DESIGNING IN SITU INTERACTION WITH UBIQUITOUS ROBOTS

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF MECHANICAL ENGINEERING
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Lawrence H. Kim
August 2020

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Sean Follmer, Primary Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

James Landay

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Allison Okamura

Approved for the Stanford University Committee on Graduate Studies.

Stacey F. Bent, Vice Provost for Graduate Education

This signature page was generated electronically upon submission of this dissertation in electronic format. An original signed hard copy of the signature page is on file in University Archives.

Abstract

We are nearing an era of ubiquitous robots. They will be in our home doing our dishes and laundry, in our workplace helping us carry heavy loads and performing repetitive tasks, and on the road driving us to our destinations and delivering packages. Due to this wide variety of potential tasks, robots will take different forms, sizes, and numbers resulting in a diverse set of robots ranging from a humanoid robot, specifically designed for interacting with humans, to a swarm of small mobile robots, used to collectively gather information in a large unstructured environment. Of these, particularly important are swarms of centimeter-scale non-anthropomorphic robots. Their appearance and size enable seamless transition between blending into the environment, becoming "invisible" to the users, and doing everyday tasks such as object manipulation through collective means and interaction with users through their motion and touch. While roboticists have made great strides in enabling swarms to move objects and control formation, the topic of *in situ human-multirobot interaction* is crucial and timely, given the rise in interaction with a group of robots in our daily environment.

My thesis centers on **ubiquitous human-centered robotics** that facilitate in situ interactions with people. I introduce the concept of Ubiquitous Robotic Interfaces (URIs), multi-robot systems that are capable of mobility, manipulation, sensing, display, and interaction both with the user and the environment. In order to better understand interaction with this new class of robots, specifically on how to enable robots to **effectively interact and communicate with people**, I develop new tabletop swarm robot platforms, design multimodal multi-robot interaction, and evaluate them through a set of human subject experiments. As non-humanoid robots lack anthropomorphic features that have been heavily relied upon for human interaction, I leverage a feature common across most, if not all, types of robots: mobility. Specifically, I study how to visually display information like expressivenss and intent to users through multi-robot motion, and how to convey information

through touch from the robots. Closing the loop, I also investigate how people naturally communicate with a swarm of robots. In summary, this thesis provides a deeper understanding of interactive ubiquitous robots as well as generating a rich set of guidelines to increase effective human-multirobot interaction.

Acknowledgments

First and foremost, I would like to thank my advisor Sean Follmer for his guidance and support. Sean, you have taught me to ask the two most important questions in research: the *why* and *how*. *Why* does it matter? *How* can this new technology benefit our society? As engineers, we often get lost in the technicality of creating technologies without these goals in mind. Thank you Sean for always reminding me the purpose of research - I will continue to put this into practice as I embark on a journey as an HCI researcher.

I am also very grateful to the other members of my committee: Allison Okamura, James Landay, Wendy Ju, and Martin Fischer. Allison - you taught me the fundamentals of research with your attention to detail, punctuality, and passion for science. Thank you for opening the door to the world of haptics and academia. James - thank you for pushing me to ask important questions and helping me see the bigger picture. Wendy - your pioneering work in design and HRI has always been a source of inspiration for me. Thank you for your creative suggestions and constructive feedback on my research. And thank you Martin, without whom my defense would not have been possible, for encouraging me to think hard about new applications of my work.

I have been so fortunate to have many other collaborators and co-authors over the years, without whom this work would not have been possible: Mathieu Le Goc, Pierre Dragicevic, Jean-Daniel Fekete, and Ali Parsaei for co-building *Zooids*, the foundational piece for this thesis; Veronika Domova, Daniel Drew, Yiwei Zhao, Jane E, Griffin Dietz, Peter Washington, and Yuqi Yao for co-designing user studies with swarm robots; and Sungjune Jang, Kesler Tanner, and Hiroshi Ishii for co-authoring my first HCI paper. I also would like to thank Ali Israr for his guidance at Facebook and for providing me with invaluable industry research experience on haptics.

Working in SHAPE Lab has been an incredible experience: we brainstormed research ideas, prepared demos together, took turns cooking for TTLs, and traveled to conferences like CHI and UIST. Thank you all for being amazing colleagues, both past and current Shapees: Inrak Choi, Alexa Siu, Eric Gonzalez, Parastoo Abtahi, Elyse Chase, Evan Strasnick, Zack Hammond, Jingyi Li, Kai Zhang, Renee Chao, Andrea Brand-Sanchez, Shenli Yuan, Seunghwan Lee, Yuqi Yao, Ernesto Ramirez, Ye Wang, Hieu Pham, Julea Chin, Allen Zhao, Abena Boadi-Agyemang and others.

I have had some of the most memorable years of my life at Stanford, and I owe that to my new and old friends: Chris Ploch, Jung Hwa Bae, Michael Lin, Daniel Shin, Jongho Kim, Peter Lee, Eric Shinn, Seohyun Kim, David Kwak, Adam Jung and many more. I also want to thank people from KSAS, KME, Tobamte, and Geumbamte groups for helping me get through graduate school with many social and recreational activities.

Last but not least, I would like to thank my family for their continuous love and support. Mom and dad-thank you for letting me pursue my passion and yet always having faith in me. I would not be here without your sacrifice and guidance. Thank you, Fred, for always being the uplifting figure in our house. Even the short video calls with you have helped me get through the roughest moments. And Jen-thank you for being my best friend, my biggest critic, and my biggest supporter.

This work was supported in part by the Samsung Scholarship and the Hasso Plattner Design Thinking Research Program.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Robots have transitioned from factories to our homes and streets. Their duties have been expanded from dirty, dangerous, and dull tasks like manufacturing or parcel sorting, to tasks in highly unstructured environments that require real-time interaction with people such as autonomous transportation or firefighting. As robots continue to permeate our environment and the cost of these robots decreases, the number of robots that we encounter and interact with will increase. Robots will take different sizes and forms depending on the application – distributed sensing may benefit from networks of small bio-inspired bots, while autonomous transportation requires larger meter-scale vehicles. As the number and types of robots expand, robots will subsequently become ubiquitous and embedded in our everyday lives.

One field that makes this future of robotics possible is swarm robotics. Swarm robotics is an emerging field within robotics that studies how to coordinate large groups of relatively simple robots through the use of local, distributed rules. A swarm of robots can collectively accomplish complex tasks while being robust and flexible due to its large degrees of freedom and homogeneity. Past researchers have built swarm robotic systems [6–8] and used such systems to complete complex tasks such as pattern formation [8], object manipulation [9], and distributed sensing [10], through collective means. These swarms of robots could be used for many different applications including search-and-rescue, agriculture, package sorting and delivery, and domestic tasks. As more robots become part of our daily lives and inhabit our environment, their ability to fluently interact with people becomes more important. However, only a limited number of researchers have explored in situ interaction between these robots and people where users communicate with robots without a separate

dedicated interface such as a joystick [11] or a graphical user interface [12]. Given the pervasive nature of such ubiquitous robots, it is important to understand how to best design and implement human-multirobot interaction. For interaction, swarms of centimeter-scale robots are especially interesting. First, they are small enough to blend into the environment enabling them to seamlessly transition between being the focus of our attention and being embedded in our surroundings. They are also small enough to collectively interact with users, but large enough for each robot to be perceived as an individual animate agent. A group of small robots can also combine to manipulate heavy objects [13, 14]. Thus, this thesis investigates how users should interact with a collection of centimeter-scale robots.

Human-Robot Interaction (HRI) researchers, Bartneck and Forlizzi, have identified some key properties of a single social robot as demonstrated in Figure 1.1 [4]. While earlier work in HRI predominantly investigated interaction with anthropomorphic or zoomorphic robots that have faces and/or limbs [15], these features are often costly and superfluous for many non-interaction tasks [16], lead users to overestimate the robots' social skills [4], and limit how easily robots can blend into user's environment. Instead of relying on robots that resemble intelligent organisms, researchers have found ways to equip simple robots with expressive capabilities by leveraging their motion [17–20]. Knight and Simmons apply the Laban Effort System used in dance [21] to generate expressive motions for a simple mobile robot with planar motion [17] and a robot head with 3-axis rotational movements [22]. In a slightly different approach, the principles of animation have also been shown to enhance the readability and expressiveness of a single robot motion [20, 23].

However, it remains unclear how the results from prior work on expressive motion of a single robot could be applied to a larger group of robots. Only a few efforts have explored how the number of robotic agents changes the dynamics of the interaction. Vázquez et al. studied the spatial patterns formed during a human group conversation with a robot in order to automate socially appropriate gaze and body orientation for a furniture robot [24,25]. Chang et al. demonstrated that people playing in groups behave more competitively towards the robots than individual human players [26]. As many tasks involve more than two members, an important question remains: how can we understand the effects of the number of social agents, ranging from one-on-one dyad interaction to one-to-multirobot settings to many-to-many interaction? While all group settings are important, I first focus on investigating the interaction between a group of small mobile robots and a single user in this thesis because

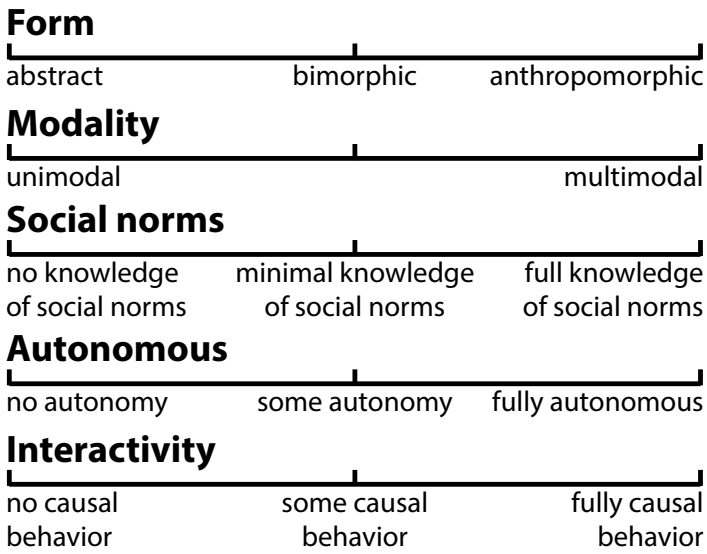


Figure 1.1: Framework for classification of social robots developed by Bartneck and Forlizzi [4] ©2004 IEEE

small robots can afford collective interaction and seamlessly transition between tasks and interaction.

1.1 Challenges

When designing in situ interaction with ubiquitous multi-robot systems, there are many challenges that arise. First, effective coordination of a large group of robots in a comprehensible manner is extremely difficult due to its high number of degrees of freedom. Ubiquitous robots also need to support interaction across different contexts, with people of varying backgrounds and expertise. Lastly, HRI designers require understanding of how different modalities should be used in the context of human-multirobot interaction. As human interaction is inherently multi-sensory [27, 28], it is important to understand how and when users use different sensory channels to control a group of robots and how the robots should leverage these modalities to convey information.

1.1.1 Interaction with *Multi-Robot* systems

A multi-robot system has significantly more degrees of freedom than a single robot. While this may enable a multi-robot system to become much more expressive and have a higher information throughput, it also presents challenges in understanding how to best coordinate these large number of robots for effective information display as well as how to design an interface for easy user-friendly control of a swarm of robots.

In an effort to display comprehensible information to users with multi-robot systems, a naive approach would be to apply findings from a single robot to each of the robots. However, that approach may not lead to the same results. Research has shown that people's perception can be affected just by a change in the number of robots [29]. Thus, researchers need to create a different approach that can truly take advantage of multi-robot systems.

Notably, despite the previous effort to develop in situ control methods for a large number of robots (i.e., without a dedicated interface like mouse and keyboard) [30, 31], none have investigated how the number of robots influence how people choose to control the robots. As users will interact with varying number of robots, designers need to understand if and how this quantity influences users' preferred control vocabulary.

1.1.2 Supporting interaction with *Ubiquitous* robots

Robots will become increasingly versatile, and be used for various tasks such as navigation, object manipulation, and interaction with people. Contrary to prior industrial robots, future robots will need to transition between task and interaction: adapting their behaviors to accommodate users, and communicating their internal states even while completing their non-interaction tasks. This is especially challenging for simple non-humanoid robots, as they lack additional degrees of freedom that humanoid robots have such as anthropomorphic features that are highly suitable for social interaction [15,32]. While recent HRI works have shown several ways to layer expressiveness and intent into a single robot motion [17, 19], a swarm of robots have additional expressive powers that warrant separate investigation.

Because these robots will interact not only with researchers in the lab but also naive users without computer science or robotics background, developers need to design an interface to support ubiquitous interaction. While there have been several proximal gesture-based control schemes for multi-agent systems [30, 31], the mapping between the proposed gestures and control commands is semi-arbitrary and not reflective of the users' preference or

intuition. To enable robots to interact fluently with the general public, we need to derive and use a user-centered interaction set.

1.1.3 Interacting through *Multimodal* sensory channels

Human interaction with the world is inherently multimodal [27, 28]. We perceive, communicate, and express our thoughts through multiple channels including aural, visual, and haptic means. Most HCI and HRI research has focused primarily on the visual and auditory senses because they are the more common and dominant communication methods. Haptics warrants more attention, because touch is an essential part of building rapport in human-human interaction [33–36]. Recent works in HRI have shown the positive effects of touch with robots as well [37,38]. Thus, it is important to leverage multiple sensory channels for both input and output in human-multirobot interaction.

As almost all robots can move, they have the ability to interact with users either visually or through direct touch. Motion has been frequently used in prior research as it can be used with any mobile robot, even at a distance from the user. In contrast, touch has only recently been studied for interaction with mostly humanoid robots. As non-humanoid robots also have the capacity for haptic interaction, further investigations are needed on how touch can be used with varying numbers of simple mobile robots.

In terms of controlling robots, it is important to understand how people use visual, aural, and haptic modalities for different commands. Most prior work related to proximal control of multi-robot systems only leverage speech and gesture, and utilize a fixed interaction set that does not take into account of user's preferred modality. In order to support users with different preferences and technical backgrounds, we need to develop a multimodal control language.

1.2 Approach

This section explains the approach adopted in this thesis to address the challenges introduced in the previous section.

1.2.1 Building *Interactive* Multi-Robot Platforms

Roboticists have developed numerous multi-robot or swarm robot platforms [6–8, 12–14, 39, 40]. However, none were designed or built for real-time interaction with people because

they cannot update robot formation at an interactive rate (ideally in the order of one second [41, 42]) or lack sensors that can detect user input for in situ interaction. To address this, Mathieu Le Goc and I built Zooids, a new swarm robot platform that can support real-time interaction enabled by the robots' fast motors and touch sensing capabilities [5]. Details are described in Chapter 3. I also built different adaptations of Zooids to either render more accurate movement [1] or output a larger range of forces for haptic applications [2]. This platform allowed me to rapidly prototype different behaviors with real robots and test my hypotheses with people.

1.2.2 Designing *Multi-Robot* Behaviors

My research on multi-robot behaviors draws on findings from relevant fields such as swarm intelligence and visual perception, and computational tools like optimization.

In order to coordinate large numbers of robots, a scalable solution is needed. Swarms in nature are able to behave in a complex manner without any central coordination through simple distributed rules [43]. Researchers have shown that people can rapidly detect different types of swarm behaviors even with high levels of noise present [44]. I see an opportunity to use collective behaviors to embed additional information even while completing a separate task. For instance, a swarm of robots rendezvousing toward an object instead of flocking in rigid formation, can elicit different perceptions from observers as well as convey additional information about the task such as the urgency of the situation or the intent of the robots. Seizing this research opportunity, I investigate the use of collective behaviors to layer expressiveness and intent on top of swarm robot motion in Chapters 4 and 5.

In contrast to collective behaviors, which are scalable methods to coordinate swarms of robots, another approach in designing multi-robot behaviors is to leverage findings from the field of visual perception. While humanoid robots closely resemble people, a collection of circular mobile robots is similar to an array of points. Thus, rather than relying on anthropomorphic features, we could potentially apply results from visual perception research to design behaviors that are optimized to human's visual processing system. Specifically, I utilized pre-attentive visual processing features [45] to design robot formations that can quickly convey intent of swarm motion in Chapter 5.

Swarms of robots have high degrees of freedom that require complex coordination. In order to find an optimal solution to the coordination problem, I used optimization frameworks to compute the optimal coordination based on the given objectives and constraints.

In Chapter 5, I used optimizations to design different types of legible motions using different objectives and constraints.

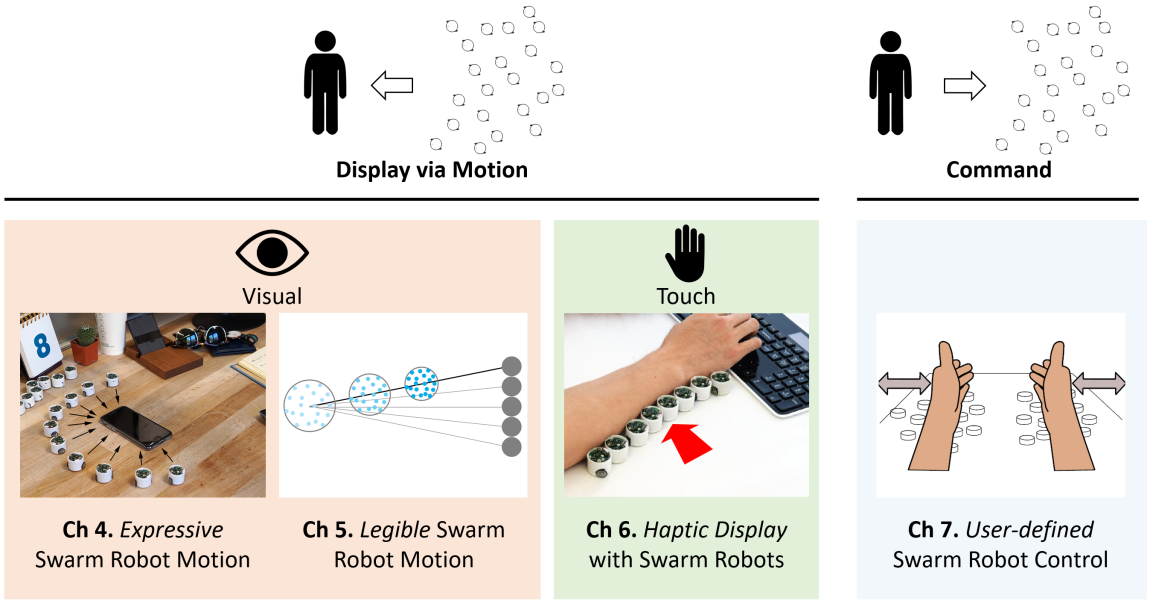


Figure 1.2: In this thesis, I explore the use of motion to display information like intent and affect visually and through touch to users, and investigate how users command a swarm of robots.

1.3 Thesis Statement

The goal of this thesis is to **design a multimodal in situ interaction paradigm for ubiquitous multi-robot systems**. To that end, I will first introduce the concept and implementation of ubiquitous robotic interfaces (URIs), multi-robot interfaces that are capable of mobility, manipulation, sensing, and interaction both with user and environment. The introduced Zooids platform, a prototype of URIs, will be used to design and evaluate different aspects of interaction with ubiquitous robots. I first broke down the interaction into two components as shown in Figure 1.2: output from the robots to the user, and input from the user to the robots. In terms of output, I focused on how to utilize the key feature of any simple non-humanoid robot: its mobility. Because these robots can move both far from and near people, I studied how to display meaningful information through their motion [1] and through direct touch [2]. In contrast, humans interact through many sensory channels

such as vision, hearing, and touch. To discover trends across people, I conducted studies to understand how people use combinations of sensory modalities for human-multirobot interaction [3].

1.4 Dissertation Outline

Chapter 2 provides an overview of the prior work, on which this dissertation builds. I describe relevant papers in ubiquitous robotics and swarm robotics. Then, I detail existing works related to various aspects of human-robot interaction such as robot motion perception, legibility of robot motion, haptic interaction with a robot, and human control of multi-agents systems.

In Chapter 3, I introduce the concept of Ubiquitous Robotic Interfaces (URIs). I outline and describe the key components of URIs. Zooids, an example implementation of URIs, is introduced with details on both its hardware and software design.

In the next three Chapters, I study how to best display information to users, leveraging the physicality of a swarm of non-humanoid robots, Zooids. Chapters 4 and 5 discuss ways to layer expressiveness and intent on top of swarm robot motion, while Chapter 6 explores the idea of using swarm robots as a haptic display and a medium for remote social touch.

Chapter 4 investigates the use of abstract motion as a display. Instead of relying on formation of patterns to convey information, we can layer expressiveness on top of existing motion by utilizing abstract motion with various swarm behaviors. Specifically, I conducted a large-scale online study to better understand how people perceive different types of swarm robot motion. This work was published in The Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT) [1] and presented at the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing.

In Chapter 5, I discuss another crucial aspect of swarm robot motion, legibility. This work builds on on prior work in legible single robot motion and applies insights from cognitive science as well as bio-inspired swarm behaviors to enhance the legibility of the swarm robot motion. It also introduces the concept of glanceability, which measures if observers can pre-attentively understand the intent of the robots, and describes an online study to compare performances of different enhancement methods in terms of legibility and glanceability.

In Chapter 6, I explore how to convey information through touch from swarm robots. While some researchers have investigated haptic interaction with a single robot [46, 47], none have investigated the use of swarm robots for haptics. After introducing the design space, I conducted a perception study to better understand the how expressive this type of system can be as well as the specific effects that each parameter of the system has on human perception. In addition, I explore the use of these robots for remote social touch application and ran an elicitation study to understand how people use robots to convey different types of social touch. This work was published and presented at the 2019 ACM Conference on Human Factors in Computing Systems (CHI) [2].

Chapter 7 shifts from the three previous chapters, where the design of the robot behaviors is the main concern, and describes a study on how humans naturally interact with a swarm of non-humanoid robots. Most of the current interfaces to control a swarm of robots in situ prioritize current input devices rather than an intuitive and user-centered system. To address this, I conducted an elicitation study to understand how humans behave without any instruction to control a swarm of robots, providing insights for the ideal interface. This work was published and presented at the 2020 ACM Conference on Human Factors in Computing Systems (CHI) [3].

Chapter 8 summarizes the dissertation and its contributions. I then discuss the implications of this work followed by potential future directions.

Chapter 2

Related Work

The upcoming swarms of ubiquitous robots present a challenge to interaction designers in two aspects. First, their non-anthropomorphic design do not include features that are traditionally associated with social interaction, such as a face or limbs. Second, these autonomous robots are often deployed at a large scale and thus will increase the number of agents involved during the interaction. Several recent efforts in HRI attempted to address this by exploring approaches to enhance the social and interactive aspect of increasingly common, task-oriented robots without any anthropomorphic or zoomorphic features [25, 48, 49]. Instead of relying on costly and often obtrusive features, researchers found ways to leverage a component that almost all robots have: mobility. The mobility of the robots has been used to implicitly layer expressiveness or intent into these non-humanoid robots [17, 19, 20]. However, most only focus on one-on-one HRI interaction instead of one-to-many interaction that is crucial when swarms of autonomous agents are involved.

Designing multimodal interaction with ubiquitous multi-robot systems is a challenge that spans across many disciplines. To address this challenge, we need to develop a hardware platform consisting of a large group of robots, support ubiquitous interaction for people both with or without expertise in robotics or HRI, and leverage multiple sensory channels to enable multimodal interaction as people socially interact through verbal and nonverbal communication. Here, I review prior work that has led to development of multi-robot systems such as ubiquitous robotics and swarm robotics. Then, I provide a overview of research in HRI with a focus on nonverbal interaction with simple mobile robots without anthropomorphic or zoomorphic features.

2.1 Ubiquitous Robotics

The shift toward ubiquitous robotics was triggered by the technologies that led to ubiquitous computing, coined by Mark Weiser, where computing devices are embedded in our environment and fully networked [50]. In addition to these connected touchscreens of various sizes, ubiquitous robotics include a non-flat physical form, mobility, actuation, and/or external sensing. Kim et al. laid out the architecture of ubiquitous robot system with three main components: Software robots, Embedded robots and Mobile robots [51]. Integrating cloud computing into stand-alone robots enhances their cognitive capabilities further [52]. Overall, ubiquitous robotics should not only provide a physical mobile platform to the existing ubiquitous computing, but also interact with both the user and the environment.

While the essential layers of ubiquitous robots have been proposed and interaction with software robots have been studied, less work has explored in situ and physical human interaction with multi-robot systems, an essential component of ubiquitous robotic interfaces. Other researchers have developed multi-robot platforms for different applications such as experimental multi-robot testbeds [53–55], smart cities [56] and surveillance [57, 58]. Additionally, some papers have explored functionality of these platforms for the applications of multi-agent coordination [59, 60], and networking and localization [61–63].

While the technical contributions of this previous work have helped build the foundation of the field, only a few researchers have explored physical interactions with multi-robot systems [64]. I build on these prior works by introducing ubiquitous robotic interfaces (URIs), interface that consists of a collection of robots that physically interact with users and the environment as explained in Chapter 3. In this thesis, I focus on use of URIs as an information display through motion and touch as well as its ideal naive user-friendly interface for control.

2.2 Swarm Robotics

Swarm robotics draws inspiration from biological swarms ranging from ants and bees to birds and fish as these organisms are able to complete complex tasks through collective means [65]. Similar to their biological counterparts, swarms robots are robust in failure and can produce complex collective behaviors with simple individual rules. Key aspects of swarm robots include swarm size, communication range, communication topology, swarm reconfigurability, and swarm composition [66].

Swarms of simple robots are particularly advantageous over a single complex robot in certain tasks. For instance, search and rescue [58, 67], autonomous vehicles [48, 68], and agriculture [69] applications require a widely distributed network of robots in order to cover a large area quickly and efficiently. Applications like collective transport [14], collective construction [39, 70] require collective efforts to manipulate large objects or build complex structures that would have been impossible or inefficient with a single robot.

One of the largest robot swarm is the Kilobots platform developed by Rubenstein et al. [71]. They can collectively form different shapes and manipulate objects through simple rules, albeit at a slow rate ($\sim$ 1cm/s) [71, 72]. Similar to the kilobots, other swarm robot platform are primarily task-oriented and are designed to complete tasks like foraging [12,69] and construction [39, 70] rather than to interact with people. Thus, I build on the past research on swarm robotics [73, 74] and apply it to ubiquitous settings with a focus on interaction with users.

2.3 Human-Robot Interaction

Human-Robot Interaction (HRI) has become increasingly important as more robots begin to appear in our environment. We see an increase in both the number of robots we interact with and types of robots, ranging from anthropomorphic robot guides in malls to non-anthropomorphic vacuum robots [75–79]. As the possibility of encountering robots in daily lives has increased, HRI researchers have investigated proximal social interaction with robots in which both verbal and nonverbal communications are used [15,23,80]. While speech recognition and production are key aspects of HRI that enable semantically rich communication, nonverbal behavior is also a crucial element to enable multimodal interaction exhibited in human communication.

To enable multimodal interaction, HRI researchers have explored other modes of interaction, mainly through varieties of nonverbal behavior such as motion [81,82], touch [46,47] pattern formation [8, 40], pointing [83, 84] and gaze [32, 84, 85]. These behaviors can help better convey subtle cues such as intent [32,82,83] or affect [47] that humans do both consciously and unconsciously. However, many of these are limited to anthropomorphic form and cues, thus may not apply as easily to abstract and multi-robot systems.

2.3.1 Human-Swarm Interaction

Researchers in human-swarm interaction (HSI) have started to develop methods for interacting with swarm robots. The majority of research in HSI focus on remote interaction, where the human teleoperates the robots from afar [12, 86–89]. Remote methods to effectively monitor and control the robots are important because one of the key motivations for swarm robots is their ability to be deployed to inaccessible or dangerous areas. The predominant interfaces for remote HSI include computer-based interfaces [86, 88] or joysticks [89, 90], where the states of the robots are visualized on a screen and the human operates via a combination of mouse and keyboard or a controller with joystick and buttons.

As robots increasingly interact with users with diverse technical backgrounds, we need to consider proximal interaction, where the humans and robots are in a shared environment without a separate interface. Alonso-Mora and colleagues investigated the use of swarm robots as physical displays [91] and recently extended their system to support interaction through sketching [92], hand-held tablet input [93] and mid-air gestures [94]. Others have also explored the use of robots as tangible interfaces for education purpose [95, 96]. Despite the growing amount of work in proximal interaction with swarm robots, more effort is needed to answer the question of how to design in situ interaction with a swarm of robots. This thesis approaches human-multirobot interaction by addressing research areas identified from one-on-one HRI literature (e.g., affective interaction, legibility, control, etc.) by leveraging features unique to swarm robots (e.g., collective behavior, pattern formation, etc.).

2.3.2 Display through Motion

One solution to enhance interaction with a group of non-humanoid robots is to leverage a feature that most robots either humanoid or non-humanoid have: mobility. Regardless of their form factor, most robots can move with respect to either the ground or other parts of their body. This degree of freedom could be used to move from point A to point B, visually convey information such as affect or intent, physically touch and manipulate objects, or any combination of them. This thesis also primarily rely on mobility of robots to display information visually and through touch.

Visual Display

The Heider-Simmel animation [97] demonstrated the enormous potential for visually conveying abstract information through motion alone. Heider and Simmel show a video in which several simple shapes (two triangles and a circle) interact nonverbally through motion. While the interpretation of the animation is not exactly the same across observers [97], many still perceive similar chains of events even without any verbal or anthropomorphic feature-based communications. By simply watching how and when these shapes move, people are able to associate actions and reasons behind these movements such as when the big triangle appear to chase the small triangle and the circle around its house.

Affect and Expressiveness

Inspired by Heider-Simmel animation, HRI researchers have explored ways to create expressive robots by modifying different motion parameters. Even across different robot platforms, researchers found that robot's speed and acceleration have a significant influence on the arousal axis in the circumplex model of affect [98–100]. On the other hand, smoothness, roundness, and perceived stability of movement were found to affect valence axis [98,99,101]. Other literature found specific relations between motion and emotion such as small and slow movement eliciting sadness and fear, while large, fast and jerky movement eliciting anger [102–104].

To generate even more expressive robot motion, researchers applied concepts from other fields such as Laban Effort System [16, 22] and animation [23]. Knight and Simmons leverage four Laban Effort features (i.e., timing, weight, space, and flow) to express different emotions with a mobile robot [17]. On the other hand, Takayama et al. apply a few key animation principles (i.e., anticipation and follow-through) to express thoughts and enhance the readability of a robot [23].

While there is an extensive literature for a single expressive robot, very few have studied how to enable expressive swarms of robots. Podevijn et al. found that increasing the number of robots alone provokes stronger responses in the psychophysiological state of humans [105]. On the other hand, Dietz et al. noticed synchronization led to higher positive affect albeit without statistical significance [106]. In Chapter 4, I study how to layer affect and expressiveness into swarm robots by understanding human's perception of different abstract swarm robot motion.

Intent

Dragan et al. introduced a novel, mathematical approach to generate legible motion that is

based on research in psychology [19]. They also differentiated the concepts of predictability and legibility based on whether the direction of the inference is "action-to-goal" or "goal-to-action" respectively. They proposed algorithm based on principle of rational action [107], Bayes' Rule, and optimization to generate the legible trajectory. Many researchers followed the basic principle with adaptations to improve the algorithm [108, 109], increase legibility for different viewpoints [110], and demonstrate that its performance depends on which manipulator is used [111]. Zhao et al. investigated the effects of different gripper orientations and found that gripper pointing towards the target in a straight path has the best performance [109]. Furthermore, Bodden et al. developed an algorithm based on "point position" heuristic that outperformed the algorithm from Dragan et al. [108].

Recently, researchers have extended the concept of legibility in the context of 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 [112]. In a Virtual Reality (VR) setting, twenty virtual robot 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 time, while the minimum-jerk trajectory was more accurate than the arc-trapezoidal trajectory. Capelli et al. ran similar study but with multiple multi-robot systems [113]. Here, they showed that trajectory and dispersion were found to impact the prediction accuracy, while harder stiffness increased prediction time contrary to previous study. In Chapter 5, I introduce and compare performances of different methods to enhance legibility of multi-robot systems such as pre-attentive processing feature-based and collective behavior-based methods.

Glanceable Display

Glanceable designs allow users to grasp information with only a quick look [114–116]. 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 in order 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 [114].

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 [117]. 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 [118]. For instance, the speed and efficiency of pre-attentive processing is contingent on the observer's current intentions and/or goals [119]. In Chapter 5, I primarily focus on controlling the salience of the robot motion to improve its glanceability.

Iconic Display

The most common way to display information visually with multi-robot systems is to form iconic shapes that resemble real-world objects [40,120]. Similar to pixels on a screen, robots serve as dynamic pixels that collectively represent an identifiable pattern. While this is a compelling and easy-to-understand way of displaying information, this thesis do not investigate iconic display in-depth as I primarily focus on display methods that allow seameless transition between task and interaction.

Haptic Display

Mobile robots can also move to physically display information to users through touch. Touch is an integral part of our daily lives from detecting objects to social interaction and affective communication. It allows us to not only discriminate shapes and textures but also is embedded in our social interactions (e.g., hand shake, high five, and huddle) [121, 122] and affective communications (e.g., hug, massaging, and holding hands) [37]. In order to build rapport and trust with people, robots should also leverage the haptic channel for information display and proximal interaction.

Relatively few works in HRI have explored touch from a robot to a human or mutual touch between robots and humans, and mostly with anthropomorphic robots. Touch from anthropomorphic robots have had mixed results. While some found it to have positive impact on human effort [46] and unfairness [123], others found people prefer touching the robots than being touched [124]. Verbal cues and perceived intent also were shown to effect human's response to a robot's touch [125].

While prior work has explored haptic display from a single anthropomorphic robot to human, none to our knowledge have looked into the haptic display from multiple, small non-anthropomorphic robots to users. As more robots enter our lives, it is important to not only study human-robot interaction in the context of dyads but also with multiple robots.

Thus in Chapter 6, I introduce a design space for haptic display with a swarm of robots and study human perception of various haptic patterns from the robots.

2.3.3 Command

Ubiquitous robots are bound to interact with a diverse set of individuals, ranging from children to expert roboticists. Thus, robots need to support interaction paradigm that is friendly to all users, or even better user-defined. Here, I review the current state of how multi-agent systems are controlled and how elicitation studies have been conducted for interaction with a few robots.

There is a good deal of research into effective computer interfaces for remotely controlling swarms and multi-agent systems [12, 126], but in this thesis I focus on proximal command that could function on an encountered basis (i.e., without dedicated interface hardware like a mouse and GUI). Most of the prior literature in this style seeks to demonstrate full end-to-end implementations in order to prove the viability of things like gesture-based command.

Gesture

A variety of possible sensor suites have been used for this purpose, including user-wearable attitude [31] or EEG sensors [127], centralized fixed vision systems [94], multi-touch screen [2, 128], and consensus-based distributed vision algorithms [129]. The most relevant work specifically investigates proximal interactions with wheeled multi-agent/swarm systems with well-defined task sets such as in [30]. In Chapter 7, I narrowly focus on single-operator multimodal interaction within an on-table environment. In contrast with tabletop swarm interface work like Reactile [130], which was explicitly a tangible interface made for only physical interactions, I let the users decide how they interact with the robots.

Techniques for User-defined Command

In an elicitation study, participants are prompted to choose and demonstrate their own preferred input method (and specific action) for a given task. Although their efficacy in uncovering natural interaction schemes has been validated in other areas like surface computing [131–133], elicitation studies in the context of swarm robot command remain rare.

There are some examples of elicitation studies for control of UAVs in the literature, but the increased low-level control complexity for safely operating high numbers of proximal

drones means that the number of robots interacted with in these studies is typically limited. A multimodal (gesture, sound, and touch) elicitation has been performed with a real single UAV [134], gesture-only elicitation for up to four real UAVs at a time [135], and for a swarm of 10 UAVs with voice and gesture multimodal input in simulation [136]. In contrast, working with on-table wheeled robots lets us deploy, without computer-rendered or VR simulation, relatively numerous groups of small robots (i.e., closer to future envisioned swarm systems) operating on a 2-dimensional workspace. This capability provides us with the unique opportunity to investigate the effects of different parameters on user input preference across a wide swath of example tasks, without worrying that our results will suffer from the documented "reality gap" that exists in simulated (as opposed to implemented in hardware) human-robot interaction studies [29].

2.4 Chapter Summary

We are approaching a phase of ubiquitous robots where we will encounter and interact with an increasing number of robots in public and personal spaces. Researchers are making great progresses in developing functional swarm robots that can navigate and sense in unstructured environments and collectively manipulate objects. However, as these robots become part of people's daily lives, it is also important to consider and design in situ interaction with multi-robot systems. As most existing literature primarily focus on one-on-one human-robot interaction, this thesis aims to broaden the scope to human-multirobot interaction. I first look at how to leverage mobility of a multi-robot system to display meaningful information to users. In Chapter 4-6, I investigate how multi-robot systems can convey information like affect or intent visually and through direct touch. In terms of input from the user to the robots, prior work either focused on building a functional interface that is not user-friendly or studied how people interact with a limited number of drones. Thus, in Chapter 7, I describe an elicitation study where I studied how users naturally interact with wheeled multi-robot systems.

Chapter 3

Ubiquitous Robotic Interfaces

Figure 3.1: I demonstrate the key elements of URIs: mobility, manipulation and display demonstrated in situ. From left to right: Zooids displaying calendar information on a vertical surface, Zooids manipulating a mobile phone, Zooids using abstract motion for a phone call notification, and user interacting with Zooids through direct touch. Adapted from [1].

Robots are becoming ubiquitous and will change how we go about our everyday lives. Prior work in ubiquitous robotics and swarm robotics provide the technical foundations of ubiquitous multi-robot systems by integrating cloud computing to enhance their cognitive capacities [51, 52] and studying effective coordination of many robots to complete complex tasks [8, 39]. While HRI researchers have explored the interaction aspect, the focus has been mostly limited to one-on-one interaction with humanoid robots that do not resemble the upcoming fleets of autonomous robots.

In this Chapter, I introduce Ubiquitous Robotic Interfaces (URIs), multi-robot interfaces that build on previous ubiquitous and swarm robots by incorporating interaction with users as well. In order to paint a picture of how I see URIs co-existing in our environment, I begin with an example scenario. I then introduce the key elements of URIs that enable ubiquitous

robotic interaction and provide details on Zooids, a multi-robot platform developed by me and Mathieu Le Goc [5]. Using this Zooids platform, I conducted user studies to prototype display and interaction aspects of URIs as described in the following Chapters. This chapter is based on my prior works [1, 5].

- ©2016 ACM. Reprinted with permission, from Le Goc, Mathieu, et al. "Zooids: Building Blocks for Swarm User Interfaces." Proceedings of the 29th Annual Symposium on User Interface Software and Technology. ACM, 2016. doi:10.1145/2984511.2984547 [5].

- ©2017 ACM. Reprinted with permission, from Kim, Lawrence H, and Follmer, Sean. "UbiSwarm: Ubiquitous Robotic Interfaces and Investigation of Abstract Motion as a Display." Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies. ACM, 2017. doi:10.1145/3130931 [1].

