Qi WANG, Zhen FAN, Wei-hua SHENG, Sen-lin ZHANG, Mei-qin LIU, 2019. Finding misplaced items using a mobile robot in a smart home environment. *Frontiers of Information Technology & Electronic Engineering,* 20(8):1036-1048. https://doi.org/10.1631/FITEE.1800275

Finding misplaced items using a mobile robot in a smart home environment

Key words: Home service robot; Smart home; Heterogeneous sensors; Autonomous robot retrieval

Corresponding author: Mei-qin LIU
E-mail: liumeiqin@zju.edu.cn

Motivation

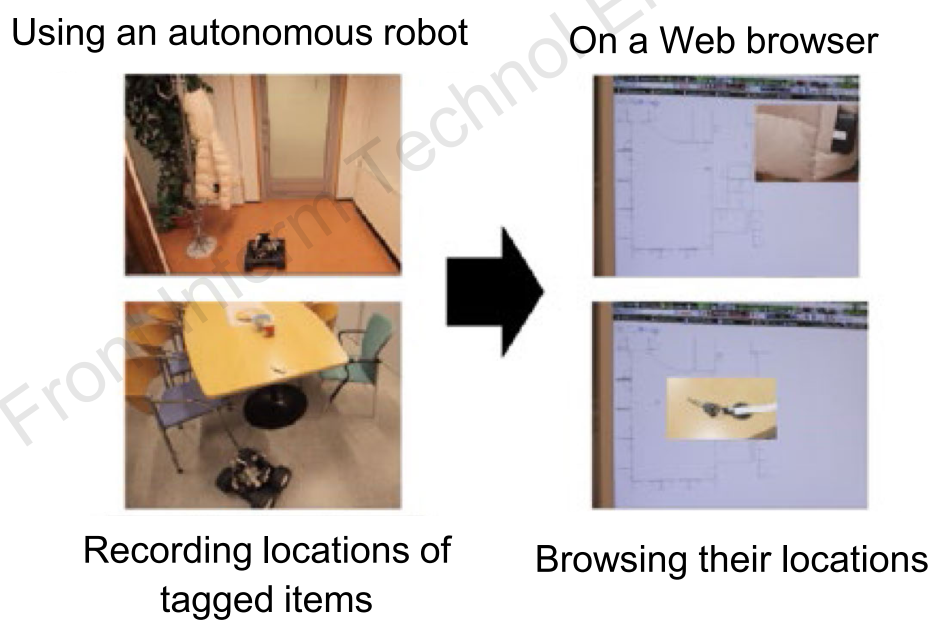
- Home service robots in smart home environments have significantly reduced our workload by undertaking some of the daily housework.
- The growth of elderly population has called for more home service robots for elderly care[1].
- For elderly care, the home service robots need to be capable of completing some more intelligent tasks.
- Finding misplaced items is a big headache for older adults[2].

Questions	Summary of Answers
Frequency of dropping	5.5 times per day(average)
Frequently dropped items	Phone/cellphone (4) Magazines or newspaper (3) TV remote (3) Pills, fork, pens (2) each
Most important to retrieve	Phone or cellphone (2) Walking cane (2) Key, pencil, fork(1 each)

[1]**Broekens J, Heerink M, Rosendal H, 2009. Assistive social robots in elderly care: a review. *Gerontechnology*, 8(2):94-103.**
[2]**Meng Q, Lee MH, 2006. Design issues for assistive robotics for the elderly. *Advanced Engineering Informatics*, 20(2):171-186.**

Motivation (Cont'd)

- An autonomous robot retrieval system is highly desirable to search for misplaced items in home environments[3].
- Traditional misplaced item finding approaches use wireless tags: IteMinder[4].



