Distance-Based Computational Models for Facilitating Robot Interaction with Children

David Feil-Seifer
Department of Computer Science, Yale University
and
Maja J Matarić
Departments of Computer Science and Pediatrics, University of Southern California

Sensing and interpreting the user's activities and social behavior, monitoring the dynamics of the social context, and selecting and producing appropriate robot action are the core challenges involved in using robots as social tools in interaction scenarios. In social human-robot interaction, speech and gesture are the commonly considered interaction modalities. In human-human interactions, interpersonal distance between people can contain significant social and communicative information. Thus, if human-robot interaction reflects this human-human interaction property, then human-robot distances also convey social information. If a robot is to be an effective social agent, its actions, including those relating to interpersonal distance, must be appropriate for the given social situation. This becomes a greater challenge in playful and unstructured interactions, such as those involving children.

This paper demonstrates the use of a distance-based model for the recognition and expression of spatial social behavior. The model was designed to classify averse social behavior of a child engaged in playful interaction with a robot and uses distance-based features to autonomously identify interaction/play, avoidance, wall-hugging, and parent-proximity behavior with 94% accuracy. The same methodology was used to model the spatial aspects of a person following a robot and use the model as part of a modified navigation planner to enable the robot to exhibit socially-aware goal-oriented navigation behavior. The model-based planner resulted in robot navigation behavior that was more effective at allowing a partner to follow the robot. This effect was demonstrated using quantitative measures of navigation performance and observer rating of the robot's behavior. These two uses of spatial models were implemented on complete robot systems and validated in evaluation studies with children with autism spectrum disorders and with neurotypical adults.

Keywords: Socially Assistive Robotics, Autism Spectrum Disorders, Machine Learning

1. Introduction

Interpersonal distance between people can contain significant social and communicative information. If human-robot interaction reflects this human-human interaction property, then human-robot distances also convey social information. If a robot is to be an effective social agent, its actions, including those relating to interpersonal distance, must be appropriate for the given social situation.

This becomes a greater challenge in playful and unstructured interactions, such as those involving children.

This work explores how an autonomous robot can use its movement and interpersonal distance to perform socially appropriate behaviors with children in different social scenarios and navigation tasks. We focus primarily on robot interaction with children with autism spectrum disorders (ASD). We posit that key elements of child-robot social interaction can be modeled using spatial features, such as interpersonal distance, and demonstrate the use of spatial models over distance-based features for enabling social behavior classification and socially-aware navigation.

Several models exist for describing how interpersonal distance is social interaction or an expression of social state. Dean et al. describe how interaction distance changes depending on military rank (Dean, Willis, & Hewitt, 1975). Chen et al. similarly describe how factors such as differences in height and orientation between two or more people can affect the perceived power in a relationship (Chen, Lee-Chai, & Bargh, 2001). One would expect human-robot interaction to reflect, in part, a power relationship. Additionally, Hall suggests that social distance is correlated with interpersonal distance, and that factors such as kinesthetics, eye contact, audio volume, perceived body heat, and even olfactory information can reflect on physical distance (Hall, 1966). Therefore, we expect that the physical distance between a child and a robot is, at least in part, a reflection of social state.

If an autonomous robot is to effectively interact with a child, it must be aware of the child's behavior and produce timely, contingent, and socially appropriate behavior in response. Observations of children with ASD have demonstrated that spatial behavior, the manipulation of interpersonal distances between social agents and aspects of the environment, can indicate social and anti-social (averse) behavior (Feil-Seifer & Matarić, 2011a). For example, a child staying in front of the robot can indicate a social response, while a child moving away from the robot and toward the wall can indicate an averse response. Furthermore, a robot's navigation toward a goal, when executing the shortest safe path, may lead to the misinterpretation of the robot's intentions (Feil-Seifer & Matarić, 2011b).

The work presented in this paper demonstrates how models of social interaction based on spatial features can be used to address play behavior classification (into social and averse) and socially-aware navigation. We present a data-driven method for utilizing spatial information as part of a play scenario to identify averse reactions of a child to the robot, evaluated using experimental data involving a robot, a child with ASD, and a parent. We also present how the same method can be used to train a spatial model of following behavior. We then use that model to inform a navigation planner for a follow-the-robot task, inspired by experimental results with children with ASD and validated using neurotypical adults.

The contributions of this paper are aimed at continued research and eventual development of therapeutic robots for ASD and other contexts; no therapeutic claims are made at this very early stage of the exploration of the methods and their effects. This paper expands on our earlier conference publications on recognition and exhibition of spatial interactive behavior and provides additional supporting experiments and analyses (Feil-Seifer & Matarić, 2011a, 2011b). This article reviews the work of those two papers while adding supplementary studies and evaluation.

The rest of this paper is organized as follows. Section 2 provides the details of the initial data collection experiment that provided the motivation and data for the work in this paper. Section 3 describes the model formulation approach for training a model of spatial social behavior. Next, Section 4 describes the method for classifying positive and averse social behavior using that model, which employed distance-based features to autonomously identify interaction/play, avoidance, wall-hugging, and parent-proximity behavior with 94% accuracy. Section 5 presents the modified navigation planner that produces autonomous socially-aware navigation behavior conforming to a spatial model incorporating information about a partner's spatial relationship when selecting the best
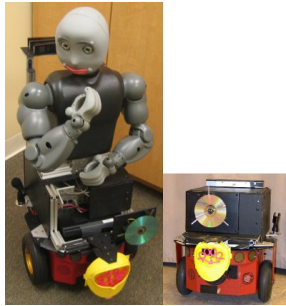
Figure 1. The two conditions used in the feasibility study. Left: the mobile humanoid robot. Right: a mobile non-biomimetic robot used in place of the humanoid robot if the child was intimidated by the larger humanoid robot.

trajectory. The model-based planner results in robot navigation behavior that shows improved effectiveness at allowing a partner to follow the robot. This effect is demonstrated using quantitative measures of navigation performance and observer rating of the robot's behavior. Finally, Section 6 summarizes the contributions of the described work.

2. Feasibility Study and Data Collection

The inspiration for the work described in this paper comes from data and insights collected from pilot human-robot interaction (HRI) experiments involving children with ASD and mobile robots. In related work, Robins et al. and Kozima et al. used table-top robots to interact socially with children with ASD (Robins, Dautenhahn, & Dickerson, 2009; Kozima, Nakagawa, & Yasuda, 2005). Instead of interacting through robot base movement, these robots interacted through drumming and dance, respectively. The systems were either reactive or interacted explicitly with turn-taking behavior. Salter, Michaud, Létourneau, Lee, and Werry used proprioceptive sensing, such as accelerometer information and orientation information, to determine how a ball-shaped robot was being played with and to select actions (Salter et al., 2007). For example, if the robot was being pushed, the robot moved quickly, while if the robot was being spun in a circle, it could play a noise. Such information is useful for reactive social interaction.

We conducted a feasibility HRI study with participant families with eight children diagnosed with ASD. It consisted of a free-play scenario involving a robot, a child, and a parent. A parent was present 1) to minimize any child distress in the unfamiliar context, and 2) to begin to explore using the robot as a catalyst for social interaction with other people, as our goal was not to isolate the child in human-machine interactions only. The goal of the study was to observe the children's reaction to the robot in a free-play setting and to evaluate the feasibility of using such autonomous robots for interaction with children with ASD. The aims of using the robot were to:

1. engage the child in social play interaction with the robot; and
2. encourage social interaction between the child and the parent.

The first aim required the child to find playing with the robot enjoyable enough to sustain an interaction throughout the length of the experiment. The second aim required the robot to direct the child to interact with the parent. The overall study was approved by the University of Southern California Institutional Review Board (IRB).

2.1 Recruitment

The participating children were recruited from Autism Speaks Autism Genetic Resource Exchange (AGRE). The AGRE program provides bio-materials and phenotype and genotype information of

families with two or more children (multiplex) with ASD to the scientific community (Geschwind et al., 2001). All participants lived in the greater Los Angeles area. Fliers were sent to eligible families. Interested families responded by phone and email. Of the 65 contacted families, 8 responded (12.3%), and 8 children from 5 families participated in the study. Because the AGRE database consists of multiplex families, three of the participating families had siblings recruited for the study.