3.1 Example Scenario with URIs

To better understand how URIs can fit into users' everyday experiences, I provide an example scenario in which an imaginary person, Logan, interacts with Zooids on a normal day. Here, I assume Zooids are truly ubiquitous.

In the morning, Logan wakes up and prepares to get dressed. He snaps his finger to get the robots' attention and verbally asks for the weather. On the wall, Zooids forms an umbrella icon and today's temperature. Logan dresses accordingly and heads to the kitchen. Zooids collectively push a plate of his favorite donuts to the center of the kitchen table. At work, he prepares a cup of tea. As soon as the tea bag touches hot water, the robots slowly circle around the cup and disperse after a minute. At work, Zooids quietly flocks toward and provide gentle haptic notifications every 30 minutes to remind Logan to stretch and take a break. Back home, he prepares dinner. In the kitchen, Logan and the robots collaboratively cook without any recipe as the robots locate and bring the necessary ingredients at the appropriate time. Before going to sleep, he decides to read a book while lying down. Robots slowly move toward the bed and turn on the reading light. He makes fine adjustments by moving them by hand. After Logan falls asleep, the robots turn off the light and slowly disperse back to their charging stations.

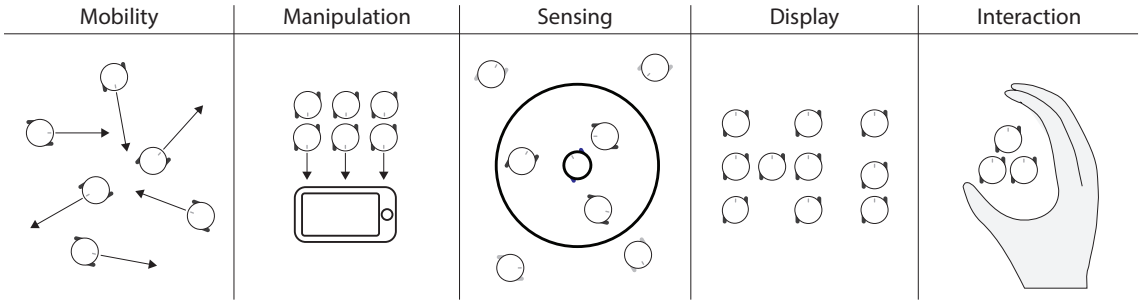


Figure 3.2: Key elements of Ubiquitous Robotic Interfaces include mobility, manipulation, sensing, display and interaction. While three former elements have been the main focus of prior works in swarm robotics, this thesis aims to address the latter two aspects. Adapted from [1].

3.2 Key Elements of Ubiquitous Robotic Interfaces

URIs are composed of many robots and have the following key elements as shown in Figure 3.2: Mobility, Manipulation, Sensing, Display and Interaction. For each element, I describe its role in ubiquitous robotic interaction.

3.2.1 Mobility

URIs need to be highly mobile. Their mobility separates URIs from traditional pixel-based interfaces. It enables URIs to initiate interaction with users and drastically increase the interaction space compared to static computing interfaces. Mobility also allows URIs to be embedded in a user's environment, moving from one place to another. I envision URIs moving from the wall to a table, to another room seamlessly. Their mobility is key not only for interaction and display, but also for carrying out other robotic tasks. The degree of mobility in terms of speed and interaction space is also important. Ideal URIs should be both fast (ideally approaching the visually perceptible refresh rate of the eye) and have infinite interaction space.

3.2.2 Manipulation

Manipulation is another element that is unique to URIs compared to traditional computing interfaces. It enables physical interaction with both the user and the environment. It can either provide direct haptic feedback to the user or manipulate objects for the user's convenience or to display contextual information (by moving or actuating a passive object).

Ideally, URIs should be able to freely manipulate all types of objects regardless of their weight or geometry.

3.2.3 Sensing

Mobility, manipulation and interaction require URIs to have sensing capability. URIs need to first sense their locations and surroundings before they can move, manipulate, or interact with any object. While ideal URIs should have these sensing abilities onboard, sensing can also come from other ubiquitous sensors such as cameras although I do not explore this extensively in this thesis.

3.2.4 Display

URIs can display information through spatial distribution and motion. These displays can be ambient, taking advantage of people's preattentive processing of motion [137, 138], or function as an interactive display with the user's full attention. As discussed by Le Goc et al., multi-robot interfaces can represent both "Things" and "Stuff" with movable elements instead of fixed pixels in screens [5]. The number of elements and identity of each element can also be varied. I envision URIs displaying information through both iconic form and abstract motion.

Iconic

Similar to screens, URIs can combine robots to form icons. Icons are an efficient way for communicating information universally without instruction when designed appropriately [139]. With a quick glance, users can understand an icon. Ideally, URIs should instantaneously form shapes of infinite resolutions to enable information display similar to current pixel-based interfaces.

Abstract Motion

Some information can be effectively communicated through motion. For instance, humans communicate their personal feelings such as emotion, intent, and affection both consciously and unconsciously through their body [140]. Researchers have shown that motion of simple shapes alone can elicit basic affective attributions [101, 103]. Abstract motion has a variety of benefits. It can be layered over pragmatic motions (e.g., moving to manipulate an object)

to provide more information, it can tap into our pre-attentive processing making it easy and fast to perceive, and finally it does not require a specific form factor or end effectors. In particular, I explore different methods to layer expressiveness and intent on top of swarm robot motion both visually and through touch.

3.2.5 Interaction

URIs can create user interfaces on demand, leveraging their mobility, when and where they are needed. URIs can also interact with users both directly and indirectly through surroundings/objects using their mobility and manipulation. Ideal URIs should have machine learning and activity tracking to allow smart and appropriate interaction. I investigate how to display comprehensible information through motion and touch, and how to design user-centered interface that can understand user's intent through multiple sensory channels.

3.3 Zooids: a Multi-robot Platform for Prototyping Display and Interaction

My colleagues and I introduced Zooids [5], a platform that consists of many cm-scale robots. While the Zooids platform is capable of the five key URI elements, its main purpose is to prototype display and interaction elements of URIs. In this section, I describe how Zooids is explicitly designed and implemented to support display and interaction with users.

3.3.1 Implementation

Despite the numerous multi-robot platforms from prior work [8,12,39,40], none are designed to support both real-time display and proximal interaction with users as those platforms cannot update robot formation at an interactive rate (ideally in the order of one second [41, 42]) or lack sensors that can detect user input. In this section, I describe the hardware and software design of the novel interactive swarm robotic platform, Zooids [5].

Hardware

The main purpose of Zooids is to enable and prototype real-time and multimodal human-multirobot interaction. Thus, our main goals were to design robots that 1) are small enough

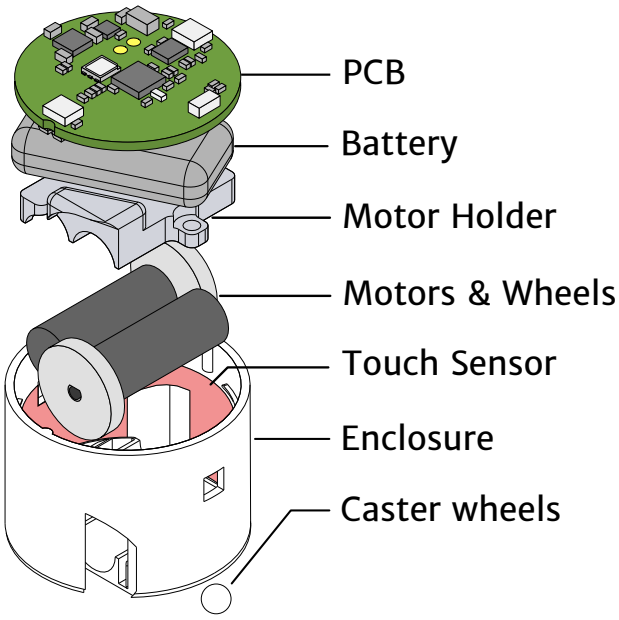


Figure 3.3: Exploded view of a Zooid. Adapted from [5].

to allow users to easily manipulate multiple robots simultaneously and blend into the environment, 2) fast enough both in terms of robot movement and communication to support real-time interaction with users (i.e., $\sim$ 1 second update rate), and 3) capable of detecting user input to afford bi-directional interaction.

Form Factor

Zooids are small custom-made robots as shown in Figure 3.3; their dimensions are 26 mm in diameter, 21 mm in height and they weight about 12 g. This form factor allows users to easily push or grab many robots at once. The robots can also more easily blend into our environment due to their small size.

Mobility

Each robot is powered by a 100 mAh LiPo battery and uses motor driven wheels. The motors are placed non-colinearly to minimize the size of the robots. Even though the motors do not rotate around the same axis, the robot has the same net force and moment as a robot with colinear motors. To drive the robot, a motor driver chip (Allegro A3901) and two micro motors (FA-GM6-3V-25) are used. With this combination, the robot has a maximum speed of approximately 74 cm/s. However, for controllability and smoothness of the motion, the robots move at a slower average speed of 44 cm/s for most of our applications.

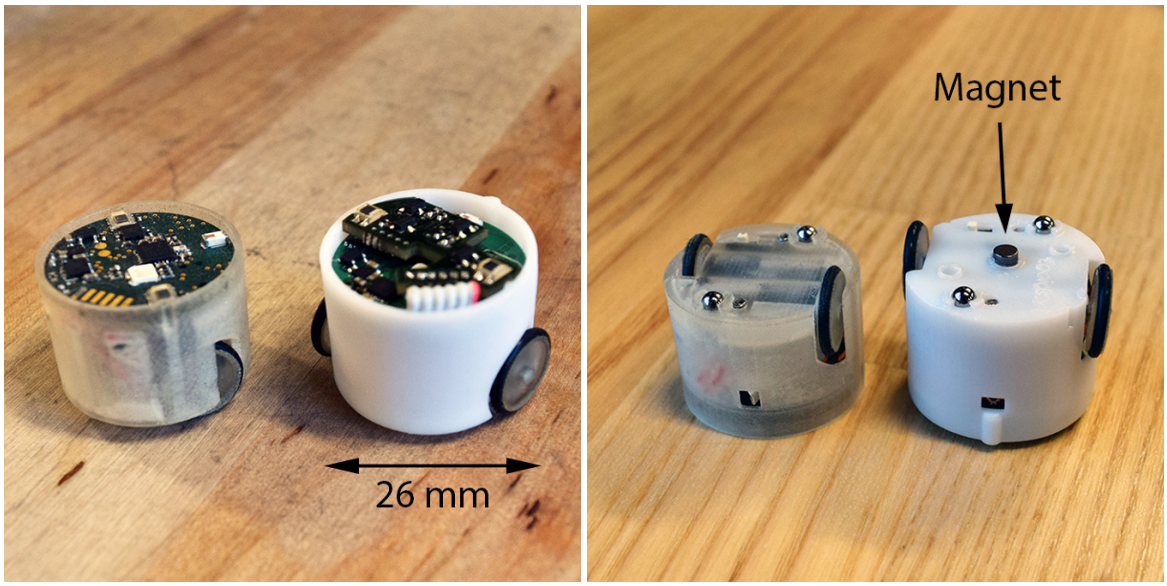


Figure 3.4: Two versions of Zooids are shown. The left robot is the original Zooid robot with lower gear ratio motors (26:1) while the right robot has higher gear ratio motors (136:1) in order to render smoother motion and stronger force output. Magnets are added to the bottom of the robots to improve mobility and force output when combined with ferromagnetic surfaces. Adapted from [1].

A modified version of the original Zooids robots is used for user studies in Chapters 4-7. As shown in Figure 3.4, the existing motors (26:1 gear ratio) in Zooids are replaced with higher gear ratio motors (136:1) in order to render more uniform and stable swarm movements, albeit at a slower speed (16cm/s vs 44cm/s).

In order to increase the mobility and force output of the robots, I added magnets to the bottom of the robots as shown in 3.4. These magnets extended the interaction space from just flat surfaces to horizontal and ferromagnetic vertical surfaces as shown in the leftmost picture in Figure 3.1. The combination of ferromagnetic surface and magnets can also increase the force output of the robots, thus widening the expressive range of haptic display with robots as described in 6.

Radio Communication

In order to make the robots interactive, robots need to both move quickly and communicate rapidly in order to update its desired goal position at an interactive rate. Thus, we chose to communicate between the robots and the central computer through radio. Each

robot communicates with the radio receiver using the NRF24L01+ chip. Using a teensy 3.1 microcontroller as the master and Arduino Pro mini as the slave, we tested the total communication times for different numbers of slaves per master and packet sizes. From the experiment, we found that the total time is linearly dependent of both packet size and number of slaves, and that we can maintain an update rate of 60 Hz with to 18 slaves per master for a packet size of 12 bytes. Zooids uses 10 slaves per master for a safety factor of about 2.

Projector-based Tracking System

Robots also need to localize swiftly in a scalable manner in order to afford real-time interaction. Thus, a projector-based tracking system similar to Lee [141] is used for robot position tracking. As opposed to camera based systems, our projector based tracking system does not add any latency from networking for the local feedback control on each robot, making position control more stable. Our system setup is demonstrated in Figure. 3.5. Using a high frame-rate (3000 Hz) projector (DLP LightCrafter) from Texas Instruments Inc., a sequence of gray-coded patterns are projected onto a flat surface. Then, the photodiodes on the robot independently decodes the gray code into a location within the projected area, and sends its position and orientation to the master computer. Due to the number of the patterns, the position refresh rate is approximately 73 Hz (1/(41 images per pattern $\times$ 1/3000)). Due to the diamond pixels of the projector, the horizontal and vertical resolutions slightly differ. In the current setup in which the projector is placed 1.25 m above the table producing a 1 m $\times$ 0.63 m projection area, the horizontal and vertical resolutions are 1.15 mm and 1.12 mm, respectively. Our current implementation utilizes a projection-based tracking [142, 143], so the applications only work in a relatively small area with a mounted projector. We envision other localization techniques such as HTC Vive's lighthouse tracking system could expand the workspace. For more details about implementation refer to [5].

User Input

In order to detect user input, we included modules for sensing touch and motion. To provide touch sensing capabilities, a flexible electrode is wrapped inside the 3D printed enclosure. An integrated capacitive touch sensing circuit is included (Atmel AT41QT1070) to detect user's touch. To detect user-induced motion such as double tap or shaking, we included an IMU sensor with 3D accelerometer and 3D gyroscope (STMicroelectronics LSM6DS3TR) inside each robot.

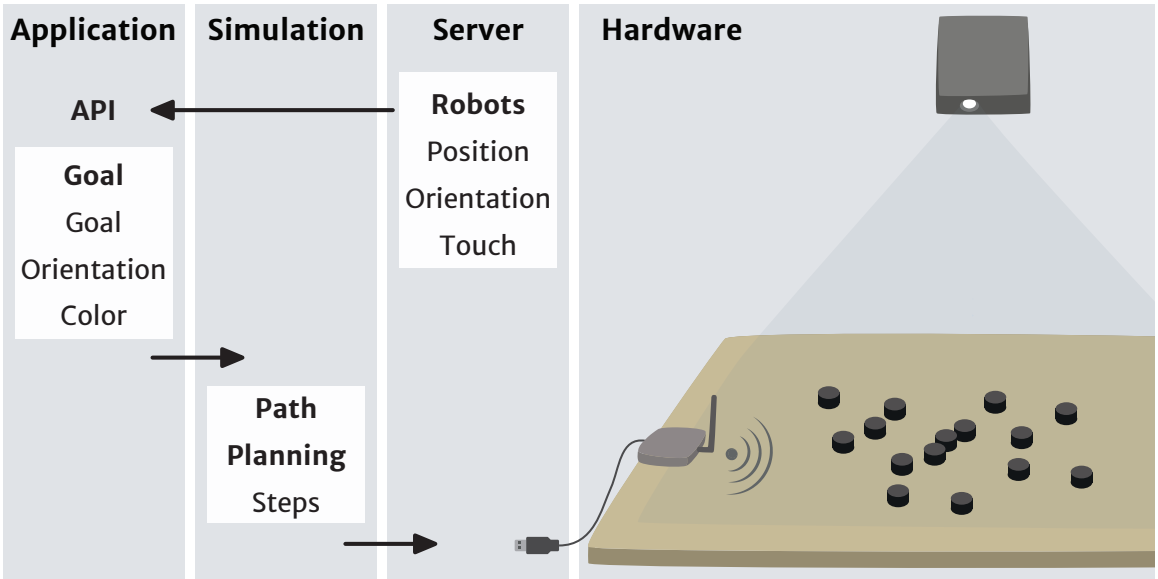


Figure 3.5: Software Architecture. Adapted from [5].

Power

Each robot is powered by a 100 mAh LiPo battery. Most of the power in the robots are consumed by (in order) the motors, radio module, micro-controller, and LED. When stationary, each robot consumes approximately 40 mA and 100 mA when moving. Thus, with a 100 mAh battery, robots are capable of continuously moving for one hour and can function even longer under normal sporadic usage.

Other Electronic Components

In addition to the parts described above, each robot consists of the following components. Embedded custom electronics, shown in the PCB layer of Figure 3.3, allows for robot control. A 48MHz ARM micro-controller (STMicroelectronics STM32F051C8) manages the overall logic computation and communicates wirelessly with the main master computer using a 2.4GHz radio chip (Nordic nRF24L01+). As part of the projector-based tracking system, two photodiodes are placed at the top of the robot. Placed between the photodiodes, a color LED is used for robot identification and visual feedback.

Software

As shown in Figure. 3.5, the communication structure consists of four main layers from highest to lowest level: Application, Simulation, Server, and Hardware.

At the application level, the desired positions of the robots are computed. These desired positions are transmitted to the simulation layer through a network socket. The application programmer can choose between two control strategies: *Proportional-Integral-Derivative* (PID) position control or *Hybrid Reciprocal Velocity Obstacles* (HRVO) combined with PID (these options are explained in the next paragraphs). Based on the chosen control strategy, the simulation layer computes the goal positions of the robots, either final positions for PID or intermediate points for HRVO, and sends them to the server. Finally, the server layer dispatches commands to the individual Zooids, while at the same time monitoring their status and positions.

Before any movement, each robot first needs to be assigned its final position. The final positions may be specific for each robot or they can be dynamically assigned to move in a more efficient manner. The Hungarian algorithm [144], a well-known optimization method for one-to-one task-agent assignment problems, can be used to assign the goal positions to robots in an optimal fashion. The cost function to be optimized is the summation of the squared distances from the initial to the final positions.

After the goal assignment step, robots need to move toward their goals, while minimizing possible collisions with each other robot. We chose to use the HRVO control strategy [145, 146] due to its fast real-time path planning capabilities. With HRVO, a robot moves at the user-defined preferred speed unless it detects possible collisions. In that case, it uses the notion of velocity obstacle, i.e., the set of all robot velocities that will result in a collision with another robot. While HRVO does not guarantee collision-free, oscillation-free control, it reduces the number of collisions dramatically compared to other velocity obstacle strategies while providing real-time updates, essential to natural and smooth user interactions. To implement HRVO, we used a slightly modified version of the HRVO library created by Snape et al. [145, 146].

With the HRVO control strategy, we can derive the incremental goal positions along a path for each robot. These positions are sequentially sent to each robot which independently controls its motion through a PID controller based on the state machine shown in Figure. 3.6. Given a final goal, the robot initially turns itself in the right direction and, once aligned, accelerates to its user-defined preferred speed. When it reaches the speed, it maintains it with a PID control on the orientation to ensure its direction towards the final goal. When a new incremental goal is given, it will still move at same speed but the PID control on orientation will direct the robot towards the new intermediate goal. When the robot arrives

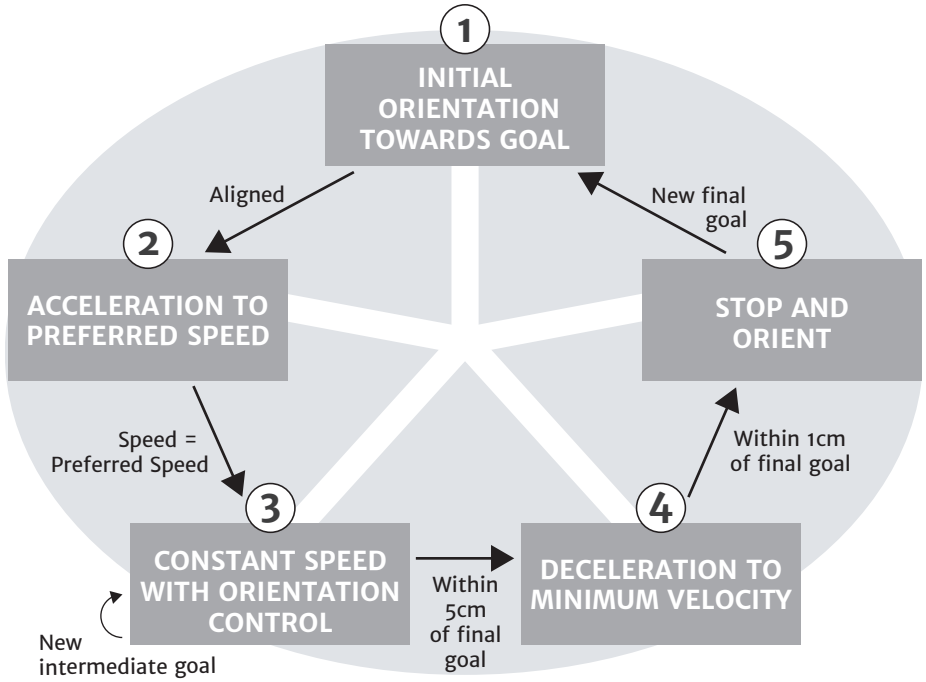


Figure 3.6: Structure of local PID position control. Adapted from [5].

within 5 cm of the final goal, it slows down to its minimum velocity and once within 1 cm of the final goal, it stops and orients itself as commanded by the application programmer. To enable smooth transitions between the incremental goal positions, robots are given their next position at 60 Hz.

The applications and robot movements for user studies in Chapters 4-7 were programmed in C++ in Visual Studio.

Open-sourced platform

All necessary material and documentation for implementing Zooids can be found at https://github.com/shapelab/SwarmUI. Several external researchers have already downloaded, built, and even published at top tier venues [147–149] using our open-source software and hardware. The Zooids platform was also used by students at the inaugural SIGCHI Summer School for Computational Fabrication and Smart Matter to develop their own applications [150].

3.4 Chapter Summary

In this chapter, I introduced my vision of Ubiquitous Robotic Interfaces. I highlighted the key components of URIs and how I anticipate them being embedded into our daily lives. I also described a prototype of URIs, Zooids [5], that was co-created with Le Goc et al. I used Zooids to explore various aspects within real-time display and proximal interaction with users. I explained the design rationals and implementation details of Zooids such as form factor, communication structure, and software. Driven by my vision of URIs, I use different versions of Zooids to study display and interaction elements of URIs in Chapters 4-7. Specifically, I first explore the use of motion and touch to display information to users in the next three chapters. Then in Chapter 7, I complete the loop and study how users naturally command multi-robot systems.

Chapter 4

Generating Expressive Abstract Multi-robot Motion

In a world with fleets of ubiquitous robots roaming around our environment, robots will have to learn how to interact with people communicating information even during a non-interaction task and understanding commands from people. In the next three chapters, I begin by investigating how to best convey information to users with multi-robot systems. In particular, I utilize a common key feature of most, if not all, multi-robot systems: mobility. When the robots are out of user's reach, they can leverage their visible motion to convey information. When they are close to the user, robots can directly interact with the user through touch. In this chapter, I study how to generate expressive swarm robot motion.

-

4.1 Introduction

A central question in the use of URIs is how to display meaningful information to users through multi-robotic motion. Much literature on human robot interaction has explored the use of limbs and facial expressions to convey intent, affect, and information [140], but this limits the ability of robots to seamlessly blend into our environments and does not

scale well when interacting with many small robots. While multi-robot systems can form iconic patterns to communicate with users [8,151], this method cannot be used during other tasks like navigation or object manipulation. On the other hand, abstract motion can be layered on top of robot motion even during other tasks. Thus in this chapter, I first explore different types of abstract swarm robot motion and study how these motions can implicitly convey different types of information. For instance, can the robots communicate how urgent their task is or how excited or sad they are about their interaction? Can we also change how users feel about the robots or the interaction experience such as animacy, likeability, or hedonic and pragmatic quality? Understanding of how different abstract swarm robot motion affects these perceptions will provide guidelines on how to best adapt multi-robot systems for particular scenarios.

Based on prior literature in natural biological swarms and swarm robotics [44,152–154], I identified three multi-robot motion parameters and used them as abstract motion parameters for the perception study: bio-inspired collective behavior (rendezvous, dispersion, random, torus, and flock), speed (fast or slow), and smoothness (smooth, synchronously jittery, or asynchronously jittery) as shown in Figure 4.2. User perception and experience are self-evaluated through measurement tools such as Self-Assessment Manikin (SAM), AttrakDiff2, and HRI metrics [155–157]. From a crowdsourced between-subjects video user study, I find that these different collective behaviors significantly impact user's perception. Finally, we apply the findings to derive example applications in everyday scenarios with the Zooids robot platform described in Chapter 3 and in paper [5].

In summary, the contributions of this chapter are:

- A crowdsourced between-subject study to investigate perception of abstract multi-robot motion,
- Design guidelines for expressive multi-robot movements, and
- Preliminary exploration and demonstration of abstract multi-robot motion in the context of URIs.

4.2 Method

To study the effects of different abstract multi-robot motions, I performed a crowdsourced between-subjects experiment using the modified version of Zooids [5] as shown in Chapter

??. The between-subjects design enabled participants to watch and rate only one single video instead of many videos that a within-subjects study design would require. This reduces both the user fatigue and the carry-over, context, and sensitization effects that can lead to improvement in the overall quality of user responses [158].

Researchers have looked at whether the video-based HRI studies yield similar results to live in-lab HRI studies. Woods et al. showed that there was a high agreement between the two studies when investigating how a robot should approach users [159]. On the other hand, Xu et al. demonstrated that physically present robots yielded greater emotional and social user feedback compared to the robots presented through video or text [160], while Bainbridge et al. showed that people trusted and provided more personal space to physically present robots than video-displayed robots [161]. However, they also showed that both physical and video-displayed robots were effective in conveying contextual information and in eliciting feedback on general attitudes [160], and were greeted and cooperated with equally [161]. From these studies, it seems that the different results occur when there is a significant interaction between the user and robot. Thus, since our study is focused solely on the perception of abstract motion and does not involve a significant interaction with the participants, I conclude that video-based trials will yield results comparable to those from an in-person study.

4.2.1 Video Preparation

The overall filming apparatus for the study is shown in Figure 4.1. The videos were filmed with ten robots as described in Chapter 3 on top of a table with a white background. The camera angle was oriented such that it matched the viewpoint of a person sitting down. Due to the limitations of the projector tracking system, the videos were filmed in a dimly lit room with a high-speed DLP projector shooting down from above. The video duration varied from 3 to 24 seconds depending on the type of abstract motion. Robot motions were filmed until the robots completed their motion for the applicable behaviors (rendezvous, dispersion, and flock), and for 10 seconds for torus and random behaviors.

4.2.2 Abstract Multi-robot Motion Parameters

With more than a single robot, it is possible to create a wider range of abstract motion than just by changing the speed and the smoothness. From the literature on swarm motion, I

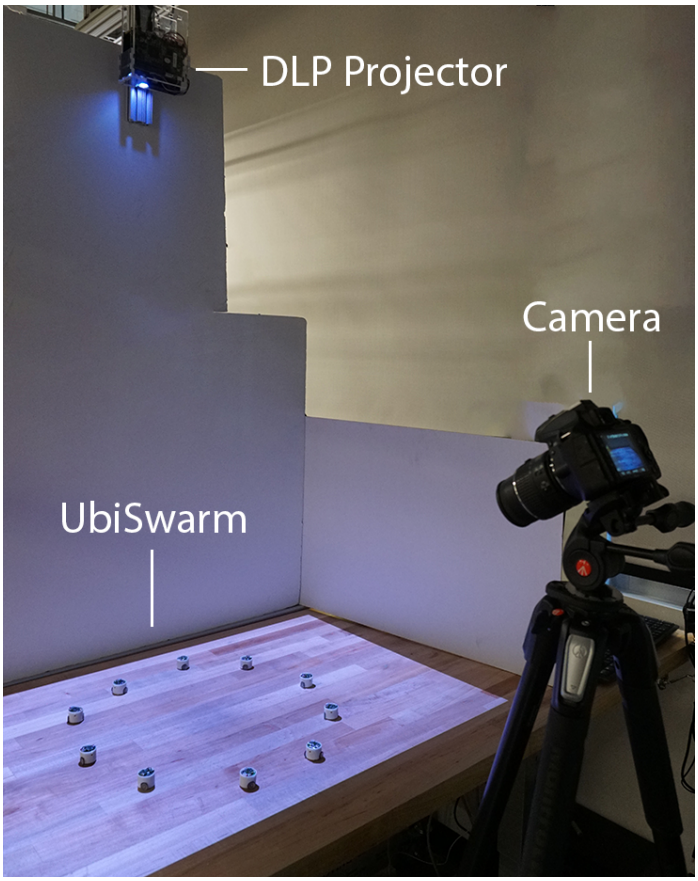


Figure 4.1: Filming apparatus for the study: a DLP projector producing the gray code pattern for localization, a modified version of Zooids as URI testbed, and a camera for capturing video. Adapted from [1].

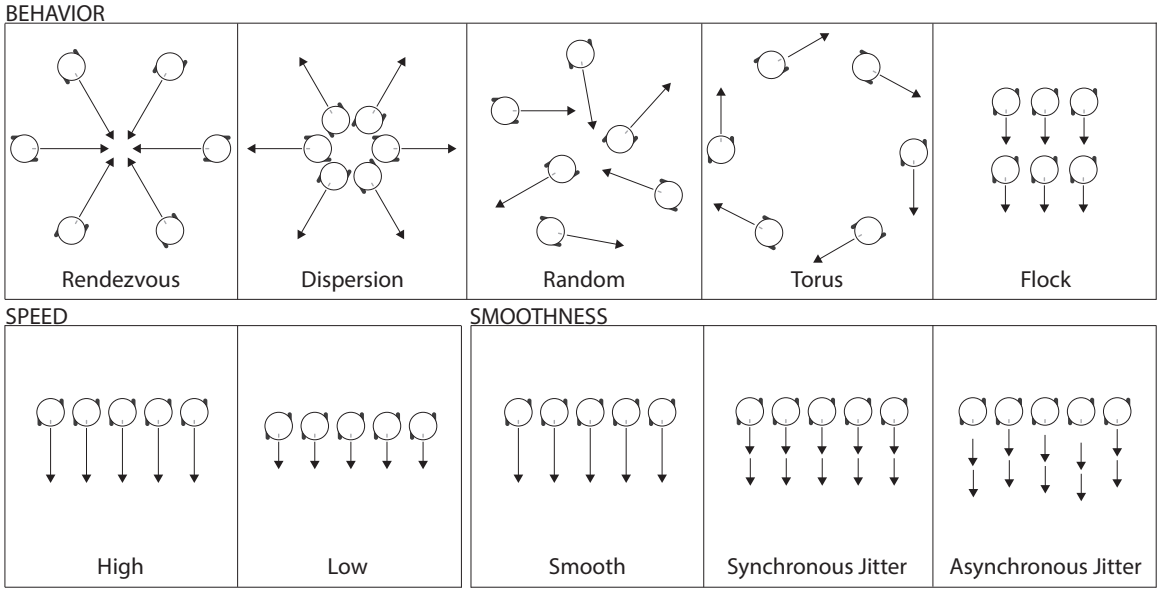


Figure 4.2: Swarm Motion Parameters. Adapted from [1].

identified bio-inspired collective behavior, an additional motion parameter unique to multi-robot systems. Thus, the following motion parameters were varied for each video: collective behavior, speed, and smoothness.

Bio-inspired collective behavior:

To leverage the additional degrees of freedom (DOF) that groups of robots have in contrast to a single robot, I looked at natural swarms for inspiration. Natural swarms exhibit complex behaviors by following a simple distributed rule. I have identified five different swarm behaviors from existing literature as depicted in Figure 4.2: rendezvous [44], dispersion [44], torus [152–154], random (swarm) [153, 154], and flock [44, 152, 153].

Rendezvous behavior is when all the robots move toward the center of the swarm, while dispersion is the opposite of rendezvous where the robots move away from the center. For torus, random, and flock behavior, the robots move in a circle, random direction, and same direction, respectively.

Speed:

Speed has consistently been found to be the most significant variable for single robot motion perception. In our study, two values of speed are chosen such that they are the most

distinguishable with our robots: high and low speeds corresponding to average values of 16 cm/s and 9 cm/s respectively.

Smoothness:

Besides speed, researchers have shown smoothness to be the second most significant parameter [98]. Previous studies involving a single robot could only change the intensity of smoothness [98]. However, in our setup with a swarm of robots, it is also possible to change the timing of the smoothness as well. For this study, I used three versions of smoothness: smooth, synchronous jitter, and asynchronous jitter. For smooth movement, a constant speed is commanded from point A to point B. For synchronous and asynchronous jitter movements, zero speed is commanded for 150 ms every 400 ms either synchronously for each robot or asynchronously (where each robot is seeded with a random starting time in the 400ms cycle).

4.2.3 Dependent Variables

To understand humans' general perception of the abstract multi-robot motion, I collected Likert scale ratings on four relevant categories: urgency, emotion, human robot interaction, and user experience. In addition, I collected users' perceived speed and smoothness rating scales to confirm their match with the commanded values. At the end, users could optionally leave comments about the study.

Urgency:

As I envision robots to be ubiquitous in the future, I expect them to be often used for notifications such as event reminder or phone calls. I adopted method used in [162] to measure urgency. Through a nine-point semantic differential, I measured the perceived urgency of the abstract motion and asked whether they will dismiss or attend to them.

Emotion:

Emotion is integral in all experiences. Emotion influences physiological, cognitive and behavioral states of users. To mediate and control emotion elicited by the robots, it is crucial to study the effect of abstract motion on users' affect. In order to measure perceived emotion, I used a seven-point scale of self-assessment mannequin, SAM [155]. SAM is a visual scale

of parameters in the PAD emotional state model [163]: valence, arousal, and dominance. Due to its reliance on pictures instead of words, it is widely used in both user experience and HRI research across different countries.

Measures for Human Robot Interaction:

HRI researchers use questionnaires specific to measuring perception of robots. Bartnet et al. designed a set of standardized measurement tools for human robot interaction (HRI) on five key concepts: Anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety. [157]. I excluded anthropomorphism since the shape of robots do not change and perceived intelligence due to lack of significant interaction between the robots and subjects. Thus, I used a nine-point semantic differential scales for rest of the three concepts that were most relevant to our perception study: animacy, likeability, and perceived safety.

User Experience:

Before releasing interactive products, it is important to study their perceived qualities or their user experience. Hassenzahl identified three major qualities that contribute to the user experience: perceived pragmatic quality, hedonic quality and attractiveness of interactive products [156]. To measure these qualities, Hassenzahl created the AttrakDiff2 questionnaire. In this study, I used a nine-point scale, ten-item abridged version of it to measure user experience on abstract multi-robot motion. AttrakDiff2 has been widely used in user experience research to assess the overall experience. It uses semantic differentials on a set of words such as tacky/stylish and unpredictable/predictable. For a complete list, refer to [156].

4.2.4 Participants

I recruited 1067 participants through Amazon Mechanical Turk for a between-subjects study. For each condition, approximately 25 participants viewed and rated the corresponding video. For quality control, only participants that satisfy the following requirements were included in the analysis:

1. Location is US
2. HIT approval rate is greater than 90

3. Number of HITs approved is greater than 50
4. He/she has not previously participated in any of our pilot studies
5. He/she is not experiencing any symptoms that may affect performance in the experiment
6. He/she has participated only once.

Requirements 1-4 were enforced through Amazon Mechanical Turk. For requirement 5, we asked whether the participants are experiencing any of the following symptoms at the end of the survey: neurological disorders, impaired vision, headache, fatigue, and any other conditions that may affect their performance. The 67 participants that checked "yes" were removed. For requirement 6, 38 participants with duplicate IP addresses were also removed leaving a total of 962 participants for the analysis.

Due to this filtering process, numbers of participants across different conditions for the analysis became uneven although equal numbers were recruited. Authors did not recruit more participants to balance the number because the numbers of participants were still relatively even ($< 8\%$ difference).

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. 21%, 63%, and 16% of participants reported education levels of middle/high school, college and advanced degrees respectively. After completing the experiment with average completion time of 3.5 minutes, each participant received $0.60 US dollars corresponding to a hourly salary of $10.30 US dollars.

4.2.5 Procedure

Before showing the abstract multi-robot motion video, I informed the participants that they would be seeing a group of robots moving in a particular manner and would be asked to rate their perception of the robot motion. Then with no training session, participants viewed their assigned video. The next button was shown only after the video finished but the participants were allowed to re-watch if desired, before moving on to the questionnaire. After watching the video, participants answered a set of questionnaires including SAM, AttrakDiff2, urgency, and HRI questionnaires. At the end, they filled out demographic information and received compensation.

4.2.6 Analysis

To examine the effects of the three independent variables including interaction effects, an n-way ANOVA was performed for each of the dependent variables. If any single independent variable or combination had statistically significant effects ($p < 0.05$), Bonferroni-corrected post-hoc tests were performed to determine which pairs of means are significantly different.

Table 4.1: Summary of the study results for emotion and user experience. Adapted from [1].

