Who will be the customer?: A social robot that anticipates people's behavior from their trajectories

Takayuki Kanda, Dylan F. Glas, Masahiro Shiomi, Hiroshi Ishiguro, Norihiro Hagita
ATR Intelligent Robotics and Communication Laboratories
2-2-2 Hikaridai, Seika-cho, Souraku-gun, Kyoto, Japan
{kanda, dylan, m-shiomi, ishiguro, hagita}@atr.jp

ABSTRACT

For a robot providing services to people in a public space such as a train station or a shopping mall, it is important to distinguish potential customers, such as window-shoppers, from other people, such as busy commuters. In this paper, we present a series of techniques for anticipating people's behavior in a public space, mainly based on the analysis of accumulated trajectories, and we demonstrate the use of these techniques in a social robot. We placed a ubiquitous sensor network consisting of six laser range finders in a shopping arcade. The system tracks people's positions as well as their local behaviors such as fast walking, idle walking, or stopping. We accumulated people's trajectories for a week, applying a clustering technique to the accumulated trajectories to extract information about the use of space and people's typical global behaviors. This information enables the robot to target its services to people who are walking idly or stopping. The robot anticipates both the areas in which people are likely to perform these behaviors, and also the probable local behaviors of individuals a few seconds in the future. In a field experiment we demonstrate that this system enables the robot to serve people efficiently.

Author Keywords

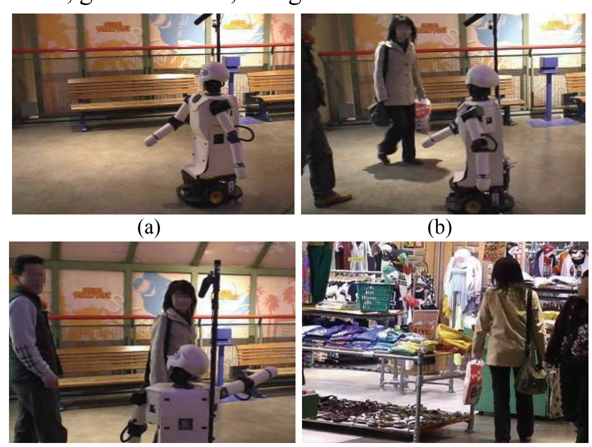
Behavior anticipation, Human-Robot Interaction

ACM Classification Keywords

H.5.2 **[Information Interfaces and Presentation]:** User Interfaces – *User-centered design, Interaction styles;*

INTRODUCTION

We believe that the robot can be a powerful device for bridging the gap between the digital and physical worlds. Since robots are mobile and embodied, they are well-suited for presenting digital information in the physical world. Humanoid robots are particularly effective in this respect, as their human-like bodies facilitate an intuitive style of communication through speech and gesture. If a humanoid robot uses body language effectively, people will communicate naturally with it, unconsciously behaving as if they were communicating with their peers. This could allow robots to perform communicative tasks in society such as presenting explanations about exhibitions or products. Previous studies in robotics have emphasized the merits of robot embodiment, showing the effectiveness of facial expression, gaze direction, and gestures.



(a) (b) (c) (d)

Figure 1. A robot successfully inviting a person to a shop

On the other hand, robots have only weak sensing capabilities. Since we aim to realize a robot that provides services in public spaces (as in Figure 1), it needs to observe the positions and motion of people. However, a robot using onboard sensors can usually recognize people only within a few meters, and its sensing is not robust. To overcome these limitations, we use a "network robot system" approach, in which a robot is supported by a ubiquitous sensor network which observes and interprets information about people. Such an approach combines the stability and wide-area sensing capability of a ubiquitous sensor network with the intuitive presentation capabilities of the robot.

This paper describes a technique for integrating a mobile robot and a ubiquitous sensor network in order to efficiently provide services in a shopping arcade. In particular, since timing is highly critical for social interactions, we focus on the problem of anticipating the motion and behavior of customers, to determine where the robot should move and which customers the robot should approach.

For example, if a robot is designed to invite customers to a shop, it should approach people who are walking slowly and possibly window-shopping. To approach those customers, we first anticipate areas where a behavior (*e.g.* walking slowly) is often observed. We then use a technique based on global behavior estimation to "pre-approach" customers who are most likely to exhibit the chosen behavior. Only when they begin actually performing that behavior does the robot commit and approach them.

RELATED WORKS

This study considers three essential types of trajectory-related information: local behavior, the use of space, and global behavior. We define the term **local behavior** to refer to basic human motion primitives, such as walking, running, going straight, and so on. The observation of these local behaviors can then reveal information about the **use of space**, that is, how people's behavior differs in different areas of the environment. Finally, for more insight into the structure of people's behaviors, we look at **global behavior,** that is, overall trajectory patterns comprised of several local behaviors in sequence, such as "entering through the north door, walking across the room, and sitting at the desk." Global behaviors are highly dependent on environments.

The detection of local behaviors and analysis of the use of space can be valuable in anticipating where behaviors are statistically likely to occur; however, an analysis of global behavior is far more powerful for predicting *individual* behavior. As people using the space have a variety of goals, an understanding of global behavior is essential in enabling the robot to anticipate the future behaviors of individuals.

Position and Local Behaviors

People's positions and trajectories have frequently been studied in robotics and computer vision (for example, [11, 25]). In ubiquitous computing, positioning devices are often used, such as GPS, or the signal strength of radio (GSM, WiFi, Bluetooth, RFID, and power line) [15, 18, 20, 21].

Ubiquitous computing technology is increasingly being used to identify people's local behavior as well. For example, Eagle and Pentland developed a Bluetooth-based device attached to a mobile phone that enables the analysis of activities such as being at home, at the office, or elsewhere [7]. Liao *et al.* also used locations obtained via GPS with a relational Markov model to discriminate location-based activities such as being at home, at the office, and out dining [16]. Subramanya *et al.* included motion states (such as stop, walk, run) and velocity into a model to estimate people's low-level activity and spatial context [24].

These techniques all used wearable or mobile personal devices. Our focus is on applications in an anonymous public space, so we chose a method independent of such devices. We measure walking motion using laser range finders, sensors often used in robotics due to their precision, simplicity, and non-invasiveness. A number of techniques exist for tracking people using multiple laser range finders [8, 22].

