# Object tracking system by integrating multi-sensored data

Kouji Murakami
Faculty of Engineering,
Kyushu Sangyo University,
Fukuoka, Japan

Tokuo Tsuji
Institute of Science and Engineering,
Kanazawa University,
Ishikawa, Japan

Tsutomu Hasegawa
National Institute of Technology,
Kumamoto College,
Kumamoto, Japan

Ryo Kurazume
Faculty of Information Science and Electrical Engineering,
Kyushu University,
Fukuoka, Japan

***Abstract*—We propose an object tracking system which recognizes everyday objects and estimates their positions by using distributed sensors in a room and mobile robots. The placement of objects is frequently changed according to human activities. Although a passive RFID tag is attached to each object for the object's recognition, the placement is often not uniquely determined due to the deficiency of measured data. We have already proposed a method for estimating the placement of objects by using the moving trajectories of objects. This estimation result is expressed as the probability distribution of the object placement. However intersections of trajectories cause the decrease of the estimation accuracy. So we propose a new method based on Bayesian inference to improve the estimation accuracy by using the size and the shape of an object measured by laser range finder. Then a mobile robot settles the placement with small workload by using the mounted sensor. The system successfully recognized and localized 10 objects in the experiment.**

## I. Introduction

It is expected that a service robot supports the daily needs of elderly people in a home. An object fetching task is one of prospective applications for a service robot. If a service robot can bring a requested everyday object, it will make people's life comfortable. A service robot needs the placement of objects in a room to perform this task. Many position tracking or estimation methods have been presented. However, most of them use vision sensors and focus on the human tracking. The sizes of most everyday objects are small. Everyday objects are frequently held, carried and housed by a resident. In addition, everyday objects may be put into hidden spaces such as a pocket or a box. So, a conventional vision system for the position tracking will not be straightly applicable to this object tracking.

Some object tracking systems have been reported. A RFID(Radio Frequency IDentifier) tag [1] is often used for the position estimation of objects. A container with a RFID reader can detect whether or not an object with a RFID tag is in it [2]. However, the position of an object with a RFID tag is not measurable when the object is out of the container. The signal strength of an active tag is used for the position estimation [3]. The presented method estimates 3-dimensional position of an object with an UHF(Ultra High Frequency) RFID tag. However, an active tag is box-type and needs a battery. An ultrasonic tag is used for the position estimation [4]. The position of an ultrasonic tag is estimated by using ultrasonic receivers installed on the ceiling. An ultrasonic tag is attached to a target object. However, an usual ultrasonic tag is box-type, and is larger than an usual thin RFID tag. The positions of residents walking in a room have been measured by using pressure sensors distributed on the floor [5]. However, the weights of everyday objects are sometimes too light to detect it. Multiple vision sensors are used for the object tracking [6]. However, a resident will feel invasion of privacy under constant surveillance by vision sensors installed in a personal space.

We have already proposed a object tracking system that recognizes everyday objects and estimates their positions in a room by using distributed sensors and mobile robots [7]. The system integrates information measured by distributed sensors for the object's recognition and position estimation. However, the deficiency of measured data causes the decrease of the estimation accuracy. In this paper, we improve the system by using the size and the shape of an object as additional information in the data integration process.

## II. Object Tracking

We propose an object tracking system which recognizes everyday objects and estimates their positions by using distributed sensors. Assumptions for this system are as follows:

1) A passive RFID tag is attached to each object. An object is uniquely identified by its tag ID.
2) Sensor systems installed in an environment measure either both the position of an object and its tag ID or only the position of an object.

Each measured position by a sensor system is labeled to track it continuously. We call this label *tracking-label*. The result of the object's recognition and position estimation is obtained as a combination of the tag ID of an object and its position. We call a set of this combination "object placement". Fig.1 shows the object placement of five object. This object placement has five combinations of a tag ID and a "tracking-label" as the position of an object. If a sensor system can measure both the position of an object and its tag ID, a combination of the tag ID and the position of an object is obtained as a measurement result. On the other hand, a sensor system can measure only the position of an object, a combination of the tag ID and the position of an object is not obtained as a measurement result. The proposed system estimates the tag ID related to the measured position by using the SIR particle filter [8]. The SIR(sampling importance resampling) particle filter integrates sensored information for this estimation.

978-1-5090-3474-1/16/$31.00 ©2016 IEEE

The object placement changes according to human activities. In the estimation process of the object placement, the system needs the transition model of the object placement to reflect the change of the object placement caused by human activities in the estimation process.

Our previous system [7] uses only the moving trajectories of objects for this transition model. However intersections of trajectories cause the decrease of the estimation accuracy. In this paper, we improve the transition model by adding the size and shape of an object in order to raise the estimation accuracy.

Fig. 1. Object placement: a combination of tag ID and *tracking-label*.

The estimation accuracy of the object placement is correlated with the number of objects whose tag IDs are unmeasurable by sensors distributed in an environment. A mobile robot with a RFID reader can move to the position of each unmeasurable object and measure its tag ID. The result of this active sensing for one unmeasurable tag ID is integrated in the estimation process and improves the estimation accuracy of the object placement. We propose a method for choosing a prospective position of unmeasurable object to improve the estimation accuracy.

