Human motion detection and attribute extraction for service engineering

Ichiro Fukuda, Soichiro Morishita, and Hajime Asama
RACE, The University of Tokyo

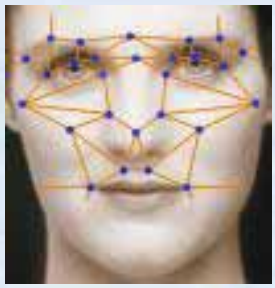
Introduction

- Many studies of intelligent environments have been conducted:
  - Robotic Room [Sato, 2003]
  - Intelligent Space [Hashimoto, 2002]
  - Smart Room [Pentland, 1996]
  - Aware Home [Kidd, 1999]
  - Intelligent Room [Brooks, 1997]
- These environments can achieve the position of a person
- For adaptive service provision to walking persons,
- Not only their position but also their attributes (age, sexuality, mother tongue and so on) are required

Photo by David Sim

Conceivable ways for personal identification

- **Identification with facial recognition**
  - Extremely difficult (not so good recognition rate),
  - Disinclined to be captured a face image



- **RFID**

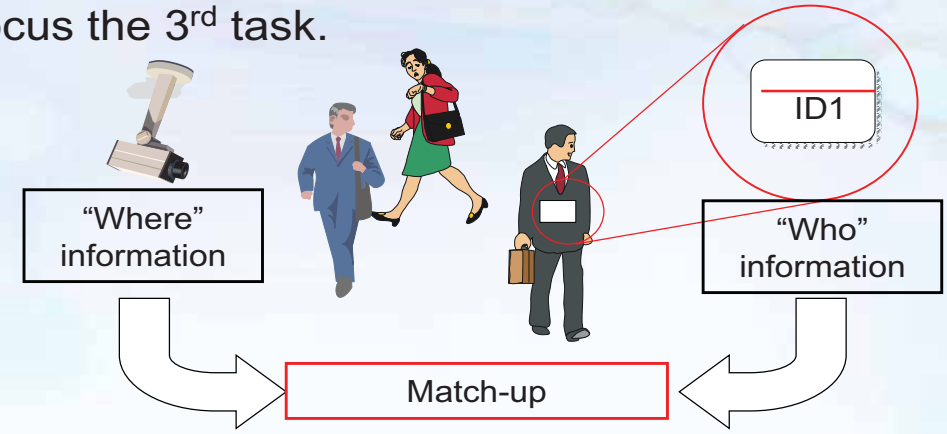
  It makes simplifies personal identification
  - **Active**
    - Internal power supply is needed
    - Difficult to maintain
  - **Passive**
    - No internal power supply
    - Easy to maintain

RFID system

- **Commonly used passive RFID system**
  - The effective range is too short
  - It imposes extraordinary operation tasks on people (such as holding RFID tag over an antenna)
- **UHF-band RFID system**
  - The effective range is reasonably long
  - It is possible merely by wearing an RFID tag around one's neck as a pendant
  - **Can not achieve the location**

Purpose of this study

- **Construction of an environment to acquire information for adaptive service provision for a walking person**
- For this discussion, information for adaptive service provision is
  1. A person's position or movement track ("Where" information)
  2. A person's attributes ("Who" information)
  3. **Match-up of "Who" and "Where".**
- We focus the 3rd task.



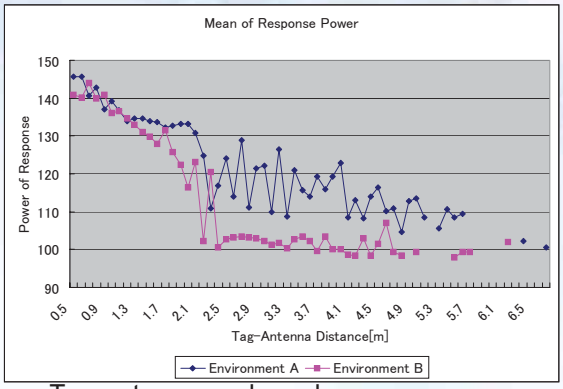
Analysis of UHF band RFID Attribute

UHF-band RFID system (RF-RW002; Mitsubishi Electric Corp.).

