2007 IEEE International Conference on
Robotics and Automation
Roma, Italy, 10-14 April 2007

Extraction of Space-Human Activity Association for Design of Intelligent Environment

Mihoko Niitsuma and Hideki Hashimoto

Abstract— **This paper presents a classification of human activities based on usage history of the spatial memory system and a numerical expression of each human activity 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, we describe two approaches to classify human activities. More specifically, the first approach is a classification of arranged SKTs by using only SKTs' information. The second approach is a classification of arranged SKTs by using both SKTs' information and usage history of the spatial memory system. As a result, the classification based on usage history of the spatial memory system showed multiple activities in the same area. On the other hand, the first approach resulted in a single cluster in the same area. We, then, give numerical values to SKT's content type, and obtained numerical expression of each human activity.**

I. INTRODUCTION

In the recent years, 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. In other words, if environmental settings will influence human activities, intended purposes for the places of human should be estimated. 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 presents 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. In this paper, we describe two approaches to classify human activities. The first approach is a classification of arranged SKTs by using only SKTs' information. The second approach is a classification of arranged SKTs by using SKTs' information and usage history of the spatial memory system.

The rest of the paper is organized as follows: Section II describes current spatial memory system. We improved the system to utilize it for a wide variety of situations. In Section III, we explain the association of space and human activities, and discuss the need to recognize such association by an intelligent environment. Section IV first defines a distance measure for clustering of arranged SKTs by using only SKT's information, and then classification results are shown. Classification of arranged SKT based on both SKT's information and observation of human activities, i.e. usage history of the spatial memory, is shown in Section V. The last section concludes the paper.

M. Niitsuma and H. Hashimoto are with Institute of Industrial Science, The University of Tokyo, Tokyo 153-8505, Japan, niitsuma@hlab.iis.u-tokyo.ac.jp, hashimoto@iis.u-tokyo.ac.jp

II. 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.

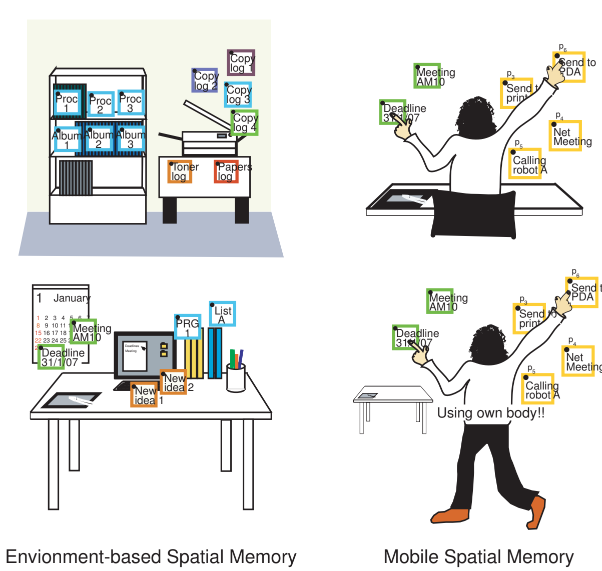

Fig. 1. Schematic concept of the spatial memory

A. Improvement of Spatial Memory

The prototype system had the following four problems.

1) the system was available for only one user,
2) only one display terminal to present accessed SKTs was available,
3) human action to access and store SKTs was recognized by using buttons, and
4) a user could not select a size of an accessible region online.

In a working environment, it is desirable that several users can utilize the spatial memory system simultaneously. For the multi-user system, of course several terminals are needed to present access SKTs. In order to attain a multi-user and multi-terminal system that solves the first and second problems, we implemented a parallel processing architecture. The system configuration of the improved spatial memory is shown in Fig. 2. When (c) SKT server presents accessed SKTs to (d) SKT client, it is important that a client which is suitable for a user to watch information is selected. The improved SKT server selects a client located most closely to the user, and delivers accessed SKTs toward the client.

To solve the third problem, the extraction method of a human indicating action, which only uses time series of the human indicator position, is implemented in the (a) Measurement unit of human indicator. Fig. 3 shows the measurement unit used in our experiments, "ZPS ultrasonic 3D positioning system (Furukawa Co., Ltd.)". The ZPS can measure the x, y, z position of an ultrasound transmitter shown in the bottom right corner of Fig. 3. A position of an ultrasound transmitter is defined as the human indicator.

