Observation of Human Activities in Intelligent Space Based on Spatial Memory

Mihoko Niitsuma[1,3], Reka Sumire Mitsui[2], Péter Korondi[2], Hideki Hashimoto[1]

[1] Institute of Industrial Science, the University of Tokyo
4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan
niitsuma@hlab.iis.u-tokyo.ac.jp

[2] Budapest University of Technology and Economics
Goldmann György tér 3, H-1111 Budapest, Hungary
korondi@elektro.get.bme.hu

[3] JSPS Research fellow

Abstract: This paper presents a classification of human activities based on usage history of the spatial memory system in order to estimate individual human's intended purpose of an environment. We named the intended purposes for a place 'space-human activity association'. Intended purposes for a place will be different depending on the purpose of activities, even if observed activities are almost the same. Therefore, estimation of space-human activity association is important for intelligent environments to design services provided for individual human according to his/her current situation. The spatial memory system enables human users to store computerized information into the real world by assigning a three-dimensional position to the information, and to retrieve the information by directly indicating the point using their own hands. In the spatial memory system, what associates computerized information with a three-dimensional position is called 'Spatial-Knowledge-Tag (SKT)'. The users actively create SKTs based on their need. As a result, arranged SKTs in a specified environment correspond to activity histories of the users. Also, users' intended purposes for the environment are reflected in the arrangement. Therefore, classification of arranged SKTs leads to classification of human activities, and accordingly estimation of intended purposes for places. In this paper, in order to investigate usefulness of the proposed classification method, SKTs that are arranged through actual activities in a real working environment are classified.

Keywords: Intelligent Space, human interface, Spatial Memory, Observation of human activities

1 Introduction

Recently, intelligent environments in which many networked sensors and actuators are installed have been widely studied [1]-[8]. In order to support and aid human activities, the intelligent environments observe humans by using distributed sensors, recognize human activities and provide various services (e.g. health care [6], cooking support [7], learning of foreign language [8] and so on). Many approaches to achieve appropriate selection of prepared services focus on successful recognition of human activities, especially their actions while doing specified activities, their positions, or objects which they are manipulating. In some situations, for example cooking, it might be possible to specify supported activities beforehand. In our daily life, however, actual activities are so wide. In addition, purposes of human activities are independent even if people are doing the same things. In fact, meanings of activities to one person are different according to his/her situation. For example, let us consider a situation when people read books. An activity such as reading books means sometimes business or work, but sometimes fun. For other examples, people sleep in a bedroom, people eat food at a dinning table. But, actually, sometimes people do exercise and read books in a bedroom, and people may play cards at a dinning table. Consequently, in order for an intelligent environment to provide various services for people in a wide range of situations, current intended purposes of the activities should be recognized. That's why we argue that such space-human activity association must exist and should be recognized.

In order to address the issue, this paper introduces extraction of space-human activity association by using the spatial memory system. The spatial memory system enables human users to store computerized information into the real world by assigning three-dimensional position to the information, and to retrieve the information by directly indicating the point using their own hands. In the spatial memory system, what associates computerized information with a three-dimensional position is called ‘Spatial-Knowledge-Tag (SKT)’. The users actively create SKTs based on their need. As a result, arranged SKTs in an environment correspond to activity histories of the users. Also, users' intended purposes for the environment are reflected in the arrangement. Therefore, classification of arranged SKTs leads to classification of human activities and estimation of intended purposes for places.

The rest of the paper is organized as follows: Section 2 describes current spatial memory system. We improved the system to utilize it for a wide variety of situations. In Section 3, we explain the association of space and human activities, and discuss the need to recognize such association by an intelligent environment. A method to extract human activties based on classification of arranged SKT is shown in Section 4. We investigate usefulness of the method through the experiment in a real environment. The last section concludes the paper.