- The response power decays according to tag-antenna distance (Effective range: 3–8 [m])
- The response power is *not* constant according to the change of environment.
  - Permittivity of architectural material
  - The effect cannot be given without actual observation

Upper: UHF-band RFID tag
Lower: IC card (for comparison)

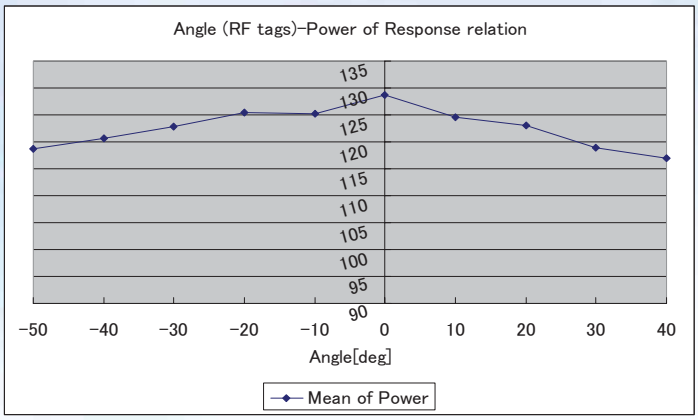
Antenna



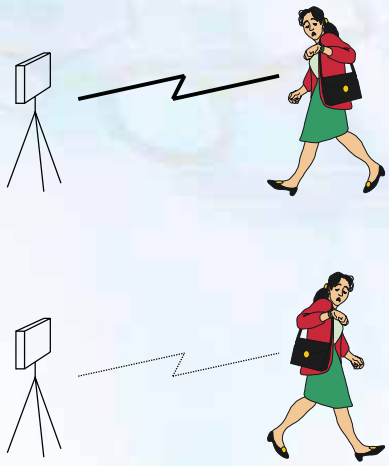
Tag-antenna angle and response power

Dependence property on tag-antenna angle

- The response power is affected by the tag-antenna angle
  - the response power is great-est when the angle of the tag and antenna are of a face-to-face position
  - Additionally, UHF radio wave is absorbed by water
  - In the case the person shows his/her back to antenna, the tag is not usually read