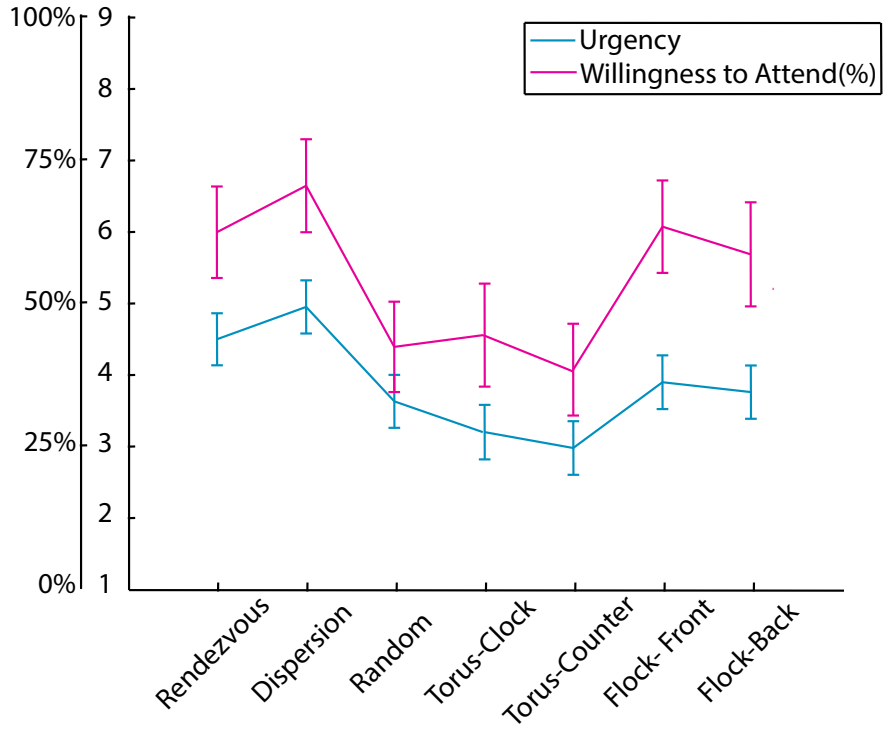
			Emotion			*User Experience*		
		N	Valence	Arousal	Dominance	Hedonic	Pragmatic	Attract
Behavior	Rendezvous	141	4.55(.17)	3.50(.24)	3.95(.23)	5.26(.27)	6.25(.22)	5.78(.25)
	Dispersion	139	4.34(.17)	3.43(.24)	3.62(.23)	5.39(.27)	6.63(.22)	5.79(.25)
	Random	132	4.29(.17)	3.12(.24)	3.52(.23)	4.73(.27)	5.37(.22)	5.38(.24)
	Torus-CW	131	4.38(.18)	2.90(.25)	3.43(.24)	4.79(.28)	6.83(.23)	5.88(.26)
	Torus-CCW	137	4.30(.18)	2.93(.25)	3.05(.23)	4.42(.27)	6.91(.23)	5.56(.25)
	Flock-Front	138	4.37(.17)	2.99(.24)	3.63(.23)	4.79(.27)	6.40(.22)	5.60(.25)
	Flock-Back	133	4.41(.18)	2.91(.25)	3.44(.23)	4.69(.28)	6.58(.23)	5.66(.25)
Speed	Slow	484	4.31(.09)	2.93(.13)	3.44(.12)	4.74(.14)	6.43(.12)	5.61(.13)
	Fast	478	4.44(.09)	3.30(.13)	3.60(.12)	4.99(.15)	6.42(.12)	5.72(.13)
Smooth -ness	Smooth	318	4.46(.12)	3.05(.16)	3.62(.15)	5.11(.18)	6.44(.15)	5.86(.16)
	Async Jitter	324	4.34(.11)	3.24(.16)	3.47(.15)	4.76(.18)	6.34(.15)	5.57(.16)
	Sync Jitter	320	4.33(.11)	3.04(.16)	3.47(.15)	4.73(.18)	6.49(.15)	5.56(.16)

4.3 Results

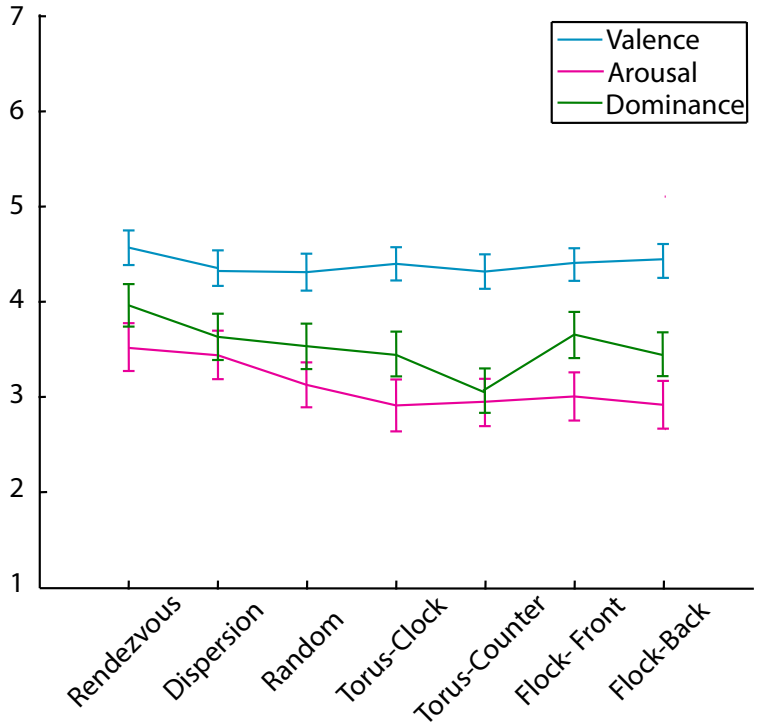
The overall results of the study are shown in Table 4.1- 4.2. They report the means of all dependent variables for each shape parameter along with their 95% confidence intervals and sample size. The interaction factors are not reported because N-way ANOVA found almost no interaction effects except for arousal and willingness to attend.

4.3.1 Urgency

Two questions were asked for urgency: semantic differential scale of not urgent-very urgent and a dichotomous question on whether to dismiss or attend to the robots. Results from both correlated very well as shown in Figure 6.12. For both questions, behavior was the most



(a) Urgency and williness to attend



(b) Emotion through SAM questionnaire

Figure 4.3: Mean ratings and 95% confidence interval for different bio-inspired behaviors. Adapted from [1].

Table 4.2: Summary of the study results for HRI metrics and perceived urgency. Adapted from [1].

			HRI Metrics			*Urgency*	
		N	Animacy	Likeability	Perceived Safety	Urgency	Attend?
Behavior	Rendezvous	141	5.24(.24)	6.01(.23)	6.07(.22)	4.46(.36)	62%(8%)
	Dispersion	139	5.13(.25)	5.96(.23)	6.28(.22)	4.91(.37)	70%(8%)
	Random	132	4.82(.24)	5.87(.23)	6.18(.22)	3.60(.36)	38%(8%)
	Torus-CW	131	4.69(.25)	6.02(.24)	6.18(.23)	3.17(.38)	44%(9%)
	Torus-CCW	137	4.36(.25)	5.83(.23)	6.07(.22)	2.95(.37)	38%(8%)
	Flock-Front	138	4.78(.25)	5.85(.23)	6.18(.22)	3.87(.37)	63%(8%)
	Flock-Back	133	4.73(.25)	5.78(.24)	6.20(.22)	3.73(.37)	58%(8%)
Speed	Slow	484	4.67(.13)	5.84(.12)	6.20(.12)	3.56(.20)	51%(4%)
	Fast	478	4.97(.13)	5.96(.13)	6.13(.12)	4.07(.20)	57%(4%)
Smoothness	Smooth	318	4.93(.16)	6.09(.15)	6.33(.14)	3.88(.24)	51%(6%)
	Async Jitter	324	4.91(.16)	5.90(.15)	6.09(.14)	3.94(.24)	57%(5%)
	Sync Jitter	320	4.62(.16)	5.73(.15)	6.08(.14)	3.63(.24)	54%(5%)

statistically significant factor ($p < 0.001$). For both, dispersion had the highest average of 4.91 (.37) and 70% (8%) respectively while counterclockwise torus had the lowest average of 2.95 (.37) and 38% (8%) respectively. Speed had statistical significance only for the urgency scale ($p < 0.001$). As expected, fast motion was perceived as more urgent than slow movements.

4.3.2 Emotion

SAM includes three variables for emotion: valence, arousal and dominance. For valence, only speed was found have statistical significance ($p < 0.05$). The Bonferroni post hoc test showed that fast motion has statistically higher valence rating than slow movement.

For arousal, both behavior and speed were statistically significant with $p < 0.001$. Similar to previous research on single robot perception, faster movement had a higher arousal rating. As for behavior, rendezvous had the highest mean rating of 3.50 (.24) while clockwise torus had the lowest mean rating of 2.90(.25) in a seven-point scale. Behavior was the only statistically significant parameter for dominance ($p < 0.001$). Again, rendezvous had highest mean rating of 3.95 (.23) while counterclockwise torus had the lowest mean of 3.05 (.23). The overall effect of collective behaviors on emotion variables is plotted in Figure 6.10.

4.3.3 HRI metrics

Out of the five key categories that Bartnet et al. developed as measurement tools for HRI, we chose to look specifically at three most relevant: animacy, likeability, and perceived safety.

All three shape parameters are found to be statistically significant for animacy in the following order: behavior, speed and smoothness ($p < .001$, $p < .01$, and $p < .05$ respectively). Rendezvous is perceived to be the most animate behavior (mean = 5.24 (.24)) while counterclockwise torus is rated the lowest animate (mean = 4.36 (.25)) as shown in Figure 4.4a. Fast movements had higher average animacy rating than slow ones. For smoothness, interestingly both smooth and asynchronously jittery movements had higher average ratings than synchronously jittery movement.

Only smoothness had a statistically significant impact on likeability. Smooth motion had a higher average likeability rating than synchronously jittery motion.

No statistically significant factor was found for perceived safety although speed was very close ($p = 0.057$).

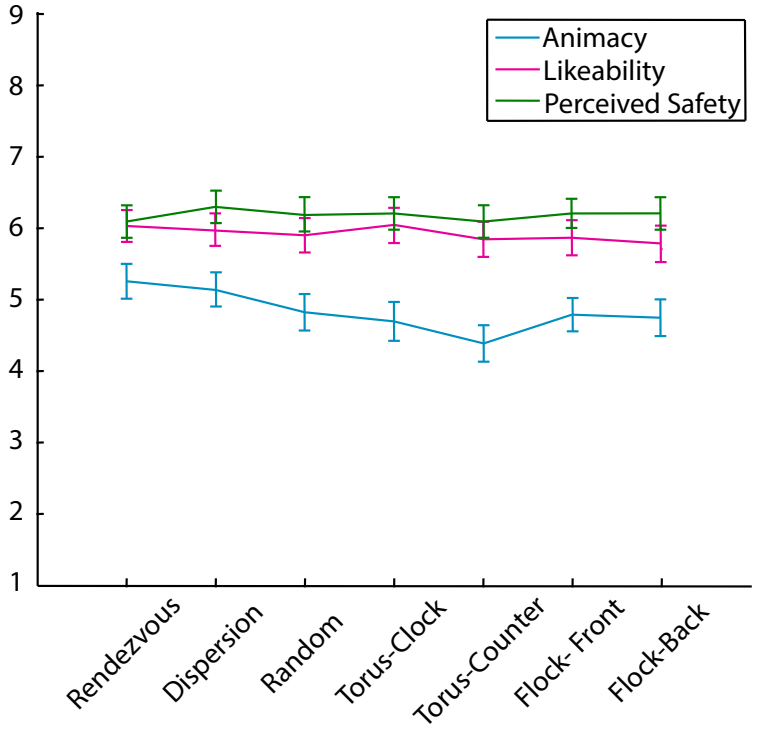
4.3.4 User Experience

User experience was evaluated with the abridged version of the AttrakDiff2 questionnaire with three key qualities: hedonic quality, pragmatic quality, and attractivity.

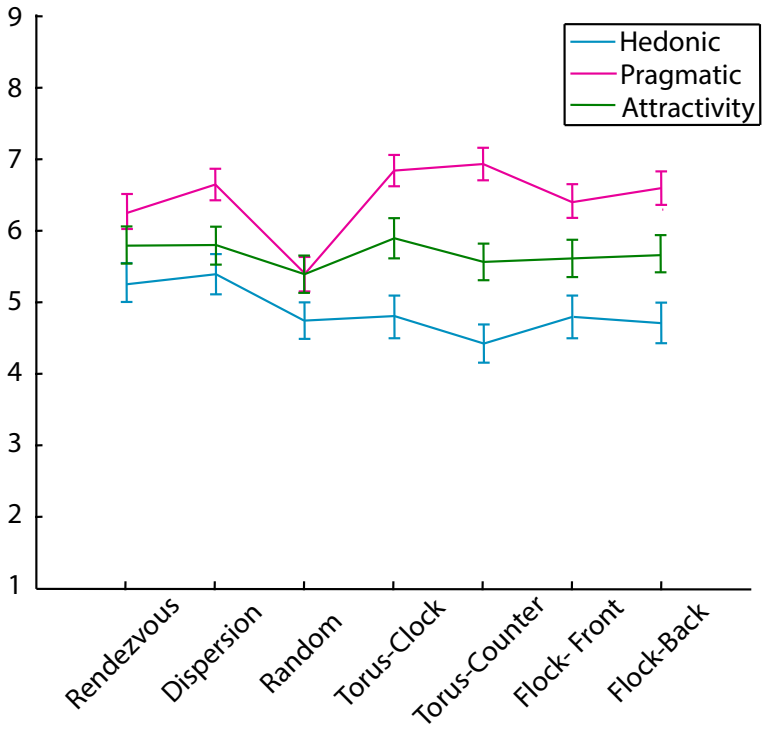
All three shape parameters had statistical significance on hedonic quality ($p < 0.05$ for speed, $p < 0.01$ for behavior and smoothness). Fast and smooth movements had higher hedonic rating than slow and jittery movements respectively. For behavior, dispersion had the highest average rating of 5.39 (.27) while counterclockwise torus had the lowest average of 4.42 (.27) as in Figure 4.4b.

For pragmatic quality, only behavior had statistical significance ($p < .001$). As expected, random behavior had the lowest average rating of 5.37 (.22) while both clockwise and counterclockwise torus behaviors had the two highest average ratings of 6.83 (.23) and 6.91 (.23) respectively.

For attractivity, smoothness was the only statistically significant factor ($p < .05$). The results show that smooth movements are rated more attractive than both synchronously and asynchronously jittery motions.



(a) Measures for Human Robot Interaction



(b) User experience through AttrakDiff2 questionnaire

Figure 4.4: Mean ratings and 95% confidence interval for different bio-inspired behaviors. Adapted from [1].

4.3.5 Perceived Speed & Smoothness

For both perceived speed and smoothness, results were as expected: speed and smoothness had the highest statistical significance respectively (both $p < 0.001$).

4.3.6 Qualitative Feedback

I also gave participants freedom to leave any additional comments. Of the 371 participants who left comments, 339 of them were not related to robots and their motion or were too general. The remaining 32 wrote their impressions of the robots and their motion. Eight of them were descriptive of their motion (e.g., "Some of the robots moved jerkily while a few moved smoothly"). Ten participants wrote robots were either "cool" or "cute" (e.g., "Robots - they are the future" and "I thought they were cute"). Two thought the robots were creepy: "robots are a little creepy but they might be able to help people with disabilities". One did not view them as robots: "they didn't look like robots, at least to me". Finally, three made metaphors (e.g., "They looked like hockey players skating!" and "The robots looked like mini trashcans :)").

4.3.7 Effect of Number of Views

As participants were allowed to re-watch the video as many times as desired, I also looked at whether the number of times participants watched the video affected their perception. Since I could not record the number of views, we used the time spent on the video page and divided it by the length of the videos. With this calculated data, an ordinal logistic regression was performed and the coefficients of the covariates along with its p-value were calculated. Out of the 10 dependent variables, only two had significant p-values (p ¡ 0.05): Animacy and Urgency. Thus, it seems that the number of times participants watched the video did not affect their perception in general.

4.4 Discussion & Design Implications

I envision that URIs can perform various tasks such as manipulation, display and interaction. The study results presented in this chapter provide guidance on how to design swarm movements for both standalone and embedded displays.

4.4.1 Collective Behavior

Collective behavior had statistically significant results in the following domains: arousal, dominance, hedonic and pragmatic qualities, animacy, urgency and willingness to attend. For these domains, some of the behaviors were perceived similarly to another. Specifically, rendezvous and dispersion behavior were rated closely, whereas for both torus and flock behaviors, direction of the motion did not influence the ratings significantly. The following discussions will be based on relative ratings among the different behaviors.

Rendezvous and dispersion behaviors were both perceived to be highly arousing, dominant, hedonic, animate, and urgent. They have the highest ratings for all categories except pragmatic quality. This suggests that these behaviors are appropriate for arousing, urgent, and hedonic notifications. The difference between the two behaviors is the direction of the motion, and thus the center of attention. For rendezvous, the focus is on the center whereas there is no focal point for dispersion behavior. Thus, rendezvous should be used in urgent and hedonic situations in which there is either a particular point or object of interest. For example, when receiving an important call, the robots rendezvous toward the phone. On the other hand, dispersion should be used in urgent and hedonic situations in which there is no particular point or object of interest. For example, when it is time to leave for an important meeting, robots disperse to alert the user.

Relative to other behaviors, users rated both directions of torus behavior to be non-arousing, non-dominant, low hedonic, inanimate, and non-urgent, but the most pragmatic behavior. This suggests torus behavior should be used in pragmatic but non-urgent scenarios such as a timer or progress bar to inform user of a status in low intensity applications. Although further research is needed, one interesting trend is that the two directions were perceived to be slightly different in some categories though not with statistical significance. Counterclockwise direction was perceived to be less dominant, less hedonic, less animate, and less urgent than clockwise torus. One potential cause may be the predominant exposure to clockwise motion in clocks in Western culture but this will need further investigation. In contrast, both forward and backward flock behaviors were perceived almost the same.

Compared to other behaviors, flock was rated average for dominance, hedonic and pragmatic quality, animacy, and urgency but high in willingness to attend and low in arousal. This suggests flock behavior to be used for average, everyday circumstances that are not urgent but nevertheless need attention. Examples include birthday reminders consisting of moving toward a picture frame of the person, and helping people cook by pointing toward

appropriate seasonings/ingredients. Since both flock directions are perceived the same, the direction will depend entirely on the direction between the robots and the point or object of interest.

Random behavior was also perceived to be average but in different categories. Compared to others, random was rated average for arousal, dominance, animacy, and urgency but low in hedonic, pragmatic, and willingness to attend. Thus, random behavior is appropriate for un-pragmatic, arousing scenarios that do not need any attention, such as a music physicalizer or exercise motivator. People will not need to attend to them but will be moderately aroused.

4.4.2 Speed

The results for speed matched well with that of prior literature. Faster speed was perceived to be more pleasant, exciting, hedonic, animate, and urgent. Thus, motion speed should be fast for urgent, arousing and pleasant events such as important calls or reminders whereas it should be slow for low intensity, non-urgent applications like ambient displays, timers, and white motion.

4.4.3 Smoothness

The results for smoothness also aligned reasonably well with that of existing work where smoothness has been shown to affect pleasantness. Although smoothness was not a statistically significant factor for valence, it was for other relevant domains such as hedonic quality, attractivity, and likeability. Smooth motion was perceived to be more hedonic, attractive, animate, and likeable than synchronously jittery motion. Thus, robots should move smoothly for positive scenarios such as birthday reminders and music physicalizers while they should move in a jittery manner for negative scenarios like low battery or an approaching deadline. Interestingly, asynchronous jitter was perceived differently than synchronous jitter in some domains: it was rated as animate as smooth motion and as average for likeability between smooth and synchronous jitter. One possible cause may be that synchrony in jitter makes the robots look more machinelike and that negatively affects the likeability of the robots. Thus, to represent negative and less animate scenarios like low battery (20%), robots should employ synchronous jittery motion while moving in an asynchronously jittery manner for negative scenarios that require you to be more animate, such as a paper deadline approaching in an hour.

4.5 Example Applications

Using the study results, I designed abstract multi-robot motion for several example applications using the modified version of Zooids [5].

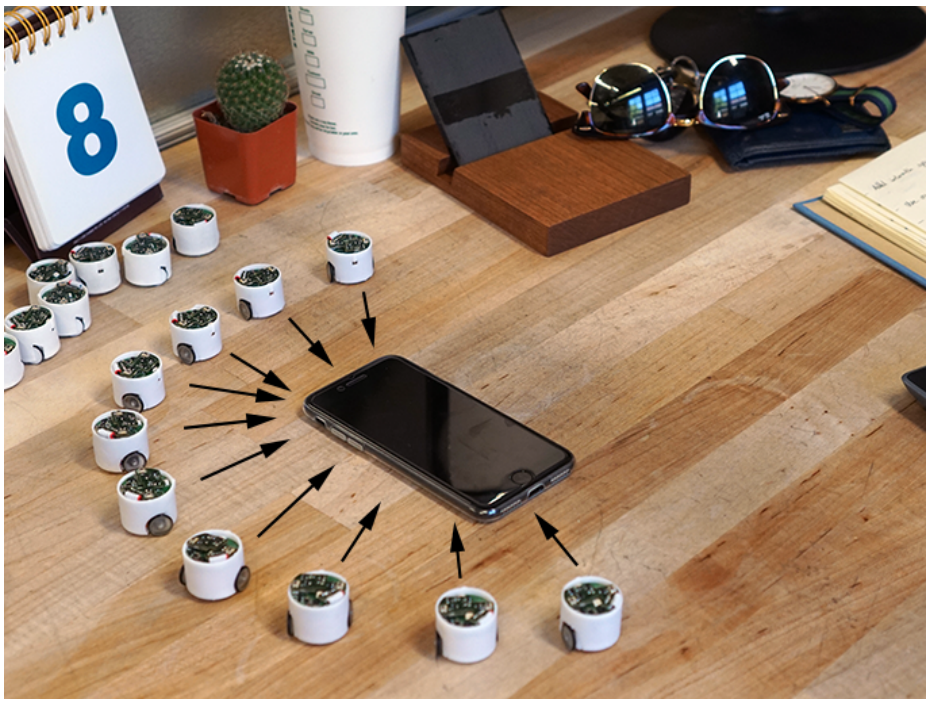


Figure 4.5: Phone call notification: modified versions of Zooids rapidly and smoothly move (rendezvous) toward a phone for notification of an urgent call. Adapted from [1].

4.5.1 Phone Call

For an important phone call that needs immediate attention, we designed a fast and smooth rendezvous behavior as shown in Figure 4.5. Out of the five behaviors, rendezvous has higher ratings for arousal, urgency, and willingness to attend. Although dispersion is also rated similarly, rendezvous is more suitable as it focuses attention to a point or in this case, to the phone. In addition, dispersion requires robots to be around the phone to begin with while robots can be anywhere for rendezvous. Fast and smooth motion is used as faster speed is perceived to be more pleasant, arousing, and urgent while smoother movement is more likeable and attractive.

Figure 4.6: Tea timer: modified versions of Zooids slowly and smoothly rotate around the tea acting as a timer and disperses when the tea is ready. Adapted from [1].

4.5.2 Tea Timer

For the tea timer application, robots move in two distinct manners: initially, a slow smooth clockwise torus behavior while tea bag is immersed as in Figure 4.6 followed by a fast smooth dispersion to let the user know that the tea is ready. During the waiting phase, a clockwise torus is used due to its low perceived urgency, willingness to attend, and arousal while being perceived as highly pragmatic. I chose the clockwise direction due to its resemblance to normal clock hands movement. Slow and smooth movement was designed to elicit calm, likeable, and attractive perception. When the tea is ready, robots disperse rapidly and smoothly to provide arousing and urgent yet attractive and likeable sensation similar to the previous phone call application but with in the opposite direction.

4.5.3 Reminder

To remind the user of a paper deadline that is approaching, UbiSwarm rapidly flocks toward a physical calendar with asynchronous jitter as in Figure 4.7. Flock behavior is used since

Figure 4.7: Event Reminder: modified versions of Zooids rapidly flock toward a calendar with asynchronous jitter to notify of an urgent deadline. Adapted from [1].

it is perceived average in terms of hedonic, pragmatic and urgency while rated high in willingness to attend. Fast movement with asynchronous jitter is designed to create an arousing, less hedonic, yet likeable movement.

4.5.4 Low Battery Status

To indicate low battery status for a phone, robots slowly rendezvous toward it with synchronous jitter as in Figure 4.8. Rendezvous motion draws attention toward the phone, while fast speed with synchronous jitter provides arousing, urgent, unpleasant, and life-less sensation to inform that the phone needs to be charged.

4.6 Limitations & Future Work

Although this perception study provides general design guidelines for abstract motion of URIs, it does not provide complete insight into abstract motion of all URIs. First, the study results are not generalizable for all sizes of robots, since only a specific size of robots was

Figure 4.8: Low battery notification: modified versions of Zooids slowly move (rendezvous) toward a phone with synchronous jitter to signal low battery. Adapted from [1].

used here. Although I speculate that similarly scaled robots will be perceived similarly, I believe that human or automotive-scale robots will most likely yield different perception results. In future work, it would be interesting to see how different sizes of URIs, ranging from hand to human to automotive scale, elicit different human perceptions.

The number of robots used during the study is also limited to ten. This was due to the size of the workspace and it is possible that different numbers of robots will yield different responses. Similar to the size of robots, I speculate similar numbers of robots will yield similar results but hundreds or thousands of robots are likely to result in different perceptions. Either by decreasing the robot size or by increasing the workspace, I could study the effect of different numbers of robots.

While the study results suggest statistical significance of motion parameters for many dependent variables, the effects were relative. When looking at the absolute values in the Likert scale, one can see that the range of different perceptions is not wide-spread except for urgency. In the future, I will look into different ways of widening the range of various perceptions such as comparing different sizes of robots and a wider range of speed.

In terms of implementation, the mobility and sensing ability of the robots can be further improved. While they can move on both horizontal and ferromagnetic vertical surfaces, the transition from one to another is not possible yet. We hope to address this by adding a ramp between the surfaces. In terms of sensing, Zooids can currently localize itself and sense user's touch but cannot detect other objects for obstacle avoidance or object manipulation. In the future, we will address this by adding either a tracking marker to the objects or by using an external camera to detect them.

Finally, there are fundamental limitations of URIs compared to ubiquitous display either through projection, screens or wearable displays. Although there are great benefits from having physical form such as ability to move around and manipulate objects, URIs are limited by their physicality. They cannot disappear instantaneously and appear elsewhere like projection or screens pixels. They are constrained by the laws of physics. We cannot simply 'copy and paste' these URIs to send them to remote locations. However, we believe that these URIs can be useful in circumstances where mobility and manipulation are required, and that URIs can complement pixel-based interfaces for other applicable scenarios. Also, just as screens have gotten more affordable over time, we believe URIs will become more affordable and truly ubiquitous in the future as costs of transistors and robots continue to fall.

4.7 Chapter Summary

Mobility is a fundamental aspect of any robotic system. It allows robots to manipulate objects and sense large environment while also enabling them to physically interact with users. In this chapter, I investigated ways to layer expressiveness such as affect and urgency on top of swarm robot motion, to help robots convey additional information potentially even while performing other tasks. As a first step towards this goal, I ran a crowdsourced between-subjects human perception study to investigate abstract multi-robot motion as a URI display. The abstract motion consisted of three motion parameters: collective behaviors, speed, and smoothness. The study results suggest that different behaviors elicit significantly different responses in arousal, dominance, hedonic and pragmatic qualities, animacy, urgency and willingness to attend. On the other hand, speed significantly affects valence, arousal, hedonic quality, urgency and animacy while smoothness affects hedonic quality, animacy, attractivity and likeability. These results serve as design guidelines for

URI-based displays and I demonstrate these through my own example applications. While there are many more methods to leverage abstract motion as a display (e.g., using principles of animation [164]), the results demonstrate the feasibility and effectiveness of collective behaviors to generate more expressive swarm robot motion. While this chapter is focused on conveying abstract information like affect and expressiveness, the next chapter will focus on delivering more task-oriented information like the intent of the robots.

Chapter 5

Generating Legible Swarm Robot Motion

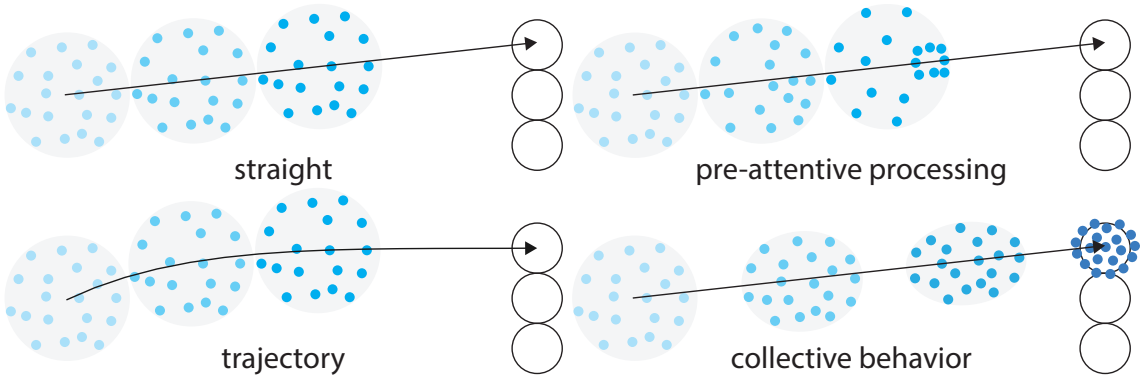


Figure 5.1: In this chapter, I design and evaluate different types of legible motion based on trajectory, pre-attentive processing features, and swarm behavior. The blue circles represent robots that are moving toward their goal represented by the white circles.

In the previous chapter, I investigated how to leverage abstract swarm robot motion to improve expressiveness and convey information like affect. In this chapter, I focus on generating legible and glanceable swarm robot motion that visually depicts the intent or the goal of the robots.

5.1 Introduction

Fleets of autonomous robots are beginning to occupy our environment. They are being used for many tasks such as transportation of goods and people, search-and-rescue, and agriculture. 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 ranging from tele-operator to supervisor to end-user.

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 [20], accomplish tasks more efficiently [108], and potentially build trust and human-robot rapport. The legible motion for a swarm of robots is especially important and relevant in the future of ubiquitous robots as robots will need to be usable even by people who have no prior experience interacting with robots.

In this chapter, I explore how we can improve the *legibility* and *glanceability* for a swarm of robots. Due to increased number of agents that all require attention, it is important to generate *glanceable* motion, that is pre-attentively legible, to expedite the robot monitoring process for observers. In order to design legible and glanceable swarm robot motion, I leverage findings from the legibility of a single robot's motion [165], swarm intelligence literature [166], and visual perception [45]. Specifically, I investigate how the trajectory-based legible motion that has been used for a single robot [19, 108] can be applied to a group of robots, how collective behavior (e.g., rendezvous) could be used to inform the user about the goal, and how pre-attentive processing features (e.g., density) could be used in designing swarm motion that rapidly directs user's attention toward the desired goal.

The underlying mechanisms behind each of these legible motions are different. Thus, I 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, I evaluated and compared the legibility and glanceability performances of the different legible motions as well as the effects of target difficulty and initial spread. I 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 chapter 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.

5.2 Legibility and Glanceability of a Robot Motion

I first define the legibility and glanceability of a swarm robot motion. For legibility, I follow the same definition as prior work in legibility of a single robot motion [19,108,111]. 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 [110, 165].

On the other hand, 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 showing the entire robot trajectory, an observer is only shown a short segment of the motion and needs to predict the intent of the robot. This 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, users will have to constantly shift focus from one group to another, spending a limited amount of time on each group. In this chapter, I propose a time limit of 250 ms as my goal is to generate robot motion that can be processed pre-attentively.

5.3 Designing and Generating Legible Swarm Motion

In order to design legible swarm robot motion, I first formulate the problem as a non-linear constrained optimization problem. Then, I 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, I explain how I leverage findings from prior works in single robot legible motion [19,108], swarm intelligence [7], and visual perception [45,117]. The objectives and design criteria are explained for trajectory, swarm behavior, and pre-attentive processing feature-based legible motions.

5.3.1 Problem Formulation

As introduced by Witkin and Kass [167] and used by Bodden et al. to generate intent-expressive motion [108], I also 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:

$$X^* = \text{argmin } g(X) \text{ subject to } c_i(X) \tag{5.1}$$

where X^* denotes the optimal trajectory. I 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, $X(t_f) = X_f$

5.3.2 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 [168, 169] or is minimal in the required energy [167]. Here, I use the setup below to compute the straight

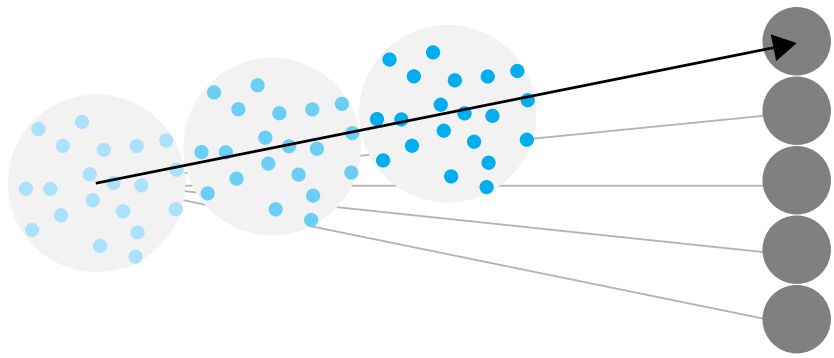


Figure 5.2: One of the common path planning algorithms minimizes the overall trajectory length resulting in a **straight motion**. For a group of robots, this results in robots synchronously following a straight trajectory toward the goal.

trajectory that minimizes the overall length resulting in trajectories as depicted in Figure 5.2, and use it as the control condition for the experiments.

$$Straight = \int_{t_o}^{t_f} \|X(t)'\|^2 \, dt \tag{5.2}$$

5.3.3 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 [19, 108, 110, 111]. 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 [112,113], both papers employed an arc-trapezoidal trajectory based

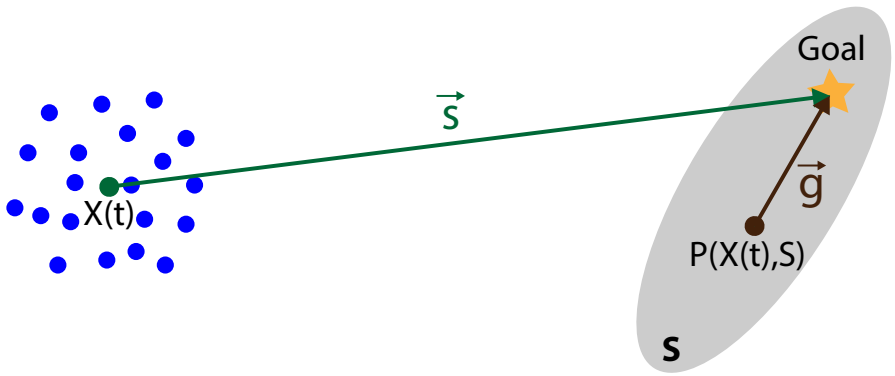


Figure 5.3: The trajectory-based legible motion is a result of the optimization of the distance from the predicted goal to the actual goal (i.e., $\vec{g}(X(t))$), the distance from the center of the robots to the goal (i.e., $\vec{s}(X(t))$), and the overall trajectory length (i.e., $\|X(t)'\|^2$).

on the animation principles of "Slow in and out" and "Arcs" from [164]. They did not use trajectory specifically designed and proven to improve legibility of a single robot motion such as [19, 108]. Thus, I propose using and evaluating the adapted algorithm from [108] to generate the legible trajectory for a swarm of robots.

I use the same optimization framework as shown below:

$$Legible = \int_{t_o}^{t_f} \alpha\|\vec{g}(X(t))\|^2 + \beta\|\vec{s}(X(t))\|^2 + \epsilon\|X(t)'\|^2 dt \tag{5.3}$$

This minimizes the following components:

$$\vec{g}(X(t)) = Goal - P(X(t), S) \tag{5.4}$$

$$\vec{s}(X(t)) = Goal - X(t) \tag{5.5}$$

As shown in Figure 5.3, $\vec{g}(X(t))$ 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 $P(X(t), S)$, I use the "point position" formula from [108], 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.

The optimization weights α and β determine the trade-off 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 [108]. The third term in Equation (3) is a regularization

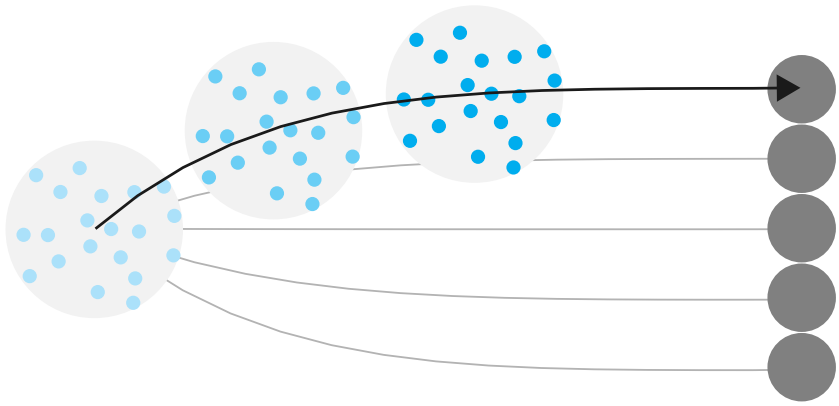


Figure 5.4: Examples of **trajectory-based legible motion** are shown. Due to space constraints, only one example is shown of how the robots follow the trajectory while maintaining the same formation.

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 5.4.

5.3.4 Collective Behavior-based Legible Motion

In contrast to the trajectory-based legible motion where I 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 [1]. 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 [1]. Thus, I 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, I use the existing literature on circle packing to derive the final configurations for the robots that are closely packed without any overlap [170]. 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

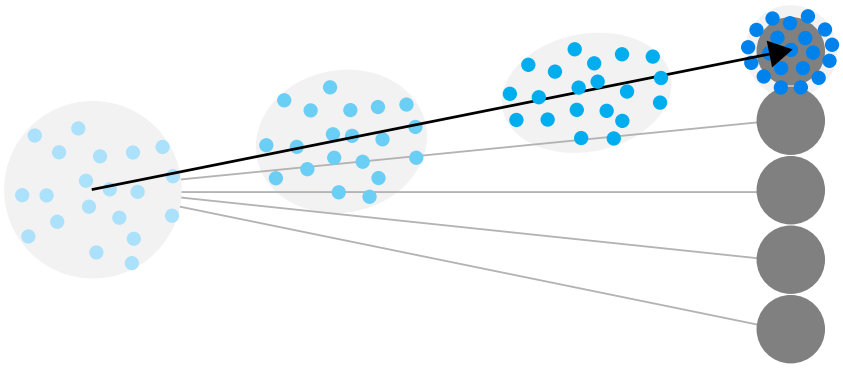


Figure 5.5: **Collective behavior-based legible motion** is plotted. Due to the space constraints, there is only one example of how the robots follow a straight trajectory while also rendezvousing toward the goal.

the trajectories, I use the Hungarian algorithm [144]. Note that the above two steps require some centralized coordination. Finally, the straight trajectory for each robot is generated with equation 5.3.2. The resulting motions for each target are shown in Figure 5.5.

5.3.5 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, I explored utilizing existing knowledge in the vision perception field, specifically about the pre-attentive visual processing features [45]. Researchers have identified visual parameters that are processed subconsciously within approximately 250 ms and they include factors such as color, shape, orientation, and density [117]. As legible swarm robot motion needs to provide hints to cue where the robots are headed, I saw potential in using pre-attentive processing features as they could be used to rapidly direct user's attention toward a desired region within the collection of the robots. Out of the many potential candidates, I became particularly interested in using density as it only requires the ability to control the distribution of the robots and does not need external light sources or shape-changing capabilities.

