Proxemic Reasoning for Group Approach

Ben Wright[1], Magdalena Bugajska[2], William Adams[2], Ed Lawson[2], J. Malcolm McCurry[3], and J. Gregory Trafton[2]

[1] NRC Postdoctoral Researcher, U.S. Naval Research Laboratory
benjamin.wright.ctr@nrl.navy.mil
[2] Navy Center for Applied Research in AI, U.S. Naval Research Laboratory
first.last@nrl.navy.mil
[3] Peraton jmccurry@peraton.com, Washington, D.C.

Abstract. How should a robot approach a group of people? To answer this, we examine a large real-world data set of approximately 1500 individuals and find that, contrary to previous models, people do not simply approach an individual within the group. They use features of the group to determine aspects of their approach. To explain and test these features, we formalize an Approach Problem and develop several strategies of determining approach to a group. These strategies are then compared to the empirical data. One strategy performed consistent with the empirical data, which we then demonstrate on a robot platform.

1 Introduction and Background

As robots become more common in everyday environments, they will need to interact with people more often; indeed the entire field of Human Robot Interaction (HRI) is working toward this goal. One key aspect of HRI is understanding how a robot should approach and join an individual or group of people. For example, if a robot has a message or information for the group, the robot will need to determine the best way to become part of the group in order to communicate with the group. Because the robot will be interacting with people, it will need to make decisions based on how people will react to it, rather than what is most optimal for the robot. For example, moving quickly may be an excellent method of traversing from one location to another, but people are typically not comfortable with people or robots approaching them at a high speed.

We believe that approaching a group may require paying attention to the group that the individual is part of in order to successfully be integrated into a group. For example, approaching a group may require coming to the outskirts of a group before being accepted into the group. The size of the group itself may also impact how close that original approach should be. In order to understand some of the features and representations of approaching a group, we analyzed a large-scale database that includes many instances of people approaching groups of different sizes in a real-world domain.

Previous researchers have focused primarily on how an individual approaches another individual. Hall proposed that individuals have a sense of space around them, and it is possible to systematically characterize some forms of relationships and familiarity based on how close another individual stands [3]. For example, an individual standing at an "intimate distance" is typically within 1 to 18 inches and is a close friend or lover. Personal distance goes from 1.5 feet to 4 feet and characterizes friends and family. Social distance (4 to 12 feet) is for interactions among acquaintances, and distances farther than 12 feet is typically for public speaking. These numbers can change a bit across cultures (e.g., Asian cultures are typically closer overall) [7] and are impacted by the auditory environment (e.g., in a library people will gather closer to speak quietly) and the social environment (e.g., at an outdoor rock concert, people will be close together). This entire field is called proxemics and has been the foundation of a great deal of work about how people and robots approach another individual.

Because of the focus on the individual in previous human empirical work, the research on robots approaching people has likewise focused on a single robot approaching a single individual [8, 6, 11]. For example, [8] provided an algorithm and description that took into account the auditory scene and incorporated that information with proxemic knowledge to approach an individual in a human-natural manner – when the environment was noisy, the robot approached closer than it would normally in order to be heard. Other researchers have focused on spatially positioning a robot. Kendon showed that many groups gather in systematic manners (e.g., a circle or a line) and that individuals within those groups typically face toward the middle and their conversational partner(s) [7]. Robots that face inward and towards conversational partners are perceived as part of the group and engender more inclusion than robots that do not [14].

While there have been previous studies on approaching pairs of people [5, 2], studies on approaches of larger groups remain scarce. Most research on approaching a group actually focuses on approaching an individual that happens to be in a group. In excellent early work, Althaus and colleagues realized that joining a group meant both approaching and joining, so their robot approached a specific individual within a group, slowed down upon approach, and then integrated into the group itself [1]. Kato and colleagues used an individual's perceived intention to approach an individual who may be part of a group[6]. Similarly, Satake and colleagues selected a target person approached the individual and non-verbally showed its intention to interact [12]. Satake also showed that it is quite difficult for a robot to simply approach a person: It is not a matter of sending a robot within X meters of an individual – pose, distance, speed, and direction are all important for whether or not the approach is successful.

2 Dataset and Analysis

We used the Stanford Drone Dataset (SDD)[4] [10] as a foundation to our understanding of how individuals approach groups. The SDD consists of an overhead view of eight different locations around the Stanford University campus, each containing video and created annotations.