Diagnosis of autism by AGRE researchers using a combination of the Autism Diagnostic Observation Schedule (ADOS) and Autism Diagnostic Inventory (ADI) established a participant's eligibility for inclusion in the study (Lord et al., 2000; Lord, Rutter & Le Couteur, 1994). Additional inclusion criteria required that the children be between the ages of five and ten, and have at least some verbal ability, defined as a score above 2.0 years of age on the communication sub-scale of the Vineland Adaptive Behavior Scale (Sparrow & Cicchetti, 1989) or an evaluation using either modules two or three of the ADOS.

2.2   Robot System Design

We equipped the experimental space with an overhead camera to detect the positions of the child, the parent, the robot, and other obstacles in the room. Given *a priori* information about the layout of the room, this sensor alone was enough to enable the robot to maneuver autonomously in the experimental space and to use the detected positions over time to sense relevant movements of the participants in order to trigger appropriate behavior in response.

Each video frame, collected at 15 frames per second, served as a single sample for detecting the locations of the interaction participants and deciding on an appropriate behavior. Each frame required less than 40% of the frame interval to compute, leaving additional time for behavior classification, supporting the real-time needs of the system. After each frame, we were able to determine the position and orientation of the robot, the position of the child, the position of the parent, and any other objects in the room. These positions comprised the data used to model the behaviors taking place during the experiment.

2.3   Experimental Protocol

The experiment participants consisted of the robot, the child, and a parent who was instructed to participate in the interaction if the child requested it. The children were given little specific instruction on what to do in the interaction. The parents were told that there was a chair available to sit in, but if the children wanted to involve them in their interaction with the robot, that they were free to move around the room and participate. Participants took part in two 5-minute experimental sessions featuring two conditions: robot and toy (see Figure 1). The humanoid robot (when present) was capable of making simple gestures and navigating autonomously without hitting obstacles. Some children had an immediate averse reaction to the humanoid robot. When that occurred, we replaced the humanoid robot with a similar robot that did not include the humanoid torso. We developed the following simple rules for the robot's behavior:

- The robot socially oriented to the child by default;
- If the child was more than 1m from the robot, the robot waved at the child, then gestured to the child to "come here;"
- If the child was more than 1m from the robot for longer than 10 seconds, the robot approached the child;
- If the child approached the robot, the robot made an approving vocalization;
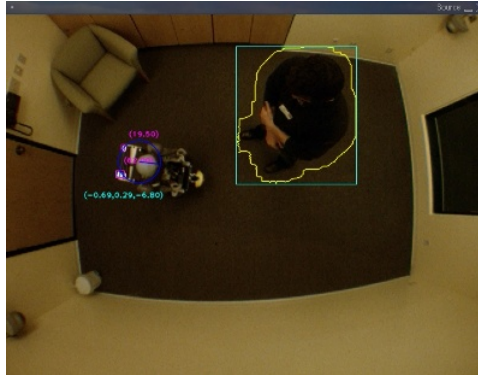- If the child moved away from the robot, the robot made a disappointed vocalization;

58

Figure 2. A view from the overhead camera with annotation about the pose of the robot and of the participant.

- If the child moved behind the robot, the robot did nothing. (This was in order to provide a safe zone where the child could hide from the robot and be ignored by it.);
- If the child pressed the button on the robot, the robot blew bubbles; and
- If the child made a vocalization to the robot, the robot blew bubbles.

The above rules were designed to motivate specific social behaviors, mostly spatial in nature. All behavior detection was autonomous, with the exception of vocalization detection, which involved an operator. Available autonomous methods for detecting vocalization cannot provide information regarding whether or not a vocalization was self-stimulation behavior, echolalic behavior, or speech directed at the parent or robot. This means that non-social vocalization could be "rewarded" with bubbles. For the purposes of autonomous behavior generation, we presume that all observed vocalizations were social and directed at the robot.

2.4 Qualitative Evaluation

A qualitative analysis of the overhead camera data from the experiment showed a clear distinction between children who had positive reactions to the robot compared to those who had averse reactions. Averse reactions ranged from trying to avoid being in front of the robot, to moving away from the robot toward a wall or the parent, to screaming at the robot and running away. The distinction between positive and averse reactions to the robot were observable from the child's spatial behavior, namely where the child chose to stand relative to the robot, parent, and walls of the room. In this section, we describe our process of experimental data analysis and coding for distance-based states.

An anthropologist specializing in ASD also coded the session data based on the children's reactions to the robot. Some children ($n = 4$, 6 sessions) had averse reactions to the robot. Many of these averse reactions could be characterized by spatial behavior, such as avoiding the robot, staying near the walls of the experimental space, or staying close to the parent without interacting with the robot. In contrast, the remainder of the children ($n = 4$, 7 sessions) had uniformly positive reactions to the robot, making attempts to engage the robot socially. Some children beckoned the robot to follow them around the room, played with the robot when it blew bubbles, and spoke to the robot in an attempt to encourage it to socially interact with them.

All of the children who had averse reactions to the robot requested to end the study early, whether due to the robot's motor sounds, suddenness of movements, or other reasons. Therefore, we classi-

fied the reaction to the robot as positive if the child was able to complete both robot sessions, and as averse if the child was unable to complete one or both of the robot sessions and ended the study early.

We also noted that, when the robot moved around the room, some children acted as if the robot were ignoring them in favor of whatever it was moving toward. In those cases, as the robot was moving away, the children temporarily stopped attempting to interact with it. This effect inspired the socially-aware navigation approach discussed later in the paper.

2.5 Discussion

When asked about what caused their averse reactions, the children stated that many features of the robot could be improved. The primary complaint was a high-pitched noise made by the mobile base when trying to hold its position. We eliminated this noise in later instances of this experiment by disabling the motors at speeds close to zero. Some children complained that the robot's arms were too loud when they moved and that they moved too suddenly. This was addressed in two ways. First, we redesigned the movements of the robot to be more smooth and continuous. Second, we designed a mobile non-anthropomorphic robot that otherwise has the same capabilities of the torso robot for children intimidated by the torso robot.

Children, especially children with ASD, will not respond well 100% of the time, no matter the robot's appearance and behavior. Therefore, it is very important that an autonomous system be able to identify and respond to a negative reaction from a child in order to facilitate the human-machine interaction. In addition, if a robot's actions create an incorrect social perception due to their lack of conformity with social precepts, then it is important to ensure that the robot's actions conform to an established social model of behavior.

In the next section we describe how we constructed and validated models of social behavior using interpersonal distance-based features.

3.   Model Formulation and Feature Selection

The insights from the above-described feasibility study indicated that spatial data could potentially be used to classify child behavior into positive and averse in this context. The next step was to design a model based on spatial features that could be used for automatic behavior classification.

Spatial-based modeling of behavior has been demonstrated using tracked positions over time. Most such approaches assume that events are related to absolute locations in the environment and use some function of the image coordinates as features in the model (Xiang & Gong, 2006; Zhong, Shi, & Visontai, 2004). However, when the target of the modeling is interpersonal interaction, proxemic information (the interpersonal distance and/or orientation between two mobile agents) is used as a supplemental or replacement feature (Kelley et al., 2008). Related approaches can be used to model regular behavior and identify sequences where the activity observed is somehow irregular (Boiman & Irani, 2007).

3.1 Recognizing Behavior From Spatial Data

Using only the spatial data we collected from the overhead camera, the children's behavior in the sessions could be classified into positive vs. averse, but also into several descriptive categories, as follows:

- **Avoiding the robot:** child is moving so as to consistently increase her/his distance to the robot.
- **Interacting with the robot or playing with bubbles:** child is moving or still and is remaining proximal to the robot. Generally, the child was in front, or to the side, of the robot.
- **Near parent:** child is touching the parent or next to the parent while not moving.
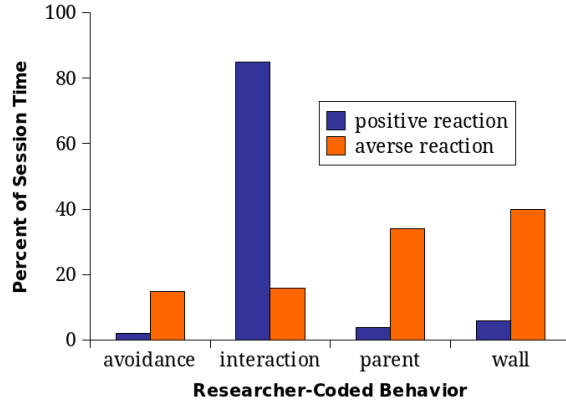
Figure 3. Coders' categorization of children's responses to the robot as positive (blue) or averse (orange). Reactions are further categorized into avoiding the robot (avoidance), interacting with the robot or bubbles (interaction), staying near or interacting with the parent (parent) or nearing the wall (wall).

