

Tracing Commodities in Indoor Environments for Service Robotics^{*}

Oscar Martinez Mozos^{*} François Chollet^{*} Kouji Murakami^{**}
Ken'ichi Morooka^{*} Tokuo Tsuji^{*} Ryo Kurazume^{*}
Tsutomu Hasegawa^{*}

^{*} Faculty of Information Science and Electrical Engineering, Kyushu University 744 Motoooka, Nishi-ku, Fukuoka 819-0395, Japan
(e-mails: omozos@irvs.ait.kyushu-u.ac.jp, francois.chollet@gmail.com,
morooka@ait.kyushu-u.ac.jp, tsuji@irvs.ait.kyushu-u.ac.jp,
hasegawa@ait.kyushu-u.ac.jp, kurazume@ait.kyushu-u.ac.jp).

^{**} Faculty of Engineering, Kyushu Sangyo University, 3-1, Matsukadai 2-Chome, Higashi-ku, Fukuoka 813-8503, Japan
(e-mail: mura.k@ip.kyusan-u.ac.jp).

Abstract: Daily life assistance for elderly people is one of the most promising scenarios for service robots in the near future. In particular, the go-and-fetch task will be one of the most demanding tasks in these cases. In this paper, we present an informationally structured room that supports a service robot in the task of daily object fetching. Our environment contains different distributed sensors including a floor sensing system and several intelligent cabinets. Sensor information is sent to a centralized management system which processes the data and makes it available to a service robot which is assisting people in the room. We additionally present the first steps of an intelligent framework used to maintain information about locations of commodities in our informationally structured room. This information will be used by the service robot to find objects under request. One of the main goal of our intelligent environment is to maintain a small number of sensors to avoid interfering with the daily activity of people and to reduce as much as possible the invasion of their privacy. In order to compensate this limited available sensor information our framework aims to exploit knowledge about people's activity and their interaction with objects to infer reliable information about the location of commodities. This paper presents simulated results that demonstrate the suitability of this framework to be applied to a service robotic environment equipped with limited sensors. In addition we discuss some preliminary experiments using our real environment and robot.

Keywords: Mobile Robots; Artificial Intelligence; Sensor Systems; Probabilities Integration; Sensor Fusion; Systems Design

1. INTRODUCTION

Inside the many applications of advanced service robotics, elderly care is one of the most promising ones both in social and economic terms. It is expected that service robots will soon be playing a role of companion to elderly people, or a role of assistant to humans with special needs at home (Kawamura and Iskarous (1994), Roy et al. (2000), Kim et al. (2009)). In these scenarios, one of the most demanding tasks by the users will be the go-and-fetch of objects that are needed for their everyday activities (Srinivasa et al. (2010), Dario et al. (1999)). There is a variant of this task that is of special interest in cases of people with special memory loss problems such as Alzheimer patients. This particular task is the retrieval of commodities that are lost by people, i.e. people forgot the location of these objects at some point in time. Finding these lost (or forgotten) objects is a challenging task for

a service robot, and it needs an indoor object tracking system that can provide information about the possible positions of objects at any time.

Several approaches to the indoor object-tracking problem have been presented before. For example Fukui et al. (2008) proposed to attach a RFID reader to a cabinet so as to keep track of what is being stored inside. However, the object cannot be tracked when it is out of the cabinet. For tracking objects outside furniture, some solutions have used ultrasonic tags attached to objects combined with receivers on the ceiling (Nishida et al. (2003)), or directional RFID signal reception (Deyle et al. (2009)). The first previous case has the inconvenience of high cost and big size of the tags, whilst in the second case the position resolution is very low. An alternative is to use a vision system that monitors the whole environment. Nelson and Green (2002) present a system to track objects by visual recognition using cameras that monitor specific locations in the home. However, fixed cameras in the environment have the problem of being frequently occluded by people.

^{*} This work is partially supported by Grants-in-Aid for Foreign Fellows from the Japan Society for the Promotion of Science (22-00362) and by Grant-in-Aid for Scientific Research (B) (23360115).

Moreover, environments that are monitored by cameras constitute a significant invasion of people's privacy.

The previous systems often assume that objects are traceable at any time. This assumption is not valid in our scenario since we aim for a limited sensor installation which results in possible loss of information about the position of objects during undefined periods of time. Furthermore, with our limited set of sensors we aim to achieve a subtle, non-invasive human support system.