The Use of Space

Humans' spatial behavior has attracted scientific interest for a long time (*e.g.*, studies of space syntax [4], and how humans establish a trail [9]). These studies required labor-intensive effort to observe people's trajectories; however, recent sensing technologies enable us to automatically accumulate trajectories and analyze people's behavior in detail. Previous studies in ubiquitous computing revealed that trajectories enable the identification of pausing points [25], traffic paths [22, 25], and frequency of stay [12].

Information on the general use of space has also been retrieved. Nurmi *et al.* applied a spectral clustering method for identifying meaningful places [19]. Aipperspach *et al.* applied clustering to UWB sensor data to identify typical places in the home [1]. Koile *et al.* conducted a clustering of spaces with a focus on the relationships between velocity and positions, which enabled a partitioning of space into "activity zones." For example, places for walking, working, and resting were separated [13]. Our work involves partitioning space in a similar manner, but based on position and local behavior. In addition, we also consider how the distribution of these zones varies as a function of time.

Global Behavior

Models of human walking have been developed for transportation engineering and architectural design. These models are usually concerned with how environmental information affects people's behavior, such as a line of sight toward environmental structures [26] and movement of individuals in a crowd [2]. Ubiquitous computing could contribute to these models by providing automated, accurate positioning.

In previous studies, ubiquitous computing technologies have been used for categorizing people, and estimating people's goals and intentions [5]. In a museum context, Sparacino developed the "museum wearable," where people were classified into three visiting patterns. Depending upon the pattern, the system adjusted the way it presented information [23]. This is a good example of the use of global behavior; however, the places and the model of global behaviors were carefully prepared by a human designer.

In contrast, we have applied a clustering technique to identify typical visiting patterns in a museum without providing any environmental information [12]. One of the novel points of our work is that the designer of the system provides information only about the *target local behavior*, with no knowledge about the structure of the space or of people's global behaviors. In addition to the previous work, this paper provides a method of online estimation of global behavior, which is indispensable for providing services.

The online estimation of global behaviors is difficult as, by definition, any global behavior being observed in real time is unfinished and thus not completely observable. Thus, it is necessary to estimate the true global behavior from a limited data set. Krumm *et al.* developed a technique they call "Predestination", which enables someone's driving destination to be estimated [14]. Liao *et al.* developed a technique for a person wearing GPS to infer her destination, transportation mode, and anomalous behavior [17].

While personal history of previous destinations was an important part of those studies, our anticipation technique for the shopping arcade assumes zero knowledge of a given person's individual history. Our technique is predicated on our observations of tens of thousands of people and the expectation that a new person's global behavior will be similar to those previously observed.

The concept of behavior anticipation is not without precedent in robotics. For example, Hoffman *et al.* demonstrated the value of anticipatory action in human-robot collaboration [10]. However, our use of global behaviors is a unique approach to behavior anticipation in this field.

Human-Robot Interaction

In the field of human-robot interaction, there have been many studies about mobile robots that provide services to people. For example, Dautenhahn *et al.* found that the robot should approach people from the side rather than the front [6]. Bennewitz *et al.* developed a technique for predicting trajectories of persons for avoiding persons around it [3]. However, none of these studies addressed the question of whom the robot should serve.

RECOGNITION SYSTEM

Figure 2 shows the overview of the system. The details will be explained in the following sections.

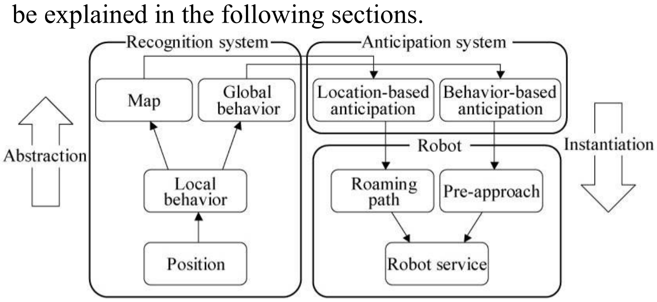


Figure 2. Overview of the system

Position

We conducted our experiments in a popular entertainment and shopping arcade located by the entrance to Universal Studios Japan, a major theme park. We operated the robot within a 20 m section of the arcade, with shops selling clothing and accessories on one side and an open balcony on the other. The motion of people through this area was monitored using a ubiquitous sensor network consisting of six SICK LMS-200 laser range finders[1] mounted around the perimeter of the trial area at a height of 85 cm (Figure 3).

A particle filtering technique was used to track people's trajectories through this space. The location of each person in the scan area was calculated based on the combined torso-level scan data from the laser range finders.

[1]In our vision, laser range finders could eventually be distributed everywhere as embedded elements of the environment, functioning as a ubiquitous sensor network infrastructure, particularly in commercial spaces.

Figure 3. The shopping arcade and laser range finders.

In our tracking algorithm, a background model is first computed for each sensor, by analyzing hundreds of scan frames to filter out noise and moving objects. Points detected in front of this background scan are grouped into segments, and segments within a certain size range persisting over several scans are registered as human detections.

Each person is then tracked with a particle filter, using a linear motion model with random perturbations. Likelihood is evaluated based on the potential occupancy of each particle's position (i.e. humans cannot occupy spaces which have been observed to be empty). By computing a weighted average across all of the particles, x-y position is calculated at a frequency of approximately 37 Hz. This tracking technique provides quite stable and reliable position data, with a position accuracy measured to be +/- 6 cm for our environment. Further details on this algorithm are presented in [8].

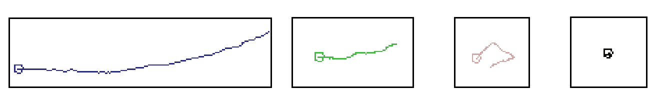
Local Behavior

As defined earlier, "local behaviors" represent basic human motion primitives. We began our analysis with the classification system, which uses SVM (support vector machine) to categorize trajectories based on their velocity, direction, and shape features.