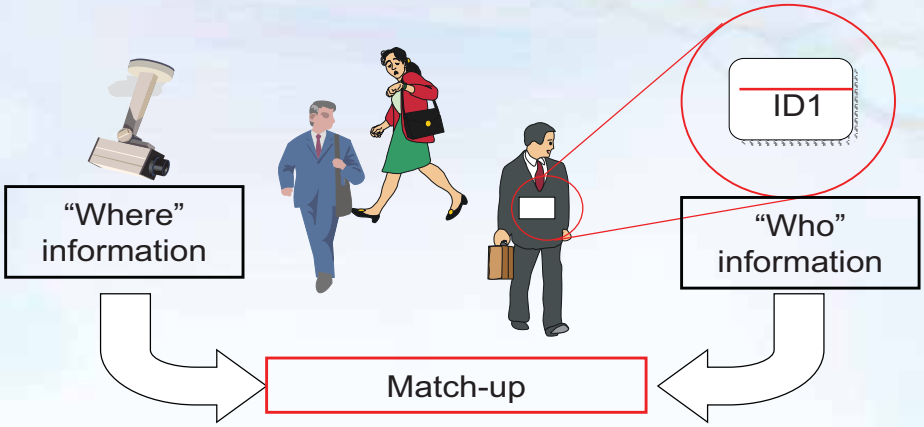


Tag-antenna distance and response power



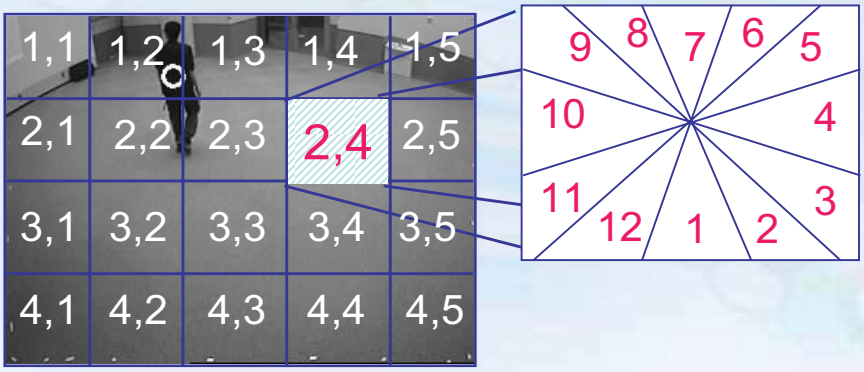
Our idea

- The tag reading data stream and moving object's location (track) are achieved separately
- Calculate **coincidence score** with probability based on actual observation for each track
- Decide correspondence of reading data and tracks
- These two pieces of information are matched up mutually



Definition of tag reading probability

- Divide the area into small rectangular areas
- Consider divided angle for each small area



x, y: the grid number (the image pixel barycenter of a moving object in the area)
θ: the angle of the moving object to RFID

Definition of tag reading probability

- A person walks around in the intended environment and we obtain tag reading data

$$\boldsymbol{r}_i = (r_1, r_2, ..., r_{T-1}, r_T), \quad r_{it} = \begin{cases} 0 \\ 1 \end{cases} \quad (t = 1,2,\ldots,T)$$

$$\boldsymbol{Q}_j = \begin{bmatrix} x_{j1} \cdots x_{jT} \\ y_{j1} \cdots y_{jT} \\ \theta_{j1} \cdots \theta_{jT} \end{bmatrix}$$

i: The index of RFID tag
j: The index of track
r_i: The reading data stream of RFID i
(0: tag was not red,
1: tag was red)
Q_j: The track

- Calculate the tag reading probability

$$p(x, y, \theta) = n_{xy\theta}$$

← The number of when $r = 1$

- Determining whether the moving track and tag reading data fit

$$S_{ij} = \sum_{t=1}^{T} \{ r_{it} p(x_{jt}, y_{jt}, \theta_{jt}) + (1 - r_{it})(1 - p(x_{jt}, y_{jt}, \theta_{jt})) \}$$

Readable ($`.` r = 1$)	Not readable ($`.` r = 0$)

- The score is considered a high value when track and reading data have good fit

Place: General Research Building 1F, Kashiwa Campus, The University of Tokyo

The tag reading probability is calculated with about 9,000 times readings

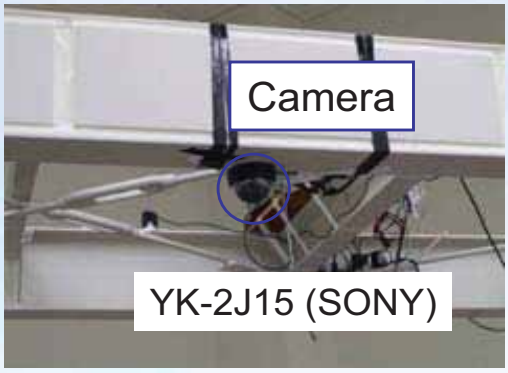
Antenna
(Height: 1.1[m])

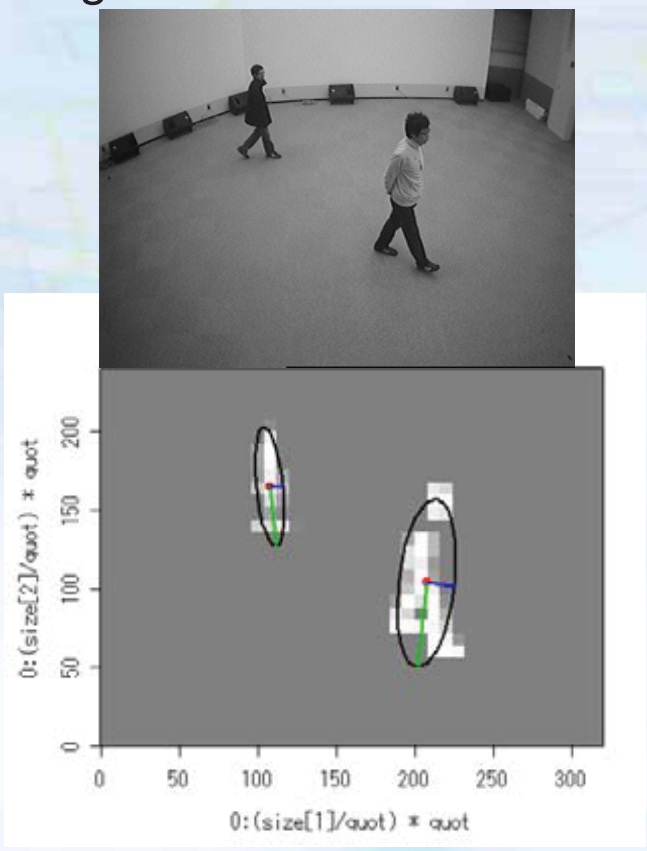
RFID R/W RF-RW002
(Mitsubishi Electric)