In this paper, we present an informationally structured room that supports a humanoid service robot in the task of daily object fetching. Our environment contains different distributed sensors including several intelligent cabinets, and a floor sensing system. The intelligent cabinets are provided with RFID sensors allowing the detection and recognition of tagged objects with high precision. The floor sensor systems is equipped with a laser range finder and it is used for people detection and tracking. Information from both sensor modalities is processed by a centralized management system which evaluates the data and provides information about possible locations for objects. This is done through a probabilistic framework which provides maps with possible location of objects at any time. In addition, the framework makes use of punctual observations made by the service robot to reduce the uncertainty in the maps.

One of the main objectives of our system is to avoid interfering with the daily activity of people, and to reduce as much as possible invasion of their privacy. For these reasons we do not use static cameras in our environment. Moreover, we keep the number of distributed sensors small and restricted to very particular places like for example cabinets. In order to compensate this limited available sensor information our framework aims to exploit knowledge about affordances of places and objects in the environment, together with people's activity and interaction with objects, to infer reliable information about the location of objects.

2. INFORMATIONALLY STRUCTURED ROOM

This section briefly describes the different components of our informationally structured environment. In particular, our scenario represents a room in a house as shown in Figure 1. The room contains two cabinets, a bed, and a desk with a chair.

The cabinets installed in this room are equipped with RFID readers which are used to detect and recognize tagged objects. Tag IDs are used to access the attributes of the objects in the data management system. An example detection of objects in one intelligent cabinet is shown in Fig. 2. Further details about our intelligent cabinets can be found in Murakami et al. (2010).

In addition to the intelligent cabinets, the room is equipped with a floor sensing system used to detect people and objects on the floor (Nohara et al. (2010), Murakami et al. (2010)). This system is composed of a laser range finder and it is extended with a lateral mirror. This configuration allows a reduction of dead angles of the LRF and it is more robust against occlusions.

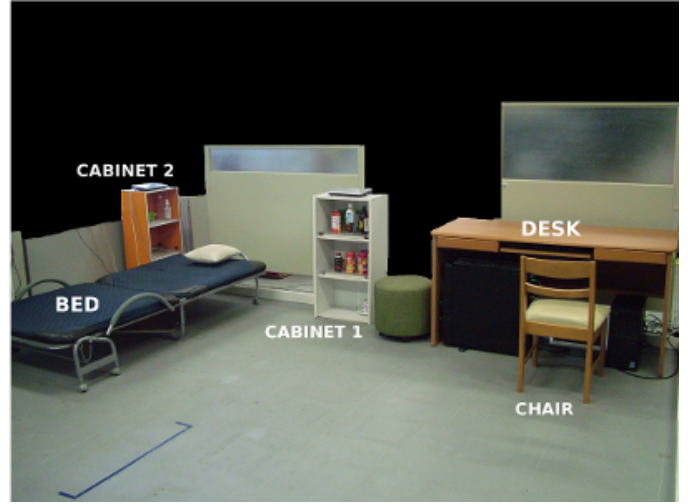


Fig. 1. Room used as scenario in our experiments. It contains two cabinets, one bed, and one desk with a chair.

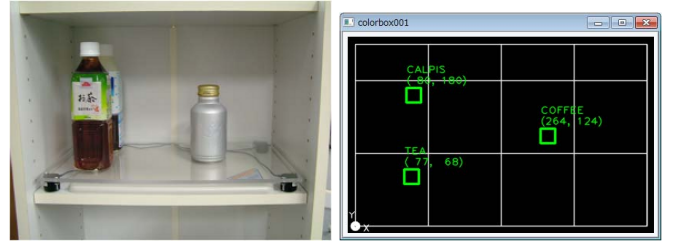


Fig. 2. Information about objects provided by the intelligent cabinet. Squares in the screen indicate the position of the different objects together with their description.

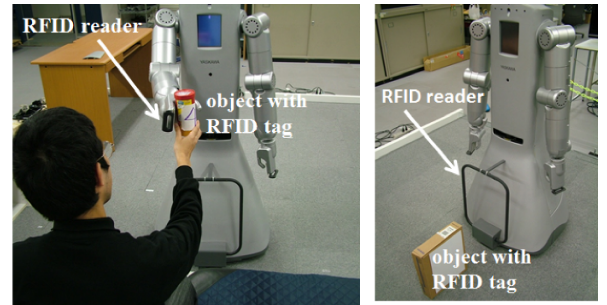


Fig. 3. Assistive humanoid robot SmartPal equipped with RFID readers.