## III. Object Tracking System

### A. System components

Several types of sensors are installed to a cabinet, the floor, and a mobile robot. The proposed system collects three types of measured information from each sensored component. These information is integrated to estimate the object placement.

*1) Measured information from a sensored cabinet:* We installed a RFID reader and load cells to a cabinet. This sensored cabinet performs the object's recognition and position estimation. Fig.2 shows the measurement result for three bottles. A sensored cabinet outputs combinations of the tag IDs and the positions of the objects stored in it. Each position is related to *tracking-label*. Each *tracking-label* is uniquely generated when an object get put into a sensored cabinet. A sensored cabinet $i$ outputs information $Z^i$ as a measurement result.

$$Z^i = \{obj_1,\ obj_2,\ \dots,\ obj_S\} \tag{1}$$

$$obj = \{tracking-label,\ tagID\} \tag{2}$$

The set of output information of all sensored cabinets at the time $t$ is expressed as $Z_t = \{Z_t^i\}$.

*2) Measured information from a sensored floor:* We installed a laser range finder (LRF) and a mirror to the floor (Fig.3). This sensored floor performs the position estimation of objects and walkers on the floor [9]. A LRF can measure the distances between the LRF and objects on the floor. The positions of objects and walkers are continuously measured by using the background subtraction.

(a) three bottles (b) top view

Fig. 2. Position estimation and recognition of three bottles in a sensored cabinet.

The contour of an object on the floor is measured as a point cloud (Fig.4). A strip of mirror placed on a wall reflects the scanning beam of a LRF. These reflected scanning beams increase the number of scanable points on the surface of an object. The measurement result for four objects on the floor is shown in Fig.5.

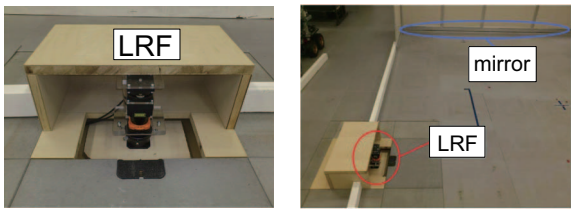



Fig. 3. The sensored floor.

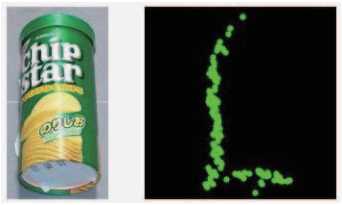


Fig. 4. A point cloud of an object measured by the sensored floor.

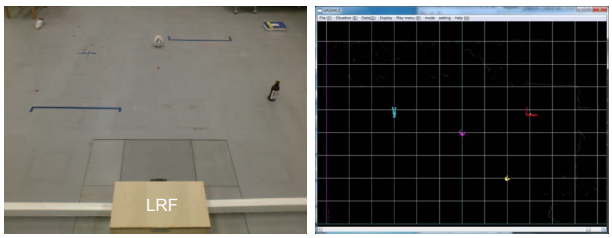



Fig. 5. Position measurement of four objects by using the sensored floor.

A sensored floor outputs combinations of the positions and $article$s of the objects on the floor. Each position is related to *tracking-label*. Each *tracking-label* is uniquely generated when an object get placed onto the floor. The sensored floor outputs information $Y$ as a measurement result.

$$Y = \{y_1,\ y_2,\ \dots,\ y_M\} \tag{3}$$

$$y = \{tracking-label,\ article\} \tag{4}$$

$$article = \{person\ or\ object\} \tag{5}$$

Output information of the sensored floor at the time $t$ is expressed as $Y_t = \{Y_t\}$.

*3) Measured information from a mobile robot with RFID readers:* A robot with a RFID reader performs the object's recognition and position estimation. This robot outputs combinations of the tag IDs and the positions of objects within the sensing are of the robot. Each position is related to *tracking-label*. A robot $i$ outputs information $A^i$ as a measurement result.

$$A^i = \{obj_1,\ obj_2,\ \ldots,\ obj_N\} \tag{6}$$

$$obj = \{tracking-label,\ tagID\} \tag{7}$$

The set of output information of all robots at the time $t$ is expressed as $A_t = \{A_t^i\}$.

### B. Estimation of object placement using SIR particle filter

We use the SIR particle filter to estimate the object placement by integrating sensored information. Each particle has the hypothesis of the object placement and its likelihood. The event $X_t$ means the object placement is $x_t$ at time $t$. The posterior probability $p(X_t|Z,\ Y,\ A)$ after the observation $Z_t$, $Y_t$, and $A_t$ is recursively estimated using particles. Here, $Z$, $Y$, and $A$ means events in which information $Z_t$, $Y_t$ and $A_t$ are obtained from sensored cabinets, the sensored floor, and mobile robots, respectively.

The procedure for using the SIR particle filter is:

1) Generation of initial particles:
$R$ particles $r_0^{(r)} = \{x_0^{(r)},\ w_0^{(r)}\}$ are generated at time 0. Here, $w_t^{(r)}$ is the weight at time $t$. Vector $x_t^{(r)} \in \Re^N$ represents the hypothesis of the object placement at time $t$ (Fig. 6). In this initial step, vector $x_0^{(r)}$ is randomly generated.