56x258x203[mm]

Read rate: 10[Hz]

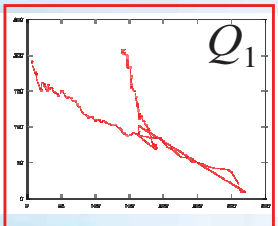
- Moving object detection with background subtraction
- Frame rate: 15[Hz]
- Image size: 320x240 [pixel]
- The number of small area:
  4x5 = 20



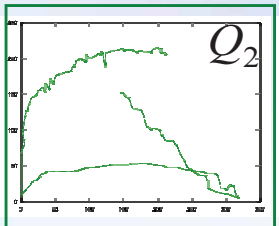


Experiment 1: with plural trajectories

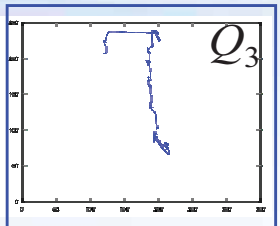
- Only a person exists in intended environment
- Five pattern tracks (each track has 200 frames of data) are chosen randomly.
- We calculate Score using them and their corresponding tag reading data



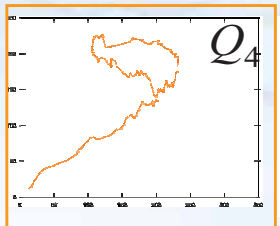
$r_1 = (0, 0, 1, 0, \ldots, r_{1T})$



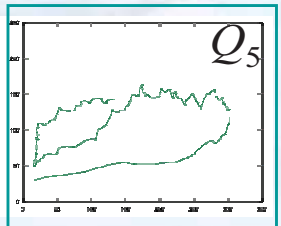
$r_2 = (0, 1, 1, 0, \ldots, r_{2T})$



$r_3 = (1, 0, 0, 0, \ldots, r_{3T})$

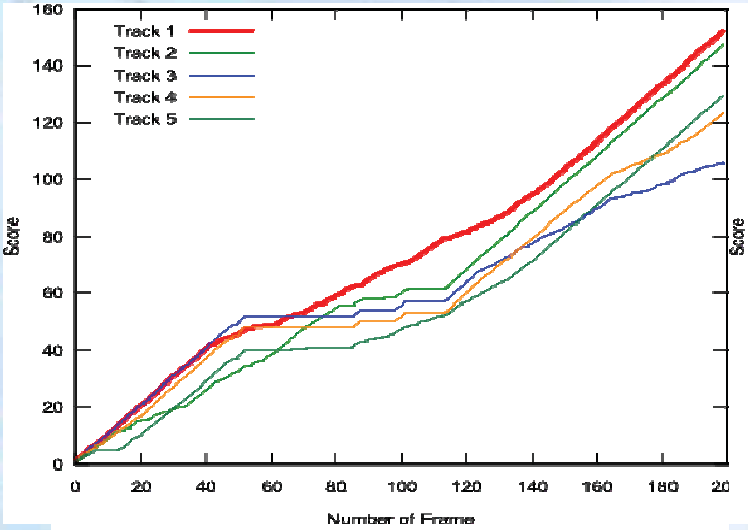


$r_4 = (0, 1, 1, 1, \ldots, r_{4T})$

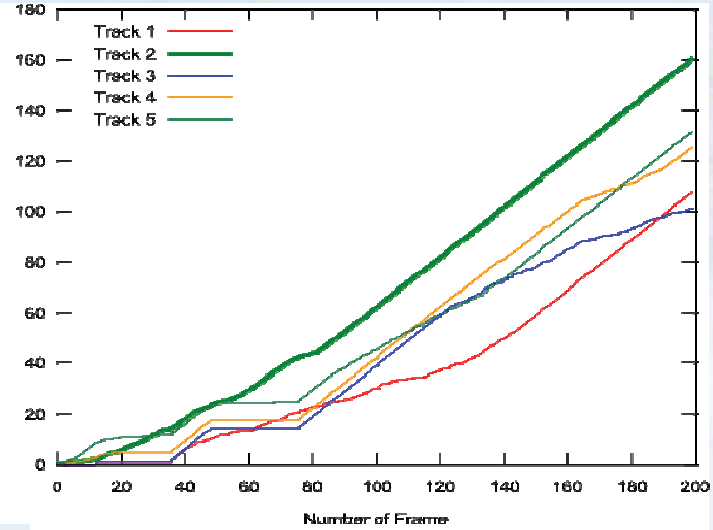


$r_5 = (1, 1, 0, 0, \ldots, r_{5T})$

Ex. 1: Transition of scores corresponding to r_1 and r_2



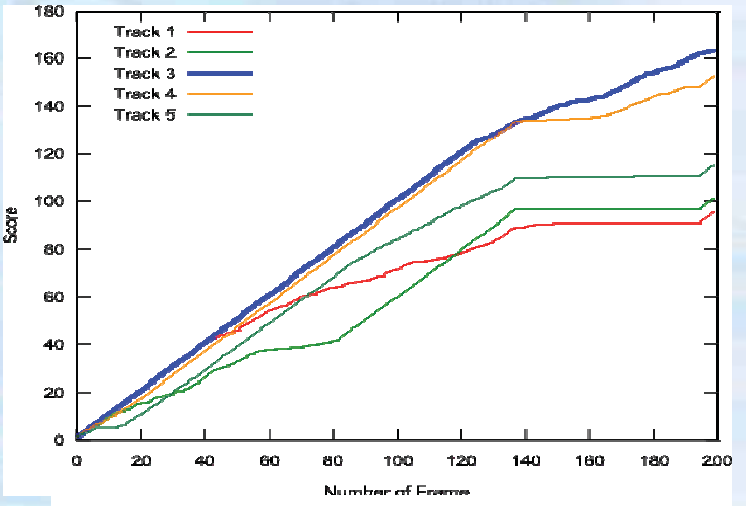
Score of each track corresponding to r_1



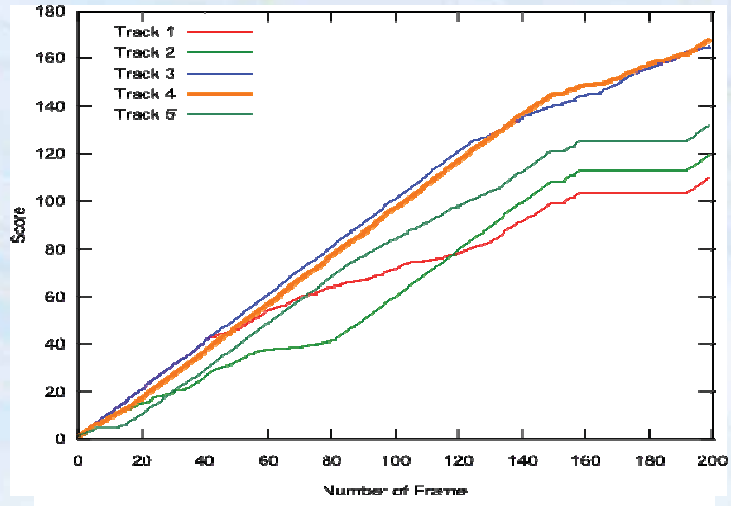
Score of each track corresponding to r_2