- **Against wall:** child is touching the wall while not moving. Most of the time, when the child was near the wall, the child was facing the robot.
- **Null/none of the above:** a catchall state used if none of the above conditions are met. Less than 5% of session time fit into this category.

Bubble-blowing is a feature particular to our robot design, and since playing with the bubbles was a sign of interacting with the robot, we used this as a definition of interaction. Other robot capabilities could be substituted for the bubble play in different contexts in order to define interaction. Furthermore, the above definition of "Near Parent" is not associated with positive social interaction with the parent. Rather, in our context, a child was typically in this state due to being upset and going to the parent for comfort.

The video data were coded for the above behaviors by three independent coders. The coders had 92% agreement (Fleiss's κ: 0.88). The coding demonstrated the spatial relationships between the child and the robot, walls, and parent. When children were observed to be in visible distress, they stood against the wall, or near a parent. The coding results suggest that some distance states are correlated with the children's averse response to the robot. Quantitatively, the children who had a positive reaction to the robot spent more than 80% (std dev 8.9%) of the session time interacting with the robot or playing with bubbles, while the children who had an averse reaction to the robot spent less than 20% (std dev 13.4%) of the session time in those states (see Figure 3). This quantitatively validated that social (positive or averse) behavior was observable from spatial information.

3.2 Feature Selection

The next step in modeling was to select the set of features for model construction. Feature selection has typically used image coordinates of detected and tracked agents as part of the feature vector. However, such features assume that the relevant information can be expressed in the form of Cartesian positions, an assumption that holds for fixed-camera activity-modeling systems (Xiang & Gong, 2006), but not for social interaction activity modeling, where distances between the social actors in the scene are of interest.

Based on the insights from the feasibility study, the important relative distances were the child-robot distance, child-wall distance, and child-parent distance. Additionally, the orientation of the
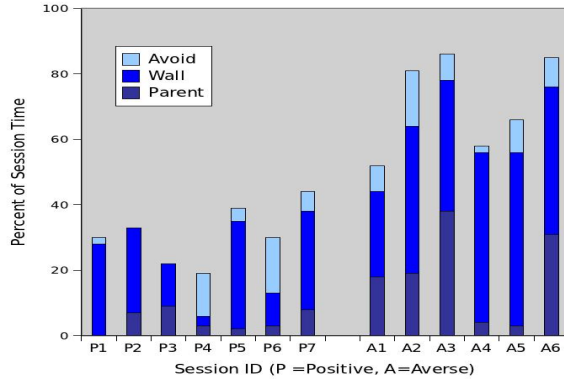
Figure 4. Heuristic classification of distance states for children that had a positive reaction to the robot (Group P, left), and those that had an averse reaction to the robot (Group A, right); Group P spent less time near the walls, the parent, or avoiding the robot.

robot to the child could be used to discern if they were facing each other (as in desirable interaction) or if the robot was moving away from the child, or the child was in front of the robot. Because behavior naturally involves movement, detecting movement and distinguishing it from static relative poses is important in our behavior classification; we thus used relative movement between the features (robot, parent, wall) as a feature itself.

To examine the feasibility of using spatial features to model a child's social behavior in our context of interest, we conducted a simple test that used distance heuristics to determine what state the child was in. For each frame: if the child was within 1.25 meters of the parent, the system recorded the child as being near the parent; if the child was within 0.3 meters of the wall, the system recorded an observation of being near the wall; finally if the child was behind the robot at any distance (greater than $135°$ or less than $-135°$ from the front of the robot) the system recorded the observation of avoiding the robot.

The above thresholds were based on visual inspection of the data and were not meant to be exact. They were not data-driven, but were based on a sampling of the data. The observation of the *avoid* state was used in this heuristic evaluation because the programmed robot behavior in the recorded experiment included the rule that caused the robot to stop all expressive social behavior if the child was detected behind the robot.

The recorded sessions were divided into two groups. Group P ($n = 4$; 7 sessions) consisted of the sessions with children who completed the experiment, characterized as liking the robot and spending a significant amount of time interacting and playing with it. Group A ($n = 4$; 6 sessions) consisted of the sessions with children who did not complete the experiment but terminated early, characterized by spending a significant amount of time avoiding the robot and/or seeking comfort from the parent.

Using the results of this automated annotation, we compared Groups P and A according to the percentage of session time that the child was either near the wall, near the parent, or avoiding the robot. We found that, in all of Group P's sessions, the wall, avoid, and parent states were observed for less than 40% of the time (mean 30%, std dev 0.07), while for Group A's sessions, these states were observed more than 50% of the time (mean 71.9%, std dev 0.15), a clear discrimination between the two groups (see Figure 4). This demonstrated that a classifier could readily be constructed based on the rule that greater than 50% time spent in the averse behaviors indicated a session wherein the

Table 1: The confusion matrix for automated classification with $m = 10$ clusters.

	Actual Behavior			
Coded Behavior (%)	Avoidance	Interaction	Parent	Wall
Avoidance	**62.30**	0.45	0	0.56
Interaction	37.47	**98.98**	4.96	10.92
Parent	0	0.45	**94.93**	2.24
Wall	0.23	0.13	0.2000	**86.27**

child was not trying to interact socially with the robot, while less than 50% indicated that the child was attempting to interact socially with the robot.

We compared the classified distance states to the human annotation for each individual video frame and achieved an accuracy of 84.98%; these heuristics only correctly identified the wall state 73% of the time and the parent state 85.03% of the time. Because these states are the most critical for identifying if the child is in distress, the accuracy of these states needed to improve to meet the classification goals.

In summary, the heuristic evaluation showed that spatial data could be used to train a model for use in discriminating positive vs. averse behavior. In the next section, we describe a machine learning approach to training just such a model.

4. Automated Classification of Positive vs. Averse Child Behavior

Our analysis demonstrated that averse child responses were observable from spatial behavior and that they could be identified by a human annotator. The heuristic-based classifier described in the previous section was sufficient for discriminating between the positive and averse child responses in our data set. However, the heuristic approach was not feasible for complex multi-dimensional feature sets. Thus, in order to develop a more general approach that would detect behavior as it is occurring without requiring human-generated heuristics, we developed a model-based automated classifier that uses interpersonal distance-based features to classify positive vs. averse behavior.

4.1 Model Formulation and Data Classification

We used an 8-dimensional feature vector: $v = \langle d_c^r, d_c^p, d_c^w, \psi_c^r, v_c, v_c^r, v_c^w, v\psi_c^r \rangle$ where d_c^r is the distance between the child and the robot, d_c^p is the distance between the child and the parent, d_c^w is the distance between the child and the nearest wall, ψ_c^r is the orientation of the child to the robot, v_c is the absolute velocity of the child, v_c^r is the velocity of the child relative to the robot, v_c^w is the velocity of the child relative to the wall, and $v\psi_c^r$ is the change in orientation of the child to the robot (all velocities are measured over $1s$).