[4] `http://cvgl.stanford.edu/projects/uav_data/`

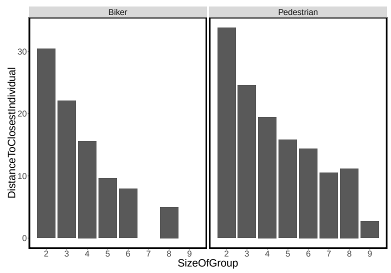
The annotations were extracted using computer vision techniques described fully in [10] and contained pedestrians, bikes, golf carts, skateboarders, and buses. Annotations consist of targets (objects moving around in the environment), their class label (e.g., pedestrian), and their x,y coordinates for each frame of the video. The dataset itself contains over 19K targets (primarily pedestrians) and consists of many different types of interactions, groups, and target motion. The SDD annotations created bounding boxes around each object; for convenience we collapsed each bounding box to its midpoint.

In order to identify when an individual approached a group, we needed to create several additional concepts and representations based on the existing dataset. Specifically, we needed to identify a *group* and when an individual *approached* the group. There are approaches to identifying when a group is formed [13, 7] but they typically require knowing what direction people are facing in addition to their x,y location. Because that information is computationally expensive, more difficult to extract and not available in the SDD, we created another approach to determining when a group has formed based on spatial proximity and temporal continuity. We defined a group as consisting of individuals who were within d distance of each other within t time, recursively. Thus, three individuals who were $\leq d$ distance of each other for $\leq t$ seconds was considered a group. This formulation prevented individuals who simply passed each other from becoming grouped together, as long as t was greater than their passing time. We also prevented subsets from being grouped together (e.g., if individuals A, B, and C were all a group, A and B were not a separate group). For this project, we set $d = 9$ feet which is near the top of the range for social distance (full social range made no difference to these results). We also set $t = 7$ seconds which allowed us to have some confidence that groups were relatively stable.

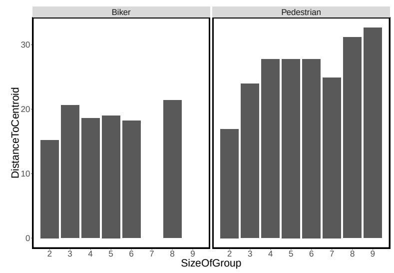
We also needed to identify when an individual approached a group. Consistent with [1], an individual who was defined as approaching a group needed to stop. To determine whether an individual had stopped, we calculated speed over the previous second. An individual who had a speed $< .2m/s$ was considered stopped (.2 seemed to capture noise from algorithm and individual swaying, since people are rarely completely motionless). Thus, an individual who had joined a group was an individual who had stopped and had just joined the edge of the group (by the group definition above).

Several other group annotations were also derived. Specifically, we calculated the *GroupSize* (the number of people in the group), the *DistanceToCentroid* (the distance of the approaching individual to the center of the group), and *DistanceToClosestIndividual* (the distance of the approaching individual to the closest individual already present in the group).

Recall that most work has focused on an individual (person or robot) approaching an individual, whether they are part of a group or not. Our hypothesis, however, is that an individual approaches a group differently than they approach an individual, taking characteristics of the group into account to change some aspect of their approach. The most obvious characteristic of a group is its size.



(a) Empirical results from SDD of distance to the closest member of the group. X-axis is size of the group. Y-axis is distance of approacher to closest member of group. Left side is bikers, right side is pedestrians.



(b) Empirical results from SDD of distance to centroid. X-axis is size of the group. Y-axis is distance of approacher to centroid of group. Left side is bikers, right side is pedestrians.

Fig. 1: Results from SDD

We therefore created two different metrics to examine how an individual approaches a group. The first measure is the DistanceToClosestIndividual (defined above). If an approaching individual takes characteristics of the group into account (e.g., the number of people in the group), there should be a systematic relationship between the group size and how close an individual approaches a current member of the group. If an approaching individual does not take group size into account (like current approaches), we should see no systematic relationship between group size and DistanceToClosestIndividual.

While there were many instances of different types of transports approaching others, we focus on two: pedestrians and bikers. We chose these two for a number of reasons. First, pedestrians were the most common, and bikers were very numerous. Second, because the results should generalize to mobile and social robots, pedestrians and bikers are most similar in form and speed to prospective robots (and in particular the robot we use later). Third, pedestrian and bikers had enough data in each group size – other transport types could be missing substantial data or only had data for limited group sizes. Finally, we should note that the results presented are remarkably consistent across transport types, suggesting that these findings should generalize to approaches of all classes. In the entire dataset, there were 323 groups that had bikers approach them and 1140 groups that had pedestrians approach them.