To recognize "access" and "store" actions of a human indicator, the measurement unit is required to extract human's indicating actions from a sequence of his/her moves. In the case of the body type human indicator, the human usually stays in the accessible region after accessing a SKT. Consequently, if an "access" action is extracted based only on the time during the stopped state, the same SKT will be presented over and over until the human exits the accessible region. On the other hand, the hand type human indicator

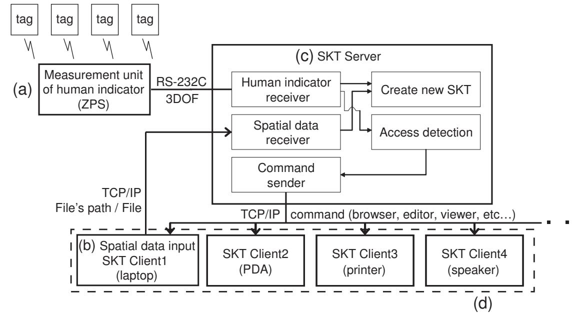

Fig. 2. System configuration

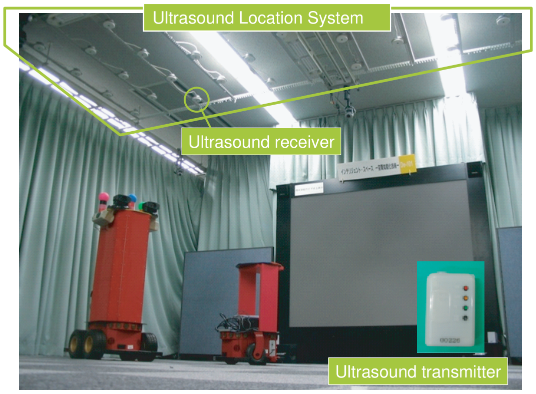

Fig. 3. Ultrasonic 3D positioning system as the measurement unit of human action

usually moves back toward a comfortable state after reaching over to an SKT. To avoid the above problem regarding repeated presentation and to recognize the human indicator action in both indicator types, a stopped state after moving a human indicator is utilized as a trigger to decide indication actions. To describe the state of a human indicator, stopped state S_{stop} is defined by

$$S_{stop} = \begin{cases} 1 & \text{if } v_x < V \text{ and } v_y < V \text{ and } v_z < V \\ & \text{and if } t_{stop} \geq T_{stop} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where v_x, v_y, v_z are velocities of the human indicator position for each axis, and t_{stop} is a time duration during which the condition on the indicator's velocities is satisfied. According to (1), the stopped state is activated when the speeds are below the threshold V for a time which exceeds T_{stop}. The values of the thresholds V and T_{stop} are experimentally determined.

Based on the states, the action of the human indicator S_{ind} is given the one of the following descriptions,

$$S_{ind} = \begin{cases} \text{INDICATE} & \text{if } S_{ind} = \text{MOVE and} \\ & \quad S_{stop} = 1 \\ \text{MOVE} & \text{if } S_{stop} = 0 \\ \text{STOP} & \text{otherwise} \end{cases} \quad (2)$$

When S_{ind} is INDICATE , the measurement unit of the human indicator carries out the determination process whether the indicating action is to access or to store.

Finally, to solve the fourth problem, the (b) Spatial memory data input unit, which is implemented as a graphical user interface application, was improved to select a size of the accessible region when a user stores an SKT. The current system allows users to select a size from the radio buttons, more specifically "large" sphere with a radius of 60 cm, "middle" sphere with a radius of 25 cm, and "small" sphere with a radius of 10 cm.

III. Association of Space and Human Activities

As shown in other approaches, location information is the main context which represents human activities, so the activities will be closely related to locations of an environment and human's location information is very useful to extract effective information about the environment [10]. However, the relationships are certainly not as simple as researchers or environment designers can predefine, since they are multitiered and underspecified. Therefore, what researchers or environment designers should consider is how to support the planning of service execution based on space-human activity association.

We found out an interesting usage of the spatial memory to observe human activities through previous experiments [11]. Fig. 4 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.