In order to uniformly distribute the robots within the two regions of different densities, I pre-computed this desired distribution by using the Centroidal Voronoi Tessellation (CVT)

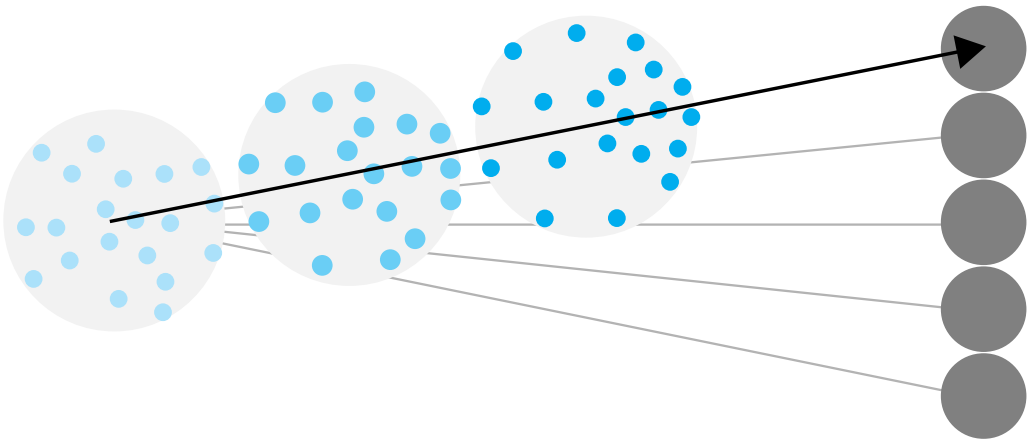


Figure 5.6: **Pre-attentive processing feature-based legible motion** is shown. The robots change formation while the following a straight trajectory overall. Due to the space constraints, the transition from the initial randomly distributed formation to the final formation is stretched.

[171] which was used to uniformly distribute robots for animation display with multiple mobile robots [40] and mobile sensing network [172].

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 based on the distance from its target similar to the method in [40]. An example of the resulting motion is shown in Figure 5.6.

5.4 Evaluation 1: Legibility

To evaluate and compare the different legible motions, I 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, I 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, I investigate the effects of initial spread of the robots and target difficulty as I 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.

5.4.1 Hypotheses

The trajectory-based multi-robot legible motion is a direct adaptation of the prior work on single robot legible motion [19,108]. As prior work have shown the effectiveness of leveraging trajectory to increase legibility of a motion, I 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, I 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, I 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, I 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 [45, 117] that can be detected rapidly by human visual processing pathways. Thus, I 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. I 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 [112], and thus I also expect

decrease in prediction time. Since the nature of different legibility cues are different, I 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, where as 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.

- *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.

5.4.2 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 5.7. 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

Table 5.1: 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

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 I 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 [5]. The values of the constant task parameters used and the initial radius are listed in Table 5.1. The different robot motions were programmed in C++ in Visual Studio using the Zooids Simulator [5].

5.4.3 Independent Variables

To test my hypotheses, I evaluate the performance of different legibility cues including control, trajectory, rendezvous, and density. As I expect the initial radius R_i to have an impact on performances of different legibility conditions, I also tested three different values of initial radius that were determined experimentally through pilot studies. Finally, I 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, I 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).

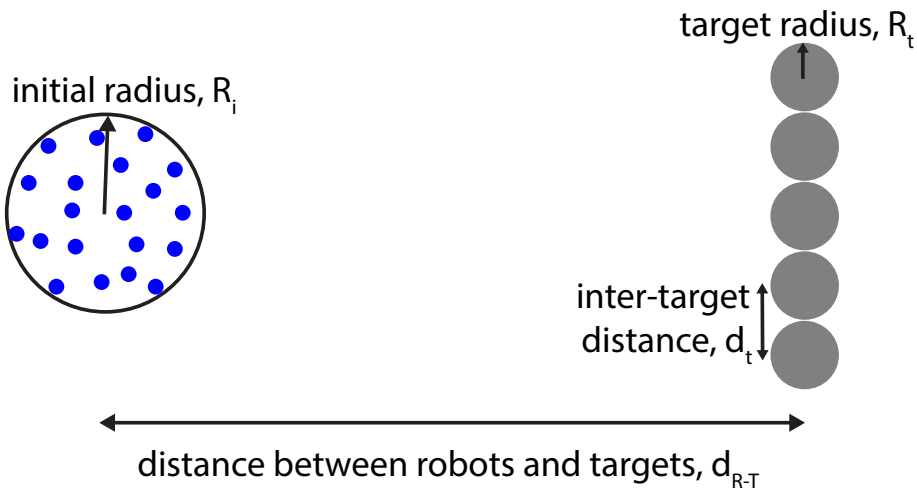


Figure 5.7: Task setup for user study with robots initially on the left side and five targets on the right side

5.4.4 Measures

For each trial, I recorded the participant's target prediction, prediction time, and prediction confidence on a Likert Scale from 1-7 as done in [110]. I also combined these three answers into a single legibility score metric similar to prior work [165] as below: If the prediction is correct,

$$Score = \frac{TotalDuration - PredictionTime}{TotalDuration} \times \frac{Confidence}{7} \tag{5.6}$$

Otherwise, I assigned a score of 0 to incorrect predictions.

I also collected subjective ratings of the robots in terms of legibility and predictability as used in prior work [108].

5.4.5 Procedure

For the legibility task, participants were asked to stop the video when they felt confident enough to make a prediction. 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 and provided demographic information.

5.4.6 Participants

I recruited 283 participants through Amazon Mechanical Turk. For each of 12 conditions (4 legibility cue conditions x 3 initial radius), approximately 20 participants viewed and rated the corresponding video. For quality control, I 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.

5.4.7 Analysis

To investigate both main and interaction effects for the overall legibility score, prediction time, prediction confidence, and accuracy, I 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 difficulty). Then, I performed the Bonferroni-corrected post-hoc tests on the statistically significant effects.

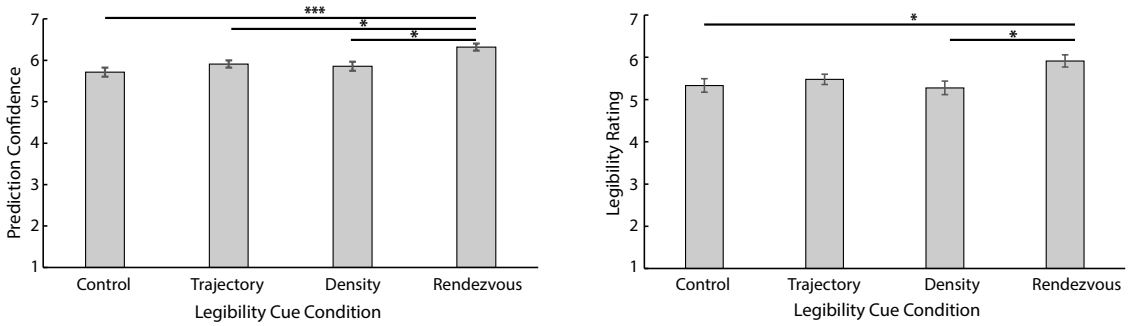
Similarly, I ran a 4×3 between-subjects ANOVA with with two between-subject factors (i.e., condition and initial radius) for the subjective ratings. Then, I performed the Bonferroni-corrected post-hoc tests on the statistically significant effects.

5.4.8 Results

All the bar graphs (figures 5.8-5.11) 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 $***$: $p < .001$.

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 Fig.5.8a and 5.8b respectively. Out of the four different legibility cues, the rendezvous behavior-based legible motion had the best legibility performance in terms of self-reported



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

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

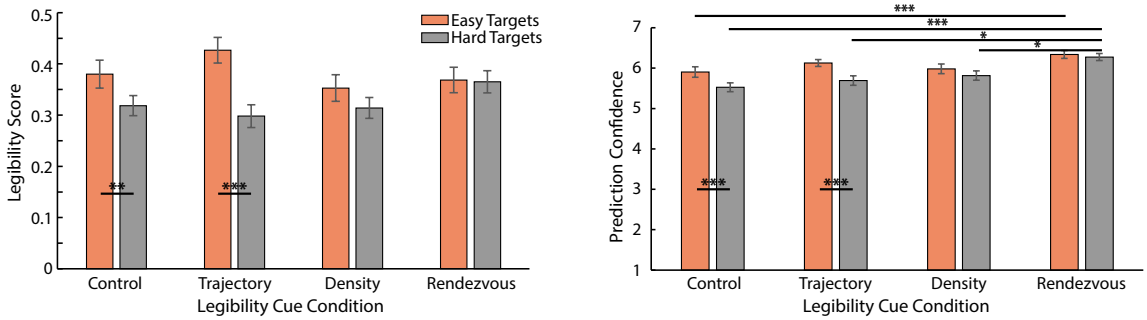
Figure 5.8: (a) Rendezvous had the highest confidence ratings among all conditions. (b) Rendezvous condition is rated higher than density and control conditions.

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. On the other hand, 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.

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 ($F(1, 214) = 34.24, p < 0.001, \eta^2 = 0.138$), and accuracy ($F(1, 218) = 3.96, p = 0.048, \eta^2 = 0.018$). Harder targets (i.e., three targets in the middle) had lower legibility score, confidence ratings, and prediction accuracy.

As shown in Fig 5.9a, 5.9b, and 5.10a, 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. I 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, I 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



(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

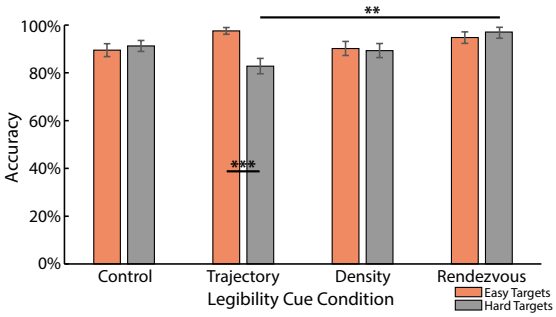
Figure 5.9: (a) There are significant difference between the target difficulty levels for the control and trajectory conditions but not for the density and rendezvous conditions. (b) For hard targets, rendezvous outperformed all other conditions while it only outperformed the control condition for easy targets. The confidence ratings are significantly different between the two target difficulty levels for control and trajectory conditions.

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.

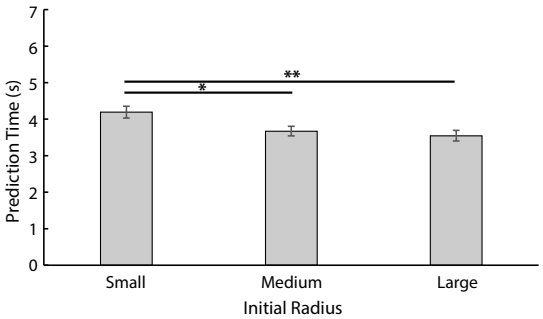
While the study results suggest that collective behavior-based motion is the most legible motion in terms of self-reported prediction confidence and self-reported legibility rating, different legible cues had varying performances based on the context or difficulty of the task. For instance, the trajectory condition had the highest overall legibility score when the target difficulty was easy (mean=0.429) albeit not significantly as shown in Fig. 5.9a. Thus, this suggest 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.

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 Fig. 5.10b 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 ($F(2, 218) = 5.23, p = 0.006, \eta^2 =$



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

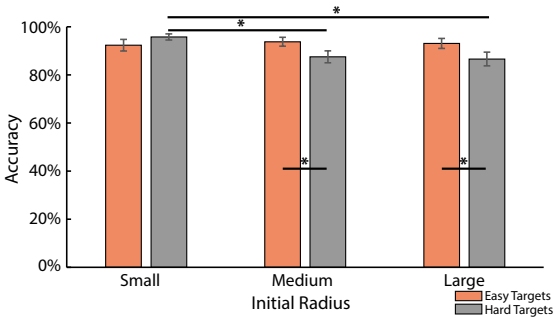


(b) Effects of initial radius on prediction time

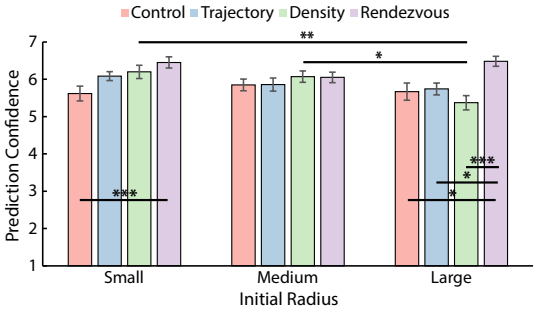
Figure 5.10: (a) There was a significant difference between the target difficulty levels for trajectory condition, and the accuracy was significantly higher for rendezvous condition than trajectory condition when the target difficulty was hard. (b) With larger initial radius values, the prediction time significantly decreased.

0.046), where the accuracy decreases with larger initial radius as shown in Fig. 5.11a. This trend of faster prediction agrees with prior work [112], 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 [112]. These results support H1.7 but also provide evidence that accuracy is decreased with larger radii.

As hypothesized in H1.8, initial radius had different effects based on the type of legibility cue. As shown in Fig. 5.11b, 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. I 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.



(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

Figure 5.11: (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 difference 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.

5.5 Evaluation 2: Glanceability

With a swarm of robots, it is unlikely that all of the robots will be performing the same task simultaneously. Rather, groups of robots will be assigned to different tasks in order to optimize the overall efficiency. In such scenarios, supervisors of these systems face an exhausting and complex task of trying to monitor multiple groups at the same time. To alleviate some of the burden, I 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 [116], in this chapter our goal is to generate robot motion that can be processed pre-attentively and thus use the threshold of 250 ms [45]. 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 previous Section, I evaluate their glanceability and the effects of task parameters such as target difficulty and initial radius.

5.5.1 Hypotheses

The trajectory-based legible motion plateaus after the initial leap toward the goal. Thus, I 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 I 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, I 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, I 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.

5.5.2 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.

5.5.3 Independent Variables

In addition to the set of independent variables used in the prior study, I varied the timing of video segment shown to the participants. As I wanted to evaluate the glanceability of the motion, I 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, I chose to show participants 250 ms clips beginning at 1 second (i.e., time 1) and 2 second (i.e., time 2) of the trajectory. As prior study indicated that most participants make the prediction around 4 second (i.e., time 4), I decided to show 250 ms at the 4 second as well.

5.5.4 Measures

There currently is no measure or instrument for glanceability. Hence, I propose a glanceability score metric which 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 [110, 165], where the legibility scores at each time stamp 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 this experiment, I recorded the participant's goal prediction and prediction confidence for each trial on a Likert Scale from 1 to 7. Then, I summed the predictions and prediction confidence ratings from 1, 2, and 4 seconds into a single glanceability score metric as below:

$$GLANCEABILITY = \sum_{T} \frac{T_{total} - T}{T_{total}} \times \frac{R_{conf}}{7} \times \mathbf{1}[Prediction == Correct] \tag{5.7}$$

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.

5.5.5 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 5 different targets and 3 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.

5.5.6 Participants

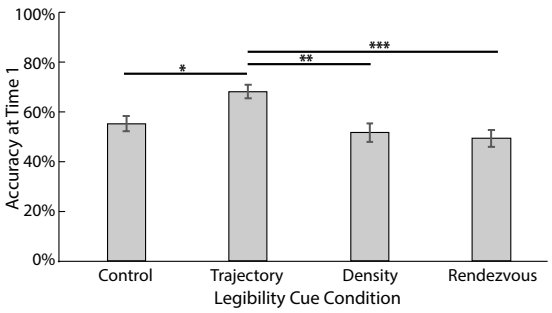
I recruited 240 participants through Amazon Mechanical Turk. For each of the 12 conditions (4 conditions x 3 initial radius values), approximately 20 participants were randomly assigned to a condition. They then viewed and rated 15 videos (5 different targets x 3 time segments) corresponding to that condition. For quality control, I 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 a hourly rate of approximately $15.00 US dollars.

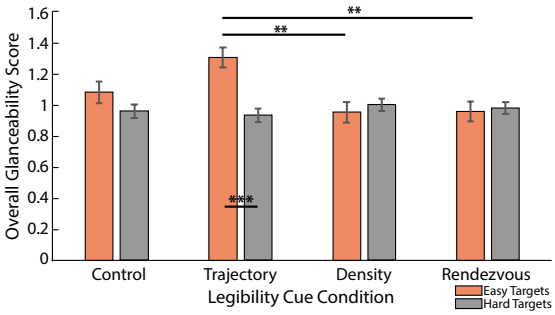
5.5.7 Analysis

To investigate both main and interaction effects for the overall legibility score, prediction time, prediction confidence, and accuracy, I 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, I performed the Bonferroni-corrected post-hoc tests on the statistically significant effects.

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



(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

Figure 5.12: (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.

5.5.8 Results & Discussion

All the bar graphs (Figure 5.12-5.15) 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 $***$: $p < .001$.

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 Fig. 5.12a. At time 1, trajectory condition had significantly higher accuracy than other conditions. Diving deeper, I see from Fig. 5.13 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 well in the beginning but not at time 4. On the other hand, 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.

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 5.12b, 5.13, and 5.14, 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 ($F(3, 248) = 3.1, p = 0.027, \eta^2 = 0.036$). 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 Fig. 5.14. This suggests that the rendezvous-based legible motion is less affected by the target difficulty than other types of legible motion.

Lastly, 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.

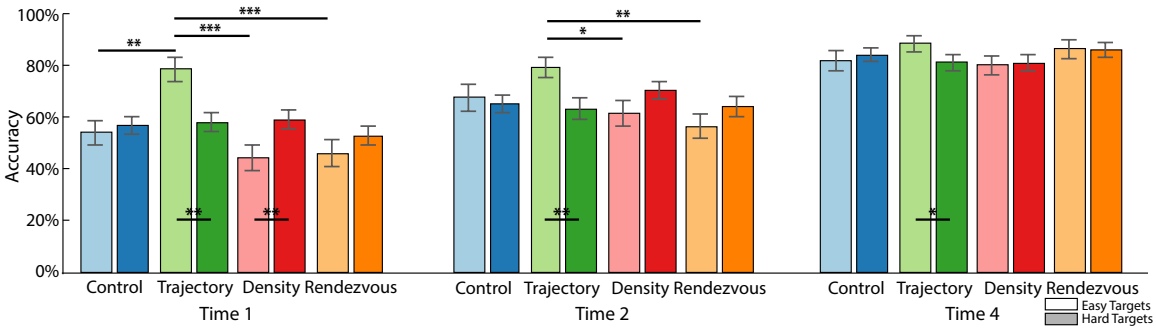


Figure 5.13: Interaction effects between target difficulty and legibility cue on prediction accuracy across time

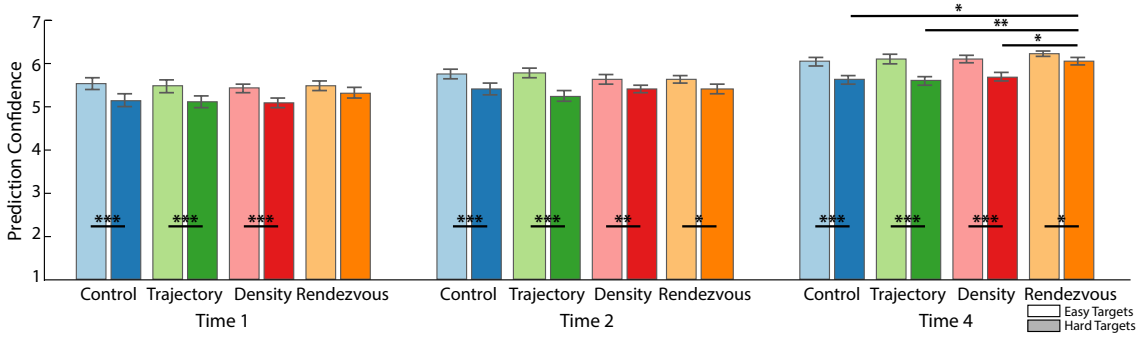


Figure 5.14: Interaction effects between target difficulty and legibility cue on prediction confidence across time

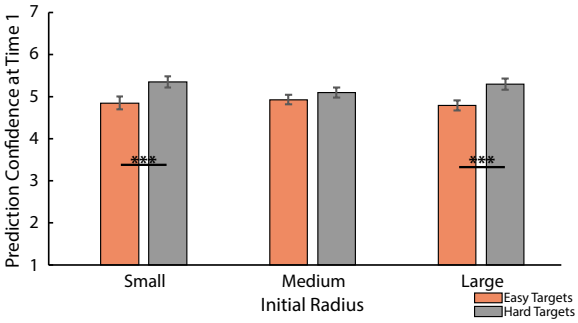


Figure 5.15: Interaction effects between target difficulty and initial radius for prediction confidence at time 1

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 Fig. 5.15. 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 I did not observe any interaction effects between radius and legibility cue, I cannot accept H2.7.

5.6 Discussion

From the two evaluations on the legibility and glanceability, I observed both similar and different trends. In general, I 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 Fig. 5.9a and 5.12b. 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 [19, 108], I 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, I saw 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 seem to help convey intent of the robots to the observers especially when exposed to short snippets of the motion.

As I 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 I did not see a significant improvement in legibility as our task may been more difficult than

those from prior work [19, 108]. 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. 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.

5.7 Limitations & Future Work

This chapter provides preliminary investigation of different ways to enhance legibility of a multi-robot system. While I varied some parameters (e.g., the initial radius and the target difficulty) in addition to the legibility cue, I 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. For the density-based legible motion, I 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 I see that even changing the absolute size impacts the prediction confidence, I 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 I did not observe any significant effect of radius

on the performance of swarm behavior-based legible motion, I did not test a wide range of initial radii. With a drastically larger radius, I 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, I used a threshold of 250 ms as I 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 this study results may serve as a starting point, further studies should be conducted to verify if the glanceability results still hold for the desired application.

While this chapter investigates these different types of legible motions independently, I 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 I observe significant effects of task parameters (i.e., target difficulty and initial radius) on legibility and glanceability, I could use the motion optimal for each setting. For instance, I 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 goal. In the real world, there will be cases when robots will need to reach and manipulate multiple objects. While Capelli et al. explored this problem, 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 [113]. In the future, I plan to explore different combinations of the features, both from this chapter and prior works, to address the multi-target problem.

5.8 Chapter Summary

With increasing number of robots that interact and collaborate with people, it is important to equip robots with intent-expressive movements that facilitate such experiences. In this chapter, I explored the use of trajectory, pre-attentive processing features, and collective behaviors in order to generate legible and glanceable swarm robot motion. I conducted 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 whereas the trajectory-based motion has the highest glanceability. I also observed significant effects of task parameters like 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 will require a more complex resolution that consists of a combination of these different legibility cues based on the context of the task. Chapter 4 and this chapter study how to best leverage the mobility of the robots to create a visual language that can communicate affect and intent of the robots. The next chapter will continue to explore output from the robots to users. However, it will explore the use of haptic channels to display information to users. I will generate the design space and conduct two user studies to evaluate human perception of different haptic stimuli from robots and how people use robots to convey different types of social touch.

Chapter 6

Haptic Display with Swarm Robots

Figure 6.1: *SwarmHaptics* uses a swarm of robots to display various haptics patterns to different body parts that are on a surface or through external objects such as a mouse. It can be used to convey notifications, social touch, directional cues, etc. Adapted from [2].

In the previous two chapters, I studied how multi-robot motion could be used to visually display abstract information such as intent and affective state of the robots. In this chapter, I investigate how multi-robot systems can communicate in proximity through direct touch with the users.

6.1 Introduction

We experience many touch events throughout our daily lives. They range from simple activities like holding a cup to more complex social interactions like consoling a friend. As more robots become a part of our personal and professional environments, it is becoming increasingly important to equip the robots with social skills for fluid in situ interaction with people. While motion or verbal communication may work most of the times, touch is an under-utilized channel that could provide urgent notification especially in noisy and crowded environment, or help connect people on a physical level rather than purely visual means.

Thus in this chapter, I explore how a swarm of robots could communicate users through touch. In particular, I focus on haptic display with multi-robot systems and show example scenarios including haptic notifications, directional cues, and remote social touch, enabling people in distant places to connect through touch. As such, I introduce SwarmHaptics, a new type of visual and haptic display with a swarm of small mobile robots. These robots can approach users' different body parts close to the surface they are on and display haptic patterns that vary in spatial, temporal, and force coordination. The motions of the robots can also serve as visual displays to provide more context and complement haptic feedback.

To better understand how people perceive SwarmHaptics and its personal space intrusion [173], I ran an in-lab within-subjects study to see the effects of different haptic stimuli using Zooids [5]. Specifically, I investigated how different haptic parameters such as the number of robots, force type, frequency of force applied, and force amplitude affect human perception of emotion, urgency, and HRI metrics. I displayed the haptic stimuli on the participants' dorsal side of the forearm as it is one of the more socially acceptable areas to touch [174] and provides ample room for multiple robots to make contact with. From the study results, we find that the number of robots, force frequency, and amplitude have significant effect on human perception, whereas force type only has interaction effects.

Lastly, to gain insights on how users would convey different types of social touch such as positive affection and ritualistic touch, we developed a platform to control multiple robots simultaneouslyand ran an elicitation study. I asked participants to generate different haptic patterns given referents relating to social interactions (e.g., greeting, hug), affective communications (e.g., happy, sad), and functional communications (e.g., notification). Although some referents elicited similar interactions, the results help demonstrate the expressiveness of SwarmHaptics for social touch.

In summary, the contributions of this chapter are:

- Exploration of the design space for SwarmHaptics,
- Perception study results on swarm haptic stimuli, and
- Elicitation study results for social touch.
-

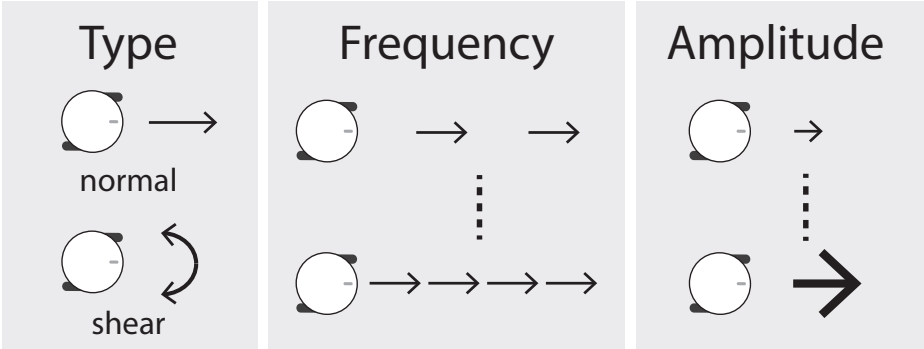


Figure 6.2: Force parameters for a single mobile robot. A robot can generate forces of different types (normal/shear), frequencies, and amplitudes to different locations. Adapted from [2].

6.2 Related Work

There are many different types of haptic devices. Examples include vibrotactile [175, 176], wearables [177–181], shape-changing displays [182–184], hand-held [185–187] and encountered-type haptic devices [188–191]. While swarm robots may not be able to provide the richest sensation compared to the other haptic devices, it can support multiple users simultaneously and it does not have to be constantly worn by the user which can cause discomfort. Also, the robots are mobile, allowing haptic designers to control when and where the haptic sensation should be delivered. Finally, a swarm of robots can be used for many other applications such as object manipulation and not just dedicated to haptic display.

6.3 Design Space of SwarmHaptics

I first explore the design space for a haptic display with swarm robots. Specifically a group of simple, mobile robots with no end-effector is chosen because it is one of the most rudimentary type of robots. Thus, the resulting design space can be more generalizable to other mobile robots. I begin with investigating the design space for a haptic display with a single robot, then broaden the scope to include a swarm of robots.

6.3.1 Haptic Parameters for a Single Robot

Force Parameters

A simple mobile robot can only move in 2-D space, and is limited to 2-D translation and 1-D rotation. Thus, the types of haptic stimuli that it can provide are also limited by its motion capabilities. When using pure translation, a simple mobile robot can apply normal force to the user. Whereas it can generate shear force through pure rotation or by moving along the skin. I can also control the magnitude and frequency of the haptic stimuli by adjusting the magnitude and frequency of the commanded speed/torque to the motors. By frequency, I refer to the rate at which the robot move forward and backward or rotate clockwise and counterclockwise to impart force on users. All forces generated by the robot are grounded to the surface that it is driving on. Overall, we can control the motion of a simple mobile robot to generate haptic stimuli with different force type, magnitude, and frequency as shown in Figure 6.2.

Contact Location

Because of its mobility, the robot can move to different accessible body parts and provide haptic stimuli. For instance, a robot on a desk can touch the user's finger, hand, wrist, elbow, and both sides of the forearm. On the other hand, a robot could provide haptic feedback to the user's feet on the ground to push away from certain areas or to the body while lying down to wake him/her up. Due to the varying mechanoreceptors and haptic sensitivities of different parts of the body, the same touch stimulus can feel different depending on the location, even just throughout the arm [192]. For example, pushing with 1N of force on a fingertip will feel much greater than pushing on the shoulder with the same force. In addition, the social appropriateness of the touch needs to be considered. Some body parts are more socially appropriate to touch such as the arm and the shoulder [174]. Thus, developers will have to carefully select the location of the haptic stimuli based on the application.

Tactile vs. Kinesthetic

Depending on the contact location and the motion of the robot, the robot can provide either tactile or kinesthetic stimuli. The stimuli will be kinesthetic when the magnitudes of the forces are great enough to move the joint angles in the body whereas they will be tactile if the forces are weak and thus only stimulate the skin. For example with our system, even

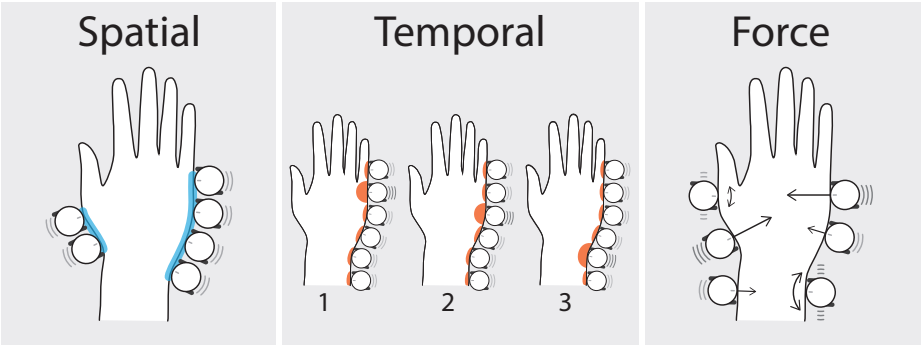


Figure 6.3: Types of coordinations possible among a swarm of robots: we can coordinate where the robots provide haptic stimuli, when they apply the forces, and which forces in terms of force type, frequency and amplitude they generate. Adapted from [2].

a single robot is sufficient to move a single finger but at least seven robots are needed for an entire arm. Thus, we need to ensure that the contact location, magnitude of force, and type of motion are properly selected to produce the desired haptic effect.

Size and Texture

The size or form factor of the robot is another important parameter that could impact the interaction experience. Prior works have shown that people perceive telepresence robots differently even when just the height is changed [193]. Similarly, any significant change in other form factor such as size and shape is highly likely to influence user's perception. On the other hand, contact material and softness have been shown to have significant effect on perceived pleasantness [194]. Thus, even with the same force type, frequency, and other haptic parameters, changing the texture of the robot will most likely effect user's perception.

6.3.2 Haptic Parameters for a Swarm of Robots

Number of Robots

The most basic parameter for swarm robots is the number of robots. As people behave differently based on the number of people they interact with [195], I conjecture that the number of robots will change how people perceive, behave, and interact with them. Also, more robots increase the degrees of freedom for haptic expressivity.

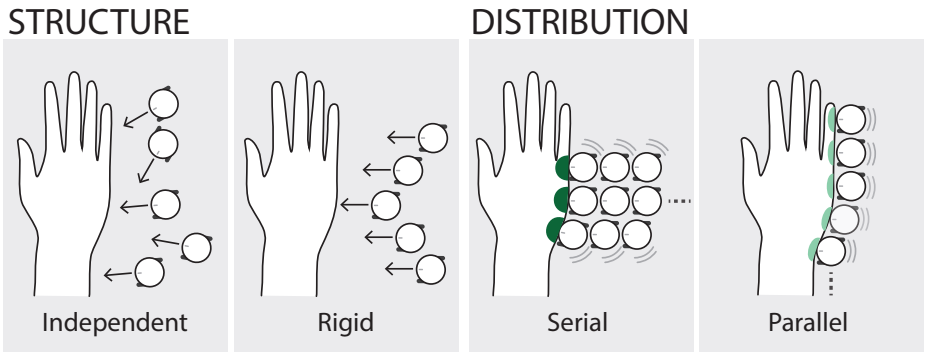


Figure 6.4: Spatial parameters: Structure between the robots, from independent robots to robots in rigid formation, affects user perception. Distribution of them either serial or parallel, determines contact area size and resulting forces. Adapted from [2].

Coordination

A more complex design parameter for a swarm of robots is the coordination between them. With more robots, it becomes not only difficult to control them [86], but also it is uncertain how to best coordinate them for different applications. As shown in Fig. 6.3, I propose three ways to coordinate the robots: spatial, temporal, and force.

Spatial distribution: With many robots, we need to determine how to spatially distribute them. There are many factors to consider such as the desired resulting force and the users' comfort. With multiple robots, it is possible to combine their forces to create different haptic patterns. For instance, with one robot, it is impossible to provide "squeeze" sensation to the user's forearm.With many robots, we can distribute the robots to both sides of the forearm and command normal forces to generate the "squeeze" sensation. At the same time, we need to consider users' comfort when touching multiple locations with the robots. Users may not be comfortable with being surrounded by the robots or being touched in particular areas [174].

Independent motion vs. Rigid structure:

In addition to the robots' relative positions to the user, the relative positions between the robots need to be considered. Inter-robot interaction can affect how the users perceive haptic stimuli from the robots. For example, haptic stimuli from a group of robots in rigid formation can feel different than one from a group of robots that move independently. This inter-robot relation has proven to have significant effect on human perception for abstract motion [1] but needs to be studied for haptic display with swarm robots.

Serial vs. Parallel distribution:

With many robots, there are different ways of distributing the robots. To maximize the contact area, we can distribute the robots in a parallel fashion; while to reduce the contact area with the same number of robots, we can place them in series as shown in Figure 6.4. We can imagine the robots as representing forces, and for higher combined force at a point, one would put the forces in serial while to spread the forces to a larger area, one would use the parallel formation. An appropriate method should be chosen based on the context. For instance, to rotate a user's forearm about her elbow, it is ideal to provide the resulting force near the wrist, the furthest point from the elbow. Then, we should distribute the robots in serial near the wrist rather than placing them in parallel across the forearm.

Spatial & temporal patterns:

To further enrich the range of expressivity, one can combine spatial and temporal coordination. A simple example is a line of robots that apply normal forces to the user's arm in a sequential manner from top to bottom as shown in Figure 6.3 which could be used to provide directional cues to the user for navigation.

Force coordination:

Finally, we can coordinate the forces that each robot generates. For instance, when providing directional cues, we can modulate the magnitude of the force in addition to the frequency to enhance the fluency of the directional cue similar to how Israr et al. used an amplitude modulation algorithm for a vibrotactile array [175]. In addition, we can vary the force type that each robot provides: normal or shear. For example, we can have some of the robots output normal force while the others produce shear force as shown in Figure 6.3. While it is still unclear how human would perceive such a combination, this flexibility adds another degree of freedom for conveying information.

Mediated Display

In addition to displaying haptics directly to the skin, the robots can provide haptic sensations indirectly through external objects. Instead of augmenting each device on the desktop like LivingDesktop [196], we can provide haptic sensations through SwarmHaptics. For instance, to help a user maintain focus, robots could push the mouse that the user is holding away from links of a distracting video. By indirectly pushing on the mouse instead, this could potentially reduce the discomfort that the users may feel from a swarm of robots touching them directly.

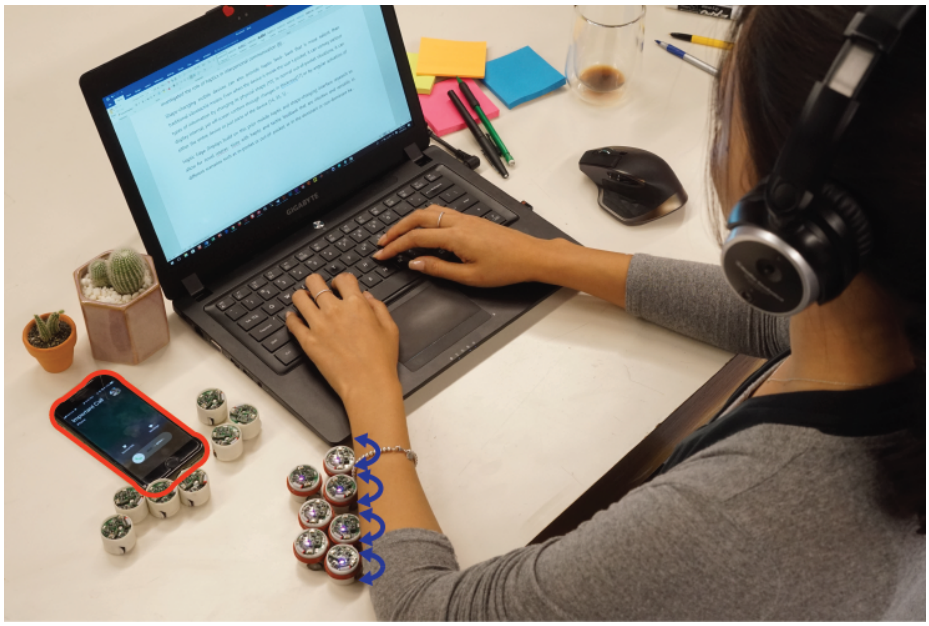


Figure 6.5: Notification Scenario: The user doesn't notice her phone (in red) is ringing because her eyes and ears are occupied. The robots approach and apply shear forces to alert her. Adapted from [2].

Visual Effects of Robots' Motion

As the robots move to produce haptic sensations, there are inherently visible motions that accompany the haptic stimuli. For instance, when the robots are providing wave-like haptic stimuli, users would also see the wave-like motion. This visual may help complement the haptic stimuli to enhance the salience of the haptic stimuli. Also, their paths and motions could help users understand the robots' intents and internal states [19].

6.4 Example Scenarios

Here I demonstrate several example scenarios of how SwarmHaptics can be used in real life situations.