Figure 1a shows two subplots of the data for a biker and a pedestrian as they approached groups of different sizes. As Figure 1a suggests, there is a strong relationship between the size of the group and how close the approacher comes to the closest individual: the bigger the group, the closer the individual approaches. This is a remarkably strong and robust finding across both types of approachers. Figure 1b shows two subplots of the data for a biker and a pedestrian approaching groups of different sizes. As Figure 1a suggests, there is a medium-strength relationship between the size of the group and how close the approacher comes to the centroid of the group: the bigger the group, the farther from the centroid the individual approaches. This is a strong effect for pedestrians, but less strong for bikers.

Both these graphs and analyses show that when an individual approaches a group, the approacher unambiguously uses the size of the existing group to determine how they approach. Unlike previous work, the approacher does not simply approach an individual that happens to be part of the group, but rather they approach the group as a whole, sensitive to group features. Because we want a robot to approach a group of people in the same manner as a human would, we will endeavor to create a model that has similar characteristics to those we have just discussed.

3 Defining our model of Approach

Let us think about a specific instance that may happen in the Stanford Drone Dataset(SDD). You are on a college campus handing out flyers. There is a group of 3 students idly chatting next to a tree and you want to approach them. Where, in relation to the group, do you decide to stop in order to interact with the group?

Definition 1 (The Approach Problem). *The objective of an Approach Problem is to find a location near a group of agents. We denote* near *and* distance *in terms of grid cells and the location, loc_x is a cell in the grid. Formally, an approach problem P can be defined as a tuple $P =< G, Ag, Obj, S >$ where G is a square grid space NxN, Ag is a set of agent group members each with a location in the grid space; that is $\forall a \in Ag, loc_a \in G$. Obj is a set of objects in the grid space that can occupy a square, but is not part of the group; $\forall o \in Obj, loc_o \in G$. Finally, S the starting location,$loc_S \in G$, for the agent approaching the group. A solution to the approaching problem P is given by a location E such that it identifies a location that approaches the locations of Ag and is not already occupied. Therefore, $E \in G \cap Ag \cap Obj$.*

Let us take the colloquial example from above, where a person was going to approach a group of three near a tree. This is denoted as:

Example 1 (Approaching a Group of 3 Near a Tree).
$P =< S, G, Ag, Obj >$, $G = 6x6$, $S = \{5, 5\}$,
$Ag = \{P1, P2, P3\}$, with locations: $loc_{P1} = \{0, 1\}$, $loc_{P2} = \{0, 2\}$, $loc_{P3} = \{0, 3\}$,
$Obj = \{Tree\}$, with location: $loc_{Tree} = \{1, 3\}$.

From this setup of P, we have viable options for the endpoint, E, to be within the following: $\{(0, 0)$, $(0, 4)$, $(0, 5)$, $(1, 0)$, $(1, 1)$, $(1, 2)$, $(1, 4)$, $(1, 5)$, $(2, 0)$, $(2, 1)$, $(2, 2)$, $(2, 3)$, $(2, 4)$, $(2, 5)$, $(3, 0)$, $(3, 1)$, $(3, 2)$, $(3, 3)$, $(3, 4)$, $(3, 5)$, $(4, 0)$, $(4, 1)$, $(4, 2)$, $(4, 3)$, $(4, 4)$, $(4, 5)$, $(5, 0)$, $(5, 1)$, $(5, 2)$, $(5, 3)$, $(5, 4)\}$ (i.e., any location not already occupied). Any of the locations are valid given the problem description. It is not clear however, which of these would be acceptable for people or a good match of the data.

4 Forming Strategies to use in the Approach Model

To address the issue from the Example, while any of the given locations would satisfy the problem, there is at some level an ordering or preference based on what we saw from the SDD. Therefore, we need to focus on some issues centered around the datasets.

Given the substantial computational resources utilized on a mobile robot platform (e.g. perception and navigation), one of our goals is to maintain strategies that are not computationally expensive. With this goal in mind, we hope to match patterns from the SDD to strategies for the approach problem. Additionally, we want to assess what features alter our strategy solutions the most.

