A Robot that Approaches Pedestrians

Satoru Satake, Takayuki Kanda, *Member, IEEE*, Dylan F. Glas, Michita Imai, *Member, IEEE*, Hiroshi Ishiguro, *Member, IEEE*, and Norihiro Hagita, *Senior Member, IEEE*

***Abstract*—When robots serve in urban areas such as shopping malls, they will often be required to approach people in order to initiate service. This paper presents a technique for human–robot interaction that enables a robot to approach people who are passing through an environment. For successful approach, our proposed planner first searches for a target person at public distance zones anticipating his/her future position and behavior. It chooses a person who does not seem busy and can be reached from a frontal direction. Once the robot successfully approaches the person within the social distance zone, it identifies the person's reaction and provides a timely response by coordinating its body orientation. The system was tested in a shopping mall and compared with a simple approaching method. The result demonstrates a significant improvement in approaching performance; the simple method was only 35.1% successful, whereas the proposed technique showed a success rate of 55.9%.**

***Index Terms*—Anticipating human behaviors, approaching people, human–robot interaction.**

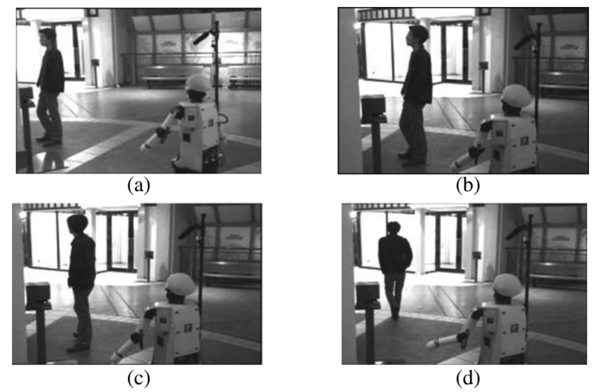


Fig. 1. What's wrong? *Unaware* failure in approach. (a) Robot approached a man looking at a map. (b) Robot started to speak. (c) Turned away from robot. (d) Left without glancing at it.

I. Introduction

Robots have started to move from laboratories to real environments, where they interact with ordinary people who spontaneously interact with them. Robots have been tested in guiding roles in museums [2]–[4] and supermarkets [5]. Social robots, such as receptionists [6] and tutors [7], have been developed to interact like humans, communicating socially with people.

The "initiating interaction" is one of the fundamental capabilities of human–robot interaction for such robots. That is, the initiating interaction would be commonly useful among these robots, while each of them would engage in task-specific interaction for each individual application after initiation. Although many robots are equipped with the capability to invite people into an interaction [8]–[12], these robots only passively wait for people to approach them.

Alternatively, a "mobile" robot can approach people (see Fig. 1) to initiate interaction. This way of providing services is more proactive than waiting, since it enables robots to find and assist people who have potential needs. For instance, imagine a senior citizen who is lost in a mall. If a robot were placed in the mall to provide route directions, it could wait to be approached for help; however, people might not know what the robot can do, or they might hesitate to ask for help. It would be more appropriate for the robot to approach and offer help. Our study presents a method to deal with this novel way of initiating interaction.

A robot's capability to approach people is important for a number of applications. We believe that one promising application is an invitation service; a robot offers shopping information and invites people to visit shops, while giving people the opportunity to interact with it, since robots remain very novel.

Manuscript received February 28, 2012; revised August 4, 2012; accepted October 19, 2012. Date of publication November 27, 2012; date of current version April 1, 2013. This paper was recommended for publication by Associate Editor T. Asfour and Editor J. P. Laumond upon evaluation of the reviewers' comments. This work was supported by the Ministry of Internal Affairs and Communications of Japan. This paper was presented in part at the 4th ACM/IEEE International Conference on Human Robot Interaction, La Jolla, CA, March 9–13, 2009 [1].

S. Satake, T. Kanda, D. F. Glas, and N. Hagita are with the Advanced Telecommunications Research Institute International, Kyoto 619-0288, Japan (e-mail: satoru@atr.jp; kanda@atr.jp; dylan@atr.jp; hagita@atr.jp).

M. Imai is with the Advanced Telecommunications Research Institute International, Kyoto 619-0288, Japan, and also with Keio University, Kanagawa 212-0056, Japan (e-mail: michita@atr.jp).

H. Ishiguro is with Advanced Telecommunications Research Institute International, Kyoto 619-0288, Japan, and also with Osaka University, Osaka 565-0871, Japan (e-mail: ishiguro@sys.es.osaka-u.ac.jp).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TRO.2012.2226387

II. Related Works

Since proactive approaching from a robot is novel, no previous study has reported an integrated method to address its whole process, although each part of the interaction has been addressed to some degree. In this section, we report related works on some aspects of proactive approaching.

A. Interaction and Distance

People engage in different types of interaction depending on the distance separating them. Hall [13] classified human interactions into four categories based on distance: "public distance" (typically >3.5 m), used for situations in which people are speaking to a group; "social distance" (typically between 1.2

and 3.5 m), characterized by situations in which people talk to each other for the first time; "personal distance" (typically between 45 cm and 1.2 m), used for interactions with familiar people; and "intimate distance" ($<$ 45 cm), used for embracing, touching, or whispering [13]. Our approach is related to interaction at both public and social distances. The robot needs to find a person with whom to talk, approach that person from a public distance, and initiate conversation from a social distance.

B. Finding a Person for Interaction

Many previous studies exist for finding and tracking people. Vision as well as distance sensors on robots have been successfully used, even in crowded exhibits [14]. Moreover, researchers have started to use sensors embedded in environments [15]–[17] that enable a robot to recognize people from a distance.

After finding people, the robot needs to identify a person with whom to interact. There are previous studies about human behaviors related to this. For example, Yamazaki *et al.* analyzed how elderly people and caregivers start conversations and found that to identify elderly people who require help, a caregiver nonverbally displays availability with body orientation, head direction, and gaze [18]. Fogarty *et al.* analyzed human interruptibility in an office environment and demonstrated that even simple silence detectors could significantly estimate interruptibility [19].

Other studies involve human–robot interaction, i.e., observing people's behavior directed toward a robot. For example, Michalowski *et al.* classified the space around a robot to distinguish such human levels of engagement as interacting and looking [11]. Bergström *et al.* classified people's motion toward a robot and categorized people into four categories: interested, indecisive, hesitating, and not interested [12]. Tasaki *et al.* developed a robot that chooses a target person based on distance [20]. Finke *et al.* developed a robot that uses a time series of human–robot distances to estimate which of the people passing in front of it are interested in interaction [21]. All of these previous studies addressed people's behavior as those who show interest in interacting with a robot, expressed within a few meters of a robot. However, our problem, making a robot approach pedestrians, requires very different perception of people's motion. It needs to observe people's walking behavior, such as their way of walking, to estimate the possibility of having a conversation.

C. Interaction at Public Distance

A public distance is too far for people to talk, even though they recognize each other's presence. At such a distance, interaction is mainly achieved by changing body position and orientation. Sisbot *et al.* developed a path-planning algorithm that considers people's positions and orientation to avoid disturbances [22]. Pacchierotti *et al.* studied passing behavior and developed a robot that waits to make room for a passing person [23]. Gockley *et al.* found the merits of a direction-following strategy for when a robot is following a person [24].

These robots only use people's current position; however, since human–robot interaction is dynamic, prediction and anticipation are crucial. Hoffman and Breazeal demonstrated the importance of anticipation in a collaborative work context [25]. However, few studies have addressed the anticipation of people's positions. Bennewitz *et al.* utilized such a prediction of position [26], but only to help a robot avoid people, not to enable interaction with them. In a previous study, we anticipated people's positions to let a robot approach them and demonstrated the importance of anticipating positions [27]; however, that work lacked a path-planning process, which is important to notify the target person of the robot's presence.

D. Initiating Conversation at Social Distances

After entering a social distance, a robot initiates a conversation with its target. People usually start conversations with greetings. Goffman suggested that social rules exist for accepting/refusing approaches, including eye-contact, which is a ritual that mutually confirms the start of a conversation [28]. Kendon suggested that friends exchange greetings twice, first nonverbally at a far distance and again at a close distance by smiling [29].

Several previous HRI studies have addressed the greeting process. The importance of greeting behavior is well highlighted in studies in human–robot interaction [6], [30], [31]. Dautenhahn *et al.* studied the comfortable direction for an approach [9] as well as the distance for talking [32]. Yamamoto and Watanabe developed a robot that performs a natural greeting behavior by adjusting the timing of its gestures and utterances [33].

These studies assume that the target person intends to talk with the robot. However, in reality, people are often indecisive about whether to talk when they see a robot for the first time. Studies have been conducted on first-time-meeting situations and making robots nonverbally display a welcoming attitude [11], [12]; however, these passive robots only waited for a person to engage in conversation. Although such a passive attitude is fine for some situations, many situations require a robot to engage in an active approach. Our study aims to allow a robot to actively approach a person to initiate conversation.

"An approach from a robot" is not an easy problem since the robot's approach needs to be acknowledged nonverbally in advance. Otherwise, the person being approached might not recognize that the robot is approaching or might perceive the robot's interruption as impolite. Humans do this well with eye gaze [28], [29], but in a real environment, it is too difficult for a robot to recognize human gaze. Instead, we use the body orientation of the target and the robot for nonverbal interaction.

E. Contingency Detection

The way to start interaction involves the process of identifying contingency, seeking whether the target person reacts in a contingent way toward initiating interaction. In other interaction contexts, the detection of contingency has been studied. Movellan and Watson proposed that information maximization is the basis of contingent behavior [34]. Methods for contingency detection have been proposed [35], [36]. While these studies aim to find a method to detect contingency in generic ways, our study addresses it in a specific but important context: the initiation of interaction.

III. Study Settings

This section introduces our environment, a shopping mall, as well as the hardware of the robot system and the robot's task.

A. Environment

Our study focuses on the initiation of interaction. We aimed to conduct our study in an environment where people could spontaneously decide whether to interact with the robot. Thus, we conducted it in a real social environment.

The robot was placed in a shopping mall that was located between a popular amusement park, Universal Studios Japan, and a train station. The primary visitors to the mall were groups of young people, couples, and families with children. The robot moved within a 20-m section of a corridor (see Fig. 1). Clothing and accessories shops were on one side and an open balcony on the other.

B. Task

The robot's task was advertising shops. The robot was designed to approach visitors and to recommend one of the 24 shops in the mall by providing such shop information as, "It is really hot today, how about an iced coffee at Seattle's Best Coffee?" or "It is really hot today, how about some ice cream?" and pointing at the shop.

Within the scope of this study, we focused on "initiating interaction" in which the robot proactively approaches people. The approaching is, as will be revealed in this study, a process that involves planning as well as exchange of nonverbal behavior. Beyond the phase of initiating interaction, we limited ourselves to only include minimum interaction: There is recommendation behavior exhibited that was a greatly simplified one. The robot did not engage in spoken dialog and kept randomly recommending shops one by one until the visitor walked away.

Visitors to the shopping mall freely interacted with the robot and could quit the interaction anytime. They did not know the existence of the robot because we had not announced our experiment. For safety, our staff monitored the robot from a distant place, which was not visible to visitors; thus, from the visitor's view, the robot seemed to move around and approach them without assistance from human experimenters. When the robot was neither approaching nor talking, it roamed along a predefined route.