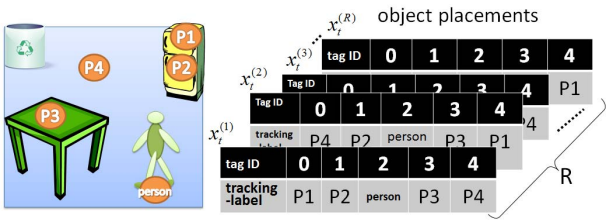



Fig. 6. Object placements of particles

2) State transition:
The object placement changes according to human activities. We modelize the change of the object placement caused by human activities. By applying this transition model to particles $r_{t-1}^{(r)}$, they shift to the next state $r_{t|t-1}^{(r)} = \{x_{t|t-1}^{(r)},\ w_t^{(r)}\}$.

**Transition model of object placement:**
We assume the change of the object placement is caused only by the following two actions:

(i) a person( or a mobile robot) puts a held object onto furniture or the floor,

(ii) a person( or a mobile robot) takes up an object.

The occurrence of the transition can be observed as the change in the total number of *tracking-labels* measured by the sensor systems.

**Transition operation(I):** When the total number of *tracking-labels* increases, the operator replaces the *tracking-label* of a person with the newly appeared *tracking-label* in vector $x_{t-1}^{(r)}$ of each particle. This operation indicates the action (i) mentioned above.
**Transition operation(II):** When the total number of *tracking-labels* decreases, the operator replaces the disappered *tracking-label* with the *tracking-label* of a person in vector $x_{t-1}^{(r)}$ of each particle. This operation indicates the action (ii) mentioned above.
Apply these transition operations to $R(1-K)$ particles $r_{t-1}^{(r)}$ chosen from the $R$ particles. The constant value $K$ $(0 < K < 1)$ depends on the reliability of each sensor system installed in an environment.

In the transition operations (I) and (II), if several persons are in a room, an operator needs to determine who has caused the action (i) or (ii). An operator chooses one of them by considering the probability which depends on the distance between the position of each person and the position of the *tracking-label* newly generated or removed.

In the transition operation (I), if a person carries some objects, their tag IDs are related to the *tracking-label* of the person in the object placement. The operator needs to determine which object has been put on the floor. However, the sensored floor can not measure its tag ID of the object. In the previous system, the operator chooses one of them with even probability. On the other hand, in the proposed method, the operator selects one of them with the selection probability of each object in order to improve the estimation accuracy. The selection probability is derived by using the contour of an object on the floor. The contour of an object is measured as a point cloud by the sensored floor (Fig.4). The detail of this transition operation is described in the next section.

3) Likelihood calculation:
The likelihood $p(Z,\ Y,\ A \mid X_{t|t-1}^{(r)})$ of each particle is calculated as
$p(Z,\ Y,\ A \mid X_{t|t-1}^{(r)}) =$

$$p(Z|X_{t|t-1}^{(r)})\ p(Y|X_{t|t-1}^{(r)})\ p(A|X_{t|t-1}^{(r)}) \tag{8}$$

where

$$p(Z|X_{t|t-1}^{(r)}) = C\prod_{s=1}^{S} g_s(x_t^{(r)}). \tag{9}$$

$g_s(x_t^{(r)})$ is a function that returns 1.0 if the *tracking-label* related to $the\ object\ s$ in $x_t^{(r)}$ corresponds to the *tracking-label* related to $the\ object\ s$ in $Z_t$, and returns the numerical constant $\alpha(0 < \alpha < 1)$ if it does not. The numeric constant $C$ $(0 < C < 1)$ depends on the reliability of the sensored cabinet. $p(A|X_{t|t-1}^{(r)})$ is also calculated from Eq.(9). $p(Y|X_{t|t-1}^{(r)})$ is the constant value.

We define the weight $w_t^{(r)}$ of a particle as

$$w_t^{(r)} = p(Z,\ Y,\ A|\ X_{t|t-1}^{(r)}). \tag{10}$$

The total weight of all particles is calculated as

$$w_t^{(all)} = \sum_{r=1}^{R} w_t^{(r)}. \tag{11}$$

4) Resampling:
$RD$ $(0 \leq D \leq 1)$ particles $r_t^{(r)}$ are chosen from $R$ particles. New $RD$ particles $r_{t+1}^{(r)}$ are generated depending to the probability $w_t^{(r)}/w_t^{(all)}$. New $R(1-D)$ particles $r_{t+1}^{(r)}$ are randomly generated.

5) Estimation of object placement:
The object placement is estimated based on the distribution of $R$ particles. The probability $p_{ij}$ that $the\ object\ i$ is related to $the\ tracking-label\ j$ is obtained as

$$p_{ij} = \frac{1}{R} \sum_{r=1}^{R} b_{ij}(r_t^{(r)}). \tag{12}$$

$b_{ij}(r_t^{(r)})$ is a function that returns the numerical constant 1.0 if the *tracking-label* related to $the\ object\ i$ in $x_t^{(r)}$ of a particle $r_t^{(r)}$ corresponds to $the\ tracking-label\ j$, and returns 0 if it is not.