Finally, the person acting in the room is assisted by a SmartPal humanoid robot (Fig. 3) from Yaskawa Electric Corporation. This robot will be responsible for fetching objects or pointing to them. The robot is composed of a mobile platform, two arms with seven joints, and one-joint grippers used as hands. Additional RFID readers are situated on the hands and front of the robot as shown in Fig. 3.

The previous sensing system and the robot itself are connected to a central database management system which processes sensor information and provides it under request (Murakami et al. (2009)). The flow of information in this system is shown in Fig. 4.

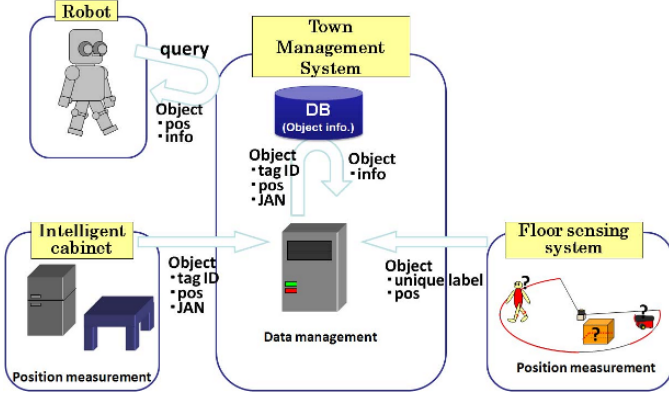


Fig. 4. Information management system.

3. PROBABILISTIC FRAMEWORK FOR OBJECT TRACING

Based on the intelligent room previously presented we now propose initial ideas of a framework for an intelligent tracing of objects with the aim of supporting a service robot. This framework uses information about activity of people and their interaction with objects, together with the sensor information, to generate hypotheses about the positions of the commodities in the environment. This hypotheses will be used by the service robot to find objects in the environment under people's requests.

The framework starts by dividing the elements contained in the environment into three classes. The first class contains the commodities to be traced by the system. Examples of such objects are books, keys, or bottles. We refer to these objects as *target objects*. The elements of the second class are called *hot-spots*. A hot-spot constitutes place which contains some sensor able to detect and recognize a target object with high reliability. In our room scenario these elements are constituted by intelligent cabinets. The rest of places in the environment are classified as *open space*. Moreover, we assume there exists a people tracking system that can provide hypotheses about the position of the person in the environment. Finally, we also make the assumption that the service robot is able to recognize target objects with high reliability when it is located sufficiently close to them. When this is the case we refer to the robot as a *mobile hot-spot*.

We have created a simulated environment with the previous characteristics and elements. This environment is shown in Fig. 5. The space is discretized into cells $x_1 \dots x_N$ where each $x_n = (i, j)$ represents the 2D-coordinates of the cell. In this scenario hot-spots represent a fridge, a shelf and a key cabinet (green cells). Open space is formed by the floor (white cells) and furniture without sensors (blue cells). The target objects used in this scenario are a bottle, a key and a book (yellow circles).

Initially objects are located at some storage hotspot. When an object disappears from its hotspot, we assume that the nearest human has taken it. We then associate the object's position to the carrier's position. We consider that at each instant, the object can be dropped by the person, either accidentally, e.g. on the floor; or intentionally, e.g. placed on a desk. In this first system we don't consider the option

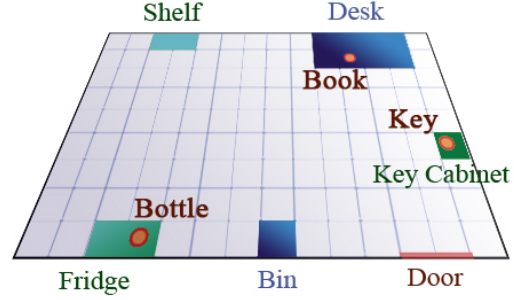


Fig. 5. Simulated environment used in our experiments. Green rectangles indicate hot-spots, yellow circles represent objects, blue rectangles indicate furniture with no sensors, and finally white squares depict the floor. The environment is discretized into cells.

that an object could be picked up later after having been dropped outside the initial storage hotspot.