C. Hardware and Infrastructure

1) Robot: We used Robovie, a communication robot, which is characterized by its human-like physical expressions. It is 120 cm high, 40 cm in diameter, and is equipped with basic computation resources as well as WiFi communication. Its locomotion platform is a Pioneer 3 DX. We set it to move at a velocity of 300 mm/s (approximately 1.0 km/h) forward and 60°/s for turns. The platform can navigate the robot faster than these parameters (up to 1600 mm/s), but we chose a lower velocity for safety.

2) Laser Range Finders: To approach people, the robot needs to robustly recognize their positions and its own position, even in distant places. We used sensors distributed in the environment to track human and robot positions. Six SICK LMS-200 laser range finders were positioned around the area's perimeter. Laser range finders were set to a maximum detection range of 80 m with a nominal precision of 1 cm, and each scanned an angular area of 180° at a resolution of 0.5°, providing readings every 26 ms.

3) People and Robot Tracking System: These laser range finders were used to track people. We used a technique that is based on the algorithm described by Glas *et al.* [17]. In this algorithm, particle filters are used to estimate people's positions and velocities, and a contouranalysis technique estimates the direction in which a person is facing. This orientation angle helps determine whether the robot should approach a person. This system covers a 20 m × 5 m area and concurrently detects over 20 people's locations. It is also used to localize the robot [37]. Estimation of people's position and localization is performed every 30 ms.

In this environment, the system tracked people with 5-cm accuracy. The localization system usually successfully tracked the robot as well and localized its position within 5-cm accuracy. On rare occasions, when the robot was surrounded by many people trying to interact with it, the robot was not observable from any of the laser range finders, causing serious occlusions and causing the system to fail to track the robot's position. For such unusual failures, a human operator manually corrected the error and restarted the robot's navigation routine.

IV. Modeling of Approach Behavior

Our first attempt to create a simple approach behavior was unsuccessful, and we developed a model of approaching behavior by analyzing the failures. We present failures and the model as motivation for the development of the technique presented in this paper. With the simple approach behavior, the robot simply approached the nearest person.

A. Simple Approach Behavior

Two computation steps were performed in this approach technique. First, the planner must receive people's positions from the *people and robot tracking system* (see Section III-C3) and select a target person to approach. These two steps are performed every 500 ms.

1) Calculating distance for each person i

$$\mathbf{dist_i} = |\mathbf{P_i} - \mathbf{P_r}|$$

where $\mathbf{P}_i$ is the current position of person i, and $\mathbf{P_r}$ is the current position of the robot.

2) Choose the person (i_{target}) whose position is closest to the robot

$$\boldsymbol{i}_{\mathbf{target}} = \arg\min_{\mathrm{i}} \mathrm{dist}_i.$$

Second, the robot executes an approach behavior. Every 30 ms, the robot receives people's positions from the *people and robot tracking system* and actuates its locomotion platform. While distance $\mathbf{dist}_{\mathbf{i_{target}}}$ is greater than 3 m, it directs its motion direction to $\mathbf{P}_{\mathbf{i_{target}}}$ and moves with its maximum velocity

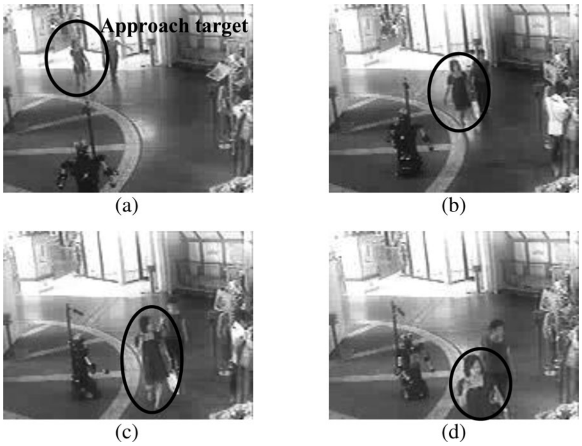


Fig. 2. *Unaware* failure. A person is walking and talking to another person. (a) Robot approached a woman. (b) Robot seemingly in her sight but she paid no attention. (c) She did not see the robot, while it approached her right side. (d) She left.

Fig. 3. *Unsure* failure. A woman unsure whether robot intended to speak to her. (a) Robot approached a person. (b) She stopped when robot started to speak. (c) She observed robot's reaction. (d) She left when robot did not immediately react.

(300 mm/s). When the distance is less than 3 m, it stops moving, initiates a greeting, and starts further conversation.

B. Lessons Learned

Many people ignored the robot when it performed this behavior. These failures, which reflected many problems in the simple approach behavior, were analyzed by watching videos and position data and then categorizing them into four categories: unreachable, unaware, unsure, and rejective.

1) Unreachable: One typical failure is a case where the robot failed to get close enough to engage the target person. This failure happened because 1) the robot was slower than the target person, and/or 2) the robot chose a person who was leaving.

2) Unaware: When a person is unaware of the robot, they do not recognize its action as initiating interaction, even when the robot is speaking to them.

Fig. 1 shows one such failure. In this case, a man was looking at a map on a wall when the robot spoke to him [see Fig. 1(b)], but he was not listening [see Fig. 1(c)] and left without even glancing at the robot [see Fig. 1(d)]. He probably did not hear the robot because the mall was quite noisy. Perhaps, he heard without recognizing that he was being addressed; he might have recognized the robot but simply ignored it.

To avoid this type of failure, the robot could approach such a person before he stopped to look at the map; by approaching from a frontal direction, while the person was still walking, the robot could more effectively make its presence known. (Alternatively, although this is beyond the focus of this paper, the robot should consider how to approach a person who is looking at a target object [38].)

Fig. 2 shows two women walking together [see Fig. 2(a)]. The robot started approaching one of them from the front and seemed to be within her sight [see Fig. 2(b)]. When the robot reached a distance to talk, it approached her right side [see Fig. 2(c)]. Unfortunately, since she was not looking at the robot but at a shop, she ignored the robot as if nothing happened and walked on. To avoid this type of failure, the robot needs to improve its notifying behavior.

3) Unsure: We labeled another type of failure as unsure. Sometimes, although people were aware of the robot, it failed to initiate interaction. They noticed the robot's behavior and recognized its utterances. However, they did not stop since they seemed unsure whether the robot intended to talk to them. Some people even tested the robot's reaction after its greeting, but since the robot was not prepared to react to such testing behaviors, it failed to provide an appropriate reaction. Thus, the robot failed to initiate interaction.

Fig. 3 shows one such failure. A woman and a man entered the environment [see Fig. 3(a)]. The robot approached and greeted the woman. She stopped walking and performed a kind of test by extending her hand to the robot's face [see Fig. 3(c)]. The robot, however, did not respond; therefore, the woman left a few seconds later.

To avoid this type of failure, the robot must unambiguously make the target person understand that it is initiating interaction. Establishing contingency with the person would be useful (e.g., facing the person, reorienting its body to the person, etc.) when the robot is going to initiate an interaction.

4) Rejective: Some people were not interested in conversation with the robot, presumably because they were too busy. These people immediately avoided the robot and refused to talk to it, although they were aware of it and knew that it was addressing them. We called such failures rejections. These people should simply be avoided.

C. Modeling

In response to the lessons learned from our failures in the simple approach behavior, we developed a model for more efficient and polite approach behavior. We propose an approach behavior consisting of the following sequence of phases: 1) finding an interaction target; 2) approaching it at a public distance; and 3) initiating conversation at a social distance.

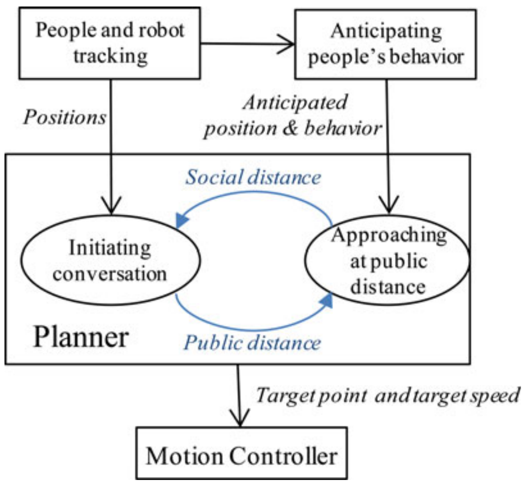


Fig. 4. System architecture.

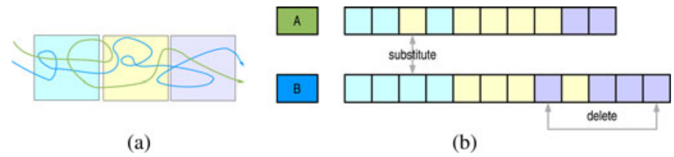


Fig. 5. Comparison of trajectories based on DP matching [27]. (a) Two trajectories. (b) Comparison of state chains of trajectories.

1) Finding an interaction target: The first phase is "finding an interaction target," which is designed to moderate unreachable and rejective failures. The robot needs to predict how people walk and estimate who can be reached with its locomotion capability. It also needs to anticipate whether people might be willing to interact with it; this requirement is especially difficult, but at least it can avoid choosing busy people who are probably unwilling to talk.

2) Approaching the target at a public distance: The second phase is "approaching" the target at a public distance, where the robot announces its presence to the target at a public distance by approaching from the front. Here, the robot must predict the target's walking course to position itself within his/her sight before starting the conversation. Thus, this phase moderates unaware failures.

3) Initiating conversation at a social distance: The last phase is initiating conversation at a social distance to moderate unsure and rejective failures. Perhaps, this can be done simply by such greetings as "hello"; however, greeting strangers is not easy. People are sometimes unaware of the robot's utterance or do not recognize that the greeting is directed at them. We focused on using nonverbal behavior to indicate the robot's intention to initiate a conversation. When the target is about to change her course, the robot faces her so that she can clearly recognize that the robot is trying to interact with her. If she stops, we assume that she is accepting interaction with the robot. After receiving such an acknowledgment, the robot starts a conversation.

V. Robot That Approaches People

A. Overview

There are four techniques involved in our proposed system (see Fig. 4). The *people and robot tracking system* (see Section III-C3) estimates the positions of people and robots using external laser range finders; *anticipating people's behavior* (see Section V-B) refers to a computation of the probabilities of the future position and future local behavior of each person.

The system includes a planner that generates an approach plan and outputs the goal point and goal direction from the approach plan to the motion controller. We followed the three steps in Section IV-C to implement the system, and therefore, the planner supports two approaching modes. In the *approaching at a public distance* mode (see Section V-C), the planner chooses a person to be approached from among the observed people. When the robot arrives within the person's social distance (since Japanese social distance seems to be smaller than Hall's original definition, thus, we set our threshold to be 3 m), it transitions to the *initiating conversation* mode (see Section V-D). Here, the robot observes the person's reaction and provides a timely response to convey the impression that it is intending to interact with the person.

B. Anticipating People's Behavior

For anticipating people's behavior, we used a technique that is based on the algorithm developed by Kanda *et al.* [27]. The basic idea of anticipation is that the future position and future local behavior of a person are likely to resemble those of other people with a similar motion history. For example, a person in a hurry may try to follow the shortest path at a high velocity, while a window shopper may move at a slower speed, following a path that passes close to shops.

Based on the aforementioned idea, a system that anticipates future position and behavior involves two key processes: *offline modeling* and *online execution*. In *offline modeling*, we construct an anticipation model by extracting typical trajectory patterns from the recorded data of visitors to the shopping mall. In the *online execution* phase, the system uses this anticipation model to calculate the probabilities of future positions and the behaviors for each person being observed.

