot swarm interaction in firefighting. In *Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 255–260.

[36] Stefanos Nikolaidis, Anca Dragan, and Siddharta Srinivasa. 2016. Viewpoint-based legibility optimization. In *Proceedings of the 11th ACM/IEEE International Conference on Human Robot Interaction (HRI'16)*. IEEE Press, 271–278.

[37] Atsuyuki Okabe and Atsuo Suzuki. 1997. Locational optimization problems solved through Voronoi diagrams. *Eur. J. Oper. Res.* 98, 3 (1997), 445–456.

[38] Ayberk Özgür, Séverin Lemaignan, Wafa Johal, Maria Beltran, Manon Briod, Léa Pereyre, Francesco Mondada, and Pierre Dillenbourg. 2017. Cellulo: Versatile handheld robots for education. In *Proceedings of the 12th ACM/IEEE International Conference on Human-robot Interaction (HRI'17)*. IEEE, 119–127.

[39] Redmond R. Shamshiri, Cornelia Weltzien, Ibrahim A. Hameed, Ian J. Yule, Tony E. Grift, Siva K. Balasundram, Lenka Pitonakova, Desa Ahmad, and Girish Chowdhary. 2018. Research and development in agricultural robotics: A perspective of digital farming. *International Journal of Agricultural & Biological Engineering* 11, 4 (2018), 1–14.

[40] Nathan Ratliff, Matt Zucker, J. Andrew Bagnell, and Siddhartha Srinivasa. 2009. CHOMP: Gradient optimization techniques for efficient motion planning. In *IEEE International Conference on Robotics and Automation*. IEEE, 489–494.

[41] Craig W. Reynolds. 1987. Flocks, herds and schools: A distributed behavioral model. In *Proceedings of the 14th Conference on Computer Graphics and Interactive Techniques*. 25–34.

[42] Eric Rosen, David Whitney, Elizabeth Phillips, Gary Chien, James Tompkin, George Konidaris, and Stefanie Tellex. 2019. Communicating and controlling robot arm motion intent through mixed-reality head-mounted displays. *Int. J. Robot. Res.* 38, 12–13 (2019), 1513–1526.

[43] Adriane E. Seiffert, Sean Timothy Hayes, Caroline E. Harriott, and Julie A. Adams. 2015. Motion perception of biological swarms. In *Proceedings of the 37nd Annual Conference of the Cognitive Science Society*. Cognitive Science Society, 2128–2133.

[44] Jamie Snape, Jan van den Berg, Stephen J. Guy, and Dinesh Manocha. 2011. The hybrid reciprocal velocity obstacle. *IEEE Trans. Robot.* 27, 4 (2011), 696–706.

[45] Daniel Szafir, Bilge Mutlu, and Terrence Fong. 2014. Communication of intent in assistive free flyers. In *Proceedings of the 9th ACM/IEEE International Conference on Human-robot Interaction (HRI'14)*. IEEE, 358–365.

[46] Daniel Szafir, Bilge Mutlu, and Terrence Fong. 2015. Communicating directionality in flying robots. In *Proceedings of the 10th ACM/IEEE International Conference on Human-robot Interaction (HRI'15)*. IEEE, 19–26.

[47] Anne Treisman. 1985. Preattentive processing in vision. *Comput. Vis., Graph. Image Proc.* 31, 2 (1985), 156–177.

[48] Michael Walker, Hooman Hedayati, Jennifer Lee, and Daniel Szafir. 2018. Communicating robot motion intent with augmented reality. In *Proceedings of the ACM/IEEE International Conference on Human-robot Interaction*. 316–324.

[49] Phillip Walker, Michael Lewis, and Katia Sycara. 2016. Characterizing human perception of emergent swarm behaviors. In *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC'16)*. IEEE, 002436–002441.

[50] Andrew Witkin and Michael Kass. 1988. Spacetime constraints. *ACM Siggraph Comput. Graph.* 22, 4 (1988), 159–168.

[51] Min Zhao, Rahul Shome, Isaac Yochelson, Kostas Bekris, and Eileen Kowler. 2016. An experimental study for identifying features of legible manipulator paths. In *Experimental Robotics*. Springer, 639–653.