To include a wide variety of movement types, we initially defined a set of 20 local behavior classes, based on walking style and walking speed, considering both 5-second trajectory segments and short-duration (2-second) segments. Each class has about 200 samples for learning, consisting of 2- or 5-second trajectory segments. We manually labeled these samples. The classification method averaged 89.6% accuracy for category estimation. The following are examples of the categories used:

Style: Classes describing style of motion: *straight*, *right-turn*, *left-turn*, *wandering*, *U-turn*, and *stop*.

Speed: Classes describing overall speed: *run*, *fast-walk*, *idle-walk*, *stop*, and *wait*.



(a) Fast-walk (b) Idle-walk (c)Wandering (d) Stop

Figure 4. Example trajectories for local behaviors

In the subsequent analysis, we merged several local behavior classes for simplicity. Within "Style", the classes *left-turn*, *right-turn*, and *U-turn* were all merged into the *wandering* category. Within "Speed", we merged *stop* and *wait* into the *stop* category. We also merged classes for short-duration and 5-second behavior. Thus, we reduced the set to the following four local behaviors: *fast-walk*, *idle-walk*, *wandering*, and *stop*. Figure 4 shows examples of these local behaviors. We define the position P_t^n of visitor *n* at time t to include the x-y coordinates (x, y) as well as Boolean variables indicating the presence or absence of local behavioral primitives $P_{fast-walk}$, $P_{idle-walk}$, $P_{wandering}$, P_{stop}.

ANALYSIS OF ACCUMULATED TRAJECTORIES

Based on the position and local behavior data thus obtained, an analysis was performed to obtain a higher-level understanding of the use of space and people's global behaviors. This analysis constitutes the foundation for the robot's ability to anticipate people's local behaviors.

Data Collection

Human motion data was collected for a week in the shopping-arcade environment, from 11am-7pm each day, including 5 weekdays and 2 weekend days. We chose this time schedule because the shops open at 11am, and the number of visitors drops after 7pm, after the theme park closes in the evening.

In this environment, the major flow consisted of customers crossing the space from the left to the upper right or vice versa, generally taking about 20 seconds to go through. We removed trajectories shorter than 20 seconds, in order to avoid noise from false detections in the position tracking system. In all, we gathered 11,063 visitor trajectories.[2]

Use of Space (Map)

The first analysis task was to identify how the space was used, and how the use of space changed over time. We applied the ISODATA clustering method to achieve this. First, we partitioned the time into one-hour segments categorized as weekday or weekend. We then partitioned the space into a 25cm grid, with each grid element containing histogram data of local behaviors: $H_{fast-walk}(i,t)$, $H_{idle-walk}(i,t)$, $H_{wandering}(i,t)$, and $H_{stop}(i,t)$, where $H_x(i,t)$ denotes the number of occurrences of local behavior *x* at time slice *t* within grid element *i,* which is normalized for each local behavior *x*.

To make the data set more manageable, we first combined time slices based on their similarity. The difference between time slices t_1 and t_2 is defined as:

$$\sum_i \sum_x | H_x(i,t_1) - H_x(i,t_2) |$$

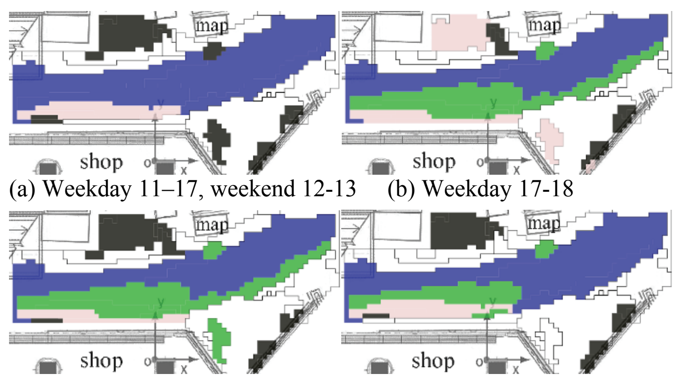
We then combined spatial grid cells where the distance was smallest and the grid was spatially connected. The distance between grid cells *i* and *j* is defined as:

$$\sum_t \sum_x | H_x(i,t) - H_x(j,t) |$$

As is usual for this type of explorative clustering, we arbitrarily set the number of partitions to intuitively understand the phenomena occurring in the environment. We set the number of spatial partitions to be 40 and temporal partitions to be 4. Figure 5 shows a visualized output of the analysis. The partitions are color-coded according to the dominant local behavioral primitive in each area. Blue (medium gray on monochrome printouts) represents the areas where the *fast-walk* behavior occurred more frequently than any other local behaviors. Thus, people tend to pass directly through this area, which can be thought of as "corridor" space.

The areas where the *idle-walk* primitive occurred most frequently are colored with green (or light gray).

In some areas, the use of space was very clearly observed to change as a function of time. The lower left area is in front of a shop. When the shopping arcade was busy in the evening, as in Figure 5 (b), with people coming back from the theme park, many people were observed to slow down in front of the shop, and the "corridor" space changed into "in front of shop" space with *idle-walk* becoming dominant; however, when there were not so many people, such as midday during the week as in Figure 5 (a), these areas disappeared and became similar to other "corridor" space. The lower right side of the map represents the side of the corridor, where people tend to walk slowly when the arcade is busy (Figure 5 (b) and (c)).



(a) Weekday 11–17, weekend 12-13 (b) Weekday 17-18

(c) Weekday 18-19 (d) Weekend 11-12, 13-19

Figure 5. Analysis of the use of space

(a) a bench (b) rest space (c) a map

Figure 6. Examples of the actual use of the space

The areas where the *stop* primitive was most frequent are colored with dark brown (or dark gray). In Figure 5, these areas can mainly be found in the upper center (photo: Figure 6 (a)) and the bottom right (photo: Figure 6 (b)). These areas contain benches, and can be considered "rest space".

---

[2] In this study, we obtained approval from 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.