1) Offline Modeling: The anticipation model was constructed in two steps: 1) extracting the typical movement patterns of the visitors at the shopping mall; and 2) calculating the pattern features for anticipation. For the first step, a clustering algorithm was applied to the trajectories that we observed in the environment over 19 days. It consists of 26 863 pieces of trajectory data from visitors who spent more than 0.5 s (average 21.1 s) in the environment where we conducted the experiment. We briefly explain the clustering algorithm here, and further details of it can be found in previous work [27].

First, a person's trajectory, which is a time series of positions represented in the $x - y$ coordinates, is sampled every 500 ms and converted to a state chain, $S^i = \{s^i_{t0}, s^i_{t1}, \ldots\}$ [see Fig. 5(a)]. Spatial partitioning s^i_t is defined as $s^i_t = \{n \in N | p^i_t \in A_n\}$, where A_n is the partition to which the point in trajectory p belongs. We used spatial partitioning based on a 50-cm grid. Second, we applied a k-means clustering algorithm. In the

TABLE I
ATTRIBUTES FOR FREQUENCY DISTRIBUTION

Attribute	Partitioning	Unit size	Symbol
Spatial	2d grid map	25cm x 25cm	s id of grid $\boldsymbol{S}$ entire set of s $\boldsymbol{cp_s}$ center point of grid s
Time	Elapsed time	500 msec	t value of time T entire set of t
Behavior	SVM	4 classes	b value of behavior B entire set of b

clustering, the trajectories were compared based on the physical distance between them at each time step using a dynamic programming (DP) matching method (widely used in many research domains, e.g., [39]) [see Fig. 5(b)], where "insert" and "delete" operation costs in the DP matching were set to be equivalent to 1.0-m distance. The k was set to be 300, and thus, we retrieved 300 clusters.

The second step is the calculation of the pattern features for anticipation. For each cluster, we computed two elements: a *center trajectory* and a *frequency.* The center trajectory is selected from the trajectories in the cluster and represents the cluster's center. For this selection, we define inner distance between a trajectory in the cluster and the cluster itself [see (1)]; in addition, the trajectory, which has the shortest inner distance, is selected as center trajectory. Equation (1) denotes the calculation of the inner distance

$$\mathrm{D}\,(traj, C) = \sum_{traj_c \in C} \mathrm{D}(traj, traj_c) \quad (1)$$

where $\mathrm{D}\,(traj, C)$ means inner distance between trajectory *traj* and cluster $C, traj_c$ means a trajectory in the cluster C, and $\mathrm{D}(traj, traj_c)$ means the distance between trajectories by using DP matching. We describe the center trajectory of cluster c as $Traj_c$.

The *frequency* of $Traj_c$ is denoted as $Freq_c(t,s,b)$ where t represents a 500-ms slice of time, s represents a 50-cm grid in a space, and b represents one of four local behaviors: *idle-walking*, *busy-walking*, *wandering*, and *stopping* (see Table I). These values were computed from the member trajectories in the cluster. For example, if in cluster c, ten trajectories show *idle-walking* behavior in a particular 50-cm grid element s_1 at time t_1 from the beginning of each trajectory, then $Freq_c(t1,s1,idle\text{-}walking)$ equals 10.

2) Online Execution (Anticipation Based on the Model): The probability of future positions and behaviors is calculated by an algorithm based on the idea that future positions and behaviors should resemble those of other people who have exhibited similar histories of positions and behaviors. The algorithm estimates the probability in two steps: 1) calculating the similarity between an observed trajectory and the center trajectory of each cluster; and 2) estimating the probability from the frequency distributions of the most similar cluster(s).

To calculate the similarity, we compared the distance between the observed trajectory and center trajectories $Traj_c$ with the DP matching method. If trajectory i is observed in the area covered by the sensors for t_{observ} seconds, the first t_{observ} seconds of $Traj_c$ are used for DP matching.

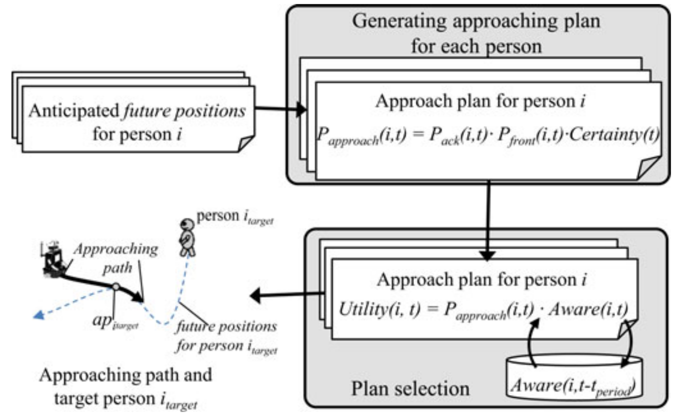


Fig. 6. Overview of generating approaching plan for person i.

Given cluster c whose center trajectory $Traj_c$ is closest to this trajectory i, the estimation of future position is computed with *frequency distribution* $Freq_c(t, s, b)$. Equation (2) denotes the computation of the estimated probability of person i at position grid s_{pred} at future time t_{pred}, with anticipated behavior b_{pred}, $P_i(t_{\mathrm{pred}}, s_{\mathrm{pred}}, b_{\mathrm{pred}})$

$$P_i(t_{\mathrm{pred}}, s_{\mathrm{pred}}, b_{\mathrm{pred}}) = \frac{Freq_c(t_{\mathrm{pred}}, s_{\mathrm{pred}}, b_{\mathrm{pred}})}{\sum_{s \in S} \sum_{b \in B} Freq_c(t_{\mathrm{pred}}, s, b)}. \quad (2)$$

To make the prediction stable, we used the five-best method: 1) selecting the five most similar clusters to the observed trajectory; and 2) averaging estimated probability $P_i(t_{\mathrm{pred}}, s_{\mathrm{pred}}, b_{\mathrm{pred}})$ over the five clusters.

C. Approaching at Public Distance

In the *approaching at a public distance* mode, the robot system selects an appropriate person among the people at a public distance for approaching and generates a path to approach its target. The computation consists of two steps [see Fig. 6]: 1) generating an approaching plan for each observed person (see Section V-C1), and 2) selecting the most promising plan (see Section V-C2).

1) Approach Plan for Each Person: Fig. 6 overviews the processes to generate an approach plan. The primary idea is to make the robot approach from the frontal direction of the target person. An example of such an approaching path is shown by the bold line at the bottom left of Fig. 6. The robot plans to go to an *anticipating point* $(ap_{i_{\mathrm{target}}})$, where the robot arrives before the person arrives, and then it goes along the anticipated trajectories toward the coming person. In the computation, the system computes the *anticipating point,* estimates the *success probability*, and chooses the plan that is most likely to succeed.

Table II shows the algorithm to compute the approach plan that consists of three parts: 1) it calculates the *future positions* of the person as the candidate of the *anticipating points*; 2) it estimates the success probability of approaching for each *future position*; and 3) it selects the future point that has the highest *success probability* as the *anticipating points*.

For calculating the *future positions*, we use the anticipated result of the person [see (2)]: The weighed mean of the positions

TABLE II
ALGORITHM OF SELECTING TURNING POINT FOR PERSON i

1	For each t, *Calculate* $\boldsymbol{fp_i}(t)$; *Calculate* $P_{approach}(i,t)$;
2	Find t_{plan} which satisfy $P_{approach}(i, t_{plan}) = max(P_{approach}(i,t))$; Set anticipating point for person i
3	$\boldsymbol{ap_i} = \boldsymbol{fp_i}(t_{plan})$; $P_{approach}(i) = P_{approach}(i,t_{plan})$;

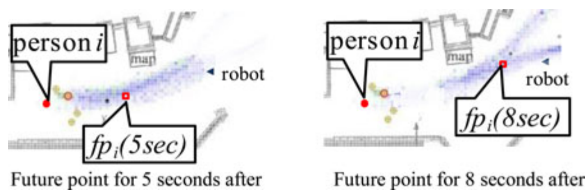


Fig. 7. Calculating future points ($fp_i(t)$). Future point for 5 s after. Future point for 8 s after.

of person i at t seconds after the current time is applied as the *future position* [see (3) and Fig. 7]

$$\boldsymbol{fp_i}(t) = \sum_{s \in S} \sum_{b \in B} \boldsymbol{cp}_s \cdot P_i(t, s, b) \tag{3}$$

where $\boldsymbol{fp}_i(t)$ means the future position of person i at t seconds later, $\boldsymbol{cp}_s$ is the center position of grid s, and $P_i(t,s,b)$ is the estimated probability of person i in grid s at time t that is given by (2). Probability $P_i(t, s, b)$ is used as the weight toward position vector cp_s (center point of grid s).

The estimation of the success of an approach plan for person i ($P_{\text{approach}}(i,t)$) is computed with the following equation:

$$P_{\text{approach}}(i,t) = P_{\text{ack}}(i,t) \cdot P_{\text{front}}(i,t) \cdot Certainty(t) \tag{4}$$

which involves three estimates: $P_{\text{ack}}(i,t)$, $P_{\text{front}}(i,t)$, and $Certainty(t)$. We explain the computation of these estimates in the following.

1) $P_{\text{ack}}(i,t)$: This represents the estimate of the probability that the target person will be willing to interact with the robot. Such an accurate estimate is difficult; instead, as we discussed in Section IV-A, our strategy chooses a person whose future behavior classes are *idle-walking, wandering,* and *stopping* rather than *busy-walking*. Thus, the likelihood value is calculated by

$$P_{\text{act}}(i,t) = 1 - P(t, \boldsymbol{fp_i}(t), busy\text{-}walking) \tag{5}$$

where $P(t, \boldsymbol{fp_i}(t), busy\text{-}walking)$ is the likelihood value of *busy-walking* of $\boldsymbol{fp_i}(t)$ at future time t.

2) $P_{\text{front}}(i,t)$: This represents the probability that the robot will be able to approach the target person from the frontal direction. To do so, the robot needs to appear in advance at the place where the person will come. We used an approximation to estimate this based on the size of the margin time to change the robot's orientation (see (6)). Thus, the margin time is the time difference from when robot arrives at $\boldsymbol{fp}_i(\boldsymbol{t})$ to when person i arrives by the following calculation:

$$t_{\text{margin}}(i,t) = \begin{cases} t - t_{\text{arrive}}(i,t), & t_{\text{arrive}}(i,t) < t \\ 0, & t_{\text{arrive}}(i,t) \geq t \end{cases} \tag{6}$$

where $t_{\text{arrive}}(i,t)$ represents the estimate of the arrival time for the robot to arrive at $\boldsymbol{fp_i}(\boldsymbol{t})$ from the current position. To notify the robot's presence at a public distance, we must choose an approach plan that has high $P_{\text{front}}(i,t)$

$$P_{\text{front}}(i,t) = \begin{cases} t_{\text{margin}}(i,t)/t_{\text{front}}, & \text{if } t_{\text{margin}}(i,t) < t_{\text{front}} \\ 1, & \text{if } t_{\text{margin}}(i,t) \geq t_{\text{front}}. \end{cases} \tag{7}$$

The value for $t_{\text{front}} = 3.6$ s was determined experimentally.