The procedures to extract space-activity association are considered as follows: first, arranged SKTs are spatially classified based on similarity between SKTs. Using the clustered SKTs, human activities are characterized by parameters such as SKTs' density, use frequency, their keywords and so on. The characteristics of human activities will associate places with the human's intended purposes. In this paper, two approaches of the spatial segmentation of human activities are shown in the following sections. The purpose of clustering of arranged SKTs is to classify them, and it is regarded as the preprocessing for estimating the activities and obtaining space-activity association based on usage history of the spatial memory.

IV. Classification of Arranged SKTs by Using Only SKT Information

In this section, we propose a classification of arranged SKTs by using only spatial memory addresses and the size of accessible regions.

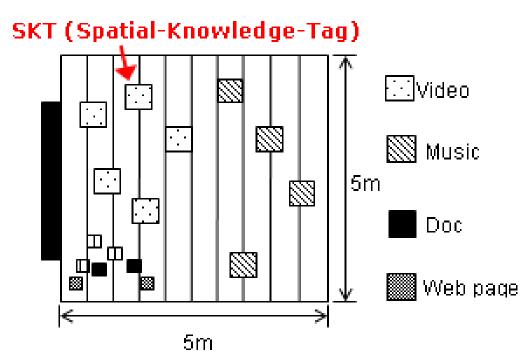

Fig. 4. An arrangement of the spatial memory

A. Distance measurement

For clustering of arranged SKTs, a distance between SKTs should be defined. From the standpoint of a human activity in an environment, it will correlate well with a position. 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 an accessible region into account. Based on this, the distance between the i-th SKT $\mathbf{x}_i$ and j-th SKT $\mathbf{x}_j$ is given by

$$D(\mathbf{x}_i,\mathbf{x}_j) = \|\mathbf{p}_i - \mathbf{p}_j\| + \alpha|r_i - r_j| \tag{3}$$

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.

Unsupervised hierarchical clustering method is used to classify arranged SKTs. 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. Clustering of arranged SKTs is performed by varying the threshold level.

B. Experiment

Fig. 5 shows the arrangement of the spatial memory used for the experiment, and the closed curves in this figure show the expected results of clusters. SKTs are arranged according to human activities in an environment, for examples, a user may work around the Desk 1 and the Desk 2. On the other hand, the demo movies may be presented for visitors in the area in front of the large-screen.

What is important here is that the segmentation can be obtained automatically – not directly specified by humans – but obtained through human activity observations.

For the clustering, the weighting factor α and the threshold level σ are given as follows; (a) ($\sigma = 1.5$, $\alpha = 0.5$), (b) ($\sigma = 2.0$, $\alpha = 0.5$) and (c) ($\sigma = 4.0$, $\alpha = 0.5$). Figs. 6 – 8 show the results of clustering of the arranged SKTs. By using the defined distance, the clustering taking the size of access regions in consideration was achieved. The result in Fig. 7 is most similar to the desired one, although of course the results are influenced by the threshold level σ and the weighting factor α. The level is considered as the corresponding value to segment human activities spatially. In the case of the larger value such as (c) $\sigma = 4.0$, SKTs were classified widely. On the other hand, in the case of the smaller value such as (a) $\sigma = 1.5$, SKTs were classified narrowly. The value of α influences what degree of the difference of accessible regions will enter in the same cluster. The degree of the difference of accessible regions may depend on the entire space size. If the classification is carried out in a large environment, the difference would have little effect on it because distance between SKTs which correspond to different activities would be relatively wide. However, in the case that there are different activities in a narrow environment, the difference of accessible regions should be important to classify the activities. Therefore, the value of α can be determined based on the size of the environment.