2 Spatial Memory

The spatial memory enables humans to store computerized information such as digital files and commands into the real world by assigning three-dimensional position as the memory address. Humans can retrieve and store such information by directly indicating the point using their own body, e.g. user's hand and user's head. That's why a point on user's body is called a human indicator. Fig. 1 shows a schematic concept of the spatial memory. The spatial memory system has the following advantages to achieve intuitively and instantaneously access and store computerized information. First, the users are able to arrange computerized information at a suitable location using their own eyes and own body action. While working in the real world, human can obtain environmental information such as arrangements of equipment including desks, file cabinets and so on. Consequently, the users can arrange information in a much easier way and memorize the whereabouts by referring to such environmental information. Also their spatial cognition capabilities such as the motion sense, are utilized for memorizing the whereabouts, and will prompt the users to recall them. Second, the users are able to store computerized information and get access to stored information without disturbing their activities to search for them. The spatial memory adopts an indication of a human body as 'store' and 'access' operation methods in order to achieve an intuitive and instantaneous access method that anyone can apply. In Fig. 1, small squares represent virtual tags called Spatial-Knowledge-Tags (SKT), which associate stored information with a spatial location. Each SKT has three important parameters, namely stored computerized information, a three-dimensional position as a memory address, and a size of an accessible region. An accessible region is necessary for human to smoothly retrieve a SKT by using the human indicator, because he/she can not indicate the exact spatial location. The size of an accessible region should be determined according to accuracy of the human indicator [9] and the type of the human indicator. We improved the system to utilize it for a wide variety of situations.

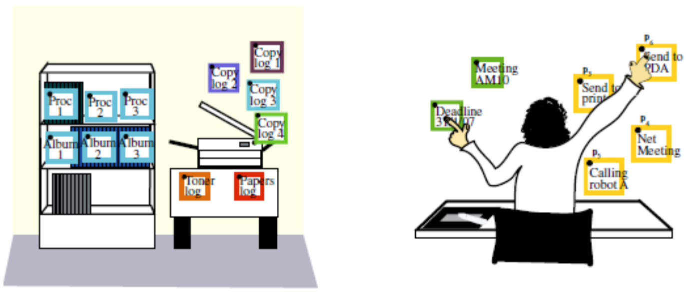


Figure 1
Schematic concept of the spatial memory

2.1 Implementation of the Spatial Memory System

The spatial memory system consists of four parts as follows: (a) Measurement unit of human indicator, (b) Spatial memory core unit, (c) Spatial memory input unit and (d) Spatial memory output unit, as shown in Fig. 2.

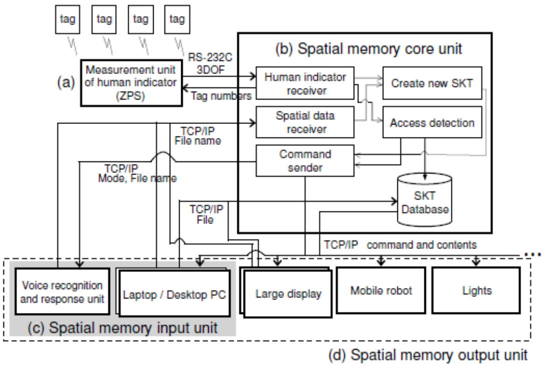


Figure 2
System configuration

2.1.1 Measurement Unit of Human Indicator

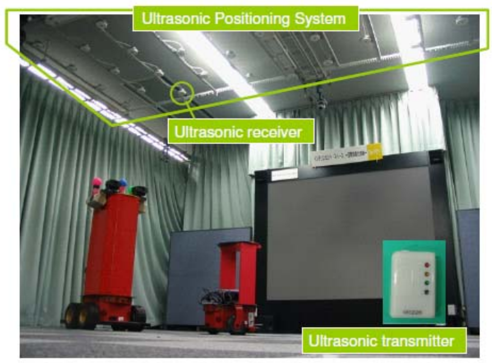


Figure 3
Ultrasonic 3D positioning system as the measurement unit of human action

Measurement unit of human indicator is implemented by using an ultrasonic 3D positioning system [10] as shown in Fig. 3. The positioning system is able to measure three-dimensional positions of ultrasonic transmitters. The transmitters are activated based on time division multiple access method. Therefore, sampling frequency depends on the number of transmitters. For applying to the spatial memory, sampling frequency is important to measure motion of human indicator.