### C. Active sensing for the tag ID of an object

The estimation accuracy of the object placement is correlated with the number of objects whose tag ID is unmeasurable by sensors installed in an environment. In order to improve the estimation accuracy, a mobile robot with a RFID reader moves to the position of this unmeasurable object and measures its tag ID. When there are several unmeasurable objects, a robot chooses one of them. For this active sensing, we propose a method for choosing the most prospective *tracking-label* to improve the estimation accuracy. This method is based on the entropy $H(E)$ of the probability distribution of the object placement. The entropy $H(E)$ is the index of insufficiently of sensored information. The entropy increases when sensored information is not enough to estimate the object placement. The proposed method chooses the most prospective *tracking-label* to reduce the entropy. The entropy $H(E)$ is calculated as

$$H(E) = \sum_{i}^{M} H(L_i) = -\sum_{i}^{M} \sum_{j}^{N} P(L_i = l_j) \cdot \log P(L_i = l_j). \tag{13}$$

Here, $M$ indicates the total number of *tracking-labels*. The probability $P(L_i = l_j)$ that $the\ object\ l_j$ is related to $the\ tracking-label\ L_i$ is calculated in the same as Eq.(12).

We defines the evaluation function $F(L_i)$ as the expectation value of mutual information after the tag ID of the object of the *tracking-label* $L_i$ is measured. The *tracking-label* $L_{max}$ which maximizes the evaluation function is chosen for the active sensing.

$$F(L_i) = \sum_{j}^{N} P(L_i = l_j) \cdot I(E; L_i = l_j) \tag{14}$$

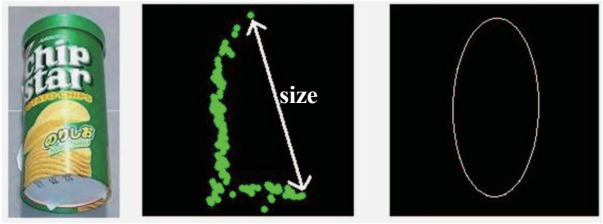
$$I(E; L_i = l_j) = H(E) - H(E|L_i = l_j) \tag{15}$$

Here, $H(E|L_i = l_j)$ is the entropy under the assumption that the object $l_j$ is related to the *tracking-label* $L_i$ .

## IV. Selection probability of an object placed on the floor

In the transition operation (I) of the proposed method, if a person holds several objects, the operator needs to determine which object of the held objects is placed on the floor. In the previous method, the operator chooses one of them with even probability. In the proposed method, the operator chooses one of them with the selection probability of each object in order to improve the estimation accuracy.

The selection probability is obtained by using the contour of an object on the floor. The contour of an object is measured as a point cloud by the sensored floor (Fig.7). We define *the object size* as the maximum distance between two points in a point cloud (Fig.7(b)). We also define *the contour image* as the elliptical approximation of this point cloud (Fig.7(c)). *The object size* and *the contour image* indicate a scale factor and an approximate shape of an object, respectively.




(a) (b) (c)

Fig. 7. A point cloud measured by the sensored floor ( (a) appearance of the object, (b) *the object size*, (c) *the contour image* ).

*The object size* and *the contour image* of each object include measurement errors. We determine the probability density function of measurement errors through the preliminary measurements. We derive the selection probability by using the probability density function, *the object size*, and *the contour image*. This procedure is divided into two parts

A) calculation of the probability density function of measurement errors, and
B) calculation of the selection probability of each object.

The contour of an object on the floor depends on its orientation (Fig.8). At the part A), we derive 1) the probability density function with respect to *the object size* and 2) the probability density function with respect to *the contour image* at each orientation of each object.

Fig. 8. *The contour images* of an object with different orientations.

At the part B), we derive the selection probability $p_i^{select}$ by using Bayesian inference. $p_i^{select}$ indicates the selection probability that the object $i$ is selected from the objects held by a person. The detail of the procedure is as follows.

### A. Probability density function for the measurement error

The contour of an object on the floor depends on its orientation. At the orientation $j$ of the object $i$, we derive (1) the probability density function $f_{ij}^{size}$ for the measurement error of *the object size*, and (2) the probability density function $f_{ij}^{contour}$ for the measurement error of *the contour image*. We prepare $f_{ij}^{size}$ and $f_{ij}^{contour}$ at each orientation $j$ of the object $i$ through the preliminary measurements. At each orientation $j$ of the object $i$, a point cloud of the object is obtained fifty times through the measurements of the sensored floor. Each data set $D_{ij}$ has fifty point clouds for the object $i$ with its orientation $j$.

*1) Probability density function for the measurement error of the object size: The object size* is calculated from each point cloud of the data set $D_{ij}$. The average $\overline{\mu}_{ij}$ and the standard deviation ${}^{size}\sigma_{ij}^2$ of *the object size* are calculated at each $D_{ij}$. The probability density function $f_{ij}^{size}$ for the measurement error of *the object size* is expressed as the Gaussian distribution $\mathcal{N}(0, {}^{size}\sigma_{ij}^2)$.