---

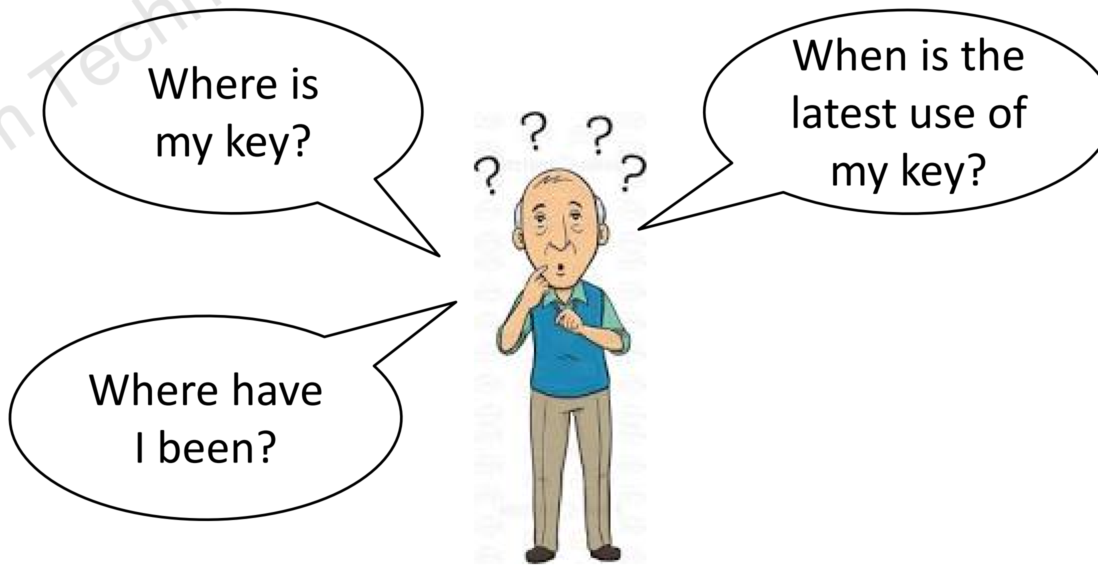
[3] **Ahern SJ, Carter J, Wilson P, 2015. Autonomous robot retrieval system. SAI Intelligent Systems Conf (IntelliSys), p.280-282.**
[4] **Komatsuzaki M, Tsukada K, Siio I, et al., 2011. IteMinder: finding items in a room using passive RFID tags and an autonomous robot (poster). Proc 13th Int Conf on Ubiquitous Computing, p.599-600.**

Motivation (Cont'd)

- Drawbacks of IteMinder:
  - Limited life time of tags
  - A full-environment search
- Alternative approaches:
  - Using visual search instead of tags
  - Introducing some complementary information to guide the robot's search movement

Motivation (Cont'd)

- The location of a daily item and the human movement are highly correlated.
- Imagine a scenario in which an older adult has misplaced his key …



Main idea

- We propose a new solution to search for misplaced items which relies on the historical information of human movement provided by the smart home environment.

- Our solution consists of the following three parts:
  1. A multi-sensor fusion method is developed to localize and track a resident.
  2. A path planning method is developed to generate the robot movement plan, which considers the knowledge of the human historical trajectory.
  3. A real-time object detector based on the convolutional neural network (CNN) is applied to detect the misplaced item.

Main idea (Cont'd)

- The general idea of our research:

  Smart home environments can provide the much needed context information for home service robots to better serve people.

- This general idea can be applied to other service tasks:
  - Human activity recognition
  - Advanced vacuum cleaning[5]

---

[5]Scassellati B, Tsui KM. Co-robots: the synergy of humans and robots operating as partners.

Main idea (Cont'd)

- Our proposed approach is to use the human historical trajectory data from the smart home environment to support misplaced item finding (MIF).
- **Conceptual design of the robotic MIF system:**
  - **A mobile service robot**
  - **A human localization sensor network**
  - **A user interface**
  - **A robot path planner**

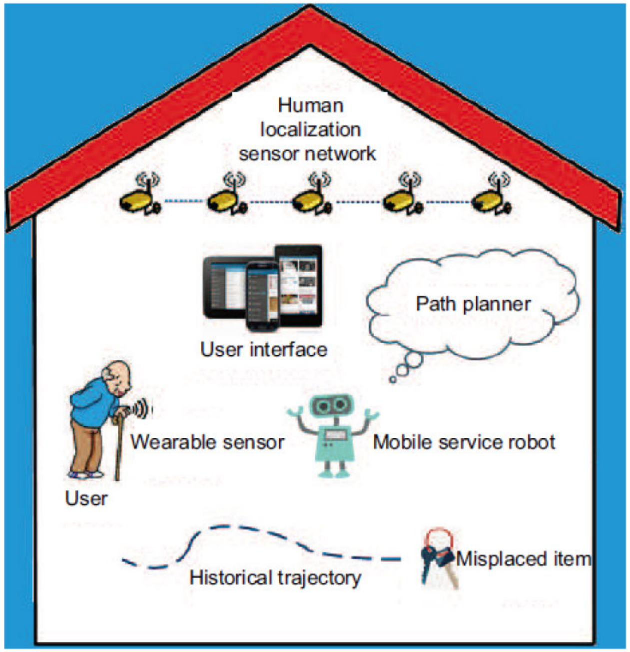


Fig. 1 Concept of the robotic misplaced item finding (MIF) system

Main idea (Cont'd)