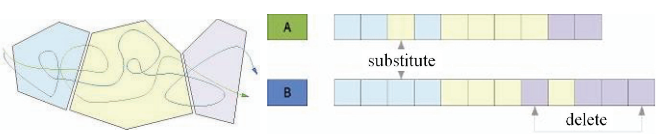
In the upper center area, below the word ‘map’, there is a small space where *stop* is the dominant primitive in Figure 5 (a) whereas *idle-walk* is dominant in (b) through (d). A map of the shopping arcade is placed on that wall. Customers sometimes slowed down, stopped, and looked at this map (Figure 6 (c)). The statistical analysis clearly revealed this phenomenon as defining a distinct behavioral space.

The areas where the *wandering* primitive was dominant are colored with pink (or very light gray). All maps in Figure 5 show the space immediately in front of the shop as having this property. The areas where none of the primitives were dominant, such as the bottom-right space, are colored white. These areas were not used so much.

To summarize, we have demonstrated that through this analysis technique, we can separate space into semantically meaningful areas such as the corridor, the space in front of the shop, the area in front of the map, and the rest space. It also reveals how usage patterns change over time, such as the change of dynamics in the space in front of the shop.

Global Behavior

To identify typical global behaviors of shopping arcade visitors, we classified trajectories with a k-means method [12]. For the clustering, the distance between two trajectories is computed by using a DP matching method. Figure 7 shows how the comparison of trajectories works.



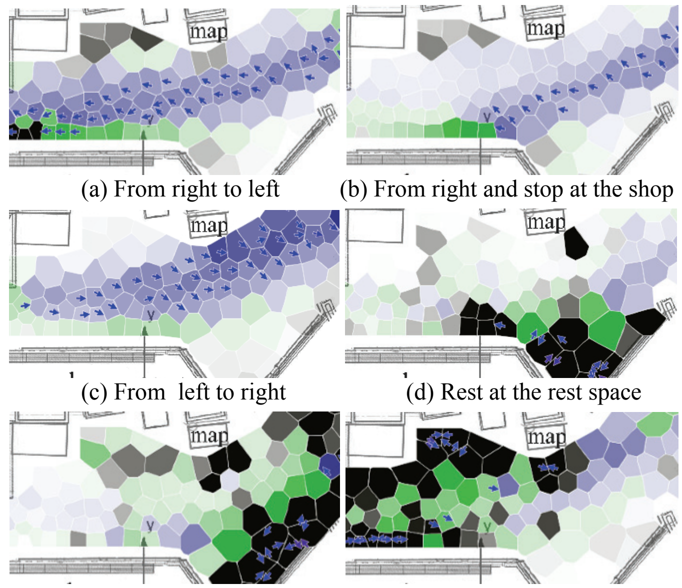
(a) two trajectories (b) comparison of trajectories

Figure 7. Comparison of trajectories based on DP matching

For the DP matching, we again partitioned the space into a 25cm grid, to easily compare trajectories. As a result, the environment was separated into 2360 grid elements. The DP matching method was chosen for its simplicity and the fact that it does not require particular tuning of parameters. Since global behaviors naturally emerge through the interactions between people and their environment, we believe that it is best to minimize the number of parameters that need to be adjusted manually, keeping the process simple and generalizable.

The trajectories are segmented into 500 ms time steps, and they are compared with each other based on the physical distance between them at each time step. To this is added a cost function, based on “insert” and “delete” operation costs in the DP matching, where we defined the cost of a single insertion or deletion to be 1.0 m.

Figure 8 shows a visualization of the global behaviors at *k*=6. We separated the space into 50 similarly-sized partitions by the k-means method for this visualization, although the actual computation used 2360 partitions. In the figure, each area is colored according to its dominant local behavior primitive. For example, blue represents *fast-walk*, and green represents *idle-walk*. Solid colors indicate a frequency of occurrence of at least one standard deviation above average, and lighter tints represent weaker dominance, down to white if the frequency is more than a standard deviation below average. Frequent transitions between adjacent areas are shown by arrows.



(a) From right to left (b) From right and stop at the shop

(c) From left to right (d) Rest at the rest space

(e) Around the rest space and right (f) Around the shop and bench

Figure 8. Six typical patterns of global behavior

The following six typical global behaviors were retrieved:

(a) Pass through from right to left (7768 people)
This pattern represents one of the major flows of people, who are coming back from the theme park (on the right) on their way to the train station (on the left). In this pattern, most of the areas are colored blue because the most frequent primitive in those areas was *fast-walk*. In front of the shop, there are some areas colored green, which represent spaces where people slow down to look at the shop.

(b) Come from the right, and stop at the shop (6104 people)
In this pattern, people come from the right side and enter the shop, as these trajectories mostly disappear at the shop.

(c) Pass through from left to right (7123 people)
This is also a major pattern, where people are coming from the train station and going in the direction of the theme park. In contrast to the patterns in (a) and (b), people rarely stopped or slowed down in front of the shop.

(d) Rest at the rest space (213 people)
In this pattern, people mostly spent time in the bottom right rest space (Figure 6 (b)) where benches were placed.

(e) Around the rest space and right (275 people)
Similar to the pattern in (d), but people moved around the right area more, and not around the shop area.

(f) Around the shop and bench (334 people)
People mainly came from the left side, walking slowly, and stopped in front of the shop as well as in front of the map.

In summary, this analysis technique has enabled us to extract typical global behavior patterns. These results show that most people simply pass through this space while a smaller number of people stop around the rest space or the map area. People tend to stop at the shop more often when they come from the right, a result which makes intuitive sense, as the shopping arcade is designed mainly to attract people coming back from the theme park.

ANTICIPATION SYSTEM

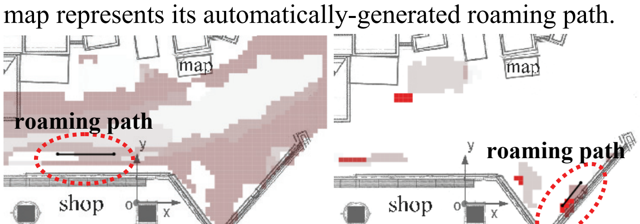
Robots differ from other ubiquitous computing systems in that they are mobile, and it takes some time for a robot to reach a person in need of its service. Thus, the ability to anticipate people's actions is important, as it enables the robot to pre-position itself so it can provide service in a timely manner.