3.1 Object Usage Maps

Whenever a person is assigned an object and moves around the room the object presence probability is spread over the cells of the environment, progressively increasing information entropy. We assume that the probability that the human will drop the object at their current position is a function of the time since the object was last picked up, and the position of the human carrier (some objects are dropped in some places more often than others). These probabilities are represented by *Object Usage Maps* describing a cycle that necessarily starts at a storage hotspot and end once the object has been dropped, either at a storage hotspot or on open space.

There are two kinds of Usage Maps: one map is over time (within a certain range), and the other is over space. Each map is object-specific, i.e., each object contains two Usage Maps: one over time (Temporal Usage Map) and one over space (Spatial Usage Map).

Usage Maps are calculated by averaging over observations made about the time and place of occurrences of object drops in the environment. In our implementation, these maps are obtained by summing Gaussians centered on data points gathered from observations. If we know that in the past the object has been dropped at cells $x_1 \dots x_N$ at times $t_1 \dots t_N$, the Spatial Usage Map is then the normalized sum of N Gaussians centered around the points x_i , and the Temporal Usage Map is the normalized sum of N Gaussians centered around the points t_i . The variance of these Gaussians is function of the incertitude range of the observations obtained by each sensor. In our experiments these variances were set by hand depending on the simulated sensor.

Usage maps are learned using the information provided by the hot-spots and the people tracker. In our system the human's behavior is modeled through a set of different activities. Each activity is represented by a sizeable Markov automaton where states are sub-steps of the activities. At each step there is a certain probability of changed the behavior (for example, forgetting what he was doing), dropping an object accidentally, or leaving an object on

some place. Repeating different behaviors several times we obtain Spatial and Temporal Usage maps for each target object. Examples of such maps are shown in Fig. 6 and Fig. 7. Each usage map is additionally assigned a *trust index* TI which quantifies how reliable the information is. The idea behind these indexes is to give more weight to usage maps with little entropy.

Alternatively, usage maps can be established a priori using the semantic knowledge and affordances in the environment. For example, desks and cabinets are places where people have higher probability of leaving objects. In a similar way we can establish prior probabilities for dropping different objects along time.

3.2 Presence Maps

Presence maps provides the service robot with hypotheses about possible locations of target objects in the environment. Presence maps are constructed using the information provided by Usage Maps combined with human position tracking information.

Each time the human carrying a target object o passes through a cell x , we augment the probability of finding o in x based on the probability that the human just dropped the object on x . This probability results from the fusion of the data from the Spatial and Temporal Usage Maps. In our initial design we decided for a normalized weighted sum using the trust indexes of each usage map

$$P_{drop}^o(x, t_a, t_b) = \frac{TI_{time}^o \int_{t_a}^{t_b} U_{time}^o(t') dt' + TI_{space}^o U_{space}^o(x)}{TI_{time}^o + TI_{space}^o} \quad (1)$$

where $P_{drop}^o(x, t_a, t_b)$ represents the probability that object o is dropped at cell x between times t_a and t_b , U_{time}^o and U_{space}^o indicate the temporal and space usage maps respectively for object o , and TI_{time}^o and TI_{space}^o are the trust indexes for temporal and space usage maps.

3.3 Activating Observations from a Mobile Hot-Spot

As previously explained, our framework uses a mobile hot-spot (the service robot in our case) to deliver observations in case the uncertainty of the presence map is high. In our approach we measure the uncertainty Q by using the following expression

$$Q = \frac{H(P^o)}{H_{max}(P^o)}, \quad (2)$$

where P^o represents the presence map of object o , and $H(.)$ is its entropy. The maximum entropy is obtained when the presence probability is homogeneous over the space considered. When Q is above a threshold, we send the service robot to make an observation in the environment. If the object is found at position x then the presence map becomes entirely determined, with a probability of 1 of finding the object at x . If the object is not found at the observe position x this position gets a value of 0 and its previous probability value is spread along the map.

The mobile hot-spot is initially sent to the cell where the probability of finding the object is higher. In our

Table 1. Possible activities and dropping places for each type of object.

Activity	Object	Places
Opening the door	key	Key cabinet, door
Drinking	Bottle	fridge, table, bin
Reading	Book	shelf, table

simulation results, we found that this can be done merely at the cost of occasional observations once the usage maps are well established (see Sect. 4).

4. EXPERIMENTS

In this section we present experiments to evaluate our informationally structured room. First we describe experiments in which we simulate our framework for object tracing. We also describe experiments using our real environment in which we applied an approach that do not take into consideration the movement of the person or the temporal usage information (Murakami et al. (2012)). In the future we aim to combine both approaches into a common framework.