The flowchart of our robotic MIF procedure

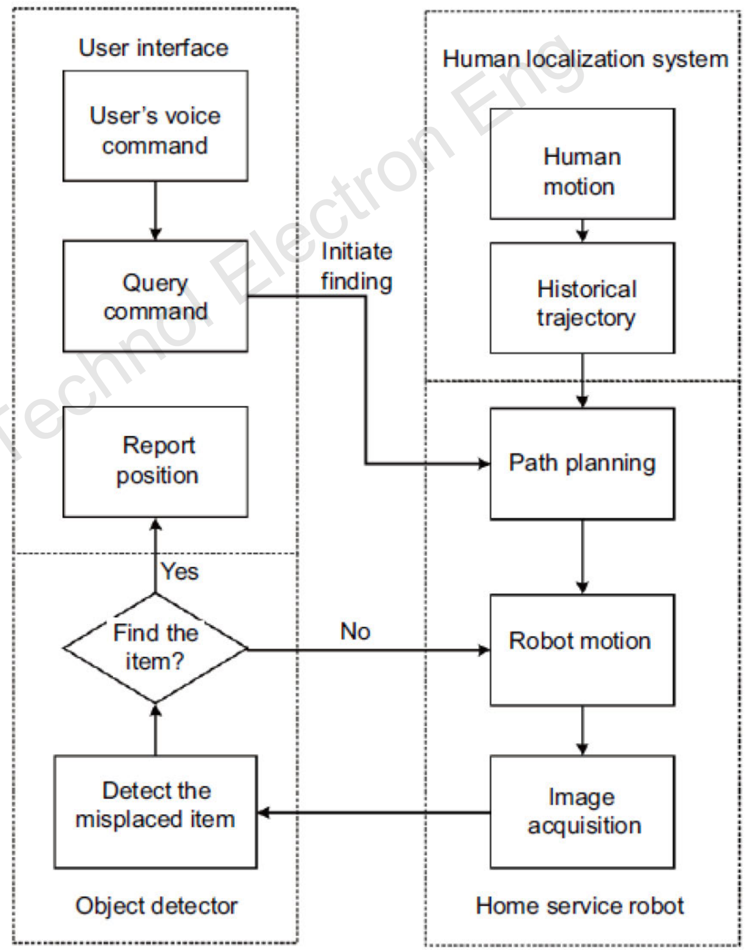


Fig. 2 Misplaced item finding (MIF) procedure

Method

- The methodology of our proposed robotic MIF system includes:
  - **Human location estimation**
    - Motion model
    - Human location based on PIR observations
    - Sensor fusion for localization
  - **Search path planning**
    - Map quantization and region partitioning
    - Integrating knowledge and path planning
  - **Vision-based object recognition**

Method (Cont'd)

- **Human location estimation**
  - **Human motion model**: provided by a wearable sensor node
  - **External observations**: provided by the PIR sensor network
  - **Particle filter based fusion framework**: obtaining an accurate human location estimate

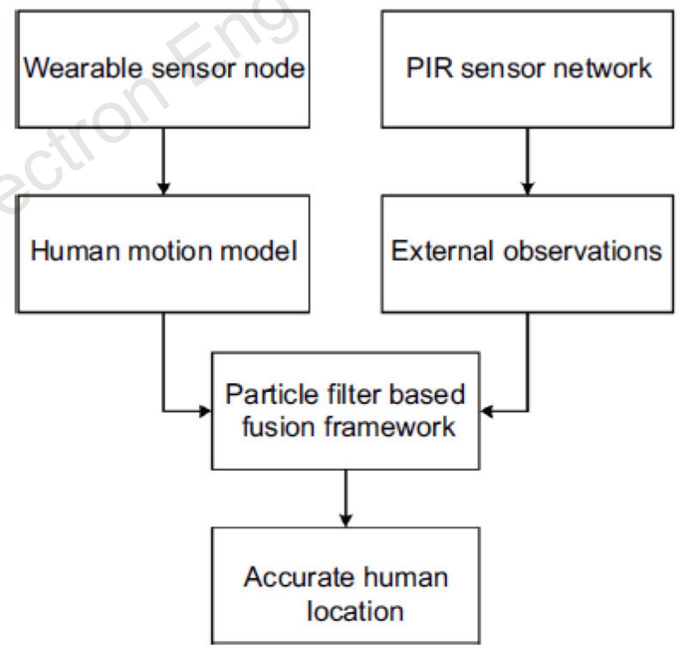


Fig. 8 Location estimation process

Method (Cont'd)

- **Motion model building**
  - Coarse human location can be estimated through the measurement of the movement distance and heading.
  - The movement distance can be calculated by counting the user's walking steps.
  - The walking step detection can be considered an acceleration detection problem.

  We use:

  X-axis acceleration a_x,

  Y-axis acceleration a_y,

  and the acceleration magnitude: $$a_v = \sqrt{(a_x)^2 + (a_y)^2 + (a_z)^2} \quad (1)$$

  from the wearable sensor as raw signals.