From the view point of the measurement of human indicator motion, the number of transmitters is fixed to four. The position of transmitters is regarded as human indicator. Based on the observation of each transmitter, indication action of each user is obtained.

2.1.2 Spatial Memory Core Unit

The spatial memory core unit implements core functions such as determination of indication action, creating SKTs, retrieving and delivering SKTs. Therefore, this unit is the only unit connected with all other units. The spatial memory core unit allows multiple spatial memory input/output units to connect with it asynchronously. When receiving information from the spatial memory input units, the core unit assigns a unique ID number to the information and creates a new SKT. SKT data structure is shown in Table. 1.

Table 1
Data structure of a Spatial-Knowledge-Tag (SKT)

SKT data
ID
Spatial memory address (*x*,*y*,*z*)
Path of spatial memory (contents)
Size of accessible region
Number of access times
Type of contents
User name
Time stamp

When delivering an SKT to a spatial memory output unit, the spatial memory core unit chooses the nearest spatial memory output unit from the human indicator which is used to access the SKT. In addition to distance, functions of spatial memory output units should be considered, because different functions are included. For example, not only computers but also mobile robots are available as spatial memory output units. In order to discriminate functions of spatial memory output units, each output unit has a label to describe attribute identifier. The label is notified by the spatial memory output unit to the core unit. The core unit chooses a spatial memory output unit to deliver an SKT according to labels of output units, distance between a user and output units, and file extensions of information.

Fig. 4 shows principal messages between units and the sequence. The notations such as '(c1)' or '(c2)' represent the operation status and the order of processes. For example, when a spatial memory input unit sends information to the spatial memory core unit, firstly the input unit sends the message (s1) to the core unit. Then, the core unit assigns a unique ID number to the information and creates an

SKT. After that, the core unit delivers the message (s2) to a spatial memory output unit. The spatial memory output unit obtains the ID number and the information. The output unit executes the information and store the information into an SKT database by assigning the ID number as the process (s3).

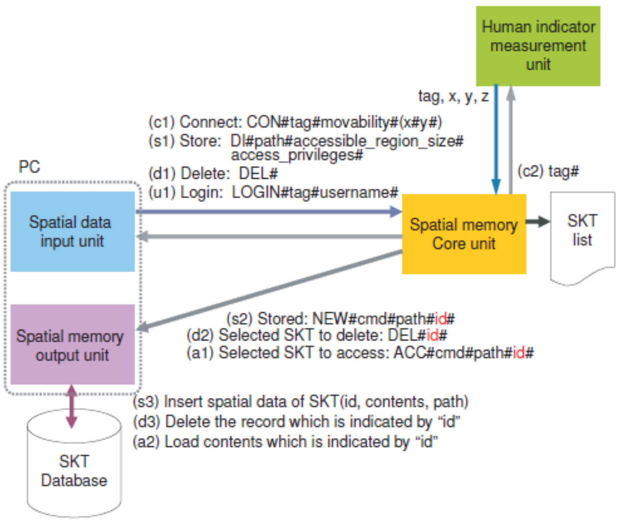


Figure 4
Connection and messages among all units

2.1.3 Spatial Memory Input Unit

Fig. 5 shows the interface of the spatial memory input unit. The interface plays a role of the front-end of the spatial memory system for users. In addition to the shown graphical interface, voice input method is also available. When starting to use the spatial memory, user identification is required. Users choose a login mode either as anonymous or as specified user. If a user choose specified user mode, the user inputs user's name into the text field as shown in Fig. 5 (u1) area. Then, the user arranges three password SKTs as his or her password. When the user is able to access all SKTs under specific conditions [11], the spatial memory system becomes available.

By using the interface, a user chooses or draws information which is stored as the spatial memory. After preparing the information, the user sets access privileges for the SKT. If the user sets access privileges as 'anonymous' for SKTs, the SKTs can be retrieved by any users. On the other hand, if the user sets access privileges as 'user' for SKTs, the SKTs can be retrieved only by the user who stored the information. Finally, the input unit sends the path of the information, accessible region size and access privileges to the core unit. Multiple input units are able to connect with the core unit asynchronously.