To address the first goal, finding computationally light strategies, a number of strategies were developed to test different possible solutions against the patterns from the SDD. These strategies broke down into two categories: consider a characteristic of the group, or determine an individual should be selected to approach. While it is our intention to approach groups as a whole, we want to utilize individual members of a group for some of the straight-forward strategies. For the individual, we had two different strategies for selection - random and closest. Random picks a member of the group randomly (RP). Closest picks the member of the group that is closest to the starting position (CP).

Once the strategy has picked an individual from the group, or uses the whole group, all strategies begin by creating an *adjacency set.*This adjacency set is effectively our satisfiable set of solutions similar to E in our Example. After calculating the adjacency set, our strategies diverge to pick a specific element in the set. These computationally light equations are strategies to pick a square from the Adjacency Set. Either pick a random square (RS) or pick the closest square (CS). Aside from these two approaches, we also developed a few group based strategies. The first being a simple, "*go to the center of the group*". Additionally, due to our Dataset results (see Figures 1a and 1b), strategies taking advantage of distance to centroid and proximity to other group members were made. The strategy related to centroid distance (CentDist) returns a square that is closest to the same distance as the group members from the center of the group. Our proximal strategy finds the minimum average distance between all group members $ProxDist(i)$ from the given Adjacency Set ($i \in AdjSet$). For most groups this is very similar to the Center strategy.Finally, we combine proximity distance and distance to centroid as a last strategy. Here, we prioritize squares that are the same distance away from the centroid as the group members while also being close to all other group members. We use a combination of both previous equations to do this. We calculate the Centroid Distance (CD) and subtract our proximal distance from that.

5 Experiment Design for Approach Strategies

With these strategies, we can begin to address the other questions posed in Section 3. That is, we need a way to see if any of our strategies match the behavior patterns found in the SDD. Additionally, does the group formation affect the strategy? What features of the problem space affect the behavior of these strategies?

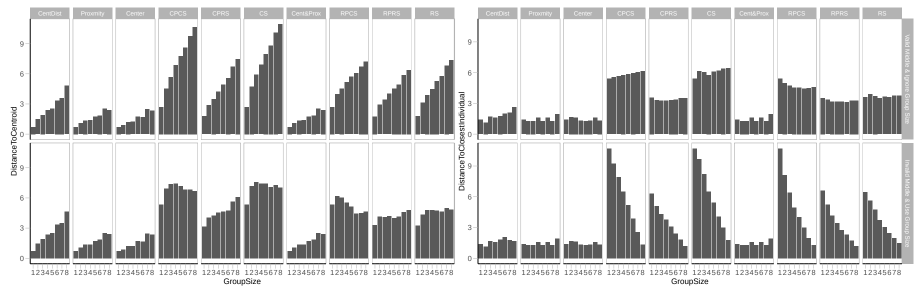
To answer these questions, we ran a simulation over different permutations. The setup for these simulations were as follows: G is a $100x100$ grid, Ag ranges from 1 agent, a, to 8 agents $\{a, b, c, d, e, f, g, h\}$, $Obj = \emptyset$, and S is randomly generated to be within G. This reasoning system was implemented in SWI-Prolog [15] and we call it the PardonMe-soner. The permutations involved 4 aspects. First, the size of the group changed. Second, the formation of the group changed. Third the spacing of the members of the group changed. Lastly, the *Closeness* which the approacher would consider was changed.

Based on our results from SDD, we found the majority of group size to be under 9 people. Therefore, we tested groups of size 1 up to size 8. To determine if the organization or formation of groups would affect our strategies, we used four different formations for groups of size 2 or larger. Each was chosen due to its difference in density, spread, and relation to the group centroid. These formations were a horizontal line, a vertical line, a circle, and a densely packed spiral. For the lines, these were in relation to the X-axis and Y-axis of the grid space respectively starting in the center of the grid (cell (49,49)). The circle formation would be the polygon of the size of the group (triangle, square, pentagon, hexagon, heptagon, octagon). For group size of 2, the circle formation was two agents diagonal to each other. All circles were centered around the center of the grid. Finally, the densely packed spiral group is a group formation that wraps around itself starting at the center and spiralling clockwise. The circle and densely packed formations were used to see how the density of a group might affect our strategies.

With each group formation there was additionally the question of "How much space exists between each member of the group?". To address this, we modified each formation so that members would go from 1 to 8 cells away from each other. To minimize randomness, all members are at the same spacing. This was to test how the density of the groups might affect strategies (i.e. very spread out vs clumped together). It also offered instances in which it is very possible to navigate into the middle of a group without walking into any group members. In real-world groups, this would be an average; some simple examples with different spacings for group members showed similar effects as described below.