3) Certainty(t): Large uncertainty exists in the prediction of the target person's trajectory in the future. If the system tries to plan further in the future, the anticipation is less likely to be accurate. Thus, we made an approximation of *Certainty*(t) based on a tendency to make it smaller when t is larger (see (8)). The value for $t_{\text{th}} = 40$ s was determined experimentally

$$Certainty(t) = \begin{cases} 1 - t/t_{\text{th}}, & t < t_{\text{th}} \\ 0, & t \geq t_{\text{th}}. \end{cases} \tag{8}$$

2) Plan Selection: The system selects a person to maximize the likelihood of a successful approach. Here, the estimated probability for the success of the approach toward person i ($P_{\text{approach}}(i,t)$) only considers information from the current moment. However, situations change over time, and thus, a robot may need to change its approach target. For this problem, the system periodically recalculates the anticipation and selects the most promising target person to approach.

However, if we only rely on the information from each current moment, the robot might sometimes frequently switch among two or more people; thus, its motion would be less stable and less efficient. To address this in the plan selection, we estimated the extent to which a person is likely to be "aware" of the robot by computing the amount of the robot's visual exposure to the pedestrian over time. In concrete, the algorithm of plan selection uses a utility function to select and update the approaching target over time; the person i, who has the maximum utility at the current time (t_{current}), is selected as approaching target. The utility function is defined as

$$\text{Utility}(i, t_{\text{current}}) = P_{\text{approach}}(i) \cdot Aware(i, t_{\text{current}}) \tag{9}$$

where *Aware*(i, t_{current}) represents the estimated degree of person i's awareness of the robot by the current time t_{current}, while $P_{\text{approach}}(i)$ is computed with the algorithm presented in Table II, representing the estimate of the probability of a successful approach considering the information from the current moment. By maximizing this utility, the system selects a person who is likely on the successful approaching path at current moment ($P_{\text{approach}}(i)$) as well as to whom the robot is likely to be exposed well over time (*Aware*(i, t)).

We consider that the *person's awareness of the robot* depends on its relative position; the likelihood that the person is aware of the robot increases 1) if the robot is *visible* to the person for a long time, and 2) if the robot is *coming* toward the person for a long time. Based on this idea, we calculate the *person's awareness of the robot* by (10). It is the sum of past awareness (considering long-time awareness) and the current *facing* and *coming* relationship

$$\begin{aligned} Aware(i,t) &= \alpha \cdot Aware(i, t - t_{\text{period}}) \\ &\quad + \beta \cdot Visible(i) \cdot Coming(i) \end{aligned} \tag{10}$$

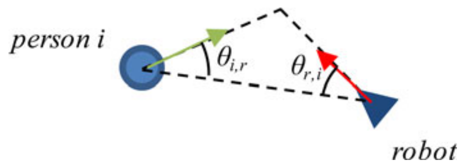


Fig. 8. Calculation of *Facing(i)* and *Coming(i)*.

where t_{period} represents the time period between the current and previous plan selection (i.e., 500 ms in our implementation). $Visible(i)$ and $Coming(i)$represent the current *visible* and *coming* relationships, respectively, that we will explain in the following. The values for $\alpha = 0.72$ and $\beta = 0.28$ were determined experimentally. Note that we set a lower boundary for $Aware\,(i,t)$; when $Aware\,(i,t)$ is smaller than small constant value $AwareTh$, we used $AwareTh$ instead of its original value. This lower boundary represents the fact that even when a person is not aware, sometimes it is possible for a robot to successfully approach.

Fig. 8 illustrates the spatial relationship of target person i and the robot. *Visible*(i) is concerned with whether the robot is within the frontal direction of the person so that it is visible to the person. We empirically decided threshold angle $\theta_{\text{visible}} = \frac{\pi}{3}\,(60^\circ)$ by assuming that the person can observe the robot if it is within the angle. *Visible*(i) is defined as

$$Visible\,(i) = \begin{cases} 1 - \dfrac{\theta_{i,r}}{\theta_{\text{th}}}, & \text{if } \theta_{i,r} < \theta_{\text{th}} \\ -\dfrac{\theta_{i,r-\theta_{\text{th}}}}{\pi - \theta_{\text{th}}}, & \text{otherwise} \end{cases} \tag{11}$$

where $\theta_{i,r}$ is the angle of the robot relative to the person's motion direction (see Fig. 8). Note that *Visible*(i) function outputs 1 if the robot is exactly in the direction of the person's motion direction, outputs 0 if the robot is in the threshold angle, and outputs -1 if the robot is directly behind the person (i.e., the opposite direction of the person's motion direction).

Coming(i) addresses whether people would perceive the robot as "coming" toward them. We observed that a robot moving away from a person creates the impression that the robot is not trying to interact with the person. Furthermore, we observed that people are not concerned with this factor if the robot is moving to their side or back. Thus, we only computed the perception of the robot to be "coming" when it is within the visible angle. It is defined as

$$Coming\,(i) = \begin{cases} 1 - \dfrac{\theta_{r,i}}{\pi/2}, & \text{if } \theta_{i,r} < \theta_{\text{th}} \\ 1, & \text{otherwise} \end{cases} \tag{12}$$

where $\theta_{i,r}$ is the angle of the robot relative to the person's motion direction, and $\theta_{r,i}$ is the angle of the person relative to the robot's motion direction (see Fig. 8).

3) Plan Execution: Finally, the system executes the plan. With the algorithm described at plan selection, target person i_{target} is chosen; thus, there is an *anticipating point* for plan $\boldsymbol{ap}_{\boldsymbol{i}_{\text{target}}}$. Every 30 ms, the robot receives people's positions from the *people and robot tracking system* and actuates its locomotion platform. If the robot has not yet arrived within 1 m from the *anticipating point,* it directs its motion direction to $\boldsymbol{ap}_{\boldsymbol{i}_{\text{target}}}$ and moves; once it has arrived within 1 m of the *anticipating point*, it directs its motion direction to the position of person $\mathbf{P}_{\boldsymbol{i}_{\text{target}}}$. When the distance to the person is within 3 m, it stops approaching and transits to the *initiating conversation* mode.

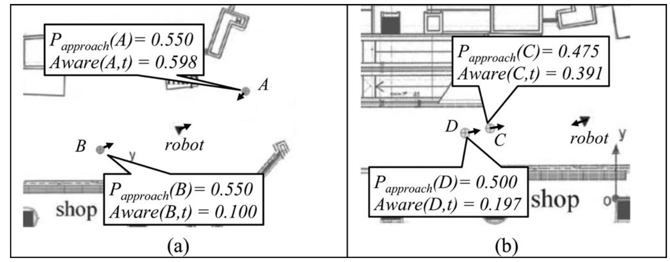


Fig. 9. Choice of approaching target considering "awareness."

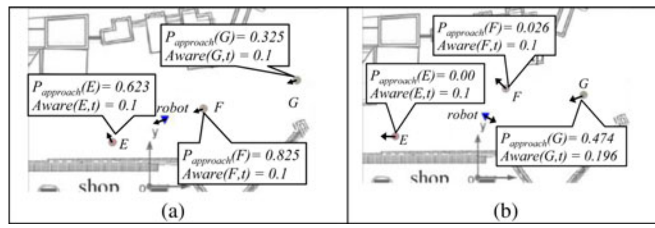


Fig. 10. Switching of the approaching target. (a) Person F is selected as approaching target. (b) Approaching target is switched to person G.

4) Example: Fig. 9 illustrates examples of how the system works. The spatial relationships of the pedestrians and the robot, as well as relevant probabilities, are illustrated. Fig. 9(a) shows a scene in which a robot started to approach target *A*. The estimate of the success of frontal approaching $(P_{\text{approach}}(i))$ is equal for both persons A and B, but since the robot and person A are already facing each other, the estimate of awareness $(Aware\,(i,t))$ is higher for person *A*. Thus, the system approached person A.

In Fig. 9(b), two people are coming from the frontal direction of the robot. In the previous moments, the robot was already facing person C; therefore, the estimate of the awareness for person C $(Aware\,(C,t))$ was higher than for person D. Thus, the robot chose to keep approaching person C.

In both cases, the awareness computation saved the robot from a situation in which it might frequently switch between two approaching targets. Once it starts approaching, and as long as the approaching goes well, the estimate of the awareness increases and stays high, and thus, the robot tends to keep approaching the same target once it has been chosen.

By contrast, the example in Fig. 10 shows the awareness computation successfully helping switch the approach target when necessary, without introducing unnecessary instability. In scene (a), the robot was initially oriented toward person E. However, the system estimated that person F's future position was more likely to be a successful approach target. Another person G was also coming but was estimated to be relatively less likely because that person G was still far away and had a relatively small *certainty* value. Thus, at this point, the robot started to rotate toward person F.

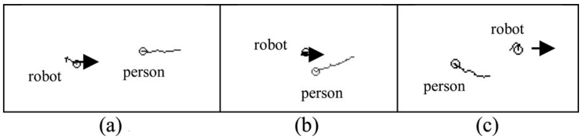


Fig. 11. Classification of reaction of approaching people. (a) Approaching. (b) Passing. (c) Leaving.

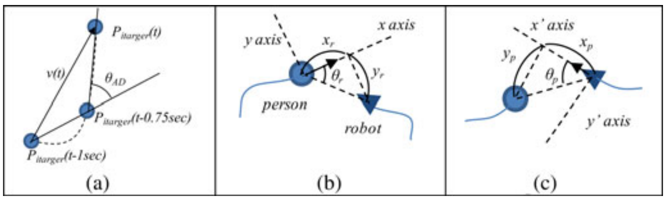


Fig. 12. Features for classification. (a) Velocity and angle deviation. (b) Relative position of the robot. (c) Relative position of the person.

After 2.9 s, however, person F changed course [in scene (b)], contrary to the previous prediction. Now, there is little possibility of success for approaching person F. In contrast, person G kept moving through the corridor, and at this point, the estimate of a successful approach was considerably high for person G. In addition, now the orientation of the robot was closer to facing person G; therefore, the *awareness* value had also increased. At this point, the robot began to approach person G.

D. Initiating Conversation

This process is for the other mode of the robot planner, in which it has finished approaching the target person and has entered the social distance zone. Here, the robot is about to start a conversation with the person. The control aim at this stage is to clearly show that the robot's intention is interaction with this target person. As our failures in our simple approach showed, simply uttering a greeting is inadequate. We implemented a behavior to express contingency toward the target person, mainly using the robot's body direction. In this behavior, the robot quickly orients its body direction toward the target person when it detects that the person is about to pass by. This involves the following two computation steps.

1) Classifying the Reaction of Approaching Person: To identify people's passing-through action, we used a support vector machine to classify the trajectory of the approaching person into four reaction classes: *approaching*, *passing*, *stopping*, and *leaving* (see Fig. 11).