The data were grouped into 3216 sequential tiles of 30 observations (2s each). A training set of 20% of the tiles was created; the other 80% was set aside for testing. The model was created by fitting a Gaussian Mixture Model (GMM) to each major class of data (avoidance, interaction, parent, and wall) using expectation-maximization (EM) to the training tiles. The initial seeding was random and was repeated five times to avoid local minima. A model order of 10 was selected, based on inspection. A naive Bayes classifier was used to determine the most likely classification for a given tile.
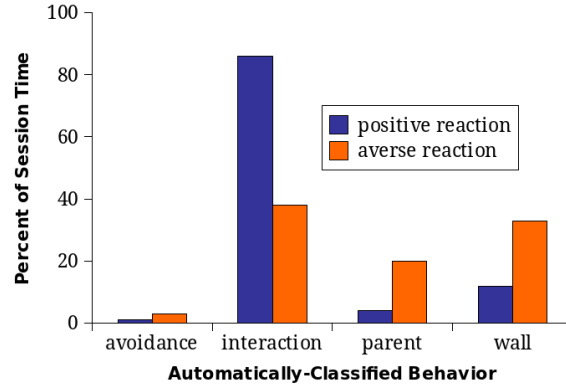
Figure 5.   Classified percentage of session time spent in each interaction state (Group P is blue, Group A is red); Group P spent more session time interacting with the robot than Group A.

4.2   Results

We classified the test data set using a 5-fold validation of the data. For this exercise, 20% of the data were used for training, 80% for testing. For simplicity, we only discuss the results for the avoidance, interaction, parent, and wall behaviors, because they account for more than 90% of session time (see Table 1).

The GMM/naive Bayes classification technique correctly identified 3023 out of a total of 3216 tiles, with the remaining 278 being incorrectly identified, for an overall correct recognition rate of 94.02%. The confusion matrices do not have an equal allocation of each class, as a much greater number of samples were from the interaction group than from the other groups, accounting for the disparity between the confusion matrices and the overall correct recognition result. These results are encouraging, especially the correct recognition of parent and wall behavior (95% and 86%, respectively), which were markedly better than the heuristic filter (85% and 73%, respectively). Because these behaviors were the main indicators of a child's averse reaction, their correct classification significantly improves the ability to automatically discriminate.

Figure 5 shows the result of the machine classification for Group P vs. Group A, demonstrating a similar trend observed from the human-rated annotation, where the children from Group P (who had a positive reaction to the robot) are observed to have spent a greater amount of time in the interaction state (78%) as compared to the parent (3%), avoidance (0%), or wall (11%) states, while the children from Group A (who had an averse reaction to the robot) spent far less time in the interaction state (36%) and more time in the other states (20%, 2.6%, and 38%). This shows that the annotation technique can be used to differentiate between children with distinct reactions to the robot.

4.3   Discussion

We developed an unsupervised classifier based on distance features, and classified behaviors of children with ASD in an experimental free-play setting. A secondary goal of the work was to apply this classification in such a way that it could be used to transcribe child behavior in real-time. Such a system could be used as an input for autonomous robot action selection.

We have shown that the GMM-based method for state clustering can efficiently and effectively cluster the 8-dimensional feature space. These states are easily labeled by using annotated training data and could be used for partial behavior transcription. Potential concerns include over-

generalization that can occur with human labeling, and over-specialization given the heterogeneity of the participant population.

Two approaches are most closely related to this work (Kelley et al., 2008; Xiang & Gong, 2006). Xiang and Gong posed a data-driven approach to action selection that used a GMM to identify clusters in a 7-dimensional feature space (Xiang & Gong, 2006). The authors used Bayes Information Criteria (BIC) (Schwarz, 1978) to select the number of recognized actions in an unsupervised manner. We attempted a similar approach for our behavior classification work. However, the features used in their work were all directly tied to either the object being tracked (size of blob, size of bounding box) or the static properties of the environment (position in the image, change in position in image coordinates). The BIC-based method for unsupervised model order selection did not provide a model order which was effective for modeling the data, because the data were not Gaussian. Kelley et al. described an approach that relied on interpersonal distances and velocities, but where the output was a detected navigation goal, not behavior classification as demonstrated in our work (Kelley et al., 2008).

Our approach could potentially scale to other environments, given certain conditions. For example, the room size used in this experiment is probably close to the minimum size usable for this approach simply because the child had very little room to maneuver. However, because the dominant feature for the parent and wall states was relative distance to the wall, it would not matter how large the room was provided the child-wall and child-parent distance remained relatively constant when the child was having an averse reaction to the robot. The interaction and avoidance behaviors, on the other hand, were mostly concerned with physical proximity to the robot and the child's movement with respect to the robot, so they would not be affected by the size of the room.

There do exist situations where the model would not work as trained. For example, in a robot-child interaction task that required the child to be behind the robot (asking the child to follow the robot, for example), the interaction behavior would need to be redefined. However, the model training methodology presented in this section would still apply, and a different interaction model would be learned. Because model training was achieved using roughly five minutes of training data, this would still be a very practical approach.

An important milestone regarding the accuracy of this system relates to inter-coder reliability. We were able to obtain about 92% agreement among the three coders. The automatic system's accuracy was marginally higher (94.02%). This means that the automated system was able to achieve roughly the same inter-rater agreement.

We compared multiple classification methods, including Gaussian Mixture Models (GMMs), Linear Discriminant Analysis (LDA), and Support Vector Machines (SVM), using a 5-fold approach, with 80% for training, 20% used for testing. Each of these methods performed better than the heuristic classifier, but the GMM approach performed the best overall, with 93.21% accuracy. The higher classification accuracy from the results presented earlier is due to the fact that a larger training set was used for this method comparison. As seen in Figure 6, the data, while not inherently Gaussian, are also not clustered in unimodal distributions. The data are also not arranged in such a way as to facilitate easy binary division between classes. This might explain the somewhat poorer performance of LDA and SVM methods for classifying the data, compared to GMM, which is multimodal.

The next section describes how the same methodology for developing a model of spatial behavior can be applied to a navigation task.

5. Socially-Aware Navigation

This section describes how the above-described approach to distance-based features for data-driven model development can be applied to a different challenge in child-robot interaction. We observed that children were misinterpreting the robot's intentions when it was navigating toward a goal with-
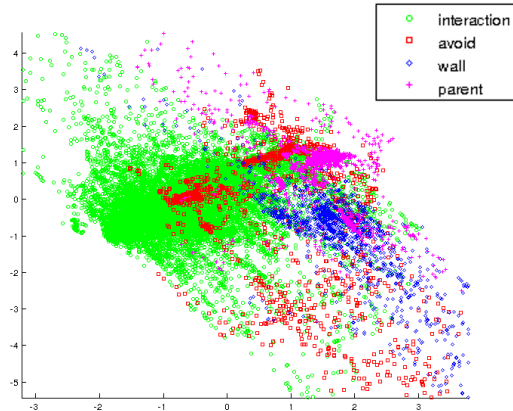
Figure 6. Scatter plot of the first two principal components of the robot-child interaction data. While not Gaussian, the data are not scattered in a fashion conducive to simple binary classification.

Table 2: A comparison of different classification algorithms used to analyze the data, Gaussian Mixture Models (GMM), Support Vector Machines (SVM), and Linear Discriminant Analysis (LDA). For each category and for overall results, GMMs produced the best results.

| Algorithm | True Recognition of Behavior (%) | | | | |
	Avoid	Interaction	Parent	Wall	Overall
GMM	62.30	98.98	86.27	94.93	**94.02**
SVM	33.96	95.34	85.50	88.48	**89.23**
LDA	29.34	96.17	71.92	85.88	**85.29**

out waiting for the child to catch up. Such behavior was interpreted as the robot disengaging from the social interaction. Such misinterpretation can "break the eggshell" (Mead, 2010) of a very fragile interaction (Mead, Atrash, & Matarić, 2011). This motivated us to develop a socially-aware navigation system, one that is capable of leading a child to a new goal location (e.g., the parent) while sustaining the ongoing interaction.

Traditional autonomous robot navigation is concerned with moving to a goal as quickly as possible while avoiding obstacles (Fox, Burgard, & Thrun, 1997), but such goal-oriented navigation may not be socially appropriate. Actions taken by the robot in a social setting inadvertently carry social meaning, so even navigation has social implications, intended or otherwise. In this section, we describe a modified trajectory planner we developed that executes navigation behavior that conforms to a model of distance-based social behavior.