  Sliding windows are used to extract the features of these raw signals, each sliding window representing the data of 3.2 s, which is sufficient to capture the cycles in the activities like walking.

Method (Cont'd)

- **Motion model building**
  - We consider it a walking step if both the signal peak and signal valley are detected within a sliding window.
  - An adaptive-threshold-based algorithm is performed in the acceleration detection of each channel to adapt to different users and different walking modes:

$$T_{\mathrm{v}} = \frac{V_k + V_{k-1} - 1}{2} + \left(K - \frac{V_k + V_{k-1} - 1}{2}\right) C_1, \tag{2}$$

$$T_{\mathrm{p}} = T_{\mathrm{v}} + \sqrt{K - T_{\mathrm{v}}} C_2, \tag{3}$$

$$K = \begin{cases} P_k, & P_k < P_{k-1}, \\ P_{k-1}, & P_k \geq P_{k-1}. \end{cases} \tag{4}$$

Method (Cont'd)

- **Motion model building**
  - The heading angle change can be estimated by integrating the angular speed over time:

$$\theta_k = \theta_{k-1} + \Delta\theta, \tag{5}$$

$$\Delta\theta = \int_{t-1}^{t} \omega_t \mathrm{d}t, \tag{6}$$

  - Combining the walking steps with the heading angle, we can derive the human motion model:

$$\boldsymbol{L}_k = \begin{bmatrix} x_k \\ y_k \end{bmatrix} = \begin{bmatrix} x_{k-1} + d_k \cos\theta_k \\ y_{k-1} + d_k \sin\theta_k \end{bmatrix} + \boldsymbol{n}_k, \tag{7}$$

Method (Cont'd)

- **External observations**
  - The probability of a human being at location L_k given a new PIR observation is

$$P(z_k^{\mathrm{PIR}} = 1|\boldsymbol{L}_k^i) = \begin{cases} 1, & \boldsymbol{L}_k^i \in \text{PIR sensor's FoV}, \\ 0, & \boldsymbol{L}_k^i \notin \text{PIR sensor's FoV}, \end{cases} \tag{8}$$

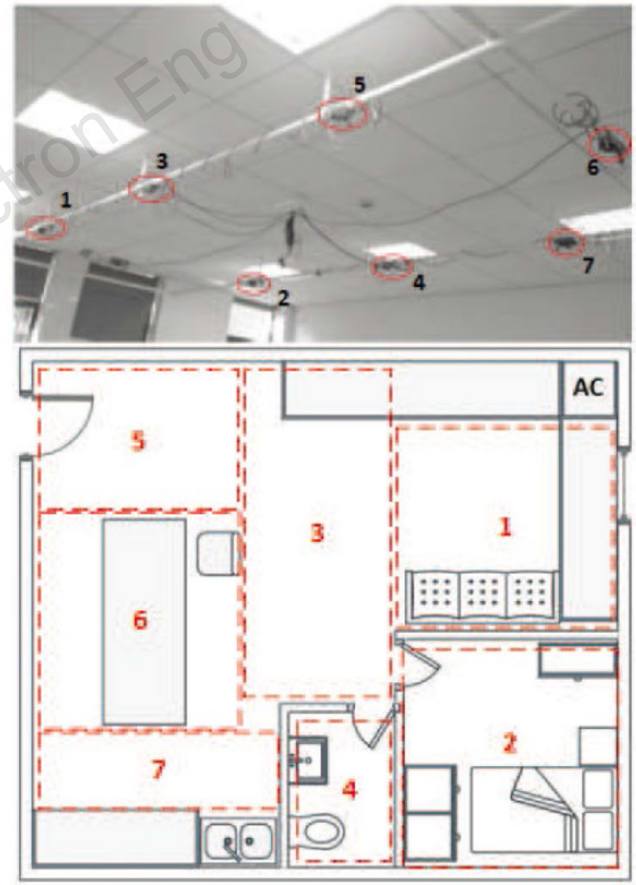


Fig. 9 Passive infrared (PIR) sensor network installation (top) and sensors' corresponding covering areas (bottom)

Method (Cont'd)