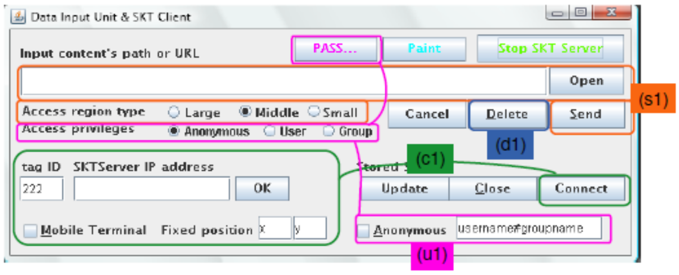


Figure 5
Spatial memory input unit

2.1.4 Spatial Memory Output Unit

Spatial memory output unit executes the information of accessed SKTs when receiving it from the core unit. Voice output units, personal computers and mobile robots are available for the output units. Also, the output unit operates an SKT database according to the messages from the core unit as shown in Fig. 4.

To share SKTs which are stored by different users and different input units, architecture of sharing SKTs is needed. Namely, an SKT database which gathers the stored information (contents of SKTs) is introduced. Contents of SKTs should be stored after assigned a unique ID number to an SKT. Therefore, a spatial memory input unit is not able to store contents into the database, because the contents have not been assigned ID.

3 Spatial Memory as Discriptions of Human Activities

Fig. 6 shows an arrangement of the spatial memory. A square represents an STK. Basically, an SKT is created by a user based on his policy and his purpose of activity in an environment. Therefore, each SKT includes both the location and the contents and describes human activity at that specific place. For example, if there are SKTs whose contents are video, it is obvious that the user would like to watch the video around there. Thus, the spatial memory can be regarded as an abstraction of human activities, and the arrangement of the spatial memory will contain human's intended purposes for places.

Usage history includes not only SKTs' information such as spatial memory addresses and the size of accessible regions but also the number of access times and time intervals between accesses. Such usage histories represent spatiotemporal histories of human's activities. Therefore, human activities can be observed by analyzing arranged SKTs.

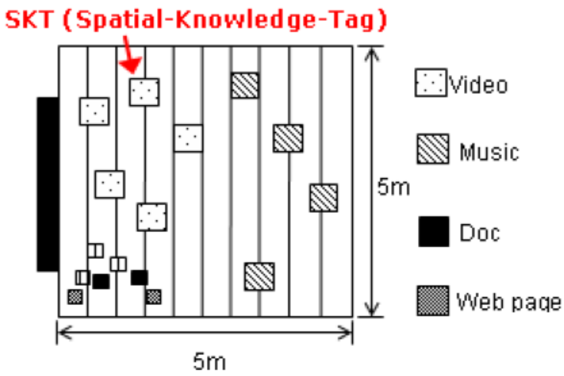


Figure 6
An arrangement of embedded Spatial-Knowledge-Tags (SKTs)

As a method of arranged SKTs analysis, we have presented a classification of arranged SKTs based on usage history of the spatial memory by using unsupervised hierarchical clustering method [13]. For the clustering, we need to define similarity between SKTs. In our previous research [14], we carried out clustering based on two types of similarities; one is defined by only using distance between SKTs, the other is defined by using distance between SKTs and usage histories of the spatial memory. As a result, we found that the latter similarity should be utilized to extract human activities. Therefore, in this paper we apply the similarity to SKT clustering.