4.1 Simulations

We implemented a Matlab simulation modeling the one-room environment of Fig. 5. Three activities are possible in this environment, each one related to one object, and to a number of preferred places for object storage and usage. This information is shown in Table 1.

Based on this simulated environment, we obtained the usage maps of Figures 6 and 7, illustrating how the usage of each object follows characteristic spatio-temporal patterns. These figures are generated from approximately 10 successful observations for each object.

The resulting presence map for each object proves to be accurate, with the object being on average at the most likely location 69.2% of the time, and at one of the top 10 most likely locations (and thus easily retrievable) 94.1% of the time (with a normalized entropy threshold of 0.2). Checks by the mobile hot-spot (robot) are very frequent when the system is still lacking appropriate usage data. At this stage, the typical system behavior pattern as shown in Fig. 8. When the usage maps are well established (typically after more than 5 successful object discoveries), place checks stabilize at about one per 5 minutes of non-static object use. This is illustrated in Fig. 9.

4.2 Experiments in the Real Environment

In addition to the previous simulation, in this section we want to present an alternative approach in which we are currently working to track objects using Sequential Importance Resampling (SIR) particle filters (Murakami et al. (2012)). We plan to integrate this system into our presented framework to take advantages of both approaches.

This experiment used a slightly modified setup in our room as shown Fig. 10. There are 20 objects in total although we only move 6 of them: OBJ0,...,OBJ5. Each object can be in one of the following states: in the cabinet, hold by the person, hold by the robot, on the desk, or on the floor. In the last case we restricted to 4 different positions on

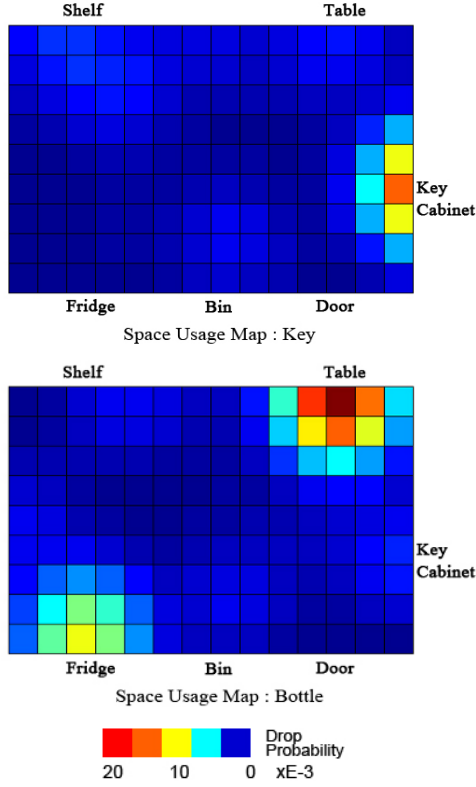


Fig. 6. Spatial usage maps for a key (top), and a bottle (bottom) after simulating the activities of Table 1.

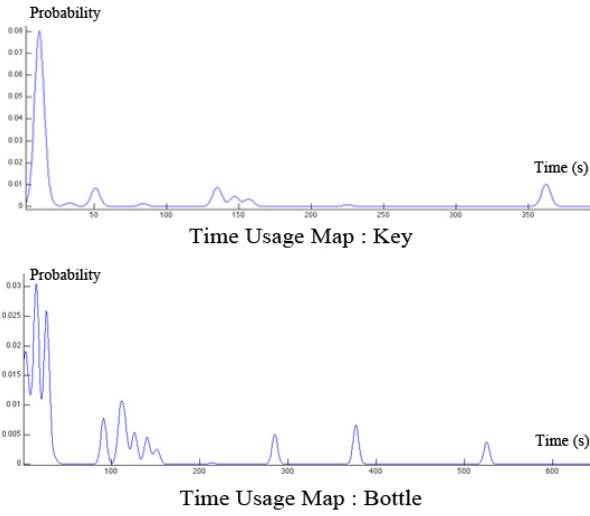


Fig. 7. Temporal usage maps for a key (top), and a bottle (bottom) after simulating the activities of Table 1.

the floor P1,...,P4. At the beginning of the experiment all objects are located on the cabinets. Then the person starts taking them and putting them on different positions on the floor. Eventually he also gives some objects to the robot. The sequence of positions of the objects during the experiment is shown in Table 2. For example, OBJ4 is in the cabinet A until it is taken by the person at step 5, then the person gives the object to the robot which keeps it during steps 6 and 7. In step 8 the robot puts OBJ4 on the desk.