We assume here that the robot's service is targeted towards people who are performing some particular local behavior, such as *stop* or *idle-walk*. The robot system uses the results of the analysis about the use of space and global behavioral primitives to anticipate the occurrence of this "target behavior". At the same time, the robot system tries to avoid people who are performing particular local behaviors, such as *fast-walk*, which we refer to as "non-target behavior". To anticipate local behaviors, we use two mechanisms: location-based anticipation and behavior-based anticipation.

Location-Based Anticipation

As shown in Figure 5, the system has use-of-space information about the frequency of the local behaviors associated with spatial and temporal partitions. The robot uses this information to estimate the locations in which people will be statistically likely to perform the target behavior.

Figure 9 shows an example anticipation map. The red areas represent areas where the system anticipates both a high likelihood of the target behavior and a low likelihood of the non-target behavior. The robot roams through this high-likelihood area looking for people. The black line on the map represents its automatically-generated roaming path.



(a) Weekday11-17, *idle-walk* (b) Weekday 11-17, *stop*

Figure 9. Example of anticipation map

In one scenario, the robot's task might be to invite people to visit a particular shop. In this case, selecting *idle-walk* as the target behavior and *fast-walk* as the non-target behavior might be appropriate, since the robot wants to attract people who have time and would be likely to visit the store. Figure 9 (a) is the anticipation map for this scenario, calculated for the behavior patterns observed on weekdays, between 11am and 5pm. Several areas away from the center of the corridor are colored, and the roaming path is set in front of the shop.

In a different scenario, the robot's task might be to entertain idle visitors who are taking a break or waiting for friends. Particularly because this shopping arcade was situated near a theme park, this is quite a reasonable expectation. In this case, it would be more appropriate to select *stop* as the target behavior and *fast-walk* as the non-target behavior. Figure 9 (b) is the anticipation map for this second scenario. In this case, only a few areas are colored. The roaming path is set to the bottom-right area.

Note that the roaming path was automatically calculated based on the anticipation map. No additional knowledge about the space was provided by designers.

Behavior-Based Anticipation

The second technique used for anticipating local behaviors is to estimate the global behaviors of people currently being observed, and then to use that information to predict their expected local behaviors a few seconds in the future.

To ensure prediction accuracy, we used a large number of clusters for the global behavior analysis. We clustered the human motion data collected earlier into 300 global behavior patterns. Next, to predict the global behavior of a new trajectory which has been observed for T seconds, the system compares the new trajectory with the first T seconds of the center trajectory of each of the 300 clusters, using the same DP matching technique applied earlier for deriving the global behaviors. The cluster with the minimum distance from the new trajectory is considered to be the best-fit global behavior for that trajectory.

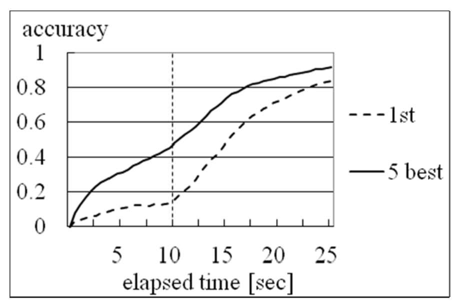


Figure 10. Accuracy of the prediction of global behavior

Figure 10 shows the prediction accuracy for observed trajectories from 0 to 25 seconds in length. Here, we used 6 of the 7 days of data to create the prediction model, and tested its ability to predict the remaining one day of the accumulated data. The accuracy accounts for only trajectories of total length greater than 20 seconds, as we filtered out shorter trajectories for calculating global behaviors. The result labeled "1st" represents the case where the best-fit global behavior at time T was the correct one (the cluster the trajectory finally fit with at completion). The result labeled "5 best" is the result if we define success to mean that correct global behavior falls within the top 5 results. Performance levels off after 20 seconds. Since there are 300 global behaviors, we believe that a success rate after 10 seconds of 45% and after 15 seconds of 71% for "5 best" represents fairly good performance.

After the most likely global behaviors are selected, the person's future position and local behavior are predicted based on an "expectation map." An expectation map is a data structure prepared *a priori* for each global behavior. For each 500-ms time step along the trajectories, a 25-cm grid representation of the observed space is added to the map. Each element of this grid contains likelihood values for each of the four local behaviors to occur in that location at any time *after* that time step. These likelihood values are empirically derived from the original observed trajectories falling within the chosen global behavior cluster, and they represent the average frequency of the occurrence of each local behavior after that time step. We used the 5-Best result to create the expectation map for the person by accumulating each expectation map of 5-Best global behaviors.

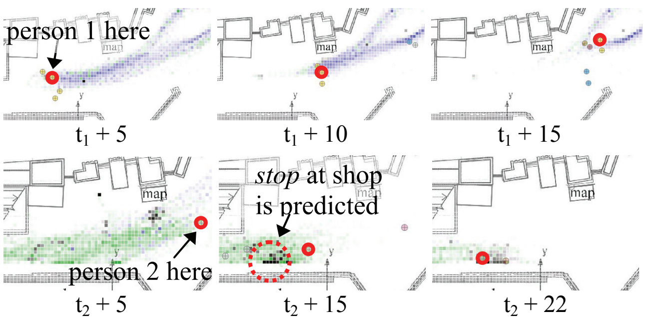


Figure 11. Example of prediction of future behaviors

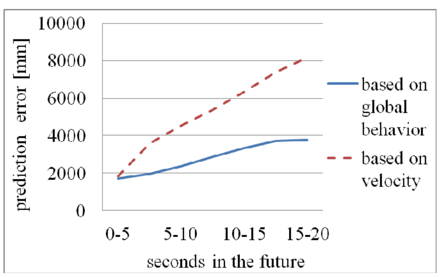


Figure 12. Prediction accuracy for position

Figure 11 shows expectation maps for various time increments. The solid circles represent the positions of people walking through the space, with the person of interest outlined in red. The expectation map for that person's estimated global behavior is shown, where the area colored blue represents the area where *fast-walk* is expected, and the green area represents the area where *idle-walk* is expected. The three figures in the top row show the trajectory for person 1, who was first observed at time t_1. The first figure shows time t_1 + 5 sec, where the expected local behaviors can be seen tracing a path through the corridor, heading toward the upper right. In fact, this course was correctly predicted, and the person followed that general path. The second line is the trajectory for person 2, first observed at time t_2. Here, since the person walked slowly, it predicted the course to the left with *idle-walk* behavior. At time t_2+15, it started to predict the possibility of *stop* at the shop, which finally came to be true at time t_2+22.