The classification used the following features.

a) Features from approaching person's trajectory

1) *Velocity of approaching person:* The average velocity during last 1.1 s is computed [see Fig. 12(a)]

$$|\mathrm{v}\,(\mathrm{t})| = \left|\mathbf{P}_{\boldsymbol{i}_{\text{target}}}\,(t) - \mathbf{P}_{\boldsymbol{i}_{\text{target}}}\,(t - 1.1\ \mathrm{s})\right|.$$

2) *Angle deviation of approaching person:* As shown in Fig. 12(a), it is computed as the difference of two angles of the person's motion vectors. This value is large if a person quickly changes his/her course within the last few hundred milliseconds

$$\begin{aligned}\theta_{\mathrm{AD}} = |&\tan^{-1}(\mathbf{P}_{\boldsymbol{i}_{\text{target}}}\,(t - 0.85\ \mathrm{s}) \\ &- \mathbf{P}_{\boldsymbol{i}_{\text{target}}}\,(t - 1.1\ \mathrm{s})) \\ &- \tan^{-1}(\mathbf{P}_{\boldsymbol{i}_{\text{target}}}\,(t) \\ &- \mathbf{P}_{\boldsymbol{i}_{\text{target}}}\,(t - 0.85\ \mathrm{s}))|.\end{aligned}$$

b) Features from each pair of trajectories

1) *Relative position of robot (x_r, y_r, and θ_r):* As shown in Fig. 12(b), the system computes its relative position of the robot in a coordinate where the person's motion direction (computed from $\mathbf{P}_{\boldsymbol{i}_{\text{target}}}\,(t) - \mathbf{P}_{\boldsymbol{i}_{\text{target}}}\,(t - 1.1\ \mathrm{s})$) is the x-axis. Thus, for instance, if the person is moving toward the robot, y_r is nearly zero. The relative direction is computed as well ($\theta_r = \tan^{-1} |y_r|/_{x_r}$).

2) *Relative position of the person from robot (x_p, y_p, and θ_p):* As shown in Fig. 12(c), the system computes the relative position of the person in a coordinate where the robot's motion direction (computed from $\mathbf{P_r}\,(t) - \mathbf{P_r}(t - 1.1\ \mathrm{s})$) is the x-axis. The relative direction is computed as well ($\theta_p = \tan^{-1} |y_p|/_{x_p}$).

We used 45 trajectory pairs where the robot approached visitors. The accuracy of the classification was tested with the leave-one-out method, which yields an 88.9% recognition rate.

2) Generating Robot's Motion: The robot changes its motion depending on each class of the reaction of the approaching target.

a) Approaching: When the person is reacting as *approaching*, the robot continues to approach the target person. We applied a position control method, whose target is the position of the approaching person, to continue approaching.

b) Passing: When the person is reacting as *passing*, the robot must immediately react to this situation. This reaction typically happens when the approaching person is unsure of the robot's intention to interact; thus, we control the robot so that its body orientation quickly faces the approaching person to show its intention to interact. We set the robot's rotation velocity to 60°/s toward the direction of the person until the robot is facing the direction of the person.

c) Stopping: This is the case where the approaching person stops in front of the robot. This typically happens when the person accepts interaction with the robot. This is the moment when the robot should start a conversation.

d) Leaving: In this situation, since the approaching person is leaving the robot, it should abandon a conversation with this person and seek another approaching target.

VI. Field Trial

We conducted a field trial at a shopping mall[1] to evaluate the effectiveness of our proposed approach behavior. The robot's

[1]In this study, we obtained approval from the shopping mall administrators for this recording under the condition that the information collected would be carefully managed and only used for research purposes. The experimental protocol was reviewed and approved by our institutional review board.

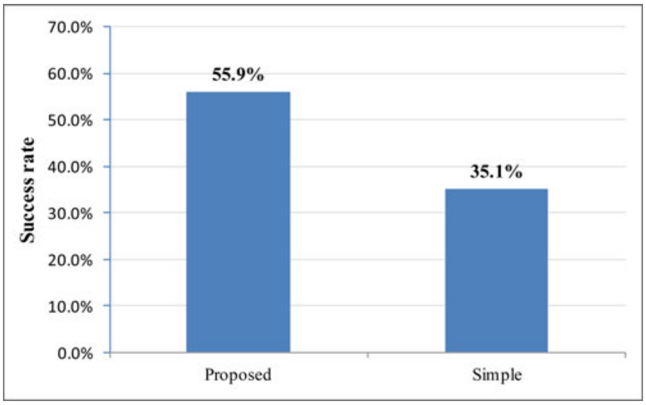


Fig. 13. Results of field trial.

task was to approach visitors to provide shopping information. The details of the environment and task are described inbrk Section III.

A. Procedure

We compared the proposed method with a simple approach behavior to evaluate its effectiveness. Since no existing method addresses the process of approaching a walking person, we used an approach behavior based on common-sense considerations for our comparison. Specifically, the "simple approach behavior" reported in Section IV was used, which is based on the assumption that a robot can be successful by simply approaching the nearest person. In both methods, the same infrastructure, the robot hardware (see Section III-C1) and people tracking and localization (see Section III-C2) were used. The details of the proposed approach behavior are reported in Section V.

For comparison, we ran the trials for several sessions to eliminate such environmental effects as the time of the trial. Each session lasted about 30 min. The two conditions, simple and proposed, were assigned into sessions whose order was counterbalanced. We ran the experiment for 2 h for each condition, and about the same number of approach behaviors (59 for the proposed method and 57 for the simple method) were conducted under each condition. The number of people in the corridor was also equivalent under both conditions. On average, there were 4.09 people for the proposed method and 4.06 people for the simple method (distributed between one and 14 persons, standard deviation (s.d.) = 2.70 and 3.68).

B. Improvement of Success Rate

Fig. 13 shows the comparison results. The approach behavior was defined as successful when the robot's approach target stopped and listened to the robot's utterance at least until the end of its greeting. In this section, we defined the term "trials" to denote actual approaches toward people and not simply the number of people passing through the area.

With the proposed method, the robot was successful in 33 approaches out of 59 trials (252 people passed through). On the other hand, with the simple method, the robot was only successful 20 times out of 57 trials (221 people passed through).

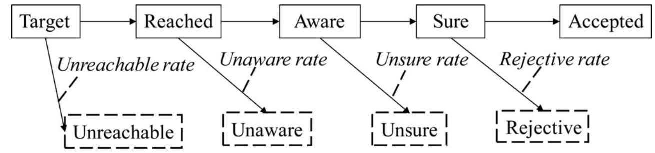


Fig. 14. Calculating failure ratio at each step.

TABLE III
RATIO OF OBSERVED FAILURE IN EACH STEP

	Proposed	Simple
Unreachable	3%	25%
Unaware	4%	14%
Unsure	18%	24%
Rejective	27%	29%

A chi-square test revealed significant differences among conditions ($\chi^2(1) = 5.076$, $p < 0.05$). Residual analysis revealed that in the proposed method, successful approaches were significantly high ($p < 0.05$), and failed approaches were significantly low ($p < 0.05$). Thus, the experimental result indicates that the proposed method contributed to greater successful approach behavior.

C. Detailed Analysis of Failures

Based on the criteria of Section IV-B, to reveal why the failures decreased in our proposed approach, a third person without knowledge of our research purpose classified them by watching videos and position data during the field trial.

These failures are consequentially related: unaware failure only happened when the robot reached the person, and unsure failure only happened when the person was aware of the robot. Only a sure person rejected the approach. Thus, we can model these processes as a probabilistic state transition. Fig. 14 summarizes the calculations at each failure category.

Table III shows the failure rate at each step under each condition. In Table III, the denominators that were used to calculate each failure are different; for example, "unreachable" failures were counted as a fraction of the number of approach attempts, but "unaware" failures were only counted among the number of people reached by the robot.

Overall, this result indicates improvement in the unreachable, unaware, and unsure steps from the simple approach behavior. Unreachable and unaware failures largely decreased.

D. Detailed Analysis of System Performance

We further analyzed how the proposed system worked, how it contributed to reduce failures, and what are the missing capabilities to further reduce the failures. Three analyses were conducted.

1) Accuracy of Anticipation for Person's Position: One of the key computations related to *unreachable* and *unaware* failures is the anticipation of people's future positions. Thus, we evaluated the anticipation accuracy.

To evaluate the effect of the anticipation, we evaluated two approach methods.

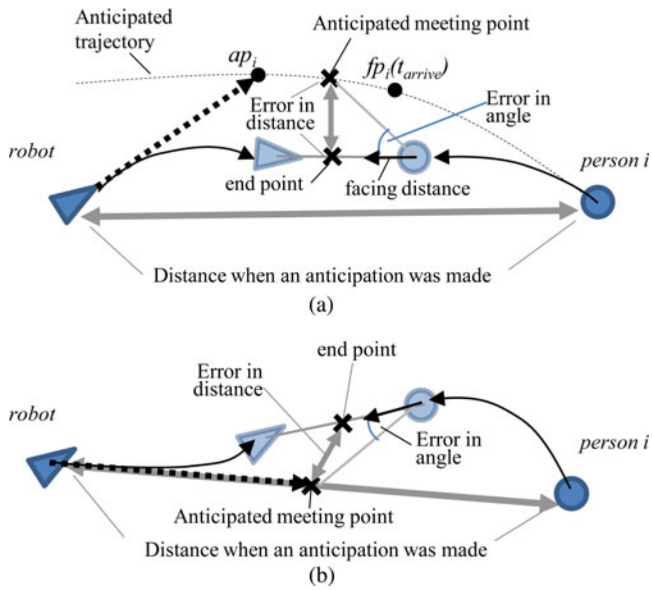


Fig. 15. Evaluating anticipating accuracy. (a) Evaluating anticipating accuracy for the proposed approach. (b) Evaluating anticipating accuracy for the simple approach.

a) Proposed approach: The robot approaches to a person by using developed techniques described in Section V. Its anticipation accuracy is evaluated.

b) Simple approach: As a comparison, we evaluated the anticipation performance of the simple approach. Although there is no explicit use of anticipation technique, it could be considered as the anticipation that the robot and target person would meet in the middle of their current position (thus, the robot directly moves toward the person).

Fig. 15 illustrates the idea behind this evaluation. The detailed computation is described in the following.

Ground truth: The "end point" is the point when the approaching behavior lasted either when 1) the initiating interaction is successful; 2) the failure of approaching because the person exit from experiment area; or 3) the robot gave up approaching because estimated change is too low.

Measurement: In the proposed method, the "anticipated" meeting point is computed as the middle point of the *anticipated point* (ap_i), which is the location at which the robot planned to arrive at time t_{arrive}, and $fp_i(t_{\mathrm{arrival}})$, which is the expected position of the person at time t_{arrive}. The facing direction [see Fig. 15(b)] is the walking direction of the person when the person actually met. In the simple approach method, the anticipated meeting point is computed as the middle point of the robot's position and the target person's position.

Anticipation error: We evaluate two types of anticipation errors: the error in distance and the error in angle. The *error in distance* is computed as the distance between the *real* and *anticipated meeting points*. The large error in distance would result in unreachable failure. The *error in angle* is computed as the degree from the facing direction to the *anticipated meeting points*. A large error in angle would result in failure in locating the robot in wrong direction from the person's view, and, thus, cause unaware failure.

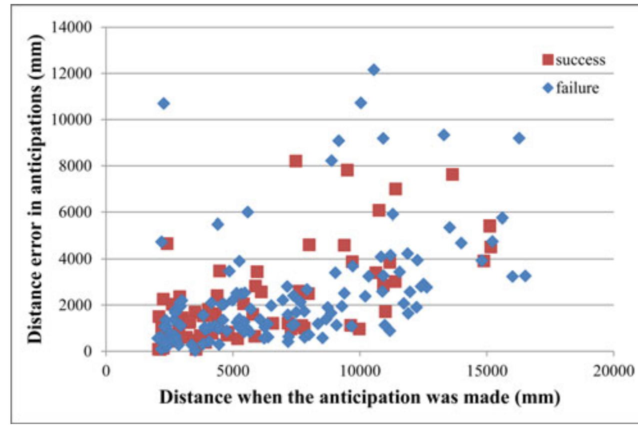


Fig. 16. Distance error in anticipations for the proposed method.