*2) Probability density function for the measurement error of the contour image:* We generate the ideal contour image $I_{ij}^{ideal}$ based on the actual shape of the object $i$ with its orientation $j$. This ideal contour image is compared with *the contour image* measured by the sensored floor. The image similarity is calculated through this comparison. We use seven Hu invariant moments for this comparison [10]. These seven Hu invariant moments are well known and independent parameters for the translation, the rotation, and the scale change. Hu invariant moments are expressed as following equations.

$$hu_1 = \eta_{20} + \eta_{02} \tag{16}$$

$$hu_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \tag{17}$$

$$hu_3 = (\eta_{30} - 3\eta)_{12}^2 + (3\eta_{21} - \eta_{03})^2 \tag{18}$$

$$hu_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \tag{19}$$

$$hu_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2]$$
$$+(3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \tag{20}$$

$$hu_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$
$$+4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \tag{21}$$

$$hu7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2]$$
$$-(\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \tag{22}$$

The image similarity $G(A, B)$ between the image A and the image B is expressed as

$$G(A, B) = \sum_{i=1}^{7} | \frac{1}{m_i^A} - \frac{1}{m_i^B} | \tag{23}$$

where

$$m_i^A = sign(h_i^A) \cdot log(h_i^A) \ , \ m_i^B = sign(h_i^B) \cdot log(h_i^B).$$

$h_i^A$ and $h_i^B$ indicates Hu invariant moments of the image A and the image B respectively.

The image similarity $G(A, B)$ is zero if the image A and the image B are same image. The calculation result of the image similarity is shown in Fig.9. The numerical values on the center and the middle image indicate the image similarity between the left image and each image.

Fig. 9. The calculation results of the image similarity between the left image and the middle or the right image.

*The contour image* $I_{ijk}^{contour}$ $(1 \leq k \leq 50)$ is obtained from each point cloud $k$ of the data set $D_{ij}$. The standard deviation ${}^{contour}\sigma_{ij}^2$ of the image similarity $G_{ijk}(I_{ij}^{ideal}, I_{ijk}^{contour})$ are calculated at each $D_{ij}$. The probability density function $f_{ij}^{contour}$ for the measurement error of *the contour image* is expressed as the Gaussian distribution $\mathcal{N}(0, {}^{contour}\sigma_{ij}^2)$.

### B. The selection probability

We derive the selection probability $p_i^{select}$ of the object $i$ from a point cloud $d_t$ of an object through the following procedure.

*1) Likelihood with respect to the object size: The object size* $S_t$ is calculated from a point cloud $d_t$ measured by the sensored floor at the time $t$. At each orientation $j$ of the object $i$, the measurement error $e_{ij}^{size}$ of *the object size* $S_t$ is expressed as

$$e_{ij}^{size} = S_t - \overline{\mu}_{ij}. \tag{24}$$

The probability $p_{ij}^{size}$ is obtained by substituting $e_{ij}^{size}$ into $f_{ij}^{size}$ (Fig.10). We determine $p_i^{size}$ as

$$p_i^{size} = \max_j(p_{ij}^{size}). \tag{25}$$

$p_i^{size}$ indicates the likelihood that the object $i$ is selected from the objects held by a person, which is based on the measurement error of *the object size*.

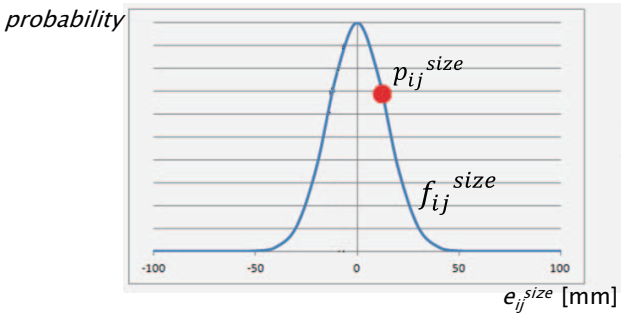



Fig. 10. Probability density function for the measurement error of *the object size*.

*2) Likelihood with respect to the contour image:* The *contour image* $C_t$ is calculated from a point cloud $d_t$ measured by the sensored floor at the time $t$. At each orientation $j$ of the object $i$, the measurement error $e_{ij}^{contour}$ of *the contour image* $C_t$ is expressed as

$$e_{ij}^{contour} = G(\ I_{ij}^{ideal},\ C_t\ ). \tag{26}$$

The probability $p_{ij}^{contour}$ is obtained by substituting $e_{ij}^{contour}$ into $f_{ij}^{contour}$ (Fig.11). We determine $p_i^{contour}$ as

$$p_i^{contour} = \max_j(p_{ij}^{contour}). \tag{27}$$

$p_i^{contour}$ indicates the likelihood that the object $i$ is selected from the objects held by a person, which is based on the measurement error of *the contour image*.

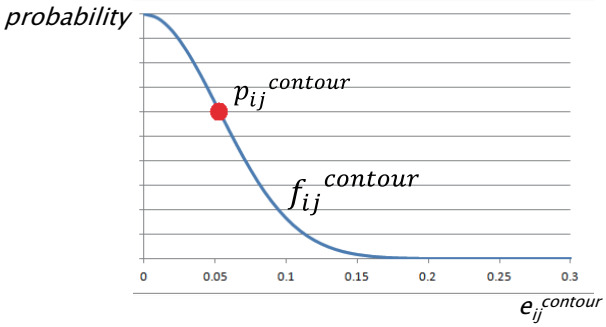



Fig. 11. Probability density function for the measurement error of *the contour image*.