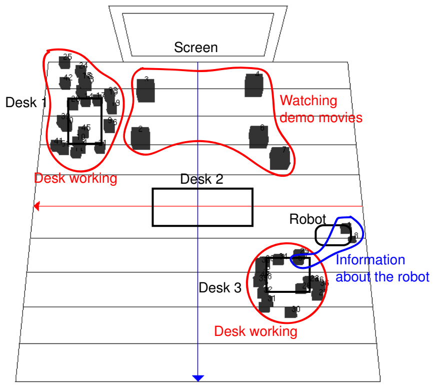

Fig. 5. Arrangement of the spatial memory used for the experiment

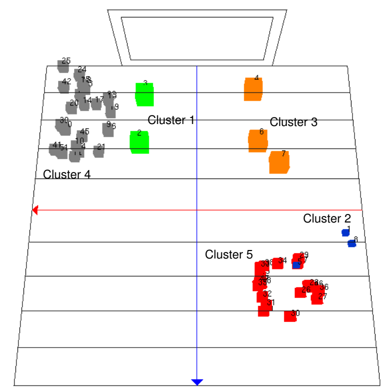

Fig. 6. Experimental result: (a) $\sigma = 1.5$, $\alpha = 0.5$

V. Classification of Arranged SKTs Based on Usage Histories of the Spatial Memory

In this section, we propose a classification of arranged SKTs based on both SKTs' information such as spatial memory addresses and the size of accessible regions, as in the previous section, and also observation of human activities. Observation of human activities can be described as usage histories of the spatial memory as mentioned in Section 3.

A. Observation of human activities

Basically, SKTs used in the same activity must have high similarity. To determine similarity distance between SKTs based on actual human activities, we define "observation

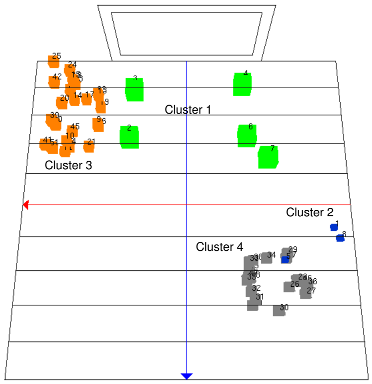

Fig. 7. Experimental result: (b) $\sigma = 2.0$, $\alpha = 0.5$

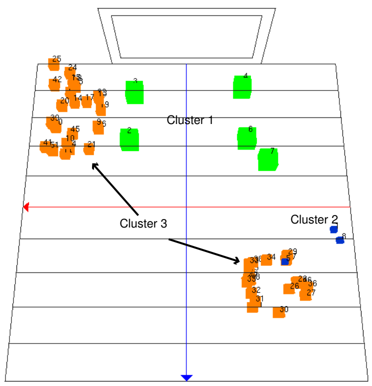

Fig. 8. Experimental result: (c) $\sigma = 4.0$, $\alpha = 0.5$

interval”, which is the time during which the user would be considered performing the same work. An observation interval is defined based on time duration and location of human activities. One of the conditions to determine an observation interval is the time interval of the spatial memory use. Another condition is the distance $d_{norm}(i,j)$ between continuously accessed SKT i and j. If the time duration while the user didn’t use the spatial memory t_s is less than a design parameter δt, and the distance $d_{norm}(i,j)$ is shorter than a design parameter $\gamma(t)$, 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 region of human activities depends on the characteristics. Therefore, by following the discussion in Section IV A, the condition value γ at time t is defined by taking the size of accessible regions into consideration as $\gamma(t) = \beta r_i(t)$ where $r_i(t)$ is the size of an accessible region of SKT i at time t and β is a design parameter.

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 (i,j) between SKTs to express the closeness is obtained by using the lists of the access times. The distance $d(i,j)$ is given by the inverse of the total number of access times $N_j(p)$ when SKT j is used with SKT i in each observation interval. Namely, $d(i,j)$ is given by

$$d(i,j) = \sum_p \sum_j N_j(p). \tag{4}$$

B. Experiment

We observed human activities using the spatial memory for about two days. The whole arrangement of SKTs is shown in Fig. 5. As in the Section IV, unsupervised hierarchical clustering method is used to classify arranged SKTs.

For the clustering, the design parameter $\beta = 6$ and the threshold level $\sigma = 2.0$ are given. Clustering of arranged SKTs based on the usage histories is performed by varying the time during which the user did not use the spatial memory.

Figs. 9 and 10 show the results of clustering. These results shows the clusters which consist of more than one SKT. As shown in the figures, the difference of the duration to classify human activities affects the classification. There are important differences between the clustering results of Section IV and Section V. By observing actual human activities, we can find that the user performs different activities in a same area. It is impossible to extract them by using only arranged SKTs and using only simple sensory information such as position and chair switches and so on. Fig. 10 is successful result to describe actual human activities.

Each cluster corresponds to each human activity. We assumed that each cluster is characterized by SKT’s contents type. Value of each contents type is given based on its modifiability as shown in Table I.