Notification. SwarmHaptics can be used to notify users through touch. This can be especially useful when the other primary senses, visual and audio, are occupied by other mediums. For example, imagine Tia is writing a paper on her laptop while listening to her favorite music through her headphones. When her collaborator calls her to discuss details

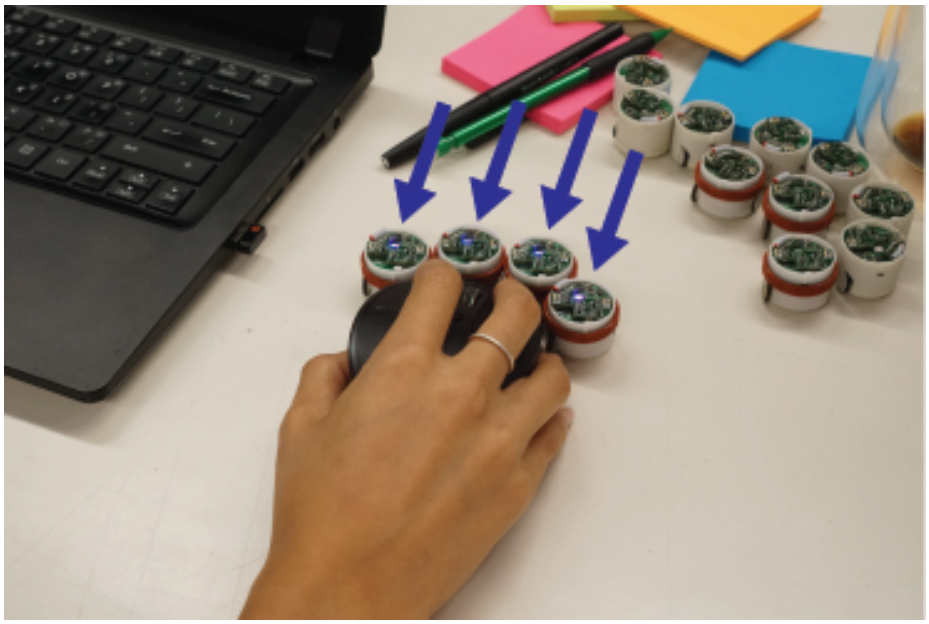


Figure 6.6: Directional Cue Scenario: the robots guide the user's mouse motion. Adapted from [2].

about the paper, she doesn't notice even though her phone is ringing. The robots then take action and approach her forearm to provide tactile notification through shear forces.

Directional Cues SwarmHaptics can be used to convey directional cues to the user. For instance, imagine Lauren is studying for final exams but is continuously tempted to watch entertaining videos instead. However, whenever she tries to move her cursor to click the link, the robots push her mouse away as shown in Fig. **??**. After a few failed attempts, she finally gives in and studies for her final.

Remote Social Touch SwarmHaptics can be used to convey social touch to a remote person. For instance, imagine that Joanne is trying to condole a friend who is having a tough time. Unfortunately, she cannot physically be there for him as they are far apart. As a way to provide comforting touch, she draws a sinusoidal wave on her phone to convey a soothing wavy haptic and visual stimuli with the robots.

6.5 Implementation

I used the the modified version of the Zooids robot as described in Chapter **??**. These robots communicate with a centralized computer and move in 2-D space via two, wheeled motors.

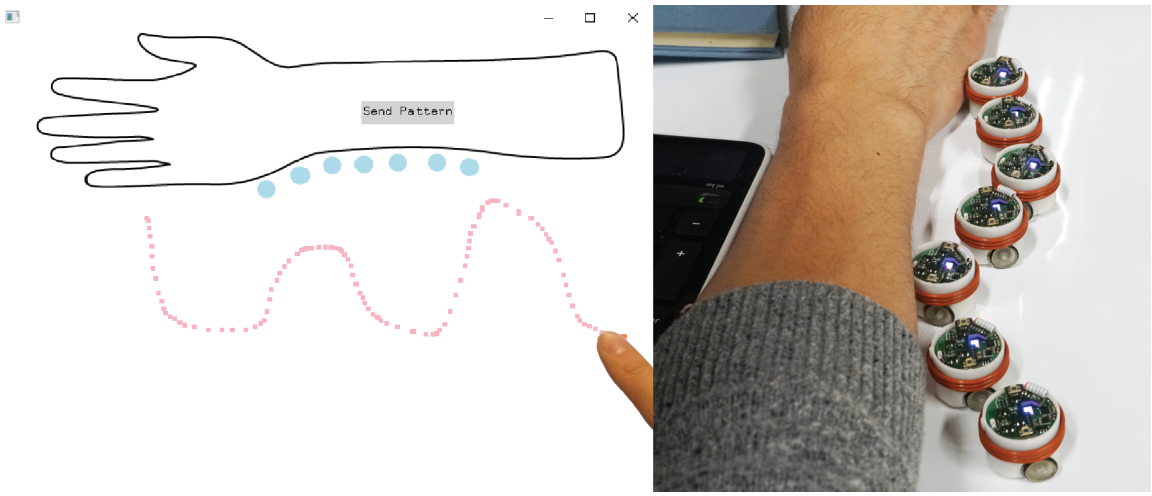


Figure 6.7: Remote Social Touch: Different haptic patterns can be drawn on a touch screen to convey remote social touch. Adapted from [2].

To track the robots, I used the same projector-based tracking [197] as in Zooids [5] or in Chapter 3.

There are several challenges with using swarm robots for a haptic display. For haptic stimuli, force amplitude is one of the most important parameters. At the minimum, it needs to be detectable by the users. To further expand the range of haptic stimuli possible, a larger range of force amplitudes is desired. The current Zooids do not output such forces, and even the modified version with stronger motors used for UbiSwarm fail to produce significantly greater forces due to the low friction between the wheels and the driving surface. To overcome this, I attached magnets to the bottoms of the robots and used a ferromagnetic surface to increase the normal force, and thus the wheel traction. This by itself increased the force output of the modified Zooids robots by a factor of 9 from 0.1 N per robot on non-ferromagnetic surfaces to approximately 0.92 N per robot on ferromagnetic surfaces as shown in Figure 6.8. The same could be achieved by increasing the mass of the robot by a factor of 9 or by using gecko-adhesives/microspines as used in [13, 198]. For this Chapter, I decided to use the ferromagnetic surface for quicker implementation and to keep the volume of the robots to a minimum allowing more robots to simultaneously interact with users' forearm. Due to cost and complexity of integrating a force sensor in a small robot, I didn't include a force sensor and the force outputs were provided in open-loop.

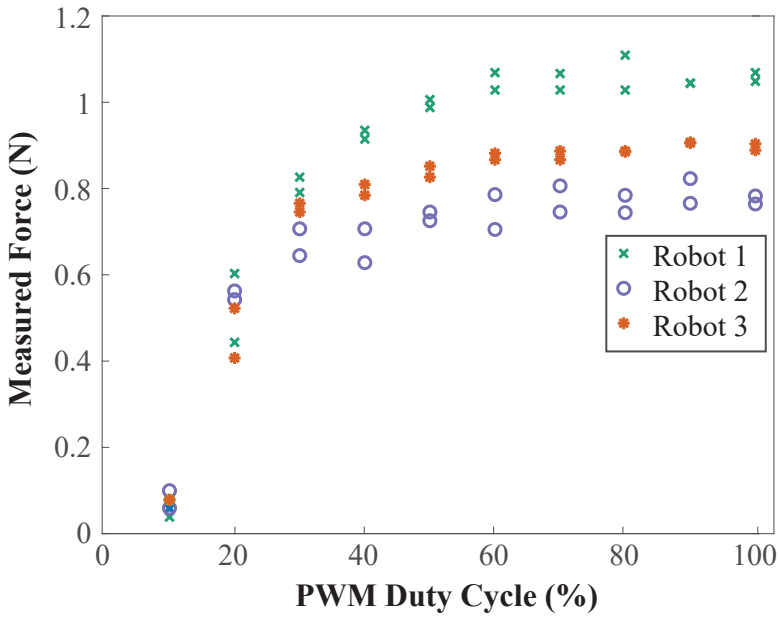


Figure 6.8: Normal force evaluation for three robots. Adapted from [2].

In addition to the friction between the robots and the ground, I need to consider the friction between the users and the robots as it determines the shear forces that users would feel. The current robots are 3D-printed plastic and thus doesn't have good traction with human skin. Studies have shown that softer materials are more pleasant than rough materials [194]. Thus, to increase friction without causing any discomfort, I added soft silicone rings around the robots.

I programmed the applications, motions, and haptic stimuli in C++ in Visual Studio using the Zooids API [5]. To track the participant's forearm, I used the position of a robot that is mounted on the participant's wrist through a wristband as shown in Figure 7.3. Based on the location of the wrist, I estimated the location of the forearm. For more details about the Zooids system, refer to [5].

6.6 User Perception of SwarmHaptics

To properly design the haptic patterns from SwarmHaptics, we first need to understand how people perceive different haptic stimuli from the robots. I first begin by studying the effects of the fundamental parameters such as force type, frequency, and amplitude with

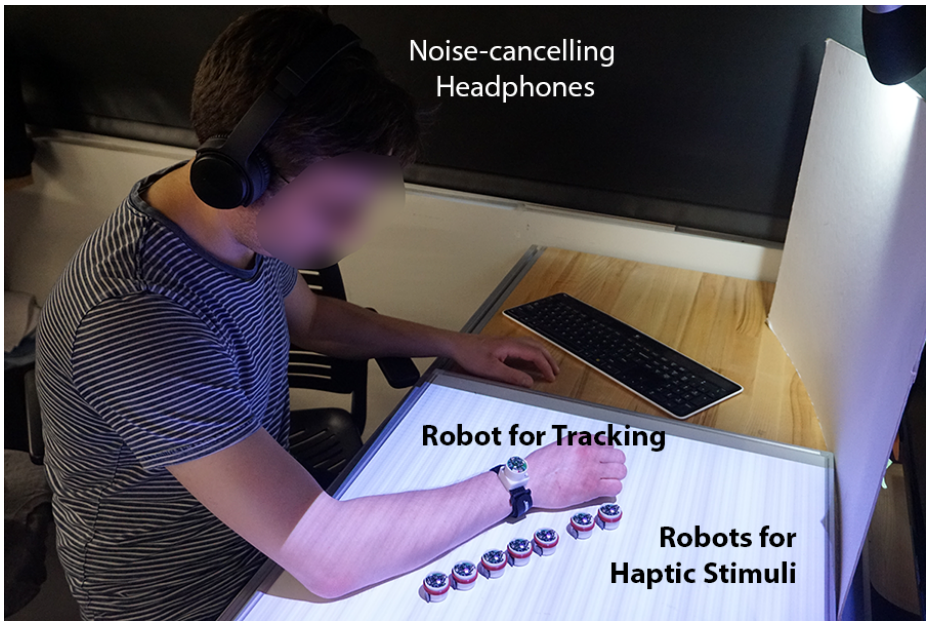


Figure 6.9: Perception study setup with robots for haptics/tracking and a user with noise-cancelling headphones. Adapted from [2].

varying number of robots. Other elements of our design space such as spatial, temporal, and force coordinations build on the fundamental parameters and are left to future work.

6.6.1 Hypotheses

The first parameter I studied was the number of robots touching the participant. No research to our knowledge has explored and tested the idea of multiple robots touching a human. I expect similarities with what Podevijn et al. has found for people observing motion of different number of robots [29]. They found increasing the number of robots in motion increased the subjects' heart rate, skin conductance level, arousal, and valence [29], I hypothesize that increasing the number of robots in contact will also increase the perceived arousal, urgency, valence, and likability.

Based on pilot testing, there seems to be significant effects from the force type. In particular, I conjecture that shear forces are more pleasant and likable than normal forces. Thus, I hypothesize that shear forces will be rated higher in valence and likability than normal forces.

Frequency is an important haptic parameter for both detectability [199] and user perception. However, the trend for perception has been unclear in prior works for vibration [200, 201] and mid-air haptics [202]. Based on the commonalities of the prior works, I hypothesize that higher frequency will elicit higher perception of arousal and urgency.

For haptic devices, controlling the force amplitude is critical not only to overcome the absolute threshold but also to mediate the users' perception. For instance, with both belt-compression and electrovibration haptic devices, higher amplitude for force or vibration has been rated lower in valence [203, 204]. Thus, I also hypothesize that the higher force amplitude will lead to lower valence and likability.

6.6.2 Method

To evaluate the human perception of haptic display with swarm robots, I provided various haptic stimuli to the users, on the dorsal side of their forearm. We chose it for three reasons. First, it is one of the more socially accepted areas for other people to make contact [174]. People typically rest their forearm on a flat table which is one of the ideal locations for the robots. Lastly, the forearm provides ample room for a swarm of robots to provide direct haptic sensations.

Independent Variables

I varied four independent variables: number of robots, force type, frequency, and amplitude. To limit the total experiment time to less than 50 min, we only used one repetition for a total of 24 trials per participant.

Number of Robots: For the study, I explored three values for the number of robots: n = 1, 3, and 7. The maximum number was limited to seven as only seven robots could touch a user's arm simultaneously in a parallel configuration.

Force Type: I also looked at the effect of different force types: normal and shear. Normal forces are generated by applying the same torque to both motors in the same direction while for shear forces, the directions are reversed.

Frequency: Frequency is an important parameter of haptic stimuli for both detectability and user perception. For instance, for vibration, the amplitude absolute threshold changes with the frequency [199]. Binary values were used for the frequency: 1 Hz or 10 Hz. 10 Hz was the highest that the robots could render without reducing the force amplitude significantly.

Amplitude: For haptic devices, controlling the force amplitude is critical not only to overcome the absolute threshold but also to mediate the users' perception. I explored the effect of binary values of the amplitude: a low value (0.8 N per robot) that we felt were just detectable and a high value (0.92 N per robot) that is the maximum amplitude possible with the current robots.

Dependent Variables

Emotion: For any experience, it is important to account for user's perceived emotion as emotion influences physiological, cognitive, and behavioral states of the users. Thus, I studied the effect of different haptic stimuli on users' affect. To measure, I used a seven-point scale of SAM [205], a visual scale of parameters in the PAD model [206]: valence, arousal, and dominance. Due to its use of pictures, SAM is a widely used to assess emotion in both user experience and HRI research across different regions.

Measures for Human-Robot Interaction (HRI): Many HRI researchers have used the questionnaire designed by Bartneck et al. that is specific to measuring perception of robots [207]. Out of the five categories of the questionnaire, I asked the participants to rate seven-point semantic differential scales on the three most relevant ones: anthropomorphism, likeability, and perceived safety. We included anthropomorphism as generating human-like touch will be meaningful and useful especially in the context of social touch. I excluded perceived intelligence and animacy as a pilot study showed these two did not vary with different haptic stimuli.

Urgency: Lastly, I envision that SwarmHaptics can be used to notify people of events with varying urgency. I adopted the method used in [208] to measure urgency. Through a seven-point semantic differential, I asked the participants to rate their perceived urgency of the haptic stimuli and their intention to either dismiss or attend to them.

Participants

Twelve participants (5 M, 7 W, Age: 21-29) were recruited. Participants had various previous haptic experiences ranging from none to extensive. None had neurological disorders, injuries to the hand/arm, or any other conditions that may have affected their performance in this experiment. They were compensated $15 for their time ($\sim$ 40 min) and the study was approved by the University's Institutional Review Board with subjects providing informed consent.

Procedure

Before the study, I informed the participants that they would be given various touch stimuli from the robots and would be asked to rate their perception. To track their arm, they were asked to wear a tracking wristband. They also wore a noise-canceling headphone to isolate the audio cues from the robots. For each trial, participants placed their arm on a designated location. Once ready, they pressed a button to start. The robots initially positioned 10 cm away from the arm, moved forward and made contact with their arm. After a second, the robots would provide the touch stimulus (500 ms) three times with a 500 ms break in-between for a total of 3 seconds. Once completed, they would move back to their initial positions. Participants would then complete a survey on a tablet and repeat for a total of 27 fully randomized trials (3 training trials + 24 conditions).

Analysis

To examine the effects of the four independent variables including interaction, a Mauchly's Test of Sphericity and a 4-way repeated measures ANOVA were performed for each dependent variable. If Mauchly's Test of Sphericity was violated, I used a Greenhouse-Geisser correction for F and p values from ANOVA indicated by F* and p*. If any independent variable or combinations had statistically significant effects ($p < 0.05$), Bonferroni-corrected post-hoc tests were used to determine which pairs were significantly different.

6.6.3 Results

Figures 6.10-6.12 report the means of all dependent variables for each haptic parameter along with their standard errors. ($* : 0.01 < p < 0.05, ** : 0.001 < p < 0.01, *** : p < 0.001$)

Emotion

All independent variables except force type had a significant effect on at least one parameter of emotion, as shown in Figure 6.10. For the arousal axis, number of robots (F(2,22)=69.4, p=3.1E-10) and force amplitude (F(1,11)=96.0, p=9E-7) had positive relation whereas they both had negative correlation with the perceived valence (F(2,22)=4.64, p=.021), (F(1,11)=26.6, p=.0003). Finally, force frequency (F(1,11)=14.2, p=.022) had positive correlation with the valence axis.

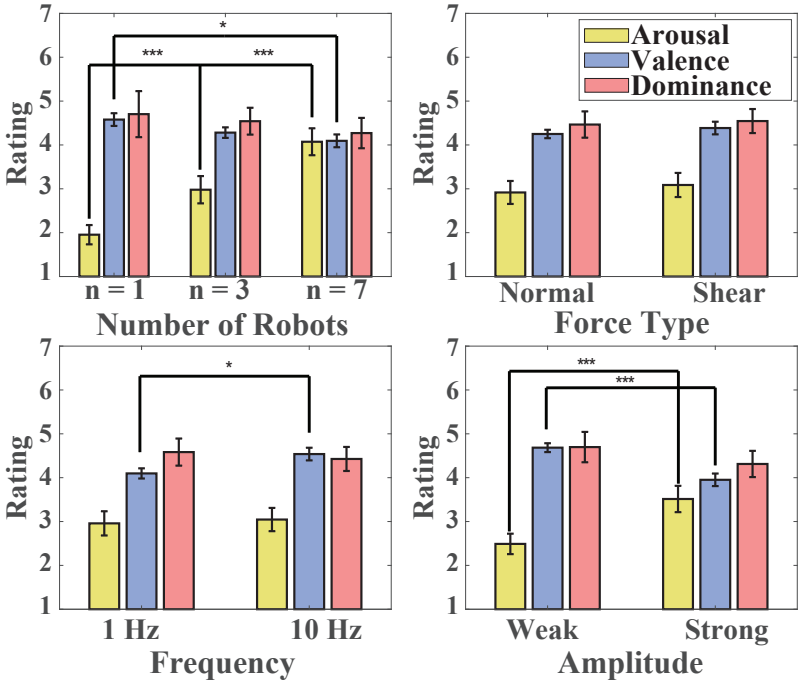


Figure 6.10: Effect of haptic parameters on emotion. Adapted from [2]. ($* : 0.01 < p < 0.05, ** : 0.001 < p < 0.01, *** : p < 0.001$)

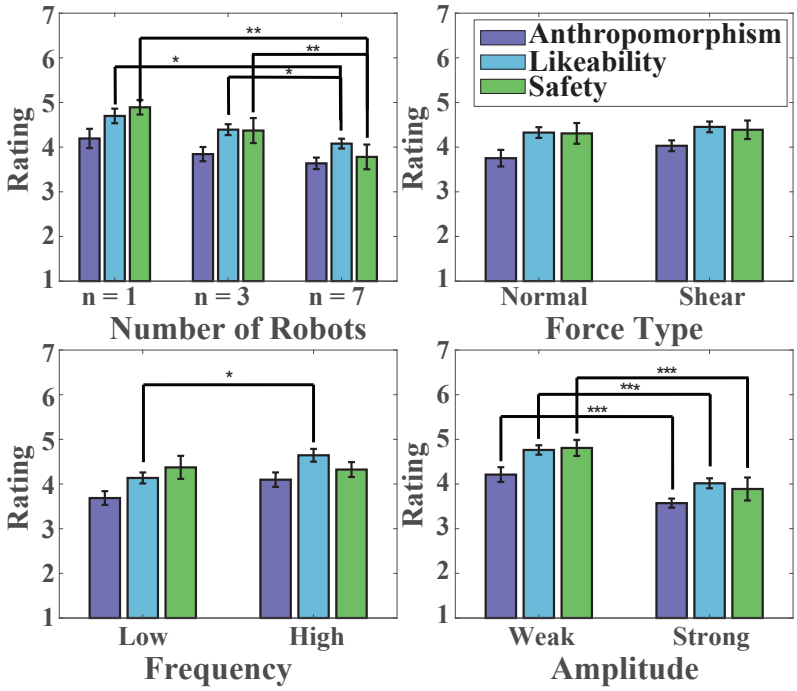


Figure 6.11: Effect of haptic parameters on HRI metrics. Adapted from [2]. ($* : 0.01 < p < 0.05, ** : 0.001 < p < 0.01, *** : p < 0.001$)

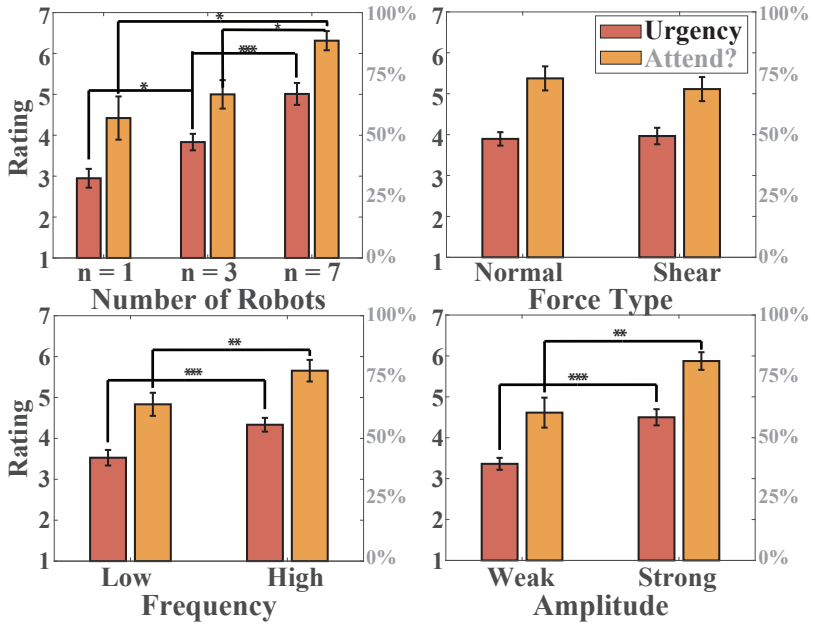


Figure 6.12: Effect of haptic parameters on perceived urgency and willingness to attend. Adapted from [2]. ($* : 0.01 < p < 0.05, ** : 0.001 < p < 0.01, *** : p < 0.001$)

HRI

All independent variables except force type had significant effect on at least one of the three HRI categories explored as shown in Figure 6.11. Number of robots and force amplitude had negative correlation with perceived anthropomorphism(F(2,22)=3.6, p=.044), (F(1,11)=19.3, p=3.5E-4), likeability (F(2,22)=6.8, p=.014), (F(1,11)=43.7, p=3.8E-5), and safety (F*(2,22)=11.2, p*=.002), (F(1,11)=36.1, p=9E-5). On the other hand, force frequency had positive correlation with perceived likeability (F(1,11)=7.0, p=.023).

For likeability (F(2,22)=5.8, p=.009) and perceived safety (F(2,22)=7.4, p=.003), there was an interaction effect among number of robots, force type, and frequency. While higher frequency usually increases likability, when there is one robot applying normal force, the frequency doesn't affect the likability. For perceived safety, when there is one robot generating normal force, lower frequency stimulus is perceived significantly safer than higher frequency one.

Urgency

All independent variables except force type had significant effect on urgency and willingness to attend as shown in Figure 6.12. Number of robots [($F^*(2,22)=24.4$, $p^*=7.2E-5$), ($F(2,22)=7.4$, $p=.004$)], force frequency [($F(1,11) = 25.9$, $p=3.5E-4$), ($F(1,11)=13.3$, $p=.004$)], and force amplitude [($F(1,11) = 69.9$, $p=4E-6$), ($F(1,11) = 16.0$, $p=0.002$)] all had positive correlations with perceived urgency and willingness to attend.

There was an interaction effect for urgency ($F(1,11)=7.2$, $p=.021$) and willingness to attend ($F(1,11)=7.3$, $p=.021$) between force type and force amplitude. While stronger amplitude is generally perceived as more urgent and has higher willingness to attend, when the robot(s) are applying shear forces, the amplitude doesn't affect the perceived urgency or the willingness to attend to the robots.

6.6.4 Discussion

Number of Robots

From the results, we can easily see the significant effects that the number of robots has on user perception. The number of robots had positive correlation with arousal, urgency, and willingness to attend while having negative correlation with valence, likability, and perceived safety. These results imply that when deciding the number of robots for haptic display, there is a tradeoff between the perceived arousal or urgency and the pleasantness or likeability of the haptic stimuli. More robots can provide more arousing and urgent sensation but at the cost of pleasantness, safety, and likeablility. Thus, I would recommend limiting the number of robots used for haptic display when conveying positive, safe or pleasant information whereas using more robots for important and urgent circumstances.

Force Type

Force type had surprisingly weak effect on user perception. There was no dependent variable which force type alone had statistically significant effect. However, interaction effects were observed for likeability, perceived safety, urgency, and willingness to attend. The results suggest that force type has complicated relationship with human perception and thus will need to be carefully combined with other parameters to elicit desired effect.

Frequency

Frequency of the haptic stimuli had positive correlation on valence, likeability, urgency, and willingness to attend. This is intriguing as some of the results correlate with previous work while some provides clarification. Specifically, the trend for perceived urgency and willingness to attend is aligned with what others found, in that higher frequency is more alarming, and arousing [200–202]. On the other hand, the results for valence and likeability provides some clarification as other works had found mixed results. However, as the values of frequency tested here (1-10Hz) are drastically different than the ones used in previous work for vibrations and ultrasonic transducers (16-175Hz), more studies are needed for further clarification.

Amplitude

Increasing the amplitude of the haptic stimuli increased the arousal, urgency, and willingness to attend but decreased valence, anthropomorphism, likeability, and perceived safety, similar to the effect of increasing the number of robots. This result is consistent with what Valenza et al. found in which higher force amplitude led to higher arousal and lower valence [204] and with Bau et al. which found higher amplitude for electrovibration was rated less pleasant [203]. Along with the number of robots, force amplitude was found to be the more influential parameter and thus will need to carefully controlled to elicit desired perception.

6.7 Elicitation Study for Social Touch

In the earlier study, I evaluated human perception of various simple haptic patterns. To generate more expressive patterns specifically for social touch with different spatial, temporal, and force coordinations, we had the participants brainstorm haptic patterns through an elicitation study. Elicitation studies with novice users have been shown to be beneficial in terms of understanding users' preferences [131–133].

6.7.1 Method

As SwarmHaptics is a novel system, I wanted to ensure there was a feedback loop in which participants could feel the haptic patterns they created and modify them as needed. Thus,

I had participants use their own non-dominant arm for feedback and the dominant arm to generate the pattern.

Referents

There exists a wide range of social touch that varies in its intent and affect. Prior work has identified six categories for symbolic meanings of touch: positive affection, control, ritualistic, playful, and task-related [209], while there are six widely accepted basic emotions (happiness, sadness, fear, anger, disgust and surprise) [210]. Based on these two concepts, I generated a list of referents, shown in Table 6.1.

Data Collection

For analysis, I videotaped participants' movements and recorded positions/velocities of the robots and touch. To understand participants' perception, I asked them to complete a questionnaire after each trial rating the clarity, ease of control, anthropomorphism, and likability of the haptic pattern they just created. I included anthropomorphism as generating human-like touch is important and relevant especially for social touch. They also wrote a textual description what they were trying to do with the robots. After the trials, they filled out a questionnaire and had short interviews about the overall experience of using SwarmHaptics for remote social touch.

Implementation

For the study, I used a set of four robots controlled by the user's finger movements on a multi-touch screen as shown in Fig 6.13a. I initially started with five robots but preliminary testing revealed that controlling five fingers independently without losing contact with the screen is uncomfortable and do not increase the expressiveness substantially. Users control the torque of the robots by dragging the blue circles with their fingers. The distance and direction of the drag controlled the torque and heading of the robot (Fig 6.13b). The sizes of the control points and spatial map were adjusted for each user to accommodate their hand sizes.

As our previous study demonstrated the importance of force amplitude, I redesigned the robots to increase force output (up to 3.6 N compared to 0.92 N) with a larger size (45 mm diameter) as shown in Fig. 6.13c. The existing motors are replaced with higher torque

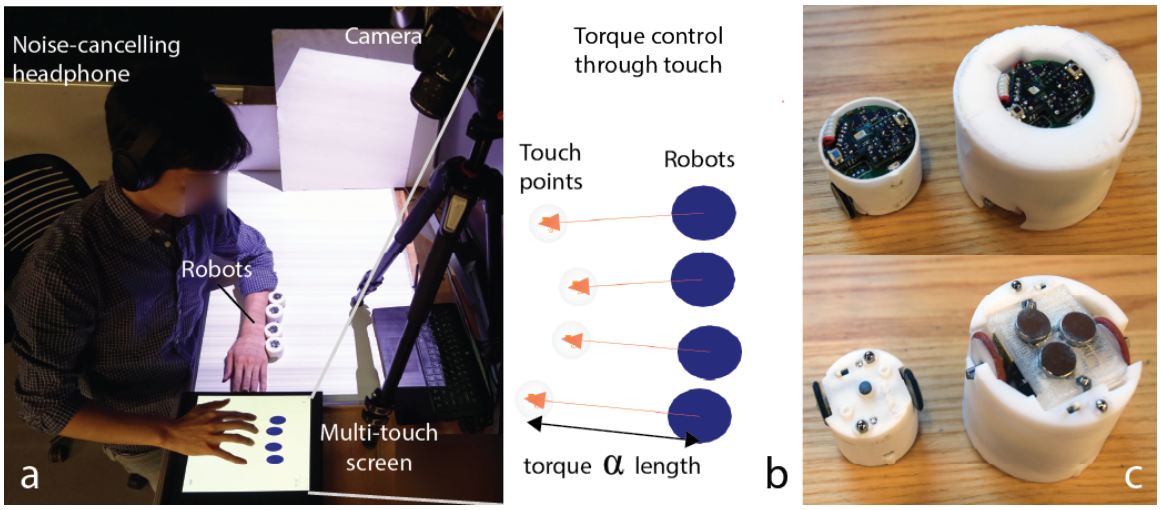


Figure 6.13: Elicitation study setup (a) with a close-up view of the multi-touch screen (b). (c) The robots for the elicitation study (right) compared to the one used in first study (left). Adapted from [2].

motor (Pololu Micro Metal Gearmotor #2365) and additional battery (850 mAh, 1C) is added to power them. Stronger magnets are added on the bottom to increase the normal force, thus the traction.

Procedure

Twelve participants (6M/6W, age: 20-33) were recruited. None had injuries, fatigue, or disorders that could impact their performance. They were compensated $15 for their time (∼45 min) and the experiment was approved by the University's Institutional Review Board.

Before the study, participants were informed that they would be given a set of referents for each of which they would create a haptic pattern using the robots and answer a questionnaire. After the instructions, three minutes were allotted for to familiarization with the interface and exploration of the range of possible haptic patterns. Then, for each referent, the participants were told to try out different patterns on themselves, decide on the best one, and record it. Participants were told to use as many robots as necessary and wore noise-canceling headphone to isolate audio cues.

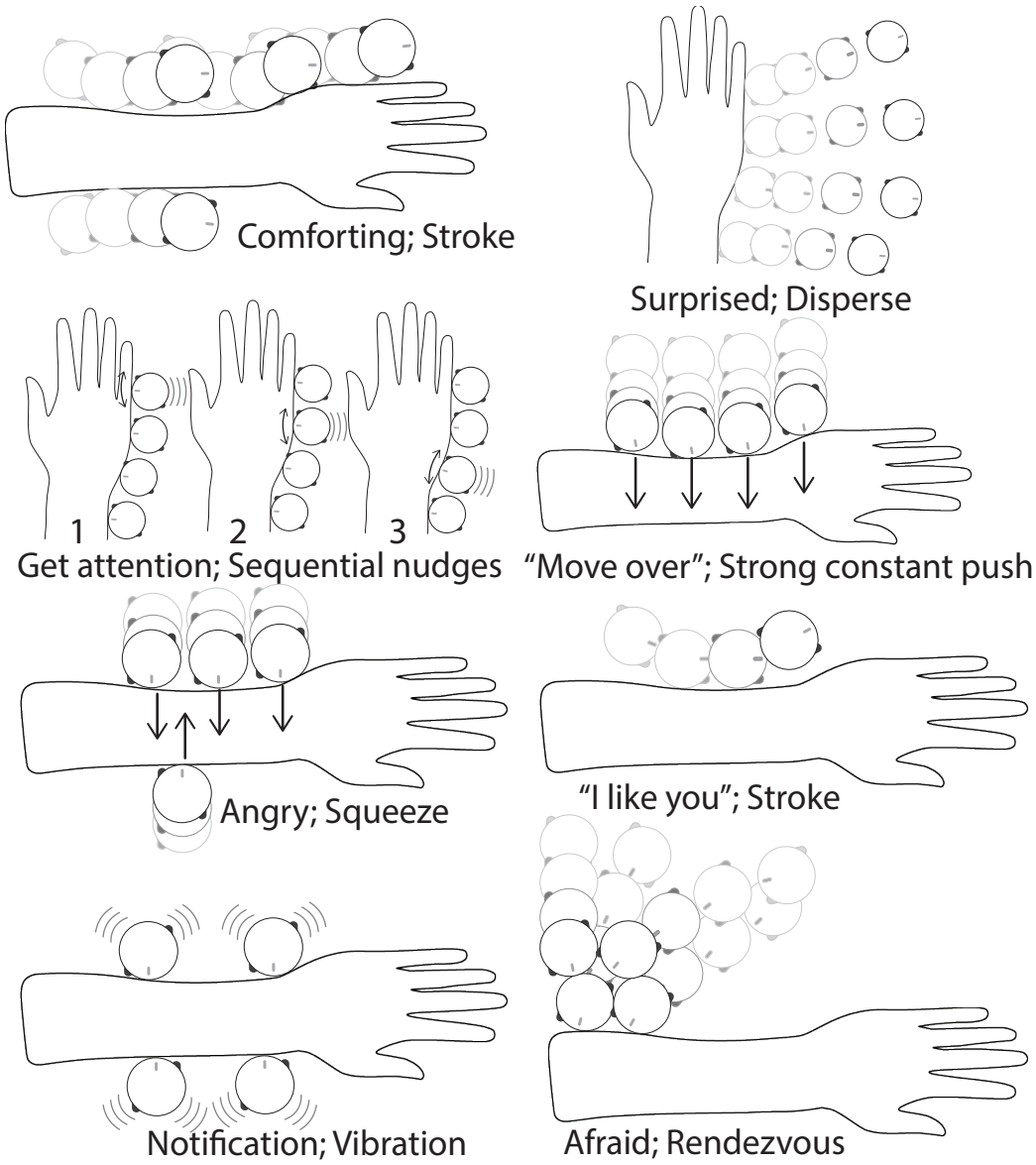


Figure 6.14: Example interactions from the elicitation study for social touch. A non-conflicting set for ten referents which may not be the most representative one is shown. Adapted from [2].

Table 6.1: Referents used for the social touch elicitation study and their corresponding interactions. Adapted from [2].

Category	Referent (Agreement)	Interaction	# robots	#
Positive Affection	"I like you" (0.39)	Stroke	1	5
		Hug	2+	3
		Touch & draw heart	1	2
	Comforting (0.31)	Stroke on both sides	2+	5
		Stroke	1	3
		Hug	2+	2
Control	"Move over" (0.25)	Strong constant push	2+	5
		Sequential push	2+	2
		Tap	1	2
	Get attention of someone (0.13)	Sequential nudges	2+	2
		Hard collide	1	2
		Sporadic taps around arm	1	2
Ritualistic	"Hello" (0.32)	Tap	1	5
		Tap	2+	2
		Fist Bump	1	2
Task-related	Notification (0.22)	Tap repeatedly	1	4
		Vibration	2+	3
		Strong Push	2+	2
Emotion	Afraid (0.13)	Nudge for long period	1	2
		Strong push back & forth	2+	2
		Touch & rendezvous	2+	2
	Angry (0.25)	Strong push	2+	5
		Squeeze	2+	2
		Strong push back & forth	2+	2
	Disgusted (0.24)	Tap & run away	2+	5
		Poke repeatedly	1	2
	Sad (0.14)	Gentle push	1	3
		Slow Stroke	1	2
	Surprised (0.17)	Strong push back & forth	2+	3
		Disperse	2+	3
	Happy (0.15)	Move/jump back & forth	2+	3
		Touch & draw smiley face	1	2
		Dancing	2+	2

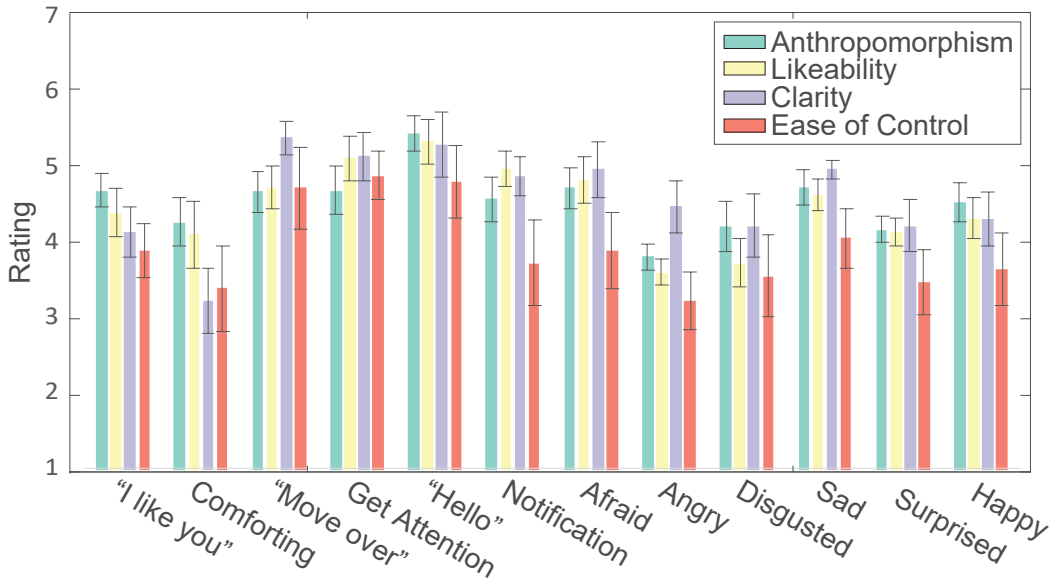


Figure 6.15: Ratings on anthropomorphism, likability, clarity, and ease of control of the robots for each referent. Error bars represent the standard errors. Adapted from [2].

6.7.2 Results & Discussions

In Table 6.1, the interactions with the number of robots used for each referent are listed. # indicates the frequency of each interaction. Only interactions with frequency greater than 2 are shown due to space constraints. Some of the example interactions are shown in Fig. 6.14.

Agreement

For each referent, I calculate the agreement score [132] as in Table. 6.1. It's interesting to note that having higher self-reported clarity rating doesn't necessarily translate to higher agreement score (Spearman's correlation p-value= 0.29). For instance, referents such as sad, afraid, and get attention of someone have high clarity ratings (> 5) but have agreement scores on the lower half (< 0.15). On the other hand, the referent with lowest clarity rating (comforting) had one of the highest agreement score.

In addition, Welch's t-test revealed that the self-reported clarity ratings of interactions that are members of larger groups of identical gestures ($\# >= 3$) are not significantly different ($p = 0.37$) from those that are not. In fact, the average rating of the former interactions

was lower. This is different from what Wobbrock et al. [132] found for surface computing gestures possibly due to the fact that we are investigating interactions for affective communication, which can be very context dependent [211] rather than to achieve a specific task.