- **External observations**
  - The probability of a human being at location L_k given a recognized human activity type (walking) is

$$P(a_k|L_k^i), \qquad (9)$$

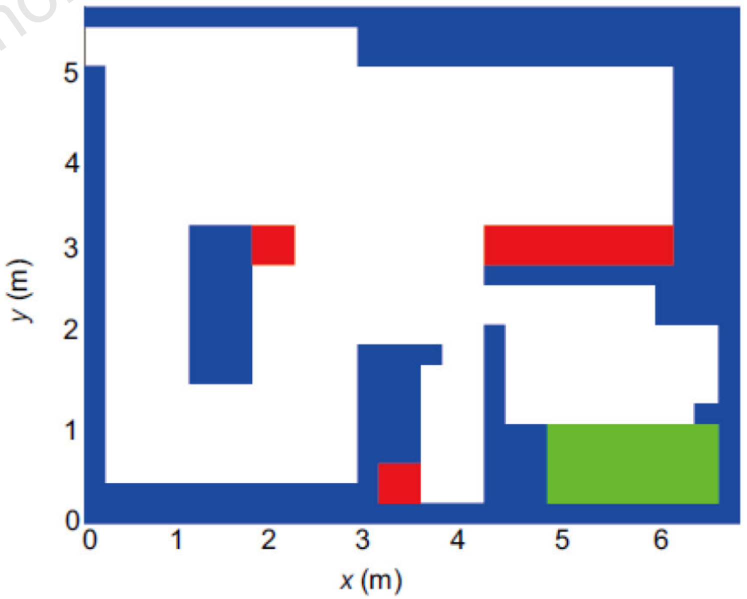


Fig. 10 Room layout (References to color refer to the online version of this figure)

Method (Cont'd)

- **Sensor fusion framework**
  - Possible human locations are represented by particles associated with likelihood weights:

$$w_k^i = w_{k-1}^i p(z_k^{\mathrm{PIR}} | \boldsymbol{L}_k^i) p(a_k | \boldsymbol{L}_k^i). \qquad (10)$$

  - We use the effective particle number N_{eff} as the trigger of a particle resampling procedure to avoid weight collapse in the filtering algorithm:

$$\omega_k^i = \omega_k^i \Big/ \sum_{i=1}^{P} \omega_k^i \quad N_{\mathrm{eff}} = 1 \Big/ \sum_{i=1}^{P} (\omega_k^i)^2.$$

  - The accurate human location can be estimated based on the mean of the particles' locations:

$$\tilde{\boldsymbol{L}}_k = E(\boldsymbol{L}_k) = \frac{\sum_{i=1}^{P} \boldsymbol{L}_k^i \omega_k^i}{\sum_{i=1}^{P} \omega_k^i}. \qquad (13)$$

Method (Cont'd)

- **Fusion algorithm for localization and tracking**

Algorithm 1 Fusion algorithm for localization and tracking

1: Initialize P particles with location vector $\boldsymbol{L}_k^i$, heading θ_k^i, and weight $\omega_k^i = 1/P$ $(i = 1, 2, \ldots, P)$

2: Recognize human activity a_k and read z_k^{PIR} from the PIR sensor network

3: **if** A walking step k is detected **then**

4: Estimate walking step length d_k and heading angle θ_k. Propagate the particles according to Eq. (7) // prediction step

5: Update the weights of the particles according to Eq. (10) // update step

6: **if** $N_{\text{eff}} < N_{\text{t}}$ (N_{eff} is calculated by Eq. (12) and N_{t} is the judgment threshold) **then**

7: Implement the resampling procedure

8: **end if**

9: Estimate human location according to Eq. (13)

10: **end if**

11: Go to step 2

Method (Cont'd)

- **Map quantization and region partitioning**
  - Converting the global map into a grid map
  - Dividing the grid map into regions