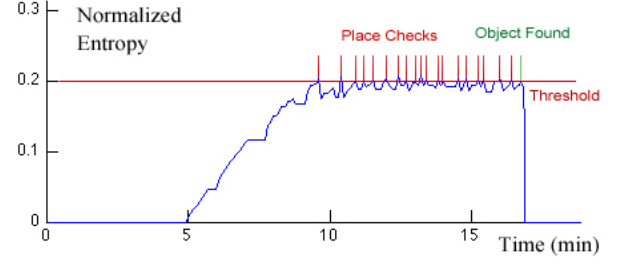


Fig. 8. System behavior when no prior data is available.

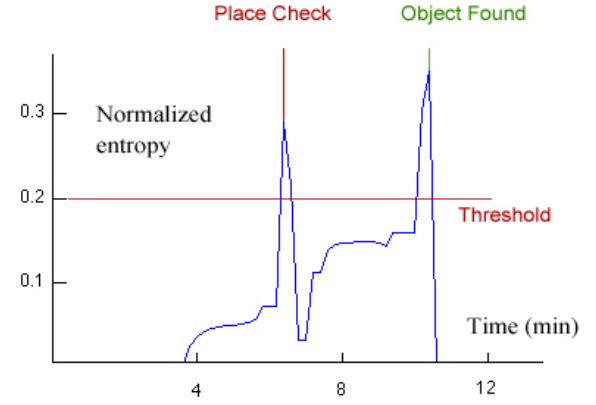


Fig. 9. System behavior when sufficient data is available.

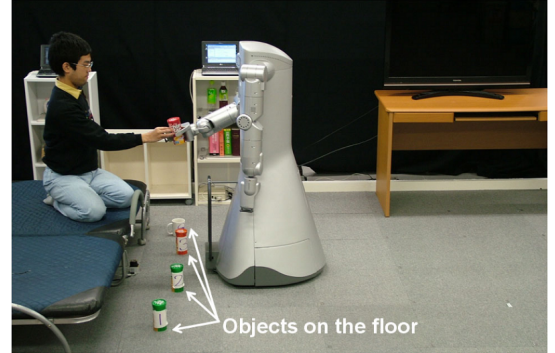


Fig. 10. Second setup for our experiments.

step	OBJ	OBJ0	OBJ1	OBJ2	OBJ3	OBJ4	OBJ5
1		(in Cab. A)	(in Cab. A)	(in Cab. A)	(in Cab. A)	(in Cab. A)	(in Cab. A)
2	person	P1	(in Cab. A)	(in Cab. A)	(in Cab. A)	(in Cab. A)	(in Cab. A)
3	person	P1	P2	(in Cab. A)	(in Cab. A)	(in Cab. A)	(in Cab. A)
4	P3	P1	P2	person	(in Cab. A)	(in Cab. A)	(in Cab. A)
5	P3	P1	P2	P4	person	(in Cab. A)	(in Cab. A)
6	P3	P1	P2	P4	robot	(in Cab. A)	(in Cab. A)
7	P3	P1	P2	P4	robot	(in Cab. A)	(in Cab. A)
8	P3	P1	P2	P4	(on desk)	(in Cab. A)	(in Cab. A)
9	P3	P1	P2	P4	(on desk)	P5	(in Cab. A)
10	P3	P1	P2	P4	(on desk)	P5	(in Cab. A)

Table 2. Position of the objects at each step.

Examples of the tracking system at different steps are shown in Fig. 11 and Fig. 12. In this figures the left image shows the measured positions of the objects. A colored circle indicates the position of an object stored in an intelligent cabinet. A colored triangle indicates the position of an object held by the robot. A colored square with a label (P1,...,P4) indicates a position on the floor.

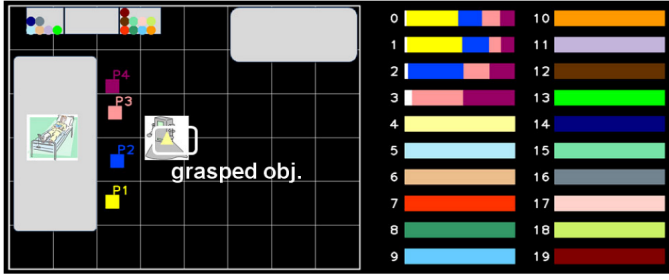


Fig. 11. States of the objects in the tracking system at step 6.