Visual Complement

One interesting trend is that many interactions relied on visual components with varying degrees. For instance, participants partially relied on the paths or motions of the robots to convey more context especially for abstract referents such as afraid (touch & rendezvous) and surprised (disperse). Others relied mostly on the visual aspects by drawing heart ("I like you") or dancing (Happy). As swarm robots inherently provide both visual and haptic cues, more studies should be done to investigate the trade offs between them and which is better for different applications.

Contact Location

Another unexpected feature was the use of contact locations. For some interactions, participants used contact location to provide more context. For instance, to convey "Hello", two participants made a robot bump toward their fist to create a "fist bump". For the "afraid" referent, one participant had a robot "hide" under the arm while another had a robot nudge between the thumb and index finger.

Metaphors

From the post-study interview and qualitative feedback after each referent, I have gathered multiple metaphors that the participants used for the robots. They mentioned that depending on the referent, they pictured the robots as being either extensions of their hand, minion/pet/living creature that delivers their message, or parts of an emoji. Though I did not measure social appropriateness, these positive metaphors suggest that people were comfortable interacting and using the robots for social touch.

Trends between Ratings of Referents

In Fig. 6.15, average ratings of perceived anthropomorphism, likability, clarity of the message, and ease of control are shown for each referent. A Welch's t-test reveals that the

self-reported clarity for functional referents (i.e., control, ritualistic, and task-related referents) is significantly higher than those for affective ones (i.e., positive affection and emotion referents) (p=.0032). This is consistent with the qualitative feedback from participants as many of them spoke about the difficulty of creating haptic patterns for abstract and emotional referents. Thus, they sometimes had to rely on the visual aspects of the robots' motions to convey the appropriate context.

6.8 Limitations & Future Work

There are a number of considerations with regards to the generalizability of the study results. First, the studies can only be generalized to systems comparable to this particular swarm robot platform. With systems of drastically different parameters such as size, form factor, force outputs, and contact material, the current study results may not hold. In the future, it would be interesting to further investigate the effect of different robot sizes and a wider variety of form factors.

In addition, the trends demonstrated with the current study results may not hold with different values of the parameters. The goal of the current study was to investigate effects of different parameters and their combinations. Thus, we only tested 2 or 3 values of each parameter. With values much greater or smaller, the trends may take unexpected turns. In the future, I suggest studies with more values of each parameter to better understand their impact on users' perception.

For the elicitation study, many participants had difficulty controlling 3+ robots simultaneously. This is partially due to the multi-touch screen that requires a constant contact for each finger and the nonholonomic nature of the robots. To alleviate this, a 3D gesture could be adopted using hand tracking sensors like Leap Motion. Also, omni-directional robots could help users express a wider range of messages.

A more fundamental limitation for SwarmHaptics is that it can only provide haptic stimuli to body parts that are in close proximity to flat surfaces. As the robots can only move in 2-D plane, they cannot provide haptic stimuli to multiple body parts that are arbitrarily separated in the 3-D space.

In addition unlike visual displays, there are visible motions that always accompany the robots when moving from point A to B. This could be both beneficial or undesirable depending on the context. As supported by the elicitation study, it could serve as a multimodal

display and provide more context or unintentionally distract users. Thus, designers will need to take this into account and either minimize this undesirable effect by slowing down the robots or take advantage of it to communicate both visually and haptically.

Finally, the example scenarios shown in this Chapter are limited to one-directional haptic display such as haptic notifications and social touch. In the future, I would like to study the use of swarm robots to provide real-time multi-point haptic feedback for interactive applications as prior work has mainly focused on using robots to provide single point feedback or as grasped puck-like devices [96].

6.9 Chapter Summary

In this chapter, I investigated how a swarm of robots could provide richer information to users in the context of haptic notifications, directional cues, and remote social touch. To do so, I introduced SwarmHaptics, a new type of visual and haptic display consisting of small mobile robots that can approach the users at will to provide different haptic patterns. I described the design space for SwarmHaptics with parameters for both each individual robots and collections of them, and demonstrated its possibilities with example scenarios. I evaluated how people perceive different haptic patterns and found that the number of robots involved is important as it increases urgency and arousal but at the sacrifice of likeability and perceived safety. A separate elicitation study on remote social touch revealed many interesting trends. For instance, many participants relied on the visual patterns that the robots create to convey more abstract social touch as haptics alone was deemed insufficient to convey comprehensible meaning. I hope that this study will spur interests and aid researchers in further exploring a haptic display with a swarm of robots. While this study looked at utilizing touch from the robots, in the next chapter I consider how people naturally command a large number of robots.

Chapter 7

User-defined Swarm Robot Control

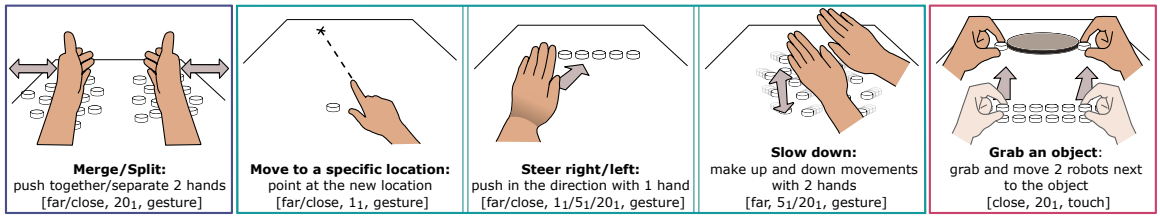


Figure 7.1: Elicitation study to understand how users control a swarm of robots. Examples of interactions with high agreement scores are shown here. People used varying numbers of fingers/hands and different interaction modalities such as gesture and touch. The first two values inside the brackets indicate the proximity and number of robots for the interaction and the last value indicates the interaction modality. The colored boxes indicate the task type that it belongs to. Blue, teal, and red boxes represent inter-robot interaction, navigation, and object manipulation task types, respectively. Adapted from [3].

Multimodal display methods from the robots to users were investigated in the Chapters 4-6. Abstract swarm robot motion was used to visually convey affect and intent, while haptic display from the robots was studied through perception and elicitation studies. To complete the loop, I investigate in this chapter how users naturally control and provide multimodal inputs to a swarm of robots.

- ©2020 ACM. Reprinted with permission, from Kim, Lawrence H, and Follmer, Sean. "SwarmHaptics: Haptic Display with Swarm Robots" Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. ACM, 2020. doi:10.1145/3290605.3300918 [3].

7.1 Introduction

Robots are increasingly being deployed across personal, commercial, and industrial sectors, with application spaces ranging from elderly-care social assistants to members of firefighting teams. We are moving towards a society where humans are actually outnumbered by autonomous and semi-autonomous agents in both their home and work lives, similar to the vision of "ubiquitous robotic interfaces" described in [1] and in Chapter 3. Some of these robots will work together in small groups, typically thought of as "multi-agent systems." For applications where things like areal distribution, low unit cost, and robustness to agent failure are critical, research has begun towards the development of swarm systems, where, as in the natural world of insects, large (>10) groups of robots must work together to become more than the sum of their parts [212]. This emerging field of swarm robotics presents many challenges in the area of human-swarm interaction (HSI), including the cognitive complexity of solving tasks with swarm systems, state estimation and visualization, and human control of the swarm [213].

While HSI researchers have developed numerous ways to control a swarm of robots in situ [30, 31, 94], they all share one important limitation: lack of consideration for user's preferences and intuition. Instead of integrating sensors that can sense a set of user-defined interaction, prior work has mostly focused on finding a set of interactions that the existing robotic sensors can detect, and then semi-arbitrarily mapping these interactions to a set of control commands. This is a problem, as such interaction vocabularies may only be effective for domain experts or the designers themselves and could present a steep learning curve for novice users. As I see a near future with wider adoption of swarm robot technologies that will constantly exist in both our public and private environments, I focus on proximal control that could function on an encountered basis even for novice users. Thus, I seek to ground HSI through user-centric approaches. While prior works have studied interaction with a single robot and ran elicitation studies on control of a team of drones [135,136], it is unclear how those results map to grounded, large agent count multi-robot systems.

To better understand how users *prefer* to interact with swarm robots, I present an elicitation study with up to 20 centimeter-scale tabletop robots. As prior work has shown that number of robots and proximity to the robots affect human's perception and behavior [214,215], I also investigate the effects of these variables on user's desired input method.

The tasks ranged a large span of possible interactions, including formation control, parameter setting, cooperative manipulation, and human-robot teaming concepts (e.g., "follow me"). Care was taken to abstract implementation details from the users in order to elicit multimodal input schemes which include gestures, touch and verbal interactions. Using the study results, I compile a user-defined interaction set with interactions based on not only referents but also number of robots and proximity. I also examine overall trends on interaction modality, taxonometric breakdown, and agreement scores to better understand how participant interact. These results can inform the design and sensing required to support rich interaction with swarm robots.

In summary, the contributions of this chapter are:

- Compilation of control commands for swarm robots,
- Taxonomy for classifying gesture, touch, and verbal interactions,
- User-defined interaction set for controlling a swarm of robots, and
- Study results on the effects of number of robots and proximity to the robots on how users interact.

7.2 Elicitation Study on In Situ Swarm Robot Control

To better understand how users prefer to interact with a swarm of robots, I conducted an elicitation study on swarm robot control. The study results can inform what types of sensors are needed to enable fluid interaction between users and a swarm of robots. To better support open science, I have pre-registered this elicitation study at OSF.io (https://osf.io/r8fnc) and all raw data along with study results are freely available at https://osf.io/dkja9/ as well.

7.2.1 Hypotheses

In addition to understanding how users interact with a swarm of robots, I investigate the effects of a few key interaction parameters: number of robots and proximity to the robot(s).

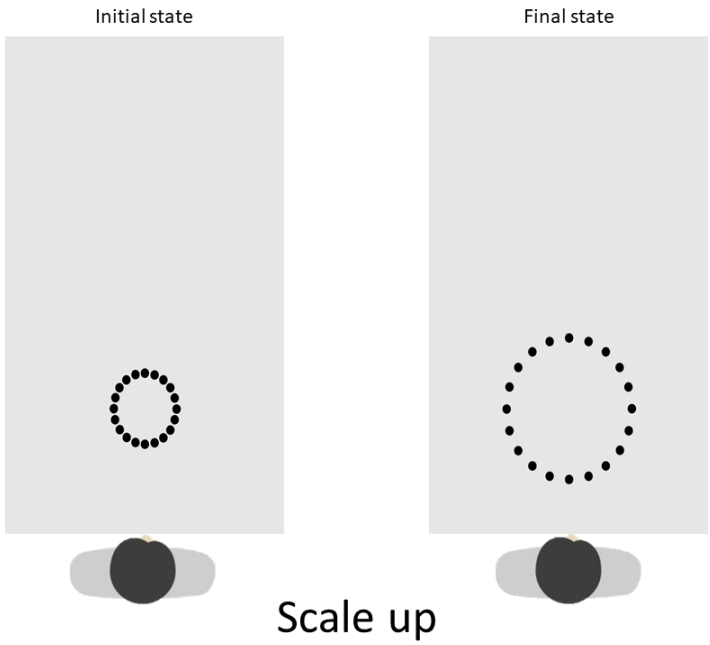


Figure 7.2: Sample pictorial prompt for the "Scale up" referent. Adapted from [3].

H1: Number of robots will affect how users interact

Researchers have shown that the number of robots can significantly alter how people perceive the robots when viewing their motion [29] or being touched by them [2]. Researchers have also developed different ways to teleoperate or remotely control a swarm of agents such as leader-follower [216], selection and beacon control [126], and physicomimetics [217]. Thus, I hypothesize that users will also adapt their interaction method for in situ control based on the number of robots.

H2: Proximity to the robot(s) will affect how users interact

Literature in Human-Robot Interaction (HRI) has shown that humans perceive robots differently based on their proximity as well as prefer robots that exhibit proxemic behavior [214, 215]. Cauchard et al. have reported that when the robots were closer, users tended to use smaller motions [49]. Thus, I also hypothesize that proximity to the robot(s) will change how users choose to interact with a swarm of robots.

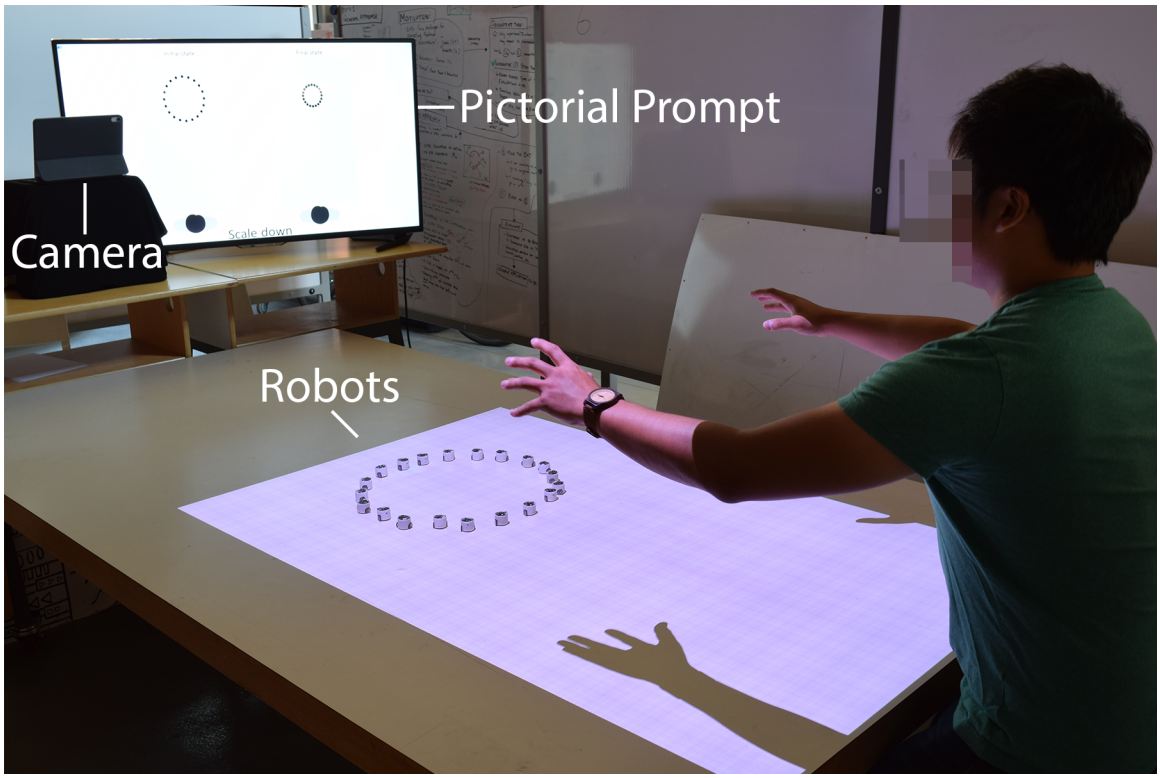


Figure 7.3: Setup for the control elicitation study: After being prompted through a television monitor, participants interact with 20 robots on a table while standing. Adapted from [3].

7.2.2 Methodology

I utilized a similar method as in [49] with slight modifications to address the real-time controllability of the robots and to improve accessibility for non-native English speakers. Instead of a complete Wizard-of-Oz (WoZ) elicitation study, I conducted a semi-WoZ study due to the difficulty of controlling a large number of robots impromptu. I pre-programmed each referent and timed the initiation once the participants completed their interaction. As shown in Fig. 7.2, I displayed pictorial [134] instead of purely textual [49] prompts for the referents as they reduce verbal biasing as well as lower language barriers for non-native English speakers. These prompts include initial and final state of the robots as well as the task title.

7.2.3 Apparatus

I used a modified version of Zooids, a wheeled multi-robot platform [5], as described in Chapter 3. As shown in Fig. 7.3, the pictorial prompts for the referents were displayed on a 50 inch television monitor while a video camera was used to record participants' interaction. Depending on the referent, up to 20 robots moved within a 1.4 x 0.89 m work space (i.e., projected space) on a table.

7.2.4 Referents

To generate the list of referents for this study, I combine the control and interaction commands from prior literature in swarm or multi-robot control [30,31,94,127,129,130,136,219, 220] and interaction with a single or multiple drones [49,64,134,135,218] as shown in Table 7.1. For referents under the "Robot Selection" and "Inter-Robot Interaction" categories as well as "Move here and there" and "Grab an object" referents, only 20 robots were used since these referents are most relevant when there are a significant number of robots. To reduce study duration and user fatigue, I combine pairs of referents from prior works that were similar and opposite of each other such as "move closer/away", "steer left/right", and "move faster/slower". For these pairs, each pair instead of each referent was presented under all 6 conditions (3 (# of robots) x 2 (proximity)).

7.2.5 Participants

15 participants were recruited (7 M, 8 F) from my institution. Age ranged from 19 to 41 (average: 29, std: 5.7). Their educational backgrounds ranged from engineering (9), computer science (2), and others (4). They were compensated $15.

7.2.6 Procedure

For each referent displayed on the screen, participants were instructed to perform any interaction method that they choose to complete the given task. They were told to focus on how they would prefer to interact as opposed to focusing on whether the robot(s) could understand their interaction. No suggestions were given on how to interact with the robot(s). After the participants completed each interaction, they explained their interaction in 1-2 sentences and rated their interaction on a 7-point Likert scale in terms of suitability (i.e., how well their interaction fit the task), simplicity (i.e., how simple their interaction was),

Table 7.1: List of referents used in the elicitation study. Adapted from [3].

Category — subcategory		**Referents [related work]**
Robot Selection		Select one robot
(20 robots)		[30, 31, 64, 94, 127, 129, 135]
		Select a group of robots
		[30, 31, 64, 94, 127, 129, 135]
		Select all robots
		[30, 31, 64, 94, 127, 129, 135]
Inter-Robot Interaction		Form a circle [127, 130, 135]
(20 robots)		Split/merge [30, 136]
		Scale up/down [94, 130, 135, 218]
		Rotate [94, 218]
User-Robot	Navigation	Move closer/away [49, 134]
Interaction		Follow me [49, 134]
(1, 5, 20 robots)	Attention	Get attention [49, 134]
Robot-Environment	Navigation	Move to a location [31, 136, 218]
Interaction		Move here and there
(1, 5, 20 robots)		(only with 20 robots)
		Steer left/right [30, 127, 136]
		Stop [30, 49, 127, 134]
		Move faster/slower [31]
		Follow trajectory [94]
	Manipulation	Grab an object
		(only with 20 robots) [219]
		Push an object [220]

and precision (i.e., how precise the interaction was). To become familiar with the process, I included 3 practice trials including one basic referent (move closer) and two referents that pilot subjects found more complex (follow me, steer right) in the beginning. They then proceeded to the actual experiment with 76 conditions in randomized order. After the

participants completed the entire study, they filled out a post-test survey and had a short interview regarding their experience.

7.3 Analysis

7.3.1 Taxonomy

To understand what types of gesture, touch, and verbal interactions were used, I analyze them using a modified version of the existing taxonomies in surface gesture [132], manipulation [221], and illocutionary acts [222] as shown in Table 7.2.

Gesture

For gesture, I labelled each interaction by the number of fingers/hands used (one or two fingers, one hand, both hands) and by the four dimensions (form, nature, binding, flow) from the taxonomy of surface gesture [132]. For the four dimensions of surface gesture [132], I modified categories within each dimension to better fit the context of this study. For the form dimension, I removed "one-point touch" and "one-point path" as the interaction space is not limited to a 2-D space. Instead, I added "deictic" in the nature dimension as well as "iconic" to better classify 3-D gestures. For the binding dimension, I removed "mixed-dependencies" as I didn't observe any corresponding interaction and added "user-centric" to better accommodate user-robot interactions.

Touch

For touch interactions, I classified each interaction by the number of fingers/hand used (one or two fingers, one hand, both hands), number of robots touched (one, few (2-4), many), number of robots touched simultaneously (one, few, many), control paradigm (leader-follower, follow crowd, control all) as well as using the "contact" part of the human manipulation taxonomy [221]. As I observed no within hand manipulation either non-prehensile or prehensile (NP, M, W or P, M, W), I excluded those categories.

Table 7.2: Taxonomy of Gesture, Touch, and Verbal Interactions. Adapted from [3].

GESTURE		
# of finger /hands	*1-2 fingers*	1-2 fingers are used.
	one hand	more than two fingers used.
	both hands	Both hands are used.
Form	*static pose*	Hand pose is held in one location.
	dynamic pose	Hand pose changes in one location.
	static pose & path	Hand pose is held as hand moves.
	dynamic pose & path	Hand pose changes as hand moves.
Nature	*deictic*	Indicative or pointing gesture.
	symbolic	Gesture visually depicts a symbol.
	physical	Gesture acts physically on objects.
	metaphoric	Gesture indicates a metaphor.
	abstract	Gesture-referent mapping is arbitrary.
	iconic	Gesture depicts aspects of spatial images, action, people or objects.
Binding	*robot-centric*	Location defined w.r.t. robots.
	user-centric	Location defined w.r.t. user.
	world-dependent	Location defined w.r.t. world.
	world-independent	Location can ignore world features.
Flow	*discrete*	Response occurs after the user acts.
	continuous	Response occurs while the user acts.
TOUCH		
# of finger /hands	*1-2 fingers*	1-2 fingers are used.
	one hand	More than two fingers used.
	both hands	Both hands are used.
# of robots touched	*one robot*	User manipulates one robot.
	few robots	User manipulates 2-4 robots.
	many robots	User manipulates more than 4 robots.
Control paradigm	*leader-follower*	User manipulates only one robot.
	follow crowd	User manipulates a subset of robots.
	control all	User manipulates all of the robots.
Manipulation type	*NP, NM*	Non-prehensile with no motion.
	NP, M, NW	NP with motion but not within hand.
	P, NM	Prehensile with no motion.
	P, M, NW	P with motion but not within hand.
VERBAL		
Illocutionary acts	*Directives*	Get the hearer to do something.
	Expressives	Express attitudes and emotions.

Verbal

Searle classified illocutionary acts into five basic categories: representatives, directives, commissives, expressives, and declarations [222]. However, as I only observed directives and expressives during the study, I labelled each verbal interaction as one or the other.

7.3.2 Reliability

Two of the authors coded the interaction based on the recorded videos. To improve agreement, all four authors discussed the coding scheme and coded some common interactions together. To measure the Inter-Rater Reliability (IRR), the two coders independently coded 15 conditions from two different participants and calculated the unweighted Cohen's Kappa for 11 items (average = 0.79, std = .17). For the remaining conditions, only one rater coded each condition.

Agreement Score

After grouping identical interactions within each referent, I computed the agreement score [223] for each referent, number of robots, and proximity.

Statistical Analysis

I used Fisher's exact test of independence to test whether the proportions of one nominal variable are different depending on the value of another nominal variable as the sample size is relatively small (n¡1000) [224]. Then, Bonferroni-corrected post-hoc tests were used to determine which pairs are significantly different. For instance, Fisher's test was used to test whether the proportions of number of hands used for gesture interactions were different based on the number of robots. To compare the means of the participant's ratings on their interaction for different number of robots, proximity, and tasks, I used N-way ANOVA followed by Bonferroni-corrected post-hoc tests.

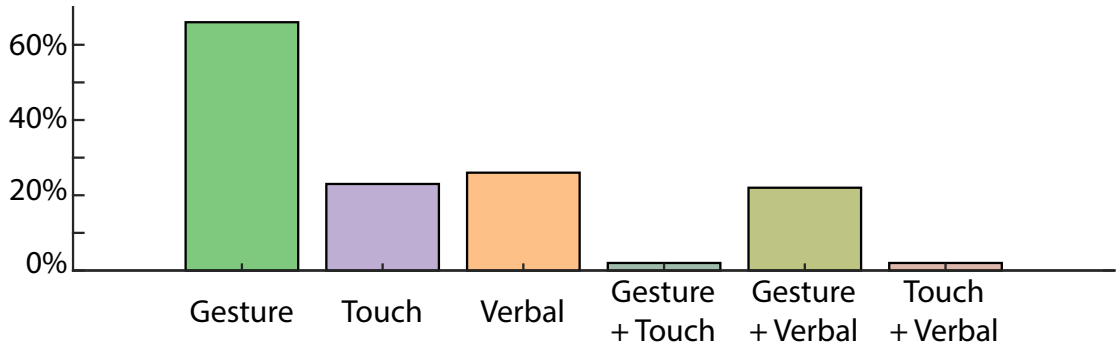


Figure 7.4: Breakdown of Interaction Modality for all conditions. Adapted from [3].

7.4 Results & Discussions

7.4.1 Overall Trends

Interaction Modality

I categorized each interaction into one of the following interaction modalities: gesture, touch, verbal commands, and combinations of them. Figure 7.4 presents the breakdown of interaction modalities used across all conditions. For multimodal interactions, they are counted in all relevant categories. For example, interactions with both gesture and verbal commands are counted in "Gesture", "Verbal", and "G+V". When looking at these overall results in the context of prior work, I see some similar trends to single robot interaction across different types of robots in terms of interaction modality - for example, the results for cm-scale wheeled robots are similar to the results found by Abtahi et al for a single caged "safe" aerial drone [134]. Yet, the results are quite different than that of uncaged "unsafe" drones, potentially due to the non hazardous nature of these small wheeled robots. However, the results are less directly comparable to other studies which did not explore the use of touch or direct manipulation for control of many robots as the study is done in a virtual environment [135, 136]. Yet, similar to [136], I also observed that the majority of the speech commands were accompanied by a gesture.

Taxonometric Breakdown

Using the taxonomies in Table 7.2, I labelled each interaction and the taxonometric breakdown is shown in Figure 7.5.

Gesture:

The majority of the gestures had static pose and path form. In terms of the nature, there is heavy reliance on the use of physical, symbolic, and deictic gestures. This suggests that similar to how *physics engine* is used for surface recognition [132, 225], swarm robot control could also benefit from a physics-based detection algorithm. In addition, it is important for the recognition algorithm to know common symbolic gestures as participants expected the robots to understand common symbolic gestures such as a "stop" gesture with palm showing toward the robot(s) or a "calming" gesture for the slow down referent with hands moving down slowly.

Most gestures were defined with respect to the robots and almost 90% were discrete. The flow was most likely influenced by two factors. First, many of the referents such as robot selection tasks and get attention task are simple with no intermediate steps. Thus, there was no need for continuous input. Second, the robots were not fully controlled impromptu but rather had pre-programmed behaviors with investigator-timed initiation. This setup did not allow any adjustments to the robots' behavior after the initiation and thus discouraged participants from using continuous interactions.

Touch:

55% of touch interactions involved one or two finger touch to physically manipulate one robot. When the task involved more robots, participants relied on different control paradigms such as leader-follower (where they only manipulate one robot and expect the rest to follow), and follow crowd (where they manipulate a subset of the entire group and expect the rest to follow) as it was difficult to grab and manipulate all of the robots at the same time.

I also observed that participants tended to use other modalities when more robots were involved. For instance, P2 wrote in the survey that "...I wasn't sure how to grab them all so it led me to think of other ways to direct them other than touching them." while P8 mentioned that "more robots there were, the more I was inclined to give a global input. Like audio."

43% of the touch interactions were non-prehensile with motion not within hand (C,NP,M,NW) while 41% were prehensile with motion not within hand (C,P,M,NW). I saw very little counts of prehensile manipulation with no motion (P, NM) as participants usually grabbed the robots to move them somewhere. Even for tasks where the robots do not need to move

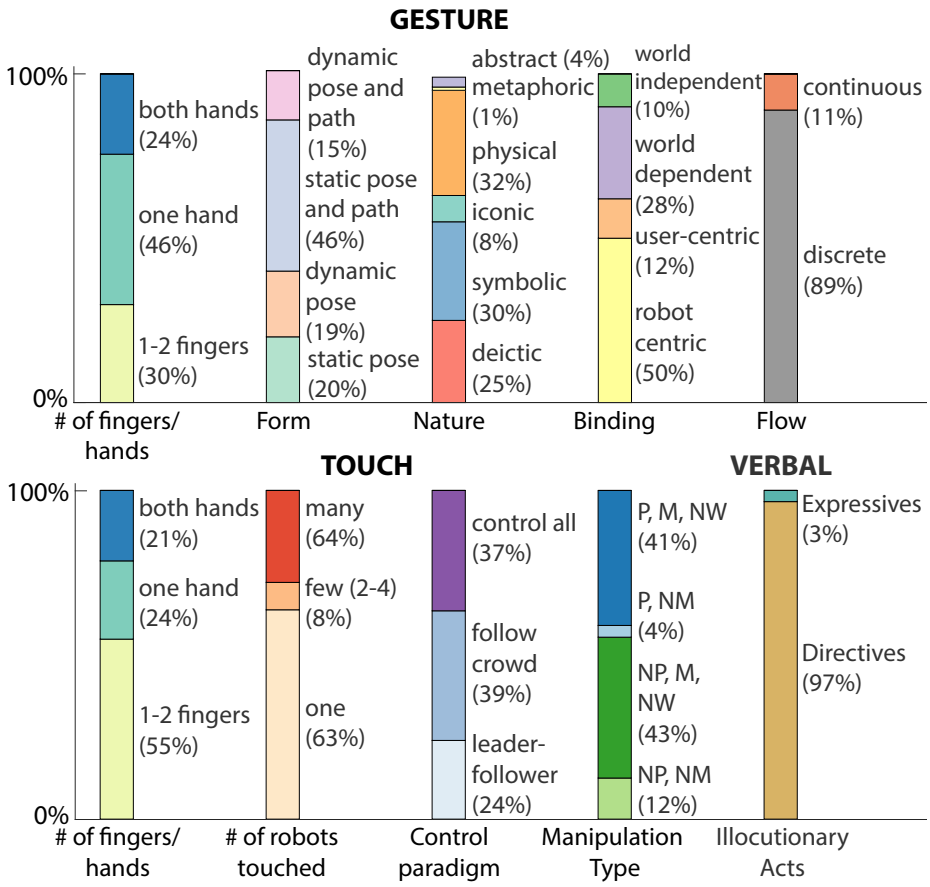


Figure 7.5: Taxonometric breakdown for Gesture, Touch, and Verbal Interactions across all conditions. Adapted from [3].

such as robot selection tasks, most interaction involved tapping or touching the robot(s) while there were only few instances of pure grasp with no motion.

Verbal:

97% of the verbal interactions were directives (i.e., commands). However, there were a few cases where the participants used expressives instead to imply what the robots should do. For instance, a participant said "you guys are moving too fast" to imply that the robots should move slower, whereas another said "you guys are too tight" for the "scale up" referent. This suggests that some users may not always explicitly communicate the desired action and that a voice recognition algorithm will need to infer the user's intention.

Agreement

The agreement scores across all interaction modalities for each referent, number of robots, and proximity are shown in Figure 7.9. The overall average agreement scores for gesture, touch, and verbal interactions independently are $A_G = 0.3$, $A_T = 0.56$, and $A_V = 0.37$.

7.4.2 User-Defined Interaction Set

The user-defined interaction set was generated by taking the most frequent interaction for each referent. If the same interaction was used for different referents thus creating conflict, the interaction was assigned to the referent with the largest group. The resulting interaction set is shown in Figures 7.1, 7.6, and 7.7. For each interaction, I describe the proximities, numbers of robots that the interaction is representative of, and the interaction modality of the interaction. These are represented by the three values inside the bracket after the description of the interaction. The subscript 1 or 2 under "far" or "20" indicates that the interaction is the first or second most frequent interaction for the "far" or "20" robot condition. For example, for the top-left interaction "draw a circle with a finger", it is the most frequent gesture interaction for both far and close proximity condition. Different task categories are indicated by the colored box around the illustration. Blue, dark green, orange, red, teal, and maroon boxes represent represent inter-robot interaction, robot selection, user-centered navigation, getting attention, navigation in environment, and object manipulation task types.

Prior work has shown aliasing significantly improves the input guessability [223,226]. In the interaction set, five referents are assigned 1 interaction, eleven referents have 2 interactions, seven referents have 3 interactions, and one referent has 5 interactions. Out of the 53 interactions in the final set, 15 (28%) are performed with one or two fingers, 21 (40%) are performed with one hand, and 17 (32%) with both hands.

7.4.3 Effects of Number of Robots

As I hypothesized, the number of robots had significant effects on the participant's interaction. It had a positive correlation with the number of hands used, affected the control paradigm they used for touch interactions as shown in Figure 7.8d, and had a negative correlation with participant's simplicity ratings of the interaction ($p < 0.05$).

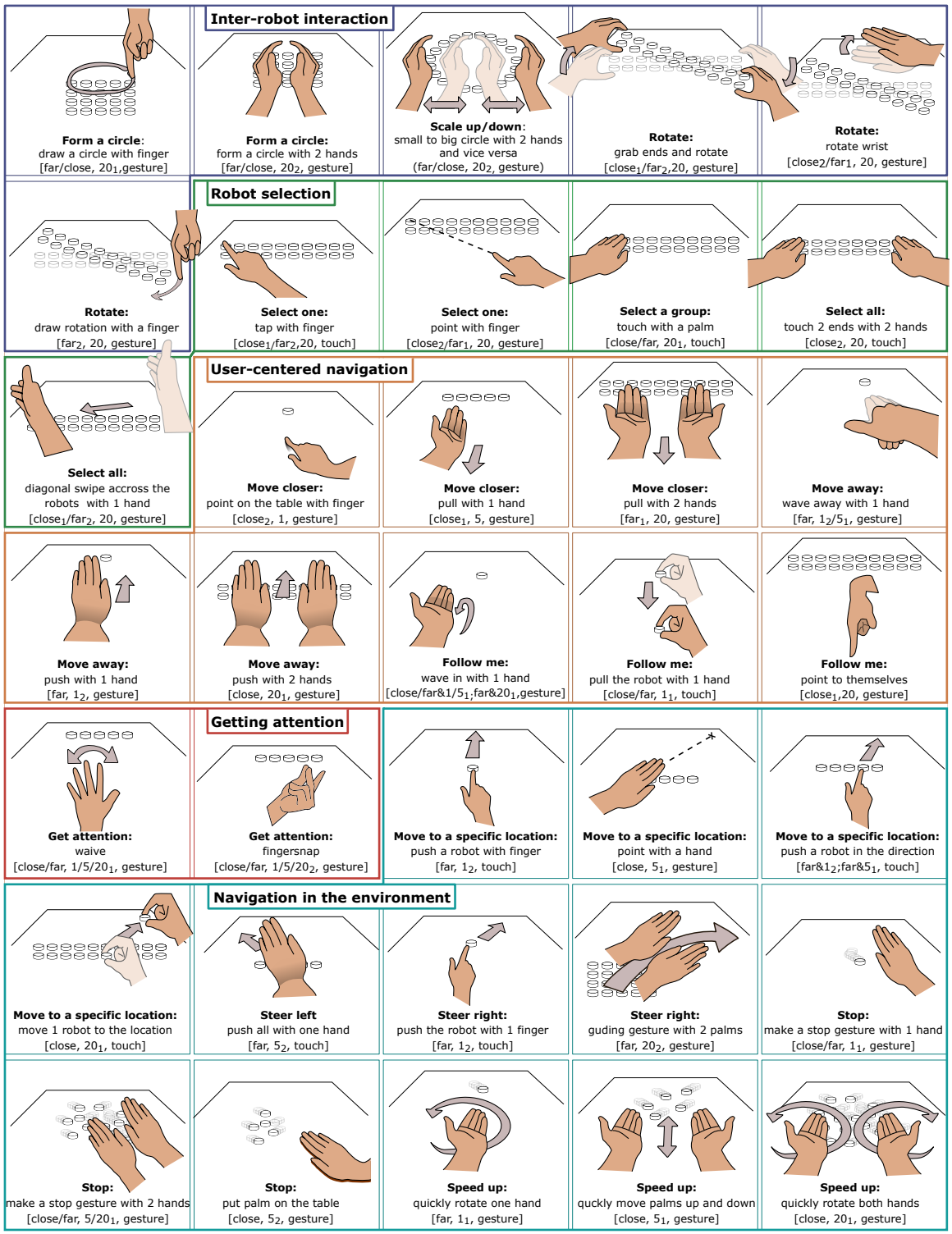


Figure 7.6: User-Defined Interaction Set. To save space, reversible gestures (split/merge, scale up/down, steer left/right) have been combined, and the interactions shown on the first page are not shown here. The set is continued in Fig. 7.7 Adapted from [3].

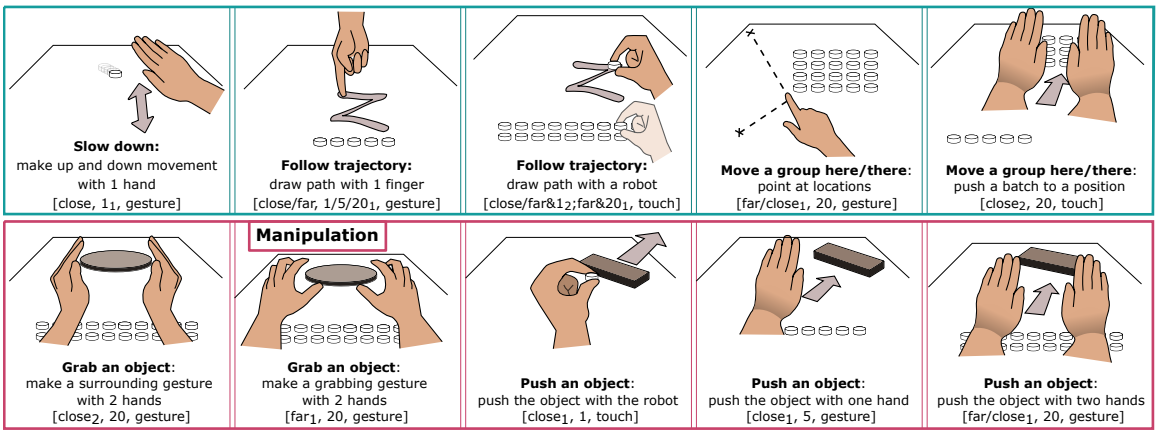


Figure 7.7: User-Defined Interaction Set, continued from Fig. 7.6. Adapted from [3].

Number of Fingers/Hands:

The number of robots had a significant effect on how many hands the participants chose to use. When interacting with more robots, participants increased the number of their hands for both their gesture and touch interactions (both $p < 0.001$) as shown in Figure 7.8a and 7.8b. To control a single robot, they used one/two fingers or a single hand whereas they relied on both hands to interact with 20 robots. The post-test survey revealed that participants were indeed mindful of how they use the number of hands. For example, P9 wrote "If there was one robot I was more likely to use one finger, versus all of the robots, I wanted to ... use two hands.", while P5 wrote "I often wanted to use both hands when interacting with a group of robots, even though I knew already a single hand could work the same." As P5 mentioned, even though there was no need or instruction to use more hands for more robots, participants felt the need to use both hands as confirmed by the study results. Although not explicitly studied, this trend is hinted in the interaction set from [135].

In addition to using both hands, I also observed that participants tried to match the size of their hands, via spreading fingers, to the number of robots. For instance, P15 wrote in the post-test survey "I tried to spread my hands wide enough to cover the whole area of the robots" while P4 mentioned that "I tended to use all my fingers with larger groups." Further investigation will be needed to confirm this.