5.1 Model Formulation

We used a 4-dimensional feature vector: $w = \langle t, \Delta_g^r, \Delta_g^p, \Delta_p^r \rangle$ where t is the normalized time spent on task ($t = 0$ is when the goal was set, $t = 1$ when the goal has been reached), Δ_g^r and Δ_g^p are the normalized distance between the robot or partner and goal (1 at the starting distance between robot or partner and goal, 0 at the goal itself). Δ_p^r is the distance between the robot and partner, in meters.
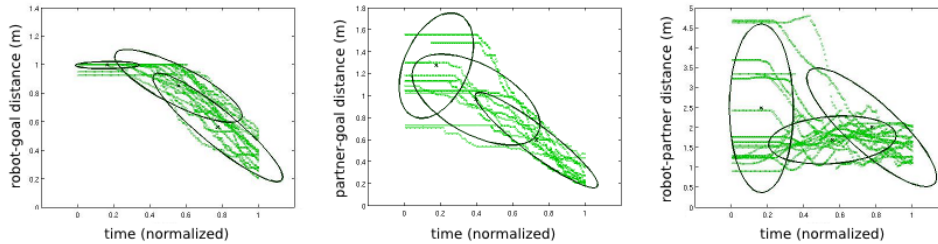
Figure 7. The training data used for the model and components of the trained GMM. Left: Δ_g^r by t. Center: Δ_g^p by t. Right: Δ_p^r by t.

These features can be normalized by completion progress, where $t = 0$ represents the start (when the robot sets a position goal) and $t = 1$ represents the end (the time when the robot reaches the goal).

The collected data were not strictly Gaussian, nor were they strictly a mixture of Gaussian distributions. However, because there were multiple distinct components of the data, a multimodal Gaussian model could be used to represent the data. In particular, three distinct ordered components (over t) could be observed:

1. The robot sets the goal, and turns in place to face the goal.
2. The robot begins moving toward the goal, but the child is not yet following.
3. The robot's goal is clearly identifiable to the child, and the child follows the robot to the goal.

These observable components result in the following progression: an initial lack of movement toward the goal starting at $t = 0$, followed by movement with inconsistent distribution of partner-robot distance, then followed by close partner robot distance as t approaches 1. We modeled these data by fitting a 3-component GMM using expectation-maximization (EM) to the training data (see Figure 7). We used OpenCV for the GMM and EM implementations. In the next section, we describe the method by which trajectories are weighted by the planner in order to conform to the model.

5.2 Modified Trajectory Planner

We implemented our system using the *nav_core* stack of the Robot Operating System (ROS) (Quigley et al., 2009), which is designed to be general-purpose, to work with a broad range of sensor inputs, and with robots using holonomic or differential drives, in order to plan and execute trajectories to a target, (x,y,θ). The stack had already been validated on a variety of existing robot systems, making it an ideal starting place for adding people-aware properties to the navigation system.

For local planning the robot uses the *base_local_planner* package in *nav_core*. This system can employ either the Trajectory Rollout System (Gerkey & Konolige, 2008) or the Dynamic Windowing Approach (Fox et al., 1997) to trajectory planning.

The ROS trajectory planner weights candidate trajectories by progress toward the goal and conformity to a planned path in order to pick the best control velocities. In order to additionally weight candidate trajectories by conformity to the given model of social behavior, we employed Gaussian Discriminant Analysis (Hastie & Tibshirani, 1996). GDA gives a probability that a candidate point belongs to a GMM model. In this section, we describe how GDA can be used as part of a modified trajectory planner to account for the social model.
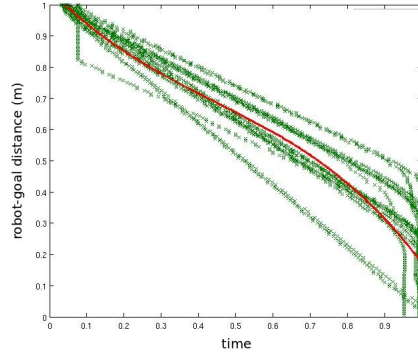
Figure 8. Δ_g^r by t, with a cubic regression shown in red.

Each candidate trajectory results in an endpoint in Cartesian space. With this potential position for the robot, along with the known positions of the partner and goal, the values of Δ_g^r, Δ_g^p, and Δ_p^r are known, only t is not known. Using a cubic regression, t can be determined as a function of Δ_g^r. Figure 8 shows that a cubic regression can accurately model the relationship between the robot's distance to the goal and the time spent moving to the goal. The candidate point w is the candidate feature vector given a trajectory (v_x, v_y, v_θ).

In order to weight these trajectories using a GDA model, we modified the cost function as follows:

$$cost(v_x, v_y, v_\theta) = \Big(1 - p'(w|\phi)\Big)\Big(\alpha(\Delta_{path}) + \beta(\Delta goal)\Big) \tag{1}$$

Candidate point w's conformity to the given model ϕ was represented as $p'(w|\phi)$. This preserved the existing balance between weights on terms in the original planner achieving good navigation behavior without requiring re-tuning of the bias parameters to accommodate the addition of the social model.

To find $p'(w|\phi)$, we used GDA, as follows. The following equation (the Mahalanobis distance) is the standardized distance of the feature vector w to a given component k of the model (Forsyth & Ponce, 2003):

$$\delta_M(w, k|\phi) = \frac{\sqrt{(w - \mu_{\phi(k)})^T \Sigma_{\phi(k)}^{-1} (w - \mu_{\phi(k)})}}{2} \tag{2}$$

This term is the standardized distance from an individual component of the GMM, taking into account the variance of that component. This value can then be used to calculate the probability that w is part of this model (Watt & Policarpo, 1998):

$$p(w, k|\phi) = \frac{1}{\sqrt{2\pi n}|\Sigma_{\phi(k)}|^{-1}} exp\Big(- \frac{\delta_M(w, k|\phi)}{2}\Big) \tag{3}$$

The probability that w conforms to a given model ϕ is the sum of the probabilities that it conforms to each of the k components of that model:

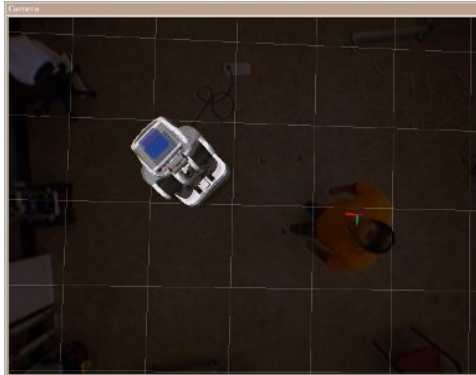$$p'(w|\phi) = \Sigma_k p(w, k|\phi) \tag{4}$$

68

Figure 9. A view of the PR2 robot and a human partner from the overhead camera with a pose overlay of the robot and axes showing the detected position of the social partner.

In summary, the modified cost function, shown in Equation 1, should have a reduced cost for trajectories that conform to the model and an increased cost for trajectories that do not conform to the model.

5.3 Experiment 1: Technical Evaluation of Navigation Planner

To evaluate the performance of this modified trajectory planner, we observed measurable values relevant to the social task. Our validation compared the effectiveness of a robot using the standard trajectory planner provided with the ROS nav_stack to a robot using the modified navigation planner described above. This validation was conducted with the Willow Garage PR2 robot (see Figure 9). This study was also approved by the University of Southern California Institutional Review Board.

5.3.1 Experimental Design The 3x2 experimental design examined two factors, the Behavior Type and the Navigation Planner Type.

The Behavior Type had three levels, each corresponding to a different method for following the robot:

1. Follow as closely as possible, keeping up with the robot (normal);
2. Follow the robot, but at a very slow rate, slower than the robot's normal speed (slow);
3. Wait a few seconds before following the robot (wait).

The different Behavior Types were intended to create a situation where the robot has to wait for the partner to catch up. Behavior 1 represents a condition where the partner followed the robot from the start and the robot did not need to wait for the partner to catch up. Behaviors 2 and 3 represent situations where the the partner initially delayed or moved too slowly so that a standard navigation planner executed a socially sub-optimal trajectory. The above conditions were designed to demonstrate the effectiveness of the planner in a variety of situations.