- For correct pairs, the largest score values are achieved

Ex. 1: Transition of scores corresponding to r_3 and r_4



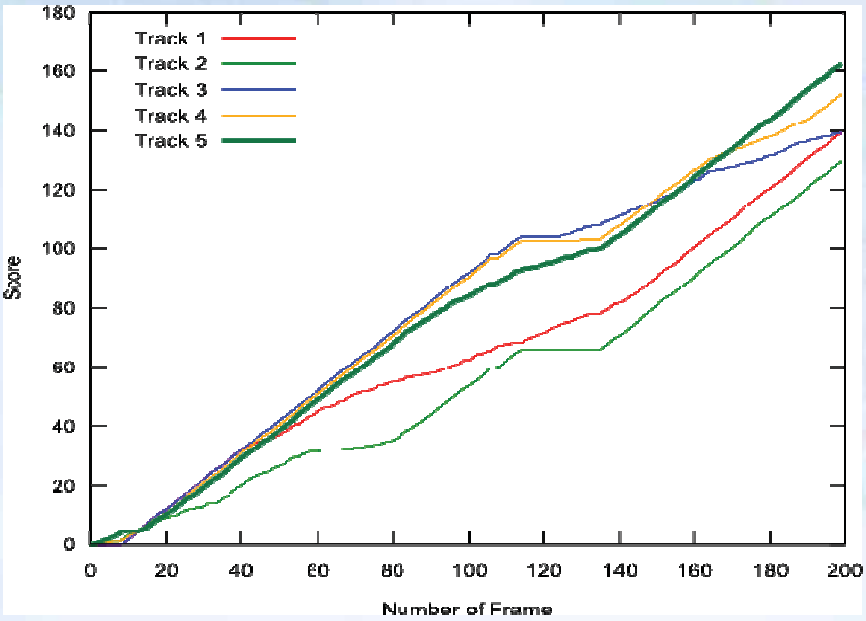
Score of each track corresponding to r_3



Score of each track corresponding to r_4

- Finally, the largest score values are achieved for correct pairs
- It is hard to distinguish these tracks.