While the proposed taxonomy does not capture the magnitude of the gesture (i.e., how big spatially the gesture is), participants also mentioned that they used "bigger gesture[s]

for larger number of robots" (P3), made "larger motions when there were more robots" (P14).

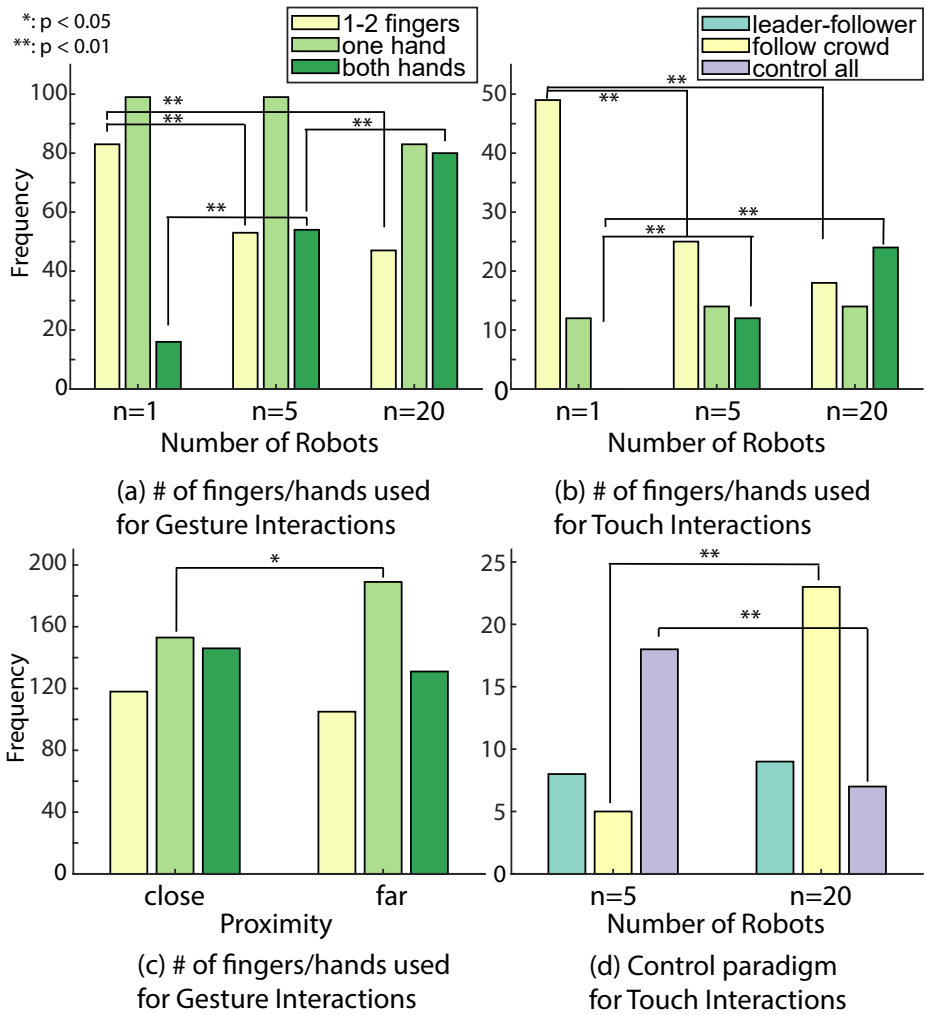


Figure 7.8: Number of robots have significant impact on the number of fingers/hands used for (a) gesture and (b) touch interactions. (c) For gesture interactions, proximity significantly affects the number of fingers/hands used. (d) For touch interactions, the number of robots has significant effects on the control paradigm. Adapted from [3].

Control Paradigm for Touch Interactions:

When interacting with many robots, participants were less likely to directly manipulate all of the robots as shown in Figure 7.8b ($p < 0.001$). To overcome this, they either used a

leader-follower or a follow crowd control paradigm, where they directly manipulate either just one or a subset of the robots respectively. I see this change as the number of robots increases from 5 to 20, as shown in Figure 7.8b.

Simplicity Ratings:

The number of robots significantly affected the participant's simplicity ratings on their interaction ($p < 0.05$). The simplicity ratings for interactions with one robot were higher than those of interactions with 20 robots.

7.4.4 Effects of Proximity

As hypothesized, proximity to the robots had significant effects on how participants chose to interact in terms of number of hands used as shown in Figure 7.8c and their self-reported ratings on how precise their interaction was ($p < 0.05$).

Number of Finger/Hands:

Proximity had significant effect ($p < 0.05$) on the number of hands used for gesture as shown in Figure 7.8c. When the robot were far away, participants used one hand more often than when the robots were close. One potential reason for this is that when the robots were far away, I found participants tended to lean forward over the table to make the gesture clearer to the robots; it may have become more convenient or stable for the users to use only one hand in such a position.

Precision Ratings:

The proximity to the robots significantly affected the participant's precision ratings on their interaction ($p < 0.05$). The precision ratings for close proximity conditions were higher than those for far proximity conditions.

7.4.5 Trends within Each Referent Category

For each referent category, I compared its data with that of the remaining referents. For instance, for the robot selection category, I compared its data with that of referents not in the robot selection category.

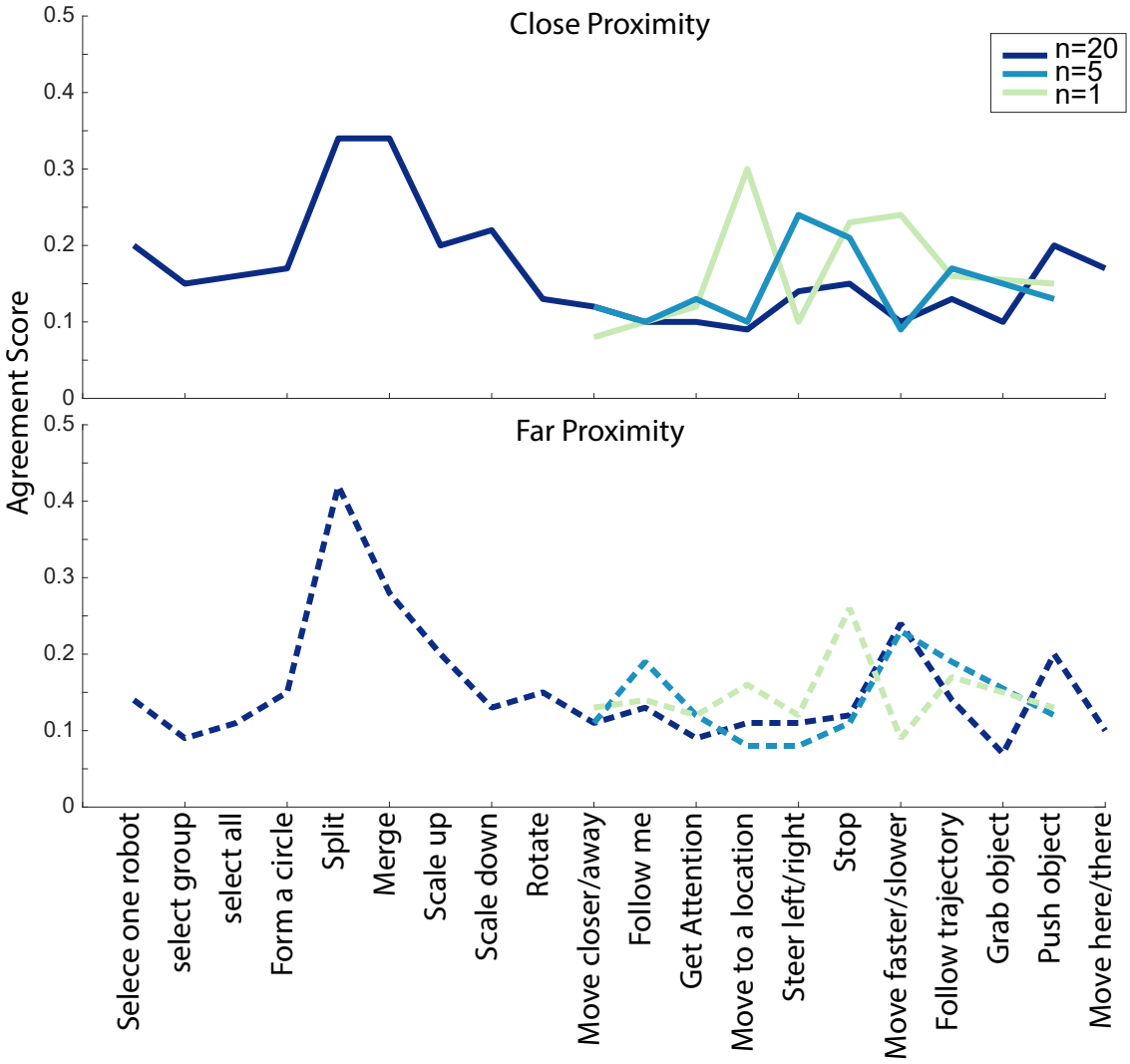


Figure 7.9: Agreements for each referent across all interaction modalities. Adapted from [3].

Robot Selection

For robot selection tasks, participants relied significantly more on touch interactions than for non-selection tasks ($p < 0.01$). Also, proximity had a significant effect on the interaction modality ($p < 0.05$). When selecting robot(s) in close proximity, participants tended to use touch interaction more frequently than when selecting remote robot(s). As shown in Figure 7.10, participants used significantly fewer two-handed ($p < 0.01$) and "static pose

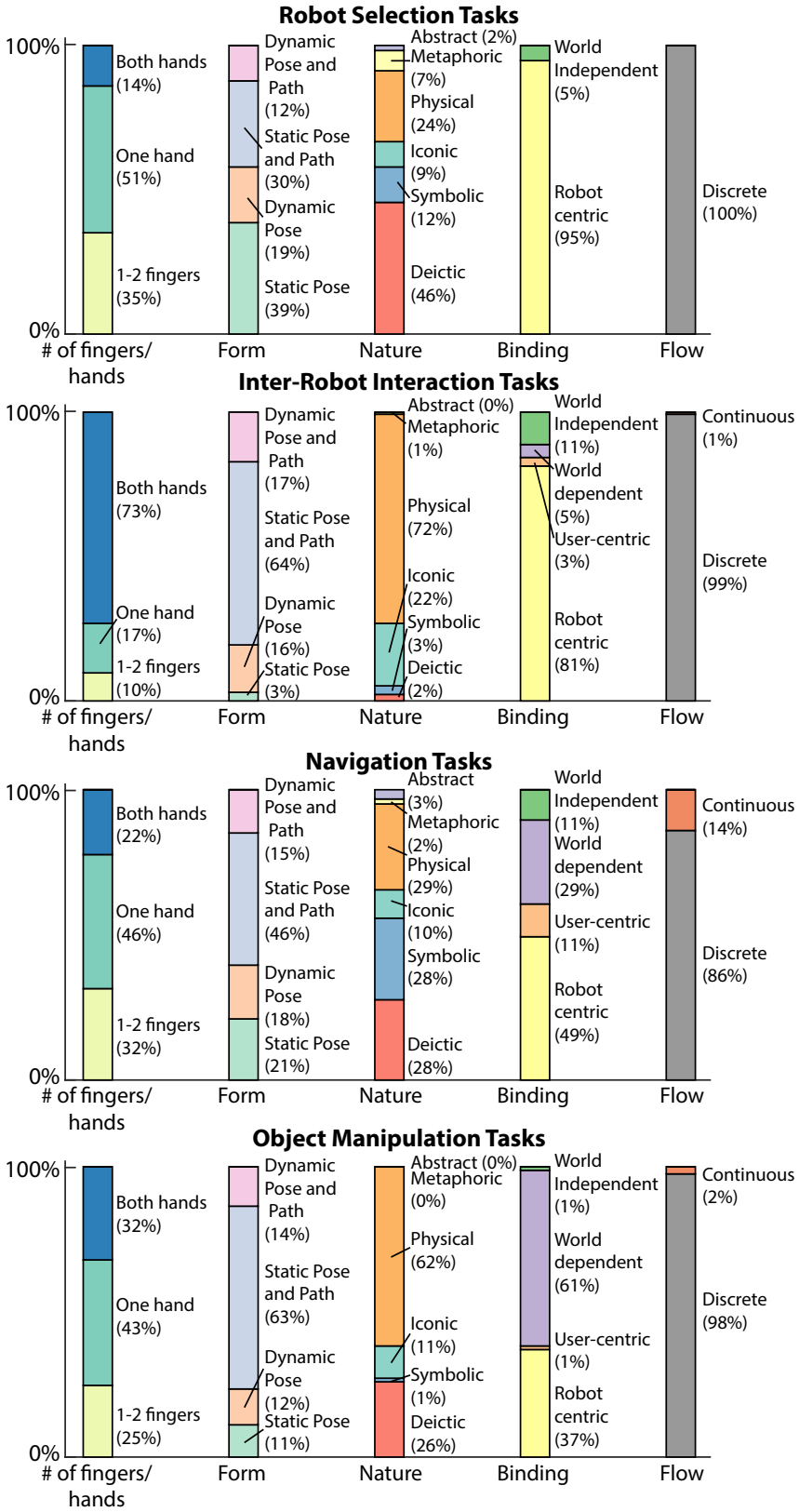


Figure 7.10: Taxonometric breakdown for different referent categories. Adapted from [3].

and path" form gestures while using more "static pose" form gestures ($p < 0.01$) compared to the non-selection tasks. The nature of the gestures was also different as there was a significant increase ($p < 0.001$) in use of deictic gestures and decreases in physical and abstract gestures. Almost all of the gestures were discrete and robots-centric. These results could inform the design of interaction techniques for selection tasks with many robots, which could be used in many applications such as infrastructure maintenance [227], search-and-rescue [67], data physicalization [228], and environmental monitoring [229].

Inter-Robot Interaction

For inter-robot interaction tasks, many participants used the shape of their hands to control the pattern formed with the robots which is also demonstrated in [135]. Contrary to the robot selection tasks, there was a significant increase ($p < 0.001$) in the use of two-handed gesture and a decrease in the use of one/two finger and one-handed gesture. Participants relied more on iconic and physical gestures and less on abstract, deictic, and symbolic gestures to control the robots' formations ($p < 0.001$). Similar to robot selection tasks, most gestures were discrete and robots-centric. These interactions can be used in applications like animation display [40] where it is critical to control the patterns formed by the robots.

Navigation

For navigation tasks many participants mapped movement of the robot(s) to hand motion, a similar trend as shown in [49, 136]. I also observed a significant increase ($p < 0.05$) in multimodal interaction, specifically gesture combined with verbal commands. As shown in Figure 7.10, I saw significant increases in one/two finger and one-handed gestures and a decrease in two-handed gesture ($p < 0.001$). Participants used more deictic and symbolic gestures and less physical gesture ($p < 0.001$). Presumably due to the nature of the tasks, there was a significant increase in continuous flow and a decrease in discrete flow ($p < 0.001$). These results can help inform the design of navigation control for deployment in search-and-rescue [67] and mining or agricultural foraging tasks [230].

Object Manipulation

Some participants explicitly communicated that they wanted the robots to push or grab the object through a tap or voice command, while others simply pushed or gestured the robots

to move toward the object. There was a significant increase in touch + verbal interactions for the object manipulation tasks ($p < 0.05$). As may be expected as the tasks involved physical manipulation of an object, I saw increases in physical nature ($p < 0.001$) and world-dependent binding ($p < 0.001$). These results are relevant for a number of different applications such as the development of domestic robots for cleaning [231] or for robotic assembly tasks [70].

7.4.6 Design Insights

Based on the results of this study, this section presents a brief series of insights towards more effective interface design for future developers of swarm systems.

The user-defined interaction set suggests that the interaction vocabulary changes depending on the state of the robots. Specifically, I observe that the number of robots as well as their proximity affects the user's interaction. This dynamic interaction vocabulary means that in addition to being able to detect input, swarm state information needs to be constantly relayed to the interface device and combined with an inference of user's intention in order to determine *contextually relevant* control inputs.

Gesture was the most commonly used interaction modality (66%) and this mirrors prior works in human-drone interaction [49, 134]. This suggests that if one were to choose only one type of sensor, one should choose a sensor that can detect different types of gestures. Interestingly, prior works in swarm robot control were able to correctly choose gesture as their main interaction modality even without such an elicitation study [31, 94]. The study results better inform what types of gestures the sensor should be able to detect in order to better accommodate user's preference. For instance, being able to sense both hands is important when the user needs to control different number of robots as the results show a positive correlation between the number of robots and the number of fingers/hands used. Simultaneously, there is a need to detect relatively fine gestures (e.g., those involving only one or two fingers) as 31% of user interactions fell in that category.

While users heavily relied on gesture, they also used touch and verbal interactions 23% and 26% of the time. An ideal interface would be able to detect various types of touch and verbal interactions in addition to gesture. This would not only better support the user-defined interaction set but also provide users with additional degrees of freedom to leverage for different operational circumstances. For instance, in a dark room where the location of

the robots is unknown, a user may find verbal interaction more useful than gesture or touch for getting attention of the robots.

Robots at the scale used in this study will struggle with the payload and energy demands of a vision system capable of user gesture identification, so even consensus-based approaches which take into account non-optimal classification may not be feasible. While centralized computer vision solutions (ideally incorporated into the infrastructure for robot path planning and control) may be the solution for tabletop and other stationary deployment environments, a gesture recognition device wearable by the operator may make the most sense for unstructured or mobile applications. Based on the study finding that operators begin to use more two-handed gestures when robot number increases, a future wearable solution must be able to accommodate use of both arms/hands.

I found a negative correlation between increasing number of robots and proximity from the robots and self-reported ratings of interaction Simplicity and Precision. This finding aligns with prior research in teleoperated swarms, where users have a difficult time predicting how their control input will propagate through swarms [136]. Future interfaces should be "predictive" [232], providing some amount of feedback in real time to the user in the form of overlaid visual output from a path-planning algorithm or haptic feedback through the interface device, in order to decrease this uncertainty.

7.5 Limitations & Future Work

The fact that the study was conducted with relatively small tabletop robots limits the generalizability of the results. For example, the size of the robots discouraged several participants (P5, P7, P8, P15) from interacting with physical touch as they were "scared to break them" [P5, P8] even though they were told not to worry about damage to the robots. Limiting the robot environment to the tabletop also sets bounds on the maximum distance from the user as well as the maximum number of robots that can be interacted with at a time. Future work should investigate interaction in a room-scale environment – not only would it add more potential robots and distance, but also another dimension to vary (i.e., workspace or group height relative to user).

There exists a "legacy effect" in elicitation studies that leads users to fall back on their early or first responses even when parameters or tasks are varied in the future [233]. A

larger participant pool would help to disambiguate user responses from this effect in future studies.

It is possible that the high percentage of users who elected for gestural control of the robots was influenced by the fact that this study was limited to tasks where the human operator is solely engaged with the robots. Future work could investigate whether preferred user input modality is changed if, for example, their visual attention is required elsewhere or hand(s) are otherwise occupied with some task.

Prior work has shown differences in preferred user input modality depending on the cultural background [234, 235]. I did not specifically investigate this effect or account for it in this elicitation study, although it is an important area for future work.

7.6 Chapter Summary

Controlling a large number of robots is a difficult task, which has primarily been managed by technicians and experts knowledgeable in robotics or control. As we approach a future in which autonomous agents will be roaming around us, it is important to lower the barriers for non-expert users. To do so, here I conduct an elicitation study aimed to dissect how people naturally interact with multi-robot systems. Mirroring the research that has shown a spectrum of feasible higher-level control strategies for swarm systems depending on their implementation details and level of autonomy, here I show that user-elicited interaction methods are closely related with the number of robots being interacted with at a time and their relative proximity. While the majority of the participants used gesture as their main control modality, it does not capture the entire range of interaction that people use. As future encountered robot swarms will be highly dynamic and mobile, this work indicates that their effective operation will also require dynamic, state-dependent, and multimodal interaction vocabularies for effective interaction.

Chapter 8

Conclusion

Robots have revolutionized the manufacturing industry. They are now in process of transforming service sectors ranging from health care, transportation, construction, entertainment, and agriculture. With the increasing complexity and larger scale of the required tasks, researchers are developing swarms of robots as means to provide robust and versatile systems suited for large-scale operations [58, 67–69]. For successful integration of robots into these sectors, it is necessary for the robots to not only successfully complete their jobs but also interact with people in the process, as well as transition quickly between these tasks. Thus, the main goal of this thesis has been to design multimodal, ubiquitous, and in situ interaction with multi-robot systems. I investigated how to leverage the mobility of these systems to display meaningful information visually through motion when the robots are distant and through touch when they are nearby users. To investigate proximal command for human-multirobot interaction, I studied how people naturally control and provide multimodal inputs to a large number of robots.

The primary contributions of this thesis are not necessarily the exact results from the studies but rather the approaches and the concepts that were used and applied to address different problems. As context is immensely important in deciding what the best solution is for any application, I believe the results from this thesis should only be used to find a starting point rather than the final solution. As I used versions of a specific robot platform, Zooids, throughout the thesis, the study findings are most applicable to similar platforms. Although interaction with robotic platforms of different scales may demonstrate different trends than those from this thesis, the methods used here are likely still applicable. For example, using swarm behaviors to embed expressive information in a scalable manner

and using preattentive processing features to design behaviors that leverage human vision perception should be directly applicable at different scales as they rely on mobility an attribute found in almost all robots.

My fundamental belief is that as robots become ubiquitous, interaction with these robots will be increasingly important. In this thesis, I provided insights on how to design interaction with small multi-robot systems and seeks to inspire more to consider this aspect.

8.1 Thesis Summary

Here, I summarize the contributions and results of the thesis.

Chapter 3. Ubiquitous Robotic Interfaces

- I described my vision for Ubiquitous Robotic Interfaces (URIs). I highlighted the key components of URIs and how I anticipate them being embedded into our daily lives. I also described a prototype of URIs, Zooids [5], a multi-robot platform designed for real-time interaction with people and co-developed with Le Goc et al. I used this platform to explore various aspects within real-time display and proximal, in situ interaction with users. I explained the design rationals and implementation details of Zooids such as form factor, communication structure, and software.

Chapter 4. Investigation of Abstract Motion as a Display

- Mobility is a fundamental aspect of any robotic system whether it is a robotic arm or a mobile robot. It enables robots to manipulate and sense the environment while also allowing them to interact with users in a physical level. I investigated ways to layer expressiveness such as affect and urgency on top of swarm robot motion such that the robots could be expressive even during navigation or manipulation. I first identified three abstract multi-robot motion parameters from prior literature: speed, smooth, and bio-inspired collective behaviors. Through a crowdsourced between-subjects user study,I evaluated human perception of these parameters, both individually and in combination. The study results indicate that the different collective behaviors elicit significantly different responses in perceived arousal, dominance, hedonic and pragmatic qualities, animacy, urgency and willingness to attend. On the other hand, speed and smoothness had similar effects as shown from prior work in human perception

of single robot motion. Speed significantly affects valence, arousal, hedonic quality, urgency and animacy while smoothness affects hedonic quality, animacy, attractivity and likeability. These results serve as design guidelines for URI-based displays and I demonstrated the use of the guidelines through a few example applications. While there are many more methods to leverage abstract motion as a display (e.g., using principles of animation [164]), the results support the feasibility and effectiveness of collective behaviors to generate more expressive swarm robot motion.

Chapter 5. Legibility and Glanceability of Swarm Robot Motion

- To facilitate human-robot interaction and collaboration, robots need to be intent-expressive in their behaviors. I explored the use of trajectory, pre-attentive processing features, and collective behaviors in order to generate legible and glanceable swarm robot motion. I conducted 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 whereas the trajectory-based motion has the highest glanceability. I also observed significant effects of task parameters like the radius of the initial circle that encompasses the robots and the location of the targets which determine the difficulty level of the task. Rather than a one-size-fits-all solution, generating legible and glanceable swarm robot motion will require a more complex solution that consists of a combination of these different legibility cues based on the context of the task.

Chapter 6. SwarmHaptics: Haptic Display with Swarm Robots

- Touch is an integral part of our lives that allow us to both detect and manipulate our environment, and help build rapport with other people. In addition to the use of motion to visually display information, I investigated how a swarm of robots could provide richer information to users through both vision and touch in the context of notification, directional cue, and remote social touch. To do so, I introduced SwarmHaptics, a new type of visual and haptic display consisting of small mobile robots that can approach the users at will to provide different haptic patterns. I described the design space for SwarmHaptics with parameters for both each individual robots and collections of them, and demonstrated its possibilities with a few example scenarios. I evaluated how people perceive different haptic patterns and found that the number of robots involved is important as it increases urgency and arousal but at the sacrifice

of likeability and perceived safety. A separate elicitation study on remote social touch revealed many interesting trends. For instance, many participants relied on the visual patterns that the robots create and the contact location to convey more abstract social touch as haptic stimulus alone was often insufficient to convey complex information. This demonstrates the need to integrate both the visual and haptic stimuli in order to provide comprehensible information to users.

Chapter 7. User-defined Swarm Robot Control

- Controlling a large number of robots is a complex task, which has mostly been managed by technicians and experts knowledgeable in robotics or control. As autonomous robotic agents are poised to become ubiquitous, it is important to lower the barriers for non-expert users. To address this, I studied how people interact with multi-robot systems through an elicitation study. Mirroring the research that has shown a spectrum of higher-level control strategies for swarm systems depending on their implementation details and level of autonomy, I demonstrated that user-defined interaction methods are closely related with the number of robots being interacted with at a time and their relative proximity to the user. While the majority of the participants used gesture as their main command modality, a significant portion also leveraged other modalities such as verbal commands and direct touch manipulation. As future encountered robot swarms will be highly dynamic and diverse with varying group size and proximity to users, the study results indicate that effective operation of multi-robot systems will also require dynamic, state-dependent, and multimodal interaction vocabularies.

8.2 Future Work

In this thesis, I introduced my vision of Ubiquitous Robotic Interfaces and sought to deepen our understanding of how to display information and command multi-robot systems. In the design of ubiquitous multimodal interaction with these robots, there are still many more challenges, both technical and design-related, that require further investigation.

8.2.1 Size and Scale

One major limitation with the studies presented in this thesis is that they were all done with the same robotic platform, Zooids [5]. With the emergence of robots that range from fleets

of autonomous vehicle to micro-robots [13, 236], we need to understand how size or scale of the robots affects the interaction experience. In Chapters 6-7, I have explored the effects of number of robots in both how people perceive their touch and how people command them, but it is easy to imagine how any drastic change in scale or size of the robots can alter the study results. For instance, the experience of a small robot touching your forearm will be significantly different when a vehicle-size robot does the same even if the force or contact area remains constant. Prior work demonstrate that even just changing the height of a telepresence robot alters the perceptions of the operator [193].

8.2.2 Combining Different Modalities

Another important future direction is making the interaction truly multimodal. In Chapters 4-6, I conducted studies to understand how to display information visually through motion of the robots and through touch. While these results could serve as building blocks, designing multimodal interaction will require a thorough understanding of how the different modalities combine. For instance, is it different when I see and feel the robots compared to when I just feel the robots? While prior work explores some aspects [237], further investigations are needed to truly leverage the aural, visual, and haptic sensory channels.

8.2.3 Bidirectional Interaction

This thesis focuses on the display and command components of human-multirobot interaction independently. In order to deploy ubiquitous robots into the real world, robots need to support bidirectional real-time interaction with people. For instance, instead of blindly following a pre-computed trajectory, robots should dynamically adjust their behaviors based on the state of the users. When people are performing delicate tasks that demand undivided attention, robots should be aware and modify their behaviors to minimize their distraction to the users.

8.2.4 Context

Context is a crucial element of any interaction. I demonstrated in Chapter 7 that even just changing the number of robots or proximity to the robots can affect how people choose to interact. Cultural background also plays a role in how people interact with robots [234, 235]. Thus, before deploying any robotic system, developers should conduct long-term field studies

in order to comprehend the culture, settings, and context in which the interaction occurs and modify robot behaviors as needed.

8.2.5 Using Robots to Support Self-Actualization

I envision that we will eventually go beyond interaction and communication with robots, and will truly leverage their physical nature to help us achieve *self-actualization*. Similar to how computers and the internet provide us with unlimited amounts of information and connections with others, robots will support us as dynamic physical tools or agents that will enable us to become the best versions of ourselves. In the future, I would like to create dynamic physical environments conducive to positive behavioral changes, and design motivational non-humanoid agents for users.

Dynamic physical intervention for positive behavioral change:

One of the primary benefits that robots have over computers is their physical nature. Unlike computers, which mainly operate in the digital world, robots have the ability to directly manipulate the physical world. For example, while computers are limited to notifying a user to drink water either visually or aurally, often via unpleasant pop-up messages, robots can either point toward the water bottle in the real world or even physically bring the water bottle to the user. My goal is to take advantage of this capacity to help people achieve their goals and generate varying degrees of physical friction against undesirable habits to help them "stay on track". For instance, robots can help users form healthy habits by preventing access to their phone until they have completed their set of tasks.

Leveraging animacy of non-humanoid robotic agents:

Robots lie on the spectrum between non-living matter and human. Depending on their visual appearance, robots can be perceived as insects, pets, or even as another human being. This creates an interesting opportunity for non-humanoid robots to emulate the presence of another animate being, but in a controlled manner and potentially without the negative side effects of having other people around. For instance, prior work in HRI revealed significant evidence of social facilitation effects for both human and humanoid robotic presences compared to an alone condition with no human or robot present [238]. When either a human or humanoid robot was present, people performed better in easy tasks

due to a heightened sense of activation, but performed worse in difficult tasks potentially due to the psychological and social pressure from other people. However, if a non-humanoid robot was used instead, the sense of being judged could potentially be mitigated while still increasing the arousal level. Once validated, this concept could be used to help people boost their own productivity whenever necessary. Non-humanoid robots provide an opportunity to inject controllable agents while avoiding the Uncanny Valley effect that humanoid robots often suffer from.

Bibliography

[1] Lawrence H Kim and Sean Follmer. Ubiswarm: Ubiquitous robotic interfaces and investigation of abstract motion as a display. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(3):66, 2017.

[2] Lawrence H Kim and Sean Follmer. Swarmhaptics: Haptic display with swarm robots. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, page 688. ACM, 2019.

[3] Lawrence H Kim, Daniel S Drew, Veronika Domova, and Sean Follmer. User-defined swarm robot control. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, page 688. ACM, 2020.

[4] Christoph Bartneck and Jodi Forlizzi. A design-centred framework for social human-robot interaction. In *RO-MAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No. 04TH8759)*, pages 591–594. IEEE, 2004.

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

[6] Marco Dorigo, Elio Tuci, Roderich Groß, Vito Trianni, Thomas Halva Labella, Shervin Nouyan, Christos Ampatzis, Jean-Louis Deneubourg, Gianluca Baldassarre, Stefano Nolfi, et al. The swarm-bots project. In *International Workshop on Swarm Robotics*, pages 31–44. Springer, 2004.

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

[8] Michael Rubenstein, Christian Ahler, and Radhika Nagpal. Kilobot: A low cost scalable robot system for collective behaviors. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pages 3293–3298. IEEE, 2012.

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

[10] Kian Hsiang Low, Wee Kheng Leow, and Marcelo H Ang Jr. Task allocation via self-organizing swarm coalitions in distributed mobile sensor network. In *AAAI*, volume 4, pages 28–33, 2004.

[11] Dingjiang Zhou and Mac Schwager. Assistive collision avoidance for quadrotor swarm teleoperation. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1249–1254. IEEE, 2016.

[12] James McLurkin, Jennifer Smith, James Frankel, David Sotkowitz, David Blau, and Brian Schmidt. Speaking swarmish: Human-robot interface design for large swarms of autonomous mobile robots. In *AAAI spring symposium: to boldly go where no human-robot team has gone before*, volume 72, 2006.

[13] David L Christensen, Elliot W Hawkes, Srinivasan A Suresh, Karen Ladenheim, and Mark R Cutkosky. μtugs: Enabling microrobots to deliver macro forces with controllable adhesives. In *Robotics and Automation (ICRA), 2015 IEEE International Conference on*, pages 4048–4055. IEEE, 2015.

[14] Zijian Wang, Guang Yang, Xuanshuo Su, and Mac Schwager. Ouijabots: Omnidirectional robots for cooperative object transport with rotation control using no communication. In *Distributed Autonomous Robotic Systems*, pages 117–131. Springer, 2018.

[15] Cynthia Lynn Breazeal. *Sociable machines: Expressive social exchange between humans and robots.* PhD thesis, Massachusetts Institute of Technology, 2000.

[16] Heather Knight. Expressive motion for low degree-of-freedom robots. 2016.

[17] Heather Knight and Reid Simmons. Expressive motion with x, y and theta: Laban effort features for mobile robots. In *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*, pages 267–273. IEEE, 2014.

[18] David Sirkin, Brian Mok, Stephen Yang, and Wendy Ju. Mechanical ottoman: how robotic furniture offers and withdraws support. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, pages 11–18. ACM, 2015.

[19] Anca D Dragan, Kenton CT Lee, and Siddhartha S Srinivasa. Legibility and predictability of robot motion. In *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction*, pages 301–308. IEEE Press, 2013.

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

[21] Rudolph Laban. Modern educational dance. revised by l. *Ullmann. London: MacDonald and Evans.(First published 1948)*, 1963.

[22] Heather Knight and Reid Simmons. Laban head-motions convey robot state: A call for robot body language. In *2016 IEEE international conference on robotics and automation (ICRA)*, pages 2881–2888. IEEE, 2016.

[23] Leila Takayama, Doug Dooley, and Wendy Ju. Expressing thought: improving robot readability with animation principles. In *2011 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 69–76. IEEE, 2011.

[24] Marynel Vázquez. *Reasoning About Spatial Patterns of Human Behavior During Group Conversations with Robots.* PhD thesis, figshare, 2017.

[25] Marynel Vázquez, Elizabeth J Carter, Braden McDorman, Jodi Forlizzi, Aaron Steinfeld, and Scott E Hudson. Towards robot autonomy in group conversations: Understanding the effects of body orientation and gaze. In *Proceedings of the 2017*

ACM/IEEE International Conference on Human-Robot Interaction, pages 42–52. ACM, 2017.

[26] Wan-Ling Chang, Jeremy P White, Joohyun Park, Anna Holm, and Selma Šabanović. The effect of group size on people's attitudes and cooperative behaviors toward robots in interactive gameplay. In *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*, pages 845–850. IEEE, 2012.

[27] Harry Bunt, Robbert-Jan Beun, and Tijn Borghuis. *Multimodal human-computer communication: systems, techniques, and experiments*, volume 1374. Springer Science & Business Media, 1998.

[28] Francis Quek, David McNeill, Robert Bryll, Susan Duncan, Xin-Feng Ma, Cemil Kirbas, Karl E McCullough, and Rashid Ansari. Multimodal human discourse: gesture and speech. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 9(3):171–193, 2002.

[29] Gaëtan Podevijn, Rehan O'grady, Nithin Mathews, Audrey Gilles, Carole Fantini-Hauwel, and Marco Dorigo. Investigating the effect of increasing robot group sizes on the human psychophysiological state in the context of human–swarm interaction. *Swarm Intelligence*, 10(3):193–210, 2016.

[30] Gaëtan Podevijn, Rehan O'Grady, Youssef SG Nashed, and Marco Dorigo. Gesturing at subswarms: Towards direct human control of robot swarms. In *Conference Towards Autonomous Robotic Systems*, pages 390–403. Springer, 2013.

[31] Boris Gromov, Luca M Gambardella, and Gianni A Di Caro. Wearable multi-modal interface for human multi-robot interaction. In *2016 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, pages 240–245. IEEE, 2016.

[32] Cynthia Breazeal, Cory D Kidd, Andrea Lockerd Thomaz, Guy Hoffman, and Matt Berlin. Effects of nonverbal communication on efficiency and robustness in human-robot teamwork. In *Intelligent Robots and Systems, 2005.(IROS 2005). 2005 IEEE/RSJ International Conference on*, pages 708–713. IEEE, 2005.

[33] Evan L Ardiel and Catharine H Rankin. The importance of touch in development. *Paediatrics & child health*, 15(3):153–156, 2010.

[34] Jeffrey D Fisher, Marvin Rytting, and Richard Heslin. Hands touching hands: Affective and evaluative effects of an interpersonal touch. *Sociometry*, 39(4):416–421, 1976.

[35] Pirkko Routasalo. Physical touch in nursing studies: a literature review. *Journal of Advanced Nursing*, 30(4):843–850, 1999.

[36] Karin Bundgaard and Karl Brian Nielsen. The art of holding hand: A fieldwork study outlining the significance of physical touch in facilities for short-term stay. *International Journal of Human Caring*, 15(3):34–41, 2011.

[37] Steve Yohanan and Karon E MacLean. The role of affective touch in human-robot interaction: Human intent and expectations in touching the haptic creature. *International Journal of Social Robotics*, 4(2):163–180, 2012.

[38] Kayako Nakagawa, Masahiro Shiomi, Kazuhiko Shinozawa, Reo Matsumura, Hiroshi Ishiguro, and Norihiro Hagita. Effect of robot's active touch on people's motivation. In *Proceedings of the 6th international conference on Human-robot interaction*, pages 465–472, 2011.

[39] Kirstin Hagelskjaer Petersen, Radhika Nagpal, and Justin K Werfel. Termes: An autonomous robotic system for three-dimensional collective construction. *Robotics: science and systems VII*, 2011.

[40] Javier Alonso-Mora, Andreas Breitenmoser, Martin Rufli, Roland Siegwart, and Paul Beardsley. Image and animation display with multiple mobile robots. *The International Journal of Robotics Research*, 31(6):753–773, 2012.

[41] Jakob Nielsen. *Usability engineering*. Morgan Kaufmann, 1994.

[42] George Robertson, Kim Cameron, Mary Czerwinski, and Daniel Robbins. Animated visualization of multiple intersecting hierarchies. *Information Visualization*, 1(1):50–65, 2002.

[43] Andreas Kolling, Phillip Walker, Nilanjan Chakraborty, Katia Sycara, and Michael Lewis. Human Interaction with Robot Swarms: A Survey. *IEEE Transactions on Human-Machine Systems*, 46(1):9–26, 2016.

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

[45] Anne Treisman. Preattentive processing in vision. *Computer vision, graphics, and image processing*, 31(2):156–177, 1985.

[46] Masahiro Shiomi, Kayako Nakagawa, Kazuhiko Shinozawa, Reo Matsumura, Hiroshi Ishiguro, and Norihiro Hagita. Does a robot's touch encourage human effort? *International Journal of Social Robotics*, 9(1):5–15, 2017.

[47] Anna Flagg and Karon MacLean. Affective touch gesture recognition for a furry zoomorphic machine. In *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*, pages 25–32. ACM, 2013.

[48] Dirk Rothenbücher, Jamy Li, David Sirkin, Brian Mok, and Wendy Ju. Ghost driver: A field study investigating the interaction between pedestrians and driverless vehicles. In *2016 25th IEEE international symposium on robot and human interactive communication (RO-MAN)*, pages 795–802. IEEE, 2016.