The Navigation Planner Type had two levels: the standard planner and the modified planner. In the standard planner condition, the stock navigation planner distributed as part of ROS was used to execute the robot's goal-oriented motor actions. No consideration of the partner's position was used as part of the planning process in this condition. In the modified planner condition, the robot used the proposed modifications to the ROS navigation planner that took into account the partner's position when weighting trajectories.

To test this implementation, we recruited neurotypical adult participants and instructed them to act as partners in the following task. In all, 7 participants were recruited (6 male / 1 female) with an average age of 20.8 years. This experiment was conducted in the USC Interaction Lab. Participants were recruited from the USC student community, ranging from those not at all experienced with robotics to those very experienced with robotics. No participants had any knowledge of the different planners used, or of the robot's intentions in the context of the experiment.

Each participant was asked to perform each of the six conditions 5 times, one at a time, for a total of 30 observations per person. The conditions were presented in a randomized order. The participant was asked to follow the PR2 according to one of the behavior types listed above. The robot moved toward a location in the experimental space using one of the two navigation planners.

After the robot reached the goal (defined as being within 0.3m radius of the goal), the distances of the partner with respect to the robot, and with respect to the goal, were recorded. The performance of the standard and modified planner were compared based on those values. Smaller partner-goal distances and partner-robot distances suggested better planner performance for the person-aware following task.
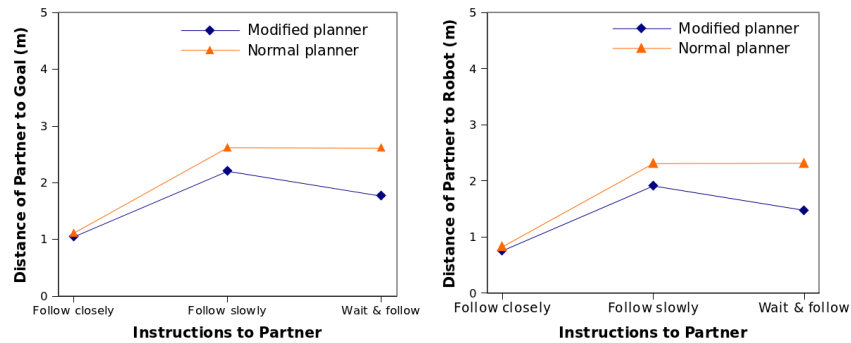

Figure 10. Average distance between partner and goal when robot arrived at goal (lower is better).

5.3.2 Results We predicted that navigation type would not matter when the partner behavior was normal. We also predicted that the partner would be closer to the robot and closer to the goal when the robot reached the goal when navigation type was modified and the blue circle was slow moving or waiting, i.e., an interaction effect. We conducted within-subjects ANOVAs with navigation type, behavior of partner, and their interaction as three fixed factors, and participant as a random factor.

For partner-goal distance, the partner was closer to the goal significantly more in the modified navigation condition (F [1, 210] = 37.43, $p < .0001$). There was a significant interaction effect (F [2, 209] = 11, $p < .0001$). As shown in Figure 10 (left), the difference between the standard and modified planner conditions was much smaller for normal partner behavior than for the slow-moving and wait conditions. A student's t test showed no significant differences between standard and modified planner for the normal partner behavior, but there were significant differences for the slow and wait behaviors ($p < .05$).

For partner-robot distance, the partner was closer to the robot significantly more in the modified navigation condition (F [1, 210] = 36.66, $p < .0001$). There was a significant interaction effect (F [2, 209] = 11, $p < .0001$). As shown in Figure 10 (right), the difference between the standard and modified planner conditions was much smaller for normal partner behavior than for the slow-moving and wait conditions. A student's t test showed no significant differences between standard
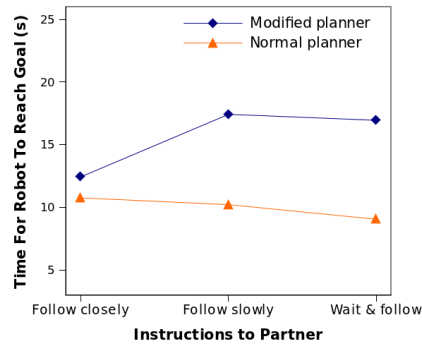
Figure 11. Average time the robot took to arrive at the goal, not related to hypothesis.

and modified planner for the normal partner behavior, but there were significant differences for the slow and wait behaviors ($p < .05$).

The differences between the planner types can be explained by the robot slowing down and stopping in order to wait for the user to catch up. Looking at the total time it took for the robot to reach the goal, we observe that for the modified planner conditions, the robot took longer to reach to the goal significantly more than for the standard conditions (F [1,210] = 68, $p < .0001$). Again, there was a significant interaction effect (F [2, 209] = 9.58, p = .0001). As shown in Figure 11, the difference between standard and modified planner conditions was much smaller for normal partner behavior than for the slow-moving and wait conditions. A student's *t* test showed that the robot in the slow and wait conditions were significantly slower ($p < .05$) than the other modified conditions.

These results support the hypotheses that the modified planner results in the partner being closer to the robot and to the goal when the robot reaches the goal. The data show that the robot takes more time to reach the goal, due to the robot slowing down or stopping to wait for the partner.

5.4 Experiment 2: Social Evaluation of the Navigation Planner

Increased proximity (reduced distance) between the robot and partner when the robot reaches the goal, as observed in the previous experiment, demonstrates that the modified navigation planner was performing as intended. However, conformity to the distance-based social model does not necessarily mean that a partner will notice and have an improved social interpretation of the navigation behavior. This section describes a survey study used to assess the social differences between the original and modified navigation planners. This study was approved by the Yale University Institutional Review Board.

5.4.1 Experiment Design Respondents took the survey using an interactive web page provided by SocialSci, an online social science research utility. Respondents were recruited from the Yale University undergraduate and graduate student community as well as with the SocialSci survey pool, a vetted collection of survey takers similar to Amazon's MTurk. In all, 67 respondents were recruited; their gender and age information was not collected. All survey respondents were over 18 years of age.

Each respondent was shown six videos of the navigation interaction in randomized order. These videos were top-down renderings inspired by Heider and Simmel videos (Heider & Simmel, 1944), but with different behavior content (see Figure 12). The videos contained a red triangle and a blue circle. The triangle and circle were simulated agents; the triangle represented the robot and the circle
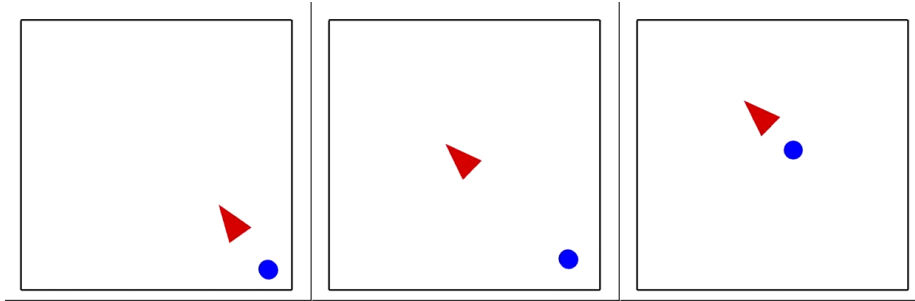
Figure 12. Heider and Simmel-inspired renderings of a simulated robot and partner used in the survey study.

represented the partner. The red triangle moved toward the goal, eventually followed by the blue circle, according to the same instructions used for Experiment 1 (Section 5.3.1).

We employed a 2x3 within-respondent design. The two factors for this design were navigation planner type, and blue circle behavior type. Navigation planner type had two levels: the standard planner and the modified planner. Blue circle behavior type had three levels: normal, slow, and wait. The blue circle's behavior is an automated agent's navigation similar to the rules given to the partner in the previous experiment. Participants saw one video for each combination of navigation planner type and blue circle behavior type.

The intent was to have the participants observe only the spatial relationships between the agents without picking up any social cues that might be observed with a more realistic representation. We used simulated agents rather than videos of the session so that viewers would only consider the navigation behavior and not any other social cues such as gestures or body posture which might be observable from video data.