*3) Selection probability of each object:* The selection probability $p_i^{select}$ indicates the probability that the object $i$ is selected from the objects held by a person. $p_i^{select}$ is obtained by applying Bayes' theorem to $p_i^{size}$ and $p_i^{contour}$. Bayes' theorem is expressed as

$$P(H_i|E) = \frac{P(E|H_i)P(H_i)}{P(E|H_1)P(H_1) + \ldots + P(E|H_n)P(H_n)}. \tag{28}$$

$H_i$ is an event that the object $i$ is selected. $P(H_i)$ is the probability of observing $H_i$. $E$ is an event that the measurement error $e$ is measured. $P(H_i|E)$ is the probability of observing $H_i$ given that $E$ has occurred. $P(H_i)$ is the prior probability. $P(H_i|E)$ is the posterior probability. $P(E|H_i)$ is the likelihood.

The number of the objects held by a person is $n$. $p_i^{select}$ is the probability that the object $i (1 \leq i \leq n)$ is selected. We derive $p_i^{select}$ through the following procedures (i) and (ii).

(i) Bayesian inference by using *the object size*: $E^{size}$ indicates observing the measurement error with respect to *the object size*. We determine $P(E^{size}|H_i)$ as

$$P(E^{size}|H_i) = p_i^{size}. \tag{29}$$

We determine $P(H_i)$ with even probability under the principle of insufficient reason. $P(H_i)$ is obtained as

$$P(H_i) = \frac{1}{n}. \tag{30}$$

$P(H_i|E^{size})$ is obtained by applying Eq.(29) and Eq.(30) to Eq.(28).

(ii) Bayesian inference by using *the contour image*: $E^{contour}$ indicates observing the measurement error with respect to *the contour image*. We determine $P(E^{contour}|H_i)$ as

$$P(E^{contour}|H_i) = p_i^{contour}. \tag{31}$$

We determine $P(H_i)$ as

$$P(H_i) = P(H_i|E^{size}). \tag{32}$$

$P(H_i|E^{contour})$ is obtained by applying Eq.(31) and Eq.(32) to Eq.(28).

We determine the selection probability $p_i^{select}$ as

$$p_i^{select} = P(H_i|E^{contour})\ ,\ (1 \leq i \leq n). \tag{33}$$

The estimation accuracy is improved by applying $p_i^{select}$ to the transition operation (I) in the proposed method.

## V. Experiment

The proposed system recognized ten objects and estimated their positions in the experiment. The experimental setup is shown in Fig.12. We installed two sensored cabinets and the sensored floor as mentioned in section III. In this experiment, we used ten everyday objects with RFID tags as shown in Fig.13. One person walked around in a room, and he moved several objects from sensored cabinets to the floor. The proposed system estimated the object placement which was frequently changed with human activity.

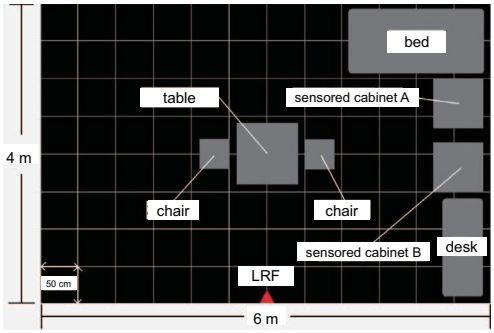



Fig. 12. Experimental setup.

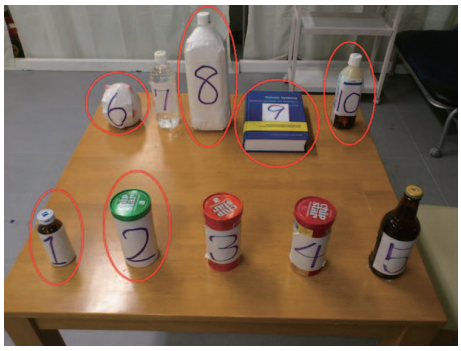


Fig. 13. Tracking target objects.

The scenario of the experiment is explained as following three steps. At the step 1, all objects are in two sensored cabinets. The estimation result of the object placement at the step 1 is shown in Fig.14. The measured positions by sensored cabinets are shown as colored circles. The numbers from 1 to 10 mean the tag IDs of objects. The position of each object is estimated as the probability of the relationship between each object and each measured position. This probability is obtained from Eq.(12). The length of each color in a bar graph indicates the value of its probability. Each color indicates a measured position. The sum of the probabilities in each bar graph is 1.0. Each bar graph is composed of mono color in Fig.14. This means that the position of each object is uniquely estimated. In addition, the object placement is correctly estimated. Because sensored cabinets can measure both the tag IDs and the positions of the objects in them.

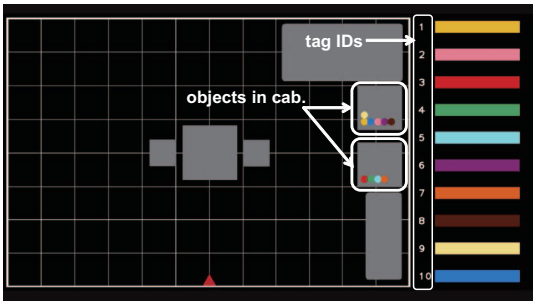



Fig. 14. Tracking result after the step 1 (Six objects are in the cabinet A, and four objects are in the cabinet B).