4 Classification of Arranged SKTs Based on Usage Histories of the Spatial Memory

4.1 Definition of the Similarity between SKTs

Basically, SKTs used in the same activity must have high similarity. To determine similarity measure between SKTs based on actual human activities, we define ‘observation interval’, which is the time during which the user would be considered performing the same work. In other words, SKTs used in the same observation interval are regarded as related with the same activity, therefore they have high similarity. An observation interval is defined based on time duration of the spatial memory system use and distance of accessed SKTs as follows:

$$p = \begin{cases} p & \text{if } \Delta t < T \text{ and } d(\mathbf{x}_i, \mathbf{x}_j) < \gamma, \\ p+1 & \text{otherwise.} \end{cases} \tag{1}$$

As shown in (1), one of the conditions to determine an observation interval is the time interval of the spatial memory use. Another condition is the distance $d(\mathbf{x}_i, \mathbf{x}_j)$ between continuously accessed SKT $\mathbf{x}_i$ and $\mathbf{x}_j$. Let us consider that i-th SKT $\mathbf{x}_i$ is accessed at time t, and j-th SKT $\mathbf{x}_j$ is accessed at time $(t$-1). If the time duration while the user didn't use the spatial memory Δt (= t-(t-1)) is less than a design parameter T, and the distance $d(\mathbf{x}_i, \mathbf{x}_j)$ is shorter than a distance γ, the p-th observation interval will continue $(p = 0, 1, \cdots)$. If either of the conditions are not satisfied, user's activity would be changed, therefore a new observation duration (p + 1)-th starts.

Here, the condition regarding the distance $d(\mathbf{x}_i, \mathbf{x}_j)$ is explained. From the standpoint of a human activity in an environment, it will correlate well with area. Therefore, Euclidean norm between SKTs should be considered to describe nearness based on a position. However, the nearness scale differs by a selected human indicator; the distance of one meter will be far if a hand is used for the human indicator, while it will be near when body is used. The human indicator is selected based on the context in which an SKT is used. For example, when SKTs are used for a desk work, a hand will be selected to retrieve them. In addition, the human will select a small accessible region. Thus, the size of an accessible region also reflects contexts of human activities. For that reason, we take also the size of accessible regions into account. The distance $d(\mathbf{x}_i, \mathbf{x}_j)$ is given by

$$d(\mathbf{x}_i, \mathbf{x}_j) = \left\| \mathbf{p}_i - \mathbf{p}_j \right\| + \alpha \left| r_i - r_j \right|, \tag{2}$$

where $\mathbf{p}_i$ and r_i are the data set attributes of the i-th SKT $\mathbf{x}_i$ and represent the position and the size of an accessible region respectively. The design parameter α is the weighting factor for the difference of the size of an accessible region.

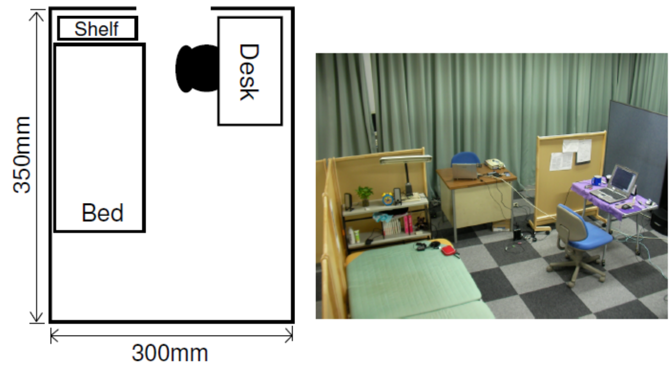
Based on an observation interval, from the observation of the spatial memory use we obtain the number of access times $N_i(p)$ of the i-th SKT for the p-th observation interval. The lists of the access times in each observation interval present usage histories of the spatial memory considered as histories of human activities. The distance $D_{obs}(\mathbf{x}_i, \mathbf{x}_j)$ between SKT $\mathbf{x}_i$ and $\mathbf{x}_j$ to express the closeness is obtained by using the lists of the access times. The distance $D_{obs}(\mathbf{x}_i, \mathbf{x}_j)$ is given by the inverse of the total number of access times $N_j(p)$ when SKT $\mathbf{x}_j$ is used with SKT $\mathbf{x}_i$ in each observation interval. Namely, $D_{obs}(\mathbf{x}_i, \mathbf{x}_j)$ is given by

$$D_{obs}(\mathbf{x}_i, \mathbf{x}_j) = \sum_p^m \sum_j^n N_j(p), \tag{3}$$

where m and n are the number of observation intervals and the number of SKTs, respectively.