We also explicitly included a way to modify how close the approacher stopped based on the size of the group. As we only tested up to group size of 8, this was calculated as: $C = 9 - N$, where N is the size of the group. Of course, with larger groups this would need to be modified.

To run these permutations against the strategies, we built a script that would iterate over the various factors described above. For each permutation we ran 100 iterations where each iteration had a new random starting point S. Each of these iterations calculated the endpoint E for each strategy.



(a) Distance to Centroid (b) Distance to Closest Person

Fig. 2: Top rows contain subgraphs whose strategies could use the middle of the group and don't take group size into account. Bottom rows can't use middle of group and do take group size into account. Each column is a different strategy tested against groups of size 1-8 (left to right). Closeness of the groups for these runs was 2.

6 Experiment Analysis

With these strategies, the question then became,"Do any of these strategies exhibit the behavior patterns noticed in SDD?". Notably, does the distance to center of the group increase as the group size increases? As well as, does the distance to closest person in the group decrease as the group size increases? In both instances from the SDD we found each pattern to be monotonic, therefore any strategy that exhibited non-monotonic behavior would be ruled out.

To examine the relationship between the different model strategies and the empirical data, we first examined the distance to the closest person. Figure 2b shows the different strategies as columns and two of the important model features as the row: the top row uses the middle of the group as valid locations and does not take into account the group size, while the bottom row does not use the middle of the group as valid locations and does take the size of the group into account when approaching. Recall that the empirical results showed a monotonically decreasing pattern. Examining the two different rows, we see a rather large difference: when the simulations are able to use the center of the group as valid locations and group size is not taken into account, the patterns are quite flat or in the opposite direction as the empirical data. When the simulations are able to use group size in their approach and do not use the center of the group, there are several strategies that decrease monotonically. Specifically, the strategies CPCS, CPRS, CS, RPCS, RPRS, and RS all show patterns across group sizes that are consistent with the empirical data.

Figure 2a uses a similar setup, but shows the results of how close the approacher comes to the centroid. Recall that the empirical data showed a monotonically increasing pattern across group size. Interestingly, when the simulations could use the middle of the group and did not take the size of the group into account, almost all of the strategies were consistent with the empirical data (though those have been ruled out by the distance to the closest person discussion above). The bottom row shows that the strategies CentDist, Proximity, Center, CPRS, and Cent&Prox all show an increasing pattern across group size (though not all of them are perfectly increasing). Comparing this list with the previous list, we see that CPRS is the only strategy that seems to be consistent with both measures.

7 Robot implementation

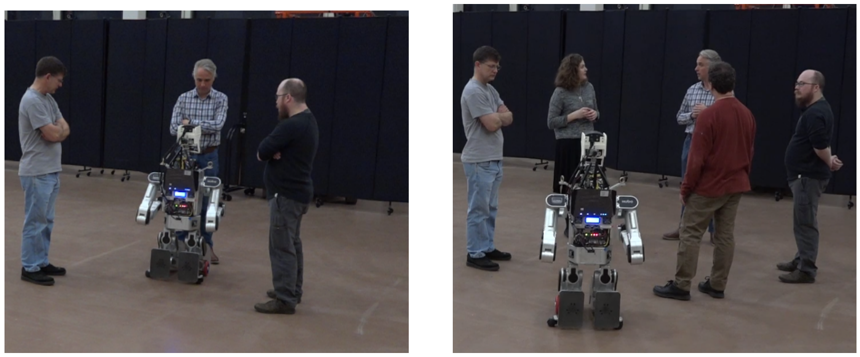
In order to demonstrate how our algorithm works on an embodied platform, we integrated several different models with our HUBO robot, Stryx. In these tests, we used Stryx's rolling mode. In rolling mode, Stryx can navigate around obstacles (including people) and travel over uneven or difficult terrain. To support sensing and interaction, Stryx has a sensor head that allows different sensors to be easily added to the platform in customized configurations [4]. Note that much of the work on approaching a group uses overhead cameras. [10] used an aerial drone and [13] used and discusses many of the datasets available for groups using aerial or off-the-ground cameras. Our work attempts to approach a group dynamically using only the cameras on the mobile robot platform. The camera in the SCIPRR head is a Carnegie Robotics Atlas head that includes a pair of stereo RGB cameras that allows detection of individuals from the robot's perspective.