We measured the accuracy of position prediction for four time windows: 0-5, 5-10, 10-15, and 15-20 seconds in the future. Predictions were begun after a trajectory had been observed for 10 seconds, as the estimation of global behavior is not stable until then. We again used 6 days of data from the accumulated trajectories to predict the data of the remaining day. Our method predicts the future position as the center-of-mass of the expectation map. Figure 12 compares our method with position prediction based on the velocity over the last second. As the velocity method cannot account for motions like following the shape of the corridor, our method performs about twice as accurately.

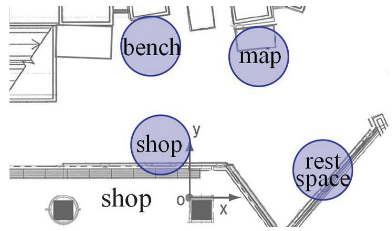
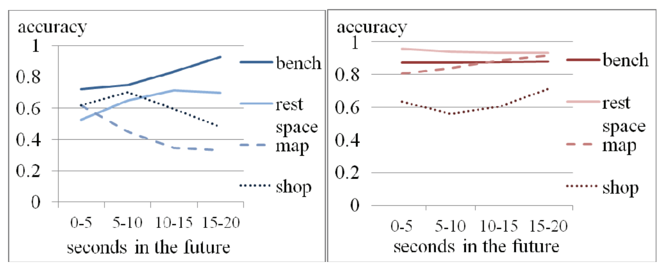


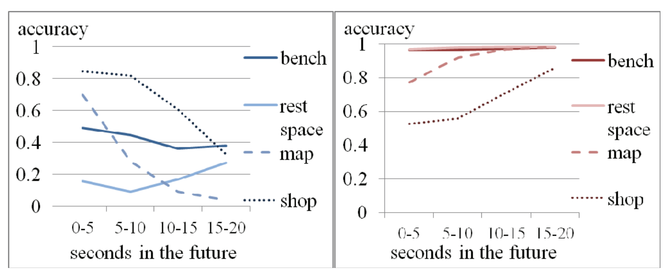
Figure 13. Places used for the measuring the performance

We then measured the correctness of the system's predictions of the future positions and local behaviors for each person, evaluated in four places (indicated by three-meter circles in Figure 13) where qualitatively distinct behaviors were observed in the use-of-space analysis. For each place, at each moment, the system predicted whether the person would exhibit each of the local behaviors at that place for forecast windows of 0-5, 5-10, 10-15, and 15-20 seconds.



Behavior occurred at the place Behavior did not occur at the place

Figure 14. Prediction accuracy for *stop* behavior



Behavior occurred at the place Behavior did not occur at the place

Figure 15. Prediction accuracy for *idle-walk* behavior

Figures 14 and 15 show the system's prediction performance. In each figure, the left graph shows the accuracy of the prediction for the case where the target local behavior occurred at each place, and the right graph show the accuracy of the prediction where the behavior did not occur. We define the occurrence of the local behavior as the case where the person appeared at the place in the predicted 5-second window (*e.g.* between 5 sec and 10 sec), and performed the target local behavior more than other local behaviors. The accuracy value used for each person is the average across all predictions made for that person, and the value shown in the graph is the average across all people.

Figure 14 shows that the prediction was fairly accurate for the *stop* behavior, particularly at the bench and the rest space. Prediction was 92% accurate at the bench even for 15-20 seconds in the future, while non-occurrence was predicted with 88% accuracy. This good performance was due to the fact that people who stay in these areas often stay for a long time. Results were more marginal at the map and shop, with 62% accuracy for occurrence and 63% for non-occurrence predicted at the shop for 0-5 seconds in the future. For 15-20 seconds in the future, the performance is still marginal, with 48% accuracy for occurrence and 71% for non-occurrence predicted at the shop.

In contrast, as Figure 15 shows, the system predicted *idle-walk* with high accuracy 0-5 seconds ahead at the map and the shop. Even for 15-20 seconds ahead, the system was able to predict 33% of the occurrences at the shop as well as 86% of the non-occurrences, which we consider to be a good result, as it is rather difficult to predict walking behavior in the future. The prediction of occurrence was not successful at the rest space, as the system mostly predicted non-occurrence, since *idle-walk* rarely happened there.

Regarding the remaining two behaviors, for *wandering*, the system predicted over 50% of occurrences and 85% of non-occurrences for 0-5 seconds ahead at all four places. For the 15-20 second window, it predicted 73% of occurrences and 93% of non-occurrences at the bench but not so well for the map and shop. It predicted *fast-walk* at map and shop well until 10 seconds; for example, it predicted 86% of occurrences and 60% of non-occurrences at the shop for 5-10 seconds in the future, though it does not predict the future well beyond 10 seconds.

We believe these anticipation results are useful for the robot. The robot is designed to wait for people in areas where it anticipates frequent occurrence of the target behavior. Behavior-based anticipation performs particularly well in areas where the anticipated behaviors occur often, such as *stop* near the benches and rest space, and *idle-walk* in the corridor in front of the map and shop. As these are the areas predicted by the location-based anticipation method, the two anticipation techniques complement each other nicely.

SERVICE FROM A SOCIAL ROBOT

In this section, we show examples where a social robot provides services using our system. A human designer defines the contents of the service as well as the context in which the robot should provide the service. Here, the notable point is that the designer only specifies the target local behavior, such as “stopping”. The robot system then automatically computes the information about space and global behavior so that the robot can efficiently wait for people in promising areas, and then proactively approach people who are anticipated to do the target local behavior.

For these services a robot has an advantage over mobile devices, in that people do not need to carry any hardware; however, there is the additional challenge that robots need to approach the person quickly enough to start the service. For this purpose, anticipation plays an important role.