SKTs which consist of each cluster in Fig. 10 are given values shown in Table I, and then average values and variance values are obtained as feature of each activity. Table II shows the result of each cluster characterization. An average value represents a type of human activity and a variance value represents wideness of the activity. Therefore, we can

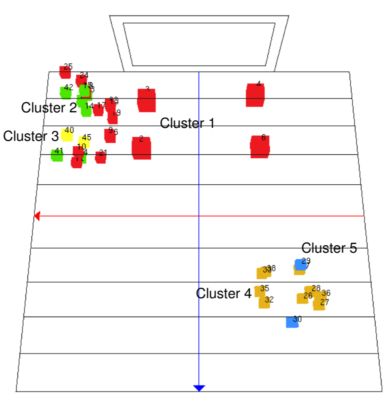

Fig. 9. Clustering result that the duration while the user did not use the spatial memory be 60 minutes

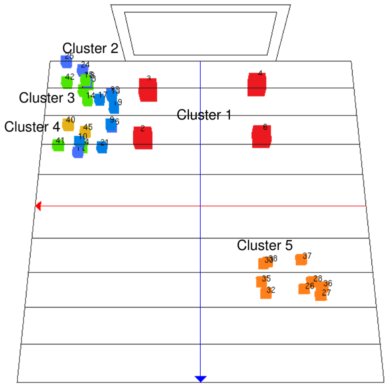

Fig. 10. Clustering result that the duration while the user did not use the spatial memory be 30 minutes

TABLE I
CHARACTERIZATION OF CONTENTS FORMAT

Format	Operation possibility	Value
Music	Listen	-3
Movie	Watch	-3
Image	Watch, Modift	-1
Webpage	Read	1
Document	Read, Write	3
Program	Read, Write	3
Application	Execution(Modify,Create)	5

consider that the activity of Cluster 1 was probably not a creative work but watching movies or listening music. Also, we can observe that the same place was utilized for different activities, such as Cluster 3 and 4. On the other hand, we can find that the activities of Cluster 3 and 5 are similar, but from the large variance of Cluster 5 we can conclude the activity there probably needs a more wide range of information than in Cluster 3. Moreover, of course, the system knows how long each activity was continued and how frequently each SKT was used. Therefore, the system has the possibility to describe features of human activities more specifically.

VI. CONCLUSION

This paper presented two approaches to classify human activities spatially and characteristically by using the spatial memory. More specifically, one is a classification of arranged SKTs by using only SKTs' information such as spatial memory addresses and the size of accessible regions. The other is a classification of arranged SKT based on observation of human activities and SKTs' information. Observation of human activities is described as usage histories of the spatial memory. Unsupervised hierarchical clustering for both classifications was applied.

TABLE II
AVERAGE VALUES AND VARIANCE VALUES OF EACH CLUSTER SHOWN IN FIG. 10

No. of cluster	Average	Variance
Cluster 1	-2	21
Cluster 2	1.333333	10.037037
Cluster 3	2.428571	29.912539
Cluster 4	-1.285714	16.314869
Cluster 5	2.571429	80.731804
Cluster 6	0	9

To classify arranged SKTs using only SKTs' information, we defined the distance between SKTs to express the closeness taking both their position and the difference of the accessible regions' size into consideration. The clustering results show that the rough classification was possible by using this distance.

To describe association between human activities and SKTs, we defined the observation method based on duration of human activities. In addition, the distance between SKTs to express the closeness was given based on usage histories of each SKT. The clustering results based on observation of actual human activities confirm that they can describe multiple human activities in certain area.

As the future work, we will investigate space-human activity association through widely experiments. Furthermore, we plan to design and implement services of an intelligent environment based on the observation of space-human activity association in order to realize flexible and adaptive effect for humans.

REFERENCES

[1] J.-H. Lee and 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 and 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 and 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, and 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 and Akiko Watanabe, "Spatial Human Interface in Working Environment – Spatial-knowledge-tags to Access the Memory of Activity–", the 30th Annual Conference of the IEEE Industrial Electronics Society (IECON 04), TC5-5, Busan Korea, November, 2004
[10] T. Sasaki and H. Hashimoto, "Human Observation Based Mobile Robot Navigation in Intelligent Space", in Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, pp. 1044-1049, 2006.
[11] M. Niitsuma, Hiroshi Hashimoto and Hideki Hashimoto, "Spatial Memory: An Aid System for Human Activity in Intelligent Space", Proceedings of The 2006 IEEE International Conference on Robotics and Automation, pp.4258–4263, 2006.