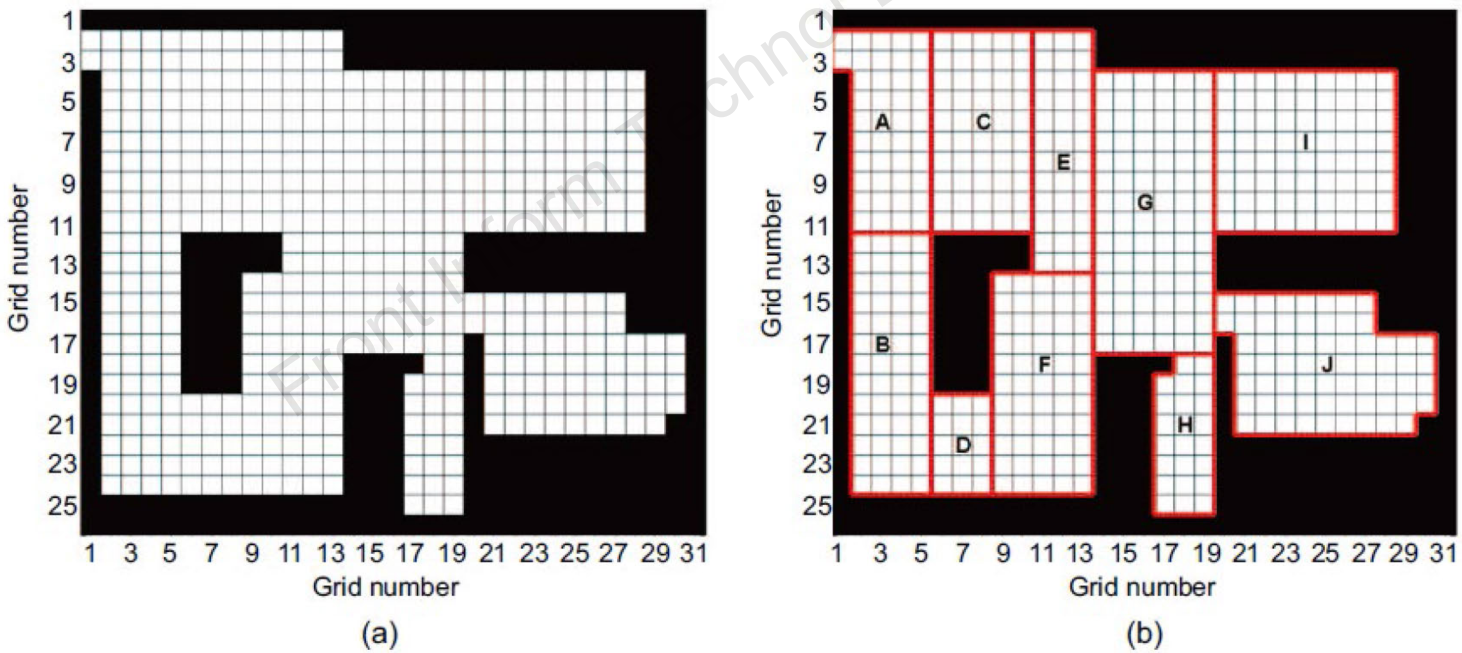


Fig. 13 Grid map (a) and region partition result (b)

Method (Cont'd)

- **Integrating knowledge and path planning**
  - The genetic algorithm (GA) is adopted to generate an optimal region transition sequence.

    A region transition sequence C→A→J→E→H→G→I→F→B→D

    can be represented by a chromosome [2, 0, 9, 4, 7, 6, 8, 5, 1, 3].
  - A new fitness function is developed to solve the planning problem:

$$f = \sum_{n=1}^{N-1} \frac{D(s_{n-1}, s_n)/\upsilon + T(s_{n-1}, s_n)/\omega}{\rho_{s_n}}, \qquad (14)$$

The probability ρ_{s_n} is dependent on the knowledge of the human historical trajectory derived in human location estimation:

$$\rho_{s_n} = \frac{L_{s_n}}{\sum_{n=0}^{N-1} L_{s_n}}, \qquad (15)$$

Method (Cont'd)

- The proposed vision-based object recognition architecture for MIF includes two modules:
  - An image acquisition module which runs on the robot
  - An image processing module which runs on a workstation

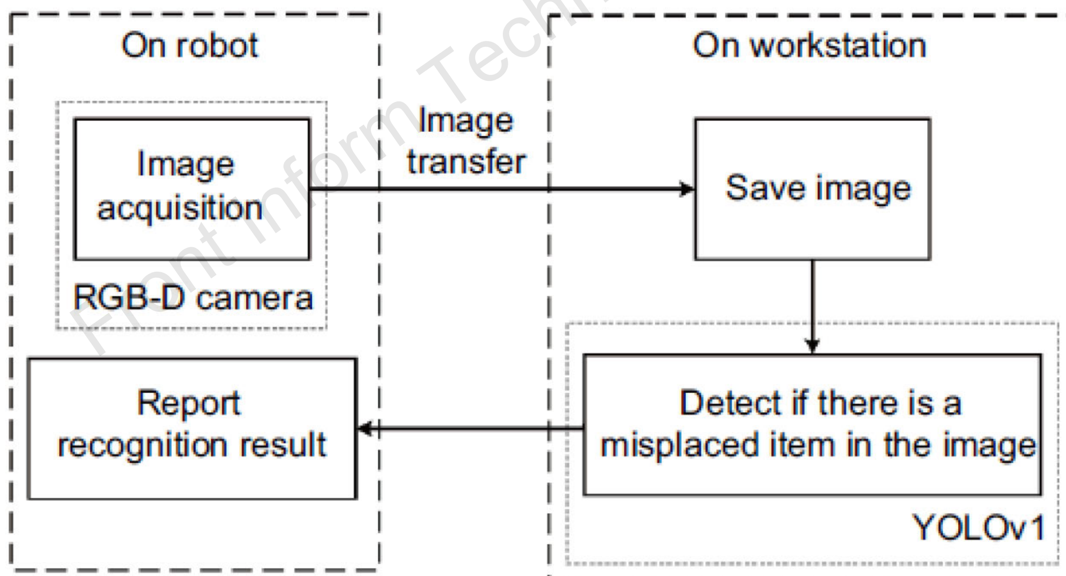


Fig. 14 Vision-based object recognition architecture

Method (Cont’d)

- Based on YOLOv1’s outputs, an evaluation flowchart is designed to detect the misplaced item within the scene images.