[49] Jessica R Cauchard, Jane L. E, Kevin Y Zhai, and James A Landay. Drone & me: an exploration into natural human-drone interaction. In *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*, pages 361–365. ACM, 2015.

[50] M Weiser. The computer for the 21st century. *Scientific American (International Edition)*, 265(3):66–75, 1991.

[51] Jong Hwan Kim, Kang Hee Lee, Yong Duk Kim, Naveen Suresh Kuppuswamy, and Jun Jo. Ubiquitous robot: A new paradigm for integrated services. *Proceedings - IEEE International Conference on Robotics and Automation*, (April):2853–2858, 2007.

[52] Abdelghani Chibani, Yacine Amirat, Samer Mohammed, Eric Matson, Norihiro Hagita, and Marcos Barreto. Ubiquitous robotics: Recent challenges and future trends. *Robotics and Autonomous Systems*, 61(11):1162–1172, 2013.

[53] Travis A. Riggs, Tamer Inanc, and Weizhong Zhang. An autonomous mobile robotics testbed: Construction, validation, and experiments. *IEEE Transactions on Control Systems Technology*, 18(3):757–766, 2010.

[54] N. Michael, J. Fink, and V. Kumar. Experimental Testbed for Large Multirobot Teams. *IEEE Robotics & Automation Magazine*, 15(March):53–61, 2008.

[55] Daniel Cruz, James McClintock, Brent Perteet, Omar A A Orqueda, Yuan Cao, and Rafael Fierro. Decentralized cooperative control: A multivehicle platform for research in networked embedded systems. *IEEE Control Systems Magazine*, 27(3):58–78, 2007.

[56] L. Chaimowicz, A. Cowley, D. Gomez-Ibanez, B. Grocholsky, M. A. Hsieh, H. Hsu, J. F. Keller, V. Kumar, R. Swaminathan, and C. J. Taylor. Deploying Air-Ground Multi-Robot Teams in Urban Environments. *Multi-Robot Systems. From Swarms to Intelligent Automata Volume*, III:223–234, 2005.

[57] Abhijeet Joshi, Trevor Ashley, Yuan R. Huang, and Andrea L. Bertozzi. Experimental validation of cooperative environmental boundary tracking with on-board sensors. *Proceedings of the American Control Conference*, pages 2630–2635, 2009.

[58] Scott A. Deloach, Eric T. Matson, and Yonghua Li. Exploiting Agent Oriented Software Engineering in Cooperative Robotics Search and Rescue. *International Journal of Pattern Recognition and Artificial Intelligence*, 17(5):817–835, 2003.

[59] J. P. How, B. Behihke, A. Frank, D. Dale, and J. Vian. Real-time indoor autonomous vehicle test environment. *IEEE Control Systems*, 28(2):51–64, April 2008.

[60] T. W. McLain and R. W. Beard. Unmanned air vehicle testbed for cooperative control experiments. In *Proceedings of the 2004 American Control Conference*, volume 6, pages 5327–5331 vol.6, June 2004.

[61] L. Cremean, W.B. Dunbar, D. Van Gogh, J. Hickey, E. Klavins, J. Meltzer, and R.M. Murray. The Caltech Multi-Vehicle Wireless Testbed. *Proceedings of the 41st IEEE Conference on Decision and Control, 2002.*, 1:3–5, 2002.

[62] Joshua A. Marshall, Terence Fung, Mireille E. Broucke, Gabriele M T D'eleuterio, and Bruce A. Francis. Experiments in multirobot coordination. *Robotics and Autonomous Systems*, 54(3):265–275, 2006.

[63] Andrew Stubbs, Vladimeros Vladimerou, Adam Thomas Fulford, Derek King, Jeffrey Strick, and Geir E. Dullerud. Multivehicle systems control over networks: A hovercraft testbed for networked and decentralized control. *IEEE Control Systems Magazine*, 26(3):56–69, 2006.

[64] Sean Braley, Calvin Rubens, Timothy R Merritt, and Roel Vertegaal. Griddrones: A self-levitating physical voxel lattice for 3d surface deformations. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, page D200. ACM, 2018.

[65] M Brambilla, E Ferrante, and M Birattari. Swarm robotics : A review from the swarm engineering perspective IRIDIA – Technical Report Series Technical Report No . *Swarm Intelligence*, 7(1):1–41, 2012.

[66] Gregory Dudek, Michael Jenkin, Evangelos Milios, and David Wilkes. A taxonomy for swarm robots. In *Intelligent Robots and Systems' 93, IROS'93. Proceedings of the 1993 IEEE/RSJ International Conference on*, volume 1, pages 441–447. IEEE, 1993.

[67] Daniel P Stormont. Autonomous rescue robot swarms for first responders. In *CIHSPS 2005. Proceedings of the 2005 IEEE International Conference on Computational Intelligence for Homeland Security and Personal Safety, 2005.*, pages 151–157. IEEE, 2005.

[68] Ricardo A Daziano, Mauricio Sarrias, and Benjamin Leard. Are consumers willing to pay to let cars drive for them? analyzing response to autonomous vehicles. *Transportation Research Part C: Emerging Technologies*, 78:150–164, 2017.

[69] Andres Milioto, Philipp Lottes, and Cyrill Stachniss. Real-time semantic segmentation of crop and weed for precision agriculture robots leveraging background knowledge in cnns. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2229–2235. IEEE, 2018.

[70] Yiwei Zhao, Lawrence H Kim, Ye Wang, Mathieu Le Goc, and Sean Follmer. Robotic assembly of haptic proxy objects for tangible interaction and virtual reality. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces*, pages 82–91. ACM, 2017.

[71] Michael Rubenstein, Christian Ahler, and Radhika Nagpal. Kilobot: A Low Cost Scalable Robot System for Collective Behaviors. *IEEE International Conference on Robotics and Automation*, pages 3293–3298, 2012.

[72] A. Becker, G. Habibi, J. Werfel, M. Rubenstein, and J. McLurkin. Massive uniform manipulation: Controlling large populations of simple robots with a common input signal. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 520–527, Nov 2013.

[73] G. Dudek, M. Jenkin, E. Milios, and D. Wilkes. A taxonomy for swarm robots. In *Intelligent Robots and Systems '93, IROS '93. Proceedings of the 1993 IEEE/RSJ International Conference on*, volume 1, pages 441–447 vol.1, Jul 1993.

[74] Frederick Ducatelle, Gianni A. Di Caro, Carlo Pinciroli, and Luca M. Gambardella. Self-organized cooperation between robotic swarms. *Swarm Intelligence*, 5(2):73–96, 2011.

[75] Jodi Forlizzi and Carl DiSalvo. Service robots in the domestic environment: a study of the roomba vacuum in the home. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, pages 258–265. ACM, 2006.

[76] Ja-Young Sung, Lan Guo, Rebecca E Grinter, and Henrik I Christensen. "my roomba is rambo": intimate home appliances. In *International conference on ubiquitous computing*, pages 145–162. Springer, 2007.

[77] Satoru Satake, Kotaro Hayashi, Keita Nakatani, and Takayuki Kanda. Field trial of an information-providing robot in a shopping mall. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1832–1839. IEEE, 2015.

[78] Takayuki Kanda, Masahiro Shiomi, Zenta Miyashita, Hiroshi Ishiguro, and Norihiro Hagita. A communication robot in a shopping mall. *IEEE Transactions on Robotics*, 26(5):897–913, 2010.

[79] Takayuki Kanda, Masahiro Shiomi, Zenta Miyashita, Hiroshi Ishiguro, and Norihiro Hagita. An affective guide robot in a shopping mall. In *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*, pages 173–180, 2009.

[80] Guy Hoffman. Anticipation in Human-Robot Interaction. *2010 AAAI Spring Symposium Series*, pages 21–26, 2010.

[81] Jong-Hoon Lee, Jin-Yung Park, and Tek-Jin Nam. Emotional interaction through physical movement. In *International Conference on Human-Computer Interaction*, pages 401–410. Springer, 2007.

[82] Drazen Brscić, Hiroyuki Kidokoro, Yoshitaka Suehiro, and Takayuki Kanda. Escaping from children's abuse of social robots. In *Proceedings of the tenth annual acm/ieee international conference on human-robot interaction*, pages 59–66. ACM, 2015.

[83] Rachel M Holladay, Anca D Dragan, and Siddhartha S Srinivasa. Legible robot pointing. In *Robot and Human Interactive Communication, 2014 RO-MAN: The 23rd IEEE International Symposium on*, pages 217–223. IEEE, 2014.

[84] Henny Admoni, Thomas Weng, Bradley Hayes, and Brian Scassellati. Robot nonverbal behavior improves task performance in difficult collaborations. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*, pages 51–58. IEEE Press, 2016.

[85] Travis J Wiltshire, Emilio JC Lobato, Anna V Wedell, Wes Huang, Benjamin Axelrod, and Stephen M Fiore. Effects of robot gaze and proxemic behavior on perceived social presence during a hallway navigation scenario. In *Proceedings of the human factors and ergonomics society annual meeting*, volume 57, pages 1273–1277. SAGE Publications Sage CA: Los Angeles, CA, 2013.

[86] Zsolt Kira and Mitchell A Potter. Exerting human control over decentralized robot swarms. In *Autonomous Robots and Agents, 2009. ICARA 2009. 4th International Conference on*, pages 566–571. IEEE, 2009.

[87] Andreas Kolling, Steven Nunnally, and Michael Lewis. Towards human control of robot swarms. In *Proceedings of the seventh annual ACM/IEEE international conference on human-robot interaction*, pages 89–96. ACM, 2012.

[88] Ellen C Haas, K Pillalamarri, CC Stachowiak, and M Fields. Multimodal controls for soldier/swarm interaction. In *2011 RO-MAN*, pages 223–228. IEEE, 2011.

[89] Jacques Penders, Lyuba Alboul, Ulf Witkowski, Amir Naghsh, Joan Saez-Pons, Stefan Herbrechtsmeier, and Mohamed El-Habbal. A robot swarm assisting a human firefighter. *Advanced Robotics*, 25(1-2):93–117, 2011.

[90] Phillip Walker, Saman Amirpour Amraii, Nilanjan Chakraborty, Michael Lewis, and Katia Sycara. Human control of robot swarms with dynamic leaders. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1108–1113. IEEE, 2014.

[91] Javier Alonso-Mora, Andreas Breitenmoser, Martin Rufli, Roland Siegwart, and Paul Beardsley. Multi-robot system for artistic pattern formation. In *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pages 4512–4517. IEEE, 2011.

[92] Sandro Hauri, Javier Alonso-Mora, Andreas Breitenmoser, Roland Siegwart, and Paul Beardsley. Multi-robot formation control via a real-time drawing interface. In *Field and Service Robotics*, pages 175–189. Springer, 2014.

[93] R Grieder, J Alonso-Mora, C Bloechlinger, R Siegwart, and P Beardsley. Multi-robot control and interaction with a hand-held tablet. In *Workshop Proc. Int. Conf. Robotics and Automation*, volume 131. Citeseer, 2014.

[94] Javier Alonso-Mora, S Haegeli Lohaus, P Leemann, R Siegwart, and P Beardsley. Gesture based human-multi-robot swarm interaction and its application to an interactive display. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5948–5953. IEEE, 2015.

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

[96] Ayberk Özgür, Wafa Johal, Francesco Mondada, and Pierre Dillenbourg. Haptic-enabled handheld mobile robots: Design and analysis. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 2449–2461. ACM, 2017.

[97] Fritz Heider and Marianne Simmel. An experimental study of apparent behavior. *The American journal of psychology*, 57(2):243–259, 1944.

[98] Jong-hoon Lee, Jin-yung Park, and Tek-jin Nam. Emotional Interaction Through Physical Movement. *Human-Computer Interaction. HCI Intelligent Multimodal Interaction Environments*, 4552:401–410, 2007.

[99] Martin Saerbeck and Christoph Bartneck. Perception of affect elicited by robot motion. *Proceeding of the 5th ACM/IEEE international conference on Human-robot interaction - HRI '10*, page 53, 2010.

[100] Jonathan Posner, James A Russell, and Bradley S Peterson. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and psychopathology*, 17(3):715–734, 2005.

[101] Marina Pavlova, Arseny A Sokolov, and Alexander Sokolov. Perceived dynamics of static images enables emotional attribution. *Perception*, 34(9):1107–1116, 2005. PMID: 16247880.

[102] Frank E. Pollick, Helena M. Paterson, Armin Bruderlin, and Anthony J. Sanford. Perceiving affect from arm movement. *Cognition*, 82(2):51–61, 2001.

[103] Bernard Rimé, Bernadette Boulanger, Philippe Laubin, Marc Richir, and Kathleen Stroobants. The perception of interpersonal emotions originated by patterns of movement. *Motivation and Emotion*, 9(3):241–260, 1985.

[104] Misako Sawada, Kazuhiro Suda, and Motonobu Ishii. Expression of emotions in dance: Relation between arm movement characteristics and emotion. *Perceptual and Motor Skills*, 97(3):697–708, 2003.

[105] Gaëtan Podevijn, Rehan O 'grady, Nithin Mathews, Audrey Gilles, Carole Fantini-Hauwel, and Marco Dorigo. Investigating the effect of increasing robot group sizes on the human psychophysiological state in the context of human–swarm interaction. *Swarm Intelligence*, 10:193–210, 2016.

[106] Griffin Dietz, Jane L. E., Peter Washington, Lawrence H Kim, and Sean Follmer. Human Perception of Swarm Robot Motion. *CHI Extended Abstracts*, 2017.

[107] Gergely Csibra and György Gergely. 'obsessed with goals': Functions and mechanisms of teleological interpretation of actions in humans. *Acta psychologica*, 124(1):60–78, 2007.

[108] Christopher Bodden, Daniel Rakita, Bilge Mutlu, and Michael Gleicher. A flexible optimization-based method for synthesizing intent-expressive robot arm motion. *The International Journal of Robotics Research*, 37(11):1376–1394, 2018.

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

[110] Stefanos Nikolaidis, Anca Dragan, and Siddharta Srinivasa. Viewpoint-based legibility optimization. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*, HRI '16, page 271–278. IEEE Press, 2016.

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

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

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

[114] Tara Matthews. Designing and evaluating glanceable peripheral displays. In *Proceedings of the 6th conference on Designing Interactive systems*, pages 343–345, 2006.

[115] S Boerema, Randy Klaassen, Hendrikus JA op den Akker, and Hermanus J Hermens. Glanceability evaluation of a physical activity feedback system for office workers. *Proceedings of EHST*, pages 52–57, 2012.

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

[117] Christopher Healey and James Enns. Attention and visual memory in visualization and computer graphics. *IEEE transactions on visualization and computer graphics*, 18(7):1170–1188, 2011.

[118] Howard E Egeth and Steven Yantis. Visual attention: Control, representation, and time course. *Annual review of psychology*, 48(1):269–297, 1997.

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

[120] Drone light shows powered by intel.

[121] Robin IM Dunbar. The social role of touch in humans and primates: behavioural function and neurobiological mechanisms. *Neuroscience & Biobehavioral Reviews*, 34(2):260–268, 2010.

[122] Tiffany Field. Touch for socioemotional and physical well-being: A review. *Developmental Review*, 30(4):367–383, 2010.

[123] Haruaki Fukuda, Masahiro Shiomi, Kayako Nakagawa, and Kazuhiro Ueda. 'midas touch'in human-robot interaction: evidence from event-related potentials during the ultimatum game. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, pages 131–132. ACM, 2012.

[124] Takahiro Hirano, Masahiro Shiomi, Takamasa Iio, Mitsuhiko Kimoto, Ivan Tanev, Katsunori Shimohara, and Norihiro Hagita. How do communication cues change impressions of human–robot touch interaction? *International Journal of Social Robotics*, 10(1):21–31, 2018.

[125] Tiffany L Chen, Chih-Hung Aaron King, Andrea L Thomaz, and Charles C Kemp. An investigation of responses to robot-initiated touch in a nursing context. *International Journal of Social Robotics*, 6(1):141–161, 2014.

[126] Andreas Kolling, Katia Sycara, Steven Nunnally, and Michael Lewis. Human-swarm interaction: An experimental study of two types of interaction with foraging swarms. *Journal of Human-Robot Interaction*, 2(2):103–129, 2013.

[127] Adrian Stoica, Theodoros Theodoridis, Huosheng Hu, Klaus McDonald-Maier, and David F Barrero. Towards human-friendly efficient control of multi-robot teams. In *2013 International Conference on Collaboration Technologies and Systems (CTS)*, pages 226–231. IEEE, 2013.

[128] Jun Kato, Daisuke Sakamoto, Masahiko Inami, and Takeo Igarashi. Multi-touch interface for controlling multiple mobile robots. In *CHI'09 Extended Abstracts on Human Factors in Computing Systems*, pages 3443–3448. ACM, 2009.

[129] Jawad Nagi, Alessandro Giusti, Luca M Gambardella, and Gianni A Di Caro. Human-swarm interaction using spatial gestures. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3834–3841. IEEE, 2014.

[130] Ryo Suzuki, Jun Kato, Mark D Gross, and Tom Yeh. Reactile: Programming swarm user interfaces through direct physical manipulation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 199. ACM, 2018.

[131] Michael D Good, John A Whiteside, Dennis R Wixon, and Sandra J Jones. Building a user-derived interface. *Communications of the ACM*, 27(10):1032–1043, 1984.

[132] Jacob O Wobbrock, Meredith Ringel Morris, and Andrew D Wilson. User-defined gestures for surface computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1083–1092. ACM, 2009.

[133] Meredith Ringel Morris, Jacob O Wobbrock, and Andrew D Wilson. Understanding users' preferences for surface gestures. In *Proceedings of graphics interface 2010*, pages 261–268. Canadian Information Processing Society, 2010.

[134] Parastoo Abtahi, David Y Zhao, Jane L. E, and James A Landay. Drone near me: Exploring touch-based human-drone interaction. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(3):34, 2017.

[135] Ekaterina Peshkova and Martin Hitz. Exploring user-defined gestures to control a group of four uavs. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 169–174. IEEE, 2017.

[136] Geraint Jones, Nadia Berthouze, Roman Bielski, and Simon Julier. Towards a situated, multimodal interface for multiple uav control. In *2010 IEEE International Conference on Robotics and Automation*, pages 1739–1744. IEEE, 2010.

[137] Hans Christoph Nothdurft. The role of features in preattentive vision: Comparison of orientation, motion and color cues. *Vision Research*, 33(14):1937–1958, 1993.

[138] Ken Nakayama and Gerald H Silverman. Serial and parallel processing of visual feature conjunctions. *Nature*, 320(6059):264–265, 1986.

[139] Martin Böcker. A multiple index approach for the evaluation of pictograms and icons. *Computer Standards and Interfaces*, 18(2):107–115, 1996.

[140] Ginevra Castellano, Santiago D. Villalba, and Antonio Camurri. Recognising Human Emotions from Body Movement and Gesture Dynamics. *Affective Computing and Intelligent Interaction*, pages 71–82, 2007.

[141] Johnny C Lee, Scott E Hudson, Jay W Summet, and Paul H Dietz. Moveable interactive projected displays using projector based tracking. In *Proceedings of the 18th annual ACM symposium on User interface software and technology*, pages 63–72. ACM, 2005.

[142] Johnny C. Lee, Paul H. Dietz, Dan Maynes-Aminzade, Ramesh Raskar, and Scott E. Hudson. Automatic projector calibration with embedded light sensors. In *Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology*, UIST '04, pages 123–126, New York, NY, USA, 2004. ACM.

[143] Woohun Lee and Jun Park. Augmented Foam: A Tangible Augmented Reality for Product Design. *Symposium on Mixed and Augmented Reality*, 2005.

[144] Harold W Kuhn. The hungarian method for the assignment problem. *Naval research logistics quarterly*, 2(1-2):83–97, 1955.

[145] Jamie Snape, Jan van den Berg, Stephen J Guy, and Dinesh Manocha. The hybrid reciprocal velocity obstacle. *Robotics, IEEE Transactions on*, 27(4):696–706, 2011.

[146] Jamie Snape, Jur P van den Berg, Stephen J Guy, and Dinesh Manocha. Independent navigation of multiple mobile robots with hybrid reciprocal velocity obstacles. In *IROS*, pages 5917–5922, 2009.

[147] Florent Levillain, David St-Onge, Elisabetta Zibetti, and Giovanni Beltrame. More than the sum of its parts: Assessing the coherence and expressivity of a robotic swarm. In *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 583–588. IEEE, 2018.

[148] David St-Onge, Ulysse Côté-Allard, Kyrre Glette, Benoit Gosselin, and Giovanni Beltrame. Engaging with robotic swarms: Commands from expressive motion. *ACM Transactions on Human-Robot Interaction (THRI)*, 8(2):11, 2019.

[149] David St-Onge, Jing Yang Kwek, and Giovanni Beltrame. Behaviours and states for human-swarm interaction studies. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, pages 42–42. ACM, 2018.

[150] Jingyi Li, Michael Wessely, Sean Follmer, and Stefanie Mueller. Summer school for computational fabrication and smart matter. *IEEE Pervasive Computing*, 16(4):50–53, 2017.

[151] J. Alonso-Mora, A. Breitenmoser, M. Rufli, R. Siegwart, and P. Beardsley. Multi-robot system for artistic pattern formation. In *2011 IEEE International Conference on Robotics and Automation*, pages 4512–4517, May 2011.

[152] Sean Kerman, Daniel Brown, and Michael A. Goodrich. Supporting human interaction with robust robot swarms. *Proceedings - 2012 5th International Symposium on Resilient Control Systems, ISRCS 2012*, pages 197–202, 2012.

[153] Ds Daniel S Brown and Michael a Ma Goodrich. Limited bandwidth recognition of collective behaviors in bio-inspired swarms. *International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 405–412, 2014.

[154] Adriane E Seiffert, Sean Timothy Hayes, Caroline E Harriott, and Julie A Adams. Motion perception of biological swarms. In *CogSci*, 2015.

[155] M Bradley and Peter J Lang. Measuring Emotion: The Self-Assessment Semantic Differential Manikin and the SEMANTIC DIFFERENTIAL. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(I):49–59, 1994.

[156] Marc Hassenzahl. The interplay of beauty, goodness, and usability in interactive products. *Human Computer Interaction*, 19(4):319–349, December 2004.

[157] Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, 1(1):71–81, 2009.

[158] Gary Charness, Uri Gneezy, and Michael A. Kuhn. Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior and Organization*, 81(1):1–8, 2012.

[159] Sarah N Woods, Michael L Walters, Kheng Lee Koay, and Kerstin Dautenhahn. Methodological issues in hri: A comparison of live and video-based methods in robot to human approach direction trials. In *Robot and Human Interactive Communication. ROMAN*, pages 51–58. IEEE, 2006.

[160] Qianli Xu, J S L Ng, Yian Ling Cheong, O Y Tan, Ji Bin Wong, B T C Tay, and Taezoon Park. Effect of Scenario Media on Elder Adults' Evaluation of Human-Robot Interaction. *Human-Robot Interaction (HRI), 2012 7th ACM/IEEE International Conference on*, (Figure 2):275–277, 2012.

[161] Wilma A. Bainbridge, Justin W. Hart, Elizabeth S. Kim, and Brian Scassellati. The benefits of interactions with physically present robots over video-displayed agents. *International Journal of Social Robotics*, 3(1):41–52, 2011.

[162] Esben W Pedersen, Sriram Subramanian, and Kasper Hornbæk. Is my phone alive?: a large-scale study of shape change in handheld devices using videos. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pages 2579–2588. ACM, 2014.

[163] Albert Mehrabian. Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in Temperament. *Current Psychology*, 14(4):261–292, 1996.

[164] John Lasseter. Principles of traditional animation applied to 3d computer animation. In *ACM Siggraph Computer Graphics*, volume 21, pages 35–44. ACM, 1987.

[165] Anca Dragan and Siddhartha Srinivasa. Generating Legible Motion. *Robotics: Science and Systems IX*, 2013.

[166] Eric Bonabeau, Marco Dorigo, and Guy Theraulaz. *Swarm intelligence: from natural to artificial systems*. Number 1. Oxford university press, 1999.

[167] Andrew Witkin and Michael Kass. Spacetime constraints. *ACM Siggraph Computer Graphics*, 22(4):159–168, 1988.

[168] Mrinal Kalakrishnan, Sachin Chitta, Evangelos Theodorou, Peter Pastor, and Stefan Schaal. Stomp: Stochastic trajectory optimization for motion planning. In *2011 IEEE international conference on robotics and automation*, pages 4569–4574. IEEE, 2011.

[169] Nathan Ratliff, Matt Zucker, J Andrew Bagnell, and Siddhartha Srinivasa. Chomp: Gradient optimization techniques for efficient motion planning. 2009.

[170] Ronald L Graham, Boris D Lubachevsky, Kari J Nurmela, and Patric RJ Östergård. Dense packings of congruent circles in a circle. *Discrete Mathematics*, 181(1-3):139–154, 1998.

[171] Atsuyuki Okabe and Atsuo Suzuki. Locational optimization problems solved through voronoi diagrams. *European Journal of Operational Research*, 98(3):445–456, 1997.

[172] Jorge Cortes, Sonia Martinez, Timur Karatas, and Francesco Bullo. Coverage control for mobile sensing networks. *IEEE Transactions on robotics and Automation*, 20(2):243–255, 2004.

[173] Jessi Stark, Roberta RC Mota, and Ehud Sharlin. Personal space intrusion in human-robot collaboration. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, pages 245–246. ACM, 2018.

[174] Juulia T Suvilehto, Enrico Glerean, Robin IM Dunbar, Riitta Hari, and Lauri Nummenmaa. Topography of social touching depends on emotional bonds between humans. *Proceedings of the National Academy of Sciences*, 112(45):13811–13816, 2015.

[175] Ali Israr and Ivan Poupyrev. Tactile brush: drawing on skin with a tactile grid display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2019–2028. ACM, 2011.

[176] Bernhard Weber, Simon Schätzle, Thomas Hulin, Carsten Preusche, and Barbara Deml. Evaluation of a vibrotactile feedback device for spatial guidance. In *World Haptics Conference (WHC), 2011 IEEE*, pages 349–354. IEEE, 2011.

[177] Karlin Bark, Jason Wheeler, Pete Shull, Joan Savall, and Mark Cutkosky. Rotational skin stretch feedback: A wearable haptic display for motion. *IEEE Transactions on Haptics*, 3(3):166–176, 2010.

[178] Artem Dementyev, Hsin-Liu Cindy Kao, Inrak Choi, Deborah Ajilo, Maggie Xu, Joseph A Paradiso, Chris Schmandt, and Sean Follmer. Rovables: Miniature on-body robots as mobile wearables. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pages 111–120. ACM, 2016.

[179] Gijs Huisman, Aduen Darriba Frederiks, Betsy Van Dijk, Dirk Hevlen, and Ben Krose. The tasst: Tactile sleeve for social touch. In *World Haptics Conference (WHC), 2013*, pages 211–216. IEEE, 2013.

[180] Alexandra Delazio, Ken Nakagaki, Roberta L Klatzky, Scott E Hudson, Jill Fain Lehman, and Alanson P Sample. Force jacket: Pneumatically-actuated jacket for embodied haptic experiences. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 320. ACM, 2018.

[181] Lawrence H Kim, Pablo Castillo, Sean Follmer, and Ali Israr. Vps tactile display: Tactile information transfer of vibration, pressure, and shear. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 3(2):51, 2019.

[182] Sean Follmer, Daniel Leithinger, Alex Olwal, Akimitsu Hogge, and Hiroshi Ishii. inform: dynamic physical affordances and constraints through shape and object actuation. In *Uist*, volume 13, pages 417–426, 2013.

[183] Sungjune Jang, Lawrence H Kim, Kesler Tanner, Hiroshi Ishii, and Sean Follmer. Haptic edge display for mobile tactile interaction. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 3706–3716. ACM, 2016.

[184] Alexa F Siu, Eric J Gonzalez, Shenli Yuan, Jason B Ginsberg, and Sean Follmer. Shapeshift: 2d spatial manipulation and self-actuation of tabletop shape displays for tangible and haptic interaction. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2018.

[185] Inrak Choi, Elliot W Hawkes, David L Christensen, Christopher J Ploch, and Sean Follmer. Wolverine: A wearable haptic interface for grasping in virtual reality. In

Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on, pages 986–993. IEEE, 2016.

[186] Jrme Pasquero, Joseph Luk, Vincent Lévesque, Qi Wang, Vincent Hayward, and Karon E MacLean. Haptically enabled handheld information display with distributed tactile transducer. *IEEE Transactions on Multimedia*, 9(4):746–753, 2007.

[187] Thomas H Massie, J Kenneth Salisbury, et al. The phantom haptic interface: A device for probing virtual objects. In *Proceedings of the ASME winter annual meeting, symposium on haptic interfaces for virtual environment and teleoperator systems*, volume 55, pages 295–300. Citeseer, 1994.

[188] Yasuyoshi Yokokohji, Ralph L Hollis, and Takeo Kanade. Wysiwyf display: A visual/haptic interface to virtual environment. *Presence*, 8(4):412–434, 1999.

[189] Pascal Knierim, Thomas Kosch, Valentin Schwind, Markus Funk, Francisco Kiss, Stefan Schneegass, and Niels Henze. Tactile drones-providing immersive tactile feedback in virtual reality through quadcopters. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pages 433–436. ACM, 2017.

[190] Muhammad Abdullah, Minji Kim, Waseem Hassan, Yoshihiro Kuroda, and Seokhee Jeon. Hapticdrone: An encountered-type kinesthetic haptic interface with controllable force feedback: Initial example for 1d haptic feedback. In *Adjunct Publication of the 30th Annual ACM Symposium on User Interface Software and Technology*, pages 115–117. ACM, 2017.

[191] Parastoo Abtahi, Benoit Landry, Jackie Yang, Marco Pavone, Sean Follmer, and James A Landay. Beyond the force: Using quadcopters to appropriate objects and the environment for haptics in virtual reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2019.

[192] Moshe Solomonow, John Lyman, and Amos Freedy. Electrotactile two-point discrimination as a function of frequency, body site, laterality, and stimulation codes. *Annals of biomedical engineering*, 5(1):47–60, 1977.

[193] Irene Rae, Leila Takayama, and Bilge Mutlu. The influence of height in robot-mediated communication. In *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction*, pages 1–8. IEEE Press, 2013.

[194] Greg K Essick, Anuj James, and Francis P McGlone. Psychophysical assessment of the affective components of non-painful touch. *Neuroreport*, 10(10):2083–2087, 1999.

[195] Gustave Le Bon. *The crowd: A study of the popular mind*. Fischer, 1897.

[196] Gilles Bailly, Sidharth Sahdev, Sylvain Malacria, and Thomas Pietrzak. Livingdesktop: Augmenting desktop workstation with actuated devices. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 5298–5310. ACM, 2016.

[197] Johnny Chung Lee. Projector-based location discovery and tracking. *Human Computer Interaction Institute. School of Computer Science*, 2008.

[198] Alan T Asbeck, Sangbae Kim, Mark R Cutkosky, William R Provancher, and Michele Lanzetta. Scaling hard vertical surfaces with compliant microspine arrays. *The International Journal of Robotics Research*, 25(12):1165–1179, 2006.

[199] Miyuki Morioka and Michael J Griffin. Absolute thresholds for the perception of fore-and-aft, lateral, and vertical vibration at the hand, the seat, and the foot. *Journal of sound and vibration*, 314(1-2):357–370, 2008.

[200] Hasti Seifi and Karon E Maclean. A first look at individuals' affective ratings of vibrations. In *World Haptics Conference (WHC), 2013*, pages 605–610. IEEE, 2013.

[201] Antonella Mazzoni and Nick Bryan-Kinns. How does it feel like? an exploratory study of a prototype system to convey emotion through haptic wearable devices. In *Intelligent Technologies for Interactive Entertainment (INTETAIN), 2015 7th International Conference on*, pages 64–68. IEEE, 2015.

[202] Marianna Obrist, Sriram Subramanian, Elia Gatti, Benjamin Long, and Thomas Carter. Emotions mediated through mid-air haptics. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 2053–2062. ACM, 2015.

[203] Olivier Bau, Ivan Poupyrev, Ali Israr, and Chris Harrison. Teslatouch: electrovibration for touch surfaces. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology*, pages 283–292. ACM, 2010.

[204] Gaetano Valenza, Alberto Greco, Mimma Nardelli, Matteo Bianchi, Antonio Lanata, Simone Rossi, and Enzo Pasquale Scilingo. Electroencephalographic spectral correlates of caress-like affective haptic stimuli. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, pages 4733–4736. IEEE, 2015.

[205] Margaret M Bradley and Peter J Lang. Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1):49–59, 1994.

[206] Albert Mehrabian. Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. *Current Psychology*, 14(4):261–292, 1996.

[207] Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics*, 1(1):71–81, 2009.

[208] Esben W Pedersen, Sriram Subramanian, and Kasper Hornbæk. Is my phone alive?: a large-scale study of shape change in handheld devices using videos. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pages 2579–2588. ACM, 2014.

[209] Stanley E Jones and A Elaine Yarbrough. A naturalistic study of the meanings of touch. *Communications Monographs*, 52(1):19–56, 1985.

[210] Paul Ekman. Are there basic emotions? 1992.

[211] Karon E MacLean. Designing with haptic feedback. In *Robotics and Automation, 2000. Proceedings. ICRA'00. IEEE International Conference on*, volume 1, pages 783–788. IEEE, 2000.

[212] Erol Şahin. Swarm robotics: From sources of inspiration to domains of application. In *International workshop on swarm robotics*, pages 10–20. Springer, 2004.

[213] Andreas Kolling, Phillip Walker, Nilanjan Chakraborty, Katia Sycara, and Michael Lewis. Human interaction with robot swarms: A survey. *IEEE Transactions on Human-Machine Systems*, 46(1):9–26, 2015.

[214] Leila Takayama and Caroline Pantofaru. Influences on proxemic behaviors in human-robot interaction. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5495–5502. IEEE, 2009.

[215] Jonathan Mumm and Bilge Mutlu. Human-robot proxemics: physical and psychological distancing in human-robot interaction. In *Proceedings of the 6th international conference on Human-robot interaction*, pages 331–338. ACM, 2011.

[216] Luca Consolini, Fabio Morbidi, Domenico Prattichizzo, and Mario Tosques. Leader–follower formation control of nonholonomic mobile robots with input constraints. *Automatica*, 44(5):1343–1349, 2008.

[217] William M Spears, Diana F Spears, Jerry C Hamann, and Rodney Heil. Distributed, physics-based control of swarms of vehicles. *Autonomous Robots*, 17(2-3):137–162, 2004.

[218] Antonio Gomes, Calvin Rubens, Sean Braley, and Roel Vertegaal. Bitdrones: Towards using 3d nanocopter displays as interactive self-levitating programmable matter. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 770–780. ACM, 2016.

[219] J Triesch. Ch. von der malsburg. a gesture interface for human-robot interaction. In *Proc. of the 2nd Int. Conf. on Automatic Face and Gesture Recognition*, pages 546–551, 1998.

[220] Shiva Shahrokhi and Aaron T Becker. Object manipulation and position control using a swarm with global inputs. In *2016 IEEE International Conference on Automation Science and Engineering (CASE)*, pages 561–566. IEEE, 2016.

[221] Ian M Bullock and Aaron M Dollar. Classifying human manipulation behavior. In *2011 IEEE International Conference on Rehabilitation Robotics*, pages 1–6. IEEE, 2011.

[222] John R Searle. A taxonomy of illocutionary acts. 1975.

[223] Jacob O Wobbrock, Htet Htet Aung, Brandon Rothrock, and Brad A Myers. Maximizing the guessability of symbolic input. In *CHI'05 extended abstracts on Human Factors in Computing Systems*, pages 1869–1872. ACM, 2005.

[224] John H McDonald. *Handbook of biological statistics*, volume 2. sparky house publishing Baltimore, MD, 2009.

[225] Andrew D Wilson, Shahram Izadi, Otmar Hilliges, Armando Garcia-Mendoza, and David Kirk. Bringing physics to the surface. In *Proceedings of the 21st annual ACM symposium on User interface software and technology*, pages 67–76. ACM, 2008.

[226] George W. Furnas, Thomas K. Landauer, Louis M. Gomez, and Susan T. Dumais. The vocabulary problem in human-system communication. *Communications of the ACM*, 30(11):964–971, 1987.

[227] Vagelis Plevris, Matthew G Karlaftis, and Nikos D Lagaros. A swarm intelligence approach for emergency infrastructure inspection scheduling. In *Sustainable and Resilient Critical Infrastructure Systems*, pages 201–230. Springer, 2010.

[228] Mathieu Le Goc, Charles Perin, Sean Follmer, Jean-Daniel Fekete, and Pierre Dragicevic. Dynamic composite data physicalization using wheeled micro-robots. *IEEE transactions on visualization and computer graphics*, 25(1):737–747, 2018.

[229] Miguel Duarte, Jorge Gomes, Vasco Costa, Tiago Rodrigues, Fernando Silva, Víctor Lobo, Mario Monteiro Marques, Sancho Moura Oliveira, and Anders Lyhne Christensen. Application of swarm robotics systems to marine environmental monitoring. In *OCEANS 2016-Shanghai*, pages 1–8. IEEE, 2016.

[230] Belkacem Khaldi and Foudil Cherif. An overview of swarm robotics: Swarm intelligence applied to multi-robotics. *International Journal of Computer Applications*, 126(2), 2015.

[231] Jordi Palacin, José Antonio Salse, Ignasi Valgañón, and Xavi Clua. Building a mobile robot for a floor-cleaning operation in domestic environments. *IEEE Transactions on instrumentation and measurement*, 53(5):1418–1424, 2004.

[232] Phillip Walker, Steven Nunnally, Mike Lewis, Andreas Kolling, Nilanjan Chakraborty, and Katia Sycara. Neglect benevolence in human control of swarms in the presence of latency. In *2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 3009–3014. IEEE, 2012.

[233] Meredith Ringel Morris, Andreea Danielescu, Steven Drucker, Danyel Fisher, Bongshin Lee, Jacob O Wobbrock, et al. Reducing legacy bias in gesture elicitation studies. *interactions*, 21(3):40–45, 2014.

[234] Jane L E, Ilene L E, James A Landay, and Jessica R Cauchard. Drone & wo: Cultural influences on human-drone interaction techniques. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 6794–6799. ACM, 2017.

[235] Dingjun Li, PL Patrick Rau, and Ye Li. A cross-cultural study: Effect of robot appearance and task. *International Journal of Social Robotics*, 2(2):175–186, 2010.

[236] Ron Pelrine, Annjoe Wong-Foy, Allen Hsu, and Brian McCoy. Self-assembly of milliscale robotic manipulators: A path to highly adaptive, robust automation systems. In *2016 International Conference on Manipulation, Automation and Robotics at Small Scales (MARSS)*, pages 1–6. IEEE, 2016.

[237] Elyse DZ Chase and Sean Follmer. Differences in haptic and visual perception of expressive 1dof motion. In *ACM Symposium on Applied Perception 2019*, pages 1–9, 2019.

[238] Nina Riether, Frank Hegel, Britta Wrede, and Gernot Horstmann. Social facilitation with social robots? In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, pages 41–48. ACM, 2012.