Fig. 12. States of the objects in the tracking system at step 7.

The right image in each figure shows the probability distribution over possible object locations. Numbers 0 to 19 indicate the IDs of the objects. In this first experiment we only move objects OBJ0 to OBJ5. The ratio of colors in each bar graph indicates the probability that an object is related to each position in the environment. Each position has its own color. The color corresponding to the person is white. Note that at step 6 (Fig. 11) the uncertainty is high for objects OBJ0 to OBJ3. In step 7 the robot checks the ID of the object located at P3. After this check the uncertainty reduces drastically as we can see in Fig. 12.

5. CONCLUSION

In this paper we have presented an informationally structured room whose goal is to support a service robot in go-and-fetch tasks. The information of the sensors is used as input for a framework that provides possible locations of objects inside the environment. This information is aimed to be used by the service robot to find objects. The framework presented in this paper presents preliminary results in simulations together with ongoing work on real scenarios.

As discussed in Section 3.1, spatial and temporal maps can be learned by observing the human behavior during certain time. However, it is clear that in home environments we can exploit more the semantic information about objects and places. For example there exist a higher probability of putting a book on the desk than inside the fridge. Our approach allows us to include this information as priors in the usage maps.

As future work we also want to study different options to fuse temporal and space information about objects usage. We further have to study the impact of the mobile hotspot in real situations when a human is living in the room. A possibility can be to send the robot when the

human activity is low (maybe at night). Finally we want to integrate and test the framework in our real environment.

REFERENCES

- Dario, P., Guglielmelli, E., Laschi, C., and Teti, G. (1999). Movaid: A personal robot in everyday life of disabled and elderly people. *Technology and Disability*, 10, 77–93.
- Deyle, T., Nguyen, H., Reynolds, M., and Kemp, C. (2009). Rf vision: Rfid receive signal strength indicator (rss) images for sensor fusion and mobile manipulation. In *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 5553–5560.
- Fukui, R., Morishita, H., Mori, T., and Sato, T. (2008). Development of a home-use automated container storage/retrieval system. In *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2875–2882.
- Kawamura, K. and Iskarous, M. (1994). Trends in service robots for the disabled and the elderly. In *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, volume 3, 1647–1654.
- Kim, M., Kim, S., Park, S., Choi, M.T., Kim, M., and Gomaa, H. (2009). Service robot for the elderly. *IEEE Robotics Automation Magazine*, 16(1), 34–45.
- Murakami, K., Hasegawa, T., Shigematsu, K., Sueyasu, F., Nohara, Y., Ahn, B.W., and Kurazume, R. (2010). Position tracking system of everyday objects in an everyday environment. In *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 3712–3718.
- Murakami, K., Matsuo, K., Hasegawa, T., and Kurazume (2012). Position tracking and recognition of everyday objects by using sensors embedded in an environment and mounted on mobile robots. In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, 2210–2216.
- Murakami, K., Hasegawa, T., Kurazume, R., and Kimuro, Y. (2009). Supporting robotic activities in informationally structured environment with distributed sensors and rfid tags. *Journal of Robotics and Mechatronics*, 21(4), 453–459.
- Nelson, R. and Green, I. (2002). Tracking objects using recognition. In *Proc. of the International Conference on Pattern Recognition*, 1025–1039.
- Nishida, Y., Aizawa, H., Hori, T., Hoffman, N., Kanade, T., and Kakikura, M. (2003). 3d ultrasonic tagging system for observing human activity. In *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, volume 1, 785–791.
- Nohara, Y., Hasegawa, T., and Murakami, K. (2010). Floor sensing system using laser range finder and mirror for localizing daily life commodities. In *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 1030–1035.
- Roy, N., Baltus, G., Fox, D., Gemperle, F., Goetz, J., Hirsch, T., Magaritis, D., Montemerlo, M., Pineau, J., Schulte, J., and Thrun, S. (2000). Towards personal service robots for the elderly. Carnegie Mellon University.
- Srinivasa, S., Ferguson, D., Helfrich, C., Berenson, D., Collet, A., Diankov, R., Gallagher, G., Hollinger, G., Kuffner, J., and VandeWeghe, M. (2010). Herb: A home exploring robotic butler. *Autonomous Robots*, 28, 5–20.