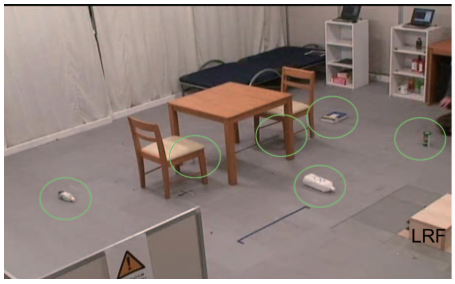



Fig. 15. Experimental process at the step 2 (Six objects are on the floor and four objects are stored in the cabinet B).

At the step 2, a person takes six objects from the sensored cabinet A. The tag IDs of these objects are 1, 2, 6, 8, 9, and 10. The person puts them one-by-one onto the floor as shown in Fig.15. At the step 2, the estimation results of the object placement are shown in Figs.16 and 17. The measured positions by the sensored floor are shown as colored squares with alphabets which indicate tracking-labels. The white color in a bar graph indicates the position of the person. So, the white color means that object is held by a person. When a bar graph is composed of several colors, the position of its object is not uniquely estimated. Because the sensored floor cannot measure tag IDs of objects on the floor. A bar graph in each lower image in Fig.17 also indicates the probability of the relationship between each object and each tracking-label. The vertical axis indicates the probability in each lower image. The horizontal axis indicates the tag IDs. From Fig.17, the accuracy rate of the position estimation of each object is ranging from 70 % to 100 % after the step 2.

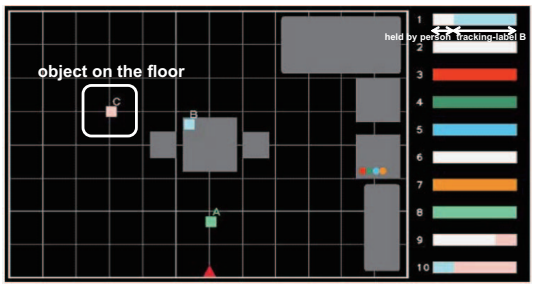



Fig. 16. Tracking result at the step 2 (Three objects are on the floor. Three objects are held by the subject. Four objects are in the cabinet B).

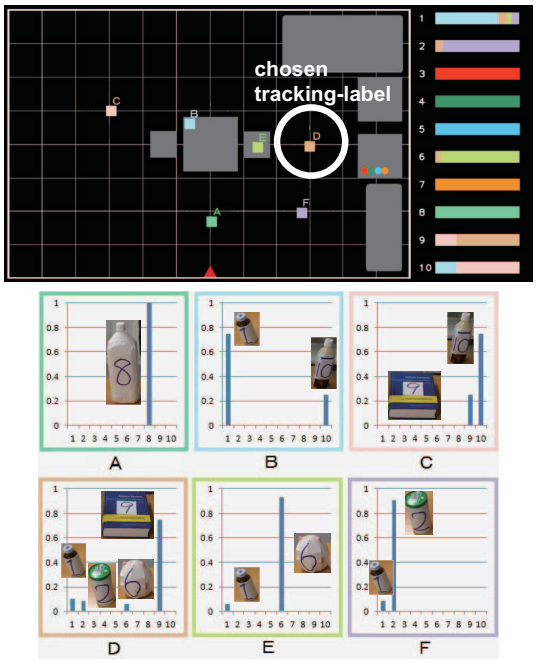



Fig. 17. Tracking result after the step 2 (Six objects are on the floor. Four objects are in the cabinet B.).

The proposed system tracks ten objects by using both the moving trajectories of objects and the size and the shape of an object placed on the floor. On the other hand, the previous system uses only the moving trajectories of objects. The estimation result of the previous system after the step 2 is shown in Fig.18. The accuracy rate of the position estimation of each object on the floor is under 20 %. This means that the moving trajectory information is not enough to determine which object is placed on the floor. So the previous system chooses one of them with even probability. This simple choice causes the decrease of the estimation accuracy.

At the step 3, the active sensing is performed to improve the accuracy rate of the position estimation. The entropy of the probability distribution of the object placement is calculated from Eq.(13). The calculation result of the entropy in the experiment is shown in Fig.19. The entropy increases when sensored information is not enough to estimate the object placement. When the entropy increases, the accuracy rate of the position estimation decreases. The entropy increases due to the handling of objects by the person in the step 2. For this active sensing, the most prospective tracking-label $D$ is chosen by applying Eq.(14). The tag ID of an object