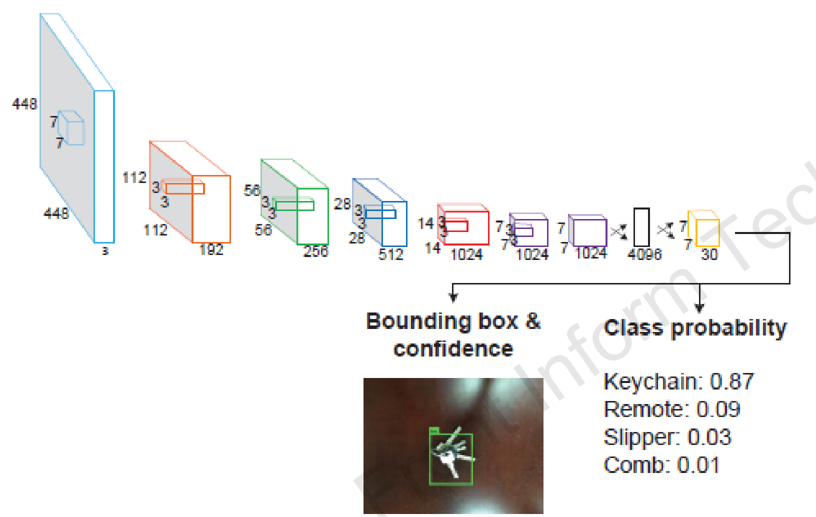


Fig. 15 Illustration of the CNN architecture (modified from Redmon et al. (2016))

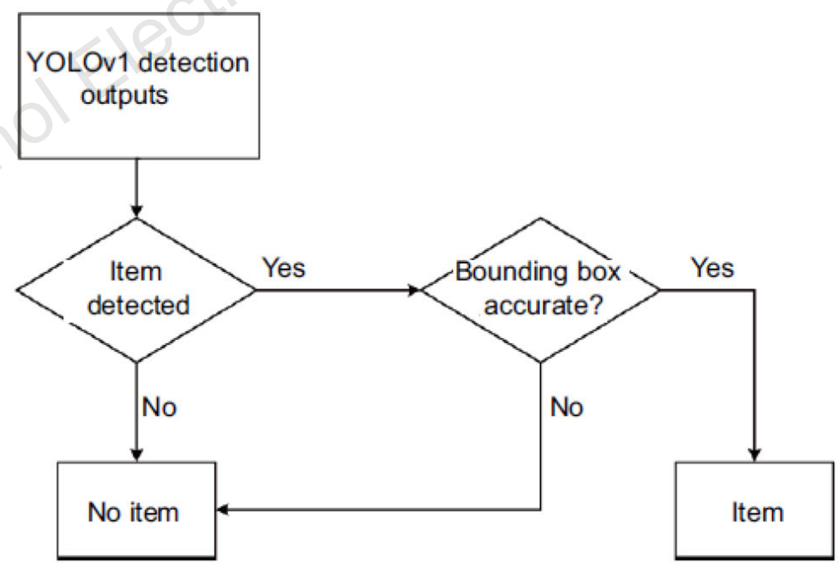


Fig. 16 Evaluation flowchart

Method (Cont'd)

- YOLOv1 is trained to detect several kinds of misplaced items, including keychains, remotes, slippers, and combs:

Fig. 17 Example of the keychain training sample

Fig. 18 Example of the remote training sample

Table 2 Sample number of the dataset for each category

Category	Number of samples			
	Keychain	Remote	Slipper	Comb
Training	867	828	638	759
Validation	186	178	137	163
Testing	186	178	137	163

The trained YOLOv1 correctly identified 90% of the sample images in the test set, and can realize real-time search with a 10 frames/s video input.

Major results

- **Experimental procedure**
  - The proposed robotic MIF system is evaluated in our smart home testbed.
  - The keychain is chosen as the misplaced item.
  - In three different scenarios, the keychain is lost in the bedroom, living room, and kitchen, respectively.

Fig. 19 Robot stops when the keychain is found

Major results (Cont'd)

- The search strategy of our robotic MIF system is compared with two benchmark methods:
  - Benchmark I: the robot search path is generated by a random region transition sequence.
  - Benchmark II: the robot search path is generated by a fixed region transition sequence.
  - Both benchmark methods are conducted without any complementary information.