Ex. 1: Transition of scores corresponding to r_5

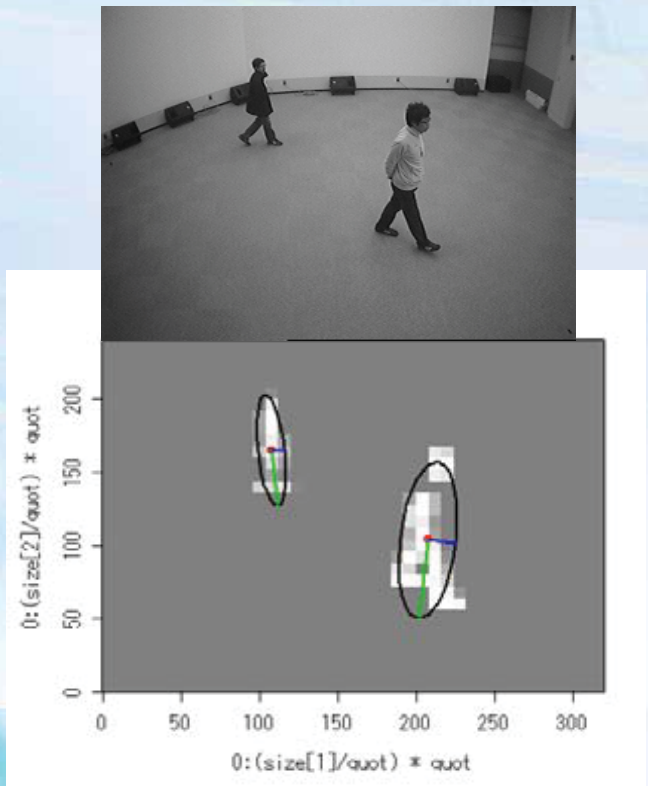


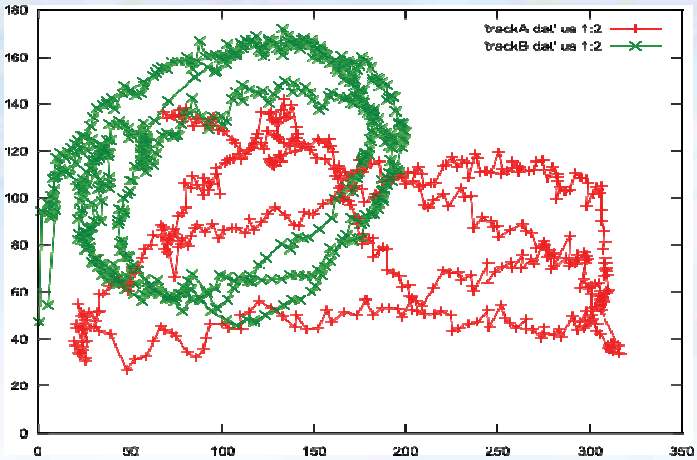
Score of each track corresponding to r_5

- Finally, the largest score values are achieved for correct pairs
- Almost always, the score is lower than others

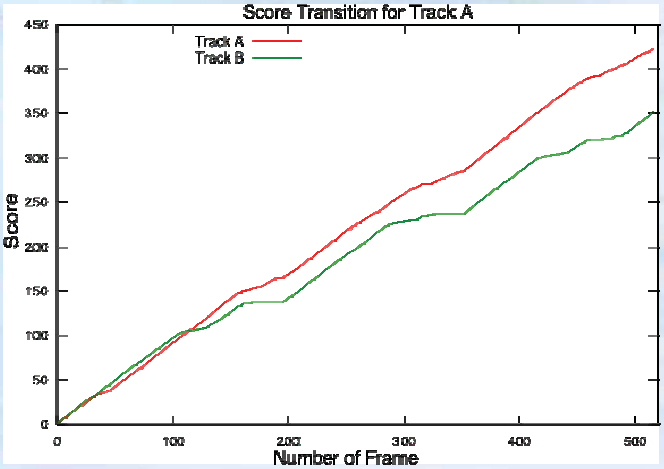
Experiment 2: Complex situation

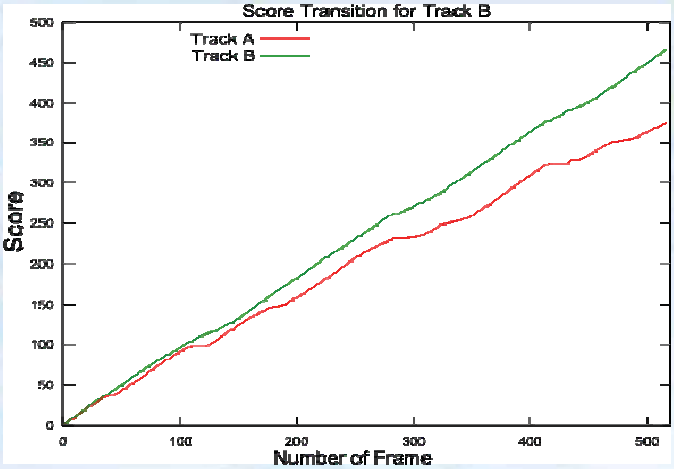
- 2 persons exists in intended environment simultaneously