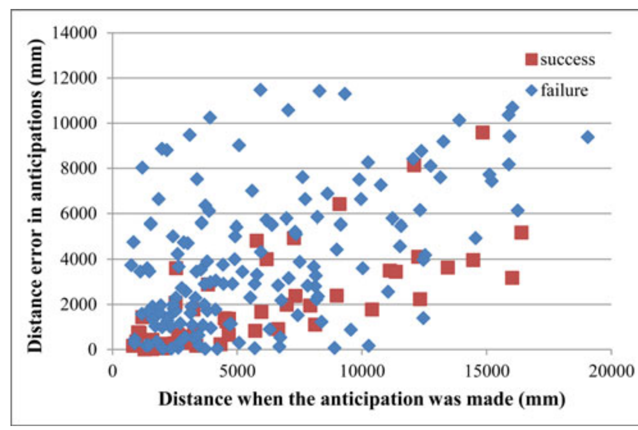


Fig. 17. Distance error in anticipations for the simple method.

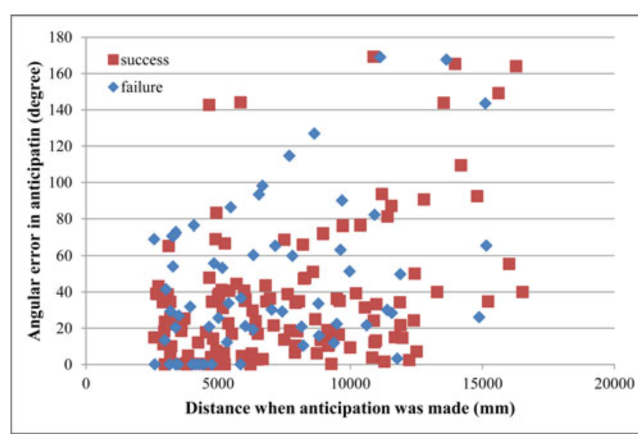


Fig. 18. Angular error in anticipations for the proposed method.

The result of anticipation error: Figs. 16–19 show the result. The horizontal axis shows the *anticipation distance*, which is the distance from the robot to the target person when the anticipation was made. The vertical axis of Figs. 16 and 17 shows

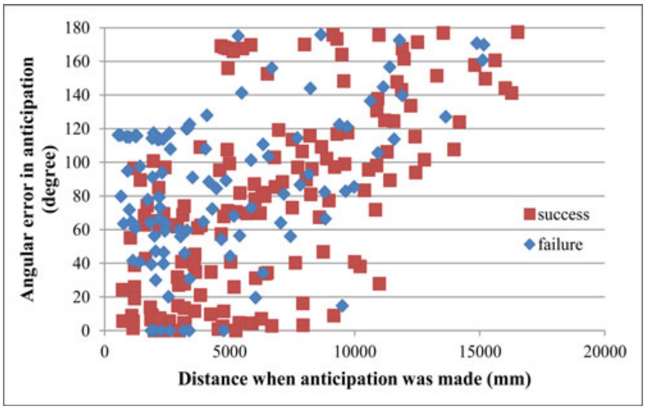


Fig. 19. Angular error in anticipations for the simple method.

the *anticipation error in distance*; the vertical axis of Figs. 18 and 19 shows the *anticipation error in angle*. The plot has two types of markers: the square markers show the result of the success approach, and the triangle markers show the result of the failure approach. 792 anticipations were made during the 59 approaches with the proposed method (see Figs. 16 and 18). 1276 anticipations were made during 57 approaches with the simple method (see Figs. 17 and 19). Note that no anticipation was made when the distance to the person was less than 3 m, since the robot transited to the "initiating conversation" mode.

We identified three findings. First, the accuracy of the proposed approach is higher than the simple one. *The average of distance error in anticipations* on the proposed was 2.27 m, and the error on the simple was 3.38 m. *The average of angular error in anticipations* on the proposed method was 39.1°, and the error on the simple method was 81.1°. It is notable that the angular error is so large in simple method. This is often caused by unreachable failure case: The simple approach sometimes tried to move toward the person who was going to left from the robot. Such a person faces toward the opposite direction, and thus, the error in angle is nearby 180°. Such a case is much fewer in the proposed method.

Second, as expected, the anticipation errors in distance were smaller when the anticipation distance was smaller, because it is rather difficult to anticipate position with more future time. For instance, when the anticipated distance was larger than 5 m, the average of the *anticipation errors in distance* was 3.05 m, and the average of *anticipation errors in angle* was 45.1°. When the anticipation distance was less than 5 m, the *anticipation errors in distance* averaged 0.648 m, and the *anticipation errors in angle* averaged 28.7°.

Third, there are rather small differences between errors in distance in the successful and failure cases. For cases when the anticipated distance was larger than 5 m, the average error was 3.21 m (s.d. 2.19 m) for success cases and 2.97 m (s.d. 2.51 m) for failure cases. When the anticipated distance was less than 5 m, the error was 1.19 m (s.d. 0.97 m) for success cases and 1.51 m (s.d. 1.67 m) for failure cases. Thus, in either successful or failure case, the anticipation error was rather similar, and the robot was able to approach nearby pedestrians using the anticipation computation.

On the other hand, as shown in Fig. 17, in case of a failure approach, it rather failed to anticipate direction of the target person. When the anticipated distance was larger than 5 m, the average error was 41.0° (s.d. 39.7°) for success cases and 54.5° (s.d. 43.9°) for failure cases. When the anticipated distance was less than 5 m, the error was 27.9° (s.d. 25.5°) for success cases and 30.2 degree (s.d. 29.2°) for failure cases. This implies that failure in anticipating direction could result in failure in approach, particularly when the robot is rather at far distance (larger than 5 m).

Note that since our anticipation method does not consider people's behavior toward the robot, one might wonder to what extent anticipation could be correctly done if people changed their trajectory when approached by the robot. First, in only 3.4% of the cases did people seem to change their course to avoid the robot. For such cases, anticipation does not work well. In contrast, when people did not change their course but only changed their speed (e.g., slowing down when the robot approached), their new trajectory resembled other patterns of stored trajectories from pedestrians who walked rather slowly. The system was able to anticipate their positions.

Overall, we believe that this anticipation computation provide much better estimate about people's future position in comparison with the simple approach method, thus reduced the number of "unreachable" failures. The proposed method also reduced angular error, meaning that the robot was more likely to approach the target person from frontal direction. It is reported that people's view angle during walking is about 80° (40° for each side) [40]. Thus, an average of 39.1° error would mean that often the robot was in the sight of the target person. In other words, the system successfully planned a frontal approach in order to announce its presence to the approached person owe to the anticipation technique. Thus, we believe that anticipation also reduced "unaware" failures, although there would be a further possibility to improve approaching performance if we can make the anticipation more accurate.

2) Accuracy of Anticipation for Person's Behavior: Aiming to reduce the number of "rejective" failures, in the computation of $P_{\rm ack}(i,t)$ in (5), we used our anticipation technique for future local behavior (i.e., one of four classes: *idle-walking*, *busy-walking*, *wandering*, and *stopping*) and computed whether a person will do *busy-walking. S*uch a busy person is less likely to accept initiation from a robot. We evaluated the anticipation accuracy of the *busy-walking* behavior.

Ground truth: For each approached person, two coders, who were not informed about the system's output, classified the pedestrian behaviors, when the robot was near the person, into two categories: *busy-walking* and *other.* Cohen's kappa coefficient from their classifications was 0.727.

Evaluation: For each category of *busy-walking* and *others*, the likelihood estimate output from the system was compared at each moment the system computed the approaching behavior, i.e., for each 500 ms.

Fig. 20 shows the likelihood output. The horizontal axis shows the *anticipation distance*, which is the distance to the position of the person when the anticipation was made. The vertical axis shows the average likelihood of *busy-walking* output by the

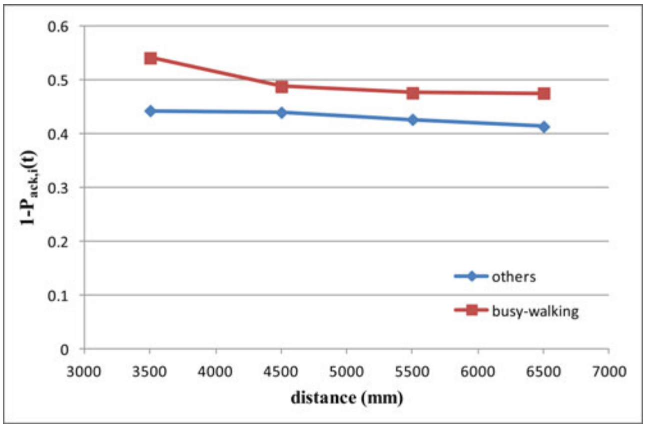


Fig. 20. Anticipation accuracy for busy-walking.

system. The two lines correspond to the ground truth categories of *busy-walking* and *other*.

The result shows that the system output higher likelihood of future *busy-walking* for people who finally performed *busy-walking* than for people who did *other* behaviors. This tendency was more prominent when the anticipation distance was within 3.5 m. As expected, the system was more accurate in anticipation for the near future than for the more distant future.

On the other hand, there was a relatively small 0.1 point difference in output value, even when the anticipation distance was 3.5 m. This is one potential reason why the system was not so successful in reducing *rejective failures*. It may have still been considering people who do *busy-walking*. However, further analysis revealed that only 5.0% of the approached targets were *busy-walking* when approached. Perhaps, the people who did *busy-walking* walked fast and, thus, were less likely to be computed as approachable.

Among the *rejective* failure cases, the ratio of *busy-walking* people was large. 15.4% of all *rejective* failure cases were people who did *busy-walking*. Yet, the remaining 84.6% *rejective* failures were cases where people did *other* behaviors. Overall, we believe that there are many other reasons for rejection beyond walking quickly, including that some were just not in the mood to stop and talk with a robot. It is unrealistic to expect that all people accept such an advertisement service from a robot.

3) Accuracy of Classifying People's Reaction in Conversational Distance: When the robot transited to the "initiating conversation" mode, it observed the target person's reaction. We evaluated the accuracy of this classification.

Ground truth: For each approached person, two coders, who were not informed about the system's output, independently classified the reactions of all the approached people. They used the same classification as the system did; they labeled people's reactions as *approaching*, *passing*, *stopping*, and *leaving*. They specified the point when people's behavior changed and provided labels in a continuous manner for the time series. Cohen's kappa coefficient from the two coders' classifications was 0.869.

Classification accuracy: The system computed this every 100 ms. For each classification the system made, we compared whether the output from the system matched the label given by the coder.

TABLE IV
ACCURACY OF CLASSIFICATION OF PEOPLE'S REACTION

	approaching	passing	leaving	stopping	total
result	75.0 %	56.1 %	71.8 %	95.2%	73.2%

Reaction delay: the classification goal is to detect the moment when people's behavior changed from *approaching* to *passing*. We evaluated the delay from the moment when the person's behavior changed to *passing* (given by the coder) to the moment when the robot physically started to change its body direction toward the person.