The software components interact largely via the Robot Operating System (ROS)[5]. To detect and provide the location for people, the robot uses the deep convolutional neural network YOLO [9]. YOLO predicts both the location and the classification of known objects in the image. For our purposes, we are only using the detected people in the environment. Once a person has been detected, Stryx needs to locate the person on the grid. For this, the stereo pair is consulted to provide the range and bearing to the person. Which is converted to a Cartesian coordinate and placed into the appropriate grid cell and published via ROS. Each grid cell was defined as a $.3x.3m^2$ square.

Unfortunately, no ROS port is available for Prolog, so an intermediary was used. The PardonMe-soner writes the goal in terms of X and Y distance and it gets published via ROS. A third-party ROS navigation stack[6] receives the goal and outputs speed and direction safe velocity commands to drive the robot.

As discussed above, Stryx's Atlas head provides stereo images that perform person detection. In addition, the sensor also provides high-resolution laser range finder data which is consumed by ROS navigation packages. Hubo wheel encoders are used to estimate robot position over time. Unfortunately, due to slippage, the estimate is very noisy. The *amcl* package compares the most recent range readings to a known, static occupancy grid of the environment to correct this. After Stryx has reached its target location, it needs to then face the group [11,

[5] `www.ros.org`
[6] `http://wiki.ros.org/navigation`



(a) Group of 3, Center (b) Group of 5, CPRS.

Fig. 3: A photo of Stryx after approaching a group using different strategies.

14]. In order to accomplish this, we sent a rotation command to Stryx to turn toward the center of the group.

8 Preliminary Demonstration Results

Stryx was tasked with approaching a group of people. Because we are focused on approaching a group, we assumed that all people that the robot could see were part of the group. Future work will integrate group detection so that Stryx will be able to choose a group to approach.

A single group was told to informally talk among themselves while Stryx approached them. There were two different group sizes used: 3 and 5. Stryx was positioned so that it could see the group. We enabled the Center and CPRS strategies separately for different runs of the approaching robot. These strategies were completely autonomous after being started (details above in Section 7). When Stryx approached a group of 3 using the Center strategy (Figure 3a), people in the group and nearby observers found the behavior *odd and slightly creepy*. Even though the robot was not moving quickly, people thought it was unusual for the robot to enter the center of the group and approach so closely to individuals in the group. When Stryx approached a group of size 5 using the CPRS strategy (Figure 3b), people *felt that the robot was becoming part of the group*. Similarly, when the robot approached a group of 3 people using the CPRS strategy, observers and group members *felt comfortable with the robot's approach* and ending location.

While these results are informal, we highlight the fact that in the future, it will be possible to perform more formal evaluations of both the system and people's impressions of the approach. Additionally, through the multiple runs performed the PardonMe-soner was never a bottleneck nor a tax on the robot's system or computational capabilities. This leads us to believe more complex representations of the group and environment are entirely possible with our current models.

9 Discussion & Conclusion

As previous researchers have noted, approaching an individual or group is not trivial. There are a variety of aspects that enter into a successful approach, including distance, orientation, spacing, and speed. While previous research focused on agents and robots approaching an individual, the few systems that can approach a group focus only on approaching an individual that is part of a group rather than using features of the group to impact the approach. By analyzing a large scale dataset, we discovered a novel empirical finding: that individuals use features of the group they are approaching to impact how they approach. Specifically, we found that the size of the group has some impact on how people approach the group. If we want robots and agents to appear naturally as they approach, these findings should be taken into account.

We created strategies to approach a group of virtual people and then compared the results of these strategies and found several aspects especially relevant to matching the empirical pattern of behavior. Specifically, we found that when approaching a non-dense group, going into the middle of the group was poor approach behavior. We also found that using an *explicit notion of group size* and using that size to control how close to approach a group was more likely to enable an acceptable and natural approach behavior. Of the strategies we tested, we found one that was especially good across multiple group shapes and sizes - Closest Person, Random Square.

We demonstrated the viability and success of the model by matching it to the empirical data and running it on an embodied platform. There is still much room for improvement. First, the current model could not differentiate individuals that were not part of a group; future work needs to be able to identify and create groups so that a decision can be made to approach a specific group. Currently, in some situations (especially with medium/small size groups), the model puts the robot on the edge of the group. While the robot itself turns toward the center of the group, it may not be considered part of the group until it has both approached (current work) and joined (future work).