For each video shown, respondents were asked to rate the behavior of the red triangle (simulated robot) for several social characteristics on a 5-point scale. For example, "To what extent is the red triangle ____ing the blue circle? Not at All (1) ... Very Much (5)." The respondents were asked the above sentence for the following adjectives: avoiding, ignoring, leading, and helping. Each question was presented as a 5-point scale. Order effects were controlled for by randomizing the video presentation order for each participant.

5.4.2 Results We predicted that navigation type would not matter when the blue circle's behavior was normal. We also predicted that the red triangle would be rated more highly for leading and helping, and lower for avoiding and ignoring, when navigation type was modified and the blue circle was slow-moving or waiting, i.e., an interaction effect. We conducted within-subjects ANOVAs with navigation type, behavior of blue circle, and their interaction as three fixed factors, and participant as a random factor.

The prediction was confirmed for three of the four dependent variables. For perceptions of the red triangle as "leading the blue circle," the red triangle was perceived as leading significantly more in the modified navigation condition (F [1, 329] = 150, $p < .0001$), and marginally more when the blue circle was normal (F [2, 329] = 2.9, $p = .06$). The interaction was significant as well (F [2, 329] = 6.3, $p = .002$). As shown in Figure 13 (left), the difference between the modified and standard condition was larger when the blue circle was slow-moving. A student's t test shows that the red triangle in the standard-planner slow-moving condition is rated as leading significantly less ($p < .05$) than the other standard conditions.
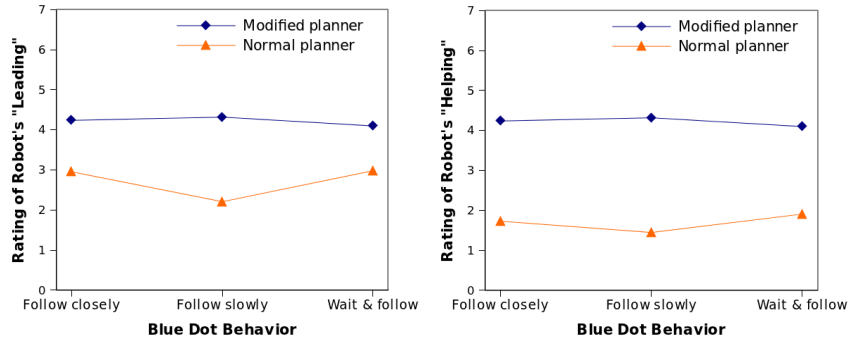
Figure 13. Graph of the differences between survey respondents' ratings of the behavior observed in videos of the standard planner compared to the modified planner for leading and helping (higher is better).
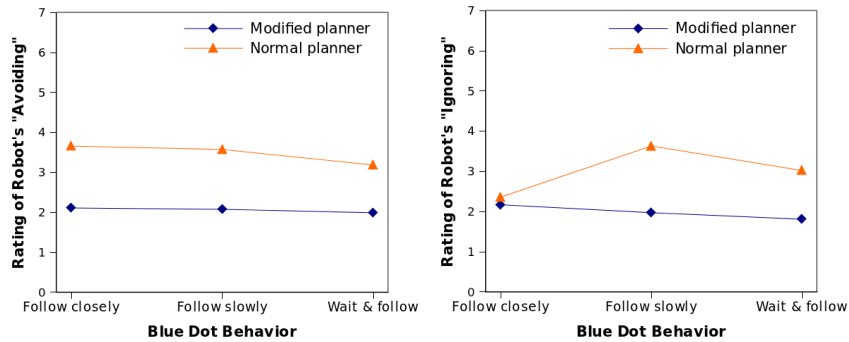

Figure 14. Graph of the differences between survey respondents' ratings of the behavior observed in videos of the standard planner compared to the modified planner for avoiding and ignoring (lower is better).

For perceptions of the red triangle as "helping the blue circle," the red triangle was perceived as helping significantly more in the modified navigation condition (F [1, 329] = 133.5, $p < .0001$), but not when the blue circle was normal (F [2, 329] = 0.22, $p = .8$). The interaction was significant as well (F [2, 329] = 4.23, $p = .015$).

For perceptions of the red triangle as "avoiding the blue circle," the red triangle was perceived as avoiding significantly more in the modified navigation condition (F [1, 329] = 111, $p < .0001$), but did not observe a significant interaction effect (F [2, 329] = .068, $p = .51$).

For perceptions of the red triangle as "ignoring the blue circle," the red triangle was perceived as ignoring significantly more in the modified navigation condition (F [1, 329] = 72, $p < .0001$), even when the blue circle was normal (F [2, 329] = 7.12, $p = .0009$). There was a significant interaction effect (F [2, 329] = 13.16, $p < .0001$). As shown in Figure 14 (right), the difference between the standard and modified planner conditions was much smaller for normal blue circle behavior than for the slow-moving and wait conditions. A student's *t* test shows that the red triangle in the modified-normal condition is rated as ignoring significantly less ($p < .05$) than the other modified conditions.

These results support the hypotheses that the modified planner exhibits less avoidance and ignorance qualities and more leading and helping qualities than the standard planner. These results also support the hypothesis that the modified planner exhibits behavior that is observed as more indicative of leading and less indicative of avoiding or ignoring the partner.

5.5   Discussion

In order to design socially-aware robot navigation behavior, we developed a modified trajectory planner to weight trajectories based on the fitness with a trained model of human following behavior. This modified planner resulted in a robot that slowed down and stopped in order to wait for a following partner to catch up. The validations showed that, while both planners were able to guide the robot to the goal, the modified planner was able to arrive at the goal with the partner much closer to both the robot and the goal. Observers of that behavior rated the modified planner as avoiding or ignoring the partner less and as leading the partner more.

One of the properties of this approach, important for generalizability, is that no parameter tuning was required to make the model work with an already existing general trajectory planner. The system was able to recognize the features in real time and to evaluate the fitness of candidate trajectories with minimal modification of the existing trajectory planning systems. The lack of need for tuning the planner bias parameters suggests that this approach could be used for other data-driven models that could have an impact on trajectory planning. This approach could be employed as implemented to be part of an unstructured play interaction controller.

A limitation of this approach is that it considers only distance to the goal, ignoring the available free space for navigation. The model could be augmented to consider free-space features, such as free space in front of each social agent, distances to walls, and distances to other obstacles, in order to be more informed. However, because the motivating example for this work concerns open-space environments, the limited approach outlined in this section is sufficient for the target social skill.

The intended application of this work was to address the challenge of socially-aware navigation in unstructured social interaction between a child and a socially assistive robot. It would be interesting to study the effectiveness of using such people-aware navigation to encourage and train children with ASD to move close to other people and thereby facilitate social interaction. This model could possibly also be used for detecting whether or not the child is following the robot at all in order to prompt some type of encouraging social intervention.

The evaluation in this section was only concerned with the distance between the robot and partner and a bystander's rating of perceived social behavior. However, a more complete evaluation of this approach would explore the use of subjective measures as well, such as how well/willingly/consistently the child follows the robot, to determine the effectiveness of this approach. The modified trajectory planner used in this section could be applied to other people-aware spatial tasks, such as walking with other people, game-playing scenarios, and turn-taking behaviors.

6.   Conclusion

This article presented two approaches to using spatial models over distance-based features. One was applied to the problem of classifying positive vs. averse child reaction to a robot and the other for generating socially-aware robot navigation behavior. The approaches were validated on fully implemented robot systems tested with with children with ASD and neurotypical adults.

The presented method can be scaled to recognition and generation of other spatial interaction tasks, but has limitations to its scalability beyond tasks which can be represented using continuous spatial data. This method could easily be applied to other movement recognition or generation tasks such as passing another person in the hallway or recognizing if a person is trying to join a two-person interaction. The major requirement for this approach is that the action be adequately

described by the provided distance-based features. This approach cannot be used if the actions are not represented as continuous features. In particular, it would be very difficult to use this approach to model actions such as speech or gesture communication without an explicit spatial or continuous component. However, if such information could be transformed into a relevant continuous variable, then those data could be used as part of a spatial model.