Table 4 Improvement over the Benchmark I method

Area	Improvement (%)		
	T_{found}	Len	Ang
Bedroom	50.37	61.00	55.82
Living room	57.78	64.97	55.87
Kitchen	34.67	55.61	30.90
Average	47.61	60.53	47.53

Table 5 Improvement over the Benchmark II method

Area	Improvement (%)		
	T_{found}	Len	Ang
Bedroom	49.43	33.95	60.92
Living room	32.72	−3.65	41.94
Kitchen	68.52	67.53	68.53
Average	50.22	32.61	57.13

Major results (Cont'd)

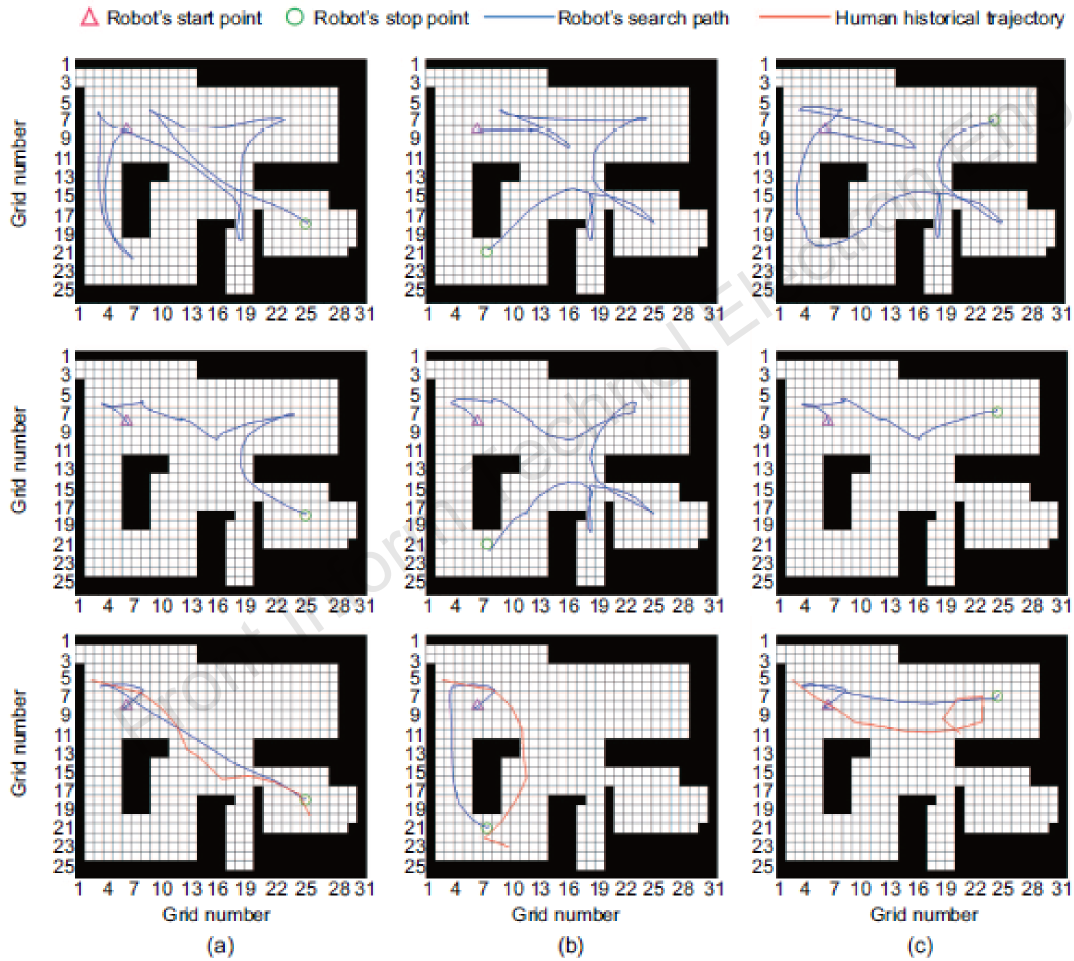


Fig. 20 Robot trajectories when the keychain was lost in the bedroom (a), kitchen (b), and living room (c). Rows 1–3 indicate the results of the Benchmark I, Benchmark II, and the proposed methods, respectively. References to color refer to the online version of this figure

Conclusions

- We developed a robotic MIF system that uses the complementary information provided by a smart home environment.
- We applied a data fusion method between a wearable sensor and multiple distributed movement detection sensors to estimate the historical trajectory of a resident. The calculation of the human trajectory can be carried out by the smart home environment, instead of the robot.
- We developed a robot path planning method, in which an optimal robot search path can be generated using the knowledge of the human historical trajectory data.
- Our work demonstrates that smart homes can provide the much needed context information for home service robots to better serve people.