4.2 Experimental Setting

Experiments of classification of SKTs are carried out by using the environmental setting as shown in Fig. 7. What is important here is that the classification can be obtained automatically – not directly specified by humans – but obtained through human activity observations. In the room Fig. 7 (a), we arranged a bed, a book shelf and a desk as shown in Fig. 7 (b). We didn't specify human activities, which included reading magazines, writing papers, arranging digital photography and so on. We focus on threshold level σ when obtaining correct classifications.



(a) (b)

Figure 7
Environmental setting

When performing hierarchical clustering method, inter-cluster distance is obtained by using the median linkage method since it is able to deal with any dissimilarity measures between SKTs. The method needs a threshold level σ to cut the clustering tree and to obtain clusters, because it is hard to determine how many activities are in the environment. Therefore, in the experiment, clustering of arranged SKTs is performed by varying the threshold level.

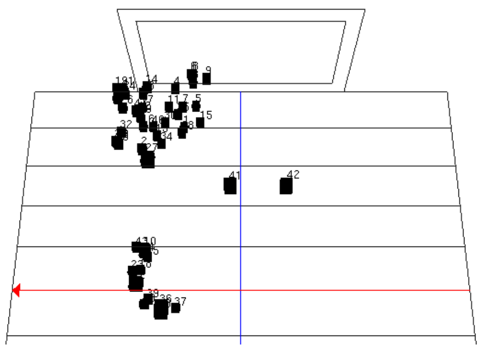
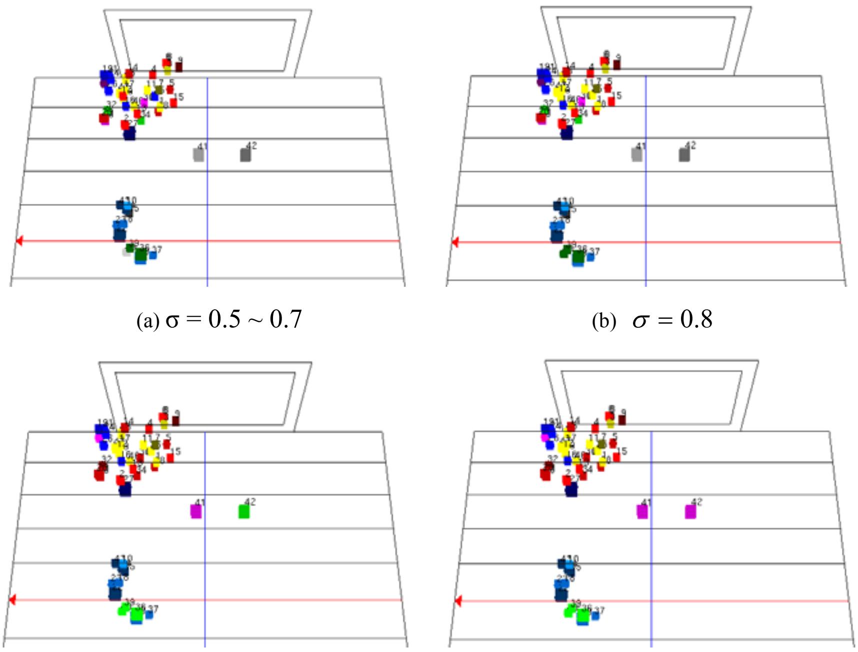


Figure 8
Arranged SKTs in two days

For the clustering, the design parameter in (1) and (2) are given by α =1.5, γ =1.3, T=30. Clustering is performed by varying the threshold level σ from 0.5 to 10.0 at 0.1 intervals. Fig. 8 shows arranged SKTs which were created by a user during more than 2 days. Similarities between SKTs are obtained by using the SKTs information and the usage histories.

4.3 Experimental Results

Fig. 9 shows the results of clustering. Filled square represents an SKT, and the size of square shows the size of the SKT. Each color represents each cluster, i.e. same colored SKTs mean same cluster. What is important here is that we can find that the user performs different activities in a same area. It is impossible to extract them by using only simple sensory information such as position and chair switches and so on.