The approaches outlined in this paper are novel in that they use interpersonal distances in addition to distances between agents and elements of the environment to model the social characteristics of spatial movement behavior. The models are then applied to social behavior recognition and socially-appropriate navigation using autonomous robot systems. Prior work has demonstrated either environment-only features or person-only features for modeling data. Also, prior work has not focused on social interpretations of the data, only task-based evaluation. Furthermore, the described social models of spatial behavior are applied to autonomous robot behavior intended for children with ASD. We presented results from experiments using data collected with human-robot interaction between a robot and a child with ASD, a novel application domain for these techniques.

The next logical step for the classification work is to apply it to a real time system whose purpose is to assess the state of the child. The robot could detect a potential problem through observations regarding the avoid, wall, or parent states. Using this information, the robot could change its behavior in response to the child's behavior. Since there was some consistency among children as to which features of the robot bothered them, this would be a signal that such dissatisfaction was occurring. An immediate application for this work is to classify interactions between a child and robot as they are occurring, so as to detect when a child is in distress and have the robot adjust its behavior to better meet the childs needs. An important practical feature of the described modeling and classification approach is its efficiency, making it readily portable to on-line use.

The next logical step for development navigation planner work is to test its performance with children with ASD. The potential study design for this is difficult, however. A good design for such an evaluation would be to compare children's response to a robot using the standard planner for moving over to another person, and a robot using the modified planner to move over to another person. The metrics used in this section would be a good measure of the planner's performance, along with a measure of the percentage of times that the robot attempted to encourage interaction with another person and the child eventually did end up interacting with the other person. Because the preconditions for the robot to begin this action are not dependent on anything manipulable by the experimenters, but rather the social state between the child and robot, it might be difficult to create controlled conditions for such autonomous behavior to occur. However, the robot could initiate such interactions when the child has been engaging in non-averse interaction with it for a certain amount of time.

We hope that the results and insights gained from this work will provide further evidence for the promise of socially assistive robots as therapeutic partners for children in the future.

7.   Acknowledgments

References

Boiman, O., & Irani, M. (2007). Detecting irregularities in images and in video. *International Journal of Computer Vision*, *74*(1), 17–31. http://dx.doi.org/10.1007/s11263-006-0009-9

Chen, S., Lee-Chai, A., & Bargh, J. (2001). Relationship orientation as a moderator of the effects of social power. *Journal of Personality and Social Psychology*, *80*(2), 173. http://dx.doi.org/10.1037//0022-3514.80.2.173

Dean, L., Willis, F., & Hewitt, J. (1975). Initial interaction distance among individuals equal and unequal in military rank. *Journal of Personality and Social Psychology*, *32*(2), 294. http://dx.doi.org/10.1037/0022-3514.32.2.294

Feil-Seifer, D. J., & Matarić, M. J. (2011a, March). Automated Detection and Classification of Positive vs. Negative Robot Interactions With Children With Autism Using Distance-Based Features. In *Proceedings of the International Conference on Human-Robot Interaction* (p. 323-330). Lausanne, Switzerland. http://dx.doi.org/10.1145/1957656.1957785

Feil-Seifer, D. J., & Matarić, M. J. (2011b, August). People-aware navigation for goal-oriented behavior involving a human partner. In *Proceedings of the International Conference on Development and Learning.* Frankfurt am Main, Germany. http://dx.doi.org/10.1109/DEVLRN.2011.6037331

Forsyth, D., & Ponce, J. (2003). *Computer vision, a modern approach*. Upper Saddle River, NJ: Prentice Hall.

Fox, D., Burgard, W., & Thrun, S. (1997). The dynamic window approach to collision avoidance. *Robotics & Automation Magazine, IEEE*, *4*(1), 23–33. http://dx.doi.org/10.1109/100.580977

Gerkey, B., & Konolige, K. (2008, May). Planning and control in unstructured terrain. In *Workshop on Path Planning on Costmaps, Proceedings of the IEEE International Conference on Robotics and Automation (ICRA).* Pasadena, California.

Geschwind, D., Sowinski, J., Lord, C., Iverson, P., Ducat, L., & Spence, S. (2001). The Autism Genetic Resource Exchange: A Resource for the Study of Autism and Related Neuropsychiatric Conditions. *American Journal of Human Genetics*, *69*, 461-463. http://dx.doi.org/10.1086/321292

Hall, E. (1966). *The hidden dimension*. Garden City, N.Y: Doubleday.

Hastie, T., & Tibshirani, R. (1996). Discriminant Analysis by Gaussian mixtures. *Journal of the Royal Statistical Society. Series B (Methodological)*, 155–176.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, *57*, 243-259. http://dx.doi.org/10.2307/1416950

Kelley, R., Tavakkoli, A., King, C., Nicolescu, M., Nicolescu, M., & Bebis, G. (2008, March). Understanding Human Intentions via Hidden Markov Models in Autonomous Mobile Robots. In *Proceedings of the Conference on Human Robot Interaction* (p. 367-374). Amsterdam, Netherlands. http://dx.doi.org/10.1145/1349822.1349870

Kozima, H., Nakagawa, C., & Yasuda, Y. (2005, August). Interactive robots for communication-care: a case-study in autism therapy. In *IEEE International Workshop on Robot and Human Interactive Communication (ROMAN)* (p. 341-346). Nashville, TN. http://dx.doi.org/10.1109/ROMAN.2005.1513802

Lord, C., Risi, S., Lambrecht, L., Jr., E. H. C., Leventhal, B. L., DiLavore, P. C., et al. (2000, June). The Autism Diagnostic Observation Schedule-Generic: A Standard Measure of Social and Communication Deficits Associated with the Spectrum of Autism. *Journal of Autism and Developmental Disorders*, *30*(3), 205-223. http://dx.doi.org/10.1023/A:1005592401947

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *24*(5), 659-685. http://dx.doi.org/10.1007/BF02172145

Mead, R. (2010, August). Space: A Social Frontier. In *Poster at the USC-UTEP Workshop on Predictive Models of Human Communication Dynamics.* Los Angeles, CA.

Mead, R., Atrash, A., & Matarić, M. (2011, May). Automated Analysis of Proxemic Behavior: Leveraging Metrics from the Social Sciences. In *Proceedings of the 2011 Robotics: Science and Systems Workshop on Human-Robot Interaction: Perspectives and Contributions to Robotics from the Human Sciences.* Los Angeles, CA.

Quigley, M., Gerkey, B., Conley, K., Faust, J., Foote, T., Leibs, J., et al. (2009, May). ROS: an Open-Source Robot Operating System. In *ICRA Workshop on Open Source Software.* Kobe, Japan.

Robins, B., Dautenhahn, K., & Dickerson, P. (2009, February). From Isolation to Communication: A Case Study Evaluation of Robot Assisted Play for Children with Autism with a Minimally Expressive Humanoid Robot. In *International Conference on Advances in Computer-Human Interaction* (pp. 205–211). Cancun, Mexico. http://dx.doi.org/10.1109/ACHI.2009.32

Salter, T., Michaud, F., Létourneau, D., Lee, D., & Werry, I. (2007, March). Using Proprioceptive Sensors for Categorizing Interactions. In *Proceedings of the International Conference on Human-Robot Interaction* (p. 105-112). Washington, D.C. http://dx.doi.org/10.1145/1228716.1228731

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2), 461–464. http://dx.doi.org/10.1214/aos/1176344136

Sparrow, S., & Cicchetti, D. (1989). *The vineland adaptive behavior scales.* Boston, MA: Allyn & Bacon.

Watt, A., & Policarpo, F. (1998). *The Computer Image*. New York, NY: ACM Press.

Xiang, T., & Gong, S. (2006). Beyond tracking: Modelling Activity and Understanding Behaviour. *International Journal of Computer Vision*, *67*(1), 21–51. http://dx.doi.org/10.1007/s11263-006-4329-6

Zhong, H., Shi, J., & Visontai, M. (2007, March). Detecting unusual activity in video. In *Proceedings of the Conference on Computer Vision and Pattern Recognition* (p. 819-826). Washington, D.C. http://dx.doi.org/10.1109/CVPR.2004.1315249

David Feil-Seifer, Department of Computer Science, Yale University, david.feil-seifer@yale.edu

Maja J Matarić, Interaction Lab, Departments of Computer Science and Pediatrics, University of Southern California, mataric@usc.edu