Table IV shows the classification accuracy results, which are separated per category in ground truth. The classification accuracy was 73.2%. It performed reasonably well for three categories: *approaching*, *leaving*, and *stopping*, but accuracy for *passing* was relatively low. Apparently, *passing* is the most difficult category to recognize. About half of the *passing* situations (19.9% of the total) were finally recognized correctly, but not exactly when they happened. This would cause a delay in the robot's reaction. The system recognized *passing* an average of 153 ms after the ground truth timing. Such a delay was relatively small in contrast with people's passing behavior, which typically takes a second or longer. Overall, the robot finally performed reactions toward 76.0% of the *passing* people as designed. Since the remaining half of the *passing* situations (24.0% of the total) were simply overlooked, the robot failed to display a reaction to those people as they passed.

In addition to the recognition delay, there were other delays. For instance, computation only happened every 100 ms, and there were delays in transferring commands and in controlling the actuators. The analysis of *reaction delay* revealed an average delay of 259 ms from the timing in the ground truth to the robot's initiation of action for the cases when the robot reacted to the *passing* behavior (76.0% of the cases).

Overall, we observed that our system worked reasonably, as designed to reduce *unsure* failures, although the improvement remains unclear, and a considerable number of failures still happened (18%) at this step. We further analyzed the *unsure* failure cases and found that the system failed to react for 61% of them. This implies that we could further reduce the 18% failure ratio at this step by improving the classification performance. In the remaining 39% of *unsure* failures, people seemingly did not recognize the robot's reaction. This would be difficult to prevent with classification improvement, but it could be improved by considering a better way of expressing reactions from the robot. Perhaps, the robot's motions were too subtle.

E. Interaction Observations After Initiating Interaction

Apart from the approaching interactions, we note a couple of interesting observations. Since the robot's role was advertisement, it talked about shops in the mall to the visitors. Its content was relatively simple since the focus was on the approaching interaction.

In one successful approach, the robot approached a young couple and said, "There is a nice restaurant named Restraint Bay in this mall. You can see Osaka Bay from it. The view is beautiful!" The women said to the man, "He says the restaurant has a good view. How about visiting it?" The information was very timely and influenced their decision.

A similar interaction happened with a child who wanted some ice cream. In this situation, the robot said, "Today is really hot! If you want something cold, how about some ice cream? I know a good shop named The Soft Cream House." The child was excited by this information and asked his mother for ice cream. They were also influenced.

These examples show that a robot can influence people by providing information.

VII. Discussion

A. Summary

The field trial demonstrated a success rate of 55.9% for our approaching technique, which we consider to be reasonably high. The targets in this study were people who were going through a shopping mall. Many reasons might explain their reluctance to interact with the robot. Even walking slowly, they might be busy chatting with their friends. They might be preoccupied with a particular shop. We cannot expect very high acceptance from them. Note that with the proposed method, unreachable and unaware failures (see Section VI-C2) greatly decreased. Unsure failure seems to have decreased, yet it remains frequent. We believe that this is because this category includes indecisive cases where they slowed down to see the robot but were not very willing to interact with it. The robot did not aggressively initiate interaction, and since the application was advertisement, the robot must not irritate potential customers.

B. Applications

A couple of possible applications could be enabled by a robot with better approach capabilities. As demonstrated in this paper, providing advertisement information is one possible application. Moreover, this approach capability enables a robot to proactively serve people who are unaware of its presence or of its capabilities, e.g., a robot can provide route direction for a person who is lost. Since people sometimes hesitate to ask for help, a proactive way of serving is also helpful. In our study, people could nonverbally reject these services if they wanted; we believe that this functionality is also useful to politely provide such a proactive service. The proposed approach model, however, is not limited to information-providing tasks. It can also be applied to such functions as porter, shop salesperson, receptionist, and security guard.

We believe that further improvement is required before we apply our approaching technique to these applications. We need to adjust a strategy to select a target person depending on the application. For providing advertisement information, as this paper addressed, the strategy can be simple: The robot approaches a person who is not "busy." For other applications, strategies may be more complex. A route-guide robot needs to identify people who appear lost, which would require observations of a person's trajectory to find certain patterns. A porter robot might need to be able to observe whether a person is carrying a piece of luggage.

C. Parameters

We experimentally determined several parameters in our algorithm. One would ask how general they are, and how the parameters should be determined if they are not generalizable. Here, we discuss how we consider the influence of parameters and to what extent they are general or must be set up. Note that due to the nature of our study that was conducted in a real field environment, we did not run rigorous and systematic experiments and instead repeatedly conducted exploratory experiments.

1) *Planning frequency* (t_{period}): This was set to 500 ms due to the limitation of computation performance, while being realistically fast enough to respond to people's walking behavior. We expect that this would work for any other environment, but it would probably be useful to make it smaller.
2) t_{front}: This was set to 3.6 s. In our exploratory experiments, we searched for a reasonable parameter, starting from small values. The smaller values resulted in approaching from the side or causing *unreachable* failures. Until the value exceeded 1.5 s, it performed very poorly. The performance improved until the value we used, i.e., 3.6 s, and did not improve after that. We expect that the parameter (3.6 s) can be used in other environments as well, which would produce a reasonable amount of time for the robot to approach from the frontal direction. On the other hand, since our experimental area was relatively small, we expect that it could be useful to use a larger value for this parameter with robots in a larger environment. It would increase the time for which people can see the robot approaching from the frontal direction; therefore, it will provide more comfortable service, providing people enough time to consider whether to accept the service.
3) t_{th}: Since this parameter reflects prediction accuracy, it depends on the nature of the environment (e.g., whether people are likely to walk in similar patterns) as well as the prediction algorithm that is used. We set this to 40 s, which is simply a large value within which we consider the prediction to be somewhat reliable. Concretely speaking, in our environment, it takes an average of 19.8 s to pass through the corridor, and people rarely spend more than 40 s if they are simply passing through. Some people stayed at benches in the environment for more than 40 s, although predicting nonwalking people's behavior in 40 s of the future is extremely difficult. Overall, we made rough estimations that prediction over 40 s in the future is useless, and that the reliability of anticipation will simply decrease linearly as the lookahead time increases up to this limit of 40 s. We believe that a similar approximation would be sufficient in other environments. We could use a distribution of the prediction accuracy as a function of time, if available.
4) *α and β in the plan selection (see Section V-C2):* These parameters control to what extent the system considers the

previous history of the person's awareness of the robot over time, in contrast with the immediate utility. With smaller α (or larger β), the robot oscillates between approach targets, and with larger α (or smaller β), it tends to keep trying to approach a target person who is no longer likely to initiate interaction (particularly when the target is still approachable in terms of distance to travel, but it has started to turn in a different direction and is no longer facing the robot, perhaps to avoid interacting). In our environment, we chose a value that does not cause oscillation of switching the targets and chose $\alpha = 0.72$.

D. Prediction Algorithm

This study is based on the prediction algorithm reported in [27], which assumes that the behavior of currently observed people will resemble that of previously observed people. It predicts future trajectory from a couple of groups that resemble the current trajectory. While this provides a rough estimation of future position, which was sufficient for our purpose, there is a limitation. Since the algorithm does not consider interaction among people or other entities, the prediction is not necessarily accurate around the robot, due to possible influence from the robot. Because our environment was relatively small, this approach did not cause a problem; however, when we consider how to extend our system for larger environments, we will probably need further study for the prediction algorithm, since people have much more interaction with other people if they travel longer.

VIII. Conclusion

We reported the development of a technique that allows a robot to approach walking people, particularly, visitors in a shopping mall. We used the failures of a simple approaching method to guide the design of a better approaching behavior. Its main concept is to anticipate people's future trajectories and plan an approach path based on the anticipated trajectory of the targeted person. In the developed system, the anticipation method extended a previous method [27] with more samples (26 863 trajectories) and improved the computation of future behavior. Moreover, when the robot approaches close enough, it changes its working mode to provide quick responses to *unsure* reactions from the target.

The developed system was tested in a real shopping mall, and the results demonstrated its effectiveness. The success rate of the approaches significantly increased. The proposed system was successful in 33 out of 59 approaches, whereas the simplistic approach was only successful in 20 out of 57 approaches. Many different applications exist for this approach behavior, and they are not limited to simple advertisement services where a robot just recommends shops, but will be connected to other services for helping people with both physical services (e.g., transporting luggage) and information-providing services.

Acknowledgment

The authors would like to thank the administrative staff at Sumisho Urban Kaihatsu Co. Ltd. for their participation. They would also like to thank Dr. Akimoto, Dr. Miyashita, and Mr. Kurumizawa for their help.

References

[1] S. Satake, T. Kanda, D. F. Glas, M. Imai, H. Ishiguro, and N. Hagita, "How to approach humans? Strategies for social robots to initiate interaction," in *Proc. ACM/IEEE Int. Conf. Human-Robot Interact.*, 2009, pp. 109–116.

[2] W. Burgard, A. B. Cremers, D. Fox, D. Hahnel, G. Lakemeyer, D. Schulz, W. Steiner, and S. Thrun, "The interactive museum tour-guide robot," in *Proc. Nat. Conf. Artif. Intell.*, 1998, pp. 11–18.

[3] S. Thrun, M. Bennewitz, W. Burgard, A. B. Cremers, F. Dellaert, D. Fox, D. Hähnel, C. Rosenberg, N. Roy, J. Schulte, and D. Schulz, "Minerva: A second-generation museum tour-guide robot," in *Proc. IEEE Int. Conf. Robot. Autom.*, 1999, pp. 1999–2005.

[4] R. Siegwart, K. O. Arras, S. Bouabdallah, D. Burnier, G. Froidevaux, X. Greppin, B. Jensen, A. Lorotte, L. Mayor, M. Meisser, R. Philippsen, R. Piguet, G. Ramel, G. Terrien, and N. Tomatis, "Robox at expo.02: A large scale installation of personal robots," *Robot. Autom. Syst.*, vol. 42, no. 3, pp. 203–222, 2003.

[5] H.-M. Gross, H. Boehme, C. Schroeter, S. Mueller, A. Koenig, E. Einhorn, C. Martin, M. Merten, and A. Bley, "Toomas: Interactive shopping guide robots in everyday use—Final implementation and experiences from long-term field trials," *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, pp. 2005–2012, 2009.

[6] R. Gockley, J. Forlizzi, and R. Simmons, "Designing robots for long-term social interaction," in *Proc. Int. Conf. Intell. Robots Syst.*, 2006, pp. 1338–1343.

[7] T. Kanda, T. Hirano, D. Eaton, and H. Ishiguro, "Interactive robots as social partners and peer tutors for children: A field trial," *Human Comput. Interact.*, vol. 19, no. 1–2, pp. 61–84, 2004.

[8] J. Schulte, C. Rosenberg, and S. Thrun, "Spontaneous, short-term interaction with mobile robots," in *Proc. IEEE Int. Conf. Robot. Autom.*, 1999, pp. 658–663.

[9] K. Dautenhahn, M. L. Walters, S. Woods, K. L. Koay, C. L. Nehaniv, E. A. Sisbot, R. Alami, and T. Siméon, "How may I serve you? A robot companion approaching a seated person in a helping context," in *Proc. Int. Conf. Human-Robot Interact.*, 2006, pp. 172–179.

[10] K. Hayashi, D. Sakamoto, T. Kanda, M. Shiomi, S. Koizumi, H. Ishiguro, T. Ogasawara, and N. Hagita, "Humanoid robots as a passive-social medium—A field experiment at a train station," in *Proc. ACM/IEEE 2nd Annu. Conf. Human-Robot Interact.*, 2007, pp. 137–144.