Ex. 2: Transition of scores corresponding to r_A and r_B

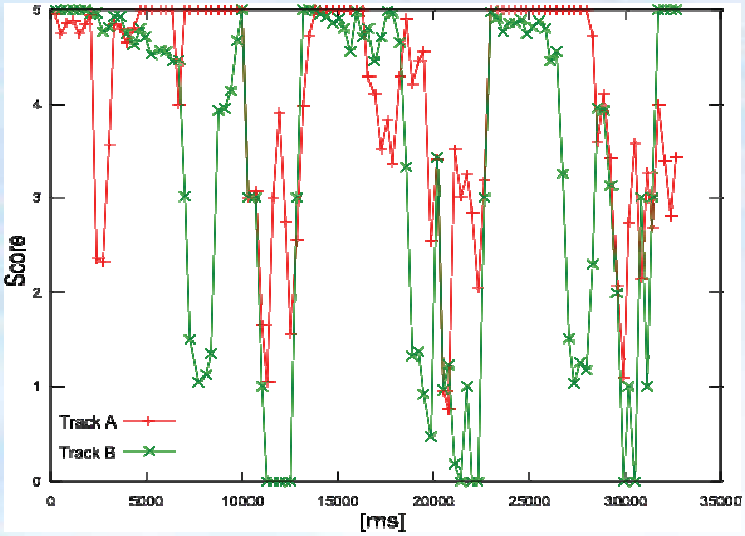




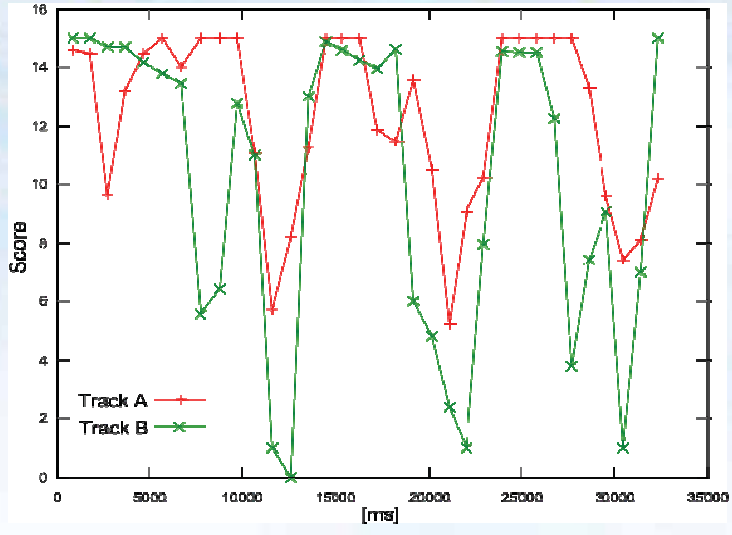
- Left: The score of correct pair is larger than another one after about the 120th frame
- Right: The score of correct pair is larger than another one after about the 20th frame.
- Track data and tag reading data are correctly matched for each pairs

Experiment 3: Transition of short length score

- The scores we have mentioned above are accumulate one
  - not useful for real-time matching
- Calculate the score per short time (5– [frames])
- This make it possible to do real-time matching



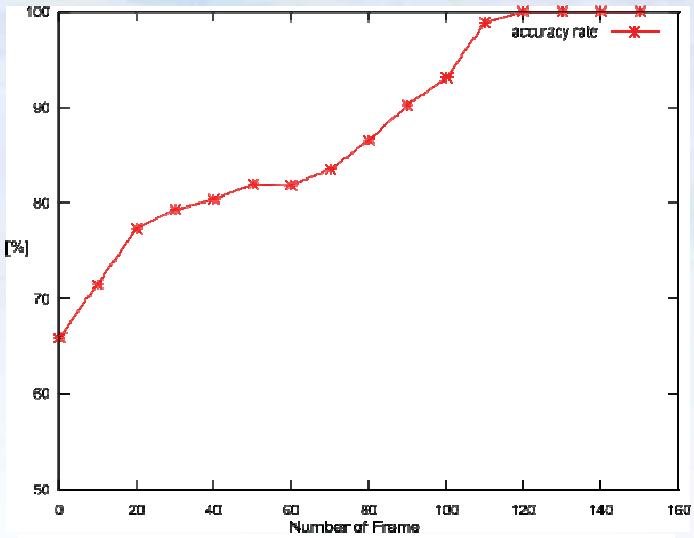
_Short length Score transition using data of Track A for 5 [frames]



Short length Score transition using data of Track A for 15 [frames]

Ex. 3: Discrimination with short length score

- The accuracy rate transition graph when the interval frame length is changed
- A higher accuracy rate when the length for a longer interval
- The accuracy rate of track and tag reading information match-up and length of interval show a tradeoff relation.



Transition of accuracy rate with r_1

Conclusion

- For adaptive service provision, we proposed this method to obtain “Who” and “Where” information simultaneously.
  - “Who” information is obtained using the UHF-band RFID system;
  - “Where” information is obtained by image processing to use background differences.
- We proposed “Score” to match up “Who” and “Where” information.
- We used experiments to analyze Score’s attributes and whether we can match up these two pieces of information using Score.
- We proposed the short length Score for real-time processing of match-up information. Results of experiments show a link between the interval frame length and accuracy.
- Using our proposed method, it was possible to obtain “Who” and “Where” information simultaneously.