Robot Hardware

We used an interactive humanoid robot (shown in Figure 1), characterized by its human-like physical expressions. It is 120 cm high and 40 cm in diameter. It is equipped with basic computation resources, and it communicates with the ubiquitous sensor network via wireless LAN.

Entertainment Application

The first example of an application that we would like to discuss is an entertainment robot, which interacts with people in the form of chatting. As mentioned earlier, the shopping arcade is next to an amusement park, so it is a reasonable for the robot to be entertaining people who have free time. In addition, we think that such an entertainment service would be reasonable for a robot in other environments as well, as robots today are still an exciting novelty.

The chat was about the attractions in the amusement park. For example, the robot says, “Hi, I'm Robovie. Yesterday, I saw the Terminator at Universal Studios. What a strong robot! I want to be cool like the Terminator. 'I’ll be back...' ”. We set the target local behavior as *stop*, and non-target as *fast-walk*, in order to serve people who are idle.

We conducted a field trial to investigate the effectiveness of the system. Based on the anticipation mechanism and its current position, the robot set its roaming path near the bench and waited for a person to approach. When the robot predicted that a detected person would probably do the *stop* behavior, the robot began positioning itself near her general area (pre-approach). When she came in front of the shop, she stopped (partly, we assume, because she was intending to stop regardless of the robot, and partly because she noticed the robot approaching her). Once she stopped, the robot approached her directly, and they had a chat. This is a typical pattern illustrating how people and the robot started to interact. Overall, people seemed to enjoy seeing a robot that approached them and spoke.

To evaluate the performance, we compared the situation with the developed system “with anticipation”, and “without anticipation”, and measured how much the anticipation mechanism improved the efficiency. In the “without anticipation” condition, the robot simply approached the nearest person who is doing the *stop* behavior. We measured the performance for one hour in total for each condition. We prepared several time slots and counter-balanced the order.

Figure 16 shows the number of people to whom the robot provided services. Due to the novelty of the robot, people often initiated interactions on their own; in such cases, the anticipation mechanism is irrelevant. Thus, we classified the robot’s interactions into two categories. The first case, "robot-initiated", is the situation where the robot initiated the service by approaching the person and entering into conversation distance. Thus, the number of "robot initiated" services indicates how the robot's anticipation system improved the efficiency of the service. The second case, "person-initiated", is the situation where the person approached the robot while it was talking to someone else. While the robot was talking with two girls, a child came from the left. When the girls left, the child stood in front of the robot to start talking with it.

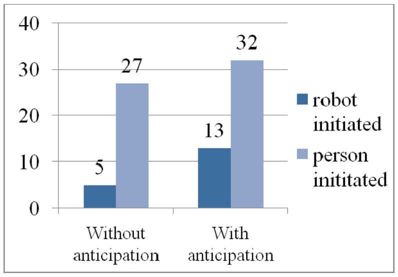


Figure 16. The number of services provided

The results in Figure 16 indicate that the number of "robot-initiated" services in "with anticipation" is much larger than "without anticipation." In other words, anticipating enables the robot to provide the service more efficiently. Due to the novelty factor of the robot, the number of "person-initiated" services is quite large. We believe that in the future when robots are no longer so novel to people, there will be less person-initiated interaction, and the results concerning anticipation will become much more significant.

Invitation Application

The second example is one in which the robot recommends and invites the customer to visit a shop. In the shopping arcade, attracting people's attention to shops and products is an important task. We believe that this is also a reasonable service to expect from a robot, as the novelty of robots makes them very effective in attracting people's attention. The contents the robot provided were simple; for example, the robot said, "Hello, I'm Robovie. Do you enjoy shopping? I'd like to recommend this shop, where they sell clothes by the kilogram!"

We chose *idle-walk* as the target local behavior, because people who are walking slowly might be window-shopping. We set the non-target local behavior as *fast-walk*, so as not to bother people who seem uninterested in shopping. We used anticipation and the pre-approach function for *idle-walk* behavior; when the robot predicts a person's future behavior as *idle-walk*, it moves towards that person's location.

We ran a field trial with the invitation robot in the shopping arcade as well. Just as in the entertainment application, the robot modified its behavior in accordance with the anticipation mechanism; the robot roamed around in front of the shop, where *idle-walk* was anticipated to be most likely, and approached people who were window-shopping.

In the demonstration, many people were interested in the robot and listened to its invitations. Figure 1 shows an impressive example where the robot approached a couple who were performing *idle-walk*. When the robot pointed to the shop and gave its recommendation (Figure 1 (c)), they smiled with surprise to see a robot performing a real business task. After the robot mentioned the shop, the woman walked directly to the shop and entered it (Figure 1 (d)). Observing such behavior indicates that such an invitation task can be a promising application. As indicated above, the robot was able to attract people's attention and redirect their interests to shops and products.

DISCUSSION

Does the presence of the robot affect global behavior?

Our model is based on data recorded without having a robot in the environment. The system tried to predict people's behavior independent of the presence of the robot. However, as a robot is still a novel object, some people were attracted by the robot, slowed down, approached the robot, and even talked to the robot. In this case, the prediction cannot be correct, since such the behaviors are not in the model.

For the application shown in this paper, this had a positive effect on the robot's ability to provide the service. Even when the prediction from the robot was incorrect, as the robot approached, sometimes the person was nevertheless attracted by the presence of the robot, and stopped, which enabled the robot to provide its service.

Other Possible Ubicomp Applications

We believe that the infrastructure shown in the paper can be useful for other ubicomp systems as well. One possible direction is to apply it to ambient intelligent environments, in which facilities (robots, display, music, illumination, etc.) are proactively controlled according to the types of users. For instance, an electronic poster could anticipate who is likely to stop nearby, and change its advertisement content in advance to something targeted to that person.

Another possibility is to combine it with mobile devices. Although GPS and WiFi have been used for locating people, laser range finders can provide more accurate positioning. The information provided by the infrastructure developed here could also complement other location-based services. For instance, if a user with a mobile device providing pedestrian navigation information entered a space with this infrastructure available, the device could then present additional information appropriate to that user's anticipated global behavior.