[11] M. P. Michalowski, S. Sabanovic, and R. Simmons, "A spatial model of engagement for a social robot," in *Proc. IEEE Int. Workshop Adv. Motion Control*, 2006, pp. 762–767.

[12] N. Bergström, T. Kanda, T. Miyashita, H. Ishiguro, and N. Hagita, "Modeling of natural human-robot encounters," in *Proc. Int. Conf. Intell. Robots Syst.*, 2008, pp. 2623–2629.

[13] E. T. Hall, *The Hidden Dimension*. Garden City, NY: Doubleday, 1966.

[14] B. Jensen, X. Greppin, A. Lorotte, L. Mayor, M. Meisser, G. Ramel, and R. Siegwart, "The interactive autonomous mobile system robox," in *Proc. Int. Conf. Intell. Robots Syst.*, 2002, pp. 1221–1227.

[15] X. Shao, H. Zhao, K. Nakamura, K. Katabira, R. Shibasaki, and Y. Nakagawa, "Detection and tracking of multiple pedestrians by using laser range scanners," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2007, pp. 2174–2179.

[16] S. Kagami, S. Thompson, Y. Nishida, T. Enomoto, and T. Matsui, "Home robot service by ceiling ultrasonic locator and microphone array," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2006, pp. 3171–3176.

[17] D. F. Glas, T. Miyashita, H. Ishiguro, and N. Hagita, "Laser tracking of human body motion using adaptive shape modeling," in *Proc. Int. Conf. Intell. Robots Syst.*, 2007, pp. 602–608.

[18] A. Yamazaki, K. Yamazaki, Y. Kuno, M. Burdelski, M. Kawashima, and H. Kuzuoka, "Prior-to-request and request behaviors within elderly day care: Implications for developing service robots for use in multiparty settings," in *Proc. 10th Eur. Conf. Comput. Supported Cooperative Work*, 2007, pp. 61–78.

[19] J. Fogarty, S. E. Hudson, C. G. Atkeson, D. Avrahami, J. Forlizzi, S. Kiesler, J. C. Lee, and J. Yang, "Predicting human interruptibility with sensors," *ACM Trans. Comput.-Human Interact.*, vol. 12, no. 1, pp. 119–146, 2005.

[20] T. Tasaki, S. Matsumoto, H. Ohba, M. Toda, K. Komatani, T. Ogata, and H. G. Okuno, "Dynamic communication of humanoid robot with multiple people based on interaction distance," in *Proc. IEEE Int. Workshop Robot Human Interact. Commun.*, 2004, pp. 71–76.

[21] M. Finke, K. Lee Koay, and K. Dautenhahn, "Hey, I'm over here—How can a robot attract people's attention," in *Proc. IEEE Int. Workshop Robot Human Interact. Commun.*, 2005, pp. 7–12.

[22] E. A. Sisbot, R. Alami, T. Simeon, K. Dautenhahn, M. L. Walters, S. N. Woods, K. L. Koay, and C. L. Nehaniv, "Navigation in the presence

of humans," in *Proc. IEEE-RAS Int. Conf. Humanoid Robots*, 2005, pp. 181–188.

[23] E. Pacchierotti, H. I. Christensen, and P. Jensfelt, "Design of an office guide robot for social interaction studies," in *Proc. Int. Conf. Intell. Robots Syst.*, 2006, pp. 4965–4970.

[24] R. Gockley, J. Forlizzi, and R. Simmons, "Natural person-following behavior for social robots," in *Proc. Int. Conf. Human-Robot Interact.*, 2007, pp. 17–24.

[25] G. Hoffman and C. Breazeal, "Effects of anticipatory action on human-robot teamwork," in *Proc. Int. Conf. Human-Robot Interact.*, 2007, pp. 1–8.

[26] M. Bennewitz, W. Burgard, G. Cielniak, and S. Thrun, "Learning motion patterns of people for compliant robot motion," *Int. J. Robot. Res.*, vol. 24, no. 1, pp. 31–48, 2005.

[27] T. Kanda, D. F. Glas, M. Shiomi, and N. Hagita, "Abstracting people's trajectories for social robots to proactively approach customers," *IEEE Trans. Robot.*, vol. 25, no. 6, pp. 1382–1396, Dec. 2009.

[28] E. Goffman, *Behavior in Public Place: Notes on the Social Organization of Gatherings*. New York: Free, 1963, pp. 83–112.

[29] A. Kendon, "Features of the structural analysis of human communicational behavior," in *Aspects of Nonverbal Communication,* W. von Raffler Engel, Ed., Liss: Swets and Zeitinger, 1986, pp. 29–43.

[30] S. Li, B. Wrede, and G. Sagerer, "A dialog system for comparative user studies on robot verbal behavior," in *Proc. IEEE Int. Symp. Robot Human Interact. Commun.*, 2006, pp. 129–134.

[31] P. Kahn, N. Freier, T. Kanda, H. Ishiguro, J. Ruckert, R. Severson, and S. Kane, "Design patterns for sociality in human-robot interaction," in *Proc. ACM/IEEE Int. Conf. Human-Robot Interact.*, 2008, pp. 97–104.

[32] M. L. Walters, D. S. Syrdal, K. L. Koay, K. Dautenhahn, and R. te Boekhorst, "Human approach distances to a mechanical-looking robot with different robot voice styles," in *Proc. IEEE Int. Symp. Robot Human Interact. Commun.*, 2008, pp. 707–712.

[33] M. Yamamoto and T. Watanabe, "Timing control effects of utterance to communicative actions on embodied interaction with a robot," in *Proc. IEEE Int. Workshop Robot Human Interact. Commun.*, 2004, pp. 467–472.

[34] J. R. Movellan and J. S. Watson, "The development of gaze following as a Bayesian systems identification problem," in *Proc. Int. Conf. Dev. Learn.*, 2002, pp. 34–40.

[35] K. Gold and B. Scassellati, "Learning acceptable windows of contingency," *Connection Sci.*, vol. 18, pp. 217–228, 2006.

[36] J. Lee, J. F. Kiser, A. F. Bobick, and A. L. Thomaz, "Vision-based contingency detection," in *Proc. ACM/IEEE Int. Conf. Human-Robot Interact.*, 2011, pp. 297–304.

[37] D. F. Glas, T. Kanda, H. Ishiguro, and N. Hagita, "Simultaneous people tracking and localization for social robots using external laser range finders," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2009, pp. 846–853.

[38] C. Shi, M. Shimada, T. Kanda, H. Ishiguro, and N. Hagita, "Spatial formation model for initiating conversation," presented at the Proc. Robot.: Sci. Syst., Los Angeles, CA, 2011.

[39] H. Sakoe, "Dynamic programming algorithm optimization for spoken word recognition," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 26, no. 1, pp. 43–49, Feb. 1978.

[40] A. Turner and A. Penn, "Evolving direct perception models of human behavior in building systems?," in *Proc. Pedestrian Evacuat. Dynam.*, 2007, pp. 411–422.

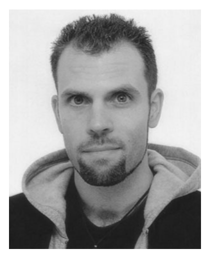
Satoru Satake received the B.Eng., M.Eng., and Ph.D. degrees from Keio University, Yokohama, Japan, in 2003, 2005, and 2009, respectively.

He is currently a Researcher with the Intelligent Robotics and Communication Laboratories, Advanced Telecommunications Research Institute International, Kyoto, Japan. His current research interests include human–robot interaction, networked robot systems, and social robots.

Takayuki Kanda (M'04) received the B.Eng., M.Eng., and Ph.D. degrees in computer science from Kyoto University, Kyoto, Japan, in 1998, 2000, and 2003, respectively.

From 2000 to 2003, he was an Intern Researcher with the Media Information Science Laboratories, Advanced Telecommunications Research Institute International (ATR), where he is currently a Senior Researcher with ATR Intelligent Robotics and Communication Laboratories, Kyoto. His current research interests include intelligent robotics, human–robot interaction, and vision-based mobile robots.

Dr. Kanda is a member of the Association for Computing Machinery, the Robotics Society of Japan, the Information Processing Society of Japan, and the Japanese Society for Artificial Intelligence.



Dylan F. Glas received the S.B. degrees in aerospace engineering and in earth, atmospheric, and planetary science in 1997 and the M.Eng. degree in aerospace engineering in 2000, all from the Massachusetts Institute of Technology (MIT), Cambridge.

Since 2005, he has been a Researcher with the Intelligent Robotics and Communication Laboratories, Advanced Telecommunications Research Institute International, Kyoto, Japan. From 1998 to 2000, he was with the Tangible Media Group, MIT Media Lab. His research interests include networked robot systems, teleoperation for social robots, human–machine interaction, ubiquitous sensing, and artificial intelligence.

Michita Imai (M'04) received the Ph.D. degree in computer science from Keio University, Tokyo, Japan, in 2002.

In 1994, he joined Human Interface Laboratories, Nippon Telegraph and Telephone Corporation. In 1997, he joined Media Integration and Communications Research Laboratories, Advanced Telecommunications Research Institute International (ATR), Kyoto, Japan. He is currently an Assistant Professor with the Faculty of Science and Technology, Keio University, and a Researcher with the Intelligence Robotics and Communications Laboratories, ATR. His research interests include autonomous robots, human–robot interaction, speech dialog systems, humanoids, and spontaneous behaviors.

Dr. Imai is a member of the Information and Communication Engineers Japan, the Information Processing Society of Japan, the Japanese Cognitive Science Society, the Japanese Society for Artificial Intelligence, the Human Interface Society, and the Association of Computing Machinery.

Hiroshi Ishiguro (M'01) received the D.Eng. degree from Osaka University, Osaka, Japan, in 1991.

In 1991, he was a Research Assistant with the Department of Electrical Engineering and Computer Science, Yamanashi University, Yamanashi, Japan. In 1992, he joined the Department of Systems Engineering, Osaka University, as a Research Assistant. In 1994, he became an Associate Professor with the Department of Information Science, Kyoto University, Kyoto, Japan, where he was involved in research on distributed vision using omnidirectional cameras. From 1998 to 1999, he was a Visiting Scholar with the Department of Electrical and Computer Engineering, University of California, San Diego. Since 1999, he has been a Visiting Researcher with the Media Information Science Laboratories, Advanced Telecommunications Research Institute International (ATR), Kyoto, where he developed Robovie: an interactive humanoid robot. In 2000, he joined the Department of Computer and Communication Sciences, Wakayama University, Wakayama, Japan, as an Associate Professor where he became a Professor in 2001. He is currently a Professor with the Department of Adaptive Machine Systems, Osaka University, and a Group Leader with the Intelligent Robotics and Communication Laboratories, ATR.

Norihiro Hagita (M'85–SM'99) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Keio University, Tsuruoka City, Japan, in 1976, 1978, and 1986, respectively.

From 1978 to 2001, he was with the Nippon Telegraph and Telephone Corporation. In 2001, he joined the Advanced Telecommunications Research Institute International (ATR), Kyoto, Japan, to establish the ATR Media Information Science Laboratories and the ATR Intelligent Robotics and Communication Laboratories in 2002. His current research interests include communication robots, network robot systems, interaction media, and pattern recognition.

Dr. Hagita is a fellow of the Institute of Electronics, Information, and Communication Engineers, Japan. He is a member of the Robotics Society of Japan, the Information Processing Society of Japan, and the Japanese Society for Artificial Intelligence. He is the Co-Chair for the IEEE Technical Committee on Networked Robots.