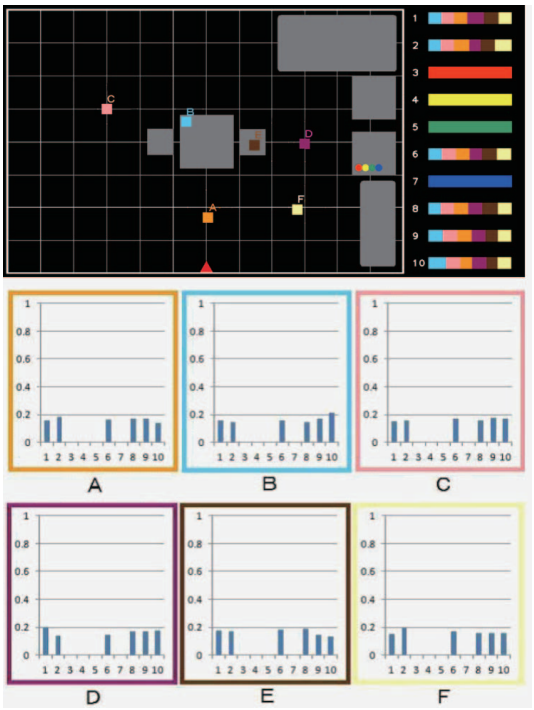



Fig. 18. Tracking result by using the previous system after the step 2 (Six objects are on the floor. Four objects are in the cabinet B.).

at the position with tracking-label $D$ was measured through the active sensing. The estimation result after the step 3 is shown in Fig.20. The accuracy rate of the position estimation of each object is nearly 100 %. The accuracy rate is clearly improved through the active sensing. The entropy in the case using additional information such as the size and the shape of an object is smaller than the entropy in the case not using it. The proposed system conducted the active sensing once at the step 3. On the other hand, the previous system required to conduct the active sensing four times in order to reduce the entropy enoughly.

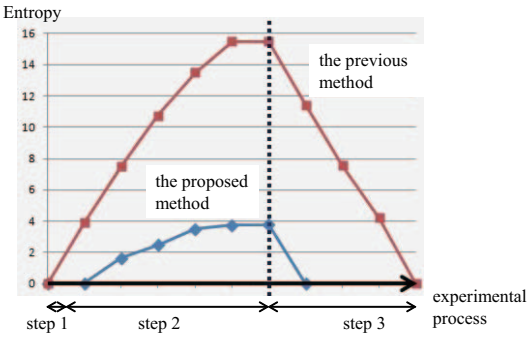



Fig. 19. The calculation result of the entropy.

## VI. Conclusion

We propose an object tracking system which recognizes everyday objects and estimates their positions by using distributed sensors in a room and mobile robots. Our previous system estimates the object placement by using only the moving trajectories of objects. This estimation result is expressed

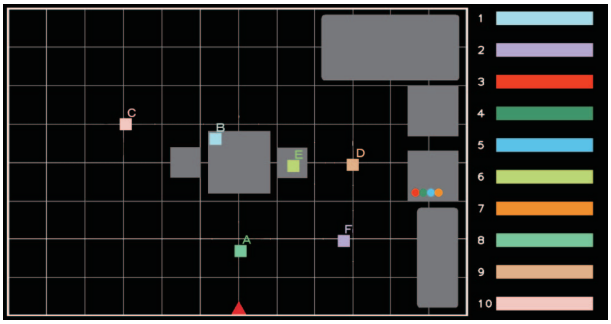


Fig. 20. Tracking result after the step 3.

as the probability distribution of the object placement. The object placement is frequently changed according to human activities. Intersections of trajectories cause the decrease of the estimation accuracy. The object placement is often not uniquely determined due to the deficiency of measured data. So we proposed a new method that improves the estimation accuracy by using the size and the shape of an object as additional information in the data integration process based on Bayesian inference. The size and the shape of an object are approximately calculated from a point cloud measured by laser range finder. Although the tag IDs of the objects on the floor are not measurable by sensors distributed in the environment, the system successfully estimated the object placement in the experiment.

## References


[1] K. Finkenzeller, *RFID handbook: fundamentals and applications in contactless smart cards and identification*, Wiley, 2003.

[2] R. Fukui, et al., "Development of a Home-use Automated Container Storage/Retrieval System", *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp.2875-2882, 2008.

[3] T. Deyle, et al., "RFID-Guided Robots for Pervasive Automation", *IEEE Pervasive Computing*, Vol.9, No.2, pp.37-45, 2010.

[4] Y. Nishida, et al., "3D Ultrasonic Tagging System for Observing Human Activity", *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp.785-791, 2003.

[5] R. Fukui, et al., "Expression Method of Human Locomotion Records for Path Planning and Control of Human-symbiotic Robot System based on Spacial Existence Probability Model of Humans," *IEEE Int. Conf. Robotics and Automation*, pp.4178-4184, 2003.

[6] S. Odashima, et al., "Household Object Management via Integration of Object Movement Detection from Multiple Cameras", *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp.3187-3194, 2010

[7] Murakami et al. , "Position Tracking and Recognition of Everyday Objects by using Sensors Embedded in an Environment and Mounted on Mobile Robots ", *Int. Conf. on Robotics and Automation*, pp.2210-2216, 2012.

[8] S. Arulampalam, et al., "A tutorial on particle filters for on-line non-linear/non-gaussian bayesian tracking", *IEEE Trans. on Signal Processing*, vol.50, no.2, pp.174-188, 2002.

[9] Y. Nohara, et al., "Floor Sensing System Using Laser Range Finder and Mirror for Localizing Daily Life Commodities," *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp.1030-1035, 2010

[10] H. Ming-Kuei, "Visual pattern recognition by momentinvariants", Information Theory, IRE Transactions, vol.8, pp.179-187, 1962.