(a) $\sigma = 0.5 \sim 0.7$ (b) $\sigma = 0.8$

(c) $\sigma = 0.9 \sim 1.0$ (d) $\sigma = 1.1 \sim 10.0$

Figure 9
Results of classification

Fig. 9 (d) is the correct result to describe actual human activities. The correct result was shown as the largest number of clustering result, even though the ranges of threshold levels were different. Therefore, we can obtain correct clusters of

SKTs by finding out the largest number of same results. However, we can not decide an exact value of the threshold level to obtain correct clusters, because it depends on arrangement of SKTs and usage histories of the SKTs.

Conclusions

In this paper, we discussed the significance of recognizing space-human activity association. We also presented the spatial memory system was suitable to observe human activities in the real world. Experiments of classification of the spatial memory to extract human activities in an environment were shown.

References

[1] J.-H. Lee, H. Hashimoto, "Intelligent Space - concept and contents", Advanced Robotics, Vol. 16, No. 3, pp. 265-280, 2002

[2] B. Brumitt, B. Meyers, J. Krumm, A. Kern, S. Shafer, "EasyLiving: Technologies for Intelligent Environments", Proc. of the Int. Conf. on Handheld and Ubiquitous Computing, pp. 12-27, 2000

[3] S. S.Intille, "Designing a House of the Future", IEEE Pervasive Computing, April-June, pp. 80-86, 2002

[4] B. Johanson, A. Fox, T. Winograd, "The Interactive Workspaces Project: Experiences with Ubiquitous Computing Rooms", IEEE Pervasive Computing Magazine, Vol. 1, No. 2, April-June, 2002

[5] L. Scanlo, "Rethinking the Computer - Project Oxygen is Turning out Prototype Computer Systems", Technology Review, 2004

[6] C. y D. Kidd, R. Orr, G. D. Abowd, C. G. Atkeson, I. A. Essa, B. MacIntyre, E. Mynatt, T. d E. Starner, W. Newstetter, "The Aware Home: A Living Laboratory for Ubiquitous Computing Research", Proc. of the Second Intl. Workshop on Cooperative Buildings, 1999

[7] Y. Nakauchi, T. Fukuda, K. Noguchi, T. Matsubara, "Intelligent Kitchen: Cooking Support by LCD and Mobile Robot with IC-Labeled Objects", Proc. of the 2005 IEEE/RSJ Int. Conf. on Intelligent Robotics and Systems, pp. 2464-2469, 2005

[8] Y. Nishida, M. Hiramoto, F. Kusunoki, H. Mizoguchi, "Learning by Doing: Space-Associate Language Learning Using a Sensorized Environment", in Proc. of IEEE Int. Conf. on Intelligent Robots and Systems, pp. 1583-1588, 2005

[9] M. Niitsuma, Hiroshi Hashimoto, Hideki Hashimoto, Akiko Watanabe, "Spatial Human Interface in Working Environment – Spatial-Knowledge-Tags to Access the Memory of Activity–", the 30th Annual Conf. of the IEEE Industrial Electronics Society, TC5-5, Busan, Korea, November, 2004

[10] ZPS(Zone Positioning System)–3D, Furukawa Industrial Machinery Systems CO., LTD, http://www.furukawakk.jp/products/ZPS_1.html

[11] M. Niitsuma, H. Hashimoto, "Spatial Memory for Collaboration Support System in Intelligent Space – User Identification Based on Spatial Memory –", the Proc. of SICE Annual Conference 2007, pp. 3073-3079, 2007

[12] T. Sasaki, H. Hashimoto, "Human Observation Based Mobile Robot Navigation in Intelligent Space", the Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, pp. 1044-1049, 2006

[13] Stephen C. Johnson, "Hierarchical Clustering Schemes", Psychometrika, Vol. 32, No. 3, pp. 241-254, 1967

[14] Mihoko Niitsuma, Hideki Hashimoto, "Extraction of Space-Human Activity Association for Design of Intelligent Environment", the Proc. of IEEE International Conference on Robotics and Automation (ICRA'07), pp. 1814-1819, 2007