Privacy Concerns

Ubiquitous computing systems operating in public spaces should be carefully designed to protect the privacy of people. In our application, the system does not identify individuals, and it finishes tracking people when they leave the environment. We believe that this is a privacy-safe application. When the system is scaled up (e.g. extended to cover a large area, or associated with personal information), privacy should be more carefully considered.

CONCLUSION

We reported a series of techniques for retrieving information about people's behavior from their trajectories, and to use it for providing services to them. The system we developed utilizes accumulated trajectories to anticipate people's behavior, and it sends a robot to provide services to people who are exhibiting a pre-defined local behavior associated with a particular service.

Results from our field trial indicate that entertainment and invitation are promising applications for the robot. People appeared excited about the presence of the robot, enjoyed interacting with it, and sometimes followed its invitations. The infrastructure developed here enables the robot to provide such services in a real shopping arcade. The area covered by the system is still small; however, we believe that this paper illustrates the powerful potential of connecting a ubiquitous sensor network and a social robot, which can then provide information specific to a person's individual situation.

ACKNOWLEDGEMENT

We wish to thank Dr. Miyashita, Mr. Nishio, and Mr. Nohara for their help. This research was supported by the Special Coordination Funds for Promoting Science and Technology of the Ministry of Education, Culture, Sports, Science and Technology, the Japanese Government.

REFERENCES

1.Aipperspach, R., Rattenbury, T., Woodruff, A., Canny, J., A Quantitative Method for Revealing and Comparing Places in the Home, *In Proc. Ubicomp 2006*, (2006), 1-18.

2.Antoninia, G., Bierlaireb, M., and Weber, M., Discrete choice models of pedestrian walking behavior, *Transportation Research Part B,* 40(8), (2006), 667-687.

3.Bennewitz, M., Burgard, W., Cielniak, G., Thrun, S., Learning Motion Patterns of People for Compliant Robot Motion, *The Int. Journal of Robotics Research,* 24(1), (2005), 31-48.

4.Batty M, Predicting where we walk, *Nature*, 388, (1997).

5.Chai X, and Yang, Q, Multiple-Goal Recognition From Low-level Signals, *In Proc. AAAI-05*, (2005), 3-8.

6.Dautenhahn, K., *et al.*, How may I serve you?: a robot companion approaching a seated person in a helping context, *In Proc. Int. Conf. on Human-robot interaction (HRI2006)*, (2006), 172 – 179.

7.Eagle, N., and Pentland, A., Reality Mining: Sensing Complex Social Systems, *Personal and Ubiquitous Computing,* Vol. 10, No. 4, (2006), 255-268.

8. Glas, D., *et al.*, Laser Tracking of Human Body Motion Using Adaptive Shape Modeling, *In Proc. Int. Conf. Intelligent Robots and Systems (IROS2007)*, (2007), 602-608.

9.Helbing D., Schweitzer, F., Keltsch, J., Molnár, P., Active walker model for the formation of human and animal trail systems, *Physical review E,*56(3), (1997), 2527-2539.

10. Hoffman, G., Breazeal, C., Effects of Anticipatory Action on Human-Robot Teamwork, *In Proc. Int. Conf. on Human-Robot Interaction (HRI2007)*, (2007), 1-8.

11. Hua, C., Wu, H., Chen, Q., Wada. T., A General Framework for Tracking People, *In Proc. Int. Conf. Automatic Face and Gesture Recognition (FG2006)*, (2006), 511-516.

12. Kanda, T., *et al*, Analysis of People Trajectories with Ubiquitous Sensors in a Science Museum, *In Proc. Int. Conf. on Robotics and Automation (ICRA2007),* (2007), 4846-4853.

13. Koile, K., *et al.*, Activity Zones for Context-Aware Computing, *In Proc. Ubicomp 2003,* (2003), 90-106.

14. Krumm J., and Horvitz, E., Predestination: Inferring Destinations from Partial Trajectories, *In Proc. Ubicomp2006,* (2006), 243-260.

15. Letchner, J., *et al.*, A., Large-Scale Localization from Wireless Signal Strength, *In Proc. AAAI-05,* (2005).

16. Liao, L., *et al.*, Location-Based Activity Recognition using Relational Markov Networks, *In Proc. IJCAI-05,* (2005).

17.Liao, L., Patterson, D., Fox, D., Kautz, H., Learning and Inferring Transportation Routines, *Artificial Intelligence,* 171(5-6), (2007), 311-331.

18.Madhavapeddy, A., and Tse, A., A study of Bluetooth propagation using accurate indoor location mapping, *In Proc. Ubicomp2005,* (2005), 105-122.

19. Nurmi, P., Koolwaaij, J., Identifying meaningful locations, *In Proc. Mobiquitous 2006,* (2006), 1-8.

20. Patel, S. N., *et al.*, PowerLine Positioning: A Practical Sub-Room-Level Indoor Location System for Domestic Use. *In Proc. Ubicomp 2006,* (2006), 441-458.

21.Schulz, D., Fox, D., and Hightower, J., People Tracking with Anonymous and ID-sensors Using Rao-Blackwellised Particle Filters, *In Proc. IJCAI-03,* (2003).

22. Shao, X., *et al.*, Detection and tracking of multiple pedestrians by using laser range scanners, *In Proc. Int. Conf. on Intelligent Robots and Systems (IROS 2007),* (2007), 2174-2179.

23.Sparacino, F., The Museum Wearable, *In Proc. Museums and the Web (MW2002),* (2002).

24. Subramanya, A., Raj, A., Bilmes, J., Fox, D., Recognizing Activities and Spatial Context Using Wearable Sensors, *In Proc. Uncertainty in Artificial Intelligence,* (2006).

25.Suzuki, N., *et al.*, Learning Motion Patterns and Anomaly Detection by Human Trajectory Analysis, *In Proc. Int. Conf. Systems, Man and Cybernetics (SMC2007)*, (2007).

26.Turner, A., Penn, A., Encoding natural movement as an agent-based system: an investigation into human pedestrian behavior in the built environment, *Environment and Planning B: Planning and Design,* 29, (2002), 473-490.