We feel it important to acknowledge that the dataset and in-person robot tests were used and generated prior to the COVID-19 pandemic. As such, certain social norms with respect to social distancing may no longer be accurate. While group distance and definitions may change in the future, the overall models and conclusion - that people use features of a group when approaching - will remain, such that the observed patterns will remain regardless.

While this work has highlighted that group features can impact and moderate how an agent should approach, we focused here on group size. There are, of course, other features to groups that likely impact how and agent should approach them. Shape, activity, and mobility all likely impact how an agent should approach. By using explicit declarations for these shapes, activity, and mobility we hope to continue using computationally straightforward representations of these features, so that reasoning never becomes a bottleneck for these robots.

Acknowledgments

This research was performed while BW held an NRC Research Associateship award at NRL. This research was funded by ONR and OSD to GT.

References

1. Althaus, P., Ishiguro, H., Kanda, T., Miyashita, T., Christensen, H.I.: Navigation for Human-Robot Interaction Tasks. In: IEEE Intl. Con. on Robotics and Automation, 2004. Proceedings. ICRA'04. 2004. vol. 2, pp. 1894–1900. IEEE (2004)
2. Ball, A., Silvera-Tawil, D., Rye, D., Velonaki, M.: Group Comfortability when a Group Approaches. In: Intl. Con. on Social Robotics. pp. 44–53. Springer (2014)
3. Hall, E.T.: The Hidden Dimension, vol. 609. Garden City, NY: Doubleday (1966)
4. Harrison, A.M., Xu, W.M., Trafton, J.G.: User-Centered Robot Head Design: A Sensing Computing Interaction Platform for Robotics Research (sciprr). In: 13th ACM/IEEE Intl. Con. on Human-Robot Interaction. pp. 215–223 (2018)
5. Karreman, D., Utama, L., Joosse, M., Lohse, M., van Dijk, B., Evers, V.: Robot Etiquette: How to Approach a Pair of People? In: 9th ACM/IEEE Intl. Con. on Human-Robot Interaction. pp. 196–197 (2014)
6. Kato, Y., Kanda, T., Ishiguro, H.: May I Help You?-Design of Human-Like Polite Approaching Behavior. In: 10th ACM/IEEE Intl. Con. on Human-Robot Interaction (HRI). pp. 35–42. IEEE (2015)
7. Kendon, A.: Conducting Interaction: Patterns of Behavior in Focused Encounters, vol. 7. CUP Archive (1990)
8. Mead, R., Matarić, M.J.: Autonomous Human–Robot Proxemics: Socially Aware Navigation Based on Interaction Potential. Autonomous Robots **41**(5), 1189–1201 (2017)
9. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You Only Look Once: Unified, Real-Time Object Detection. In: Proc. of the IEEE Con. on Computer Vision and Pattern Recognition. pp. 779–788 (2016)
10. Robicquet, A., Sadeghian, A., Alahi, A., Savarese, S.: Learning Social Etiquette: Human Trajectory Understanding in Crowded Scenes. In: European Con. on Computer Vision. pp. 549–565. Springer (2016)
11. Rousseau, V., Ferland, F., Létourneau, D., Michaud, F.: Sorry to Interrupt, But May I Have Your Attention?: Preliminary Design and Evaluation of Autonomous Engagement in HRI. Journal of Human-Robot Interaction **2**(3), 41–61 (2013)
12. Satake, S., Kanda, T., Glas, D.F., Imai, M., Ishiguro, H., Hagita, N.: How to Approach Humans? Strategies for Social Robots to Initiate Interaction. In: 4th ACM/IEEE Intl. Con. on Human Robot Interaction. pp. 109–116 (2009)
13. Setti, F., Russell, C., Bassetti, C., Cristani, M.: F-Formation Detection: Individuating Free-Standing Conversational Groups in Images. PloS One **10**(5), e0123783 (2015)
14. Shi, C., Shimada, M., Kanda, T., Ishiguro, H., Hagita, N.: Spatial Formation Model for Initiating Conversation. Proc. of Robotics: Science and Systems VII pp. 305–313 (2011)
15. Wielemaker, J., Schrijvers, T., Triska, M., Lager, T.: SWI-Prolog. Theory and Practice of Logic Programming **12**(1-2), 67–96 (2012)